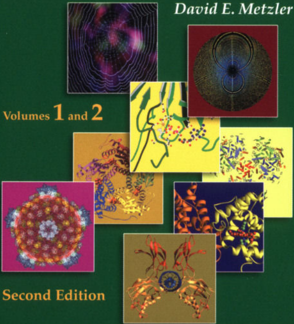
ELSEVIER
ACADEMIC
PRESS

BIOCHEMISTRY

The Chemical Reactions of Living Cells

David E. Metzler



Volumes 1 and 2

Second Edition

VISIT…

LANZAROTE
Caliente.com

Table of Contents
Volume 1

Chapter 1. The Scene of Action *1*

Chapter 2. Amino Acids, Peptides, and Proteins *39*

Chapter 3. Determining Structures and Analyzing Cells *95*

Chapter 4. Sugars, Polysaccharides, and Glycoproteins *161*

Chapter 5. The Nucleic Acids *199*

Chapter 6. Thermodynamics and Biochemical Equilibria *281*

Chapter 7. How Macromolecules Associate *325*

Chapter 8. Lipids, Membranes, and Cell Coats *379*

Chapter 9. Enzymes: The Catalysts of Cells *455*

Chapter 10. An Introduction to Metabolism *505*

Chapter 11. The Regulation of Enzymatic Activity and Metabolism *535*

Chapter 12. Transferring Groups by Displacement Reactions *589*

Chapter 13. Enzymatic Addition, Elimination, Condensation, and Isomerization: Roles for Enolate and Carbocation Intermediates *677*

Chapter 14. Coenzymes: Nature's Special Reagents *719*

Chapter 15. Coenzymes of Oxidation – Reduction Reactions *765*

Chapter 16. Transition Metals in Catalysis and Electron Transport *837*

Table of Contents
Volume 2

Chapter 17. The Organization of Metabolism *938*

Chapter 18. Electron Transport, Oxidative Phosphorylation, and Hydroxylation *1012*

Chapter 19. The Chemistry of Movement *1088*

Chapter 20. Some Pathways of Carbohydrate Metabolism *1128*

Chapter 21. Specific Aspects of Lipid Metabolism *1180*

Chapter 22. Polyprenyl (Isoprenoid) Compounds *1226*

Chapter 23. Light and Life *1272*

Chapter 24. The Metabolism of Nitrogen and Amino Acids *1358*

Chapter 25. Metabolism of Aromatic Compounds and Nucleic Acid Bases *1420*

Chapter 26. Biochemical Genetics *1472*

Chapter 27. Organization, Replication, Transposition, and Repair of DNA *1528*

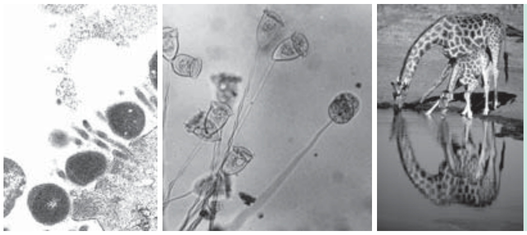
Chapter 28. The Transcription of Genes *1602*

Chapter 29. Ribosomes and the Synthesis of Proteins *1668*

Chapter 30. Chemical Communication Between Cells *1740*

Chapter 31. Biochemical Defense Mechanisms *1830*

Chapter 32. Growth and Development *1878*



Left: Cells of the pathogenic O157:H7 strain of *Escherichia coli* attached to the surface epithelium of the cecum of a neonatal piglet. Electron-dense filaments (presumably polymerized actin) in the host cytoplasm can be seen subjacent to attached bacteria. The bacteria have effaced most micro-villi but some remain between the bacterial cells. Courtesy of Evelyn A. Dean-Nystrom, National Animal Disease Center, USDA, Agricultural Research Service, Ames, IA. Center: Many unicellular organisms such as these *Vorticella* inhabit wet and moist environments throughout the biosphere. Invertebrates have evolved as long as humans and have complex specializations such as the contractile stem of these protozoa. Courtesy of Ralph Buchsbaum. Right: Although 97% of animals are invertebrates, ~ 3% of the several million known species have backbones. Giraffe: © M. P. Kahl, Photo Researchers

Contents

2 .... **A. The Simplest Living Things**
3 .... 1. *Escherichia coli*
3 .... 2. The Bacterial Genome
5 .... 3. Ribonucleic Acids (RNA) and the Transcription and Translation of Genetic Information
5 .... 4. Membranes and Cell Walls
6 .... 5. Flagella and Pili
6 .... 6. Classification and Evolution of Bacteria
8 .... 7. Nutrition and Growth of Bacteria
9 .... 8. Photosynthetic and Nitrogen-Fixing Prokaryotes
11 .... **B. Eukaryotic Cells**
11 .... 1. The Nucleus
11 .... 2. The Plasma Membrane
12 .... 3. Vacuoles, Endocytosis, and Lysosomes
13 .... 4. The Endoplasmic Reticulum and Golgi Membranes
14 .... 5. Mitochondria, Plastids, and Peroxisomes
15 .... 6. Centrioles, Cilia, Flagella, and Microtubules
15 .... 7. Cell Coats, Walls, and Shells
15 .... **C. Inheritance, Metabolic Variation, and Evolution of Eukaryotes**
17 .... 1. A Changing Genome
17 .... 2. Genetic Recombination, Sex, and Chromosomes
18 .... 3. Haploid and Diploid Phases
18 .... **D. Survey of the Protists**
18 .... 1. Protozoa
20 .... 2. Fungi
20 .... 3. Algae
23 .... **E. The Variety of Animal Forms**
23 .... 1. Major Groups of Multicellular Animals
25 .... 2. Cell Types and Tissues
26 .... *Tissues*
26 .... *Blood cells*
26 .... *Cell culture*
26 .... 3. Communication
28 .... *Cell contacts and junctions*
29 .... *Cell recognition*
29 .... **F. Higher Plants and Plant Tissues**
30 .... **G. The Chemical Composition of Cells**

34 .... **References**
36 .... **Study Questions**

Boxes

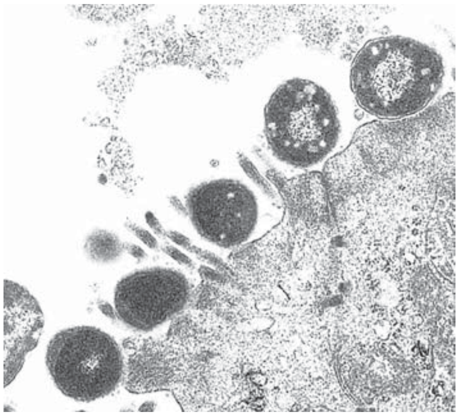
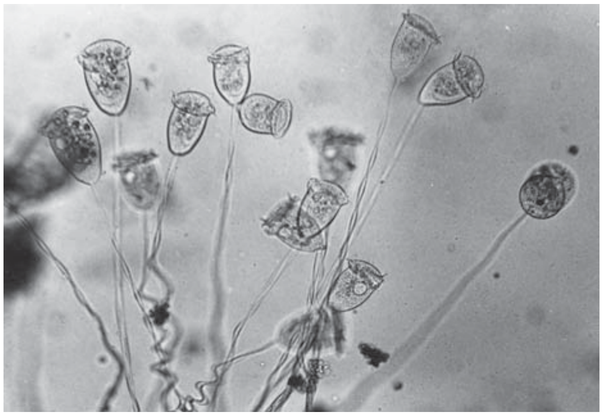
2 .... Box 1-A About Measurements
4 .... Box 1-B Relative Molecular Mass, M_r, and Daltons
9 .... Box 1-C In the Beginning
16 .... Box 1-D Inherited Metabolic Diseases
25 .... Box 1-E Errors, Misconceptions, and Speculation
32 .... Box 1-F About the References

Tables

7 .... Table 1-1 A Systematic Classification Scheme for Bacteria
11 .... Table 1-2 Approximate Sizes of Some Cells
12 .... Table 1-3 Haploid Genome Sizes for Several Organisms
31 .... Table 1-4 Approximate Composition of Metabolically Active Cells and Tissues

The Scene of Action

1

This book is about the chemistry of living cells with special emphasis on the trillions of cells that make up your own body. Every aspect of life depends upon the chemical makeup of cells and on the chemical properties of the remarkable molecules found within the cells. The information presented here will give the reader a solid foundation for understanding not only the chemical basis of life but also the revolutionary developments in molecular biology, biochemical genetics, medicine, and agriculture which dominate today's scientific news and which will play an increasingly important role in our lives.

The first theme of the book is **biomolecular structure**. We'll look carefully at the complex structures of proteins, carbohydrates, RNA, DNA, and many other substances. We'll not only examine in-depth their molecular architecture but also study the chemical properties that make life possible.

A second theme is **metabolism**, the unceasing, complex network of thousands of chemical reactions by which cells grow and reproduce, take up foods and excrete wastes, move, and communicate with each other. Within cells we have a **steady state**, a condition in which the complex chemical constituents of cells are continuously being synthesized in one series of reactions and degraded in another. The result is a marvelous system of self-renewal or "turnover" of tissues. We'll examine the chemical reactions involved in these processes as well as the ways in which they are controlled. We will consider both the reaction sequences and the techniques such as cloning of genes, isotopic labeling, X-ray diffraction, and nuclear magnetic resonance spectroscopy, which are used today to study metabolism.

Human beings are surrounded by many other living creatures whose activities are important to us. Photosynthetic organisms obtain energy from sunlight and synthesize compounds that the human body requires but cannot make. Microorganisms cause decay of organic matter and convert it into forms usable by plants. This book deals with the chemical reactions occurring in all of these organisms. We'll look at strange and unusual reactions, along with those metabolic sequences common to most living things.

Each one of the thousands of chemical reactions of metabolism is catalyzed by an **enzyme**. Most of these enzymes are proteins, but others are made from **RNA** (ribonucleic acid). In both cases enzymes are very large molecules with precise three-dimensional structures. The study of the properties of enzymes and of **enzymatic catalysis** is a third theme of the book. Not only are the chemical mechanisms by which enzymes act of interest but also enzymes are often targets for useful drugs. Incorrectly formed enzymes can result in serious diseases.

The sequences of the amino acids in the chains from which proteins are constructed are encoded in the nucleotide sequences of **DNA** (deoxyribonucleic acid). The coding sequence for a protein in the DNA is found in the **structural gene** for that protein. The RNA enzymes are also encoded by DNA genes. A fourth major theme of the book deals with the nature of the **genetic code** used in DNA and with the processes by which cells read and interpret the code. It also includes study of the methods by which thousands of genes have been mapped to specific positions in chromosomes, isolated, cloned, and sequenced.

A large number of proteins present in the outer surfaces of cells serve as **receptors** that receive chemical messages and other signals from outside the cell. The receptors, which are sometimes enzymes, respond by generating internal signals that control metabolism and cell growth. Such **molecular signaling** is another major area of contemporary biochemistry.

Biologists have described over a million species, and several millions of others probably exist.[1] Many of these organisms have very specialized ways of life. However, they all have much chemistry in common. The same 20 amino acids can be isolated from proteins of plants, animals, and microorganisms. Formation of lactic acid in both bacteria and human muscle requires the same enzymes. Except for some small variations, the genetic code is universal—the same for all organisms. Thus, there is a unity of life and we can study metabolism as the entirety of chemical transformations going on in all living things. However, the differences among species are also impressive. Each species has its own gene for almost every protein.

When the enzyme that catalyzes a particular metabolic reaction is isolated from a number of different organisms, it is usually found to have similar properties and a similar mechanism of catalysis, regardless of the source. However, the *exact* sequence of amino acids in the enzyme will be almost unique to the organism that produced it. When the three-dimensional structures are compared it is found that differences between species often affect only the peripheral parts of an enzyme molecule. The interior structure of the protein, including the catalytic machinery, is highly conserved. However, the surface regions, which often interact with other macromolecules, vary greatly. Such interactions help to control metabolism and may account for many differences in the metabolism among living beings.

Variations in protein structures are not limited to differences between species. Individuals differ from one another. Serious genetic diseases sometimes result from the replacement of a single amino acid unit in a protein by a different amino acid. Genetic deviations from the "normal" structure of a protein result from **mutations**. Many mutations, whether they occurred initially in our own cells or in those of our ancestors, are detrimental.

However, such mutations also account for variation among individuals of a species and allow for evolution. The chemical nature and consequences of mutations and their significance to health, medicine, and agriculture are dealt with throughout the book. We now have reliable methods for inducing in the laboratory mutations at any specific place in a protein sequence and also for synthesizing new DNA sequences. These techniques of **genetic engineering** have given biochemists the ability to modify protein structures freely, to create entirely new proteins, and to provide a basis for the rapidly developing field of **genetic therapy**.

It should be clear from this introduction that **biochemistry** deals with virtually every aspect of life. The distinguishing feature of the science is that it approaches biological questions in terms of the underlying chemistry. The term **molecular biology** is often regarded as synonymous with biochemistry.

However, some scientists use it to imply a more biological approach. These molecular biologists also emphasize structure and function but may have a goal of understanding biological relationships more than chemical details. **Biophysics**, a closely related science, encompasses the application of physical and mathematical tools to the study of life.

A. The Simplest Living Things

The simplest organisms are the **bacteria.**[2–5] Their cells are called **prokaryotic** (or procaryotic) because no membrane-enclosed nucleus is present. Cells of all other organisms contain nuclei separated from the cytoplasm by membranes. They are called **eukaryotic**. While viruses (Chapter 5) are sometimes regarded as living beings, these amazing parasitic objects are not complete organisms and have little or no metabolism of their own. The smallest bacteria are the **mycoplasmas**.[6–8] They do not have the rigid cell wall characteristic of most bacteria. For this reason they are easily deformed and often pass through filters designed to stop bacteria. They are nutritionally fussy and are usually, if not always, parasitic. Some live harmlessly in mucous membranes of people, but others cause diseases.

BOX 1-A ABOUT MEASUREMENTS

In 1960 the International General Conference on Weights and Measures adopted an improved form of the metric system, **The International System of Units** (SI). The units of mass, length, and time are the kilogram (kg), meter (m), and second (s). The following prefixes are used for fractions and multiples:

10^{-18}, *atto* (a)
10^{-15}, *femto* (f)
10^{-12}, *pico* (p)
10^{-9}, *nano* (n)
10^{-6}, *micro* (μ)
10^{-3}, *milli* (m)
10^{3}, *kilo* (k)
10^{6}, *mega* (M)
10^{9}, *giga* (G)
10^{12}, *tera* (T)
10^{15}, *peta* (P)
10^{18}, *exa* (E)

There is an inconsistency in that the prefixes are applied to the gram (g) rather than to the basic unit, the kilogram.

SI units have been used throughout the book whenever possible. There are no feet, microns, miles, or tons. Molecular dimensions are given uniformly in nanometers rather than in angstrom units (Å; 1Å = 0.1 nm). Likewise the calorie and kilocalorie have been replaced by the SI unit of energy, the **joule** (J; 1 calorie = 4.184 J).

Throughout the book frequent use is made of the following symbols:

, "approximately equal to"
~, "approximately" or "about"

For example, M*ycoplasma pneumoniae* is responsible for primary atypical pneumonia.

Cells of mycoplasmas sometimes grow as filaments but are often spherical and as small as 0.3 micrometer (μm) in diameter. Their outer surface consists of a thin **cell membrane** about 8 nanometers (nm) thick. This membrane encloses the **cytoplasm**, a fluid material containing many dissolved substances as well as submicroscopic particles. At the center of each cell is a single, highly folded molecule of DNA, which constitutes the bacterial chromosome. Besides the DNA there may be, in a small spherical mycoplasma, about 1000 particles ~20 nm in diameter, the **ribosomes**. These ribosomes are the centers of protein synthesis. Included in the cytoplasm are many different kinds of proteins, but there is room for a total of only about 50,000 protein molecules. Several types of RNA as well as many smaller molecules are also present. Although we don't know what minimum quantities of proteins, DNA, and other materials are needed to make a living cell, it is clear that they must all fit into the tiny cell of the mycoplasma.

1. *Escherichia coli*

The biochemist's best friend is *Escherichia coli*, an ordinarily harmless inhabitant of our intestinal tract. This bacterium is easy to grow in the laboratory and has become the best understood organism at the molecular level.[4,9] It may be regarded as a typical true bacterium or **eubacterium**. The cell of *E. coli* (Figs. 1-1, 1-2) is a rod ~2 μm long and 0.8 μm in diameter with a volume of ~1 μm^3 and a density of ~1.1 g/cm^3. The mass is ~1×10^{-12} g, i.e., 1 picogram (pg) or ~0.7×10^{12} daltons (Da) (see Box 1-B).[4] It is about 100 times bigger than the smallest mycoplasma but the internal structure, as revealed by the electron microscope, resembles that of a mycoplasma.

Each cell of *E. coli* contains from one to four identical DNA molecules, depending upon how fast the cell is growing, and ~15,000–30,000 ribosomes. Other particles that are sometimes seen within bacteria include food stores such as fat droplets and granules (Fig. 1-3). The granules often consist of **poly-β-hydroxybutyric acid**[10] accounting for up to 25% of the weight of *Bacillus megaterium*. **Polymetaphosphate**, a highly polymerized phosphoric acid, is sometimes stored in "metachromatic granules." In addition, there may be droplets of a separate aqueous phase, known as **vacuoles**.

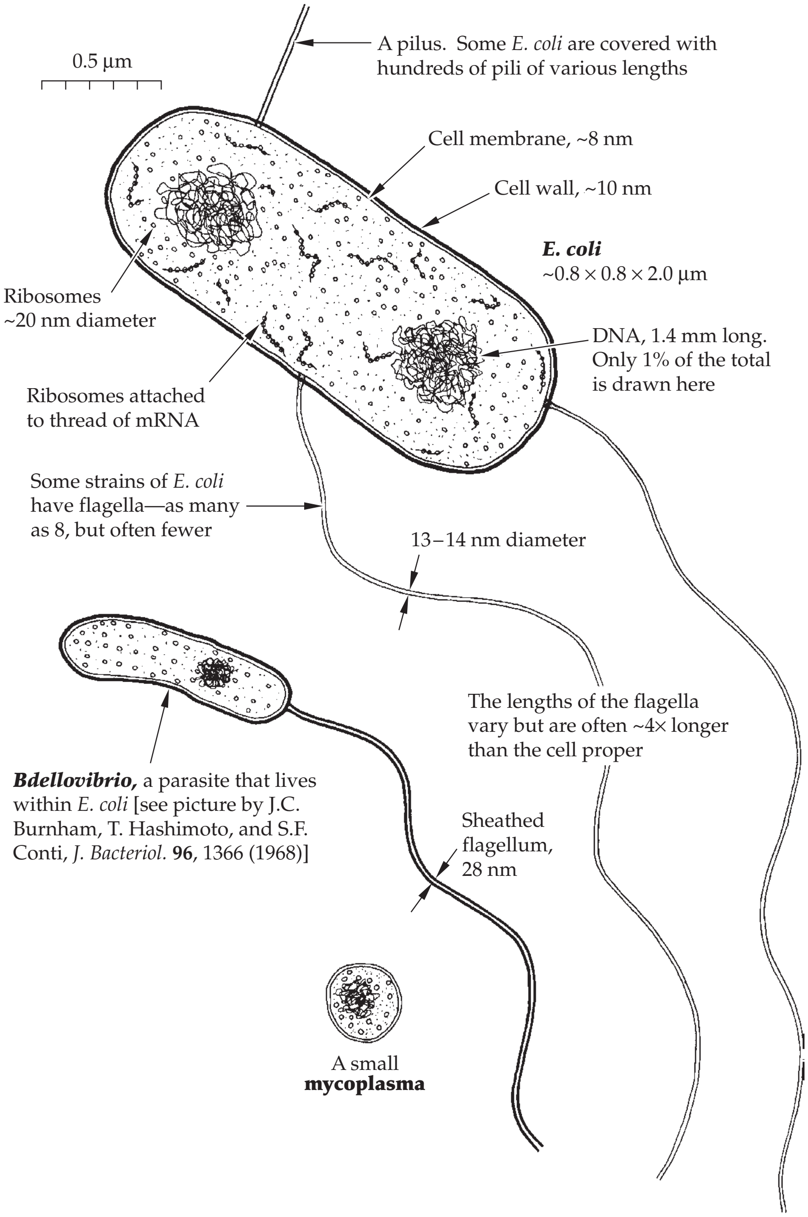

0.5 μm
A pilus. Some *E. coli* are covered with hundreds of pili of various lengths
Cell membrane, ~8 nm
Cell wall, ~10 nm
E. coli
~0.8 × 0.8 × 2.0 μm
Ribosomes ~20 nm diameter
DNA, 1.4 mm long. Only 1% of the total is drawn here
Ribosomes attached to thread of mRNA
Some strains of *E. coli* have flagella—as many as 8, but often fewer
13–14 nm diameter
The lengths of the flagella vary but are often ~4× longer than the cell proper
Bdellovibrio, a parasite that lives within *E. coli* [see picture by J.C. Burnham, T. Hashimoto, and S.F. Conti, *J. Bacteriol.* **96**, 1366 (1968)]
Sheathed flagellum, 28 nm
A small **mycoplasma**

Figure 1-1 *Escherichia coli* and some smaller bacteria.

2. The Bacterial Genome

The genetic instructions for a cell are found in the **DNA molecules**. All DNA is derived from four different kinds of monomers, which we call **nucleotides**. DNA molecules are double-stranded: two polymer chains are coiled together, their nucleotide units being associated as **nucleotide pairs** (see Fig. 5-7). The genetic messages in the DNA are in the form of sequences of nucleotides. These sequences usually consist of a series of code "words" or **codons**. Each codon is composed of three successive nucleotides and specifies which one of the 20 different kinds of amino acids will be used at a particular location in a protein. The sequence of codons in the DNA tells a cell how to order the amino acids for construction of its many different proteins.

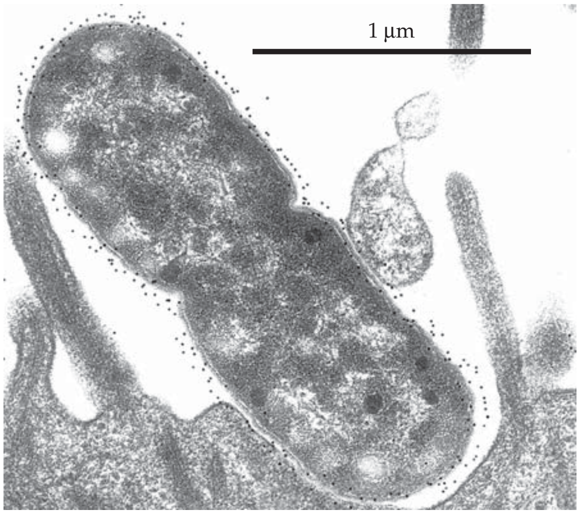


Figure 1-2 Transmission electron micrograph of a dividing cell of *Escherichia coli* O157:H7 attached to the intestinal epithelium of a neonatal calf. These bacteria, which are able to efface the intestinal microvilli, form characteristic attachments, and secrete shiga toxins, are responsible for ~73,000 illnesses and 60 deaths per year in the U. S.[10a] After embedding, the glutaraldehyde-fixed tissue section was immunostained with goat anti-O157 IgG followed by protein A conjugated to 10-nm gold particles. These are seen around the periphery of the cell bound to the O-antigen (see Fig. 8-28). Notice the two microvilli of the epithelium. Courtesy of Evelyn A. Dean-Nystrom, National Animal Disease Center, USDA, Agricultural Research Service, Ames, IA.

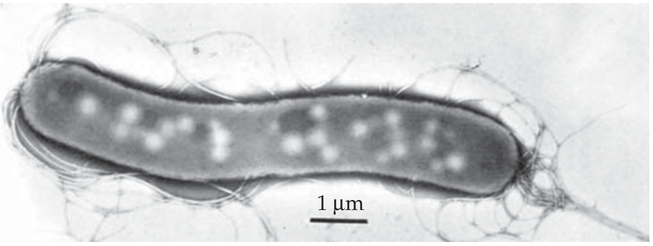


Figure 1-3 A cell of a *Spirillum* negatively stained with phosphotungstic acid. Note the tufts of flagella at the ends, the rough appearance of the outer surface, the dark granules of poly-β-hydroxybutyric acid and the light-colored granules of unknown nature. Courtesy of F. D. Williams, Gail E. VanderMolen, and C. F. Amstein.

Assume that a typical protein molecule consists of a folded chain of 400 amino acids. Its structural gene will therefore be a sequence of 1200 nucleotide pairs. Allowing a few more nucleotides to form spacer regions between genes we can take ~1300 as the number of nucleotide pairs in a typical bacterial gene. However, some genes may be longer and some may be much shorter. The **genome** is the quantity of DNA that carries a complete set of genetic instructions for an organism. In bacteria, the genome is a single chromosome consisting of one double-stranded DNA molecule. *Mycoplasma genitalium* is the smallest organism for which the DNA sequence is known.[11] Its genome is a double-helical DNA circle of 580,070 nucleotide pairs and appears to contain about 480 genes (an average of ~1200 nucleotides per gene).

The average mass of a nucleotide pair (as the disodium salt) is 664 Da. It follows that the DNA of *M. genitalium* has a mass of ~385 x 10^6 Da. The relative molecular mass (M_r) is 0.385 x 10^9 (See Box 1-B for definitions of dalton and M_r). The DNA of *E. coli* is about seven times larger with a mass of ~2.7 x 10^9 Da. It contains ~4.2 x 10^6 nucleotide pairs and encodes over 4000 different proteins (see Table 1-3).

Each nucleotide pair contributes 0.34 nm to the length of the DNA molecule; thus, the total length of DNA of an *E. coli* chromosome is 1.4 mm. This is about 700 times the length of the cell which contains it. Clearly, the molecules of DNA are highly folded, a fact that accounts for their appearance in the electron microscope as dense aggregates called **nucleoids**, which occupy about one-fifth of the cell volume (Fig. 1-4).

BOX 1-B RELATIVE MOLECULAR MASS, M_r, AND DALTONS

Atomic and molecular masses are assigned relative to the mass of the carbon isotope, ^{12}C, whose atomic weight is defined as exactly 12. The actual mass of a single atom of ^{12}C is defined as 12 **daltons**, one dalton being 1.661 x 10^{-24} g. The mass of a molecule can be given in daltons **(Da)** or kilodaltons **(kDa)**. This molecular mass in daltons is numerically equivalent to the relative molecular mass **(M_r)** or molecular weight **(MW)**[a] and also to the molar mass **(g/mol)**. However, it is not correct to use the dalton for the unitless quantity M_r. Masses of structures such as chromosomes, ribosomes, mitochondria, viruses, and whole cells as well as macromolecules can be given in daltons.[b]

---

[a] The Union of Pure and Applied Chemistry renamed molecular weight as **relative molecular mass** with the symbol M_r; M_r = MW.

[b] J. T. Edsall (1970) *Nature* (*London*) **228**, 888.

Each bacterial nucleoid contains a complete set of genetic "blueprints" and functions independently. Each nucleoid is **haploid**, meaning that it contains only a single complete set of genes. In addition to their chromosome, bacteria often contain smaller DNA molecules known as **plasmids**. These plasmids also carry genetic information that may be useful to bacteria. For example, they often encode proteins that confer resistance to antibiotics. The ability to acquire new genes from plasmids is one mechanism that allows bacteria to adapt readily to new environments.[12] Plasmids are also used in the laboratory in the cloning of genes and in genetic engineering (Chapter 26).

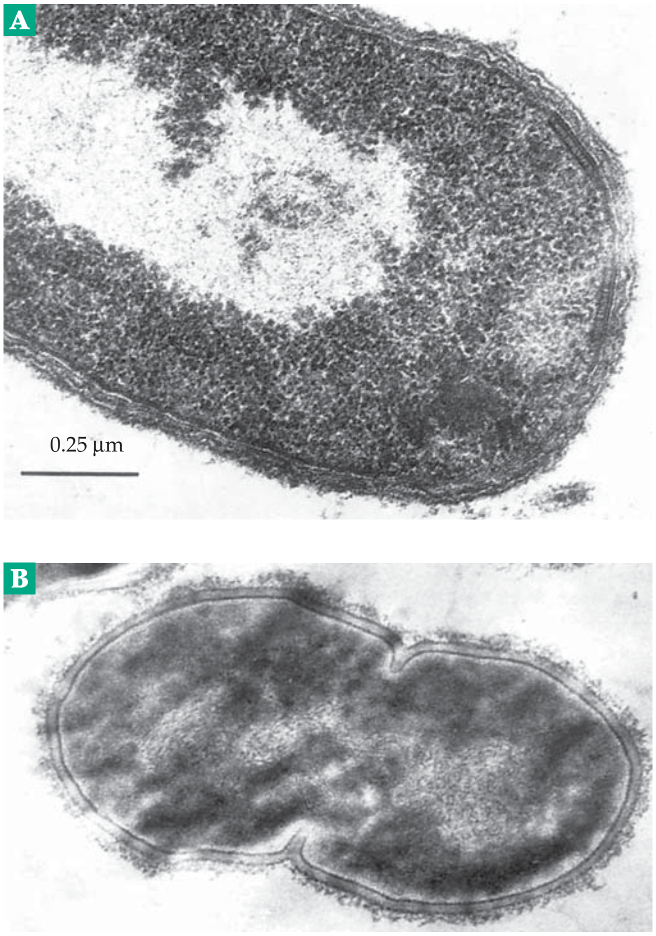


Figure 1-4 (A) Thin (~60 nm) section of an aquatic gram-negative bacterium, *Aquaspirillum fasciculus*. Note the light-colored DNA, the dark ribosomes, the double membrane characteristics of gram-negative bacteria (Chapter 8, Section E), and the cell wall. In addition, an internal "polar membrane" is seen at the end. It may be involved in some way in the action of the flagella. (B) A thin section of dividing cell of *Streptococcus*, a gram-positive organism. Note the DNA (light-stranded material). A portion of a mesosome is seen in the center and septum can be seen forming between the cells. Micrographs courtesy of F. D. Williams, Gail E. VanderMolen, and C. F. Amstein.

3. Ribonucleic Acids (RNA) and the Transcription and Translation of Genetic Information

The genetic information in the DNA is not utilized directly by the protein-synthesizing machinery of cells. Instead, molecules of ribonucleic acid (RNA) are synthesized according to the instructions encoded in the DNA, a process called **transcription**. Although they differ from DNA significantly in their structure, these RNA molecules carry the same coded information as is found in a length of DNA that contains one or a few genes. If DNA is regarded as the "master blueprint" of the cell, molecules of RNA are "secondary blueprints." This concept is embodied in the name **messenger RNA** (mRNA) which is applied to a small, short-lived fraction of RNA that carries information specifying amino acid sequences of proteins. Each molecule of mRNA carries the genetic message from one or more genes to the ribosomes where the proteins are made.

Ribosomes are extraordinarily complex little protein-synthesizing machines. Each ribosome of *E. coli* has a mass of 2.7×10^6 Da and contains 65% of a stable **ribosomal RNA** and ~35% protein. About 50 different kinds of protein molecules are present as parts of the ribosomal structure. Working together with a variety of **transfer RNA** molecules and enzymes, the ribosomes are able to read the genetic messages from mRNA and to accurately assemble any kind of protein molecule that a gene may specify. This process is called **translation** of the genetic message.

4. Membranes and Cell Walls

Like the mycoplasma, the *E. coli* cell is bounded by an 8-nm membrane which consists of ~50% protein and 50% lipid. When "stained" (e.g., with permanganate) for electron microscopy, this single membrane appears as two very thin (2.0 nm) dark lines separated by an unstained center band (~3.5 nm) (Fig. 1-4; see also Fig. 8-4). Single membranes of approximately the same thickness and staining behavior occur in all cells, both of bacteria and of eukaryotes.

A cell membrane is much more than just a sack. It serves to control the passage of small molecules into and out of the cell. Its outer surface carries receptors for recognition of various materials. The inside surface of bacterial membranes contains enzymes that catalyze most of the oxidative metabolism of the cells. Bacterial cell membranes are sometimes folded inward to form internal structures involved in photosynthesis or other specialized reactions of metabolism such as oxidation of ammonia to nitrate.[2] In *E. coli* replication of DNA seems to occur on certain parts of the membrane surface, probably under the control of membrane-bound enzymes. The formation of the new membrane which divides multiplying cells proceeds synchronously with the synthesis of DNA.

A characteristic of true bacteria (**eubacteria**) is a rigid **cell wall** which surrounds the cell membrane. The 40-nm-thick wall of *E. coli* is a complex, layered structure five times thicker than the cell membrane. Its chemical makeup is considered in Chapter 8. One of the layers is often referred to as the **outer membrane**. In some bacteria the wall may be as much as 80 nm thick and may be further surrounded by a thick **capsule** or **glycocalyx** (slime layer).[13] The main function of the wall seems to be to prevent osmotic swelling and bursting of the bacterial cell when the surrounding medium is hypotonic.

If the osmotic pressure of the medium is not too low, bacterial cell walls can sometimes be dissolved, leaving living cells bounded only by their membranes. Such **protoplasts** can be produced by action of the enzyme lysozyme on gram-positive bacteria such as *Bacillus megaterium*. Treatment of cells of gram-negative bacteria with penicillin (Box 20-G) produces **spheroplasts**, cells with partially disrupted walls. Spheroplasts and protoplasts are useful in biochemical studies because substances enter cells more readily when the cell wall is absent. Strains of bacteria lacking rigid walls are known as **L forms**.

5. Flagella and Pili

Many bacteria swim at speeds of 20–60 µm/s, ten or more body lengths per second! Very thin thread-like **flagella** of diameter 13–20 nm coiled into a helical form are rotated by the world's smallest "electric motors" to provide the motion.[14] While some bacteria have a single flagellum, the corkscrew-like *Spirillum* (Fig. 1-3) synchronously moves tufts of flagella at both ends. Some strains of *E. coli* have no flagella, but others contain as many as eight flagella per cell distributed over the surface. The flagella stream out behind in a bundle when the bacterium swims. The flagella of the helical **spirochetes** are located inside the outer membrane.[15,16]

In addition to flagella, extremely thin, long, straight filaments known as **pili** or **fimbriae** (Fig. 1-2) project from the surfaces of many bacteria.[14] The "sex pili" (F pili and I pili) of *E. coli* have a specific role in sexual conjugation. The similar but more numerous common pili or fimbriae range in thickness from 3 to 25 nm and in length from 0.2 to 2 µm. Pili are involved in adhesion of bacteria to surrounding materials or to other bacteria and facilitate bacterial infections.[17–19] A typical *E. coli* cell has 100–300 pili.[5]

6. Classification and Evolution of Bacteria

Bacteria vary greatly in their chemistry and metabolism, and it is difficult to classify them in a rational way. In higher organisms species are often defined as forms that cannot interbreed and produce fertile offspring, but such a criterion is meaningless for bacteria whose reproduction is largely asexual and which are able readily to accept "visiting genes" from other bacteria.[12] The classification into species and genera is therefore somewhat arbitrary. A currently used scheme (Table 1-1)[20] classifies the prokaryotes into 35 groups on the basis of many characteristics including shape, staining behavior, and chemical activities. Table 1-1 also includes genus names of most of the bacteria discussed in this book.

Bacteria may have the shape of spheres or straight or curved rods. Some, such as the **actinomycetes**, grow in a branching filamentous form. Words used to describe bacteria often refer to these shapes: a **coccus** is a sphere, a **bacillus** a rod, and a **vibrio** a curved rod with a flagellum at one end. A **spirillum** is screw-shaped. These same words are frequently used to name particular genera or families. Other names are derived from some chemical activity of the bacterium being described.

The **gram stain** provides an important criterion of classification that depends upon differences in the structure of the cell wall (see Chapter 20). Bacterial cells are described as **gram-positive** or **gram-negative** according to their ability to retain the basic dye crystal violet as an iodine complex. This difference distinguishes two of four large categories of bacteria.[20] Most actinomycetes, the spore-forming bacilli, and most cocci are gram-positive, while *E. coli*, other enterobacteria, and pseudomonads are gram-negative. A third category consists of eubacteria that lack cell walls, e.g. the mycoplasma.

Comparisons of amino acid sequences of proteins and the nucleotide sequences of DNA and RNA have provided a new approach to classification of bacteria.[21–28] Although the origins of life are obscure, we can easily observe that the genome changes with time through mutation and through the enzyme-catalyzed process of **genetic recombination**. The latter gives rise to the deletion of some nucleotides and the insertion of others into a DNA chain. When we examine sequences of closely related species, such as *E. coli* and *Salmonella typhimurium,* we find that the sequences are very similar. However, they differ greatly from those of many other bacteria. Consider the 23S ribosomal RNA, a molecule found in the ribosomes of all bacteria. It contains ~3300 nucleotides in a single highly folded chain. The basic structure is highly conserved but between any two species of bacteria there are many nucleotide substitutions caused by mutations as well as deletions and insertions. By asking what is the minimum number of mutations that could have converted one 23S RNA into another and by assuming a more or less constant rate of mutation over millions of years it is possible to construct a **phylogenetic tree** such as that shown in Fig. 1-5.

One conclusion from these comparisons is that the methane-producing bacteria, the **methanogens**,[24] are only distantly related to most other bacteria. Methanogens together with the cell wall-less *Thermoplasma*,[28] some salt-loving **halobacteria**, and some **thermophilic** (heat-loving) sulfur bacteria form a fourth major category. They are often regarded as a separate kingdom, the **archaeobacteria**,[25] which together with the kingdom of the eubacteria form the superkingdom prokaryota. Certain archaeobacteria have biochemical characteristics resembling those of eukaryotes and

TABLE 1-1
A Systematic Classification Scheme for Bacteria[a,b]

Kingdom Procaryotae

The bacteria are classified according to the following 35 groups. Within these groups many genera are classified into subgroups or families. A few genera, most of which are mentioned elsewhere in this book, are listed here by name. Members of a single subgroup are placed together and are separated by semicolons from members of other subgroups.

1. The spirochetes (long bacteria, up to 500 µm, that are propelled by the action of filaments wrapped around the cell between the membrane and wall). *Borrelia* (*B. burgdorferi*, Lyme disease), *Leptospira*, *Treponema* (*T. pallidum*, syphilis)
2. Aerobic spiral and curved motile gram-negative bacteria. *Bdellovibrio*, *Campylobacter* (*C. jejuni*, diarrhea), *Helicobacter* (*H. pylori,* gastric ulcers), *Spirillum*
3. Nonmotile gram-negative curved bacteria
4. Gram-negative, aerobic rods and cocci. *Acetobacter, Agrobacterium, Azotobacter, Brucella* (*B. abortus*, brucellosis), *Flavobacterium, Gluconobacter, Legionella* (*L. pneumophila*, Legionnaire's disease), *Methylomonas*, *Neisseria* (*N. gonorrhea*, gonorrhea), *Pseudomonas*, *Rhizobium*, *Thermus*, *Xanthomonas*, *Rochalimaea* (*R. henselae*, cat scratch disease)
5. Gram-negative, facultatively anaerobic rods. *Enterobacter*,[c] *Proteus*, *Yersinia* (*Y. pestis*, plague), *Escherichia, Klebsiella, Salmonella* (*S. typhi*, typhoid fever), *Serratia, Shigella* (*S. dysenteriae*, bacterial dysentery), *Haemophilus; Vibrio* (*V. cholerae*, Asiatic cholera); *Zymomonas*
6. Gram-negative, anaerobic bacteria. *Butyrivibrio*
7. Dissimilatory sulfate- or sulfur-reducing bacteria. *Desulfovibrio*
8. Anaerobic gram-negative cocci. *Veillonella*
9. The rickettsias (parasitic bacteria with exacting nutritional requirements and small genome sizes) and chlamydias. *Chlamydia* (*C. trachomatis*, trachoma), *Rickettsia* (*R. rickettsii*, Rocky Mountain spotted fever)
10. Anoxygenic photosynthetic bacteria. Green sulfur bacteria. *Chlorobium, Prosthecochloris;* purple nonsulfur bacteria: *Rhodopseudomonas, Rhodospirillum;* purple sulfur bacteria: *Chromatium, Thiospirillum*
11. Oxygenic photosynthetic bacteria. Cyanobacteria (blue-green algae): *Synechocystis; Anabaena, Nostoc; Oscillatoria*
12. Aerobic, chemolithotrophic bacteria. Colorless sulfur bacteria: *Thiobacillus;* iron or manganese-oxidizing bacteria, magnetotactic bacteria; nitrifying bacteria: *Nitrobacter, Nitrosomonas*
13. Budding and/or appendaged bacteria. *Caulobacter*
14. Sheathed bacteria
15. Nonphotosynthetic, nonfruiting gliding bacteria. *Beggiatoa* (a filamentous bacterium containing sulfur granules)
16. Fruiting, gliding bacteria. *Myxococcus*
17. Gram-positive cocci. *Leuconostoc, Micrococcus, Peptococcus, Staphylococcus* (*S. aureus,* boils, infections), *Streptococcus* (*S. pyogenes,* scarlet fever, throat infections, *S. pneumoniae,* pneumonia)
18. Endospore-forming gram-positive rods and cocci. Aerobic: *Bacillus* (*B. anthracis,* anthrax), anaerobic: *Clostridium* (*C. tetani,* tetanus; *C. botulinum,* botulism)
19. Regular nonsporing gram-positive rods. *Lactobacillus*
20. Irregular nonsporing gram-positive rods. *Actinomyces, Bifidobacterium, Corynebacterium* (*C. diphtheriae,* diphtheria), *Propionibacterium*
21. Mycobacteria. *Mycobacterium* (*M. tuberculosis,* tuberculosis; *M. leprae,* leprosy)

22-29. Actinomycetes

30. Mycoplasmas. *Acholeplasma, Mycoplasma*
31. Methanogens. *Methanobacterium; Methanosarcina; Methanospirillum*
32. Archaeal sulfate reducers
33. Halobacteria. *Halobacterium*
34. Cell wall-less archaeobacteria. *Thermoplasma*
35. Very thermophilic S^0-Metabolizers. *Sulfolobus; Thermococcus*

---

[a] From Bergey's Manual of Systematic Bacteriology, 9th ed. J. G. Holt, N. R. Krieg, P. H. A. Sneath, J. T. Staley and S. T. Williams, Eds. (1994) Williams and Wilkins, Baltimore, Maryland. For another recent list see http://www.ncbi.nlm.nih.gov/

[b] The human diseases caused by some species are also listed.

[c] Formerly *Aerobacter*.

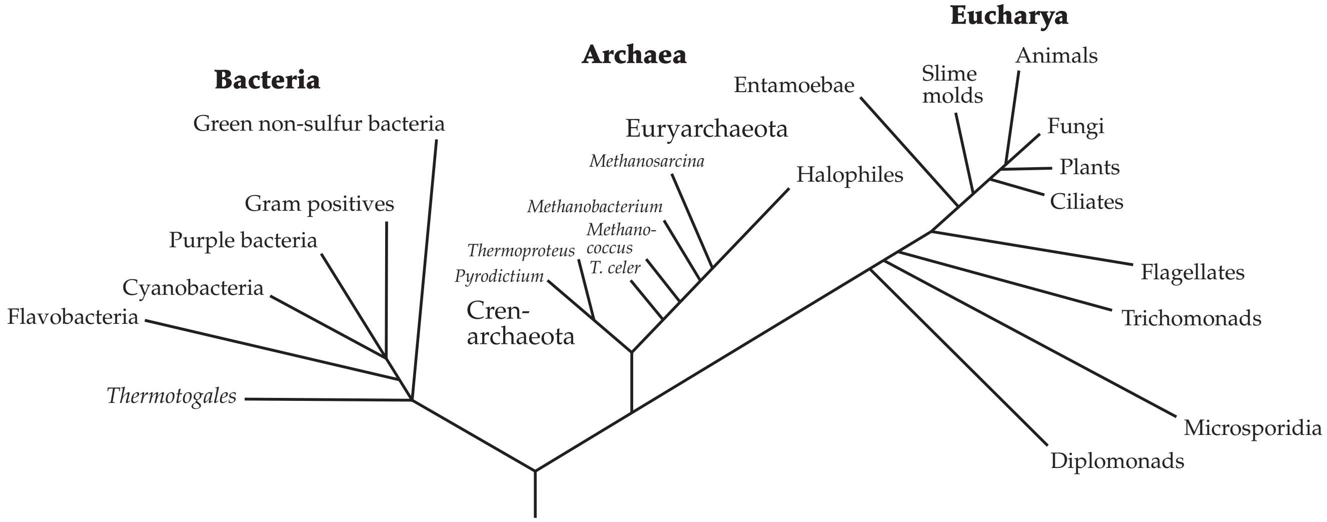


Figure 1-5 Universal phylogenetic tree. From Wheelis *et al.*[29]

some biologists therefore classify them as **archaea** and rank their kingdom as equal to that of the bacteria and the eukaryotes (Fig. 1-5).[27,29,30,30a,30b] Others disagree.[31] In Table 1-1, the archaeobacteria are found in groups 31–35. Most bacteria are very small in size but there are species large enough to be confused with eukaryotic protozoa. The record for bacteria seems to be held by *Epulopiscium fishelsoni,* a parasite of the surgeonfish intestinal tract. A single cell measured > 600 µm by 80 µm diameter, over 10^6 times larger in volume than a cell of *E. coli*.[32] The organism is a gram-positive bacterium as judged by analysis of its cloned ribosomal RNA genes.

7. Nutrition and Growth of Bacteria

Autotrophic (self-nourishing) bacteria can synthesize all of their organic cell constituents from carbon dioxide, water, and inorganic forms of nitrogen and sulfur. The **photoautotrophs** extract their energy from sunlight, while the **chemoautotrophs** obtain energy from inorganic chemical reactions. For example, the hydrogen bacteria oxidize H_2 to H_2O and sulfur bacteria oxidize H_2S to H_2SO_4. Like the fungi and animals, most bacteria are **chemoheterotrophic**; they obtain energy from the breakdown of organic compounds. Some of these heterotrophic bacteria are **anaerobes** which live without O_2. Many of them metabolize complex organic substances such as sugars in the absence of oxygen, a process called **fermentation**. Others oxidize organic compounds with an inorganic oxidant such as nitrite or sulfate. Members of the genus *Clostridium* are poisoned by oxygen and are known as **obligate anaerobes**. Others, including *E. coli*, are **facultative anaerobes**, able to grow either in the presence or in the absence of oxygen. **Obligate aerobes** depend for energy upon combustion of organic compounds with oxygen.

One of the largest groups of strictly aerobic heterotrophic bacteria, the pseudomonads (*Pseudomonas* and related genera), are of interest to biochemists because of their ability to oxidize organic compounds, such as alkanes, aromatic hydrocarbons, and steroids, which are not attacked by most other bacteria. Often, the number of oxidative reactions used by any one species of bacteria is limited. For example, the acetic acid bacteria that live in wine and beer obtain all of their energy by oxidation of ethanol to acetic acid:

$$CH_3CH_2OH + O_2 \rightarrow CH_3COOH + H_2O$$

Bacteria can grow incredibly fast. Under some conditions, it takes a bacterial cell only 10–20 min to double its size and to divide to form two cells.[4] An animal cell may take 24 h for the same process. Equally impressive are the rates at which bacteria transform their foods into other materials. One factor contributing to the high rate of bacterial metabolism may be the large surface to volume ratio. For a small spherical bacterium (coccus) of diameter 0.5 µm, the ratio of the surface area to the volume is 12 x 10^6 m^{-1}, while for an ameba of diameter 150 µm the ratio is only 4 x 10^4 m^{-1} (the ameba can increase this by sticking out some pseudopods). Thimann[33] estimated that for a 90-kg human, the ratio is only 30 m^{-1}.

When food is limited, some bacteria such as the *Bacillus* form **spores**. These are compact little cells that form inside the vegetative cell and are therefore called **endospores**. They sometimes have only 1/10 the volume of the parent cell. Their water content is very low, their metabolic rate is near zero, and they are extremely resistant to heat and further desiccation. Under suitable conditions, the spores can "germinate" and renew their vegetative growth. Spore formation is one of several examples of the development of specialized cells or **differentiation** among prokaryotes.

8. Photosynthetic and Nitrogen-Fixing Prokaryotes

It is likely that the earth was once a completely anaerobic place containing water, ammonia, methane, formaldehyde, and more complicated organic compounds. Perhaps the first forms of life, which may have originated about 3.5 x 10^9 years ago, resembled present-day anaerobic bacteria. The purple and green photosynthetic bacteria may be related to organisms that developed at a second stage of evolution: those able to capture energy from sunlight. Most of these gram-negative photosynthetic bacteria are strict anaerobes. None can make oxygen as do higher plants. Rather, the hydrogen needed to carry out the reduction of carbon dioxide in the photosynthetic process is obtained by the splitting of inorganic compounds, such as H_2S, thiosulfate, or H_2, or is taken from organic compounds. Today, photosynthetic bacteria are found principally in sulfur springs and in deep lakes, but at one time they were probably far more abundant and the only photosynthetic organisms on earth.

Before organisms could produce oxygen a second complete photosynthetic system, which could cleave H_2O to O_2, had to be developed. The simplest oxygen-

BOX 1-C IN THE BEGINNING

No one knows how life began. Theories ranging from the biblical accounts to recent ideas about the role of RNA are plentiful but largely unsatisfying. In the 1800s the great physical chemist Arrhenius was among scientists that preferred the idea held by some scientists today that a "seed" came from outer space. Until recently the only concrete data came from fossils. Making use of a variety of isotopic dating methods it can be concluded that cyanobacteria were present 2.2×10^9 years ago and eukaryotes 1.4×10^9 years ago. About 0.5×10^9 years ago the "Cambrian explosion" led to the appearance of virtually all known animal phyla. Many of these then became extinct about 0.2×10^9 years ago.

New insights published in 1859[a] were provided by Charles Darwin. However, his ideas were only put into a context of biochemical data after 1950 when sequencing of proteins and later nucleic acids began. From an astonishingly large library of sequence data available now we can draw one firm conclusion: *Evolution can be observed;*[b] *it does involve mutation of DNA*. Comparisons of sequences among many species allow evolutionary relationships to be proposed.[c-e] In general these are very similar to those deduced from the fossil record. They support the idea that evolution occurs by natural selection and that duplication of genes and movements of large pieces of DNA within the genome have occurred often. As many as 900 "ancient conserved regions" of DNA in the *E. coli* genome corresponding to those in human, nematode, and yeast DNA are thought to date back perhaps 3.5×10^9 years.[f] However, nobody has explained how life evolved before there was DNA.

One of the first scientists to devote his career to biochemical evolution was I. V. Oparin,[g] who published a book on the "origin of life" in 1924. Oparin and J. B. S. Haldane, independently, proposed that early life was anaerobic and that energy was provided by fermentation. In 1951 Stanley Miller built an apparatus that circulated CH_4, NH_3, H_2O, and H_2, compounds thought to be present in a primitive atmosphere, past an electric discharge. He found glycine, alanine, β-alanine, and other amino acids among the products formed.[h] Schrödinger pointed out that a flux of energy through a system will tend to organize the system. The solar energy passing through the biosphere induces atmospheric circulation and patterns of weather and ocean currents.[i,j] Perhaps in the primordeal oceans organic compounds arose from the action of light and lightning discharges. These compounds became catalysts for other reactions which eventually evolved into a rudimentary cell-less metabolism. It is a large jump from this to a cell! Among other problems is the lack of any explanation for the development of individual cells or of their genomes. However, because it helps to correlate much information *we will always take an evolutionary approach in this book* and will discuss the "beginnings" a little more in later chapters.

---

[a] Maynard-Smith, J. (1982) *Nature (London)* **296**, 599–601

[b] Lenski, R. E., and Travisano, M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6808–6814

[c] Wilson, A. C. (1985) *Sci. Am.* **253**(Oct), 164–173

[d] Eigen, M., Gardiner, W., Schuster, P., and Winkler-Oswatitsch, R. (1981) *Sci. Am.* **244**(Apr), 88–118

[e] Doolittle, R. F. (1992) *Protein Sci.* **1**, 191–200

[f] Green, P., Lipman, D., Hillier, L., Waterston, R., States, D., and Claverie, J.-M. (1993) *Science* **259**, 1711–1716

[g] Broda, E. (1980) *Trends Biochem. Sci.* **5**, IV–V

[h] Miller, S. L. (1953) *Science* **117**, 528–529

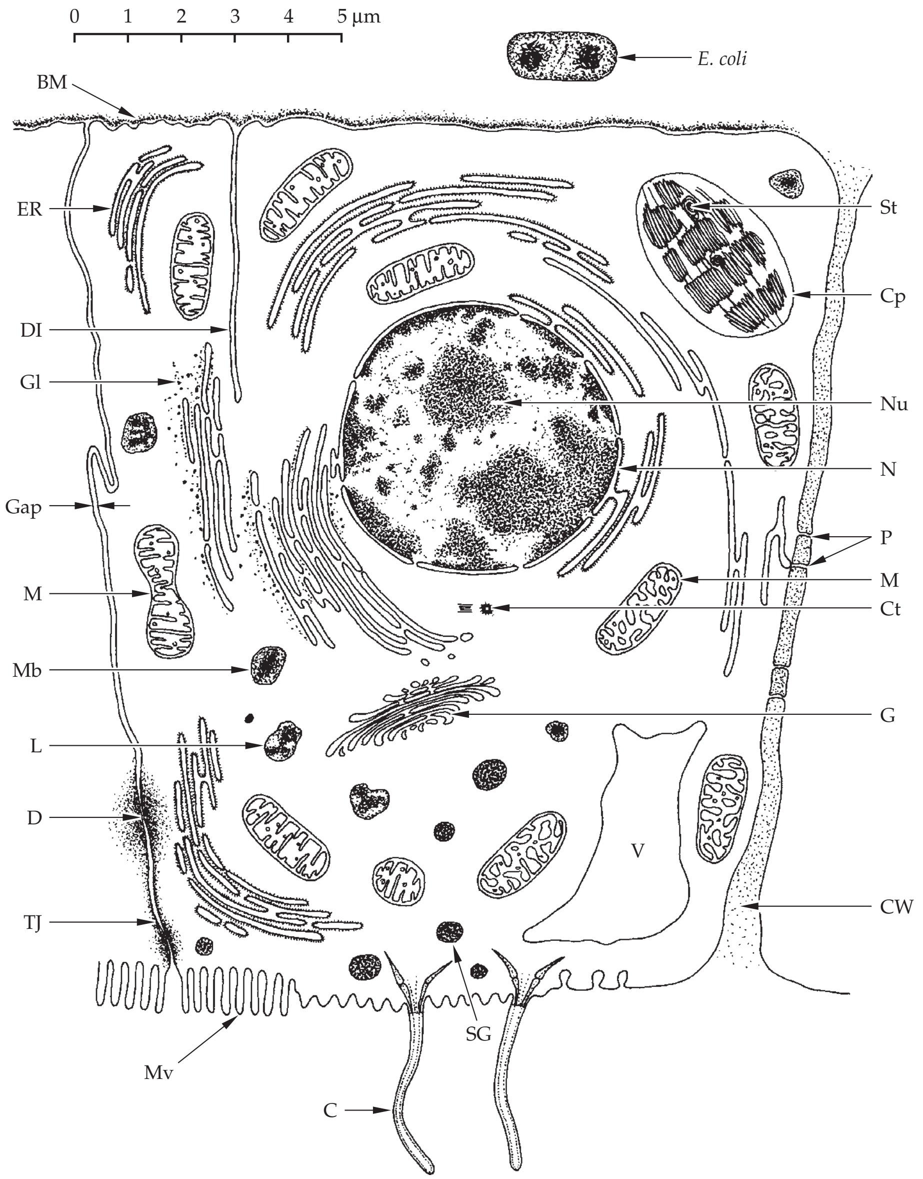
[i] Mason, S. (1993) *Trends Biochem. Sci.* **18**, 230–231

[j] Welch, G. R. (1996) *Trends Biochem. Sci.* **21**, 452

producing creatures existing today are the **cyanobacteria**,[34] also known as blue-green algae. Many cyanobacteria are unicellular, but others such as *Oscillatoria*, a slimy "plant" that often coats the inside walls of household aquaria, consist of long filaments about 6 µm in diameter (see Fig. 1-11). All cyanobacteria contain two groups of pigments not found in other prokaryotes: **chlorophyll a** and **β-carotene**, pigments that are also found in the chloroplasts of true algae and in higher plants. A recently discovered group of bacteria, the **prochlorophytes**, are even closer to chloroplasts in their pigment composition.[35]

In addition to pigmented cells, some cyanobacteria contain paler cells known as **heterocysts**. They have a specialized function of fixing molecular nitrogen. The development of the ability to convert N_2 into organic nitrogen compounds represents another important evolutionary step. Because they can both fix nitrogen and carry out photosynthesis, the blue-green algae have the simplest nutritional requirements of any organisms. They need only N_2, CO_2, water, light, and minerals for growth.

Evolution of the photosynthetic cleavage of water to oxygen was doubtless a major event with far-reaching consequences. Biologists generally believe that as oxygen accumulated in the earth's atmosphere, the obligate anaerobes, which are poisoned by oxygen, became limited to strictly anaerobic environments. Meanwhile, a new group of bacteria, the **aerobes**, appeared with mechanisms for detoxifying oxygen



Abbreviations:

BM, basement membrane
ER, rough endoplasmic reticulum (with ribosomes attached; smooth ER is depicted nearer the nucleus and on the right side of the cell.)
DI, deep indentation of plasma membrane
GI, glycogen granules
Gap, space ~10-20 nm thick between adjacent cells
M, mitochondrion
Mb, microbody
L, lysosome
D, desmosome
TJ, tight junction
Mv, microvilli
C, cillium
SG, secretion granule
V, vacuole
Nu, nucleolus
G, Golgi apparatus
CW, cell wall (of a plant)
Ct, centrioles
P, plasmodesmata
N, nucleus
Cp, chloroplast
St, starch granule

Figure 1-6 The "average" eukaryotic cell. This composite drawing shows the principal organelles of both animal and plant cells approximately to the correct scale. (Adapted from a drawing by Michael Metzler.)

and for using oxygen to oxidize complex organic compounds to obtain energy.

B. Eukaryotic Cells

Cells of the **eukaryotes** contain true nuclei and are much larger and more complex internally than are those of prokaryotes. The nucleus of a cell contains most of its DNA and is separated from the cytoplasm by membranes. Within the cytoplasm are various **organelles** with characteristic structures. These include **mitochondria**, **lysosomes**, **peroxisomes,** and **centrioles**. Eukaryotic cells come in so many sizes and shapes and with so many specialized features that it is impossible to say what is typical. Nevertheless, Fig. 1-6 is an attempt to portray some sort of "average" cell, partly plant and partly animal.

As can be seen from Table 1-2, which lists the diameters and volumes of several roughly spherical cells, there is a great variation in size. However, a diameter of 10–20 µm may be regarded as typical for both plants and animals. For growth of a large cell such as the ovum, many adjacent cells assist in synthesis of foodstuffs which are transferred to the developing egg cell. Plant cells are often large but usually 90% or more of the cell is filled with a **vacuole** or **tonoplast**,[36] which is drawn unrealistically small in Fig. 1-6. The metabolically active protoplasm of plant cells often lies in a thin layer at their peripheries.

Many cells are far from spherical; for example, human red blood cells are discs 8 x 8 x 1 to 2 µm with a volume of 80 μm^3. Plant fiber cells may be several millimeters in length. Nerve cells of animals have long extensions, the **axons**, which in the human sometimes attain a length of a meter. Muscle cells fuse to give very long multinucleate fibers.

TABLE 1-2
Approximate Sizes of Some Cells

Cell	Diameter (µm)	Approximate volume (μm^3)
E. coli	1	1.0
Small thymus cell	6	120
Liver cell	20	4,000
Human ovum (mature)	120	500,000
Hen's egg (white excluded)	20,000	4×10^{12}
Yeast cell	10	500
Onion root (meristematic cell)	17	2,600
Parenchyma cell of a fruit	1,000	1×10^8

1. The Nucleus

In a typical animal cell the nucleus has a diameter of ~5 µm and a volume of 65 μm^3. Except at the time of cell division, it is densely and almost uniformly packed with DNA. The amount of DNA present is larger than that in bacteria as is indicated in Table 1-3. Yeast contains about three times as much genetic matter as *E. coli* and a human being or a mouse about 700 times as much. However, genes are sometimes duplicated in higher organisms and large amounts of **repetitive DNA** of uncertain significance are often present. Some amphibians have 25 times *more* DNA per cell than do humans. The fruit fly *Drosophila* contains about 13,600 functioning genes and a human being perhaps 50,000.[37]

Because of its acidic character, DNA is stained by basic dyes. Long before the days of modern biochemistry, the name **chromatin** was given to the material in the nucleus that was colored by basic dyes. At the time of cell division, the chromatin is consolidated into distinct **chromosomes** which contain, in addition to 15% DNA, about 10% RNA and 75% protein.

Nearly all of the RNA of the cell is synthesized (transcribed) in the nucleus, according to the instructions encoded in the DNA. Some of the RNA then moves out of the nucleus into the cytoplasm where it functions in protein synthesis and in some other ways. Many eukaryotic genes consist of several sequences that may be separated in the DNA of a chromosome by **intervening sequences** of hundreds or thousands of base pairs. The long RNA transcripts made from these **split genes** must be cut and **spliced** in the nucleus to form the correct messenger RNA molecules which are then sent out to the ribosomes in the cytoplasm.

Each cell nucleus contains one or more dense **nucleoli**, regions that are rich in RNA and may contain 10–20% of the total RNA of cells. Nucleoli are sites of synthesis and of temporary storage of ribosomal RNA, which is needed for assembly of ribosomes. The **nuclear envelope** is a pair of membranes, usually a few tens of nanometers apart, that surround the nucleus. The two membranes of the pair separate off a thin **perinuclear space** (Fig. 1-7). The membranes contain "pores" ~130 nm in diameter with a complex structure (see Fig. 27-8).[38,39] There is a central channel ~42 nm in diameter, which provides a route for controlled passage of RNA and other large molecules from the nucleus into the cytoplasm and also from the cytoplasm to the nucleus. Smaller ~10 nm channels allow passive diffusion of ions and small molecules.

TABLE 1-3
Haploid Genome Sizes for Several Organisms

Organism; see footnotes for sequence information	Millions of nucleotide base pairs (Mb)	Number of chromosomes (haploid)	Estimated number of genes
Mycoplasma genitalium[a–c]	0.580	1	482
Rickettsia prowazekii[d]	1.11	1	834
Haemophilus influenzae[c,e–g]	1.83	1	1,709
Methanococcus janaschii (an archaeon)[h]	1.66	1	1,738
Bacillus subtilis[i]	4.16	1	
Escherichia coli[j,k]	4.64	1	4,288
Myxococcus xanthis[f]	9.2	1	
Synechocystis sp. (a cyanobacterium)[l]	3.57	1	3,169
Saccharomyces cerevisiae (a yeast)[k,m-p,t]	13.5	17	6,241
Giardia lamblia (a protozoan)[q]	12		
Plasmodium falciparum (malaria parasite)[v]	25-30	14	
Dictyostelium discoideum (a slime mold)[r]	34		
Caenorhabditis elegans (a nematode)[s,t]	97	6	18,424
Sea Urchin	900		
Drosophila melanogaster (fruit fly)[u,t]	180	4	13,601
Danio rerio (zebrafish)[w]	1,700	2.5	
Fugu rubripes (pufferfish)[x]	400		60,000
S. African lungfish[y]	102,000	19	
Mus musculus (mouse)[z]	~3,000	20	80,000
Bos (cow)[aa]	~3,000	30	80,000
Homo sapiens (human)[bb,cc]	~3,000	23	50,000–150,000
Arabidopsis thaliana (green plant)[dd,ee]	115.4	5	25,498
Rice[ff-hh]	450	12	
Maize[ii] or Wheat[ff]	~2,700		
Lily[jj]	>100,000		

2. The Plasma Membrane

The thin (8 nm) outer cell membrane or "plasma-lemma" (Fig. 1-7) controls the flow of materials into and out of cells, conducts impulses in nerve cells and along muscle fibrils, and participates in chemical communication with other cells. Deep infoldings of the outer membrane sometimes run into the cytoplasm. An example, is the "T system" of tubules which functions in excitation of muscle contraction (Figs. 19-7, 19-21). Surfaces of cells designated to secrete materials or to absorb substances from the surrounding fluid, such as the cells lining kidney tubules and pancreatic secretory cells, are often covered with very fine projections or **microvilli** which greatly increase the surface area. In other cases projections from one cell interdigitate with those of an adjacent cell to give more intimate contact.

3. Vacuoles, Endocytosis, and Lysosomes

Cells often contain vacuoles or smaller vesicles that are separated from the cytosol by a *single* membrane. Their content is often quite

---

[a] Fraser, C. M., and 28 other authors (1995) *Science* **270**, 397–403
[b] Goffeau, A. (1995) *Science* **270**, 445–446
[c] Brosius, J., Robison, K., Gilbert, W., Church, G. M., and Venter, J. C. (1996) *Science* **271**, 1302–1304
[d] Andersson, S. G. E., Zomorodipour, A., Andersson, J. O., Sicheritz-Pontén, T., Alsmark, U. C. M., Podowski, R. M., Näslund, A. K., Eriksson, A.-S., Winkler, H. H., and Kurland, C. G. (1998) *Nature (London)* **396**, 133–140
[e] Fleischmann, R. D., and 39 other authors (1995) *Science* **269**, 496–512
[f] He, Q., Chen, H., Kupsa, A., Cheng, Y., Kaiser, D., and Shimkets, L. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9584–9587
[g] Mrázek, J., and Karlin, S. (1996) *Trends Biochem. Sci.* **21**, 201–202
[h] Bult, C. J., and 39 other authors; corresponding author Venter, J. C. (1996) *Science* **273**, 1058–1073
[i] Azevedo, V., Alvarez, E., Zumstein, E., Damiani, G., Sgaramella, V., Ehrlich, S. D., and Serror, P. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6047–6051
[j] Blattner, F. R. and 16 other authors (1997) *Science* **277**, 1453–1462
[k] Winzeler, E. A., and 51 other authors (1999) *Science* **285**, 901–906
[l] Kaneko, T., and 23 other authors (1996) *DNA Res.* **3**, 109–136. See also http://www.kazusa.or.jp/cyano
[m] Dujon, B., and 107 other authors (1994) *Nature (London)* **369**, 371–378
[n] Taguchi, T., Seko, A., Kitajima, K., Muto, Y., Inoue, S., Khoo, K.-H., Morris, H. R., Dell, A., and Inoue, Y. (1994) *J. Biol. Chem.* **269**, 8762–8771
[o] Johnston, M., and 34 other authors. (1994) *Science* **265**, 2077–2082
[p] Williams, N. (1995) *Science* **268**, 1560–1561
[q] Sogin, M. L., Gunderson, J. H., Elwood, H. J., Alonso, R. A., and Peattie, D. A. (1989) *Science* **243**, 75–77
[r] Loomis, W. F., and Insall, R. H. (1999) *Nature (London)* **401**, 440–441
[s] The *C. elegans* Sequencing Consortium (1998) *Science* **282**, 2012–2018 (See this article for list of authors.)
[t] Rubin, G. M., and 54 other authors. (2000) *Science* **287**, 2204–2215
[u] Adams, M. D., and 194 other authors. (2000) *Science* **287**, 2185–2195
[v] Su, X.-z, Ferdig, M. T., Huang, Y., Huynh, C. Q., Liu, A., You, J., Wootton, J. C., and Wellems, T. E. (1999) *Science* **286**, 1351–1353
[w] Postlethwait, J. H., Johnson, S. L., Midson, C. N., Talbot, W. S., Gates, M., Ballinger, E. W., Africa, D., Andrews, R., Carl, T., Eisen, J. S., Horne, S., Kimmel, C. B., Hutchinson, M., Johnson, M., and Rodriguez, A. (1994) *Science* **264**, 699–703
[x] Fishman, M. C. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10554–10556
[y] Kornberg, A., and Baker, T. A. (1992) *DNA Replication*, 2nd ed., pp. 19–21, Freeman, New York
[z] Dietrich, W. F., Copeland, N. G., Gilbert, D. J., Miller, J. C., Jenkins, N. A., and Lander, E. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10849–10853
[aa] Anonymous (1994) *Nature (London)* **368**, 167
[bb] Schuler, G. D. and 102 other authors (1996) *Science* **274**, 540–546
[cc] Koonin, S. E. (1998) *Science* **279**, 36–37
[dd] Olson, M. V. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4338–4344
[ee] The Arabidopsis Genome Initiative, (2000) *Nature* **408**, 796–815
[ff] Stevens, J. E. (1994) *Science* **266**, 1186–1187
[gg] Shimamoto, K. (1995) *Science* **270**, 1772–1773
[hh] Singh, K., Ishii, T., Parco, A., Huang, N., Brar, D. S., and Khush, G. S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6163–6168
[ii] Carels, N., Barakat, A., and Gernardi, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11057–11060
[jj] Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) *Molecular Biology of the Cell*, 3rd ed., Garland, New York

acidic.[40] Small vesicles sometimes bud inward from the plasma membrane in a process called **endocytosis**. In this manner the cell may engulf particles (**phagocytosis**) or droplets of the external medium (**pinocytosis**). The resulting endocytotic vesicles or **endosomes** often fuse with **lysosomes**, which are small acidified vesicles containing a battery of enzymes powerful enough to digest almost anything in the cell. In cells that engulf bits of food (e.g., ameba) lysosomes provide the digestive enzymes. Lysosomes also take up and digest denatured or damaged proteins and may digest "worn out" or excess cell parts including mitochondria. Lysosomes are vital components of cells,[41] and several serious human diseases result from a lack of specific lysosomal enzymes.

4. The Endoplasmic Reticulum and Golgi Membranes

Although cytoplasm is fluid and in some organisms can undergo rapid streaming, the electron microscope

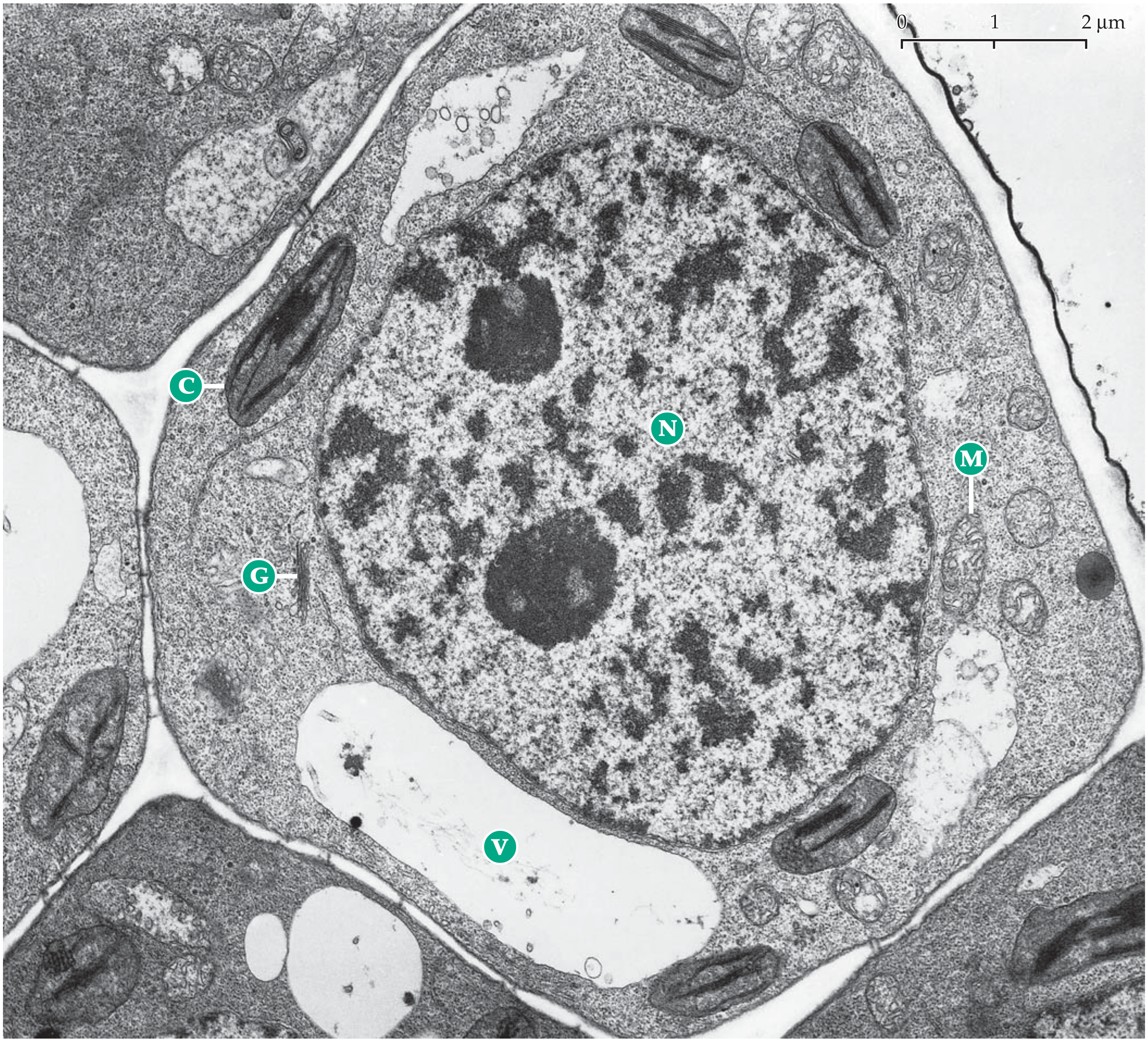
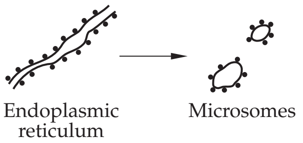


Figure 1-7 Electron micrograph of a thin section of a young epidermal cell of a sunflower. The tissue was fixed and stained with uranyl acetate and lead citrate. Clearly visible are the nucleus (N), mitochondria (M), chloroplasts (C), a Golgi body dictyosome (G), endoplasmic reticulum, vacuole (V), cell wall, plasmodesmata, and cuticle (upper right, thin dark layer). Micrograph courtesy of H. T. Horner.

has revealed that within the liquid portion, the **cytosol**, there is a complex network of membranes known as the **endoplasmic reticulum** (ER). The membranes of the ER form tubes, vesicles, and flattened sacs called **cisternae**. The intracisternal spaces appear to connect with the perinuclear space and to a series of 3–12 flattened, slightly curved disk-shaped membranes known as the **Golgi apparatus** (Figs. 1-7, 20-8).[42,43] This organelle was first reported by Camillo Golgi in 1898.[44] Its existence was long doubted, but it is known now to play a vital role in metabolism.

The ER, the Golgi membranes, and secretion granules apparently represent an organized system for synthesis of secreted protein and formation of new membranes. Parts of the ER, the **rough endoplasmic reticulum** are lined with many ribosomes of 21–25 nm diameter. While resembling those of bacteria, these eukaryotic ribosomes are about 50% heavier (4×10^6 Da). The **smooth endoplasmic reticulum** lacks ribosomes but proteins made in the rough ER may be modified in the smooth ER, e.g., by addition of carbohydrate chains. Small membrane vesicles break off from the smooth ER and pass to the Golgi membranes which lie close to the smooth ER on the side toward the center of the cell. Here additional modification reactions occur (Chapter 20). At the outer edges the membranes of the Golgi apparatus pinch off to form vacuoles which are often densely packed with enzymes or other proteins. These **secretion granules** move to the surface and are released from the cell. In this process of **exocytosis** the membranes surrounding the granules fuse with the outer cell membrane. The rough ER appears to contribute membrane material to the smooth ER and Golgi apparatus, while material from Golgi membranes can become incorporated into the outer cell membrane and into lysosomes. Outer mitochondrial membranes and membranes around vacuoles in plant cells may also be derived directly from the ER. Outer membrane materials are probably "recycled" by endocytosis.

The term **microsome**, frequently met in the biochemical literature, refers to small particles of 50–150 nm diameter which are mostly fragments of the ER together with some material from the plasma membrane. Microsomes are formed when cells are ground or homogenized. Upon centrifugation of the disrupted cells, nuclei and other large fragments sediment first, then the mitochondria. At very high speeds (e.g., at 100,000 times the force of gravity) the microsomes, whose masses are 10^8–10^9 Da, settle. With the electron microscope we see that in the microsomes the membrane fragments have closed to give small sacs to the outside of which the ribosomes still cling:



5. Mitochondria, Plastids, and Peroxisomes

Mitochondria, complex bodies about the size of bacteria and bounded on the outside by a double membrane (Figs. 1-6 and 1-7), are present in all eukaryotic cells that use oxygen for respiration. The numbers per cell appear to vary from the *one* for certain tiny trypanosomes to as many as 3×10^5 in some oocytes. Liver cells often contain more than 1000 mitochondria apiece.[45] Study of ultrathin serial sections of a single yeast cell by electron microscopy has shown that under some growth conditions all of the yeast mitochondria are interconnected.[46] More recent evidence from new imaging procedures, e.g. using the green fluorescent protein (Box 23-A)[46a, 46b] also supports the idea that mitochondria are interconnected in a reticulum that can become fragmented under some conditions. The inner membrane of a mitochondrion is often highly folded to form the **cristae** (crests). The outer membrane is porous to small molecules but the passage of substances into and out of the inner space of the mitochondrion, known as the **matrix,** is tightly controlled by the inner membrane. Although some of the oxidative chemical activites of the cells are located in the ER and in peroxisomes, the major energy-yielding reactions for aerobic organisms are found in the mitochondria, which are also the principal site of utilization of oxygen. Within each mitochondrion is a small circular molecule of DNA whose genes encode only a few of the many proteins needed in this organelle. Also present within mitochondria are ribosomes of a size similar to those of bacteria and smaller than those lining the rough ER.

Plastids are organelles of plant cells that serve a variety of purposes.[45] Most important are the **chloroplasts**, the chlorophyll-containing sites of photosynthesis. Like mitochondria they contain folded internal membranes (see Fig. 23-19) and several small molecules of DNA.

Fragile organelles, the **peroxisomes** or **microbodies,** occur in many cells.[47–50] In green leaves they may occur in numbers up to one-third those of mitochondria. Peroxisomes are often about the size of mitochondria but have only a single membrane and do not contain DNA. They often contain an apparently crystalline "core." The single membrane of peroxisomes is porous to small molecules such as sucrose. This permits these organelles to be separated from mitochondria by centrifugation in a sucrose gradient where the microbodies assume a density of about 1.25 g/cm^3 compared to 1.19 for the impervious mitochondria.

Peroxisomes are rich in enzymes that produce and decompose hydrogen peroxide. They often make a major contribution to the oxidative metabolism of cells. In germinating oilseeds **glyoxysomes**, a type of peroxisome, contain enzymes that catalyze reactions of the biosynthetic "glyoxylate pathway" of metabolism.[51] Organelles that resemble peroxisomes in appearance but which are functionally more closely related to mitochondria are the **hydrogenosomes** of anaerobic protozoa.[52] As the name suggests, these organelles are the site of formation of molecular hydrogen, a common product of anaerobic metabolism.

6. Centrioles, Cilia, Flagella, and Microtubules

Many cells contain **centrioles**,[53] little cylinders about 0.15 µm in diameter and 0.5 µm long, which are *not* enclosed by membranes. Each centriole contains a series of fine **microtubules** of 25 nm diameter. A pair of centrioles are present near the nucleus in most animal cells and play an important role in cell division. Together with surrounding materials they form the **centrosome**. However, centrioles have never been observed in plant cells.

Related in structure to centrioles are the long **flagella** and shorter **cilia** (the two words are virtually synonymous) which are commonly present as organelles of locomotion in eukaryotic cells. Stationary cells of our own bodies also often have cilia. For example, there are 10^9 cilia/cm^2 in bronchial epithelium.[54] Modified flagella form the receptors of light in our eyes and of taste in our tongues. Flagella and cilia have a diameter of about 0.2 µm and a characteristic internal structure. Eleven hollow microtubules of ~24 nm diameter are usually arranged in a "9 + 2" pattern with nine pairs of fused tubules surrounding a pair of single tubules (Figs. 1-8 and 19-23). Each microtubule resembles a bacterial flagellum in appearance, but there are distinct and significant chemical differences. The **basal body** of the flagellum, the **kinetosome** (Fig. 1-8), resembles a centriole in structure, dimensions, and mode of replication. Recently a small 6–9 megabase pair DNA has been found in basal bodies of the protozoan *Chlamydomonas*.[55,56]

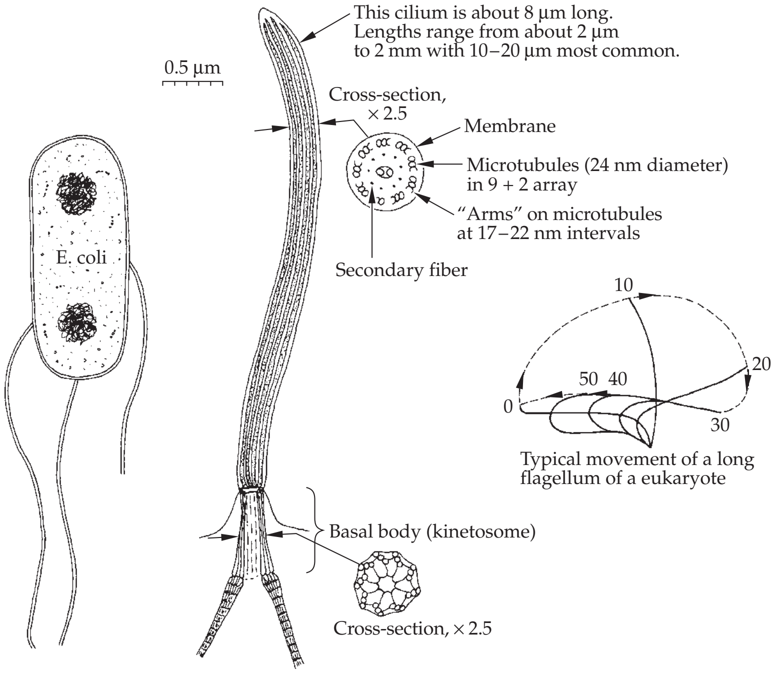

This cilium is about 8 µm long. Lengths range from about 2 µm to 2 mm with 10–20 µm most common.
0.5 µm
Cross-section, × 2.5
Membrane
Microtubules (24 nm diameter) in 9 + 2 array
"Arms" on microtubules at 17–22 nm intervals
Secondary fiber
E. coli
10
20
50 40
0
30
Typical movement of a long flagellum of a eukaryote
Basal body (kinetosome)
Cross-section, × 2.5

Figure 1-8 Structure of cilia and flagella of eukaryotes. After P. Satir.[57]

Microtubules similar to those found in flagella are also present in the cytoplasm. Together with thinner **microfilaments** of several kinds they form an internal **cytoskeleton** that provides rigidity to cells.[58,59] Microtubules also form the "spindle" of dividing cells. In nerve axons (Chapter 30) the microtubules run parallel to the length of the axons and are part of a mechanical transport system for cell constituents.

7. Cell Coats, Walls, and Shells

Like bacteria, most cells of higher plants and animals are surrounded by extracellular materials. Plants have rigid walls rich in cellulose and other carbohydrate polymers. Outside surfaces of plant cells are covered with a **cuticle** containing layers of a polyester called **cutin** and of wax (Fig. 1-6). Surfaces of animal cells are usually lined with carbohydrate molecules which are attached to specific surface proteins to form **glycoproteins**. Spaces between cells are filled with such "cementing substances" as **pectins** in plants and **hyaluronic acid** in animals. Insoluble proteins such as **collagen** and **elastin** surround connective tissue cells. Cells that lie on a surface (epithelial and endothelial cells) are often lined on one side with a thin, collagen-containing **basement membrane** (Figs. 1-6 and 8-31). Inorganic deposits such as calcium phosphate (in bone), calcium carbonate (eggshells and spicules of sponges), and silicon dioxide (shells of diatoms) are laid down, often by cooperative action of several or many cells.

C. Inheritance, Metabolic Variation, and Evolution of Eukaryotes

The striking differences between eukaryotic and prokaryotic cells have led to many speculations about the evolutionary relationship of these two great classes of living organisms. A popular theory is that mitochondria, which are characteristic of most eukaryotes, arose from aerobic bacteria. After cyanobacteria had developed and oxygen

BOX 1-D INHERITED METABOLIC DISEASES

Photomicrograph of human male metaphase chromosomes. © Photo Researchers

In 1908 Archibald Garrod[a,b] proposed that **cystinuria** (Chapter 8) and several other defects in amino acid and sugar metabolism were "inborn errors of metabolism", i.e. inherited diseases. Since that time the number of recognized genetic defects of human metabolism has increased at an accelerating rate to ~4000.[c–e] Hundreds of other genetic problems have also been identified. For over 800 of these the defective gene has been mapped to a specific chromosome.[f] An example is **sickle cell anemia** (Box 7-B) in which a defective hemoglobin differs from the normal protein at one position in one of its constituent polypeptide chains. Many other defects involve loss of activity of some important enzyme.

Most genetic diseases are rare, affecting about one person in 10,000. However, **cystic fibrosis** affects one in 2500. There are so many metabolic diseases that over 0.5% of all persons born may develop one. Many die at an early age. A much greater number (>5%) develop such conditions as diabetes and mental illness which are, in part, of genetic origin. Since new mutations are always arising, genetic diseases present a problem of continuing significance.

At what rate do new mutations appear? From the haploid DNA content (Table 1-2) we can estimate that the total coding capacity of the DNA in a human cell exceeds two million genes (actually two million *pairs* of genes in diploid cells). However, only a fraction of the DNA codes for proteins. There are perhaps 50,000 pairs of structural genes in human DNA. The *easily detectable* rate of mutation in bacteria is about 10^{-6} per gene, or 10^{-9} per base per replication.[g] As a result of sophisticated "proofreading" and repair systems, it may be as low as 10^{-10} per base in humans.[h] Thus, in the replication of the 3×10^9 base pairs in diploid human chromosomes we might anticipate about one mistake per cell division. Only about 1/50 of these would be in structural genes and potentially harmful. Thus, if there are 10^{16} division cycles in a normal life span[h] each parent may pass on to future generations about 2 mutations in protein sequences. The ~10^{14} body cells (somatic cells) also undergo mutations which may lead to cancer and to other problems of aging. Most mutations may be harmless or nearly so and a few may be beneficial. However, many are damaging and some are *lethal*. If a mutation is lethal, a homozygote will not survive and will be lost in an early (and usually undetected) spontaneous abortion. Healthy individuals carry as many as ten lethal recessive mutations as well as at least 3–5 autosomal recessive mutations of a seriously harmful type. Harmful dominant mutations are also frequent in the population. These include an elevated lipoprotein content of the blood and an elevated cholesterol level which are linked to early heart disease.

Biochemical disorders are also important because of the light they shed on metabolic processes. No other species is observed as carefully as *Homo sapiens*. As a consequence frequent reference will be made to genetic diseases throughout the book. A goal is to find ways to prevent or ameliorate the effects of these disorders. For example, in the treatment of **phenylketonuria** (Chapter 25) or of **galactosemia** (Chapter 20), a change in the diet can prevent irreversible damage to the brain, the organ most frequently affected by many of these diseases. Injection of a missing enzyme is giving life to victims of Gaucher's disease (Chapter 20). In many other cases no satisfactory therapy is presently available, but the possibilities of finding some way to supply missing enzymes or to carry out "genetic surgery" are among the most exciting developments of contemporary medical biochemistry (Chapter 26).

---

[a] Garrod, A. E. (1909) *Inborn Errors of Metabolism*, Oxford, London

[b] Bearn, A. G. (1993) *Archibald Garrod and the Individuality of Man*, Oxford, New York

[c] Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds. (1995) *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., McGraw-Hill, New York

[d] Davies, K. E. (1992) *Molecular Basis of Inherited Disease*, Oxford, New York

[e] McKusick, V. A. (1994) *Mendelian Inheritance in Man*, 11th ed., Johns Hopkins Univ. Press, Baltimore, Maryland

[f] McKusick, V. A., Amberger, J., and Steinberg, J. (1994) *J. NIH Res.* **6**, 115–134

[g] Watson, J. D. (1976) *Molecular Biology of the Gene*, 3rd ed., Benjamin, Menlo Park, California (p. 254)

[h] Koshland, D. E., Jr. (1994) *Science* **266**, 1925

had become abundant, a **symbiotic** relationship could have arisen in which small aerobic bacteria lived within cells of larger bacteria that had previously been obligate anaerobes. Sequence similarities of proteins suggest that these symbionts may have been related to present-day methanogens[60] and thermophilic sulfur bacteria.[61] The aerobes presumably used up any oxygen present, protecting the surrounding anaerobic organisms from its toxicity. The relationship became permanent and led eventually to the mitochondria-containing eukaryotic cell.[62–65] Further symbiosis with cyanobacteria or prochlorophytes could have led to the chloroplasts of the eukaryotic plants.

A fact that supports such ideas is the existence among present-day organisms of many endosymbiotic relationships. For example, the green paramecium (*Paramecium bursaria*) contains, within its cytoplasm, an alga (*Chlorella*), a common green plant that is quite capable of living on its own. Perhaps by accident it took up residence within the paramecium.[62] Some dinoflagellates (Fig. 1-9) contain endosymbiotic cyanobacteria[66] and recently a ciliate that contains endosymbiotic purple photosynthetic bacteria has been discovered.[67] These bacteria do not produce O_2 but utilize products of the host ciliates' metabolism such as acetate, lactate, and H_2 as electron donors for photosyntheses. They also utilize O_2 for respiration and may protect their hosts from the toxicity of O_2, just as may have happened in the distant past. According to this theory the symbionts would eventually have lost their photosynthetic ability and have become mitochondria. The relationship of mitochondria to bacteria is also supported by many biochemical similarities.

Fossils of bacteria and blue-green algae have been obtained from rocks whose age, as determined by geochemical dating, is more than 3×10^9 years.[68,69] However the first eukaryotic cells may have appeared about 1×10^9 years ago[70] and started to evolve into the more than one million species that now exist.[1,71,72]

1. A Changing Genome

How is it possible for the genome of an organism to increase in size as it evolved from a lower form to a higher one? Simple mutations that cause alterations in protein sequences could lead to changes in form and behavior of the organisms but could not, by themselves, account for the increase in genetic material that accompanied evolution. As a result of new techniques of genetic mapping and determining the sequence of nucleotides in DNA we are rapidly acquiring a detailed knowledge of the organization of the genome. It has been found that genes are often present as duplicate but not entirely identical copies. This suggests that there are mechanisms by which cells can acquire extra copies of one or more genes. Indeed it seems probable that at some time in the past the entire genome of bacteria was doubled and that it was later doubled again.[73] Evidence for this is that the masses of bacterial chromosomes group around values of 0.5, 1.4, and 2.7×10^9 Da. Genes can also be duplicated during the process of genetic recombination, which is discussed in Chapter 27. In addition, the size of the genome may have increased by incorporation of genetic material from extrachromosomal plasmids.

A possible advantage to a cell possessing an extra copy of a gene is that the cell would survive even when mutations rendered unusable the protein encoded by one of the copies. As long as one of the genes remained "good," the organism could grow and reproduce. The extra, mutated gene could be carried for many generations. As long as it produced only harmless, nonfunctioning proteins there might be little selection pressure to eliminate it and it might undergo repeated mutations. After many mutations and many generations later, the protein for which it coded could prove useful to the cell in some new way.

An example of evolution via gene duplication is provided by the oxygen-carrying proteins of blood. It appears that about a billion years ago, the gene for an ancestral **globin**, the protein of hemoglobin, was doubled. One gene evolved into that of present-day **globins** and the other into the gene of the muscle protein **myoglobin**. Still later, the globin gene again doubled leading to the present-day α and β chains of hemoglobin (Chapter 7). These are two distinctly different but related protein subunits whose genes are not even on the same chromosome. To complicate the picture further, most human beings have two or more copies of their α chain gene[74] as well as genes for fetal and embryonic forms of hemoglobin. However, some populations have lost one or more α chain genes. Thus, the genome changes in many details, even today.

2. Genetic Recombination, Sex, and Chromosomes

Bacteria usually reproduce by simple fission. The single DNA molecule of the chromosome is duplicated and the bacterium divides, each daughter cell receiving an identical chromosome. However, genetic recombination, which is accomplished in several ways by bacteria (Chapter 27), provides a deliberate process for mixing of genes. This process has been most fully developed in eukaryotic organisms that undergo sexual reproduction. The growth of a multicelled individual begins with the fusion of two haploid **gametes**, an egg and a spermatozoon. Each gamete carries a complete set of genetic instructions, and after the nuclei fuse the fertilized egg or **zygote** is **diploid**. Each diploid cell contains *two* complete sets of genetic blueprints of quite different origin. Even if a gene from one parent is defective, the chances are that the gene from the other parent will be good. Sexual reproduction and the associated genetic recombination also provide a means for mixing of genes.

When eukaryotic cells prepare to divide in the process called **mitosis** (Fig. 26-11), the DNA molecules of the nucleus, which become spread out through a large volume, coil and fold. Together with proteins and other molecules they form the compact bodies known as chromosomes. Some organisms, such as *Ascaris* (a roundworm), have only two chromosomes, a **homologous pair**, one inheritied from the father and one from the mother. Both chromosomes divide in every mitotic cell division so that every cell of the organism has the homologous pair. Higher organisms usually have a larger number of chromosomes. Thus, humans have 23 homologous pairs. The mouse has 20, the toad 11, onions 8, mosquitos 3, and *Drosophila* 4. Human chromosomes vary in size but are usually 4–6 µm long and ~1 µm in diameter.

By the successive divisions of mitosis, a single fertilized eukaryotic egg cell can grow to an adult. Less than 50 successive mitotic divisions will produce the $\sim10^{14}$ cells of a human. However, formation of gametes, which are haploid, requires the special process of **meiosis** (Fig. 26-12), by which the number of chromosomes is divided in half. During meiosis one chromosome of each of the homologous pairs of the diploid cell is passed to each of the gametes that are formed. In an organism such as *Ascaris*, which contains only a single pair of chromosomes, a gamete receives either the chromosome of maternal origin or that of paternal origin but not both. In organisms that have several pairs of chromosomes, one chromosome of each pair is passed to the gamete in a random fashion during meiosis. Most gametes receive some chromosomes of maternal and some of paternal origin. An important feature of meiosis is the genetic recombination that occurs during **crossing-over**. In this process, the strands of DNA are cut and genetic material is exchanged between the chromosomes of maternal and paternal origin. Thus, crossing-over breaks the **linkage** between genes and provides for greater variability in the offspring than would otherwise be possible. Each of us receives half of our genes from our mother and half from our father, but some of these genes have been inherited from each grandparent on both sides of the family, some from each great-grandparent, etc.

Many genes are passed down through many generations without substantial change, but others are evidently designed to be scrambled readily within somatic cells. Cell surface proteins[75] and antibody molecules are among the proteins whose genes undergo alteration during growth and differentiation of the tissues of the body (Chapter 32).

3. Haploid and Diploid Phases

In human beings and other higher animals, meiosis leads directly to formation of the gametes, the egg and sperm cells. These fuse to form a diploid nucleus and the adult develops by repeated mitosis of the diploid cells. While meiosis also occurs in the life cycle of all eukaryotic creatures, it is not always at a point corresponding to that in the human life cycle. Thus, the cells of many protozoa and of fungi are ordinarily haploid. When two haploid nuclei fuse to form a diploid cell, meiosis quickly occurs to produce haploid individuals again. Among lower plants and animals there is often an alternation of haploid and diploid phases of the life cycle. For example, gametes of ferns fall to the ground and germinate to form a low-growing green mosslike haploid or **gametophyte** form. The latter produces motile haploid gametes which fuse to a diploid zygote that grows into the larger and more obvious **sporophyte** form of the fern.

It is presumably the ability to survive as a heterozygote, even with one or more highly deleterious mutations, that has led to the dominance of the diploid phase in higher plants and animals.[76] However, to the biochemical geneticist organisms with a haploid phase offer experimental advantages because recessive mutants can be detected readily.

D. Survey of the Protists

Unicellular eukaryotes have traditionally been grouped together with multicellular organisms in which all cells have similar functions, with little or no differentiation into tissues, as the kingdom **Protista**.[77,78] The fungi may also be included or may be regarded as a separate kingdom.[79] With present-day emphasis on DNA sequence comparisons the traditional classification is changing, however.[26]

1. Protozoa

Among the best known of the animal-like protista is the **ameba** (subphylum Sarcodina or Rhizopoda). The most striking feature of the ameba (Fig. 1-9) is its method of locomotion, which involves the transformation of cytoplasm from a liquid state to a semisolid gel. As the ameba moves, the cytoplasm at the rear liquifies and flows to the front and into the extending pseudopodia where it solidifies along the edges. The ameba poses several important biochemical questions: What chemistry underlies the reversible change from liquid to solid cytoplasm? How can the cell membranes break and reform so quickly when an ameba engulfs food particles?[80]

Relatives of the ameba include the **Radiolaria**, marine organisms of remarkable symmetry with complex internal skeletons containing the carbohydrate polymer chitin together with silica (SiO_2) or strontium sulfate. The **Foraminifera** deposit external shells of calcium carbonate or silicon dioxide. Over 20,000 species are known and now as in the distant past their minute shells fall to the bottom of the ocean and form limestone deposits.

Tiny ameboid parasites of the subphylum **Sporozoa** attack members of all other animal phyla. Several genera of **Coccidia** parasitize rabbits and poultry causing enormous damage. Humans are often the victims of species of the genus *Plasmodium* (Fig. 1-9) which invade red blood cells and other tissues to cause **malaria**, one of our most serious ailments on a worldwide basis.[81–84] Throughout history malaria has probably killed more persons than any other disease. *Toxoplasma gondii* is another parasite which, in its haploid phase, is found throughout the world in wild animals and in humans. Although its presence usually elicits no symptoms, it sometimes causes blindness and mental retardation in children and can be fatal to persons with AIDS. Its sexual cycle occurs exclusively in cats.[85,86]

Another subphylum of protozoa, the Mastigophora, are propelled by a small number of flagella and are intermediate between animals and the algae. One of these is *Euglena viridis*, a small freshwater organism with a long flagellum in front, a flexible tapered body, green chloroplasts, and a light-sensitive "eyespot" which it apparently uses to keep itself in the sunshine (Fig. 1-9). Euglena is also able to live as a typical animal if there is no light. Treatment with streptomycin (Box 20-B) causes *Euglena* to lose its chloroplasts and to become an animal permanently. The **dinoflagellates** (Fig. 1-9), some colorless and some green, occur in great numbers among the plankton of the sea. *Giardia lamblia* is a troublesome intestinal parasite.

The **hemoflagellates** are responsible for some of our most terrible diseases. Trypanosomes (genus *Trypanosoma*) invade the cells of the nervous system causing African sleeping sickness. Mutating their surface proteins frequently by gene-scrambling mechanisms, these and other parasites are able to evade the immune response of the host.[87,88] For the same reason it is difficult to prepare vaccines against them. Other flagellates live in a symbiotic relationship within the alimentary canals of termites (Fig. 1-9) and roaches. Termites depend upon bacteria that live within the cells of these symbiotic protozoans to provide the essential enzymes needed to digest the cellulose in wood.

Members of the subphylum Ciliophora, structurally the most complex of the protozoa, are covered with a large number of cilia which beat together in an organized pattern.[89] The following question immediately comes to mind: How

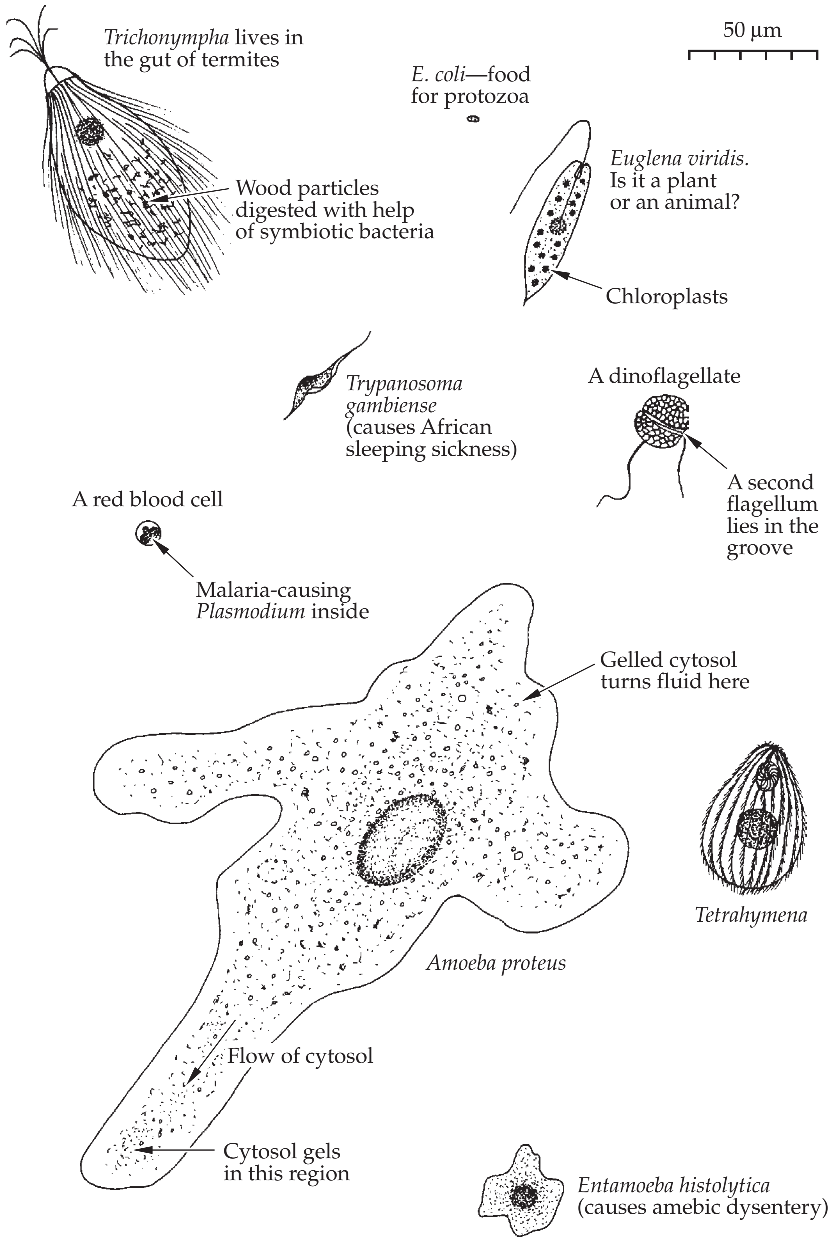

Trichonympha lives in the gut of termites
50 µm
E. coli—food for protozoa
Euglena viridis. Is it a plant or an animal?
Wood particles digested with help of symbiotic bacteria
Chloroplasts
Trypanosoma gambiense (causes African sleeping sickness)
A dinoflagellate
A second flagellum lies in the groove
A red blood cell
Malaria-causing *Plasmodium* inside
Gelled cytosol turns fluid here
Tetrahymena
Amoeba proteus
Flow of cytosol
Cytosol gels in this region
Entamoeba histolytica (causes amebic dysentery)

Figure 1-9 A few well-known protists.

are the cilia able to communicate with each other to provide this organized pattern? Two ciliates that are often studied by biochemists are *Tetrahymena* (Fig. 1-9), one of the simplest, and *Paramecium*, one of the more complex.

The **Myxomycetes** or "slime molds" are more closely related to protozoa than to fungi.[90] Members of the family Acrasieae, the best studied member of which is *Dictyostelium discoideum*, start life as small amebas. After a time, when the food supply runs low, some of the amebas begin to secrete pulses of a chemical attractant **cyclic AMP**. Neighboring amebas respond to the pulses of cyclic AMP by emitting their own pulses about 15 s later, then moving toward the original source.[91,92] The ultimate effect is to cause the amebas to stream to centers where they aggregate and form fungus-like fruiting bodies. Asexual spores are formed and the life cycle begins again. Other Myxomycetes grow as a multinucleate (diploid) **plasmodium** containing millions of nuclei but no individual cell membranes. *Physarum polycephalum*, a species whose plasmodium may spread to a diameter of 30 cm, has become popular with biochemists. The 800,000 nuclei per square millimeter all divide synchronously.

2. Fungi

Lacking photosynthetic ability, living most often in soil but sometimes in water, the fungi are represented by almost half as many species (~10^5) as are the vascular plants.[93] The distinguishing characteristics of fungi are the lack of chlorophyll and growth as a series of many branched tubules (usually 6–8 µm diameter), the **hyphae**, which constitute the **mycelium**. The hyphae are not made up of separate cells but contain a mass of protoplasm with many nuclei. Only occasional septa divide the tubules. Most fungi are saprophytic, living on decaying plants or animal tissues. However, others are parasites that produce serious and difficult-to-treat infections in humans. An important medical problem is the lack of adequate antibiotics for treating fungal infections (mycoses).[94–96] On the other hand, fungi produce important antibiotics such as **penicillin**. Still others form some of the most powerful toxins known!

The lower fungi or **Phycomycetes** include simple aquatic molds and mildew organisms. Higher fungi are classified as **Ascomycetes** or **Basidiomycetes** according to the manner in which the sexual spores are born. In the Ascomycetes these spores are produced in a small sac called an **ascus** (Fig. 1-10). Each ascus contains four or eight spores in a row, a set of four representing the results of a single pair of meiotic divisions. A subsequent mitotic division will give eight spores. This is one of the features that has made *Neurospora crassa* (Fig. 1-10) a favorite subject for genetic studies.[97] The ascospores can be dissected out in order from the ascus and cultivated separately to observe the results of crossing-over during meiosis.

Neurospora also reproduces via haploid spores called **conidia**. The haploid mycelia exist as two mating types and conidia or mycelia from one type can fertilize cells in a special body (the protoperithecium) of the other type to form zygotes. The latter immediately undergo meiosis and mitosis to form the eight ascospores. Among other Ascomycetes are the highly prized edible truffles and morels. However, most mushrooms and puffballs are fruiting bodies of Basidiomycetes. Other Basidiomycetes include the **rusts**, which cause enormous damage to wheat and other grain crops.

Yeasts are fungi adapted to life in an environment of high sugar content and which usually remain unicellular and reproduce by budding (Fig. 1-10). Occasionally the haploid cells fuse in pairs to form diploid cells and sexual spores. Some yeasts are related to the Ascomycetes, others to Basidiomycetes. *Saccharomyces cerevisiae*, the organism of both baker's and brewer's yeast, is an Ascomycete. It can grow indefinitely in either the haploid or diploid phase. The genetics and biochemistry of this yeast have been studied extensively.[98–102] The genome is relatively small with 13.5 x 10^6 base pairs in 17 chromosomes. The sequence of the 315,000 base pairs of chromosome III was determined in 1992[101,102] and the sequence of the entire genome is now known.[103]

Fungi often grow in symbiotic association with other organisms. Of special importance are the **mycorrhizae** (fungus roots) formed by colonization of fine roots by beneficial soil fungi. Almost all plants of economic importance form mycorrihizae.[104]

3. Algae

Algae are chlorophyll-containing eukaryotic organisms which may be either unicellular or colonial.[105] The colonial forms are usually organized as long filaments, either straight or branched, but in some cases as blades resembling leaves. However, there is little differentiation among cells. The gold-brown, brown, and red algae contain special pigments in addition to the chlorophylls.

The euglenids (**Euglenophyta**) and dinoflagellates (**Pyrrophyta**), discussed in the protozoa section, can equally well be regarded as algae. The bright green **Chlorophyta**, unicellular or filamentous algae, are definitely plants, however. Of biochemical interest is *Chlamydomonas*, a rather animal-like creature with two flagella and a carotenoid-containing eyespot or **stigma** (Fig. 1-11). *Chlamydomonas* contains a single chloroplast. The "pyrenoid", a center for the synthesis of starch, lies, along with the eyespot, within the chloroplast. The organism is haploid with "plus" and "minus" strains and motile gametes. Zygotes immediately undergo meiosis to form haploid spores. With a well-established genetic map, *Chlamydomonas* is another important organism for studies of biochemical genetics.[106]

The filamentous *Ulothrix* shows its relationship to the animals through formation of asexual spores with four flagella and biflagellate gametes. Only the zygote is diploid. On the other hand, the incomparably beautiful *Spirogyra* (Fig. 1-11) has no motile cells. The ameboid male gamete flows through a tube formed between the two mating cells, a behavior suggesting a relationship to higher green plants.

Some unicellular algae grow to a remarkable size. One of these is *Acetabularia* (Fig. 1-11), which lives in the warm waters of the Mediterranean and other tropical seas. The cell contains a single nucleus which lies in the base or rhizoid portion. In the mature alga, whose life cycle in the laboratory is 6 months, a cap of characteristic form develops. When cap development is complete, the nucleus divides into about 10^4 secondary nuclei which migrate up the stalk and out into the rays of the cap where they form cysts. After the cap decays and the cysts are released, meiosis occurs and the flagellated gametes fuse in pairs to form zygotes which again grow into diploid algae. Because of its large size and the location of the nucleus in the base, the cells can be cut and grafted. Nuclei can be removed or transplanted and growth and development can be studied in the presence or absence of a nucleus.[107–110] The green algae **Volvox** live in wheel-like colonies of up to several thousand cells and are useful for biochemical studies of differentiation.[111]

Look through the microscope at almost any sample of algae from a pond or aquarium and you will see little boatlike **diatoms** slowly gliding through the water. The most prominent members of the division Chrysophyta, diatoms are characterized by their external "shells" of silicon dioxide. Large and ancient deposits of diatomaceous earth contain these durable silica skeletons which are finely marked, often with beautiful patterns (Fig. 1-11). The slow motion of diatoms is accomplished by streaming of protoplasm through a groove on the surface of the cell. Diatoms are an important part of marine plankton, and it is estimated that three-fourths of the organic material of the world is produced by diatoms and dinoflagellates. Like the brown algae, Chrysophyta contain the pigment **fucoxanthin**.

Other groups of algae are the brown and red marine algae or seaweed. The former (**Phaeophyta**) include the giant kelps from which the polysaccharide **algin** is obtained. The **Rhodophyta** are delicately branched plants containing the red pigment **phycoerythrin**. The polysaccharides, **agar** and **carrageenin**,

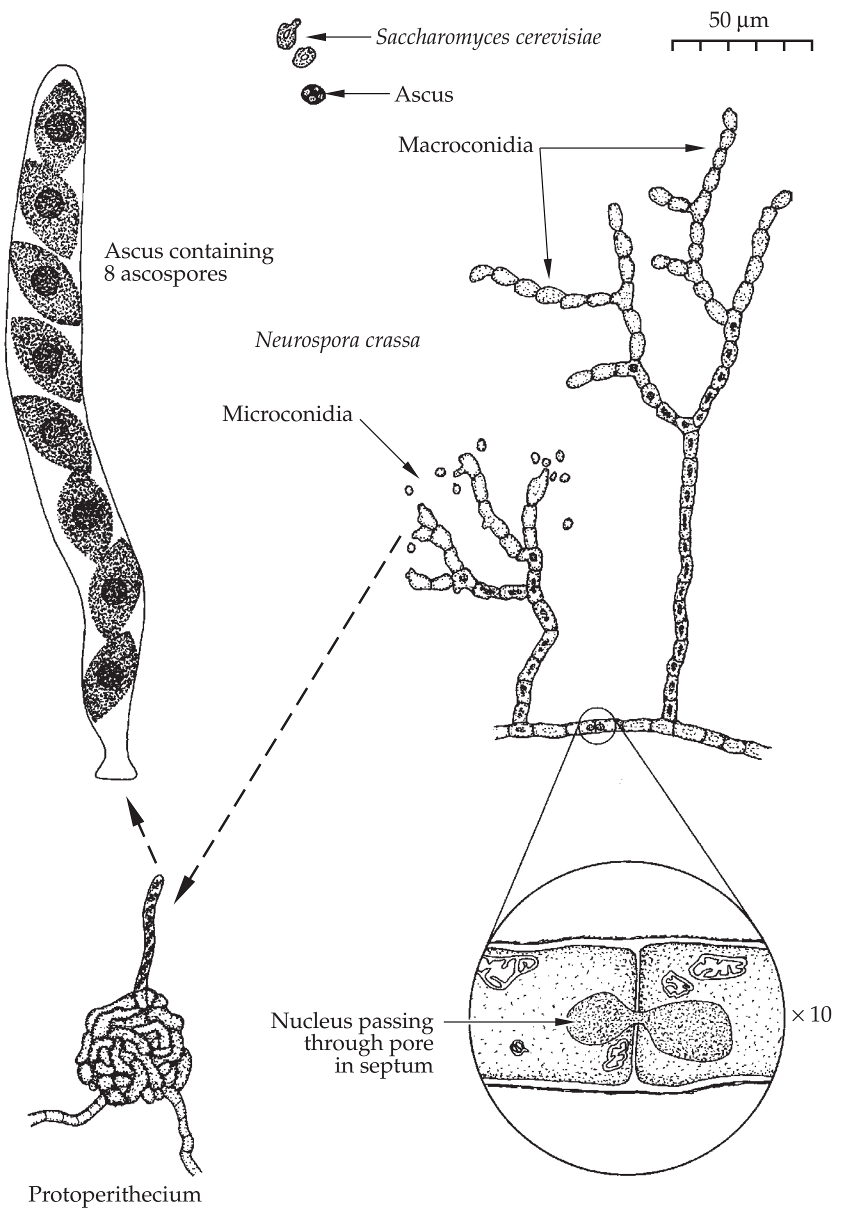


Figure 1-10 Two frequently studied fungi. Top (including ascus): the yeast *Saccharomyces cerevisiae*. Below: *Neurospora crassa* showing various stages. After J. Webster.[93]

a popular additive to chocolate drinks and other foods, come from red algae.

Symbiotic associations of fungi with either true algae or with cyanobacteria are known as **lichens**. Over 15,000 varieties of lichens grow on rocks and in other dry and often cold places. While the algae appear to benefit little from the association, the fungi penetrate the algae cells and derive nutrients from them.[112] Although either of the two partners in a lichen can be cultured separately, the combination of the two is capable of producing special pigments and phenolic substances known as **depsides** which are not formed by either partner alone.

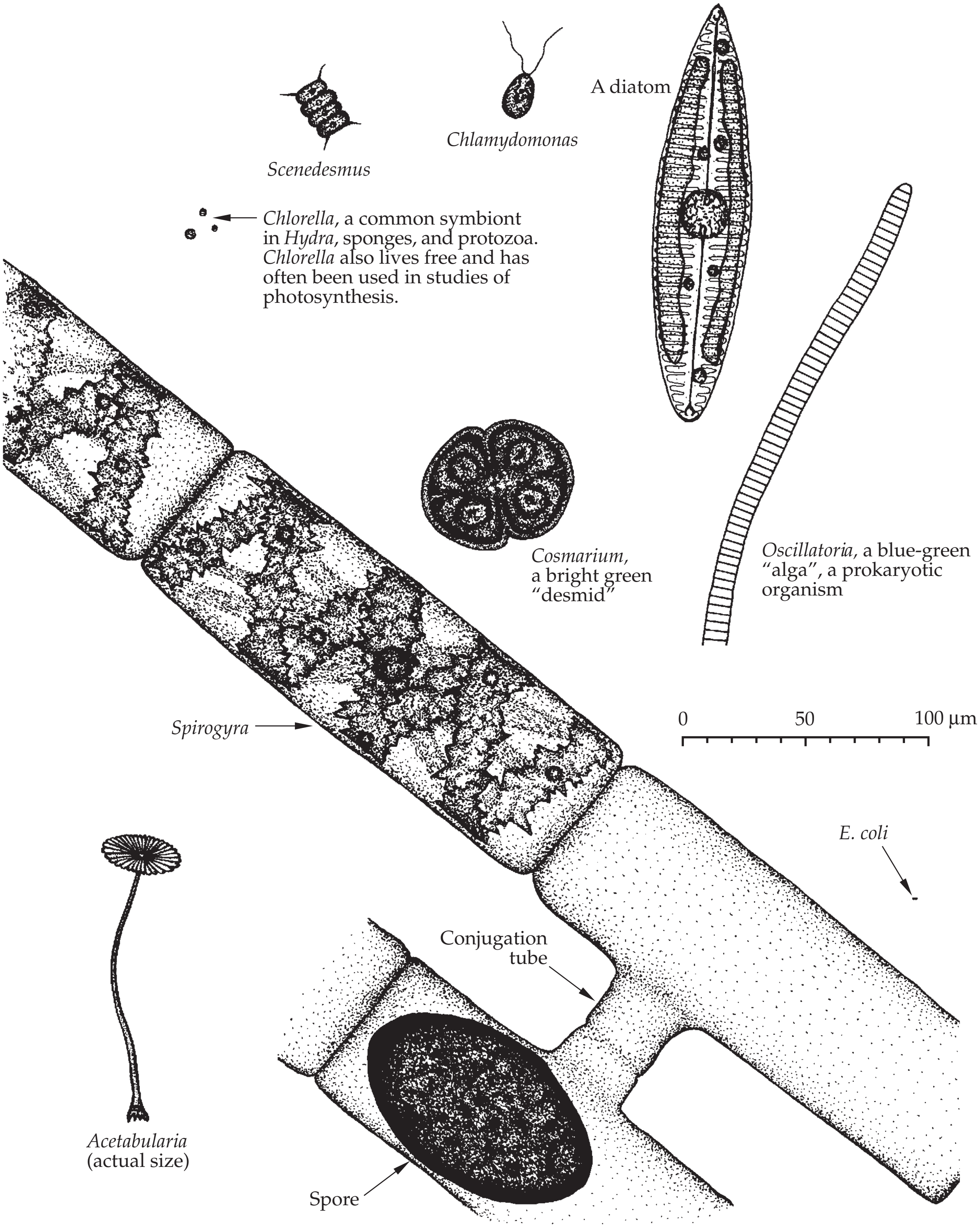

Scenedesmus
Chlamydomonas
A diatom
Chlorella, a common symbiont in Hydra, sponges, and protozoa. Chlorella also lives free and has often been used in studies of photosynthesis.
Cosmarium, a bright green "desmid"
Oscillatoria, a blue-green "alga", a prokaryotic organism
0
50
100 µm
Spirogyra
E. coli
Conjugation tube
Acetabularia (actual size)
Spore

Figure 1-11 A few species of algae.

E. The Variety of Animal Forms

In this section, we will consider only a few biochemical and other aspects of multicellular animals or **Metazoa**. The sudden appearance of a large number of Metazoans about 0.5 x 10^9 years ago[113,114] may have been an outcome of the appearance of split genes (see Section B, 1). As a result of gene duplication the coding pieces of split genes, the **exons**, could be moved to new locations in a chromosome where they could have become fused with other pieces of DNA to form entirely new genes.[115]

1. Major Groups of Multicellular Animals

The simplest metazoa are tiny symbiotic worms of the phylum (or subkingdom) **Mesozoa**, which live in the kidneys of deep sea-dwelling cephalopods (octopi and squid). Each worm is made up of only 25 cells in a single layer enclosing one or a small number of elongated axial cells (Fig. 1-12). Mesozoa have been regarded as parasitic, but they appear to facilitate excretion of NH_3 by the host through acidification of the urine.[118,119]

Porifera or sponges are the most primitive of multicelled animals.[120] They lack distinct tissues but contain several specialized types of cells. The body is formed by stationary cells that pump water through the pores to bring food to the sponge. Within the body **amebocytes** work in groups to form the **spicules** of calcium carbonate, silicon dioxide, or the protein **spongin** (Fig. 1-12). Sponges appear to lack a nervous system.

Individuals of the next most complex major phylum, **Cnidaria** (formerly Coelenterata), are radially symmetric with two distinct cell layers, the **endoderm** and **ectoderm**. Many species exist both as a polyp or **hydra** form (Fig. 1-13) and as a **medusa** or jellyfish. The jellyfish apparently has no brain but the ways in which its neurons interconnect in a primitive radial net are of interest. The Cnidaria have a very simple body form with remarkable regenerative powers. The freshwater hydra, a creature about 1 cm long (Fig. 1-13), contains a total of ~10^5 cells. A complete hydra can be regenerated from a small piece of tissue if the latter contains some of both the inner and the outer cell layers.[121,122]

The body of flatworms (phylum **Platyhelminthes**) consists of two external cell layers (endoderm and ectoderm) with a third layer between. A distinct excretory system is present. In addition to a nerve net resembling that of the Cnidaria, there are a cerebral ganglion and distinct eyes. One large group of flatworms, the **planarians** (typically about 15 mm in length, Fig. 1-14), inhabit freshwater streams. They are said to be the simplest creatures in which *behavior* can be studied.

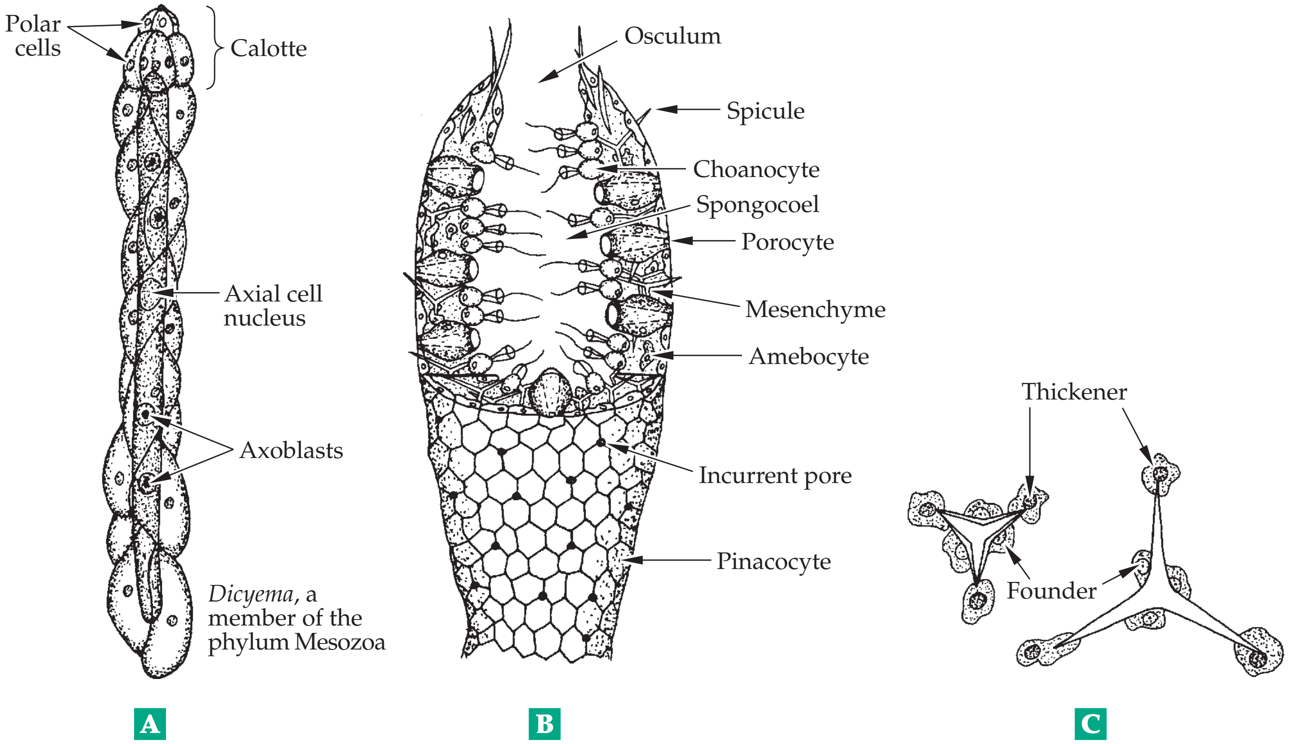


Figure 1-12 Some lower forms of Metazoa. (A) Mesozoa (25 cells). After C. P. Hickman.[116] (B) A small asconoid sponge. After C. A. Villee, W. F. Walker, Jr., and R. D. Barnes.[117] (C) Ameboid cells of a sponge forming spicules. After Hickman.

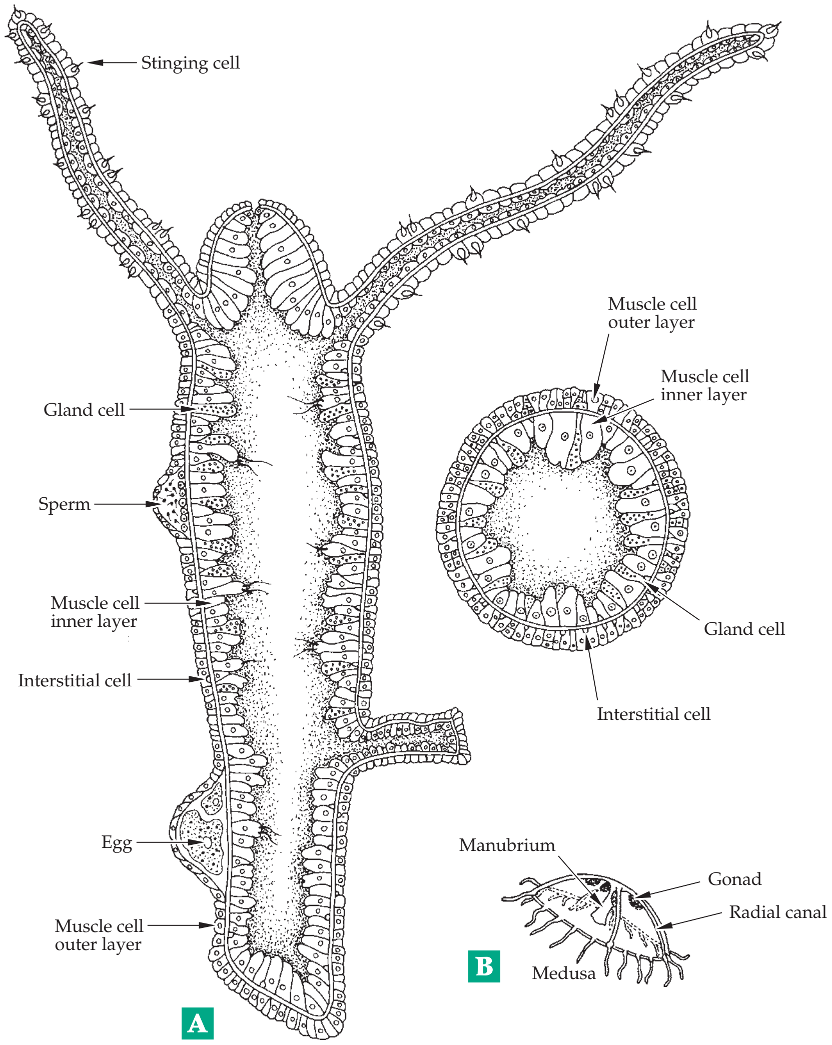


Figure 1-13 (A) Hydra. After Loomis.[123] (B) The medusa stage of *Obelia*, a hydroid coelenterate.[123a]

Many parasitic flatworms (tapeworms and flukes) attack higher organisms.[124] Among them are the **Schistosoma**, tiny worms that are transmitted to humans through snails and which attack the blood vessels. The resulting **schistosomiasis** is one of the most widespread debilitating diseases on earth today, affecting 200 million people or more.[125,126]

The roundworms (**Nematoda**)[127–129] have, in addition to the **enteron** (alimentary tract), a separate body cavity. Free-living nematodes abound in water and soil but many species are parasitic. They do enormous damage to plants and to some animal species. Trichina, hookworms[129a], and filaria worms attack humans. However, in the laboratory the 1-mm-long, 810-cell nematode, *Caenorhabditis elegans* (Fig. 1-14) has become an important animal. In 1963 Sydney Brenner launched what has become a worldwide effort to make this tiny worm the equivalent in the animal kingdom of *E. coli* in the bacterial world.[129] The 10^8 nucleotides in the worm's six chromosomes contain ~13,600 genes. *C. elegans* has become an important animal in which to study differentiation. Already the exact lineage of every cell has been traced, as has every connection among the 302 neurons in the animal's nervous system. The related **rotifers**,[130] with whirling "wheels" of cilia on their heads (Fig. 1-14) and transparent bodies, are a delight to the microscopist. Like nematodes, they are "cell constant" organisms. The total number of cells in the body is constant as is that in almost every part of every organ. Part of the developmental plan of such organisms is a "programmed cell death" (Chapter 32).

The **Annelida** (segmented worms)[131] are believed to be evolutionary antecedents of the arthropods. Present-day members include earthworms, leeches, and ~10^5 species of marine **polychaetes**. Annelids have a true body cavity separate from the alimentary canal and lined by a peritoneum. They have a well-developed circulatory system and their blood usually contains a type of hemoglobin.

About 10^6 species of **arthropods** (80% of all known animals) have been described. Most are very small.[72] These creatures, which have a segmented exoskeleton of **chitin** and other materials, include the horseshoe crabs, the Arachnida (scorpions, spiders, and mites), the Crustacea, Myriopoda (centipedes and millipedes), and the Insecta. Important biochemical problems are associated with the development and use of insecticides and with our understanding of the metamorphosis that occurs during the growth of arthropods.[132] The fruit fly *Drosophila melanogaster* has provided much of our basic knowledge of genetics and continues to be the major species in which development is studied.[133–134a]

Among the molluscs (phylum **Mollusca**) the squids and octopuses have generated the most interest among biochemists. The neurons of squid contain giant axons, the study of which has led to much of our knowledge of nerve conduction. Octopuses show signs of intelligence not observed in other invertebrates whose nervous reactions seem to be entirely "preprogrammed." The brains of some snails contain only 10^4 neurons, some of which are unusually large. The

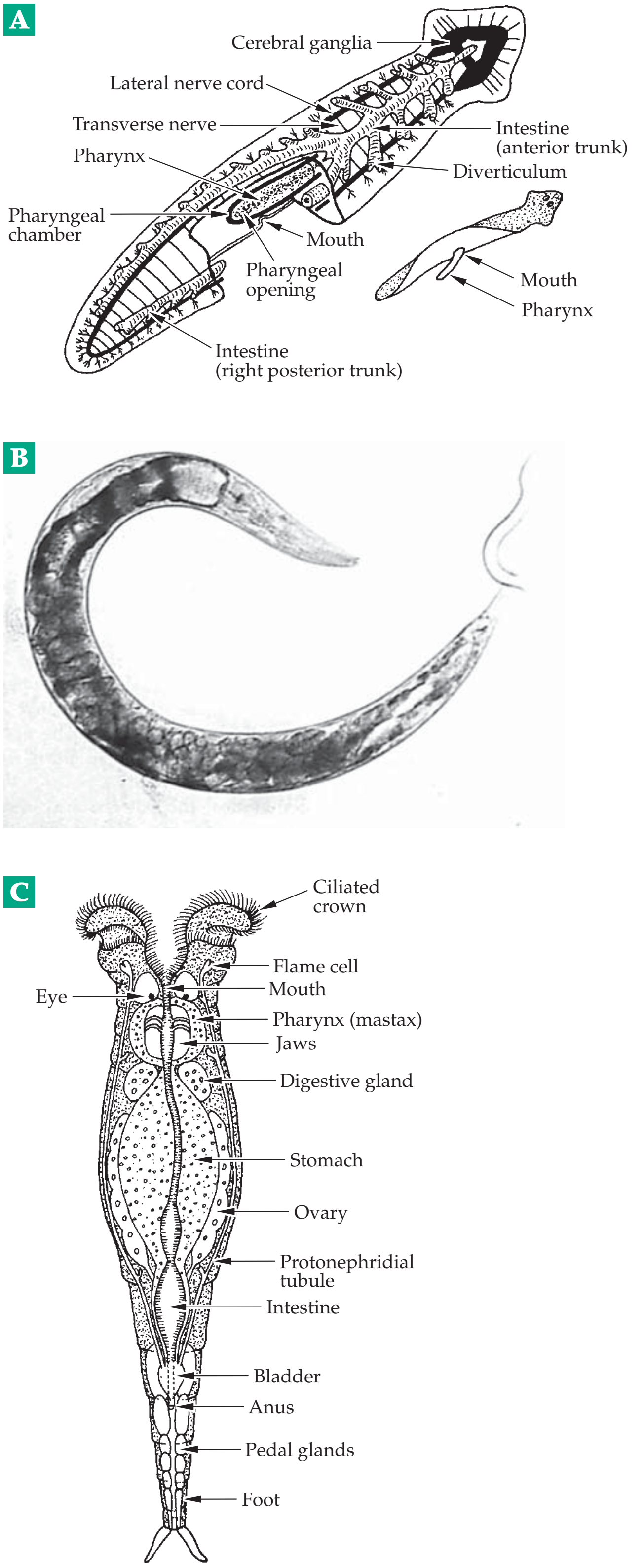

A
Cerebral ganglia
Lateral nerve cord
Transverse nerve
Pharynx
Pharyngeal chamber
Intestine (anterior trunk)
Diverticulum
Mouth
Pharyngeal opening
Mouth
Pharynx
Intestine (right posterior trunk)
B
C
Ciliated crown
Flame cell
Eye
Mouth
Pharynx (mastax)
Jaws
Digestive gland
Stomach
Ovary
Protonephridial tubule
Intestine
Bladder
Anus
Pedal glands
Foot

Figure 1-14 (A) A planarian, length 15 mm. After Hickman.[116] Diagram of digestive and nervous sytems; cutaway section shows ventral mouth. Small drawing shows pharynx extended through ventral mouth. (B) The nematode *Caenorhabditis elegans*. *Ascaris* is very similar in appearance. From Buchsbaum.[77] (C) A rotifer, *Philodina* (~10^3 cells). After C. A. Villee *et al.*[117]

Echinodermata or spiny-skinned animals (starfish, sea urchins, and sea cucumbers) are regarded as a highly advanced phylum. Their embryological development has been studied intensively.

The phylum **Chordata**, to which we ourselves belong, includes not only the vertebrates but also more primitive marine animals that have a spinal cord. Among these primitive species, which may be related to early ancestral forms, are the **tunicates** or sea squirts. They have a very high concentration of vanadium in their blood.

2. Cell Types and Tissues

Isolated animal cells in tissue culture, no matter how highly differentiated, tend to revert quickly to one of three basic types known as **epitheliocytes**, **mechanocytes**, and **amebocytes**. Epitheliocytes are closely adherent cells derived from epithelial tissues and thought to be related in their origins to the two surface layers of the embryonic blastula. Mechanocytes, often called **fibroblasts** or **fibrocytes**, are derived from muscle, supporting, or connective tissue. Like the amebocytes, they arise from embryonic mesenchymal tissue cells that have migrated inward from the lower side of the blastula (Chapter 32). **Neurons**, **neuroglia**, and **lymphocytes** are additional distinct cell types.

BOX 1-E ERRORS, MISCONCEPTIONS, AND SPECULATION

Warning: Not everything in this book is true. Despite all efforts to get it right, there are unintentional errors and misinterpretations of experimental results. Indeed, the history of biochemistry is replete with accounts of experimental findings that were interpreted incorrectly. Yet, the ideas expressed often stimulated others to develop a more correct picture later. The same is true today. Students should be critical, should look at experimental details, and *consult original literature as much as possible.*

Progress in science depends both upon careful observations and measurements and upon imaginative interpretations of unexpected findings. Speculative ideas, a number of which are mentioned in this book, provide an important stimulus in science. They should neither be ignored nor accepted as facts. I have tried to write in such a way that established facts will not often be confused with speculation.

Tissues. Cells aggregate to form four major kinds of tissue. Epithelial tissues line the primary surfaces of the body: the skin, the digestive tract, urogenital tract, and glands. External skin is composed of flat platelike squamous epithelial cells whereas internal surfaces are often formed by colummar epithelial cells. Glands (sweat, oil, mammary, and internal secretory) as well as the sensory organs of the tongue, nose, and ear are all composed of epithelial cells. Epithelial cells are among the most highly polarized of cells. One side of each cell faces the outside, either air or water, while the other side is often directly against a basement membrane.

Supporting and connective tissues include the fatty **adipose tissue** as well as **cartilage** and **bone**. Both of the latter contain large amounts of intercellular material or **ground substance** consisting largely of complex polymers. Embryonic fibroblasts differentiate into white fibers, which produce collagen, and yellow fibers, which form elastin. The fibrils of both of these proteins are assembled in the intercellular space where they are embedded in the ground substance. **Osteoblasts** form bone by deposition of calcium phosphate in 3–7 µm thick layers within a ground substance that contains special proteins.

A third tissue is **muscle**, which is classified into three types: **striated** (voluntary skeletal muscle), **cardiac** (involuntary striated muscle), and **smooth** (involuntary) muscle. There are two major groups of cells in **nervous tissue**, the fourth tissue type. **Neurons** are the actual conducting cells whose cell membranes carry nerve impulses. Several kinds of **glial cells** lie between and around the neurons.

Blood cells. Blood and the linings of blood vessels may be regarded as a fifth tissue type.[135,135a] The human body contains 5 x 10^9 **erythrocytes** or red blood cells per ml, a total of 2.5 x 10^{13} cells in the five liters of blood present in the body. Erythrocytes are rapidly synthesized in the bone marrow. The nucleus is destroyed, leaving a cell almost completely filled with hemoglobin. With an average lifetime of 125 days, human red blood cells are destroyed by leukocytes in the spleen and liver.

The white blood cells or **leukocytes** are nearly a thousandfold less numerous than red cells. About 7 x 10^6 cells are present per ml of blood. There are three types of leukocytes: **lymphocytes** (~26% of the total), **monocytes** (~7% of the total), and **polymorphonuclear leukocytes** or **granulocytes** (~70% of the total). Lymphocytes are about the same size as erythrocytes and are made in lymphatic tissue. Individual lymphocytes may survive for as long as ten years. They function in antibody formation and are responsible for maintenance of long-term immunity.

Monocytes, two times larger, are active in ingesting bacteria. These cells stay in the blood only a short time before they migrate into the tissues where they become **macrophages**,[136] relatively fixed phagocytic cells. Macrophages not only phagocytize and kill invading bacteria, protozoa, and fungi but also destroy cancer cells. They also destroy damaged cells and cellular debris as part of the normal turnover of tissues. They play an essential role in the immune system by "processing" antigens and in releasing stimulatory proteins.

Granulocytes of diameter 9–12 µm are formed in the red bone marrow. Three types are distinguished by staining: **neutrophils**, **eosinophils**, and **basophils**. Neutrophils are the most numerous phagocytic cells of our blood and provide the first line of defense against bacterial infections. The functions of eosinophils and basophils are less well understood. The number of eosinophils rises during attacks of hay fever and asthma and under the influence of some parasites, while the basophil count is increased greatly in leukemia and also by inflammatory diseases. Granules containing histamine, heparin, and leukotrienes are present in the basophils. Blood **platelets** or **thrombocytes** are tiny (2–3 µm diameter) cell-like bodies essential for rapid coagulation of blood. They are formed by fragmentation of the cytoplasm of bone marrow **megakaryocytes**. One mature megakaryocyte may contribute 3000 platelets to the 1–3 x 10^8 per ml present in whole blood.

Cell culture. Laboratory growth of isolated animal cells has become very important in biochemistry.[137] Sometimes it is necessary to have many cells with as nearly as possible identical genetic makeup. Such bacterial cells are obtained by plating out the bacteria and selecting a small colony that has grown from a single cell to propagate a "pure strain." Similarly, single eukaryotic cells may be selected for tissue culture and give rise to a **clone** of cells which remains genetically identical until altered by mutations.

The culture of embryonic fibroblasts is used to obtain enough cells to perform prenatal diagnosis of inherited metabolic diseases (Box 1-D). Tissue culture is easiest with embryonic or cancer cells, but many other tissues can be propagated. However, the cells that grow best and which can be propogated indefinitely are not entirely normal; the well-known **HeLa** strain of human cancer cells which was widely grown for many years throughout the world contains 70–80 chromosomes per cell compared with the normal 46.

3. Communication

Plants are able to maintain their form because the cells are surrounded by thick walls that cement the cells together. However, animal cells lack rigid walls and must be held together by specialized contacts.[138,139] Contacts between cells of both plants and animals are

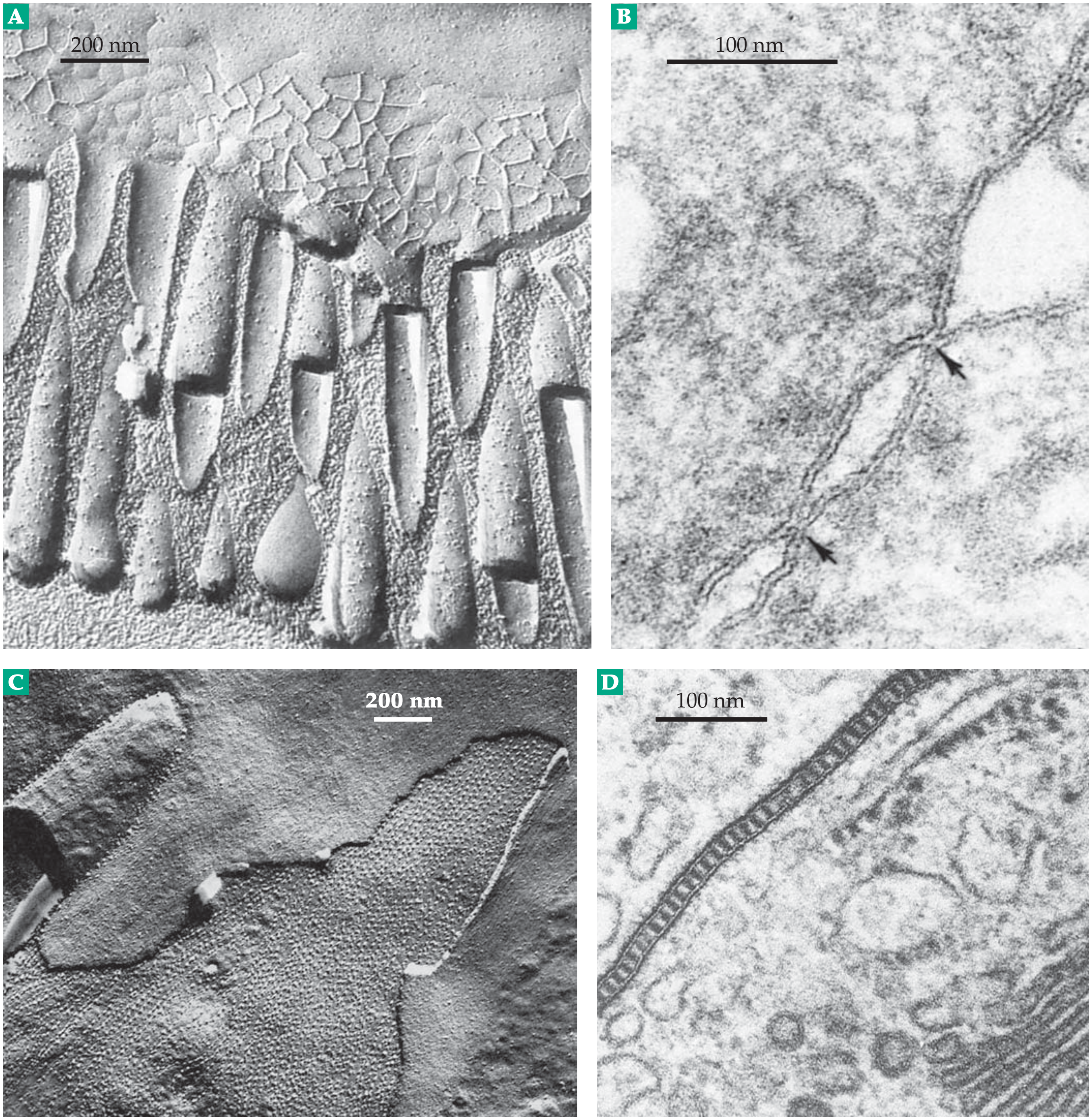
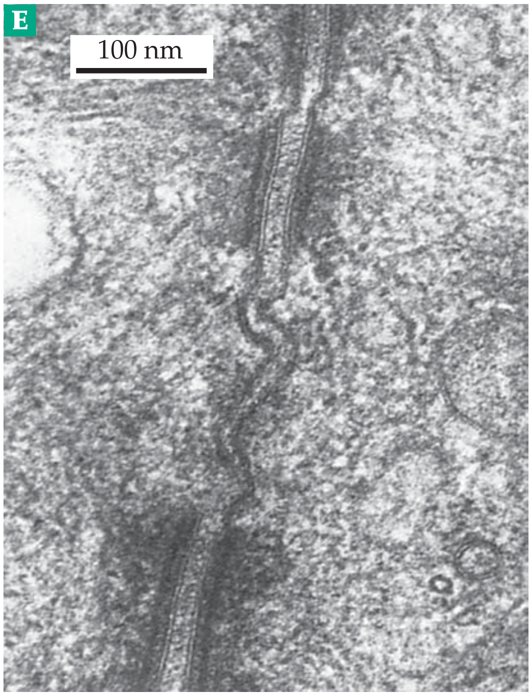
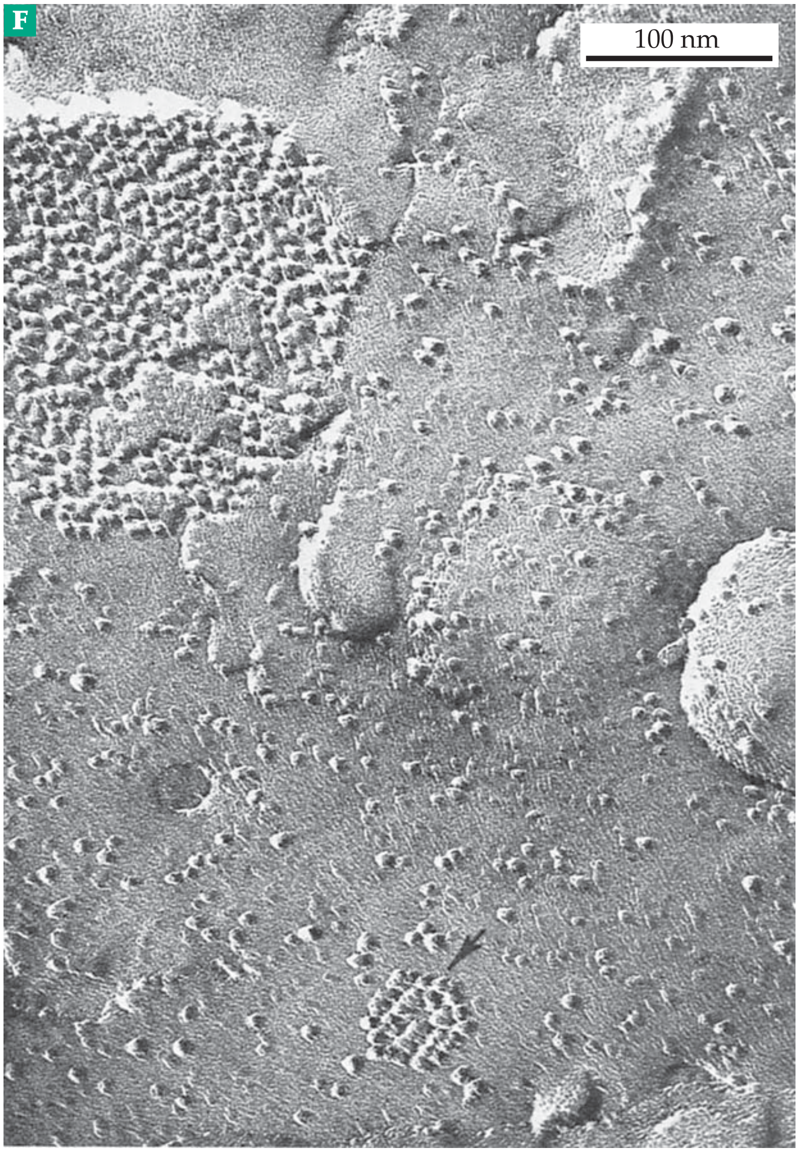


Figure 1-15 Electron micrographs of cell junctions of three types. (A) Freeze-fractured zona occludens (occlusion zone) between epithelial cells of the rat small intestine. The tight junctions are repesented as a meshwork of ridges (in the P or protoplasmic fracture face) or grooves (in the E fracture face which looks toward the extracellular space). These represent the actual sites of membrane fusion. Microvilli are seen in the lower part of the photograph. From D. S. Friend and N. B. Gilula.[141] (B) Thin cross section of tight junction between mouse hepatocytes. The arrows indicate points of membrane fusion. From Gilula.[142] Copyright 1975 by The Williams & Wilkins Co., Baltimore. (C) A freeze-fractured septate junction from ciliated epithelium of a mollusc. This type of junction forms a belt around the cells. Fracture face P (central depressed area) contains parallel rows of membrane particles that correspond to the arrangmenet of the intracellular septa seen in thin sections. The surrounding fracture face E contains a complementary set of grooves. Particles in nonjunctional membrane regions (upper right corner) are randomly arranged. (D) Thin section of a septate junction of the type shown in (C). The plasma membranes of the two cells are joined by a periodic arrangement of electron-dense bars or septa, which are present within the intercellular space. Note the Golgi membranes in the lower right part of the photograph. (C) and (D) are from N. B. Gilula[143]





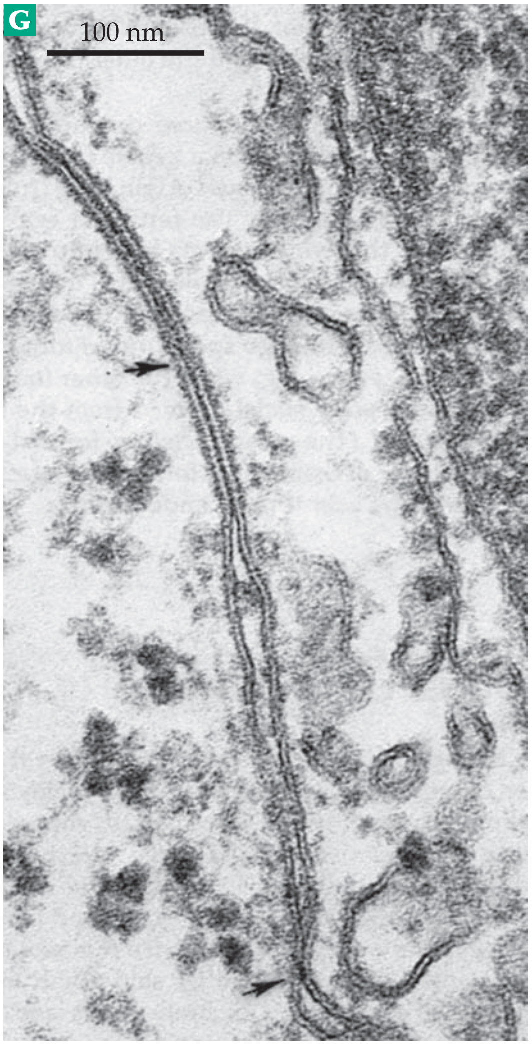


Figure 1-15 (continued) (E) Desmosomes (macula adherens) in rat intestinal epithelium. Features include a wide (25–35 nm) intercellular space containing dense material, two parallel cell membranes, a dense plaque associated with the cytoplasmic surface and cytoplasmic tonofilaments that converge on the dense plaque. From Gilula.[143] (F) Freeze-fractured surface through gap junctions between communicating cells in culture. Both a large junction and a smaller one below (arrow) can be seen. (G) Gap junctions in thin section. (F) and (G) are from N. B. Gilula.[144]

important for a second reason: Cells must communicate, one with another.

Cell contacts and junctions. Many epithelial cells, e.g., those lining the border of kidney tubules and secretory glands, form **tight junctions** with adjacent cells. Electron microscopy shows that in these junctions the outer portions of the membranes actually fuse in some places (Fig. 1-15). One way in which this has been demonstrated is to freeze a tissue rapidly and to fracture it in the frozen state within a vacuum chamber. The fractured tissue is kept at about –100°C in a vacuum for a short time while water molecules evaporate from the fractured surfaces. A thin plastic replica is then made of the etched surfaces, which sometimes pass through tight junctions revealing details of their structure (Fig. 1-15, A, E). Study of electron micrographs of such surfaces shows that some cells are completely surrounded by belts of tight junctions, sometimes referred to as **occlusion zones** or **terminal bars**. Tight junctions between endothelial cells of blood capillaries in the brain prevent free diffusion of compounds from the blood stream into brain cells and form the **blood–brain barrier**.[140] Tight junctions between neurons and adjacent cells surround the **nodes of Ranvier** (Chapter 30).

Contacts of another type, known as **septate desmosomes** or **adhesion discs**, form a belt around the cells of invertebrate epithelia. In these contacts a space of ~18 nm between adjacent cell membranes is bridged in a number of places by thin walls. Behind the desmosomes the membrane is often backed up at these points by an electron-dense region to which are attached many fine microfilaments of ~6–10 nm diameter (Fig. 1-15, D).

One method of communication between cells is by passage of chemical substances through special junctions which, because of their appearance in electron micrographs of thin sections (Fig. 1-15, G) are known as **gap junctions**.[139,145,146] Gap junctions may cover substantial areas of the cell interface. In cross section, a thin 3–4 nm gap between the adjacent cell membranes is bridged by a lattice-like structure, which may appear in freeze-fractured surfaces as a hexagonal array of particles (Fig. 1-15, F, lower junction). These particles or **connexons** are each thought to be composed of six protein subunits. A central channel in the connexon is able to pass molecules of molecular mass up to about 500 Da.[147,148] Small molecules may be able to pass freely from one cell to another through the gap junction. Because of their low electrical resistance, gap junctions allow "electrical coupling" of cells. Such junctions form the **electrotonic synapses** that link some neurons to other excitable cells. Heart cells are all electrically coupled through gap junctions.[149]

Another type of communicating junction is also found in **synapses** of the nervous system. At these specialized contacts a nerve impulse transmitted along the membrane of one neuron triggers the release of a **neurotransmitter**, a chemical substance that passes across the gap between cells of the synapse and initiates a nerve impulse in the second neuron (Chapter 30).

Cell recognition. Cells of higher organisms are able to recognize other cells as identical, as belonging to another tissue, or as being "foreign." This ability is developed most highly by cells of the immune system but is possessed to some extent by others. For example, cells of sponges can be separated by partial digestion of the protein "cement" that holds them together. When dissociated cells from orange sponges were mixed with those from yellow sponges, the cells clumped together to reform small sponges.[150,151] Furthermore, orange cells stuck to orange cells and yellow to yellow cells. Similar results have been obtained using a mixture of cultivated liver, kidney, and embryonic brain cells. When a wound heals, epithelial cells grow and move across the wound surface but they stop when they meet. Cells in tissue culture and growing on a glass surface experience this same **contact inhibition**[152] and spread to form a unicellular layer. Cancer cells in culture do not stop but climb one on top of the other, apparently lacking proper recognition and communication. Many chemical signals appear to pass between cells. An important goal of contemporary biochemistry is to understand how cells recognize each other and respond to signals that they receive.

F. Higher Plants and Plant Tissues

Botanists recognize two divisions of higher plants. The **Bryophyta** or moss plants consist of the Musci (mosses) and Hepaticae (liverworts). These plants grow predominantly on land and are characterized by swimming sperm cells and a dominant gametophyte (haploid) phase. **Tracheophyta**, or vascular plants, contain conducting tissues. About 2×10^5 species are known. The ferns (class Filicineae, formerly Pteridophyta) are characterized by a dominant diploid plant and alternation with a haploid phase. Seed plants are represented by two classes: **Gymnosperms** (cone-bearing trees) and **Angiosperms**, the true flowering plants.

Genetically the simplest of the angiosperms is the little weed *Arabidopsis thaliana*, whose generation time is as short as five weeks. Its five chromosomes contain only 10^8 base pairs in all, the smallest known genome among angiosperms[153] and one whose complete nucleotide sequence is being determined. Its biochemistry, physiology, and developmental biology are under intensive study. It may become the "fruit fly" of the plant kingdom.

There are several kinds of plant tissues. Undifferentiated, embryonic cells found in rapidly growing regions of shoots and roots form the **meristematic tissue**. By differentiation, the latter yields the simple tissues, the parenchyma, collenchyma, and sclerenchyma. **Parenchyma** cells are among the most abundant and least specialized in plants. They give rise through further differentiation to the **cambium layer**, the growing layer of roots and stems. They also make up the pith or pulp in the center of stems and roots, where they serve as food storage cells.

The **collenchyma**, present in herbs, is composed of elongated supporting cells and the **sclerenchyma** of woody plants is made up of supporting cells with hard lignified cell walls and a low water content. This tissue includes **fiber cells**, which may be extremely long; e.g., pine stems contain fiber cells of 40 µm diameter and 4 mm long.

Two complex tissues, the **xylem** and **phloem**, provide the conducting network or "circulatory system" of plants. In the xylem or woody tissue, most of the cells are dead and the thick-walled tubes (**tracheids**) serve to transport water and dissolved minerals from the roots to the stems and leaves. The phloem cells provide the principal means of downward conduction of foods from the leaves. Phloem cells are joined end to end by **sieve plates**, so-called because they are perforated by numerous minute pores through which cytoplasm of adjoining sieve cells appears to be connected by strands 5–9 µm in diameter.[154] Mature sieve cells have no nuclei, but each sieve cell is paired with a nucleated "companion" cell.

Epidermal tissue of plants consists of flat cells, usually containing no chloroplasts, with a thick outer wall covered by a heavy waxy **cuticle** about 2 µm thick. Only a few specialized cells are found in the epidermis. Among them are the paired **guard cells** that surround the small openings known as **stomata** on the undersurfaces of leaves and control transpiration of water. Specialized cells in the root epidermis form **root hairs**, long extensions (~1 mm) of diameter 5–17 µm. Each hair is a single cell with the nucleus located near the tip.

Figure 1-16 shows a section from a stem of a typical angiosperm. Note the thin cambium layer between the phloem and the xylem. Its cells continuously undergo differentiation to form new layers of xylem increasing the woody part of the stem. New phloem cells are also formed, and as the stem expands all of the tissues external to the cambium are renewed and the older cells are converted into bark.

Plant **seeds** consist of three distinct portions. The **embryo** develops from a zygote formed by fusion of a sperm nucleus originating from the pollen and an egg cell. The fertilized egg is surrounded in the gymnosperms by a nutritive layer or **endosperm** which is **haploid** and is derived from the same gametophyte tissue that produced the egg. In angiosperms *two* sperm nuclei form; one of these fertilizes the egg, while the other fuses with *two* haploid **polar nuclei** derived from the female gametophyte. (The polar nuclei are formed by the same mitotic divisions that formed the egg.) From this develops a *3n* **triploid** endosperm.

G. The Chemical Composition of Cells

Water is the major component of living cells, but the amount varies greatly. Thus, the pig embryo is 97% water; at birth a new-born pig is only 89% water. A lean 45-kg pig may contain 67% water but a very fat 135-kg animal only 40% water. Similar variations are encountered with other constitutents.

The water content of a tissue is often determined by thoroughly drying a weighed sample of tissue at low temperature in vacuum and then weighing it a second time. The solid material can then be extracted with a solvent that will dissolve out the fatty compounds. These are referred to collectively as **lipids**. After evaporation of the solvent the lipid residue may be weighed. By this procedure a young leafy vegetable might be found to contain 2–5% lipid on a dry weight

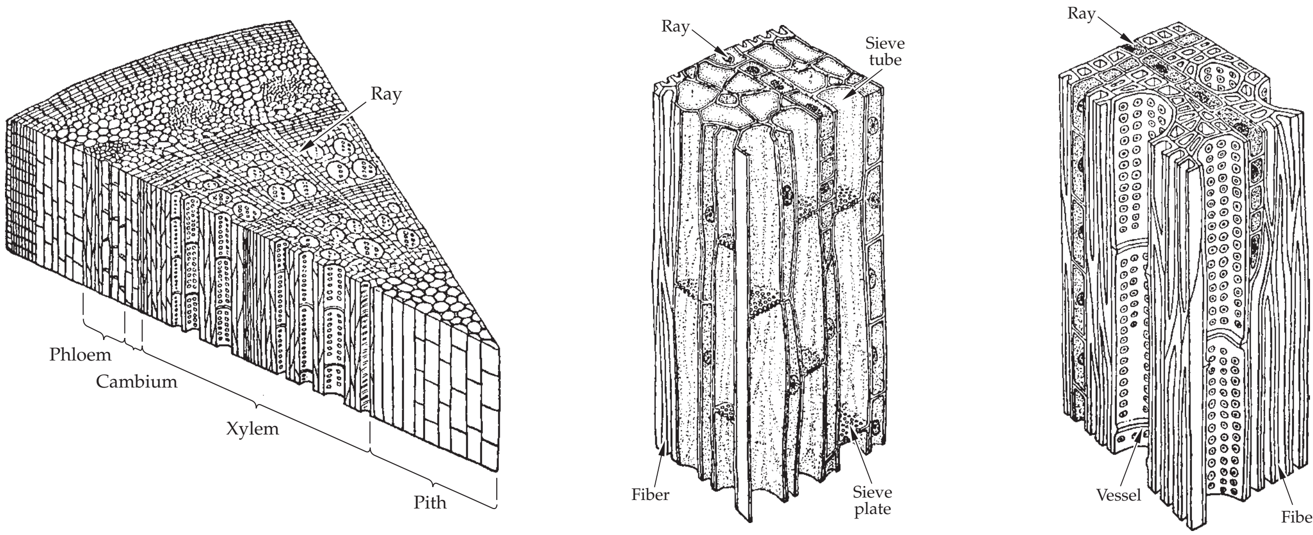


Figure 1-16 Section of the stem of an angiosperm. Enlarged sections showing tubes of the phloem (left) and xylem (right). From S. Biddulph and O. Biddulph.[155] Drawn by Bunji Tagawa.

basis. Even very lean meats contain 10–30% lipid.

The residue remaining after removal of the lipid consists predominately of three groups of compounds: **proteins**, **nucleic acids**, and **carbohydrates**. Most of the nitrogen present in tissues is found in the proteins and the protein content is sometimes estimated by determining the percentage of nitrogen and multiplying by 6.25. In a young green plant, 20–30% of the dry matter may be protein, while in very lean meat it may reach 50–70%.

TABLE 1-4
Approximate Composition of Metabolically Active Cells and Tissues[a]

Component	*E. coli*[b] (%)	Green plant (spinach, *Spinacia oleracea*)[c]	Rat liver[d] (%)
H_2O	70	93	69
Protein	15	2.3	21
Amino acids	0.4		
DNA	1		0.2
RNA	6		1.0
Nucleotides	0.4		
Carbohydrates	3	3.2	
Cellulose		0.6	
Glycogen			3.8
Lipids	2	0.3	6
Phospholipids			3.1
Neutral lipids			1.6
Sterols			0.3
Other small molecules	0.2		
Inorganic ions	1	1.5	
K^+			0.4

	Equivalents per liter in rat liver
Amino acid residues	2.1
Nucleotide units	0.03
Glycogen (glucose units)	0.22
K^+	0.1

[a] Data were not readily available for spaces left blank

[b] From J. D. Watson (1976) *Molecular Biology of the Gene*, 3rd ed., p. 69, Benjamin, New York The amounts of amino acids, nucleotides, carbohydrates, and lipids include precursors present in the cell.

[c] From B. T. Burton (1976) *Human Nutrition*, 3rd ed., McGraw-Hill, New York (p. 505)

[d] From C. Long, ed., (1961) *Biochemists' Handbook*, pp. 677–679, Van Nostrand-Reinhold, Princeton, New Jersey

A dried tissue sample may be burned at a high temperature to an **ash**, which commonly amounts to 3–10% and is higher in specialized tissues such as bone. It is a measure of the inorganic constituents of tissues.

The carbohydrate content can be estimated by the difference of the sum of lipid, protein, and ash from 100%. It amounts to 50–60% in young green plants and only 2–10% in typical animal tissues. In exceptional cases the carbohydrate content of animal tissues may be higher; the glycogen content of oysters is 28%.

The amount of nucleic acid in tissues varies from 0.1% in yeast and 0.5–1% in muscle and in bacteria to 15–40% in thymus gland and sperm cells. In these latter materials of high nucleic acid content it is clear that multiplication of % N by 6.25 is not a valid measure of protein content. For diploid cells of the body the DNA content per cell is nearly constant.

Table 1-4 compares the composition of a bacterium, of a green plant, and of an active animal tissue (rat liver). Although the solid matter of cells consists principally of C, H, O, N, S, and P, many other chemical elements are also present. Among the cations, Na^+, K^+, Ca^{2+}, and Mg^{2+} are found in relatively large amounts. Thus, the body of a 70 kg person contains 1050 g Ca (mostly in the bones), 245 g K, 105 g Na, and 35 g Mg. Iron (3 g), zinc (2.3 g), and rubidium (1.2 g) are the next most abundant. Of these iron and zinc are essential to life but rubidium is probably not. It is evidently taken up by the body together with potassium.

The other metallic elements in the human body amount to less than 1 g each, but at least seven of them play essential roles. They include copper (100 mg), manganese (20 mg), and cobalt (~5 mg). Others, such as chromium (<6 mg), tin, and vanadium, have only recently been shown essential for higher animals.[156,157] Nickel, lead, and others may perhaps be needed.

Nonmetallic elements predominating in the ash are phosphorus (700 g in the human body), sulfur (175 g), and chlorine (105 g). Not only are these three elements essential to all living cells but also selenium, fluorine, silicon (Box 4-B), iodine, and boron are needed by higher animals and boron by plants (Fig. 1-17). Iodine deficiency may affect one billion human beings and may cause 20 million cases per year of **cretinism**, or less severe brain damage.[158]

What is the likelihood that other elements will be found essential? Consider a human red blood cell, an object of volume ~80 μm^3 and containing about 3×10^8 protein molecules (mostly hemoglobin). About 7×10^5 atoms of the "trace metal" copper and 10^5 atoms of the nutritionally essential tin are present in a single red cell. Also present are 2×10^4 atoms of silver, a toxic metal. Its concentration, over 10^{-7} M, is sufficient that it could have an essential catalytic function. However, we know of none and it may simply have gotten into our bodies from handling money, jewelry, and other

BOX 1-F ABOUT THE REFERENCES

The lists of references at the ends of chapters are provided to encourage readers to look at original research articles. The lists are neither complete nor critically selected, but they do increase the information given in this book many-fold. I apologize for the important papers omitted. However, the references that are here will help a student to get started in reading the literature. Each reference contains other references and names of persons active in the field. By searching recent journal indices or a computer database it is easy to find additional articles by the same authors or on the same subject.

Look at the various types of scientific articles including reviews, preliminary reports and full research papers. Be sure to examine those in the *primary source journals* which publish detailed research results. These articles have always been sent to referees, active scientists, who check to see that the experiments are described accurately, that the authors have cited relevant literature, and that the conclusions are logical. Some journals, e.g., *Biochem. Biophys. Res. Comm.* and *FEBS Letts.* are dedicated to rapid publication of short reports but are also refereed. Other journals provide mostly reviews or a mixture of reviews. Periodical review series, such as *Advances in Nucleic Acid Chemistry and Related Topics*, often appear annually. Every student who intends to become a professional biochemist should consider purchasing the *Annual Review of Biochemistry* each year. This indispensible source of current information on most aspects of biochemistry is available to students at a very low price.

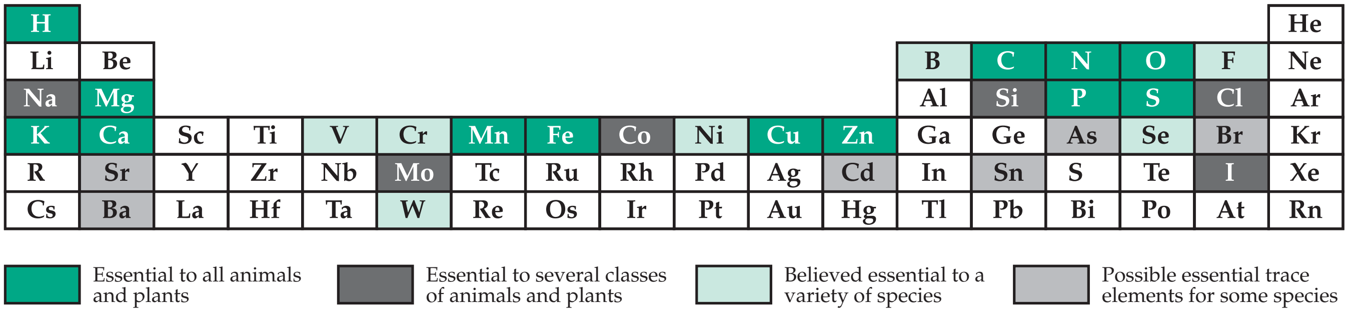
Many journal papers are difficult to read. To start, pick papers that have an understandable introduction. Choose reviews that are short, such as those in *Trends in Biochemical Sciences.* Then go on to the more comprehensive ones. Never sit back and hope that your computer will automatically fetch just what you need! Many journals carry papers of biochemical importance. Those specializing in biochemistry include the following:

Full Title	**Abbreviation**
Advances in Carbohydrate Chemistry and Biochemistry[a]	Adv. Carbohydr. Chem. Biochem.
Advances in Protein Chemistry[a]	Adv. Protein Chem.
Analytical Biochemistry	Anal. Biochem.
Annual Review of Biophysics and Biomolecular Structure[a]	Ann. Rev. Biophys. Biomolec. Struct.
Annual Review of Biochemistry[a]	Ann. Rev. Biochem.
Archives of Biochemistry and Biophysics	Arch. Biochem. Biophys.
Biochemical and Biophysical Research Communications	Biochem. Biophys. Res. Commun.
Biochemical Journal	Biochem. J.
Biochemistry	
Biochimica et Biophysica Acta	Biochim. Biophys. Acta
Bioorganic Chemistry	Bioorg. Chem.
Carbohydrate Research	Carbohydr. Res.
EMBO Journal[b]	EMBO J.
European Journal of Biochemistry	Eur. J. Biochem.
FASEB Journal[c]	FASEB J.
Journal of Bacteriology	J. Bacteriol.
Journal of Biochemistry	J. Biochem.
Journal of Biological Chemistry	J. Biol. Chem.
Journal of Lipid Research	J. Lipid Res.
Journal of Molecular Biology	J. Mol. Biol.
Journal of the American Chemical Society	J. Am. Chem. Soc.
Journal of Theoretical Biology	J. Theor. Biol.
Methods in Enzymology[a]	Methods Enzymol.
Nature	
Nucleic Acids Research	Nucleic Acids Res.
Proceedings of the National Academy of Sciences, USA	Proc. Natl. Acad. Sci. U.S.A.
Science	
Structure	
Trends in Biochemical Sciences	Trends Biochem. Sci. or TIBS

[a] These are not journals but series of review and reference books. There are many other series of “Advances in...” and “Annual Reviews of ...” that are not listed here.

[b] EMBO – European Molecular Biology Organization

[c] FASEB – Federation of American Societies for Experimental Biology



H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
R	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	S	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn

🟩 Essential to all animals and plants

⬛ Essential to several classes of animals and plants

▨ Believed essential to a variety of species

░ Possible essential trace elements for some species

Figure 1-17 Elements known to be essential to living things (after da Silva and Williams[157]). Essential elements are enclosed within shaded boxes. The 11 elements–C, H, O, N, S, P, Na, K, Mg, Ca, and Cl–make up 99.9% of the mass of a human being. An additional 13 are known to be essential for higher animals in trace amounts. Boron is essential to higher plants but apparently not to animals, microorganisms, or algae.

silver objects. The red blood cell also contains boron and aluminum (3×10^5 atoms each), arsenic (7×10^5 atoms), lead (7×10^4 atoms), and nickel (2×10^4 atoms). Of the elements (uranium and below) in the periodic table, only four (Ac, Po, Pa, and Ra) are present, on the average, in quantities less than one atom per cell.[156]

Of the apparently nonessential elements, several, e.g., Cs, Rb, Sr, and Ni (possibly essential) are not toxic at low concentrations. Others, such as Sb, As, Ba, Be, Cd, Pb, Hg, Ag, Tl, and Th, are highly toxic.

The ionic compositions of tissues and of body fluids vary substantially. Blood of marine organisms is similar to that of seawater in its content of Na^+, Cl^-, Ca^{2+}, and Mg^{2+}. Blood of freshwater and terrestrial organisms contains about ten times less Na^+ and Cl^- and several times less Ca^{2+} and Mg^{2+} than is present in seawater, but it is nevertheless relatively rich in these ions.

In general, cells are rich in K^+ and Mg^{2+}, the K^+ predominating by far, and are poor in Na^+ and Ca^{2+}. Chloride is the principal inorganic anion, but organic carboxylate and phosphate groups contribute most of the negative charges (Table 1-4), many of which are fixed to proteins or other macromolecules. Ling estimated that cells typically contain about 1.66 M of amino acid residues in their proteins. Of these residues, 10% have negatively charged side chains and 8% positively charged. The difference is a net negative charge amounting to 33 mM within cells.[159.]

References

1. May, R. M. (1992) *Sci. Am.* **267**(Oct), 42–48
2. Stanier, R. Y., Ingraham, J. L., Wheelis, M. L., and Painter, P. R. (1986) *The Microbial World*, 5th ed., Prentice-Hall, Englewood Cliffs, New Jersey
3. Brock, T. D., Smith, D. W., and Madigan, M. T. (1988) *Biology of Microorganisms*, 5th ed., Prentice-Hall, Englewood Cliffs, New Jersey
4. Neidhardt, F. C., ed. (1987) *Escherichia coli and Salmonella typhimurium*, Am. Soc. for Microbiology, Washington, D.C.
5. Neidhardt, F. C., Ingraham, J. L., and Schaechter, M. (1990) *Physiology of the Bacterial Cell*, Sinauer, Sunderland, Mass.
6. Maniloff, J., and Morowitz, H. J. (1972) *Bacteriol. Rev.* **36**, 263–290
7. Razin, S. (1978) *Microbiol. Rev.* **42**, 414–470
8. Barile, M. F., Razin, S., Tully, J. G., and Whitcomb, R. F., eds. (1979) *The Mycoplasmas*, Vol. I, Academic Press, New York
9. Watson, J. D. (1976) *Molecular Biology of the Gene*, 3rd ed., Benjamin, Menlo Park, California (p. 61)
10. Horowitz, D. M., and Sanders, J. K. M. (1994) *J. Am. Chem. Soc.* **116**, 2695–2702

10a. Gansheroff, L. J., and O'Brien, A. D. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2959–2961

11. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J.-F., Dougherty, B. A., Bott, K. F., Hu, P.-C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchison, C. A., III, and Venter, J. C. (1995) *Science* **270**, 397–403
12. Sonea, S. (1988) *Nature (London)* **331**, 216
13. Costerton, J. W., Geesey, G. G., and Cheng, K.-J. (1978) *Sci. Am.* **238**(Jan), 86–95
14. Macnab, R. M. (1987) in *Escherichia coli and Salmonella typhimurium* (Niedhardt, F. C., ed), pp. 70–83, Am. Soc. for Microbiology, Washington, D.C.
15. Margulis, L., Ashen, J. B., Sonea, S., Solé, M., and Guerrero, R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6966–6970
16. Goldstein, S. F., Charon, N. W., and Kreiling, J. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3433–3437
17. Levine, M. M. (1985) *N. Engl. J. Med.* **313**, 445–447
18. Kuehn, M. J., Heuser, J., Normark, S., and Hultgren, S. J. (1992) *Nature (London)* **356**, 252–255
19. Alper, J. (1993) *Science* **262**, 1817
20. Holt, J. G., Krieg, N. R., Sneath, P. H. A., Staley, J. T., and Williams, S. T., eds. (1994) *Bergey's Manual of Determinative Bacteriology*, 9th ed., Williams & Wilkins, Baltimore, Maryland
21. Doolittle, R. F. (1992) *Protein Sci.* **1**, 191 – 200
22. Schwartz, R. M., and Dayhoff, M. O. (1978) *Science* **199**, 395–403
23. Cavalier-Smith, T. (1986) *Nature (London)* **324**, 416–417
24. Jones, W. J., Nagle, D. P., Jr., and Whitman, W. B. (1987) *Microbiol. Rev.* **51**, 135–177
25. Woese, C. R. (1987) *Microbiol. Rev.* **51**, 221–271
26. Sogin, M. L., Gunderson, J. H., Elwood, H. J., Alonso, R. A., and Peattie, D. A. (1989) *Science* **243**, 75–77
27. Garrett, R. A., Dalgaard, J., Larson, N., Kjems, J., and Mankin, A. S. (1991) *Trends Biochem. Sci.* **16**, 22–26
28. Searcy, D. G. (1982) *Trends Biochem. Sci.* **7**, 183–185
29. Wheelis, M. L., Kandler, O., and Woese, C. R. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2930–2934
30. Woese, C. R., Kandler, O., and Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576–4579

30a. Doolittle, W. F. (1999) *Science* **284**, 2124–2128

30b. Pennisi, E. (1999) *Science* **284**, 1305–1307

31. Cavalier-Smith, T. (1992) *Nature (London)* **356**, 570
32. Angert, E. R., Clements, K. D., and Pace, N. R. (1993) *Nature (London)* **362**, 239–241
33. Thimann, K. V. (1963) *The Life of Bacteria*, 2nd ed., Macmillan, New York
34. Stanier, R. Y., and Cohen-Bazire, G. (1977) *Ann. Rev. Microbiol.* **31**, 225–274
35. Morden, C. W., and Golden, S. S. (1989) *Nature (London)* **337**, 382–385
36. Marme, D., Marre, E., and Hertel, R., eds. (1982) *Plasmalemma and Tonoplast: Their Functions in the Plant Cell*, Elsevier, Amsterdam
37. Collins, F. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10821–10823
38. Hinshaw, J. E., Carraghen, B. O., and Milligan, R. A. (1992) *Cell* **69**, 1133–1141
39. Dingwall, C., and Lasky, R. (1992) *Science* **258**, 942–947
40. Nolta, K. V., Padh, H., and Steck, T. L. (1991) *J. Biol. Chem.* **266**, 18318–18323
41. de Duve, C. (1975) *Science* **189**, 186–194
42. Griffiths, G., and Simons, K. (1986) *Science* **234**, 438–443
43. Hartley, S. M. (1992) *Trends Biochem. Sci.* **17**, 325–327
44. Golgi, C. (1898) *Arch. Ital. Biol.* **30**, 60 and 278
45. Kirk, J. T. O., and Tilney-Bassett, R. A. E. (1978) *The Plastids: Their Chemistry, Structure, Growth, and Inheritance*, 2nd ed., Elsevier-North Holland Biomedical Press, Amsterdam
46. Hoffmann, H. P., and Avers, C. J. (1973) *Science* **181**, 749–751

46a. Capaldi, R. A. (2000) *Trends Biochem. Sci.* **25**, 212–214

46b. Rutter, G. A., and Rizzuto, R. (2000) *Trends Biochem. Sci.* **25**, 215–221

47. Fukui, S., and Tanaka, A. (1979) *Trends Biochem. Sci.* **4**, 246–249
48. Kindl, H., and Lazarow, P. B., eds. (1982) *Peroxisomes and Glyoxysomes*, Vol. 386, Annals of the New York Academy of Sciences, New York
49. de Duve, C. (1983) *Sci. Am.* **248**(May), 74–84
50. Masters, C., and Crane, D. (1995) *The Peroxisome: A Vital Organelle*, Cambridge Univ. Press, London
51. Lord, J. M., and Roberts, L. M. (1980) *Trends Biochem. Sci.* **5**, 271–274
52. Cerkasov, J., Cerkasovova, A., Kulda, J., and Vilhelmova, D. (1978) *J. Biol. Chem.* **253**, 1207–1214
53. Glover, D. M., Gonzalez, C., and Raff, J. W. (1993) *Sci. Am.* **268**(Jun), 62–68
54. Olmsted, J. B., and Borisy, G. G. (1973) *Ann. Rev. Biochem.* **42**, 507–540
55. Hall, J. L., Ramanis, Z., and Luck, D. J. L. (1989) *Cell* **59**, 121–132
56. Hyams, J. S. (1989) *Nature (London)* **341**, 485–486
57. Satir, P. (1961) *Sci. Am.* **204**(Feb), 108–116
58. Lazarides, E. (1980) *Nature (London)* **283**, 244–256
59. Amos, L. A., and Amos, W. B. (1991) *Molecules of the Cytoskeleton*, Guilford, New York
60. Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., and Miyata, T. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9355–9359
61. Lake, J. A. (1991) *Trends Biochem. Sci.* **16**, 46–50
62. Margulis, L. (1981) *Symbiosis in Cell Evolution*, Freeman, San Francisco, California
63. Hurst, L. D. (1994) *Nature (London)* **369**, 451
64. Gupta, R. S., and Golding, G. B. (1996) *Trends Biochem. Sci.* **21**, 166–171
65. de Duve, C. (1996) *Sci. Am.* **274** (Apr), 50–57
66. Markert, C. L., Shaklee, J. B., and Whitt, G. S. (1975) *Science* **189**, 102–114
67. Fenchel, T., and Bernard, C. (1993) *Nature (London)* **362**, 300
68. Schopf, J. W. (1993) *Science* **260**, 640–646
69. Horgan, J. (1993) *Sci. Am.* **269**(Aug), 24
70. Doolittle, W. F. (1980) *Trends Biochem. Sci.* **5**, 146–149
71. May, R. M. (1986) *Nature (London)* **324**, 514–515
72. Seger, J. (1989) *Nature (London)* **337**, 305–306
73. Riley, M., and Anilionis, A. (1978) *Ann. Rev. Microbiol.* **32**, 519–560
74. Goossens, M., Dozy, A. M., Embury, S. H., Zachariades, Z., Hadjiminas, M. G., Stamatoyannopoulos, G., and Kan, Y. W. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 518–521
75. Williams, R. O., Young, J. R., and Majiwa, P. A. O. (1979) *Nature (London)* **282**, 847–849
76. Kondrashov, A. S. (1988) *Nature (London)* **336**, 435–440
77. Buchsbaum, R., Buchsbaum, M., Pearse, J., and Pearse, V. (1987) *Animals Without Backbones*, 3rd ed., Univ. Chicago Press, Chicago, Illinois
78. Margulis, L., and Schwartz, K. V. (1999) *Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth*, 3rd ed., Freeman, New York
79. Laver, W. G., Air, G. M., Dopheide, T. A., and Ward, C. W. (1980) *Nature (London)* **283**, 454–457
80. McKanna, J. A. (1973) *Science* **179**, 88–90
81. Miller, L. H. (1992) *Science* **257**, 36–37
82. Wyler, D. J. (1993) *N. Engl. J. Med.* **329**, 31–37
83. Miller, L. H., Good, M. F., and Milon, G. (1994) *Science* **264**, 1878–1883
84. Nussenzweig, R. S., and Long, C. A. (1994) *Science* **265**, 1381–1382
85. Sibley, L. D., and Boothroyd, J. C. (1992) *Nature (London)* **359**, 82–85
86. Soldati, D., and Boothroyd, J. C. (1993) *Science* **260**, 349–352
87. Donelson, J. E., and Turner, M. J. (1985) *Sci. Am.* **252**(Feb), 44–51
88. Blum, M. L., Down, J. A., Gurnett, A. M., Carrington, M., Turner, M. J., and Wiley, D. C. (1993) *Nature (London)* **362**, 603–609
89. Gall, J. G., ed. (1986) *The Molecular Biology of Ciliated Protozoa*, Academic Press, Orlando, Florida
90. Olive, L. S. (1975) *The Mycetozoans*, Academic Press, New York
91. Kimmel, A. R., ed. (1989) *Molecular Biology of Dictyostelium Development*, Liss, New York
92. Traynor, D., Kessin, R. H., and Williams, J. G. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8303–8307
93. Webster, J. (1980) *Introduction to Fungi*, 2nd ed., Cambridge Univ. Press, New York
94. Davies, D. A. L., and Pope, A. M. S. (1978) *Nature (London)* **273**, 235–236
95. Georgopapadakou, N. H., and Walsh, T. J. (1994) *Science* **264**, 371–373
96. Sternberg, S. (1994) *Science* **266**, 1632–1634
97. Davis, R. H., and DeSerres, F. J. (1970) *Meth. Enzymol.* **17A**, 79–148
98. Botstein, D., and Fink, G. R. (1988) *Science* **240**, 1439–1443
99. Pringle, J. R., Broach, J. R., and Jones, E. W., eds. (1991–1994) *The Molecular and Cellular Biology of the Yeast Saccharomyces*, Cold Springs Harbor Laboratory Press, Cold Spring Harbor, New York (3 Vols.)

References

100. Wheals, A. E., Rose, A. H., and Harrison, S. J., eds. (1995) *The Yeasts*, 2nd ed., Vol. 6, Academic Press, San Diego, California
101. Gimeno, C. J., Ljungdahl, P. O., Styles, C. A., and Fink, G. R. (1992) *Cell* **68**, 1077–1090
102. Oliver, S. G., and 146 other authors (1992) *Nature (London)* **357**, 38–46
103. Williams, N. (1995) *Science* **268**, 1560–1561
104. Ruehle, J. L., and Marx, D. H. (1979) *Science* **206**, 419–422
105. Canter-Lund, H., and Lund, J. W. G. (1995) *Freshwater Algae: Their Microscopic World Explored*, Biopress, Bristol, England
106. Harris, E. H. (1989) *The Chlamydomonas Sourcebook*, Academic Press, San Diego, California
107. Brachet, J. L. A. (1965) *Endeavour* **24**, 155–161
108. Bonotto, S., Kefeli, V., and Puiseux-Dao, S. (1979) *Developmental Biology of Acetabularia*, Elsevier-North Holland, Amsterdam
109. Berger, S., and Kaever, M. J. (1992) *Dasycladales: An Illustrated Monograph of a Fascinating Algal Order*, Oxford Univ. Press, New York
110. Menzel, D., and Elsner-Menzel, C. (1990) *Protoplasm* **157**, 52–63
111. Jaenicke, L. (1982) *Trends Biochem. Sci.* **7**, 61–64
112. Ahmadjian, V. (1962) in *Physiology and Biochemistry of Algae*, Vol. 2 (Lewin, R. A., ed), pp. 817–822, Academic Press, New York
113. Morris, S. C. (1993) *Nature (London)* **361**, 219–221
114. Levinton, J. S. (1992) *Sci. Am.* **267**(Nov), 84–91
115. Keese, P. K., and Gibbs, A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9489–9493
116. Hickman, C. P. (1973) *Biology of the Invertebrates*, Mosby, St. Louis, Missouri
117. Villee, C. A., Walker, W. F., Jr., and Barnes, R. D. (1973) *General Zoology*, Saunders, Philadelphia, Pennsylvania
118. Lapan, E. A., and Morowitz, H. (1972) *Sci. Am.* **227**(Dec), 94–101
119. Lapan, E. A. (1975) *Comp. Biochem. Physiol.* **52A**, 651–657
120. Bergquist, P. R. (1978) *Sponges*, Hutchinson, London
121. Grimmelikhuijzen, C. J. P., and Schaller, H. C. (1979) *Trends Biochem. Sci.* **4**, 265–267
122. Lenhoff, H. M., and Lenhoff, S. G. (1988) *Sci. Am.* **258**(Apr), 108–113
123. Loomis, W. F. (1959) *Sci. Am.* **200**(Apr), 150
123a. Hickman, C. P. *Integrated Principles of Zoology* (1966) Mosby, St. Louis
124. Trager, W. (1986) *Living Together: The Biology of Animal Parasitism*, Plenum, New York
125. Cherfas, J. (1991) *Science* **251**, 630–631
126. Kolberg, R. (1994) *Science* **264**, 1859–1861
127. Wood, W. B., ed. (1988) *The Nematode Caenorhabditis elegans*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
128. Kenyon, C. (1988) *Science* **240**, 1448–1453
129. Roberts, L. (1990) *Science* **248**, 1310–1313
129a. Hotez, P. J., and Pritchard, D. I. (1995) *Sci. Am.* **272**(Jun), 68–74
130. Dumont, H. J., and Green, J., eds. (1980) *Rotatoria*, p. 42, Junk, The Hague
131. Mill, P. J., ed. (1978) *Physiology of Annelids*, Academic Press, New York
132. Florkin, M., and Scheer, B. T., eds. (1970, 1971) *Chemical Zoology*, Vol. 5 and 6, Academic Press, New York
133. Goldstein, L. S. B., and Fyrberg, E. A., eds. (1994) *Drosophila melanogaster: Practical Uses in Cell and Molecular Biology*, Academic Press, San Diego, California
134. Lawrence, P. (1992) *The Making of a Fly: The Genetics of Animal Design*, Blackwell, Oxford
134a. Rubin, G. M., and Lewis, E. B. (2000) *Science* **287**, 2216–2218
135. Maddy, A. H. (1992) *Trends Biochem. Sci.* **17**, 125–126
135a. Harris, J. R., ed. (1991) *Blood Cell Biochemistry*, Plenum, New York (Three volumes)
136. Johnston, R. B., Jr. (1988) *N. Engl. J. Med.* **318**, 747–754
137. Freshney, R. I. (1994) *Culture of Animal Cells*, 3rd ed., Wiley-Liss, New York
138. Feldman, J., Gilula, N. B., and Pitts, J. D., eds. (1978) *Intercellular Junctions and Synapses*, Halsted Press, New York
139. Staehelin, L. A., and Hull, B. E. (1978) *Sci. Am.* **238**(May), 141–153
140. Bradbury, M. (1979) *The Concept of a Blood-Brain Barrier*, Wiley, Chichester, New York
141. Friend, D. S., and Gilula, N. B. (1972) *J. Cell Biol.* **53**, 771
142. Gilula, N. B. (1975) *Cellular Membranes and Tumor Cells*, Williams & Wilkins Co., Baltimore, Maryland (p. 221)
143. Gilula, N. B. (1974) in *Cell Communication* (Cox, R. P., ed), pp. 1–29, Wiley, New York
144. Gilula, N. B. (1972) *Nature (London)* **235**, 262–265
145. Zimmer, D. B., Green, C. R., Evans, W. H., and Gilula, N. B. (1987) *J. Biol. Chem.* **262**, 7751–7763
146. Hertzberg, E. L., and Johnson, R. G., eds. (1988) *Gap Junctions*, Liss, New York
147. Robinson, S. R., Hampson, E. C. G. M., Munro, M. N., and Vaney, D. I. (1993) *Science* **262**, 1072–1074
148. Stauffer, P. L., Zhao, H., Luby-Phelps, K., Moss, R. L., Star, R. A., and Muallem, S. (1993) *J. Biol. Chem.* **268**, 19769–19775
149. Veenstra, R. D., and DeHaan, R. L. (1986) *Science* **233**, 272–274
150. Moscona, A. A. (1961) *Sci. Am.* **205**(Sep), 143–162
151. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1983) *Molecular Biology of the Cell*, pp. 679–680, Garland Publ., New York
152. Heaysman, J. E. M. (1978) *Intl. Rev. Cytol.* **55**, 49–66
153. Palca, J. (1989) *Science* **245**, 131
154. Lloyd, A. M., Barnason, A. R., Rogers, S. G., Byrne, M. C., Fraley, R. T., and Horsch, R. B. (1986) *Science* **234**, 464–466
155. Biddulph, S., and Biddulph, O. (1959) *Sci. Am.* **200**(Feb), 44–49
156. Bowen, H. J. M. (1966) *Trace Elements in Biochemistry*, Academic Press, New York
157. Fraústo da Silva, J. J. R., and Williams, R. J. P. (1991) *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*, Clarendon Press, Oxford
158. Hetzel, B. S. (1994) *N. Engl. J. Med.* **331**, 1770–1771
159. Ling, G. N. (1962) *A Physical Theory of the Living State*, Ginn (Blaisdell), Boston, Massachusetts

Study Questions

1. Describe the principal structural or organizational differences between prokaryotic and eukaryotic cells.

2. Describe two or more principal functions of proteins within cells, one function of DNA, two or more functions of RNA, and one function of lipids.

3. Compare the chemical makeup of ribosomes, of cell membranes, and of bacterial flagella.

4. Assume the following dimensions: Mycoplasma, sphere, 0.33 µm diameter; *E. coli*, cylinder, 0.8 µm diameter x 2 µm; liver cell, sphere 20 µm; root hair, cylinder, 10 µm diameter x 1 mm.

   a. Calculate for each cell the total volume, the mass in grams and in daltons (assume a specific gravity of 1.0).

   b. Assume that bacterial ribosomes are approximately spherical with a diameter of 23 nm. What is their volume? If the mass of a bacterial ribosome is 2.7×10^6 daltons, what is its apparent density (divide mass by volume)? Experimentally the buoyant density of bacterial ribosomes in a cesium chloride gradient (Chapter 5) is about 1.6 g/cm^3. How can this difference be explained? If eukaryotic ribosomes are 1.17 times larger than bacterial ribosomes in linear dimensions, what is the volume of a eukaryotic ribosome?

   c. What fraction of volume of *E. coli* consists of cell wall, of plasma membrane, of ribosomes (assume 15,000 are present)? If a cell of *E. coli* is 80% water, what fraction by weight of the total solids consists of ribosomes? Of DNA (assuming 2 chromosomes per cell)?

   d. What fraction by volume of a liver cell is composed of ribosomes, of plasma membrane, of mitochondria (assume 1000 mitochondria)? What fraction is accounted for by the nucleus?

5. a. What is the molar concentration of an enzyme of which only one molecule is present in an *E. coli* cell?

   b. Assume that the concentration of K^+ within an *E. coli* cell is 150 mM. Calculate the number of K^+ ions in a single cell.

   c. If the pH inside the cell is 7.0, how many H^+ ions are present?

6. If chromosomes (and chromatin) are 15% DNA, what will be the mass of 23 pairs of chromosomes in a human diploid cell? If the nucleus has a diameter of 5 mm and a density of 1.1 g/cm^3, what fraction by weight of the nucleus is chromatin?

7. Compare the surface to volume ratios for an *E. coli* cell, a liver cell, the nucleus of a eukaryotic cell, a root hair. If a cell of 20 µm diameter is 20% covered with microvilli of 0.1 µm diameter and 1 µm length centered on a 0.2 µm spacing, how much will the surface/volume ratio be increased?

8. It has been shown that the code for specifying a particular amino acid in a protein is determined by a sequence of three nucleotides (a codon) in a DNA chain. There are four different kinds of nucleotide units in DNA. How many different codons exist? Note that this is larger than the number of different amino acids (20) that are incorporated into proteins plus the three stop (termination) codons (see Tables 5-5 and 5-6 for a list of codons).

Study Questions

9. State two similarities and two differences between cyanobacteria (blue-green algae) and green algae.

10. Compare the sizes and structures of bacterial and eukaryotic flagella.

11. How much larger in volume is a typical eukaryotic cell compared to a bacterium?

12. Compare the structure and properties of mitochondria, chloroplasts, and peroxisomes.

13. What are the possible origins of mitochondria and chloroplasts? What evidence can you cite to support your answer?

14. How many different kinds of polymers, e.g. proteins, RNA, that are present in or around living cells, can you name? Can you name some subgroups in any of your categories?

15. Compare the composition of these three, especially with respect to C, H, O, N, S, P, Fe, Cu, Al, and Si:

    a. The earth's crust

    b. Ocean water

    c. Cytoplasm

A spider's orb-web is formed by extrusion of a concentrated protein solution and stretching of the resulting fiber. The cross-strands, which are stronger than steel, resemble silkworm silk. The molecules contain microcrystalline β sheet domains that are rich in Gly-Ala repeats as well as polyalanine segments. The capture spiral is formed from much more elastic molecules that contain many β-turn-forming sequences. These assume a spring-like β spiral. See Box 2-B.

© John Gerlach/DRK PHOTO

Contents

39 .... **A. Structural Principles for Small Molecules**
39 .... 1. Bond Angles
40 .... 2. Bond Lengths
40 .... 3. Contact Distances
41 .... 4. Asymmetry: Right-Handed and Left-Handed Molecules
42 .... *The D- and L-families of amino acids*
42 .... *The RS notation for configuration*
43 .... *Diastereoisomers*
43 .... *Geometrical isomers*
43 .... 5. Conformations: The Shapes that Molecules can Assume
45 .... 6. Tautomerism and Resonance
46 .... **B. Forces between Molecules and between Chemical Groups**
46 .... 1. Van der Waals Forces
47 .... 2. Attraction between Charged Groups (Salt Linkages)
47 .... 3. Hydrogen Bonds
49 .... 4. The Structure and Properties of Water
50 .... 5. Hydration of Polar Molecules and Ions
50 .... 6. Hydrophobic Interactions
51 .... **C. Amino Acids and Peptides**
51 .... 1. Properties of α-Amino Acids
55 .... 2. Acidic and Basic Side Chains
55 .... 3. The Peptide Unit
56 .... 4. Polypeptides
59 .... **D. The Architecture of Folded Proteins**
59 .... 1. Conformations of Polypeptide Chains
61 .... 2. The Extended Chain β Structures
62 .... *Pleated sheets*
63 .... *Twisted sheets*
64 .... *Properties of β sheets*
65 .... *Cylinders and barrels*
67 .... *Beta propellers*
68 .... 3. Helices
69 .... *Properties of helices*
70 .... *Stacking of helices in proteins*
70 .... *Coiled coils*
71 .... *Helix bundles*
72 .... 4. Polyglycine II and Collagen
72 .... 5. Turns and Bends
74 .... 6. Domains, Subunits, and Interfaces
75 .... 7. Packing of Side Chains
75 .... 8. The Network of Internal Hydrogen Bonds
76 .... **E. Folding Patterns and Protein Families**
76 .... 1. Complex Folding Patterns
77 .... 2. Symmetry
78 .... 3. Effects of Sequence on Folding
79 .... **F. Chemical Modification and Crosslinking**
80 .... 1. Disulfide Crosslinkages
80 .... 2. Other Crosslinkages
81 .... **G. Dynamic Properties of Proteins**
81 .... 1. Motion of Backbone and Side Chains
81 .... 2. Conformational Changes
82 .... 3. Denaturation and Refolding
83 .... 4. Effects of pH and Solvent
84 .... 5. Irreversible Damage to Proteins
85 .... **H. Design and Engineering of Proteins**

86 .... **References**
91 .... **Study Questions**

Boxes

58 .... Box 2-A Proteins of Blood Plasma
67 .... Box 2-B Silks
83 .... Box 2-C The Nobel Prizes

Tables

41 .... Table 2-1 The Sizes of Some Atoms
52 .... Table 2-2 Structure and Chemical Properties of Side Chain Groups of Amino Acids
61 .... Table 2-3 Approximate Torsion Angles for Some Regular Peptide Structures
78 .... Table 2-4 Classification of Protein Residues According to Their Tendencies to Form α Helix, β Structure, and β Turns

Amino Acids, Peptides, and Proteins

2

Thousands of different proteins make up a very large fraction of the "machinery" of a cell. Protein molecules catalyze chemical reactions, carry smaller molecules through membranes, sense the presence of hormones, and cause muscle fibers to move. Proteins serve as structural materials within cells and between cells. Proteins of blood transport oxygen to the tissues, carry hormones between cells, attack invading bacteria, and serve in many other ways. No matter what biological process we consider, we find that a group of special proteins is required.

The amino acid units that make up a protein molecule are joined together in a precise sequence when the protein is made on a ribosome. The chain is then folded, often into a very compact form. Sometimes the chain is then cut in specific places. Pieces may be discarded and parts may be added. A metal ion, a coenzyme derived from a vitamin, or even a single methyl group may be attached to form the biologically active protein. The final product is a complex and sophisticated machine, often with moving parts, that is exquisitely designed for its particular role.

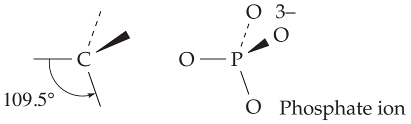
The biological functioning of a protein is determined both by the properties of the chemical groups in the amino acids that are joined to form the protein chain and by the way the chain is folded. The ways in which the different parts of the protein interact with each other and with other molecules are equally important. These interactions play a major role in determining the folding pattern and also provide much of the basis for the biological functioning of proteins. Similar considerations apply also to carbohydrates, nucleic acids, and other biopolymers. For these reasons it is appropriate to review some fundamentals of molecular structure and geometry.

A. Structural Principles for Small Molecules

Stable organic molecules are held together by covalent bonds which are usually very strong. The standard Gibbs energies of formation $(\Delta G_f^\circ)^\dagger$ of many covalent single bonds are of the order of −400 kJ/mol (96 kcal/mol). The bonds have definite directions, which are measured by **bond angles** and definite **bond lengths**.

1. Bond Angles

Because of the tetrahedral arrangement of the four bonds around single-bonded carbon atoms and most phosphorus atoms, all six of the bond angles about the central atom have nearly the same tetrahedral angle of 109.5°.



Bond angles within chains of carbon atoms in organic compounds vary only slightly from this, and even atoms that are attached to fewer than four groups usually have similar angles; for example, the H–O–H angle in a water molecule is 105°, and the H–N–H angles of ammonia are 107°. In ethers the C–O–C angle is 111°. However, bond angles of only 101° are present in H_2O_2 and of 92° in H_2S and PH_3.

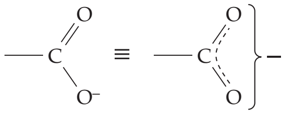
[†] **See Chapter 6 for a review of thermodynamics**

The presence of **double bonds** leads to **planarity** and to compounds with bond angles of 120°, the internal angle in a hexagon. The planar geometry imposed upon an atom by a double bond is often transmitted to an adjacent nitrogen or oxygen atom as a result of **resonance** (Section 6). For example, the amide groups that form the peptide linkage in proteins (see Fig. 2-5) are nearly planar and the angles all fall within four degrees of 120°.

2. Bond Lengths

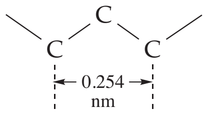
Chemists describe bond lengths as the distances between the nuclei of bonded atoms. The C–C single bond has a length of 0.154 nm (1.54 Å). The C–O bond is ~0.01 nm shorter (0.143 nm), and the typical C–H bond has a length of ~0.109 nm. The C–N bond distance is halfway between that for C–C and C–O (0.149 nm). Other lengths, such as that of O–H, can be estimated from the covalent radii given in Table 2-1.

The length of a double bond between any two atoms (e.g., C=C) is almost exactly 0.020 nm less than that for a single bond between the same atoms. If there is resonance, hence only partial double bond character, the shortening is less. For example, the length of the C–C bond in benzene is 0.140 nm; the C–O distances in the carboxylate anion are 0.126 nm.

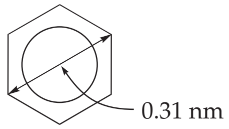


Carboxylate ion

Using simple geometry, it is easy to calculate overall lengths of molecules; here are two distances worth remembering:

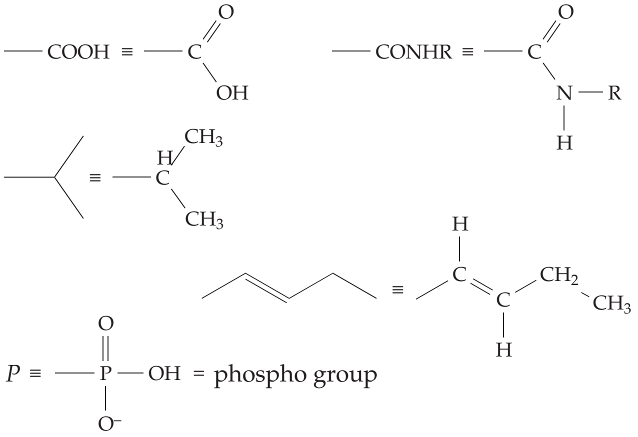


Distance between alternate C atoms in fully extended hydrocarbon chain



Distance across a benzene ring

In the preceding simplified structural formula for benzene the six hydrogen atoms have been omitted. Resonance between the two possible arrangements of the three double bonds[1] is indicated by the circle. Chemical shorthand of the following type is used throughout the book. Carbon atoms may be represented by an angle or the end of a line, but other atoms will always be shown.



3. Contact Distances

Covalent bond distances and angles tell us how the atomic nuclei are arranged in space but they do not tell us anything about the outside surfaces of molecules. The distance from the center of an atom to the point at which it contacts an adjacent atom in a packed structure such as a crystal (Fig. 2-1) is known as the **van der Waals radius**. The ways in which biological molecules fit together are determined largely by the van der Waals contact radii. These, too, are listed in Table 2-1. In every case they are approximately equal to the *covalent radius plus 0.08 nm*. Van der Waals radii

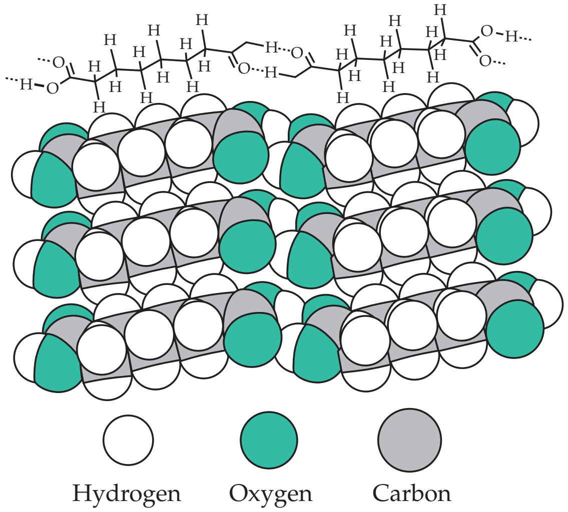


Figure 2-1 Packing of molecules of suberic acid HOOC–$(CH_2)_6$–COOH in a crystal lattice as determined by neutron diffraction.[2] Notice the pairs of hydrogen bonds that join the carboxyl groups at the ends of the molecules and also the close contact of hydrogen atoms between the chains. Only the positions of the hydrogen nuclei were determined; the van der Waals radii have been drawn around them. However, the radii were originally determined from X-ray and neutron diffraction data obtained from many different crystalline compounds.

are not as constant as covalent radii because atoms can be "squeezed" a little, but only enough to decrease the contact radii by 0.005–0.01 nm. The radii of space-filling molecular models are usually made a little smaller than the actual scaled van der Waals radii to permit easier assembly.

4. Asymmetry: Right-Handed and Left-Handed Molecules

The left hand looks much like the right hand, but they are different. One is the mirror image of the other. A practical difference is that your right hand will not fit into a left-handed glove. Despite our daily acquaintance with "handedness" it may seem difficult to explain in words how a right and a left hand differ. However, since most biochemical compounds are asymmetric,[3] it is important to be able to visualize these molecules in three dimensions and to draw their structures on paper. One of the best ways of learning to do this is to study molecular models. You may learn the most by making your own models (see Appendix).

TABLE 2-1
The Sizes of Some Atoms[a–c]

Element	Covalent radius (nm)[a,b,d]	van der Waals radii (nm)[d] ***a*** Sideways contact[c]	 ***b*** Polar contact[c]
H	0.030	0.12	
F	0.064	0.138	0.130
C	0.077	0.16	
N	0.070	0.160	0.160
O	0.066	0.154	0.154
Cl	0.099	0.178	0.158
Si	0.117		
P	0.110	0.19	
S	0.104	0.203	0.160
Br	0.114	0.184	0.154
I	0.133	0.213	0.176
Se		0.215	0.170
"Radius" of methyl group		0.20	
Half-thickness of aromatic molecules		0.170	

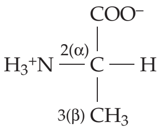
[a] From Pauling, L. (1960) *The Nature of the Chemical Bond*, 3rd ed., Cornell Univ. Press, Ithaca, New York (pp. 224–227 and 260).

[b] Covalent radii for two atoms can be summed to give the interatomic distance. The van der Waals radii determine how closely molecules can pack. The closest observed contacts between atoms in macromolecules are approximately 0.02 nm less than the sum of the van der Waals radii. From Sasisekharian, V., Lakshminarayanan, A. V., and Ramachandran, G. N. (1967) in *Conformation of Biopolymers*, Vol. 2 (Ramachandran, G. N., ed), Academic Press, New York, (p. 641)

[c] Nyburg, S. C and Faerman, C. H. (1985) *Acta Crystal*. **B41**, 274–279 Shapes of many atomic surfaces are elliptical. The major radius ***a*** applies to sideways contacts and the minor radius ***b*** to "polar" contacts along a covalent bond axis. Distances are for atoms singly bonded to C and may differ slightly if bonds are to other atoms.

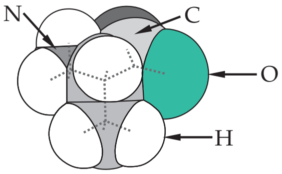
[d] For distances in Å multiply by 10.

Whenever four different groups are bonded to a central carbon atom, the molecule is asymmetric and the four groups can be arranged in two different **configurations**. Consider alanine, one of the alpha (α)-amino acids from which proteins are built.

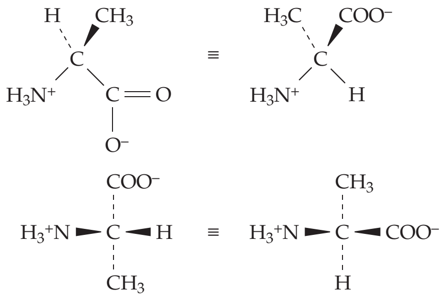


$$\begin{array}{c} \mathrm{COO^-} \\ | \\ \mathrm{H_3{}^+N} \overset{2(\alpha)}{—} \mathrm{C} — \mathrm{H} \\ | \\ 3(\beta)\ \mathrm{CH_3} \end{array}$$

α–Alanine
According to the Fisher projection formula it is L-alanine.



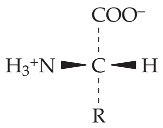
It is called an α-amino acid because the amino group is attached to the α (or number 2) carbon atom. To indicate its three-dimensional structure on a flat piece of paper, the bonds that project out of the plane of the paper and up toward the reader are often drawn as elongated triangles, while bonds that lie behind the plane of the paper are shown as dashed lines. The isomer of alanine having the configuration about the α-carbon atom shown in the following structural formulas is called *S*-alanine or L-alanine. The isomer which is a mirror image of *S*-alanine is *R*-alanine or D-alanine. Pairs of *R* and *S* compounds (see Section B for definitions) are known as **enantiomorphic** forms or **enantiomers**.



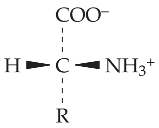
S-Alanine (L-Alanine)

Notice that in the foregoing drawings, the carboxyl group ($—\mathrm{C}\genfrac{}{}{0pt}{}{\mathrm{O}}{\mathrm{OH}}$), abbreviated –COOH, is shown as having lost a proton to form the carboxylate ion $–\mathrm{COO^-}$. Likewise, the amino group ($–\mathrm{NH_2}$) has gained a proton to form the $–\mathrm{NH_3{}^+}$ ion. The resulting **dipolar ionic** or **zwitterionic** structure is the one that actually exists for amino acids both in solution and in crystals.

The D- and L- families of amino acids. The amino acids of which proteins are composed are related to L-alanine but have various side chains (R groups) in place of the methyl group of alanine. In the preceding section the structure of L-alanine was given in four different ways. To recognize them all as the same structure, we can turn them in space to an orientation in which the carboxyl group is at the top, the side chain ($-CH_3$) is down, and both project behind the paper. The amino group and hydrogen atom will then project upward from the paper at the sides as shown below. According to a convention introduced at the beginning of this century by Emil Fischer, *an amino acid is L if, when oriented in this manner, the amino group lies to the left and D if it lies to the right.*



L-Amino acid



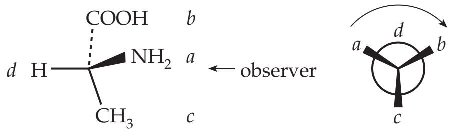
D-Amino acid

Fischer further proposed that the amino acid in this orientation could be projected onto the paper and drawn with ordinary lines for all the bonds. This gives the previously shown **Fischer projection formula** of L-alanine.

Although the D and L system of designating configuration is old it is still widely used. Remember that D and L refer to the absolute configurations about a selected reference atom in the molecule; for an amino acid this is the number 2 or α-carbon. A quantity that is related to the asymmetry of molecules is the experimentally measurable **optical rotation** (Chapter 23). The sign of the optical rotation (+ or –) is sometimes given together with the name of a compound, e.g., D(+)-glucose. The older designations *d* (dextro) and *l* (levo) indicated + and –, respectively. However, compounds with the D configuration may have either + or – optical rotation.

In older literature optical isomerism of the type represented by D and L pairs was usually discussed in terms of "asymmetric carbon atoms" or "asymmetric centers." Now the terms **chiral** (pronounced *ki-ral*) **molecules**, chiral centers, and **chirality** (Greek: "handedness") are preferred.

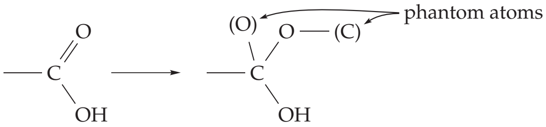
The RS notation for configuration. This notation, devised by Cahn, Ingold, and Prelog, provides an unambiguous way of specifying configuration at any chiral center.[4,5] It is especially useful for classes of compounds for which no well-established DL system is available. The groups or atoms surrounding the central carbon atom, or other central atom, are ranked according to a **priority sequence**. The priority of a group is determined by a number of sequence rules, the first of which is (1) *Higher atomic number precedes lower*. In the following illustration, the priorities of the groups in D-alanine are indicated by the letters *a* > *b* > *c* > *d*. The highest priority (*a*) is assigned to the NH_2 groups which contain nitrogen bonded to the central atom. To establish the configuration, the observer views the molecule down the axis connecting the central atom to the group having the lowest priority, i.e., to group *d*. Viewed in this way, the sequence of groups *a*, *b*, and *c* can either be that of a right-handed turn (clockwise) as shown in the drawing or that of a left-handed turn (counterclockwise).



R-Alanine (D-Alanine)

The view down the axis and toward the group of lowest priority (d), which lies behind the page. The right-handed turn indicates the configuration R (rectus = right); the opposite configuration is S (sinister = left).

To establish the priority sequence of groups first look at the atoms that are bonded directly to the central atom, arranging them in order of decreasing atomic number. Then if necessary, move outward to the next set of atoms, again comparing atomic numbers. In the case of alanine, groups *b* and *c* must be ordered in this way because they both contain carbon directly bonded to the central atom. When double bonds are present at one of the atoms being examined, e.g., the carboxyl group in alanine, imagine that **phantom atoms** that replicate the real ones are present at the ends of the bonds:



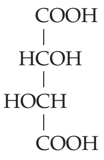
These phantom atoms fill out the valences of the atoms involved in the multiple bonds and are considered to have zero atomic number and zero mass. They are not considered in establishing priorities.

If the first rule and the expansion of multiple bonds are not sufficient to establish the priority, use these additional rules: (2) *Higher atomic mass precedes lower.* (3) When a double bond is present *Z precedes E* (see Geometrical isomers). For ring systems a *cis* arrangement of the highest priority substituents precedes *trans*. (4) When a pair of chiral centers is present *R,R* or *S,S* precedes *R,S* or *S,R*. (5) *An R chiral center precedes S.* For further details see Eliel *et al.*[5] and Bentley.[6] The following groups are ordered in terms of *decreasing priority*[6]: SH > OR > OH > $NH-COCH_3$ > NH_2 > COOR > COOH > CHO > CH_2OH > C_6H_5 > CH_3 > ^3H > ^2H > H.

Although the *RS* system is unambiguous, closely related compounds that belong to the same configurational family in the DL system may have opposite configurations in the *RS* system. Thus, L-cysteine (side chain $-CH_2SH$) has the *R* configuration. This is one of the reasons that the DL system is still used for amino acids and sugars.

Diastereoisomers. Whereas compounds with one chiral center exist as an enantiomorphic pair, molecules with two or more chiral centers also exist as diastereoisomers (diastereomers). These are pairs of isomers with an opposite configuration at one or more of the chiral centers, but which are not complete mirror images of each other. An example is L-threonine which has the 2*S*, 3*R* configuration. The diastereoisomer with the 2*S*, 3*S* configuration is known as L-*allo*-threonine. L-isoleucine, whose side chain is $-CH(CH_3)\ CH_2CH_3$, has the 2*S*, 3*R* configuration. It can be called 2(*S*)-amino-3(*R*)-methyl-valeric acid but the simpler name L-isoleucine implies the correct configuration at both chiral centers.

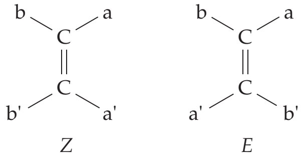
Sometimes the subscript s or g is added to a D or L prefix to indicate whether the chirality of a compound is being related to that of serine, the traditional configurational standard for amino acids, or to that of glyceraldehyde. In the latter case the sugar convention (Chapter 4) is followed. In this convention the configurations of the chiral centers furthest from C1 are compared. Ordinary threonine is L_s- or D_g-threonine. The configuration of dextrorotatory (+)-tartaric acid can be described as 2*R*, 3*R*, or as D_s, or as L_g.


$$\begin{array}{c} COOH \\ | \\ HCOH \\ | \\ HOCH \\ | \\ COOH \end{array}$$

2*R*, 3*R*-tartaric acid

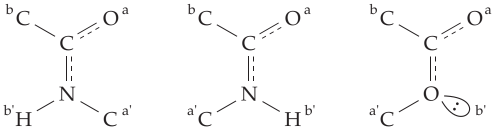
Biochemical reactions are usually stereospecific and a given enzyme will catalyze reactions of molecules of only a single configuration. A related fact is that proteins ordinarily consist entirely of amino acids of the L series.

Geometrical isomers. The *RS* system also gives an unambiguous designation of geometrical isomers containing a double bond.[5,7] At each end of the bond, select the group of highest priority. If these two groups lie on the same side of the double bond the configuration is ***Z*** (from the German **zusammen**, "together"); if on opposite sides ***E*** (**entgegen**, "opposite").



a and a' higher priority than b or b'

Configurations of amide or ester linkages may also be specified in this manner. This is possible because the C–N bond of an amide has partial double-bond character, as to a lesser extent does the C–O bond to the bridge oxygen in an ester. In this case, assign the lowest priority to the unshared electron pair on the ester bridge oxygen.



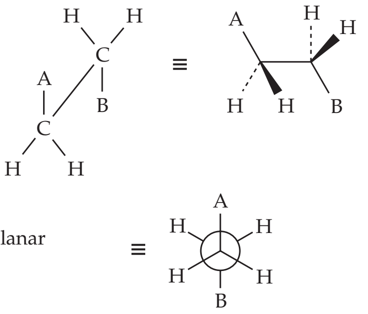
Z or *trans* *E* or *cis*

An amide of the *Z* configuration is ordinarily referred to as *trans* in protein chemistry because the main chain atoms are trans.

5. Conformations: The Shapes That Molecules Can Assume

As important to biochemists as configurations, the stable arrangements of bonded atoms, are **conformations**, the various orientations of groups that are caused by rotation about single bonds.[5,8] In many molecules such rotation occurs rapidly and freely at ordinary temperatures. We can think of a $-CH_3$ group as a kind of erratic windmill, turning in one direction, then another. However, even the simplest molecules have *preferred conformations*, and in more complex structures rotation is usually very restricted.

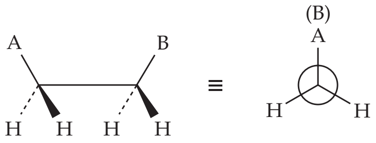
Consider a molecule in which groups A and B are joined by two CH_2 (methylene) groups. If A and B are pulled as far apart as possible, the molecule is in its fully extended **anti** or **staggered** conformation:



The antiperiplanar conformation

View down the axis joining the C atoms (Newton projection)

Groups A and B are said to be **antiperiplanar** (*ap*) in this conformation. Not only are A and B as far apart as possible but also all of the hydrogen atoms are at their maximum distances one from the other. This can be seen by viewing the molecule down the axis joining the carbon atoms (Newman projection). Rotation of the second carbon atom 180° around the single bond yields the **eclipsed** conformation in which groups A and B are synperiplanar.

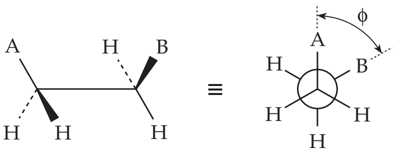


Eclipsed or synperiplanar conformation

If A and B are large bulky groups they will bump together, attainment of the eclipsed conformation will be almost impossible, and rotation will be severely restricted. Even if A and B are hydrogen atoms (ethane), there will be a rotational barrier in the eclipsed conformation which amounts to ~12 kJ (3 kcal) per mole because of the crowding of the hydrogen atoms as they pass each other.[5,9] This can be appreciated readily by examination of space-filling molecular models.

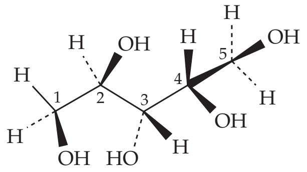
If groups A and B are methyl groups (butane), the steric hindrance between A and B leads to a rotational barrier of ~25 kJ (6 kcal) per mole. The consequence of this simple fact is that in fatty acids and related substances and in polyethylene the chains of CH_2 groups tend to assume fully extended zigzag conformations.

In addition to this extended conformation there are two **gauche** (**skewed** or synclinal) conformations which are only slightly less stable than the staggered conformation and in which A and B interfere only if they are very bulky. In one of the two gauche conformations B lies to the right of A and in the other to the left of A when viewed down the axis.



One of the two gauche or synclinal conformations

These two conformations are related to right-handed and left-handed screws, respectively. The threads on an ordinary right-handed household screw, when viewed down the axis from either end, move backward from left to right in the same fashion as do the groups A and B in the illustration. The angle ϕ is the **torsion angle** and is positive for right-handed conformations. Gauche conformations are important in many biological molecules; for example, the sugar alcohol **ribitol** stacks in crystals in a “sickle” conformation,[6] in which the chain starts out (at the left) in the zigzag arrangement but shifts to a gauche conformation around the fourth carbon atom, thereby minimizing steric interference between the OH groups on the second and fourth carbons.



Ribitol in sickle conformation

The complete series of possible conformations is shown in Fig. 2-2.

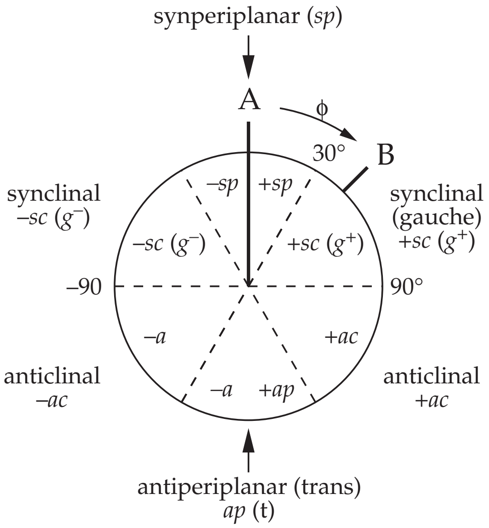
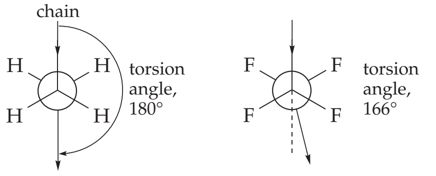


Figure 2-2 Description of conformations about a single bond in the terminology of Klyne and Prelog[10,11] using the Newman projection. Group A is on the front atom at the top: the conformation is given for each possible position of group B on the other atom.

In the chain of methylene units, the hydrogen atoms on alternate carbon atoms of the fully extended chain barely touch (Fig. 2-1) but larger atoms cannot be accommodated. Thus, when fluorine atoms of van der Waals radius 0.135 nm replace the hydrogen atoms of radius 0.12 nm, a fully extended chain is no longer possible. For this reason the torsion angle in polyfluoroethylene is changed from the 180° of polyethylene to 166°, enough to relieve the congestion but not enough to cause severe eclipsing of the fluorines on adjacent carbons. The resulting **helical structure** is reminiscent of those occurring in proteins and other biopolymers. We see that helix formation can be a natural result of steric hindrance between groups of atoms.

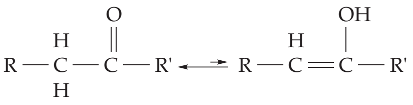


Polyethylene Polyfluoroethylene (Teflon)

Conformations of ring-containing molecules are dealt with in Chapter 4.

6. Tautomerism and Resonance

Many simple organic compounds exist as mixtures of two or more rapidly interconvertible isomers or **tautomeric forms**. Tautomers can sometimes be separated one from the other at low temperatures where the rate of interconversion is low. The classic example is the **oxo-enol** (or keto-enol) equilibrium (Eq. 2-1).

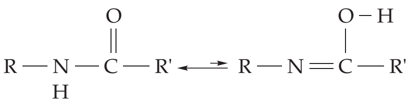


$$R-\overset{H}{\underset{H}{C}}-\overset{O}{\overset{\|}{C}}-R' \rightleftharpoons R-\overset{H}{C}=\overset{OH}{\overset{|}{C}}-R'$$

Oxo form Enol (2-1)

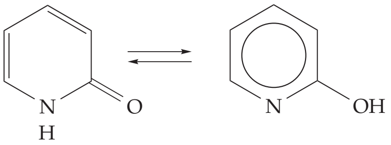
Although usually less stable than the oxo (keto) form, the enol is present in a small amount. It is formed readily from the oxo tautomer by virtue of the fact that hydrogen atoms attached to carbon atoms that are immediately adjacent to carbonyl (C=O) groups are remarkably acidic. Easy dissociation of a proton is a prerequisite for tautomerism. Since most hydrogen atoms bound to carbon atoms do not dissociate readily, tautomerism is unusual unless a carbonyl or other "activating group" is present.

Since protons bound to oxygen and nitrogen atoms usually *do* dissociate readily, tautomerism also exists in amides and in ring systems containing O and N (Eqs. 2-2 to 2-5).

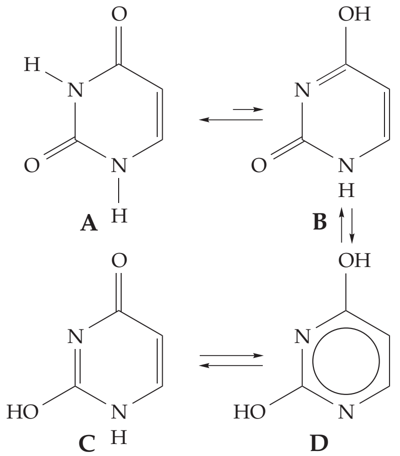


$$R-\underset{H}{N}-\overset{O}{\overset{\|}{C}}-R' \rightleftharpoons R-N=\overset{O-H}{\overset{|}{C}}-R'$$

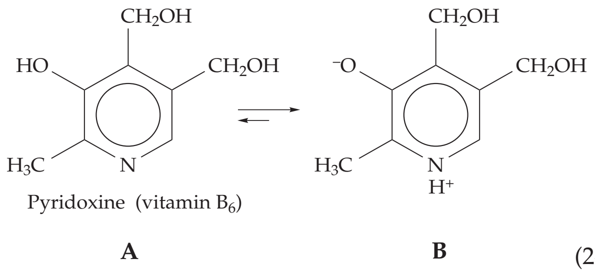
A **B** (2-2)



2-Pyridone 2-Hydroxypyridine (2-3)



(2-4)



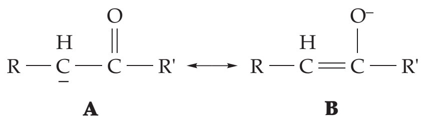
Pyridoxine (vitamin B_6)

(2-5)

The tautomerism in Eq. 2-2 is the counterpart of that in the oxo-enol transformation. However, the equilibrium constant for aqueous conditions favors form A very strongly. 2-Pyridone is tautomerized to 2-hydroxypyridine (Eq. 2-3) to a greater extent. Pyrimidines (Eq. 2-4) and purines can form a variety of tautomers. The existence of form D of Eq. 2-4 is the basis for referring to uracil as dihydroxypyrimidine. However, the di-oxo tautomer A predominates. Pyridoxine (vitamin B_6) exists in water largely as the dipolar ionic tautomer B (Eq. 2-5) but in methanol as the uncharged tautomer A. In a pair of tautomers, a hydrogen atom always moves from one position to another and the lengths and bond character of these bonds also change.

The equilibrium constant for a tautomeric interconversion is simply the ratio of the mole fractions of the two forms; for example, the ratio of enol to oxo forms of acetone[12] in water at 25°C is 6.0×10^{-9}, while that for isobutyraldehyde is 1.3×10^{-4}. The ratio of 2-hydroxypyridine to 2-pyridone is about 10^{-3} in water but increases to 0.6 in a hydrocarbon solvent and to 2.5 in the vapor phase.[13,14] The ratio of dipolar ion to uncharged pyridoxine (Eq. 2-5) is ~4 at 25°C in water.[15] The ratios of tautomers B, C, and D to the tautomer A of uracil (Eq. 2-4) are small, but it is difficult to measure them quantitatively.[16] These tautomeric ratios are defined for given overall states of protonation (see Eq. 6-82). The constants are independent of pH but will change if the overall state of protonation of the molecule is changed. They may also be altered by changes in temperature or solvent or by binding to a protein or other molecule.

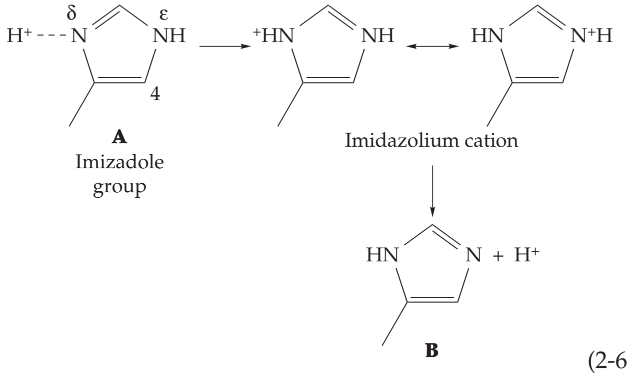
It is important to distinguish tautomerism from **resonance**, a term used to indicate that the properties of a given molecule cannot be represented by a *single* valence structure but can be represented as a hybrid of two or more structures in which all the nuclei remain in the same places. Only bonding electrons move to convert one resonance form into another. Examples are the **enolate anion**, which can be thought of as a hybrid of structures A and B, and the amide linkage, which can be represented by a similar pair of resonance forms.



Enolate anion

A double-headed arrow is often used to indicate that two structures drawn are resonance structures rather than tautomers or other separable isomers.

Although they are distinctly different, tautomerism and resonance are related. Thus, the acidity of carbon-bound hydrogens in ketones, which allows formation of enol tautomers, results from the fact that the enolate anion produced by dissociation of one of these hydrogens is stabilized by resonance. Similarly, tautomerism in the imidazole group of the amino acid histidine is related to resonance in the imidazolium cation. Because of this resonance, if a proton approaches structure A of Eq. 2-6 and becomes attached to the left-hand nitrogen atom (N^{δ}), the positive charge in the resulting intermediate is distributed over both nitrogen atoms. This makes the proton on N^{ε} acidic, permitting it to dissociate to tautomer B.



(2-6)

Note: The nitrogen atom designated N^{δ} (or ND1)[11] in Eq. 2-6 may also be called N1 or N^{π} (*pros*-N). Likewise, N^{ε} (NE2) may be designated N3 or N^{τ} (*tele*- N). Since N^{σ} has sometimes also been called N3, it is best not to use the numerical designations for the nitrogen atoms.

The tautomeric ratio of B to A for histidine in water (Eq. 2-6) has been estimated, using ^{15}N- and ^{13}C-NMR, as 5.0 when the α-amino group is protonated and as 2.5 when at high pH it is unprotonated.[17] This tautomerism of the imidazole group is probably important to the function of many enzymes and other proteins; for example, if N^{ε} of structure A (Eq. 2-6) is embedded in a protein, a proton approaching from the outside can induce the tautomerism shown with the release of a proton in the interior of the protein, perhaps at the active site of an enzyme. The form protonated on N^{δ} (B of Eq. 2-6), which is the minor form in solution, predominates in some positions within proteins.[18]

B. Forces between Molecules and between Chemical Groups

The structure of living cells depends very much on the covalent bonds within individual molecules and on covalent crosslinks that sometimes form between molecules. However, weaker forces acting between molecules and between different parts of the same molecule are responsible for many of the most important properties of biochemical substances. These are described as **van der Waals forces**, **electrostatic forces**, **hydrogen bonds**, and **hydrophobic interactions**. In the discussion that follows the thermodynamic quantities ΔH, ΔS, and ΔG will be used. If necessary, please see Chapter 6 for definitions and a brief review.

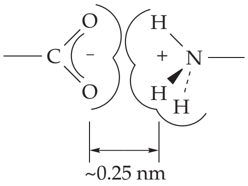
1. Van der Waals Forces

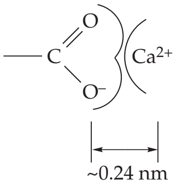
All atoms have a weak tendency to stick together, and because of this even helium liquifies at a low enough temperature. This is a result of the van der Waals or "London dispersion forces" which act strongly only at a very short distance. These forces arise from electrostatic attraction between the positively charged nucleus of one atom and the negatively charged electrons of the other.[19–21] Because nuclei are screened by the electron clouds surrounding them, the force is weak. The energy (enthalpy) of binding of one methylene ($-CH_2-$) unit into a monomolecular layer of a fatty acid is about $-\Delta H^{\circ} = 1.7$ kJ/mol.[22] Although this is a small quantity, when summed over the 16 or more carbon atoms of a typical fatty acid the binding energy is substantial. When a methylene group is completely surrounded in a crystalline hydrocarbon, its van der Waals energy, as estimated from the heat of sublimation, is 8.4 kJ/mol; that of H_2O in liquid water at the melting point of ice is 15 kJ/mol.[22]

While van der Waals forces between individual atoms act over very short distances, they can be felt at surprisingly great distances when exerted by large molecules or molecular aggregates.[23] Forces between very smooth surfaces have been measured experimentally at distances as great as 10 nm and even to 300 nm.[23a] However, these "long-range van der Waals forces" probably depend upon layers of oriented water molecules on the plates[23] (see also Section 5).

2. Attraction between Charged Groups (Salt Linkages)

Fixed positive and negative charges attract each other strongly. Consider a carboxylate ion in contact with $-NH_3^+$ or with an ion of calcium:





From the van der Waals radii of Table 2-1 and the ionic crystal radius of Ca^{2+} of 0.10 nm, we can estimate an approximate distance between the centers of positive and negative charge of 0.25 nm in both cases. It is of interest to apply Coulomb's law to compute the force F between two charged particles which are almost in contact. Let us choose a distance of 0.40 nm (4.0 Å) and apply Eq. 2-7.

$$F = 8.9875 \times 10^9 \times \frac{qq'}{r^2\varepsilon} \text{ newtons} \tag{2-7}$$

In this equation r is the distance in meters, q and q' are the charges in coulombs (one electronic charge = 1.6021×10^{-19} coulombs), ε is the dielectric constant, and F is the force in newtons (N). The force per mole is NF where N is Avogadro's number.

An uncertainty in this kind of calculation is in the dielectric constant ε, which is 1.0 for a vacuum, about 2 for hydrocarbons, and 78.5 for water at 25°C. If ε is taken as 2, the force for $r = 0.40$ nm is 4.3×10^{14} N/mol. The force would be twice as great for the $Ca^{2+} - COO^-$ case. To move two single charges further apart by just 0.01 nm would require 4.3 kJ/mol, a substantial amount of energy. However, if the dielectric constant were that of water, this would be reduced almost 40-fold and the electrostatic force would not be highly significant in binding. It is extremely difficult to assign a dielectric constant for use in the interior of proteins.[23b] For charges spaced far apart within proteins the effective dielectric constant is usually as high as 30–60.[24] For closely spaced charges in hydrophobic niches it may be as low as 2–4.[25–27]

A calculation that is often made is the work required to remove completely two charges from a given distance apart (e.g., 0.40 nm) to an infinite distance (Eq. 2-8).

$$\begin{aligned} W\ (\text{kJ mol}^{-1}) &= 8.9875 \times 10^6 \times \frac{qq'}{r\varepsilon} N \\ &= \frac{138.9}{\varepsilon r \text{ (in nm)}} \end{aligned} \tag{2-8}$$

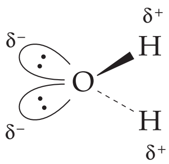
If $\varepsilon = 2$, this amounts to 174 kJ/mol for single charges at a distance of 0.40 nm; 69 kJ/mol at 1 nm; and only 6.9 kJ/mol at 10 nm, the distance across a cell membrane. We see that very large forces exist between closely spaced charges.

Electrostatic forces are of great significance in interactions between molecules and in the induction of changes in conformations of molecules. For example, attraction between $-COO^-$ and $-NH_3^+$ groups occurs in interactions between proteins. Calcium ions often interact with carboxylate groups, the doubly charged Ca^{2+} bridging between two carboxylate or other polar groups. This occurs in carbohydrates such as agarose, converting solutions of these molecules into rigid gels (Chapter 4).

Individual macromolecules as well as cell surfaces usually carry a net negative charge at neutral pH. This causes the surfaces to repel each other. However, at a certain distance of separation the van der Waals attractive forces will balance the electrostatic repulsion.[21] Protruding hydrophobic groups may then interact and may "tether" bacteria or other particles at a fixed distance, often ~5 nm, from a cell surface.[28]

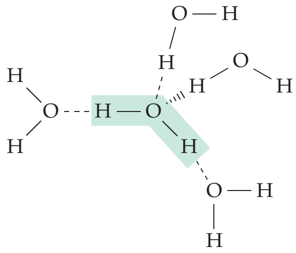
3. Hydrogen Bonds

One of the most important weak interactions between biologically important molecules is the hydrogen bond (H-bond).[29,30,30a] These "bonds" are the result of electrostatic attraction caused by the uneven distribution of electrons within covalent bonds. For example, the bonding electron pairs of the H–O bonds of water molecules are attracted more tightly to the oxygen atoms than to the hydrogen atoms. A small net positive charge is left on the hydrogen and a small net negative charge on the oxygen. Such **polarization** of the water molecules can be indicated in the following way:



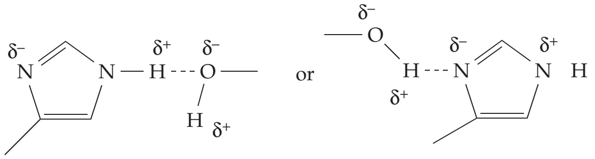
Here the δ+ and δ– indicate a fraction of a full charge present on the hydrogen atoms and on the nonbonded electron pairs of the oxygen atom, respectively. Molecules such as H_2O, with strongly polarized bonds, are referred to as **polar molecules** and functional groups with such bonds as **polar groups**. They are to be contrasted with such nonpolar groups as the $-CH_3$ group in which the electrons in the bonds are nearly equally shared by carbon and hydrogen.

A hydrogen bond is formed when the positively charged end of one of the **dipoles** (polarized bonds) is attracted to the negative end of another dipole. Water molecules tend to hydrogen bond strongly one to another; each oxygen atom can be hydrogen-bonded to two other molecules and each hydrogen to another water molecule. Thus, every water molecule can have up to four hydrogen-bonded neighbors.



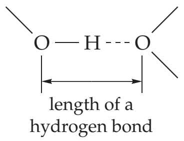
A water molecule hydrogen bonded to four other water molecules; note the tetrahedral arrangement of bonds around the central oxygen.

Many groups in proteins, carbohydrates, and nucleic acids form hydrogen bonds to one another and to surrounding water molecules. For example, an imidazole group of a protein can bond to an OH group of an amino acid side chain or of water in the following ways:



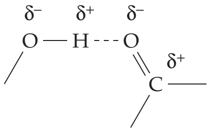
Remember that hydrogen bonds are always formed between pairs of groups, with one of them, often C=O or C=N-, containing the negative end of a dipole and the other providing the proton. The proton acceptor group, often OH or NH and occasionally SH, and even CH in certain structures,[31,31a,31b] donates an unshared pair of electrons. Dashed arrows are sometimes drawn from the hydrogen atom to the electron donor atom to indicate the direction of a hydrogen bond. Do not confuse these arrows with the curved arrows that indicate flow of electrons in organic reactions.

The strength of hydrogen bonds, as measured by the bond energy, varies over the range 10–40 kJ/mol. The stronger the hydrogen bond the shorter its length. Because hydrogen atoms can usually not be seen in X-ray structures of macromolecules, the lengths of hydrogen bonds are often measured between the surrounding heavy atoms:



A typical —OH---O hydrogen bond will have a length of about 0.31 nm; a very strong hydrogen bond may be less than 0.28 nm in length, while weak hydrogen bonds will approach 0.36 nm, which is the sum of the van der Waals contact distances plus the O–H bond length. Beyond this distance a hydrogen bond cannot be distinguished easily from a van der Waals contact.

Hydrogen bonds are strongest when the hydrogen atom and the two heavy atoms to which it is bonded are in a straight line. For this reason hydrogen bonds tend to be linear. However, *the dipoles forming the hydrogen bond do not have to be colinear for strong hydrogen bonding:*[32,32a] There is some preference for hydrogen bonding to occur in the direction of an unshared electron pair on the oxygen or nitrogen atom.[33–35]



A linear O–H- - -O hydrogen bond with dipoles at an angle one to another.

Both ammonia, NH_3, and the $-NH_2$ groups of proteins are good electron donors for hydrogen bond formation. However, the hydrogen atoms of uncharged $-NH_2$ groups tend to be poor proton donors for H-bonds.[36] Do hydrogen bonds have some covalent character? The answer is controversial.[36a,36b,36c]

Hydrogen bonding is important both to the internal structure of biological macromolecules and in interactions between molecules. Hydrogen bonding often provides the specificity necessary to bring surfaces together in a complementary way. Thus, the location of hydrogen-bond forming groups in surfaces between molecules is important in ensuring an exact alignment of the surfaces.[37] The hydrogen bonds do not always have to be strong. For example, Fersht and coworkers, who compared a variety of mutants of an enzyme of known three-dimensional structure, found that deletion of a side chain that formed a good hydrogen bond to the substrate weakened the binding energy by only 2–6 kJ/mol. However, loss of a hydrogen bond to a charged group in the substrate caused a loss of 15–20 kJ/mol of binding energy.[37] Study of mutant proteins created by genetic engineering is now an important tool for experimentally investigating the biological roles of hydrogen bonding.[37–39]

4. The Structure and Properties of Water

Water is the major constituent of cells and a remarkable solvent whose chemical and physical properties affect almost every aspect of life. Many of these properties are a direct reflection of the fact that most water molecules are in contact with their neighbors entirely through hydrogen bonds.[40–48] Water is the only known substance for which this is true.

In ordinary ice all of the water molecules are connected by hydrogen bonds, six molecules forming a hexagonal ring resembling that of cyclohexane. The structure is extended in all directions by the formation of additional hydrogen bonds to adjacent molecules (Fig. 2-3). As can be seen in this drawing, the molecules in ice assume various orientations in the hexagonal array, and frequently rotate to form their hydrogen bonds in different ways. This randomness remains as the temperature is lowered, and ice is one of few substances with a residual entropy at absolute zero.[49,50] Ice is unusual also in that the molecules do not assume closest packing in the crystal but form an open structure. The hole through the middle of the hexagon and on through the hexagons lying below it is ~0.06 nm in diameter.

The short hydrogen-bond length (averaging 0.276 nm) in ice indicates of strong bonding. The heat of sublimation ($\Delta H°$) of ice is –48.6 kJ/mol. If the van der Waals dispersion energy of –15 kJ/mol is subtracted from this, the difference of –34 kJ/mol can be attributed entirely to the hydrogen bonds—two for each molecule. Their average energy is 17 kJ/mol apiece. However, some of the hydrogen bonds are stronger and others weaker than the average.[51]

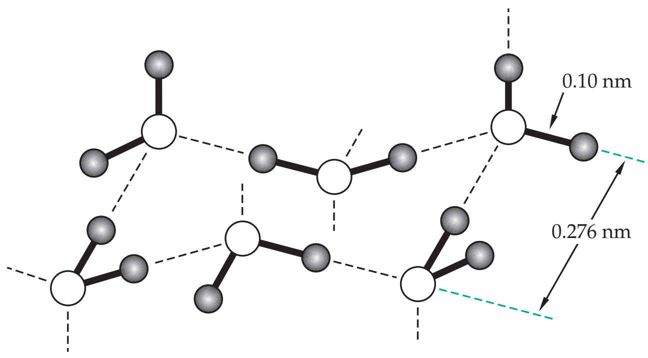
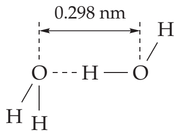
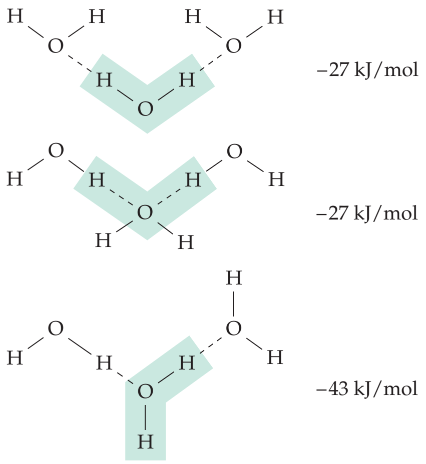


Figure 2-3 Six water molecules in the lattice of an ice crystal. The hydrogen bonds, which connect protons with electron pairs of adjacent molecules, are shown as dashed lines.

In a gaseous water dimer the hydrogen bond is linear, a fact that suggests some covalent character.[52]



Its length is distinctly greater than that in ice. This is one of a number of pieces of evidence suggesting cooperativity in formation of chains of hydrogen bonds.[40,53,54] Consider the following three trimers for which theoretical calculations have predicted the indicated hydrogen bond energies.[40] In the first case the central water molecule *donates* two protons for hydrogen-bond formation; in the second it *accepts* the protons. In the third case it is both an electron acceptor and a donor. The OH dipoles are oriented "head to tail" and the hydrogen bonds are stronger than in the other cases. Long chains of similarly oriented hydrogen bonds exist in ice and this may account for the short hydrogen bond lengths. Closed rings of hydrogen bonds oriented to give a maximum cooperative effect also exist in liquid water clusters[55] and within proteins, carbohydrates, and nucleic acids.[53,54]



The nature of liquid water is still incompletely understood,[46-48,56] but we know that water contains ice-like clusters of molecules that are continually breaking up and reforming in what has been called a "flickering cluster" structure. Judging by the infrared spectrum of water, about 10% of the hydrogen bonds are broken when ice melts.[41] A similar conclusion can be drawn from the fact that the heat of melting of ice is –5.9 kJ/mol. It has been estimated that at 0°C the average cluster contains about 500 water molecules.[41] At 50°C there are over 100 and at the boiling point about 40. Although most molecules in liquid water are present in these clusters, the hydrogen bonds are rapidly broken and reformed in new ways, with the average lifetime of a given hydrogen bond being ~10^{-12} s.

5. Hydration of Polar Molecules and Ions

Water molecules are able to hydrogen bond not only to each other but also to polar groups of dissolved compounds. Thus, every group that is capable of forming a hydrogen bond to another organic group is also able to form hydrogen bonds of a somewhat similar strength with water. For this reason, hydrogen bonding is usually not a significant force in holding small molecules together in aqueous solutions. Polar molecules that stick together through hydrogen bonding when dissolved in a nonpolar solvent often do not associate in water. How then can biochemists assert that hydrogen bonding is so important in biochemistry? Part of the answer is that proteins and nucleic acids can be either properly folded with hydrogen bonds formed internally or denatured with hydrogen bonds from those same groups to water. The Gibbs energy change between these two states is small.

Every ion in an aqueous solution is surrounded by a shell of oriented water molecules held by the attraction of the water dipoles to the charged ion. The hydration of ions has a strong influence on all aspects of electrostatic interactions and plays a dominant role in determining such matters as the strength of acids and bases, the Gibbs energy of hydrolysis of ATP, and the strength of bonding of metal ions to negatively charged groups. For example, the previously considered interaction between carboxylate and calcium ions would be much weaker if both ions retained their hydration shells.

Consider the following example. $\Delta G°$ for dissociation of acetic acid in water is +27.2 kJ/mol (Table 6-5). The enthalpy change $\Delta H°$ for this process is almost zero (–0.1 kJ/mol) and $\Delta S°$ is consequently $-91.6\ \mathrm{J\ K^{-1}}$. This large entropy decrease reflects the increased amount of water that is immobilized in the hydration spheres of the H^+ and acetate$^-$ ions formed in the dissociation reaction. In contrast, dissociation of NH_4^+ to NH_3 and H^+ converts one positive ion to another. $\Delta H°$ is large (+52.5 kJ/mol) but the entropy change $\Delta S°$ is small ($-2.0\ \mathrm{J\ K^{-1}}$, Table 6-5).

Although effects of hydration are important in almost all biochemical equilibria, they are difficult to assess quantitatively. It is hard to know how many molecules of water are freed or immobilized in a given reaction. Charged groups in proteins are often hydrated. However, if they are buried in the interior of the protein, they may be solvated by polarizable protein side chain groups such as –OH or by backbone or side chain amide groups.[57,58]

6. Hydrophobic Interactions

Fats, hydrocarbons, and other materials whose molecules consist largely of nonpolar groups have a low solubility in water and a high solubility in nonpolar solvents. Similarly, the long alkyl groups of fatty acid esters aggregate within membranes and nonpolar side chains of proteins are often packed together in the centers of protein molecules. Because it is as if the nonpolar groups "fear" water, this is known as the **hydrophobic** effect.[59–70] The terms hydrophobic forces, hydrophobic interactions, and hydrophobic bonding have also been used. However, the latter term can be misleading because the hydrophobic effect arises not out of any special attraction between nonpolar groups but primarily from the strong internal cohesion of the hydrogen-bonded water structure.

How strongly do nonpolar groups "attract" each other in water? A partial answer can be obtained by measuring the standard Gibbs energy ΔG of transfer of a hydrocarbon molecule from a dilute aqueous solution into a dilute solution in another hydrocarbon. By studying a series of alkanes, Abraham[62] calculated the Gibbs energy change per CH_2 unit (Eq. 2-9) as: $\Delta G° = -3.8$ kJ/mol.

$$-CH_2-(\text{water}) \longrightarrow -CH_2-(\text{hexane}) \qquad \Delta G° = -3.8\ \text{kJ/mol} \tag{2-9}$$

This equation is a quantitative statement of the fact that the CH_2 group prefers to be in a nonpolar environment than to be surrounded by water. A similar Gibbs energy change would be expected to accompany the bringing together of a methylene unit from a small molecule and a hydrophobic surface on a protein molecule. However, in the latter case the accompanying decrease in entropy would make $\Delta G°$ less negative.

What causes the decrease in Gibbs energy when nonpolar groups associate in water? Jencks[60] suggested that we think of the transfer of a nonpolar molecule from a nonpolar solvent into water in two steps: (1) Create a cavity in the water of about the right size to accommodate the molecule. Since many hydrogen bonds will be broken, the Gibbs energy of cavity formation will be high. It will be principally an enthalpy (ΔH) effect. (2) Allow the water molecules in the solvent to make changes in their orientations to accommodate the nonpolar molecule that has been placed in the cavity. The water molecules can move to give good van der Waals contacts and also reorient themselves to give the maximum number of hydrogen bonds. Since hydrogen bonds can be formed in many different ways in water, there may be as many or even more hydrogen bonds after the reorientation than before. This will be true especially at low temperature where most water exists as large icelike clusters. For dissolved hydrocarbons, the enthalpy of formation of the new hydrogen bonds often almost exactly balances the enthalpy of creation of the cavity initially so that ΔH for the overall process (transfer from inert solvent into water) is small. For the opposite transfer (from water to hydrocarbon hexane; Eq. 2-9) $\Delta H°$ is usually a small positive number for aliphatic hydrocarbons and is nearly zero for aromatic hydrocarbons.

Since $\Delta G° = \Delta H° - T\Delta S°$ (see Chapter 6), it follows that the negative value of $\Delta G°$ for hydrophobic interactions must result from a positive entropy change, which may arise from the restricted mobility of water molecules that surround dissolved hydrophobic groups. When two hydrophobic groups come together to form a "hydrophobic bond," water molecules are freed from the structured region around the hydrophobic surfaces and the entropy increases. The $\Delta S°$ for Eq. 2-9 is about 12 J deg^{-1} mol^{-1}. Attempts have been made to relate this value directly to the increased number of orientations possible for a water molecule when it is freed from the structured region.[64] However, interpretation of the hydrophobic effect is complex and controversial.[65–71a]

The formation constant K_f for hydrophobic associations often increases with increasing temperature. This is in contrast to the behavior of K_f for many association reactions that involve polar molecules and for which $\Delta H°$ is often strongly negative (heat is released). An example of the latter is the protonation of ammonia in an aqueous solution (Eq. 2-10).

$$\begin{gathered} NH_3 + H^+ \rightarrow NH_4^+ \\ \Delta H° = -52.5 \text{ kJ/mol} \end{gathered} \qquad (2\text{-}10)$$

Since $R1nK_f = -\Delta G°/T = -\Delta H°/T + \Delta S°$, K_f decreases with increasing temperature if $\Delta H°$ is negative. Because for a hydrophobic interaction with a positive value of $\Delta H°$ K_f *increases* with increasing temperature, an increase in stability at higher temperatures is sometimes used as a criterion for hydrophobic bonding. However, this criterion does not always hold. For example, base stacking interactions in polynucleotides (Chapter 5), whose strength does not increase with increasing temperature, are still thought to be hydrophobic.

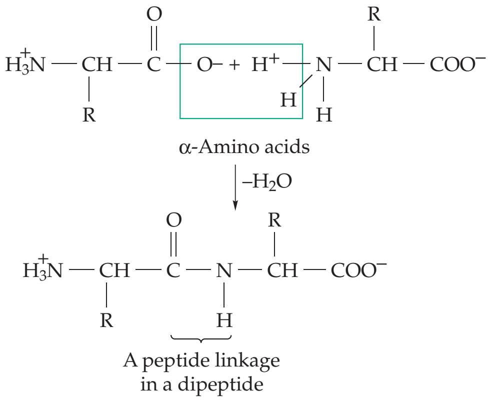
The water molecules that are in immediate contact with dissolved nonpolar groups are partially oriented. They form a cagelike structure around each hydrophobic group. When particles surrounded by such hydration layers are 1–2 nm apart, they sometimes experience either a fairly strong repulsion or an enhanced attraction caused by these hydration layers.[21,64–66,72] Direct experimental measurements have shown that these effects extend to distances of 10 nm[21,63] and can account for the previously mentioned long-range van der Waals forces.

Various efforts have been made to develop scales of **hydrophobicity** that can be used to predict the probability of finding a given amino acid side chain buried within a protein or in a surface facing water.[59,73] A new approach has been provided by the study of mutant proteins. For example, deletion of a single $-CH_2-$ group from an interior hydrophobic region of a protein was observed to decrease the stability of the protein by 4.6 kJ/mol.[74]

C. Amino Acids and Peptides

Twenty α-amino acids are the monomers from which proteins are made. These amino acids share with other biochemical monomers a property essential to their role in polymer formation: *They contain at least two different chemical groups able to react with each other to form a covalent linkage*. In the amino acids these are the protonated amino (NH_3^+) and carboxylate (COO^-) groups. The characteristic linkage in the protein polymer is the **peptide** (amide) linkage whose formation can be imagined to occur by the splitting out of water between the carboxyl of one amino acid and the amino group of another (Eq. 2-11).

This equation is not intended to imply a mechanism for peptide synthesis. The equilibrium position for this reaction in an aqueous solution favors the free amino acids rather than the peptide. Therefore, both biological and laboratory syntheses of peptides usually do not involve a simple splitting out of water. Since the dipeptide of Eq. 2-11 still contains reactive carboxyl and amino groups, other amino acid units can be joined by additional peptide linkages to form **polypeptides**. These range from short-chain **oligomers** to polymers of from ~50 to several thousand amino acid units, the proteins.[75–77]



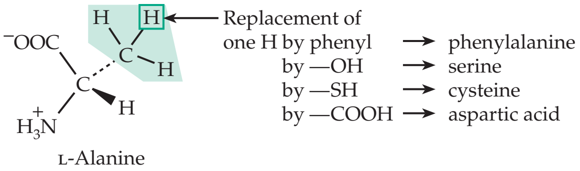
(2-11)

1. Properties of α-Amino Acids

The amino acids have in common a dipolar ionic structure and a chiral center. They are differentiated, one from another, by the structures of their **side chain groups**, designated R in the foregoing formulas. These groups are of varying size and chemical structure. The side chain groups fill much of the space in the interior of a protein molecule and also protrude from the external surfaces of the protein where they determine many of the chemical and physical properties of the molecule.

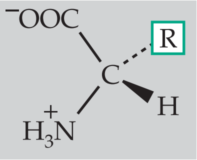
Table 2-2 shows the structures of the side chains of the amino acids commonly found in proteins. The complete structure is given for proline. Both the three-letter abbreviations and one-letter abbreviations used in describing sequences of amino acids in proteins are also given in this table. Amino acids of groups **a**–**c** of Table 2-2 plus phenylalanine and methionine are sometimes grouped together as *nonpolar*. They tend to be found in a hydrophobic environment on the inside of a protein molecule. Groups **f** and **i** contain *polar*, *charged* side chains which usually protrude into the water surrounding the protein. The rest are classified as *polar* but *noncharged*.

To get acquainted with amino acid structures, learn first those of **glycine**, **alanine**, **serine**, **aspartic acid**, and **glutamic acid**. The structures of many other amino acids can be related to that of alanine (R=CH$_3$) by replacement of a β hydrogen by another group. Metabolic interrelationships will make it easier to learn structures of the rest of the amino acids later.



Since the –COOH groups of glutamic and aspartic acids are completely dissociated to –COO$^-$ at neutral pH, it is customary in the biochemical literature to refer to these amino acids as **glutamate** and **aspartate** without reference to the nature of the cation or cations present as counter ions. Such "-ate" endings are also used for most other acids (e.g., malate, oxaloacetate,

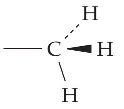
TABLE 2-2
Structure and Chemical Properies of Side Chain Groups (R) of Amino Acids



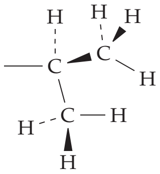
a. Glycine "side chain" = —H (Gly, G)[a]

Strictly a link in the peptide chain, glycine provides a minimum of steric hindrance to rotation and to placement of adjacent groups.

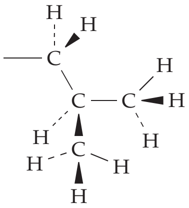
b. Amino acids with alkyl groups as side chains



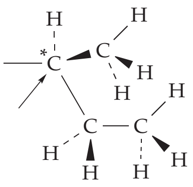
Alanine (Ala, A)



Valine (Val, V)



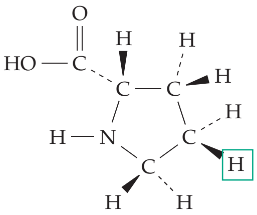
Leucine (Leu, L)



Isoleucine (Ile, I)

These bulky groups of distinctive shapes participate in hydrophobic interactions in protein interiors and in forming binding sites of specific shapes.

c. The imino acid proline. Because the side chain is fused to the α-amino group, the entire structure is given, not just the side chain, is shown.

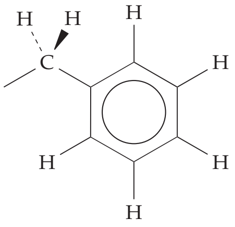


Proline (Pro, P)

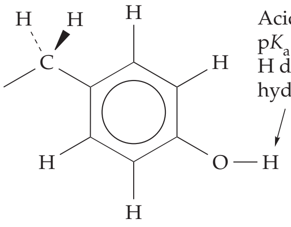
H replaced by —OH in hydroxyproline (Hyp) which occurs in collagen

Note the secondary amino group and the relatively rigid conformation. The presence of proline strongly influences the folding of protein chains.

d. Aromatic amino acids

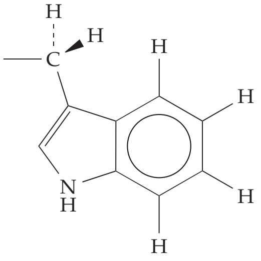


Phenylalanine (Phe, F)



Acidic protein, pK_a ~9.5–10.9. H donor in hydrogen bonds

Tyrosine (Tyr, Y)

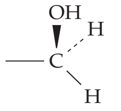


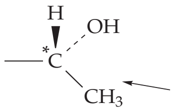
Tryptophan (Trp, W)

Tyrosine, phenylalanine, and tryptophan can form hydrophobic bonds and may be especially effective in bonding to other flat molecules.

TABLE 2-2
(continued)

e. Amino acid alcohols

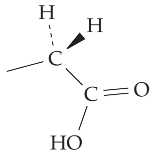

Serine
(Ser, S)

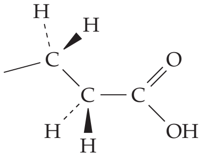

Threonine
(Thr, T)

Note the second chiral center. The L-amino acid with the opposite configuration in the side chain is L-*allo*-threonine.

The —OH group is very weakly acidic (pK_a ~13.6). It can form esters with phosphoric acid or organic acids and is a site of attachment of sugar rings in glycoproteins. Hydroxyl groups of serine are found at the active centers of some enzymes.

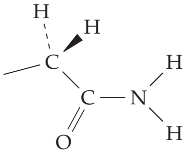
f. Acidic amino acids

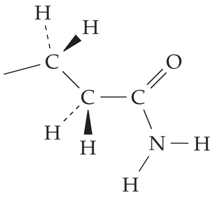

Aspartic acid (Asp, D)


Glutamic acid (Glu, E)

Carboxyl groups of these side chains are dissociated at neutral pH (pK_a values are 4.3–4.7) and provide anionic (–) groups on the surfaces of proteins.

g. Amides of aspartic acid and glutamic acid

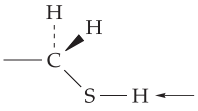

Asparagine (Asn, N)


Glutamine (Gln, Q)

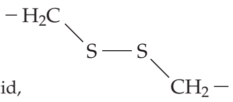
The amide group is not acidic but is polar and participates in hydrogen bonding.

If it is uncertain whether a position in a protein is occupied by aspartic acid or asparagine, it may be designated Asx or B. If glutamic or glutamine, it may be designated as Glx or Z.

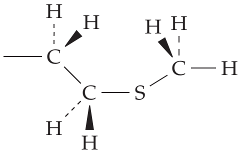
h. Sulfur-containing amino acids


Cysteine (Cys, C)

A weak acid, pK_a ~8.5



Two cysteine SH groups can be oxidatively joined to form a disulfide bridge in the "double-headed" amino acid cystine.


Methionine (Met, M)

Key

☐ **Essential in the human diet**

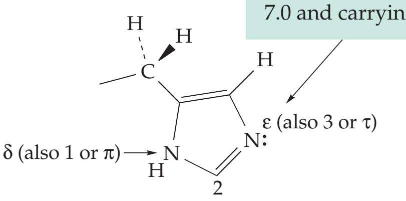
⬚ **Essential if phenylalanine or cysteine is inadequate**

▨ **Essential for rapid growth**

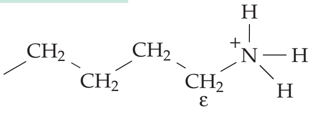
[a] The three-letter and one-letter abbreviations used for the amino acid residues in peptides and proteins are given in parentheses. B, J, U, X, and Z can be used to indicate modified or unusual amino acids.

i. Basic amino acids

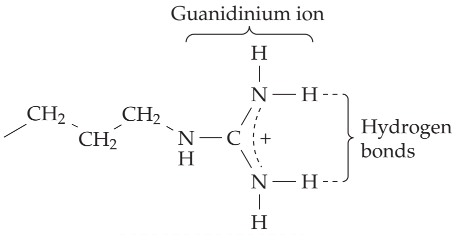
This basic site accepts a proton to form a conjugate acid of pK_a ~6.4–7.0 and carrying a positive charge.



Histidine (His, H)



Lysine (Lys, K)



Arginine (Arg, R)

The **imidazole** groups in histidine side chains are parts of the active sites of many enzymes. Like other basic groups in proteins they also may bind metal ions.

A flexible side arm with a potentially reactive amino group at the end. The high pK_a of ~10.5 means that lysine side chains are ordinarily protonated in neutral solutions.

The guanidinium group has a high pK_a of over 12 and remains protonated under most circumstances. It is stabilized by resonance as indicated by the dashed line. Guanidinium groups are often sites for binding of phosphate or carboxylate groups by pairs of hydrogen bonds.

phosphate, and adenylate) and in names of enzymes (e.g., lactate dehydrogenase).

During the formation of polypeptides, the α-amino and carboxyl groups of the amino acids are converted into the relatively unreactive and uncharged amide (peptide) groups except at the two chain termini. In many cases the terminal amino and carboxyl groups are also converted within cells into uncharged groups (Chapter 10). Immediately after the protein is synthesized its terminal carboxyl group is often converted into an amide. The N terminus may be acetylated or cyclized to a pyroglutamyl group. Sometimes a cyclic peptide is formed (Fig. 2-4).

The properties of polypeptides and proteins are determined to a large extent by the chemistry of the side chain groups, which may be summarized briefly as follows. Glycine in a peptide permits a maximum of conformational mobility. The nine relatively nonpolar amino acids–alanine, valine, leucine, isoleucine, proline, methionine, phenylalanine, tyrosine, and tryptophan–serve as building blocks of characteristic shape. Tyrosine and tryptophan also participate in hydrogen bonding and in aromatic: aromatic interactions within proteins.

Much of the chemistry of proteins involves the side chain functional groups –OH, –SH, $-COO^-$, $-NH_3^+$, and imidazole (Eq. 2-6) and the guanidinium group of arginine. The side chains of asparagine and glutamine both contain the amide group $CONH_2$, which is relatively inert chemically but which can undergo hydrogen-bonding interactions. The amide linkages of the polypeptide backbone must also be regarded as important functional groups. Most polar groups are found on the outside surfaces of proteins where they can react chemically in various ways. When inside proteins they form H-bonds to the peptide backbone and to other polar groups.

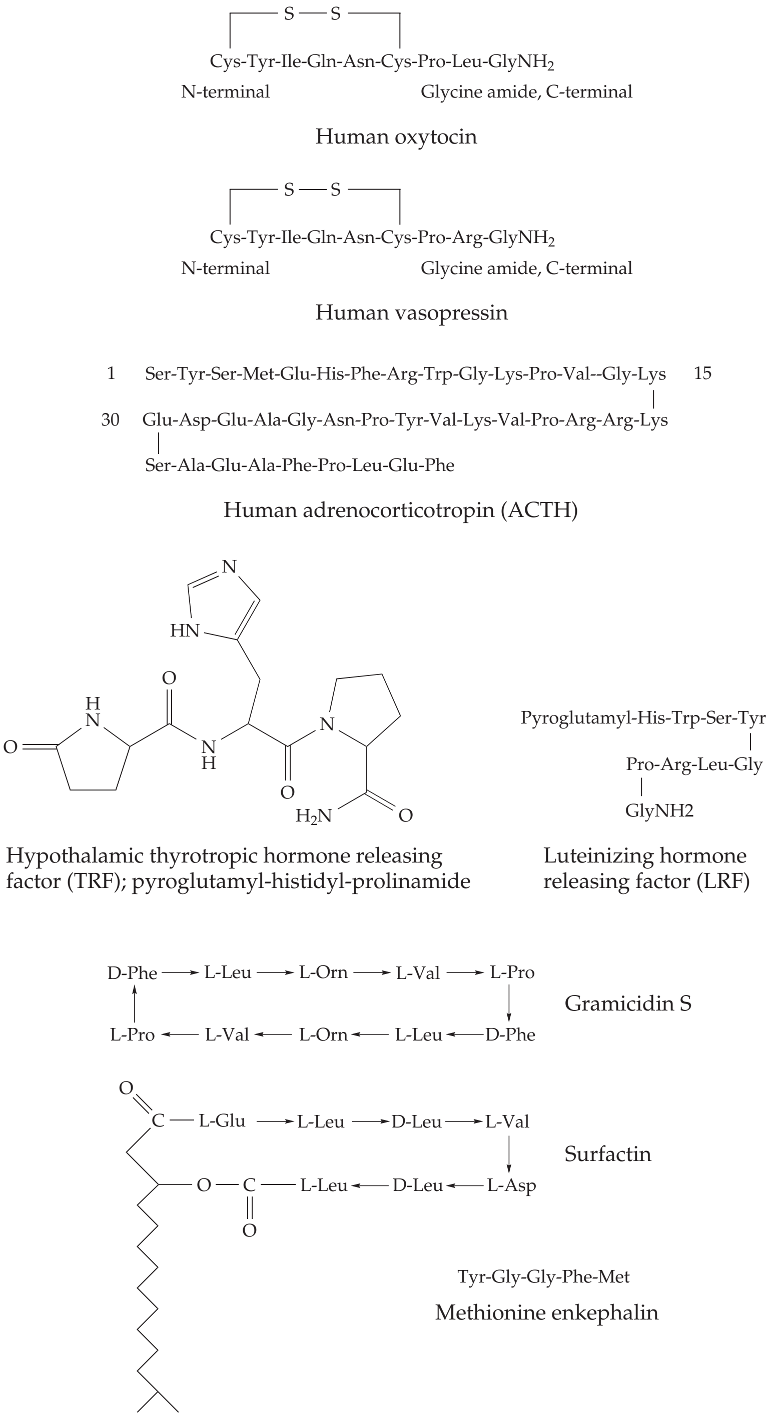
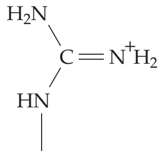

S—S
Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-GlyNH2
N-terminal
Glycine amide, C-terminal
Human oxytocin
S—S
Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Arg-GlyNH2
N-terminal
Glycine amide, C-terminal
Human vasopressin
1 Ser-Tyr-Ser-Met-Glu-His-Phe-Arg-Trp-Gly-Lys-Pro-Val--Gly-Lys 15
30 Glu-Asp-Glu-Ala-Gly-Asn-Pro-Tyr-Val-Lys-Val-Pro-Arg-Arg-Lys
Ser-Ala-Glu-Ala-Phe-Pro-Leu-Glu-Phe
Human adrenocorticotropin (ACTH)
N
HN
O
H
N
O
N
H
N
O
H2N
O
Hypothalamic thyrotropic hormone releasing factor (TRF); pyroglutamyl-histidyl-prolinamide
Pyroglutamyl-His-Trp-Ser-Tyr
Pro-Arg-Leu-Gly
GlyNH2
Luteinizing hormone releasing factor (LRF)
D-Phe L-Leu L-Orn L-Val L-Pro
L-Pro L-Val L-Orn L-Leu D-Phe
Gramicidin S
O
C L-Glu L-Leu D-Leu L-Val
O C L-Leu D-Leu L-Asp
O
Surfactin
Tyr-Gly-Gly-Phe-Met
Methionine enkephalin

Figure 2-4 Structures of some naturally occurring peptides. Oxytocin and vasopressin are hormones of the neurohypophysis (posterior lobe of the pituitary gland). Adrenocorticotropin is a hormone of the adenohypophysis (anterior pituitary). Hormones of the adenohypophysis are released under the influence of releasing factors (regulatory factors) produced in the neighboring hypothalamus (a portion of the brain) in response to neural stimulation. Structures of two releasing factors are shown. Note that the γ-carboxyl groups of the N-terminal glutamine residues have reacted, with loss of NH_3, with the neighboring terminal $-NH_2$ groups to form cyclic amide (pyroglutamyl) groups.[78] Gramicidin S is an antibiotic made by *Bacillus brevis*, and surfactin is a depsipeptide (containing an ester linkage), a surface active antibiotic of *Bacillus subtilis*. Methionine enkephalin is a brain peptide with opiate-like activity.[79]

2. Acidic and Basic Side Chains

The side chains of aspartic and glutamic acids carry negatively charged carboxylate groups at pH 7 while those of lysine and arginine carry the positively charged $–NH_3^+$ and guanidinium ions, respectively.



Guanidinium group

At pH 7 the weakly basic imidazole group of histidine may be partially protonated. Both the –SH group of cysteine and the phenolic –OH of tyrosine are weakly acidic and will dissociate and thereby acquire negative charges at a sufficiently high pH.

The number of positive and negative charges on a protein at any pH can be estimated approximately from the acid dissociation constants (usually given as pK_a values) for the amino acid side chains. These are given in Table 2-2. However, pK_a values of buried groups are often greatly shifted from these, especially if they associate as **ion pairs**. In addition, many proteins have free amino and carboxyl-terminal groups at the opposite ends of the peptide chain. These also participate in acid–base reactions with approximately the following pK_a values.

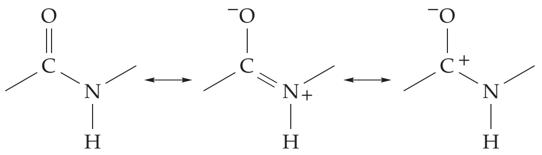
terminal —COOH, $pK_a = 3.6–3.7$
terminal $—NH_3^+$, $pK_a = 7.5–7.9$

The acid–base properties of an amino acid or of a protein are described by titration curves of the type shown in Figs. 3-1 and 3-2. In these curves the number of equivalents of acid or base that have reacted with an amino acid or protein that was initially at neutral pH are plotted against pH. The net negative or positive electrical charge on the molecule can be read directly from the curves. Both the net electrical charges and the distribution of positively and negatively charged groups are often of crucial importance to the functioning of a protein.

Additional aspects of the acid–base chemistry of amino acids and proteins are considered in Chapter 3, Section A and Chapter 6, Section E. The student may find it appropriate to study these sections at this time and to work the associated study problems.

3. The Peptide Unit

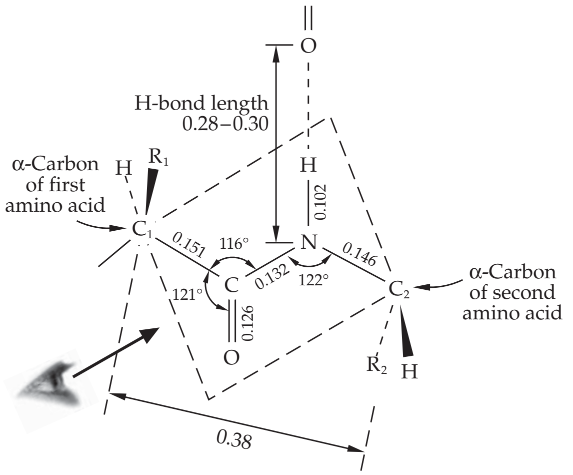
The very ability of a protein to exist as a complex three-dimensional structure depends upon the properties of the amide linkages between the amino acid units. Many of these properties follow from the fact that an amide can be viewed as a resonance hybrid of the following structures. Because of the partial double-bond character, the C–N bond is shorter than that of a normal single bond and the C=O bond is lengthened.



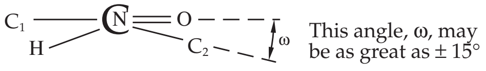
The observed lengths in nanometers determined by X-ray diffraction measurements are given in Fig. 2-5 (top). The partial double-bond character of the C–N bond has important consequences. The peptide unit is nearly planar as is indicated by the dashed parallelogram in Fig. 2-5.

However, the bonds around the nitrogen retain some pyramidal character (Fig. 2-5, bottom). Even more important is the fact that there is flexibility. As a result, the torsion angle ω may vary over a range of ± 15° or even more from that in the planar state.[80,81,81a]

The resonance stabilization of the amide linkage is thought to be about 85 kJ/mol. Rotation around the C–N bond through 90° would be expected to require



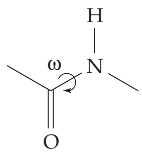
Distance (nm) between successive α-carbon atoms in a protein.



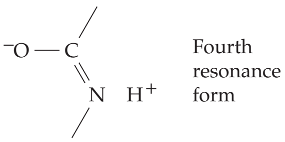
View down C—N axis
(as indicated by heavy arrow above)

Figure 2-5 Dimensions of the peptide linkage. Interatomic distances in nm, including the hydrogen bond length to an adjacent peptide linkage, are indicated. The atoms enclosed by the dotted lines all lie *approximately* in a plane. However, as indicated in the lower drawing, the nitrogen atom tends to retain some pyramidal character.

about this much energy. This fact immediately suggests a way in which proteins may sometimes be able to store energy—by having one or more peptide units twisted out of complete planarity.[69]

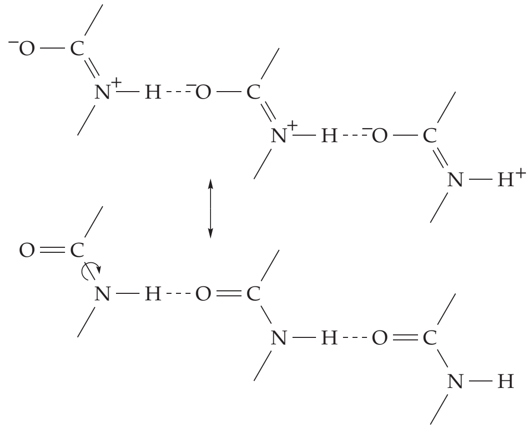


An important effect of the resonance of the amide linkage is that the oxygen atom acquires some negative charge and the NH group some positive charge. Some of the positive charge is usually depicted as residing on the nitrogen, but some is found on the hydrogen atom. The latter can be pictured as arising from a contribution of a fourth resonance form that contains no bond to hydrogen.



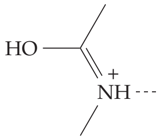
Nevertheless, this picture is inadequate. Various evidence indicates that the nitrogen actually carries a net negative charge.[81a]

The positive and negative ends of the dipoles in the amide group tend to associate to form strong hydrogen bonds. These hydrogen bonds together with the connecting amide linkages can form chains that may run for considerable distances through proteins. The tendency for cooperativity in hydrogen bond formation may impart unusual stability to these chains. As with individual amide linkages, these chains of hydrogen-bonded amides can also be thought of as resonance hybrids:

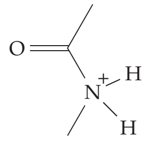


The two structures pictured are extreme forms, the true structure being something in between. In the lower form, rotation about the C–N bond would be permitted but then the charge separation present in the upper structure would no longer exist. Thus, the hydrogen bonds would be weakened. We can conclude that if an amide linkage in such a chain becomes twisted, the hydrogen bonds that it forms will be weakened. If there is cooperativity, the hydrogen bonds will all be strongest when there is good planarity in all of the amides in the chain.

Amides have very weak basic properties and protonation is possible either on the oxygen (A) or on the nitrogen (B).[82,83]

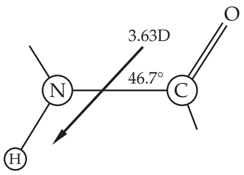


A



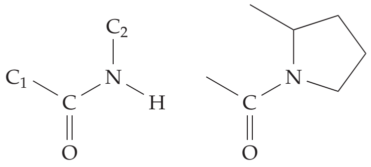
B

The pK_a values for such protonation are usually less than zero, but it is possible that a correctly placed acidic group in a protein could protonate either oxygen or nitrogen transiently during the action of a protein. Protonation on oxygen would strengthen hydrogen bonds from the nitrogen whereas protonation on nitrogen would weaken hydrogen bonds to oxygen and might permit rotation. The amide group has a permanent dipole moment of 3.63 debyes oriented as follows:



Here the arrow points toward the *positive* end of the dipole.

While the *trans* peptide linkage shown in Fig. 2-4 is usual, the following *cis* peptide linkage, which is ~8 kJ/mol less stable than the *trans* linkage, also occurs in proteins quite often. The nitrogen atom is usually but not always from proline.[81,84]

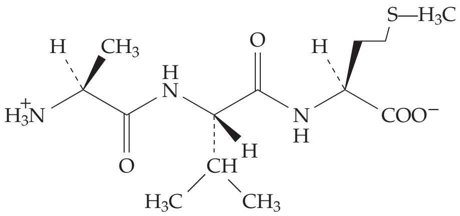


cis Peptide linkages

4. Polypeptides

The chain formed by polymerization of amino acid molecules provides the **primary structure** of a protein. Together with any covalent crosslinkages and other modifications, this may also be called the

covalent structure of the protein. Each monomer unit in the chain is known as an amino acid **residue**. This term acknowledges the fact that each amino acid has lost one molecule of H_2O during polymerization. To be more precise, the number of water molecules lost is *one less* than the number of residues. Peptides are named according to the amino acid residues present and beginning with the one bearing the terminal amino group. Thus, L-alanyl-L-valyl-L-methionine has the following structure:



Like amino acids, this tripeptide is a dipolar ion. The same structure can be abbreviated Ala-Val-Met or, using one-letter abbreviations, AVM. It is customary in describing amino acid sequences to place the amino-terminal (N-terminal) residue at the left end and the carboxyl-terminal (C-terminal) residue at the right end. Residues are numbered sequentially with the N-terminal residue as 1. An example is shown in Fig. 2-6.

The sequence of amino acid units in a protein is always specified by a gene. The sequence determines how the polypeptide chain folds and how the folded protein functions. For this reason much effort has gone into "sequencing," the determination of the precise order of amino acid residues in a protein. Sequences of several hundreds of thousands of proteins and smaller peptides have been established and the number doubles each year.[75,87,88,88a] Most of these

A

```
1                 10                  20                  30                  40
A P P S V F A E V P Q A Q P V L V F K L I A D F R E D P D P R K V N L G V G A Y
                  50                  60                  70                  80
R T D D C Q P W V L P V V R K V E Q R I A N N S S L N H E Y L P I L G L A E F R
                  90                  100                 110                 120
T C A S R L A L G D D S P A L Q E K R V G G V Q S L G G T G A L R I G A E F L A
                  130                 140                 150                 160
R W Y N G T N N K D T P V Y V S S P T W E N H D G V F T T A G F K D I R S Y R Y
                  170                 180                 190                 200
W D T E K R G L D L Q G F L S D L E N A P E F S I F V L H A C A H N P T G T D P
                  210                 220                 230                 240
T P E Q W K Q I A S V M K R R F L F P F F D S A Y Q G F A S G N L E K D A W A I
                  250                 260                 270                 280
R Y F V S E G F E L F C A Q S F S K N F G L Y N E R V G N L T V V A K E P D S I
                  290                 300                 310                 320
L R V L S Q M Q K I V R V T W S N P P A Q G A R I V A R T L S P D E L F H E W T
                  330                 340                 350                 360
G N V K T M A D R I L S M R S E L R A R L E A L K T P G T W N H I T D Q I G M F
                  370                 380                 390                 400
S F T G L N P K Q V E Y L I N E K H I Y L L P S G R I N M C G L T T K N L D Y V
                  410
A T S I H E A V T K I Q
```

B

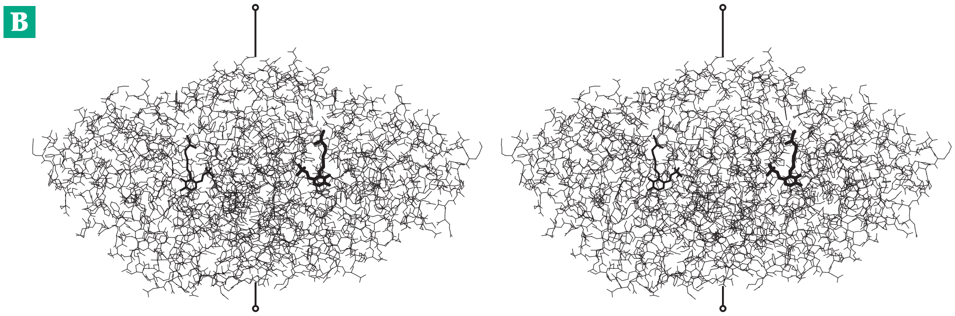


Figure 2-6 (A) The complete amino acid sequence of the cytoplasmic enzyme aspartate aminotransferase from pig heart. The peptide has the composition Lys_{19}, His_8, Arg_{26}, $(Asp + Asn)_{42}$, Ser_{26}, Thr_{26}, $(Glu + Gln)_{41}$, Pro_{24}, Gly_{28}, Ala_{32}, Cys_5, Val_{29}, Met_6, Ile_{19}, Leu_{38}, Tyr_{12}, Phe_{23}, Trp_9. The molecular mass is 46.344 kDa and the complete enzyme is a 93.147-kDa dimer containing two molecules of the bound coenzyme pyridoxal phosphate attached to lysine-258 (enclosed in box).[85,86] (B) A stereoscopic view of a complete enzyme molecule which contains two identical subunits with the foregoing sequence. Coordinates from Arthur Arnone. In this "wire model" all the positions of all of the nearly 7000 atoms that are heavier than hydrogen are shown. The > 8000 hydrogen atoms have been omitted. The view is into the active site of the subunit on the right. The pyridoxal phosphate and the lysine residue to which it is attached are shown with heavy lines. The active site of the subunit to the left opens to the back side as viewed here. The drawing may be observed best with a magnifying viewer available from Abrams Instrument Corp., Lansing, Michigan or Luminos Photo Corp., Yonkers, New York. However, with a little practice, it is possible to obtain a stereoscopic view unaided. Hold the book with good illumination about 20–30 cm from your eyes. Allow your eyes to relax as if viewing a distant object. Of the four images that are visible, the two in the center can be fused to form the stereoscopic picture. Drawings by program MolScript (Kraulis, 1991).

BOX 2-A PROTEINS OF BLOOD PLASMA

Among the most studied of all proteins are those present in blood plasma.[a–f] Their ready availability and the clinical significance of their study led to the early development of electrophoretic separations. Electrophoresis at a pH of 8.6 (in barbital buffer) indicated six main components. The major and one of the fastest moving proteins is **serum albumin**. Trailing behind it are the **α_1-, α_2-**, and **β-globulins**, **fibrinogen**, and **γ-globulins**. Each of these bands consists of several proteins and two-dimensional separation by electrophoresis and isoelectric focusing (Chapter 3) reveals over 30 different proteins.[e] Many of these contain varying numbers of attached carbohydrate units and appear as families of spots.

Fractionation of large quantities of plasma together with immunochemical assays has led to identification of over 200 different proteins. Sixty or more are enzymes, some in very small quantitites which may have leaked from body cells. Normally plasma contains 5.7 – 8.0 g of total protein per 100 ml (~1 mM). Albumin accounts for 3.5 – 4.5 g/100 ml. An individual's liver synthesizes about 12 g each day. Next most abundant are the **immunoglobulins**. One of these (IgG or γ-globulin) is present to the extent of 1.2–1.8 g/100 ml. Also present in amounts greater than 200 mg per 100 ml are **α-** and **β-lipoproteins**, the **α_1 antitrypsin**, **α_2-macroglobulin**, **haptoglobin**, **transferrin**, and fibrinogen.

Plasma proteins have many functions. One of them, fullfilled principally by serum albumin, is to impart enough osmotic pressure to plasma to match that of the cytoplasm of cells. The heart-shaped human serum albumin consists of a single 65 kDa chain of 585 amino acid residues coiled into 28 helices.[g] Three homologous repeat units or domains each contain six disulfide bridges, suggesting that gene duplication occurred twice during the evolution of serum albumins. The relatively low molecular mass and high density of negative charges on the surface make serum albumin well adapted for the role of maintaining osmotic pressure. However, serum albumin is not essential to life. Over 50 mutant forms have been found and at least 30 persons have been found with no serum albumin in their blood.[h,i] These analbuminemic individuals are healthy and have increased concentrations of other plasma proteins.

Human serum albumin. From He and Carter.[g]

A second major function of plasma proteins is transport. Serum albumin binds to and carries many sparingly soluble metabolic products, including fatty acids, tryptophan, cysteine, steroids, thyroid hormones, Ca^{2+}, Cu^{2+}, Zn^{2+}, other metal ions, bilirubin, and various drugs. There are also many more specialized transporter proteins. **Transferrin** carries iron and **ceruloplasmin** (an α_2 globulin) transports copper. **Transcortin** carries corticosteroids and progesterone, while another protein carries sex hormones. **Retinol-binding protein** carries vitamin A and **cobalamin-binding proteins** vitamin B_{12}. **Hemopexin** carries heme to the liver, where the iron can be recovered.[j] Haptoglobin binds hemoglobin released from broken red cells and also assists in the recycling of the iron in the heme.[k] **Lipoproteins** (see Table 21-1) carry phospholipids, neutral lipids, and cholesterol esters. Most of the mass of these substances is lipid.

Immunoglobulins, α_1-trypsin inhibitor and **α_2-macroglobulin**,[k] ten or more blood clotting factors; and proteins of the **complement system** all have protective functions that are discussed elsewhere in this book. Hormones, many of them proteins, are present in the blood as they are carried to their target tissues. Many serum proteins have unknown or poorly understood functions. Among these are the **acute phase proteins**, whose concentrations rise in response to inflammation or other injury.

---

[a] Allison, A. C., ed. (1974) *Structure and Function of Plasma Proteins*, Vol. 1, Plenum, New York

[b] Allison, A. C., ed. (1976) *Structure and Function of Plasma Proteins*, Vol. 2, Plenum, New York

[c] Putnam, F. W., ed. (1975) *The Plasma Proteins*, 2nd ed., Vol. 1 and 2, Academic Press, New York

[d] Blomback, B., and Hanson, L. A., eds. (1979) *Plasma Proteins*, John Wiley, Chichester, Oklahoma

[e] Geisow, M. J., and Gordon, A. H. (1978) *Trends Biochem. Sci.* **3**, 169–171

[f] Smith, E. L., Hill, R. L., Lehman, I. R., Lefkowitz, R. J., Handler, P., and White, A. (1983) in *Principles of Biochemistry, Mammalian Biochemistry*, 7th ed., McGraw-Hill, New York, pp. 3–37

[g] He, X. M., and Carter, D. C. (1992) *Nature* **358**, 209–215

[h] Peters, T., Jr. (1996) *All About Albumin: Biochemistry, Genetics and Medical Applications*, Academic Press, San Diego, California

[i] Madison, J., Galliano, M., Watkins, S., Minchiotti, L., Porta, F., Rossi, A., and Putnam, F. W. (1994) *Proc. Natl. Acad. Sci: USA* **91**, 6476–6480

[j] Satoh, T., Satoh, H., Iwahara, S.-I., Hrkal, Z., Peyton, D. H., and Muller-Eberhard, U. (1994) *Proc. Natl. Acad. Sci: U.S.A.* **91**, 8423–8427

[k] Feldman, S. R., Gorias, S. L., and Pizzo, S. V. (1985) *Proc. Natl. Acad. Sci: U.S.A.* **82**, 5700–5704

have been deduced from the sequences of nucleotides in DNA. Sequences of some small peptide hormones and antibiotics are shown in Fig. 2-4 and that of a 412-residue protein in Fig. 2-6. The molecular mass of a protein can be estimated from the chain length by assuming that each residue adds 100–115 Da.

The amino acid composition varies greatly among proteins. A typical globular protein contains all or most of the 20 amino acids. The majority are often present in roughly similar amounts but His, Cys, Met, Tyr, and Trp tend to be less abundant than the others. Specialized proteins sometimes have unusual amino acid compositions. For example, collagen of connective tissue contains 33 mole% glycine and 21% of proline + hydroxyproline residues; the major proteins of saliva contain 22% of glutamate + glutamine and 20–45% proline.[89] Cell walls of plants contain both high proline and high glycine polypeptides. One from petunias is 67% glycine.[90] Silk fibroin contains 45% glycine and 29% alanine. A DNA repair protein of yeast has 13 consecutive aspartate residues.[91] The tough eggshell (chorion) of the domesticated silkmoth *Bombyx mori* contains proteins with ~30% cysteine.[92] Many proteins consist, in part, of repeated short sequences. For example, the malaria-causing *Plasmodium falciparum* in its sporozoite stage is coated with a protein that contains 37 repeats of the sequence NANP interspersed with 4 repeats of NDVP.[93] These two sequences have been indicated with single-letter abbreviations for the amino acids.

With a large number of protein and DNA sequences available, it has become worthwhile to compare sequences of the same protein in different species or of different proteins within the same or different species. Computer programs make it possible to recognize **similarities** and **homologies** between sequences even when deletions and insertions have occurred.[88,88a,94–97] The term homology has the precise biological definition "*having a common evolutionary origin*," but it is often used to describe any close similarity in sequence.[98] Among a pair of homologous proteins, a change at a given point in a sequence may be either **conservative**, meaning that a residue of similar character (large, small, positively charged, nonpolar, etc.) has been substituted, or it may be **nonconservative**.

D. The Architecture of Folded Proteins

All proteins are made in the same way but as the growing peptide chains peel off from the ribosome, each of the thousands of different proteins in a living cell folds into its own special **tertiary structure**.[88a] The number of possible conformations of a protein chain is enormous. Consider a 300-residue polypeptide which could stretch in fully extended form for ~100 nm. If the chain were folded back on itself about 13 times it could form a 7-nm square sheet about 0.5 nm thick. The same polypeptide could form a thin helical rod 45 nm long and ~1.1 nm thick. If it had the right amino acid sequence it could be joined by two other similar chains to form a collagen-type triple helix of 87 nm length and about 1.5 nm diameter (Fig. 2-7).

The highly folded **globular proteins** vary considerably in the tightness of packing and the amount of internal water of hydration.[99,100] However, a density of ~1.4 g cm^{-3} is typical. With an average mass per residue of 115 Da our 300-residue polypeptide would have a mass of 34.5 kDa or 5.74 x 10^{-20} g and a volume of 41 nm^3. This might be approximated by a cube 3.45 nm in width, a "brick" of dimensions 1.8 x 3.6 x 6.3 nm, or a sphere of diameter 4.3 nm. Although protein molecules are usually very irregular in shape,[101] for purposes of calculation idealized ellipsoid and rod shapes are often assumed (Fig. 2-7).

It is informative to compare these dimensions with those of the smallest structures visible in cells; for example, a bacterial flagellum is ~13 nm in diameter and a cell membrane ~8–10 nm in thickness. Bricks of the size of the 300-residue polypeptide could be used to assemble a bacterial flagellum or a eukaryotic microtubule. Helical polypeptides may extend through cell membranes and project on both sides, while a globular protein of the same chain length may be almost completely embedded in the membrane.

1. Conformations of Polypeptide Chains

To understand how a polypeptide chain folds we need to look carefully at the possible conformations of the peptide units. Since each peptide unit is nearly planar, we can think of a polypeptide as a chain of flat units fastened together as in Fig. 2-8. Every peptide unit is connected to the next by the α-carbon of an amino acid. This carbon provides two single bonds to the chain and rotation can occur about both of them (except in the cyclic amino acid proline). To specify the conformation of an amino acid unit in a polypeptide chain, we must describe the torsion angles about both of these single bonds.[11,76,102] These angles are indicated by the symbols **ϕ (phi)** and **ψ (psi)** and are assigned the value 180° for the fully extended chain as shown in Fig. 2-8. Each angle is taken as zero for the impossible conformation in which the two chain ends are in the eclipsed conformation. By the same token, the torsion angle **ω (omega)** around the C–N bond of the amide is 0° for a planar *cis* peptide linkage and 180° for the usual *trans* linkage.

Since both ϕ and ψ can vary for each residue in a protein, there are a large number of possible conformations. However, many are excluded because they bring certain atoms into collision. This fact can be established readily by study of molecular models.

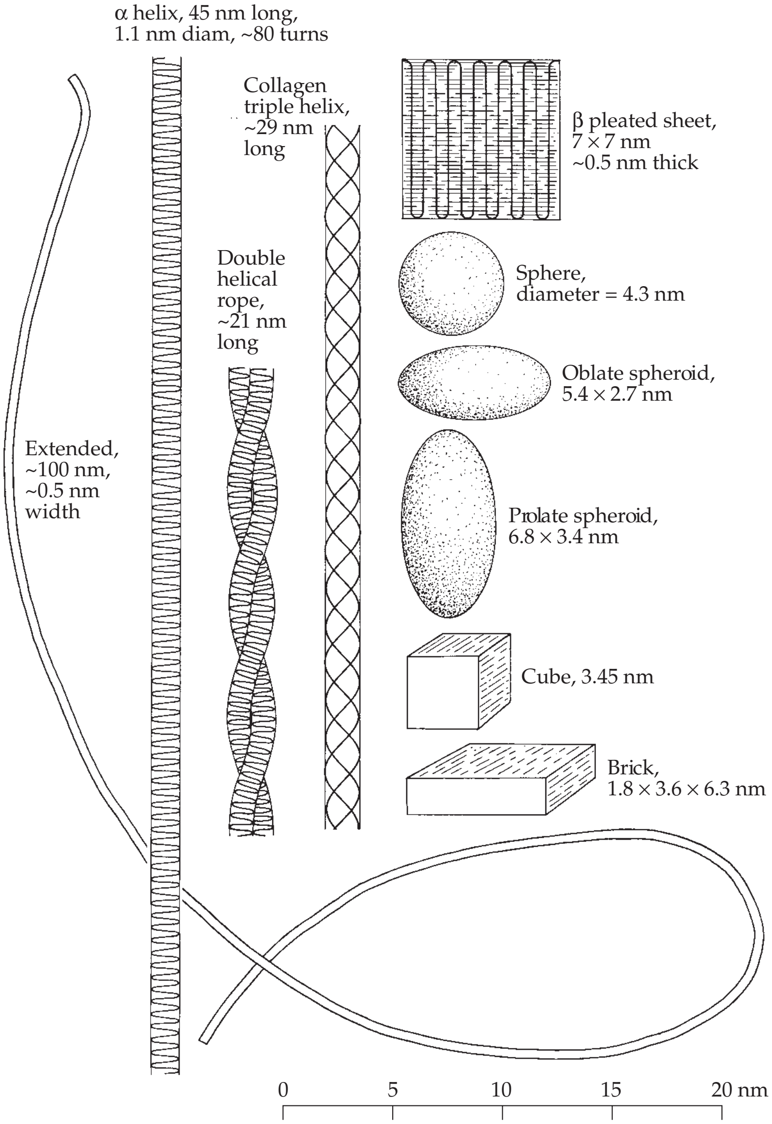

α helix, 45 nm long,
1.1 nm diam, ~80 turns
Collagen triple helix, ~29 nm long
β pleated sheet, 7 × 7 nm ~0.5 nm thick
Sphere, diameter = 4.3 nm
Double helical rope, ~21 nm long
Oblate spheroid, 5.4 × 2.7 nm
Extended, ~100 nm, ~0.5 nm width
Prolate spheroid, 6.8 × 3.4 nm
Cube, 3.45 nm
Brick, 1.8 × 3.6 × 6.3 nm
0
5
10
15
20 nm

Figure 2-7 Some idealized shapes that a 34.5 kDa protein molecule of 300 amino acids might assume.

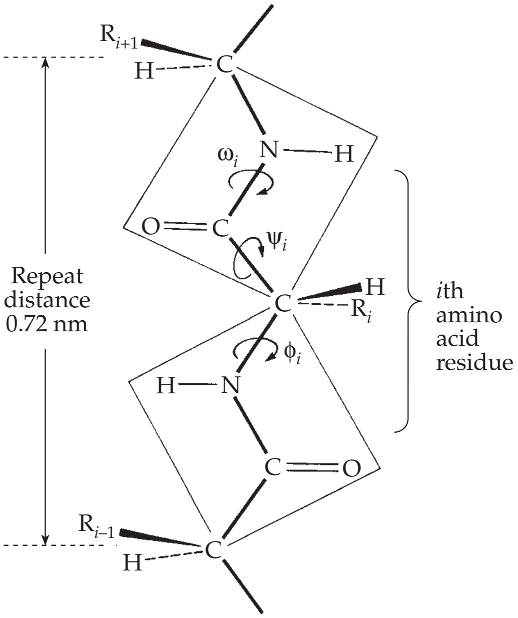

R_{i+1}
H
C
N
H
ω_i
O
C
ψ_i
C
H
R_i
ϕ_i
H
N
C
O
R_{i-1}
H
C
Repeat distance 0.72 nm
*i*th amino acid residue

Figure 2-8 Two peptide units in the completely extended β conformation. The torsion angles ϕ_i, ψ_i, and ω_i are defined as 0° when the main chain atoms assume the *cis* or eclipsed conformation. The angles in the completely extended chain are all 180°. The distance from one α carbon atom (C_α) to the next in a peptide chain is always 0.38 nm, no matter how the chain is folded.

Using a computer, it is possible to study the whole range of possible combinations of ϕ and ψ. This has been done for the peptide linkage by Ramachandran. The results are often presented as plots of ϕ vs ψ (Ramachandran plots or **conformational maps**)[102a,103,103a] in which possible combinations of the two angles are indicated by blocked out areas. The original Ramachandran plots were made by representing the atoms as hard spheres of appropriate van der Waals radii, but the version shown in Fig. 2-9 was calculated using a complex potential energy function to represent the van der Waals attraction and the repulsion from close contact.[103] This map was calculated for poly-L-alanine but it would be very similar for most amino acids.

Notice the two areas connected by a higher energy bridge on the left side of Fig. 2-9. The upper area contains the pairs of torsion angles for the extended **β structures** as well as for **collagen**. The lower area contains allowed conformations for the *right-handed helices*. As can be seen from Fig. 2-10, most of the observed conformations of peptide units in a real protein fall into these regions. Glycyl residues are an exception. Since glycine has no β-carbon atom, the conformations are less restricted. Out of nearly 1900 non-glycine residues in well-determined protein structures, 66 were found in disallowed areas of the Ramachandran diagram.[104a] These were often accommodated by local distortions in bond angles. The positions at which such steric strain occurs are often in regions concerned with function.[104b] One residue, which lies in a disallowed region in Fig. 2-10 is asparagine 297 of aspartate aminotransferase (Fig. 2-6). It is located

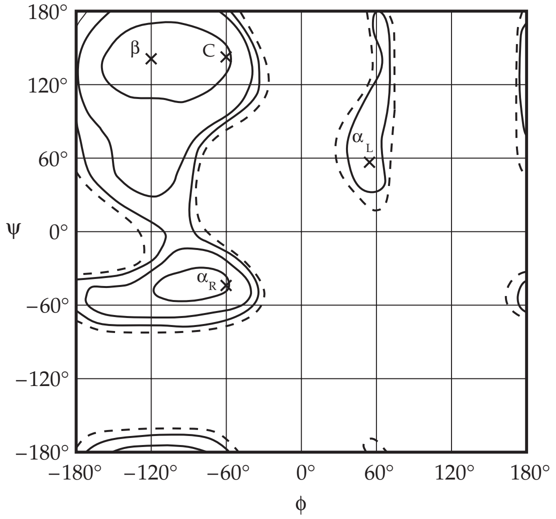


Figure 2-9 Potential energy distribution in the ϕ–ψ plane for a pair of peptide units with alanyl residues calculated using potential parameters of Scheraga and Flory. Contours are drawn at intervals of 1 kcal (4.184 kJ) per mol going down from 0 kcal per mol. The zero contour is dashed. From Ramachandran *et al.*[104] The points marked x are for the four ideal structures: twisted β structure (β), collagen (C), right-handed α helix (α_R), and the less favored left-handed α helix (α_L).

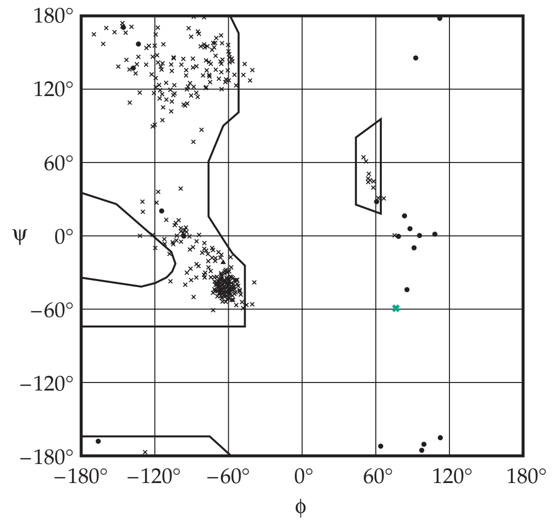


Figure 2-10 Ramachandran plot for cytosolic aspartate aminotransferase. The angles ϕ and ψ were determined experimentally from X-ray diffraction data at 0.16 nm resolution and model building. The majority of conformations are those of α helices or of β structure. Glycine residues are indicated by filled circles, while all other residues are denoted by an "×". One of these (green) lies quite far from an allowed area and must give rise to localized strain.[104b] Extreme lower limit "allowed" regions by the hard-sphere criteria are shown in outline. From coordinates of Arthur Arnone *et al.* (unpublished).[106]

adjacent to the coenzyme site. The possible conformations of proline residues are limited. The angle ϕ is always −60 ± 20°, while ψ for the residue adjacent to the proline N can be either ~150° or ~ −30°.[105] Typical ϕ, ψ angles for some regular peptide structures are given in Table 2-3.

Conformations of side chain groups are designated by a series of torsion angles designated χ_i.[11] Within proteins there are strong preferences for certain χ_i,χ_2 pairs.[107] Torsion angles within proline rings are considered in Chapter 5,A,6.

2. The Extended Chain β Structures

As was first pointed out by Pauling and Corey,[108,109] an important structural principle is that within proteins the maximum possible number of hydrogen bonds involving the C=O and N–H groups of the peptide chain should be formed. One simple way to do this is to line up fully extended chains (ϕ = ψ = 180°) and to form hydrogen bonds between them. Such a structure

TABLE 2-3
Approximate Torsion Angles for Some Regular Peptide Structures[a]

Structure	ϕ (deg.)	ψ (deg.)
Hypothetical fully extended polyglycine chain[b]	– 180	+180
β-Poly(L-alanine) in antiparallel-chain pleated sheet	– 139	+135
Parallel-chain pleated sheet	– 119	+113
Twisted β strand	– 120	+140
Polyglycine II	– 80	+150
Poly(L-proline) II	– 78	+149
Collagen[c]	– 60 ± 15	– 140 ± 15
Right-handed α helix[d]	– 57	– 47
Left-handed α helix	+57	+47
β Bends: Type I, residue 2	– 60	– 30
residue 3	– 90	0
Type II, residue 2	– 60	120
residue 3	80	0
Type III, residue 2 & 3	– 60	– 30
β Bulges[e]: "Classical" β bulge residue 1	– 100	– 45
G1 bulge residue 1	85	0

[a] From Liébecq, C., ed. (1992) *Biochemical Nomenclature*, Portland Press, London and Chapel Hill (for the International Union of Biochemistry and Molecular Biology).

[b] Torsion angles for the fully extended chain can be designated either +180° or –180°, the two being equivalent. They are given as ϕ = –180° and ψ = +180° to facilitate comparison with the other structures.

[c] Ramachandran, G. N. (1967) *Treatise on Collagen*, Vol. 1, p. 124, Academic Press, New York

[d] Both ϕ and ψ are quite variable but ϕ + ψ = –104°.

[e] For residues other than these indicated the torsion angles are about those of typical β structures.

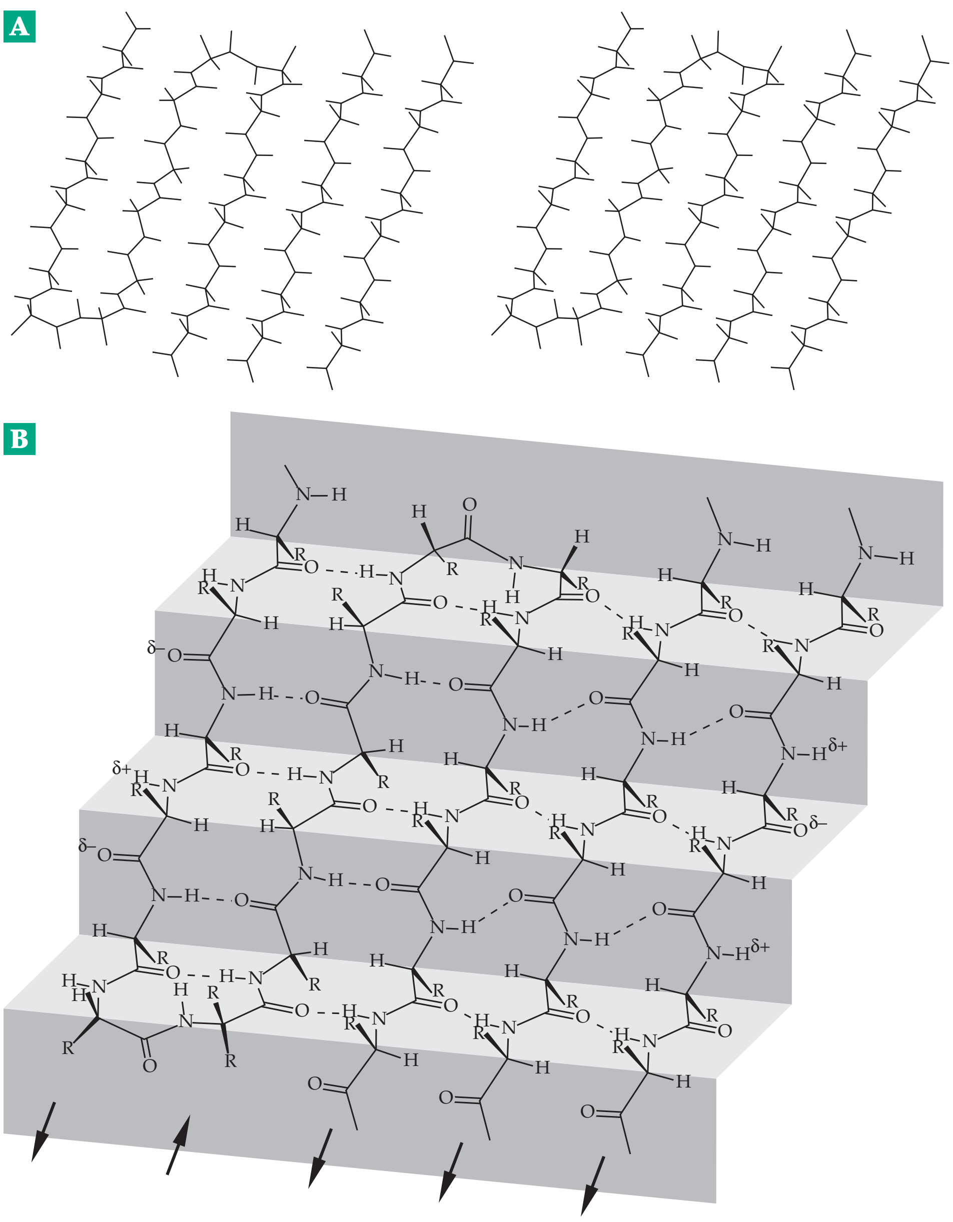


Figure 2-11 The extended chain β pleated sheet structures. (A) Stereoscopic drawing without atomic symbols. (B) Drawing with atomic symbols. At the left is the *antiparallel* structure. The 0.70 nm spacing is slightly decreased from the fully extended length. The amino acid side chains (R) extend alternately above and below the plane of the “accordion pleated” sheet. The pairs of linear hydrogen bonds between the chains impart great strength to the structure. The chain can fold back on itself using a “β turn” perpendicular to the plane of the pleated sheet. The *parallel* chain structure (right side) is similar but with a less favorable hydrogen bonding arrangement. Arrows indicate chain directions.

exists for polyglycine and resembles that in Fig. 2-11. Notice that on the left side of this figure, the adjacent chains run in opposite directions; hence, the term **antiparallel β structure**. The antiparallel arrangement not only gives the best hydrogen bond formation between chains but also permits a single chain to fold back on itself giving a compact hairpin loop.

Pleated sheets. While a fully extended polyglycine chain is possible, the side chains of other amino acids cannot be accommodated without some distortion of the structure. Thus, the peptide chains in silk fibroin have a repeat distance of 0.70 nm compared with the 0.72 nm for the fully extended chain (Fig. 2-8). Pauling and Corey[108] showed that this shortening of the chain could result from rotation of angle ϕ by ~40° (to –140°) and rotation of ψ in the opposite direction by ~45° (to +135°) to give a slightly puckered chain. The resulting multichain structure (shown in Fig. 2-11) is known as a **pleated sheet**. As in this figure, both

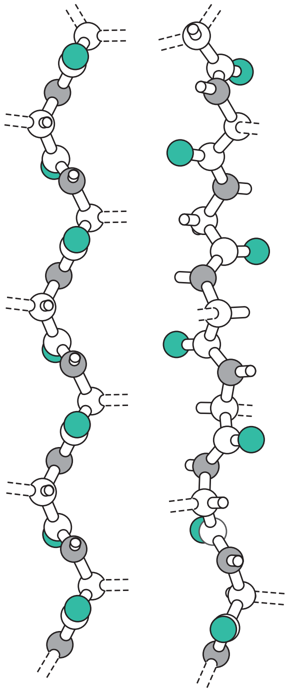


Figure 2-12 Straight (left) and twisted (right) peptide chains in extended β conformations. From Chothia.[110]

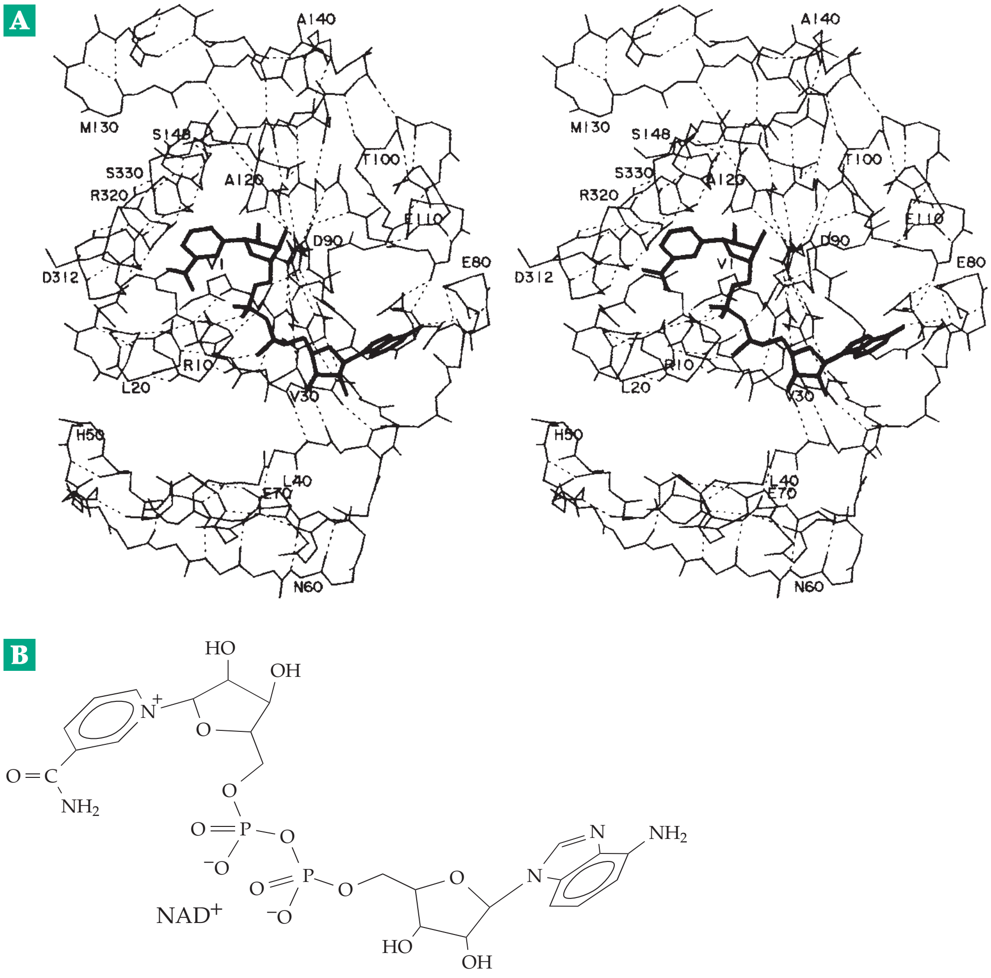


Figure 2-13 (A) Stereoscopic view of the nucleotide binding domain of glyceraldehyde phosphate dehydrogenase. The enzyme is from *Bacillus stearothermophilus* but is homologous to the enzyme from animal sources. Residues are numbered 0–148. In this wire model all of the main chain C, O, and N atoms are shown but side chains have been omitted. The large central twisted β sheet, with strands roughly perpendicular to the page, is seen clearly; hydrogen bonds are indicated by dashed lines. Helices are visible on both sides of the sheet. The coenzyme NAD^+ is bound at the end of the β sheet toward the viewer. Note that the two phosphate groups in the center of the NAD^+ are H-bonded to the N terminus of the helix beginning with R10. From Skarzynski *et al.*[111a] (B) Structural formula for NAD^+.

parallel and antiparallel strands are often present in a single β sheet within a protein.

The β structure is one of the most important **secondary structures** in proteins. It occurs in about 80% of the soluble globular proteins whose structures have been determined. In many cases almost the entire protein is made up of β structure. Single strands of extended polypeptide chain are sometimes present within globular proteins but more often a chain folds back on itself to form a hairpin loop. A second fold may be added to form an antiparallel "β meander"[102] and additional folds to form β sheets. Beta structures are found in silk fibers (Box 2-B) as well as in soluble proteins.

Twisted sheets. X-ray diffraction studies have shown that β pleated sheets are usually not flat but are twisted. In a twisted sheet the individual polypeptide chains make a shallow left-handed helix. However, when successive carbonyl groups are viewed along the direction of the chain, a **right-handed twist** is seen (Fig. 2-12).[110] Such twisted β sheets are often found in the globular proteins. An example (Fig. 2-13) is the "nucleotide-binding" domain of a dehydrogenase enzyme. The twist of the sheet is seen clearly in this stereoscopic view. When such chains are associated into β sheets, whether parallel or antiparallel, and are viewed in a direction perpendicular to the chains and looking down the edge of the sheet, a left-handed "propeller" is seen. Such a propeller is visible in the drawing of carboxypeptidase A (Fig. 2-14).

The cause of the twist in β sheets appears to lie in noncovalent interactions between hydrogen atoms on the β-carbon atoms of side chains and the peptide backbone atoms. For side chains of most L- amino acids these interactions provide a small tendency towards the observed right-handed twist.[111] Nonplanarity in the amide groups (Fig. 2-5) may also contribute. Interstrand interactions seem to be important.[111b]

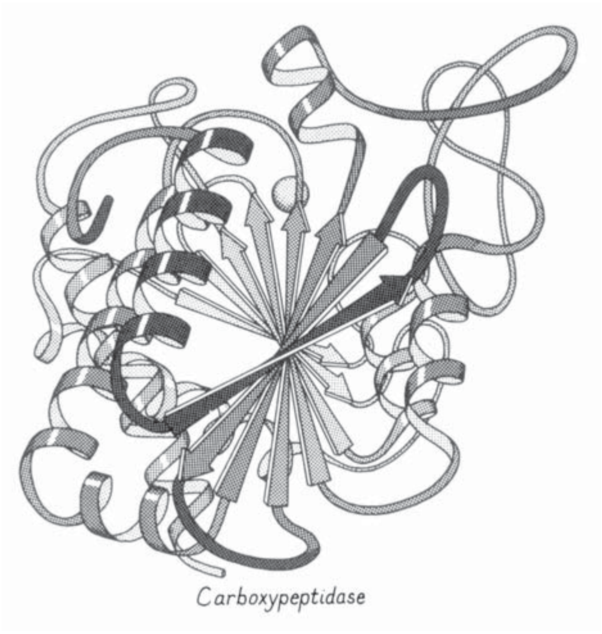


Figure 2-14 A "ribbon" drawing of the 307- residue protein-hydrolyzing enzyme carboxypeptidase A. In this type of drawing wide ribbons are used to show β strands and helical turns while narrower ribbons are used for bends and loops of the peptide chains. The direction from the N terminus to C terminus is indicated by the arrowheads on the β strands. No individual atoms are shown and side chains are omitted. Courtesy of Jane Richardson.[117]

Properties of β sheets. In antiparallel β sheets, nonpolar residues are often present on one side of the sheet and polar residues on the other. The nonpolar side may be buried in the protein, perhaps backed up against another β sheet of similar structure as in the carbohydrate-binding **lectin** shown in Fig 2-15 to give a **β sandwich**. To accommodate this packing arrangement, nonpolar and polar residues tend to alternate in the amino acid sequence. The facing nonpolar surfaces of the two sheets are essentially smooth. The β strands in one sheet lie at an angle of ~30° to those in the other sheet. The twist of the strands allows pairs of nonpolar side chains from the two sheets to maintain good van der Waals contact with each other for a considerable distance along the strands. Beta sheets of silks (Box 2-B) and of immunoglobulin domains (Fig 2-16B) are also thought to be associated in a back-to-back fashion. *Parallel* β structures have been found only when there are five or more strands, some of which may be antiparallel to the others. The parallel β structure is apparently less stable than the antiparallel structure. Parallel strands are usually buried in the protein, being surrounded by either other extended strands or helices.

Chains of hydrogen bonds and amide linkages pass across β sheets perpendicular to the chain direction. There are partial positive and negative charges along the outside edges of the sheet where these hydrogen bond chains terminate (Fig. 2-11). The polarity of these chains alternates and the positive end of one peptide bond is relatively near to the negative end of the next along the edges of the sheets. These "unsatisfied ends" of hydrogen bond chains are often sites of

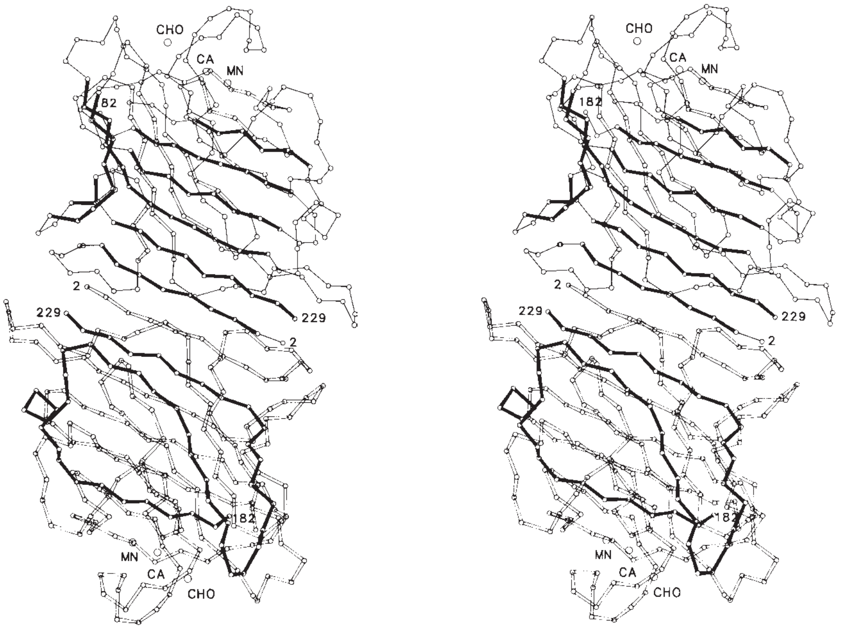


Figure 2-15 A stereoscopic **alpha-carbon plot** showing the three-dimensional structure of favin, a sugar-binding lectin from the broad bean (*Vicia faba*). In this plot only the α-carbon atoms are shown at the vertices. The planar peptide units are represented as straight line segments. Side chains are not shown. The protein consists of two identical subunits, each composed of a 20-kDa α chain and a 20-kDa β chain. The view is down the twofold rotational axis of the molecule. In the upper subunit the residues involved in the front β sheet are connected by double lines, while those in the back sheet are connected by heavy solid lines. In the lower subunit the α chain is emphasized. Notice how the back β *sheet* (not the chain) is continuous between the two subunits. Sites for bound Mn^{2+} (MN), Ca^{2+} (CA), and sugar (CHO) are marked by larger circles. From Reeke and Becker.[112]

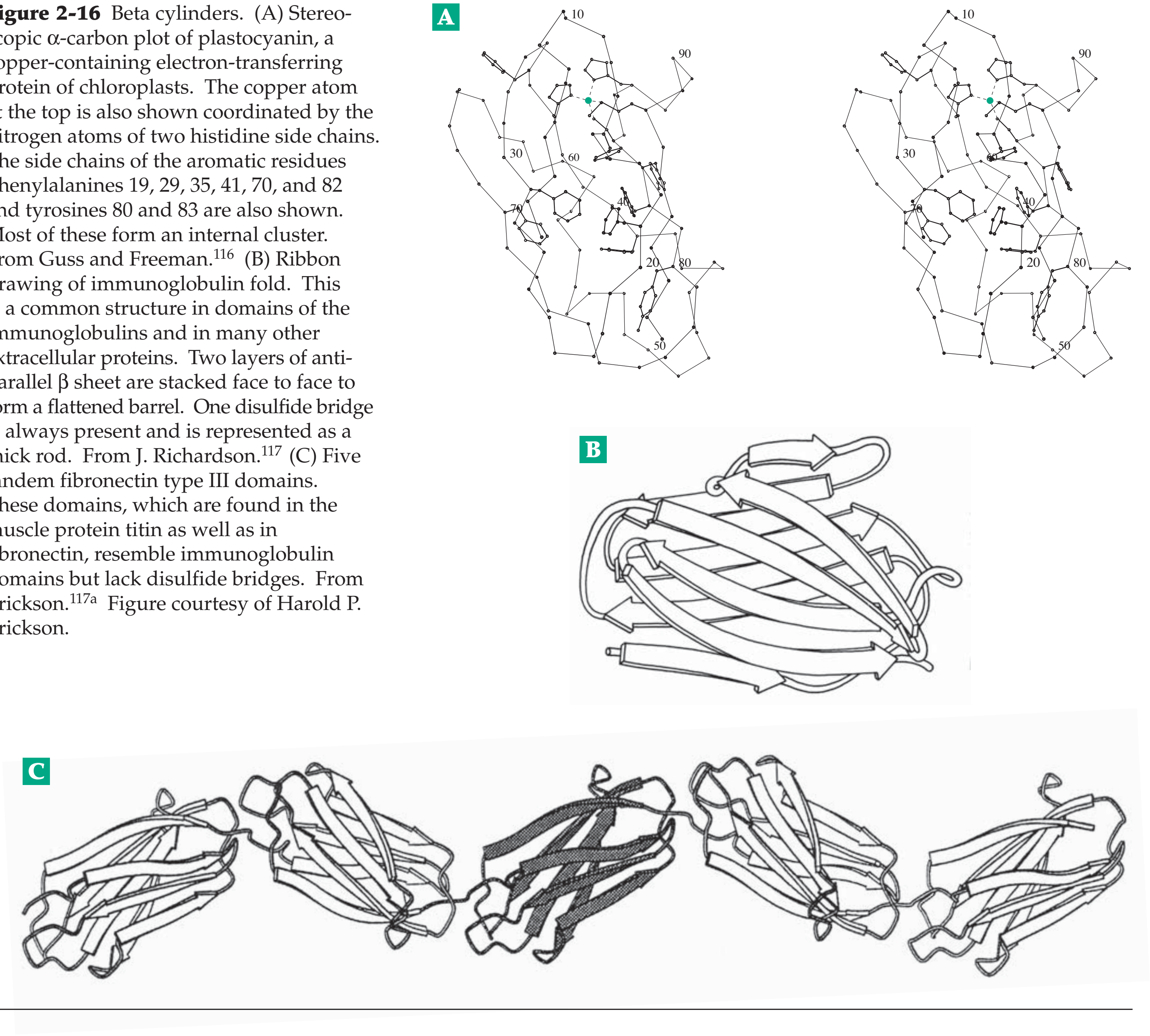


Figure 2-16 Beta cylinders. (A) Stereoscopic α-carbon plot of plastocyanin, a copper-containing electron-transferring protein of chloroplasts. The copper atom at the top is also shown coordinated by the nitrogen atoms of two histidine side chains. The side chains of the aromatic residues phenylalanines 19, 29, 35, 41, 70, and 82 and tyrosines 80 and 83 are also shown. Most of these form an internal cluster. From Guss and Freeman.[116] (B) Ribbon drawing of immunoglobulin fold. This is a common structure in domains of the immunoglobulins and in many other extracellular proteins. Two layers of antiparallel β sheet are stacked face to face to form a flattened barrel. One disulfide bridge is always present and is represented as a thick rod. From J. Richardson.[117] (C) Five tandem fibronectin type III domains. These domains, which are found in the muscle protein titin as well as in fibronectin, resemble immunoglobulin domains but lack disulfide bridges. From Erickson.[117a] Figure courtesy of Harold P. Erickson.

interaction of polar groups from side chains. Thus, OH groups of serine or threonine residues, amide groups of asparagine and glutamine residues, etc., often fold back and hydrogen-bond to these ends.

Edges of β sheets can also serve as binding sites for other polar molecules. For example, substrates bind to an edge of a β sheet in the active sites of trypsin and other proteases (Chapter 12). Some proteins, e.g. the lectin shown in Fig. 2-15, form dimers by joining identical edges of a β sheet in antiparallel orientation.[112]

Cylinders and barrels. The twisted β sheets of proteins are often curved to form structures known as **β cylinders** or **β barrels** (Fig. 2-16).[113,114] Simple cylinders formed by parallel β strands form the backbones of the electron transport protein plastocyanin, the enzyme superoxide dismutase, the oxygen carrier hemerythin (Fig. 16-20), transporter proteins that carry hydrophobic ligands,[115] and the immunoglobulins in which each domain contains a β barrel.

The eight-stranded β cylinder of plastocyanin (Fig. 2-16A) is somewhat flattened and can also be regarded as a β sandwich.[116,118] However, the β barrel of triose phosphate isomerase (see Fig. 2-28) is surrounded by eight α helices which provide additional stability and a high symmetry. Bacterial outer membranes contain pores created by very large β cylinders within proteins called **porin**s.[119,120] The one shown in Fig. 8-20 has 16 strands.

A single β strand can also be wound into a cylinder with the hydrogen bonds running parallel to the helix axis. A right-handed **parallel β helix** of this type has been found in the bacterial enzyme **pectate lyase**.[121,122] The polypeptide chains of the 353-residue protein contain seven complete turns of about 22

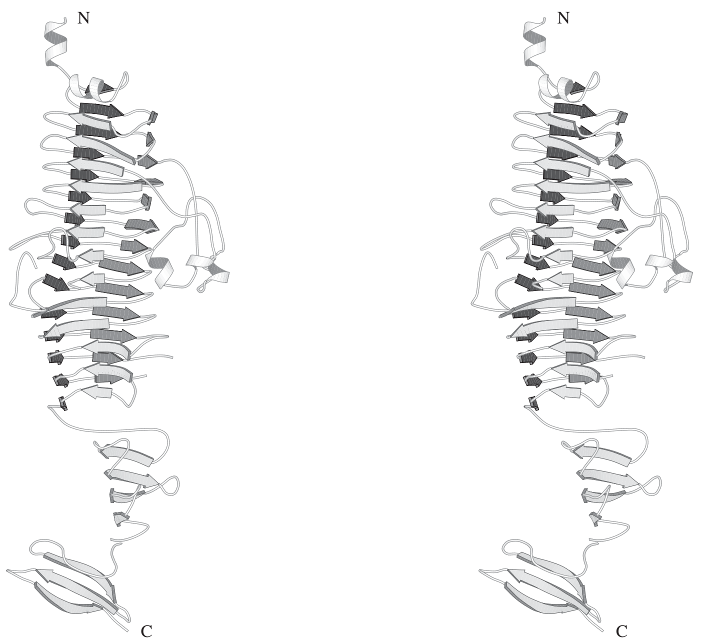


Figure 2-17 Wire model of the tailspike protein of bacteriophage P22 of *Salmonella*. Three of these fish-shaped molecules associate as a trimer to form the spike. From Steinbacher *et al*.[123]

residues each which form a cylindrical parallel β sheet. The cylinder is folded inward into roughly an L-shape with packed side chains filling the remaining inside space. The tailspikes of bacteriophage P22, a virus of *Salmonella,* are formed from 666-residue protein subunits which contain 13 turns of parallel β helix (Fig. 2-17).[123]

The tendency for a β sheet to fold into a cylinder is encouraged in antiparallel β structures by the existence of a common irregularity called the β **bulge**.[124,125] As illustrated in Fig. 2-18, a β bulge contains an extra residue inserted into one of the chains. In the second type of bulge shown in Fig. 2-18, the extra residue is glycine with torsion angles of about $\phi = 85°$, $\psi = 0°$, which are possible only for glycine. In the two β cylinders of trypsin, chymotrypsin, and elastase (Fig. 12-9), there are seven β bulges.

Beta structures are found in many small peptides. The hormone oxytocin (Fig. 2-4), the antibiotics gramicidin S (Fig. 2-4) and valinomycin (Fig. 8-22), and the mushroom peptide antamanide (Box 28-B) are among these. The cyclic structures of these compounds favor formation of antiparallel β strands with sharp turns at the ends. Polypeptide antibiotics that have alternating

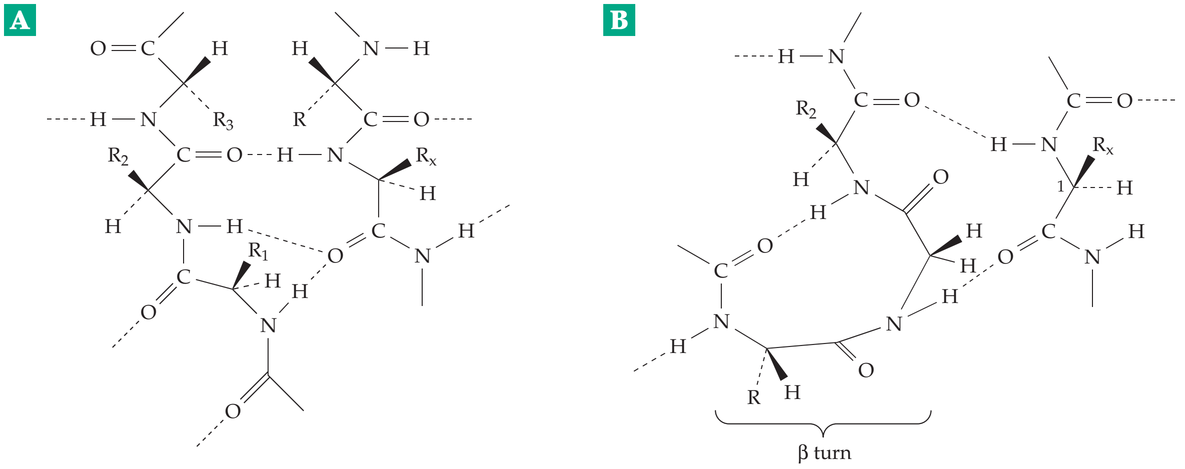


Figure 2-18 Typical β bulges in antiparallel pleated sheets. The residues R_1, R_2, and R_x identify the bulges. (A) A "classic" β bulge, in which ϕ_1 and ψ_1 are nearly those of an α helix while other torsion angles are approximately those of regular β structures. (B) The G1 bulge in which the first residue is glycine with $\phi_1 = 85°$, $\psi_1 = 0°$. It is attached to a type II β turn of which the glycine (labeled 1) is the third residue.

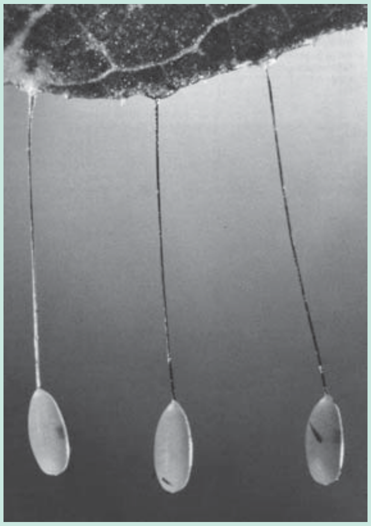
sequences of D and L residues can be coiled into a β helix[126,127] or a pair of polypeptide chains arranged in an antiparallel fashion can form a double stranded β helix with an 0.3 to 0.4-nm hydrophilic pore through the center. These peptide antibiotics appear to exert their antibacterial action by creating pores through cell membranes and allowing ions to pass through without control.

Beta propellers. Another major folding pattern is a circular array of four to eight "blades" that form a propeller-like structure. Each blade is a small, roughly triangular four-stranded antiparallel β sheet (See Figs. 11-7 and 15-23). Sequences that fold into these blades can often, but not always, be recognized as **WD repeats**. These are typically 44- to 60-residue sequences that have the sequence GH (Gly-His) about 11–24 residues from the N terminus and WD (Trp-Asp) at the C terminus.[127a,127b] This repeat sequence encodes the

BOX 2-B SILKS

The green lacewing fly *Chrysopa flava* lays its eggs on 1 cm silk stalks glued to the undersides of leaves. It has been proposed, as for other silks, that the peptide chains in the stalks are aligned perpendicular to the long dimension of the fiber and are folded back on themselves many times to form a β sheet with only ~8 residues between folds.[a] The chains of **silk fibroin**, the major protein of silkworm silk, contain 50 repeats of the sequence:[b]

GAGAGSGAAG(SGAGAG)$_8$Y


Courtesy of Ralph Buchsbaum

All of the alanine and serine side chains of this sequence presumably protrude on one side of a β sheet while the other side has only the hydrogen atoms of the glycine. This permits an efficient stacking of the sheets with interdigitation of the side chains of the alanine and serine residues of adjacent sheets.

However, DNA sequences of silk genes have revealed a greater complexity. The silk fibroin sequence suggests that at bends between β strands there are often S–S bridges and the crystalline β sheet domains are interspersed with 100–200 residue segments of amorphous protein[b–d] whose coiled chains can be stretched greatly. The protein from *Bombyx mori* consists of 350-kDa heavy chains linked by a disulfide bond to 25-kDa light chains.[e] A silkworm **housing silk** contains polyalanine sequences, e.g. (Ala)$_{10-14}$. These evidently form α-helical regions.[f] Silk molecules are synthesized and stored in glands as a concentrated solution of apparently globular molecules. The silk is extruded through a spinneret, whose diameter is ~10 μm, after which the silkworm stretches the silk, causing it to form a stiff fiber. Apparently the stretching of the folded polypeptide chains permits the β-sheet-forming sequences to find each other and form oriented crystalline regions.[d]

Spiders may produce silk from as many as seven different types of glands. Dragline silk, by which a spider itself may descend, is stronger than steel but because of the coiled amorphous regions may be stretched by 35%.[g–l] The most elastic silks are those of the catching spirals of orb-webs.[i,j,l] They can be stretched 200% and contain a variety of repeated sequences, including GPCC(X)$_n$. The latter is similar to sequences in elastin (Section 4), and is able to form type II β bends, which may have proline in position 2 and must have glycine in position 3 (Fig. 2-24). A chain of repeated β bend motifs can form a flexible spring, a β spiral as proposed for elastin. Genes for spider silk have been cloned and are being used to engineer new proteins with commercial uses, e.g., to help anchor cells in regenerating body tissues.[d,h,j]

---

[a] Geddes, A. J., Parker, K. D., Atkins, E. D. T., and Beighton, E. (1968) *J. Mol. Biol.* **32**, 343–368

[b] Garel, J.-P. (1982) *Trends Biochem. Sci.* **7**, 105–108

[c] Vollrath, R. (1992) *Sci. Am.* **266**(Mar), 70–76

[d] Calvert, P. (1998) *Nature (London)* **393**, 309–311

[e] Mori, K., Tanaka, K., Kikuchi, Y., Waga, M., Waga, S., and Mizuno, S. (1995) *J. Mol. Biol.* **251**, 217–228

[f] van Beek, J. D., Beaulieu, L., Schäfer, H., Demura, M., Asakura, T., and Meier, B. H. (2000) *Nature (London)* **405**, 1077–1079

[g] Simmons, A. H., Michal, C. A., and Jelinski, L. W. (1996) *Science* **271**, 84–87

[h] Hinman, M. B., and Lewis, R. V. (1992) *J. Biol. Chem.* **267**, 19320–19324

[i] Hayashi, C. Y., and Lewis, R. V. (1998) *J. Mol. Biol.* **275**, 773–784

[j] Guerette, P. A., Ginzinger, D. G., Weber, B. H. F., and Gosline, J. M. (1996) *Science* **272**, 112–115

[k] Spek, E. J., Wu, H.-C., and Kallenbach, N. R. (1997) *J. Am. Chem. Soc.* **119**, 5053–5054

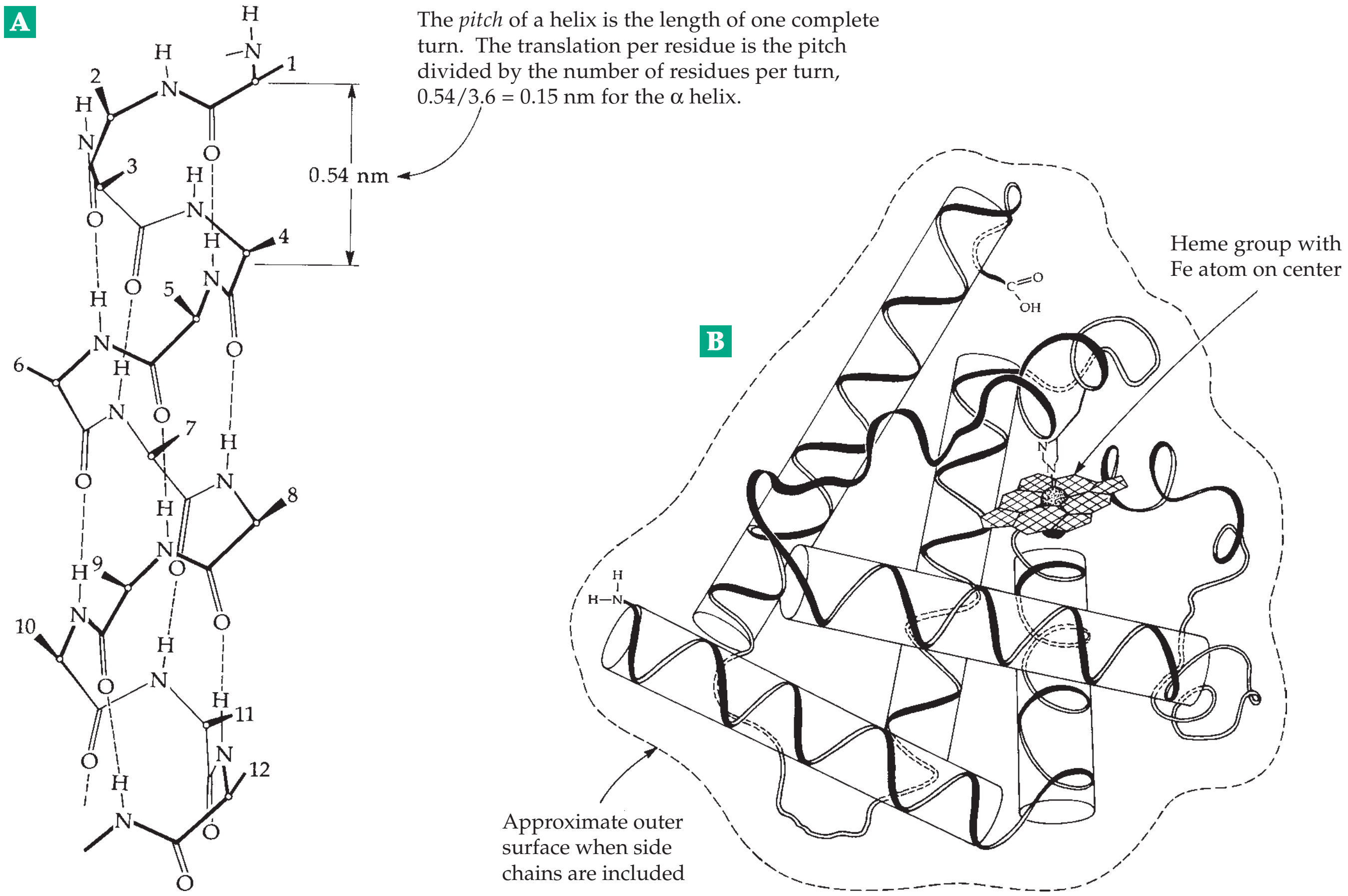
[l] Hayashi, C. Y., and Lewis, R. V. (2000) *Science* **287**, 1477–1479

fourth β strand of one blade followed by the first three strands of the next blade. This overlap snaps the blades together to form the propeller. At least one tight turn is present in each blade. A seven-bladed propeller is present in the β subunits of the regulatory GTP-hydrolyzing **G proteins**, which couple extracellular signals to intracellular enzymes and ion channels (Fig. 11-7). A similar β propeller is present in **clathrin**, which forms cage-like enclosures around endocytic vesicles (Chapter 8).[127c] Six-bladed propellers are predicted to occur in many extracellular proteins.[127d] A β propeller binds the coenzyme PQQ in bacterial dehydrogenases (Fig. 15-23).

3. Helices

The **alpha helix** represents the second major structural element of soluble proteins[108,128] and is also found in many fibrous proteins, including those of muscle and hair. In the α helix both ϕ and ψ are about −50 to −60° and except at the N-terminal helix end, *each NH is hydrogen bonded to the fourth C=O further down the chain*. All of the N–H and C=O groups of the peptide linkages lie roughly parallel to the helix axis; the N–H groups point toward the N terminus of the chain and the carbonyl groups toward the C terminus (Fig. 2-19A).

The number of amino acid units per turn of the helix is ~3.6, with five turns of the helix containing 18 residues. The **pitch** (repeat distance) of the helix, which can be determined experimentally from X-ray diffraction data, is 0.54 nm. Polar coordinates for the α helix have been tabulated.[130] With L-amino acids, the right-handed helix, is more stable than the left-handed helix which has so far not been found in proteins. Frequently, however, a few residues have the ϕ, ψ angles of this helix. The ϕ, ψ angles of the α helix are given in Table 2-3 as −57°, −47°, but are much more variable in real helices. In erythrocruorin, for which an accurate structure determination has been made,[131]



The *pitch* of a helix is the length of one complete turn. The translation per residue is the pitch divided by the number of residues per turn, 0.54/3.6 = 0.15 nm for the α helix.

Heme group with Fe atom on center

Approximate outer surface when side chains are included

Figure 2-19 The α helix. (A) The right-handed α helix with vertical hydrogen bonds indicated by dotted lines. The positions of the amino acid side chains are indicated by the numbers. (B) The conformation of the peptide backbone of myoglobin.[129] Five long α helices are indicated as rods. Several shorter helices can also be seen. The overall size of the molecule is approximately 4.4 × 4.4 × 2.5 nm.

the average values are $\phi = -64°$, $\psi = -40°$. More important is the observation that $\phi + \psi = 104°$ within about ± 10° for most residues in helices of this and other proteins.[132,133] The deviation from the ideal dimensions results in part from the hydrogen bonding of water molecules or polar protein side chains to the carbonyl oxygen atoms in the helix.[131]

Helix formation is spontaneous for peptides as short as 13 residues in water.[134,135] Although the difference in thermodynamic stability between an α helix and an unfolded **"random coil"** conformation is small, poly-L-alanine peptides form helices in water. Glycine destabilizes helices, presumably because of the increased entropy of the unfolded chain which results from the wider range of the conformational angles ϕ and ψ for glycyl residues. Proline destabilizes helices even more because its restricted ϕ and ψ angles cause the helix to be kinked.[136,137] However, prolyl residues are often present at ends of helices. Other amino acids all fit into helices but may stabilize or destabilize the helix depending upon immediately neighboring groups.[138–140] Bulky side chain groups with a low dielectric constant stabilize helices by strengthening the hydrogen bonds within the helices.[141]

Helices of smaller and larger diameter than that of the α helix are possible. The most important is the **3_{10} helix** (or 3.0_{10} helix), which has exactly three residues per turn.[140,142–144] Each NH forms a hydrogen bond to the third C=O on down the chain; thus, the 3_{10} helix is tighter than the α helix. Although long 3_{10} helices are seldom found in proteins, a single turn of this tighter helix frequently occurs as a "defect" at the end of an α helix. A polymer of α-aminoisobutyric acid forms long 3_{10} helices because the α-dimethyl side chains constrain ϕ and ψ to appropriate values.[145] Short helical peptide chains in water may exist as a mixture of α and 3_{10} forms in equilibrium.[146] The **π helix**, with 4.4 residues per turn is of a larger diameter than the α helix and has only been found in proteins as a single turn, usually at a C terminus.[132,132a]

Properties of helices. The dipoles of the backbone amide linkages of an α helix are all oriented in the same direction. The positive end of each dipole is associated through hydrogen bonding with the negative end of another. This leaves three partial positive charges at the N terminus of the helix and three partial

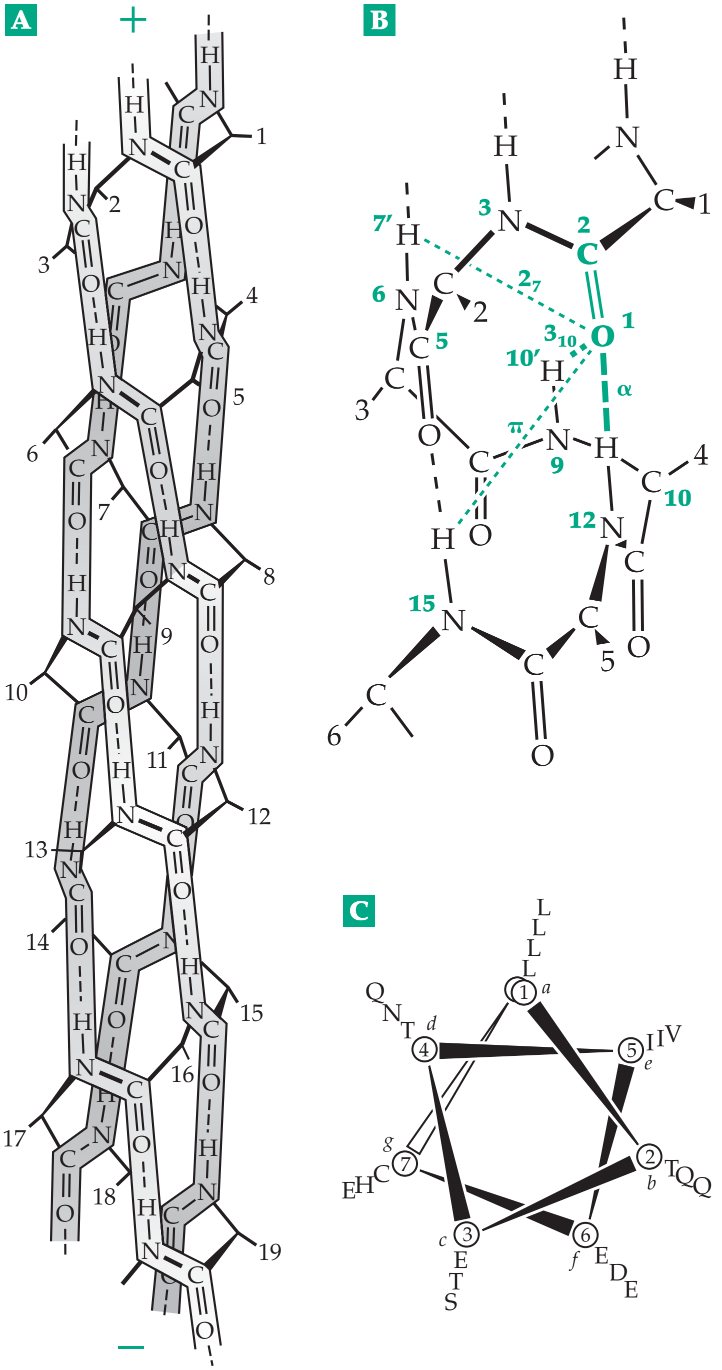


Figure 2-20 (A) The α helix showing the three chains of interconnected amide groups and hydrogen bonds with partial net positive and negative charges at the ends. These chains run across the turns of the helical polypeptide backbone. (B) Scheme illustrating hydrogen-bonding pattern for 2_7 ribbon, 3_{10} helix, α helix, and π helix. (C) The α helix represented as a helical wheel. Imagine viewing the helix from the N-terminal end of the segment with the lines corresponding to the backbone of the peptide. Residues, which are designated by single letters, are spaced 100° apart since there are 3.6 residues/turn. The peptide shown is a 22-residue sequence from a "leucine-zipper" domain of a protein that participates in gene regulation. From Fathallah-Shaykh *et al.*[145a]

negative charges at the C terminus and creates a “macrodipole.” It has been estimated that each end of the helix carries one-half an elementary unit of charge.[147–148b] However, this may be an overestimate. The partial charges are connected by three chains of hydrogen bonds that run across the turns of the helix as is indicated in Fig. 2-20A. These chains are polarized and also possess a large **hyperpolarizability**,[149] i.e., the polarization of the helix increases more rapidly than in direct proportion to an applied electrical field.

In a 3_{10} helix, there are just two chains of hydrogen bonds across the polypeptide chain and running the length of the helix. The characteristic hydrogen bonding of the α helix and that of the 3_{10} helix are often both present within a helix. Irregularities with 3_{10} type hydrogen bonds, arising from interactions with amino acid side chains or with solvent molecules, may cause a helix to be kinked or curved.[132] Side chains of polar residues, including those of Asn, Asp, Thr, and Ser (and less often Glu, Gln, or His), frequently fold back and hydrogen bond to the NH and CO groups that carry the partial charges at the helix ends (Fig. 2-20). The side chain of the third residue in the helix may also hydrogen bond to the NH of the first residue.[133,150–157] The hydrogen bonding of a negatively charged side chain group to the N-terminal end of the helix or of a positively charged group to the C-terminal end provides an **N-cap** or a **C-cap** which helps to stabilize the helix by strengthening its hydrogen bonds.[158] However, the most frequent residue at the C-terminal end is glycine.[151,155,159] Helices often point toward active sites of enzymes and interactions of the helix dipoles with substrates undergoing reaction may be important to the mechanism of action of these catalysts.[148,149,160]

Stacking of helices in proteins. Many proteins are made up almost entirely of α helices. One of these, **myoglobin**, was the first protein for which the complete three-dimensional structure was worked out by X-ray crystallography.[129] Myoglobin is a small oxygen-carrying protein of muscle. Its structure is closely related to that of hemoglobin of blood. Its 153 amino acid residues are arranged in eight different α-helical segments containing from 7 to 26 residues each. These rodlike helices are stacked together in an irregular fashion as shown in Fig. 2-19. Serum albumin (Box 2-A) has 28 helices organized into three homologous domains. In contrast, the filamentous bacterial viruses have protein coats made up of small subunits, each coiled as a single α helix. These are packed in a regular manner to form the rodlike virus coats (Fig. 7-7).

Coiled coils. In a large family of proteins, two right-handed α helices are coiled around each other in a left-handed **superhelix** (Fig. 2-21).[161–167] This **coiled coil** structure was first suggested by Francis Crick[166] to account for the fact that the X-ray diffraction pattern of the **keratins** of skin and hair indicated a pitch of 0.51 nm rather than the 0.54 nm expected for an α helix.

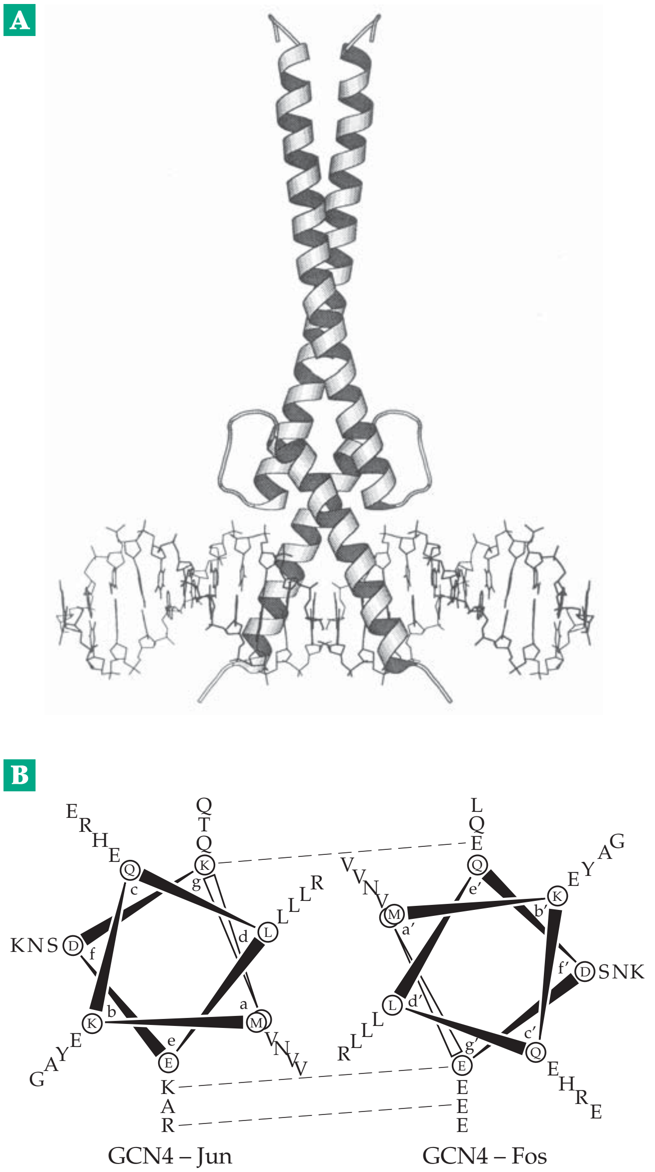


Figure 2-21 (A) Ribbon drawing of the transcription factor called Max in a complex with DNA. The C termini of the peptide chains are at the top.[169] Courtesy of S. K. Burley. (B) Helical wheel representation of residues 2–31 of the coiled coil portion of the leucine zipper (residues 249–281) of the related transcription factor GCN4 from yeast. The view is from the N terminus and the residues in the first two turns are circled. Heptad positions are labeled *a*–*g*. Leucine side chains at positions *d* interact with residues *d*′ and *e*′ of the second subunit which is parallel to the first. However, several residues were altered to give a coiled coil that mimics the structure of the well-known heterodimeric oncoproteins Fos and Jun (see Chapter 11). This dimer is stabilized by ion pairs which are connected by dashed lines. See John et al.[172]

Crick suggested that two supercoiled α helices inclined at an angle of 20° to each other would produce the apparent shortening in the helix pitch and would also permit nonpolar side chains from one strand to fit into gaps in the surface of the adjacent strand, a **knobs-into-holes** bonding arrangement. Helical strands tend to coil into ropes because a favorable interstrand contact can be repeated along the length of the strands only if the strands coil about each other. A coiled coil can often be recognized by **heptad repeats**:

$$(\mathbf{a}\text{-b-c-}\mathbf{d}\text{-e-f-}\mathbf{g})_n$$
$$1\ 2\ 3\ 4\ 5\ 6\ 7$$

Here, residues *a*, *d*, and *g* (1, 4, and 7) often carry nonpolar side chains. These come together in the coiled coil as is illustrated in the **helical wheel** representations in Fig. 2-21B and provide a longitudinal hydrophobic strip along the helix.[167] Charged groups are often present in other locations and in such a way as to provide electrostatic stabilization through interactions between residues within a single α helix[170-171a] or between the pair of helices. The latter type of interaction also determines whether the coiled-coil consists of a parallel or antiparallel pair and whether the two helices are of identical or of differing amino acid sequence.[171b]

Recent attention has been focused on a DNA-binding structure called the **leucine zipper**. A pair of parallel α helices are joined as a coiled coil at one end but flare out at the other end to bind to DNA. In the yeast transcription factor GCN4, whose three-dimensional structure has been determined to high resolution (Fig. 2-21B),[172] the *d* position of the coiled coil is occupied by leucine and the *e* and *g* positions are often occupied by charged groups that form stabilizing ion pairs. Residues at positions *b*, *c*, and *f* are generally on the outside and exposed to solvent.[168,171,173] The coiled coil flares out at the C-terminal ends and carries DNA-binding groups. The structure of a related transcription factor is shown in Fig. 2-21A.[169]

The muscle proteins **myosin**[174] and **tropomyosin** also both consist of pairs of identical chains oriented in the same direction. The two 284-residue tropomyosin chains each contain 40 heptads and are linked by a single disulfide bridge. X-ray crystallographic studies[175,176] and electron microscopy[177] show that the molecule is a rod of 2.0 nm diameter and 41 nm length, the dimensions expected for the coiled coil. However, as with other "regular" protein structures, there are some irregularities. Myosin chains (Chapter 19) contain 156 heptads.

Another group of proteins have parallel coiled coils flanked by nonhelical domains in subunits that associate as filaments. These include the **keratins** of skin as well as the intermediate fiber proteins of the cytoskeleton (Chapter 7).[164,178] Natural coiled coils often have a parallel orientation, but synthetic peptides have been designed to form antiparallel coiled coils.[179,179a]

Helix bundles. A third peptide chain can be added to a coiled coil to form a triple-stranded bundle.[180–183] An example is the glycoprotein **laminin** found in basement membranes. It consists of three peptide chains which, for ~600 residues at their C-terminal ends, form a three-stranded coil with heptad repeats.[182,184] Numerous proteins are folded into four helical segments that associate as **four-helix bundles** (Fig. 2-22).[185–188] These include electron carriers, hormones, and structural proteins. The four-helix bundle not only is a simple packing arrangement, but also allows interactions between the + and – ends of the macro-dipoles of the helices.

Membranes contain many largely α-helical proteins. Cell surface receptors often appear to have one, two, or several membrane-spanning helices (see Chapter 8). The single peptide chain of the bacterial light-operated ion pump **bacteriorhodopsin** (Fig. 23-45) folds back upon itself to form seven helical rods just long enough to span the bacterial membrane in which it functions.[189] Photosynthetic reaction centers contain an α helix bundle which is formed from two different protein subunits (Fig. 23-31).[190] A recently discovered α,α **barrel** contains 12 helices. Six parallel helices form an inner barrel and 6 helices antiparallel to the first 6 form an outer layer (see Fig. 2-29).[191–193]

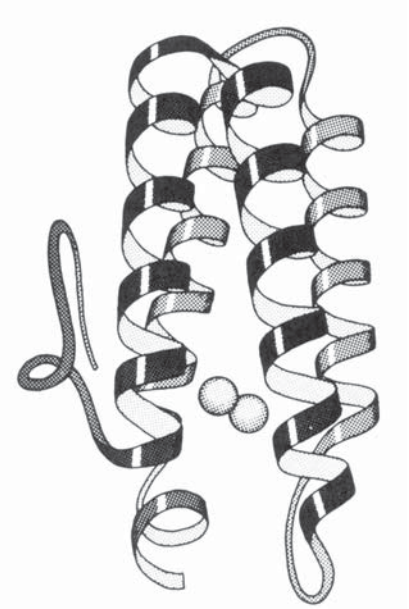


Figure 2-22 Ribbon drawing of an up-and-down four-helix bundle in **myohemerythrin**. The two spheres represent the two iron atoms which carry an O_2 molecule. They are coordinated by histidine and aspartate side chains. Courtesy of J. Richardson.[117]

4. Polyglycine II and Collagen

In a second form of polyglycine each amino acid residue is rotated 120° from the preceding one about a 3-fold screw axis as is shown in the end view of Fig. 2-23A. The angle ψ is about 150° while ϕ is about –80° for each residue. The distance along the axis is 0.31 nm/residue and the repeat distance is 0.93 nm. The molecules can coil into either a right-handed or a left-handed helix. In this structure, the N–H and C=O groups protrude perpendicular to the axis of the helix and, as in the β structure, can form H-bonds with adjacent chains.

Poly (L-proline) assumes a similar helical structure. However, the presence of the bulky side chain groups induces a *left-handed* helical twist. **Collagen**, the principal protein of connective tissue, basement membranes, and other structures, is the most abundant protein in the animal body. Its fundamental unit of structure is a triple helix of overall dimensions 1.4 x 300 nm which resembles the structure of polyglycine II but contains only three chains.[195] The left-handed helices of the individual chains are further wound into a right-handed superhelix (Fig. 2-23B,C). Collagen contains 33% glycine and 21% (proline + hydroxyprolines). The reason for the high glycine content is that bulky side chain groups cannot fit inside the triple helix. A hydrogen-bonded cylinder of hydration surrounds each triple helix and is anchored by hydrogen bonds to the –OH groups of the hydroxyproline residues.[196] Sequences of several collagen chains have been established. One of these contains ~1050 residues, ~1020 of which consist of ~340 triplets of sequence GlyXY where Y is usually proline and X is often hydroxyproline (Hyp).[197,198] The commonest triplet is Gly-Hyp-Pro.

There are several types of collagen. In one type, two identical chains of one kind are coiled together with a third dissimilar chain to form the triple helix. Several of these triple helices associate to form 8-nm microfibrils (Fig 2-23D).[199] Once synthesized, collagen is extensively modified and crosslinked. See Chapter 8.

Collagen-like triple helices also occur within other proteins. One of these is protein C1q, a component of the **complement system** of blood (Chapter 31). This protein interacts with antibodies to trigger a major aspect of the immune response. C1q has six subunits, each made up of three different polypeptide chains of about 200 residues apiece. Beginning a few residues from the N termini, there are over 80 residues in each chain with collagen-like sequences. The three chains apparently form a triple helix within each subunit. However, the C-terminal portions are globular in nature.[200] Collagen-like tails also are present on some forms of the enzyme acetylcholinesterase (see Chapter 12C,10). The **extensins** of plant cell walls contain 4-hydroxyproline and evidently have a structure related to that of collagen.[201,202] Shorter 4- to 8-residue segments of left-handed polyproline helix are found in many proteins.[203]

5. Turns and Bends

To form a globular protein, a polypeptide chain must repeatedly fold back on itself. The turns or bends by which this is accomplished can be regarded as a third major secondary structural element in proteins. Turns often have precise structures, a few of which are illustrated in Fig. 2-24. As components of the loops of polypeptide chains in active sites, turns have a special importance for the functioning of enzymes and other proteins. In addition, tight turns are often sites for modification of proteins after their initial synthesis (Section F).

The **β turn** (Fig. 2-24) is often found in hairpin or reverse turns at the edges of β sheets (Fig. 2-11) and at other locations.[204–212] If all four residues that contribute to β bends are counted, they constitute about one-third of the amino acid residues in most proteins.[124] In many β turns, the C=O of the first residue hydrogen bonds to the NH of the fourth residue. This hydrogen bond may be part of the hydrogen bond network of a β pleated sheet. The peptide unit between α-carbon atoms 2 and 3 of the turn is perpendicular to the sheet. There are two possibilities for the orientation of this peptide unit. In a *type I turn*, the C=O is down when the turn is viewed as in Fig. 2-24, while the side chains of residues 2 and 3 point upwards or outward on the opposite side of the bend. In a *type II turn*, the C=O is up and the NH down. Residue 3 is always glycine in a type II turn because the side chain would collide with the C=O group if any other amino acid were present. As is seen in Fig. 2-24, a *trans*-proline can fit at position 2 in a type II turn[206] as well as in type I turns. A *cis*-proline residue can fit at position 2 or position 3 in a type I β turn.[137,213]

The *type III β turn* is similar to a type I turn but has the ϕ, ψ angles of a 3_{10} helix and the two chains emerging from the turn are not as nearly parallel as they are in type I turns. Beta turns of the less common types I′, II′, and III′ have a left-handed twist. As can be seen in Fig. 2-24, this permits a better match to the twist of strands in a β sheet. Unless glycine is present, these bends are less stable because of steric hindrance.[214,215] Polar side chain groups such as those of aspartate or asparagine often form hydrogen bonds to the central peptide units of β turns.[214]

The tight γ turn[215] and the proline-containing β turn shown in Fig. 2-24 are thought to be major components of the secondary structure of **elastin**.[216–218] This stretchable polymer, which consists largely of nonpolar amino acids, is the most abundant protein of the elastic fibers of skin, lungs, and arteries. The

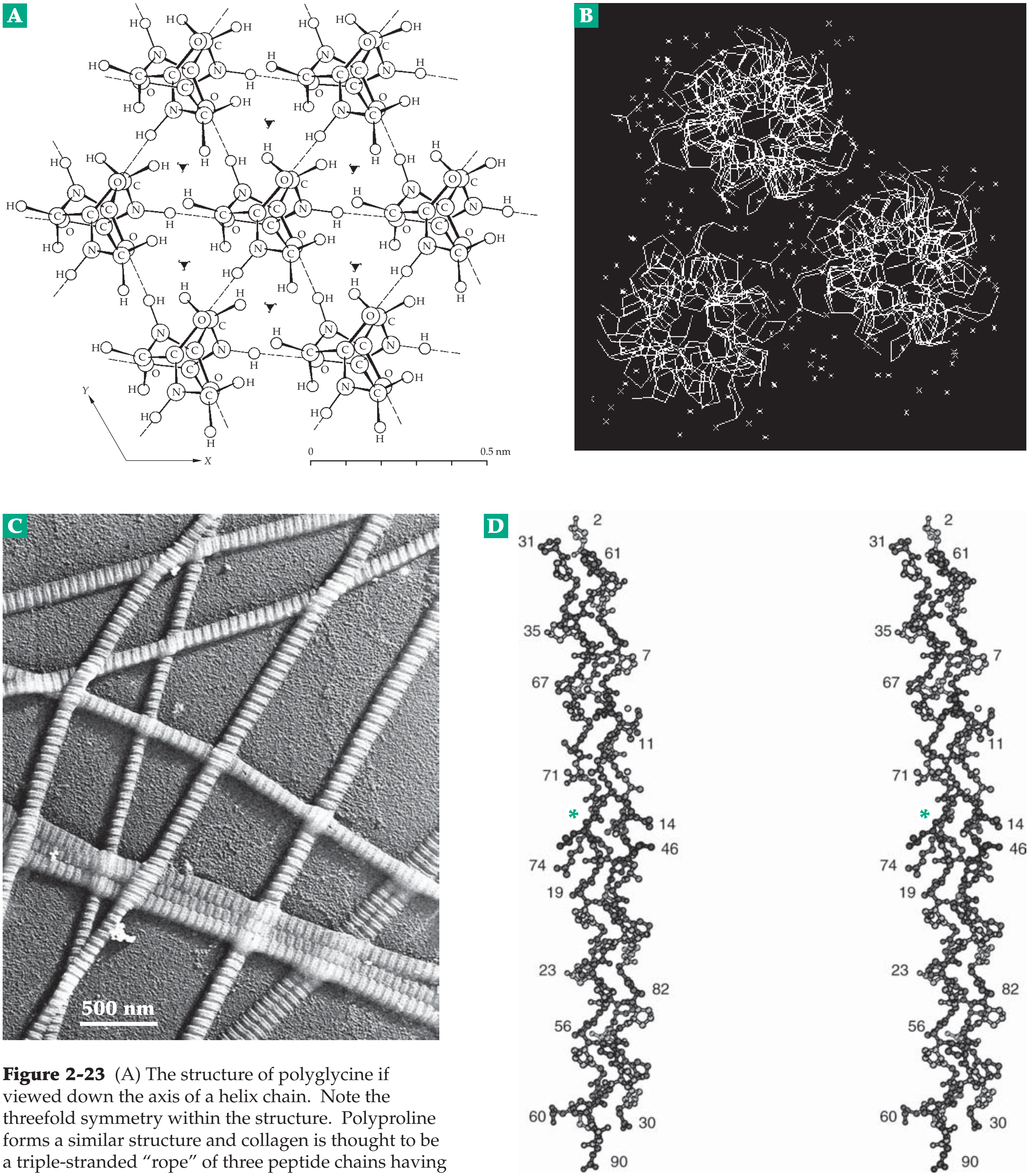
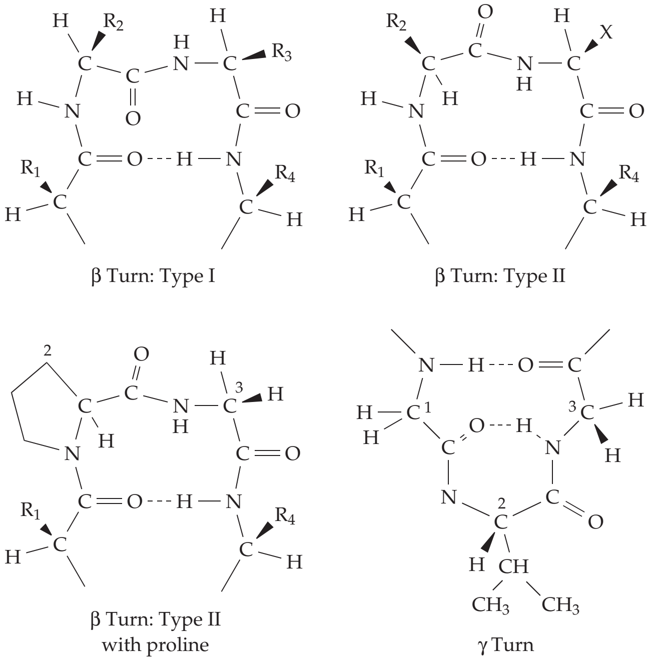


Figure 2-23 (A) The structure of polyglycine if viewed down the axis of a helix chain. Note the threefold symmetry within the structure. Polyproline forms a similar structure and collagen is thought to be a triple-stranded "rope" of three peptide chains having essentially the same structure but in addition containing a right-handed supercoil. This is illustrated by the crystal structure of a collagen-like peptide shown in B and C. (B) View similar to that in (A), but showing how three chains form the triple-stranded ropes separated by a cylinder of hydration. The structure is shown as a wire model. The x's are water molecules. They form an extensive network of H-bonds to one another and to peptide groups. From Bella *et al.*[194] C) Electron micrograph of collagen fibrils, each of which consists of many triple-helical units. Deposited from suspension and shadow cast with chromium. Courtesy of Jerome Gross, M.D. (D) Stereoscopic view of a collagen-like model peptide. Each of the three parallel 30-residue chains contains a (Pro-Hyp-Gly)$_3$ "cap" at each end and the 12-residue sequence Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-Pro-Hyp-Gly in the center. This sequence is found in human type III collagen and includes a site (green asterisk) of known mutations (see Chapter 8). From Kramer *et al.*[194a] Courtesy of Helen Berman.


β Turn: Type I

β Turn: Type II

β Turn: Type II
with proline

γ Turn

Figure 2-24 Tight turns found in polypeptide chains. Two types of β turn are shown. A third variant, the type III or 3_{10} turn resembles the type I turn but has the ϕ, ψ angles of a 3_{10} helix. Type II β turns containing proline and tighter γ turns are thought to be major structural components of elastin. Another β turn, lacking the hydrogen bond has a *cis*-proline residue at position 3.

70-kDa chains are thought to have 70- to 75-residue regions in which the polypeptide folds back on itself repeatedly with a large number of bends in a broad left-handed **β spiral**.[216,218a] The consensus sequence VPGVP, which tends to form a type II β turn with proline in position 2 (Fig. 2-24) is present in long tandem repeats e.g., $(VPGVG)_{11}$. These extensible regions alternate with short α helices which are crosslinked to other chains. Similar structures are present in silks (Box 2-B) and in proteins of wheat gluten.[217a,218b]

Besides hairpin turns and broader U-turns, a protein chain may turn out and fold back to reenter a β sheet from the opposite side. Such **crossover connections**, which are necessarily quite long, often contain helices. Like turns, crossover connections have a handedness and are nearly always right-handed (Fig. 2-25).[117,219] Most proteins also contain poorly organized loops on their surfaces. Despite their random appearance, these loops may be critical for the functioning of a protein.[220] In spite of the complexity of the folding patterns, peptide chains are rarely found to be knotted.[221]

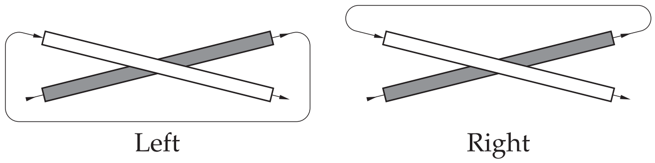


Figure 2-25 Right- and left-handed crossover connections in proteins. These connections are nearly always right-handed. The broad arrows represent β strands. The crossover often contains a helix. Units of two adjacent β strands (βαβ units) with an α helix between are found frequently in globular proteins.

6. Domains, Subunits, and Interfaces

Many proteins are organized into tightly folded globular "domains" consisting of 50–150 amino acid residues.[117,222–227] Smaller proteins may have 2 or 3 domains but large proteins may have more. For example, the immunoglobulin "heavy chains" (Chapter 31) have four or five domains similar to that shown in Fig. 2-16B. The enormous 3000-kDa muscle protein **titin** contains 260 domains, many of which are of the immunoglobulin typeFig. 2-16,C.[228,117a] In most proteins the polypeptide chain folds to form one domain, then passes through a "hinge" to the next. In others, the C terminus or the N terminus of the polypeptide folds back across two or more domains as a kind of "strap" that helps to hold the protein together. Even when a protein contains only one domain, it often consists of two distinct lobes with a cleft between them. Many proteins, e.g. hemoglobin (Fig. 7-25), consist of subunits about the size of the globular domains in larger proteins.

Structural domains of proteins are sometimes encoded by a single coding segment of DNA i.e., by a single exon in a split gene. Domains of this type may have served as **evolutionarily mobile modules** that have spread to new proteins and multiplied during evolution. For example, the immunoglobulin structural domain is found not only in antibodies but also in a variety of cell surface proteins.[229–232]

Whether we deal with domains connected by a flexible hinge or with subunits, there are **interfaces** between the different parts of the protein. These interfaces are often formed largely of nonpolar groups. However, they frequently contain a small number of hydrogen bonds that bridge between one domain and another or between one subunit and another. In the case of hemoglobin, important changes occur in this hydrogen bonding and movement occurs along one of the interfaces between two subunits. Likewise the active sites of enzymes are often located at interfaces between domains. During catalysis, movement and reorganization of the hydrogen bonds and side chain packing in the interfaces may take place.[232a]

7. Packing of Side Chains

In Fig. 2-19B myoglobin is pictured as a cluster of α-helical rods surrounding the heme core. This picture is incomplete because the space between the rods and inside the molecule is tightly packed with amino acid side chains almost all of which are hydrophobic. The same is true for the β barrels of Fig. 2-16, which are filled largely with nonpolar side chains. As the structures of more and more proteins have been determined, a consistent pattern has emerged. Within the interior of proteins the side chain groups are packed together remarkably well.[99,100,222,233–236] Although occasional holes are present, they are often filled by water molecules.[99,237a] The **packing density**, the volume enclosed by the van der Waals envelope divided by the total volume, is ~0.75 for the interior of the lysozyme and ribonuclease molecules compared with the theoretical value of 0.74 for close-packed spheres. However, regions with many hydrogen bonds may be less tightly packed.[238]

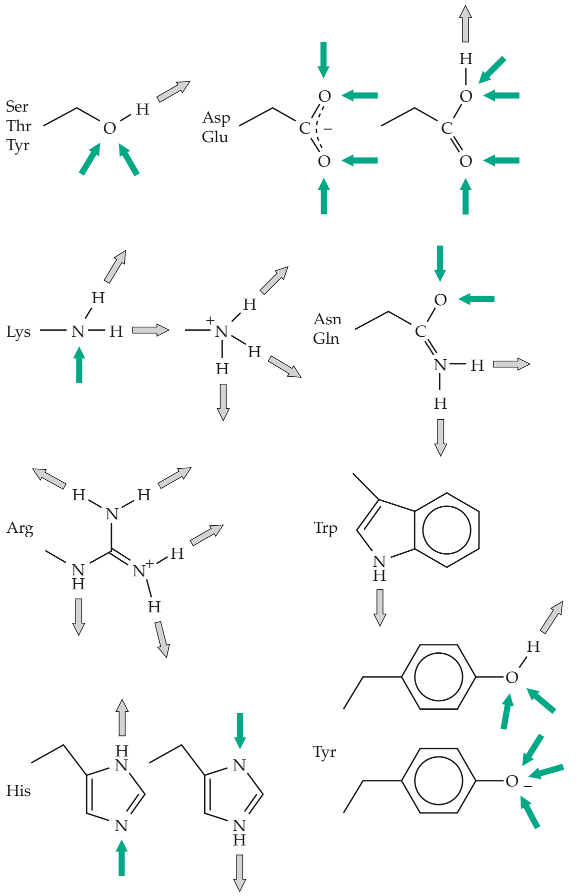


Figure 2-26 Some of the possibilities for hydrogen bonding of side chain groups in proteins. Oxygen atoms can and frequently do form up to three hydrogen bonds at once. Open arrows point *from* H-atoms and *toward* electron donor pairs.

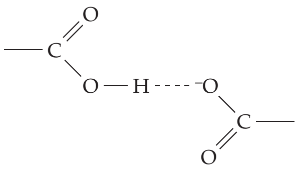
The interiors of proteins often contain large numbers of aromatic side chains which are frequently associated as pairs or as **aromatic clusters**.[239,240] For example, see the structure of plastocyanin in Fig. 2-16A. Rings may lie perpendicular one to another or be "stacked" face to face but offset. Oxygen atoms often lie in contact with the edges of aromatic rings.[241] The planar guanidinium groups of arginine side chains often stack against aromatic rings and amide groups may sometimes do the same.[242] It has been suggested that both aromatic: aromatic and aromatic:oxygen interactions may be associated with additional stabilization of the protein by ~4- to 8- kJ/mol. Tyrosine side chain –OH groups often stabilize ends of β strands by forming H-bonds to backbone atoms.[243]

Most polar groups are on the surfaces of proteins, and those that are not are almost always hydrogen bonded to other groups in the interior.[244] While most nonpolar groups are inside proteins, they are also present in the outer surfaces where they are often clustered into **hydrophobic regions** or "**patches**." The latter may be sites of interaction with other proteins or with lipid portions of membranes.

8. The Network of Internal Hydrogen Bonds

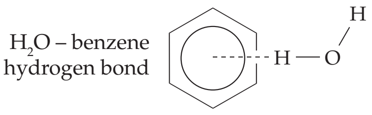
The fact that nonpolar groups tend to be buried in the interior of proteins suggests that the inside of a protein might be a flexible blob of oily material. In fact, the nonpolar groups tend to be densely packed and aromatic rings often impart considerable rigidity to the hydrophobic cores. Buried polar groups, which are invariably hydrogen bonded to other side chain groups, to amide groups of the polypeptide backbone,[244–247] or to buried water molecules, form a well-defined internal network. When a series of closely related proteins, for example, those having the same function in several different species, are compared, the hydrogen bonded network is often nearly invariant. This suggests a functional significance. The hydrogen bonding possibilities for some of the side chain groups in proteins are indicated in Fig. 2-26.

Charged side chains are sometimes buried in the interior of proteins, usually together with an ion of the opposite charge to form a **hydrogen-bonded ion pair**.[248,248a,248b] However, there is sometimes a single charge that is "neutralized" by the interaction with dipoles of polar groups.[217] Sometimes an undissociated carboxyl group is hydrogen bonded to a carboxylate ion.[249]



A buried carboxyl group of this type will display a pK_a value far higher than the normal value.

Arginine side chains are large and able to form multiple hydrogen bonds[247,250] as well as salt linkages[251] to different parts of a folded peptide. Cations of arginine and lysine are never buried unless in ion pairs. Protons from strong acids have long been known to bind to centers of benzene rings and water can form weak hydrogen bonds to the centers of aromatic rings.[252,253,253a] Such bonds also occur within proteins and often involve the binding of guanidinium groups or of inorganic cations to indole rings of tryptophans.[253b,253c] Protonated imidazole groups may also bind to aromatic rings.[253d]



Another important aspect of the structure of proteins is the presence of **hydrogen-bonded water molecules** in pockets and cracks. These molecules, as well as a much larger number of water molecules bound at the outer surface, are clearly visible from X-ray studies. They often occur singly, bonded to the ends of amide groups, especially to the carbonyl ends. Internal bends of the peptide chain are almost always hydrated.[254] These water molecules often make two or more hydrogen bonds to different parts of the protein or to other water molecules. Clusters of water molecules,[49,255,255a] sometimes in the form of pentagonal rings,[256] are often present. NMR spectroscopy has shown that water molecules bound to protein *surfaces* exchange rapidly with the bulk water in which the protein is dissolved; the "residence time" on the protein is typically less than a nanosecond. *Interior* water molecules have much longer residence times of 10^{-2} to 10^{-8} s for a small protein.[257] They may be regarded as part of the protein structure.

E. Folding Patterns and Protein Families

Proteins are folded in many ways. We have already considered several simple patterns: the antiparallel β cylinder (Fig. 2-16), the 2-helix coiled coil (Fig. 2-21) and the 3- and 4-helix bundles (Fig. 2-22). Another simple motif that has been found repeatedly is the **helix–turn–helix** or **helix–loop–helix** in which two helices at variable angles, one to another and with a turn or short loop between them, form a structural unit. DNA-binding **repressors** and **transcription factors** (see Fig. 2-21 and also Chapter 5) often contain this motif as do many Ca^{2+}-binding proteins. Proteins containing 3–6 helical segments, often fold into a roughly polyhedral shape.[258,259] An example is myoglobin (Fig. 2-19B).

1. Complex Folding Patterns

Proteins often contain elements of both α and β structure. One of the first of the complex folding motifs to be recognized was a nucleotide-binding domain identified by Rossmann and associates.[260–262] This Rossman fold contains six parallel β strands which alternate with six helices. The result of the parallel β structure is that the helices are also parallel and that their amino-terminal ends, which carry partial positive charges, are aligned in approximately the same directions. The positive end of one of the helices lies behind the negatively charged phosphate groups[148] which characteristically bind at the edge of the sheet containing the C termini of the β strands. This can be seen in Fig. 2-13, which shows binding of the coenzyme NAD to the nucleotide binding domain of **glyceraldehyde phosphate dehydrogenase**. Similar nucleotide-binding domains are found in many other dehydrogenases whose members constitute a **protein family**.

Figure 2-27 depicts **topology diagrams** for the Rossman fold and for two related families of proteins. These families bind the nucleotides called GTP and ATP, respectively. Both are structural relatives of NAD. A major part of the structure of all of the proteins in these families consist of β–α units, each one containing a β strand followed by a helix. They are

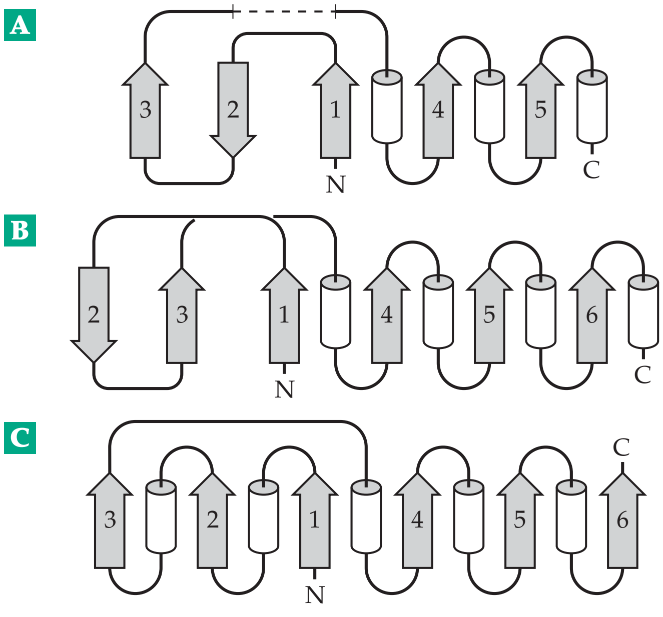


Figure 2-27 Topologies of the folds of three families of nucleotide binding α/β proteins. Cylinders represent α helices and arrows β strands. (A) The ATPase fold for the clathrin-uncoating ATPase; (B) The G-protein fold that binds GTP and is found in *ras* proteins; (C) The Rossmann fold that binds NAD in several dehydrogenases. From Brändén.[262]

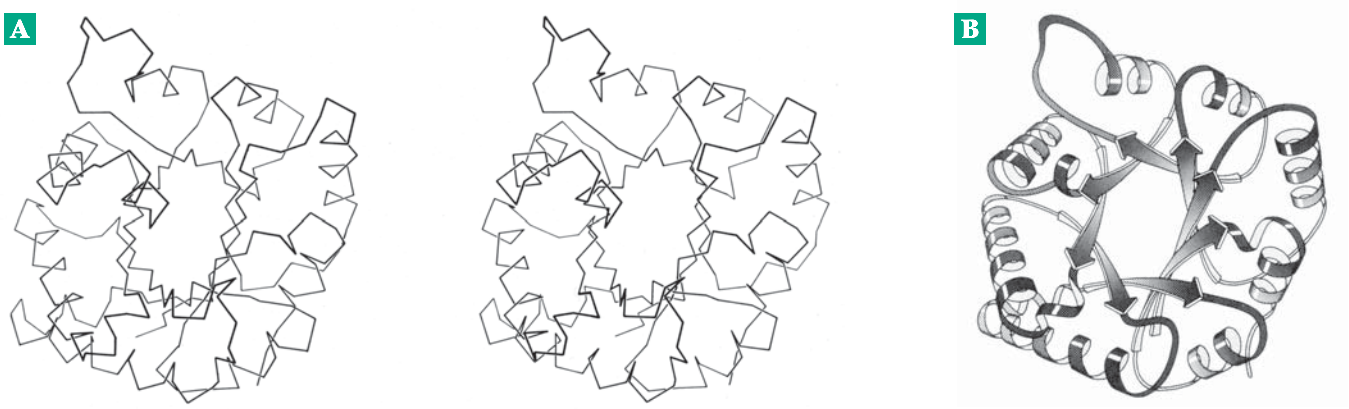


Figure 2-28 The eight-fold α/β barrel structure of triose phosphate isomerase. From Richardson. (A) Stereoscopic view. (B) Ribbon drawing. Courtesy of Jane Richardson.[117]

classified as α/β proteins. The Rossman fold is composed of six β–α units. Recently a ribonuclease inhibitor protein with 15 consecutive β–α units has been characterized.[263] Each β–α unit contains several residues of leucine. This **leucine-rich repeat** occurs in many other proteins as well.[264,264a,264b]

The **α/β barrel** shown in Fig. 2-28 consists of 8 consecutive β–α units in a symmetric array.[265,266] By 1995 over 40 of these barrels had been identified in a diverse group of enzymes. One bifunctional enzyme contains two α/β barrels. Although the nature of the reaction catalyzed varies, the active site is always found in the center of the barrel at the C-terminal ends of the 8 parallel β strands and therefore between the N termini of the surrounding helices. The enzyme sequences show no homology and frequent occurrence of the 8-stranded barrel may reflect the fact that it is a natural packing arrangement of β–α units. However, a 10-stranded barrel of this type has also been found.[267]

This barrel can be compared with that of the 12-helix α,α barrel of a fungal glucoamylase whose structure is shown in Fig. 2-29. Numerous more complex folding patterns have been discovered. They have been classified by Jane Richardson.[117,122] Many of the proteins described by these folding patterns can be grouped into families and "superfamilies."[227] Chothia suggested that there may be about 1,000 families in nature;[268] over 700, with over 360 distinct folds have been identified.[268a]

2. Symmetry

A sometimes puzzling feature of protein structure is the widespread occurrence of an approximate two-fold axis of symmetry. This often arises as a natural result of association of a pair of irregular subunits (Chapter 7). The association is such that rotation

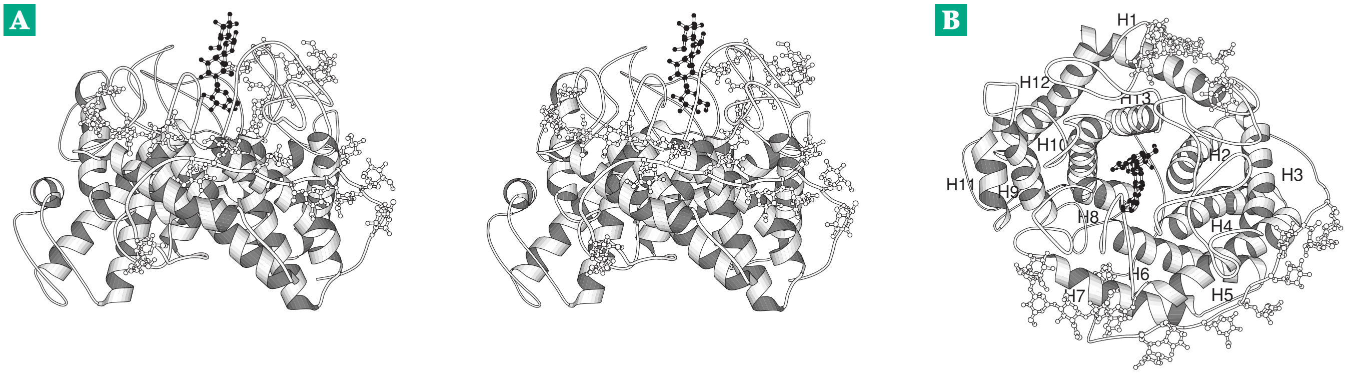


Figure 2-29 Structure of the α,α barrel of a fungal enzyme glucoamylase. (A) side view (stereoscopic); (B) top view. The active site, which cleaves glucose units from the ends of starch chains, is in the depression in the center of the barrel. Here it is occupied by an inhibitor. See Aleshin *et al.*[192] Courtesy of Alexander Aleshin.

about the twofold axis will cause the two subunits to exchange positions and to remain in an identical chemical environment. Approximate symmetry is often observed also *within* single peptide chains. For example, in the Rossman fold (Figs. 2-13, 2-27), an approximate twofold axis passes between the center strands of the β sheet residues and relates the two flanking helices, which begin with residues R10 and T100, respectively. The bound NAD$^+$ also possesses an approximate twofold axis, but it is not quite symmetrically placed at the end of the β sheet. Both phospho groups are seen to interact with the N terminus of the helix beginning at residue 10. The small bacterial protein ferredoxin (Fig. 16-16B) contains two iron-sulfur clusters related by an approximate 2-fold axis. The two β cylinders of elastase (Fig. 12-9) as well as the two sides of the flattened β barrel of copper–zinc superoxide dismutase[269] are approximately related by twofold axes. The enzyme thiosulfate: cyanide sulfurtransferase (Eq. 24-46) is remarkably symmetric but the active site is located in just one half. The widespread existence of this approximate symmetry suggests a biological significance that remains to be discovered.

3. Effects of Sequence on Folding

Studies of synthetic polypeptides as well as examination of known protein structures reveal that some amino acids, e.g., Glu, Ala, Leu, tend to promote α helix formation. Others, such as Tyr, Val, and Ile, are more often present in β structure, while Gly, Pro, and Asn are likely to be found in bends.[270,270a,270b] The frequencies with which particular amino acids appear in helices, β structure, or turns were first compiled by

TABLE 2-4
Classification of Protein Residues According to Their Tendencies to Form α Helix, β Structure, and β Turns[a]

Amino acid	P_α	Helix-forming tendency	P_β	β structure-forming tendency	P_t
Glu$^-$	1.51	++	0.37	br+	0.44
Met	1.45	++	1.05	+	0.67
Ala	1.42	++	0.83	i	0.57
Leu	1.21	++	1.30	+	0.53
Lys+	1.16	+	0.74	br	1.01
Phe	1.13	+	1.38	+	0.71
Gln	1.11	+	1.10	+	0.56
Trp	1.08	+	1.37	+	1.11
Ile	1.08	+	1.60	++	0.58
Val	1.06	+	1.70	++	0.30
Asp$^-$	1.01	w	0.54	br+	1.26
His$^+$	1.00	w	0.87	i	0.69
Arg$^+$	0.98	i	0.93	i	1.00
Thr	0.83	i	1.19	+	1.00
Ser	0.77	i	0.75	br	1.56
Cys	0.70	i	1.19	+	1.17
Tyr	0.69	br	1.47	++	1.25
Asn	0.67	br	0.89	i	1.68
Pro	0.57	br+	0.55	br+	1.54
Gly	0.57	br+	0.75	br	1.68

++ = strong former
+ = former
w = weak former
i = indifferent
br = breaker
br+ = strong breaker

[a] The conformational parameters P_α, P_β, and P_t (β turn) are the frequencies of finding a particular amino acid in an α helix, β structure, or β turn (in 29 proteins of known structure) divided by the average frequency of residues in those regions. Residues are arranged in order of decreasing tendency toward helix formation. From Chou, P. V. and Fasman, G. D (1974) *Biochemistry* **13**, 222–245.

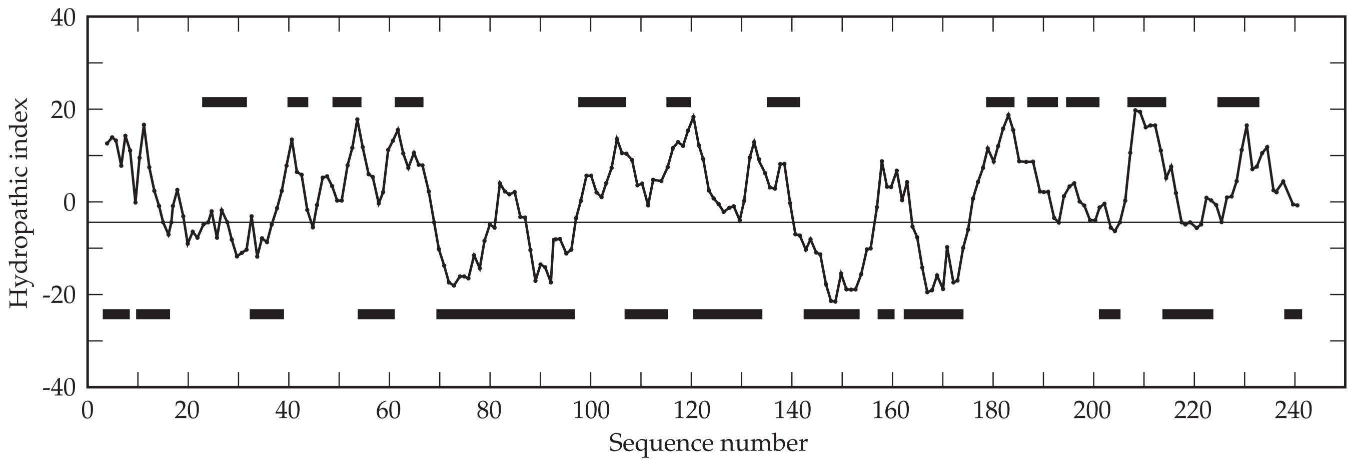


Figure 2-30 Plot of hydropathy index versus sequence number for bovine chymotrypsinogen. The indices for individual residues have been averaged nine at a time. The solid bars at the top of the plot mark interior regions as determined by crystallography. The solid bars below the plot indicate regions that are on the outside of the molecule. From Kyte and Doolittle.[280]

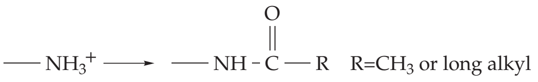
Chou and Fasman[270] (Table 2-4). These frequencies also differ significantly between parallel and antiparallel β sheets. Some of the preferences are readily understandable. For example, glycine is too flexible to stabilize a helix, but it can occur in helices between residues that are better helix formers. Because the peptide linkage to its nitrogen lacks an NH group, proline can fit into only one of three positions at the N terminus of a helix. For the strong helix formers, the β-CH_2 groups fit into the helix well and provide stabilization through van der Waals interactions. However, the β-methyl groups of valine and isoleucine cause crowding. For these residues, a β structure is more favorable than a helix.[271] Side chains of Asp, Asn, Ser, Thr, and Cys can hydrogen bond to backbone amide groups and can either stabilize or destabilize a helix or β sheet depending upon their location.

It appears that the folding pattern of a peptide is encoded in the sequence itself.[272–274] Thus, when several residues that favor helix formation are clustered together, a helix may form. Chou and Fasman suggested that when four helix formers out of six residues are clustered, nucleation of a helix takes place. The helix can then be elongated in both directions until terminated by a proline or other “helix breaker”; additional folding can then occur. If three out of five β formers are clustered, a β strand may form. If random folding brings two or more of these strands together they may associate to form the nucleus for a β sheet. Some success has been achieved using this approach in predicting whether a given residue will be found in a helix, a β strand or a loop.[275] However, prediction of complete folding patterns is much more difficult. Many new approaches are being explored.[275–279d] The problem is an important one. Although three-dimensional structures have been determined for thousands of proteins, sequences are known for hundreds of thousands. The number is growing rapidly. Being able to predict correctly a three-dimensional protein structure will be a major scientific accomplishment with many practical consequences.

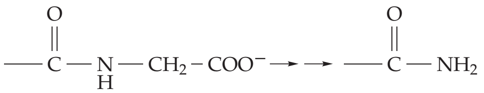
A way of examining the entire sequence for polar or nonpolar character of the side chains was introduced by Kyte and Doolittle.[280,281] A **hydropathy index** based on the polarity of the side chains of a given residue and of its neighbors in the sequence is plotted against residue number (Fig. 2-30). Helices are often found to have an **amphipathic** character, hydrophobic on one side and hydrophilic on the other. Such helices can be characterized by a plot such as that in Fig. 2-30. The hydrophobic side of an amphipathic helix can pack against a hydrophobic core of a protein, can lie against a membrane, or can be aligned with other helices to form coiled coils (Fig. 2-21) or to give a hydrophilic channel with an outer hydrophobic surface that can fit into a cell membrane.

F. Chemical Modification and Crosslinking

Some newly synthesized proteins, upon folding, are ready to “go to work” immediately but others must be modified. Pieces are frequently cut from the ends or out of the center of a folded peptide chain. Sometimes the amino and carboxyl termini are converted to nonionic groups, e.g.,

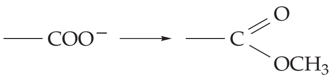


$$—NH_3^+ \longrightarrow —NH-\overset{\overset{\displaystyle O}{\|}}{C}—R \quad R{=}CH_3 \text{ or long alkyl} \tag{2-12}$$

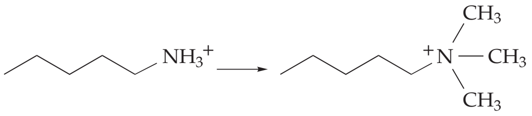


$$—\overset{\overset{\displaystyle O}{\|}}{C}—\underset{\displaystyle H}{N}—CH_2-COO^- \longrightarrow \longrightarrow —\overset{\overset{\displaystyle O}{\|}}{C}—NH_2 \tag{2-13}$$

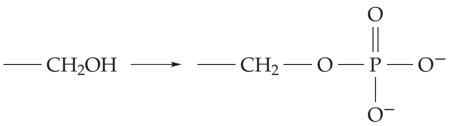
Side chains may be modified in a very large number of different ways.[282–284] These include acetylation and other kinds of acylation (Eq. 2-12),[285–287] **methylation** (Eqs. 2-14, 2-15), **phosphorylation** (Eq. 2-16), **phosphoadenylation**,[288] formation of **sulfate esters** (Eq. 2-17),[289,290] and hydrolysis (Eqs. 2-18, 2-19).[291] In at least a few proteins some L-amino acid residues are converted to D-.[292–294]



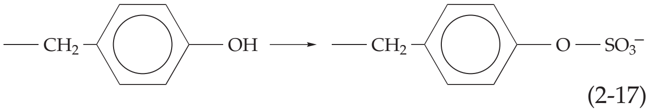
(2-14)



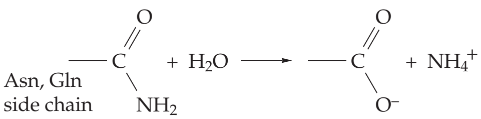
(2-15)



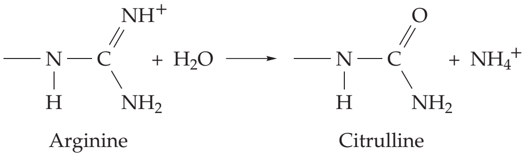
$$—CH_2OH \longrightarrow —CH_2—O—\overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle O^-}{|}}{P}}—O^- \tag{2-16}$$



(2-17)

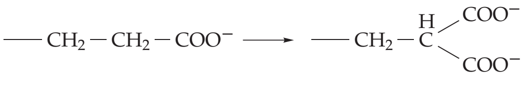


(2-18)



(2-19)

Usually, only one or a small number of side chains of a protein is modified. However, glycoproteins may contain many different attached sugar molecules. An example is the glucoamylase shown in Fig. 2-29. In several proteins involved in blood clotting (Fig. 12-17), as many as 10 glutamic acid side chains are carboxylated (Eq. 2-20).

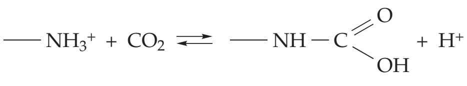


$$—CH_2—CH_2—COO^- \longrightarrow —CH_2—C(H)(COO^-)_2 \qquad (2\text{-}20)$$

The 216-residue hen egg yolk storage protein **phosvitin** contains 123 serine residues, most of which have been phosphorylated (Eq. 2-16).[295] A basic protein of the myelin sheath of neurons contains as many as 6 specific residues of **citrulline** (Eq. 2-19).[296] An adhesive protein from the foot of a marine mollusk contains ~80 repeated sequences containing hydroxy-proline 2,3-dihydroxyproline and **3,4-dihydroxyphenylalanine** (Dopa).[297,298]

Some modification reactions alter the electrical charge on a side chain and, as a consequence, can affect the ways in which the protein interacts with other molecules. Negative charges added in formation of clusters of γ-carboxyglutamates (Eq. 2-20) create strong **calcium ion-binding centers** in the modified proteins. Acylation of N termini or of Ser, Thr, or Cys side chains by long-chain fatty acids provides hydrophobic "tails" able to anchor proteins to membrane surfaces.[285,286] Addition of polyprenyl groups to cysteine side chains near the C-termini has a similar effect.[299,300] These and more complex membrane anchors are considered in Chapter 8. While modified amino acids can be found at many places in a protein, they are often located at turns. For example, serine or threonine residues in turns are often phosphorylated or glycosylated. Modification of proteins is dealt with further in Chapter 10 and at other points in the book. Glycosylation of proteins is considered in Chapter 20.

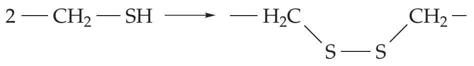
In addition to deliberate enzyme-catalyzed processes, there are nonenzymatic processes that alter proteins. These include the degradative reactions described in Section 5 and also *reversible* reactions that may be physiologically important. For example, the N-terminal amino groups of peptides, and other amino groups of low pK_a can form **carbamates** with bicarbonate (Eq. 2-21).[301–303] This provides an important mechanism of carbon dioxide transport in red blood cells (Chapter 7) and a way by which CO_2 pressure can control some metabolic processes.



$$—NH_3^+ + CO_2 \rightleftharpoons —NH—C(=O)—OH + H^+ \qquad (2\text{-}21)$$

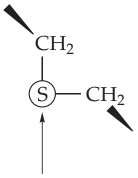
1. Disulfide Crosslinkages

The linking together of two different parts of a peptide chain or of two different peptide chains is extremely important to living beings.[304,305] One of the most widespread of crosslinkages is the **disulfide bridge**. It forms spontaneously when two –SH groups of cysteine side chains are close together and are oxidized by O_2 or some other reagent (Eq. 2-22).



$$2\ —CH_2—SH \longrightarrow —H_2C—S—S—CH_2— \qquad (2\text{-}22)$$

The disulfide group is inherently chiral, with the preferred torsion angle about the S–S bond being ~90°. Both right- and left-handed disulfide bridges occur in proteins.[305,306]

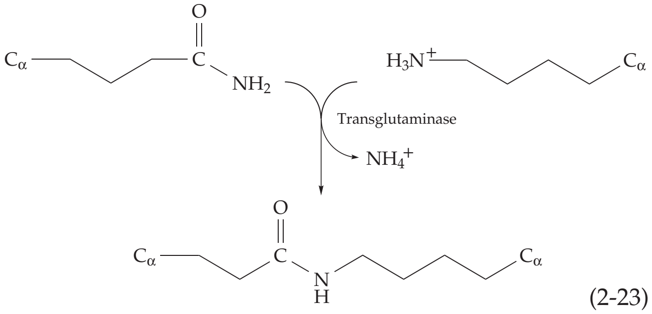


Two S atoms viewed down S—S axis

Disulfide linkages are frequently present in proteins that are secreted from cells but are less common in enzymes that stay within cells. Perhaps because the latter are in a protected environment, the additional stabilization provided by disulfide bridges is not needed. Disulfide bridges are not only found within single-subunit proteins but they also link different peptide chains. For example, the four chains of each immunoglobulin molecule (Chapter 31) are joined by disulfide bridges and each domain within the chains is stabilized by a disulfide bridge. One of the most highly crosslinked proteins known occurs in the keratin matrix of hair. Breakage of the –S–S– linkages of this protein is an essential step in the chemical "permanent wave" process. A thiol compound is used to reductively cleave the crosslinks and after the hair is reset new crosslinks are formed by air oxidation.

2. Other Crosslinkages

Another common crosslink is an amide formed between the γ-carboxyl group of a glutamic acid side chain and an amino group from a lysine residue.[307] This **isopeptide** linkage is formed from a residue of gluta-mine through the action of the enzyme **transglutaminase** (Eq. 2-23). Isopeptide crosslinks are found in hair, skin, connective tissue, and blood clots.



(2-23)

Occasionally an isopeptide linkage joins amino groups of lysine side chains to the C-terminal carboxyl groups of other peptide chains to give branched chains, e.g. see **ubiquitin** (Box 10-C). Elaborate crosslinks derived from lysine are found in collagen and elastin (Chapter 8). Dityrosine linkages formed by oxidative joining of the aromatic rings of tyrosine are found in insect cuticle and in the plant cell wall extensins[202] (Chapters 20, 25).

G. Dynamic Properties of Proteins

Sometimes using energy from the cleavage of chemical bonds, sometimes depending only upon energy provided by the random bombardment by solvent molecules, proteins perform their specific functions with amazing speed. A question that has long intrigued biochemists is "to what extent do proteins stretch or flex, unfold and refold, or undergo other conformational movements during their action?" To answer this and related questions, many techniques are being applied to study the dynamic properties of proteins.[308,309]

1. Motion of Backbone and Side Chains

Even in the *crystalline state* there is evidence of movement. In the images constructed from X-ray or neutron diffraction experiments side chains on the surfaces of protein molecules are often not clearly visible because of rapid rotational movement. Some segments of the polypeptide chain may be missing from the image. However, side chain groups within the core of a domain are usually seen clearly. They probably move only in discrete steps. However, they may sometimes shift rapidly between different conformations, all of which maintain a close-packed interior.[310–312]

Studies of nuclear magnetic resonance spectra (Chapter 3) and of polarization of fluorescence (Chapter 23), have shown that there is rapid though restricted rotational movement of side chains of proteins *in solution*. Even buried phenylalanine and tyrosine side chains often rotate rapidly whereas movement of the bulkier tryptophan rings is more limited. Peptide NH protons in unfolded polypeptide chains undergo rapid acid or base catalyzed exchange with H_2O.[313] However, in globular proteins, the rate of exchange of buried NH protons is often orders of magnitude slower.[314–316] For most proteins there appears to be little tendency to unfold completely and then refold; the major domains hold together tightly. However, there may be local unfolding, for example, of a helix at the surface of a protein, which will allow more rapid exchange. Using NMR techniques (Chapter 3), rates of exchange of all of the individual peptide NH protons within small proteins can be measured.[315] Cracks may open up in proteins. This is suggested by the fact that O_2, I_2, and certain other small molecules are apparently able to penetrate the protein freely and to quench the fluorescence of buried tryptophan side chains (Chapter 23).

Since packing density tends to be lower at active sites than in the bulk of the protein, it is probable that more conformational alterations occur near active sites than elsewhere.[317] Lumry and Rosenberg[318] suggested that the "defects" of poor packing and poor hydrogen bonding of some regions of a folded peptide chain provide a store of potential energy that can be important to the functioning of a protein. Even in a very well-packed protein domain there are defects. Some atoms are compressed by the folding of the peptide chain and are closer together by over 0.04 nm than predicted by the van der Waals radii. These packing defects have been estimated to destablize the protein by as much as 250 kJ/mol.[319] Details of structural heterogeneity within several proteins for which very precise structural data are available have been described.[311]

2. Conformational Changes

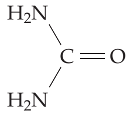
We have seen that some polypeptides assume an extended β conformation while others form helices. In some cases, the same protein can do both. For example, hair can be stretched greatly, the α helices of the keratin molecules uncoiling and assuming a β conformation with hydrogen bonds *between* chains instead of *within* a single chain. Thus, a polymer may have more than one conformational state in which the folding and hydrogen bonding are different.[320,321] With soluble proteins more than one folded conformation is possible with different sets of hydrogen bonds and internal hydrophobic interactions. Some of the conformations of a globular protein are more stable than others, and a protein will ordinarily assume one of the energetically most favorable conformations. However, there may be other conformations of almost equal energy.

A large body of evidence suggests that many proteins do exist in two or more different but well-defined conformational states and that the ability of a protein to undergo easy conversion from one to another is of great biological significance. In some cases such as that of hemoglobin (Chapter 7), there are changes in the interactions between subunits. Alterations in the hinge regions between domains of immunoglobulins have been seen. In many enzymes, including the dehydrogenases (Chapter 15), kinases (Chapter 12), and aspartate aminotransferase (Fig. 2-6), a cleft between two domains appears to open and close.[322] In other cases, more subtle alterations in conformational state, involving mainly changes in the internal hydrogen bonded network together with small localized changes in chain folding, have been observed.

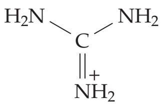
Conformational alterations in proteins are probably facilitated by the fact that some hydrogen-bonded groups are found within the hydrophobic interior. *All of the buried hydrogen atoms suitable for hydrogen bond formation are ordinarily hydrogen bonded to an electron donor group. However, because oxygen atoms in proteins each have two unshared electron pairs, there are, in general, more electron donor groups than there are hydrogen atoms to which they can bind. This sets the stage for a competition between electronegative centers for particular proton suitable for hydrogen bonding and provides a molecular basis for the easy triggering of conformational changes.*[323]

3. Denaturation and Refolding

An extreme conformational alteration is the **denaturation** of proteins, which may be caused by heating or by treatment with reagents such as strong acids and bases, **urea**, **guanidinium chloride**, and **sodium dodecyl sulfate** (SDS).



Urea



Guanidinium chloride

$CH_3—(CH_2)_{11}—OSO_3^- Na^+$

Sodium dodecyl sulfate (SDS)

Denaturation leads to unfolding of a protein to a more random conformation. In the denatured state the amide groups of the peptide chain form hydrogen bonds with surrounding water molecules and with denaturants such as urea or the guanidinium ion rather than with each other.[324,325] Denaturants also diminish the strength of the hydrophobic interactions that promote folding.[326] Characteristic biochemical activities are lost and physical properties such as sedimentation constant, viscosity, and light absorption are altered. The ease of denaturation of proteins and the fact that denaturation is sometimes reversible show that the energy differences between the folded conformations and the open random coil conformation are usually not great.[274,327,328] However, it has been difficult to establish the amount of stabilization of a folded polypeptide chain provided by buried hydrogen bonds[328a,328b,328c] or the role of cooperatively formed hydrogen-bonded chains[328c,328d] or clusters.[328c,328d,328e]

Complete denaturation of a protein was generally regarded as an irreversible process prior to 1956 when Anfinsen showed that denatured ribonuclease (Chapter 12) could refold spontaneously.[273] This 124-residue protein contains four disulfide (–S–S–) bridges and thus is tied firmly together. When these bridges are broken reductively in the presence of a denaturing agent, the enzyme becomes inactive. Anfinsen found that upon reoxidation under appropriate conditions, complete activity reappeared. The molecules had folded spontaneously into the correct conformation, the one in which the correct *one* of 105 (7 x 5 x 3 x 1) possible pairings of the eight –SH groups present needed to reform the four disulfide bridges had taken place. This observation has had an important influence on thinking about protein synthesis and folding of polypeptide chains into biologically active molecules.

A puzzling problem was posed by Levinthal many years ago.[329] We usually assume that the peptide chain folds into one of the most stable conformations possible. However, proteins fold very rapidly. Even today, no computer would be able, in our lifetime, to find by systematic examination the thermodynamically most stable conformation.[328] It would likewise be impossible for a folding protein to "try out" more than a tiny fraction of all possible conformations. Yet folded and unfolded proteins often appear to be in a thermodynamic equilibrium! Experimental results indicate that denatured proteins are frequently in equilibrium with a **compact denatured state** or "molten globule" in which hydrophobic groups have become clustered and some secondary structures exists.[330–336] From this state the polypeptide may rearrange more slowly through other **folding intermediates** to the final "native conformation."[336a,336b]

It is generally assumed that within cells the folding of the peptide chain commences while the chain is still being synthesized on a ribosome. The growing chain probably folds rapidly in a random way until it finds one or more stable conformations which serve as folding intermediates for slower conversion to the finished protein.[337,338] However, any process within a cell is affected by the complex intracellular environment.[339] Folding can be catalyzed or inhibited by proteins known as **molecular chaperones**. Folding may sometimes require isomerization of one or a few proline residues from *trans* to *cis*.[340,341] For example, during the refolding of ribonuclease the isomerization of Pro 93 appears to be a rate-limiting step.[342,343] Such isomerizations may be assisted by **peptidyl-prolyl-(*cis–trans*) isomerases** (Box 9-F). Disulfide linkages are sometimes formed incorrectly.[344,345] A **protein-disulfide isomerase** catalyzes cleavage and reformation of these bridges, allowing the protein to find the most stable crosslinking arrangement. The actions of these enzymes and of molecular chaperones are considered further in Chapters 10 and 12.

4. Effects of pH and Solvent

Because polypeptide chains contain many acidic and basic amino acid side chains, the properties of proteins are greatly influenced by pH. At low pH carboxylates, $-S^-$, and imidazole groups accept protons causing the overall net charge on the macromolecule to be strongly positive. At high pH protons are lost and the protein becomes negatively charged. Electrostatic repulsion between like charges may cause proteins to denature at low or high pH. More stable proteins may be very soluble at low or high pH. Proteins often have a minimum solubility and a maximum stability near the **isoelectric point,** the pH at which the net charge is zero.[336,346,347] Activities of enzymes, abilities to bind specifically to other proteins, and

BOX 2-C THE NOBEL PRIZES

Many young scientists dream of one day winning a Nobel Prize. Although denounced by some, the much sought and highly publicized award has, since 1901, been given to an outstanding group of scientists. Many of these have made major contributions to biochemistry or to techniques important to biochemists. Here is a partial list.

Year	Name	Prize[a]	Discovery or subject of study
1901	Wilhelm C. Röntgen	Physics	Discovery of X-rays
1902	Emil H. Fischer	Chemistry	Synthesis of sugars and purines
1903	Svante A. Arrhenius	Chemistry	Electrolytic dissociation; a founder of physical chemistry
1903	Antoine Henri Becquerel; Marie S. Curie, and Pierre Curie	Physics	Discovery and study of radioactivity
1906	Camillo Golgi and S. Ramon y Cajal	Physiology/Medicine	Discovery of Golgi apparatus
1907	Edward Buchner	Chemistry	Biochemistry, cell-free fermentation
1910	Albrecht Kossel	Physiology/Medicine	Isolation of nucleic acid bases
1914	Max von Laue	Physics	Discovery of X-ray diffraction by crystals
1915	Richard M. Willstätter	Chemistry	Plant pigments, especially chlorophyll
1915	William H. and William L. Bragg	Physics	Analysis of crystal structure by X-rays
1919	Jules Bordet	Physiology/Medicine	Discovery of blood complement; complement fixation test
1920	Walther H. Nernst	Chemistry	Thermochemistry
1922	Archibald V. Hill and Otto F. Meyerhof	Physiology/Medicine	Chemistry of muscle contraction
1923	Frederick G. Banting and John J. R. Macleod	Physiology/Medicine	Discovery of insulin and treatment of diabetes
1926	Theodor Svedburg	Chemistry	Study of high M_r compounds, development of ultracentrifuge
1927	Heinrich O. Wieland	Chemistry	Bile acids
1928	Adolf O. R. Windaus	Chemistry	Sterols and vitamins
1929	Frederick G. Hopkins and Christiaan Eijkman	Physiology/Medicine	Discovery of vitamins, tryptophan, vitamin B_1
1929	Arthur Harden and Hans A. S. von Euler-Chelpin	Chemistry	Fermenting enzymes, fermentation of sugars
1930	Karl Landsteiner	Physiology/Medicine	Blood groups A, B, O, Rh
1930	Hans Fischer	Chemistry	Structures and chemistry of porphyrins, chlorophyll
1931	Otto H. Warburg	Physiology/Medicine	Respiratory enzymes
1933	Thomas H. Morgan	Physiology/Medicine	Chromosome theory and chromosome maps
1934	George R. Minot, William P. Murphy, and George H. Whipple	Physiology/Medicine	Treatment of pernicious anemia
1936	Henry H. Dale and Otto Loewi	Physiology/Medicine	Acetylcholine release at nerve endings
1937	Albert von Szent-Györgyi	Physiology/Medicine	Vitamin C
1937	Walter N. Haworth and Paul Karrer	Chemistry	Carbohydrate structures, structures of carotenoids, flavins, vitamin A
1938	Richard Kuhn	Chemistry	Carotenoids and vitamins
1939	Gerhard Domagk	Physiology/Medicine	Prontosil, first antibacterial sulfa drug
1939	Adolf F. J. Butenandt and L. Ruzicka	Chemistry	Isolation and study of sex hormones, study of polymethylenes, terpenes
1943	E. A. Doisy and Carl Henrik Dam	Physiology/Medicine	Isolation and study of vitamin K
1945	Alexander Fleming, Ernst B. Chain, and Howard W. Florey	Physiology/Medicine	Discovery and structure of penicillin
1945	A. J. Virtanen	Chemistry	Nutritional chemistry
1946	James B. Sumner, J. H. Northrop, and W. M. Stanley	Chemistry	Crystallization of enzymes and virus proteins
1947	Carl F. Cori and Gerty T. Cori	Physiology/Medicine	Glycogen metabolism, the Cori cycle
1947	B. A. Houssay and Robert Robinson	Chemistry	Investigation of plant alkaloids
1948	Arne W. K. Tiselius	Chemistry	Electrophoresis, study of serum proteins
1950	Phillip S. Hench, Edward C. Kendall, and Tadeus Reichstein	Physiology/Medicine	ACTH
1952	Selman A. Waksman	Physiology/Medicine	Discovery of streptomycin
1952	A. J. P. Martin and Richard L. M. Synge	Chemistry	Paper chromatography
1953	Fritz A. Lipmann and Hans A. Krebs	Physiology/Medicine	Coenzyme A, citric acid cycle
1954	Linus C. Pauling[b]	Chemistry	The nature of the chemical bond
1955	A. H. T. Theorell	Physiology/Medicine	Oxidative enzymes
1955	Vincent du Vigneaud	Chemistry	Synthesis of biotin and oxytocin
1957	Daniel Bovet	Physiology/Medicine	First antihistamine drug
1957	Alexander R. Todd	Chemistry	Work on nucleotides coenzymes
1958	Joshua Lederberg, George W. Beadle, and Edward L. Tatum	Physiology/Medicine	One gene-one enzyme hypothesis from genetic studies of neurospora
1958	Frederick Sanger[b]	Chemistry	Protein sequencing, insulin

numerous physiological processes are controlled by effects of pH on proteins.

The solvent for most proteins in nature is water. However, many enzymes function well in organic solvents if they retain only a small amount of essential structural and catalytic water.[348]

5. Irreversible Damage to Proteins

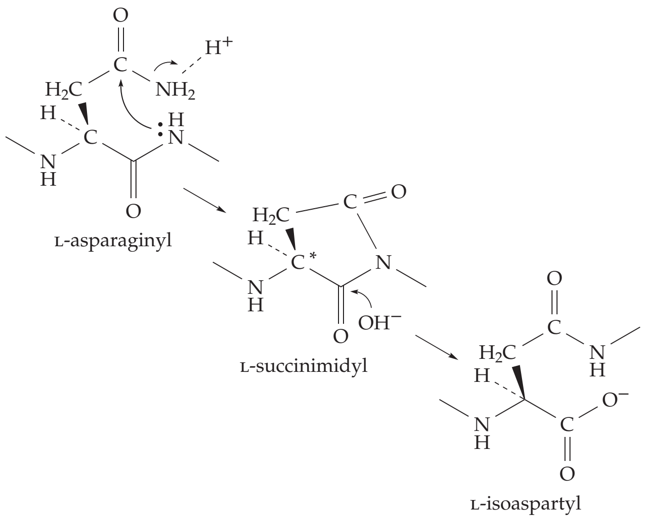
Every protein in an organism has its own characteristic lifetime. No sooner is it synthesized than degradation begins. Enzyme-catalyzed cleavage of the peptide linkages leads to turnover of proteins but before this occurs a number of spontaneous damaging reactions may alter the protein. Prevalent among these is **deamidation** of asparaginyl residues to aspartyl or

BOX 2-C THE NOBEL PRIZES (continued)

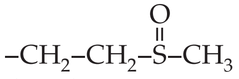
Year	Name	Prize[a]	Discovery or subject of study
1959	Arthur Kornberg and Severo Ochoa	Physiology/Medicine	Enzymatic synthesis of DNA
1960	F. M. Burnet and P. Medawar	Physiology/Medicine	Immunological tolerance in animals
1961	Melvin Calvin	Chemistry	Photosynthesis using $^{14}CO_2$
1962	Francis H. C. Crick, James D. Watson, and Maurice H. F. Wilkins	Physiology/Medicine	Molecular structure of DNA
1962	John C. Kendrew and Max F. Perutz	Chemistry	Structures of protein by X-ray diffraction
1962	Linus C. Pauling[b]	Peace	Ending atmospheric testing of nuclear weapons
1964	Konrod E. Bloch and Feodor Lynen	Physiology/Medicine	Pathways of cholesterol biosynthesis
1964	Dorothy M. Crowfoot Hodgkin	Chemistry	X-ray structures, vitamin B_{12}
1965	Andre Lwolf, Jacques Monod, and Francois Jacob	Physiology/Medicine	Messenger RNA, regulation of transcription
1967	Ragnar Granit, Haldan Keffer Hartline, and George Wald	Physiology/Medicine	Chemistry of vision
1967	Manfred Eigen and Norrish Porter	Chemistry	Study of high-speed chemical reactions
1968	Robert W. Holley	Physiology/Medicine	RNA sequence
1968	H. Gobind Khorana and Marshall W. Nirenberg	Physiology/Medicine	The genetic code
1969	Max Delbrück, Alfred D. Hershey, and Salvador E. Luria	Physiology/Medicine	Replication and genetic structures of viruses
1970	Julius Axelrod, Bernard Katz, and Ulf von Euler	Physiology/Medicine	Transmission of nerve impulses, noradrenaline
1970	Luis F. Leloir	Chemistry	Role of sugar nucleotides in biosynthesis of carbohydrates
1971	Earl W. Sutherland, Jr.	Physiology/Medicine	Cyclic AMP
1972	Gerald M. Edelman and Rodney R. Porter	Physiology/Medicine	Structure of antibodies
1972	Christian B. Anfinsen, Sanford Moore, and William H. Stein	Chemistry	Ribonuclease, structure and activity
1974	Albert Claude, Christian R. de Duve, and George E. Palade	Physiology/Medicine	Cell structure
1975	David Baltimore, Renato Dulbecco, and Howard M. Temin	Physiology/Medicine	Reverse transscriptase
1975	John W. Cornforth and Vladimir Prelog	Chemistry	Stereochemistry of organic molecules and of enzymatic reactions
1977	Rosalyn S. Yalow, Roger C. L. Guillemin, and Andrew V. Schally	Physiology/Medicine	Radioimmunoassay thyrotropin-releasing hormone
1978	Daniel Nathans, Werner Arber, and Hamilton O. Smith	Physiology/Medicine	Restriction enzymes
1978	Peter Mitchell	Chemistry	Biological energy transfer
1980	Paul Berg, Walter Gilbert, and Frederick Sanger[b]	Chemistry	Recombinant DNA, methods of sequence determination for DNA
1982	Sune K. Bergstrom, Bergt I. Samuelsson, and John R. Vane	Physiology/Medicine	Isolation and study of prostaglandins
1982	Aaron Klug	Chemistry	Development of crystallographic electron microscopy
1983	Barbara McClintock	Physiology/Medicine	Gene transposition
1984	Niels K. Jerne, Georges J. F. Köhler, and Cesar Milstein	Physiology/Medicine	Cellular basis of immunology
1984	R. Bruce Merrifield	Chemistry	Solid-phase synthesis of peptides
1985	Joseph L. Goldstein and Michael S. Brown	Physiology/Medicine	Control of cholesterol metabolism
1986	Rita Levi-Montalcini and Stanley Cohen	Physiology/Medicine	Nerve growth factor
1987	Susumu Tonegawa	Physiology/Medicine	Genetics of antibody formation
1988	Johan Diesenhofer, Robert Huber, and Hartmut Michel	Chemistry	Three-dimensional structure of a photosynthetic reaction center
1988	Gertrude Elion, George Hitchings, and James Black	Physiology/Medicine	Principles of drug treatment and design of many important drugs
1989	J. Michael Bishop, Harold E. Varmus, and Joseph E. Murray	Physiology/Medicine	Origin of retroviral oncogenes
1989	Sidney Altman and Thomas R. Cech	Chemistry	Catalytic RNA
1991	Erwin Neher and Bert Sakmann	Physiology/Medicine	Functioning of single ion channels in cells
1991	Richard R. Ernst	Chemistry	High-resolution NMR
1992	Edmond H. Fischer and Edwin G. Krebs	Physiology/Medicine	Reversible protein phosphorylation in biological regulation
1992	Rudolph A. Marcus	Chemistry	Theory of electron transfer reactions
1993	Richard J. Roberts and Phillip A. Sharp	Physiology/Medicine	Discovery of split genes
1993	Michael Smith and Kary B. Mullis	Chemistry	Oligonucleotide-directed mutagenesis and polymerase chain reaction
1994	Alfred Gilman and Martin Rodbell	Physiology/Medicine	“G-proteins”
1995	Edwin B. Lewis, Christiane Nusslein-Volhard, and Eric Wieschaus	Physiology/Medicine	Homeotic mutations in *Drosophila*
1995	Paul Crutzen, Sherwood Rowland, and Mario Molina	Chemistry	Damage to the stratospheric ozone layer
1996	Peter C. Doherty and Rolf M. Zinkernagel	Physiology/Medicine	Specificity of cell-mediated immune response
1997	Paul D. Boyer and John E. Walker	Chemistry	Mechanism of ATP synthesis
1997	Jens C. Skou	Chemistry	Discovery of Na$^+$, K$^+$-ATPase
1997	Stanley B. Prusiner	Physiology/Medicine	Discovery of prions
1998	Robert F. Furchgott, Louis J. Ignarro, and Ferid Murad	Physiology/Medicine	Nitric oxide as a signaling molecule
1999	Günter Blobel	Physiology/Medicine	Intrinsic signals that govern transport and localization of proteins
2000	Arvid Carlsson	Physiology/Medicine	Identification of dopamine as signaling molecule in brain
	Paul Greengard	Physiology/Medicine	Discovery of the dopamine signaling cascade
	Eric R. Kandel	Physiology/Medicine	Molecular basis of learning

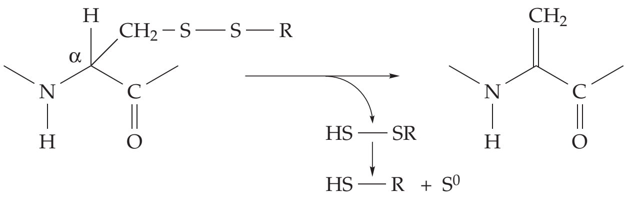
[a] Prizes are given in Physics, Chemistry, Physiology or Medicine, Literature and Peace.
[b] Sanger and Pauling each have been awarded two Nobel Prizes.

isoaspartyl groups (Eq. 2-24).[349,350] Aspartyl residues can undergo the same type of cyclization at low pH. The sequence Asn-Gly is especially susceptible to rearrangement according to Eq. 2-24. However, the peptide torsion angles may be more important in determining whether deamidation occurs. The intermediate succinimide may be racemized easily at the chiral center marked by the asterisk in Eq. 2-24. Thus, the Asn-Gly sequence represents a weak linkage, which is nevertheless present in many proteins.



(2-24)

Cystine residues in disulfide bridges are also not completely stable but undergo β elimination[305,351] (Chapter 13) at slightly alkaline pH values according to Eq. 2-25. The free thiol group formed (HSR) is still attached to the protein but may sometimes be in a position to enter into thiol-disulfide exchange reactions with other S–S bridges causing further degradation. Methionine side chains in proteins can be oxidized to sulfoxides:  $-CH_2-CH_2-\overset{\overset{O}{\|}}{S}-CH_3$ and hydroxyl groups can be introduced into aromatic rings by oxidation. Hydrogen peroxide and other oxidants may be responsible for such oxidation within cells (see Chapter 18). The very long-lived proteins of the lens of the eye are especially susceptible to deamidation, racemization, oxidation, and accumulation of covalently attached blue fluorescent compounds.[352–354]



(2-25)

H. Design and Engineering of Proteins

Methods of chemical synthesis of polypeptides and of cloning and mutating genes now allow us to alter peptide sequences at will and to design completely new proteins.[355–356b] The methods are discussed in Chapters 3, 5, and 26. The following are examples.

Peptides that form α helices that associate as coiled coils,[357] or as three- or four-helix tetrameric bundles[179a,358–360]or amphipathic helices that associate with lipid bilayers have been made.[355,361] More difficult has been the design of proteins that form β sheets.[362–364a] Many efforts are being made to understand protein stability[365–367] by systematic substitutions of one residue for another. Addition of new disulfide linkages at positions selected by study of three-dimensional structures sometimes stabilizes enzymes.[368–371] On the other hand elimination of unnecessary cysteine residues can enhance stability by preventing β elimination[351] and replacement of asparagine by threonine can improve the thermostability of enzymes by preventing deamidation.[372,373]

Artificial mutants of subtilisin[373] and of many other enzymes are helping us to understand the mechanisms of catalysis. Artificially prepared temperature-sensitive mutants (Chapter 26) and naturally occurring mutants are providing new insight into pathways of folding of proteins whose three-dimensional structures are known. For these studies it is necessary to follow a rational strategy in deciding which of the astronomical number of possible mutant forms may be of interest. For example, residues within active sites may be changed. Introduction of mutations at random locations can also be useful in finding regions of interest.[374] When mutant proteins can be crystallized, an exact understanding of the effect of the mutation on the structure is possible. Sometimes two or more mutations may have to be introduced to obtain the desired modification.

It is also possible to incorporate a range of "unnatural" amino acids in specific sites in a polypeptide sequence and to observe resulting effects on a protein's properties.[375] A quite different approach is to design polypeptides that mimic a natural peptide but consist of D-amino acids. The peptide chain is *reversed*, i.e., the N terminus becomes the C terminus and every peptide linkage is also reversed. The amino acid side chains preserve their relationships one to another and the backbone atoms tend to preserve their hydrogen bonding pattern. Peptides made in this way tend to be resistant to cleavage by enzymes. Some may be useful as drugs.[376,377]

References

1. Glendening, E. D., Faust, R., Streitwieser, A., Vollhardt, K. P. C., and Weinhold, F. (1993) *J. Am. Chem. Soc.* **115**, 10952–10957

2. Gao, Q., Weber, H.-P., Craven, B. M., and McMullan, R. K. (1994) *Acta Crystallogr.* **B50**, 695–703

3. Hegstrom, R. A., and Kondepudi, D. K. (1990) *Sci. Am.* **262**(Jan), 108–115

4. Cahn, R. S., Ingold, C., and Prelog, V. (1966) *Angew. Chem. Int. Ed. Engl.* **5**, 385–415

5. Eliel, E. L., Wilen, S. H., and Mander, L. N. (1994) *Stereochemistry of Organic Compounds*, Wiley, New York

6. Bentley, R. (1969) *Molecular Asymmetry in Biology*, Vol. 1, Academic Press, New York (pp. 49–56)

7. Blackwood, J. E., Gladys, C. L., Loening, K. L., Petrarca, A. E., and Rush, J. E. (1968) *J. Am. Chem. Soc.* **90**, 509–510

8. Hanson, K. R. (1966) *J. Am. Chem. Soc.* **88**, 2731–2742

9. Lambert, J. B. (1970) *Sci. Am.* **222**(Jan), 58–70

10. Klyne, W., and Prelog, V. (1960) *Experientia* **16**, 521–523

11. Liébecq, C., ed. (1992) *Biochemical Nomenclature*, Portland Press, London and Chapel Hill, North Carolina (for the International Union of Biochemistry and Molecular Biology)

12. Chiang, Y., Kresge, A. J., and Tang, Y. S. (1984) *J. Am. Chem. Soc.* **106**, 460–462

13. Beak, P., Fry, F. S., Jr., Lee, J., and Steele, F. (1976) *J. Am. Chem. Soc.* **98**, 171–179

14. Beak, P. (1977) *Acc. Chem. Res.* **10**, 186–192

15. Metzler, D. E., Harris, C. M., Johnson, R. J., Siano, D. B., and Thomson, J. A. (1973) *Biochemistry* **12**, 5377–5392

16. Katritzky, A. R., and Karelson, M. (1991) *J. Am. Chem. Soc.* **113**, 1561–1566

17. Blomberg, F., Maurer, W., and Rüterjans, H. (1977) *J. Am. Chem. Soc.* **99**, 8149–8159

18. Walters, D. E., and Allerhand, A. (1980) *J. Biol. Chem.* **255**, 6200–6204

19. Fersht, A. (1977) *Enzyme Structure and Mechanism*, Freeman, San Francisco, California

20. Gabler, R. (1978) *Electrical Interactions in Molecular Biophysics*, Academic Press, New York (p. 188)

21. Israelachvili, J. N., and McGuiggan, P. M. (1988) *Science* **241**, 795–800

22. Salem, L. (1962) *Can. J. Biochem. Physiol.* **40**, 1287–1298

23. Tsao, Y.-H., Fennell Evans, D., and Wennerström, H. (1993) *Science* **262**, 547–550

23a. Kurihara, K., and Kunitake, T. (1992) *J. Am. Chem. Soc.* **114**, 10927–10933

23b. Yang, L., Valdeavella, C. V., Blatt, H. D., and Pettitt, B. M. (1996) *Biophys. J.* **71**, 3022–3029

24. Warshel, A. (1987) *Nature (London)* **330**, 15–16

25. Honig, B., and Nicholls, A. (1995) *Science* **268**, 1144–1149

26. Meot-Ner, M. M. (1988) *J. Am. Chem. Soc.* **110**, 3075–3080

27. Loewenthal, R., Sancho, J., Reinikainen, T., and Fersht, A. R. (1993) *J. Mol. Biol.* **232**, 574–583

28. Gristina, A. G., Oga, M., Webb, L. X., and Hobgood, C. D. (1985) *Science* **228**, 990–993

29. Jeffrey, G. A., and Saenger, W. (1991) *Hydrogen Bonding in Biological Structures*, Springer-Verlag, Berlin

30. Gilli, P., Bertolasi, V., Ferretti, V., and Gastone, G. (1994) *J. Am. Chem. Soc.* **116**, 909–915

30a. Jeffrey, G. A. (1997) *An Introduction to Hydrogen Bonding*, Oxford Univ. Press, New York

31. Derewenda, Z. S., Derewenda, U., and Kobos, P. M. (1994) *J. Mol. Biol.* **241**, 83–93

31a. Wahl, M. C., and Sundaralingam, M. (1997) *Trends Biochem. Sci.* **22**, 97–102

31b. Bella, J., and Berman, H. M. (1996) *J. Mol. Biol.* **264**, 734–742

32. Schulz, G. E., and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer, New York (p. 36)

32a. Platts, J. A., Howard, S. T., and Bracke, B. R. F. (1996) *J. Am. Chem. Soc.* **118**, 2726–2733

33. Taylor, R., Kennard, O., and Versichel, W. (1983) *J. Am. Chem. Soc.* **105**, 5761–5766

34. Murray-Rust, P., and Glusker, J. P. (1984) *J. Am. Chem. Soc.* **106**, 1018–1025

35. Görbitz, C. H., and Etter, M. C. (1992) *J. Am. Chem. Soc.* **114**, 627–631

36. Nelson, D. D., Jr., Fraser, G. T., and Klemperer, W. (1987) *Science* **238**, 1670–1674

36a. Ghanty, T. K., Staroverov, V. N., Koren, P. R., and Davidson, E. R. (2000) *J. Am. Chem. Soc.* **122**, 1210–1214

36b. Cornilescu, G., Ramirez, B. E., Frank, M. K., Clore, G. M., Gronenborn, A. M., and Bax, A. (1999) *J. Am. Chem. Soc.* **121**, 6275–6279

36c. Benedict, H., Shenderovich, I. G., Malkina, O. L., Malkin, V. G., Denisov, G. S., Golubev, N. S., and Limbach, H.-H. (2000) *J. Am. Chem. Soc.* **122**, 1979–1988

37. Fersht, A. R., Shi, J., Knill-Jones, J., Lowe, D. M., Witkinson, A. J., Blow, D. M., Brick, P., Carter, P., Woye, M. M. Y., and Winter, G. (1985) *Nature (London)* **314**, 235–238

38. Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) *Science* **255**, 178–183

38a. Takano, K., Yamagata, Y., Kubota, M., Funahashi, J., Fujii, S., and Yutani, K. (1999) *Biochemistry* **38**, 6623–6629

39. Lim, W. A., Hodel, A., Sauer, R. T., and Richards, F. M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 423–427

40. Schuster, P., Zundel, G., and Sandorfy, C. (1976) *The Hydrogen Bond*, North Holland Publ. Co., Amsterdam (3 vols.)

41. Luck, W. A. P., and Kleeberg, H. (1978) in *Photosynthetic Oxygen Evolution* (Metzner, H., ed), pp. 1–29, Academic Press, London

42. Edsall, J. T. (1983) *Trends Biochem. Sci.* **8**, 29–31

43. Grunwald, E. (1986) *J. Am. Chem. Soc.* **108**, 5719–5726

44. Westhof, E. (1993) *Water and Biological Macromolecules*, CRC Press, Boca Raton, Florida

45. Kusalik, P. G., and Svishchev, I. M. (1994) *Science* **265**, 1219–1221

46. Libnau, F. O., Toft, J., Christy, A. A., and Kvalheim, O. M. (1994) *J. Am. Chem. Soc.* **116**, 8311–8316

47. Benson, S. W., and Siebert, E. D. (1992) *J. Am. Chem. Soc.* **114**, 4269–4276

48. Liu, K., Cruzan, J. D., and Saykally, R. J. (1996) *Science* **271**, 929–933

49. Pauling, L. (1935) *J. Am. Chem. Soc.* **57**, 2680

50. Wittebort, R. J., Usha, M. G., Ruben, D. J., Wemmer, D. E., and Pines, A. (1988) *J. Am. Chem. Soc.* **110**, 5668–5671

51. Li, J., and Ross, D. K. (1993) *Nature (London)* **365**, 327–329

52. Dyke, T. R., Mack, K. M., and Muenter, J. S. (1977) *J. Chem. Phys.* **66**, 498–510

53. Saenger, W. (1979) *Nature (London)* **279, 280**, 343–344, 848

54. Saenger, W. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 93–114

55. Pribble, R. N., and Zwier, T. S. (1994) *Science* **265**, 75–79

56. Gregory, J. K., Clary, D. C., Liu, K., Brown, M. G., and Saykally, R. J. (1997) *Science* **275**, 814–817

57. Quiocho, F. A., Sack, J. S., and Vyas, N. K. (1987) *Nature (London)* **329**, 561–564

58. Pflugrath, J. W., and Quiocho, F. A. (1988) *J. Mol. Biol.* **200**, 163–180

59. Kauzmann, W. (1959) *Adv. Prot. Chem.* **14**, 1–63

60. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology*, Dover, Mineola, New York (p. 393)

61. Tanford, C. (1980) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, 2nd ed., Wiley, New York

62. Abraham, M. H. (1980) *J. Am. Chem. Soc.* **102**, 5910–5912

63. Ben-Náim, A. (1980) *Hydrophobic Interactions*, Plenum, New York

64. Murphy, K. P., Privalov, P. L., and Gill, S. J. (1990) *Science* **247**, 559–561

65. Muller, N. (1992) *Trends Biochem. Sci.* **17**, 459–463

66. Spolar, R. S., Livingstone, J. R., and Record, J., MT. (1992) *Biochemistry* **31**, 3947–3955

67. Stites, W. E., Meeker, A. K., and Shortle, D. (1994) *J. Mol. Biol.* **235**, 27–32

68. Hecht, D., Tadesse, L., and Walters, L. (1993) *J. Am. Chem. Soc.* **115**, 3336–3337

69. Gill, S. J., and Wadsö, I. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2955–2958

70. Sharp, K. A., Nicholls, A., Fine, R. F., and Honig, B. (1991) *Science* **252**, 106–109

71. Van Oss, C. J. (1994) *Interfacial Forces in Aqueous Media*, Dekker, New York

71a. Hummer, G., Garde, S., García, A. E., Pohorille, A., and Pratt, L. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8951–8955

72. Israelachvili, J., and Wennerström, H. (1996) *Nature (London)* **379**, 219–225

73. Radzicka, A., and Wolfenden, R. (1988) *Biochemistry* **27**, 1664–1670

74. Kellis, J. T., Jr., Nyberg, K., Sali, D., and Fersht, A. R. (1988) *Nature (London)* **333**, 784–786

75. Sanger, F. (1988) *Ann. Rev. Biochem.* **57**, 1–28

76. Creighton, T. E. (1993) *Proteins, Structures and Molecular Principles*, 2nd ed., Freeman, New York

77. Kyte, J. (1995) *Structure in Protein Chemistry*, Garland Publ., New York

78. Guillemin, R., and Burgus, R. (1972) *Sci. Am.* **227**(Nov), 24–33

79. Hochster, R. M., Kates, M., and Quastel, J. H., eds. (1973) *Metabolic Inhibitors*, Vol. 3, Academic Press, New York (p. 312)

80. Ramachandran, G. N., and Kolaskar, A. S. (1973) *Biochim. Biophys. Acta.* **303**, 385–388

81. MacArthur, M. W., and Thornton, J. M. (1996) *J. Mol. Biol.* **264**, 1180-1195

81a. Milner-White, E. J. (1997) *Protein Sci.* **6**, 2477–2482

82. Moffat, J. B. (1973) *J. Theor. Biol.* **40**, 247–258

83. Fersht, A. R. (1971) *J. Am. Chem. Soc.* **93**, 3504–3515

84. Karle, I. L. (1974) *Biochemistry* **13**, 2155–2162

85. Ovchinnikov, Y. A. (1973) *FEBS Lett.* **29**, 31–33

86. Doonan, S. (1974) *FEBS Lett.* **38**, 229–233

87. Bhown, A. S. (1990) *Handbook of Proteins*, A & M Publ., Birmingham, Alabama

88. Boguski, M., and McEntyre, J. (1994) *Trends Biochem. Sci.* **19**, 71

88a. Branden, C., and Tooze, J. (1991) *Introduction to Protein Structure*, Garland Publ., New York

89. Clements, S., Mehansho, H., and Carlson, D. M. (1975) *J. Biol. Chem.* **260**, 13471–13477

90. Condit, C. M., and Meagher, R. B. (1986) *Nature (London)* **323**, 178–181

91. Reynolds, P., Weber, S., and Prakash, L. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 168–172

92. Rodakis, G. C., and Kafatos, F. C. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3551–3555

93. Young, J. F., Hockmeyer, W. T., Gross, M., Ballou, W. R., Wirtz, R. A., Trosper, J. H., Beaudoin, R. L., Hollingdale, M. R., Miller, L. H., Diggs, C. L., and Rosenberg, M. (1985) *Science* **228**, 958–962

References

94. Doolittle, R. F. (1987) *Of URFS and ORFS: A Primer on How to Analyze Derived Amino Acid Sequences*, Univ. Science Books, Mill Valley, California
95. Doolittle, R. F. (1989) *Trends Biochem. Sci.* **14**, 244–245
96. Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2444–2448
97. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233**, 123–138
98. Lewin, R. (1987) *Science* **237**, 1570
99. Fleming, P. J., and Richards, F. M. (2000) *J. Mol. Biol.* **299**, 487–498
100. Gekko, K., and Hasegawa, Y. (1986) *Biochemistry* **25**, 6563–6571
101. Goodsell, D. S., and Olson, A. J. (1993) *Trends Biochem. Sci.* **18**, 65–68
102. Schultz, G. E., and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York

102a. Ramachandran, G. N., Ramakrishnam, C., and Sasisekharan, V. (1963) *J. Mol. Biol.*, **7**, 95–99

103. Ramachandran, G. N., Venkatachalam, C. M., and Krimm, S. (1966) *Biophys. J.* **6**, 849–872

103a. Karplus, P. A. (1996) *Protein Sci.* **5**, 1406–1420

104. Sasisekharian, V., Lakshminarayanan, A. V., and Ramachandran, G. N. (1967) in *Conformation of Biopolymers*, Vol. 2 (Ramachandran, G. N., ed), p. 641–654, Academic Press, New York

104a. Gunasekaran, K., Ramakrishnan, C., and Balaram, P. (1996) *J. Mol. Biol.* **264**, 191–198

104b. Herzberg, O., and Moult, J. (1991) *Proteins* **11**, 223–229

105. Milner-White, E. J., Bell, L. H., and Maccallum, P. H. (1992) *J. Mol. Biol.* **228**, 725 – 734
106. Rhee, S. (1990) *High resolution X-ray diffraction studies of cytosolic aspartate aminotransferase* [Ph.D. Dissertation], University of Iowa, Iowa City, Iowa
107. Dunbrack, R. L., Jr., and Karplus, M. (1993) *J. Mol. Biol.* **230**, 543–574
108. Pauling, L., and Corey, R. B. (1951) *Proc. Natl. Acad. Sci. U.S.A.* **37**, 729–740
109. Rich, A. (1994) *Nature (London)* **371**, 285
110. Chothia, C. (1973) *J. Mol. Biol.* **75**, 295–302
111. Chou, K.-C., Némethy, G., and Scheraga, H. A. (1983) *Biochemistry* **22**, 6213–6221

111a. Skarzynski, T., Moody, P. C. E., and Wonacott, A. J. (1987) *J. Mol. Biol.* **193**, 171–187

111b. Wang, L., O'Connell, T., Tropsha, A., and Hermans, J. (1996) *J. Mol. Biol.* **262**, 283–293

112. Reeke, G. N., Jr., and Becker, J. W. (1986) *Science* **234**, 1108–1111
113. Murzin, A. G., Lesk, A. M., and Chothia, C. (1994) *J. Mol. Biol.* **236**, 1369–1381
114. Murzin, A. G., Lesk, A. M., and Chothia, C. (1994) *J. Mol. Biol.* **236**, 1382–1400
115. LaLonde, J. M., Bernlohr, D. A., and Ranaszak, L. J. (1994) *FASEB J.* **8**, 1240–1241
116. Guss, J. M., and Freeman, H. C. (1983) *J. Mol. Biol.* **169**, 521–563
117. Richardson, J. S. (1981) *Adv. Prot. Chem.* **34**, 167–339

117a. Erickson, H. P. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10114–10118

118. Redinbo, M. R., Cascio, D., Choukair, M. K., Rice, D., Merchant, S., and Yeates, T. O. (1993) *Biochemistry* **32**, 10560–10567
119. Weiss, M. S., and Schulz, G. E. (1993) *J. Mol. Biol.* **231**, 817–824
120. Kreusch, A., and Schulz, G. E. (1994) *J. Mol. Biol.* **243**, 891–905
121. Yoder, M. D., and Jurnak, F. (1995) *FASEB J.* **9**, 335–342
122. Richardson, J. S. (1994) *FASEB J.* **8**, 1237–1239
123. Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) *Science* **265**, 383–385
124. Richardson, J. S., Getzoff, E. D., and Richardson, D. C. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2574–2578
125. Chan, A. W. E., Hutchinson, E. G., Harris, D., and Thornton, J. M. (1993) *Protein Sci.* **2**, 1574–1590
126. Urry, D. W. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1610–1614
127. Nicholson, L. K., and Cross, T. A. (1989) *Biochemistry* **28**, 9379–9385

127a. Smith, T. F., Gaitatzes, C., Saxena, K., and Neer, E. J. (1999) *Trends Biochem. Sci.* **24**, 181–185

127b. Garcia-Higuera, I., Fenoglio, J., Li, Y., Lewis, C., Panchenko, M. P., Reiner, O., Smith, T. F., and Neer, E. J. (1996) *Biochemistry* **35**, 13985–13994

127c. ter Naar, E., Harrison, S. C., and Kirchhausen, T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1096–1100

127d. Springer, T. A. (1998) *J. Mol. Biol.* **283**, 837–862

128. Pauling, L. (1993) *Protein Sci.* **2**, 1060–1063
129. Kendrew, J. C. (1969) *Sci. Am.* **205**(Dec), 96–110
130. Arnott, S., and Dover, S. D. (1967) *J. Mol. Biol.* **30**, 209–212
131. Steigemann, W., and Weber, E. (1979) *J. Mol. Biol.* **127**, 309–338
132. Barlow, D. J., and Thornton, J. M. (1988) *J. Mol. Biol.* **201**, 601–619

132a. Rajashankar, K. R., and Ramakumar, S. (1996) *Protein Sci.* **5**, 932 –946

133. Harper, E. T., and Rose, G. D. (1993) *Biochemistry* **32**, 7605–7609
134. Chakrabartty, A., Schellman, J. A., and Baldwin, R. L. (1991) *Nature (London)* **351**, 586–588
135. Creamer, T. P., and Rose, G. D. (1995) *Protein Sci. U.S.A.* **4**, 1305–1314
136. Dempsey, C. E. (1992) *Biochemistry* **31**, 4705–4712
137. Ramachandran, G. N., and Mitra, A. K. (1976) *J. Mol. Biol.* **107**, 85–92
138. Blaber, M., Zhang, X.-j, and Matthews, B. W. (1993) *Science* **260**, 1637–1639
139. Chakrabartty, A., Kortemme, T., and Baldwin, R. L. (1994) *Protein Sci.* **3**, 843–852
140. Toniolo, C., and Benedetti, E. (1991) *Trends Biochem. Sci.* **16**, 350–353
141. Avbelj, F., and Moult, J. (1995) *Biochemistry* **34**, 755–764
142. Karpen, M. E., De Haseth, P. L., and Neet, K. E. (1992) *Protein Sci.* **1**, 1333–1342
143. Fiori, W. R., Miick, S. M., and Millhauser, G. L. (1993) *Biochemistry* **32**, 11957–11962
144. Smythe, M. L., Huston, S. E., and Marshall, G. R. (1993) *J. Am. Chem. Soc.* **115**, 11594–11595
145. Bavosa, A., Benedetti, E., Di Blasio, B., Pavone, V., Pedone, C., Toniolo, C., and Bonora, G. M. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1988–1992

145a. Fathallah-Shaykh, H., Wolf, S., Wong, E., Posner, J. B., and Furneaux, H. M. (1991) *Proc. Natl. Acad. Sci. U.S.A* **88**, 3451–3454

146. Miick, S. M., Martinez, G. V., Flori, W. R., Todd, A. P., and Millhauser, G. L. (1992) *Nature (London)* **359**, 653–655
147. Wada, A. (1976) *Adv. Biophys.* **9**, 1–63
148. Hol, W. G. J., van Duijnen, P. T., and Berendsen, H. J. C. (1978) *Nature (London)* **273**, 443–446

148a. Hol, W. G. J., Halie, L. M., and Sander, C. (1981) *Nature (London)* **294**, 532–536

148b. Lockhart, D. J., and Kim, P. S. (1992) *Science* **257**, 947–951

149. Levine, B. F., and Bethea, C. G. (1976) *J. Chem. Phys.* **65**, 1989–1993
150. Gray, T. M., and Matthews, B. W. (1984) *J. Mol. Biol.* **175**, 75–81
151. Richardson, J. S., and Richardson, D. C. (1988) *Science* **240**, 1648–1652
152. Nicholson, H., Anderson, D. E., Dao-pin, S., and Matthews, B. W. (1991) *Biochemistry* **30**, 9816–9828
153. Åqvist, J., Luecke, H., Quiocho, F. A., and Warshel, A. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2026–2030
154. Lyu, P. C., Wemmer, D. E., Zhou, H. X., Pinker, R. J., and Kallenbach, N. R. (1993) *Biochemistry* **32**, 421–425
155. Viguera, A. R., and Serrano, L. (1995) *J. Mol. Biol.* **251**, 150–160
156. Doig, A. J., Chakrabartty, A., Klingler, T. M., and Baldwin, R. L. (1994) *Biochemistry* **33**, 3396–3403
157. Doig, A. J., Macarthur, M. W., Stapley, B. J., and Thornton, J. M. (1997) *Protein Sci.* **6**, 147–155
158. Sali, D., Bycroft, M., and Fersht, A. R. (1988) *Nature (London)* **335**, 740–743
159. Aurora, R., and Rose, G. D. (1998) *Protein Sci.* **7**, 21–38
160. Metzler, D. E. (1979) *Adv. Enzymol.* **50**, 1–40
161. Lupas, A. (1996) *Trends Biochem. Sci.* **21**, 375–382
161. Lupas, A. (1996) *Trends Biochem. Sci.* **21**, 375–382
162. Kohn, W. D., Mant, C. T., and Hodges, R. S. (1997) *J. Biol. Chem.* **272**, 2583–2586
163. Fraser, R. D. B., MacRae, T. P., Parry, D. A. D., and Suzuki, E. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1179–1183
164. Johnson, L. D., Idler, W. W., Zhou, X.-M., Roop, D. R., and Steinert, P. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1896–1900
165. Cooper, D., and Sun, T.-T. (1986) *J. Biol. Chem.* **261**, 4646–4654
166. Crick, F. H. C. (1953) *Acta Crystallogr.* **6**, 689–697
167. Vazquez, S. R., Kuo, D. Z., Bositis, C. M., Hardy, L. W., Lew, R. A., and Humphreys, R. E. (1992) *J. Biol. Chem.* **267**, 7406–7410
168. O'Shea, E. K., Llemm, J. D., Kim, P. S., and Alber, T. (1991) *Science* **254**, 539–544
169. Ferré-D'Amaré, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) *Nature (London)* **363**, 38–44
170. Lyu, P. C., Marky, L. A., and Kallenbach, N. R. (1989) *J. Am. Chem. Soc.* **111**, 2733–2734
171. Graddis, T. J., Myszka, D. G., and Chaiken, I. M. (1993) *Biochemistry* **32**, 12664–12671

171a. Fairman, R., Chao, H.-G., Lavoie, T. B., Villafranca, J. J., Matsueda, G. R., and Novotny, J. (1996) *Biochemistry* **35**, 2824–2829

171b. Kohn, W. D., Kay, C. M., and Hodges, R. S. (1998) *J. Mol. Biol.* **283**, 993–1012

172. John, M., Briand, J.-P., Granger-Schnarr, M., and Schnarr, M. (1994) *J. Biol. Chem.* **269**, 16247–16253
173. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992) *Biochemistry* **31**, 5739–5746
174. Cohen, C., and Parry, D. A. D. (1986) *Trends Biochem. Sci.* **11**, 245–248
175. Phillips, G. N., Jr., Fillers, J. P., and Cohen, C. (1986) *J. Mol. Biol.* **192**, 111–131
176. Mo, J., Holtzer, M. E., and Holtzer, A. (1993) *Protein Sci.* **2**, 128–130
177. Xie, X., Rao, S., Walian, P., Hatch, V., Phillips, G. N., Jr., and Cohen, C. (1994) *J. Mol. Biol.* **236**, 1212–1226
178. Eckert, R. L. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1114–1118
179. Gernert, K. M., Surles, M. C., Labean, T. H., Richardson, J. S., and Richardson, D. C. (1995) *Protein Sci.* **4**, 2252 –2260

179a. Monera, O. D., Zhou, N. E., Lavigne, P., Kay, C. M., and Hodges, R. S. (1996) *J. Biol. Chem.* **271**, 3995–4001

180. Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., and Eisenberg, D. (1993) *Science* **259**, 1288–1293

References

181. Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature (London)* **371**, 80–83

182. Efimov, V. P., Nepluev, I. V., Sobolev, B. N., Zurabishvili, T. G., Schulthess, T., Lustig, A., Engel, J., Haener, M., Aebi, U., Venyaminov, S. Y., Potekhin, S. A., and Mesyanzhinov, V. V. (1994) *J. Mol. Biol.* **242**, 470–486

183. Lieberman, M., Tabet, M., and Sasaki, T. (1994) *J. Am. Chem. Soc.* **116**, 5035–5044

184. Utani, A., Nomizu, M., Timpl, R., Roller, P. P., and Yamada, Y. (1994) *J. Biol. Chem.* **269**, 19167–19175

185. Robinson, C. R., and Sligar, S. G. (1993) *Protein Sci.* **2**, 826–837

186. Steif, C., Weber, P., Hinz, H.-J., Flossdorf, J., Cesareni, G., and Kokkinidis, M. (1993) *Biochemistry* **32**, 3867–3876

187. Redfield, C., Boyd, J., Smith, L. J., Smith, R. A. G., and Dobson, C. M. (1992) *Biochemistry* **31**, 10431–10437

188. Harris, N. L., Presnell, S. R., and Cohen, F. E. (1994) *J. Mol. Biol.* **236**, 1356–1368

189. Chou, K.-C., Carlacci, L., Maggiora, G. M., Parodi, L. A., and Schulz, M. W. (1992) *Protein Sci.* **1**, 810–827

190. Rees, D. C., Komiya, H., Yeates, T. O., Allen, J. P., and Feher, G. (1989) *Ann. Rev. Biochem.* **58**, 607–633

191. Aleshin, A. E., Hoffman, C., Firsov, L. M., and Honzatko, R. B. (1994) *J. Mol. Biol.* **238**, 575–591

192. Aleshin, A. E., Firsov, L. M., and Honzatko, R. B. (1994) *J. Biol. Chem.* **269**, 15631–15639

193. Harris, E. M. S., Aleshin, A. E., Firsov, L. M., and Honzatko, R. B. (1993) *Biochemistry* **32**, 1618–1626

194. Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994) *Science* **266**, 75–81

194a. Kramer, R. Z., Bella, J., Mayville, P., Brodsky, B., and Berman, H. M. (1999) *Nature Struct. Biol.* **6**, 454–457

195. Brodsky, B., and Shah, N. K. (1995) *FASEB J.* **9**, 1537–1546

196. Redeker, V., Levilliers, N., Schmitter, J.-M., Le Caer, J.-P., Rossier, J., Adoutte, A., and Bré, M.-H. (1994) *Science* **266**, 1688–1691

197. Miller, A. (1982) *Trends Biochem. Sci.* **7**, 13–18

198. Eyre, D. R. (1980) *Science* **207**, 1315–1322

199. Parry, D. A. D., and Craig, A. S. (1979) *Nature (London)* **282**, 213–215

200. Porter, R. R., and Reid, K. B. M. (1979) *Adv. Prot. Chem.* **33**, 1–71

201. Ashford, D., and Neuberger, A. (1980) *Trends Biochem. Sci.* **5**, 245–248

202. Chen, J., and Varner, J. E. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4399–4403

203. Adzhubei, A. A., and Sternberg, M. J. E. (1994) *Protein Sci.* **3**, 2395–2410

204. Richardson, J. S. (1985) *Nature (London)* **316**, 102–103

205. Hutchinson, E. G., and Thornton, J. M. (1994) *Protein Sci.* **3**, 2207–2216

206. Venkatachalam, C. M. (1968) *Biopolymers* **6**, 1425–1436

207. Perczel, A., Foxman, B. M., and Fasman, G. D. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8210–8214

208. Yan, Y., Tropsha, A., Hermans, J., and Erickson, B. W. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7898–7902

209. Perczel, A., McAllister, M. A., Császár, P., and Csizmadia, I. G. (1993) *J. Am. Chem. Soc.* **115**, 4849–4858

210. Mattos, C., Petsko, G. A., and Karplus, M. (1994) *J. Mol. Biol.* **238**, 733–747

211. Hynes, T. R., Hodel, A., and Fox, R. O. (1994) *Biochemistry* **33**, 5021–5030

212. Haque, T. S., Little, J. C., and Gellman, S. H. (1994) *J. Am. Chem. Soc.* **116**, 4105–4106

213. Gierasch, L. M., Deber, C. M., Madison, V., Niu, C.-H., and Blout, E. R. (1981) *Biochemistry* **20**, 4730–4738

214. Rees, D. C., Lewis, M., and Lipscomb, W. N. (1983) *J. Mol. Biol.* **168**, 367–387

215. Sapse, A.-M., Mallah-Levy, L., Daniels, S. B., and Erickson, B. W. (1987) *J. Am. Chem. Soc.* **109**, 3526–3529

216. Urry, D. W. (1993) *Angew. Chem. Int. Ed. Engl.* **32**, 819–841

217. Sandberg, L. B., Soskel, N. T., and Leslie, J. G. (1981) *N. Engl. J. Med.* **304**, 566–579

217a. Tatham, A. S., and Shewry, P. R. (2000) *Trends Biochem. Sci.* **25**, 567–571

218. Gotte, L., Volpin, D., Horne, R. W., and Mammi, M. (1976) *Micron* **7**, 95–102

218a. Urry, D., Luan, C.-H., and Peng, S. (1995) *Ciba. Found. Symp.* **192**, 4–30

218b. van Dijk, A. A., de Boef, E., Bekkers, A., van Wijk, L. L., van Swieten, E., Hamer, R. J., and Robillard, G. T. (1997) *Protein Sci.* **6**, 649–656

219. Sternberg, M. J. E., and Thornton, J. M. (1976) *J. Mol. Biol.* **105**, 367–382

220. Leszczýnski, J. F., and Rose, G. D. (1986) *Science* **234**, 849–855

221. Liang, C., and Mislow, K. (1994) *J. Am. Chem. Soc.* **116**, 11189–11190

222. Chothia, C. (1984) *Ann. Rev. Biochem.* **53**, 537–572

223. Resnick, D., Pearson, A., and Krieger, M. (1994) *Trends Biochem. Sci.* **19**, 5–8

224. Shoyab, M., Plowman, G. D., McDonald, V. L., Bradley, J. G., and Todaro, G. J. (1989) *Science* **243**, 1074–1076

225. Doolittle, R. F. (1995) *Ann. Rev. Biochem.* **64**, 287–314

226. Berman, A. L., Kolker, E., and Trifanov, E. N. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4044–4047

227. Orengo, C. A., Jones, D. T., and Thornton, J. M. (1994) *Nature (London)* **372**, 631–634

228. Pan, K.-M., Damodaran, S., and Greaser, M. L. (1994) *Biochemistry* **33**, 8255–8261

229. Doolittle, R. F., and Bork, P. (1993) *Sci. Am.* **269**(Oct), 50–56

230. Baron, M., Norman, D. G., and Campbell, I. D. (1991) *Trends Biochem. Sci.* **16**, 13–17

231. Mulichak, A. M., Tulinsky, A., and Ravichandran, K. G. (1991) *Biochemistry* **30**, 10576–10588

232. Bork, P., Holm, L., and Sander, C. (1994) *J. Mol. Biol.* **242**, 309–320

232a. Bogan, A. A., and Thorn, K. S. (1998) *J. Mol. Biol.* **280**, 1–9

233. Tsai, J., Taylor, R., Chothia, C., and Gerstein, M. (1999) *J. Mol. Biol.* **290**, 253–266

233a. Gerstein, M., and Chothia, C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10167–10172

234. Varadarajan, R., and Richards, F. M. (1992) *Biochemistry* **31**, 12315–12327

235. Allewell, N. (1987) *Trends Biochem. Sci.* **12**, 417–418

236. Lim, W. A., Farruggio, D. C., and Sauer, R. T. (1992) *Biochemistry* **31**, 4324–4333

237. Hubbard, S. J., and Argos, P. (1994) *Protein Sci.* **3**, 2194–2206

237a. Hubbard, S. J., and Argos, P. (1996) *J. Mol. Biol.* **261**, 289–300

238. Hunt, N. G., Gregoret, L. M., and Cohen, F. E. (1994) *J. Mol. Biol.* **241**, 214–225

239. Hunter, C. A., and Sanders, K. M. (1990) *J. Am. Chem. Soc.* **112**, 5525–5534

240. Burley, S. K., and Petsko, G. A. (1985) *Science* **229**, 23–28

241. Thomas, K. A., Smith, G. M., Thomas, T. B., and Feldmann, R. J. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4843–4847

242. Flocco, M. M., and Mowbray, S. L. (1994) *J. Mol. Biol.* **235**, 709–717

243. Hemmingsen, J. M., Gernert, K. M., Richardson, J. S., and Richardson, D. C. (1994) *Protein Sci.* **3**, 1927–1937

244. Teeter, M. M. (1991) *Ann. Rev. Biophys. Biophys. Chem.* **20**, 577–600

245. McDonald, I. K., and Thornton, J. M. (1994) *J. Mol. Biol.* **238**, 777–793

246. Bordo, D., and Argos, P. (1994) *J. Mol. Biol.* **243**, 504–519

247. Shimoni, L., and Glusker, J. P. (1995) *Protein Sci.* **4**, 65–74

248. Dao-pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., and Matthews, B. W. (1991) *Biochemistry* **30**, 11521–11529

248a. Kumar, S., and Nussinov, R. (1999) *J. Mol. Biol.* **293**, 1241–1255

248b. Strop, P., and Mayo, S. L. (2000) *Biochemistry* **39**, 1251–1255

249. Sawyer, L., and James, M. N. G. (1982) *Nature (London)* **295**, 79–80

250. Borders, C. L., Jr., Broadwater, J. A., Bekeny, P. A., Salmon, J. E., Lee, A. S., Eldridge, A. M., and Pett, V. B. (1994) *Protein Sci.* **3**, 541–548

251. Mrabet, N. T., Van den Broeck, A., Van den brande, I., Stanssens, P., Laroche, Y., Lambeir, A.-M., Matthiissens, G., Jenkins, J., Chiadmi, M., van Tilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., De Maeyer, M., and Wodak, S. J. (1992) *Biochemistry* **31**, 2239–2253

252. Dougherty, D. A. (1996) *Science* **271**, 163–168

253. Mitchell, J. B. O., Nandi, C. L., McDonald, I. K., Thornton, J. M., and Price, S. (1994) *J. Mol. Biol.* **239**, 315–331

253a. De Wall, S. L., Meadows, E. S., Barbour, L. J., and Gokel, G. W. (1999) *J. Am. Chem. Soc.* **121**, 5613–5614

253b. Gallivan, J. P., and Dougherty, D. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9459–9464

253c. Wouters, J. (1998) *Protein Sci.* **7**, 2472–2475

253d. Fernández-Recio, J., Romero, A., and Sancho, J. (1999) *J. Mol. Biol.* **290**, 319–330

254. Rose, G. D., Young, W. B., and Gierasch, L. M. (1983) *Nature (London)* **304**, 654–657

255. Watenpaugh, K. D., Margulis, T. N., Sieker, L. C., and Jensen, L. H. (1978) *J. Mol. Biol.* **122**, 175–190

255a. Nakasako, M. (1999) *J. Mol. Biol.* **289**, 547–564

256. Teeter, M. M. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6014–6018

257. Otting, G., Liepinsh, E., and Wüthrich, K. (1991) *Science* **254**, 974–980

258. Murzin, A. G., and Finkelstein, A. V. (1988) *J. Mol. Biol.* **204**, 749–770

259. Chothia, C. (1989) *Nature (London)* **337**, 204–205

260. Buehner, M., Ford, G. C., Moras, D., Olsen, K. W., and Rossmann, M. G. (1974) *J. Mol. Biol.* **90**, 25–49

261. Rossmann, M. G., Liljas, A., Branden, C.–I., and Banaszak, L. J. (1975) *The Enzymes*, 3rd ed., Vol. 11, Academic Press, New York (p. 61)

262. Brändén, C. (1990) *Nature (London)* **346**, 607–608

263. Kobe, B., and Deisenhofer, J. (1995) *Nature (London)* **374**, 183–186

264. Kobe, B., and Deisenhofer, J. (1994) *Trends Biochem. Sci.* **19**, 415–421

264a. Kajava, A. V. (1998) *J. Mol. Biol.* **277**, 519–527

264b. Kobe, B., and Kajava, A. V. (2000) *Trends Biochem. Sci.* **25**, 509–515

265. Farber, G. K., and Petsko, G. A. (1990) *Trends Biochem. Sci.* **15**, 228–234

266. Reardon, D., and Farber, G. K. (1995) *FASEB J.* **9**, 497–503

267. Uhlin, U., and Eklund, H. (1994) *Nature (London)* **370**, 533–539

268. Chothia, C. (1992) *Nature (London)* **357**, 543–544

268a. Zhang, C., and DeLisi, C. (1998) *J. Mol. Biol.* **284**, 1301–1305

269. McLachlan, A. D. (1980) *Nature (London)* **285**, 267–268

References

270. Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry* **13**, 222–245
270a. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1997) *Biochemistry* **36**, 10923–10929
270b. Spek, E. J., Olson, C. A., Shi, Z., and Kallenbach, N. R. (1999) *J. Am. Chem. Soc.* **121**, 5571–5572
271. Minor, D. L., Jr., and Kim, P. S. (1994) *Nature (London)* **367**, 660–663
272. Richards, F. M. (1991) *Sci. Am.* **264**(Jan), 54–63
273. Anfinsen, C. B. (1973) *Science* **181**, 223–230
274. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D., and Chan, H. S. (1995) *Protein Sci.* **4**, 561–602
275. Rost, B., Schneider, R., and Sander, C. (1993) *Trends Biochem. Sci.* **18**, 120–123
276. Wako, H., and Blundell, T. L. (1994) *J. Mol. Biol.* **238**, 693–708
277. Yi, T.-M., and Lander, E. S. (1993) *J. Mol. Biol.* **232**, 1117–1129
278. Monge, A., Friesner, R. A., and Honig, B. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5027–5029
279. Smith, C. K., Withka, J. M., and Regan, L. (1994) *Biochemistry* **33**, 5510–5517
279a. Lacroix, E., Viguera, A. R., and Serrano, L. (1998) *J. Mol. Biol.* **284**, 173–191
279b. Gotoh, O. (1996) *J. Mol. Biol.* **264**, 823–838
279c. Dandekar, T., and Argos, P. (1996) *J. Mol. Biol.* **256**, 645–660
279d. Jones, D. T. (1999) *J. Mol. Biol.* **292**, 195–202
280. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132
281. Reyes, V. E., Phillips, L., Humphreys, R. E., and Lew, R. A. (1989) *J. Biol. Chem.* **264**, 12854–12858
282. Wold, F. (1986) *Trends Biochem. Sci.* **11**, 58–59
283. Graves, D. J., Martin, B. L., and Wang, J. H. (1994) *Co-and post-translational modification of proteins*, Oxford Univ. Press, New York
283a. Angeletti, R. H., ed. (1998) *Protein, Analysis and Design*, Academic Press, San Diego, California
284. Yan, S. C. B., Grinnell, B. W., and Wold, F. (1989) *Trends Biochem. Sci.* **14**, 264–268
285. Berthiaume, L., Deichaite, I., Peseckis, S., and Resh, M. D. (1994) *J. Biol. Chem.* **269**, 6498–6505
286. McIlhinney, R. A. J. (1990) *Trends Biochem. Sci.* **15**, 387–391
287. Violand, B. N., Schlittler, M. R., Lawson, C. Q., Kane, J. F., Siegel, N. R., Smith, C. E., Kolodziej, E. W., and Duffin, K. L. (1994) *Protein Sci.* **3**, 1089–1097
288. Hilz, H., Fanick, N., and Klapproth, K. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6267–6271
289. Hortin, G., Fok, K. F., Toren, P. C., and Strauss, A. W. (1987) *J. Biol. Chem.* **262**, 3082–3085
290. Bateman, A., Solomon, S., and Bennett, H. P. J. (1990) *J. Biol. Chem.* **265**, 22130–22136
291. Takahara, H., Tsuchida, M., Kusubata, M., Akutsu, K., Tagami, S., and Sugawara, K. (1989) *J. Biol. Chem.* **264**, 13361–13368
292. Kreil, G. (1994) *J. Biol. Chem.* **269**, 10967–10970
293. Mor, A., Amiche, M., and Nicolas, P. (1992) *Trends Biochem. Sci.* **17**, 481–485
294. Kreil, G. (1994) *Science* **266**, 996–997
295. Prescott, B., Renugopalakrishnan, V., Glimcher, M. J., Bhushan, A., and Thomas, G. J., Jr. (1986) *Biochemistry* **25**, 2792–2798
296. Wood, D. D., and Moscarello, M. A. (1989) *J. Biol. Chem.* **264**, 5121–5127
297. Waite, J. H., Housley, T. J., and Tanzer, M. L. (1985) *Biochemistry* **24**, 5010–5014
298. Taylor, S. W., Waite, J. H., Ross, M. M., Shabanowitz, J., and Hunt, D. F. (1994) *J. Am. Chem. Soc.* **116**, 10803–10804
299. Clarke, S. (1992) *Ann. Rev. Biochem.* **61**, 355–386
300. Marshall, C. J. (1993) *Science* **259**, 1865–1866
301. Mroz, E. A. (1989) *Science* **243**, 1615
302. Rothgeb, T. M., England, R. D., Jones, B. N., and Gurd, R. S. (1978) *Biochemistry* **17**, 4564–4571
303. Lorimer, G. H. (1983) *Trends Biochem. Sci.* **8**, 65–68
304. Friedman, M. (1977) *Protein Crosslinking*, Vol. 2, Plenum, New York
305. Torchinsky, Y. M. (1981) *Sulfur in Proteins*, Pergamon, Oxford
306. Harrison, P. M., and Sternberg, M. J. E. (1994) *J. Mol. Biol.* **244**, 448–463
307. Folk, J. E., and Finlayson, J. S. (1977) *Adv. Prot. Chem.* **31**, 1–133
308. Gurd, F. R. N., and Rothgeb, T. M. (1979) *Adv. Prot. Chem.* **33**, 73–165
309. Karplus, M., and McCammon, J. A. (1983) *Ann. Rev. Biochem.* **53**, 263–300
310. Caspar, D. L. D., Clarage, J., Salunke, D. M., and Clarage, M. (1988) *Nature (London)* **332**, 659–662
311. Smith, J. L., Hendrickson, W. A., Honzatko, R. B., and Sheriff, S. (1986) *Biochemistry* **25**, 5018–5027
312. Frauenfelder, H., Sligar, S. G., and Wolynes, P. G. (1991) *Science* **254**, 1598–1603
313. Buck, M., Radford, S. E., and Dobson, C. M. (1994) *J. Mol. Biol.* **237**, 247–254
314. Kim, K.-S., Fuchs, J. A., and Woodward, C. K. (1993) *Biochemistry* **32**, 9600–9608
315. Rohl, C. A., and Baldwin, R. L. (1994) *Biochemistry* **33**, 7760–7767
316. Spyracopoulos, L., and O'Neil, J. D. J. (1994) *J. Am. Chem. Soc.* **116**, 1395–1402
317. Huber, R. (1979) *Trends Biochem. Sci.* **4**, 271–276
318. Lumry, R., and Rosenberg, A. (1974) *Colloques Internationaux du C.N.R.S.* **246**, 53
319. Mao, B., Pear, M. R., McCammon, J. A., and Northrup, S. H. (1981) *J. Mol. Biol.* **151**, 199–202
320. Brack, A., and Spach, G. (1981) *J. Am. Chem. Soc.* **103**, 6319–6323
321. Frauenfelder, H., Parak, F., and Young, R. D. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 451–479
322. Gerstein, M., Lesk, A. M., and Chothia, C. (1994) *Biochemistry* **33**, 6739–6749
323. Kretsinger, R. H., and Nockolds, C. E. (1973) *J. Biol. Chem.* **248**, 3313–3326
324. Liepinsh, E., and Otting, G. (1994) *J. Am. Chem. Soc.* **116**, 9670–9674
325. Scholtz, J. M., Barrick, D., York, E. J., Stewart, J. M., and Baldwin, R. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 185–189
326. Kamoun, P. P. (1988) *Trends Biochem. Sci.* **13**, 424–425
327. Dill, K. A., and Shortle, D. (1991) *Ann. Rev. Biochem.* **60**, 795–825
328. Baker, D., and Agard, D. A. (1994) *Biochemistry* **33**, 7505–7509
328a. Sharp, K. A., and Englander, S. W. (1994) *Trends Biochem. Sci.* **19**, 526–529
328b. Pace, C. N., Shirley, B. A., McNutt, M., and Gajiwala, K. (1996) *FASEB J.* **10**, 75–83
328c. Sippl, M. J. (1996) *J. Mol. Biol.* **260**, 644–648
328d. Meyer, E. (1992) *Protein Sci.* **1**, 1543–1562
328e. Hill, R. B., Hong, J.-K., and DeGrado, W. F. (2000) *J. Am. Chem. Soc.* **122**, 746–747
329. Levinthal, C. (1968) *J. Chim. Phys.-Chim. Biol.* **65**, 44–45
330. Lattman, E. E., Fiebig, K. M., and Dill, K. A. (1994) *Biochemistry* **33**, 6158–6166
331. Mark, A. E., and van Gunsteren, W. F. (1992) *Biochemistry* **31**, 7745–7748
332. Peng, Z., and Kim, P. S. (1994) *Biochemistry* **33**, 2136–2141
333. Hagihara, Y., Tan, Y., and Goto, Y. (1994) *J. Mol. Biol.* **237**, 336–348
334. Yu, Y., Makhatadze, G. I., Pace, C. N., and Privalov, P. L. (1994) *Biochemistry* **33**, 3312–3319
335. Dobson, C. M., Evans, P. A., and Radford, S. E. (1994) *Trends Biochem. Sci.* **19**, 31–37
336. Yang, A.-S., and Honig, B. (1994) *J. Mol. Biol.* **237**, 602–614
336a. Baker, D. (2000) *Nature (London)* **405**, 39–42
336b. Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000) *Trends Biochem. Sci.* **25**, 331–339
337. Udgaonkar, J. B., and Baldwin, R. L. (1988) *Nature (London)* **335**, 694–699
338. Roder, H., Elöve, G. A., and Englander, S. W. (1988) *Nature (London)* **335**, 700–704
339. Gething, M.-J., and Sambrook, J. (1992) *Science* **355**, 33–45
340. Odefey, C., Mayr, L. M., and Schmid, F. X. (1995) *J. Mol. Biol.* **245**, 69–78
341. Chazin, W. J., Kördel, J., Drakenberg, T., Thulin, E., Brodin, P., Grundström, T., and Forsén, S. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2195–2198
342. Lin, L.-N., and Brandts, J. F. (1987) *Biochemistry* **26**, 3537–3543
343. Herning, T., Yutani, K., Taniyama, Y., and Kikuchi, M. (1991) *Biochemistry* **30**, 9882–9891
344. Chatrenet, B., and Chang, J.-Y. (1993) *J. Biol. Chem.* **268**, 20988–20996
345. Talluri, S., Rothwarf, D. M., and Scheraga, H. A. (1994) *Biochemistry* **33**, 10437–10449
346. Pace, C. N. (1990) *Trends Biochem. Sci.* **15**, 14–17
347. Yang, A.-S., and Honig, B. (1993) *J. Mol. Biol.* **231**, 459–474
348. Hartsough, D. S., and Merz, K. M., Jr. (1993) *J. Am. Chem. Soc.* **115**, 6529–6537
349. Aswad, D. W., and Johnson, B. A. (1987) *Trends Biochem. Sci.* **12**, 155–158
350. Brennan, T. V., and Clarke, S. (1993) *Protein Sci.* **2**, 331–338
351. Volkin, D. B., and Klibanov, A. M. (1987) *J. Biol. Chem.* **262**, 2945–2950
352. Wells-Knecht, M. C., Huggins, T. G., Dyer, D. G., Thorpe, S. R., and Baynes, J. W. (1993) *J. Biol. Chem.* **268**, 12348–12352
353. Luthra, M., Ranganathan, D., Ranganathan, S., and Balasubramanian, D. (1994) *J. Biol. Chem.* **269**, 22678–22682
354. Lubec, G., Weninger, M., and Anderson, S. R. (1994) *FASEB J.* **8**, 1166–1169
355. Kaiser, E. T. (1987) *Trends Biochem. Sci.* **12**, 305–309
355a. Beasley, J. R., and Hecht, M. H. (1997) *J. Biol. Chem.* **272**, 2031–2034
356. Oxender, D. E., and Fox, C. F., eds. (1987) *Protein Engineering*, Liss, New York
356a. Dahiyat, B. I., and Mayo, S. L. (1997) *Science* **278**, 82–87
356b. DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F., and Lombardi, A. (1999) *Ann. Rev. Biochem.* **68**, 779–819
357. Bryson, J. W., Betz, S. F., Lu, H. S., Suich, D. J., Zhou, H. X., O'Neil, K. T., and DeGrado, W. F. (1995) *Science* **270**, 935–941
358. Nautiyal, S., Woolfson, D. N., King, D. S., and Alber, T. (1995) *Biochemistry* **34**, 11645–11651
359. Betz, S. F., Liebman, P. A., and DeGrado, W. F. (1997) *Biochemistry* **36**, 2450–2458
360. Handel, T. M., Williams, S. A., and DeGrado, W. F. (1993) *Science* **261**, 879–885
361. Struthers, M. D., Cheng, R. P., and Imperiali, B. (1996) *Science* **271**, 342–345
362. Yan, Y., and Erickson, B. W. (1994) *Protein Sci.* **3**, 1069–1073
363. Quinn, T. P., Tweedy, N. B., Williams, R. W., Richardson, J. S., and Richardson, D. C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8747–8751
364. Hecht, M. H. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8729–8730
364a. Kortemme, T., Ramírez-Alvarado, M., and Serrano, L. (1998) *Science* **281**, 253–256

References

365. Kellis, J. T., Jr., Nyberg, K., and Fersht, A. R. (1989) *Biochemistry* **28**, 4914–4922
366. Goldenberg, D. P. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 481–507
367. Alber, T. (1989) *Ann. Rev. Biochem.* **58**, 765–798
368. Katz, B. A., and Kossiakoff, A. (1986) *J. Biol. Chem.* **261**, 15480–15485
369. Pantoliano, M. W., Ladner, R. C., Bryan, P. N., Rollence, M. L., Wood, J. F., and Poulos, T. L. (1987) *Biochemistry* **26**, 2077–2082
370. Wetzel, R. (1987) *Trends Biochem. Sci.* **12**, 478–482
371. Mitchinson, C., and Wells, J. A. (1989) *Biochemistry* **28**, 4807–4815
372. Ahern, T. J., Casal, J. I., Petsko, G. A., and Klibanov, A. M. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 675–679
373. Estell, D. A., Graycar, T. P., Miller, J. V., Powers, D. B., Burnier, J. P., Ng, P. G., and Wells, J. A. (1986) *Science* **233**, 659–663
374. Loeb, D. D., Swanstrom, R., Everitt, L., Manchester, M., Stamper, S. E., and Hutchison, C. A., III. (1989) *Nature (London)* **340**, 397–440
375. Mendel, D., Ellman, J. A., Chang, Z., Veenstra, D. L., Kollman, P. A., and Schultz, P. G. (1992) *Science* **266**, 1798–1802
376. Brady, L., and Dodson, G. (1994) *Nature (London)* **368**, 692–694
377. Liu, N., Deillon, C., Klauser, S., Gutte, B., and Thomas, R. M. (1998) *Protein Sci.* **7**, 1214–1220

Study Questions

1. Name all of the isometric tripeptides which could be formed from one molecule each of tyrosine, alanine, and valine.

2. What functional groups are found in protein side chains? Of what importance to protein structure and function are (a) hydrophobic groups, (b) acidic and basic groups, (c) sulfyhydryl groups?

3. If placed in water and adjusted to a pH of 7, will the following migrate toward the anode or the cathode if placed in an electrical field?
   (a) Aspartic acid, (b) alanine, (c) tyrosine, (d) lysine, (e) arginine, (f) glutamine.

4. Draw the following hydrogen-bonded structures:
   (a) A dimer of acetic acid.
   (b) A tyrosine–carboxylate bond in the interior of a protein.
   (c) A phosphate–guanidinium ion pair in an enzyme–substrate complex.

5. Contrast the properties of the amino acids with those of the saturated fatty acids with respect to solubility in water and in ether and to physical state.

6. Describe in as much detail as you can the characteristic properties of (a) β sheets, (b) α helices, (c) turns in peptide chains, and (d) collagen.

7. Predict whether the following peptide segments will be likely to exist as an α helix or as part of a β structure within a protein:
   (a) Poly-L-leucine
   (b) Poly-L-valine
   (c) Pro-Glu-Met-Val-Phe-Asp-Ile
   (d) Pro-Glu-Ala-Leu-Phe-Ala-Ala

8. Describe three ways in which a side chain of a serine residue can fold back and hydrogen bond to a C=O or N–H group of the backbone and two ways by which an asparagine side chain can do the same. There are yet other possibilities.

9. Compare structural features and properties of the following proteins: silk fibroin, α-keratin, collagen, and bovine serum albumin.

10. In what way do the solubilities of proteins usually vary with pH? Why?

11. Compare the following: the diameters of (a) a carbon atom in an organic molecule (b) a bacterial cell, e.g. of *E. coli* (c) a human red blood cell (d) a ribosome (e) the length of a peptide unit in an extended polypeptide chain (f) the length of the carbon atom chain in an 18-carbon fatty acid.

12. Compare: (a) the length of a peptide unit (residue) in a polypeptide in an extented (β) conformation. (b) the length by which an α helix is extended by the addition of one amino acid unit (c) the length of one turn of an α helix. (d) the diameter of an α helix (both using atom centers in the backbones and using van der Waals radii) for a poly-L-alanine helix.

13. What are: albumins, globulins, protamines, scleroproteins, glycoproteins, lipoproteins?

14. Where are the following found and what are their functions? Gamma globulin, hemocyanin, pepsin, glucagon, ferritin, phosphorylase.

15. List the nutritionally essential amino acids for human beings. Compare these needs with those of other species, including lactic acid bacteria, malaria parasites, green plants, etc.

16. Define: chiral, enantomer, diastereomer, epimer, anomer (see Chapter 4), prochiral (see Chapter 9). What is meant by the statement that biochemical reactions are stereochemically specific? Why is such stereospecificity to be expected in organisms (which are constructed from asymmetric units)? See Chapter 9 for further discussion.

17. What are disulfide bridges and of what significance are they in protein structure?

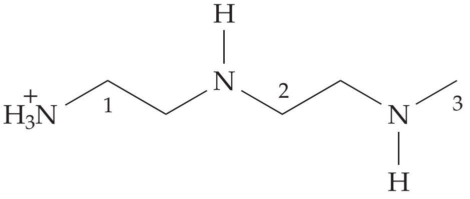
18. What is meant by "denaturation" of a protein? Mention several ways in which denaturation can be brought about. How is denaturation explained in terms of structure?

19. A chain of L-amino acids can form either a right-handed or a left-handed helix. From the Ramachandran diagram in Fig. 2-9, can you say anything about the relative stabilities of right and left-handed helices? What do you predict for polyglycine?

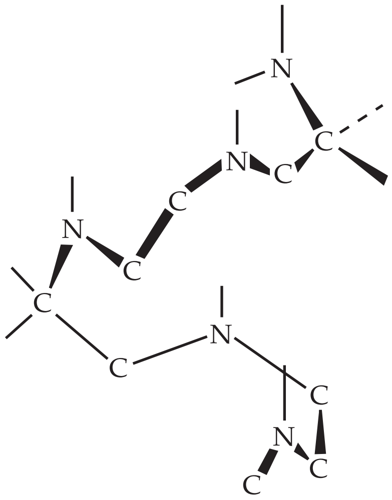
20. What similarities and differences would you predict for two proteins of identical amino acid sequence but one made from all L-amino acids and the other from all D-amino acids?

Study Questions

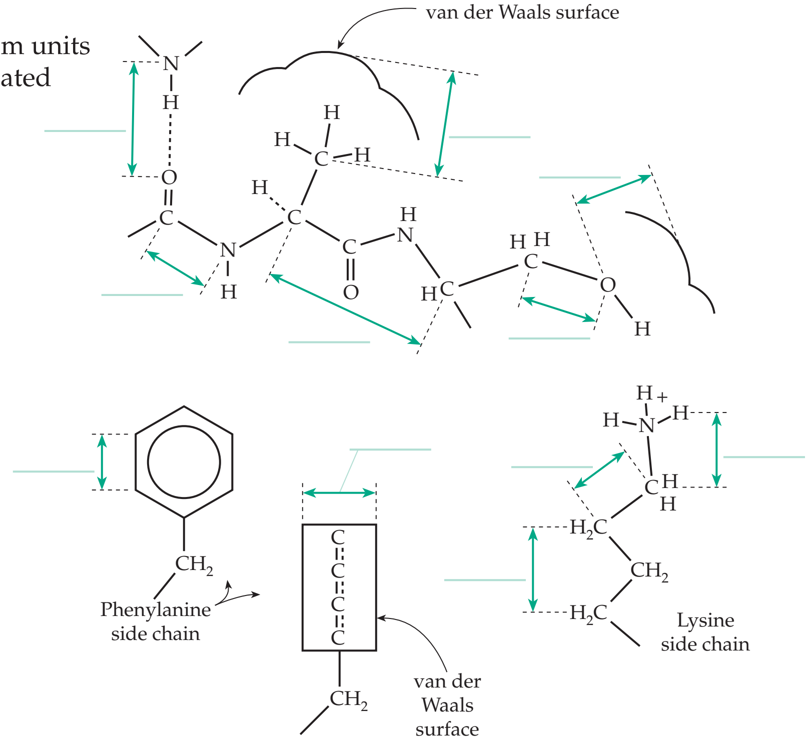
21. Complete the following peptide structure for L-seryl-L-valyl-L-asparaginyl—etc. Extend the chain in the C-terminal direction to form a **beta turn** with the chain coming back to form a **beta sheet**. Add a third segment of peptide *parallel* to the folded back chain to form a 3-stranded beta sheet. Indicate all hydrogen bonds correctly. Draw the side chains of the seryl and asparaginyl residues in positions 1 and 3 so that they form proper hydrogen bonds to groups in the peptide backbone.


H
N
$H_3\overset{+}{N}$
1
2
3
N
H

22. Complete the following structure to form a short segment of alpha helix. Extend the chain by at least three residues. Form all hydrogen bonds correctly. Add at least three side chains with the correct chirality at the alpha carbon positions. Add one electrically charged side chain and show how it interacts with the peptide backbone at either the C- or N-terminus to help *stabilize* the helix.


N
C
N
C
C
N
C
C
C
N
C
N
C
C

23. What is the relationship between Ångstrom units (Å) and nanometers (nm)? Give the indicated distance in Å or nm.

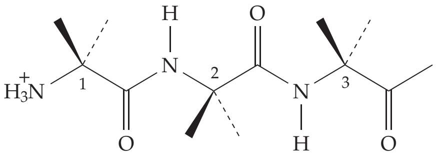

van der Waals surface
N
H
O
C
N
H
H
C
H
H
C
C
O
H
N
H
C
H H
C
O
H
CH₂
Phenylanine side chain
C
C
C
C
C
CH₂
van der Waals surface
H+
H–N–H
C H
H
H₂C
CH₂
H₂C
Lysine side chain

24. Draw three residues of a polypeptide chain constructed of L-amino acids using the top template of the two below. Now, using the second template, whose polypeptide chain begins at the right, use the same three residues, *numbering from the C-terminus* using D-amino acids. This is know as a *retro*-inverso polypeptide.

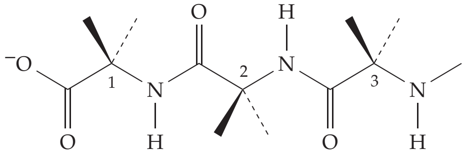
    Compare the outer surface of of the standard polypeptide and its *retro*-inverso analog. How would your answer be affected by the presence of threonine or isoleucine in the peptides?

    If the two polypeptides were folded into a hairpin loop of β structure how would the exterior surfaces compare? How would the hydrogen bonding compare?

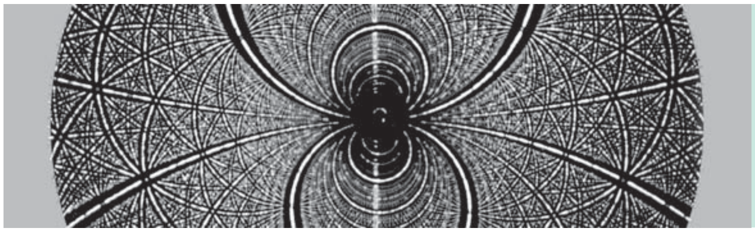
    What possible value can you imagine for *retro*-inverso polypeptides in design of drugs? See Brady and Dodson, *Nature* **368**, 692–694 (1994) or Guichard *et al., Proc. Natl. Acad. Sci. USA* **91**, 9765–9769 (1994)



Polypeptide



Retro-inverso polypeptide



A Laue X-ray diffraction pattern from a protein crystal. A stationary crystal is irradiated with very intense white, multiwavelength X rays from a synchrotron source. The diffraction pattern is rich in information. A single 0.1 ms X-ray pulse may provide a pattern with enough information to determine a three-dimensional structure. The pattern consists of thousands of diffraction spots arranged on intersecting rings. The coordinates of the diffraction spots together with their measured intensities provide the necessary information for structure determination. Courtesy of Louise Johnson.

Contents

- 95 **A. Understanding pH and Electrical Charges on Macromolecules**
  - 95 1. Strengths of Acids and Bases: the pKa's
  - 96 2. Titration Curves
  - 97 3. Buffers
- 98 **B. Isolating Compounds**
  - 98 1. Fractionation of Cells and Tissues
  - 100 2. Separations Based on Molecular Size, Shape, and Density
    - 100 *Dialysis, ultrafiltration, and perfusion chromatography*
    - 100 *Centrifugation*
  - 101 3. Separations Based on Solubility
  - 102 4. Separation by Partition
  - 103 5. Ion Exchange Chromatography
  - 104 6. Affinity Chromatography
  - 106 7. Electrophoresis and Isoelectric Focusing
- 108 **C. Determining the Relative Molecular Mass, M_r**
  - 108 1. Ultracentrifugation
    - 108 *Sedimentation velocity*
  - 112 2. Gel Filtration and Gel Electrophoresis
  - 112 3. Mass Spectrometry
- 115 **D. Determining Amino Acid Composition and Sequence**
  - 115 1. Cleavage of Disulfide Bridges
  - 116 2. Hydrolysis and Other Chain Cleavage Reactions
    - 116 *Selective enzymatic hydrolysis*
    - 117 *Nonenzymatic cleavages*
    - 118 *Separating the peptides*
  - 118 3. Determining Amino Acid Sequence
    - 118 *The Edman degradation*
    - 119 *Protein sequences from the genes*
    - 119 *Mass spectrometry in sequencing*
  - 119 4. Locating Disulfide Bridges
  - 120 5. Detecting Products
  - 122 6. Absorption of Light
- 123 **E. Quantitative Determinations and Modification Reactions of Side Chain Groups**
  - 123 1. Reactions of Amino Groups
  - 125 2. Reactions of SH Groups
  - 126 3. Reactions of Other Side Chains
  - 127 4. Affinity Labeling
- 127 **F. Synthesis of Peptides**
  - 127 1. Solid-Phase Peptide Synthesis
  - 128 2. Combinatorial Libraries
- 129 **G. Microscopy**
- 132 **H. X-ray and Neutron Diffraction**
- 137 **I. Nuclear Magnetic Resonance (NMR)**
  - 137 1. Basic Principles of NMR Spectroscopy
    - 138 *Band widths*
    - 138 *The chemical shift*
    - 139 *Scalar coupling (J coupling)*
  - 140 2. Nuclei Other Than Hydrogen
    - 140 *Deuterium (2H)*
    - 140 *Carbon - 13*
    - 140 *Nitrogen - 15*
    - 140 *Phosphorus - 31*
    - 140 *Fluorine - 19*
    - 141 *Some Other Nuclei*
  - 141 3. Fourier Transform Spectrometers and Two-Dimensional NMR
    - 141 *Free induction decay*
    - 141 *Relaxation times T1 and T2*
    - 142 *Two-dimensional and multidimensional NMR spectra*
  - 145 4. Three-Dimensional Structures and Dynamics of Proteins
    - 145 *Assignment of resonances*
    - 145 *NOEs and distance constraints*
  - 147 5. Other Information from NMR Spectra
    - 147 *NMR titrations*
    - 148 *Observing exchangable protons*
    - 148 *Exchange rates of amide protons*
    - 149 *Solid-state NMR and other topics*
- 149 **J. The Protein Data Bank, Three-Dimensional Structures, and Computation**

- 150 **References**
- 157 **Study Questions**

Boxes

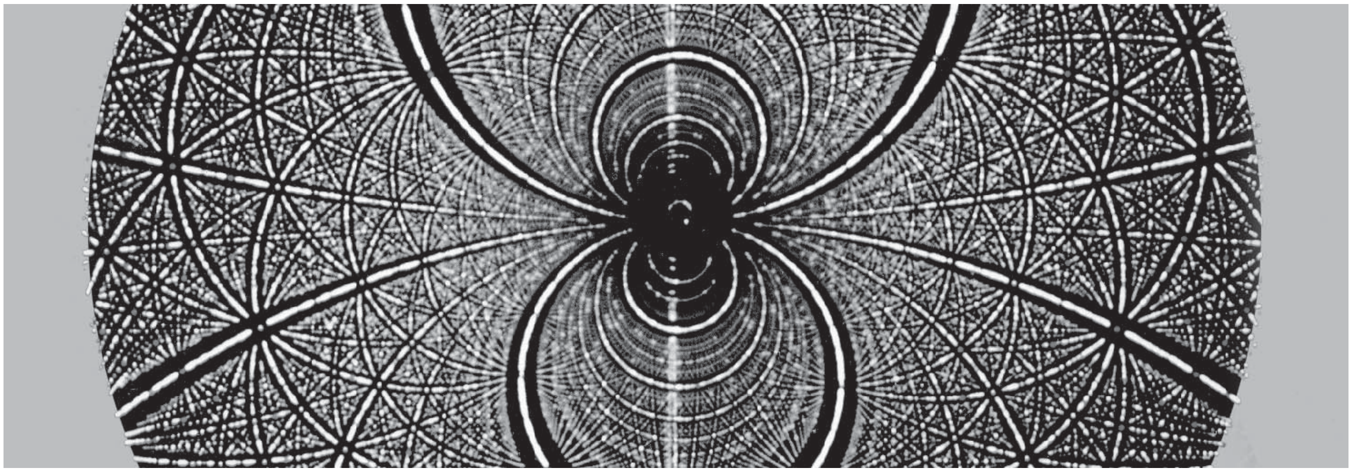
- 102 Box 3-A Quantitative Estimation of Protein Concentrations
- 107 Box 3-B Sorting and Analyzing Single Cells
- 110 Box 3-C Isotopes in Biochemical Investigations
- 116 Box 3-D The Proteome
- 121 Box 3-E Ninhydrin
- 122 Box 3-F Biosensors and Electronic Noses

Tables

- 99 Table 3-1 Practical pK_a Values for Some Useful Buffer Compounds at 25°C and Ionic Strength 0.1
- 117 Table 3-2 Specifities of Commonly Used Protein-Hydrolyzing Enzymes
- 141 Table 3-3 Approximate Chemical Shift Ranges in ^{1}H- and in ^{15}N- NMR Spectra
- 149 Table 3-4 Selected World Wide Web Servers Related to Protein Structures and Sequences

Determining Structures and Analyzing Cells

3



How have chemists deduced the thousands of structural formulas that we write for the substances found in nature? The answer is far too complex to give here in detail. However, the separation of compounds, the analysis of mixtures, and the unraveling of structures remain essential parts of biochemistry. A "minireview" of methods, with emphasis on proteins, follows. Additional procedures having to do primarily with carbohydrates, nucleic acids, or lipids are given in Chapters 4, 5, and 8, respectively.

A. Understanding pH and Electrical Charges on Macromolecules

Proteins, nucleic acids, and carbohydrates all contain acidic or basic functional groups. The strengths of the acidic groups vary over a broad range from that of the strongly acidic phosphate and sulfate esters to that of the very weakly acidic alcoholic – OH group. The net electrical charge, as well as the spatial distribution of the charged groups, affects the properties of a macromolecule greatly. Therefore, it will be worthwhile for us to consider some aspects of acid–base chemistry before discussing other topics.

1. Strengths of Acids and Bases: the pK_a's

The strength of an acid is usually described by the acid dissociation constant K_a

$$HA = A^- + H^+ \tag{3-1}$$

$$K_a = [A^-][H^+] / [HA] \tag{3-2}$$

For strong acids K_a is high and for weak acids it is low. Since the values of K_a vary by many orders of magnitude it is customary to use as a measure of the acid strength pK_a. This is the negative logarithm of K_a ($pK_a = -\log K_a$) . For very strong acids pK_a is less than zero, while very weak acids have pK_a values as high as 15 or more.

In the biochemical literature *the strength of a base is nearly always given by the* pK_a *of the conjugate acid.* Thus, A^- in Eq. 3-1 is a base and HA its conjugate acid. The base could equally well be uncharged A and its conjugate acid HA^+. In both cases Eq. 3-2 would hold. This defines the strength of both the acid HA and the base A^-. It follows that strong bases have weak conjugate acids with high pK_a values, while weak bases have strong conjugate acids with low pK_a values.

For a compound containing several acidic groups we define a series of consecutive dissociation constants K_{1a}, K_{2a}, K_{3a}, etc. For the sake of simplicity we will omit the a's and call these K_1, K_2, K_3, ⋯.

$$H_3A \xrightarrow{K_1} H_2A \xrightarrow{K_2} HA \xrightarrow{K_3} A \tag{3-3}$$

If there are n consecutive dissociation constants there will be $n + 1$ ionic species H_3A, H_2A, etc. Notice that H_3A could be a neutral molecule with H_2A, HA, and A bearing changes of –1, –2, and –3, respectively. Alternatively, H_3A might carry 1, 2, or 3 positive changes. In every case the mathematical expressions will be the same. For this reason the charges have been deliberately omitted from Eq. 3-3 and others that follow. Each constant in Eq. 3-3 is defined as in Eq. 3-2, i.e., $K_2 = [HA][H^+] / [H_2A]$, etc. Keep in mind that there are significant differences between the *apparent equilibrium constants* (concentration equilibrium constants) that we ordinarily use and thermodynamic equilibrium constants that are obtained by extrapolation to zero ionic strength.[1] A related complication is the uncertainity associated with the measurement of pH. This is often considered a measurement of hydrogen ion activity but this is not a correct statement. (The matter is considered briefly in Chapter 6). However, for all practical purposes, in the range of about pH 4–10 the pH can be equated with $-\log [H^+]$.

Often only one of the ionic forms of Eq. 3-3 will be important in a biochemical reaction. A particular ionic species may be the substrate for an enzyme. Likewise, an enzyme–substrate complex in only a certain state of protonation may react to given products. In these cases, and whenever pH affects an equilibrium, it is useful to relate the concentration $[A_i]$ of a given ionic form of a compound to the total of all ionic forms $[A]_t$ using Eqs. 3-4 and 3-5.

$$[A_i] = [A]_t / F_i \tag{3-4}$$

$$[A]_t = [A] + [HA] + [H_2A] + \cdots [H_nA] \tag{3-5}$$

For Eq. 3-4, $A_1 = H_nA$, $A_2 = H_{n-1}A$, etc. and F_1, F_2, etc. are the *Michaelis pH functions*,[2,3] which were proposed by L. Michaelis in 1914. For the case represented by Eq. 3-3 there are four ionic species and therefore four Michaelis pH functions which have the following form (Eq. 3-6). Here, K_1, K_2, etc. are the usual consecutive acid dissociation constants.

$$\begin{aligned}
F_1 &= 1 + K_1/[H^+] + K_1K_2/[H^+]^2 + K_1K_2K_3/[H^+]^3 \\
F_2 &= [H^+]/K_1 + 1 + K_2/[H^+] + K_2K_3/[H^+]^2 \\
F_3 &= [H^+]^2/K_1K_2 + [H^+]/K_2 + 1 + K_3/[H^+] \\
F_4 &= [H^+]^3/K_1K_2K_3 + [H^+]^2/K_2K_3 + [H^+]/K_3 + 1
\end{aligned} \tag{3-6}$$

If there are only three ionic forms the first three of these equations will apply if the final term is dropped from each. The student should be able to verify these equations and to write the appropriate pH functions for other cases. Since these relationships are met so often in biochemistry it is worthwhile to program a computer to evaluate the Michaelis pH functions and to apply them as needed. From Eq. 3-4 it can be seen that *the reciprocal of the Michaelis pH function for a given ionic form represents the fraction of the total compound in that form* and that the sum of these reciprocals for all the ionic forms is equal to one. Examples of the use of the Michaelis pH functions in this book are given in Eq. 6-50, which relates the Gibbs energy of hydrolysis of ATP to the pH, and in Eqs. 9-55 to 9-57, which describe the pH dependence of enzymatic action.

Notice that in Eq. 3-3 single arrows have been used rather than the pairs ($\rightleftarrows$) that are often employed to indicate reversible equilibria. This is done so that *the direction of the arrow indicates whether we are using a dissociation constant or an association constant*. The use of single arrows in this manner to indicate how the equilibrium constants are to be written is a good practice when dealing with complex equilibria.

2. Titration Curves

When a neutral amino acid is titrated with acid, the carboxylate groups become protonated and acid is taken up. Likewise titration with base removes protons from the protonated amino groups and base is taken up. If we plot the number of equivalents of acid or base that have reacted with the neutral amino acid versus pH, a titration curve such as that shown in Fig. 3-1 is generated. The curve for histidine contains three steps; the first corresponds to the titration of the carboxylate group with acid, the second to the titration of the protonated imidazole of the side chain, and the third to the titration of the protonated amino group with base. Each step is characterized by a midpoint that is equal to the pK_a value for the group being titrated. The ends of the curve at low and high pH, which are drawn with a dashed line, are obtained only after corrections have been applied to the data. If only the equivalents of acid or base *added* rather than the number *reacted* are plotted, we obtain the solid line shown in Fig. 3-1. We see that at the low pH end there is no distinct end point. As we add more acid to try to complete the titration, the correction that must be applied to the data becomes increasingly greater. The difference between the dashed and solid lines reflects the fact that at the low pH end much of the acid added is used to simply lower the pH. Therefore, we have a large free $[H^+]$. Similarly, at the high pH end we have a high free $[OH^-]$.

The exact shapes of the ends of the titration curve depend heavily on the total concentration. Likewise, the magnitude of the correction required to obtain a plot of equivalents of acid or base reacted varies with the concentration and is smaller the higher the concentration of the substance being titrated (see problems 2 and 3 at the end of this chapter). An important rule in doing acid-base titrations, especially when very small samples are available, is to use the highest possible concentration of sample in the smallest possible volume and to titrate with relatively concentrated acid or base. Because of the difficulty of adequately correcting titration curves at low pH it is hard to estimate the pK_a values of the carboxyl groups of amino acids accurately from titration. An additional experimental difficulty exists at the high pH end because of the tendency for basic

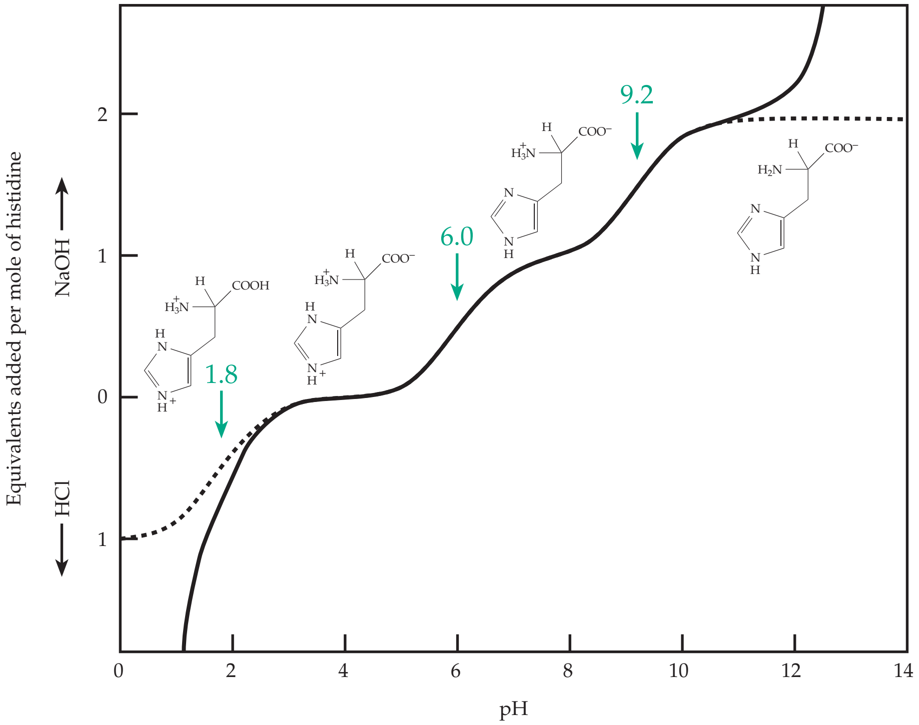


Figure 3-1 Titration curve for histidine. The solid line represents the uncorrected titration curve for 3 mM histidine monohydrochloride titrated with 0.2 M HCL to lower pH or with 0.2 M NaOH to higher pH assuming pK_a values of 1.82, 6.00, and 9.17. The dashed line represents the corrected curve showing the number of protons bound or lost per mole of histidine monohydrochloride.

solutions to absorb CO_2 from the air. Sometimes formaldehyde is added to shift the apparent pK_a of the amino groups to lower values and make the titration more satisfactory.

Despite the difficulties, real proteins can be titrated successfully to estimate the numbers of carboxyl, histidine, tyrosine, and other groups.[4] An example is shown in Fig. 3-2. The experimental data have been fitted with a theoretical curve based on the pK_a's of the carboxyl, histidine, tyrosine, and amino groups as determined by NMR measurements.[5] Account has been taken of the effect of the electrical field created by the many charged groups in distorting the curve from that obtained by summing the theoretical titration curves of the component groups. However, when there are multiple acid–base groups that are close together in a protein, a more complex situation involving tautomerism arises. This is discussed in Chapter 6.

Titration curves based on plots of light absorption versus pH or of NMR chemical shifts versus pH (see Fig. 3-29) are often useful. They have the important advantage that no special correction for free acid or base is needed at low or high pH.

3. Buffers

A mixture of a weak acid HA and of its **conjugate base** A constitutes a buffer which resists changes in pH. This can be seen most readily by taking logarithms of both sides of Eq. 3-2. By replacing log K with −pK_a and log [H^+] with −pH and rearranging we obtain Eq. 3-7 (the Henderson–Hasselbalch equation). It is sometimes useful to rewrite this as Eq. 3-8, where α is the fraction of the acid that is dissociated at a given pH.

$$\text{pH} - \text{p}K_a = \log([\text{A}]/[\text{HA}]) \qquad (3\text{-}7)$$

$$\text{pH} - \text{p}K_a = \log[\alpha/(1-\alpha)];$$

$$\alpha = [\text{A}]\ /\ [\text{HA}] + [\text{A}] = \frac{10^{(\text{pH}-\text{p}K_a)}}{1 + 10^{(\text{pH}-\text{p}K_a)}} \qquad (3\text{-}8)$$

Logarithms to the base 10 are used in both equations. These equations are useful in preparing buffers and in thinking about what fraction of a substance exists in a given ionic form at a particular value of pH. From Eq. 3-7 it is easy to show that when the pH is near the pK_a relatively large amounts of acid or base must be added to change the pH if the concentrations of the buffer pair A and HA are high.

Buffers are often added to maintain a constant pH in biochemical research[6] and naturally occurring buffer systems within body fluids and cells are very important (Box 6-A). Among the most important natural buffers are the proteins themselves, with the imidazole groups of histidine side chains providing much of the buffering capacity of cells around pH 7 (Figs. 3-1 and 3-2). Table 3-1 lists some useful biochemical buffers and their pK_a values. Here are a few

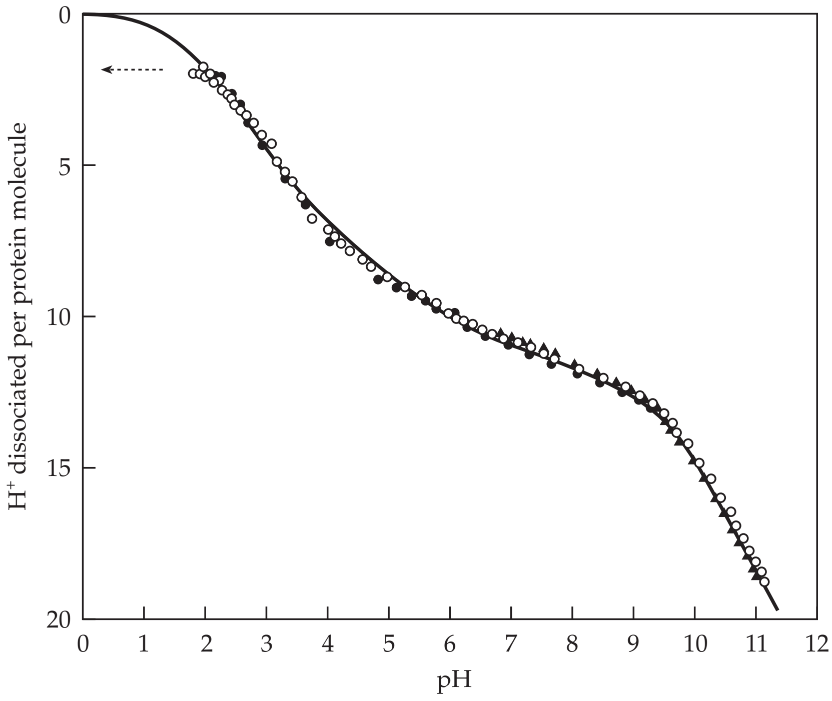


Figure 3-2 Acid–base titration curve for hen lysozyme at 0.1 ionic strength and 25°C. ○, initial titration from the pH attained after dialysis; ●, back titration after exposure to pH 1.8; ▲, back titration after exposure to pH 11.1. The solid curve was constructed on the basis of "intrinsic" pK_a values based on NMR data. From Kuramitsu and Hamaguchi[5]

practical hints about buffer preparation. Buffers containing monovalent ions tend to change pH with dilution less than do those with multivalent ions such as HPO_4^{2-} or $HP_2O_7^{3-}$. The pK_a values of carboxylic acids and of phosphoric acid, or of its organic derivatives, change very little with change in temperature. The pH of a buffer prepared with such components is nearly independent of temperature (Table 3-1). However, the pK_a of the $-NH_3^+$ group changes greatly with temperature. Buffer composition can be calculated readily from Eq. 3-8 and pK_a values from Table 3-1. It is convenient to keep in the laboratory standardized (to ~1% error) 1 M HCl and 1 M NaOH for use in buffer preparation. Compositions calculated from Eq. 3-8 will usually yield buffers of pH very close to those expected. Final adjustment with HCl may be needed if the pH is more than one unit away from a pK_a value. When two buffering materials are present, the composition should be calculated independently for each. The measurement of pH should always be done with great care because it is easy to make errors. Everything depends upon the reliability of the standard buffers used to calibrate the pH meter.[7] Often, especially during isolation of small compounds, it is desirable to work in the neutral pH region with volatile buffers, e.g., trimethylamine and CO_2 or ammonium bicarbonate, which can be removed by vacuum evaporation or lyophilization.[8]

B. Isolating Compounds

Before structural work can begin, pure substances must be separated from the complex mixtures in which they occur in cells and tissues.[4,13–24] Often, a substance must be isolated from a tissue in which it is present in a very low concentration. After it is isolated in pure form, if it is a large molecule, it must often be cut up into smaller pieces which are separated, purified, and identified. Accurate quantitative analysis is required to determine the ratios of these fragments. Considerable ingenuity may then have to be exercised in putting the pieces of the jigsaw puzzle back together to determine the structure of the "native" molecule. Many books, a few of which are cited here,[4,13–23,25–45] provide instructions. There are also journals and other periodicals dedicated to biochemical methods.[46–53]

1. Fractionation of Cells and Tissues

A fresh tissue or a paste of packed cells of a microorganism, usually collected by centrifugation, may be the starting material.[23,54–57] Tissue is often ground in a kitchen-type blender or, for gentler treatment, in a special **homogenizer**. The popular Potter–Elvehjem homogenizer is a small apparatus in which a glass or plastic pestle rotates inside a tight-fitting mortar tube (see standard laboratory equipment catalogs for pictures). Microbial cells are frequently broken with supersonic oscillation (**sonication**) or in special pressure cells. It is important to pay attention to the pH, buffer composition, and, if subcellular organelles are to be separated, the osmotic pressure. To preserve the integrity of organelles, 0.25 M sucrose is frequently used as the suspending medium, and $MgCl_2$ as well as a metal complexing agent such as ethylenediaminetetraacetate (EDTA) (Table 6-10) may be added. Soluble enzymes are often extracted without addition of sucrose, but reducing compounds such as glutathione (Box 11-B), mercaptoethanol, or dithiothreitol (Eq. 3-23) may be added. The crude **homogenate** may be strained and is usually centrifuged briefly to remove cell fragments and other "debris." Large-scale purification of

TABLE 3-1
Practical pK_a Values for Some Useful Buffer Compounds at 25°C and Ionic Strength 0.1[a]

Compound	pK_a	Grams per mole	d(pK_a) / dT	Charge on conjugate base
Citric acid (pK_1)		192		−1
Formic acid	3.7	0	−1	
Citric acid (pK_2)	4.45	192	−0.0016	−2
Acetic acid	4.64	60	0.0002	−1
Succinic acid (pK_2)	5.28	118	0	−2
Citric acid (pK_3)	5.80	192	0	−3
3,3-Dimethylglutaric acid	5.98	160	0.006	−2
Piperazine (pK_1) (5.68)	6.02	86		0
Cacodylic acid (dimethylarsinic acid)	6.1	138		−1
MES[b]	6.1	195	−0.011	
BIS-TRIS[b]	6.41	209	−.017	
Carbonic acid (pK_1)	6.4[c]			
Pyrophosphoric acid (pK_3)	6.76	178		-2
Phosphoric acid (pK_2)	6.84	98	−.0028	−2
PIPES[b]	6.90	353	−.0085	−2
Imidazole	7.07	68	−.020	0
BES[b]	7.06	213	−.016	−1
Diethylmalonic acid	7.2	136		
TES[b]	7.37	279	−.020	−1
HEPES[b]	7.46	238	−.014	−1
N-Ethylmorpholine	7.79	115	−.022	0
Triethanolamine	7.88	149	−.020	0
TRICINE[b]	8.02	178	−.021	−1
TRIS[b]	8.16	121	−.031	0
Glycylglycine	8.23	132	−.028	−1
BICINE[b]	8.26	163	−.018	−1
4-Phenolsulfonic acid	8.70	174	−.013	−2
Diethanolamine	9.00	105	−.024	0
Ammonia	9.2		−0.031	0
Boric acid	9.2		−0.008	mixed
Pyrophosphoric acid (pK_4)	9.41	178		-3
Ethanolamine	9.62	61	−.029	0
Glycine (pK_2)	9.8	75	−0.025	−1
Piperazine (pK_2)	9.82	86		0
Carbonic acid (pK_2)	10.0		−0.009	−2
Piperidine	11.1	85		0

[a] Based on compilation by Ellis and Morrison[9] with additional data from Good *et al.*[10,11] and Dawson *et al.* [12] The Good buffers have dipolar ionic constituents. Since no form is without electrically charged groups they are unlikely to enter and disrupt cells.

[b] Abbreviations used:

BES	*N*,*N*-Bis(2-hydroxyethyl)-2-aminoethanesulfonic acid
BICINE	*N*,*N*-Bis(2-hydroxyethyl)glycine
BIS-TRIS	Bis(2-hydroxyethyl)iminotris(hydroxymethyl)methane
HEPES	*N*-2-Hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid
MES	2-(*N*-Morpholino)ethanesulfonic acid
PIPES	Piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)
TES	*N*-Tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid
TRICINE	*N*-Tris(hydroxymethyl)methylglycine
TRIS	Tris (hydroxymethyl) aminomethane

[c] For CO_2 (solid) + $H_2O \rightarrow H^+ + HCO_3^-$, apparent p$K_a$.

proteins is often initiated with such a crude homogenate.

Cell organelles are also often separated by centrifugation. In one procedure a homogenate in 0.25 M sucrose (**isotonic** with most cells) is centrifuged for 10 min at a field of 600–1000 times the force of gravity (600–1000 *g*) to sediment nuclei and whole cells. The supernatant fluid is then centrifuged another 10 min at ~10,000 *g* to sediment mitochondria and lysosomes. Finally, centrifugation at ~100,000 *g* for about an hour yields a pellet of microsomes (p. 14),[58] which contains both membrane fragments and ribosomes. Each of the separated components can be resuspended and recentrifuged to obtain cleaner preparations of the organelles. The sedimented particles can often be solubilized by chemical treatment, for example, by the addition of either ionic or nonionic detergents. Membrane proteins can be isolated following solubilization in this way (Chapter 8). The soluble supernatant fluid remaining after the highest speed centrifugation provides the starting material for isolation of soluble enzymes and many small molecules.

2. Separations Based on Molecular Size, Shape, and Density

The simplest way to separate very large dissolved molecules is to let the small ones pass through a suitable sieve which may be a membrane with holes or a bed of gel particles. If the size of the particles approaches that of the holes in the sieve, the rate of passage will depend upon shape as well as size.

Dialysis, ultrafiltration, and perfusion chromatography. In dialysis[59] and ultrafiltration,[60] a thin membrane, e.g., made of cellulose acetate (cellophane) and containing holes 1–10 nm in diameter (typically 5 nm), is used as a semipermeable barrier. Small molecules pass through but large ones are retained. Dialysis depends upon diffusion and can be hastened by adequate stirring. Ultrafiltration requires a pressure difference across the membrane. The more sophisticated procedures of gel filtration and perfusion chromatography were introduced in 1959.[61–65] A column is packed with material such as the crosslinked dextran Sephadex, polyacrylamide gels (such as the Bio-Gel P Series), or agarose gels (e.g., Bio-Gels A and Sepharoses). These come in the form of soft beads, the interior network of which is a three-dimensional network of polymer strands (Fig. 4-10). Recently porous beads of hard crosslinked polystyrene, glass, or various other silicate materials have been employed.[66–68] The interstices between strands, whose size depends upon the degree of crosslinking introduced chemically into the gel, are small enough to exclude large molecules but to admit smaller ones. If a mixture of materials of different molecular size is passed through such a column the smaller molecules are retarded because of diffusion into the gel, while the larger molecules pass through unretarded (Fig. 3-3). Sephadex G-25 excludes all but salts and compounds no larger than a simple sugar ring. Sephadex G-200, which is much less crosslinked, permits separation of macromolecules in the range of 5–200 kDa. As is explained in Section B, gel filtration also provides an important way of estimating M_r for proteins and other macromolecules.

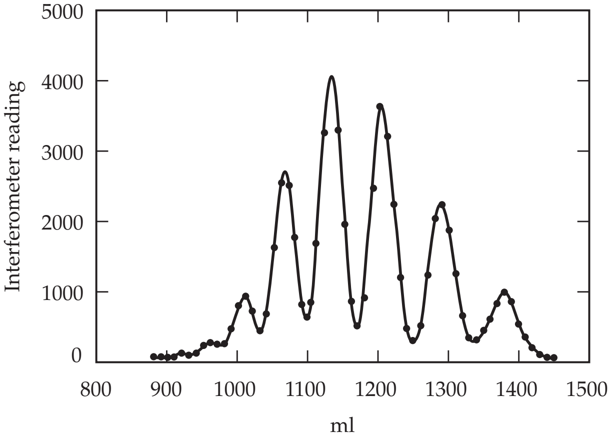


Figure 3-3 Separation of oligosaccharides by gel filtration. The sugars dissolved in distilled water were passed through a column of Sephadex G-25. The peaks contain (right to left) glucose, cellobiose, cellotriose, etc. From Flodin and Aspberg.[64]

Centrifugation. Centrifuges of many sizes and speeds are used in the laboratory to remove debris as well as to collect precipitated proteins and other materials at various steps in a purification scheme. The most remarkable are the **ultracentrifuges** which produce forces greater than 4 x 10^5 times that of gravity. They can be used both for separation of molecules and for determination of M_r (see Section C).

When macromolecules in a solution are subjected to an ultracentrifugal field they are accelerated rapidly to a constant velocity of sedimentation. This is expressed as a **sedimentation constant *s***, which is the rate (cm/s) per unit of centrifugal force. The unit of *s* is the second but it is customary to give it in Svedberg units, S (1S = 10^{-13}*s*). Sizes of particles are often cited by their S values. The sedimentation constant is affected by the sizes, shapes, and densities of the particles as is discussed further in Section C. If carried out at constant velocity an equilibrium will eventually be attained in which sedimentation is just balanced by diffusion and a smooth concentration gradient forms from the top to the bottom of the centrifuge cell or tube. Concentration gradients can also be formed by centrifuging a concentrated solution of small molecules.[13,69,70] Such a concentration gradient is also a **density gradient** which can be made very steep. For example, a gradient with over a 10% increase in density from top to bottom can be created using 6 M **cesium chloride** (CsCl) and is widely used in DNA separations. If DNA is added prior to centrifugation it will come to rest in a narrow band or bands determined by the **buoyant densities** of the species of DNA present (see Chapter 5). After centrifugation, which is usually done in a plastic tube, a hypodermic needle is inserted through the bottom of the tube and the contents are pumped or allowed to flow by gravity into a fraction collector.

Another type of gradient centrifugation (**zone centrifugation**) utilizes a preformed gradient to stabilize bands of cell fragments, organelles, or macromolecules as they sediment.[71–74] For example, RNA can be separated into several fractions of differing sedimentation constants in a centrifuge tube that contains **sucrose** ranging in concentration from 25% at the bottom to 5% at the top. This is prepared by a special mixing device or "gradient maker" prior to centrifugation. The solution of RNA is carefully layered on the top, the tube is centrifuged at a high speed for several hours, and the different RNA fractions separate into slowly sedimenting sharp bands. A 20–60% gradient of sucrose or glycerol may be used in a similar way to separate organelles.[58]

3. Separations Based on Solubility

Some fibrous proteins are almost insoluble in water and everything else can be dissolved away. More often soluble proteins are precipitated from aqueous solutions by adjustment of the pH or by addition of large amounts of salts or of organic solvents. The solubility of any molecule is determined both by the forces that hold the molecules together in the solid state and by interactions with solvent molecules and with salts or other solutes that may be present.[58a]

Proteins usually have many positively and negatively charged groups on their surfaces. If either a positive or a negative charge predominates at a given pH the protein particles will tend to repell each other and to remain in solution. However, near the isoelectric point (pI), the pH at which the net charge is zero (see Section 7), the solubility will usually be at a minimum. The pH of a tissue extract may be adjusted carefully to the pI of a desired protein. Any protein that precipitates can be collected by centrifugation and redissolved to give a solution enriched in the protein sought. Some proteins, such as those classified (by an old system) as **globulins**, are insoluble in water but are readily **"salted in"** by addition of low concentrations (e.g., up to 0.1 M) of salts. Low concentrations of salts increase the solubility of most proteins because the salt ions interact with the charged groups on the protein surfaces and interfere with strong electrostatic forces that are often involved in binding protein molecules together in the solid state. Some salts, including $CaCl_2$ and NaSCN, which bind to proteins, are especially effective in salting in.[75,76] Addition of *high* concentrations of salt causes precipitation of most proteins from aqueous solutions. The most effective and most widely used materials for this **"salting out"** of proteins are $(NH_4)_2SO_4$ and Na_2SO_4. Because the salt ions interact so strongly with water, the protein molecules interact less with water and more with each other. A similar intramolecular effect may cause the stabilization of proteins of halophilic bacteria by 1–4 M KCl.[77]

Different proteins precipitate at different concentrations of an added salt. Hence, a fraction of proteins precipitating between two different concentrations of salt can be selected for further purification (Fig. 3-4). Protein concentrations can be estimated as described in Box 3-A.

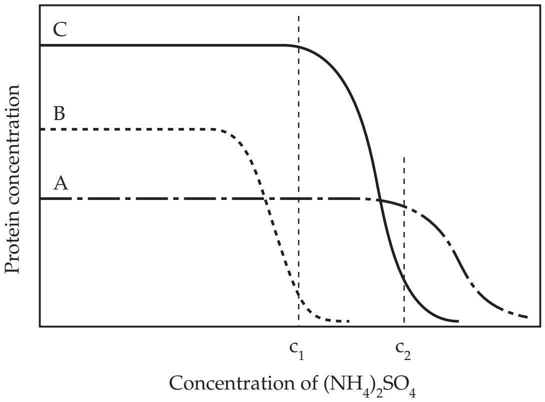


Figure 3-4 Hypothetical behavior of a solution containing three proteins, A, B, and C, upon ammonium sulfate fractionation. The concentration of protein remaining in the solution is plotted against ammonium sulfate concentration (usually expressed as % saturation). Addition of ammonium sulfate to concentration c_1 will precipitate largely protein B, which can be removed by centrifugation. Addition of additional salt to c_2 will precipitate largely protein C, while A remains in solution.

Precipitation methods are popular first steps in purification of proteins because they can be carried out on a large "batch" scale. The amounts of ammonium sulfate used are often expressed as percentage saturation, i.e., as a percentage of the amount required to saturate the solution (4.1 M at 25°). Convenient tables are available[12,78] that permit one to weigh out the correct amount of solid ammonium sulfate to give a desired percentage saturation or to go from one percentage saturation to a higher one.

Proteins are often stabilized by low concentrations of simple alcohols or ketones[76] and by higher concentrations of polyhydroxy alcohols, such as glycerol[77] and sucrose,[78] and also by certain inert, synthetic polymers such as **polyethyleneglycol** (PEG).[79] The latter is a widely used precipitant. The polyhydroxyalcohols and PEG are all hydrated but tend not to interact strongly with the protein molecules. On the other hand, simple alcohols may denature proteins by their interaction with nonpolar regions.[77]

4. Separation by Partition

Many of the most important separation methods are based on repeated equilibration of a material between two separate phases, at least one of which is usually liquid. Small molecules may be separated by **countercurrent distribution** in which a material is repeatedly partitioned between two immiscible liquid phases, one more polar than the other. New portions of both liquids are moved by a machine in a "countercurrent manner" between the equilibration steps or are moved continuously through coiled tubes.[80–82,82a]

A similar result is accomplished by using as one phase a solid powder or fine "beads" packed in a vertical column or spread in a thin layer on a plate of glass. The methods are usually referred to as **chromatography**, a term proposed by Tswett to describe separation of materials by color. In 1903 Tswett passed solutions of plant leaf pigments (chlorophylls and carotenes) in nonpolar solvents such as hexane through columns of alumina and of various other adsorbents and observed separation of colored bands which moved down the column as more solvent was passed through. Individual

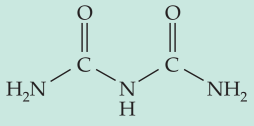
BOX 3-A QUANTITATIVE ESTIMATION OF PROTEIN CONCENTRATIONS

Biochemists often need to estimate the content of protein in a sample. For example, in devising a purification procedure for an enzyme it is customary to estimate the number of units of enzyme activity (as defined in Chapter 9) per milligram of protein (U/ mg). As progress is made in the purification this ratio increases. It becomes constant with respect to additional purification attempts, when a homogeneous enzyme is obtained.

One of the most widely used and most sensitive protein assays (for 0.1–1 mg/ml of protein) is the colorimetric procedure of **Lowry**.[a–c] It makes use of a phosphomolybdic–phosphotungstic acid reagent (the Folin–Ciocalteu reagent) which is reduced by proteins in the presence of alkaline Cu^{2+} to characteristic "blue oxides" whose color can be monitored at 750 nm. Much of the color comes from the reducing action of tyrosine and tryptophan. The color yield varies greatly from protein to protein and users may have trouble with reproducibility. A related method utilizes **bicinchoninic acid** which forms a purple color (measured at 362 nm) with the Cu^{+1} that is formed by reduction of alkaline Cu^{2+} by the protein.[d–f] This reagent is easier to use than that of the Lowry procedure and gives stable and reproducible readings.

A third widely used procedure, introduced by **Bradford**[g] and modified by others, measures the binding of the dye Coomassie brilliant blue whose peak absorption shifts from 465 nm to 595 nm upon binding. The change occurs within two minutes and is stable. However, the color yield varies from one protein to another.

Less sensitive but very simple and precise is measurement of the light absorption around 280 nm. This is discussed in the main text in Section D,6. For a typical protein an absorbance of 1.0 at 280 nm corresponds to a protein concentration of 1 mg / ml.[h] The very old **biuret method** is also useful for samples containing 1–10 mg / ml of protein. The violet color that arises upon addition of copper sulfate to an alkaline solution of a peptide or protein is recorded at 540–560 nm.[h] The color is especially intense for longer polypeptides. The name of the method arises from the fact that biuret gives a similar color[i] (see also Eq. 6-85).



Biuret

---

[a] Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275

[b] Peterson, G. L. (1979) *Anal. Biochem.* **100**, 201–220

[c] Larson, E., Howlett, B., and Jagendorf, A. (1986) *Anal. Biochem.* **155**, 243–248

[d] Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Kenk, D. C. (1985) *Anal. Biochem.* **150**, 76–85

[e] Davis, L. C., and Rodke, G. A. (1987) *Anal. Biochem.* **161**, 152–156

[f] Hill, H. D., and Straka, J. G. (1988) *Anal. Biochem.* **170**, 203–208

[g] Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254

[h] Fruton, J. S., and Simmonds, S. (1958) *General Biochemistry*, 2nd ed., Wiley, New York (p. 130)

[i] Layne, E. (1957) *Methods Enzymol.* **III**, 450–451

pure pigments could be **eluted** from the column by continued passage of solvent. This important method is called **adsorption chromatography**. It is assumed that the pigments are absorbed on the surface rather than being disssolved in the solid material. A related method is **hydrophobic interaction chromatography** of proteins.[83–86] The packing material is similar to that described in Section 6 (Affinity Chromatography) but bears long-chain alkyl groups that can interact with the hydrophobic patches on surfaces of proteins. A very different adsorbent that is very useful in separation of proteins is carefully prepared microcrystalline **hydroxylapatite.**[87,88] It presumably functions in part by ion exchange.

Column packing materials such as **silica gel** contain a large amount of water, and separation involves partition between an immobilized aqueous phase in the gel and a mobile, often organic, solvent flowing through the column. Usually materials elute sooner when they are more soluble in the mobile phase than in the aqueous phase. These methods are closely related to perfusion chromatography, which is described in Section 2.

Aromatic amino acids, lipids, and many other materials can be separated on **reversed-phase** columns in which nonpolar groups, usually long-chain alkyl groups, are covalently attached to silica gel, alumina, or other inert materials.[66,80] The mobile phase is a more polar solvent, often aqueous, and gradually made less polar by addition of an organic solvent. In reversed-phase chromatography more polar compounds migrate faster through the system than do nonpolar materials, which experience hydrophobic interaction with the solid matrix.

Many traditional chromatographic methods including reversed-phase chromatography have been adapted for use in automatic systems which employ columns of very finely divided solid materials such as silica, alumina, or ion exchange materials coated onto fine glass beads.[89–91] These **high-performance liquid chromatographic** (**HPLC**) systems often utilize pressures as high as 300 atmospheres. Separations are often sharper and faster than with other chromatographic methods.[92–97] Reversed-phase columns in which the solid matrix may carry long (e.g. C_{18}) hydrocarbon chains have been especially popular for separation both of peptides and of proteins. Proteins may also be separated by gel filtration, ion exchange, or other procedures with HPLC equipment.

A sheet of high-quality filter paper containing adsorbed water serves as the stationary phase in **paper chromatography**. However, **thin-layer chromatography**, which employs a layer of silica gel or other material spread on a glass or plastic plate, has often supplanted paper chromatography because of its rapidity and sharp separations (Fig. 3-5).[16,96a,98–100] An approach that requires no stationary phase at all is **field flow fractionation**.[101] Here a suitable external field (e.g., electrical, magnetic, and centrifugal) or a thermal gradient is imposed on the particles flowing through a narrow channel.

For volatile materials **vapor phase chromatography** (gas chromatography) permits equilibration between the gas phase and immobilized liquids at relatively high temperatures. The formation of volatile derivatives, e.g., methyl esters or trimethylsilyl derivatives of sugars, extends the usefulness of the method.[103,104] A method which makes use of neither a gas nor a liquid as the mobile phase is **supercritical fluid chromatography**.[105] A gas above but close to its critical pressure and temperature serves as the solvent. The technique has advantages of high resolution, low temperatures, and ease of recovery of products. Carbon dioxide, N_2O, and xenon are suitable solvents.

5. Ion Exchange Chromatography

Separation of molecules that contain electrically charged groups is often accomplished best by ion exchange chromatography.[105a] The technique depends upon interactions between the charged groups of the molecules being separated and fixed ionic groups on

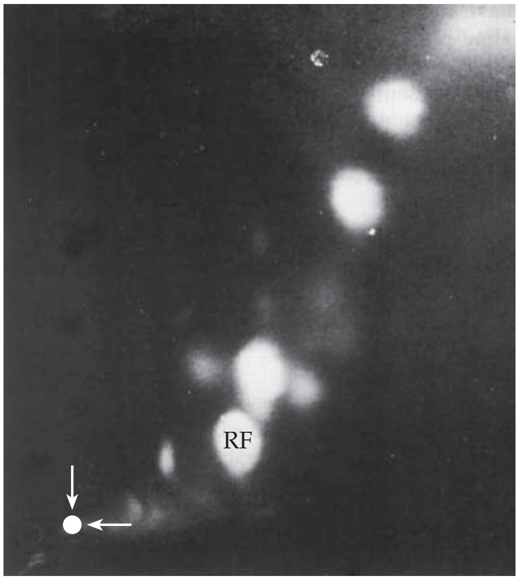


Figure 3-5 Photograph of a two-dimensional thin layer (silica gel) chromatogram of a mixture of flavins formed by irradiation of ~10 μg of the vitamin riboflavin. The photograph was made by the fluorescence of the compounds under ultraviolet light. Some riboflavin (RF) remains. The arrows indicate the location of the sample spot before chromatography. Chromatography solvents: a mixture of acetic acid, 2-butanone, methanol, and benzene in one direction and *n*-butanol, acetic acid, and water in the other. See Treadwell *et al*.[102]

an immobile matrix. Separation depends upon small differences in pK_a values and net charges and upon varying interactions of nonpolar parts of the molecules being separated with the matrix. Since changes in pH can affect both the charges on the molecules being separated and those of the ion exchange material, the affinities of the molecules being separated are strongly dependent on pH. For example, proteins and most amino acids are held tightly by cation exchangers at low pH but not at all at high pH.

Aqueous solutions are usually employed and the columns are packed with beads of **ion exchange resins**, porous materials containing bound ionic groups such as $-SO_3^-$, $-COO^-$, $-NH_3^+$, or quaternary nitrogen atoms. Synthetic resins based on a cross-linked polystyrene are usually employed for separation of small molecules. For larger molecules chemical derivatives of cellulose or of crosslinked dextrans (Sephadex), agarose, or polyacrylamide are more appropriate. Positively charged ions, such as amino acids in a low pH solution, are placed on a cation exchange resin such as Dowex 50, which contains dissociated sulfonic acid groups as well as counter ions such as Na^+, K^+, or H^+. The adsorbed amino acids are usually eluted with buffers of increasing pH containing sodium or lithium ions. The procedure, which was developed by Moore and Stein,[106–109] is widely used for automatic quantitative analysis of amino acid mixtures obtained by hydrolysis of a protein or peptide (Fig. 3-6).[110–113]

Ion exchange chromatography of proteins and peptides is often done with such ion exchange materials as carboxymethyl-Sephadex and phosphocellulose, which carry negatively charged side chains or diethylaminoethylcellulose (DEAE-cellulose), which carries positively charged amino groups.[114] These materials do not denature proteins or entrap them and have a large enough surface area to provide a reasonable absorptive capacity. The mobile phase is usually buffered. For anion exchangers the pH should be above ~4.4 to keep most carboxylate side chains on the proteins ionized. The pH may be increased to pH 7–10, where most histidine imidazolium ions have dissociated, increasing the mobility of many proteins. For cation exchange the pH is usually buffered below pH 6 or 7 (Fig. 3-6).[115,116]

6. Affinity Chromatography

In this technique the chromatographic absorbent is designed to make use of specific biochemical inter-

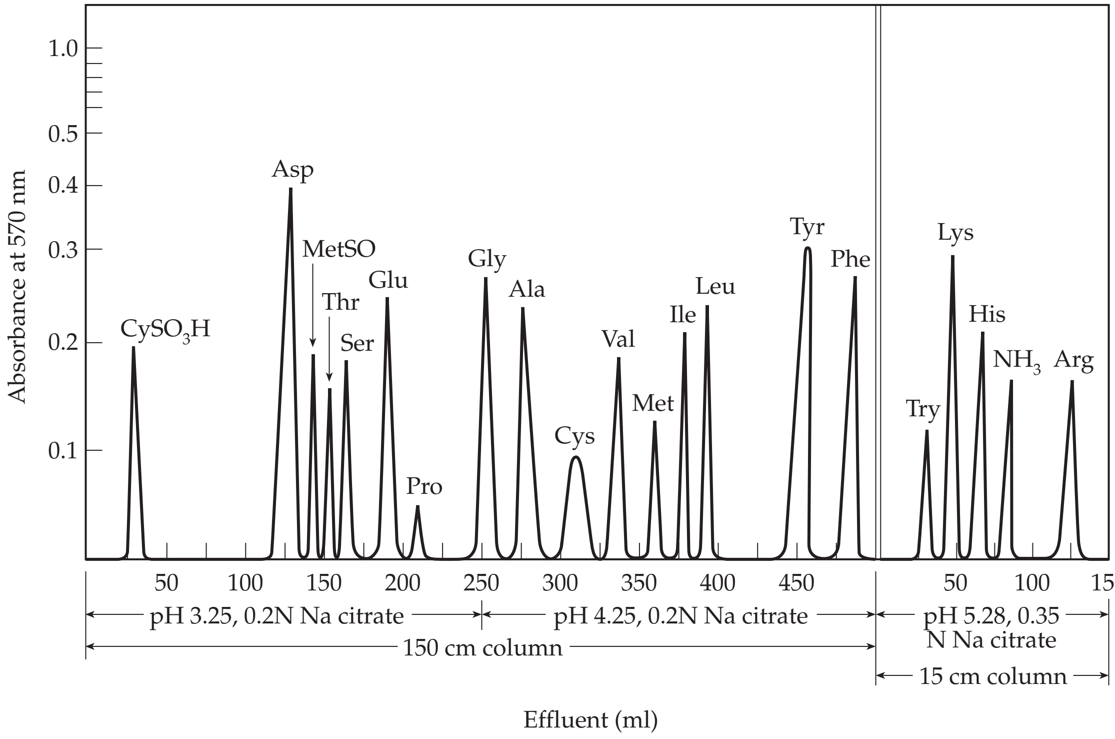
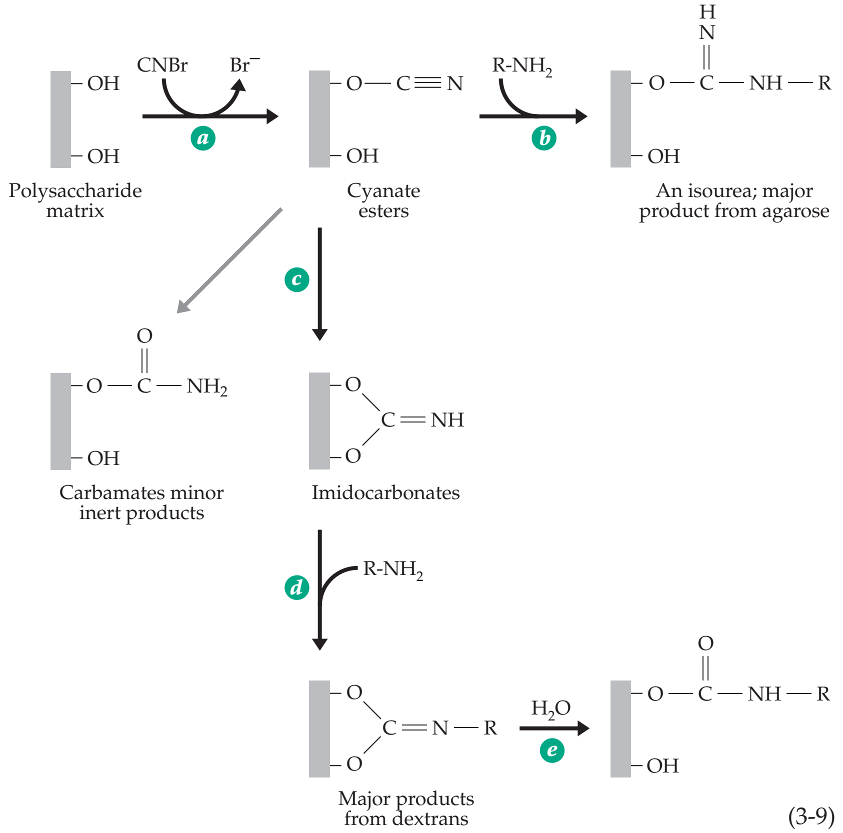


Figure 3-6 Separation of amino acids by cation-exchange chromatography on a sulfonated polystyrene resin in the Na^+ form by the method of Moore and Stein.[110] The amino acids were detected by reaction with ninhydrin (Box 3-C); areas under the peaks are proportional to the amounts. Two buffers of successively higher pH are used to elute the amino acids from one column, while a still higher pH buffer is used to separate basic amino acids on a shorter column. From Robyt and White.[13]

actions to "hook" selectively a particular macromolecule or group of macromolecules.[117–122] Affinity chromatography is used in many ways, including the purification of enzymes, antibodies, and other proteins that bind tightly to specific small molecules.

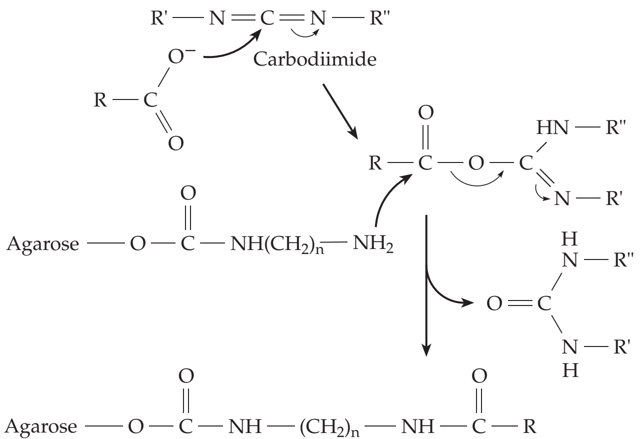
Because of their open gel structure (Fig. 4-10) agarose derivatives in bead form provide a good solid support matrix. The hydroxyl groups of the agarose are often linked to amino compounds. In one widely used procedure[119] the agarose is treated with cyanogen bromide ($Br-C\equiv N$) in base to "activate" the carbohydrate (Eq. 3-9, step *a*). Then the amino compound is added (step *b*). The isourea product shown is the major one with agarose gels but dextran-based matrices tend to react by steps *c*–*e*.

Absorbents containing a large variety of R groups of shapes specifically designed to bind to the desired proteins can be made in this way. If the coupling is done with a diamine [$R = (CH_2)_n—NH_2$], the resulting ω-aminoalkyl agarose can be coupled with other compounds by reaction with **carbodiimide** (Eq. 3-10). For reaction in nonaqueous medium dicyclohexylcarbodiimide (R″ = cyclohexyl) is often used, but for linking groups to agarose a water-soluble reagent such as 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide is recommended.[120] Carbodiimides are widely used for forming amide or phosphodiester linkages in the laboratory. The formation of an amide with a side chain of ω-aminoalkyl agarose can be pictured as in Eq. 3-10.



CNBr
Br⁻
a
Polysaccharide matrix
OH
OH
O—C≡N
OH
Cyanate esters
R-NH₂
b
H
N
O—C—NH—R
OH
An isourea; major product from agarose
c
O
O—C—NH₂
OH
Carbamates minor inert products
O
C=NH
O
Imidocarbonates
d
R-NH₂
O
C=N—R
O
Major products from dextrans
H₂O
e
O
O—C—NH—R
OH

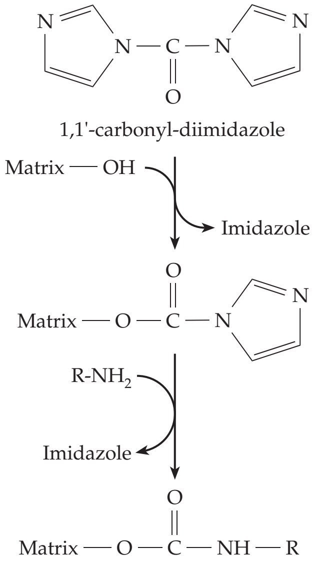
(3-9)



R'—N=C=N—R"
Carbodiimide
O⁻
R—C
O
O
R—C—O—C
HN—R"
N—R'
O
Agarose—O—C—NH(CH₂)ₙ—NH₂
H
N—R"
O=C
N—R'
H
O
Agarose—O—C—NH—(CH₂)ₙ—NH—C—R

(3-10)

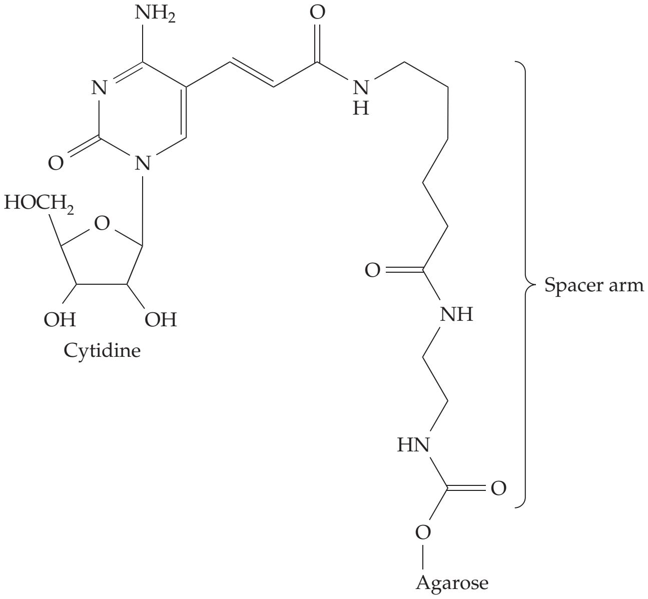
Many other means of preparation of adsorbents for affinity chromatography are also available.[121,122] For example, 1,1'-carbonyl-diimidazole can be used to couple a diamine to the matrix (Fig. 3-11). This reagent has the advantage that it does not depend upon the relatively unstable isourea linkages formed by Eq. 3-9 to hold the specific affinity ligands.[121]



N
N—C—N
N
O
1,1'-carbonyl-diimidazole
Matrix—OH
Imidazole
O
Matrix—O—C—N
N
R-NH₂
Imidazole
O
Matrix—O—C—NH—R

(3-11)

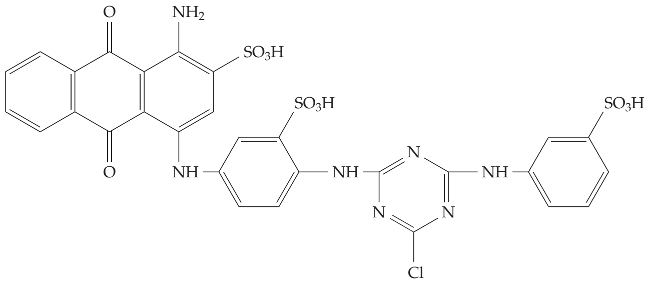
An example of a successful application of affinity chromatography is the isolation of the enzyme cytidine deaminase from cells of *E. coli*. Cytidine was linked covalently via long **spacer** arms to the agarose beads as in the following diagram:



A cell extract was subjected to ammonium sulfate fractionation and the dialyzed protein was then poured through the affinity column which held the cytidine deaminase molecules because of their affinity for the cytidine structures that were bound to the agarose. The protein was eluted with a borate buffer; the borate formed complexes with the adjacent hydroxyl groups of the cytidine and thereby released the protein. After passage of the protein through an additional column of DEAE-Sephadex the deaminase had been purified 1700-fold compared to the crude extract.[123]

Another technique is to engineer genes to place a polyhistidine "tag" at the C terminus of a protein chain. Commercial cloning vehicles and kits are available for this purpose.[124] The protein produced when the engineered gene is expressed can be captured by the affinity of the polyhistidine tag for Cu^{2+}, Ni^{2+}, Co^{2+}, or Zn^{2+} held in chelated form on an affinity column.[124–127]

A surprising discovery was that certain dyes, for example Cibachron blue, when covalently coupled to a suitable matrix, often bind quite specifically to proteins that have a nonpolar binding pocket near a positive charge.[128] This includes many enzymes that act on nucleotides.



Cibacron blue F3GA

7. Electrophoresis and Isoelectric Focusing

The methods considered in this section make use of movement of molecules in an electrical field. Separation depends directly upon differences in the net charge carried by molecules at a fixed pH. The net charge for compounds containing various combinations of acidic and basic groups can be estimated by considering the pK_a of each group and the extent to which that group is dissociated at the selected pH using Eqs. 3-3 to 3-5. At some pH, the **isoelectric point** (pI), a molecule will carry no net charge and will be immobile in an electric field. At any other pH it will move toward the anode (+) or cathode (–). The pH at which the protein carries no net charge in the complete absence of added electrolytes is called the **isoionic point**.[129]

Electrophoresis, the process of separating molecules, and even intact cells[130] (Box 3-B), by migration in an electrical field, is conducted in many ways.[28,131–140] In **zone electrophoresis**, a tiny sample of protein solution, e.g., of blood serum, is placed in a thin line on a piece of paper or cellulose acetate. The sheet is moistened with a buffer and electrical current is passed through it. An applied voltage of a few hundred volts across a 20-cm strip suffices to separate serum proteins in an hour. To hasten the process and to prevent diffusion of low-molecular-weight materials, a higher voltage may be used. Two to three thousand volts may be applied to a sample cooled by water-chilled plates. Large-scale electrophoretic separations may be conducted in beds of starch or of other gels.

One of the most popular and sensitive methods for separation of proteins is electrophoresis in a column filled with **polyacrylamide** or **agarose gel** or on a thin layer of gel on a plate. The method depends upon both electrical charge and molecular size and has been referred to as **electrophoretic molecular sieving**.[28,133,135–137,141–143] Polyacrylamide gel electrophoresis is often carried out in the presence of ~1% of the denaturing detergent sodium docecyl sulfate which coats the polypeptide chain rather evenly. This method, which is often referred to as **SDS–PAGE,** has the advantage of breaking up complex proteins composed of more than one subunit and sorting the resultant monomeric polypeptide chains according to molecular mass (see Box 3-C). A disulfide-reducing reagent (see Eq. 3-23) such as ~1% 2-mercaptoethanol is usually present but may be omitted to permit detection of crosslinked peptides.

Capillary electrophoresis is increasingly popular[144 – 149a] and can be used to separate attomole amounts.[150] It can be used not only for separation of proteins but also for rapid estimation of the net charge on a protein.[151] The separation is conducted in tubes with internal diameters as small as 10–15 µm and as short as 1 cm. Multiple channels cut into a glass chip

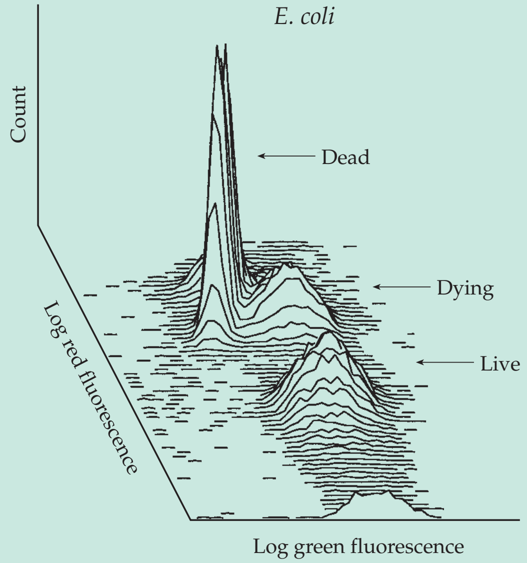
BOX 3-B SORTING AND ANALYZING SINGLE CELLS

It is often important to examine and analyze individual cells.[a] For example, large numbers of single blood cells can be tested for the presence of specific antigenic determinants that arise by mutation. This permits assessment of the frequency of these mutations.[b] The complex chemical processing of neuropeptides can be studied on the contents of a single neuron (Chapter 30) using mass spectrometry.[c]

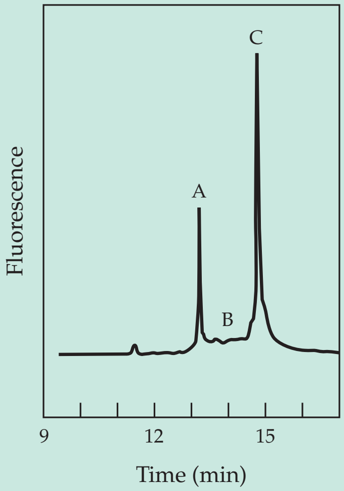
Several methods for separating cells have been devised. These include electrophoresis[d] or use of magnetic microspheres.[b] Micromanipulation can sometimes be used to select single cells for analysis. The most impressive technique is **flow cytometry**,[e,f] which is used daily on human blood samples in clinical laboratories. A suspension of cells is passed at a high rate of flow through a narrow capillary of ~0.2 mM diameter. The sample stream, which is surrounded by a larger "sheath" stream, has a smaller diameter of ~20 µm. One or more laser beams are used to record information about each cell over a period of a few microseconds as the cells pass by at a rate of as much as 10^5 cells / s.

Flow cytometers developed from simpler cell counters, but now they are used to record cell size (from light scattering), optical absorbance, fluorescence, and phosphorescence. The optical properties are often enhanced by staining. The use of two dyes that fluoresce at different wavelengths permits the construction of two-dimensional plots as in the accompanying figure.

Capillary electrophoresis is one of the techniques able to separate constituents of single cells and is illustrated in the second figure.



Flow cytometric histogram of fluorescently labeled live and dead *E. coli* bacteria. The dye kit (BacLightTM) that was used stains membrane-compromised dead bacteria with a red fluorescing dye and live bacteria with a green fluorescing dye. Cells were analyzed on an EPICS XL cytometer (Coulter Corporation). By integrating the area under each population it was possible to discern the percentage of dead (54), dying (16), and live (30) bacteria within a mixed population. The use of fluorescent dyes such as these has proven useful for studying various mechanisms employed by the food industry for killing microorganisms in food products and for studying a variety of bacteria derived from seawater, soil, plant materials, and laboratory-grown cultures. Courtesy of Kristi Harkins.



Electropherogram of major proteins from a single erythrocyte. Peaks A, B, and C are carbonic anhydrase (~7 amol), methemoglobin (~ 5 amol), and hemoglobin A_0 (~ 450 amol), as identified from migration times relative to standards.[a] Courtesy of Edward S. Yeung.

---

[a] Yeung, E. S. (1994) *Acc. Chem. Res.* **27**, 409–414

[b] Jovin, T. M., and Arndt-Jovin, D. J. (1980) *Trends Biochem. Sci.* **5**, 214–219

[c] Li, K. W., Hoek, R. M., Smith, F., Jiménez, C. R., van der Schors, R. C., van Veelen, P. A., Chen, S., van der Greef, J., Parish, D. C., Benjamin, P. R., and Geraerts, W. P. M. (1994) *J. Biol. Chem.* **269**, 30288–30292

[d] Bauer, J. (1994) *Cell Electrophoresis*, CRC Press, Boca Raton, Florida

[e] Shapiro, H. M. (1995) *Practical Flow Cytometry*, 3rd ed., Wiley-Liss, New York

[f] Darzynkiewicz, Z., Robinson, J. P., and Crissman, H. A., eds. (1994) *Flow Cytometry*, 2nd ed., Academic Press, San Diego

can be used.[152]

Whereas in conventional zone electrophoresis most of the electrical current is carried by the buffer, in **isotachophoresis**[153,154] the ions being separated carry most of the current. In **isoelectric focusing**,[28,155–157] a pH gradient is developed electrochemically in a vertical column or on a thin horizontal plate between an anode and a cathode. The pH gradient in a column is stabilized by the presence of a density gradient, often formed with sucrose, and the apparatus is maintained at a very constant temperature. Proteins within the column migrate in one direction or the other until they reach the pH of the isoelectric point where they carry no net charge and are "focused" into a narrow band. As little as 0.01 pH unit may separate two adjacent protein bands which are located at positions corresponding to their isoelectric points. A newer development is the use of very narrow pH gradients that are immobilized on a polyacrylamide matrix.[158–160] With this technique some hemoglobin mutants differing only in substitution of one neutral amino acid for another have been separated.[161] Special techniques are needed for highly basic proteins.[162]

A two-dimensional method in which proteins are separated by isoelectric focusing (preferably with an immobilized pH gradient) in the first dimension and by SDS-gel electrophoresis in the second has become a popular and spectacularly successful method for studying complex mixtures of proteins (Box 3-C).[163–166] Over 2000 proteins can be separated on a single plate. A similar procedure but without SDS can be used to examine undenatured proteins.[167–169] Computer-assisted methods are being developed to catalog the thousands of proteins being identified in this way[170–172] and also to allow rapid identification of spots by mass spectrometry. The technique can be applied to intact proteins in subpicomole quantities, even in whole cell lysates,[150,173,173a] or an enzyme such as trypsin can be used to cut the proteins into pieces on the gel plate and the mixtures of peptides can be analyzed by mass spectrometry.[174 – 176] Capillary electrophoresis or capillary isoelectric focusing can be applied before samples are sent to the mass spectrometer.

C. Determining the Relative Molecular Mass, M_r

The evaluation of M_r is often of critical importance. Minimum values of M_r can often be computed from the content of a minor constituent, e.g., the tryptophan of a protein or the iron of hemoglobin. However, physicochemical techniques provide the basis for most measurements.[177] Observations of osmotic pressure or light scattering can also be used and provide determinations of M_r that are simple in principle, but which have pitfalls.[178]

1. Ultracentrifugation

Some of the most reliable methods for determining M_r depend upon **analytical ultracentrifuges**. These instruments, capable of generating a centrifugal field as much as 4 x 10^5 times that of gravity, were developed in the 1920s and 1930s by T. Svedburg and associates in Uppsala, Sweden.[179–181] Driven by oil turbines, the instruments were expensive and difficult to use, but by 1948 a reliable electrically driven machine, the Beckman Model E ultracentrifuge, came into widespread use. It has had a major impact on our understanding of proteins, on methods of purification of proteins, and on our understanding of interactions of protein molecules with each other and with small molecules.[182,183] Nevertheless, it was not until 1990 that a truly "user-friendly" analytical untracentrifuge became available.[180,184–186] The Beckman Model XL-A centrifuge has a very small rotor driven by an air-cooled induction motor and is computer controlled. Data are recorded automatically in digital form and computer programs are available to carry out the necessary computations. The instrument can record ultraviolet-visible spectra at multiple radial positions in the sample cell (Fig. 3-7).

A straightforward determination of M_r is obtained by centrifuging until an equilibrium distribution of the molecules of a protein or other macromolecular material is obtained and by recording the variation in concentration from the center to the periphery of the centrifuge cell[177,183,185,187–190] (see also Section A,2). Using short cells, this **sedimentation equilibrium** can be attained in 1–5 hours instead of the 1–2 days needed with older instruments. For a single component system the concentration distribution at equilibrium is given by Eq. 3-12.

$$c(r) = c(a) \exp \left[M (1 - \bar{v}\rho)\, \omega^2 (r^2 - a^2) / 2RT \right] \qquad (3\text{-}12)$$

Here $c(r)$ is the concentration c at the radial position r (measured from the centrifuge axis), a is the radial distance of the meniscus, M is the molecular mass in daltons, and $\bar{v}$ is the partial specific volume in ml/gram. For most proteins $\bar{v}$ varies from 0.69–0.75. It is the reciprocal of the density of the particle. Rho (ρ) is the density (g/ml) of the solvent. A plot of log $c(r)$ against r^2 is a straight line of slope $M (1 - \bar{v}\rho) / 2RT$. The computer can also accommodate mixtures of proteins of differing molecular masses, interacting mixtures, etc.[185,191]

Sedimentation velocity. The relative molecular mass M_r can also be measured from observation of the velocity of movement of the boundary (or boundaries for multicomponent systems) between solution and solvent from which the macromolecules have sedi-

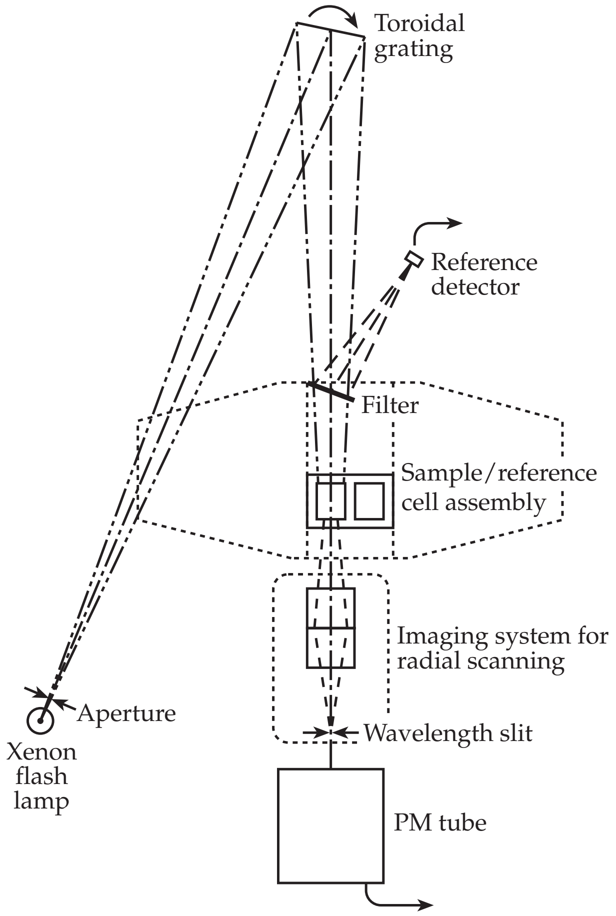


Figure 3-7 The scanning absorption optical system of the Beckman Optima™ XL-A ultracentrifuge. Courtesy of Beckman Coulter.

mented. This boundary, which can be visualized by optical methods, is quite sharp initially, but it broadens with time, because of diffusion, as the macromolecules sediment.

A molecule in a centrifuge is acted upon not only by the applied centrifugal force but also by an opposing **buoyant force** that depends upon the difference in density of the sedimenting particles and the solvent and by a frictional drag, which is proportional to a **frictional coefficient** *f*. Setting the sum of these forces to zero for the hydrodynamic steady state yields Eq. 3-13, which defines the **sedimentation constant** *s*.

$$s = \frac{v}{\omega^2 r} = \frac{M(1-\bar{v}\rho)}{Nf} \tag{3-13}$$

Here *f* is the frictional coefficient which is difficult to predict or to measure but is often assumed to be the same as the frictional coefficient that affects **diffusion**. It can be obtained from the diffusion coefficient *D* using Eq. 3-14.

$$D(\mathrm{cm}^2\ \mathrm{s}^{-1}) = k_B T/f = RT/Nf \tag{3-14}$$

By combining Eqs. 3-13 and 3-14 we obtain the Svedberg equation:

$$M = \frac{RTs}{D(1-\bar{v}\rho)} \tag{3-15}$$

Here *R* is in the cgs (cm-gram-second) unit of 8.31 x 10^7 erg mol^{-1} deg^{-1}. Using this equation the relative molecular mass M_r, which is numerically the same as *M*, can be evaluated from the sedimentation constant *s*. Since *s*, *D*, and $\bar{v}$ must all be measured with care, the method is demanding. It is often necessary to measure *s* and *D* at several concentrations and to extrapolate to infinite dilution. It is also customary to correct the data to give the values $s°_{20,w}$ and $D°_{20,w}$ expected at 20°C in pure water at infinite dilution.

To the extent that we can regard protein molecules as spherical we can substitute for *f* in Eq. 3-13 the frictional coefficient of a sphere:

$$f_{\mathrm{sphere}} = 6\pi\eta r_h \tag{3-16}$$

Here r_h is the hydrated radius or **Stokes radius** of the protein. On this assumption *s* will be expected to increase with the relative molecular mass approximately as $M_r^{2/3}$. A plot of log *s* against log M_r should be a straight line. Figure 3-8 shows such a plot for a number of proteins. The plots for nucleic acids, which can often be approximated as rods rather than spheres, fall on a different line from those of proteins. Furthermore, the sedimentation constant falls off more rapidly with increasing molecular mass than it should for spheres.

From analysis of a variety of well-characterized proteins, Squire and Himmel[192] observed that if proteins are assumed to contain 0.53 g H_2O per gram of protein and to have a mean value for $\bar{v}$ of 0.730 g / cm^3 the value of M_r can be predicted by Eq. 3-17 with the standard deviation indicated. Here, S is the sedimentation constant in Svedberg units (10^{-13}s).

$$M_r = 6850\ S^{3/2}\ )\ 0.090\ M_r \tag{3-17}$$

For proteins with various values of $\bar{v}$, Eq. 3-18 applies.

$$M_r = 922\ [S/(1-\bar{v}\rho)]^{3/2}\ )\ \pm 0.066\ M_r \tag{3-18}$$

BOX 3-C ISOTOPES IN BIOCHEMICAL INVESTIGATIONS

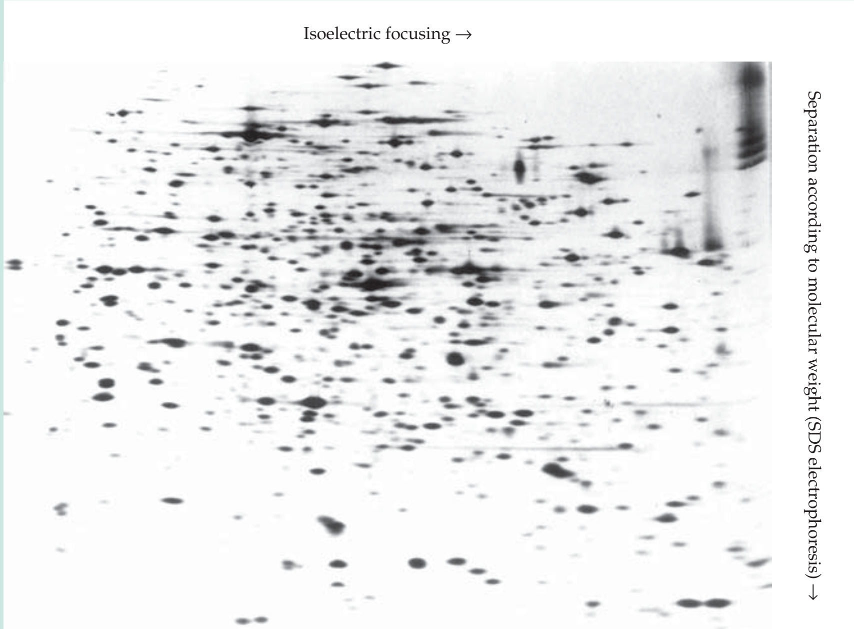
Both stable[a] and radioactive[b–e] isotopes are widely used in chemical and biological investigations. The study of metabolism was revolutionized by the introduction of isotopic tracers. In one of the first biological experiments with the stable isotope ^{15}N (detected by mass spectrometry), Schoenheimer and associates in 1937 established the previously unsuspected turnover of protein in living tissues (Chapter 24, Section B). In 1937 Ruben *et al.* reported the uptake of radioactive $^{11}CO_2$ by plants.[f] A few years later Calvin and associates first traced the pathway of carbon in photosynthesis using the much longer lived $^{14}CO_2$ (Box 17-F). Wood and Werkman, in 1941, employed the stable isotope ^{13}C in studies of bacterial and mammalian metabolism (Box 17-C). The radioactive ^{32}P and ^{35}S have served to elucidate the metabolism of phosphorus and sulfur. Titrium (^{3}H) has been used to label many organic substances including thymine, which has been used extensively in the study of nucleic acid metabolism. Radioactive isotopes provide the basis for sensitive analytical procedures such as **radioimmunoassays** of minute quantities of hormones (Box 31-D). Through **radioautography** these isotopes facilitate numerous analytical procedures (see accompanying photo) and have provided the basis for important end-group methods used in sequence determination of polynucleotides (Eq. 5-24).

Several isotopes used in biochemistry are listed in the following table. For each radioactive isotope, the half-life is given, as is the type of particle emitted, and the energy of the particle. Gamma rays, such as those given off in decay of ^{125}I or ^{131}I, are very penetrating and easy to count precisely, as is the energetic β radiation from ^{32}P. On the other hand, ^{3}H (tritium) is relatively difficult to detect[g] but its weak β particle, which can travel only a short distance through a sample, makes it uniquely suitable for radioautography on a microscopic scale. Positrons (β^+) travel some distance, e.g., up to a few millimeters in the case of ^{13}N. They are then destroyed by reacting with an electron to produce a pair of γ rays of energy 0.511 MeV, equal to the sum of the rest masses of an electron plus a positron. The half-life (Eq. 9-4) determines the isotopic abundance needed to achieve a given radiation rate, a practical matter in providing a sufficient rate of decay to permit counting with an acceptably low statistical error. Even very short-lived isotopes such as ^{13}N, have proved useful as tracers.[h] The amount of an isotope giving 3.7×10^{10} disintegrations per second (this is 1 g of pure radium, 0.3 mg of ^{3}H, or 0.22 g of ^{14}C) is known as the **curie** (Ci). One millicurie (mCi) provides 2.22×10^{9} disintegrations / min (dpm). Radiolabeled substances ordinarily contain only a small fraction of the unstable isotope together with a larger number of unlabeled molecules. Compounds are usually sold in millicurie or microcurie quantities and with a stated specific activity as mCi $mmol^{-1}$. For example, a compound labeled at a single position with ^{3}H and having a specific activity of 50 mCi $mmol^{-1}$ would contain about 0.17% ^{3}H at that position.

(MeV)		Maximum energy of radiation	
Isotope	Half-life	β	γ
^{2}H (deuterium)	Stable		
^{3}H (tritium)	12.26 years	0.018	
^{11}C	20.4 min 0.511		(β^+)
^{13}C	Stable		
^{14}C	5730 years	0.156	
^{13}N	9.96 min 0.511	1.2	(β^+)
^{15}N	Stable		
^{15}O	20.4 min 0.511		(β^+)
^{18}O	Stable		
^{18}F	110 min 0.511		(β^+)
^{22}Na	2.6 years 1.28	0.55	
^{32}P	14.3 days	1.71	
^{35}S	87.2 days	0.167	
^{36}Cl	3×10^{5} years	0.716	
^{40}K	1.3×10^{9} years 1.5	1.4	
^{45}Ca	165 days	0.26	
^{59}Fe	45 days 1.1	0.46	
^{65}Zn	250 days 1.14	0.32	
^{90}Sr	29 years	0.54	
^{125}I	60 days	0.036	
^{131}I	8.06 days 0.36	0.61	

Because of the development of new NMR techniques and improvements in mass spectrometry stable isotopes, such as ^{2}H, ^{13}C, ^{19}F, and ^{31}P, are being used more frequently to study metabolism.[j] Carbon-13 containing compounds can fulfill many

BOX 3-C ISOTOPES IN BIOCHEMICAL INVESTIGATIONS (continued)


Isoelectric focusing →
Separation according to molecular weight (SDS electrophoresis) →

Radioautogram showing the separation of proteins of *E. coli* labeled with ^{14}C amino acids. From O'Farrell.[i] Twenty-five μl of sample containing 180,000 cpm and ~10 μg of protein were subjected to isoelectric focusing in a 2.5 x 130-mm tube containing polyacrylamide gel to separate proteins according to isoelectric point. The gel was then extruded from the column and was placed on one edge of a slab of polyacrylamide gel. Then SDS electrophoresis in the second dimension separated the proteins according to size. Over 1000 spots could be seen in the original radioautogram, which was obtained by placing a piece of photographic film over the gel slab and exposing it to the radiation for 875 hours. For details see O'Farrell.[i]

of the same tracer functions as ^{14}C. Even the radioactive ^{3}H nucleus can be utilized for *in vivo* NMR.[k,l] Although radioisotope labeling is very sensitive, it gives little information unless compounds are isolated and laboriously degraded to determine the positions of the labels. NMR spectroscopy is less sensitive but can give direct chemical information about the positions of ^{13}C in compounds within living cells. A compound containing only ^{13}C in one or in many positions can be safely administered to human individuals as well as to other organisms and spectra of products that arise can be observed. High-resolution deuterium NMR spectroscopy has been used to follow ^{2}H incorporated at C-1, C-2, or C-6 positions in glucose.[m]

As a result of metabolic reactions an isotope may appear at more than one position in a product, yielding two or more isotope isomers or **isotopomers**. These are seen individually by NMR spectroscopy and the concentration and isotope labeling patterns of the labeled compounds can be followed over a period of time. The use of this **isotopomer analysis** in studies of the citric acid cycle is illustrated in Box 17-C and its use in studies of glucose metabolism is considered in Chapter 17, Section L.

A change in isotopic mass, especially from ^{1}H to ^{2}H or ^{3}H, often produces a strong effect on reaction rates and the study of **kinetic isotope effects** has provided many insights into the mechanisms of enzymatically catalyzed reactions. Isotopes have permitted a detailed understanding of the stereochemistry of enzymatic reactions, an impressive example being the synthesis and use of chiral acetate (Chapter 13)[n] and chiral phosphate groups (Chapter 12). Specific isotopic properties provide the basis for NMR (Section G).

---

[a] Matwiyoff, N. A., and Ott, D. G. (1973) *Science* **181**, 1125–1132

[b] Wang, C. H., Willis, D. L., and Loveland, W. D. (1975) *Radiotracer Methodology in the Biological, Environmental, and Physical Sciences*, Prentice-Hall, Englewood Cliffs, New Jersey

[c] Wang, Y., ed. (1969) *Handbook of Radioactive Nuclides*, CRC Press, Cleveland, Ohio

[d] Thornburn, C. C. (1972) *Isotopes and Radiation in Biology*, Butterworth, London

[e] Slater, R. J., ed. (1990) *Radioisotopes in Biology: A Practical Approach*, IRL Press, Oxford

[f] Ruben, S., Hassid, W. Z., and Kamen, M. D. (1939) *J. Am. Chem. Soc.* **61**, 661–663

[g] Bransome, J., ed. (1970) *Liquid Scintillation Counting*, Grune & Stratton, New York

[h] Cooper, A. J. L. (1985) *Adv. Enzymol.* **57**, 251–356

[i] O'Farrell, P. H. (1975) *J. Biol. Chem.* **250**, 4007–4021

[j] Wolfe, R. R. (1992) *Radioactive and Stable Isotope Tracers in Biomedicine*, Wiley, New York

[k] Newmark, R. D., Un, S., Williams, P. G., Carson, P. J., Morimoto, H., and Klein, M. P. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 583–587

[l] Bergerat, A., Guschlbauer, W., and Fazakerley, G. V. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 6394–6397

[m] Aguayo, J. B., Gamcsik, M. P., and Dick, J. D. (1988) *J. Biol. Chem.* **263**, 19552–19557

[n] Cornforth, J. W. (1976) *Science* **193**, 121–125

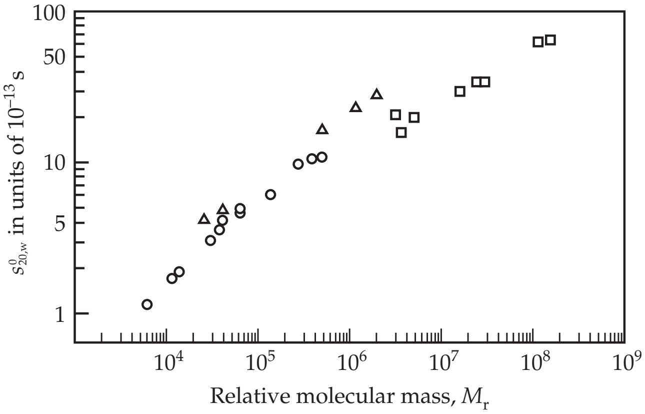


Figure 3-8 Plots of the logarithm of the sedimentation constant *s* against the logarithm of the molecular weight for a series of proteins and nucleic acids: (○) globular proteins, (△) RNA, and (□) DNA. Proteins include a lipase (milk), cytochrome *c*, ribonuclease (pancreatic), lysozyme (egg white), follicle-stimulating hormone, bacterial proteases, human hemoglobin, prothrombin (bovine), malate dehydrogenase, γ-globulin (horse), tryptophanase (*E. coli*), glutamate dehydrogenase (chicken), and cytochrome *a*. Double-stranded DNA molecules are those of bacteriophage ϕX174 (replicative form), T_7, λ_{b2}, T_2, and T_4, and that of a papilloma virus. The RNA molecules are tRNA, rRNA, and mRNA of *E. coli*, and that of turnip yellow mosaic virus.[190a,191a]

2. Gel Filtration and Gel Electrophoresis

Several newer methods of molecular mass determination were developed in the 1960s–1980s. One is gel filtration. A column of gel beads such as Sephadex is prepared carefully and is calibrated by passing a series of protein solutions through it. The volume V_e at which a protein peak emerges from the column can be expressed as the sum of two terms (Eq. 3-19) in

$$V_e = V_o + \sigma V_i \qquad (3\text{-}19)$$

which V_o is the **void volume**, i.e., the elution volume that is observed for very large particles that are completely excluded from the gel, and V_i is the internal volume within the beads of gel. The value of σ is inversely related to the diffusion constant *D*, which for a spherical particle *D* is related by Eq. 3-20 to the Stokes radius r_h. This equation comes directly from Eqs. 3-16 and 3-14.

$$r_h = \frac{k_B T}{6\pi\eta D} \qquad (3\text{-}20)$$

This suggests a proportionality between σ and the molecular radius. In fact, Eq. 3-21, in which a and b are constants provides a fairly good approximation for σ and V_e is correlated approximately with log M_r as shown in Fig. 3-9.[193,194]

$$\sigma = a \log r_h + b \qquad (3\text{-}21)$$

A series of reference proteins of known molecular masses are used to calibrate the column and M_r for an unknown protein is estimated from its position on the graph.[195,196] Another modification of the method depends upon chromatography in a high concentration of the denaturing salt guanidinium chloride. The assumption is made that proteins are denatured into random coil conformations in this solvent.[196]

Probably the most widely used method for determining the molecular mass of protein subunits is gel electrophoresis in the presence of the denaturing detergent sodium dodecyl sulfate (SDS). The protein molecules are not only denatured but also all appear to become more or less evenly coated with detergent.[197] The resulting rodlike molecules usually show a uniform dependence of electrophoretic mobility on molecular mass (plotted as log M_r). An example is shown in Fig. 3-10. Again, the molecular mass of the protein under investigation is estimated by comparison of its rate of migration with that of a series of marker proteins.[195,198]

3. Mass Spectrometry

Mass spectrometry has played a role in biochemistry since the early 1940s when it was introduced for use in following isotopic labels during metabolism.[199–200c] However, it was not until the 1990s that suitable commercial instruments were developed to permit mass spectrometry using two new methods of ionization. The techniques are called **matrix-assisted laser desorption / ionization time-of-flight (MALDI-TOF)** and **electrospray ionization (ESI)** mass spectrometry.

In the MALDI technique a pulsed laser beam strikes a solid sample and heats, vaporizes, and ionizes compounds with little decomposition.[201–209] Proteins or other biopolymers are mixed with a "matrix" that absorbs the heat of the laser beam. The protein sample together with the matrix is dried. Most proteins form crystals and the laser beam is directed toward individual protein crystals or aggregates. Various materials are used for the matrix. Compounds as simple as glycerol, succinic acid, or urea can be used with an infrared laser. For proteins an ultraviolet nitrogen laser tuned to 337 nm is usually employed with an ultraviolet light-absorbing matrix such as hydroxybenzoic acid, 2,5-dihydroxybenzoic acid, α-hydroxy-

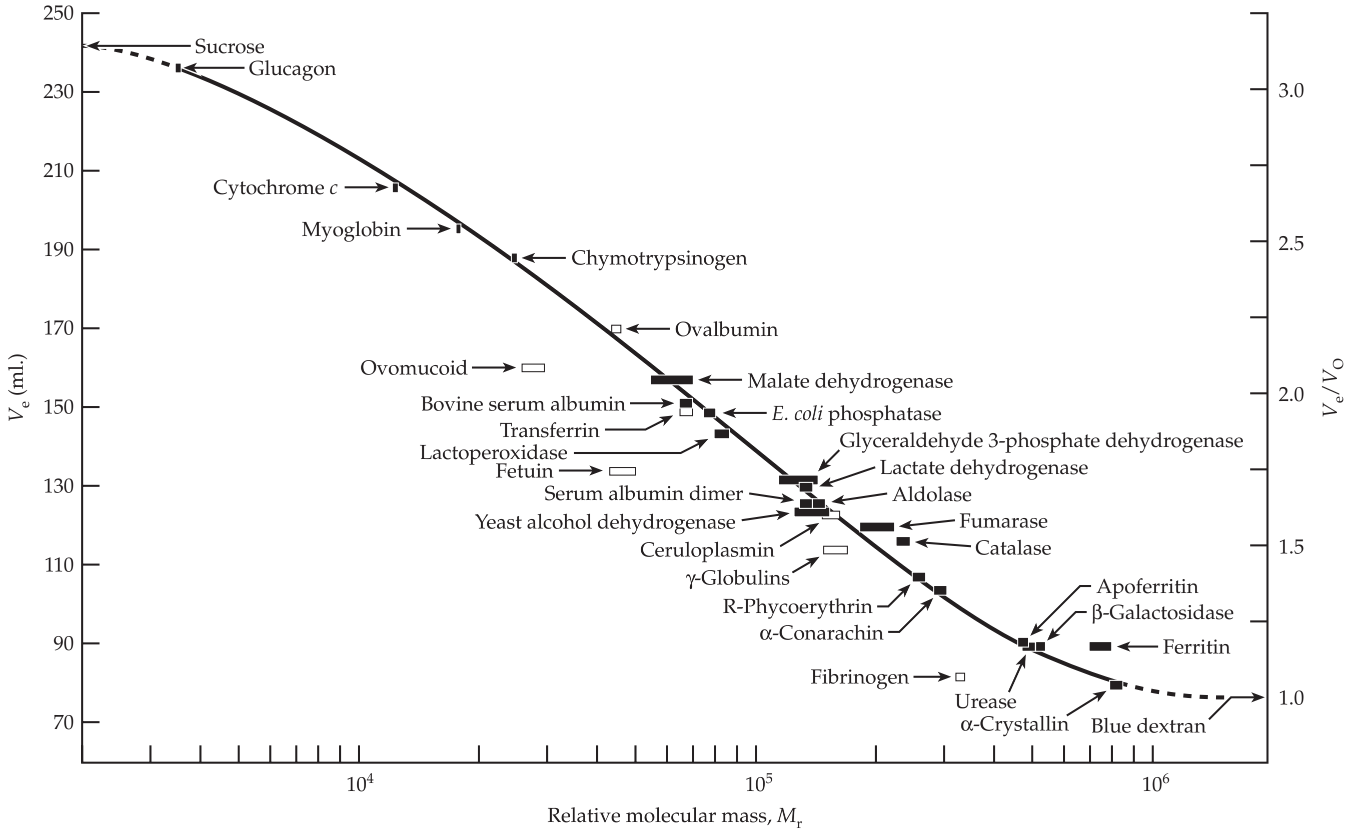


Figure 3-9 Elution volume of various proteins on a column of Sephadex G-200 as a function of molecular mass. The right-hand vertical axis shows the ratio of the elution volumes to that of blue dextran, a high-molecular-mass polysaccharide that is excluded from the internal volume. After Andrews.[193]

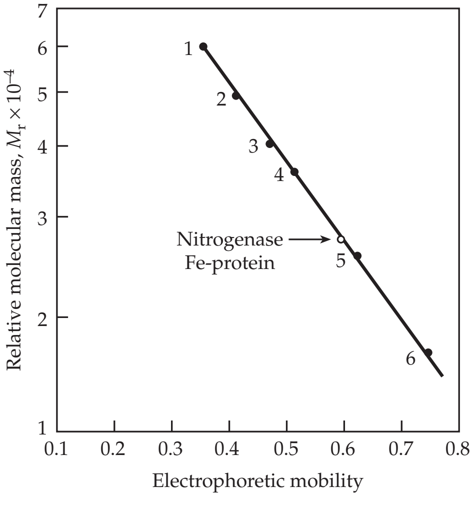


Figure 3-10 Estimation of the molecular mass of the polypeptide chain of the nitrogenase Fe-protein using SDS-polyacrylamide electrophoresis; from a set of four standard curves. The marker proteins are (1) catalase, (2) fumarase, (3) aldolase, (4) glyceraldehyde-phosphate dehydrogenase, (5) α-chymotrypsinogen A, and (6) myoglobin. (o) indicates position of azoferredoxin. From Nakos and Mortenson.[195]

cinnamic acid, or sinapinic acid (Chapter 25). The matrix ionizes, desorbs from the surface, and transfers energy to the crystalline protein, causing it to ionize and desorb from the surface. Oligosaccharides and oligonucleotides can be ionized in a similar way.

MALDI spectra are relatively simple (Fig. 3-11), often containing a single major peak corresponding to the singly charged molecular ion $[M + H]^+$ of mass $m + 1$ and perhaps a doubly charged molecular ion $[M + 2H]^{2+}$ of mass $(m + 2)/2$. For oligomeric proteins the major peak is often that of the monomer with weaker peaks for oligomers. The instrument can also be adjusted to generate negative ions whose detection is useful for study of phosphorylated peptides, many oligosaccharides, and oligonucleotides. With a TOF spectrometer there is no upper limit to the mass range and masses of over 100 kDa can be measured to about ±0.1%. Femtomole quantities can be detected.

The MALDI method is especially useful for complex mixtures of peptides and can be utilized in peptide sequencing. The technique is also appropriate for studying mixtures of glycoproteins. Negative-ion MALDI can be applied to oligonucleotide mixtures. Further improvements in resolution in both MALDI and ESI methods are anticipated as a result of developement of Fourier transform mass spectrometers.[202]

In ESI mass spectrometry[201,203 – 205,210 – 213] the sample, dissolved in an appropriate solvent (usually a 50:50 mixture of methanol and water for proteins), is infused directly into the ionization chamber of the spectrometer through a fused silica capillary. At the end of the capillary the solution is subjected to electrical stress created by a voltage difference of about 5 kV between the electrospray needle and the sampling orifice (the counter-electrode). The process results in the formation of singly and / or multiply charged molecular ions which are guided into the analyzer for mass analysis. For proteins every arginine, lysine, and histidine may bind a hydrogen ion to form a variety of positive ions. A 100-kDa protein may easily bind 100 protons bringing

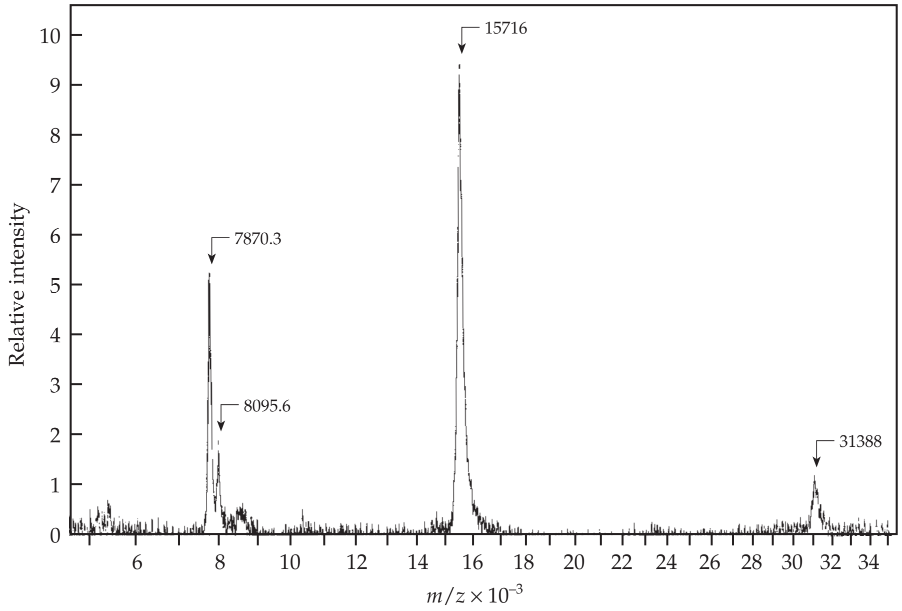


Figure 3-11 Matrix-assisted laser desorption / ionization time-of-flight (MALDI-TOF) mass spectrum of bovine erythrocyte Cu-Zn superoxide dismutase averaged over ten shots with background smoothing. One-half µl of solution containing 10 pmol of the enzyme in 5 mM ammonium bicarbonate was mixed with 0.5 µl of 50 mM α-cyanohydroxycinnamic acid dissolved in 30% (v / v) of acetonitrile-0.1% (v / v) of trifluoroacetic acid. The mixture was dried at 37° C before analysis. The spectrum shows a dimer of molecular mass of 31,388 Da, singly charged and doubly charged molecular ions at 15,716, and 7870 Da, respectively. The unidentified ion at mass 8095.6 may represent an adduct of the matrix with the doubly charged molecular ion. Courtesy of Louisa Tabatabai.

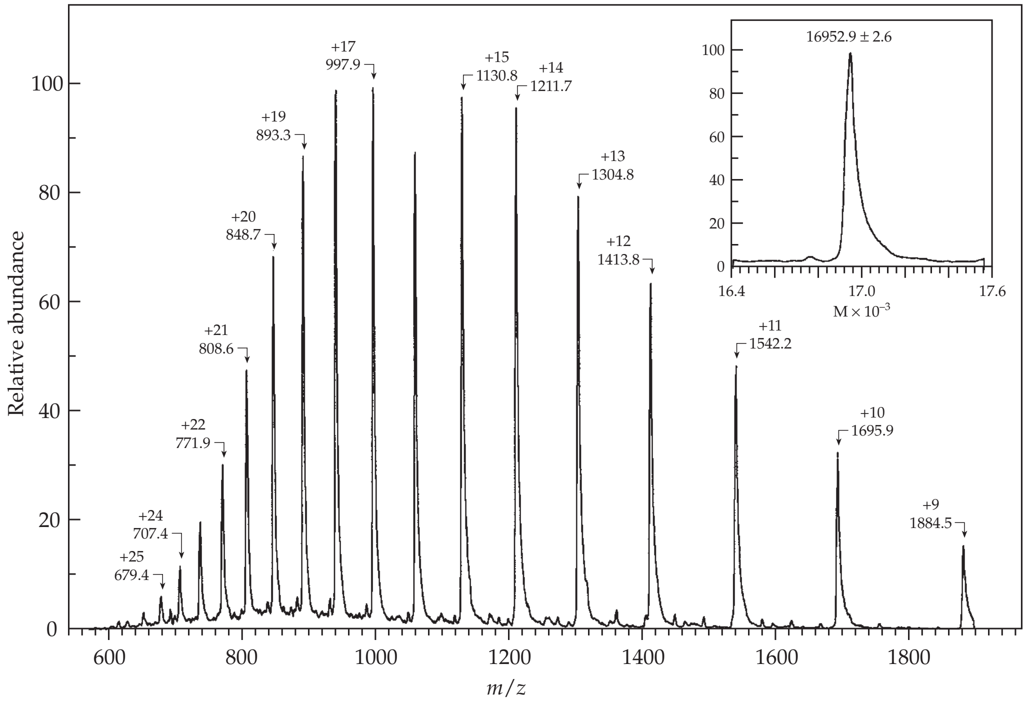


Figure 3-12 Positive ion electrospray mass spectrum of horse apomyoglobin (M_r 16,950.4). The net charge on each ion as well as the mass to charge ratio m/z is indicated at the top of each peak. The inset shows a computer “deconvolution” of the spectrum with the calculated value of molecular mass. Courtesy of Kamel Harrata.

the *m*/*z* ratio for the fully charged protein to 1000, well below the maximum *m*/*z* ratio of ~2400 for a typical quadrupole mass spectrometer. Since not all basic groups are protonated the spectra consist of families of peaks of differing *m*/*z* (Fig. 3-12). For a single pure protein the molecular mass can be calculated from the ratio of *m*/*z* values from any pair of adjacent peaks. It is better, especially if there is a mixture of proteins, to use a computer to extract M_r.[214] The accuracy can be quite high, typically ±0.01%: one mass unit in 10,000.

Some complexity arises from the fact that each carbon atom in a protein contains about 1% of ^{13}C. This means that for a protein of mass >10 kDa there will be a confusing array of peaks in the mass spectrum and it may be difficult to pick out the relatively minor "monoisotopic" peak that arises from molecules containing only ^{12}C, ^{1}H, ^{14}N, ^{16}O, and ^{32}S. In fact, the peak representing the most abundant mass will be a few mass units higher than the monoisotopic peak.[215,216] (see Study Question 15). New computer programs have been devised to assist in the analysis. Use of **^{13}C and ^{15}N-depleted nutrients** also extends the applicability of mass spectrometry.[217]

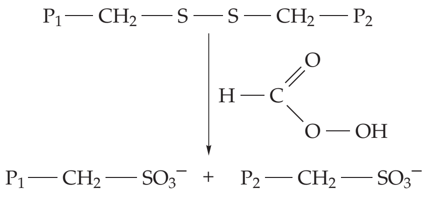
Electrospray mass spectrometry utilizes a soft ionization technique at nearly atmospheric pressure. As a result, intact molecular ions are formed in high yield. The instrument can be interfaced readily to HPLC or capillary electrophoresis columns and subfemtomole amounts of proteins can be detected. A disadvantage is that salt concentrations must be kept low (< mM) and that the protein tends to bind Na^+, K^+, and anions that may confuse interpretation of spectra.

D. Determining Amino Acid Composition and Sequence

When they are first isolated, proteins are usually characterized by M_r, isoelectric point, and other easily measured properties. Among these is the amino acid composition[112] which can be determined by completely hydrolyzing the protein to the free amino acids. Later, it is important to establish the primary structure or amino acid sequence.[51] This has been accomplished traditionally by cutting the peptide chain into smaller pieces that can be characterized easily. However, most protein sequences are now deduced initially from the corresponding DNA sequences, but further chemical characterization is often needed.

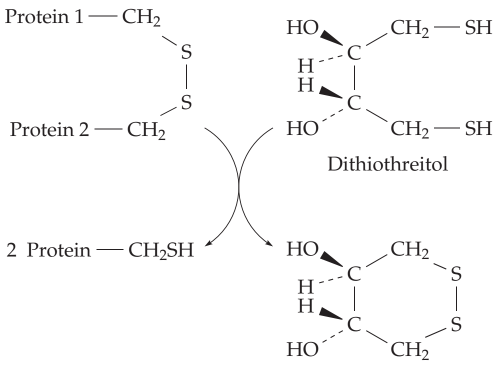
1. Cleavage of Disulfide Bridges

Before a polypeptide chain can be degraded it is usually necessary to break any disulfide bridges.[218–220] For some proteins such as the keratins of hair, these linkages must be broken even to get the protein into solution. Oxidation with performic acid (Eq. 3-22) has been used on ribonuclease but is not often employed because the performic acid also oxidizes tryptophan residues.


P₁ — CH₂ — S — S — CH₂ — P₂
H — C(=O) — O — OH
P₁ — CH₂ — SO₃⁻ + P₂ — CH₂ — SO₃⁻

(3-22)

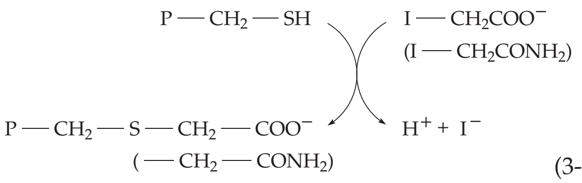
Reduction by **dithiothreitol** or **dithioerythritol** (Eq. 3-23) is usually successful.



(3-23)

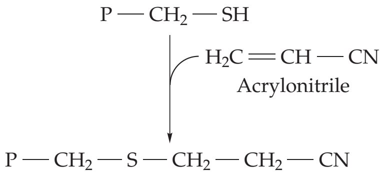
Upon oxidation these dithiols cyclize to form stable disulfides, driving the reaction to completion. These same compounds are also widely used to protect SH groups in enzymes against accidental oxidation by oxygen and to dissolve highly crosslinked insoluble proteins. Mercaptoethanol, $HS-CH_2-CH_2-OH$, may be used for the same purposes but requires higher concentrations and has a disagreeable odor.

Once cleaved, disulfide bridges may be prevented from reforming by conversion of the resulting thiol groups to stable derivatives, e.g., with iodoacetate and iodacetamide (Eq. 3-24).



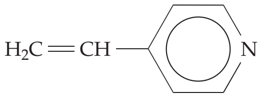
(3-24)

A better reagent is acrylonitrile (Eq. 3-25).



(3-25)

In modern sequencing methods vinylpyridine, which reacts in a similar way, is often used. It can be detected during amino acid analysis or sequencing after derivatization with phenylisothiocyanate (Eq. 3-30).



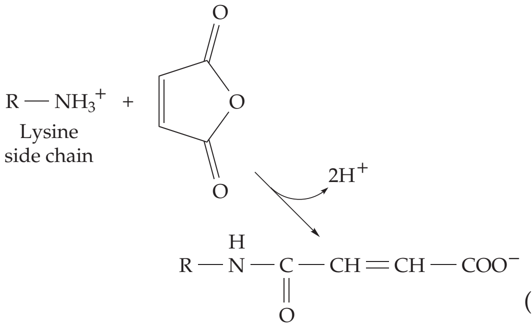
Vinylpyridine

2. Hydrolysis and Other Chain Cleavage Reactions

Most biopolymers are inherently unstable with respect to cleavage to monomer units by reaction with water. Hydrolysis can be catalyzed by protons, by hydroxyl ions, or by the protein-hydrolyzing enzymes that are discussed in Chapter 12. Complete hydrolysis of proteins is usually accomplished by heating under nitrogen with 6 M HCl at 150°C for 65 min. Some amino acids, especially tryptophan, are destroyed and the amides in the side chains of asparagine and glutamine are converted to the free acids. Some peptide linkages such as Val-Val are very resistant and tend to be incompletely hydrolyzed. No procedure has been found which gives the ideal complete hydrolysis. Use of 4 M methanesulfonic acid containing 3-(2-aminoethyl)indole instead of 6 M HCl gives less decomposition of tryptophan.[221] Base-catalyzed hydrolysis of proteins also gives good yields of tryptophan but causes extensive racemization of amino acids.[222]

Complete enzymatic digestion of proteins can be accomplished with a mixture of enzymes including proteases produced by fungi (Pronase). However, the enzymes attack each other, making quantitative analysis difficult. The problem can be circumvented by immobilizing the hydrolytic enzymes in a column of agarose gel. The protein to be hydrolyzed is passed through the gel and the constituent amino acids emerge from the bottom of the column.[223,224]

Selective enzymatic hydrolysis. The traditional strategy in sequence determination is to cut protein chains into smaller pieces which can be separated by chromatography or electrophoresis and sequenced individually. Enzymatic cleavage is especially useful because of its specificity. **Trypsin**, a so-called **endopeptidase**, cleaves peptide chains at a rapid rate only if the carbonyl group of the amide linkage cleaved is contributed by one of the basic amino acids lysine, arginine, or aminoethylcysteine (see Fig. 12-10). If the protein is treated with maleic anhydride (Eq. 3-26),



(3-26)

BOX 3-D THE PROTEOME

The ability to separate rapidly and detect minute amounts of proteins has spawned a new concept: The **proteome** is envisioned as a record of all proteins being actively synthesized by a cell – or of all genes being actively "expressed."[a–c] The concept evolved from efforts to automate the cataloging of spots on two-dimensional gels such as that shown in Box 3-B where each spot represents a single protein. The ability to unambiguously identify the spots by mass spectroscopy has brought new optimism to the attempt to use gels automatically to analyze all of the proteins formed by a cell.[a,c] The methods are potentially very important to developing new diagnostic procedures for human medicine. Watching changes in the proteome, including posttranslational modifications in proteins, as cells develop and grow will provide new insights into biochemical regulation.

A similar concept is that of a **complete transcriptional map**, a record of all of the different RNA molecules being synthesized by a cell.[d,e] These include many different mRNAs, each of which may give rise to more than one protein, as well as many RNAs with other functions.

---

[a] Kahn, P. (1995) *Science* **270**, 369–370

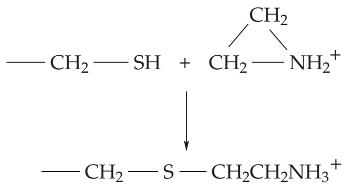
[b] Swinbanks, D. (1995) *Nature (London)* **378**, 653

[c] Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko.A, Boucherie, H., and Mann, M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14440–14445

[d] Richard, G.-F., Fairhead, C., and Dujon, B. (1997) *J. Mol. Biol.* **268**, 303–321

[e] Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P. O., and Davis, R. W. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10614–10619

the lysine residues are protected and trypsin will cleave only at the Arg-X positions. If the resultant peptides are separated and held at pH 3.5 overnight, the blocking groups are hydrolyzed off and a second trypsin treatment can be used to cleave at the Lys-X positions.[225] The number of cleavage sites for trypsin can be increased by converting an –SH group to a positively charged one by aminoethylation (Eq. 3-27).[226] The reaction can be accomplished either with ethyleneimine (caution: carcinogen) as shown in this equation or by bromoethylamine, which eliminates Br^- to form ethyleneimine.



$$—CH_2—SH + \text{(}CH_2—CH_2—NH_2^+\text{, ring)} \rightarrow —CH_2—S—CH_2CH_2NH_3^+ \qquad (3\text{-}27)$$

Because of its specificity for basic residues, trypsin converts a protein into a relatively small number of **tryptic peptides** which may be separated and characterized. Trypsin acts primarily on denaturated proteins, and to obtain good results the disulfide bridges must be broken first. **Chymotrypsin** is less specific than trypsin and **pepsin** is even less specific (Table 3-2). Nevertheless, they can be used to cut a peptide chain into smaller fragments whose sequences can be determined. To establish the complete amino acid sequence for a protein, "overlapping" peptide fragments must be found that contain sequences from ends of two different tryptic fragments. In this way the tryptic peptides can be placed in the order in which they occurred in the native protein. This tedious procedure is rarely used today. Peptide sequencing is still important but is usually coordinated with gene sequencing, X-ray structure determination, or mass spectroscopy which minimize the need for overlapping fragments.

While trypsin cuts the peptide linkages Lys-X and Arg-X, a fungal protease cleaves only X-Lys.[227] A protease from the submaxillary glands of mice cleaves only Arg-X,[228] one from *Staphylococcus* specifically at Glu-X,[229,230] and one from kidneys at Pro-X.[231]

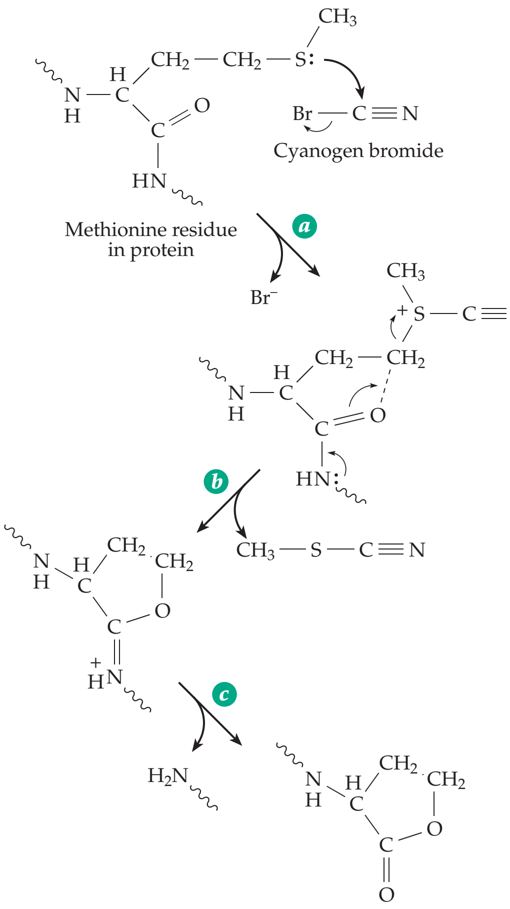
Several enzymes catalyze stepwise removal of amino acids from one or the other end of a peptide chain. **Carboxypeptidases**[232] remove amino acids from the carboxyl-terminal end, while **aminopeptidases** attack the opposite end. Using chromatographic methods, the amino acids released by these enzymes may be examined at various times and some idea of the sequence of amino acids at the chain ends may be obtained. A **dipeptidyl aminopeptidase** from bovine spleen cuts dipeptides one at a time from the amino terminus of a chain. These can be converted to volatile trimethylsilyl derivatives and identified by mass spectrometry.[233] If the chain is shortened by one residue using the Edman degradation (Section 3) and the dipeptidyl aminopeptidase is again used, a different set of dipeptides that overlaps the first will be obtained and a sequence can be deduced. Carboxypeptidase Y can be used with MALDI mass spectrometry to deduce the C-terminal amino acid sequence for a peptide. However, Ile and Leu cannot be distinquished.

TABLE 3-2
Specificities of Commonly Used Protein-Hydrolyzing Enzymes

Trypsin	Lys-X, Arg-X	X not Pro
Chymotrypsin		
rapidly:	Phe-X, Tyr-X, Trp-X	X not Pro
slowly:	Y-X Y=Leu, Asn, Gln, His, Met, Ser, Thr	X not Pro
Staphylococcus aureus		
protease V-8	Glu-X	X not Pro
Clostripain	Arg-X	
Pepsin		
preferentially:	X-Phe-X, X-Tyr-X, X-Leu-X	
less so:	X-Ala-X	
Thermolysin		
rapidly:	X-Y Y=Ile, Leu, Val, Ala, Phe, Met	
slowly:	X-Y Y=Tyr, Gly, Thr, Ser	

Nonenzymatic cleavages. Of the various nonenzymatic methods that have been proposed, one has been outstandingly useful. Cyanogen bromide, $N \equiv C - Br$, cleaves peptide chains adjacent to methionine residues. The sulfur of methionine displaces the bromide ion (Eq. 3-28) and because of a favorable spatial relationship, the resulting sulfonium compound undergoes C–S bond cleavage through participation of the adjacent peptide group (Eq. 3-28, step *b*). The C=N of the product is then hydrolyzed with cleavage of the peptide chain in step *c*.

The linkage Asp-Gly can often be cleaved specifically by treatment with hydroxylamine at high pH.[195] Procedures for specific cleavage of tryptophanyl bonds have been de-



(3-28)

Peptidyl homoserine lactone

vised.[234] The Asp-Pro linkage is susceptible to cleavage by trifluoroacetic acid, which is used in the automated Edman degradation employed in peptide sequencing (Eq. 3-30). Cleavage by trifluoroacetic acid can be used to generate peptides for subsequent sequence determination.

Separating the peptides. A procedure that has been very important in the development of protein chemistry is **peptide mapping** or "fingerprinting." The procedure begins with cleavage of the disulfide linkages, denaturation, and digestion with trypsin or some other protease. The sizes and amino acid compositions of the resulting series of peptides are characteristic of the protein under study. The mixture of peptides is placed on a thin layer plate and subjected to chromatography in one direction, then to electrophoresis in the other direction, with the peptides separating into a characteristic pattern or fingerprint. Fingerprinting has been especially useful in searching for small differences in protein structure, for example, between genetic variants of the same protein (Fig. 7-27). Currently, the peptides are usually separated on ion exchange or gel filtration columns, by reversed-phase HPLC, or by capillary electrophoresis and are then often passed, in subpicomole amounts, into a mass spectrometer.[176,235,236]

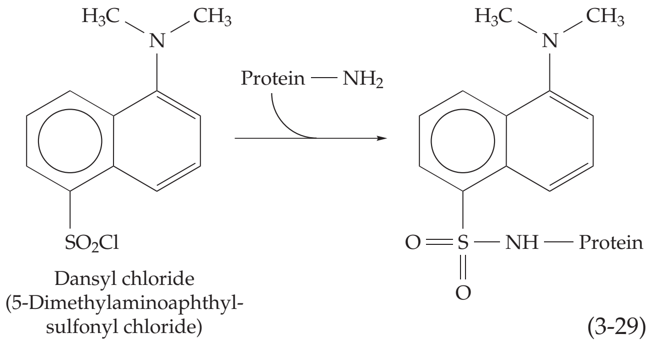
3. Determining Amino Acid Sequence

The covalent structure of insulin was established by Frederick Sanger in 1953 after a 10-year effort. This was the first protein sequence determination.[237,238] Sanger used partial hydrolysis of peptide chains whose amino groups had been labeled by reaction with 2,4-dinitrofluorobenzene[239] to form shorter end-labeled fragments. These were analyzed for their amino acid composition and labeled and hydrolyzed again as necessary. Many peptides had to be analyzed to deduce the sequence of the 21-residue and 30-residue chains that are joined by disulfide linkages in insulin.[237,238]

The Sanger method is mainly of historic interest, although end-labeling may still be used for various purposes. A more sensitive labeling reagent than was used by Sanger is **dansyl chloride**. It reacts to form a sulfonamide linkage that is stable to acid hydrolysis and is brilliantly fluorescent (Eq. 3-29). The related reagent dimethylaminoazobenzene-4'-sulfonyl chloride gives highly colored derivatives easily seen on thin-layer chromatography plates.[240]

The Edman degradation. One of the most important reagents for sequence analysis is **phenylisothiocyanate,** whose use was developed by P. Edman.[241–243] This reagent also reacts with the N-terminal amino group of peptides (Eq. 3-30, step *a*). The resulting adduct undergoes cyclization with cleavage of the peptide linkage (Eq. 3-30, step *b*) under acidic conditions. After rearrangement (step *c*) the resulting **phenylthiohydantoin** of the N-terminal amino acid can be identified. The procedure can then be repeated on the shortened peptide chain to identify the amino acid residue in the second position. With careful work the Edman degradation can be carried down the chain for several tens of residues.

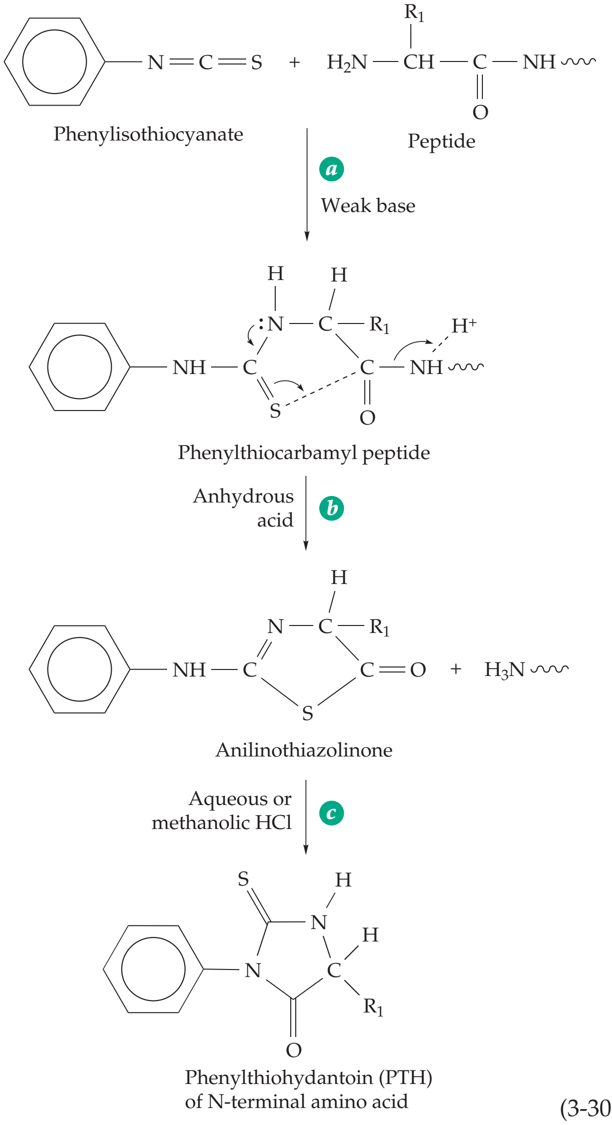
Ingenious **protein sequenators** have been devised to carry out the Edman degradation automatically.[242,244–246] Each released phenylthiohydantoin is then identified by HPLC or other techniques. Commercial sequenators have often required 5–20 nmol of peptide but new microsequenators can be used with amounts as low as 5–10 picomoles or less.[247,248]



(3-29)

Microsequencers permit sequence analysis on minute amounts of protein. Microsequencing can be used in conjunction with two-dimensional electrophoretic separations of proteins such as that shown in Box 3-C. The proteins in the polyacrylamide gel are electrophoretically transferred onto a porous sheet (membrane) of an inert material such as polyvinyl difluoride.[249–251] After staining, a selected spot is cut out and placed into the sequencer. To avoid the problems associated with blocked N termini, the protein may be treated with proteases on the membrane and the resulting peptide fragments may then be separated on a narrow-bore HPLC column and sequenced.[240]

Because many proteins are modified at the N terminus, blocking application of the Edman degradation, it would be useful to have a similar method for sequencing from the C terminus. It has been difficult to devise a suitable strategy, but there has been some success.[252–254]


Phenylisothiocyanate
Peptide
R₁
a
Weak base
Phenylthiocarbamyl peptide
H⁺
Anhydrous acid
b
Anilinothiazolinone
c
Aqueous or methanolic HCl
Phenylthiohydantoin (PTH) of N-terminal amino acid

(3-30)

Protein sequences from the genes. Complete sequences of large numbers of genes have been determined and the corresponding sequences of proteins can be read directly from those of the corresponding genes. One method for sequencing a gene is to isolate a specific messenger RNA, which does not contain intervening sequences. A DNA copy (cDNA) is made from the mRNA and is used to ascertain the sequence of the encoded protein. The genomic DNA is also often sequenced. Introns are recognized by the nucleotide sequences at their ends and the correct amino acid sequence for the encoded protein is deduced.

In many instances, however, a gene can be identified only after part of the protein, often an N-terminal portion has been sequenced. This knowledge permits synthesis of an **oligonucleotide probe** that can be used to locate the gene (Chapter 5). Nucleotide sequences can be verified by comparison with sequences of tryptic or other fragments of a protein. Similarly, protein sequences are often checked by sequencing the corresponding genes as well as by study of X-ray structures. Substantial numbers of errors are made in sequencing of both DNA and protein so that checking is important.

Mass spectrometry in sequencing. Proteins can also be sequenced by mass spectrometry or by a combination of Edman degradation and mass spectrometry.[213,255,255a] Until recently the peptides had to be converted to volatile derivatives by extensive methylation and acetylation or by other procedures. However, newer ionization methods including MALDI (Fig. 3-11) and ESI (Fig. 3-12) have made it possible to obtain mass spectra on unmodified peptides. In one procedure a nonspecific protease cleaves a peptide chain into a mixture of small oligopeptides which are separated by HPLC into 20–40 fractions, each of which may contain 10–15 peptides but which can be sent directly into the ionization chamber of the mass spectrometer.[256] Peptides can be generated from a protein using immobilized enzymes, separated on a chromatographic column, and introduced sequentially into the mass spectrometer. Examination of peptide mixtures by mass spectrometry provides a way of verifying sequences deduced from DNA sequencing.[257] Mass spectrometry is also used widely to study covalently modified proteins.[173,216,257a,257b] As a rule, these cannot be recognized from gene sequences.

4. Locating Disulfide Bridges

A final step in sequencing is often the location of S–S bridges. The reduced and alkylated protein can be cleaved enzymatically (e.g., with elastase, pepsin, or thermolysin) to relatively small fragments, each of which contains no more than one modified cysteine.

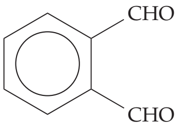
The same enzymatic cleavage can then be applied to the unreduced enzyme. Pairs of peptide fragments remain linked by the S–S bridges. These crosslinked pairs can be separated, the disulfide bridges cleaved, and the resulting peptides identified, each as one of the already sequenced fragments. Mass spectrometry provides a rapid method for their identification.[258]

Another elegant way of locating S–S bridges employs **diagonal electrophoresis**. Electrophoresis of the digest containing the crosslinked pairs is conducted in one direction on a sheet of filter paper. Then the paper is exposed to performic acid vapor to cleave the bridges according to Eq. 3-22 and electrophoresis is conducted in the second direction and the paper is sprayed with ninhydrin. The spots falling off the diagonal are those that participated in S–S bridge formation. They can be associated in pairs from their positions on the paper and can be identified with peptides characterized during standard sequencing procedures.[259]

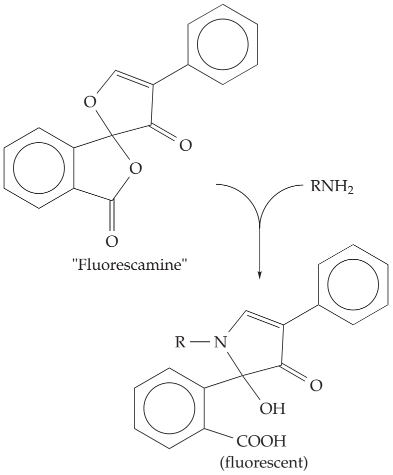
Diagonal electrophoresis and its relative diagonal chromatography are useful for other purposes as well. After electrophoresis or chromatography is conducted in one direction, the paper or thin-layer plate may be sprayed with a reagent that will react with some components or may be irradiated with light before the separation is repeated in the second direction (Fig. 3-5).[102,260]

5. Detecting Products

Important to almost all biochemical activity is the ability to detect, and to measure quantitatively, tiny amounts of specific compounds. “Color reagents,” which develop characteristic colors with specific compounds, are especially popular. For example, **ninhydrin** (Box 3-E) can be used as a “spray reagent” to detect a small fraction of a micromole of an amino acid or peptide in a spot on a chromatogram. It can also be used for a quantitative determination, the color being developed in a solution. More sensitive than absorption of light (color) is fluorescence. **Fluorescamine** (Eq. 3-31) reacts with any primary amine to form a highly fluorescent product. As little as 50 pmol of amino acid can be determined quantitatively.[261] A yet more sensitive fluorogenic reagent for detection of amino acids, peptides, and amines of all types is *o*-phthaldialdehyde.[262,263]

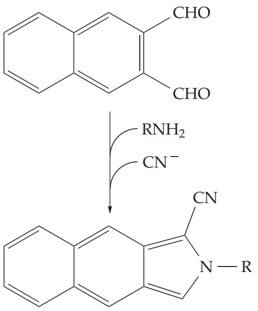


o-Phthaldialdehyde



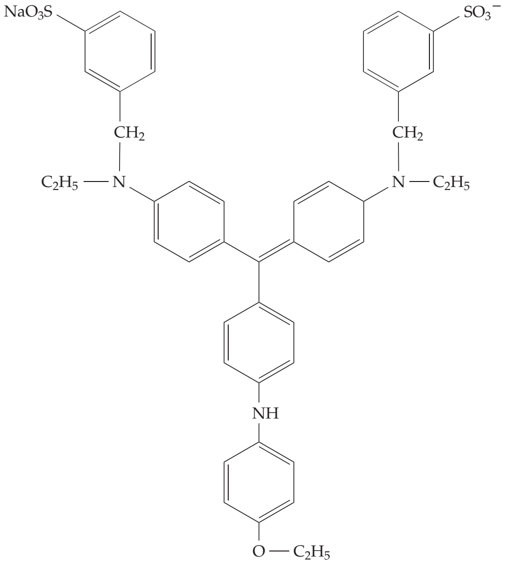
(3-31)

Reaction with naphthalene 2,3-dicarboxaldehyde (Eq. 3-32) increases the limit of detection 100-fold or more.[264]



(3-32)

Detection of proteins on thin-layer plates, gel slabs, or membranes is often accomplished by staining with a dye,[265–267] the most widely used being Coomassie brilliant blue.[268] Various silver-containing stains may also be used. After separation of a protein mixture by electrophoresis and transfer to an inert membrane,

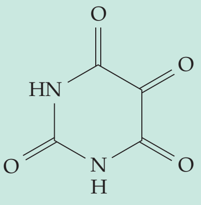


Coomassie brilliant blue

BOX 3-E NINHYDRIN

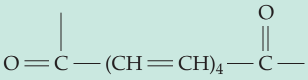
Ninhydrin (1,2,3-indantrione monohydrate) forms Schiff bases (ketimines) with amino acids. These react in ways similar to those of Schiff bases of pyridoxal phosphate (Chapter 14). Decarboxylation of the ketimines followed by hydrolysis of the resulting aldimines yields an intermediate amine that can couple with a second molecule of ninhydrin to form a characteristic purple color.[a,b] First reported by Ruhemann in 1910, the intermediate amine can also be hydrolyzed to free ammonia. Therefore, to ensure maximum color yield ninhydrin solutions for quantitative analysis usually contain reduced ninhydrin, which can react with free NH_3 and ninhydrin to form Ruhemann's purple (see scheme). The reaction has been widely used in chromatography and in quantitative amino acid analysis and also as a convenient spray reagent for paper and thin-layer chromatography. While α-amino acids react most readily, primary amines and peptides also form Ruhemann's purple. In these cases a proton rather than CO_2 is lost from the ketimine. When pyridoxamine (Chapter 14) on chromatograms reacts, a bright orange product, presumably the aldimine, appears. Secondary amines, such as proline, give a yellow color.

Both the ninhydrin reaction and pyridoxal phosphate-catalyzed decarboxylation of amino acids (Chapter 14) are examples of the **Strecker degradation**. Strecker reported in 1862 that alloxan causes the decarboxylation of alanine to acetaldehyde, CO_2, and ammonia.[c]



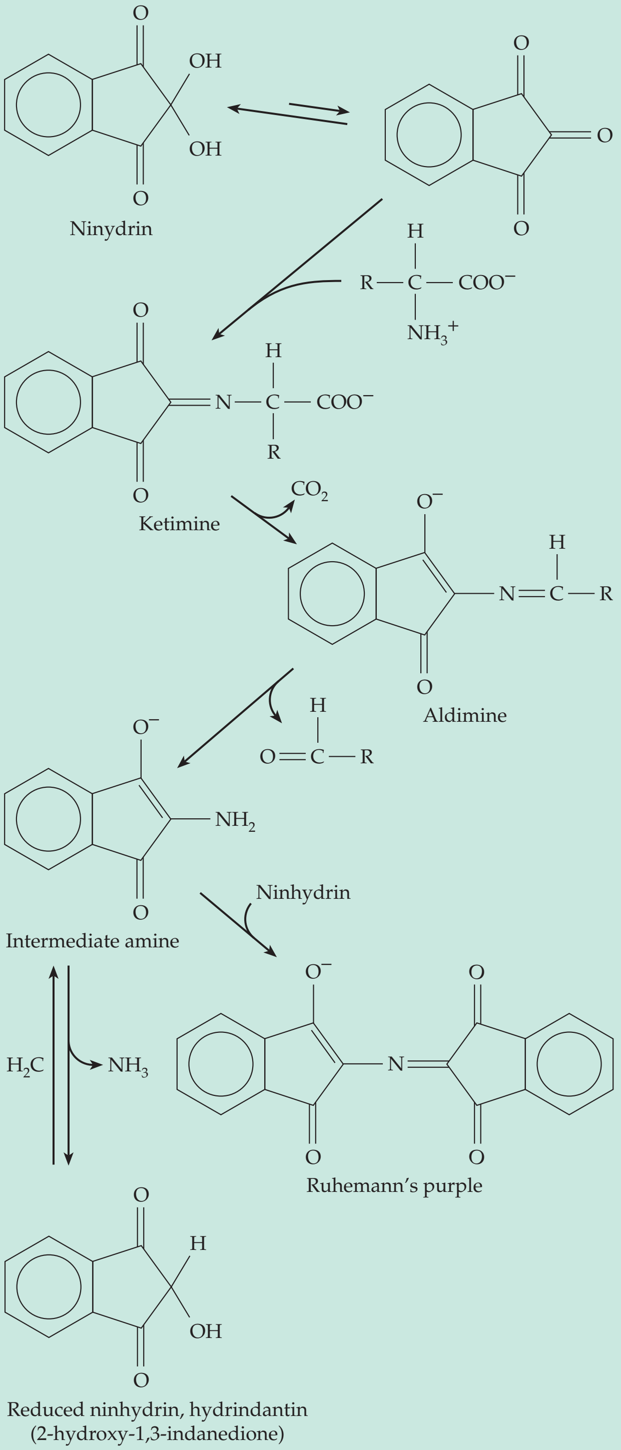
Alloxan

Many other carboxyl compounds, e.g., those of the general structure



$$O{=}\overset{|}{C}{-}(CH{=}CH)_4{-}\overset{O}{\overset{\|}{C}}{-}$$

and *p*-nitrosalicylaldehyde also cause the Strecker degradation.[d]



---

[a] Wigfield, D. C., and Croteau, S. M. (1980) *Biochem. Edu.* **8**, 26–27

[b] Friedman, M., and Williams, L. D. (1974) *Bioorg. Chem.* **3**, 267–280

[c] Strecker, A. (1862) *Annalen* **123**, 363–365

[d] Schonberg, A., and Moubacher, R. (1952) *Chem. Rev.* **50**, 261–277

the resulting protein “blots” can be stained with specific antibodies.[269,270] Flame ionization detectors can measure as little as a few picomoles of almost any substance leaving a vapor-phase chromatographic column. The importance of developing new, more sensitive analytical methods by which the quantity of material investigated can be scaled down can hardly be overemphasized. Increasingly sensitive methods of detection, including mass spectrometry, now permit measurement of fmol (10^{-15} mol) quantities in some cases. With this ability the output of neurotransmitters from a single neuron in the brain can be measured and the contents of single cells can be analyzed.

6. Absorption of Light

Side chains of the three aromatic amino acids phenylalanine, tyrosine, and tryptophan absorb ultraviolet light in the 240- to 300-nm region, while histidine and cystine absorb to a lesser extent. Figure 3-13 shows the absorption spectrum of a “reference compound” for tyrosine. There are three major absorption bands, the first one at 275 nm being a contributor to the well-known 280-nm absorption band of proteins. There is a much stronger absorption band at about 240 nm. Sensitive methods for estimating protein concentration depend upon the measurement of this absorption together with that from other side chains at around 280 or 230 nm.[13,271–274] There is an even stronger absorption band at 192 nm. However, at these wavelengths even air absorbs light and experimental difficulties are extreme. At 280 nm, and even more at 230 nm, it is easy to contaminate samples with traces of light-absorbing material invisible to the eye. Therefore, most estimations of protein concentration from light absorption depend upon the 280-nm band.

Figure 3-14 shows the spectra of *N*-acetyl ethyl esters of all three of the aromatic amino acids and of cystine. To a first approximation, the absorption spectra of proteins can be regarded as a summation of the spectra of the component amino acids. However, the absorption bands of some residues, particularly of tyrosine and tryptophan, are shifted to longer wavelengths than those of the reference compounds in water. This is presumably a result of being located within nonpolar regions of the protein. Notice that the spectra for tyrosine, phenylalanine, and cystine in Fig.

BOX 3-F BIOSENSORS AND ELECTRONIC NOSES

A new approach to detection of molecules of biological interest is the development of biosensors. These are small devices that detect the binding of specific molecules to a **receptor** which is in intimate contact with a specially prepared surface that serves as a **transducer**. The receptor might be a layer of enzyme, antibody, hormone receptor, lectin, or oligonucleotide. Binding of substrate, antigen, hormone, sugar, or complementary polynucleotide strand, respectively, induces a response consisting of some kind of electrical or optical signal.[a–d] If the sensor is constructed on a semiconductor chip changes in an imposed potential difference may be detected.[a,e] However, changes in optical properties are more often observed. Fluorescence of dyes incorporated into the tranducing layer may be induced by binding of a molecule to a protein that undergoes an allosteric modification (see Chapter 9).[f] Many biosensors measure **surface plasmon resonance**, a change in the evanescent wave that develops in a surface when a light beam at the angle of total reflectance strikes the surface. This induces a change in the dielectric constant which can be measured.[f–m] Biosensors are used to estimate binding constants and also rate constants. However, read the article by Schuck and Milton[m] for tests of the validity of kinetic data. Biosensors can serve as “electronic noses.” One possible application is in the analysis of compounds in human breath as an aid to medical diagnosis. Over 400 volatile organic compounds have been identified in breath using gas chromatography and mass spectrometry.[n]

---

[a] Briggs, J. (1987) *Nature (London)* **329**, 565–566

[b] Zurer, P. (1997) *Chem. Eng. News* **September 15**, 7

[c] Kress-Rogers, E., ed. (1997) *Handbook of Biosensors and Electronic Noses: Medicine, Food, and the Environment*, CRC Press, Boca Raton, Florida

[d] Cunningham, A. J. (1998) *Introduction to Bioanalytical Sensors*, Wiley, New York

[e] McConnell, H. M., Owicki, J. C., Parce, J. W., Miller, D. L., Baxter, G. T., Wada, H. G., and Pitchford, S. (1992) *Science* **257**, 1906–1912

[f] Marvin, J. S., Corcoran, E. E., Hattangadi, N. A., Zhang, J. V., Gere, S. A., and Hellinga, H. W. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4366–4371

[g] Raether, H. (1988) *Surface Plasmons, Springer Tracts in Modern Physics*, Vol. 111, Springer-Verlag, Berlin

[h] Peterlinz, K. A., Georgiadis, R. M., Herne, T. M., and Tarlov, M. J. (1997) *J. Am. Chem. Soc.* **119**, 3401–3402

[i] Hendrix, M., Priestley, E. S., Joyce, G. F., and Wong, C.-H. (1997) *J. Am. Chem. Soc.* **119**, 3641–3648

[j] Salamon, Z., Brown, M. F., and Tollin, G. (1999) *Trends Biochem. Sci.* **24**, 213–219

[k] Chao, H., Houston, M. E., Jr., Grothe, S., Kay, C. M., O’Connor-McCourt, M., Irvin, R. T., and Hodges, R. S. (1996) *Biochemistry* **35**, 12175–12185

[l] McNally, A. J., Mattsson, L., and Jordan, F. (1995) *J. Biol. Chem.* **270**, 19744–19751

[m] Schuck, P., and Minton, A. P. (1996) *Trends Biochem. Sci.* **21**, 458–460

[o] Phillips, M. (1992) *Sci. Am.* **267**(July), 74–79

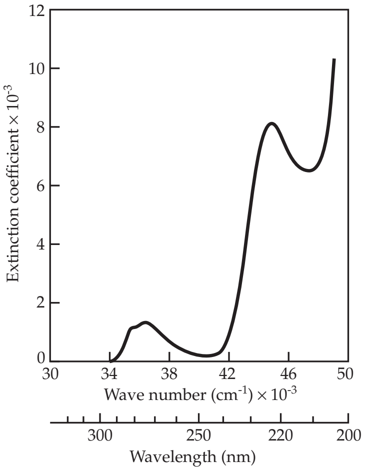


Figure 3-13 The absorption spectrum of *N*-acetyltyrosine ethyl ester in an aqueous phosphate buffer of pH 6.8. Absorbance (as molar extinction coefficient, Eq. 23-5) is plotted against increasing energy of light quanta in units of wave number. The more commonly used wavelength scale is also given. Spectra are most often presented with the low wavelength side to the left. In the convention adopted here the energy of a quantum increases to the right. There are three $\pi-\pi^*$ electronic transitions that give rise to absorption bands of increasing intensity. The third $\pi-\pi^*$ transition of the aromatic ring is at ~ 52,000 cm^{-1} (192 nm) and reaches a molar extinction coefficient of ~ 40,000. The $n-\pi^*$ and $\pi-\pi^*$ transitions of the amide group in this compound also contribute to the high energy end of the spectrum (see Chapter 23 for further discussion).

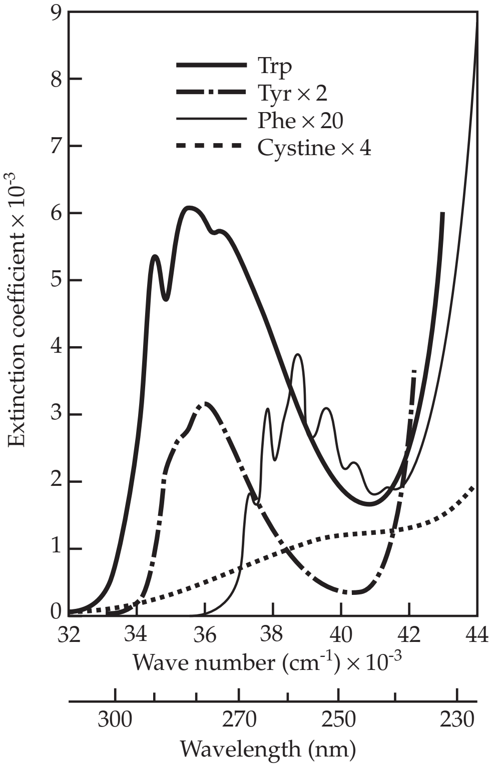


Figure 3-14 The spectra of the first electronic transitions of the *N*-acetyl derivatives of the ethyl esters of phenylalanine, tyrosine, and tryptophan together with that of the dimethyl ester of cystine in methanol at 25°C. The spectra for the Tyr, Phe, and cystine derivatives have been multiplied by the factors given on the graph.[272]

3-14 have been multiplied by factors of 2 to 20. It is evident that if all of the light-absorbing side chains were present in equal numbers tryptophan would dominate the absorption band and that phenylalanine would contribute little except some small wiggles. The molar extinction coefficient ε can be estimated from the numbers of residues of each type per molecule as follows:[274]

$$\varepsilon_{280}\ (M^{-1}cm^{-1}) = 5500\ (\text{no. Trp}) + 1490\ (\text{no. Tyr}) + 125\ (\text{no. cystine}) \qquad (3\text{-}33)$$

For proteins of unknown composition, a useful approximation is that a solution containing 1 mg / ml of protein has an absorbance at 280 nm of about 1.0.

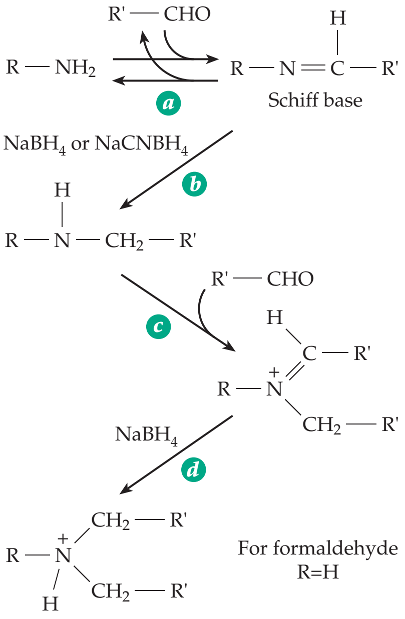
E. Quantitative Determinations and Modification Reactions of Side Chain Groups

The functional groups present in the side chains of proteins include $-NH_2$, $-SH$, $S-S$, $-OH$, $-COO^-$, the imidazole group of histidine, the guanidine group of arginine, the phenolic group of tyrosine, the indole ring of tryptophan, and the $-S-CH_3$ group of methionine. These are able to enter into a great variety of chemical reactions, most of which make use of the nucleophilic properties of these groups. The reactions are most often those of nucleophilic **addition** or nucleophilic **displacement**. The basic chemistry of these reactions often parallels biochemical reactions that are discussed in Chapters 12 and 13. In many instances, the reactions are nonspecific; amino, thiol, and hydroxyl groups may all react with the same reagent. The usefulness of the reactions depends to a large extent on the discovery of conditions under which there is some selectivity. It is also important that the reactions be complete. Only a few reactions will be considered here; these and others have been reviewed by Glazer *et al.*[225,275]

1. Reactions of Amino Groups

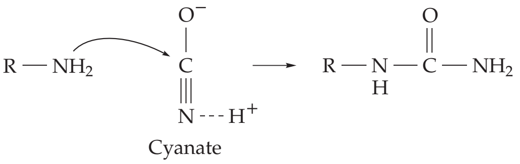
The numerous amino groups of lysine residues and of the N termini of peptide chains usually protrude into the aqueous surroundings of a protein. Chemical modification can be done in such a way as to preserve the net positive charge which amino groups carry at most pH values, to eliminate the positive charge leaving a neutral side chain, or to alter the charge to a negative value. Alterations of these charges can greatly affect interactions of the protein molecules with each other and with other substances.

Amino groups react reversibly with carbonyl compounds to form Schiff bases.[275–277] Reduction of the latter by sodium borohydride or sodium cyanoborohydride causes an irreversible change (Eq. 3-34, steps *a* and *b*). Cyanoborohydride is specific for Schiff bases and does not reduce the carbonyl compound. However, side products may cause problems.[277,278] Depending upon which carbonyl compound is used, the net positive charge on the amino group may be retained or it may be replaced with a different charge by this "reductive alkylation" sequence. Formaldehyde will react according to Eq. 3-34 in two steps to give a dimethyl amino group with no change of net charge.[279] Pyridoxal phosphate (Chapter 14) is converted by Eq. 3-34 into a fluorescent label. With a limited amount of pyridoxal phosphate only one or a few lysine residues may be labeled, often at active centers of enzymes. Schiff bases formed from glyceraldehyde in Eq. 3-34 can undergo the Amadori rearrangement (Eq. 4-8) to form stable products which, however, can be reconverted to the original amino groups upon acid hydrolysis. The borohydride reduction product of Eq. 3-34 with glyceraldehyde can be reconverted to the original amine by periodate oxidation (Eq. 4-11).[277]

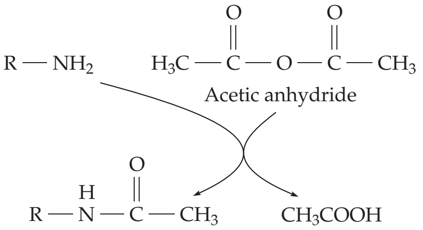


(3-34)

Another addition reaction of amino groups is **carbamoylation** with sodium cyanate (Eq. 3-35). A displacement reaction by an amino group on an acid anhydride such as acetic anhydride (Eq. 3-36) leads to **acylation**, a nonspecific reaction which is also undergone by thiol, hydroxyl, and other groups. When acetic anhydride is used, the net positive charge of an amino group is lost. However, the product obtained with succinic anhydride or maleic anhydride (Eq. 3-26) carries a negative charge. In the latter case, the modification can readily be reversed by altering the pH.

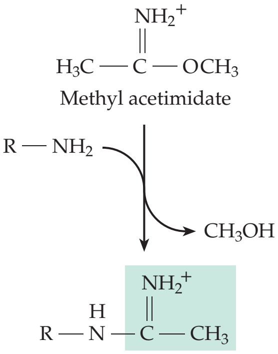


(3-35)

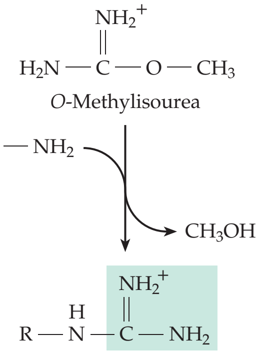


(3-36)

Both **amidination** (Eq. 3-37) and **guanidination** (Eq. 3-38) lead to retention of the positive charge.

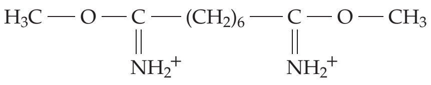


(3-37)



(3-38)

Bifunctional imidoesters such as **dimethylsuberimidate** may be used to establish whether or not two different proteins or subunits are close together in a complex or in a supramolecular structure such as a membrane or ribosome.



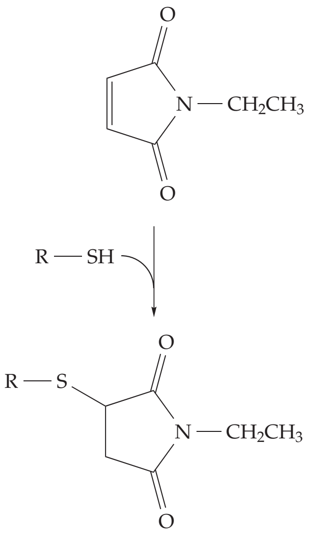
$$H_3C-O-\underset{NH_2^+}{\underset{\|}{C}}-(CH_2)_6-\underset{NH_2^+}{\underset{\|}{C}}-O-CH_3$$

Dimethylsuberimidate

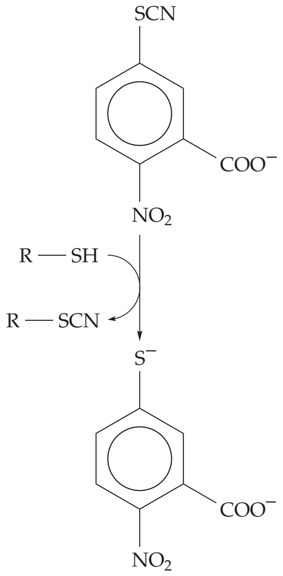
Another useful reaction of amino side chains is that with dansyl chloride (Eq. 3-29). Many lysine derivatives can be determined quantitatively by amino acid analysis.[280]

2. Reactions of SH Groups

In addition to the alkylation with iodoacetate (Eq. 3-24), sulfhydryl groups can react with *N*-ethylmaleimide (Eq. 3-39).[281] This reaction blocks the SH groups irreversibly and has often been used in attempts to establish whether or not a thiol group plays a role in the functioning of a protein. Loss of function in the presence of this “sulfhydryl reagent” may mean that an SH group has an essential role or it could be a result of the bulk of the group added. The *N*-ethylmaleimide group is large and could prevent proper contact between an enzyme and substrate or between two proteins. To avoid the possible effect of excessive bulk, it is useful to convert the SH to the small thiocyanate group – SCN (Eq. 3-40).[282]

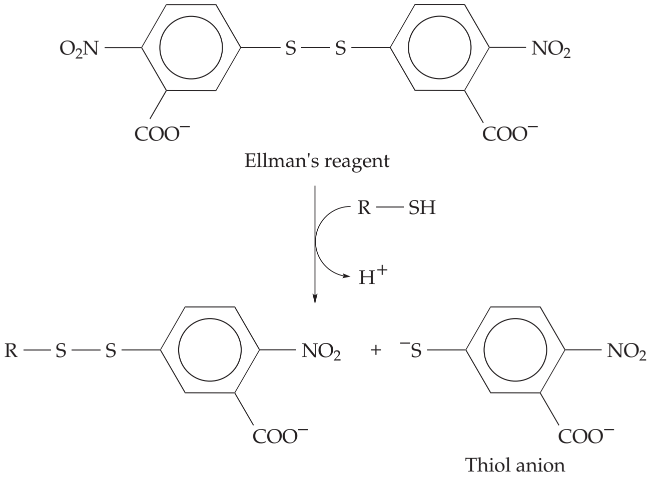


(3-39)



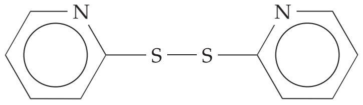
(3-40)

Ellman’s reagent 5,5'-dithiobis(2-nitrobenzoic acid; DTNB) reacts quantitatively with – SH groups (Eq. 3-41) to form mixed disulfides with release of a thiolate anion that absorbs light at 412 nm with a molar extinction coefficient of 14,150 M^{-1} cm^{-1}.[283] While DTNB has been widely used to determine the content of –SH groups in proteins, there are some disadvantages.

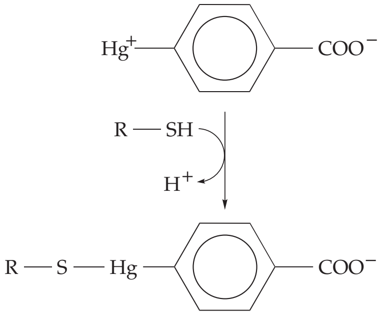


(3-41)

Pyridyldisulfides such as 2-pyridyldisulfide or the isomeric 4-pyridyldisulfide react more completely and with greater selectivity.[281]



Thiol groups have a high affinity for mercury ions including organic mercury derivatives, which are widely used in the determination of protein structures by X-ray crystallography (Section F). Titration of SH groups in proteins is often accomplished with *p*-mercuribenzoate (Eq. 3-42). The reaction may be followed spectrophotometrically at 250 – 255 nm, a region in which the mercaptide product absorbs strongly.

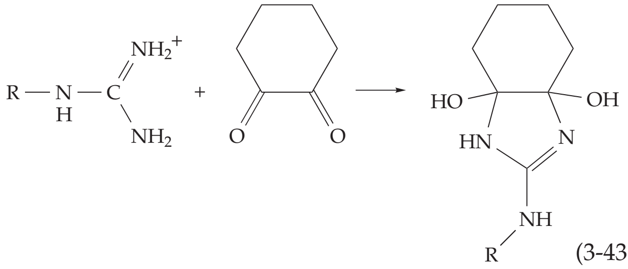


(3-42)

3. Reactions of Other Side Chains

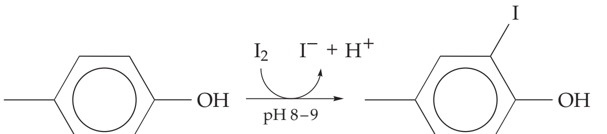
There are no highly selective reactions for –OH, –COO⁻, or imidazole groups. However, some hydroxyl groups in active sites of enzymes are unusually reactive in nucleophilic addition or displacement and can be modified by acylation, phosphorylation, or in other ways. Carboxyl groups, which are exceedingly numerous on protein surfaces, can be modified by treating with a water-soluble carbodiimide (Eq. 3-10) in the presence of a high concentration of an amine such as the ethyl ester of glycine. The imidazole groups of residues of histidine can often be selectively destroyed by dye-sensitized photooxidation (Ch. 12, Section D,5) or can be acylated with ethoxyformic anhydride.

Compounds with two adjacent carbonyl groups such as 1,2-cyclohexanedione (Eq. 3-43) react selectively with guanidinium groups from arginine residues in proteins. Under certain conditions the product indicated in Eq. 3-43 predominates. Related reagents are derived from camphorquinone.[284]



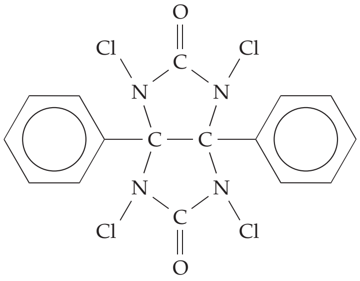
(3-43)

The phenolic group of tyrosine undergoes iodination (Eq. 3-44), acylation, coupling with diazonium compounds, and other reactions.



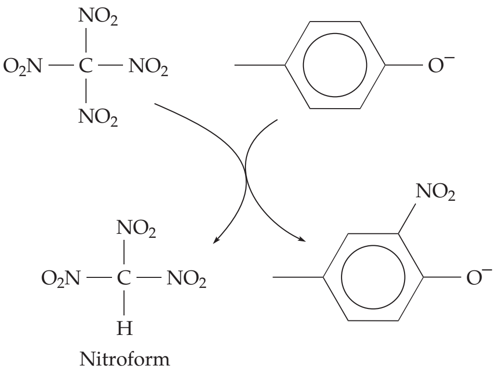
(3-44)

The following sparingly soluble chloroamide together with I⁻ will also iodinate tyrosine and can be used to incorporate radiolabeled iodine into proteins.[285,286]



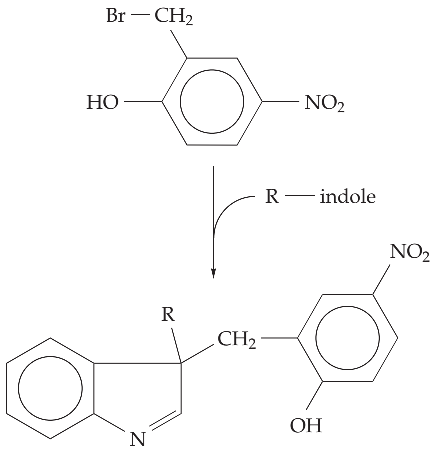
1,3,4,6-Tetrachloro-3α,6α-diphenylglycouril

Tetranitromethane reacts slowly with tyrosyl groups to form 3-nitrotyrosyl groups (Eq. 3-45). The by-product **nitroform** is intensely yellow with ε_{350} = 14,400. The reagent also oxidizes SH groups and reacts with other anionic groups.



(3-45)

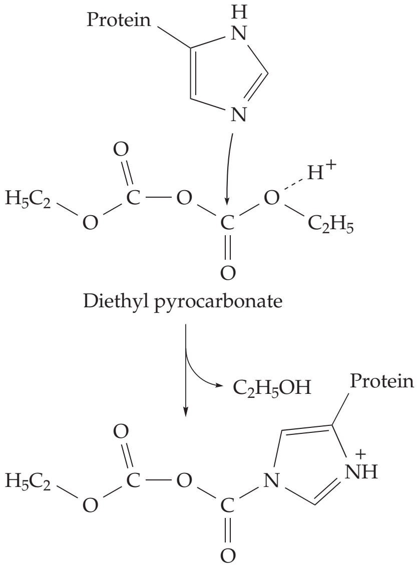
Koshland devised the following reagent for the indole rings of tryptophan residues (Eq. 3-46).



(3-46)

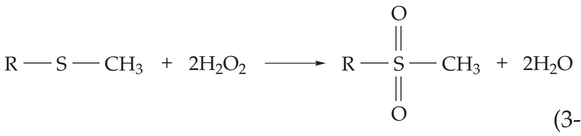
Imidazole, lysine amino groups, and tyrosine hydroxyl groups react with **diethylpyrocarbonate**

(Eq. 3-47) at low enough pH (below 6) that the reaction becomes quite selective for histidine.[70,287] Reactivity with this reagent is often used as an indication of histidine in a protein.[288–290]. The reaction may be monitored by observation of NMR resonances of imidazole rings.[288,290]



(3-47)

The thioether side chains of methionine units in proteins can be oxidized with hydrogen peroxide to the corresponding sulfones (Eq. 3-48). They can also be alkylated, e.g., by CH_3I to form $R{-}S^+(CH_3)_2$.



$$\mathrm{R{-}S{-}CH_3 + 2H_2O_2 \longrightarrow R{-}\overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle O}{\|}}{S}}{-}CH_3 + 2H_2O} \qquad (3\text{-}48)$$

4. Affinity Labeling

To identify groups that are part of or very near to the active site of a protein, reagents can be designed that carry a reactive chemical group into the active site.[291] The related **photoaffinity labeling**[292,293] is also widely used (see also Chapter 23).

F. Synthesis of Peptides

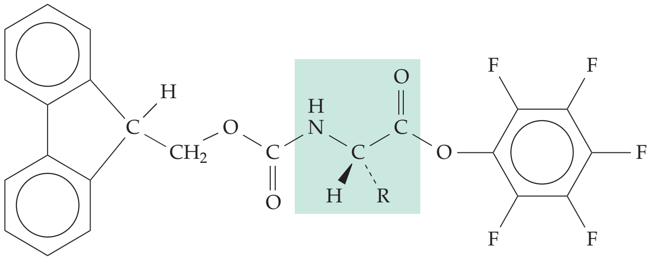
The synthesis of peptides of known sequence in the laboratory is extremely important to biochemical research. For example, we might want to know how the effects of a peptide hormone are altered by replacement of one amino acid in a particular position by another. The synthetic methods must be precise[294–298] and because there are so many steps the yield should be 98% or better for every step. Even so, it is still impractical to synthesize very large peptides. Those that have been made, such as the hormone insulin and the enzyme ribonuclease, have been obtained in low yields and have been difficult to purify.[298] It is usually more practical to obtain large peptides from natural sources. It is often practical to clone a suitable piece of DNA in a bacterial plasmid and to set up biological production of the desired peptide (Chapter 26). On the other hand, for smaller peptides, laboratory synthesis is feasible. Even for large peptides it is useful because it permits incorporation of unnatural amino acids as well as isotopic labels.

The general procedure for making a peptide in the laboratory is to “block” the amino group of what will become the N-terminal amino acid with a group that can be removed later. The subsequent amino acid units “activated” at their carboxyl end are then attached one by one. The chemical activation is often accomplished by conversion of the carboxyl group of the amino acid to an anhydride. At the end of the synthesis, the blocking group must be removed from the N terminus and also from various side chain groups such as those of cysteine and lysine residues. In many respects, this procedure is analogous to the biological synthesis of proteins whose basic chemistry is discussed in Chapter 29.

1. Solid-Phase Peptide Synthesis

Modern methods of peptide synthesis began with the solid-phase method introduced by Merrifield[299] in 1962 (Fig. 3-15). To begin the synthesis a suitably protected amino acid is covalently linked to a polystyrene bead. The blocking *t*-butoxycarbonyl (Boc) group is removed as isobutene by an elimination reaction to give a bound amino acid with a free amino group. This can then be coupled to a second amino acid with a blocked amino group using dicyclohexylcarbodiimide (Eq. 3-10). The removal of the blocking group and addition of a new amino acid residue can then be repeated as often as desired. The completed peptide is removed from the polystyrene by action of a strong acid such as HF.

Advantages of the Merrifield procedure are that the peptide is held tightly and can be washed thoroughly at each step. Problems arise from repeated use of trifluoroacetic acid and the need to use HF or other strong acid to cleave the peptide from the matrix and also to remove blocking benzyl groups that must be present on many side chain groups. Newer variations of the procedure include a more labile linkage to a polyamide type of polymer and use of blocked amino acids.[297,300–301a] These “active esters” will spontaneously condense with the free amino group of the growing peptide and with suitable catalysis will eliminate


H
C
CH_2
O
C
O
H
N
C
H
R
O
C
O
F
F
F
F
F

Blocked amino acid

pentafluorophenol. The fluorenylmethoxycarbonyl (Fmoc) blocking group is removed under mildly basic conditions. The whole procedure has been automated in commercially available equipment.

Smaller peptides may be joined to form longer ones.[298] Also useful is enzymatic synthesis. Protein-hydrolyzing enzymes under appropriate conditions will form peptide linkages, for example, joining together oligopeptides.[302,303] Other new methods have been devised to join unprotected peptides.[304,305]

Semisynthetic approaches can also be used to place unnatural amino acids into biologically synthesized proteins through the use of suppressor transfer RNAs (Chapter 29).[306]

2. Combinatorial Libraries

Many chemists devote all of their efforts to the synthesis of new compounds, including polypeptides, that might be useful as drugs. Traditionally, this has involved the tedious preparation of a large number of compounds of related structure which can be checked individually using various biochemical or biological tests. In recent years a new approach using "combinatorial chemistry" has become very popular and is continually being adapted for new purposes.[307–310] There are several approaches to creating a combinatorial library.

In "split synthesis" procedures a solid-phase

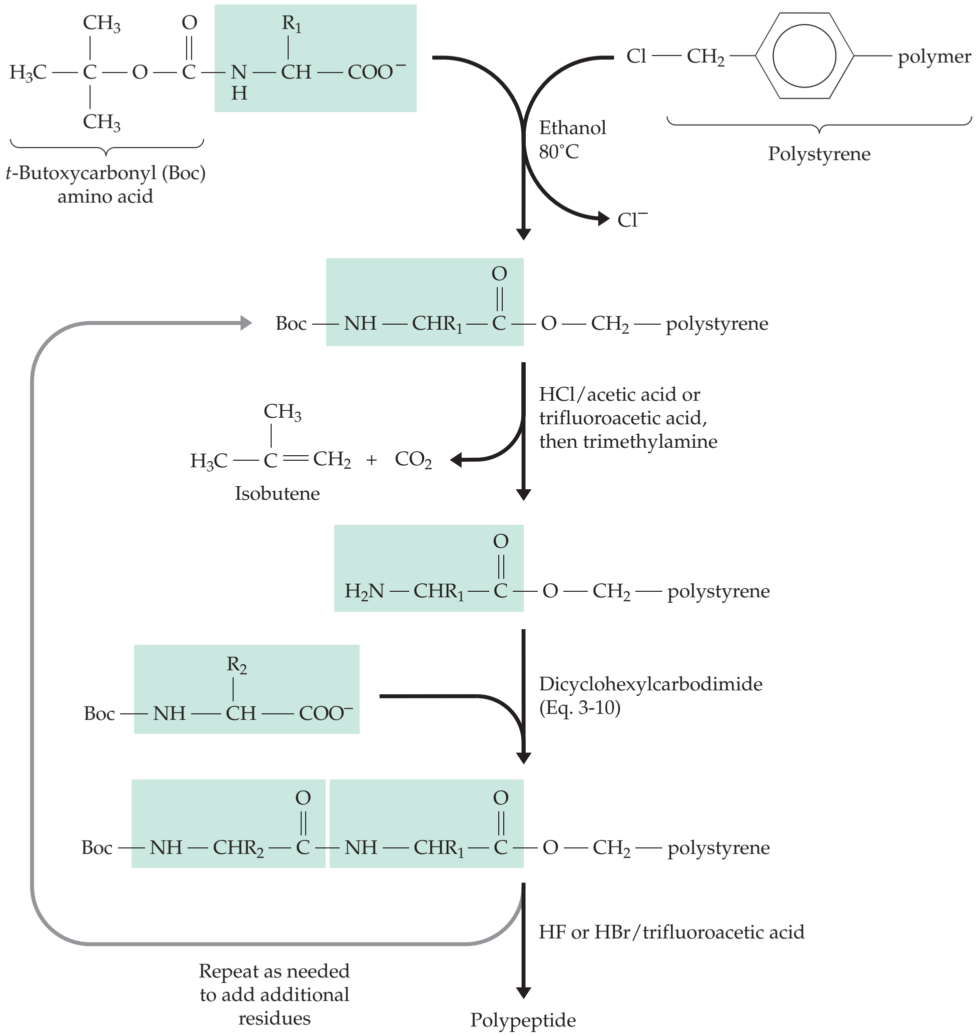

$H_3C-C(CH_3)_2-O-C(=O)-NH-CH(R_1)-COO^-$
t-Butoxycarbonyl (Boc) amino acid
Cl — CH_2 — polymer
Polystyrene
Ethanol
80°C
Cl^-
Boc — NH — CHR_1 — C(=O) — O — CH_2 — polystyrene
HCl/acetic acid or trifluoroacetic acid, then trimethylamine
$H_3C-C(CH_3)=CH_2$ + CO_2
Isobutene
H_2N — CHR_1 — C(=O) — O — CH_2 — polystyrene
Boc — NH — CH(R_2) — COO^-
Dicyclohexylcarbodimide (Eq. 3-10)
Boc — NH — CHR_2 — C(=O) — NH — CHR_1 — C(=O) — O — CH_2 — polystyrene
HF or HBr/trifluoroacetic acid
Repeat as needed to add additional residues
Polypeptide

Figure 3-15 Procedure for solid-phase peptide synthesis devised by Merrifield.[299]

synthesis is conducted on beads. For example, a family of peptides, each with the same C terminus, can be started on a large number of beads. After the first amino acid residue is attached the beads are divided into up to 20 equal portions and different amino acids are added to each portion. The beads can then be mixed and again subdivided. The third residue will again contain many different amino acids attached to each of the different amino acids in the second position. By repeating the procedure again, perhaps for many steps, a "library" of random peptides with each bead carrying a single compound will be formed.

To test whether a polypeptide or other compound carried on a given bead has a derived biological activity, such as the ability to inhibit a certain enzyme, various assays that require only one bead can be devised. However, if a particular bead carries a compound of interest, how can it be identified? The bead carries only a small amount of compound but it may be possible using microsequencing procedures to identify it. An alternative procedure is to use an encoding method to identify the beads.

An alternative to the "one bead–one peptide" approach is to incorporate random sequences of a DNA segment into a gene that can be used to "display" the corresponding peptide sequence. This is illustrated in Fig. 3-16. A protein segment (which may be a random sequence) can be displayed either on the major coat proteins along the shaft or on the minor coat proteins at the end of the bacterial virus fd (see Fig. 7-7). In the case of random insertions, each virus particle may display a different sequence (as many as 10^8). Peptides may be selected by binding to a desired receptor or monoclonal antibody and the DNA encapsulated in the virus particle can be used to produce more peptides for identification purposes.[311–316] Many other ingenious systems for constructing and testing libraries of peptides[317–320] and other molecules[319,321,322] are being devised. One of these involves a photolithography procedure for immobilizing macromolecules in a regular addressable array, e.g., in a 0.5-mm checkerboard pattern, on a flat surface.[323,324]

G. Microscopy

The light microscope[325] was developed around 1600 but serious studies of cell structure (histology) did not begin until the 1820s. By 1890, microscope lenses had reached a high state of perfection[326] but the attainable resolution was limited by the wavelength of light. For 450 nm blue light the limit is about 300 nm and for ultraviolet light, viewed indirectly, about 200 nm.[325,327] By the 1940s the electron microscope with its far superior resolving power had overshadowed the light microscope.

For both light and electron microscopy, the preparation of thin sections of cells is a very important technique. Only with very thin sections is the image sufficiently focused. However, **confocal scanning optical microscopy**, invented in the 1950s but not used commercially until much later,[328,329] provided an alternative solution to the focusing problem. A conical beam of light focused to a point is scanned across the sample and the transmitted light (or light emitted by fluorescence) passes through a small "pinhole" aperture located in the primary image plane to a photomultiplier tube where its intensity is recorded. The illuminating beam is moved to scan the entire field sequentially. A series of pinholes in a spinning disk may accomplish the same result. The focal plane can

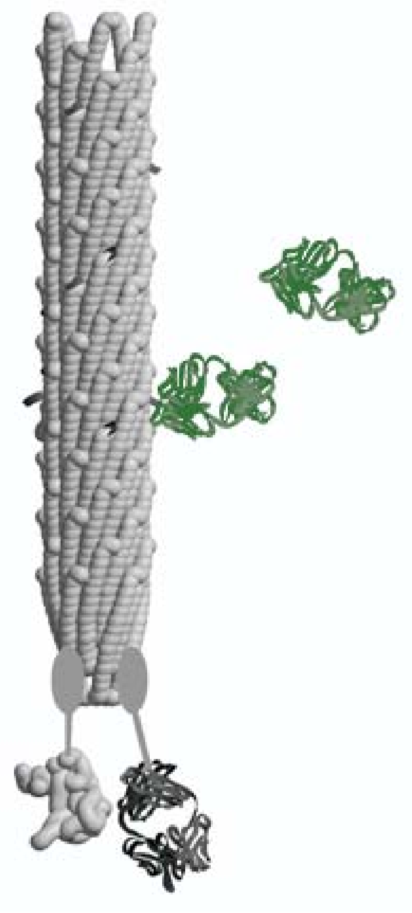


Figure 3-16 Model of bacteriophage fd engineered to display peptides as inserts in the coat proteins of the virus. The native virus structure is shown in gray; proteins not present in the native virus are shown black or green. Inserted near the N-termini of some major coat proteins is a 6-residue peptide. To one of these peptides a specific Fab antibody fragment (green) has bound from solution, and a second Fab is shown nearby. The N-terminal region of a minor coat protein at the end of the virion has been engineered to display a (different) Fab fragment. Steric constraints are less stringent for inserts in the minor proteins, but fewer copies per virion are possible. Reprinted with permission from Barbas, *et al.*[313a]

be varied so that an image of a thick object such as a cell can be optically sectioned into layers of less than 1 µm thickness. Stereoscopic pairs can also be generated (Fig. 3-17).[330]

A newer development in confocal microscopy is the use of two-photon and three-photon excitation of the fluorescent molecules that occur naturally within cells using short pulses of short-wavelength high-energy laser light.[331] Distribution of such compounds as NADH,[332] DNA, and the neurotransmitter serotonin[333] can be observed without damaging cells. Individual storage granules, each containing ~5 x 10^8 molecules of serotonin in a concentration of ~50 mM, can be seen.[334] Another new instrument, the **near-field scanning optical microscope** (NSOM),[335,336] is a lensless instrument in which the illuminating beam passes through a very small (e.g., 100 nm diameter) hole in a probe that is scanned in front of the sample. It may extend the limit of optical microscopy to ~1/50 the wavelength of the light.

Since it first became commercially available in 1939, the electron microscope has become one of the most important tools of cell biology.[337,338] The practical resolution is about 0.4 nm, but recent developments in scanning electron microscopy have resulted in resolution of 0.14 nm.[339] Of major importance was the development around 1950 of microtomes and knives capable of cutting thin (20–200 nm) sections of tissues embedded in plastic.[340] A bacterium such as *E. coli* can be sliced into as many as 10 thin longitudinal slices (see Fig. 1-4) and a eukaryotic cell of 10 µm diameter into 100 slices. Serial sections can be examined to determine three-dimensional structures. For some results see Bubel.[341]

If a slice of fresh (frozen) tissue is examined directly, little is seen because most of the atoms found in cells are of low atomic mass and scatter electrons weakly and uniformly. Therefore, thin sections must be "stained' with atoms of high atomic mass, e.g., by treatment with potassium permanganate or osmium tetroxide. Tissues must also be "fixed" to prevent disruption of cell structures during the process of removal of water and embedding in plastic. Fixatives such as formaldehyde react with amino groups and other groups of proteins and nucleic acids. Some proteins are precipitated in place and digestive enzymes that otherwise would destroy much of the fine structure of the cell are inactivated. Glutaraldehyde (a five-carbon dialdehyde) is widely used to fix and crosslink protein molecules in the tissue. The methods continue to be improved.[342]

Small particles, including macromolecules, may be "shadowed." Chromium or platinum can be evaporated in a vacuum from an angle onto the surface of the specimen, Individual DNA molecules can be "seen" in this way.[343] In fact, only the "shadows" are seen and they are 2–3 times wider than the DNA molecules. In the **negative contrast** method a thin layer of a solution containing the molecules to be examined, together with an electron-dense material such as 1% sodium phosphotungstate, is spread on a thin carbon support film. Upon drying, a uniform electron-dense layer is formed. Where the protein molecules lie, the phosphotungstate is excluded, giving an image of the protein molecule.

Surfaces of cells, slices, or intact bacteria can be coated with a deposit of platinum or carbon. The coating, when removed, provides a "negative" **replica** which can be examined in the microscope. Alternatively, a thin plastic replica can be made and can be shadowed to reveal topography. In "freeze fracturing" and "freeze etching," fresh tissue, which may contain glycerol to prevent formation of large ice crystals, is frozen rapidly. Such frozen cells can often be revived; hence, they may be regarded as still alive until the moment that they are sliced! The frozen tissue is placed in a vacuum chamber within which it is sliced or fractured with a cold knife. If desired, the sample can be kept in the vacuum chamber at about –100°C for a short time, during which some water molecules evaporate from the surface. The resultant etching reveals a fine structure of cell organelles and membranes in sharp relief. After etching, a suitable replica is made and examined (Fig. 1-15A and E). Fracturing tends to take place through lipid portions of cell membranes.

Small viruses, bacterial flagella, ribosomes, and even molecules can be seen by electron microscopy. However, to obtain a clear image in three dimensions requires a computer-based technique of **image reconstruction** or **electron microscope tomography**, which was developed initially by Aaron Klug and associates.[344–349] A sample is mounted on a goniometer, a device that allows an object to be tilted at exact angles. Electron

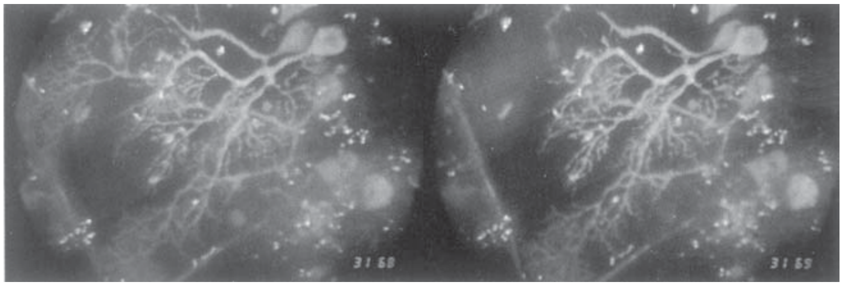


Figure 3-17 Confocal micrograph showing a forty-micrometer stereo slice in a 90-µm thick section of mouse cerebellum.[330] Courtesy of A. Boyde.

micrographs are prepared with the sample untilted and tilted in several directions at various angles, e.g., up to 90° in 10° increments. The micrographs are digitized and a computer is used to reconstruct a three-dimensional image.

In **electron crystallography**[350] micrographs of two-dimensional crystalline arrays of molecules or larger particles are prepared. A Fourier transform of the micrograph gives a diffraction pattern which can be treated in a manner similar to that usual for X-ray diffraction to give a three-dimensional image. An important milestone in use of this technique was the determination of the structure of bacteriorhodopsin (Fig. 23-45) at 0.3-nm resolution. Bear in mind that X-ray crystallography can also be viewed as a form of microscopy.

Invention of the **scanning tunneling microscope** (STM) by Binnig and Roher[351–353] initiated a new revolution in microscopy. The STM and similar scanned probe microscopes examine surfaces by moving a fine probe mounted in a piezoelectric x,y,z-scanner[354] across the surface to be examined. The tiny tungsten probe of the STM is so fine that its tip may consist of a single atom. When a small voltage is applied a minuscule quantum mechanical tunnelling current flows across the small gap between the probe and the surface and a high-resolution image, sometimes at atomic resolution, is created from the recorded variation in current.[354-359] The STM theoretically responds only to surfaces that conduct electrons, but nonconducting samples have been imaged at high humidity; presumably by conductance of electrons or ions through the surface water layer. The success of the STM spurred the development of many other types of scanned probe microscopes. Among these, the **atomic force microscope** (AFM; Fig. 3-18) has been especially useful for biological materials, including proteins and nucleic acids. The AFM moves a fine-tipped stylus directly across the sample surface or, alternatively, vibrates the probe above the surface. The small up-and-down movements of the stylus are recorded[359–366] and thereby create a topographic or force-field map of the sample. AFM images contain three-dimensional information and can be used to view individual molecules (Fig. 3-19).[367,367a] **Chemical force microscopy** is sensitive to adhesion and friction as a function of the interaction between defined chemical groups on the tip and sample.[368]

An emerging field is force spectroscopy, in which the AFM measures interaction forces between and within individual molecules.[369–371a] Under development are NMR microscopes (Section I). There is continual effort to see small objects more directly and more clearly!

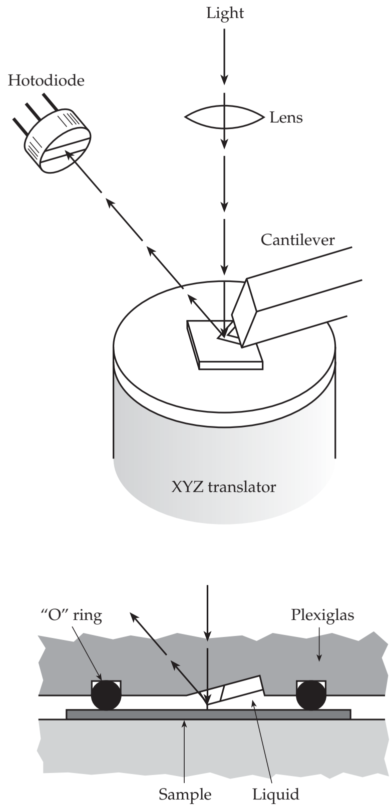


Figure 3-18 Schematic diagram of the atomic force microscope.[360] Courtesy of Paul Hansma.

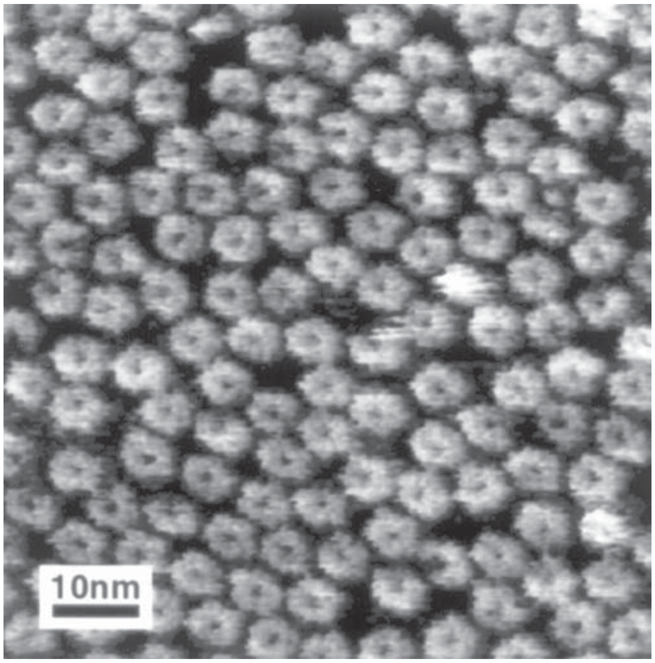


Figure 3-19 AMF images of cholera toxin[364] (see also Box 11-A). Courtesy of Z. Shao.

H. X-ray and Neutron Diffraction

One of the most important techniques by which we have learned bond lengths and angles and precise structures of small molecules is X-ray diffraction. Today, this technique, which involves measurements of the scattering of X-rays by crystalline arrays of molecules, is being used with spectacular success to study macromolecules of biochemical and medical importance.[372–378]

X-rays were described by Röntgen in 1896 but there was uncertainty as to their wave nature.[274] It was not until 1912 that the wavelengths of X-rays had been measured and it was recognized that they were appropriate for the use of X-rays in structure determination. Consider the fact that with a conventional light microscope we cannot distinguish two small objects that are much closer together than the distance represented by the wavelength of the light with which we observe them. This is about 460 nm for blue light. By comparison the wavelength of the copper K_α radiation, which is used in protein crystallography, is 0.1542 nm entirely appropriate for seeing the individual atoms of which matter is composed. Recognizing this, W. L. Bragg in Cambridge, England, in 1913 used X-ray diffraction to establish the structures of NaCl, KCl, and KBr in the crystalline state. The science of X-ray crystallography had been founded.[379]

In 1926, James Summer crystallized the enzyme urease (Chapter 16) and crystallization of other enzymes soon followed.[380] In 1934, J. B. Bernal brought back to Cambridge from Uppsala, Sweden, some crystals of pepsin almost 2 mm long that had been grown in T. Svedberg’s laboratory. Bernal and Crowfoot showed that these delicate crystals, which contained almost 50% water, gave a sharp diffraction pattern when they were protected by enclosure in a narrow capillary tube containing some of the mother liquor from which the crystals had been grown.[381,382] After this, diffraction patterns were obtained for many protein crystals, and in 1937 Max Perutz chose for his thesis work at Cambridge the X-ray crystallography of hemoglobin. The project seemed hopeless at times, but in 1968, 31 years later, Perutz had determined the structure of hemoglobin.[383–386]

If X-rays could be focused easily, could one build an X-ray microscope that would permit the immediate viewing of molecular structures? X-ray holograms at the molecular level have been obtained.[387,388] However, currently the only practical X-ray microscope for protein structures involves the measurement of diffraction patterns created by the scattering of X-rays (or of neutrons) from the crystalline lattice. The details of this procedure can be found in other sources.[372–375,378,389–390] It is sufficient to point out here that a pattern of many

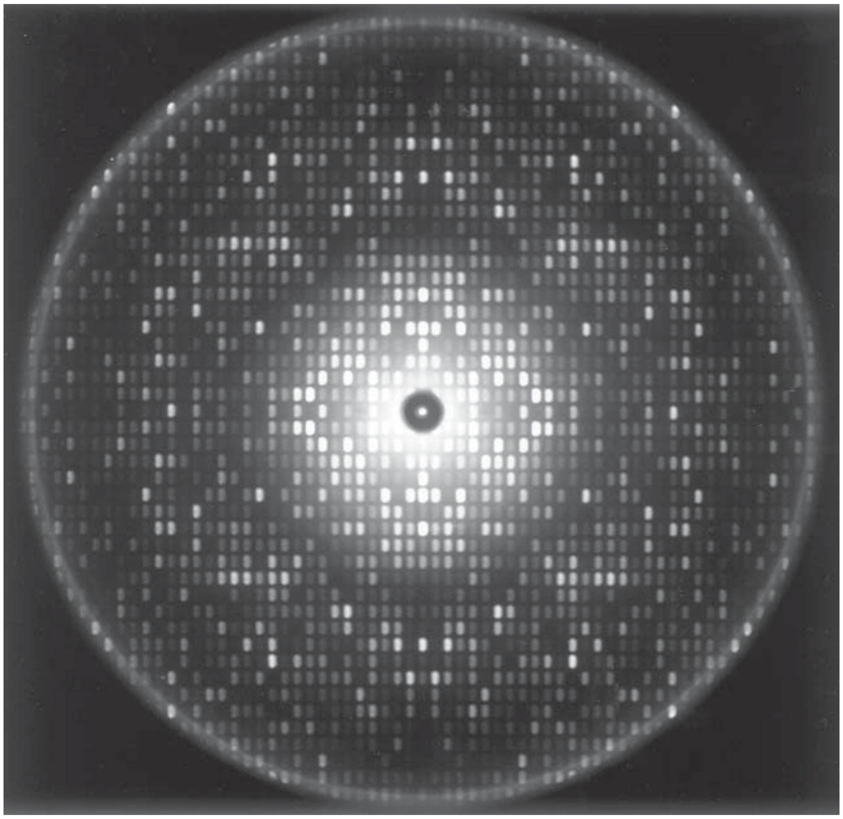


Figure 3-20 An X-ray diffraction photograph such as was used to determine the structure of hemoglobin. This precession photograph was obtained from a crystal of human deoxyhemoglobin by rotating the crystal along two different axes in a defined manner before a narrow X-ray beam. The film also was moved synchronously. The periodicity observed in the photograph is a result of a diffraction phenomenon arising from the periodic arrangement of atoms in the crystal. The distances of the spots from the origin (center) are inversely related to the distances between planes of atoms in the crystal. In this photograph (which shows only two dimensions of the three-dimensional diffraction pattern) the spots at the periphery represent a spacing of 0.28 nm. If the intensities of the spots are measured, and if the phases of the harmonic functions required for the Fourier synthesis can be assigned correctly, the structure can be deduced to a resolution of 0.28 nm from a set of patterns of this type. In the case of human deoxyhemoglobin, a complete set of data would consist of about 16,000 spots. With modern equipment the resolution limit can be extended to better than 0.15 nm (or more than 100,000 unique spots). Photograph courtesy of Arthur Arnone.

spots, such as that in Fig. 3-20, is obtained. In a typical determination of a protein structure from 10,000 to several times that number of spots must be measured. The needed information is contained in the coordinates of the spots (i.e., in the angles through which the scattering occurs) and in the intensities of the spots. The structure is obtained by a Fourier synthesis in which a large number of mathematical functions in the form of three-dimensional sinusoidal waves are summed. The wavelengths of these functions are determined by the positions of the spots in the diffraction pattern. The amplitudes of the waves are related to the intensities of the spots.[374] However, the waves are not all in phase and the phases must be learned in some other way. This presents a difficult problem. Mathematical methods have been devised that automatically determine the phases for small molecules and usually allow the structures to be established quickly. For proteins it is more difficult.

In 1954, Perutz introduced the **isomorphous replacement method** for determining phases. In this procedure a heavy metal, such as mercury or platinum, is introduced at one or more locations in the protein molecule. A favorite procedure is to use mercury derivatives that combine with SH groups. The resulting heavy metal-containing crystals must be isomorphous with the native, i.e., the molecules must be packed the same and the dimensions of the crystal lattice must be the same. However, the presence of the heavy metal alters the intensities of the spots in the diffraction pattern and from these changes in intensity the phases can be determined. Besides the solution to the phase problem, another development that was absolutely essential was the construction of large and fast computers. It would have been impossible for Perutz to determine the structure of hemoglobin in 1937, even if he had already known how to use heavy metals to determine phases.

The first protein structure to be learned was that of myoglobin, which was established by Kendrew *et al*. in 1960.[391–393] That of the enzyme lysozyme was deduced by Blake *et al*. in 1965.[394] Since then, new structures have appeared at an accelerating rate so that today we know the detailed architecture of over 6000 different proteins[395] with about 300 distinctly different folding patterns.[396] New structures are being determined at the rate of about one per day. X-ray diffraction has also been very important to the study of naturally or artifically oriented fibrous proteins[397] and provided the first experimental indications of the β structure of proteins.

Suppose that you have isolated a new protein. How can you learn its three-dimensional structure? The first step is crystallization of the protein, something that a biochemist may be able to do. Crystallization is done in many ways, often by the slow diffusion of one solution into another or by the slow removal of solvent through controlled evaporation.[374,398] Ammonium sulfate and polyethylene glycol are two commonly used precipitants. The presence of the neutral detergent β-octyl glucoside improves some crystals.[399] Droplets of protein solution mixed with the precipitant are often suspended on microscope cover glasses in small transparent wells or are placed in depression plates within closed plastic boxes.[374,400,401] In either case, a reservoir of a solution with a higher concentration of precipitant is present in the same compartment. Water evaporates from the samples into the larger reservoir, concentrating the protein and causing its crystallization. Crystals grow slightly better in a spacecraft than on Earth.[402,403] Some proteins, notably myosin from muscle, crystallize well only after reductive methylation of all lysine side chains to dimethyllysine with formaldehyde and sodium borohydride (Eq. 3-34).[278]

The next step is for a protein crystallographer to mount a small perfect crystal in a closed silica capillary tube and to use an X-ray camera to record diffraction patterns such as that in Fig. 3-20. These patterns indicate how perfectly the crystal is formed and how well it diffracts X-rays. The patterns are also used to calculate the dimensions of the unit cell and to assign the crystal to one of the seven **crystal systems** and one of the 65 enantiomorphic **space groups**. This provides important information about the relationship of one molecule to another within the unit cell of the crystal. The unit cell (Fig. 3-21) is a parallelopiped

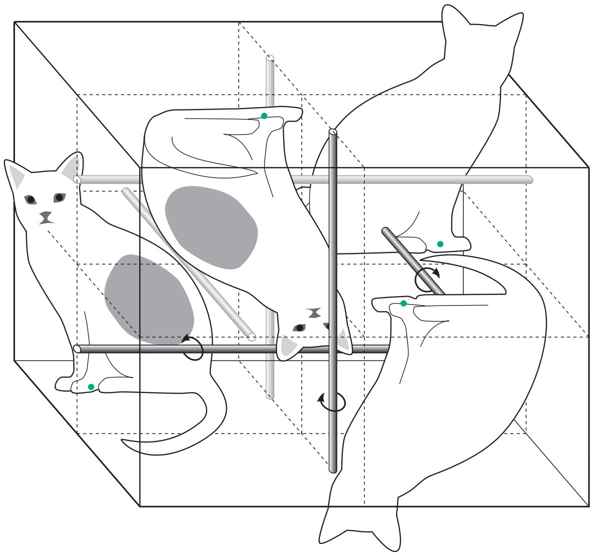


Figure 3-21 Diagram showing an asymmetric unit as it might appear in a unit cell of space group $P2_12_12_1$. This unit cell has three pairs of nonintersecting twofold **screw axes** which are marked by the shaded rods. These are designated by arrows and the symbol. Two asymmetric units are related one to another by rotation around a twofold axis together with translation by one-half the dimension of the unit cell.

whose sides are parallel to crystallographic axes and which, by translation in three dimensions, gives rise to the entire crystal. The unit cell must contain at least one **asymmetric unit**, the smallest unit of structure that lacks any element of crystal symmetry. The asymmetric unit may consist of one or a small number of protein subunits.

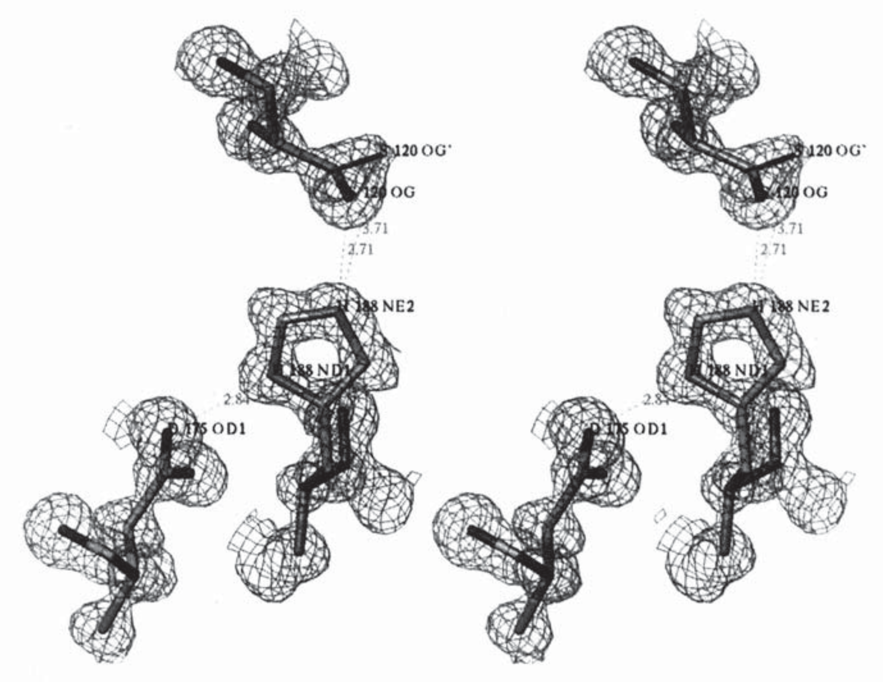


Figure 3-22 Stereoscopic view of a section of the structure of cutinase from the fungus *Fusarium solani* determined to a resolution of 0.10 nm. The three amino acid residues shown are serine 120 (top), histidine 188, and aspartate 175 (lower left). The structure is presented as a contour map with a "wire mesh" drawn at a "cutoff" level of density equal to 1 σ above the average, where σ is the root mean square density of the entire map. The side chains of these three residues constitute the "catalytic triad" in the active site of this enzyme (see Chapter 12). At this resolution more than one conformation of a group may often be seen. For example, the gamma oxygen (OG) of S120 is seen in two positions, the major one being toward His 188. When the map is drawn with a lower contour level the N-H proton on His 188 that is hydrogen bonded to Asp 175 can also be seen.[410] Courtesy of Christian Cambillau.

In crystals of the triclinic system all of the asymmetric units are aligned in the same manner and there are no axes of symmetry. Monoclinic crystals have a single **twofold** or **dyad axis** (see Fig. 7-11). The unit cell might contain two molecules related one to another by the dyad or one dimeric molecule with the dyad located as in Fig. 7-11C. Orthorhombic crystals have three mutually perpendicular twofold axes and the unit cell is a rectangular solid. Trigonal, tetragonal, and hexagonal crystal systems have three, four and sixfold axes of symmetry, respectively, while the cubic crystal contains four threefold axes along with diagonals of the cube as well as two-fold axes passing through the faces (see Fig. 15-14). Within each crystal system there are several space groups. An example is the orthorhombic space group $P2_12_12_1$, which is often met with small organic molecules and proteins. In this space group the unit cell contains three mutually perpendicular but nonintersecting twofold screw axes (Fig. 3-21). The position of one molecule in the unit cell is related to the next by both a 180° degree rotation about the twofold axes and a translation of one-half the length of the unit cell. There is such a screw axis for each of the three directions.[372]

The third step in the structure determination is collection of the X-ray diffraction data. This may be done with a **diffractometer** in which a narrow collimated pencil source of X-rays is aimed at the crystal and the intensities and positions of the diffracted beams are measured automatically. The computer-controlled diffractometer is able to measure the angles to within less than one-hundredth of a degree. If sufficient time is allowed, very weak spots can be counted. Today, diffractometers are more likely to be used for preliminary measurements, while the major data collection is done with an **area detector**, an instrument that collects many reflections simultaneously. There are some difficulties. For example, during the long periods of irradiation needed to measure the weak spots with a diffractometer the protein crystals decompose and must be replaced frequently. Data collection may last for months. The newest methods utilize more powerful X-ray sources and often **synchrotron radiation**, which delivers very short and extremely intense pulses of X-rays and allows data to be collected on very small well-formed crystals.[404]

The fourth step is the preparation of isomorphous crystals of heavy metal-containing derivatives. The heavy metal may be allowed to react with the protein before crystallization or may be diffused into preformed crystals. A variety of both cationic and anionic metal complexes, even large $Ta_6Br_{12}^{2+}$ tantalum clusters, have been used.[405] Two or more different heavy metal derivatives are often required for calculation of the phases. The heavy metal atoms must be present at only a very small number of locations in the unit cell.

An entire data set must be collected for each of these derivatives. The evaluation of the phases from these data is a complex mathematical process which usually involves the calculation first of a “difference Patterson projection.”[406] This is derived by Fourier transformation of the differences between the scattering intensities from the native and heavy atom-containing crystals. The Patterson map is used to locate the coordinates of the heavy metal atoms which are then refined and used to compute the phases for the native protein.

Alternative methods of solving the phase problem are also used now. When a transition metal such as Fe, Co, or Ni is present in the protein, anomolous scattering of X-rays at several wavelengths (from synchrotron radiation) can be used to obtain phases. Many protein structures have been obtained using this multiple wavelength anomalous diffraction (MAD phasing) method.[404,407,408] Selenocysteine is often incorporated into a protein that may be produced in bacteria using recombinant DNA procedures. Crystals are prepared both with protein enriched in Se and without enrichment. In some ways it is better to incorporate tellurium (^{127}Te) in telluromethionine.[409]

In the fifth step of an X-ray structure determination the **electron density map** is calculated using the intensities and phase information. This map can be thought of as a true three-dimensional image of the molecule revealed by the X-ray microscope. It is usually displayed as a stereoscopic view on a computer graphics system (Fig. 3-22). It is also often prepared in the form of a series of transparencies mounted on plastic sheets. Each sheet represents a layer, perhaps 0.1 nm thick, with contour lines representing different levels of electron density.

Using the electron density map a three-dimensional model of the protein can be built. For years, the customary procedure was to construct a model at a scale of 2 cm = 0.1 nm using an **optical comparator**, but

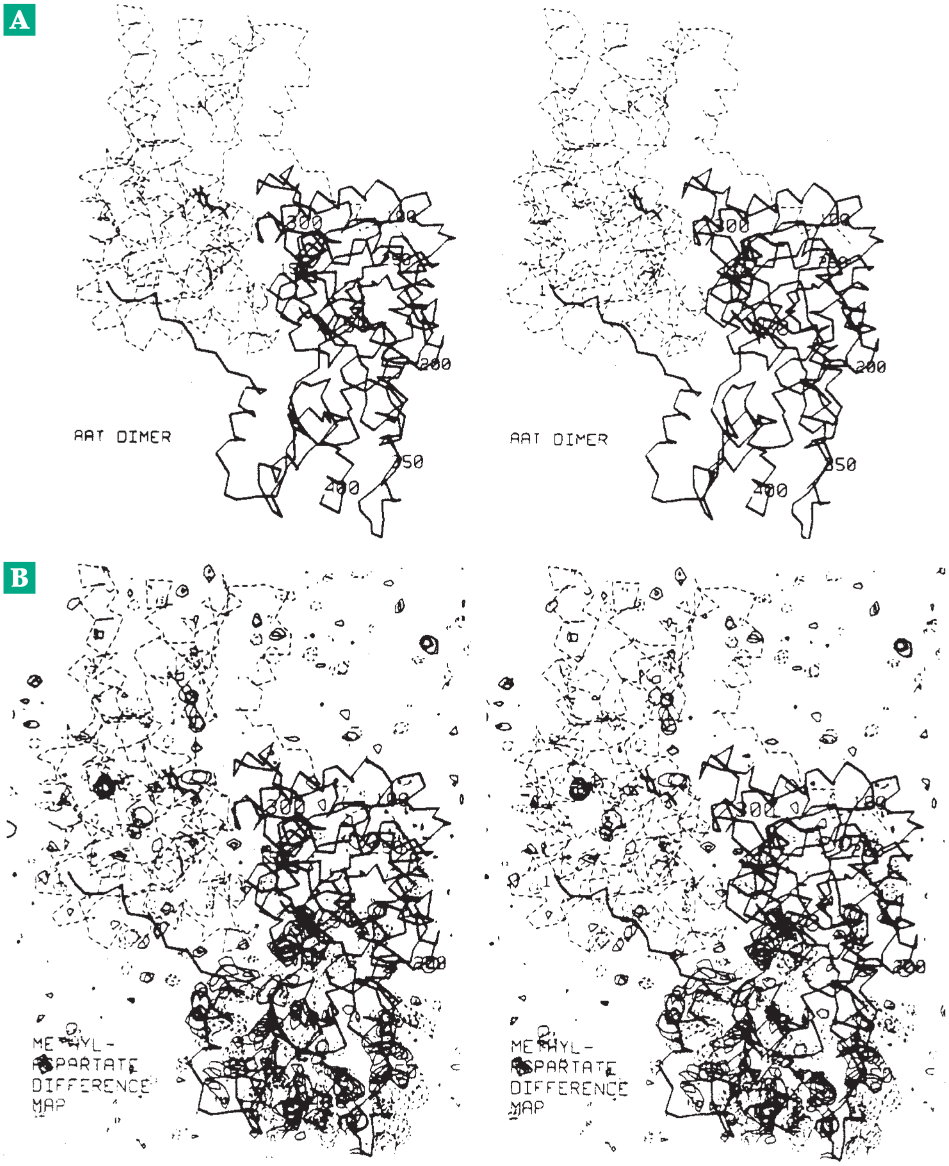


Figure 3-23 (A) Stereoscopic α-carbon plot of the cystolic aspartate aminotransferase dimer viewed down its dyad symmetry axis. Bold lines are used for one subunit (subunit 1) and dashed lines for subunit 2. The coenzyme pyridoxal 5'-phosphate (Fig. 3-24) is seen most clearly in subunit 2 (center left). (B) Thirteen sections, spaced 0.1 nm apart, of the 2-methylaspartate difference electron density map superimposed on the α-carbon plot shown in (A). The map is contoured in increments of ± 2σ (the zero level omitted), where σ = root mean square density of the entire difference map. Positive difference density is shown as solid contours and negative difference density as dashed contours. The alternating series of negative and positive difference density features in the small domain of subunit 1 (lower right) show that the binding of L-2-methylaspartate between the two domains of this subunit induces a right-to-left movement of the small domain. (Continues)

crystallographers now use computer graphics systems. The three-dimensional image of the electron density map and a computer-generated atomic model are superimposed on the computer screen. An example is shown in Fig. 3-22. When the superposition has been completed, the coordinates of all atoms are present in the memory of the computer.

The final step in the structure determination is **refinement** using various mathematical methods. From the coordinates of the atoms in the model the expected diffraction pattern is computed and is compared with that actually observed. The differences between predicted and observed density are squared and summed. The sum of the squares constitutes an error function which is then minimized by moving the various atoms in the model short distances while keeping bond lengths, angles, and van der Waals distances within acceptable limits and recalculating the error function. This complex "refinement" procedure must be repeated literally hundreds of thousands of times with every part of the structure being varied. Structures at very high resolution (0.07–0.1 nm) may reveal multiple positions for hydrogen-bonded side chains[410a,b] as well as hydrogen atoms (Fig. 3-22) and even bonding electrons.[410b]

Once the three-dimensional structure of the protein is known, further experiments are usually done using the X-ray diffraction technique. Since protein crystals contain channels of solvent between the packed molecules, it is usually possible to diffuse small molecules into the crystal. These may be substrates, inhibitors, or allosteric effectors. Diffraction data are collected after diffusion of the small molecules into the crystal and a **difference electron density map** may be calculated. This may show exactly where those molecules were bound. An example is shown in Fig. 3-23. Sometimes difference maps not only show the binding to an enzyme of a substrate or other small molecule but also reveal conformational changes in proteins. Such is the case for Fig. 3-23, which shows the binding of α-methylaspartate, an inhibitor that behaves initially like a substrate and goes part way through the reaction sequence for aspartate aminotransferase until further reaction is blocked by the methyl group. This difference map also shows that the part of the protein to the left of the binding site in the figure has moved.[411] Using the X-ray data it was possible subsequently to deduce the nature of the conformational change.

Examination of the effect of temperature on the diffraction pattern of a protein can give direct information about the mobility of different parts of the mole-

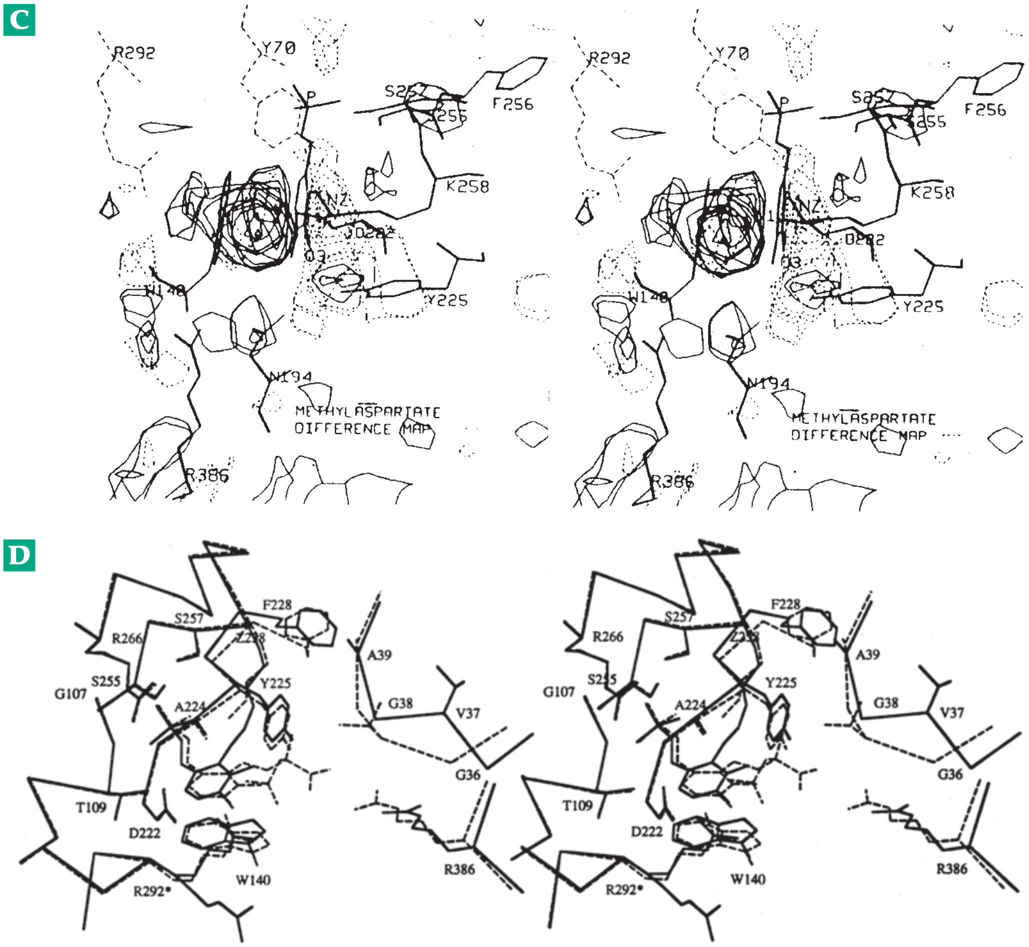
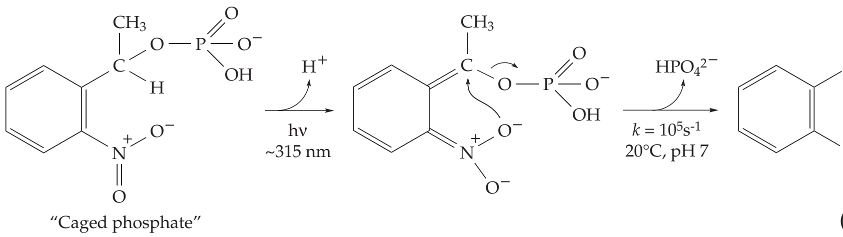


Figure 3-23 (Cont.) (C) Nine sections, spaced 0.1 nm apart, of a part of the 2-methylaspartate difference map superimposed on the atomic model shown in (A) and (D). The coenzyme is shown as the internal aldimine with Lys 258 (see Fig. 14-6, 14-10). The positive and negative contours on the two sides of the coenzyme ring indicate that the coenzyme tilts over to form the external aldimine when substrates react.[413] (D) Superimposed structure of the active site of the enzyme in its free form as in (A) (bold lines) and the refined structure of the α-methylaspartate complex, (dashed lines).[411] This illustrates the tilting of the coenzyme ring, which is also shown in Eq. 14-39 and Fig. 14-10. Courtesy of Arthur Arnone and Sangkee Rhee.

cules.[412] This complements information obtainable in solution from NMR spectra.

Because hydrogen atoms contain only one electron, and therefore scatter X-rays very weakly, they are usually not seen at all in X-ray structures of proteins. However, neutrons are scattered strongly by hydrogen atoms and **neutron diffraction** is a useful tool in protein structure determination.[414,415] It has been used to locate tightly bonded protons that do not exchange with ^2H_2O as well as bound water (^2H_2O).

The development of synchrotron radiation as an X-ray source[404,416–418] has permitted accumulation of data for electron density difference maps in less than 1 s and it is expected that such data can eventually be acquired in ~1 ps.[419–421] If a suitable photochemical reaction can be initiated by a picosecond laser flash, a substrate within a crystalline enzyme can be watched as it goes through its catalytic cycle. An example is the release of inorganic phosphate ions from a "caged phosphate" (Eq. 3-49) and study of the reaction of the released phosphate with glycogen phosphorylase (Chapter 12).[422,423]



(3-49)

However, caged substrates usually must diffuse some distance before reacting, so very rapid events cannot be studied. An alternative approach is to diffuse substrates into crystals at a low temperature at which reaction is extremely slow but a substrate may become seated in an active site ready to react. In favorable cases such "frozen" Michaelis complexes may be heated by a short laser pulse to a temperature at which the reaction is faster and the steps in the reaction may be observed by X-ray diffraction.[424,425]

I. Nuclear Magnetic Resonance (NMR)

Organic chemists and biochemists alike have long relied on NMR spectroscopy to assist in identification and determination of structures of small compounds. Most students have some familiarity with this technique and practical information is available in many places.[426–428] Measurements can be done on solids, liquids, or gases but are most often done on solutions held in special narrow NMR tubes. Volumes of samples are typically 0.5 ml (for a 1–5-mM solution containing 1–5 µmol of protein) but less for small molecules and higher concentrations. Newer techniques allow use of a volume as small as 5 nl and containing <0.1 nmol of sample.[429] For many years progress in biochemical application of NMR was slow, but a dramatic increase in the power of the spectrometers, driven in great measure by the revolution in computer technology, has made NMR spectroscopy a major force in the determination of structures and functions of proteins and nucleic acids.[430,431]

1. Basic Principles of NMR Spectroscopy

The basis of NMR spectroscopy lies in the absorption of electromagnetic radiation at radiofrequencies by atomic nuclei.[426,427,432–437] All nuclei with odd mass numbers (e.g., ^{1}H, ^{13}C, ^{15}N, ^{17}O, ^{19}F, and ^{31}P), as well as those with an even mass number but an odd atomic number, have magnetic properties. Absorption of a quantum of energy $E = h\nu$ occurs only when the nuclei are in the strong magnetic field of the NMR spectrometer and when the frequency ν of the applied electromagnetic radiation is appropriate for "resonance" with the nucleus being observed. In the widely used "500-megahertz" NMR spectrometers the liquid helium-cooled superconducting electromagnet has a field strength of 11.75 tesla (T). In this field a proton resonates at ~500 megahertz (MHz) and nuclei of ^{31}P, ^{13}C, and ^{15}N at ~202, 125, and 50 MHz, respectively. At 500 MHz the energy of a quantum is only $E = 3.3 \times 10^{-33} \times 10^{8}$ J = 0.2 J mol^{-1}, more than four orders of magnitude less than the average energy of thermal motion of molecules (3.7 kJ mol^{-1}). Thus, the spin transitions induced in the NMR spectrometer have no significant effect on the chemical properties of molecules.

The resonance frequency ν at which absorption occurs in the spectrometer is given by Eq. 3-50, where H_o is the strength of the external magnetic field, μ is the magnetic moment of the nucleus being investigated, and h is Planck's constant. The basis for NMR spectroscopy lies in the fact that nuclei in different positions in a molecule resonates at slightly different frequencies. In a protein each one of the hundreds or thousands of protons resonate at its own frequency. With older NMR instruments a spectrum at the constant magnetic field H_o can be obtained by varying the frequency and observing the values at which absorption occurs, much as is done for ultraviolet, visible, and infrared spectra (Chapter 23). With newer pulsed NMR spectrometers the measurement is done differently but the spectra

$$\nu = \mu H_o / h \qquad (3\text{-}50)$$

look the same. The higher the magnetic field, the greater the variation in resonance frequency and the higher the sensitivity. The most powerful commercial NMR spectrometers currently available operate at about 750 MHz for ^1H and a few higher frequency instruments have been built.

The **proton NMR** spectrum of the coenzyme pyridoxal phosphate in ^2H_2O is shown in Fig. 3-24 as obtained with a 60-MHz spectrometer. Four things can be measured from such a spectrum: (1) the **intensity** (area under the band). In a proton NMR spectrum, areas are usually proportional to the numbers of equivalent protons giving rise to absorption bands; (2) the **chemical shift**, the difference in frequency between the peak observed for a given proton and a peak of some standard reference compound. In Fig. 3-24 the reference peak is at the right edge; (3) the **width** at half-height (in hertz), a quantity that can provide information about molecular motion and about chemical exchange; and (4) **coupling constants** which measure interactions between nearby magnetic nuclei. These are extremely important to the determination of structures of both small and large molecules.

With a magnetic field of H_o = 11.75 T and a 500-MHz oscillator, the positions of proton resonances in organic compounds are spread over a range of ~10,000 Hz. This is 20 parts per million (ppm) relative to 500 MHz. Positions of individual resonances are usually given in ppm and are always measured in terms of a shift from the resonance position of some standard substance. For protons this is most often **tetramethylsilane (TMS)**, an inert substance that can be added directly to the sample in its glass tube. Biochemists often use ^2H_2O as solvent and the water-soluble sodium 3-trimethylsilyl 1-propane sulfonate (**DSS** or Tier's salt) as a standard. Its position is insignificantly different from that of TMS. For NMR spectra measured in ^2H_2O, the "pD" of the medium is sometimes indicated. It has often been taken as the pH meter reading plus 0.4. However, because of uncertainty about the meaning of pD, most workers cite the apparent pH measured with a glass electrode and standardized against aqueous buffers.[438] It is important to describe how the measurement was made when publishing results.

Band widths. The narrowness of a band in an NMR spectrum is limited by the **Heisenberg uncertainty principle**, which states that $\Delta E \times \Delta t = h/2\pi$, where h is Planck's constant, ΔE is the uncertainty in the energy, and Δt is the lifetime of the magnetically excited state. Since $E = h\nu$ for electromagnetic radiation, E is directly proportional to the width of the absorption band (customarily measured at one-half its full height). The magnetic nucleus is well shielded from external influences and the lifetime of its excited state tends to be long. Hence, $\Delta\nu$ is small, often amounting to less than 0.2 Hz. This fact is very favorable for the success of high-resolution proton magnetic resonance. However, bands are often much broader for large macromolecules.

The chemical shift. In a molecule such as TMS, the electrons surrounding the nuclei "shield" the nucleus so that it does not experience the full external magnetic field. For this reason, absorption occurs at a high frequency (high energy). Protons that are bound

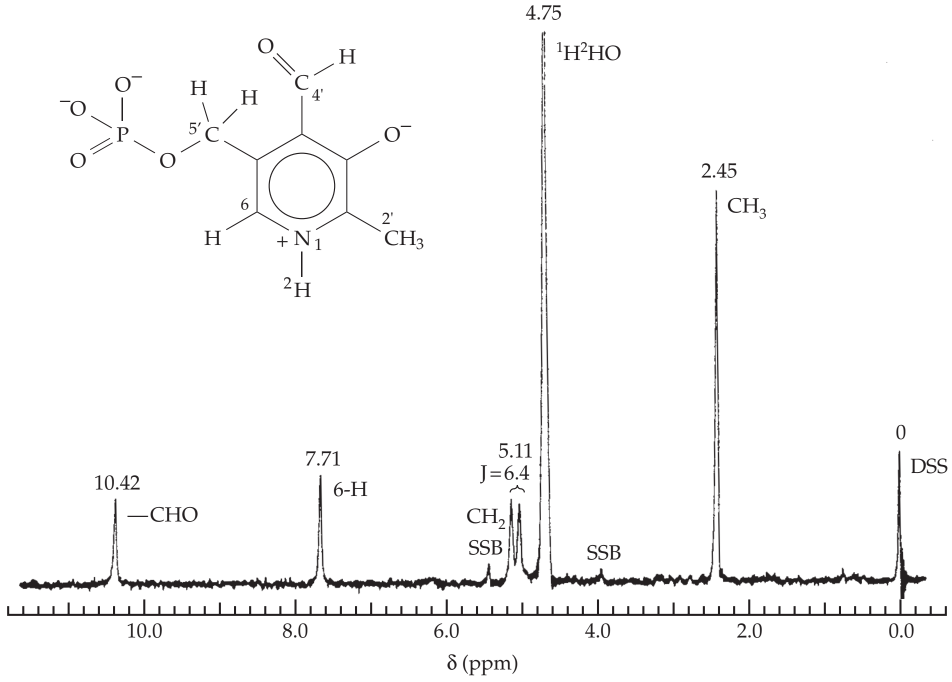


Figure 3-24 The 60-MHz proton magnetic resonance spectrum of pyridoxal 5'-phosphate at neutral pH (apparent pH = 6.65). The internal standard is DSS. Chemical shifts in parts per million are indicated beside the peaks. Spectrum courtesy of John Likos.

to an atom deficient in electrons (because of attachment to electron withdrawing atoms or groups) are **deshielded**. The greater the deshielding, the further **downfield** from the TMS position is the NMR peak.

The magnitude of this chemical shift may be stated in hertz, but it is most often expressed in ppm as δ (Eq. 3-51). The value of δ is the shift in frequency relative to frequency of the oscillator in parts per million and is independent of the field strength. It still depends upon use of a particular reference standard which must be stated when a δ value is given. In the spectrum shown in Fig. 3-24, the methyl protons appear 2.45 ppm below the DSS peak but still at a relatively high field. Characteristic chemical shift ranges for other protons (Table 3-3) extend to ~20 ppm.

$$\delta\,(\text{ppm}) = \frac{\Delta\nu\,(\text{Hz}) \times 10^6}{\nu\,(\text{Hz})\text{ of oscillator}} \tag{3-51}$$

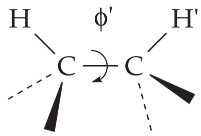
Aromatic rings lead to strong deshielding of attached protons because of a **ring current** induced in the circulating π electrons. Thus, in Fig. 3-24 the peaks of the methylene protons which are adjacent to the aromatic ring occur at 5.10 and 5.12 ppm. The 6-H, which is bound directly to the ring, is more strongly deshielded and appears at 7.71 ppm. The hydrogen of the aldehyde groups is deshielded as a result of a similar "diamagnetic electronic circulation" in the carbonyl group. Its peak is even further downfield at 10.4 ppm. Ring current and other effects on chemical shifts are important in NMR spectroscopy of proteins. Aromatic proton resonances sometimes stand out because they have been shifted far downfield. Ring current effects on chemical shifts can be predicted quite accurately if three-dimensional structures are known.[439] Additionally, computer programs are available for predicting them.[440–443]

Hydrogen bonding has a very large effect on the chemical shift of protons. The resonance of a strongly hydrogen-bonded proton is usually shifted downfield from its position in non-hydrogen-bonding media. This is especially true for hydrogen bonds to charged groups, e.g., the NH of a histidine or tryptophan side chain hydrogen bonded to a carboxylate group, a situation often met in the active sites of proteins. The chemical shift of ^{13}C in the carboxyl group is also affected.[444,445] While **ring current shifts** can be predicted quite well, it is much more difficult to predict the total chemical shift.[446,447]

Scalar coupling (J coupling). The energy of the spin transition of a hydrogen nucleus is strongly influenced by the local presence of other magnetic nuclei, e.g., other protons that are covalently attached to the same or an adjacent atom. These neighboring protons can be in either of the two spin states, a fact that results in easily measured differences in the energy of the NMR transition under consideration. This spin–spin interaction (coupling) leads to a splitting of NMR bands of protons into two or more closely spaced bands. The ethyl group often appears in NMR spectra as a "quartet" of four evenly spaced peaks that arise from the CH_2 group and a triplet of peaks arising from the CH_3 protons. Protons attached to the same carbon (**geminal** protons), and in similar environments, do not ordinarily split each other's peaks, while the protons on the neighboring carbon do.

The **coupling constant** *J* is the difference in hertz between the successive peaks in a multiplet. It is a field-independent quantity and the same no matter what the frequency of the spectrometer. In Fig. 3-24 the peak of the methylene protons is split by ^{1}H–^{31}P coupling, with a value of J ~6.4 Hz. While spin–spin coupling is most pronounced when magnetic nuclei are close together in a structure, the effect can sometimes be transmitted through up to five covalent bonds. The technique of **double irradiation** or **spin decoupling** can be used to detect spin coupling. The sample is irradiated at the resonance frequency of one of the nuclei involved in the coupling, while the spectrum is observed in the frequency region of the other nucleus of the coupled pair. Under these conditions the multiplet collapses into a singlet and the mutual coupling of the two nuclei is established. The coupling can be seen directly in appropriate two-dimensional NMR spectra.

The coupling constant between two **vicinal protons** which are attached to adjacent carbon atoms (or other atoms) depends upon the torsion angle.



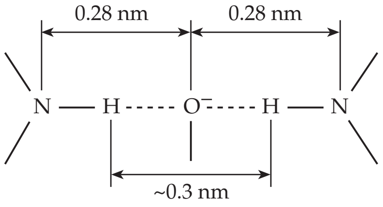
The **Karplus equation** (Eq. 3-52) relates *J* to the torsion angle ϕ′ (so labeled to distinguish it from peptide torsion angle ϕ; Fig. 2-8).

$$J_{H,H'} \quad A\,\cos^2\phi' + B\cos\phi' + C \tag{3-52}$$

This equation was predicted on theoretical grounds, but the constants, *A*, *B*, and *C* are empirical.[450] Other forms of the equation, some of them simplified, have also been proposed.[451] The Karplus relationship is often used to estimate time-averaged torsion angles in peptides. For a $C_\alpha H$–NH torsion angle the parameters *A*, *B* and *C* of Eq. 3-52 are ~ 6.4, 1.4, and 1.9, respectively. For a $C_\alpha H$ – $C_\beta H$ they are ~ 9.5, 1.6, and

1.8, respectively.[452] Coupling constants between ^{1}H and ^{13}C or ^{15}N are of importance in determination of three-dimensional structure of proteins.[453]

The nuclear Overhauser effect (NOE) is the result of transfer of magnetization from one nucleus to a nearby nucleus directly through space rather than via *J*-coupling.[427,454] This was observed first as a result of irradiation of a resonance in a one-dimensional spectrum resulting in an increased intensity of the resonance of the nearby nucleus. Magnetization transfer can occur from a given nucleus to one or more nearby nuclei. Each such transfer that is detected is usually referred to simply as an NOE. For an NOE to be observable the two nuclei must be very close together, <0.5 nm (5 Å). The strength of the magnetization transfer falls off approximately as the sixth power of the interatomic distance. Consider two hydrogen atoms, both tightly hydrogen-bonded to an intervening oxygen atom, e.g., of a carboxylate or phosphate group. The expected H–H distance would be ~0.3 nm and a strong NOE between them would be anticipated. Two nonbonded hydrogen atoms (e.g., on methyl groups of amino acid side chains) can be as close together as 0.24 nm at van der Waals contact and could show a very strong NOE. However, contact is rarely this close in proteins unless in a hydrogen bond.



2. Nuclei Other than Hydrogen

Deuterium (^{2}H). The natural abundance is very low so that use of ^{2}H-labeled compounds is practical for study of metabolism, e.g., for following an ^{2}H label in glucose into products of fermentation[455] or in mammalian blood flow.[456] Deuterium NMR has been used extensively to study lipid bilayers (Chapter 8).

Carbon - 13. Use of ^{13}C in NMR developed slowly because of the low natural abundance of this isotope. Another complication was the occurrence of ^{13}C–^{1}H coupling involving the many protons normally present in organic compounds. The latter problem was solved by the development of **wide-band proton decoupling** (noise decoupling). With a natural abundance of only 1.1%, ^{13}C is rarely present in a molecule at adjacent positions. Thus, ^{13}C–^{13}C coupling does not introduce complexities and in a noise-decoupled natural abundance spectrum each carbon atom gives rise to a single peak. Even so, ^{13}C NMR spectroscopy was not practical until pulsed Fourier transform (FT) spectrometers were developed (Section 2).

Chemical shifts in ^{13}C spectra are often 200 ppm or more downfield relative to TMS. The effects of substituents attached to a carbon atom are often additive when two or more substituents are attached to the same atom.[457,458] It is often necessary (but costly) to prepare compounds enriched in ^{13}C beyond the natural abundance. For proteins this may be done by growing an organism on a medium containing [^{13}C]glucose or a single amino acid enriched in ^{13}C.[459,460] Using metabolites enriched in ^{13}C, it is also possible to observe metabolism of living tissues directly. For example, glycogen synthesis from ^{13}C-containing glucose has been observed in a human leg muscle using a wide-bore magnet and surface coils for transmitting and receiving.[461] This topic is discussed further in Box 17-C.

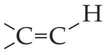
Nitrogen - 15. Despite difficulties associated with low natural abundance (0.37%) and low sensitivity, ^{15}N NMR is practical and with isotopically enriched samples has become very important. Proteins with a high content of ^{15}N can be produced easily and inexpensively from cloned genes in bacterial plasmids. For example, cells of *E. coli* can be grown on a minimal medium containing [^{15}N] NH_4Cl. Since ^{13}C can also be introduced in a similar way it is possible to incorporate both isomers simultaneously. Production of uniformly labeled protein containing ^{15}N and / or ^{13}C provides the basis for multidimensional isotope-edited spectra necessary for protein structure determination (next section) and for study of tautomerization of histidine rings (Eq. 2-6).[460,462–464] ^{15}N chemical shifts of groups in proteins are spread over a broad range (Table 3-3).[465]

Phosphorus - 31. NMR spectroscopy using ^{31}P, the ordinary isotope of phosphorus, also has many uses.[466] Application of ^{31}P NMR to living tissues has been extraordinarily informative[467] and is dealt with in Chapter 6. The many phosphorus nuclei in nucleotides, coenzymes, and phosphorylated metabolites and proteins are all suitable objects of investigation by NMR techniques.

Fluorine - 19. Although not abundant in nature, ^{19}F gives an easily detected NMR signal and can be incorporated in place of hydrogen atoms into many biochemical compounds including proteins.[468,469] In one study genetic methods were used to place 3-fluorotyrosine separately into eight positions in the lac repressor protein (Chapter 29).[470] Measurements of ^{19}F NMR spectra were used to study domain movement. Active site groups of enzymes can be modified to incorporate ^{19}F.[471] Binding of fluorinated substrates can be studied.[472] Nontoxic ^{19}F-containing compounds are useful as intracellular pH indicators, the NMR spectrometer serving as the pH meter (Box 6-A).[473,474] An atom of fluorine attached to an aromatic ring is highly and predictibly sensitive to inductive effects of substituents in the para position to the fluorine.[475] For example, fluorine at the 6-position in pyridoxal phosphate (Fig. 14-4) can be observed in enzymes and reports changes in coenzyme structure.[476,477]

Some other nuclei. Here are a few reported uses of NMR on other nuclei. ^{3}He, binding into little cavities in fullerenes;[478] ^{11}B, binding of boronic acids to active sites;[479] ^{23}Na, measurement of intracellular [Na$^+$];[480–482] ^{35}Cl and ^{37}Cl, binding to serum albumin;[483] ^{113}Cd, reporter that can replace Zn^{2+} (no magnetic moment) in active sites of many enzymes and in nonenzymatic systems as well;[484] Tl, replacing K$^+$ in enzyme binding sites;[485] ^{17}O, study of dynamics of protein hydration;[486] and ^{77}Se, observation of acetylchymotrypsin intermediate.[487]

TABLE 3-3
Approximate Chemical Shift Ranges in ^{1}H- and in ^{15}N- NMR Spectra

Group	^{1}H chemical shift (ppm from TMS)	^{15}N chemical shift (ppm from liquid NH_3)
$-CH_3$	0–4.0	
$-CH_2-$	1.1–4.4	
$-CH$	2.4–5	
Peptide αH random coil[a]	3.9–5.0	
$-OH$[b]	~ 5–6	
	5–8	
$-NH_2$		31– 37
NH		
Peptide	7–12	103 – 142
Aromatic H	7–9	
Imidazole		
$C^{\varepsilon1}-H$	~7.7	
$C^{\delta2}-H$	~7.0	
N–H	~ 10	165 – 180
Imidazolium		
$C^{\varepsilon1}-H$	~8.7	
$C^{\delta2}-H$	~7.4	
N–H	10–18	
$-CHO$ Aldehyde	9.4–10.4	
$-COOH$	11.3–12.2	
Indole NH	~ 10	130 – 145
Guanidinium		
$-N^{\eta}H_2$		69 –77
$-N^{\varepsilon}H$		31–37

[a] See Wishart *et al.*[448]
[b] See Linderstrøm-Lang[449]

3. Fourier Transform Spectrometers and Two-Dimensional NMR

Although NMR spectroscopy was widely used by the 1950s it was revolutionized by two developments, pioneered by Richard Ernst in the mid 1960s and 1970s.[488] The first of these was pulsed Fourier transform (FT) spectroscopy, which permits rapid accumulation of high-resolution spectra. In an FT NMR spectrometer a strong pulse of radiofrequency (RF) radiation is delivered to the sample over a period of a few microseconds and its effects are observed at all frequencies simultaneously. Although 1–2 or more seconds must be allowed before the next pulse is delivered, one complete NMR spectrum is obtained with each pulse. Often the results of hundreds, thousands, or even hundreds of thousands of pulses are added to provide greater sensitivity. With good temperature control this may be accomplished over periods of minutes to days.[489,490]

Free induction decay. The strong exciting RF pulse is delivered with an orientation at right angles to that of the static field H_o of the magnet and whose direction defines the z axis. As a result of this pulse, the magnetization of a nucleus is tilted away from the z axis and *precesses around the z axis* at its resonance (Larmor) frequency, which is ~500 MHz for ^{1}H in a 500-MHz spectrometer. The frequency that is measured is actually a difference from the “carrier frequency” of the RF pulse. The precessing magnetization of the nuclei has a component in the xy plane which induces an electrical signal in the coil of the NMR probe which defines the y axis. This signal, which contains the Larmor frequencies of all of the nuclei of a given element, is recorded as a function of time over a period of a few seconds. The signal decays away exponentially. However, this curve of **free induction decay** (FID) is not smooth but contains within it all of the Larmor frequencies. If enough points (perhaps 500 in a 2-s acquisition) are recorded and stored in the computer’s memory, Fourier transformation of the data will produce the frequency-dependent NMR spectrum.[489,490]

Relaxation times T_1 and T_2. When a very strong pulse of electromagnetic radiation is applied in the NMR spectrometer, virtually all of the nuclei are placed in the magnetically excited state. If another pulse were applied immediately, little energy would be absorbed because the system is **saturated**. In the older "continuous-wave" NMR spectrometers, the energy is always kept small so that little saturation occurs. However, in FT NMR instruments, the strong pulses lead to a high degree of saturation. Application of repeated pulses would produce no useful information were it not for the fact that the excited nuclei soon relax back to their equilibrium energy distribution. Relaxation occurs through interactions of the nuclei with fluctuating magnetic fields in the environment. For organic molecules in solution the fluctuations that are most often effective in bringing about relaxation are the result of moving electrical dipoles in the immediate vicinity. Even so, relaxation of protons in water requires seconds.

Relaxation of nuclear magnetic states is characterized by two relaxation times. The longitudinal or **spin-lattice relaxation time** T_1 measures the rate of relaxation of the net magnetic vector of the nuclei in the direction of H_o. The transverse or **spin–spin relaxation time** T_2 measures the relaxation in the xy plane perpendicular to the direction of H_o. The two relaxation times can be measured independently. In general, $T_2 < T_1$. For solids, T_2 is quite short (~10^{-5} s) whereas relaxation times of seconds are observed in solutions. This lengthening of the lifetime of the excited state in going from solid to liquid leads to a narrowing of absorption lines and explains why NMR bands in liquids are often narrow. However, an increase in viscosity or a loss of fluidity in a membrane leads to broadening.

How can T_2 and T_1 be measured? T_2 for fluids can often be estimated from the width of the band $\Delta\nu$ at half-height (Eq. 3-53). However, pulsed NMR methods are usually employed, the measurement of T_1 being especially easy.

$$T_2 \approx 1/\pi\,\Delta\nu \tag{3-53}$$

An attempt is often made to relate T_1 and T_2 to the molecular dynamics of a system. For this purpose a relationship is sought between T_1 or T_2 and the **correlation time τ_c** of the nuclei under investigation. The correlation time is the time constant for exponential decay of the fluctuations in the medium that are responsible for relaxation of the magnetism of the nuclei. In general, $1/\tau_c$ can be thought of as a rate constant made up of the sum of all the rate constants for various independent processes that lead to relaxation. One of the most important of these ($1/\tau_1$) is for molecular tumbling.

$$1/\tau_1 \approx (3k_BT)/4\pi\eta r^3 \tag{3-54}$$

This equation is closely related to that of rotational diffusion (Eq. 9-35). Another term is the reciprocal of the **residence time** τ_m, the mean time that a pair of dipoles are close enough together to lead to relaxation.

In the usual solvents at room temperature, τ_c is of the order of 10^{-12} to 10^{-10} s. Thus, relaxation rates in solutions are considerably faster than the frequencies of radiation absorbed in the NMR spectrometer (~10^8 s^{-1}). Relaxation is relatively ineffective and T_1 and T_2 are usually large and equal. Bands remain sharp. As the correlation time increases (as happens, for example, if the viscosity is increased), T_1 and T_2 decrease with T_1 reaching a minimum when $\tau_c \sim v$, the frequency of the absorbed radiation. Lines are broadened and hyperfine lines (from coupling between nuclei) cannot be resolved. As τ_c is increased further, T_2 reaches a constant low value, while T_1 rises again. NMR measurements can be made in the region where τ_c exceeds ν, a circumstance that is favored by the use of high-frequency spectrometers. On the other hand, in fluids it is more customary to work in the range of "extreme motional narrowing" at low values of τ_c. Both T_1 and T_2 rise as the mobility of the molecules increases.

A limitation of use of NMR measurements of proteins comes from the increase in tumbling time with increasing size of the molecules. Since $1/\tau_r$ is often the most important term in the relaxation rate constant, only small proteins of mass <20 kDa give very sharp bands. Nevertheless, usable spectra are often obtainable on proteins ten times this size.

A practical problem in ^{13}C NMR arises from slow relaxation (long T_1). Partial saturation is attained and signal intensities are reduced for those carbon atoms for which relaxation is especially ineffective. Relaxation times can be measured separately for each carbon atom in a molecule and can yield a wealth of information about the **segmental motion** of groups within a molecule. Although the relationships between relaxation times and molecular motion are complex, they are often relatively simple for ^{13}C. Carbon atoms are usually surrounded by attached hydrogen atoms, and dipole–dipole interactions with these hydrogen atoms cause most of the nuclear relaxation. For a carbon atom attached to N equivalent protons in a molecule undergoing rapid tumbling, Eq. 3-55 holds. Where $h = h/2\pi$ and γ_C and γ_H are the magnetogyric ratios of carbon and hydrogen nuclei.

$$1/T_1 \approx \frac{Nh^2\gamma_c{}^2\gamma_H{}^2\tau_{\text{eff}}}{r^6} \tag{3-55}$$

This equation permits a calculation of an effective correlation time τ_{eff} for each carbon atom.[491] T_1 and T_2 can also be evaluated for individual ^{15}N or ^{13}C nuclei in labeled proteins.[491]

Two-dimensional and multidimensional NMR spectra. Proteins have such complex NMR spectra that, except for small regions at the upfield and downfield ends (see Fig. 3-26A), it is impossible to interpret one-dimensional spectra. A solution to this problem came from the development by Jeener, Ernst, and Freeman of methods of displaying NMR spectra in two dimensions.[428,490,492,493] All two-dimensional and multidimensional NMR methods make use of one basic procedure: After the initial RF pulse, a second or a series of subsequent RF pulses are introduced. This is done before the nuclei have had time to relax completely. The time t_1 from the initial pulse to the second pulse is called the **evolution period** and allows accumulation of information about NOEs or *J*- coupling, whether homonuclear (e.g. $^1H-^1H$ or $^{13}C-^{13}C$) or heteronuclear (e.g., $^1H-^{13}C$ or $^1H-^{15}N$).

Two-dimensional spectra usually require hours or days of acquisition because separate FIDs are collected for a series of many different values of t_1. This provides a second timescale for the experiment. The second pulse is often used to rotate the directions of magnetization of the nuclei that are being observed from the *xy* plane into the *yz* plane but with the *z* component at 180° to the H° vector. One very important type of two-dimensional plot is the **NOESY** (NOE spectroscopy) spectrum, which detects NOEs between all excited nuclei that are close enough together (Fig. 3-25B). The transfer of magnetization occurs during a **mixing time** Δ, which follows the second pulse. For a protein Δ may be 25–300 ms. A third 90° pulse returns the *z* components of the magnetization to be parallel with the *y* axis and the FID is collected over the period t_2. The amplitude of the resonances detected is modulated by the frequencies that existed during the evolution

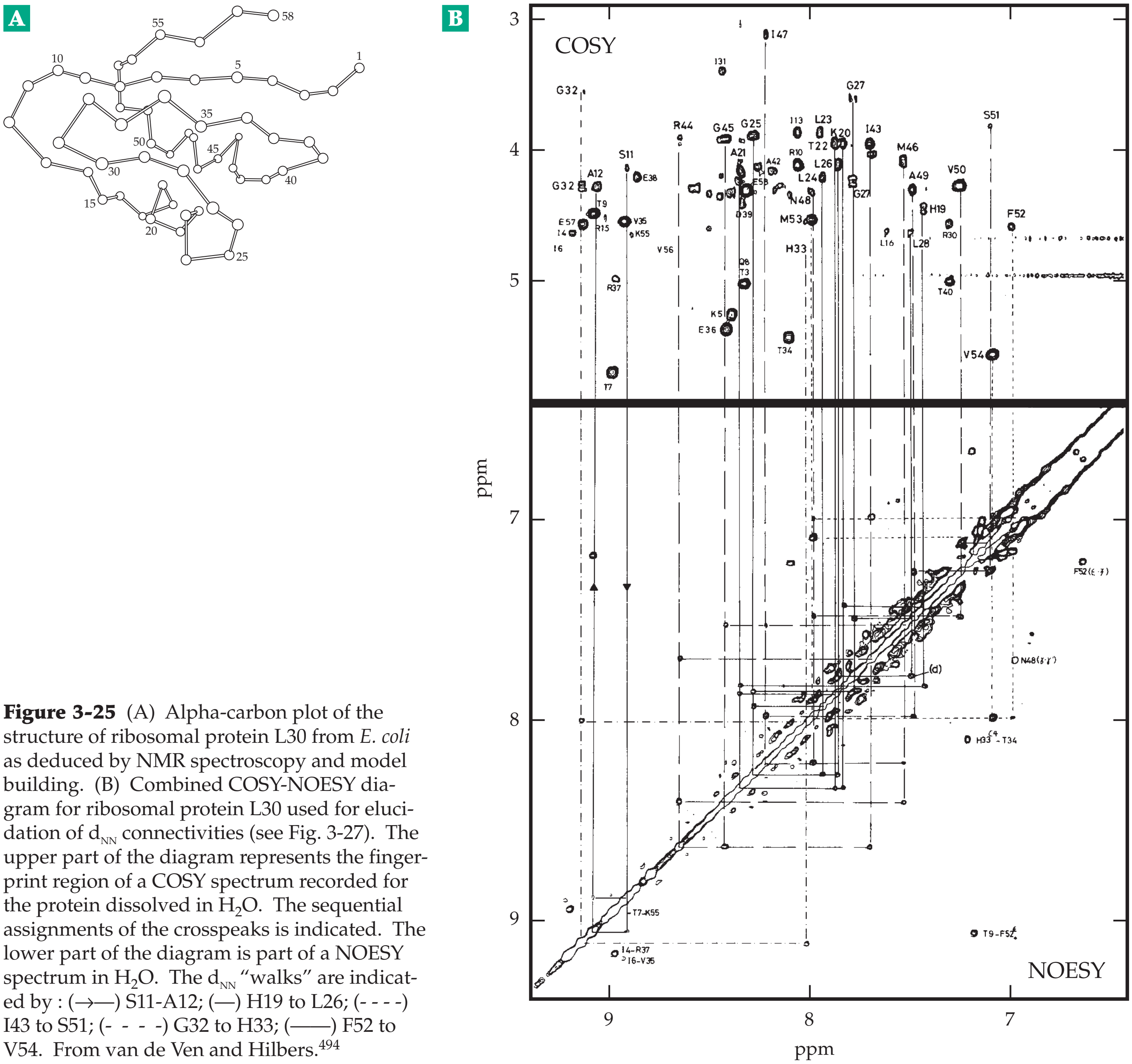


Figure 3-25 (A) Alpha-carbon plot of the structure of ribosomal protein L30 from *E. coli* as deduced by NMR spectroscopy and model building. (B) Combined COSY-NOESY diagram for ribosomal protein L30 used for elucidation of d_{NN} connectivities (see Fig. 3-27). The upper part of the diagram represents the fingerprint region of a COSY spectrum recorded for the protein dissolved in H_2O. The sequential assignments of the crosspeaks is indicated. The lower part of the diagram is part of a NOESY spectrum in H_2O. The d_{NN} "walks" are indicated by : (→—) S11-A12; (—) H19 to L26; (- - - -) I43 to S51; (- - - -) G32 to H33; (——) F52 to V54. From van de Ven and Hilbers.[494]

period. After Fourier analysis[436,490] a two-dimensional plot with two frequency axes is generated and is usually displayed as a contour plot. Along the diagonal of Fig. 3-25B are peaks representing the one-dimensional spectrum. All of the peaks off of the diagonal are NOEs which can be related back to the peaks on the diagonal as shown by the horizontal and vertical lines.

A second important two-dimensional method is **correlation spectroscopy (COSY)**, in which the pairs of off-diagonal peaks result from spin–spin coupling. A related method called **TOCSY** provides correlations that extend through more than three bonds. The COSY plot in Fig. 3-26B is for the synthetic cyclic decapeptide *cyclo*-(Δ^3-Pro–D-*p*-Cl-Phe–D-Trp–Ser–Tyr–D-Trp–N-Me-Leu–Arg–Pro–β-Ala). It was obtained in six hours on a 500-MHZ instrument. The region marked I reveals couplings of protons on α-carbons to those on adjacent β carbons within the same residue ($J_{\alpha\beta}$; Fig. 3-27) and other couplings within the side chain. Each amino acid has a characteristic pattern. From careful study of this region it is possible to correlate each α-H resonance with a particular amino acid side chain. However, some residues are difficult to distinguish, e.g., His, Trp, Phe, and Tyr have similar $J_{\alpha\beta}$ values. Region II of Fig. 3-26B reveals connectivities of α and β hydrogens to N-H protons in the 7– 9 ppm region. Each α-H is coupled to the N-H of the same residue ($J_{\alpha H}$, Fig. 3-27). A section of a COSY plot is also shown in Fig. 3-25B and indicates how resonances can be related to those in the NOESY plot made on the same sample.

Of great importance in the determination of protein structures is the use of ^{15}N- or ^{13}C-enriched samples to obtain **isotope-edited** spectra. For example in **HSQC** or in **^{15}N-multiple quantum cohenence (HMQC)** spectra we see only NH protons in a plot of ^{1}H chemical shift in one dimension versus the ^{15}N chemical shift of the attached

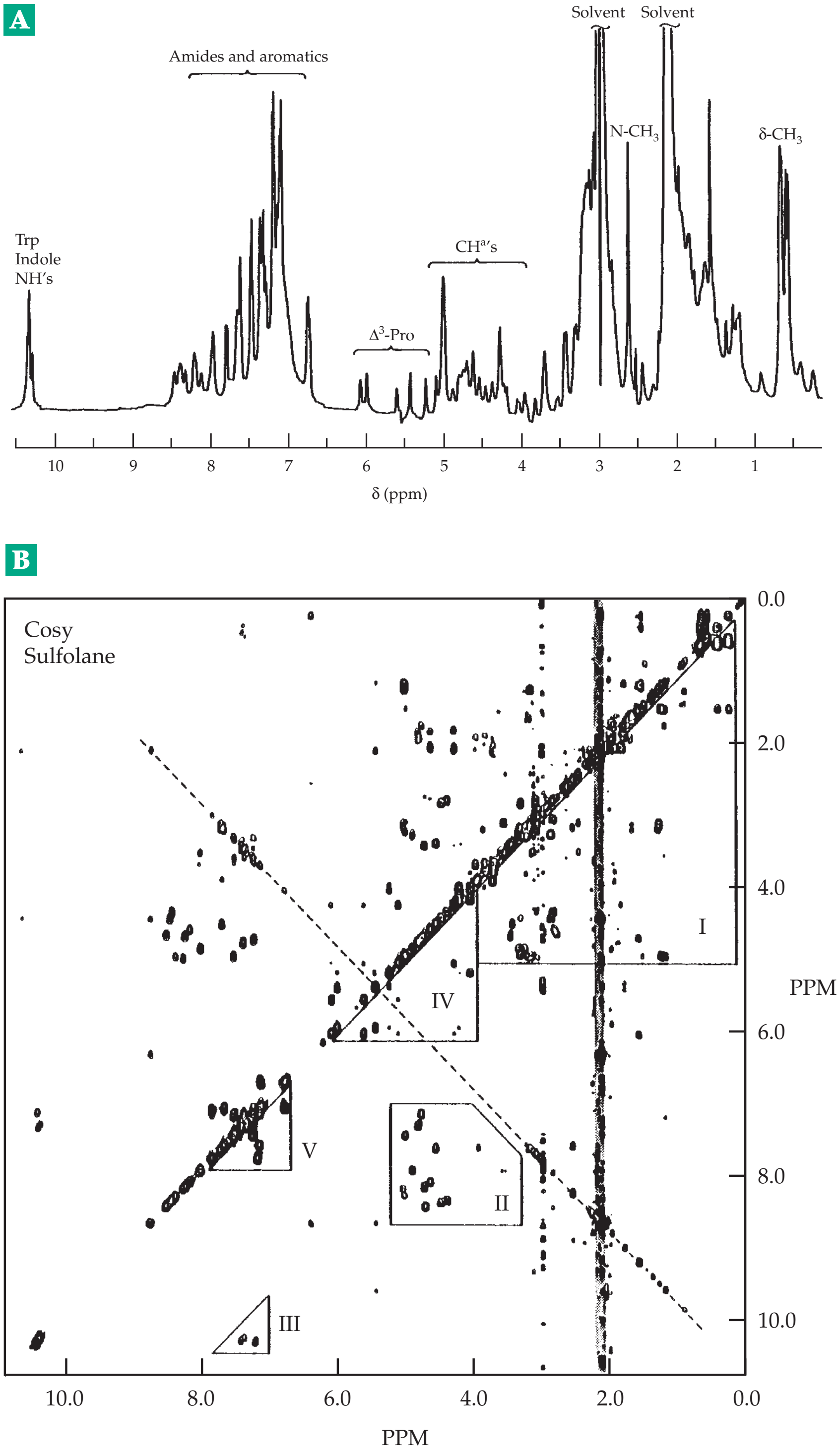


Figure 3-26 Proton NMR spectra of a cyclic decapeptide analog of the gonadotropin-releasing hormone in the solvent sulfolane at 500 MHz. (A) One-dimensional spectrum. This figure also illustrates upfield methyl group regions, α-hydrogen, amide, aromatic, and downfield (10–20 ppm) regions. The indole NH resonances are shifted downfield by the ring current of the indole. (B) COSY spectrum plotted as a contour map. The outlined areas represent five unique *J*-coupled regions; area I, the C$^{\alpha}$ to C$^{\beta}$ to C$^{\gamma}$ etc., side chain connectivities; area II, the NH to C$^{\alpha}$ connectivities; area III, the indole NH of tryptophan; area IV, the connectivities of the Δ^3-Prol residue; and area V, the aromatic resonances. Peaks that appear on the solid vertical line and the dashed diagonal lines are artifacts. From Baniak *et al.*[495]

nitrogen atom in the other (Fig. 3-28). Furthermore, as shown in this figure, particular types of NH bonds (peptide, $-NH_2$, imidazole, indole, amide, and guanidinium) appear in different regions. There is only one peptide NH per residue and, for a small protein, each may be separately visible.

4. Three-Dimensional Structures and Dynamics of Proteins

The first three-dimensional structure of a small protein was determined solely from NMR measurements in 1984. To date hundreds of "NMR structures" have been deduced. For small proteins of $M_r < 10{,}000$ two-dimensional COSY and NOESY spectra can suffice. For larger proteins use of ^{15}N- and/or ^{13}C-enriched proteins is essential.[428,430,496] This permits generation of a third and even a fourth frequency axis and three- and four dimensional NMR. For example, using various complex pulse sequences an ^{15}N-correlated $^1H - {}^2H$ NOESY spectrum can be generated. This shows directly which NOEs arise from NH protons with resonances in the amide region of an HSQC or HMQC spectrum (Fig. 3-28). With both ^{15}N and ^{13}C present, many additional coupling patterns and J values can be observed.[497–502] It is also possible to measure NOEs from atoms in a protein to those of a relatively weakly bound ligand such as a coenzyme or substrate analog and to determine the conformation of the bound ligand from this **transferred NOE**.[503–505]

Assignment of resonances. After acquisition of the necessary data, which may require 10–15 mg of protein, the observed resonances must be assigned to specific amino acid residues in the peptide chain. The connectivities of the individual CH and NH groups that have been identified in the COSY spectrum and information about the relationship of one atom to another atom nearby in space are required. The closest neighbors to either αH or peptide NH protons are often protons in a neighboring residue ($d_{\alpha N}$, d_{NN}, Fig. 3-27). As was pointed out in the preceding section, NOE correlations obtained from plots such as that in Fig. 3-25B provide much of the information needed to establish which resonances belong to each residue in a known sequence.[494] The three-dimensional structure of the ribosomal protein L30 of *E. coli*. (Fig. 3-25A) was deduced entirely by NMR spectroscopy. A downfield part of the NOESY and COSY plots used is shown in Fig. 3-25B. This figure also shows how cross-peaks in the NOESY spectrum were correlated with identified COSY peaks. It is helpful initially to hunt for unique dipeptides that can be identified in the NOESY spectrum. Ambiguities that arise can be resolved by use of various additional techniques.

NOEs and distance constraints. NOESY plots also contain the essential information needed to determine which side chains *distant* in the sequence are close together in space. A NOE observed for a pair of nuclei falls off as the inverse sixth power of the distance between them. For this reason, NOEs are observed only for pairs of atoms closer than about 0.4 nm. It is possible, in principle, to calculate distances between nuclei from the NOE intensity, but this is not accurate. Often, the NOE cross-peaks are grouped into three categories that correspond to maximum possible distances of 0.25, 0.30, and 0.40 nm. These can be related for the most part to intraresidue (e.g., $d_{N\alpha}$, of Fig. 3-27) sequential (e.g., $d_{\alpha N}$, and d_{NN}, Fig. 3-27) and long range backbone–backbone distances. These values constitute a series of **distance constraints** which are applied while making an automated computer search for a folding pattern that will meet these constraints and at the same time have acceptable torsion angles and good side chain packing throughout. This process makes use of distance geometry algorithms and other methods.[430,506,507] An early success was the solution of a 75-residue amylase inhibitor independently by crystallographers[508] and NMR spectroscopists.[509] The NMR structure was based on 401 NOE distance constraints, 168 distance constraints imposed by hydrogen bonds and 50 torsion angles deduced from J values. Recently, refinement of NMR structures has been done as in X-ray crystallography.[507] Some structures have been refined using both NMR and X-ray data.[442]

The spectra in Figs. 3-25, 3-26 and 3-28 are for relatively small proteins. Spectra of larger proteins are more complex and lines are broader. Many techniques are used to simplify spectra. The NMR spectrum of a protein is simplified considerably if the protein is denatured by heating, and ^1H NMR spectra of "random coil" proteins can be predicted well from tables of standard chemical shifts for the individual amino acids.[510] Many amide NH protons exchange with solvent rapidly, making it easier to assign the remaining peaks. However, nearly all NH peaks will be seen in an HMQC or HSQC (Fig. 3-28) spectrum. Partial, or even complete, substitution of deuterium for hydrogen will also simplify spectra.[511–513] Microorganisms that will grow in a medium rich in D_2O can be used as sources of partially deuterated proteins. Because the remaining protons usually have ^2H rather than ^1H as a neighbor, dipolar line broadening is reduced and sharper resonances are observed.[514] Substitution of ^{15}N for ^{14}N in the backbone amide groups can also yield spectra with narrower lines.[515] **Isotope-edited** NMR spectra allow simplification of complex two-dimensional spectra by observation of only those protons attached to an isotopically labeled nucleus, e.g., ^{13}C or ^{15}N.[516–518] Measurement of $^{15}NH\text{-}C_{\alpha H}$ J couplings facilitates structure determinations.[519]

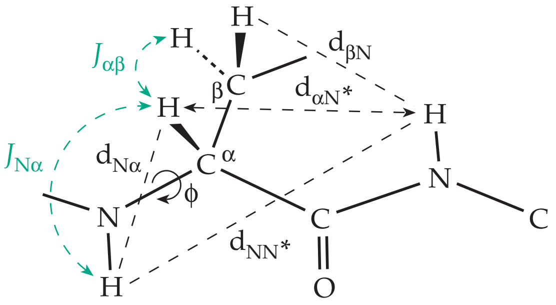


Figure 3-27 Illustration of some distances (d$_{NN}$, d$_{\alpha N}$ and d$_{\beta N}$ obtained from NOESY spectra and some coupling constants ($J_{\alpha\beta}$ and $J_{N\alpha}$ obtained from ^{1}H COSY spectra or J-resolved spectra. The coupling constants and proton chemical shifts provide a "fingerprint" for each residue and $J_{N\alpha}$ may also provide an estimated value for torsion angle f. The distances establish residue-to-residue connectivities as well as distance constraints that may permit a calculation of three-dimensional structure. Additional coupling constants can be measured for ^{15}N- or ^{13}C- enriched proteins. Coupling from β-hydrogens to other side chain hydrogens provides "fingerprint" information about individual residues.

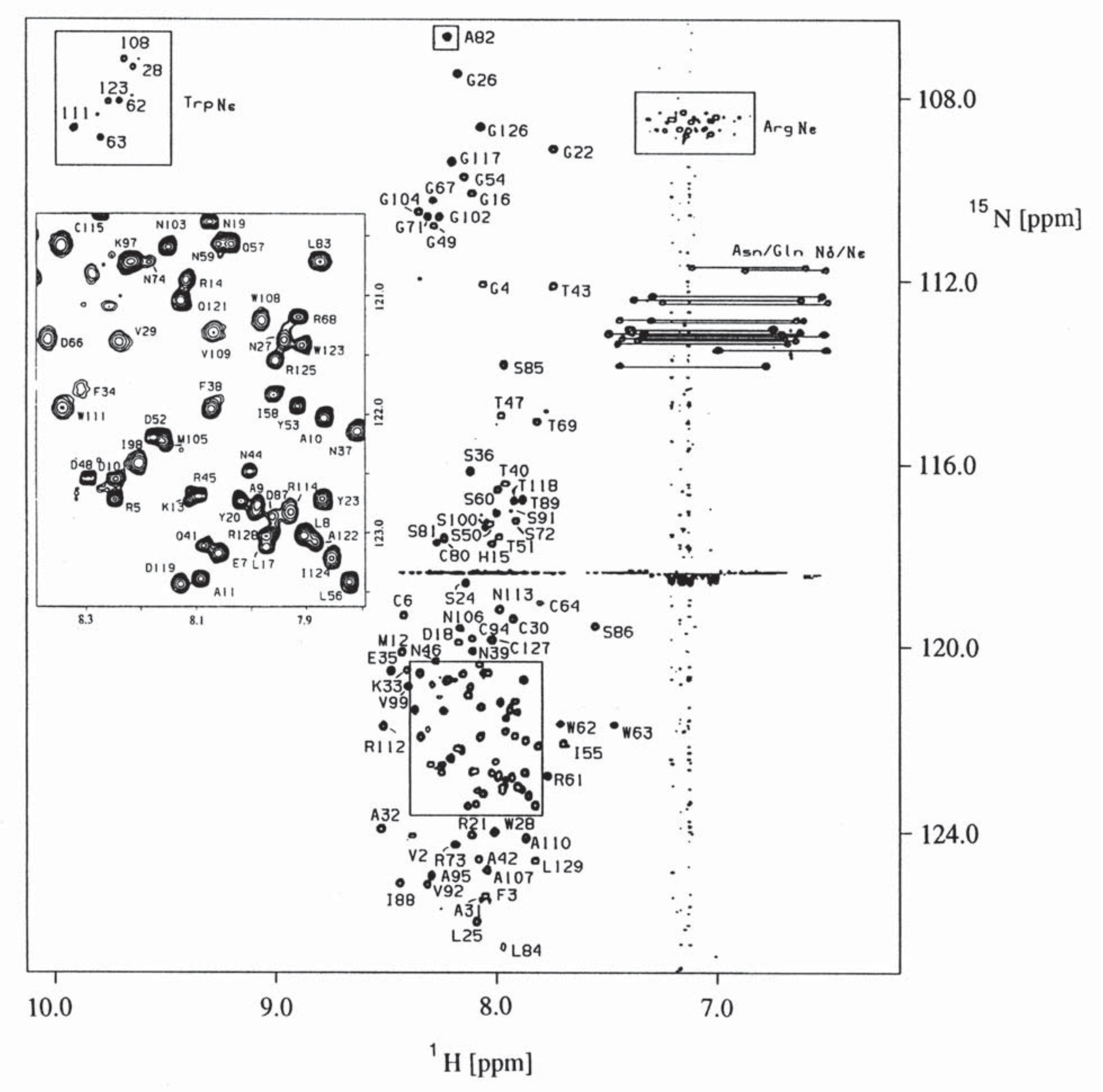


Figure 3-28 A ^{15}N – ^{1}H HSQC spectrum of partially denatured 129-residue hen lysozyme. Boxes enclose the tryptophan indole region (upper left), the arginine side chain N$^{\varepsilon}$ region (upper left), and a portion of the amide NH region (lower center and enlarged in the insert). Resonances of pairs of hydrogen atoms in side chain (Asn and Gln) amide groups are indicated by horizontal lines. From Buck *et al.*[524]

Many aspects of the dynamics of the action of enzymes of this size can be studied by NMR spectroscopy. New approaches allow study of proteins up to ~30 kDa in size.[519a] Computer programs can analyze data and make automatic assignments of resonances.[520,521]

Also useful are techniques for measuring NMR spectra on solids, including microcrystalline proteins[522] (see also Box 9-C). In some cases, e.g., for heme proteins,[522,523] the crystals can be oriented in a magnetic field permitting measurements of NMR spectra with more than one orientation of the crystals. This potentially affords more information than the usual techniques. For example, growth of bacteria in ^2H_2O media containing ^1H-containing pyruvate yields proteins with almost complete deuteration in the C-α an C-β positions but with highly protonated methyl groups. This gives rise to good $-CH_3$ to $-CH_3$ NOEs and provides other advantages in both NMR spectroscopy and mass spectrometry.[512] Isotopically enriched amino acids, e.g., ^{13}C-enriched leucine,[525] or isoleucine containing ^{15}N, ^{13}C, and ^2H as well[526] can be fed to growing bacteria. The use of paramagnetic shifts by ions such as Gd^{3+} may be helpful.[527] Very large shifts are sometimes induced in heme proteins by the Fe^{3+} of the heme that is embedded within the protein.[528]

5. Other Information from NMR Spectra

NMR titrations. How do pH changes affect NMR resonances? Resonances of ^1H nuclei close enough to a proton with a pK_a in the pH region under study will experience a shift, which may be either upfield or downfield when the proton dissociates. The resonance of ^{13}C in a carboxyl group will shift downfield when the proton on the carboxyl group dissociates. If the proton that dissociates is tightly hydrogen bonded in a protein, its rate of dissociation may be *slow* compared to the NMR frequency used. If so, the original peak will decrease in value as the pH is raised and a new peak will appear at a position characteristic of the dissociated form. However, protons attached to N or O are usually in *rapid* exchange with the solvent. In this case, the NMR resonance of the nucleus being observed will move continuously from one chemical shift value at low pH to a different one at high pH. In an intermediate case the resonance will shift and broaden.

Both the $C^{\delta 2}$-H and $C^{\varepsilon 1}$-H protons of histidine can often be seen in proteins (Fig. 3-29A).[529–531] As is shown in Fig. 3-29 A and B, their chemical shifts are strongly dependent upon the state of protonation of the ring nitrogen atoms. At low pH, the positive charge that is shared by the two NH groups attracts electrons away from both CH positions, causing deshielding of the CH protons. The effect is greater for the $C^{\varepsilon 1}$-H than for $C^{\delta 2}$-H protons (Table 3-3). If the groups being titrated do not interact strongly with other nearby basic or

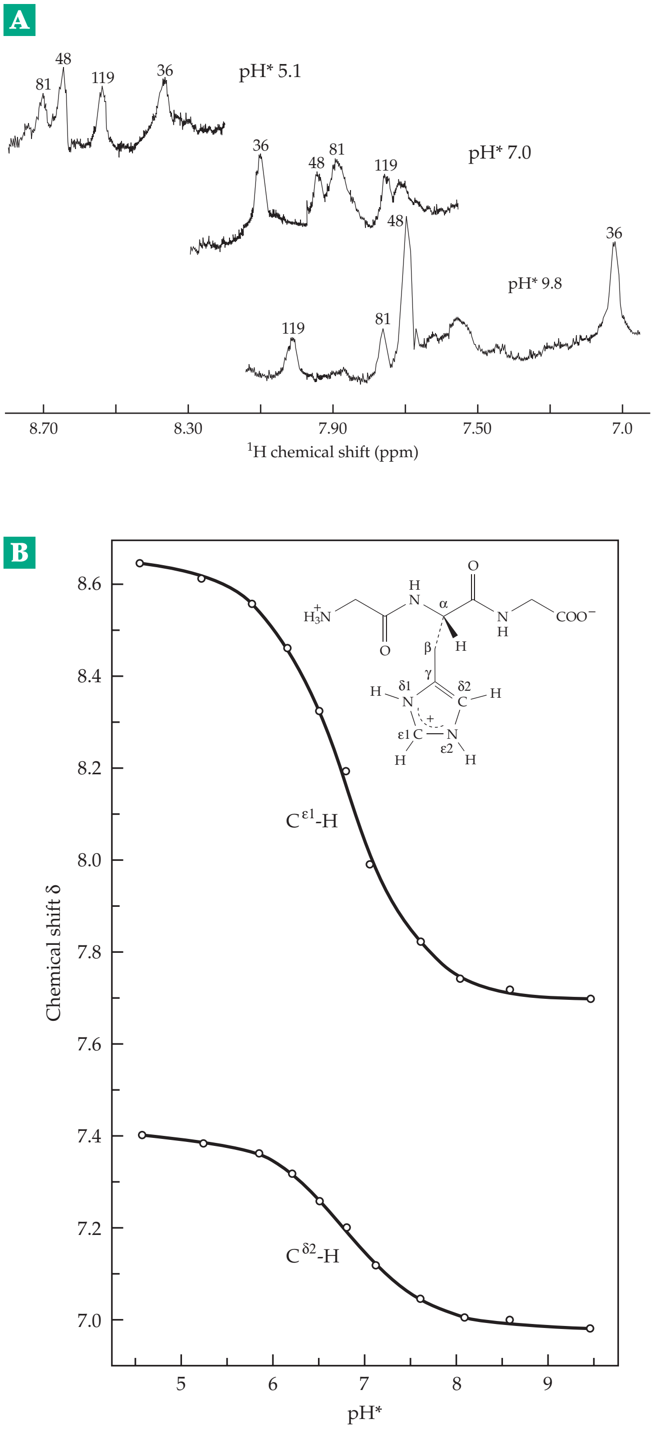


Figure 3-29 (A) ^1H NMR spectra of human myoglobin in D_2O showing the C-$H^{\varepsilon 1}$ resonances of the imidazole rings of histidines 36, 48, 81, and 119 at three values of the apparent pH (pH*). From Bothelho and Gurd.[529] (B) ^1H NMR titration curves for the histidine $C^{\varepsilon 1}$ and C-$^{\delta 2}$ ring protons of glycyl-L-histidylglycine. Data obtained at 100 MHz using 0.1 M tripeptide in D_2O containing 0.3 M NaCl. The pH values (labeled pH^*) are uncorrected glass electrode pH meter readings of D_2O solutions using an electrode standardized with normal H_2O buffers. Chemical shifts are downfield from TMS. Courtesy of J. L. Markley.[438]

acidic groups, pK_a values can be estimated for individual histidines.[530,532] With ^{15}N-containing proteins the tautomeric states (NH$^{\varepsilon 2}$ vs NH$^{\delta 1}$) of the imidazole rings can also be deduced.[533] Observation of imidazole rings shows that in proteins pK_a values of imidazolium groups are sometimes less than five and sometimes greater than ten.[463,534]

Observing exchangable protons. ^{1}H spectra of proteins are often recorded in D$_2$O because of interference from the very strong absorption of H$_2$O at approximately 4.8 ppm. However, resonances of some rapidly exchanging protons are lost. Special pulse sequences, as well as improvements in spectrometer design, allow many of these resonances to be seen in H$_2$O.[535–537] At the far downfield end (δ>10) of ^{1}H NMR spectra there are often weak peaks arising from NH protons of imidazole or indole side chains that can be observed in H$_2$O.[463,538–540] These resonances are often shifted 2–5 ppm downfield from the positions given in Table 3-3. The ^{1}H resonance for a carboxyl (– COOH) proton is shifted to 20 ppm or more in spectra of very strongly hydrogen-bonded anionic complexes such as the malonate dianion.[541–544] (see Chapter 9, Section D,4). The strongest hydrogen bonds cause the greatest downfield shifts because the negative charge pulls the ^{1}H proton away from the electrons of the atom to which it is attached, deshielding the proton.[545]

A good example is provided by the imidazole NH proton of the active site of trypsin and related serine proteases (Chapter 12), which is seen at ~16 ppm. Another example is provided by aspartate aminotransferase, whose ^{1}H NMR spectrum in a dilute aqueous phosphate buffer is shown in Fig. 3-30. The peak labeled A is the resonance of the NH proton on the ring of the pyridoxal phosphate coenzyme (marked in Fig. 14-6) and peak B belongs to an adjacent imidazole group of histidine 143. Both of these protons move with pH changes around a pK_a of ~ 6.2 which is associated with the Schiff base proton 6–8 nm away.[534,546–551] Peak A moves upfield 2.0 ppm and peak B downfield 1.0 ppm when the pH is raised around this pK_a. These hydrogen-bonded protons act as sensitive "reporters" of the electronic environment of the active site. Many proteins contain carboxylate or phosphate groups in their active sites and observation of NMR resonances of protons hydrogen bonded to them or to groups in substrates or inhibitors may be a useful technique for study of many enzymes and other proteins.

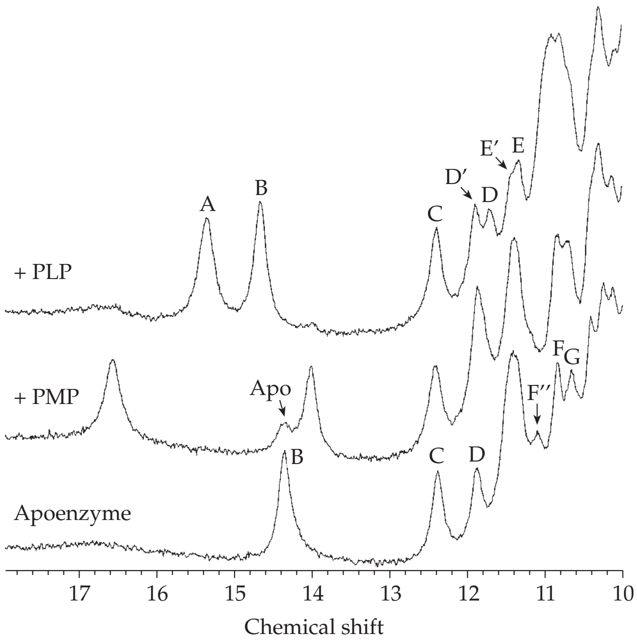


Figure 3-30 Spectra of the pyridoxal phosphate (PLP), pyridoxamine phosphate (PMP) and apoenzyme forms of pig cytosolic aspartate aminotransferase at pH 8.3, 21°C. Some excess apoenzyme is present in the sample of the PMP form. Spectra were recorded at 500 MH$_2$. Chemical shift values are in parts per million relative to that of H$_2$O taken as 4.80 ppm at 22°C. Peak A is from a proton on the ring nitrogen of PLP or PMP, peaks B and D are from imidazole NH groups of histidines 143 and 189 (see Fig. 14-6), and peaks C and D′ are from amide NH groups hydrogen bonded to carboxyl groups.

Exchange rates of amide protons. The NH protons of the peptide backbone can be observed in H$_2$O in the 6 –11 ppm region. If the spectrum is recorded for a sample in D$_2$O, many of these resonances disappear gradually as the protons of the peptide units exchange with the deuterium ions of the medium. A study of the observed exchange rates can shed light on the dynamics of protein molecules in solution. Exchange of amide protons in proteins with ^{2}H or ^{3}H has been studied on a relatively slow timescale by such techniques as observation of infrared vibrations of the amide group (see Chapter 23).[552,553] However, the development of two-dimensional NMR dramatically improved the ability to study proton exchange.[552,554–556] The measurements may involve use of quenching by rapid solvent change,[556] special pulse sequences,[555] and study of T_1 relaxation rates (seconds timescale). Recently, electrospray mass spectrometry has also been exploited.[557,558] Exchange patterns for large proteins may be followed,[559] and proteolytic fragmentation into short peptides may be used after various lengths of exchange time to investigate the exchange in specific regions of a protein.[216,560] Unfolding of a protein under denaturing conditions can be studied, as can refolding of a completely denatured protein.[561]

Amide hydrogen exchange is usually discussed in terms of a model proposed by Linderstrϕm-Lang in 1955.[554,556] He suggested that portions of protein molecules unfold sporadically to allow rapid exchange which can be catalyzed by H$^+$, HO$^-$, or other acids or bases. The pH dependence of exchange for a given amide NH can be described as follows:

$$k_{ex} = k_H [H^+] + k_{OH} [OH^-] + k_w \tag{3-56}$$

where k_H, k_{OH}, and k_w are rate constants for acid-catalyzed, base-catalyzed, and the very slow water-catalyzed exchange.[554,556,562] While many amide protons exchange rapidly, hydrogen-bonded NH protons in well-packed hydrophobic core regions exchange slowly,[554,563] sometimes remaining in the protein for years in a D_2O solution. Some unusually stable small proteins such as the seed protein **crambin** show little exchange. The $C^{\varepsilon 1}$ protons of histidine imidazoles also exchange slowly with D_2O from the medium. The exchange rates[529,564] are rapid at higher temperatures, with an average half-life of about 11 min at 65°C. Different residues may exchange at different rates. Binding of substrates or inhibitors can stabilize the protein, slowing all exchange rates of both amide and imidazole groups.

Solid-state NMR and other topics. Little can be said about these topics, but NMR measurement on solid crystalline materials is now practiced and is providing a wealth of information. It is being applied more often to biochemically related problems.[465,542,565 – 567]

NMR spectroscopy of nucleic acids is discussed briefly in Chapter 5. An important medical application of NMR is in **imaging**, a topic dealt with in Box 30-A.

J. The Protein Data Bank, Three-Dimensional Structures, and Computation

The x,y,z coordinates of all atoms in published, refined three-dimensional structures have been deposited in the Protein Data Bank (Table 3-4).[568–571] Many other related databases are available,[572] e.g., covering molecular modeling,[573] gene sequences, proteome data,[574] and much, much more. A good way to keep up to date is to read the "computer corner" in *Trends in Biochemical Sciences (TIBS).* Most databases can be reached on the World Wide Web.[572] A selected list is

TABLE 3-4
Selected World Wide Web Servers Related to Protein Structures and Sequences[a]

Question / area	Tool	Access / URL
Database search by comparison of 3D structures	Dali server	*http://www.embl-heidelberg.de/dali/dali.html*
Structural classification of proteins	SCOP	*http://www.bio.cam.ac.uk/scop/*
Summary and analysis of PDP structures	PDBSum	*http://www.biochem.ucl.ac.uk/bsm/pdbsum*
Retrieve 3D coordinates	Protein Data Bank	*http://www.rcsb.org/pdb/*
SWISS-PROT sequence database, Swissmodel homology modeling, etc.	ExPASy	*http://expasy.hcuge,ch/*
Molecular graphics viewer for PCs and workstations	RasMol	*http://www.bernstein-plus-sons.com/software/rasmol*
World Wide Web -Entrez and Molecular Modeling Database access		*http://www.ncbi.nlm.nih.gov*
Protein Science Kinemages	MAGE and PREKIN	*http://www.prosci.uci.edu/kinemages/ KinemageIndex.html*
Pedro's Biomolecular Research Tools		*http://www.fmi.ch/biology/research_tools.html*
Predict secondary structure from sequence	Predict Protein server	*http://www.sander.embl-heidelberg.de*
Browse databanks in molecular biology	SRS server	*http://www.embl-heidelberg/srs/srsc*
The Human Genome Database		*http://gdbwww.gdb.org/*
Image Library of Biological Macromolecules		*http://www.imb-jena.de/IMAGE.html*
Bacterial Nomenclature		*http://www.gbf-braunschweig.de/DSMZ/ bactnom/bactnam.htm*

[a] From Holm and Sander,[570] Hogue *et al.*,[573] Laskowski *et al.*,[580] and Walsh *et al.*[572]

given in Table 3-4. The widely used viewer called RasMol can be used with your PC or MacIntosh computer[575–576a] or UNIX workstation[569,575]. Another way to view macromolecules is to use the *Protein Science* Kinemages ("kinetic images") using the program MAGE.[577,578]

Although it is mentioned in a few places, this book does not begin to describe the rapid growth of computation in biochemistry, biophysics, and biology in general. Very fast methods of protein structure determination are being developed.[579,582] One of the major goals in current computation is to predict folding patterns of proteins from their sequences.[582,583] By comparing sequences we can often guess an approximate structure but accurate predictions are still not possible. Having a structure, we would like to predict properties and reactivities and to be able to guess how two more macromolecules interact to form macromolecular complexes. We would like to understand the complex chain of nonpolar and electrostatic interactions that underlie the fundamental properties of catalysis, movement, and responsiveness of organisms. Many computers and ingenious minds are working to help us match theory with reality in these areas. Read the current journals!

References

1. Strang, R. (1981) *Trends Biochem. Sci.* **6**, VII – VIII
2. Michaelis, L. (1922) *Die Wasserstoffionenkonzentration*, p. 48, Springer-Verlag, Berlin
3. Dixon, M., and Webb, E. C., eds. (1979) *Enzymes*, 3rd ed., pp.138 – 163, Academic Press, New York
4. Dryer, R. L., and Lata, G. F. (1989) *Experimental Biochemistry*, Oxford Univ. Press, New York
5. Kuramitsu, S., and Hamaguchi, K. (1980) *J. Biochem.* **87**, 1215 – 1219
6. Beynon, R. J., and Easterby, J. S. (1996) *Buffer Solutions: The Basics*, IRL Press, Oxford
7. Bates, R. G. (1964) *Determination of pH*, Wiley, New York
8. Stoll, V. S., and Blanchard, J. S. (1990) in *Guide to Protein Purification* (Deutscher, M. P., ed), pp. 24 – 38, Academic Press, San Diego, California
9. Ellis, K. J., and Morrison, J. F. (1982) *Methods Enzymol.* **87**, 405 – 426
10. Good, N. E., Winget, G. D., Winter, W., Connolly, T. N., Izawa, S., and Singh, R. M. M. (1966) *Biochemistry* **5**, 467 – 477
11. Ferguson, W. J., Braunschweiger, K. I., Braunschweiger, W. R., Smith, J. R., McCormick, J. J., Wasmann, C. C., Jarvis, N. P., Bell, D. H., and Good, N. E. (1980) *Anal. Biochem.* **104**, 300 – 310
12. Dawson, R. M. C., Elliot, D. C., Elliot, W. H., and Jones, K. M. (1986) *Data for Biochemical Research*, 3rd ed., Oxford Science Pub., London
13. Robyt, J. F., and White, B. J. (1990) *Biochemical Techniques, Theory and Practice*, 2nd ed., Waveland Press, Inc., Prospect Height, Ill
14. Cooper, T. G. (1977) *The Tools of Biochemistry*, Wiley, New York
15. Alexander, R. R., and Griffiths, J. M. (1992) *Basic Biochemical Methods*, 2nd ed., Wiley, New York
16. Wilson, K., and Walker, J., eds. (1994) *Principles and Techniques of Practical Biochemistry*, 4th ed., Cambridge Univ. Press, New York
17. Boyer, R. F. (1992) *Modern Experimental Biochemistry*, 2nd ed., Benjamin-Cummings Publ., Redwood City, California
18. le Maire, M., Chabaud, R., and Herve, G. (1991) *Laboratory Guide to Biochemistry, Enzymology, and Protein Physical Chemistry*, Plenum, New York
19. Fini, C., Floridi, A., and Finelli, V. N., eds. (1989) *Laboratory Methodology in Biochemistry*, CRC Press, Boca Raton, Florida
20. Scopes, R. K. (1993) *Protein Purification: Principles and Practice*, 3rd ed., Springer, New York
21. Harris, E. L. V., and Angal, S. (1990) *Protein Purification: A Practical Approach*, IRL Press, Oxford, 2 volumes
22. Deutscher, M. P., ed. (1990) *Guide to Protein Purification: Methods in Enzymology*, Vol. 182, Academic Press, San Diego, California
23. Celis, J. E., ed. (1994) *Cell Biology: A Laboratory Handbook*, Academic Press, New York (3 volumes)
24. Fasman, G. D., ed. (1989) *CRC Practical Handbook of Biochemistry and Molecular Biology*, CRC Press, Boca Raton, Florida
25. Rosenberg, I. M. (1996) *Protein Analysis and Purification*, Birkhäuser, Cambridge, Massachusetts
26. Walker, J. M., ed. (1996) *The Protein Protocols Handbook*, Humana Press, Totowa, New Jersey
27. Marshak, D. R., Kadonaga, J. T., Burgess, R. R., Knuth, M. W., Brennan, W. A., Jr., and Lin, S.-H. (1996) *Strategies for Protein Purification and Characterization*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
28. Creighton, T. E., ed. (1989) *Protein Structure: A Practical Approach*, IRL Press, Oxford
29. Marshak, D. R., Kadonaga, J. T., Burgess, R. R., Knuth, M. W., Brennan, W. A., and Lin, S.-H. (1995) *Strategies for Protein Purification and Characterization*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
30. Crabb, J. W., ed. (1995) *Techniques in Protein Chemistry*, Vol. VI, Academic Press, San Diego, California
31. Walker, J. M., ed. (1994) *Basic Protein and Peptide Protocols*, Humana Press, Totowa, New Jersey
32. Bollag, D. M., and Edelstein, S. J. (1990) *Protein Methods*, Wiley-Liss, New York
33. Creighton, T. E., ed. (1989) *Protein Structure and Protein Function: A Practical Approach*, IRL Press, Oxford
34. Copeland, R. A. (1994) *Methods for Protein Analysis. A Practical Guide to Laboratory Protocols*, Chapman and Hall, New York
35. Matsuidaira, P. T., ed. (1995) *A Practical Guide to Protein and Peptide Purification for Microsequencing*, Academic Press, San Diego, CA
36. Hugli, T. E., ed. (1989) *Techniques in Protein Chemistry*, Academic Press, New York
37. Means, G. E., and Feeney, R. E. (1973) *Chemical Modification of Proteins*, Holden-Day, San Francisco
38. Keleti, G., and Lederer, W. H. (1974) *Handbook of Micromethods for the Biological Sciences*, Van Nostrand-Reinhold, New York
39. Switzer, R. L., and Garrity, L. F. (1999) *Experimental Biochemistry:Theory and Exercises in Fundamental Methods*, 3rd ed., Freeman, New York
40. Schulz, G. E., and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York
41. Scopes, R. K. (1994) *Protein Purification. Principles of Protein Structure*, Springer-Verlag, New York
42. Villafranca, J., ed. (1990) *Current Research in Protein Chemistry: Techniques, Structure, and Function*, Academic Press, New York
43. Coligan, J., Dunn, B., Ploegh, H., Speicher, D., and Wingfield, P., eds. (1995) *Current Protocols in Protein Science*, Wiley, New York
44. Matsudaira, P. T., ed. (1989) *A Practical Guide to Protein and Peptide Purification for Microsequencing*, Academic Press, San Diego, California
45. Doonan, S., ed. (1996) *Protein Purification Protocols*, Humana Press, Totowa, New Jersey
46. Little, P., and Hughes, S., eds. (1992) *Technique*, Vol. 4, Academic Press, London
47. *Methods: A Companion to Methods in Enzymology*, 1990–present, Academic Press, San Diego, California
48. Wankat, P. C., Van Oss, C. J., and Henry, J. D., Jr., eds. (1992) *Separation and Purification Methods*, Dekker, New York
49. *protein expression and purification*, (1990 – present), Academic Press, San Diego, California
50. Coligan, J., Dunn, B., Ploegh, H., Speicher, D., and Wingfield, P., eds. (1996–present) *Current Protocols in Protein Science*, Wiley, New York
51. Hirs, C. H. W., and Timasheff, S. N., eds. (1983) *Methods in Enzymology, Enzyme Structure*, Vol. 91, Academic Press, New York
52. Doolittle, R. F., ed. (1990) *Methods in Enzymology, Molecular Evolution: Computer Analysis of Protein and Nucleic Acid*, Vol. 183, Academic Press, New York
53. Bollag, D. E., Rozycki, M. D., and Edelstein, S. J., eds. (1996) *Protein Methods*, 2nd ed., Wiley, New York
54. Birnie, G. D., and Fox, S. M., eds. (1969) *Subcellular Components*, Butterworth, London
55. Estabrook, R. W., and Pullman, M. E., eds. (1967) *Methods in Enzymology*, Vol. 10, Academic Press, New York
56. Kaback, H. R. (1971) in *Methods in Enzymology*, Vol. 22 (Jakoby, W. B., ed), pp. 99 – 120, Academic Press, New York
57. Dingle, J. T. (1972) *Lysosomes. A Laboratory Handbook*, North-Holland Publ., Amsterdam
58. Rasmussen, N. (1996) *Trends Biochem. Sci.* **21**, 319 – 321

58a. Jenkins, W. T. (1998) *Protein Sci.* **7**, 376 – 382

59. McPhie, P. (1971) in *Methods in Enzymology*, Vol. 22 (Jakoby, W. B., ed), pp. 23 – 32, Academic Press, New York
60. Blatt, W. F. (1971) in *Methods in Enzymology*, Vol. 22 (Jakoby, W. B., ed), pp. 39 – 49, Academic Press, New York
61. Porath, J., and Flodin, P. (1959) *Nature (London)* **183**, 1657 – 1659

62. Porath, J. (1978) *Trends Biochem. Sci.* **3**, N100
63. Lathe, G. H. (1978) *Trends Biochem. Sci.* **3**, N99 – N100
64. Flodin, P., and Aspberg, K. (1961) *Biol. Struct. Funct. Proc. IUB/IUBS Int. Sym*, 1st ed., Vol. 1, pp. 345-349 (p. 346)
65. Weir, M., and Sparks, J. (1987) *Biochem. J.* **245**, 85
66. Regnier, F. E. (1991) *Nature (London)* **350**, 634 – 635
67. Afeyan, N. B., Gordon, N. F., and Mazsaroff, I. (1990) *J. Chromatogr.* **519**, 1 – 29
67a. Wu, C.-s. (1999) *Column Handbook for Size Exclusion Chromatography*, Academic Press, San Diego, California
68. Liapis, A. I., and McCoy, M. A. (1992) *J. Chromatogr.* **599**, 87 – 104
69. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry, Part II*, Freeman, San Francisco, California (pp. 549 – 685)
70. Kyte, J. (1995) *Structure in Protein Chemistry*, Garland Publ., New York
71. Brakke, M. K. (1953) *Arch. Biochem. Biophys.* **45**, 275 – 290
72. Britten, R. J., and Roberts, R. B. (1960) *Science* **131**, 32 – 33
73. Martin, R. G., and Ames, B. N. (1961) *J. Biol. Chem.* **236**, 1372 – 1379
74. Meselson, M., Stahl, F. W., and Vinograd, J. (1957) *Proc. Natl. Acad. Sci. U.S.A.* **43**, 581 – 588
75. da Silva, J. R. R. F. (1978) in *New Trends in Bio-organic Chemistry* (Williams, R. J. P., and da Silva, J. R. R. F., eds), Academic Press, London
76. Arakawa, T., and Timasheff, S. N. (1982) *Biochemistry* **21**, 6545 – 6552
77. Rao, J. K. M., and Argos, P. (1981) *Biochemistry* **20**, 6536 – 6543
78. Green, A. A., and Hughes, W. L. (1955) in *Methods in Enzymology*, Vol. 1 (Colowick, S. P., and Kaplan, N. O., eds), pp. 67 – 90, Academic Press, New York
79. Dawson, R. M. C., Elliott, D. C., Elliott, W. H., and Jones, K. M. (1969) *Data for Biochemical Research*, 2nd ed., Clarendon Press, Oxford
80. Ito, Y. (1982) *Trends Biochem. Sci.* **7**, 47 – 50
81. Ito, Y. (1987) *Nature (London)* **326**, 419 – 420
82. Mandava, N. B., and Ito, Y., eds. (1988) *Countercurrent Chromatography*, Dekker, New York
82a. Menet, J.-M., and Thiébaut, D., eds. (1999) *Countercurrent Chromatography*, Dekker, New York
83. Er-el, Z., Zaidenzaig, Y., and Shaltiel, S. (1972) *Biochem. Biophys. Res. Commun.* **49**, 383 – 390
84. Arakawa, T., and Timasheff, S. N. (1982) *Biochemistry* **21**, 6536 – 6544
85. Atha, D. H., and Ingham, K. C. (1981) *J. Biol. Chem.* **256**, 12108 – 12117
86. Kenney, A., and Fowell, S., eds. (1992) *Practical Protein Chromatography*, Humana Press, Totowa, New Jersey
87. Tiselius, A., Hjerten, S., and Levin, O. (1956) *Arch. Biochem. Biophys.* **65**, 132
88. Wolf, W. J., and Sly, D. A. (1964) *J. Chromatogr.* **15**, 247 – 250
89. Regnier, F. E. (1987) *Science* **238**, 319 – 323
90. Hutchens, T. W., and Porath, J. (1987) *Biochemistry* **26**, 7199 – 7204
91. Molnár, I., Horváth, C., and Jatlow, P. (1978) *Chromatographia* **11**, 260 – 265
92. Hamilton, R. J., and Sewell, P. A. (1982) *Introduction to High Performance Liquid Chromatography*, 2nd ed., Chapman and Hall, London
93. Berridge, J. C. (1985) *Techniques for the Automated Optimization of HPLC Separations*, Wiley, New York
94. Simpson, C. F. (1982) *Techniques in Liquid Chromatography*, Wiley, New York
95. Mant, C. T., and Hodges, R. S. (1991) *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation*, CRC Press, Boca Raton, Florida
96. Gooding, K. M., and Regnier, F. E., eds. (1990) *HPLC of Biological Macromolecules*, Dekker, New York
96a. Beesley, T. E., Buglio, B., and Scott, R. P. W. (2000) *Quantitative Chromatographic Analysis*, Dekker, New York
97. Oliver, R. W. A., ed. (1989) *HPLC of Macromolecules: A Practical Approach*, IRL Press, Oxford
98. Fried, B., and Sherma, J., (1999) *Thin-Layer Chromatography*, 4th ed., Dekker, New York
99. Zlatkis, A., and Kaiser, R. E. (1977) *High Performance Thin-Layer Chromatography*, Elsevier Scientific-North Holland, Amsterdam
100. Sherma, J., and Fried, B., eds. (1996) *Practical Thin-Layer Chromatography: A Multidisciplinary Approach*, CRC Press, Boca Raton, Florida
101. Liu, M.-K., Li, P., and Giddings, J. C. (1993) *Protein Sci.* **2**, 1520 – 1531
102. Treadwell, G. E., Cairns, W. L., and Metzler, D. E. (1968) *J. Chromatogr.* **35**, 376 – 388
103. Novotny, M. V. (1989) *Science* **264**, 51 – 57
104. Schomvurg, G. (1990) *Gas Chromatography: A Practical Course*, VCH Publ., Cambridge, UK
105. Lee, M. L., and Markides, K. E. (1987) *Science* **235**, 1342 – 1347
105a. Muraviev, D., Gorshkov, V., and Warshawsky, A. (1999) *Ion Exchange*, Dekker, New York
106. Moore, S., and Stein, W. H. (1949) *Cold Spring Harbor Symp. Quant. Biol.* **14**, 179 – 190
107. Moore, S., and Stein, W. H. (1949) *J. Biol. Chem.* **178**, 53 – 77
108. Manning, J. M. (1993) *Protein Sci.* **2**, 1188 – 1191
109. Svasti, J. (1980) *Trends Biochem. Sci.* **5**, 8 – 9
110. Moore, S., and Stein, W. H. (1951) *J. Biol. Chem.* **192**, 663 – 681
111. Moore, S., and Stein, W. H. (1954) *J. Biol. Chem.* **211**, 893 – 906
112. Moore, S., and Stein, W. H. (1963) in *Methods in Enzymology*, Vol. 6 (Colowick, S. P., and Kaplan, N. O., eds), pp. 819 – 831, Academic Press, New York
113. Yamamoto, S., Nakanishi, K., and Matsuno, R. (1988) *Ion Exchange Chromatography of Proteins*, Dekker, New York
114. Peterson, E. A. (1970) *Cellulosic Ion Exchangers*, Elsevier-North Holland, Amsterdam
115. Nölting, B., Golbik, R., Neira, J. L., Soler-Gonzalez, A. S., Schreiber, G., and Fersht, A. R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 826 – 830
116. Lundell, N., and Markides, K. (1992) *Chromatographia* **34**, 369 – 375
117. Dean, P. D. G., Johnson, W. S., and Middle, F. A., eds. (1985) *Affinity Chromatography, A Practical Approach*, IRL Press, Oxford
118. Parikh, I., and Cuatrecasas, P. (1985) *Chem. Eng. News* **63 (**Aug 26), 17 – 32
119. Porath, J. (1968) *Nature (London)* **218**, 834 – 838
120. Inman, J. K. (1974) in *Methods in Enzymology*, Vol. 34 (Jakoby, W. B., and Wilchek, M., eds), pp. 30 – 58, Academic Press, New York
121. Bethell, G. S., Ayers, J. S., Hearn, M. T. W., and Hancock, W. S. (1981) *J. Chromatogr.* **219**, 361 – 372
122. Wilchek, M., and Miron, T. (1987) *Biochemistry* **26**, 2155 – 2161
123. Ashley, G. W., and Bartlett, P. A. (1984) *J. Biol. Chem.* **259**, 13615 – 13620
124. Hengen, P. N. (1995) *Trends Biochem. Sci.* **20**, 285 – 286
125. Loo, T. W., and Clarke, D. M. (1995) *J. Biol. Chem.* **270**, 21449 – 21452
126. Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R. A., Nordheim, A., and Stunnenberg, H. G. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8972 – 8976
127. Hochuli, E., Dobeli, H., and Schacher, A. (1987) *J. Chromatogr.* **411**, 177 – 184
128. Turner, A. J. (1981) *Trends Biochem. Sci.* **6**, 171 – 173
129. Vesterberg, O. (1971) in *Methods in Enzymology*, Vol. 22 (Jakoby, W. B., ed), pp. 389 – 412, Academic Press, New York
130. Bauer, J. (1994) *Cell Electrophoresis*, CRC Press, Boca Raton, Florida
131. Allison, S. A., and Tran, V. T. (1995) *Biophys. J.* **68**, 2261 – 2270
132. Shuster, L. (1971) in *Methods in Enzymology*, Vol. 22 (Jakoby, W. B., ed), pp. 412 – 433, Academic Press, New York
133. Chrambach, A., and Rodbard, D. (1971) *Science* **172**, 440 – 451
134. Andrews, A. T. (1986) *Electrophoresis: Theory, Techniques and Biochemical and Clinical Applications*, 2nd ed., Clarendon-Oxford Univ. Press, Oxford
135. Gordon, A. H. (1980) *Electrophoresis of Proteins in Polyacrylamide and Starch Gels*, Revised ed., Elsevier-North Holland, Amsterdam
136. Chrambach, A., Dunn, M. J., and Radola, B. J., eds. (1987) *Advances in Electrophoresis*, Vol. I, VCH Publ., Cambridge, UK
137. Serwer, P., Easom, R. A., Hayes, S. J., and Olson, M. S. (1989) *Trends Biochem. Sci.* **14**, 4 – 7
138. *electrophoresis*, (1995 – present), VCH Publ., New York
139. *advances in electrophoresis*, (1987 – present), VCH Publ., Cambridge
140. Hawcroft, D. M. (1997) *Electrophoresis: The Basics*, IRL Press, Oxford
141. Hengen, P. N. (1995) *Trends Biochem. Sci.* **20**, 202 – 203
142. Garfin, D. E. (1990) *Methods Enzymol.* **182**, 425 – 441
143. Hames, B. D., and Rickwood, D., eds. (1990) *Gel Electrophoresis of Proteins: A Practical Approach*, 2nd ed., IRL Press, Oxford
144. Righetti, P. G., ed. (1996) *Capillary Electrophoresis in Analytical Biotechnology*, CRC Press, Boca Raton, Florida
145. Camilleri, P., ed. (1993) *Capillary Electrophoresis: Theory and Practice*, CRC Press, Inc., Boca Raton
146. Altria, K. D., ed. (1996) *Capillary Electrophoresis Guidebook*, Humana Press, Totowa, New Jersey
147. Landers, J. P., ed. (1994) *Handbook of Capillary Electrophoresis*, CRC Press, Boca Raton, Florida
148. Landers, J. P. (1993) *Trends Biochem. Sci.* **18**, 409 – 414
149. Wehr, T., Rodríguez-Díaz, R., and Zhu, M. (1998) *Capillary Electrophoresis of Proteins*, Dekker, New York
149a. Yao, S., Anex, D. S., Caldwell, W. B., Arnold, D. W., Smith, K. B., and Schultz, P. G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5372 – 5377
150. Valaskovic, G. A., Kelleher, N. L., and McLafferty, F. W. (1996) *Science* **273**, 1199 – 1202
151. Gao, J., Gomez, F. A., Härter, R., and Whitesides, G. M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12027 – 12030
152. Harrison, D. J., Fluri, K., Seiler, K., Fan, Z., Effenhauser, C. S., and Manz, A. (1993) *Science* **261**, 895 – 897
153. Radola, B. J., and Graesslin, D., eds. (1977) *Electrofocusing and Isotachophoresis*, de Gruyter, Berlin
154. Bocek, P., Dem, M., Gebauer, P., and Dolnik, V. (1987) *Analytical Isotachophoresis*, VCH Publ., Cambridge, UK

References

155. Righetti, P. G. (1983) *Isoelectric Focusing: Theory, Methodology and Applications*, Elsevier - North Holland Biomedical Press, Amsterdam
156. Catsimpoolas, N., and Drysdale, J., eds. (1977) *Biological and Biomedical Applications of Isoelectric Focusing*, Plenum, New York
157. Egen, N. B., Bliss, M., Mayersohn, M., Owens, S. M., Arnold, L., and Bier, M. (1988) *Anal. Biochem.* **172**, 488–494
158. Bjellqvist, B., Ek, K., Righetti, P. G., Gianazza, E., Görg, A., Westermeier, R., and Postel, W. (1982) *J. Biochem. Biophys. Methods* **6**, 317–339
159. Righetti, P. G., Gianazza, E., and Gelfi, C. (1988) *Trends Biochem. Sci.* **13**, 335–338
160. Righetti, P. G. (1990) *Immobilized pH Gradients: Theory and Methodology*, Elsevier, Amsterdam
161. Whitney, J. B., III, Cobb, R. R., Popp, R. A., and O'Rourke, T. W. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7646–7650
162. Madden, M. S. (1995) *Anal. Biochem.* **229**, 203–206
163. Anderson, N. (1979) *Nature (London)* **278**, 122–123
164. Wade, N. (1981) *Science* **211**, 33–35
165. Henslee, J. G., and Srere, P. A. (1979) *J. Biol. Chem.* **254**, 5488–5491
166. Hanash, S. M., and Strahler, J. R. (1989) *Nature (London)* **337**, 485–486
167. Neukirchen, R. O., Schlosshauer, B., Baars, S., Jackle, H., and Schwarz, U. (1982) *J. Biol. Chem.* **257**, 15229–15234
168. Racine, R. R., and Langley, C. H. (1980) *Nature (London)* **283**, 855–857
169. Celis, J. E., and Bravo, R., eds. (1984) *Two-Dimensional Gel Electrophoresis of Proteins: Methods and Applications*, Academic Press, New York
170. Celis, J. E., and Bravo, R. (1981) *Trends Biochem. Sci.* **6**, 197–201
171. Garrels, J. I. (1989) *J. Biol. Chem.* **264**, 5269–5282
172. Krauss, M. R., Collins, P. J., and Blose, S. H. (1989) *Nature (London)* **337**, 669–670
173. Wood, T. D., Chen, L. H., Kelleher, N. L., Little, D. P., Kenyon, G. L., and McLafferty, F. W. (1995) *Biochemistry* **34**, 16251–16254
173a. McLafferty, F. W., Fridriksson, E. K., Horn, D. M., Lewis, M. A., and Zubarev, R. A. (1999) *Science* **284**, 1289 – 1290
174. Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko.A, Boucherie, H., and Mann, M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14440–14445
175. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) *Nature (London)* **379**, 466–469
176. Hess, D., Covey, T. C., Winz, R., Brownsey, R. W., and Aebersold, R. (1993) *Protein Sci.* **2**, 1342–1351
177. Van Holde, K. E. (1985) *Physical Biochemistry*, 2nd ed., Prentice-Hall, Engelwood Cliffs, New Jersey
178. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, Freeman, San Francisco, California
179. Svedburg, T., and Pederson, K. O. (1940) *The Ultracentrifuge*, Oxford Univ. Press, London
180. Schachman, H. K. (1989) *Nature (London)* **341**, 259–260
181. Schachman, H. K. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., eds), pp. 3–15, Royal Society for Chemistry, Cambridge, UK
182. Schachman, H. K. (1959) *Ultracentrifugation in Biochemistry*, Academic Press, New York
183. Rivas, G., and Minton, A. P. (1993) *Trends Biochem. Sci.* **18**, 284–287
184. Harding, S. E., Rowe, A. J., and Horton, J. C., eds. (1992) *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, CRC Press, Boca Raton, Florida
185. Hansen, J. C., Lebowitz, J., and Demeler, B. (1994) *Biochemistry* **33**, 13155–13163
186. Ralston, G. (1993) *Introduction to Analytical Ultracentrifugation*, Beckman Instruments, Fullerton, California
187. Stafford, W. F., and Schuster, T. M. (1995) in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (Glasel, J. A., and Deutscher, M. P., eds), pp. 111–145, Academic Press, San Diego
188. Bowen, T. J. (1970) *An Introduction to Ultracentrifugation*, Wiley (Interscience), New York
189. Bothwell, M. A., Howlett, G. J., and Schachman, H. K. (1978) *J. Biol. Chem.* **253**, 2073–2076
190. Pollet, R. J., Haase, B. A., and Standaert, M. L. (1979) *J. Biol. Chem.* **254**, 30–33
190a. *CRC Handbook of Biochemistry*, (1968) Chem. Rubber Publ. Co., Cleveland, Ohio
191. Jacobsen, M. P., Wills, P. R., and Winzor, D. J. (1996) *Biochemistry* **35**, 13173–13179
191a. Colowick, S. P., and Kaplan, N. O., eds., (1968) *Methods Enzymol.* **12B**, 388–389
192. Squire, P. G., and Himmel, M. E. (1979) *Arch. Biochem. Biophys.* **196**, 165–177
193. Andrews, P. (1965) *Biochem. J.* **96**, 595–606
194. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, Freeman, San Fransico, California (pp. 670–684)
195. Nakos, G., and Mortenson, L. (1971) *Biochemistry* **10**, 455–458
196. Mann, K. G., and Fish, W. W. (1972) *Methods Enzymol.* **26C**, 28–42
197. Igou, D. K., Lo, J. T., Clark, D. S., Mattice, W. L., and Younathan, E. S. (1974) *Biochem. Biophys. Res. Commun.* **60**, 140–145
198. Weber, K., Pringle, J. R., and Osborn, M. (1972) *Methods Enzymol.* **26C**, 3–29
199. Wood, H. G., Werkman, C. H., Hemingway, A., and Nier, A. O. (1941) *J. Biol. Chem.* **139**, 377–381
200. Biemann, K. (1995) *Protein Sci.* **4**, 1920–1927
200a. Niessen, W. M. A. (1998) *Liquid Chromatography - Mass Spectrometry*, 2nd ed., Dekker, New York
200b. Chapman, J. R., ed. (2000) *Mass Spectrometry of Proteins and Peptides*, Hamana Press, Totowa, New Jersey
200c. Burlingame, A. L., Carr, S. A., and Baldwin, M. A., eds. (2000) *Mass Spectrometry in Biology and Medicine*, Hamana Press, Totowa, New Jersey
201. Senko, M. W., and McLafferty, F. W. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 763–785
202. Li, Y., Hunter, R. L., and McIver, R. T., Jr. (1994) *Nature (London)* **370**, 393–395
203. Siuzdak, G. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11290–11297
204. Caprioli, R. M., and Suter, M. J.-F. (1995) in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (Glasel, J. A., and Deutscher, M. P., eds), pp. 147-204, Academic Press, San Diego
205. Burlingame, A. L., and Carr, S. A., eds. (1996) *Mass Spectrometry in the Biological Sciences*, Humana Press, Totowa, New Jersey
206. Karas, M., and Hillenkamp, F. (1988) *Anal. Chem.* **60**, 2299–2303
207. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., and Yoshida, T. (1988) *Rapid Commun. Mass Spectrom.* **2**, 151–156
208. Beavis, R. C., and Chait, B. T. (1989) *Rapid Commun. Mass Spectrom.* **3**, 233–237
209. Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T. (1991) *Anal. Chem.* **63**, 1193A–1198A
210. Smith, R. D., Loo, J. A., Edmonds, C. G., Barinaga, C. J., and Udseth, H. R. (1990) *Anal. Chem.* **62**, 882–899
211. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1990) *Mass Spectrom. Rev.* **9**, 37
212. Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D., and Alice, M. B. (1968) *J. Chem. Phys.* **49**, 2240
213. Mann, M., and Wilm, M. (1995) *Trends Biochem. Sci.* **20**, 219–224
214. Mann, M., Meng, C. K., and Fenn, J. B. (1989) *Anal. Chem.* **61**, 1702–1708
215. Covey, T. R., Bonner, R. F., Shushan, B. I., and Henion, J. (1988) *Rapid Commun. Mass Spectrom.* **2**, 249–256
216. Miranker, A., Kruppa, G. H., Robinson, C. V., Aplin, R. T., and Dobson, C. M. (1996) *J. Am. Chem. Soc.* **118**, 7402–7403
217. Marshall, A. G., Senko, M. W., Li, W., Li, M., Dillon, S., Guan, S., and Logan, T. M. (1997) *J. Am. Chem. Soc.* **119**, 433–434
218. Torchinsky, Y. M. (1981) *Sulfur in Proteins*, Pergamon, Oxford
219. Friedman, M. (1973) *The Chemistry and Biochemistry of the Sulfhydryl Group in Amino Acids, Peptides and Proteins*, Pergamon, Oxford
220. Jocelyn, P. C. (1972) *Biochemistry of the SH Group*, Academic Press, New York
221. Simpson, R. J., Neuberger, M. R., and Liu, T.-Y. (1976) *J. Biol. Chem.* **251**, 1936–1940
222. Hugli, T. E., and Moore, S. (1972) *J. Biol. Chem.* **247**, 2828–2834
223. Chin, C. C. Q., and Wold, F. (1974) *Anal. Biochem.* **61**, 379–391
224. Royer, G. P., Schwartz, W. E., and Liberatore, F. A. (1977) *Methods Enzymol.* **47E**, 40–45
225. Glazer, A. N. (1976) in *The Proteins*, 3rd ed., Vol. II (Neurath, H., and Hill, R. L., eds), pp. 1–103, Academic Press, New York
226. Plapp, B. V., Raftery, M. A., and Cole, R. D. (1967) *J. Biol. Chem.* **242**, 265–270
227. Doonan, S., and Fahmy, H. M. A. (1975) *Eur. J. Biochem.* **56**, 421–426
228. Mitchell, W. M. (1977) *Methods Enzymol.* **47E**, 165–169
229. Gadasi, H., Maruta, H., Collins, J. H., and Korn, E. D. (1979) *J. Biol. Chem.* **254**, 3631–3636
230. Drapeau, G. R. (1977) *Methods Enzymol.* **47E**, 189–191
231. Koida, M., and Walter, R. (1976) *J. Biol. Chem.* **251**, 7593–7599
232. Hayashi, R. (1977) *Methods Enzymol.* **47E**, 84–93
233. Krutzsch, H. C., and Pisano, J. J. (1977) *Methods Enzymol.* **47E**, 391–404
234. Mahoney, W. C., and Hermodson, M. A. (1979) *Biochemistry* **18**, 3810–3814
235. Hancock, W. S., ed. (1996) *New Methods in Peptide Mapping for the Characterization of Proteins*, CRC Press, Boca Raton, Florida
236. Lundblad, R. L. (1995) *Techniques in Protein Modification*, CRC Press, Boca Raton, Florida
237. Sanger, F. (1949) *Biochem. J.* **45**, 563–574
238. Sanger, F. (1988) *Ann. Rev. Biochem.* **57**, 1–28
239. Ruthen, R. (1993) *Sci. Am.* **269**(Oct), 30–31
240. Aebersold, R. H., Leavitt, J., Saavedra, R. A., Hood, L. E., and Kent, S. B. H. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6970–6974
241. Edman, P. (1956) *Acta Chem. Scand.* **10**, 761–768
242. Edman, P., and Begg, C. (1967) *Eur. J. Biochem.* **1**, 80–91
243. Niall, H. D. (1977) *Nature (London)* **268**, 279–280
244. Schlesinger, D. H., ed. (1988) *Macromolecular Sequencing and Synthesis*, Liss, New York
245. Shively, J. E., ed. (1986) *Methods of Protein Microcharacterization*, Humana Press, Clifton, New Jersey

References

246. Bhown, A. S. (1988) *Protein/Peptide Sequence Analysis: Current Methodoligies*, CRC Press, Boca Raton, Florida
247. Hunkapiller, M. W., and Hood, L. E. (1983) *Science* **219**, 650–659
248. Totty, N. F., Waterfield, M. D., and Hsuan, J. J. (1992) *Protein Sci.* **1**, 1215–1224
249. Gershoni, J. M. (1985) *Trends Biochem. Sci.* **10**, 103–106
250. Moos, M., Jr., Nguyen, N. Y., and Liu, T.-Y. (1988) *J. Biol. Chem.* **263**, 6005–6008
251. Baldo, B. A., and Tovey, E. R., eds. (1989) *Protein Blotting*, S. Karger, Farmington, CT
252. Wellner, D., Panneerselvam, C., and Horecker, B. L. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1947–1949
253. Bailey, J. M., Nikfarjam, F., Shenoy, N. R., and Shively, J. E. (1992) *Protein Sci.* **1**, 1622–1633
254. Bailey, J. M., Shenoy, N. R., Ronk, M., and Shively, J. E. (1992) *Protein Sci.* **1**, 68–80
255. Johnson, R. S., and Walsh, K. A. (1992) *Protein Sci.* **1**, 1083–1091
255a. Keough, T., Youngquist, R. S., and Lacey, M. P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7131–7136
256. Hunt, D. F., Shabanowitz, J., Yates, J. R., III, Zhu, N.-Z., Russell, D. H., and Castro, M. E. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 620–623
257. Takao, T., Hitouji, T., Shimonishi, Y., Tanabe, T., Inouye, S., and Inouye, M. (1984) *J. Biol. Chem.* **259**, 6105–6109
257a. Jaffe, H., Veeranna, and Pant, H. C. (1998) *Biochemistry* **37**, 16211–16224
257b. Whitelegge, J. P., Gundersen, C. B., and Faull, K. F. (1998) *Protein Sci.* **7**, 1423–1430
258. Yazdanparast, R., Andrews, P. C., Smith, D. L., and Dixon, J. E. (1987) *J. Biol. Chem.* **262**, 2507–2513
259. Brown, J. R., and Hartley, B. S. (1966) *Biochem. J.* **101**, 214–228
260. Svobodova, S., Hais, I. M., and Kostir, J. V. (1953) *Chem. Listy* **47**, 205
261. Chen, R. F., Smith, P. D., and Maly, M. (1978) *Arch. Biochem. Biophys.* **189**, 241–250
262. Lee, K. S., and Drescher, D. G. (1979) *J. Biol. Chem.* **254**, 6248–6251
263. Chen, R. F., Scott, C., and Trepman, E. (1979) *Biochim. Biophys. Acta.* **576**, 440–455
264. Lunte, S. M. *et al*. (1989) *Anal. Biochem.* **178**, 202–207
265. Sharma, Y., Rao, C. M., Rao, S. C., Krishna, A. G., Somasundaram, T., and Balasubramanian, D. (1989) *J. Biol. Chem.* **264**, 20923–20927
266. Wilson, C. M. (1983) *Methods Enzymol.* **91**, 236–247
267. Merril, C. R. (1990) *Methods Enzymol.* **182**, 477–488
268. Finn, F. M., and Hofmann, K. (1976) in *The Proteins*, 3rd ed., Vol. II (Neurath, H., and Hill, R. L., eds), pp. 105–253, Academic Press, New York
269. Burnette, W. N. (1981) *Anal. Biochem.* **112**, 195–203
270. Karey, K. P., and Sirbasku, D. A. (1989) *Anal. Biochem.* **178**, 255–259
271. Strickland, E. H., Billups, C., and Kay, E. (1972) *Biochemistry* **11**, 3657–3662
272. Metzler, D. E., Harris, C., Yang, I.-Y., Siano, D., and Thomson, J. A. (1972) *Biochem. Biophys. Res. Commun.* **46**, 1588–1597
273. Morton, R. A. (1975) *Biochemical Spectroscopy*, Wiley, New York
274. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci.* **4**, 2411–2423
275. Glazer, A. N., Delange, R. J., and Sigman, D. S. (1976) *Chemical Modification of Proteins: Selected Methods and Analytical Procedures*, Elsevier, New York
276. Feeney, R. E., Blankenhorn, G., and Dixon, H. B. F. (1975) *Advances in Protein Chemistry*, Vol. 29, Academic Press, New York (pp. 135–203)
277. Acharya, A. S., and Manjula, B. N. (1987) *Biochemistry* **26**, 3524–3530
278. Rypniewski, W. R., Holden, H. M., and Rayment, I. (1993) *Biochemistry* **32**, 9851–9858
279. Jentoft, N., and Dearborn, D. G. (1979) *J. Biol. Chem.* **254**, 4359–4365
280. Hennecke, M., and Plapp, B. N. (1984) *Anal. Biochem.* **136**, 110–118
281. Brocklehurst, K. (1979) *Int. J. Biochem* **10**, 259–274
282. Degani, Y., and Patchornik, A. (1974) *Biochemistry* **13**, 1–11
283. Riddles, P. W., Blakeley, R. L., and Zerner, B. (1983) *Methods Enzymol.* **91**, 49–60
284. Pande, C. S., Pelzig, M., and Glass, J. D. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 895–899
285. Fraker, P. J., and Speck, J. C., Jr. (1978) *Biochem. Biophys. Res. Commun.* **80**, 849–857
286. Tolan, D. R., Lambert, J. M., Boileau, G., Fanning, T. G., Kenny, J. W., Vassos, A., and Traut, R. R. (1980) *Anal. Biochem.* **103**, 101–109
287. Miles, E. W. (1977) *Methods Enzymol.* **47**, 431–442
288. Kawata, Y., Sakiyama, F., Hayashi, F., and Kyogoku, Y. (1990) *Eur. J. Biochem.* **187**, 255–262
289. Secundo, F., Carrea, G., D'Arrigo, P., and Servi, S. (1996) *Biochemistry* **35**, 9631–9636
290. Miura, S., Tomita, S., and Ichikawa, Y. (1991) *J. Biol. Chem.* **266**, 19212–19216
291. Colman, R. F. (1983) *Ann. Rev. Biochem.* **52**, 67–91
292. Chowdhry, V., and Westheimer, F. H. (1979) *Ann. Rev. Biochem.* **48**, 293 – 325
293. Jelenc, P. C., Cantor, C. R., and Simon, S. R. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3564–3568
294. Simmonds, R. J. (1992) *Chemistry of Biomolecules*, CRC Press, Boca Raton, Florida
295. Bodanszky, M., Bodanszky, A., and Trost, B. M., eds. (1994) *The Practice of Peptide Synthesis*, 2nd ed., Springer-Verlag, New York
296. Lloyd-Williams, P., Albericio, F., and Giralt, E. (1997) *Chemical Approaches to the Synthesis of Peptides and Proteins*, CRC Press, Boca Raton, Florida
297. Kent, S. B. H. (1988) *Ann. Rev. Biochem.* **57**, 957–989
298. Kaiser, E. T., Mihara, H., Laforet, G. A., Kelly, J. W., Walters, L., Findeis, M. A., and Sasaki, T. (1989) *Science* **243**, 187–192
299. Merrifield, B. (1986) *Science* **232**, 341–347
300. Atherton, E., Logan, C. J., and Sheppard, R. C. (1981) *J. Chem. Soc., Perkin, I.*, 538–546
301. Atherton, E., and Sheppard, R. C. (1985) *J. Chem. Soc. Chem. Commun.*, 165–166
301a. Miller, S. C., and Scanlan, T. S. (1998) *J. Am. Chem. Soc.* **120**, 2690–2691
302. You-shang, Z. (1983) *Trends Biochem. Sci.* **8**, 16–17
303. Kullman, W. (1987) *Enzymatic Peptide Synthesis*, CRC Press, Boca Raon, Florida
304. Liu, C.-F., and Tam, J. P. (1994) *J. Am. Chem. Soc.* **116**, 4149–4153
305. Tam, J. P., Lu, Y.-A., Liu, C.-F., and Shao, J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12485–12489
306. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) *Science* **244**, 182–188
307. Borman, S. (1996) *Chem. Eng. News* **74** (Feb 12), 29–35
308. Youngquist, R. S., Fuentes, G. R., Lacey, M. P., and Keough, T. (1995) *J. Am. Chem. Soc.* **117**, 3900–3906
309. Cortese, R., ed. (1996) *Combinatorial Libraries: Synthesis, Screening and Application Potential*, Walter de Gruyter & Co., New York
310. Ellman, J., Stoddard, B., and Wells, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2779–2782
311. Kishchenko, G., Batliwala, H., and Makowski, L. (1994) *J. Mol. Biol.* **241**, 208–213
312. Scott, J. K. (1992) *Trends Biochem. Sci.* **17**, 241–245
312a. Barbas, C. F., Burton, D. B., Scott, J. K., and Silverman, G. J. (2000) *Phage Display: A Laboratory Manual* Cold Spring Harbor Press, Cold Spring Harbor, New York
313. Marvin, D. A., (1998) *Current Opinions in Structural Biology* **8**, 150–158
313a. Barbas, C. F., Burton, D. B., Scott, J. K., and Silverman, G. J. (2000) *Phage Display: A Laboratory Manual* Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
314. McConnell, S. J., and Hoess, R. H. (1995) *J. Mol. Biol.* **250**, 460–470
315. Katz, B. A. (1995) *Biochemistry* **34**, 15421–15429
316. Allen, J. B., Walberg, M. W., Edwards, M. C., and Elledge, S. J. (1995) *Trends Biochem. Sci.* **20**, 511–516
317. Gates, C. M., Stemmer, W. P. C., Kaptein, R., and Schatz, P. J. (1996) *J. Mol. Biol.* **255**, 373–386
318. Bastos, M., Maeji, N. J., and Abeles, R. H. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6738–6742
319. Han, H., Wolfe, M. M., Brenner, S., and Janda, K. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6419–6423
320. Chu, Y.-H., Dunayevskiy, Y. M., Kirby, D. P., Vouros, P., and Karger, B. L. (1996) *J. Am. Chem. Soc.* **118**, 7827–7835
321. Déprez, B., Williard, X., Bourel, L., Coste, H., Hyafil, F., and Tartar, A. (1995) *J. Am. Chem. Soc.* **117**, 5405–5406
322. Kim, R. M., Manna, M., Hutchins, S. M., Griffin, P. R., Yates, N. A., Bernick, A. M., and Chapman, K. T. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10012–10017
323. Sundberg, S. A., Barrett, R. W., Pirrung, M., Lu, A. L., Kiangsoontra, B., and Holmes, C. P. (1995) *J. Am. Chem. Soc.* **117**, 12050–12057
324. Baum, R. M. (1991) *Chem. Eng. News* **69**, 21–22
325. Slayter, E. M., and Slayter, H. S. (1992) *Light and Electron Microscopy*, Cambridge Univ. Press, New York
326. Bracegirdle, B. (1989) *Trends Biochem. Sci.* **14**, 464–468
327. Matthews, C. K., and van Holde, K. E. (1996) *Biochemistry*, 2nd ed., Benjamin, Menlo Park, California
328. Lichtman, J. W. (1994) *Sci. Am.* **271**(Aug), 40–45
329. Shotton, D., and White, N. (1989) *Trends Biochem. Sci.* **14**, 435–439
330. Boyde, A. (1985) *Science* **230**, 1270–1276
331. Gura, T. (1997) *Science* **276**, 1988–1990
332. Bennett, B. D., Jetton, T. L., Ying, G., Magnuson, M. A., and Piston, D. W. (1996) *J. Biol. Chem.* **271**, 3647–3651
333. Pennisi, E. (1997) *Science* **275**, 480–481
334. Maiti, S., Shear, J. B., Williams, R. M., Zipfel, W. R., and Webb, W. W. (1997) *Science* **275**, 530–532
335. Pool, R. (1988) *Science* **241**, 25–26
336. Betzig, E., and Chichester, R. J. (1993) *Science* **262**, 1422–1425
337. Dykstra, M. J. (1992) *Biological Electron Microscopy*, Plenum, New York
338. Hand, A. R. (1995) in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (Glasel, J. A., and Deutscher, M. P., eds), pp. 205-260, Academic Press, San Diego
339. Nellist, P. D., McCallum, B. C., and Rodenburg, J. M. (1995) *Nature (London)* **374**, 630–632

References

340. Schliwa, M. (1997) *Nature (London)* **387**, 764
341. Bubel, A., ed. (1989) *Microstructure and Function of Cells*, 1st ed., Ellis Harwood, New York
342. Penman, S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5251–5257
343. Morel, G., ed. (1995) *Visualization of Nucleic Acids*, CRC Press, Boca Raton, Florida
344. Caspar, D. L. D., and DeRosier, D. J. (1982) *Science* **218**, 653–655
345. Frank, J., ed. (1992) *Electron Tomography: Three-Dimensional Imaging with the Transmission Electron Microscope*, Plenum, New York
346. Skoglund, U., and Daneholt, B. (1986) *Trends Biochem. Sci.* **11**, 499–503
347. Lambert, O., Boisset, N., Taveau, J.-C., and Lamy, J. N. (1994) *J. Mol. Biol.* **244**, 640–647
348. Taveau, J.-C., Boisset, N., Lamy, J., Lambert, O., and Lamy, J. N. (1997) *J. Mol. Biol.* **266**, 1002–1015
349. De Rosier, D. J. (1997) *Nature (London)* **386**, 26–27
350. Dorset, D. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1791–1794
351. Binnig, G., and Rohrer, H. (1987) *Rev. Mod. Phys.* **59**, 615–625
352. Binnig, G., and Rohrer, H. (1985) *Sci. Am.* **253**, 50–56
353. Wiesendanger, R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12749–12750
354. Wickramasinghe, H. K. (1989) *Sci. Am.* **261**(Oct), 98–105
355. Hansma, P. K., Elings, V. B., Marti, O., and Bracker, C. E. (1988) *Science* **242**, 209–216
356. Pool, R. (1990) *Science* **247**, 634–636
357. Edstrom, R. D., Yang, X., Lee, G., and Evans, D. F. (1990) *FASEB J.* **4**, 3144–3151
358. Engel, A. (1991) *Ann. Rev. Biophys. Biophys. Chem.* **20**, 79–108
359. Amato, I. (1997) *Science* **276**, 1982–1985
360. Prater, C. B., Butt, H. J., and Hansma, P. K. (1990) *Nature (London)* **345**, 839–840
361. Hansma, H. G., and Hoh, J. H. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 115–139
362. Radmacher, M., Tillmann, R. W., Fritz, M., and Gaub, H. E. (1992) *Science* **257**, 1900–1905
363. Karrasch, S., Hegerl, R., Hoh, J. H., Baumeister, W., and Engel, A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 836–838
364. Mou, J., Yang, J., and Shao, Z. (1995) *J. Mol. Biol.* **248**, 507–512
365. Yip, C. M., and Ward, M. D. (1996) *Biophys. J.* **71**, 1071–1078
366. Mou, J., Sheng, S., Ho, R., and Shao, Z. (1996) *Biophys. J.* **71**, 2213–2221
367. Kuznetsov, YuG., Malkin, A. J., Land, T. A., DeYoreo, J. J., Barba, A. P., Konnert, J., and McPherson, A. (1997) *Biophys. J.* **72**, 2357–2364

367a. Czajkowsky, D. M., and Shao, Z. (1998) *FEBS Letters* **430**, 51– 54

368. Noy, A., Frisbie, C. D., Rozsnyai, L. F., Wrighton, M. S., and Lieber, C. M. (1995) *J. Am. Chem. Soc.* **117**, 7943–7951
369. Boland, T., and Ratner, B. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5297–5301
370. Miles, M. (1997) *Science* **277**, 1845–1847
371. Hansma, H. G., Laney, D. E., Bezanilla, M., Sinsheimer, R. L., and Hansma, P. K. (1995) *Biophys. J.* **68**, 1672–1677

371a. Fisher, T. E., Oberhauser, A. F., Carrion-Vazquez, M., Marszalek, P. E., and Fernandez, J. M. (1999) *Trends Biochem. Sci.* **24**, 379 –384

372. Blundell, T. L., and Johnson, L. N. (1976) *Protein Crystallography*, Academic Press, New York
373. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, Freeman, San Francisco, California (pp. 687–811)
374. McPherson, A. (1982) *Preparation and Analysis of Protein Crystals*, Wiley, New York
375. Wyckoff, H. W., Hirs, C. H. W., and Timasheff, S. N., eds. (1985) *Methods in Enzymology*, Vol. 114,115, Academic Press, Orlando, Florida
376. Allewell, N. M., and Trikha, J. (1995) in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (Glasel, J. A., and Deutscher, M. P., eds), pp. 381–431, Academic Press, San Diego
377. Verlinde, C. L. M. J., Merritt, E. A., Van Den Akker, F., Kim, H., Feil, I., Delboni, L. F., Mande, S. C., Sarfaty, S., Petra, P. H., and Hol, W. G. J. (1994) *Protein Sci.* **3**, 1670–1686
378. Woolfson, N. M. (1997) *An Introduction to X-ray Crystallography*, 2nd ed., Cambridge Univ. Press, New York
379. Thomas, J. M. (1993) *Nature (London)* **364**, 478–482
380. Dounce, A. L., and Allen, P. Z. (1988) *Trends Biochem. Sci.* **13**, 317–320
381. Bernal, J. D., and Crowfoot, D. (1934) *Nature (London)* **133**, 794–795
382. Glusker, J. P. (1994) *Protein Sci.* **3**, 2465-2469
383. Perutz, M. F., Muirhead, H., Cox, J. M., Goaman, L. G. G., Mathews, F. S., McGandy, E. L., and Webb, L. E. (1968) *Nature (London)* **219**, 29–32
384. Perutz, M. (1992) *Protein Structure: New Approaches to Disease and Therapy*, W.H. Freeman, San Fransico and New York
385. Eisenberg, D. (1994) *Protein Sci.* **3**, 1625–1628
386. Perutz, M. F. (1978) *Sci. Am.* **239**(Dec), 92–125
387. Tegze, M., and Faigel, G. (1996) *Nature (London)* **380**, 49–51
388. Fadley, C. S., and Len, P. M. (1996) *Nature (London)* **380**, 27–28
389. Rhodes, G. (1993) *Crystallography Made Crystal Clear*, Academic Press, New York

389a. Glusker, J. P., and Trueblood, K. N. (1985) *Crystal Structure Analysis A Primer*, 2nd ed., Oxford Univ. Press, Oxford

389b. Drenth, J. (1994) *Principles of Protein X-ray Crystallography*, Springler-Verlag, New York

389c. McRee, D. E. (1999) *Practical Protein Crystallography*, Academic Press, San Diego, California

390. Sweet, R. M. (1985) *Methods Enzymol.* **114**, 19–46
391. Kendrew, J. C. (1963) *Science* **139**, 1259–1266
392. Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., and Shore, V. C. (1960) *Nature (London)* **185**, 422–427
393. Rossman, M. G. (1994) *Protein Sci.* **3**, 1731–1733
394. Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1965) *Nature (London)* **206**, 757–761
395. Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L., and Thornton, J. M. (1997) *Trends Biochem. Sci.* **22**, 488–490
396. Gerstein, M., and Levitt, M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11911–11916
397. Marvin, D. A., and Nave, C. (1982) in *Structural Molecular Biology* (Davies, D. B., Saenger, W., and Danyluk, S. S., eds), pp. 3–44, Plenum, New York
398. Ducruix, A., and Giegé, R., eds. (1999) *Crystallization of Nucleic Acids and Proteins*, 2nd ed., IRL Press, Oxford
399. McPherson, A., Koszelak, S., Axelrod, H., Day, J., Williams, R., Robinson, L., McGrath, M., and Cascio, D. (1986) *J. Biol. Chem.* **261**, 1969–1975
400. McPherson, A., Jr. (1976) *J. Biol. Chem.* **251**, 6300–6303
401. Ray, W. J., Jr., and Puvathingal, J. M. (1986) *J. Biol. Chem.* **261**, 11544–11549
402. Day, J., and McPherson, A. (1992) *Protein Sci.* **1**, 1254–1268
403. Koszelak, S., Day, J., Leja, C., Cudney, R., and McPherson, A. (1995) *Biophys. J.* **69**, 13–19
404. Service, R. F. (1999) *Science* **285**, 1342–1346
405. Knäblein, J., Neuefeind, T., Schneider, F., Bergner, A., Messerschmidt, A., Löwe, J., Steipe, B., and Huber, R. (1997) *J. Mol. Biol.* **270**, 1–7
406. Glusker, J. P. (1984) *Trends Biochem. Sci.* **9**, 328–330
407. Guss, J. M., Merritt, E. A., Phizackerley, R. P., Hedman, B., Murata, M., Hodgson, K. O., and Freeman, H. C. (1988) *Science* **241**, 806–811
408. Walter, R. L., Ealick, S. E., Friedman, A. M., Blake, R. C., II, Proctor, P., and Shoham, M. (1996) *J. Mol. Biol.* **263**, 730–751
409. Budisa, N., Karnbrock, W., Steinbacher, S., Humm, A., Prade, L., Neuefeind, T., Moroder, L., and Huber, R. (1997) *J. Mol. Biol.* **270**, 616–623
410. Longhi, S., Czjzek, M., Lamzin, V., Nicolas, A., and Cambillau, C. (1997) *J. Mol. Biol.* **268**, 779–799

410a. Yamano, A., Heo, N.-H., and Teeter, M. M. (1997) *J. Biol. Chem.* **272**, 9597–9600

410b. Lamzin, V. S., Morris, R. J., Dauter, Z., Wilson, K. S., and Teeter, M. M. (1999) *J. Biol. Chem.* **274**, 20753–20755

411. Rhee, S., Silva, M. M., Hyde, C. C., Rogers, P. H., Metzler, C. M., Metzler, D. E., and Arnone, A. (1997) *J. Biol. Chem.* **272**, 17293–17302
412. Frauenfelder, H., Petsko, G. A., and Tsernoglou, D. (1979) *Nature (London)* **280**, 558–563
413. Arnone, A., Rogers, P. H., Hyde, C. C., Briley, P. D., Metzler, C. M., and Metzler, D. E. (1985) in *Transaminases* (Christen, P., and Metzler, D. E., eds), pp. 138–155, Wiley, New York
414. Wlodawer, A. (1982) *Prog. Biophys. and Mol. Biol.* **40**, 115–159
415. Kossiakoff, A. A. (1985) *Ann. Rev. Biochem.* **54**, 1195–1227
416. Butler, D. (1994) *Nature (London)* **371**, 469
417. Hellemans, A. (1997) *Science* **277**, 1214–1215
418. Glanz, J. (1995) *Science* **267**, 1904–1906
419. Pool, R. (1988) *Science* **241**, 295
420. Eisenberger, P., and Suckewer, S. (1996) *Science* **274**, 201–202
421. Neutze, R., and Hajdu, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5651–5655
422. Hajdu, J., and Johnson, L. N. (1990) *Biochemistry* **29**, 1669–1678
423. Duke, E. M. H., Wakatsuki, S., Hadfield, A., and Johnson, L. N. (1994) *Protein Sci.* **3**, 1178–1196
424. Jacoby, M. (1997) *Chem. Eng. News* **Dec 8**, 5
425. Rischel, C., Rousse, A., Uschmann, I., Albouy, P.-A., Geindre, J.-P., Audebert, P., Gauthier, J.-C., Förster, E., Martin, J.-L., and Antonetti, A. (1997) *Nature (London)* **390**, 490–492
426. Rattle, H. (1995) *An NMR Primer for Life Scientists*, Partnership Press, Tuckerton, NJ
427. Roberts, G. C. K., ed. (1993) *NMR of Macromolecules. A Practical Approach*, IRL Press, Oxford
428. Evans, J. N. S. (1995) *Biomolecular NMR Spectroscopy*, Oxford University Press, New York
429. Olson, D. L., Peck, T. L., Webb, A. G., Magin, R. L., and Sweedler, J. V. (1995) *Science* **270**, 1967–1969
430. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York
431. Nageswara Rao, B. D., and Kemple, M. D., eds. (1996) *NMR as a structural tool for macromolecules: Current status and future directions*, Plenum, New York
432. Bovey, F. A. (1969) *Nuclear Magnetic Resonance Spectroscopy*, Academic Press, New York

References

433. Sanders, J. K. M., and Hunter, B. K. (1993) *Modern NMR Spectroscopy*, 2nd ed., Oxford Univ. Press, London
434. Dybowski, C., and Lichter, R. L., eds. (1987) *NMR Spectroscopy Techniques*, Vol. 5, Dekker, New York
435. Bangerter, B. W. (1995) in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (Glasel, J. A., and Deutscher, M. P., eds), pp. 317-379, Academic Press, San Diego
436. Freeman, R. (1997) *A Handbook of Nuclear Magnetic Resonance*, 2nd ed., Addison Wesley Longman Limited, Essex, England
437. Emsley, J. W., ed. (1996) *Encyclopedia of NMR*, John Wiley & Sons, Ltd., Sussex, England
438. Markley, J. L. (1975) *Acc. Chem. Res.* **8**, 70–80
439. Perkins, S. J. (1982) in *Biological Magnetic Resonance*, Vol. 4 (Berliner, L. J., and Reuben, J., eds), pp. 193-334, Plenum Press, New York and London
440. Osapay, K., and Case, D. A. (1991) *J. Am. Chem. Soc.* **113**, 9436–9444
441. Szilágyi, L. (1995) *Prog. Nucl. Magn. Reson. Spectrosc.* **27**, Part 4, 326–443
442. Williamson, M. P., Kikuchi, J., and Asakura, T. (1995) *J. Mol. Biol.* **247**, 541–546
443. Sitkoff, D., and Case, D. A. (1997) *J. Am. Chem. Soc.* **119**, 12262–12273
444. Gu, Z., Zambrano, R., and McDermott, A. (1994) *J. Am. Chem. Soc.* **116**, 6368–6372
445. Gu, Z., Ridenour, C. F., Bronnimann, C. E., Iwashita, T., and McDermott, A. (1996) *J. Am. Chem. Soc.* **118**, 822–829
446. de Dios, A. C., Pearson, J. G., and Oldfield, E. (1993) *Science* **260**, 1491–1496
447. Sulzbach, H. M., Schleyer, P. V. R., and Schaefer, H. F., III (1994) *J. Am. Chem. Soc.* **116**, 3967–3972
448. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry* **31**, 1647–1651
449. Linderstrøm-Lang, K. (1955) *Chemical Society Special Publication*, No. 2, 1–20
450. Barfield, M., and Karplus, M. (1969) *J. Am. Chem. Soc.* **91**, 1–16
451. Karplus, M. (1963) *J. Am. Chem. Soc.* **85**, 2870–2871
452. Bax, A. (1989) *Ann. Rev. Biochem.* **58**, 223–256
453. Wang, A. C., and Bax, A. (1996) *J. Am. Chem. Soc.* **118**, 2483–2494
454. Stonehouse, J., Adell, P., Keeler, J., and Shaka, A. J. (1994) *J. Am. Chem. Soc.* **116**, 6037–6038
455. Aguayo, J. B., Gamcsik, M. P., and Dick, J. D. (1988) *J. Biol. Chem.* **263**, 19552–19557
456. Ackerman, J. J. H., Ewy, C. S., Kim, S.-G., and Shalwitz, R. A. (1987) *Ann. N.Y. Acad. Sci.* **508**, 89–98
457. Breitmaier, E., and Voelter, W. (1987) *Carbon-13 NMR Spectroscopy*, 3rd ed., VCH Publishers, New York
458. Beckmann, N. (1995) *Carbon-13 NMR Spectroscopy of Biological Systems*, Academic Press, New York
459. Gronenborn, A. M., Clore, G. M., Schmeissner, U., and Wingfield, P. T. (1986) *Eur. J. Biochem.* **161**, 37–43
460. Clore, G. M., Bax, A., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) *Biochemistry* **29**, 8172–8184
461. Jue, T., Rothman, D. L., Shulman, G. I., Tavitian, B. A., DeFronzo, R. A., and Shulman, R. G. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4489–4491
462. Markley, J. L. (1989) *Methods Enzymol.* **176**, 12–64
463. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) *Protein Sci.* **2**, 543–558
464. Rajagopal, P., Waygood, E. B., and Klevit, R. E. (1994) *Biochemistry* **33**, 15271–15282
465. Solum, M. S., Altmann, K. L., Strohmeier, M., Berges, D. A., Zhang, Y., Facelli, J. C., Pugmire, R. J., and Grant, D. M. (1997) *J. Am. Chem. Soc.* **119**, 9804–9809
466. Verkade, J. G., and Quin, L. D., eds. (1987) *Phosphorus-31 NMR Spectroscopy in Stereochemical Analysis*, VCH Publishers, Deerfield Beach, Florida
467. Ugurbil, K., Kingsley-Hickman, P. B., Sako, E. Y., Zimmer, S., Mohanakrishnan, P., Robitaille, P. M. L., Thoma, W. J., Johnson, A., Foker, J. E., From, A. H. L. (1987) *Ann. N.Y. Acad. Sci.* **508**, 265–286
468. Feeney, J., McCormick, J. E., Bauer, C. J., Birdsall, B., Moody, C. M., Starkmann, B. A., Young, D. W., Francis, P., Havlin, R. H., Arnold, W. D., and Oldfield, E. (1996) *J. Am. Chem. Soc.* **118**, 8700–8706
469. Lau, E. Y., and Gerig, J. T. (1997) *Biophys. J.* **73**, 1579–1592
470. Jarema, M. A. C., Lu, P., and Miller, J. H. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2707–2711
471. Slebe, J. C., and Martinez-Carrion, M. (1990) *J. Biol. Chem.* **253**, 2093–2097
472. Briley, P. A., Eisenthal, R., Harrison, R., and Smith, G. D. (1990) *Biochem. J.* **163**, 325–331
473. Roberts, J. K. M., Wade-Jardetzky, N., and Jardetzky, O. (1981) *Biochemistry* **20**, 5389–5394
474. Deutsch, C. J., and Taylor, J. S. (1987) *Ann. N.Y. Acad. Sci.* **508**, 33
475. Gutowsky, H. S., McCall, D. W., McGarvey, B. R., and Meyer, L. H. (1952) *J. Am. Chem. Soc.* **74**, 4809–4817
476. Chang, Y. C., and Graves, D. J. (1985) *J. Biol. Chem.* **260**, 2709–2714
477. Scott, R. D., Chang, Y.-C., Graves, D. J., and Metzler, D. E. (1985) *Biochemistry* **24**, 7668–7681
478. Saunders, M., Jiménez-Vázquez, H. A., Cross, R. J., Mroczkowski, S., Freedberg, D. I., and Anet, F. A. L. (1994) *Nature (London)* **367**, 256–257
479. Adebodun, F., and Jordan, F. (1988) *J. Am. Chem. Soc.* **110**, 309–310
480. Springer, C. S., Jr. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 375–399
481. Wittenberg, B. A., and Gupta, R. K. (1985) *J. Biol. Chem.* **260**, 2031–2034
482. Jelicks, L. A., and Gupta, R. K. (1989) *J. Biol. Chem.* **264**, 15230–15235
483. Price, W. S., Ge, N.-H., Hong, L.-Z., and Hwang, L.-P. (1993) *J. Am. Chem. Soc.* **115**, 1095–1105
484. Wang, S. M., and Gilpin, R. K. (1983) *Anal. Chem.* **55**, 493–497
485. Markham, G. D. (1986) *J. Biol. Chem.* **261**, 1507–1509
486. Denisov, V. P., and Halle, B. (1995) *J. Mol. Biol.* **245**, 682–687
487. Mullen, G. P., Dunlap, R. B., and Odom, J. D. (1986) *Biochemistry* **25**, 5625–5632
488. Aldhous, P. (1991) *Nature (London)* **353**, 689
489. Becker, E. D., and Farrar, T. C. (1972) *Science* **178**, 361–368
490. Bax, A., and Lerner, L. (1986) *Science* **232**, 960–967
491. Yamazaki, T., Muhandiram, R., and Kay, L. E. (1994) *J. Am. Chem. Soc.* **116**, 8266–8278
492. Farrar, T. C., and Becker, E. D. (1971) *Pulse and Fourier Transform NMR*, Academic Press, New York
493. Wüthrich, K. (1989) *Science* **243**, 45–50
494. van de Ven, F. J. M., and Hilbers, C. W. (1986) *J. Mol. Biol.* **192**, p. 419-441
495. Baniak, E. L., Rivier, J. E., Struthers, R. S., Hagler, A. T., and Gierasch, L. M. (1987) *Biochemistry* **26**, 2642–2656
496. Wüthrich, K. (1990) *J. Biol. Chem.* **265**, 22059–22062
497. Bax, A., and Grzesiek, S. (1993) in *NMR of Proteins* (Clore, G. M., and Gronenborn, A. M., eds), pp. 33–52, CRC Press, Boca Raton, Florida
498. Clore, G. M., and Gronenborn, A. M. (1994) *Protein Sci.* **3**, 372–390
499. Hoitnik, C. W. G., Driscoll, P. C., Hill, H. A. O., and Canters, G. W. (1994) *Biochemistry* **33**, 3560–3571
500. Yamazaki, T., Tochio, H., Furui, J., Aimoto, S., and Kyogoku, Y. (1997) *J. Am. Chem. Soc.* **119**, 872–880
501. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994) *J. Am. Chem. Soc.* **116**, 11655–11666
502. Hu, J.-S., and Bax, A. (1997) *J. Am. Chem. Soc.* **119**, 6360–6368
503. Song, S., Velde, D. V., Gunn, C. W., and Himes, R. H. (1994) *Biochemistry* **33**, 693–698
504. Barsukov, I. L., Lian, L.-Y., Ellis, J., Sze, K.-H., Shaw, W. V., and Roberts, G. C. K. (1996) *J. Mol. Biol.* **262**, 543–558
505. Wu, W.-J., Anderson, V. E., Raleigh, D. P., and Tonge, P. J. (1997) *Biochemistry* **36**, 2211–2220
506. Metzler, W. J., Hare, D. R., and Pardi, A. (1989) *Biochemistry* **28**, 7045–7052
507. McDowell, L. M., Lee, M., McKay, R. A., Anderson, K. S., and Schaefer, J. (1996) *Biochemistry* **35**, 3328–3334
508. Pflugrath, J. W., Wiegand, G., Huber, R., and Vértesy, L. (1986) *J. Mol. Biol.* **189**, 383–386
509. Kline, A. D., Braun, W., and Wüthrich, K. (1986) *J. Mol. Biol.* **189**, 377–382
510. McDonald, C. C., and Phillips, W. D. (1969) *J. Am. Chem. Soc.* **91**, 1513–1521
511. Nietlispach, D., Clowes, R. T., Broadhurst, R. W., Ito, Y., Keeler, J., Kelly, M., Ashurst, J., Oschkinat, H., Domaille, P. J., and Laue, E. D. (1996) *J. Am. Chem. Soc.* **118**, 407–415
512. Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T., and Kay, L. E. (1996) *J. Mol. Biol.* **263**, 627–636
513. Venters, R. A., Farmer, B. T., II, Fierke, C. A., and Spicer, L. D. (1996) *J. Mol. Biol.* **264**, 1101–1116
514. LeMaster, D. M., and Richards, F. M. (1988) *Biochemistry* **27**, 142–150
515. Bax, A., Kay, L. E., Sparks, S. W., and Torchia, D. A. (1989) *J. Am. Chem. Soc.* **111**, 408–409
516. Torchia, D. A., Sparks, S. W., and Bax, A. (1989) *Biochemistry* **28**, 5509–5524
517. Lowry, D. F., Redfield, A. G., McIntosh, L. P., and Dahlquist, F. W. (1988) *J. Am. Chem. Soc.* **110**, 6885–6886
518. Stockman, B. J., Reily, M. D., Westler, W. M., Ulrich, E. L., and Markley, J. L. (1989) *Biochemistry* **28**, 230–236
519. Qian, H., Mayo, S. L., and Morton, A. (1994) *Biochemistry* **33**, 8167–8171

519a. Riek, R., Pervushin, K., and Wüthrich, K. (2000) *Trends Biochem. Sci.* **25**, 462–468

520. Günter, P., Braun, W., Billeter, M., and Wüthrich, K. (1989) *J. Am. Chem. Soc.* **111**, 3997–4004
521. Oh, B.-H., Westler, W. M., and Markley, J. L. (1989) *J. Am. Chem. Soc.* **111**, 3083–3085
522. Cross, T. A., and Opella, S. J. (1983) *J. Am. Chem. Soc.* **105**, 306–308
523. Rothgeb, T. M., and Oldfield, E. (1981) *J. Biol. Chem.* **256**, 1432–1446
524. Buck, M., Schwalbe, H., and Dobson, C. M. (1995) *Biochemistry* **34**, 13219–13232
525. Nicholson, L. K., Kay, L. E., Baldisseri, D. M., Arango, J., Young, P. E., Bax, A., and Torchia, D. A. (1992) *Biochemistry* **31**, 5253–5263
526. Gardner, K. H., and Kay, L. E. (1997) *J. Am. Chem. Soc.* **119**, 7599–7600
527. Lenkinski, R. E., Dallas, J. L., and Glickson, J. D. (1979) *J. Am. Chem. Soc.* **101**, 3071–3077

References

528. Carver, J. A., and Bradbury, J. H. (1984) *Biochemistry* **23**, 4890 – 4905
529. Bothelho, L. H., and Gurd, F. R. N. (1978) *Biochemistry* **17**, 5188 – 5196
530. Bothelho, L. H., Friend, S. H., Matthew, J. B., Lehman, L. D., Hanania, G. I. H., and Gurd, F. R. N. (1978) *Biochemistry* **17**, 5197 – 5205
531. Morino, Y., Yamasaki, M., Tanase, S., Nagashima, F., Akasaka, K., Imoto, T., and Miyazawa, T. (1984) *J. Biol. Chem.* **259**, 3877 – 3882
532. Cocco, M. J., Kao, Y.-H., Phillips, A. T., and Lecomte, J. T. J. (1992) *Biochemistry* **31**, 6481 – 6491
533. Bachovin, W. W., Wong, W. Y. L., Farr-Jones, S., Shenvi, A. B., and Kettner, C. A. (1988) *Biochemistry* **27**, 7689 – 7697
534. Mollova, E. T., Metzler, D. E., Kintanar, A., Kagamiyama, H., Hayashi, H., Hirotsu, K., and Miyahara, I. (1997) *Biochemistry* **36**, 615 – 625
535. Guéron, M., Plateau, P., and Decorps, M. (1991) *Progr. in NMR Spectr.* **23**, 135 – 209
536. Plateau, P., and Guéron, M. (1982) *J. Am. Chem. Soc.* **104**, 7310 – 7311
537. Sklenar, V., and Bax, A. (1987) *J. of Magnetic Resonance* **74**, 469 – 479
538. Bullitt, E., and Makowski, L. (1995) *Nature (London)* **373**, 164 – 167
539. Griffin, J. H., Cohen, J. S., and Schechter, A. N. (1973) *Biochemistry* **12**, 2096 – 2099
540. Tong, H., and Davis, L. (1995) *Biochemistry* **34**, 3362 – 3367
541. Gunnarsson, G., Wennerström, H., Egan, W., and Forsén, S. (1976) *Chem. Phys. Lett.* **38**, 96 – 99
542. McDermott, A., and Ridenour, C. F. (1996) in *Encyclopedia of NMR* (Emsley, J. W., ed), pp. 3820 – 3824, J. Wiley & Sons, Ltd., Sussex, UK
543. Zhao, Q., Abeygunawardana, C., Gittis, A. G., and Mildvan, A. S. (1997) *Biochemistry* **36**, 14616 – 14626
544. Perrin, C. L., and Nielson, J. B. (1997) *J. Am. Chem. Soc.* **119**, 12734 – 12741
545. Tjandra, N., and Bax, A. (1997) *J. Am. Chem. Soc.* **119**, 8076 – 8082
546. Kintanar, A., Metzler, C. M., Metzler, D. E., and Scott, R. D. (1991) *J. Biol. Chem.* **266**, 17222 – 17229
547. Metzler, C. M., Metzler, D. E., Kintanar, A., Scott, R. D., and Marceau, M. (1991) *Biochem. Biophys. Res. Commun.* **178**, 385 – 392
548. Firsov, L. M., Neustroev, K. N., Aleshin, A. E., Metzler, C. M., Metzler, D. E., Scott, R. D., Stoffer, B., Christensen, T., and Svensson, B. (1994) *Eur. J. Biochem.* **223**, 293 – 302
549. Metzler, D. E., Metzler, C. M., Mollova, E. T., Scott, R. D., Tanase, S., Kogo, K., Higaki, T., and Morino, Y. (1994) *J. Biol. Chem.* **269**, 28017 – 28026
550. Metzler, D. E., Metzler, C. M., Scott, R. D., Mollova, E. T., Kagamiyama, H., Yano, T., Kuramitsu, S., Hayashi, H., Hirotsu, K., and Miyahara, I. (1994) *J. Biol. Chem.* **269**, 28027 – 28033
551. Metzler, D. E. (1997) *Methods Enzymol.* **280**, 30 – 40
552. Englander, S. W., Mayne, L., Bai, Y., and Sosnick, T. R. (1997) *Protein Sci.* **6**, 1101 – 1109
553. Kossiakoff, A. A. (1982) *Nature (London)* **296**, 713 – 721
554. Rohl, C. A., and Baldwin, R. L. (1994) *Biochemistry* **33**, 7760 – 7767
555. Gemmecker, G., Jahnke, W., and Kessler, H. (1993) *J. Am. Chem. Soc.* **115**, 11620 – 11621
556. Zhang, Y.-Z., Paterson, Y., and Roder, H. (1995) *Protein Sci.* **4**, 804 – 814
557. Zhang, Z., and Smith, D. L. (1993) *Protein Sci.* **2**, 522 – 531
558. Wagner, D. S., Melton, L. G., Yan, Y., Erickson, B. W., and Anderegg, R. J. (1994) *Protein Sci.* **3**, 1305 – 1314
559. Zhang, Z., Post, C. B., and Smith, D. L. (1996) *Biochemistry* **35**, 779 – 791
560. Zappacosta, F., Pessi, A., Bianchi, E., Venturini, S., Sollazzo, M., Tramontano, A., Marino, G., and Pucci, P. (1996) *Protein Sci.* **5**, 802 – 813
561. Kim, K.-S., Fuchs, J. A., and Woodward, C. K. (1993) *Biochemistry* **32**, 9600 – 9608
562. Eriksson, M. A. L., Härd, T., and Nilsson, L. (1995) *Biophys. J.* **69**, 329 – 339
563. Spyracopoulos, L., and O'Neil, J. D. J. (1994) *J. Am. Chem. Soc.* **116**, 1395 – 1402
564. Markley, J. L. (1975) *Biochemistry* **14**, 3546 – 3554
565. Schmidt-Rohr, K., and Spiess, H. W. (1994) *Multidimensional Solid-State NMR and Polymers*, Academic Press, San Diego, California
566. Schmidt-Rohr, K. (1996) *J. Am. Chem. Soc.* **118**, 7601 – 7603
567. Wu, C. H., Ramamoorthy, A., Gierasch, L. M., and Opella, S. J. (1995) *J. Am. Chem. Soc.* **117**, 6148 – 6149
568. Meyer, E. F. (1997) *Protein Sci.* **6**, 1591 – 1597
569. Stampf, D. R., Felder, C. E., and Sussman, J. L. (1995) *Nature (London)* **374**, 572 – 573
570. Holm, L., and Sander, C. (1995) *Trends Biochem. Sci.* **20**, 478 – 480
571. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol.* **247**, 536 – 540
572. Walsh, L., Newell, M., Ruczaj, H., Reid, T., and Desmond, P., eds. (1997) *Trends Guide to the Internet*, Elsevier, Cambridge, UK
573. Hogue, C. W. V., Ohkawa, H., and Bryant, S. H. (1996) *Trends Biochem. Sci.* **21**, 226 – 229
574. Wilkins, M. R., Hochstrasser, D. F., Sanchez, J.-C., Bairoch, A., and Appel, R. D. (1996) *Trends Biochem. Sci.* **21**, 496 – 497
575. Peitsch, M. C., Wells, T. N. C., Stampf, D. R., and Sussman, J. L. (1995) *Trends Biochem. Sci.* **20**, 82 – 84
576. Sayle, R. A., and Milner-White, E. J. (1995) *Trends Biochem. Sci.* **20**, 374 – 376

576a. Bernstein, H. J. (2000) *Trends Biochem. Sci.* **25**, 453 – 454

577. Sokolik, C. W. (1995) *Trends Biochem. Sci.* **20**, 122 – 124
578. Hunt, T. (1996) *Trends Biochem. Sci.* **21**, 74
579. Service, R. F. (2000) *Science* **287**, 1954–1956
580. Laskowski, R. A., Hutchinson E. G., Michie, A. D., Wallace, A. C., Jones, M. L., and Thornton, J. M. (1997) *Trends Biochem. Sci* **22**, 488 – 490
581. Cohn, E. J., and Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, Van Nostrand-Reinhold, Princeton, New Jersey (pp. 90 – 93)
582. Abbott, A. (2000) *Nature (London)* **408**, 130 – 132
583. Simmerling, C., Lee, M. R., Ortiz, A. R., Kolinski, A., Skolnick, J., and Kollman, P. A. (2000) *J. Am. Chem. Soc.* **122**, 8392 – 8402

Study Questions

1. a. From the expression for the dissociation constant of an acid, HA (Eq. 3-1), derive the logarithmic form $pH - pK_a = \log[A^-] / [HA] = \log_{10} \alpha/(1-\alpha)$ (Eq. 3-11), where α is the fraction of the acid in the ionized form.

   b. The apparent dissociation constant K_a for the $H_2PO_4^-$ ion at 25°C and 0.5 M total phosphate concentration is 1.380×10^{-7} M. What will be the pH of a solution 0.025 M in KH_2PO_4 and 0.25 M in $Na_2\ HPO_4$? This is a National Bureau of Standards buffer (see Bates[7]).

   c. Suppose that you wanted to prepare a buffer of pH = 7.00 at 25°C from anhydrous KH_2PO_4 (M_r = 136.09) and Na_2HPO_4 (M_r = 141.98). If you placed 3.40 g of KH_2PO_4 in a 1 liter volumetric flask, how much anhydrous Na_2HPO_4 would you have to weigh and add before making to volume to obtain the desired pH? If you wanted to have the correct pH to ± 0.01 unit, how accurately would you have to weigh your salts? NOTE: It is quicker to prepare a buffer of precise pH this way than it is to titrate a portion of buffer acid to the desired pH with sodium hydroxide.

2. The apparent pK_a for 0.1 M formic acid is 3.70 at 25°C.

   a. Concentrated HCl was added to a liter of 0.1 M sodium formate until a pH of 1.9 was attained. Calculate the concentration of formate ion and that of unionized formic acid.

   b. Calculate the hydrogen ion concentration.

   c. How many equivalents of HCl had to be added in part a to bring the pH to 1.9?

3. Exactly 0.01-mol portions of glycine were placed in several 100-ml volumetric flasks. The following exact amounts of HCl or NaOH were added to the flasks, the solutions were made to volume with water, mixed, and the pH measured. Calculate the pK_a values for the carboxyl and amino groups from the following, making as many independent calculations of each pKa as the data permit. At low pH values *you must correct for the free hydrogen ion concentration* (see question 2c).[581]

Flask No.	Mol HCl	Mol NaOH	pH
1	0.010		1.71
2	0.009		1.85
3	0.006		2.25
4	0.002		2.94
5		0.002	9.00
6		0.004	9.37
7		0.005	9.60

4. Using the pK_a values from problem 3, construct the theoretical titration curve showing the equivalents of H^+ or OH^- reacting with 1 mol of glycine as a function of pH. Note that the shape of this curve is independent of the pK_a. Sketch similar curves for glutamic acid (pK_a's equal 2.19, 4.25, and 9.67), histidine (pK_a's equal 1.82, 6.00, and 9.17) and lysine (pK_a's equal 2.18, 8.95, and 10.53).

   Compare your plot for glycine with a plot of 1 M acid or base added to 0.01 mol of glycine in 100 ml of water. You may also compare your curves with those for glycine published in other textbooks.

5. Make a table of characteristic pK_a values for acidic and basic groups in proteins. Which of these groups contribute most significantly to the titration curves of proteins?

6. If placed in water and adjusted to a pH of 7, will the following migrate toward the anode or the cathode if placed in an electrical field? (a) Aspartic acid, (b) alanine, (c) tyrosine, (d) lysine, (e) arginine, and (f) glutamine

7. The tripeptide L-Ala – L-His – L-Gln had the following pK_a values: 3.0 (α-COOH), 9.1 (α-NH_3^-), and 6.7 (imidazolium).

   a. What is the isoelectric pH (pI) of the peptide, i.e., the pH at which it will carry no net charge? Hint, the pI for amino acids is usually given approximately as the arithmetic mean of two pK_a values.[581]

   b. Draw the structures of the ionic forms of the peptide that occur at pH 5 and at pH 9. At each pH compute the fraction of the peptide in each ionic form.

Study Questions

8. a. Write the structure for glycyl-L-tryptophanyl-L-prolyl-L-seryl-L-lysine.

   b. What amino acids could be isolated from it following acid hydrolysis?

   c. Following alkaline hydrolysis?

   d. After nitrous acid treatment followed by acid hydrolysis

   e. In an electrolytic cell at pH 7.0 would the peptide migrate toward the cathode or toward the anode? What is the approximate isoelectric point of the peptide?

   f. If a solution of this peptide were adjusted to pH 7, and then titrated with sodium hydroxide in the presence of 10% formaldehyde, how many equivalents of base would be required per mole of peptide to raise the pH to 10?

9. A peptide was shown to contain only L-lysine and L-methionine. Titration of the peptide showed 3 free amino groups for each free carboxyl group present, and each amino group liberated 1 mole of N_2 when the peptide was treated with HNO_2 in the Van Slyke apparatus. When the deaminated peptide was hydrolyzed completely in acid and the hydrolyzate again treated with HNO_2, the same amount of N_2 is liberated as that derived from the intact peptide. A sample of the original peptide was treated with excess dinitrofluorobenzene to give a dinitrophenyl (DNP) peptide, which was shown spectrophotometrically to contain three DNP groups per free carboxyl group. When this DNP-peptide was completely hydrolyzed, the following products were found: a colorless compound containing S (A_1) and a yellow compound containing S (A_3). Partial hydrolysis of the DNP-peptide yields A_1, A_2, A_3, plus 4 additional yellow compounds, B_1, B_2, B_3, and B_4. On complete hydrolysis, B_1 yields A_1, A_2, and A_3; B_2 yields A_1 and A_2; B_3 yields A_1 and A_3; and B_4 yields A_3 only.

   What is the most probable structure of the original peptide?

10. (a) Explain two advantages of the isotope dilution method of analysis. (b) From the following data, calculate the amount of cyclic AMP (cAMP) present per ml of human gluteus maximus muscle cells. Cells were treated with ^{32}P-enriched cAMP, S.A. = 50 µCi / µmol for 0.2 h (all cAMP was taken up by cells), cells were homogenized, and the soluble cAMP was isolated and purified. The specific activity of the isolated cAMP = 10 µCi / µmol. The total amount of cAMP added was 1.0×10^{-7} mol per ml of cells.

11. The figure in Box 3-C shows a high-resolution separation of the soluble proteins of *E. coli*. The investigator labeled the proteins with ^{14}C-containing amino acids.

    a. How would you carry out the labeling experiment?

    b. What other isotope(s) could be used to label proteins? What chemical form(s) would you use? What limitations might there be?

    c. What soluble components of an *E. coli* cell sonicate might interfere with the two-dimensional separation, and how could they be removed?

    d. What technique(s) other than radioactive labeling could be used for locating proteins?

    e. If *all* the soluble proteins of *E. coli* were detected, about how many separate proteins would you expect to see?

    f. Indicate two or more properties of the resolution technique which are most significant in making it applicable to a system containing a very large number of proteins.

12. ^{35}S is a beta emitter, with no gamma or other type of radiation. It has the following properties: $t_{1/2}$ = 86.7 days, ε_{max} = 0.168 MeV.

    a. Write the equation for the radiochemical decomposition of ^{35}S.

    b. Discuss the advantages and limitation of the use of ^{35}S as an isotopic tracer.

Each cotton fiber is a single cell seed hair, ~30 mm in length. The dry fiber is ~95% cellulose, which constitutes the secondary cell wall (See Fig. 20-4,D) and is also present in the primary cell wall. The fibers, which are ~30 nm in length, consist of many parallel chains (Fig. 4-5) ~5 µm in length, each containing ~10,000 glucose residues. Van der Waals forces, together with one hydrogen bond per glucose, contribute to the stability of the tightly packed fiber. Cotton is one of the major agricultural crops, ~87 million bales, each ~220 kg, being produced annually world-wide. Photo and information from A. D. French and M. A. Godshall, Southern Regional Research Center, USDA, New Orleans, LA.

Contents

161 .... **A. Structures and Properties of Simple Sugars**
162 .... 1. The Variety of Monosaccharides
163 .... *Ways of indicating configuration*
164 .... *Natural derivatives of sugars*
166 .... 2. Conformations of Five- and Six-Membered Ring Forms
167 .... 3. Characteristic Reactions of Monosaccharides
167 .... **B. Glycosides, Oligosaccharides, Glycosylamines, and Glycation**
169 .... **C. Polysaccharides (Glycans)**
170 .... 1. Conformations of Polysaccharide Chains
172 .... 2. The Glucans
175 .... 3. Other Homopolysaccharides
175 .... 4. Heteropolysaccharides of the Animal Body
177 .... 5. Plant Heteropolysaccharides
179 .... 6. Polysaccharides of Bacterial Surfaces
180 .... **D. Glycoproteins and Proteoglycans**
181 .... 1. *O*-Linked Glycosyl Groups
182 .... 2. Asparagine-Linked Oligosaccharides
186 .... 3. Glycoproteins in Biological Recognition
186 .... *Lectins and other carbohydrate-binding proteins*
186 .... *Carbohydrate-binding sites*
186 .... *Binding of viruses and bacteria to cells*
187 .... *Aggregation and adherence of cells*
187 .... *Growth and differentiation*
188 .... *Recognition and adhesion by leukocytes*
188 .... **E. Some Special Methods**
188 .... 1. Release of Oligosaccharides from Glycoproteins
189 .... 2. Hydrolysis
189 .... 3. Methylation
190 .... 4. Periodate Oxidation (Smith Degradation)
190 .... 5. Nuclear Magnetic Resonance

192 .... **References**
196 .... **Study Questions**

Boxes

171 .... Box 4-A Cyclodextrins
178 .... Box 4-B Silicon: An Essential Trace Element
184 .... Box 4-C The Blood Group Determinants
191 .... Box 4-D Antifreeze and Ice-Nucleation Proteins

Tables

170 .... Table 4-1 Some of the Many Polysaccharides Found in Nature

Sugars, Polysaccharides, and Glycoproteins

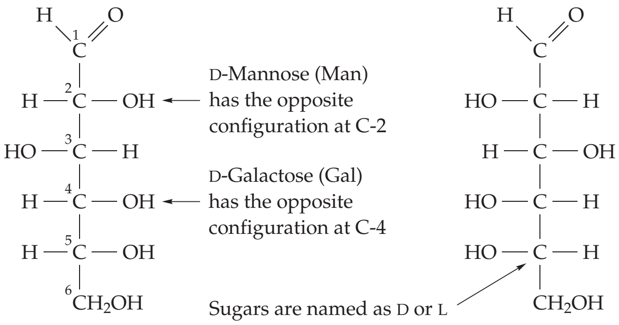
4

We are all familiar with sugars, important components of our diet which are present in fruits, honey, table sugar, and syrups. Within our bodies the simple sugar **D-glucose** is an essential source of energy. It is present in blood at a relatively constant concentration of 5.5 mM and is carried to all tissues. A **polysaccharide** called **glycogen**, a polymeric form of glucose, provides a reserve of readily available energy within our cells. **Starches** and other polysaccharides store energy within plants. Polysaccharides also have major *structural* functions in nature. **Cellulose**, another polymer of glucose, forms the fibers of cotton, plant cell walls, and wood. Both the tough exoskeletons of anthropods and the cell walls of fungi depend for their strength on the nitrogen-containing polysaccharide **chitin**. Polysaccharides that carry many negative charges, such as **hyaluronan,** form a protective layer between animal cells, while **pectins** play a similar role in plants.

A third function of sugar residues (**glycosyl groups**) is in biological recognition and communication. The outer surfaces of cells are nearly covered by covalently attached **oligosaccharides**, small polymeric arrays of sugar rings. Some are attached to side chain groups of proteins and others to lipids to form **glycoproteins** and **glycolipids,** respectively. Because of the variety of different sugars, the various ways in which they can be linked, and their ability to form oriented hydrogen bonds, these oligosaccharides provide much of the chemical code for identifying cells. This coding enables cells to attach to each other in correct ways during development of a multicelled organism. It helps to activate our immune system to attack parasites and also helps bacteria to attack us! We are only beginning to understand the numerous critical functions of the glycosyl groups of cell surfaces.

A. Structures and Properties of Simple Sugars

The simple sugars, or **monosaccharides**, are polyhydroxyaldehydes (**aldoses**) or polyhydroxyketones (**ketoses**).[1–5] All have the composition $(CH_2O)_n$, hence the family name **carbohydrate**. A typical sugar, and the one with the widest distribution in nature, is glucose.



D-Mannose (Man) has the opposite configuration at C-2

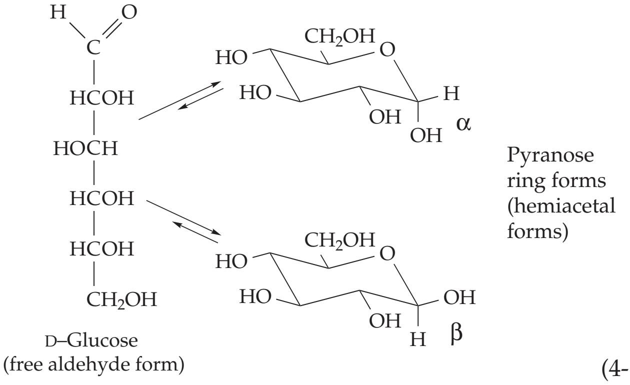
D-Galactose (Gal) has the opposite configuration at C-4

Sugars are named as D or L according to the configuration about this carbon atom, one removed from the terminal position

D-Glucose (Glc) showing numbering of atoms

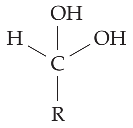
L-Glucose

The carbonyl group of this and other sugars is highly reactive and a characteristic reaction is **addition** of electron-rich groups such as –OH. If a sugar chain is long enough (4–6 carbon atoms) one of the hydroxyl groups of the same molecule can add to the carbonyl group to form a cyclic **hemiacetal** or ring form, which reaches an equilibrium with the free aldehyde or ketone form (Eq. 4-1). The six-membered rings formed in this way (**pyranose** rings) are especially stable, but five-membered **furanose** rings also exist in many carbohydrates.

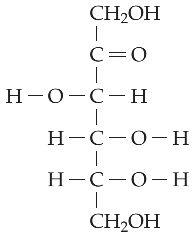


(4-1)

The ring forms of the sugars are the monomers used by cells to form polysaccharides. Indeed, it is the natural tendency of 5- and 6-carbon sugars (**pentoses** and **hexoses**) to cyclize that permits formations of stable sugar polymers from the reactive and unstable monomers. When a sugar cyclizes a new chiral center is formed at the **anomeric carbon atom**, the atom that was present in the original carbonyl group. The two configurations about this carbon atom are designated α and β as indicated in Eq. 4-1. In an equilibrium mixture, *ring forms of most sugars predominate over open chains.*[6–13] Thus, at 25°C in water glucose reaches an equilibrium with ~0.001% free aldehyde, 0.004% covalent hydrate of the aldehyde (see also Eq. 13-1), 39% α-pyranose form, 61% β-pyranose form, and 0.15% each of the much less stable α and β furanose forms.[9,12] The ketose sugar **fructose** in solution exists as ~73% β-pyranose, 22% β-furanose, 5% α-furanose, and 0.5% open chain form (Fig. 4-1).[6,10,13] Although polysaccharides are composed almost exclusively of sugar residues in ring forms, the open chain forms are sometimes metabolic intermediates.



Covalent hydrate
of R—CHO



D-Fructose (Fru), a ketose that is a close structural and metabolic relative of D-glucose. It occurs in honey and fruit juices in free form, in the disaccharide sucrose (table sugar) as a 5-membered furanose ring, and in other oligosaccharides and polysaccharides.

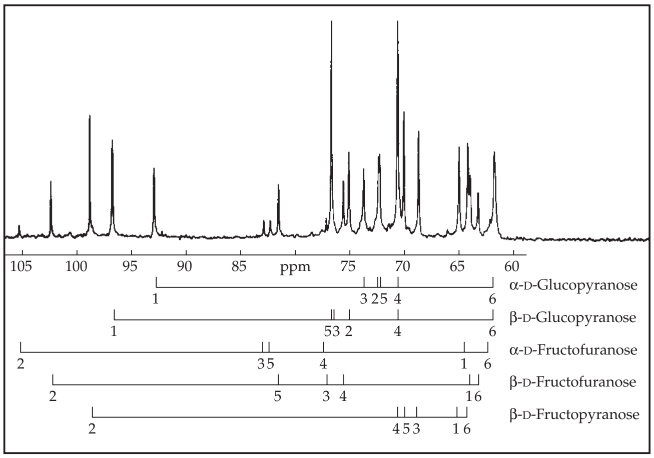


Figure 4-1 Natural abundance ^{13}C-NMR spectrum of honey showing content of α and β pyranose ring forms of glucose and both pyranose and furanose ring forms of fructose. Open chain fructose, with a peak at 214 ppm, was present in a trace amount. From Prince *et al*.[6]

Because of their many polar hydroxyl groups, most sugars are very soluble in water. However, hydrogen bonds between molecules stabilize sugar crystals making them insoluble in nonpolar solvents. Intermolecular hydrogen bonds between chains of sugar rings in cellulose account for much of the strength and insolubility of these polysaccharides.

1. The Variety of Monosaccharides

Sugars contain several chiral centers and the various diastereomers are given different names. The commonly occurring sugars **D-glucose**, **D-mannose**, and **D-galactose** are just three of the 16 diastereomeric aldohexoses. The Fischer projection formulas for the entire family of eight D-aldoses with 3–6 carbon atoms are given in Fig. 4-2. Several of these occur only rarely in nature.

Monosaccharides are classified as D or L according to the configuration at the chiral center farthest from the carbonyl group. If the –OH group attached to this carbon atom lies to the right when the sugar is oriented according to the Fischer convention, the sugar belongs to the D family. The simplest of all the chiral sugars is glyceraldehyde. The family of aldoses in Fig. 4-2 can be thought of as derived from D-glyceraldehyde by upward extension of the chain. Besides D-glucose, D-mannose, and D-galactose, the most abundant naturally occurring sugar shown in Fig. 4-2 is **D-ribose**, a major component of RNA. Another abundant aldose is **D-xylose**, a constituent of the polysaccharides of wood.

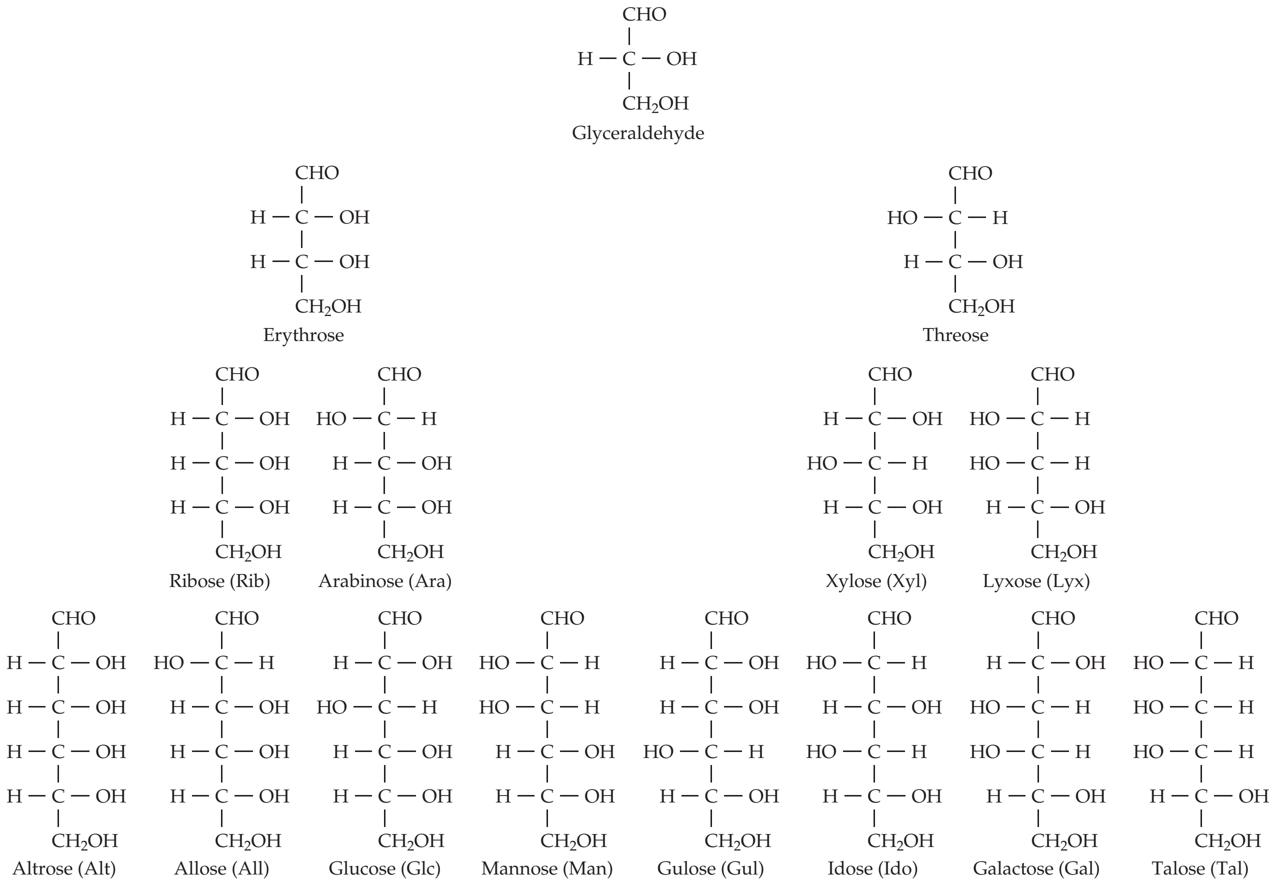
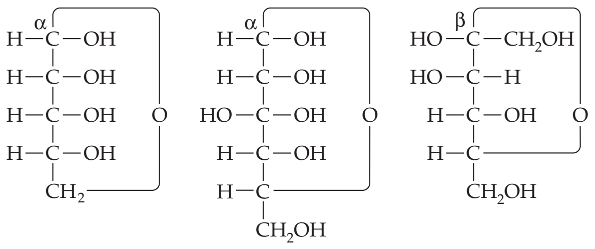


Figure 4-2 Formulas for the D-aldoses. Prefixes derived from the names of these aldoses are used in describing various other sugars including ketoses. The prefixes include *erythro*, *threo*, *arabino*, *ribo*, *galacto*, *manno*, etc., D-Fructose can be described in this manner as D-*arabino*-hexulose. The prefixes refer to the configurations of a series of consecutive but not necessarily contiguous chiral centers.[2] Thus, 3-*arabino*-hexulose is an isomer of fructose with the carbonyl group at the 3 position. The prefix deoxy, which means "lacking oxygen," is often used to designate a modified sugar in which an –OH has been replaced by –H, e.g., 2-deoxyribose.

For each D sugar there is an L sugar which is the complete mirror image or enantiomer of the D form. Although **L-arabinose** occurs widely in plants, and some derivatives of L sugars are present in glycoproteins, most naturally occurring sugars have the D configuration. A pair of sugars, such as glucose and galactose, differing in configuration at only one of the chiral centers are known as **epimers**. The D-ketose sugars with the carbonyl group in the 2 position (Fig. 4-3) are also abundant in nature, often occurring in the form of phosphate esters as intermediates in metabolism.

Ways of indicating configuration. The Fischer projection formulas used in Figs. 4-2 and 4-3 are convenient in relating the sugar structures by their individual carbon configurations to each other, but they give an unrealistic three-dimensional picture. According to the Fischer convention each carbon atom must be viewed with both vertical bonds projecting behind the atom viewed. In fact, the molecule cannot assume such a conformation because the chain folds back on itself, bringing many atoms into collision.

Ring forms of sugars are also often drawn according to the Fischer convention; making use of elongated bent lines to represent ordinary simple bonds:



α-D-Ribopyranose α-D-Glucopyranose β-D-Fructofuranose

In this representation the hydroxyl at the anomeric

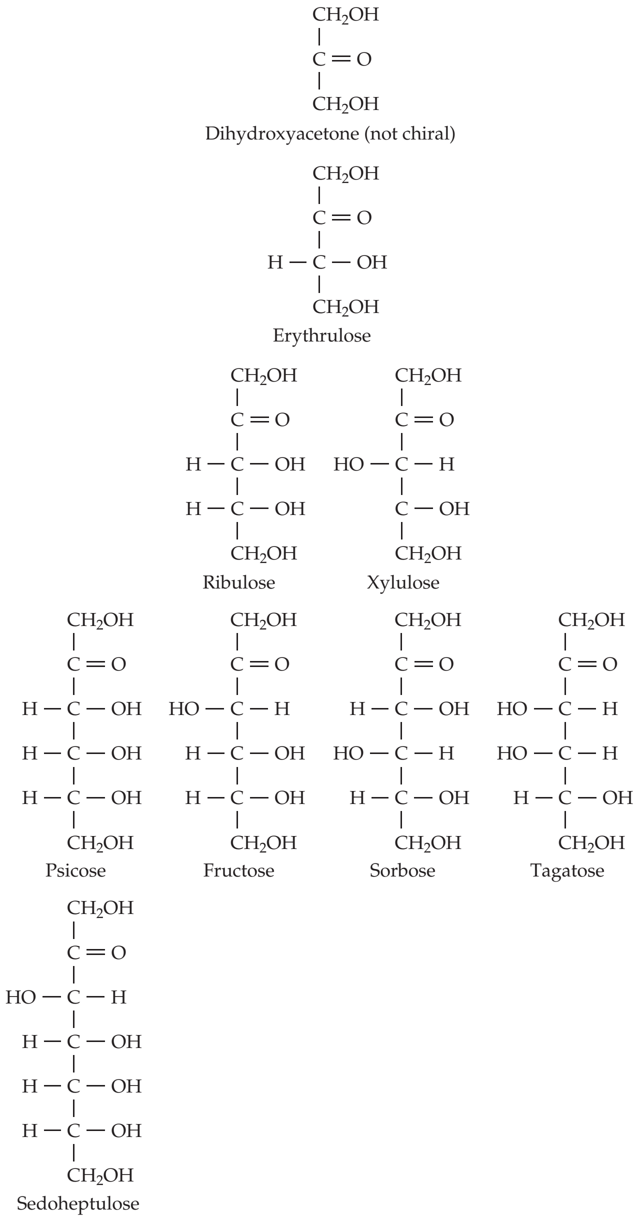
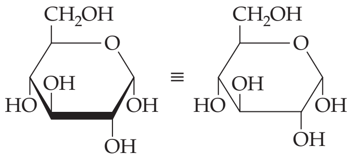


Figure 4-3 Formulas for the open forms of the D-ketoses.

carbon atom is always on the right side of the molecules for α forms in the D series of sugars and at the left side for the β forms. For the L sugars the opposite is true. For example, since α-D-glucopyranose and α-L-glucopyranose are enantiomers, their Fischer formulas must be mirror images.

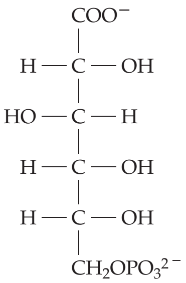
Simplified sugar rings are often drawn with **Haworth structural formulas**. The lower edge of the ring, which may be shown as a heavy line, is thought of as projecting out toward the reader and the other edge as projecting behind the plane of the paper.



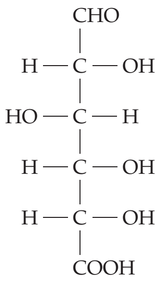
α-D-Glucopyranose (Haworth structural formula)

Haworth structures are easy to draw and unambiguous in depicting configurations,[14] but they also do not show the spatial relationships of groups attached to other rings correctly. For this reason conformational formulas of the type described in Section 2 and shown in Fig. 4-4 are used most often in this book.

Natural derivatives of sugars. The aldehyde group of an aldose can be oxidized readily to a carboxyl group to form an **aldonic acid**. Among the several aldonic acids that occur naturally is 6-phosphogluconic acid, which is pictured here as the 6-phosphogluconate ion:



6-Phosphogluconate ion



Glucuronic acid (GlcA)

The enzymatically formed **uronic acids** have –COOH in the terminal position. This is position 6 in **glucuronic acid**, whose structure is given here and as a ring in Fig. 4-4. Sugar chains with –COOH at both ends are called **aldaric acids**, e.g., glucaric acid. The –OH group in the 2 position of glucose may be replaced by $-NH_2$ to form 2-amino-2- deoxyglucose, commonly called **glucosamine** (GlcN), or by $-NH-CO-CH_3$ to form ***N*-acetylglucosamine** (GlcNAc). Similar derivatives of other sugars exist in nature. In many poly-saccharides, sulfate groups are attached in ester linkage to the sugar units. The sulfo ($-SO_3^-$) sugar **6-sulfo-α-D-quinovose** (Fig. 4-4) is found in lipids of photosynthetic membranes.[15,16]

The sugar alcohols, in which the carbonyl group has been reduced to –OH, also occur in nature. For example, **D-glucitol** (D-sorbitol), the sugar alcohol obtained by reducing either D-glucose or L-sorbose (Eq. 4-2), is a major product of photosynthesis and widely distributed in bacteria and throughout the eukaryotic kingdom. It is present in large amounts in berries of the mountain ash and in many other fruits. It exists in a high concentration in human semen and

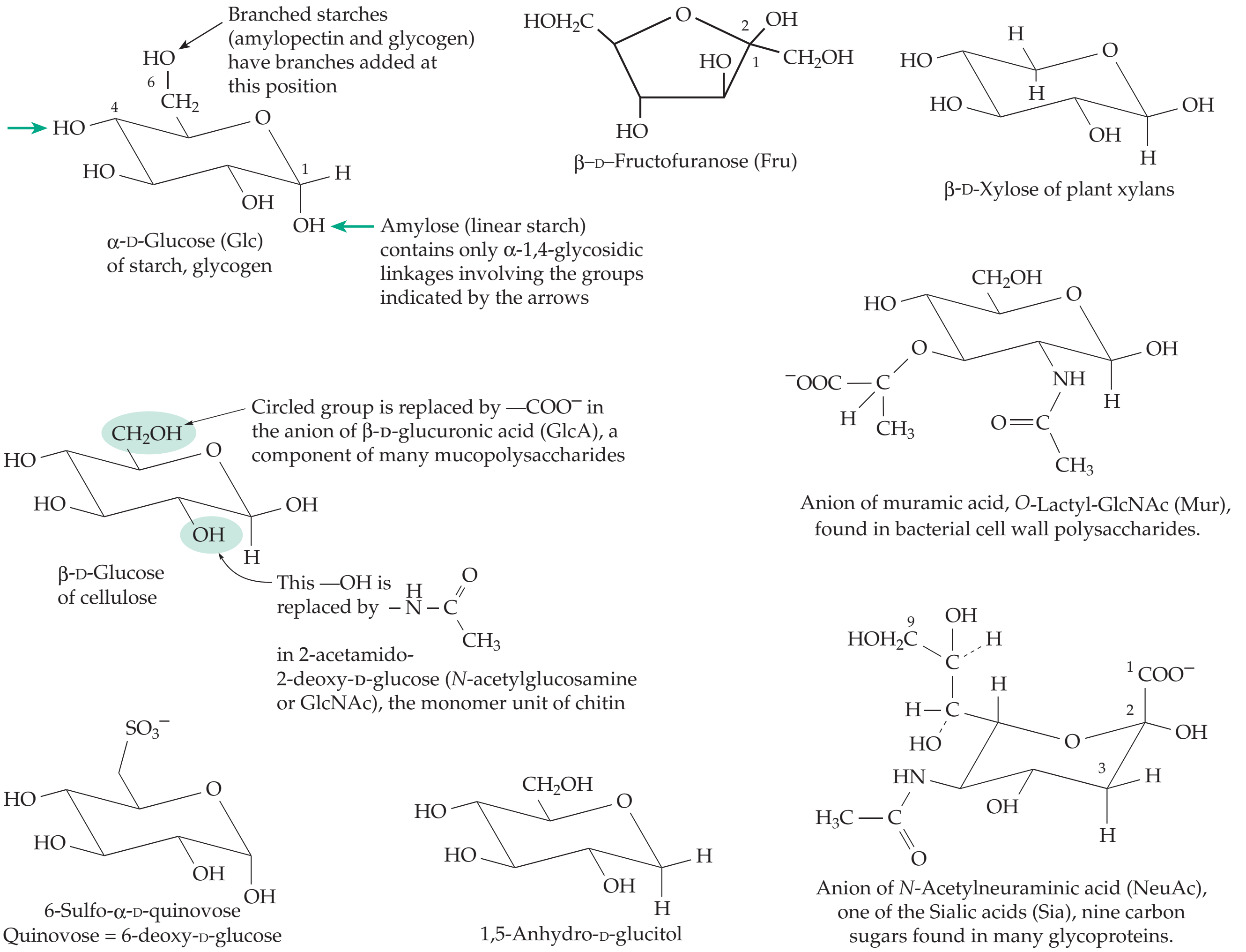
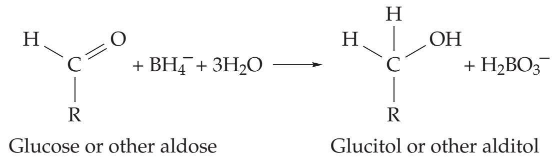

Branched starches (amylopectin and glycogen) have branches added at this position
HO
6
CH2
4
HO
O
HO
1
H
OH
OH
α-D-Glucose (Glc) of starch, glycogen
Amylose (linear starch) contains only α-1,4-glycosidic linkages involving the groups indicated by the arrows
HOH2C
O
2
OH
HO
1
CH2OH
HO
β-D-Fructofuranose (Fru)
H
O
HO
HO
H
OH
OH
H
β-D-Xylose of plant xylans
CH2OH
HO
O
O
OH
−OOC
C
NH
H
H
CH3
O=C
CH3
Anion of muramic acid, O-Lactyl-GlcNAc (Mur), found in bacterial cell wall polysaccharides.
Circled group is replaced by —COO− in the anion of β-D-glucuronic acid (GlcA), a component of many mucopolysaccharides
CH2OH
HO
O
HO
OH
OH
H
β-D-Glucose of cellulose
This —OH is replaced by −N−C(=O)CH3 (H on N)
in 2-acetamido-2-deoxy-D-glucose (N-acetylglucosamine or GlcNAc), the monomer unit of chitin
9
OH
HOH2C
H
C
H
H—C
1
COO−
HO
2
O
OH
HN
3
H
H3C—C
OH
O
H
Anion of N-Acetylneuraminic acid (NeuAc), one of the Sialic acids (Sia), nine carbon sugars found in many glycoproteins.
SO3−
HO
O
HO
OH
OH
6-Sulfo-α-D-quinovose
Quinovose = 6-deoxy-D-glucose
CH2OH
HO
O
HO
H
OH
H
1,5-Anhydro-D-glucitol

Figure 4-4 Some simple sugars and sugar derivatives in ring forms. Most of these are present in polysaccharides.

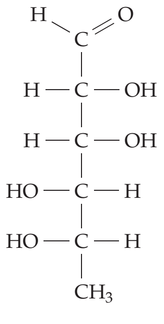


$$\mathrm{RCHO} + BH_4^- + 3H_2O \longrightarrow \mathrm{RCH_2OH} + H_2BO_3^- \qquad (4\text{-}2)$$

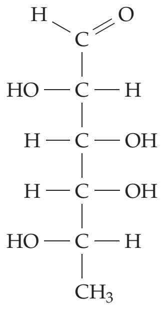
Glucose or other aldose
Glucitol or other alditol

accumulates in lenses of diabetics. D-Glucitol and other sugar alcohols arise in some fungi during metabolism of the corresponding sugars.[17] Mannitol, another product of photosynthesis, is also present in many organisms.[18] Another polyol that is found in human blood in significant concentrations (0.06–0.25 mM) is 1,5-anhydro-D-glucitol (Fig. 4-4). It is largely of dietary origin.[19]

Two common 6-deoxy sugars which lack the hydroxyl group at C-6 are **rhamnose** and **fucose**. Both are of the "unnatural" L configuration but are derived metabolically from D-glucose and D-mannose, respectively.

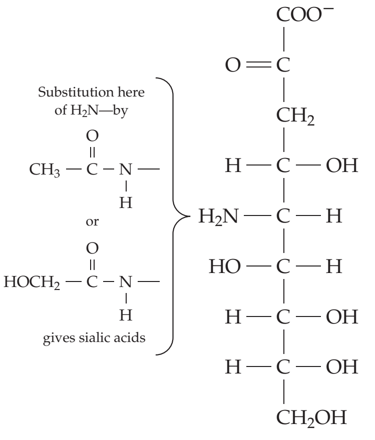


L-Rhamnose (Rha)



L-Fucose (Fuc)

Vitamin C (ascorbic acid, Box 18-D) is another important sugar derivative. **Neuraminic acid** is a 9-carbon sugar made by transferring a 3-carbon piece onto a hexosamine. Its *N*-acetyl and *N*-glycolyl derivatives are called **sialic acids**. Their names may be abbreviated NeuAc and NeuGl, respectively, or simply as Sia (see also Fig. 4-4).

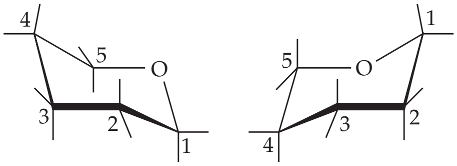


Neuraminic acid anion (Neu)

The sialic acids are prominent constituents of the glycoproteins of cell surfaces. More than 30 modified forms, for example, with added methyl or acetyl groups, are known.[20–24]

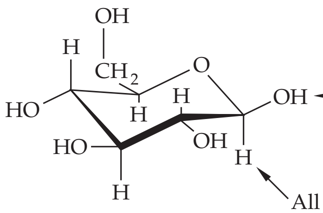
2. Conformations of Five- and Six-Membered Ring Forms

Single-bonded six-membered molecular rings, such as those in cyclohexane and in sugars, most often assume a **chair (C)** conformation. For sugars in pyranose ring forms, there are two possible chair conformations,[4,25,26] which are designated $^4C_1$(or C1) and $^1C_4$ (or 1C). The superscript and subscript numbers on the designations indicate which atoms are above and below the plane of the other four ring atoms. *These conformers are not easily interconvertible unless the ring is opened by a chemical reaction,* e.g., that of Eq. 4-1. However, manipulation with an atomic force microscope has shown that stretching of single polysaccharide chains can cause interconversion of the two chair conformations of pyranose rings, which are separated by an energy barrier of ~46 kJ/mol.[26a]



$^4C_1$ (or C1) $^1C_4$ (or 1C)

Most sugars occur in the chair conformation that places the largest number of substituents in equatorial positions and is therefore most stable thermodynamically. For D-aldoses this is usually the $^4C_1$ conformation:

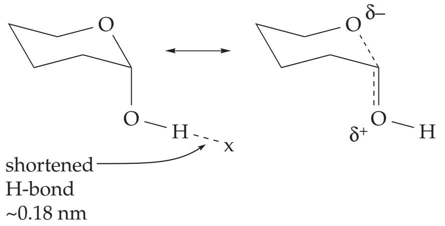


All bulky substituents, —OH and —CH_2OH, are in equatorial positions projecting out from the ring.

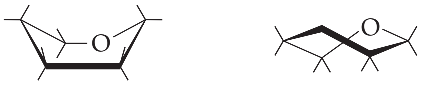
All of the hydrogens attached to ring carbons project downward or upward in axial positions.

β-D-Glucose, a pyranose ring form in a *chair* conformation.

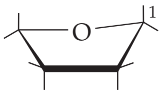
For L-aldoses it is the $^1C_4$ form. There are exceptions to this rule. For example, α-D-idopyranose as well as iduronate rings assume the $^1C_4$ conformation because this conformation places the maximum number of bulky groups in equatorial positions.[27,28] It is noteworthy that electronegative substituents on the anomeric carbon atom of a sugar ring often prefer an axial orientation. This **anomeric effect** is also reflected in a shortening by about 0.01 nm of hydrogen bonds involving the hydrogen atom of the anomeric axial OH group. These effects can be explained partially as a result of coulombic repulsion of the two C–O dipoles and of resonance of the following type:[29–31]



In addition to the chair conformations of six-membered rings the less stable **boat (B)** conformations are also possible. The six boat forms are smoothly interconvertible through intermediate **twist** (**T**) forms, which are also called **skew** (**S**) forms.[3,5] Since the internal angle in a pentagon is 108° (close to the tetrahedral angle) we might anticipate a nearly planar five-membered ring. However, eclipsing of hydrogen atoms on adjacent carbons prevents formation of such a flat structure. One of the atoms may be buckled out of the plane of the other four about 0.05 nm, into an **envelope conformation** (e.g., ^1E), or only three atoms may be in a plane, as in a twist conformation.



Boat Skew or twist

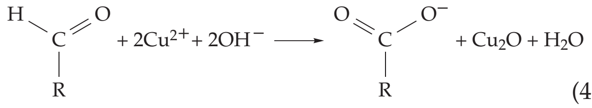


The ^1E envelope conformation

Any one of the five atoms of the ring can be either above or below the plane defined by the other four in the envelope conformation. The energy barriers separating them are very low, and in cyclopentane or in proline all of the envelope conformations are freely interconvertible through intermediate skew forms.[32] Furanose sugar rings are very flexible but the presence of the bulky substituents reduces the number of possible conformations.[33–36a] See Chapter 5 for further discussion.

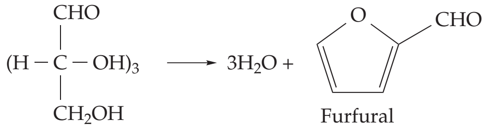
3. Characteristic Reactions of Monosaccharides

The aldehyde group of aldoses can either be oxidized or reduced and ketoses can be reduced. The best laboratory reagent for reduction is **sodium borohydride** which acts rapidly in neutral aqueous solutions (see Eq. 4-2). Since both NaB^3H_4 and NaB^2H_4 are available, radioactive or heavy isotope labels can be introduced in this way. The aldehyde groups can be oxidized by a variety of agents to the corresponding aldonic acids, a fact that accounts for the reducing properties of these sugars. In alkaline solution aldoses reduce Cu^{2+} ions to cuprous oxide (Eq. 4-3), silver ions to the free metal, or hexacyanoferrate (III) to hexacyanoferrate (II). These reactions provide the basis for sensitive analytical procedures. Even though the aldoses tend to exist largely as hemiacetals (Eq. 4-1) the reducing property is strongly evident. Oxidation by metal-containing reagents is usually via the free aldehyde, but oxidation by hypobromite BrO^- (Br_2 in alkaline solution) yields the lactone, as does enzymatic oxidation (Eq. 15-10).

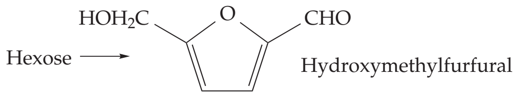


(4-3)

Sugars are unstable in acid. Boiling with concentrated HCl or H_2SO_4 converts pentoses to furfural (Eq. 4-4a) and hexoses to hydroxymethylfurfural (Eq. 4-4b).



(4-4a)

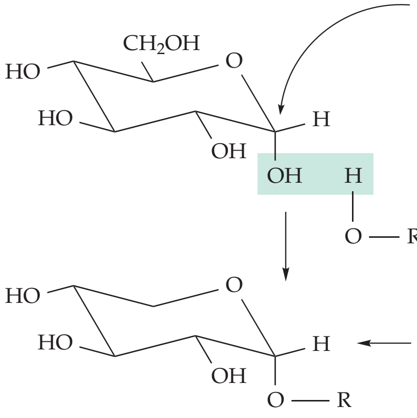


(4-4b)

The aldehydes produced in these reactions can be condensed with various phenols or quinones to give colored products useful in quantitative estimation of sugar content or in visualizing sugars on thin-layer chromatographic plates. Phenol and sulfuric acid yield a product whose absorbance at 470 nm can be used as a measure of the total carbohydrate content of most samples. Resorcinol (1,3-dihydroxybenzene) in 3 M HCl (Seliwanoff's reagent) gives a red precipitate with ketoses. Orcinol (5-methylresorcinol) reacts rapidly with pentoses and with ribonucleosides and ribonucleotides. Since 2-deoxy sugars react slowly this can be used as a test for RNA. Diphenylamine with H_2SO_4 gives a blue-green color specifically with 2-deoxy sugars and can be used to test for DNA if the sample is first hydrolyzed.[37]

B. Glycosides, Oligosaccharides, Glycosylamines, and Glycation

The hydroxyl group on the anomeric carbon atom of the ring forms of sugars is reactive and can be replaced by another nucleophilic group such as –O–R from an alcohol. The product is a glycoside (Eq. 4-5). The reaction with methanol occurs readily with acid catalysis under dehydrating conditions, e.g., in 100% methanol.

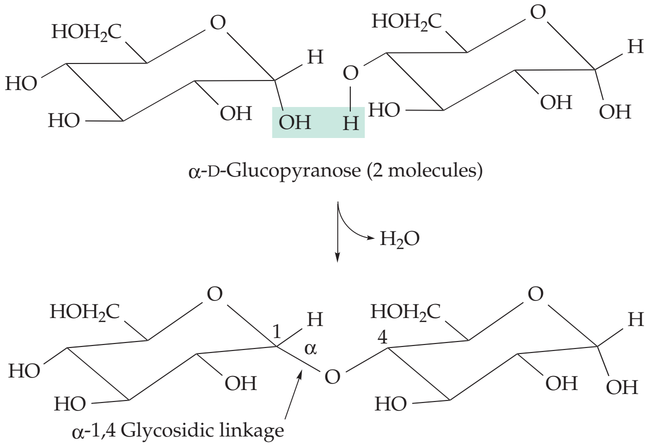


The carbon atom that is attached directly to *two* oxygen atoms in the hemiacetal group is the **anomeric** carbon atom. The attached OH group can be replaced to form a glycoside.

A glycoside (an α-D-glucopyranoside in this instance). The molecule HO–R is called the ***aglycone*** of the glycoside.

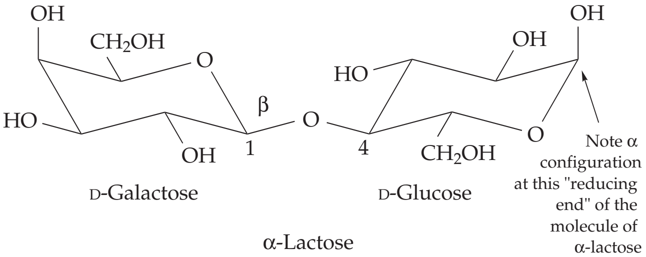
(4-5)

The alcohol in this equation can be a simple one such as methanol or it can be any of the –OH groups of another sugar molecule. For example, two molecules of α-D-glucopyranose can be joined, in an indirect synthesis, to form **maltose** (Eq. 4-6). Maltose is formed by the hydrolysis of starch and is otherwise not found in nature. There are only three abundant naturally occurring **disaccharides** important to the metabolism of plants and animals.[38] They are **lactose** (milk), **sucrose** (green plants), and **trehalose** (fungi and insects).



Maltose, a disaccharide (4-6)

Disaccharides are linked by **glycosidic** (acetal) linkages. The symbol α-1,4, used in Eq. 4-6, refers to the fact that in maltose the glycosidic linkage connects carbon atom 1 (the anomeric carbon atom) of one ring with C-4 of the other and that the configuration about the anomeric carbon atom is α. While the α and β ring forms of free sugars can usually undergo ready interconversion, the configuration at the anomeric carbon atom is "frozen" when a glycosidic linkage is formed. To describe such a linkage, we must state this configuration together with the positions joined in the two rings (see Eq. 4-6). Lactose, whose structure follows, can be described as a disaccharide containing one galactose unit in a β-pyranose ring form and whose anomeric carbon atom (C-1) is joined to the 4 position of glucose, giving a β-1,4 linkage:

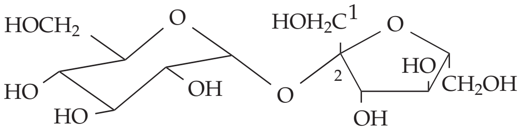

OH
CH_2OH
O
HO
OH
β
1
O
4
HO
OH
OH
CH_2OH
O
D-Galactose
D-Glucose
Note α configuration at this "reducing end" of the molecule of α-lactose

α-Lactose

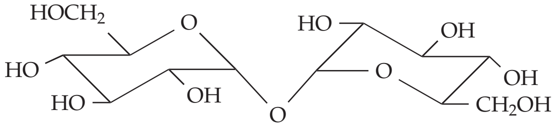
The systematic name for α-lactose, *O*-β-D-galactopyranosyl-(1→4)-α-D-glucopyranose, provides a complete description of the stereochemistry, ring sizes, and mode of linkages. This name may be abbreviated β-D-Gal*p*-(1→4)-α-D-Glc*p*. Since pyranose rings are so common and since most natural sugars belong to the D family, the designations D and *p* are often omitted. It may be assumed that in this book sugars are always D unless they are specifically designated as L. When linkages remain uncertain abbreviated formulas are given. Because the glucose ring in lactose is free to open to an aldehyde and to equilibrate (in solution) with other ring forms, the name lactose does not imply a fixed ring structure for the glucose half. Thus, the name lactose can be abbreviated as βGal-(1→4)-Glc or, more succinctly, as Galβ1→4Glc. However, in crystalline form the sugar exists either as α-lactose or β-lactose. The latter is more soluble and sweeter than α-lactose, which sometimes crystallizes in ice cream upon prolonged storage and produces a "sandy" texture. An isomer of lactose, Galβ1→6Glc or **allolactose**, is an important inducer of transcription in cells of *E. coli* (Chapter 28).

Notice that in the drawing of the lactose structure the glucose ring has been "flipped over" with respect to the orientation of the galactose ring, a consequence of the presence of the β-1,4 linkage. For maltose, where the linkage is α-1,4, the two rings are usually drawn with the same orientation (Eq. 4-6). Maltose can be described as α-D-Glc*p*-(1→4)-D-Glc*p* or more simply as Glcα1→4Glc.

In sucrose and in α,α-trehalose the reducing groups of two rings are joined. Each of these sugars exists in a single form. Sucrose serves as the major transport sugar in green plants, while trehalose plays a similar role in insects, as does D-glucose in our blood.



Sucrose: Glc*p*α1→β2Fru*f*



α, α-Trehalose: Glc*p*α1→α1Glc*p*

Trehalose, or "mushroom sugar," is found not only in fungi but also in many other organisms.[39–41]. It serves as the primary transport sugar in the hemolymph of insects and also acts as an "antifreeze" in many species. It accounts for up to 20% of the dry weight of **anhydrobiotic organisms**, which can survive complete dehydration. These include spores of some fungi, yeast cells, macrocysts of *Dictyostelium*, brine shrimp cysts (dried gastrulas of *Artemia salina*), some nematodes, and the resurrection plant. These organisms can remain for years in a dehydrated state. Hydrogen bonding between the trehalose and phosphatidylcholine may stabilize the dry cell membranes.[18,40,41] Although they can be desiccated, fungal spores remain dormant even when considerable water is present. One of the first detectable changes when the spores germinate is a rapid increase in the activity of the enzyme **trehalase** which hydrolyzes trehalose to glucose.[42,43]

Disaccharides, as well as higher oligosaccharides and polysaccharides, are thermodynamically unstable with respect to hydrolysis, for example, for lactose in aqueous solution:

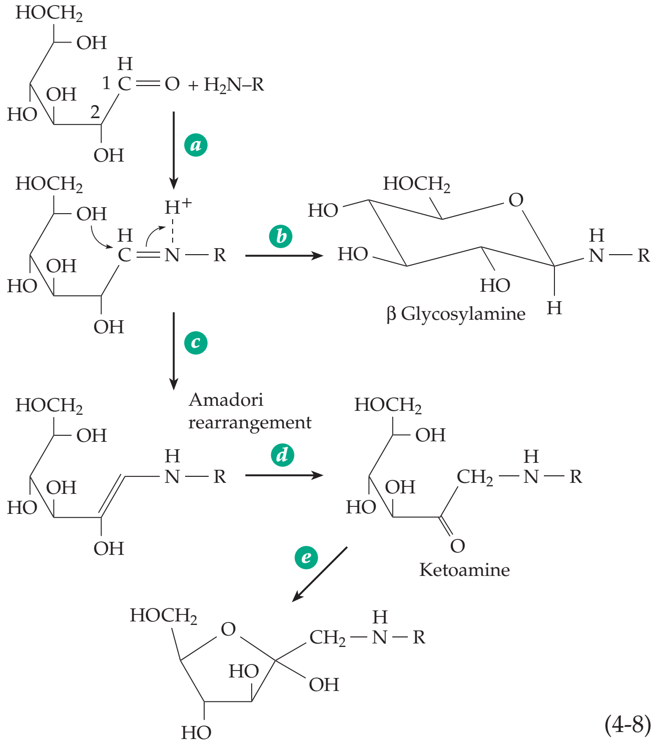
$$\text{Lactose} + H_2O \rightarrow \textsc{d}\text{-glucose} + \textsc{d}\text{-galactose}$$
$$\Delta G° \approx -8.7 \pm 0.2 \text{ kJ mol}^{-1} \text{ at } 25°C \qquad (4\text{-}7)$$

The corresponding equilibrium constant $K = [\textsc{d}\text{-glucose}][\textsc{d}\text{-galactose}]/[\text{lactose}] \approx 34$ M. For other oligosaccharides K varies from 17 to 500 M.[44,45] However, sucrose is far less stable with $K = 4.4 \times 10^4$ M and $\Delta G° = -26.5 \pm 0.3$ kJ mol^{-1} for its hydrolysis.[46] Sucrose is also less stable kinetically and undergoes rapid acid-catalyzed hydrolysis. This fact is exploited when vinegar is used to convert sucrose to the less crystallizable mixture of glucose and fructose during candy-making. In marked contrast, trehalose is extremely resistant to acid-catalyzed hydrolysis.

The joining of additional sugar rings through glycosidic linkages to a disaccharide leads to the formation of **oligosaccharides**, which contain a few residues, and to **polysaccharides**, which contain many residues. Among the well-known oligosaccharides are the substituted sucroses **raffinose**, Gal*p*(1→6) Glc*p*(1→2) Fru*f*, and **stachyose**, Gal*p*(1→6) Gal*p*(1→6) Glc*p*(1→2) Fru*f*. Both sugars are found in many legumes and other green plants in which they are formed by attachment of the galactose rings to sucrose. Oligosaccharides have many functions. For example, gram-negative bacteria often synthesize oligosaccharides of 6–12 glucose units in β-1,2 linkage joined to *sn*-1-phosphoglyceryl groups. They are found in the periplasmic space between the inner and outer cell membranes and may serve to control osmotic pressure.[47] Oligosaccharides of 10–14 α-1,4-linked D-galacturonic acid residues serve as signals of cell wall damage to plants and trigger defensive reactions against bacteria in plants.[48]

Just as alcohols can be linked to sugars by glycoside formation (Eq. 4-5), amines can react similarly to give **glycosylamines** (*N*-glycosides). In this instance it is usually the free aldehyde that reacts via formation of a Schiff base (Eq. 4-8*a*). The latter can cyclize (Eq. 4-8*b*) to the glycosylamine with either the α or β configuration. Another important reaction of Schiff bases of sugars is the **Amadori rearrangement** (Eq. 4-8*c*,*d*), which produces a secondary ketoamine with a 1-oxo-2-deoxy structure.

This can cyclize as in step *e* of Eq. 4-8. In addition, epimerization at position 2 occurs through the reversal of step *d*. The overall reaction of Eq. 4-8 is often called **glycation** to distinguish it from **glycosylation,** the transfer of a glycosyl group. The Amadori rearrangement is important in nitrogen metabolism and in nonenzymatic reactions of sugars. For example, small



(4-8)

amounts of glycosylated hemoglobin and other proteins modified by glycation of the protein amino groups are normally present in the blood.[49–51] People with diabetes, who have a high concentration of blood glucose, have increased amounts of glycated protein. High concentration of either glucose or fructose[41,52] may cause serious problems. For example, modification of the protein **crystallins** of the lens of the eye may lead to cataracts. Similar problems with galactose may accompany galactosemia (Chapter 20).[53] Ketoamines formed by glycation of proteins may undergo crosslinking reactions with side chains of other protein molecules and this may be one cause of aging.[42,43] Other reactions of the ketoamines lead to formation of fluorescent and colored products.[41,43,50,54–56] Oxidation products of sugars also participate in these reactions,[57] and nucleic acid bases also react.[56]

C. Polysaccharides (Glycans)

Polymers of sugars are present in all cells and serve a variety of functions.[58–60] The simple sugars commonly used in the assembly of polysaccharides include D-glucose, D-mannose, D-galactose, D-fructose, D-xylose, L-arabinose, related uronic acids, and amino sugars (Fig. 4-4). These monomer units can be put together in many ways, either as **homopolysaccharides** containing a single kind of monomer or as **heteropolysaccharides** containing two or more different monomers. Because there are many sugars and many ways in which they can be linked, there is a bewildering

TABLE 4-1
Some of the Many Polysaccharides Found in Nature

Name	Source	Monomer	Main linkage	Branch linkages
Starch	Green plants			
Amylose		D-Glucose	α1,4	
Amylopectin		D-Glucose	α1,4	α1,6
Glycogen	Animals, bacteria	D-Glucose	α1,4	α1,6
Cellulose	Green plants, some bacteria	D-Glucose	β1,4	
Dextrans	Some bacteria	D-Glucose	α1,6	α1,3
Pullulan	Yeast	D-Glucose	α1,6 + α1,4	
Callose	Green plants	D-Glucose	β1,3	
Yeast glucan	Yeast	D-Glucose	β1,3	
Schizophyllan, curdlan, paramylon		D-Glucose	β1,3	β1,6 on every third residue
Mannans	Algae	D-Mannose	1,4	
	Yeast	D-Mannose	α1,6	
Xylans	Green plants Brown seaweed	D-Xylose	β1,3	
Inulin	Some plant tubers	D-Fructose	β2,6	
Chitin	Fungi, anthropods	*N*-acetyl-D-Glucosamine	β1,4	
Alternating polysaccharides				
Hyaluronan	Animal connective tissue	Glucuronic acid + *N*-acetylglucosamine	β1,4	
Chondroitin sulfate		D-Glucosamine *N*-acetyl-D-Galactosamine	β1,3 + β1,4	
Dermatan sulfate		α-L-Iduronate + *N*-acetyl-D-Galactosamine	β1,3 + β1,4	
Pectin	Higher plants	D-Galactunonate + others	β1,4 + others	
Alginate	Seaweed	D-Mannuronate + L-Guluronate	β1,4 + α1,4	
Agar-agar	Red seaweed	Galactose	β1,4 and α1,3	
Carageenan	Red seaweed	Galactose-4-sulfate + 3,6-anhydro-D-Galactose-2-sulfate	β1,4 + α1,3	
Murein	Bacterial cell wall	*N*-acetyl-D-Glucosamine + *N*-acetyl-D-Muramic acid	β1,4	

variety of different polysaccharides. Their chains can be linear or helical. The most numerous functional groups present are the free hydroxyl groups, some of which may form additional glycosidic linkages to produce **branched chains**. Polysaccharides may also contain $-COOH$, $-NH_2$, $-NHCOCH_3$, and other groups. After polymerization, hydroxyl groups are sometimes methylated or converted to sulfate esters or to ketals formed with pyruvic acid. Structural characteristics of some of the major polysaccharides are listed in Table 4-1.

1. Conformations of Polysaccharide Chains

Despite the variety of different monomer units and kinds of linkage present, the conformational possibilities for carbohydrate chains are limited. The sugar ring is a rigid unit and the connection of one unit to the next can be specified by means of two torsion angles ϕ and ψ just as with peptides.[61–63] To specify a torsion angle four atoms must be selected–the two at the ends of the bond about which rotation is being considered and two others. There is more than one way to define the zero angle for ϕ and ψ. As illustrated in the drawing on p. 172, ϕ may be taken as the H1–C1–O–C4' dihedral angle and ψ as the H4' –C4'–O–C1 angle. The zero angle is when H1 and H4' are eclipsed. A related alternative is to take ϕ and ψ as 0° when the two midplanes of the sugar rings are coplanar.

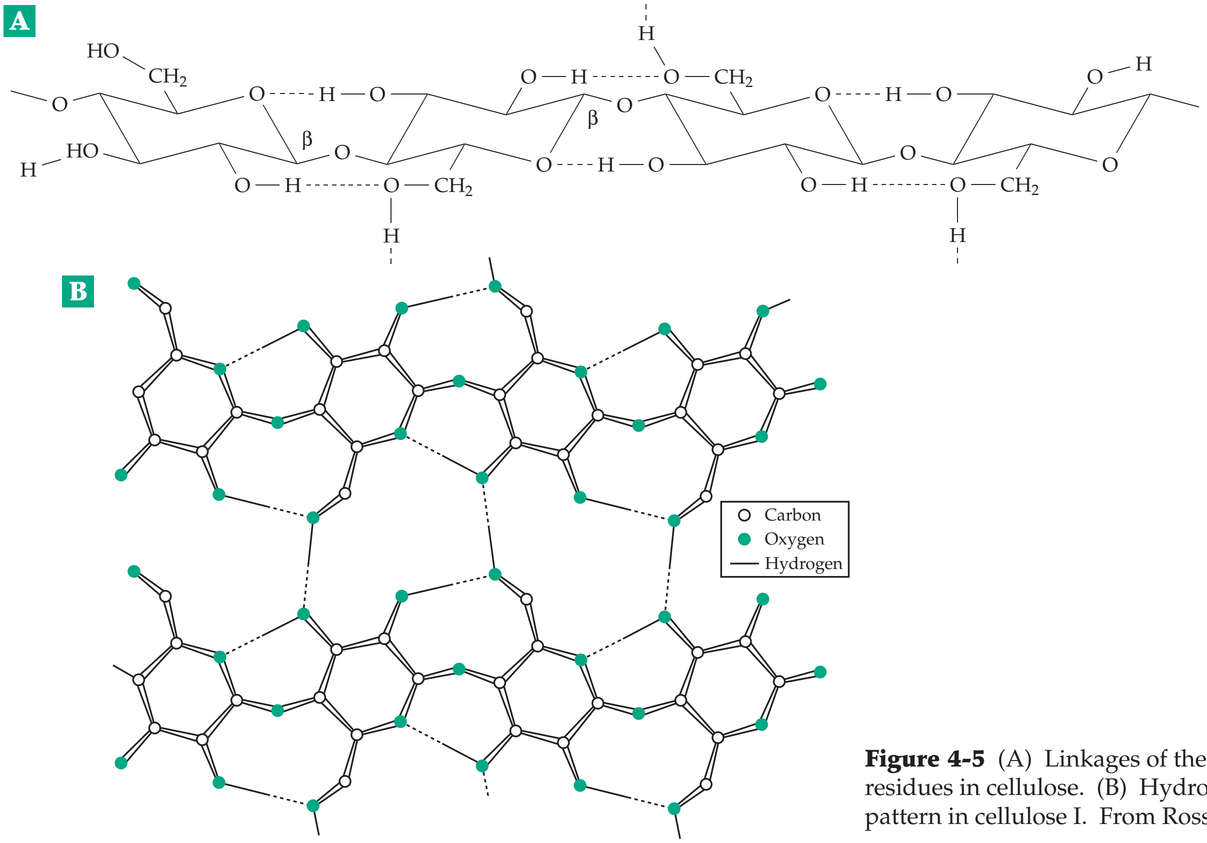
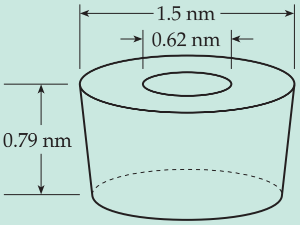


Figure 4-5 (A) Linkages of the D-glucose residues in cellulose. (B) Hydrogen-bonding pattern in cellulose I. From Ross *et al.*[66]

BOX 4-A CYCLODEXTRINS

An enzyme produced by some bacteria of the genus *Bacillus* cuts chains of amylose and converts them into tiny rings consisting of six (α-cyclodextrin), seven (β-cyclodextrin), eight (γ-cyclodextrin), or more glucose units. The cyclodextrins, which were first isolated by F. Schardinger in 1903, have intrigued carbohydrate chemists for many years by their unusual properties.[a] They are surprisingly resistant to acid hydrolysis and to attack by amylases and cannot be fermented by yeast.

The torus-like cyclodextrin molecules have an outer polar surface and an inner nonpolar surface. The small hydrophobic cavities within the cyclodextrins have diameters of 0.50, 0.62, and 0.79 nm, respectively, for the α, β, and γ dextrins.[b] Cavities are potential binding sites for a great variety of both inorganic and organic molecules.[a–g] Complexes of simple alcohols, polyiodides,[c] ferrocene,[d] and many other compounds have been observed. Cyclodextrins can be used to "encapsulate" food additives[b] and their complexes may be useful in separation of enantiomers of drugs.[f,g] They can be chemically modified by adding catalytic groups to serve as enzyme models[h,i] or as color-change indicators sensitive to binding of organic molecules.[j] They can be linked together to form molecular **nanotubes**. Polymer chains can even be threaded through the tubes.[k] What practical applications may yet come from this?



---

[a] French, D. (1957) *Adv. Carbohydr. Chem. Biochem.* **12**, 189–260

[b] Korpela, T., Mattsson, P., Hellman, J., Paavilainen, S., and Mäkelä, M. (1988–89) *Food Biotechnology* **2**, 199–210

[c] Noltemeyer, M., and Saenger, W. (1980) *J. Am. Chem. Soc.* **102**, 2710–2722

[d] Menger, F. M., and Sherrod, M. J. (1988) *J. Am. Chem. Soc.* **110**, 8606–8611

[e] Hamilton, J. A., and Chen, L. (1988) *J. Am. Chem. Soc.* **110**, 4379–4391

[f] Armstrong, D. W., Ward, T. J., Armstrong, R. D., and Beesley, T. E. (1986) *Science* **232**, 1132–1135

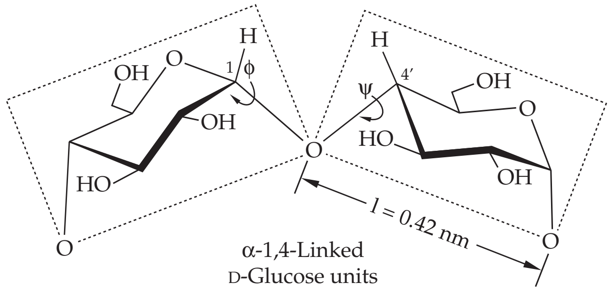
[g] Lipkowitz, K. B., Raghothama, S., and Yang, J. (1992) *J. Am. Chem. Soc.* **114**, 1154–1162

[h] Anslyn, E., and Breslow, R. (1989) *J. Am. Chem. Soc.* **111**, 8931–8932

[i] Granados, A., and de Rossi, R. H. (1995) *J. Am. Chem. Soc.* **117**, 3690–3696

[j] Ueno, A., Kuwabara, T., Nakamura, A., and Toda, F. (1992) *Nature* **356**, 136–137

[k] Harada, A., Kamachi, J. L., and Kamachi, M. (1993) *Nature* **364**, 516–518



The IUB has proposed another convention.[64]

2. The Glucans

From glucose alone various organisms synthesize a whole series of polymeric glucans with quite different properties. Of these, **cellulose,** an unbranched β-1,4-linked polyglucose (Fig. 4-5A), is probably the most abundant. It is the primary structural polysaccharide of the cell walls of most green plants.[65] For the whole Earth, plants produce ~10^{14} kg of cellulose per year.

A systematic examination of the possible values for ϕ and ψ shows that for cellulose these angles are constrained to an extremely narrow range which places the monomer units in an almost completely extended conformation.[62] Each glucose unit is flipped over 180° from the previous one. The polymer has a twofold screw axis and there is a slight zigzag in the plane of the rings.[66–68] Remember that in the chair form of glucose all of the –OH groups lie in equatorial positions and are able to form hydrogen bonds with neighboring chains. This feature, together with the rigidity of conformation imposed by the β configuration of the monomer units, doubtless accounts for the ability of cellulose to form strong fibers.

It has been impossible to obtain large single crystals of cellulose. However, from 60 to 90% of native cellulose is thought to be clustered to form the needle-like crystalline microfibrils. These microcrystals of **cellulose I** can be separated from other plant materials by prolonged boiling with dilute NaOH and HCl. Their structure has been established by electron diffraction[67] and by comparison with high-resolution X-ray diffraction structure of the tetrasaccharide β-D-cellotetraose.[66,69] Two closely similar parallel-chained, hydrogen-bonded structures appear to be present. One is shown in Fig. 4-5B. The hydrogen bonds and van der Waals forces bind the chains into sheets which are stacked to form fibers. A typical fiber of plant cellulose has a diameter of 3.5–4 nm and contains 30–40 parallel chains, each made up of 2000–10,000 glucose units. The chain ends probably overlap to form essentially endless fibers that can extend for great distances through the cell wall. They interact with other polysaccharides as is illustrated in Fig. 4-14. A single cotton fiber may be 2–3 cm in length.

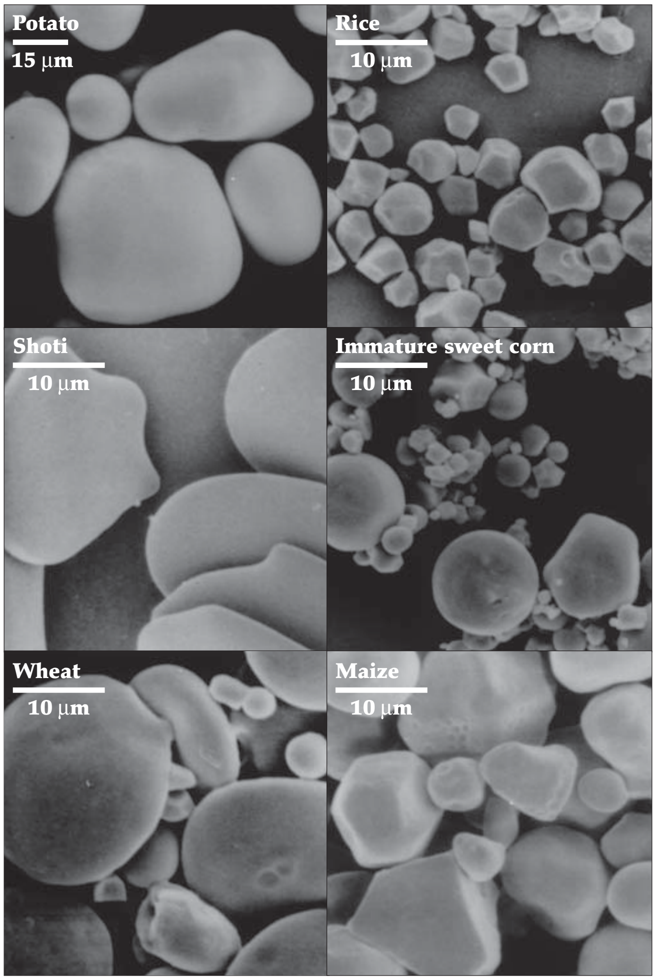


Figure 4-6 Scanning electron micrographs of starch granules. Magnification 2500× and 3000×. From Jane *et al.*[73]

Cotton thread treated with concentrated NaOH shrinks and has an increased luster. The resulting "Mercerized" cellulose has changed into other crystalline forms. The major one is **cellulose II**, in which the chains in the sheets are antiparallel.[69a] Cellulose II may also occur to some extent in nature. Many other modified celluloses, e.g., **methylcellulose**, in which some –OH groups have been converted to methyl ethers[70] are important commercial products.

Starch, another of the most abundant polymers of glucose, is stored by most green plants in a semi-crystalline form in numerous small granules. These granules, which are usually formed within colorless membrane-bounded plastids, have characteristic shapes and appearances (Fig. 4-6) that vary from plant to plant. One component of starch, **amylose**, is a linear polymer of many α-D-glucopyranose units in 1,4 linkage (Fig. 4-7) as in maltose. Starch granules always contain a second kind of molecule known as **amylopectin**.[58]

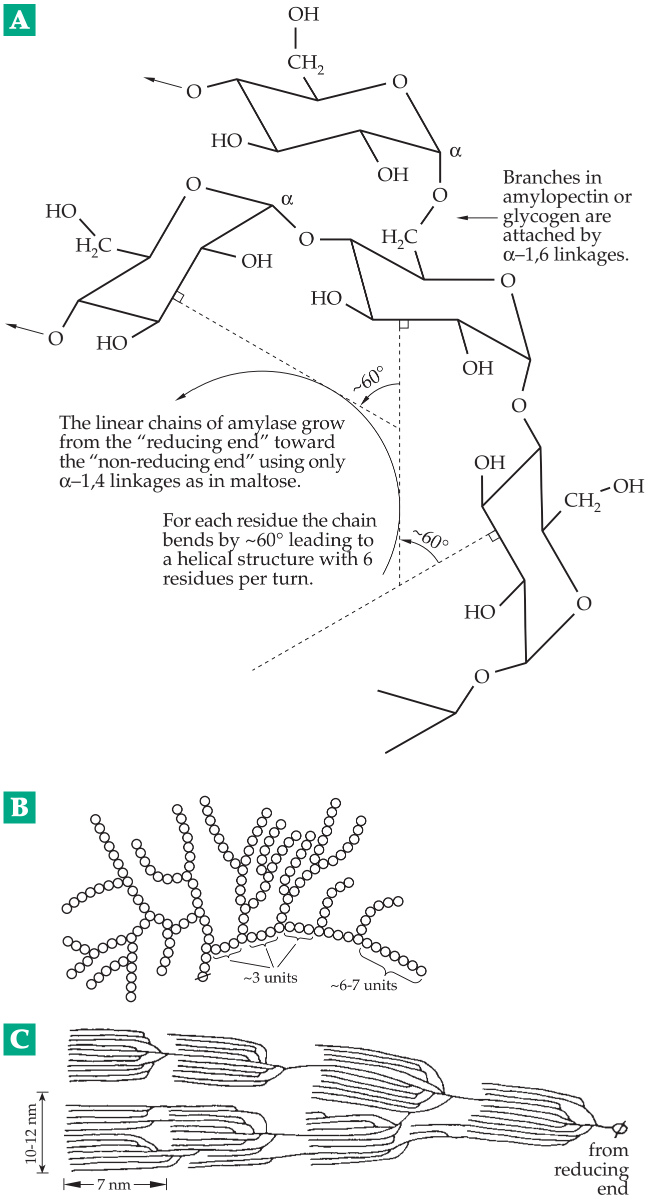

A
OH
CH2
O
O
HO
α
OH
O
α
HO
H2C
O
OH
H2C
O
HO
O
HO
OH
Branches in amylopectin or glycogen are attached by α–1,6 linkages.
~60°
O
The linear chains of amylase grow from the "reducing end" toward the "non-reducing end" using only α–1,4 linkages as in maltose.
OH
CH2
OH
For each residue the chain bends by ~60° leading to a helical structure with 6 residues per turn.
~60°
HO
O
O
B
~3 units
~6-7 units
C
10-12 nm
7 nm
from reducing end

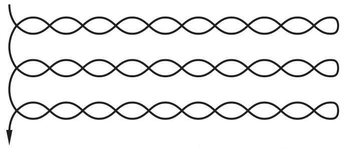
Figure 4-7 (A) Linkages of the glucose residues in starches and in glycogen. (B) Schematic diagram of the glycogen molecule as proposed originally by K. H. Meyer.[74] The circles represent glucose residues which are connected by α-1,4 linkages and, at the branch points, by α-1,6 linkages. The symbol ϕ designates the reducing group. From D. French.[75] (C) Proposed broomlike clusters in amylopectin. After D. French.[71]

Both amylopectin and **glycogen** (animal starch) consist of highly branched bushlike molecules. Branches are attached to α-1,4-linked chains through α-1,6 linkages (Fig. 4-7). When glucose is being stored as amylopectin or glycogen the many "nonreducing" ends appear to grow like branches on a bush, but when energy is needed the tips are eaten back by enzymatic action (Fig. 11-2). There is only one **reducing end** of a starch molecule where one might expect to find a hemiacetal ring in equilibrium with the free aldehyde. However, this end may be attached to a protein.

Starch granules show a characteristic pattern of growth rings of 0.3- to 0.4-μm thickness with thinner dense layers about 7 nm apart. Study of this layered structure suggests that the individual amylopectin molecules have branches close together in broomlike clusters (Fig. 4-7C).[61,71,72] The amylopectin chains are 120–400 nm long and their relative molecular masses may reach 15–30 million. In addition, starch granules usually contain molecules of the straight-chain amylose, each containing several hundred glucose units and having molecular masses of ~100 kDa. Most starches contain 20–21% amylose but there are special varieties of plants that produce starch with 50–70% amylose. On the other hand, the "waxy" varieties of maize form only amylopectin and lack amylose.

In both starch and glycogen the glucose units of the main chains are linked with α-1,4 linkages. An extended conformation is not possible and the chains tend to undergo helical coiling. One of the first helical structures of a biopolymer to be discovered (in 1943)[76,77] was the left-handed helix of amylose wound around molecules of pentaiodide (I_5^-) in the well-known blue starch–iodine complex[78] (Fig. 4-8). The helix contains six residues per turn, with a pitch of 0.8 nm and a diameter of nearly 14 nm. Amylose forms complexes of similar structure with many other small molecules.[79]

Another more tightly coiled double-helical form of amylose has been proposed.[80] Each chain would contain six glucose units per turn and the two chains could be arranged in either parallel or antiparallel directions. The average amylose molecule contains 2000 glucose units and could be stretched to a slender chain over 1 μm long, longer than the crystalline regions observed in starch granules. Thus, the chains within the granules would have to fold back on themselves, possibly in hairpin fashion:



Crystalline array of hairpin folds proposed for linear starch molecules

X-ray diffraction studies support the double-helical structure but suggest a *parallel* orientation of the amylose chains.[81] Since amylose has not been obtained as single crystals the diffraction data do not give a definitive answer. However, if double helices are formed by adjacent branches in amylopectin and glycogen the two strands would be parallel. Starch granules also contain amorphous starch which appears to contain single helices, possibly wrapped around lipid materials.[82]

Glycogen is stored in the cytoplasm of animal cells and to some extent in the lysosomes as enormous 100- to 200-MDa particles. These appear in the electron microscope as aggregates of smaller particles of molecular masses up to 20–40 MDa. A laminated internal structure with surface bumps is suggested by STM microscopy (Fig. 4-9). Biosynthesis of glycogen may be initiated by a 37-kDa protein called **glycogenin**, which remains covalently attached to the reducing end of the glycogen (Chapter 20).[83] Despite the huge molecular masses of glycogen particles, both ^1H and ^{13}C NMR resonances are sharp, indicating a high degree of mobility of the glycosyl units.[84]

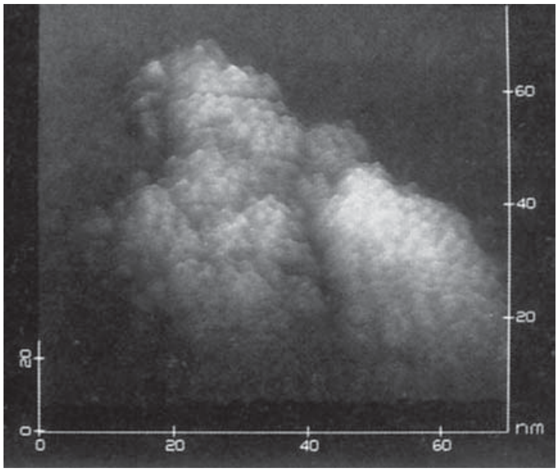


Figure 4-9 A scanning tunneling microscopic (STM) image of three glycogen molecules. The corrugated surface suggests a laminated structure. These molecules have been purified from tissues by treatment with strong alkali, which breaks the larger aggregates into particles of $M_r = 1-10 \times 10^6$ and diameter 25–30 nm. Courtesy of Fennell Evans.[93]

Beta-1,3-linked glucans occur widely in nature. When a new green plant cell is formed the first polysaccharide to be synthesized is not cellulose but the β-1,3-linked glucose polymer **callose**. Cellulose appears later. Callose is also produced in some specialized plant tissues, such as pollen tubes,[85] and is formed in massive amounts at the site of wounds or of attack by pathogens. The major structural component of the yeast cell wall is a β-1,3-linked glucan with some β-1,6 branches.[86] **Schizophyllan** is a β-1,3-linked glucan with a β-1,6-linked glucosyl group attached to every third residue. A glucan from the coleoptiles of oats contains 30% β-1,3 linkages in a linear chain that otherwise has the structure of cellulose.[87] Other β-1,3-linked glucans serve as energy storage molecules in lower plants and in fungi. Among these are β-1,3-linked glucans such as **paramylon**,[88,89] which is stored by the euglena. A similar polysaccharide, **curdlan**, is formed by certain bacteria.[90]

Some other bacteria, e.g., *Leuconostoc mesenteroides*, make 1,6-linked poly-D-glucose or **dextrans**.[91,92] These always contain some α-1,3-linked branches and may also have α-1,4 and α-1,2 linkages, the structures varying from species to species. Dextrans formed by bacteria such as *Streptococcus mutans* growing on the surfaces of teeth are an important component of dental plaque. Bacterial

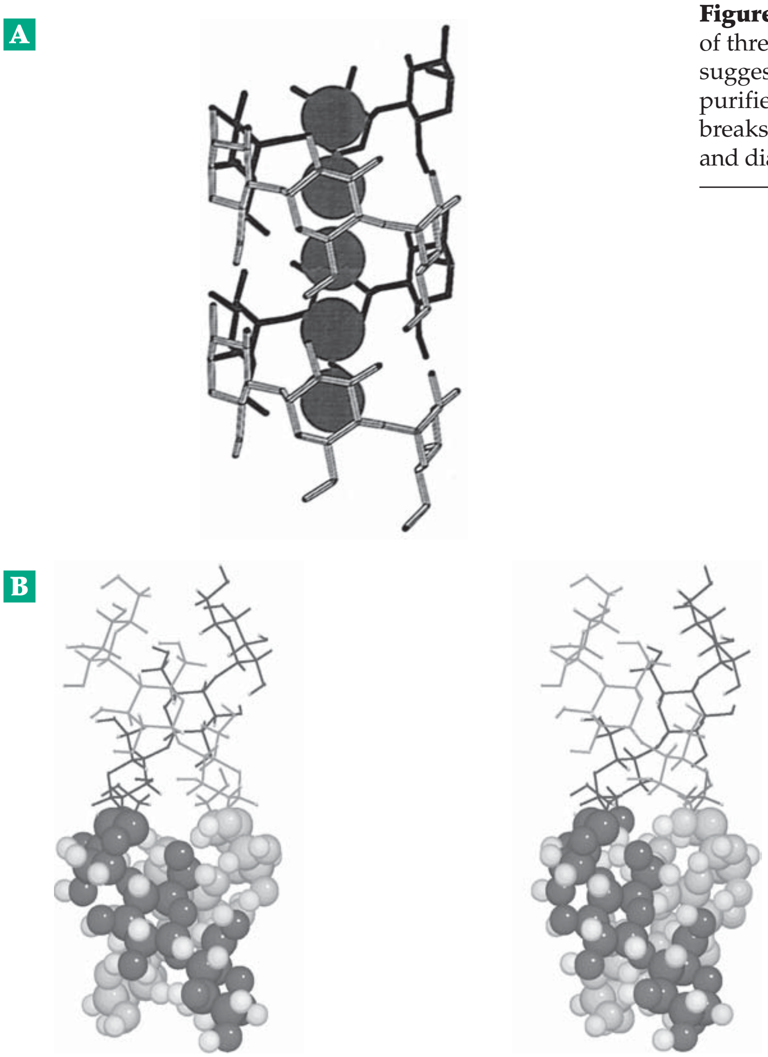


Figure 4-8 (A) Structure of the helical complex of amylose with I_3^- or I_5^-. The iodide complex is located in the interior of the helix having six glucose residues per turn. (B) Model of a parallel-stranded double helix. There are six glucose units per turn of each strand. The repeat period measured from the model is 0.35 nm per glucose unit. Courtesy of Alfred French.

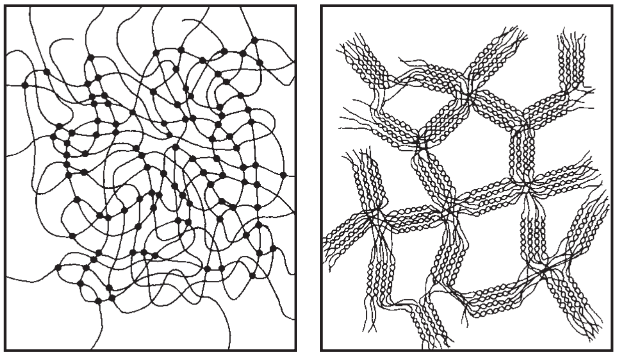


Figure 4-10 A schematic representation of the gel networks of Sephadex (left) and agarose (right). Note that the aggregates in agarose gels may actually contain 10–10^4 helices rather than the smaller numbers shown here. From Arnott *et al*.[95]

dextrans are also produced commercially and are chemically crosslinked to form gels (Sephadex) which are widely used in biochemical separation procedures (Figs. 3-1 and 4-10). A yeast polysaccharide **pullulan** is a regular polymer of maltotriose units with α-1,4 linkages joined in a single chain by α-1,6 linkages.[94] One glucan, called **alternan** is a linear polymer with alternate α-1,6 and 1,3 linkages.

3. Other Homopolysaccharides

Cell walls of yeasts contain **mannans** in which the main α-1,6-linked chain carries short branches of one to three mannose units joined in α-1,2, α-1,3, and sometimes α-1,6 linkages.[96] These are covalently linked to proteins.[97] A β-1,4-linked mannan forms microfibrils in the cell wall of some algae such as *Acetabularia* (Fig. 1-11) which do not make cellulose.[98,99] The cell walls of some seaweeds contain a β-1,3-linked **xylan** instead of cellulose. This polysaccharide forms a three-stranded right-handed triple helix.[99,100] Even though xylose is a five-carbon sugar, the polymer contains the thermodynamically more stable six-membered pyranose rings. On the other hand, fructose, a 6-carbon sugar, is present as five-membered furanose rings in **inulin**, the storage polysaccharide of the Jerusalem artichoke and other Compositae, and also in sweet potatoes. The difference has to do with biosynthetic pathways. Furanose rings arise both in inulin and in sucrose because the biosynthesis occurs via the 6-phosphate ester of fructose, making it impossible for the phosphate derivative to form a 6-membered ring.

The major structural polysaccharide in the exoskeletons of arthropods and of other lower animal forms is **chitin**, a linear β-1,4-linked polymer of *N*-acetylglucosamine whose structure resembles that of cellulose. In β chitin the individual parallel chains are linked by hydrogen bonds to form sheets in which parallel chains are held together by NH - - - O = C hydrogen bonds between the carboxamide groups. In the more abundant α chitin the chains in alternate sheets have opposite orientations,[101,102] possibly a result of hairpin folds in the strands. Native chitin exists as microfibrils of 7.25 nm diameter. These contain a 2.8-nm core consisting of 15–30 chitin chains surrounded by a sheath of 27-kDa protein subunits. The microfibrils pack in a hexagonal array, but the structure is not completely regular. Several proteins are present; some of the glucosamine units of the polysaccharide are not acetylated and the chitin core is often calcified.[103] The commercial product **chitosan** is a product of alkaline deacetylation of chitin but it also occurs naturally in some fungi.[102] Chitin is also present in cell walls of yeasts and other fungi. It is covalently bonded to a β-1,3-linked glycan which may, in turn, be linked to a mannoprotein (see Section D,2).[97]

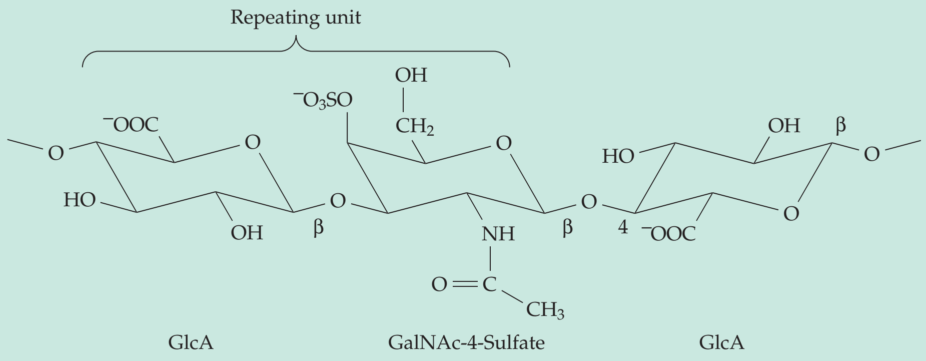
4. Heteropolysaccharides of the Animal Body

Many polysaccharides contain repeating units consisting of more than one different kind of monomer.[104,105] Some of these are composed of two sugars in a simple alternating sequence. Examples are **hyaluronan** (hyaluronic acid) and the **chondroitin**, **dermatan**, **keratan**, and **heparan sulfates**. They are important components of the "ground substance" or intracellular cement of connective tissue in animals. Hyaluronan,[106,107] which is abundant in synovial fluid and the vitreous humor of the eye, is a repeating polymer of glucuronic acid and *N*-acetylglucosamine with the structure shown in Fig. 4-11 and M_r of several million. The chondroitin sulfates and dermatan sulfate are similar polymers but with substitution by *N*-acetylgalactosamine and α-L-iduronic acid, respectively, and with sulfate ester groups in the positions indicated in Fig. 4-11.

Hyaluronan solutions are remarkably viscous and at a concentration of only 0.1% can have over 80% of the typical viscosity of biological fluids. This property may result from the presence of hydrogen bonds between the carboxylate, carboxamide, and hydroxyl groups of adjacent sugar residues as in Fig. 4-12. The hydrogen bonds stiffen the chain to give a slender rod. The tetrasaccharide shown in the figure is the repeating unit in a threefold helix.[109] However, the charged molecules do not associate to form strong fibrils like those of cellulose. The chain can be bent easily with breakage of the H-bonds at various positions to give a random coil structure.

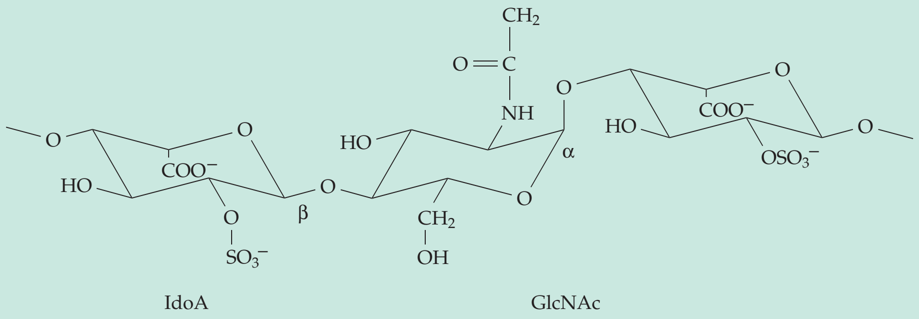
While hyaluronan is not covalently attached to proteins, it is usually anchored to cell surfaces and to

Type I Chain. Alternating 1→3 and 1→4 linkages; all β, (equatorial).


Repeating unit
$^-$OOC
HO
O
OH
β
^-O_3SO
OH
CH_2
O
NH
O=C
CH_3
β
O
4
$^-$OOC
HO
OH
β
O
GlcA
GalNAc-4-Sulfate
GlcA

Name	Repeating unit	Variations
Hyaluronan	[–4GlcAβ1→3GlcNAcβ1→]$_n$	None. Homogeneous.
Chondroitin sulfate	[–4GlcAβ1→3GalNAc(SO_3^-]β1→)$_n$	Some GlcA-2-SO_3^-, GalNAc-4- or 6-SO_3^-or both.
Dermatan sulfate	[–4IdoAα1→3GalNAc(4-SO_3^-]β1→)$_n$	Some L-IdoA-2-SO_3^-.
Keratan sulfate	[–3Galβ1→4GlcNAc(6-SO_3^-]β1→)$_n$	Some Gal-6-SO_3, some Sia, Man, Fuc.

Type II Chain. All 1→4 linkages, alternating β and α (axial).


CH_2
O=C
NH
O
HO
O
O
COO^-
O
SO_3^-
β
O
HO
CH_2
OH
α
O
HO
COO^-
O
O
OSO_3^-
IdoA
GlcNAc

Name	Repeating unit	Variations
Heparan sulfate	[–4-GlcAβ1→4GlcNAc]$_n$	Some IdoA, sulfation.
Heparin	[–4-L-IdoA(SO_3^-)α1→4GlcN(SO_3^-)α1→]$_n$	Some IdoA-2-SO_3^-, GlcNAc. 24 different disaccharides possible.

Figure 4-11 The repeating disaccharide units of hyaluronan and other glycosaminoglycans. See Fransson[108] and Hardingham and Fosang.[107]

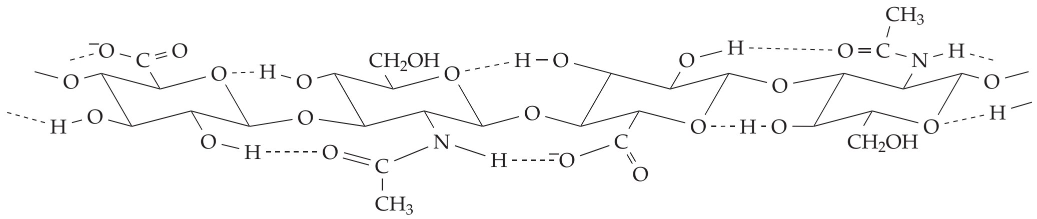

$^-$O
C=O
O
H
O
O
H
CH$_2$OH
O
H–O
O
H
O=C
CH$_3$
N
H
O
H–O
O
O
O
H
O
N
C
CH$_3$
H
$^-$O
C
O
O
H
O
CH$_2$OH
O
H

Figure 4-12 Proposed hydrogen-bonding scheme for the "native" conformation of hyaluronan. See Morris *et al.*[109]

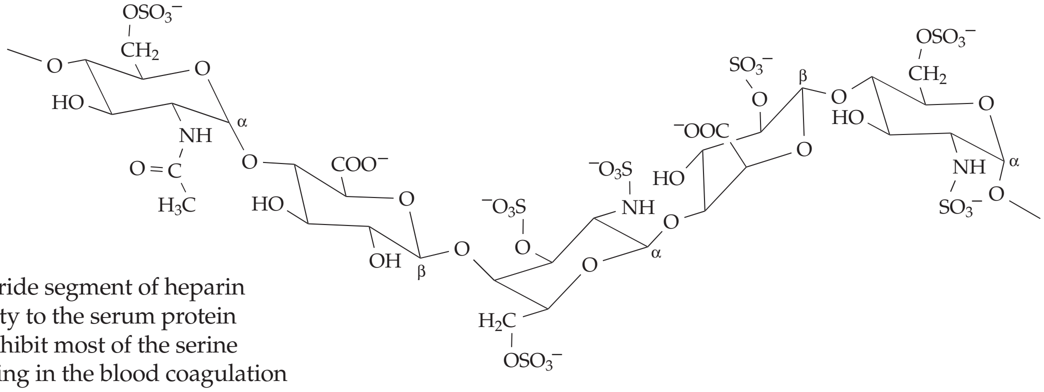


Figure 4-13 A pentasaccharide segment of heparin which binds with high affinity to the serum protein antithrombin causing it to inhibit most of the serine protease enzymes participating in the blood coagulation process (see Chapter 12). See Lindahl *et al*.[118]

---

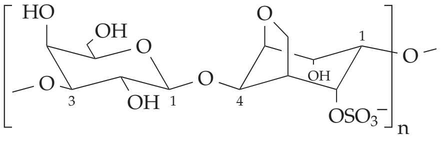
protein "receptors" within the extracellular matrix. In contrast, chondroitin, dermatan, keratan, and heparan sulfates are attached covalently to the proteins at the reducing ends of the polymer chains (see Section D). The attached polymers undergo enzyme-catalyzed chemical alteration. In dermatan most of the glucuronate residues found in chondroitin have been epimerized to iduronate and sulfate groups in ester linkages have been added. Chondroitin sulfate is especially abundant in cartilage; dermatan sulfate is concentrated in skin. Heparan sulfates are more heterogeneous than the other polymers of this group and have been described as "the most complex polysaccharides on the surface of mammalian cells."[110]

In the mast cells of lungs, liver, and other cells heparan sulfate chains are modified further and released into the bloodstream as **heparin**, a material with important anticoagulant properties. Both the amino groups and the 6-hydroxyls of the glucosamine residues of heparin carry sulfate groups. In some units D-glucuronic acid or glucuronate-2-sulfate[111] is present in α-1,4 linkage, but more often L-iduronic acid-2-sulfate is the first unit in the disaccharide.[112–114] The iduronate ring appears to have unusually high conformational flexibility which may influence the biological activity of heparin and related polysaccharides.[115] Because of its anticoagulant property, heparin is an important drug for prevention of blood clot formation.[116] In the United States, in 1976 six metric tons of heparin were used to treat 10 million patients.[117] The anticoagulant activity resides in large part in a nonrepeating pentasaccharide (Fig. 4-13).[112,118] This portion of the heparin, especially if part of a larger octasaccharide,[113] binds to several proteins including the enzyme inhibitors **heparin cofactor II** and **plasma antithrombin III** (Chapter 12). Heparin greatly accelerates the rate at which these proteins bind and inactivate blood clotting factors. See also Section D,1. Lower invertebrates, as well as marine brown algae, contain heavily sulfated fucans which are largely 1,3-linked.[119]

5. Plant Heteropolysaccharides

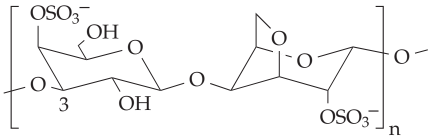
Fibers of cellulose, which run like rods through the amorphous matrix of plant cell walls, appear to be coated with a monolayer of **hemicelluloses**. Predominant among the latter is a **xyloglucan**, which has the basic cellulose structure but with α-1,6-linked xylose units attached to three-fourths of the glucose residues.[104,120,121] L-Fucose may also be present in trisaccharide side chains: L-Fucα1→2Galβ1→ 2Xylα1→.[122] **Pectins** of higher plants contain β-1,4-linked polygalacturonates interrupted by occasional 1,2-linked L-rhamnose residues. Some of the carboxyl groups of these **rhamnogalacturonan** chains are methylated.[123] **Arabinans** and **galactans** are also present in pectin. A possible arrangement of cellulose fibers, hemicelluloses, and pectic materials in a cell wall has been proposed (Fig. 4-14).

Agarose, an alternating carbohydrate polymer consisting of ~120-kDa chains, is the principal component of agar and the compound that accounts for most of the gelling properties of that remarkable substance. A solid agar gel containing 99.5% water can be formed. Agarose molecules form left-handed double helices with a threefold screw axis, a pitch of 1.90 nm, and a central cavity containing water molecules.[124,125]



Agarose:
[-3-D-Galβ1→4-(3,6-anhydro)-L-Galα1→]$_n$

A similar structure has been established for the gel-forming **carrageenans** from red seaweed. The X-ray data suggest that three of the disaccharide units form one turn of a right-handed helix with a pitch of 2.6 nm. A second chain with a parallel orientation, but displaced by half a turn, wraps around the first helix.[124] Such double-helical regions provide “tie points” for the formation of gels (Fig. 4-10).[104,127,128]

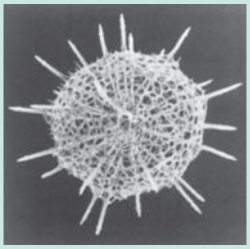


ι-Carrageenan:
[3Gal (4 sulfate) β1→4 (3,6-anhydro) Gal (2-sulfate) α1→]$_n$

Sulfate groups protrude from the structure in pairs and provide binding sites for calcium ions, which stabilize the gel. The presence of occasional extra sulfate groups in these polymers causes kinks in the chains because the derivatized pyranose rings reverse their conformation to the other chair form. This prevents the entire polysaccharide chain from assuming a regular helical structure.[100,104]

Alginates, found in cell walls of some marine algae and also formed by certain bacteria, consist in part of a linear β-1,4-linked polymer of D-mannuronate with a cellulose-like structure. Alginates also contain

BOX 4-B SILICON: AN ESSENTIAL TRACE ELEMENT

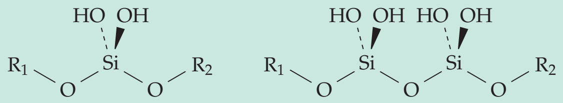


Silica skeleton of a radiolarian[e]

No one can doubt that diatoms, which make their skeletons from SiO_2, have an active metabolism of silicon. They can accumulate as much as 0.7 µM soluble silicon, possibly attached to proteins.[a-d]

The radiolaria and sponges often accumulate silicon and limpets make opal base plates for their teeth. Silicon may account for as much as 4% of the solids of certain grasses. Although silicon is usually not considered an essential nutrient for all plants, there is much evidence that it is essential to some and that it is often beneficial.[f] Silicon is found in soil primarily as silicic acid, H_4SiO_4, whose concentration ranges from 0.1–0.6 mM. Most of the silicon taken up by plants is deposited within cells, in cell walls, between cells, or in external layers as hydrated SiO_2. Presumably the organic components of the plant control the deposition. The SiO_2 in some plants takes the form of sharp particles which may have a defensive function. They abrade the enamel surfaces of the teeth of herbivores and can cause other illnesses.[g]

Silicon is essential for growth and development of higher animals,[h-l] and it has been suggested that humans may require 5–20 mg per day.[m] In the chick, silicon is found in active calcification sites of young bone.[i] Silicon-deficient animals have poorly calcified bones and also an elevated aluminum content in their brains.[m] Silicon is present in low amounts in the internal organs of mammals but makes up ~0.01% of the skin, cartilage, and ligaments, in which it is apparently bound to proteoglycans such as chondroitin-4-sulfate, dermatan sulfate, and heparan sulfate (Fig. 4-11).[m,n] These polymers contain ~0.04% silicon or one atom of silicon per 130–280 repeating units of the polysaccharides. Plant pectins contain about five times this amount. The silicon is apparently bound tightly in *ether* linkage. Perhaps orthosilicic acid, $Si(OH)_4$, reacts with hydroxyl groups of the carbohydrates to form bridges between two chains as follows:



In each of these formulas additional free OH groups are available on the silicon so that it is possible to crosslink more than two polysaccharide chains. Silicon may function as a biological crosslinking agent in connective tissue. **Silaffins**, small polypeptides containing polyamine side chains of modified lysine residues, apparently initiate silica formation from silicic acid in diatoms.[o]

---

[a] Robinson, D. H., and Sullivan, C. W. (1987) *Trends Biochem. Sci.* **12**, 151–154

[b] Round, F. E. (1981) in *Silicon and Siliceous Structures in Biological Systems* (Simpson, T. L., and Volcani, B. E., eds), pp. 97–128, Springer, New York

[c] Evered, D., and O’Connor, M. (1986) *Silicon Biochemistry*, Wiley, New York

[d] Round, F. E., Crawford, R. M., and Mann, D. G. (1990) *The Diatoms*, Cambridge Univ. Press, Cambridge UK

[e] Buchsbaum, R., Buchsbaum, M., Pearse, J., and Pearse, V. (1987) *Animals Without Backbones*, 3rd ed., Univ. Chicago Press, Chicago

[f] Epstein, E. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11–17

[g] McNaughton, S. J., and Tarrants, J. L. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 790–791

[h] Schwarz, K. (1970) in *Trace Element Metabolism in Animals* (Mills, F., ed), pp. 25–38, Livingstone, Edinburgh, UK

[i] Schwarz, K., and Milne, D. B. (1972) *Nature* **239**, 333–334

[j] Carlisle, E. M. (1972) *Science* **278**, 619–621

[k] Hoekstra, W. H., Suttie, J. W., Ganther, H. E., and Mertz, W., eds. (1974) *Trace Element Metabolism in Animals-2*, University Park Press, Baltimore, Maryland

[l] Carlisle, E. M. (1988) *Science Total Environment* **73**, 95–106

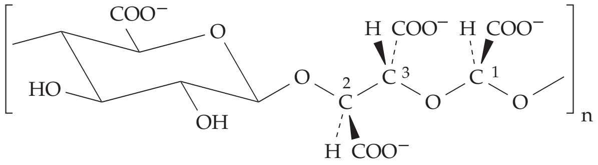
[m] Nielsen, F. H. (1991) *FASEB J.* **5**, 2661–2667

[n] Schwarz, K. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1608–1612

[o] Kröger, N., Deutzmann, R., and Sumper, M. (1999) *Science* **286**, 1129–1132

α-L-guluronate, sometimes in homopolymeric "blocks" and sometimes alternating with the mannuronate residues. Groups of adjacent guluronate units are thought to impart calcium-binding properties to alginates.[129]

Polysaccharides with calcium-binding sites may also serve to initiate deposition of calcium carbonate. For example, the unicellular alga *Pleurochrysis carterae* contains an unusual polysaccharide with the following highly negatively charged repeating unit:



[-4GlcAβ1→2-*meso*-tartrate-3→1-glyoxylate-1→]$_n$

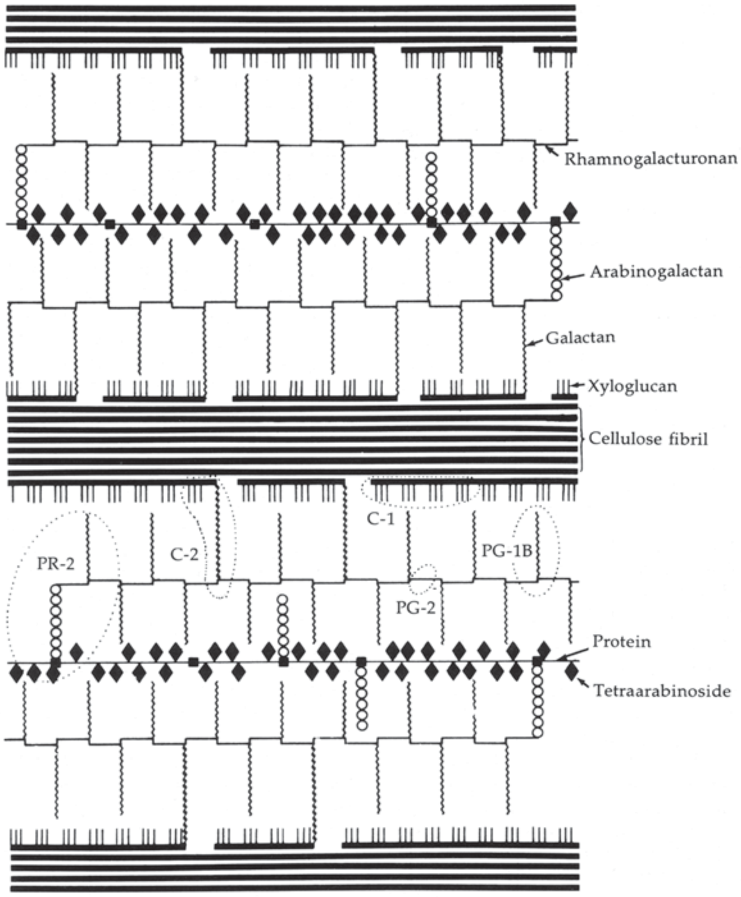
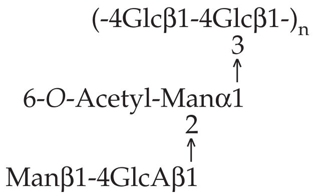


Figure 4-14 Tentative structure of the walls of suspension-cultured sycamore cells. The wall components are in approximately proper proportions but the distance between cellulose elementary fibrils is expanded to allow room to present the interconnecting structure. There are probably between 10 and 100 cellulose elementary fibrils across a single primary cell wall. From Albersheim *et al.*[126]

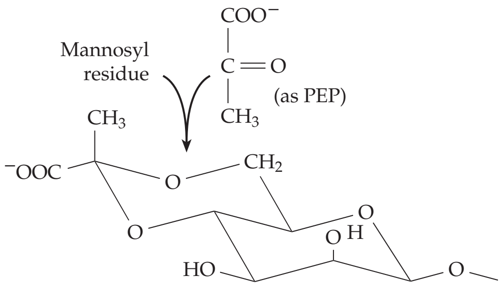
Meso-tartrate is joined to glucuronic acid in glycosidic linkage and by acetal formation to the aldehyde **glyoxylic acid**, HOOC—CHO, which is also joined in an ether linkage to the next repeating unit.[130] Similar **open acetal linkages** join monosaccharide units in some bacterial lipopolysaccharides and may occur more widely.[130a]

Bacteria form and secrete a variety of heteropolysaccharides, several of which are of commercial value because of their useful gelling properties. **Xanthan gum** (formed by *Xanthomonas campestris*) has the basic cellulose structure but every second glucose residue carries an α-1,3-linked trisaccharide consisting of 6-*O*-acetylmannose, glucuronic acid, and mannose in the following repeating unit:[131,132]



$$\begin{array}{c} (\text{-4Glc}\beta 1\text{-4Glc}\beta 1\text{-})_n \\ 3 \\ \uparrow \\ 6\text{-}O\text{-Acetyl-Man}\alpha 1 \\ 2 \\ \uparrow \\ \text{Man}\beta 1\text{-4GlcA}\beta 1 \end{array}$$

The polymer is further modified by reaction of about half of the mannosyl residues with pyruvate to form ketals (Eq. 4-9).

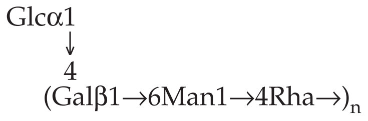


(4-9)

Acetan of *Acetobacter xylinum* has pentasaccharide side chains that contain L-rhamnose.[122] A helical structure for the strands has been observed by atomic force microscopy.

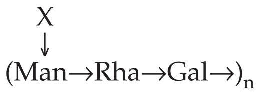
6. Polysaccharides of Bacterial Surfaces

The complex structure of bacterial cell walls is discussed in Chapter 8. However, it is appropriate to mention a few bacterial polysaccharides here. The innermost layer of bacterial cell walls is a porous network of a highly crosslinked material known as **peptidoglycan** or **murein** (see Fig. 8-29). The backbone of the peptidoglycan is a β-1,4-linked alternating polymer of *N*-acetyl-D-glucosamine and *N*-acetyl-D-muramic acid. Alternate units of the resulting chitin-like molecule carry unusual peptides that are attached to the lactyl groups of the *N*-acetyl-muramic acid units and crosslink the polysaccharide chains. In *E. coli* and other gram-negative bacteria the peptidoglycan forms a thin (2 nm) continuous network around the cell. This "baglike molecule" protects the organism from osmotic stress. In addition, gram-negative bacteria have an outer membrane and on its outer surface a complex lipopolysaccharide. The projecting ends of the lipopolysaccharide molecules consist of long carbohydrate chains with repeating units that have antigenic properties and are called **O antigens**. Specific antibodies can be prepared against these polysaccharides, and so varied are the structures that 1000 different "serotypes" of *Salmonella* are known. These are classified into 17 principal groups. For example, group E3 contains the following repeating unit, where *n* may be ~50 on the average. Rha = L-rhamnose.



$$\begin{array}{c}\text{Glc}\alpha 1\\ \downarrow\\ 4\\ (\text{Gal}\beta 1\to 6\text{Man}1\to 4\text{Rha}\to)_n\end{array}$$

Polysaccharides of groups A, B, and D contain the repeating unit



$$\begin{array}{c}\text{X}\\ \downarrow\\ (\text{Man}\to\text{Rha}\to\text{Gal}\to)_n\end{array}$$

where X is a 3,6-dideoxyhexose: **paratose** in type A, **abequose** in type B, and **tyvelose** in type D (Fig. 4-15). The existence of the many serotypes depends on the variety of components, on the many types of linkage (α and β, 1→2, 1→3, 1→4, and 1→6) in the repeating units, and on further structural variations at the chain ends.

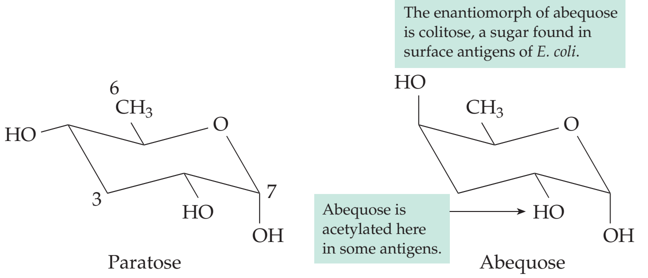
At the inner end of the O antigen is a shorter polysaccharide "core" whose structure is less varied than that of the outer ends but which is remarkable in containing two sugars found only in bacterial cell walls: a seven-carbon heptose and an eight carbon α-oxo sugar acid, **ketodeoxyoctonate** (KDO). The structures are given in Fig. 4-15 and the arrangement of these sugars in the *Salmonella* lipopolysaccharide is shown in Fig. 8-30. That figure also shows the manner in which the oligosaccharide that bears the O antigen is attached to a lipid anchoring group that is embedded in the outer membrane of the bacteria.

A great variety of polysaccharides are present in the outer layers of other types of bacteria. For example, the mycobacteria have an alternating 5- and 6-linked β-D-Gal*f* polymer attached to their peptidoglycan. Attached to this galactan are branched penta-D-arabinose units:

$$(\text{Ara}f\beta 1\to 2\text{Ara}f\alpha 1)_2\to 3{,}5\text{Ara}f\alpha\text{—}$$

These are further modified by esterification with mycolic acids.[133] Information on some other cell wall components is given in Chapter 20.

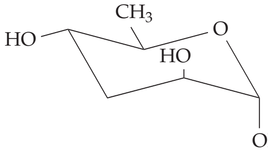
The enantiomorph of abequose is colitose, a sugar found in surface antigens of *E. coli*.



Abequose is acetylated here in some antigens.

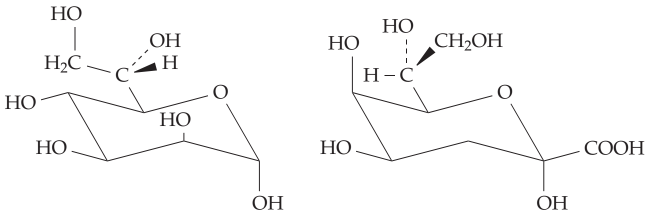
Paratose (3,6-dideoxy-D-ribohexose)

Abequose (3,6-dideoxy-D-xylohexose)



The enantiomorph of tyvelose is ascarylose, found not only in certain bacterial antigens but also in the eggs of *Ascaris*.

Tyvelose (3,6-dideoxy-D-arabohexose)



L-*Glycero*-D-*manno*heptose

3-Deoxy-D-mannooctulosonic acid, also called 2-keto-3-deoxyoctonate or KDO

Figure 4-15 Structures of special sugars found in the "antigens" of the outer cell walls of gram-negative bacteria.

D. Glycoproteins and Proteoglycans

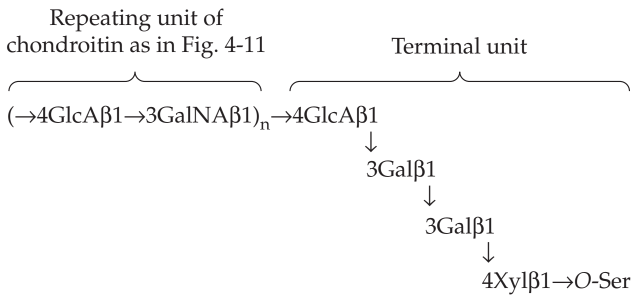
Many proteins, including almost all of those that are secreted from cells and many that are components of cell surfaces, carry covalently attached oligosaccharides.[134–139] These **glycoproteins** may carry just one or a few, often highly branched, oligosaccharide chains. For example, ribonuclease B has a structure identical to that of ribonuclease A (Fig. 12-25) except for the presence of an oligosaccharide on asparagine 34.[140] In other instances proteins carry a large number of such chains and the carbohydrate may account for over half of the mass of a glycoprotein. Most carbohydrate chains are attached either as *O*-glycosides with the hydroxyl groups of the side chains of serine, threonine, or other hydroxyamino acid residues or as *N*-glycosyl groups through linkage to the amide groups of asparagine side chains. Both types of linkage may be present in a single protein. Here are some examples.

Xylβ1→*O*-Ser(Thr)	Proteoglycans of connective tissue; thyroglobulins
$\to\underset{H}{\overset{\varepsilon}{N}}$—Lys	Some dermatan sulfates
Galβ1→*O*-Hydroxylysine (Hydroxyproline)	Collagen, extension
L-Araα1→*O*-4-Hydroxy-proline	Plants
GalNAcα1→*O*-Ser(Thr)	Many glycoproteins
GlcNAcα1→*O*-Ser(Thr)	Glycoproteins of cytoplasmic surfaces
GlcNAcβ1→NH-βCH_2-Asn	Many glycoproteins Some dermatan and heparan sulfates

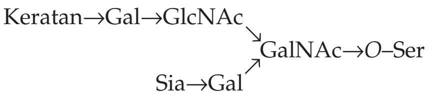
Linkage of a glycosyl group to a carbon atom of an indole ring of tryptophan has also been demonstrated.[141]

1. *O*-Linked Glycosyl Groups

In the *O*-linked glycoproteins the sugar that is attached directly to the protein is usually either xylose, galactose, or *N*-acetylgalactosamine, all in the pyranose ring form. Xylose is found only in the intercellular **proteoglycans** which carry the chondroitin, dermatan, and related sulfated polysaccharide chains of connective tissues.[141a,b] Since amino sugars are a major constituent, proteoglycans are often called **glycosaminoglycans**. Chondroitin, dermatan, and heparan sulfates are all attached to "core" proteins by the same linkage, which is illustrated here for chondroitin

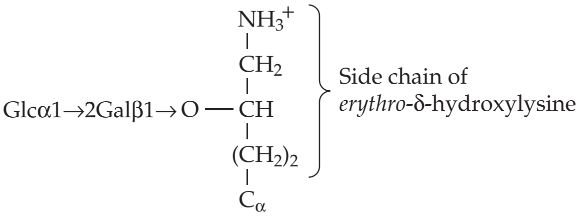

Repeating unit of chondroitin as in Fig. 4-11
Terminal unit
(→4GlcAβ1→3GalNAβ1)n→4GlcAβ1
↓
3Galβ1
↓
3Galβ1
↓
4Xylβ1→O-Ser

Some IdoA may be present in the terminal unit of dermatan sulfate.[142] Keratan sulfate has its own core proteins[143] and has different terminal units including the following:[107]



The large proteoglycan of human cartilage is built upon the 246 kDa protein **aggrecan**. In the central half of the peptide chain are many Ser-Gly sequences to which about one hundred 10- to 25-kDa chondroitin sulfate chains are attached. About 30 keratan sulfate chains as well as other oligosaccharide groups are also present. These proteoglycan subunits are joined with the aid of 44- to 49-kDa **link protein** to molecules of hyaluronan[144–146] (Fig. 4-16). Several types and sizes of proteoglycan are known.[143,145–147] Dermatan sulfates may be linked to these through either serine or asparagine, depending upon the tissue. The polysaccharide chains of the proteoglycans also bind to collagen fibrils to form a "fiber-reinforced composite material" between cells. Chondroitin and heparan sulfates may be attached at different Ser-Gly sites in a single peptide chain.[148] Degradation of heparan proteoglycans may lead to the shorter free carbohydrate chains found in the circulating heparin. Commercial heparin preparations used as anticoagulants are produced by oxidative destruction of the attached proteins.[113]

A quite different situation holds for **collagen** in which β-galactosyl units and glucosyl-β-galactosyl disaccharide units are attached to side chains of hydroxylysine formed by postsynthetic modification of the original procollagen chain.



A great deal of variation in the amount of glycosylation is observed from one species to another. The human α-1(II) chains of collagen usually carry four disaccharides and four monosaccharide units. In the related collagen-like **extensins**, which are found in plant cell walls, the hydroxyproline (Hyp) side chains are *O*-glycosylated, largely by short oligosaccharides of arabinose in furanose ring form,[150–152] e.g., Ara*f* β1→2Ara*f* β1→2Ara*f* β1→4Hyp. There are as many as 25 repeats of Ser-Pro-Pro-Pro-Pro encoded in an extensin gene. Most of the prolines are hydroxylated and glycosylated. The presence of two or more contiguous proline residues seems to be the signal for the hydroxylation reaction to take place.[152]

Another distinct family of *O*-linked glycoproteins are the **mucins**, which are present in saliva and other

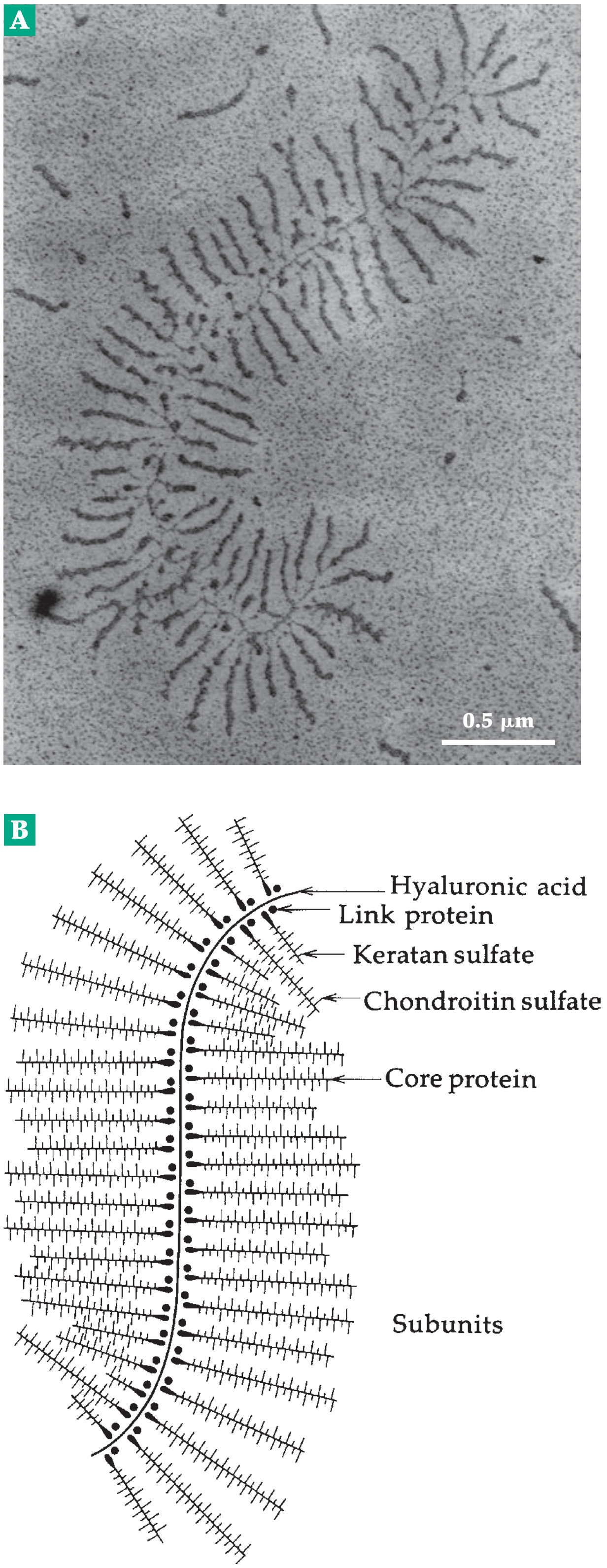


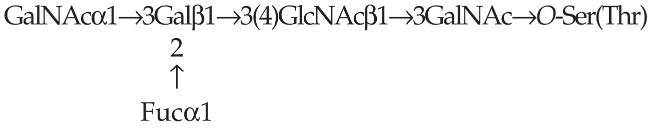
Figure 4-16 (A) Dark field electron micrograph of a proteoglycan aggregate from bovine articular cartilage (from bearing surfaces of joints). Courtesy of Joseph A. Buckwalter. The filamentous backbone consists of hyaluronic acid, as in (B). The proteoglycan subunits extend from the backbone. From Rosenberg.[149]

mucous secretions. The polypeptides consist largely of serine, threonine, proline, and glycine, with up to one-third of the residues being Ser + Thr. Mucins may contain 70 to 85% carbohydrate linked through *N*-acetylgalactosamine residues to the serine and threonine hydroxyls.[153–160a] Salivary mucins of cows, sheep, and dogs contain largely the following disaccharide, which may contain either *N*-acetyl- or *N*-glycolyl-neuraminic acid:

$$\text{NeuNAc}\alpha 2\rightarrow 6\text{GlcNAc}\beta\rightarrow\text{Ser(Thr)}$$

Up to 800 disaccharides are present on a single large protein which is composed of disulfide-linked subunits. The many negatively charged sialic acid (*N*-glycolylneuraminic acid) groups are thought to cause expansion and rigidity which increase the viscosity of the protein. Some other mucins contain predominantly large oligosaccharides, which often carry **blood group determinants**.[161,162]

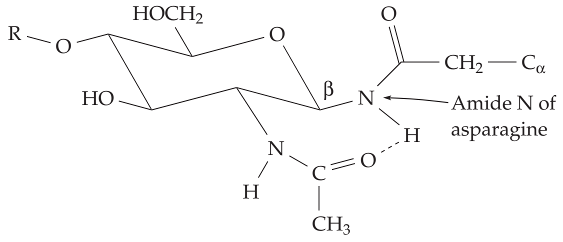
The ABO(H) family of blood group determinants (Box 4-C) are oligosaccharide groups assembled from D-galactose, *N*-acetyl-D-galactosamine, and L-fucose. They are carried on the nonreducing ends of *O*-linked oligosaccharides which may be attached to cell surface proteins, to mucins, or to the sphingolipid known as ceramide (Chapter 8). Attachment to proteins is usually via GalNAc. An example of a blood type A determinant on an *O*-linked oligosaccharide is the following, where the linkage between the galactose and *N*-acetylglucosamine may be either β-1,3 or β-1,4:



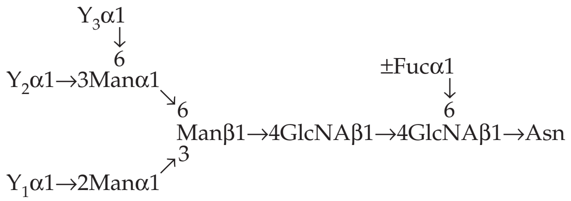
$$\text{GalNAc}\alpha 1\rightarrow 3\underset{\substack{2\\ \uparrow\\ \text{Fuc}\alpha 1}}{\text{Gal}\beta 1}\rightarrow 3(4)\text{GlcNAc}\beta 1\rightarrow 3\text{GalNAc}\rightarrow O\text{-Ser(Thr)}$$

2. Asparagine-Linked Oligosaccharides

In many glycoproteins oligosaccharides are linked through *N*-acetylglucosamine to the side chain of asparagine.[134,163,164]



This structure also illustrates one of the hydrogen-bonding possibilities through which the sugar can interact with the protein. Since the site of glycosylation is often at β bends in the surface of the protein, the amide groups of the *N*-acetylglucosamine may alternatively hydrogen bond to amide groups of the peptide.[165] This asparagine linkage is very common. For example, it is present in 97% of glycoproteins of blood plasma[166] and it is also predominent in the glycoproteins of tissue surfaces.[167] There are numerous structures for asparagine-linked oligosaccharides but many contain the following core to which additional glycosyl groups may be attached:



$$\begin{array}{l}
\quad\; Y_3\alpha 1 \\
\qquad \downarrow \\
\qquad 6 \\
Y_2\alpha 1\rightarrow 3\mathrm{Man}\alpha 1 \qquad\qquad\qquad\qquad \pm\mathrm{Fuc}\alpha 1\\
\qquad\qquad\qquad \searrow_6 \qquad\qquad\qquad\qquad\qquad\quad \downarrow 6 \\
\qquad\qquad\qquad\quad \mathrm{Man}\beta 1\rightarrow 4\mathrm{GlcNA}\beta 1\rightarrow 4\mathrm{GlcNA}\beta 1\rightarrow \mathrm{Asn}\\
\qquad\qquad\qquad \nearrow^3 \\
Y_1\alpha 1\rightarrow 2\mathrm{Man}\alpha 1
\end{array}$$

Here the ±Fuc indicates that this residue is present only on some of the chains.

Notice the three mannose residues on the left side. In the **high mannose type** oligosaccharide, Y_1, Y_2, and Y_3 are additional mannose units. In many instances $Y_1 = Y_2 = Y_3$ = Manα→Man. These "extra" mannose units are put onto the oligosaccharide during the original biosynthesis and before it is attached to the protein (Chapter 20). Some of the mannose units may then be removed during the "processing" of the oligosaccharide in the endoplasmic reticulum and residues of glucosamine, galactose, and sialic acid (Sia) may be added. Thus, Y_1 and Y_2 in the foregoing structure often become

Siaα2→3(6)Galβ1→4GlcNAcβ1→

Here, the sialic acid may be linked either 2,3 or 2,6 and the GlcNAc either 1,2 or 1,3. Both Y groups may consist of trisaccharides of this type in "biantennary" oligosaccharides and a third trisaccharide (Y_3) may be added to form a "triantennary" molecule. An additional *N*-acetylglucosamine is often linked by β-1,4 linkage to the central mannose of the core and fucosyl residues in α-1,6 linkage are often linked to the *N*-acetylglucosamine next to the asparagine.[168] The Galβ1→4GlcNAc disaccharide unit in the above Y group is also called ***N*-acetyllactosamine** because of its relationship to lactose. It is often repeated in long Y groups, e.g., as $(\rightarrow 3\mathrm{Gal}\beta 1\rightarrow 4\mathrm{Glc\text{-}NAc}\beta 1—)_n$, the oligosaccharides being called poly-*N*-acetyllactosamino-glycans. These structures are also principal carriers of the ABO blood type determinants on erythrocyte surfaces.[169]

Many *N*-linked oligosaccharides are highly branched. For example, in ovomucoid, a protease inhibitor of hen eggs, "pentaantennary" oligosaccharides have two and three *N*-acetylglucosamine rings, respectively, attached to the terminal mannose units of the oligosaccharide core in 1,2, 1,4 and 1,6 linkages. Another large

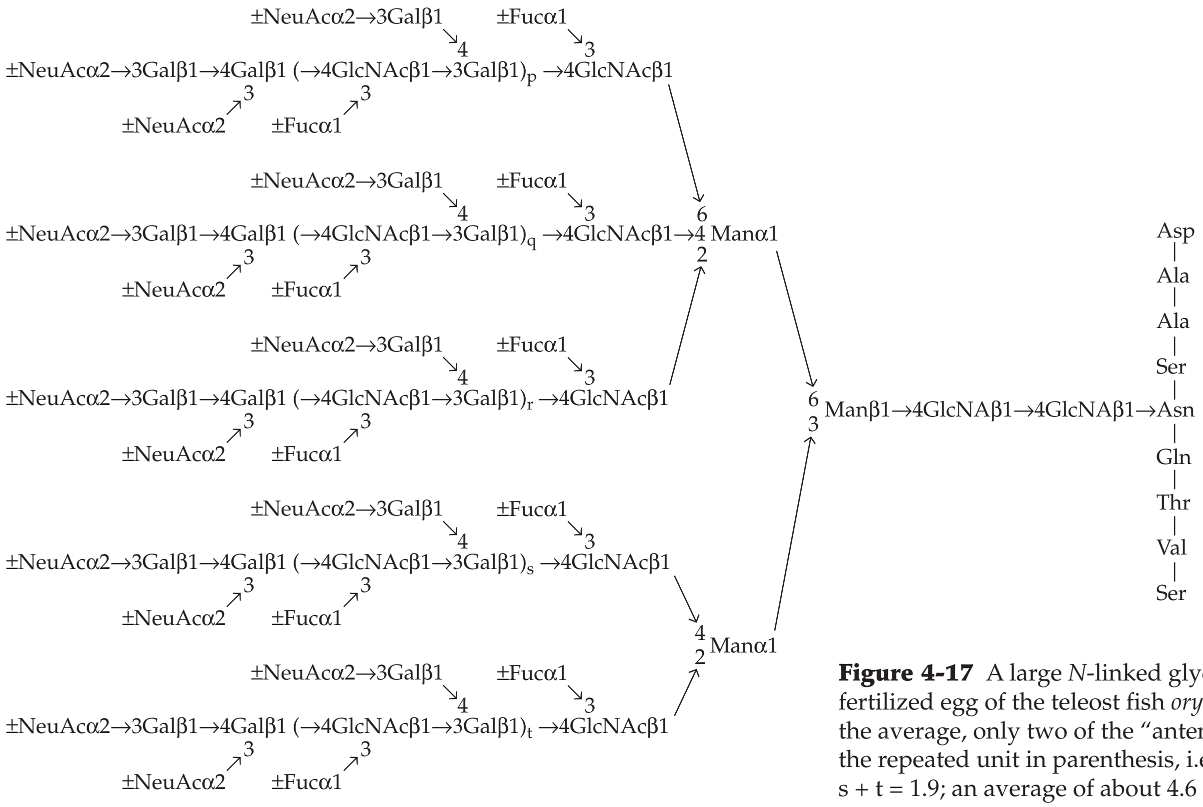

±NeuAcα2→3Galβ1 ±Fucα1
4 3
±NeuAcα2→3Galβ1→4Galβ1 (→4GlcNAcβ1→3Galβ1)p →4GlcNAcβ1
3 3
±NeuAcα2 ±Fucα1
±NeuAcα2→3Galβ1 ±Fucα1
4 3
±NeuAcα2→3Galβ1→4Galβ1 (→4GlcNAcβ1→3Galβ1)q →4GlcNAcβ1→4 Manα1
6
2
3 3
±NeuAcα2 ±Fucα1
±NeuAcα2→3Galβ1 ±Fucα1
4 3
±NeuAcα2→3Galβ1→4Galβ1 (→4GlcNAcβ1→3Galβ1)r →4GlcNAcβ1
3 3
±NeuAcα2 ±Fucα1
6
3
Manβ1→4GlcNAβ1→4GlcNAβ1→Asn
Asp
Ala
Ala
Ser
Gln
Thr
Val
Ser
±NeuAcα2→3Galβ1 ±Fucα1
4 3
±NeuAcα2→3Galβ1→4Galβ1 (→4GlcNAcβ1→3Galβ1)s →4GlcNAcβ1
3 3
±NeuAcα2 ±Fucα1
4
2
Manα1
±NeuAcα2→3Galβ1 ±Fucα1
4 3
±NeuAcα2→3Galβ1→4Galβ1 (→4GlcNAcβ1→3Galβ1)t →4GlcNAcβ1
3 3
±NeuAcα2 ±Fucα1

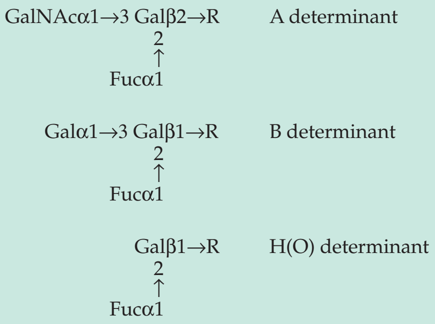
Figure 4-17 A large *N*-linked glycan from the fertilized egg of the teleost fish *oryzias*.[171] On the average, only two of the "antenna" contain the repeated unit in parenthesis, i.e., p + q + r + s + t = 1.9; an average of about 4.6 residues of sialic acid are present.

BOX 4-C THE BLOOD GROUP DETERMINANTS

The role of carbohydrates in biological communication is well illustrated by the human blood types.[a,b] According to the ABO system first described by Landsteiner in 1900, individuals are classified into types A, B, AB, and O. Blood of individuals of the same type can be mixed without clumping of cells, but serum from a type O individual contains antibodies that agglutinate erythrocytes of persons of types A and B. Serum of persons of type B causes type A cells to clump and vice versa. Individuals of none of the four types have antibodies against type O erythrocytes. For this reason, persons with type O blood are sometimes inaccurately described as "universal donors."

The ABO blood types are determined by specific **blood group determinants** which are attached to the nonreducing ends of *O*-linked oligosaccharides of surface glycoproteins, mucins, glycolipids, and, to a lesser extent, *N*-linked oligosaccharides. The blood group determinants are found on erythrocytes and all endothelial cells of the body. In about 80% of the population they are also present on glycoproteins of the saliva and other secretions.

The minimal determinant structures, attached to "carrier" R, are as follows:



$$\mathrm{GalNAc}\alpha1\rightarrow3\ \underset{\substack{2\\ \uparrow\\ \mathrm{Fuc}\alpha1}}{\mathrm{Gal}\beta2}\rightarrow\mathrm{R}$$

A determinant

$$\mathrm{Gal}\alpha1\rightarrow3\ \underset{\substack{2\\ \uparrow\\ \mathrm{Fuc}\alpha1}}{\mathrm{Gal}\beta1}\rightarrow\mathrm{R}$$

B determinant

$$\underset{\substack{2\\ \uparrow\\ \mathrm{Fuc}\alpha1}}{\mathrm{Gal}\beta1}\rightarrow\mathrm{R}$$

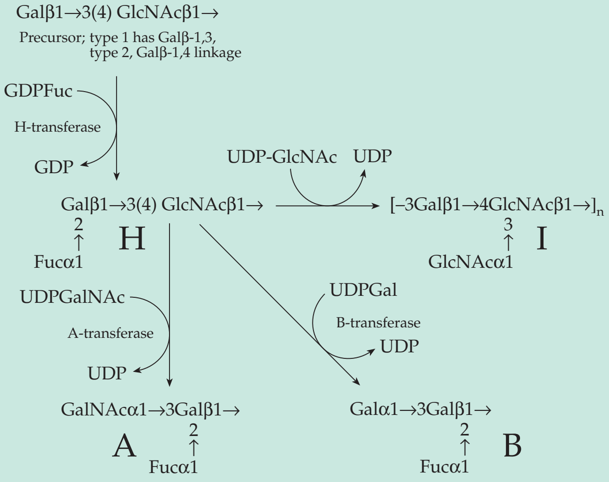
H(O) determinant

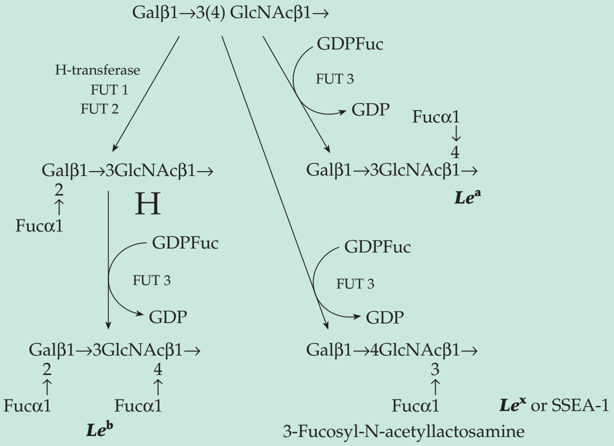
Here R refers to a "carrier oligosaccharide" which can be as simple as

→3(4)GlcNAcβ1→3Galβ1→4Glcβ1—*O*—Ser (Thr, ceramide)

or may be a complex lactosaminoglycan or an oligosaccharide such as that in Fig. 4-17. The minimal determinants can be linked to the carrier oligosaccharide by either a β-1,3 (type I chain) or β-1,4 (type II chain) glycosidic bond.

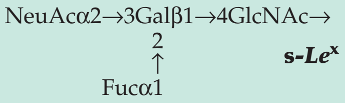
The genetic basis for the ABO blood groups is well understood. There are three **alleles**, variants of a gene, that encodes a **glycosyltransferase**. In A type individuals, this enzyme transfers *N*-acetylgalactosamine from a carrier molecule, called UDP (Chapter 17), onto the terminal positions of the H(O) determinant. The enzyme specified by the *B* allele transfers galactose. The two enzymes differ in only four amino acid residues but the result is an altered substrate specificity.[a,c,d] The *O* allele produces inactive enzyme as a result of a single base deletion in the gene.[a] The H gene has been identified as that of an α-1,2 fucosyltransferase that transfers α-L-fucose from the carrier GDP to the galactose unit in the foregoing structure.[e] Persons with an inactive *H* transferase gene may have the rare type I, which results from addition of glucosamine branches to the repetitive H antigen structures of polylactosaminoglycans.[f] More often though, the H antigen is acted upon by a different fucosyltransferase encoded by the *Le* gene, which determines the **Lewis blood group**. This enzyme places α-1,4-linked L-fucose onto the H antigen to give the *Le*^{b} antigen. The same fucosyltransferase (which is


Galβ1→3(4) GlcNAcβ1→
Precursor; type 1 has Galβ-1,3, type 2, Galβ-1,4 linkage
GDPFuc
H-transferase
GDP
UDP-GlcNAc
UDP
Galβ1→3(4) GlcNAcβ1→
2
Fucα1
H
[–3Galβ1→4GlcNAcβ1→]n
3
GlcNAcα1
I
UDPGalNAc
A-transferase
UDP
UDPGal
B-transferase
UDP
GalNAcα1→3Galβ1→
2
Fucα1
A
Galα1→3Galβ1→
2
Fucα1
B


Galβ1→3(4) GlcNAcβ1→
H-transferase
FUT 1
FUT 2
GDPFuc
FUT 3
GDP
Fucα1
4
Galβ1→3GlcNAcβ1→
Lea
Galβ1→3GlcNAcβ1→
2
Fucα1
H
GDPFuc
FUT 3
GDP
GDPFuc
FUT 3
GDP
Galβ1→3GlcNAcβ1→
2
4
Fucα1
Fucα1
Leb
Galβ1→4GlcNAcβ1→
3
Fucα1
Lex or SSEA-1
3-Fucosyl-N-acetyllactosamine

BOX 4-C Continued

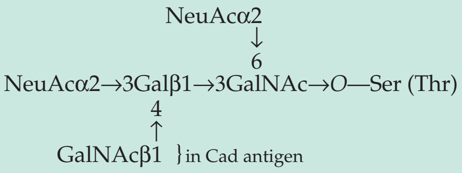
designated FUT 3) acting on type 2 precursor chains forms the Le^{a} antigen. The same transferase adds a β-1,3-linked fucose to type 2 chains to give the Le^{x} antigen, which is also called SSEA-1.[g,h] Lactosamine-type precursors carrying α-2,3-linked sialic acid at their nonreducing ends on surfaces of granulocytes, monocytes, and natural killer cells can be converted to sialylated Le^{x} (s-Le^{x}).[i,j]



$$\text{NeuAc}\alpha 2\rightarrow 3\underset{\substack{2\\ \uparrow\\ \text{Fuc}\alpha 1}}{\text{Gal}\beta 1}\rightarrow 4\text{GlcNAc}\rightarrow \qquad \textbf{s-}\boldsymbol{Le^{x}}$$

Individuals with active gene *Se* (for secretion) secrete glycoproteins bearing the blood group substance into saliva and other body fluids. The *Se* gene is another H-transferase present in epithelial cells and salivary glands. In individuals with active *Se* genes (~80% of most populations) most soluble H antigens are converted to Le^{b} oligosaccharides.[k]

There are at least 12 other well established human blood groups, several of which involve oligosaccharides attached to glycoproteins or glycolipids. The MN antigens consist of a sequence of amino acids near the N terminus of the protein glycophorin (Chapter 8) with attached sialic acid-containing *O*-linked oligosaccharides,[l] e.g., as in the following structure:



$$\text{NeuAc}\alpha 2\rightarrow 3\underset{\substack{4\\ \uparrow\\ \text{GalNAc}\beta 1\ \}\ \text{in Cad antigen}}}{\text{Gal}\beta 1}\rightarrow 3\overset{\substack{\text{NeuAc}\alpha 2\\ \downarrow\\ 6}}{\text{GalNAc}}\rightarrow O\text{—Ser (Thr)}$$

The Cad antigen, also found on glycophorin as well as on gangliosides, has an additional β-1,4-linked GalNAc at the indicated position.[m] The Kell blood group antigens are carried on a 93 kDa glycoprotein of erythrocyte surfaces.[n] The P blood group depends upon surface carbohydrates such as the tetrasaccharide part of the ganglioside called globoside; see Fig. 20-9. Its characteristic antigenic activities are destroyed by treatment with dilute periodate.

Other antigens, including those of the Rh and LW groups, are represented by exposed parts of proteins on the erythrocyte surfaces.[o]

---

[a] Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* **345**, 229–233

[b] Frevert, J., and Ballou, C. E. (1985) *Biochemistry* **24**, 753–759

[c] Feizi, T. (1990) *Trends Biochem. Sci.* **15**, 330–331

[d] Yamamoto, F., and Hakomori, S. (1990) *J. Biol. Chem.* **265**, 19257–19262

[e] Larsen, R. D., Ernst, L. K., Nair, R. P., and Lowe, J. B. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6674–6678

[f] van den Eijnden, D., Koenderman, A., and Schiphorst, W. (1988) *J. Biol. Chem.* **263**, 12461–12471

[g] Mollicone, R., Reguigne, I., Kelly, R. J., Fletcher, A., Watt, J., Chatfield, S., Aziz, A., Cameron, H. S., Weston, B. W., Lowe, J. B., and Oriol, R. (1994) *J. Biol. Chem.* **269**, 20987–20994

[h] Nishihara, S., Narimatsu, H., Iwasaki, H., Yazawa, S., Akamatsu, S., Ando, T., Seno, T., and Ikuyo, N. (1994) *J. Biol. Chem.* **269**, 29271–29278

[i] Natsuka, S., Gersten, K. M., Zenita, K., Kannagi, R., and Lowe, J. B. (1994) *J. Biol. Chem.* **269**, 16789–16794

[j] Murray, B. W., Wittmann, V., Burkart, M. D., Hung, S.-C., and Wong, C.-H. (1997) *Biochemistry* **36**, 823–831

[k] Kelly, R. J., Rouquier, S., Giorgi, D., Lennon, G. G., and Lowe, J. B. (1995) *J. Biol. Chem.* **270**, 4640–4649

[l] Adamany, A., Blumenfeld, O., Sabo, B., and McCreary, J. (1983) *J. Biol. Chem.* **258**, 11537–11545

[m] Gillard, B., Blanchard, D., Bouhours, J.-F., Cartron, J.-P., van Kuik, J. A., Kamerling, J., Vliegenthart, J., and Marcus, D. (1988) *Biochemistry* **27**, 4601–4606

[n] Redman, C., Avellino, G., Pfeffer, S., Mukherjee, T., Nichols, M., Rubinstein, P., and Marsh, W. L. (1986) *J. Biol. Chem.* **261**, 9521–9525

[o] Bloy, C., Hermand, P., Blanchard, D., Cherif-Zahar, B., Goossens, D., and Cartron, J.-P. (1990) *J. Biol. Chem.* **265**, 21482–21487

N-linked glycan is pictured in Fig. 4-17. Like many others, it has a number of sialic acid residues at the nonreducing ends and also contains *N*-acetyllactosamine units. The major component of the cell walls of yeast (*S. cerevisiae*) is a mannoprotein that carries long *N*-linked oligosaccharides with highly branched outer chains of over 100 mannose residues[170] (see also Section C,3).

The special importance of sialic acids in glycoproteins may lie in the negative electrical charges and the resultant Ca^{2+}-binding properties that they impart to cell surfaces. Some glycoproteins carry sulfate or phosphate groups that have similar effects. The glycoproteins of the slime mold *Dictyostelium* contain both mannose-6-*P* and mannose-6-sulfate residues.[163] Oligosaccharides on the protein subunits of flagella of halobacteria contain sulfate esters of glucuronic acid.[172] Some *N*-linked oligosaccharides carry chains of keratan sulfate in place of the sialic acid in the Y_1 and Y_2 groups of the oligosaccharide.[173]

3. Glycoproteins in Biological Recognition

The clusters of sugar rings that form the oligosaccharides on glycoproteins play a vital role in many aspects of biological recognition.[139,174–176] A good example is provided by the human blood groups whose characteristics are determined largely by oligosaccharides (Box 4-C). The adhesion of viruses, bacteria, and eukaryotic parasites to cell surfaces and of one cell to another in multicellular organisms is also dependent on carbohydrates. Recently, it has become clear that the oligosaccharides of cell surfaces change during growth and development and provide an important mechanism by which cells can recognize each other and respond. Why are carbohydrates used for this purpose? It has been pointed out by Sharon and Lis[175] that four different nucleotides can make only 24 distinct tetranucleotides but that four different monosaccharides can make 35,560 unique tetrasaccharides.

Lectins and other carbohydrate-binding proteins. Much of biological carbohydrate-dependent recognition is a result of interaction of individual glycosyl groups or of oligosaccharides on a glycoprotein with a second protein. In some cases that protein is referred to as a **receptor**; in other cases the glycosyl groups may be called the receptor. Carbohydrate-binding proteins include antibodies, enzymes, and carriers that help sugars to cross cell membranes. In addition, there is a large group of carbohydrate binding proteins called **lectins** (from the latin *lectus*; to select).[24,174,177,178]

The first lectins discovered were proteins of plant seeds with specific sugar-binding properties and the ability to agglutinate erythrocytes. **Ricin**, a very toxic protein from castor beans, was isolated in 1888.[178,179] Perhaps the best known lectin is **concanavalin A**, a protein crystallized by Sumner in 1919.[180] Concanavalin A makes up 2–3% of the protein of the jack bean. It is one of a family of legume lectins that resemble favin, whose structure is shown in Fig. 2-15.[177,181,182] Many lectins, including ricin, have quite different three-dimensional structures but share the common characteristic of having a selective binding site for one or more glycosyl rings. Concanavalin A binds to α-D-mannopyranose or α-D-glucopyranose with unmodified hydroxyl groups at C-3, C-4, and C-6.[183,184] Tighter binding is observed if additional mannose residues are present in an oligosaccharide.[178,183] The protein also has specific binding sites for Ca^{2+} and for a transition metal ion such as Mn^{2+}. Soybean lectin binds D-*N*-acetylgalactosamine and D-galactose units, while wheat lectin is specific for D-*N*-acetylglucosamine.

Animal cells also produce lectins.[24,105,185–193] The amebas of the cellular slime mold *Dictyostelium* synthesize a classical lectin called **discoidin I** that binds GalNAc or Gal. It is absent from cells until they are ready to differentiate into an aggregating form (Chapter 1). Then it is produced in abundance.[178] Discoidin I has a second binding site specific for the peptide sequence Arg-Gly-Asp (RGD) which is known to be involved in cell adhesion and which binds to such surface proteins as fibronectin and laminin (Chapter 8). Many animal tissues contain soluble lactose-binding lectins known as **galectins** or S-Lac lectins. The best known member is a dimer of 14-kDa subunits;[187,188] many other related lectins have been found.[189,192] Another family are Ca^{2+}-dependent or C-type lectins which are specific for mannose, L-fucose, or other sugars.[105,193] **Lectin domains** are being discovered in many proteins.

Carbohydrate-binding sites. The structures of the sites that recognize and bind sugar rings in lectins, enzymes, transporter proteins, and other carbohydrate-binding proteins vary greatly, as does the tightness of binding. However, there are certain common features: Sugar rings are bound by hydrogen bonds, which are often numerous. An example is the galactose chemoreceptor protein from *E. coli*. It binds both α and β anomers of either D-glucose or D-galactose and is utilized by the bacteria in searching for food (see Chapter 19). The structure of D-glucose bound to this protein is shown in Fig. 4-18. Notice the many hydrogen bonds. Two of the –OH groups have the maximum of three hydrogen bonds apiece. There are three negatively charged aspartate side chains and one positively charged guanidinium group. These provide strong ion–dipole interactions which add strength to the bonds. The presence of ionized groups in varying numbers and constellations is another common feature of protein–carbohydrate interactions. A third common feature is the presence of aromatic rings, which often lie against one face of the sugar. The stereoscopic drawing of Fig. 4-18A shows an indole ring of a tryptophan residue in front and a phenylalanine side chain behind the sugar.[194] Sugars bind to lectins,[181,184,195] to enzymes (Chapter 12), and to antibodies[196] through similar interactions.

Binding of viruses and bacteria to cells. The cholera toxin and a related toxic protein from *E. coli* bind to Sia→Gal groups attached to glycolipids (gangliosides, Chapter 8) of erythrocytes and other cells.[197] The influenza virus gains access to our body cells by first binding through a viral surface **hemagglutinin**. This is a protein that binds specifically to NeuAcα2→6Gal or NeuAcα2→3Gal of cell surface oligosaccharides.[174,175,198] Removal of sialic acid from erythrocyte surfaces abolishes the ability of the influenza and some other viruses to bind. It seems somewhat surprising that a second surface protein on the influenza virus is a **neuraminidase** (or **sialidase**) which catalyzes the removal of sialic acid from cell surface proteins, destroying unoccupied virus receptors.[199,200] This may facilitate movement of a virus particle through the mucin layer surrounding a cell. Bacteria and other invading parasites also produce neuraminidases.[201] *Trypanosamas cruzi,* the causative agent of chagas disease, employs a **transsialidase** to transfer sialic acid from a Galβ on the host cell onto a protein on the parasite surface. This is essential for successful invasion of the host.[202]

The adherence of cells of *E. coli* to mannosyl units of cell surface proteins may initiate the infections that sometimes occur with this bacterium.[174,203] However, cells of *E. coli* from strains that cause urinary infections bind to Galα1→4Gal on glycolipids that carry the blood group P antigens (Box 4C).[174,204] Neuropathogenic strains of *E. coli* or of *Neisseria meningitidis*, which may cause neonatal meningitis, bind to α-2,8-polysialic acid chains on nerve cells.[135,205,206] *Helicobacter pylori*, the stomach ulcer bacterium, binds to the human Lewis^b blood group antigen (Box 4-C).[207] *Entamoeba histolytica*, which causes amebic dystentery, binds to Gal and GalNAc-containing oligosaccharides such as GalNAcα1→3Galα.[208]

Aggregation and adherence of cells. Differences among cell surfaces are fundamental to the formation of multicellular organisms and to many physiological processes. Proteins, carbohydrates, and lipids all contribute material to exposed cell surfaces. The adhesion of one cell to another is mediated by a group of **adhesion proteins** together with oligosaccharide groups and sometimes polysaccharides. Families of adhesion proteins include **integrins** and **cadhedrins** (discussed in Chapter 8), various members of the **immunoglobulin superfamily**, the **cell differentiation antigens**, (often designated CD44, etc.), the C-type lectins known as **selectins**, and proteoglycans.

Here are two of many known examples of specific cell–cell adhesion. The species-specific reaggregation of dissociated cells of marine sponges (Chapter 1) depends upon a 20-kDa proteoglycan of unique structure[209–211] together with a cell surface receptor protein and calcium ions. The recognition of egg cell surfaces by sperm[212–214] is species specific and depends upon interaction of sperm receptors with *O*-linked oligosaccharides of the extracellular coat of the ovum.

Growth and differentiation. The exact structures of the oligosaccharides and polysaccharides of cell surfaces vary not only with cell and tissue type but also with the position of a cell and with time. Actions of numerous glycosyl transferases alter these saccharide groups as an organism grows and develops. Other enzymes alter them by hydrolytic removal of sugars, by isomerizaion, oxidation, and addition of other components such as phospho, sulfo, and acetal groups. For example, the presence of the H-antigen determinant, whose structure is shown in Box 4-C, is strictly regulated, both temporally and spatially, during vertebrate development.[215,216] The relative amounts of the H determinant vs Le^x, sLe^x (Box 4-C)

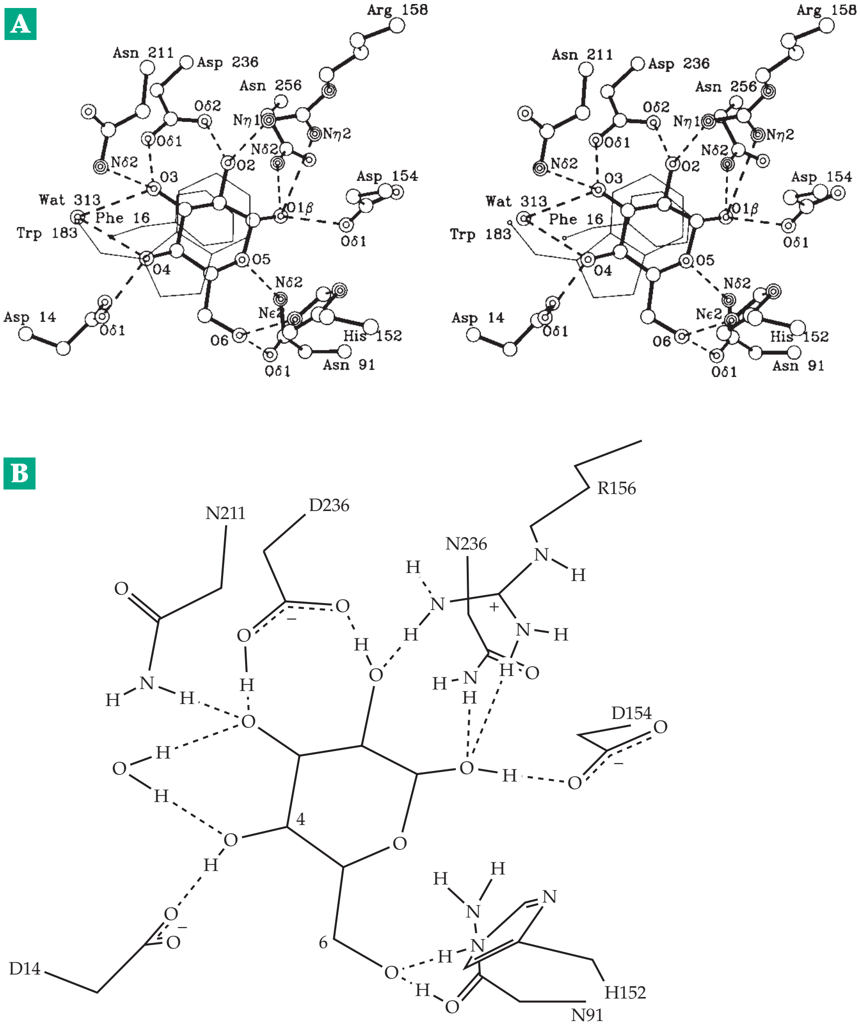


Figure 4-18 (A) Stereoscopic view of the interactions between the *E. coli* galactose chemoreceptor protein and a bound molecule of D-glucose. (B) Schematic drawing showing many of these interactions and the state of ionization deduced for the aspartate and arginine side chains. From Vyas *et al.*[194]

and other surface groupings is controlled by fucosyltransferases, sialyltransferases, etc. Human cancers often accumulate large amounts of fucosylated glycoproteins and glycolipids carrying Le^a, Le^x, and sialyl-Le^x antigens[175,216,217] and sialomucins.[218] Glycoproteins help to control many metabolic processes. For example, a protein **calnexin** (a chaperonin, Chapter 10) helps glycoproteins to fold correctly.[219] On the other hand, removal of terminal sialic acid residues from blood plasma proteins leads to rapid removal of the proteins from circulation and to catabolism by liver cells.[175] This process depends upon a receptor protein specific for oligosaccharides with terminal galactosyl residues (Chapter 20).

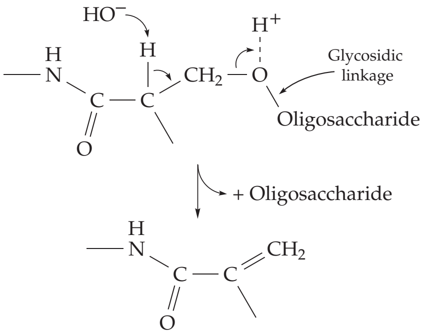
Recognition and adhesion by leukocytes. A group of three calcium-dependent lectins known as **selectins** bind the sialyl Lewis x (sLe^x) antigen and play important roles in adhesion to cells of the vascular endothelium and leukocytes[175,220–224] and also to platelets. Although all of the selectins bind the sLe^x antigen, the binding is weak and these multidomain proteins may simultaneously bind to other ligands such as heparan sulfate.[225] The leukocyte L-selectin (CD62L) binds very tightly to the 3'-sialyl-6'-sulfo-Le^x determinant[220,226] which occurs on mucin-like glycoproteins. The interaction with P selectin helps leukocytes to bind to surfaces on endothelial cells in lymph nodes and sites of chronic or acute inflammation. In a similar way E selectin is synthesized in vascular endothelial cells that have been transiently "activated" by cytokines in response to injury and other inflammatory stimuli. The E selectin binds the sulfated sLe^x antigen on surfaces of neutrophils, monocytes, eosinophils, and basophils.[220,222,227–230] P selectin is stored in secretory granules of platelets and endothelial cells and is released to the cell membrane upon activation by thrombin.[220]

E. Some Special Methods

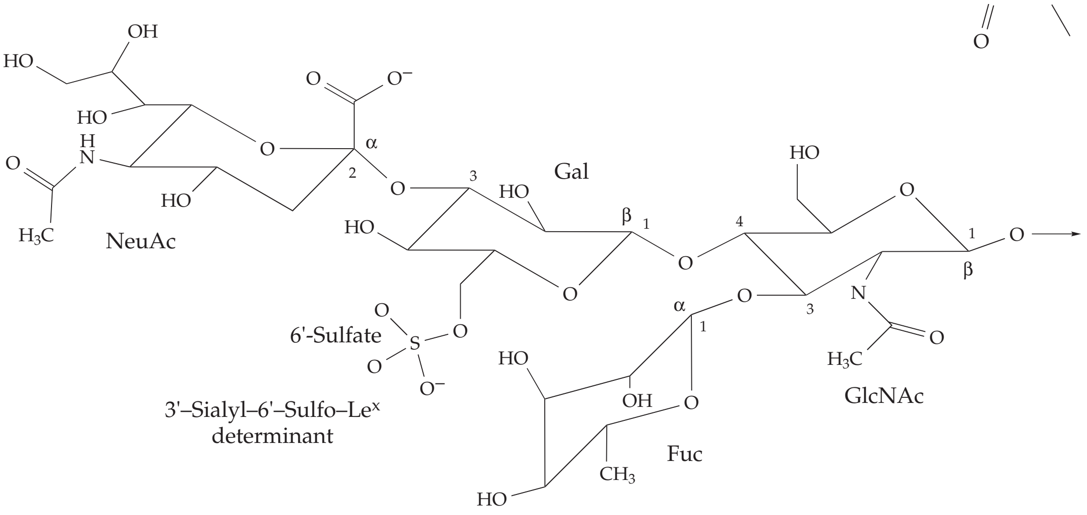
Small monosaccharides and oligosaccharides can be separated readily from polymeric constituents and can be purified further by chromatographic procedures including gel filtration as is illustrated in Fig. 3-3. However, polysaccharides and complex oligosaccharides are harder to purify. A few of these, such as cellulose and glycogen, are sufficiently stable that other materials can be dissolved away from them by prolonged boiling in strongly basic solutions. Complex carbohydrates are usually cut into smaller oligosaccharides or glycopeptides. These may be separated by HPLC, capillary electrophoresis, or thin-layer chromatography[134,231–235] (Fig. 4-19) or by chromatography on immobilized lectins.[234] Quantities of less than 25 picomoles can be separated by use of mass spectrometry with liquid chromatography.[236,237] High resolution Fourier transform mass spectroscopy is very useful in the study of posttranslational glycosylation of proteins.[237a,b]

1. Release of Oligosaccharides from Glycoproteins

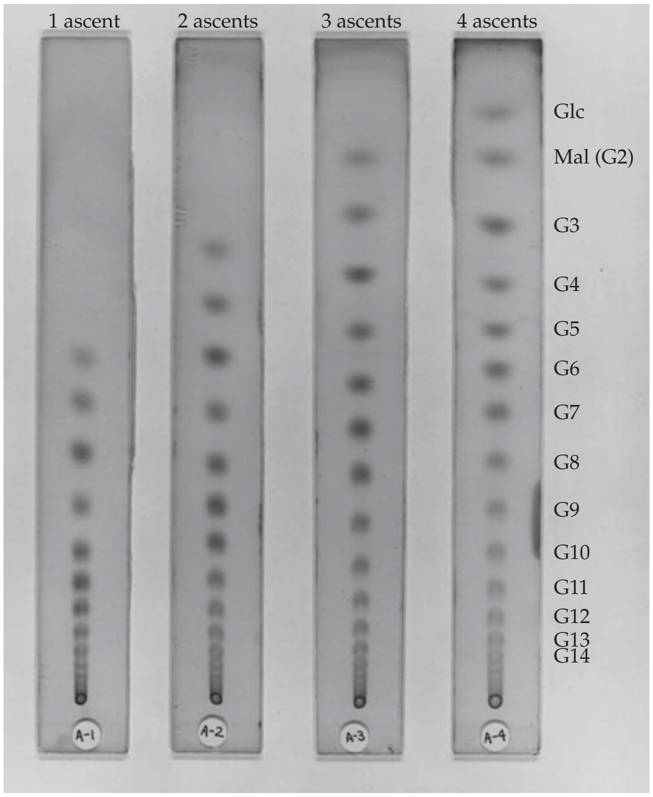
The *O*-linked oligosaccharides of glycoproteins or glycolipids can be split off from the proteins by β elimination (see also Chapter 13):



(4-10)



3'–Sialyl–6'–Sulfo–Le^x
determinant

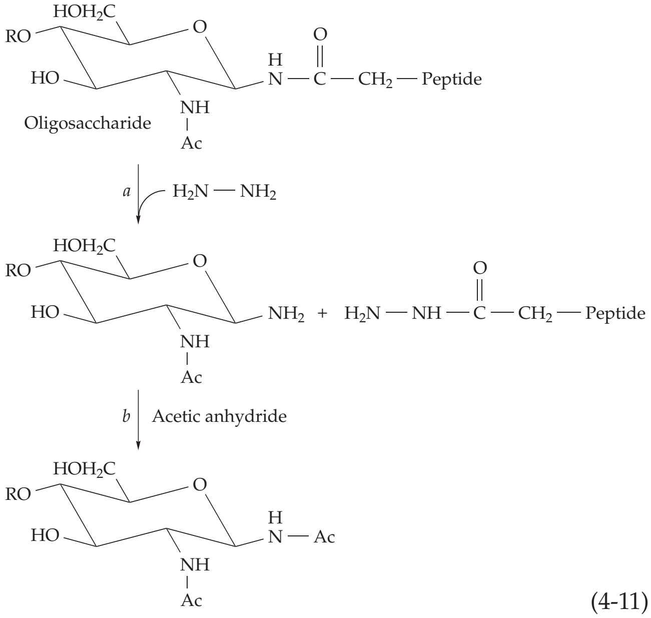


Solvent: MeCN/EtOAc/PrOH-1/H_2O
85 20 50 50

Figure 4-19 Ascending thin-layer chromatography of a mixture of maltooligosaccharides obtained by the hydrolysis of linear starch. G2 (maltose), G3 (maltotriose), G4 (maltotetraose), G5 (maltopentaose) – G14 represent oligosaccharides with the indicated number of glucose residues, all in α-1,4 linkage. In the multiple ascent technique the chromatographic solvent, whose composition (by volume) is indicated, is allowed to ascend the thin-layer plate repeatedly with the plate allowed to dry between ascents. The separation of the higher oligosaccharides is distinctly improved by a larger number of ascents. Photograph courtesy of John Robyt.

Treatment with 0.1–0.5 M NaOH for several hours will completely liberate the oligosaccharides, whose released carbonyl groups may then be reduced with $NaBH_4$, NaB^2H_4, or ^3H-labeled borohydride to form stable sugar alcohols (Eq. 4-2).[238] Asparagine-linked oligosaccharides are often recovered as glycopeptides prepared by complete proteolytic digestion of the denatured glycoprotein. These can be separated by high voltage electrophoresis in borate buffers.[239] The oligosaccharides can be released from the glycopeptides by enzymes such as endo-*N*-acetylglucosaminidase[240] or glycopeptidyl amylase of almonds.[241,242] These release the oligosaccharide as a 1-amino derivative.

When a dry glycoprotein is heated at 105° with anhydrous hydrazine for 8–12 h all of the asparagine-linked oligosaccharide chains are released (Eq. 4-11).[239,243–245] In addition to the glycosylamine product of step *a*, the corresponding 1–OH and –NH–NH_2 compounds are also formed. These are all converted to the stable acetyl derivative (step *b*), after which the oligosaccharides may be characterized by mass spectrometry and high-field NMR spectoscopy.[234,246–248] Crystallization of oligosaccharides in complex mixtures is difficult to impossible.



(4-11)

2. Hydrolysis

Most glycosidic linkages are hydrolyzed readily by heating with 1 N mineral acids. The mechanism of the hydrolytic reaction is similar to that employed by the enzyme lysozyme (Chapter 12). Some linkages are unusually sensitive to acid and a few are very resistant. Thus, a variety of conditions may be applied for partial acid hydrolysis as an aid to characterization.[249] Acetolysis, cleavage by acetic anhydride, is also of value.[250] A battery of hydrolytic enzymes specific for sugars that are joined in a given type of glycosidic linkage are available.[251] These are useful in determining sequences of oligosaccharides released from glycoproteins. Radioactive tracer techniques can also be applied.[252]

3. Methylation

An important general method in characterization of carbohydrates is the classical **exhaustive methylation** (permethylation). Repeated treatment with a methylating agent such as methyl iodide converts all free OH groups to OCH_3 groups. Then, complete acid hydrolysis, followed by separation of the methylated sugars and their quantitative determination, reveals the relative amounts of **end units** (containing four methoxyl groups), straight **chain units** (containing three methoxyl groups), and **branch points** (containing two methoxyl groups). In addition, the structure of the methylated derivatives provides information on the positions of the linkages in the sugar rings.

After a methylated polysaccharide is subjected to partial hydrolysis, the newly exposed hydroxyl groups (or those created by borohydride reduction of carbonyl groups) can be labeled by ethylation or propylation. One procedure for sequencing complex carbohydrates makes use of high-resolution reversed-phase liquid chromatography to separate the many alkylated oligosaccharides produced by methylation followed by partial acid hydrolysis, reduction, and ethylation.[233] Reductive cleavage of the glycosidic linkages in methylated polysaccharides allows uniquivocal determination of ring size.[253] Branch points may be located by methanolysis of the permethylated polysaccharide followed by conversion of free –OH groups to *p*-bromobenzoate esters. The latter are separated and the circular dichroism (Chapter 23) is measured. Mass spectrometry has also been applied successfully.[246] A simple procedure that can be conducted in any laboratory using thin-layer chromatography is illustrated in Fig. 4-20.[254]

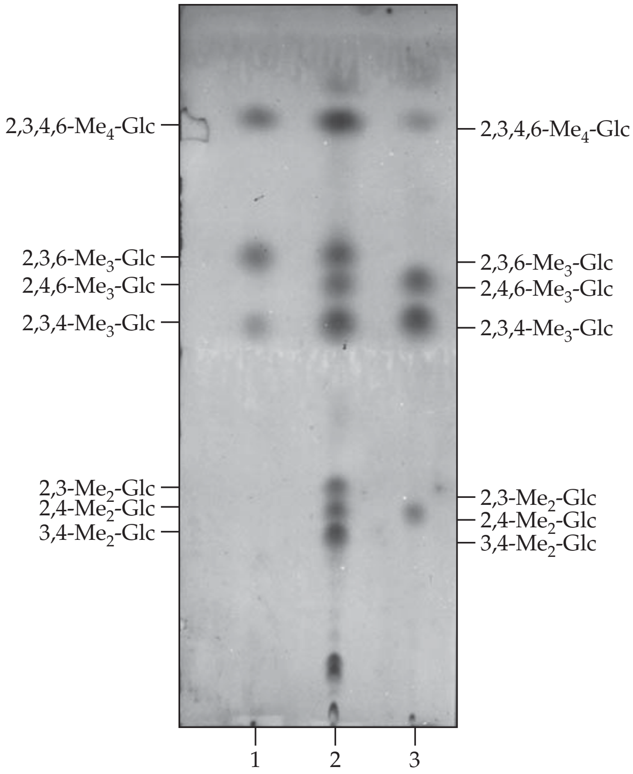
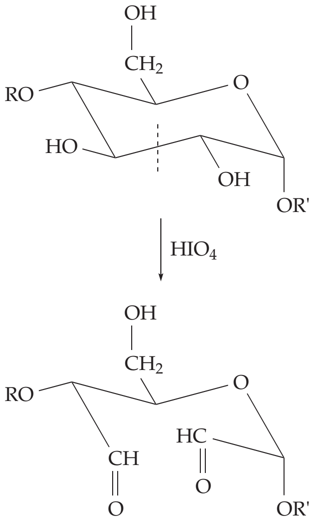


Figure 4-20 Separation of a mixture of *O*-methylated glucoses by ascending thin-layer chromatography. Whatman K6 TLC plates were used with two ascents of the solvent acetonitrile/chloroform/methanol in the ratio 3/9/2, V/V/V. Courtesy of John Robyt.

4. Periodate Oxidation (Smith Degradation)

One of the most important reagents in investigations of carbohydrate structure is periodic acid (or sodium periodate).[255] This reagent oxidatively cleaves C–C bonds bearing adjacent OH or NH_2 groups to form dialdehydes (Eq. 4-12). The method is quantitative. After some hours of reaction, excess periodate not consumed in the oxidation can be determined. If three consecutive carbon atoms bear hydroxyl or amino groups, formic acid is liberated from the central atom and can also be measured quantitatively. After destruction of excess periodate the dialdehyde can be reduced by addition of solid sodium borohydride to form stable CH_2OH groups. Following mild acid hydrolysis to split the acyclic acetal linkages, the fragments can be separated and identified. The sequence of reactions is known as the Smith degradation. If sodium borotritide (NaB^3H_4) is used for the reduction the fragments will be radioactive and can be located on chromatograms by fluorography. Periodate oxidation can also be used to alter all surface oligosaccharides.[99] Removal of *O*-linked oligosaccharides from glycoproteins can be accomplished by periodate oxidation and alkaline β elimination from the dialdehyde product of Eq. 4-12 if OR′ is part of a serine or other hydroxyamino acid side chain in a protein.[256]



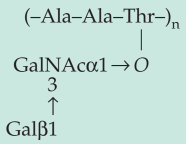
(4-12)

5. Nuclear Magnetic Resonance

As in other areas of biochemistry, NMR has become an extremely important tool in carbohydrate research. The mixtures of anomers of various ring

BOX 4-D ANTIFREEZE AND ICE-NUCLEATION PROTEINS

Fish living in Arctic and Antarctic waters may encounter temperatures as low as –1.9°C. The freezing point depression provided by dissolved salts and proteins in the blood is insufficient to protect the fish from freezing. As winter approaches, they synthesize and accumulate in their blood serum a series of eight or more special antifreeze proteins.[a–d] One type of antifreeze glycoprotein from winter flounder contains the following unit repeated 17–50 times.



$$\begin{array}{l} (\text{–Ala–Ala–Thr–})_n \\ \qquad\qquad\qquad\;\; | \\ \text{GalNAc}\alpha 1 \rightarrow O \\ \qquad 3 \\ \qquad \uparrow \\ \text{Gal}\beta 1 \end{array}$$

Destruction of the galactosyl residues by oxidation with periodate, acetylation of the free hydroxyl groups of the oligosaccharides, or their removal by β elimination all lead to loss of antifreeze activity.

The same fish contain a second series of alanine-rich antifreeze polypeptides that are *not* glycosylated but which exist as amphipathic helices. One of these (Type I) contains ~40 residues in a single helix.[e–i] A third family of antifreeze proteins (Type II), found in the sea raven are globular proteins, rich in cysteine and β structures. They are members of the lectin family.[d,j] A fourth type (Type III) found in the sea pout and some other fishes are 62- to 66-residue globular proteins containing an orthogonal β sandwich structure.[k,l] Messenger RNA molecules coding for the antifreeze proteins are found in the livers of flounder in the winter but are absent in the summer.[m]

Antifreeze proteins, that are 3 – 4 times as effective as those in fish, have been isolated from some insects and other arthropods.[m,n,o] They help beetle larvae to overwinter.[m] The insect proteins have a parallel β helix structure resembling that in Fig. 2-17 and stabilized by S — S bridges.[o,p] Some plants also synthesize antifreeze proteins.[n,q,r] One of these, isolated from carrots, is a member of the leucine-rich-repeat family.[q]

How do antifreeze proteins work? The major effect is to greatly slow the freezing rather than to decrease the freezing point. The proteins apparently accomplish this by binding to the surfaces of small ice crystals and preventing their growth.[d,f,h,i,k,l,s,t] This provides the fish with enough time for the blood to pass back into the liver, in which a high enough temperature is maintained to melt any microcrystals before the blood again circulates through the colder tissues. Some of the proteins have clusters of polar side chains that bind to specific faces of the ice crystals and inhibit growth.[s]

A few fishes tolerate a high internal osmotic pressure and accumulate glycerol in their blood up to a concentration of 0.4 M.[t] Insects may accumulate up to 3 M glycerol and some species utilize various other cryoprotectants, such as mannitol, sorbitol, erythritol, threitol, trehalose, glucose, fructose, proline, and alanine.[u] Some amphibians and reptiles can survive freezing and recover fully. For the most studied wood frog, rapid freezing is fatal, but slow freezing leaves the frog, whose heart ceases to function, with a 200-fold increased glucose concentration and a decreased water content in its organs. It resumes normal activities within 14–24 h of thawing.[t]

Having an effect opposite to that of the antifreeze proteins are surface proteins of some bacteria of the genera *Pseudomonas*, *Erwinia*, and *Xanthomonas*. These proteins provide nuclei for growth of ice crystals from supercooled water.[v,w]

---

[a] Feeney, R. E., Burcham, T. S., and Yeh, Y. (1986) *Ann. Rev. Biophys. Biophys. Chem.* **15**, 59–78

[b] Eastman, J. T., and DeVries, A. L. (1986) *Sci. Am.* **255** (Nov), 106–114

[c] Davies, P. L., and Hew, C. L. (1990) *FASEB J.* **4**, 2460–2468

[d] Gronwald, W., Loewen, M. C., Lix, B., Daugulis, A. J., Sönnichsen, F. D., Davies, P. L., and Sykes, B. D. (1998) *Biochemistry* **37**, 4712–4721

[e] Chakrabartty, A., Ananthanarayanan, V. S., and Hew, C. L. (1989) *J. Biol. Chem.* **264**, 11307–11312

[f] Yang, D. S. C., Sax, M., Chakrabartty, A., and Hew, C. L. (1988) *Nature (London)* **333**, 232–237

[g] Madura, J. D., Wierzbicki, A., Harrington, J. P., Maughon, R. H., Raymond, J. A., and Sikes, C. S. (1994) *J. Am. Chem. Soc.* **116**, 417–418

[h] Sicheri, F., and Yang, D. S. C. (1995) *Nature (London)* **375**, 427–431

[i] Wierzbicki, A., Taylor, M. S., Knight, C. A., Madura, J. D., Harrington, J. P., and Sikes, C. S. (1996) *Biophys. J.* **71**, 8–18

[j] Ng, N. F. L., and Hew, C. L. (1992) *J. Biol. Chem.* **267**, 16069–16075

[k] Chao, H., Sönnichsen, F. D., DeLuca, C. I., Sykes, B. D., and Davies, P. L. (1994) *Protein Sci.* **3**, 1760–1769

[l] Jia, Z., DeLuca, C. I., Chao, H., and Davies, P. L. (1996) *Nature (London)* **384**, 285–288

[m] Gourlie, B., Lin, Y., Price, J., DeVries, A. L., Powers, D., and Huang, R. C. C. (1984) *J. Biol. Chem.* **259**, 14960–14965

[n] Li, N., Chibber, B. A. K., Castellino, F. J., and Duman, J. G. (1998) *Biochemistry* **37**, 6343–6350

[o] Graether, S. P., Kuiper, M. J., Gagné, S. M., Walker, V. K., Jia, Z., Sykes, B. D., and Davies, P. L. (2000) *Nature (London)* **406**, 325 – 328

[p] Liou, Y.-C., Tocilj, A., Davies, P. L., and Jia, Z. (2000) *Nature (London)* **406**, 322–324

[q] Worrall, D., Elias, L., Ashford, D., Smallwood, M., Sidebottom, C., Lillford, P., Telford, J., Holt, C., and Bowles, D. (1998) *Science* **282**, 115–117

[r] Sidebottom, C., Buckley, S., Pudney, P., Twigg, S., Jarman, C., Holt, C., Telford, J., McArthur, A., Worrall, D., Hubbard, R., and Lillford, P. (2000) *Nature (London)* **406**, 256

[s] Knight, C. A., (2000) *Nature (London)* **406**, 249–251

[t] Costanzo, J. P., Lee, R. E., Jr., DeVries, A. L., Wang, T., and Layne, J. R., Jr. (1995) *FASEB J.* **9**, 351–352

[u] Storey, K. B., and Storey, J. M. (1990) *Sci. Am.* **263** (Dec), 92–97

[v] Wolber, P., and Warren, G. (1989) *Trends Biochem. Sci.* **14**, 179–182

[w] Gurian-Sherman, D., and Lindow, S. E. (1993) *FASEB J.* **7**, 1338 – 1343

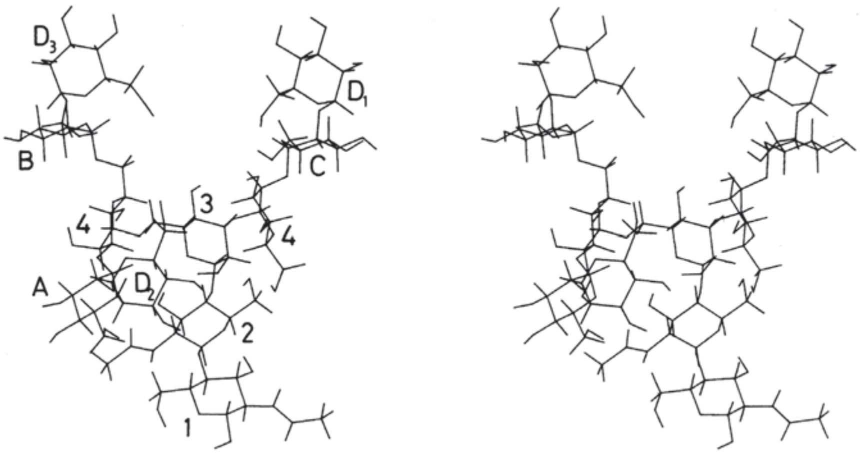


Figure 4-21 Stereoscopic view of an N-linked oligosaccharide whose structure has been deduced by two-dimensional NMR spectroscopy and energy calculations. This is one of a range of allowed conformations. From Homans *et al.*[257]

sizes (Eq. 4-1) can be analyzed with proton or ^{13}C NMR.[10,257] A variety of newer NMR techniques, some of which have been described in Chapter 3, have been applied.[28,258–267] A problem with NMR spectroscopy of these compounds has been the lack of the large number of nuclear Overhauser enhancements (NOEs) that can be observed.[268,269] Nevertheless, when combined with energy calculations it is possible to use NMR measurements to deduce conformations of sugar rings, three-dimensional structures, and the degree of conformational flexibility in various parts of N-linked oligosaccharides (Fig. 4-21). Measurement of C–O–C–C spin-coulping constants is also of value.[270] Use of multidimensional NMR has permitted analysis of mixtures of cellulose oligosaccharides.[269]

References

1. Pigman, W. W., and Horton, D., eds. (1972) *The Chemistry of Carbohydrates*, 2nd ed., Academic Press, New York
2. Guthrie, R. D. (1974) *Guthrie and Honeyman's Introduction to Carbohydrate Chemistry*, 4th ed., Oxford Univ. Press (Clarendon), London and New York (p. 17)
3. Binkley, R. W. (1988) *Modern Carbohydrate Chemistry*, Dekker, New York
4. Shalenberger, R. S. (1982) *Advanced Sugar Chemistry: Principles of Sugar Stereochemistry*, AVI Publ. Co., Westport, Connecticut
5. El-Khadem, H. S. (1988) *Carbohydrate Chemistry*, Academic Press, San Diego, California
6. Prince, R. C., Gunson, D. E., Leigh, J. S., and McDonald, G. G. (1982) *Trends Biochem. Sci.* **7**, 239–240
7. Wertz, P. W., Garver, J. C., and Anderson, L. (1981) *J. Am. Chem. Soc.* **103**, 3916–3922
8. Ha, S., Gao, J., Tidor, B., Brady, J. W., and Karplus, M. (1991) *J. Am. Chem. Soc.* **113**, 1553–1557
9. Angyal, S. J. (1979) in *Asymmetry of Carbohydrates* (Harmon, R. E., ed), p. 15, Decker, New York
10. Gray, G. R. (1976) *Acc. Chem. Res.* **9**, 418
11. Pierce, J., Serianni, A. S., and Barker, R. (1985) *J. Am. Chem. Soc.* **107**, 2448–2456
12. Maple, S. R., and Allerhand, A. (1987) *J. Am. Chem. Soc.* **109**, 3168–3169
13. Goux, W. J. (1985) *J. Am. Chem. Soc.* **107**, 4320–4327
14. Barker, S. A., and Baggett, N. (1978) *Trends Biochem. Sci.* **3**, 140–141
15. Benson, A. A. (1963) *Adv. Lipid Res.* **1**, 387–394
16. Benning, C., Beatty, J. T., Prince, R. C., and Somerville, C. R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1561–1565
17. Dijkema, C., Kester, H. C. M., and Visser, J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 14–18
18. Tarczynski, M. C., Jensen, R. G., and Bohnert, H. J. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2600–2604
19. Yamanouchi, T., Tachibana, Y., Sekino, N., Akanuma, H., Akaoka, I., and Miyashita, H. (1994) *J. Biol. Chem.* **269**, 9664–9668
20. Nadano, D., Iwasaki, M., Endo, S., Kitajima, K., Inoue, S., and Inoue, Y. (1986) *J. Biol. Chem.* **261**, 11550–11557
21. Schauer, R. (1985) *Trends Biochem. Sci.* **10**, 357–360
22. Faillard, H. (1989) *Trends Biochem. Sci.* **14**, 237–241
23. Roth, J., Kempf, A., Reuter, G., Schauer, R., and Gehring, W. J. (1992) *Science* **256**, 673–675
24. Powell, L. D., and Varki, A. (1995) *J. Biol. Chem.* **270**, 14243–14246
25. Stoddart, J. F. (1971) *Stereochemistry of Carbohydrates*, Wiley-Interscience, New York
26. Brady, J. W. (1986) *J. Am. Chem. Soc.* **108**, 8152–8160

26a. Marszalek, P. E., Pang, Y.-P., Li, H., Yazal, J. E., Oberhauser, A. F., and Fernandez, J. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7894–7898

27. Venkataraman, G., Sasisekharan, V., Cooney, C. L., Langer, R., and Sasisekharan, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6171–6175
28. Hajduk, P. J., Horita, D. A., and Lerner, L. E. (1993) *J. Am. Chem. Soc.* **115**, 9196–9201
29. Wolfe, S., Whangbo, M.-H., and Mitchell, D. J. (1979) *Carbohydrate Res.* **69**, 1–26
30. Kirby, A. J. (1983) *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, Springer, Berlin
31. Perrin, C. L., Armstrong, K. B., and Fabian, M. A. (1994) *J. Am. Chem. Soc.* **116**, 715–722
32. Cui, W., Li, F., and Allinger, N. L. (1993) *J. Am. Chem. Soc.* **115**, 2943–2951
33. Westhof, E., and Sundaralingam, M. (1980) *J. Am. Chem. Soc.* **102**, 1493–1500

33a. French, A. D., Dowd, M. K., and Reilly, P. J. (1997) *J. Mol. Struct. (Theochem)*, 395–396

34. Gabb, H. A., and Harvey, S. C. (1993) *J. Am. Chem. Soc.* **115**, 4218–4227
35. Plavec, J., Tong, W., and Chattopadhyaya, J. (1993) *J. Am. Chem. Soc.* **115**, 9734–9746
36. Ellervik, U., and Magnusson, G. (1994) *J. Am. Chem. Soc.* **116**, 2340–2347

36a. French, A. D., and Murphy, V. G. (1973) *Carbohydr. Res.* **27**, 391–406

37. Robyt, J. F., and White, B. J. (1996) *Biochemical Techniques, Theory and Practice*, 2nd ed., Waveland Press, Prospect Heights, Illinois
38. Lipták, A., Fügedi, P., Szurmai, Z., and Harangi, J. (1990) *CRC Handbook of Oligosaccharides. Vol. I: Disaccharides*, CRC Press, Boca Raton, Florida
39. Cerami, A. (1986) *Trends Biochem. Sci.* **11**, 311–314
40. Watkins, N. G., Neglia-Fisher, C. I., Dyer, D. G., Thorpe, S. R., and Baynes, J. W. (1987) *J. Biol. Chem.* **262**, 7207–7212
41. Súarez, G., Rajaram, R., Oronsky, A. L., and Gawinowicz, M. A. (1989) *J. Biol. Chem.* **264**, 3674–3679

References

42. Takata, K., Horiuchi, S., Araki, N., Shiga, M., Saitoh, M., and Morino, Y. (1988) *J. Biol. Chem.* **263**, 14819 – 14825
43. Hayase, F., Nagaraj, R. H., Miyata, S., Njoroge, F. G., and Monnier, V. M. (1989) *J. Biol. Chem.* **264**, 3758 – 3764
44. Goldberg, R. N., and Tewari, Y. B. (1989) *J. Biol. Chem.* **264**, 9897 – 9900
45. Tewari, Y. B., and Goldberg, R. N. (1989) *J. Biol. Chem.* **264**, 3966 – 3971
46. Goldberg, R. N., Tewari, Y. B., and Ahluwalia, J. C. (1989) *J. Biol. Chem.* **264**, 9901 – 9904
47. Weissborn, A. C., Rumley, M. K., and Kennedy, E. P. (1991) *J. Biol. Chem.* **266**, 8062 – 8067
48. Reymond, P., Grünberger, S., Paul, K., Müller, M., and Farmer, E. E. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4145 – 4149
49. Nacharaju, P., and Acharya, A. S. (1992) *Biochemistry* **31**, 12673 – 12679
50. Baynes, J. W., and Monnier, V. M., eds. (1989) *the Maillard Reaction in Aging, Diabetes and Nutrition*, Liss, New York
51. Larsen, M. L., Hørder, M., and Mogensen, E. F. (1990) *N. Engl. J. Med.* **323**, 1021 – 1025
52. Luthra, M., and Balasubramanian, D. (1993) *J. Biol. Chem.* **268**, 18119 – 18127
53. Urbanowski, J. C., Cohenford, M. A., and Dain, J. A. (1982) *J. Biol. Chem.* **257**, 111 – 115
54. Grandhee, S. K., and Monnier, V. M. (1991) *J. Biol. Chem.* **266**, 11649 – 11653
55. Brownlee, M., Cerami, A., and Vlassara, H. (1988) *N. Engl. J. Med.* **318**, 1315 – 1320
56. Papoulis, A., Al-Abed, Y., and Bucala, R. (1995) *Biochemistry* **34**, 648 – 655
57. Wells-Knecht, K. J., Zyzak, D. V., Litchfield, J. E., Thorpe, S. R., and Baynes, J. W. (1995) *Biochemistry* **34**, 3702 – 3709
58. French, D. (1975) in *International Review of Biochemistry*, Vol. 5 (Whelan, W. J., ed), p. 269, Univ. Park Press, Baltimore, Maryland (Series One)
59. Atkins, E. D. T., ed. (1986) *Polysaccharides, Topics in Structure and Morphology*, VCH Publisher, New York
60. Whistler, R. L., and BeMiller, J. N., eds. (1993) *Industrial Gums*, 3rd ed., Academic Press, San Diego, California
61. Rao, V. S. R., and Sathyanarayana, B. K. (1972) *Biopolymers* **11**, 1379 – 1394
62. French, A. D., and Dowd, M. K. (1993) in *Cellulostics: Chemical,Biochemical and Materials Aspects* (Kennedy, J. F., Phillips, G. O., and Williams, P. A., eds), pp. 51 – 56, Ellis Harwood, New York
63. Duda, C. A., and Stevens, E. S. (1993) *J. Am. Chem. Soc.* **115**, 8487 – 8488
64. Liébecq, C., ed. (1992) *Biochemical Nomenclature*, Portland Press, London and Chapel Hill, North Carolina (for the International Union of Biochemistry and Molecular Biology)
65. Preston, R. D. (1986) *Trends Biochem. Sci.* **11**, 377 – 380
66. Ross, P., Mayer, R., and Benziman, M. (1991) *Microbiol. Rev.* **55**, 35 – 58
67. Sugiyama, J., Vuong, R., and Chanzy, H. (1991) *Macromolecules* **24**, 4168 – 4175
68. Dudley, R. L., Fyfe, C. A., Stephenson, P. J., Deslandes, Y., Hamer, G. K., and Marchessault, R. H. (1983) *J. Am. Chem. Soc.* **105**, 2469 – 2472
69. Gebler, K., Kraub, N., Steiner, T., Betzel, C., Sandmann, C., and Saenger, W. (1994) *Science* **266**, 1027 – 1029

69a. Langan, P., Nishiyama, Y., and Chanzy, H. (1999) *J. Am. Chem. Soc.* **121**, 9940 – 9946

70. Grover, J. A. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 475 – 504, Academic Press, San Diego, California
71. French, D. (1973) *J. Animal Sci.* **37**, 1048 – 1061
72. Maddelein, M. L., Libessart, N., Bellanger, F., Delrue, B., D'Hulst, C., Van den Koornhuyse, N., Fontaine, T., Wieruszeski, J. M., Decq, A., and Ball, S. (1994) *J. Biol. Chem.* **269**, 25150 – 25157
73. Jane, J., Kasemsuwan, T., Leas, S., Zobel, H., and Robyt, J. F. (1994) *Starch/Stärke* **46**, 121 – 129
74. Meyer, K. H. (1943) *Adv. Enzymol.* **3**, 109 – 136
75. French, D. (1969) in *Symposium on Foods: Carbohydrates and their Roles* (Schultz, H. W., ed), pp. 26 – 54, AVI Publ. Co., Westport, Connecticut
76. Rundle, R. E., and French, D. (1943) *J. Am. Chem. Soc.* **65**, 1707 – 1710
77. Hybl, A., Rundle, R. E., and Williams, D. E. (1965) *J. Am. Chem. Soc.* **87**, 2779 – 2788
78. Teitelbaum, R. C., Ruby, S. L., and Tobin, J. M. (1980) *J. Am. Chem. Soc.* **102**, 3322 – 3328
79. Winter, W. T., and Sarko, A. (1974) *Biopolymers* **13**, 1461 – 1482
80. Kainuma, K., and French, D. (1972) *Biopolymers* **11**, 2241 – 2250
81. Wu, H.-C., H, and Sarko, A. (1978) *Carbohydr. Res.* **61**, 7 – 25
82. Gidley, M. J., and Bociek, S. M. (1988) *J. Am. Chem. Soc.* **110**, 3820 – 3829
83. Lomako, J., Lomako, W. M., Whelan, W. J., Dombro, R. S., Neary, J. T., and Norenberg, M. D. (1993) *FASEB J.* **7**, 1386 – 1393
84. Chen, W., Avison, M. J., Zhu, X. H., and Shulman, R. G. (1993) *Biochemistry* **32**, 11483 – 11487
85. Ohana, P., Delmer, D. P., Steffens, J. C., Matthews, D. E., Mayer, R., and Benziman, M. (1991) *J. Biol. Chem.* **266**, 13742 – 13745
86. Mol, P. C., Park, H., Mullins, J. T., and Cabib, E. (1994) *J. Biol. Chem.* **269**, 31267 – 31274
87. Yamamoto, R., and Nevins, D. J. (1978) *Carbohydrate Res.* **67**, 275 – 280
88. Marchessault, R. H., and Deslandes, Y. (1979) *Carbohydrate Res.* **75**, 231 – 242
89. Kang, M. S., and Cabib, E. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5808 – 5812
90. Harada, T., Terasaki, M., and Harada, A. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 427 – 445, Academic Press, San Diego, California
91. Walker, G. J. (1978) in *International Review of Biochemistry*, Vol. 16 (Manners, D. J., ed), p. 75, Univ. Park Press, Baltimore, Maryland
92. de Belder, A. N. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 399 – 425, Academic Press, San Diego, California
93. Yang, X., Miller, M. A., Yang, R., Evans, D. F., and Edstrom, R. D. (1990) *FASEB J.* **4**, 3140 – 3143
94. Tsujisaka, Y., and Mitsuhashi, M. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 447 – 460, Academic Press, San Diego, California
95. Arnott, S., Fulmer, A., Scott, W. E., Dea, I. C. M., Moorehouse, R., and Rees, D. A. (1974) *J. Mol. Biol.* **90**, 269 – 284
96. Shibata, N., Ikuta, K., Imai, T., Satoh, Y., Satoh, R., Suzuki, A., Kojima, C., Kobayashi, H., Hisamichi, K., and Suzuki, S. (1995) *J. Biol. Chem.* **270**, 1113 – 1122
97. Kollár, R., Petráková, E., Ashwell, G., Robbins, P. W., and Cabib, E. (1995) *J. Biol. Chem.* **270**, 1170 – 1178
98. Preston, R. D. (1968) *Sci. Am.* **218**(Jun), 102 – 108
99. Turvey, J. R. (1978) in *International Review of Biochemistry*, Vol. 16 (Manners, D. J., ed), Univ. Park Press, Baltimore, Maryland
100. Kirkwood, S. (1974) *Ann. Rev. Biochem.* **43**, 401 – 417
101. Minke, R., and Blackwell, J. (1978) *J. Mol. Biol.* **120**, 167 – 181
102. Roberts, G. A. F. (1992) *Chitin Chemistry*, MacMillan, London
103. Blackwell, J., and Weih, M. A. (1980) *J. Mol. Biol.* **137**, 49 – 60
104. Rees, D. A., and Welsh, E. J. (1977) *Angew. Chem. Int. Ed. Engl.* **16**, 214 – 224
105. Lee, Y. C. (1992) *FASEB J.* **6**, 3193 – 3200
106. Laurent, T. C., and Fraser, J. R. E. (1992) *FASEB J.* **6**, 2397 – 2404
107. Hardingham, T. E., and Fosang, A. J. (1992) *FASEB J.* **6**, 861 – 870
108. Fransson, L. (1987) *Trends Biochem. Sci.* **12**, 406 – 411
109. Morris, E. R., Rees, D. A., and Welsh, E. J. (1980) *J. Mol. Biol.* **138**, 383 – 400
110. David, G. (1993) *FASEB J.* **7**, 1023 – 1030
111. Yamada, S., Murakami, T., Tsuda, H., Yoshida, K., and Sugahara, K. (1995) *J. Biol. Chem.* **270**, 8696 – 8705
112. Lane, D. A., and Björk, I., eds. (1992) *Heparin And Related Polysaccharides*, Plenum, New York
113. Edens, R. E., Fromm, J. R., Linhardt, R. J., and Weiler, J. M. (1995) *Biochemistry* **34**, 2400 – 2407
114. Linhardt, R. J., Ampofo, S. A., Fareed, J., Hoppensteadt, D., Mulliken, J. B., and Folkman, J. (1992) *Biochemistry* **31**, 12441 – 12445
115. Casu, B., Petitou, M., Provasoli, M., and Sinaÿ, P. (1988) *Trends Biochem. Sci.* **13**, 221 – 225
116. Oates, J. A., and Wood, A. J. J. (1991) *N. Engl. J. Med.* **324**, 1565 – 1574
117. Rosenberg, R. D., and Lam, L. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1218 – 1222
118. Lindahl, U., Kusche-Gullberg, M., and Kjelln, L. (1998) *J. Biol. Chem.* **273**, 24979 – 24982
119. Mulloy, B. A., Ribeiro, A.-C., Alves, A.-P., Vieira, R. P., and Mourao, P. A. S. (1994) *J. Biol. Chem.* **269**, 22113 – 22123
120. Gibeaut, D., and Carpita, N. C. (1994) *FASEB J.* **8**, 904 – 915
121. Hayashi, T., and Matsuda, K. (1981) *J. Biol. Chem.* **256**, 11117 – 11122
122. Millane, R. P. (1992) in *Frontiers in Carbohydrate Research-2* (Chandrasekaran, R., ed), pp. 168 – 190, Elsevier, London
123. Rolin, C. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 257 – 293, Academic Press, San Diego, California
124. Arnott, S., Scott, W. E., Rees, D. A., and McNab, C. G. A. (1974) *J. Mol. Biol.* **90**, 253 – 267
125. Selby, H. H., and Whistler, R. L. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 87 – 103, Academic Press, San Diego, California
126. Albersheim, P., Bauer, W. D., Keestra, K., Talmadge, K. W. (1973) *Biogenesis of Plant Cell Wall Polysaccharides*, (Loewus, F., ed.), Academic Press, New York
127. Rees, D. A. (1972) *Biochem. J.* **126**, 257 – 273
128. Therkelsen, G. H. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 145 – 180, Academic Press, San Diego, California
129. Clare, K. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 105 – 143, Academic Press, San Diego, California
130. Marsh, M. E., Chang, D. C., and King, G. C. (1992) *J. Biol. Chem.* **267**, 20507 – 20512

130a. Vinogradov, E., and Bock, K. (1999) *Angew. Chem. Ed. Engl.* **38**, 671 – 674

131. Sutherland, I. W. (1979) *Trends Biochem. Sci.* **4**, 55 – 59

References

132. Kang, K. S., and Pettitt, D. J. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 341–397, Academic Press, San Diego, California
133. Wolucka, B. A., McNeil, M. R., de Hoffmann, E., Chonjnacki, T., and Brennan, P. J. (1994) *J. Biol. Chem.* **269**, 23328–23335
134. Kobata, A. (1984) in *Biology of Carbohydrates*, Vol. 2 (Ginsburg, V., and Robbins, P. W., eds), pp. 87–161, Wiley, New York
135. Lennarz, W. J., ed. (1980) *The Biochemistry of Glycoproteins and Proteoglycans*, Plenum, New York
136. Montreul, J. (1982) in *Comprehensive Biochemistry*, 1st ed., Vol. 19B, Part II (Florkin, M., and Stotz, E. H., eds), pp. 1–188, Elsevier, Amsterdam
137. Rademacher, T. W., Parekh, R. B., and Dwek, R. A. (1988) *Ann. Rev. Biochem.* **57**, 785–838
138. Roberts, D. D., and Mecham, R. P., eds. (1993) *Cell Surface and Extracellular Glycoconjugates: Structure and Function*, Academic Press, San Diego, California
138a. Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., and Marth, J., eds. (1999) *Essentials of Glycobiology*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
139. Fukuda, M., and Hindsgaul, O., eds. (1994) *Molecular Glycobiology*, Oxford Publishing, Oxford
140. Williams, R. L., Greene, S. M., and McPherson, A. (1987) *J. Biol. Chem.* **262**, 16020–16031
141. Hofsteenge, J., Müller, D. R., de Beer, T., Löffler, A., Richter, W. J., and Vliegenthart, J. F. G. (1994) *Biochemistry* **33**, 13524–13530
141a. Iozzo, R. V. (1998) *Ann. Rev. Biochem.* **67**, 609–652
141b. Bernfield, M., Götte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincecum, J., and Zako, M. (1999) *Ann. Rev. Biochem.* **68**, 729–777
142. Sugahara, K., Ohkita, Y., Shibata, Y., Yoshida, K., and Ikegami, A. (1995) *J. Biol. Chem.* **270**, 7204–7212
143. Funderburgh, J. L., Funderburgh, M. L., Brown, S., Vergnes, J.-P., Hassell, J. R., Mann, M. M., and Conrad, G. W. (1993) *J. Biol. Chem.* **268**, 11874–11880
144. Neame, P. J., Christner, J. E., and Baker, J. R. (1987) *J. Biol. Chem.* **262**, 17768–17778
145. Scott, J. E. (1987) *Trends Biochem. Sci.* **12**, 318–321
146. Rosenberg, L., Tang, L.-H., Pal, S., Johnson, T. L., and Choi, H. U. (1988) *J. Biol. Chem.* **263**, 18071–18077
147. Choi, H. U., Johnson, T. L., Pal, S., Tang, L.-H., Rosenberg, L., and Neame, P. J. (1989) *J. Biol. Chem.* **264**, 2876–2884
148. Kokenyesi, R., and Bernfield, M. (1994) *J. Biol. Chem.* **269**, 12304–12309
149. Rosenberg, L. (1975) in *Dynamics of Connective Tissue Macromolecules* (Burleigh, P. M. C., and Poole, A. R., eds), p. 107, North-Holland Publ., Amsterdam
150. Chen, J., and Varner, J. E. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4399–4403
151. Kieliszewski, M., deZacks, R., Leykam, J. F., and Lamport, D. T. A. (1992) *Plant Physiol.* **98**, 919–926
152. Kieliszewski, M. J., O'Neill, M., Leykam, J., and Orlando, R. (1995) *J. Biol. Chem.* **270**, 2541–2549
153. Eckhardt, A. E., Timpte, C. S., Abernethy, J. L., Toumadje, A., Johnson, W. C., and Hill, R. L. (1987) *J. Biol. Chem.* **262**, 11339–11344
154. Timpte, C. S., Eckhardt, A. E., Abernethy, J. L., and Hill, R. L. (1988) *J. Biol. Chem.* **263**, 1081–1088
155. Gupta, R., and Jentoft, N. (1989) *Biochemistry* **28**, 6114–6121
156. Carter, S. R., Slomiany, A., Gwozdzinski, K., Liau, Y. H., and Slomiany, B. L. (1988) *J. Biol. Chem.* **263**, 11977–11984
157. Jentoft, N. (1990) *Trends Biochem. Sci.* **15**, 291–294
158. Butenhof, K. J., and Gerken, T. A. (1993) *Biochemistry* **32**, 2650–2663
159. Toribara, N. W., Roberton, A. M., Ho, S. B., Kuo, W.-L., Gum, E., Hicks, J. W., Gum, J. R., Byrd, J. C., Siddiki, B., and Kim, Y. S. (1993) *J. Biol. Chem.* **268**, 5879–5885
160. Sadler, J. E. (1984) in *"Biology of Carbohydrates"*, Vol. 2 (Ginsburg, V., and Robbins, P. W., eds), pp. 199–288, Wiley, New York
160a. Perez-Vilar, J., and Hill, R. L. (1999) *J. Biol. Chem.* **274**, 31751–31754
161. Fukuda, M., Carlsson, S. R., Klock, J. C., and Dell, A. (1986) *J. Biol. Chem.* **261**, 12796–12806
162. Gerken, T. A., and Jentoft, N. (1987) *Biochemistry* **26**, 4689–4699
163. Freeze, H. H., and Wolgast, D. (1986) *J. Biol. Chem.* **261**, 127–134
164. Carver, J. P., and Brisson, J.-R. (1984) in *Biology of Carbohydrates*, Vol. 2 (Ginsburg, V., and Robbins, P. W., eds), pp. 289–329, Wiley, New York
165. Bush, C. A., Duben, A., and Ralapati, S. (1980) *Biochemistry* **19**, 501–504
166. Finne, J., and Krusius, T. (1979) *Eur. J. Biochem.* **102**, 583–588
167. Järnefelt, J., Finne, J., Krusius, T., and Rauvala, H. (1978) *Trends Biochem. Sci.* **3**, 110–114
168. Kornfeld, R., and Kornfeld, S. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W. J., ed), Plenum, New York
169. Fukuda, M., Dell, A., and Fukuda, M. N. (1984) *J. Biol. Chem.* **259**, 4782–4791
170. Hernández, L. M., Olivero, I., Alvarado, E., and Larriba, G. (1992) *Biochemistry* **31**, 9823–9831
171. Taguchi, T., Seko, A., Kitajima, K., Muto, Y., Inoue, S., Khoo, K.-H., Morris, H. R., Dell, A., and Inoue, Y. (1994) *J. Biol. Chem.* **269**, 8762–8771
172. Wieland, F., Paul, G., and Sumper, M. (1985) *J. Biol. Chem.* **260**, 15180–15185
173. Funderburgh, J. L., Funderburgh, M. L., Mann, M. M., and Conrad, G. W. (1991) *J. Biol. Chem.* **266**, 14226–14231
174. Sharon, N., and Lis, H. (1989) *Science* **246**, 227–229
175. Sharon, N., and Lis, H. (1993) *Sci. Am.* **268**(Jan), 82–89
176. Opdenakker, G., Rudd, P. M., Ponting, C. P., and Dwek, R. A. (1993) *FASEB J.* **7**, 1330–1337
177. Sharon, N. (1993) *Trends Biochem. Sci.* **18**, 221–226
178. Barondes, S. H. (1988) *Trends Biochem. Sci.* **13**, 480–482
179. Sphyris, N., Lord, J. M., Wales, R., and Roberts, L. M. (1995) *J. Biol. Chem.* **270**, 20292–20297
180. Sumner, J. B., and Howell, S. F. (1936) *J. Bacteriol.* **32**, 227–237
181. Bourne, Y., Rougés, P., and Cambillau, C. (1990) *J. Biol. Chem.* **265**, 18161–18165
182. Sharon, N., and Lis, H. (1990) *FASEB J.* **4**, 3198–3208
183. Mandal, D. K., and Brewer, C. F. (1993) *Biochemistry* **32**, 5116–5120
184. Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb, A. J., Dodson, E. J., Papiz, M. Z., Wan, T., and Campbell, J. (1989) *EMBO J.* **8**, 2189–2193
185. Drickamer, K. (1988) *J. Biol. Chem.* **263**, 9557–9560
186. Krusius, T., Gehlsen, K. R., and Ruoslahti, E. (1987) *J. Biol. Chem.* **262**, 13120–13125
187. Gitt, M. A., and Barondes, S. H. (1991) *Biochemistry* **30**, 82–89
188. Barondes, S. H., Cooper, D. N. W., Gitt, M. A., and Leffler, H. (1994) *J. Biol. Chem.* **269**, 20807–20810
188a. Varela, P. F., Solís, D., Díaz-Maurino, T., Kaltner, H., Gabius, H.-J., and Romero, A. (1999) *J. Mol. Biol.* **294**, 537–549
189. Chiu, M. L., Parry, D. A. D., Feldman, S. R., Klapper, D. G., and O'Keefe, E. J. (1994) *J. Biol. Chem.* **269**, 31770–31776
190. Madsen, P., Rasmussen, H. H., Flint, T., Gromov, P., Kruse, T. A., Honoré, B., Vorum, H., and Celis, J. E. (1995) *J. Biol. Chem.* **270**, 5823–5829
191. Cho, M., and Cummings, R. D. (1995) *J. Biol. Chem.* **270**, 5207–5212
192. Gitt, M. A., Wiser, M. F., Leffler, H., Herrmann, J., Xia, Y.-R., Massa, S. M., Cooper, D. N. W., Lusis, A. J., and Barondes, S. H. (1995) *J. Biol. Chem.* **270**, 5032–5038
193. Weis, W. I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W. A. (1991) *Science* **254**, 1608–1615
194. Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1988) *Science* **242**, 1290–1295
195. Shaanan, B., Lis, H., and Sharon, N. (1991) *Science* **254**, 862–865
196. Cygler, M., Rose, D. R., and Bundle, D. R. (1991) *Science* **253**, 442–445
197. Ångström, J., Teneberg, S., and Karlsson, K.-A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11859–11863
198. Sauter, N. K., Hanson, J. E., Glick, G. D., Brown, J. H., Crowther, R. L., Park, S.-J., Skehel, J. J., and Wiley, D. C. (1992) *Biochemistry* **31**, 9609–9621
199. Janakiraman, M. N., White, C. L., Laver, W. G., Air, G. M., and Luo, M. (1994) *Biochemistry* **33**, 8172–8179
200. Bossart-Whitaker, P., Carson, M., Babu, Y. S., Smith, C. D., Laver, W. G., and Air, G. M. (1993) *J. Mol. Biol.* **232**, 1069–1083
201. Crennell, S. J., Garman, E. F., Laver, E. G., Vimr, E. R., and Taylor, G. L. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9852–9856
202. Colli, W. (1993) *FASEB J.* **7**, 1257–1264
203. Ofek, I., Beachey, E. H., and Sharon, N. (1978) *Trends Biochem. Sci.* **3**, 159–160
204. Lindberg, F., Lund, B., and Normark, S. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1891–1895
205. Cho, J.-W., and Troy, F. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11427–11431
206. Baumann, H., Brisson, J., Michon, F., Pon, R., and Jennings, H. J. (1993) *Biochemistry* **32**, 4007–4013
207. Falk, P. G., Bry, L., Holgersson, J., and Gordon, J. I. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1515–1519
208. Adler, P., Wood, S. J., Lee, Y. C., Lee, R. T., Petri, W. A., and Schnaar, R. L. (1995) *J. Biol. Chem.* **270**, 5164–5171
209. Misevic, G. N., and Burger, M. M. (1990) *J. Biol. Chem.* **265**, 20577–20584
210. Spillmann, D., Hård, K., Thomas-Oates, J., Vliegenthart, J. F. G., Misevic, G., Burger, M. M., and Finne, J. (1993) *J. Biol. Chem.* **268**, 13378–13387
211. Spillmann, D., Thomas-Oates, J. E., van Kuik, J. A., Vliegenthart, J. F. G., Misevic, G., Burger, M. M., and Finne, J. (1995) *J. Biol. Chem.* **270**, 5089–5097
212. Rosati, F., and De Santis, R. (1980) *Nature (London)* **283**, 762–764
213. Seppo, A., Penttilä, L., Niemelä, R., Maaheimo, H., and Renkonen, O. (1995) *Biochemistry* **34**, 4655–4661
214. Litscher, E. S., Juntunen, K., Seppo, A., Penttilä, L., Niemelä, R., Renkonen, O., and Wassarman, P. M. (1995) *Biochemistry* **34**, 4662–4669

References

215. Hitoshi, S., Kusunoki, S., Kanazawa, I., and Tsuji, S. (1995) *J. Biol. Chem.* **270**, 8844–8850
216. de Vries, T., Srnka, C. A., Palcic, M. M., Swiedler, S. J., van der Eijnden, D. H., and Macher, B. A. (1995) *J. Biol. Chem.* **270**, 8712–8722
217. Stroud, M. R., Levery, S. B., Mårtensson, S., Salyan, M. E. K., Clausen, H., and Hakomori, S. (1994) *Biochemistry* **33**, 10672–10680
218. Hull, S. R., Sugarman, E. D., Spielman, J., and Carraway, K. L. (1991) *J. Biol. Chem.* **266**, 13580 –13586
219. Ware, F. E., Vassilakos, A., Peterson, P. A., Jackson, M. R., Lehrman, M. A., and Williams, D. B. (1995) *J. Biol. Chem.* **270**, 4697–4704
220. Leppänen, A., Mehta, P., Ouyang, Y.-B., Ju, T., Helin, J., Moore, K. L., van Die, I., Canfield, W. M., McEver, R. P., and Cummings, R. D. (1999) *J. Biol. Chem.* **274**, 24838 – 24848
221. Chandrasekaran, E. V., Jain, R. K., Larsen, R. D., Wlasichuk, K., and Matta, K. L. (1995) *Biochemistry* **34**, 2925–2936
222. Kogan, T. P., Revelle, B. M., Tapp, S., Scott, D., and Beck, P. J. (1995) *J. Biol. Chem.* **270**, 14047–14055
223. Waddell, T. K., Fialkow, L., Chan, C. K., Kishimoto, T. K., and Downey, G. P. (1995) *J. Biol. Chem.* **270**, 15403–15411
224. Bertozzi, C. R., Fukuda, S., and Rosen, S. D. (1995) *Biochemistry* **34**, 14271–14277
225. Norgard-Sumnicht, K., and Varki, A. (1995) *J. Biol. Chem.* **270**, 12012–12024
226. Hemmerich, S., Leffler, H., and Rosen, S. D. (1995) *J. Biol. Chem.* **270**, 12035–12047
227. Yuen, C.-T., Bezouska, K., O'Brien, J., Stoll, M., Lemoine, R., Lubineau, A., Kiso, M., Hasegawa, A., Bockovich, N. J., Nicolaou, K. C., and Feiz, T. (1994) *J. Biol. Chem.* **269**, 1595–1598
228. Patel, T. P., Goelz, S. E., Lobb, R. R., and Parekh, R. B. (1994) *Biochemistry* **33**, 14815–14824
229. Graves, B. J., Crowther, R. L., Chandran, C., Rumberger, J. M., Li, S., Huang, K.-S., Presky, D. H., Familletti, P. C., Wolitzky, B. A., and Burns, D. K. (1994) *Nature (London)* **367**, 532–538
230. Lasky, L. A. (1995) *Ann. Rev. Biochem.* **64**, 113–139
231. Ginsburg, V. E. (1982) *Meth. Enzymol.* **83**, all
232. Churms, S. C. (1990) *CRC Handbook of Chromatography Carbohydrates*, Vol. 2, CRC Press, Boca Raton, Florida
233. Valent, B. S., Darvill, A. G., McNeil, M., Robertsen, B. K., and Albersheim, P. (1980) *Carbohydrate Res.* **79**, 165–192
234. Chaplin, M. F., and Kennedy, J. F., eds. (1994) *Carbohydrate Analysis A Practical Approach*, Oxford University, Oxford
235. Robyt, J. F., and Mukerjea, R. (1994) *Carbohydr. Res.* **251**, 187–202
236. Green, E. D., and Baenziger, J. U. (1989) *Trends Biochem. Sci.* **14**, 168–172
237. Carr, S. A., Huddleston, M. J., and Bean, M. F. (1993) *Protein Sci.* **2**, 183–196

237a. Fridriksson, E. K., Beavil, A., Holowka, D., Gould, H. J., Baird, B., and McLafferty, F. W. (2000) *Biochemistry* **39**, 3369–3376

237b. Burlingame, A. L., Carr, S. A., and Baldwin, M. A., eds. (2000) *Mass Spectrometry in Biology and Medicine*, Hamana Press, Totowa, New Jersey

238. Downs, F., and Pigman, W. (1969) *Biochemistry* **8**, 1760–1766
239. Narasimhan, S., Harpaz, N., Longmore, G., Carver, J. P., Grey, A. A., and Schacter, H. (1980) *J. Biol. Chem.* **255**, 4876–4884
240. Cohen, R. E., and Ballou, C. E. (1980) *Biochemistry* **19**, 4345–4358
241. Plummer, T. H., Jr., and Tarentino, A. L. (1981) *J. Biol. Chem.* **256**, 10243–10246
242. Risley, J. M., and Van Etten, R. L. (1985) *J. Biol. Chem.* **260**, 15488–15494
243. Takasaki, S., Mizuochi, T., and Kobata, A. (1982) *Meth. Enzymol.* **83**, 263–268
244. Parente, J. P., Wieruszeski, J.-M., Strecker, G., Montreuil, J., Fournet, B., van Halbeek, H., Dorland, L., and Vliegenthart, J. F. G. (1982) *J. Biol. Chem.* **257**, 13173–13176
245. Patel, T., Bruce, J., Merry, A., Bigge, C., Wormald, M., Jaques, A., and Parekh, R. (1993) *Biochemistry* **32**, 679–693
246. Sweeley, C. C., and Nunez, H. A. (1985) *Ann. Rev. Biochem.* **54**, 765–801
247. Dua, V. K., Roa, B. N. N., Wu, S.-S., Dube, V. E., and Bush, C. A. (1986) *J. Biol. Chem.* **261**, 1599–1608
248. Tsai, P.-K., Dell, A., and Ballou, C. E. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4119–4123
249. Lindberg, B., Lönngren, L., and Svensson, S. (1975) *Adv. Carbohydr. Chem. Biochem.* **31**, 185–240
250. Guthrie, R. D., and McCarthy, J. F. (1967) *Adv. Carbohydr. Chem. Biochem.* **22**, 11–23
251. Edge, C., Parekh, R., Rademacher, T., Wormald, M., and Dwek, R. (1992) *Nature (London)* **385**, 693–694
252. Varki, A. (1991) *FASEB J.* **5**, 226–235
253. Rolf, D., and Gray, G. R. (1982) *J. Am. Chem. Soc.* **104**, 3539–3541
254. Mukerjea, R., Kim, D., and Robyt, J. F. (1996) *Carbohydrate Res.* **292**, 11–20
255. Bobbitt, J. M. (1956) *Adv. Carbohydr. Chem. Biochem.* **11**, 1–41
256. Rich, A., Nordheim, A., and Wang, A. H.-J. (1984) *Ann. Rev. Biochem.* **53**, 791–846
257. Homans, S. W., Dwek, R. A., and Rademacher, T. W. (1987) *Biochemistry* **26**, 6571–6578
258. Homans, S. W., Pastore, A., Dwek, R. A., and Rademacher, T. W. (1987) *Biochemistry* **26**, 6649 –6655
259. Cumming, D. A., Shah, R. N., Krepinsky, J. J., Grey, A. A., and Carver, J. P. (1987) *Biochemistry* **26**, 6655–6663
260. Reuben, J. (1985) *J. Am. Chem. Soc.* **107**, 1747–1755
261. Cumming, D. A., Dime, D. S., Grey, A. A., Krepinsky, J. J., and Carver, J. P. (1986) *J. Biol. Chem.* **261**, 3208–3213
262. Vuister, G. W., de Waard, P., Boelens, R., Vliegenthart, J. F. G., and Kaptein, R. (1989) *J. Am. Chem. Soc.* **111**, 772–774
263. Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry* **28**, 5978–5984
264. Abeygunawardana, C., and Bush, C. A. (1991) *Biochemistry* **30**, 8568–8577
265. Rivière, M., and Puzo, G. (1992) *Biochemistry* **31**, 3575–3580
266. van Duynhoven, J. P. M., Goudriaan, J., Hilbers, C. W., and Wijmenga, S. S. (1992) *J. Am. Chem. Soc.* **114**, 10055–10056
267. Wu, J., Bondo, P. B., Vuorinen, T., and Serianni, A. S. (1992) *J. Am. Chem. Soc.* **114**, 3499–3505
268. Homans, S. W. (1990) *Biochemistry* **29**, 9110–9118
269. Bose, B., Zhao, S., Stenutz, R., Cloran, F., Bondo, P. B., Bondo, G., Hertz, B., Carmichael, I., and Serianni, A. S. (1998) *J. Am. Chem. Soc.* **120**, 11158–11173
270. Flugge, L. A., Blank, J. T., and Petillo, P. A. (1999) *J. Am. Chem. Soc.* **121**, 7228 – 7238

Study Questions

1. A nonreducing disaccharide gives an octamethyl derivative with dimethyl sulfate and alkali. On acid hydrolysis, this derivative yields 1 mol of 2,3,4,6-tetramethyl-D-glucose and 1 mol of 2,3,4,6-tetramethyl-D-galactose. The disaccharide is hydrolyzed rapidly by either maltase or lactase (a β-galactosidase).

   Give an adequately descriptive name of the disaccharide, and draw its Haworth projection formula.

2. An aldopentose (A) of the D-configuration on oxidation with concentrated nitric acid gives a 2,3,4-trihydroxypentanedioic acid (a trihydroxyglutaric acid) (B) which is optically inactive. (A) on addition of HCN, hydrolysis, lactonization, and reduction gives two stereoisomeric aldohexoses (C) and (D). (D) on oxidation affords a 2,3,4,5-tetrahydroxy-hexanedioic acid (a saccharic acid) (E) which is optically inactive. Give structures of compounds (A)-(E).

3. What products are formed when periodic acid reacts with sorbitol?
4. A 10.0 g sample of glycogen gave 6.0 millimol of 2,3-di-*O*-methylglucose on methylation and acid hydrolysis.

   a. What percent of the glucose residues in glycogen have chains substituted at the α–1→6 position?

   b. What is the average number of glucose residues per chain?

   c. How many millimols of 2,3,6-tri-*O*-methylglucose were formed?

   d. If the molecular weight of the polysaccharide is 2×10^6, how many glucose residues does it contain?

   e. How many nonreducing ends are there per molecule or equivalently how many chains are there per molecule?

5. D-Mannitol is a symmetric molecule, yet it is optically active. Explain.

6. When D-glucose is treated with acidic methanol, the first products which can be isolated are mainly methyl furanosides, but after extensive reaction the furanosides disappear and methyl glucopyranosides accumulate. Why?

7. Write the structural formulas for (a) 1,6 anhydro β-D-glucopyranose; (b) 1,6 anhydro β-D-altrose. Compound (b) is many times more stable than compound (a). Explain this on stereochemical grounds.

8. The disaccharide *nigerose* is α-D-Glu*p*-(1→3)-D-glu. Write out its structure. How would you prove this structure using methylation, periodate oxidation, and other methods.

9. Inositol is 1,2,3,4,5,6-hexahydroxycyclohexane. Draw configurational formulas for all possible stereoisomers and indicate which would be expected to be optically active.

10. Why do you suppose that the major form of D-fructose in solution is the pyranose form but D-fructose in sucrose is in the furanose form?

11. D-Xylose is an easily prepared sugar, potentially available in enormous quantity. What is a common source? How can it be obtained from this source?

12. How can xylitol be obtained from xylose? Discuss the stereochemical properties of xylitol.

13. The enzyme xylose isomerase is important industrially. Why?

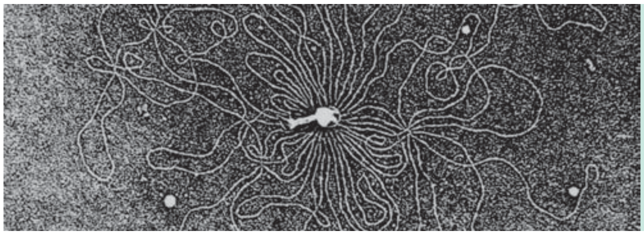
14. Glucose reacts non-enzymatically with amino acids and proteins, including hemoglobin, egg-white proteins and serum albumin. For example, if glucose is not removed prior to drying, dried egg whites slowly turn brown and develop off-flavors and odors. What do you propose as the most likely first step in the non-enzymatic glucose-protein chemical reaction? How can the first product transform spontaneously into a ketose derivative?

15. What characteristics would you expect in a binding site for a sugar ring in an enzyme, lectin, or other proteins?

16. Using structural formulas, descirbe the two major types of linkage of carbohydrate chains or clusters to proteins to form glycoproteins or proteoglycans.

17. What products would you expect from cellulose as a result of methylation analysis? Periodate oxidation? The Smith degradation? The action of an alpha amylase?

18. When glycoproteins are treated with alkaline brohydride, amino acid analysis often indicates a decrease in the amount of serine and a corresponding increase in the amount of alanine, or a decrease in threonine with the appearance of α-aminobutyric acid. Explain.



DNA spreading from the broken head of a bacteriophage T2 phage. This classic electron micrograph, published by A. K. Kleinschmidt and coworkers in 1962 (*Biochem. Biophys. Acta.* **61**, 857–864, 1962) was prepared by spreading the phage particles suspended in a protein–salt solution as a mixed monolayer on a water-air interface. The resultant osmotic shock burst the head and confined the DNA as a single thread near the phage ghost. After transfer to a suitable carbon surface, removal of water, and shadowing with platinum, the micrograph was obtained. Courtesy of Albrecht K. Kleinschmidt

Contents

203 **A. Structure and Chemistry of Nucleotides**
203 1. Names and Abbreviations
204 2. Acid–Base Chemistry and Tautomerism
205 3. Absorption of Ultraviolet Light
207 4. Chemical Reactions of the Bases
207 5. Base Pairs, Triplets, and Quartets
209 *Strengths of base pairs*
209 *Stacking of bases*
211 *Tautomerism and base pairing*
211 6. Conformations of Nucleotides
213 **B. Double Helices**
213 1. The B Form of DNA
213 2. Other Double-Helical Forms of DNA
216 3. The Conformational Flexibility of Double Helices
217 *Rotational and translational movements of bases*
217 *Bends and bulges*
218 *Interactions with ions*
218 **C. The Topology and Dynamics of Nucleic Acids**
218 1. Rings, Catenanes, and Knots
219 2. Supercoiled DNA
222 3. Intercalation
226 4. Polynucleotides with Three or Four Strands
226 *H-DNA*
226 *R-DNA*
227 *Designed third strands*
227 *Tetraplex (quadruplex) structures*
228 5. Junctions
230 **D. Ribonucleic Acids (RNA)**
230 1. RNA Loops and Turns
231 2. Transfer RNAs
232 3 . Properties and Structures of Some Other RNA Molecules
234 4. Modified Nucleotides
234 5. RNA Aptamers
235 **E. The Genetic Code**
236 1. The "Reading Frames"
237 2. Variations
238 3. Palindromes and Other Hidden Messages
239 4. The Base Composition of DNA and RNA
239 **F. Interaction of Nucleic Acids with Proteins**
239 1. The Helix–Turn–Helix Motif
241 2. Other DNA-binding Motifs
241 *Leucine zipper proteins*
241 *Zinc fingers*
241 *Beta ribbons*
243 *The winged helix family*
243 3. RNA-Binding Proteins
244 **G. Viruses and Plasmids**
244 1. Viruses with Single-Stranded DNA
244 2. Viruses with Double-Stranded DNA
247 3. Viruses Containing RNA
248 4. Viruses without Nucleic Acid?
248 5. Life Cycles
248 6. Plasmids and Transposable Genetic Elements
249 **H. Methods of Study**
249 1. Isolation and Separation of Nucleic Acids
249 2. Hydrolysis and Analysis
251 3. Characteristic Reactions of the Bases and Backbone
251 *Reactions of nucleophiles*
253 *Reactions with electrophilic reagents*
254 *Reactions causing cleavage of the sugar-phosphate backbone*
255 4. Melting, Hybridization, and Polynucleotide Probes
258 5. Synthesis of Oligonucleotides and Polynucleotides
260 6. The Polymerase Chain Reaction (PCR)
260 7. Sequence Determination
260 *Preparing the DNA*
260 *Restriction maps and Southern blots*
262 *The Sanger dideoxy method*
264 *The method of Maxam and Gilbert*
264 *Sequencing RNA*
265 *Nearest neighbor analysis*
265 *Understanding sequences*
266 8. Protein–DNA Interactions
266 9. Nuclear Magnetic Resonance

270 **References**
278 **Study Questions**

Boxes

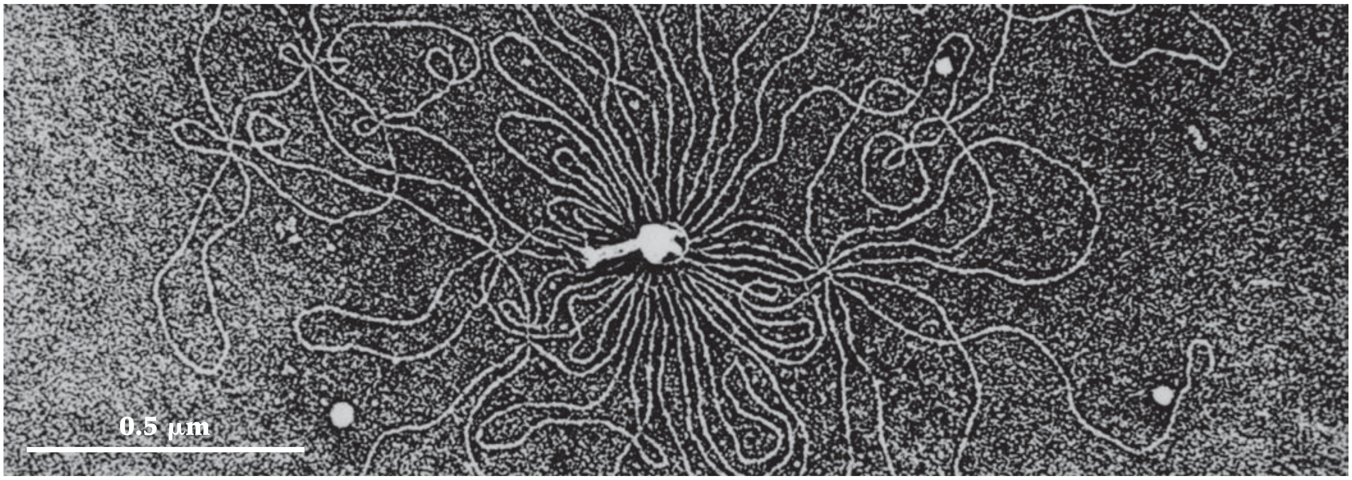
206 Box 5-A The Alkali Metal Ions
224 Box 5-B Antitumor DNA Drugs
235 Box 5-C The RNA World
259 Box 5-D DNA Fingerprinting

Tables

203 Table 5-1 Names of Pyrimidine and Purine Bases, Nucleosides, and 5'-Nucleotides
210 Table 5-2 Gibbs Energies of Formation ΔG_f° at 25°C for Addition of One Base Pair to an Existing RNA Helix
216 Table 5-3 Helix Parameters for Three Types of DNA
232 Table 5-4 Some Specialized RNA Molecules
236 Table 5-5 The Genetic Code
237 Table 5-6 The Sixty-Four Codons of the Genetic Code
245 Table 5-7 Characteristics of Some Individual Viruses and Groups of Viruses

5

The Nucleic Acids



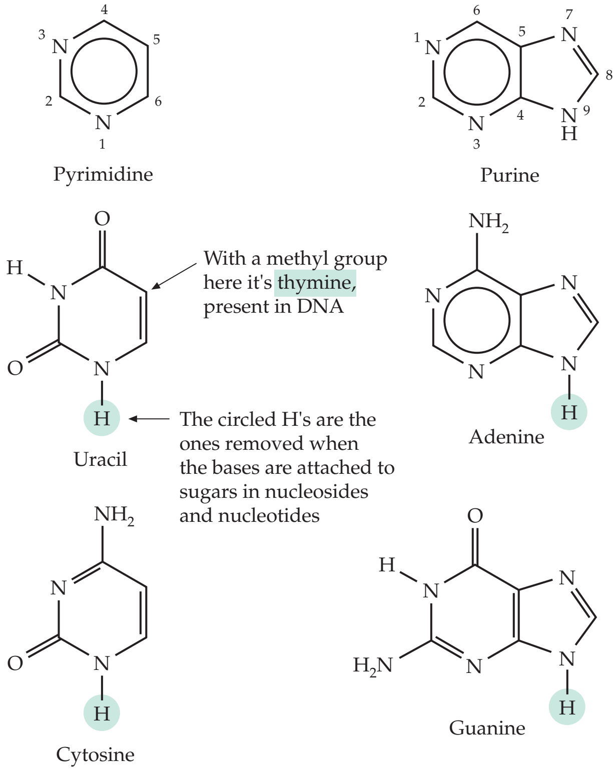
The phosphorus- and nitrogen-containing materials that came to be known as nucleic acids were first isolated from cells around 1870 by Friedrich Miescher but were long regarded as something of a curiosity.[1] Nevertheless, the structures of the monomer units, the **nucleotides**, were established by 1909 and the correct **polynucleotide** structure of the chains of DNA and RNA was proposed by Levene and Tipson in 1935.[2,3]

The nucleotides are made up of three parts:

1. One of the **pyrimidine** or **purine "bases"** uracil, cytosine, adenine, or guanine (Fig. 5-1). All four of these bases are present in RNA, while DNA contains thymine instead of uracil. Atoms in the bases are numbered 1–6 or 1–9.
2. A **sugar**, either D-ribose or D-2-deoxyribose. Carbon atoms in sugars are numbered 1’–5.’
3. **Phosphoric acid**

Although the biological synthesis is indirect, we can imagine that nucleotides are formed from these parts by elimination of two molecules of water as indicated in Eq. 5-1. In nucleic acids the nucleotides are combined through phosphodiester linkages between the 5'-hydroxyl of the sugar in one nucleotide and the 3'-hydroxyl of another. Again, we can imagine that these linkages were formed by the elimination of water (Eq. 5-2). The structures of a pair of short polynucleotide strands in DNA are shown in Fig. 5-2. That of a segment of double-helical DNA is shown in Fig. 5-3 and that of a transfer RNA in Figs. 5-30 and 5-31.

Despite the fact that Levene had deduced the correct structure for polynucleotides, he was thrown off the trail of a deeper understanding by the roughly equal amounts of the four bases found in either DNA or RNA. He assumed that nucleic acids must be regular repeating polymers for which there was no obvious biological function. It was not until 1944 that there

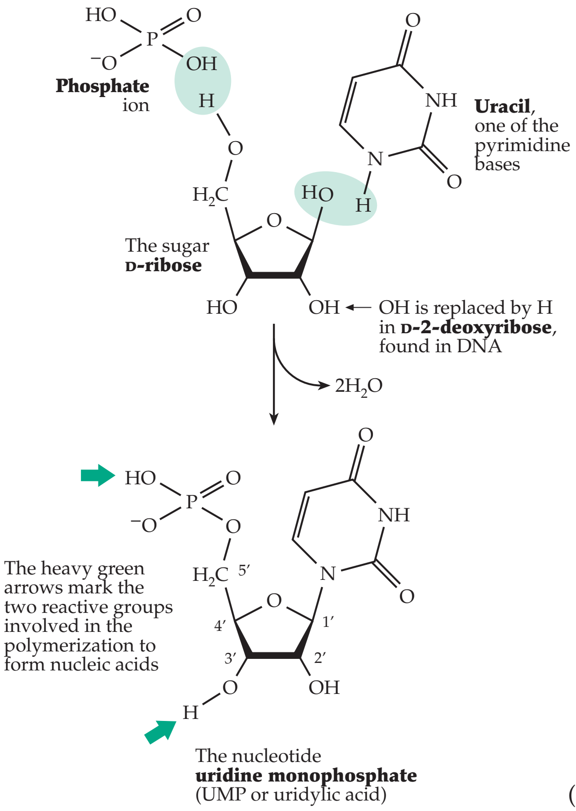


With a methyl group here it's thymine, present in DNA

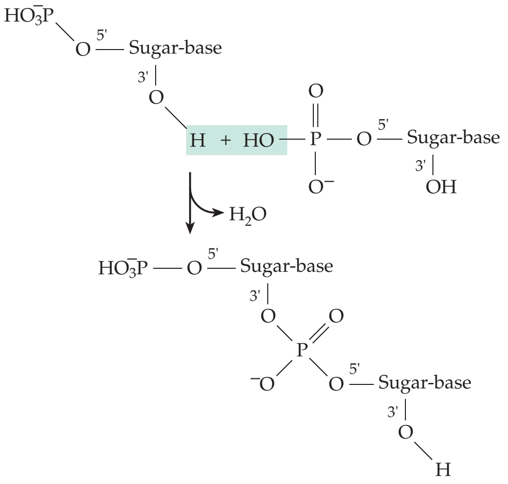
The circled H's are the ones removed when the bases are attached to sugars in nucleosides and nucleotides

Note: all of these molecules are almost perfectly flat!

Figure 5-1 Structures of the major pyrimidine and purine bases of DNA and RNA.


HO
P
O
−O
OH
H
Phosphate ion
O
NH
Uracil, one of the pyrimidine bases
N
O
O
H_2C
HO
H
O
The sugar **D-ribose**
HO
OH ← OH is replaced by H in **D-2-deoxyribose**, found in DNA
$2H_2O$
O
HO
P
O
−O
O
NH
The heavy green arrows mark the two reactive groups involved in the polymerization to form nucleic acids
H_2C 5′
N
O
O
4′
1′
3′
2′
O
OH
H
The nucleotide **uridine monophosphate** (UMP or uridylic acid)

(5-1)


HO_3^-P
O 5′ Sugar-base
3′
O
H + HO — P — O 5′ Sugar-base
O
O^-
3′
OH
H_2O
HO_3^-P — O 5′ Sugar-base
3′
O
O
P
−O
O 5′ Sugar-base
3′
O
H

(5-2)

was concrete evidence that DNA carried genetic information (see Chapter 26). However, it was James Watson and Francis Crick's recognition of the double-helical structure of DNA[4-9] in 1953 and the mechanism of replication that this structure implied that captured the imagination of biologists and chemists alike and paved the way for the present-day explosion of knowledge of DNA, RNA, and of the encoded proteins.

Watson and Crick proposed that DNA is a double helix of two antiparallel polynucleotide chains (Figs. 5-2 and 5-3). The structure was deduced from model building together with knowledge of the X-ray diffraction data of Maurice F. Wilkins and Rosalind Franklin[9a] on artificially formed DNA fibers. An additional key piece of information was the discovery by Erwin Chargaff that *in all double-stranded DNA the content of adenine equals that of thymine and the content of guanine equals that of cytosine.*

The most significant feature of the proposed structure was the pairing of bases in the center through hydrogen bonding. The pairs and triplets of hydrogen bonds (Fig. 5-2) could form in the manner shown only if adenine (A) was paired with thymine (T) and cytosine (C) with guanine (G) at every point in the entire DNA structure. Thus, *the nucleotide sequence in one chain is complementary to but not identical to that in the other chain.* It was apparent almost immediately that the sequence of bases in a DNA chain must convey the encoded genetic information. The complementarity of the two strands suggested a simple mechanism for replication of genes during cell divisions. The two strands could separate and a complementary strand could be synthesized along each strand to give two molecules of the DNA, one for each of the two cells.

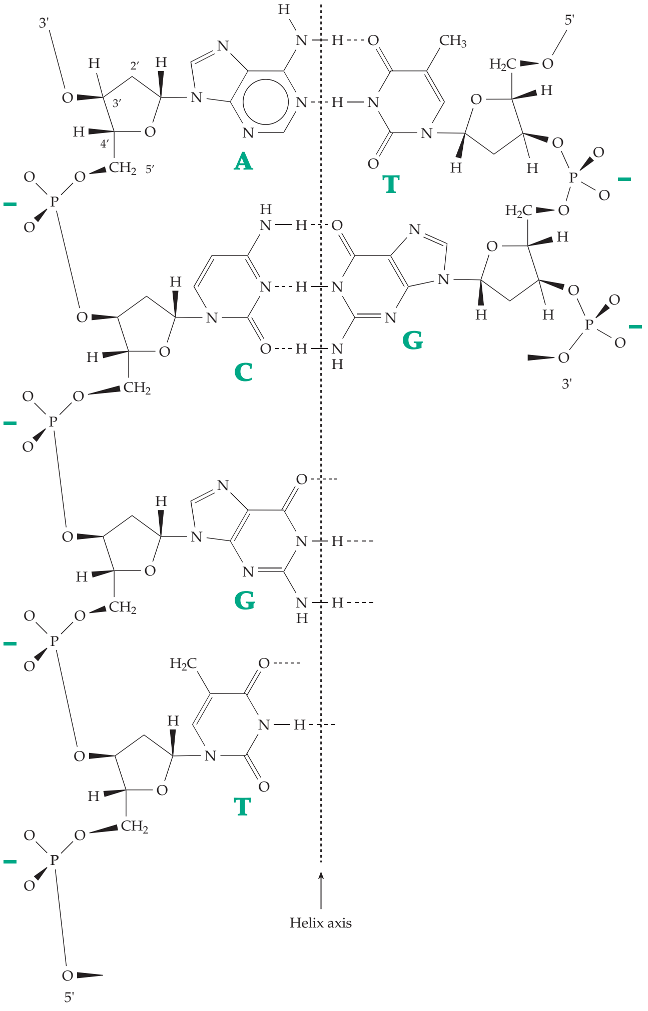

3′
5′
A
T
C
G
G
T
CH_3
H_2C
CH_2
5′
3′
–
Helix axis
3′
5′

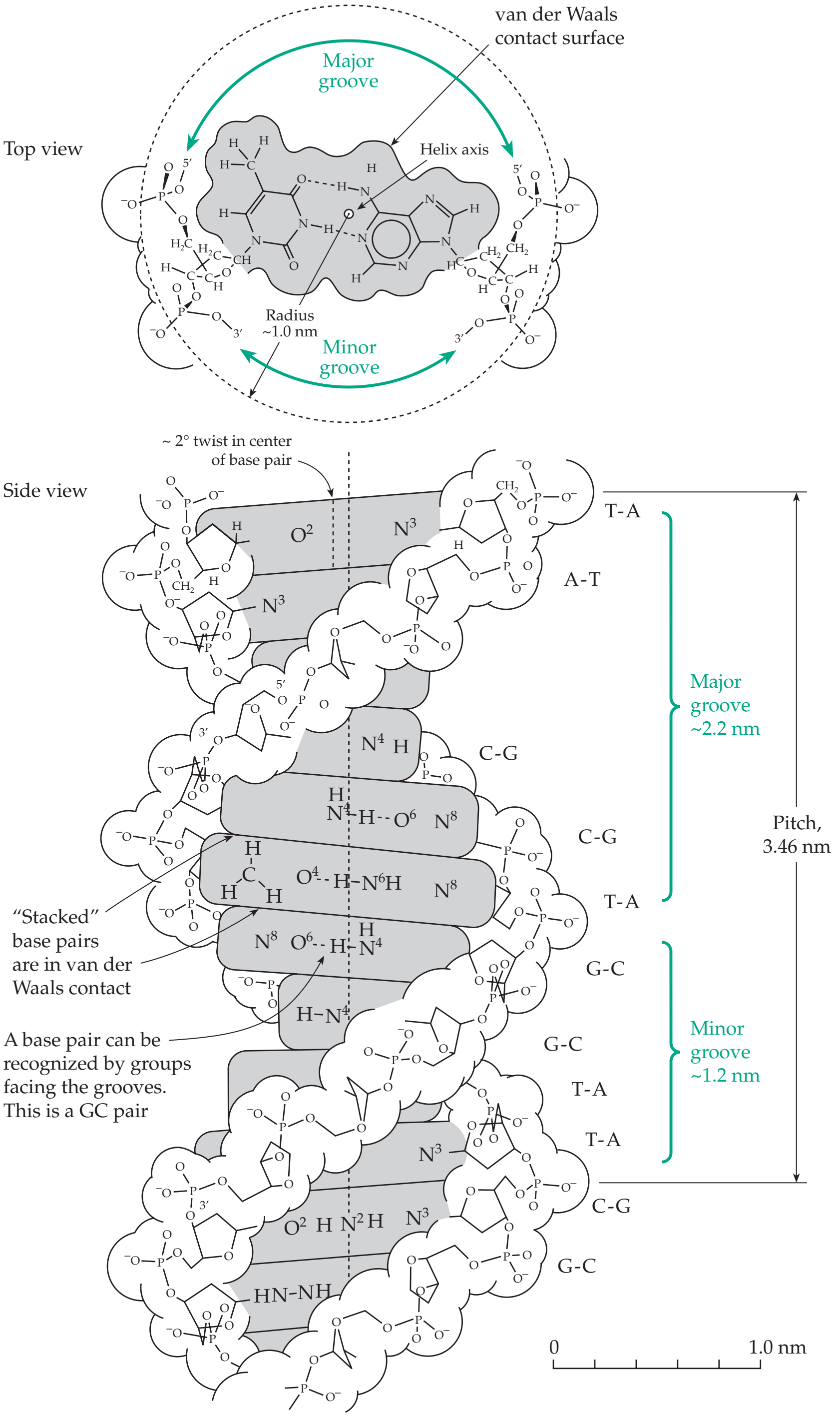
Figure 5-2 A distorted (flattened) view of the Watson–Crick structure of DNA showing the hydrogen-bonded base pairs.

The essential correctness of the concept has been proved.

Two extremely important developments came in the 1970s: (1) Methods were found for cutting and rejoining DNA fragments and for **cloning** them in bacteria and (2) ways were devised for rapid determination of **nucleotide sequences**. The application of these techniques is now providing startling advances in biology and medicine. In 1970 we knew virtually nothing about the sequences of nucleotides in genes but today, we know the sequences for many thousands of genes of all types. By the 1980s sequences had been established for hemoglobin,[11] γ-globulins,[12] collagen,[13] and for many enzymes. An example is the gene sequence for mitochondrial aspartate aminotransferase (Fig. 5-4).[14] Its coding regions consist of 1299 pairs of mononucleotides, the bases being paired as in Fig. 5-2 in double helical form. Complete DNA sequences are known for numerous viruses including the 9740-base pair (bp) DNA provirus form of the RNA virus that causes AIDS[15] and for bacterial viruses such as T7 bacteriophage (39,936 bp).[16] Also determined in the 1980s were sequences of human mitochondrial DNA (16,598 bp),[17] and of chloroplast DNA from the tobacco plant (155,844 bp).[18]

In 1995 the first complete sequences of bacterial genomes were obtained (Table 1-3).[19,20] These were followed by sequences of many other bacterial genomes,[21] including the 4.2×10^6 bp *E. coli* genome (Table 1-3). Sequencing of the 16 chromosomes of the 12.07 Mbp genome of yeast[22,23] containing ~6300 genes was completed in 1996[24] and that of the 97 Mbp genome of the nematode *Caenorhabditis elegans* in 1998. By 2000 the sequence of the 180 Mbp genome of the fruit fly *Drosophila melanogaster* was largely completed, and most sequences of the 3×10^9 bp human genome were known.[24a]

Cloning of DNA has not only provided an essential step in sequence determination but also has given birth to a new industry devoted to producing proteins from genes cloned in bacteria, yeast, or other cells. Human insulin and the antiviral protein interferon were two of the first proteins produced in this way. Methods now in use permit us to introduce at will alterations at any point in a DNA sequence. We are able to locate and study the genes responsible for many genetic defects.



"Stacked" base pairs are in van der Waals contact

A base pair can be recognized by groups facing the grooves. This is a GC pair

Figure 5-3 The double-helical structure of DNA. The structure shown is that of the B form and is based on coordinates of Arnott and Hukins.[10] The major and minor grooves, discussed on p. 213, are marked.

```
                                     -200         .         .         .         .       -150       .        .         .
                                 5'- CATGACCTCCGTGGAGGCGTCGTCGTTCTACCCCCGACCGCAAGGGGCCGCTGACATGATTCGCTCCAGCGCAGGGCCGCT
       .      -100         .         .         .         .       -50         .          .         .        .         +1
           ↓        ↓         ↓     ↓                              ↓                                                  Me
CTAGCCACGCCCAGGGAAGTCACTGTCCTCACCTTTTAGGAGCCCGCGCCTCGGTTCCAGCGGACGCTTCCCCAGATCTCGGCTCTACCACCATCCACTGCCGTCTTACCGCCCACC AT

t Ala Leu Leu His Ser Ser Arg Ile Leu Ser Gly Met Ala Ala Ala Phe His Pro Gly Leu Ala Ala Ala Ala Ser Ala Arg Ala Se
G GCC CTC CTG CAC TCC AGC CGC ATC CTC TCC GGG ATG GCT GCT GCC TTT CAC CCT GGC CTA GCT GCT GCA GCC TCT GCC AGA GCC AG GTG
                                                                        A
                                                                       INTRON 1                                    r Ser
AGCCGAGGGATACAGAGATGCAGCGGCACCGGGCCTGCCCTGCCAGCCGAAGTGTGGACCCTGTGAA----(10.2 kb)----ATGGCTGGGGTAACCTATTTCTCATTTCAG C TCC

Trp Trp Thr His Val Glu Met Gly Pro Pro Asp Pro Ile Leu Gly Val Thr Glu Ala Phe Lys Arg Asp Thr Asn Ser Lys Lys Met Asn
TGG TGG ACC CAT GTT GAA ATG GGA CCT CCA GAT CCC ATC CTG GGC GTT ACC GAA GCC TTC AAG AGA GAT ACC AAC AGC AAG AAG ATG AAC

Leu Gly Val Gly Ala Tyr Arg Asp Asp Asn Gly Lys Pro Tyr Val Leu Pro Ser Val Arg Lys A
CTG GGA GTT GGT GCC TAC CGG GAT GAT AAC GGA AAA CCT TAC GTG CTC CCC AGT GTC CGG AAG G GTGAGCTTGGCACTCGTCTCCTGCCAGCTAGGAT

          INTRON 2                                    la Glu Ala Gln Ile Ala Ala Lys Asn Leu Asp Lys Glu Tyr Leu Pro Ile
GTGGAA----(1.4 kb)----AAAAACTAAGACTTATGATTTTCTTGTCTAG CA GAG GCC CAG ATT GCT GCA AAA AAT TTG GAC AAA GAA TAC CTG CCC ATT

Gly Gly Leu Ala Glu Phe Cys Lys Ala Ser Ala Glu Leu Ala Leu Gly Glu Asn Asn Glu Val Asn Glu Val Leu Lys Ser Gly Arg
GGG GGA CTG GCT GAA TTC TGT AAG GCT TCT GCA GAA CTG GCC CTG GGC GAG AAC AAT GAA GTG AAT GAA GTG TTG AAA AGC GGC CGG GTAA

                                INTRON 3                                             Phe Val Thr Val Gln Thr Ile Ser Gly
GCCAGCGGAGTCGGGCTTGAGCTTGATA----(1.2 kb)----TGTAAAGGGTAGAGAGAGTGACTCTGTGTATCCCCTGCAG TTC GTC ACT GTG CAG ACC ATT TCC GGG

Thr Gly Ala Leu Arg Val Gly Ala Ser Phe Leu                                             INTRON 4
ACT GGA GCC TTA AGG GTC GGG GCC AGT TTT CTG GTCAGTGGAACTCTTTCAAGAATGAATCTTTTGGGGTGGC----(2.2 kb)----CTTTCTTCATTTTCTCATTC

                     Gln Arg Phe Phe Lys Phe Ser Arg Asp Val Phe Leu Pro Lys Pro Ser Trp Gly Asn His Thr Pro Ile Phe Arg
CCTTTTATCCCGACTTTTAG CAA AGG TTT TTT AAG TTC AGC CGA GAT GTC TTT CTG CCC AAA CCA TCC TGG GGA AAT CAC ACG CCC ATC TTC AGG

Asp Ala Gly Met Gln Leu Gln Gly Tyr Arg Tyr Tyr Asp Pro Lys Thr Cys Gly Phe Asp Phe Ser Gly Ala Leu Glu Asp Ile Ser
GAT GCC GGC ATG CAG CTC CAA GGT TAT CGC TAC TAT GAC CCC AAG ACT TGC GGT TTT GAC TTC TCC GGA GCC CTA GAA GAC ATA TCA GTAA

                              INTRON 5                                           Lys Ile Pro Glu Gln Ser Val Leu Leu Leu
GTGTGGCTTTCCAGGCCCGACTTCTG----(0.2 kb)----GAAGCTGCACAGCCAAAAATCTCGATGTTTCTCCTTAG AAA ATC CCA GAG CAG AGT GTC CTC CTC CTG

His Ala Cys Ala His Asn Pro Thr Gly Val Asp Pro Arg Pro Glu Gln Trp Lys Glu Ile Ala Ser Val Val Lys
CAT GCC TGC GCT CAC AAC CCC ACC GGC GTG GAC CCG CGT CCC GAG CAG TGG AAG GAG ATA GCG TCC GTG GTG AAG GTGAGGAGGATGAAGCGTCA

                      INTRON 6                                   Lys Lys Asn Leu Phe Ala Phe Phe Asp Met Ala Tyr Gln Gly
GGAGCTGGTTGCTTAACC----(0.2 kb)----TCAAACACTGGTCTTCTCATTCCTTCCCAG AAA AAG AAT CTC TTC GCA TTC TTT GAC ATG GCC TAC CAA GGC

Phe Ala Ser Gly Asp Gly Asp Lys Asp Ala Trp Ala Val Arg His Phe Ile Glu Gln Gly Ile Asn Val Cys Leu Cys Gln Ser Tyr Ala
TTT GCC AGC GGT GAT GGT GAT AAG GAT GCC TGG GCC GTG CGG CAC TTC ATC GAG CAG GGC ATC AAT GTC TGC CTC TGC CAA TCG TAT GCC

Lys Asn Met Gly Leu Tyr G                                    INTRON 7                                             ly Glu
AAG AAC ATG GGC CTG TAC G GTAAGCCAGAGGGCCCAATATAAGGTGTGTT----(1.2 kb)----GTTTCAGCCTGCTAGAAGACAGTATCCTCTGTCTTTTCAG GT GAG

Arg Val Gly Ala Phe Thr Val Val Cys Lys Asp Ala Glu Glu Ala Lys Arg Val Glu Ser Gln Leu Lys Ile Leu Ile Arg Pro Leu Tyr
CGT GTG GGA GCC TTC ACG GTG GTC TGC AAA GAT GCA GAA GAA GCC AAA AGG GTG GAG TCA CAG CTG AAG ATC TTG ATC CGT CCC CTG TAT

Ser Asn Pro Pro Leu Asn Gly Ala Arg Ile Ala Ala Thr Ile Leu Thr Ser Pro Asp Leu Arg Lys Gln Tr
TCC AAC CCA CCT CTC AAT GGG GCC CGG ATC GCA GCA ACC ATC CTG ACT TCT CCA GAC TTC CGG AAG CAA TG GTAACGATTACTAGCTGTATACCGT

                    INTRON 8                                   p Leu Gln Glu Val Lys Gly Met Ala Asp Arg Ile Ile Ser Met
GACTACAGCTCCATGA----(2.4 kb)----GAGCACGGCATCCCTCTGCTTTCCTCACAG G TTG CAA GAG GTG AAA GGC ATG GCT GAC CGC ATC ATC AGC ATG

Arg Thr Gln Leu Val Ser Asn Leu Lys Lys Glu Gly Ser Ser His Asn Trp Gln His Ile Thr Asp Gln Ile Gly Met Phe Cys Phe Thr
AGG ACC CAG CTG GTC TCC AAC CTG AAG AAA GAG GGC TCT TCC CAC AAC TGG CAG CAC ATC ACC GAC CAG ATC GGC ATG TTC TGT TTC ACC

Gly Leu Lys Pro Glu Gln                                         INTRON 9                                         Val Glu
GGC CTA AAG CCA GAG CAG GTGAGTGGGCTCTGGTTCCCCACAGCCAACCCCCGC----(2.0 kb)----ACCAGAAGTGGTTATCTAGCTCTTTTCTTTCTGCAG GTA GAG

Arg Leu Thr Lys Glu Phe Ser Val Tyr Met Thr Lys Asp Gly Arg Ile Ser Val Ala Gly Val Thr Ser Gly Asn Val Gly Tyr Leu Ala
CGG CTG ACC AAG GAG TTC TCG GTC TAC ATG ACA AAG GAT GGC CGA ATC TCC GTG GCA GGG GTC ACC TCT GGC AAT GTG GGC TAC CTT GCC

His Ala Ile His Gln Val Thr Lys  *
CAT GCC ATT CAC CAG GTC ACC AAG TAA TTCCCAAGTGCAAAGGAACAGAGACCACTTTCCCGACAGCCTTTGCGCTCGTGAGCGTCACGTGCAGGGTGAGGGAGGGTGGGT
                                          G
GGTGGTGAATAGATCCGGTTTCCAACCACGGTGCATAACTCAGCGACTGAACTCGCTCCTCGGAAGAGAGGTAGGGCAGAGGCTTCCTCGGCTGGTATCTGGAACTTCGTCGGCTCTAAA

CCAAACTCCCTCATCCTTTTGTCTCCAGCTTTTCTGAAAGTTTACACACGCAAGAAAAATCACAGCACCAAACACCTGTCAGCCATGGCACTGGAATAGG|←   620 bases   →|
                                                                    ↓
CACCATCTGCTCTAATCATGTAGACGTACTGCCGCCTGGTTTCTCTGTTACAATAAAATTACTATAGACCCAGCCACTGTCTGCTTTTCTACTTACTGAGAGGGAAGGGGTAAGAGTGGA -3'
```

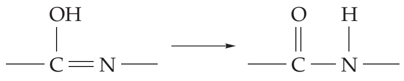
Figure 5-4 The nucleotide sequence of the gene for mitochondrial aspartate aminotransferase from the mouse. From Tsuzuki *et al.*[14] The gene encodes a 433-residue protein requiring 1299 nucleotide pairs (1.3 kb). However, 29 residues are cut off from the N terminus to form the mature mitochondrial protein. In addition, the gene is split by nine introns which vary in length from 0.2–10.2 kb. The sequences at the ends of the introns are shown. There is also an "upstream" region (of which 200 nucleotides are shown) at the 5'-end of the gene. It contains two binding sites for **transcription factor Sp1** (boxed). At the 3' end the 993 additional nucleotides contain signals (boxed) surrounding the **polyadenylation site** (green arrow) for 3′ processing and termination of transcription. The mature messenger RNA is about 2400 nucleotides (2.4 kb) in length but the gene, with introns, occupies about 25 kb. The +1 marks the position of the first nucleotide of the initiation codon ATG (encoding methionine) and the asterisk (*) the termination codon TAA. These are AUG and UAA in the mRNA. Other codons are indicated by amino acid abbreviations.

The first successes in using virus-like particles to carry new pieces of DNA into the cells to help correct these defects have been reported. Our ability to breed new varieties of plants and microorganisms has been enormously enhanced. We can foresee the production of artificially designed enzymes to conduct many industrial chemical processes. These are among the many reasons for the excitement today in the fields of nucleic acid chemistry and molecular genetics.

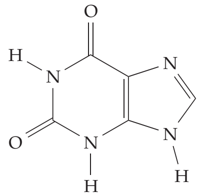
A. Structure and Chemistry of Nucleotides

1. Names and Abbreviations

The purine and pyrimidine ring compounds found in nucleic acids are known as "bases," even though some of them have almost no basic character. **Nucleosides** are the *N*-glycosyl derivatives of the bases with ribose or 2-deoxyribose. The **nucleotides** are phosphate esters of nucleosides. Similar names are applied to related compounds such as adenosine triphosphate (ATP) that are not present in DNA or RNA. The names of the principal nucleotides from which the nucleic acids are formed are given in Table 5-1. The novice may find these confusing! Even worse than the names in the table is **hypoxanthine** (Hyp), which is derived from adenine by replacement of its $-NH_2$ group with –OH and tautomerization:



The nucleoside formed from hypoxanthine and ribose is known as **inosine** (Ino or I) and the corresponding nucleotide as **inosinic acid**. Further substitution at C-2 of –H by –OH and tautomerization yields **xanthine** (Xan). Its nucleoside is xanthosine (Xao, X). A similar hydroxylation at C-7 converts xanthine to **uric acid**, an important human urinary excretion product derived from nucleic acid bases.


Xanthine

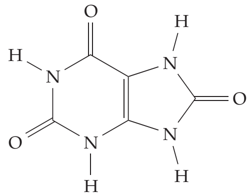

Uric acid

TABLE 5-1
Names of Pyrimidine and Purine Bases, Nucleosides, and 5'-Nucleotides[a]

A Nucleotide units of RNA (abbreviations in parentheses)[b]

Base:	Uracil (Ura)	Cytosine (Cyt)	Adenine (Ade)	Guanine (Gua)
Nucleoside:	Uridine (Urd or U)	Cytidine (Cyd or C)	Adenosine (Ado or A)	Guanosine (Guo or G)
5'-Nucleotide:	Uridine 5'-phosphate or 5'-uridylic acid (Urd-5'-*P* or UMP)	Cytidine 5'-phosphate or 5'-cytidylic acid (Cyd-5'-*P* or CMP)	Adenosine 5'-phosphate or 5'-adenylic acid (Ado-5'-*P* or AMP)	Guanosine 5'-phosphate or 5'-guanylic acid (Guo-5'-*P* or GMP)

B Nucleotide units of DNA

These contain 2-deoxyribose and the nucleosides and nucleotides are called deoxyadenosine (dAdo or dA), deoxyadenosine 5'-phosphate (dAMP), etc.

DNA contains thymine (Thy) rather than uracil. The deoxyribose derivatives are thymidine (dThd or dT) and thymidine 5'-phosphate. The ribose derivatives of thymine are the nucleoside ribosylthymidine (Thd) and ribosylthymidine 5'-phosphate (Thd-5'-*P*).

C Abbreviations used in describing polynucleotide sequences

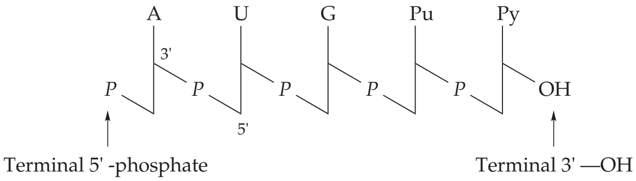
U,T,C,A,G	Uracil, thymine, cytosine, adenine, guanine
Y or Pyr	Pyrimidine (T or C)
R or Pur	Purine (A or G)
M	Amino base (A or C)
K	Keto base (G or T)
S	Strongly pairing (G or C)
W	Weakly pairing (A or T)
H	Not G (any other base)
B	Not A
V	Not T or U
D	Not C
N	Any base

---

[a] From "Biochemical Nomenclature"Liébecq, C., ed.[25]

[b] Isomers of the 5'-nucleotides, in which the phosphate is attached to the oxygen on C-3', are the 3'-nucleotides. Care must be taken to avoid ambiguity. The simple abbreviations UMP, CMP, AMP, and GMP always refer to the 5'-nucleotides.

Nucleic acid structures are abbreviated in several ways. For example, the sugar rings may be portrayed by vertical lines. The abbreviations A, C, U, T, and G for the individual bases or Pu (purine) and Py (pyrimidine) are placed at the upper ends of the lines, and slanted lines with *P* in the centers represent the 3'–5' phosphodiester linkages in a polynucleotide.



The same structure can be further abbreviated

5' end pApUpGpRpY 3' end
or 5' end A U G R Y 3' end

Here purine is abbreviated R and pyrimidine Y. By convention the 5' end of a polynucleotide is ordinarily placed to the left in these formulas. Lengths of nucleic acid chains are usually given as a number of bases or kilobases (kb). For double-stranded DNA (dsDNA) the length is given as base pairs (bp), kilobase pairs (kbp), or megabase pairs (Mbp). However, in most places, including this book, the abbreviations kb and Mb will be used for a length of DNA whether single or double stranded.

For double-stranded DNA, one strand, usually the **coding strand,** from which the amino acid sequence can be read using the code in Tables 5-4 or 5-5 (Section C), has the 5' end at the left while the complementary strand has the 3' end at the left, e.g.,

$$5'\ \underbrace{GCC}_{\text{Ala}}\,\underbrace{TAT}_{\text{Tyr}}\,\underbrace{GGA}_{\text{Gly}}\,\underbrace{CCT}_{\text{Pro}}\,\underbrace{TTC}_{\text{Phe}}\,\underbrace{TGC}_{\text{Cys}}\,\underbrace{CAG}_{\text{Gln}}\,\underbrace{TGA}_{\text{Termination}}\ 3'$$

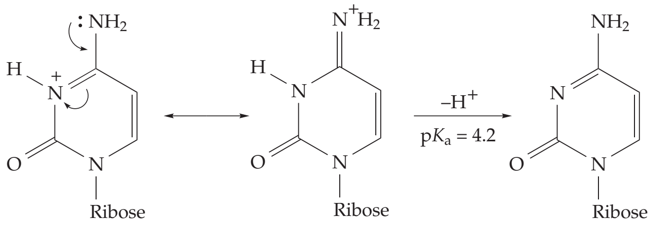
3' C G G A T A C C T G G A A A G A C G G T C A C T 5'

2. Acid–Base Chemistry and Tautomerism

The ionized phosphate groups of the polymer "backbone" give nucleic acid molecules a high negative charge. For this reason DNA in cells is usually associated with basic proteins such as the **histones** or **protamines** (in spermatozoa), with polycations of amines such as **spermidine** ($H_3^+NCH_2CH_2CH_2CH_2NH_2^+CH_2CH_2CH_2NH_3^+$), or with alkaline earth cations such as Mg^{2+}. If the pH of a solution containing double-helical DNA is either lowered to ~3 or raised to ~12, the two strands unravel and can be separated.

Over the entire range of pH the alternating sugar-phosphate backbone of the polymeric chains remains negatively charged. However, depending on the pH, the bases can be protonated or deprotonated with a resultant breaking of the hydrogen bonds that hold the pairs of bases together.[26,27]

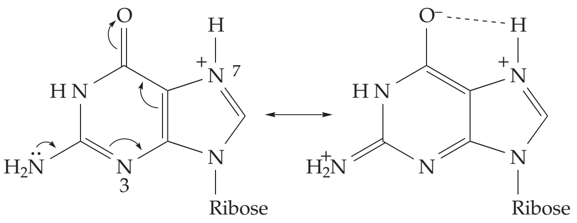
Pyrimidines and purines, which contain the $-NH_2$ group, are weakly basic. The cationic protonated conjugate acid forms of cytidine, adenosine, and guanosine have pK_a values of 4.2, 3.5, and 2.7, respectively. Similar values are observed for the 5'-nucleotides. In these compounds it is not the $-NH_2$ group that binds the proton but an adjacent nitrogen atom in the ring (Eq. 5-3).



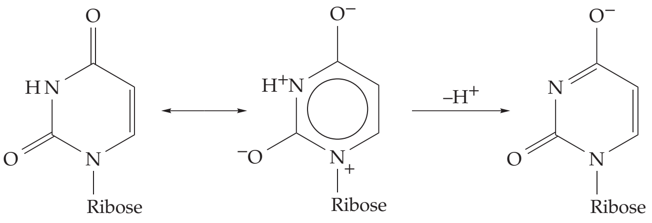
(5-3)

We can understand this if we recognize that the bases have substantial aromatic character.[26,28] In aniline (aminobenzene) electrons are withdrawn from the amino group into the aromatic ring with a strong decrease in basicity of this $-NH_2$ group (the pK_a is 4.6). Similarly, electrons are withdrawn from the NH_2 groups of cytosine, adenosine, and guanine into the pyrimidine and purine rings as is indicated by the small curved arrows on the left-hand structure of Eq. 5-3. The effect is even stronger than in aniline, largely because of the presence of the nitrogen atoms in the rings. In cytosine it is primarily N-3 that serves as the electron acceptor. As a consequence this nitrogen becomes more basic than the $-NH_2$ group and is the major site of protonation. However, as is indicated in Eq. 5-3, the positive charge on the cation is shared by resonance with the exocyclic amino group.

Adenosine is similar to cytosine in its acid–base chemistry; N-1, adjacent to the $-NH_2$ group, is the principal site of protonation. A tautomer of the cation protonated at N-3 is formed in smaller amounts. Guanosine is electronically more complex, being protonated mainly at N-7 and to a lesser extent at N-3[29]. This can be understood in terms of electronic interaction with the adjacent oxygen as indicated in the resonance structure to the right in the following diagram:



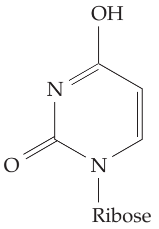
Under basic conditions the proton on N-3 of uridine or thymine or on N-1 of guanosine can dissociate with a pK_a of ~9.2. These "bases" are actually weak acids!



(5-4)

Look at the "Kekule" resonance structure shown in the center of Eq. 5-4. The two negative charges flanking the dissociable proton are sufficient to hold it firmly to the ring at low pH. The proton is half dissociated only when the pH is raised to 9.2. At still higher values of pH (pK_a ~12.4) a proton dissociates from a ribose hydroxyl group.

The tautomerism of pyrimidines was discussed in Chapter 2, Section A,6. The tautomeric forms shown in Fig. 5-1 predominate. However, it is possible that minor tautomers such as the following are sometimes preferentially bound into active sites of enzymes where the dielectric constant may be low and where the geometrical arrangement of amino acid functional groups may favor protonation on oxygen rather than nitrogen.[30,31]



A minor tautomer of uridine

3. Absorption of Ultraviolet Light

Nucleic acids strongly absorb ultraviolet light of wavelengths below about 300 nm, with an absorption maximum at ~260 nm and a stronger one below 200 nm. This property is biologically important because of the resultant induction of mutations, a natural result of exposure to sunlight. In the laboratory the same property is useful in identification and quantitative analysis of nucleotides. The ultraviolet spectra of four nucleosides are shown in Fig. 5-5. Notice the changes that accompany protonation or deprotonation of the ring. The absorption bands, which are related to those of benzene[26,32] (see Chapter 23), provide another indication of the partial aromatic character of the bases. Chapter 23 also provides information on photochemical reactions of the pyrimidines.

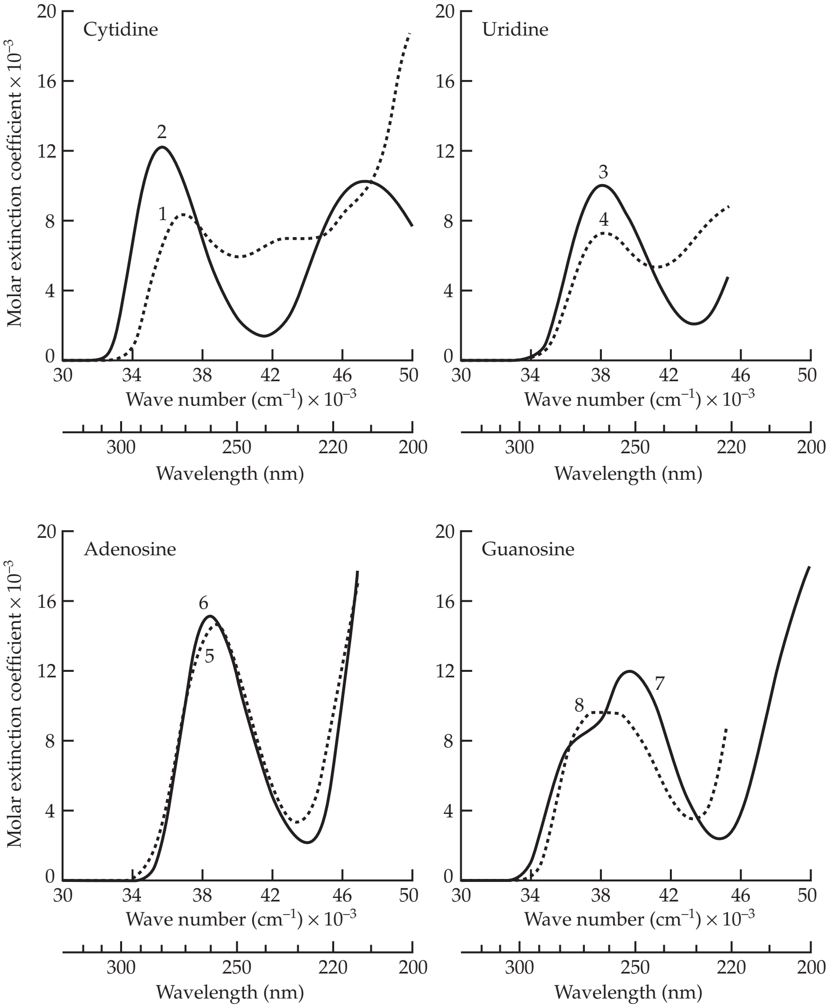


Figure 5-5 Near ultraviolet absorption spectra of cytidine, uridine, adenosine, and guanosine. 1. Monoprotonated form of cytidine (for which pK_a = 4.2). 2. Neutral form (pH ~ 7) of cytidine. 3. Neutral form of uridine (for which pK_a = 9.2). 4. Monoanionic form of uridine. 5. Monoprotonated form of adenosine (pK_a = 3.5). 6. Neutral form of adenosine. 7. Neutral form of guanosine (pK_a = 9.2). 8. Monoanion of guanosine.

BOX 5-A THE ALKALI METAL IONS

The many negative charges along a nucleic acid backbone interact with all of the cations in a cell. This box discusses some of these ions with emphasis on the group IA metal ions. Although sodium and potassium occur in similar amounts in the crust of the earth, living cells all accumulate potassium ions almost to the exclusion of sodium.[a-c] Sodium ions may be required only by certain marine organisms and by multicellular animals that regulate their internal body fluids. Most nonmarine plants have no demonstrable need for sodium.

The tendency to accumulate K^+ is even more remarkable since seawater is ~0.46 M in Na^+ and only 0.01 M in K^+. Other alkali metals occur in even smaller amounts, e.g., 0.026 mM Li^+, 0.001 mM Rb^+, and a trace of Cs^+. Soil water is ~0.1 mM in K^+ and 0.65 mM in Na^+. Again, strong discrimination in favor of potassium is observed in uptake by plants.

Intracellular concentrations of K^+ range from 200 mM in *E. coli* and 150 mM in mammalian muscle to ~30 mM in freshwater invertebrates such as clams, hydra, and some protozoa. While K^+ cannot be replaced by Na^+, a partial replacement by Rb^+ and to a lesser extent by Cs^+ is usually possible. In many microorganisms rubidium can almost completely replace potassium, and even a rat can survive for a *short* time with almost complete substitution of K^+ by Rb^+. Protons replace most K^+ in brown algae.[d] The human nutritional requirement for potassium is high, amounting to ~2 g/day. Present populations may suffer a chronic deficiency of potassium as a result of food processing and boiling of vegetables.[e]

Sodium is also essential to higher animals, and rats die on a sodium-free diet. The sodium content of cells varies among species, but it is usually no more than 0.1–0.2 times that of K^+. A measurement of the $[Na^+]$ within heart cells gave a concentration of ~9 mM, which was increased by a factor of ~2.5 in a low Ca^{2+} insulin-containing medium.[f] In this measurement the NMR resonance of the abundant external Na^+ was shifted by use of a paramagnetic reagent (e.g., a dysprosium (III) complex), that remained outside the cell. The signal from the internal Na^+ was then seen clearly. In blood, the relationship between Na^+ and K^+ concentrations is reversed from that within cells. Human plasma is 0.15 M in Na^+ and 0.005 M in K^+. Curiously, the taste for salt in the diet appears to be largely an acquired one.[g]

It is not immediately obvious why K^+ is the preferred counterion within tissues, but a fundamental reason may lie in the differences in hydration between Na^+ and K^+ (Chapter 6). On the other hand, the relationship of these ions to the excitability of membranes (Chapter 30) may be of paramount importance, even in bacteria. The concentration differences in the two ions across membranes represent a readily available source of Gibbs energy for a variety of membrane-associated activities. Cells actively pump Na^+ out and K^+ into cells (Chapter 8).

Many intracellular enzymes require K^+ for activity.[b,c] These include those promoting phosphorylation of carboxyl groups or enolate anions and elimination reactions yielding enols as well as some enzymes dependent upon the coenzyme pyridoxal phosphate.[h–j] In all of these enzymes NH_4^+, Rb^+, or Tl^+ can usually replace K^+. This permits study of the binding site for Tl^+ by the very sensitive ^{205}Tl NMR spectroscopy.[k] The discovery that K^+ is preferentially bound in some tetraplex DNA structures (see Fig. 5-8) further emphasizes the significant difference in biological properties of the alkali metal ions. Various synthetic macrocyclic compounds are also able to selectively bind specific alkali metal ions.[l]

The following tabulation gives concentrations not only of K^+ and Na^+ but also of the other principal ionic constituents in human blood plasma and within cells of skeletal muscle.[m] Units are mmol/kg H_2O.

Ion	Blood plasma	Skeletal muscle (intracellular)
Na^+	150	14
K^+	5	150
Mg^{2+}	0.9	8
Ca^{2+}	2.5	1
Cl^-	105	16
HCO_3^-	27	10
Proteins$^-$	17*	50*
Other anions†	6	146

* Milliequivalents/kg H_2O

† Phosphates and other nonprotein anions.

[a] Kernan, R. P. (1965) *Cell K*, Butterworth, London

[b] Suelter, C. H. (1974) in *Metal Ions in Biological Systems*, Vol. 3 (Sigel, H., ed), pp. 201–251, Dekker, New York

[c] Suelter, C. H. (1970) *Science* **168**, 789–795

[d] Steinbach, H. B. (1962) *Comp. Biochem. Physiol.* **4**, 677–720

[e] Weber, C. E. (1970) *J. Theor. Biol.* **29**, 327–328

[f] Wittenberg, B. A., and Gupta, R. K. (1985) *J. Biol. Chem.* **260**, 2031–2034

[g] Kaunitz, H. (1956) *Nature (London)* **178**, 1141–1144

[h] Toney, M. D., Hohenester, E., Cowan, S. W., and Jansonius, J. N. (1993) *Science* **261**, 756–759

[i] Antson, A. A., Demidkina, T. V., Gollnick, P., Dauter, Z., Von Tersch, R. L., Long, J., Berezhnoy, S. N., Phillips, R. S., Harutyunyan, E. H., and Wilson, K. S. (1993) *Biochemistry* **32**, 4195–4206

[j] Metzler, C. M., Viswanath, R., and Metzler, D. E. (1991) *J. Biol. Chem.* **266**, 9374–9381

[k] Markham, G. D. (1986) *J. Biol. Chem.* **261**, 1507–1509

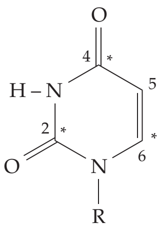
[l] Christensen, J. J., Hill, J. O., and Izatt, R. M. (1971) *Science* **174**, 459–467

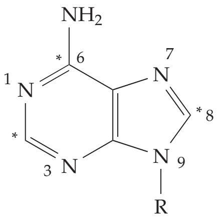
[m] Composite data from Muntwyler, E. (1968) *Water and Electrolyte Metabolism and Acid–Base Balance*, p. 14. Mosby, St. Louis, Missouri; White, A., Handler, P., and Smith, E. L. (1973) *Principles of Biochemistry*, 5th ed., p. 802. McGraw-Hill, New York; Long, C. (1961) *Biochemist's Handbook*, p. 670. Van Nostrand, Princeton, New Jersey. Reported ranges for some constituents are very wide.

4. Chemical Reactions of the Bases

The purines and pyrimidines are relatively stable compounds with considerable aromatic character. Nevertheless, they react with many different reagents and, under some relatively mild conditions, can be completely degraded to smaller molecules. The chemistry of these reactions is complex and is made more so by the fact that a reaction at one site on the ring may enhance the reactivity at other sites. The reactions of nucleic acids are largely the same as those of the individual nucleosides or nucleotides, the rates of reaction are often influenced by the position in the polynucleotide chain and by whether the nucleic acid is single or double stranded. The reactions of nucleosides and nucleotides are best understood in terms of the electronic properties of the various positions in the bases.[26,33] Most of the chemical reactions are nucleophilic addition or displacement reactions of types that are discussed in Chapters 12 and 13.

Positions 2, 4, and 6 of pyrimidine bases are deficient in electrons and are therefore able to react with nucleophilic reagents. The 6 position is especially reactive toward additions, while the 2 position is the least reactive. The corresponding electron-deficient positions in the purine bases are 2, 6, and 8. These positions, which are marked by asterisks on the following structures, have electrophilic character in all of the commonly occurring pyrimidines and purines.





All of the oxygen and nitrogen atoms in the pyrimidines, as well as the 5 position of the ring, have nucleophilic character and can therefore react with electrophilic centers of various reagents. A number of specific reactions that have been found useful to biochemists are described in Section H,3.

5. Base Pairs, Triplets, and Quartets

The purine and pyrimidine bases are the "side chains" of the nucleic acids. The polar groups that are

$$>C=O, \quad >NH, \text{ and } -NH_2$$

present in the bases can form hydrogen bonds to other nucleic acid chains, e.g., in the base pairs of the DNA double helix, and also to proteins. Figure 5-6 shows the shapes and the hydrogen bonding groups available in the bases. The number of both electron donor groups and proton donors available for hydrogen bonding is large and more than one mode of base pairing is possible.

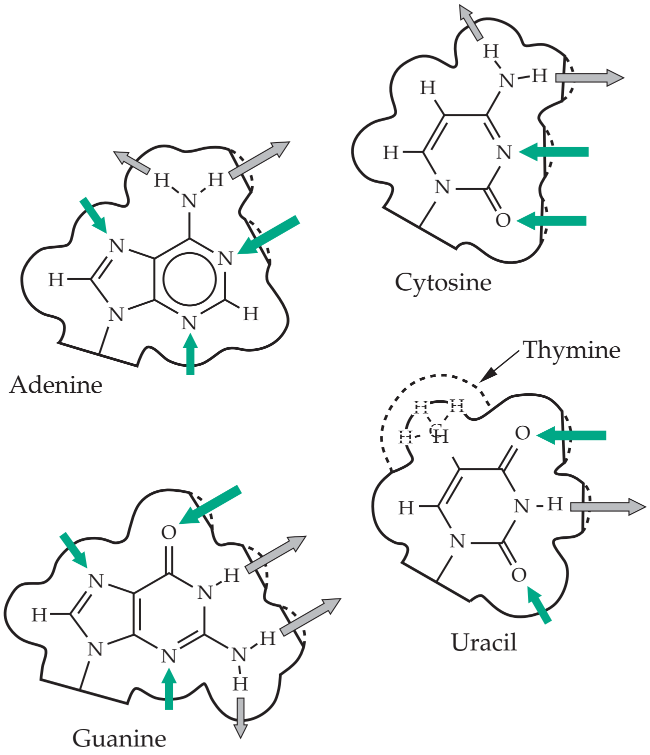
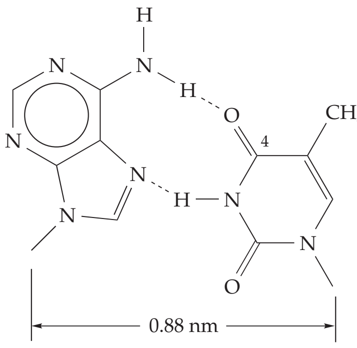


Figure 5-6 Outlines of the purine and pyrimidine bases of nucleic acids showing van der Waals contact surfaces and some of the possible directions in which hydrogen bonds may be formed. Large arrows indicate the hydrogen bonds present in the Watson–Crick base pairs. Smaller arrows indicate other hydrogen bonding possibilities. The directions of the green arrows are from a suitable hydrogen atom in the base toward an electron pair that serves as a hydrogen acceptor. This direction is *opposite* to that in the first edition of this book to reflect current usage.

The base pairs proposed by Watson and Crick are shown in Fig. 5-2 and again in Fig. 5-7. While X-ray diffraction studies indicate that it is these pairs that usually exist in DNA, other possibilities must be considered. For example, Hoogsteen proposed an alternative A-T pairing using the 6-NH_2 and N-7 of adenine.[34] Here the distance spanned by the base pair, between the C-1' sugar carbons, is 0.88 nm, less than the 1.08 nm of the Watson–Crick pairs. Duplexes of certain substituted poly (A) and poly (U) chains contain only Hoogsteen base pairs[35] and numerous X-ray structure determinations have established that Hoogsteen pairs do occur in true nucleic acids.[36] Figure 5-7 shows the structures of two Hoogsteen pairs. In each case the purine component has formed a Watson–Crick pair with a third base to give a **base triplet**. The first of these triplets may be designated T∗A•T, where ∗ represents the Hoogsteen hydrogen bonding and • represents the Watson–Crick bonding. The second triplet in Fig. 5-7, C^+∗G•C, can form only with the N-1 protonated form (low pH form) of cytosine.

Reversed Watson–Crick or Hoogsteen AU or AT pairs are formed if the 2-carbonyl rather than the 4-carbonyl of the U or T makes a hydrogen bond with the amino group of adenine.[27] Because of the resulting arrangement of the ribose rings the base pairs cannot fit into the ordinary Watson–Crick double-stranded DNA structure. A reversed Watson–Crick pair can also be formed between G and C but with only two hydrogen bonds, while a reversed Hoogsteen pair can form only if a minor tautomer of cytosine is used.

Hoogsteen pairs were first observed in nature in transfer RNA molecules (Fig. 5-31). These molecules contain mostly Watson–Crick base pairs but there are also two reversed Hoogsteen pairs. One of them, between U8 and A14, is invariant in all tRNAs studied. Hoogsteen pairing also occurs in four-stranded DNA, which has important biological functions. A **G quartet** from a DNA tetraplex held together by Hoogsteen base pairs is shown in Fig. 5-8.



Hoogsteen A∗T base pair

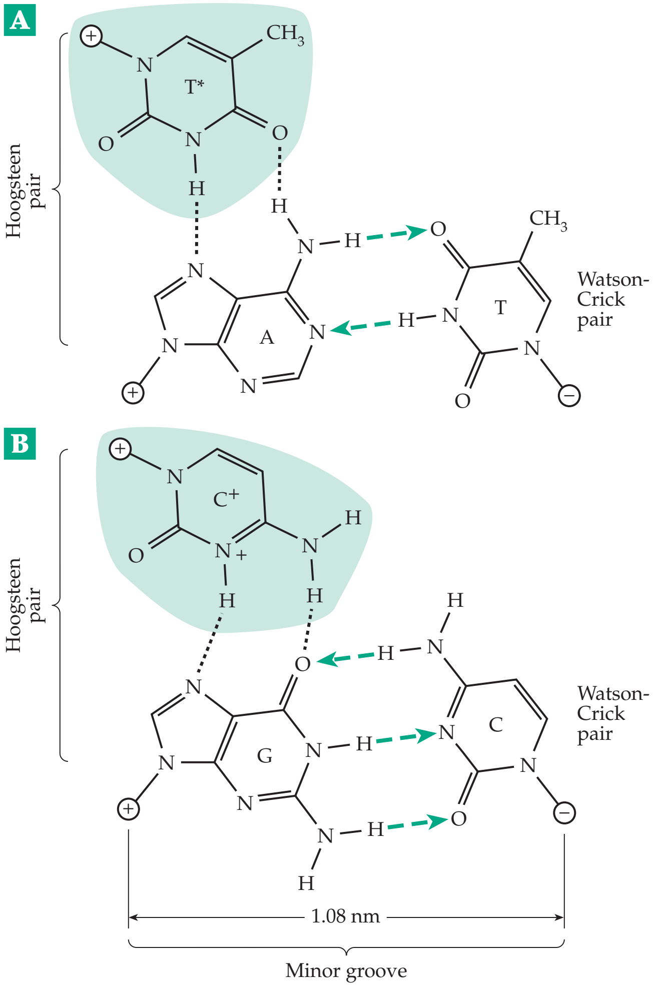


Figure 5-7 Two base triplets that form in triple-stranded DNA and involve both Watson–Crick and Hoogsteen base pairing. (A) The triplet T∗A•T, where the T (marked T∗) of the third strand is hydrogen bonded as a Hoogsteen pair (·······) to an adenine of a Watson–Crick AT pair (whose hydrogen bonds are indicated (– – –). (B) The triplet C^+∗G•C, where C^+ is cytosine in its N-1 protonated (low pH) form. The Watson–Crick strands are antiparallel, as indicated by the ⊕ and ⊖ signs. The third strand may have either orientation, but when it contains largely pyrimidines it is parallel to a purine-rich strand. An example is shown in Fig. 5-24.

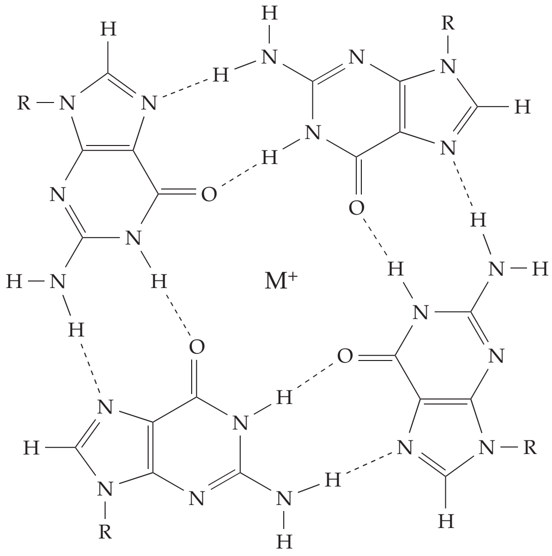
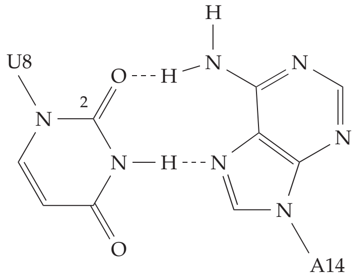
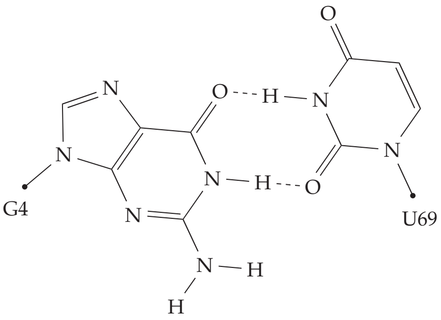


Figure 5-8 A guanine quartet held together by Hoogsteen base pairing. This structure is found in the telomeres at the ends of linear chromosomes. Four segments of DNA, each of which may be part of a single folded strand, (see p. 227) give structures in which four or more of these G quartets are stacked one above the other. Monovalent ions, usually K^+ or Na^+, are bound in the center, although not always in the plane of the bases. See Fang and Cech[38] and Gellert *et al.*[39]



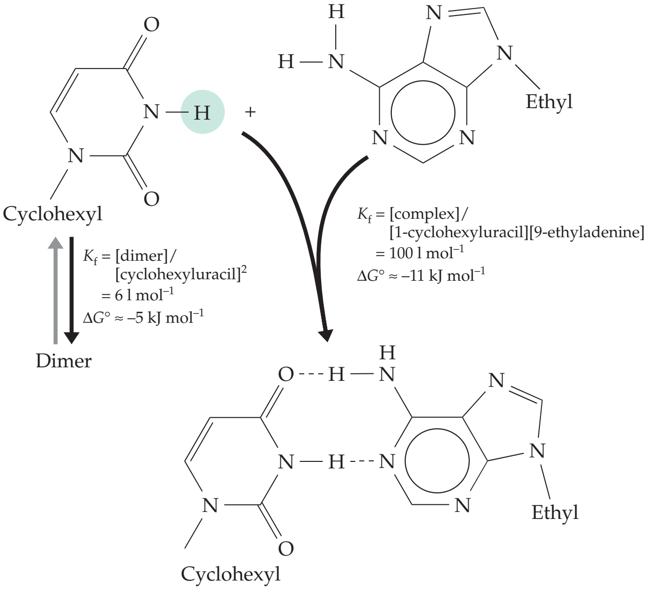
A reversed Hoogsteen base pair present in tRNA molecules

Another pairing that occurs in tRNAs allows guanine to pair with uracil, e.g., G4 with U69. This was originally proposed to account for codon–anticodon interactions betweentRNA molecules and messenger RNA (Chapter 29). It is commonly called **wobble pairing** because the uracil must wobble away from its orientation in the normal Watson–Crick pair.[27,37]



A "wobble" base pair

Strengths of base pairs. How strong are the bonds between pairs of bases in DNA? The question is hard to answer because of the strong interaction of the molecules with polar solvents through hydrogen bonding and hydrophobic effects. Some insight has come from studies of the association of bases in nonpolar solvents. Thus, 1-cyclohexyluracil forms a dimer involving either hydrogen bonding or stacking, but the association is weak with the Gibbs energy of formation $\Delta G_f° = 5$ kJ mol^{-1}. When the same compound was mixed with 9-ethyladenine, a base-paired complex formed between the two compounds with a formation constant over ten-fold greater than that for the dimer (Eq. 5-5).[40] When the circled hydrogen atom in Eq. 5-5 was replaced by $-CH_3$, which blocked the pairing, K_f fell below 1 kJ/mol ($\Delta G° > 0$). The difference in ΔG_f in the two cases was only 7 kJ mol^{-1}. Many other estimates have been made.[40a,40b] The small energies summed over the many base pairs present in the DNA molecule help provide stability to the structure.


Cyclohexyl
Ethyl
K_f = [dimer]/[cyclohexyluracil]^2 = 6 l mol^{-1}
$\Delta G° \approx -5$ kJ mol^{-1}
Dimer
K_f = [complex]/[1-cyclohexyluracil][9-ethyladenine] = 100 l mol^{-1}
$\Delta G° \approx -11$ kJ mol^{-1}
Cyclohexyl
Ethyl
Complex (base pair)

(5-5)

Stacking of bases. The purines and pyrimidines of nucleic acids, as well as many other compounds with flat ring structures and containing both polar and nonpolar regions, are sparingly soluble in either water or organic solvents. Molecules of these substances prefer neither type of solvent but adhere tightly to each other in solid crystals. Both experimental measurements and theoretical computations[41] suggest that hydrogen bonding is the predominant force in the pairing of bases in a vacuum or in nonpolar solvents. However, in water stacking becomes important.[42] In a fully extended polynucleotide chain consecutive bases are 0.7 nm apart, twice the van der Waals thickness of a pyrimidine or purine ring, but in double-helical DNA of the B type (Figs. 5-3, 5-12) the distance between consecutive base pairs is only 0.34 nm. They are touching.

One effect of stacking is a decrease in the expected intensity of light absorption. The molar extinction coefficient of a solution of double-helical DNA or RNA is always less by up to 20–30% than that predicted from the spectra of the individual nucleosides (Fig. 5-5). This **hypochromic effect** is considered further in Chapter 23.

Because both hydrogen bonding and stacking are involved, the thermodynamics of base pairing in nucleic acids is complicated.[43] The hydrophobic parts of exposed bases tend to induce an ordering of the surrounding water molecules and therefore a decrease in their entropy. However, hydrogen bonding of the polar groups of the bases to the solvent causes a decrease in water structure. This is greater than the increase in structure around the hydrophobic regions and the stacking of bases leads to a net decrease in entropy. The entropy change ΔS for addition of a base pair to the end of a double-stranded RNA helix in a hairpin loop such as that displayed in Fig. 5-9 ranges from -0.05 to -0.15 kJ/degree per base pair.[44,45]

The enthalpy change ΔH tends to be small and positive for association of alkyl groups in water and nearly zero for association of aromatic hydrocarbons

(Chapter 2). However, Δ*H* is distinctly negative for association of heterocyclic bases. This has also been attributed to a decrease in the ordering of solvent around the bases as a result of exclusion of water. Attraction or repulsion of partial charges on the polar groups comprising the purine and pyrimidine bases may also be an important factor.[43,46–48] For addition of a base pair to an RNA helix, the change in enthalpy, Δ*H*, varies from about –24 to –60 kJ/mol.[44,45]

Since $\Delta G = \Delta H - T\Delta S$, the net result is a negative value of Δ*G*, a "hydrophobic effect" that favors association of bases. Substantial efforts have been made to estimate quantitatively the Gibbs energies of formation of helical regions of RNA molecules in hairpin stem-loops such as that of Fig. 5-9.[44,45,49–51] Table 5-2 shows the observed increments in $\Delta G_f°$ of such a helix upon addition of one base pair at the end of an existing helix. Addition of an AU pair supplies only –4 to –5 kJ to $\Delta G_f°$. The exact amount depends upon whether an A or a U is at the 5' end in the existing helix. If an AU pair is added to the helix terminating in CG or GC, about –9 kJ/mol is added. Larger increases in $-\Delta G_f°$ result from addition of GC pairs, which contain three hydrogen bonds between the bases versus the two in AU pairs.[27,41,52]

UG pairs provide a very small amount of stabilization to an RNA double helix, while the presence of unpaired bases has a destabilizing effect. The most stable hairpin loops contain four or five bases. Depending upon whether the loop is "closed" by CG or AU, the helix is destabilized by 20–30 kJ/mol. "Bulge loops," which protrude from one side of a helix, have a smaller destabilizing effect. An example of the way in which Table 5-2 can be used to estimate the energies of formation of a loop in a straight-chain RNA is illustrated in Fig. 5-9. Similar analysis of base pairing in DNA can also be done.[53–55]

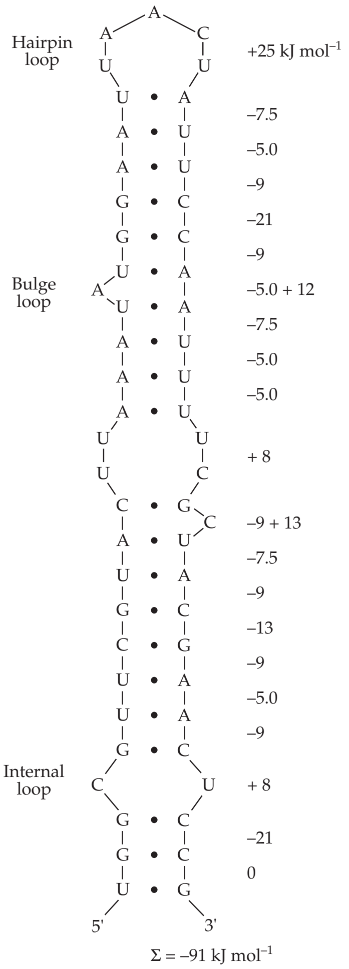


Figure 5-9 The contribution of base-paired regions and loops to the Gibbs energy of a possible secondary structure for a 55 base fragment from R17 virus. The stem-loop structure shown here is part of a larger one considered by Tinoco *et al.*[58]

TABLE 5-2
Gibbs Energies of Formation $\Delta G_f°$ at 25°C for Addition of One Base Pair to an Existing RNA Helix[a,b]

Base pair at end of existing helix	Base pair added	$\Delta G_f°$ (kcal mol^{-1} ± 10%)	$\Delta G_f°$ (kJ mol^{-1} ± 10%)
A • U	A • U	–1.2	–5.0
U[c] • A	A • U	–1.8	–7.5
C•G or G[d]•C	A • U	–2.2	–9
C • G	G • C	–3.2	–13
G • C	C•G or G•C	–5.0	–21
G • U	U • G	–0.3	–1
Hairpin loops of 4 or 5 bases			
Closed by GC		+5	+21
Closed by AU		+7	+29
Bulge loops			
1 base		+3	+12
4–7 bases		+5	+21

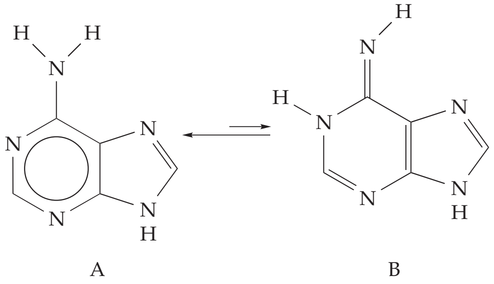
[a] Table modified from that of Tinoco *et. al.*[58]

[b] All base pairs in the table are oriented as follows: 5'—A→3' • 3'←U—5'

[c] $\Delta G_f°$ is the same for U•A added to an A•U end.

[d] $\Delta G_f°$ is the same for C•G or G•C added to an A•U end.

Tautomerism and base pairing. Tautomerism has an interesting relationship to the formation of the pairs and triplets of hydrogen bonds in DNA or RNA. Each base exists predominately as one preferred tautomer, but at any moment a very small fraction is present as less stable tautomers. Watson and Crick suggested that this fact may be responsible for the occurrence of some mutations.[56,57] Thus, tautomer B of Eq. 5-6 would not be able to pair with thymine, its proper pairing partner, but could pair with cytosine.



(5-6)

Similarly, tautomer B of uracil in Eq. 2-4 could pair with guanine instead of its proper partner adenine. If a similar event occurred to an AT pair during gene replication an incorrect copy of the gene, differing in a single "code letter," would be formed. However, because the tautomerism is affected so strongly by the environment (Chapter 2), the extent to which it may cause mispairing while replication enzymes act is uncertain.[59] Both bases in a pair could be tautomerized by synchronous transfer of protons in two parallel hydrogen bonds. However, theoretical calculations predict a high energy barrier to this process.[60] Proton transfer, which can also be induced by light, has been studied on a femtosecond scale.[61] After photochemical transfer of one proton a second is transferred within a few picoseconds.

6. Conformations of Nucleotides

The furanose ring of ribose or deoxyribose is flexible and can be interconverted smoothly among an infinite number of envelope (E) and skew or twist (T) conformations. See Chapter 4, Section A,2. However, there are limits set by steric and anomeric effects.[27,62–64] Conformations are often described as in Fig. 5-10 by stating which atom in an envelope conformation lies mostly out of the plane of the other four atoms. If this atom lies above the ring, i.e., toward the base, the ring

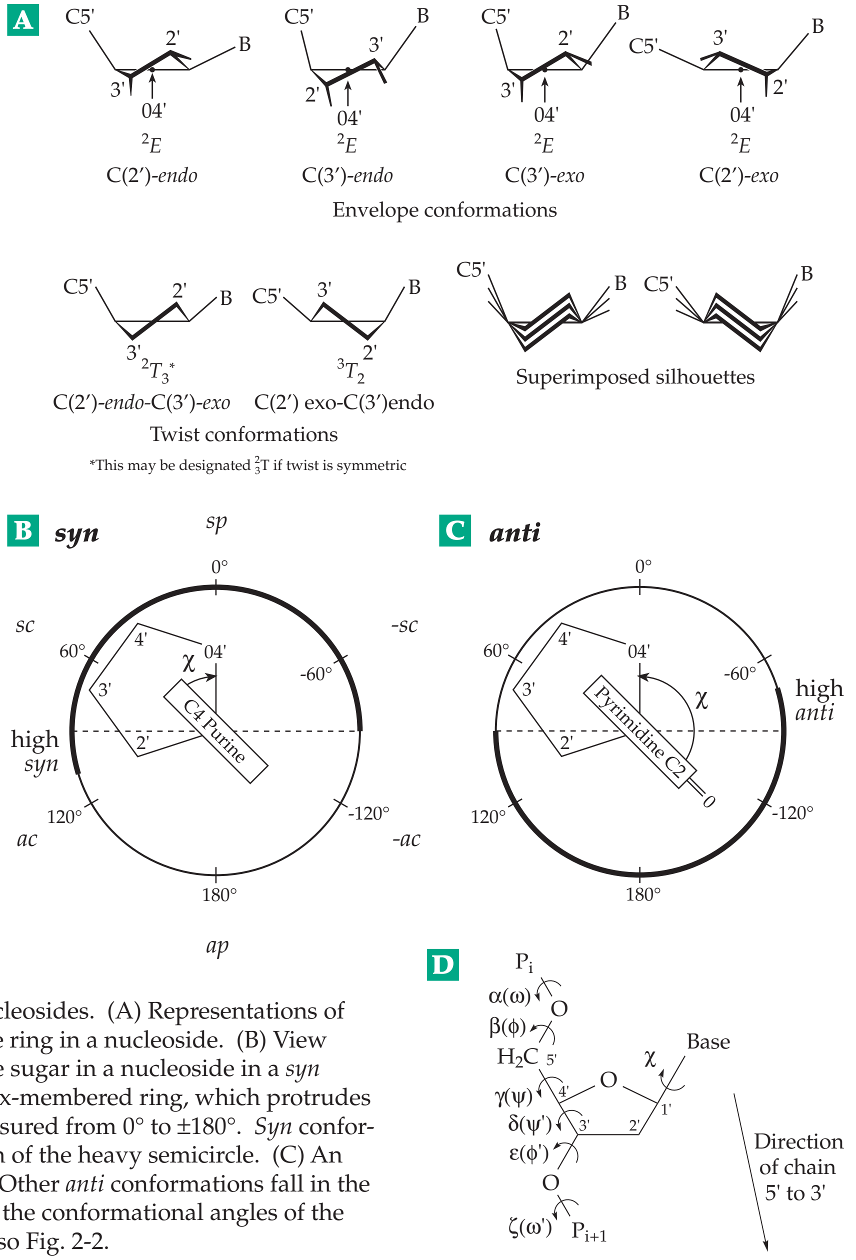
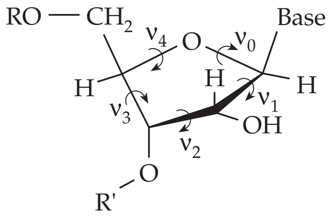

A
C5'
2'
B
3'
04'
^2E
C(2′)-*endo*
C(3′)-*endo*
C(3′)-*exo*
C(2′)-*exo*
Envelope conformations
$^2T_3^*$
$^3T_2$
C(2′)-*endo*-C(3′)-*exo*
C(2′) exo-C(3′)endo
Twist conformations
Superimposed silhouettes
*This may be designated ^2_3T if twist is symmetric
B *syn*
sp
0°
sc
-sc
60°
-60°
4'
04'
χ
3'
C4 Purine
high *syn*
120°
-120°
ac
-ac
180°
ap
C *anti*
Pyrimidine C2
O
high *anti*
D
P_i
α(ω)
O
β(ϕ)
H_2C 5'
Base
χ
γ(ψ)
4'
1'
δ(ψ')
3'
2'
ε(ϕ')
ζ(ω')
P_{i+1}
Direction of chain 5' to 3'

Figure 5-10 Conformational properties of nucleosides. (A) Representations of several conformations of a ribose or deoxyribose ring in a nucleoside. (B) View down the N–C axis joining a purine base to the sugar in a nucleoside in a *syn* conformation. The atom marked C4 is in the six-membered ring, which protrudes further over the sugar ring. The angle χ is measured from 0° to ±180°. *Syn* conformations are those for which χ falls in the region of the heavy semicircle. (C) An *anti* conformation of a pyrimidine nucleoside. Other *anti* conformations fall in the region of the heavy semicircle. (D) Labeling of the conformational angles of the main chain in a polynucleotide.[25,27,37,37a] See also Fig. 2-2.

conformation is known as ***endo***; when below the ring, it is known as ***exo***. The C(2')-*endo* (^{2}E) and C(3')-*endo* (^{3}E) conformations are most commonly approximated in nucleotides and nucleic acids.[65] The C(3')-*exo* conformation is designated E_3, the twist conformation C(2')-*endo*-C(3')-*exo* as $^2T_3$, etc. A conformation can be specified more precisely by the five torsion angles ν_0 to ν_4 (which have also been designated τ_0 to τ_4). All of the envelope and twist conformers can be interconverted readily. The interconversions can be imagined to occur in a systematic way by **pseudorotation**, a rotation of the pucker around the sugar ring. In a commonly used convention a pseudorotation phase angle *P* ranges from 0° to 360° as the pucker moves *twice* around the ribose ring to restore the original conformation. *P* is taken as 0° for the symmetric $^3T_2$ twist, −18° for E_2, +18° for ^{3}E, +54° for ^{4}E, +198° for E_3, etc. The five torsion angles ν_j, for j = 0 – 4 as defined in the foregoing diagram, are related to *P* as follows:



$$\nu_j = \nu_{max} \cdot \cos[P + j \cdot \phi] \qquad (5\text{-}7)$$

where $\phi = 720°/5 = 144°$. The maximum torsion angle, ν_{max}, is about 40°. See Saenger[27] for details.

Conformational alterations of ribose and deoxyribose rings can occur within polynucleotides and are of biochemical importance. An interesting consequence of changes in ring conformation is that the distance between the C(5') and N atoms attached to the sugar ring of a nucleoside may vary by as much as 0.05 nm (see Fig. 5-10A). The orientation of a base with respect to the sugar is specified by the angle χ. In one convention (Fig. 5-10B,C) its zero value is taken as that in which the N(1)–C(2) bond of a pyrimidine or the N(9)–C(4) bond of a purine is *cis* to the C(1')–O(4') bond, [sometimes called the C(1')–O(1') bond]. However, other definitions have been used.[27] Typical values of χ for nucleotides and nucleosides vary between −75° and −165°. In these ***anti*** conformations the CO and NH groups in the 2 and 3 positions of the pyrimidine ring (or in positions 1, 2, and 6 of the purine ring) are *away* from the sugar ring, while in the ***syn*** conformations they lie over the ring. *Anti* conformations are more often present than *syn* in nucleic acids. A different view of *syn* and *anti* pairs of nucleotides is shown in Fig. 5-11.

An additional five torsion angles are needed to specify the backbone conformation of a polynucleotide. According to the convention adopted by the IUB the six angles are desig-

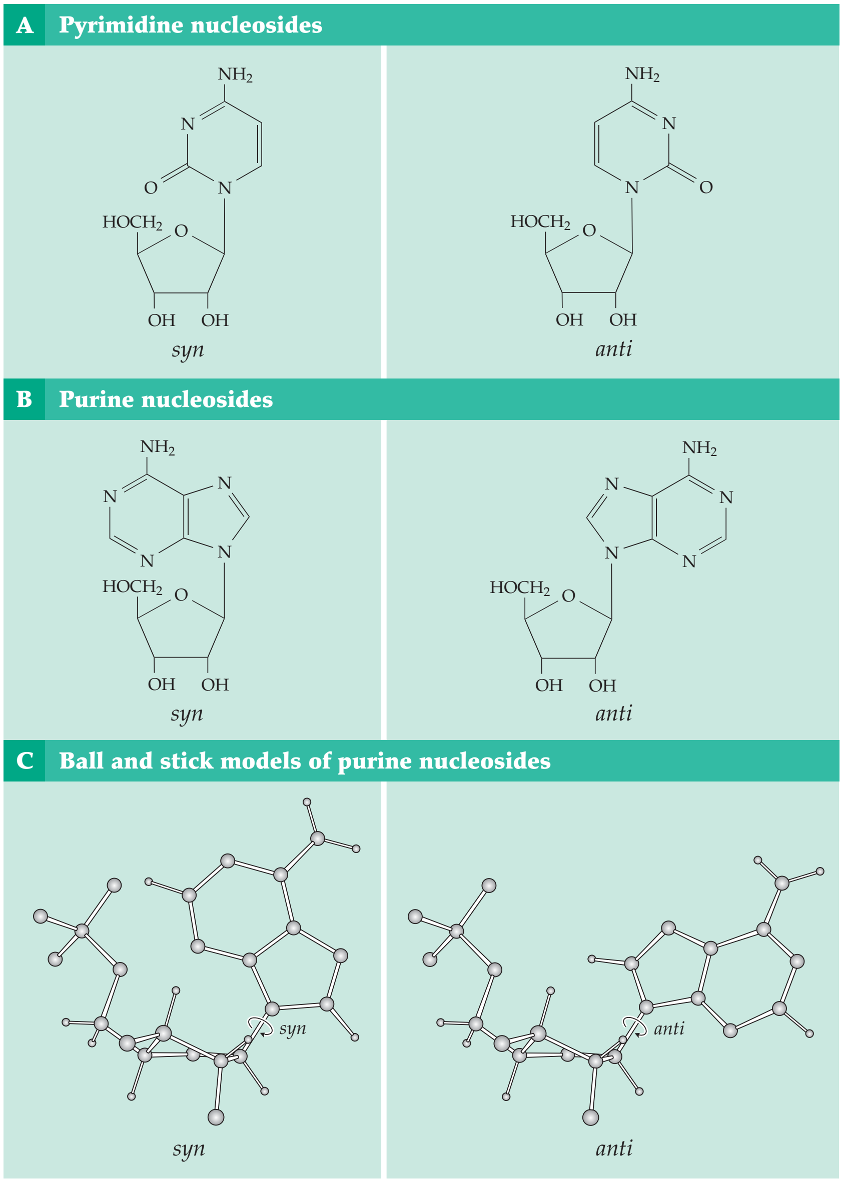


Figure 5-11 Views of a nucleoside in *syn* and *anti* conformations.[66] (C) Courtesy of Dr. Muttaiya Sundaralingam.

nated α, β, γ, δ, ε, and ζ, as is indicated in Fig. 5-10D.[25,27] In an older but much used convention,[66] starting at any phosphorus atom the angles ω, ϕ, and ψ specify the next three torsion angles as one moves toward the 3' end of the chain, while ω′, ϕ′, and ψ′ specify the angles lying toward the 5' end (these are shown in parentheses in Fig. 5-10D). Notice that $\delta(\psi') = \nu_3 + 120°$.

B. Double Helices

1. The B Form of DNA

This form of DNA, whose structure is depicted in Figs. 5-3 and 5-12, is stable at high humidity and is thought to approximate that of most DNA in cells.[27,37,67] If we look directly at the axis of the double helix and perpendicular to one of the base pairs, and ignore the fact that the base pair is asymmetric, we see that the nucleotide unit in one chain is related to the nucleotide unit lying across from it in the opposite plane by a two-fold axis of rotation (**dyad axis**). This symmetry element, which arises from the antiparallel arrangement of the chains, makes the DNA molecules from the outside look nearly identical whether viewed from one end or the other – and whether viewed as a model by the human eye or through contact with an enzyme which might act on the molecule. Actually, *the two chains are not identical*, and the genetic information can be read off from the functional groups exposed in two **grooves** in the surface of the helix (Figs. 5-3 and 5-7). The broader groove in the B form, which is referred to as the **major groove,** is about 0.85 nm deep and 1.1–1.2 nm wide when allowance is made for the van der Waals radii of the atoms. In some other forms of DNA the major groove is narrow, but it can always be identified by the larger of the arcs that can be drawn between the two N– C bonds of the nucleosidic linkages in a nucleotide pair. The **minor groove** or narrow groove, which is defined by the smaller arc between the two N–C bonds, is ~ 0.75 nm deep and 0.6 nm wide in B-DNA.

The diameter of the double helix of B-DNA, measured between phosphorus atoms, is just 2.0 nm. The rise per turn, the **pitch**, is 3.4 nm. There are about ten base pairs per turn (9.7 and 10.6 in two different crystal forms).[68,69] Thus, the rise per base pair is 0.34 nm, just the van der Waals thickness of an aromatic ring (Table 2-1). It is clear that the bases are stacked in the center of the helix. A 1000-bp (1-kb) gene would be a segment of DNA rod about 340 nm long, about 1/40 the length of the molecule in the electron micrograph of Fig. 5-13.

As is appropriate for the cell's master blueprint, DNA in the double helix is stable. Factors contributing to this stability are (1) the pairs and triplets of hydrogen bonds between the bases; (2) the van der Waals attraction between the flat bases which stack together; (3) the fact that on the outside of the molecule are many oxygen atoms, some negatively charged, which are able to form strong hydrogen bonds with water, with small ions, or with proteins that surround the DNA; and (4) the ability to form superhelices (see Section C,3). Nevertheless, the long DNA chains present in our chromosomes are frequently broken and an elaborate system of repair enzymes is needed to preserve the reliability of this master code for the cell.[73,74]

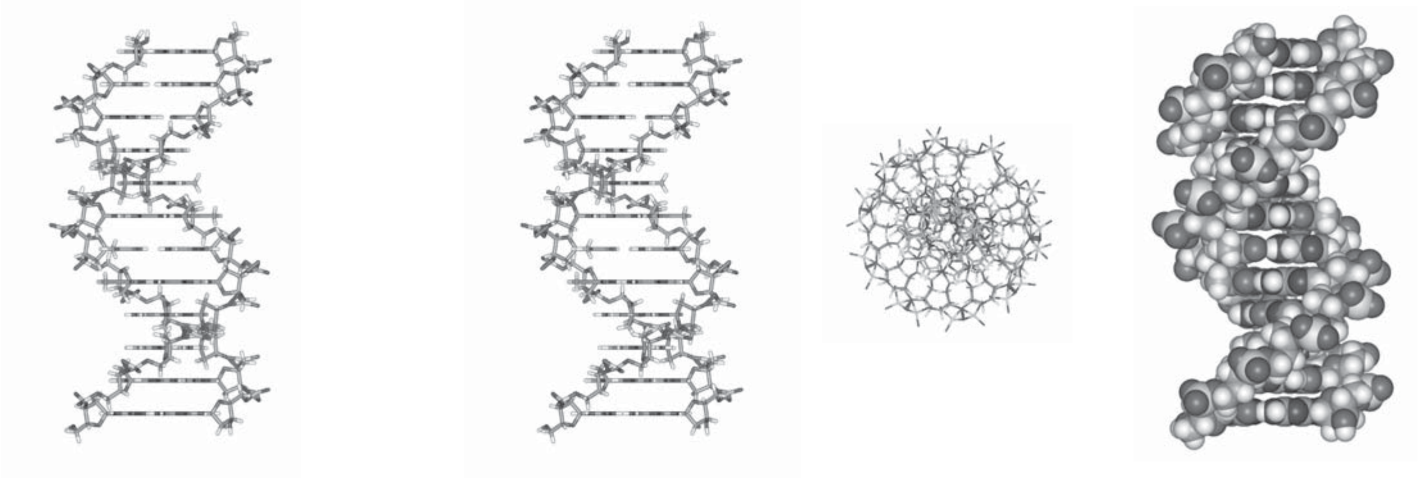
Today it is generally accepted that most DNA exists in nature as a double helix resembling the B form. However, doubts were expressed as recently as 1980. The reason for the uncertainty lay in the fact that X-ray data from the stretched "paracrystalline" DNA fibers used in earlier studies are much less precise than those from true crystals.[10] The X-ray data for B-DNA fibers could also be interpreted in terms of an alternative "side-by-side" structure which would permit easier separation of strands during replication.[75] However, numerous high-resolution X-ray structure determinations on single crystals of synthetic DNA fragments have confirmed the double-helical structure and the presence of Watson–Crick base pairs.[68,69,76–84] Similarly, fragments related to the double helices of RNA have been crystallized and the structures determined to atomic resolution.[85–87] The right-handed helical structure of DNA in solution has also been confirmed by independent methods based on electron microscopy,[88] scanning probe microscopy,[89,90] fluorescence resonance energy transfer,[91] and computation.[92]

Dickerson and associates discovered[78,79] that B-DNA has a "spine" of water in the minor groove of regions rich in A-T base pairs. Two water molecules per base pair are hydrogen bonded to form a long chain. Half of the water molecules in the chain also form hydrogen bonds to oxygen and nitrogen atoms that are exposed in the minor groove while the others hydrogen-bond to a second chain of water molecules, forming a ribbon of hydration (Fig. 5-14A).[81,93-95] Additional water binds to polar groups in the major groove. The hydration pattern is largely *local*, i.e., each base has characteristic hydration sites (Fig. 5-14B).[94] Some of the hydration sites may be occupied partially by the nomovalent ions, Na^+ or K^+, depending upon the medium. Bound divalent metal ions are also sometimes seen in X-ray structures.[81,83a] However, most cations are thought to remain mobile.

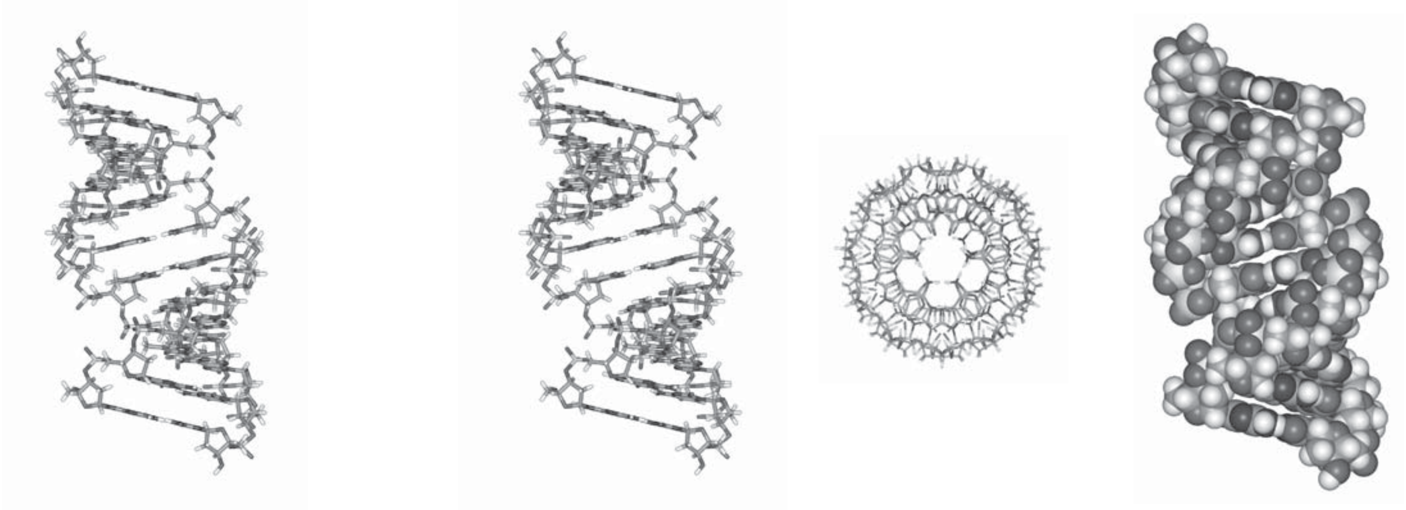
2. Other Double-Helical Forms of DNA

The B form of fibrous DNA is stable under conditions of high (~ 93%) humidity but at 75% humidity it is converted into **A-DNA** in which the base pairs are inclined to the helix axis by about 13° and in which the ribose rings are primarily C3'-*endo* rather than C2'-*endo*.

B-form



A-form



Z-form

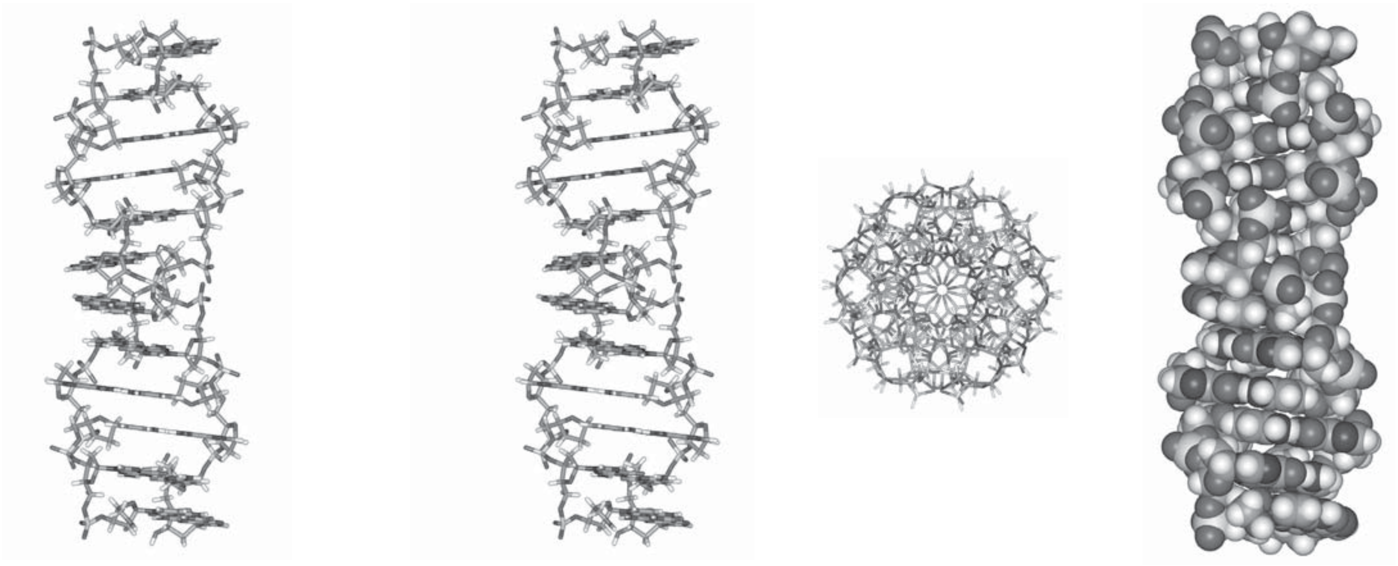


Figure 5-12 (Left) Stereoscopic skeleton models of the B, A, and Z forms of double helical DNA. See Schlick.[70] (Center) End views (Right) Space-filling models of the same three DNA forms: B, A, and Z. Courtesy of Tamar Schlick.

There are 10.9 base pairs per turn[93,96–98] (Fig 5-12; Table 5-3). In the A form the major groove is very deep (~ 1.35 nm) and narrower (~ 0.27 nm) than in the B form and extensively hydrated.[99] The minor groove is wider (~ 1.1 nm) and shallower (~ 0.28 nm) than in the B form. The crystal structure for a form intermediate between A and B has also been reported.[99a] Paracrystalline forms of DNA known as B′, C, and D have also been observed[100,101] and others have been proposed.[102,103]

The most interesting additional form, **Z-DNA**, was discovered by X-ray studies of the alternating oligodeoxyribonucleotides d(CpGp-CpGpCpG) and d(CpGpCpG).[104] The helix of Z-DNA is *left-handed*.[71,72,105–112] The repeating unit consists of two Watson–Crick base pairs, the backbone following a zigzag pattern (Fig. 5-12). There are 12 base pairs per turn. The cytosine groups have the usual *anti* conformation but the guanosine groups are *syn*. Some of the ribose rings have the C2'-*endo* conformation characteristic of B-DNA but some are C3'-*endo*. The alternating CpG sequence is not essential for Z-DNA formation. However, alternating *syn* and *anti* conformations are important and purines assume the *syn* conformation more readily than do pyrimidines. The major groove of Z-DNA is shallow (Fig. 5-12), allowing it to accomodate bulky substituents at C8 of purines or C5 of pyrimidines. Such substituents favor Z-DNA.

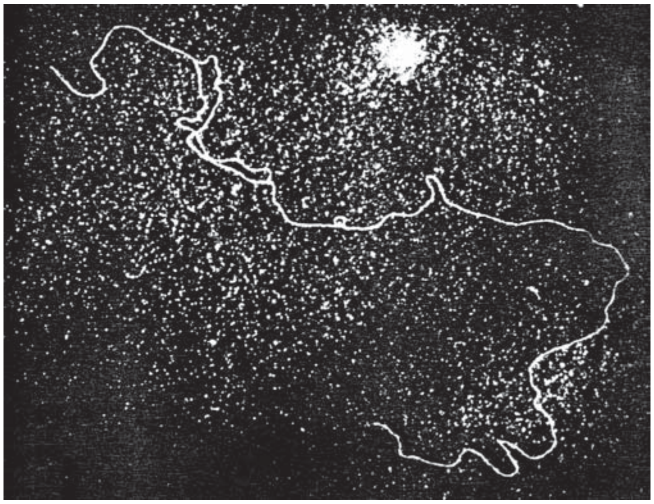
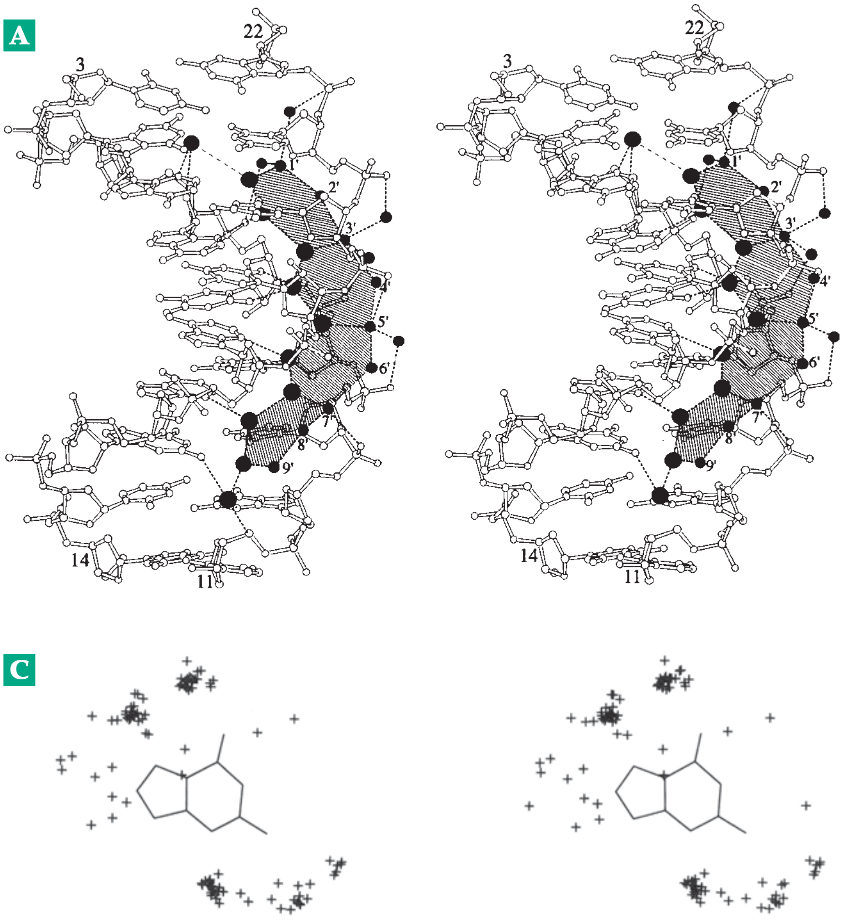
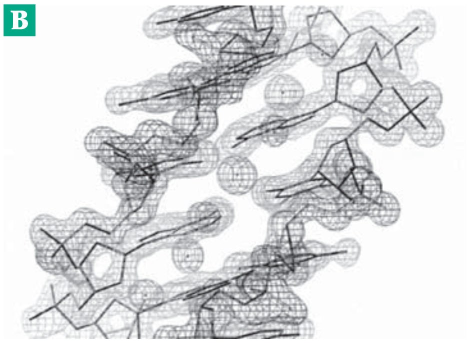


Figure 5-13 Electron micrograph of a DNA molecule (from a bacterial virus bacteriophage T7) undergoing replication. The viral DNA is a long (~ 14 μm) duplex rod containing about 40,000 base pairs. In this view of a replicating molecule an internal "eye" in which DNA has been duplicated is present. The DNA synthesis was initiated at a special site (origin) about 17% of the total length from one end of the duplex. The DNA was stained with uranyl acetate and viewed by dark field electron microscopy. Micrograph courtesy J. Wolfson and D. Dressler.

Figure 5-14 (A) Stereoscopic drawing showing two layers of water molecules that form a "spine" or "ribbon" of hydration in the minor groove of B-DNA. The inner layer is shown as larger filled circles; water molecules of the outer layer are depicted with smaller dots and are numbered. Hydrogen bonds are shown as dashed lines. (B) Electron density map. (A) and (B) from Tereshko *et al.*[95] (C) Stereoscopic representation of the superimposed electron densities of 101 water molecules observed to hydrate 14 guanine rings found in 14 B-DNA molecules for which high-resolution X-ray structures were available. Positions of 101 water molecules within 0.34 nm from any atom of the 42 guanines are plotted. From Schneider and Berman.[94]





The B form is the most hydrated and most stable form of DNA under conditions of high humidity but even in solution it can be converted to A-DNA and Z-DNA by a high concentration of NaCl.[113] This is presumably because the salt dehydrates the DNA. Saenger and coworkers pointed out that the oxygen atoms of successive phosphate groups in the polynucleotide backbone of B-DNA are at least 0.66 nm apart, too far apart to be bridged by a water molecule.[114] However, the phosphates are individually hydrated. On the other hand, in A-DNA and Z-DNA the oxygen atoms on successive phosphates are as close as 0.53 and 0.44 nm, respectively. This allows one H_2O molecule to bridge between two phosphates, stabilizing these forms in environments of low humidity. This may be one factor that affects the B to A and Z transitions.[115–117] In the narrow Z-DNA helix repulsion between negative charges on the phospho groups of opposite strands is strong. By shielding these charges high salt concentrations also help to stabilize Z-DNA.

Since the Z form of DNA is favored in regions rich in G-C pairs,[118,119] it is reasonable to expect that it may occur in nature. Antibodies have been prepared which bind specifically to regions of DNA in the Z form and have been used to identify many such regions.[111,120] Genetic studies in *E. coli* have also provided strong evidence that left-handed DNA sequences are formed *in vivo* in that bacterium.[121,122] Z-DNA-forming sequences have also been found in the *Halobacterium* genome.[119] Segments of Z-DNA may occur in control regions called **enhancers** (Chapter 28)[121–124] and Z-DNA may also form behind RNA polymerase molecules that are moving along a gene while synthesizing mRNA. The associated negative supercoiling of the DNA could cause it to assume the Z form. This region of Z-DNA can, in turn, be a site for interaction with specific proteins.

In the Watson–Crick structure the two strands are *antiparallel*, an essential for replication. However, stable segments of double-stranded DNA with *parallel* strands can also be formed and may occur in specialized regions of the genome.[125–128]

TABLE 5-3
Helix Parameters for Three Types of DNA[a]

	Form of DNA or RNA (A-form)		
Parameter	B	A	Z
Helical twist, degrees mean	28–42 36 ± 4	16–44 33 ± 6	GC −51 ± 2 CG −8.5 ± 1
Base pairs per turn	10.0 (9.7–10.6)	11-12	12
Helix rise per base pair, nm	0.34 ± .04	0.29 ± .04	GC 0.35 ± .02 CG 0.41 ± .02
Base inclination, degrees	−2.0 ± 5	13 ± 2	8.8 ± .7
Propeller twist, degrees	12 ± 5	15 ± 6	4.4 ± 3
Base roll degrees	−1.0 ± 5	6 ± 5	3.4 ± 2
Predominant conformation of deoxyribose	C2′-*endo*	C3′-*endo*	C C2′-*endo* G C3′-*endo*
Depth of grooves (nm)			
Major	0.85	1.35	very shallow
Minor	~0.75	~0.28	very deep
Width of grooves (nm)			
Major	1.1–1.2	0.27	broad
Minor	0.6	~1.1	narrow

[a] See R. E. Dickerson[78,78a]; based on single-crystal X-ray analysis.

3. The Conformational Flexibility of Double Helices

Local variations in the sequence of nucleotides affect the conformation of a DNA molecule and it is clear that the helix is not uniformly coiled throughout the entire length.[80,103,124,129–134] While most helix segments probably have a right-handed twist others may be left-handed. Most DNA is probably in the B form but there are segments in the A form. These may arise from formation of hybrid duplexes with RNA, which assume the A conformation and are also favored by certain base sequences. Rules for predicting the DNA conformation from the nucleotide sequence have been proposed.[116,135] In the simplest case[135] we consider each pair of adjacent nucleotides in the double helix. There are 16 possible pairs in one chain. These can be designated as in the following examples: (AA,TT), (CG,CG), and (AG,CT). The first two letters within the parentheses represent the sequence (from 5' to 3') in one chain while the second pair of letters represent the sequence (again from 5' to 3') in the complementary chain. The rules state that (AA,TT) or (TT,AA) repeated in a sequence will stabilize the B form of DNA. Repetition of (CC,GG) or (GG,CC) will favor conversion to the A form. Repetitions of (CG,CG) favor the Z form, especially if an alternating sequence of purines and pyrimidines is present throughout the (G + C)-rich region.

Even within the regular B structure the

torsion angles χ and δ (Fig. 5-10) are variable. Their changes are highly correlated.[136,136a] As χ ranges from $-140°$ to $-90°$, ζ varies between 80° and 160°. Other pairs of torsion angles are also correlated. Besides allowing for changes between B, A, and Z conformations, this flexibility of the DNA helix together with cooperativity with adjacent base pairs may allow transmission of conformational effects for some distances along a DNA helix.[137,138] Supercoiling, discussed in Section C,3, also affects the helical conformation.

DNA can be stretched into yet another form or forms. Application of a force of 65 – 70 piconewtons (pN) stretches B-DNA by 70%. This "overstretched DNA" may also be important biologically.[130,139–141a]

Rotational and translational movements of bases. In considering the conformational flexibility of a polynucleotide it is useful to define the parameters associated with movement of base pairs or individual bases. The possible movements are indicated in Fig. 5-15 and in Table 5-3. A **long axis** is established for each base pair in a structure for which X-ray data are available. Making use of these axes, we can move along the helix and measure the **angle of twist** from one base pair to the next.[142–145] While the mean value of the twist for B-DNA is 36°, values of 28–42° have been observed in structures of oligonucleotides determined by X-ray analysis. The **base tilt**, which is nearly zero for B-DNA, is less variable. Within each base pair there is also a **propeller twist** around the long axis.[146] It averages about 12° for B-DNA. In addition, the whole base pair may **roll**, i.e., be rotated around its long axis by several degrees. These motions relieve steric interferences in the center of the helix, for example, those that would arise between purines in adjacent base pairs if the same helix parameters were imposed on each base pair.[142,147] A more detailed analysis reveals the large range of motions, mostly small, which are described in Fig. 5-15.[145,148] The mathematics needed to deal with computer-based modeling of polynucleotide structures has been developed.[80,149,150] Its application showed that there is a strong correlation between twist and roll. Typical average values of base tilt, propeller twist, and roll are given in Table 5-3 for B-, A-, and Z-DNA.

Bends and bulges. If we overlook the ridges and grooves on their surfaces the DNA structures shown in Figs. 5-3 and 5-12 are straight rods. However, real DNA rods are crooked and may contain distinct bends. The

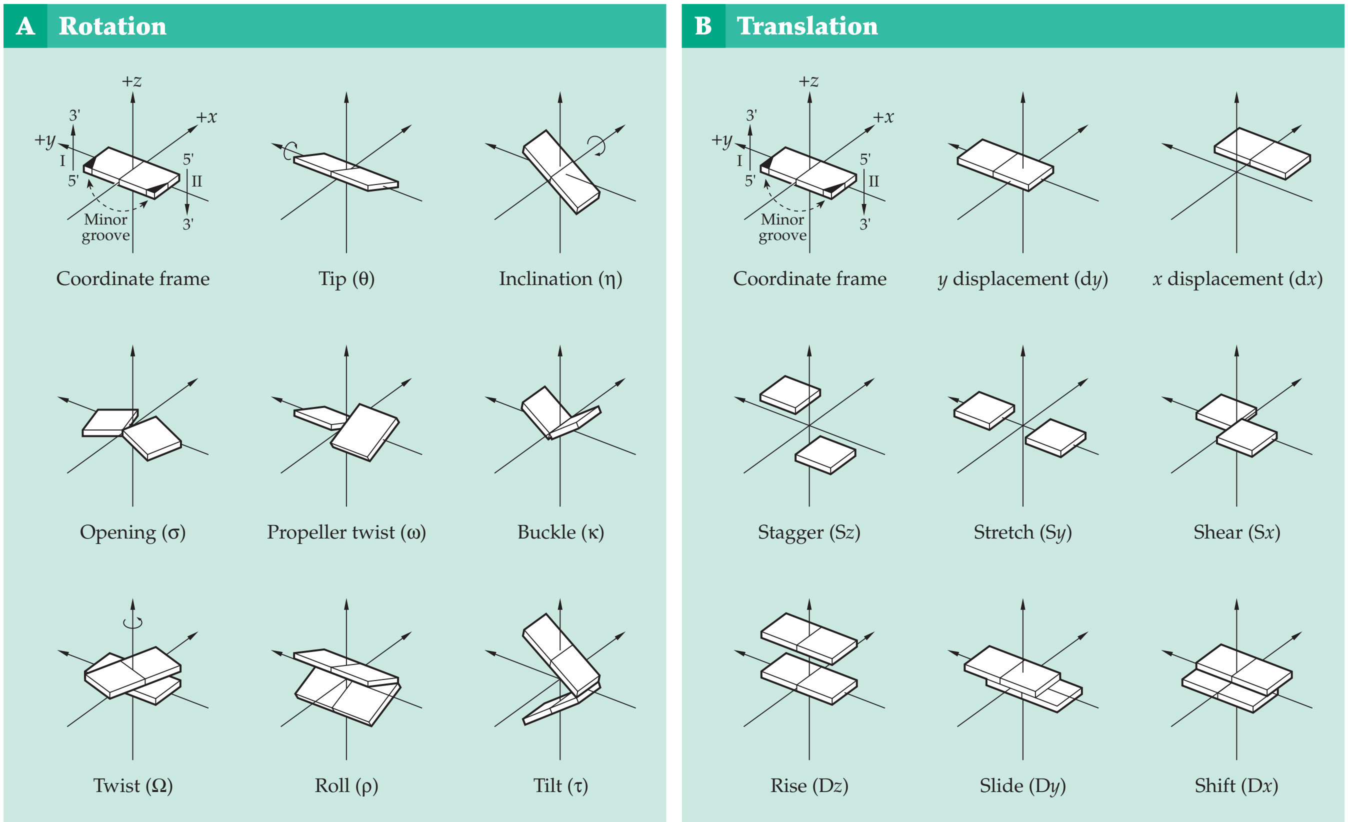


Figure 5-15 (A) Drawing illustrating various rotational movements of bases in polynucleotides. *Upper two rows*: rotations of two bases of a pair. *Lower row*: rotations involving two successive base pairs. (B) Translational movements. *Upper two rows*: involving two bases of a pair. *Lower row*: two successive base pairs. From Diekmann.[145]

existence of **bent DNA** in nature was first discovered from study of fragments enzymatically cut from kinetoplast DNA of tropical parasites.[151,152] The bent fragments moved anomolously slowly during electrophoresis in polyacrylamide gels of small pore size. Evidently, the bent shape impedes movement through the pores. Since the initial discovery, regions of bent DNA have been found in origins of replication[152–154] and other specialized sites. Bending of DNA because of local sequence variations may also be important to the positioning of nucleosomes (Fig. 27-3) on DNA.[155]

Several causes for DNA bending have been identified.[124,138,142,156–164] For example, a sharp 26° bend is expected at the junction between B- and A-DNA segments.[158] The presence of a thymine photodimer (Eq. 23-26) may cause a 30° bend.[165] Some naturally occurring bent DNAs contain repeated (A+T)-rich sequences. This suggested that the sequence (AA,TT), containing adjacent thymine methyl groups, can be thought of as a wedge, consisting of both roll and tilt components. Such a wedge repeated at intervals averaging 10.5 bp along the helix might cause the helix to bend.[129,144] However, computer modeling experiments suggest that thymine methyl groups do not distort the DNA helix. The bending may be a result of the greater tendency for AT pairs to roll and to enhance the van der Waals and electrostatic attractions between the 2'-O of cytosine and the 2'-NH_2 of guanine in the CG base pairs of bent DNA.[157,166] Bending can occur in both (A+T)- and (G+C)-rich regions and is often induced by binding to proteins or protein complexes that act in replication, transcription, and recombination.[167]

Various errors are made during replication or recombination of DNA. Incorporation of an incorrect nucleotide will cause a mismatched base pair in which proper hydrogen bonds cannot be formed. Most of the mispaired bases that result from mistakes in replication are removed by repair processes (Chapter 27). Those that remain can often assume alternative pairings that distort the helix only slightly. For example, a G-T wobble pair fits readily into an oligodeoxyribonucleotide double helix. Even a bulky G-A base pair causes little perturbation of the helix.[168] Incorporation of an extra nucleotide into one strand of the DNA will create a bulge in the helix.[169] NMR spectroscopic studies on bulged oligonucleotides have shown that the extra base can be stacked into the helix causing a sharp bend.[170,171] However, some oligonucleotides with mismatched bases crystallize as straight helices with the extra nucleotide looped out.[171–173] Mismatched base pairs tend to destabilize helices.[174]

Interactions with ions. Because each linking phospho group carries a negative charge (two charges per base pair, Fig. 5-2) the behavior of polynucleotides is strongly affected by cations of all kinds. The predominant small counterions within cells are K^+ and Mg^{2+}. They are attracted to the negative charges on the polynucleotide backbone and, although they remain mobile, they tend to occupy a restricted volume.[175–177] Some may bind in well-defined locations as in Fig. 5-8. Because of the presence of these positive ions the interactions of nucleic acids with cationic groups of proteins are strongly affected by the salt concentration.

Organic cations compete with the simple counterions K^+ and Mg^{2+}. Among these, the polyamines are predominant.[178] Crystal structures have revealed that spermine binds across the deep grooves of tRNA and the major groove of B-DNA. Spermine also binds into the deep groove of A-DNA interacting by hydrogen bonding with bases in GTG sequence in both strands.[179] It binds tightly to CG-rich sequences in the minor groove of Z-DNA, where it tightens the structure and shortens the helix.[110] At higher concentrations of DNA, as occur in cell nuclei, spermidine induces the conversion of the DNA into liquid crystalline phases.[180]

Heavier metal ions and metal complexes can find sites on nitrogen atoms of the nucleic acid bases. Examples are the platinum complex **cisplatin** and the DNA-cleaving antibiotic **neocarzinostatin** (Box 5-B). Can metals interact with the π electrons of stacked DNA bases? A surprising result has been reported for intercalating complexes of ruthenium (Ru) and rhodium (Rh). Apparent transfer of electrons between Ru (II) and Rh (III) over distances in excess of 4.0 nm, presumably through the stacked bases, has been observed,[181] as has electron transfer from other ions.[181a] Stacked bases are apparently semiconductors.[182]

C. The Topology and Dynamics of Nucleic Acids

1. Rings, Catenanes, and Knots

While a DNA molecule may exist as a straight rod, the two ends are often covalently joined. Thus, the chromosomes of *E. coli* and of other bacteria are single closed circles. Circular DNA molecules are also found in mitochondria, chloroplasts, and many viruses. Further complexity arises from the fact that the circles of DNA are sometimes interlocked in chainlike fashion (**catenated**). An unusual example of this phenomenon is the presence of thousands of small catenated DNA circles in the single mitochondrion of a trypanosome (Fig. 5-16).[183] Sometimes circular DNA is **knotted** as in Fig. 5-17.[184–186] Knots and catenanes often appear as intermediate forms during replication and recombination, especially involving circular DNA.[187,188]

Methods have been devised for synthesis of even very complex DNA knots.[185,186] Let's look briefly at the topology of knots. The three simple knots shown here have a chirality beyond that of the nucleotide

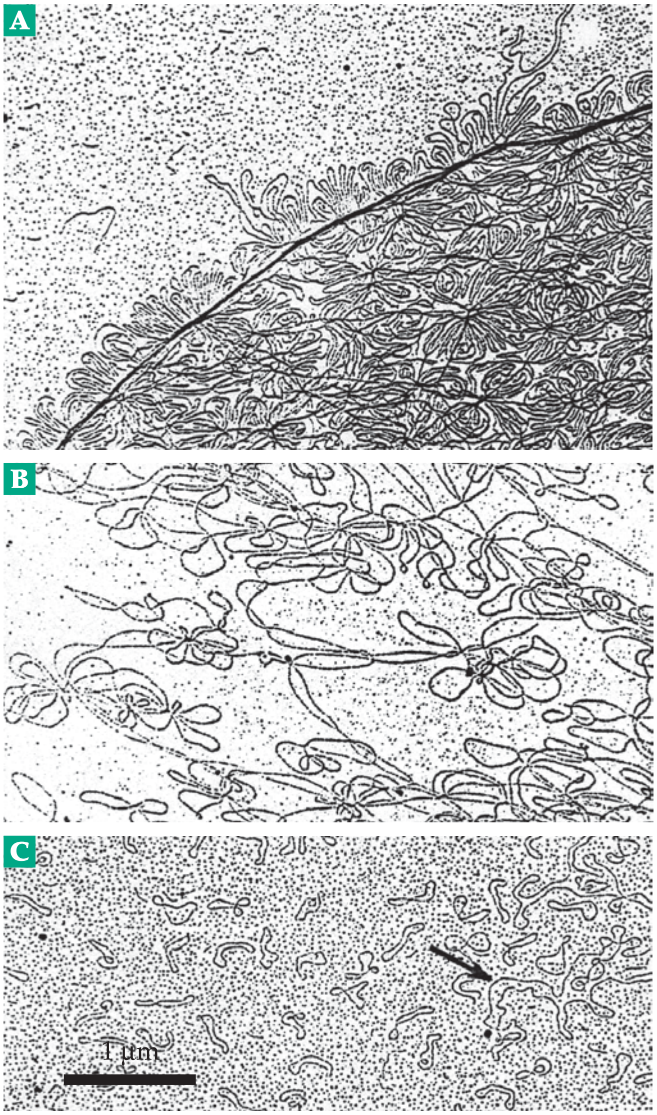
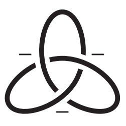


Figure 5-16 (A) Electron micrograph of the network of catenated DNA circles in the mitochondrion of the trypanosome *Crithidia fasciculata*. (B) and (C) The same network after treatment with a **topoisomerase** from bacteriophage T4 that catalyzes a decatenation to form individual covalently closed circles (Chapter 27). Five times as much enzyme was added in (C) as in (B). Two sizes of circles are present. Most are "minicircles", each containing about 2300 bp but a smaller number of larger ~35-kb "maxicircles" are also present. One of these is marked by the arrow. From Marini, Miller, and Englund.[183]

units.[189,190] This can be expressed by indicating the sign of each node in the knot:



Trefoil knot (–)

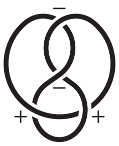
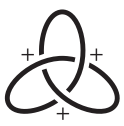


Figure-8 knot



Trefoil knot (+)

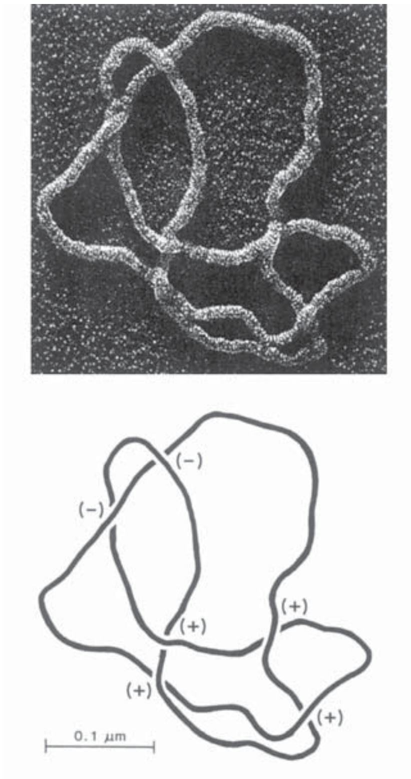
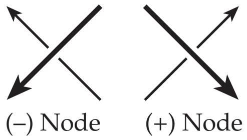


Figure 5-17 Electron micrograph of a six-noded knot made by the Tn3 resolvase which is involved in movement of the Tn3 transposon (Chapter 27) from one location to another within the genome. Putative six-noded knot DNA was isolated by electroelution from an agarose gel. The knots, which are nicked in one strand, were denatured to allow the nicked strand to slide away and leave a ssDNA knot. This was coated with *E. coli* recA protein (Fig. 27-24) to greatly thicken the strand and to permit the sign of each node (designated in the tracing) to be seen. From Wasserman *et al.*[184]

The node is negative if the crossing is like that in a left-handed supercoil and positive if like that in a right-handed supercoil:[191,192]



2. Supercoiled DNA

Double-stranded DNA in solutions of low salt content usually assume the B-DNA conformation with 10.4–10.5 base pairs per turn. If the two ends are joined the resulting covalently closed circular DNA will be "relaxed." However, there are topoisomerases that act on this form of DNA by cutting both strands, holding the ends, and twisting the two chains (Chapter 27).

The energy source for the process is provided by cleavage of ATP. **DNA gyrase** untwists relaxed circular dsDNA one turn at a time and reseals the cut ends. A **reverse gyrase** from certain bacteria twists the relaxed DNA more tightly. In both cases the change causes the DNA to form **superhelical turns**.[67,193–197] These may be either **solenoidal** or **plectonemically interwound** (as a twisted thread; Fig. 5-18).

The geometric and topological properties of closed supercoiled DNA molecues may be described by three quantities: The **linking number** (***Lk*** also called the winding number, α), the **twist** (***Tw***), and the **writhe** (***Wr***). If a segment of double helical DNA were laid on a flat surface and the ends were joined to form a relaxed circle both *Lk* and *Tw* would equal the number of helical turns in the DNA. The writhe *Wr* would be zero. The linking number is a topological property. It has an integral value which is unchanged if the DNA molecule is distorted. It can be changed only for DNA with open ends. When the ends are joined the linking (winding) number is constant unless one or both chains are cleaved. However, twist and writhe are geometric properties, which can change according to Eq. 5-8 while the *Lk* remains constant. The twist is related to the number of helical turns while the writhe is related to the number of superhelical turns.[27,67,198–200]

$$Lk = Tw + Wr \tag{5-8}$$

However, *Tw* and *Wr* do not usually have integral values. It is hard to define the relaxed state for a circular DNA. For example, changing the ionic composition will alter *Tw* and *Wr* according to equation 5-8. Both *Lk* and *Tw* are taken as positive for a right-handed **toroidal** (solenoidal) supercoil or a left-handed interwound twist.

For relaxed B-DNA, *Lk* is equal to the total number of base pairs in the circle divided by 10.5, $Wr = 0$, and $Tw = Lk$. Both *Lk* and *Tw* are positive in the right-handed B- and A-DNA forms, but *Tw* is negative for Z-DNA. In closed circular DNA the value of *Wr* is usually negative, the secondary structure being a fully formed Watson–Crick helix but with right-handed interwound superhelical turns or left-handed toroidal superhelical turns. The helix is said to be **underwound** ($Lk < Tw$).

Some of the topological properties of double-stranded DNA can be demonstrated by twisting together two pieces of flexible rubber tubing whose ends can be joined with short rods to form a closed circle.[196] Twist the tubing in a right-handed fashion as tightly as possible without causing supercoiling ($Lk = Tw$; $Wr = 0$). If the ends are now joined the circle will be relaxed. Now twist one turn tighter before joining the ends. A right-handed toroidal supercoil will be formed ($\Delta Wr = 1$; $\Delta Lk = 1 + Tw$; *Tw* is the same as before). If twisting is continued until several supercoils appear before the ends are joined, two interconvertible forms result. One form has right-handed toroidal supercoils and the other left-handed interwound supercoils as in Fig. 5-19. If relaxed circular DNA is unwound by one turn by cutting and resealing one chain, a single right-handed interwound supercoil will be formed ($\Delta Wr = -1$, $Lk = Tw - 1$). On the other hand, if the two chains in a closed relaxed helix are pried apart, as happens during intercalation (Section 3), *Lk* must remain constant, *Tw* will decrease, and *Wr* will increase with appearance of left-handed supercoiling.

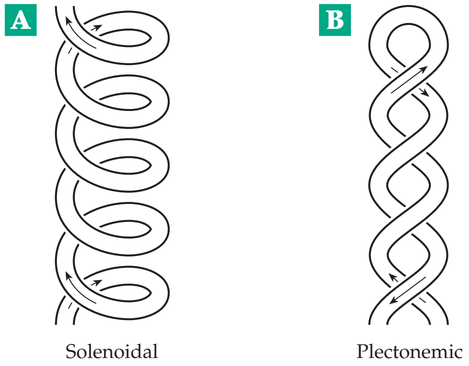


Figure 5-18 Two forms of supercoiling of a DNA duplex: (A) Solenoidal. (B) Plectonemically interwound. These are both negatively supercoiled, as can be deduced from the arrows at the nodes, but the solenoid is left-handed and the plectonemic form right-handed. From Wasserman and Cozzarelli.[187]

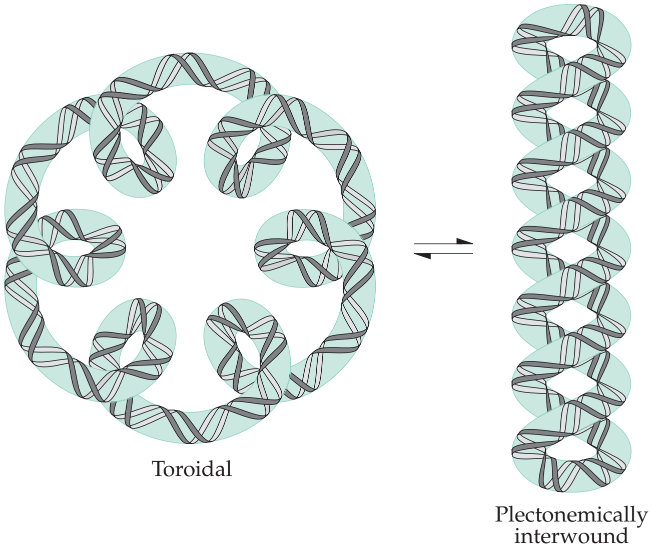


Figure 5-19 Topological equivalence of toroidal (solenoidal) and plectonemically interwound forms of a circular DNA. These two forms have a constant value of the linking number *Lk* (or α), the twist *Tw*, and writhing number *Wr*.

During replication of DNA (Chapter 27) pairing of some bases associated with the replication apparatus is prevented. Upon release of the constraint the newly replicated DNA forms base pairs and becomes supercoiled.

Since Lk is constant in circularly closed DNA a change of one turn of B-DNA ($Tw = 1$) into a turn of Z-DNA ($Tw = -1$) will cause the writhing number Wr to change by −2. Conversely, if the writhing number is forced to change by −2, a turn of Z-DNA may develop somewhere in a suitable (G + C)-rich region of the DNA.

Some of the information about supercoiled DNA can be summarized as follows:

Relaxed circular DNA	$Lk = Tw$, $Wr = 0$
Underwound circular DNA	$Lk < Tw$; Wr is negative; right-handed interwound supercoils or left-handed toroidal supercoils may be present. Alternatively $Lk = Tw$ and left-handed Z-DNA regions may appear.
DNA with intercalated molecules	DNA is partially untwisted; Tw is lowered; Wr is increased with decrease in number of negative supercoils. As Tw approaches Lk the DNA becomes relaxed.

The **superhelix density** of a DNA molecule is often expressed as $\sigma = Wr/\,Tw \approx$ number of superhelical turns per 10 bp.[31,198,199a,199b,201] In most naturally occurring circular DNA molecules σ is negative, a typical value being −0.05 (~5 negative superhelical turns per 1000 bp). The presence of superhelices in circular DNA molecules can be recognized readily by its effect upon the sedimentation constant of the DNA. Naturally occurring supercoiled DNA from polyoma virus sediments rapidly but after nicking of one of the strands of the double helix by brief exposure to a DNA-hydrolyzing enzyme the resulting relaxed form of the molecule sediments more slowly. Supercoiling lowers the viscosity of solutions of DNA and increases the electrophoretic mobility (Fig. 5-20) and may also be recognized by electron microscopy.

Naturally occurring or artificially prepared supercoiled DNA molecules can often be separated by electrophoresis into about ten forms, each differing from the other by one supercoiled turn and by $\Delta Lk = \pm 1$ (Fig. 5-20). The relative amounts of these **topological isomers** form an approximately Gaussian distribution. The isomers apparently arise as a result of thermal fluctuations in the degree of supercoiling at the time that the circles were enzymatically closed.[202]

Why is DNA in cells supercoiled? One effect of supercoiling is to contract the very long, slender double helices into more compact forms. In eukaryotic cells much of the DNA exists in **nucleosomes**. Each bead-like nucleosome consists of a core of eight subunits of proteins called **histones** around which an ~140 bp length of DNA is coiled into two negative, left-handed toroidal superhelical turns (Fig. 5-21).[203,204] There is some spacer DNA between nucleosomes. Otherwise a nucleosome providing two superhelical turns per 200 bp would produce a superhelix density of −0.1, twice that observed. A detailed analysis of the geometric properties of DNA wrapped around nucleosomes or other protein particles has been developed.[199−199c] Nucleosome formation is thought to protect DNA and to keep it in a more compact state than when it is fully active and involved in transcription. Nucleosomes, discussed further in Chapter 27, are also important in the regulation of transcription. Even though nucleosomes as such are absent, bacterial DNA also has a superhelix density of − 0.05, apparently a result of interaction with other proteins such as the histonelike HU.[205] Interaction with smaller molecules can also affect supercoiling.[206]

The presence of naturally supercoiled DNA and a variety of topoisomerases suggests that the control of DNA supercoiling is biologically important. In fact,

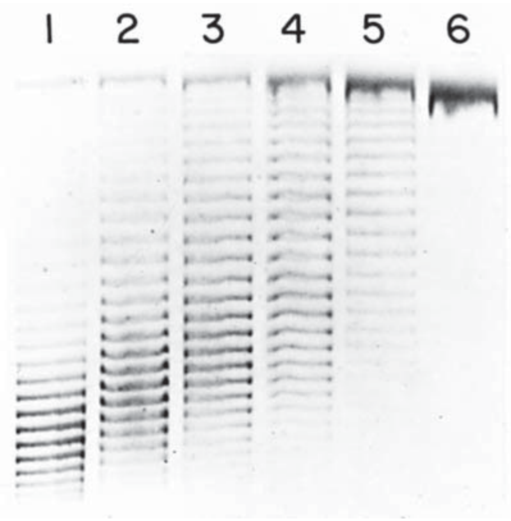


Figure 5-20 Electrophoresis of DNA from the SV40 virus with varying numbers of superhelical turns. Molecules of the native DNA (lane 1) move rapidly toward the anode as a series of bands, each differing from its neighbors by one superhelical turn. The average number of superhelical turns is about 25. Incubation with a topoisomerase from human cells causes a stepwise removal of the superhelical turns by a cutting and resealing of one DNA strand. The DNA incubated with this enzyme at 0°C for periods of 1, 3, 6, 10, and 30 min (lanes 2−6) is gradually converted to a form with an average of zero supercoils. For details see Keller.[202,212] Electrophoresis was carried out in an 0.5% agarose −1.9% polyacrylamide slab gel (17 x 18 x 0.3 cm). The bands of DNA were visualized by staining with the fluorescent intercalating dye ethidium bromide. Photograph courtesy of Walter Keller.

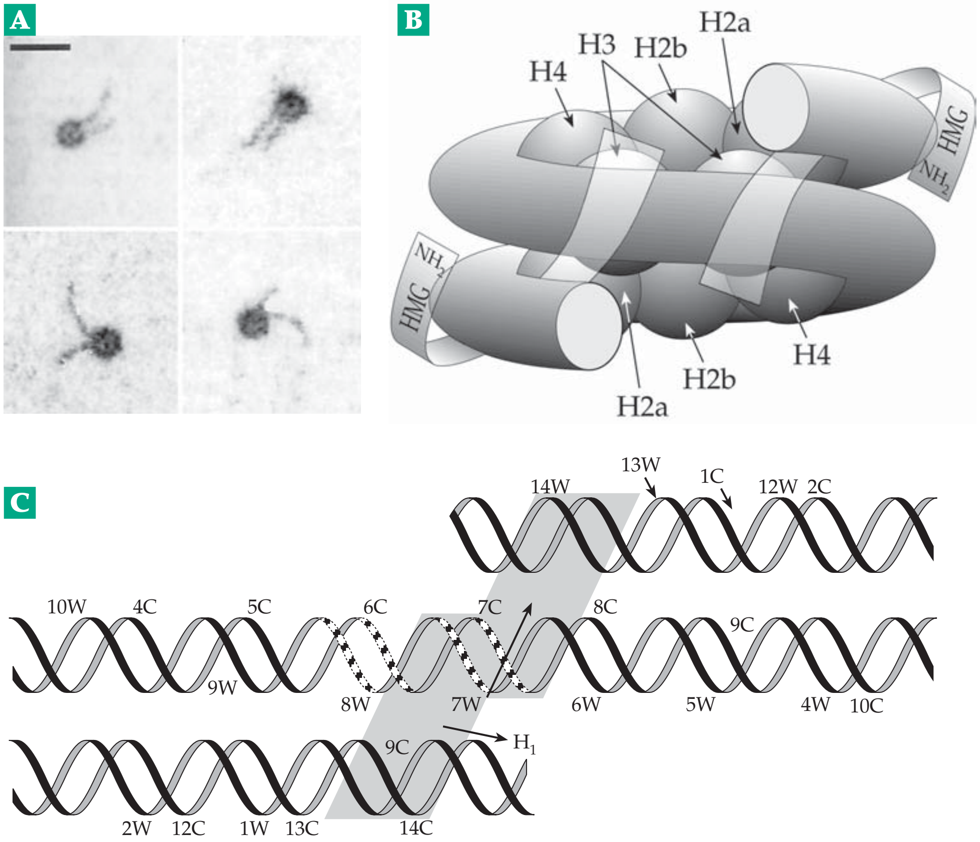


Figure 5-21 Nucleosomes. (A) Electron micrographs of individual nucleosomes reconstituted from 256-bp DNA fragments and separated proteins. From Hamiche *et al.*[213] Courtesy of Ariel Prunell. (B) Model of a nucleosome core. The 1.75-turn (145-bp) DNA superhelix winds around the histone octomer which consists of two subunits apiece of histones H2A, H2B, H3, and H4. In addition, two elongated molecules of proteins HMG-14 or HMG-17 are indicated (see also Chapter 27). (C) Schematic radial projection of the double-helical DNA showing areas protected from cleavage by hydroxyl radicals (see Fig. 5-50) by the bound proteins. The shaded areas are those protected by HMGs. The zigzag lines near the dyad axis indicate the most prominent regions of protection. (B) and (C) are from Alfonso *et al.*[214]

---

topoisomerases play essential roles in both replication of DNA and in transcription of genes. Supercoiling requires energy and topoisomerases that induce supercoiling must provide energy, e.g., by cleavage of ATP. Conversely, supercoiled DNA can be a source of energy for biological processes. It has been estimated that the ΔH, ΔS, and ΔG per mole for formation of a single superhelical turn are 35.9 kJ, 68 J/ °K, and 14.6 kJ, respectively.[207] A reduction in superhelix density is accompanied by a decrease in Gibbs energy and can therefore be coupled to other processes that have positive values of ΔG. An example is conversion of (G + C)-rich regions into the Z form of DNA, which is favored by negative supercoiling.[207,208] Supercoiling affects binding of various proteins to DNA[197,209,210] as well as intercalation (discussed in the next section) and formation of cruciform structures (Section D,3).

The folding of DNA into compact forms, such as that in chromosomes, is also influenced by supercoiling. In the absence of nucleosomes supercoiled DNA may assume the plectonemic form (Figs. 5-18, 5-19). Segments of the resulting rods may aggregate side-by-side in a liquid crystalline state.[200,211] The relatively high cation concentration within cells favors this transformation.[215–217] Polyamines such as spermidine are especially effective in promoting aggregation of DNA, in formation of Z-DNA,[110] and possibly in facilitating cooperative processes that require that two DNA molecules interact with each other.[218,219]

3. Intercalation

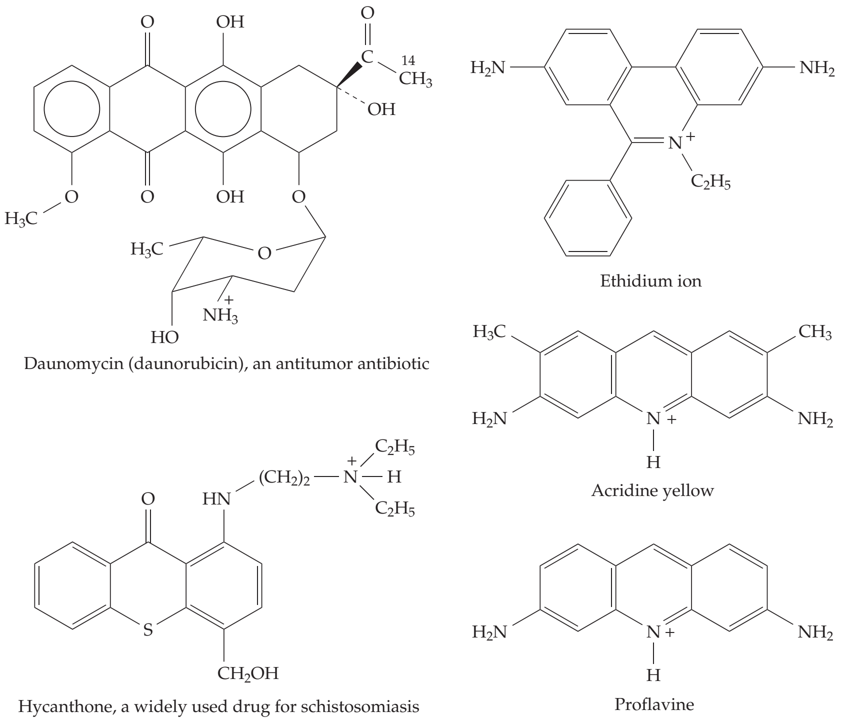
Flat, aromatic, hydrophobic rings are often able to insert themselves between the base pairs of a DNA duplex. Such **intercalation** is observed for many antibiotics, drugs, dyes, and environmental pollutants. Among them are proflavine, ethidium bromide, actinomycin (Box 28-A), hycanthone (Fig. 5-22), and daunomycin (Fig. 5-23). Hycanthone, employed in the treatment of schistosomiasis, is one of the most widely used drugs in the world. Since intercalating agents can be mutagenic, such drugs are not without their hazards.

Intercalation is often used to estimate the amount of negative supercoiling of DNA molecules. Varying amounts of the intercalating agent are added, and the sedimentation constant or other hydrodynamic property of the DNA is observed. As increasing intercalation occurs, the secondary turns of DNA are unwound (the value of *Tw* in Eq. 5-8 decreases). Each intercalated ring causes an unwinding of the helix of ~26°. Since for a closed covalent duplex the value of *Lk* in Eq. 5-8 is constant, the decrease in *Tw* caused by increased intercalation leads to an increase in the value of *Wr*, which is usually negative for natural DNA. When sufficient intercalation has occurred to raise *Wr* to zero, a minimum sedimentation rate is observed. Addition of further intercalating agent causes a positive supercoiling.

When the "replicative form" of DNA of the virus ϕX174 (Chapter 27), a small circular molecule containing ~5000 bp, was treated with proflavine,[221] the binding of 0.06 mol of proflavine per mole of nucleotides reduced *Wr* to zero. From this it was estimated that $\sigma = 0.055$, corresponding to −27 superhelical turns at 25°, pH 6.8, ionic strength ~0.2. Changes in temperature, pH, and ionic environment strongly influence supercoiling. In general σ becomes less negative by $\sim 3.3 \times 10^{-4}$ per degree of temperature increase.[222] For example, the observed value[221] of σ for ϕX174 DNA was −0.059 at 15°C and −0.040 (−20 superhelical turns) at 75°C at an ionic strength of ~0.2.

The exact ways in which intercalating substances can fit between the base pairs of nucleic acids are being revealed by X-ray diffraction studies of complexes with nucleosides, dinucleotides, and other oligonucleotides.[220,223–225] The structure of a complex in which daunomycin is intercalated between two GC base pairs in DNA is shown as in Fig. 5-23. Daunomycin and other related anthracycline antibiotics also have an amino sugar ring that binds into the minor groove of the DNA, providing both electrostatic stabilization and hydrogen bonding. Substituents on this aliphatic ring also hydrogen bond to DNA bases.[226]

Does intercalation of flat molecules into nucleic acid chains have a biochemical function? Aromatic rings of amino acid side chains in proteins designed to interact with nucleic acids may sometimes intercalate into nucleic acid helices serving a kind of "bookmark" function.[227] Changes in superhelix density caused by such intercalation may be important in the orderly handling of DNA by enzymes within cells.



Daunomycin (daunorubicin), an antitumor antibiotic

Ethidium ion

Acridine yellow

Hycanthone, a widely used drug for schistosomiasis

Proflavine

Figure 5-22 Structures of some substances that tend to "intercalate" into DNA structures. See also Fig. 5-23.

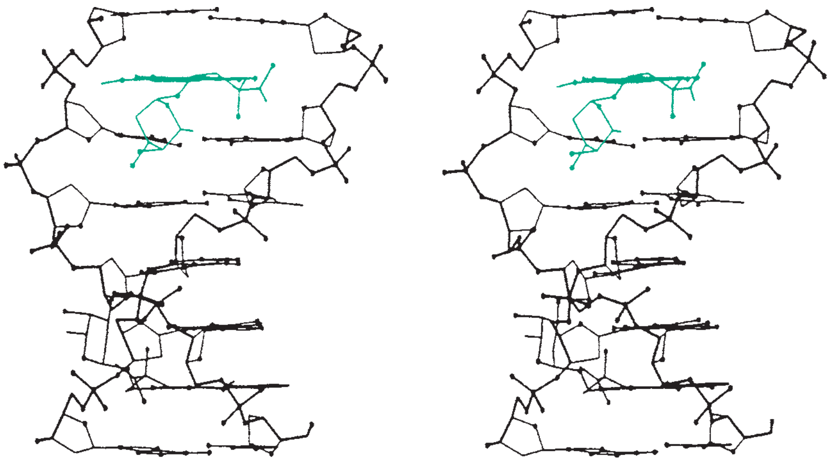
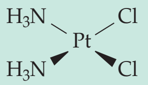


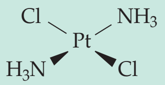
Figure 5-23 Stereoscopic drawing showing a molecule of daunomycin (Fig. 5-22) intercalated between two base pairs in a molecule of double-helical DNA, d(CGTACG). Nitrogen and oxygen atoms are shown as dots. From Quigley *et al.*[220] Both daunomycin and adriamycin (doxorubicin; 14-hydroxydaunomycin) are important but seriously toxic anticancer drugs.

BOX 5-B ANTITUMOR DNA DRUGS

Chemotherapy of cancer at present involves simultaneous use of two or more drugs. For example, antifolates (Chapter 15) or nucleoside analogs such as 5-fluorouridine may be used together with a drug that binds directly to DNA and inhibits the replication of cancer cells. In 1963, it was discovered accidentally that platinum ions released from supposedly inert platinum electrodes inhibited the division of *E. coli* cells. This led Rosenberg and associates to test platinum compounds against animal cells.[a] Among many compounds tested *cis*-dichlorodiammineplatinum(II) (*cis*-DDP or **cisplatin**) emerged as an important anticancer drug that is especially effective against testicular and ovarian cancers.[b–d] The *trans* isomer, however, is inactive.

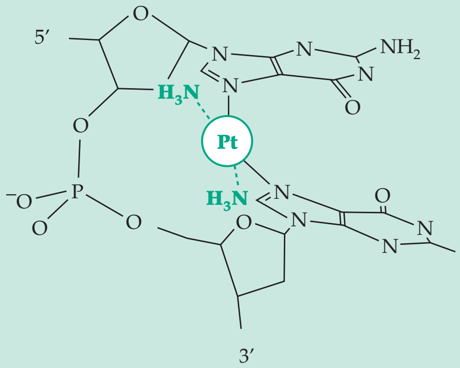


cis-DDP
(Cisplatin)



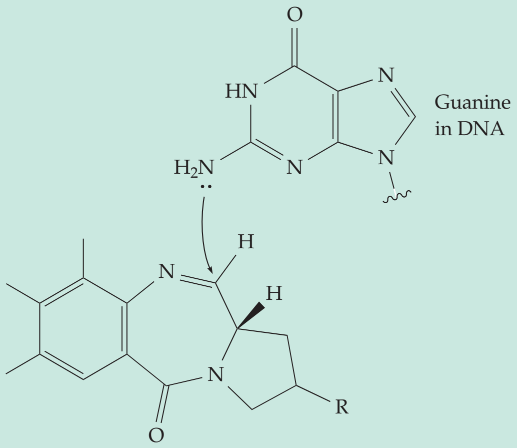
trans-DDP
(inactive)

cis-DDP binds to adjacent deoxyguanosines within one strand of ds or ssDNA. The stacking of adjacent bases is disrupted as the platinum binds to N-7 nitrogen atoms of the two guanine rings by replacement of the two chloride ions. The product has the following structure,[c,e] with the Pt lying in the minor groove of dsDNA:



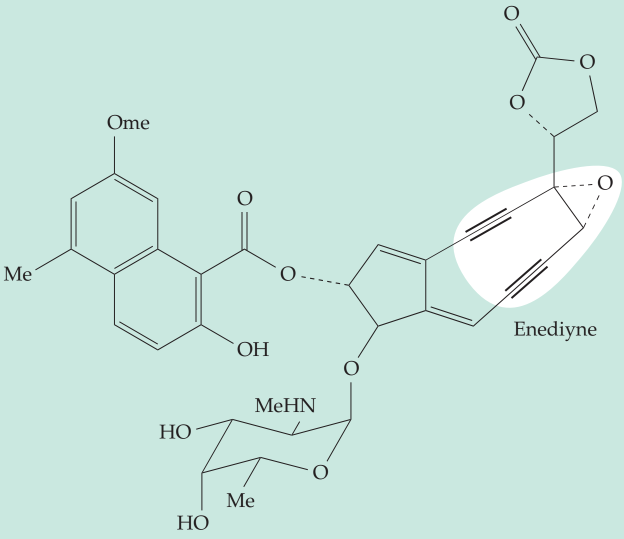
The two guanines are no longer stacked, but the structure is impossible for the *trans* isomer. The cisplatin adducts appear to prevent proper DNA repair and to induce programmed cell death (apoptosis).[f]

Since the 1950s, using a different approach, the U.S. National Cancer Institute, as well as agencies in other countries, has sought to find natural anticancer compounds in plants, fungi, microorganisms, and marine invertebrates.[g] Among these are many antibiotics that intercalate into DNA helices, e.g., **daunomycin** (Figs. 5-22, 5-23), **menogaril**,[h] **triostin A**,[i] and the antitrypanosomal drug **berenil.**[j] Some of these are also alkylating agents that contain double bonds to which such groups as the 2-NH$_2$ of guanine may add:



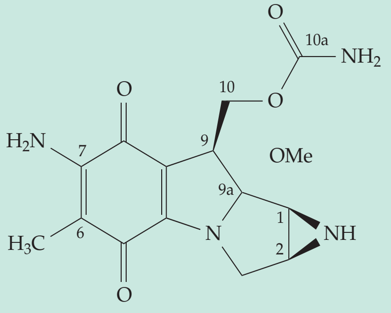
Tomaymycin and related antibiotics[k]

Diol epoxides and cyclic imines such as **mitomycin** also form adducts specifically with guanine 2-NH$_2$ groups:[l] **Neocarzinostatin** is an antitumor protein with a nonprotein "chromophore." After intercalation and binding into the minor groove of bulged DNA, it undergoes "activation" by addition of a thiol group. The enediyne structure undergoes rearrangement with formation of a reactive diradical that attacks the DNA.[m] A family of related antitumor enediynes has also been discovered.[n]

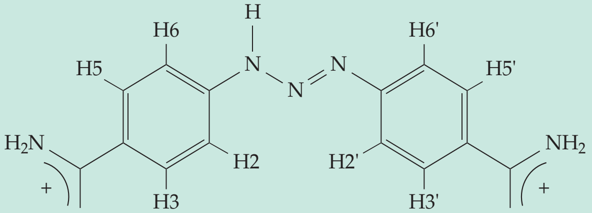


Neocarzinostatin chromophore

BOX 5-B (continued)



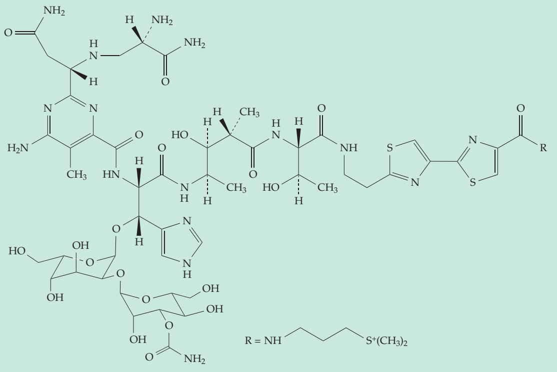
Mitomycin



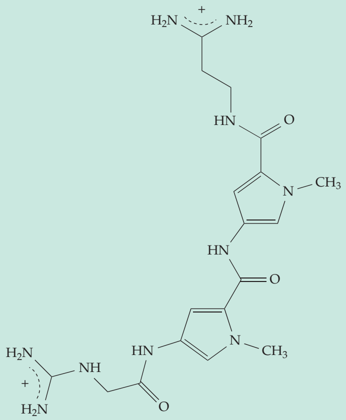
Berenil

Another group of drugs occupy extended binding sites in the minor grooves of DNA double helices, often with specificity for a particular base sequence. Examples are the antitrypanosomal drug **berenil**,[j] toxic *Streptomyces* antibiotics **netropsin, distamycin**, and related synthetic compounds.[o–q] Netropsin lies within the minor groove in regions with two or more consecutive AT pairs, displacing the spine of hydration as shown in the following stereoscopic drawing. Binding depends upon both electrostatic interactions and formation of specific hydrogen bonds involving the amide groups of the antibiotics.

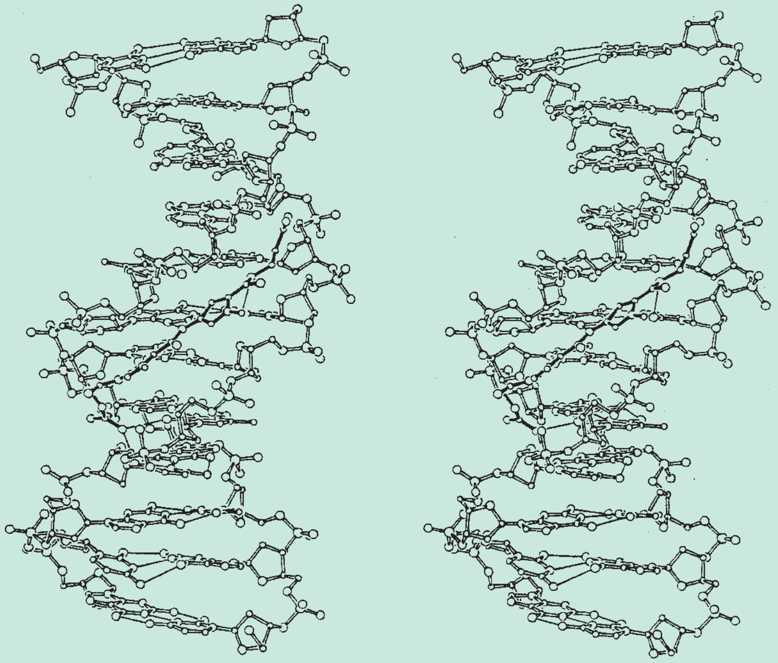
The antibiotic **bleomycin**, which also binds in the minor groove of B-DNA with some specificity for G-C Sites, forms an iron (II) complex. It can be oxygenated to form an Fe(II)–O_2 complex (see Chapter 16) which cleaves the DNA chain.[s,t] Synthetic compounds that do the same thing have been made by connecting an EDTA–iron, or other iron chelate complex covalently to the DNA-binding compound.[n,u–x] A goal is to direct drugs to selected target sites in DNA and to induce bond cleavage at those sites in a manner analogous to that observed with restriction endonucleases. Most current chemotherapeutic agents are very toxic. Present research is designed to target these drugs more precisely to specific DNA sequences and to identify target sequences peculiar to cancers. See also **Designed third strands** in main text.



Bleomycin A_2



Netropsin



Netropsin lying in the minor groove of B-DNA hydrogen-bonded to bases in the central ATAT tetranucleotide. From Coll *et al*.[r]

BOX 5-B Continued

[a] Rosenberg, B., Van Camp, L., and Krigas, T. (1965) *Nature (London)* **205**, 698–699
[b] Zamble, D. B., and Lippard, S. J. (1995) *Trends Biochem. Sci.* **20**, 435–439
[c] Pilch, D. S., Dunham, S. U., Jamieson, E. R., Lippard, S. J., and Breslauer, K. J. (2000) *J. Mol. Biol.* **296**, 803–812
[d] Temple, M. D., McFadyen, W. D., Holmes, R. J., Denny, W. A., and Murray, V. (2000) *Biochemistry* **39**, 5593–5599
[e] Huang, H., Zhu, L., Reid, B. R., Drobny, G. P., and Hopkins, P. B. (1995) *Science* **270**, 1842–1845
[f] Zlatanova, J., Yaneva, J., and Leuba, S. H. (1998) *FASEB J.* **12**, 791–799
[g] Booth, W. (1987) *Science* **237**, 969–970
[h] Chen, H., and Patel, D. J. (1995) *J. Am. Chem. Soc.* **117**, 5901–5913
[i] Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., and Rich, A. (1984) *Science* **225**, 1115–1121
[j] Pilch, D. S., Kirolos, M. A., Liu, X., Plum, G. E., and Breslauer, K. J. (1995) *Biochemistry* **34**, 9962–9976
[k] Barkley, M. D., Cheatham, S., Thurston, D. E., and Hurley, L. H. (1986) *Biochemistry* **25**, 3021–3031
[l] Kumar, G. S., Lipman, R., Cummings, J., and Tomasz, M. (1997) *Biochemistry* **36**, 14128 –14136
[m] Yang, C. F., Stassinopoulos, A., and Goldberg, I. H. (1995) *Biochemistry* **34**, 2267–2275
[n] Nicolaou, K. C., Dai, W.-M., Tsay, S.-C., Estevez, V. A., and Wrasidlo, W. (1992) *Science* **256**, 1172–1178
[o] Goodsell, D. S., Ng, H. L., Kopka, M. L., Lown, J. W., and Dickerson, R. E. (1995) *Biochemistry* **34**, 16654–16661
[p] Rentzeperis, D., Marky, L. A., Dwyer, T. J., Geierstanger, B. H., Pelton, J. G., and Wemmer, D. E. (1995) *Biochemistry* **34**, 2937–2945
[q] Tanious, F. A., Ding, D., Patrick, D. A., Tidwell, R. R., and Wilson, W. D. (1997) *Biochemistry* **36**, 15315–15325
[r] Coll, M., Aymami, J., van der Marel, G. A., van Boom, J. H., Rich, A., and Wang, A. H.-J. (1989) *Biochemistry* **28**, 310–320
[s] Burger, R. M., Drlica, K., and Birdsall, B. (1994) *J. Biol. Chem.* **269**, 25978–25985
[t] Kane, S. A., Hecht, S. M., Sun, J.-S., Garestier, T., and Hélène, C. (1995) *Biochemistry* **34**, 16715–16724
[u] Veal, J. M., and Rill, R. L. (1988) *Biochemistry* **27**, 1822 –1827
[v] Campisi, D., Morii, T., and Barton, J. K. (1994) *Biochemistry* **33**, 4130–4139
[w] Mack, D. P., and Dervan, P. B. (1992) *Biochemistry* **31**, 9399–9405
[x] Han, H., Schepartz, A., Pellegrini, M., and Dervan, P. B. (1994) *Biochemistry* **33**, 9831–9844

4. Polynucleotides with Three or Four Strands

Some nucleotide sequences in DNA favor the formation of a regular triple-helical (**triplex**) structure. This is possible when there are long stretches of adjacent pyrimidines having any sequence of C and T in one strand of double-helical DNA. The other strand of the DNA will contain the correct purines for formation of Watson–Crick base pairs. With such a structure it is always possible to add a third strand of a polypyrimidine using Hoogsteen base pairing. Triads formed in this way contain either two T's and one A or two C's and one G, as is shown in Fig. 5-7. In the latter case one of the C's must be protonated to allow formation of the pair of hydrogen bonds.[227a] These triplex structures can be formed only for stretches of DNA containing all pyrimidines in one strand (a **homopyrimidine** strand) and all purines in the other (a **homopurine** strand). A poly (AAU) triplex can also be formed, using the hydrogen bond pattern of the first triplet in Fig. 5-7. Triple-stranded synthetic polynucleotides of these types were prepared by Felsenfeld and others as early as 1957.[228–230] The third strand in a triplex can be either parallel or antiparallel to the homopurine strand. The third strand is either homopurine or a mixture of purine and thymine. The triplets are G•C*C, A*A•T, or T*A•T and are formed by Watson–Crick (•) and **reversed Hoogsteen** (*) pairing.[231] Recently, there has been a renewed interest in DNA triplets because of their occurrence in natural DNA, their possible importance in genetic recombination, and the potential for design of powerful inhibitors of replication and transcription that function via triplex formation.[35,232–238]

H-DNA. Strands of DNA contain many homopyrimidine "tracts" consisting of repeated sequences of pyrimidines, e.g., $d(T\text{-}C)_n$, which may be abbreviated more simply as $(TC)_n$. In this example the complementary strand would contain the two purines in the repeated sequence $(GA)_n$. At low pH, where protonation of the cytosine rings occurs, or in negatively supercoiled DNA, the two strands of the repeating sequence may separate, with the $(TC)_n$ strand folding back to form a triple helix in which the base triplets have the hydrogen bonding pattern of Fig. 5-7. The resulting structure, which is shown in Fig. 5-24, is known as **H-DNA**.[239–241] A variety of related "nodule" and looped forms of DNA can also be formed.[35,242]

R-DNA. A different type of triplex DNA may be formed during genetic recombination. A Watson–Crick duplex is brought together by one or more proteins with a single strand that is, for at least a considerable distance, an exact copy of one of the strands of the duplex. It is within such a triplex that cutting of a strand of the duplex takes place to initiate recombination (see Chapter 27 for a detailed discussion). Can a triplex structure containing two identical chains be formed? Possible base triplets include the following C•G*G triplet which occurs in a crystalline oligonucleotide structure.[243] This triplet contains a variation on Hoogsteen pairing and is related to the first triplet in Fig. 5-7. The other triplets needed for the proposed **R-DNA** recombination intermediates are G•C*C, T•A*A, and A•T*T. While there is keen interest in R-DNA[35,244–247] the formation of a stable intermediate triplex is still uncertain.[35,247]

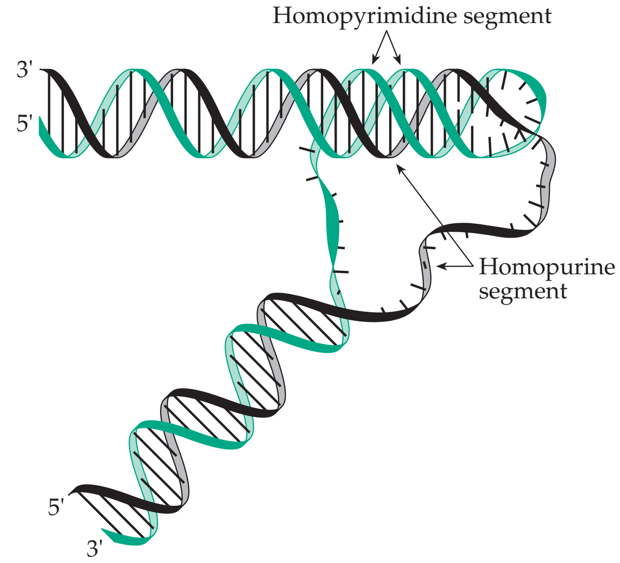
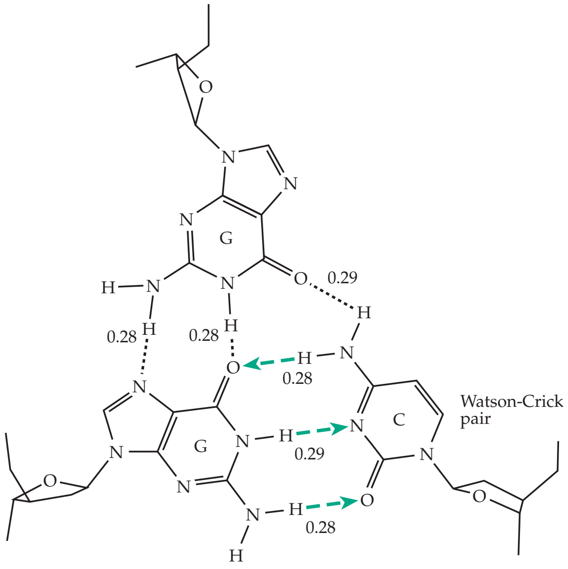
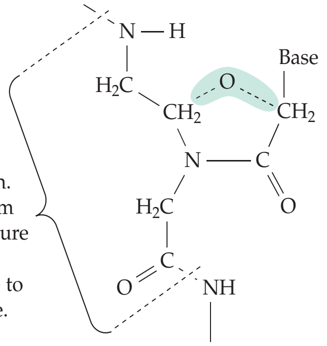


Figure 5-24 Proposed structure of H-DNA which can be formed when a homopyrimidine segment and its complementary homopurine segment separate as a result of protonation of the cytosine rings or of negative supercoiling stress.[239,240] The triple-helical portion contains base triplets of the kind shown in Fig. 5-7.



Designed third strands. If triplex DNA segments can form naturally it should be possible to design oligonucleotides that will bind into the major groove of a DNA duplex to form a triplex at a specific location or locations in the genome. Dervan and associates are studying this approach systematically.[248–250] Such oligonucleotides may be modified chemically to provide stronger binding to targeted locations and may prove to be useful therapeutic agents. Among these are synthetic phosphoramidates in which the phosphodiester linkage between nucleosides is replaced by $3'\text{-}NH{-}PO_3^{-}{-}O\text{-}5'$. The resulting oligonucleotide is resistant to digestion by phosphodiesterases present in cells. Its geometry favors binding to A-DNA or RNA.[251] Another oligonucleotide mimic, dubbed PNA (peptide nucleic acid), contains monomer units of the following type[252–253b] and is able to form a triple helix:

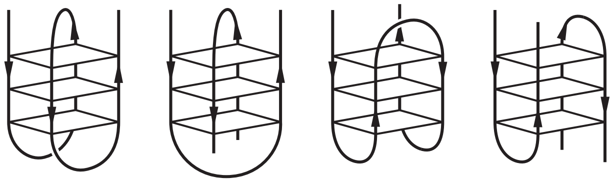


Residue of a PNA chain. The phantom green structure shows the relationship to a nucleotide.

Yet another approach is to synthesize a "hairpin polyamide" that contains pyrrole, hydroxypyrrole, and imidazole groups in a sequence that favors tight binding to a specific dsDNA sequence.[254] Some possible geometries for triplex DNA are illustrated in Fig. 5-25. These are based on computer-assisted modeling.[238]

Tetraplex (quadruplex) structures. The ends of linear chromosomes, the **telomeres**, have unusual nucleotide sequences repeated hundreds or thousands of times.[255 – 258] There is usually a **guanine-rich strand** running 5' to 3' toward the end of the chromosome and consisting of sequences such as TTAGGG in vertebrates, TTTTGGGG in the ciliate *Oxytricha*, and TG_{1-3} in *Saccharomyces cerevisiae*. The complementary strand is cytosine rich. The guanine-rich strand is longer than its complement, "overhanging" by about 2 repeat units. The significance for the replication of chromosomes is discussed in Chapter 27.

The thing that has attracted most attention to telomeric DNA is the unusual structure of the G-rich strand that was signaled by the first NMR studies.[259] Subsequent investigation[260,261] revealed the presence of G quartets (Fig. 5-8) which are apparently stacked in folding patterns such as the following:[255,262–268]



These structures are stabilized by the presence of univalent cations, K⁺ being more effective than Na⁺,

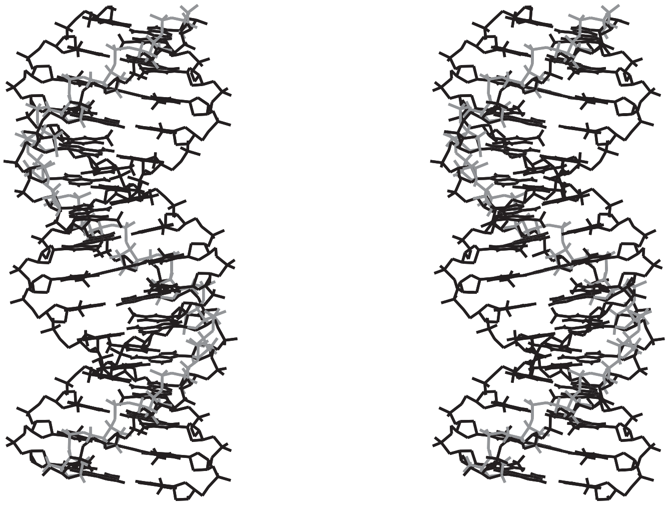
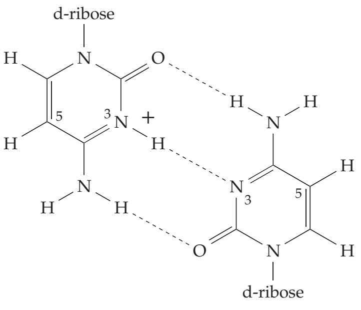


Figure 5-25 Two superposed stereoscopic diagrams illustrating the lowest energy triple helical conformations of one pyrimidine, one purine, and one Hoogsteen DNA strand (in black). Superposed is an RNA purine strand in gray. See Srinivasan and Olson.[238] Courtesy of Wilma K. Olson.

Li^+, Rb^+, or Cs^+.[260,269–271] A protein isolated from *Tetrahymena* binds specifically to the G4-DNA regions.[272] The presence of 5-methylcytosine (m^5c) also stabilizes quadruplex structures.[273] The oligoribonucleotide UG_4U forms G_4 quartets so stable that it takes days at 40°C in D_2O for the hydrogen-bonded NH protons forming the quartet to be exchanged for ^2H.[274]

Tetraplex structures have also been observed for G-rich repeating sequences associated with the human **fragile X syndrome**.[275,276] This is the most common cause of inherited mental retardation and appears to arise as a result of the presence of an excessive number of repeats of the trinucleotide sequence $(CGG)_n$. For normal persons n = 60 or less; for healthy carriers n may be as high as 200 but for sick individuals it may be much higher.[275] The structure in solution, as determined by NMR spectroscopy, is shown in Fig. 5-26. Another variant of four-stranded DNA, which arises from cytosine-rich DNA, contains $C \cdot CH^+$ pairs such as the following at low pH.[277–279]



Hemiprotonated $d(CH^+ \cdot C)$

In sequences such as $d(TC_5)$ and $d(C_3T)$ these $C \cdot CH^+$ pairs are intercalated as is shown in Fig. 5-27. This intercalated DNA (**I-DNA**) may provide an alternative conformation for some telomeric sequences.[281,282] An I-DNA motif has also been identified in oligonucleotides from the DNA of human centromeres.[283] The seemingly unusual forms of DNA described in this section may represent only a fraction of the naturally occurring DNA structures of biological significance.

5. Junctions

Special structural features may be found at junctions between different types of DNA, e.g., between A-DNA and B-DNA.[284–286] However, the most interesting junctions are *branched*.[287–290] For example, Fig. 5-28 shows a four-way junction in which all of the bases form Watson–Crick pairs. This junction is better known as a **Holliday junction** because it was proposed by Holliday in 1964 as an intermediate in genetic recombination.[291] As shown at the top of Fig. 5-28A the junction is formed from *two homologous DNA duplexes.* These are identical except for the boxed and shaded base pairs. The ends of the first duplex are marked I and II and those of the second III and IV. The Holliday junction appears to arise by cleavage of one strand of each duplex with rejoining of the strands as indicated by the green arrows. Rotation gives the untwisted Holliday junction structure

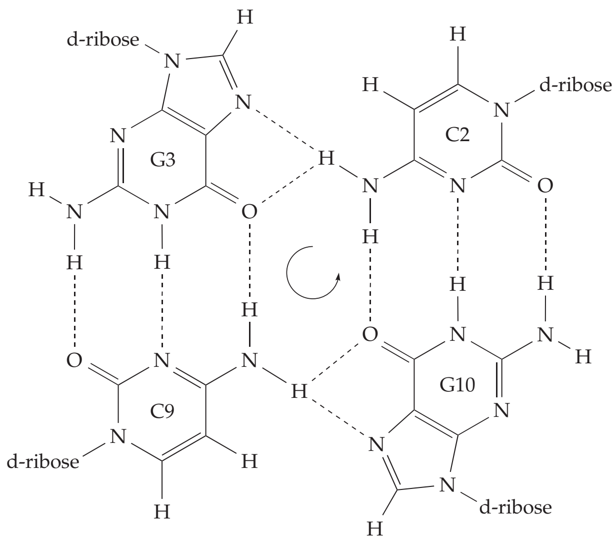


Figure 5-26 Structure of a G•C∗G•C∗ tetrad present in a quadruplex structure formed by the oligonucleotide d(GCG-CTTTGCGC) in Na^+-containing solution. See Kettani *et al*.[276]

shown. The postulated three-dimensional structure of the junction is indicated in Fig. 5-28B.[292–294]

An important characteristic of Holliday junctions formed from homologous duplexes is that they can move by a process called **branch migration**.[295] Because of the twofold symmetry of the branched structure the hydrogen bonds of one base pair can be broken while those of a new base pair are formed, the branch moving as shown in Fig. 5-28. Notice that, in this example, the nonhomologous (boxed) base pairs TA and GC have become *mispaired* as TG and AC after branch migration. More significantly, the junction may be cut by a **resolvase** at the points marked

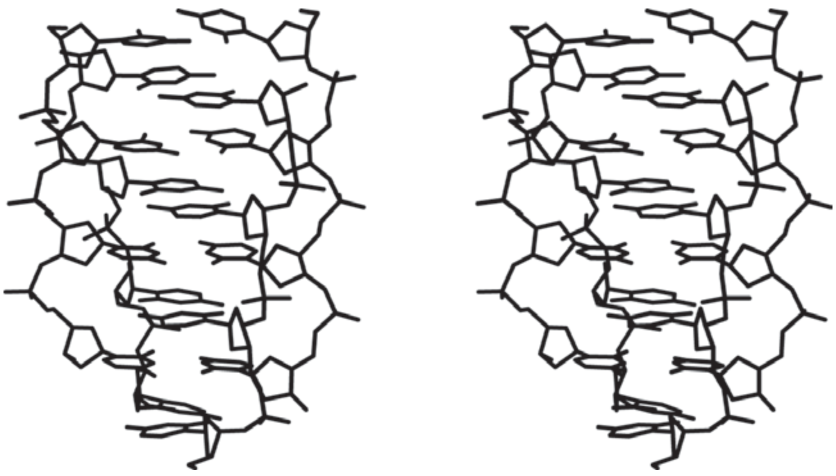
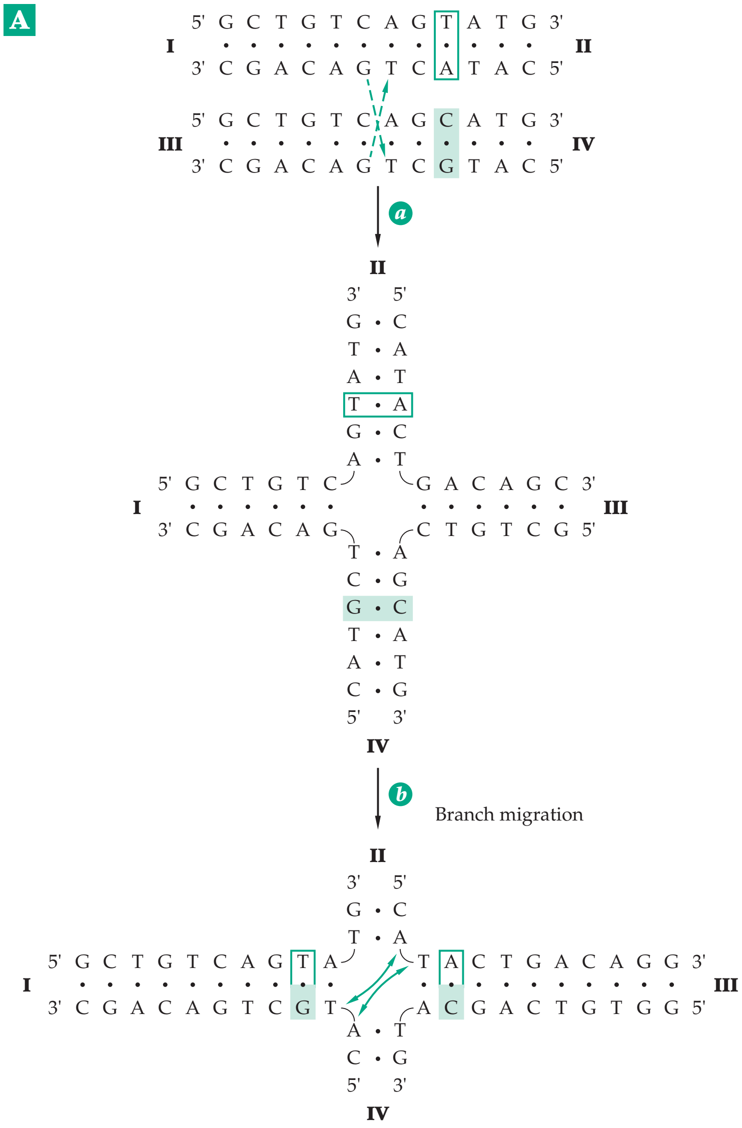


Figure 5-27 Stereoscopic view of a four-stranded intercalated DNA or I-DNA formed from d(C_4). Two parallel duplexes with C•CH^+ pairs are intercalated into each other. From Chen *et al*.[280]


A
5' G C T G T C A G T A T G 3'
I
II
3' C G A C A G T C A T A C 5'
5' G C T G T C A G C A T G 3'
III
IV
3' C G A C A G T C G T A C 5'
a
II
3' 5'
G • C
T • A
A • T
T • A
G • C
A • T
5' G C T G T C
G A C A G C 3'
I
III
3' C G A C A G
C T G T C G 5'
T • A
C • G
G • C
T • A
A • T
C • G
5' 3'
IV
b
Branch migration
II
3' 5'
G • C
T • A
5' G C T G T C A G T A
T A C T G A C A G G 3'
I
III
3' C G A C A G T C G T
A C G A C T G T G G 5'
A • T
C • G
5' 3'
IV

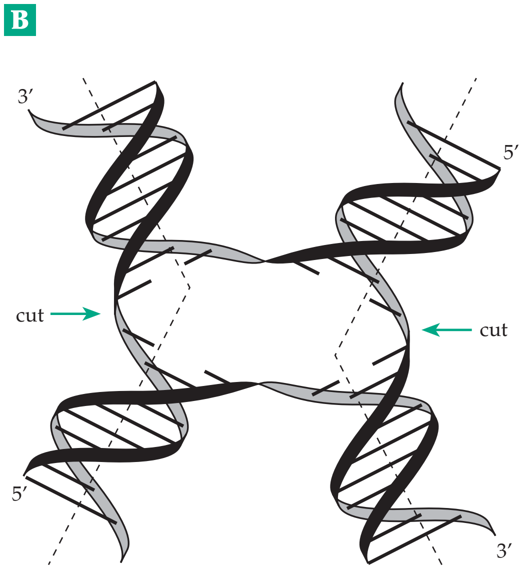

B
3′
5′
cut
cut
5′
3′

Figure 5-28 (A) Abbreviated reaction sequence for formation of a four-way Holliday junction between two homologous DNA duplexes. In step *a* strands are cut and rejoined with movement of the strands to a roughly antiparallel orientation. The resulting structure is thought to resemble that shown below the four-stranded representation. In step *b* branch migration takes place, separating the nonhomologous base pairs TA and CG and causing mismatched pairs which will be subject to repair. (B) Proposed three-dimensional structure (after drawing by Bennett and West).[292]

by the green double-headed arrow. If the strand break is then resealed, and the DNA strands are replicated, the boxed base pair in duplex I–II will have been transferred into a strand replicated from III–IV and genetic recombination will have been accomplished. Branch migration can occur over much longer distances than are indicated in this figure, so the alteration transferred may be far from the site of the initial cleavage and whole genes or groups of genes can be transferred. Recombination is considered in more detail in Chapter 27.

The Holliday junctions formed during recombination are mobile, but synthetic **immobile Holliday junctions** can be synthesized by using nonhomologous base sequences or by locking the junctions.[288] This has permitted careful physical study of these and other more elaborate synthetic junctions. With suitable choices of base sequences for the oligonucleotides from which they are made, such junctions will assemble spontaneously. Double-stranded DNA with immobile junctions is a very suitable construction material on a "nanochemical" scale. It has been assembled into knots, rings, cubes, and more complex polyhedra.[296–297b]

D. Ribonucleic Acids (RNA)

The best known forms of RNA are: (1) the long chains of messenger RNA (mRNA), which carry genetic messages copied from DNA to the ribosomes where proteins are made; (2) the much shorter transfer RNAs (tRNAs) which participate in reading the genetic code, correctly placing each amino acid in its sequence in the proteins; and (3) ribosomal RNAs (rRNA), which provide both structural material and a catalytic center for peptide bond formation. In addition there are numerous small RNAs that function in the splicing and editing[298 – 300] of mRNA, processing of tRNA precursors, methylation of ribosomal RNA,[301] transfer of proteins across membranes, and replication of DNA.[302] The genomes of many viruses consist of RNA. There are doubtless additional as yet undiscovered types of RNA.

Unlike DNA, which exists largely as double helices, the single chains of RNA can fold into complex forms containing many bulges and loops of the sort depicted in Fig. 5-9.[300,303,304] These loops are closed by double-stranded **stems** which have the A conformation. The B conformation is impossible because of the presence of the 2'-hydroxyl groups on the ribose rings in RNA. Even one ribonucleotide in a 10-nucleotide oligomer prevents formation of the B structure.[55,83] The 2'-OH groups not only keep the RNA in the A form but also engage in hydrogen-bond formation. Hydrogen bonds may form between the 2'-OH and the oxygen atom in the next ribose ring in the 3' → 5' direction. The –OH groups also hydrogen bond to water molecules which form a network within the minor groove.[305,306] These bound water molecules, in turn, can bond to associated protein and to other atoms of the complex loops found in RNA molecules. The 2'-OH groups also act as ligands for divalent metal ions in some tRNAs and in some RNA catalytic sites. Transient **hybrid DNA–RNA double helices** also exist within cells and they too usually have the overall shape of A-DNA.[55,307–309] However, the minor groove is intermediate in width between that expected for the A and B forms.

1. RNA Loops and Turns

Like polypeptides, polynucleotide chains have preferred ways of bending or turning. The loops at the ends of the hairpin turns of RNA molecules sometimes consist of a trinucleotide such as UUU,[304] but are usually larger. In tRNA there are typically seven bases that do not participate in regular Watson–Crick pairing (see Figs. 5-30, 5-31). The tetranucleotide **UUCG** is frequently present, and the sequence 5'-GGAC**UUCG**GUCC forms an unusually stable hairpin.[85,310] Other tetranucleotides, such as UGAA,[311] CCCG (also found in DNA loops),[312] GCAA,[313] and GAAA,[314] occur often. The latter are members of a larger group of loop structures with the consensus sequence GNRA, where N is any nucleotide and R is a purine.[315,315a] These sequences are very common in highly folded structures of ribosomal RNAs. Until recently high-resolution X-ray structures were available for only a few tRNAs and ribozymes. To help remedy this deficiency the structures of a great variety of oligonucleotide stem/loop (hairpin) structures are being determined, most by NMR spectroscopy.[304,316,317] The sharp turns in the loops involve mostly rotation about the two torsion angles around the phosphorus atom of the third, from the 5' end, of the seven nucleotides. Base bulges on stems (Fig. 5-9) not only introduce kinks and bends in RNA stems[169] but also provide well-defined hydrogen-bonded binding sites for proteins. Numerous branched three-way and more complex junctions provide other important motifs in folded RNA.[318]

Among the new RNA structures are those of RNA–antisense RNA pairs in "kissing" hairpin complexes.[87,319] Another interesting complex folding pattern in RNA is the **pseudoknot**, a structural feature that has been identified in many RNA sequences.[85,320–327] A pseudoknot can be formed if nucleotide sequences favorable to formation of two short RNA stems are overlapped as shown in Fig. 5-29. After stem 2 in this drawing is formed (step *a*) additional base pairing can lead to formation of stem 1 (step *b*). The base pairs of the two stems can stack coaxially to form the pseudoknot (step *c*).

For the formation of base-paired stems the RNA must contain antiparallel sequences that allow Watson

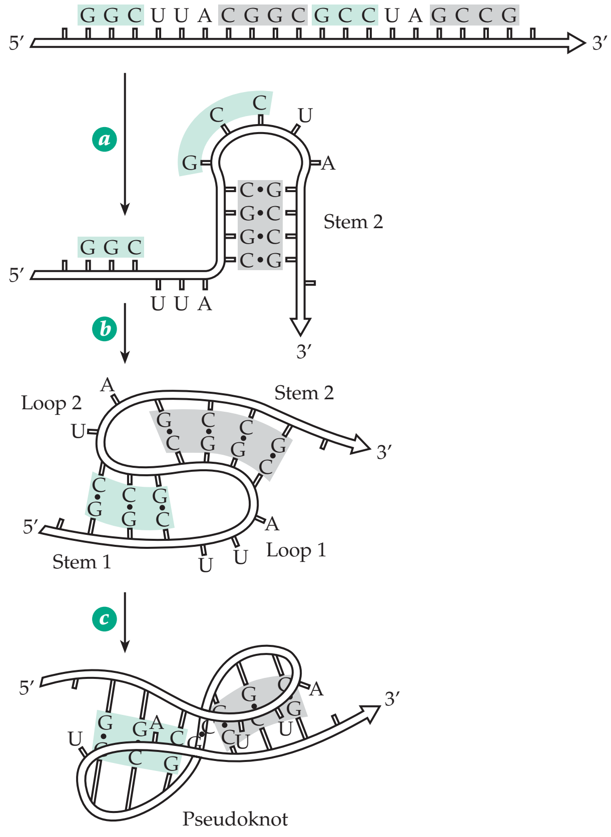


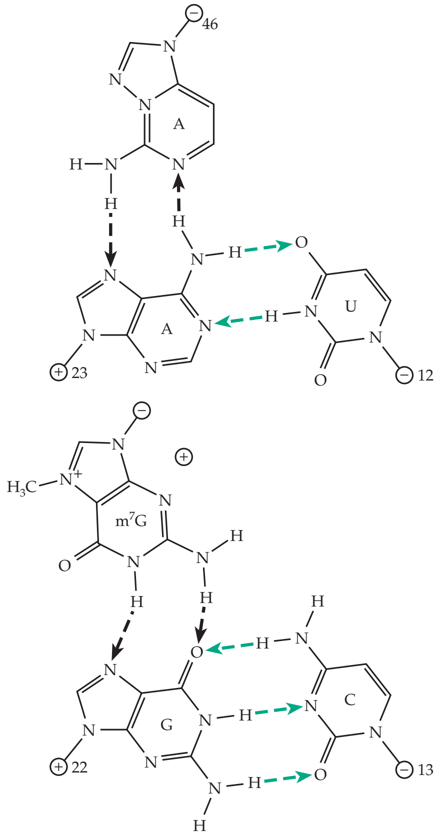
Figure 5-29 Formation of a pseudoknot in an RNA chain. After Puglisi *et al.*[325]

–Crick or wobble (GU) base pairing in the stems. This requires certain relationships in the sequences of the DNA in the genes that encode these molecules as discussed in Section E,3. Because of the base-pairing requirements, some of the bases in the stems protrude as bulges and fail to form pairs. Often, it is possible to find more than one reasonable structure, each having some bases unpaired.[328]

2. Transfer RNAs

In all tRNAs the bases can be paired to form "cloverleaf" structures with three hairpin loops and sometimes a fourth as is indicated in Fig. 5-30.[329 – 331] This structure can be folded into the L-shape shown in Fig. 5-31. The structure of a phenylalanine-carrying tRNA of yeast, the first tRNA whose structure was determined to atomic resolution by X-ray diffraction, is shown.[170,332–334] An aspartic acid-specific tRNA from yeast,[335] and an *E. coli* chain-initiating tRNA, which places *N*-formylmethionine into the N-terminal position of proteins,[336,337] have similar structures. These molecules are irregular bodies as complex in conformation as globular proteins. Numerous NMR studies show that the basic structure is conserved in all tRNAs. However, animal mitochondrial tRNAs often lack some of the usual stem-loop "arms" as well as the invariant nucleotides in the dihydrouridine and TψC loops (Fig. 5-30).[307,338,339] At the bottom of the structure as shown in Figs. 5-30 and 5-31 is the **anticodon**, a triplet of bases having the correct structures to permit pairing with the three bases of the codon specifying a particular amino acid (see Table 5-5), in this case phenylalanine.

While tRNAs consist largely of loops and stems containing Watson–Crick base pairs, they also contain Hoogsteen pairs, wobble pairs, and triplets such as the following.



The first of these contains a Watson–Crick A•U pair with a second A bound to form a reversed Hoogsteen A∗A pair. The second contains a G•C Watson–Crick pair with a Hoogsteen N-7-guanine∗G (m^7G∗G) pair. Among the complex base associations present in tRNAs are some that also involve hydrogen bonding to the 2'-hydroxyl groups of ribose rings and to at least one of the phosphate groups. There are over 100 internal hydrogen bonds, a large proportion of which are relatively invariant among the known tRNAs.

3. Properties and Structures of Some Other RNA Molecules

A few specialized RNA molecules are listed in

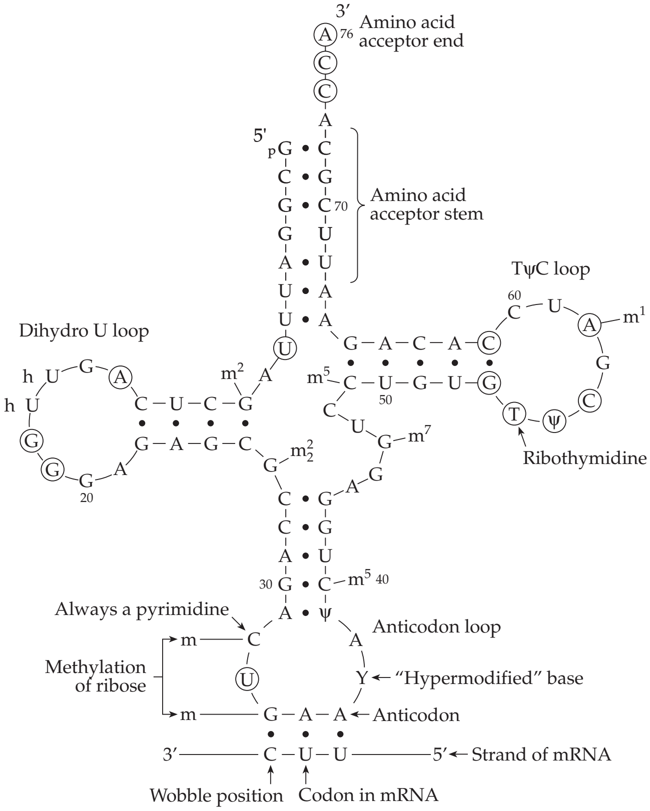


Figure 5-30 Schematic cloverleaf structure of a phenylalanine-specific transfer RNA (tRNA^{Phe}) of yeast. The dots represent pairs or triplets of hydrogen bonds. Nucleosides common to almost all tRNA molecules are circled. Other features common to most tRNA molecules are also marked. The manner in which the anticodon may be matched to a codon of mRNA is indicated at the bottom.

Table 5-4. There are many others.[302] In addition, there are thousands of different mRNAs within a cell. The most abundant RNA molecules in cells are those of the ribosomes. Ribosomes consist of two elaborate RNA-protein subunits, a large subunit with sedimentation constant ~ 30S in bacteria and ~ 40S in eukaryotes. The small subunit contains 16S or 18S RNA and the large subunit 23S or 28S as well as smaller 5S and 5.8S RNAs (Table 5-4; Table 28-1). A proposed three-dimensional structure[341,342] of a bacterial 16S ribosomal RNA and the corresponding ribosomal subunit with its 21 proteins are shown in Fig. 5-32. It might seem impossible that the folding pattern of the RNA was deduced correctly before an X-ray structure was available. However, a **phylogenetic approach**, the comparison of nucleotide sequences among several species, suggested that the stem structures of rRNAs are highly conserved (see discussion in Chapter 29). This fact, together with a variety of other chemical

TABLE 5-4
Some Specialized RNA Molecules

Kind of RNA	Number of nucleotides
Transfer RNAs (Figs. 5-30, 5-31)	60 – 85
Ribosomal RNAs	
5S	~ 120
5.8S (rat)	158
16S (*E. coli*; Fig. 5-32)	1542
18S (rat)	1874
23S (*E. coli*)	2904
28S (rat)	4718
Telomerase guide RNA[a-d]	159 (*Tetrahymena*)
M1 RNA of Ribonuclease P[e,f]	350 – 410
	377 (*E. coli*)
Tetrahymena Intron ribozyme (Fig. 12-26)[g]	413
Viroid hammerhead ribozyme (Fig. 12-27)[h]	~ 55
Signal recognition particle 7S RNA[i]	295 (human)
Small nuclear RNAs[j,k]	65 – 1200
RNA-editing guide RNA[l]	~ 60
Thermotolerance factor (G8 RNA)[m]	~ 300
Viroid RNA (Fig. 28-19)	240 – 380
Virus MS2 genome (Chapter 29)	3569

[a] Greider, C. W., and Blackburn, E. H. (1989) *Nature (London)* **337**, 331–337

[b] Bhattacharyya, A., and Blackburn, E. H. (1994) *EMBO J.* **13**, 5721–5731

[c] Singer, M. S., and Gottschling, D. E. (1994) *Science* **266**, 404–409

[d] Feng, J., Funk, W. D., Wang, S.-S., Weinrich, S. L., Avilion, A. A., Chiu, C.-P., Adams, R. R., Chang, E., Allsopp, R. C., Yu, J., Le, S., West, M. D., Harley, C. B., Andrews, W. H., Greider, C. W., and Villeponteau, B. (1995) *Science* **269**, 1236–1241

[e] Stark, B. C., Kole, R., Bowman, E. J., and Altman, S. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3717–3721

[f] Mattsson, J. G., Svärd, S. G., and Kirsebom, L. A. (1994) *J. Mol. Biol.* **241**, 1–6

[g] Cech, T. R. (1987) *Science* **236**, 1532–1539

[h] Hertel, K. J., Herschlag, D., and Uhlenbeck, O. C. (1994) *Biochemistry* **33**, 3374–3385

[i] Li, W.-Y., Reddy, R., Henning, D., Epstein, P., and Busch, H. (1982) *J. Biol. Chem.* **257**, 5136–5142

[j] Maxwell, E. S., and Fournier, M. J. (1995) *Ann. Rev. Biochem.* **35**, 897–934

[k] Nicoloso, M., Qu, L.-H., Michot, B., and Bachellerie, J.-P. (1996) *J. Mol. Biol.* **260**, 178–195

[l] Kable, M. L., Seiwert, S. D., Heidmann, S., and Stuart, K. (1996) *Science* **273**, 1189–1195

[m] Fung, P. A., Gaertig, J., Gorovsky, M. A., and Hallberg, R. L. (1995) *Science* **268**, 1036–1039

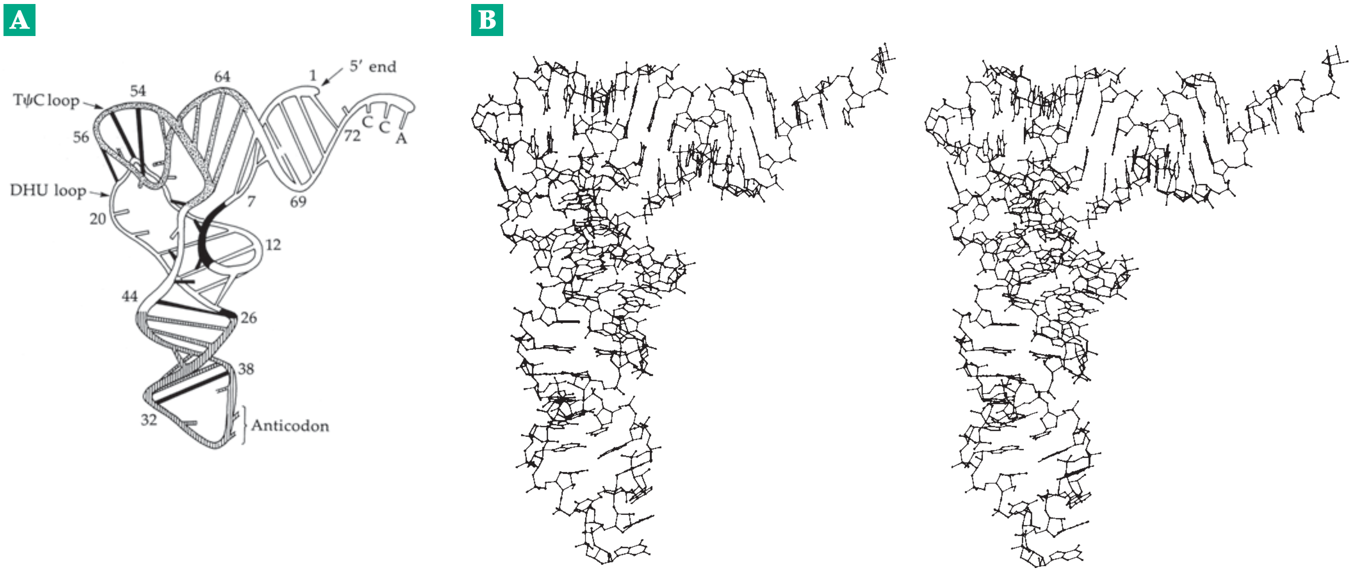


Figure 5-31 The three-dimensional structure of a phenylalanine-specific transfer RNA of yeast. (A) Perspective diagram of folding of polynucleotide chain. The ribose phosphate backbone is drawn as a continuous cylinder with bars to indicate hydrogen-bonded base pairs. The positions of single bases are indicated by rods which are intentionally shortened. The TψC arm is heavily stippled, and the anticodon arm is marked by vertical lines. Tertiary structure interactions are illustrated by black rods. Redrawn from Quigley and Rich.[340] (B) Stereoscopic view of the structure of yeast tRNA^{Phe} as revealed by X-ray crystallography. The acceptor stem with the protruding ACCA sequence at the 3' end is to the right. The anticodon GAA is at the bottom right side of the drawing. The guanine ring is clearly visible at the very bottom. The middle adenine of the anticodon is seen exactly edge-on as is the "hypermodified" base Y (see Fig. 5-33) which lies just above the anticodon. Its side chain is visible at the back of the drawing. Preceding the anticodon on the left (5' side) are two unpaired bases (C and U). The 2'-hydroxyl groups of the cytosine and of the guanosine in the anticodon are methylated. Moving far up the anticodon stem one can see two groups of base triplets which utilize both Watson–Crick and Hoogsteen types of base pairing. Drawing courtesy of Alexander Rich.

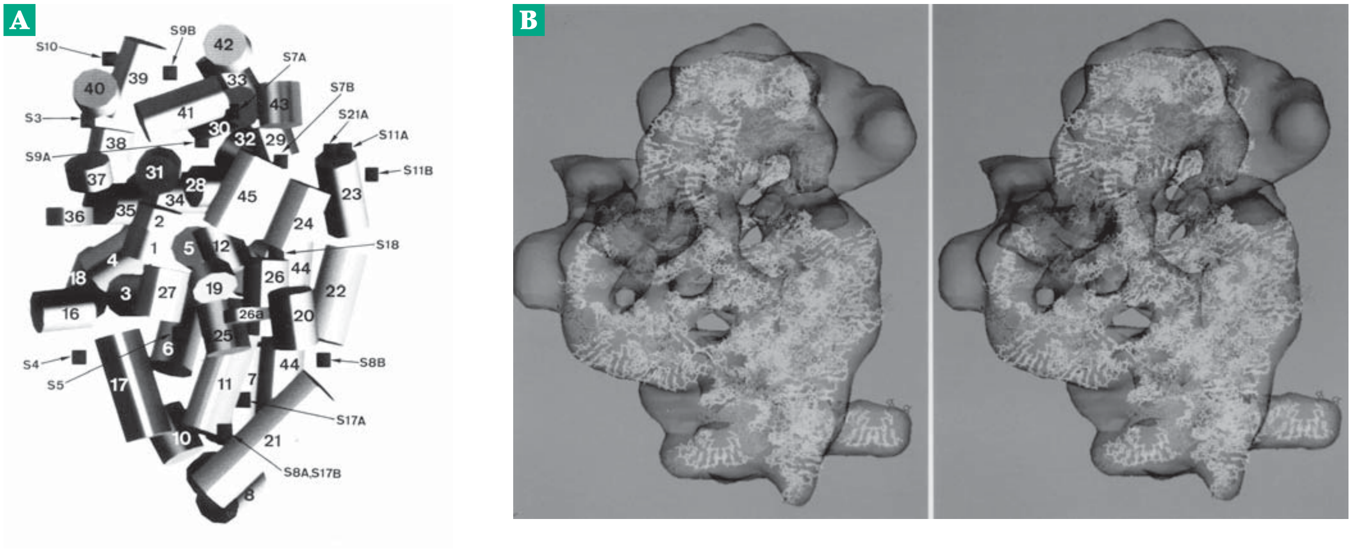


Figure 5-32 (A) A three-dimensional computer graphics model proposed by Brimacombe *et al.*[341] for the single chain of *E. coli* 16S ribosomal RNA. The helices are depicted as cylinders, which are all connected. The small dark squares denote the positions of artificially formed RNA–protein crosslinks, marked with the appropriate protein number. For proteins exhibiting more than one crosslink site (e.g., S17), the sites are denoted A or B, in each case A being the site nearer to the 5' terminus of the 16S RNA. (B) Stereoscopic view of tentative atomic model of 16S RNA in the 30S ribosomal subunit. The viewing direction is different from that in (A). From Mueller and Brimacombe.[342] Courtesy of Richard Brimacombe.

evidence,[341,343,344] allowed the prediction of the structure shown as well as many characteristics of the rRNA of the large subunit.[345] Culminating decades of effort[346–350] (Chapter 29), complete structures of bacterial ribosomes were established by 2000[342,351–354] and the peptidyl transferase center was identified as a ribozyme[355–357] (discussed in Chapter 12).

4. Modified Nucleotides

The picture of DNA or RNA as chains of only four kinds of nucleotides is not quite accurate. DNA contains a significant number of methylated bases among which are **5-methylcytosine** (5mC or m^{5}C) and 6-methyladenine (6mA). The former is regularly present in the nuclear DNA of higher animals and plants. In mammals from two to seven percent of the cytosine is methylated. It is likely that these methylated bases mark special points in the genetic blueprint. As is discussed in Chapter 27, methylation, which is accomplished after the synthesis of the polynucleotide, may block the expression of certain genes.[358] This appears to happen when one of the two X-chromosomes becomes inactivated in cells of females. Methylation may also be used to **imprint** certain genes, i.e. to mark them as coming from a specific one of the two parents.[359]

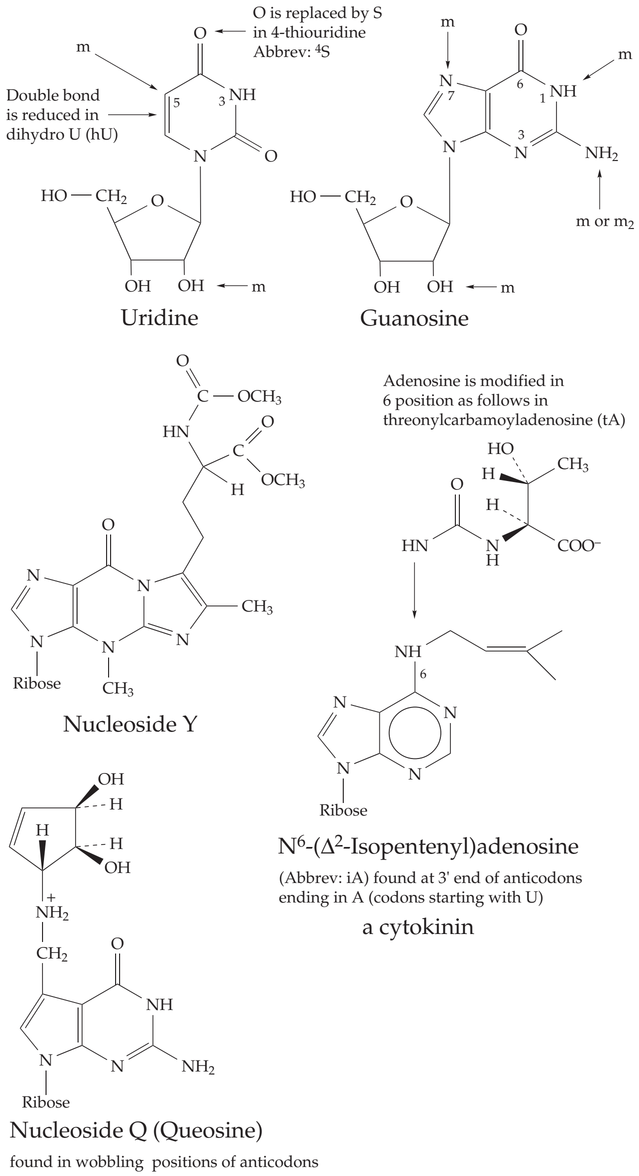
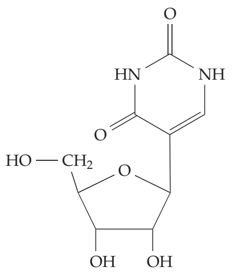

O is replaced by S
in 4-thiouridine
Abbrev: ^{4}S
m
Double bond
is reduced in
dihydro U (hU)
O
NH
N
O
5
3
HO — CH$_2$
OH
OH
m
Uridine
m
O
N
7
6
1
NH
3
N
N
NH$_2$
m
m or m$_2$
HO — CH$_2$
OH
OH
m
Guanosine
O
C — OCH$_3$
HN
C
O
OCH$_3$
H
O
N
N
N
N
N
CH$_3$
Ribose
CH$_3$
Nucleoside Y
Adenosine is modified in
6 position as follows in
threonylcarbamoyladenosine (tA)
HO
H
CH$_3$
O
H
HN
N
H
COO$^-$
NH
6
N
N
N
N
Ribose
N^6-(Δ^2-Isopentenyl)adenosine
(Abbrev: iA) found at 3' end of anticodons
ending in A (codons starting with U)
a cytokinin
OH
H
H
H
OH
NH$_2$
CH$_2$
O
NH
N
N
NH$_2$
Ribose
Nucleoside Q (Queosine)
found in wobbling positions of anticodons

Figure 5-33 Structures of some nucleosides containing modified bases and found in tRNA molecules. Positions where methylation may occur are designated m. Some other abbreviations are e (ethyl), ac (acetyl), and f (formyl).[25]

Another function of methylation is to protect DNA from attack by enzymes (restriction endonucleases) formed in response to invading viruses (Chapter 26). Some viruses, notably the bacteriophage of the T-even series that attack *E. coli* (Box 7-C), have developed their own protective devices. They contain **5-hydroxymethylcytosine** (HOMeC) in place of cytosine. The extra hydroxyl groups provided in this fashion often carry one or two glucose units in glycosidic linkage.[360] A bacteriophage attacking *Bacillus subtilis* substitutes hydroxymethyluracil for uracil and 5-dihydroxypentyluracil for thymine, and phage W14 of *Pseudomonas acidovorans* substitutes the 5-methyl of thymine with $-CH_2-NH-(CH_2)_4-NH_3^+$.[360,361]

The modifications carried out on RNA molecules are more varied and more extensive than those of DNA. Sixty or more modification reactions are known for tRNA, with the number and the extent of modification depending upon the species. Structures of some of the modified bases are indicated in Fig. 5-33. Uridine can be methylated either on the base or on the 2'-hydroxyl of the sugar. Methylation at the 5 position of uridine yields **ribothymidine**. Cytidine can be modified in the same positions. Reduction of the 5,6 double bond of uridine gives **dihydrouridine** (hU). Replacement of the oxygen at position 4 by sulfur gives **4-thiouridine** (^{4}sU). Positions in the guanosine structure that can be methylated are also indicated in Fig. 5-33. The symbol m is commonly used to designate methylation in nucleic acid bases; m$_2$ indicates dimethylation, e.g., 6,6-dimethyladenine is abbreviated m$_2{}^6$A.

A remarkable transformation is that of uridine into pseudouridine (ψ).


O
HN
NH
O
HO — CH$_2$
O
OH
OH

Pseudouridine (ψ)

BOX 5-C THE RNA WORLD

The discovery in the 1980s that RNA molecules often have catalytic properties and may serve as true enzymes (ribozymes; Chapter 12) stimulated new thinking about evolution. Although RNA catalysts are not as fast as the best enzymes they are able to catalyze a wide variety of different reactions. Could it be that in the early evolution of organisms RNA provided both the genetic material and catalysts? The "RNA world" would have been independent of both DNA and protein.[a,b] Later DNA could have been developed as a more stable coding molecule and proteins could have evolved as more efficient catalysts. Plausible reactions by which both cytosine and uracil could have arisen in drying ponds on early Earth have been demonstrated.[c]

A major objection to the RNA world is the lack of stability of ribose and the inability to demonstrate the nonenzymatic synthesis of ribose in significant amounts. Even if ribose were present, it would be largely in the pyranose ring forms. Initial formation of the 5-phosphate would be required to allow formation of a nucleotide with a furanose ring. These and other obstacles to the RNA world have led to the suggestion that some *other* genetic material preceded RNA and DNA.[d,e] One possibility is a peptide-like RNA analog.[e] A simple coding system could also have been used, e.g. one based on only two bases, such as C and G, instead of four.[f,g]

Perhaps it is more probable that formation of proteins *and* the present coding system evolved simultaneously? The major metabolic cycles (Chapter 10) could also have developed at the same time. RNAs could not have been the *first catalysts*. Hydrogen ions, hydroxyl ions, ammonium, cyanide, and other simple ions as well as amines and peptides could all have played a role in prebiotic chemistry. Another speculation suggests an "iron-sulfur world" in which organic materials would be formed on mineral surfaces through reactions involving reduction of bicarbonate by iron sulfide and H_2S.[h]

If the RNA world did exist, has it left us with any real clues? Benner *et al.* suggest that modern metabolism is a palimpsest of the RNA world, a parchment that has been inscribed two or more times, with previous texts imperfectly erased and therefore still partially legible.[i] If so, can we find a way to read the text of the ancient RNAs?

---

[a] Gesteland, R. F., Cech, T. R., and Atkins, J. F., eds. (1999) *The RNA World*, 2nd ed. Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York

[b] Orgel, L. E. (1994) *Sci. Am.* **271** (Oct), 77 – 83

[c] Robertson, M. P., and Miller, S. L. (1995) *Nature (London)* **375**, 772 – 774

[d] Larralde, R., Robertson, M. P., and Miller, S. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8158 – 8160

[e] Böhler, C., Nielsen, P. E., and Orgel, L. E. (1995) *Nature (London)* **376**, 578 – 581

[f] Sievers, D., and von Kiedrowski, G. (1994) *Nature (London)* **369**, 221 – 224

[g] Piccirilli, J. A. (1995) *Nature (London)* **376**, 548 – 549

[h] Maden, E. H. (1995) *Trends Biochem. Sci.* **20**, 337 – 341

[i] Benner, S. A., Ellington, A. D., and Tauer, A. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7054 – 7058

Pseudouridine is formed by enzymatic rearrangement of uridine in the original transcript (Eq. 28-3). It can form a base pair with adenine in the same manner as does uracil. Pseudouridine is found not only in tRNA but also in several places in both large and small ribosomal RNA subunits. For example, it is present at position 516 in the *E. coli* 16S RNA,[364] at a specific position in the 23S RNA, and at many more locations in eukaryotic rRNA.

The bases called "Y" and "Q" are highly modified guanines (Fig. 5-33). Q is found at the 5' end of some anticodons in the "wobble position" (see Fig. 5-30). Two **hypermodified** adenosines are also shown in Fig. 5-33. The N^6-isopentenyladenosine is found at the 3' end of the anticodons that pair with codons starting with U. This compound is also a plant hormone, a **cytokinin** (Chapter 30). Another highly modified purine, threonylcarbamoyladenine, occurs adjacent to the end of anticodons pairing with codons starting with A. The function of these hypermodified bases is uncertain, but they appear to promote proper binding to ribosomes. The modifications are often not absolutely essential for function.

Another source of modified bases in both DNA and RNA is spontaneous or "accidental" alteration. Nucleic acids encounter many highly reactive and mutagenic materials including hydroxyl radicals, formed from O_2, and are able to convert guanine rings into 7,8-dihydro-8-oxoguanine.[362] Other reactive and carcinogenic compounds can form adducts with nucleic acid bases.[363] See Eq. 5-18 and also Chapter 27.

5. RNA Aptamers

Ellington and Szostak[365] synthesized a random "pool" of ~10^{15} different oligodeoxyribonucleotides, each ~100 nucleotides in length. They "amplified" these using the polymerase chain reaction (PCR; Section H,6) and prepared a mixture of the corresponding RNAs by *in vitro* transcription. From the ~10^{13} different sequences still present they selected individual oligonucleotides by affinity chromatography on columns that contained well-defined immobilized ligands such as organic dyes. They called the selected RNAs **aptamers**. Their approach is being used to find RNA sequences that bind to such ligands as ATP, FMN,[366] the bronchodilator theophylline,[367] aminoglycoside antibiotics,[368] arginine,[369] etc.[370] Many of the selected aptamers bind their ligands very tightly and studying them may shed light on interactions of RNA with proteins and on the catalytic activities of RNA, which are discussed in Chapter 12.

E. The Genetic Code

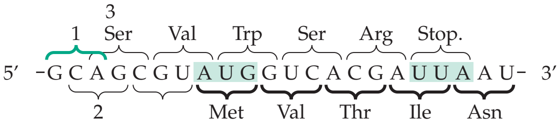
The general nature of the genetic code was suggested by the structure of DNA. Both DNA and proteins are linear polymers. Thus, it was logical to suppose that the sequence of the bases in DNA codes for the sequence of amino acids. There are only four bases in DNA but 20 different amino acids in proteins at the time of their synthesis. It is obvious that each amino acid must be specified by some combination of more than one base. While 16 pairs of bases are possible, this is still too few to specify 20 different amino acids. Therefore, it appeared that at least a triplet group of three nucleotides would be required to code for one amino acid.[371] Sixty-four (4^3) such triplet **codons** exist, as is indicated in Tables 5-5 and 5-6.

Simplicity argues that the genetic blueprint specifying amino acid sequences in proteins should consist of consecutive, nonoverlapping triplets. This assumption turned out to be correct, as is illustrated by the DNA sequence for a gene shown in Fig. 5-5. In addition to the codons that determine the sequence of amino acids in the protein, there are **stop codons** that tell the ribosomal machinery when to terminate the polypeptide chain. One methionine codon serves as an **initiation codon** that marks the beginning of a polypeptide sequence. One of the valine codons sometimes functions in the same way.

How does a cell read the code? This question is dealt with in detail in Chapters 28 and 29. A key step is the positioning of each amino acid on the ribosome in proper sequence. This is accomplished by the pairing of codons of messenger RNA with the anticodons of the appropriate transfer RNA molecules as is indicated at the bottom of Fig. 5-30. Each tRNA carries the appropriate "activated" amino acid at its 3' end ready to be inserted into the growing peptide.

1. The "Reading Frames"

It is immediately obvious that there are three ways of reading the genetic code in mRNA depending upon which nucleotide is used to start each codon. For example, in the following mRNA sequence either codons GCA, CAG, or AGC could be selected as first.



These codons define the three reading frames or phases in which the code may be read. Here, the term *frame*

TABLE 5-5
The Genetic Code[a]

Amino acid	Codons	Total number of codons
Alanine	GCX	4
Arginine	CGX, AGA, AGG	6
Asparagine	AAU, AAC	2
Aspartic acid	GAU, GAC	2
Cysteine	UGU, UGC	2
Glutamic acid	GAA, GAG	2
Glutamine	CAA, CAG	2
Glycine	GGX	4
Histidine	CAU, CAG	2
Isoleucine	AAU, AUC, AUA	3
Leucine	UUA, UUG, CUX	6
Lysine	AAA, AUG	2
Methionine (also initiation codon)	AUG	1
Phenylalanine	UUU, UUC	2
Proline	CCX	4
Serine	UCX, AGU, AGC	6
Threonine	ACX	4
Tryptophan	UGG	1
Tyrosine	UAU, UAC	2
Valine (GUG is sometimes an initiation codon)	GUX	4
Termination	UAA (*ochre*) UAG (*amber*) UGA	3
Total		64

[a] The codons for each amino acid are given in terms of the sequence of bases in messenger RNA. From left to right, the sequence is from the 5' end to the 3' end. The symbol X stands for any one of the four RNA bases. Thus, each codon symbol containing X represents a group of four codons.

does not designate a single codon, although frame does designate a single exposure in a motion picture film. **Reading frame** designates which of the three possible sets of codons we are using. In the foregoing sequence the codons in reading frames 2 and 3 are labeled. Reading frame 2 contains the initiation codon AUG (shaded) which could mark the beginning of an encoded protein sequence. Reading frame 3 contains a termination (stop) codon which, when the mRNA transcript is read by ribosomes, will terminate polypeptide synthesis. It may be in the position shown but not have any real function. However, it could represent the end of the coding sequence that is marked if genes for the two proteins overlap a little at the ends, a situation that actually occurs in nature.

A reading frame in a specified part of a DNA sequence is said to be **"open"** if there is an initiation codon preceded by suitable regulatory signals (an **operator** region). This means that it *could* encode a protein. The reading frame of a sequence is open until the next termination codon. Recently another usage has appeared. Many writers refer to an **open reading frame** as a segment of DNA in which any one of the three reading frames is open. Another complexity in the reading of genetic messages arises because splicing of RNA may sometimes cause a shift in the reading frame. For example, a mRNA being transcribed from the sequence in reading frame 2 in the foregoing example may skip over a nucleotide part of the time to form an RNA in which the first part is encoded by reading frame 2 and the second by reading frame 3 (a + 1 frameshift). Alternatively, a nucleotide could be read twice with a – 1 frameshift with the sequence of reading frame 1 for the latter part of the mRNA. Frameshifts can also occur during protein synthesis as the mRNA is being read.

TABLE 5-6
The Sixty-Four Codons of the Genetic Code

5′–OH Terminal base	Middle base				3′–OH Terminal base
	U(T)	C	A	G	
U(T)	Phe	Ser	Tyr	Cys	U(T)
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Term[c]	Term[d]	A
	Leu	Ser	Term	Trp	G
C	Leu	Pro	His	Arg	U(T)
	Leu	Pro	His	Arg	C
	Leu[a]	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U(T)
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met[b]	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U(T)
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val[b]	Ala	Glu	Gly	G

[a] The codon CUA (CTA) encodes threonine and the codon AUA (ATA) methionine in mammalian mitochondria.

[b] Initiation codons. The methionine codon AUG is the most common starting point for translation of a genetic message but GUG can also serve. In such cases it codes for methionine rather than valine.

[c] The "termination codon" UAA (TAA) encodes glutamine in *Tetrahymena*.

[d] The termination codon UGA (TGA) encodes tryptophan in mitochondria and selenocysteine in some contexts in nuclear genes.

In the present example we have examined the sequence in mRNA. In the DNA there are two strands. One is the **coding strand** (also called the nontranscribing or nontranscribed strand), which has a sequence that corresponds to that in the mRNA and the one that is given in Fig. 5-4. The second antiparallel and complementary strand can be called the **template strand** or the noncoding, transcribing, or transcribed strand.[372] The mRNA that is formed is sometimes referred to as a **sense strand**. The complementary mRNA, which corresponds in sequence to the noncoding strand of DNA, is usually called **antisense RNA**.

2. Variations

Is the genetic code "universal" or does it vary from one organism to another? Studies with bacteria, viruses, and higher organisms including humans have convinced us that the code is basically the same for all oganisms. However, there are some variations. For example, in mitochondria of both humans and yeast the codon TGA is not a termination codon but represents tryptophan. In mammalian mitochondria CTA represents threonine rather than leucine, and ATA encodes methionine instead of isoleucine. These differences in the code are related to the fact that mitochondria contain their own piece of DNA. It encodes not only several proteins but also tRNA molecules whose anticodon structures are altered to accommodate the changed meanings of the codons of the mitochondrial DNA and mRNA.[373,374]

Variations in the code for cytoplasmic proteins have been found. In *Tetrahymena* and other ciliates the codon TAA represents glutamine rather than being a termination codon.[375] A few proteins, including some in the human body, contain **selenocysteine**, the selenium-containing analog of cysteine. Selenocysteine is encoded by termination codon TGA. See Chapter 29 for details. However, even though TGA is occasionally used in this way, it serves as a termination codon for most proteins within the same cells.[376] Thus, the *context* in which the codon TGA occurs determines how it is read by the ribosomal machinery.

3. Palindromes and Other Hidden Messages

The sentence "Madam, I'm Adam" reads the same either forward or backward. Such sentences, known as **palindromes**,[377] are infrequent in the English language. However, most DNA contains many palindromes, sequences of base pairs that read the same in forward and reverse directions. Consider, for example, the gene that specifies the sequence of nucleotides for the tRNA molecules of Fig. 5-30. It is a double-stranded DNA segment in which one strand has a sequence identical to that of Fig. 5-30 except for the substitution of T for U and for ψ (pseudouridine) and for the lack of methylation and other base modifications. The second strand is the exact complement. Figure 5-34 shows the part of this gene (residues 49–76) that corresponds to the 3' end of the tRNA molecule. This DNA segment could exist in a second conformation having a loop on each side of the molecule (Fig. 5-34). The stems of the two loops in this **cruciform** conformation are identical and symmetrically disposed around the center of the molecule. If we overlook the seven nucleotides in the center of the loop, the message in the stems reads the same in both directions, as is indicated by the green arrows.

Palindromes are often imperfect as is the one shown in Fig. 5-34. Here the two stems in the cruciform structure are related by an exact twofold rotational symmetry but the loops at the ends of the stems are not. Unpaired bases may bulge at various points in double-stranded stems of longer palindromes. These imperfect palindromes in the DNA are responsible for much of the tertiary structure of the various kinds of RNA. The tertiary structure, in turn, often determines the interaction of the RNA with enzymes and other proteins.

Special properties may be observed for palindromes containing homopurine tracts in one strand, and therefore homopyrimidine tracts in the other. If two identical palindromes of this type occur close together it is possible that the pyrimidine-containing strand of one can join with a hairpin loop of the other to form a triplex base structure. A related triplex structure may be formed when inverted repeat sequences occur within a homopyrimidine tract in one chain of DNA (Fig. 5-34B).[378] These have been called **mirror-repeats** or **H-palindromes** to distinguish them from true palindromes.[237,379] Each base triplex structure contains one set of Watson–Crick hydrogen bonds and one set of Hoogsteen hydrogen bonds. The triplexes of H-DNA are all either TAT or CGC$^+$, where one C is protonated (see also Fig. 5-24).

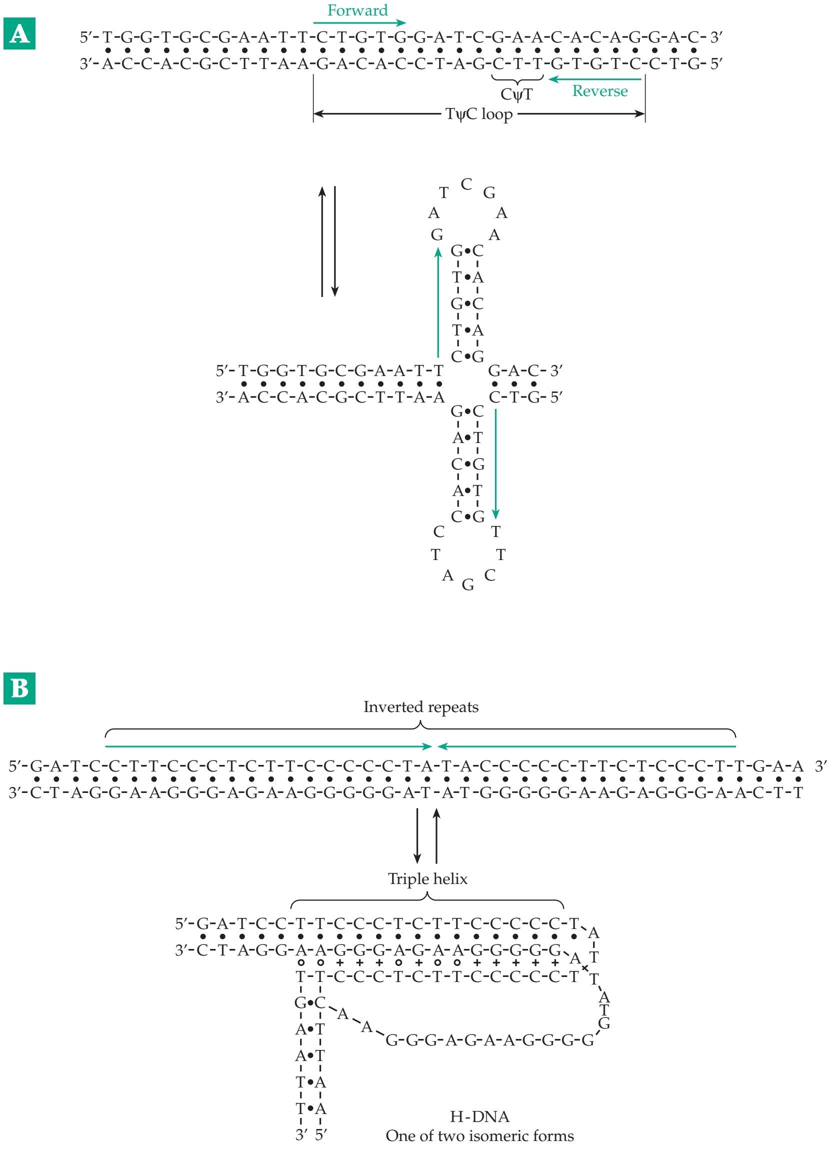

A
Forward
5′-T-G-G-T-G-C-G-A-A-T-T-C-T-G-T-G-G-A-T-C-G-A-A-C-A-C-A-G-G-A-C-3′
3′-A-C-C-A-C-G-C-T-T-A-A-G-A-C-A-C-C-T-A-G-C-T-T-G-T-G-T-C-C-T-G-5′
CψT
Reverse
TψC loop
B
Inverted repeats
5′-G-A-T-C-C-T-T-C-C-C-T-C-T-T-C-C-C-C-C-T-A-T-A-C-C-C-C-C-T-T-C-T-C-C-C-T-T-G-A-A 3′
3′-C-T-A-G-G-A-A-G-G-G-A-G-A-A-G-G-G-G-G-A-T-A-T-G-G-G-G-G-A-A-G-A-G-G-G-A-A-C-T-T
Triple helix
H-DNA
One of two isomeric forms

Figure 5-34 (A) Two conformations of a segment of the yeast phenylalanine tRNA gene. The segment shown codes for the 3' end of the tRNA molecule shown in Fig. 5-30, including the TψC loop. (B) Formation of H-DNA (Fig. 5-24) proposed for a sequence in plasmid pGG32. The major element of the structure is the triplex, which is formed from the Watson–Crick duplex (•) associated with the homopyrimidine loop through Hoogsteen base pairing (∘, +). One of the two possible "isomeric" forms is shown. See Mirkin *et al.*[378]

Segments of H-DNA might block transcription, and proteins that bind to the triplex H-DNA may be involved in transcriptional regulation.[379]

Until rather recently there had been little to indicate that DNA actually assumes cruciform conformations in cells. However, strong experimental evidence suggests that some cruciform structures do form naturally.[380] Their formation from palindromic DNA [like the formation of Z-DNA from (G + C)-rich sequences] is a way of relieving torsional strain induced by supercoiling. Whether or not cruciform structures occur frequently within cells, there is no doubt that palindromic sequences are of great importance in the interaction of nucleic acids with symmetric dimeric and tetrameric protein molecules such as the gene repressor protein shown in Fig. 5-35.[381–383]

DNA contains numerous other protein binding sites which are not palindromes but whose sequences represent additional encoded information. The RNA transcripts likewise contain sequences that direct the catalytic machinery involved in splicing, that bind to ribosomal proteins, that control rates of transcription, and that cause termination of transcription.

4. The Base Composition of DNA and RNA

The nucleotide composition of DNA is surprisingly variable. The sum of the percentage cytosine plus the percentage guanine (C + G) for bacteria varies from 22 to 74%. That for *E. coli* is 51.7%. Among eukaryotic organisms, the range is somewhat narrower (28 to 58%; for humans, 39.7%). The fact that bacterial DNA molecules are more varied than those of higher organisms is not surprising. The prokaryotes have evolved for just as many more million of years as have we. Perhaps because of their simpler structure and rapid rate of division, nature has done more experimentation with genetic changes in bacteria than in people.

Comparisons of the C + G content of organisms have been used as a basis for establishing genetic relatedness.[384] However, since thymine is especially susceptible to photochemical alteration by ultraviolet light, bacteria with a high (C + G) content may have evolved in environments subject to strong sunlight or high temperatures, whereas those with a low (C + G) content have developed in more protected locations.[385,386]

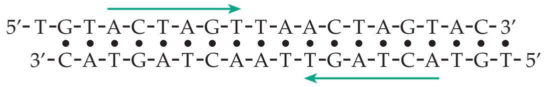
F. Interaction of Nucleic Acids with Proteins

Most functions of DNA and RNA are dependent on proteins. Cells contain a vast array of polypeptides that bind to nucleic acids in many specific ways.[387–390] We have seen (Fig. 5-21) that the histones hold supercoiled DNA in the form of nucleosomes[391] in eukaryotic nuclei and similar proteins bind to and protect the double-stranded DNA of bacteria.[205,392] Sperm cells "package" a large amount of DNA into a small space with the help of small arginine-rich proteins called **protamines.**[393] Cells always have some single-stranded DNA segments as well as single-stranded DNA binding (**SSB**) proteins. Among the latter are proteins from *E. coli*[394,395] and from viruses.[396–399] Specialized proteins bind to DNA sequences in telomeres[38,272] and centromeres.[400] A large number of proteins interact with RNA in ribosomes, spliceosomes, and other complexes.

A host of enzymes, which are described elsewhere in the book, act on DNA and RNA. They include hydrolytic nucleases, methyltransferases, polymerases, topoisomerases, and enzymes involved in repair of damaged DNA and in modifications of either DNA or RNA. While most of these enzymes are apparently proteins, a surprising number are **ribozymes**, which consist of RNA or are RNA–protein complexes in which the RNA has catalytic activity.

1. The Helix–Turn–Helix Motif

Much current interest in DNA–protein interactions is focused on regulatory processes. In prokaryotes the initiation of transcription of a large fraction of the genes is blocked by the binding of proteins known as **repressors**. While their structures are varied, one large group of repressors have DNA-binding domains with a similar helix–turn–helix architectural motif. They bind with high affinity to specific control regions which contain palindromic DNA sequences such as the following one, which defines a binding site for the *E. coli* **trp** (tryptophan) **repressor**:


5′-T-G-T-A-C-T-A-G-T-T-A-A-C-T-A-G-T-A-C-3′
3′-C-A-T-G-A-T-C-A-A-T-T-G-A-T-C-A-T-G-T-5′

These repressor proteins form dimers joined through a rigid central domain with a pair of arms, each containing two helices that form the helix–turn–helix motif. One helix of each pair fits into the major groove of the DNA helix and forms a "reading head" that carries a specific arrangement of amino acid side chains able to locate the symmetric nucleotide sequences flanking the twofold axis of the palindrome. In the case of the *trp* repressor, these are the ACTAGT hexanucleotides (marked by the arrows in the preceding structure). The interaction is depicted in Fig. 5-35.[401–405] The repressor has a high affinity for the DNA only if one molecule of L-tryptophan is bound to each subunit at a specific site near the DNA helix. Binding of the tryptophan causes a conformational change, which is pictured in Fig. 5-35. If tryptophan is absent the repressor binds to the palindromic DNA only weakly and transcription of genes needed for tryptophan biosynthesis occurs freely. However, if tryptophan accumulates within the cell it binds to the repressor molecules causing them to bind firmly to DNA and prevent transcription. There are at least three of these palindromic sequences in *E. coli*, each one regulating a set of genes (**operons**) involved in tryptophan synthesis. In contrast to the effect of tryptophan, the tryptophan analog indole-3-propionate, which lacks the amino group of tryptophan, *derepresses* the same operons. It also binds to the *trp* repressor, but with the indole ring flipped over by 180° so that its carboxylate group contacts the phosphate groups of the DNA repelling them through both electrostatic and steric effects.[406,407] The picture in Fig. 5-35 is an oversimplification. In fact, at some binding sites (operator sites) more than one dimeric repressor binds in a tandem fashion.[405]

The helix–turn–helix motif is also found in many other proteins. One of these is the bacterial **lac** (lactose) **repressor** which controls the *lac* operon and for which

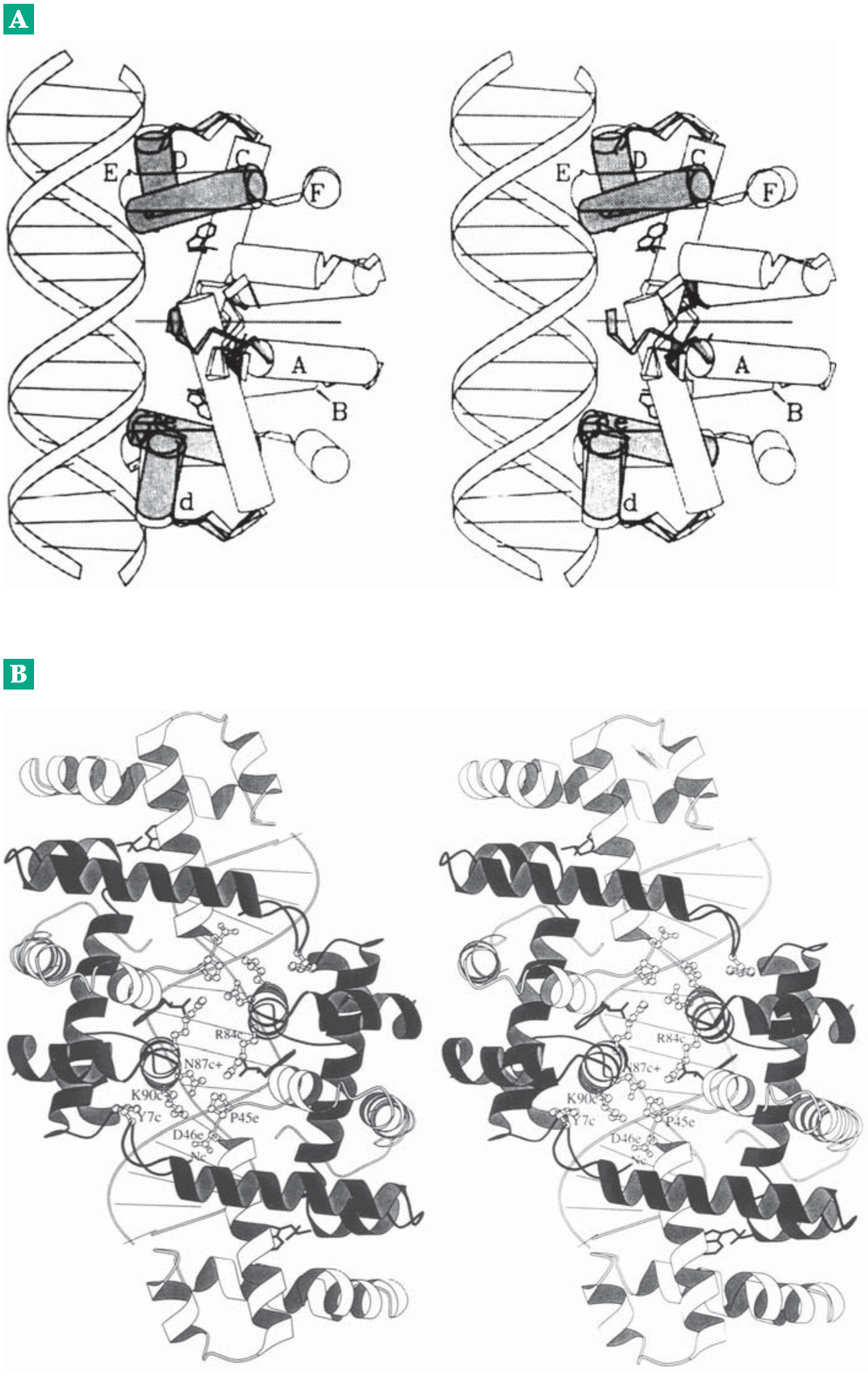


Figure 5-35 Stereoscopic drawings illustrating the binding of a dimeric molecule of the Trp repressor protein to a palindromic sequence in DNA. (A) Schematic view showing structures of the aporepressor (partly shaded gray) and the holorepressor with bound tryptophan (unshaded) are superimposed. Cylinders represent the α helices in (B). From Zhang *et al.*[402] (B) MolScript ribbon diagram with a few side chains that interact with the DNA shown. Two tandemly bound dimeric repressor molecules are shown. Two bound molecules of tryptophan are visible in each dimer. The DNA is drawn as a double helix with lines representing the base pairs. From Lawson and Carey.[405]

the terms operator, promoter, repressor, and operon were first introduced (Chapter 28).[408–410] Some bacterial viruses, such as **phage lambda (λ)** of *E. coli*, encode repressors that allow the virus to reside in the bacteria without immediately destroying them. The λ repressor and other closely related proteins also utilize the helix–turn–helix motif[411–415] as do some proteins that *activate* transcription (see Chapter 28).

2. Other DNA-Binding Motifs

Nobody knows how many different DNA-binding structures may be discovered. However, most of those that are designed to recognize specific DNA sequences have some part that fits into the major groove of B-DNA.[416] Usually the DNA structure must be in a specific form: B-, A- or Z-, but it may sometimes be bent or distorted. The DNA recognition motifs in the proteins may consist of helices, β strands, or loops.

Leucine zipper proteins. Several transcription factors have the leucine zipper structure, which has been described in Chapter 2 and was illustrated there by the structure of transcription factor Max (Fig. 2-21).[417] Related structures include those of the transcription activators c-Jun and c-Fos[418] and of the yeast transcription factor GCN4.[419,420] The latter is a dimeric protein in which the C-terminal halves of the two monomers form the helices of the leucine zipper. The helices, which are continuous for over 60 residues, fan out to interact with the DNA double helix as in Fig. 2-21. A 19-residue basic domain of each protein helix crosses the major groove of the DNA with side chain groups interacting with the DNA as is illustrated in Fig. 5-36.

Zinc fingers. Another large group of transcription factors contain a bound zinc ion in a "finger" motif. The Zn^{2+} is held by two cysteine $-S^-$ groups present on a loop of β structure and by two imidazole groups present on an α helix (Fig. 5-37).[421–424] Amino acid sequences found in many proteins that regulate transcription tend to form zinc fingers and to interact with DNA. These were recognized first in the transcription factor TFIIIA from *Xenopus laevis* in 1985.[421,425,426] This protein contains a zinc finger motif repeated nine times. X-ray structures, one of which is shown in Fig. 5-38, are known for proteins with three[423] and five[427] zinc fingers. (See Chapter 28.)

Beta ribbons. An antiparallel double-stranded β ribbon can fit into the major groove of DNA and form

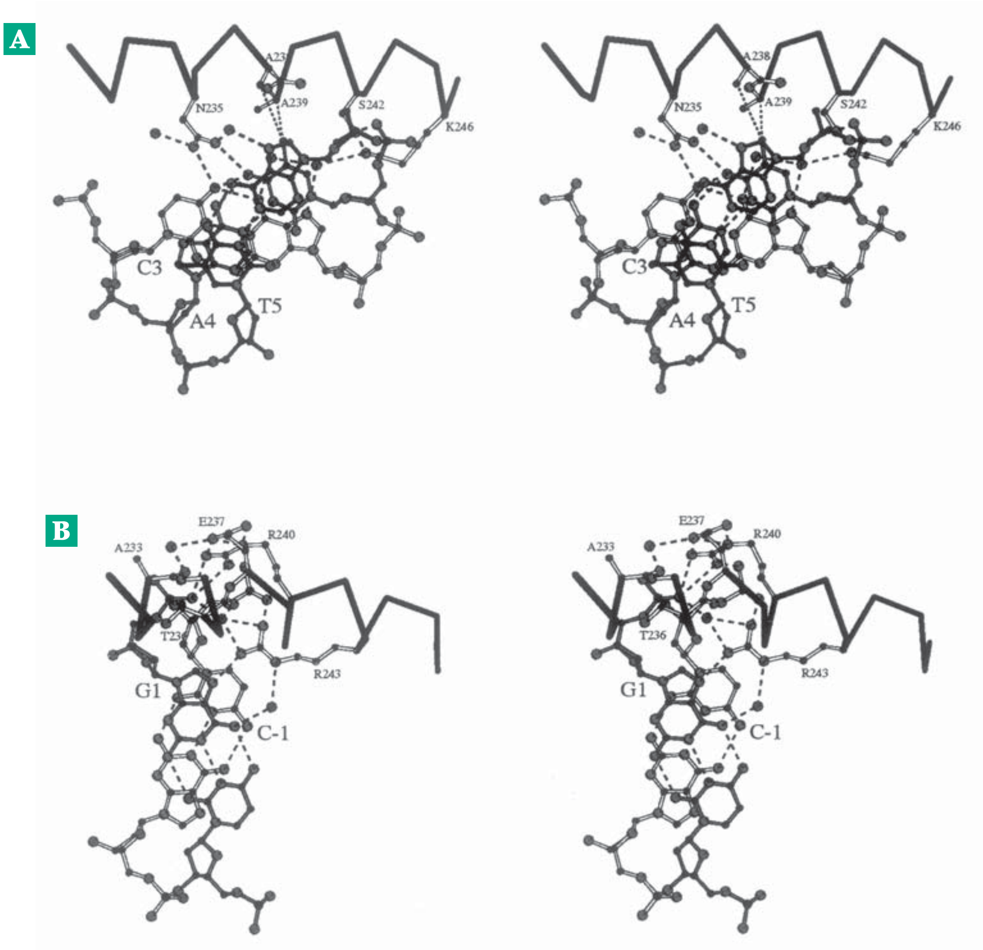


Figure 5-36 Stereoscopic diagrams showing some of the interactions between an N-terminal helical domain of the yeast transcription factor GCN4-bZIP, a leucine zipper protein, and a specific palindromic DNA binding site:

−5 −1 1 5
5′ –A T G A C G T C A T

(The bases are numbered outward from the central C and G.) The small solid spheres are water molecules. Notice the water mediated interactions of the basic arginine and lysine side chains with the nucleic acid bases and also the interaction of R240 and R243 (in B) with a backbone phosphate. The overall structure of the protein is similar to that of another leucine zipper shown in Fig. 2-21. From Keller *et al.*[419] Drawings courtesy of Timothy J. Richmond.

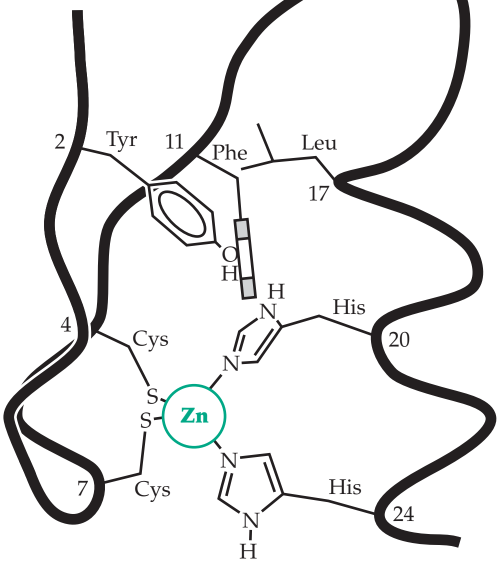

2
Tyr
11
Phe
Leu
17
O
H
H
N
His
4
Cys
20
N
S
S
Zn
N
7
Cys
His
24
N
H

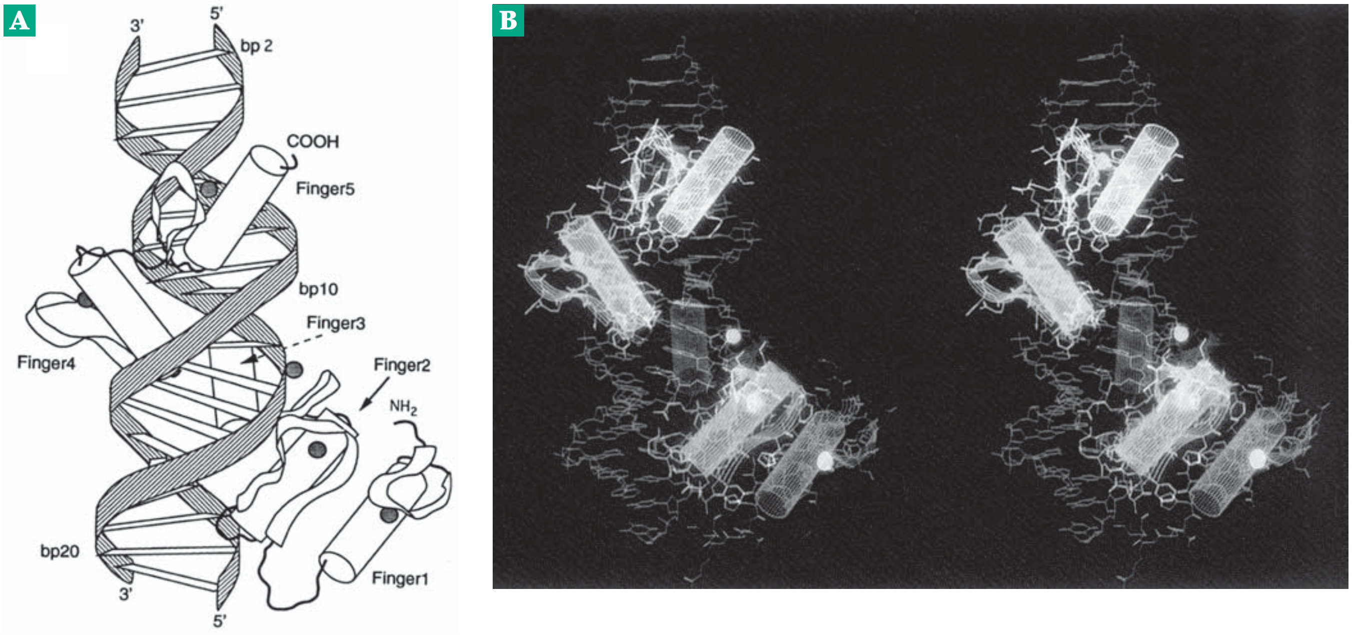
Figure 5-37 Three-dimensional structure of a zinc finger. This is formed by the binding of Zn^{2+} to the following sequence in a protein:

$$\underbrace{\phi \; X \; C}_{\beta} \; X_{2-4} \; \underbrace{C \; X \; X \; X \; F}_{\beta} \; X_5 \; \underbrace{L \; X \; X \; H \; X_{3,4} \; H}_{\alpha \text{ helix}}$$

$$\xrightarrow{} \text{turn} \xrightarrow{} \text{turn} \xrightarrow{}$$

$$\beta \qquad\qquad \beta \qquad\qquad \alpha \text{ helix}$$

Here ϕ is a hydrophobic amino acid and X may be any amino acid. After Krizek *et al.*[428]


A
3'
5'
bp 2
COOH
Finger5
bp10
Finger3
Finger4
Finger2
NH_2
bp20
3'
5'
Finger1
B

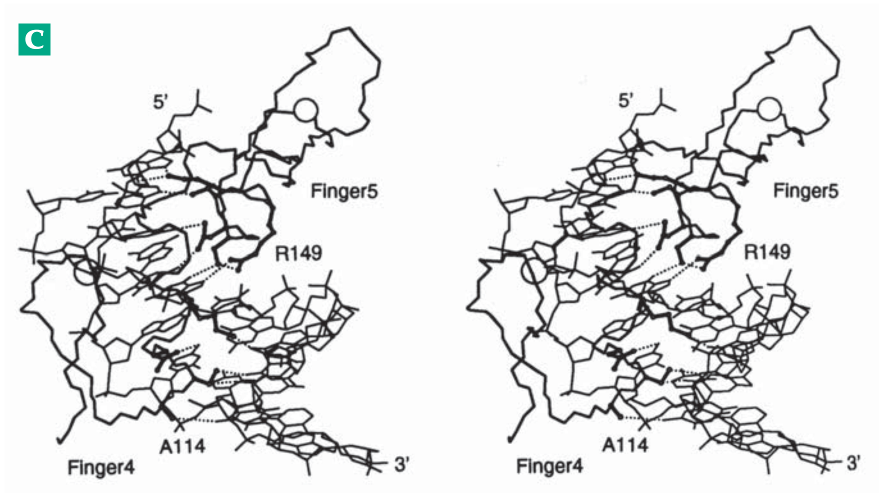

C
5'
Finger5
R149
A114
Finger4
3'
5'
Finger5
R149
A114
Finger4
3'

Figure 5-38 (A) Sketch of a complex of a five-zinc finger protein bound to a high-affinity site in DNA. This protein is a human oncogene called GL1. Its gene is amplified in many tumors, but it is similar to many other transcription factors. The cylinders and ribbons highlight the α helices and β sheets. Cobalt ions, which replaced the Zn^{2+} in the crystals, are shown as spheres. (B) Stereoscopic view of the complex in a similar orientation. (C) Stereoscopic view emphasizing the contacts of side chains from fingers 5 and 6 with the DNA. The view is similar to that in (A) and (B) but the structure has been tilted back. Drawings courtesy of Nikola Pavletich.[427]

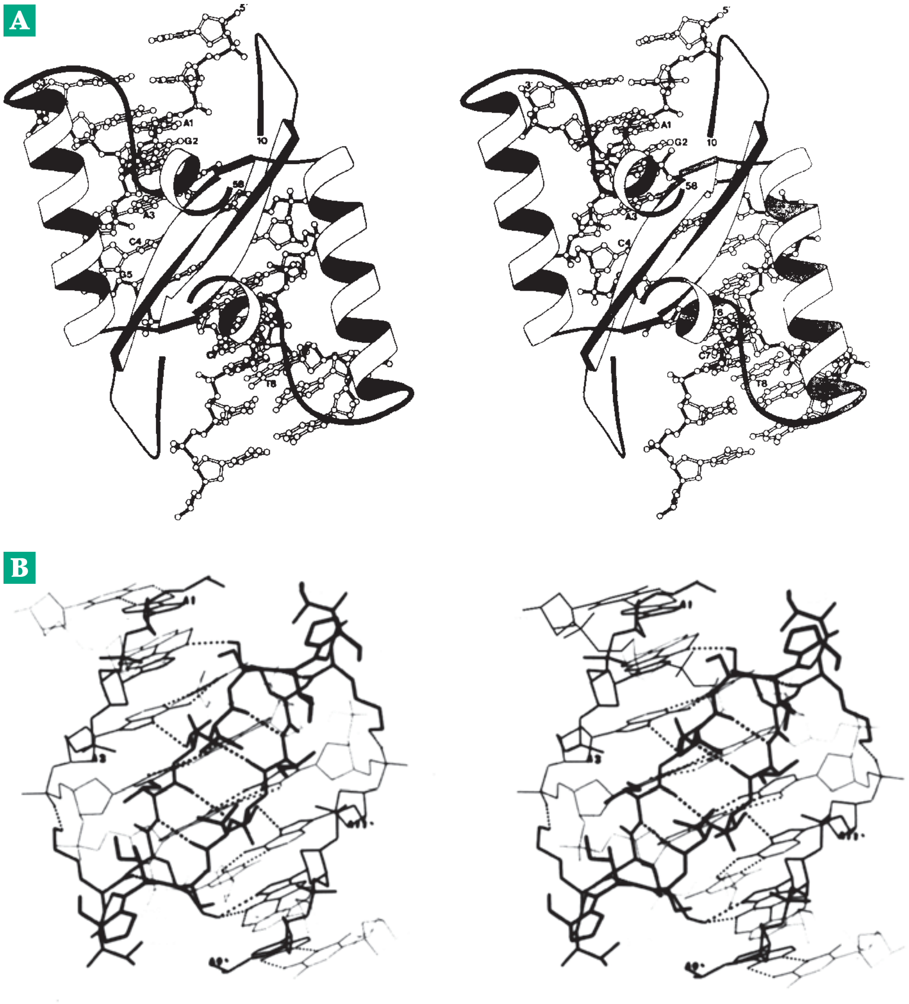


Figure 5-39 (A) Stereoscopic ribbon representation of the *E. coli* methionine repressor–operator complex. Two subunits form a dimer with a double-stranded antiparallel β ribbon that fits into the major groove of the DNA in the B form. One strand is shaded more darkly than the other. (B) View of the β ribbon and its interactions with the DNA. Notice the direct hydrogen bonds from amino acid side chains of lysine and threonine residues to bases in the specific palindromic sequence AGACGTCT. From Somers and Phillips.[429] Drawings courtesy of S. E. V. Phillips.

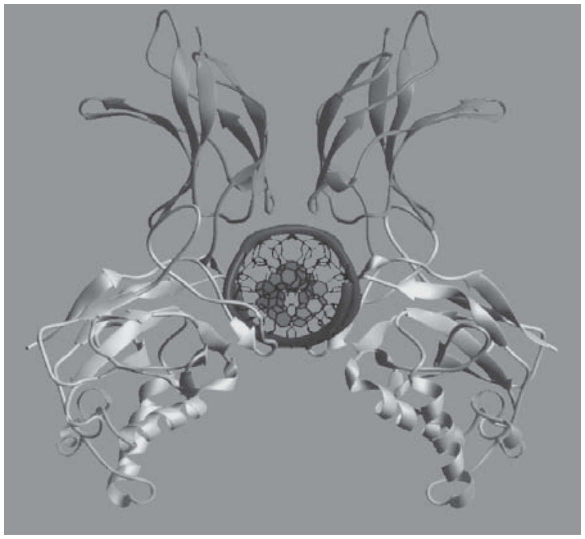


Figure 5-40 Structure of a protein known as transcription factor NF-κB bound to its DNA target. Each subunit of the dimeric protein contains two β barrel domains. The loops at the ends of the barrels interact with the DNA in the center. From Müller *et al.*[433] Courtesy of Stephen C. Harrison.

hydrogen bonds and other interactions with the DNA. This is the basis for specific recognition of the operator sequence for the *E. coli* **met** (methionine) **repressor** as is illustrated in Fig. 5-39.[429] The same binding motif is employed by some other repressors[430,431] and also by the abundant bacterial DNA-binding protein **HU**.[432] It has also been utilized in designing the previously mentioned hairpin polyamide DNA-binding compounds.

The winged helix family. A group of large protein transcription factors contain an N-terminal DNA-binding domain with the striking winged helix motif shown in Fig. 5-40.[433–436] It occurs in proteins from a wide range of organisms from yeast to human.

3. RNA-Binding Proteins

Among the many proteins that bind to RNA molecules[437–439] are the aminoacyl-tRNA synthetases, a variety of other well known enzymes,[440] the ribosomal proteins discussed in Chapter 29, and various proteins with dual functions of catalysis and regulation of translation. A widely found RNA-binding **ribonucleoprotein domain** (RNP domain), occurs in hundreds of proteins including many of the RNA processing proteins considered in Chapter 28.[437,441,442] This 70–90 residue αβ module binds an RNA strand against a β sheet surface. Another ~ 70-residue protein motif binds ds RNA.[443]

G. Viruses and Plasmids

Attacking every living thing from the smallest mycoplasma to human beings the nucleoprotein particles known as viruses have no metabolism of their own. However, they "come alive" when the nucleic acid that they contain enters a living cell. Viruses are significant to us not only because of the serious disease problems that result from their activities but also as tools in the study of molecular biology. A mature virus particle or **virion** consists of one or more nucleic acid molecules in a protein coat or **capsid**, usually of helical or icosahedral form. The capsid is made up of morphological subunits called **capsomers**. They can sometimes be seen clearly with the electron microscope. The capsomers in turn are usually composed of a number of smaller protein subunits. Some of the larger viruses are surrounded by membranous envelopes. Others, such as the T-even **bacteriophage** which attack *E. coli*, have extraordinarily complex structures (Box 7-C).

Most viruses contain a genome of either double-stranded DNA or single-stranded RNA, but some small viruses have single-stranded DNA and others have double-stranded RNA. The number of nucleotides in a virus genome may vary from a few thousand to several hundred thousand and the number of genes from 3 to over 200. Sometimes the nucleic acid molecules within the virion are circular, but in other cases they are linear. Table 5-7 lists a few of the known types of viruses as well as some individual viruses.[72,444] The size of the genome, in kilobases (kb), or kilobase pairs (kbp) is indicated. The number of genes in a virus is often somewhat more than one per kbp of DNA. A vast amount of information is available about these infectious particles. Only a few reference sources are cited here.[72,444–446] The architectures of some helical and icosahedral protein coats that surround genomic DNA or RNA are described further in Chapter 7.

1. Viruses with Single-Stranded DNA

The very small **helical bacteriophages** of the Ff family, such as fd, f1, and M13 (see Fig. 7-7) resemble thin bacterial pili but each virus particle contains a molecule of single-stranded circular ~6400 nucleotide DNA of M_r ~2 x 10^6 which encodes only ten proteins.[447–449] Bacteriophage ϕX174 (Fig. 5-41), an **icosahedral** DNA-containing virus 25 nm in diameter, is only three times as thick as the thinnest cell membrane. Its DNA contains just 5386 nucleotides.[450–452] The similar bacteriophage G4 contains 5577.[453,454]

It is remarkable that such tiny viruses are able to seize control of the metabolic machinery of the cell and turn it all in the direction of synthesis of more virus particles. There are only 11 known proteins encoded by the genes of ϕX or G4. Three genes encode the three kinds of protein subunits of the virus coat. Sixty copies of each are needed, as are 12 copies of a "pilot" protein.[451] Eight of the genes are spaced closely together, occupying most of the DNA. The other three genes are embedded within some of the first eight but in different reading frames. In one short region of the G4 chromosome the same nucleotides are part of three different genes, using all three possible reading frames. In addition to their own genes, these small viruses make use of many components of the cell that they infect. A large group of animal viruses, the **parvoviruses**, are similar in size and architecture to bacteriophage ϕX174.

Canine parvovirus, first identified in 1978, is now endemic.[455,456] Childhood **fifth disease** is also caused by a parvovirus.[456,457] When these single-stranded DNA viruses infect cells a double-stranded **replicative form** of DNA arises by synthesis of the complementary **negative strand** alongside the original **positive** DNA strand. Many copies of the replicative form are then synthesized. The negative strands of the replicative forms serve as templates for synthesis of numerous new positive strands that are incorporated into the progeny viruses. The whole process may take only 20 minutes. Some parvoviruses are unable to reproduce unless the cell is also infected by a larger adenovirus.

While most plant viruses contain dsRNA, the **geminiviruses**,[458] which cause a number of plant diseases, contain single-stranded DNA. The virus particle consists of a fused pair of incomplete icosahedra, evidently containing a single ~2500-bp DNA strand. Replication may require coinfection with two virus particles of differing sequence.

2. Viruses with Double-Stranded DNA

One of the smallest viruses containing double-stranded DNA is the 3180-nucleotide human **hepatitis B virus**. It infects millions of people throughout the world causing chronic hepatitis and often liver cancer.[459–461] The circular DNA is surrounded by an envelope consisting of proteins, carbohydrate, and a lipid bilayer. Many icosahedral viruses also contain dsDNA. Among them are the **papovaviruses**, some of which cause warts and others malignant tumors. Much studied by biochemists is the 5386-nucleotide **simian virus 40 (SV40)**, a monkey virus capable of inducing tumors in other species.[462,463] Closely related

TABLE 5-7
Characteristics of Some Individual Viruses and Groups of Viruses

Type of genome[a] and group or individual virus name	Shape[b]	Diameter[c] (nm)	Masses in daltons x 10^{-6}		Thousands of bases or base pairs (kilobases)
			Total	DNA or RNA	
DNA, single-stranded					
Bacteriophages *fd*, *f1*, *M13*	H	~6 x ~880 (length)	17.6	2.1	6.4
Bacteriophages ϕ174, G4	I	25	6.2	1.8	5.4–5.6
Parvoviruses	I	18–25		1.8	5.5
Geminiviruses	I (fused pairs)				
DNA, double-stranded					
Hepatitis B virus	I (enveloped)			2.1	3.18
Papoviruses	I	45–55			
SV40 (monkey)	I		17.6	3.5	5.22
Polyoma (mouse)	I	45	23.6	3.3	4.96
BK virus (human)	I			3.4	4.96
Papilloma (human wart)	I	56		5.3	8.0
Bacteriophage ϕ29	T			12	19.3
Adenoviruses	I	70		20–30	30–45
Bacteriophage Mu				25	38
Bacteriophage T7	T(short)				39.9
Bacteriophage P22	T			28.5	43.2
Bacteriophage λ	T			32	48.6
T-even bacteriophage	T	100 x 80 (head)	215	130	166
Baculoviruses of insects	I	70–130			
Herpesviruses	I	100			
	core	78	~1000	80–120	80–140
	envelope	150–200			
Pox viruses, e.g.,					
Smallpox, vaccinia (cowpox)	C	160 x 250	~4000	150–240	240–300
Cauliflower mosaic virus	I				8.0
RNA, single-stranded					
Unsheathed					
Potato spindle tuber viroid				0.116	0.30
Hepatitis delta virus					1.7
Sheathed, *plus-strand*					
Tobacco necrosis satellite	I	18	1.7	0.4	1.20
Small bacteriophages					
R17, MS2, Qβ	I	23–26	3.6–4.0	1.2–1.5	3.5–4.5
Picornaviruses	I		8.4	2.6	7.9
Polioviruses	I	27	6.4	2	6.1
Rhinoviruses	I	27–30	7–8	2.2–2.8	6.7–8.5
Turnip yellow mosaic virus	I	28	5.0–6.0	2.0	6.1
Tobacco mosaic virus	H	18 x 300	40	2.2	6.7
Togaviruses	I	20–40		4	11
Negative-strand viruses					
Influenza virus	I	80–100	200	2.0	6.1
Bullet-shaped viruses					
Rhabdoviruses	C	20 x 130			
Retroviruses	I	80–100		7–10	20–30
RNA, double-stranded					
Reoviruses	I	55–60		11–12	16–18
Mobillivirus	I (enveloped)	38			
Rotavirus	Wheel-shaped	70			
Leishmania RNA virus					5.28

[a] Complete nucleotide sequences are known for most of these viruses.

[b] Shapes are indicated as I, icosahedral; H, helical; T, a tailed phage; C, complex.

[c] The second dimension given for some helical and complex viruses is the length (nm).

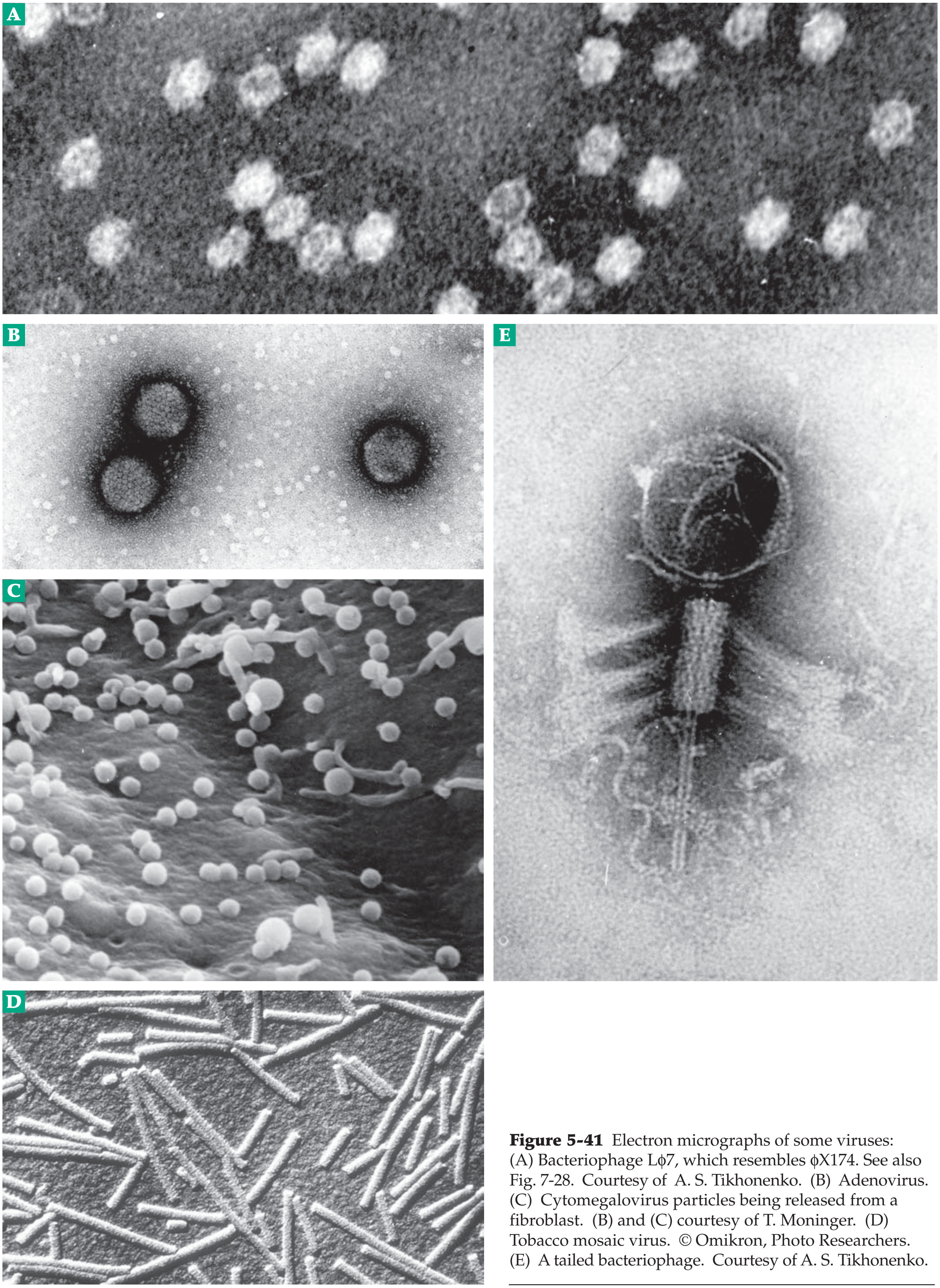


Figure 5-41 Electron micrographs of some viruses: (A) Bacteriophage Lϕ7, which resembles ϕX174. See also Fig. 7-28. Courtesy of A. S. Tikhonenko. (B) Adenovirus. (C) Cytomegalovirus particles being released from a fibroblast. (B) and (C) courtesy of T. Moninger. (D) Tobacco mosaic virus. © Omikron, Photo Researchers. (E) A tailed bacteriophage. Courtesy of A. S. Tikhonenko.

is the **polyomavirus** of the mouse[464] and the **BK virus** of humans,[465] also suspected of causing cancer. The **papillomaviruses** cause warts and perhaps cancer[466] and the larger (70 nm diameter) **adenoviruses** (Fig. 5-41),[467,468] cause respiratory infections. Some 32 types infect humans. Many of the details of eukaryotic transcription were first studied using the adenovirus.[469] The very large **herpesviruses** are enveloped by a lipid-containing membrane.[470–472] Among them are herpes simplex viruses, which infect human mucous membranes and varicella-zoster virus, which causes chicken pox and shingles. Other herpesviruses include cytomegaloviruses (Fig. 5-41), another common human pathogen, and the Epstein–Barr virus, which causes mononucleosis and is suspected of causing cancer.[473] Another herpesvirus has been associated with multiple sclerosis.[474] The very large icosahedral **baculoviruses** cause polyhedroses in insects. One that infects the fly *Tipula* measures 130 nm in diameter. **Poxviruses** are also large and complex.[475] The tailed bacteriophages (Box 7-D) range in size from the small ~29-kDa phage P22 to the very complex T-even phage. Some plant viruses also contain dsDNA. One of these, the **cauliflower mosaic virus**,[476,477] is transmitted by aphids. It has proved useful as a gene-transfer vehicle for genetic engineering.

3. Viruses Containing RNA

Several plant diseases including the **potato spindle tuber disease** are caused by **viroids**, molecules of single-stranded RNA only 240–380 nucleotides in length with a folded structure.[478–481] Such an RNA could code for a protein containing only about 100 amino acids. However, the known 359-nucleotide sequence of the potato spindle tuber virus contains no AUG initiation codon. It seems impossible that the virus carries any gene for a protein. Conserved features of viroid sequences suggest a close relationship to the intervening DNA sequences known as type I introns (see Fig 28-19). Whatever its genetic message, a plant viroid causes the plant cell to replicate many copies of the viroid molecule, which may then be transmitted to other plants by aphids, or on the surface of tools, by humans. A larger 1678-nucleotide viroid-like RNA (**hepatitis delta virus)** has been identified in human patients with severe chronic hepatitis and who were also infected with hepatitis B virus. The ssRNA in this virus does encode at least one protein.[482–484]

One of the smallest of the encapsulated RNA-containing viruses is the **satellite tobacco necrosis** virus. It replicates only when the plant is also infected with the larger tobacco necrosis virus. The satellite virus, whose three-dimensional structure is known from X-ray diffraction studies,[485] contains a 1200-nucleotide strand of RNA which encodes a 195-residue protein. Sixty copies of the latter are assembled around the RNA in an icosahedral array (Fig. 7-14) to form the virion. The structure of the similar satellite tobacco mosaic virus has also been described in detail.[486,487]

RNAs of the small bacteriophages f2, R17, MS2 and Qβ contain 3500 – 4500 nucleotides and are enclosed in icosahedral shells made up of 90 identical subunits.[488–490] Initially only three genes were evident but a fourth small gene in a different reading frame has since been found in at least one of them.[491] The RNA molecules in these and other **positive-strand** viruses serve as mRNA within the host cells. Another group of small RNA viruses are the **picornaviruses** (picoRNA, meaning very little RNA). Many of these icosahedral viruses of 15–30 nm diameter attack humans. Among them are the **enteroviruses** including the **polioviruses**,[492,493] the **hepatitis A virus**,[494] the **coxsackieviruses**, and some of the **echoviruses**. A second class of picornaviruses include **rhinoviruses** which cause the common cold. More than 100 types are known.[495–497] The **foot–and–mouth disease** virus,[498–500] which attacks the cloven-footed animals, and the **Mengo virus**,[501] which can cause a fatal encephalitis in mice, are also picornaviruses. The three-dimensional structures of polio virus, human rhinoviruses, Mengo virus, and many other eicosahedral viruses are known. Their architecture is discussed in Chapter 7.

The **togaviruses**, which are a little larger than the picornaviruses, have an icosahedral core surrounded by a lipid membrane. Yellow fever and rubella (German measles) are both caused by togaviruses. Other togaviruses, such as **Sindbis virus**[502] and **Semliki Forest virus**,[503] have become important in biological research.

A large number of icosahedral RNA viruses of diameter 28–30 nm (Fig. 7-14) attack plants, causing diseases such as tomato bushy stunt,[504] southern bean mosaic,[505] or turnip yellow mosaic. Best known of the helical RNA viruses is the **tobacco mosaic virus** (Figs. 5-41, 7-8).[506-507a] Its genome contains 6395 nucleotides as linear ssRNA. Many strains are known. Related viruses cause cucumber green mottle[508] and other plant diseases.

Large viruses of 80 – 100 nm diameter bearing 8–10 spikes at the vertices of the icosahedra cause influenza,[509,510] mumps, measles, and related diseases. The internal structure must be complex. Only 1% of the virus is RNA, and that consists of several relatively small pieces. These are **negative strand viruses** whose RNA is of the opposite polarity to the mRNA. The latter must be formed by transcription from the negative strand. The viruses carry their own RNA polymerase for this purpose. Of even more complex structure are the bullet-shaped **rhabdoviruses** which cause rabies and vesicular stomatitis.[511] The diameter of these viruses is 65–90 nm and the length 120–500 nm. The internal structure includes a helical arrangement of nucleoprotein. They also are negative strand viruses.

Among the **retroviruses**[512–514] are types B and C **oncoviruses** which induce malignant tumors in mammals and birds[515] and the **human immunodeficiency virus** (HIV), the apparent causative agent of AIDS. Their RNA functions in a surprising way. Each virion contains a **reverse transcriptase**, an enzyme that transcribes copies of circular dsDNA copies from the one or two mRNA-like molecules that make up the virus genome. Following action of the reverse transcriptase, one of the transcribed DNA circles becomes covalently spliced into the host's own cellular DNA. There it remains permanently as a **provirus**. RNA molecules transcribed from the provirus serve as mRNA for virus-encoded proteins and also as the genomes for new virus particles.

Double-stranded RNA is unusual in nature but constitutes the genome of the **reoviruses**.[516] The RNA of these viruses fragments into segments upon infection. One member of the group is thought to be the cause of acute diarrhea of infants.[517]

4. Viruses without Nucleic Acid?

The cause of the slow, fatal neurological disease of sheep known as **scrapie** has been a mystery for many years. Similar human diseases include **kuru** and **Creutzfeldt-Jakob** disease.[518 – 520] Scrapie can be transmitted by injection and this has permitted isolation of the apparent infective agent, a 27- to 30-kDa hydrophobic protein particle[521] which is devoid of DNA or RNA. Prusiner[521] suggested the name **prion** (proteinaceous infectious particle) for the scrapie agent. However, mRNA for the prion is present in normal as well as infected brains, and protein produced in mouse cells from cloned prion genes did not cause scrapie infections. Therefore, there was doubt about the causative agent for the disease. The prion concept is now generally accepted and is considered further in Chapter 29. There are still some who are looking for a nucleic acid component.[519,522–524]

5. Life Cycles

Viruses have many modes of life. They enter cells in various ways. Some enter through coated pits from which they are taken into lysosomes via endocytosis. Others are literally injected into the cells (See Box 7-C). Within cells some viruses are assembled in the nucleus, some in the cytoplasm, and some in membranes. The typical life cycle of a virus leads to rapid formation of large numbers of progeny. Within 20 minutes after entrance into a bacterial cell, a bacteriophage can induce the formation of 100–200 new bacteriophage particles. One of the bacteriophage genes encodes a protein that is also synthesized by the host and which induces lysis of the cell membrane and destruction of the cell. Many animal viruses destroy cells in a similar fashion.

Temperate bacteriophage, the best known being phage λ, have a very different life cycle. Their DNA usually becomes integrated at a specific point into the genome of the bacterium (Chapter 27). Only rarely is an infected cell lysed. The retroviruses that attack mammals and birds have a similar characteristic. Their DNA is also integrated into the host genome. Some viruses that usually produce lysis of cells, e.g., SV40, adenoviruses, herpes viruses, and hepatitis B virus, can occasionally be integrated into the DNA of the host. If such integration occurs in the middle of a gene, that gene will be mutated. This is one way in which such viruses may induce cancers.

One of the most important results of integration of viral DNA into the host genome is that the integrated genes are replicated as part of the genome and are transmitted from one generation to the next. Among these are the cancer-causing **viral oncogenes** (*v-onc*), which are discussed in Box 11-D and in Chapter 11, Section H. While viruses are important causes of cancer in some animals, relatively few human cancers are thought to result directly from the action of viruses. However, the Epstein–Barr virus, which causes mononucleosis, can sometimes be integrated into epithelial cells of nasal regions and can evidently cause cancer. The same virus appears to be responsible for Burkitt's lymphoma, a common cancer in certain areas in Africa.[525]

6. Plasmids and Transposable Genetic Elements

In addition to their chromosomal DNA, bacteria often carry extra small pieces of DNA as permanent parts of their genome. These **plasmids** (sometimes called **episomes**), which are about the size of the DNA of viruses, replicate independently of the host chromosomes. Each bacterial cell usually contains more than a single copy of the plasmid. For example, the "colicinogenic" plasmid **ColE1**, that infects *E. coli* is a circular piece of DNA of molecular mass 4.2 x 10^6 Da. Over 20 copies are normally found per cell but in the presence of a suitable concentration of the drug chloramphenicol the number may rise to 1000–2000.

Plasmids carry a variety of genes which are often useful to bacteria. Some proteins encoded by plasmid genes confer drug resistance to a bacterium. Some are antibiotics. For example, a protein encoded by a gene in plasmid ColE1 is toxic to other strains of *E. coli.* Some plasmids carry genes for enzymes needed for the oxidation of hydrocarbons. Some plasmids contain genes for the **restriction endonucleases** which have become essential to present-day molecular biology and genetic engineering (Section H, 2).

As with some viruses, the DNA of many plasmids can become integrated into the genome of the host. An example is provided by the large 62-kDa plasmids known as **sex factors**. They contain genes encoding the protein subunits of the sex pili (Chapter 7) and can become integrated into the bacterial chromosome. Bacteria containing integrated sex factors are "male" and are able to transfer genes not only of the sex factor but also of virtually the entire bacterial genome into other susceptible bacterial cells. This provides bacteria with the means for sexual reproduction. The transfer of DNA between the bacteria may occur via the sex pili (see Chapter 26). In this respect the sex factors are similar to viruses such as M13 that also appear to gain entrance to bacteria via sex pili.[72]

Integrated viruses are also related to **transposable genetic elements** (transposons). These are segments of DNA that allow genes to move from place to place within the chromosomes (Chapter 27).

H. Methods of Study

Many of the methods discussed in Chapter 3 are directly applicable to nucleic acids. A few additional methods will be considered in this section.[525a]

1. Isolation and Separation of Nucleic Acids

RNA is often extracted from lysed cells or tissues, separated ribosomes, mitochondria, plastids, or nuclei by warming with aqueous phenol and a detergent such as sodium dodecyl sulfate (SDS). Proteins are denatured by this treatment and are dissolved by the phenol, while RNA remains in the lighter aqueous layer. Depending on the conditions DNA may either remain in the aqueous layer or be removed.[526–528] Various precipitation and extraction procedures may be used to separate the RNA in the aqueous layer from polysaccharides, from DNA (if present), and from their components.[526,528,529] DNA may be extracted from cells or nuclei as a nucleic acid–protein complex using 1 M NaCl. The protein can then be denatured with an organic solvent, by detergents, or by phenol. It is desirable to digest proteins away with a nonspecific protein-hydrolyzing enzyme such as proteinase K.[528] After removal of proteins DNA is often precipitated with cold ethanol.

During isolation of RNA, bentonite (a type of clay) or other inhibitors of ribonuclease are often added. For the same reason, chelating agents that complex metal ions needed for the action of deoxyribonucleases are used to protect DNA. Care is necessary to avoid shearing of the very long, narrow strands of DNA. Even rapid pipetting of solutions will cause such breakage.

Extracted RNA molecules may be separated from each other by centrifugation in a sucrose gradient (Chapter 3).[528] Fragments of DNA are purified in the same way or by equilibrium centrifugation in **CsCl gradients**.[37] Concentration gradients in the dense salt solution are stable, and the sharpness of banding of particles is ensured by use of a high centrifugal field. Single-stranded DNA may be separated from double-stranded DNA, and DNAs of differing G + C content can be separated. The latter separation is based on differences in buoyant densities ρ in CsCl which are approximately shown in Eq. 5-9.

$$\rho = 1.660 + 0.098 \text{ (mole fraction C + G)} \qquad (5\text{-}9)$$

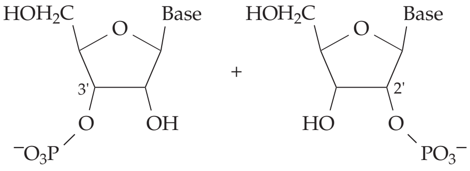
One of the most important methods for separating either RNA or DNA mixtures is zone electrophoresis through polyacrylamide or agarose gels. The separated bands may be visualized by scanning in ultraviolet light or by fluorescence of intercalated dyes such as ethidium bromide (Figs. 5-20, 5-22). This method is being displaced to some extent by HPLC using DEAE type ion exchange columns[530,531] for small lengths of DNA including plasmids. The procedure called **pulsed field electrophoresis** makes it possible to isolate very large pieces of DNA, up to several million base pairs in length.[532–534] The separation is carried out in agarose gels, through which the long DNA rods must move in a snakelike fashion.[534] The current is delivered in a pulse and then, after a period of a second to several minutes, a second pulse in a different direction, usually at 90° to the first. The procedure is repeated many times. The size of the DNA seems to affect the time required to reorient the molecules and to start moving in the second direction. Intact DNA from small chromosomes can be separated (Fig. 5-42). To prevent breakage of the DNA by shearing, intact cells are suspended in liquid agarose and allowed to gel into a block about 2 x 5 x 10 mm in size. The block is treated with enzymes and detergents to lyse the cells and to remove all protein and RNA.[534] The block, containing the residual DNA molecules, is then embedded in the electrophoresis gel. Other methods of DNA separation include chromatography on hydroxyl-apatite and gel filtration.

2. Hydrolysis and Analysis

Both DNA and RNA are easily broken down by acid-catalyzed hydrolysis. Thus, heating at 100°C for one hour in 12 M $HClO_4$ is sufficient to hydrolyze nucleic acids to their constituent bases. However, for analysis of RNA it is better to heat in 1 N HCl for 1 h at

100°C. The products are adenine, guanine, cytidine-5'-phosphate, and uridine-5'-phosphate.[535] As is suggested by this distribution of products, the glycosylamine linkages to purines are more labile than those to pyrimidines. The linkages are also less stable in DNA than in RNA. A procedure based on these differences and useful in sequencing by the Maxam-Gilbert method, is to leave DNA overnight in the cold at pH 2 to cleave off all of the purine bases. The resulting polymer is known as an **apurinic** acid.

In alkaline solutions RNA is hydrolyzed to a mixture of 2'- and 3'-nucleotides.



The mechanism involves participation of the free 2'-OH of the ribose groups and formation of cyclic 2', 3'-phosphates and is similar to that of pancreatic ribonuclease (Chapter 12). Because deoxyribose lacks the free 2'-OH, the phosphodiester linkages in DNA are quite stable in base.

Hydrolytic cleavages of nucleic acids by the enzymes known as **nucleases** are of great practical value. Pancreatic ribonuclease, an **endonuclease**, cuts a chain adjacent to a pyrimidine in nearly random fashion, leaving phospho groups attached to the 3' position in the nucleotide products (Fig. 5-43). **Exonucleases** cleave from the ends of chains. For example, the phosphodiesterase of snake venom cleaves from the 3' end, which must have a free 3'-OH group, to give 5'-nucleotides. On the other hand, the phosphodiesterase from spleen has the opposite polarity, cleaving chains from the 5' end to give 3'-nucleotides. Similar variations in specificity are found among enzymes that cleave DNA. For example, pancreatic DNAase I, which cleaves preferentially between adjacent purines and pyrimidines, yields 5' mononucleotides whereas DNAase II gives 3'-mononucleotides. Various hydrolytic cleavage reactions of polynucleotides are summarized in Fig. 5-43.

The most striking specificity in DNA hydrolysis is displayed by the **restriction endonucleases** which are discussed further in Chapter 26. These fussy catalysts cleave only at points within or close to a defined sequence of several nucleotides in double-stranded DNA. For example, the enzyme EcoR I cuts only at the following palindromic sequence:

$$\begin{array}{c} 5'-\overset{\downarrow}{G}\ A\ A\ T\ T\ C-3' \\ \bullet\ \ \bullet\ \ \bullet\ \ \bullet\ \ \bullet\ \ \bullet \\ 3'-C\ T\ T\ A\ A\ \underset{\uparrow}{G}-5' \end{array}$$

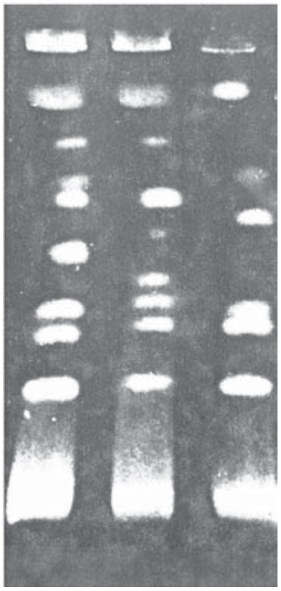
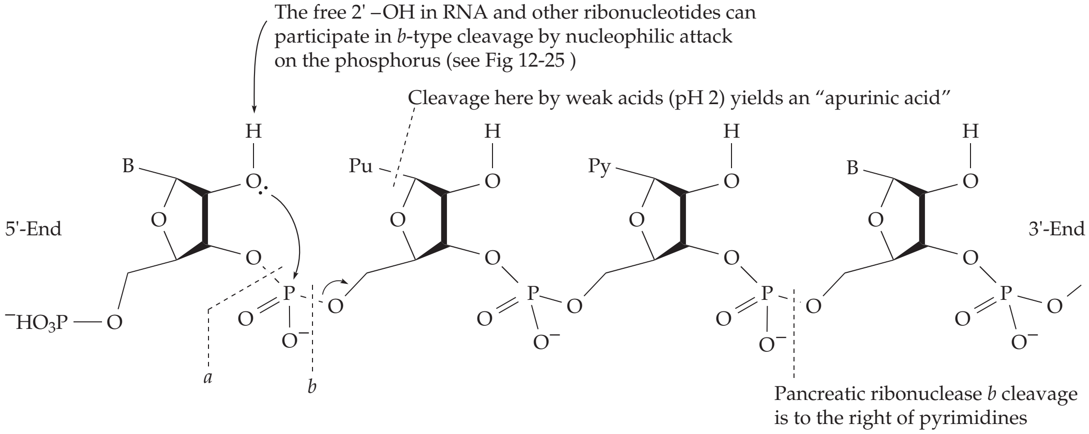


Figure 5-42 Intact DNA from the chromosomes of three strains of the malaria parasite *Plasmodium falciparum*, ranging from 750 Kb to 5 Mb, separated by pulsed-field gel electrophoresis. Courtesy of C. Smith and T. E. Wellems. Reproduced by permission of Amersham Pharmacia Biotech Inc.

The cuts in the two strands are made at the points indicated by the arrows. This one endonuclease will cut almost any DNA into long pieces averaging about 5000 base pairs each. These pieces can in turn be cleaved by other restriction endonucleases to form smaller fragments. Since there are about 2400 of these enzymes known, with 188 different specificities,[536] it is possible to cut any piece of DNA down to a size of 100–500 base pairs, ideal for sequencing.[537–539] Each fragment has known sequences at the two ends. Some restriction enzymes cleave outside their specific recognition sequence (see Table 26-2). Some recognize 16-nucleotide palindromes and cut at rare sites.

It is sometimes desirable to cut a large DNA molecule at only a few points. One approach is to protect most sites of a restriction enzyme's action by methylating them (see Chapter 26) while protecting the desired cleavage site, for example by a repressor protein[540] or by a PNA molecule (p. 227) of specific sequence designed to "clamp" the site chosen for protection.[541] Ribozymes (Chapter 12) have been engineered to be as specific or more specific than endonucleases.[542,543] Other new approaches are being developed.[544]

The base composition of either RNA or DNA can be determined after hydrolysis catalyzed by 98% formic acid at 175°C for 30 min or by 12 M perchloric acid at 100°C for 1 h.[545] The bases can then be separated by ion exchange chromatography on a sulfonated polystyrene resin. RNA can be hydrolyzed to a mixture of nucleoside 2′- and 3′-phosphates by 0.3 M NaOH at


The free 2' –OH in RNA and other ribonucleotides can participate in *b*-type cleavage by nucleophilic attack on the phosphorus (see Fig 12-25)

Cleavage here by weak acids (pH 2) yields an "apurinic acid"

Pancreatic ribonuclease *b* cleavage is to the right of pyrimidines

A. Cleavage at point *a* is catalyzed
  1. Throughout the molecule by endonucleases
     Pancreatic deoxyribonuclease I
  2. Only at the 3' end by exonucleases
     Venom diesterase, nonspecific, attacks DNA and RNA. A free 3'-OH is essential

B. Cleavage at point *b* is catalyzed
  1. Randomly throughout the molecule by endonucleases and by bases (nonenzymatically)
     Pancreatic ribonuclease cleaves only to the right of a pyrimidine-containing nucleotide
     Ribonuclease T1 of *Aspergillus oryzae* cleaves to the right of a guanine-containing residue (3'-guanylate) in ssRNA
     Ribonuclease T2 of *Aspergillus oryzae* cleaves to the right of an adenine-containing residue (3'-adenylate) in ssRNA
     Pancreatic deoxyribonuclease (DNase) II
     Micrococcal DNase
  2. Only at the 5' end by exonucleases
     Bovine spleen phosphodiesterase hydrolyzes both polyribo- and polydeoxyribonucleotides

Figure 5-43 Some hydrolytic cleavage reactions of polynucleotides. Reactions of both RNA and DNA are included.

37°C for 16 h and DNA can be hydrolyzed to nucleotides enzymatically. The negatively charged nucleotides can then be separated by ion exchange chromatography on a quaternary base-type resin (Chapter 3). Periodate cleavage (Eq. 4-12) and reduction of the resulting dialdehydes by [^{3}H]NaBH$_4$ to trialcohols allows introduction of a radioactive label (Fig. 5-44). Alternatively, the dialdehydes can be reductively alkylated by an amine plus NaCNBH$_3$.[546]

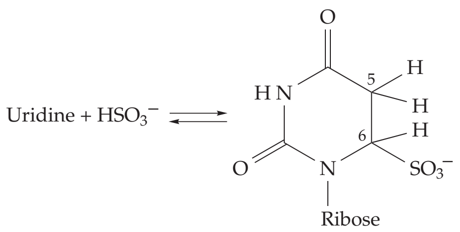
The total content of RNA + DNA in tissues may be estimated from the phosphorus content or by color reactions of the sugars.[37,545] These reactions depend upon dehydration to furfural or deoxyfurfural by concentrated sulfuric acid or HCl (Eq. 4-4). Furfural formed from RNA reacts with orcinol (3,5-dihydroxytoluene) and ferric chloride to produce a green color useful in colorimetric estimation of RNA. A similar reaction of DNA with diphenylamine yields a blue color.

Quantitative determination of over 90 free nucleotide compounds found within cells can be accomplished by thin layer chromatographic procedures on as few as 10^6 bacterial cells (~2 µg) labeled by growth in a ^{32}P$_i$-containing medium.[547]

3. Characteristic Reactions of the Bases and Backbone

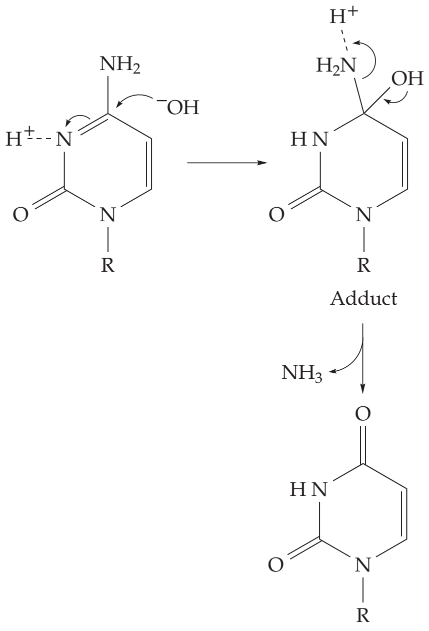
Reactions of nucleophiles. A number of nucleophilic reagents add reversibly at the 6 position of pyrimidines. Thus, bisulfite adds to uridine (Eq. 5-10).[528] Hydroxylamine (HONH$_2$) adds in a similar fashion to give a compound with –HNOH in the 6 positions.[528] Sodium borohydride (NaBH$_4$), which can be viewed as a donor of a hydride ion (H$^-$), reduces uridine to the 5,6-dihydro derivative. This presumably occurs by attack of the hydride ion at position 6 in a manner analogous to the reaction of bisulfite in Eq. 5-10.

Cytidine reacts in the same way, but the bisulfite addition compound is unstable. These C5–C6 adducts of cytidine all have a greatly enhanced reactivity at C4,



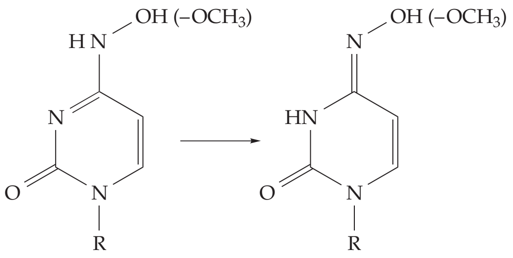
(5-10)

presumably because of the lessened aromatic character of the ring. Cytidine is slowly deaminated by base, presumably as a result of attack by hydroxyl ion on the electrophilic center at C4 and subsequent elimination of NH_3 (Eq. 5-11). The reaction is catalyzed by buffer salts and by bisulfite and hydroxylamine. Catalysis probably occurs, at least in part, as a result of addition of these nucleophiles at the 6 position to form compounds with increased nucleophilic reactivity at C4.[548] Hydroxylamine and methoxyamine (NH_2OCH_3) participate in reactions parallel to that of the hydroxyl ion in Eq. 5-11. Products contain $-NHOH$ or $-NH-OCH_3$ in place of $-NH_2$ but tautomerize to the more stable forms shown in Eq. 5-12. Similar substitution reactions occur with other amines.[33] The C6 adduct with hydrazine can undergo ring cleavage (Eq. 5-13). The initial product then undergoes β elimination, leaving ribosylurea or deoxyribosylurea. The same reaction can be carried out on intact strands of DNA and is widely used in determination of nucleotide sequences.



(5-11)

The conversion of 5-hydroxymethylcytosine to the



(5-12)

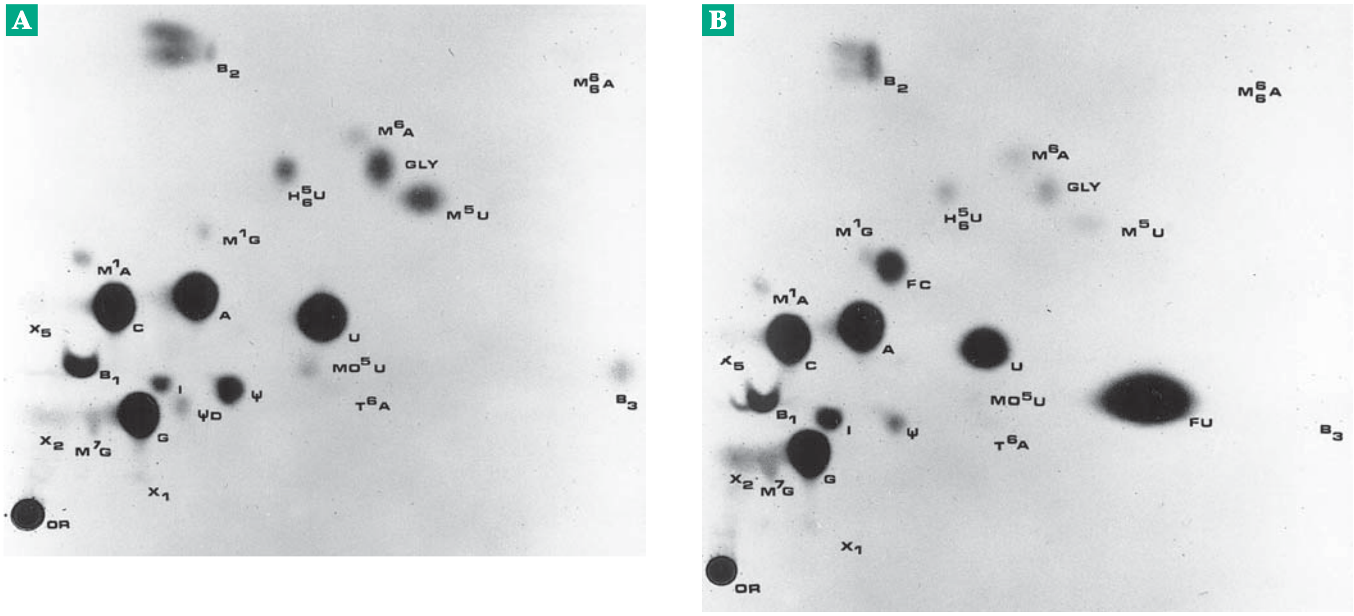
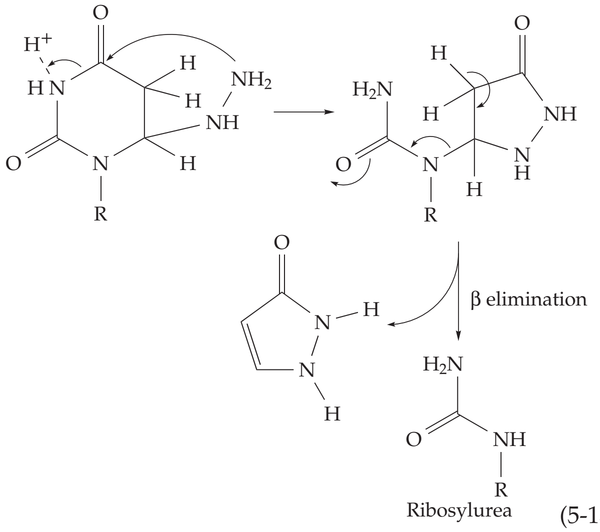


Figure 5-44 Fluorographs of 2′3′-[^{3}H] nucleoside trialcohols from *Bacillus subtilis* grown in the absence (A) and presence (B) of 5-fluorouracil. About 2.3 nmol of nucleosides from each sample was chromatographed and exposed to X-ray film for 90 h. at –80 °C. Or (origin) B_1, B_2, and B_3 contain unidentified material present in a reaction mixture lacking RNA. Abbreviations used: FU, 5-fluorouridine; FC, 5-fluorocytidine; U, uridine; C, cytidine; G, guanosine; A, adenosine; I, inosine; m′G, 1-methylguanosine; m^{7}G, 7-methylguanosine; m′A, 1-methyladenosine; m^{6}A, 6-methyladenosine; m$^6_6$A, 6,6-dimethyladenosine; t^{6}A, N-[9-(β-D-ribofuranosyl) purin-6-yl carbamoyl] threonine; H$^5_6$U, 5,6-dihydrouridine; ψ, pseudouridine; ψ_D, decomposition product of ψ; m^{5}U, 5-methyluridine (ribosylthymine); mo^{5}U, 5-methoxyuridine; N′, a nucleoside trialcohol obtained by reduction of a nucleoside dialdehyde with [^{3}H]$NaBH_4$; FU-5 and FU-20 samples correspond to tRNAs from cells grown at that final concentration of 5-fluorouracil in µg / ml. Courtesy of Ivan Kaiser.

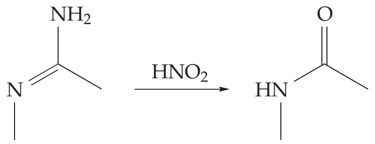


(5-13)

5-methylenesulfonate (5-$CH_2SO_3^-$) by reaction with bisulfite should also be mentioned.[549] This is a nucleophilic displacement on the electron-deficient methylene group of this base.

Reactions with electrophilic reagents. Reactions of nucleic acids with the simplest electrophile, the proton, have been considered in Section A2. Somewhat similar are the reactions by which metal ions bind at many sites on both the bases and the phosphate groups of the backbone.[550]

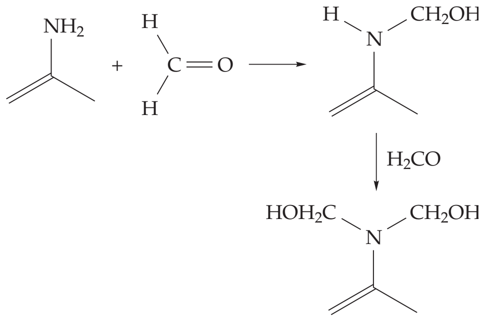
An important reaction is the deamination of amines by dilute **nitrous acid**. This reagent, by a complex mechanism, converts the amino groups of cytidine, adenosine, and guanosine to hydroxyl groups; hydroxy compounds tautomerize to the corresponding amides (Eq. 5-14).



(5-14)

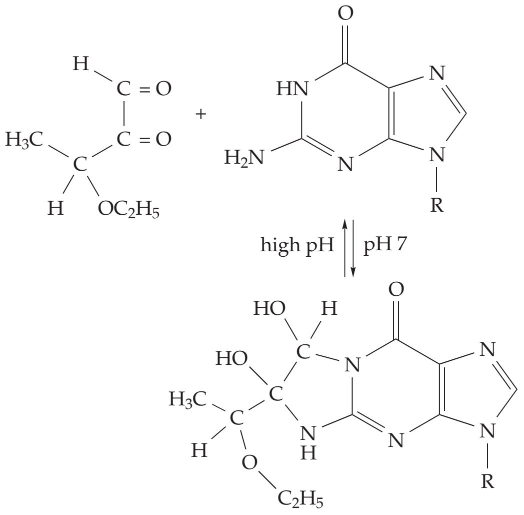
Cytidine reacts more rapidly than does adenosine which in turn reacts more rapidly than guanosine. The reaction converts cytosine into uracil and adenine into hypoxanthine. The changes are mutagenic because during replication the modified bases of the DNA pair differently than do the original bases. Guanine is converted to xanthine but this is not likely to be highly mutagenic. Nitrous acid can also convert uridine to 5-nitrouridine.

The amino groups of the bases react reversibly with aldehydes but to a lesser extent than do the more strongly basic amino groups of the amino acids. Formaldehyde forms adducts containing either one or two molecules of the aldehyde (Eq. 5-15).



(5-15)

These are reversible reactions. A more nearly irreversible crosslinking can occur by elimination of water between one of these products and a nucleophilic group in another base. A dicarbonyl reagent that is widely used because of its specificity toward guanine is **kethoxal** (Eq. 5-16).



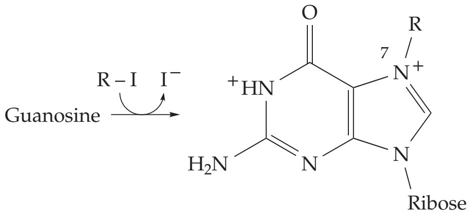
(5-16)

Formation of the cyclic product is a consequence of the presence of the adjacent amino and NH groups in the guanine ring.

Pyrimidines undergo **halogenation** at position 5 (Eq. 5-17), while guanine reacts at position 8. Adenine is quite unreactive. Elemental halogens or a variety of other halogenating reagents may be used. Of special value is iodination with ^{131}I or ^{125}I, by which a high level of radioactivity may be introduced into nucleic acids.

$$Br_2 + \text{cytidine} \rightarrow \text{5-Br-cytidine} + H^+ + Br^- \qquad (5\text{-}17)$$

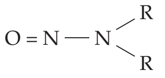
Alkylation reactions are not only of use in structural studies but also provide the basis for the action of a large class of mutagenic compounds.[528] Treatment of a nucleoside, nucleotide, or nucleic acid with an alkyl iodide or a dialkylsulfate converts residues of guanosine to an N^7-alkyl-guanosine (Eq. 5-18).



(5-18)

Reaction occurs at other nitrogen atoms as well as the oxygen atom of the base and of the ribose ring to a lesser extent. Adenine is alkylated preferentially at N-1 and cytosine and thymine at the corresponding position (N-3) almost exclusively. Uridine and thymidine react very slowly. Adenine is also alkylated at N-3, N-7 and at the exocyclic N-6.

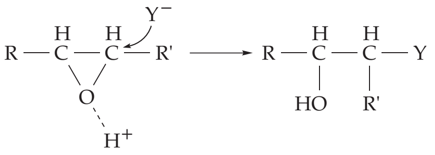
Other alkylating reagents include the powerful mutagens dialkylnitrosamines and alkylnitrosoureas.



$$\mathrm{O{=}N{-}N}\begin{matrix}\mathrm{R}\\ \mathrm{R}\end{matrix}$$

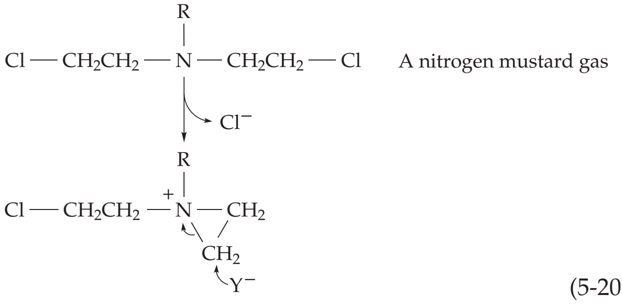
Dialkylnitrosamines

Epoxides alkylate by a nucleophilic displacement reaction that opens the ring (Eq. 5-19).



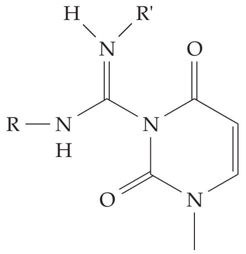
(5-19)

The nitrogen and sulfur mustards undergo internal ring closure to an iminium ion (Eq. 5-20) which can then open by attack of a nucleophilic atom of the nucleic acid.



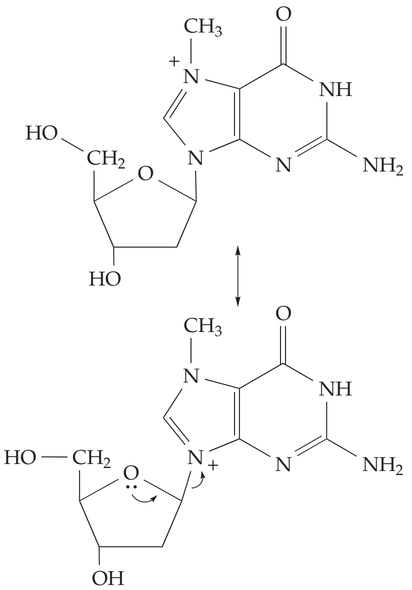
(5-20)

Other alkylating agents react through nucleophilic addition to a carbon–carbon double bond. Thus, acrylonitrile reacts with the nitrogen or oxygen atoms of nucleic acids in the same manner as does the SH group in Eq. 3-25. The water-soluble carbodiimides react as in the first step of Eq. 3-10 to form adducts of the following type:



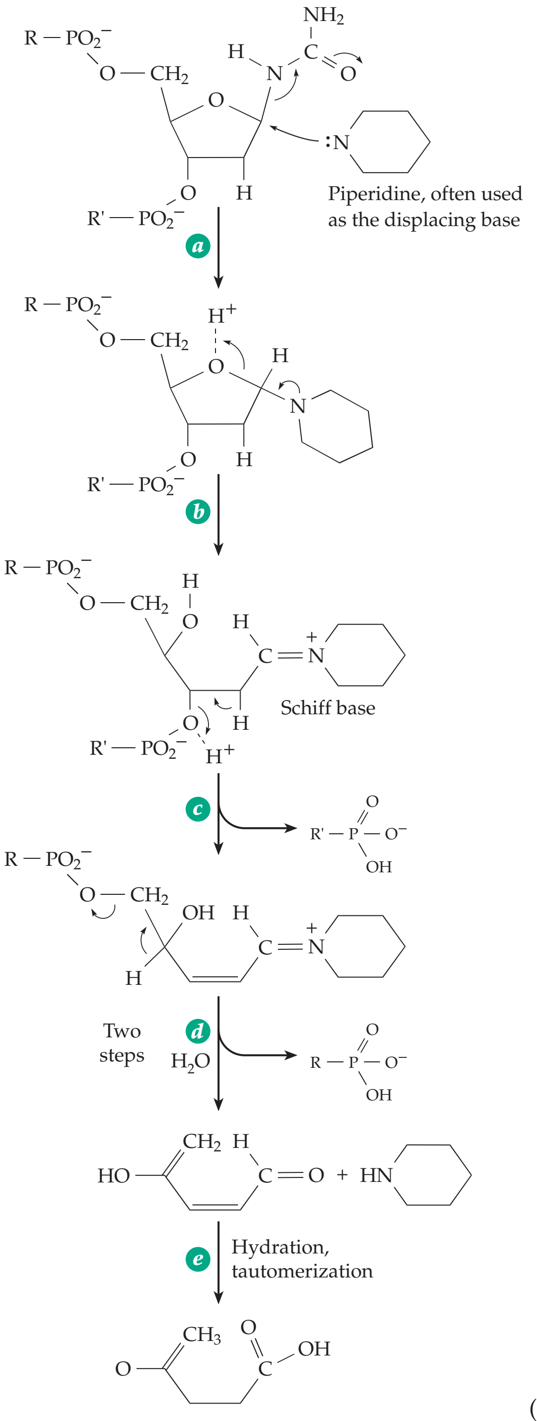
There are many other alkylating agents which often display widely varying reactivity and specificity toward particular nucleic acid bases and particular nucleotide sequences.

A striking effect of alkylation of guanine in nucleic acids is the labilization of the *N*-glycosyl linkage to the ribose or deoxyribose. This effect can be understood in terms of the induction by resonance of a partial positive charge on the nitrogen of the glycosyl linkage.



As is indicated by the small arrows on the right-hand structure, the positive charge assists in an elimination reaction that produces an oxycarbocation. The latter can then react with a hydroxyl ion from water.

Reactions causing cleavage of the sugar-phosphate backbone. Treatment of DNA with 16–18 M hydrazine (Eq. 5-13) leads to the destruction of the pyrimidine rings. The reaction can be made somewhat specific for cytosine by carrying it out in the presence of a high concentration of chloride.[551] The remaining polymer, an apyrimidinic acid, contains residues of ribosylurea. These undergo an amino-catalyzed displacement and a β elimination sequence that cleaves the polynucleotide chain (Eq. 5-21). Hydration of the aldehyde (Eq. 13-1) and several


Piperidine, often used as the displacing base
a
b
Schiff base
c
Two steps
d
H_2O
e
Hydration, tautomerization

(5-21)

tautomerization steps are involved in step *e* of this equation. This reaction is very useful in sequence determination (Section 6). Notice that "tracts" of purine nucleotides remain intact after this treatment. A similar base-catalyzed reaction sequence can be used to displace N^7-methylguanine and to cleave the polynucleotide. Ethylnitrosourea, in its reaction with purines, is useful as a structural probe of RNA.

4. Melting, Hybridization, and Polynucleotide Probes

Like proteins, nucleic acids can undergo denaturation. The strands of the double helix of DNA are separated and the double-stranded regions of RNA molecules "melt." Denaturation can be accomplished by addition of acids, bases, and alcohols or by removal of stabilizing counter ions such as Mg^{2+}. The product is a random coil and denaturation can be described as a helix → coil transition. Denaturation of nucleic acids by heat, like that of proteins, is cooperative (Chapter 7, Section A,3) and can be described by a characteristic **melting temperature**.

A plot of the optical absorbance at 260 nm (the wavelength of maximum light absorption by nucleic acids) versus temperature is known as a **melting curve** (Fig. 5-45). The absorbance is lower, by up to 40%, for native than for denatured nucleic acids. This **hypochromic effect** (Chapter 23) is a result of the interaction between the closely stacked bases in the helices of the native molecules. The melting temperature T_m is taken as the midpoint of the increase in absorbance (Fig. 5-45). As the percentage of G + C increases, the nucleic acid becomes more stable toward denaturation because of the three hydrogen bonds in each GC pair. T_m increases almost linearly with increases in the G + C content. In the "standard" citrate buffer (0.15 M NaCl + 0.015 M sodium citrate, pH 7.0) Eq. 5-22 holds. The exact numerical relationship depends strongly upon the ionic composition and pH of the medium.[37,72,552,553]

$$\%(\text{G} + \text{C}) = 2.44(T_m - 69.3);\ T_m \text{ in } ^\circ\text{C} \qquad (5\text{-}22)$$

The curves in Fig. 5-45 appear simple, but using newer apparatus and plotting the first derivative of the melting curve yields a complex pattern that depends on the sequence of bases.[555]

Complete denaturation of DNA leads to separation of the two complementary strands. If a solution of denatured DNA is cooled quickly, the denatured strands remain separated. However, if the temperature is held for some time just below T_m (a process known as **annealing**), the native double-stranded structure can be reformed. An important tool for studying DNA has been the measurement of the **kinetics of reassociation** of separated strands of relatively short DNA fragments.[72,556,557]

Because it depends upon the concentration of two separated strands, reassociation obeys second-order kinetics (Chapter 9) and Eq. 5-23, which is readily derived by integrating Eq. 9-8 for [A] = [B] = C from time 0 to *t*:

$$C / C_0 = 1 / (1 + k\,C_0 t) \qquad (5\text{-}23)$$

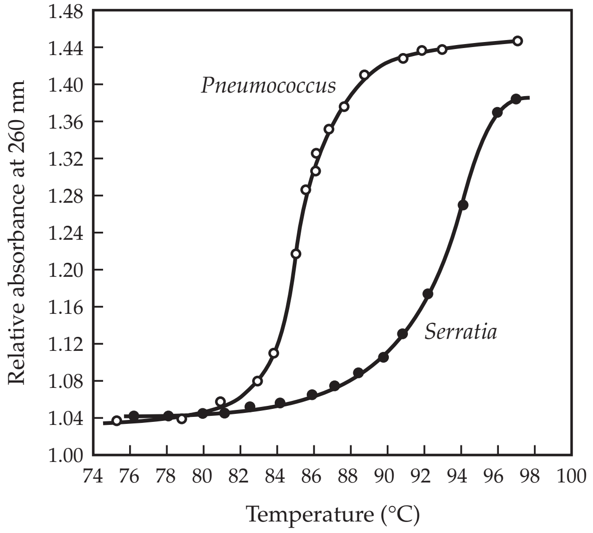


Figure 5-45 A melting curve for DNA molecules from two different sources. From Davidson.[554]

The initial concentration of denatured DNA, C_0, is related in this way to the concentration C of DNA remaining dissociated at time *t*. A plot of the fraction of molecules remaining single-stranded versus the logarithm of C_0t (Fig. 5-46A) is a convenient way of displaying data. As indicated in Fig. 5-46B, the value of C_0t increases in direct proportion to the length of the DNA chain in the genome, but it is very much decreased if the sequence of bases is highly repetitive [poly(T) and poly(A)]. The slope of the plot at the midpoint gives an indication of the heterogeneity of the DNA fragments in a solution.

Denatured DNA fragments can sometimes reassociate with DNA from a different source to form **hybrid duplexes**. Such double helices, in which one strand comes from one strain of an organism and the other strand from a genetic variant of the same organism or from a different species, are known as **heteroduplex**. Some mutations consist of **deletions** or **additions** of one or a substantial number of bases to a DNA chain. Heteroduplexes prepared from DNA of such mutants hybridized with that from a nonmutant strain have normal hydrogen-bonded Watson–Crick base pairs for the most part. However, they may have single-stranded loops in regions where long deletions or additions prevent complementary base pairing.

Hybridization measurements have been used in many studies of **homology** of nucleic acids from different species. A nucleic acid is cut (e.g., by sonic oscillation) into pieces of moderate length (~ 1000 nucleotides) and is denatured. The denatured DNA fragments are mixed with denatured DNA of another species. Nucleotide sequences that are closely similar between species tend to hybridize, whereas sequences that are drastically different between two species do not (Fig. 5-46). One way to do such an experiment is to immobilize the long-chain denatured DNA from the one organism by embedding it in an agar gel[558] or by absorbing it onto a nitrocellulose filter.[559,560] The DNA fragments from the second organism are passed through a column containing "beads" of the DNA-containing agar or through the filter with adsorbed DNA. Pairing of fragments with complementary sequences occurs

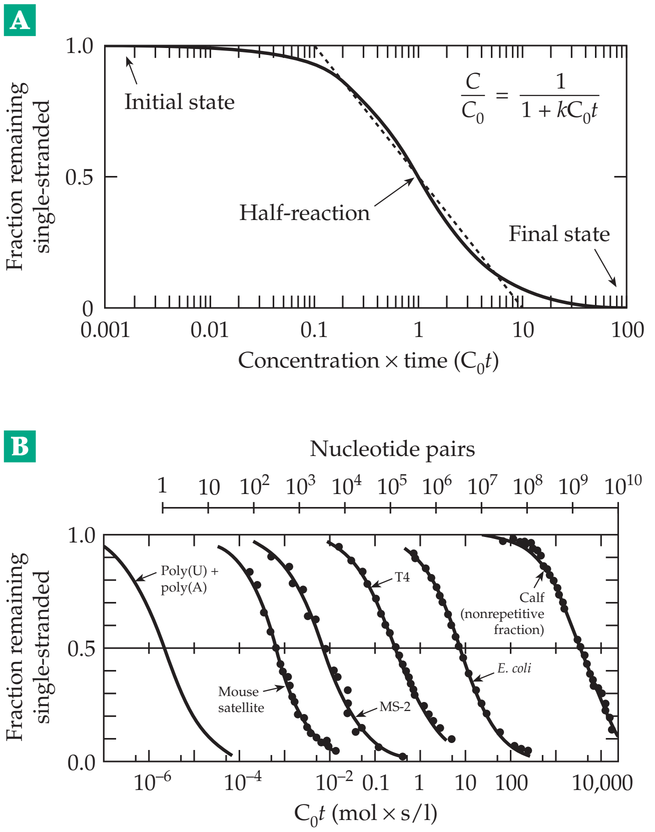


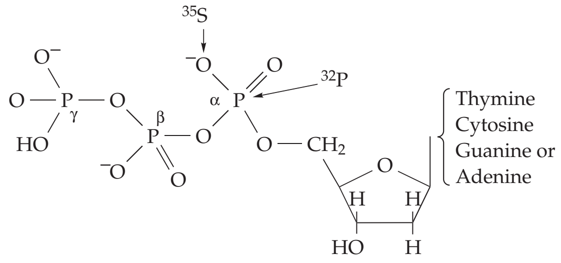
Figure 5-46 Reassociation curves for DNA from Britten and Kohne.[556,561] (A) Time course of an ideal, second-order reaction to illustrate the features of the log C_0t plot. The equation represents the fraction of DNA which remains single-stranded at any time after the initiation of the reaction. For this example, *k* is taken to be 1.0, and the fraction remaining single-stranded is plotted against the product of total concentration and time on a logarithmic scale. (B) Reassociation of double-stranded nucleic acids from various sources. The genome size is indicated by the arrows near the upper nomographic scale. Over a factor of 10^9, this value is proportional to the C_0t required for half-reaction. The DNA was sheared, and the other nucleic acids are reported to have approximately the same fragment size (about 400 nucleotides, single-stranded). Correction has been made to give the rate that would be observed at 0.18 M sodium-ion concentration. The temperature in each case was optimal, i.e., ~30°C below the melting temperature T_m. The extent of reassociation was established by measuring optical rotation (calf thymus DNA), ribonuclease resistance (MS-2), or hypochromicity.

and such paired fragments are retained while strands that do not pair pass on through the column (or filter).

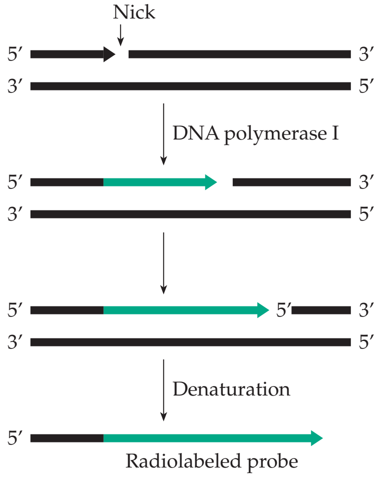
Both DNA hybrids and **DNA–RNA hybrid duplexes** are very important to present day genetic research.[560,562] Molecules of mRNA that represent transcripts of a particular gene will hybridize only with one of the two separated strands of DNA for that gene.

A major use of hybridization is to locate a gene or other DNA or RNA sequence by means of a **synthetic probe**.[563] This is a small piece of DNA or RNA which is labeled in some way, e.g., with a radioisotope such as ^{3}H, ^{32}P, or ^{125}I. Alternatively, the probe may carry a highly fluorescent dye or a "tag" that can be recognized by a specific antibody.[564,565] An example of the latter is the use of the vitamin biotin and the specific binding protein **avidin** (see Box 14-B).[566] Related procedures employ labeling with **digoxigenin** and often employ chemiluminescent detection.[567–569] Several methods for preparation of probes are in use. Some are enzymatic but the direct chemical synthesis of oligonucleotide probes is probably used the most.

One of the first methods devised for making a highly radioactive DNA probe is called **nick translation**.[570] A piece of dsDNA, e.g., a "restriction" fragment cut from a larger piece of DNA by restriction endonucleases, is selected. A small amount of pancreatic DNase I is added. It creates "nicks" in which one strand has been cut and some nucleotides have been removed leaving a gap. Now the DNA is incubated with DNA polymerase I (pol I) and a mixture of the four mononucleotide triphosphates, the precursors of biological synthesis of DNA (Chapter 27). Usually a ^{32}P or ^{35}S label with high radioactivity is present in one of the nucleotide triphosphates as indicated in the following structure. The polymerase fills the gap, adding nucleotides to the exposed 3' end of the nicked chain. *E. coli* pol I has a second enzymatic activity which allows it to digest a polynucleotide chain from the 5' end. Thus, as synthesis proceeds at the 3' side the nick is "translated" as shown in (Eq. 5-24).



The four nucleotide triphosphate precursors of DNA



(5-24)

Probes may also consist of DNA copied from mRNA. This is known as **cDNA** and is also widely used to determine indirectly the sequences of mRNA molecules. Messenger RNA may be isolated from the total cellular RNA by affinity chromatography on bound poly (dT) or poly (U). These materials selectively hold RNA with the poly (A) tails characteristic of most eukaryotic mRNA (see Chapter 28). Another source of mRNA is polyribosomes (polysomes), which are "reading" mRNA and actively making proteins. An antibody to the protein for which mRNA is desired will often bind to the protein chains being synthesized and precipitate the polysomes. The mRNA can be recovered and used as a template for cDNA.

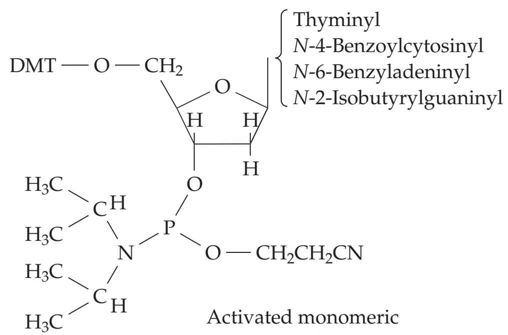
Synthesis of cDNA, usually in radiolabeled form is accomplished with **reverse transcriptase**, the enzyme from retroviruses that synthesize a DNA–RNA hybrid from ssRNA.[570–572] A short oligo (dT) primer is usually hybridized to the 3' poly (A) tail to initiate synthesis. Reverse transcriptase also has ribonuclease (RNase H) activity and will digest away the RNA. If desired, synthesis of the second strand can be carried out by a DNA polymerase to give a complete DNA duplex. Many gene sequences have been deduced from cDNA copies.

Often the most practical approach to obtaining a DNA probe is synthesis of a mixture of short oligonucleotides, often in radioactive form as described in the next section. The "redundancy" in the genetic code, i.e., the existence of two or more codons for most amino acids, presents a problem in designing an oligonucleotide probe based on amino acid sequence information. Examination of Table 5-5 suggests part of the solution. Whereas only Met and Trp have single unambiguous codons nine amino acids have only two codons each. We should try to find an amino sequence that contains Met and Trp and as many of the nine others as possible. We should avoid sequences that contain Ser, Leu, or Arg because each has six codons. We can then make a mixture of oligonucleotides, using the various coding possibilities. Mixtures of as many as 1024 (2^{10}) oligonucleotides have been used. One of these may bind tightly and specifically to the desired DNA segment. Instead of such a complex mixture it may be more useful to incorporate a modified base at the most ambiguous positions. For example, inosine, which occurs in the wobble position in anticodons (Fig. 5-30), can pair with A, C, or T.[573] Substitution of 2-aminoadenosine can cause a probe to bind more tightly because a third hydrogen bond will be present in each AT pair.[574]

Another important procedure is labeling ends of polynucleotides. Most often the 5' end is labeled with a radioisotope or by covalent attachment of a fluorescent dye. For example, a **polynucleotide kinase** can be used to transfer a radioactive γ-phospho group from ATP to the 5' end of a polynucleotide that has a free 5'-OH group.

5. Synthesis of Oligonucleotides and Polynucleotides

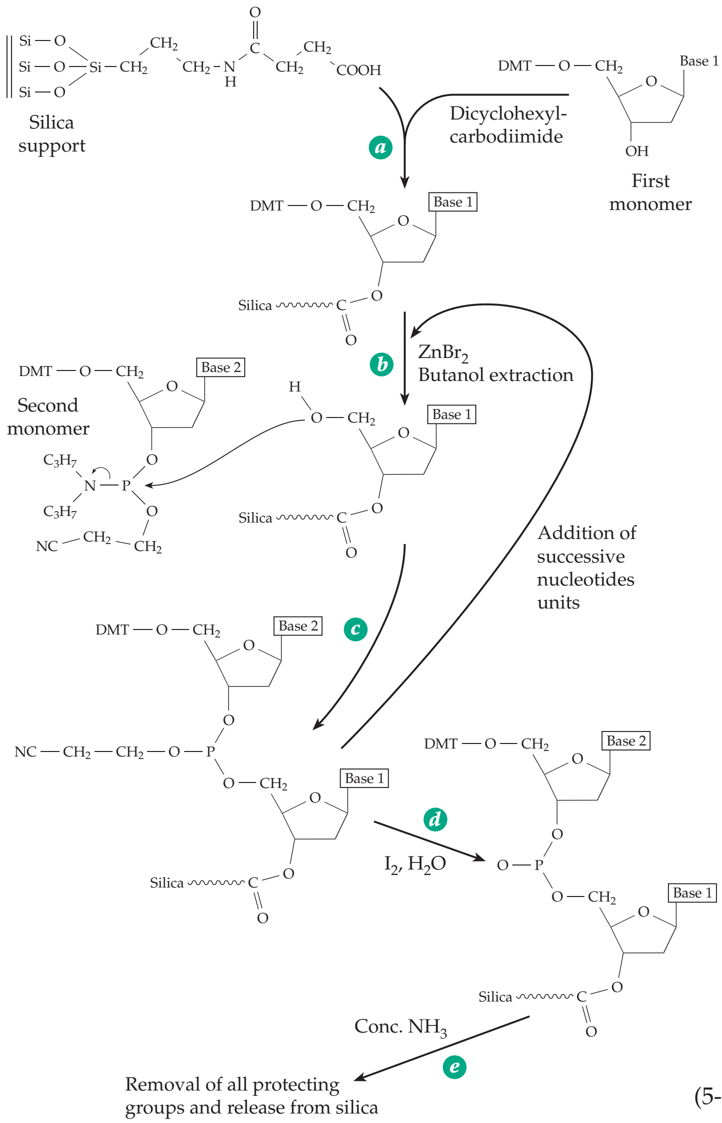
Efficient solid-phase methods of synthesis analogous to those for polypeptides (Fig. 3-15) have been devised. The pioneering work was done by H. G. Khorana, who made the first synthetic gene[575] and later synthesized a gene for the visual pigment rhodopsin (Chapter 23). Several synthetic approaches have been developed.[576] Currently the most popular method involves the use of phosphite esters. Most nucleophilic groups of the monomers are derivatized with removable blocking groups. For example, N-4 of cytosine and N-6 of adenosine may carry benzoyl groups. The 5'– OH of each nucleotide is blocked by a di-*p*-anisylphenylphenylmethyl (also called dimethoxytrityl, DMTO) group. The 3'–OH is converted to one of a number of activated derivatives such as the following *N*, *N*-diisopropylamino phosphines.[575,577–579]


DMT — O — CH₂
Thyminyl
N-4-Benzoylcytosinyl
N-6-Benzyladeninyl
N-2-Isobutyrylguaninyl
O
H
H
H
O
P
N
O — CH₂CH₂CN
H₃C
CH
H₃C
H₃C
CH
H₃C

Activated monomeric nucleotides for synthesis of polynucleotides

Solid-phase synthesis is usually done on a silica support with a covalently attached succinamide as shown in Eq. 5-25. The first nucleotide at the 3' end of the chain to be synthesized is attached by an ester linkage to the bound succinamide (step *a*, Eq. 5-25). The 5'-protecting group is removed in step *b* and the 5'-OH reacts with the activated phosphine of the second nucleotide (step *c*, Eq. 5-25). Steps *b* and *c* are then repeated as often as necessary to complete the chain. The finished polynucleotide can be removed from the solid support, the cyanoethyl groups removed from the phosphorus atoms by β elimination and all of the other blocking groups removed by treatment with concentrated NH_3. The whole procedure has been automated.[580–582]

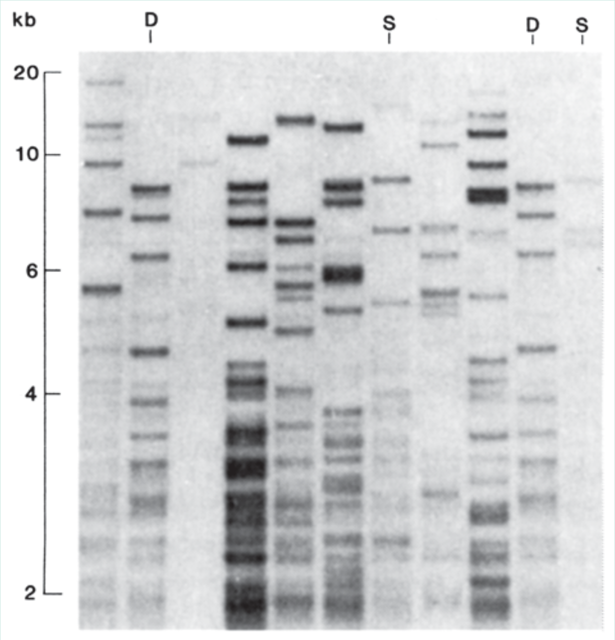
If a large piece of DNA is needed several oligonucleotides can be joined end to end enzymatically (Eq.


Si — O
Si — O — Si — CH₂
Si — O
CH₂
CH₂ – N
H
O
C
CH₂
CH₂
COOH
Silica support
DMT — O — CH₂
Base 1
O
OH
First monomer
Dicyclohexyl-carbodiimide
a
DMT — O — CH₂
Base 1
Silica
C
O
b
ZnBr₂
Butanol extraction
DMT — O — CH₂
Base 2
Second monomer
C₃H₇
N — P
C₃H₇
NC — CH₂ — CH₂
H
O — CH₂
Base 1
Silica
Addition of successive nucleotides units
c
DMT — O — CH₂
Base 2
NC — CH₂ — CH₂ — O — P
O — CH₂
Base 1
Silica
d
I₂, H₂O
DMT — O — CH₂
Base 2
O — P
O — CH₂
Base 1
Silica
Conc. NH₃
e
Removal of all protecting groups and release from silica

(5-25)

BOX 5-D DNA FINGERPRINTING

Jeffreys *et al.*[a–c] digested human DNA to completion with *Hin* f I and *Sau*3A restriction endonucleases. Certain fragments, which originated from "minisatellite" bands of repetitive DNA showed a very high degree of **polymorphism** among the population. Many different fragments sharing these repeated sequences were formed in the restriction digest. If a suitable labeled probe was used, it hybridized with as many as 80 different bands.[a,d–f] The resulting pattern appeared, like a fingerprint, to be different for every individual, as is shown in the accompanying photo. Unlike a fingerprint the DNA pattern also contains information that often allows deductions about parentage.



DNA "fingerprints" made from one or two drops of blood from ten different individuals. DNA was isolated, digested to completion with restriction endonuclease *Hin* f I, and subjected to electrophoresis in a 20-cm-long agarose gel until all DNA fragments smaller than 1.5 kb in length had passed off the gel. The DNA was then transferred to a nitrocellulose filter by Southern's method and hybridized with a ^{32}P-labeled single-stranded DNA probe prepared from cloned human minisatellite DNA. The probe used had the "consensus" composition (AGAGGTGGGCAGGTGG). Within the 29 tandem repeats in this 0.46-kb probe there are various sequences close to the one shown. Filters were then autoradiographed for four days. Two duplicate samples (marked D) were taken from the same individual and two others (marked S) from two sisters. A number of bands in common are evident. From Jeffreys *et al.*[c]

The technique has come into widespread use in forensic analysis, with DNA typing being possible from a single hair.[g] In a famous early case it was used to allow an immigrant child to be reunited with his mother[h] and it is being used regularly to protect innocent persons accused of rape or murder.[i] It is also widely used to provide evidence of guilt. However, the very small chance of a close match between unrelated persons prevents the use of DNA typing alone as proof of guilt. Because DNA samples are often "amplified" by PCR (Section H,6), there is also a possibility of contamination and forensic use of DNA typing is still controversial.[i–m]

However, DNA typing continues to be improved[f,n,o] and to be applied in a great variety of ways. For example, the skeletal remains of a murder victim were identified by DNA fingerprints after being buried for eight years.[p] DNA typing is also used to study mating habits of birds,[q] the genetic variability of populations of whales sampled by biopsy,[r] etc.

---

[a] Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985) *Nature (London)* **314**, 67–73

[b] Lewin, R. (1986) *Science* **233**, 521–522

[c] Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985) *Nature (London)* **316**, 76–79

[d] Vassart, G., Georges, M., Monsieur, R., Brocas, H., Lequarre, A. S., and Christophe, D. (1987) *Science* **235**, 683–684

[e] Huang, L.-S., and Breslow, J. L. (1987) *J. Biol. Chem.* **262**, 8952–8955

[f] Kirby, L. T. (1990) *DNA Fingerprinting*, Stockton Press, New York

[g] Higuchi, R., von Beroldingen, C. H., Sensabaugh, G. F., and Erlich, H. A. (1988) *Nature (London)* **332**, 543–546

[h] Jeffreys, A. J., Brookfield, J. F. Y., and Semeonoff, R. (1985) *Nature (London)* **317**, 818 – 819

[i] Balding, D. J., and Donnelly, P. (1994) *Nature (London)* **368**, 285–286

[j] Neufeld, P. J., and Colman, N. (1990) *Sci. Am.* **262**(May), 46–53

[k] Lewontin, R. C., and Hartl, D. L. (1991) *Science* **254**, 1745–1750

[l] Lander, E. S., and Budowle, B. (1994) *Nature (London)* **371**, 735–738

[m] Lewontin, R. C. (1994) *Nature (London)* **372**, 398

[n] Uitterlinden, A. G., Slagboom, P. E., Knook, D. L., and Vijg, J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2742–2746

[o] Jeffreys, A. J., MacLeod, A., Tamaki, K., Neil, D. L., and Monckton, D. G. (1991) *Nature (London)* **354**, 204–209

[p] Hagelberg, E., Gray, I. C., and Jeffreys, A. J. (1991) *Nature (London)* **352**, 427 – 429

[q] Burke, T., and Bruford, M. W. (1987) *Nature (London)* **327**, 149–152

[r] Hoelzel, A. R., and Amos, W. (1988) *Nature (London)* **333**, 305

27-5). For example, a functional 17-bp gene for the 53-residue human epidermal growth factor was synthesized by joining ten oligonucleotides of lengths 11–59 bp.[575] DNA is often synthesized enzymatically using methods described in Chapter 26. Cloned sequences of synthetic DNA can also be transcribed to produce **polyribonucleotides** of any desired sequence.[583] New nonenzymatic methods for RNA synthesis have also been devised.[583–587]

6. The Polymerase Chain Reaction (PCR)

This important technique was first described in 1971–1974 by Khorana and associates[588,589] but was not used until it was rediscovered in 1983 by Mullis.[590 – 593] It was quickly developed[591,592,594–596] and has played a major role in biochemistry ever since. It continues to be applied in numerous ways.[589,597–599a]

The PCR technique provides a way of “amplifying” a small number of DNA molecules, i.e., to produce many copies. This is often done by cloning but PCR offers a quick and easy way to obtain millions of copies of a desired relatively short segment of DNA. Standard PCR can be used for up to about 5000-nucleotide pieces. More recently modified procedures have allowed 35-kb segments to be amplified.[600]

The basic PCR procedure is initiated by hybridizing two oligonucleotide primers onto opposite strands of denatured DNA, one at each end of the section chosen for amplification (Fig. 5-47). A DNA polymerase is then used to convert each of the separated strands into a duplex. The mixture of products is then heated to denature the two new duplexes. After cooling, the primers, which are present in great excess, hybridize to all four strands. In a second cycle of polymerase action these are all converted to duplexes, etc. After 20 cycles millions of copies will be made. At first, copies with tails extending beyond the limits specified by the oligonucleotide primers will be formed. However, it is easy to see that after a few cycles, most molecules will be of just the desired length. A heat-stable polymerase from *Thermus aquaticus* (*Taq* polymerase) is used so that the enzyme is not denatured by the repeated cycles of heating and cooling, which are conducted automatically by a simple apparatus.

The polymerase chain reaction is being used to speed up prenatal diagnosis of genetic diseases, to detect viral infections, for tissue typing needed for organ transplantation, in forensic procedures, and in the study of the DNA of ancient tissues such as those of frozen wooly mammoths.[601 – 603] If suitable restriction enzyme sites are present in the primers, the amplified DNA can be cloned readily.[604] The 3.3 x 10^{-9} fmol of a DNA sequence found in a diploid chromosome pair in a single cell can be amplified in 50 cycles to 5–500 fmol, enough to study by hybridization with radioactive probes.[605] A large sample of a few pg of DNA can be amplified in 20 cycles to micrograms. By placing sequencing primers within the amplified segments, it is possible to generate DNA that can be sequenced directly using the dideoxy sequencing technique (Chapter 5) without cloning.[606] The PCR technique has also been used to amplify cDNA molecules formed from RNA transcripts present in very low abundance. One of the problems with the PCR is that priming may occur by DNA fragments other than the added primers. Contamination must be scrupulously avoided. Another problem is that errors are introduced into DNA during amplification by PCR. Perhaps 1 in 200 of the copies will contain an incorrect base.[607] If such a molecule is cloned the error will be perpetuated. Good practice requires that more than one clone is selected and sequenced to allow such errors to be avoided.

7. Sequence Determination

Satisfactory (but slow) methods for determining sequences of RNA molecules have been known for over 30 years. The procedures are somewhat parallel to those used in sequencing proteins. However, no similar method could be devised for DNA. Little progress was made until rather recently when new approaches led to extremely rapid procedures for sequencing DNA. As a consequence, it is now much easier to learn the sequences of genes than it is to sequence the proteins which they encode!

Preparing the DNA. The first step is to obtain a sample of enough identical DNA molecules to permit sequence analysis. This in itself may be a complex undertaking. Perhaps we want to know the sequence of one particular gene in the 3,500,000 kilobase pairs of DNA present in a single human cell. How can this gene be found and the DNA be obtained for analysis? Three techniques have been essential: cutting the DNA with restriction endonucleases, hybridization, and cloning. More recently PCR and related methods[608] have simplified the sample preparation. DNA can often be amplified using primers that contain sequences that will later serve as **sequencing primers**.

Restriction maps and Southern blots. Although it doesn’t require the synthesis of a primer, the Maxam–Gilbert procedure usually demands that a “restriction map” of the DNA be prepared to help keep track of the fragments being sequenced.[609] See also Chapter 26. Figure 5-48[610 – 612] shows the restriction map of the mitochondrial DNA gene, *oxi3* from yeast. This gene, which encodes one of the subunits of cytochrome oxidase (Chapter 18), consists of 9979 base pairs. It was cloned in a suitable plasmid after which the restriction map (Fig. 5-48) was prepared by cutting with 18 different restriction enzymes.[611] The protein subunit contains 510 residues and therefore requires a coding capacity in the DNA of 1530 base pairs. This is only 16% of the total length of the gene, the majority of whose DNA is found in the three large introns.

One way to select a desired segment of DNA from a digest of chromosomal DNA is to sort out the "restriction fragments" by gel electrophoresis.[613] The DNA from the gel can be transferred to a nitrocellulose sheet while retaining the separation pattern using a method devised by Southern.[560,614,615] In this **Southern blot** technique, solvent flows from a pool beneath the gel up through the gel and the nitrocellulose sheet into paper towels. The DNA is trapped on the nitrocellulose in the same pattern observed in the electropherogram. A suitably labeled probe such as cDNA with incorporated ^{32}P is flowed repeatedly across the nitrocellulose sheet under conditions that favor formation of hybrids. Only the DNA complementary to the cDNA probe will retain the label. This DNA can then be located with the help of a autoradiogram. It is important that single-stranded DNA be used. If double-stranded restriction fragments are separated on the electropherogram they must be denatured while in place in the gel before hybridization is attempted.

Once the desired piece of DNA has been identified it is usually necessary to increase its amount. The conventional approach is to incorporate the DNA fragment into a plasmid and clone by the methods described in Chapter 26. The selected DNA can usually be cut cleanly from the plasmid used for cloning with the same restriction endonuclease originally used in

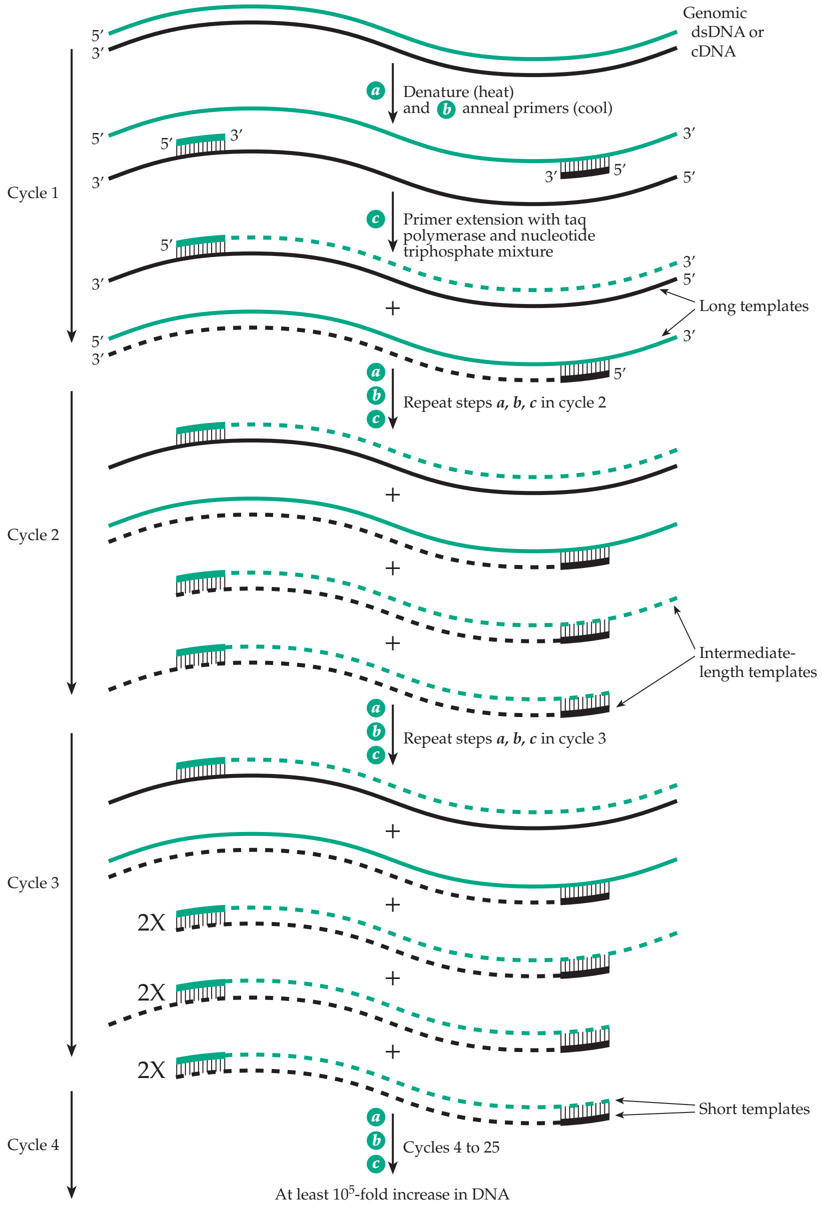

Genomic dsDNA or cDNA
5′
3′
a Denature (heat) and b anneal primers (cool)
5′
5′
3′
3′
3′
3′
5′
5′
Cycle 1
c Primer extension with taq polymerase and nucleotide triphosphate mixture
5′
3′
3′
5′
+
Long templates
5′
3′
3′
5′
a b c Repeat steps a, b, c in cycle 2
+
Cycle 2
+
+
Intermediate-length templates
a b c Repeat steps a, b, c in cycle 3
+
Cycle 3
2X
+
2X
+
2X
Short templates
Cycle 4
a b c Cycles 4 to 25
At least 10^5-fold increase in DNA

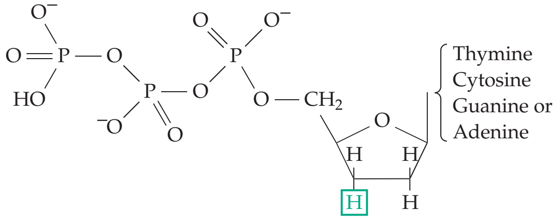
Figure 5-47 Amplification of DNA using the polymerase chain reaction (PCR). Double-stranded DNA is denatured by heating to 90–99° C (step *a*) and oligonucleotide primers complementary to short 12–18 nucleotide sequences at the two ends of the piece of DNA to be amplified are annealed to the separated strands by cooling to 40–75° C (step *b*). The two DNA strands serve as templates for synthesis of new complementary strands using a heat-stable DNA polymerase and a mixture of the four nucleotide triphosphates. Nucleotide units are added to the 3' ends of the primers, with the new chains growing in the 5' → 3' direction (step *c*). Steps *a*, *b*, and *c* are then repeated as many as 30 times using a thermal cycler device that periodically raises and lowers the temperature with a cycle time of a few minutes. The polymerase is unharmed by the heating and is reused in each cycle. An excess of the primer and of the nucleotide triphosphates sufficient for all of the cycles is present initially. In the early cycles new long and intermediate length templates are created. However, the number of short templates increases exponentially and the final product consists predominately of the short selected DNA segment (short templates).

fragmenting the DNA. An alternative procedure is to clone a mixture of DNA fragments and then sort colonies of bacteria containing the cloned fragments using DNA–RNA hybridization.[616] Bacteria from the selected colonies are then propagated to produce large amounts of the plasmid DNA. Alternatively, PCR can be used directly on the selected DNA fragment. This is more often the preferred choice.[604,606]

The Sanger dideoxy method. The rapid sequencing methods all depend upon the fact that single-stranded DNA fragments under denaturing conditions migrate on electrophoresis in polyacrylamide gels strictly according to their length. Thus, if a mixture contains all lengths of radiolabeled polynucleotides from very short oligonucleotides to fragments containing 200 or 300 bases, the polynucleotides will all appear, one above the other, as a series of bands that can be visualized by radioautography. The first of these methods was published by Sanger and Coulson[617] in 1975 and was followed in 1977 by the method which is now used.[618–621]

A sample of double-stranded DNA is denatured. One of the resulting single strands is used as a template to direct the synthesis of a complementary strand of radioactive DNA using a suitable DNA polymerase. The "Klenow fragment" of *E. coli*, DNA polymerase I, reverse transcriptase from a retrovirus, bacteriophage T7 DNA polymerase, *Taq* polymerase, and specially engineered enzymes produced from cloned genes have all been used.

Before the sequencing begins it is necessary to prepare a short **primer** that is complementary to a sequence at one end of the DNA strand to be sequenced. This may be prepared enzymatically,[622,623] or by nonenzymatic synthesis. The short primer is annealed to the end of the DNA and the resulting molecule is incubated with a DNA polymerase and a mixture of the four mononucleotide triphosphates, one of which is radiolabeled in this position. Four reaction mixtures are prepared. Each mixture contains all four nucleoside triphosphates and also one of four different **chain-terminating inhibitors**, the most popular of which are the 2', 3'–dideoxyribonucleoside triphosphates:



2',3'-Dideoxyribonucleotide triphosphates

These inhibitors are added in a ratio of about 100:1 with the natural substrates. In this ratio they are incorporated into the growing DNA chain about once in 200 times on the average. However, in the various growing DNA chains they are incorporated at different points ranging from the very first nucleotide to the last. Since the incorporated dideoxy monomer lacks the 3-hydroxyl group needed for polymer formation, chain growth is terminated abruptly. Synthesized polynucleotides are denatured and subjected to electrophoresis in four adjacent lanes. The resulting patterns contain bands corresponding to all of the successive oligonucleotides but not all in the same lane. A given lane will contain only the bands of the oligonucleotides terminated by the particular inhibitor used. The other bands will be found in the other three lanes. Each band will have been terminated by the inhibitor employed in that lane. The nucleotide sequence can be read directly from the banding pattern as is shown in Fig. 5-49. Arabinosyl nucleotide triphosphates have also been used as chain-terminating inhibitors. A sequence determined by the Sanger method is usually checked by also sequencing the complementary strand.

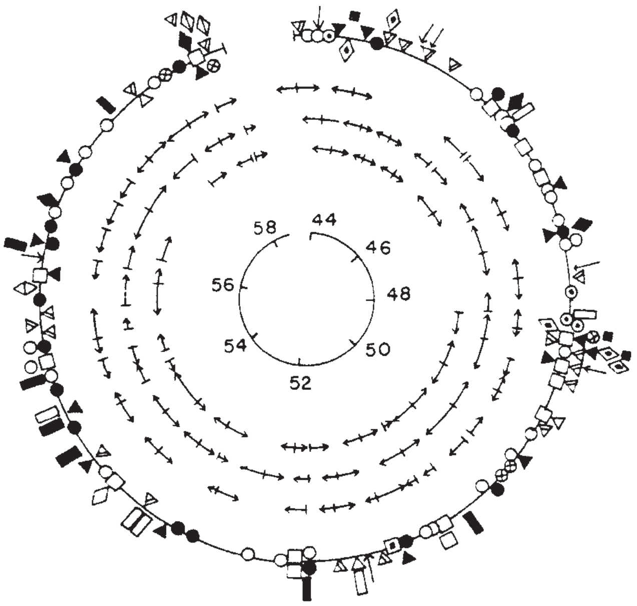


Figure 5-48 A physical map of the *oxi3* locus of yeast mitochondrial DNA. The restriction fragments used for DNA sequencing are indicated by the *arrows*. The extent to which the sequences were read is represented by the lengths of the arrows. The map units are shown in the *inner circle*. The following symbols, together with the names of the restriction enzymes (Chapter 26), are used for the restriction sites:

▲ *Hinf* I
△ *Hpa* II
▯ *Hae* III
◆ *Taq* I
● *Mbo* I
□ *Mbo* II
○ *Alu* I
⟐ *Pvu* II
◈ *Hinc* II
⬖ *Hind* III
⊗ *Eco* RI
↓ *Eco* RII
⊙ *Hha* I
◬ *Rsa* I
▮ *Hph* I
⊡ *Blg* II
◇ *Bam* HI
■ *Hpa* I

From Bonitz *et al.*[611]

A major factor in the success of the dideoxy sequencing method has been the development of cloning techniques that provide ssDNA in a form ready for use. In any cloning procedure the DNA that is to be sequenced has been covalently ligated to the end of a DNA strand of the **cloning vehicle**, a modified plasmid or virus (Chapter 26). Sequencing is often done on DNA cloned in a modified ssDNA bacterial virus such as M13. Although dsDNA is ligated to the ds replicating form of viral DNA, the virus particles produced when the virus is propagated in *E. coli* cells are single stranded. Cleavage of the viral DNA with its incorporated "passenger DNA" with an appropriate restriction endonuclease releases the passenger DNA (to be sequenced) with a short piece of DNA from the cloning vehicle attached at the 3' end. Since the sequence of this small piece of the cloning vehicle is known, a suitable primer of length ~12–18 nucleotides can be synthesized (or purchased) and annealed to the DNA. This serves as the primer and allows the sequence to be read from the 5' end to the 3' end of the synthesized complementary strand. Double-stranded DNA attached to vehicles such as the pUC plasmids can also be sequenced directly if the DNA is denatured by alkali treatment. After neutralization and precipitation an appropriate primer is annealed to one or the other of the two strands.[624]

Since its introduction, many modifications and improvements have been developed. The sequencing gels have been improved. The use of ^{35}S labeling has given sharper autoradiographs.[625] Alternatively, a silver stain can be used with unlabeled primers.[626] GC-rich DNA sequences are often difficult to sequence, probably because even in the denaturing polyacrylamide gels used for sequencing they tend to form hairpin loops, perhaps as a result of formation of Hoogsteen base pairs (Fig. 5-7). Formation of these loops results in uneven spacing between the adjacent bands in the sequencing gel, so-called "compression artifacts." Use of a 7-deaza-dGTP in place of dGTP in the sequencing reaction ameliorates this problem.[627] Sensitivity can be improved by use of "cycle sequencing" in which a heat-stable polymerase such as *Taq* polymerase is used, and after heating the same template DNA is used repeatedly to give a higher yield of labeled fragments.[628,629]

About 200 – 400 bases can be successfully sequenced manually in a single run. By prolonging the time of electrophoresis in a second run, the sequence can be extended considerably. By using the just obtained sequence information, it is possible to select a new start point 200 or more nucleotides further along the template chain and to synthesize an oligonucleotide primer to anneal to the template at this point. In this way it is possible to "walk" along the template adding additional sequences at each step. Another procedure is to delete by mutation various segments of the cloned DNA above the sequence that binds the primer. This

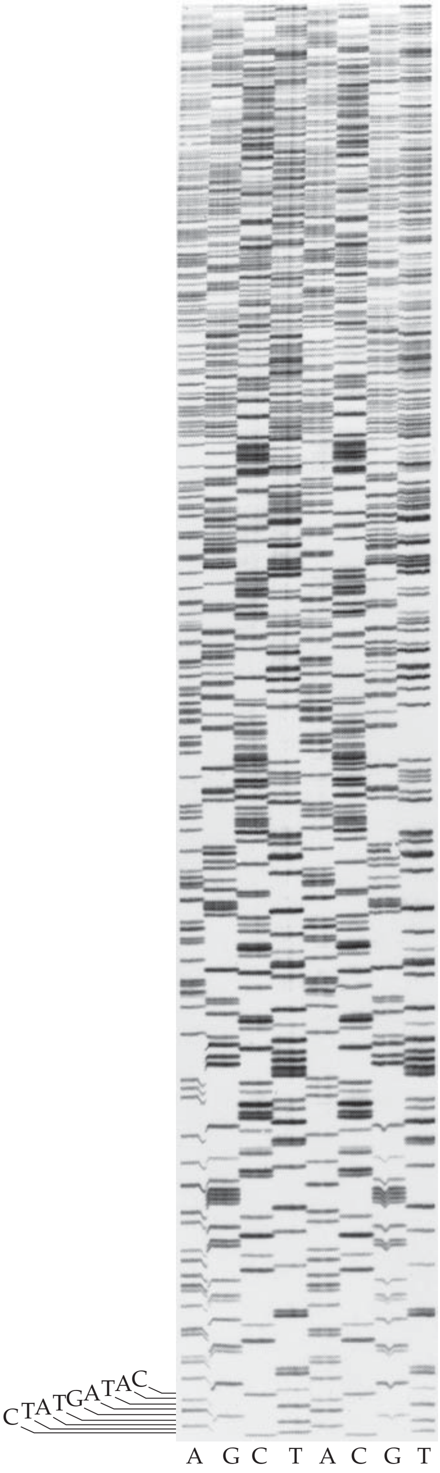


Figure 5-49 A DNA sequencing gel obtained using a segment of DNA from salmon sperm selected by suitable oligonucleotide primers, amplified by PCR, and sequenced with a ^{35}S label in the primer. Four samples were used, one with each of the four dideoxy chain terminators (A, G, C, T, A, C, G, T from left to right). After electrophoresis the shorter fragments are at the lower end of the gel. The sequence of the strand complementary to the template strand whose sequence is being determined is read from the bottom of the gel. Here it starts CTATGATAC. Reproduced by permission of Amersham Pharmacia Biotech, Limited.

permits analysis of the whole cloned fragment via an overlapping set of sequences from the deletion mutants.[630,631]

Many DNA sequences continue to be determined manually by the well-developed long gel procedures as illustrated in Fig. 5-49. However, sequencing whole genomes has depended upon the development of high-speed automated procedures.[632] Instead of radiolabeling, fluorescent dyes may be joined to the primer to allow detection of the chain fragments produced during sequencing. Automatic sequencers use four dyes that fluoresce with different colors.[633,634] A different dye is used for each of the reaction mixtures. Then the four samples are mixed together and the DNA fragments are subjected to electrophoresis in a single lane. A laser beam excites the fluorescence, scanning several lanes with different samples as the electrophoresis progresses. A photomultiplier tube records the fluorescence intensity of each band through a series of four filters in a rotating wheel. This allows automatic recognition of the four different colors of fluorescence and therefore of the nucleic acid base present in each position in the sequence. Improved strategies for "primer walking"[635] and for "shotgun sequencing"[19,632,636] have been devised. In the shotgun strategy, whole bacterial genomes have been cut by restriction enzymes into large numbers of overlapping fragments which have been separated and sequenced.[637] A computer program is used to analyze and assemble the sequences into a complete genomic sequence. An example is provided by the genome of the *Methanococcus jannaschii*. Its large circular chromosome contains 1,664,976 bp and there are two additional pieces to the genome, one containing 58,407 bp and the other 16,550 bp. The sequences were deduced from 36,718 individual sequencing runs on high-speed automatic sequencers. For each run, on average, 481 bp could be read.[636] To sequence the human genome faster methods are needed.[638] Capillary electrophoresis with a single laser beam scanning the output of 24 capillaries has been demonstrated.[639,640] Extremely rapid sequencing of oligonucleotides up to 100 bp in length can be accomplished by mass spectrometry.[641] This may be an important technique for diagnosis of genetic defects (Chapter 26).

The method of Maxam and Gilbert. The nonenzymatic method devised by Maxam and Gilbert[642–644] can be used to sequence either ss or dsDNA. Before the sequencing is begun, a radioactive label is incorporated, usually at the 5' end. This is often done by cleaving off any phosphate groups present on the 5' end with alkaline phosphatase and then transferring a new radioactive phospho group with the assistance of the enzyme polynucleotide kinase and radioactive γ-^{32}p-labeled ATP. If dsDNA is used the strands are separated so that each has a label only at one end.

The key step in sequencing by the Maxam–Gilbert procedure is to cleave chemically the DNA at random locations using reagents that have some specificity for particular bases. The cleavage process involves three distinct steps: (1) chemical modification, as specific as is possible for the chosen base; (2) displacement of the modified base from the sugar; and (3) elimination and chain cleavage using amine catalysis (Eq. 5-21). Two consecutive steps can often be combined. There are several versions of the method; one involves dimethyl sulfate as the specific reagent for guanine. It forms N^7-methylguanosine (see Eq. 5-18) which upon heating with the strong base piperidine at 90°C undergoes addition of hydroxyl ion with ring opening and displacement of the modified base according to Eq. 5-26. The product, a glycosylamine of piperidine, is in equilibrium with a Schiff base which can undergo chain cleavage as in Eq. 5-21.

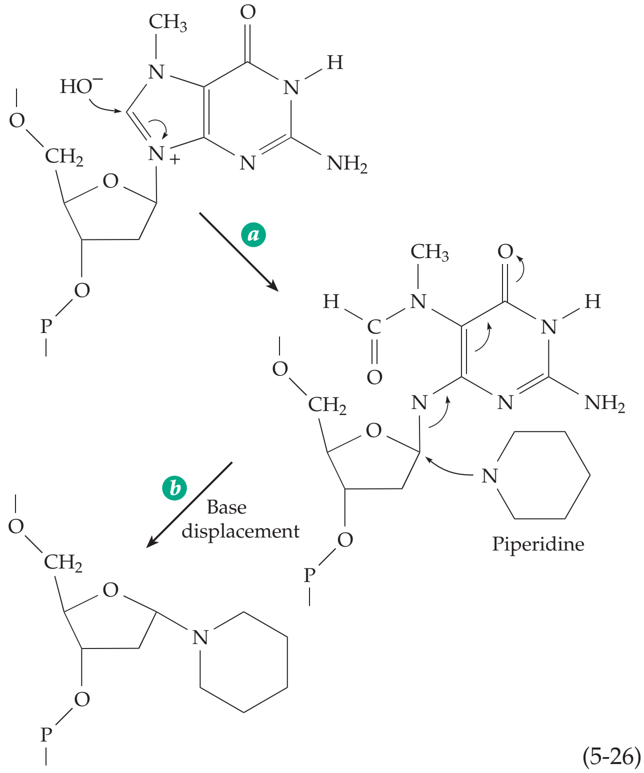
A second sample of DNA is treated with a piperidine-formate buffer of pH 2 in the cold. This promotes the acid depurination of both guanine and adenosine. A third sample is treated with hydrazine, with both cytidine and thymidine being cleaved to ribosylurea according to Eq. 5-13. Again, this is followed by displacement and β elimination (Eq. 5-21) catalyzed by piperidine. The fourth sample is also treated with hydrazine but in the presence of a high concentration of NaCl which inhibits the reaction with thymidine, by lowering the pK_a of the thymine, and allows the cleavage to be more nearly specific for cytidine. Each of the reactions is conducted in such a way that on the average only one cleavage event occurs per molecule of DNA. Since the cleavages occur at many different points, a family of nested radioactively labeled oligonucleotides, one from each original molecule, is produced. When these are sorted by polyacrylamide gel electrophoresis, the pattern of the oligonucleotides in the four adjacent channels allows the nucleotide sequence to be read directly from the autoradiogram.

The Maxam–Gilbert method doesn't require synthesis of a primer and it sometimes works well for sequences that are difficult to obtain with the Sanger–Coulson procedure. The two methods may both be used to provide additional certainty about a sequence. The Maxam–Gilbert method is very convenient for sequencing small oligonucleotides which often react poorly with the polymerase used for the chain termination method. The method usually requires that a restriction map be prepared.

Sequencing RNA. The first known RNA sequence, that of an alanine tRNA, was determined by Holley and associates in 1965. The RNA was subjected to partial hydrolysis with pancreatic ribonuclease and ribonuclease T_1 (Fig. 5-43). The small oligonucleotide fragments were separated by ion exchange chromatography under denaturing conditions (7 M urea) and were then characterized individually.[645] The availability of additional enzymes such as ribonuclease U_2, the *B cereus* ribonuclease and ribonuclease Phy M of *Physarum* (Fig. 5-43) and the use of radiolabeling and of two-dimensional fingerprinting of digests have made the procedures more versatile.[646] A valine tRNA was sequenced independently by Bayev[647] and Campbell.[648]

Since the development of the rapid methods for sequencing DNA, many mRNA sequences have been determined by using reverse transcriptase to make a cDNA strand complementary to the RNA. The cDNA is then sequenced.[649] Rapid sequencing methods parallel to those used for DNA have also been devised.[650–652]

Nearest neighbor analysis. A technique developed by Kornberg and associates before the availability of sequencing methods is the **nearest neighbor sequence analysis**. Using a single radioactive ^{32}P-containing nucleoside triphosphate together with the three other unlabeled nucleoside triphosphates, a primer chain of DNA is elongated from the 3' end along a ssDNA template chain using a DNA polymerase. The incorporation of ^{32}P from the α position of the nucleotide triphosphate occurs in the bridge phosphates that connect the nucleotide originally carrying the ^{32}P to the 3' position of the neighboring nucleotide. Cleavage of the ^{32}P-containing product of the reaction with a mixture of micrococcal DNase and spleen phosphodiesterase, which catalyze *b*-type cleavage (Fig. 5-43), gives fragments in which the ^{32}P will now be attached to what was the 5′ nearest neighbor to the radioactive nucleotide in the DNA.[72,653] Measurement of the radioactivity in each of the 3' nucleotides of thymine, cytosine, adenine, and guanine gives the frequencies of the adjacent pairs, TA, CA, AA, and GA. Using the other radioactive nucleoside triphosphates one at a time in separate experiments, all of the nearest neighbor frequencies can be obtained. From such an experiment it was possible to deduce that the strands in the double helix were oriented in an antiparallel fashion, as predicted by Watson and Crick. If the strands had been parallel, different nearest neighbor frequencies would have been observed.



(5-26)

Understanding sequences. Sequences of over 20 million nucleotides from hundreds of organisms had been determined by 1988 and the number is doubling each 2–3 years.[654] Sequencing the human genome has required rates of millions of bases per day. With the massive amount of data already available it has become of great interest to compare sequences of genes, whether they encode similar or dissimilar proteins, to make comparisons between species, and to search for sequences that bind specific proteins or that encode particular regulatory signals. Relationships of common evolutionary origin or homology as well as other sequence similarities are often sought.

To handle the mass of existing data, powerful computer programs have been developed and various graphical procedures have also been developed to help the human mind comprehend the results.[654,655] One important problem is to define and locate what are called **consensus sequences**. The problem is best illustrated by examples.[654] The cleavage site for the *Eco*RI restriction endonuclease is **GAATTC**. There is no ambiguity. In a DNA of random sequence this would be expected to occur by chance in about $(1/4)^6$ nucleotides (4 kb). On the other hand, the *Hin*II restriction endonuclease cleaves within the consensus sequence **GTYRAC** where Y = C or T and R = A or G. It would be expected to occur by chance in about $1/4^5$ nucleotides. Many binding sites for RNA polymerase, the so called **promoters** (Chapter 28) contain the consensus sequence **TAtAaT**, at position -10, ahead of the 5' end of the sequence that is transcribed into mRNA. The lower case t and a used here imply that other nucleotides may often replace T or A at these positions. There are many promoters and over 70% of those described have this consensus sequence. All have the less restricted sequence **TAxxxT**, where x may be any nucleotide. Our definition of consensus sequence is somewhat arbitrary. Now consider the problem of locating a −35 site whose consensus sequence is **TTGACA** but which may, for different genes, be shifted backward or forward by a nucleotide or two. This is a consensus sequence. Therefore, in many cases one or more substitutions in the sequence will have been made. The result is that the sequence of nucleotides in which the consensus sequence is to be found is likely to appear entirely random. Sophisticated computer programs are helpful in locating it.[654]

8. Protein–DNA Interactions

The most detailed information about interactions of proteins with DNA is coming from X-ray crystallographic studies. Examples are seen in Figs. 5-35 to 5-40. Several other methods have also been very useful. Much has been learned from the effects of mutations in DNA-binding proteins or in regions of DNA to which a protein binds. Binding of proteins to DNA can also be recognized by its effects on the mobility of DNA during gel electrophoresis.[656,657] Chemical[658] or laser-induced crosslinking can show that within a complex a specific residue in a protein is adjacent to a certain sequence in the DNA.

The technique of **protection mapping** or **"footprinting"** is widely used to determine which nucleotides in a sequence are covered by a bound protein.[659] A reagent which attacks and cleaves DNA nearly randomly is used. DNase I was first introduced for this purpose[659] and has been used widely. An important finding is that certain sites that are readily cleaved (**hypersensitive sites**) are frequently located in chromatin undergoing transcription. Footprinting has also been accomplished with other nucleases, with dimethylsulfate (which acts on A and C), with carbodiimides (which act on U and G), and with the Maxam–Gilbert guanine-specific cleavage (Eq. 5-26). One of the most popular methods employs cleavage by hydroxyl radicals.[660–663] **Photofootprinting** depends upon decreased or increased sensitivity to ultraviolet light at sites bound by proteins.[214,662,664] In footprinting experiments the DNA to be studied is radioactively labeled at one end of one strand. In the absence of the protecting protein, denaturation and electrophoresis of the cleaved fragments yields a nearly random "ladder" of DNA fragments. In the presence of the binding protein some cleavage products will be missing from the ladder. The bound protein leaves a "footprint" (Fig. 5-50A,C).

Related methods are being applied to the determination of the secondary structure of RNA molecules[665,666] and to the study of interactions with proteins. For example, treatment with dimethyl sulfate under appropriate conditions methylates bases that are not paired, giving largely 1-methyladenosine and 3-methylcytidine.[667]

9. Nuclear Magnetic Resonance

Much of the initial effort to study polynucleotides by NMR spectroscopy was directed toward transfer RNAs, only a few of which have been crystallized in a form suitable for X-ray diffraction. Study of the other tRNAs by NMR techniques has established that all of the tRNAs have a similar architecture and that the structures observed in the crystals are preserved in solution.[668] Figure 5-51 shows the low-field end of the NMR spectrum of a valine-specific tRNA from *E. coli.* The spectrum is run in H_2O rather than D_2O so that exchangeable hydrogens in the hydrogen bonds of the Watson–Crick base pairs can be observed.[669] The protons giving rise to the downfield resonances are primarily those attached to nitrogen atoms of the rings and in hydrogen-bonded positions. These protons are shielded by adjacent electron-donating groups and by their attachment to the semiaromatic rings of the bases. The NMR signals are further shifted downfield to varying degrees depending upon whether or not the proton being observed is attached to a base that is stacked with other bases. The size of the shift also depends upon which neighboring bases are present. The proton on N-3 of AU base pairs is deshielded more than the proton on the N-1 of GC base pairs. Therefore, the AU protons appear further downfield than the GC protons. The stronger ring current in A than in C enhances this separation.

All of the resonances in Fig. 5-51 have been assigned to particular bases. This was done in part by varying the temperature, changing the magnesium ion concentration, and predicting shifts caused by ring currents in adjacent bases making use of the X-ray crystal structures. NMR spectra of hairpin helical fragments also provided essential information. From integration of the areas under the peaks it was concluded that the 20 resonances seen below –11 ppm represent 27 protons. Twenty of these are in Watson-Crick base pairs and correspond to those expected from the X-ray structure. Six more belong to protons involved in tertiary interactions, such as base pair triplets or non-Watson–Crick pairs. One of these, labeled "G" in the figure, is in the dihydrouridine stem and involves the ring proton of N-1 of m^7G46 which is hydrogen bonded to N-7 of G22 in the major groove of the RNA. G22 is located at the beginning of the "extra loop."

Measurements of the NOE of nearby protons in both small RNA molecules[668,669] and DNA oligonucleotides[670,671] provided much additional information. Figures 5-51B and C show NOESY spectra of the $tRNA^{Val}$ and the way in which weak cross-peaks between the H-bonded imino protons in adjacent base pairs (Fig. 5-51B) can be used to establish connectivities.[669] Beginning with resonance B, it is possible to establish the sequence of the NH groups giving rise to these resonances as OBUGJNT. Using other data as well, it was concluded that these represent the seven base pairs of the acceptor stem (see Fig. 5-30), with resonance C representing the first GC pair, resonance B the second, etc. Resonance A was identified as coming from the base triplet containing a Hoogsteen

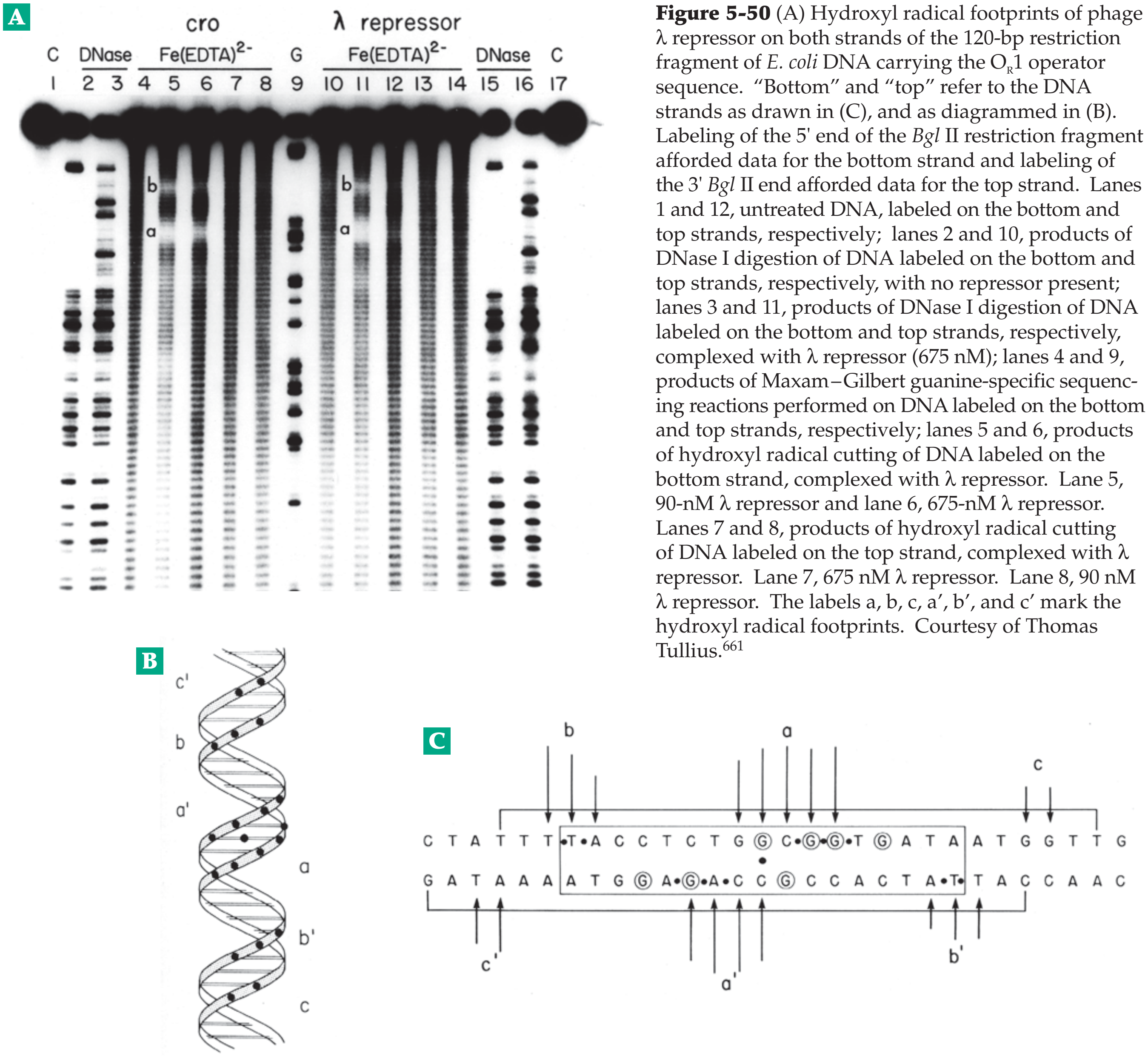

A
cro
λ repressor
C DNase Fe(EDTA)2- G Fe(EDTA)2- DNase C
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
b
a
b
a
B
c'
b
a'
a
b'
c
C
b
a
c
C T A T T T•T•A C C T C T G Ⓖ C•Ⓖ•Ⓖ•T Ⓖ A T A A T G G T T G
G A T A A A A T G Ⓖ A•Ⓖ•A•C C Ⓖ C C A C T A•T• T A C C A A C
c'
a'
b'

Figure 5-50 (A) Hydroxyl radical footprints of phage λ repressor on both strands of the 120-bp restriction fragment of *E. coli* DNA carrying the O_R1 operator sequence. "Bottom" and "top" refer to the DNA strands as drawn in (C), and as diagrammed in (B). Labeling of the 5' end of the *Bgl* II restriction fragment afforded data for the bottom strand and labeling of the 3' *Bgl* II end afforded data for the top strand. Lanes 1 and 12, untreated DNA, labeled on the bottom and top strands, respectively; lanes 2 and 10, products of DNase I digestion of DNA labeled on the bottom and top strands, respectively, with no repressor present; lanes 3 and 11, products of DNase I digestion of DNA labeled on the bottom and top strands, respectively, complexed with λ repressor (675 nM); lanes 4 and 9, products of Maxam–Gilbert guanine-specific sequencing reactions performed on DNA labeled on the bottom and top strands, respectively; lanes 5 and 6, products of hydroxyl radical cutting of DNA labeled on the bottom strand, complexed with λ repressor. Lane 5, 90-nM λ repressor and lane 6, 675-nM λ repressor. Lanes 7 and 8, products of hydroxyl radical cutting of DNA labeled on the top strand, complexed with λ repressor. Lane 7, 675 nM λ repressor. Lane 8, 90 nM λ repressor. The labels a, b, c, a', b', and c' mark the hydroxyl radical footprints. Courtesy of Thomas Tullius.[661]

base pair of 4-thiouracil at position 8 with A14 (see Fig. 5-7). Its connectivity to the sequence KCEO is also outlined in Fig. 5-51B. However, it could not be established without additional data which also helped to identify the sequence O-K as residues 10–13 of the dihydrouracil stem. Peak O is a multiproton peak representing not only GC 10 but also UA 7. In general, the GC protons are at the higher field side of the spectrum and the AU protons at the lower side. However, AU 7 is shifted to an anomolously high position. Cross-peaks between imino protons of uracil and the nearby C2 protons of adenine in Watson–Crick AU base pairs or C8 protons of Hoogsteen AU pairs can be observed in the 6.5 – 9 ppm region as shown in Fig. 5-35B. This region also contains information about other protons bound to the nuclei acid bases.

Similar techniques are being used for the study of DNA.[672] The presence of a second hydrogen in the 2' position of the deoxyribose rings of DNA adds several H-H distances (Fig. 5-52) that can be measured in addition to those seen in RNAs. Characteristic differences are seen in the NOESY plots of A, B, and Z forms of DNA.[670,671,673,674] Although detailed structural information has been obtained for short segments of DNA, spectra of larger oligonucleotides are impossible to analyze with two-dimensional methods because of extensive overlap of resonances.[665] The difficulty is already apparent in the 17 base pair DNA segment for which a one-dimensional spectrum as well as COSY and NOESY spectra are shown in Fig. 5-53.

Some help with the complexity can be obtained by incorporation of ^{13}C-enriched methyl groups into

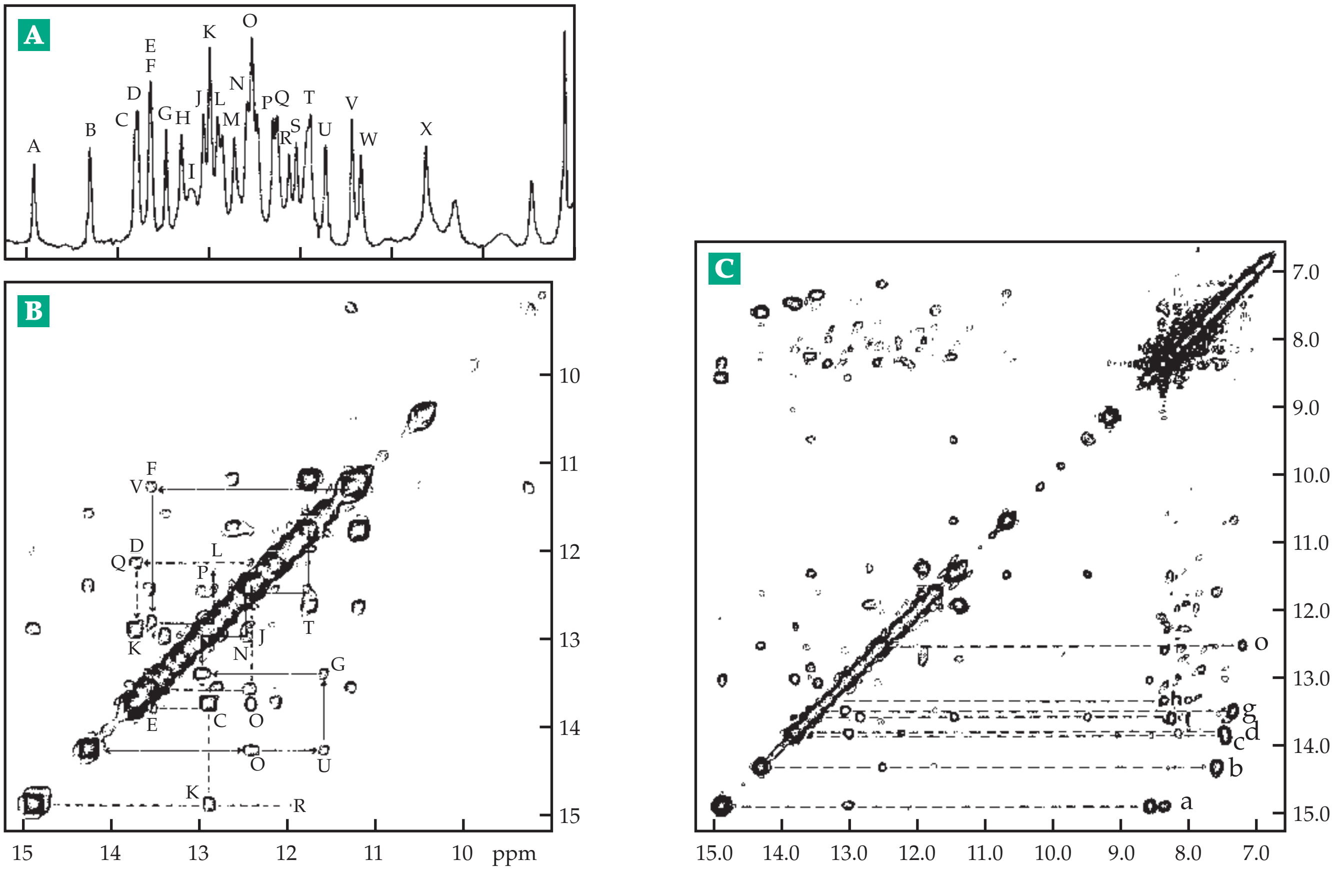


Figure 5-51 (A) The low-field region of the one-dimensional ^{1}H NMR spectrum of *E. coli* tRNA$_1^{\mathrm{Val}}$ at 27°C in H$_2$O. Resonances are identified by letters A – X. (B) NOESY spectrum of the same tRNA under similar conditions showing the imino-imino NOEs. In the lower right sector the connectivity traces of the acceptor helix and dihydrouridine helix are shown as solid and dotted lines, respectively. In the NOESY sample the two protons in peak EF are partially resolved whereas the two protons in peak T have coalesced. (C) NOESY spectrum of *E. coli* tRNA$_1^{\mathrm{Val}}$ at 32°C showing the imino and aromatic proton regions. AU-type imino protons have been connected horizontally by a dotted line to the cross-peak of their proximal C2-H or C8-H in the 7 to 9 ppm region, which has been labeled with the corresponding lower-case letter. From Hare *et al.*[669] Courtesy of Brian Reid.

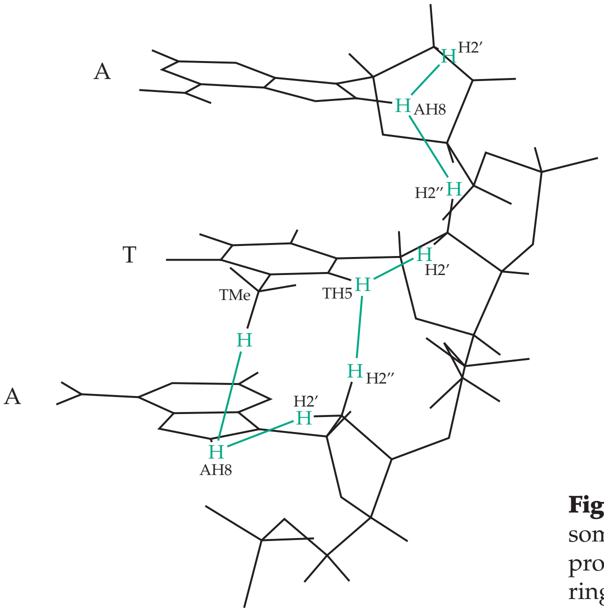


Figure 5-52 Segment of a DNA chain in the B conformation illustrating some intrachain NOEs that may be observed. The close juxtaposition of proton pairs is provided by the H2'-*endo* (^{2}E) conformation of the sugar rings with *anti* base conformation (Fig. 5-11). After Cohen.[675]

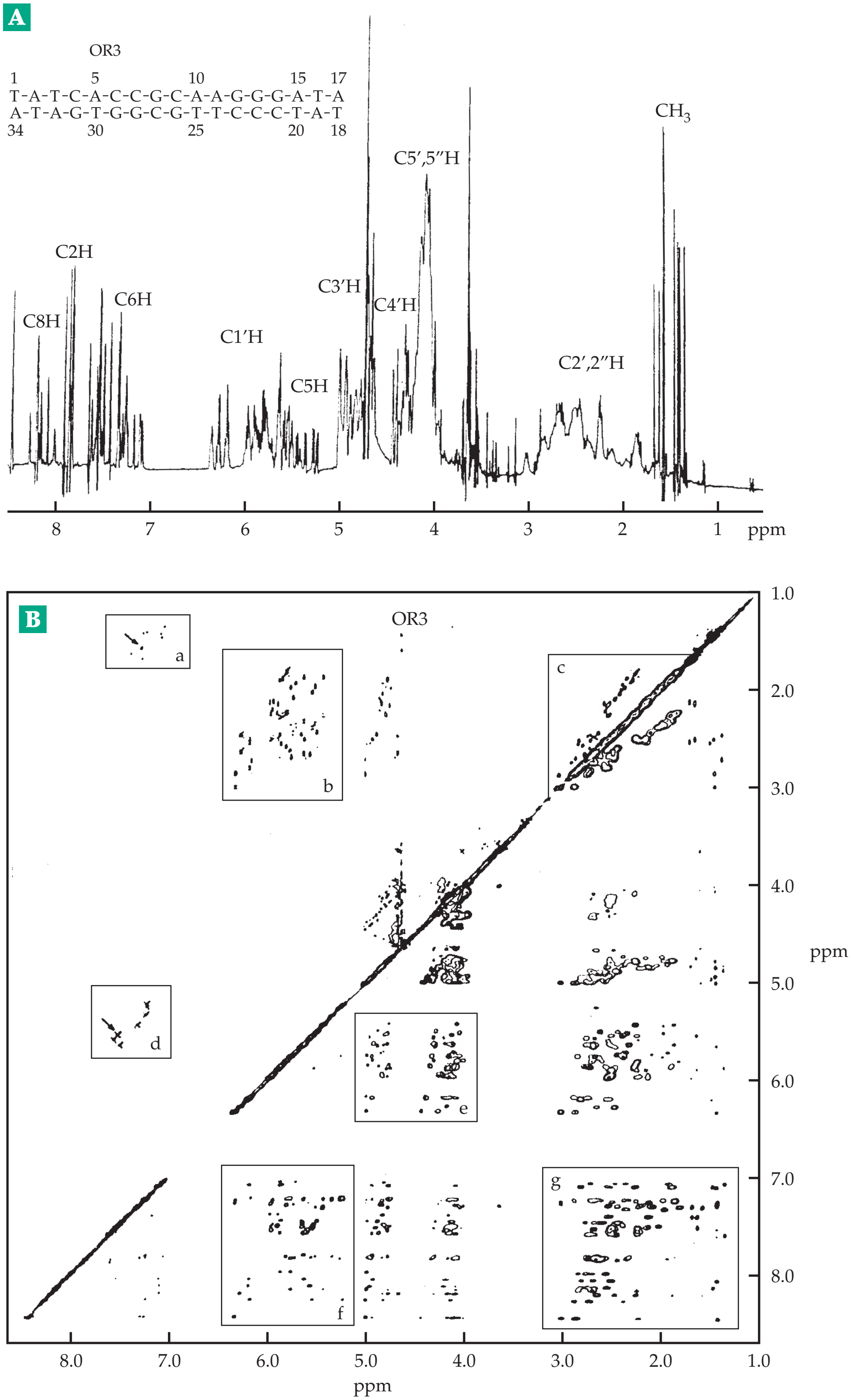


Figure 5-53 (A) ^{1}H NMR spectrum of a 17 base-pair DNA segment from the operator sequence OR3 from bacteriophage λ in D_2O at 37°C. (B) Combined COSY above the diagonal and NOESY (below the diagonal) spectra. C5H and C6H J coupling is established from cross-peaks in box d for cytosines and in box a for thymines. Two unresolved cross-peaks give rise to the more intense spots marked by arrows. Box b contains cross-peaks from scalar coupling of the two H2' protons to the H1' protons of the deoxyribose rings. Most of the aromatic proton resonances could be assigned using the NOE cross-peaks in box f. For further details see Wemmer *et al.*[676] See also Bax and Lerner.[672] Courtesy of B. Reid.

polynucleotides. One way to do this is to grow yeast on a medium containing methionine with ^{13}C in its methyl group. The most intense peaks in ^{13}C NMR spectra of tRNA molecules from this yeast will represent methyl groups in modified bases in the tRNA. Incorporated ^{13}C also permits study of internal motion within tRNA or other oligonucleotides by NMR methods.[677] However, as with protein NMR spectroscopy the major recent advances have come from systematic incorporation of both ^{13}C and ^{15}N into nucleic acids and the development of three- and four-dimensional NMR methods.[678–684b] Also important are methods for replacing some hydrogen atoms with deuterium to simplify spectra.[685,686]

The very sensitive ^{19}F nucleus can be introduced into tRNAs by incorporation of 5-fluorouracil in place of uracil[687] (Fig. 5-54A,B). Phosphorus 31 NMR spectra (Fig. 5-54C) can provide information about conformations of the chain.[688]

A few recent NMR investigations of polynucleotides include studies of triple-helical DNA,[689] Holliday junctions,[290] double-stranded oligonucleotides containing adducts of carcinogens,[690,691] of hairpin loops with sheared A•A and G•G pairs,[692] and of proton exchange in both imino and amino groups.[693]

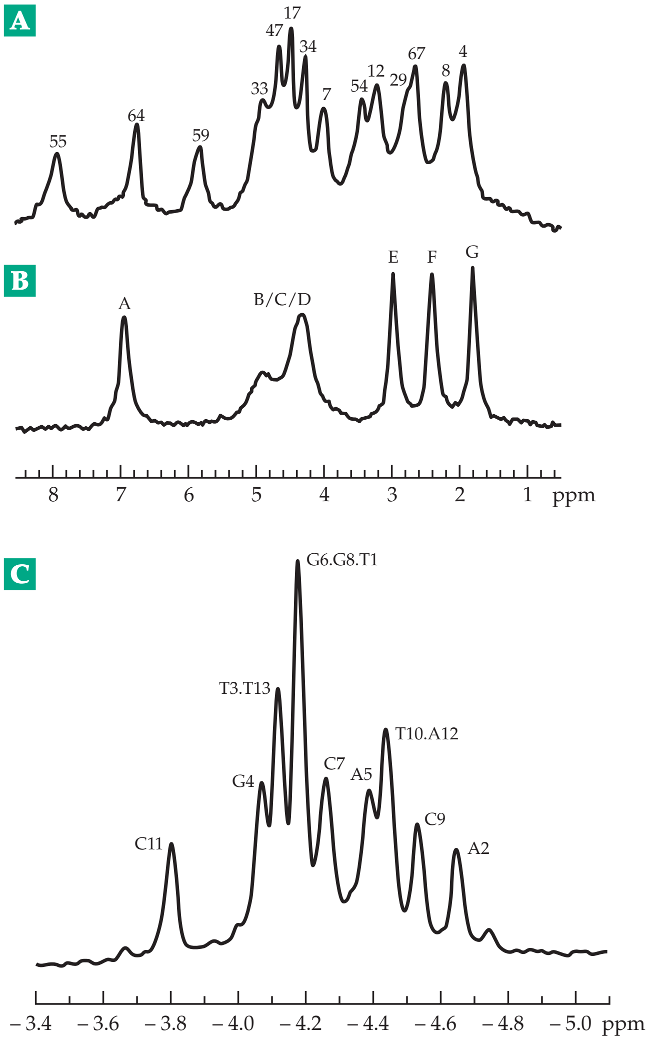


Figure 5-54 (A) An ^{19}F NMR spectrum of the 76-residue *E. coli* $tRNA^{Val}$ containing 5-fluorouracil in 14 positions. Recorded at 47°C. The numbers above the resonances indicate the position in the sequence. (The sequence is not identical to that for the yeast tRNA shown in Fig. 5-30.) Modified from Chu *et al.*[694] Courtesy of Jack Horowitz. (B) A similar spectrum for a 35-residue "minihelix" that contains the acceptor stem of the $tRNA^{Val}$ and seven fluorouracils. The broad peaks B, D, and E are shifted far upfield by reaction with bisulfite (Eq. 5-11) suggesting that they are not hydrogen bonded and are present in the loop of the stem–loop structure. Peaks A, E, F, and G correspond to resonances 64, 7, 67, and 4, respectively, in (A) and represent fluorouracil in the stem structure. From Chu *et al.*[694] Courtesy of Jack Horowitz. (C) A ^{31}P NMR spectrum of a synthetic 14 base-pair DNA segment related to the *E. coli lac* operator. The palindromic sequence is TCTGAGCGCTCAGA. The numbers refer to the positions from the 5' end. From Schroeder *et al.*[688]

References

1. Portugal, F. H., and Cohen, J. S. (1977) *A Century of DNA: A History of the Discovery of the Structure and Function of the Genetic Substance*, MIT Press, Cambridge, Massachusetts
2. Levene, P. A., and Tipson, R. S. (1935) *J. Biol. Chem.* **109**, 623–630
3. Schlenk, F. (1988) *Trends Biochem. Sci.* **13**, 67–69
4. Watson, J. D., and Crick, F. H. C. (1953) *Nature (London)* **171**, 737–738
5. Watson, J. D. (1968) *The Double Helix*, Atheneum, New York
6. Crick, F. (1988) *What a Mad Pursuit: A Personal View of Scientific Discovery*, Basic Books, New York
7. Olby, R. (1974) *The Path to the Double Helix*, Macmillan, London
8. Wilson, H. R. (1988) *Trends Biochem. Sci.* **13**, 275–278
9. Chambers, D. A., Reid, K. B. M., and Cohen, R. L. (1994) *FASEB J.* **8**, 1219–1226
9a. Piper, A. (1998) *Trends Biochem. Sci.* **23**, 151–154
10. Arnott, S., and Hukins, D. W. L. (1973) *J. Mol. Biol.* **81**, 93–105
11. Proudfoot, N. J., Shander, M. H. M., Manley, J. L., Gefter, M. L., and Maniatis, T. (1980) *Science* **209**, 1329–1336
12. Tucker, P. W., Liu, C.-P., Mushinski, J. F., and Blattner, F. R. (1980) *Science* **209**, 1353–1360
13. Yamada, Y., Mudryj, M., Sullivan, M., and de Crombrugghe, B. (1983) *J. Biol. Chem.* **258**, 2758–2761
14. Tsuzuki, T., Obaru, K., Setoyama, C., and Shimada, K. (1987) *J. Mol. Biol.* **198**, 21–31
15. Ratner, L., and 18 other authors (1985) *Nature (London)* **313**, 277–284
16. Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol.* **166**, 477– 535
17. Peden, K. W. C. (1983) *Gene* **22**, 277–280

References

18. Anderson, S., Bankier, A. T., Barrell, B. G., deBruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981) *Nature (London)* **290**, 457 – 470

19. Fraser, C. M., and 28 other authors (1995) *Science* **270**, 397 – 403

20. Fleischmann, R. D., and 39 other authors (1995) *Science* **269**, 496 – 512

20a. Doolittle, R. F. (1998) *Nature (London)* **392**, 339 – 342

21. Nowak, R. (1995) *Science* **267**, 172 – 174

22. Dujon, B., and 107 other authors (1994) *Nature (London)* **369**, 371 – 378

23. Johnston, M., and 34 other authors (1994) *Science* **265**, 2077 – 2082

24. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., and Oliver, S. G. (1996) *Science* **274**, 546 – 557

24a. Macilwain, C. (2000) *Nature (London)* **405**, 983 – 984

25. Liébecq, C., ed. (1992) *Biochemical Nomenclature*, Portland Press, London and Chapel Hill, North Carolina (for the International Union of Biochemistry and Molecular Biology)

26. Shabarova, Z., and Bogdanov, A. (1994) *Advanced Organic Chemistry of Nucleic Acids*, VCH Publ., Weinheim

27. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer, New York

28. Hurst, D. T. (1980) *An Introduction to the Chemistry and Biochemistry of Pyrimidines, Purines and Pteridines*, Wiley, New York

29. Markowski, V., Sullivan, G. R., and Roberts, J. D. (1977) *J. Am. Chem. Soc.* **99**, 714 – 718

30. Katritzky, A. R., and Karelson, M. (1991) *J. Am. Chem. Soc.* **113**, 1561 – 1566

31. Szczesniak, M., Leszczynski, J., and Person, W. B. (1992) *J. Am. Chem. Soc.* **114**, 2731 – 2733

32. Scanlan, M. J., and Hillier, I. H. (1984) *J. Am. Chem. Soc.* **106**, 3737 – 3745

33. Brown, D. M. (1974) in *Basic Principles in Nucleic Acid Chemistry*, Vol. II (Tso, P. O. P., ed), pp. 2 – 90, Academic Press, New York

34. Hoogsteen, K. (1963) *Acta Crystallogr.* **16**, 907 – 916

35. Frank-Kamenetskii, M. D., and Mirkin, S. M. (1995) *Ann. Rev. Biochem.* **64**, 65 – 95

36. Portugal, J. (1989) *Trends Biochem. Sci.* **14**, 127 – 130

37. Adams, R. L. P., Knowler, J. T., and Leader, D. P. (1992) *The Biochemistry of the Nucleic Acids*, 11th ed., Chapman & Hall, London

37a. Saenger, W. (1973) *Angew. Chem. Int. Ed. Engl.* **12**, 591 – 601

38. Fang, G., and Cech, T. R. (1993) *Biochemistry* **32**, 11646 – 11657

39. Gellert, M., Lipsett, M. N., and Davies, D. R. (1962) *Proc. Natl. Acad. Sci. U.S.A.* **48**, 2013 – 2018

40. Kennell, D. E. (1971) *Prog. Nucleic Acid Res. Mol. Biol.* **11**, 259 – 301

40a. Stofer, E., Chipot, C., and Lavery, R. (1999) *J. Am. Chem. Soc.* **121**, 9503 – 9508

40b. Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J., Jr. (1999) *Biochemistry* **38**, 3468 – 3477

41. Pranata, J., Wierschke, S. G., and Jorgensen, W. L. (1991) *J. Am. Chem. Soc.* **113**, 2810 – 2819

42. Dang, L. X., and Kollman, P. A. (1990) *J. Am. Chem. Soc.* **112**, 503 – 507

43. Newcomb, L. F., and Gellman, S. H. (1994) *J. Am. Chem. Soc.* **116**, 4993 – 4994

44. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) *J. Mol. Biol.* **86**, 843 – 853

45. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9373 – 9377

46. Sarai, A., Mazur, J., Nussinov, R., and Jernigan, R. L. (1988) *Biochemistry* **27**, 8498 – 8502

47. Hunter, C. A. (1993) *J. Mol. Biol.* **230**, 1025 – 1054

48. Friedman, R. A., and Honig, B. (1995) *Biophys. J.* **69**, 1528 – 1535

49. DePrisco Albergo, D., Marky, L. A., Breslauer, K. J., and Turner, D. H. (1981) *Biochemistry* **20**, 1409 – 1413

50. Privalov, P. L., and Filimonov, V. V. (1978) *J. Mol. Biol.* **122**, 447 – 464

51. Privalov, P. L., and Filimonov, V. V. (1978) *J. Mol. Biol.* **122**, 465 – 470

52. Petersen, S. B., and Led, J. J. (1981) *J. Am. Chem. Soc.* **103**, 5308 – 5313

53. Petruska, J., and Goodman, M. F. (1995) *J. Biol. Chem.* **270**, 746 – 750

54. SantaLucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996) *Biochemistry* **35**, 3555 – 3562

55. Egli, M., Usman, N., and Rich, A. (1993) *Biochemistry* **32**, 3221 – 3237

56. Watson, J. D., and Crick, F. H. C. (1953) *Nature (London)* **171**, 964 – 967

57. Topal, M. D., and Fresco, J. R. (1976) *Nature (London)* **263**, 285 – 289

58. Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1973) *Nature New Biol.* **246**, 40 – 41

59. Fazakerley, G. V., Gdaniec, Z., and Sowers, L. C. (1993) *J. Mol. Biol.* **230**, 6 – 10

60. Florián, J., Hrouda, V., and Hobza, P. (1994) *J. Am. Chem. Soc.* **116**, 1457 – 1460

61. Douhal, A., Kim, S. K., and Zewail, A. H. (1995) *Nature (London)* **378**, 260 – 263

62. Gabb, H. A., and Harvey, S. C. (1993) *J. Am. Chem. Soc.* **115**, 4218 – 4227

63. Ellervik, U., and Magnusson, G. (1994) *J. Am. Chem. Soc.* **116**, 2340 – 2347

64. Thibaudeau, C., Plavec, J., and Chattopadhyaya, J. (1994) *J. Am. Chem. Soc.* **116**, 8033 – 8037

65. Santos, R. A., Tang, P., and Harbison, G. S. (1989) *Biochemistry* **28**, 9372 – 9378

66. Sundaralingam, M. (1969) *Biopolymers* **7**, 821 – 860

67. Vologodskii, A. (1992) *Topology and Physics of Circular DNA*, CRC Press, Boca Raton, Florida

68. Baikalov, I., Grzeskowiak, K., Yanagi, K., Quintana, J., and Dickerson, R. E. (1993) *J. Mol. Biol.* **231**, 768 – 784

69. Lipanov, A., Kopka, M. L., Kaczor-Grzeskowiak, M., Quintana, J., and Dickerson, R. E. (1993) *Biochemistry* **32**, 1373 – 1389

70. Schlick, T., Hingerty, B. E., Peskin, C. S., Overton, M. L., and Broyde, S. (1991) *Theoretical Biochemistry and Molecular Physics,* Vol. 1 (Beveridge, D. L. and Lavery, R., eds), pp. 39 – 58, Adenine Press, Guilderland, New York

71. Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G., and Rich, A. (1979) *Nature (London)* **282**, 680 – 686

72. Kornberg, A., and Baker, T. A. (1992) *DNA Replication*, 2nd ed., Freeman, New York

73. Kaufmann, W. K., and Paules, R. S. (1996) *FASEB J.* **10**, 238 – 247

74. Lindahl, T. (1993) *Nature (London)* **362**, 709 – 715

75. Rodley, G. A., and Bates, R. H. T. (1980) *Trends Biochem. Sci.* **5**, 231 – 233

76. Kennard, O., and Salisbury, S. A. (1993) *J. Biol. Chem.* **268**, 10701 – 10704

77. Drew, H. R., and Dickerson, R. E. (1981) *J. Mol. Biol.* **151**, 535 – 556

78. Dickerson, R. E. (1983) *Sci. Am.* **249**(Dec), 94 – 111

78a. Dickerson, R. E. (1997) in *International Tables for Macromolecular Crystallography* (Rossmann, M. G., and Arnold, E., eds), (Chapter 23, Part c)

79. Shui, X., McFail-Isom, L., Hu, G. G., and Williams, L. D. (1998) *Biochemistry* **37**, 8341 – 8355

80. Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1995) *J. Mol. Biol.* **247**, 34 – 48

81. Tereshko, V., Minasov, G., and Egli, M. (1999) *J. Am. Chem. Soc.* **121**, 470 – 471

82. Yuan, H., Quintana, J., and Dickerson, R. E. (1992) *Biochemistry* **31**, 8009 – 8021

83. Ban, C., Ramakrishnan, B., and Sundaralingam, M. (1994) *J. Mol. Biol.* **236**, 275 – 285

83a. Kielkopf, C. L., Ding, S., Kuhn, P., and Rees, D. C. (2000) *J. Mol. Biol.* **296**, 787 – 801

84. Johansson, E., Parkinson, G., and Neidle, S. (2000) *J. Mol. Biol.* **300**, 551 – 561

85. Shen, L. X., Cai, Z., and Tinoco, I., Jr. (1995) *FASEB J.* **9**, 1023 – 1033

86. Rosenberg, J. M., Seeman, N. C., Day, R. O., and Rich, A. (1976) *J. Mol. Biol.* **104**, 145 – 167

87. Marino, J. P., Gregorian, R. S., Jr., Csankovszki, G., and Crothers, D. M. (1995) *Science* **268**, 1448 – 1454

88. Iwamoto, S., and Hsu, M.-T. (1983) *Nature (London)* **305**, 70 – 72

89. Morel, G., ed. (1995) *Visualization of Nucleic Acids*, CRC Press, Boca Raton, Florida

90. Arscott, P. G., Lee, G., Bloomfield, V. A., and Evans, D. F. (1989) *Nature (London)* **339**, 484 – 486

91. Jares-Erijman, E. A., and Jovin, T. M. (1996) *J. Mol. Biol.* **257**, 597 – 617

92. Duan, Y., Wilkosz, P., Crowley, M., and Rosenberg, J. M. (1997) *J. Mol. Biol.* **272**, 553 – 572

93. Auffinger, P., and Westhof, E. (2000) *J. Mol. Biol.* **300**, 1113 – 1131

93a. Soler-López, M., Malinina, L., Liu, J., Huynh-Dinh, T., and Subirana, J. A. (1999) *J. Biol. Chem.* **274**, 23683 – 23686

94. Schneider, B., and Berman, H. M. (1995) *Biophys. J.* **69**, 2661 – 2669

95. Tereshko, V., Minasov, G., and Egli, M. (1999) *J. Am. Chem. Soc.* **121**, 3590 – 3595

96. Franklin, R. E., and Gosling, R. G. (1953) *Nature (London)* **172**, 156 – 157

97. Ramakrishnan, B., and Sundaralingam, M. (1995) *Biophys. J.* **69**, 553 – 558

98. Ramakrishnan, B., and Sundaralingam, M. (1993) *J. Mol. Biol.* **231**, 431 – 444

99. Eisenstein, M., and Shakked, Z. (1995) *J. Mol. Biol.* **248**, 662 – 678

99a. Ng, H.-L., Kopka, M. L., and Dickerson, R. E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2035 – 2039

100. Selsing, E., Wells, R. D., Alden, C. J., and Arnott, S. (1979) *J. Biol. Chem.* **254**, 5417 – 5422

101. Mahendrasingam, A., Forsyth, V. T., Hussain, R., Greenall, R. J., Pigram, W. J., and Fuller, W. (1986) *Science* **233**, 195 – 197

102. Hopkins, R. C. (1981) *Science* **211**, 289 – 291

103. Gupta, G., Bansal, M., and Sasisekharan, V. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6486 – 6490

104. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W., and Ratliff, R. L. (1980) *Nature (London)* **283**, 743 – 745

105. Herbert, A., and Rich, A. (1996) *J. Biol. Chem.* **271**, 11595 – 11598

106. Gessner, R. V., Frederick, C. A., Quigley, G. J., Rich, A., and Wang, A. H.-J. (1989) *J. Biol. Chem.* **264**, 7921 – 7935

107. Zhou, G., and Ho, P. S. (1990) *Biochemistry* **29**, 7229 – 7236

References

108. Parkinson, G. N., Arvanitis, G. M., Lessinger, L., Ginell, S. L., Jones, R., Gaffney, B., and Berman, H. M. (1995) *Biochemistry* **34**, 15487–15495

109. Sadasivan, C., and Gautham, N. (1995) *J. Mol. Biol.* **248**, 918–930

110. Bancroft, D., Williams, L. D., Rich, A., and Egli, M. (1994) *Biochemistry* **33**, 1073–1086

111. Rich, A., Nordheim, A., and Wang, A. H.-J. (1984) *Ann. Rev. Biochem.* **53**, 791–846

112. Moore, M. H., Van Meervelt, L., Salisbury, S. A., Lin, P. K. T., and Brown, D. M. (1995) *J. Mol. Biol.* **251**, 665–673

113. Wang, Y., Thomas, G. A., and Peticolas, W. L. (1987) *Biochemistry* **26**, 5178–5186

114. Saenger, W., Hunter, W. N., and Kennard, O. (1986) *Nature (London)* **324**, 385–388

115. Frank-Kamenetskii, M. (1986) *Nature (London)* **324**, 305

116. Basham, B., Schroth, G. P., and Ho, P. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6464–6468

117. Misra, V. K., and Honig, B. (1996) *Biochemistry* **35**, 1115–1124

118. McLean, M. J., Blaho, J. A., Kilpatrick, M. W., and Wells, R. D. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5884–5888

119. Kim, J.-m, and DasSarma, S. (1996) *J. Biol. Chem.* **271**, 19724–19731

120. Nordheim, A., Pardue, M. L., Lafer, E. M., Möller, A., Stollar, B. D., and Rich, A. (1981) *Nature (London)* **294**, 417–422

121. Jaworski, A., Hsieh, W.-T., Blaho, J. A., Larson, J. E., and Wells, R. D. (1987) *Science* **238**, 773–777

122. Zacharias, W., O'Conner, T. R., and Larson, J. E. (1988) *Biochemistry* **27**, 2970–2978

123. Nordheim, A., and Rich, A. (1983) *Nature (London)* **303**, 674–679

124. Wells, R. D. (1988) *J. Biol. Chem.* **263**, 1095–1098

125. Rentzeperis, D., Kupke, D. W., and Marky, L. A. (1994) *Biochemistry* **33**, 9588–9591

126. Liu, K., Miles, H. T., Frazier, J., and Sasisekharan, V. (1993) *Biochemistry* **32**, 11802–11809

127. Zhou, N., Germann, M. W., van de Sande, J. H., Pattabiraman, N., and Vogel, H. J. (1993) *Biochemistry* **32**, 646–656

128. Otto, C., Thomas, G. A., Rippe, K., Jovin, T. M., and Peticolas, W. L. (1991) *Biochemistry* **30**, 3062–3069

129. Trifonov, E. N. (1991) *Trends Biochem. Sci.* **16**, 467–470

130. Rich, A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13999–14000

131. Hageman, P. J. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 265–286

132. Poncin, M., Hartmann, B., and Lavery, R. (1992) *J. Mol. Biol.* **226**, 775–794

133. MacKerell, A. D., Jr., Wiórkiewicz-Kuczera, J., and Karplus, M. (1995) *J. Am. Chem. Soc.* **117**, 11946–11975

134. Kosikov, K. M., Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1999) *J. Mol. Biol.* **289**, 1301–1326

135. Peticolas, W. L., Wang, Y., and Thomas, G. A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2579–2583

136. Dickerson, R. E., and Drew, H. R. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7318–7322

136a. Packer, M. J., and Hunter, C. A. (1998) *J. Mol. Biol.* **280**, 407–420

137. Hogan, M., Dattagupta, N., and Crothers, D. M. (1979) *Nature (London)* **278**, 521–524

138. Mitra, C. K., Sarma, M. H., and Sarma, R. H. (1981) *J. Am. Chem. Soc.* **103**, 6727–6737

139. Cluzel, P., Lebrun, A., Heller, C., Lavery, R., Viovy, J.-L., Chatenay, D., and Caron, F. (1996) *Science* **271**, 792–794

140. Smith, S. B., Cui, Y., and Bustamante, C. (1996) *Science* **271**, 795–799

141. Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A., and Croquette, V. (1996) *Science* **271**, 1835–1837

141a. Allemand, J. F., Bensimon, D., Lavery, R., and Croquette, V. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14152–14157

142. Dickerson, R. E. (1983) *J. Mol. Biol.* **166**, 419–441

143. Tung, C.-S., and Harvey, S. C. (1986) *J. Biol. Chem.* **261**, 3700–3709

144. Ulanovsky, L. E., and Trifonov, E. N. (1987) *Nature (London)* **326**, 720–722

145. Diekmann, S. (1989) *EMBO J.* **8**, 1–4

146. El Hassan, M. A., and Calladine, C. R. (1996) *J. Mol. Biol.* **259**, 95–103

147. Calladine, C. R. (1982) *J. Mol. Biol.* **161**, 343–352

148. Hunter, C. A., and Lu, X.-J. (1997) *J. Mol. Biol.* **265**, 603–619

149. Babcock, M. S., Pednault, E. P. D., and Olson, W. K. (1994) *J. Mol. Biol.* **273**, 125–156

150. Lu, X.-J., El Hassan, M. A., and Hunter, C. A. (1997) *J. Mol. Biol.* **273**, 668–680

151. Kitchin, P. A., Klein, V. A., Ryan, K. A., Gann, K. L., Rauch, C. A., Kang, D. S., Wells, R. D., and England, P. T. (1986) *J. Biol. Chem.* **261**, 11302–11309

152. Koo, H.-S., Wu, H.-M., and Crothers, D. M. (1986) *Nature (London)* **320**, 501–506

153. Zahn, K., and Blattner, F. B. (1987) *Science* **236**, 416–422

154. Snyder, M., Buchman, A. R., and Davis, R. W. (1986) *Nature (London)* **324**, 87–89

155. Travers, A. A. (1987) *Trends Biochem. Sci.* **12**, 108–112

156. Grzeskowiak, K., Goodsell, D. S., Kaczor-Grzeskowiak, M., Cascio, D., and Dickerson, R. E. (1993) *Biochemistry* **32**, 8923–8931

157. Allewell, N. (1988) *Trends Biochem. Sci.* **13**, 193–195

158. Haran, T. E., Kahn, J. D., and Crothers, D. M. (1994) *J. Mol. Biol.* **244**, 135–143

159. Strauss, J. K., and Maher, L. J., III. (1994) *Science* **266**, 1829–1834

160. Olson, W. K., Marky, N. L., Jernigan, R. L., and Zhurkin, V. B. (1993) *J. Mol. Biol.* **232**, 530–554

161. Price, M. A., and Tullius, T. D. (1993) *Biochemistry* **32**, 127–136

162. DiGabriele, A. D., and Steitz, T. A. (1993) *J. Mol. Biol.* **231**, 1024–1039

163. Chuprina, V. P., Fedoroff, O. Y., and Reid, B. R. (1991) *Biochemistry* **30**, 561–568

164. Crothers, D. M. (1994) *Science* **266**, 1819–1820

165. Husain, I., Griffith, J., and Sancar, A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2558–2562

166. Srinivasan, A. R., Torres, R., Clark, W., and Olson, W. K. (1987) *J. Biomol. Struct. Dyn.* **5**, 459–496

167. Lutter, L. C., Halvorson, H. R., and Calladine, C. R. (1996) *J. Mol. Biol.* **261**, 620–633

168. Privé, G. G., Heinemann, V., Chandrasegaran, S., Kan, L.-S., Kopka, M. L., and Dickerson, R. E. (1987) *Science* **238**, 498–504

169. Lilley, D. M. J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7140–7142

170. Woodson, S. A., and Crothers, D. M. (1988) *Biochemistry* **27**, 3130–3141

171. Miller, M., Harrison, R. W., Wlodawer, A., Appella, E., and Sussman, J. L. (1988) *Nature (London)* **334**, 85–86

172. Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D., and Kennard, O. (1987) *J. Biol. Chem.* **262**, 9962–9970

173. Lane, A., Martin, S. R., Ebel, S., and Brown, T. (1992) *Biochemistry* **31**, 12087–12095

174. Li, Y., and Agrawal, S. (1995) *Biochemistry* **34**, 10056–10062

175. Manning, G. S. (1978) *Q. Rev. Biophys.* **11**, 179–246

176. Mascotti, D. P., and Lohman, T. M. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3142–3146

177. Li, A. Z., Huang, H., Re, X., Qi, L. J., and Marx, K. A. (1998) *Biophys. J.* **74**, 964–973

178. Frydman, L., Rossomando, P. C., Frydman, V., Fernandez, C. O., Frydman, B., and Samejima, K. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9186–9190

179. Jain, S., Zon, G., and Sundaralingam, M. (1989) *Biochemistry* **28**, 2360–2364

180. Pelta, J., Durand, D., Doucet, J., and Livolant, F. (1996) *Biophys. J.* **71**, 48–63

181. Rajski, S. R., Kumar, S., Roberts, R. J., and Barton, J. K. (1999) *J. Am. Chem. Soc.* **121**, 5615–5616

181a. Aich, P., Labiuk, S. L., Tari, L. W., Dalbaere, L. J. T., Roesler, W. J., Falk, K. J., Steer, R. P., and Lee, J. S. (1999) *J. Mol. Biol.* **294**, 477–485

182. Fink, H.-W., and Schönenberger, C. (1999) *Nature (London)* **398**, 407–410

183. Marini, J. C., Miller, K. G., and Englund, P. T. (1980) *J. Biol. Chem.* **255**, 4976–4979

184. Wasserman, S. A., Dungan, J. M., and Cozzarelli, N. R. (1985) *Science* **229**, 171–174

185. Du, S. M., Wang, H., Tse-Dinh, Y.-C., and Seeman, N. C. (1995) *Biochemistry* **34**, 673–682

186. Du, S. M., Stollar, B. D., and Seeman, N. C. (1995) *J. Am. Chem. Soc.* **117**, 1194–1200

187. Wasserman, S. A., and Cozzarelli, N. R. (1986) *Science* **232**, 951–960

188. White, J. H., and Cozzarelli, N. R. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3322–3326

189. Cipra, B. (1992) *Science* **255**, 403

190. Adams, C. C. (1994) *The Knot Book*, Freeman, New York

191. Krasnow, M. A., Stasiak, A., Spengler, S. J., Dean, F., Koller, T., and Cozzarelli, N. R. (1983) *Nature (London)* **304**, 559–560

192. Du, S. M., and Seeman, N. C. (1992) *J. Am. Chem. Soc.* **114**, 9652–9655

193. Wang, J. C. (1980) *Trends Biochem. Sci.* **5**, 219–221

194. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry, Part 3*, Freeman, San Francisco, California (pp. 1265–1290)

195. Lebowitz, J. (1990) *Trends Biochem. Sci.* **15**, 202–207

196. Bates, A. D., and Maxwell, A. (1993) *DNA Topology*, Oxford Univ. Press, London

197. Yang, Y., Westcott, T. P., Pedersen, S. C., Tobias, I., and Olson, W. K. (1995) *Trends Biochem. Sci.* **20**, 313–319

198. Vinograd, J., Lebowitz, J., and Watson, R. (1968) *J. Mol. Biol.* **33**, 173–197

199. White, J. H., Cozzarelli, N. R., and Bauer, W. R. (1988) *Science* **241**, 323–327

199a. Vologodskii, A. V., and Cozzarelli, N. R. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 609–643

199b. Cozzarelli, N. R., Boles, T. C., and White, J. H. (1990) in *DNA Topology and its Biological Effects*, pp. 139–184, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York

199c. Podtelezhnikov, A. A., Cozzarelli, N. R., and Vologodskii, A. V. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12974–12979

200. Reich, Z., Levin-Zaidman, S., Gutman, S. B., Arad, T., and Minsky, A. (1994) *Biochemistry* **33**, 14177–14184

201. Eriksson, M. A. L., Härd, T., and Nilsson, L. (1995) *Biophys. J.* **68**, 402–426

202. Keller, W. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2550–2554

203. Kornberg, R. D., and Klug, A. (1981) *Sci. Am.* **244**(Feb), 52–64

204. Wang, B.-C., Rose, J., Arents, G., and Moudrianakis, E. N. (1994) *J. Mol. Biol.* **236**, 179–188

205. Malik, M., Bensaid, A., Rouviere-Yaniv, J., and Drlica, K. (1996) *J. Mol. Biol.* **256**, 66–76
206. Serrano, M., Salas, M., and Hermoso, J. M. (1993) *Trends Biochem. Sci.* **18**, 202–206
207. Seidl, A., and Hinz, H.-J. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1312–1316
208. Rahmouni, A. R., and Wells, R. D. (1989) *Science* **246**, 358–363
209. Spolar, R. S., and Record, M. T., Jr. (1994) *Science* **263**, 777–784
210. White, J. H., Gallo, R. M., and Bauer, W. R. (1992) *Trends Biochem. Sci.* **17**, 7–12
211. van Holde, K., and Zlatanova, J. (1995) *J. Biol. Chem.* **270**, 8373–8376
212. Keller, W. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4876–4880
213. Hamiche, A., Schultz, P., Ramakrishnan, V., Oudet, P., and Prunell, A. (1996) *J. Mol. Biol.* **257**, 30–42
214. Alfonso, P. J., Crippa, M. P., Hayes, J. J., and Bustin, M. (1994) *J. Mol. Biol.* **236**, 189–198
215. Bednar, J., Furrer, P., Stasiak, A., Dubochet, J., Egelman, E. H., and Bates, A. D. (1994) *J. Mol. Biol.* **235**, 825–847
216. Kumar, A. (1995) *Biochemistry* **34**, 12921–12925
217. Stigter, D. (1995) *Biophys. J.* **69**, 380–388
218. Baeza, I., Gariglio, P., Rangel, L. M., Chavez, P., Cervantes, L., Arguello, C., Wong, C., and Montañez, C. (1987) *Biochemistry* **26**, 6387–6392
219. Schmid, M. B. (1988) *Trends Biochem. Sci.* **13**, 131–135
220. Quigley, G. J., Wang, A. H.-J., Ughetto, G., van der Marel, G., van Boom, J. H., and Rich, A. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7204–7208
221. Campbell, A. M., and Jolly, D. J. (1973) *Biochem. J.* **133**, 209–226
222. Depew, R. E., and Wang, J. C. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4275–4279
223. Westof, E., and Sundaralingam, M. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1852–1856
224. Sobell, H. M. (1974) *Sci. Am.* **231**(Aug), 82–91
225. Chen, H., and Patel, D. J. (1995) *J. Am. Chem. Soc.* **117**, 5901–5913
226. Chaires, J. B., Satyanarayana, S., Suh, D., Fokt, I., Przewloka, T., and Priebe, W. (1996) *Biochemistry* **35**, 2047–2053
227. Rajeswari, M. R., Bose, H. S., Kukreti, S., Gupta, A., Chauhan, V. S., and Roy, K. B. (1992) *Biochemistry* **31**, 6237–6241

227a. Leitner, D., Schröder, W., and Weisz, K. (2000) *Biochemistry* **39**, 5886–5892

228. Felsenfeld, G., Davies, D., and Rich, A. (1957) *J. Am. Chem. Soc.* **79**, 2023–2024
229. Felsenfeld, G., and Miles, H. T. (1967) *Ann. Rev. Biochem.* **36**, 407–448
230. Arnott, S., and Selsing, E. (1974) *J. Mol. Biol.* **88**, 509–521
231. Wang, E., Koshlap, K. M., Gillespie, P., Dervan, P. B., and Feigon, J. (1996) *J. Mol. Biol.* **257**, 1052–1069
232. Radhakrishnan, I., and Patel, D. J. (1994) *Biochemistry* **33**, 11405–11416
233. Radhakrishnan, I., and Patel, D. J. (1994) *J. Mol. Biol.* **241**, 600–619
234. Rhee, S., Han, Z.-j, Liu, K., Miles, H. T., and Davies, D. R. (1999) *Biochemistry* **38**, 16810–16815
235. Roberts, R. W., and Crothers, D. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4320–4325
236. van Dongen, M. J. P., Heus, H. A., Wymenga, S. S., van der Marel, G. A., van Boom, J. H., and Hilbers, C. W. (1996) *Biochemistry* **35**, 1733–1739
237. Klysik, J. (1995) *J. Mol. Biol.* **245**, 499–507
238. Srinivasan, A. R., and Olson, W. K. (1998) *J. Am. Chem. Soc.* **120**, 484–491
239. Hanvey, J. C., Shimizu, M., and Wells, R. D. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6292–6296
240. Htun, H., and Dahlberg, J. E. (1989) *Science* **243**, 1571–1576
241. Mirkin, S. M., and Frank-Kamenetskii, M. D. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 541–576
242. Singh, S., Patel, P. K., and Hosur, R. V. (1997) *Biochemistry* **36**, 13214–13222
243. Van Meervelt, L., Vlieghe, D., Dautant, A., Gallois, B., Précigoux, G., and Kennard, O. (1995) *Nature (London)* **374**, 742–744
244. Reddy, G., Jwang, B., Rao, B. J., and Radding, C. M. (1994) *Biochemistry* **33**, 11486–11492
245. Kim, M. G., Zhurkin, V. B., Jernigan, R. L., and Camerini-Otero, R. D. (1995) *J. Mol. Biol.* **247**, 874–889
246. Cox, M. M. (1995) *J. Biol. Chem.* **270**, 26021–26024
247. Jain, S. K., Cox, M. M., and Inman, R. B. (1995) *J. Biol. Chem.* **270**, 4943–4949
248. Best, G. C., and Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117**, 1187–1193
249. Greenberg, W. A., and Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117**, 5016–5022
250. Hashem, G. M., Pham, L. P., Vaughan, M. R., and Gray, D. M. (1998) *Biochemistry* **37**, 61–72
251. Gryaznov, S. M., Lloyd, D. H., Chen, J.-K., Schultz, R. G., DeDionisio, L. A., Ratmeyer, L., and Wilson, W. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5798–5802
252. Betts, L., Josey, J. A., Veal, J. M., and Jordan, S. R. (1995) *Science* **270**, 1838–1841
253. Tomac, S., Sarkar, M., Ratilainen, T., Wittung, P., Nielsen, P. E., Nordén, B., and Gräslund, A. (1996) *J. Am. Chem. Soc.* **118**, 5544–5552

253a. Ray, A., and Nordén, B. (2000) *FASEB J.* **14**, 1041–1060

253b. Izvolsky, K. I., Demidov, V. V., Nielsen, P. E., and Frank-Kamenetskii, M. D. (2000) *Biochemistry* **39**, 10908–10913

254. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. (1998) *Nature (London)* **391**, 468–471
255. Marsh, T. C., and Henderson, E. (1994) *Biochemistry* **33**, 10718–10724
256. Blackburn, E. H. (1991) *Nature (London)* **350**, 569–573
257. Blackburn, E. H. (1991) *Trends Biochem. Sci.* **16**, 378–381
258. Moyzis, R. K. (1991) *Sci. Am.* **265**(Aug), 48–55
259. Henderson, E., Hardin, C. C., Wolk, S. K., Tinoco, I., Jr., and Blackburn, E. (1987) *Cell* **51**, 899–908
260. Laughlan, G., Murchie, A. I. H., Norman, D. G., Moore, M. H., Moody, P. C. E., Lilley, D. M. J., and Luisi, B. (1994) *Science* **265**, 520–527
261. Kang, C. H., Zhang, X., Ratliff, R., Moyzis, R., and Rich, A. (1992) *Nature (London)* **356**, 126–131
262. Smith, F. W., and Feigon, J. (1993) *Biochemistry* **32**, 8682–8692
263. Balagurumoorthy, P., and Brahmachari, S. K. (1994) *J. Biol. Chem.* **269**, 21858–21869
264. Williamson, J. R. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 703–730
265. Frank-Kamenetskii, M. (1992) *Nature (London)* **356**, 105
266. Wang, K. Y., Swaminathan, S., and Bolton, P. H. (1994) *Biochemistry* **33**, 7517–7527
267. Wang, Y., and Patel, D. J. (1995) *J. Mol. Biol.* **251**, 76–94
268. Laporte, L., and Thomas, G. J., Jr. (1998) *Biochemistry* **37**, 1327–1335
269. Ross, W. S., and Hardin, C. C. (1994) *J. Am. Chem. Soc.* **116**, 6070–6080
270. Sen, D., and Gilbert, W. (1990) *Nature (London)* **344**, 410–414
271. Miura, T., Benevides, J. M., and Thomas, G. J., Jr. (1995) *J. Mol. Biol.* **248**, 233–238
272. Schierer, T., and Henderson, E. (1994) *Biochemistry* **33**, 2240–2246
273. Hardin, C. C., Corregan, M., Brown, B. A., II, and Frederick, L. N. (1993) *Biochemistry* **32**, 5870–5880
274. Cheong, C., and Moore, P. B. (1992) *Biochemistry* **31**, 8406–8414
275. Chen, F.-M. (1995) *J. Biol. Chem.* **270**, 23090–23096
276. Kettani, A., Kumar, R. A., and Patel, D. J. (1995) *J. Mol. Biol.* **254**, 638–656
277. Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., and Rich, A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11636–11640
278. Gehring, K., Leroy, J.-L., and Guéron, M. (1993) *Nature (London)* **363**, 561–565
279. Lacroix, L., Mergny, J.-L., Leroy, J.-L., and Hélène, C. (1996) *Biochemistry* **35**, 8715–8722
280. Chen, L., Cai, L., Zhang, X., and Rich, A. (1994) *Biochemistry* **33**, 13540–13546
281. Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., and Rich, A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3874–3878
282. Benevides, J. M., Kang, C., and Thomas, G. J., Jr. (1996) *Biochemistry* **35**, 5747–5755
283. Gallego, J., Chou, S.-H., and Reid, B. R. (1997) *J. Mol. Biol.* **273**, 840–856
284. Selsing, E., Wells, R. D., Alden, C. J., and Arnott, S. (1979) *J. Biol. Chem.* **254**, 5417–5422
285. Nishizaki, T., Iwai, S., Ohkubo, T., Kojima, C., Nakamura, H., Kyogoku, Y., and Ohtsuka, E. (1996) *Biochemistry* **35**, 4016–4025
286. Salazar, M., Champoux, J. J., and Reid, B. R. (1993) *Biochemistry* **32**, 739–744
287. Seeman, N. C., and Kallenbach, N. R. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 53–86
288. Zhang, S., Fu, T.-J., and Seeman, N. C. (1993) *Biochemistry* **32**, 8062–8067
289. Pikkemaat, J. A., van den Elst, H., van Boom, J. H., and Altona, C. (1994) *Biochemistry* **33**, 14896–14907
290. Carlström, G., and Chazin, W. J. (1996) *Biochemistry* **35**, 3534–3544
291. Holliday, R. (1964) *Genet. Res. Camb.* **5**, 282–304
292. Bennett, R. J., and West, S. C. (1995) *J. Mol. Biol.* **252**, 213–226
293. Grainger, R. J., Murchie, A. I. H., and Lilley, D. M. J. (1998) *Biochemistry* **37**, 23–32
294. Lilley, D. M. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9513–9515
295. Panyutin, I. G., Biswas, I., and Hsieh, P. (1995) *EMBO J.* **14**, 1819–1826
296. Chen, J., and Seeman, N. C. (1991) *Nature (London)* **350**, 631–633
297. Zhang, Y., and Seeman, N. C. (1994) *J. Am. Chem. Soc.* **116**, 1661–1669

297a. Liu, F., Sha, R., and Seeman, N. C. (1999) *J. Am. Chem. Soc.* **121**, 917–922

297b. Fahlman, R. P., and Sen, D. (1999) *J. Am. Chem. Soc.* **121**, 11079–11085

298. Feagin, J. E. (1990) *J. Biol. Chem.* **265**, 19373–19376
299. Stuart, K. (1991) *Trends Biochem. Sci.* **16**, 68–72
300. Simons, R. W., and Grunberg-Manago, M., eds. (1997) *RNA Structure and Function*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
301. Nicoloso, M., Qu, L.-H., Michot, B., and Bachellerie, J.-P. (1996) *J. Mol. Biol.* **260**, 178–195
302. Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995) *Ann. Rev. Biochem.* **64**, 763–797
303. Draper, D. E. (1996) *Trends Biochem. Sci.* **21**, 145–149
304. Sich, C., Ohlenschläger, O., Ramachandran, R., Görlach, M., and Brown, L. R. (1997) *Biochemistry* **36**, 13989–14002

References

305. Egli, M., Portmann, S., and Usman, N. (1996) *Biochemistry* **35**, 8489–8494
306. Auffinger, P., and Westhof, E. (1997) *J. Mol. Biol.* **274**, 54–63
307. Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, N. S., and Reid, B. R. (1993) *Biochemistry* **32**, 4207–4215
308. Fedoroff, O. Y., Salazar, M., and Reid, B. R. (1993) *J. Mol. Biol.* **233**, 509–523
309. Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G., and Wilson, W. D. (1994) *Biochemistry* **33**, 5298–5304
310. Varani, G., Cheong, C., and Tinoco, I., Jr. (1991) *Biochemistry* **30**, 3280–3289
311. Butcher, S. E., Dieckmann, T., and Feigon, J. (1997) *J. Mol. Biol.* **268**, 348–358
312. van Dongen, M. J. P., Wijmenga, S. S., van der Marel, G. A., van Boom, J. H., and Hilbers, C. W. (1996) *J. Mol. Biol.* **263**, 715–729
313. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) *Science* **256**, 217–219
314. Murphy, F. L., and Cech, T. R. (1994) *J. Mol. Biol.* **236**, 49–63
315. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. M., and Pardi, A. (1996) *J. Mol. Biol.* **264**, 968–980

315a. Gutell, R. R., Cannone, J. J., Konings, D., and Gautheret, D. (2000) *J. Mol. Biol.* **300**, 791–803

316. Fountain, M. A., Serra, M. J., Krugh, T. R., and Turner, D. H. (1996) *Biochemistry* **35**, 6539–6548
317. Huang, S., Wang, Y.-X., and Draper, D. E. (1996) *J. Mol. Biol.* **258**, 308–321
318. Nowakowski, J., and Tinoco, I., Jr. (1996) *Biochemistry* **35**, 2577–2585
319. Chang, K.-Y., and Tinoco, I., Jr. (1997) *J. Mol. Biol.* **269**, 52–66
320. Rosendahl, G., Hansen, L. H., and Douthwaite, S. (1995) *J. Mol. Biol.* **249**, 59–68
321. Pleij, C. W. A. (1990) *Trends Biochem. Sci.* **15**, 143–147
322. ten Dam, E., Pleij, K., and Draper, D. (1992) *Biochemistry* **31**, 11665–11676
323. Shen, L. X., and Tinoco, I., Jr. (1995) *J. Mol. Biol.* **247**, 963–978
324. Pinard, R., Payant, C., and Brakier-Gingras, L. (1995) *Biochemistry* **34**, 9611–9616
325. Puglisi, J. D., Wyatt, J. R., and Tinoco, I., Jr. (1988) *Nature (London)* **331**, 283–286
326. Qiu, H., Kaluarachchi, K., Du, Z., Hoffman, D. W., and Giedroc, D. P. (1996) *Biochemistry* **35**, 4176–4186

326a. Kolk, M. H., van der Graaf, M., Wijmenga, S. S., Pleij, C. W. A., Heus, H. A., and Hilbers, C. W. (1998) *Science* **280**, 434–438

327. Shi, P.-Y., Brinton, M. A., Veal, J. M., Zhong, Y. Y., and Wilson, W. D. (1996) *Biochemistry* **35**, 4222–4230
328. Hubbard, J. M., and Hearst, J. E. (1991) *Biochemistry* **30**, 5458–5465
329. Soll, D., and Abelson, J., eds. (1979) *Transfer RNA*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
330. Zimmermann, R. A. (1996) *Science* **271**, 1240–1241
331. Steinberg, S., Gautheret, D., and Cedergren, R. (1994) *J. Mol. Biol.* **236**, 982–989
332. Holbrook, S. R., Sussman, J. L., Warrant, R. W., and Kim, S.-H. (1978) *J. Mol. Biol.* **123**, 631–660
333. Rich, A., and Kim, S. H. (1978) *Sci. Am.* **238** (Jan), 52–62
334. Rich, A. (1978) *Trends Biochem. Sci.* **3**, 34–37
335. Moras, D., Comarmond, M. B., Fischer, J., Weiss, R., Thierry, J. C., Ebel, J. P., and Giegé, R. (1980) *Nature (London)* **288**, 669–674
336. Woo, N. H., Roe, B. A., and Rich, A. (1980) *Nature (London)* **286**, 346–351
337. Wakao, H., Romby, P., Westhof, E., Laalami, S., Grunberg-Manago, M., Ebel, J.-P., Ehresmann, C., and Ehresmann, B. (1989) *J. Biol. Chem.* **264**, 20363–20371
338. Hou, Y.-M., and Schimmel, P. (1992) *Biochemistry* **31**, 4157–4160
339. Leehey, M. A., Squassoni, C. A., Friederich, M. W., Mills, J. B., and Hagerman, P. J. (1995) *Biochemistry* **34**, 16235–16239
340. Quigley, G. J., and Rich, A. (1976) *Science* **194**, 796–806
341. Brimacombe, R., Atmadja, J., Stiege, W., and Schüler, D. (1988) *J. Mol. Biol.* **199**, 115–136
342. Mueller, F., and Brimacombe, R. (1997) *J. Mol. Biol.* **271**, 524–544
343. Turner, D. H., and Sugimoto, N. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 167–192
344. Brimacombe, R. (1988) *Biochemistry* **27**, 4207–4213
345. Schnare, M. N., Damberger, S. H., Gray, M. W., and Gutell, R. R. (1996) *J. Mol. Biol.* **256**, 701–719
346. Yonath, A. (1984) *Trends Biochem. Sci.* **9**, 227–230
347. Pennisi, E. (1999) *Science* **285**, 2048–2051
348. Svergun, D. I., Koch, M. H. J., Pedersen, J. S., and Serdyuk, I. N. (1994) *J. Mol. Biol.* **240**, 78–86
349. Vladimirov, S. N., Druzina, Z., Wang, R., and Cooperman, B. S. (2000) *Biochemistry* **39**, 183–193
350. Gregory, S. T., and Dahlberg, A. E. (1999) *J. Mol. Biol.* **285**, 1475–1483
351. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 905–920
352. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., and Noller, H. F. (1999) *Science* **285**, 2095–2104
353. Culver, G. M., Cate, J. H., Yusupova, G. Z., Yusupov, M. M., and Noller, H. F. (1999) *Science* **285**, 2133–2135
354. Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000) *Nature (London)* **407**, 340–348
355. Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 920–930
356. Muth, G. W., Ortoleva-Donnelly, L., and Strobel, S. A. (2000) *Science* **289**, 947–950
357. Cech, T. R. (2000) *Science* **289**, 878–879
358. Pennisi, E. (1996) *Science* **273**, 574–575
359. O'Gara, M., Klimasauskas, S., Roberts, R. J., and Cheng, X. (1996) *J. Mol. Biol.* **261**, 634–645
360. Gommers-Ampt, J. H., and Borst, P. (1995) *FASEB J.* **9**, 1034–1042
361. Maltman, K. L., Neuhard, J., and Warren, R. A. J. (1981) *Biochemistry* **20**, 3586–3591
362. Lipscomb, L. A., Peek, M. E., Morningstar, M. L., Verghis, S. M., Miller, E. M., Rich, A., Essigmann, J. M., and Williams, L. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 719–723
363. Leonard, G. A., McAuley-Hecht, K. E., Gibson, N. J., Brown, T., Watson, W. P., and Hunter, W. N. (1994) *Biochemistry* **33**, 4755–4761
364. Wrzesinski, J., Bakin, A., Nurse, K., Lane, B. G., and Ofengand, J. (1995) *Biochemistry* **34**, 8904–8913
365. Ellington, A. D., and Szostak, J. W. (1990) *Nature (London)* **346**, 818–822
366. Fan, P., Suri, A. K., Fiala, R., Live, D., and Patel, D. J. (1996) *J. Mol. Biol.* **258**, 480–500
367. Jenison, R. D., Gill, S. C., Pardi, A., and Polisky, B. (1994) *Science* **263**, 1425–1429
368. Hamasaki, K., Killian, J., Cho, J., and Rando, R. R. (1998) *Biochemistry* **37**, 656–663
369. Connell, G. J., Illangesekare, M., and Yarus, M. (1993) *Biochemistry* **32**, 5497–5502
370. Patel, D. J., Suri, A. K., Jiang, F., Jiang, L., Fan, P., Kumar, R. A., and Nonin, S. (1997) *J. Mol. Biol.* **272**, 645–664
371. Crick, F. H. C. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 3–9
372. Cornish-Bowden, A. (1996) *Trends Biochem. Sci.* **21**, 155
373. Hall, B. D. (1979) *Nature (London)* **282**, 129–130
374. Barrell, B. G., Bankier, A. T., and Drouin, J. (1979) *Nature (London)* **282**, 189–194
375. Horowitz, S., and Gorovsky, M. A. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2452–2455
376. Chambers, I., and Harrison, P. R. (1987) *Trends Biochem. Sci.* **12**, 255–256
377. Gardner, M. (1970) *Sci. Am.* **223**(Aug), 110–112
378. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., and Frank-Kamenetskii, M. D. (1987) *Nature (London)* **330**, 495–497
379. Guieysse, A.-L., Praseuth, D., and Hélène, C. (1997) *J. Mol. Biol.* **267**, 289–298
380. Panayotatos, N., and Fontaine, A. (1987) *J. Biol. Chem.* **262**, 11364–11368
381. Pabo, C. O., and Sauer, R. T. (1984) *Ann. Rev. Biochem.* **53**, 293–321
382. Berg, O. G., and von Hippel, P. H. (1988) *Trends Biochem. Sci.* **13**, 207–211
383. Marmorstein, R. Q., Joachimiak, A., Sprinzl, M., and Sigler, P. B. (1987) *J. Biol. Chem.* **262**, 4922–4927
384. Muto, A., and Osawa, S. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 166–169
385. Singer, C. E., and Ames, B. N. (1979) *Science* **170**, 822–826
386. Singer, C. E., and Ames, B. N. (1972) *Science* **175**, 1391
387. Harrison, S. C. (1991) *Nature (London)* **353**, 715–719
388. Churchill, M. E. A., and Travers, A. A. (1991) *Trends Biochem. Sci.* **16**, 92–97
389. Travers, A. (1993) *DNA-Protein Interactions*, Chapman & Hall, New York
390. Jones, S., van Heyningen, P., Berman, H. M., and Thornton, J. M. (1999) *J. Mol. Biol.* **287**, 877–896
391. Arents, G., and Moudrianakis, E. N. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11170–11174
392. Starich, M. R., Sandman, K., Reeve, J. N., and Summers, M. F. (1996) *J. Mol. Biol.* **255**, 187–203
393. Hud, N. V., Milanovich, F. P., and Balhorn, R. (1994) *Biochemistry* **33**, 7528–7535
394. Lohman, T. M., and Ferrari, M. E. (1994) *Ann. Rev. Biochem.* **63**, 527–570
395. Overman, L. B., and Lohman, T. M. (1994) *J. Mol. Biol.* **236**, 165–178
396. Folmer, R. H. A., Nilges, M., Folkers, P. J. M., Konings, R. N. H., and Hilbers, C. W. (1994) *J. Mol. Biol.* **240**, 341–357
397. Olah, G. A., Gray, D. M., Gray, C. W., Kergil, D. L., Sosnick, T. R., Mark, B. L., Vaughan, M. R., and Trewhella, J. (1995) *J. Mol. Biol.* **249**, 576–594
398. Kanellopoulos, P. N., Tsernoglou, D., van der Vliet, P. C., and Tucker, P. A. (1996) *J. Mol. Biol.* **257**, 1–8
399. Bogdarina, I., Fox, D. G., and Kneale, G. G. (1998) *J. Mol. Biol.* **275**, 443–452
400. Pluta, A. F., Mackay, A. M., Ainsztein, A. M., Goldberg, I. G., and Earnshaw, W. C. (1995) *Science* **270**, 1591–1594
401. Schevitz, R. W., Otwinowski, Z., Joachimiak, A., Lawson, C. L., and Sigler, P. B. (1985) *Nature (London)* **317**, 782–786
402. Zhang, R.-G, Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., and Sigler, P. B. (1987) *Nature (London)* **327**, 591–597

References

403. Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., and Sigler, P. B. (1988) *Nature (London)* **335**, 321–329
404. Shakked, Z., Guzikevich– Guerstein, G., Frolow, F., Rabinovich, D., Joachimiak, A., and Sigler, P. B. (1994) *Nature (London)* **368**, 469–473
405. Lawson, C. L., and Carey, J. (1993) *Nature (London)* **366**, 178–182
406. Lawson, C. L., and Sigler, P. B. (1988) *Nature (London)* **333**, 869–871
407. Guenot, J., Fletterick, R. J., and Kollman, P. A. (1994) *Protein Sci.* **3**, 1276–1285
408. Rastinejad, F., Artz, P., and Lu, P. (1993) *J. Mol. Biol.* **233**, 389–399
409. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., and Miller, J. H. (1994) *J. Mol. Biol.* **240**, 421–433
410. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) *Science* **271**, 1247–1254
411. Brennan, R. G., and Matthews, B. W. (1989) *J. Biol. Chem.* **264**, 1903–1906
412. Jordan, S. R., and Pabo, C. O. (1988) *Science* **242**, 893–899
413. Wolberger, C., Dong, Y., Ptashne, M., and Harrison, S. C. (1988) *Nature (London)* **335**, 789–795
414. Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M., and Harrison, S. C. (1988) *Science* **242**, 899–907
415. Bell, A. C., and Koudelka, G. B. (1993) *J. Mol. Biol.* **234**, 542–553
416. Mandel-Gutfreund, Y., Schueler, O., and Margalit, H. (1995) *J. Mol. Biol.* **253**, 370–382
417. Ferré-D'Amaré, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) *Nature (London)* **363**, 38–44
418. Junius, F. K., O'Donoghue, S. I., Nilges, M., Weiss, A. S., and King, G. F. (1996) *J. Biol. Chem.* **271**, 13663–13667
419. Keller, W., König, P., and Richmond, T. J. (1995) *J. Mol. Biol.* **254**, 657–667
420. Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature (London)* **371**, 80–83
421. Klug, A., and Schwabe, J. W. R. (1995) *FASEB J.* **9**, 597–604
422. Rhodes, D., and Klug, A. (1993) *Sci. Am.* **268** (Feb), 56–65
423. Pavletich, N. P., and Pabo, C. O. (1991) *Science* **252**, 809–817
424. Mackay, J. P., and Crossley, M. (1998) *Trends Biochem. Sci.* **23**, 1–4
425. Miller, J., McLachlan, A. D., and Klug, A. (1985) *EMBO J.* **4**, 1609–1614
426. Pieler, T., and Theunissen, O. (1993) *Trends Biochem. Sci.* **18**, 226–230
427. Pavletich, N. P., and Pabo, C. O. (1993) *Science* **261**, 1701–1707
428. Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L., and Berg, J. M. (1991) *J. Am. Chem. Soc.* **113**, 4518–4523
429. Somers, W. S., and Phillips, S. E. V. (1992) *Nature (London)* **359**, 387–393
430. Phillips, S. E. V. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 671–701
431. Breg, J. N., van Opheusden, H. J., Burgering, M. J. M., Boelens, R., and Kaptein, R. (1990) *Nature (London)* **346**, 586–589
432. Bonnefoy, E., Takahashi, M., and Yaniv, J. R. (1994) *J. Mol. Biol.* **242**, 116–129
433. Müller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L., and Harrison, S. C. (1995) *Nature (London)* **373**, 311–317
434. Ghosh, G., Van Duyne, G., Ghosh, S., and Sigler, P. B. (1995) *Nature (London)* **373**, 303–310
435. Baltimore, D., and Beg, A. A. (1995) *Nature (London)* **373**, 287–288
436. Kaufmann, E., Müller, D., and Knöchel, W. (1995) *J. Mol. Biol.* **248**, 239–254
437. Burd, C. G., and Dreyfuss, G. (1994) *Science* **265**, 615–621
438. Nagai, K., and Mattaj, I. W., eds. (1994) *RNA-Protein Interactions*, IRL Press, Oxford
439. Draper, D. E. (1995) *Ann. Rev. Biochem.* **64**, 593–620
440. Hentze, M. W. (1994) *Trends Biochem. Sci.* **19**, 101–103
441. Allain, F. H.-T., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D., and Varani, G. (1996) *Nature (London)* **380**, 646–650
442. Kenan, D. J., Query, C. C., and Keene, J. D. (1991) *Trends Biochem. Sci.* **16**, 214–220
443. Fierro-Monti, I., and Mathews, M. B. (2000) *Trends Biochem. Sci.* **25**, 241–246
444. Voyles, B. A. (1993) *The Biology of Viruses*, Mosby, St. Louis, Missouri
445. Fields, B. N., and Knipe, D. M., eds. (1990) *Fields Virology*, 2nd ed., Raven Press, New York
446. Cann, A. (1993) *Principles of Molecular Virology*, Academic Press, San Diego, California
447. Endemann, H., and Model, P. (1995) *J. Mol. Biol.* **250**, 496–506
448. Williams, K. A., Glibowicka, M., Li, Z., Li, H., Khan, A. R., Chen, Y. M. Y., Wang, J., Marvin, D. A., and Deber, C. M. (1995) *J. Mol. Biol.* **252**, 6–14
449. Wen, Z. Q., Overman, S. A., and Thomas, G. J., Jr. (1997) *Biochemistry* **36**, 7810–7820
450. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., III, Slocombe, P. M., and Smith, M. (1977) *Nature (London)* **265**, 687–695
451. McKenna, R., Ilag, L. L., and Rossmann, M. G. (1994) *J. Mol. Biol.* **237**, 517–543
452. McKenna, R., Xia, D., Willingmann, P., Ilag, L. L., Krishnaswamy, S., Rossmann, M. G., Olson, N. H., Baker, T. S., and Incardona, N. L. (1992) *Nature (London)* **355**, 137–143
453. Godson, G. N. (1978) *Trends Biochem. Sci.* **3**, 249–253
454. McKenna, R., Bowman, B. R., Ilag, L. L., Rossmann, M. G., and Fane, B. A. (1996) *J. Mol. Biol.* **256**, 736–750
455. Wu, H., and Rossmann, M. G. (1993) *J. Mol. Biol.* **233**, 231–244
456. Tsao, J., Chapman, M. S., Agbandje, M., Keller, W., Smith, K., Wu, H., Luo, M., Smith, T. J., Rossmann, M. G., Compans, R. W., and Parrish, C. R. (1991) *Science* **251**, 1456–1464
457. Chipman, P. R., Agbandje-McKenna, M., Kajigaya, S., Brown, K. E., Young, N. S., Baker, T. S., and Rossmann, M. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7502–7506
458. Howarth, A. J., and Goodman, R. M. (1982) *Trends Biochem. Sci.* **7**, 180–182
459. Tiollais, P., Pourcel, C., and Dejean, A. (1985) *Nature (London)* **317**, 489–495
460. Tiollais, P., and Buendia, M.-A. (1991) *Sci. Am.* **264**(Apr), 116–123
461. Böttcher, B., Wynne, S. A., and Crowther, R. A. (1997) *Nature (London)* **386**, 88–91
462. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L., and Weissman, S. M. (1978) *Science* **200**, 494–502
463. Liddington, R. C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T. L., and Harrison, S. C. (1991) *Nature (London)* **354**, 278–284
464. Griffith, J. P., Griffith, D. L., Rayment, I., Murakami, W. T., and Caspar, D. L. D. (1992) *Nature (London)* **355**, 652–654
465. Yang, R. C. A., and Wu, R. (1979) *Science* **206**, 456–462
466. Howley, P. M. (1986) *N. Engl. J. Med.* **315**, 1089–1090
467. Hess, M., Cuzange, A., Ruigrok, R. W. H., Chroboczek, J., and Jacrot, B. (1995) *J. Mol. Biol.* **252**, 379–385
468. Athappilly, F. K., Murali, R., Rux, J. J., Cai, Z., and Burnett, R. M. (1994) *J. Mol. Biol.* **242**, 430–455
469. Witkowski, J. A. (1988) *Trends Biochem. Sci.* **13**, 110–113
470. Sugden, B. (1991) *Trends Biochem. Sci.* **16**, 45–46
471. Zhou, Z. H., Prasad, B. V. V., Jakana, J., Rixon, F. J., and Chiu, W. (1994) *J. Mol. Biol.* **242**, 456–469
472. Trus, B. L., Booy, F. P., Newcomb, W. W., Brown, J. C., Homa, F. L., Thomsen, D. R., and Steven, A. C. (1996) *J. Mol. Biol.* **263**, 447–462
473. Henle, W., Henle, G., and Lennette, E. T. (1974) *Sci. Am.* **241**(Jul), 48–59
474. Berardelli, P. (1997) *Science* **278**, 1710
475. Senkevich, T. G., Bugert, J. J., Sisler, J. R., Koonin, E. V., Darai, G., and Moss, B. (1996) *Science* **273**, 813–816
476. Hohn, T., Hohn, B., and Pfeiffer, P. (1985) *Trends Biochem. Sci.* **10**, 205–209
477. Rosa, P., Mantovani, S., Rosboch, R., and Huttner, W. B. (1992) *J. Biol. Chem.* **267**, 12227–12232
478. Diener, T. O. (1991) *FASEB J.* **5**, 2808–2813
479. Diener, T. V. (1984) *Trends Biochem. Sci.* **9**, 133–136
480. Diener, T. O. (1979) *Viroids and Viroid Diseases*, Wiley, New York
481. Martínez-Soriano, J. P., Galindo-Alonso, J., Maroon, C. J. M., Yucel, I., Smith, D. R., and Diener, T. O. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9397–9401
482. Polson, A. G., Bass, B. L., and Casey, J. L. (1996) *Nature (London)* **380**, 454–456
483. Wang, K.-S., Choo, Q.-L., Weiner, A. J., Ou, J.-H., Najarian, R. C., Thayer, R. M., Mullenbach, G. T., Denniston, K. J., Gerin, J. L., and Houghton, M. (1986) *Nature (London)* **323**, 508–514
484. Lai, M. M. C. (1995) *Ann. Rev. Biochem.* **64**, 259–286
485. Unge, T., Liljas, L., Strandberg, B., Vaara, I., Kannan, K. K., Fridborg, K., Nordman, C. E., and Lentz, P. J., Jr. (1980) *Nature (London)* **285**, 373–377
486. Larson, S. B., Koszelak, S., Day, J., Greenwood, A., Dodds, J. A., and McPherson, A. (1993) *Nature (London)* **361**, 179–182
487. Larson, S. B., Koszelak, S., Day, J., Greenwood, A., Dodds, J. A., and McPherson, A. (1993) *J. Mol. Biol.* **231**, 375–391
488. Golmohammadi, R., Valegård, K., Fridborg, K., and Liljas, L. (1993) *J. Mol. Biol.* **234**, 620–639
489. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1995) *Biochemistry* **34**, 10600–10606
490. Stonehouse, N. J., Valegård, K., Golmohammadi, R., van den Worm, S., Walton, C., Stockley, P. G., and Liljas, L. (1996) *J. Mol. Biol.* **256**, 330–339
491. Beremarand, M. N., and Blumenthal, T. (1979) *Cell* **18**, 257–266
492. Hogle, J. M., Chow, M., and Filman, D. J. (1985) *Science* **229**, 1358–1365
493. Hogle, J. M., Chow, M., and Filman, D. J. (1987) *Sci. Am.* **256**(Mar), 42–49
494. Najaran, R., Caput, D., Gee, W., Potter, S. J., Renard, A., Merryweather, J., Van Nest, G., and Dina, D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2627–2631
495. Rossmann, M. G., Arnold, E., Griffith, J. P., Kamer, G., Luo, M., Smith, T. J., Vriend, G., Rueckert, R. R., Sherry, B., McKinlay, M. A., Diana, G., and Otto, M. (1987) *Trends Biochem. Sci.* **12**, 313–318

References

496. Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S., and Rossmann, M. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 507–511
497. Hadfield, A. T., Oliveira, M. A., Kim, K. H., Minor, I., Kremer, M. J., Heinz, B. A., Shepard, D., Pevear, D. C., Rueckert, R. R., and Rossmann, M. G. (1995) *J. Mol. Biol.* **253**, 61–73
498. Küpper, H., Keller, W., Kurz, C., Forss, S., Schaller, H., Franze, R., Stohmaier, K., Marquardt, O., Zaslavsky, V. G., and Hofschneider, P. H. (1981) *Nature (London)* **289**, 555–559
499. Brown, F. (1981) *Trends Biochem. Sci.* **6**, 325–327
500. Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D., and Brown, F. (1989) *Nature (London)* **337**, 709–716
501. Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossmann, M. G., Boege, U., Scraba, D. G., Duke, G. M., and Palmenberg, A. C. (1987) *Science* **235**, 182–191
502. Choi, H.-K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G., and Wengler, G. (1991) *Nature (London)* **354**, 37–43
503. Simons, K., Garoff, H., and Helenius, A. (1982) *Sci. Am.* **246**(Feb), 58–66
504. Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K., and Bricogne, G. (1978) *Nature (London)* **276**, 368–373
505. Erickson, J. W., Silva, A. M., Murthy, M. R. N., Fita, I., and Rossmann, M. G. (1985) *Science* **229**, 625–629
506. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R., and Klug, A. (1978) *Nature (London)* **276**, 362–368
507. Holmes, K. C. (1980) *Trends Biochem. Sci.* **5**, 4–7

507a. Lauffer, M. A. (1984) *Trends Biochem. Sci.* **9**, 369–371

508. Wang, H., and Stubbs, G. (1994) *J. Mol. Biol.* **239**, 371–384
509. Kaplan, M. M., and Webster, R. G. (1977) *Sci. Am.* **237**(Dec), 88–106
510. von Itzstein, M., and 17 other authors. (1993) *Nature (London)* **363**, 418–423
511. Clarke, D. K., Duarte, E. A., Elena, S. F., Moya, A., Domingo, E., and Holland, J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4821–4824
512. Varmus, H. (1988) *Science* **240**, 1427–1434
513. Darlix, J.-L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B. P. (1995) *J. Mol. Biol.* **254**, 523–537
514. Löwer, R., Löwer, J., and Kurth, R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5177–5184
515. Stephenson, J. R., ed. (1980) *Molecular Biology of RNA Tumor Viruses*, Academic Press, New York
516. Compans, R. W., and Bishop, D. H. L., eds. (1983) *Double-Stranded RNA Viruses*, Elsevier, Biomedical, New York
517. Thornton, A., and Zuckerman, A. J. (1975) *Nature (London)* **254**, 557–558
518. Baldwin, M. A., Cohen, F. E., and Prusiner, S. B. (1995) *J. Biol. Chem.* **270**, 19197–19200
519. Manuelidis, L., Sklaviadis, T., Akowitz, A., and Fritch, W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5124–5128
520. Prusiner, S. B. (1996) *Trends Biochem. Sci.* **21**, 482–487
521. Prusiner, S. B. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13363–13383
522. Weissmann, C. (1999) *J. Biol. Chem.* **274**, 3–6
523. Chesebro, B. (1998) *Science* **286**, 660–662
524. Balter, M. (1999) *Science* **286**, 660–662
525. Temin, H. M. (1978) *Trends Biochem. Sci.* **3**, N80–N83

525a. Sambrook, J., and Russell, D. (2000) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Lab. Press, Plainview, New York

526. Grossman, L., and Moldave, K., eds. (1967) *Methods in Enzymology*, Vol. 12A, Academic Press, New York (pp. 531–708)
527. Robyt, J. F., and White, B. J. (1996) *Biochemical Techniques, Theory and Practice*, 2nd ed., Waveland Press, Prospect Heights, Illinois
528. Blackburn, G. M., and Gait, M. J., eds. (1996) *Nucleic Acids in Chemistry and Biology*, 2nd ed., Oxford Univ. Press, Oxford
529. Rajeswari, M. R., Montenay-Garestier, T., and Helene, C. (1987) *Biochemistry* **26**, 6825–6831
530. Hecker, R., Colpan, M., and Riesner, D. (1985) *Journal of Chromatography* **326**, 251–261
531. Edwardson, P. A. D., Atkinson, T., Lowe, C. R., and Small, D. A. P. (1986) *Anal. Biochem.* **152**, 215–220
532. Lawrance, S. K., Smith, C. L., Srivastava, R., Cantor, C. R., and Weissman, S. M. (1987) *Science* **235**, 1387–1390
533. Cantor, C. R., Smith, C. L., and Matthew, M. K. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 287–304
534. Smith, C. L., and Cantor, C. R. (1987) *Trends Biochem. Sci.* **12**, 284–287
535. Kochetkov, N. K., and Budovskii, E. I. (1972) *Organic Chemistry of Nucleic Acids*, Plenum, New York (Part A, pp. 137–147)
536. Bozic, D., Grazulis, S., Siksnys, V., and Huber, R. (1996) *J. Mol. Biol.* **255**, 176–186
537. Smith, H. O. (1979) *Science* **205**, 455–462
538. Air, G. M. (1979) *CRC Crit. Revs. Biochem.* **6**, 1–33
539. Roberts, R. J. (1980) *Methods Enzymol.* **65**, 1–15
540. Koob, M., and Szybalski, W. (1990) *Science* **250**, 271–273
541. Veselkov, A. G., Demidov, V. V., Frank-Kamenetskii, M. D., and Nielsen, P. E. (1996) *Nature (London)* **379**, 214
542. Zaug, A. J., Been, M. D., and Cech, T. R. (1986) *Nature (London)* **324**, 429–433
543. Uhlenbeck, O. C. (1987) *Nature (London)* **328**, 596–600
544. Landgraf, R., Chen, C.-hB., and Sigman, D. S. (1994) *Biochemistry* **33**, 10607–10615
545. Adams, R. L. P., Burdon, R. H., Campbell, A. M., and Smellie, R. M. S. (1976) *Davidson's The Biochemistry of the Nucleic Acids*, 8th ed., Academic Press, New York (pp. 50–82)
546. Rayford, R., Anthony, D. D., Jr., O'Neill, R. E., Jr., and Merrick, W. C. (1985) *J. Biol. Chem.* **260**, 15708–15713
547. Bochner, B. R., and Ames, B. N. (1982) *J. Biol. Chem.* **257**, 9759–9769
548. Chen, H., and Shaw, B. R. (1994) *Biochemistry* **33**, 4121–4129
549. Hayatsu, H., and Shiragami, M. (1979) *Biochemistry* **18**, 632–647
550. Spiro, T. G., ed. (1980) *Nucleic Acid-Metal Ion Interactions*, Wiley, New York
551. Dische, Z. (1955) *Nucleic Acids* **1**, 755
552. Marmur, J., and Doty, P. (1962) *J. Mol. Biol.* **5**, 109–118
553. Korolev, N., Lyubartsev, A. P., and Nordenskiöld, L. (1998) *Biophys. J.* **75**, 3041–3056
554. Davidson, J. N. (1972) *The Biochemistry of Nucleic Acids*, 7th ed., Academic Press, New York (p. 148)
555. Wada, A., Yabuki, S., and Husimi, Y. (1980) *Crit. Revs. Biochem.* **9**, 87–144
556. Britten, R. J., and Kohne, D. E. (1968) *Science* **161**, 529–540
557. Wilson, D. A., and Thomas, C. A., Jr. (1974) *J. Mol. Biol.* **84**, 115–144
558. Bendich, A. J., and Bolton, E. T. (1968) in *Methods in Enzymology*, Vol. 12B (Grossman, L., and Moldave, K., eds), pp. 635–640, Academic Press, New York
559. Gillespie, D. (1968) in *Methods in Enzymology*, Vol. 12B (Grossman, L., and Moldave, K., eds), pp. 641–668, Academic Press, New York
560. Hall, B. D., Haarr, L., and Kleppe, K. (1980) *Trends Biochem. Sci.* **5**, 254–256
561. Britten, R. J., and Kohne, D. E. (1967) *Carnegie Instit. Washington Yearbook*, 65th ed., Carnegie Instit., Washington, D. C. (pp. 78–106)
562. Marmur, J. (1994) *Trends Biochem. Sci.* **19**, 343–346
563. Szabo, P., and Ward, D. C. (1982) *Trends Biochem. Sci.* **7**, 425–427
564. Kessler, C., ed. (1992) *Nonradioactive Labeling and Detection of Biomolecules*, Springer-Verlag, New York
565. Leitch, A. R., Chwarzacher, T., Jackson, D., and Leitch, I. J. (1994) *In Situ Hybridization: a Practical Guide*, Bios, Oxford
566. Kumar, A., Tchen, P., Roullet, F., and Cohen, J. (1988) *Anal. Biochem.* **169**, 376–
567. Isaac, P. G., ed. (1994) *Protocols for Nucleic Acid Analysis by Nonradioactive Probes*, Humana Press, Totowa, New Jersey
568. Gillevet, P. M. (1990) *Nature (London)* **348**, 657–658
569. Negro, F., Pacchioni, D., Shimizu, Y., Miller, R. H., Bussolati, G., Purcell, R. H., and Bonino, F. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2247–2251
570. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
571. Gubler, U., and Hoffman, B. J. (1983) *Gene* **25**, 263–269
572. Watson, C. J., and Jackson, J. F. (1985) in *DNA Cloning Vol. I: A Practical Approach* (Glover, D. M., ed), pp. 79–100, IRL Press, Washington, DC
573. Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y., and Matsubara, K. (1985) *J. Biol. Chem.* **260**, 2605–2608
574. Warner, B. D., Warner, M. E., Karus, G. A., Ku, L., Brown-Shimer, S., and Urdea, M. S. (1983) *DNA* **4**, 401–411
575. Sekiya, T., Takeya, T., Brown, E. L., Belagaje, R., Contreras, R., Fritz, H.-J., Gait, M. J., Lees, R. G., Ryan, M. J., Khorana, H. G., and Norris, K. E. (1979) *J. Biol. Chem.* **254**, 5787–5801
576. Wada, T., Sato, Y., Honda, F., Kawahara, S.-i, and Sekine, M. (1997) *J. Am. Chem. Soc.* **119**, 12710–12721
577. Matteucci, M. D., and Caruthers, M. H. (1981) *J. Am. Chem. Soc.* **103**, 3185–3191
578. Beaucage, S. L., and Caruthers, M. H. (1981) *Tetrahedron Letters* **22**, 1859–1862
579. Air, G. M. (1979) *CRC Crit. Revs. Biochem.* **6**, 1–33
580. Alvarado-Urbina, G., Sathe, G. M., Liu, W.-C., Gillen, M. F., Duck, P. D., Bender, R., and Ogilvie, K. K. (1981) *Science* **214**, 270–274
581. Caruthers, M. H. (1985) *Science* **230**, 281–285
582. Lashkari, D. A., Hunicke-Smith, S. P., Norgren, R. M., Davis, R. W., and Brennan, T. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7912–7915
583. Usman, N., Ogilvie, K. K., Jiang, M.-Y., and Cedergren, R. J. (1987) *J. Am. Chem. Soc.* **109**, 7845–7854
584. Ebe, K., Schöed, M., Rossi, J. J., and Wallace, R. B. (1987) *DNA* **6**, 497–504
585. Usman, N., and Cedergren, R. (1992) *Trends Biochem. Sci.* **17**, 334–339
586. Eaton, B. E., and Pieken, W. A. (1995) *Ann. Rev. Biochem.* **64**, 837–863
587. Rohatgi, R., Bartel, D. P., and Szostak, J. W. (1996) *J. Am. Chem. Soc.* **118**, 3340–3344
588. Kleppe, K., Ohtsuka, E., Kleppe, R., and Khorana, H. G. (1971) *J. Mol. Biol.* **56**, 341–361
589. Panet, A., and Khorana, H. G. (1974) *J. Biol. Chem.* **249**, 5213–5221

590. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) *Science* **230**, 1350 – 1354
591. Mullis, K. B. (1990) *Sci. Am.* **262**(Apr), 56 – 65
592. Rabinow, P. (1996) *Making PCR: A Story of Biotechnology*, Univ. Chicago Press, Chicago, Illinois
593. Mullis, K., Ferré, F., and Gibbs, R., eds. (1994) *The Polymerase Chain Reaction*, Birkhäuser, Boston, Massachusetts
594. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) *Science* **239**, 487 – 491
595. Erlich, H. A., Gelfand, D. H., and Saiki, R. K. (1988) *Nature (London)* **331**, 461 – 462
596. Appenzeller, T. (1990) *Science* **247**, 1030 – 1032
597. Newton, C., and Graham, A. (1997) *PCR*, 2nd ed., Springer-Verlag, New York
598. Bloch, W. (1991) *Biochemistry* **30**, 2735 – 2747
599. White, B. A., ed. (1993) *PCR Protocols*, Humana Press, Totowa, New Jersey
599a. Innis, M. A., Gelfand, D. H., and Sninsky, J. J., eds. (1999) *PCR Applications*, Academic Press, San Diego, California
600. Cohen, J. (1994) *Science* **263**, 1564 – 1565
601. Pääbo, S. (1993) *Sci. Am.* **269**(Nov), 86 – 92
602. Pääbo, S., Higuchi, R. G., and Wilson, A. C. (1989) *J. Biol. Chem.* **264**, 9709 – 9712
603. Höss, M., Pääbo, S., and Vereshchagin, N. K. (1994) *Nature (London)* **370**, 333
604. Scharf, S. J., Horn, G. T., and Erlich, H. A. (1986) *Science* **233**, 1076 – 1078
605. Li, H., Gyllensten, U. B., Cui, X., Saiki, R. K., Erlich, H. A., and Arnheim, N. (1988) *Nature (London)* **335**, 414 – 417
606. Engelke, D. R., Hoener, P. A., and Collins, F. S. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 544 – 548
607. Karlovsky, P. (1990) *Trends Biochem. Sci.* **15**, 419
608. Hengen, P. N. (1995) *Trends Biochem. Sci.* **20**, 372 – 373
609. Wong, G. K.-S., Yu, J., Thayer, E. C., and Olson, M. V. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5225 – 5230
610. Bernardi, G. (1978) *Nature (London)* **276**, 558 – 559
611. Bonitz, S. G., Coruzzi, G., Thalenfeld, B. E., Tzagoloff, A., and Macino, G. (1980) *J. Biol. Chem.* **255**, 11927 – 11941
612. Martin, N. C., Miller, D. L., and Donelson, J. E. (1979) *J. Biol. Chem.* **254**, 11729 – 11734
613. Fischer, S. G., and Lerman, L. S. (1979) *Methods Enzymol.* **68**, 183 – 191
614. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503 – 517
615. Southern, E. (1979) *Methods Enzymol.* **68**, 152 – 176
616. Grunstein, M., and Wallis, J. (1979) *Methods Enzymol.* **68**, 379 – 389
617. Sanger, F., and Coulson, A. R. (1975) *J. Mol. Biol.* **94**, 441 – 448
618. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 – 5467
619. Alphey, L. (1997) *DNA Sequencing*, Springer-Verlag, New York
620. Ansorge, W., Voss, H., and Zimmerman, J., eds. (1997) *DNA Sequencing Strategies - Automated and Advanced Approaches*, Wiley, New York
621. Griffin, H. G., and Griffin, A. M., eds. (1993) *DNA Sequencing Protocols*, Humana Press, Totowa, New Jersey
622. Smith, A. J. H. (1980) *Methods Enzymol.* **65**, 560 – 580
623. Gillam, S., and Smith, M. (1980) *Methods Enzymol.* **65**, 687 – 701
624. Chen, E. Y., and Seeburg, P. H. (1985) *DNA* **4**, 165 – 170
625. Ornstein, D., Moen, P. T., and Kashdan, M. A. (1985) *DNA* **4**, 94
626. Bassam, B. J., Caetano-Anolles, G., and Gresshoff, P. M. (1991) *Anal. Biochem.* **196**, 80 – 83
627. Mizusawa, S., Nishimura, S., and Seela, F. (1986) *Nucleic Acids Res.* **14**, 1319 – 1324
628. Gerken, T. A., Gupta, R., and Jentoft, N. (1992) *Biochemistry* **31**, 639 – 648
629. Hengen, P. N. (1996) *Trends Biochem. Sci.* **21**, 33 – 34
630. Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., and Surrey, S. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4298 – 4302
631. Henikoff, S. (1984) *Gene* **28**, 351 – 359
632. Adams, M. D., Kerlavage, A. R., Kelley, J. M., Gocayne, J. D., Fields, C., Fraser, C. M., and Venter, J. C. (1994) *Nature (London)* **368**, 474 – 475
633. Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H., and Hood, L. E. (1986) *Nature (London)* **321**, 674 – 679
634. Ju, J., Ruan, C., Fuller, C. W., Glazer, A. N., and Mathies, R. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4347 – 4351
635. Fu, D.-J., Broude, N. E., Köster, H., Smith, C. L., and Cantor, C. R. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10162 – 10166
636. Bult, C. J., and 39 other authors. (1996) *Science* **273**, 1058 – 1073
637. Sutcliffe, J. G. (1995) *Trends Biochem. Sci.* **20**, 87 – 90
638. Rowen, L., Mahairas, G., and Hood, L. (1997) *Science* **278**, 605 – 607
639. Mathies, R. A., Huang, X. C., and Quesada, M. A. (1992) *Anal. Chem.* **64**, 2149 – 2154
640. Yeung, E. S., and Li, Q. (1998) in *High Performance Capillary Electrophoresis* (Khaldi, M. G., ed), Wiley, New York
641. Little, D. P., Aaserud, D. J., Valaskovic, G. A., and McLafferty, F. W. (1996) *J. Am. Chem. Soc.* **118**, 9352 – 9359
642. Maxam, A. M., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560 – 564
643. Gilbert, W. (1981) *Science* **214**, 1305 – 1312
644. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499 – 560
645. Brownlee, G. G. (1972) in *Determination of Sequences in RNA* (Work, T. S., and Work, E., eds), North-Holland Publ., Amsterdam
646. Pilly, D., Niemeyer, A., Schmidt, M., and Bargetzi, J. P. (1978) *J. Biol. Chem.* **253**, 437 – 445
647. Bayev, A. A. (1995) *Comprehensive Biochemistry* **38**, 439 – 479
648. Campbell, P. N. (1995) *Trends Biochem. Sci.* **20**, 259 – 260
649. Myers, T. W., and Gelfand, D. H. (1991) *Biochemistry* **30**, 7661 – 7666
650. Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R., and Guilley, H. (1977) *Nature (London)* **269**, 833 – 836
651. Stanley, J., and Vassilenko, S. (1978) *Nature (London)* **274**, 87 – 89
652. Kitamura, N., and Wimmer, E. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3196 – 3200
653. Josse, J., Kaiser, A. D., and Kornberg, A. (1961) *J. Biol. Chem.* **236**, 864 – 875
654. Stormo, G. D. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 241 – 263
655. Goad, W. B. (1986) *Ann. Rev. Biophys. Biophys. Chem.* **15**, 79 – 95
656. Garner, M. M., and Revzin, A. (1986) *Trends Biochem. Sci.* **11**, 395 – 396
657. Berger, R., Duncan, M. R., and Berman, B. (1993) *BioTechniques* **15**, 650 – 652
658. Welsh, J., and Cantor, C. R. (1984) *Trends Biochem. Sci.* **9**, 505 – 508
659. Galas, D., and Schmitz, A. (1978) *Nucleic Acids Res.* **5**, 3157 – 3170
660. Tullius, T. D. (1987) *Trends Biochem. Sci.* **12**, 297 – 300
661. Tullius, T. D., and Dombroski, B. A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5469 – 5473
662. Prigodich, R. V., and Martin, C. T. (1990) *Biochemistry* **29**, 8017 – 8019
663. Flaus, A., Luger, K., Tan, S., and Richmond, T. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1370 – 1375
664. Tornaletti, S., and Pfeifer, G. P. (1995) *J. Mol. Biol.* **249**, 714 – 728
664a. Chaulk, S. G., Pezacki, J. P., and MacMillan, A. M. (2000) *Biochemistry* **39**, 10448 – 10453
665. Jaeger, J. A., SantaLucia, J., Jr., and Tinoco, I., Jr. (1993) *Ann. Rev. Biochem.* **62**, 255 – 287
666. Hüttenhofer, A., and Noller, H. F. (1994) *EMBO J.* **13**, 3892 – 3901
667. Laughrea, M., and Tam, J. (1992) *Biochemistry* **31**, 12035 – 12041
668. Reid, B. R. (1981) *Ann. Rev. Biochem.* **50**, 969 – 996
669. Hare, D. R., Ribeiro, N. S., Wemmer, D. E., and Reid, B. R. (1985) *Biochemistry* **24**, 4300 – 4306
670. Ikuta, S., Chattopadhyaya, R., Ito, H., Dickerson, R. E., and Kearns, D. R. (1986) *Biochemistry* **25**, 4840 – 4849
671. Lefèvre, J.-F., Lane, A. N., and Jardetzky, O. (1987) *Biochemistry* **26**, 5076 – 5090
672. Bax, A., and Lerner, L. (1986) *Science* **232**, 960 – 967
673. Patel, D. J., Shapiro, L., and Hare, D. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 423 – 454
674. Flynn, P. F., Kintanar, A., Reid, B. R., and Drobny, G. (1988) *Biochemistry* **27**, 1191 – 1197
675. Cohen, J. S. (1987) *Trends Biochem. Sci.* **12**, 133 – 135
676. Wemmer, D. E., Chou, S.-H., and Reid, B. R. (1984) *J. Mol. Biol.* **180**, 41 – 60
677. Schmidt, P. G., Sierzputowska-Gracz, H., and Agris, P. F. (1987) *Biochemistry* **26**, 8529 – 8534
678. Farmer, B. T., Jr., and Müller, L. (1993) *J. Am. Chem. Soc.* **115**, 11040 – 11041
679. Nikonowicz, E. P., and Pardi, A. (1992) *Nature (London)* **355**, 184 – 186
680. Tate, S.-i, Ono, A., and Kainosho, M. (1994) *J. Am. Chem. Soc.* **116**, 5977 – 5978
681. Marino, J. P., Prestegard, J. H., and Crothers, D. M. (1994) *J. Am. Chem. Soc.* **116**, 2205 – 2206
682. Moore, P. B. (1995) *Acc. Chem. Res.* **28**, 251 – 256
683. Pardi, A. (1995) *Nature Struct. Biol.* **1**, 846 – 849
684. Mer, G., and Chazin, W. J. (1998) *J. Am. Chem. Soc.* **120**, 607 – 608
684a. Nikonowicz, E. P., Michnicka, M., and DeJong, E. (1998) *J. Am. Chem. Soc.* **120**, 3813 – 3814
684b. Liu, A., Majumdar, A., Hu, W., Kettani, A., Skripkin, E., and Patel, D. J. (2000) *J. Am. Chem. Soc.* **122**, 3206 – 3210
685. Tolbert, T. J., and Williamson, J. R. (1996) *J. Am. Chem. Soc.* **118**, 7929 – 7940
686. Xu, J., Lapham, J., and Crothers, D. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 44 – 48
687. Horowitz, J., Ofengand, J., Daniel, W. E. Jr., and Cohn, M. (1977) *J. Biol. Chem.* **252**, 4418 – 4420
688. Schroeder, S. A., Roongta, V., Fu, J. M., Jones, C. R., and Gorenstein, D. G. (1989) *Biochemistry* **28**, 8292 – 8303
689. Radhakrishnan, I., Patel, D. J., and Gao, X. (1992) *Biochemistry* **31**, 2514 – 2523
690. Schurter, E. J., Sayer, J. M., Oh-hara, T., Yeh, H. J. C., Yagi, H., Luxon, B. A., Jerina, D. M., and Gorenstein, D. G. (1995) *Biochemistry* **34**, 9009 – 9020
691. Mao, H., Deng, Z., Wang, F., Harris, T. M., and Stone, M. P. (1998) *Biochemistry* **37**, 4374 – 4387
692. Chou, S.-H., Zhu, L., Gao, Z., Cheng, J.-W., and Reid, B. R. (1996) *J. Mol. Biol.* **264**, 981 – 1001
693. Kettani, A., Guéron, M., and Leroy, J.-L. (1997) *J. Am. Chem. Soc.* **119**, 1108 – 1115
694. Chu, Wen-C., Liu, J. C.-H., and Horowitz, J. (1997) *Nucleic Acids Res.* **25**, 3944 – 3949

Study Questions

1. Describe the typical distribution pattern of RNA and DNA in bacterial cells and in eukaryotic cells.

2. Draw the structures of the Watson–Crick base pairs guanine–cytosine (GC) and adenine–thymine (AT). Also draw the GU pair, which is not a Watson–Crick pair.

3. Draw the tautomeric structures possible for the cation formed by protonation of 9-methyladenine.

4. What unusual base pairs could arise from a minor tautomer of cytosine or from a minor tautomer of guanine?

5. The minor imino tautomeric form of adenosine occurs infrequently in DNA. Can this cause mutations? Explain; draw structures to illustrate your answer.

6. Will the substitution of hypoxanthine for adenine in DNA result in mutation? Explain.

7. Why is the methylation of DNA to form O^6-methyl guanine mutagenic?

8. Draw the structure of a dinucleotide that might be obtained by the partial hydrolysis of RNA. Indicate the following:
   a) The 5' end
   b) The 3' end
   c) The torsion angle χ
   d) The point of cleavage by pancreatic ribonuclease
   e) The point of cleavage by periodic acid
   f) Two points at which the structure might be methylated by modifying enzymes acting on a polynucleotide

9. Draw the structure of guanosine-5'-phosphate in such a way that the configurations of the sugar ring and of the glycosidic linkage are clearly indicated. State whether you have drawn a *syn* or an *anti* conformer. Circle the most acidic proton in the guanine ring and indicate its approximate pK_a. Which is the most basic center? What is the approximate pK_a of the conjugate acid?

10. What are the chemical functional groups in DNA? In RNA?

11. Electrophoresis of a mixture of the dinucleotides ApC and ApU at pH 3.5 separates two components. Identify these and explain the order of migration. Be as quantitative as possible.

12. Draw the structure of the predominant form of pGpC as it occurs at pH 3.5.

13. Why is DNA denatured at pH 11?

14. Draw a schematic representation of the polynucleotide portion of a DNA molecule and of an RNA molecule and indicate positions of cleavage by the following treatments:
   a) Mild HCl
   b) More vigorous HCl
   c) Mild NaOH
   d) More vigorous NaOH
   e) Pancreatic RNase
   f) Pancreatic DNase
   g) Splenic DNase
   h) Splenic phosphodiesterase
   i) Snake venom phophodiesterase
   j) Dnase from *Micrococcus*

15. A sample of DNA from a virus was hydrolyzed by acid and was found to have the following base composition (in mol%): adenine, 30; thymine, 39; guanine, 18; cytosine, 13. This differs from that of most DNA preparations. Offer a possible explanation. Sketch the expected temperature-absorbance profile of this DNA. Do you expect much hyperchromicity? Explain your answer.

16. Adenine is found to constitute 16.3% of the nucleic acid bases in a sample of bacterial DNA. What are the percentages of the other three bases?

17. For the following DNA sequence

    3'-CGATACGGCTATGCCATAGGC-5'

    write a) the sequence of the complementary DNA strand;
    b) the sequence of the corresponding segment of mRNA formed using the DNA segment above as the template;
    c) the amino acid sequence encoded by this segment.

18. What is the molecular mass of a segment of B-DNA that encodes a 386-residue protein? What is the length in nm? in Å? Do not make allowance for introns.

19. Complete the following table:

Name	Monomer	Linkage	Range of molecular masses
Protein			
Polysaccharide			
Nucleic acid			
Teichoic acid			
Poly-β-hydroxybutyrate			

20. What is meant by the T_m of a DNA sample? How does T_m vary with base composition and what is the explanation of this?

21. Isolated "naked" bacterial DNA, from which proteins have been removed, is supercoiled. DNA in the bacterial chromosome is also supercoiled. When naked DNA is nicked, its supercoiling is abolished. In contrast nicking the chromosomal DNA does not abolish its supercoiling. Explain.

22. A closed circular duplex DNA has a 90 base-pair segment of alternating G and C residues. Upon transfer to a solution containing a high salt concentration, this segment undergoes a transition from the B conformation to the Z conformation.
    a) Explain why the high salt concentration induces a B→Z transformation.
    b) What changes would you expect in (1) the linking number *Lk*, (2) the writhe *Wr*, and (3) the twist *Tw* of the DNA as a result of this transition.

23. Name two or more characteristics of a DNA sequence or of its environment that will favor conversion of B-DNA into Z-DNA.

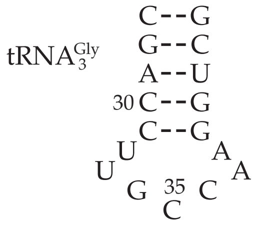
24. Suppose one double helical turn of a superhelical DNA molecule changes from a B conformation to the Z conformation. Calculate the approximate changes in (1) the linking ΔLk, (2) the writhe ΔWr, and (3) the twist ΔTw of the DNA as a result of this transition. Show your calculations and explain your answers. For this problem assume that the B form of DNA has 10.4 bp per turn. Why is the B→Z transition favored in naturally occurring supercoiled DNA?

25. A circular DNA plasmid of length 1144 bp is supercoiled with a twist (*Tw*) of 110. Assume that the DNA has 10.4 bp per turn in its relaxed state.
    a) What is the linking number *Lk* and the writhe *Wr* in the plasmid?
    b) Is the plasmid negatively or positively supercoiled?
    c) Ethidium bromide is an intercalating agent that inserts between the stacked base pairs, separating the stacks and causing local unwinding that decreases the value of *Tw*. What effect would ethidium bromide have on the migration rate of the plasmid during electrophoresis?
    d) If part of the plasmid were to undergo a transition from B-DNA to Z-DNA, what would be the effect on *Lk*, *Tw*, and *Wr*?

26. You have been given a sample of nucleic acid, describe two ways you could determine whether it is RNA or DNA and two ways to determine whether it is single- or double-stranded.

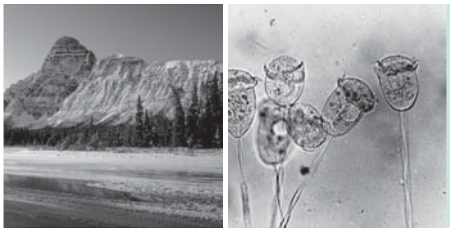
27. What conclusions can you draw about the nature of the protein binding site on DNA from the observation that methylation of cytosine residues in the protein-recognition sequence inhibits protein binding?

28. The anticodon loop of one of the tRNA Gly molecules from *E. coli* is as follows. Identify the anticodon, reading from 3' to 5'. This tRNA recognizes two different Gly codons. What are they? Write them from 5' to 3'.



The complete tRNA contains 75 nucleotides. Sketch the rest of the molecule in the cloverleaf representation. Label the 5' and 3' ends and the dihydrouridine and TψC loops. What are the last three nucleotides at the 3' end?

29. Are viruses alive? Explain your answer.



Ice and water are in equilibrium at 0°C and atmospheric pressure. When ice melts under these conditions the heat Q absorbed from the surroundings is the **enthalpy change**, ΔH, which equals 6.008 kJ mol^{-1}. For a reversible reaction the **entropy change**, ΔS, equals $\Delta H/T$ and the entropy increases when the ice melts by 6.008×10^3 J / 273.16 = 22.0 JK^{-1}. Along the frozen edges of the river the ice and water are not in a true equilibrium but in a steady state. It is an **open system** in which water flows and energy is exchanged with the surroundings. Similarly, the cells of *Vorticella* represent open systems through which water and nutrients flow, and energy is exchanged with the surrounding water. Many chemical reactions within the cells are near equilibrium while the organism maintains its own structure. Photos: Frosty River, Banff Natl. Park, Alberta © Stephen J. Kraseman; *Vorticella* courtesy of Ralph Buchsbaum.

Contents

281 **A. Thermodynamics**
282 1. The First Law of Thermodynamics
282 2. Enthalpy Changes and Thermochemistry
282 *Processes at constant pressure*
282 *Enthalpies of combustion and physiological fuel values*
284 3. The Second Law of Thermodynamics
284 *The thermodynamic temperature*
285 *Entropies from measurement of heat capacities*
285 4. A Criterion of Spontaneity: The Gibbs Energy
286 5. Practical Thermochemistry
286 *Summing changes in Gibbs energy*
287 *Reactions in solution*
287 *$\Delta G°$ and the equilibrium constant*
288 *Activity coefficients and concentration equilibrium constants*
289 *Changes in equilibria with temperature*
289 6. Thermodynamics and Life Processes
292 **B. Tables of $\Delta G°$ Values for Biochemical Compounds**
292 1. Gibbs Energies of Formation
292 2. Gibbs Energies of Dissociation of Protons
292 3. Group Transfer Potentials
293 4. "Constants" That Vary with pH and Magnesium Ion Concentrations
297 5. A New Standard for Biochemical Thermodynamic?
297 6. Bond Energies and Approximate Methods for Estimation of Thermodynamic Data
297 7. Gibbs Energies of Combustion by O_2 and by NAD^+
300 **C. Electrode Potentials and Gibbs Energy Changes for Oxidation–Reduction Reactions**
302 **D. The Adenylate System**
302 1. Storage and Utilization of Energy
303 2. Synthesis of ATP
303 3. Creatine Phosphate, an Energy Buffer
303 4. Phosphorus-31 NMR
304 **E. Complex Biochemical Equilibria**
305 1. Effects of pH on Equlibria
305 2. Microscopic Dissociation Constants and Tautomerization
307 3. The Binding of Metal Ions
307 *Factors affecting the strength of binding of a metal in a complex*
311 *How properties of the metal ion affect chelation*
311 *Metal binding sites in cells*
312 *Calcium-binding proteins*

319 **References**
321 **Study Questions**

Boxes

295 Box 6-A Measurement of Intracellular pH
299 Box 6-B Magnesium
308 Box 6-C Linear Gibbs Energy Relationships
314 Box 6-D Calcium
317 Box 6-E Metallothioneins

Tables

283 Table 6-1 Units of Energy and Work and the Values of Some Physical Constants
283 Table 6-2 Caloric Values of Food Components
285 Table 6-3 Entropies of Selected Substances
290 Table 6-4 Gibbs Energies of Formation and of Oxidation at 25°C for Compounds of Biochemical Interest
293 Table 6-5 Values of pKa, $\Delta G°$, and $\Delta H°$ for Ionization of Acids at 25°C
294 Table 6-6 Gibbs Energies of Hydrolysis at 25°C (in kJ mol^{-1})
298 Table 6-7 Empirical Bond Energies and Resonance Energies
301 Table 6-8 Reduction Potentials of Some Biologically Important Systems
310 Table 6-9 Logarithms of Binding Constants for Some 1:1 Metal Complexes at 25°C
311 Table 6-10 Ionic Radii in Nanometers for Some Metallic and Nonmetallic Ions

Thermodynamics and Biochemical Equilibria

6

We all know from experience the importance of energy to life. We know that we must eat and that hard work not only tires us but also makes us hungry. Our bodies generate heat, an observation that led Lavoisier around 1780 to the conclusion that respiration represented a slow combustion of foods within the body. It soon became clear that respiration must provide the energy for both the mechanical work done by muscles and the chemical synthesis of body constituents. All organisms require energy and the ways in which living things obtain and utilize energy is a major theme of biology.

The discovery of the first and second laws of thermodynamics permitted the development of precise, quantitative relationships between heat, energy, and work. It also allowed **chemical equilibria** to be understood. Modern biochemical literature abounds with references to the thermodynamic quantities **energy** E, **enthalpy** H, **entropy** S, and **Gibbs energy** (also called free energy) G. The purposes of this chapter are: (1) to provide a short review of thermodynamic equations, (2) to provide tables of thermodynamic quantities for biochemical substances and to explain the use of these data in the consideration of equilibria in biochemical systems, and (3) to introduce the **adenylate system**, which consists of adenosine triphosphate (ATP), adenosine diphosphate (ADP), adenosine mono-phosphate (AMP), and inorganic phosphate (P_i). This system plays a central role in energy metabolism, and (4) to provide a quantitative understanding of the effects of pH and of metal ions on biochemical equilibria. *Many readers will want to go directly to Section D, which deals with the adenylate system and its significance for life.*

A. Thermodynamics

Thermodynamics is concerned with the quantitative description of heat and energy changes and of chemical equilibria.[1–10] Knowledge of *changes* in thermodynamic quantities, such as ΔH and ΔS, enables us to predict the equilibrium positions in reactions and whether or not under given circumstances a reaction will or will not take place. Furthermore, the consideration of thermodynamic quantities provides insight into the nature of forces responsible for bonding between molecules, enzymatic catalysis, functioning of DNA and RNA, and many other phenomena.

It is important to realize that while thermodynamic information will tell us whether or not a reaction can take place it says nothing about the rate of the reaction. It will not even say whether a reaction will proceed at all within a given period of time. This has led to the occasional assertion that thermodynamics is not relevant to biochemistry. This is certainly not true; it is important to understand energy relationships in biochemical reactions. At the same time, one should avoid the trap of assuming that thermodynamic calculations appropriate for *equilibrium* situations can always be applied directly to the *steady state* found in a living cell.

Thermodynamics is an exact science and its laws deal with measurable quantities whose values are determined only by the **state** of the **system** under consideration. For example, the system might be the solution in a flask resting in a thermostated bath. To specify its state we would have to say whether it is pure solid, liquid, or gas, or a solution of specified composition and give the temperature and pressure. The flask, the bath, and everything else would be called **surroundings** or **environment**. The system plus surroundings is sometimes referred to as **the universe**.

1. The First Law of Thermodynamics

The first law of thermodynamics asserts the conservation of energy and also the equivalence of **work** and **heat**. Work and heat are both regarded as energy in transit. Heat may be absorbed by a system from the surroundings or evolved by a system and absorbed in the surroundings. Work can be done *by* a system on the surroundings or it can be done *on* a system. The first law postulates that there is an **internal energy** E (also designated U), which is dependent only on the present state of the system and in no way is dependent upon the history of the system. The first law states that *E can be changed only by the flow of energy as heat or by work*. In other words, energy can neither be created nor destroyed.

In mathematical form, the first law is given as follows:

$$\Delta E = E\,(\text{products}) - E\,(\text{reactants}) = Q - W \quad (6\text{-}1)$$

Here Q is the heat absorbed by the system from the surroundings and W is the work done by the system on the surroundings. Energy, heat, and work are all measured in the same units. Chemists have traditionally used the **calorie** (cal) or **kilocalorie** (kcal) but are switching to the SI unit, the **joule** (Table 6-1).[10a] Work done by the system may be **mechanical** (e.g., by changing the volume of the surroundings), **electrical** (e.g., by charging of a battery), or **chemical** (e.g., by effecting the synthesis of a polypeptide from amino acids).

2. Enthalpy Changes and Thermochemistry

We are most often interested in the *changes* in the thermodynamic functions when a chemical reaction takes place; for example, the heat absorbed by the system within a **bomb calorimeter** where the volume stays constant (Q_V) is a direct measure of the change in E:

$$Q_V = \Delta E \quad (6\text{-}2)$$

To measure ΔE for combustion of a biochemical compound, the substance may be placed in a bomb together with gaseous oxygen and the mixture ignited within the calorimeter by an electric spark. In this case, heat will be evolved from the bomb and will pass into the surroundings. Q_V and ΔE will be negative. The bomb calorimeter is designed to measure Q_V and thereby to give us a way of determining ΔE for reactions.

Processes at constant pressure. Chemical and biochemical reactions are much more likely to be conducted at constant pressure (usually 1 atm) than they are at constant volume. For this reason, chemists tend to use the **enthalpy *H*** more often than the internal energy E.

$$H = E + PV \quad (6\text{-}3)$$

It follows from Eq. 6-3 that *if the pressure is constant*, ΔH_P is equal to $\Delta E_P + P\,\Delta V$. Since in a process at constant pressure, $P\,\Delta V$ is exactly the pressure–volume work done on the surroundings, the heat absorbed at constant pressure (Q_P) is a measure of ΔH_P.

$$Q_P = \Delta E_P + P\,\Delta V = \Delta H_P \quad (6\text{-}4)$$

Since enthalpy changes can be obtained directly from measurement of heat absorption at constant pressure, even small values of ΔH for chemical and biochemical reactions can be measured using a microcalorimeter.[11,12] Using the technique of pulsed acoustic calorimetry, changes during biochemical processes can be followed on a timescale of fractions of a millisecond. An example is the laser-induced dissociation of a carbon monoxide–myoglobin complex.[13]

The term *enthalpy* was coined to distinguish H from E, but we sometimes tend to be careless about language and many discussions of energy in the literature are in fact about enthalpy. The difference is often not significant because if the pressure–volume work is negligible, E and H are the same.

Enthalpies of combustion and physiological fuel values. The **heat of combustion** ($-\Delta H_c$) of an organic substance is usually determined from ΔE_c, which is measured in a bomb calorimeter. Since ΔE_V and ΔE_P are nearly identical, it follows that $\Delta H_P = \Delta E_V + P\,\Delta V$. Here ΔV is the volume change which would have occurred if the reaction were carried out at constant pressure P; thus, ΔH_P can be estimated by calculation. Since ΔH is desired for combustion to carbon dioxide, water, elemental nitrogen (N_2), and sulfur, correction must be made for the amounts of the latter elements converted into oxides. By these procedures, it has been possible to obtain highly accurate values of ΔH_c both for biochemical compounds and for mixed foodstuffs. In nutrition, $-\Delta H_c$ is sometimes referred to as the **gross energy**. Values are usually expressed in kilocalories (kcal) by chemists but often as **Cal** (with a capital C) in the nutritional literature.

Caloric values of foods (physiological fuel values) are enthalpies of combustion but with an opposite sign, ($-\Delta H_c$), and corrected for energy lost in urine (e.g., as urea) and feces. While enthalpies of combustion of foods are all negative, the caloric values are given as positive numbers. Caloric values for proteins are

TABLE 6-1
Units of Energy and Work and the Values of Some Physical Constants

The joule, SI unit of energy 1 J = 1 kg m^2 s^{-2} = 1 N m (newton meter) = 1 W s (watt second) = 1 C V (coulomb volt)
Thermochemical calorie 1 cal = 4.184 J
Large calorie 1 Cal = 1 kcal = 4.184 kJ
Work required to raise 1 kg 1 m on earth (at sea level) = 9.807 J
Gibbs energy of hydrolysis of 1 mole of ATP at pH 7, millimolar concentrations = –12.48 kcal = –52.2 kJ
Work required to concentrate 1 mole of a substance 1000-fold, e.g., from 10^{-6} to 10^{-3} M = 4.09 kcal = 17.1 kJ
Avogadro's number, the number of particles in a mole $N = 6.0220 \times 10^{23}$
Faraday 1 F = 96,485 C mol^{-1} (coulombs per mole)
Coulomb 1 C = 1 A s (ampere second) = 6.241×10^{18} electronic charges
The Boltzmann constant $k_B = 1.3807 \times 10^{-23}$ J deg^{-1}
The gas constant, $R = N\, k_B$ $R = 8.3144$ J deg^{-1} mol^{-1} = 1.9872 cal deg^{-1} mol^{-1} = 0.08206 1 atm deg^{-1} mol^{-1} and at 25°C $RT = 2.479$ kJ mol^{-1}
The unit of temperature is the kelvin (K); 0°C = 273.16 K
ln x = 2.3026 log x
One atmosphere (atm) = 101.325 kilopascals (kPa)

calculated for the conversion of the nitrogen to urea, the major nitrogenous excretion product in mammals, rather than to elemental nitrogen. Typical values are shown in Table 6-2.

TABLE 6-2
Caloric Values of Food Components

Component	Caloric values per gram	
Carbohydrates	4.1 kcal	17 kJ
Pure glucose	3.75 kcal	15.7 kJ
Lipids	9.3 kcal	39 kJ
Proteins[a]	4.1 kcal	17 kJ

[a] Nitrogen excreted as urea.

From a thermochemical viewpoint, can a human or animal be regarded as just a catalyst for the combustion of foodstuffs? To answer this question, large calorimeters were constructed into which an animal or a human being was placed. If, while in the calorimeter, the subject neither gained nor lost weight, the heat evolved should have been just equal to $-\Delta H$ for combustion of the food consumed to CO_2, water, and urea. That this prediction was verifed experimentally does not seem surprising, but at the time that the experiments were first done in the early years of the century there may have been those who doubted that the first law of thermodynamics applied to mammals.

In practice, animal calorimetry is quite complicated because of the inherent difficulty of accurate heat measurements, uncertainties about the amount of food stored, and the necessity of corrections for ΔH_c of the waste products. However, the measurement of energy metabolism has been of considerable importance in nutrition and medicine. Indirect methods of calorimetery have been developed for use in measuring the **basal metabolic rate** of humans. For a good discussion see White *et al.*[14]

The basal metabolic rate is the rate of heat evolution in the resting, postabsorptive state, in which the subject has not eaten recently. In this condition, stored foods provide the energy and are oxidized at a relatively constant rate. The basal metabolic rate tends to be *proportional to the surface area*; which can be approximated (in units of m^{-2}) as 1/60 [height (cm) × mass (kg)].[15] For a young adult female, the basal metabolic rate is typically ~154 kJ h^{-1} m^{-2} and for a young adult male ~172 kJ h^{-1} m^{-2}. This is ~320 – 360 kJ h^{-1} for a 70-kg person. Note that 360 kJ h^{-1} is the same as the power output of a 100-watt lightbulb. While there is considerable variation among individuals, basal metabolic rates far below or above normal may indicate a pathological condition such as an insufficiency or oversupply of the thyroid hormone thyroxine. Metabolic rates fall somewhat below the basal value during sleep and are much higher than basal during hard exercise. A human may attain rates as high as 2500 kJ (600 kcal) per hour. At a basal rate of 320 kJ (76 kcal) per hour, a person requires 7680 kJ (1835 kcal) each 24 h to supply his or her basal needs, plus additional energy during periods of muscular exercise. Routine light exercise as in the office or during housework increases metabolism to about double the basal rate. Although the caloric values in Table 6-2 are reliable for prediction of metabolic energy needs, they must be adjusted to predict the efficiency of utilization for growth. In one study[16] a group of rats deposited 28% of the available energy from sucrose as body protein and fats, but fats were deposited with an efficiency of 36%.

3. The Second Law of Thermodynamics

Why does heat flow from a warm body into a cold one? Why doesn't it ever flow in the reverse direction? We can see that differences in temperature control the direction of flow of heat, but this observation raises still another question: What *is* temperature? Reflection on these questions, and on the interconversion of heat and work, led to the discovery of the second law of thermodynamics and to the definition of a new thermodynamic function, the **entropy *S***.

Consider the melting of ice. This is a phase transition that usually takes place at constant temperature and pressure. At a temperature just above 0°C ice melts completely, but at a temperature just below 0°C it does not melt at all. At 0°C we have an equilibrium. In the language of thermodynamics, the melting of ice at 0°C is a **reversible reaction**. What criterion could be used to predict this behavior for water? For many familiar phenomena, e.g., combustion, a spontaneous reaction is accompanied by the evolution of a large amount of heat, i.e., ΔH is negative. However, when ice melts it *absorbs* heat. The ΔH of fusion amounts to 6.008 kJ mol^{-1} at 0°C and is nearly the same just below 0°C, where the ice does not melt, and just above 0°C, where the ice melts completely. In the latter case, the melting of ice is a spontaneous reaction for which ΔH is positive. It is clear from such facts that *the sign of the enthalpy change does not serve as a criterion of spontaneity*.

A correct understanding of the ice–water transition came when it was recognized that when ice melts not only does H increase by 6.008 kJ mol^{-1}, as the molecules acquire additional internal energy of translation, vibration, and rotation, but also *the molecules become more disordered*. Although historically entropy was introduced in a different context, it is now recognized to be a measure of "microscopic disorder." When ice melts, the entropy S increases because the structure becomes less ordered.

The second law of thermodynamics is stated in many different ways, but the usual mathematical formulation asserts that for the universe (or for an isolated system)

$$\Delta S \text{ (system + surroundings)} = 0$$
$$\text{for reversible processes}$$

$$\Delta S > 0$$
$$\text{for real (nonreversible) processes} \qquad \text{(6-5)}$$

The second law is sometimes stated in another way: *The entropy of the universe always increases.*

The second law also defines both S and the thermodynamic temperature scale as follows:

$$dS_{\text{reversible}} = q/T \qquad \text{(6-6)}$$

Here q is an infinitesimal quantity of heat absorbed from the surroundings by the system and T is measured in kelvins (K). For a *reversible phase transition* such as the melting of ice at constant pressure and temperature, the change in entropy of the H_2O is just $\Delta H/T$.

$$\Delta S)_{P,T,\text{ reversible}} = Q/T = \Delta H/T \qquad \text{(6-7)}$$

Entropy is measured in units of joules per kelvin (or °C) or calories per K, the latter sometimes being abbreviated as e. u. (entropy units). Since the melting of ice at 0°C is a reversible process, the second law asserts that the entropy of the surroundings decreases by the same amount that the entropy of the water increases. The value of $T\,\Delta S$ is numerically equal to the heat of fusion, 6.008 kJ mol^{-1} in the case of water at 0°C. Thus, the entropy increase in the ice as it melts at 0°C is 6.008 x 10^3 J/273.16 K = 22.0 J K^{-1}.

The thermodynamic temperature. The definition of thermodynamic temperature in kelvins (Eq. 6-8) also follows from Eq. 6-6. See textbooks of thermodynamics for further treatment.

$$T = (\partial E/\partial S)_V = (\partial H/\partial S)_P \qquad \text{(6-8)}$$

The entropy of a substance can be given a precise mathematical formulation involving the degree of molecular disorder (Eq. 6-9).

$$S = k_{\text{B}} \ln \Omega \qquad \text{(6-9)}$$

Here k_{B} is the **Boltzmann constant** (see Table 6-1) and Ω is given precisely as the number of microscopic states (different arrangements of the particles) of the system corresponding to a given macroscopic state, i.e., to a given temperature, pressure, and quantity. It increases as volume or temperature is increased and in going from solid to liquid to gaseous states. Equation 6-9 is not part of classical thermodynamics (which deals only with macroscopic systems, i.e., with large collections of molecules). However, using the methods of statistical thermodynamics,[17] this equation can be used to predict the entropies of gases.

The *racemization of an amino acid* provides a biochemical example that can be related directly to Eq. 6-9. A solution of an L-amino acid will be efficiently changed into the racemic mixture of 50% D and 50% L by the action of an enzyme (a **racemase**) with no uptake or evolution of heat. Thus, $\Delta H = 0$ and the only change is an entropy change. Let us designate Ω for the pure isomer as Ω'. Since there are just two choices of configuration for each of the N molecules in 1 mole of the racemate we see that for the racemate

$$\Omega = 2^N \,\Omega' \qquad \text{(6-10)}$$

Applying Eq. 6-9 we calculate ΔS as follows:

$$\begin{aligned} \Delta S &= k_{\mathrm{B}}\,(\ln 2^N + \ln \Omega') - k_{\mathrm{B}} \ln \Omega' \\ &= Nk_{\mathrm{B}} \ln 2 = R \ln 2 = 5.76 \text{ J K}^{-1} \text{ mol}^{-1} \end{aligned} \tag{6-11}$$

Entropies from measurement of heat capacities.

It follows from Eq. 6-9 that $S = 0$ when $T = 0$ for a perfect crystalline substance in which no molecular disorder exists. The *third law of thermodynamics* asserts that as the thermodynamic temperature T approaches 0 K the entropy S also approaches zero for perfect crystalline substances. From this it follows that at any temperature above 0 K, the entropy is given by Eq. 6-12.

$$S = \int_0^T C_p \; d \; \ln \; T \tag{6-12}$$

In this equation C_p is the heat capacity at constant pressure:

$$C_P = (\partial H / \partial T)_P \tag{6-13}$$

If C_p is measured at a series of low temperatures down to near zero K, Eq. 6-12 can be used to evaluate the absolute entropy S. If phase transitions occur as the temperature is raised, entropy increments given by Eq. 6-7 must be added to the value of S given by Eq. 6-12. For a few compounds, such as water (Chapter 2), molecular disorder is present in the crystalline state even at 0 K. For these substances a term representing the entropy at 0 K must be added to Eq. 6-12.

TABLE 6-3
Entropies of Selected Substances[a]

Substance	State[b]	Entropy S	
		cal $\mathrm{K^{-1}\,mol^{-1}}$	J $\mathrm{K^{-1}\,mol^{-1}}$
C (diamond)	s	0.55	2.3
C (graphite)	s	1.36	5.7
Cu	s	8.0	33
Na	s	12.2	51
H_2O (ice)	s	9.8	41
H_2O	l	16.7	70
H_2O	g (1 atm)	45.1	189
He	g	30.1	126
H_2	g	31.2	131
N_2	g	45.8	192
CO_2	g	51.1	214
Benzene	g	64.3	269
Cyclohexane	g	71.3	298

[a] All values are given in entropy units (e.u.) of calories per Kelvin per mole and in joules per Kelvin per mole at 25°C (298.16 K).
[b] Here s stands for solid, l for liquid, and g for gaseous.

The entropies of a few substances are given in Table 6-3. Notice how the entropy increases with increasing complexity of structure, with transitions from solid to liquid to gas, and with decreasing hardness of solid substances.

Measurements of C_p versus temperature for solutions of macromolecules or for biological membranes (Fig. 8-9) over a narrower temperature range are also of interest. These can be obtained with a **differential scanning calorimeter**[18,19] or by an indirect procedure.[20] Denaturation of polymers or phase changes in membranes may be observed. Larger values of C_p are observed for open, denatured, or random-coil structures that are usually present at higher temperatures than for tightly folded molecules.

4. A Criterion of Spontaneity: The Gibbs Energy

We have seen that while many spontaneous processes, e.g., combustion of organic compounds, are accompanied by liberation of heat (negative ΔH), others are accompanied by absorption of heat from the surroundings (positive ΔH). An example of the latter is the melting of ice at a temperature just above 0°C, during which there is a large increase in the entropy of the water. As we have seen, at 0°C at equilibrium $T\,\Delta S$ is just equal to $-\Delta H$ (Eq. 6-7).

The recognition that $\Delta H - T\,\Delta S = 0$ for a system at equilibrium led J. W. Gibbs to realize that the proper thermodynamic function for determining the spontaneity of a reaction is what is now known as the **Gibbs energy** or Gibbs function **G** (Eq. 6-14).

$$G = H - TS \tag{6-14}$$

In the older literature the Gibbs energy was usually called the **free energy** or Gibbs free energy and was often given the symbol F. For a process at constant temperature and pressure the change in G is given by Eq. 6-15 in which all quantities refer to the system.

$$\Delta G_{T,P} = \Delta H - T\,\Delta S \tag{6-15}$$

For a reversible (equilibrium) process doing only pressure–volume work:

$$\Delta G)_{T,\text{reversible}} = \Delta H - T\,\Delta S = 0 \tag{6-16}$$

It can also be shown readily that ΔG *is negative for any spontaneous* (irreversible) *process*. Such a process is called **exergonic**. Likewise, if ΔG is positive, a given reaction will *not* proceed spontaneously and is called **endergonic**. The magnitude of the decrease in the Gibbs energy ($-\Delta G$) is a direct measure of the maximum work which could be obtained from a given chemical reaction if that reaction could be coupled in some fashion reversibly to a system able to do work. It represents the maximum amount of electrical work that could be extracted or the maximum amount of muscular work or osmotic work obtainable from a reaction in a biological system. In any real system, the amount of work obtainable is necessarily less than $-\Delta G$ because real processes are irreversible, i.e., entropy is created.

Returning to the older assumption that the magnitude of ΔH might be an index of work obtainable, we note that $T\,\Delta S$ amounts to only a few kilojoules for most reactions. Therefore, if ΔH is large, as in the combustion of foodstuffs, it is not greatly different from ΔG for the same process. Therefore, we can justify use of the caloric value of a food as an approximate measure of the work obtainable from its metabolism in the body.

5. Practical Thermochemistry

For thermodynamic data to be useful in chemical calculations, we must agree upon **standard states** for elements and compounds. If we wish to talk about the change in the Gibbs energy that occurs when one or more pure compounds are converted to other pure substances, we must agree upon a state (crystalline, liquid, gaseous, or in solution) and upon a pressure (especially when gases are involved) at which the data apply. The standard pressure is usually 1 atm. Standard states of the elements are the pure *crystalline*, *solid*, or *gaseous materials*, e.g., C (graphite), S (crystalline, rhombic), P (crystalline, white), and N_2, O_2, and H_2 (gaseous). It is also essential to specify the temperature. Thermodynamic data are most often given for 25°C, but there is a standard state for each substance at each temperature.

It is usually impractical to measure the values of G or H, but ΔG and ΔH for a chemical reaction can be evaluated. Changes in Gibbs energy can be calculated from tables of ΔG for formation of compounds from the elements (Eq. 6-17). These values of ΔG_f can be obtained experimentally by measuring ΔH of combustion for the compound of interest and for H_2, elemental carbon, and other elements present in the compound and also by obtaining entropies from heat capacity measurements. Many other tabulated ΔG_f values have been obtained indirectly utilizing data from equilibrium constants. The resulting **standard Gibbs energies of formation** are given the symbol $\Delta G_f°$. *The values of $\Delta G_f°$ for the elements in their standard states are all exactly zero.*

Summing changes in Gibbs energy. A convenient feature of thermodynamic calculations is that if two or more chemical equations are summed, ΔG for the resulting overall equation is just the sum of the ΔG's for the individual equations as illustrated in Eqs. 6-17 to 6-20. The same applies for ΔH and ΔS.

$$\mathrm{CH_3COOH\ (l) \longrightarrow 2\,C + 2\,H_2 + O_2}$$
$$-\Delta G_f° \text{ for acetic acid} = +396.4 \text{ kJ mol}^{-1} \quad (6\text{-}17)$$

$$\mathrm{2\,O_2 + 2\,C \longrightarrow 2\,CO_2}$$
$$2\,\Delta G_f° \text{ for } \mathrm{CO_2} = -788.8 \text{ kJ mol}^{-1} \quad (6\text{-}18)$$

$$\mathrm{O_2 + 2\,H_2 \longrightarrow 2\,H_2O}$$
$$2\,\Delta G_f° \text{ for } \mathrm{H_2O} = -474.4 \text{ kJ mol}^{-1} \quad (6\text{-}19)$$

$$\overline{\mathrm{CH_3COOH\ (l) + 2\,O_2 \longrightarrow 2\,CO_2 + 2\,H_2O\ (l)}}$$
$$\Delta G_c° \text{ for acetic acid} = -866.8 \text{ kJ mol}^{-1} \quad (6\text{-}20)$$

In this example an equation for the decomposition of acetic acid into its elements (Eq. 6-17) has been summed with Eqs. 6-18 and 6-19, which represent the formation of the proper number of molecules of CO_2 and H_2O from the elements. The sum of the three equations gives the equation for the combustion of acetic acid to CO_2 and water, and the sum of the ΔG values for the three equations gives ΔG for combustion of acetic acid. The resulting value of ΔG is for combustion of pure liquid acetic acid by oxygen at 1 atm to give CO_2 at 1 atm and pure liquid water, all reactants and products being in their standard states.

The process described in the preceding paragraph is represented by Eq. 6-21, which is a general equation for calculation of $\Delta G°$ for any reaction from $\Delta G_f°$ of products and reactants.

$$\Delta G° = \Sigma \Delta G_f° \text{(products)} - \Sigma \Delta G_f° \text{(reactants)} \quad (6\text{-}21)$$

How does the change in Gibbs energy vary if we go from the standard state of a compound to some other state? Consider a change of pressure in a gas. It is easy to show (see any thermodynamics text) that

$$(\partial G/\partial P)_T = V \quad (6\text{-}22)$$

Using Eq. 6-22 together with the perfect gas law, we obtain the relationship (Eq. 6-23) between the Gibbs energy $\overline{G}$ of one mole of a substance at pressure P and the standard Gibbs energy $\overline{G}°$ at pressure $P°$.

$$\overline{G} - \overline{G}° = RT \ln \frac{P}{P°} = RT\ \Delta \ln P \quad (6\text{-}23)$$

Here the bar over the symbol G indicates that the Gibbs energy is for one mole of substance. Since $P°$ is by definition 1 atm, the Gibbs energy change per mole upon changing the pressure from $P°$ to P is just $RT \ln P$.

Reactions in solution. It is customary in books on thermodynamics to develop most of the important thermodynamic equations as applied to a perfect gas, but we will move at this point to a consideration of biochemical substances in solution. Biochemists are usually interested in the behavior of substances dissolved in relatively dilute aqueous solutions but also in cytoplasm, in which some concentrations may be very high. Sometimes the interest may be in nonaqueous solutions. In any case, it is necessary to establish a standard state for the solute. The standard state of a substance in aqueous solution is customarily taken as a strictly hypothetical one **molal** solution (one mole of solute per kilogram of water) *whose properties are those of a solute at infinite dilution*. An equation exactly analogous to Eq. 6-23 can be written relating the Gibbs energy of one mole of dissolved solute $\overline{G}_i$ to the Gibbs energy $\overline{G}_i°$ in the hypothetical standard state of unit activity and to the **activity** a_i of the solute (Eq. 6-24).

$$\overline{G}_i = \overline{G}_i° + RT \ln a_i \tag{6-24}$$

Here the subscript i designates a particular component in a solution which also contains solvent and, perhaps, other components. To be precise, $\overline{G}_i$ is a *partial molar Gibbs energy*, i.e., the changes in total Gibbs energy of a very large volume of solution when one mole of the component is added.

From Eq. 6-24 it follows that the *Gibbs energy change for dilution* from one activity a_1 to another a_2 is:

$$\Delta\overline{G} \text{ (dilution from } a_1 \text{ to } a_2) = RT \ln (a_2/a_1) \tag{6-25}$$

Equation 6-24 and the equations that follow from it apply to molal activities. However, the concentration can be substituted for activity in very dilute solution where the behavior of the dissolved molecules approximates that of the hypothetical ideal solution for which the standard state is defined. For any real solution, the activity can be expressed as the product of an activity coefficient and the concentration (Eq. 6-26).

$$a = \gamma c \tag{6-26}$$

where a = activity, γ = activity coefficient, and c = molal concentration. Thus, to use the tabulations of thermodynamic functions for substances in solution to predict behavior in other than very dilute solution, we must multiply the concentration of every component by the appropriate activity coefficient. For the approximate calculations which are often of interest to biochemists, it is customary to equate concentration with activity. Furthermore, in dilute solutions, the more usual **molar** concentrations (moles per liter) are nearly equal to molal concentrations. The same equations can be used with *mole fractions* rather than molal concentrations.

ΔG° and the equilibrium constant. Consider the following generalized chemical equation (Eq. 6-27) for reaction of a moles of A with b moles of B to give products C and D, etc.

$$a\text{A} + b\text{B} + \cdots = c\text{C} + d\text{D} \cdots \tag{6-27}$$

The standard Gibbs energy change $\Delta G°$ for the process is given by Eq. 6-28:

$$\Delta G° = \text{c}\,\overline{G}°(\text{C}) + \text{d}\,\overline{G}°(\text{D}) + \cdots - \text{a}\,\overline{G}°(\text{A}) - \text{b}\,\overline{G}°(\text{B}) \cdots \tag{6-28}$$

The symbol $G°$ (A) designates the Gibbs energy of A, etc. The value of ΔG for any desired concentrations of reactants or products can be related to this $\Delta G°$ by applying to each component Eq. 6-24 with the following result:

$$\Delta G = \Delta G° + RT \ln \frac{a_C{}^c\, a_D{}^d \cdots}{a_A{}^a\, a_B{}^b \cdots} \tag{6-29}$$

Here a_C represents the activity of component C, etc. This useful equation permits us to calculate ΔG for the low concentrations usually found in biochemical systems. These are more often in the millimolar range or less rather than approaching the hypothetical 1 M of the standard state. Often concentrations are substituted in Eq. 6-29 for activities:

$$\Delta G \approx \Delta G° + RT \ln \frac{[\text{C}]^c[\text{D}]^d}{[\text{A}]^a[\text{B}]^b} \tag{6-30}$$

Equation 6-29 is used in another way by noting that $\Delta G = 0$ when a system is at equilibrium and that at equilibrium the product $a_C{}^c a_D{}^d \ldots / a_A{}^a a_B{}^b \ldots$ is just the equilibrium constant K. It follows that

$$\begin{aligned} \Delta G° &= -RT \ln K = -2.303RT \log K \\ &= -19.145T \log K \text{ J mol}^{-1} \\ &= -5.708 \log K \text{ kJ mol}^{-1} \text{ at } 25°\text{C} \\ &= -1.364 \log K \text{ kcal mol}^{-1} \text{ at } 25°\text{C} \end{aligned} \tag{6-31}$$

Although the units of $\Delta G°$ are kJ mol^{-1}, the Gibbs energy change in Eq. 6-31 is that for the reactions of a moles of

A, b moles of B, etc., as in the equation used to define K (Eq. 6-27 in this instance). It is also important to realize that the log term in Eq. 6-31 must be unitless. Although we usually write ln K or log K, K here represents K / Q, where $Q = 1$ because it has the same form as K but with all components in their standard states. Since the units of K and Q are the same, log K is unitless. Similar considerations apply to expressions of K in exponential form.

Activity coefficients and concentration equilibrium constants. Strictly speaking, Eq. 6-31 applies only to thermodynamic equilibrium constants – that is, to constants that employ activities rather than concentrations. The experimental determination of such constants requires measurements of the apparent equilibrium constant or **concentration equilibrium constant**[21] K_c at a series of different concentrations and extrapolation to infinite dilution (Eq. 6-32).

$$\begin{aligned} K_c &= \text{concentration equilibrium constant} \\ &= \frac{[\mathrm{C}]^c[\mathrm{D}]^d}{[\mathrm{A}]^a[\mathrm{B}]^b} \text{ at equilibrium} \end{aligned} \tag{6-32}$$

Extrapolation of K_c to infinite dilution to give K is usually easy because the activity coefficients of most ionic substances vary in a regular manner with **ionic strength** and follow the **Debye–Hückel** equation (Eq. 6-33) in very dilute solutions (ionic strength < 0.01).

$$\log \gamma = -0.509 z_1 z_2 \sqrt{\mu} \tag{6-33}$$

The integers z_1 and z_2 are the numbers of charges (valences) for the cation and anion of the salt. The ionic strength (μ, or I) is evaluated as follows:

$$\mu = \tfrac{1}{2} \Sigma_i c_i z_i^2 \tag{6-34}$$

Here c_i are the molar concentrations of the ions. The summation is carried out over all the ions present. The activity coefficient γ (Eq. 6-33) is the mean activity coefficient for both the cation and anion.

Equation 6-33 suggests that extrapolation of equilibrium constants to infinite dilution is done appropriately by plotting log K_c vs $\sqrt{\mu}$. For example, Fig. 6-1 shows plots of pK'_a for dissociation of $H_2PO_4^-$, AMP^-, and ADP^{2-}, and ATP^{3-} vs $\sqrt{\mu}$. The variation of pK'_a with $\sqrt{\mu}$ at low concentrations (Eq. 6-35) is derived by application of the Debye-Hückel equation (Eq. 6-33):

$$\mathrm{p}K_a' = \mathrm{p}K_a - 0.509\,(z_A^{\,2} - z_{HA}^{\,2})\,\sqrt{\mu} \tag{6-35}$$

Straight lines of slope – 0.509 ($z_A^{\,2} - z_{HA}^{\,2}$) are expected. The observed (negative) slopes (Fig. 6-1) are ~1.5 for $H_2PO_4^-$ and AMP^-, ~2.5 for ADP^{2-}, and ~3.5 for ATP^{3-}. The data over the entire range of ionic strength are fitted by empirical relationships of the type of Eq. 6-35a:

$$\mathrm{p}K_a' = \mathrm{p}K_a - a\sqrt{\mu} + b\mu \quad \text{for } \mu < 0.2 \tag{6-35a}$$

in which a and b are empirically determined constants. For example, for $H_2PO_4^-$ $a = 1.52$ and $b = 1.96$. The value of pK_a found was 7.18, about 0.22 greater than the value at $\mu = 0.2$, an ionic strength more commonly used in the laboratory and close to that found in tissues. Note that the difference between the extrapolated pK_a for ATP^{3-} of 7.68 and the observed value of ~ 7.04 at μ = 0.2 is even greater. Serious errors can be introduced into calculations by using extrapolated values for K for solutions of appreciable ionic strength. The errors will be maximal for ions of high charge type such as ATP^{3-} and ATP^{4-}.

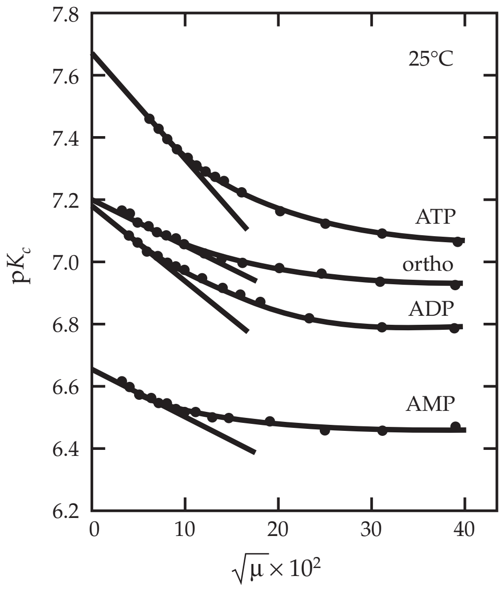


Figure 6-1 The apparent pK_a values for secondary ionizations of AMP, ADP, ATP, and H_3PO_4 (abbreviated ortho) plotted against $\sqrt{\mu}$. Temperature: 25°C. From R. C. Phillips *et al.*[22]

Another problem with equilibrium constants for reactions that use or produce hydrogen ions is that there is no rigorous relationship between pH and a_{H^+} or [H⁺]. Indeed, the concept of an activity of a single ion has little meaning in thermodynamics. Nevertheless, in the pH range of interest to biochemists, results that are very close to those obtained by more rigorous methods are achieved by assuming that the pH meter responds to hydrogen ion activity. The almost universal practice of biochemists is to assume the pH meter reading obtained with a glass electrode equal to – log a_{H^+} and to substitute the value of a_{H^+} so obtained for [H^+] in defining the concentration equilibrium constant, K_c. It is also customary in most branches of chemistry to use values of equilibrium constants and of Gibbs energy which have *not* been extrapolated to $\mu = 0$. Thus most values of K and ΔG, including many of those in this book, are actually of K_c and ΔG_c. An international Commission on Biothermodynamics[21] recommended that values of K_c be measured with the lowest effective buffer concentration and that the ionic strength be brought to 0.1 with KCl.

Changes in equilibria with temperature. At constant pressure ΔG varies with absolute temperature as follows:

$$\frac{d(\Delta G/T)}{dT} = -\Delta H/T^2 \qquad (6\text{-}36)$$

The corresponding variation in K is described by the **van't Hoff equation**:

$$\frac{d \ln K}{dT} = \frac{\Delta H°}{RT^2}$$

or

$$\frac{d \ln K}{d(1/T)} = \frac{-\Delta H°}{R}$$

or

$$\Delta H°(\text{kJ mol}^{-1}) = \frac{-0.01914\ d \log_{10} K}{d(1/T)} \qquad (6\text{-}37)$$

If $\Delta H°$ can be assumed constant over the temperature range of an experiment, a plot of ln K vs $1/T$ provides a convenient estimate of $\Delta H°$ (or $\Delta H'$ if ln K' is plotted). The slope of the line will be $-\Delta H°/R$. Since $\Delta G°$ can be calculated from K, the method also permits evaluation of $\Delta S°$ using Eq. 6-15. However, unless great care is taken the method is of low accuracy[23] and it is preferable to establish ΔH by direct calorimetry. Also, especially for proteins, the assumption that $\Delta H°$ is constant over a significant temperature range may be erroneous.

From observations at only two temperatures, T_1, and T_2, Eq. 6-37 becomes

$$\Delta H° = R \ln (K_2/K_1)[T_1T_2/(T_2 - T_1)] \text{ kJ mol}^{-1} \qquad (6\text{-}37a)$$

6. Thermodynamics and Life Processes

Can thermodynamics be applied to living organisms? Classical thermodynamics deals with equilibria, but living beings are never in equilibrium. The laws of thermodynamics are usually described as statistical laws. How can such laws apply to living things, some of which contain all of their genetic information in a single molecule of DNA?[24,25] The ideal, reversible reactions of classical thermodynamics occur at infinitesimal speeds. How can thermodynamics be applied to the very rapid chemical reactions that take place in organisms? One answer is that thermodynamics can be used to decide *whether or not a reaction is possible* under given conditions. Thus, if we know the steady-state concentrations of reactants and products within a cell, we can state whether a reaction will or will not tend to go in a given direction.

We may still ask whether there are generalities comparable to the laws of thermodynamics that apply to the kind of steady state or "dynamic equilibrium" that exists in organisms. Lars Onsager showed that such relationships can be found for conditions that are *near equilibrium*. Ilya Prigogine and associates extended Onsager's findings and showed that under conditions that are *far from equilibrium* the system tends to become unstable and to spontaneously develop new structures, which Prigogine calls **dissipative structures**.[26–28] Vortices in flowing water, tornados and hurricanes are examples of dissipative structures. The maintenance of dissipative structures depends upon a flow of energy and matter through the system. The flow of energy is sufficient to "organize" a system. Thus, if a flask of water is placed on a hot plate, a cycle is established. The water moves via cyclic convection currents that develop as a result of the flow of energy through the system. Morowitz reckoned that the 6 ± 3 x 10^{18} kJ/year of solar energy that falls on the earth supplies the organizing principle for life.[28] Just as it drives the great cycles within the atmosphere and within the seas, it gives rise to the branching and interconnecting cycles of metabolism. This idea may even make the spontaneous development of the organized systems that we call life from inanimate precursors through evolution seem a little more understandable.

A characteristic of this nonequilibrium or irreversible thermodynamics is that time is explicitly introduced. Furthermore, open systems, in which materials and energy flow into and out of the system, are considered. Clearly, a living organism is an open system not a closed one of classical thermodynamics. Because of the flow of materials concentration gradients are set up and transport phenomena often become of primary importance. Articles and books that provide an introduction to nonequilibrium thermodynamics and to the literature in the field include the following.[10,26,28–34] Whether these methods can be applied in a practical way to metabolic systems has been debated.[35,36]

TABLE 6-4

Gibbs Energies of Formation and of Oxidation at 25°C for Compounds of Biochemical Interest[a,b]

				For oxidation by NAD^+		
Compound	Formula	$\Delta G_f°$ (kJ mol^{-1})	$\Delta G_c°$ (kJ mol^{-1})	$\Delta G°_{ox}$ (kJ mol^{-1})	$\Delta G°_{ox}$ (pH 7) (kJ mol^{-1})	Number of electrons
Acetaldehyde	C_2H_4O	–139.7	–1123.5	171.5	–28.3	10
Acetic acid	$C_2H_4O_2$	–369.4	–866.8	169.2	9.3	8
Acetate$^-$		–369.2	–894.0	142.0	22.1	8
Acetyl-CoA		–374.1*	–889.1*	146.9*	–13.0*	8
Acetyl-*P*		–1218.4	–901.7	134.3	–25.6	8
Acetylene[c]	C_2H_2	209.2	–1235.2	59.8	–140.0	10
Acetoacetate$^-$	$C_4H_5O_3^-$	–493.7	–1795.4	276.5	–3.2	16
Acetone	C_3H_6O	–161.2	–1733.6	338.4	18.7	16
cis-Aconitate^{3-}	$C_6H_3O_6^{3-}$	–920.9	–2157.0	173.9	–65.8	18
L-Alanine	$C_3H_7O_2N$	–371.3	–1642.0	300.4	0.8	15
L-Asparagine	$C_4H_8O_3N_2$	–526.6	–1999.7	331.2	–28.4	18
L-Aspartate$^-$	$C_4H_6O_4N^-$	–700.7	–1707.0	235.4	–24.3	15
n-Butanol	$C_4H_{10}O$	–171.8	–2591.7	516.2	36.7	24
n-Butyric acid	$C_4H_8O_2$	–380.2	–2146.1	443.8	44.2	20
n-Butyrate$^-$	$C_4H_7O_2^-$	–352.6	–2173.7	416.2	56.6	20
Butyryl-CoA		–357.5*	–2168.8*	421.1*	21.5*	20
Caproate$^-$	$C_6H_{11}O_2^-$	–329.7	–3459.7	684.1	84.7	32
CO_2 (g)		–394.4	0.0	0.0	0.0	0
CO_2 (aq)		–386.2	–8.2	–8.2	–8.2	0
HCO_3^-		–587.1	–44.5	–44.5	–4.6	0
CO (g)		–137.3	–257.1	1.9	–38.1	2
Citrate^{3-}	$C_6H_5O_7^{3-}$	–1166.6	–2148.4	182.5	–57.3	18
Creatine	$C_4H_9O_2N_3$	–264.3	–2380.6	338.8	–80.8	21
Creatinine	$C_4H_7ON_3$	–28.9	–2378.8	340.6	–79.0	21
Crotonate$^-$	$C_4H_5O_2^-$	–275.7	–2013.4	317.5	–2.1	18
Cysteine	$C_3H_7O_2NS$	–339.8	–2178.3	541.1	121.5	21
Cystine	$C_6H_{12}O_4N_2S_2$	–665.3	–4133.8	1046.0	246.9	40
Dihydroxyacetone-*P*[d]		–1293.2	–1458.4	95.5	–144.2	12
Erythrose 4-*P*[d]		–1439.1	–1944.1	127.8	–191.9	16
Ethanol	C_2H_6O	–181.5	–1318.8	235.1	–4.6	12
Ethylene (g)[c]	C_2H_4O	68.1	–1331.3	222.7	–17.1	12
Formaldehyde	CH_2O	–130.5	–501.0	16.9	–63.0	4
Formic acid	CH_2O_2	–356.1	–275.5	–16.5	–56.5	2
Formate$^-$	CHO_2^-	–350.6	–281.0	–22.0	–22.0	2
Fructose	$C_6H_{12}O_6$	–915.4	–2874.1	233.8	–245.7	24
Fructose 6-*P*[d]		–1758.3	–2888.1	219.8	–259.7	24
Fructose di-*P*[d]		–2600.8	–2902.5	205.4	–274.1	24
Fumaric acid	$C_4H_4O_4$	–647.1	–1404.8	149.2	–90.6	12
Fumarate$^-$	$C_4H_3O_4^-$	–604.2	–1447.7	106.2	–93.6	12
α-D-Galactose	$C_6H_{12}O_6$	–923.5	–2865.9	242.0	–237.5	24
α-D-Glucose	$C_6H_{12}O_6$	–917.2	–2872.2	235.6	–243.8	24
Glucose 6-*P*		–1760.3	–2886.0	221.8	–257.6	24
L-Glutamate$^-$	$C_5H_8O_4N^-$	–696.8	–2342.5	376.9	–2.7	21
L-Glutamine	$C_5H_{10}O_3N_2$	–524.8	–2633.1	474.8	–4.7	24
3-*P* glycerate$^-$[d]		–1515.7	–1235.9	59.0	–100.8	10
2-*P* glycerate$^-$[d]		–1509.9	–1241.8	53.2	–106.6	10
Glyceraldehyde 3-*P*[d]		–1285.6	–1466.0	87.9	–151.8	12
Glycerol	$C_3H_8O_3$	–488.5	–1643.4	169.5	–110.2	14
Glycerol-*P*		–1336.2	–1652.6	160.3	–119.3	14
Glycine	$C_2H_5O_2N$	–373.5	–1008.3	157.2	–22.6	9
Glycogen	$C_6H_{10}O_5$	–665.3	–2887.0	220.9	–258.6	24
Glycolate$^-$	$C_2H_3O_3^-$	–523.4	–739.7	37.2	–42.7	6
Glyoxylate$^-$	$C_2HO_3^-$	–461.1	–564.9	–46.9	–86.9	4
H_2O (l)		–237.2	0.0	0.0	0.0	0
OH^-		–157.3	–79.9	–79.9	–39.9	0
H^+		0.0	0.0	0.0		0
H_2 (g)		0.0	–237.2	21.8	–18.2	2
H_2O_2		–136.8	–100.4	–359.4	–319.4	–2
H_2S		–27.4	–714.6	321.3	161.5	8
HS^-		12.6	–754.5	281.4	161.5	8
β-Hydroxybutyric acid	$C_4H_8O_3$	–531.4	–1994.9	336.0	–23.6	18
β-Hydroxybutyrate$^-$	$C_4H_7O_3^-$	–506.3	–2020.0	310.9	–8.8	18
Hydroxypyruvate	$C_3H_4O_4$	–615.9	–1041.6	–5.7	–165.5	8

TABLE 6-4 (continued)

				For oxidation by NAD$^+$		
Compound	Formula	$\Delta G_f°$ (kJ mol^{-1})	$\Delta G_c°$ (kJ mol^{-1})	$\Delta G°_{ox}$ (kJ mol^{-1})	$\Delta G°_{ox}$ (pH 7) (kJ mol^{-1})	Number of electrons
Hypoxanthine	C_5H_6O	89.5	–2773.0	334.8	–144.6	24
Isocitrate^{3-}	$C_6H_5O_7^{3-}$	–1160.0	–2155.1	175.8	–63.9	18
α-Ketoglutarate^{2-}	$C_5H_4O_5^{2-}$	–798.0	–1885.5	186.4	–53.3	16
Lactate$^-$	$C_3H_5O_3^-$	–516.6	–1378.1	175.9	–23.9	12
α-Lactose	$C_{12}H_{22}O_{11}$	–1515.2	–5826.5	389.3	–569.7	48
L-Leucine	$C_6H_{13}O_2N$	–356.3	–3551.7	721.6	62.3	33
Mannitol	$C_6H_{14}O_6$	–942.6	–3084.0	282.8	–236.6	26
Malate^{2-}	$C_4H_4O_5^{2-}$	–845.1	–1444.0	109.9	–49.9	12
Methane (g)	CH_4	–50.8	–818.0	218.0	58.2	8
Methanol	CH_4O	–175.2	–693.5	83.4	–36.4	6
NH_4^+		–79.5	–276.3	112.2	12.3	3
NO_2^-		–34.5	–84.1	–472.6	–372.7	–3
NO (g)		86.7	–86.7	–345.7	–305.7	–2
NO_3^-		–110.5	–8.1	–655.6	–515.7	–5
Oxalate^{2-}	$C_2O_4^{2-}$	–674.9	–351.1	–92.1	–52.1	2
Oxaloacetate^{2-}	$C_4H_2O_5^{2-}$	–797.2	–1254.7	40.2	–79.7	10
H_3PO_4 (aq)[e]		–1147.3	0.0	0.0	0.0	0
$H_2PO_4^-$ (aq)[e]		–1135.1	–12.1	–12.1	27.8	0
HPO_4^{2-} (aq)[e]		–1094.1	–53.1	–53.1	26.8	0
n-Propanol	C_3H_8O	–175.8	–1956.1	374.8	15.2	18
Isopropanol	C_3H_8O	–185.9	–1946.0	384.9	25.3	18
Propionate$^-$	$C_3H_5O_2^-$	–360.0	–1534.7	278.2	38.5	14
Pyruvate$^-$	$C_3H_3O_3^-$	–474.5	–1183.1	111.9	–47.9	10
Phosphoenolpyruvate^{3-}		–1269.5	–1245.0	50.0	–109.8	10
Ribose 5-*P*[d]		–1599.9	–2414.9	175.0	–224.6	20
Ribulose 5-*P*[d]		–1597.6	–2417.1	172.8	–226.8	20
Sedoheptulose 7-*P*[d]		–1913.3	–3364.6	261.2	–298.2	28
Sedoheptulose di-*P*[d]		–2755.8	–3379.0	246.9	–312.5	28
Sorbitol	$C_6H_{14}O_6$	–942.7	–3083.9	282.9	–236.5	26
Succinate^{2-}	$C_4H_4O_4^{2-}$	–690.2	–1598.9	214.1	14.3	14
Succinyl-CoA		–686.7*	–1602.4*	210.6*	–29.2*	14
Sucrose	$C_{12}H_{22}O_{11}$	–1551.8	–5789.9	425.9	–533.1	48
SO_4^{2-}		–742.0	0.0	0.0	79.9	0
SO_3^{2-}		–497.1	–244.9	14.1	54.0	2
$S_2O_3^{2-}$		–513.4	–733.4	302.5	222.6	8
L-Threonine	$C_4H_9O_3N$	–514.6	–2130.3	330.1	–49.5	19
L-Tyrosine	$C_9H_{11}O_3N$	–387.2	–4466.8	842.5	23.4	41
Urea	CH_4ON_2	–203.8	–664.9	112.0	–7.8	6
Uric acid	$C_5H_4O_3N_4$	–356.9	–2089.4	241.5	–118.1	18
L-Valine	$C_5H_{11}O_2N$	–360.0	–2916.5	579.9	40.5	27
Xanthine	$C_5H_5O_2N_4$	–139.3	–2425.6	293.8	–125.7	21
D-Xylulose	$C_5H_{10}O_5$	–748.1	–2409.8	180.1	–219.5	20

[a] The quantities tabulated are $\Delta G_f°$, the standard free energy of formation from the elements; $\Delta G_c°$, the standard free energy of combustion; $\Delta G°_{ox}$, the standard free energy of oxidation by NAD$^+$ to products NADH + H$^+$, CO_2, H_2O, N_2, HPO_4^{2-}, and SO_4^{2-}; $\Delta G'_{ox}$ (pH 7), the apparent standard free energy change at pH 7. All values are in kJ mol^{-1} at 25°C in aqueous solution unless indicated otherwise. If a compound is designated (g) the values are for the gaseous phase at 1 atm pressure. The number of electrons involved in complete oxidation to CO_2, H_2O, N_2, and H_2SO_4 is given in the final column. If this number is negative, the compound must be reduced to obtain the products, e.g., $2\ NO_3^- + 10\ e^- + 12\ H^+ \rightarrow N_2 + 6\ H_20$. The data for phosphate esters refer to the compounds with completely dissociated phosphate groups ($-O-PO_3^{2-}$). The values of $\Delta G_f°$ for many of these compounds were calculated as $\Delta G_f°$ (nonphosphorylated compound) – $\Delta G°$ for hydrolysis (to HPO_4^{2-}, Table 6-6) $-\Delta G_f°$ for H_2O (one molecule for each phosphate ester formed) + $\Delta G_f°$ for HPO_4^{2-} (from this table). Data from Bassman and Krause[d] were used directly. For acyl-CoA derivatives CoA (–SH) is treated as an "element," i.e., the values of $\Delta G_f°$ given and designated with an asterisk (*) are for formation from the elements plus free CoA. The values of $\Delta G_c°$ and ΔG_{ox} are for oxidation to the usual products plus CoA. Values of $\Delta G°$ of hydrolysis (Table 6-6) were used in computing $\Delta G_f°$ for each of these compounds from that of the corresponding alcohol or carboxylate anion. Another source containing an extensive table of Gibbs energy values is Wilhoit, R. C. (1969) in *Biochemical Microcalorimetry* (Brown, H. D. ed.), pp. 305–317. Academic Press, New York

[b] The major source is Long, C., ed. (1961) *Biochemists Handbook*, pp. 90–92. Van Nostrand, Reinhold, Princeton, New Jersey. Most of the values in this collection are from Burton, K. (1957) *Ergeb. Physiol., Biol. Chem. Exp. Pharmakol.* **49**, 275–298

[c] From Stull, D.R., Westrum, E. F., Jr., and Sinke, G. C. (1969) *The Chemical Thermodynamics of Organic Compounds*. Wiley, New York

[d] Bassham, J. A. and Krause, G. H. (1969) *Biochim. Biophys. Acta.* **189**, 207–221

[e] Van Wazer, J. R. (1958) *Phosphorus and Its Compounds*, Vol. I, p. 889. Wiley (Interscience), New York

B. Tables of $\Delta G°$ Values for Biochemical Compounds

1. Gibbs Energies of Formation

Table 6-4 gives, in the first column, standard values of Gibbs energies of formation from the elements $\Delta G_f°$ for a variety of pure solids, gases, and liquids as well as values for substances in solution at the hypothetical 1 M activity. As an example, consider the value of $\Delta G_f°$ for pure liquid acetic acid, –389.1 kJ mol^{-1}. The equation for its formation from the elements is:

$$2\,C\,(s) + 2\,H_2\,(g, 1\,atm) + O_2\,(g, 1\,atm) \longrightarrow C_2H_4O_2\,(l)$$
$$\Delta G_f° = -389.1\ \text{kJ mol}^{-1} \qquad (6\text{-}38)$$

To obtain the Gibbs energy of formation in aqueous solution, we must have solubility data as well as activity coefficients of acetic acid at various concentrations. From these data the change in Gibbs energy for solution of the liquid acetic acid in water to give aqueous acetic acid in the hypothetical 1 molal standard state (Eq. 6-39) can be obtained.

$$\text{Acetic acid (l)} \longrightarrow \text{acetic acid (aq)}$$
$$\Delta G = -7.3\ \text{kJ mol}^{-1} \qquad (6\text{-}39)$$

Summing Eqs. 6-38 and 6-39 we obtain:

$$2\,C + 2\,H_2 + O_2 \longrightarrow \text{acetic acid (aq)}$$
$$\Delta G_f° = -396.4\ \text{kJ mol}^{-1} \qquad (6\text{-}40)$$

In many computations it is convenient to have ΔG values for single ions, e.g., for acetate$^-$. We can obtain $\Delta G_f°$ of acetate$^-$ (aq) from that of acetic acid (aq) by making use of $\Delta G°$ of dissociation (Eq. 6-41).

$$\text{Acetic acid} \longrightarrow H^+ + \text{acetate}^-$$
$$\Delta G° = -5.708 \log K_a = 5.708\ pK_a$$
$$= +27.2\ \text{kJ mol}^{-1}\ \text{at}\ 25°C \qquad (6\text{-}41)$$

By convention *we define the Gibbs energy of formation of H⁺ as zero*. Then, by summing Eqs. 6-40 and 6-41 we obtain $\Delta G_f°$ of acetate$^-$ = –369.2 kJ mol^{-1}.

2. Gibbs Energies of Dissociation of Protons

Table 6-5 gives thermodynamic dissociation constants and values of $\Delta G°$ and $\Delta H°$ for a number of acids of interest in biochemistry. Some of these values were used in obtaining the values of $\Delta G_f°$ for the ions of Table 6-4. The data of Table 6-5 can also be used in evaluation of Gibbs energy changes for reactions of ionic forms not given in Table 6-4.

3. Group Transfer Potentials

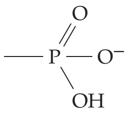
Recall that the equilibria for reactions by which monomers are linked to form biopolymers (whether amides, esters, phosphodiesters, or glycosides) usually favor *hydrolysis* rather than formation (condensation). The equilibrium positions depend on the exact structures. Some linkages are formed easily if monomer concentrations are high enough, but others are never formed in significant concentrations. Likewise, hydrolysis may be partial at equilibrium or it may be 99.9% or more complete.

Let us compare the hydrolysis of the two organic phosphates **adenosine triphosphate** (ATP) and **glucose 6-phosphate** (Eqs. 6-42 and 6-43).

$$HATP^{3-} + H_2O \longrightarrow HADP^{2-} + H_2PO_4^-$$
$$\Delta G° \approx -32.9\ \text{kJ mol}^{-1}$$
$$\Delta H° \approx -22.6\ \text{kJ mol}^{-1}\ \text{at}\ 25°C,\ \mu = 0.25 \qquad (6\text{-}42)$$

$$\alpha\text{-D-Glucose 6-phosphate}^- + H_2O \longrightarrow \text{glucose} + H_2PO_4^-$$
$$\Delta G° = -16.4\ \text{kJ mol}^{-1}\ \text{at}\ 25°C \qquad (6\text{-}43)$$

The decrease in Gibbs energy upon hydrolysis is twice as large for ATP as it is for glucose 6-phosphate. Glucose phosphate is thermodynamically more stable than ATP. It would be easier to form than would ATP by a reversal of the hydrolysis reaction and also easier to form biosynthetically. From the Gibbs energies of hydrolysis it follows that a **phospho group** could be transferred spontaneously from ATP to glucose in the presence of a suitable catalyst but not vice versa.



Phospho (phosphoryl) group

Because it reflects quantitatively the thermodynamic tendency for a group to be transferred to another

TABLE 6-5
Values of pK_a, $\Delta G°$, and $\Delta H°$ for Ionization of Acids at 25°C[a,b]

Acid	pK_a	$\Delta G°$ (kJ mol^{-1})	$\Delta H°$ (kJ mol^{-1})
Formic acid	3.75	21.4	0.04
Acetic acid	4.76	27.2	–0.1
Propionic acid	4.87	27.8	–0.6
Lactic acid	3.97 (35°C)	23.4	2.2
Pyruvic acid	2.49	14.2	12.1
NH_4^+	9.25	52.8	52.2
$CH_3NH_4^+$	10.59	60.4	55.4
Alanine			
–COOH	2.35	13.4	3.1
$–NH_3^+$	9.83	56.1	45.4
β-Alanine			
–COOH	3.55	20.3	4.5
$–NH_3^+$ (apparent)	10.19	58.2	
L-Alanyl-L-alanine	3.34	19.1	–0.5
Aspartic acid			
–COOH	2.05	11.7	7.7
–COOH	3.87	22.1	4.0
$–NH_3^+$	10.60	60.5	38.8
H_2CO_3, pK_1	6.35[c]	36.2	9.4
pK_2	10.33	59.0	15.1
H_3PO_4, pK_1	2.12	12.1	–7.9
pK_2	7.18[d]	41.0	3.8
(apparent)	6.78[e]	38.7	3.3
pK_3	12.40	70.8	17.6
Glycerol 1-phosphate, pK_2	6.66	38.0	–3.1
Glucose 6-phosphate	6.50	37.1	–1.8
Pyrophosphoric acid, $H_4P_2O_7$			
pK_3	6.7	38.1	–1.3
(apparent)	6.12[e]	34.9	0.5
pK_4	9.4	53.6	–7.1
(apparent)	8.95[e]	51.2	1.7
Adenosine	3.5	20.1	13.0
AMP			
pK_1 (ring, apparent)	3.74[e]	21.3	4.2
pK_2 (phosphate)	6.67[d]	38.1	3.6
(apparent)	6.45[e]	36.8	3.6
ADP			
pK_2 (ring, apparent)	3.93[e]	22.4	4.2
pK_3 (diphosphate)	7.20[d]	41.1	–5.7
(apparent)	6.83[f]	39.0	–5.7
ATP			
pK_3 (ring, apparent)	4.06	23.2	0
pK_4 (triphosphate)	7.68[d]	43.8	–7.0
(apparent)	7.06[d]	40.2	–7.0
Pyridine	5.17	29.5	20.1
Phenol	9.98	56.9	23.6

[a] These are thermodynamic values (infinite dilution) except for those labeled apparent. The latter apply at an ionic strength of 0.2–0.25.

[b] Most data are from Jencks, W. P. and Regenstein, J. (1976) in *Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol. I (Fasman, G. D. ed.), pp. 305–351, CRC Press, Cleveland, Ohio.

[c] Here, pK_1 is for $K_1 = [H^+][HCO_3^-] / [CO_2] + [H_2CO_3]$. From Forster, R. E., Edsall, J. T., Otis, A. B., and Roughton, F. J. W., eds. (1969) *NASA Spec. Publ.* **188**.

[d] From Phillips, R. C., George, P., and Rutman, R. J. (1963) *Biochemistry* **2**, 501–508.

[e] From Alberty, R. A. (1969) *J. Biol. Chem.* **244**, 3290–3302.

[f] Values used by Alberty, R. A. (1972) *Horizons of Bioenergetics*, pp. 135–147, Academic Press, New York, calculated for 0.2 ionic strength from equations of Phillips, R. C., George, P., and Rutman, R. J. (1966) *J. Am. Chem. Soc.* **88**, 2631–2640.

nucleophile (see Chapter 12), *the Gibbs energy decrease (–ΔG°) upon hydrolysis* is sometimes called the **group transfer potential**. During the hydrolysis of ATP (Eq. 6-42) the phospho group of ATP is transferred to a hydroxyl ion from water with a value $\Delta G° = -32.9$ kJ mol^{-1}. The group transfer potential of this phospho group is 32.9 kJ/mol and that of the phospho group of glucose 6-phosphate is only 27.6 kJ mol^{-1}. While the choice of water as the reference nucleophile for expression of the group transfer potential is somewhat arbitrary, it is customary. Transfer of groups is important in energy metabolism and in biosynthesis of polymers. Gibbs energies of hydrolysis are given in Table 6-6 for several compounds.

4. "Constants" That Vary with pH and Magnesium Ion Concentrations

Equation 6-42 is written for hydrolysis of $HATP^{3-}$ to $HADP^{2-} + H_2PO_4^-$, a stoichiometry that applies well in the pH range around 6. However, at a pH above ~7 most of the ATP is in the form ATP^{4-} and is cleaved to HPO_4^{2-} according to Eq. 6-44.

$$\begin{gathered} ATP^{4-} + H_2O \rightarrow ADP^{3-} + HPO_4^{2-} + H^+ \\ \Delta G° \approx +5.4 \text{ kJ mol}^{-1} \text{ at } 25°\text{C}, \mu = 0.25 \\ \Delta H° \approx -19.7 \text{ kJ mol}^{-1} \end{gathered} \tag{6-44}$$

The value of $\Delta G° = +5.4$ kJ mol^{-1} for this reaction is hardly the large negative number expected for a highly spontaneous reaction. What is the matter? The problem is that H^+ is produced and that the standard state of H^+ is 1 M, not 10^{-7} M. Because of this, biochemists often prefer to use another kind of **apparent dissociation constant** and an **apparent ΔG** such that the standard state of H^+ is taken as that of the pH at which the experiments were done, usually pH 7. The symbol K' has often be used and is used in this book to represent the following pH-dependent equilibrium constant (Eq. 6-45) which will be a constant only at a single pH.

$$K' = \frac{[\text{ADP}^{3-}][\text{HPO}_4^{\ 2-}]}{[\text{ATP}^{4-}]} \qquad (6\text{-}45)$$

If one proton is produced in the reaction as in Eq. 6-44, the following relationship will hold.

$$\Delta G' = \Delta G^\circ - 5.708 \text{ x pH kJ mol}^{-1} \text{ at } 25°\text{C} \qquad (6\text{-}46)$$

Note that $\Delta G' = -RT \ln K'$ and that $[H^+]$ does not appear in the expression for K' given by Eq. 6-45. From the value $\Delta G^\circ = +5.4$ kJ mol^{-1} and applying Eq.

TABLE 6-6
Gibbs Energies of Hydrolysis at 25°C (in kJ mol^{-1})[a]

Compound	Products	ΔG°	ΔG′(pH 7)	ΔH°
ATP^{4-}	$ADP^{3-} + HPO_4^{\ 2-} + H^+$	5.41[b]	–34.54	–19.71
ATP^{4-}	$AMP^{2-} + HP_2O_7^{\ 3-} + H^+$	2.54[c]	–37.4	–19.0
$MgATP^{2-}$	$MgADP^- + HPO_4^{\ 2-} + H^+$	16.0[d]	–24.0	–14.2
ADP^{3-}	$AMP^{2-} + HPO_4^{\ 2-} + H^+$	3.67[c]	–36.3	–13.5
AMP^{2-}	Adenosine + $HPO_4^{\ 2-}$	–9.6[e]	–9.6	0
ATP^{4-}	Adenosine + $HP_3O_{10}^{\ 4-}$	–36.0[e]	–36.0	–7.9
$HP_2O_7^{\ 3-}$	2 $HPO_4^{\ 2-} + H^+$	6.54[c]	–33.4	–12.6
Acetyl phosphate^{2-}	Acetate$^-$ + $HPO_4^{\ 2-} + H^+$	–7.7[f]	–47.7	
1,3-Diphosphoglycerate^{4-}	3-Phosphoglycerate^{3-} + $HPO_4^{\ 2-} + H^+$	–14.5[g]	–54.5	
Phosphoenolpyruvate^{3-}	Pyruvate$^-$ + $HPO_4^{\ 2-}$	–61.9	–61.9	–25.1
Carbamoyl phosphate^{2-}	$H_2N-\overset{\overset{\displaystyle O}{\Vert}}{C}-O^- + HPO_4^{\ 2-} + H^+$	–11.5	–51.5	
Creatine phosphate$^-$	Creatine$^+$ + $HPO_4^{\ 2-}$	–43.1	–43.1	
Phosphoarginine$^-$	Arginine + $HPO_4^{\ 2-}$		–38.1[h]	(Mg^{2+} present)
Glycerol phosphate^{2-}	Glycerol + $HPO_4^{\ 2-}$	–9.2	–9.2	
α-D-Glucose 6-phosphate^{2-}	α-D-Glucose + $HPO_4^{\ 2-}$	–13.8	–13.8	–2.5
Glucose 1-phosphate^{2-}	Glucose + $HPO_4^{\ 2-}$	–20.9	–20.9	
Maltose (or glycogen)	2-Glucose	–16.7	–16.7	
Sucrose	Glucose + fructose	–29.3	–29.3	
UDP glucose^{2-}	Glucose + $UDP^{3-} + H^+$	9.4	–30.5	
N^{10}-Formyltetrahydrofolic acid	Formate$^-$ + H^+ + tetrahydrofolic acid	14.1	–25.9[h]	
Acetic anhydride	2-Acetate$^-$ + $2H^+$	31.1	–48.9	
Acetyl-CoA	Acetate$^-$ + H^+ + CoA	4.9	–35.1[i]	
Succinyl-CoA$^-$	Succinate^{2-} + H^+ + CoA	–3.5	–43.5[j]	
Ethyl acetate	Ethanol + acetate$^-$ + H^+	20.2	–19.7	
Asparagine	Aspartate$^-$ + $NH_4^{\ +}$	–15.1	–15.1	
Glycine ethyl ester$^+$ (39°C)	Glycine + ethanol + H^+	4.9	–35.1	
Valyl-tRNA$^+$	Valine + tRNA + H^+	4.9	–35.1	

---

[a] Unless indicated otherwise, the values are based on tables from Jencks, W. P. (1976) *Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol I (Fasman,G.D. ed.), pp. 296–304. CRC Press, Cleveland, Ohio. For a reaction producing one proton at pH 7 $\Delta G' = -39.96$ kJ mol$^{-1.}$

[b] Guynn, R.W. and Veech, R.L. (1973) *J. Biol. Chem.* **248**, 6966–6972.

[c] Based on +11.80 kcal mol^{-1} for hydrolysis to $P_2O_7^{\ 4-}$ plus ΔG° of dissociation of $HP_2O_7^{\ 3-}$ as quoted by Alberty, R. A. (1969) *J.Biol.Chem.* **244**, 3290–3324. However, 1.017 kcal mol^{-1} was added to the value of 11.80 to make it consistent with that for hydrolysis of ATP to ADP. Reevaluation by Frey and Arabshahi (1995) *Biochemistry* **34**, 11307–11310, indicates that ΔG′ (pH 7) for hydrolysis of ATP to AMP + PP_i is ~ 10 kJ mol^{-1} more negative than is shown here.

[d] Alberty, R.A. (1972) *Horizons of Bioenergetics*, Academic Press, New York, pp. 135–147.

[e] George, P. , Witonsky, R.J., Trachtman, M., Wu, C., Dorwart, W., Richman, L., Richman, W., Shurayh, F., and Lentz, B. (1970) *Biochim. Biophys. Acta.* **223**, 1–15.

[f] Based on $\Delta G^\dagger = 3.0$ kcal mol^{-1} for acetyl phosphate + CoA → acetyl-CoA + P_i from Stadtman, E. R. (1973) *The Enzymes*, (Boyer, P.D., ed.), 3rd ed., Vol. 8. pp. 1–49. Academic Press, New York, together with ΔG′ (pH 7) for hydrolysis of acetyl-CoA.

[g] Estimated from ΔG° for ATP hydrolysis + $\Delta G' = -19.9$ kJ mol^{-1} for the 3-phosphoglycerate kinase reaction: Burton, K. and Krebs, H. A. (1953) *Biophys. J.* **54**, 94–107; and (1955) *Biophys. J.* **59**, 44–46.

[h] Estimated from data at pH 7.7 or 8.0 (tables of Jencks).

[i] Guynn, R. W., Gelberg, H. J., and Veech, R. L. (1973) *J. Biol. Chem.* **248**, 6957–6965 found $\Delta G^\circ = -35.75$ kJ mol^{-1} at 38°C. Without data on ΔH, correction to 25°C is difficult. Burton, K., (1955) *Biophys. J.* **59**, 44–46, gave ΔG′ (pH 7) for ATP^{4-} + acetate$^-$ + CoA → ADP^{3-} + acetyl-CoA + $HPO_4^{\ 2-}$ as approximately zero at 25°C. Guynn *et al.* found –0.56 kJ mol^{-1} at 38°C. This same value (–0.56 kJ) at 25°C was assumed to obtain the figure given here. This is equivalent to assuming ΔH° of hydrolysis as almost the same for ATP and acetyl-CoA, an unsupported assumption.

[j] Assumed 2 kcal mol^{-1} more negative than that of acetyl-CoA as in tables of Jencks.

6-46, we obtain for the hydrolysis of ATP at 25°C, $\mu = 0.25$:

$$\begin{aligned}\Delta G' \text{ (pH 7)} &= -34.5 \text{ kJ mol}^{-1} \\ &= -8.26 \text{ kcal mol}^{-1} \qquad (6\text{-}47)\end{aligned}$$

An additional set of standard states is frequently met in the biochemical literature. An equilibrium constant, designated in this book as $K^\dagger$, is used to relate the *total concentrations of all ionic forms* of the components present at the pH of the experiment. Thus,

$$K^\dagger = \frac{[\text{ADP, all forms}][\text{phosphate, all forms}]}{[\text{ATP, all forms}]} \qquad (6\text{-}48)$$

BOX 6-A MEASUREMENT OF INTRACELLULAR pH

What is the pH within a cell? Is it constant or does it vary with physiological conditions? Do all cells operate at similar pH? The answers to these important questions have been sought using a variety of techniques.[a] Tiny microelectrodes with tips only 1 μm in diameter have been inserted into cells. Indicator dyes have been diffused into cells and either light absorption or fluorescence[b] measured. The distribution of a suitable radiolabeled weak acid or weak base, such as [^{14}C]methylamine, that is able to permeate cells can be used to calculate the difference between internal and external pH.[c] The activity of the enzyme carbonic anhydrase, which is very pH sensitive, can be used to monitor the pH of mitochrondia.[d] Since 1973 NMR methods have been popular.[e–l] The chemical shifts of ^{31}P in inorganic phosphate (pK_a = 6.9), ATP, glucose 6-phosphate, and of other metabolites of ^{13}C in citrate and bicarbonate,[l] give direct estimates of pH. However, caution must be exercised if internal pH values good to ± 0.1 unit are to be obtained.[g] More sensitive measurements over the pH range 1.3 – 9.1 can be made by diffusing one of a series of aminophosphonates into cells and measuring the ^{31}P chemical shift.[h] Fluorinated probes such as dimethylfluoroalanine (pK_a = 7.3) are also useful because of their low toxicity and high sensitivity of the ^{19}F NMR signal. These alanine derivatives can be diffused into cells as their methyl esters, which are rapidly cleaved within cells, allowing the fluorinated amino acids to accumulate.[i,m]

The pH within cells appears to be tightly controlled although small variations are sometimes observed. Red blood cells, thymocytes, liver, skeletal muscles, and intact hearts all maintain a pH in the range 7.0 – 7.3.[b,h,i] However, the pH can fall to 6.2 within 13 minutes of oxygen deprivation (ischemia) and to 6.1 after exhaustive exercise.[n,o] The ^{31}P NMR technique permits the monitoring of pH as well as the state of the adenylate system (Section D) in human limbs suffering from circulatory insufficiency.[c]

In higher plants the pH of cytoplasm is 7.4–7.5 but vacuoles are acidic with a pH of 4.5–6.[l] The cytoplasm of maize root tips has a pH of 7.1 but the vacuoles are at a pH of 5.5.[f] The pH of granules of the chromaffin cells of the adrenal cortex, which accumulate high concentrations of ATP and catecholamines, is also low, ~5.7.[p] The bacterium *Streptococcus faecalis* maintains a higher internal pH of ~8.0, even when the pH of the medium varies from 6.5 to 8.0[q] while *E. coli* operates at pH 7.6, the extremes of variation being 7.4 – 7.8 when the external pH changes from 5.5–9.[r]

Changes of internal pH during developmental events such as fertilization of sea urchin eggs (+0.3 unit) have been recorded. However, the significance of pH changes in metabolic regulation remains uncertain.[c]

---

[a] Kotyk, A., and Slavik, J. (1989) *Intracellular pH and its Measurement*, CRC Press, Boca Raton, Florida

[b] Rogers, J., Hesketh, T. R., Smith, G. A., and Metcalfe, J. C. (1983) *J. Biol. Chem.* **258**, 5994–5997

[c] Nuccitelli, R., and Deamer, D. W., eds. (1982) *Intracellular pH: Its Measurement, Regulation and Utilization in Cellular Functions*, Liss, New York

[d] Dodgson, S. J., Forster, R. E., II, and Storey, B. T. (1982) *J. Biol. Chem.* **257**, 1705–1711

[e] Moon, R. B., and Richards, J. H. (1973) *J. Biol. Chem.* **248**, 7276–7278

[f] Roberts, J. K. M., Ray, P. M., Wade-Jardetsky, N., and Jardetsky, O. (1980) *Nature (London)* **283**, 870–872

[g] Avison, M. J., Hetherington, H. P., and Shulman, R. G. (1986) *Ann. Rev. Biophys. Biophys. Chem.* **15**, 377–402

[h] Pietri, S., Miollan, M., Martel, S., Le Moigne, F., Blaive, B., and Culcasi, M. (2000) *J. Biol. Chem.* **275**, 19505–19512

[i] Taylor, J. S., and Deutsch, C. (1983) *Biophys. J.* **43**, 261–267

[j] Bailey, I. A., Williams, S. R., Radda, G. K., and Gadian, D. G. (1981) *Biochem. J.* **196**, 171–178

[k] Barton, J. K., Den Hollander, J. A., Lee, T. M., MacLaughlin, A., and Shulman, R. G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2470–2473

[l] Gout, E., Bligny, R., and Douce, R. (1992) *J. Biol. Chem.* **267**, 13903–13909

[m] Deutsch, C. J., and Taylor, J. S. (1987) *Ann. N.Y. Acad. Sci.* **508**, 33

[n] Garlick, P. B., Radda, G. K., and Seeley, P. J. (1979) *Biochem. J.* **184**, 547–554

[o] Pan, J. W., Hamm, J. R., Rothman, D. L., and Shulman, R. G. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7836–7839

[p] Pollard, H. B., Shindo, H., Creutz, C. E., Pazoles, C. J., and Cohen, J. S. (1979) *J. Biol. Chem.* **254**, 1170–1177

[q] Kobayashi, H., Murakami, N., and Unemoto, T. (1982) *J. Biol. Chem.* **257**, 13246–13252

[r] Slonczewski, J. L., Rosen, B. P., Alger, J. R., and MacNab, R. M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6271–6275

and

$$\Delta G^{\dagger} = -RT \ln K^{\dagger} \tag{6-49}$$

The Gibbs energy change $\Delta G^{\dagger}$ can be related to $\Delta G'$ by considering the relationship of $K^{\dagger}$ to K'. For ATP hydrolysis in the pH range of 2–10, $K^{\dagger}$ is given by Eq. 6-50.

$$K^{\dagger} = \frac{K'\left(1 + \dfrac{[\mathrm{H}^+]}{K_{\mathrm{HADP}^{2-}}} + \dfrac{[\mathrm{H}^+]^2}{K_{\mathrm{HADP}^{2-}}K_{\mathrm{H_2ADP}^{-}}}\right)\left(1 + \dfrac{[\mathrm{H}^+]}{K_{\mathrm{HPO_4}^{2-}}}\right)}{\left(1 + \dfrac{[\mathrm{H}^+]}{K_{\mathrm{HATP}^{3-}}} + \dfrac{[\mathrm{H}^+]^2}{K_{\mathrm{HATP}^{3-}}K_{\mathrm{H_2ATP}^{2-}}}\right)} \tag{6-50}$$

In this equation $K_{\mathrm{HADP}2-}$, etc., are consecutive dissociation constants as given in Table 6-4. The expressions in parentheses are the **Michaelis pH functions**, which were considered in Chapter 3 (Eqs. 3-4 to 3-6). In Eq. 6-50 they relate the total concentration of each component to the concentration of the most highly dissociated form. Thus, for the pH range 2–10

$$[\mathrm{P_i}]_{\mathrm{total}} = [\mathrm{HPO_4}^{2-}](1 + [\mathrm{H}^+] / K_{\mathrm{H_2PO_4}^-}) \tag{6-51}$$

Using apparent pK_a values ($\mu = 0.2$) for $\mathrm{H_2PO_4}^-$, HADP^{2-}, and HATP^{3-} of 6.78, 6.83, and 7.06 (Table 6-5) and taking $\Delta G'$ at pH 7 as –34.5 kJ mol^{-1}, we compute $\Delta G^{\dagger} = -35.0$ kJ mol^{-1} at pH 7. The difference between $\Delta G'$ and $\Delta G^{\dagger}$ in this case is small, but it would be larger if the ionic forms in Eq. 6-44 were not the ones predominating at pH 7.

To obtain the Gibbs energy change for a reaction under *other than standard conditions*, Eq. 6-29 must be applied. Thus, at pH 7 and 0.01 M activities of ADP^{3-}, ATP^{4-}, and $\mathrm{HPO_4}^{2-}$, ΔG for hydrolysis of ATP according to Eq. 6-44 is $-34.5 - (2 \times 5.71) = -45.9$ kJ mol^{-1} = –11.0 kcal mol^{-1}. We see that at concentrations existing in cells (usually in the millimolar range) ATP has a substantially higher group transfer potential than under standard conditions.

The reader should bear in mind that there is no accepted standard usage of K' and that $K^{\dagger}$ is just for this book! An international committee[21] has recommended that K' be used with the same meaning as $K^{\dagger}$ in this book and more changes may be coming (see the next section).

To obtain ΔG at a temperature other than 25°C, we must know ΔH for the reaction. Using Eq. 6-37a it is easy to show that ΔG at temperature T_2 is related to that at T_1 as follows:

$$\Delta G_2 \approx \frac{T_2\,\Delta G_1 - (T_2 - T_1)\,\Delta H}{T_1} \tag{6-52}$$

The enthalpy of hydrolysis of ATP according to Eq. 6-44 is approximately –19.7 kJ mol^{-1}. Using this value and applying Eq. 6-52 we can calculate that $\Delta G'$ (pH 7) for the hydrolysis of ATP at 38°C is –35.2 kJ mol^{-1}. The value of $\Delta G'$ (25°C, pH 7), used in obtaining this answer, is – 35.54 kJ mol^{-1} which was computed from the value of $\Delta G'$ (pH 7), 38°C of 35.19 kJ mol^{-1} reported by Guynn and Veech[37] using Eq. 6-52. Bear in mind that all of the foregoing Gibbs energy changes are *apparent* values applying to solutions of ionic strength ~0.25.

Both ADP and ATP as well as inorganic pyrophosphate form complexes with metal ions. Since the magnesium complexes are often the predominant forms of ADP and ATP under physiological conditions, we must consider the following Gibbs energy changes. These are apparent values for $\mu = 0.2$ at 25°C.

$$\begin{gathered}\mathrm{ATP}^{4-} + \mathrm{Mg}^{2+} \rightarrow \mathrm{Mg\,ATP}^{2-}\\ \Delta G^{\circ} = -26.3;\ \Delta H^{\circ} = 13.8\ \mathrm{kJ\ mol^{-1}}\end{gathered} \tag{6-53}$$

$$\begin{gathered}\mathrm{ADP}^{3-} + \mathrm{Mg}^{2+} \rightarrow \mathrm{Mg\,ADP}^{-}\\ \Delta G^{\circ} = -19.8;\ \Delta H^{\circ} = 15.1\ \mathrm{kJ\ mol^{-1}}\end{gathered} \tag{6-54}$$

Combining the apparent ΔG° values for Eqs. 6-47, 6-53, and 6-54, we obtain

$$\begin{gathered}\mathrm{MgATP}^{2-} + \mathrm{H_2O} \rightarrow \mathrm{MgADP}^{-} + \mathrm{HPO_2}^{2-} + \mathrm{H}^+\\ \Delta G'\ (\mathrm{pH}\ 7) = -28.0\ \mathrm{kJ\ mol^{-1}}\end{gathered} \tag{6-55}$$

The stoichiometry of Eq. 6-55 never holds exactly. Some of the Mg^{2+} dissociates from the MgADP^-; both protons and Mg^{2+} bind to $\mathrm{H_2PO_4}^-$; and HATP^{3-} and HADP^{2-} are present and bind Mg^{2+} weakly.[38] Thus, the observed value of $\Delta G^{\dagger}$ varies with both pH and magnesium concentration as well as with changes in ionic strength. Tables and graphs showing the apparent value of $\Delta G^{\dagger}$ under various conditions have been prepared by Alberty[38] and by Phillips *et al.*[39] An example is shown in Fig. 6-2. From this graph we find that $\Delta G^{\dagger}$ for hydrolysis of ATP at pH 7, 25°C, $\mu = 0.2$, and 1 mM Mg^{2+} (a relatively high intracellular concentration[40]) is –30.35 kJ mol^{-1} (–7.25 kcal mol^{-1}).

Figure 6-2 was drawn using the equations of Alberty, but the value of $\Delta G'$ (pH 7) of hydrolysis of ATP = 34.54 kJ mol^{-1} at $[\mathrm{Mg}^{2+}] = 0$ based on results of Guynn and Veech[37] was used. The pK_a values and formation constants of Mg^{2+} complexes were those of Alberty.[38] Note that George *et al.*[40] provided formation constants of these complexes at infinite dilution where the values of ΔG° of formation are considerably more negative than those given in Eqs. 6-53 to 6-55.

From the foregoing considerations we see that complexing with Mg^{2+} somewhat decreases the group transfer potential of the phospho group of ATP. Furthermore, changes in the concentration of free Mg^{2+} with time and between different regions of a cell may

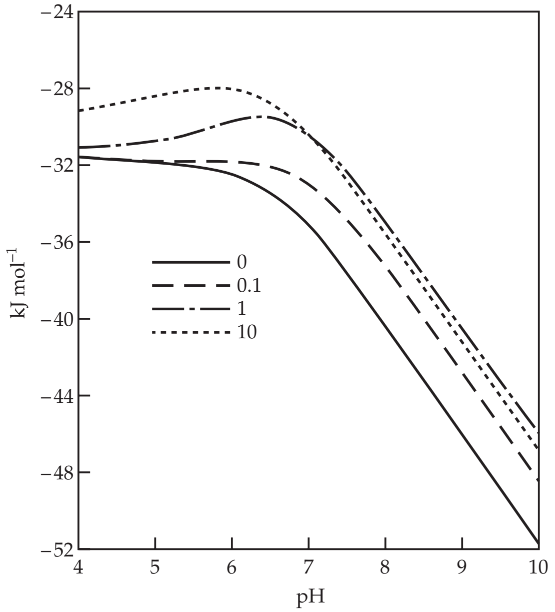


Figure 6-2 Plots of the apparent Gibbs energy $\Delta G^{\dagger}$ for hydrolysis of ATP as a function of pH at a series of different concentrations of free magnesium ions. Millimolar [Mg^{2+}] is indicated by the numbers by the curves. Computer-drawn graphs courtesy of Carol M. Metzler.

---

have significant effects.[41] While Mg^{2+} is a principal cation in tissues, it is by no means the only one. Thus, Ca^{2+}, Mn^{2+}, and even K^+ will affect equilibria involving polyphosphates such as ATP.

It is not easy to measure the group transfer potential of ATP and published values vary greatly. George *et al.*[40] reported $\Delta G°$ for Eq. 6-42 as −39.9 kJ mol^{-1} at 25°C, $\mu = 0.2$, and as −41.25 kJ mol^{-1} at infinite dilution. They regarded these values as good to ±4 kJ mol^{-1}. However, most other estimates have been at least 4 kJ mol^{-1} less negative.[42–44] The self-consistent set of thermodynamic data used throughout this book are based in part on the value of ΔG for hydrolysis of ATP obtained by Guynn and Veech.[37]

5. A New Standard for Biochemical Thermodynamics?

Because of the complexities of the equilibria involved in biochemical reactions Alberty and others[45–47] and a Panel on Biochemical Thermodynamics[48] have suggested that tables of thermodynamic properties for biochemical use be tabulated for the following conditions: $T = 298.15$ K, $P = 1$ bar (0.1 MP$_a$ ≡ 0.987 atm), pH = 7, pMg = 3 ([Mg^{2+}] = 10^{-3} M), and ionic strength I (μ in this chapter) = 0.25 M. There appears to be both advantages and disadvantages to this. The proponents suggest that the symbols $\Delta G'$, K', etc. be used for this new standard. To change all of the numbers in this book, which uses a self-consistent set of thermodynamic quantities, is impractical. However, it is worthwhile to compare the value of Alberty and Goldberg for $\Delta G'°$ for hydrolysis of ATP to ADP + P$_i$ under the proposed biochemical standard conditions with other values given in this chapter, all at 25°C.

$\Delta G'° = -32.49$ kJ mol^{-1} (new proposed biochemical standard)
$\Delta G^{\dagger}$ (pH 7) −35.0 kJ mol^{-1} (Eq. 6-50)
$\Delta G'$ (pH 7) −34.5 kJ mol^{-1} (for $ATP^{4-} \rightarrow ADP^{3-} + HPO4^{2-}$; Eq. 6-47)
$\Delta G'$ (pH 7) −28.0 kJ mol^{-1} (for $MgATP^{2-} \rightarrow MgADP^- + HPO4^{2-}$; Eq. 6-55)

6. Bond Energies and Approximate Methods for Estimation of Thermodynamic Data

For approximate estimation of enthalpy changes during reactions, use can be made of empirical bond energies (Table 6-7) which represent the approximate enthalpy changes ($-\Delta H°$) for formation of compounds in a gaseous state from atoms in the gas phase. Other more comprehensive methods of approximation have been developed.[49,50]

7. Gibbs Energies of Combustion by O_2 and by NAD^+

Since oxidation processes are so important in the metabolism of aerobic organisms, it is often convenient to discuss Gibbs energies of combustion. These are easily derived from the Gibbs energies of formation. For example, ΔG_c for acetate ion (aqueous) may be obtained as follows:

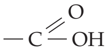
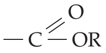
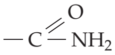
$$\text{Acetate}^- + H^+ \rightarrow 2\,C + 2\,H_2 + O_2$$
$$\Delta G° = +369.2 \text{ kJ mol}^{-1} = -\Delta G_f° \text{ of acetate (from Table 6-3)}$$

$$2\,H_2 + O_2 \rightarrow 2\,H_2O$$
$$\Delta G° = 2 \times -237.2 = -474.4 \text{ kJ mol}^{-1} = 2 \times \Delta G_f° \text{ of } H_2O \qquad (6\text{-}56)$$

$$2\,C + 2\,O_2 \rightarrow 2\,CO_2$$
$$\Delta G° = 2 \times -394.4 = -788.8 \text{ kJ mol}^{-1} = 2 \times \Delta G_f° \text{ of } CO_2$$

$$\text{Acetate}^- (1\text{ M}) + H^+ (1\text{ M}) + 2\,O_2 (1\text{ atm}) \rightarrow 2\,CO_2 (1\text{ atm}) + H_2O\text{ (l)}$$
$$\Delta G° = -894.0 \text{ kJ mol}^{-1} = \Delta G_c° \text{ of acetate}^-$$

TABLE 6-7
Empirical Bond Energies and Resonance Energies[a]

Energy values for single bonds $(\mathrm{kJ\ mol^{-1}})$		Energy values for multiple bonds $(\mathrm{kJ\ mol^{-1}})$		Empirical resonance energy values $(\mathrm{kJ\ mol^{-1}})$	
H–H	436	C=C	615	Benzene	155
C–C	346	N=N	418	Naphthalene	314
Si–Si	177	O=O	402 ('Δ state)		
N–N	161	C=N	615	Styrene	155 + 21
O–O	139	C=O	686 (formaldehyde)	Phenol	155 + 29
S–S	213		715 (aldehydes)		
C–H	413		728 (ketones)	Benzaldehyde	155 + 17
Si–H	295	C=S	477	Pyridine	180
N–H	391	C≡C	812	Pyrrole	130
P–H	320	N≡N	946 (N_2)	Indole	226
O–H	463	C≡N	866 (HCN)		
S–H	339		891 (nitriles)	 $-\mathrm{C}(=\mathrm{O})-\mathrm{OH}$	117
C–Si	290				
C–N	292			 $-\mathrm{C}(=\mathrm{O})-\mathrm{OR}$	100
C–O	351				
C–S	259				
C–F	441			 $-\mathrm{C}(=\mathrm{O})-\mathrm{NH_2}$	88
C–Cl	328				
C–Br	276			Urea	155
C–I	240			CO	439
Si–O	369			CO_2	151

[a] From Pauling, L. (1960) *The Nature of the Chemical Bond*, 3rd ed., pp. 85, 189, and 195–198, Cornell Univ. Press, Ithaca, New York

Table 6-4 lists values of $\Delta G_c°$ as well as those of $\Delta G_f°$. Besides CO_2 and H_2O, the other products assumed in calculating the values of $\Delta G_c°$ in Table 6-4 are N_2 (1atm), H_3PO_4 (1 M), and H_2SO_4 (1 M).

Much of the oxidation occurring in cells is carried out by the biological oxidizing agent **nicotinamide adenine dinucleotide** (**NAD$^+$**) or by the closely related NADP$^+$ (Chapter 10). It is convenient to tabulate values of $\Delta G°$ for complete oxidation of compounds to CO_2 using NAD$^+$ rather than O_2 as the oxidizing agent. These values, designated $\Delta G°_{ox}$ and $\Delta G'_{ox}$ (pH 7), are also given in Table 6-3. Notice that these values are relatively small, corresponding to the fact that little energy is made available to cells by oxidation with NAD$^+$; for example (Eq. 6-57):

$$\text{Acetate}^- + 2\ H_2O + 4\ NAD^+ \rightarrow 2\ CO_2 + 4\ NADH + 3\ H^+\ (10^{-7}\ M)$$
$$\Delta G'_{ox}\ (\text{pH 7}) = +22.1\ \text{kJ mol}^{-1} \qquad (6\text{-}57)$$

When the reduced NAD$^+$ (designated NADH) so formed is reoxidized in mitochondria (Eq. 6-58), a large amount of energy is made available to cells.

$$4\ NADH + 4\ H^+ + 2\ O_2 \rightarrow 4\ H_2O + 4\ NAD^+$$
$$\Delta G'_{ox}\ (\text{pH 7}) = -876.1\ \text{kJ mol}^{-1} \qquad (6\text{-}58)$$

The sum of Eqs. 6-57 and 6-58 is the equation for combustion of acetate by O_2 (Eq. 6-59) and the two ΔG values sum to ΔG_c for acetate$^-$.

$$\text{Acetate}^- + H^+ + 2\ O_2 \rightarrow 2\ CO_2 + 2\ H_2O$$
$$\Delta G_c° = -894.0\ \text{kJ mol}^{-1}$$
$$\Delta G_c'\ (\text{pH 7}) = -854.0\ \text{kJ mol}^{-1} \qquad (6\text{-}59)$$

The values of ΔG_{ox} (Table 6-4) not only give an immediate indication of the relative amounts of energy available from oxidation of substrate with NAD$^+$ but also are very convenient in evaluating ΔG for fermentation reactions. For example, consider the fermentation of glucose to ethanol (Eq. 6-60):

$$\alpha\text{-D-Glucose} \rightarrow 2\ CO_2 + 2\ \text{ethanol} \qquad (6\text{-}60)$$

The Gibbs energy change $\Delta G'$ (pH 7) for fermentation of glucose to ethanol and CO_2 can be written immediately from the data of Table 6-4 (Eq. 6-61).

$$\Delta G'\ (\text{pH 7}) = -243.8 - 2(-4.6)$$
$$= -234.6\ \text{kJ mol}^{-1} \qquad (6\text{-}61)$$

The values of ΔG_{ox} for H_2O, CO_2, and H^+ are always zero and need not be considered. The same computation can be made using the awkwardly large values of $\Delta G_c°$ or using values of $\Delta G_f°$ The latter are also large, and CO_2 and water must be considered in the equations. Table 6-4 can be used to obtain values of $\Delta G°$ for many metabolic reactions considered later in the book. Data from any column in the table may be used for this purpose, but for simplicity try the values in the $\Delta G'_{ox}$ column.

Note, however, that for reactions involving oxidation by O_2 or by any oxidant, other than NAD$^+$, not appearing in Table 6-4, the following two-step procedure is necessary if the ΔG_{ox} values are used. From the ΔG_{ox} values compute $\Delta G°$ or $\Delta G'$ for the reaction under consideration using NAD$^+$ as the oxidant. Then add to this the Gibbs energy of oxidation by O_2 (or other oxidant) of the appropriate number of moles of NADH formed. The latter is given for O_2 in Table 6-8 and can also be evaluated for a number of other oxidants, such as Fe^{3+} and cytochrome *c*, from the data in Table 6-8.

BOX 6-B MAGNESIUM

After potassium ion Mg^{2+} is the most abundant cation present in tissues[a,b]. The average adult ingests ~10–12 mmol of magnesium ion daily (~1/4 g). Of this, about one-third is absorbed from the digestive tract. An equivalent amount is excreted in the urine to maintain homeostasis. Sixty percent of the magnesium in the body is found in the bones. The total Mg^{2+} concentration in serum is ~1 mM, while various tissues contain as much as 5 – 17 mM.[c] The concentration of *free* Mg^{2+} is difficult to measure. A variety of measurements using ^{31}P NMR spectroscopy of ATP,[d–f] ^{19}F NMR of added fluorine-containing chelators,[g] fluorescent chelators,[h] ion selective electrodes,[c] and other indirect procedures[g,i,j] indicate that the free magnesium concentration is almost the same in extracellular fluids, cytosol, and mitochondria.[h] Most values have been about 0.5 mM, somewhat less in erythrocytes. However, the most recent estimates indicate an intracellular $[Mg^{2+}]$ of 0.8–1.1 mM.[c,f]

The additional Mg^{2+} present in cells is bound to proteins, nuclei acids, and soluble compounds such as ATP, ADP, citrate,[k] and other phosphate- and carboxylate-containing substances. The binding is reversible and equilibration is usually rapid. It has been suggested that $[Mg^{2+}]$, like $[H^+]$, remains relatively constant within cells and that these two ions are in free equilibrium in the blood serum.[l] Nevertheless, there are instances in which at least temporary alterations in concentrations of both free Mg^{2+} and H^+ occur.[m] During rapid catabolism of carbohydrates the formation of lactic acid by glycolysis leads to acidification of muscle cells, the pH falling from 7.3 to as low as 6.3. This drop in pH causes a large decrease in the extent of binding of Mg^{2+} to molecules such as ATP and to a transient increase in $[Mg^{2+}]$. Similarly, the release of bisphosphoglycerate from hemoglobin upon oxygenation leads to a decreased concentration of free Mg^{2+} as the latter coordinates with bisphosphoglycerate.[n] The 0.25 mM $[Mg^{2+}]$ of aerobic red blood cells rises to 0.67 mM under anaerobic conditions.[d] Such changes in the free Mg^{2+} concentrations will affect many equilibrium[o] and may be of significance in metabolic regulation.

The magnesium ion has a smaller radius than Ca^{2+}, a fact that may account for its more ready entry into cells. Mg^{2+} can often be replaced by Mn^{2+} with full activity for enzymes that require it. On the other hand, high concentrations of Ca^{2+} are often antagonistic to Mg^{2+}. This antagonism is clearly seen in the effect of the two ions on irritability of protozoa.[p] Both deficiency of magnesium and excess of calcium in the surrounding medium cause increased irritability. Excess magnesium leads to anesthesia. The Mg^{2+} concentration is high in hibernating animals.

Over 300 enzymes are dependent upon Mg^{2+}, the largest single group being the phosphotransferases, for which MgATP complex may be regarded as the substrate.[q] Magnesium has a special role in photosynthesis as a component of chlorophyll.

One of the most toxic metals is beryllium. It has been suggested that Be^{2+} competes with Mg^{2+} at many enzyme sites, including those of phosphoglucomutase and of phosphatases. However, as pointed out by Petsko,[r] because of its small size beryllium tends not to form Be^{2+} ions but, rather, covalent complexes such as $BeF_3•OH_2^-$, BeF_4^{2-}, and MgADP•-$BF_2(OH)$. In the latter complex beryllium is covalently linked to the oxygen atom of the β phospho group of ADP to give an analog of MgATP, which inhibits many enzymes. See for example, Fig. 19-16A, in which MgADP•BeF_x occupies the active site of myosin.

---

[a] Cowan, J. A., ed. (1995) *The Biological Chemistry of Magnesium*, VCH Publ., New York

[b] Strata, P., and Carbone, E., eds. (1991) *Mg^{2+} and Excitable Membranes*, Springer-Verlag, Berlin and New York

[c] McGuigan, J. A. S., Blatter, L. A., and Buri, A. (1991) in *Mg^{2+} and Excitable Membranes* (Strata, P., and Carbone, E., eds), Springer-Verlag, Berlin and New York

[d] Gupta, R. K., and Moore, R. D. (1980) *J. Biol. Chem.* **255**, 3987–3993

[e] Garfinkel, L., and Garfinkel, D. (1984) *Biochemistry* **23**, 3547–3552

[f] Clarke, K., Kashiwaya, Y., King, M. T., Gates, D., Keon, C. A., Cross, H. R., Radda, G. K., and Veech, R. L. (1996) *J. Biol. Chem.* **271**, 21142–21150

[g] Levy, L. A., Murphy, E., Raju, B., and London, R. E. (1988) *Biochemistry* **27**, 4041–4048

[h] Jung, D. W., Apel, L., and Brierley, G. P. (1990) *Biochemistry* **29**, 4121–4128

[i] Corkey, B. E., Duszynski, J., Rich, T. L., Matschinsky, B., and Williamson, J. R. (1986) *J. Biol. Chem.* **261**, 2567–2574

[j] Magneson, G. R., Puvathingal, J. M., and Ray, W. J., Jr. (1987) *J. Biol. Chem.* **262**, 11140–11148

[k] Kwack, H., and Veech, R. L. (1992) *Curr. Top. Cell. Regul.* **33**, 185–207

[l] Veloso, D., Guynn, R. W., Oskarrson, M., and Veech, R. L. (1973) *J. Biol. Chem.* **248**, 4811–4819

[m] Purich, D. L., and Fromm, H. J. (1972) *Curr. Top. Cell. Regul.* **6**, 131–167

[n] Bunn, H. F., Ransil, B. J., and Chao, A. (1971) *J. Biol. Chem.* **246**, 5273–5279

[o] Cornell, N. W. (1979) *J. Biol. Chem.* **254**, 6522–6527

[p] Meli, J., and Bygrave, F. L. (1972) *Biochem. J.* **128**, 415–420

[q] Vink, R., McIntosh, T. K., and Faden, A. I. (1991) in *Mg^{2+} and Excitable Membranes* (Strata, P., and Carbone, E., eds), Springer-Verlag, Berlin and New York

[r] Petsko, G. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 538–540

C. Electrode Potentials and Gibbs Energy Changes for Oxidation–Reduction Reactions

We live under a blanket of the powerful oxidant O_2. By cell respiration oxygen is reduced to H_2O, which is a very poor reductant. Toward the other end of the scale of oxidizing strength lies the very weak oxidant H^+, which some bacteria are able to convert to the strong reductant H_2. The O_2–H_2O and H^+ – H_2 couples define two biologically important oxidation–reduction (**redox**) systems. Lying between these two systems are a host of other pairs of metabolically important substances engaged in oxidation–reduction reactions within cells.

There are two common methods for expressing the oxidizing or reducing powers of redox couples in a quantitative way. On the one hand, we can list values of ΔG for oxidation of the reduced form of a couple to the oxidized form by O_2. A compound with a large value of $-\Delta G$ for this oxidation will be a good reductant. An example is H_2 for which ΔG of combustion at pH 7 (Table 6-4) is – 237 kJ / mol. Poor reductants such as Fe^{2+} are characterized by small values of ΔG of oxidation (-8.5 kJ mol^{-1} for $2\,Fe^{2+} \rightarrow 2\,Fe^{3+}$). The Gibbs energies of oxidation of biological hydrogen carriers, discussed in Chapter 15, for the most part fall between those of H_2 and Fe^{2+}.

A second way of expressing the same information is to give **electrode potentials** (Table 6-8). Electrode potentials are also important in that their direct measurement sometimes provides an experimental approach to the study of oxidation–reduction reactions within cells. To measure an electrode potential it must be possible to reduce the oxidant of the couple by flow of electrons (Eq. 6-62) from an electrode surface, often of specially prepared platinum.

$$A + 2\,H^+ + 2\,e^- \rightarrow AH_2 \qquad (6\text{-}62)$$

Equation 6-62 represents a reversible reaction taking place at a single electrode. A complete electrochemical cell has two electrodes and the reaction occurring is the sum of two half-reactions. The electrode potential of a given half-reaction is obtained from the measured electromotive force of a complete cell in which one half-reaction is that of a standard **reference electrode** of known potential. Figure 6-3 indicates schematically an experimental setup for measurement of an electrode potential. The standard hydrogen electrode consists of platinum over which is bubbled hydrogen gas at one atmosphere pressure. The electrode is immersed in a solution containing hydrogen ions at unit activity ($a_{H^+} = 1$). The potential of such an electrode is conventionally taken as zero. In practice it is more likely that the reference electrode will be a calomel electrode or some other electrode that has been established experimentally as reliable and whose potential is accurately known.

The standard electrode is connected to the experimental electrode compartment by an electrolyte-filled bridge. In the experimental compartment the reaction represented by Eq. 6-62 occurs at the surface of another electrode (often platinum). The voltage difference between the two electrodes is measured with a potentiometer. The difference between the observed voltage and that of the reference electrode gives the electrode potential of the couple under investigation. It is important that the electrode reaction under study be strictly reversible. When the electromotive force (emf) of the experimental cell is balanced with the potentiometer against an external voltage source, no current flows through the cell. However, for a reversible reaction a slight change in the applied voltage will lead to current flow. The flow will be in either of the two directions, depending upon whether the applied voltage is raised or lowered.

Not all redox couples are reversible. This is especially true of organic compounds; for example, it is not possible to determine readily the electrode potential for an aldehyde–alcohol couple. In some cases, e.g., with enzymes, a readily reducible dye with a potential similar to that of the couple being measured can be added. A list of suitable dyes has been described by Dutton.[51] If the dye is able to rapidly exchange electrons with the couple being studied, it is still possible to measure the electrode potential directly. In many cases electrode potentials appearing in tables have been calculated from Gibbs energy data. The student should be able to calculate many of the potentials in Table 6-8 from Gibbs energy data from Table 6-4. If A, H^+, and AH_2 are all present at unit activity in the experimental cell, the observed potential for the half-reaction will be the **standard electrode potential** $E°$. If the emf of the hypothetical cell with the standard hydrogen electrode is positive when electron flow is in the direction indicated by the arrow in Fig. 6-3, the potential of the couple A / AH_2 is also taken as positive (and is often called a **reduction potential**). This is the convention used in establishing Table 6-8, but potentials of exactly the opposite sign (oxidation potentials) are used by some chemists. To avoid confusion in reading it is best to be familiar with values of one or two potentials such as those of the O_2–H_2O and NAD^+–NADH couples.

When electrons flow in the external circuit the maximum amount of work that can be done per mole of electrochemical reaction ($-\Delta G$) is given by Eq. 6-63

$$\begin{aligned}\Delta G &= nEF = nE \times 96.487 \text{ kJ mol}^{-1}\text{ V}^{-1}\\ &= -\,nE \times 23.061 \text{ kcal mol}^{-1}\text{ V}^{-1}\end{aligned} \qquad (6\text{-}63)$$

where F equals the number of coulombs per mole of electrons (Avogadro's number multiplied by the charge on the electron = 96,487 coulombs). E is measured in volts and represents the difference of the electrode potentials of the two half-cells. In the case of a cell using the standard hydrogen electrode, E is the electrode potential of the experimental couple. The number of moles of electrons transferred in the reaction equation (n) is usually 1 or 2 in biochemical reactions (2 for Eq. 6-62).

Since the reactants and products need not be at unit activity, we must define the observed electrode potential E as a function of $E°$ and the activities (concentrations) of A, AH_2, and H^+ (Eq. 6-64).

$$E = E° + \frac{RT}{nF} \ln \frac{[A][H^+]^2}{[AH_2]}$$

If $n = 2$

$$E = E° + 0.0296 \log \frac{[A][H^+]^2}{[AH_2]} \quad \text{volts at 25°C} \tag{6-64}$$

In the biochemical literature values of the apparent standard electrode potential at pH 7 ($E°'$) are usually tabulated instead of values of $E°$ (Table 6-8, second column). Note that $E°'$ (pH 7) for the hydrogen electrode is not zero, as it is at pH 0, but −0.414 V. Values of E are related to $E°'$ by Eq. 6-64 with $[H^+]^2$

TABLE 6-8
Reduction Potentials of Some Biologically Important Systems[a,b]

Half-reaction	$E°$ (V)	$E°'$ (pH 7) (V)	$-\Delta G'$ (pH 7) (kJ mol^{-1}) for oxidation by O_2 (per 2 electrons)
$O_2 + 4\,H^+ + 4e^- \rightarrow 2\,H_2O$	+1.229	+0.815	0.0
$Fe^{3+} + e^- \rightarrow Fe^{2+}$	0.771	0.771	8.5
$NO_3^- + 2\,H^+ + 2e^- \rightarrow NO_2^- + H_2O$0.42176.0			
Cytochrome *f* (Fe^{3+}) + e^- → cytochrome *f* (Fe^{2+})		0.365	86.8
Fe $(CN)_6^{3-}$ (ferricyanide) + e^- → Fe $(CN)_6^{4-}$		0.36	87.8
$O_2 + 2H^+ + 2\,e^- \rightarrow H_2O_2$	0.709	0.295	100.3
Cytochrome *a* (Fe^{3+}) + e^- → Cytochrome *a* (Fe^{2+})		0.29	101.3
p-Quinone + 2 H^+ + 2 e^- → hydroquinone	0.699	0.285	102.3
Cytochrome *c* (Fe^{3+}) + e^- → cytochrome *c* (Fe^{2+})		0.254	108.3
Adrenodoxin (Fe^{3+}) + e^- → adrenodoxin (Fe^{2+})		0.15	128.3
Cytochrome b_2 (Fe^{3+}) + e^- → cytochrome b_2 (Fe^{2+})		0.12	134.1
Ubiquinone + 2 H^+ + 2 e^- → ubiquinone H_2		0.10	138.0
Cytochrome *b* (Fe^{3+}) + e^- → cytochrome *b* (Fe^{2+})		0.075	142.8
Dehydroascorbic acid + 2 H^+ + 2 e^- → ascorbic acid		0.058	146.1
Fumarate^{2-} + 2 H^+ + 2 e^- → succinate^{2-}		0.031	151.3
Methylene blue + 2 H^+ + 2 e^- → leucomethylene blue (colorless)		0.011	155.2
Crotonyl-CoA + 2 H^+ + 2 e^- → butyryl-CoA		−0.015	160.2
Glutathione + 2 H^+ + 2 e^- → 2-reduced glutathione		~ −0.10	176.6
Oxaloacetate^{2-} + 2 H^+ + 2 e^- → malate^{2-}		−0.166	189.3
Pyruvate$^-$ + 2 H^+ + 2 e^- → lactate^{1-}		−0.185	193.0
Acetaldehyde + 2 H^+ + 2 e^- → ethanol		−0.197	195.3
Riboflavin + 2 H^+ + 2 e^- → dihydroriboflavin		−0.208	197.4
Acetoacetyl-CoA + 2 H^+ + 2 e^- → β-hydroxybutyryl-CoA		−0.238 (38°C)	203.2
$S + 2\,H^+ + 2\,e^- \rightarrow H_2S$	0.14	−0.274	210.2
Lipoic acid + 2 H^+ + 2 e^- → dihydrolipoic acid		−0.29	213.2
$NAD^+ + H^+ + 2\,e^- \rightarrow NADH$	−0.113	−0.32	219.0
$NADP^+ + H^+ + 2\,e^- \rightarrow NADPH$		−0.324	219.8
Ferredoxin (Fe^{3+}) + e^- → ferredoxin (Fe^{2+}) (*Clostridia*)		−0.413	237.0
$2\,H^+ + 2\,e^- \rightarrow H_2$	0	−0.414	237.2
$CO_2 + H^+ + 2\,e^- \rightarrow$ formate$^-$		−0.42 (30°C)	238.3
Ferredoxin (Fe^{3+}) + e^- → ferredoxin (Fe^{2+}) (spinach)		−0.432	240.6

---

[a] A compound with a more positive potential will oxidize the reduced form of a substance of lower potential with a standard free energy change $\Delta G° = -n\text{F}\,\Delta E° = -n\,\Delta E° \times 96.49$ kJ mol^{-1} where n is the number of electrons transferred from reductant to oxidant. The temperature is 25°C unless otherwise indicated. $E°$ refers to a standard state in which the hydrogen ion activity = 1; $E°'$ refers to a standard state of pH 7, but in which all other activites are unity.

[b] The major source is Loach, P. A. (1976) in *Handbook of Biochemistry and Molecular Biology* 3rd ed. Vol. I (Fasman, G. D. ed.), pp. 122–130, CRC Press, Cleveland, Ohio.

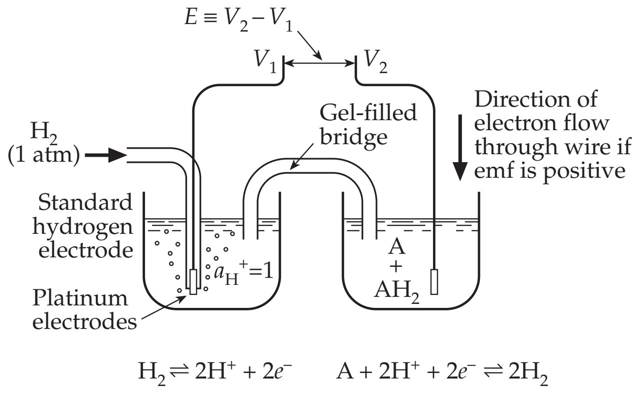
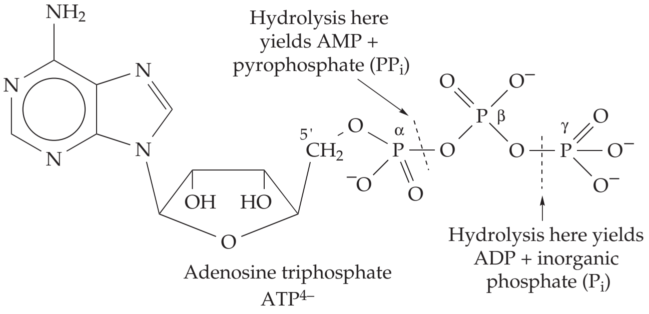


Figure 6-3 Device for measurement of electrode potentials. The electrode reactions are indicated below each half-cell. The maximum electrical work that can be done by such a cell on its surroundings is $-\Delta G = nEF$, where $E = V_2 - V_1$ as measured by a potentiometer. If A is reduced to AH_2 by H_2, electrons will flow through an external circuit as indicated. A will be reduced in the right-hand cell. H_2 will be oxidized to H^+ in the left-hand cell. Protons will flow through the gel bridge from left to right as one of the current carriers in the internal circuit.

deleted from the numerator (since the term in log $[H^+]$ is contained in $E°'$). On the scale of $E°'$ (pH 7) the potential of the oxygen–water couple is 0.815 V, while that of the NAD^+–NADH couple is −0.32 V.

D. The Adenylate System


NH2
N
N
N
N
5'
CH2
O
OH HO
O
α
P
O
O−
O
β
P
O
O−
O
γ
P
O
O−
O−
Hydrolysis here yields AMP + pyrophosphate (PPi)
Hydrolysis here yields ADP + inorganic phosphate (Pi)
Adenosine triphosphate
ATP4−

Of central importance to the energy metabolism of all cells is the **adenylate system** which consists of adenosine 5'-triphosphate (**ATP**), adenosine 5'-diphosphate (**ADP**), and adenosine 5'-monophosphate (**AMP**) together with **inorganic phosphate** (P_i), **pyrophosphate** (PP_i), and **magnesium** ions. Remember that P_i refers to the mixture of ionic forms of phosphoric acid present under experimental conditions. Between pH 4 and pH 10 this will be mainly $H_2PO_4^-$ ($pK_a = 6.8$) and HPO_2^{2-}. Likewise, the symbols AMP, ADP, and ATP refer to mixtures of ionic forms and PP_i refers to a mixture of the ions of pyrophosphoric acid. Above pH 4.4 only $H_2P_2O_7^{2-}$ ($pK_a = 6.1$), $HP_2O_7^{3-}$ ($pK_a = 9.0$), and $P_2O_7^{4-}$ contribute appreciably to PP_i.

1. Storage and Utilization of Energy

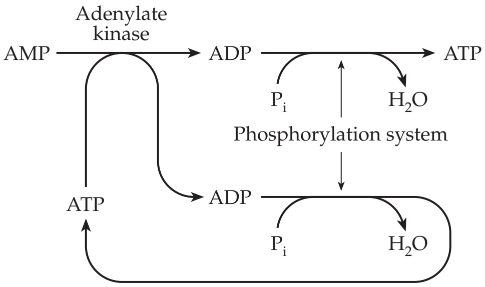
ATP is a thermodynamically unstable molecule with respect to hydrolysis to either ADP or AMP as is indicated in the foregoing diagram. The standard Gibbs energy of hydrolysis, $\Delta G'$, of ATP^{4-} to ADP^{3-} + HPO_4^{2-} at pH 7 is −34.5 kJ mol^{-1} and that of hydrolysis of ADP^{3-} to AMP^{2-} + HPO_4^{2-} is −36.3 kJ mol^{-1} at 25°C (Table 6-6). The exact value of these changes in Gibbs energy depends on pH and on the concentration of Mg^{2+} as is detailed in Sections B, 4, 5. Rates of reaction of components may also depend upon metal ions. Magnesium ion is especially important and complexes such as $MgATP^{2-}$ are regarded as the true substrates for many ATP-utilizing enzymes.

The large *Gibbs energy decreases upon hydrolysis (high group transfer potentials)* enable cells to use ATP and ADP as stores of readily available energy. Energy from the adenylate system is used for many purposes including biosynthesis, transport of ions and molecules across membranes, and for doing mechanical work. The mechanisms by which this energy is utilized are considered later (see Chapters 10, 12, and 17–19). The first step most often involves transfer of either the terminal γ-phospho group of ATP to a site on a different molecule or transfer of the entire AMP portion of the molecule onto another group. Thus the products of cleavage of ATP in these energy-utilizing processes may be either ADP + P_i or AMP + PP_i. In the latter case the pyrophosphate is usually hydrolyzed rapidly to two molecules of P_i by pyrophosphatases. This process, too, serves an essential function in the adenylate system, because it removes a product of the initial ATP cleavage reaction, shifting the overall equilibrium of the reaction sequence in the direction of the products.

The adenylate system provides the major store of rapidly available energy but the whole family of **nucleoside triphosphates** that are related to ATP in structure have similar functions. These include guanosine triphosphate (GTP), uridine triphosphate (UTP), cytidine triphosphate (CTP), and deoxyribonucleotide triphosphates (Table 5-1). These compounds are formed by successive transfer to the nucleoside monophosphate (GMP, UMP, CMP, dGMP, etc.) from two different molecules of ATP of two phospho groups. The resulting compounds are used to provide energy for a variety of specific biosynthetic processes, including synthesis of RNA and DNA. Inorganic **polyphosphates**, linear polymers of orthophosphate (P_i), are present in nearly all living forms.[52–54] Like ATP, they can also store energy and in some organisms substitute for ATP in certain enzymatic reactions.

2. Synthesis of ATP

The **phosphate anhydride** (pyrophosphate) linkages of ATP are generated by the joining of ADP and inorganic phosphate by means of special **phosphorylation** reactions. The most important of the latter occur in the photosynthetic membranes of chloroplasts (**photosynthetic phosphorylation**) and in oxygen-utilizing membranes of bacteria and of mitochondria (**oxidative phosphorylation**). Conversion of AMP to ADP is accomplished by transfer of the terminal phospho group from an ATP molecule to AMP in a reaction catalyzed by the extremely active enzyme **adenylate kinase** (Chapter 12) which is found in all cells. The following equations indicate how one of the special phosphorylation reactions must be used twice for the conversion of one molecule of AMP into one molecule of ATP.



(6-65)

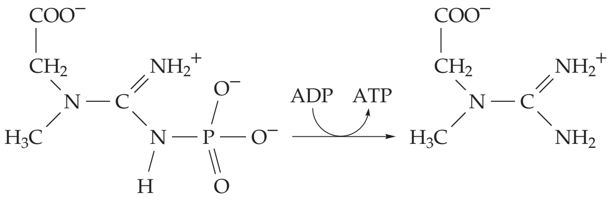
Various measures of the phosphorylating potential of the adenylate system within cells have been proposed. One measure is the product [ATP] / [ADP][P_i], which will be called the **phosphorylation state ratio** or R_p in this book. It is also sometimes called the "phosphorylation potential." This ratio directly affects the Gibbs energy of hydrolysis of ATP, as is shown by Eq. 6-29. The value of R_p may be as high as 10^4 to 10^5 M^{-1} within cells[55] adding – 22.8 kJ mol^{-1} to ΔG of hydrolysis of ATP. Another quantity that is sometimes cited is the "energy charge," the mole fraction of adenylic acid "charged" by conversion to ATP. ADP is regarded as "half-charged."[56,57]

$$\text{Energy charge} = \frac{[\text{ATP}] + \frac{1}{2}[\text{ADP}]}{[\text{ATP}] + [\text{ADP}] + [\text{AMP}]} \tag{6-66}$$

The energy charge varies from 0 if only AMP is present to 1.0 if all of the AMP is converted to ATP. Measurements on a variety of cells and tissues show that the energy charge is usually between 0.75 and 0.90. Although it is easy to calculate its numerical value, the energy charge cannot be used in chemical equations and the idea that the energy charge of a cell plays a key role in regulation of metabolism has been challenged.[58]

3. Creatine Phosphate, an Energy Buffer

Although ATP provides the immediate source of energy for operating muscle, its concentration is only about 5 mM. However, muscle contains, in addition, a **phosphagen**, an *N*-phospho derivative of a guanidinium compound. In mammalian muscle, the phosphagen is **creatine phosphate**. Related compounds including **arginine phosphate** serve in various invertebrates. The group transfer potential ($-\Delta G°'$ of hydrolysis) for creatine phosphate is 43.1 kJ mol^{-1}. Thus, the transfer of a phospho group from creatine phosphate to ADP to form ATP (Eq. 6-67) is spontaneous with $\Delta G' = -8.6$ kJ mol^{-1}. Recent values under a variety of conditions have been reported by Taugue and Dobson.[59] Creatine phosphate, which is present at a concentration of 20 mM, provides a reserve of high-energy phospho groups and keeps the adenylate system of muscle buffered at a high value of R_p.



Creatine phosphate

Creatine

$\Delta G' = -8.6$ kJ mol^{-1} (6-67)

4. Phosphorus-31 NMR

A spectacular development is the ability to observe components of the adenylate system as well as phosphocreatine and other phosphate esters in living cells by ^{31}P NMR.[60–61a] Spectra can be recorded on suspensions of cells[60] or of mitochondria,[62] on individual excised muscles (Fig. 6-4),[63,64] or on perfused organs[65,66] including beating rat hearts and surgically exposed animal organs.[61,61a] Metabolism can be observed in human erythrocytes[67] and even in human limbs, liver, hearts, and brain.[68–70] The method can play a valuable role in understanding human diseases.[63,67,70] As is seen in Fig. 6-4, each phosphorus atom of ATP gives a distinct resonance. The area of the P_γ peak provides a direct measure of the ATP concentration, and phosphocreatine (creatine-*P*) and P_i can be estimated in a similar way. Knowing the concentrations of ATP, P_i, and creatine-*P*, the amount of ADP present, usually too low to be estimated by NMR, may be calculated.[61] Barnacle muscle contains, instead of phosphocreatine, a high concentration of phosphoarginine which has also been measured by NMR.[71]

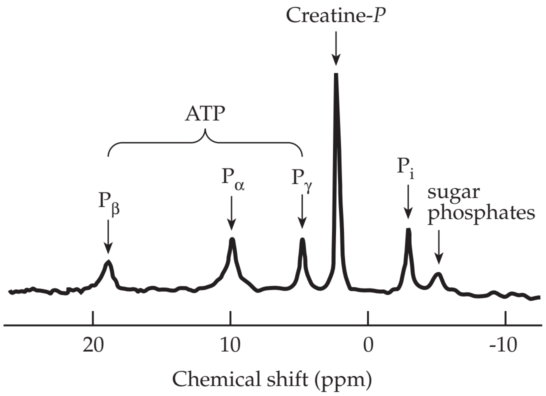


Figure 6-4 Phosphorus-31 NMR spectrum of an excised rat muscle (*vastus lateralis*) in Ringer solution at 15°C. The spectrum represents the accumulation of 400 scans. From P. J. Seeley *et al.*[64]

Usually no ADP can be seen in NMR spectra, although there may be about 0.5 mM ADP according to chemical analysis of rapidly frozen tissues. It has been concluded that most ADP is bound to proteins. In muscle the proteins myosin and actin (Chapter 20) hold most of the ADP leaving only about 0.02 mM free.[62,72] The same conclusion has been reached on the basis of other evidence.[55]

If a perfused heart in an NMR spectrometer is deprived of oxygen, the level of phosphocreatine drops rapidly and that of P_i increases while that of ATP remains relatively constant until the phosphocreatine is gone. Then it too falls and becomes undetectable after 17 min.[72] Similar changes occur when a tourniquet is placed on a human arm in the NMR spectrometer. Study of the rate and extent of recovery of the adenylate system when oxygen is readmitted is helping to provide a better means of protecting kidneys and other organs during transplantation operations. Diagnosis of circulatory ailments in human limbs by ^{31}P NMR may soon be routine. Use of radiofrequency coils placed on the body surface allows monitoring of the adenylate system in the heart, brain, and other tissues deep within the body.[69,73] Changes of concentration in the adenylate system and of phosphocreatine can be monitored very rapidly and evenly throughout the cardiac heartbeat cycle.[74]

E. Complex Biochemical Equilibria

The binding of small molecules to larger ones is basic to most biological phenomena. Substrates bind to enzymes and hormones bind to receptors. Metal ions bind to ATP, to other small molecules, and to metalloproteins. Hydrogen ions bind to amino acids, peptides, nucleotides, and most macromolecules. In this section we will consider ways of describing mathematically the equilibria involved.

The strength of bonding between two particles can be expressed as a **formation constant** (or **association constant)** K_f. Consider the binding of a molecule X to a second molecule P, which may be a protein or some other macromolecule. If there is on the surface of P only one single binding site for X, the process can be described by Eq. 6-68 and the equilibrim constant K_f by Eq. 6-69.

$$X + P \rightarrow PX \tag{6-68}$$

$$K_f = [PX] / [P][X] \tag{6-69}$$

The units for K_f are liters per mole (M^{-1}). The constant K_f is a direct measure of the strength of the binding: The higher the constant, the stronger the interaction. This fact can be expressed in an alternative way by giving the standard Gibbs energy change ($\Delta G°$) for the reaction (Eq. 6-70). The more negative $\Delta G_f°$, the stronger the binding.

$$\begin{aligned}\Delta G_f° &= -RT \ln K_f = -2.303RT \log K_f \\ &= -5.708 \log K_f \text{ kJ mol}^{-1} \text{ at } 25°\text{C}\end{aligned} \tag{6-70}$$

To avoid confusion, it is important to realize that the frequently used **dissociation constants** (K_d) are reciprocal association constants (Eq. 6-71). The use of

$$K_d = 1 / K_f \tag{6-71}$$

association constants and of dissociation constants is firmly entrenched in different parts of the chemical literature; be sure to keep them straight. Dissociation constants are customarily used to describe acid–base chemistry, while formation constants are more often employed to describe complexes with metal ions or associations of macromolecules (Section C). However, both types of equilibria can be described using either formation constants or dissociation constants.

Logarithms of formation constants, which are proportional to the Gibbs energies of association, are often tabulated. The logarithms of formation constants and pK_a values of dissociation constants are identical (Eq. 6-72) and are a measure of the standard

$$\log K_f = -\log K_d = pK_d \tag{6-72}$$

Gibbs energy decrease in the association reaction. The difference in $\Delta G°$ corresponding to a change of one unit in log K_f or pK_d is –5.7 kJ mol^{-1}, a handy number to remember.

1. Effects of pH on Equilibria

Compounds that contain several acidic or basic groups can exist in a number of different ionic forms, H_nA, $H_{n-1}A$, H_2A, HA, A, etc., as is indicated in the following equation, which is identical in form to Eq. 3-3.

$$H_nA \xrightarrow{K_1} H_{n-1}A \longrightarrow \cdots\cdots H_2A \xrightarrow{K_{n+1}} HA \xrightarrow{K_n} A \tag{6-73}$$

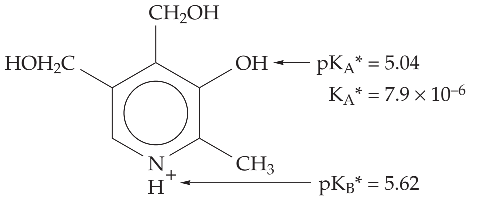
The dissociation constants K_1 ---- K_n for a multiprotic acid H_nA are defined as stepwise or **macroscopic constants** (also called molecular constants). For some compounds, e.g. alanine, the pK_a values are far apart (pK_1 and pK_2 are 2.4 and 9.8, respectively). The macroscopic constants can be assigned specifically, K_1 to the carboxyl group and K_2 to the protonated amino group. At the isoelectric pH of 6.1 the alanine exists almost entirely as the dipolar ion. However, for compounds in which the macroscopic pK_a values are closer together, they cannot be assigned to specific groups. We will consider some specific examples in the next section.

2. Microscopic Dissociation Constants and Tautomerization

A microscopic constant applies to a single site. Consider the dissociation of a simple carboxylic acid:

$$R{-}COOH \longrightarrow R{-}COO^- + H^+ \tag{6-74}$$

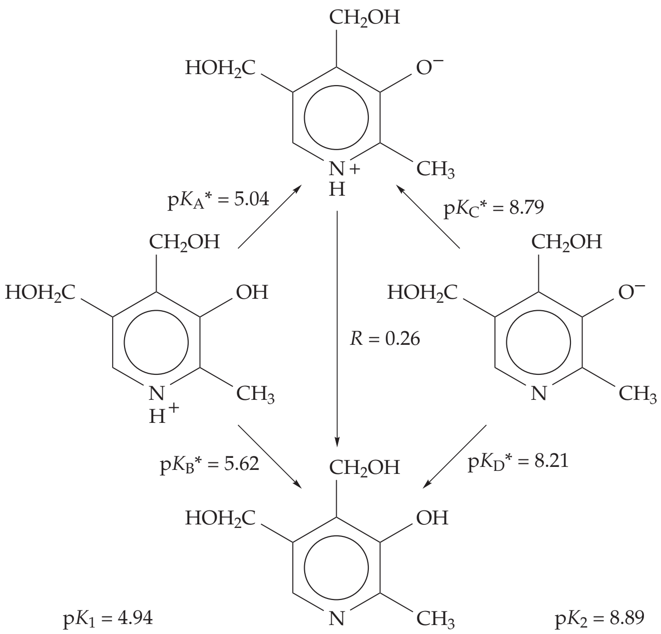
The dissociation constant is about 1.7 x 10^{-5} and pK_a = 4.8. Since there is only one proton, the observed dissociation constant is also the microscopic dissociation constant. Now consider the cation of pyridoxine which has two dissociable protons bound to distinctly different sites, the phenolic oxygen and the ring nitrogen.



Pyridoxine
(vitamin B_6)

Either of the two protons might dissociate first as the pH is raised (Eq. 6-75). However, the two microscopic dissociation constants pK_A* and pK_B* are distinctly different. The result is that at 25°C in the neutral (monoprotonated) form 80% of the molecules carry a proton on the N, while the other 20% are protonated on the less basic $-O^-$. Notice that the subscripts a and b used in this discussion do *not* refer to acidic and basic but to the individual dissociation steps shown in Eq. 6-75. Microscopic constants will always be indicated with asterisks in this discussion.

The two monoprotonated forms of pyridoxine are the tautomeric pair shown in Eq. 6-75 and whose concentrations are related by the **tautomeric ratio**, R = [neutral form]/[dipolar ion], a *pH-independent equilibrium constant* with a value of 0.204/0.796 = 0.26 at 25°C.[75] Evaluation of microscopic constants for dissociation of protons from compounds containing non-identical groups depends upon measurement of the tautomeric ratio, or ratios if more than two binding sites are present. In the case of pyridoxine, a spectrophotometric method was used to estimate R.



(6-75)

To calculate microscopic constants from stepwise constants and tautomeric ratios, consider Eq. 6-76 in which $[HP]_A$ and $[HP]_B$ are the concentrations of the two tautomers and K_1 is the first **stoichiometric** or **macroscopic** dissociation constant for the diprotonated species H_2P.

$$K_1 = \frac{([HP]_A + [HP]_B)[H^+]}{[H_2P]} = K_A{}^* + K_B{}^*$$

$$pK_1 = 4.94;\ K_1 = 1.15 \times 10^{-5} = 9.1 \times 10^{-6} + 2.4 \times 10^{-6} \tag{6-76}$$

We see that K_1 is just the sum of the two microscopic constants K_A^* and K_B^* for dissociation of H_2P to the pair of tautomers HP(A) and HP(B). In a similar fashion it can be shown that the second stoichiometric constant K_2 is related to the microscopic constants K_C^* and K_D^* for dissociation of HP(A) and HP(B) to form P (Eq. 6-77).

$$1 / K_2 = 1 / K_C^* + 1 / K_D^* \qquad (6\text{-}77)$$

Since the tautomeric ratio R equals $[HP]_B / [HP]_A$, Eqs. 6-76 and 6-77 can be rearranged to Eqs. 6-78 to 6-81. These allow the evaluation of all of the microscopic constants from the two stoichiometric constants K_1 and K_2 plus the tautomeric ratio R.

$$pK_A^* = pK_1 + \log(1 + R) \qquad (6\text{-}78)$$

$$pK_B^* = pK_A^* - \log R \qquad (6\text{-}79)$$

$$pK_C^* = pK_2 - \log(1 + R) \qquad (6\text{-}80)$$

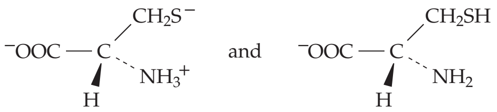
$$pK_D^* = pK_C^* + \log R \qquad (6\text{-}81)$$

For pyridoxine pK_1 and pK_2 were determined spectrophotometrically as 4.94 and 8.89. These values, together with that of R given above, were used to estimate the microscopic constants that are given in Eq. 6-74.[75] Notice that the microscopic constants of Eq. 6-74 are not all independent; if any three of the five equilibrium constants are known the other two can be calculated readily. In describing and measuring such equilibria it is desirable to select one pathway of dissociation, e.g., $H_2P \rightarrow HP(A) \rightarrow P$, and to relate the species HP(B) to it via the pH-independent constant R.

Often more complex situations arise in which additional tautomers or other forms arise via pH-independent reactions. These can all be related back to the reference ionic species by additional ratios R, which may describe equilibria for tautomerization, hydration, isomerization, etc. (Eq. 6-82).[76] In the case illustrated, only one of the ratios, namely R_2 or R_3, is likely to be a tautomerization constant because, as a rule, H_2P and P will not have tautomers. Equations analogous to Eqs. 6-76 to 6-82 can be written easily to derive K_C^*, K_D^* and any other microscopic constants desired from the stoichiometric constants plus the ratios R_1 to R_4. While it is easy to describe tautomerism by equations such as Eqs. 6-76 and 6-82 *it is often difficult to measure the tautomeric ratios R*.[77] In favorable cases measurements of spectra of one kind or another allow their evaluation. However, because tautomerism may be extremely rapid, NMR spectra will often show only one peak for a proton present in a mixture of tautomers.

As was pointed out in Chapter 2, tautomerization ratios are often affected strongly by changes in solvent. Tautomerism among monoprotonated forms of cysteine, glutathione, and histidine (Eq. 2-6) has received considerable attention by biochemists[77–79] as has tautomerism in binding of protons and other small ligands to proteins.[80,81] For cysteine,[78,79] for which the following species coexist in the alkaline pH range, the distribution of the various ionic species including the two tautomers is shown in Fig. 6-5. A similar situation is met in the small protein thioredoxin (Box 15-C) which has a buried aspartate carboxylate that interacts with a nearby cysteine –SH group,[82] in papain where –SH and imidazole groups interact (Fig. 12-15),[83] and in carbohydrases where two or more carboxyl and carboxylate side chains interact (Chapter 12).[84]

$$\begin{tikzcd}
H_2P(A) \arrow[rr, "K_a^*"] \arrow[d, "R_1"] & & HP(A) \arrow[rr, "K_b^*"] \arrow[dl, "R_2"'] \arrow[dr, "R_3"] & & P(A) \arrow[d, "R_4"] \\
H_2P(B) & HP(B) & & HP(C) & P(B)
\end{tikzcd} \qquad (6\text{-}82)$$



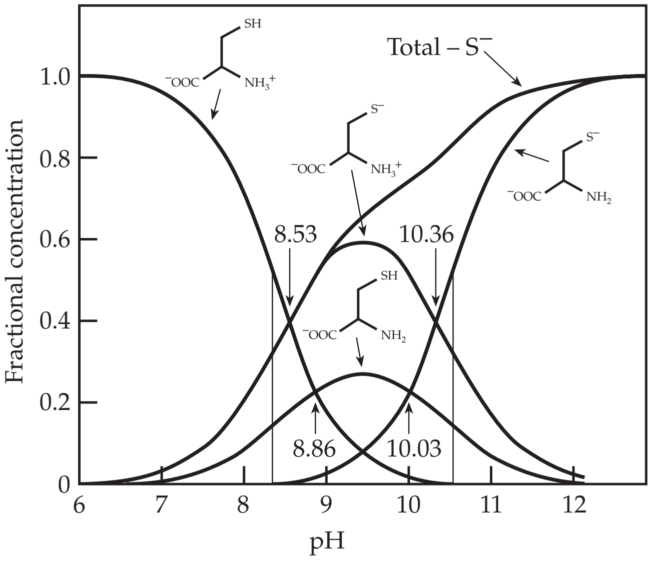
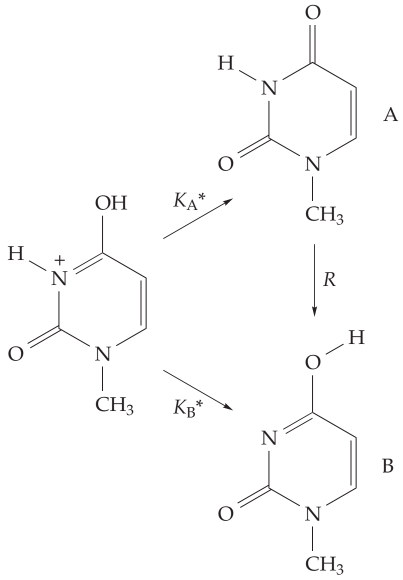


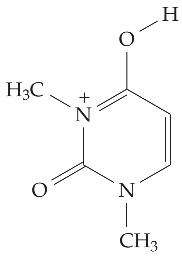
Figure 6-5 Distribution of various ionic species of cysteine as a function of pH. The function in each ionic and tautomeric form is plotted. Microscopic pK_a values are given numerically and macroscopic pK_a's are indicated by the vertical lines. From Dixon and Tipton.[79]

Because it is usually difficult to measure tautomeric ratios, dissociation constants are frequently assumed identical to those of compounds in which one of the acidic groups has been modified by methylation or esterification or in some other manner to prevent dissociation of a proton. For example, the cation of 1-methyluracil can be dissociated to the two tautomers shown in Eq. 6-83.

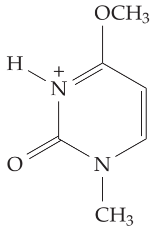


(6-83)

The apparent pK_a values were estimated for dissociation of the following two dimethylated cations which resemble that of Eq. 6-83.



pK_a = –3.25 ≈ pK_A*



pK_a = +0.65 ≈ pK_A*

It can reasonably be assumed that these pK_a values approximate pK_A* and pK_B* as indicated. Thus, applying Eq. 6-85, $\log R \approx -3.25 - 0.65 + 3.90$. This result indicates that tautomer A of Eq. 6-89 is overwhelmingly predominant in water.[85,86] It also suggests that, within experimental error, the pK_a for 1-methyluracil will equal pK_A*, namely, –3.25. In fact, it is close to this (–3.40).[85]

3. The Binding of Metal Ions

Equilibria in the formation of complex ions with metals are treated exactly as is the binding of small molecules and ions to macromolecules.[87–89] Stepwise constants are defined for the formation of complexes containing one, two, or more ligands L bound to a central metal ion M. The binding constants K_f's are usually referred to as β's as in Eq. 6-84.

$$\beta_1 = K_{f1} \qquad \beta_n = K_{f1}\, K_{f2} \ldots K_{fn} \tag{6-84}$$

Many important questions can be asked about the binding of metal ions within living cells. For example, What fraction of a given metal ion is free and what fraction is bound to organic molecules? To what ligands is a metal bound? Since many metal ions are toxic in excess, it is clear that homeostatic mechanisms must exist. How do such mechanisms sense the free metal ion activity within cells? How does the body get rid of unwanted metal ions? Answers to all these questions depend upon the quantitative differences in the binding of metal ions to the variety of potential binding sites found within a cell.

Table 6-9 gives formation constants for 1:1 complexes of several metal ions and a number of inorganic as well as organic ligands.[89] Only the values of log K_1 are given when a series of stepwise constants have been established. However, in many cases two or more ligands can bind to the same metal ion. Thus for cupric ion and ammonia there are four constants.

$$\mathrm{Cu^{2+} + NH_3} \qquad \log K_i = 4.0,\ 3.3,\ 2.7,\ 2.0$$

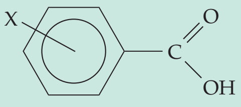
They are all separated by more than the "statistical distance," which in this case is less than the 0.6 logarithmic units for two equivalent binding sites (see Chapter 7, Section A,2). Thus, the second ligand binds less tightly than the first and **anticooperativity** in binding of successive ligands is observed. Most metal ions will also bind two or three successive amino acids. In the case of copper, whose preferred coordination number is four, two ligands may be bound. Again, a distinct anticooperativity is evident in the spread of the two constants.

$$\mathrm{Cu^{2+}} + \text{alanine} \qquad \log K_i = 8.1,\ 6.8$$

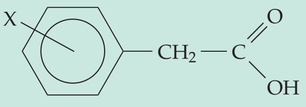
Factors affecting the strength of binding of a metal in a complex. More basic ligands tend to bind metal ions more tightly just as they do protons. However, the strength of bonding to metal ions to a group is more nearly proportional to the **nucleophilic character** (Chapter 12) which is only partly determined by basicity to protons.

BOX 6-C LINEAR GIBBS ENERGY RELATIONSHIPS

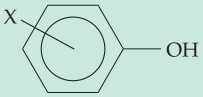
Organic functional groups exert characteristic electronic effects upon other groups to which they are attached. The quantitative expression of such effects can sometimes be correlated by linear Gibbs energy relationships. The best known of these is the **Hammett equation**, which deals with the transmission of electronic effects across a benzene or other aromatic ring. Consider the acid dissociation constants of three classes of compounds:



Benzoic acids



Phenylacetic



Phenol

The values of pK_a given in the following table have been observed for the parent compounds and for the *meta*-chloro- and the *meta*-nitro-substituted compounds.

Values of pK_a for Unsubstituted and Substituted Benzoic Acids, Phenylacetic Acids, and Phenols

Meta substituent	Benzoic acids*	Phenylacetic acids	Phenols
— H (parent compound)	4.202	4.31	9.92
— Cl	3.837	4.13	9.02
— NO_2	3.460	3.97	8.39

* For an independent set of pK_a values for substituted benzoic acids see Bolton and Fleming.[a] The values of σ calculated from them are slightly different from those given here.

Since pK_a is the negative logarithm of a dissociation constant, it follows from Eq. 6-30 that values of pK_a are directly proportional to values of $\Delta G°$ for dissociation of protons. In the Hammett treatment differences in pK_a values, rather than differences in $\Delta G°$, are considered. When a hydrogen atom in the meta position of benzoic acid is replaced by the electron-withdrawing Cl or NO_2, pK_a is lowered, i.e., the basicity of the conjugate base of the acid is decreased. The decrease in pK_a amounts to −0.365 for *m*-chlorobenzoic acid and −0.742 for *m*-nitrobenzoic acid. The Hammett treatment asserts that these changes in pK_a are a measure of the electron-withdrawing power of the meta substituent.[b] Thus, the nitro group is about twice as strong as the chloro group in this respect. The numerical values of these changes in the dissociation constant of benzoic acid define the **substituent constants** σ, which are used in the Hammett equation and are given in the following table. In this equation we use the symbol pK_0 to represent the pK_a of the unsubstituted parent compound and pK to represent the pK_a of the substituted molecule.

$$\text{For substituted benzoic acids: } pK_0 - pK = \sigma$$

The decreases in the pK of phenylacetic acid occasioned by replacement of the *meta*-hydrogen with Cl or NO_2 are −0.18 and −0.34, respectively, substantially less than for the benzoic acids. On the other hand, for the phenols the differences amount to −0.90 and −1.53, much *greater* than for the benzoic acids. The Hammett equation asserts that for reactions such as the dissociation of protons from phenylacetic acids or from phenols, the changes in ΔG occasioned by meta substitutions are proportional to the σ values, i.e., to the changes in ΔG for the standard reaction – dissociation of a proton from benzoic acid.[b–d]

Substituent Constants for Use in the Hammett and Taft Equations[b-d]

Substituent	σ_m	σ_p	σ_p^-	σ_p^+	σ* (Taft)
$-O^-$	−0.71	−0.52			
$-NH_2$	−0.16	−0.66		−1.11	0.62
$-OH$	0.121	−0.37		−0.85	1.34
$-OCH_3$	0.115	−0.27		−0.78	1.81
$-CH_3$	−0.069	−0.17		−0.31	0.00
$-NH-COCH_3$	0.21	−0.01		−0.25	
$-H$	0	0		0	0.49
$-CH_2OH$	0.08	0.08			0.56
$-COO^-$	−0.10	0.00		0.11	−1.06
$-SO_3$	$^{-}$0.05	0.09		0.12	
$-SH$	0.25	0.15		0.019	1.68
$-CH_2Cl$		0.18			1.05
$-CONH_2$	0.28	0.36			
$-F$	0.337	0.06	0.02	−0.07	
$-I$	0.352	0.18		0.135	
$-Cl$	0.373	0.227		0.114	
$-CHO$	0.36	0.22	0.37		
$-COCH_3$	0.376	0.502	0.85		
$-COOH$	0.37	0.45		0.42	2.08
$-COOCH_3$	0.39	0.31		0.49	
$-SO_2NH_2$	0.55	0.62		0.61	
$-CN$	0.56	0.66	0.89	0.66	
$-C\equiv CH$					2.18
$-CF_3$	0.43	0.54		0.61	2.61
$-CCl_3$					2.65
$-NO_2$	0.710	0.778	1.25	0.790	4.0
$-NH_3^+$	1.13	1.70			3.76
N (pyridine)	0.73	0.83			
NH^+ (pyridine)	2.18	2.42		4.0	

BOX 6-C (continued)

$$\log (K / K_0) = pK_0 - pK = \rho\sigma$$

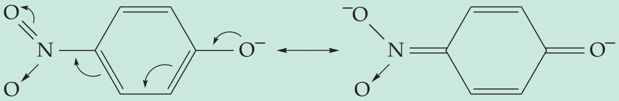
The Hammett equation

The proportionality constant ρ, which also appears in the equation, is a measure of the *sensitivity* of the reaction to the presence of electron-withdrawing or electron-donating substituents in the ring. For benzoic acid, ρ is taken as 1.00. Using the data from the accompanying table together with many other data, an average value of $\rho = 0.49$ has been found for phenylacetic acids. Likewise, $\rho = 2.23$ for phenols, and $\rho = 5.7$ for dissociation of protons from substituted pyridinium ions. The sensitivity to substituent changes is highest (highest ρ) in the latter case where proton dissociation is directly from an atom in the ring and is lowest when the basic center is removed farthest from the ring (phenylacetic acid).

Through knowledge of the value of ρ for a given reaction, it is possible to predict the effect of a substituent on pK_a using the tabulated values of σ. In many cases the effects are additive for multisubstituted compounds.

$$pK_a = pK_0 - \rho\Sigma\sigma$$

Since substituents in *o*, *m*, and *p* positions have quantitatively different influences, different substituent constants σ are defined for each position. Moreover, since special complications arise from ortho interactions, it is customary to tabulate σ values only for meta and para positions. These are designated σ_m and σ_p. Apparent σ values for *ortho* substituents are also available.[d] An additional complication is that certain reactions are unusually sensitive to para substituents that are able to interact by resonance directly across the ring. An example is the acid dissociation of phenols. While σ_p for the nitro group is ordinarily 0.778, a correct prediction of the effect of the *p*-nitro group on dissociation of phenol is given only if σ_p is taken as 1.25. This higher σ value is designated σ_p^-. The resonance in the phenolate anion giving rise to this enhanced effect of the nitro group may be indicated as follows:



(Here the curved arrows represent the direction of flow of electrons needed to convert from the one resonance structure to the other.) For similar reasons, some reactions require the use of special σ_p^+ constants for strongly electron-donating substituents such as OH which are able to interact across the ring by resonance. Thus σ_p for the OH group is -0.37, while σ_p^+ is -0.85.[e] For the methoxyl group (—OCH_3) $\sigma_p = -0.27$, $\sigma_p^+ = -0.78$, and $\sigma_m = +0.12$.

The use of different kinds of substituent constants complicates the application of the Hammett equation and over 20 different sets of σ values have been proposed. A simplification is the representation of substituent constants as linear combinations of two terms, one representing "field" or "inductive" effects and the other resonance effects.[e,f]

Many other linear Gibbs energy relationships have been proposed; for example, the acid strengths of aliphatic compounds can be correlated using the "Taft polar substituent constants" σ*.

$$\log (K/K_0) = \rho^* \sigma^*$$

For example, the following give good approximations of pK_a values.[d]

for R—COOH $pK_a = 4.66 - 1.62\ \sigma^*$

for R—CH_2—COOH $pK_a = 5.16 - 0.73\ \sigma^*$

While the examples chosen here concern only dissociation of protons, the Hammett equation has a much broader application. Equilibria for other types of reactions can be treated. Furthermore, since rates of reactions are related to Gibbs energies of activation, many rate constants can be correlated. For these purposes the Hammett equation can be written in the more general form in which *k* may be either an equilibrium constant or a rate constant.[b] The subscript *j* denotes the reaction under consideration and *i* the substituent influencing the reaction.

$$\log k_{ij} - \log k_{0j} = \rho_i \sigma_j$$

An example of a linear Gibbs energy relationship that is widely used in discussing mechanisms of enzymatic reactions is the Brönsted plot (Eqs. 9-90 and 9-91).

---

[a] Bolton, P. D., Fleming, K. A., and Hall, F. M. (1972) *J. Am. Chem. Soc.* **94**, 1033–1034

[b] Hammett, L. P. (1970) *Physical Organic Chemistry*, 2nd ed., p. 356, McGraw-Hill, New York

[c] Wells, P. R. (1963) *Chem. Rev.* **63**, 171–219

[d] Barlin, G. B., and Perrin, D. D. (1966) *Q. Rev., Chem. Soc.* **20**, 75–101

[e] Swain, C. G., and Lupton, E. C., Jr. (1968) *J. Am. Chem. Soc.* **90**, 4328–4337

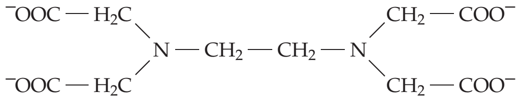
[f] Hansen, L. D., and Hepler, L. G. (1972) *Can. J. Chem.* **50**, 1030–1035

TABLE 6-9
Logarithms of Binding Constants for Some 1:1 Metal Complexes at 25°C[a]

Ligand	H^+	Mg^{2+}	Ca^{2+}	Mn^{2+}	Cu^{2+}	Zn^{2+}
Hydroxide, OH^-	14.0	2.5	1.4		6.5	4.4
Acetate$^-$	4.7	~0.65	0.5	~1.0	2.0	1.5
Lactate$^-$	3.8	~1.0	~1.2	1.3	3.0	2.2
Succinate^{2-}	5.2	1.2	1.2		3.3	2
NH_3	9.3	~0	–0.2	0.8	4.0	2.4
Ethylenediamine	10.2	0.4		2.8	10.8	6.0
Glycine$^-$	9.6	2.2	1.4*	2.8	8.2	5.0
Glycine amide	8.1			~1.5	5.3	3.3
Alanine$^-$	9.7	2.0*	1.2*	3.0*	8.1	4.6
Aspartate^{2-}	9.6	2.4	1.6		8.6	5.8
Glycylglycine$^-$	8.1		1.2*	2.2*	5.5	3.4
Pyridine	5.2			0.1	2.5	1
Imidazole	7.5			1.6	4.6	2.6
Histidine	9.1			3.3	10.2	6.6
Adenine	9.8				8.9	6.4
Citrate^{3-}	5.6	3.2	4.8	3.5	~4	4.7
EDTA^{4-} [b]	10.2	8.8	10.6	13.8	18.7	16.4
EGTA^{4-} [c]	9.4	5.3	10.9	12.2	17.6	12.6
ATP^{4-}	6.5	4.2	3.8	4.8	6.1	4.9

[a] All values are for log K_1. Included is the highest pK_a for protons. Data for amino acids are from Martell, A.E. and Smith, R.M. (1974, 1975) *Critical Stability Constants*, 3 vols., Plenum, New York. Others are from Sillén, L.G. and Martell, A.E. (1964) *Stability Constants of Metal–Ion Complexes*, Spec. Publ. No. 17, Chemical Society, London. Most constants for amino acids are for ionic strength 0.1. Some (designated by asterisks) are for zero ionic strength. The values shown for other ligands have been selected from a large number reported without examination of the original literature.

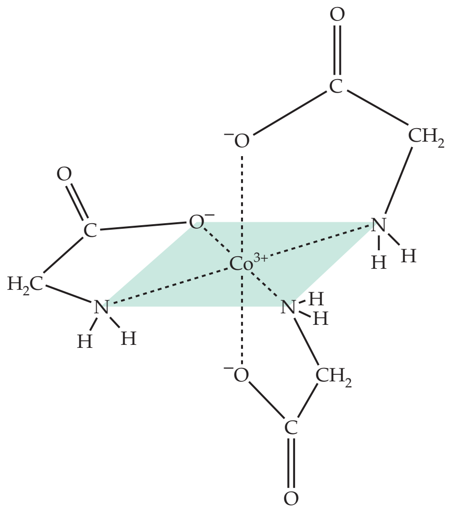
[b] Ethylenediaminetetraacetic acid (EDTA), a chelating agent widely used by biochemists for preventing unwanted reactions of metal ions. The high formation constants ensure that most metal ions remain bound to the EDTA.

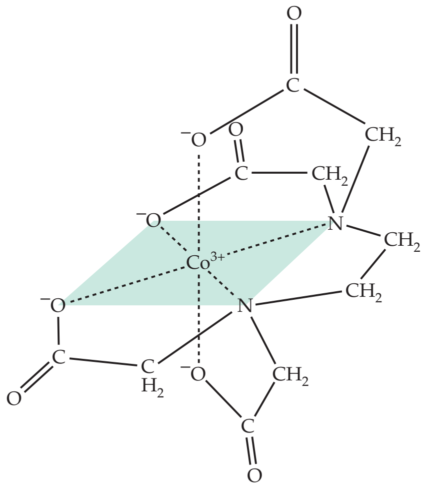


[c] EGTA is similar to EDTA but has the group $-CH_2-CH_2-O-CH_2-CH_2-O-CH_2-CH_2-$ joining the two nitrogen atoms in place of $-CH_2-CH_2-$ of EDTA. Note that EGTA has a higher selectivity for Ca^{2+} compared to Mg^{2+} than does EDTA.

The pH of the medium always has a strong effect on metal binding. Competition with protons means that metal complexes tend to be of weak stability at low pH. Anions of carboxylic acids are completely protonated below a pH of ~4 and a metal can combine only by displacing a proton. However, at pH 7 or higher, there is no competition from protons. On the other hand, in the case of ethylenediamine, whose pK_a values are 10.2 and 7.5 (Table 6-9), protons are very strong competitors at pH 7, even with a strongly complexing metal ion such as Cu^{2+}. At high pH there may be competition between the ligand and hydroxyl ion. At pH 7 about one-half of Cu^{2+} dissolved in water is complexed as $CuOH^+$. Aluminum forms soluble complexes $Al(OH)^{2+}$, $Al(OH)_2^+$, and $Al(OH)_4^+$ (Box 12-F).

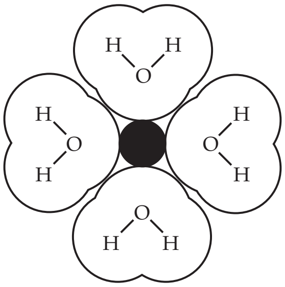
One of the most important factors in determining the affinity of organic molecules for metal ions is the **chelate effect**. The term chelate (pronounced "keel-ate") is from a Greek word meaning crab's claw. It refers to the greatly enhanced binding of metal ions resulting from the presence of two or more complexing groups in the same organic molecule. The chelate effect has been exploited by nature in the design of many important metal-binding molecules, including porphyrins (Fig. 16-5), chlorophyll (Fig. 23-20), the siderophores (Fig. 16-1), and metal-binding proteins. Structures of two **chelate complexes** are shown here. Notice from Table 6-9 that many simple compounds, such as the α-amino acids and citric acid, often form strong chelate complexes with metal ions.





How properties of the metal ion affect chelation. The **charge**, the **ionic radius** (Table 6-10), the **degree of hydration**, and the **geometry of orbitals** used in covalent bonding between metal and chelating groups all affect the formation constants of a complex. Multicharged ions form stronger complexes than do monovalent ions, which have a lower charge density.

Among ions of a given charge type (e.g., Na^+ vs K^+; Mg^{2+} vs Ca^{2+}), the smaller ions are more strongly hydrated than are the larger ions in which the charge is dispersed over a greater surface area. Most cations, except for the large ones, have a primary hydration sphere containing about six molecules of water. Four molecules of water can be placed around the ion in one plane as shown in the following drawing of water molecules coordinated to Mg^{2+}.



One additional water molecule can be bound above and another below to provide six molecules in an array of **octahedral** geometry. An alternative tetrahedral arrangement of four molecules of water around the ion has been suggested for Li^+ and Na^+.[90] In either case additional solvent molecules are held in a looser secondary sphere. For instance, electrochemical transference experiments indicate a total of ~16 molecules of water around Na^+ and ~10 around K^+.

TABLE 6-10
Ionic Radii in Nanometers for Some Metallic and Nonmetallic Ions[a]

		Mn^{2+}	0.080		
Li^+	0.060	Fe^{2+}	0.076	H^-	0.21
Na^+	0.095	Co^{2+}	0.074	F^-	0.136
K^+	0.133	Ni^{2+}	0.069	Br^-	0.195
Rb^+	0.148	Cu^{2+}	0.072[b]	I^-	0.216
		Zn^{2+}	0.074		
		Cd^{2+}	0.097		
Be^{2+}	0.031				
Mg^{2+}	0.065	Al^{3+}	0.050		
Ca^{2+}	0.099	Fe^{3+}	0.064		
Sr^{2+}	0.113	Mo^{4+}	0.070		
Ba^{2+}	0.135	Mo^{6+}	0.062		

[a] Radii are calculated according to the method of Pauling and are taken from Cotton, F.A. and Wilkinson, G. (1972) *Advanced Inorganic Chemistry*, 3rd ed. Wiley (Interscience), New York .

[b] From Ahrens, L. H. as given by Sienko, M. J. and Plane, R. A. (1963) *Physical Inorganic Chemistry*, pp. 68–69. Benjamin, New York.

To form a chelate complex a metal ion must usually lose most of its hydration sphere. For this reason, the larger, less hydrated metal ions often bind more strongly than do the smaller, more hydrated ones. For example, Ca^{2+} binds to EDTA more tightly than does Mg^{2+} (Table 6-9). However, the reverse order may be observed with negatively charged ligands such as OH^- in which the charge in highly concentrated. The same is true for ATP^{4-}, which binds Mg^{2+} more tightly than Ca^{2+} (Table 6-9; Section B,5).

Differences in both the charge density and the hydration of ions determine the **Hofmeister series** (lyotropic series).[91]

$$Ba^{2+} > Ca^{2+} > Mg^{2+} > Li^+ > Na^+ > K^+ > Cs^+ > NH_4^+$$

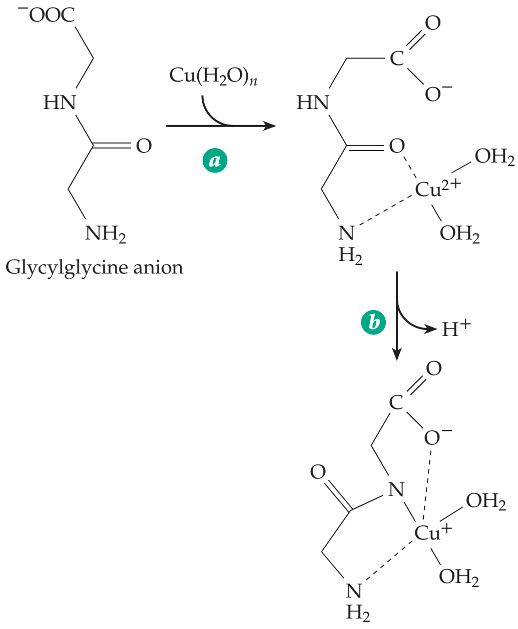
This was originally defined as the order of effectiveness in precipitating colloids or protein molecules. The ions to the left are less hydrated than those to the right. A similar series can be defined for anions. The following, a well-known sequence of the stabilities of complexes of metals of the first transition series, applies to many different types of complexes including those of the α-amino acids, as is shown in Fig. 6-11.

$$Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$$

Simple electrostatic theory based upon differences in the ionization potentials or electronegativity of the ions would predict a gradual monotonic increase in chelate stability from manganese to zinc. In fact, with nitrogen-containing ligands Cu^{2+} usually forms by far the strongest complexes. Cobalt, nickel, and iron ions also show an enhanced tendency to bind to nitrogen-containing ligands. The explanation is thought to lie in the ability of the transition metals to supply *d* orbitals which can participate in covalent bond formation by accepting electrons from the ligands. Iron, copper, and cobalt are often located in the centers of nitrogen-containing structures such as the heme of our blood (iron, Fig. 16-5) and vitamin B_{12} (Box 16-B).

Metal binding sites in cells. Functional groups that participate in metal binding include negatively charged carboxylate $-COO^-$, thiolate $-S^-$, phenolate $-O^-$, and phosphate$^-$ as well as uncharged amino, imidazole, –OH, and the polarizable carbonyl groups of peptide and amide side chains. To which of these ligands will specific ions tend to bind? The alkali metal ions Na^+ and K^+ are mostly free within cells

(Box 5-A) but are in part bound to defined sites in proteins. Both Ca^{2+} and Mg^{2+} tend to remain partially free and complexed with the numerous phosphate and carboxylate ions present in cells. However, they may find very specific tight binding sites such as that of Mg^{2+} in chlorophyll (Fig. 23-20). The heavier metal ions, including those of zinc, copper, iron, and other transition metals, usually bind to nitrogen or sulfur atoms.[92] For example, small peptides react with Cu^{2+} to form chelate complexes in which the peptide carbonyl oxygen binds to the metal (Eq. 6-85, step *a*).[93,94] By losing a proton the peptide NH can sometimes also function as a metal ligand (Eq. 6-85, step *b*).



(6-85)

In many **metalloproteins** the metals are found in **prosthetic groups** such as the porphyrin of heme proteins and the molybdopterin of molybdenum containing enzymes (Fig. 16-31). Very often clusters of carboxylate, imidazole, and other groups are used to create binding sites. In carboxypeptidase A (Fig. 12-16) two imidazole groups and one carboxylate of a glutamate site chain hold Zn^{2+}. In carbonic anhydrases, three imidazoles hold a zinc ion (Fig. 13-1), while in one kind of superoxide dismutase both Zn^{2+} and Cu^{2+} are bound in adjacent locations with six imidazoles and one carboxylate group participating in the binding. In contrast, Zn^{2+} and Cd^{2+} are bound in metallothioneins by clusters of thiol groups from cysteine side chains (Box 6-E). While Fe^{2+} is bound by from four to six nitrogen atoms in heme proteins, it is also found attached to oxygen atoms of tyrosine side chains, as is shown for a transferrin in Fig. 16-2. One imidazole, one carboxylate group, and a bicarbonate ion also bind to the iron. This site is also quite satisfactory for Al^{3+} (Box 12-F), which tends to occupy a fraction of the transferrin sites in blood, although it binds much more weakly than does Fe. Several other binding sites for transition metals are pictured in Chapter 16.

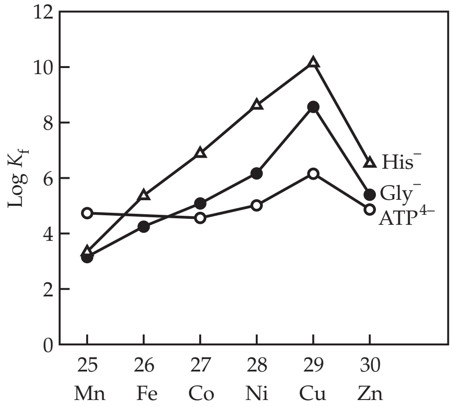


Figure 6-6 Logarithms of formation constants of metal complexes of histidine, glycine, and ATP plotted against the atomic number of elements from manganese to zinc.

Calcium-binding proteins. Much information about the functions of transition metal ions is given in Chapter 16. Here and in Box 6-D we consider calcium ions which, because of their broad distribution and range of functions, will be discussed in virtually every chapter of this book. The concentration of Ca^{2+} varies greatly in different parts of a cell and also with time (Box 6-D). Many **calcium-binding proteins** participate in mediating the physiological effects of these changes and also in buffering the calcium ion concentration. These include a large family of **helix–loop–helix** or **EF-hand proteins**.[95–97] The first of these to be discovered was **parvalbumin**, a 108-residue protein from carp muscle.[96,98-99a] The structure of the pair of metal-binding sites of the protein is shown in Fig. 6-7.

Each consists of two helices that are almost perpendicular and connected by a loop that forms the Ca^{2+}-binding site. In the site at the left side in Fig. 6-7A the Ca^{2+} is bound by four carboxylate groups from aspartate and glutamate side chains, a hydroxyl group of serine, and the residue 57 carbonyl oxygen of the peptide backbone.[99b] The same peptide group is hydrogen bonded to a carbonyl group of another segment of peptide chain near the second Ca^{2+} site (to the right in Fig. 6-7A). This site contains four carboxylate ions, one of which coordinates the Ca^{2+} with both oxygen atoms, and another peptide carbonyl group. By attaching itself to several different side chain groups, the metal ion can induce a substantial change in conformation from that present in the calcium-free protein.[97]

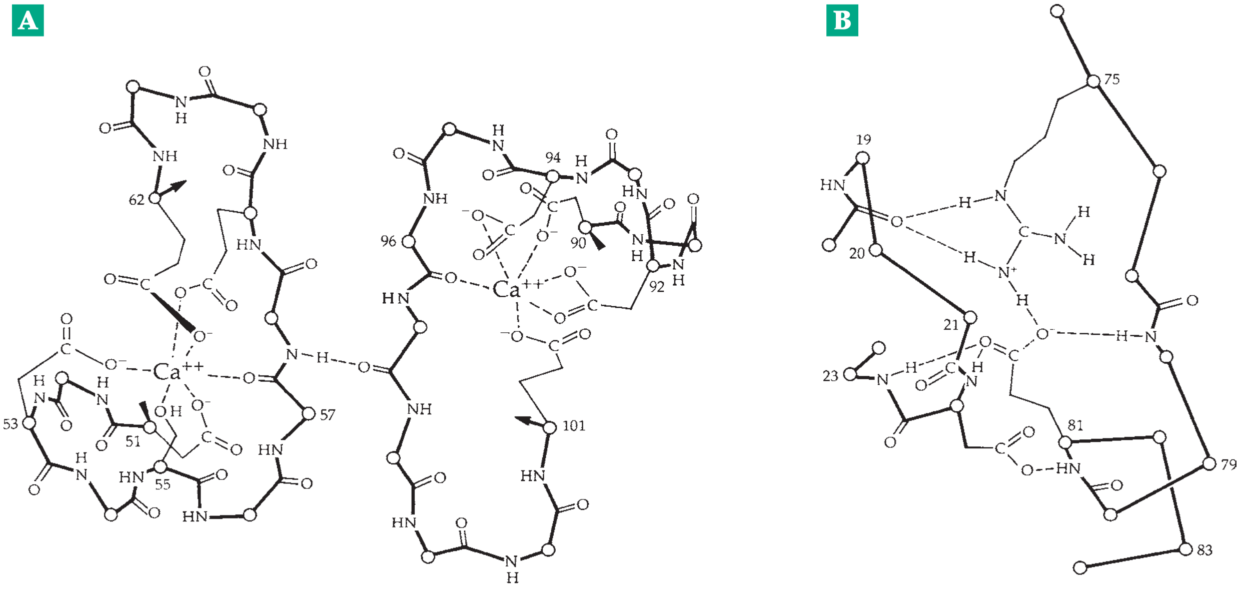


Figure 6-7 (A) Part of the 108-residue peptide chain of the calcium-binding protein of carp muscle. The two calcium-binding loops are shown together with a hydrogen bond between them. (B) A view of the intricate network of hydrogen bonds linking two segments of the peptide chain in the interior of the molecule. Note especially the bonding of the guanidinium group from arginine-75 to the carboxylate of glutamic acid 81 and to the peptide carbonyl of residue 18. Note that the carboxylate group also interacts with two different peptide NH groups. From Kretsinger and Nockolds.[98]

Kretzinger, who discovered the structure of parvalbumin, named the Ca^{2+}-binding helix–loop–helix motif an **EF-hand** because it is formed by helices E and F and resembles a hand with pointer finger extended along the E helix and the thumb in the direction of the F helix, the flexed middle finger forming the Ca^{2+}-binding loop. Helices C and D also form a hand. The resulting pair of hands can be visualized better at the top of the calmodulin structure in Fig. 4-8. The EF-hand motif has been identified in over 1000 proteins.[100]

Parvalbumins, which are also found in other vertebrates, are high-affinity Ca^{2+}-buffers.[99] Additional calcium buffers with EF-hand structures are the vitamin D-induced **calbindins**. One 9-kDa calbindin is found in mammalian intestinal tissue and in skin. It has two helix–loop–helix Ca^{2+}-binding sites of differing affinity[101,102] that presumably function in the absorption of calcium. A 28-kDa vitamin D-dependent protein from chicken intestine contains six similar Ca^{2+}-binding loops.[97,103]

Another group of calcium-buffering and storage proteins with remarkable Ca^{2+}-binding properties are the 40- to 45-kDa **calsequestrins**, which are found in the lumen of the ER (sarcoplasmic reticulum) of skeletal muscle. Calsequestrins are not typical EF-hand proteins but have a high content of glutamate and aspartate. They bind ~50 Ca^{2+} per molecule of protein with K_d ~ 1 mM.[104,105] Similar proteins called **calreticulins** are found in most non-muscle cells.[106,107]

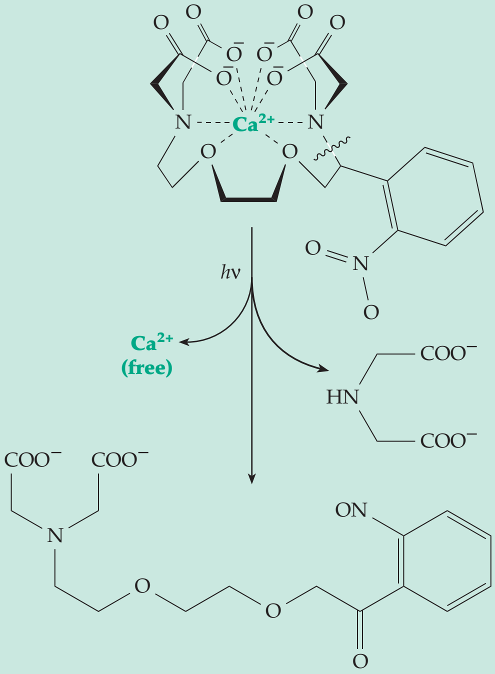
A very large number of EF-hand proteins have signaling functions. The best known of these is the 148-residue **calmodulin**, which regulates many enzymes and cellular processes (Box 6-D; also Chapter 11).[108,109] The protein, which is present in all eukaryotes, has a conserved sequence that forms two pairs of helix–loop–helix Ca^{2+}-binding sites that are separated by a long helix (Fig. 6-8).[108,110,111] Two of the sites bind Ca^{2+} tightly and cooperatively,[110,112] with K_d values in the micromolar range. Calmodulin from almost all species contains the modified amino acid ε-*N*-trimethyllysine at position 115. However, octopus calmodulin has ordinary lysine at this position and seems to function well.[108] Calmodulin's controlling functions result from Ca^{2+}-induced conformational changes that modify its affinity for other proteins whose activity may be increased or decreased by bound calmodulin.[109,113] A protein with a similar dumbell shape and structure is **troponin C** of skeletal muscles.[114,115] Troponin C binds to a complex of proteins that assemble on the thin actin filaments of muscle fibers and control contraction in response to changes in the calcium ion concentration (Chapter 19).[116] Other proteins that contain EF-hand motifs and are therefore responsive to Ca^{2+} include **spectrin** of cell membranes,[117] **clathrin light chains** from coated vesicles,[118,119] the extracellular **osteonectin** of bones and teeth,[120] and a birch pollen antigen.[121] Another group of 17 or more small **S100 EF-hand** proteins play a variety of other roles.[122–123a] One of these, which has a high affinity for Zn^{2+}, has been named **psoriasin** because of its 5-fold or greater

BOX 6-D CALCIUM

The essentiality of calcium ions to living things was recognized in the last century by S. Ringer, who showed that ~1 µM Ca^{2+} was needed to maintain the beat in a perfused frog heart. Later, calcium was shown essential for repair of ruptures in the cell membrane of the protozoan *Stentor* and for the motion of amebas. The animals quickly died in its absence. The role in the frog heart was traced to transmission of the nerve impulse from the nerve to the heart muscle. More recently it has been recognized that Ca^{2+} is required for blood clotting and is involved in triggering many responses by cells. Calcium ions are also an integral part of many enzymes and have structural roles in proteins, carbohydrate gels, and biological membranes.

Like Na^+, the calcium ion is actively excluded from cells. Indeed, 99% of the calcium in the human body is present in the bones.[a–d] The blood serum concentration of Ca^{2+} is ~3 mM, of which ~1.5 mM is free. The rest is chelated by proteins, carbohydrates, and other materials. Within cells the concentration of free Ca^{2+} is < 1 µM and typically ~0.05–0.2 µM for unexcited cells.[d–f] However, the total intracellular Ca^{2+} is considerably higher and may be in excess of 1 mM. Approximate total concentrations are: red blood cells, 20 µM; liver, 1.6 mM; and heart, 4 mM. A gradient in $[Ca^{2+}]$ of 10^3 or more is maintained across membranes by the calcium ion pump (Chapter 8). The action of this pump is counteracted by a very slow diffusion of the external Ca^{2+} back through the membrane via an Na^+–Ca^{2+} exchange into the cells.[g]

Free Ca^{2+} lacks spectroscopic properties suitable for its direct measurement at the low concentrations present in cells. However, it can be measured indirectly by the use of chelating agents that are relatively specific for Ca^{2+} and which have a measurable property that changes upon calcium-binding. The fluorescent photoprotein **aequorin** (Chapter 23), which may be injected into cells or synthesized within cells from transferred genes is often employed.[h,i] Various synthetic Ca^{2+} indicators also fluoresce brilliantly upon chelation.[j–n] Others contain fluorine or a suitably placed atom of ^{13}C which changes its NMR chemical shift upon chelation with Ca^{2+}.[o] These compounds may be carried into cells in the form of esters which pass through membranes but are then hydrolyzed leaving the indicators trapped in the cytoplasm.[j] Chelate compounds that bind Ca^{2+} within cells and release it rapidly upon irradiation with ultraviolet light have also been developed.[p,q]



Calcium chelate compound that releases free calcium ion within cells upon irradiation with ultraviolet light.[p,q]

Consistent with their role in signaling, calcium ions are unevenly distributed within cells. Mitochondria, endoplasmic reticulum, Golgi, and nuclei may all take up calcium ions. Cytoplasmic Ca^{2+} may sometimes be sequestered in microvesicles (**calciosomes**) or in intracellular granules[c,r,s] in which the $[Ca^{2+}]$ is ~0.5–1 mM but may reach 5–10 mM.[f] Many regulatory mechanisms exist. For example, **calcitonin** and **parathyroid hormone** interact with **vitamin D** and its metabolites in the small intestine, bones, and kidneys to control the deposition of calcium in bones, a topic considered in Box 22-C.

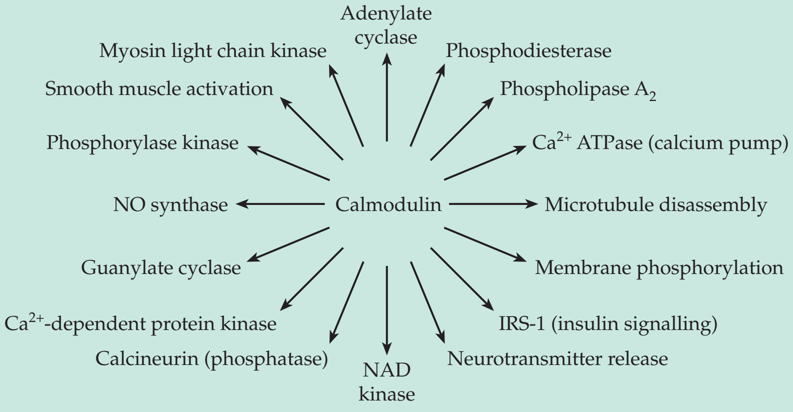
A characteristic function of Ca^{2+} in living things is **activation** of various metabolic processes. This occurs when a sudden change in permeability of the plasma membrane or in the membranes of the endoplasmic reticulum (ER) allows Ca^{2+} to diffuse into the cytoplasm. For example, during the contraction of muscle, the Ca^{2+} concentration rises from ~0.1 to ~10 µM as a result of release from storage in the calciosomes of the ER. The calcium ions bind to **troponin C** initiating muscle contraction (Chapter 19).[t] The ER membranes of muscle are rich in a Ca^{2+} pump protein[u,v] (Fig. 8-26), and in a series of calcium-binding proteins such as **calsequestrin** (see text).[f] Their combined action soon restores the $[Ca^{2+}]$ to the original low value.

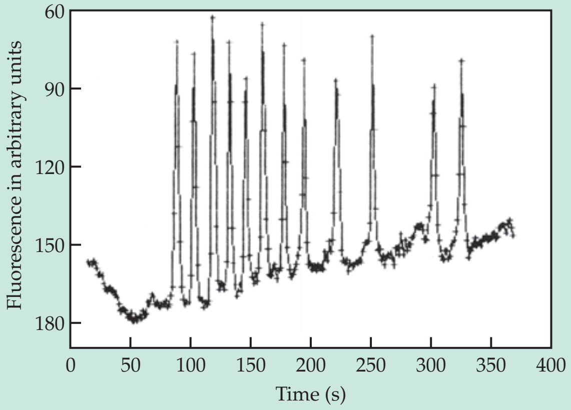
Skeletal muscle is activated by nerve impulses which induce Ca^{2+} release through the action of a **voltage sensor**, a protein also known as the **dihydropyridine receptor**,[w,x] together with a **calcium release channel** known as the **ryanodine**

BOX 6-D (continued)

receptor.[w,y–aa] Inositol trisphosphate (Fig. 11-9), cyclic ADP-ribose, as well as pH changes[bb] are involved in controlling these channels,[cc,dd] a topic discussed in Chapter 11, E. The release of Ca^{2+} can be visualized using high-speed digital imaging microscopy and fluorescent [Ca^{2+}] indicators.[ee,ff] Release of Ca^{2+} stored in sea urchin eggs is induced by cyclic ADP-ribose and by nicotinic acid adenine dinucleotide phosphate (NAADP$^+$, Eq. 15-16).[dd,gg]

In a similar manner, when a nerve impulse reaches a neuron ending (synapse) calcium ions are released into the cytoplasm and provide the trigger that causes stored neurotransmitters to be dumped into the narrow synaptic cleft that separates the endings of two communicating neurons.[y,hh] The released neurotransmitter initiates an impulse in the "postsynaptic neuron" usually again with an inflow of Ca^{2+} (see Chapter 30).[e] Hormones and various other compounds often stimulate the flow of calcium ions into cells.[ii] There Ca^{2+} regulates enzymes[jj,kk] (Chapter 11), microtubules, clathrin of coated vesicles (Chapter 8), K$^+$ channels in nerve membranes, and events within mitochondria,[ll,mm] in the ER,[op] and in the nucleus.[nn,oo] A substantial fraction of these responses are mediated by **calcium-binding regulatory proteins**. Among the most prominent of these is **calmodulin** (Fig. 6-8), which, upon binding of Ca^{2+} activates a host of metabolic processes[pp] as indicated in the following scheme, which is modified from that of Cheung.[qq]


Calmodulin
Adenylate cyclase
Phosphodiesterase
Phospholipase A$_2$
Ca^{2+} ATPase (calcium pump)
Microtubule disassembly
Membrane phosphorylation
IRS-1 (insulin signalling)
Neurotransmitter release
NAD kinase
Calcineurin (phosphatase)
Ca^{2+}-dependent protein kinase
Guanylate cyclase
NO synthase
Phosphorylase kinase
Smooth muscle activation
Myosin light chain kinase



Calcium oscillations observed with six cultured pancreatic β cells after a single infusion of 0.2 mM carbamoylcholine. The fluorescence intensity of the Ca^{2+} indicator dye fura 2, with excitation at 380 nm, was recorded versus time. From Pretki *et al.*[ss]

In many cases metabolic control by Ca^{2+} is modulated by phosphorylation and dephosphorylation or other covalent modification of proteins. Such modification may alter the affinity for Ca^{2+}, allowing the latter to either bind and induce a conformational change or remain unbound,[rr] without a change in [Ca^{2+}]. Nevertheless, the free Ca^{2+} concentration within cells can change greatly and very rapidly. For example, the oscillatory change in intracellular [Ca^{2+}] shown above was observed in pancreatic insulin-secreting β cells responding to stimulation by the agonist carbamoylcholine. The free [Ca^{2+}] was evaluated from fluorescence measurements using the Ca^{2+} indicator dye fura 2 (From Prentki *et al*[ss]). Oscillations in [Ca^{2+}] have been observed under many circumstances.[ff,oo,tt–ww] These are of particular interest because of the possible relationship to the initiation of oscillatory nerve conduction (Chapter 30).[uu] Released Ca^{2+} often appears to move across cells in waves and sometimes to be released as "puffs" or "sparks."[ee,ff,oo]

---

[a] Bianchi, C. P. (1968) *Cell Ca^{++}*, Appleton, New York

[b] Means, A. R., ed. (1994) *Calcium Regulation of Cellular Function*, Lippincott-Raven, Hagerstown, Maryland

[c] Volpe, P., Krause, K.-H., Hashimoto, S., Zorzato, F., Pozzan, T., Meldolesi, J., and Lew, D. P. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1091–1095

[d] Carafoli, E. (1987) *Ann. Rev. Biochem.* **56**, 395–433

[e] Ghosh, A., and Greenberg, M. E. (1995) *Science* **268**, 239–247

[f] Meldolesi, J., and Pozzan, T. (1998) *Trends Biochem. Sci.* **23**, 10–14

[g] Khananshvili, D. (1991) *J. Biol. Chem.* **266**, 13764–13769

(continued)

BOX 6-D (continued)

[h] Rizzuto, R., Simpson, A. W. M., Brini, M., and Pozzan, T. (1992) *Nature (London)* **358**, 325–327

[i] Knight, M. R., Campbell, A. K., Smith, S. M., and Trewavas, A. J. (1991) *Nature (London)* **352**, 524–526

[j] Tsien, R. Y., Pozzan, T., and Rink, T. J. (1984) *Trends Biochem. Sci.* **9**, 263–266

[k] Davies, P. L., and Hew, C. L. (1990) *FASEB J.* **4**, 2460–2468

[l] Ganz, M. B., Rasmussen, J., Bollag, W. B., and Rasmussen, H. (1990) *FASEB J.* **4**, 1638–1644

[m] Lloyd, Q. P., Kuhn, M. A., and Gay, C. V. (1995) *J. Biol. Chem.* **270**, 22445–22451

[n] Etter, E., Minta, A., Poenie, M., and Fay, F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5368–5373

[o] Robitaille, P.-M. L., and Jiang, Z. (1992) *Biochemistry* **31**, 12585–12591

[p] Ellis-Davies, G. C. R., and Kaplan, J. H. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 187–191

[q] Ellis-Davies, G. C. R., Kaplan, J. H., and Barsotti, R. J. (1996) *Biophys. J.* **70**, 1006–1016

[r] Burgoyne, R. D., and Cheek, T. R. (1991) *Trends Biochem. Sci.* **16**, 319–320

[s] Golovina, V. A., and Blaustein, M. P. (1997) *Science* **275**, 1643–1648

[t] Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry* **36**, 4386–4392

[u] MacLennan, D. H., Rice, W. J., and Green, N. M. (1997) *J. Biol. Chem.* **272**, 28815–28818

[v] Yonekura, K., Stokes, D. L., Sasabe, H., and Toyoshima, C. (1997) *Biophys. J.* **72**, 997–1005

[w] Yano, M., El-Hayek, R., and Ikemoto, N. (1995) *J. Biol. Chem.* **270**, 3017–3021

[x] Mitterdorfer, J., Sinnegger, M. J., Grabner, M., Striessnig, J., and Glossmann, H. (1995) *Biochemistry* **34**, 9350–9355

[y] McPherson, P. S., and Campbell, K. P. (1993) *J. Biol. Chem.* **268**, 13765–13768

[z] Sakube, Y., Ando, H., and Kagawa, H. (1997) *J. Mol. Biol.* **267**, 849–864

[aa] Zorzato, F., Fujii, J., Otsu, K., Phillips, M., Green, N. M., Lai, F. A., Meissner, G., and MacLennan, D. H. (1990) *J. Biol. Chem.* **265**, 2244–2256

[bb] Donoso, P., Beltrán, M., and Hidalgo, C. (1996) *Biochemistry* **35**, 13419–13425

[cc] Lee, M. G., Xu, X., Zeng, W., Diaz, J., Wojcikiewicz, R. J. H., Kuo, T. H., Wuytack, F., Racymaekers, L., and Muallem, S. (1997) *J. Biol. Chem.* **272**, 15765–15770

[dd] Graeff, R. M., Franco, L., De Flora, A., and Lee, H. C. (1998) *J. Biol. Chem.* **273**, 118–125

[ee] Isenberg, G., Etter, E. F., Wendt-Gallitelli, M.-F., Schiefer, A., Carrington, W. A., Tuft, R. A., and Fay, F. S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5413–5418

[ff] Horne, J. H., and Meyer, T. (1997) *Science* **276**, 1690–1693

[gg] Lee, H. C., and Aarhus, R. (1997) *J. Biol. Chem.* **272**, 20378–20383

[hh] Miller, R. J. (1992) *J. Biol. Chem.* **267**, 1403–1406

[ii] Quitterer, U., Schröder, C., Müller-Esterl, W., and Rehm, H. (1995) *J. Biol. Chem.* **270**, 1992–1999

[jj] Suzuki, K. (1987) *Trends Biochem. Sci.* **12**, 103–105

[kk] Martin, B. L., and Graves, D. J. (1986) *J. Biol. Chem.* **261**, 14545–14550

[ll] Moudy, A. M., Handran, S. D., Goldberg, M. P., Ruffin, N., Karl, I., Kranz-Eble, P., DeVivo, D. C., and Rothman, S. M. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 729–733

[mm] Kasumi, T., Tsumuraya, Y., Brewer, C. F., Kersters-Hilderson, H., Claeyssens, M., and Hehre, E. J. (1987) *Biochemistry* **26**, 3010–3016

[nn] Santella, L., and Carafoli, E. (1997) *FASEB J.* **11**, 1091–1109

[oo] Lipp, P., Thomas, D., Berridge, M. J., and Bootman, M. D. (1997) *EMBO J.* **16**, 7166–7173

[op] Corbett, E. F., and Michalak, M. (2000) *Trends Biochem. Sci.* **25**, 307–311

[pp] Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) *Science* **256**, 632–638

[qq] Cheung, W. Y. (1980) *Science* **207**, 19–27

[rr] Geisow, M. J. (1987) *Trends Biochem. Sci.* **12**, 120–121

[ss] Prentki, M., Glennon, M. C., Thomas, A. P., Morris, R. L., Matschinsky, F. M., and Corkey, B. E. (1988) *J. Biol. Chem.* **263**, 11044–11047

[tt] Tang, Y., and Othmer, H. G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7869–7873

[uu] Putney, J. W., Jr. (1998) *Science* **279**, 191–192

[vv] Berridge, M. J. (1997) *Nature (London)* **386**, 759–760

[ww] Hajnóczky, G., and Thomas, A. P. (1997) *EMBO J.* **16**, 3533–3543

increase above the normal level in keratinocytes of patients with the skin disease psoriasis.[123b]

Calcium ions are often involved in holding negatively charged groups together, for example, in the binding of proteins to phospholipid membranes. Among these membrane-associated proteins are the vitamin K-dependent proteins, all of which contain several residues of the chelating amino acid γ-carboxyglutamate (Chapter 15) at their calcium-binding sites.[124,125] Many of these are involved in the clotting of blood (Chapter 12). Some of the membrane-binding proteins called **annexins** (Chapter 8) are also Ca^{2+}-dependent ion channels.[126,127] Other Ca^{2+}-requiring lipid-binding proteins include the lipocortins, calpactins, and calelectrins.[128] Cadherins bind Ca^{2+} and help provide cohesion between cells.[129] Many proteins contain bound Ca^{2+} in precisely defined sites where it plays a structural role. These include the galactose-binding protein of bacterial transport and chemotaxis (Fig. 4-18)[130] and α-lactalbumin of milk. Although α-lactalbumin does not contain the helix–loop–helix pattern, its Ca^{2+}-binding site does consist of three carboxylate groups and two backbone carbonyl oxygen atoms.[131] The α-amylases, thermolysin, and staphylococcal nuclease (Chapter 12), and the lectin favin (Fig. 2-15) all contain bound Ca^{2+}. Calcium ions also bind to anionic groups in carbohydrates, e.g., to the sulfate groups in carageenin gels (Chapter 4) where they provide structural stability.

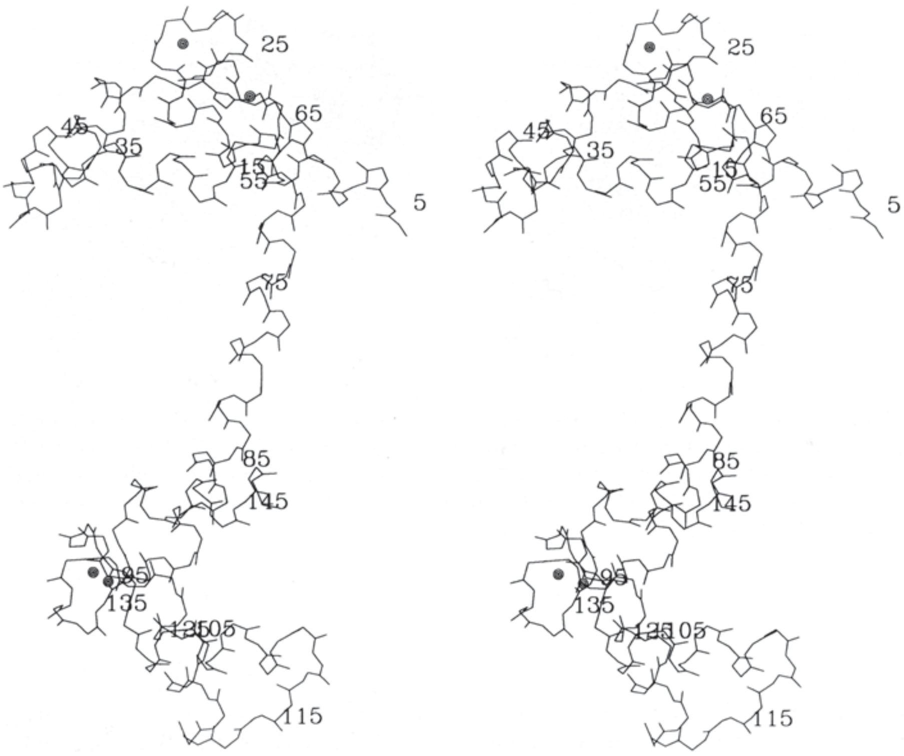
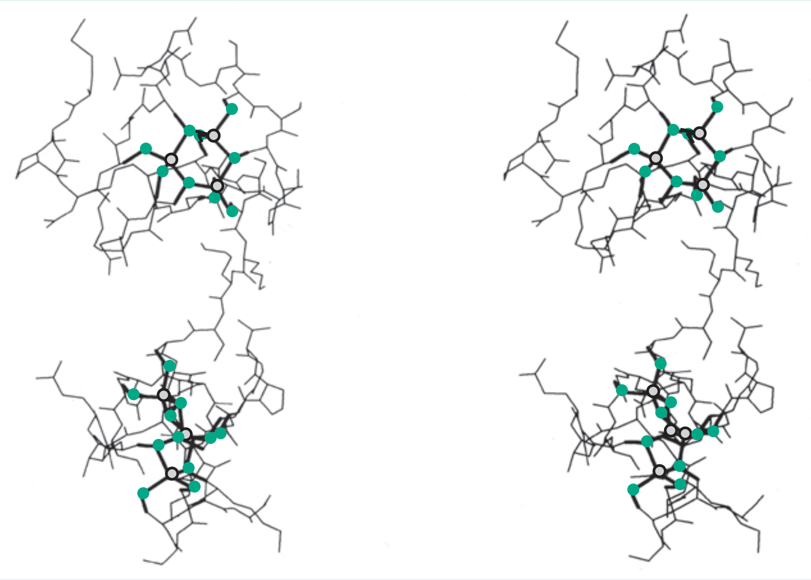


Figure 6-8 Stereoscopic backbone trace of a 148-residue recombinant calmodulin. The two helix–turn–helix (EF-hand) loops and their bound Ca^{2+} (as concentric circles) are at the top and the two near the C terminus are at the bottom. The long central helix, seen in this crystal structure, may undergo conformational changes during the functioning of this Ca^{2+}-sensing molecule.[132] From Chattopadhyaya *et al.*[111] Courtesy of F. A. Quiocho.

BOX 6-E METALLOTHIONEINS



Metallothionein containing bound Cd^{2+} and Zn^{2+}. From Robbins *et al.*[f]

If animals ingest excessive amounts of Zn(II), Cd (II), Hg(II), or Cu(I) their livers and kidneys accumulate these metals as complexes of proteins called metallothioneins.[a–e] In mammals at least three related genes encode these metal-binding proteins. The best known, metallothionein II, has a highly conserved 61-residue sequence containing 20 cysteine residues and no aromatic residues. The protein is organized into two domains, each able to bind a cluster of metal ions via thiolate side chains. The three-dimensional structure of rat liver metallothionein containing five Cd^{2+} and two Zn^{2+} ions is shown in the accompanying stereoscopic diagram.[f] The 61 alpha carbons, the beta carbon and sulfur atoms (green) of cysteine residues and the bound metal ions are indicated. The N-terminal domain (residues 1–29) contains one Cd^{2+} and two Zn^{2} and nine cysteine sulfurs which bind the metal ions. Three of the sulfur atoms form bridges between pairs of metals. The second cluster contains four Cd^{2+} held by 11 cysteine sulfur atoms, five of which bridge between pairs of metals. All of the metal ions are tetrahedrally coordinated. The polypeptide chains of metallothioneins consist predominantly of beta turns.[g] Important techniques in the study of these proteins include ^{113}Cd NMR,[h,i]

BOX 6-E METALLOTHIONEINS (continued)

spectroscopic methods,[j,k] and X-ray absorption.[l]

Transcription of metallothionein genes is induced by metal ions, and toxic metals such as Cd and Hg accumulate as metallothionein complexes, suggesting that one function is to protect against metal toxicity.[m] However, synthesis is also induced by glucocorticoid hormones,[n] and accumulation of a high concentration of copper and zinc in fetal metallothionein suggests a role in storage of these essential metals.[c] Another metal that binds to metallothioneins is Au(I),[o] which is widely used in thiolate salts as a chemotherapeutic agent for rheumatoid arthritis.

Metallothioneins are also found in insects, lower invertebrates, and even in bacteria.[i,p] Nevertheless, there are other metal-binding proteins.[q,r] For example, albacore tuna contain a 66-kDa glycoprotein that contains eight mole percent histidine and binds three Zn^{2+}, each by a cluster of three His.[s] Plants and fungi contain **phytochelins**, peptides consisting largely of repeated γ-glutamylcysteine units.[a,t,u] These appear to protect plants against toxicity of cadmium in the same manner as do the metallothioneins in our bodies.

---

[a] Mehra, R. K., Garey, J. R., Butt, T. R., Gray, W. R., and Winge, D. R. (1989) *J. Biol. Chem.* **264**, 19747–19753

[b] Kägi, J. H. R., and Kojima, Y., eds. (1987) *Metallothionein II*, Birkhäuser, Basel, Berlin

[c] Fischer, E. H., and Davie, E. W. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3333–3334

[d] Kägi, J. H. R., and Schäffer, A. (1988) *Biochemistry* **27**, 8509–8515

[e] Hamer, D. H. (1986) *Ann. Rev. Biochem.* **55**, 913–951

[f] Robbins, A. H., McRee, D. E., Williamson, M., Collett, S. A., Xuong, N. H., Furey, W. F., Wang, B. C., and Stout, C. D. (1991) *J. Mol. Biol.* **221**, 1269–1293

[g] Pande, J., Pande, C., Gilg, D., Vasák, M., Callender, R., and Kägi, J. H. R. (1986) *Biochemistry* **25**, 5526–5532

[h] Cismowski, M. J., Narula, S. S., Armitage, I. M., Chernaik, M. L., and Huang, P. C. (1991) *J. Biol. Chem.* **266**, 24390–24397

[i] Vasák, M., Hawkes, G. E., Nicholson, J. K., and Sadler, P. J. (1985) *Biochemistry* **24**, 740–747

[j] Willner, H., Vasák, M., and Kägi, J. H. R. (1987) *Biochemistry* **26**, 6287–6292

[k] Lu, W., and Stillman, M. J. (1993) *J. Am. Chem. Soc.* **115**, 3291–3299

[l] George, G. N., Byrd, J., and Winge, D. R. (1988) *J. Biol. Chem.* **263**, 8199–8203

[m] Sadhu, C., and Gehamu, L. (1988) *J. Biol. Chem.* **263**, 2679–2684

[n] Karin, M., Haslinger, A., Holtgreve, H., Richards, R. I., Krauter, P., Westphal, H. M., and Beato, M. (1984) *Nature (London)* **308**, 513–519

[o] Laib, J. E., Shaw, C. F., III, Petering, D. H., Eidsness, M. K., Elder, R. C., and Garvey, J. S. (1985) *Biochemistry* **24**, 1977–1986

[p] Freedman, J. H., Slice, L. W., Dixon, D., Fire, A., and Rubin, C. S. (1993) *J. Biol. Chem.* **268**, 2554–2564

[q] Waalkes, M. P., and Peratoni, A. (1986) *J. Biol. Chem.* **261**, 13097–13103

[r] Willuhn, J., Schmitt-Wrede, H. P., Greven, H., and Wunderlich, F. (1994) *J. Biol. Chem.* **269**, 24688–24691

[s] Dyke, B., Hegenauer, J., Saltman, P., and Laurs, R. M. (1987) *Biochemistry* **26**, 3228–3234

[t] Reese, R. N., and Winge, D. R. (1988) *J. Biol. Chem.* **263**, 12832–12835

[u] Clemens, S., Kim, E. J., Neumann, D., and Schroeder, J. I. (1999) *EMBO J.* **18**, 3325–3333

References

1. Lewis, G. N., and Randall, M. (1961) *Thermodynamics and the Free Energy of Chemical Substances*, 2nd ed., McGraw-Hill, New York
2. Mahan, B. H. (1963) *Elementary Chemical Thermodynamics*, Benjamin, New York
3. Everdell, M. H. (1965) *Introduction to Chemical Thermodynamics*, Norton, New York
4. Dickerson, R. E. (1969) *Molecular Thermodynamics*, Benjamin, New York
5. Jones, M. N., ed. (1988) *Biochemical Thermodynamics*, 2nd ed., Elsevier, Amsterdam
6. Hinz, H.-J., ed. (1986) *Thermodynamic Data for Biochemistry and Biotechnology*, Springer-Verlag, Berlin
7. Gutfreund, H., and Edsall, J. T. (1983) *Biothermo-dynamics*, Wiley, New York
8. Harold, F. M. (1986) *The Vital Force: A Study of Bioenergetics*, Freeman, San Francisco, California
9. Garby, L., and Larsen, P. S. (1995) *Bioenergetics; Its Thermodynamic Foundations*, Cambridge Univ. Press, London and New York
10. Jou, D., and Llebot, J. E., eds. (1990) *Introduction to the Thermodynamics of Biological Processes*, Prentice Hall, Englewood Cliffs, New Jersey

10a. Wadsö, I. (1985) *Eur. J. Biochem.* **153**, 429–434

11. Sturtevant, J. M. (1971) in *Techniques of Chemistry*, Vol. I, Part V (Weissberger, A., ed), pp. 347–425, Wiley, New York
12. Brown, H. D., ed. (1969) *Biochemical Microcalorimetry*, Academic Press, New York
13. Peters, K. S., and Snyder, G. J. (1988) *Science* **241**, 1053–1057
14. White, A., Handler, P., and Smith, E. L. (1968) *Principles of Biochemistry*, 4th ed., McGraw-Hill, New York (pp. 291–301)
15. Mosteller, R. D. (1988) *N. Engl. J. Med.* **318**, 1130
16. Donata, K., and Hegsted, D. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4866–4870
17. Ben-Naim, A. (1992) *Statistical Thermodynamics for Chemists and Biochemists*, Plenum, New York
18. Donovan, J. W. (1984) *Trends Biochem. Sci.* **9**, 340–344
19. Tanaka, A., Flanagan, J., and Sturtevant, J. M. (1993) *Protein Sci.* **2**, 567–576
20. Pace, C. N., and Laurents, D. V. (1989) *Biochemistry* **28**, 2520–2525
21. Interunion Commission on Biothermodynamics. (1976) *J. Biol. Chem.* **251**, 6879–6885
22. Phillips, R. C., George, P., and Rutman, R. J. (1963) *Biochemistry* **2**, 501–508
23. Liu, Y., and Sturtevant, J. M. (1995) *Protein Sci.* **4**, 2559–2561
24. Halling, P. J. (1989) *Trends Biochem. Sci.* **14**, 317–318
25. Örstan, A. (1990) *Trends Biochem. Sci.* **15**, 137–138
26. Nicolis, G., and Prigogine, I. (1977) *Self-organization in Non-equilibrium Systems: From Dissipative Structure to Order Through Fluctuations*, Wiley-Interscience, New York
27. Procaccia, I., and Ross, J. (1977) *Science* **198**, 716–717
28. Morowitz, H. J. (1968) *Energy Flow in Biology*, Academic Press, New York
29. Caplan, S. R., and Essig, A. (1983) *Bioenergetics and Linear Nonequilibrium Thermodynamics/ The Steady State*, Harvard Univ. Press, Cambridge, Massachusetts
30. Peacocke, A. R. (1983) *An Introduction to the Physical Chemistry of Biological Organization*, Clarendon Press, Oxford
31. Blumenfeld, L. A. (1983) *Physics of Bioenergetic Processes*, Springer-Verlag, Berlin
32. Katchalsky, A., and Curran, P. F. (1965) *Non-equilibrium Thermodynamics in Biophysics*, Harvard Univ. Press, Cambridge, Massachusetts
33. Prigogine, I. (1967) *Introduction to Thermodynamics of Irreversible Processes*, 3rd ed., Wiley, New York
34. Coveney, P. V. (1988) *Nature (London)* **333**, 409–415
35. Wilson, D. F. (1982) *Trends Biochem. Sci.* **7**, 275–278
36. Westerhoff, H. V. (1982) *Trends Biochem. Sci.* **7**, 275–279
37. Guynn, R. W., and Veech, R. L. (1973) *J. Biol. Chem.* **248**, 6966–6972
38. Alberty, R. A. (1969) *J. Biol. Chem.* **244**, 3290–3302
39. Phillips, R. C., George, P., and Rutman, R. J. (1969) *J. Biol. Chem.* **244**, 3330–3342
40. George, P., Phillips, R. C., and Rutman, R. J. (1963) *Biochemistry* **2**, 508–512
41. Purich, D. L., and Fromm, H. J. (1972) *Curr. Top. Cell. Regul.* **6**, 131–167
42. Jencks, W. P. (1968) in *Handbook of Biochemistry* (Sober, H. A., ed), pp. J-148, CRC, Cleveland, Ohio
43. Bassham, J. A., and Krause, G. H. (1969) *Biochim. Biophys. Acta.* **189**, 207–221
44. Rosing, J., and Slater, E. C. (1972) *Biochim. Biophys. Acta.* **267**, 275–290
45. Alberty, R. A. (1992) *Biophysical Chem.* **42**, 117–131
46. Alberty, R. A., and Cornish-Bowden, A. (1993) *Trends Biochem. Sci.* **18**, 288–291
47. Alberty, R. A., and Goldberg, R. N. (1992) *Biochemistry* **31**, 10610–10615
48. Alberty, R. A. (1994) *Pure & Appl. Chem.* **66**, 1641–1666
49. Stokes, G. B. (1988) *Trends Biochem. Sci.* **13**, 422–424
50. Mavrovouniotis, M. L. (1991) *J. Biol. Chem.* **266**, 14440–14445
51. Dutton, P. L. (1971) *Biochim. Biophys. Acta.* **226**, 63–80
52. Pepin, C. A., and Wood, H. G. (1986) *J. Biol. Chem.* **261**, 4476–4480
53. Wood, H. G., and Clark, J. E. (1988) *Ann. Rev. Biochem.* **57**, 235–260
54. Crooke, E., Akiyama, M., Rao, N. N., and Kornberg, A. (1994) *J. Biol. Chem.* **269**, 6290–6295
55. Veech, R. L., Lawson, J. W. R., Cornell, N. W., and Krebs, H. A. (1979) *J. Biol. Chem.* **254**, 6538–6547
56. Swedes, J. S., Sedo, R. J., and Atkinson, D. E. (1975) *J. Biol. Chem.* **250**, 6930–6938
57. Chapman, A. G., Fall, L., and Atkinson, D. E. (1971) *J. Bacteriol.* **108**, 1072–1086
58. Purich, D. L., and Fromm, H. J. (1973) *J. Biol. Chem.* **248**, 461–466
59. Teague, W. E., Jr., and Dobson, G. P. (1992) *J. Biol. Chem.* **267**, 14084–14093
60. Cohen, S. M., Ogawa, S., Rottenberg, H., Glynn, P., Yamane, T., Brown, T. R., and Shulman, R. G. (1978) *Nature (London)* **273**, 554–556

60a. Ugurbil, K., Kingsley-Hickman, P. B., Sako, E. Y., Zimmer, S., Mohanakrishnan, P., Robitaille, P. M. L., Thoma, W. J., Johnson, A., Foker, J. E., and From, A. H. L. (1987) *Ann. N.Y. Acad. Sci.* **508**, 265 –286

61. Brosnan, M. J., Chen, L., Van Dyke, T. A., and Koretsky, A. P. (1990) *J. Biol. Chem.* **265**, 20849–20855

61a. Saupe, K. W., Spindler, M., Hopkins, J. C. A., Shen, W., and Ingwall, J. S. (2000) *J. Biol. Chem.* **275**, 19742 –19746

62. Hutson, S. M., Williams, G. D., Berkich, D. A., LaNoue, K. F., and Briggs, R. W. (1992) *Biochemistry* **31**, 1322–1330
63. Burt, C. T., Glonek, T., and Bárány, M. (1977) *Science* **195**, 145–149
64. Seeley, P. J., Sehr, P. A., Gadian, D. G., Garlick, P. B., and Radda, G. K. (1977) in *NMR in Biology* (Dwek, R. A., Campbell, I. D., Richards, R. E., and Williams, R. J. P., eds), pp. 247–275, Academic Press, London
65. Masson, S., and Quistorff, B. (1992) *Biochemistry* **31**, 7488 – 7493
66. Jeffrey, F. M., Storey, C. J., Nunnally, R. L., and Malloy, C. R. (1989) *Biochemistry* **28**, 5323–5326
67. Kagimoto, T., Higaki, T., Nagata, K., Morino, Y., and Takatsuki, K. (1989) *NMR in Biomedicine* **2**, 93–97
68. Park, J. H., Brown, R. L., Park, C. R., Cohn, M., and Chance, B. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8780–8784
69. Radda, G. K. (1992) *FASEB J.* **6**, 3032–3038
70. Radda, G. K. (1986) *Science* **233**, 640–645
71. Burt, C. T. (1985) *Trends Biochem. Sci.* **10**, 404–406
72. Grove, T. H., Ackerman, J. J. H., Radda, G. K., and Bore, P. J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 299–302
73. Bottomley, P. A. (1985) *Science* **229**, 769–772
74. Fossel, E. T., Morgan, H. E., and Ingwall, J. S. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3654–3658
75. Metzler, D. E., Harris, C. M., Johnson, R. J., Siano, D. B., and Thomson, J. A. (1973) *Biochemistry* **12**, 5377–5392
76. Johnson, R. J., and Metzler, D. E. (1970) *Methods Enzymol.* **18**, 433–471
77. Dixon, H. B. F. (1992) *Essays in Biochemistry* **27**, 161–176
78. Vander Jagt, D. L., Hansen, L. D., Lewis, E. A., and Han, L. B. (1972) *Arch. Biochem. Biophys.* **153**, 55–61
79. Dixon, H. B. F., and Tipton, K. F. (1973) *Biochem. J.* **133**, 837–842
80. Klotz, I. M., and Hunston, D. L. (1975) *J. Biol. Chem.* **250**, 3001–3009
81. Weber, G. (1975) *Adv. Prot. Chem.* **29**, 1–83
82. Chivers, P. T., Prehoda, K. E., Volkman, B. F., Kim, B.-M., Markley, J. L., and Raines, R. T. (1997) *Biochemistry* **36**, 14985–14991
83. Johnson, F. A., Lewis, S. D., and Shafer, J. A. (1981) *Biochemistry* **20**, 44–48
84. McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Körner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) *Biochemistry* **35**, 9958–9966
85. Katritzky, A. R., and Waring, A. (1962) *J. Chem. Soc.*, 1540–1548
86. Katritzky, A. R., and Karelson, M. (1991) *J. Am. Chem. Soc.* **113**, 1561–1566
87. Bjerrum, J. (1941) *Metal Amine Formation in Aqueous Solution*, Haase & Son, Copenhagen
88. Eichhorn, G. L., ed. (1973) *Inorganic Biochemistry*, 2 vols, Elsevier, Amsterdam
89. Martell, A. E., and Smith, R. M. (1975) *Critical Stability Constants*, Plenum, New York
90. Michaelian, K. H., and Moskovits, M. (1978) *Nature (London)* **273**, 135–136
91. Lewsin, S. (1974) *Displacement of Water and its Control in Biochemical Reactions*, Academic Press, London
92. Karlin, S., Zhu, Z.-Y., and Karlin, K. D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14225–14230
93. Freeman, H. C., Healy, M. J., and Scudder, M. L. (1977) *J. Biol. Chem.* **252**, 8840–8847
94. Torrado, A., Walkup, G. K., and Imperiali, B. (1998) *J. Am. Chem. Soc.* **120**, 609–610
95. Nakayama, S., and Kretsinger, R. H. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 473–507

References

96. Kretsinger, R. H. (1976) *Ann. Rev. Biochem.* **45**, 239–266

96a. Pawlowski, K., Bierzynski, A., and Godzik, A. (1996) *J. Mol. Biol.* **258**, 349–366

97. Ikura, M. (1996) *Trends Biochem. Sci.* **21**, 14–17

98. Kretsinger, R. H., and Nockolds, C. E. (1973) *J. Biol. Chem.* **248**, 3313–3326

99. McPhalen, C. A., Sielecki, A. R., Santarsiero, B. D., and James, M. N. G. (1994) *J. Mol. Biol.* **235**, 718–732

99a. Declercq, J.-P., Evrard, C., Lamzin, V., and Parello, J. (1999) *Protein Sci.* **8**, 2194 – 2204

99b. Biekofsky, R. R., Martin, S. R., Browne, J. P., Bayley, P. M., and Feeney, J. (1998) *Biochemistry* **37**, 7617 – 7629

100. Drake, S. K., Zimmer, M. A., Miller, C. L., and Falke, J. J. (1997) *Biochemistry* **36**, 9917–9926

101. Akke, M., Forsén, S., and Chazin, W. J. (1995) *J. Mol. Biol.* **252**, 102–121

102. Denisov, V. P., and Halle, B. (1995) *J. Am. Chem. Soc.* **117**, 8456–8465

103. Heizmann, C. W., and Hunziker, W. (1991) *Trends Biochem. Sci.* **16**, 98–103

104. Ikemoto, N., Ronjat, M., Mészáros, L. G., and Koshita, M. (1989) *Biochemistry* **28**, 6764–6771

105. Meldolesi, J., and Pozzan, T. (1998) *Trends Biochem. Sci.* **23**, 10–14

106. Randolph, J. T., McClure, K. F., and Danishefsky, S. J. (1995) *J. Am. Chem. Soc.* **117**, 5712–5719

107. Fliegel, L., Burns, K., MacLennan, D. H., Reithmeier, R. A. F., and Michalak, M. (1989) *J. Biol. Chem.* **264**, 21522–21528

108. Cohen, P., and Klee, C. B., eds. (1988) *Calmodulin*, Elsevier, Amsterdam

109. James, P., Vorherr, T., and Carafoli, E. (1995) *Trends Biochem. Sci.* **20**, 38–42

110. Lafitte, D., Capony, J. P., Grassy, G., Haiech, J., and Calas, B. (1995) *Biochemistry* **34**, 13825–13832

111. Chattopadhyaya, R., Meador, W. E., Means, A. R., and Quiocho, F. A. (1992) *J. Mol. Biol.* **228**, 1177–1192

112. Pedigo, S., and Shea, M. A. (1995) *Biochemistry* **34**, 10676–10689

113. da Silva, A. C. R., and Reinach, F. C. (1991) *Trends Biochem. Sci.* **16**, 53–57

114. Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry* **36**, 4386–4392

115. Trigo-Gonzalez, G., Awang, G., Racher, K., Neden, K., and Borgford, T. (1993) *Biochemistry* **32**, 9826–9831

116. Spyracopoulos, L., Li, M. X., Sia, S. K., Gagné, S. M., Chandra, M., Solaro, R. J., and Sykes, B. D. (1997) *Biochemistry* **36**, 12138–12146

117. Travé, G., Lacombe, P.-J., Pfuhl, M., Saraste, M., and Pastore, A. (1995) *EMBO J.* **14**, 4922–4931

118. Näthke, I., Hill, B. L., Parham, P., and Brodsky, F. M. (1990) *J. Biol. Chem.* **265**, 18621–18627

119. Pley, U. M., Hill, B. L., Alibert, C., Brodsky, F. M., and Parham, P. (1995) *J. Biol. Chem.* **270**, 2395–2402

120. Engel, J., Taylor, W., Paulsson, M., Sage, H., and Hogan, B. (1987) *Biochemistry* **26**, 6958–6965

121. Engel, E., Richter, K., Obermeyer, G., Briza, P., Kungl, A. J., Simon, B., Auer, M., Ebner, C., Rheinberger, H.-J., Breitenbach, M., and Ferreira, F. (1997) *J. Biol. Chem.* **272**, 28630–28637

122. Schäfer, B. W., and Heizmann, C. W. (1996) *Trends Biochem. Sci.* **21**, 134–140

123. Drohat, A. C., Baldisseri, D. M., Rustandi, R. R., and Weber, D. J. (1998) *Biochemistry* **37**, 2729 – 2740

123a. Gribenko, A. V., and Makhatadze, G. I. (1998) *J. Mol. Biol.* **283**, 679 – 694

123b. Brodersen, D. E., Nyborg, J., and Kjeldgaard, M. (1999) *Biochemistry* **38**, 1695 – 1704

124. Christiansen, W. T., Tulinsky, A., and Castellino, F. J. (1994) *Biochemistry* **33**, 14993–15000

125. Schwalbe, R. A., Ryan, J., Stern, D. M., Kisiel, W., Dahlbäck, B., and Nelsestuen, G. L. (1989) *J. Biol. Chem.* **264**, 20288–20296

126. Demange, P., Voges, D., Benz, J., Liemann, S., Göttig, P., Berendes, R., Burger, A., and Huber, R. (1994) *Trends Biochem. Sci.* **19**, 272–276

127. Saurel, O., Cézanne, L., Milon, A., Tocanne, J.-F., and Demange, P. (1998) *Biochemistry* **37**, 1403–1410

128. Moss, S. E., and Crumpton, M. J. (1990) *Trends Biochem. Sci.* **15**, 11–12

129. Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grübel, G., Legrand, J.-F., Als-Nielsen, J., Colman, D. R., and Hendrickson, W. A. (1995) *Nature (London)* **374**, 327–336

130. Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1988) *Science* **242**, 1290–1295

131. Stuart, D. I., Acharya, K. R., Walker, N. P. C., Smith, S. G., Lewis, M., and Phillips, D. C. (1986) *Nature (London)* **324**, 84–87

132. Williams, R. J. P. (1991) *Trends Biochem. Sci.* **16**, 206

Study Questions

1. a) From $\Delta G°$ for hydrolysis of sucrose (Table 6-6) calculate the equilibrium constant

$$K = [\text{glucose}][\text{fructose}] / [\text{sucrose}]$$

at 25°C. Call this hydrolysis reaction 1.

b) Is the sucrose in a 1 M solution stable? Explain.

c) If acid is added to a 1 M sucrose solution to catalyze its hydrolysis, what will be the final sucrose concentration at equilibrium? (Assume that concentrations equal activities for the purpose of these calculations.)

d) Reaction 2 is the hydrolysis of α-D-glucose 1-phosphate to glucose and inorganic phosphate (P_i). Using $\Delta G°$ for this reaction (Table 6-6) calculate the equilibrium constant.

e) Sucrose phosphorylase from the bacterium *Pseudomonas saccharophila* catalyzes the following reaction (reaction 3):

$$\text{Sucrose} + P_1 \rightarrow$$
$$\alpha\text{-D-glucose 1-phosphate} + \text{fructose}$$

Calculate the equilibrium constant and the standard Gibbs energy change at 25°C for reaction 3 from the equilibrium constants obtained above for reactions 1 and 2. Show that $\Delta G°$ for reaction 3 = $\Delta G°$ of reaction 1 - $\Delta G°$ of reaction 2.

f) Could the bacterium carry out reaction 3 in the following two consecutive steps? Explain.

$$\text{Sucrose} \rightarrow \text{glucose} + \text{fructose}$$
$$\text{Glucose} + P_i \rightarrow \text{glucose 1-phosphate}$$

2. For each of the following reactions, state whether the equilibrium constant will be between 0.1 and 10 (i.e., about one), greater than 100, or less that 0.01. Assume that the pH is constant at 7.0.
   a) $2\ ADP^{3-} \rightarrow ATP^{4-} + AMP^{2-}$
   b) ATP^{3-} + glucose → glucose 6-phosphate^{2-} + $ADP^{2-} + H^+$
   c) $ADP^{2-} + HPO_4^{2-} + H^+ \rightarrow ATP^{3-}$
   d) Glucose 6-phosphate^{2-} → fructose 6-phosphate^{2-}
   e) Phosphoenolpyruvate^{3-} + glucose → glucose 1-phosphate^{2-} + pyruvate^{-}

3. The combustion of 1 mol of solid urea to liquid water and gaseous carbon dioxide and nitrogen (N_2) in a bomb colorimeter at 25°C (constant volume) liberated 666 kJ of heat energy. Calculate ΔH, the change in heat content (enthalpy), for this reaction.

4. Using data of Table 6-4 calculate $\Delta G'$ (pH 7) for the following reactions:
   a) Glucose → 2 lactate$^-$ + $2\ H^+$
   b) $2\ NH_4^+ + HCO_3^-$ → urea + $2\ H_2O + H^+$
   c) 2-oxoglutarate^{2-} + $1/2\ O_2$ + CoA + H^+ → succinyl-CoA + $H_2O + CO_2$

5. What is the ionic strength of a 0.2 M solution of NaCl? of 0.2 M Na_2SO_4?

6. The [ATP] / [ADP] ratio in an actively respiring yeast cell is about 10. What would be the intracellular [3-phosphoglycerate] / [1,3-bisphosphoglycerate] ratio have to be to make the phosphoglycerate kinase reaction (Fig. 9-7, reaction 7) proceed toward 1,3-bisphosphoglycerate synthesis at 25°C, pH 7?

7. a) Using data from Table 6-8 determine the equilibrium constant for the reaction between malate and methylene blue, assuming all reactants present initially at the same concentration. Indicate clearly the direction of the reaction for which the Gibbs energy change is written.

b) Calculate the percentage of the reduced (leuco) form of methylene blue present at pH 7 and 25°C in a system for which the measured electrode potential is 0.065 V.

8. NAD^+ is a coenzyme for both pyruvate dehydrogenase and ethanol dehydrogenase. Using the values of E_o' from Table 6-8 calculate the Gibbs energy change and the equilibrium constant for the reaction.

Lactate$^-$ + acetaldehyde → pyruvate$^-$ + ethanol

9. Consider the oxidation of acetate at 25°C:

$$CH_3COO^- (0.1\ M) + 2\ O_2\ (g, 0.2\ atm) +$$
$$H^+ (10^{-7}\ M) \rightarrow 2\ H_2O\ (l) + 2\ CO_2\ (g)$$

a) What is the equilibrium pressure of CO_2 if the reaction is not coupled to any other reaction?

b) What is the equilibrium pressure of CO_2 if the reaction is coupled to the formation of 0.01 M ATP from 0.02 M ADP and 0.01 M HPO_4^{2-} in the citric acid cycle?

c) What do the above calculations tell you about the prospects of gaining 100% efficiency of energy storage in ATP from the citric acid cycle?

d) If the actual pressure of CO_2 is 0.01 atm, what is the efficiency of energy storage under the conditions in (b)?

Study Questions

10. The equilibrium constant for the following reaction, which is catalyzed by creatine kinase, has been determined by chemical analysis. The data are given below. [S. A. Kuby and E. A. Noltman, in *The Enzymes*, 2nd ed. (P. D. Boyer, H. Lardy, and K. Myrbäck, eds), Vol. VI, pp. 515–602. Academic Press, New York, 1962]

$$ATP^{4-} + \text{creatine} \rightarrow ADP^{3-} + \text{creatine phosphate}^{2-} + H^+$$

t (°C)	K
20	6.30×10^{-9}
30	5.71×10^{-9}
38	5.47×10^{-9}

a) What are $\Delta G°$, $\Delta H°$, and $\Delta S°$ for the reaction at 25°C?
b) What are $\Delta G'$, $\Delta H'$, and $\Delta S'$ (pH 7) for the reaction at 25°C?
c) What are $\Delta G'$, $\Delta H'$, and $\Delta S'$ (pH 7) for the hydrolysis of creatine phosphate at 25°C?

11. The following reaction was carried out in a calorimeter at 25°C in 0.1 M phosphate buffer at pH 7.4 in the presence of a particulate suspension containing the mitochondrial electron transport system [M. Poe, H. Gutfreund, and R. W. Estabrook, *ABB* **122**, 204–211 (1967)]:

$$NADH + H^+ + 1/2\ O_2\ (\text{aq., sat.}) \rightarrow NAD^+ + H_2O$$

The oxygen consumption was monitored continuously with an oxygen electrode. The temperature was monitored simultaneously with a thermocouple immersed in the solution. At the start of the reaction 96 µmol NADH was added to 29.0 ml buffer containing O_2. A nearly zero-order reaction was observed with the rate of O_2 consumption of 6.87 µmol/min and the rate of temperature rise of 0.01171 K/min. The heat capacity of the calorimeter and contents was 254.6 J/K. What is ΔH for the above reaction? NOTE: The H^+ is supplied by the phosphate buffer, which has a ΔH of dissociation of 5.4 kJ mol^{-1}.

12. Enthalpy and Gibbs energy changes for the following reaction at 25°C are given in Table 6-6.

$$ATP^{4-}\ (1\ M) + H_2O \rightarrow ADP^{3-}\ (1\ M) + H^+\ (10^{-7}\ M) + HPO_4^{2-}\ (1\ M)$$

a) How much heat is evolved at constant temperature and pressure if the reaction takes place in a test tube without doing any work other than $p\,\Delta V$ work?
b) How much heat is evolved or absorbed by the foregoing reaction if it is coupled with 100% efficiency to an endergonic reaction?
c) What efficiency of coupling to an endergonic reaction is required in order that the foregoing reaction neither evolve nor absorb heat?

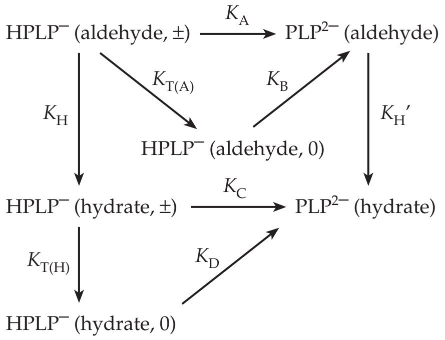
13. Microorganisms use a great variety of fermentation reactions for obtaining energy. Could the following reaction be used for such a purpose? Explain the reasons for your answer.

$$C_4H_6O_4\ (\text{succinic acid}) + H_2O \rightarrow C_3H_8O_3\ (\text{glycerol}) + CO_2$$

14. a) Calculate the work done in kJ and in kcal by a 70 kg person in climbing up stairs three stories (13 m).
b) Calculate how much ATP (in mmol and in grams) will be needed for the climb if muscles can use the ATP with 50% efficiency and if the phosphorylation state $R_p = [ATP]/[ADP][P_i]$ is $10^3\ M^{-1}$. The standard value of $\Delta G^{°'}$ (pH 7) for hydrolysis of MgATP to MgADP and P_i may be taken as −31 kJ/mol at 37°C.
c) How much of this ATP could be provided by transfer of phospho groups from the stored creatine phosphate (Cr-*P*) to ADP in muscle? Assume that [Cr-*P*] = 20 mM and may fall to 10 mM during the climb. $\Delta G^{°'}$ (pH 7) for hydrolysis of creatine phosphate is about −43 kJ/mol. If the creatine kinase reaction attains equilibrium what will the value of R_p be for the adenylate system?

15. Acid–base titration gave the following three pK_a values for cysteine: 1.70, 8.36, and 10.53. Spectrophotometric data allowed H. B. F. Dixon and K. F Tipton (1973, *Biochem J.* **133**, 837–842) to estimate the following ratio of tautomeric species at pH 9.4 $[{}^-OOC-C(NH_3^+)-CH_2S^-]/[{}^-OOC-C(NH_2)-CH_2SH] = 2.12$, as is also shown in Fig. 6-5. Verify and assign the four microscopic pK_a values.

16. The coenzyme pyridoxal phosphate (PLP; structure on p. 740) has pK_a values of 3.62, 6.10, and 8.33, as determined by acid–base titration or by spectrophotometric titration. The pK_a of 6.10 belongs primarily to the phosphate group and is nearly independent of the others. However, the pK_a values of 3.62 and 8.33 are shared by the protonated ring nitrogen and the phenolic –OH group. PLP exists as an equilibrium mixture of aldehyde together with its covalent hydrate (see Eq. 13-1). The equilibrium constants for hydrate formation (K_h = [hydrate] / [aldehyde]) are independent of pH but differ for each ionization state of the ring. Consider only the equilibria in the pH range 7–12.



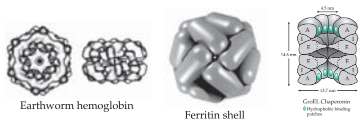
The UV-visible absorption spectrum of the monoprotonated form of PLP was divided mathematically into individual bands for the aldehyde with dipolar ionic ring (±), the aldehyde tautomer with an uncharged ring (0) and the hydrate of the dipolar ion. The following fractions were estimated (Harris *et al.*, 1976, *Biochem. Biophys. Acta.* **521**, 181–194.

Aldehyde (±)	52%
Aldehyde (0)	28%
Hydrate (±)	16%
Hydrate (0)	4%, estimated indirectly

From 300 MHz ^{1}H NMR spectra areas for the following hydrogen atoms were estimated by Robitaille.

	pH 7	pH 12
4'-H aldehyde	.86	.900
4'-H hydrate	.14	.012
6-H, aldehyde	.96	1.059
6-H, hydrate	.17	.014
2'-CH_3, aldehyde	2.49	2.96
2'-CH_3, hydrate	.51	.038

Evaluate the tautomerization constants $K_{T(A)}$ and $K_{T(H)}$, the hydration constants K_H and $K_{H'}$, and the microscopic pK_a values pK_A, pK_B, pK_C, and pK_D in the foregoing scheme. Assume that the hydration ratios K_h and $K_{h'}$ are the same in H_2O (spectrophotometric data) and in D_2O (NMR data).



Some ways in which protein subunits associate. (Left) The 3.66 MDa hemoglobin of the earthworm *Lumbricus terrestris* contains 144 globin subunits organized as 12 cylindrical disulfide-linked dodecamers. Two 6-dodecamer layers, each a ring with 6-fold cyclic symmetry, lie back-to-back. This reconstructed particle also contains three types of linker proteins in the center region. From Lamy *et al.* (2000) *J. Mol. Biol.* **298**, 638. (Center) The iron storage protein ferritin is formed from 24 19- to 21-kDa 4-helix-bundle subunits with cubic symmetry. As many as 4500 atoms of iron, as hydrated iron oxide, may be stored in the internal cavity. See Fig. 7-13. From Trikha *et al.* (1995) *J. Mol. Biol.* **248**, 954. Courtesy of Elizabeth Theil. (Right) The 2 MDa molecular chaperone GroEL consists of two back-to-back 7-subunit rings, each subunit formed from domains E, I and A. A 7-subunit cap of the smaller GroES may cover either end to form a compartment in which polypeptides fold. See Box 7-A.

Contents

325 .... **A. Describing Binding Equilibria**
325 .... 1. Analyzing Data
327 .... 2. Multiple Binding Sites on a Single Molecule
328 .... *Microscopic binding constants and statistical effects*
329 .... *Electrostatic repulsion: anticooperativity*
330 .... 3. Cooperative Processes
332 .... **B. Complementarity and the Packing of Macromolecules**
332 .... 1. Rings and Helices
332 .... *Molecules with cyclic symmetry*
334 .... *Helical structures*
334 .... *Filamentous bacteriophages*
334 .... *A rod-shaped plant virus*
335 .... *Bacterial pili*
336 .... *The thin filaments of muscle*
337 .... 2. Oligomers with Twofold (Dyad) Axes
337 .... *Paired interactions*
338 .... *Dihedral symmetry*
342 .... *Oligomers with cubic symmetry (polyhedra)*
344 .... *Asymmetry and quasi-equivalence in oligomers*
345 .... *Quasi-equivalence in virus coats*
348 .... *Regulatory subunits and multienzyme complexes*
349 .... **C. Cooperative Changes in Conformation**
349 .... 1. Unequal Binding of Substrate and "Induced Fit"
350 .... 2. Binding Equilibria for a Dimerizing Protein
350 .... *The Monod–Wyman–Changeux (MWC) model*
351 .... *The induced fit model*
352 .... *One conformational state dissociated*
352 .... 3. Higher Oligomers
353 .... **D. The Oxygen-Carrying Proteins**
353 .... 1. Myoglobin and Hemoglobin
353 .... *The binding of oxygen*
355 .... *Structural changes accompanying oxygen binding*
357 .... *The Bohr effect and allosteric regulators*
358 .... *Carbon monoxide, cyanide, and nitric oxide*
359 .... 2. Abnormal Human Hemoglobins
362 .... 3. Comparative Biochemistry of Hemoglobin
362 .... **E. Self-Assembly of Macromolecular Structures**
362 .... 1. Bacteriophages
367 .... 2. "Kringles" and Other Recognition Domains and Motifs
368 .... **F. The Cytoskeleton**
369 .... 1. Intermediate Filaments
369 .... 2. Microfilaments
370 .... 3. Microtubules

373 .... **References**
377 .... **Study Questions**

Boxes

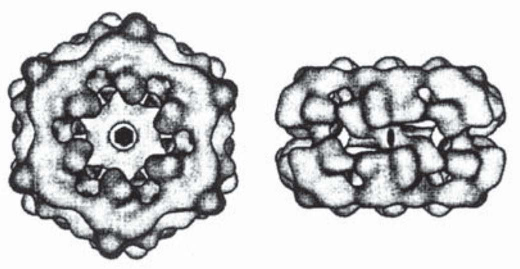
339 .... Box 7-A Life and Death for Proteins: Chaperonins and Proteasomes
360 .... Box 7-B Sickle Cell Disease, Malaria, and Blood Substitutes
363 .... Box 7-C The T-Even Bacteriophages
371 .... Box 7-D Mitosis, Tetraploid Plants, and Anticancer Drugs

Tables

329 .... Table 7-1 Binding Constants of Protons to Dianions of Dicarboxylic Acids
355 .... Table 7-2 Thermodynamic Functions for Oxygenation of Hemoglobin
367 .... Table 7-3 A Few Well-Known Structural Domains

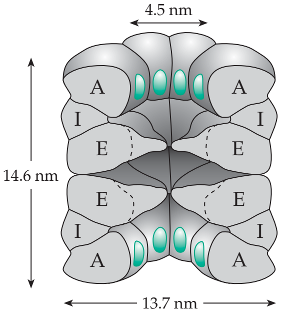
How Macromolecules Associate

7



Earthworm hemoglobin

Ferritin shell



GroEL Chaperonin
🟩 Hydrophobic binding patches

The complicated shapes and internal structures of cells are determined to a large extent by the way in which proteins and other macromolecules are bonded one to another. In addition, intimate association of macromolecules is essential to such biological processes as the motion of flagella, the contraction of muscle, the action of antibodies, the transmission of nerve impulses, the replication of DNA, and the synthesis of proteins. Equally important is the binding of small molecules to large ones. In this chapter we will first examine methods of measuring binding with an emphasis on protons and small molecules. Then we will consider the ways in which macromolecules stick together as well as the role of conformational changes within macromolecules.

A. Describing Binding Equilibria

In previous discussions of pH we have dealt with dissociation constants, but in this section we will use formation constants K_f, where $K_f = 1/K_d$. Measurement of the strength of association of molecules is an everyday aspect of modern biochemical research. It may be important to know how strongly a hormone binds to a receptor in a cell membrane or how well a feedback inhibitor binds to an enzyme to determine whether the interaction is significant physiologically. The binding of O_2 to hemoglobin and other oxygen carriers is vitally important, but the description of these oxygenation reactions is mathematically complex. This is especially so because we must consider effects of pH changes and of changing concentrations of allosteric effectors on the binding equilibria.

In considering such equilibria we must first examine the individual interactions of different domains of a protein, one with another. These can be described by association constants or, alternatively, by the Gibbs energy changes for the association reaction.[1,2] The average kinetic energy of motion of a molecule in solution is about $3/2k_BT$, where k_BT is Boltzmann's constant. For one mole the kinetic energy is $3/2RT$ or 3.7 kJ (0.89 kcal) mol^{-1} at 25°C. Thus, if $K_f = 10\ M^{-1}$ ($\Delta G° = -5.7$ kJ mol^{-1} or -1.36 kcal mol^{-1}) the binding energy is only slightly in excess of the thermal energy of the molecules and the complex is weakly bound. In this instance, if X and P are both present in 10^{-4} molar concentrations (typical enough for biochemical systems), only 0.1% of the molecules will exist as the complex ([complex] = K_f[X][P]). If the formation constant is higher by a factor of 1000, i.e., $K_f = 10^4\ M^{-1}$ ($\Delta G° = -22.8$ kJ mol^{-1}), 38% of the molecules will exist as the complex; while if $K_f = 10^7\ M^{-1}$ (extremely strong binding, $\Delta G° = -40$ kJ or -9.55 kcal mol^{-1}), 97% of the molecules will be complexed.

1. Analyzing Data

The extent of binding of a molecule X to another molecule P (Eqs. 7-1, 7-2) is measured by varying the concentrations of X and P and observing changes in the concentration of the complex [PX]. The first

$$X + P \rightleftarrows PX \tag{7-1}$$

$$K_f = [PX]/[P][X]) \tag{7-2}$$

prerequisite is to find a measurable property that is different for the complex than for either of the free components. For example, the complex may be colored and the components colorless. More commonly, the complex simply has a different light absorbance (A) at a certain wavelength than do the components. Likewise, the circular dichroism or the chemical shift of a peak in the NMR spectrum may change. If P is an enzyme, only the complex PX will undergo decomposition to products. Sometimes (but not always) the rate of breakdown of PX (the enzyme–substrate complex) to form products will be relatively slow compared to the rate at which the equilibrium between X, P, and PX is established. In this case the concentration of complex PX will be proportional to the observed rate of formation of product.

Whatever change of property is measured, its value will increase with increasing concentrations of X if the total concentration of the macromolecule P is kept constant. In the usual experimental design, the molar concentration of P is small and it is possible to increase the concentration of X to quite large values. When this is done, it is usually observed that at high enough values of [X] almost all of the P is converted to PX, and the change being measured (e.g., ΔA for increased light absorption) no longer increases. This effect is known as **saturation** and is observed in most binding studies and also in many physiological phenomena.

The property being measured (ΔA) reaches a maximum value ΔA_{max} at saturation and when all of compound P has been converted to PX. The ratio of [PX] to the total concentration of all forms of P present $[P]_t$ is known as the **saturation fraction** and is often given the symbol Y. If P has more than one binding site for X, Y is defined as the fraction of the total binding sites occupied. If n is the number of sites per molecule, the total number of sites is n[P]. The value of Y is often taken as $\Delta A/\Delta A_{max}$, an equality that holds for multisite macromolecules only if the change in A is the same for each successive molecule of X added. This is not always true, but when it is Eq. 7-3 is followed.

$$\sum_i i\frac{[PX_i]}{n[P]_t} = Y = \frac{\Delta A}{\Delta A_{max}} \tag{7-3}$$

Here i represents the number of ligands X bound to P and may vary from 0 to n. When $n = 1$ the saturation fraction Y and ΔA are related to the concentration of free unbound X and the formation constant as follows:

$$Y = \frac{K_f[X]}{1 + K_f[X]} \qquad \Delta A = \frac{\Delta A_{max}\, K_f[X]}{1 + K_f[X]} \tag{7-4}$$

A plot of Y or ΔA against [X] is shown in Fig. 7-1. This kind of plot is sometimes called an **adsorption isotherm** because it describes binding only at a constant temperature. Notice, from both Fig. 7-1 and Eq. 7-4, that Y reaches a value of 0.5 when [X] is just equal to $1/K_f$ (or to K_d). Note also that as [X] increases saturation is reached slowly and that even at the point representing the highest concentration of X ($8/K_f$ in Fig. 7-1) saturation is less than 90%. Since in the usual experimental situation, we do not know Y but only ΔA, it is difficult to estimate the limiting value ΔA_{max} from a plot of this type unless K_f is very high. However, we need to know ΔA_{max} to evaluate K_f. For this reason, plots like that of Fig. 7-1 are seldom used, this one being included mainly to illustrate a point of nomenclature. The curve shown in Fig. 7-1 is a rectangular hyperbola, and the type of saturation curve shown is frequently referred to as **hyperbolic**. This is in contrast to certain other binding curves (Section 3) which, when plotted in this way, are **sigmoidal** (S-shaped).

A better type of plot is often that of Y against log [X] (Fig. 7-2). It has the following features. (1) The curve is symmetric about the midpoint at log [X] = log K_f. (2) No matter how high or low the concentration range used in the experiments, it is easy to choose a scale that puts all the points on the same sheet of paper. (3) Spacing between points tends to be more uniform than in a plot against [X]; e.g., compare Figs. 7-1 and 7-2 for which the experimental points represent the same data and for which values of [X] for successive points are each twofold greater than the preceding one. (4) The same logarithmic scale can be used for all compounds, no matter how strong or weak the binding, and the same shape curve is obtained for all 1:1 complexes. The midpoint slope, $dY/d\log [X]$, is 0.576; the change in log [X] in going from 10 to 90% saturation is 1.81. The curve is familiar to most chemists because it is frequently used for pH titration curves in which pH substitutes for –log [X]. To represent a complex with tighter binding, the curve is simply moved to the left, and for weaker binding, it is moved to the right.

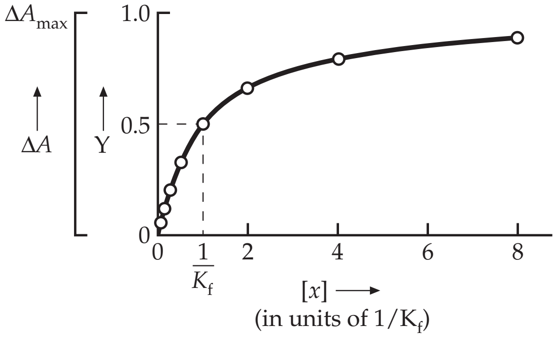


Figure 7-1 An adsorption isotherm, a plot of the saturation fraction Y or of some change in a measured property ΔA vs [X], the concentration of a substance that binds reversibly to a macromolecule. The curve is hyperbolic and $[X] = 1/K_f$ when Y = 0.5.

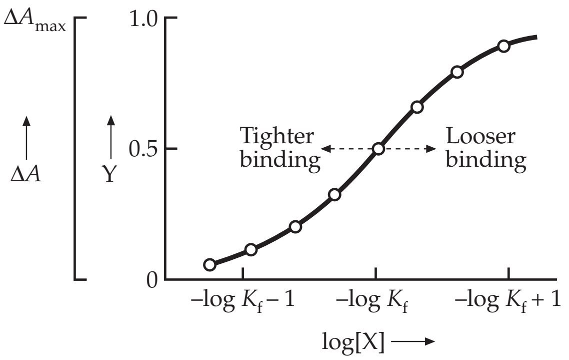


Figure 7-2 A saturation curve plotted on a logarithmic scale for [X]. The data points are the same as those used in Fig. 7-1.

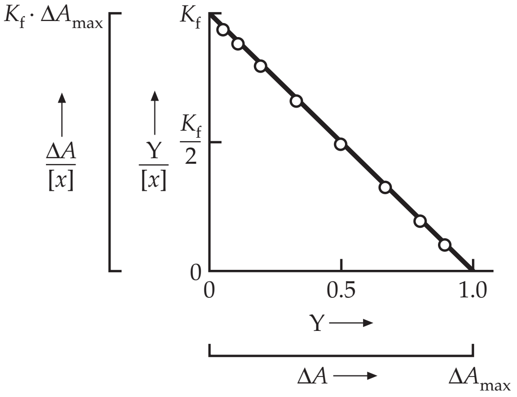


Figure 7-3 A Scatchard plot of the same data shown in Figs. 7-1 and 7-2. This is the best of the linear plots for studying binding.

Saturation data are often plotted in yet another form known as the **Scatchard plot** (Fig. 7-3). The value of $\Delta A / [X]$ (or of $Y / [X]$) is plotted against ΔA (or Y) and a straight line is fitted to the points, preferably using the "method of least squares." The intercept on the *x* axis and the slope of the fitted line give values of $\Delta A_{max} / K_f$ and K_f, respectively, as indicated by Eq. 7-5, which follows directly from Eq. 7-4.

$$Y / [X] = K_f - YK_f$$
$$\Delta A / [X] = \Delta A_{max} K_f - \Delta A K_f \qquad (7\text{-}5)$$

The Scatchard plot is the best of the various linear transformations of the saturation equation and is preferred to "double reciprocal plots" analogous to that shown in Fig. 9-3.

Scatchard's original equation was formulated to deal with the binding of two or more ligands to a single macromolecule.[3–5] If we let $[X]_b$ represent the concentration of bound X and $[P]_t$ the total molar concentration of the protein or other macromolecule and if there is only one binding site on the protein, $[X]_b / [P]_t$ will equal Y. However, if there are *n* independent binding sites that have the same binding constant K_f Eq. 7-5a will hold.

$$\frac{[X]_b}{[P]_t} \cdot \frac{1}{[X]} = nK_f - \frac{[X]_b}{[P]_t} K_f \qquad (7\text{-}5a)$$

If $[X]_b / [P]_t[X]$ is plotted against $[X]_b / [P]_t$ the resulting linear plot will have an intercept of K_f on the *y* axis and *n* on the *x* axis. Thus, *n* is directly apparent, whereas in Eq. 7-5 it is incorporated into Y. A problem arises if, as discussed in the next section, the multiple binding sites are not independent but interact. Curved Scatchard plots result and attempts to extract more than one binding constant can lead to very large errors. Before measuring saturation curves, the student should read additional articles or books on the subject.[2,6–10]

2. Multiple Binding Sites on a Single Molecule

A macromolecule may often be able to bind several molecules of a second compound X. Consider the case in which the macromolecule P binds successively one molecule of X, then a second, and a third, up to a total of *n*. We define the stepwise formation constants, $K_1, K_2, \ldots, K_n$, as follows:

$$P + X \xrightarrow{K_1} PX$$
$$PX + X \xrightarrow{K_2} PX_2 \ldots$$
$$PX_{n\text{-}1} + X \xrightarrow{K_n} PX_n \qquad (7\text{-}6)$$

Remember that these are reversible reactions even though unidirectional arrows are used. The general expression for the *i*th stepwise formation constant is given by Eq. 7-7.

$$K_i = \frac{[PX_i]}{[PX_{i-1}][X]} \qquad (7\text{-}7)$$

Remember that Y is the fraction of total binding sites saturated. The number of moles of X bound per mole of P is *n*Y and is obtained by summing the concentrations $[PX] + 2[PX_2] + \ldots$ and dividing the sum of all the forms of P:

For two binding sites $(n = 2)$

$$2Y = \frac{[PX] + 2[PX_2]}{[P] + [PX] + [PX_2]} \tag{7-8}$$

For the general case

$$nY = \left(\sum_{i=1}^{n} i[PX_i] \right) \Big/ \left([P] + \sum_{i=1}^{n} [PX_i] \right) \tag{7-9}$$

The summations are over all of the integral values of i from 1 to n. Now, by expressing each concentration, $[PX_i]$, in terms of the concentrations [X] and [P] of *free* X and P, together with the stepwise formation constants, we obtain Eq. 7-10.

For $n = 2$

$$2Y = \frac{K_1[X] + 2K_1K_2[X]^2}{1 + K_1[X] + K_1K_2[X]^2} \tag{7-10}$$

A similar equation can be written for the general case. Note that the concentration of P does not appear in Eq. 7-10 and that Y is a function only of [X] and the stepwise formation constants. Such equations define the isotherms for binding of two or more molecules of X to P. From an experimental plot of Y (or of ΔA) vs [X] or log [X], it is possible in favorable cases to determine the stepwise constants $K_1, K_2, \ldots, K_n$. However, this becomes quite complicated. To simplify Eq. 7-10 and the corresponding equation for the general case, we can group the constants together and designate the products of constants (K_1, K_1K_2, $K_1K_2K_3$, etc.) as $\psi_1, \psi_2, \ldots, \psi_n$. Our equations are now as follows:

For $n = 2$

$$2Y = \frac{\psi_1[X] + 2\psi_2[X]^2}{1 + \psi_1[X] + \psi_2[X]^2} \tag{7-11}$$

For the general case

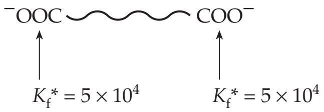
$$nY = \sum_{i=1}^{n} i\psi_i[X]^i \Big/ \left(1 + \sum_{i=1}^{n} \psi_i[X]^i \right) \tag{7-12}$$

From experimental data, it is usually easiest to first determine the ψ's (there are n of them), and then to calculate from the ψ's the stepwise constants. For example:

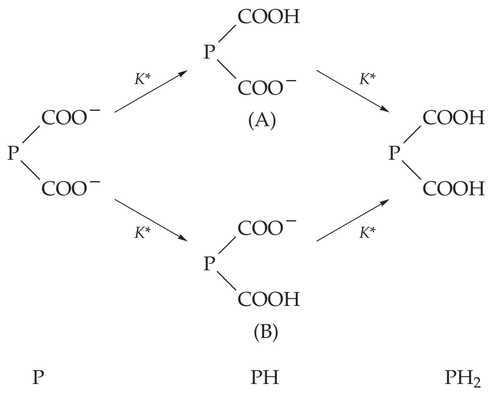
$$K_1 = \psi_1 \qquad K_2 = \psi_2/K_1 \text{, etc.} \tag{7-13}$$

While Eq. 7-12, known as the **Adair equation**,[11] might seem to provide a complete description of the binding process, it usually does not. In many cases, there is more than one kind of binding site on a macromolecule and Eq. 7-12 tells us nothing about the distribution of the ligand X among different sites in complex PX. To consider this problem we must examine the *microscopic binding constants*.

Microscopic binding constants and statistical effects. As discussed in Chapter 6, Section E,2, microscopic binding constants represent the constants for binding to specific individual sites. Now, consider a straight-chain dicarboxylic acid which has two *identical* binding sites for protons. If the chain connecting the two carboxylate anions is long enough, the carboxylate groups will be far enough apart that they do not influence each other through electrostatic interaction.



Each group will have a microscopic binding constant (K_f*) of 5 x 10^4 M^{-1}. The constant K_f* can also be called an **intrinsic binding constant**, because it is characteristic of a carboxylate group that is free of interactions with other groups. Intuition tells us (correctly) that, in its binding of protons, a solution of this dicarboxylic acid dianion will behave exactly like a solution of the monovalent anion $R{-}COO^-$ at twice the concentration. A single intrinsic binding constant suffices to describe both binding sites. It may seem surprising then that the **stepwise formation constants** (also called **stoichiometric** or **macroscopic** formation constants) K_1 and K_2 differ: K_1 = 10 x 10^4 M^{-1} and K_2 = 2.5 x 10^4 M^{-1}. This fact reflects the so-called statistical effect. Either of the two carboxylate groups in the molecule can bind a proton in the first step to give two indistinguishable molecules, PH:



(7-14)

If we label the two forms of PH as A and B (Eq. 7-14) and consider that each one of them is independently in equilibrium with P through formation constant K_f^*, we obtain Eq. 7-15 (which may be compared with Eqs. 6-75 and 6-76, which are written for *dissociation* constants).

$$K_1 = \frac{[\mathrm{PH}]_A + [\mathrm{PH}]_B}{[\mathrm{P}][\mathrm{H}^+]} = 2K^* \qquad \text{and} \qquad K_2 = K^*/2 \tag{7-15}$$

This result is related to probability and arises for the same reason that if you reach into a barrel containing 50% white balls and 50% black balls, you will pull out one of each just twice as often as you will pull out a pair of white or a pair of black. In the general case of *n* equivalent binding sites, the microscopic formation constants K_i^* are related[12,13] to the stepwise constants K_i as follows:

$$K_i = \frac{(n + 1 - i)}{i} K_i^* \tag{7-16}$$

It is also easy to show,[14] using Eqs. 7-12 and 7-16 that for *n completely equivalent and independent* binding sites Eqs. 7-17 and 7-18 hold:

$$\mathrm{Y} = \frac{K^*[\mathrm{X}](1 + K^*[\mathrm{X}])^{n-1}}{(1 + K^*[\mathrm{X}])^n} \tag{7-17}$$

or

$$\mathrm{Y} = \frac{K^*[\mathrm{X}]}{1 + K^*[\mathrm{X}]} \tag{7-18}$$

In this case the microscopic association constants are all identical and represent a single **intrinsic** constant applicable to all of the sites. In fact, Eq. 7-18 is identical to that for association of a single proton (or other ligand) with a single binding site, satisfying our intuitive notion that a set of *n* completely independent binding sites should behave just like a solution of an *n*-fold more concentrated compound with a single binding site. Thus, our arithmetic has led us to a conclusion that was already obvious. However, it is rarely true that binding sites on a single macromolecule are completely independent; there is almost always *interaction* between them, and the equations that we have derived for evaluation of stepwise and intrinsic constants cannot be applied without modification.

TABLE 7-1
Binding Constants of Protons to Dianions of Dicarboxylic Acids[a]

Acid dianion	No. of CH_2 groups	log K_1 (pK_2)	log K_2 (association) (pK_1) (dissociation)
Hypothetical dianion with log K^* = 4.8	∞	5.1	4.5
Azelaic	7	5.41	4.55
Adipic	4	5.41	4.42
Succinic	2	5.48	4.19
Malonic	1	5.69	2.83

[a] From R. P. Bell, (1973) *The Proton in Chemistry*, 2nd ed., p. 96. Cornell Univ. Press, Ithaca, New York

Electrostatic repulsion: anticooperativity. As we have seen, a hypothetical acid with an infinite distance between the carboxylate groups and log K_f^* = 4.8 would have two macroscopic binding constants separated by the statistical distance (log 4 = 0.6). Compare these values with the observed binding constants for protons with the dianions of acids containing 7, 4, 2, and 1 CH_2 groups given in Table 7-1. For the longest chain (that of azelaic acid) the log K_f values are not very different from those of the hypothetical long-chain acid. However, as the groups come closer together, the first binding constant is increased markedly because of the additional electrostatic attraction and the second is decreased. The spread between the two log K_f values increases from 0.6 to as much as 2.9 as a result of interaction between the binding sites.

In malonic and succinic acids the first proton bound can be shared by both carboxyl groups through formation of a hydrogen bond. (See discussion in Chapter 9, Section D.) Additional factors operate in oxalic acid where the carboxyl groups are connected directly and for which pK_a values (pK_a = log K_f) are 4.19 and 1.23. In all of these examples the binding of the first proton makes it harder to bind a second proton. Such negative interaction or **anticooperativity** between binding sites is very common and always leads to a spread of the formation constants and a broadening of the curve of Y vs log [X]. This is shown graphically on the right side of Fig. 7-4 where the binding curves for protons with acetate ion and with succinic acid dianion are compared. Notice that binding of protons *increases* as log $[H^+]$ increases, giving the curves an unfamiliar appearance when compared with the more familiar curve of *dissociation vs* pH.

Can we predict the pK_a values in Table 7-1? With an appropriate dielectric constant chosen Eq. 2-8 can be applied. The difference between the two successive log K_f values reduced by 0.6 (the statistical factor) is a

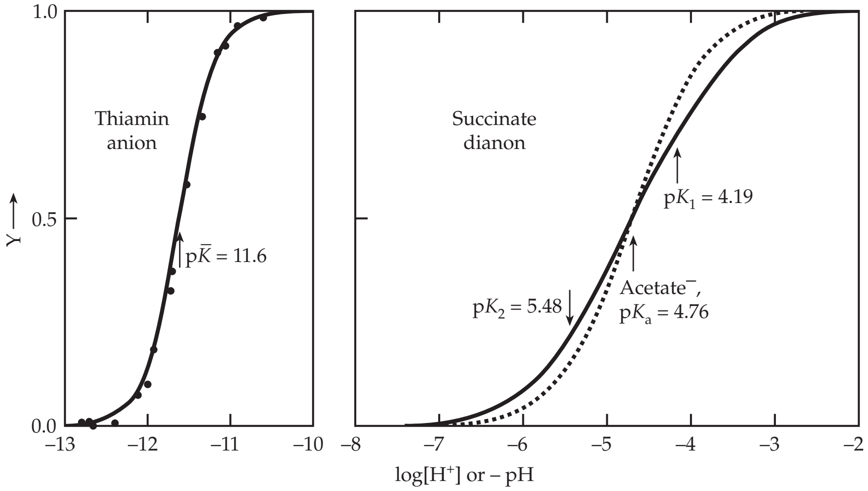
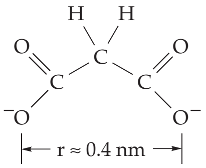


Figure 7-4 Binding of protons to the thiamin anion, the succinate dianion, and the acetate anion. Acetate (dashed line) binds a single proton with a normal width binding curve. Succinate dianion binds two protons with anticooperativity, hence a broadening of the curve. The thiamin anion (yellow form, see Eq. 7-19) binds two protons with complete cooperativity and a steep binding curve.

measure of the electrostatic effect. For malonic acid this ΔpK_a is 2.25 and for succinic acid it is 0.69 (Table 7-1). In 1923, N. Bjerrum proposed that the value of ΔpK_a could be equated directly with the work needed to bring the two negative charges together to a distance representing the charge separation in the malonate dianion.



Thus, applying Eq. 6-31, $\Delta G = 5.708\ \Delta pK_a$ kJ mol^{-1} = 12.84 kJ mol^{-1} for malonate. Equating this with W in Eq. 2-8 and assuming a dielectric constant of 78.5 (that of water), the distance of charge separation r is calculated to be 0.138 nm. This is much too small. The computation was improved by Westheimer and Kirkwood, who assumed a dielectric constant of 2.0 *within* the molecule. By approximating the molecule as an ellipsoid of revolution, they were able to make reasonably accurate calculations of electrostatic effects on pK_a values.[15] Thus, for malonic acid Westheimer and Shookhoff[16] predicted $r = 0.41$ nm for malonic acid dianion. Recently more sophisticated calculations[17] have been used to predict pK_a values for the compounds in Table 7-1 and others.[18]

Electrostatic theory has also been used successfully to interpret titration curves of proteins in which the net negative or positive charge distributed over the surface of the protein varies continuously from high pH to low as more protons are added.[19]

Electrostatic effects can be transmitted extremely effectively through aromatic ring systems, a fact that explains some of the significance of heterocyclic aromatic systems in biochemical molecules. Consider the microscopic binding constant of the phenolate anion of pyridoxine as influenced by the state of protonation of the ring nitrogen. These are shown in Eq. 6-75, where $pK_a^* = 4.94$ and $pK_d^* = 8.20$ define the binding constants for protonation of a phenolate ion when the ring nitrogen is protonated or unprotonated, respectively. We see that $\Delta\ pK_a = 3.26$, even greater than that of malonic acid.

3. Cooperative Processes

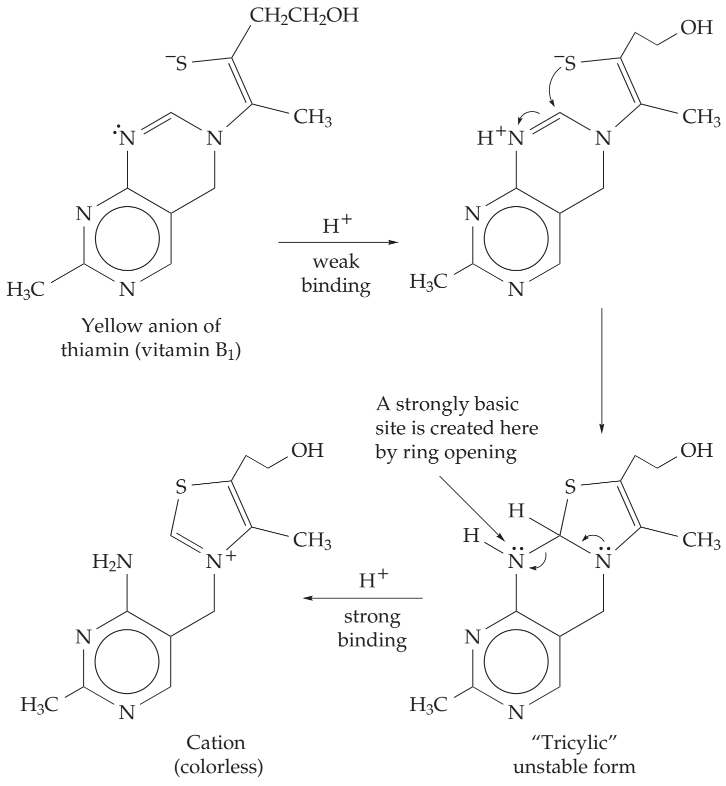
Can it ever happen that interaction between groups leads to a *decrease* from the statistical separation between values of the stepwise constants instead of to an increase? At first glance, the answer seems to be no. A decreased separation would imply that the intrinsic binding constant for the second proton bound is higher than that for the first, but common sense tells us that the first proton ought to bind at the site with the highest binding constant. However, look at the experimental binding curve of protons with the anion of thiamin shown in Fig. 7-4. Instead of being broadened from the curve of acetate, it is just half as wide. The explanation depends upon some rather amusing chemistry of thiamine. Under suitable conditions, this vitamin can be crystallized as a yellow sodium salt, the structure of whose anion is shown in Eq. 7-19. Weak binding of a proton to one of the nitrogens as shown in Eq. 7-19 creates an electron deficiency at the adjacent carbon and the $-S^-$anion adds to the C=N group, closing the ring to an unstable tricyclic form of thiamin. This tricyclic form can be observed in methanol and can be crystallized. It is unstable in water because the central ring can open, with the electrons flowing as indicated by the small arrows to create a strongly basic site on the same nitrogen. A second proton combines at this basic nitrogen with a high binding constant to form a cation.

The key to the reversed order of strength of the binding *constants lies in the molecular rearrangements intervening between the two binding steps.*[20,21] In this particular case, we cannot measure the successive binding constants K_1 and K_2 directly from the titration curve because K_2 is almost two orders of magnitude larger than K_1. Consequently, the binding curve shown in Fig. 7-4 is (within the experimental error) twice as steep at the center (the slope is 2 x 0.576) as that for acetate ion and is accurately represented in Eq. 7-20. Comparison of Eq. 7-20 with Eq. 7-10 shows how the latter has been simplified because no significant concentration of the form PX is present in the cooperative case.

$$Y = \frac{\bar{K}^2[X]^2}{1 + \bar{K}^2[X]^2} \tag{7-20}$$

This is only part of the story about the acid–base chemistry of thiamin. For the rest, see Chapter 14, Section D,1.

The binding of protons by the thiamin anion is an example of a **cooperative process**, so named because binding of the first proton makes binding of the second easier. Although relatively rare among small molecules, cooperative processes are very common and important in biochemistry.[22,23] A cooperative binding curve is sometimes referred to as **sigmoidal** because the plot of Y against [X] (the binding isotherm) is S shaped. The maximum possible cooperativity is observed when the binding of the first ligand enhances the affinity of all other sites so much that no species other than P and PX_n are present in significant concentration.



Yellow anion of thiamin (vitamin B_1)

A strongly basic site is created here by ring opening

Cation (colorless)

“Tricyclic” unstable form

(7-19)

It is easy to show that, for *n* binding sites with such *completely cooperative binding*, the saturation fraction is:

$$Y = \frac{\bar{K}^n[X]^n}{1 + \bar{K}^n[X]^n} \tag{7-21}$$

where $\bar{K} = (K_1 \ldots K_n)^{1/n}$. The midpoint slope in the binding curve (Y vs log [X]) is 0.576*n* and the change, Δlog [X], between Y = 0.1 and Y = 0.9 is 1.81/*n*.

Equation 7-21 can be rewritten as

$$Y/(1 - Y) = \bar{K}^n[X]^n \tag{7-22}$$

Taking logarithms (Eq. 7-22)

$$\log[Y/(1 - Y)] = n \log \bar{K} + n \log[X] \tag{7-23}$$

A plot of log [Y/(1 – Y)] vs log [X] is known as a **Hill plot**. According to Eq. 7-22, it is linear with a slope of *n*. Remember that this equation was derived for an ideal case of completely cooperative binding at *n* sites.[24] However, Hill plots are often used to plot experimental data for systems in which cooperativity is incomplete. Thus, the experimentally measured slope of a Hill plot (n_{Hill}) is not an integer and is usually less than *n*, the number of binding sites. A comparison of n_{Hill} with *n* is often used as a measure of the degree of cooperativity: $n_{Hill}/n = 1.00$ for complete cooperativity but is less than one if cooperativity, is incomplete. An example of a very high degree of cooperativity is provided by the hexameric enzyme glutamate dehydrogenase, whose saturation curve for substrate displays n_{Hill} approaching six.[25] It is not necessary to make a Hill plot to get n_{Hill}. From the usual binding curve of Y (or Δ*A*) vs log [X] the midpoint slope can be measured with satisfactory precision. Alternatively, the difference, Δlog [X], between 0.1 and 0.9 saturation can be evaluated and n_{Hill} calculated from Eq. 7-24.

$$n_{Hill} = \frac{\text{midpoint slope}}{0.576} = \frac{1.81}{\Delta \log[X]} \tag{7-24}$$

Binding curves sometimes show more than one step; in such cases Hill plots are not linear and no simple measure of cooperativity can be defined.

A second example of cooperativity is provided by the reversible denaturation of coiled peptide chains. Some proteins can be brought to a pH of 4 by addition of acid but without protonation of buried groups with intrinsic pK_a values greater than four. When a little more acid is added, some less basic group is protonated, permitting the protein to unfold and to expose the more basic hidden groups. Thus, cooperative proton binding is observed. As in the case of thiamin the cooperativity depends upon the occurrence of a conformational change in the molecule linked to protonation of a particular group.

The reversible transformation between an α helix and a random coil conformation is also cooperative. In this case, once a helix is started, additional turns form rapidly and the molecule is completely converted into the helix. Likewise, once it unfolds it tends to unfold completely. Melting of DNA (Chapter 5) or, indeed, of any crystal is cooperative.[8] The stacking of nucleotides alongside a template polynucleotide can also be cooperative. For example, the binding of an adenylate residue to two strands of polyuridylic acid leads to cooperative formation of a triple-helical complex (Chapter 5, Section C,4). Here the stacking interactions make helix growth energetically easier than initiation of new helical regions.[26]

B. Complementarity and the Packing of Macromolecules

Because the forces acting between them are weak, two molecules will cling together tightly only if there is a close fit between their surfaces. For a firm bond to be formed many atoms must be in contact and the two molecular surfaces must be *complementary* one to the other. If a "knob," such as a $-CH_3$ group, is present on one surface, there must be an appropriate hollow in the complementary surface. A positive charge in one surface is likely to be opposite a negative charge in the other. A proton donor group can form a hydrogen bond only if it is opposite a group with unshared electrons; nonpolar (hydrophobic) groups must be opposite each other if hydrophobic interaction is to occur. An important principle is that *two molecules with complementary surfaces tend to join together and interact, whereas molecules without complementary surfaces do not interact*. Watson called this "selective stickiness."[27] Selective stickiness permits the **self-assembly** of biological macromolecules having surfaces of complementary shape into fibers, tubes, membranes, and polyhedra. It also provides the means for specific pairing of purine and pyrimidine bases during the replication of DNA and during the synthesis of RNA and of proteins.

Complementarity of surfaces is equally important to the chemical reactions of cells. Each of these reactions is catalyzed by an enzyme, which contains reactive chemical groupings in the right places and in the right orientations to interact with and promote a chemical change in another molecule, the **substrate**. Specific catalysis is one of the most basic characteristics of living things. Enzymatic catalysis provides the basis not only for the reactions of metabolism but also for the movement of muscle fibers, the flowing of the cytoplasm in the ameba, and virtually all other biological responses. To understand these phenomena requires an examination of the structures of the macromolecules involved and of the ways in which they can fit together.

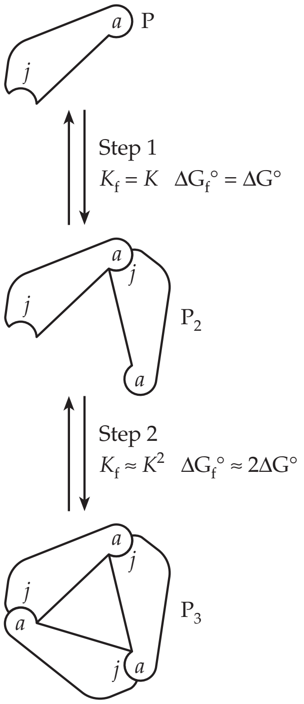
Just as the amino acids, sugars, and nucleotides are the building blocks for formation of proteins, polysaccharides, and nucleic acids, these three kinds of macromolecule are the units from which larger subcellular structures are assembled. Fibers, microtubules, virus "coats," and small symmetric groups of **subunits** in **oligomeric proteins** all result from the packing of macromolecules in well-defined ways, something that is often called **quaternary structure**.

1. Rings and Helices

Consider first the aggregation of identical protein subunits. While many protein molecules are nearly spherical, they are nevertheless asymmetric. In the drawings that follow the asymmetry is exaggerated, but the principles illustrated are valid. One easily observed lesson from nature is that even though living things are made up of asymmetric materials, a great deal of symmetry is evident.[28] At the molecular level the symmetry of crystalline arrays of atoms or molecules is described mathematically by the elements of symmetry present in **space groups** (p. 133). There are 230 of these but only 65 accommodate asymmetric objects (Chapter 3).[29] Two of the natural ways for identical asymmetric subunits to interact lead to rings and helices, respectively.

Molecules with cyclic symmetry. Consider a subunit (**protomer**) of the shape shown in Eq. 7-25 and containing a region *a* that is complementary to the surface *j* on another part of the same molecule. Two such protomers will tend to stick together to form a dimer, region *a* of one protomer sticking to region *j* of the other. The dimer will still contain a free region *a* at one end and a region *j* at the other which are not involved in bonding. Other protomers can stick to these free ends. In some instances long chains can be formed. However, if the geometry is just right, a third subunit can fit in to form a closed ring (a trimer). Depending on the geometry of the subunits, the ring can be even smaller (a dimer) or it can be larger (a tetramer, pentamer, etc.). The bonding involved is between two different regions (*a* and *j*) of a subunit and is sometimes described as **heterologous**.[30] To obtain a closed ring of subunits, the angle between the bonding groups *a* and *j* must be correct or the ring cannot be completed.

A ring formed using exclusively heterologous interactions possesses **cyclic symmetry**. The trimer in Eq. 7-25 has a **threefold axis**: Each subunit can be superimposed on the next by rotation through 360°/3. The oligomer is said to have C_3 symmetry. Many real proteins, including all of those with 3, 5, or another uneven number of identical protomers, appear to be formed of subunits arranged with cyclic symmetry. An example is the cholera toxin from *Vibrio cholerae*, which forms a pentamer with an outer ring of subunits with C_5 symmetry (Fig. 7-5).


a
P
j
Step 1
$K_f = K \quad \Delta G_f^\circ = \Delta G^\circ$
a
j
j
P$_2$
a
Step 2
$K_f \approx K^2 \quad \Delta G_f^\circ \approx 2\Delta G^\circ$
a
j
j
a
P$_3$
j
a

(7-25)

Now consider the quantitative aspects of heterologous interactions with ring formation. Let K_f be the formation constant and ΔG° the Gibbs energy change for the reaction of the *j* end of protomer P with the *a* end of a second protomer to form the dimer P_2 (Eq. 7-25).

In the second step (Eq. 7-25) a third protomer combines. It forms *two* new *aj* interactions. If we assume for this step that ΔG_f° is $2\,\Delta G^\circ$, K_f will be equal to K^2. The overall association constant for formation of a trimer from three protomers will be given by Eq. 7-26.

$$\begin{gathered} 3\text{ P} \rightarrow \text{P}_3 \\ K_f = K^3 \\ \Delta G_f^\circ = 3\,\Delta G^\circ \end{gathered} \tag{7-26}$$

This will be true only if ΔG° for formation of both new *aj* bonds in the trimers is exactly the same as that for formation of the *aj* bond in the dimer. The reader may wish to criticize this assumption[30a,b,c] and to suggest conditions that might lead to overestimation or underestimation of K_f for the trimer as calculated previously.

Now consider a hypothetical example: Protomer P is continuously synthesized by a cell. At the same time some subunits are degraded to a nonaggregating form via a second metabolic reaction. The two reactions are balanced so that [P] is always present at a steady state value of 10^{-5} M. Suppose that a value for a single *aj* interaction of $K = 10^4$ (and $\Delta G^\circ = -22.8$ kJ mol^{-1}) governs aggregation to form dimers and trimeric rings. What concentration of dimers and of trimer rings will be present in the cell in equilibrium with the 10^{-5} concentration of P. Using Eq. 7-25 we see that the concentration of dimers $[P_2]$ is $10^4 \times (10^{-5})^2 = 10^{-6}$ M. (Note that the amount of material in this concentration of dimer is equivalent to 2×10^{-6} M of the monomer units.) The concentration of rings $[P_3]$ is $(10^4)^3 \times (10^{-5})^3 = 10^{-3}$ M, equivalent to 3×10^{-3} M of the monomer units. Thus, of the *total* P present in the cell ($10^{-5} + 0.2 \times 10^{-5} + 300 \times 10^{-5}$ M), 99.6% is associated with trimers, 0.33% is still monomers, and only 0.07% exists as dimers. Thus, the formation of two heterologous bonds simultaneously to complete a ring imparts a high degree of cooperativity to the association reaction of Eq. 7-25. We will find in a cell mostly either rings or monomer but little dimer.

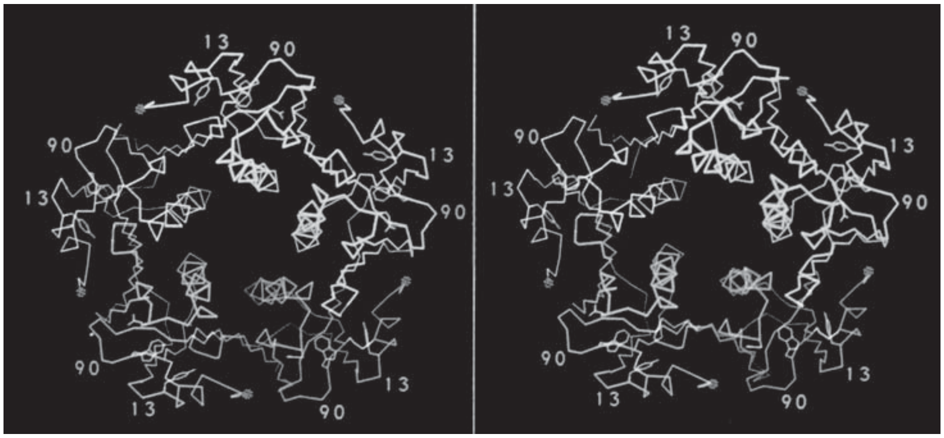


Figure 7-5 Stereoscopic view of the B_5 pentamer of cholera toxin B. The pentamer, known as **choleragenoid**, has a central hole of ~1.5 nm diameter into which a helix from the A subunit is inserted. As viewed here, the front surface of the pentamer has binding sites for the oligosaccharide chains of ganglioside GM_1, which serves as the toxin receptor. The back side binds the A subunit. See also Box 11-A. From Zhang *et al.*[31]

Now consider what will happen to the little rings within the cell if the process that removes P to a nonassociating form suddenly becomes more active so that [P] falls to 10^{-6} M. If K is still 10^4, what will be the percentages of P, P_2, and P_3 at equilibrium? Here we note a characteristic of cooperative processes: A higher than first power dependence on a concentration.

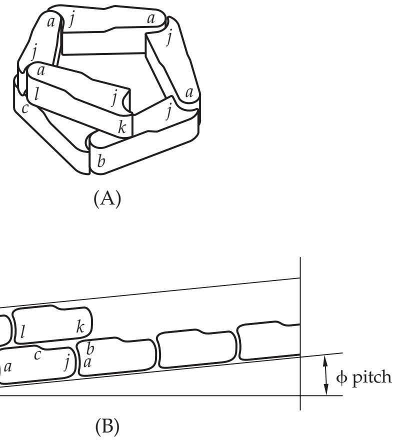


Figure 7-6 (A) Heterologous bonding of subunits to form a helix. (B) Radial projection of subunits arranged as in helix A. Different bonding regions of the subunit are designated *a*, *b*, *c*, *j*, *k*, and *l*.

Helical structures. If the angle at the interface *aj* is slightly different, instead of a closed ring, we obtain a helix as shown in Fig. 7-6A. The helix may have an integral number of subunits per turn or it may have a nonintegral number, as in the figure. The same type of heterologous interaction *aj* is involved in joining each subunit to the preceding one, but in addition other interactions occur. If the surfaces involved in these additional interactions are complementary and the geometry is correct, groups from two different parts of the molecule (e.g., *b* and *k*) may fit together to form another heterologous bond. Still a third heterologous interaction *cl* may be formed between two other parts of the subunit surfaces. If interactions *aj*, *bk*, and *cl* are strong (i.e., if the surfaces are highly complementary over large areas), extremely strong microtubular structures may be formed, such as those in the flagella of eukaryotic organisms (Fig. 1-8). If the interactions are weaker, labile microfilaments and microtubules, such as are often observed to form and dissociate within cells, may arise.

The geometry of subunits within a helix is often advantageously displayed by imagining that the surface of the structure can be unfolded to give a radial projection (Fig. 7-6B). Here subunits corresponding to those in the helix in Fig. 7-6A are laid out on a plane obtained by slitting the cylinder representing the surface of the helix and laying it out flat. In the example shown, the number of subunits per turn is about 4.8 but it can be an integral number. The interactions *bk* between subunits along the direction of the fiber axis may sometimes be stronger than those (*aj*) between adjacent subunits around the spiral. In such cases the microtubule becomes frayed at the ends through breaking of the *aj* interactions. This phenomenon can be observed under the electron microscope for the microtubules from flagella of eukaryotic organisms. Figures 7-7 to 7-10 show four helical structures from the molecular domain. They are a filamentous bacteriophage, a plant virus, a bacterial pilus, and an actin microfibril. Each is composed largely of a single kind of protomer. A larger and more complex helical structure, the microtubule, is shown in Fig. 7-34.

Filamentous bacteriophages. Bacteriophages of the Ff family include the fd, f1, and M13 strains.[31a,32–36] Phage M13 is widely used in cloning genes and for many other purposes (Chapter 26). The genome is a circular, single-stranded DNA of ~6400 nucleotides which is held in an elongated double-stranded form by a helical sheath of about 2700 subunits of a 50-residue protein. The rod is about 6 nm in diameter and 880 nm long and it is capped by two specialized proteins at one end and a different pair of proteins at the other end. The five coat proteins are encoded by five of the 11 genes present in these little viruses.[37,38]

Each coat subunit in the Ff viruses is coiled into an α-helical rod of 7 nm length. These are arranged in the virus in a right-handed helical pattern with a pitch of 1.5 nm and with 4.4 subunits per turn (Fig. 7-7). The protein rods are inclined to the helix axis and extend inward. This arrangement permits a “knobs-in-holes” hydrophobic bonding between subunits. The helix of pitch 1.5 nm is the **primary** or **one-start helix**. However, in every regular helical structure we can also trace a two-start helix, a three-start helix, etc. In this instance the five-start helix is easiest to see.

The protein coat of these viruses provides an elongated cylindrical cavity to protect the circular, single-stranded DNA molecule that is the genome. Although there are two antiparallel strands of DNA, a regular base-paired structure is impossible and the DNA is probably not present in a highly ordered form.[38a] There are about 2.4 nucleotides in the DNA per protein subunit. However, there are related viruses with ratios as low as one nucleotide per protein subunit and containing more highly extended DNA.[34,39]

A rod-shaped plant virus. The tobacco mosaic virus (Figs. 5-41, 7-8) is a 300-nm-long rod constructed from 2140 identical wedge-shaped subunits whose detailed molecular structure is known.[40] Each 158-residue subunit contains five helices and a small β sheet. A single strand of RNA containing 6395 nucleotides (~ 3 per protein subunit) lies coiled in a groove where it interacts with side chains from two of the helices (Fig. 7-8B).[41–44a] The virus is assembled by the binding of a region of the RNA 800 – 1000 nucleotides from the 3' end to a two-turn helix of subunits that appears to form spontaneously. Additional subunits then add at each end, binding to the RNA as well as the adjacent protein subunits.[42,44,45] A relative with a very similar structure is cucumber green mottle mosaic virus.[46]

Bacterial pili. The adhesion pili, or fimbriae,[48] of bacteria are also helical arrays of subunits. The P pili of *E. coli* are encoded by a cluster of 11 genes in the *pap* (pilus associated with pyelonephritis) cluster. They are needed to allow the bacteria to colonize the human urinary tract. The bulk of the ~1-µm-long pilus is made up of about 1000 subunits of a 185-residue protein encoded by gene *PapA.* They form a right-handed helix of ~7 nm diameter with 3.28 subunits per turn and a pitch of ~2.5 nm (Fig. 7-9A).[49–51] The rod is anchored to the bacterial outer membrane by a protein encoded by gene *PapH,* while subunits encoded by *PapE* and *PapF* fasten the adhesin protein (PapG) to the tips of the pili.[49,52,53,53a] The adhesin binds to the Galα1→4Gal ends of glycolipids in

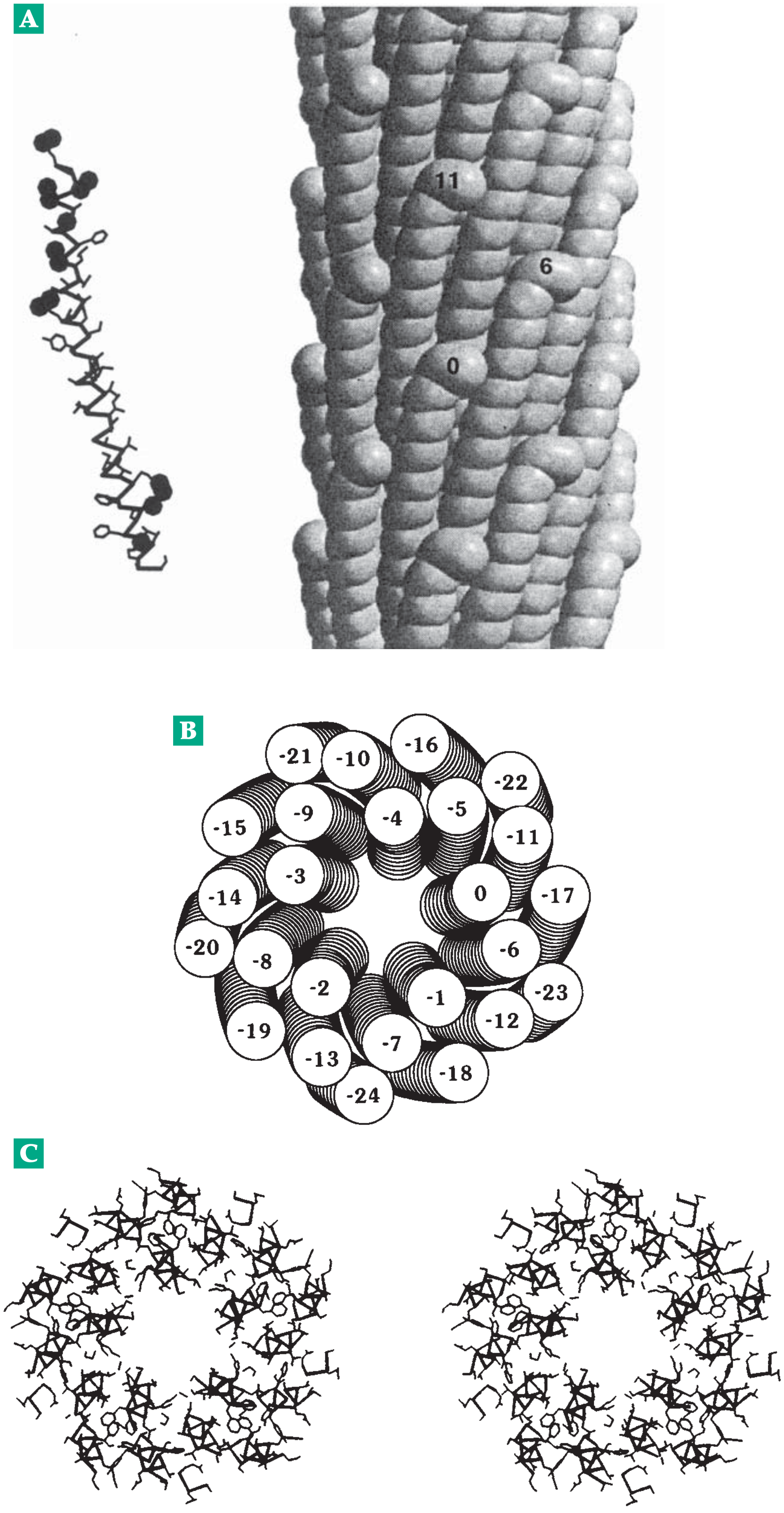


Figure 7-7 Structure of the virus fd protein sheath. (A) Left. A single coat subunit, with its N terminus towards the top, as if moved from the left side of the sheath. The dark circles represent charged atoms of Asp, Glu, and Lys side chains. The backbone of the protein is a C$^\alpha$ diagram. The positively charged atoms near the C terminus line the inner surface of the sheath neutralizing the negative charge of the DNA core. Right. Each subunit is represented by a helical tube through successive C$^\alpha$ atoms. Three nearest neighbors, indexed as 0, 6, and 11, are indicated. The axial slab shown represents ~1% of the total length of the virion. From Marvin.[31a] (B) A 2.0 nm section through the virus coat with the helices shown as curved cylinders. The view is down the axis from the N-terminal ends of the rods. The rods extend upward and outward. The rods with indices 0 to –4 start at the same level, forming a five-start helical array. The rods with more negative indices start at lower levels and are therefore further out when they are cut in this section. (C) The same view but with "wire models" of the atomic structure of the rods. From Marvin *et al.*[32]

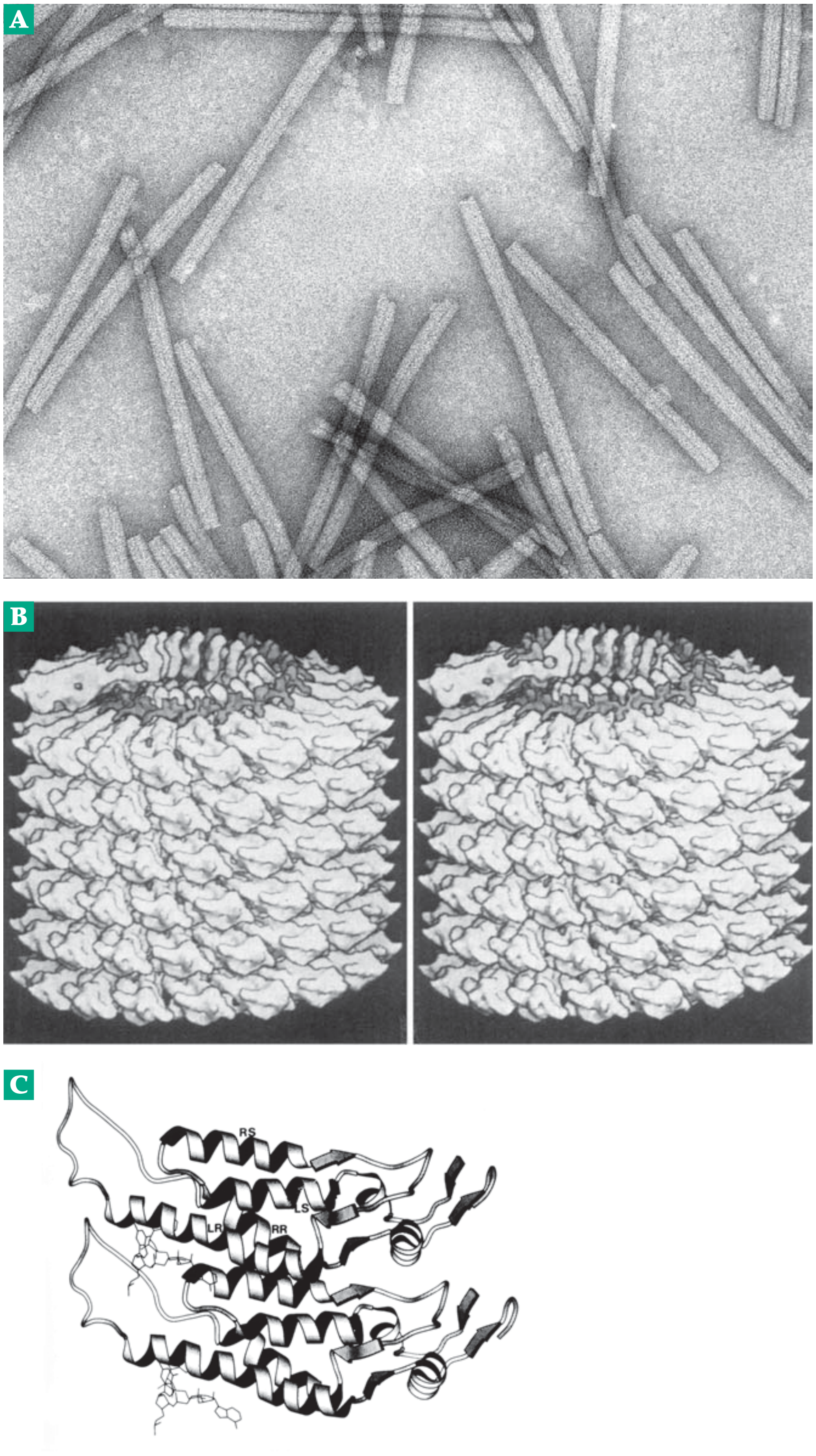


Figure 7-8 (A) Electron micrograph of the rod-shaped particles of tobacco mosaic virus. © Omikron, Photo Researchers. See also Butler and Klug.[42] (B) A stereoscopic computer graphics image of a segment of the 300 nm long tobacco mosaic virus. The diameter of the rod is 18 nm, the pitch of the helix is 2.3 nm, and there are 16 1\3 subunits per turn. The coat is formed from ~2140 identical 17.5-kDa subunits. The 6395-nucleotide genomic RNA is represented by the dark chain exposed at the top of the segment. The resolution is 0.4 nm. From Namba, Caspar, and Stubbs.[47] (C) A MolScript ribbon drawing of two stacked subunits. From Wang and Stubbs.[46]

the kidney.[51,54] The PapE, F, and G subunits form a thin ~2.5-nm-thick by 15 nm "fibrillum" which is attached by an adapter protein encoded by gene *PapK*. A special chaperonin (*PapD* gene) is also required for pilus assembly in *E. coli*[53a] as well as other bacteria.[55] Another *E. coli* pilus adheres to mannose oligosaccharides.[53b]

Similar pili of *Neisseria gonnorheae* are used by that bacterium. The three-dimensional structure of the 158-residue pilin subunit is that of a globular subunit with an 8.5 nm α-helical spine at one end.[56,57] A proposed model of the intact pilus shaft is shown in Fig. 7-9B,C. Notice the similarity of the packing of the α-helical spines in the center to the packing arrangement in the bacteriophage coat in Fig. 7-7. Similar features may be present in the P pilus rod shown in Fig. 7-9A. However, there is uncertainty about the packing arrangement. The *E. coli* type 1 pilus subunits contain immunoglobulin folds that are completed by donation of an N-terminal strand from a neighboring subunit.[53a] In thin fimbriae of *Salmonella* extended, parallel β helices may be formed (see Fig. 2-17)[53e,53f]

Other types of pili are also well-known.[53c,d] F pili or conjugative pili are essential for sexual transfer of DNA between bacterial cells (Chapter 26). F^+ strains of *E. coli* form hollow pili of 8.5 nm diameter with a 2.0-nm central hole.[58,59] Their 90-residue subunits apparently form rotationally symmetric pentamers which stack to form the pili.[59] These pili are essential to establishing the initial contact between conjugating bacterial cells.

The thin filaments of muscle.

An essential component of skeletal muscle (discussed further in Chapter 19) is filamentous **actin** (F-actin). It is composed of 375-residue globular subunits of a single type and with a highly conserved sequence.[60,61] It is found not only in muscle but also in other cells where it is a component of the cytoskeleton. The actin microfilament has the geometry of a left-

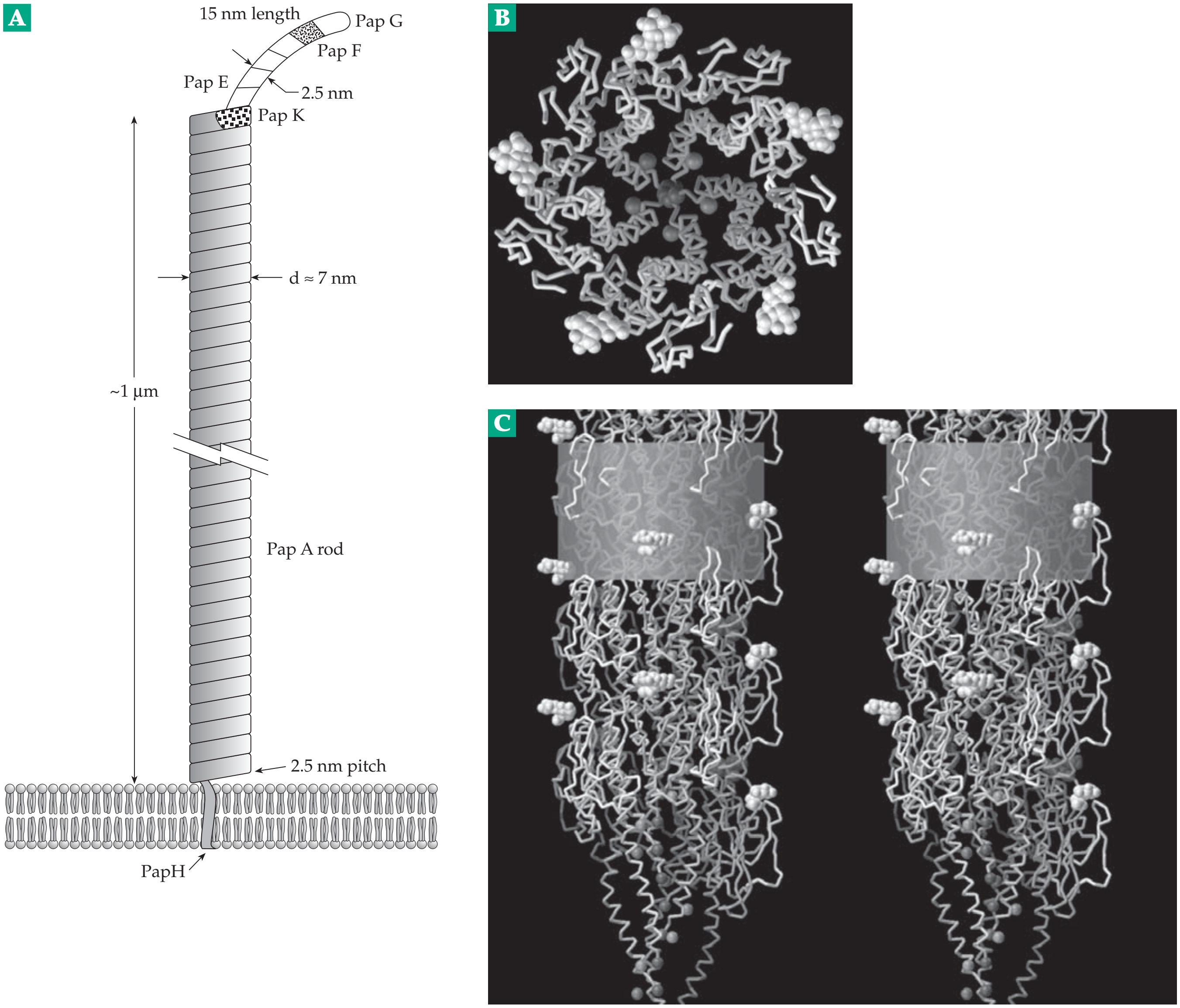


Figure 7-9 (A) Schematic diagram of a bacterial P pilus. The ~1-µm-long helical rod is anchored to the outer cell membrane by protein Pap H. The adhesion Pap G binds to galactosyl glycolipids of the host. (B, C) The structure of pilus fiber from *Neisseria gonnorrheae* modeled from the atomic structure of the 54-residue pilin subunit. The exact structure of the fiber is uncertain, but the model generated here by trying various possible helical packings matches the dimensions obtained from fiber diffraction patterns and electron microscopic images. (B) Cross section. (C) Stereoscopic view. The experimental dimensions of 4.1-nm pitch and 6.0-nm diameters are shown by the "transparent" ring in (B). From Parge *et al.*[57]

---

handed one-start or **primary helix** with a pitch of only 0.54 nm and with approximately two subunits per turn (Fig. 7-10).[62,63]
It can also be described as a right-handed two-start helix in which *two chains* of subunits coil around one another with a long pitch (Fig. 7-10).

2. Oligomers with Twofold (Dyad) Axes

Paired interactions. If two subunits are held together with interactions *aj* and are related by a twofold axis of rotation as shown in Fig. 7-11, we obtain an **isologous dimer**. Each point such as *a* in one subunit is related to the same point in the other subunit by reflection through the axis of rotation. In the center, along the twofold axis, points *c* and *c′* are *directly opposite the same points* in the other subunit. Figure 7-11

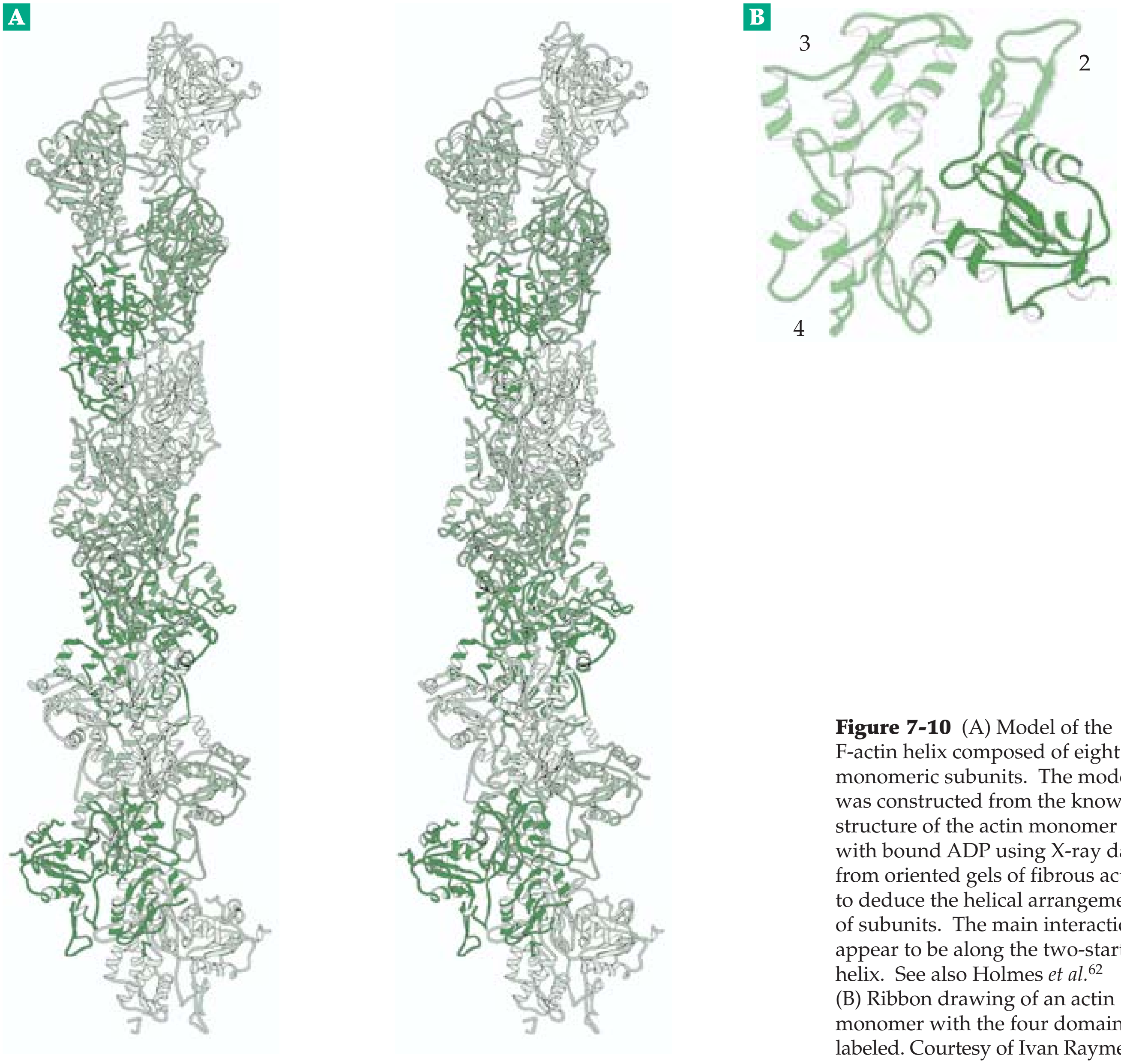


Figure 7-10 (A) Model of the F-actin helix composed of eight monomeric subunits. The model was constructed from the known structure of the actin monomer with bound ADP using X-ray data from oriented gels of fibrous actin to deduce the helical arrangement of subunits. The main interactions appear to be along the two-start helix. See also Holmes *et al.*[62] (B) Ribbon drawing of an actin monomer with the four domains labeled. Courtesy of Ivan Rayment.

is drawn with a hole in the center so that groups *c* and *c′* do not actually touch, and it is the paired interactions such as *aj* of groups not adjacent to the axis that contribute most to the bonding. However, a real protein dimer may or may not have such a hole. The pair of identical interactions in an isologous dimer may be referred to as a single **isologous bond**. Such a bond always contains the paired interactions between complementary groups (*aj*) and has pairs of identical groups along the axis. However, because they are identical those groups usually cannot interact in a specific complementary manner.

Isologous bonding is very important in oligomeric enzymes, and it has been suggested that isologous interactions evolved early. Initially there may not have been much complementarity in the bonding but two "hydrophobic spots" on the surface of the subunits came together in a nonspecific association.[64] Later in evolution the more specific paired interactions could have been added.

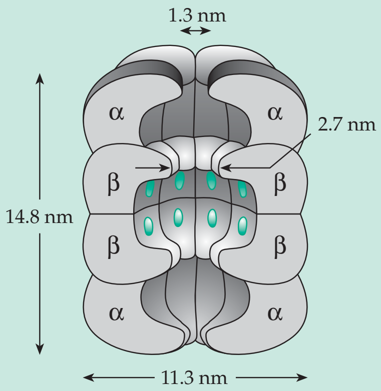
Dihedral symmetry. Isologous dimers can serve as subunits in the formation of larger closed oligomers and helices; for example, an isologous pair of the sort shown in Fig. 7-11A can be flipped over onto the top of another similar pair as shown in Figs. 7-11B and 7-11C. Again, if the proper complementary surfaces exist, bonds can form as shown (*bk* in Fig. 7-11B and *bk* and *cl* in Fig. 7-11C). Both structures in Figs. 7-11B and 7-11C possess dihedral (D_2) symmetry.[65] In addition to the twofold axis of rotation lying perpendicular to the two rings, there are two other twofold axes of rotation as

BOX 7-A LIFE AND DEATH FOR PROTEINS: CHAPERONINS AND PROTEASOMES

In 1968, a tiny cylindrical particle, which appeared to be a stack of 11-nm rings, was observed by electron microscopy of an extract of erythrocytes.[a,b] Later, a similar particle was found in both the nucleus and the cytoplasm of other cells of many organisms. The particles were soon recognized as a new type of protein-hydrolyzing enzyme, a large 700-kDa particle consisting of 20–30 subunits of several different types which came to be known as the **multicatalytic protease** or **20S proteasome**.[c,d] Electron microscopy and X-ray diffraction showed that the particle is formed from four stacked rings, each of which consists of seven subunits whose molecular masses range from 21 to 31 kDa.[e–j]

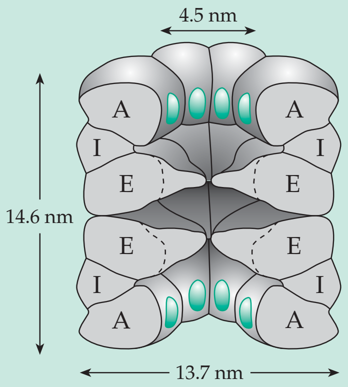
Proteasomes are strikingly similar in architecture, though not in peptide sequences, to another particle found in both bacteria and eukaryotes: a molecular "chaperone" or **chaperonin**. The chaperonins, of which there are several types, protect proteins while they fold or undergo translocation within cells.[k] One of the best studied members is the *E. coli* protein **GroEL**, which is also composed of double rings of 14 subunits with seven-fold rotational symmetry and with two of these assemblies associated back-to-back with dihedral symmetry.[l] The dimensions of GroEL and 20S proteasomes are nearly the same. However, GroEL has only two rings of ~60-kDa subunits, more than twice the size of proteosomal subunits. The accompanying sketch illustrates this fact and also the basic structural similarity of 20S proteosomes with GroEL. The αβ pairs of the proteosome, correspond to single subunits of the chaperonin, but these subunits have three distinct domains–apical, intermediate, and equatorial (labeled A, I, and E , respectively, in the drawing).[m] After a protein, whether correctly, incorrectly, or only partially folded, enters a cavity in GroEL, a second protein **GroES** of smaller size (~10 kDa) but with seven-fold symmetry binds to one end of the chaperonin.[p] Seven molecules of ATP also bind to sites on the GroEL ring to which GroES binds (the *cis* ring). The binding of the ATP and GroES evidently induces a major conformational change in the GroEL subunits[l,m,n,q] which causes the binding cavity to expand to over twice the original volume. This change (see drawing) also causes hydrophobic surfaces of the cavity to become buried and hydrophilic side chains to be exposed. The cavity surface was initially largely hydrophobic and able to bind many proteins nonspecifically, but upon expansion it becomes hydrophilic and less likely to bind. This releases the encased protein to complete its folding or to partially unfold and refold without interference from other proteins.

While a protein is adjusting its folding in the *cis* compartment another protein molecule may become trapped in the *trans* compartment. After some time the bound ATP molecules are hydrolyzed. As in the contraction of muscle, which is discussed in Chapter 19, the loss of inorganic phosphate (P_i) and ADP from the active site can be accompanied by movement. In the chapenonin this involves a conformational switch so that the ES heptamer is released and the conformation of the *trans* ring of EL is switched to that of the initial *cis* ring and vice versa. The new *cis* ring is ready to receive an ES cap and the new *trans* ring can release the folded protein.[r] A variety of experimental approaches are being used in an



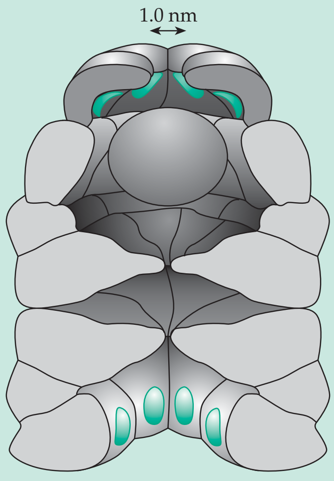
20*S* Proteasome

● Catalytic sites



GroEL Chaperonin

● Hydrophobic binding patches



Expanded GroEL-GroES complex with enclosed folding polypeptide[o]

From Weissman, Sigler and Horwich[g]. Illustrations by K. Sutliff.
Right: From Mayhew and Hartl[o]

BOX 7-A LIFE AND DEATH FOR PROTEINS: CHAPERONINS AND PROTEASOMES (continued)

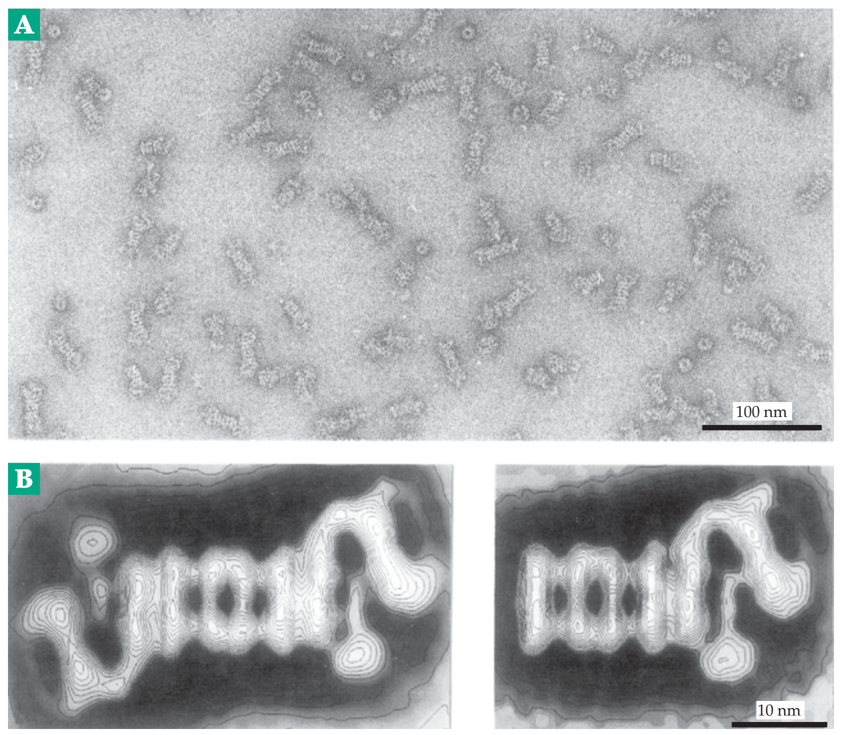
effort to further understand the action of GroEL.[s–x] The chaperonin may function repeatedly before a protein becomes properly folded.[t]

While chaperonins assist proteins to fold correctly proteasomes destroy unfolded chains by partial hydrolysis, cutting the chains into a random assortment of pieces from 3 to 30 residues in length with an average length of ~ 8 residues.[y] Proteasomes destroy not only unfolded and improperly folded proteins but also proteins marked for destruction by the ubiquitin system described in Box 10-C. It has been hard to locate true proteosomes in most bacteria. However, they do contain protease particles with similar characteristics[z–bb] and archaeons, such as *Thermoplasma acidophilum,* have proteasomes similar to those of eukaryotes.[cc]

The *Thermoplasma* proteasome contains only two kinds of subunits, α and β, which have similar amino acid sequences. These form α_7 and β_7 rings which associate in $\alpha_7\beta_7$ pairs with two of these double rings stacked back-to-back with dihedral D7 symmetry: $\alpha_7\beta_7\beta_7\alpha_7$. The crystal structure has been determined for this 20S proteasome from *T. acidophilum*[g,dd] and for the corresponding proteasome from yeast (*Saccharomyces cerevisiae*).[f,ee] The accompanying drawings illustrate top and side views of the *T. acidophilum* proteasome. The particle contains three internal cavities. The outer two are formed between the α_7 and β_7 rings and the inner is formed between the two β_7 rings. A channel only 1.3 nm in diameter permits the entrance of peptide chains into the compartments.

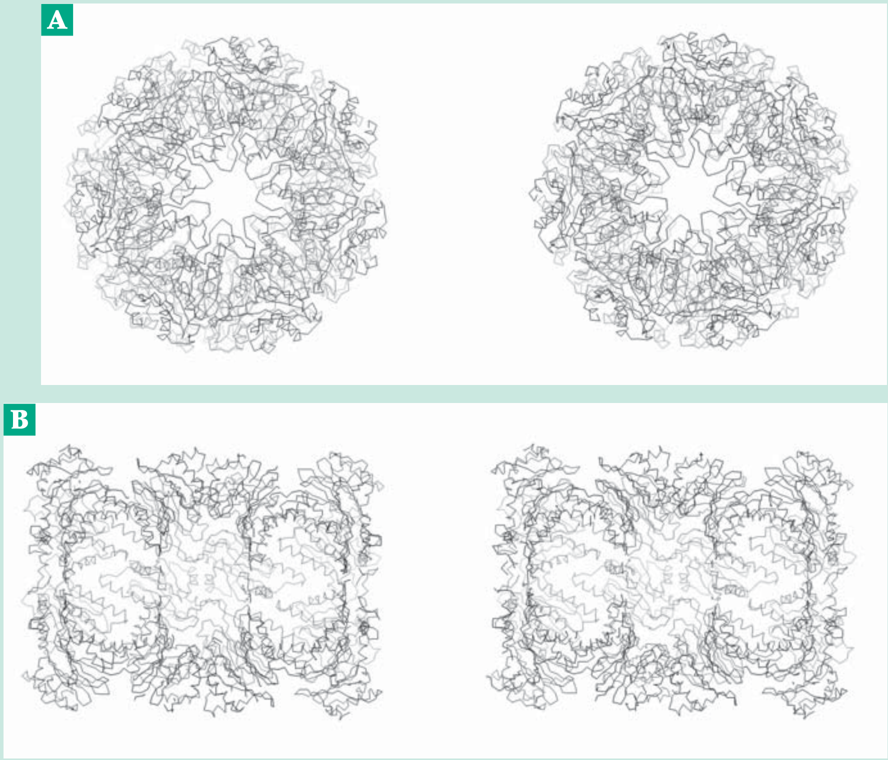
The active sites of the enzymes[ff] are located in the β subunits in the central cavity.[dd] While the yeast and human proteasomes are similar to those of *Thermoplasma,* the β subunits consist of seven different protein-hydrolyzing enzymes whose catalytic activities and mechanisms are considered in Chapter 12. There are also seven different α subunits, all of whose sequences are known.[i,gg] To make the story more complex, additional subunits, some of which catalyze ATP hydrolysis, form a 600- to 700-kDa cap which adds to one or both ends of a 20S proteosome to give a larger **26S proteasome**.[b,d,e,w] These larger proteasomes carry out an ATP-dependent cleavage of proteins selected for degradation by the ubiquitin system (Box 10-C; Chapter 12). Some of the short peptide segments formed by proteasomes may leave cells and participate in intercellular communication. For example, pieces of antigenic peptides are used by cells of the immune system for "antigen presentation" (Chapter 31),[hh] an important process by which the immune system recognizes which cells are "self" and which are foreign or malignant and must be killed.

The structure of the caps on the 26S proteasome ends is complex. At least 20 different regulatory subunits have been identified.[ii,jj]



(A) Electron micrograph of 26S proteasomes from *Xenopus* oocytes negatively stained with 2% uranyl acetate. (B) Image of the 26S proteasome (left) and a 20S proteasome with only one end cap. These views were obtained by correlation averaging of 527 individual images of the 26S proteasome and 395 images of the single-ended form. From Peters *et al*.[e] Courtesy of Wolfgang Baumeister.

BOX 7-A (continued)



(A) Top view of the 20S proteasome as an α-carbon plot showing the seven-fold symmetry. The α subunits are in front of the β subunits. (B) Side view showing the proteasome cut open along its seven-fold axis. From Löwe *et al.*[dd] Courtesy of Robert Huber.

---

[a] Harris, J. R. (1968) *Biochim. Biophys. Acta.* **150**, 534–537

[b] Peters, J.-M. (1994) *Trends Biochem. Sci.* **19**, 377–382

[c] Bosch, G., Baumeister, W., and Essen, L.-O. (2000) *J. Mol. Biol.* **301**, 19–25

[d] Ferrell, K., Wilkinson, C. R. M., Dubiel, W., and Gordon, C. (2000) *Trends Biochem. Sci.* **25**, 83–88

[e] Peters, J.-M., Cejka, Z., Harris, J. R., Kleinschmidt, J. A., and Baumeister, W. (1993) *J. Mol. Biol.* **234**, 932–937

[f] Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D., and Huber, R. (1997) *Nature (London)* **386**, 463–471

[g] Weissman, J. S., Sigler, P. B., and Horwich, A. L. (1995) *Science* **268**, 523–524

[h] DeMartino, G. N., and Slaughter, C. A. (1999) *J. Biol. Chem.* **274**, 22123–22126

[i] Gerards, W. L. H., de Jong, W. W., Bloemendal, H., and Boelens, W. (1998) *J. Mol. Biol.* **275**, 113–121

[j] Schmidtke, G., Schmidt, M., and Kloetzel, P.-M. (1997) *J. Mol. Biol.* **268**, 95–106

[k] Netzer, W. J., and Hartl, F. U. (1998) *Trends Biochem. Sci.* **23**, 68–73

[l] Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A., and Horwich, A. L. (1998) *Ann. Rev. Biochem.* **67**, 581–608

[m] Kawata, Y., Kawagoe, M., Hongo, K., Miyazaki, T., Higurashi, T., Mizobata, T., and Nagai, J. (1999) *Biochemistry* **38**, 15731–15740

[n] Betancourt, M. R., and Thirumalai, D. (1999) *J. Mol. Biol.* **287**, 627–644

[o] Mayhew, M., and Hartl, F. U. (1996) *Science* **271**, 161–162

[p] Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L., and Deisenhofer, J. (1996) *Nature (London)* **379**, 37–45

[q] Cliff, M. J., Kad, N. M., Hay, N., Lund, P. A., Webb, M. R., Burston, S. G., and Clarke, A. R. (1999) *J. Mol. Biol.* **293**, 667–684

[r] Rye, H. S., Burston, S. G., Fenton, W. A., Beechem, J. M., Xu, Z., Sigler, P. B., and Horwich, A. L. (1997) *Nature (London)* **388**, 792–798

[s] Chatellier, J., Buckle, A. M., and Fersht, A. R. (1999) *J. Mol. Biol.* **292**, 163–172

[t] Nieba-Axmann, S. E., Ottiger, M., Wüthrich, K., and Plückthun, A. (1997) *J. Mol. Biol.* **271**, 803–818

[u] Gervasoni, P., Gehrig, P., and Plückthun, A. (1998) *J. Mol. Biol.* **275**, 663–675

[v] Torella, C., Mattingly, J. R., Jr., Artigues, A., Iriarte, A., and Martinez-Carrion, M. (1998) *J. Biol. Chem.* **273**, 3915–3925

[w] Horwich, A. L., Weber-Ban, E. U., and Finley, D. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11033–11040

[x] Buckle, A. M., Zahn, R., and Fersht, A. R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3571–3575

[y] Kisselev, A. F., Akopian, T. N., and Goldberg, A. L. (1998) *J. Biol. Chem.* **273**, 1982–1989

[z] Kessel, M., Maurizi, M. R., Kim, B., Kocsis, E., Trus, B. L., Singh, S. K., and Steven, A. C. (1995) *J. Mol. Biol.* **250**, 587–594

[aa] Shin, D. H., Lee, C. S., Chung, C. H., and Suh, S. W. (1996) *J. Mol. Biol.* **262**, 71–76

[bb] Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F., and Baumeister, W. (1996) *Science* **274**, 1385–1389

[cc] Maupin-Furlow, J. A., and Ferry, J. G. (1995) *J. Biol. Chem.* **270**, 28617–28622

[dd] Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) *Science* **268**, 533–539

[ee] Stuart, D. I., and Jones, E. Y. (1997) *Nature (London)* **386**, 437–438

[ff] Voges, D., Zwickl, P., and Baumeister, W. (1999) *Ann. Rev. Biochem.* **68**, 1015–1068

[gg] Schmidt, M., and Kloetzel, P.-M. (1997) *FASEB J.* **11**, 1235–1243

[hh] Goldberg, A. L., and Rock, K. L. (1992) *Nature (London)* **357**, 375–379

[ii] Adams, G. M., Falke, S., Goldberg, A. L., Slaughter, C. A., DeMartino, G. N., and Gogol, E. P. (1997) *J. Mol. Biol.* **273**, 646–657

[jj] Knowlton, J. R., Johnston, S. C., Whitby, F. G., Realini, C., Zhang, Z., Rechsteiner, M., and Hill, C. P. (1997) *Nature (London)* **390**, 639–643

indicated in the drawings. Again, the interactions are paired and isologous; of many possible contacts two *bk* interactions and two *cl* interactions are marked for each pair of subunits in Fig. 7-11. There are a total of six pairs of these interactions, one between each combination of two subunits. This may be a little more difficult to see in Fig. 7-11B than in Fig. 7-11C because in the former the subunits are arranged in a more or less square configuration. Nevertheless, a pair of interactions between the left-hand subunit in the top ring and the subunit in the lower ring at the right does exist, even if it is only electrostatic and at a distance. An example of a tetrameric enzyme with perfect dihedral symmetry of the type shown in Fig. 7-11B is **lactate dehydrogenase** (Chapter 15). The plant agglutinin **concanavalin A** has a quaternary structure resembling that in Fig. 7-11C.

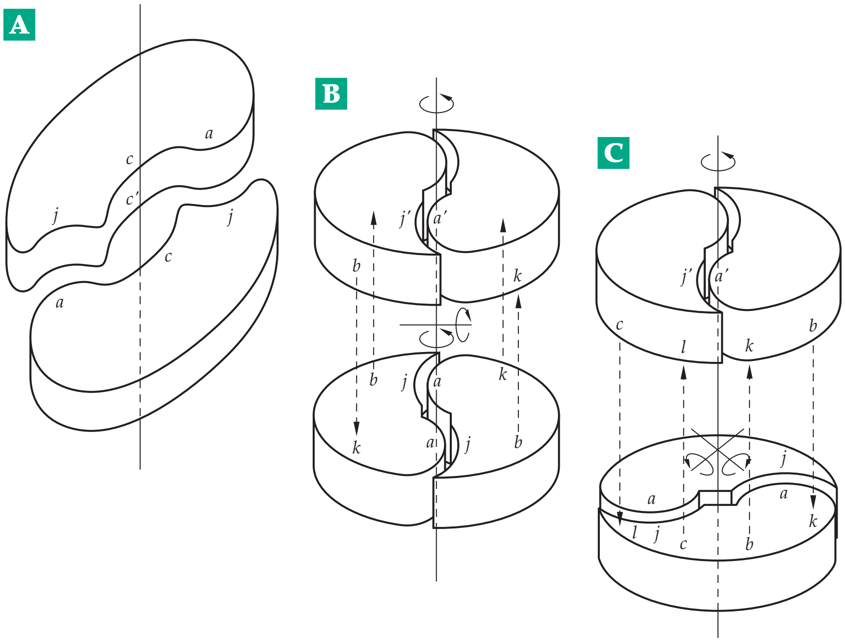


Figure 7-11 (A) Isologous bonding between pairs of subunits; (B) an "isologous square" arrangement of subunits; (C) an apparently "tetrahedral" arrangement of subunits. Note the three twofold axes.

Square arrays of four subunits can be formed using either heterologous or isologous interactions. Both types of bonding can occur in larger aggregates. For example, two trimers such as that shown in Eq. 7-25 can associate to a hexamer having dihedral (D_3) symmetry; a heterologous "square tetramer" can dimerize to give a dihedral (D_4) octamer.[65] The enzymes **ornithine decarboxylase** (Fig. 7-12)[66] and **glutamine synthetase** (Chapter 24)[67] each consist of double rings of six subunits each. The upper ring is flipped over onto the lower giving dihedral symmetry (D_6) with one 6-fold axis and six 2-fold axes at right angles to it.

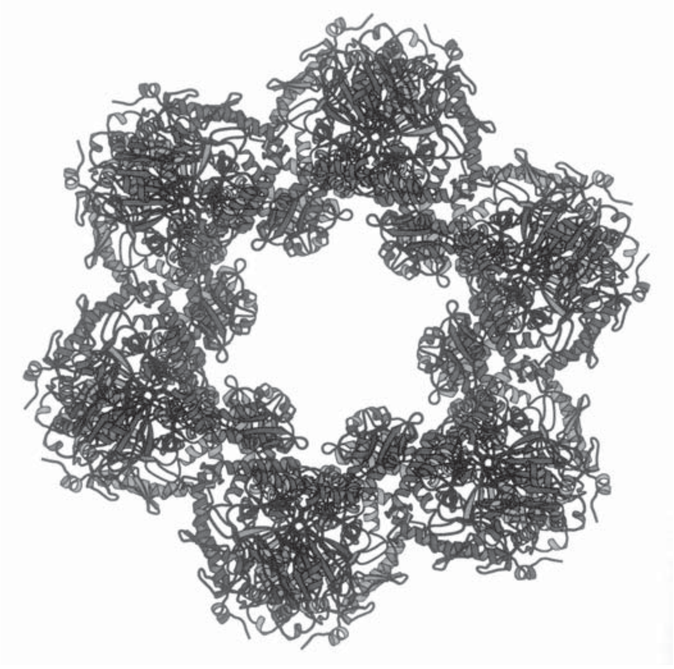


Figure 7-12 A ribbon representation of the ornithine decarboxylase dodecamer. Six dimers of the 730-residue subunits are related by C_6 crystallographic symmetry. MolScript drawing from Momany *et al.*[66] Courtesy of Marvin Hackert.

Oligomers with cubic symmetry (polyhedra). Symmetrical arrangements containing more than one axis of rotation of order higher than 2-fold are said to have cubic symmetry. The **tetrahedron** is the simplest example. It contains four 3-fold axes which pass through the vertices and the centers of the faces and three 2-fold axes which pass through the midpoints of the six edges. *Since protein subunits are always asymmetric, a tetrameric protein cannot possess cubic symmetry*. As we have already seen, tetrameric enzymes have dihedral symmetry. However, a heterologous trimer with 3-fold symmetry can form a face of a tetrahedron containing a total of 12 asymmetric subunits. Twenty-four subunits can interact to form a **cube**. Three 4-fold axes pass through the centers of the faces, four 3-fold axes pass through the vertices, and six 2-fold axes pass through the edges (see Figs. 7-13 and 16-3).

The largest structure of cubic symmetry that can

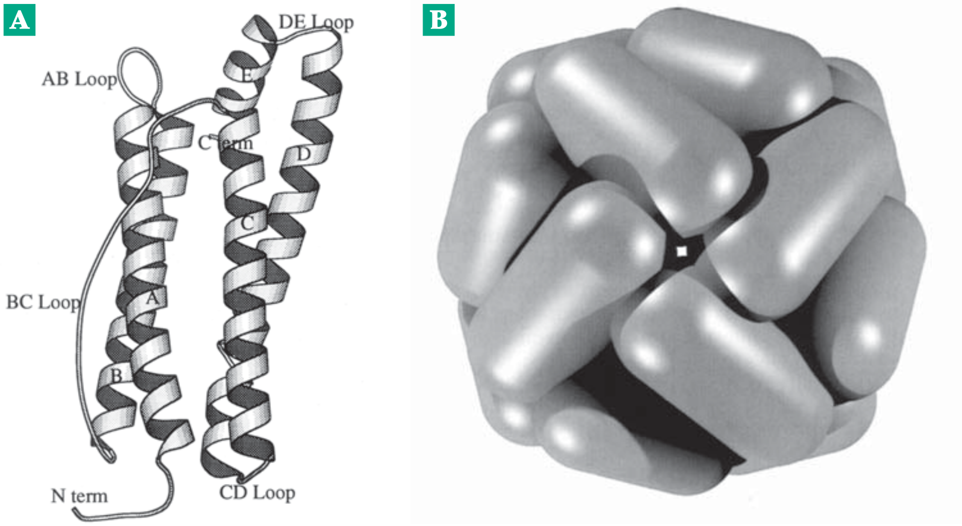


Figure 7-13 (A) MolScript ribbon drawing of a subunit of the iron oxide storage proteins L-ferritin from amphibian red cells. This 4-helix bundle is represented by cylinders of 1.3 nm diameter in the oligomer. (B) Helices A and C of the monomer are on the outer surfaces of the oligomer and helices B and D are on the inner surface. The oligomer consists of a shell of 24 subunits and is viewed down a 4-fold axis illustrating its 423 (cubic) symmetry. The molecule is illustrated further in Fig. 16-3. From Trikha *et al.*[74] Courtesy of Elizabeth Theil.

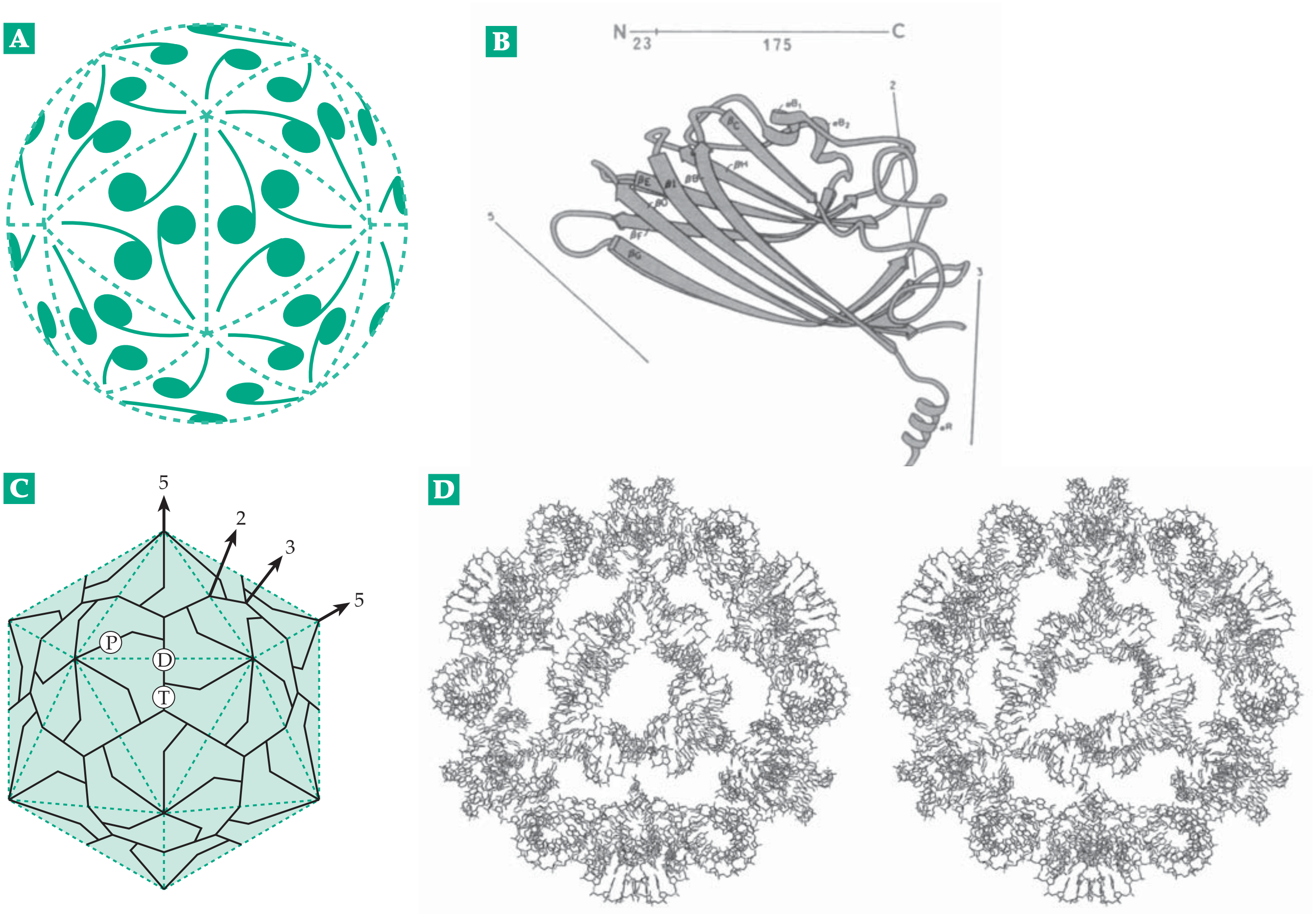


Figure 7-14 (A) Schematic drawing illustrating an icosahedrally symmetric structure with sixty identical asymmetric subunits all in equivalent positions. The 5-fold axes are located at the vertices of the icosohedron and the 2-fold and 3-fold axes can readily be seen. (B) Ribbon drawing of the 195-residue polypeptide chain of the coat subunit of satellite tobacco necrosis virus. The protein folds into an inwardly projecting N-terminal segment and a "β-jellyroll" domain. The packing of this subunit in the virus particle is shown schematically in (C). The symmetry axes drawn next to the subunit diagram (B) can be used to position it in the structure. Contacts between subunits are labeled D, T, and P ('dimer', 'trimer', 'pentamer'). Diagrams courtesy of Drs. Strandberg, Liljas, and Harrison.[68] (D) The distribution of RNA helical segments in a hemisphere of a virion of a similar small virus, the satellite tobacco mosaic virus. The virion viewed down a 3-fold axis from the virus exterior. The helical axes of the RNA segments are along icosahedral edges. From Larson *et al.*[75]

be made is the **icosahedron**, a regular solid with 20 triangular faces. Sixty subunits, or some multiple of 60, are required and at each vertex they form a heterologous pentamer. As with the tetrahedron, each face contains a heterologous trimer, while isologous bonds across the edges form dimers (Fig. 7-14C). Many viruses consist of roughly spherical protein shells (coats) containing DNA or RNA inside.[68–70] As with the filamentous viruses, the protein coats consist of many identical subunits, a fact that can be rationalized in terms of economy from the genetic viewpoint. *Only one gene is needed to specify the structure of a large number of subunits*.[70,71] Under the electron microscope the viruses often have an icosahedral appearance (Figs. 5-41A, 7-14), and chemical studies show that the number of the most abundant subunits is usually a multiple of 60. An example is the tiny **satellite tobacco necrosis virus**,[72] diameter ~18 nm, whose coat contains just 60 subunits of a 195-residue protein. Its genome is a 1239-nucleotide molecule of RNA. The structure of the coat has been determined to 0.25-nm resolution.[73]

Many virus coats have 180 subunits or a number that is some other multiple of 60. However, in these coats the subunits cannot all be in identical environments. Two cases may be distinguished. If all of the subunits have identical amino acid sequences they probably exist in more than one distinct conformation that permit them to pack efficiently. (Next section) Alternatively, two or more subunits of differing sequence and structure may associate to form 60 larger subunits that do pack with icosahedral symmetry. For example, the polioviruses (diameter 25 nm) contain three major coat proteins (α, β, and γ or VP1, VP2, and VP3). These are formed by cleavage of a large precursor protein into at least four pieces.[76,77] The three largest pieces of ~33-, 30-, and 25-kDa mass (306, 272, and 238 residues, respectively) aggregate as $(\alpha\beta\gamma)_{60}$. Sixty copies of a fourth subunit of 60 residues are found within the shell. Related picorna viruses such as human rhinoviruses (Fig. 7-15),[69,78,79] foot-and-mouth disease virus, parvovirus,[80] and Mengo virus[81] have similar architectures. The small (diameter 25 nm) single-stranded DNA bacteriophages such as ϕX174 also have three different coat proteins, one of which forms small hollow spikes at the vertices of the icosahedral shell (Fig. 5-41A).[82]

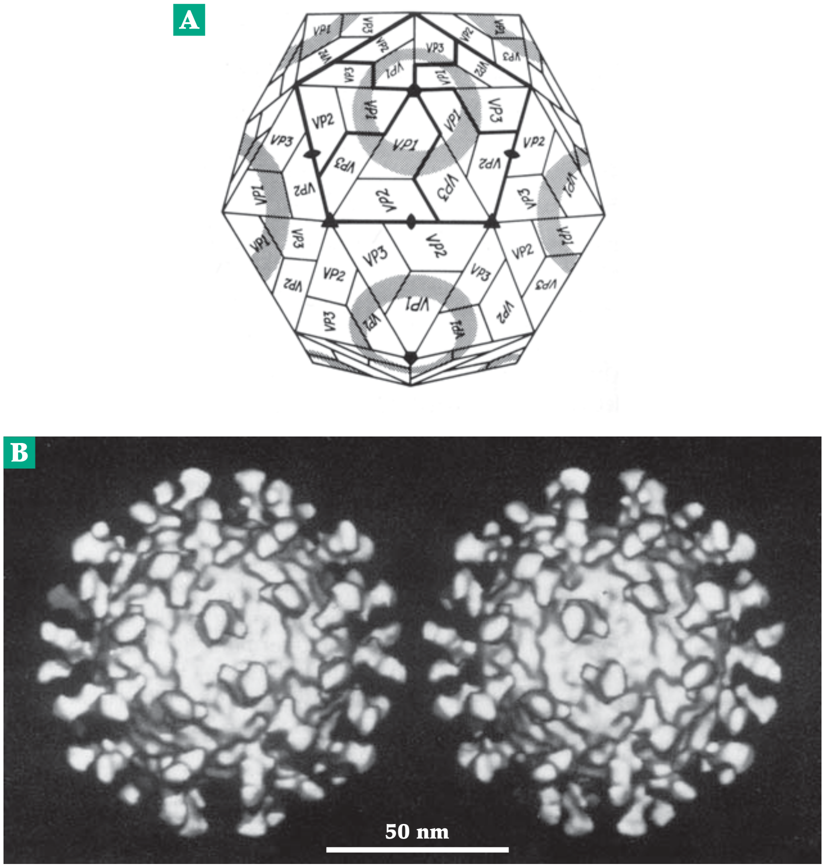


Figure 7-15 (A) Schematic diagram of the icosahedral shell of a human rhinovirus showing the arrangement of the three subunits VP1, VP2, and VP3, each present as 60 copies. (B) Stereoscopic view of an image of the virus "decorated" by the binding of two immunoglobulinlike domains of the intercellular adhesion molecule ICAM-1, a natural receptor for the virus. Part of this receptor binds into a groove or "canyon," which in marked in (A) by the dark bands. From Olson *et al.*[78] Courtesy of Michael Rossmann.

Asymmetry and quasi-equivalence in oligomers. It is natural to think about association of subunits in symmetric ways. Consequently the observation of square, pentagonal, and hexagonal arrangements of subunits directly with the electron microscope led to a ready acceptance of the idea that protomers tend to associate symmetrically. However, consider the predicament of the two molecules shown in Fig. 7-16. They might get together to form an isologous dimer if it were not for the fact that their "noses" are in the way. Despite the obvious steric hindrance, an isologous dimer can be formed in this case if one subunit is able to undergo a small change in conformation (Fig. 7-16). In the resulting dimer the two subunits are only **quasi-equivalent**.

Unsymmetrical dimerization of proteins appears to be a common phenomenon that is often observed in protein crystals. For example, the

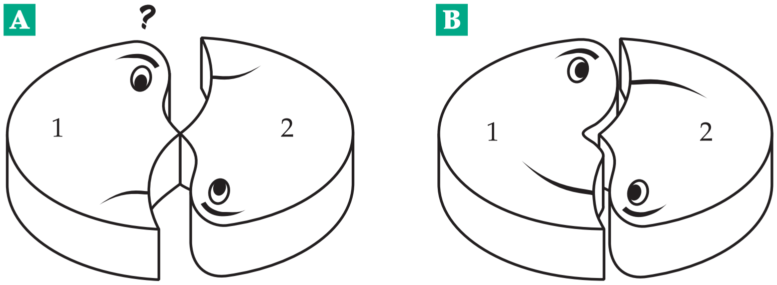


Figure 7-16 Nonsymmetric bonding in a dimer. (A) Two molecules which cannot dimerize because of a bad fit at the center. (B) A solution: Molecule 1 has refolded its peptide chain a little, changing shape enough to fit to molecule 2.

enzymes malic dehydrogenase and glyceraldehyde phosphate dehydrogenase (Chapter 15) are both tetramers of approximate dihedral symmetry but X-ray crystallography revealed distinct asymmetries[83,84] which include a weaker binding of the coenzyme NAD^+ in one subunit. This may simply reflect differences in environment within the crystal lattice. However, negative cooperativity in coenzyme binds has also been revealed by kinetic experiments.[85]

The polypeptide hormone **insulin** is a small protein made up of two chains (designated A and B) which are held together by disulfide bridges (Fig. 7-17A). Figure 7-17B is a sketch of the structure as revealed by X-ray crystallography,[86,87] with only the backbone of the peptide chains and a few side chains shown. In the drawing, the B chain lies behind the A chain. Beginning with the N-terminal Phe 1 of the B chain the peptide backbone makes a broad curve, and then falls into an helix of three turns lying more or less in the center of the molecule. After a sharp turn, it continues upward on the left side of the drawing in a nearly completely extended β structure. The A chain has an overall U shape with two roughly helical portions. The U shape is partly maintained by a disulfide bridge running between two parts of the A chain. Two disulfide bridges hold the A and B chains together, and hydrophobic bonding of internal side chain groups helps to stabilize the molecule.

Insulin in solution dimerizes readily, the subunits occupying quasi-equivalent positions. Figure 7-17C shows some details of the bonding between the subunits in the insulin dimer with a view from the outside of the molecule down the 2-fold axis (marked by the X in the center of the Phe 25 ring in the right-hand chain) through the dimer. The C-terminal ends of the B chains are seen in an extended conformation. The two antiparallel chains form a β structure with two pairs of hydrogen bonds. If there were perfect isologous bonding, the two pairs would be entirely equivalent and symmetrically related one to the other. A straight line drawn from a position in one chain and passing through the twofold axis (X) would also pass approximately through the corresponding position in the other chain. However, there are many deviations from perfect symmetry, the most striking of which is at the center where the Phe 25 from the right-hand chain projects upward and to the left. If the symmetry were perfect the corresponding side chain from the left-hand chain would project upward and to the right and the two phenylalanines would collide, exactly as do the "noses" in Fig. 7-16. In insulin one phenylalanine side chain has been flipped back out of the way.

Under proper conditions, three insulin dimers associate to form a hexamer of approximate dihedral (D_3) symmetry that is stabilized by the presence of two zinc ions. Figure 7-17D is a crude sketch of the hexamer showing the three dimers, the 3-fold axis of symmetry, and the two pseudo 2-fold axes, one passing between the two subunits of the dimer and the other between two adjacent dimers. Figure 7-18 is a stereoscopic ribbon diagram of the atomic structure, with the A chains omitted, as obtained by X-ray diffraction.[87] The structure has also been obtained by NMR spectroscopy.[88] Note that each of the two zinc atoms lies on the threefold axis and is bound by three imidazole rings from histidines B-10. The significance of the zinc binding is uncertain but these hexamers readily form rhombohedral crystals, even within the pancreatic cells that synthesize insulin. The structure illustrates a feature that is common to many oligomers of circular or dihedral symmetry. A central "channel" is often quite open and protruding side chain groups, such as the imidazole groups in insulin, form handy nests into which ions or molecules regulating activity of proteins can fit. Conformational differences in insulin are induced by the binding of phenol. In Fig. 7-18A the C-terminal ends of the chains are extended but in the phenol complex (B) they have coiled to extend the α helices.

Quasi-equivalence in virus coats. A large number of icosahedral viruses have coats consisting of 180 identical subunits. For example, the small RNA-containing bacteriophage MS2 consists of an eicosahedral shell of 180 copies of a 129-residue protein that encloses one molecule of a 3569-residue RNA.[89] The virus also contains a single molecule of a 44-kDa protein, the A protein, which binds the virus to a bacterial pilus to initiate infection. Related bacteriophages GA, fr, f2, and Qβ[90,91] have a similar architecture. Many RNA-containing viruses of plants also have 180 subunits in their coats.[68] Much studied are the **tomato bushy stunt virus** (diameter ~33 nm, 40-kDa subunits),[68] and the related **southern bean mosaic virus**.[92] The human **wart virus** (diameter ~ 56 nm) contains 420 subunits, seven times the number in a regular icosahedron. **Adenoviruses** (diameter ~100 nm) have 1500 subunits, 25 times more than the 60 in a regular icosahedron.[93,94] Caspar and Klug[95] proposed a theory of quasi-equivalence of subunits according to which the distances between the centers of subunits are preserved in a family of **icosadeltahedra** containing subunits in multiples of 20. However, the angles must vary somewhat from those in a regular icosahedron (compare with geodesic shells in which the angles are constant but the distances are not all the same). The resulting polyhedra contain hexamers as well as pentamers at vertices; for example, the shells of the 180-subunit viruses contain clusters of subunits forming 12 pentamers and 20 hexamers. There are

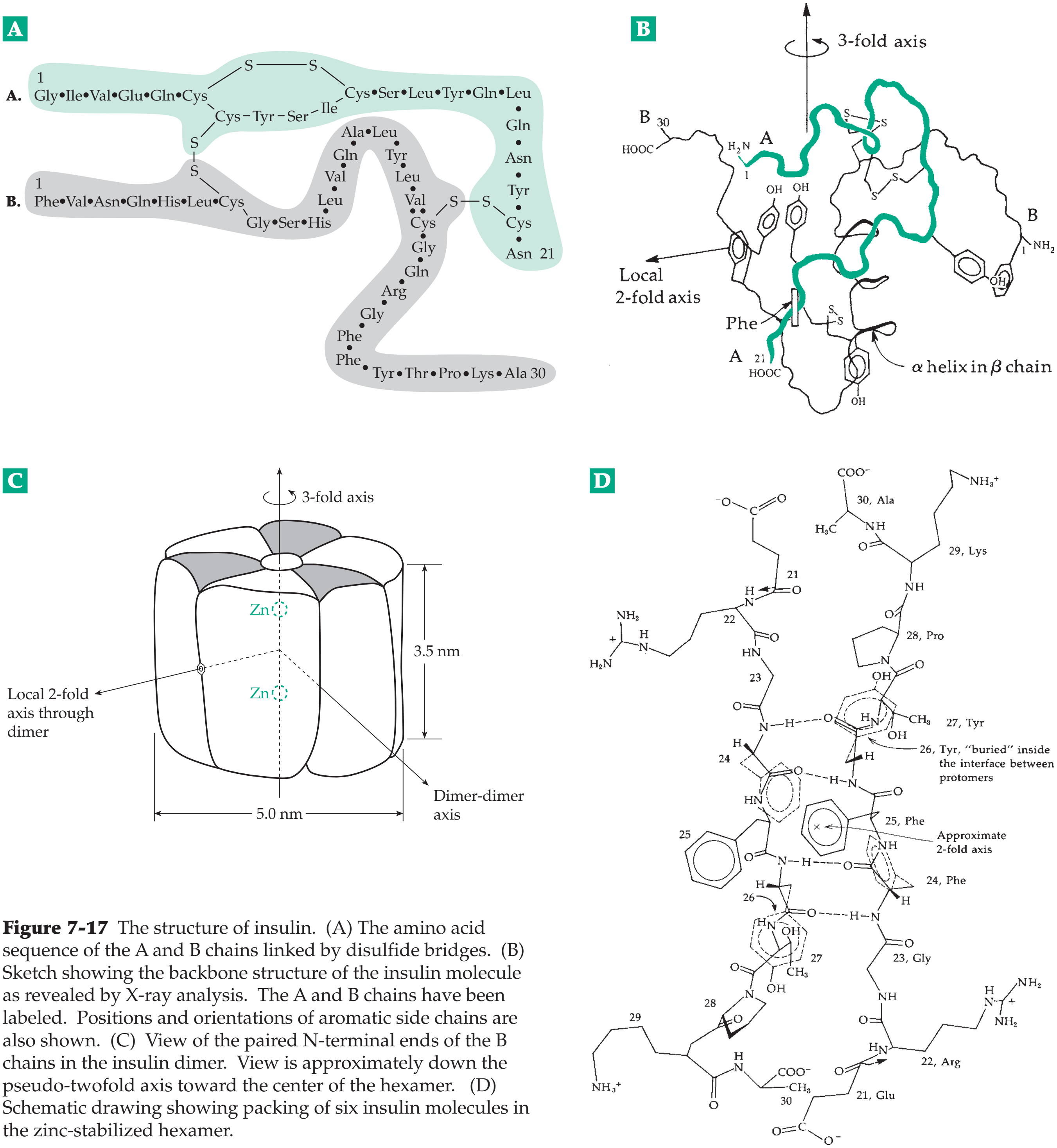

A
A. 1 Gly•Ile•Val•Glu•Gln•Cys
S — S
Cys–Tyr–Ser–Ile
Cys•Ser•Leu•Tyr•Gln•Leu
Gln
Asn
Tyr
Cys
Asn 21
S
S
B. 1 Phe•Val•Asn•Gln•His•Leu•Cys
Gly•Ser•His
Leu
Val
Gln
Ala•Leu
Tyr
Leu
Val
S — S
Cys
Gly
Gln
Arg
Gly
Phe
Phe
Tyr•Thr•Pro•Lys•Ala 30
B
3-fold axis
B 30
HOOC
H2N A
1
S S
S
S
OH OH
B
1 NH2
Local 2-fold axis
OH
Phe
S S
A 21 HOOC
OH
α helix in β chain
C
3-fold axis
Zn
Zn
3.5 nm
Local 2-fold axis through dimer
Dimer-dimer axis
5.0 nm
D
COO⁻
NH3⁺
⁻O C O
30, Ala
H3C NH
O
29, Lys
21
H O
N
NH
NH2
22
O
28, Pro
H
+ N
O
H2N
HN
N
OH O
23
O
CH3
27, Tyr
N—H---O
HN
OH
H
26, Tyr, "buried" inside the interface between protomers
24
H
O---H—N
O
HN
25, Phe
25
Approximate 2-fold axis
NH
O
N—H---O
24, Phe
H
H
26
O---H—N
HN OH
O
27
23, Gly
O
CH3
OH
NH2
N
H
NH
N +
28
O
NH2
29
O
HN
22, Arg
H
N
COO⁻
O
NH3⁺
CH3
30
21, Glu
O
O=C
O⁻

Figure 7-17 The structure of insulin. (A) The amino acid sequence of the A and B chains linked by disulfide bridges. (B) Sketch showing the backbone structure of the insulin molecule as revealed by X-ray analysis. The A and B chains have been labeled. Positions and orientations of aromatic side chains are also shown. (C) View of the paired N-terminal ends of the B chains in the insulin dimer. View is approximately down the pseudo-twofold axis toward the center of the hexamer. (D) Schematic drawing showing packing of six insulin molecules in the zinc-stabilized hexamer.

also 60 trimers (on the faces) and 90 dimers (across the edges) (Fig. 7-19). Such structures can be formed only for certain values of *T* where the number of subunits is 60*T* and there are 12 pentamers (pentons) and 10 (*T*-1) hexamers hexons. *T* can assume values of $h^2 + hk + k^2$ where h and k may be positive integers or zero. Some allowed *T* values are 1, 3, 4, 7, 9, 13, 25.[68,70,96,97]

The subunits in virus coats with *T* greater than one are not all in equivalent positions. For example, the three subunits labeled A, B, and C in Fig. 7-19 are each slightly differently positioned with respect to neighboring subunits. Since virus coats are usually tightly packed the subunits must assume more than one conformation. One kind of conformational change that allows quasi-equivalence of subunits is observed in the tomato bushy stunt virus. Two structural domains are connected by a hinge which allows an outer protruding domain to move slightly to preserve good isologous interactions with a corresponding domain in another subunit.[68]

The southern bean mosaic virus has an eight-stranded antiparallel β-barrel structure closely similar to that of the major domain of the bushy stunt viruses but lacking the second hinged domain. The problem of quasi-equivalence is resolved by the presence of an N-terminal extension that binds onto a subunit across the quasi-six-fold axis to give a set of three subunits (labeled C in Fig. 7-19) that associate with true three-fold symmetry and another set (B) with a slightly different conformation fitting between them.[68,92] The subunits A, which have a third conformation, fit together around the five-fold axis in true cyclic symmetry.

A surprising finding is that the polyoma virus coat, which was expected to contain 420 (7 x 60) subunits, apparently contains only 360. The result is that the hexavalent morphological unit is a pentamer and that quasi-eqivalence appears

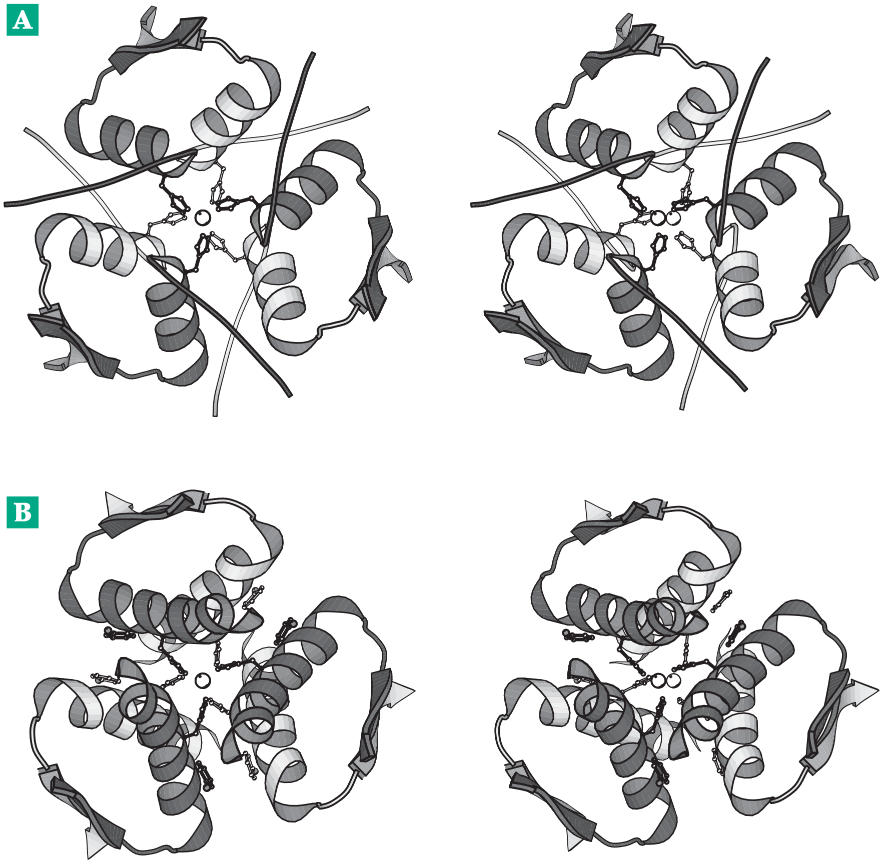


Figure 7-18 Stereoscopic MolScript ribbon drawings of the B chains (A chains omitted) of (A) hexameric 2-zinc pig insulin. (B) A phenol complex of the same protein. Within each dimer the B chains are shaded differently. The Zn^{2+} ions are represented by white spheres and the coordinating histidine side chains are shown. Six noncovalently bound phenol molecules can be seen, as can several conformational differences. From Whittingham *et al*.[87] Courtesy of Peter C. E. Moody.

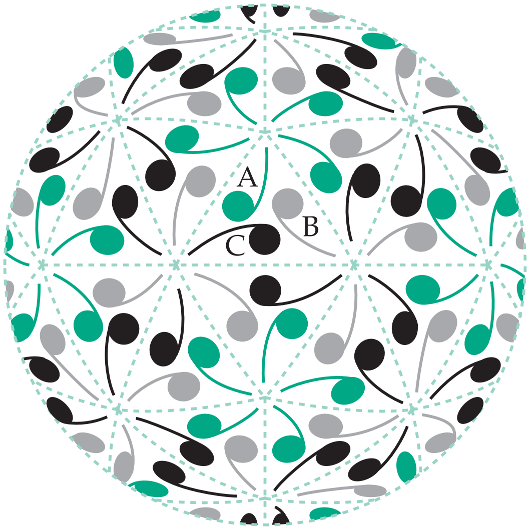


Figure 7-19 Schematic icosahedrally symmetric structure with 180 subunits. The quasi-equivalent units A, B, and C are necessarily somewhat differently positioned with respect to their neighbors and must therefore assume different conformations in order to fit together tightly. From Harrison.[68]

to be violated.[98,99] Flexible arms tie the pentamers together.

Quasi-equivalence of subunits also provides the supercoil in bacterial flagella (Chapter 19) and accounts for some interesting aspects of the structure of tobacco mosaic virus. The protein subunits of the virus can exist either as a helix with 16.3 subunits per turn (Fig. 7-8) or as a flat ring of 17 subunits.[100] A very small conformational difference is involved. These rings dimerize but do not form larger aggregates. What is surprising is that the dimeric rings do not have dihedral symmetry, all of the subunits in the dimeric disk being oriented in one direction but with two different conformations. The disk may serve as an intermediate in virus assembly. The inner portions of the quasi-equivalent disk subunits have a jawlike appearance as if awaiting the incorporation of RNA. As the RNA becomes bound, the disks could dislocate to a "lock-washer" conformation to initiate and to propagate growth of the helical virus particle.[100,44a] However, there is uncertainty about this interpretation.[45,101]

Some enzymes, such as yeast hexokinase and creatine kinase (Chapter 12), associate in extremely asymmetric ways.[102] A dimer is formed by means of heterologous interactions but steric hindrance prevents the unsatisfied sets of interacting groups from joining with additional monomers to form higher polymers. As Galloway pointed out, many biological structures are not completely ordered but nevertheless possess well-defined and functionally important local relationships.[103]

Regulatory subunits and multienzyme complexes. Proteins are often organized into large complexes, sometimes for the purpose of regulating metabolism. An example is **aspartate carbamoyltransferase** which catalyzes the first step in the synthesis of the pyrimidine rings of DNA and RNA (Chapter 25). The 310-kDa enzyme from *E. coli* can be dissociated into two 100-kDa trimers, referred to as **catalytic subunits**, and three 34-kDa dimers, the **regulatory subunits** which alter their conformations in response to changes in the ATP, UTP, and CTP concentrations.[104–107] The molecule is roughly triangular in shape[47,108] with a thickness of 9.2 nm and a length of the triangular side of 10.5 nm (Fig. 7-20). The symmetry is 3:2, i.e., it is dihedral with one 3-fold axis of rotation and three 2-fold axes. The two trimers of catalytic subunits lie face-to-face with the dimeric regulatory subunits fitting between them into the grooves around the edges of the trimers (Fig. 7-20). The dimers are not aligned exactly parallel with the 3-fold axis, but to avoid eclipsing, the upper half of the array is rotated around the 3-fold axis with respect to the lower half. In the center is an aqueous cavity of dimensions ~2.5 x 5.0 x 5.0 nm. The active sites of the enzyme are inside this cavity which is reached through six ~1.5-nm opening around the sides.

Many other oligomeric enzymes and other complex assemblies of more than one kind of protein subunit are known. For example, the **2-oxoacid dehydrogenases** are huge 2000- to 4000-kDa complexes containing three different proteins with different enzymatic activities in a cubic array (Fig. 15-14). The filaments of striated muscle (Chapter 19), antibodies and complement of blood (Chapter 31), and the tailed bacteriophages (Box 7-C) all have complex molecular architectures.

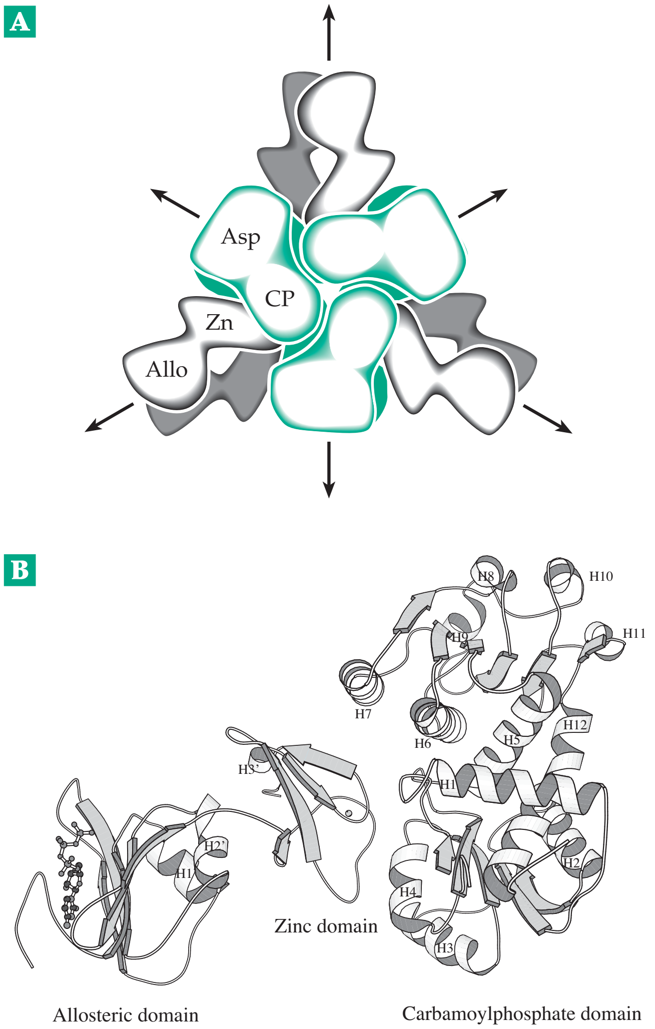


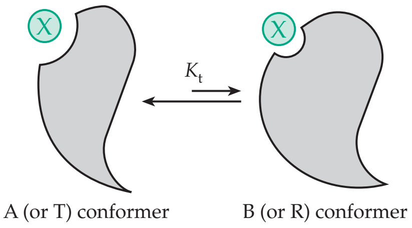
Figure 7-20 (A) Subunit assembly of two C_3 catalytic trimers (green) and three R_2 regulatory dimers around the periphery in aspartate carbamoyltransferase. After Krause *et al*.[109] Courtesy of William N. Lipscomb. The aspartate- and carbamoylphosphate-binding domains of the catalytic subunits are labeled Asp and CP, respectivley, while the zinc and allosteric domains of the regulatory subunits are labeled Allo and Zn, respectively. (B) Ribbon drawing of a single pair of regulatory (left) and catalytic (right) subunits with the structural domains labeled. MolScript drawing from Thomas *et al*.[110]

C. Cooperative Changes in Conformation

A substrate will bind better to some conformations of a protein than it will to others. This simple fact, together with the tendency for protein monomers to associate into clusters, allows for cooperative changes in conformation within oligomeric proteins. These changes provide the basis for important aspects of the regulation of enzymes and of metabolism. They impart cooperativity to the binding of small molecules such as that of oxygen to hemoglobin and of substrates and regulating molecules to enzymes. Many of the most fundamental and seemingly mysterious properties of living things are linked directly to cooperative changes within the fibrils, membranes, and other structures of the cell.

In 1965 a simple, appealing mathematical description of cooperative phenomena was suggested by Monod, Wyman, and Changeux[30,110a,110b] and focused new attention on the phenomenon. They suggested that conformational changes in protein subunits, which could be associated with altered binding characteristics, occur cooperatively within an oligomer. For example, binding of phenol to hexameric 2-zinc insulin (Fig. 7-18) could induce all six individual subunits to change their conformation together, preserving the D_6 symmetry of the complex. (In fact, it is more complex than this.[111]) The four subunits of hemoglobin could likewise change their conformation and affinity for O_2 synchronously. This is very nearly true and is of major physiological significance.

Consider an equilibrium (Eqs. 7-27 and 7-28) between protein molecules in two different conformations A and B (T and R in the MWC terminology) and containing a single binding site for molecule X. In the Monod–Wyman–Changeux (MWC) model the conformations are designated T (tense) and R (relaxed) but in the interest of providing a more general treatment the terminology used in this book is that of Koshland *et al.*[13,112–115]



(7-27)

$$K_t = [B] / [A] \qquad (7\text{-}28)$$

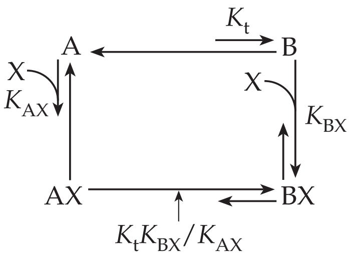
If the equilibrium constant K_t is approximately 1, the two **conformers** have equal energies, but if $K_t < 1$, A is more stable than B.

1. Unequal Binding of Substrate and "Induced Fit"

Assume that conformer B binds X more strongly than does conformer A (as is suggested by the shapes of the binding sites in Eq. 7-27). The intrinsic binding constants to the A and B conformers K_{AX} and K_{BX} (or K_T and K_R) are defined by Eq. 7-29:

$$\begin{aligned} K_{AX} &= [AX] / [A][X] \\ K_{BX} &= [BX] / [B][X]) \end{aligned} \qquad (7\text{-}29)$$

The entire set of equilibria for this system are shown in Eq. 7-30. Note that the constant relating

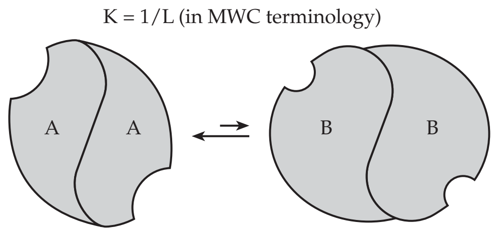


(7-30)

BX to AX is not independent of the other three constants but is given by the expression K_tK_{BX} / K_{AX}. Now consider the following situation. Suppose that A predominates in the absence of X but that X binds more tightly to B than to A. There will be largely either free A or BX in the equilibrium mixture with smaller amounts of AX and B. An interesting kinetic question arises. By which of the two possible pathways from A to BX (Eq. 7-30) will the reaction take place? The first possibility, assumed in the MWC model, is that X binds only to preformed B, which is present in a small amount in equilibrium with A. The second possibility is that X can bind to A but that AX is then rapidly converted to BX. We could say that *X induces a conformational change* that leads to a better fit. This is the basis for the **induced fit** theory of Koshland. Bear in mind that the equilibrium constants can give us the equilibrium concentrations of all four forms in Eq. 7-30. However, rates of reaction are often important in metabolism and we cannot say *a priori* which of the two pathways will be followed.

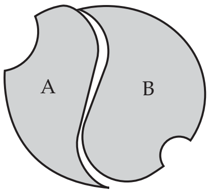
If K_{BX} / K_{AX} is very large, an insignificant amount of AX will be present at equilibrium. In such a case there is no way experimentally to determine K_{AX}. The two constants K_t and K_{BX} are sufficient to describe the *equilibria* but an induced fit mechanism may still hold.

Now consider the association of A and B to form oligomers in which the intrinsic binding constants K_{AX} and K_{BX} have the same values as in the monomers. Since more enzymes apparently exist as isologous dimers than as any other oligomeric form,[116] it is appropriate to consider the behavior of such dimers in some detail. Monod *et al.* emphasized that both conformers A and B (T and R) can associate to form isologous dimers in which symmetry is preserved (Eq. 7-31).



(7-31)

On the other hand, association of B and A would lead to an unsymmetric dimer in which bonding between subunits might be poor:



Mixed AB dimer
which associates weakly

In the MWC treatment, the assumption is made that the mixed dimer AB can be neglected entirely. However, a more general treatment requires that we consider all dimeric forms. The formation constants K_{AA}, K_{BB}, and K_{AB} are defined as follows[13,113] (Eqs. 7-32 to 7-34; note the statistical factor of 2 relating K_{AB} to the association constant K_f):

$$2A \rightleftarrows A_2 \qquad K_{AA} = \frac{[A_2]}{[A]^2} \tag{7-32}$$

$$2B \rightleftarrows B_2 \qquad K_{BB} = \frac{[B_2]}{[B]^2} = \frac{[B_2]}{K_t{}^2[A]^2} \tag{7-33}$$

$$A + B \rightleftarrows AB \qquad K_f = 2K_{AB} = \frac{[AB]}{[A][B]} = \frac{[AB]}{K_t[A]^2} \tag{7-34}$$

2. Binding Equilibria for a Dimerizing Protein

All of the equilibria of Eqs. 7-28 through 7-34 involved in formation of dimers A_2, AB, and B_2 and in the binding of one or two molecules of X per dimer are depicted in Fig. 7-21. Above each arrow the microscopic constant associated with that step is shown multiplied by an appropriate statistical factor. The fractional saturation Y is given by Eq. 7-35. Each of the nine terms in the numerator gives the concentration of

$$2Y \text{ (based on dimer)} = \frac{\begin{array}{r}[AX] + [BX] + [A_2X] + 2\,[A_2X_2] + [ABX] \\ + [AXB] + 2\,[ABX_2] + [B_2X] + 2\,[B_2X_2]\end{array}}{\begin{array}{r}\frac{1}{2}\,(\,[A] + [AX] + [B] + [BX]\,) + [A_2] + [A_2X] \\ + [A_2X_2] + [AB] + [ABX] + [AXB] \\ + [ABX_2] + [B_2] + [B_2X] + [B_2X_2]\end{array}} \tag{7-35}$$

bound X represented by one of the nine forms containing X in Fig. 7-21. The 14 terms in the denominator represent the concentration of protein in each form including those containing no bound X. Protein concentrations are given in terms of the molecular mass of the dimer; hence, some of the terms in the denominator are multiplied by 1/2.

All of the terms in both the numerator and the denominator of Eq. 7-35 can be related back to [X], using the microscopic constants from Fig. 7-21 to give an equation (comparable to Eq. 7-8) which presents Y in terms of [X], K_{AX} and K_{BX}, K_t, and the interaction constants K_{AA}, K_{AB}, and K_{BB}. Since the equation is too complex to grasp immediately, let us consider several specific cases in which it can be simplified.

The Monod-Wyman-Changeux (MWC) model. If both K_{AA} and K_{BB} are large enough, there will be no dissociation into monomers. The transition between conformation A and conformation B can occur cooperatively within the dimer or higher oligomer, and the mathematical relationships shown in Fig. 7-21 are still appropriate. One further restriction is needed to describe the MWC model. Only symmetric dimers are allowed. That is, K_{AA} and $K_{BB} >> K_{AB}$ (see Eq. 7-31), and only those equilibria indicated with green arrows in Fig. 7-21 need be considered.[30] In the absence of ligand X, the ratio $[B_2]/[A_2]$ is a constant, 1/L in the MWC terminology (Eq. 7-36; see also Eq. 7-31).

$$\frac{[B_2]}{[A_2]} = \frac{1}{L} = \frac{K_{BB}}{K_{AA}}\,K_t{}^2 \tag{7-36}$$

Both of the association constants K_{AA} and K_{BB} and the transformation constant K_t affect the position of the equilibrium. Thus, a low ratio of $[B_2]$ to $[A_2]$ could result if K_{BB} and K_{AA} were similar but K_t was small. If K_t were ~1 a low ratio could still arise because $K_{AA} >$

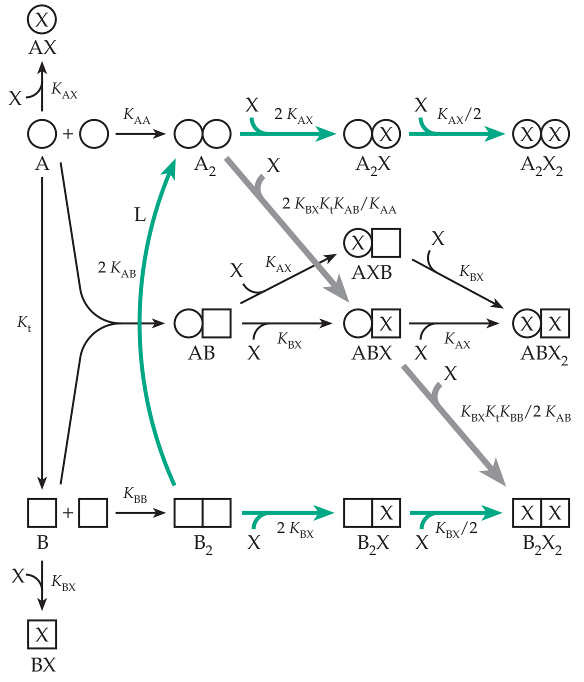


Figure 7-21 Possible forms of a dimerizing protein existing in two conformations with a single binding site per protomer for X. Green arrows indicate equilibria considered by MWC. Solid arrows indicate equilibria considered by Koshland *et al.*[13,114] Heavy gray arrows are for the simplest induced fit model with no dissociation of the dimer. Note that all equilibria are regarded as reversible (despite the unidirectional arrows). K_{AX} and K_{BX} are assumed the same for subunits in monomeric and dimeric forms.

K_{BB}, i.e., because the subunits are associated more tightly in A_2 than in B_2. For this case Eq. 7-35 simplifies to Eq. 7-37.

$$2Y = \frac{[A_2X] + 2\,[A_2X_2] + [B_2X] + 2\,[B_2X_2]}{[A_2] + [A_2X] + [A_2X_2] + [B_2] + [B_2X] + [B_2X_2]}$$

$$= \frac{2K_{AA}K_{AX}[X] + 2K_{AA}K_{AX}{}^2[X]^2 + 2K_{BB}K_{BX}K_t{}^2[X] + 2K_{BB}K_{BX}K_t{}^2[X]^2}{K_{AA} + 2K_{AA}K_{AX}[X] + K_{AA}K_{AX}{}^2[X]^2 + K_{BB}K_t{}^2 + 2K_{BB}K_{BX}K_t{}^2[X] + K_{BB}K_{BX}{}^2K_t{}^2[X]^2} \tag{7-37}$$

Substituting from Eq. 7-36 into Eq. 7-37 we obtain (Eq. 7-38):

$$Y\ \text{(for dimer)} = \frac{L \cdot K_{AX}[X](1 + K_{AX}[X]) + K_{BX}[X](1 + K_{BX}[X])}{L(1 + K_{AX}[X])^2 + (1 + K_{BX}[X])^2} \tag{7-38}$$

For an oligomer with n subunits Monod *et al*. assumed that all sites in either conformer are independent and equivalent. The equation for Y (based on Eq. 7-17) is

$$Y = \frac{L \cdot K_{AX}[X](1 + K_{AX}[X])^{n-1} + K_{BX}[X](1 + K_{BX}[X])^{n-1}}{L(1 + K_{AX}[X])^n + (1 + K_{BX}[X])^n} \tag{7-39}$$

We assume initially that B_2 binds X more strongly than does A_2. Hence, if the equilibrium in Eq. 7-36 favors B_2 strongly (L is small), the addition of X to the system will not shift the equilibrium between the two conformations and binding will be noncooperative (Eq. 7-39 will reduce to Eq. 7-18). However, if the equilibrium favors A_2 (L is large), addition of X will shift the equilibrium in favor of B_2 (which binds X more tightly). Furthermore, since the expression for Y (Eq. 7-39) contains a term in $K^2_{BX}[X]^2$ in the numerator, binding will tend to be cooperative. In the extreme case that L is large and $K_{AX} \sim 0$, most of the terms in Eq. 7-39 drop out and it approaches the equation previously given for completely cooperative binding (Eq. 7-21) with $K = K_{BX}{}^2L$. With other values of K_{AX}, K_{BX}, and L incomplete cooperativity is observed.[112]

Further development of the MWC theory as it applies to enzyme kinetics is given in Chapter 9, Section B.

The induced fit model. In this model, only A_2, ABX, and B_2X_2 are considered (heavy arrows in Fig. 7-15).[13,114] The expression for 2Y is:

$$2\,Y = \frac{[ABX] + 2\,[B_2X_2]}{[A_2] + [ABX] + [B_2X_2]}$$

$$= \frac{2\,K_{BX}K_t\dfrac{K_{AB}}{K_{AA}}[X] + 2\,(K_{BX}K_t)^2\dfrac{K_{BB}}{K_{AA}}[X]^2}{1 + 2\,K_{BX}K_t\dfrac{K_{AB}}{K_{AA}}[X] + (K_{BX}K_t)^2\dfrac{K_{BB}}{K_{AA}}[X]^2} \tag{7-40}$$

The constants used here are defined by Eqs. 7-8 through 7-10 and differ from those of Koshland, who sometimes arbitrarily set $K_{AA} = 1$ and redefined K_{BB} as an *interaction constant* equal to K_{BB} / K_{AA}. Although this simplifies the algebra it is appropriate only for completely associated systems and might prove confusing.

When K_{AB} is small (no "mixed" dimer) Eq. 7-16 also simplifies to Eq. 7-45 for completely cooperative binding with the value K given by Eq. 7-17. On the other hand, if K_{AB} is large compared to K_{AA} and K_{BB}, anticooperativity (negative cooperativity) will be observed. The saturation curve will contain two separate steps just as in the binding of protons by succinate dianion (Fig. 7-5).

$$\overline{K} = K_{\mathrm{BX}}{}^2 K_{\mathrm{t}}{}^2 \frac{K_{\mathrm{BB}}}{K_{\mathrm{AA}}} = \frac{K_{\mathrm{BX}}{}^2}{L} \qquad (7\text{-}41)$$

One conformational state dissociated. It may happen that A_2 is a dimer but that B_2 dissociates into monomers because K_{BB} is very small. In such a case binding of X leads to dissociation of the protein. A well-known example is provided by hemoglobin of the lamprey which is a dimer and which dissociates to a monomer upon binding of oxygen.[117] Equation 7-11 simplifies to Eq. 7-42. The reader may wish to consider whether the weakly cooperative binding of oxygen by lamprey hemoglobin is predicted by this equation.

$$2\mathrm{Y} = \frac{[\mathrm{BX}]}{[\mathrm{A}_2] + \frac{1}{2}[\mathrm{BX}]} \qquad (7\text{-}42)$$

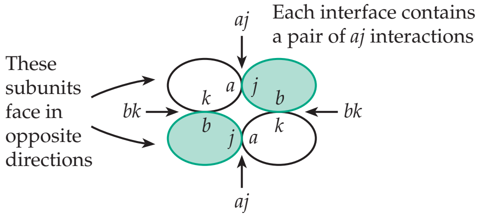
Look again at the expression for L, the constant determining the relative amounts of a protein in conformations A and B in the absence of ligand. From Eq. 7-36 we see that a large value of L (conformer A favored) can result either because K_{t} is very small or because $K_{\mathrm{BB}} << K_{\mathrm{AA}}$. Thus, if $K_{\mathrm{t}} \sim 1$ and L is large, the subunits must associate much more weakly in B_2 than in A_2 and the chances are that binding of X will dissociate the molecule as in the case of lamprey hemoglobin. On the other hand, if K_{t} is very small, implying that the molecule is held in conformation A because of some intrinsically more stable folding pattern in that conformation, K_{BB} might exceed K_{AA} very much; if K_{AA} were low enough A_2 could be completely dissociated. Binding of ligand would lead to association and to cooperative binding. This can be verified by writing down the appropriate terms from Eq. 7-35.

3. Higher Oligomers

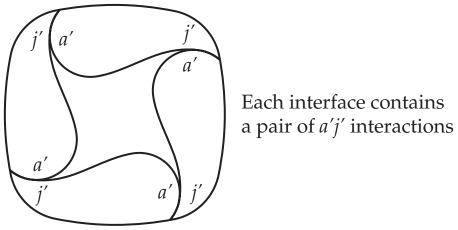
Mathematical treatment of binding curves for oligomers containing more than two subunits is complex, but the algebra is straightforward. A computer can be programmed to do necessary calculations. Avoid picking an equation from the literature and assuming that it will be satisfactory. Consider the two tetrameric structures in Fig. 7-22. In the isologous square (also shown in Fig. 7-11) separate contributions to the free energy of binding can be assigned to the individual pairs of interactions *aj* and *bk*.

Thus, following Cornish-Bowden and Koshland[114] for assembly of the tetramer (Eq. 7-43):

$$\begin{aligned} \Delta G_{\mathrm{f}} &= 2\,\Delta G_{aj\mathrm{AA}} + 2\,\Delta G_{bk\mathrm{AA}} \\ K_{\mathrm{f}} &= K_{aj\mathrm{AA}}^2 K_{bk\mathrm{AA}}^2 \end{aligned} \qquad (7\text{-}43)$$


aj
Each interface contains a pair of aj interactions
These subunits face in opposite directions
bk
k
a
j
b
bk
b
j
a
k
aj

Isologous square



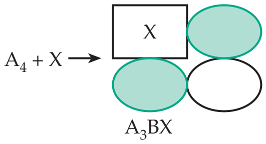
Heterologous square

Figure 7-22 Comparison of the interactions in isologous (dihedral) and heterologous (cyclic) square configurations of subunits.

Since Gibbs energies are additive, the formation constant will be the product of formation constants representing the individual interactions; thus, $K_{aj\mathrm{AA}}$ represents the formation constant of a dimer in which only the *aj* pair of bonds is formed.

In the isologous tetrahedron (Fig. 7-5) the third set of paired interactions *cl* must be taken into account. (However, the third interaction constant will not be an association constant of the type represented by $K_{aj\mathrm{AA}}$ and $K_{bk\mathrm{AA}}$ but a dimensionless number.) On the other hand, the heterologous square has only a single interaction constant.

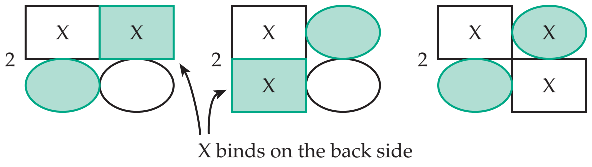
Now consider the binding of one molecule of X to the isologous tetramer with a conformational change in one subunit (Eq. 7-44). We see that one pair of *aj* interactions and one pair of *bk* interactions have been altered. The equilibrium constant for the binding of X to the tetramer will be (Eq. 7-45) in which the 4 is a



(7-44)

$$K = 4\,\frac{K_{aj\mathrm{AB}} K_{bk\mathrm{AB}}}{K_{aj\mathrm{AA}} K_{bk\mathrm{AA}}}\,K_{\mathrm{BX}} K_{\mathrm{t}} \qquad (7\text{-}45)$$

statistical factor arising from the fact that there are four different ways in which to form A_3BX. When a second molecule of X is added three geometrical arrangements are possible:



Each one can be formed in two ways. It is a simple matter to write down the microscopic constants for addition of the second molecule of X as the sum of three terms. Because values of the constants for *aj* and *bk* differ, it will be clear that the three ways of adding the second molecule of X are not equally probable. Thus, the oligomer will show preferred orders of "loading" with ligand X.

Two different geometries for the heterologous tetramer are possible in form $A_2B_2X_2$. Again, the different arrangements need not be equally probable and the relative distribution of each will be determined by the specific values of the interaction constants. In the heterologous tetramer A_4 only one type of an interaction is present between subunits. However, as soon as a single molecule of X is bound and one subunit of conformation B is present, two kinds of *aj* interactions exist. (One in which group *a* is present in conformation A and the other in which it is present in conformation B.) Since these interactions always occur in equal numbers they can be lumped together.

While the foregoing may seem like an unnecessarily long exercise, it should provide a basic approach which can be applied to specific problems. However, remember that mathematical models require simplification. Real proteins often have more than two stable conformations.[117] The entire outside surface of a protein is made up of potential binding sites for a number of different molecules, both small and large. Filling of almost any of these sites can affect the functioning of a protein.

D. The Oxygen-Carrying Proteins

1. Myoglobin and Hemoglobin

The most studied example of a conformational change in a multisubunit protein induced by binding of a small molecule is provided by the cooperative binding of oxygen to hemoglobin.[118–120] Mammalian hemoglobin is an $\alpha_2\beta_2$ tetramer of ~16-kDa subunits, each containing 140 – 150 residues. Within each subunit the peptide chain folds in a characteristic largely α-helical pattern around a single large flat iron-containing ring structure called **heme** (Fig. 7-23). The folding is essentially the same in all hemoglobins, both in the α and β subunits and in the monomeric muscle oxygen storage protein, myoglobin. Amino acid residues are customerily designated by their position in one of the eight helices A–H. The imidazole group of histidine F-8 is coordinated with the iron in the center of the heme on the "proximal" side. The other side of the iron atom (the "distal" side) is the site of binding of a single molecule of O_2.

Although the folding of the peptide chain is almost the same in both subunits, and almost identical to that of myoglobin,[118,121,122] there are numerous differences in the amino acid sequence. If it were not for these differences, hemoglobin would be a highly symmetric molecule with the bonding pattern indicated in Fig. 7-5 with three 2-fold axes of rotation. In fact, hemoglobin has one true axis of rotation and two pseudo-twofold axes. There are two sets of true isologous interactions (those between the two α subunits and between the two β subunits) and two pairs of unsymmetrical interactions (between α and β subunits). The nearly symmetric orientation of different portions of the peptide backbone is clearly seen in the beautiful drawings of Geis.[119]

The contact region involved in one pair of interactions in hemoglobin ($\alpha_1\beta_1$) is more extensive than the other. There is close contact between 34 different amino acid side chains and 110 atoms lie within 0.4 nm of each other.[118] Hydrophobic bonding is the principal force holding the two subunits together, and only a few reciprocal contacts of the type found in a true isologous bond remain. The second contact designated $\alpha_1\beta_2$ involves only 19 residues and a total of 80 atoms. Because this interaction is weaker, hemoglobin dissociates relatively easily into αβ dimers held together by the $\alpha_1\beta_1$ contacts and motion occurs along the $\alpha_1\beta_2$ contacts during oxygenation. The truly isologous interactions (i.e., αα and ββ) are weak because the identical protomers hardly touch each other.

The binding of oxygen. Curves of percentage oxygenation (Y) vs the partial pressure of O_2 are given in Fig. 7-24 and illustrate the high degree of cooperativity. Depending upon conditions, values of n_{Hill} (Eq. 7-24) may be as high as 3. As a result of this cooperativity the hemoglobin, in the capillaries of the lungs at a partial pressure of oxygen of ~100 mm of mercury, is nearly saturated with oxygen. However, when the red cells pass through the capillaries of tissues in which oxygen is utilized the partial pressure of oxygen falls to about 5 mm of mercury. The cooperativity means that the oxygen is more completely "unloaded" than it would be if all four heme groups acted independently.

Deoxyhemoglobin has a low affinity for O_2, but the observed cooperativity in binding implies that in the fully oxygenated state the O_2 is held with a high affinity. The monomeric myoglobin also has a high affinity for oxygen, as does the abnormal **hemoglobin H**, which is made up of four β subunits. The latter also completely lacks cooperativity in binding.[123] These results can be interpreted according to the MWC model to indicate that deoxyhemoglobin exists in the T (A) conformation, whereas oxyhemoglobin is in the R (B) conformation. Myoglobin stays in the R conformation in *both* states of oxygenation as do the separated α and β chains of hemoglobin. The subunits of hemoglobin H also appear to be frozen in the R conformation, even though the quaternary structure is similar to that of deoxyhemoglobin.[123,124]

Oxygenation curves of hemoglobin are often fitted with the Adair equation (Eq. 7-12). Thus, at pH 7.4 under the conditions given in Table 7-2, Imai[125] found for the successive formation constants $K_1 = 0.004$, $K_2 = 0.009$, $K_3 = 0.002$ and $K_4 = 0.95$ in units of mm Hg^{-1}. From the definition of a formation constant the oxygen pressure PO_2 required for 50% oxygenation in the first step will be at $PO_2 = 1/K_f$ or $\log PO_2 = \log(1/.004) = 2.4$. This is a high oxygen pressure, far to the right side of the oxygenation curve in Fig. 7-24A. However, log K_4 is about 0.02, well to the left on the oxygenation curve. From these formation constants we can say that after three of the subunits have become oxygenated the affinity of the remaining subunit has increased about 300-fold when the concentration of the effector 2,3-bisphosphoglycerate is present at the normal physiological concentraton (see Section 4).[125–127]

However, we must ask what uncertainties are present in the data used to obtain these constants. To extract four successive binding constants from a curve like that in Fig. 7-24A is extremely difficult.[129,130] This fact has encouraged the widespread use of the simpler MWC model.[30a,127a] When the same data were treated by Imai[131] using the MWC model it was found that $L = 2.8 \times 10^6$ and $c = K_{f(T)} / K_{f(R)} = 0.0038$. Changes in

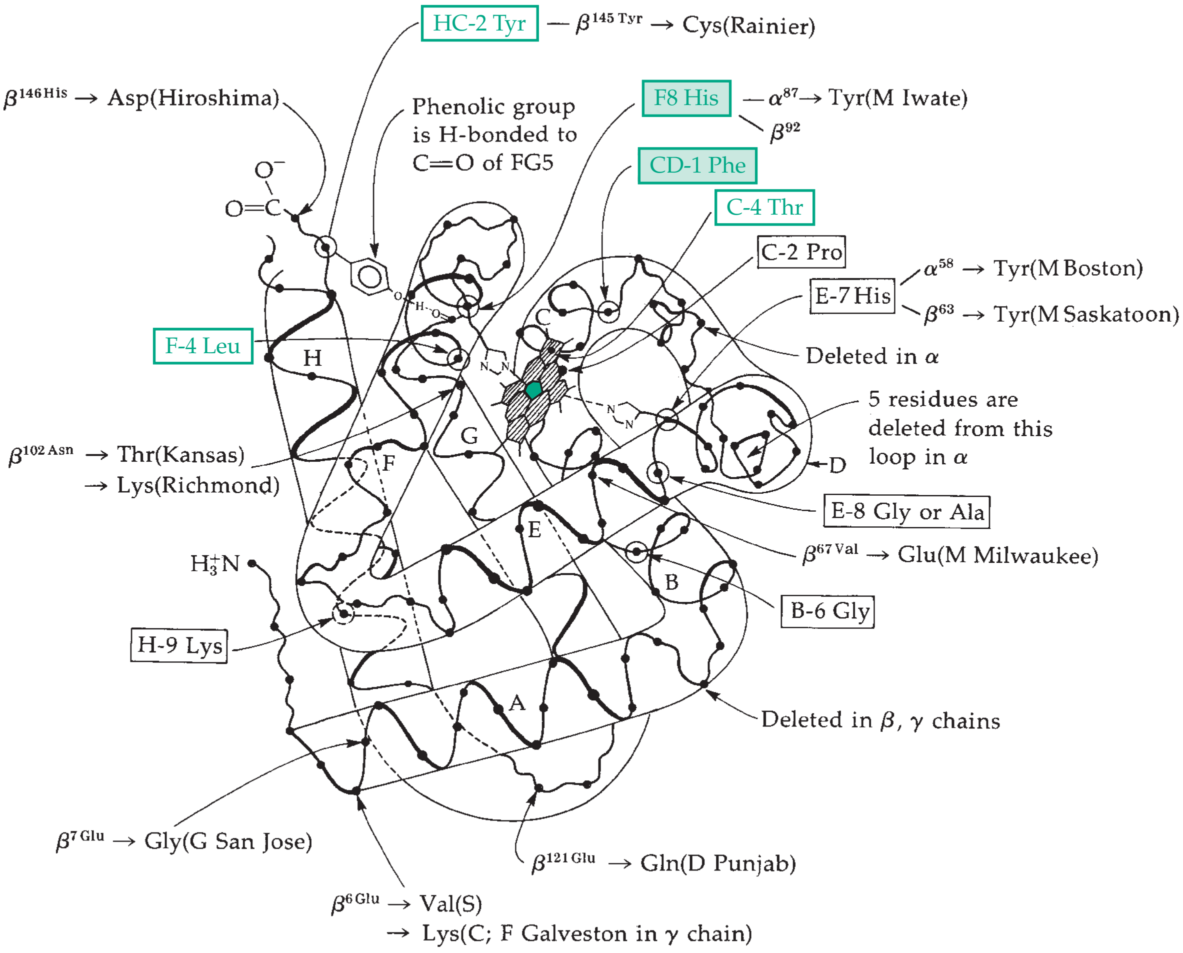

HC-2 Tyr
— $\beta^{145\,Tyr}$ → Cys(Rainier)
$\beta^{146\,His}$ → Asp(Hiroshima)
Phenolic group is H-bonded to C═O of FG5
O^-
O═C
F8 His
— α^{87} → Tyr(M Iwate)
β^{92}
CD-1 Phe
C-4 Thr
C-2 Pro
α^{58} → Tyr(M Boston)
E-7 His
β^{63} → Tyr(M Saskatoon)
Deleted in α
F-4 Leu
H
C
N
N
5 residues are deleted from this loop in α
D
$\beta^{102\,Asn}$ → Thr(Kansas)
→ Lys(Richmond)
G
F
E-8 Gly or Ala
E
$\beta^{67\,Val}$ → Glu(M Milwaukee)
H_3^+N
B
B-6 Gly
H-9 Lys
A
Deleted in β, γ chains
$\beta^{7\,Glu}$ → Gly(G San Jose)
$\beta^{121\,Glu}$ → Gln(D Punjab)
$\beta^{6\,Glu}$ → Val(S)
→ Lys(C; F Galveston in γ chain)

Figure 7-23 Folding pattern of the hemoglobin monomers. The pattern shown is for the β chain of human hemoglobin. Some of the differences between this and the α chain and myoglobin are indicated. Evolutionarily conserved residues are indicated by boxes, ☐ highly conserved, 🟩 invariant. Other markings show substitutions observed in some abnormal human hemoglobins. Conserved residues are numbered according to their location in one of the helices A–H, while mutant hemoglobins are indicated by the position of the substitution in the entire α and β chain.

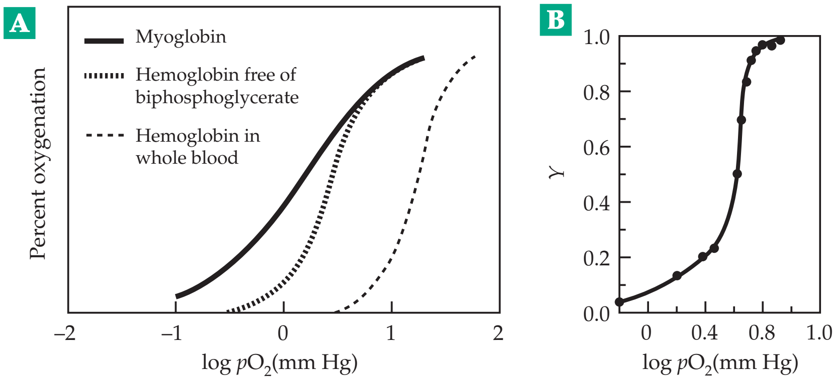


Figure 7-24 Cooperative binding of oxygen by hemoglobins. (A) Binding curve for myoglobin (noncooperative) and for hemoglobin in the absence and presence (in whole blood) of 2,3-bisphosphoglycerate. Oxygen affinity is decreased by bisphosphoglycerate. (B) Saturation curve for hemoglobin (erythrocruorin) of *Arenicola*, a spiny annelid worm. The molecule contains 192 subunits and 96 hemes. It shows very strong cooperativity with $n_{Hill} \sim 6$. From (A) Benesch and Benesch,[123] and (B) Waxman.[128]

enthalpy, entropy, and Gibbs energy are given in Table 7-2. Kinetic data[131a] as well as O_2-binding measurements with single crystals[132] are partially consistent with the MWC model.[133] However, the discovery of a third quaternary structure of hemoglobin, similar to the R state but distinct from it,[134–136a] emphasizes the complexity of this allosteric molecule.

Hemoglobin tetramers tend to stay tightly associated but some dissociation of oxyhemoglobin into dimers does occur ($K_f = 7 \times 10^5$ M).[137,138] Deoxyhemoglobin is about 40,000 times more tightly associated. All of the equilibria involved are strongly affected by pH and by the presence of salts such as NaCl.[126,139] This is in part a reflection of the strong role of ionic interactions in holding together the subunits in the T state as is discussed in the following sections.[127]

Structural changes accompanying oxygen binding. Perutz and associates, using X-ray crystallography, found *small* but real differences in the conformation of the subunits of deoxy- and oxy- hemoglobin.[140,141] More striking is the fact that upon oxygenation, both α and β subunits undergo substantial amounts of *rotation*, the net result being that the hemes of the two β subunits move about 0.07 nm closer together in the oxy form than in the deoxy form. Within the $\alpha_1\beta_1$ contacts (Fig. 7-25) little change is seen. On the other hand, contact $\alpha_1\beta_2$, the "allosteric interface," is altered drastically. As Perutz expressed it, there is a "jump in the dovetailing" of the CD region of the subunit relative to the FG region of the β subunit. The hydrogen-bonding pattern is also changed. A major difference is seen in the hydrogen-bonded salt bridges present at the ends of the molecules of deoxyhemoglobin. The $-NH_3^+$ group of Lys H-10 in each α subunit is hydrogen bonded to the carboxyl group of the C-terminal arginine of the opposite α chain. The guanidinium group of each C-terminal arginine is hydrogen bonded to the carboxyl group of Asp H-9 in the opposite α chain. It is also hydrogen bonded to an inorganic anion (phosphate or Cl^-), which in turn is hydrogen-bonded to the α amino group of Val 1 of the opposite α chain[142] forming a pair of isologous interactions.

At the other end of the molecule, the C-terminal group of His 146 of each β chain binds to the amino group of Lys C-6 of the α chain, while the imidazole side chain binds to Asp FG-1 of the same β chain (Fig. 7-25). These salt bridges appear to provide extra stability to

TABLE 7-2
Thermodynamic Functions for Oxygenation of Hemoglobin[a]

Reaction	ΔH (kJ mol^{-1})	ΔS (J°K^{-1} mol^{-1})	ΔG (kJ mol^{-1})	K_f
$T \rightarrow T(O_2)_4$	-51 ± 1	-154 ± 4	-5.0 ± 1.7	7.5
$R \rightarrow R(O_2)_4$	-62 ± 2	-146 ± 6	-19 ± 2	2.0×10^3
$T \rightarrow R$ (unoxygenated)	-70 ± 7	-111 ± 25	-37 ± 10	3.6×10^{-7}
$T(O_2)_4 \rightarrow R(O_2)_4$ (oxygenated)			-19	2.1×10^3

Parameters for MWC model

$L = (3.6 \times 10^{-7})^{-1} = 2.8 \times 10^6$

$c = K_{f(T)}/K_{f(R)} = 0.0038$

[a] From Imai, K. (1979) *J. Mol. Biol.* **133**, 233–247. The measurements were made at pH 7.4 in the presence of 0.1 M chloride ion and 2 mM 2,3-bisphosphoglycerate to mimic physiological conditions. The values of ΔH, ΔS, and ΔG given are per mole of heme, i.e., per monomer unit. They must be multiplied by 4 to correspond to the reactions as shown for the tetramer.

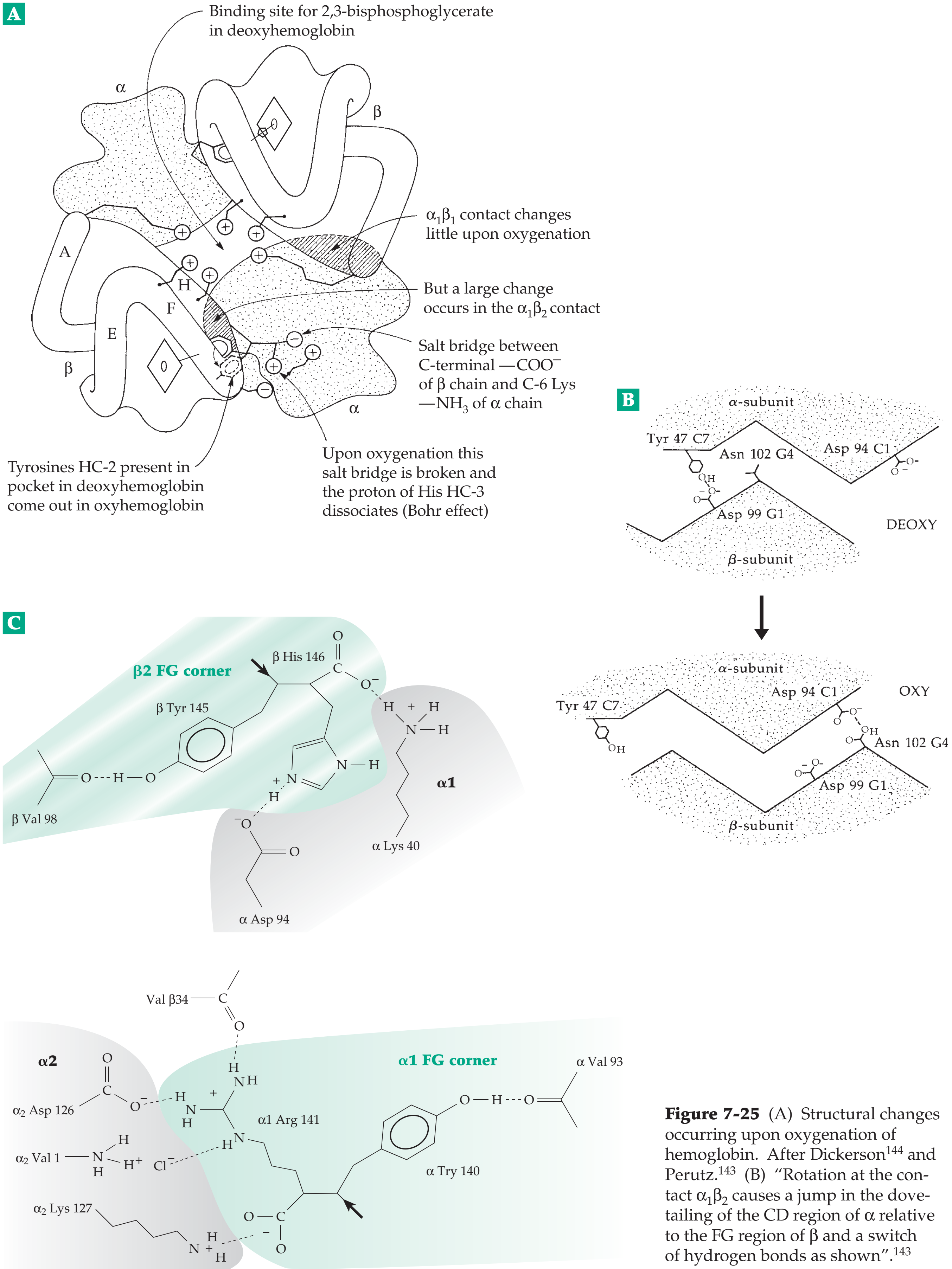

A
Binding site for 2,3-bisphosphoglycerate in deoxyhemoglobin
α
β
A
H
F
E
β
α
$\alpha_1\beta_1$ contact changes little upon oxygenation
But a large change occurs in the $\alpha_1\beta_2$ contact
Salt bridge between C-terminal —COO⁻ of β chain and C-6 Lys —NH_3 of α chain
Tyrosines HC-2 present in pocket in deoxyhemoglobin come out in oxyhemoglobin
Upon oxygenation this salt bridge is broken and the proton of His HC-3 dissociates (Bohr effect)
B
α-subunit
Tyr 47 C7
Asn 102 G4
Asp 94 C1
Asp 99 G1
β-subunit
DEOXY
α-subunit
Asp 94 C1
OXY
Tyr 47 C7
Asn 102 G4
Asp 99 G1
β-subunit
C
β2 FG corner
β His 146
β Tyr 145
β Val 98
α1
α Lys 40
α Asp 94
Val β34
α2
α_2 Asp 126
α_2 Val 1
α_2 Lys 127
α1 FG corner
α1 Arg 141
α Val 93
α Try 140
Cl⁻

Figure 7-25 (A) Structural changes occurring upon oxygenation of hemoglobin. After Dickerson[144] and Perutz.[143] (B) "Rotation at the contact $\alpha_1\beta_2$ causes a jump in the dovetailing of the CD region of α relative to the FG region of β and a switch of hydrogen bonds as shown".[143] (C) Some details of the salt bridges.

deoxyhemoglobin and account for the high value of the constant L. In deoxyhemoglobin the side chain of the highly conserved Tyr HC-2 lies tucked into a pocket between the H and F helices and is hydrogen bonded to the main chain carbonyl of residue FG-5 (Figs. 7-23 and 7-25). Upon oxygenation this tyrosine in each subunit is released from its pocket; the salt bridges at the ends of the molecules are broken and the subunit shifts into the new bonding pattern characteristic of oxyhemoglobin.[143,144] Cooperativity in O_2 binding is absent or greatly decreased in mutant hemoglobins with substitutions in the residues involved in these salt bridges[145–147] or in residues lying in the $\alpha_1\beta_2$ interface.[148,149]

How does the binding of O_2 to the iron of heme trigger the conformational change in hemoglobin? An enormous amount of effort by many people has been expended in trying to answer this question. As is pointed out in Chapter 16, the iron atom in deoxyhemoglobin lies a little outside the plane of the heme rings. When oxygenation occurs the iron atom moves toward the oxygen and into the plane of the heme.[150,151] This movement probably amounts to only about 0.05 nm. Nevertheless, this small displacement evidently induces the other structural changes that are observed. The iron pulls the side chain of histidine F-8 with it and moves helix F which is also hydrogen bonded to this imidazole ring. Because of the tight packing of the various groups this motion cannot occur freely but is accompanied by a movement of the F helix by 0.1 nm across the heme plane. These movements may induce additional structural changes in the irregularly folded FG corners that allow the subunits to shift to the new stable position of the R state. All four subunits appear to change conformation together. This conformational change must also cause the affinity for oxygen of any unoxygenated subunits to rise dramatically, presumably by shifting the iron atoms into the planes of the heme rings. This ensures the cooperative loading of the protein by O_2.

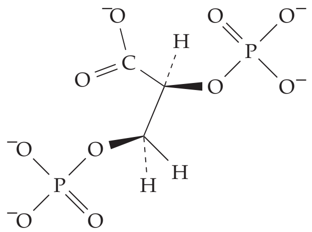
While there is no doubt that the iron atom moves upon oxygenation, it is not obvious that this will lead to the observed cooperativity. Oxygenated heme has some of the characteristics of an Fe(III)–peroxide anion complex.[152] The iron atom acquires an increased positive charge upon oxygenation by donating an electron for bond formation.

$$\mathrm{Fe}^{2+}\ (\mathrm{deoxy}) + \mathrm{O}_2 \longrightarrow \mathrm{Fe}^{3+}\!-\mathrm{O}_2^{-} \qquad (7\text{-}46)$$

This may transmit an electronic effect through either His F-8 or the heme ring to the nearby $\alpha_1\beta_2$ interface and also affect the subunit interactions. The reaction of various heme proteins with oxygen is discussed further in Chapter 16.

The Bohr effect and allosteric regulators. The breaking of the salt bridges at the ends of the hemoglobin molecule upon oxygenation has another important result. The pK_a values of the N-terminal valines of the α subunits and of His HC-3 of the β subunits are abnormally high in the deoxy form because they are tied up in the salt bridges. In the oxy form in which the groups are free, the pK_a values are lower. If hemoglobin is held at a constant pH of 7, these protons dissociate upon oxygenation. This **Bohr effect**, described in 1904,[153–156] is important because acidification of hemoglobin stabilizes the deoxy form. In capillaries in which oxygen pressure is low and in which carbon dioxide and lactic acid may have accumulated, the lowering of the pH causes oxyhemoglobin to release oxygen more efficiently. These effects are also strongly dependent on the presence of chloride ions.[139,142,143,157,158]

Just as the conformational equilibria in hemoglobin can be shifted by attachment of oxygen to the heme groups, so the binding of certain other molecules at different sites can also affect the conformation. Such compounds are called **allosteric effectors** or **regulators** because they bind at a site other than the "active site." They are considered in more detail in Chapter 9. An important allosteric effector for human hemoglobin is **2,3-bisphosphoglycerate**, a compound found in human red blood cells in a high concentration approximately equimolar with that of hemoglobin. One molecule of bisphosphoglycerate binds to a hemoglobin tetramer in the deoxy form with $K_f = 1.4 \times 10^5$ but has only half this affinity for oxyhemoglobin.[159] X-ray crystallography shows that bisphosphoglycerate binds between the two β chains of deoxyhemoglobin directly on the twofold axis (Fig. 7-26).[159] Because of the presence of 2,3-bisphosphoglycerate in erythrocytes the affinity of oxygen for hemoglobin in whole blood is less than that for isolated hemoglobin[160,161] (Fig. 7-24). This is important because it allows a larger fraction of the oxygen carried to be unloaded from red corpuscles in body tissues. The bisphosphoglycerate level of red cells varies with physiological conditions, e.g., people living at higher elevations have a higher concentration.[161] It has been suggested that artificial manipulation of the level of this regulatory substance in erythrocytes may be of clinical usefulness for disorders in oxygen transport.



2,3-Bisphosphoglycerate

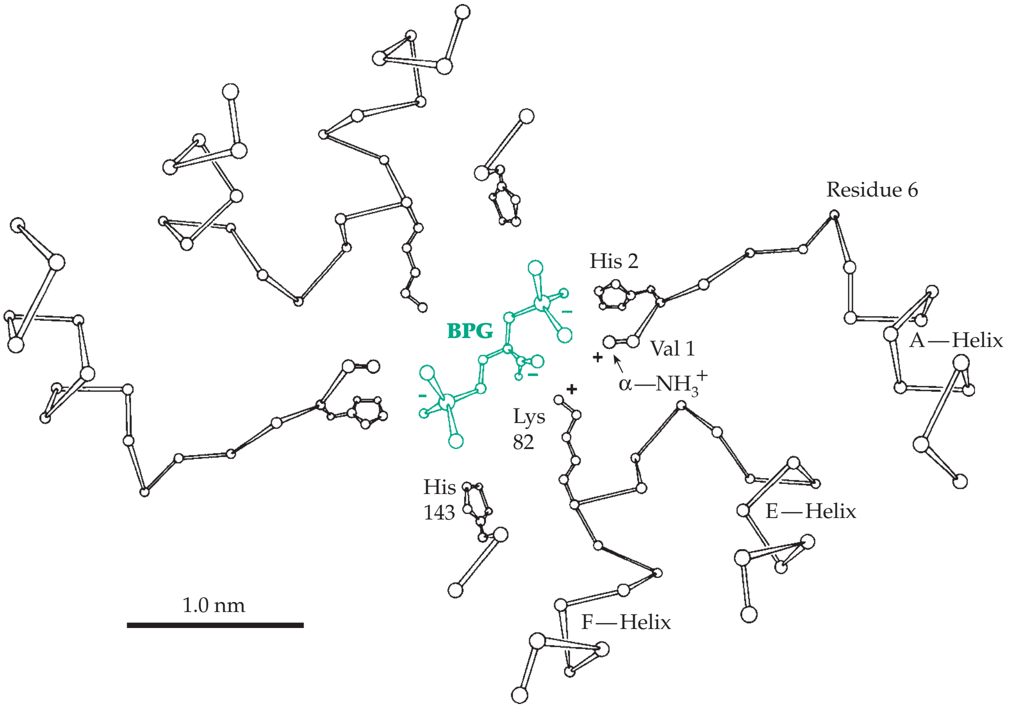
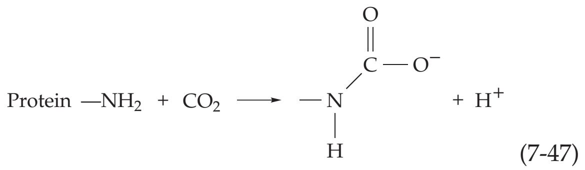


Figure 7-26 The allosteric effects of 2,3-bisphosphoglycerate (BPG) bound to the β chains of human deoxyhemoglobin. The phosphate groups of the BPG form salt bridges with valines 1 and histidines 2 and 143 of both β chains and with lysine 82 of one chain. This binding pulls the A helix and residue 6 toward the E helix. From Arnone.[159]

Not all species contain 2,3-bisphosphoglycerate in their erythrocytes. In birds and turtles its function appears to be served by inositol pentaphosphate. In crocodiles the site between the two β chains that binds organic phosphates in other species has been modified so that it binds bicarbonate ion, HCO_3^-, specifically. This ion, which accumulates in tissues as crocodiles lie under water, acts as an allosteric regulator in these animals.[162,163] It allows the animals to more completely utilize the O_2 from the hemoglobin and to remain under water longer. The Bohr effect, which was considered in the preceding section, can be viewed as resulting from the action of **protons** as allosteric effectors that bind to the amino and imidazole groups of the salt linkages. **Carbon dioxide** also acts as a physiological effector in mammalian blood by combining reversibly with NH_2-terminal groups of the α and β subunits to form **carbamino** (– NH – COO$^-$) groups (Eq. 7-47).[119,164] It is the N-terminal amino groups rather than lysyl side chain groups that undergo this reaction. Because of their relatively low pK_a values there is a significant fraction of unprotonated $-NH_2$ groups at the pH of blood. The affinity for CO_2 is highest in deoxygenated hemoglobin. Consequently, unloading of O_2 is facilitated in the CO_2-rich respiring tissues. Hemoglobin carries a significant fraction of CO_2 to the lungs, and there the oxygenation of hemoglobin facilitates the dissociation of CO_2 from the carbamino groups. Hemoglobin is also one of the major pH buffers of blood.



$$\text{Protein}-NH_2 + CO_2 \longrightarrow -\underset{H}{\underset{|}{N}}-\overset{O}{\overset{\|}{C}}-O^- + H^+ \qquad (7\text{-}47)$$

Carbon monoxide, cyanide, and nitric oxide. A danger to hemoglobin and other heme proteins is posed by competing ligands such as CO, CN^-, and NO. All of these are present within organisms and both CO and NO act as hormones. Hemoglobin and myoglobin are partially protected from carbon monoxide by the design of the binding site for O_2. The distal imidazole of histidine E7 hydrogen bonds to O_2 but not to the nonpolar CO. The site also accommodates the geometry of the bound O_2 better than that of CO.[165] Bound CO can be released from hemes by the action of light. Using X-ray diffraction[166,166a] and X-ray absorption measurements[167] at cryogenic temperatures, it has been possible to observe the motions of both the released CO and the heme in myoglobin, motions which may shed light on the normal oxygen transport cycle. Cooperativity in the binding of CO to hemoglobins has been studied extensively,[158,168,169] as has binding to model heme compounds.[170] Cyanide ions bind most tightly and also cooperatively[171–173] to the oxidized Fe^{3+} form, which is called **methemoglobin**.

Nitric oxide is a reactive, paramagnetic gaseous free radical which is formed in the human body and in other organisms by an enzymatic oxidation of L-arginine (Eq. 18-65). Since about 1980, NO has been recognized as a hormone with a broad range of effects

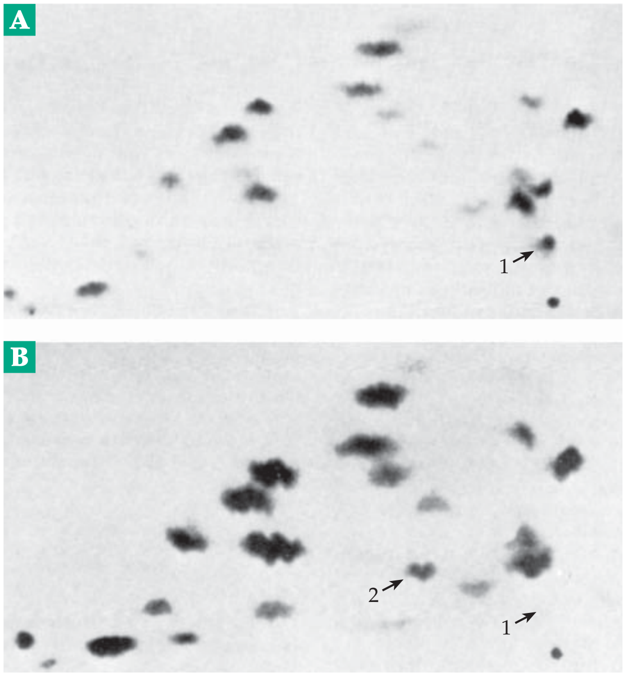


Figure 7-27 "Fingerprints" of human hemoglobins. The denatured hemoglobin was digested with trypsin and the 28 resulting peptides were separated on a sheet of paper by electrophoresis in one direction (horizontal in the figures; anode to the left) and by chromatography in the other direction (vertical in the figure). The peptides were visualized by spraying with ninhydrin or with specific reagents for histidine or tyrosine residues. Since trypsin cuts only next to lysine, which occurs infrequently, the petide pattern provides a fingerprint, characteristic for any pure protein. (A) The fingerprint of normal adult hemoglobin A. (B) Fingerprint of hemoglobin S (sickle cell hemoglobin). One histidine-containing peptide (1) is missing and a new one (2) is present. This altered peptide contains the first eight residues of the N-terminal chain of the subunit of the protein. From H. Lehmann and R. G. Huntsman, *Man's Haemoglobin*.[187]

(Chapters 11, 18). It binds to the iron of heme groups in either the Fe^{2+} or Fe^{3+} form and also reacts with thiol groups of proteins and small molecules to form *S*-nitrosothiols (R–S–N=O).[174–176] It reacts with the heme iron of myoglobin and hemoglobin and, by transfer of one electron, can oxidize the iron of hemoglobin to the Fe^{3+} methemoglobin with formation of the nitroxyl ion NO^-.[177,178] This reaction may be a major cause of methemoglobin formation.

One of the major effects of NO is to induce the relaxation of smooth muscle of blood vessels, an important factor in the regulation of blood pressure. Hemoglobin can carry NO both on its heme and on the thiol group of cysteine β93.[174,175] The affinity for NO is high in the T state and low in the R state. This allows hemoglobin to carry NO from the lungs to tissues, where it can be released and participate in the regulation of blood pressure.[174,179]

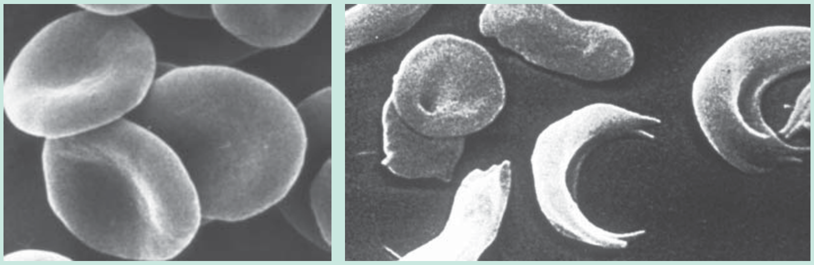
A cytoplasmic hemoglobin of the clam *Lucina pectinata* has evolved to carry oxygen to symbiotic chemoautotrophic bacteria located within cells of the host's gills. It is also readily oxidized to the Fe^{3+} methemoglobin form which binds **sulfide ions** extremely tightly[180–182] and is thought to transport sulfide to the bacteria.

2. Abnormal Human Hemoglobins

Many alterations in the structure of hemoglobin have arisen by mutations in the human population. It is estimated that one person in 20 carries a mutation that will cause a hemoglobin disorder in a homozygote.[183] There are also many unrecognized and harmless substitutions of one amino acid for another. However, substitutions near the heme group often adversely affect the binding of oxygen and substitutions in the interfaces between subunits may decrease the cooperative interaction between subunits.[184] One of the most common and serious abnormal hemoglobins is **hemoglobin S**, which is present in individuals suffering from **sickle cell disease** (see Box 7-B). In Hb S, glutamic acid 6 of the β chain is replaced by valine. Replacement of the same amino acid by lysine leads to **Hb C**[185] and is associated with a mild disease condition . A few of the many other substitutions that have been studied are indicated in Fig. 7-23. The locations of the defects in the hemoglobin structure have been established with the aid of protein "fingerprinting" (Fig. 7-27).

A group of serious defects are represented by the **hemoglobins M**. Only heterozygotic individuals survive. Their blood is dark because in Hb M the iron in half of the subunits is oxidized irreversibly to the ferric state. The resulting methemoglobin is present in normal blood to the extent of about 1%. While normal methemoglobin is reduced by a **methemoglobin reductase** system (Box 15-H), methemoglobins M cannot be reduced. All of the five hemoglobins M result from substitutions near the heme group. In four of them, one of the heme-linked histidines (either F-8 or E-7) of either the α or the β subunits is substituted by tyrosine. In the fifth, valine 67 of the β chains is substituted by glutamate. The two hemoglobins M that carry substitutions in the α subunits (M_{Boston} and M_{Iwate}) are frozen in the T (deoxy) conformation and therefore have low oxygen affinities and lack cooperativity.

BOX 7-B SICKLE CELL DISEASE, MALARIA, AND BLOOD SUBSTITUTES



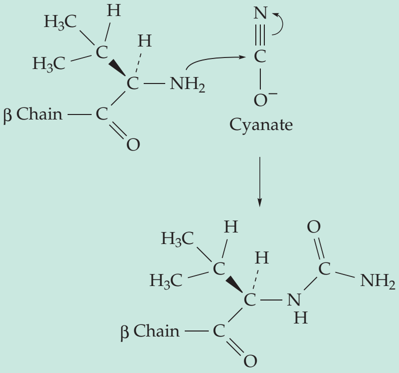
Left: Normal erythrocytes, © Biophoto Associates, Photo Researchers. Right: Sickled erythrocytes, © Nigel Calder.

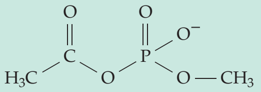
Many persons, especially if they are of west African descent, suffer from the crippling and often lethal sickle cell disease.[a,b] In 1949, Pauling, Itano, and associates discovered that hemoglobin from such individuals migrated unusually rapidly upon electrophoresis.[c] Later, Ingram devised the method of protein fingerprinting illustrated in Fig. 7-27 and applied it to hemoglobin.[d] He split the hemoglobin molecule into 15 tryptic peptides which he separated by electrophoresis and chromatography. From these experiments the abnormality in sickle cell hemoglobin (hemoglobin S; Hb S) was located at position 6 in the β chain (see Fig. 7-23). The glutamic acid present in this position in hemoglobin A was replaced by valine in Hb S. This was the first instance in which a genetic disease was traced directly to the presence of a single amino acid substitution in a specific protein. The DNA of the normal gene for the β globin chain has since been sequenced and found to have the glutamic acid codon GAG at position 6. A single base change to GTG (see Table 5-5) causes the sickle cell mutation. Persons homozygous for this altered gene have sickle cell disease, while the much more numerous heterozygotes have, at most, minor problems.

When HbS is deoxygenated it tends to “crystallize” in red blood cells, which contain 33% by weight hemoglobin. The crystallization (actually gel formation) distorts the cells into a sickle shape and these distorted corpuscles are easily destroyed, leading to anemia. The introduction of the hydrophobic valine residue in Hb S at position 6 near the end of the molecule helps form a new bonding domain by which the hemoglobin tetramers associate to form long semicrystalline microfilamentous arrays.[b,e–g]

Why is there such a high incidence of the sickle cell gene, estimated to be present in three million Americans? The occurrence and spread of the gene in Africa was apparently the result of a balance between its harmful effects and a beneficial effect under circumstances existing there. The malaria parasite, the greatest killer of all time, lives in red blood cells during part of its life cycle (see Fig. 1-9). Red cells that contain Hb S as well as Hb A are apparently less suitable than cells containing only Hb A for growth of the malaria organism. Thus, heterozygotic carriers of the sickle cell gene survived epidemics of malaria but at the price of seeing one-fourth of their offspring die of sickle cell disease.

What is the outlook for the many (50,000 in the United States alone) sufferers of sickle cell disease today? Careful medical care, including blood transfusion, can prolong life greatly[h] and intense efforts are under way to find drugs that will prevent Hb S from crystallizing.[i] The problem arises from a hydrophobic interaction of valine B6 with phenylalanine B85 and leucine B88 of another molecule in the filaments of Hb S. The latter two residues are on the outside surface of helix F (see Fig. 7-23). It is difficult to modify one of these residues chemically but various alterations at the nearby N-termini of the β chains do inhibit sickling. Cyanate does so by specifically carbamoylating these amino groups. However, although it was tested in humans, cyanate is too toxic for use.[j] Another approach employs an aldehyde that will form Schiff bases (Eq. 13-4)[k] with the same amino groups.[k,l] A third approach is to use an acylating reagent. For example, methylacetyl phosphate[m] acetylates the same β Lys 82 amino groups that react with bisphosphoglycerate and with cyanate.

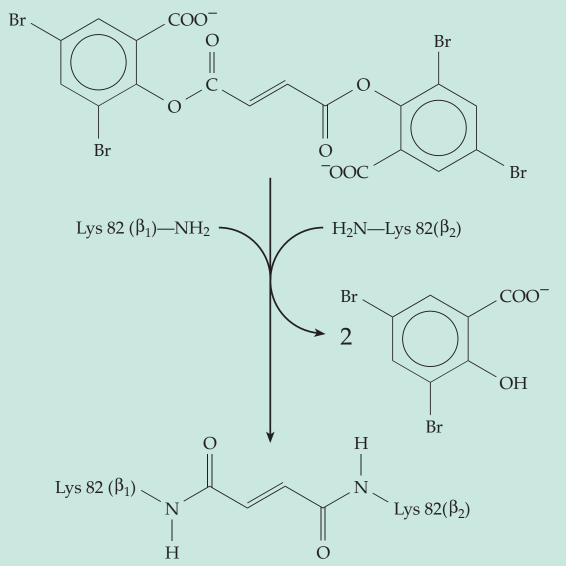




Aspirin (2-acetoxybenzoic acid) is also a mild acetylating reagent and “two-headed” aspirins such

BOX 7-B (continued)

as the following react specifically to crosslink the hemoglobin β chains.



These compounds bind into the bisphosphoglycerate binding site (Fig. 7-26) and prevent the chains from spreading apart as far as they normally do in the deoxy (T) state. Since it is only the latter that crystallizes, the compounds have a powerful antisickling action.[n,o] Various other crosslinking reagents have been developed and more than one could be used together.[p–r] New drugs that serve as allosteric modifiers in the same fashion as bisphosphoglycerate may also be useful.[s]

A fouth approach to treatment of sickle cell disease is gene therapy. This might allow patients to produce, in addition to Hb S, an engineered hemoglobin with compensating mutations that would mix with the Hb S and prevent gelling.[t,u,v] This is impractical at present but there is another approach. Persons with sickle cell disease sometimes also have the disorder of hereditary persistence of fetal hemoglobin. They continue to make Hb F into adulthood. Great amelioration of sickle cell disease is observed in patients with 20–25% Hb F.[t] Hydroxyurea stimulates a greater production of Hb F and in patients with hereditary persistence of Hb F hydroxyurea may raise its level in erythocytes to ~50% of the total hemoglobin.[w]

Crosslinking of the *alpha* chains of normal deoxyhemoglobin through lysines 99 yields a hemoglobin with normal oxygen-binding behavior and an increased stability.[o] It makes a practical emergency blood substitute, whereas unmodified hemoglobin is unsatisfactory.[v] Unless encapsulated in an erythrocyte the hemoglobin tetramers tend to dissociate to dimers, losing cooperativity and escaping through kidneys. Suitable crosslinking helps to solve this problem.[x,y,z] Both α and β chains can be produced from cloned genes and reassembled to form hemoglobin.[aa,bb] This will probably allow genetic engineering to form more stable but suitably cooperative hemoglobins that can be used to avoid hazards of transmission of viruses by transfusion.

---

[a] Weatherall, D. J., Clegg, J. B., Higgs, D. R., and Wood, W. G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3417–3484, McGraw-Hill, New York

[b] Harrington, D. J., Adachi, K., and Royer, W. E., Jr. (1998) *J. Biol. Chem.* **273**, 32690–32696

[c] Strasser, B. J. (1999) *Science* **286**, 1488–1490

[d] Ingram, V. M. (1957) *Nature (London)* **180**, 326–328

[e] Cretegny, I., and Edelstein, S. J. (1993) *J. Mol. Biol.* **230**, 733–738

[f] Padlan, E. A., and Love, W. E. (1985) *J. Biol. Chem.* **260**, 8272–8279

[g] Cao, Z., and Ferrone, F. A. (1996) *J. Mol. Biol.* **256**, 219–222

[h] Acquaye, C., Wilchek, M., and Gorecki, M. (1981) *Trends Biochem. Sci.* **6**, 146–149

[i] Klotz, I. M., Haney, D. N., and King, L. C. (1981) *Science* **213**, 724–731

[j] Harkness, D. R. (1976) *Trends Biochem. Sci.* **1**, 73–

[k] Acharya, A. S., Sussman, L. G., and Manning, J. M. (1983) *J. Biol. Chem.* **258**, 2296–2302

[l] San George, R. C., and Hoberman, H. D. (1986) *J. Biol. Chem.* **261**, 6811–6821

[m] Ueno, H., Pospischil, M. A., and Manning, J. M. (1989) *J. Biol. Chem.* **264**, 12344–12351

[n] Walder, J. A., Walder, R. Y., and Arnone, A. (1980) *J. Mol. Biol.* **141**, 195–216

[o] Chatterjee, R., Welty, E. V., Walder, R. Y., Pruitt, S. L., Rogers, P. H., Arnone, A., and Walder, J. A. (1986) *J. Biol. Chem.* **261**, 9929–9937

[p] Benesch, R. E., and Kwong, S. (1988) *Biochem. Biophys. Res. Commun.* **156**, 9–14

[q] Kluger, R., Wodzinska, J., Jones, R. T., Head, C., Fujita, T. S., and Shih, D. T. (1992) *Biochemistry* **31**, 7551–7559

[r] Jones, R. T., Shih, D. T., Fujita, T. S., Song, Y., Xiao, H., Head, C., and Kluger, R. (1996) *J. Biol. Chem.* **271**, 675–680

[s] Abraham, D. J., Wireko, F. C., Randad, R. S., Poyart, C., Kister, J., Bohn, B., Liard, J.-F., and Kunert, M. P. (1992) *Biochemistry* **31**, 9141–9149

[t] McCune, S. L., Reilly, M. P., Chomo, M. J., Asakura, T., and Townes, T. M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9852–9856

[u] Cole-Strauss, A., Yoon, K., Xiang, Y., Byrne, B. C., Rice, M. C., Gryn, J., Holloman, W. K., and Kmiec, E. B. (1996) *Science* **273**, 1386–1388

[v] May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K. M. L., Luzzatto, L., and Sadelain, M. (2000) *Nature (London)* **406**, 82–86

[w] Eaton, W. A., and Hofrichter, J. (1995) *Science* **268**, 1142–1143

[x] Snyder, S. R., Welty, E. V., Walder, R. Y., Williams, L. A., and Walder, J. A. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7280–7284

[y] Dick, L. A., Heibel, G., Moore, E. G., and Spiro, T. G. (1999) *Biochemistry* **38**, 6406–6410

[z] Manjula, B. N., Malavalli, A., Smith, P. K., Chan, N.-L., Arnone, A., Friedman, J. M., and Acharya, A. S. (2000) *J. Biol. Chem.* **275**, 5527–5534

[aa] Yamaguchi, T., Pang, J., Reddy, K. S., Witkowska, H. E., Surrey, S., and Adachi, K. (1996) *J. Biol. Chem.* **271**, 26677–26683

[bb] Jeong, S. T., Ho, N. T., Hendrich, M. P., and Ho, C. (1999) *Biochemistry* **38**, 13433–13442

In hemoglobins Rainier and Nancy the usually invariant C-terminal tyrosine 145 of the β chains is substituted by cysteine and by aspartate, respectively. Oxygen affinity is high and cooperativity is lacking.[186] Hemoglobin Kansas, in which the β^{102} asparagine is substituted by threonine, also lacks cooperativity and has a very low oxygen affinity, while hemoglobin Richmond, in which the same amino acid is substituted by lysine, functions normally. In hemoglobin Creteil the β89 serine is replaced by asparagine with the result that the adjacent C-terminal peptide carrying tyrosine 145 becomes disordered. In hemoglobin Hiroshima the C-terminal histidine in the β chain is replaced by aspartic acid. This histidine is one that donates a Bohr proton and in the mutant hemoglobin the oxygen affinity is increased 3-fold and the Bohr effect is halved.[188] In hemoglobin Suresnes the C-terminal arginines 141 of the α chains are replaced by histidine with loss of one of the anion-binding sites mentioned in Section 3.[189]

3. Comparative Biochemistry of Hemoglobin

Even within human beings there are several hemoglobins. In addition to myoglobin, a brain protein neuroglobin,[189a] and adult hemoglobin A (**Hb A**, $\alpha_2\beta_2$), there is a minor hemoglobin A_2 ($\alpha_2\delta_2$). Prior to birth the blood contains **fetal hemoglobin**, also called hemoglobin F (**Hb F**, $\alpha_2\gamma_2$). In the presence of 2,3-bisphosphoglycerate Hb F has a 6-fold higher oxygen affinity than Hb A as befits its role in obtaining oxygen from the mother's blood.[190–192] Hemoglobin F disappears a few months after birth and is replaced by Hb A. Each of the hemoglobins differs from the others in amino acid sequence.

In other species the amino acid composition of hemoglobins varies more, as do the interactions between subunits. Hemoglobins and myoglobins are found throughout the animal kingdom and even in plants.[192a] The **leghemoglobins**[193–195] are formed in root nodules of legumes and are involved in nitrogen fixation by symbiotic bacteria. Other hemoglobins apparently function in the roots of plants.[196–197] Hemoglobins or myoglobins are found in some cyanobacteria[198] and in many other bacteria.[197,199] The globin fold of the polypeptide is recognizable in all of these.[200]

The quaternary structure of hemoglobin also varies. Myoglobin is a monomer, as is the leghemoglobin. Hemoglobin of the sea lamprey dissociates to monomers upon oxygenation.[201,201a] The clam *Scapharca inaequivalvis* has a dimeric hemoglobin that binds O_2 cooperatively even though the interactions between subunits are very different from those in mammalian hemoglobins.[202–204] Hemoglobin of the nematode *Ascaris* is an octamer.[205,206] It has puzzling properties, including a very high affinity for O_2 and a slow dissociation rate. The distal His E-7 is replaced by glutamine, which has a hydrogen-bonding ability closely similar to that of histidine.

Earthworms,[207] polychaete worms,[208,209] and leeches[210] have enormous hemoglobin molecules consisting of as many as 144 globin chains arranged into 12 dodecamers and held together by 36–42 linker chains. These hemoglobins are often called **erythrocruorins**. In a few families of polychaetes chloroheme (Fig. 16-5) substitutes for heme and the proteins are called **chlorocruorins**.[208]

What is common to all of the hemoglobins? The same folding pattern of the peptide chain is always present. The protein is always wrapped around the heme group in an identical or very similar manner. In spite of this striking conservation of overall structure, when animal hemoglobins are compared, *there are only ten residues that are highly conserved*. They are indicated in Fig. 7-23 by the boxes. The two glycines (or alanine) at B-6 and E-8 are conserved because the close contact between the B and E helices does not permit a larger side chain. Proline C-2 helps the molecule turn a corner. Four of the other conserved residues are directly associated with the heme group. Histidine E-7 and His F-8 are the "heme-linked" histidines. Tyrosine HC-2, as previously mentioned, plays a role in the cooperativity of oxygen binding. Only Lys H-9 is on the outside of the molecule. The reasons for its conservation are unclear.[211] When sequences from a broader range of organisms were determined five residues (see Fig. 7-23) were found to be highly conserved; *only two are completely conserved.* These are His F-8 and Phe CD-1, which binds the heme noncovalently.[212] Hemoglobins are not the only biological oxygen carriers. The two-iron **hemerythrins** (Fig. 16-20) are used by a few phyla of marine invertebrates, while the copper containing **hemocyanins** (Chapter 16, Section D,4) are used by many molluscs and arthropods.

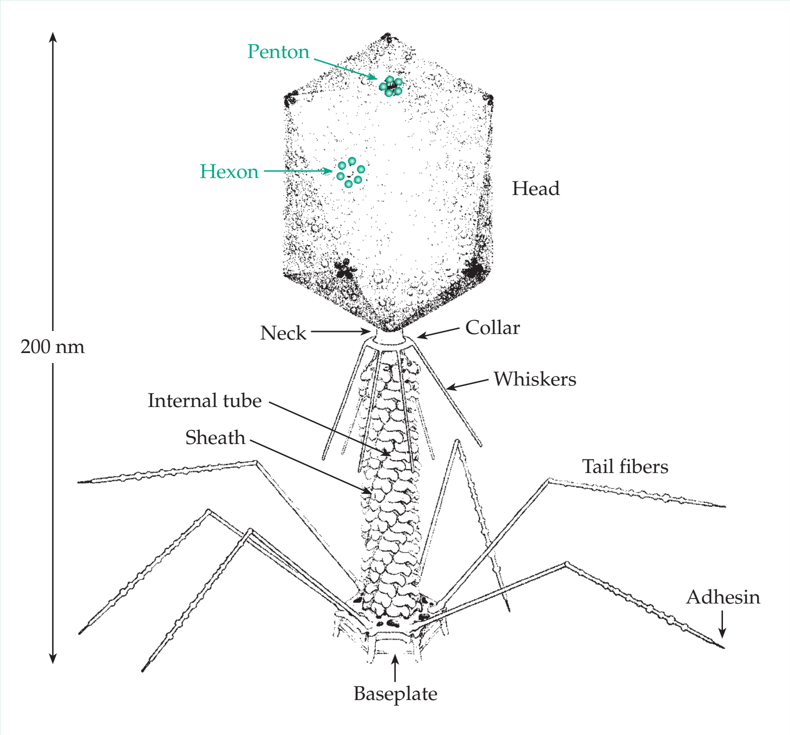
E. Self-Assembly of Macromolecular Structures

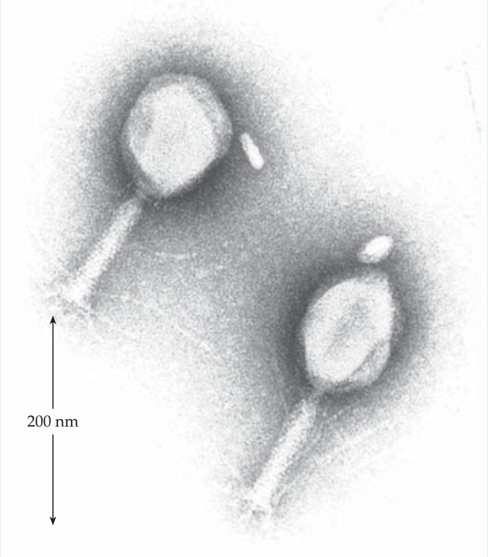
While it is easy to visualize the assembly of oligomeric proteins, it is not as easy to imagine how complex objects such as eukaryotic cilia (Fig. 1-8) or the sarcomeres of muscle (Fig. 19-6) are formed. However, study of the assembly of bacteriophage particles and other small biological objects has led to the concepts of **self-assembly** and **assembly pathways**, concepts that are now applied to every aspect of the architecture of cells.

1. Bacteriophages

A remarkable example of self-assembly is that of the T-even phage (Box 7-C).[213–215] From genetic analysis (Chapter 26) at least 22 genes are known to be required for formation of the heads, 21 genes for the tails, and 7 genes for the tail fibers. Many of these genes encode

BOX 7-C THE T-EVEN BACTERIOPHAGES





Bacteriophage T4. × 240,000
Micrograph courtesy of Tom Moninger.

Drawing courtesy of F. Eiserling and the American Society for Microbiology.[a]

Among the most remarkable objects made visible by the electron microscope are the T-even bacteriophage (T2, T4, and T6) which attack *E. coli*.[a–e] While it is not often evident how a virus gains access to a cell, these "molecular syringes" literally inject their DNA through a hole dissolved in the cell wall of the host bacterium. The viruses, of length ~200 nm and mass ~225 x 10^6 Da each, contain 130 x 10^6 Da of DNA in a 100 x 70 nm head of elongated icosahedral shape. The head surface appears to be formed from ~840 copies of a 45-kDa protein known as gp23 (gene product 23; it is encoded by gene 23; see map in Fig. 26-2). These protein molecules are arranged as 140 hexamers (hexons) and together with ~55 copies of protein gp24, arranged as 11 pentamers (pentons), make up the bulk of the shell.[f] The head also contains at least nine other proteins, including three internal, basic proteins that enter the bacterium along with the DNA. Additional proteins form the **neck**, **collar**, and **whiskers**. The phage **tail**, which fastens to the collar via a **connector** protein,[g] contains an **internal tube** with a 2.5-nm hole, barely wide enough to accommodate the flow of the DNA molecule into the bacterium. The tube is made up of 144 subunits of gp19. The 8 x 10^6-Da **sheath** that surrounds the tail tube is made up of 144 subunits of gp18, each of mass 55 kDa, arranged in the form of 24 rings of six subunits each.[h,i] The sheath has contractile properties. After the virus has become properly attached to the host it shortens from ~80 to ~30 nm, forcing the inner tube through a hole etched in the wall of the bacterium. At the end of the tail is a **baseplate**, a hexagonal structure bearing six short **pins**, each a trimer of a 55 kDa zinc metalloprotein.[j] One of the ten proteins known to be present in the baseplate is the enzyme T4 lysozyme (Chapter 12). The baseplate also contains six molecules of the coenzyme 7,8-dihydropteroylhexaglutamate (Chapter 15, Section D).

Six elongated, "jointed" **tail fibers** are attached to the baseplate. The proximal segment of each fiber is a trimer of the 1140-kDa protein gp34. A globular domain attaches it to the baseplate. The distal segment is composed of three subunits of the 109-kDa gp37, three subunits of the 23-kDa gp36, and a single copy of the 30-kDa gp35.[k] The C-terminal ~140 residues of the 1026-residue gp37 are the specific **adhesin** that binds to a lipopolysaccharide of *E. coli* type B cells or to the outer membrane protein OmpC (Chapter 8).[l] Among the smaller molecules present in the virus are the polyamines **putrescine** and **spermidine** (Chapter 24), which neutralize about 30% of the basic groups of the DNA.

How is infection by a T-even virus initiated? Binding of the tail fibers to specific receptor sites on the bacterial surface triggers a sequence of conformational changes in the fibers, baseplate, and

BOX 7-C THE T-EVEN BACTERIOPHAGES (continued)

sheath. The lysozyme is released from the baseplate and etches a hole in the bacterial cell wall. Contraction of the sheath is initiated at the baseplate and continues to the upper end of the sheath. The tail tube is forced into the bacterium and the DNA rapidly flows through the narrow hole into the host cell.

During contraction the subunits of the sheath undergo a remarkable rearrangement into a structure containing 12 larger rings of 12 subunits each.[h] Thus, a kind of mutual "intercalation" of subunits occurs. In its unidirectional and irreversible nature the shortening of the phage tail differs from the contraction of muscle. The protein subunits of the sheath seem to be in an unstable high energy state when the tail sheath of the phage is assembled. The stored energy remains available for later contraction.

[a] Mathews, C. K., Kutter, E. M., Mosig, G., and Berget, P. B., eds. (1983) *Bacteriophage T_4*, Am. Soc. Microbiology, Washington, D.C.

[b] Wood, W. B., and Edgar, R. S. (1967) *Sci. Am.* **217**(Jul), 60–74

[c] Mathews, C. K. (1971) *Bacteriophage Biochemistry*, Van Nostrand-Reinhold, Princeton, New Jersey

[d] Voyles, B. A. (1993) *The Biology of Viruses*, Mosby, St. Louis, Missouri

[e] Tikhonenko, A. S. (1970) *Ultrastructure of Bacterial Viruses*, Plenum, New York

[f] Branton, D., and Klug, A. (1975) *J. Mol. Biol.* **92**, 559–565

[g] Cerritelli, M. E., and Studier, F. W. (1996) *J. Mol. Biol.* **258**, 299–307

[h] Moody, M. F., and Makowski, L. (1981) *J. Mol. Biol.* **150**, 217–244

[i] Müller, D. J., Engel, A., Carrascosa, J. L., and Vélez, M. (1997) *EMBO J.* **16**, 2547–2553

[j] Zorzopulos, J., and Kozloff, L. M. (1978) *J. Biol. Chem.* **253**, 5543–5547

[k] Cerritelli, M. E., Wall, J. S., Simon, M. N., Conway, J. F., and Steven, A. C. (1996) *J. Mol. Biol.* **260**, 767–780

[l] Tétart, F., Repoila, F., Monod, C., and Krisch, H. M. (1996) *J. Mol. Biol.* **258**, 726 – 731

---

sequences of proteins that are incorporated into the mature virus, but several specify enzymes needed in the assembly process. Several mutant strains of the viruses are able to promote synthesis of all but one of the structural proteins of the virion. Proteins accumulating within these defective bacterial hosts have no tendency to aggregate spontaneously. However, when the missing protein (synthesized by bacteria infected with another strain of virus) is added complete virus particles are formed rapidly. Investigations resulting from this and other observations have led to the conclusion that during assembly *each different protein is added to the growing aggregate in a strictly specified sequence or assembly pathway.* The addition of each protein creates a binding site for the next protein. In some cases the protein that binds is an enzyme that cuts off a piece from the growing assembly of subunits and thereby creates a site for the next protein to bind.

Before considering this complex process further, let's look at the assembly of simpler filamentous bacteriophages (Fig. 7-7) and bacterial pili (Fig. 7-9). The filamentous bacteriophages are put together from hydrophobic protein subunits and DNA. After their synthesis the protein subunits are stored within the cytoplasmic membrane of the infected bacteria.[216,217] These small, largely α-helical rods can easily fit within the membrane and remain there until a DNA molecule also enters the membrane.[218,219] Two additional proteins (gene I proteins) also enter the membrane. One has a 348-residue length, while the second is a 107-residue protein formed by translational initiation at a later point in the DNA sequence of the gene. These proteins help to create an assembly site at a place where the inner and outer membrane of the host bacterium are close together.[220] It isn't clear how the process is initiated, but it is likely that each subunit of the viral coat contains a nucleotide binding site that interacts with the DNA. Adjacent sides of the subunits are hydrophobic and interact with other subunits to spontaneously coat the DNA. As the rod is assembled, the hydrophobic groups become "buried." It is postulated that the remaining side chain groups on the outer surface of the virus are hydrophilic and that the formation of this hydrophilic rod provides a driving force for automatic extrusion of the phage from the membrane.[216]

Bacterial pili appear to be extruded in a similar manner. They arise rapidly and may possibly be retracted again into the bacterial membrane. The P pilus in Fig. 7-9A is made up of subunits PapA, G, F, E, and K which must be assembled in the correct sequence. A chaperonin PapD is also required as is an "**usher protein**," PapC,[50] and also the disulfide exchange protein DsbA (Chapter 10). DsbA helps PapD to form the correct disulfide bridges as it folds and PapD binds and protects the various pilus subunits as they accumulate in the periplasmic space of the host. The usher protein displaces the chaperonin PapD and "escorts" the subunits into the membrane where the extrusion occurs.[50,55]

Because eicosahedra are regular geometric solids and the faces can be made up of hexons and pentons of identical subunits, it might seem that self-assembly of eicosahedral viruses would occur easily. However, the subunits usually must be able to assume three or more different conformations and the shells can easily be assembled incorrectly. Several stategies are employed to avert this problem.[221,222] Some viruses assemble an empty shell into which the DNA flows, but many others first form an internal **scaffolding** or **assembly core** around which the shell is assembled. An external scaffold may also be needed.[223] The RNA virus MS2 forms its T=3 capsid by using the RNA molecule as the assembly core.[224] Other viruses may have one or more core proteins which dissociate from the completed shell or are removed by the action of proteases. This is a feature of the small ϕX phage (Fig. 7-28),[225,226,230] the tailed phages,[227,228] and double-stranded RNA viruses including human reoviruses.[229] Bacteriophage PRD1, another virus of *E. coli* and *Salmonella typhimurium*, has a membrane inside the capsid apparently playing a role in assembly.[232]

A general concept that seems to hold in all cases is one of "**local rule**." Several conformers of a virus subunit may equilibrate within a cell. However, *they*

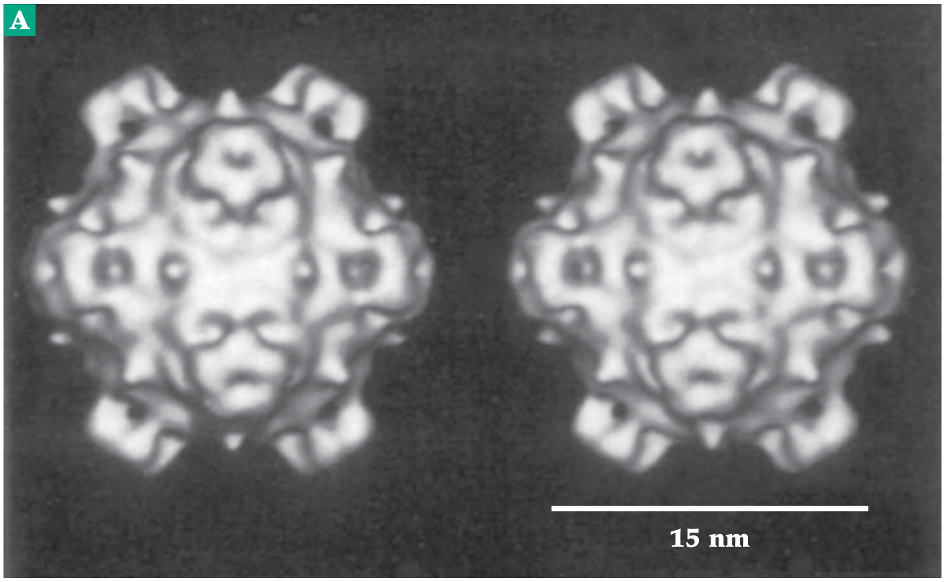
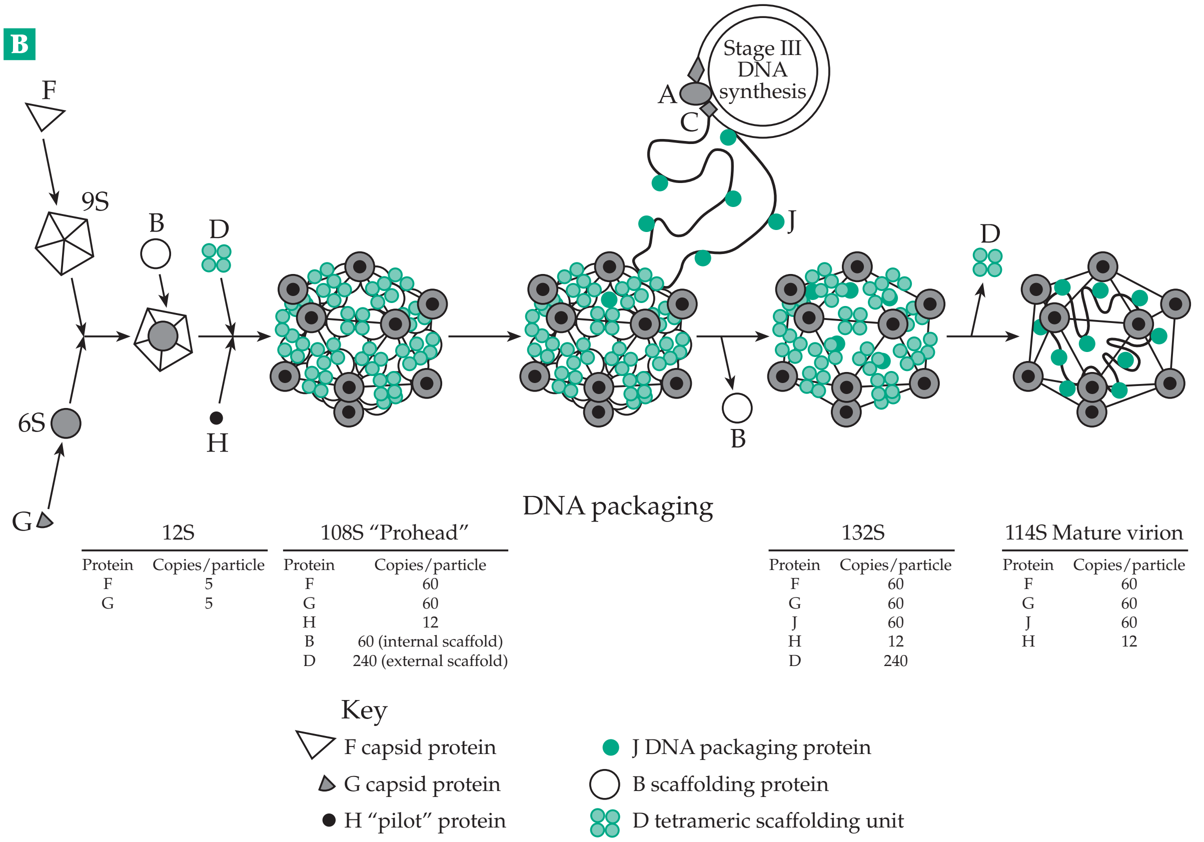


Figure 7-28 (A) Stereoscopic view of the ϕX174 114 S mature virion viewed down a twofold axis after a cryoelectron microscopy reconstruction. From McKenna *et al.*[225] (B) Morphogenesis of ϕX174 (based on a report of Hayashi *et al.*[231]). Proteins A and C are required for DNA synthesis. Drawing from McKenna *et al.*[225] Courtesy of Michael G. Rossmann.


B
F
9S
B
D
6S
H
G
A
C
Stage III DNA synthesis
J
D
B
DNA packaging
Key
F capsid protein
G capsid protein
H "pilot" protein
J DNA packaging protein
B scaffolding protein
D tetrameric scaffolding unit

12S

Protein	Copies/particle
F	5
G	5

108S "Prohead"

Protein	Copies/particle
F	60
G	60
H	12
B	60 (internal scaffold)
D	240 (external scaffold)

132S

Protein	Copies/particle
F	60
G	60
J	60
H	12
D	240

114S Mature virion

Protein	Copies/particle
F	60
G	60
J	60
H	12

can associate only through surfaces that are complementary. A conformer that allows pentons to form cannot assemble into a hexon and only certain combinations of other conformers can give rise to hexons, etc.[233] If there is only one conformation and the shape is right a T=1 shell will be formed. If there are three conformers a T= 3 shell may arise. Another generalization is that in most cases the **procapsid** or **prohead** that is formed initially is fragile. Subunits may still be undergoing conformational changes. However, a final conformational alteration, which may include chain cleavage by a protease, usually occurs. This often expands the overall dimensions of the capsid and creates new intersubunit interactions which greatly strengthen the mature capsid.[234] Scaffolding proteins are then removed and DNA or RNA enters the capsid, again in a precise sequence. There are many variations and the detail that is known about virus assembly is far too great to describe here.

Figure 7-29 illustrates the assembly pathway for the very small ϕX174, a T=1 virus. The major capsid protein F is a 426-residue eight-stranded β-barrel. The 175-residue G protein forms pentameric spikes while 60 copies of the internal scaffolding protein and 240 copies of the external scaffolding protein D and 12 copies of the pilot protein are required to form the prohead. The single-stranded DNA enters along with 60 copies of a DNA packaging protein J.

Assembly of the tailed bacteriophages (Box 7-C) is even more complex. The genome of the viruses is large. The 166-kb circular dsDNA of phage T4 contains ~250 genes, many of which encode proteins of the virion or enzymes or chaperonins needed in assembly. The assembly pathway for the bacteriophage heads requires at least 22 gene products.[235] Seven of these form the assembly core which serves as a scaffolding around which the 840 copies of gp23 and 55 copies of gp24 (see Box 7-C) are added to give the elongated icosahedral prohead I. Most of the internal proteins are then dissolved by proteases, one of which is the phage-specified gp21. A protease also cuts a piece from each molecule of gp23 to form the smaller gp23*, the major protein of the mature **prohead II**. This cleavage also triggers the conformational change leading to head expansion. The empty proheads are now filled with DNA in a process which is assisted by another series of catalytic proteins.

The T4 phage tail is assembled in a separate sequence. Six copies of each of three different proteins form a "hub" with hexagonal symmetry (Fig. 7-29). In another assembly sequence, seven different proteins form wedge-shaped pieces, six of which are then joined to the hub to form the hexagonal baseplate. Two more proteins then add to the surface of the base plate and activate it for the growth of the internal tail tube. Only after assembly of the internal tube begins does the sheath

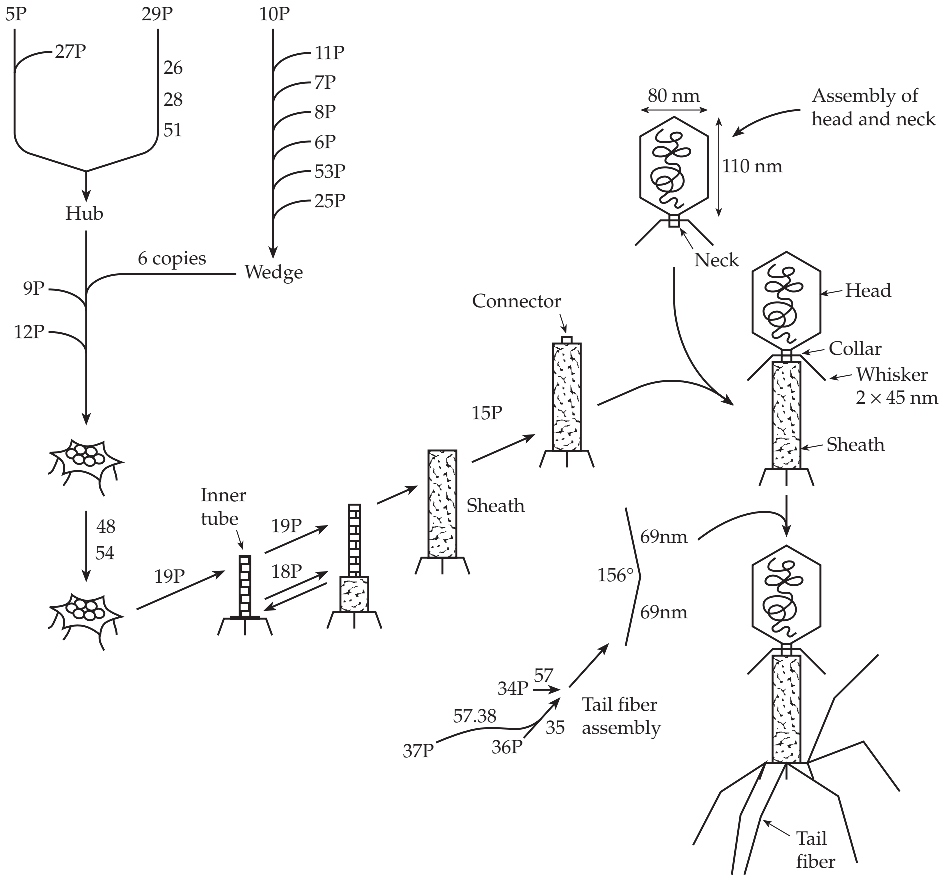


Figure 7-29 Assembly sequence for bacteriophage T4 with details for the tail. The numbers refer to the genes in the T4 chromosome map (Fig. 26-2). A "P" after the number indicates that the protein gene product is incorporated into the phage tail. Other numbers indicate gene products that are thought to have essential catalytic functions in the assembly process. Adapted from King and Mykolajewycz[236] and Kikuchi and King.[214]

begin to grow, and only when both of these tubular structures have reached the correct length, is a cap protein placed on top. The DNA-filled head is then attached by a special connector and only then do the tail fibers, which have been assembled separately, join at the opposite end.

How can each step in this complex assembly process set the stage for the next step? Apparently the structure of each newly synthesized protein monomer is stable only until a specific interaction with another protein takes place. The binding energy of this interaction is sufficient to induce a conformational alteration that affects a distant part of the protein surface and generates complementarity toward a binding site on the next protein that is to be added. Every one of the baseplate proteins must have such self-activating properties! Sometimes proteolytic cleavage of a subunit is required. If it occurs at an appropriate point in the sequence it provides thermodynamic drive for the assembly process.

The induction of a change in one protein by interaction with another protein is a phenomenon that is met also in the construction of microtubules, ribosomes, cilia, and myofibrillar assemblies of muscle. It is basic to the assembly of the many labile but equally real cascade systems of protein–protein interactions such as that involved in the clotting of blood (Chapter 12) and signaling at membrane surfaces.

2. "Kringles" and Other Recognition Domains and Motifs

The assemble of either transient or long lasting complexes of proteins is often dependent upon the presence of conserved structural domains of 30–100 residues. A similar domain may occur in many different proteins and often two or more times within a single protein. The sequences within such a domain are homologous, allowing it to be recognized from protein or gene sequences alone.[237,238] Domains are often named after the protein in which they were first discovered. For example, **EGF-like domains** resemble the 53-residue epidermal growth factor. **SH2** and **SH3** domains are src-homology domains, named after **Src** (c-src), the protein encoded by the *src* protooncogene (Table 11-3).[239] The SH2 and SH3 domains are found near the N terminus of this 60-kDa protein. They are also found in many other proteins. An adapter protein called **Grb2**, important in cell signaling, consists of nothing but one SH2 domain and two SH3 domains (Figs. 11-13, 11-14).[240] The SH2 domains bind to phosphotyrosyl side chains of various proteins, while SH3 domains bind to a polyproline motif. Another phosphotyrosyl binding domain, the plekstrin homology or **PH** domain, is named for the protein in which it was discovered. **Kringle** and **apple** describe the appearances of the folded proteins in those domains. Structural domains often function to hold two proteins together or to help anchor them at a membrane surface by binding to specific protein groups, such as phosphotyrosyl, or calcium ions. Table 7-3 lists a few well-known folding domains and Fig. 7-30 shows three-dimensional structures of two of them.

Recognition domains often function transiently. For example, SH2 domains are often found in proteins that interact with phosphotyrosyl groups of "activated" cell surface receptors. The receptors become activated

TABLE 7-3
A Few Well-Known Structural Domains

Name	Length in amino acid residues	Specific ligands
EGF-like[241–244]	~45	Ca^{2+}
SH2[238,245–247]	~100	Phosphotyrosine
Structure[248–253]		
SH3[239,246,254]	~60	Polyproline, PXXP
Structure[255–257]		
PTB[238,258,259]		Proline-rich sequence
PH (plekstrin homology)[260,261]		Phosphotyrosine
Structure[262,263]		
PDZ[264–266]	80–100	C-terminal XS / TXV–COO^-
Immunoglobulin repeat (Fig. 2-16)[267]	~100	
Kringles, blood clotting proteins[268–270]	80-85	Calcium binding
Apple, Blood clotting Factor X[271]	90	Calcium binding
WW (Trp–Trp)[272]	~38	Proline
Serine protease[273]		
P (Trefoil)[274]	~50	
TPR (Tetratrico peptide repeat)[247,275,276]		
ZBD (Zinc-binding domains)		
Zinc finger (Fig. 5-37)[277]		
Others[277–280]		

by conformational alteration resulting from the binding. The src protein is a tyrosine kinase and, when activated, uses ATP to phosphorylate tyrosyl groups of other proteins, and using its SH2 domains it will bind to such groups forming and passing an intracellular message to them.[249]

F. The Cytoskeleton

The cytoplasm of eukaryotic cells contains a complex network of slender rods and filaments that serve as a kind of internal skeleton. The properties of this **cytoskeleton** affect the shape and mechanical properties of cells. For example, the cytoskeleton is responsible

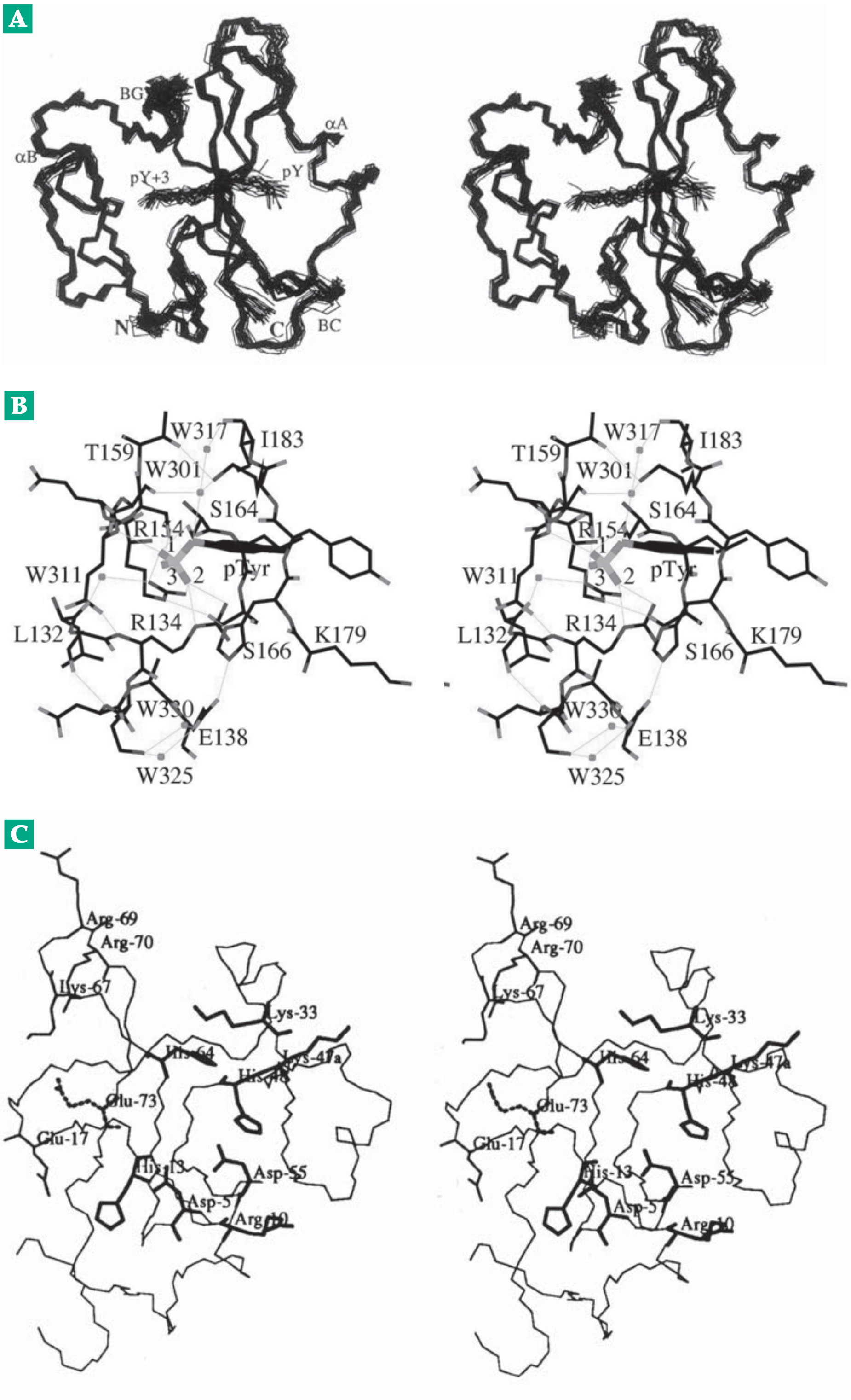

A
BG
αA
αB
pY+3
pY
N
C
BC
B
W317
I183
T159
W301
S164
R154
1
3
2
pTyr
W311
L132
R134
K179
S166
W330
E138
W325
C
Arg-69
Arg-70
Lys-67
Lys-33
His-64
Lys-47a
His-48
Glu-73
Glu-17
His-13
Asp-55
Asp-57
Arg-10

Figure 7-30 (A) Stereoscopic MolScript view showing 30 superposed solution structures of the SH2 / phosphopeptide complex from protein Shc calculated from NMR data. The N and C termini of the protein as well as the phosphotyrosine (pY) and (pY + 3) residues of the phosphopeptide are indicated. From Zhou *et al*.[252] Courtesy of Stephen W. Fesik. (B) View of the phosphotyrosine side chain of the peptide pYEEI bound to a high-affinity SH2 domain from the human *src* tyrosine kinase called p56^{lck}. The phosphate group forms a series of hydrogen bonds with groups in the protein and with water molecules (small dots) and an ion pair with the guanidinium group of R134. From Tong *et al*.[248] Courtesy of Liang Tong. (C) Structure of kringle 2 from human tissue plasminogen activator (see Chapter 12). From de Vos *et al*.[269] Courtesy of Abraham M. de Vos.

for the biconcave disc shape of erythrocytes and for the ameba's ability to rapidly interconvert gel-like and fluid regions of the cytoplasm.[281–283]

Three principal components of the cytoskeleton are **microfilaments** of ~6 nm diameter, **microtubules** of 23–25 nm diameter, and **intermediate filaments** of ~10 nm diameter. A large number of associated proteins provide for interconnections, for assembly, and for disassembly of the cytoskeleton. Other proteins act as **motors** that provide motion. One of these motors is present in **myosin** of muscle. This protein is not only the motor for muscular work but also forms **thick filaments** of 12–16 nm diameter, which are a major structural component of muscle (see Fig. 19-6).

1. Intermediate Filaments

In most cells the intermediate filaments provide the scaffolding for the cytoskeleton[284–286] They may account for only 1% of the protein in a cell but provide up to 85% of the protein in the tough outer layers of skin. Intermediate filament proteins are encoded by over 50 human genes[286] which specify proteins of various sizes, structures, and properties. However, all of them have central 300- to 330-residue α-helical regions through which the molecules associate in parallel pairs to form coiled-coil rods with globular domains at the ends (Fig. 7-31). Some of these proteins, such as the **keratin** of skin, are insoluble. Others, including the nuclear **lamins** (Chapter 27)[287] and **vimentin**,[288–289a] dissociate and reform filaments reversibly.

Vimentin is found in most cells and predominates in fibroblasts and other cells of mesenchymal origin. **Desmin** (55-kDa monomer) is found in both smooth and skeletal muscle.[289b,290,290a] In the latter, it apparently ties the contractile myofibrils to the rest of the cytoskeletal network and the individual myofibrils to each other at Z disc (see Fig. 19-6). The **glial filaments** from the astroglial cells of the brain are composed mainly of a single type of 55-kDa subunits of the **glial fibrillar acidic protein** but the **neurofilaments** of mammalian neurons are composed of three distinct subunits of 68-, 150-, and 200-kDa mass.[291–293] The larger subunits have C-terminal tails that are not required for filament formation but which can be phosphorylated and form bridges to neighboring neurofilaments and other cytoskeletal components and organelles. Keratin filaments, which eventually nearly fill the highly differentiated epidermal cells, are also made up of several different subunits.[294] Extensions of the keratin chains are rich in cysteine side chains which form disulfide crosslinkages to adjacent molecules to provide a network that can be dehydrated to form hair and the tough outer layers of skin.[286] Elastin-associated microfibrils are important constituents of elastic tissues of blood vessels, lungs, and skin.[295]

A common architecture of intermediate filaments is a staggered head-to-tail and side-by-side association of pairs of the coiled-coil dimers into 2- to 3-nm protofilaments and further association of about eight protofilaments to form the 10-nm intermediate filaments.[286,290,296]

2. Microfilaments

The most abundant microfilaments are composed of fibrous actin (F-actin; Fig. 7-10). The **thin filaments** of F-actin are also one of the two major components of the contractile fibers of skeletal muscle. There is actually a group of closely related actins encoded by a multigene family. At least four vertebrate actins are specific to various types of muscle, while two (β- and γ-actins) are cytosolic.[298,299] Actins are present in all animal cells and also in fungi and plants as part of the cytoskeleton. The microfilaments can associate to

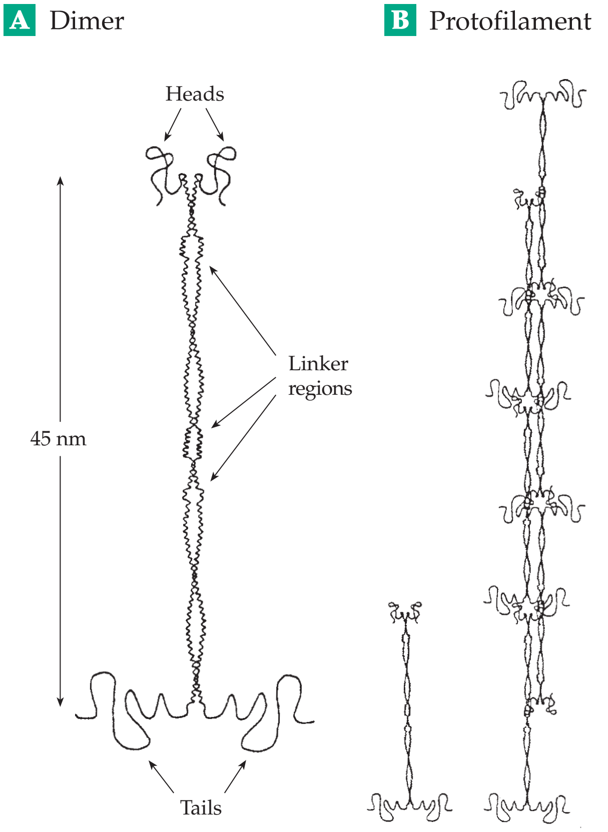


Figure 7-31 A model for the structure of keratin microfibrils of intermediate filaments. (A) A coiled-coil dimer, 45-nm in length. The helical segments of the rod domains are interrupted by three linker regions. The conformations of the head and tail domains are unknown but are thought to be flexible. (B) Probable organization of a protofilament, involving staggered antiparallel rows of dimers. From Jeffrey A. Cohlberg[297]

form larger arrays and actin often exists as thicker "cables," some of which form the **stress fibers** seen in cultured cells adhering to a glass surface. In the red blood cells the **spectrin–actin** meshwork (Fig. 8-14), which lies directly beneath the plasma membrane, together with the proteins that anchor it to the membrane, form the cytoskeleton.[284,300,301] Its mechanical properties appear to be responsible for the biconcave disc shape of the cell.

The **acrosomal process** of some invertebrate sperm cells is an actin cable that sometimes forms almost instantaneously by polymerization of the actin monomers and shoots out to penetrate the outer layers of the egg during fertilization (Chapter 32). The **stereocilia**, the "hairs" of the hair cells in the inner ear, contain bundles of actin filaments.[302] Motion of the stereocilia caused by sound produces changes in the membrane potential of the cells initiating a nerve impulse. In certain lizards each hair cell contains about 75 stereocilia of lengths up to 30 µm and diameter 0.8 µm and containing more than 3000 actin filaments in a semicrystalline array. Microvilli (Fig. 1-6) contain longitudinal arrays of actin filaments.

In every instance, groups of microfilaments are held together by other proteins. Stress fibers of higher eukaryotes contain the "muscle proteins" **tropomyosin**, **α-actinin**, and myosin (Chapter 19), although the latter is usually not in fibrillar form. **Filamin** (250-kDa) and a 235-kDa protein are associated with actin in platelets.[303] The high-molecular-weight **synemin** crosslinks vimentin and desmin filaments,[304] while the smaller, highly polar **filaggrin** provides a matrix around the keratin filaments in the external layers of the skin.[305]

Postsynthetic modifications of cytoskeletal microfilaments can also occur. For example, epidermal keratin has been found to contain lanthionine, (γ-glutamyllysine) and lysinoalanine, both presumably arising from crosslinkages.[306]

3. Microtubules

A prominent component of cytoplasm consists of microtubules which appear under the electron microscope to have a diameter of 24 ± 2 nm and a 13 - to 15-nm hollow core.[307–310] However, the true diameter of a hydrated microtubule is about 30 nm and the microtubule may be further surrounded by a 5–20 nm low density layer of associated proteins. Microtubules are present in the most striking form in the flagella and cilia of eukaryotic cells (Fig. 1-8). The **stable microtubules** of cilia are integral components of the machinery causing their motion (Chapter 19). **Labile microtubules**, which form and then disappear, are often found in cytoplasm in which motion is taking place, for example, in the pseudopodia of the ameba. The mitotic spindle consists largely of microtubules which function in the movement of chromosomes in a dividing cell (Box 7-D and Chapter 26).

Microtubules in the long axons of nerve cells function as "rails" for the "fast transport" of proteins and other materials from the cell body down the axons. In fact, microtubules appear to be present throughout the cytoplasm of virtually all eukaryotic cells (Fig. 7-32) and also in spirochetes.[311] Motion in microtubular systems depends upon motor proteins such as **kinesin**, which moves bound materials toward what is known as the "negative" end of the microtubule,[312] **dyneins** which move toward the positive end.[310] These motor proteins are driven by the Gibbs energy of hydrolysis of ATP or GTP and in this respect, as well as in some structural details (Chapter 19), resemble the muscle protein myosin. Dynein is present in the arms of the microtubules of cilia (Fig. 1-8) whose motion results from the sliding of the microtubules driven by the action of this protein (Chapter 19).

Microtubules are assembled from ~ 55-kDa **tubulins**, which are mixed dimers of α subunits (450 residues) and β subunits (445 residues) with 40% sequence identity. The αβ dimers, whose structure is shown in

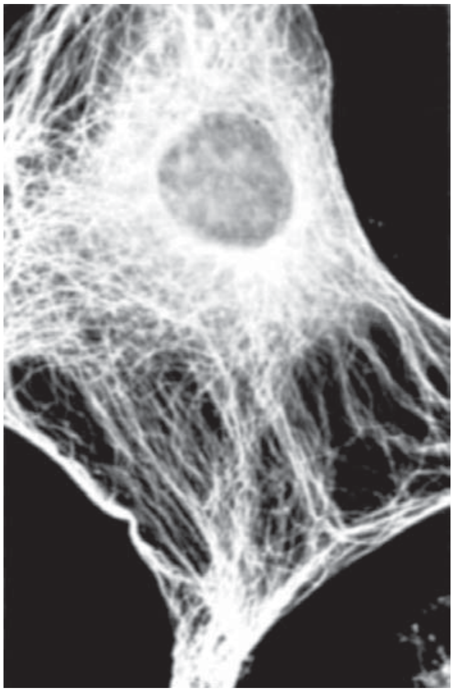
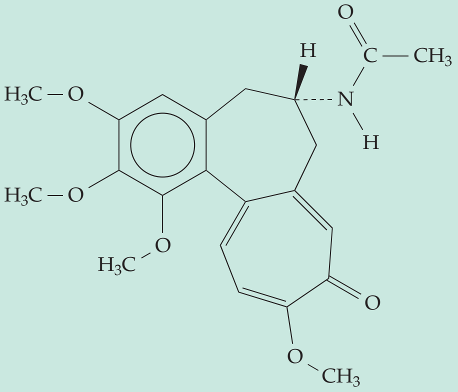


Figure 7-32 Micrograph of a mouse embryo fibroblast was obtained using indirect immunofluorescence techniques.[313] The cells were fixed with formaldehyde, dehydrated, and treated with antibodies (formed in a rabbit) to microtubule protein. The cells were then treated with fluorescent goat antibodies to rabbit γ-globulins (see Chapter 31) and the photograph was taken by fluorescent light emission. Courtesy of Klaus Weber.

BOX 7-D MITOSIS, TETRAPLOID PLANTS, AND ANTICANCER DRUGS

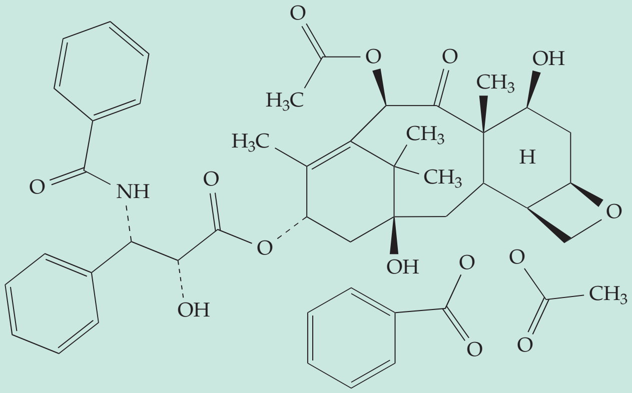
Microtubules in cells undergoing mitosis are the target of several important drugs. One of these is the alkaloid **colchicine** which is produced by various members of the lily family and has been used since ancient Egyptian times for the alleviation of the symptoms of gout.[a,b]



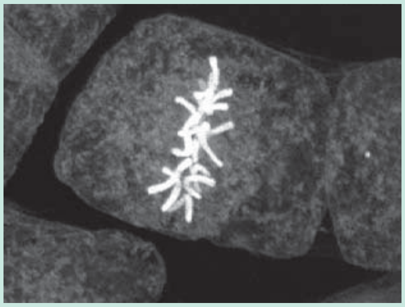
Colchicine

This compound, with its tropolone ring system, binds specifically and tightly and prevents assembly of microtubules, including those of the mitotic spindle.[b] Colchicine forms a complex with soluble tubulin,[c] perhaps a dimeric αβ complex of the two subunits.[d] Dividing cells treated with colchicine appear to be blocked at metaphase (Chapter 26) and daughter cells with a high degree of polyploidy are formed. This has led to the widespread use of colchicine in inducing formation of tetraploid varieties of flowering plants. Similar effects upon microtubules are produced by the antitumor agents **vincristine** and **vinblastine**, alkaloids formed by the common plant *Vinca* (periwinkle), and also by a variety of other drugs.[e]

The more recently discovered Taxol (paclitaxel) was extracted from the bark of the Pacific yew.[f] It *stabilizes* microtubules, inhibiting their disassembly.[g] Taxol also blocks mitosis and causes the cells which fail to complete mitosis to die. Taxol has been synthesized[f] and is a promising drug that is being used in treatment of breast, ovarian, and other cancers.[f] Binding sites for the compound have been located in β tubulin subunits (Fig. 7-33).[h,i] Attempts are being made to develop "taxoids" and other drugs more effective than taxol against cancer cells.[j,k]



Taxol (paclitaxel)



Laser scanning confocal micrograph of chromosomes at metaphase. Courtesy of Tom Moninger

Another group of drugs that bind to microtubules are **benzimidazole** and related compounds. These have been used widely to treat infection by parasitic nematodes in both humans and animals. Unfortunately resistance has developed rapidly. In a nematode that infects sheep a single tyrosine to phenylalanine mutation at position 200 in the β-tubulin subunit confers resistance.[l]

---

[a] Margulis, T. N. (1974) *J. Am. Chem. Soc.* **96**, 899–902

[b] Chakrabarti, G., Sengupta, S., and Bhattacharyya, B. (1996) *J. Biol. Chem.* **271**, 2897–2901

[c] Panda, D., Daijo, J. E., Jordan, M. A., and Wilson, L. (1995) *Biochemistry* **34**, 9921–9929

[d] Shearwin, K. E., and Timasheff, S. N. (1994) *Biochemistry* **33**, 894–901

[e] Hastie, S. B., Williams, R. C., Jr., Puett, D., and Macdonald, T. L. (1989) *J. Biol. Chem.* **264**, 6682–6688

[f] Nicolaou, K. C., Nantermet, P. G., Ueno, H., Guy, R. K., Couladouros, E. A., and Sorensen, E. J. (1995) *J. Am. Chem. Soc.* **117**, 624–633

[g] Derry, W. B., Wilson, L., and Jordan, M. A. (1995) *Biochemistry* **34**, 2203–2211

[h] Rao, S., He, L., Chakravarty, S., Ojima, I., Orr, G. A., and Horwitz, S. B. (1999) *J. Biol. Chem.* **274**, 37990–37994

[i] Makowski, L. (1995) *Nature (London)* **375**, 361–362

[j] Nicolaou, K. C., Guy, R. K., and Potier, P. (1996) *Sci. Am.* **274**(Jun), 94–98

[k] Kowalski, R. J., Giannakakou, P., and Hamel, E. (1997) *J. Biol. Chem.* **272**, 2534–2541

[l] Kwa, M. S. G., Veenstra, J. G., Van Dijk, M., and Roos, M. H. (1995) *J. Mol. Biol.* **246**, 500–510

Fig. 7-33[314–316] are thought to be packed into an imperfect helix as indicated in Fig. 7-34. The structure can also be regarded as an array of longitudinal protofilaments. Naturally formed microtubules usually have precisely 13 protofilaments and a discontinuity in the helical stacking of subunits as shown in Fig. 7-34. When grown in a laboratory the microtubules usually have 14 protofilaments[317] and rarely 10 or 16 protofilaments with regular helical packing.[318] Microtubules of some moths and also of male germ cells of *Drosophila* have 16-protofilament microtubules without a discontinuity, an architecture that is specified by the geometry of a specific β-tubulin isoform.[319]

Each tubulin dimer binds one molecule of GTP strongly in the α subunit and a second molecule of GTP or GDP more loosely in the β subunit. In this respect, tubulin resembles actin, whose subunits are about the same size. However, there is little sequence similarity. Labile microtubules of cytoplasm can be formed or disassembled very rapidly. GTP is essential for the fast growth of these microtubules and is hydrolyzed to GDP in the process.[320] However, nonhydrolyzable analogs of GTP, such as the one containing the linkage $P{-}CH_2{-}P$ between the terminal and central phosphorus atoms of the GTP, also support polymerization.[321] Since microtubules have a distinct **polarity**, the two ends have different tubulin surfaces exposed, and polymerization and depolymerization can occur at different rates at the two ends. As a consequence, microtubules often grow at one end and disassemble at the other. Such **"treadmilling"** may be important in movement of chromosomes in neuronal migration[322] and in fast axonal transport of macromolecules (Chapter 30).[323] During mitosis the minus ends of the microtubules are believed to be tightly anchored at the centrosome while subunit exchanges can occur at the plus ends[323,324] where the β subunits are exposed. Using a phage display system (see Fig. 3-16) it could be shown that the N termini of the α subunits are exposed at the minus ends.[325] Kinesin can bind to the β subunits all along the microtubule.[326] Microtubules are formed by growth from microtubule nucleation sites in **microtubule organizing centers** found in centrosomes, spindle poles, and other locations.[327] Several proteins, including **γ-tubulin**, are required.[317,328,329] A proposed assembly pathway is illustrated in Fig. 7-34.

Isolated microtubules always contain small amounts of larger ~300-kDa **microtubule-associated proteins** (MAPS).[330] These elongated molecules may in part lie in the grooves between the tubulin subunits and in part be extended outward to form a low-density layer around the tubule.[283,309] Nerve cells that contain stable microtubules have associated stabilizing proteins.[331] A family of proteins formed by differential splicing of mRNA are known as **tau**. The tau proteins are prominent components of the cytoskeleton of neurons. They not only interact with microtubules but also undergo reversible phosphorylation. Hyperphosphorylated tau is the primary component of the paired helical filaments found in the brains of persons with Alzheimer disease.[330]

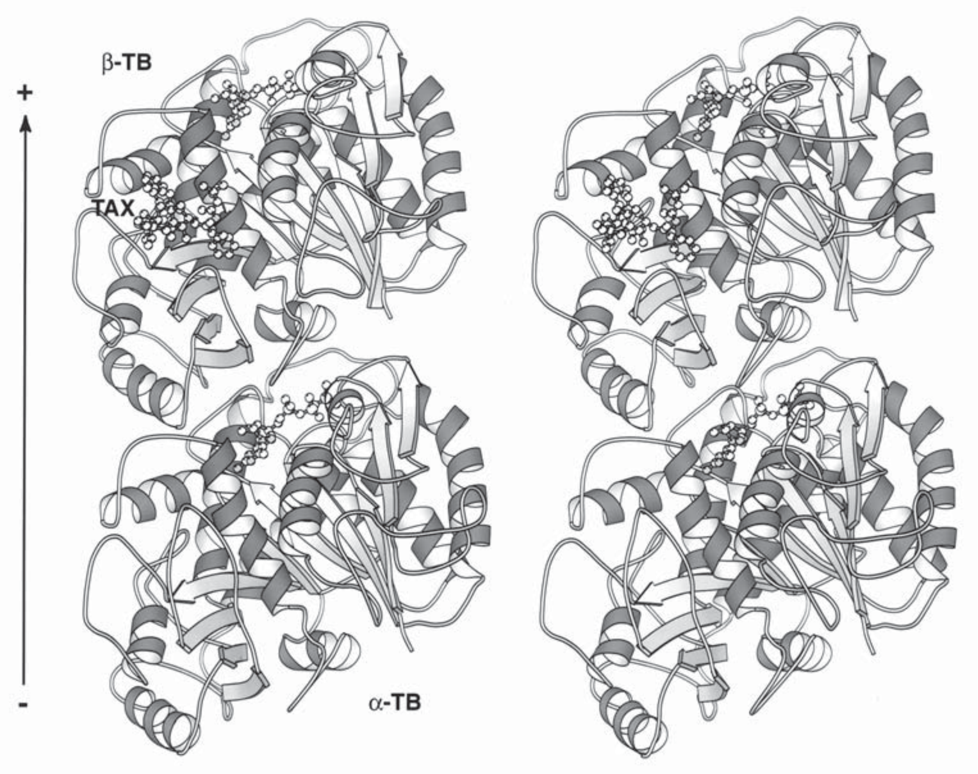


Figure 7-33 Stereoscopic ribbon diagram of the tubulin dimer with α-tubulin with bound GTP at the top and β-tubulin with bound GDP at the bottom. The β-tubulin subunit also contains a bound molecule of taxotere (see Box 7-D) which is labeled TAX. This model is based upon electron crystallography of zinc-induced tubulin sheets at 0.37-nm resolution and is thought to approximate closely the packing of the tubulin monomers in microtubules.[315] The arrow at the left points toward the plus end of the microtubule. Courtesy of Kenneth H. Downing.

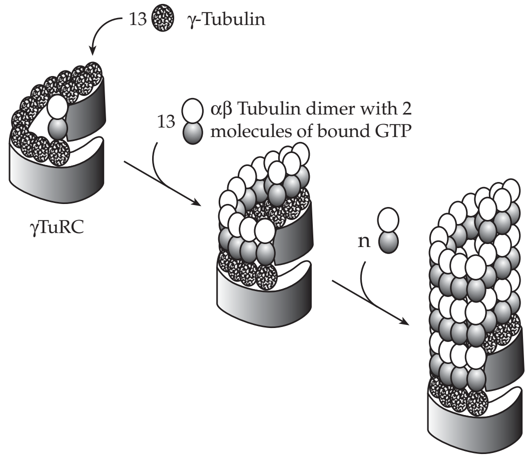


Figure 7-34 Growth of a microtubule from a γ-tubulin ring complex (γTURC). The helical γ-tubulin rings are formed in the microtubule organizing centers which, in animal cells, are the centrosomes. Thirteen γ-tubulin subunits are shown in a hypothetical array formed together with a base of other molecules of unknown structure. The microtubule grows by addition of successive layers of α/β-tubulin dimers, each a split ring of 13 dimers with the β-tubulin subunits toward the base, the **negative end**, and the α-tubulin subunits toward the growing **positive end**. After Zheng *et al*.[317]

References

1. Williams, D. H., Searle, M. S., Mackay, J. P., Gerhard, U., and Maplestone, R. A. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1172–1178
2. Weber, G. (1992) *Protein Interactions*, Chapman and Hall, London
3. Scatchard, G. (1949) *Ann. N.Y. Acad. Sci.* **51**, 660–672
4. Zierler, K. (1989) *Trends Biochem. Sci.* **14**, 314–317
5. Munson, P. J., Rodbard, D., and Klotz, I. M. (1983) *Science* **220**, 979–981
6. Deranleau, D. A. (1969) *J. Am. Chem. Soc.* **91**, 4044–4049
7. Dowd, J. E., and Riggs, D. S. (1965) *J. Biol. Chem.* **240**, 863–869
8. Feldman, H. A. (1983) *J. Biol. Chem.* **258**, 12865–12867

8a. Dunford, H. B. (1984) *J. Chem. Educ.* **61**, 129–132

8b. Klotz, I. M. (1997) *Ligand-Receptor Energetics*, Wiley, New York

9. Conners, K. A. (1987) *Binding Constants*, Wiley, New York
10. Motulsky, H. J., and Ransnas, L. A. (1987) *FASEB J.* **1**, 365–374
11. Schejter, A., and Margoliash, E. (1985) *Trends Biochem. Sci.* **10**, 490–492
12. Steinhardt, J., and Reynolds, J. A. (1969) *Multiple Equilibria in Proteins*, Academic Press, New York
13. Koshland, D. E., Jr. (1970) in *The Enzymes*, 3rd ed., Vol. 1 (Boyer, P. D., ed), pp. 341–396, Academic Press, New York
14. Tanford, C. (1961) *Physical Chemistry of Macromolecules*, Wiley, New York (pp. 533–534)
15. Edsall, J. T., and Wyman, J. (1958) *Biophysical Chemistry*, Academic Press, New York
16. Westheimer, F. H., and Shookhoff, M. W. (1939) *J. Am. Chem. Soc.* **61**, 555–560
17. Honig, B., and Nicholls, A. (1995) *Science* **268**, 1144–1149
18. Rajasekaran, E., Jayaram, B., and Honig, B. (1994) *J. Am. Chem. Soc.* **116**, 8238–8240
19. Tanford, C. (1961) *Physical Chemistry of Macromolecules*, Wiley, New York (Chapter 8)
20. Maier, G. D., and Metzler, D. E. (1957) *J. Am. Chem. Soc.* **79**, 4386–4391
21. Metzler, D. E. (1960) in *The Enzymes*, 2nd ed., Vol. 2 (Boyer, P. D., Lardy, H., and Myrback, K., eds), pp. 295–337, Academic Press, New York
22. Applequist, J. (1977) *J. Chem. Educ.* **54**, 417–419
23. Hill, T. (1962) *Introduction to Statistical Thermodynamics*, Addison–Wesley, Reading, Massachusetts (pp. 235–241)
24. Weiss, J. N. (1997) *FASEB J.* **11**, 835–841
25. Wang, X.-G., and Engel, P. C. (1995) *Biochemistry* **34**, 11417–11422
26. Damle, V. N. (1972) *Biopolymers* **11**, 1789–816
27. Watson, J. D. (1976) *Molecular Biology of the Gene*, 3rd ed., Benjamin, Menlo Park, California (p. 98)
28. Bernal, J. D. (1967) *J. Mol. Biol.* **24**, 379–390
29. Blundell, T. L., and Johnson, L. N. (1976) *Protein Crystallogrophy*, Academic Press, New York
30. Monod, J., Wyman, J., and Changeux, J. D. (1965) *J. Mol. Biol.* **12**, 88–118

30a. Dill, K. A. (1997) *J. Biol. Chem.* **272**, 701–704

30b. Boresch, S., and Karplus, M. (1995) *J. Mol. Biol.* **254**, 801–807

30c. Tamura, A., and Privalov, P. L. (1997) *J. Mol. Biol.* **273**, 1048–1060

31. Zhang, R.-G., Westbrook, M. L., Westbrook, E. M., Scott, D. L., Otwinowski, Z., Maulik, P. R., Reed, R. A., and Shipley, G. G. (1995) *J. Mol. Biol.* **251**, 550–562

31a. Marvin, D. A. (1998) *Current Opinion in Structural Biology* **8**, 150–158

32. Marvin, D. A., Hale, R. D., Nave, C., and Citterich, M. H. (1994) *J. Mol. Biol.* **235**, 260–286
33. Glucksman, M. J., Bhattacharjee, S., and Makowski, L. (1992) *J. Mol. Biol.* **226**, 455–470
34. Day, L. A., Marzec, C. J., Reisberg, S. A., and Casadevall, A. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 509–539
35. Williams, K. A., Glibowicka, M., Li, Z., Li, H., Khan, A. R., Chen, Y. M. Y., Wang, J., Marvin, D. A., and Deber, C. M. (1995) *J. Mol. Biol.* **252**, 6–14
36. Overman, S. A., Tsuboi, M., and Thomas, G. J., Jr. (1996) *J. Mol. Biol.* **259**, 331–336
37. Makowski, L. (1992) *J. Mol. Biol.* **228**, 885–892
38. Rakonjac, J., and Model, P. (1998) *J. Mol. Biol.* **282**, 25–41

38a. Welsh, L. C., Marvin, D. A., and Perham, R. N. (1998) *J. Mol. Biol.* **284**, 1265–1271

39. Liu, D. J., and Day, L. A. (1994) *Science* **265**, 671–674
40. Bhyravbhatla, B., Watowich, S. J., and Caspar, D. L. D. (1998) *Biophys. J.* **74**, 604–615
41. Fraenkel-Conrat, H. (1964) *Sci. Am.* **211**(Oct), 47–54
42. Butler, P. J. G., and Klug, A. (1978) *Sci. Am.* **239**(Nov), 62–69
43. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R., and Klug, A. (1978) *Nature (London)* **276**, 362–368
44. Holmes, K. C. (1980) *Trends Biochem. Sci.* **5**, 4–7

44a. Eisenstein, M., Shariv, I., Koren, G., Friesem, A. A., and Katchalski-Katzir, E. (1997) *J. Mol. Biol.* **266**, 135–143

45. Raghavendra, K., Kelly, J. A., Khairallah, L., and Schuster, T. M. (1988) *Biochemistry* **27**, 7583–7588
46. Wang, H., and Stubbs, G. (1994) *J. Mol. Biol.* **239**, 371–384
47. Namba, K., Casper, D. L. D., and Stubbs, G. J. (1985) *Science* **227**, 773–776
48. Eisenstein, B. I. (1987) Escherichia coli *and* Salmonella typhimurium, Am. Soc. Microbiology, Washington, D.C., FC Niedhardt, ed. (pp. 84–90)
49. Bullitt, E., and Makowski, L. (1995) *Nature (London)* **373**, 164–167
50. Bullitt, E., Jones, C. H., Striker, R., Soto, G., Jacob-Dubuisson, F., Pinkner, J., Wick, M. J., Makowski, L., and Hultgren, S. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12890–12895
51. Bullitt, E., and Makowski, L. (1998) *Biophys. J.* **74**, 623–632
52. Gong, M., and Makowski, L. (1992) *J. Mol. Biol.* **228**, 735–742
53. Kuehn, M. J., Ogg, D. J., Kihlberg, J., Slonim, L. N., Flemmer, K., Bergfors, T., and Hultgren, S. J. (1993) *Science* **262**, 1234–1241

53a. Sauer, F. G., Fütterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J., and Waksman, G. (1999) *Science* **285**, 1058–1061

53b. Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J., and Knight, S. D. (1999) *Science* **285**, 1061–1065

53c. Skerker, J. M., and Shapiro, L. (2000) *EMBO J.* **19**, 3223–3234

53d. Hazes, B., Sastry, P. A., Hayakawa, K., Read, R. J., and Irvin, R. T. (2000) *J. Mol. Biol.* **299**, 1005–1017

53e. Collinson, S. K., Parker, J. M. R., Hodges, R. S., and Kay, W. W. (1999) *J. Mol. Biol.* **290**, 741–756

53f. Murphy, F. V., IV, Sweet, R. M., and Churchill, M. E. A. (1999) *EMBO J.* **18**, 6610–6618

54. Hanson, M. S., and Brinton, C. C., Jr. (1988) *Nature (London)* **332**, 265–268
55. St. Geme, J. W., III, Pinkner, J. S., Krasan, G. P., Heuser, J., and Bullitt, E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11913–11918
56. Parge, H. E., Bernstein, S. L., Deal, C. D., McRee, D. E., Christensen, D., Capozza, M. A., Kays, B. W., Fieser, T. M., Draper, D., So, M., Getzoff, E. D., and Tainer, J. A. (1990) *J. Biol. Chem.* **265**, 2278–2285

References

57. Parge, H. E., Forest, K. T., Hickey, M. J., Christensen, D. A., Getzoff, E. D., and Tainer, J. A. (1995) *Nature (London)* **378**, 32–38
58. Folkhard, W., Leonard, K. R., Malsey, S., Marvin, D. A., Dubochet, J., Engel, A., Achtman, M., and Helmuth, R. (1979) *J. Mol. Biol.* **130**, 145–160
59. Paiva, W. D., Grossman, T., and Silverman, P. M. (1992) *J. Biol. Chem.* **267**, 26191–26197
60. Rubenstein, P. A. (1990) *BioEssays* **12**, 309–315
61. Venkatesh, B., Tay, B. H., Elgar, G., and Brenner, S. (1996) *J. Mol. Biol.* **259**, 655–665
62. Holmes, K. C., Popp, D., Gebhard, W., and Kabsch, W. (1990) *Nature (London)* **347**, 44–49
63. Lorenz, M., Poole, K. J. V., Popp, D., Rosenbaum, G., and Holmes, K. C. (1995) *J. Mol. Biol.* **246**, 108–119
64. Valentine, R. C. (1969) in *Symmetry and Function of the Biological Systems at the Macromolecular Level* (Engstrom, A., and Strandberg, B., eds), p. 165, Wiley, New York (11th Nobel Symp.)
65. Klotz, I. M., Darnall, D. W., and Langerman, N. R. (1975) in *The Proteins*, 3rd ed., Vol. 1 (Neurath, H., and Hill, R. L., eds), p. 293, Academic Press, New York
66. Momany, C., Ernst, S., Ghosh, R., Chang, N.-L., and Hackert, M. L. (1995) *J. Mol. Biol.* **252**, 643–655
67. Almassy, R. J., Janson, C. A., Hamlin, R., Xuong, N.-H., and Eisenberg, D. (1986) *Nature (London)* **323**, 304–309
68. Harrison, S. C. (1984) *Trends Biochem. Sci.* **9**, 345–351
69. Rossmann, M. G., and Johnson, J. E. (1989) *Ann. Rev. Biochem.* **58**, 533–573
70. Johnson, J. E., and Speir, J. A. (1997) *J. Mol. Biol.* **269**, 665–675
71. Watson, J. D. (1976) *Molecular Biology of the Gene*, 3rd ed., Benjamin, Menlo Park, California (p. 107)
72. Unge, T., Liljas, L., Strandberg, B., Vaara, I., Kannan, K. K., Fridborg, K., Nordman, C. E., and Lentz, P. J., Jr. (1980) *Nature (London)* **285**, 373–377
73. Küpper, H., Keller, W., Kurz, C., Forss, S., Schaller, H., Franze, R., Stohmaier, K., Marquardt, O., Zaslavsky, V. G., and Hofschneider, P. H. (1981) *Nature (London)* **289**, 555–559
74. Trikha, J., Theil, E. C., and Allewell, N. M. (1995) *J. Mol. Biol.* **248**, 949–967
75. Larson, S. B., Day, J., Greenwood, A., and McPherson, A. (1998) *J. Mol. Biol.* **277**, 37–59
76. Hogle, J. M., Chow, M., and Filman, D. J. (1985) *Science* **229**, 1358–1365
77. Hogle, J. M., Chow, M., and Filman, D. J. (1987) *Sci. Am.* **256** (Mar), 42–49
78. Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S., and Rossmann, M. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 507–511
79. Hadfield, A. T., Oliveira, M. A., Kim, K. H., Minor, I., Kremer, M. J., Heinz, B. A., Shepard, D., Pevear, D. C., Rueckert, R. R., and Rossmann, M. G. (1995) *J. Mol. Biol.* **253**, 61–73
80. Wu, H., and Rossmann, M. G. (1993) *J. Mol. Biol.* **233**, 231–244
81. Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossmann, M. G., Boege, U., Scraba, D. G., Duke, G. M., and Palmenberg, A. C. (1987) *Science* **235**, 182–191
82. Tikhonenko, A. S. (1970) *Ultrastructure of Bacterial Viruses*, Plenum, New York
83. Birktoft, J. J., Rhodes, G., and Banaszak, L. J. (1989) *Biochemistry* **28**, 6065–6081
84. Skarzynski, T., Moody, P. C. E., and Wonacott, A. J. (1987) *J. Mol. Biol.* **193**, 171–187
85. Zimmerle, C. T., and Alter, G. M. (1993) *Biochemistry* **32**, 12743–12748
86. Hodgkin, D. C. (1975) *Nature (London)* **255**, 103
87. Whittingham, J. L., Chaudhuri, S., Dodson, E. J., Moody, P. C. E., and Dodson, G. G. (1995) *Biochemistry* **34**, 15553–15563
88. Chang, X., Jorgensen, A. M. M., Bardrum, P., and Led, J. J. (1997) *Biochemistry* **36**, 9409–9422
89. Golmohammadi, R., Valegård, K., Fridborg, K., and Liljas, L. (1993) *J. Mol. Biol.* **234**, 620–639
90. Liljas, L., Fridborg, K., Valegård, K., Bundule, M., and Pumpens, P. (1994) *J. Mol. Biol.* **244**, 279–290
91. Tars, K., Bundule, M., Fridborg, K., and Lijas, L. (1997) *J. Mol. Biol.* **271**, 759–773
92. Abad-Zapatero, C., Abdel-Meguid, S. S., Johnson, J. E., Leslie, A. G. W., Rayment, I., Rossmann, M. G., Suck, D., and Tsukihara, T. (1980) *Nature (London)* **286**, 33–39
93. Athappilly, F. K., Murali, R., Rux, J. J., Cai, Z., and Burnett, R. M. (1994) *J. Mol. Biol.* **242**, 430–455
94. Stewart, P. L., Chiu, C. Y., Huang, S., Muir, T., Zhao, Y., Chait, B., Mathias, P., and Nemerow, G. R. (1997) *EMBO J.* **16**, 1189–1198
95. Caspar, D. L. D., and Klug, A. (1962) *Cold Spring Harbor Symposia on Quantitative Biology* **27**, 1–24
96. Johnson, J. E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 27–33
97. Steven, A. C., Trus, B. L., Booy, F. P., Cheng, N., Zlotnick, A., Caston, J. R., and Conway, J. F. (1997) *FASEB J.* **11**, 733–742
98. Rayment, I., Baker, T. S., Caspar, D. L. D., and Murakami, W. T. (1982) *Nature (London)* **295**, 110–115
99. Liddington, R. C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T. L., and Harrison, S. C. (1991) *Nature (London)* **354**, 278–284
100. Champness, J. N., Bloomer, A. C., Bricogne, G., Butler, P. J. G., and Klug, A. (1976) *Nature (London)* **259**, 20–24
101. Fukuda, M., and Okada, Y. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4035–4038
102. McDonald, R. C., Engelman, D. M., and Steita, T. A. (1979) *J. Biol. Chem.* **254**, 2942–2943
103. Galloway, J. (1984) *Trends Biochem. Sci.* **9**, 233–238
104. Kantrowitz, E. R., and Lipscomb, W. N. (1990) *Trends Biochem. Sci.* **15**, 53–59
105. Zhou, B.-B., and Schachman, H. K. (1993) *Protein Sci.* **2**, 103–112
106. Xi, X. G., De Staercke, C., Van Vliet, F., Triniolles, F., Jacobs, A., Stas, P. P., Ladjimi, M. M., Simon, V., Cunin, R., and Hervé, G. (1994) *J. Mol. Biol.* **242**, 139–149
107. Fetler, L., Vachette, P., Hervé, G., and Ladjimi, M. M. (1995) *Biochemistry* **34**, 15654–15660
108. Honzatko, R. B., Crawford, J. L., Monaco, H. L., Ladner, J. E., Edwards, B. F. D., Evans, D. R., Warren, S. G., Wiley, D. C., Ladner, R. C., and Lipscomb, W. N. (1982) *J. Mol. Biol.* **160**, 219–263
109. Krause, K. L., Volz, K. W., and Lipscomb, W. N. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1643–1647
110. Thomas, A., Field, M. J., and Perahia, D. (1996) *J. Mol. Biol.* **261**, 490–506

110a. Brunori, M. (1999) *Trends Biochem. Sci.* **24**, 158–161

110b. Di Cera, E., ed. (1998) *Advances in Protein Chemistry, Linkage Thermodynamics of Macromolecular Interactions*, Vol. 51, Academic Press, San Diego, California

111. Bloom, C. R., Choi, W. E., Brzovic, P. S., Ha, J. J., Huang, S.-T., Kaarsholm, N. C., and Dunn, M. F. (1995) *J. Mol. Biol.* **245**, 324–330
112. Hammes, G. G., and Wu, C.-W. (1971) *Science* **172**, 1205–1211
113. Koshland, D. E., Jr., Némethy, G., and Filmer, D. (1966) *Biochemistry* **5**, 365–385
114. Cornish-Bowden, A. J., and Koshland, D. E., Jr. (1970) *J. Biol. Chem.* **245**, 6241–6250
115. Cornish-Bowden, A. J., and Koshland, D. E., Jr. (1971) *J. Biol. Chem.* **246**, 3092–3102
116. Darnall, D. W., and Klotz, I. M. (1972) *Arch. Biochem. Biophys.* **149**, 1–14
117. Doci, Y., Sugita, Y., and Yoneyama, Y. (1973) *J. Biol. Chem.* **248**, 2354–2363
118. Perutz, M. F., Muirhead, H., Cox, J. M., and Goaman, L. C. G. (1968) *Nature (London)* **219**, 131–139
119. Dickerson, R. E., and Geis, I. (1983) *Hemoglobin: Structure, Function, Evolution and Pathology*, Benjamin / Cummings Publ., Redwood City, California
120. Antonini, E., and Brunori, M. (1971) *Hemoglobin and Myoglobin in Their Reactions with Ligands*, North-Holland, Publ., Amsterdam
121. Kendrew, J. C. (1961) *Sci. Am.* **205**(Dec), 96–110
122. Fermi, G., Perutz, M. F., Shaanan, B., and Fourme, R. (1984) *J. Mol. Biol.* **175**, 159–174
123. Benesch, R., and Benesch, R. E. (1974) *Science* **185**, 905–908
124. Borgstahl, G. E. O., Rogers, P. H., and Arnone, A. (1994) *J. Mol. Biol.* **236**, 831–843
125. Imai, K. (1979) *J. Mol. Biol.* **133**, 233–247
126. Amiconi, G., Antonini, E., Brunori, M., Wyman, J., and Zolla, L. (1981) *J. Mol. Biol.* **152**, 111–129
127. Flanagan, M. A., Ackers, G. K., Matthew, J. B., Hanania, G. I. H., and Gurd, F. R. N. (1981) *Biochemistry* **20**, 7439–7449

127a. Shibayama, N. (1999) *J. Mol. Biol.* **285**, 1383–1388

128. Waxman, L. (1971) *J. Biol. Chem.* **246**, 7318–7327
129. Doyle, M. L., Di Cera, E., and Gill, S. J. (1988) *Biochemistry* **27**, 820–824
130. Edelstein, S. J. (1996) *J. Mol. Biol.* **257**, 737–744
131. Imai, K. (1983) *J. Mol. Biol.* **167**, 741–749

131a. Gibson, Q. H. (1999) *Biochemistry* **38**, 5191–5199

132. Paoli, M., Liddington, R., Tame, J., Wilkinson, A., and Dodson, G. (1996) *J. Mol. Biol.* **256**, 775–792
133. Henry, E. R., Jones, C. M., Hofrichter, J., and Eaton, W. A. (1997) *Biochemistry* **36**, 6511–6528
134. Silva, M. M., Rogers, P. H., and Arnone, A. (1992) *J. Biol. Chem.* **267**, 17248–17256
135. Jayaraman, V., and Spiro, T. G. (1995) *Biochemistry* **34**, 4511–4515
136. Schumacher, M. A., Zheleznova, E. E., Poundstone, K. S., Kluger, R., Jones, R. T., and Brennan, R. G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7841–7844

136a. Tame, J. R. H. (1999) *Trends Biochem. Sci.* **24**, 372–377

137. Manning, L. R., Jenkins, W. T., Hess, J. R., Vandegriff, K., Winslow, R. M., and Manning, J. M. (1996) *Protein Sci.* **5**, 775–781
138. Holt, J. M., and Ackers, G. K. (1995) *FASEB J.* **9**, 210–218
139. Perutz, M. F., Shih, D. T.-b, and Williamson, D. (1994) *J. Mol. Biol.* **239**, 555–560
140. Perutz, M. F., Muirhead, H., Cox, J. M., Goaman, L. C. G., Mathews, F. S., McGandy, L. E., and Webb, L. E. (1968) *Nature (London)* **219**, 29–32
141. Perutz, M. (1990) *Mechanisms of Cooperativity and Allosteric Regulation in Proteins*, Cambridge Univ. Press, London
142. O'Donnell, S., Mandaro, R., Schuster, T. M., and Arnone, A. (1979) *J. Biol. Chem.* **254**, 12204–12208
143. Perutz, M. F. (1970) *Nature (London)* **228**, 726–739
144. Dickerson, R. E. (1972) *Ann. Rev. Biochem.* **41**, 815–842

145. Ishimori, K., Hashimoto, M., Imai, K., Fushitani, K., Miyazaki, G., Morimoto, H., Wada, Y., and Morishima, I. (1994) *Biochemistry* **33**, 2546–2553
146. Chao, H., Sönnichsen, F. D., DeLuca, C. I., Sykes, B. D., and Davies, P. L. (1994) *Protein Sci.* **3**, 1760–1769
147. Kim, H.-W., Shen, T.-J., Ho, N. T., Zou, M., Tam, M. F., and Ho, C. (1996) *Biochemistry* **35**, 6620–6627
148. Kavanaugh, J. S., Weydert, J. A., Rogers, P. H., and Arnone, A. (1998) *Biochemistry* **37**, 4358–4373
149. Kiger, L., Klinger, A. L., Kwiatkowski, L. D., De Young, A., Doyle, M. L., Holt, J. M., Noble, R. W., and Ackers, G. K. (1998) *Biochemistry* **37**, 4336–4345
150. Hoard, J. L., and Scheidt, W. R. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3919–3922
151. Fermi, G., Perutz, M. F., and Shulman, R. G. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6167–6168
152. De Baere, I., Perutz, M. F., Kiger, L., Marden, M. C., and Poyart, C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1594–1597
153. Bohr, C., Hasselbalch, K., and Krogh, A. (1904) *Skand. Arch. Physiol.* **16**, 402–
154. Wyman, J., Jr. (1948) *Adv. Prot. Chem.* **4**, 407–531
155. Benesch, R. E., and Benesch, R. (1974) *Adv. Prot. Chem.* **28**, 211–237
156. Fang, T.-Y., Zou, M., Simplaceanu, V., Ho, N. T., and Ho, C. (1999) *Biochemistry* **38**, 13423–13432
157. Fronticelli, C., Pechik, I., Brinigar, W. S., Kowalczyk, J., and Gilliland, G. L. (1994) *J. Biol. Chem.* **269**, 23965–23969
158. Perrella, M., Ripamonti, M., and Caccia, S. (1998) *Biochemistry* **37**, 2017–2028
159. Arnone, A. (1972) *Nature (London)* **237**, 146–149
160. Benesch, R., and Benesch, R. E. (1969) *Nature (London)* **221**, 618–622
161. Brewer, G. J., and Eaton, J. W. (1971) *Science* **171**, 1205–1211
162. Perutz, M. F., Bauer, C., Gros, G., Leclercq, F., Vandecasserie, C., Schnek, A. G., Braunitzer, G., Friday, A. E., and Joysey, K. A. (1981) *Nature (London)* **291**, 682–684
163. Komiyama, N. H., Miyazaki, G., Tame, J., and Nagai, K. (1995) *Nature (London)* **373**, 244–246
164. Fantl, W. J., Di Donato, A., Manning, J. M., Rogers, P. H., and Arnone, A. (1987) *J. Biol. Chem.* **87**, 12700–12713
165. Sage, J. T., and Jee, W. (1997) *J. Mol. Biol.* **274**, 21–26
166. Teng, T.-Y., Srajer, V., and Moffat, K. (1997) *Biochemistry* **36**, 12087–12100
166a. Ostermann, A., Waschipky, R., Parak, F. G., and Nienhaus, G. U. (2000) *Nature (London)* **404**, 205–208
167. Chance, M. R., Miller, L. M., Fischetti, R. F., Scheuring, E., Huang, W.-X., Sclavi, B., Hai, Y., and Sullivan, M. (1996) *Biochemistry* **35**, 9014–9023
168. Ösapay, K., Theriault, Y., Wright, P. E., and Case, D. A. (1994) *J. Mol. Biol.* **244**, 183–197
169. Martin, K. D., and Parkhurst, L. J. (1990) *Biochemistry* **29**, 5718–5726
170. Tetreau, C., Lavalette, D., Momenteau, M., Fischer, J., and Weiss, R. (1994) *J. Am. Chem. Soc.* **116**, 11840–11848
171. Huang, Y., and Ackers, G. K. (1995) *Biochemistry* **34**, 6316–6327
172. Paoli, M., Dodson, G., Liddington, R. C., and Wilkinson, A. J. (1997) *J. Mol. Biol.* **271**, 161–167
173. Daugherty, M. A., Shea, M. A., and Ackers, G. K. (1994) *Biochemistry* **33**, 10345–10357
174. Stamler, J. S., Jia, L., Eu, J. P., McMahon, T. J., Demchenko, I. T., Bonaventura, J., Gernert, K., and Piantadosi, C. A. (1997) *Science* **276**, 2034–2037
175. Upmacis, R. K., Hajjar, D. P., Chait, B. T., and Mirza, U. A. (1997) *J. Am. Chem. Soc.* **119**, 10424–10429
176. Miller, L. M., Pedraza, A. J., and Chance, M. R. (1997) *Biochemistry* **36**, 12199–12207
177. Gow, A. J., and Stamler, J. S. (1998) *Nature (London)* **391**, 169–173
178. Lutter, L. C., Halvorson, H. R., and Calladine, C. R. (1996) *J. Mol. Biol.* **261**, 620–633
179. Jia, L., Bonaventura, C., Bonaventura, J., and Stamler, J. S. (1996) *Nature (London)* **380**, 221–226
180. Rizzi, M., Wittenberg, J. B., Coda, A., Fasano, M., Ascenzi, P., and Bolognesi, M. (1994) *J. Mol. Biol.* **244**, 86–99
181. Rizzi, M., Wittenberg, J. B., Coda, A., Ascenzi, P., and Bolognesi, M. (1996) *J. Mol. Biol.* **258**, 1–5
182. Nguyen, B. D., Zhao, X., Vyas, K., La Mar, G. N., Lile, R. A., Brucker, E. A., Phillips, G. N., Jr., Olson, J. S., and Wittenberg, J. B. (1998) *J. Biol. Chem.* **273**, 9517–9526
183. Weatherall, D. J., Clegg, J. B., Higgs, D. R., and Wood, W. G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3417–3484, McGraw-Hill, New York
184. Bank, A., Mears, J. G., and Ramirez, F. (1980) *Science* **207**, 486–493
185. Hirsch, R. E., Juszczak, L. J., Fataliev, N. A., Friedman, J. M., and Nagel, R. L. (1999) *J. Biol. Chem.* **274**, 13777–13782
186. Arnone, A., Thillet, J., and Rosa, J. (1981) *J. Biol. Chem.* **256**, 8545–8552
187. Lehmann, H., and Huntsman, R. G. (1966) *Man's Haemoglobins*, North-Holland Publ. Co., Amsterdam (p. 246)
188. Olson, J. S., Gibson, Q. H., Nagel, R. L., and Hamilton, H. B. (1972) *J. Biol. Chem.* **247**, 7485–7493
189. Poyart, C., Bursaux, E., Arnone, A., Bonaventura, J., and Bonaventura, C. (1980) *J. Biol. Chem.* **255**, 9465–9473
189a. Burmester, T., Weich, B., Reinhardt, S., and Hankeln, T. (2000) *Nature* **407**, 520–523
190. Tomita, S. (1981) *J. Biol. Chem.* **256**, 9495–9500
191. Giardina, B., Scatena, R., Clementi, M. E., Cerroni, L., Nuutinen, M., Brix, O., Sletten, S. N., Castagnola, M., and Condò, S. G. (1993) *J. Mol. Biol.* **229**, 512–516
192. Clementi, M. E., Scatena, R., Mordente, A., Condò, S. G., Castagnola, M., and Giardina, B. (1996) *J. Mol. Biol.* **255**, 229–234
192a. Burr, A. H. J., Hunt, P., Wagar, D. R., Dewilde, S., Blaxter, M. L., Vanfleteren, J. R., and Moens, L. (2000) *J. Biol. Chem.* **275**, 4810–4815
193. Lee, H. C., Wittenberg, J. B., and Peisach, J. (1993) *Biochemistry* **32**, 11500–11506
194. Harutyunyan, E. H., Safonova, T. N., Kuranova, I. P., Popov, A. N., Teplyakov, A. V., Obmolova, G. V., Rusakov, A. A., Vainshtein, B. K., Dodson, G. G., Wilson, J. C., and Perutz, M. F. (1995) *J. Mol. Biol.* **251**, 104–115
195. Hargrove, M. S., Barry, J. K., Brucker, E. A., Berry, M. B., Phillips, G. N., Jr., Olson, J. S., Arredondo-Peter, R., Dean, J. M., Klucas, R. V., and Sarath, G. (1997) *J. Mol. Biol.* **266**, 1032–1042
196. Trevaskis, B., Watts, R. A., Andersson, C. R., Llewellyn, D. J., Hargrove, M. S., Olson, J. S., Dennis, E. S., and Peacock, W. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12230–12234
196a. Hargrove, M. S., Brucker, E. A., Stec, B., Sarath, G., Arredondo-Peter, R., Klucas, R. V., Olson, J. S., and Phillips, G. N., Jr. (2000) *Structure* **8**, 1005–1014
197. Hardison, R. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5675–5679
198. Potts, M., Angeloni, S. V., Ebel, R. E., and Bassam, D. (1992) *Science* **256**, 1690–1692
199. Yeh, S.-R., Couture, M., Ouellet, Y., Guertin, M., and Rousseau, D. L. (2000) *J. Biol. Chem.* **275**, 1679–1684
200. Aronson, H.-E. G., Royer, W. E., Jr., and Hendrickson, W. A. (1994) *Protein Sci.* **3**, 1706–1711
201. Honzatko, R. B., Hendrickson, W. A., and Love, W. E. (1985) *J. Mol. Biol.* **184**, 147–164
201a. Qiu, Y., Maillett, D. H., Knapp, J., Olson, J. S., and Riggs, A. F. (2000) *J. Biol. Chem.* **275**, 13517–13528
202. Condon, P. J., and Royer, W. E. J. (1994) *J. Biol. Chem.* **269**, 25259–25267
203. Royer, W. E., Jr., Heard, K. S., Harrington, D. J., and Chiancone, E. (1995) *J. Mol. Biol.* **253**, 168–186
204. Mozzarelli, A., Bettati, S., Rivetti, C., Rossi, G. L., Colotti, G., and Chiancone, E. (1996) *J. Biol. Chem.* **271**, 3627–3632
205. Yang, J., Kloek, A. P., Goldberg, D. E., and Mathews, F. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4224–4228
206. Huang, S., Huang, J., Kloek, A. P., Goldberg, D. E., and Friedman, J. M. (1996) *J. Biol. Chem.* **271**, 958–962
207. Lamy, J., Kuchumov, A., Taveau, J.-C., Vinogradov, S. N., and Lamy, J. N. (2000) *J. Mol. Biol.* **298**, 633–647
208. de Haas, F., Taveau, J.-C., Boisset, N., Lambert, O., Vinogradov, S. N., and Lamy, J. N. (1996) *J. Mol. Biol.* **255**, 140–153
209. de Haas, F., Zal, F., You, V., Lallier, F., Toulmond, A., and Lamy, J. N. (1996) *J. Mol. Biol.* **264**, 111–120
210. de Haas, F., Boisset, N., Taveau, J.-C., Lambert, O., Vinogradov, S. N., and Lamy, J. N. (1996) *Biophys. J.* **70**, 1973–1984
211. Zuckerkandl, E. (1965) *Sci. Am.* **212**(May), 110–118
212. Perutz, M. F. (1986) *Nature (London)* **322**, 405
213. Wood, W. B., and Edgar, R. S. (1967) *Sci. Am.* **217**(Jul), 60–74
214. Kikuchi, Y., and King, J. (1975) *J. Mol. Biol.* **99**, 645–672
215. Casjens, S., and King, J. (1975) *Ann. Rev. Biochem.* **44**, 555–611
216. Pollard, T. D., and Craig, S. W. (1982) *Trends Biochem. Sci.* **7**, 55–58
217. Ito, K., Date, T., and Wickner, W. (1980) *J. Biol. Chem.* **255**, 2123–2130
218. McDonnell, P. A., Shon, K., Kim, Y., and Opella, S. J. (1993) *J. Mol. Biol.* **233**, 447–463
219. Sanders, J. C., Haris, P. I., Chapman, D., Otto, C., and Hemminga, M. A. (1993) *Biochemistry* **32**, 12446–12454
220. Guy-Caffey, J. K., and Webster, R. E. (1993) *J. Biol. Chem.* **268**, 5496–5503
221. Duda, R. L., Martincic, K., and Hendrix, R. W. (1995) *J. Mol. Biol.* **247**, 636–647
222. Xie, Z., and Hendrix, R. W. (1995) *J. Mol. Biol.* **253**, 74–85
223. Marvik, C. J., Dokland, T., Nøkling, R. H., Jacobsen, E., Larsen, T., and Lindqvist, B. H. (1995) *J. Mol. Biol.* **251**, 59–75
224. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1995) *Biochemistry* **34**, 10600–10606
225. McKenna, R., Ilag, L. L., and Rossmann, M. G. (1994) *J. Mol. Biol.* **237**, 517–543
226. Dokland, T., McKenna, R., Ilag, L. L., Bowman, B. R., Incardona, N. L., Fane, B. A., and Rossmann, M. G. (1997) *Nature (London)* **389**, 308–313
227. Thuman-Commike, P. A., Greene, B., Jakana, J., Prasad, B. V. V., King, J., Prevelige, P. E., Jr., and Chiu, W. (1996) *J. Mol. Biol.* **260**, 85–98
228. Parker, M. H., Stafford, W. F., III, and Prevelige, P. E., Jr. (1997) *J. Mol. Biol.* **268**, 655–665

References

229. Butcher, S. J., Dokland, T., Ojala, P. M., Bamford, D. H., and Fuller, S. D. (1997) *EMBO J.* **16**, 4477–4487
230. Olson, N. H., Baker, T. S., Willingmann, P., and Incardona, N. L. (1992) *J. Structural Biol.* **108**, 168–175
231. Hayashi, M., Aoyama, A., Richardson, D. L., Jr., and Hayashi, M. N. (1988) in *The Bacteriophages* (Calendar, R., ed), pp. 1–71, Plenum, New York
232. Butcher, S. J., Bamford, D. H., and Fuller, S. D. (1995) *EMBO J.* **14**, 6078–6086
233. Berger, B., Shor, P. W., Tucker-Kellogg, L., and King, J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7732–7736
234. Steven, A. C., Greenstone, H. L., Booy, F. P., Black, L. W., and Ross, P. D. (1992) *J. Biol. Chem.* **228**, 870–884
235. van Driel, R. (1980) *J. Mol. Biol.* **138**, 27–42
236. King, J., and Mykolajewycz, N. (1973) *J. Mol. Biol.* **75**, 339–358
237. Feller, S. M., Ren, R., Hanafusa, H., and Baltimore, D. (1994) *Trends Biochem. Sci.* **19**, 453–458
238. Pawson, T. (1995) *Nature (London)* **373**, 573–580
239. Koch, C. A., Anderson, D., Moran, M. F., Ellis, C., and Pawson, T. (1991) *Science* **252**, 668–674
240. Maignan, S., Guilloteau, J.-P., Fromage, N., Arnoux, B., Becquart, J., and Ducruix, A. (1995) *Science* **268**, 291–293
241. Ullner, M., Selander, M., Persson, E., Stenflo, J., Drakenberg, T., and Teleman, O. (1992) *Biochemistry* **31**, 5974–5983
242. Meininger, D. P., Hunter, M. J., and Komives, E. A. (1995) *Protein Sci.* **4**, 1683–1695
243. Bersch, B., Hernandez, J.-F., Marion, D., and Arlaud, G. J. (1998) *Biochemistry* **37**, 1204–1214
244. Rand, M. D., Lindblom, A., Carlson, J., Villoutreix, B. O., and Stenflo, J. (1997) *Protein Sci.* **6**, 2059–2071
245. Müller, K., Gombert, F. O., Manning, U., Grossmüller, F., Graff, P., Zaegel, H., Zuber, J. F., Freuler, F., Tschopp, C., and Baumann, G. (1996) *J. Biol. Chem.* **271**, 16500–16505
246. Fry, M. J., Panayotou, G., Booker, G. W., and Waterfield, M. D. (1993) *Protein Sci.* **2**, 1785–1797
247. Malek, S. N., Yang, C. H., Earnshaw, W. C., Kozak, C. A., and Desiderio, S. (1996) *J. Biol. Chem.* **271**, 6952–6962
248. Tong, L., Warren, T. C., King, J., Betageri, R., Rose, J., and Jakes, S. (1996) *J. Mol. Biol.* **256**, 601–610
249. Xu, R. X., Word, J. M., Davis, D. G., Rink, M. J., Willard, D. H., Jr., and Gampe, R. T., Jr. (1995) *Biochemistry* **34**, 2107–2121
250. Mikol, V., Baumann, G., Zurini, M. G. M., and Hommel, U. (1995) *J. Mol. Biol.* **254**, 86–95
251. Pascal, S. M., Yamazaki, T., Singer, A. U., Kay, L. E., and Forman-Kay, J. D. (1995) *Biochemistry* **34**, 11353–11362
252. Zhou, M.-M., Meadows, R. P., Logan, T. M., Yoon, H. S., Wade, W. S., Ravichandran, K. S., Burakoff, S. J., and Fesik, S. W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7784–7788
253. Metzler, W. J., Leiting, B., Pryor, K., Mueller, L., and Farmer, B. T., II. (1996) *Biochemistry* **35**, 6201–6211
254. Viguera, A. R., Arrondo, J. L. R., Musacchio, A., Saraste, M., and Serrano, L. (1994) *Biochemistry* **33**, 10925–10933
255. Liang, J., Chen, J. K., Schreiber, S. L., and Clardy, J. (1996) *J. Mol. Biol.* **257**, 632–643
256. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) *Nature (London)* **372**, 375–379
257. Guruprasad, L., Dhanaraj, V., Timm, D., Blundell, T. L., Gout, I., and Waterfield, M. D. (1995) *J. Mol. Biol.* **248**, 856–866
258. van der Geer, P., and Pawson, T. (1995) *Trends Biochem. Sci.* **20**, 277–280
259. Zhou, M.-M., Ravichandran, K. S., Olejniczak, E. T., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J., and Fesik, S. W. (1995) *Nature (London)* **378**, 584–592
260. Rameh, L. E., Arvidsson, A.-k, Carraway, K. L., III, Couvillon, A. D., Rathbun, G., Crompton, A., VanRenterghem, B., Czech, M. P., Ravichandran, K. S., Burakoff, S. J., Wang, D.-S., Chen, C.-S., and Cantley, L. C. (1997) *J. Biol. Chem.* **272**, 22059–22066
261. Hemmings, B. A. (1997) *Science* **275**, 1899
262. Blomberg, N., Baraldi, E., Nilges, M., and Saraste, M. (1999) *Trends Biochem. Sci.* **24**, 441–445
263. Koshiba, S., Kigawa, T., Kim, J.-H., Shirouzu, M., Bowtell, D., and Yokoyama, S. (1997) *J. Mol. Biol.* **269**, 579–591
264. Saras, J., and Heldin, C.-H. (1996) *Trends Biochem. Sci.* **21**, 455–458
265. Cabral, J. H. M., Petosa, C., Sutcliffe, M. J., Raza, S., Byron, O., Poy, F., Marfatia, S. M., Chishti, A. H., and Liddington, R. C. (1996) *Nature (London)* **382**, 649–652
266. Ponting, C. P. (1997) *Protein Sci.* **6**, 464–468
267. Resnick, D., Pearson, A., and Krieger, M. (1994) *Trends Biochem. Sci.* **19**, 5–8
268. Chang, Y., Mochalkin, I., McCance, S. G., Cheng, B., Tulinsky, A., and Castellino, F. J. (1998) *Biochemistry* **37**, 3258–3271
269. de Vos, A. M., Ultsch, M. H., Kelley, R. F., Padmanabhan, K., Tulinsky, A., Westbrook, M. L., and Kossiakoff, A. A. (1992) *Biochemistry* **31**, 270–279
270. De Serrano, V. S., and Castellino, F. J. (1992) *Biochemistry* **31**, 11698–11706
271. Sun, M.-F., Zhao, M., and Gailani, D. (1999) *J. Biol. Chem.* **274**, 36373–36377
272. Macias, M. J., Hyvönen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996) *Nature (London)* **382**, 646–649
273. Bogusky, M. J., Dobson, C. M., and Smith, R. A. G. (1989) *Biochemistry* **28**, 6728–6735
274. Hoffmann, W., and Hauser, F. (1993) *Trends Biochem. Sci.* **18**, 239–243
275. Lamb, J. R., Tugendreich, S., and Hieter, P. (1995) *Trends Biochem. Sci.* **20**, 257–259
276. Das, A. K., Cohen, P. T. W., and Barford, D. (1998) *EMBO J.* **17**, 1192–1199
277. Barlow, P. N., Luisi, B., Milner, A., Elliott, M., and Everett, R. (1994) *J. Mol. Biol.* **237**, 201–211
278. Lichtarge, O., Yamamoto, K. R., and Cohen, F. E. (1997) *J. Mol. Biol.* **274**, 325–337
279. Saurin, A. J., Borden, K. L. B., Boddy, M. N., and Freemont, P. S. (1996) *Trends Biochem. Sci.* **21**, 208–214
280. Pérez-Alvarado, G. C., Kosa, J. L., Louis, H. A., Beckerle, M. C., Winge, D. R., and Summers, M. F. (1996) *J. Mol. Biol.* **257**, 153–174
281. Porter, K. R., and Tucker, J. B. (1981) *Sci. Am.* **244**(Mar), 57–67
282. Weber, K., and Osborn, M. (1985) *Sci. Am.* **253**(Oct), 110–120
283. Amos, L. A., and Amos, B. W. (1991) *Molecules of the Cytoskeleton*, The Guilford Press, New York
284. Lazarides, E. (1980) *Nature (London)* **283**, 249–256
285. Fuchs, E., and Weber, K. (1994) *Ann. Rev. Biochem.* **63**, 345–382
286. Fuchs, E., and Cleveland, D. W. (1998) *Science* **279**, 514–519
287. Kaufmann, S. H. (1989) *J. Biol. Chem.* **264**, 13946–13955
288. Ip, W., Hartzer, M. K., Pang, Y.-Y. S., and Robson, R. M. (1985) *J. Mol. Biol.* **183**, 365–375
289. Herrmann, H., Häner, M., Brettel, M., Müller, S. A., Goldie, K. N., Fedtke, B., Lustig, A., Franke, W. W., and Aebi, U. (1996) *J. Mol. Biol.* **264**, 933–953
289a. Herrmann, H., Strelkov, S. V., Feja, B., Rogers, K. R., Brettel, M., Lustig, A., Häner, M., Parry, D. A. D., Steinert, P. M., Burkhard, P., and Aebi, U. (2000) *J. Mol. Biol.* **298**, 817–832
289b. Heimburg, T., Schuenemann, J., Weber, K., and Geisler, N. (1996) *Biochemistry* **35**, 1375–1382
290. Geisler, N., Schünemann, J., and Weber, K. (1992) *Eur. J. Biochem.* **206**, 841–852
290a. Bellin, R. M., Sernett, S. W., Becker, B., Ip, W., Huiatt, T. W., and Robson, R. M. (1999) *J. Biol. Chem.* **274**, 29493–29499
291. Leterrier, J. F., Käs, J., Hartwig, J., Vegners, R., and Janmey, P. A. (1996) *J. Biol. Chem.* **271**, 15687–15694
292. Betts, J. C., Blackstock, W. P., Ward, M. A., and Anderton, B. H. (1997) *J. Biol. Chem.* **272**, 12922–12927
293. Carter, J., Gragerov, A., Konvicka, K., Elder, G., Weinstein, H., and Lazzarini, R. A. (1998) *J. Biol. Chem.* **273**, 5101–5108
294. Osborn, M., and Weber, K. (1986) *Trends Biochem. Sci.* **11**, 469–472
295. Gibson, M. A., Kumaratilake, J. S., and Cleary, E. G. (1989) *J. Biol. Chem.* **264**, 4590–4598
296. Steinert, P. M. (1991) *J. Structural Biol.* **107**, 157–174
297. Cohlberg, J. A. (1993) *Trends Biochem. Sci.* **18**, 360–362
298. Nakajima-Iijima, S., Hamada, H., Reddy, P., and Kakunaga, T. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6133–6137
299. Orlova, A., Chen, X., Rubenstein, P. A., and Egelman, E. H. (1997) *J. Mol. Biol.* **271**, 235–243
300. Fulton, A. B. (1984) *The Cytoskeleton: Cellular Architecture and Choreography*, Chapman and Hall, New York
301. Bershadsky, A. D., and Vasiliev, J. M. (1988) *Cytoskeleton*, Plenum, New York
302. DeRosier, D. J., Tilney, L. G., and Egelman, E. (1980) *Nature (London)* **287**, 291–296
303. Collier, N. C., and Wang, K. (1982) *J. Biol. Chem.* **257**, 6937–943
304. Moon, R. T., and Lazarides, E. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5494–5499
305. Meek, R. L., Lonsdale-Eccles, J. D., and Dale, B. A. (1983) *Biochemistry* **22**, 4867–4877
306. Steinert, P. M., and Idler, W. W. (1979) *Biochemistry* **18**, 5664–5669
307. Dustin, P. (1980) *Sci. Am.* **243**(Aug.), 67–76
308. Mandelkow, E., and Mandelkow, E.-M. (1994) *Current Opinion in Structural Biology* **4**, 171–179
309. Hyams, J. S., and Lloyd, C. W., eds. (1994) *Microtubules*, Wiley-Liss, New York
310. Sosa, H., and Milligan, R. A. (1996) *J. Mol. Biol.* **260**, 743–755
311. Margulis, L., To, L., and Chase, D. (1978) *Science* **200**, 1118–1124
312. Thormählen, M., Marx, A., Müller, S. A., Song, Y.-H., Mandelkow, E.-M., Aebi, U., and Mandelkow, E. (1998) *J. Mol. Biol.* **275**, 795–809
313. Weber, K., Pollack, R., and Bibring, T. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 459–463
314. Wolf, S. G., Nogales, E., Kikkawa, M., Gratzinger, D., Hirokawa, N., and Downing, K. H. (1996) *J. Mol. Biol.* **262**, 485–501
315. Nogales, E., Wolf, S. G., and Downing, K. H. (1998) *Nature (London)* **391**, 199–203
316. Pennisi, E. (1998) *Science* **279**, 176–177
317. Zheng, Y., Wong, M. L., Alberts, B., and Mitchison, T. (1995) *Nature (London)* **378**, 578–583
318. Hackney, D. D. (1995) *Nature (London)* **376**, 215–216

References

319. Raff, E. C., Fackenthal, J. D., Hutchens, J. A., Hoyle, H. D., and Turner, F. R. (1997) *Science* **275**, 70–73
320. Angelastro, J. M., and Purich, D. L. (1992) *J. Biol. Chem.* **267**, 25685–25689
321. Terry, B. J., and Purich, D. L. (1980) *J. Biol. Chem.* **255**, 10532–10536
322. Rakic, P., Knyihar-Csillik, E., and Csillik, B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9218–9222
323. Rodionov, V. I., and Borisy, G. G. (1997) *Science* **275**, 215–218
324. Hoenger, A., and Milligan, R. A. (1996) *J. Mol. Biol.* **263**, 114–119
325. Fan, J., Griffiths, A. D., Lockhart, A., Cross, R. A., and Amos, L. A. (1996) *J. Mol. Biol.* **259**, 325–330
326. Hirose, K., Fan, J., and Amos, L. A. (1995) *J. Mol. Biol.* **251**, 329–333
327. Knop, M., and Schiebel, E. (1997) *EMBO J.* **16**, 6985–6995
328. Moritz, M., Braunfeld, M. B., Sedat, J. W., Alberts, B., and Agard, D. A. (1995) *Nature (London)* **378**, 638–640
329. Berridge, M. J. (1990) *J. Biol. Chem.* **265**, 9583–9586
330. Arnold, C. S., Johnson, G. V. W., Cole, R. N., Dong, D. L.-Y., Lee, M., and Hart, G. W. (1996) *J. Biol. Chem.* **271**, 28741–28744
331. Bosc, C., Cronk, J. D., Pirollet, F., Watterson, D. M., Haiech, J., Job, D., and Margolis, R. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2125–2130

Study Questions

1. Rewrite Equations 6-75 through 6-77 in terms of dissociation constants. These may be labeled K_1, K_2, K_i, etc., as is conventional, but you may prefer to use K_{1d}, K_{2d}, K_{id}, etc., to avoid confusion.

2. A molecule has two identical binding sites for a ligand X. The Gibbs energy of interaction between ligands bound to the same molecule, ε, is defined as the change in Gibbs energy of binding of the ligand to the molecule that results from the prior binding of a ligand at the adjacent site. If the saturation fraction is Y, show from the equation for the binding isotherm that the following equation holds when Y = 1/2:

$$dY/d\ln[X] = \frac{1}{2(1+e^{\varepsilon/RT})}$$

3. The hydrogen ion binding curve for succinate is shown is Fig. 7-4. From the curve estimate ε and the microscopic association constants.

4. A linear chain molecule has a very large number of identical binding sites for a ligand X. The Gibbs energy of interaction between ligands bound to adjacent sites is ε. Interactions between non-nearest neighbors are considered negligible. If the binding constant for a site adjacent to unoccupied sites is K_r, the binding isotherm is given by

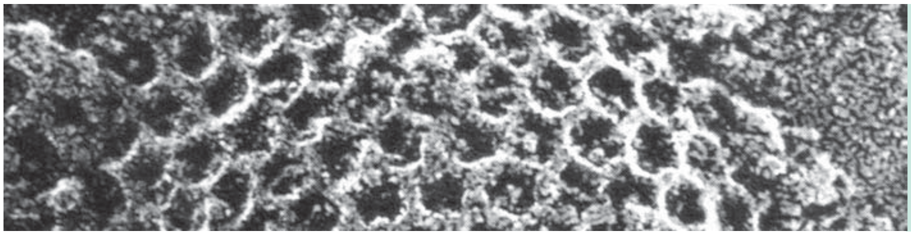
$$Y = \frac{1}{2} + \frac{K[X]e^{-\varepsilon/RT} - 1}{2\{K[X]e^{-\varepsilon/RT} - 1)^2 + 4K[X]\}^{1/2}}$$

[Applequist, J. (1977) *J. Chem. Ed.* **54**, 417]. Show from the equation for the binding isotherm that the following equation holds at Y = 1/2;

$$dY/d\,ln\,[X] = 1/4e^{\varepsilon/2RT}$$

5. The binding of adenosine to polyribouridylic acid [poly(U)] has been studied by the method of equilibrium dialysis [Huang and Ts'o (1966) *J. Mol. Biol.* **16**, 523]. The table below gives the fraction of poly(U) sites occupied, Y at various molar concentrations of free adenosine [A] at 5°C. Assuming that the nearest-neighbor interaction model is correct, determine the intrinsic association constant for the binding of adenosine to poly U and the free energy of interaction of adjacent bound adenosines. Do the bound molecules attract or repel each other?

$[A] \times 10^3$	Y	$[A] \times 10^3$	Y
0.51	0	3.07	0.72
2.10	0	4.00	0.92
2.70	0.15	6.50	0.93
2.96	0.36	8.50	0.93
3.01	0.52	10.00	1.00



Portion of an endocytic vacuole forming in the plasma membrane of a cultured fibroblast. The view is from the inside of the cell and shows a large clathrin cage assembling to form a coated vesicle. The overall diameter of the vacuole is ~0.2 µm. Clathrin cages vary in size and in the number of faces but are typically ~0.1 µm in diameter (see Fig. 8-27 and associated references). Courtesy of Barbara Pearse.

Contents

380 .... **A. Lipid Structures**
380 .... 1. Fatty Acids, Fatty Alcohols, and Hydrocarbons
382 .... 2. Acylglycerols, Ether Lipids, and Waxes
382 .... 3. Phospholipids
387 .... 4. Glycolipids
389 .... 5. Sphingolipids
389 .... 6. Sterols and Other Isoprenoid Lipids
390 .... **B. Membranes**
390 .... 1. The Structure of Membranes
392 .... *Lipids of membranes*
392 .... *Liquid crystals, liposomes, and artificial membranes*
394 .... *Physical properties of membrane lipids*
396 .... *Functions of phospholipid head groups*
397 .... *Non-bilayer structures of phospholipids*
397 .... *Membrane fluidity and life*
400 .... *Electrical properties of membranes*
401 .... *The two sides of a membrane*
401 .... 2. Membrane Proteins
401 .... *Integral membrane proteins*
402 .... *Anchors for proteins*
403 .... *Analyzing erythocyte membranes*
404 .... *Glycoproteins*
405 .... *Connections to the cytoskeleton*
405 .... *Integrins and focal adhesions*
406 .... *Other cell adhesion molecules*
407 .... *Peripheral proteins of the outer membrane surface*
409 .... 3. Enzymes and Membrane Metabolism
410 .... **C. The Transport of Molecules through Membranes**
410 .... 1. Facilitated Diffusion and Active Transport
411 .... 2. Pores, Channels, and Carriers
411 .... *Porins*
411 .... *Aquaporins*
412 .... *Ion channels*
414 .... *Channel-forming toxins and antibiotics*
414 .... *Ionophores and other mobile carriers*
415 .... 3. The 12-Helix Major Facilitator Superfamily
415 .... *Entrance of sugars into cells*
416 .... *Cotransport of sugars and other nutrients with H^+ or Na^+*
417 .... 4. Active Transport Systems
417 .... *Periplasmic permeases*
419 .... *The bacterial phosphotransferase system*
420 .... 5. Transport of Ions
420 .... *Anions*
421 .... *Cation channels*
422 .... *Active transport of cations*
422 .... *ATP-driven ion pumps*
425 .... 6. Exocytosis, Endocytosis, and the Flow of Membrane Constituents
427 .... **D. Communication**
427 .... **E. The Extracellular Matrix and Cell Walls**
427 .... 1. The Structure of Bacterial Cell Walls
428 .... *The outer membrane of gram-negative bacteria*
431 .... *Teichoic and teichuronic acids*
431 .... *Mycobacteria*
431 .... *Other bacterial coats*
431 .... 2. The Surroundings of Animal Cells
432 .... *The collagens*
436 .... *Elastic fibers*
436 .... *Cartilage and basement membranes*
440 .... *Fibrillin and Marfan's syndrome*
440 .... *The cuticles of invertebrates*
440 .... *Bones, teeth, and shells*
443 .... 3. Cell Walls of Fungi and Green Plants

444 .... **References**
452 .... **Study Questions**

Boxes

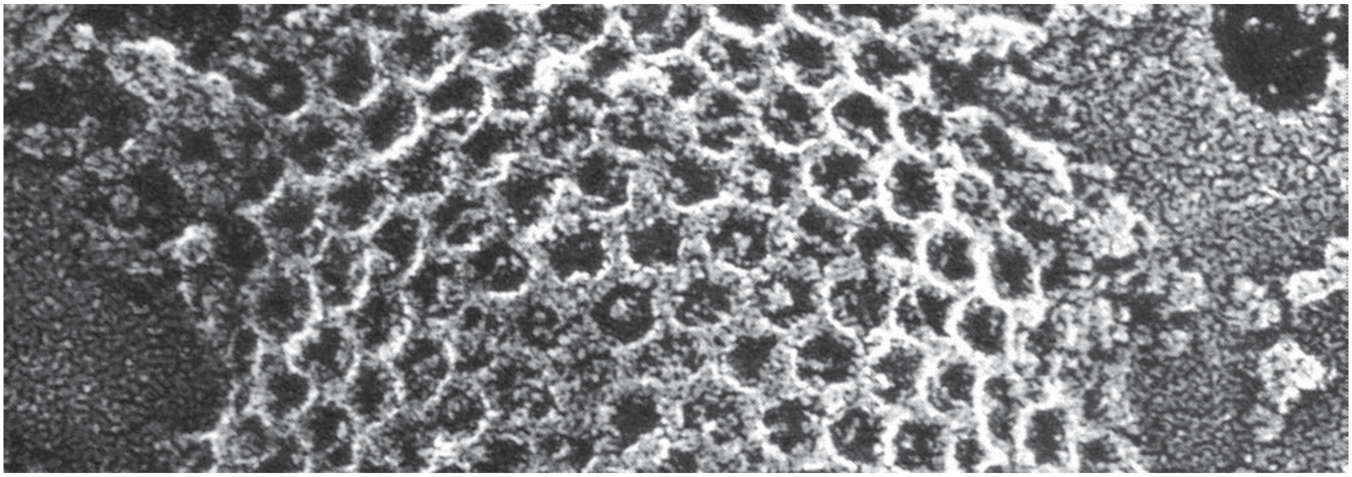
384 .... Box 8-A The Platelet-Activating Factor
386 .... Box 8-B Dipalmitoylphosphatidylcholine and the Surfactant System of the Lungs
398 .... Box 8-C Electron Paramagnetic Resonance (EPR) Spectra and "Spin Lables"
418 .... Box 8-D Colicins: Antibiotic Proteins
438 .... Box 8-E Genetic Defects of Collagen Structure
439 .... Box 8-F Skin
442 .... Box 8-G The Biochemistry of Teeth

Tables

380 .... Table 8-1 Some Important Fatty Acids
381 .... Table 8-2 Fatty Acid Composition (in %) of Some Typical Fats and Oils
393 .... Table 8-3 Estimated Chemical Compositions of Some Membranes
435 .... Table 8-4 Types of Vertebrate Collagen

Lipids, Membranes, and Cell Coats

8



The boundary between a living cell and its surroundings is the incredibly thin (7–10 nm) plasma membrane. This vital partition, which controls the flow of materials into and out of a cell and which senses and controls the response of cells to hormones and other external signals, consists largely of **phospholipids** together with embedded proteins. The nonpolar chains of the phospholipids stick together to form a double molecular layer or **bilayer** which provides the basic structure of almost all biological membranes.

Phospholipids, together with other natural materials that have a high solubility in apolar solvents or are structurally related to compounds with such solubility properties, are classified as **lipids**. The most abundant lipids are the fats, compounds that are stored by animals and by many plants as an energy reserve (Fig. 8-1). Other lipids form the outer cuticle of plants and yet others serve as protective coatings on feathers and hair. Vitamins A, D, K, and E and ubiquinone are all lipids as are a variety of hormones and such light-absorbing plant pigments as the chlorophylls and carotenoids. Many of these compounds are dissolved in or partially embedded in the plasma membrane of bacteria or in the mitochondrial and chloroplast membranes of higher organisms. Membranes serve many purposes. The most obvious is to divide space into compartments. Thus, the plasma membrane forms cell boundaries and mitochondrial membranes separate the enzymes and metabolites of mitochondria from those of the cytosol. Membranes are semipermeable and regulate the penetration into cells and organelles of both ionic and nonionic substances. Many of these materials are brought into the cell against a concentration gradient. Hence, osmotic work must be done in a process known as **active transport**. Many enzymes,

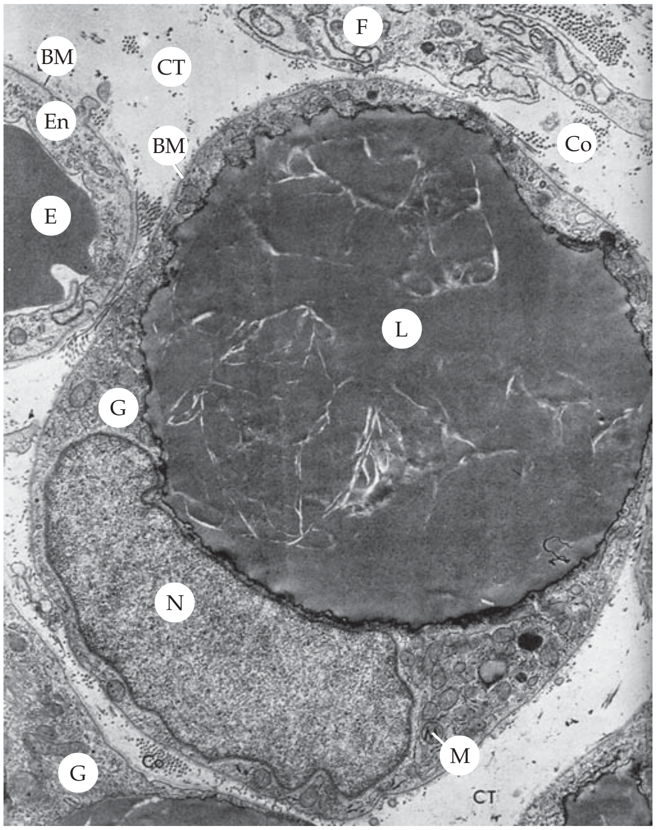


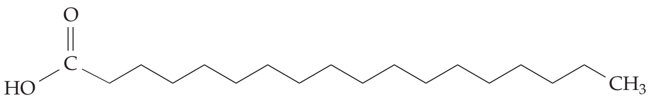
Figure 8-1 Electron micrograph of a thin section of a fat storage cell or adipocyte. L, the single large fat droplet; N, nucleus; M, mitochondria; En, endothelium of a capillary containing an erythrocyte (E); CT, connective tissue ground substance which contains collagen fibers (Co) and fibroblasts (F). The basement membranes (BM) surrounding the endothelium and the fat cell are also marked. From Porter and Bonneville.[6] Courtesy of Mary Bonneville.

including those responsible for most of the oxidative metabolism of cells, are found in membranes of bacteria and of mitochondria. Within the chloroplasts of green leaves, highly folded membranes containing chlorophyll absorb energy from the sunlight. Thin membranes contain the photoreceptor proteins that function in vision. Electrical impulses are transmitted along the membranes of nerve cells.

The outer surfaces of membranes are designed to interact with the cell's external world. Special receptors sense the presence of hormones. Binding proteins await the arrival of needed nutrients and help to bring them into cells. Highly individual arrangements of protein and of the carbohydrate "fuzz" of glycoproteins and sphingolipids screen the outer surface, helping to prevent attack by foreign bacteria, viruses, and toxins.

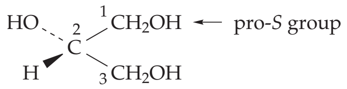
A. Lipid Structures

Unlike proteins, polysaccharides, and nucleic acids, most lipids are not polymers. However, they are made by linking together smaller molecules.[1–5] Among the "building blocks" of lipids are **fatty acids, glycerol, phosphoric acid**, and **sugars**. Many lipids have both polar and nonpolar regions. This gives them an **amphipathic** character, i.e., a tendency toward both hydrophobic and hydrophilic behavior, and accounts for their tendency to aggregate into membranous structures. Notice that in the following structure the carbon atoms of glycerol have been numbered 1 to 3. Although glycerol is *not* chiral, the positioning of the two hydroxymethyl groups is not equivalent. If the *priority*, used in the *RS* system (page 42) for the C-1 group is raised (e.g., by ester formation) to be higher than that for the C-3 group, the molecule would have the *S* configuration. The C-1 group is said to be pro-*S* and the C-3 group pro-*R*. According to the stereochemical numbering (*sn*) system, which is discussed further in Chapter 9, the carbon in the pro-*S* position is numbered 1.



A fatty acid

Stearic acid, 18 carbon atoms, major component of animal triacylglycerols



Glycerol

Labeled by the stereochemical numbering (*sn*) system

1. Fatty Acids, Fatty Alcohols, and Hydrocarbons

The names and structures of some fatty acids are summarized in Table 8-1. Notice that these acids have straight carbon chains and may contain one or more double bonds. Except for the smallest members of the series, which are soluble in water, fatty acids are strongly hydrophobic. However, they are all acids with pK_a values in water of ~4.8. To the extent that free fatty acids occur in nature, they are likely to be found in interfaces between lipid and water with the carboxyl groups dissociated and protruding into the water. However, most naturally occurring fatty acids

TABLE 8-1
Some Important Fatty Acids

No. carbon atoms	Systematic name	Common name	Abbreviation[a]	Common name of acyl group
Saturated fatty acids				
1	Methanoic	Formic		Formyl
2	Ethanoic	Acetic		Acetyl
3	Propanoic	Propionic		Propionyl
4	Butanoic	Butyric	4:0	Butyryl
12	Dodecanoic	Lauric	12:0	Lauroyl[b]
14	Tetradecanoic	Myristic	14:0	Myristoyl
16	Hexadecanoic	Palmitic	16:0	Palmitoyl
18	Octadecanoic	Stearic	18:0	Stearoyl
20	Eicosanoic	Arachidic	20:0	
22	Docosanoic	Behenic	22:0	
24	Tetracosanoic	Lignoceric	24:0	
Unsaturated fatty acids[c]				
4		Crotonic	4:1(2*t*)	Crotonoyl
16		Palmitoleic	16:1(9*c*)	
18		Oleic	18:1(9*c*)	Oleoyl
18		Vaccenic	18:1(11*c*)	
18		Linoleic	18:2(9c,12*c*)	
18		Conjugated linoleic	18:2(9c,11*t*)	
18		Linolenic	18:3(9*c*,12*c*,15*c*)	
20		Arachidonic	20:4(5*c*,8*c*,11*c*,14*c*)	

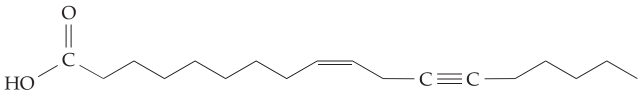
[a] The number of carbon atoms is given first, then the number of double bonds. The positions of the lowest numbered carbon of each double bond and whether the configuration is *cis* (*c*) or *trans* (*t*) are indicated in parentheses.

[b] Official IUPAC names of these and other acyl groups have been designated by the Commission of the Nomenclature of Organic Chemistry in *Pure and Applied Chemistry* **10**, 111–125 (1965). In a number of cases IUPAC inserted an *o* in the traditional name, e.g., palmityl became palmitoyl and crotonyl became crotonoyl. However, acetyl was not changed. In many cases the systematic names, e.g., hexadecanoyl (from hexadecanoic acid), are preferable and IUPAC–IUB recommends that alkyl radicals always be designated by systematic names, e.g., hexadecyl, *not* palmityl alcohol. The older use of palmityl for both acyl and alkyl radicals was one reason for IUPAC's adoption of new names for acyl radicals.

[c] Systematic names are not often used because of their complexity, e.g., linolenic acid is *cis,cis,cis*-9,12,15-octadecatrienoic acid.

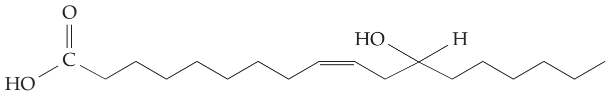
are esterified or combined via amide linkages in complex lipids. For example, ordinary fats are largely the fatty acid esters of glycerol called **triacylglycerols** (triglycerides).

There is a seemingly endless variety of fatty acids, but only a few of them predominate in any single organism. Most fatty acid chains contain an even number of carbon atoms. In higher plants the C_{16} **palmitic acid** and the C_{18} unsaturated **oleic** and **linoleic acids** predominate. The C_{18} saturated **stearic acid** is almost absent from plants and C_{20} to C_{24} acids are rarely present except in the outer cuticle of leaves. Certain plants contain unusual fatty acids which may be characteristic of a taxonomic group. For example, the Compositae (daisy family) contain acetylenic fatty acids and the castor bean contains the hydroxy fatty acid **ricinoleic acid**.



Crepenynic acid 18:2 (9c, 12a)

Accounts for 60% of the fatty acids in seeds of *Crepis foetida*, a member of the compositae family



Ricinoleic acid (12-hydroxyoleic acid)

Accounts for up to 90% of the fatty acids of *Ricinus communis* (castor bean)

Like plants, animals contain palmitic and oleic acids. In addition, large amounts of stearic acid and small amounts of the C_{20}, C_{22}, and C_{24} acids are also present. Phospholipids of photoreceptor membranes of the retina contain fatty acid chains as long as C_{36}.[7] The variety of fatty acids found in animals is greater than in a given plant species. A large fraction of the fatty acids present in most higher organisms are *unsaturated* and contain strictly *cis* double bonds. Table 8-2 shows the fatty acid composition of some typical triacylglycerol mixtures.

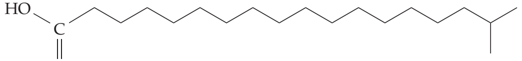
TABLE 8-2
Fatty Acid Composition (in %) of Some Typical Fats and Oils[a]

	No. of carbon atoms and (following colon) the number of double bonds					
Fats and oils	14	16	18	16:1	18:1	18:2
Human depot fat	3	23	4	8	45	10
Beef tallow	4	30	25	5	36	1
Corn oil		13	2		31	54
Lard	1	28	15	3	42	9

[a] From Gunstone.[8]

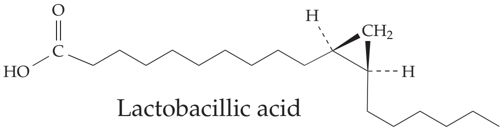
Bacteria usually lack polyunsaturated fatty acids but often contain branched fatty acids, cyclopropane-containing acids, hydroxy fatty acids, and unesterified fatty acids. Mycobacteria, including the human pathogen *Mycobacterium tuberculosis*, contain **mycolic acids**. In these compounds the complex grouping R contains a variety of functional groups including –OH, $-OCH_3$, C=O, –COOH, cyclopropane rings, methyl branches, and C=C bonds. Each species of *Mycobacterium* contains about two dozen different mycolic acids[9,10] as well as other complex $C_{30}-C_{56}$ fatty acids (see Eq. 21-5).[11]

Branched fatty acids of the anteiso series have a branch here and a 5-carbon "starter piece" derived from isoleucine



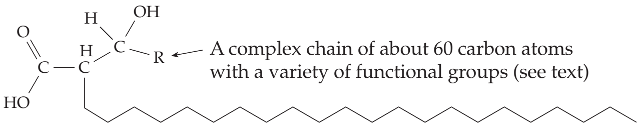
Branched fatty acids of the iso series contain a 5-carbon "starter piece" derived from leucine or a 4-carbon piece derived from valine

Branched fatty acids



Lactobacillic acid

A major fatty acid of lactobacilli

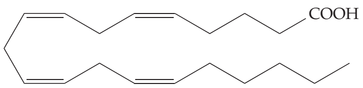


A complex chain of about 60 carbon atoms with a variety of functional groups (see text)

C_{22} or C_{24} alkyl group

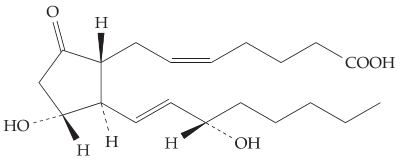
A mycolic acid

Certain polyunsaturated fatty acids are essential in the human diet (see Box 21-B). One of these, **arachidonic acid** (which may be formed from dietary linoleic acid), serves as a precursor for the formation of the hormones known as **prostaglandins** and a series of related **prostanoids**. Lipids of animal origin also



Arachidonic acid

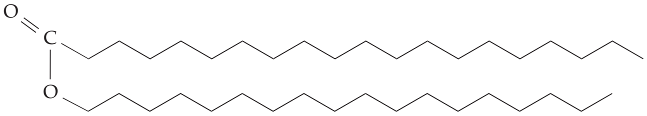
A nutritionally essential fatty acid (Box 21-A) is shown here in a folded conformation



Prostaglandin PGE_2

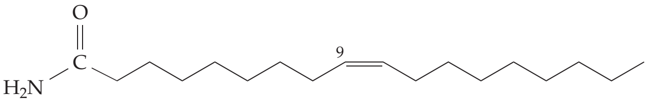
One of a family of hormones, this one is synthesized in tissues from arachidonic acid contain unusual unsaturated fatty acids. Among them, **conjugated linoleic acids** are receiving attention for their possible cancer-preventive action. The predominant form in meats, dairy products, and the human body is the C_{18} 9-*cis*, 11-*trans* isomer whose two double bonds are conjugated.[11a]

Other lipid components include the **fatty alcohols** which are formed by reduction of the acids. These are esterified with fatty acids to form **waxes**. Both fatty alcohols and free fatty acids occur in waxes together with the esterified forms. These mixtures are found on exterior surfaces of plants and animals. Plants and, to a limited extent, animals are able to decarboxylate fatty acids in a multistep process to **alkanes** and these too are important constituents of some waxes. Small amounts of fatty acid amides such as *cis*-9,10-octadecenoamide are present in low concentrations in the cerebrospinal fluid of cats and rats as well as humans. This compound accumulates in cats that are deprived of sleep. When the synthetic compound was injected into rats they fell into apparently normal sleep.[12]



A wax

An ester of a fatty acid and a fatty alcohol

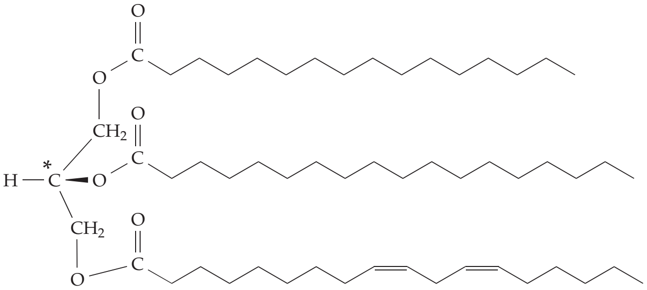


cis-9,10-Octadecenoamide (Oleic acid amide)

Insects make unsaturated as well as saturated hydrocarbons. The former as well as long-chain alcohols and their esters often form the volatile **pheromones** with which insects communicate. Thus, the female pink bollworm attracts a male with a sex pheromone consisting of a mixture of the *cis,cis* and *cis,trans* isomers of 7,11-hexadecadienyl acetate,[13] and European corn borer males are attracted across the cornfields of Iowa by *cis*-11-tetradecenyl acetate.[14] Addition of a little of the *trans* isomer makes the latter sex attractant much more powerful. Since more than one species uses the same attractant, it is possible that the males can distinguish between different ratios of isomers or of mixtures of closely related substances.

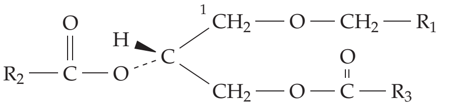
2. Acylglycerols, Ether Lipids, and Waxes

The components of complex lipids are linked in a variety of ways. Often, glycerol acts as the central unit, e.g., combining in ester linkage with three fatty acids to form **triacylglycerols** (triglycerides), the common fats of adipose tissues and plant oils. Diacyl- and monoacylglycerols (diglycerides, and monoglycerides) are present to a lesser extent. In addition, small amounts of **alkyl ethers** or **alkenyl ethers** are often present in isolated lipids. They are especially abundant in fish liver oils.

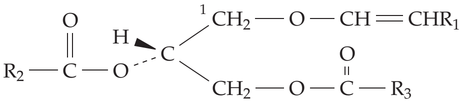


A triacylglycerol (fat)

Notice the chiral center, designated by the asterisk



1-Alkyl-2,3-diacyl-*sn*-glycerol, an alkyl ether lipid



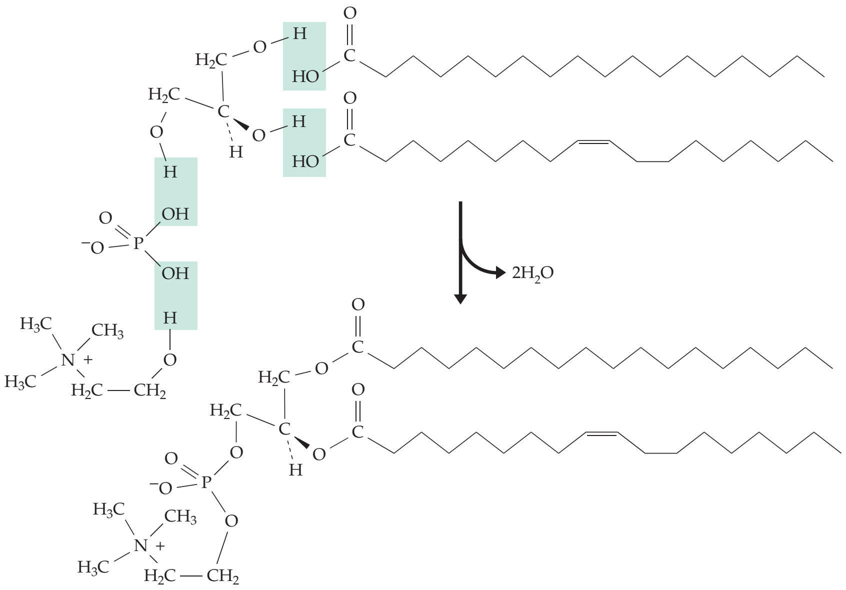
1-Alkenyl-2,3-diacyl-*sn*-glycerol, an alkenyl ether

These ether lipids are all chiral molecules with an *R* configuration but are derivatives of the nonchiral glycerol. The carbon atoms of glycerol are numbered using the stereochemical system which is described on p. 470. The ether linkage is to the *sn*-1 carbon atom. Most phospholipids are derivatives of the *sn*-3 phosphate ester of glycerol.

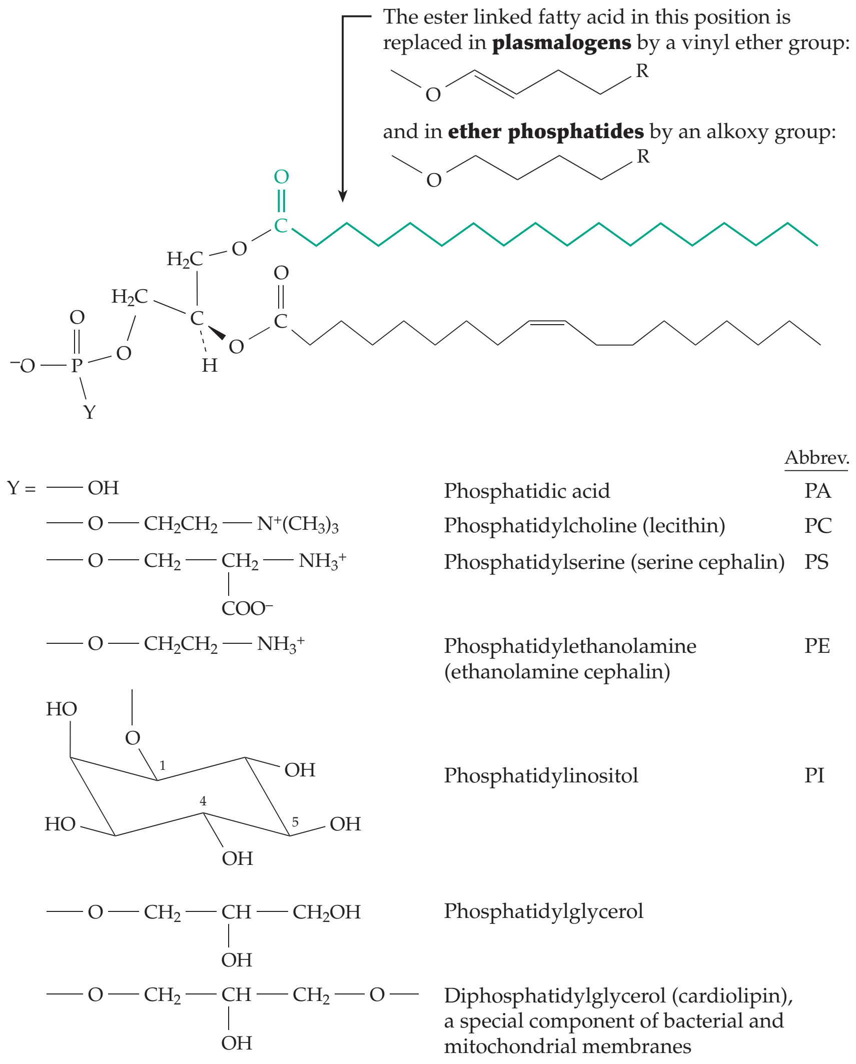
Triacylglycerols and the ether lipids described in the previous section are classified as **neutral lipids**. Other neutral lipids are alcohols, waxes, aldehydes, and hydrocarbons derived from fatty acids. These sometimes have specific biological functions. For example, fatty aldehydes are important in the bioluminescence of bacteria (Eq. 23-47).

3. Phospholipids

As major constituents of biological membranes, phospholipids play a key role in all living cells. The two principal groups of phospholipids are the



A phosphatidylcholine (lecithin). (8-1)

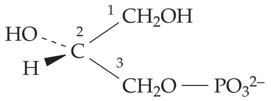

The ester linked fatty acid in this position is replaced in **plasmalogens** by a vinyl ether group:

and in **ether phosphatides** by an alkoxy group:

Abbrev.
Y = —OH Phosphatidic acid PA
Phosphatidylcholine (lecithin) PC
Phosphatidylserine (serine cephalin) PS
Phosphatidylethanolamine (ethanolamine cephalin) PE
Phosphatidylinositol PI
Phosphatidylglycerol
Diphosphatidylglycerol (cardiolipin), a special component of bacterial and mitochondrial membranes

Figure 8-2 Structures of some phosphatides (glycerophospholipids).

glycerophospholipids (glycerophosphatides; Fig. 8-2) which contain the alcohol glycerol and the **sphingophospholipids** which contain the alcohol **sphingosine** (Section 5). The glycerophospholipids can be thought of as arising from the building blocks glycerol, fatty acids, the dihydrogen phosphate ion $H_2PO_4^-$, and the appropriate alcohol by removal of four molecules of water (Eq. 8-1). They are derivatives of *sn*-glycerol-3-phosphate. Esterification of this alcohol with two fatty acids gives a **phosphatidic acid** (Fig. 8-2). Formation of a phosphate diester linkage to one of the alcohols choline, serine, or ethanolamine yields a glycerophospholipid. The resulting three groups of phospholipids are called **phosphatidylcholine** (lecithin), **phosphatidylserine**, and **phosphatidylethanolamine**, respectively (Fig. 8-2).



sn-Glycerol 3-phosphate
(L-α-glycerol phosphate)

The phosphate and choline, ethanolamine, or serine portions of the phosphatide are electrically charged and provide a polar "head" for the molecule. In all three cases the positively charged group is able to fold back and form an ion pair with the negatively charged phosphate group. However, the methyl groups surrounding the nitrogen in phosphatidylcholine prevent a very close approach and with phosphatidylserine the adjacent carboxylate group weakens this electrostatic interaction. Unlike the triacylglycerols, most of which are liquid at body temperature, phospholipids are solid at this temperature. This property, like the ionic properties of the phosphatides, is doubtless related to their suitability for functioning in biological membranes.

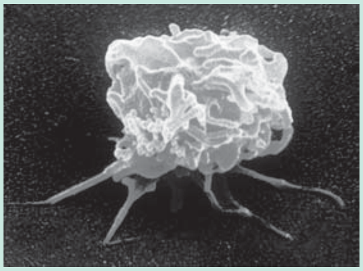
Lecithins and related phospholipids usually contain a saturated fatty acid in the C-1 position but an unsaturated acid, which may contain from one to four double bonds, at C-2. Arachidonic acid is often present here. Hydrolysis of the ester linkage at C-2 yields a 1-acyl-3-phosphoglycerol, better known as a **lysophosphatidylcholine**. The name comes from the powerful detergent action of these substances which leads to lysis of cells. Some snake venoms contain phospholipases that form lysophosphatidylcholine. Lysophosphatidic acid (1-acyl-glycerol-3-phosphate) is both an intermediate in phospholipid biosynthesis (Chapter 21) and also a signaling molecule released into the bloodstream by activated platelets.[15]

Another group of phosphatides contain the hexahydroxycyclohexane known as **inositol** (Fig. 8-2, see also Chapter 21).[16] Phosphatidylinositol, as well as smaller amounts of phosphatides derived from phosphate esters of inositol are present in membranes of all eukaryotes and have a specific role in regulating responses of cells to hormones and other external agents. See Chapter 11 for details. Phosphatidylinositol also forms part of "anchors" used to hold certain proteins onto membrane surfaces (see Fig. 8-13).

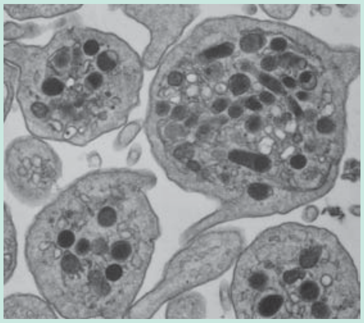
Bacteria and plants often make the anionic **phosphatidyglycerol** in which the second glycerol is esterified at its *sn*-1 position with the phosphate. Bacteria, as well as mitochondria, contain diphosphatidylglycerol (**cardiolipin**) in which phosphatidyl groups are attached at both the 1 and 3 positions of glycerol (Fig. 8-2). Ether phospholipids, analogous to the ether lipids described in Section 2, are also widely distributed. The alkenyl ether analogs of phosphatidylcholine (Fig. 8-2) are called **plasmalogens**.[17] In neutrophils the 1-*O*-alkyl ethers contain the major share of the cell's arachidonic acid, which is esterified in the 2 position.[18,19]

In halophilic (salt loving), thermophilic, and methanogenic bacteria, most of the lipids present are either

BOX 8-A THE PLATELET-ACTIVATING FACTOR



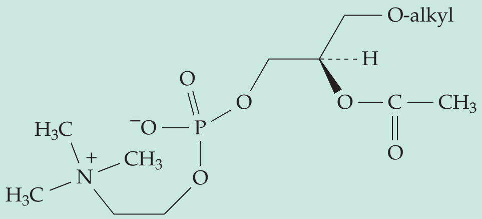
Scanning EM of activated blood platelets. © Quest, Photo Researchers



Transmission EM of thin section through activated blood platelets. © Photo Researchers

Part of the immune response consists of the release from stimulated neutrophils, macrophages, and other cells of a **platelet activating factor (PAF)**, a material that "activates" blood platelets. Activated platelets aggregate, a response that initiates clot formation. They may also be lysed and release stored substances that include platelet-derived growth factors (see Chapter 11) and fibrin stabilizing factor, a proenzyme that is converted to the protein-crosslinking enzyme transglutaminase (Eq. 2-17). See Fig. 12-17. However, the principal interest in the platelet-activating factor has arisen from its powerful effect in inducing **inflammation** in surrounding tissues.

PAF has been identified as the following simple ether phospholipid:[a–d]

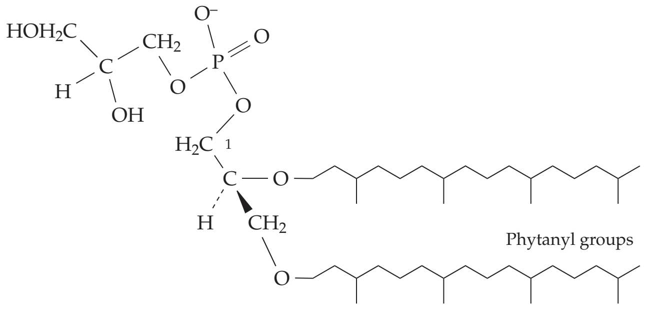


Platelet-activating factor (PAF)
1-*O*-alkyl-2-acetyl-*sn*-glycerophosphocholine

A remarkably potent compound, its effects on platelets are observed at concentrations of 10^{-11} to 10^{-10} M. Both lyso-PAF and the glycero-1- phosphocholine enantiomer are inactive. Specific receptors for this factor are evidently present on platelet surfaces.[e–g]

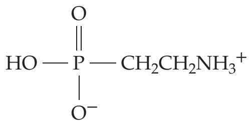
One effect of PAF on platelets is to induce a rapid (5–10s) cleavage of phosphatidylinositol 4,5-bisphosphate by phospholipase C to give **diacylglycerol** and **inositol 1,4,5-trisphosphate**. The subsequent effects of these two substances in causing a rapid influx of Ca^{2+} and in inducing a series of secondary responses are outlined in Fig. 11-9. Among these responses are the release of the materials stored in the platelet's granules. PAF also appears to inhibit adenylate cyclase[b] and causes vasodilation, a property not expected for a compound that stimulates clotting. Receptors for PAF are also present in the brain, where this phosholipid may function in regulation of development.[h]

phospholipids and glycolipids containing the C_{20} isoprenoid **phytanyl** group or the C_{40} **diphytanyl** group[20–25] (see also Section 4), related isoprenoid alcohols, or long-chain 1,2-diols.[25] An example of a diphytanylglycerophospholipid is the following:



A phospholipid derived from 2,3-di-*O*-phytanyl-*sn*-glycerol

Many other phospholipids are present in small amounts or in a limited number of species. These include **phosphonolipids**, which contain a C–P bond and are abundant in ciliate protozoa such as *Tetrahymena* and in some other invertebrates.[26] Phosphonoethylamine replaces phosphoethanolamine in these lipids. A consequence of this structural alteration is a high degree of resistance to the action of the enzyme phospholipase C. The phosphonolipids of the external membrane of *Tetrahymena* are also ether lipids with an alkoxy group in the *sn*-1 position. This makes them



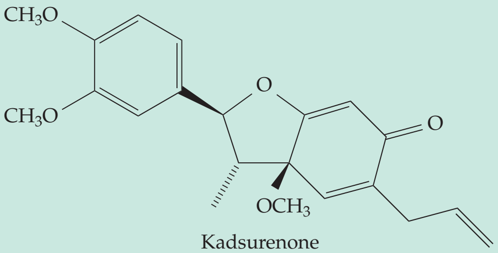
$$HO-\overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle O^-}{|}}{P}}-CH_2CH_2NH_3^+$$

Phosphonoethylamine

BOX 8-A (continued)

Stimulated platelets release arachidonic acid rapidly from their phospholipids, apparently as a result of activation of phospholipase A_2. The released arachidonate can in turn be metabolized to endoperoxides and thromboxane A_2 (Chapter 21). These compounds are also potent activators of platelets and cause a self-activating or **autocrine** effect.[i,j] While PAF has a beneficial function, it can under some conditions contribute in a dangerous way to inflammation and to allergic responses including **anaphylaxis**,[j] **asthma**[g] and **cold-induced urticaria**.[k] Although the effect of PAF is separate from those of histamine and of leukotrienes, these agents may act cooperatively to induce inflammation.[l]

The biosynthesis of PAF is discussed in Chapter 21, Section C. One pathway is deacylation at the glycerol C-2 of a longer chain alkyl phospholipid followed by acetylation at the same position.[m,n] Platelets can inactivate PAF by the inverse sequence: deacetylation followed by acylation with arachidonic acid. PAF can also be hydrolyzed by a PAF acetylhydrolase, a phospholipase.[o] Absence of this enzyme in the brain may be related to a human brain malformation.[h] Some tissues may contain a lipid that inhibits binding of PAF.[p] The following compound from a Chinese herb binds to PAF receptors and may be the forerunner of useful drugs.[q]



Kadsurenone

---

[a] Cusack, N. J. (1980) *Nature (London)* **285**, 193

[b] Hanahan, D. J. (1986) *Ann. Rev. Biochem.* **55**, 483–509

[c] Prescott, S. M., Zimmerman, G. A., and McIntyre, T. M. (1990) *J. Biol. Chem.* **265**, 17381–17384

[d] Winslow, C. M., and Lee, M. L., eds. (1987) *New Horizons in Platelet Activating Factor Research*, Wiley, New York

[e] Hwang, S.-B., Lam, M.-H., and Pong, S.-S. (1986) *J. Biol. Chem.* **261**, 532–537

[f] Ishii, I., Izumi, T., Tsukamoto, H., Umeyama, H., Ui, M., and Shimizu, T. (1997) *J. Biol. Chem.* **272**, 7846–7854

[g] Bazan, N. G. (1995) *Nature (London)* **374**, 501–502

[h] Hattori, M., Adachi, H., Aoki, J., Tsujimoto, M., Arai, H., and Inoue, K. (1995) *J. Biol. Chem.* **270**, 31345–31352

[i] Bussolino, F., Sironi, M., Bocchietto, E., and Mantovani, A. (1992) *J. Biol. Chem.* **267**, 14598–14603

[j] Darius, H., Lefer, D. J., Smith, J. B., and Leefer, A. M. (1986) *Science* **232**, 58–60

[k] Grandel, K. E., Farr, R. S., Wanderer, A. A., Eisenstadt, T. C., and Wasserman, S. I. (1985) *N. Engl. J. Med.* **313**, 405–409

[l] Tomeo, A. C., Egan, R. W., and Durán, W. N. (1991) *FASEB J.* **5**, 2850–2855

[m] Billah, M. M., Eckel, S., Myers, R. F., and Siegel, M. I. (1986) *J. Biol. Chem.* **261**, 5824–5831

[n] Lee, T.-c, Ou, M.-c, Shinozaki, K., Malone, B., and Snyder, F. (1996) *J. Biol. Chem.* **271**, 209–217

[o] Stafforini, D. M., McIntyre, T. M., Zimmerman, G. A., and Prescott, S. M. (1997) *J. Biol. Chem.* **272**, 17895–17898

[p] Miwa, M., Hill, C., Kumar, R., Sugatan, J., Olson, M. S., and Hanahan, D. J. (1987) *J. Biol. Chem.* **262**, 527–530

[q] Hwang, S.-B., Lam, M.-H., Biftu, T., Beattie, T. R., and Shen, T.-Y. (1985) *J. Biol. Chem.* **260**, 15639–15645

BOX 8-B DIPALMITOYLPHOSPHATIDYLCHOLINE AND THE SURFACTANT SYSTEM OF THE LUNGS

When air is exhaled the small alveoli of the lungs could collapse if it were not for the surface active material (surfactant) present in the fluid that coats the lungs.[a–e] In fact, the lack of adequate surfactant is the cause of **respiratory distress syndrome**, a major cause of death among premature infants and a disease that may develop in acute form in adults. The surfactant material forms a thin film of high fluidity at the air–liquid interface and lowers the surface tension from the 72 mN/m of pure water to <10 mN/m.[f,g] (Pay attention to the definition of surface tension.[h]) About 65% by weight of the surfactant is lecithin, mostly dipalmitoylphosphatidylcholine (see Fig. 8-4), a phospholipid resistant to attack by oxygen. Phosphatidylglycerol, in an unusually high concentration, accounts for ~ 12% of the human surfactant. Other phospholipids, plasmalogen,[i] cholesterol, proteins, and calcium ions are also present.

The surfactant contains four unique proteins, designated SP-A, SP-B, SP-C, and SP-D.[c,e] The major protein (SP-A) is a sialic acid-rich glycoprotein derived from a 26-kDa peptide, which contains a short collagen-like domain.[e,j,k] Like collagen, this domain contains glycosylated hydroxyproline. The C-terminal domain is a Ca^{2+}-dependent C-type lectin (Chapter 4), while the N-terminal domain is involved in oligomer formation through disulfide bridges. The overall structure is similar to that of the complement protein C1q (Chapter 31).[e,j] Protein D is also collagen-like[l] but evidently plays a very different functional role than SP-A. The latter associates with the major surfactant lipids but SP-D does not. It does bind phosphatidylinositol[m] and glucosylceramide, lipids that are present in small amounts. Perhaps SP-D helps to remove these polar lipids which might interfere with surfactant action.[e] Both proteins A and D may also have functions in the immune system.[l]

Proteins SP-B and SP-C are small extremely hydrophobic polypeptides consisting of 79 and 35 amino acid residues, respectively.[n,o] Aliphatic branched amino acids constitute 23 of the 35 residues of the C-terminal part of protein C, which is also palmitoylated on two cysteine residues. SP-B is formed from a large 381-residue precursor. The mature protein contains seven cysteines and disulfide bridges. Both proteins have major effects on the properties of the surfactant mixture. They promote rapid reorganization of lipid layers, an important consideration for the functioning of the surfactant. Infants lacking SP-B suffer severe respiratory failure with high mortality.[e]

The properties of the surfactant allow for rapid formation of a large area of lipid monolayer–air interface. The low surface tension and the ability to rapidly spread the mixture of lipids and proteins are essential.[o] At the end of the expiration stage of breathing the surfactant is present in the interface as a strong, tightly packed monolayer whose properties reflect the rigidity of the dipalmitoylphosphatidylcholine. The excess surfactant in the alveolar fluid forms liposome-like bilayer structures and also associates with proteins and calcium ions to form a lattice-like material called **tubular myelin**. Lipid in this form must be transferred rapidly into the air–liquid interface during inspiration.[p] The ability of phospholipids to pass through a hexagonal phase (Fig. 8-12) may also be important for this transition.[q,r]

One enzyme present in the surfactant fluid is an acid phosphatase able to hydrolyze phosphatidylglycerol phosphate, perhaps functioning in the final step of biosynthesis of the phosphatidylglycerol present in the surfactant.[c] Study of the action of the natural lung surfactant has led to development of artificial surfactant mixtures that are being used effectively to save many lives.[d]

---

[a] Rooney, S. A. (1979) *Trends Biochem. Sci.* **4**, 189–191

[b] Persson, A., Chang, D., Rust, K., Moxley, M., Longmore, W., and Crouch, E. (1989) *Biochemistry* **28**, 6361–6367

[c] Rooney, S. A., Young, S. L., and Mendelson, C. R. (1994) *FASEB J.* **8**, 957–967

[d] Jobe, A. H. (1993) *N. Engl. J. Med.* **328**, 861–868

[e] Kuroki, Y., and Voelkers, D. R. (1994) *J. Biol. Chem.* **269**, 25943–25946

[f] Shiffer, K., Hawgood, S., Haagsman, H. P., Benson, B., Clements, J. A., and Goerke, J. (1993) *Biochemistry* **32**, 590–597

[g] Pastrana-Rios, B., Flach, C. R., Brauner, J. W., Mautone, A. J., and Mendelsohn, R. (1994) *Biochemistry* **33**, 5121–5127

[h] Bangham, A. D. (1992) *Nature (London)* **359**, 110

[i] Rana, F. R., Harwood, J. S., Mautone, A. J., and Dluhy, R. A. (1993) *Biochemistry* **32**, 27–31

[j] McCormack, F. X., Calvert, H. M., Watson, P. A., Smith, D. L., Mason, R. J., and Voelker, D. R. (1994) *J. Biol. Chem.* **269**, 5833–5841

[k] Taneva, S., McEachren, T., Stewart, J., and Keough, K. M. W. (1995) *Biochemistry* **34**, 10279–10289

[l] Crouch, E., Persson, A., Chang, D., and Heuser, J. (1994) *J. Biol. Chem.* **269**, 17311–17319

[m] Ogasawara, Y., Kuroki, Y., and Akino, T. (1992) *J. Biol. Chem.* **267**, 21244–21249

[n] Korimilli, A., Gonzales, L. W., and Guttentag, S. H. (2000) *J. Biol. Chem.* **275**, 8672–8679

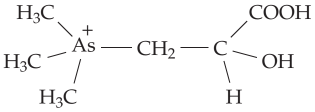
[o] Pastrana, B., Mautone, A. J., and Mendelsohn, R. (1991) *Biochemistry* **30**, 10058–10064

[p] Lipp, M. M., Lee, K. Y. C., Zasadzinski, J. A., and Waring, A. J. (1996) *Science* **273**, 1196–1199

[q] Perkins, W. R., Dause, R. B., Parente, R. A., Minchey, S. R., Neuman, K. C., Gruner, S. M., Taraschi, T. F., and Janoff, A. S. (1996) *Science* **273**, 330–332

[r] Discher, B. M., Maloney, K. M., Grainger, D. W., Sousa, C. A., and Hall, S. B. (1999) *Biochemistry* **38**, 374–383

resistant to phospholipase A_1 as well. These two properties appear to protect the naked cell membranes of the protozoa from their own phospholipases which may be secreted into the environment.[26] Marine algae form an arsenic-containing phospholipid *O*-phosphatidyltrimethylarsonium lactic acid.[27]

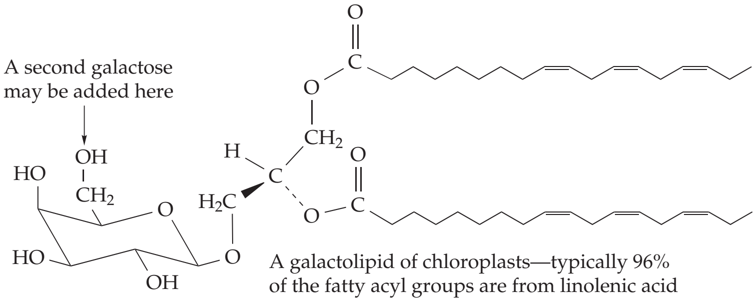


Trimethylarsonium lactic acid

They apparently do this as part of a scheme for detoxifying arsenate taken up with phosphate from the phosphate-poor ocean water.

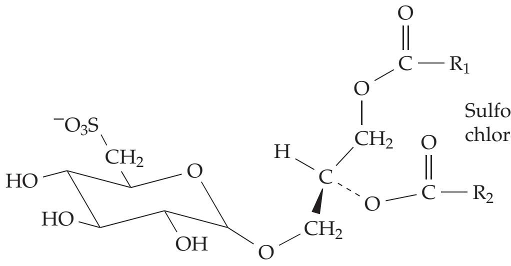
4. Glycolipids

The polar heads of the **glycoglycerolipids** lack phospho groups but contain *sugars* in glycosidic linkage.[28] Large amounts of the galactolipids shown in the following structure are found in chloroplasts.[29] The monogalactosyl diacylglycerol is said to be the most abundant polar lipid in nature.



A galactolipid of chloroplasts—typically 96% of the fatty acyl groups are from linolenic acid

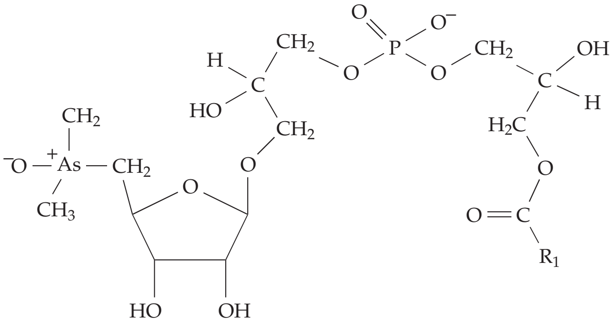
Chloroplasts also contain the following **sulfolipid**, an anionic sulfonate.



Sulfolipid of chloroplasts

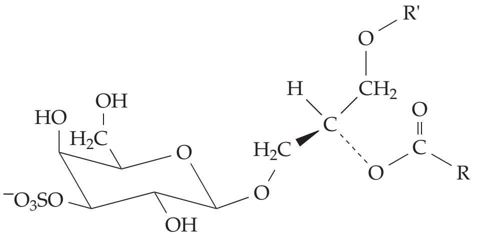
6-Sulfo-6-deoxy-α-D glucopyranosyl diglyceride

Marine algae as well as aquatic higher plants accumulate **arsenophospholipids**.[30]



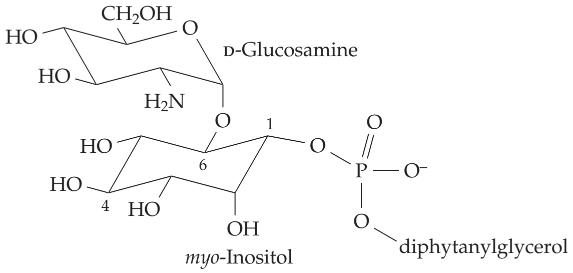
Arsenophospholipid of aquatic plants

The plasma membrane of mammalian male germ cells contains the following **sulfogalactosylglycerolipid**. It is found only in spermatozoa and testes, in which it accounts for 5–8% of total lipid, and in the brain, in which it accounts for only 0.2% of total lipid.[31]

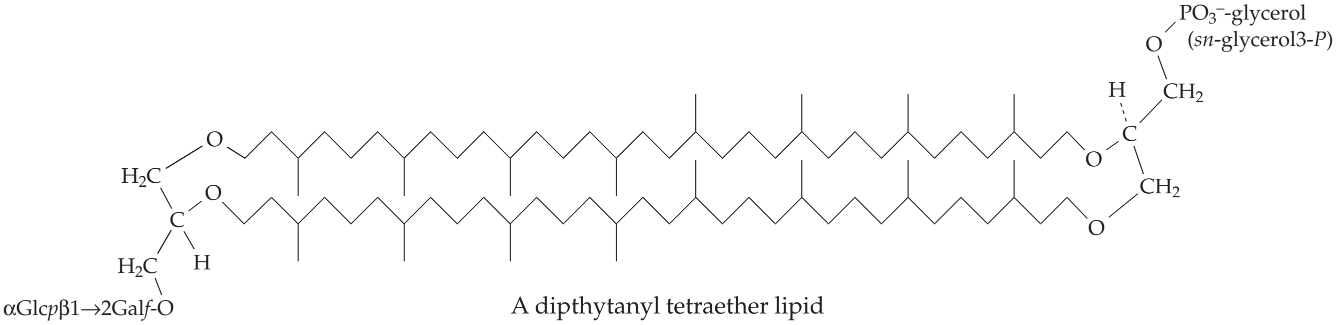


Sulfogalactosylglycerolipid, a sulfate ester

A variety of acylated glucolipids and phosphoglucolipids, including monoglucosyl and diglucosyl diacylglycerols, have been identified in membranes of the cell-wall-less bacterium *Acholeplasma laidlawii*.[32,32a] The following glycolipid from the methanogen *Methanosarcina*[33] is identical to the core structure of eukaryotic glycosylphosphatidylinositol membrane protein anchors (Fig. 8-13).



In addition to the previously mentioned phytanyl ether phospholipids, methanogens contain **diphytanyl tetraether lipids** that are both glycerophospholipids and glycolipids.

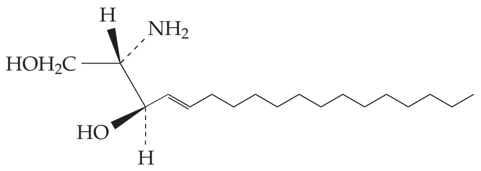


A dipthytanyl tetraether lipid

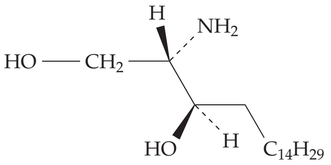
From a methanogenic bacterium.[21]

5. Sphingolipids

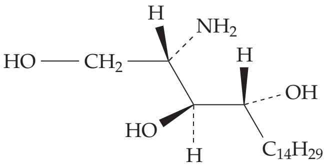
The backbone of the sphingolipids is the basic alcohol **sphingosine** (sphingenine) or a related long chain base. At least 60 such bases have been identified.[34,35] They vary in chain length from C_{14} to C_{22} and include members of the branched iso and anteiso series. Up to two double bonds may be present. Sphingosine contains 18 carbon atoms and is formed from palmitic acid and serine (Fig. 21-6). An intermediate in the formation of sphingosine is the saturated **sphinganine** (dihydrosphingosine), which is also a common component of animal sphingolipids. Hydroxylation of sphinganine to **phytosphingosine** occurs in both plants and animals, especially within glycolipids. The name comes from the fact that phytosphingosine was first discovered in plants.



Sphingosine (sphingenine)



Sphinganine (dihydrosphingosine)

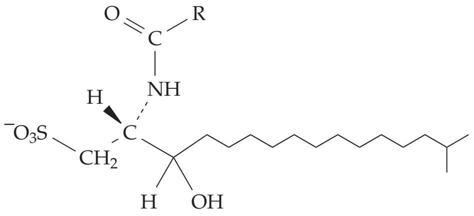


Phytosphingosine

Sphingosine-containing lipids are classified as **sphingophospholipids** (sphingomyelins) and **sphingoglycolipids**. In both cases the sphingosine is combined in amide linkage with a fatty acid to form a **ceramide** (Fig. 8-3) which still contains a free hydroxyl group able to combine with another component. In the sphingomyelins, which were first isolated from human brain by Thudicum in 1884, the additional component is usually phosphocholine (Fig. 8-3). Ceramide aminoethylphosphonates and related glycolipids occur in some invertebrates.[36,37]

The **cerebrosides** are glycosides of ceramide containing galactose or glucose. They are found in relatively large amounts in the brain where monogalactosylceramide predominates. Cerebrosides also occur in other animal tissues and to a lesser extent in plants. Many glycosphingolipids contain oligosaccharides of various sizes. When the oligosaccharide contains one or more residues of sialic acid the compound is known as a **ganglioside**.[38] Sialic acids are never found in plants. However, plants and fungi contain **phytoglycolipids**, which resemble gangliosides. Some contain inositol phosphates as well as sugars.[39–41] The structures of gangliosides may be very complex. Like glycoproteins, these substances are often located at the outer surface of cells where they may act as receptors for toxins, viruses, and hormones (see Section D).[42] Some of them carry attached blood group antigens (Box 4-C).[43,44] The sulfate esters of cerebrosides, known as **sulfatides** (Fig. 8-3), are also important components of membranes.[45]

Gliding bacteria such as *Capnocytophaga* contain sulfonolipids:[46]



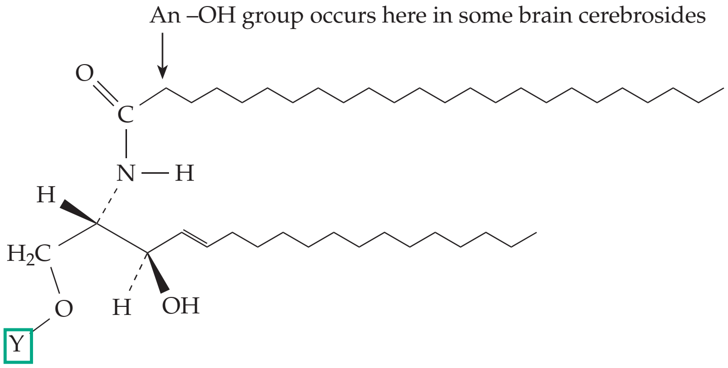
N-acylcapnine

A 1-deoxy-1-sulfonate analog of ceramides from gliding bacteria. R is a long-chain alkyl group.

These are analogs of the ceramides. The sulfonolipids seem to be necessary for the gliding movement of these bacteria across solid surfaces.[47]

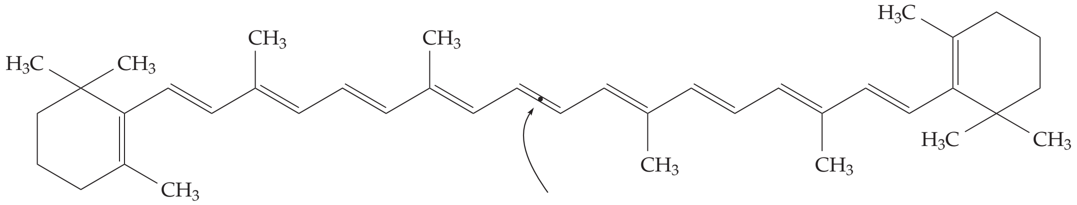
6. Sterols and Other Isoprenoid Lipids

A large group of isoprenoid lipids, including **sterols**, **terpenes**, and **carotenoid compounds**, are often present in membranes or in extracted lipids. Among these are the fat-soluble vitamins A, D, E, and K and the high polymers **rubber** and **gutta-percha**. The phytyl group of chlorophyll (Fig. 23-30), the phytanyl and diphytanyl groups of the lipids of methanogens, and the side chain of the pigment heme *a* (Fig. 16-5) are all related and are all derived from the precursor **prenyl pyrophosphate** (isopentenyl diphosphate). The whole group of compounds are often referred to as **polyprenyl** or by the older designation **isoprenoid**. The major discussion of polyprenyl compounds is found in Chapter 22 but the role of cholesterol and related compounds in membranes and the function of polyprenyl groups as membrane anchors for some proteins will be considered in this chapter.


An –OH group occurs here in some brain cerebrosides

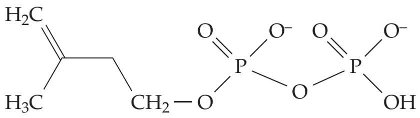
Y = —H	The compounds in which Y = H are **ceramides**. Sphingolipids are often named as ceramide derivatives.
Y = $-O-\underset{O^-}{\overset{O}{\overset{\|}{\underset{\|}{P}}}}-O-CH_2CH_2-N^+(CH_3)_3$ Phosphocholine	**Sphingomyelins**
Y = $-\underset{O^-}{\overset{O}{\overset{\|}{\underset{\|}{P}}}}-CH_2CH_2NH_3^+$	Ceramide aminoethyl phosphonates
Y = D-Galβ or D-Galβ1→4Glc– ↑ The galactose bears a 3-sulfate group in cerebroside sulfatides, e.g., in lactosyl ceramide sulfate	**Cerebrosides** or ceramide mono- and oligosaccharides
Y = GalNAc1→3Gal1→4Gal1→4Glc	Present in sphingolipid of red blood cell membranes
Y = $-O-\underset{O^-}{\overset{O}{\overset{\|}{\underset{\|}{P}}}}-O-$Inositol — Mannose	Present in **phytoglycolipid** of yeast

Figure 8-3 Structures of some sphingolipids.

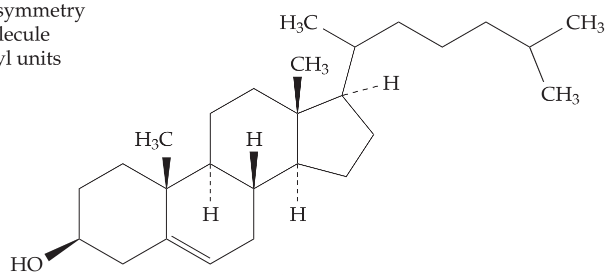


β-Carotene, a yellow pigment of carrots and other plants; converted in the human body to vitamin A.

Notice the center of symmetry in this 40-carbon molecule derived from 8 prenyl units



Prenyl diphosphate (pyrophosphate) (isopentenyl pyrophosphate), precursor of steroid,carotenoid, and other isoprenoid substances



Cholesterol, a sterol

B. Membranes

In Chapter 7 we examined ways in which protein subunits can be stacked to form helices and closed oligomers. Another important arrangement of cell constituents is that of flat sheets or membranes.[48–54] Chemists, physicists, and biologists have mounted a sustained effort to understand these thin but remarkably tough outer surfaces of cells. However, consider the fact that a 7- to10-nm-thick plasma membrane of a cell of 10-µm diameter has less than 1/1000 the thickness of the cell and occupies only 0.5% of the total volume. The technical difficulties in studying such a membrane are great and are compounded by the fact that a cell contains more than one kind of membrane.

Membranes from many sources have been studied. One of these is **myelin**, the multilayered insulation that surrounds the axons of many nerve cells.[55–57] Myelin is derived from the plasma membrane of **Schwann cells** which lie adjacent to many neurons. Schwann cells literally wrap themselves around neuronal axons. Their cytoplasm is squeezed out leaving little but tightly packed membrane layers. Myelin membranes are the most stable known and also have the highest lipid content (80%). Another readily available experimental material is the plasma membrane of the human red bood cell, which can be prepared by osmotic rupture of the cells. The remaining **erythrocyte ghosts** contain ~1% of the dry matter of the cell and may have been studied more than any other membrane. A much investigated specialized membrane is the outer portion of the visual receptor cells known as **rods** (Chapter 23), which contains a closely packed and regular array of flat discs, each one consisting of a pair of membranes. Both membranes and cell walls of many kinds of bacteria have also been investigated.

1. The Structure of Membranes

Membranes consist largely of protein and lipid. The ratio (by weight) of protein to lipid varies from 0.25 in myelin to ~3.0 in bacterial membranes. In membranes of erythrocytes it is about 1.2 and a ratio of about 1.0 may be regarded as typical for animal cells. Small amounts of carbohydrates (<5%) are present, as are traces of RNA (<0.1%).

In 1926, Gorter and Grendel calculated that the erythrocyte ghost contained just enough lipid to form a 3.0- to 4.0-nm-thick layer around the cell. Apparently they reached this correct conclusion only because their measurements of pressures of surface films contained compensating errors.[58] Nevertheless, this information, together with the known propensity of lipids to aggregate in **micelles** in which the hydrocarbon "tails" clustered together and the polar "heads" protruded into the surrounding water,[59] led Danielli and Davsen in 1935 to propose the **lipid bilayer** structure for membranes.[60] Its essential features are indicated in Fig. 8-4. Hydrophobic bonding holds the extended hydrocarbon chains together, while the polar groups of the phospholipid molecules may interact with proteins on the sides of the bilayer. The original proposal assumed an extended β structure for the proteins, which would allow them to coat the bilayer uniformly on both sides. However, this is not correct. Proteins are sometimes embedded in the bilayer, sometimes protrude through it, and sometimes are attached on one surface, most often the cytoplasmic surface of the plasma membrane. These concepts were brought together by Singer and Nicolson in 1972 in the **fluid mosaic model** of membrane structure[61] (Fig. 8-5). The lipid bilayer still provides the basic structure upon which the complex membranes of living organisms are assembled.[52,54,62–65] The term "fluid" refers to the fact

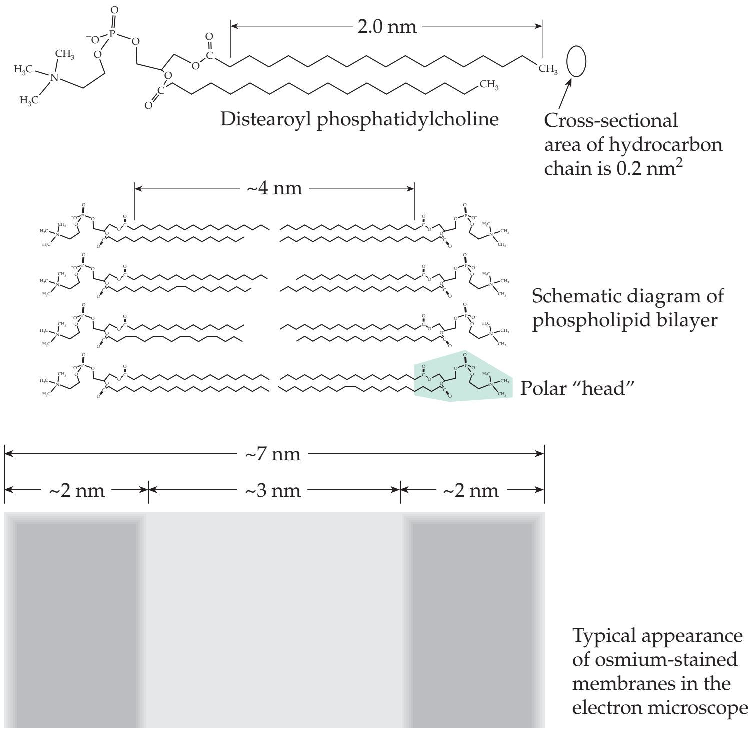

2.0 nm
CH₃
H₃C
N
H₃C
CH₃
O
P
O
O
C
Distearoyl phosphatidylcholine
Cross-sectional area of hydrocarbon chain is 0.2 nm²
~4 nm
Schematic diagram of phospholipid bilayer
Polar "head"
~7 nm
~2 nm
~3 nm
~2 nm
Typical appearance of osmium-stained membranes in the electron microscope

Figure 8-4 Bimolecular lipid layers and membranes. (Top) A molecule of phosphatidylcholine. (Center) Lipid bilayer structure. (Bottom) Bilayer structure as seen by the electron microscope with osmium tetroxide staining.

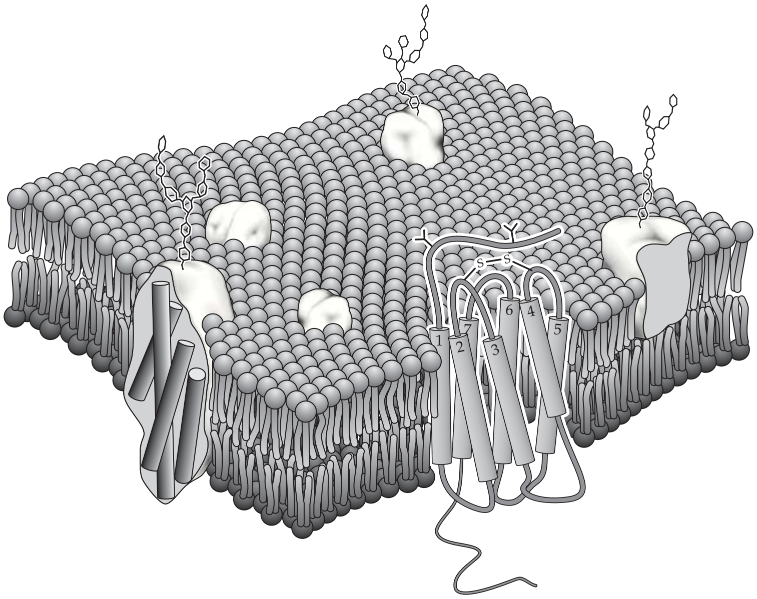


Figure 8-5 The fluid mosaic model of Singer and Nicolson.[61] Some integral membrane proteins, which are shown as irregular solids, are dissolved in the bilayer. Transmembrane proteins protrude from both sides. One of these is pictured as a seven-helix protein, a common type of receptor for hormones and for light absorption by visual pigments. Other proteins adhere to either the outer or the inner surface. Many membrane proteins carry complex oligosaccharide groups which protrude from the outer surface (Chapter 4). A few of these are indicated here as chains of sugar rings.

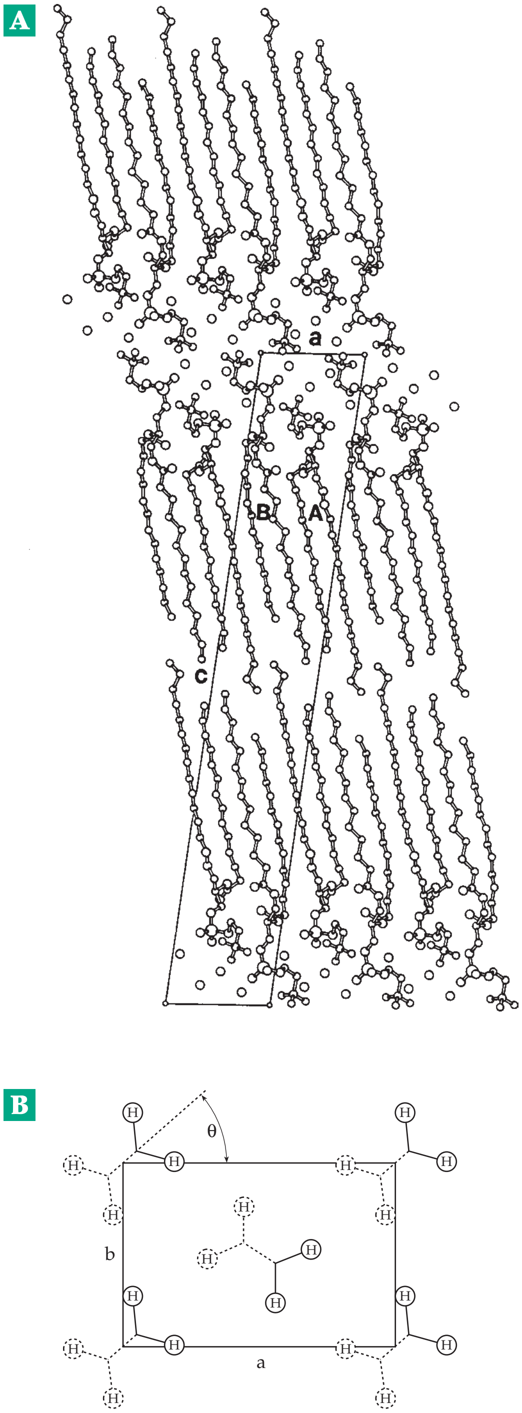

A
a
B
A
c
B
θ
H
b
a

Figure 8-6 (A) Molecular packing of 2,3-dimyristoyl-D-glycero-1-phosphocholine dihydrate. The two molecules in the asymmetric unit are labeled 1 and 2. The position of the water molecules is indicated either by W1–W4 or by small open circles. Hydrogen bonds are represented by dotted lines. From Pascher *et al.*[66] (B) Two-dimensional "orthorhombic" arrangement of hydrocarbon chains in a crystalline alkane. The *a*–*b* plane corresponds to the plane of the bilayer surface; the long axes of the acyl chains project from the page. From Cameron *et al.*[67]

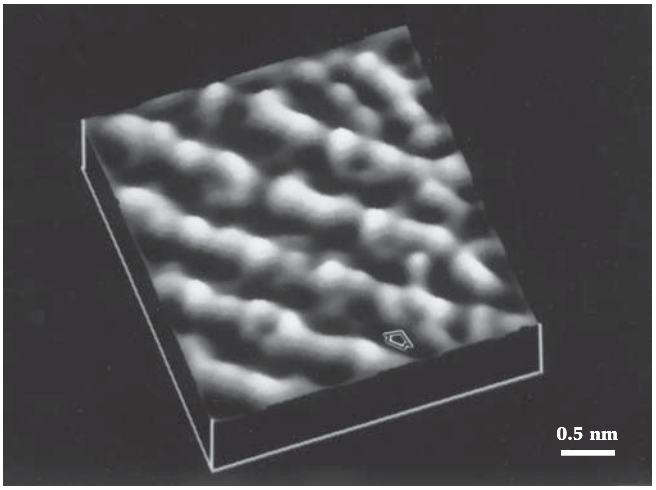


Figure 8-7 Atomic force microscope image of a dimyristoylphosphatidylethanolamine bilayer deposited by the Langmuir–Blodgett technique (see Fig. 8-8) at a specific molecular area of 0.41 nm^2 and a surface pressure of 40 mN/m on a freshly cleaved mica substrate. The images were taken under water. The long, uniformly spaced rows are roughly 0.7–0.9 nm in spacing. The modulation along the rows, with rounded bright spots roughly every 0.5 nm, corresponds to the individual headgroups of the phosphatidylethanolamine molecules. The lattice spacing is identical to that measured by X-ray diffraction at the air–water interface. The area per molecule in the AFM image is ~0.4 nm^2. From Zasadzinski *et al.*[68]

that at temperatures suitable for growth and metabolism the hydrocarbon chains are not rigidly packed in the center of the bilayer but are "molten" (see Section 2). However, at a low enough temperature they become rigid and pack together in a manner similar to that of the chains in crystals of the phosphatidylcholine shown in Fig. 8-6A. These crystals consist of stacked bilayers of thickness 5.5 nm.[66,69] The scanning tunneling and atomic force microscopes have provided direct views of a similar arrangement of side chains in a monomolecular fatty acid layer[68,70,71] (Fig. 8-7). Measurements on multilamellar vesicles of dipalmitoylphosphatidylcholine give bilayer thicknessess from 5.4 nm for dehydrated vesicles to 6.7 nm for the biologically relevant fully hydrated bilayers.[72]

Lipids of membranes. Approximately 1500 different lipids have been identified in the myelin of the central nervous system of humans. About 30 of these are present in substantial amounts.[73] The distribution of the different lipids varies markedly between membranes from different sources (Table 8-3) making generalization difficult. However, phospholipids are apparently always present and, except in chloroplasts, make up from 40% to over 90% of the total lipid (Table 8-3).

Five kinds of phospholipid predominate: phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylglycerols, and sphingomyelin. Usually there are also small amounts of phosphatidylinositol. The major phospholipid in animal cells is phosphatidylcholine, but in bacteria it is phosphatidylethanolamine. The phospholipids of *E. coli* consist of 80% phosphatidylethanolamine, 15% phosphatidylglycerol, and 5% diphosphatidylglycerol (cardiolipin). Significant amounts of cardiolipin are found only in bacteria and in the inner membrane of mitochondria. Sphingomyelin is almost absent from mitochondria, endoplasmic reticulum, or nuclear membranes.

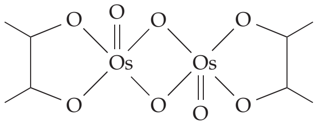
Glycolipids are important constituents of the plasma membranes, of the endoplasmic reticulum, and of chloroplasts. The cerebrosides and their sulfate esters, the sulfatides, are especially abundant in myelin. In plant membranes, the predominant lipids are the galactosyl diglycerides.[29,74] The previously described ether phospholipids (archaebacteria), ceramide aminoethylphosphonate (invertebrates), and sulfolipid (chloroplasts) are also important membrane components.

Cholesterol makes up 17% of myelin and is present in plasma membranes. However, it usually does not occur in bacteria and is present only in trace amounts in mitochondria. Related sterols are present in plant membranes. Esters of sterols occur as transport forms but are not found in membranes. Membrane bilayers, likewise, contain little or no triacylglycerols, the latter being found largely as droplets in the cytoplasm.

Quantitatively minor membrane components with important biological functions include **ubiquinone**, which is present in the inner mitochondrial membrane, and the **tocopherols**. Plant chloroplast membranes contain chlorophyll, carotenes, and other lipid-soluble pigments.

Liquid crystals, liposomes, and artificial membranes. Phospholipids dissolve in water to form true solutions only at very low concentrations (~10^{-10} M for distearoyl phosphatidylcholine). At higher concentrations they exist in **liquid crystalline phases** in which the molecules are partially oriented. Phosphatidylcholines (lecithins) exist almost exclusively in a **lamellar** (smectic) phase in which the molecules form bilayers. In a warm phosphatidylcholine–water mixture containing at least 30% water by weight the phospholipid forms **multilamellar vesicles**, one lipid bilayer surrounding another in an "onion skin" structure. When such vesicles are subjected to ultrasonic vibration they break up, forming some very small vesicles of diameter down to 25 nm which are surrounded by a single bilayer. These unilamellar vesicles are often used for study of the properties of bilayers. Vesicles of both types are often called **liposomes**.[75–77]

When liposomes are stained with osmium tetroxide or potassium permanganate, embedded, and sectioned for electron microscopy, their membranes show a characteristic three layered structure similar to that observed for biological membranes. Two darkly stained lines ~2–2.5 nm thick are separated by a clear space ~2.5–3.5 nm wide in the center. Both myelin and the retinal rod outer segments show closely spaced pairs of such membranes with a combined width of 18 nm. These results seemed to support the original Davsen–Danielli model. However, many questions must be raised about the interpretation of these results. Why does OsO_4 stain only the outer protein layer when it is known to react also with double bonds of hydrocarbon side chains of lipids to form osmate esters which are readily reduced to a diol and osmium?[77a-c,78] Membranes from which most of the lipid has been extracted still stain with OsO_4 to give three-layered electron micrographs. Perhaps little can be concluded from the three-layered appearance. We have learned that it is difficult to determine even the thickness, let alone the complete structure of an object that is only 6–10 nm thick.



An osmate ester formed
from two unsaturated groups

Strong support for the lipid bilayer model comes from the preparation of another type of artificial membrane which can be made from a solution of phosphatidylcholine or of a mixture of phospholipids plus cholesterol in a hydrocarbon solvent. A droplet of solution is placed on a small orifice in a plastic sheet, separating two compartments filled with an aqueous medium (Fig. 8-8). The solution in the orifice quickly drains, just as does a soap bubble, and the resulting film eventually becomes so thin that the bright colors disappear and a **"black membrane"** is formed. Similar membranes, but without a residual content of hydrocarbon solvent, have been created by apposition of two lipid monolayers formed at an air–water interface.[79,80] The thickness of such artificial membranes is thought to be only 6–9 nm. Resilient and self-sealing, the membranes can be stained with OsO_4 to give a typical three-layered pattern.

TABLE 8-3

Estimated Chemical Compositions of Some Membranes

	Percentage of total dry weight of membrane[a]					
Compound	Myelin (bovine)	Retinal rod	Plasma membrane (human erythrocyte)	Mitochondrial membranes	*E. coli*[b,c,d] (inner and outer membranes)	Chloroplasts[e]
Protein	22	59	60	76	75	48
Total lipid	78	41	40	24	25	52[f]
Phosphatidylcholine	7.5	13	6.9	8.8		
Phosphatidylethanolamine	11.7	6.5	6.5	8.4	18	
Phosphatidylserine	7.1	2.5	3.1			
Phosphatidylinositol	0.6	0.4	0.3	0.75		
Phosphatidylglycerol					4	
Cardiolipin[g]		0.4		4.3	3	
Sphingomyelin	6.4	0.5	6.5			
Glycolipid	22.0	9.5	Trace	Trace		23
Cholesterol	17.0	2.0	9.2	0.24		
Total phospholipid	33	27	24	22.5	25	4.7
Phospholipid as a percentage of total lipid	42	66	60	94	>90%	9

---

[a] Dewey, M. M. and Barr, L. (1970) *Curr. Top. Membr. Transp.* **1**, 6.

[b] Kaback, H. R. (1970) *Curr. Top. Membr. Transp.* **1**, 35–99.

[c] Mizushima, S. and Yamada, H. (1975) *Biochim. Biophys. Acta.* **375**, 44–53 .

[d] Yamato, I. Anraku, Y. and Hirosawa, H. (1975) *J. Biochem. (Tokyo)* **77**, 705–718. These investigators found 67% protein, 21% lipids, 10% carbohydrate, and 2% RNA.

[e] Lichtenthaler, H. K. and Park, R. B. (1963) *Nature (London)* **198**, 1070–1072.

[f] About 14% is accounted for by chlorophyll, carotenoids, and quinones.[e]

[g] Diphosphatidylglycerol (Fig. 8-2).

The study of monolayers formed on a water surface has also provided important information. A thin film of an amphiphilic (containing both polar and nonpolar groups) compound such as a fatty acid is prepared. This is done by depositing a small quantity of the compound dissolved in a volatile solvent on a clean aqueous surface between the barriers of a **Langmuir trough** (Fig. 8-8).[81,82] The difference in surface tension (π) across the barriers is measured with a suitable device[81] for different areas of the monolayer, i.e., for different positions of the moveable barrier. The value of π is low for expanded monolayers and falls to nearly zero when the surface is no longer completely covered. The pressure reaches a plateau when a compact monolayer is formed, after which it rises again (Fig. 8-8B). At very high values of π the monolayer collapses (buckles). Both the cross-sectional area per molecule in the monolayer and the collapse pressure can be determined. For typical fatty acids, regardless of chain length, the area covered is only ~0.2 nm^2 per molecule indicating that the fatty acid chains are stacked vertically to the surface in the monolayer. The collapse pressure is higher for longer molecules as a result of the greater number of van der Waals interaction between the chains. **Langmuir–Blodgett layers** are prepared by transferring one or more monolayers onto a smooth solid surface (Fig. 8-7).[82,83]

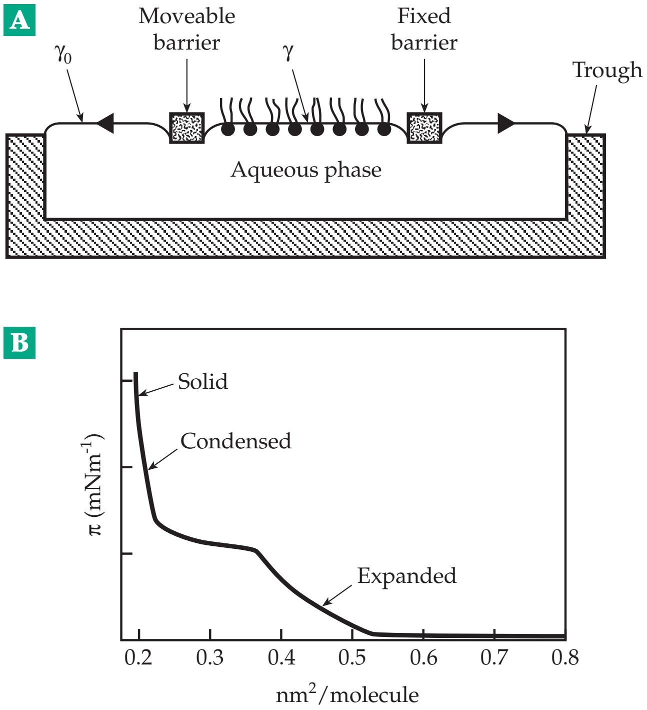
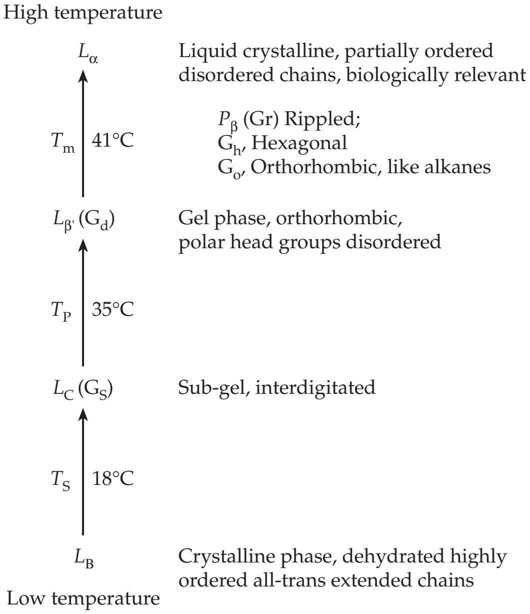


Figure 8-8 (A) The Langmuir–Adam film balance. Tension on the moveable barrier is recorded for different areas of the surface between the barriers. This gives the surface pressure π, which is the difference between the surface tension (γ_o) of a clean aqueous surface and that of a spread monolayer (γ): $\pi = \gamma_o - \gamma$. Courtesy of Jones and Chapman.[81] (B) Surface pressure (π)–area per molecule isotherm for a typical fatty acid (e.g., pentadecanoic acid $C_{14}H_{29}CO_2H$) at the aqueous–air interface. From Knobler.[81a]

Physical properties of membrane lipids. A completely extended C_{18} fatty acid chain as shown in Fig. 8-4 has a length of ~2.0 nm and occupies, either in crystals or in monolayers, when viewed "end-on," an area of ~0.2 nm^2. The hydrocarbon layer in a lipid bilayer containing such chains would have a thickness of about 4.0 nm; that determined by X-ray diffraction for myelin is ~3.5 nm. However, for artificial black membranes the thickness of the hydrocarbon layer can be as little as 3.1 nm when all solvent is removed.[84] These and many other results[85] indicate that the hydrocarbon chains are to some extent folded and that the membrane is expanded over that expected according to the simplest model.

Structure determinations on crystalline alkanes confirm that the chains exist in a completely extended conformation and that adjacent chains often pack together in the orthorhombic arrangement shown in Fig. 8-6B. As the temperature of such crystals is raised a series of solid–solid phase transitions is observed below the melting point of the crystals.[86] These can be detected by changes in the infrared absorption spectrum or by small amounts of heat absorption revealed by **differential scanning calorimetry** (Fig. 8-9). Each new phase permits a greater degree of mobility for the hydrocarbon chains. Thus, at a high enough temperature but below the melting point, the chains are able to rotate freely about their own axes in a so-called **hexagonal phase**. Now the chains are packed in a hexagonal array instead of the orthorhombic array of Fig. 8-6B. At intermediate temperatures, some of the chains may assume nonplanar conformations and changes in the tilt of the hydrocarbon chains (Fig. 8-6) may occur.

Similar phase transitions are observed for bilayers.[88–90] For dipalmitoyl phosphatidylcholine the first detectable **subtransition**[91] is centered at a temperature T_s of 18°C. The second, known as the **pretransition**, occurs at 35°C (T_p). The structure below T_s may be described as rigid or crystalline and that above T_s as a **gel** in which the hydrocarbon side chains twist and turn much more freely but in which the orthorhombic packing is maintained.[86] Above T_p the head groups become disordered. Although the orthorhombic packing of the tails may be maintained, there are several distinct phases,[92,93] including one or more in which the gel is thought to assume a structure analogous to that in the hexagonal phase of hydrocarbons. At 41°C the **main transition** occurs.


High temperature
L_α
Liquid crystalline, partially ordered disordered chains, biologically relevant
T_m 41°C
P_β (Gr) Rippled;
G_h, Hexagonal
G_o, Orthorhombic, like alkanes
$L_{\beta'}$ (G_d)
Gel phase, orthorhombic, polar head groups disordered
T_P 35°C
L_C (G_S)
Sub-gel, interdigitated
T_S 18°C
L_B
Crystalline phase, dehydrated highly ordered all-trans extended chains
Low temperature

This is a sharper transition with a well-defined melting temperature designated T_m. Above T_m the lipid is in the lamellar liquid crystalline or L_α state. The bilayer continues to hold together, but the fatty acid chains have melted and are now free to rotate and undergo twisting movements more freely than at lower temperatures (Fig. 8-11). The main transition is highly, but not completely, cooperative. Thus, the melting of the membrane occurs over a range of several degrees. The presence in biological membranes of a variety of different components containing a variety of fatty acid chains leads to a broadening of the melting range.

The behavior of bilayers is strongly influenced by the lipid composition. Phospholipids containing saturated, long-chain fatty acids have high transition temperatures. The presence of unsaturated fatty acyl groups with *cis* double bonds in membrane lipids encourages folding of the hydrocarbon chains and lowers T_m. Even a single double bond lowers T_m, the decrease being greatest when the double bond is near the center of the chain.[94–96] While T_m for dipalmitoyl phosphatidylcholine is 41°C, that of 1,2-dipalmitoyl phosphatidyl-*sn*-glycerol, which lacks the phosphocholine head group, is 70°C. This falls to 11.6°C for the polyunsaturated 1-stereoyl-2-linoleoyl-*sn*-glycerol, whose melting curve is shown in Fig. 8-9.[87] This lipid also shows a complex phase behavior and a melting point for the stable, crystalline β′ phase higher than that of the α phase.

Inclusion of other molecules of irregular shape within membranes also lowers T_m. However, a molecule of cholesterol can pack into a bilayer with a cross-sectional area of 0.39 nm^2, just equal to that of two hydrocarbon chains.[49] It tends to harden membranes above T_m but increases mobility of hydrocarbon chains below T_m.[97–100] A complex of cholesterol and phosphatidylcholine may form a separate phase within the membrane.[101,102] The ether-linked plasmalogens may account for over 30% of the phosphoglycerides of the white matter of the brain and of heart and ether linked phospholipids are the major lipids of many anaerobic bacteria.[103] Their T_m values are a few degrees higher than those of the corresponding acyl phospholipids.[104]

Between the pretransition temperature and T_m solid and liquid regions may coexist within a bilayer.[101] The term **lateral phase separation** has been applied to this phenomenon.[105,106] Since changes in the equilibrium between solid and liquid can be induced readily, e.g., by changes in the ionic environment surrounding the bilayer, lateral phase separation may be of significance in such phenomena as nerve conduction.[107]

The phase transitions in bilayers can be recognized in many ways. Differential scanning colorimetry has already been mentioned. Another approach is to measure the spacing between molecules by X-ray diffraction. The cross-sectional area occupied by a phospholipid in a bilayer is always greater than the 0.40 nm^2 expected for closest packing of a pair of extended hydrocarbon chains.[39,85]

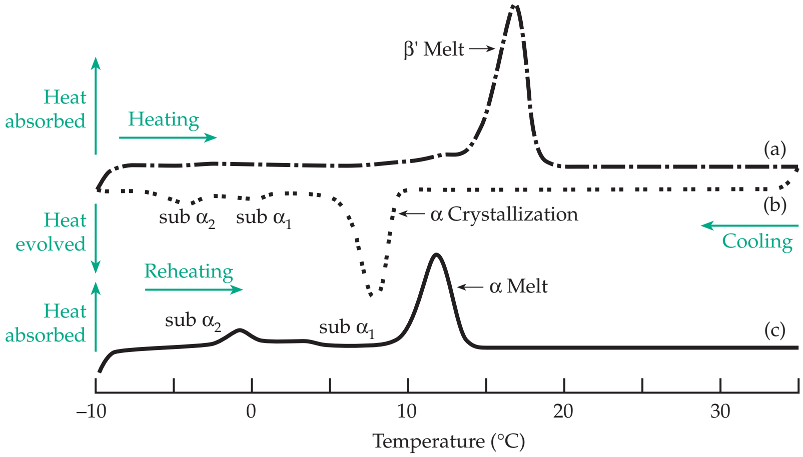


Figure 8-9 Differential scanning calorimetric curves for 1-stearoyl-2-linoleoyl-*sn*-glycerol. (A) Crystals of the compound grown from a hexane solution were heated from –10° to 35°C at a rate of 5°C per minute and the heat absorbed by the sample was recorded. (B) The molten lipid was cooled from 35° to –10°C at a rate of 5° per minute and the heat evolved was recorded as the lipid crystallized in the α phase and was then transformed through two sub-α phases. (C) The solid was reheated. From Di and Small.[87] Courtesy of Donald M. Small.

Below T_m the spacing between chains is about 0.42 nm corresponding to close packing of the fatty acid chains in a hexagonal array with an area per phospholipid of 0.41 nm^2. As the temperature is raised above T_m the spacing increases[85] to give an average area per phospholipid of 0.64–0.73 nm^2. Another technique (Box 8-C) is to study a **spin label** by EPR while yet another is to observe the fluorescence of a **polarity-dependent fluorescence probe** such as *N*-phenylnaphthylamine or other fluorescent probes[108] (see Chapter 23). The compound is incorporated into the membrane and undergoes changes in the intensity of its fluorescence when the state of the membrane is altered.

A variety of NMR techniques are being applied[109–113] both to liposomes and to natural membranes.[111,114] Incorporation of ^{13}C or ^{2}H into various positions in the hydrocarbon chains has allowed measurements of the relative degree of mobility of the chains at different depths in the bilayer (Fig. 8-10).[109,115–117] The results are in agreement with statistical mechanical predictions that configurational freedom increases with depth toward the midplane of the bilayer. Separation of a bilayer into two or more phases can be observed using ^{2}H- or ^{31}P- NMR.[118–120] The orientation and dynamic behavior of various head groups has been explored,[110,121] as have effects of mixing into the bilayer other lipids such as glycosphingolipids[122] and cholesterol.[123,124] Crystalline phospholipids are being investigated by solid-state NMR.[125]

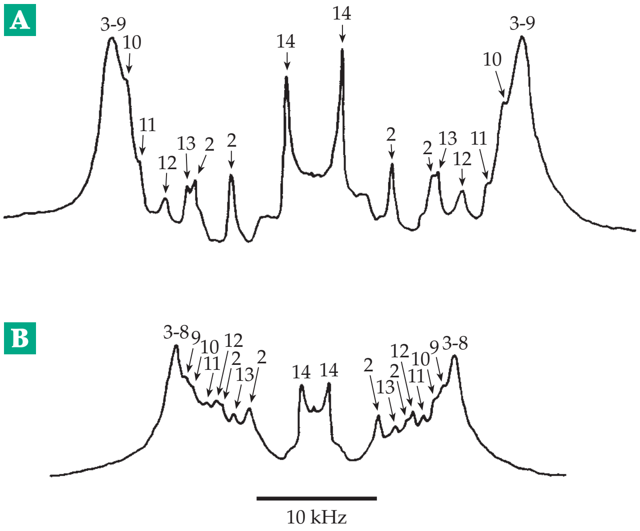
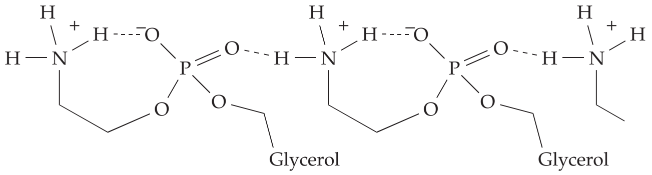


Figure 8-10 ^{2}H NMR spectra of dimyristoyl phosphatidylcholine-d_{27}/water in lamellar phases at 40°C. One chain of the phosphatidylcholine is fully deuterated, containing 27 atoms of ^{2}H. The mole ratios of water to lipid were 5.0 in (A) and 25.0 in (B). The average interfacial areas per alkyl chain as measured by X-ray diffraction were 0.252 nm^2 for (A) and 0.313 nm^2 for (B). ^{2}H NMR spectra are presented as "powder patterns" because the lipid molecules are randomly oriented in the magnetic field of the spectrometer as if in a powder. This gives rise to pairs of peaks symmetrically located on both sides of the origin. The separation distances are a measure of the quadrupole splitting of the NMR absorption line caused by the ^{2}H nucleus. The various splittings of the resonances of the 13 $-CH_2-$ and one $-CH_3$ groups reflect differences in mobility.[109] The peaks have been assigned tentatively as indicated. From Boden, Jones, and Sixl.[115] Courtesy of N. Boden.

Fourier transform infrared spectroscopy[126,127] also provides information about conformation of both hydrocarbon chains and head groups. EPR spectroscopy (Box 8-C) with doxyl probes on carbon atoms at different depths within the bilayer has also been employed.[128]

In recent years **molecular dynamics simulations** have been used to predict behavior of membranes. As is indicated in Fig. 8-11, the molten interior of the liquid crystalline L_α state is portrayed clearly.[129–131] In the gel state the hydrocarbon chains maintain a closer packing and undergo coordinated movement.[88] It is difficult to know how realistic the simulations are. To calibrate the method efforts are made to correctly predict a series of known properties such as density and area per lipid (0.61 nm).[130]

Functions of phospholipid head groups. The dipolar ionic head groups of phosphatidylcholine and phosphatidylethanolamine occupy about the same cross-sectional area as the two hydrocarbon tails. Thus, they are in rather close contact with each other. In crystals chains of hydrogen-bonded atoms may be formed. In phosphatidylethanolamine the phosphate and $-NH_3^+$ ions may alternate in these chains.[132]



In phosphatidylcholine, in which the nitrogen is surrounded by methyl groups and cannot form this kind of chain, water molecules bridge between the phosphates but the positive charges still interact with the adjacent negative charges.

The chains of hydrogen bonds between the head groups of phosphatidylethanolamine help to stabilize the bilayer and are apparently responsible for the elevation of T_m by 10–30° above that observed for phosphatidylcholine.[132] In contrast, the negatively charged carboxyl groups of phosphatidylserine make the membrane less stable. The melting point is increased if the pH is lowered, protonating these groups. Their presence also makes the membrane sensitive to the concentration of cations.[133] The same is true of phosphatidylglycerol, whose head group contains a negatively charged phosphate without an attached counterion. Addition of calcium ions increases T_m greatly and causes either phosphatidyglycerol or phosphatidylserine to form a separate phase with a more crystalline-like packing of the hydrocarbon side chains.[134] Hydrogen bonding between head groups also occurs with glycolipids.[135]

Non-bilayer structures of phospholipids. Under appropriate conditions some aqueous phospholipids can exist in non-bilayer phases, a fact that may be of considerable biological importance.[119,136,137] In the presence of Ca^{2+} some pure phospholipids can be converted to the **inverted hexagonal** or H_{II} phase (Fig. 8-12).[136,138–140] In this phase the phospholipid heads are clustered together in cylindrical "inverted" micelles which pack in a hexagonal array. The ease with which this transition can occur is increased by the presence of small amounts of diacylglycerols or lysolecithins.[141] Some lipids, such as the galactosyldiacylglycerol of chloroplasts, do not form bilayers but prefer the hexagonal phase structure.[29,32a] This is thought to be a result of very high curvature of a bilayer that arises from the sizes and packing of their head groups. Another phase, even though it is liquid, has a three-dimensional cubic symmetry.[142–144a] It apparently consists of a complex arrangement of polyhedral bilayer surfaces with interpenetrating water channels between them.[143]

Membrane fluidity and life. In agreement with the known behavior of bilayers, the lipids of most membranes in all organisms are partially liquid at those temperatures suitable for life. Organisms have developed at least three distinct means of ensuring that membrane lipids remain liquid.[145] (1) In our bodies (as well as in *E. coli*) the unsaturated fatty acids that are present lower the melting point. Mutants of *E. coli* that are unable to synthesize unsaturated fatty acids cannot live unless these materials are supplied in the medium.[146] (2) In *Bacillus subtilis*, which contains no unsaturated fatty acids when grown at 37°C, and in other gram-positive bacteria, more than 70% of membrane

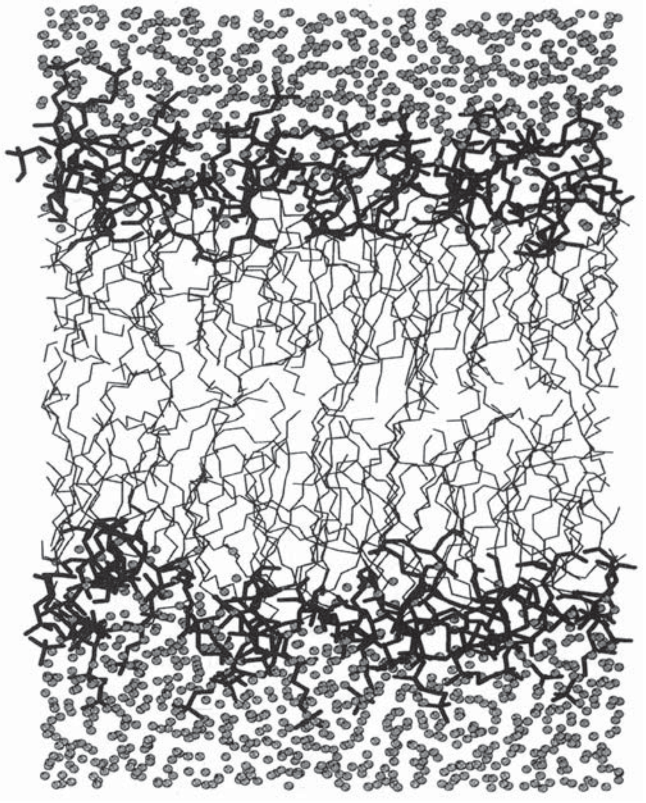


Figure 8-11 Results of simulated motion in a lipid bilayer consisting of 64 molecules of dipalmitoylphosphatidylcholine and 23 water molecules per lipid at a pressure of 2 atm and 50°C. The view is that observed after 500 ps of simulation. Bold lines represent the head group and glycerol parts of the structures and the thin lines the hydrocarbon chains. The gray spheres represent water molecules. From Berger, Edholm, and Jähnig.[130] Courtesy of Dr. Olle Edholm.

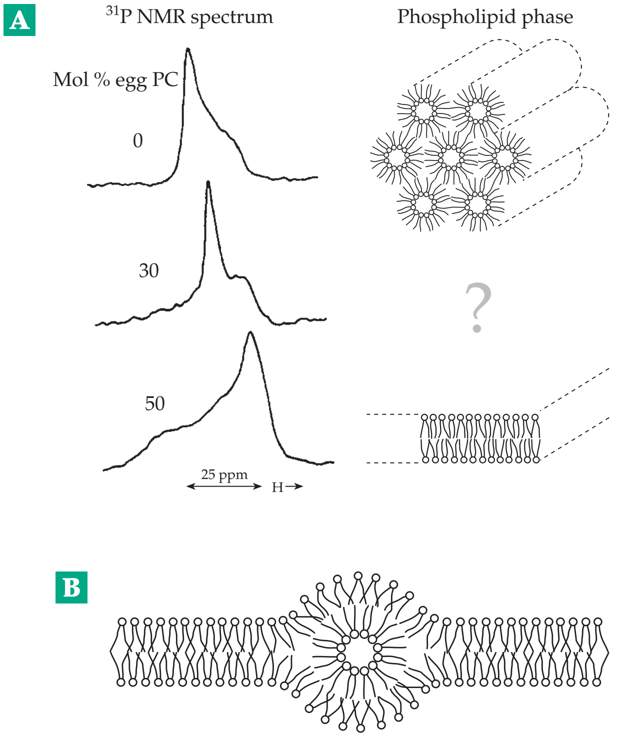


Figure 8-12 (A) ^{31}P NMR spectra of different phospholipid phases. Hydrated soya phosphatidylethanolamine adopts the hexagonal H_{II} phase at 30°C. In the presence of 50 mol% of egg phosphatidylcholine only the bilayer phase is observed. At intermediate (30%) phosphatidylcholine concentrations an isotropic component appears in the spectrum. (B) Inverted micelles proposed to explain "lipidic particles" seen in freeze fracture micrographs of bilayer mixture of phospholipids, e.g., of phosphatidylethanolanine + phosphatidylcholine + cholesterol. From de Kruijft *et al.*[119] Courtesy of B. de Kruijft.

BOX 8-C ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTRA AND "SPIN LABELS"

Unpaired electrons have magnetic moments and are therefore suitable objects for magnetic resonance spectroscopy. The technique is similar to NMR spectroscopy, but microwave frequencies of ~10^{10} Hz are employed, the energies being ~100 times greater than those used in NMR.[a,b] Unpaired electrons are found in organic free radicals and in certain transition metal ions, both of which are important to many enzymatic processes. Furthermore, **spin labels** in the form of stable organic free radicals, can be attached to macromolecules at many different points.[a,c–e] Coupling of such artificially introduced unpaired electrons with the magnetic moments of other unpaired electrons or of magnetic nuclei can often be observed by EPR techniques.

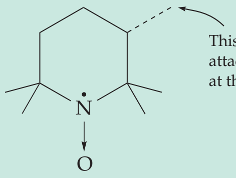
The conditions for absorption of energy in the EPR spectrometer are given by the equation

$$h\nu = g\beta H_o$$

which is identical in form to that for NMR spectroscopy. Here H_o is the external magnetic field strength and β is a constant called the Bohr magneton. The value of g, the **spectroscopic splitting factor**, is one of the major characteristics needed to describe an EPR spectrum. The value of g is exactly 2.000 for a free electron but may be somewhat different in radicals and substantially different in transition metals. One factor that causes g values to vary with environment is **spin-orbit coupling** which arises because the p and d orbitals of atoms have directional character. For the same reason g sometimes has three discrete values for the three different directions (g value anisotropy). The g value parallel to the direction of H_o ($g_{||}$) often differs from that in the perpendicular direction ($g_\perp$). Both values can be ascertained experimentally.

A second feature of an EPR spectrum is **hyperfine structure** which results from coupling of the magnetic moment of the unpaired electron with nuclear spins. The coupling is analogous to the spin–spin coupling of NMR (Chapter 3). The hyperfine splitting constant A, like the coupling constant J of NMR spectroscopy, is given in Hertz. Splitting may be caused by a magnetic atomic nucleus about which the electron is moving or by some adjacent nucleus or other unpaired electron. Sometimes important chemical conclusions can be drawn from the presence or absence of splitting. Thus, the EPR spectrum of a metal ion in a complex will be split by nuclei in the ligand only if covalent bonding takes place.

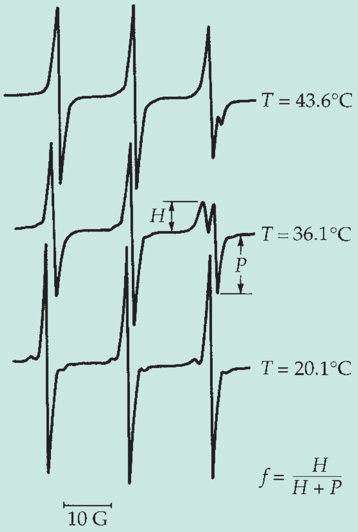
It is customary in EPR spectroscopy to plot the first derivative of the absorption rather than the absorption itself. Thus, for the paramagnetic nitroxide 2,2,6,6-tetramethylpiperidine-1-oxyl the EPR spectrum consists of three equally spaced bands whose peaks are marked at the points where the steep lines in the middle of the first derivative plots cross the horizontal axis.



This "spin" label is often attached by covalent linkage at this point to larger molecules.

2,2,6,6-Tetramethylpiperidine-1-oxyl

Coupling with the ^{14}N nuclear spin causes splitting into three lines as shown in the accompanying figure.

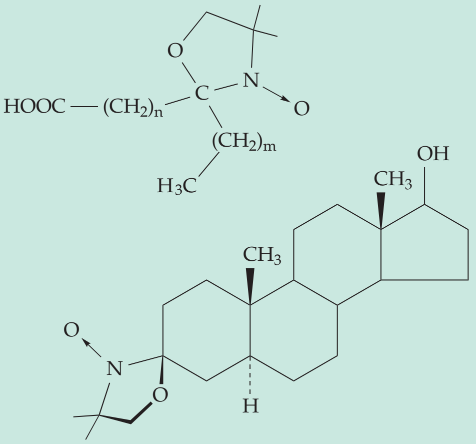


EPR spectrum of tetramethylpiperidine-1-oxyl dissolved in an aqueous dispersion of phospholipids. (Top) above the main bilayer transition temperature T_t; (center) between T_t and pretransition temperature; (bottom) below pretransition temperature. From Shimshick and McConnell.[f]

This nitroxide is more soluble in liquid regions of bilayers than it is in solid regions. As bilayers are warmed in the EPR spectrometer, the solubility of this spin-labeled compound in the lipid can be followed (see figure). The lower of the three spectra approximates that of the spin label in water alone, while the others are composite spectra for which part of the spin label has dissolved in the phospholipid bilayers.

BOX 8-C (continued)

Since frequencies for EPR spectroscopy are ~100 times higher than those for NMR spectroscopy, correlation times (Chapter 3) must be less than $\sim 10^{-9}$ s if sharp spectra are to be obtained. Sharp bands may sometimes be obtained for solutions, but samples are often frozen to eliminate molecular motion; spectra are taken at very low temperatures. For spin labels in lipid bilayers, both the bandwidth and shape are sensitively dependent upon molecular motion, which may be either random or restricted. Computer simulations are often used to match observed band shapes under varying conditions with those predicted by theories of motional broadening of lines. Among the many spin-labeled compounds that have been incorporated into lipid bilayers are the following:



Much of the interpretation of the observed changes in EPR spectra of spin labels is empirical. For example, the spectra in the accompanying figure can be interpreted to indicate that the spin label dissolves in the lipid to a greater extent at higher temperatures. The ratio *f* (defined in the figure) is an empirical quantity whose change can be monitored as a function of temperature. Plots of *f* vs *T* have been used to identify transition and pretransition temperatures in bilayers.[f]

EPR spectroscopy is used widely in the study of proteins and of lipid–protein interactions.[c] It has often been used to estimate distances between spin labels and bound paramagnetic metal ions.[g] A high-resolution EPR technique that detects NMR transitions by a simultaneously irradiated EPR transition is known as electron-nuclear double resonance (ENDOR).[h]

---

[a] Berliner, L. J., and Reuben, J., eds. (1989) *Spinlabeling. Theory and Applications*, Vol. 8, Plenum, New York ((Biological Magnetic Resonance Series)

[b] Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, Freeman, San Francisco, California (pp. 525–536, 1352–1362)

[c] Marsh, D. (1983) *Trends Biochem. Sci.* **8**, 330–333

[d] Esmann, M., Hideg, K., and Marsh, D. (1988) *Biochemistry* **27**, 3913–3917

[e] Millhauser, G. L. (1992) *Trends Biochem. Sci.* **17**, 448–452

[f] Shimshick, E. J., and McConnell, H. M. (1973) *Biochemistry* **12**, 2351–2360

[g] Voss, J., Salwinski, L., Kaback, H. R., and Hubbell, W. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12295–12299

[h] Lubitz, W., and Babcock, G. T. (1987) *Trends Biochem. Sci.* **12**, 96–100

fatty acids contain methyl branches (Chapter 21)[147,148] which can decrease the melting point and increase the monolayer surface area by a factor of as much as 1.5. (3) Yet another mechanism for lowering the melting point of fats is the incorporation of **cyclopropane-containing fatty acids** (Chapter 21).

On the other hand, as we have already seen, cholesterol tends to reduce the mobility of molecules in membranes and causes phospholipid molecules to occupy a smaller area than they would otherwise. Myelin is especially rich in long-chain sphingolipids and cholesterol, both of which tend to stabilize artificial bilayers. Within our bodies, the bilayers of myelin tend to be almost solid. Bilayers of some gram-positive bacteria growing at elevated temperatures are stiffened by biosynthesis of bifunctional fatty acids with covalently joined “tails” that link the opposite sides of a bilayer.[149]

Why must membrane lipids be mobile? One reason is probably to be found in the participation of membranes in many vital transport processes. Biological membranes have a relatively high permeability to neutral molecules (including H_2O),[64,150] and it has been suggested that above T_m fatty acid chains are free to rotate by 120° around single bonds from trans to gauche conformations. When such rotation occurs about adjacent, or nearly adjacent single bonds, **kinks** are formed. If a kink originates near the bilayer surface, as will usually be the case, a small molecule may jump into the void created. Since the kink can easily migrate through the bilayer, a small molecule may be carried through with it.[151,152] The same factors may assist larger protein molecules which function in membrane transport. They probably also account for the substantial degree of **hydration** of bilayers which involves both the polar head groups and water diffusing through the nonpolar interior.[153]

Not only can molecules diffuse through membranes but also membrane lipids and proteins can move with respect to neighboring molecules. The rates of **lateral**

diffusion of lipids in bilayers and of antigenic proteins on cell surfaces are rapid.[80,154] If diffusion of phospholipids is assumed to occur by a pairwise exchange of neighboring molecules, the frequency of such exchanges can be estimated[155] as ~10^7 s^{-1}. However proteins may meet many obstacles to free diffusion.[156] Lateral diffusion is often measured by the technique of **fluorescence recovery after photobleaching**. One small spot in a bilayer that contains a dye attached to a lipid or a protein is bleached by a laser beam. Lateral diffusion of nearby unbleached molecules into the bleached spot can then be observed.[80] Lateral diffusion can also be observed by NMR spectroscopy[157] and by single-particle tracking.[158,159] In addition to diffusion there may often be a flow of membrane constituents in directions dictated by metabolism.[160] Although lateral diffusion is fast a "flip-flop" transfer of a phospholipid from one side of the bilayer to the other may require many seconds.[161] However, a sudden increase in the calcium ion concentration, an important intramolecular signal (Chapter 11), activates a "scramblase" protein which promotes a rapid transbilayer movement of phospholipids.[162]

Electrical properties of membranes. Biological membranes serve as barriers to the passage of ions and polar molecules, a fact that is reflected in their high electrical resistance and capacitance. The electrical resistance is usually 10^3 ohms cm^{-2}, while the capacitance is 0.5–1.5 microfarad (μF) cm^{-2}. The corresponding values for artificial membranes are ~10^7 ohms cm^{-2} and 0.6 – 0.9 μF cm^{-2}. The lower resistance of biological membranes must result from the presence of proteins and other ion-carrying substances or of pores in the membranes. The capacitance values for the two types of membrane are very close to those expected for a bilayer with a thickness of ~2.5 nm and a dielectric constant of 2.[54,84,163] The electrical potential gradient is steep.

Outer cell surfaces usually carry a net negative charge, the result of phosphate groups of phospholipids, of carboxylate groups on proteins, and of sialic acids attached to glycoproteins. This negatively charged surface layer attracts ions of the opposite charge (counterions), including protons, and repels those of the same charge. The result is development of a diffuse **electrical double layer** consisting of the fixed negative charges on the surface and a positive **ionic atmosphere** extending into the solution for a distance that depends upon the ionic strength.[164–166] This ionic atmosphere is analogous to that postulated by the Debye–Hückel theory (Chapter 6). At the physiological ionic strength of 0.145 M the thickness of the double layer, taken as the distance at which the electrical potential falls to a certain fraction of that at the cell surface,[164,165] is about 0.8 nm. However, the double-layer thickness increases to about three times this value at an ionic strength of 10^{-3} M and to still greater distances at lower ionic strengths.

The net surface charge of a cell and the associated electrical double layer are important in interactions between cells and may influence the development of extracellular structure such as basement membranes. The net negative charge on cells also gives rise to an experimentally measurable electrophoretic mobility.

A characteristic of living cells is the maintenance of **ionic gradients** across the plasma membrane. Thus almost all cells accumulate K^+, even "pumping" it from very dilute external solutions. Cells also exclude sodium, pumping it out from the cytoplasm by mechanisms considered in Section C,2. If a microelectrode is inserted through a cell membrane and the potential difference is measured between the inside and outside of the cell, a **resting potential** which, in nerve cells, may be as high as 90 mV is observed. The origin of the potential appears to lie in the concentration differences of ions. From the value of ΔG for dilution of an ion (Eq. 6-25) and the relationship between ΔG and electrode potential (Eq. 6-63), the Nernst equation (Eq. 8-2) can be derived. According to this equation, which applies to a single ion for which the membrane is permeable,

$$E_m = \frac{RT}{nF} \ln\left(\frac{c_1}{c_2}\right) = \frac{0.059}{n} \log\left(\frac{c_1}{c_2}\right) \quad \text{at } 25°\text{C} \tag{8-2}$$

a 10-fold concentration difference across the membrane for a monovalent ion ($n = 1$) would lead to a 59-mV membrane potential, E_m. Since membranes are relatively impermeable to sodium ions, it is generally conceded that for many membranes the origin of the membrane potential lies mainly with the potassium ion concentration difference which is maintained by the Na^+, K^+-ATPase (Section C). A more complete equation takes account of K^+, Na^+, and Cl^- together with their respective permeabilities.[167–169] Note also that Eq. 6-64 is also often called the Nernst equation.[170]

Protons are also pumped across cytoplasmic and inner mitochondrial membranes, a topic of Chapter 18. The flow of protons from inside to outside also contributes to the membrane potential. The positive charges of H^+, K^+, and other cations associated with the external membrane surface are balanced by the negative charges of protein molecules as well as Cl^- and phosphate anions that are in or near to the inner surface of the membranes.

Another possibility for proton flow has intrigued biophysicists for years. Membranes often display a substantial electrical conductivity in a lateral direction along the membrane surface.[171–173] Electrical conduction may involve movement of protons along hydrogen-bonded lines, e.g., involving ethanolamine head groups or phosphate groups and bridging water as previously discussed (see also Eq. 9-96). Alternatively, conduction may depend upon membrane-associated proteins.[174] This lateral proton conduction may be important to many proton-driven membrane processess, such as rotation of bacterial flagella, ATP synthesis, and pumping of ions (Chapter 18).

The two sides of a membrane. Many observations indicate great differences between the inside and outside of the membranes that surround cells.[51,175,176] Bretscher and Raff[62] observed that, among the phospholipids of the erythrocyte membrane, phosphatidylcholine predominates in many mammals but is replaced by sphingomyelin in ruminants. Sheep erythrocytes are resistant to cobra venom phospholipase A, which is known to remove the fatty acid from the central position on the glycerol of phosphatidylcholine, causing lysis of the cells. The resistance of sheep erythrocytes suggested that the sphingomyelin is on the outside of the membrane while the phosphatidylethanolamine and other phospholipids are inside. By inference, phosphatidylcholine is also largely on the outside of plasma membranes. Supporting this conclusion is the observation that most of the reactive amino groups of phosphatidylethanolamine and phosphatidylserine are found on the inside (cytoplasmic) surfaces.[177] Since the total content of phosphatidylcholine and sphingomyelin often exceeds that of phosphatidylethanolamine and phosphatidylserine, the bilayer would be incomplete on the inside of the membrane were it not for the presence of proteins, which contribute more to the inside than to the outside surface.

Glycolipids are usually on the *outside* of plasma membranes with the attached sugar chains projecting into the surrounding water. An important generalization is that *sugar groups attached either to lipids or to proteins tend to be on outer cell surfaces or on materials that are being exported from cells*. An exception is found in the abundant galactolipids of chloroplasts.

2. Membrane Proteins

The many proteins present within or attached to membranes have a variety of functions. Some are obviously structural, tying other proteins to a membrane or providing a base for projecting fimbriae, flagella, and other appendages. Some proteins of the outer surface act as anchoring points for macromolecules that lie between cells. The inner surfaces of membranes are attached to the cytoskeleton (Chapter 7). Many of the proteins embedded in membranes control the passage of materials across membranes. Others serve as receptors that sense the presence of specific compounds or of light. Membranes may also contain foreign proteins such as subunits of virus coats. Proteins that are deeply embedded in membranes are referred to as **intrinsic** or **integral membrane proteins**. Proteins that are more loosely associated with the membrane, principally at the inner surface, are called **peripheral**.[61]

Integral membrane proteins. Membrane proteins are hard to crystallize[178] and precise structures are known for only a few of them.[179–181] A large fraction of all of the integral membrane proteins contain one or more **membrane-spanning helices** with loops of peptide chain between them. Folded domains in the cytoplasm or on the external membrane surface may also be present. The best-known structure of a transmembrane protein is that of the 248-residue bacteriorhodopsin. It consists of seven helical segments that span the plasma membrane (Fig. 23-45) and serves as a light-activated proton pump. Other proteins with similar structures act as hormone receptors in eukaryotic membranes. A seven-helix protein embedded in a membrane is depicted in Fig. 8-5 and also, in more detail, in Fig. 11-6.

The most hydrophobic integral membrane proteins can be extracted into organic solvents such as mixtures of chloroform and methanol. One such **proteolipid protein**, the 23.5-kDa lipophilin, accounts for over half the protein of myelin.[57,182] The purified protein from rat brain contains 66% of nonpolar amino acids and six molecules of covalently bound palmitic acid and other fatty acids per peptide chain in thioester linkage to cysteine side chains. This protein evidently has four transmembrane helical segments with the six fatty acid chains incorporated into the membrane bilayer. It also has cytoplasmic and extracellular loops, one of which binds inositol hexakisphosphate (Ins *P*-6). (Fig. 11-9).[183] The myelin proteolipid is an essential component of the myelin sheath and defects in this protein are associated with some demyelinating diseases[57] which are discussed in Chapter 30.

There are many known topologies for helix-bundle membrane proteins with the number of membrane-spanning helices ranging from 1 to 14 or more e.g., see Fig. 8-23.[184–186] A topology can often be predicted using suitable computer programs.[187–191] A first step is to identify all sequences of 20 or more residues that could form a helix sufficiently hydrophobic to allow good nonpolar interactions with the bilayer core. Sequences that can form **amphipathic** (amphiphilic) helices must also be considered because two or more of these can pack together in a membrane with their hydrophilic sides together, sometimes forming pores (see Section C,1).[192]

Predictions of membrane protein structure are in part based on the **positive-inside rule** which states that positively charged lysine and arginine residues will not pass through a membrane but will remain on the negatively charged cytoplasmic surface. Often, N-terminal parts of a protein will pass through a membrane and will contain glutamate or aspartate residues whose side chains carry negative charges. These tend to remain outside of the membrane where they are attracted to the positive charges on the membrane outer surface. Positively charged residues may interact with phosphate groups of phospholipids and tyrosine and tryptophan side chains may interact with the carbonyl groups of the ester linkages.[192a,b]

After probable transmembrane helical regions have been identified a residue-by-residue attempt can be made to identify cytosolic and extracellular loops. Chemical reactivities of the naturally occurring side chains can also be examined. Residues within the helix bundle will be protected. Mutations can be prepared systematically by "**scanning mutagenesis**." For example, presumed extracellular loops in the erythrocyte band 3 protein (next section) have been mutated by introduction of *N*-glycosylation acceptor sites (Asn-X-Ser / Thr). If the sequence is at least 12–14 residues away from transmembrane sequences it will probably be glycosylated in a suitable laboratory test system.[193] Within a transmembrane domain amino acid residues can be systematically replaced with alanine ("alanine scanning mutagenesis") or with cysteine and effects on the protein can be observed. Substitution with cysteine allows another possibility: Two cysteines on adjacent transmembrane helices may be linked as disulfides if they are close enough.[194,195] Discovery of such neighboring pairs can be very valuable in attempting to establish relationships of one helix to another. A cloned gene can also be split into two pieces prior to crosslinking of cysteine side chains. This may facilitate mapping of tertiary interactions within transmembrane proteins.[196] Molecular dynamics methods for modeling helix bundle proteins are also being developed.[197]

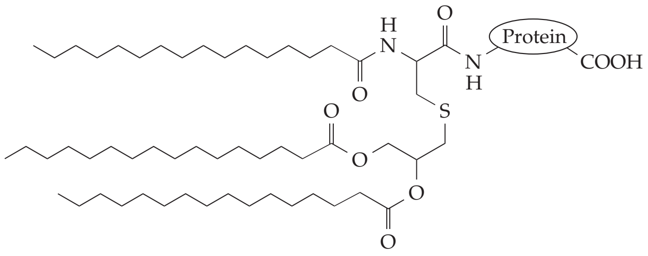
Not all integral membrane proteins have a helix bundle structure. Some of the simplest transmembrane proteins are the subunits of bacterial viruses and pili (Figs. 7-7 and 7-9) The 7-nm α-helical rods of phage M13 have a 20-residue hydrophobic section (residues 25-46) which is preceded by a negatively charged sequence that appears to form a short *amphipathic* helix that lies on the external surface of the membrane and helps to anchor the subunit. A positively charged cluster at the opposite end of the hydrophobic helical region remains in the cytoplasm.[180,186,198] The **porins**, which form pores in outer membranes of bacteria and ribosomes, are large 16-strand β cylinders (Section C,1).[179,180] It has been suggested that this different basic architecture may be related to the fact that the porin polypeptides must be exported through the inner membrane before being refolded in the periplasm.[199] Keep in mind that a β strand of nine residues can span the 33-nm bilayer core just as well as can a 22-residue helix. Beta structures as well as shorter helices may well be present in proteins that also contain membrane-spanning helices.

A third important structural pattern involves extensive use of amphipathic helices that lie partially embedded in a membrane surface. For example, the blood lipoproteins are lipid particles partially coated by amphipathic helices (Chapter 21).[200,201]

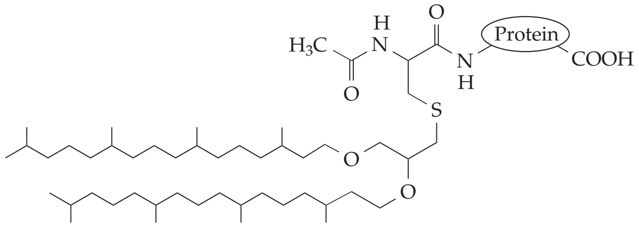
Anchors for proteins. Proteins with membrane-spanning sequences are usually anchored into the membranes with the help of many polar side chains, often including glycosylated residues, in the ends and loops that protrude on the two sides of the membranes. Other proteins are anchored by insertion into membranes of nonpolar groups. These include fatty acyl groups at N termini, fatty acid ester groups on serine, threonine, or cysteine side chains, and polyprenyl groups attached to cysteine side chains as thioethers.[202] Many membrane-anchored proteins carry the saturated 14-carbon **myristoyl** group in amide linkage with an N-terminal glycine of the protein.[203–206] The fatty acid chain is added at the time of the protein synthesis, i.e., **cotranslationally**. It can intercalate into the membrane bilayer but provides a relatively weak anchor.[207] The 16-carbon **palmitoyl** group is typically added to a cysteine thiol in recognition sequences at various positions in a protein.[182,202,208,209] The modification occurs after protein synthesis, i.e., **posttranslationally**. Because thioesters are relatively unstable, palmitoylation is regarded as a reversible modification.[210]

A third lipid anchor is provided by the polyprenyl **farnesyl** (15-carbon) and **geranylgeranyl** (20-carbon) groups in thioether linkage to cysteine residues. These must be present in specific recognition sequences at the C termini of proteins, most often with the sequence CAAX.[211–215] The prenylation (also called isoprenylation) reaction is followed by proteolytic removal of the last three residues (AAX) and methylation of the new C-terminal carboxyl group as is discussed in Chapter 11, Section D,3. See also Chapter 22, Section A,4.

A lipoprotein present in the periplasmic space of *E. coli* is anchored to the outer bacterial membrane by a triacylated modified N-terminal cysteine containing a glyceryl group in thioether linkage as shown in the following structure (see also Section E,1).



A related anchor that uses diphytanylglycerylation is found in certain proteins of archaeobacteria.[216]



A series of **glycosylphosphatidylinositol** (GPI) anchors that are covalently linked to a variety of proteins utilize diacylglycerol or alkylacylglycerol for attachment to a bilayer. The proteins are joined through their C-terminal carboxyl groups to the diacylglycerol by a chain of covalently linked ethanolamine, phosphate, mannose, glucosamine, and *myo*-inositol as shown in Fig. 8-13.[217–223] The proteins are linked to the diacylglycerol through the conserved structure: H$_2$N-protein→ethanolamine→*P*→6Manα1→2Manα1→6Manα1→4GlcNAc→Ins→diacylglycerol.

The structure in Fig. 8-13, which anchors the small Thy-1 antigen to surfaces of rat thymocytes,[224] contains additional mannose, *N*-acetylgalactosamine, and ethanolamine phosphate. These groups may be missing or substituted by other groups in other anchors.[225] For example, that of human erythrocyte acetylcholinesterase lacks the extra mannose and GalNAc of the Thy-1 anchor but contains a palmitoyl group attached to an oxygen atom of the inositol. This provides an additional hydrocarbon tail that can enter the bilayer.[218] Other PI-anchored proteins include enzymes, such as alkaline phosphatase and lipoprotein lipase, and surface proteins of the parasites *Trypanosoma*,[225]*Leishmania,* and *Toxoplasma*.[220] Some adhesion molecules and a variety of other outer surface proteins are similarly attached to membranes. Analysis of the genome of *Caenorhabditis elegans* suggests that the nematode contains over 40 GPI-tailed proteins and perhaps more than 120.[225a]

Analyzing erythrocyte membranes. The proteins of red blood cell membranes were among the first to be studied. Because membrane proteins are present in small amounts and tend to be hard to dissolve without denaturation, they have been difficult to study. They usually do not dissolve readily in water, but red cell membranes can be almost completely solubilized in water using a 5 x 10^{-3} M solution of the chelating agent EDTA (Table 6-9) or by 0.1 M tetramethyl ammonium bromide.[226] These observations suggested that ionic linkages between proteins, or between proteins and phospholipids, are important to membrane stability. Nonionic detergents such as those of the Triton X series or β-octylglucopyranoside also solubilize most membrane proteins,[178,227–229] whereas ionic detergents such as sodium dodecyl sulfate (SDS) often cause unfolding and denaturation of peptide chains.

Gel electrophoresis of plasma membrane proteins in SDS solution yields ~10 prominent bands and at

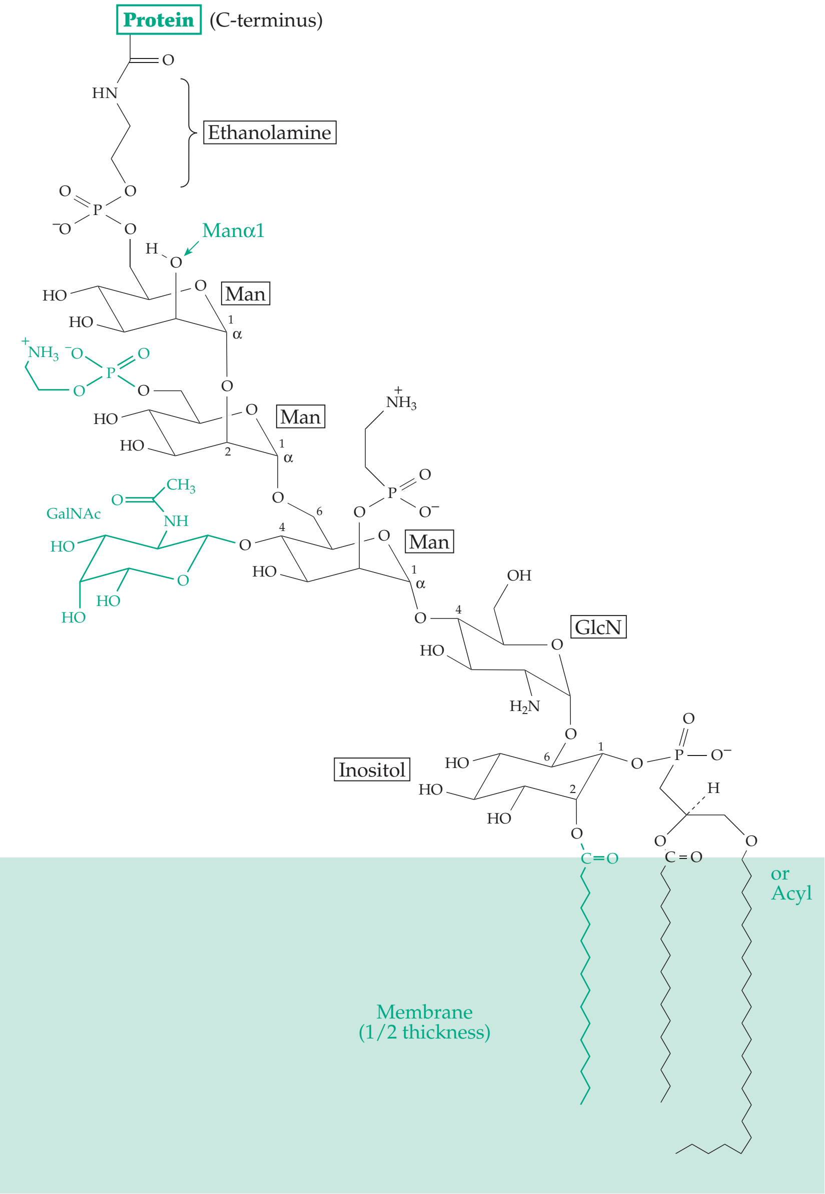
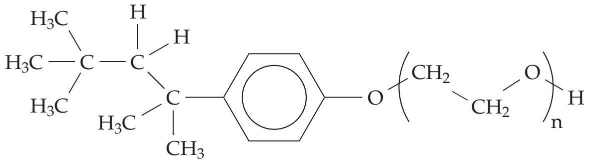


Figure 8-13 Structure of glycosylphosphatidylinositol (also called phosphatidylinositol-glycan) membrane anchors. The core structure is shown in black. The green parts are found in the Thy-1 protein and / or in other anchors.



Polyoxyethylene p-t-octyl phenol: the Triton X series of detergents. n = 9–10 for Triton X-100 and 7–8 for Triton X-114.

least 30 less intense bands ranging in molecular mass from 10 to 360 kDa.[230,231] Proteins of erythrocyte ghosts, and the usual system for numbering them, are shown in Fig. 8-14. Some very important proteins known to be present in this membrane, such as $(Na^+ + K^+)$-activated ATPase (Section C,2), are found in such low quantities (e.g., a few hundred molecules in a single red blood cell)[62] that they do not show up in electropherograms. Mitochondrial membranes appear to be more complex than plasma membranes, but myelin has a somewhat simpler composition.

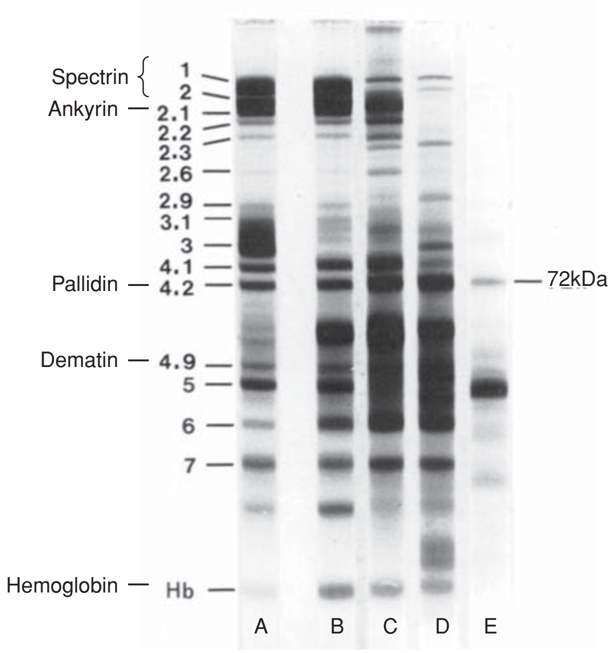


Figure 8-14 SDS-polyacrylamide gel electrophoresis of human erythrocyte ghosts. (A) From untreated cells. (B) From cells digested externally with chymotrypsin. (C) Inside-out vesicles prepared from cells pretreated with chymotrypsin. (D) The same inside-out vesicles after further treatment with chymotrypsin. (E) Polypeptides released by the chymotryptic treatment of the inside-out vesicles. The peptides are numbered according to the system of Steck[232]; Hb, hemoglobin. From Luna *et al.*[233]

Glycoproteins. Many of the integral proteins of membranes are glycoproteins.[234–236] These may sometimes be recognized in electropherograms because they are stained by the periodic acid-Schiff (PAS) procedure. At least 20 glycoproteins are present in erythrocyte ghosts, and glycoproteins appear to be prominent protein components of the plasma membranes of all eukaryotes and of primitive archaebacteria such as *Halobacterium*.

The most abundant glycoprotein of red blood cell membranes is the 95-kDa PAS-reactive **band 3 protein** (Fig. 8-14) which makes up ~25% of the total membrane protein.[236–240] A variety of asparagine-linked oligosaccharides based on the core hexasaccharide structure shown in Chapter 4, Section D,2 are present and apparently project into the surrounding medium. Electron micrographs of freeze-fractured surfaces through the membrane bilayer (Fig. 8-15) show ~4200 particles of 8 nm diameter per square micrometer, randomly distributed and apparently embedded in the membrane. These probably represent dimers of the glycoprotein. The amino acid sequence of the 911-residue protein suggests 13 membrane-spanning helices in the 550-residue C-terminal domain and that the 41-kDa N-terminal domain projects inward into the cytoplasm.[236,240] Among the first 31 residues are 16 of aspartate or glutamate. These provide a highly negatively charged tail that is able to interact electrostatically with other proteins.[236] Among these are components of the cytoskeleton, which appears to be anchored to the membrane via the band 3 protein. The band 3 protein is also a substrate for transglutaminase (Eq. 2-23) which creates covalent crosslinks to other proteins.[237] Another major function of the band 3 glycoprotein is to form channels for the transport of anions (Section C,2).

Another integral glycoprotein of erythrocytes, the 31-kDa **glycophorin A** (PAS-1),[235,241–244] is ~60% by weight carbohydrate. Its 131-residue chain has a single ~23-residue membrane-spanning helix. The first 50 residues at the N terminus, which project from the outside surface of the membrane, include many serines and threonines. Their side chains carry 15 O-linked tetrasaccharides and one complex N-linked oligosaccharide. There are a total of ~ 160 sugar residues per peptide chain, largely *N*-acetylgalactosamine, galactose, and sialic acid. Some of the oligosaccharides contain MNO blood group determinants[234,241] (Box 4-C). Because of a high content of sialic acid, glycophorin also carries a large negative charge. The 35-residue C-terminal domain is hydrophilic and rich in proline, glutamate, and aspartate. It probably extends into the cytoplasm and may bind calcium ions or interact with $-NH_3^+$ groups on phospholipid heads.

If all the sugar residues of the glycophorin molecules in an erythrocyte were spread over the surface of the cell they could cover approximately one-fifth of its surface in a loose network. However, it is more likely that they form bushy projections of a more localized sort. These oligosaccharides not only act as immunological determinants but also serve as receptors for influenza viruses. Other glycoproteins related to glycophorin A occur in smaller amounts.[244]

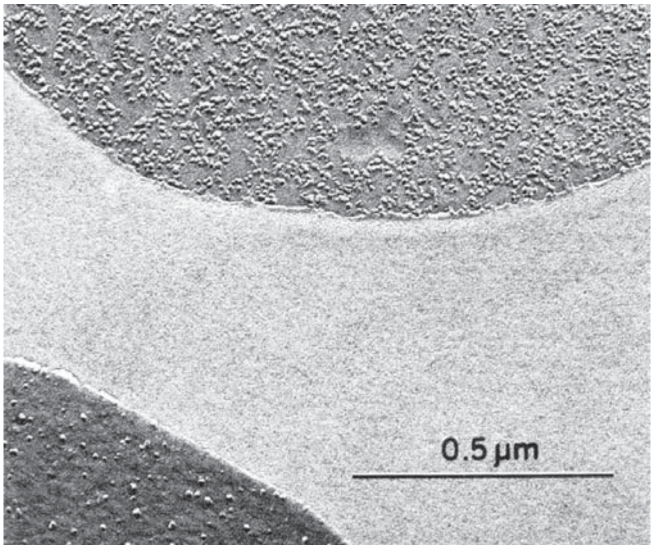


Figure 8-15 Freeze-fractured membranes of two erythrocyte "ghosts." The upper fracture face (PF) shows the interior of the membrane "half" closest to the cytoplasm. The smooth region is lipid and contains numerous particles. The lower face, the extracellular half (EF), possesses fewer particles. The space between the two is nonetched ice. See Figs. 1-4 and 1-15 for electron micrographs of sections through biological membranes. Courtesy of Knute A. Fisher.

Connections to the cytoskeleton. About one-third of the protein of the red blood cell membrane is accounted for by a pair of larger hydrophobic peptides called **spectrin** with molecular masses of 280 kDa (α chain) and 246 kDa (β chain).[245–250a] These are found in bands 1 and 2 of Fig. 8-14. The spectrin monomers consist largely of 106- to 119- residue repeat sequences each of which folds into a short triple-helical bundle (Fig. 8-16B). The beaded-chain monomers associate readily to $\alpha\beta$ dimers, long ~100 nm thin flexible rods which associate further to $(\alpha\beta)_2$ tetramers. The latter, in turn, bind to monomers or to small oligomers of actin. In red blood cells the actin crosslinks the spectrin tetramers into a two-dimensional "fishnet" (Fig. 8-16A), the ~85,000 spectrin tetramers uniformly covering the entire inner surface (130 μm^2) of the erythrocyte.[245,251] The inner location of spectrin was established by the fact that chemical treatments[252] that covalently label groups on proteins exposed on the outer surface of erythrocytes (e.g., iodination with lactoperoxidase; Chapter 16) did not label spectrin.[257]

Spectrin also binds to one domain of another large 215-kDa peripheral protein called **ankyrin** (band 2.1, in Fig. 8-14) which anchors the spectrin network to the membrane.[258,259] Ankyrin is actually a multigene family of related proteins that are present in many metazoan tissues.[258,260] These are modular proteins with separate binding domains for spectrin and band 3 protein. The latter domain contains 24 **ankyrin repeats**, 33-residue modules, also found in a variety of other proteins.[258,259]

Ankyrin binds firmly to the band 3 glycoprotein which is embedded in the membrane. The 78 kDa band 4.1 protein is another major component of the erythrocyte **membrane skeleton**.[261,262] Protein 4.1 binds both to ankyrin and also to **glycophorin C**, providing another anchor to an integral membrane protein (Fig. 8-16C). Spectrin:actin:protein 4.1 in a 1:2:1 ratio are the major components of the membrane skeleton.[262] Other less abundant proteins include **adducin**, protein 4.2 (**pallidin**, which interacts with band 3 protein),[263] protein 4.9 (**dematin**, an actin bundling protein),[264] and the muscle proteins tropomyosin and tropomodulin (Chapter 19). While spectrin and ankyrin of erythrocyte membranes have been studied most intensively, related proteins occur in other cells.[265] Spectrin of brain and other tissues is also known as **fodrin**.[266] **Dystrophin** and **α-actinin**, actin-crosslinking proteins of muscle, are also members of the spectrin superfamily.[248] Protein 4.1 also occurs in various organisms,[267] in various tissues, and in various locations in cells.[268]

What is the function of the membrane skeleton? There is a group of hereditary diseases including **spherocytosis** in which erythrocytes do not maintain their biconcave disc shape but become spherical or have other abnormal shapes and are extremely fragile.[269–272] Causes of spherocytosis include defective formation of spectrin tetramers and defective association of spectrin with ankyrin or the band 4.1 protein.[265,273] Thus, the principal functions of these proteins in erythrocytes may be to strengthen the membrane and to preserve the characteristic shape of erythrocytes during their 120-day lifetime in the bloodstream. In other cells the spectrins are able to interact with microtubules, which are absent from erythrocytes, and to microtubule-associated proteins of the cytoskeleton (Chapter 7, Section F).[270] In nerve terminals a protein similar to erythrocyte protein 4.1 may be involved in transmitter release.[274] The cytoskeleton is also actively involved in transmembrane signaling.

Integrins and focal adhesions. Mature erythrocytes have no nucleus and lack the microtubules and actin filaments that span other cells. In nonerythroid cells the major connections of the cytoskeleton to the membrane are through large $\alpha\beta$ heterodimeric membrane-spanning proteins called integrins (Fig. 8-17). These proteins, as well as the ends of the actin filaments, tend to be concentrated in regions long observed and described by microscopists as focal adhesions.[275–279] These are also sites of interaction with the external proteins that form the **extracellular matrix (ECM)**. There are at least 16 different integrin

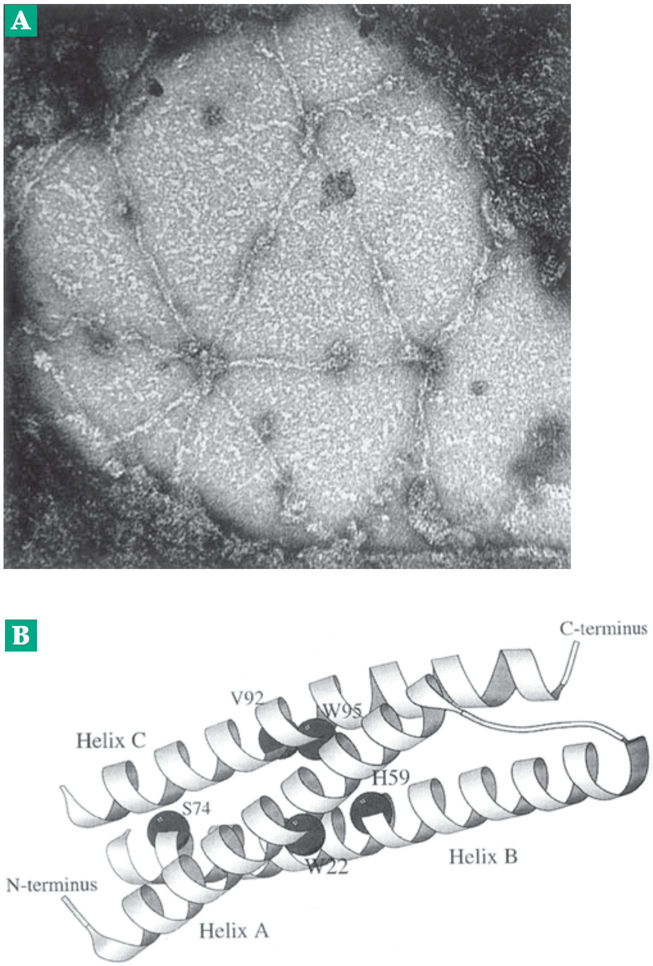
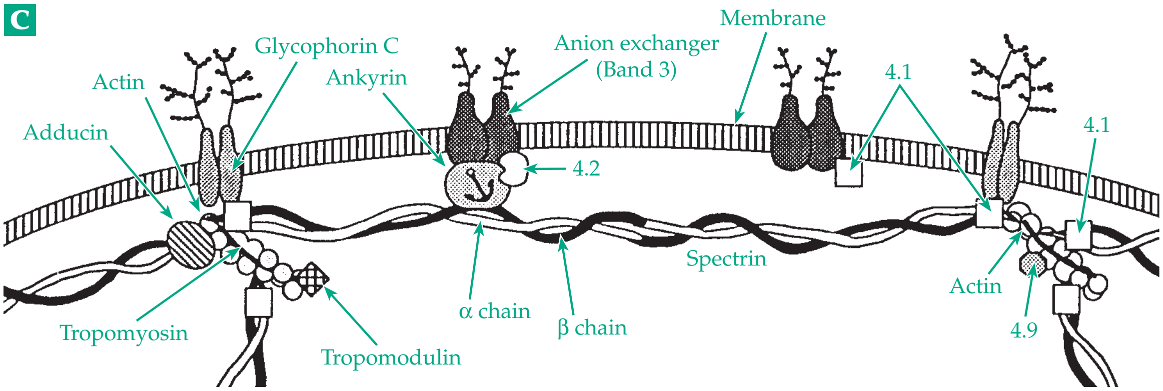


Figure 8-16 The erythrocyte membrane skeleton. (A) Electron micrograph showing a region of the membrane skeleton (negatively stained, X 200,000) and artificially spread to a surface area nine to ten times as great as in the native membrane. Spreading makes it possible to obtain clear images of the skeleton whose protein components are so densely packed and so subject to thermal flexing on the native, unspread membrane that it is difficult to visualize the individual molecules and the remarkably regular way that they are connected. The predominantly hexagonal and pentagonal network is composed of spectrin tetramers cross-linked by junctions containing actin oligomers and band 4.1 protein. Band 4.9 protein and tropomyosin are probably also bound to the oligomers, whose length (13 actin monomers long) corresponds to the length of a tropomyosin molecule. From Byers and Branton.[253,254] Courtesy of Daniel Branton. (B) Proposed triple α-helical structure of a single spectrin repeat unit.[255] Courtesy of Ruby I. MacDonald and Alfonso Mondragón. Each α spectrin chain consists of 20 and each β spectrin chain of 17 such repeats, which have only partially conserved sequences. The α and β chains are thought to associate in a side-by-side fashion and the αβ heterodimers in an end-to-end fashion to give tetramers. These are the rod-like structures seen in (A) and (C). (C) Cross section of the unspread membrane and cytoskeleton as pictured by Luna and Hitt.[256] Major interactions among components of the cytoskeleton are shown. Apparent sizes of the protein subunits, based on migration positions in SDS-polyacrylamide gel electrophoresis are: spectrin (260 and 225 kDa), adducin (105 and 100 kDa), band 3 (90 to 100 kDa), protein 4.1 (78 kDa), protein 4.2 (pallidin, 72 kDa), dematin (protein 4.9, 48 kDa), glycophorin C (~25 kDa), actin (43 kDa), and tropomyosin (29 and 27 kDa). Courtesy of Elizabeth J. Luna.



α subunits, whose external domains consist of up to 1114 residues, at least 8 β subunits with external domains of up to 678 residues, and at least 22 distinctly different αβ heterodimers.[280–282] The N-terminal part of each α subunit contains seven repeats of ~60 residues each, probably arranged as a β-propeller (see Fig. 11-7D or 15-23).[281] The integrins are structurally complex. Some contain a nucleotide-binding domain (see Figs. 2-13 and 2-27C).[283] Integrins have differing and quite exacting specificities toward the proteins of the external matrix to which they bind. They span the cell membrane and appear to be actively involved in communication between the cytoskeleton and external proteins.[275,277,284] They are often described as **receptors** for the proteins that bind to them.

Other cell adhesion molecules. Long before the discovery of integrins another class of transmembrane adhesion molecules were recognized as members of the immunoglobulin family. These were called cell

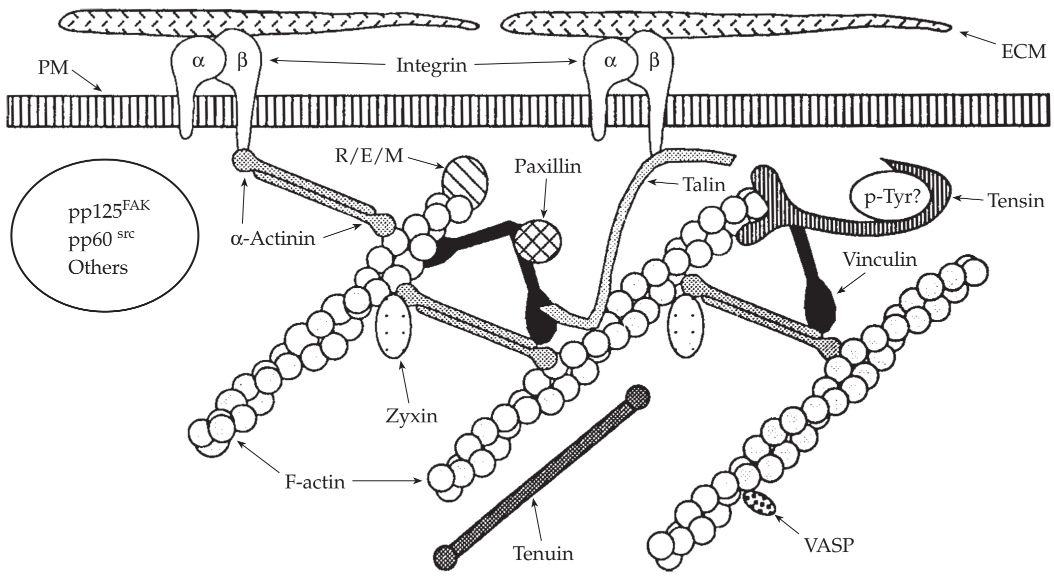


Figure 8-17 Working model of the protein–protein interactions in focal adhesions determined by *in vitro* binding experiments and immunolocalization. In addition, several interactions are of relatively low affinity in solution but may be enhanced at the membrane surface. Abbreviations are: ECM, extracellular matrix; PM, plasma membrane; p-Tyr-?, unknown phosphotyrosine-containing protein; R/E/M, member of the radixin/ezrin/moesin family; VASP, vasodilator-stimulated phosphoprotein. Diagram is modified from Simon *et al.*[285]

adhesion molecules (**CAMs**) by Edelman.[286,287] These molecules go by names such as intercellular adhesion molecule-1 (ICAM-1), neural cell adhesion molecule (NCAM),[288,289] liver cell adhesion molecule (LCAM), vascular cell adhesion molecule (VCAM), and platelet endothelial cell adhesion molecule (PECAM).[290,291] Many of these proteins, including the **T-cell antigen CD2**,[291a] were first recognized on leukocyte surfaces as differentiation antigens and are often designated by the "Cluster of Differentiation" names (Chapter 31). For example, ICAM-1 is also called CD54 and PECAM CD31.[290,292] The genes for these proteins are often expressed differentially in various tissues and the mRNA molecules formed undergo alternative splicing.[287,289] The extracellular domains of these adhesion molecules consist largely of Ig domains and most have a single transmembrane helix and a small cytoplasmic C-terminal domain. Some of the alternatively spliced forms are attached to the membrane by PGI tails.[287,289] ICAM-1 (Fig. 8-18) has five Ig domains and VCAM-1 has seven,[293] but some CAMs have only two.[294] The CAMs are glycoproteins, often with large *N*-linked oligosaccharides attached. The widely distributed N-CAM contains long $\alpha 2 \rightarrow 8$ linked polysialic acid chains on two of the three *N*-glycosylation sites in the fifth Ig domain.[289] The CAMs are often referred to as **receptors**. Their ligands include surface proteins such as fibronectin (next section) but also the integrins, which are also called receptors. Integrins are **coreceptors** for receptors of the Ig superfamily. Each of the coreceptors in a pair has binding sites for other ligands as well (Fig. 8-18). The CAMs and many other adhesion molecules are most abundantly expressed in embryonic tissues in which cells often migrate to new locations and for which the communication with neighboring cells is especially active. Many adhesion molecules bind only weakly and reversibly to their ligands, allowing cells to move.

The **cadherins** are calcium-dependent adhesion proteins that mediate direct cell–cell interactions.[295,296] The external parts of the cadherins also have repeated structural domains with the Ig fold.[297-298b] They have high affinity for each other, allowing cadherins from two different cells to interact and tie the cells together with a zipper-like interaction that is stabilized by the bound Ca^{2+} ions,[297,300] and may be relatively long-lived. The gene for cadherin E is often mutated in breast cancers and may be an important **tumor suppressor gene** (Box 11-D).[301]

Peripheral proteins of the outer membrane surface. Many integral membrane glycoproteins have their sugar-bearing portions exposed on the outer surface of the plasma membrane. Among these are receptors, ion pumps, and biochemical markers of individuality. In addition to these proteins, which are actually embedded in the bilayer, there are external peripheral proteins. One of the best known of these is

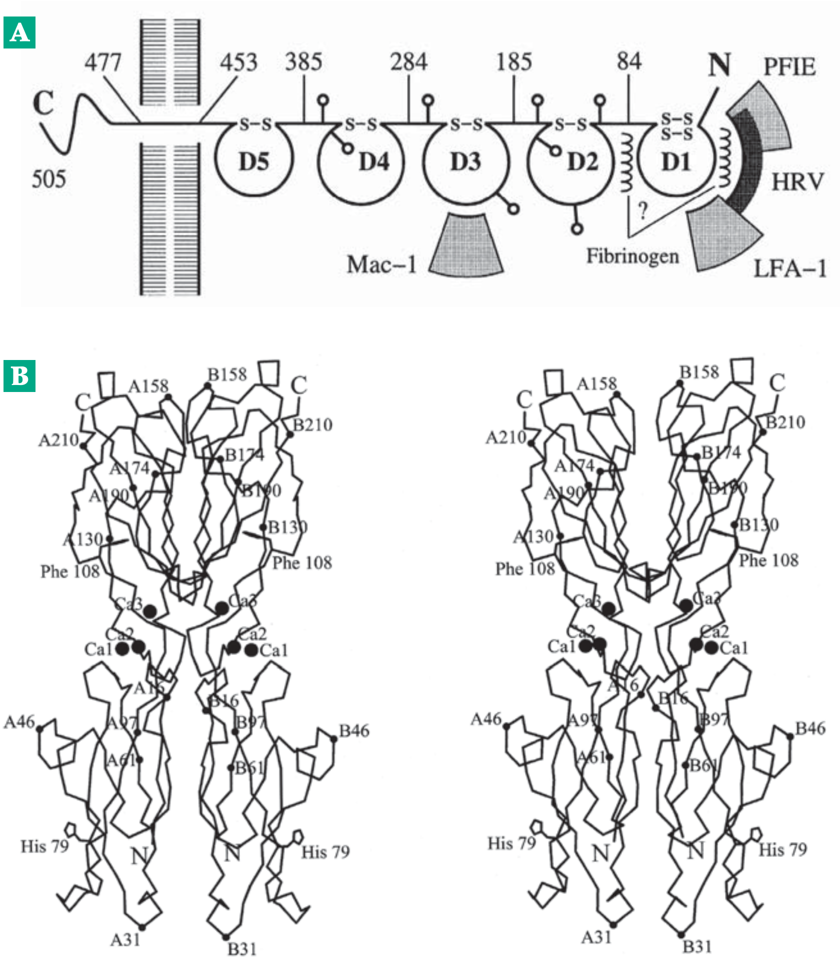
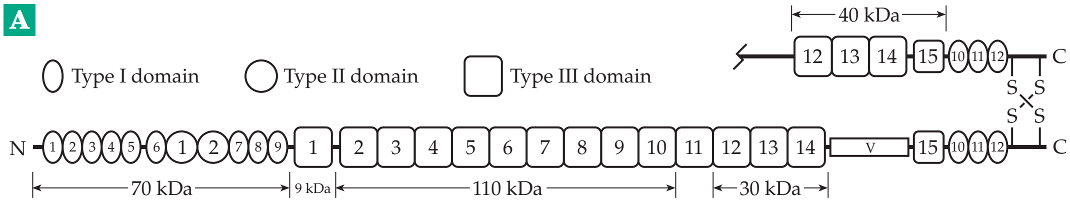


Figure 8-18 (A) Diagram of the ICAM-1 molecule. The structures labeled D1–D5 are the Ig domains. The glycosylation sites are labeled with small lollipops and approximate sites for binding of the α chain (CD11a) of the integrin called LFA-1, for the macrophage antigen Mac-1, for fibrinogen, for a ligand from erythrocytes infected by the malaria organism *Plasmodium falciparum* (PFIE), and for human rhinoviruses (HRV) are labeled. The binding sites are indicated schematically but each one is a complex interacting surface complementary to its ligand. From Bella *et al.*[299] Courtesy of Michael Rossman. (B) Structure of two N-terminal domains of E-cadherin. The molecules of the dimer are related by a noncrystallographic twofold symmetry axis running vertically in the plane of the page. Clusters of three calcium ions are bound in the linker regions, connecting the N- and C-terminal domains of each molecule and, in this view, are separated by the twofold axis. The N- and C-terminal domains are composed of seven-stranded β-barrels showing the same topology and similar three-dimensional structures. From Nagar *et al*.[296]

fibronectin (from *fibra*, "fiber", and *nectare*, "to bind").[302–309] This very large 470-kDa glycoprotein is a disulfide-linked dimer. Appearing under the electron microscope as having two 60-nm arms,[305] fibronectin molecules join together and surround animal cells, anchoring other proteins and carbohydrates of the ECM to the cells. Fibronectin binds tightly to several different cell surface integrins and also to collagen and to glycosaminoglycans of the matrix. It binds to the blood-clotting protein fibrinogen, to actin, and also to staphylococci and other bacteria. Each peptide chain of the fibronectin dimer is organized as several domains. Fibrin and staphylococci bind to the N-terminal domain. A second domain binds collagen and a domain near the C-terminus binds heparin. Several different integrins of cell surfaces bind to the region of the 8th, 9th, and 10th type III repeats.[308] The specificity of the binding depends to a large extent on the presence of the specific tripeptide sequence Arg-Gly-Asp (**RGD**) in a type III repeat (Fig. 8-19B). Peptide sequences such as the PHSRN (Fig. 8-19B) and others[310] also participate in binding to specific integrins. Initial interactions are noncovalent but fibrin and collagen gradually become covalently attached to fibronectin through isopeptide linkages formed by transglutaminase (Eq. 2-23). This enzyme is abundant



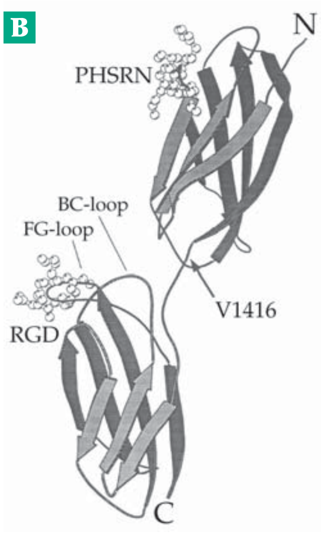


Figure 8-19 (A) Schematic diagram of a human fibronectin molecule showing one complete ~ 250-kDa chain consisting largely of 29 type I, II, or III fibronectin domains with a 12-kDa type III connecting segment or "V-region." The second chain, for which only the C-terminal portion is shown, is identical to the first except for the absence of the V-region as a result of alternative splicing of the mRNA. The two chains are joined in antiparallel fashion by a pair of –S–S– bridges near the C terminus.[313] The kDa sizes refer to fragments generated by the action of thermolysin. After Ingham *et al*.[306] (B) MolScript ribbon drawing of the structure of the ninth and tenth type III repeats in human fibronectin. The biologically active peptide segments PHSRN and RGD are represented in a stick and ball mode. The arrow marks valine 1416, next to which a polyglycine linker of various lengths has been inserted experimentally in order to study interactions between modules.[308] From Spitzfaden *et al*.[314]

in many developing tissues but often disappears as the tissues mature.[311,312] Fibronectin was formerly known as the "large external transformation-sensitive protein," a name derived from the fact that its quantity is greatly reduced in many virus-transformed cells. This might explain the loss of adhesiveness and of "contact inhibition" observed for cancer cells (Box 11-D).

There are 20 isoforms of human fibronectin. These arise by alternative splicing of the primary gene transcript. Their formation is developmentally regulated.[308,315] One of the isoforms is present as a soluble protein (cold-insoluble globulin) in blood plasma.[302] The fibronectins belong to a larger family of **cytoadhesins**, among which are the blood-clotting proteins **fibrinogen** and **von Willebrand factor** (Chapter 12), and also **thrombospondin**, **vitronectin**,[316] **tenascin**,[317–320] **laminin**,[321,322] **osteopontin**,[323] and **collagens**. These proteins all have a modular construction with repeated domains, often of several different types.[324] For example, tenascin contains 14 ½ EGF domains (Table 7-3), 16 fibronectin type III repeats, and a C-terminal segment homologous to fibrinogens.[325] Laminin (see Fig. 8-33) also has EGF-like modules. Most of these proteins also have in common the presence of the RGD sequence, which binds the cytoadhesins to those integrins and other molecules that carry RGD receptor sites.[304,326–328] The extracellular matrix is discussed further in Section E,2.

3. Enzymes and Membrane Metabolism

Many of the proteins of membranes are enzymes. For example, the entire electron transport system of mitochondria (Chapter 18) is embedded in membranes and a number of highly lipid-soluble enzymes have been isolated. Examples are **phosphatidylserine decarboxylase**, which converts phosphatidylserine to phosphatidylethanolamine in biosynthesis of the latter, and **isoprenoid alcohol phosphokinase**, which participates in bacterial cell wall synthesis (Chapter 20). A number of **ectoenzymes** are present predominantly on the outsides of cell membranes.[329] Enzymes such as phospholipases (Chapter 12), which are present on membrane surfaces, often are relatively inactive when removed from the lipid environment but are active in the presence of phospholipid bilayers.[330,331] The distribution of lipid chain lengths as well as the cholesterol content of the membrane can affect enzymatic activities.[332]

Why are membranes so important to cells? Besides their obvious importance in enclosing and defining the limits of cells, membranes are the result of a natural aggregation of amphipathic molecules. They also represent a natural arrangement for boundaries between different aqueous phases within a cell. In addition, membranes are the "habitat" for many relatively nonpolar molecules formed by metabolism. These include proteins with hydrophobic surfaces and those with membrane anchors. The semiliquid interior of the membrane permits distortion of the bilayer and the addition or subtraction of proteins (and low-molecular-mass materials) in response to metabolic processes in the adjacent cytoplasm.

The principal factor providing stability to macromolecules and membranes is the hydrophobic nature of reduced organic compounds. This characteristic leads to the separation of lipids, proteins, and other molecules from the aqueous cytoplasm into oligomeric aggregates and membranes. However, the best catalysts, including most enzymes, are soluble in water. Thus, membranes represent thin regions of relative stability adjacent to aqueous regions in which chemical reactions occur readily and which tend to contain the more polar, the smaller, and the more water-soluble materials. The stability of membrane surfaces provides a means of bringing together reactants and of promoting complex sequences of biochemical reactions. For example, membranes contain both oxidative enzymes and reactive, dissolved quinones. The membrane–cytoplasmic interfaces may often be the metabolically most active regions of cells.

Despite their stability, membrane components have a metabolism of their own which is related to the high concentrations of oxidizing enzymes located in or on membranes. Oxidative reactions provide a mechanism for modification of hydrophobic membrane constituents. For example, sterols, prostaglandins, and other regulatory molecules are initially synthesized as hydrophobic chains attached to water-soluble "head groups" (Chapter 21). The hydrophobic products of these synthetic reactions tend to be deposited in membranes. However, attack by oxygen leads to introduction of hydroxyl groups and to a gradual increase in water solubility. As the hydrophilic nature of the compound is increased through successive enzymatic hydroxylation reactions, the hydrophobic membrane constituents eventually redissolve in the water and are completely metabolized. Another process that actively degrades membrane lipids is attack by hydrolytic enzymes such as the phospholipases.

C. The Transport of Molecules through Membranes

Small neutral molecules, such as water or ethanol, can penetrate membranes by **simple diffusion**.[64,150] The rate is determined by the solubility of a substance in the membrane, by its diffusion coefficient (see Eq. 9-24) in the membrane, and by the difference in its concentration between the outside and the inside of the cell. This concentration difference is commonly referred to as the **concentration gradient** across the membrane. The ease of diffusion through a membrane is described quantitatively by a **permeability coefficient** P which is related to the diffusion coefficient D (Eq. 8-3).

$$J = -D_m Kc / \Delta x = -P\Delta c \qquad (8\text{-}3)$$

Here J is the **flux** of molecules across the membrane, i.e., the number of molecules crossing one cm^2 per second. D_m is the diffusion coefficient in the bilayer, while K is a **partition coefficient**, the ratio of the concentration of the diffusing solute in the bilayer to that in water. The concentration gradient of the solute across the membrane is Δc, while Δx is the membrane thickness in centimeters. The permeability coefficient P for H_2O through biological membranes[333] is about $1 - 10$ µm s^{-1}. For H^+ and OH^- P is 0.1 µm s^{-1}. For halide ions diffusing across liposome bilayers P ranges from 10^{-5} to 10^{-3} µm s^{-1},[334] fast enough to be of some biological significance. However, for most other ions P is less than 10^{-6} µm s^{-1}. Because of their high lipid content membranes are quite permeable to nonpolar materials. For example, anesthetics usually have a high solubility in lipids, enabling them to penetrate nerve membranes.

1. Facilitated Diffusion and Active Transport

While simple diffusion may account for the entrance of water, carbon dioxide, oxygen, and anesthetic molecules into cells, movement of most substances is facilitated by protein **channels** and **transporters**.[335] Genes of 76 families of such proteins have been located in the genomes of 18 prokaryotes.[335a] Some of these provide for **facilitated diffusion**.[169,335] Like simple diffusion, it depends upon a concentration gradient and molecules always flow from a higher to a lower concentration. A distinguishing feature of facilitated diffusion is a **saturation effect**, i.e., a tendency to reach a maximum rate of flow through the membrane as the concentration of the diffusing substance, on the high concentration side, is increased. In this characteristic it is similar to enzymatic action (Chapter 9).

In **active transport** a material is carried across a membrane against a concentration gradient, i.e., from a lower concentration to a higher concentration. This process necessarily has a positive Gibbs energy change (as given by Eq. 6-25) of approximately 5.71 log c_2 / c_1 kJ mol^{-1}, where c_2 and c_1 are the higher and lower concentrations, respectively. The transport process must be coupled with a spontaneous exergonic reaction. In **primary active transport** there is a direct coupling to a reaction such as the hydrolysis of ATP to "pump" the solute across the membrane. **Secondary active transport** utilizes the energy of an **electrochemical gradient** established for a second solute; that is, a second solute is pumped against a concentration gradient and the first solute is then allowed to cross the membrane through an exchange process with the second solute (**antiport** or **exchange diffusion**).

Alternatively, both the first and the second solutes may pass through the membrane bound to the same carrier (**cotransport** or **symport**). Another form of active transport is **group translocation**, a process in which the substance to be transported undergoes covalent modification, e.g., by phosphorylation. The modified product enters the cell and within the cell may be converted back to the unmodified substance. Transport processes, whether facilitated or active, often require the participation of more than one membrane protein. Sometimes the name **permease** is used to describe the protein complexes utilized.

Like facilitated diffusion, active transport depends upon conformational changes in carrier or pore proteins, the equilibrium between the two conformations depending upon the coupled energy-yielding process. Thus, if ATP provides the energy a phosphorylated carrier will probably have a different conformation than the unphosphorylated protein. A carrier with Na^+ bound at one site may have a different affinity for glucose than the same carrier lacking Na^+. A hypothetical example of the kind of cycle that can function in active transport is provided by the picture of the “sodium pump” given in Fig. 8-25.

2. Pores, Channels, and Carriers

To accommodate the rapid diffusion that is often needed to supply food, water, and inorganic ions to cells, membranes contain a variety of small pores and channels. The pores may be nonspecific or they may be selective for anions or cations or for some other chemical characteristics. They may be permanently open or sometimes closed and referred to as **gated**. The gating may be controlled by the membrane electrical potential, by a hormone, by the specific ligand, or by other means. Some pores may be small enough to allow only small molecules such as H_2O to pass through. Others may be large enough to allow for nonspecific simple diffusion of molecules of low molecular mass. Structures are known for only a few.

Large pores tend to be nonspecific, but when the solute approaches the pore diameter in size the specificity increases. Furthermore, diffusion of ions through pores is influenced strongly by any charged groups in or near the pore. Thus, a cation will not enter a pore containing a net positive charge in its surface. Any electrical potential difference across the membrane, resulting from accumulation of excess negative ions within the cell, will also affect the diffusion of ions.[336,337]

Porins. The outer membranes of gram-negative bacteria contain several 34- to 38-kDa proteins known as porins. They form a large number of trimeric pores which allow molecules and ions with molecular masses <600 Da to enter. However, even small proteins are excluded. This appears to be a means for protecting the bacteria against enzymes such as lysozyme (Chapter 12). Four distinct porins are among the most abundant proteins of the outer membrane of *E. coli.*[338] They are designated according to the names of their genes. **OmpF** (a nonspecific, open channel) and OmpC (**osmoporin)** are encoded by *OmpF* (outer membrane protein F) and *OmpC* genes, respectively. **Maltoporin** (LamB) is selective for maltodextrins but also allows other small molecules and ions to pass.[338–341a] Its name comes from its original discovery as a receptor for bacteriophage lambda. PhoE (**phosphoporin)** is a porin with a preference for anions such as sugar phosphates while OmpF prefers cations.[342] Porin **FepA** is ligand gated, opening to take up the chelated iron from ferric enterochelin (Chapter 16). OmpF, OmpC, and PhoE all have 16-stranded β-barrel structures (Fig. 8-20).[199,343] Maltoporin forms a quite similar 18-stranded barrel. Similar porins are present in many bacteria.[199,344]

The porin monomers associate to form trimeric channels as is shown in Fig. 8-20B. They all have a central water-filled, elliptical channel that is constricted in the center to an “eye” ~0.8 x 1.1 nm in size. In this restriction zone the channel is lined with polar residues that provide the substrate discrimination and gating. For example, in OmpF and PhoE there are many positively and negatively charged side chains that form the edge of the eye (Fig. 8-20C). The electrostatic potential difference across the outer membrane is small, but apparently determines whether the porins are in an open or a closed state.[344a] The voltage difference has opposite effects on OmpF and PhoE, apparently as a result of the differing distribution of charged groups.[342,345,346] A key role in determining the voltage dependence may be played by Lys 18 (Fig. 8-20C).[342] Polyamines (Chapter 24), which are present in the outer membrane, induce closing of porin channels by binding to specific aspartate and tyrosine side chains.[347]

The most *abundant* protein in the *E. coli* outer membrane is OmpA. It appears to form a transmembrane helical bundle. Although it is regarded primarily as a structural protein it too acts, in monomeric form, as an inefficient diffusion pore.[350] Mitochondrial outer membranes contain nonspecific pores (mitochondrial porins) that allow passage of sucrose and other saccharides of molecular mass up to 2 to 8 kDa.[351,352] Similar pore-forming proteins have been found in plant peroxisomes.[353]

Aquaporins. Many biological membranes are not sufficiently permeable to water to allow for rapid osmotic flow. For example, the kidney membranes in portions of Henle's loop have permeabilities as high as 2500 μm s^{-1} (compared to 10–20 μm s^{-1} in a 1:1 cholesterol / phospholipid bilayer).[354-355a] This high permeability is provided by aquaporin-1 (**AQP-1**), formerly called CHIP (channel-forming integral membrane protein) or AQP-CHIP. Aquaporin-1 was first identified in erythrocyte membranes and is present in many tissues. The 28-kDa subunits of the protein form six-helix bundles, each with a pore in the center. These are associated as tetramers in the membrane.[356–358]

Other aquaporins with related sequences occur broadly. There are at least ten in mammals.[359,359a] Plants, which must accommodate to heavy loss of water in hot dry weather, have aquaporins in both plasma membranes and tonoplasts.[360] Bacteria also have aquaporins.[356,361] A defect in aquaporin-2 of the kidney collecting duct leads to **nephrogenic diabetes insipidus**, in which the kidneys fail to concentrate urine in response to secretion of the hormone vasopressin.[355a,362,363]

Ion channels. Most organisms contain a large number of ion channels. One of these, which plays a key role in nerve conduction, is the **voltage-gated K^+ channel**. It is closed most of the time but opens when a nerve impulse arrives, dropping the membrane potential from its resting 50 - to 70-mV (negative inside) value to below zero. This voltage change opens the channel, allowing a very rapid outflow of K^+ ions.[364–367] The channels then close spontaneously. There are many different K^+ channels but most have a similar architecture.[368–370] The three-dimensional structure has been determined for the membrane-spanning part of the K^+ channel of *Streptomyces lividans* (Fig. 8-21). The funnel-shaped tetrameric molecule has a narrow **conduction channel** which contains the **selectivity filter**.

The conduction channel is lined largely with hydrophobic groups. The selectivity filter, which discriminates between K^+ and Na^+, is a short (~1.2-nm-long) narrow (~1.0-nm-diameter) portion of the

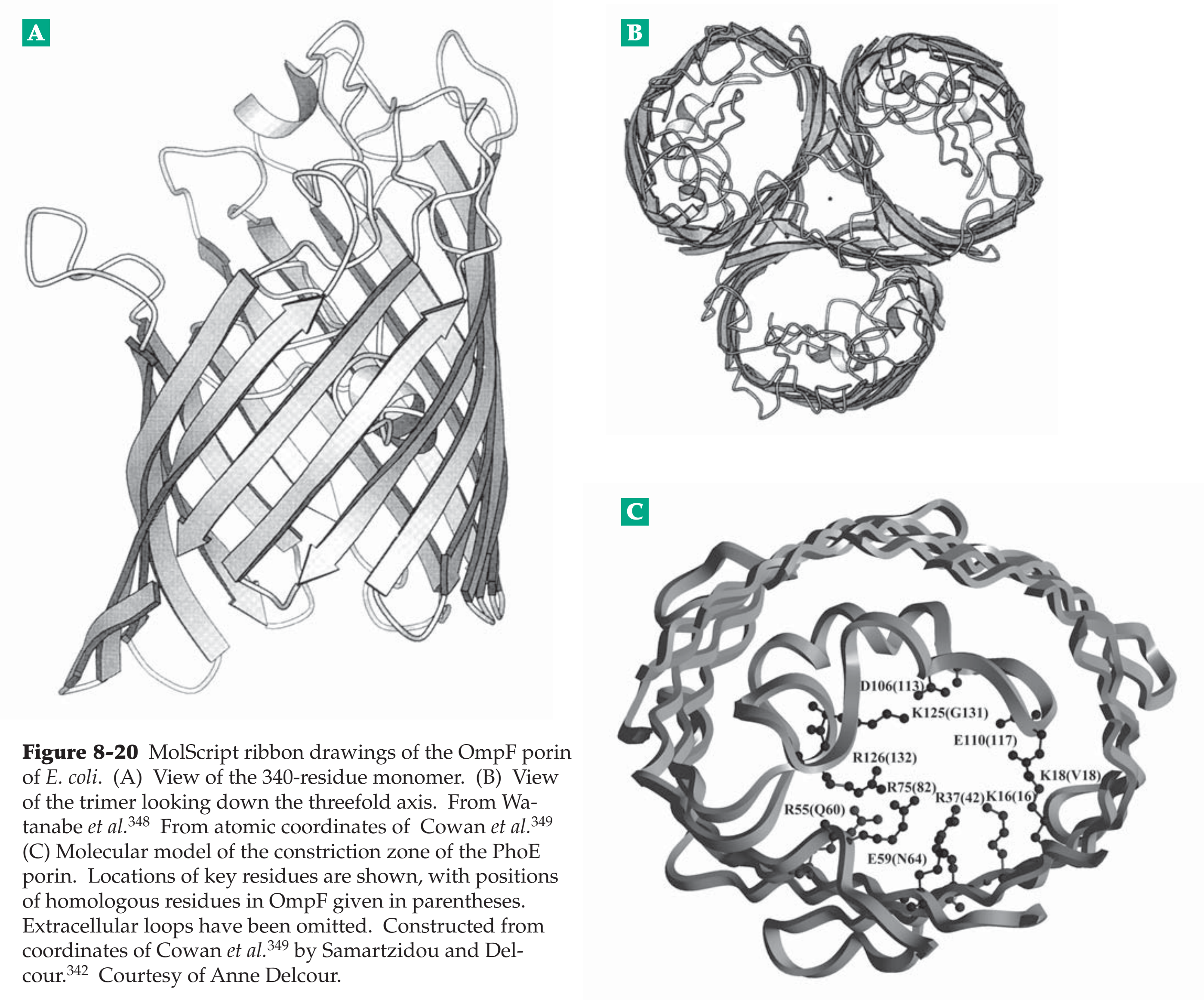


Figure 8-20 MolScript ribbon drawings of the OmpF porin of *E. coli*. (A) View of the 340-residue monomer. (B) View of the trimer looking down the threefold axis. From Watanabe *et al.*[348] From atomic coordinates of Cowan *et al.*[349] (C) Molecular model of the constriction zone of the PhoE porin. Locations of key residues are shown, with positions of homologous residues in OmpF given in parentheses. Extracellular loops have been omitted. Constructed from coordinates of Cowan *et al.*[349] by Samartzidou and Delcour.[342] Courtesy of Anne Delcour.

pore that is aligned roughly with the center of the bilayer. It is formed by four extended peptide chains, one from each subunit, each having the "**signature sequence**" for K^+ channels, TVGYG. In this sequence the peptide carbonyl groups all point into the channel. The consecutive groups of four carbonyls along the channel form binding sites for K^+, whose ionic diameter (Table 6-10) is 0.27 nm. The K^+ must lose its hydration sphere to fit into the 1.0-nm channel. The site of strongest binding, occupied by K^+ in Fig. 8-21C, lies just at the C-terminal ends of four helices, and the partial negative charges of the helix dipoles probably contribute to the binding. The pore may constrict to strengthen the bonds. When a second K^+ ion enters the channel it appears to bind ~0.75 nm from the central K^+, repelling it and weakening its interaction with the filter, allowing it to pass through the pore. Rb^+ (0.30 nm diameter) and Cs (0.34 nm diameter) also pass through.

How are the smaller Na^+ (0.19 nm diameter) and Li^+ (0.12 nm diameter) excluded from the pore? The pore is too small for the hydrated ions and perhaps too large to bind the dehydrated ion well enough to let it escape from its hydration sphere. Four negatively charged side chains in the cytoplasmic mouth of the pore presumably discourage anions from entering. At the other end, the extracellular entryway is a site that can be blocked specifically by 35- to 40-residue scorpion toxins.[368]

The *S. lividans* K^+ channel is *not* voltage gated. Voltage-gating mechanisms must be learned from study of other channels! The voltage-dependent K^+ channels from the rat have an $\alpha_4\beta_4$ composition. The α and β subunits coassemble in the endoplasmic reticulum and remain as a permanent complex,[370a–c] After insertion into the plasma membrane the α subunits form a channel as in Fig. 8-21. However, an additional intracellular domain of

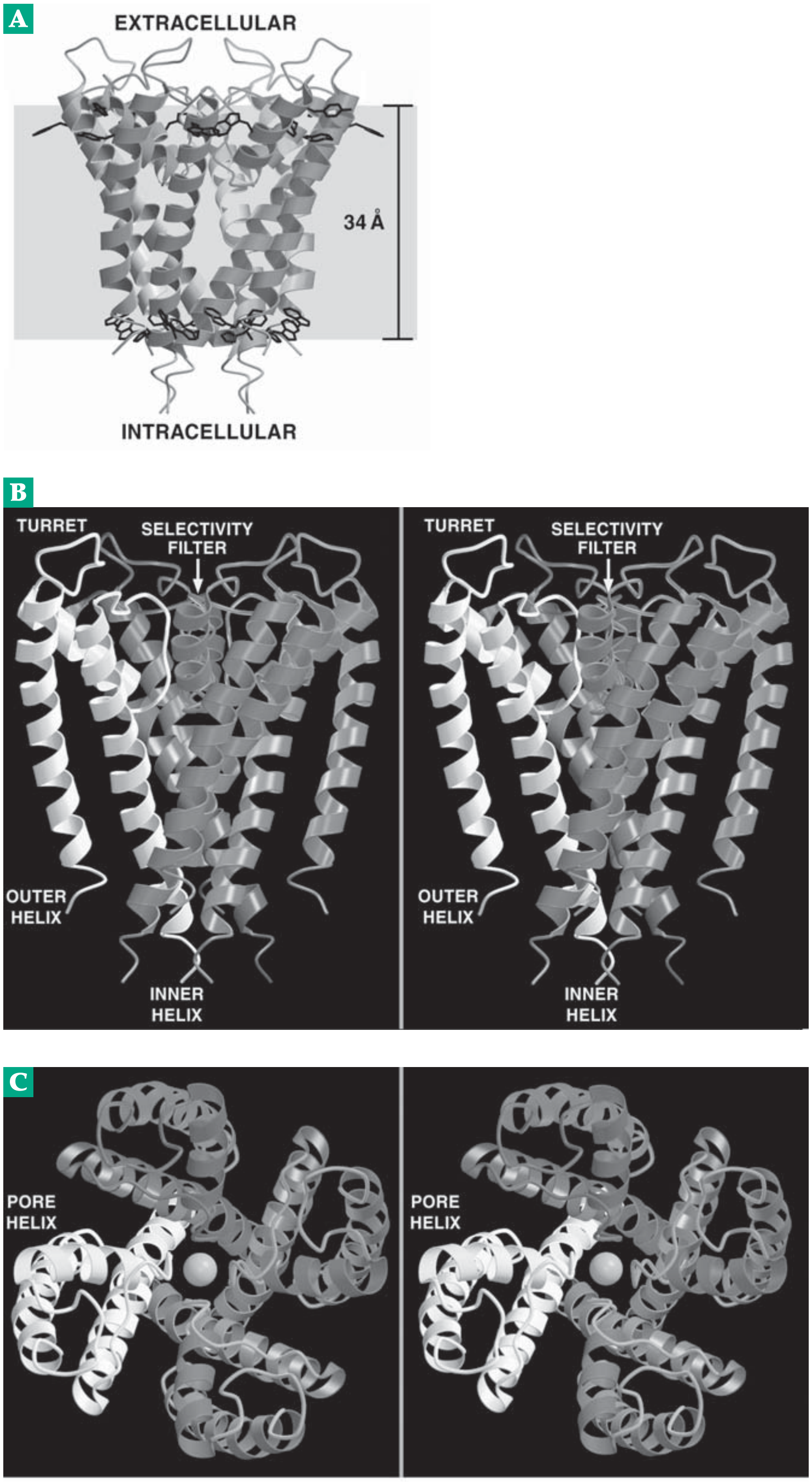


Figure 8-21 Views of the tetrameric K^+ channel from *Streptococcus lividans*. (A) Ribbon representation as an integral membrane protein. Aromatic amino acids on the membrane-facing surface are also shown. (B) Stereoscopic view. (C) Stereoscopic view perpendicular to that in (B) with a K^+ ion in the center. From Doyle *et al.*[366]

the α subunits, together with the β_4 complex, provide an elaborate and as yet poorly understood internal structure. The β subunits are oxidoreductases containing bound NADH, whose function is also uncertain.

There are many more pores and carriers of various kinds in biological membranes. Some, like the *S. lividans* K^+ channel, facilitate diffusion of a single ion or compound. They are **uniporters**. Others promote cotransport in which an ion such as H^+ or Na^+ also passes through the carrier. In some cases a pore is formed by a single molecule. Other pores follow a twofold, threefold, or fourfold axis (Fig. 8-21) of an oligomeric protein.[371] Two different conformations for the protein, one in which a "gate" opens to one side of the membrane and one in which it opens to the other, are usually involved. Interconversion between the two conformations is spontaneous but the equilibrium between them may be influenced by the binding of the solute, by the membrane potential, or by binding of inhibitors or activators. The latter may bind differently to the parts of the carrier (pore) protein exposed on the two sides of the membrane.

Channel-forming toxins and antibiotics. Some of the bacterial toxins known as **colicins** (Box 8-D) kill susceptible bacteria by creating pores that allow K^+ to leak out of the cells. One part of the **complement system** of blood (Chapter 31) uses specific proteins to literally punch holes in foreign cell membranes. **Mellitin**, a 26-residue peptide of bee venom,[372,373] as well as other hemolytic toxins and antibiotic peptides of insects, amphibians, and mammals (Chapter 31) form amphipathic helices which associate to form voltage-dependent anion-selective channels in membranes.[374–377]

The polypeptide antibiotic **gramicidin A** consists of 15 nonpolar residues of alternating D- and L-configuration; two molecules can form a channel with a right-handed β helix structure.[378,379] The central 0.48-nm hole in the β helix is large enough to allow unhydrated cations such as Na^+ or K^+ to pass through. The same peptide, under other conditions, forms left-handed helical channels whose structures are known.[379] Aggregates of β_{10} helices may form channels between helices of **suzukacillin**.[380] For this antibiotic as well as for **alamethicin** (Chapter 30) the conductance depends upon the membrane potential, a characteristic shared with the pores of nerve membranes. A variety of synthetic channel-forming peptides have been prepared. Some of these form α-helical bundles with a central pore. A five helix bundle of this type[381] has a structure reminiscent of the acetylcholine receptor channel of neurons.[382]

Ionophores and other mobile carriers. Facilitated diffusion of a molecule or ion is sometimes accomplished by binding to a **mobile carrier**. An example is the diffusion of a complex of K^+ with the low-molecular-mass lipid-soluble carrier, or ionophore **valinomycin** (Fig. 8-22). The K^+–valinomycin complex diffuses the short distance to the other side of the membrane and discharges the bound ion. If the rates of binding to a carrier and of release from the carrier are greater than those of the diffusion process, Michaelis–Menten kinetics are observed. The maximum velocity V_{max} and Michaelis constant K_m can be defined as in Eq. 9-15 for enzymatic catalysis.

Valinomycin is a **depsipeptide** which contains ester linkages as well as amide linkages. The antibiotic is made up of D- and L-valine, L-lactic acid, and D-hydroxyisovaleric acid. When incorporated into an artificial membrane bathed in a K^+-containing medium, valinomycin increases the conductance greatly and when it is added to a suspension of *Streptococcus faecalis* cells the high ratio of $[K^+]_i$ / $[K^+]_o$ falls rapidly.[383] The loss of K^+ from cells probably explains the antibiotic activity. However, under suitable conditions, with a high external $[K^+]$, the bacteria will continue to grow and reproduce in the presence of the antibiotic.[384]

Uncomplexed valinomycin has a more extended conformation than it does in the potassium complex.[385,386] The conformational change results in the breaking of a pair of hydrogen bonds and formation of new hydrogen bonds as the molecule folds around the potassium ion. Valinomycin facilitates potassium transport in a passive manner. However, there are cyclic changes between two conformations as the carrier complexes with ions, diffuses across the membrane, and releases ions on the other side. The rate of transport is rapid, with each valinomycin molecule being able to carry $\sim 10^4$ potassium ions per second across a membrane. Thus, a very small amount of this ionophore is sufficient to alter the permeability and the conductance of a membrane.

Because the stability constant of its complex with potassium is much greater than that with sodium, valinomycin is a relatively specific potassium ionophore. In contrast, the mushroom peptide **antamanide** has a binding cavity of a different geometry and shows a strong preference for *sodium* ions.[388,390] The structure of the Na^+–antamanide complex is also shown in Fig. 8-22B. The *Streptomyces* polyether antibiotic **monensin** (Fig. 8-22D),[389,391] a popular additive to animal feeds, is also an ionophore. However, its mode of action, which involves disruption of Golgi functions, is uncertain.[392]

Anions of lipid-soluble phenols such as 2,4-dinitrophenol can serve as effective carriers of *protons* (Chapter 18). However, proteins usually serve as the natural carriers, both of protons and of other ions. A protein is sometimes pictured as rotating to present the solute-binding surface first to one side, then to the other side of a membrane. However, gated pores or channels are probable for most biological transport.

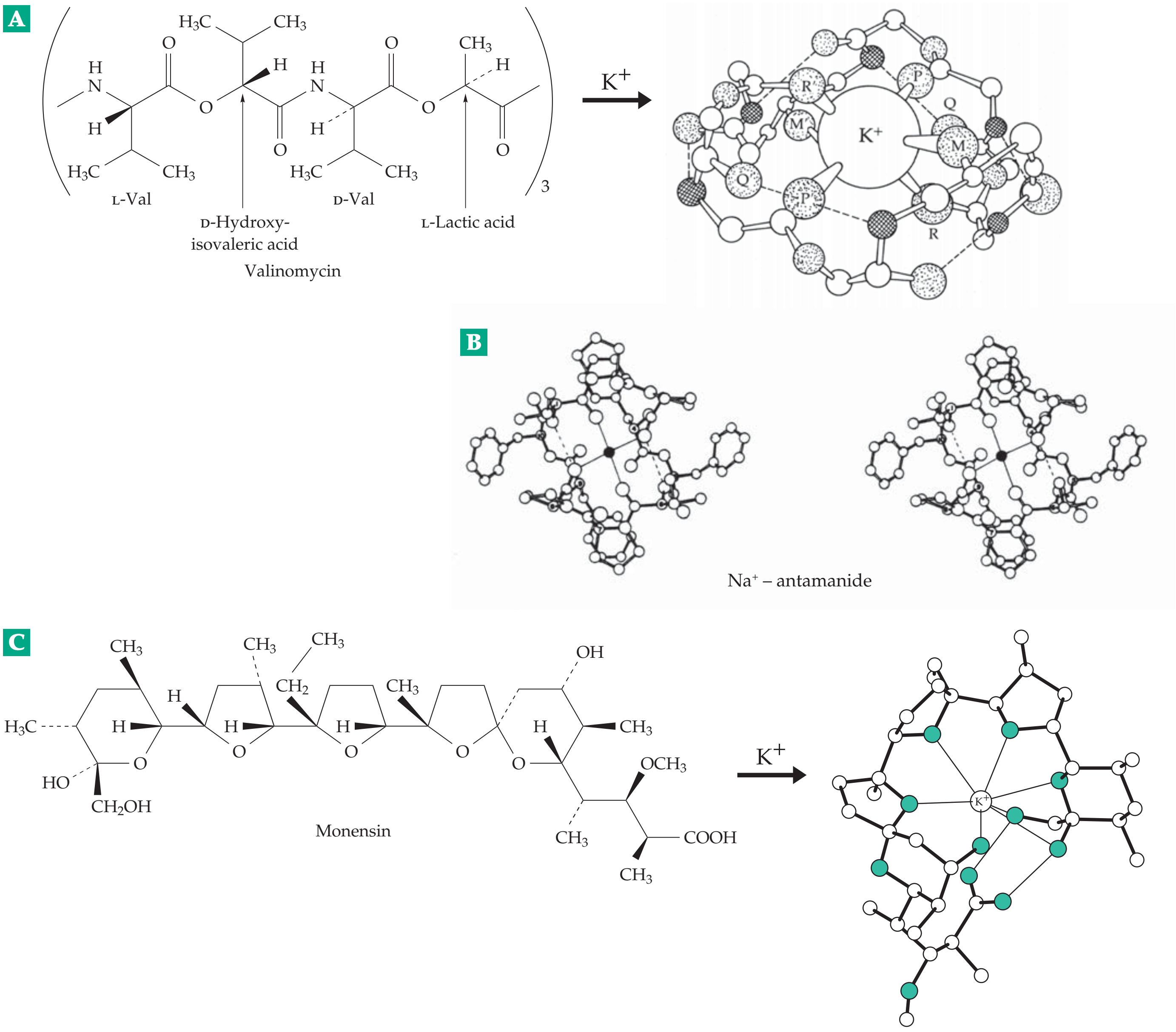


Figure 8-22 (A) Uncomplexed valinomycin and its complex with K+ (from Duax *et al.*[387]). (B) Stereodiagram of the Na–[Phe^4, Val^6]antamanide complex. A molecule of C_2H_5OH, which forms the fifth ligand to the Na+, is omitted for clarity. From Karle *et al.*[388] (C) Monensin and its complex with K+. From Pangborn *et al.*[389]

3. The 12-Helix Major Facilitator Superfamily

A large family of transmembrane facilitators from bacteria and eukaryotes appear to consist largely of 12 transmembrane helices with intervening cytosolic and extracellular loops. Some of these transporters facilitate simple uniport diffusion, but others participate in active transport of the symport or antiport type.[393,394] Several hundred members of the family are known.[395]

Entrance of sugars into cells. It is important that sugars be able to enter cells rapidly. However, the permeability coefficient *P* for D-glucose across a lipid bilayer is only 10^{-6} to 10^{-5} µm s^{-1}. For an intact erythrocyte, *P* is much greater: ~1 µm s^{-1}. This is the result of facilitated diffusion by a transport protein with a high specificity for hexose and pentose sugars having a pyranose ring in a C1 chain conformation.[396] This human erythrocyte glucose transporter, now known as **GLUT1**, has a K_m of 1.6 mM for D-glucose but of >3 M for L-glucose. It is a 55-kDa intrinsic membrane glycoprotein migrating in band 4.5 of Fig. 8-14. From the sequence of its cloned gene the unglycosylated carrier was deduced to be a 54-kDa peptide of 492 residues.

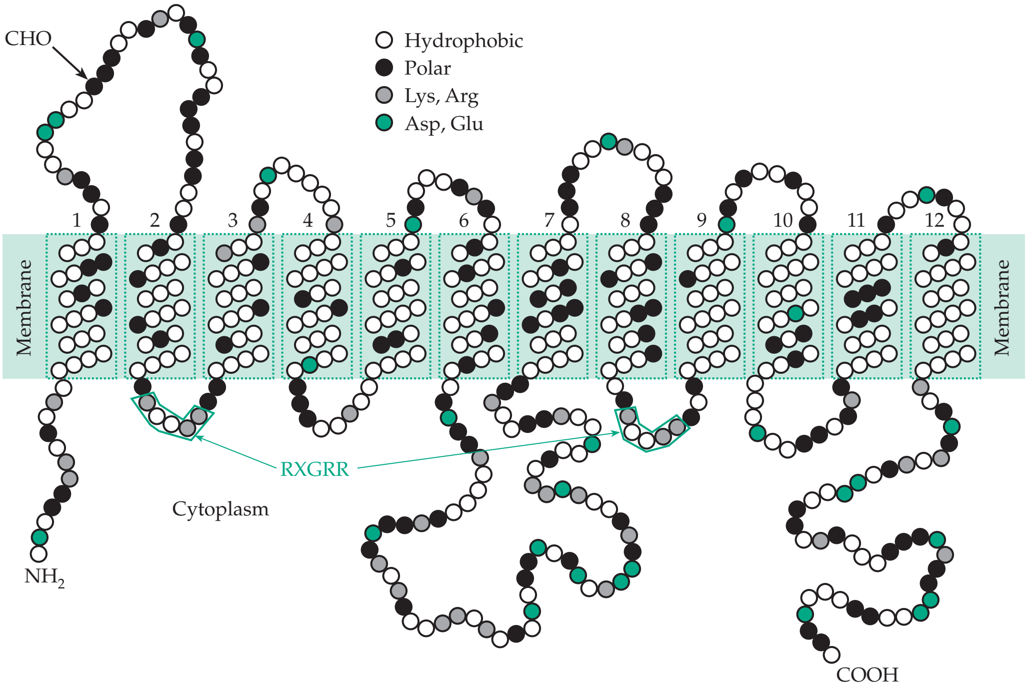
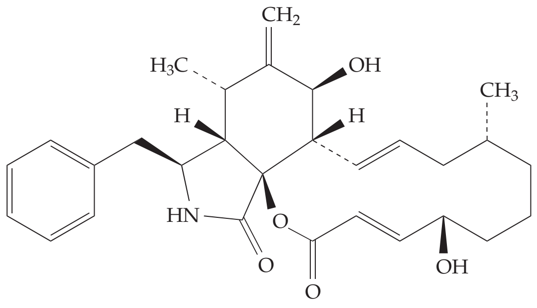


Figure 8-23 Predicted topology of the human glucose transporter GLUT1. The 12 predicted helices are numbered and the single external N-linked oligosaccharide is marked CHO. The sequence RXGRR (marked) is found in many 12-helix transporters. Its occurrence in these two positions suggests that the transporters may have evolved by duplication of a 6-helix motif. However, the human glucose transporters otherwise show no sequence similarity to other 12-helix transporters. After Bell *et al*.[398] See also Muekler.[400]

The glucose transporter is specifically inhibited by the fungal metabolite **cytochalasin B** which can also be used for photoaffinity labeling of the transporter.[397]



Cytochalasin B

There are at least six closely related facilitative sugar transporters in the human body.[398–400] GLUT1 is found not only in red cells but also in brain[401] and other tissues. GLUT2 is the principal liver transporter and GLUT3 is found along with GLUT1 and GLUT5 in the brain.[401] GLUT5 is primarily a fructose transporter.[402,403] The latter is present in spermatozoa, in which fructose transport is especially important (Box 20-A), and also in the small intestine. Both the GLUT1 and GLUT5 genes are overexpressed in breast cancer.[403] GLUT6 is an unexpressed pseudogene (see Chapter 27) and GLUT7 has been found only in liver microsomes. GLUT4 of skeletal and cardiac muscle and adipose tissue has received a great deal of attention because of its response to insulin[400,404-405a] and to exercise[406] (see Chapter 17).

All of the GLUT family appear to be 12-helix transmembrane-regulated, gated-pore proteins[407,408] with relatively short cytosolic N- and C-terminal ends. The proposed topology of GLUT1, which has been supported by much experimental data,[400,406] is shown in Fig. 8-23. However, the helices are thought to be bundled with the glucose channel centered within a helix bundle. The structure is unknown, but it could resemble the α,α-barrel of glucoamylase (Fig. 2-29), which has 12 helices, a glucose-binding site in the center, and a structure that could easily be modified to form a central pore.

Dehydroascorbate, the oxidized form of vitamin C (Box 18-D) is also transported into cells by GLUT1 and GLUT3.[409] A related transporter carries L-fucose into mammalian cells.[410] Another facilitates the uptake of galactose in yeast.[411]

Cotransport of sugars and other nutrients with H^+ or Na^+. Epithelial cells of the small intestine or of kidney tubules must take up glucose at low concentrations and discharge it into the bloodstream at a higher concentration.[412] This active transport is accomplished by cotransport of glucose with Na^+ in a 1:2 ratio, with the sodium ion concentration gradient across the membrane providing a usable source of energy amounting to 5.8 kJ / equivalent of Na^+.[413] The 662-residue human, sodium-dependent transporter **SGLT1** may have 14 transmembrane helices, 5 of which have been proposed to provide the sugar pathway.[414]

Cotransport with Na^+ is also observed for transport of many other sugars, amino acids, neurotransmitters, and cofactors.[415] A confusing variety of transporter molecules have been identified and are now being classified into families based on gene sequences as well as function.[416,417] Transporters from intestinal mucosal cells, kidney membranes, and synaptic endings of neurons have been studied most and have been the source for many of the cloned genes.[418] The Na^+-dependent transporters of neutral amino acids from these tissues have long been classified as system A, system B, and system ASC (alanine, serine, cysteine) according to substrate specificities.[417,419-420a] Both cationic amino acids and cystine are taken up by kidney tubules and intestinal epithelial cells by another Na^+-dependent transporter which is defective in human **cystinuria**, a common metabolic genetic problem.[421,422]

The brain contains several transporters specialized for rapid uptake of neurotransmitters glutamate and aspartate,[423–425] glycine,[426] γ-aminobutyrate (Gaba),[427,428] and catecholamines and also for taurine, L-proline,[429] serotonin,[430] and other substances. Many of these are not only Na^+ dependent but also require cotransport of Cl^-.[428–430] There are several different glutamate transporter genes with specialized distribution in the brain and other tissues.[423,424]

Cotransport of sugars with H^+ is especially common in bacteria[431] but also occurs in eukaryotes. For example, the alga **chlorella** employs hexose / H^+ symporters.[432] The most investigated H^+ cotransporter is probably the **lactose** (*lac*) **permease** from *E. coli* which enables *E. coli* to take up lactose and other β-galactosides from very dilute solutions.[433–436] From a variety of measurements it has been possible to propose a stacking arrangement for the 12 helices[434] and to identify groups that are essential for function of the 417-residue protein. Glutamates as well as an arginine, and a histidine, all in transmembrane regions, are essential and may be involved in the gating and transport functions.[436] The 469-residue **melibiose permease** of *E. coli* transports α-D-galactopyranosides, including meliobiose (Gal*p*α1 → 6 Glc) and raffinose (Gal*p*α1 → 6 Glc*p*α1 → 2 Fru*f*).[437,438] This permease will couple sugar uptake to either the H^+ or Na^+ gradient (or to a gradient of Li^+ but *not* of K^+).[437] Cells of *E. coli* also contain H^+ symporters for D-galactose, D-xylose, and L-arabinose that are homologous to the mammalian GLUT proteins.[439]

A series of specific H^+-linked cotransporters are found in the brush border membranes of small intestine and in kidney epithelial cells.[440,441] In green plants H^+-linked cotransport of amino acids is used in the distribution of amino acids synthesized in the roots and leaves to other parts of the plants.[442,443]

Antiporter or ion exchange transporters are also common. For example, *E. coli* uses a metal ion–tetracycline / H^+ transporter to carry the antibiotic tetracycline out of cells. This protein, when present, provides a high level of antibiotic resistance to the bacteria.[444]

4. Active Transport Systems

Both bacteria and eukaryotes also possess complex active transport systems for uptake of sugars, amino acids, and other nutrients and for pumping out toxic xenobiotics. One of the most important groups are the high affinity **ABC** (ATP-binding cassette) **transporters**, also called **traffic ATPases**.[445–448a] The *E. coli* genome contains genes for 80 ABC transporters, 54 of which had been identified before the genome sequence was completed.[447] In the human body an ABC transporter enables eukaryotic cells to pump out a large number of different drugs and other foreign compounds. This **multidrug resistance protein** (or P-glycoprotein) not only protects cells but also can seriously interfere with drug treatment. For example, cancer cells with increased amounts of this transporter often arise during chemotherapy. The transporter protein is a single 1280-residue, 170-kDa chain which probably has 12 transmembrane helices and two ATP-binding domains.[448,449] Other human ABC transporters include the antigen processing transporter **TAP** (Chapter 31), the 1480-residue anion transporter **CFTR**, which is defective in cystic fibroses (Box 26-C),[450] the erythrocyte glutathione-conjugate exporter (Box 11-B), and a long-chain fatty acid transporter of peroxisomes. ABC transporters usually consist of four domains. Two are hydrophobic intrinsic membrane domains, each with six membrane-spanning helices and two are peripheral membrane ATP binding domains. All four domains may be in a single peptide chain, as in CFTR, or they may be separate smaller proteins as in bacterial periplasmic permeases.[431,446]

Periplasmic permeases. Gram-negative bacteria contain numerous ABC transporters with components located on the periplasmic surfaces of their plasma membranes. Many of these can be dissociated from the surfaces by **osmotic shock**, i.e., by sudden changes in the osmotic pressure of the medium.[451,452] For example, cells of *E. coli* suspended in 0.5 M sucrose, treated with 10^{-4} M EDTA for 10 min, and then diluted with cold water release ~50 **binding proteins** that hold sugars, amino acids, ions, and other substances tightly with $K_f = 10^6–10^8\ M^{-1}$.[453,454] An example is the **L-arabinose binding protein** of *E. coli*, a 306-residue peptide organized into two large α / β units with a deep cleft between them. The sugar is bound into this cleft[455,456] by an extensive network of hydrogen-bonding interactions between polar groups in the sugar and in the protein similar to the network that binds either D-galactose (Fig. 4-18) or D-glucose to another of the periplasmic binding proteins.[457,458] Proteins with the same general architecture bind D-ribose, L-arabinose,[458] and maltodextrins.[459] The α1,4-linked maltodextrins, as well as maltose, are major nutrients for *E. coli* and enter the periplasmic space via the previously mentioned LamB porin. Other periplasmic proteins bind histidine,[460] basic amino acids,[461] branched chain amino acids, only leucine, oligopeptides, polyamines,[462] and the tetrahedral anions phosphate[454] and sulfate.[463] The sulfate^{2-} anion is held to its binding protein by seven hydrogen bonds – one from a serine –OH, one from an indole ring NH, and five from peptide NH groups, three of which are at the positive ends of α helices. No permanently charged groups nor cations nor water molecules come into contact with the SO_4^{2-}.[463] The phosphate-binding protein also binds a tetrahedral dianion HPO_4^{2-} but it doesn't bind sulfate. It forms numerous hydrogen bonds with its ligand and also an ion pair with a guanidinium group. An aspartate carboxylate hydrogen bonds to the –OH

BOX 8-D COLICINS: ANTIBIOTIC PROTEINS

Certain strains of *E. coli* and related bacteria synthesize proteins known as colicins that kill cells of other susceptible strains.[a–c] Three kinds of colicins, each encoded in its own small DNA plasmid (colicinogenic factor), are known. **Colicin E3** is a 58-kDa ribonuclease (RNase) that attacks 26S ribosomal RNA of susceptible bacteria;[d] **colicin E2** is a deoxyribonuclease (DNase) that cleaves the bacterial chromosome.[e] **Colicin E1** and its relatives, colicins A, B and Ia, Ib and N, attack the bacterial inner membranes and form lethal pores which allow K^+ and other ions to flow out of the cell. The presence of a single channel will kill the bacterium. The effect is similar to that of valinomycin.[f,g] The small colicin V is an 88-residue peptide antibiotic that is secreted by a dedicated ABC export system.[h] Colicins are members of a larger group of **bacteriocins**. One of these proteins, megacin Cx from *Bacillus megaterium*, kills bacteria of sensitive strains by blocking protein synthesis.[i]

The channel-forming colicins bind to bacterial surface molecules that serve as their receptors. For example, colicin N binds to the abundant *E. coli* surface protein OmpF. Interaction with a complex of membrane proteins known as tol Q, R, A, and B then leads to translocation across both outer and inner membranes and refolding of the colicin in a pore-forming conformation.[g] This mechanism is also used by colicins A, E, and K. The single-stranded bacteriophages M13, fd, and f1 "parasitize" the same transport system. Colicins B, D, I, and M enter bacteria with the aid of a second transport system consisting of proteins TonB, ExbB, and ExbD which participate in uptake of chelated iron (Chapter 16) and of vitamin B_{12}. This system is also parasitized by bacteriophages T1, T5, and φ80.[j]

The N-terminal portion of the 522-residue polypeptide chain of colicin E1 appears to be required for transport into the membrane and the central part for binding to the receptor; the channel-forming property is characteristic of the C-terminal region.[k] A similar organization has been established for the smaller colicin N:translocation domain, (residues 1–66), receptor domain, (residues 67–182), and pore-forming domain (residues 183–387).

The colicin E1 plasmid is a 4.43 MDa circular double stranded DNA molecule consisting of 6646 base pairs.[l] Only one site is susceptible to cleavage by the restriction endonuclease ECoR1 (Chapter 26) This feature has led to its widespread use in cloning of genes.

---

[a] Luria, S. E. (1975) *Sci. Am.* **233**(Dec), 30–37

[b] Parker, M. W., Tucker, A. D., Tsernoglou, D., and Pattus, F. (1990) *Trends Biochem. Sci.* **15**, 126–129

[c] Parker, M. W., and Pattus, F. (1993) *Trends Biochem. Sci.* **18**, 391–395

[d] Escuyer, V., Boquet, P., Perrin, D., Montecucco, C., and Mock, M. (1986) *J. Biol. Chem.* **261**, 10891–10898

[e] Dvhsllrt, K., and Nomura, M. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3989–3993

[f] Wiener, M., Freymann, D., Ghosh, P., and Stroud, R. M. (1997) *Nature (London)* **385**, 461–464

[g] Evans, L. J. A., Labeit, S., Cooper, A., Bond, L. H., and Lakey, J. H. (1996) *Biochemistry* **35**, 15143–15148

[h] Fath, M. J., Zhang, L. H., Rush, J., and Kolter, R. (1994) *Biochemistry* **33**, 6911–6917

[i] Brusilow, W. S. A., and Nelson, D. L. (1981) *J. Biol. Chem.* **256**, 159–164

[j] Derouiche, R., Bénédetti, H., Lazzaroni, J.-C., Lazdunski, C., and Lloubès, R. (1995) *J. Biol. Chem.* **270**, 11078–11084

[k] Griko, Y. V., Zakharov, S. D., and Cramer, W. A. (2000) *J. Mol. Biol.* **302**, 941–953

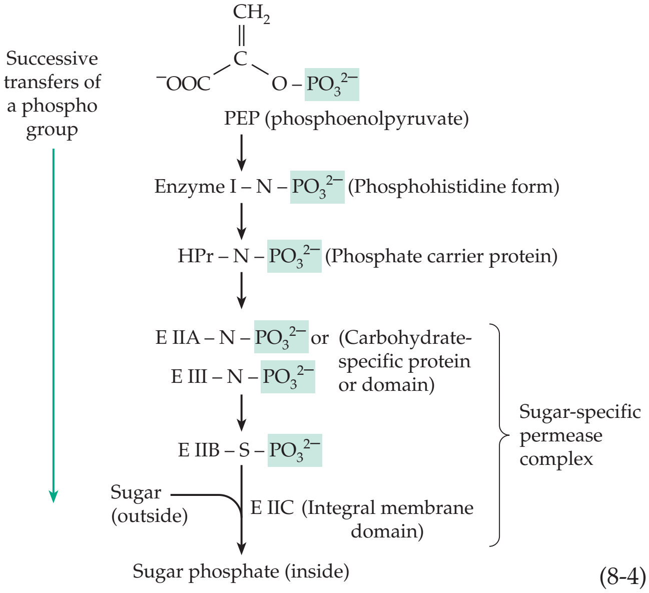
[l] Chan, P. T., Ohmori, H., Tomizawa, J.-I., and Lebowitz, J. (1985) *J. Biol. Chem.* **260**, 8925–8935

of the phosphate. Since the sulfate ion lacks the necessary H to form this bond it is excluded from the binding site.[454]

All of these binding proteins have similar architectures. The ligands fit into a groove between two domains as is seen in Fig. 8-24 for the histidine binding protein. The histidine is held by formation of carboxylate–arginine and $-NH_3^+$–aspartate ion pairs and an additional hydrogen bond to the imidazole. The proteins are able to bend in the hinge region between the two domains to give a better fit to their ligands.

The periplasmic binding proteins function together with the other subunits of the ABC transporter system. One of the best understood systems is encoded by the histidine transport operon of *Salmonella typhimurium*.[445,453] There are four genes: *hisJ* (encoding the histidine-binding protein), *hisQ*, *hisM*, and *hisP*. The Q and M proteins are hydrophic integral membrane proteins which interact with two copies of the P protein, which contains the ATP-binding motif, to form a His QMP_2 membrane complex. The soluble HisJ transfers its bound histidine to this complex, which with the hydrolysis of ATP supplying the driving force transfers the histidine across the membrane, presumably via a channel.[445] In *E. coli* the gene *malE* encodes the periplasmic maltose binding protein, while *malF* and *malG* are genes for the integral plasma membrane components. The *malK* gene encodes an ATP-binding protein homologous to the *hisP* protein. As with the histidine permease, a multiprotein complex $MalFGK_2$ is formed. It accepts maltose or a maltodextrin and ATP is hydrolyzed to drive the transport.[459] *MalT* encodes a positive regulatory protein that stimulates transcription of the mal genes.[464] Membrane transport systems are made more complex by the fact that several of the binding proteins serve also as receptors for stimulation of **chemotaxis** (Chapter 19).

The bacterial phosphotransferase system. A third system for uptake of sugars is utilized by *E. coli* and by many other bacteria. This phosphoenolpyruvate-dependent phosphotransferase system converts glucose, mannose, fructose, other sugars, or mannitol into their 6-phosphate esters, at the same time transporting the latter across the membrane (a group translocation).[431,465] Four proteins form a cascade (Eq. 8-4). Phosphoenolpyruvate, an intermediate in sugar metabolism whose high energy of hydrolysis provides the


Successive transfers of a phospho group
CH_2
C
^-OOC
$O-PO_3^{2-}$
PEP (phosphoenolpyruvate)
Enzyme I – N – PO_3^{2-} (Phosphohistidine form)
HPr – N – PO_3^{2-} (Phosphate carrier protein)
E IIA – N – PO_3^{2-} or
E III – N – PO_3^{2-}
(Carbohydrate-specific protein or domain)
E IIB – S – PO_3^{2-}
Sugar (outside)
E IIC (Integral membrane domain)
Sugar-specific permease complex
Sugar phosphate (inside)

(8-4)

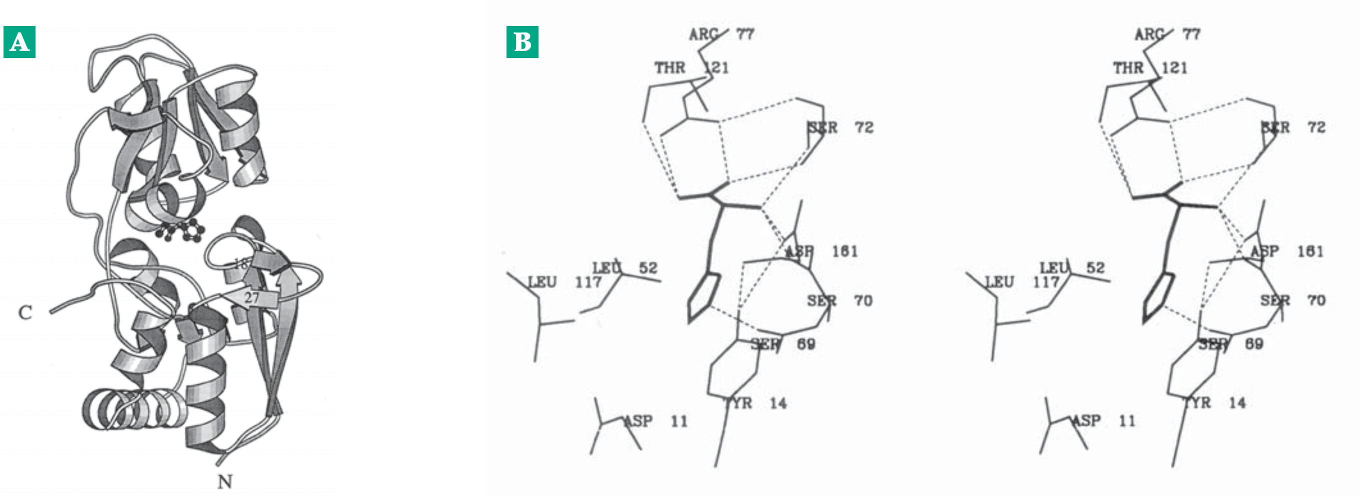

A
C
N
18
27
B
ARG 77
THR 121
SER 72
ASP 161
LEU 52
LEU 117
SER 70
SER 69
TYR 14
ASP 11
ARG 77
THR 121
SER 72
ASP 161
LEU 52
LEU 117
SER 70
SER 69
TYR 14
ASP 11

Figure 8-24 (A) MolScript ribbon drawing of the periplasmic histidine-binding protein HisJ, a component of an ABC transporter system of *Salmonella*. The bound L-histidine is shown as a ball-and-stick model. (B) Stereoscopic view of the histidine-binding site showing hydrogen-bonding interactions of protein side chains with the histidine. From Oh *et al.*[460] Courtesy of Giovanna Ferro-Luzzi Ames.

driving force, phosphorylates $N^{\varepsilon 2}$ of a histidine side chain on **enzyme I**, a large 64-kDa soluble membrane-associated protein.[466-467a]

The phospho group is then transferred sequentially to the small 88-residue dimeric phosphate carrier protein **HPr**, to the carbohydrate-specific membrane proteins IIA and IIB, and to the sugar being transported. A histidine side chain at position 15 of HPr is phosphorylated by PEP to form $N^{\delta 1}$-phosphohistidine.[468] The $NH^{\varepsilon 2}$ of the same histidine forms a hydrogen bond to C-terminal glutamate 85, which is also hydrogen bonded to an arginine side chain. This chain of interacting groups may function in the phospho transfer reactions.[469] Both enzyme I and HPr function in the transport of many ligands but there are specific **enzymes II** (EIIs) for each sugar or other ligand.[470] In *E. coli* there are at least 13 different PTS transporters.[471] A complete EII usually consists of three domains – EIIA (or EIII), EIIB, and EIIC. In some cases all three domains are in a single polypeptide chain, but in other cases they are individual proteins or some combination of individual and bifunctional proteins. The 637-residue mannitol-specific EII contains all three domains. Domains A and B are cytoplasmic while the N-terminal segments form the integral membrane C domain. The two phosphorylation sites are His 554 (EIIA) and Cys 384 (EIIB).[472] The glucose-specific EII from *E. coli* consists of two subunits, IIA and IICB.[471,473–475] The mannose transporter has three subunits representing domains IIAB, IIC, and an additional integral membrane subunit IID.[476] In addition to their direct transport functions, components of the PTS system play regulatory roles in chemotaxis, transcription, and control of other transporters.[477,478]

5. Transport of Ions

Cell membranes are impermeable to most ions. Only a small number of ions can enter cells readily and these usually do so with the assistance of protein channels or pores. The principal anion of plasma (Box 5A) is Cl^-, which passes through membranes readily by virtue of the presence of channel-forming proteins. Chloride ions are often distributed across membranes passively according to Eq. 8-5, which describes the **Donnan equilibrium**.[167,479,480]

$$[K^+]_i[Cl^-]_i = [K^+]_o[Cl^-]_o \tag{8-5}$$

Here the subscripts i and o refer to the inside and the outside of the cell, respectively. The potassium ion concentration within a cell is maintained at a high value by the operation of the $Na^+ + K^+$ pump and by the presence of nondiffusible anions within the cell. According to Eq. 8-5, the internal chloride concentration must be low, with the product of $[K]_i[Cl]_i$ equaling that of the low exterior $[K^+]$ and high exterior $[Cl^-]$. The internal $[Na^+]$ and $[Ca^{2+}]$ are both low, while the internal $[K^+]$ is high. These differences are also linked to the membrane potential (Eq. 8-2), which is ordinarily expressed as a negative voltage of the interior of a cell, mitochondrion, plastid, etc. with respect to a reference electrode in the external medium.

The maintenance of both the membrane potential and the steep gradients of ionic concentrations is essential to cells, both as a means of coupling metabolic energy to transport and other processes and for electrical signaling. The effects are most pronounced for mitochondrial membranes for which E_m may attain –140 to –170 mV and for plasma membranes of excitable cells such as neurons (E_m = –70 to –90 mV). For liver and kidney cells E_m of plasma membranes may be approximately – 35 mV and for erythrocytes only – 9 mV.[480]

In excitable cells electrical impulses are initiated by opening or closing ion channels. They are propagated along an axon by a complex sequence of opening and closing of voltaged-gated channels,[481] a process that is described in Chapter 30. All cells appear to also contain ATP-driven ion pumps as well as simple channels, cotransporter proteins, and ion exchangers.

Anions. Cell membranes have long been known to be relatively permeable to Cl^- and other small anions. However, the molecular basis of this permeability is quite complex.[482] Voltage-gated selective anion channels, often called **chloride channels**, are important in electrically excitable membranes, where they ensure a high resting chloride conductance and stability.[483] The gene for one of these channel proteins was first cloned from the electric ray *Torpedo*[484,485] and is designated *Clc*-0. Similar channels have been found in organisms ranging from bacteria to yeast, green plants, and vertebrate animals. The yeast genome contains just one *Clc* gene[486] but mammals have at least nine. One of these, *Clc*-2, is defective in **myotonia congenita**, a human disease of impaired muscle relaxation; and in similar diseases of mice, goats, and horses.[483,487] A mutation in the chloride channel *Clc*-5 causes kidney problems, including proteinuria, hypercalciuria, and kidney stones.[483]

Erythrocyte membranes permit rapid transport of anions to allow for the exchange of Cl^- within the red cells for HCO_3^- generated by tissue metabolism. The HCO_3^- binds to deoxyhemoglobin (Eq. 7-47) and is carried to the lungs, where it is released upon oxygenation of the hemoglobin. Then the reverse exchange of internal HCO_3^- for Cl^- occurs. This electroneutral ion exchange is mediated by the **band 3 protein** (Fig. 8-14) which contains a channel that allows anions but not cations to pass.[238,239,488] The band 3 protein, also known as **AE1** (anion exchanger 1), is found principally in red blood cells but is also present in kidney tubules.[489]

Related proteins occur in other tissues.[488] The 911-residue band 3 protein consists of two distinct parts of nearly equal size. The N-terminal portion is attached to the membrane skeleton (Fig. 8-16). The C-terminal part, which is embedded in the membrane, is thought to form 14 transmembrane helices and to contain the ion exchange channel or channels.[489a] As previously mentioned, defects in the N-terminal portion cause spherocytosis. The mutation Arg 589 His in the C-terminal half causes **renal tubular acidosis** in which the kidneys do not adequately remove acids from the body.[238,489] Band 3 proteins can also exchange phosphate, sulfate, and phosphoenolpyruvate for Cl^- or bicarbonate.

Another chloride channel, which is regulated by cyclic AMP (Chapter 11), functions in secretory epithelia. Its regulation is faulty in **cystic fibrosis** (Box 26-A), one of the most common human genetic defects, especially among persons of European descent.[490] As was previously mentioned, this **cystic fibrosis transmembrane conductance regulator** (CFTR) is a member of the ABC superfamily of transporters. The large 1480-residue protein apparently has two 6-helix membrane-spanning domains, two cytoplasmic nucleotide-binding domains,[491] and another large cytoplasmic regulatory domain.

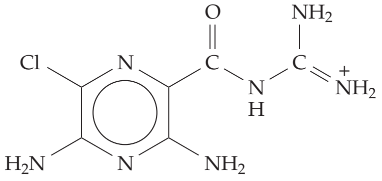
In addition to AE1 (band 3 protein), kidneys depend upon other modes of reabsorption of HCO_3^- from the proximal tubules. These include a Na^+/HCO_3^- cotransporter, which seems to be related to the AE family of ion exchangers. However, it transfers three HCO_3^- ions per Na^+ and is therefore highly **electrogenic**.[492,493] Transporters for phosphate, sulfate, and small organic anions are found in many organisms. Bacterial periplasmic transporters have already been described. Plants employ H^+/phosphate,[494] H^+/sulfate,[495] and H^+/nitrate[496] cotransporters. Phosphate transporters are probably essential to all organisms. One of the best known is the mitochondrial P_i/H^+ cotransporter which carries phosphate ions originating from hydrolysis of ATP to ADP + P_i back into the mitochondria.[497,498] See also Table 18-8. A human Na^+/P_i cotransporter in the kidney is also essential. An X-linked trait leading to inadequate synthesis of the transporter causes hypophosphatemic vitamin D-resistant **rickets**.[499]

Monocarboxylates such as lactate and pyruvate enter animal cells with the aid of monocarboxylate / H^+ cotransporters of low specificity.[500] A Cl^-/oxalate transporter is one of several ion exchange proteins in the kidney.[501] Transport systems for ADP, phosphate, dicarboxylates, and other anions are very active in mitochondrial membranes (Chapter 18).

Cation channels. When a nerve impulse passes along an axon gated pores or channels specifically permeable to Na^+ and K^+ open for short periods of time as a result of changes in membrane potential induced by the advancing wave of the action potential (Chapter 30). We have already considered the structure of a potassium ion channel (Fig. 8-21). However, there are at least 30 types of K^+ channels that can be distinguished.[502] Many of these appear to have similar channel structures but to serve a variety of purposes. While the K^+ channels of neurons are voltage gated many others are controlled by hormones, neurotransmitters, or mechanical stimuli.[503] One of the most investigated K^+ channels, known as K_{ATP}, sets the resting potential in the insulin-secreting β cells of the pancreas by facilitating a flow of K^+ *into* cells. Such channels, which help equilibrate intracellular and extracellular K^+ at near equilibrium, are found in cells with low negative values of E_m. They are called *inwardly rectifying*. When the internal glucose concentration in the β cells rises it initiates a complex signaling sequence involving blockage of the K_{ATP} channels by ATP, opening of voltage-sensitive Ca^{2+} channels, and insulin secretion.[504–506] The K_{ATP} channel is also blocked by **sulfonylureas**. As a result, these compounds induce insulin release and are useful in treatment of diabetes meltitus (Box 17-G). When its gene was cloned the **sulfonylurea receptor** was found to be a transmembrane protein of the ABC transporter family.[505] The potassium channel protein is another subunit of the transporter. Similar K_{ATP} channels are found in the kidneys[507] and also in embryonic cells in which they may participate in regulation of the cell cycle.[508]

Voltage-regulated **sodium channels** are the major participants in propagation of nerve impulses. The large 260-kDa α subunit of the sodium channel of nerve membranes contains four homologous repeat sequences, each of which may form transmembrane helices and also contain a loop that may participate in forming a pore similar to the K^+ pore of Fig. 8-21.[509-510a] However, the structure is uncertain.[511] The channel complex also contains 36- or 33-kDa β_1 and β_2 subunits that appear to be members of the Ig superfamily.

Epithelial cells contain a quite different Na^+ channel that participates in reabsorption of urinary Na^+ and in control of blood pressure (Box 22-D). The channel consists of at least three structurally similar subunits, each with two predicted transmembrane helices and a large extracellular domain.[512] These channels are blocked specifically by the diuretic compound **amiloride**.[513]



Amiloride, a diuretic

The best known sodium channel is, in fact, a general cation channel which is part of the "nicotinic" **acetylcholine receptor**. Acetylcholine is the principal excitatory transmitter in the peripheral nervous system and upon release from synaptic endings or neuromuscular junctions occupies receptors on the "postsynaptic" membranes of one or more adjacent neurons. When acetylcholine binds in the ion pore, the acetylcholine receptor opens and cations flow out, depolarizing the membrane. Under favorable circumstances, this initiates a nerve impulse (action potential) in the postsynaptic neuron. The receptor gene was first cloned from *Torpedo* and the receptor protein has been studied extensively.[382] The 290-kDa protein consists of five similar-sized subunits with an $\alpha_2\beta\gamma$ structure. These form a nearly symmetric fivefold oligomer with a pore in the center. Images of both the open and the closed states, obtained by electron microscopy at a resolution of 0.9 nm, suggest that the inner pore of ~2.6 nm diameter is formed by five α helices, one from each subunit. It is open when acetylcholine binds to the two α subunits and closes to a much smaller diameter when the acetylcholine leaves (and is destroyed by hydrolysis).

Calcium channels are a third major group of cation-selective channels.[514] As pointed out in Box 6-D, calcium ions are involved in a very wide range of signaling functions. These are discussed in several places in this book. Several of these functions depend upon voltage-gated Ca^{2+} channels. Muscle is rich in **L-type** or DHP-sensitive channels (Box 6-D) which play a role in transmission of nerve impulses to muscles by allowing rapid flow of calcium ions into cells from outside.[515]

The structure appears to be homologous to that of voltage-gated Na^+ channels with a large 170-kDa subunit with a fourfold repeat plus smaller subunits. The **ryanodine receptors** of muscle control the release of Ca^{2+} from stores in the endoplasmic reticulum[516,517] (see also Chapter 19). These receptors are ligand gated, being activated by cyclic ADP-ribose (Chapter 11). Two additional types of voltage-sensitive Ca^{2+} channels, **N** and **P**, are found in the central nervous system.[514,518] Another ligand gated calcium channel has been found in endothelial cells.[519] It is activated by **sphingosylphosphocholine**[519] rather than by cAMP-ribose or inositol triphosphate.[520] A sodium / calcium ion exchanger and cotransporter[521] utilizes the Na^+ gradient to exchange three external Na^+ ions for one internal Ca^{2+}.

Many aspects of calcium function are poorly understood. Among these is the role of a group of proteins known as **annexins** (formerly lipocortins, calpactins, endonexins, etc.).[522] The ten or more members of the annexin family[523–527] share the property of binding to phospholipid membranes in the presence of Ca^{2+}. One of the several proposed functions of annexins is formation of Ca^{2+} channels, a function that is suggested by the modular three-dimensional structures.[525,526] Other suggested functions include roles in membrane fusion, exocytosis, and adhesion.

Active transport of cations. Most organisms take up ions from their surroundings by active transport. Green plants extract essential nutrients from the extremely dilute solutions in contact with their roots. Microorganisms such as yeast and bacteria have the same ability, and specific concentrating systems for many ions such as K^+, Ca^{2+}, sulfate, and phosphate have been identified. The skin of a frog can take up Na^+ from a 10^{-5} M solution of NaCl and extrude it into the internal fluids whose Na^+ concentration may be greater than 0.1 M. Ions can also be concentrated from internal fluids and excreted at higher concentrations. Some seabirds and marine animals rid their bodies of excess salt by secretion from salt glands. The lining of the human stomach is able to concentrate hydrogen ions in gastric juice to ~0.16 M.

Organelles within cells have their own ion-concentrating mechanisms. Thus, mitochondria can concentrate K^+, Ca^{2+}, Mg^{2+}, and other divalent metal ions as well as dicarboxylic acids (Chapter 18). The entrance and exit of many substances from mitochondria appear to occur by exchange diffusion, i.e., by secondary active transport. Such ion exchange processes may also occur in other membranes.

ATP-driven ion pumps. Within virtually all cells the sodium concentration is relatively low, while that of potassium is high (Box 5-A). One theory[528] regards the cytoplasm as analogous to an ion exchange resin with fixed charges in a lattice. Highly crosslinked ion exchange resins exhibit specificity toward binding of certain ions; e.g., sulfonic acid resins tend to bind K^+ preferentially, while phosphonic acid resins tend to bind Na^+. Do proteins also prefer K^+ to Na^+?

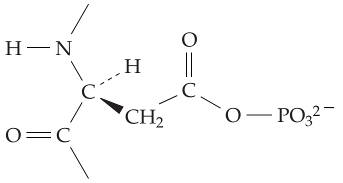
In contrast to the ion exchange theory, much evidence indicates that cells have an active **ion pump** that removes Na^+ from cells and introduces K^+. For example, the cytoplasm of the giant axons of nerves of squid can be squeezed out and replaced by ionic solutions. Erythrocyte ghosts can be allowed to reseal with various materials inside. Ion transport into or out of cells has been demonstrated with such preparations and also with intact cells of many types. Such transport is blocked by such inhibitors as cyanide ion, which prevents nearly all oxidative metabolism. However, the cyanide block can be relieved by introduction into the cells of ATP and other phosphate compounds of high group-transfer potential.

Uptake of K^+ by cells and extrusion of Na^+ from cells are also specifically blocked by "cardiac glycosides" such as **ouabain** (Fig. 22-12). Ouabain labeled with ^3H binds to the outer surface of cells, and from this binding it was estimated that erythrocytes possess 100–200 ion pumping sites per cell (~1 site / µm^2).[529] For the HeLa cell (a widely studied strain of human cancer cells) 10^5 to 10^6 sites / cell (~10 / µm^2) were found. Further experiments showed that in the presence of Na^+ within the cell and K^+ on the outside of the cell, ATP is hydrolyzed. The rate of hydrolysis was directly related to the concentrations of these two ions and to the number of ouabain binding sites and also required the presence of Mg^{2+}. These observations led to the concept of an **(Na^+ + K^+)- activated ATPase** (often abbreviated **Na^+, K^+-ATPase**) as synonymous with the membrane-bound ion pump. Within the cell Na^+ must be located on one side of the membrane and K^+ on the other to activate this enzyme. However, the purified enzyme would be expected to hydrolyze ATP in the test tube in the presence of $Na^+ + K^+ + Mg^{2+}$. Such a protein was isolated from several sources and has been studied intensively. It is an αβ mixed dimer with molecular masses of ~113 kDa for the α chains and ~55 kDa for the glycoprotein β chains.[530,531] The proteins may associate to $\alpha_2\beta_2$ tetramers in membranes. The genes for various isoforms of the proteins from several sources have been cloned and sequenced.[531,532] The large α subunit may span the bilayer of the membrane as many as ten times; the glycoprotein β subunit is thought to be largely on the outer surface and may have only one membrane-spanning helix.[533,534] A small 68-residue γ subunit copurifies with the pump protein. It may be involved in control of the ATPase, which has complex regulatory properties.[535,536] Sulfatides (ceramide galactose-3-sulfate) may also play a role in the enzymatic activity.[535]

The sodium–potassium pump displays a curious stoichiometry. *Three sodium ions are pumped from the inside and two potassium ions from the outside of a cell for each molecule of ATP cleaved*. Thus, an excess of positive ions is pumped out with the result that a negative charge develops inside the cell and a positive charge accumulates on the outside. This action of the Na^+, K^+-ATPase is the primary source of the membrane potential for most eukaryotic cells and is said to be **electrogenic**. Because the cell membrane is somewhat permeable to K^+, outward diffusion of K^+ through the "leaky" membrane along its concentration gradient helps to maintain the membrane potential as does inward leakage of Cl^-. At the same time, Na^+ diffuses inward, aided by the membrane potential. Even though the permeability of Na^+ is low, a steady state is reached at which the rate of passive inward diffusion of cations just balances the membrane potential set up by the active transport.

The energy for transport of Na^+ and K^+ by the ion pump is supplied by ATP. The Na^+,K^+- ATPase does not merely catalyze the hydrolysis of ATP but also couples its cleavage to the pumping of the ions. The pumping of sodium and potassium ions is one of the most important energy-requiring activities of cells. It is said to account for 23% of the ATP utilization in a resting human. *Thus, it constitutes an important fraction of the basal metabolic activity.*

The Na^+,K^+- ATPase is one of a family of over 50 ion pumps that are characterized by transfer of a phospho group from ATP to an aspartate side chain carboxylate in the invariant sequence DKTG to give an intermediate phosphoenzyme + ADP.[530,537]



Phosphoenzyme intermediate of Na^+,K^+- ATPase

These **P-type ATPases** are characterized by phosphoenzyme intermediates, by a conserved consensus sequence, and through inhibition by vanadate ion.[537–539] The structures are poorly known. Some consist of single chains (perhaps dimerized) and some have more than one chain. However, the major subunit always appears to have about ten transmembrane helices with a large ~430-residue cytoplasmic domain between the fourth and fifth helices. This domain contains the ATP binding site and the phosphoaspartyl group of the phosphoenzyme.[534] This is Asp 369 for the Na^+,K^+- ATPase.

In addition to the Na^+,K^+- ATPases there is a very active **Ca^{2+}-ATPase** which transports two Ca^{2+} from the inside of cells to the outside while returning two H^+ from outside per ATP.[540–543a] This is the primary transporter by which cells maintain a low internal [Ca^{2+}]. During its action it becomes phosphorylated on Asp 351. However, in neurons, in which the membrane potential is maintained at a high negative value by the sodium pump, an Na^+ / Ca^{2+} ion exchange plays an even more important role.[540]

Other P-type ATPases include the gastric H^+, K^+- ATPase, which acidifies the stomach and has a high degree of sequence homology with the Na^+, K^+- ATPase.[544] Secretion of HCl into the stomach apparently involves diffusion of K^+ together with Cl^- from the bloodstream through the cells lining the stomach. The K^+ is then pumped back into these cells in exchange for H^+ by H^+,K^+-ATPase.[545] The chloride channel may be in the same protein as the (K^+ + H^+) pump.[546] The kidney is the principle acid excretory organ of the body and as such also contains proton pumps. An electrogenic H^+- ATPase pumps H^+ alone outward through the plasma membranes of fungi and of green plants.[547,548] The resulting proton gradient may be used to provide energy for transport of other materials into cells. A group of metal ion P-type transporters carry copper and other nutrient ions into cells and extrude Cd^{2+} and other toxic ions.[539] Some **alkaliphilic bacteria** pump Na^+ to create a sodium ion gradient.[539a] All of these ion pumping systems require MgATP as the source of energy and function via phosphoenzyme intermediates. The ionic gradients generated can be used to move other ions or nonionic compounds into or out of cells by exchange or cotransport processes. For example, internal H^+ may be exchanged for external Na^+ in an exchanger-mediated process that assists in control of cytoplasmic pH.[549,549a] The reverse process in *E. coli*[549b] and many other bacteria provides the principal mechanism by which those cells export Na^+. This exchange is driven by the electrochemical gradient of the H^+ ion created by oxidative phosphorylation (Chapter 18). The Na^+ ion, in turn, can be used by bacterial cells to drive other uptake processes, e.g., sugar or amino acid-Na^+ cotransport.

What is the mechanism by which ATPase transporters function? We still do not know.[550] The pumping cycles for the Na^+,K^+- ATPase and the Ca^{2+}- ATPase are similar although different in details. The ATPases are reversible and with suitable ionic gradients will work as ATP synthases.[551] A strictly hypothetical model for the Na^+, K^+-ATPase is shown in Fig. 8-25. There are at least two conformations of the ion pump proteins.[552,552a] In one conformation the protein binds three sodium ions tightly, while in the other conformation it binds two potassium ions. The ATP operates the "motor" that carries out the conformational changes. In Fig. 8-25 the ion pump, in conformation A, is shown embedded in a membrane. In the center, perhaps between three or more transmembrane helices, there is a narrow cavity, perhaps resembling that of the K^+ channel (Fig. 8-21), into which chelating groups (e.g., C=O groups of the peptide chain) protrude. These groups form the three binding sites for the 0.19-nm-diameter Na^+ ion. The spontaneous binding of the sodium ions triggers a phosphorylation reaction by which a phospho group from the $MgATP^{2-}$ complex is transferred to the side chain carboxyl of the active site aspartate. This phosphorylation in turn triggers a change to the second conformation in which the channel to the outside is open and that to the inside is closed. At the same time the affinity for Na^+ is decreased over 100-fold and the sodium ions dissociate on the outside. The affinity for Na^+ may decrease because the diameter of the pore is increased to accommodate the larger 0.27-nm diameter of K^+ ions, perhaps by a twisting motion of the peptide chains that form the channel.

The next step is loading with two K^+ ions. The affinity for K^+ in the second conformation is high. The return to conformation 1 with release of K^+ to the inside is triggered by hydrolytic removal of the phospho group as inorganic phosphate (P_i). It may seem surprising that a channel could be opened and closed so readily with synchronous changes in the number and specificity of ion binding sites. However, recall the type of structural alteration occurring upon oxygenation of hemoglobin (Fig. 7-25). Rotation of the hemoglobin subunits with respect to one another causes small changes in the geometrical relationships of groups protruding into the central channel. This strongly

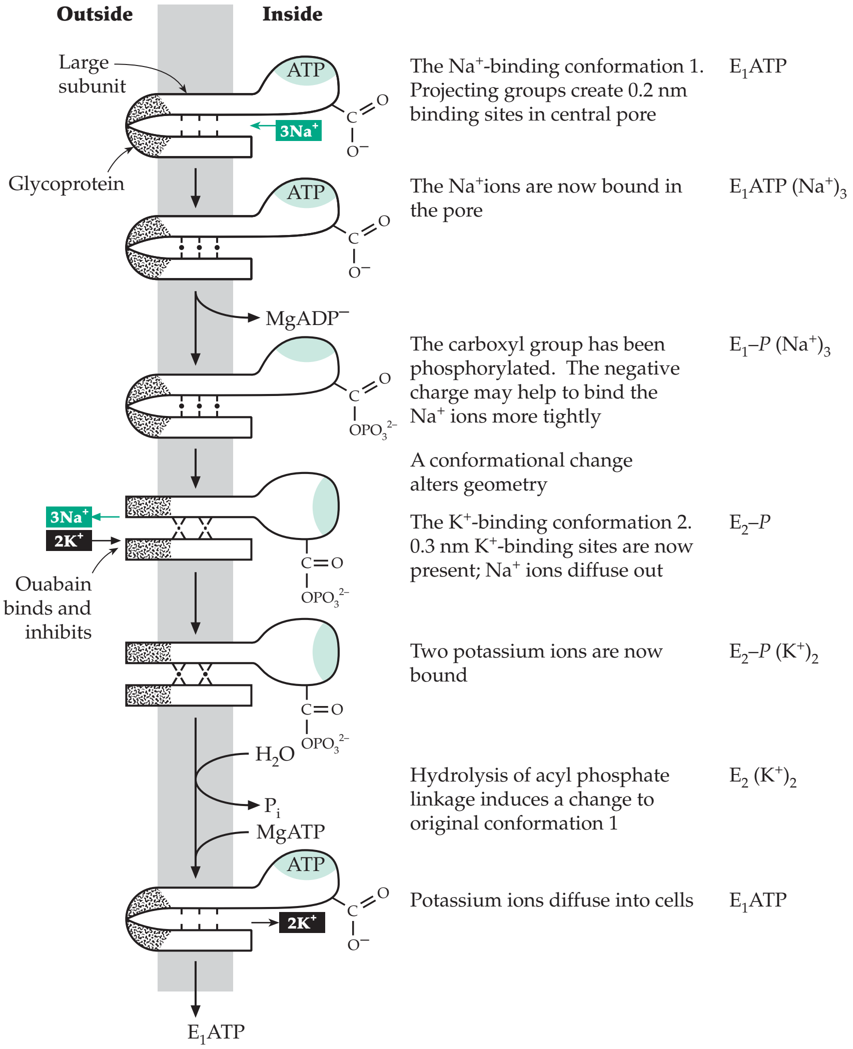

Outside
Inside
Large subunit
Glycoprotein
ATP
3Na⁺
C=O
O⁻
The Na⁺-binding conformation 1. Projecting groups create 0.2 nm binding sites in central pore
E₁ATP
The Na⁺ions are now bound in the pore
E₁ATP (Na⁺)₃
MgADP⁻
The carboxyl group has been phosphorylated. The negative charge may help to bind the Na⁺ ions more tightly
E₁–P (Na⁺)₃
OPO₃²⁻
A conformational change alters geometry
3Na⁺
2K⁺
Ouabain binds and inhibits
The K⁺-binding conformation 2. 0.3 nm K⁺-binding sites are now present; Na⁺ ions diffuse out
E₂–P
Two potassium ions are now bound
E₂–P (K⁺)₂
H₂O
Pᵢ
MgATP
Hydrolysis of acyl phosphate linkage induces a change to original conformation 1
E₂ (K⁺)₂
2K⁺
Potassium ions diffuse into cells
E₁ATP
E₁ATP

Figure 8-25 A strictly hypothetical model of a $Na^+ + K^+$ pump which operates by ATP-driven opening and closing of a channel at opposite ends and with alternate tight binding of Na^+ and K^+.

affects the binding of 2,3-bisphosphoglycerate. Very small movements could open up the Na^+ binding groups and create new binding sites for the larger K^+ ion, using in part the same chelating groups.

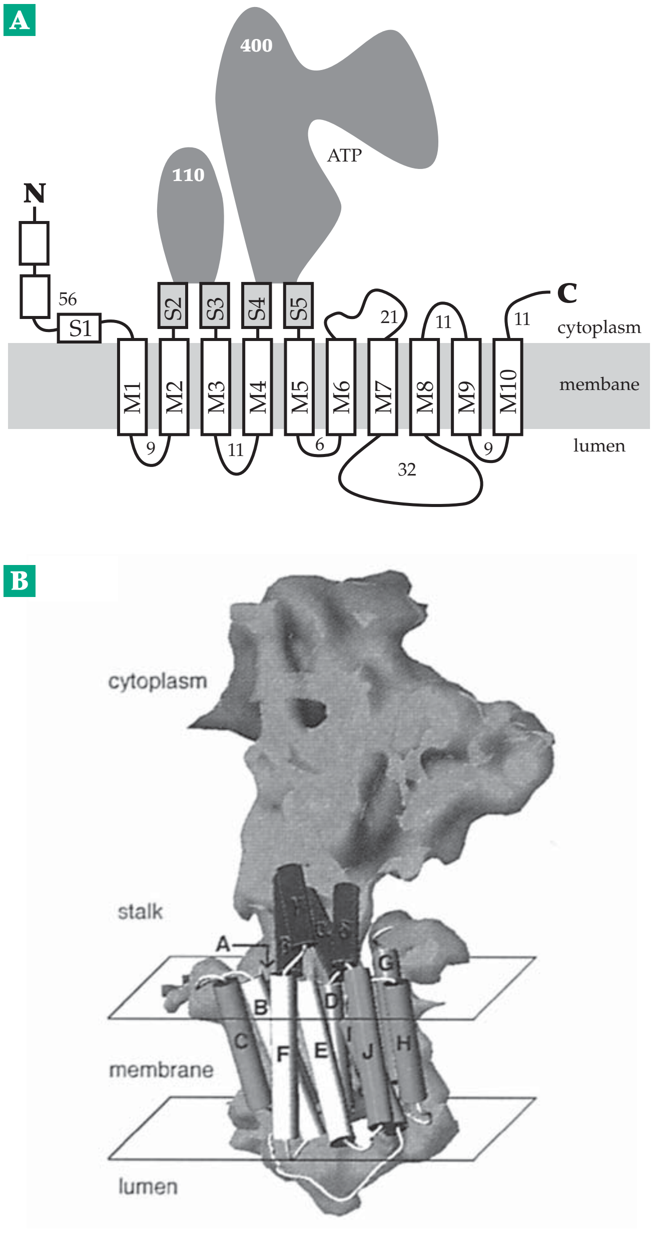


Figure 8-26 The structure of the 994-residue Ca^{2+}-ATPase of the endoplasmic reticulum of rabbit muscle at 0.8-nm resolution. (A) Predicted topology diagram organized to correspond to the electron density map prepared by electron crystallography of frozen-hydrated tubular crystals. The number of amino acid residues in each connecting loop is marked. (B) The electron density map with the predicted structure embedded. The relationships of the helices in (B) to those in (A) are not unambiguous. The helices marked B, D, E, and F in (B) may form the Ca^{2+} channel. The large cytoplasmic loops, which are black in (A), were not fitted. From Zhang *et al.*[553] Courtesy of David L. Stokes.

Images of both the Ca^{2+}-ATPase (Fig. 8-26)[553] and the H^+-ATPase of *Neurospora* plasma membranes[548] at 0.8 nm resolution reveal similar transmembrane regions and large cytoplasmic domains which are somewhat differently organized. The picture in Fig. 8-26 has been greatly clarified by determination of the structure by X-ray diffraction to a resolution of 0.26 nm.[553a,b,c] Two calcium-binding sites have been located in the transmembrane domain between the helices marked M4, M5, M6, and M8. The Ca^{2+} ions are apparently coordinated by side chains of Asp, Glu, Gln, and Thr. There are three cytosolic domains. The site of phosphorylation, Asp 351, lies within a large ~ 27 kDa **P** (phosphorylation) domain adjacent to the membrane. The ATP is held by a nucleotide-binding **N** domain which must at some point in the cycle move close to Asp 351 for phosphorylation to occur. The third cytosolic (**A**, actuator) domain is thought to be involved in control of the conformational alterations. The nucleotide-binding domain lacks the “P loop” characteristic of many ATPases and GTPases (see p. 648), but is homologous in its sequence to L-2-haloacid dehalogenase (Eq. 12-2)

Two other types of proton-pumping ATPases are considered in Chapter 18. One is the mitochondrial F_1F_0 ATPase, which ordinarily operates in the reverse direction as the body’s principal **ATP synthase**. The other type, which in some ways resembles the mitochondrial F_1F_0 ATPase, is the **vacuolar ATPase** (V-ATPase). These are true proton pumps which acidify vacuoles of plants and also lysosomes and phagocytic vacuoles.[554,555] They are also considered in Chapter 18.

6. Exocytosis, Endocytosis, and the Flow of Membrane Constituents

Observation of cells under the microscope with time-lapse photography reveals that the plasma membrane as well as the mitochondria and other organelles are in a constant state of motion. Mitochondria twist and turn and the surface membrane undulates continuously. Vesicles empty their contents to the outside of the cells, while materials are taken into cells through endocytosis. In addition, chemical evidence indicates a directed flow of the materials of which membranes are constructed from the endoplasmic reticulum (ER) to the Golgi vesicles, excretion granules, and plasma membrane (see Fig. 10-8). Along this route new materials are inserted from the cytoplasmic side of the membrane, while enzymes within the vesicles add glycosyl units and make other modifications. The plasma membrane surface area grows quite rapidly. In secretory cells fusion of secretion granules with the plasma membrane also adds additional material to the membrane.

Counterbalancing this expansion of the plasma membrane is active endocytosis of fluids and of solid materials from outside the cell. This not only brings new materials into the cell but also accomplishes removal of material from the plasma membrane and partial recycling of its components. One form of endocytosis is seen with the ameba.[555a] The cytoplasm flows around a smaller organism or other particle of food enclosing it in an internal membrane-bound compartment (**endocytic vacuole** or **endosome**). This vacuole then fuses with lysosomes which supply the necessary enzymes to digest the food. In a similar way phagocytic cells of our bodies engulf microorganisms or other particles to form **phagosomes** which, over a period of more than 24 hours, undergo extensive biochemical changes.[556] They acquire digestive enzymes, vacuolar ATPase,[557] other proteins needed to kill bacteria, and other parasites.

Uptake of smaller particles including protein molecules occurs by **micropinocytosis**, a process that can be seen only by electron microscopy. This often takes place via **coated pits**, indentations of ~ 0.3 µm diameter underlain by a thickened membrane.[558–559b] The pit membrane is also coated with protein molecules and appears to have many short bristles or spikes protruding into the cytoplasm (Fig. 8-27A). After endocytosis the coated pits become **coated vesicles** of 0.15 – 0.25 nm (Fig. 8-27B). Within a few seconds, however, these vesicles lose their coat and become endosomes.

The major protein making up the coat is the 180-kDa **clathrin**, but smaller 33- to 36-kDa peptides of several types also contribute.[561] The coat forms a "basket" with pentagonal and hexagonal faces surrounding the lipid bilayer of the vesicle. At each vertex of the basket is a "triskelion," a trimer of clathrin together with an equal number of the smaller chains. The smallest baskets consist of 12 pentagons plus 4, 8 or more hexagons, a relationship that allows formation of a variety of larger baskets.[562,563] Additional 50- and 100-kDa accessory proteins form a shell around the clathrin cage. That clathrin is essential for normal cell growth has been established by the observation that deletion of its structural gene from yeast is lethal.[564] The addition of more trimer units from a reserve of soluble clathrin in the cytoplasm allows the vesicles to develop and break off from the membrane. From studies with inhibitors it is evident that metabolic energy is required to drive the process.

Other vesicles are surrounded by nonclathrin membrane coats. Some of these originate from **caveolae** (little caves), which act in endocytosis, exocytosis, and transmembrane signaling.[564a,b,c] A **coatomer** complex of eight subunits with molecular masses from 20 to 60 kDa coats vesicles involved in transport between compartments of the Golgi.[565–567]

What is inside a coated vesicle? Cells take up a variety of peptide hormones and proteins. This usually occurs with the aid of specific receptor proteins located in or on the outside of the plasma membrane. Some of these, e.g., receptors for the low-density lipoprotein of plasma (Chapter 22), are clustered in coated pits. Other receptors, such as those for insulin or epidermal growth factor, are spread more evenly across the membrane but collect in coated pits when the hormone binds. Endocytosis provides a means for the cell to take up and in some cases destroy the hormone or the receptor or both.

Transmembrane proteins, including hormone receptors, are incorporated into coated vesicles with the help of **clathrin adapter proteins** (APs). These

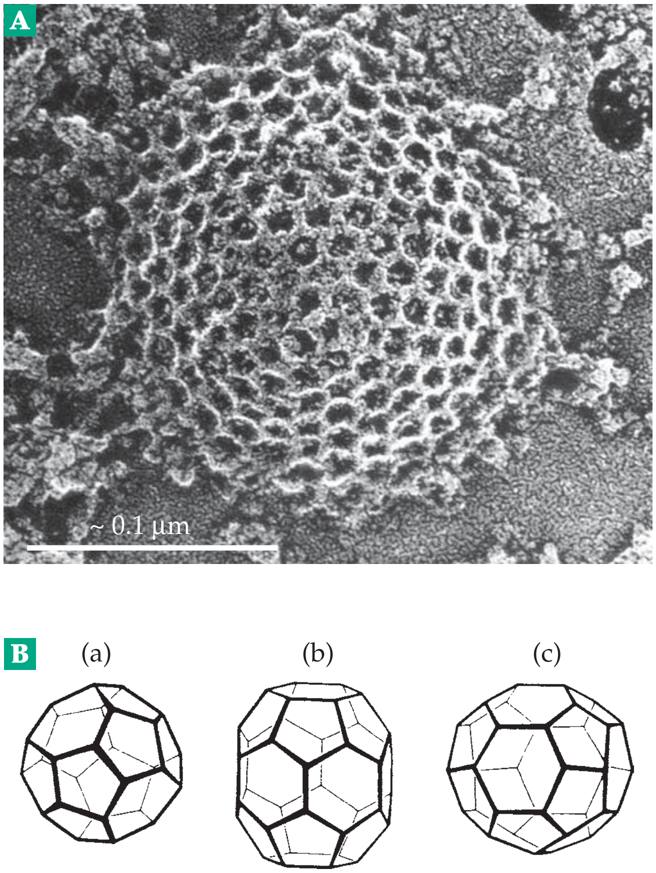


Figure 8-27 (A) Region of a coated membrane from fibroblasts at an intermediate stage of the budding process, demonstrated by deep etching and rotary replication (by J. E. Heuser[559b]). From Pearse and Bretscher.[560] Courtesy of Barbara Pearse. (B) Three structures identified among the smallest coated vesicles. Structure (*a*) contains 12 pentagons and four hexagons, the latter lying at the vertices of a tetrahedron; (*b*) has a barrel shape built of 12 pentagons and eight hexagons; structure (*c*) also has twelve pentagons and eight hexagons, but the latter are arranged in two arcs of four, related in the same way as the two parts of a tennis ball. Larger coats seem to be constructed on similar principles, with the addition of further hexagons. From Pearse and Bretscher.[560]

complex oligomeric proteins bind to recognition or "sorting" sequences such as dileucine on YXXϕ (Y = Tyr, X = any amino acid, ϕ = bulky hydrophobic).[567a–d] The adapter proteins also bind to clathrin, the N-terminal β-propeller domain associating with the sequence LϕXϕD / E of some AP adapters.[567c] Completion of a coated vesicle requires membrane fusion as the vesicle is pinched off from the membrane surface. A GTPase (Chapter 11) called **dynamin** as well as another protein **endophilin I** are required. Endophilin I is an acyltransferase able to transfer the fatty acyl group of arachidonoyl-coenzyme A to lysophosphatidic acid in the cytosolic surface of the membrane. This may change the curvature of the membrane, assisting in vesicle formation.[567e]

Once inside a cell the vesicles lose their coats to become endosomes which may then fuse with lysosomes or with Golgi membranes. The removal of a clathrin coat requires ATP as well as the chaperonin Hsp 70 (Chapter 10) and a coat protein called **auxilin**.[568] Triskelion is distorted and displaced from the clathrin cage. The interior of the newly formed endosome is quickly acidified by the action of a proton pump in the vesicle walls.[554,569] This sometimes leads to dissociation of enclosed receptors from their ligands and permits recycling of receptors and lipids of the vesicle membranes to the cell surface. This is the case for the low-density lipoprotein receptor.[570,571]

Exocytosis, by which the content of a secretion vesicle is released to the outside of a cell, is just as important as endocytosis. The process is sometimes very specialized. For example, the release of a nematocyst from *Hydra* (Fig. 1-13) can occur in about 3 ms.[572] Exocytosis involves **fusion** of membranes,[562,573] a process also occurring during the movement of endosomes along the endocytic pathway,[574] during vesicular transport between Golgi compartments, and in many other biological processes. Exocytosis is often triggered by the binding of Ca^{2+} to specific proteins of the vesicle wall and of the cytoskeleton (Chapter 7).[575] The fusion of membranes at several stages in the vesicle-mediated transport of materials between Golgi compartments requires a specific protein known as the *N*-ethylmaleimide-sensitive fusion protein (**NSF**).[576,577] A host of other specialized proteins are also involved[562,578] and are discussed in Chapter 10. See also Chapters 20 and 29.

D. Communication

External coats and cell walls help to control the access of materials to a cell. However, it is the outer surface of the plasma membrane that makes the cell's first contact with nutrients, hormones, and other important chemicals. The membrane must often not only detect these materials but also send signals to the interior of the cell and sometimes to adjacent cells. These signals may be about changes in pH or nutrient concentration or the presence of hormones, neurotransmitters, or harmful materials. For these reasons cell membranes contain many embedded receptors and signaling complexes. These are discussed in Chapter 11 and later sections of the book.

The plasma membrane also contains many "markers" of the individuality of the species or of an individual. These are chemical groupings that, in higher animals, can be recognized by the immune system as "self" rather than as a foreign invader. These surface markers may also be used by parasites as camouflage to evade the immune response of the host. Such chemical groupings on cell surfaces are often described as **antigenic determinants** and the molecules that carry them as **antigens**. Each antigenic determinant elicits the production of antibodies that will bind specifically to it. Over 250 different antigenic groups have already been described for the surface of the red blood cell. They determine the blood type. Groups on the surfaces of other cells determine whether a transplanted tissue will be rejected. Various proteins from plant and other sources act as **agglutinins** by binding to surface groups much as do antibodies. Viruses that attack cells may also become adsorbed onto specific surface molecules, which act as receptors.

Immunoglobulins can also be receptors. For example, molecules of IgE bound to basophils and the related **mast cells** of tissues serve as receptors for allergens. Binding of an allergen to the IgE molecules stimulates the release of granules containing histamine and other substances (Chapter 31).

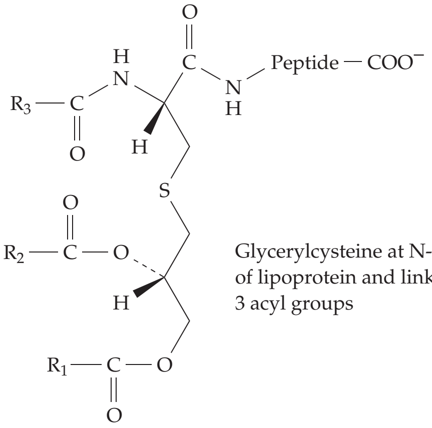
E. The Extracellular Matrix and Cell Walls

The surroundings of cells are extremely complex and vary from one organism to another and from one tissue to another. The principal function of cell walls and other surface coats is to protect cells against attack by organisms and against physical disruption.

1. The Structure of Bacterial Cell Walls

The plasma membrane of bacterial cells, other than the wall-less mycoplasmas and some archaebacteria, is surrounded by a multilayered wall which may be separated from the membrane by a thin **periplasm** (or periplasmic space). This can be seen most clearly in suitably prepared thin sections of cells of *E. coli* or other gram-negative bacteria as a relatively empty space of 11- to 25-nm thickness (Fig. 8-28).[579–581] The volume of this space (which may be filled with gelled material) depends upon the osmotic pressure of the medium. In *E. coli* it contains 20–40% of the total cell water. In gram-negative bacteria the innermost layer of the walls lies within the periplasm (Fig. 8-28). It is a porous network of a highly crosslinked material known as **peptidoglycan** or **murein**. The backbone of the peptidoglycan is a β-1,4-linked polymer of alternating *N*-acetylglucosamine and *N*-acetylmuramic acid residues. Alternate units of the resulting chitin-like molecule carry unusual peptides attached to the lactyl groups of the *N*-acetylmuramic acid units (Fig. 8-29). The peptide side chains are crosslinked as indicated in the figure. The peptides vary considerably in structure and crosslinkages.[582,583] In *E. coli* and other gram-negative bacteria the peptidoglycan forms a thin (2-nm) continuous network around the cell, but in gram-positive bacteria the highly crosslinked peptidoglycan forms a layer as much as 10 nm thick.[584] The peptidoglycan layer is surrounded by other layers whose structures vary from one organism to another, with the outermost antigenic layers being the most variable.

The outer membrane of gram-negative bacteria. Outside the murein layer of *E. coli* and other gram-negative organisms is a phospholipid-containing **outer membrane** which has the thickness and something of the structure of a typical biological membrane (Fig. 8-28).[585,586] This membrane is attached to the peptidoglycan layer with the aid of a small hydrophobic 58-residue **lipoprotein** whose N terminus contains a **glycerylcysteine** which carries three fatty acids connected by ester and amide linkages:



Glycerylcysteine at N-terminus of lipoprotein and linked to 3 acyl groups

The fatty acid chains are evidently embedded in the outer membrane as an anchor. About one-third of the lipoprotein molecules are attached covalently to the peptidoglycan through an amide linkage between the side chain amino group of the C-terminal lysine of the protein and a diaminopimelic acid residue of the peptidoglycan (Fig. 8-29). Thus, the protein replaces one of the terminal D-alanine residues of about one in ten of the murein peptides. There are ~2.5 x 10^5 molecules of the bound form of the lipoprotein per cell spread over a surface area of peptidoglycan of ~3 μm^2. They appear to be associated as trimers located primarily in the periplasmic space.[589]

About 10^5 copies per cell of the previously mentioned (Section C,2) larger 325-residue structural protein, **OmpA protein**,[350] after its gene symbol *ompA*

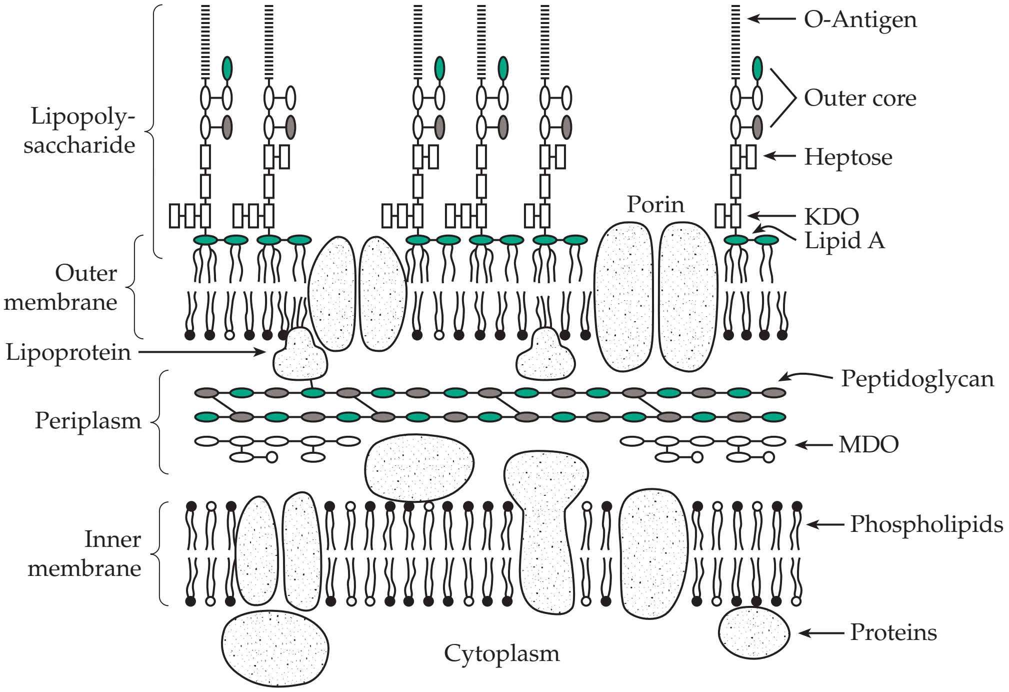


Figure 8-28 Schematic molecular structure of the *E. coli* envelope. Sugar residues are represented by ovals and rectangles. Circles represent polar head groups of phospholipids. MDO, membrane-derived oligosaccharides; KDO, 3-deoxy-*manno*-octulosonic acid (structures for KDO and heptose are in Fig. 4-15). From Raetz and Dowhan.[587]

(outer membrane protein A), are also present. The outer membrane contains almost the same number of molecules of two porins. Together with phospholipids or with the lipopolysaccharide discussed in the next paragraphs, the porins and OmpA protein associate in hexagonal arrays which provide the basic framework structure of the outer membrane.[585] Two of the outer membrane proteins have been shown to contain some of the modified lysine α-aminoadipic acid 5-semialdehyde (**allysine**).[590] The aldehyde groups of allysine are able to form crosslinks to other proteins as has been well established for collagen and elastin in the human body.

A characteristic feature of the outer surface of gram-negative bacteria is a **lipopolysaccharide**[586] that is anchored in the outer membrane. It was discussed briefly in Chapter 4 where the structures of the repeating oligosaccharides known as O-antigens are given. Figures 8-28 and 8-30 show the manner in which the oligosaccharide bearing the O antigen is attached to a lipophilic anchoring group that is embedded in the outer membrane of the bacteria.[591] The anchor, which is called **lipid A**, is a β-1,6-linked disaccharide of *N*-acetylglucosamine. It is also linked both to phosphate groups and to the fatty acyl groups that fit into the lipid bilayer of the membrane. In *E. coli* and *Salmonella typhimurium* four molecules of **3-D-hydroxymyristic** acid are joined by ester and amide linkages to the two GlcN units (Fig. 8-30). Other fatty acids, including lauric, myristic, and palmitic acids, are esterified to the hydroxyl groups of two or three residues of hydroxymyristic acid (Fig. 8-30). Lipid A from other gram-negative bacteria is similar but with variations in the fatty acyl group composition and linkages to the carbohydrate.[592–594]

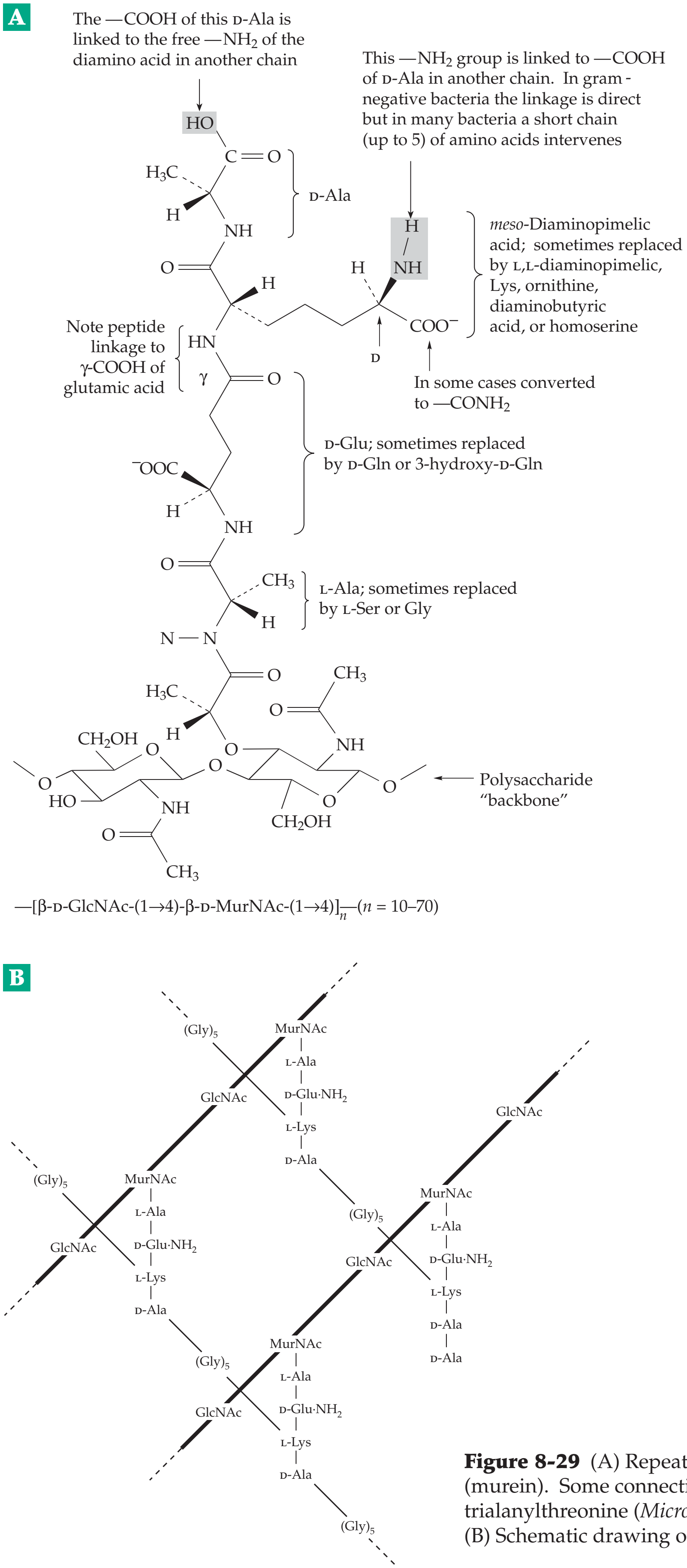

A
The —COOH of this D-Ala is linked to the free —NH2 of the diamino acid in another chain
HO
C=O
H3C
H
NH
D-Ala
This —NH2 group is linked to —COOH of D-Ala in another chain. In gram-negative bacteria the linkage is direct but in many bacteria a short chain (up to 5) of amino acids intervenes
H
NH
H
meso-Diaminopimelic acid; sometimes replaced by L,L-diaminopimelic, Lys, ornithine, diaminobutyric acid, or homoserine
O
H
COO−
D
In some cases converted to —CONH2
Note peptide linkage to γ-COOH of glutamic acid
HN
γ
O
D-Glu; sometimes replaced by D-Gln or 3-hydroxy-D-Gln
−OOC
H
NH
O
CH3
H
L-Ala; sometimes replaced by L-Ser or Gly
N—N
O
CH3
H3C
O
H
NH
CH2OH
O
O
O
O
HO
NH
O
CH2OH
O
Polysaccharide "backbone"
O
CH3
—[β-D-GlcNAc-(1→4)-β-D-MurNAc-(1→4)]n—(n = 10–70)
B
(Gly)5
MurNAc
L-Ala
GlcNAc
D-Glu·NH2
L-Lys
D-Ala
GlcNAc
(Gly)5
MurNAc
L-Ala
GlcNAc
D-Glu·NH2
L-Lys
D-Ala
(Gly)5
MurNAc
L-Ala
GlcNAc
D-Glu·NH2
L-Lys
D-Ala
D-Ala
(Gly)5
MurNAc
L-Ala
GlcNAc
D-Glu·NH2
L-Lys
D-Ala
(Gly)5

Figure 8-29 (A) Repeating unit of structure of a bacterial peptidoglycan (murein). Some connecting bridges are pentaglycine (*Staphylococcus aureus*), trialanylthreonine (*Micrococcus roseum*), and polyserine (*S. epidermis*). (B) Schematic drawing of the peptidoglycan of *S. aureus*. From Osborn.[588]

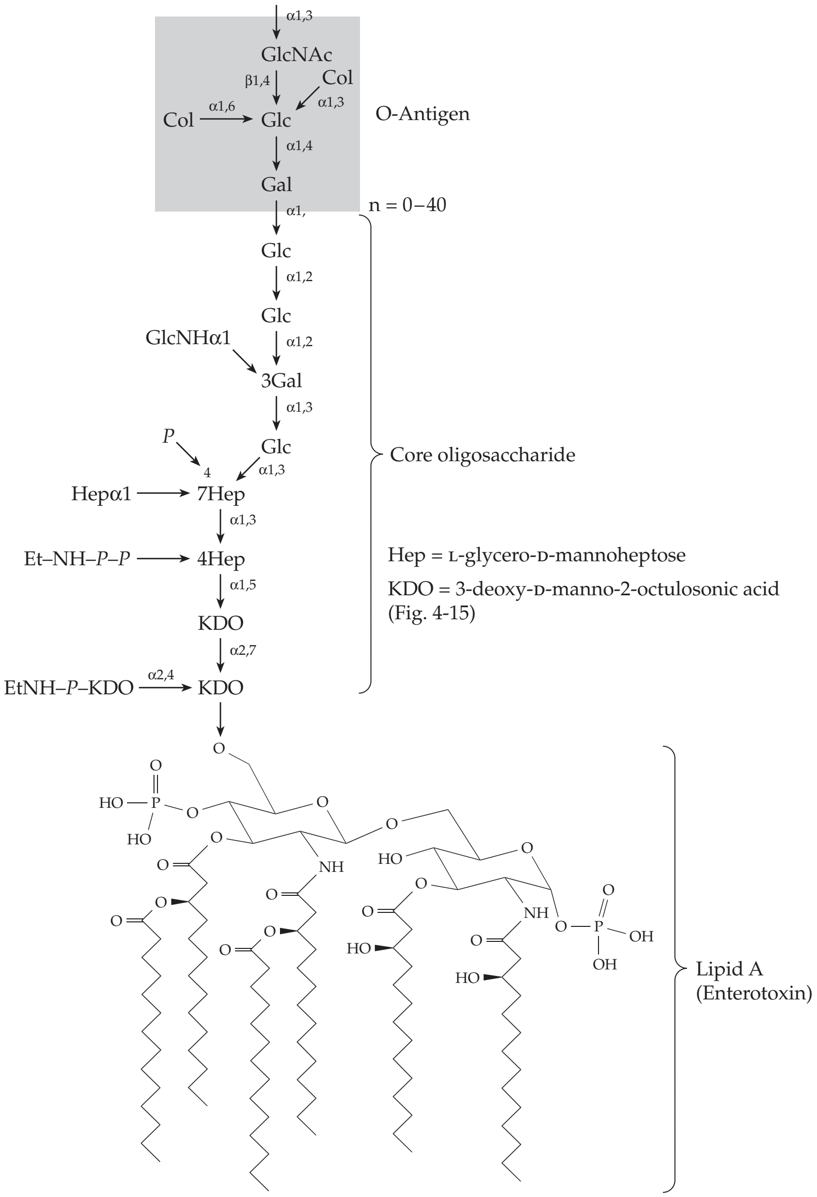
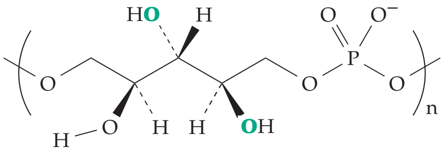

α1,3
GlcNAc
β1,4
Col
α1,3
α1,6
Col
Glc
O-Antigen
α1,4
Gal
α1,
n = 0–40
Glc
α1,2
Glc
GlcNHα1
α1,2
3Gal
α1,3
P
Glc
4
α1,3
Hepα1
7Hep
α1,3
Et–NH–P–P
4Hep
α1,5
KDO
α2,7
α2,4
EtNH–P–KDO
KDO
Core oligosaccharide
Hep = L-glycero-D-mannoheptose
KDO = 3-deoxy-D-manno-2-octulosonic acid (Fig. 4-15)
O
HO
P
HO
NH
HO
NH
OH
OH
Lipid A (Enterotoxin)

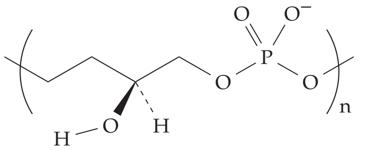
Figure 8-30 Structures of the lipopolysaccharides of the outer membrane of *E. coli* and *S. typhimurium* including the bilayer anchor lipid A. For structures of L-glycero-D-mannoheptulose, KDO, and colitose, see Fig. 4-15.

In *Bacillus subtilis* D-alanine is attached to one of the green oxygen atoms in at least half of the units


HO
H
O
O⁻
O
P
O
O
n
H
O
H
H
OH

Ribitolteichoic acids

In *B. subtilis* β-D-glucose is attached here



Glycerolteichoic acids

In *Lactobacillus arabinosus* D-alanine in ester linkage occupies this position but is replaced by β-D-glucose on about one unit in nine

Many differences are also found in the "core" oligosaccharide (Fig. 8-30) and in the O-antigens.[594–596]

Why do cells of *Salmonella* have a thousand distinguishable surface antigens, many based on differences in the O-antigens? The ends of these carbohydrate clusters are the groups to which the antibodies in animals clamp themselves if the bacteria enter the bloodstream. Mutants known as R forms (because of the growth as rough colonies on agar plates) completely lack the outer O-antigen. The R mutants of *Salmonella* are nonpathogenic, whereas the smooth strains with intact O-antigen often cause illness. Perhaps, if the O-antigen has the right cluster of sugar rings at the ends, the host organism does not recognize it as dangerous. This is only part of a continuous battle between the immune system of the body and camouflaged surfaces of attacking pathogens.

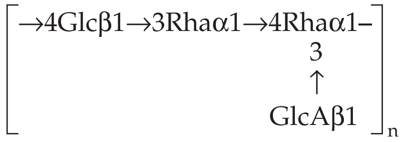
The lipopolysaccharides of bacterial surfaces have been identified as the "**endotoxins**" that cause many of the worst effects of gram-negative bacterial infections and that are often lethal.[597] Although the name endotoxin came from the assumption that there was an internal heat-resistant toxin that was released from the bacteria, the toxin is part of the bacterial cell wall. While the O-antigens provide for innumological recognition, it is the *lipid A* part that is responsible for the toxicity and unusually strong fever-inducing properties of the lipopolysaccharide.[592–594,597]

Teichoic and teichuronic acids. The cell walls of gram-positive bacteria are composed of a thick peptidoglycan layer which also contains proteins and additional polymers known as **teichoic acids** and **teichuronic acids**. In some species these account for 50% of the dry weight of the cells.[598,599] Teichoic acids are high polymers of the following general types: They are often attached through phosphodiester linkages to *N*-acetylglucosamine or a disaccharide which, in turn, is attached to muramic acid residues of the peptidoglycan. Since the teichoic acid is uniformly distributed in its attachment to the peptidoglycan, it must be intimately associated with peptidoglycan throughout the wall.

Teichoic acids are often covalently attached to glycolipids which are part of the plasma membrane. For example, the glycerolteichoic acid of *Streptococcus faecium* contains about 28 monomer units of glycerol phosphate, approximately 60% of which carry residues of kojibiose (Gluα1 → 2Glu) as a phosphatidylkojibiosyl diacylglycerol membrane anchor.[600] Teichuronic acids contain uronic acids:

$[\rightarrow 4(N\text{-acetyl})\text{–ManA}\beta 1\rightarrow 6\text{Glc}\alpha 1\text{–}]_n$

From *Micrococcus luteus*[601]



$$\left[\rightarrow 4\text{Glc}\beta 1\rightarrow 3\text{Rha}\alpha 1\rightarrow \underset{\underset{\text{GlcA}\beta 1}{\underset{\uparrow}{3}}}{4\text{Rha}\alpha 1}\text{–}\right]_n$$

From *Bacillus megaterium*[602]

Nearly 800 of the foregoing tetrasaccharide units are joined to give a large, densely packed, ~500-kDa polymer. The reducing end of the polymer is covalently attached to the peptidoglycan layer.

Mycobacteria. Most mycobacterial species seem to be harmless saprophytic soil bacteria of the gram-positive group. However, tuberculosis (caused by *Mycobacterium tuberculosis*) may infect one-third of the inhabitants of the earth[603] and kills about three million people per year.[604] Leprosy, caused by *M. leprae*, affects 10 million or more.[605] Most antibiotics are not effective against mycobacteria because of the unusual nature of their multilayered cell walls. The muramic acid of their peptidoglycans contains glycolyl groups instead of acetyl groups, perhaps providing extra hydrogen bonding within the wall.[604] Most characteristically, they are rich in lipid materials. Among these are a unique polysaccharide, a highly branched **arabinogalactan**, that is covalently attached to muramic acid residues of the peptidoglycan. Clusters of arabinofuranosyl units at the nonreducing ends of the chains are esterified with **mycolic acids** (Section A,1).[10,604,605] Mycobacteria also contain **lipoarabinomannans** that act as major antigens.[603,606] Some mycobacterial species synthesize small glycopeptidolipids that may disrupt cell membranes of hosts.[607] Mycobacteria don't produce toxins of usual types but cause slow, long-lasting infections.

Other bacterial coats. Archaebacteria not only have unusual plasma membranes that contain phytanyl and diphytanyl groups (Section A,3)[608] but also have special surface layers (**S-layers**) that may consist of many copies of a single protein that is anchored in the cell membrane.[609] The surface protein of the hypothermic *Staphylothermus marius* consists of a complex structure formed from a tetramer of 92-kDa rods with an equal number of 85-kDa "arms."[610,611] S-layers are often formed not only by archaebacteria but also by eubacteria of several types and with quite varied structures.[612–614] While many bacteria carry adhesins on pili, in others these adhesive proteins are also components of surface layers.[615] Additional sheaths, capsules, or slime layers, often composed of dextrans (Chapter 4) and other carbohydrates, surround some bacteria.

2. The Surroundings of Animal Cells

Cells in the external epithelial layers are always surrounded by a protective covering of some kind. Our own skin is made up of specialized cells which become filled with microfibrils of keratin as they move outward and become the relatively dry nonliving external surface (Box 8-F). Internal epithelial cells secrete protein and carbohydrate materials that form a thin **basement membrane** around the exposed parts of the cells. The **connective tissue** that lies between organs and which also includes tendons, cartilage, and bone consists of a relatively small number of cells surrounded by a "matrix" consisting of the protein fibers collagen and elastin in a "ground substance" rich in proteoglycans (Chapter 4).[616–618] In bone, the calcium phosphate is deposited within this matrix.

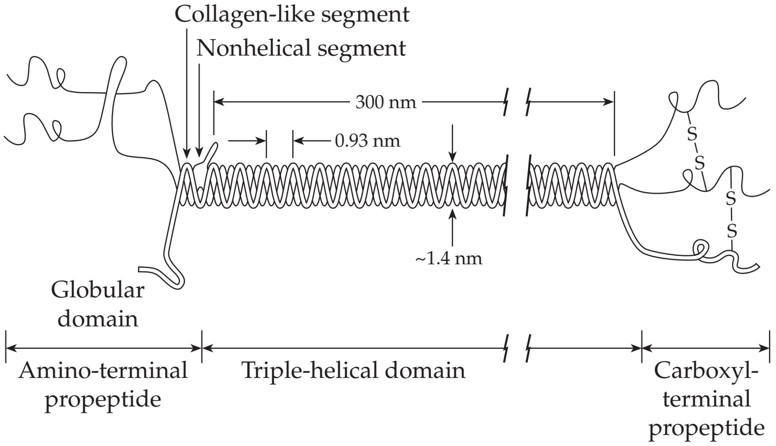
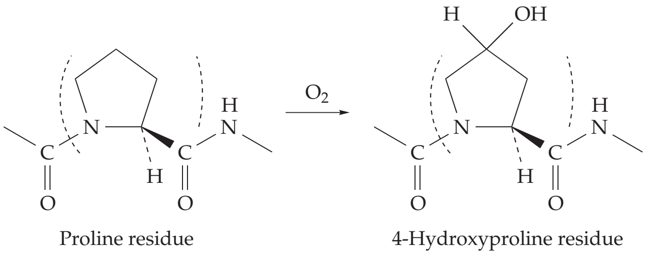


Figure 8-31 Schematic representation of type I procollagen. The molecule is composed of two identical proα1 chains (solid lines) and one proα2 chain (dashed line). In addition to the central triple-helical region that gives rise to the collagen molecule, as portrayed in Fig. 2-23(C), the precursor contains amino- and carboxyl-terminal non-triple-helical domains. The amino-terminal domain is composed of a presumably globular region, a short collagen-like segment, and a non-triple-helical region in which cleavages by the amino-terminal protease occur. Interchain disulfide bonds are limited to the carboxyl-terminal domain. The short telopeptides at the ends of collagen α chains represent the residual sequences of the linkage regions between the collagen helix and the terminal domains. After Prockop[628] and Byers.[625]

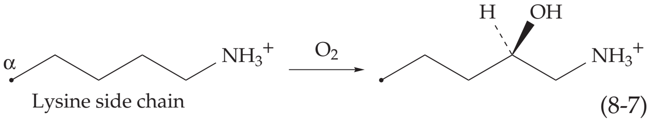
The collagens. The most abundant proteins in the body are the collagens,[619–624] a family of closely related materials that account for 20% of the total protein in higher animals. Collagens make up much of the organic mass of skin, tendons, blood vessels, bone, the cornea, vitreous humor of the eye, and basement membranes. They are found in every metazoan phylum studied. There are at least 16 types in the human body.[625] Type I collagen, which accounts for 90% of the total, is the major form occurring in skin, tendon, and bone. It is synthesized by the fibroblasts and is excreted into the extracellular space where it is polymerized into a durable long-lived material. Collagen II is found exclusively in cartilage and the vitreous humor of the eyes. Form III is located in blood vessels and intestines and is prominent in embryonic tissues. Collagen IV is the major form found in basement membranes. Some other collagens and characteristic features are listed in Table 8-4.

All collagens contain the triple-helical structure shown in Fig. 2-23. Collagen I consists of two chains of one kind (α1) and one of another (α2), while most other collagens have three identical chains. In these chains over 1000 residues have the characteristic GlyXY sequence.[624,626] At each end of the rodlike molecules short segments of the peptides fold into globular domains. For type I collagen 16 residues at the N termini and 25 at the C termini form these domains. Collagens are synthesized as intracellular precursors known as **procollagens**. The three chains of procollagens are much longer than in the mature proteins and have larger non-triple-helical ends (Fig. 8-31). The C-terminal extensions are crosslinked by S–S bridges.[625,627] Synthesis of the collagen chains requires at least six minutes after which they are released into the cisternal space of the ER, associate, and become crosslinked (Fig. 8-32). This crosslinking ensures that the three chains remain in proper register in the triple helix while the procollagen is converted into collagen, a process involving additional crosslinking.

Before this “maturation” can occur there must be other modifications to procollagen. These begin while the peptide chains are still attached to ribosomes of the rough ER. Hydroxylases (Chapter 18) localized in the membranous vesicles of the ER convert some of the proline and lysine residues of the procollagen chains into **4-hydroxyproline**[625,629,630] and **hydroxylysine** (Eqs. 8-6 and 8-7). Lesser amounts of 3-hydroxyproline are formed. About 100 molecules of 4-hydroxyproline and 50 of 5-hydroxylysine are created in each α1 chain.



(8-6)



(8-7)

Galactosyl units are then transferred onto some of the hydroxyl groups of the hydroxylysine side chains, and glucosyl groups are transferred onto some of the galactosyl groups. This glycosylation may prevent incorrect association of the procollagen molecules.

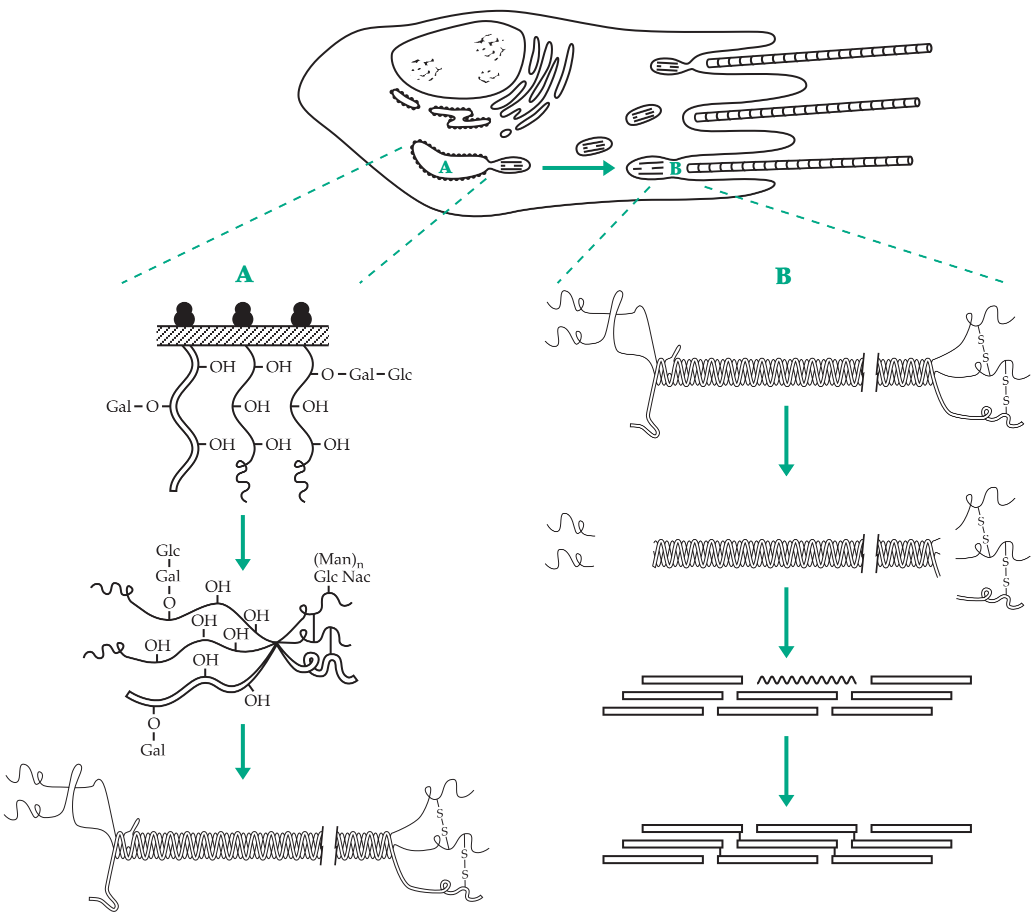
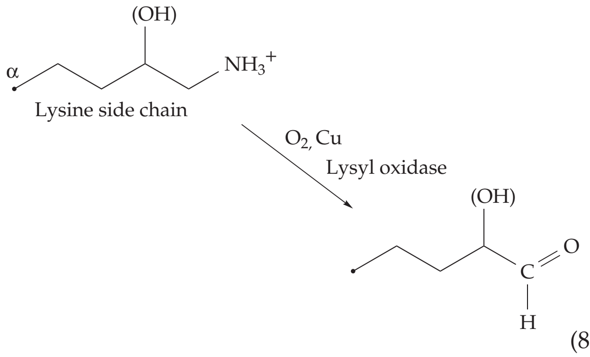


Figure 8-32 Scheme summarizing the biosynthesis of a fibrillar collagen by a fibroblast. (A) Assembly of pro-α chains in cisternae of the rough endoplasmic reticulum; posttranslational hydroxylations and glycosylations; association and disulfide bonding of C-propeptides; and zipper-like folding of the triple helix by nucleated growth. (B) Proteolytic processing of procollagen to collagen; self-assembly of fibrils by nucleated growth; and covalent crosslinking of the fibrils. The collagen molecule is first shown as a triple helix and then either as a *wavy line* to depict a molecule assembling on the surface of a fibril or as a *rectangle* to depict the quarter-staggered assay of monomers in a fibril. The proteolytic processing of procollagen and assembly of fibrils may occur within crypts of fibrils as shown here or perhaps at some distance from the cell. After D. J. Prockop.[628]

Within the extracellular space two **procollagen peptidases** act to cleave a 35-kDa peptide from the C terminus[631] and a 20-kDa peptide from the N-terminal end of each of the three chains of the secreted procollagen. The amino acid composition of the peptides removed is quite unlike that of the remaining collagen monomer (also called tropocollagen) which contains one-third glycine and much proline.

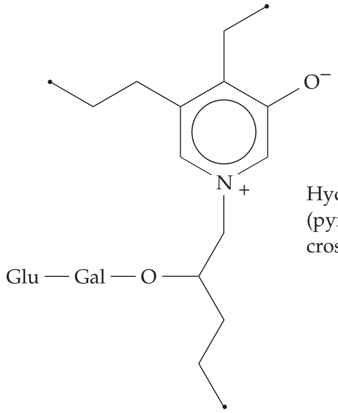
The three-stranded monomers of collagens I–III are rods of dimensions ~1.4 x 300 nm (Fig. 8-23). When they reach their final location they associate and become crosslinked to form strong fibers with diameters ranging from 8 nm to 0.5 m. Tendons tend to contain large fibrils, while those in bone are small. The smallest 8-nm fibrils must contain about 20 triple helices in a single cross section but the successive monomers are staggered by 6.4–6.7 nm (234 residues for type I collagen) in such a way that 3-nm gaps are left between the ends (Fig. 8-32). These gaps give rise to the characteristic banding pattern seen in the fibrils in Fig. 2-23D. Finer analysis of the bands together with the known sequences shows that the bands also reflect the locations of residues with charged side chains.[632] These ionized groups are thought to provide electrostatic stabilization to the fibrils.[633] The exact packing of the collagen molecules into sheetlike or microfibrillar crystalline arrays is still uncertain.[634] However, there is agreement on the staggered arrangement[634a] and upon the fact that precisely formed crosslinks with neighboring rods prevent the gaps from weakening the fibrils.

Crosslinking of collagen is initiated by oxidation of some of the lysyl and hydroxylysyl side chains from amino groups to aldehyde groups under the action of a copper-containing oxidase (Eq. 8-8, Chapter 18). The aldehyde groups enter into a variety of reactions that lead to crosslinking of the collagen monomers and to the formation of insoluble fibers.[635] One reaction is an aldol condensation followed by elimination of water (Eq. 8-9, step *a*). If one of the two aldehydes involved in the condensation is derived from hydroxylysine and the other from lysine, two isomeric condensation products are formed. The aldol condensation product can react further: An imidazole group from a histidine side chain can add to the carbon–carbon double bond and (either before or after this reaction) another lysine side chain can form a Schiff base with the free aldehyde. The results of these two processes are summarized in Eq. 8-9, step *b*. The final product **histidinohydroxymerodesmosine** links four different side chain groups. In other instances, simple Schiff bases (**aldimines**) are formed between aldehyde and ε-amino groups. If there is an adjacent hydroxyl group these can isomerize to ketoamines (Eq. 8-10). Two residues of hydroxylysine, one of which is glycosylated as shown in Eq. 8-10, are often involved. Borohydride reduction (Eq. 4-2) and hydrolysis leads to isolation of **dihydroxylysinonorleucine**, the predominant product of such treatment of bone or cartilage.

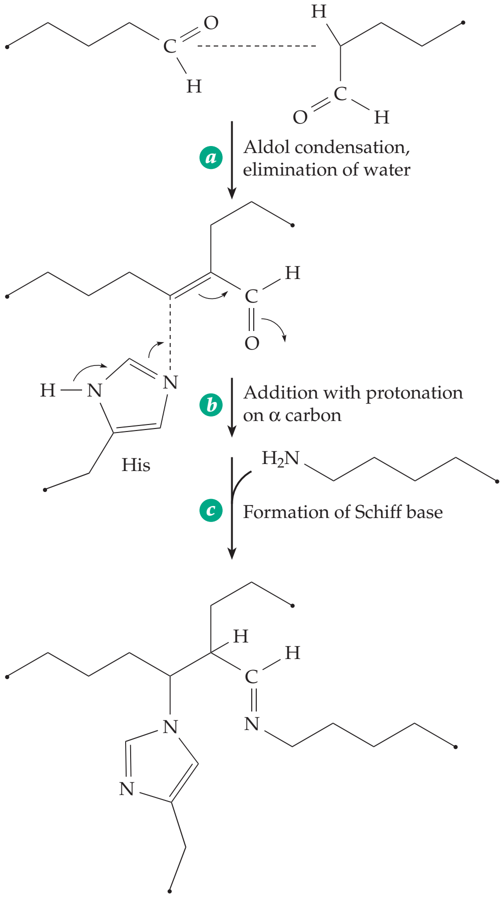


(8-8)

Crosslinkages reducible with borohydride are characteristic of newly formed collagen but these disappear and are replaced by more stable crosslinks as collagen matures. For example, a **3-hydroxypyridine** that joins three triple helices may be formed from the reaction of two ketoamine groupings (with elimination of one glycosylated hydroxylysine residue).[636–639] Similar chemistry can also produce pyrrole crosslinks.[639]



Hydroxypyridine (pyridinoline) crosslink


O
C
H
H
C
O
H
a
Aldol condensation, elimination of water
H
C
O
H—N
N
His
b
Addition with protonation on α carbon
H_2N
c
Formation of Schiff base
H
H
C
N
N
N

Histidinohydroxymerodesmosine

(8-9)

The crosslinkages in collagen are not located at random but are found in certain positions, often toward the ends of the collagen monomers. Thus, histidine residues are found only at positions 89, 929, and 1034 in the α1(I) peptide chains. Residue 9 in the N-terminal globular portion of one chain is linked to residue 946 of another while residue 103 is linked to 1047 in the globular C-terminal peptide.[640] The variety and number of crosslinkages vary among different species. As collagen ages through a lifetime, glycation (Eq. 4-8) leads to crosslinkages in which two ketoamines are formed via glycation cyclize.[641]

Collagens I, II, and III form fibrils with similar structures. However, other collagens are longer or shorter and aggregate in different ways. A pepsin-resistant part of the collagen V molecules resembles collagens I, II, and III but it may contain an additional non-collagenous segment at the N terminus. Collagens V and XI are quantitatively minor components of the extracellular matrix but are thought to provide a core around which the fibrils of collagens I and III may form.[642] Types XII, XIV, IX, and XVI collagens contain interruptions in the helix which create bends, flexible sites, and sites of increased proteolytic susceptibility. They may link the fibrils to other components of the surrounding matrix.[643]

TABLE 8-4
Types of Vertebrate Collagen

Type number	Location	Characteristics	Gene location: human chromosome
Forming quarter-staggered fibrils			
I[a,b]	Skin, bone, tendon, dentin	Most abundant, banded quarter-staggered fibrils	7, 17
III[a,c]	Skin, blood vessels	Abundant, small banded fibrils	2
V[a,d]	Most interstitial tissues; cartilage, bone	Abundant, small fibrils	2
Predominant in cartilage and bone			
II[a]	Hyaline cartilage, vitreous humor	Very abundant, small banded fibrils	12
XI[a]	Hyaline cartilage		
With interrupted triple helices			
XII[a,e]	Embryonic tendon, periodental ligaments	Fibril associated	
XIV[a,e]	Fetal skin, tendon	Fibril associated	
IX[a,e]	Cartilage, vitreous humor	Minor, contains attached glycosaminoglycan, fibril associated	
XVI[f]	Cartilage	Fibril associated	
Forming sheets and networks			
IV[a,g]	All basement membranes	Nonfibrillar network	13
X[a]	Mineralizing cartilage, growth plate	Short chain	
VIII[a,h]	Endothelial cells; Descemet's membrane of the cornea	Small helices linked in hexagonal arrays	
Forming beaded filaments			
VI[a]	Most interstitial tissues, intervertebral discs	Beaded microfilaments	
Forming anchoring fibrils			
VII[a,i]	Basement membranes	Long-chain, antiparallel dimers, anchoring fibrils	

---

[a] Martin, G. R., Timpl, R., Müller, P. K., and Kühn, K. (1985) *Trends Biochem. Sci.* **10**, 285-287
van der Rest, M., and Garrone, R. (1991) *FASEB J.* **5**, 2814-2823
Prockop, D. J., and Kivirikko, K. I. (1995) *Ann. Rev. Biochem.* **64**, 403-434

[b] Prockop, D. J. (1990) *J. Biol. Chem.* **265**, 15349-15352

[c] Nah, H.-D., Niu, Z., and Adams, S. L. (1994) *J. Biol. Chem.* **269**, 16443-16448

[d] Myers, J. C., Loidl, H. R., Stolle, C. A., and Seyer, J. M. (1985) *J. Biol. Chem.* **260**, 5533-5541

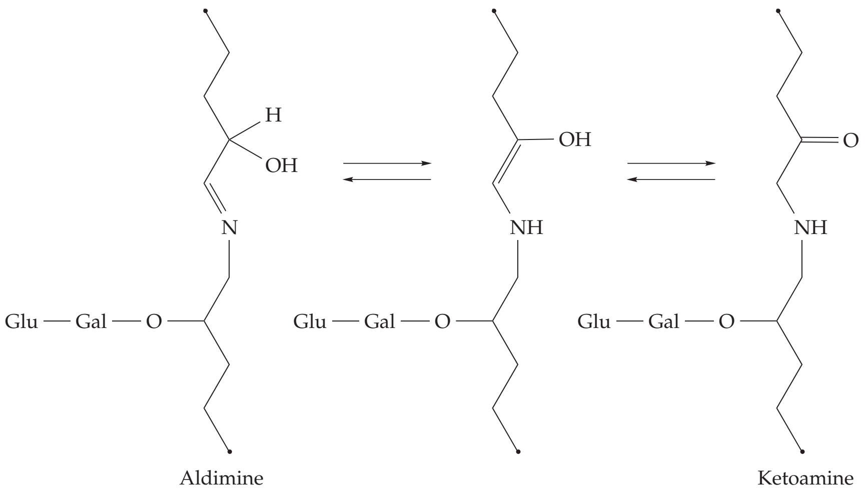
[e] Shaw, L. M., and Olsen, B. R. (1991) *Trends Biochem. Sci.* **16**, 191-194

[f] Myers, J. C., Yang, H., D'Ippolito, J. A., Presente, A., Miller, M. K., and Dion, A. S. (1994) *J. Biol. Chem.* **269**, 18549-18557

[g] Hudson, B. G., Reeders, S. T., and Tryggvason, K. (1993) *J. Biol. Chem.* **268**, 26033-26036

[h] Benya, P. D., and Radilla, S. R. (1986) *J. Biol. Chem.* **261**, 4160-4169

[i] Lunstrum, G. P., Sakai, L. Y., Keene, D. R., Morris, N. P., and Burgeson, R. E. (1986) *J. Biol. Chem.* **261**, 9042-9048



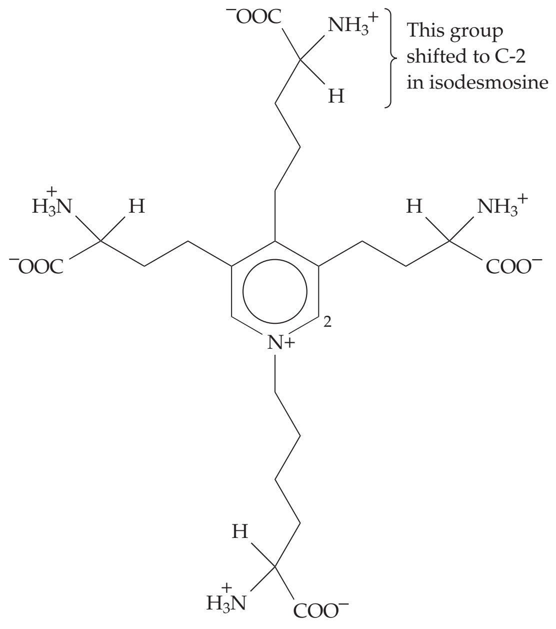
(8-10)

The basement membrane collagen IV forms a nonfibrillar network.[644,645] The 400-nm-long molecules aggregate via their identical ends. Four molecules are held together through their triple-helical N termini, while the C-terminal globular domains connect pairs of molecules. Type IV collagen is also a proteoglycan with a glycoaminoglycan chain attached to one of its nonhelical domains. It may become covalently attached to type II collagen via a hydroxypyridine linkage.[646] Dimers of the microfibrillar collagen VI are formed from 105-nm-long monomers by antiparallel and staggered alignment, with the 75-nm overlapping helical segments twisting around each other to form coiled dimers. After a symmetrical association of dimers to tetramers, fibrillar structures are formed by end-to-end aggregation. The connection of monomers to dimers, tetramers, and polymers occurs by disulfide bonds between triple-helical segments and globular domains. The 450-nm-long collagen VII molecules associate into antiparallel dimeric structures which show a 60-nm overlap and which subsequently assemble laterally with their ends in register.[622]

At least 32 genes encode the α peptide chains of vertebrate collagens.[647,648] These chains are assembled into the 19 known types of collagen. Alternative splicing of the mRNAs provides additional isoforms.[649] The collagen α2(I) gene from both the chick and human DNA is ~38 kb in length and consists of 52 exons separated by introns ranging in length from 80 to 2000 base pairs.[647,650] At least nine of the exons that encode the triple-helical regions have exactly 54 bp. All are multiples of 9 bp, i.e., the length needed to encode one Gly-X-Y triplet (Chapter 2, Section D,4). The significance of these observations is unclear but the presence of so many introns does suggest ways in which collagen sequences could have been transferred into the genes for such proteins as acetylcholineesterase and the C1q component of complement.[647] The human α1(I) collagen gene also consists of 51 segments but they lie within a shorter 18-kb length of DNA.[651] A collagen gene from *Drosophila* is much less fragmented.

Collagens are found in all metazoan organisms.[652] Invertebrate collagens play a variety of specialized roles.[653] For example, minicollagens from *Hydra* strengthen the walls of their nematocysts.[572]

Elastic fibers. The elastic properties of lung, skin, and large blood vessels are provided by elastic fibers in the extracellular matrix.[654] The fibers consist of amorphous material together with the insoluble protein **elastin**, which is rich in glycine, proline, and hydrophobic amino acids. Its special structure (Chapter 2) provides elasticity to the fibers. A 72-kDa precursor **tropoelastin** is secreted into the extracellular space where it is acted upon by lysyl oxidase (Eq. 8-7) and crosslinked into a rubber-like network.[654] Remarkable crosslinkages are formed. Three aldehyde groups derived by oxidation of lysine side chains combine with one lysine amino group through aldol condensations, dehydration, and oxidation to form residues of **desmosine** and **isodesmosine**.



Desmosine

Cartilage and basement membranes. Tendons consist largely of collagen, but in most tissues the collagen fibrils are embedded in a matrix of proteoglycans and various other proteins.[655–657] Both the core proteins of the proteoglycans and the attached

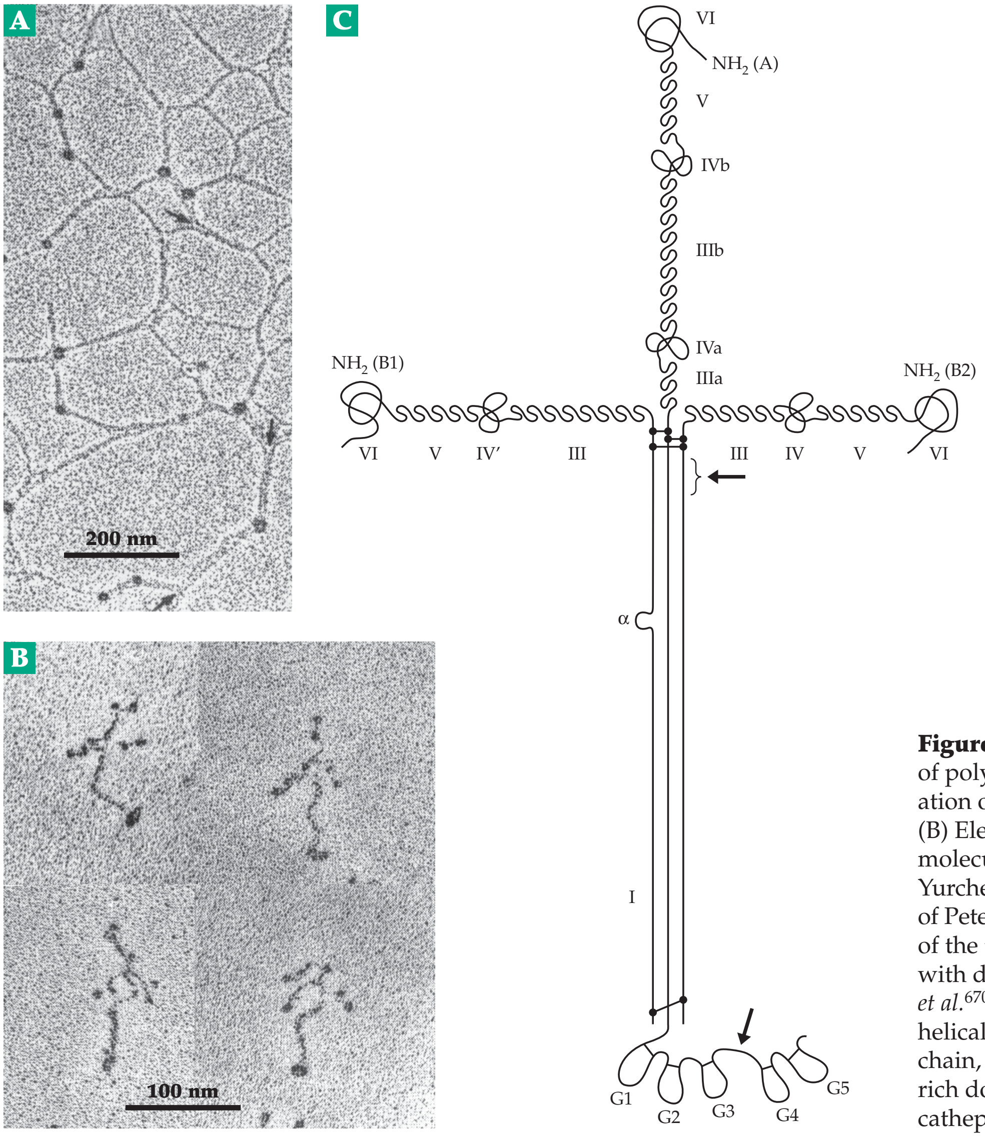


Figure 8-33 (A) Electron micrograph of polygonal network formed by association of collagen type IV monomers. (B) Electron micrographs of single molecules of laminin. (A) and (B) from Yurchenko and Schittny.[663,663a] Courtesy of Peter Yurchenko (C) Structural model of the three-chain laminin molecule with domain designations. From Beck *et al.*[670] Domains I and II are a triple-helical coiled coil rod which, in the B1 chain, is interrupted by a small cysteine-rich domain α. Sites of cleavage by cathepsin are marked by arrows.

polysaccharides interact with collagen fibrils and with fibronectin and other (previously discussed) cell surface proteins. The cartilage matrix[658,659] consists largely of proteoglycans (Fig. 4-16)[660] and of several difficult to study, insoluble proteins. One of these is the 148-kDa **cartilage matrix protein**, which yields 52-kDa subunits upon reduction.[661,662] It interacts with both proteoglycans and collagen and may help to integrate the cartilage matrix.

Basement membranes (Fig. 1-6)[663] function in part as an exoskeleton that helps keep cells positioned. However, the thick basement membranes of the capillary walls of the glomeruli of the kidney provide the ultrafilters that prevent most proteins from entering the urine. Basement membranes contain large amounts of collagen IV, which forms a polygonal network (Fig. 8-33A). A second macromolecular network is formed by the very large 950-kDa cross-shaped multisubunit protein called **laminin** (Fig. 8-33B).[664–666] Laminin is one of a series of extracellular proteins which appear to have arisen by shuffling of structural modules during evolution.[667] It contains sites for binding to heparin, to integrins,[668] to the heparin sulfate proteoglycan **agrin** (see also Fig. 4-11), and to the 150-kDa sulfated glycoprotein **nidogen** (entactin).[669] At least seven isoforms of laminin, with varying tissue distributions, are formed, in part as a result of alternative splicing of the mRNA transcripts.[666] Laminin is rich in EGF-like modules.[665] An X-ray structure of three of them shows that they form a continuous rod of complex structure.[321] A smaller 100 kDa basement membrane **fibulin** also contains multiple EGF-like repeats.[671] As with other extracellular structures, crosslinking of laminin and other components of basement membranes by transglutaminase provides additional stability.[672]

BOX 8-E GENETIC DEFECTS OF COLLAGEN STRUCTURE

Any major protein of the body is likely to be associated with a number of genetic problems. In the case of collagen, the possibility for harmful mutations is enhanced by the existence of a large number of genes that encode the more than 16 types of collagen which are expressed differently in different tissues. There are known human disorders resulting from defects in synthesis, secretion, or structure of types I, II, III, IV, and VII collagens. Other defects involve lysyl hydroxylase and procollagen *N*-proteinase.[a–d]

In the severe lethal form of **osteogenesis imperfecta** (brittle bone syndrome) the victims' collagen I may contain an α1 chain lacking as many as 100 residues or a shortened α2 chain. In other cases a cysteine, arginine, or other amino acid has been substituted for glycine in the triple-helical region of an α1 chain.[e,f] The cloning of collagen genes has permitted an exact description and precise location of the defects in these genes. Alterations toward the N-terminus of the α1 chain or in the α2 chain often cause a milder type of osteogenesis imperfecta.[g] Some patients have deletions in the pro-α2 chains of collagen I causing the chain to be incorporated into the collagen without removal of the N-terminal or C-terminal peptide to give a protein with poor stability.[h] In other cases amino acid substitutions in the α2 chain cause formation of chains with excessive posttranslational modification. Sometimes the α2 chain is not incorporated into the triple helix and the collagen I formed contains three α1 chains.

Another well established abnormality of collagen is found in cattle suffering from **dermatosparaxis**, a disease in which the skin is extremely brittle. The collagen chains are disorganized and have poor fiber-forming properties. The procollagen peptidase that cleaves a peptide from the N termini of the chains of procollagen is apparently defective. A similar human collagen disease is the **Ehlers–Danlos syndrome**, which in some instances is accompanied with recurrent joint dislocations and curvature of the spine. At least ten different types of the disorder are known.[d] The procollagen peptidase is sometimes lacking.[i] In other cases a person synthesizes an abnormal pro-α2 chain that is resistant to cleavage by the peptidase because of deletion of the normal cleavage site. In others collagen is formed in only small amounts or is degraded rapidly. Some individuals lack lysyl hydroxylase and others have a defect in mRNA splicing which causes loss of an exon from the mRNA and synthesis of shortened pro-α2 chains.[j]

Somewhat similar symptoms are observed in **lathyrism**, a disease which arises when animals ingest seeds of *Lathyris odoratus*, the common sweetpea. Since lathyris peas form part of the diet of some peoples, the condition is also known in humans and often causes curvature of the spine and rupture of the aorta. The biochemical problem has been traced to the presence in the seeds of **β-cyanoalanine** and of its decarboxylation product **β-aminopropionitrile**.

$$\mathrm{N{\equiv}C - CH_2CH_2 - NH_3^+}$$

Although the mode of action is not certain, this compound is an inhibitor of lysyl oxidase essential to the crosslinking of both collagen and elastin. A hereditary defect with a similar effect in the mouse involves a defect in lysyl oxidase.[k,l]

Collagen defects account for a variety of other skeletal problems[m] including some cases of the common **osteoarthritis**.[n] Mice lacking the α1 chain of collagen IX develop a degenerative joint disease resembling human osteoarthritis.[o] An inherited defect in the basement membrane collagen IV is responsible for the inherited **Alport disease** in which kidney filtration is defective.[p,q] Similar symptoms are observed with the acute autoimmune disease **Goodpasture syndrome** and in **diabetic nephropathy**.[p]

---

[a] Prockop, D. J. (1990) *J. Biol. Chem.* **265**, 15349–15352

[b] Kuivaniemi, H., Tromp, G., and Prockop, D. J. (1991) *FASEB J.* **5**, 2052–2060

[c] Prockop, D. J., and Kivirikko, K. I. (1995) *Ann. Rev. Biochem.* **64**, 403–434

[d] Byers, P. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4029–4077, McGraw-Hill, New York

[e] Cohen-Solal, L., Zylberberg, L., Sangalli, A., Gomez Lira, M., and Mottes, M. (1994) *J. Biol. Chem.* **269**, 14751–14758

[f] Lightfoot, S. J., Atkinson, M. S., Murphy, G., Byers, P. H., and Kadler, K. E. (1994) *J. Biol. Chem.* **269**, 30352–30357

[g] Marini, J. C., Lewis, M. B., Wang, Q., Chen, K. J., and Orrison, B. M. (1993) *J. Biol. Chem.* **268**, 2667–2673

[h] Mundlos, S., Chan, D., Weng, Y. M., Sillence, D. O., Cole, W. G., and Bateman, J. F. (1996) *J. Biol. Chem.* **271**, 21068–21074

[i] Holmes, D. F., Watson, R. B., Steinmann, B., and Kadler, K. E. (1993) *J. Biol. Chem.* **268**, 15758–15765

[j] Weil, D., D'Alessio, M., Ramirez, F., Steinmann, B., Wirtz, M. K., Glanville, R. W., and Hollister, D. W. (1989) *J. Biol. Chem.* **264**, 16804–16809

[k] Pope, F. M., Martin, G. R., Lichtenstein, J. R., Penttinen, R., Gerson, B., Rowe, D. W., and McKusick, V. A. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1314–1316

[l] Rowe, D. W., McGoodwin, E. B., Martin, G. R., and Grahn, D. (1977) *J. Biol. Chem.* **252**, 939–942

[m] Freisinger, P., Ala-Kokko, L., LeGuellec, D., Franc, S., Bouvier, R., Ritvaniemi, P., Prockop, D. J., and Bonaventure, J. (1994) *J. Biol. Chem.* **269**, 13663–13669

[n] Ala-Kokko, L., Baldwin, C. T., Moskowitz, R. W., and Prockop, D. J. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6565–6568

[o] Fässler, R., Schnegelsberg, P. N. J., Dausman, J., Shinya, T., Muragaki, Y., McCarthy, M. T., Olsen, B. R., and Jaenisch, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5070–5074

[p] Zhou, J., Hertz, J. M., Leinonen, A., and Tryggvason, K. (1992) *J. Biol. Chem.* **267**, 12475–12481

[q] Gunwar, S., Ballester, F., Noelken, M. E., Sado, Y., Ninomiya, Y., and Hudson, B. G. (1998) *J. Biol. Chem.* **273**, 8767–8775

BOX 8-F SKIN

Mammalian skin must be tough, water-resistant, self-renewing, and rapidly healing. The outer layers of cells or **epidermis** consist principally of **keratinocytes**, epithelial cells specialized for formation of keratin (Fig. 7-31). In the inner layer of the epidermis the **basal stem cells** divide, providing a constant outward flow of cells which become progressively flattened, dehydrated, and filled with keratin fibrils.[a] The outer layers contain only dead cells which are finally sloughed or abraded from the surface. Human epidermis is completely renewed in about 28 days!

About 25 different human genes encode the keratins of skin and other soft tissues. Others specify the keratins of hair and nails.[b] Both of these hard tissues as well as claws, hoofs, beaks, horns, scales, quills, and feathers are largely keratin. However, there are additional constituents. During the final stages of keratinocyte differentiation a 15-nm-thick crosslinked sheath of protein, the **cornified cell envelope** (CE), forms beneath the plasma membrane.[c] Crosslinkages between keratin and other proteins are formed by the action of transglutaminases.[d–f] A specialized protein **involucrin**, which contains glutamine-rich repeating sequences, provides many of the side chain amide groups for the crosslinking reaction (Eq. 2-23).[g] **Loricin**, a protein containing glycine-rich flexible loops,[h] is also a major partner in these crosslinkages.[c,h] The histidine-rich **filaggrin** undergoes complex processing before binding to keratin fibrils to provide another form of crosslinkage.[i,j] Small proline-rich proteins, desmosonal proteins, and others are also present in the CE.[c]

As the final outer **stratum corneum** is formed the phospholipid bilayer deteriorates and intercellular lipid layers are formed.[k,l] These contain principally ceramides, cholesterol, and free fatty acids. Some sphingolipids are covalently attached to proteins.[a]

Also present in the epidermis are embedded macrophage-like Langerhans cells as well as pigmented **melanocytes**, cells with highly branched dendrites, which lie just above the basal stem cell layer. Each melanocyte contains hundreds of pigmented organelles called **melanosomes**. They contain not only the black or reddish **melanin** pigments but also the enzymes needed to form them (Chapter 25).[n,o]

The dendrites of a melanocyte contact about 36 keratinocytes and are able to transfer melanosomes to these adjacent cells. The numbers and sizes of the melanosomes as well as melanin structure determine differences in skin color.[m] Similar cells in amphibians, the **melanophores**, also contain light receptors.[p] Their melanosomes are not transferred to other cells but may be either clustered near the center of the cell or dispersed. The location can be changed quickly by transport of the melanosomes along a network of microtubules allowing the animals to change in response to changes in light color.[q] Various stimuli, including ultraviolet irradiation of melanocytes, cause increased synthesis of melanin with a resultant tanning[o] and added protection against sunburn.

Beneath the basement membrane of the epidermis is the **dermis**, a thick, tough, collagen-rich connective tissue. Blood vessels and nerve endings are found in this layer, as are roots of hairs and oil and sweat glands.[r]

Skin suffers from a variety of ailments including serious hereditary diseases.[a] One group of **keratinization disorders**, known as **ichthyoses**, are characterized by thickened, scaly skin. Some hereditary ichthyoses result from defects in type II

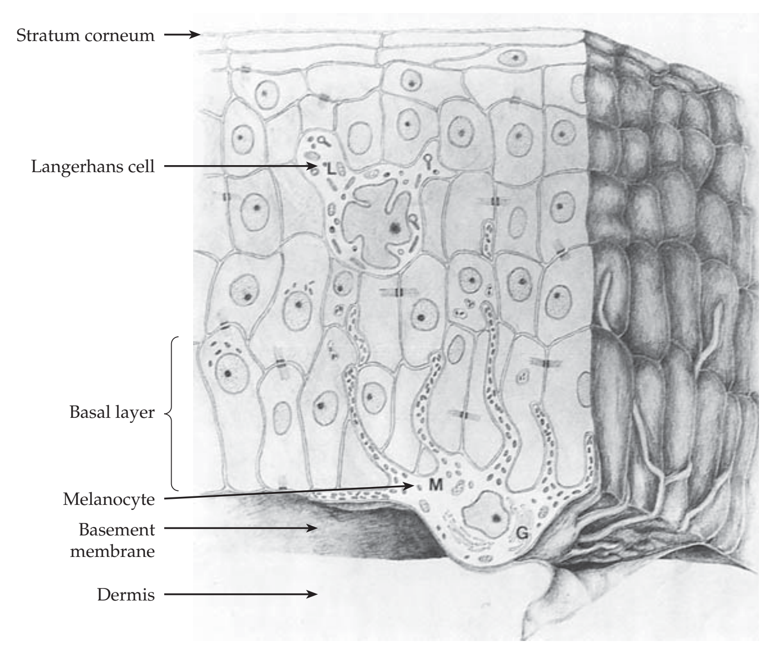


Diagram of a dendritic melanocyte surrounded by satellite keratinocytes. The Golgi area (G), where the melanosomes are synthesized, is shown around the nucleus. The other branched cell, higher in the epidermis, is a Langerhans cell with its tennis racquet-shaped granules. Courtesy of Dr. W. Quevedo, Jr. From Montagna *et al*.[m]

BOX 8-F SKIN (continued)

keratin.[b] **Lamellar ichthyosis** reflects defects in the crosslinking enzyme transglutaminase.[e,f] **Epidermolysis bullosa** is a heterogeneous group of disorders characterized by easy formation of blisters. One form has been traced to a defect in the anchoring fibrils of type VII collagen, which tie cells of the basal layer to the basement membrane.[s,t] Others are defects in keratins of the basal or intermediate layers.[a] Yet others involve the lipid metabolism of skin, e.g. a steroid sulfatase deficiency.[a]

The most frequent skin disorder, which affects about 2% of the world's population is **psoriasis**. The thickened, scaly patches can cover much of the skin and become disabling. The inflammation and excessive epidermal growth are usually a T-cell mediated immunologic response to antigenic stimuli.[p,u,v] However, there is a hereditary form.[u] Other common skin disorders include actinic keratosis induced by light and cancer.

---

[a] Roop, D. (1995) *Science* **267**, 474–475

[b] Takahashi, K., Paladini, R. D., and Coulombe, P. A. (1995) *J. Biol. Chem.* **270**, 18581–18592

[c] Robinson, N. A., Lapic, S., Welter, J. F., and Eckert, R. L. (1997) *J. Biol. Chem.* **272**, 12035–12046

[d] Kim, S.-Y., Chung, S.-I., and Steinert, P. M. (1995) *J. Biol. Chem.* **270**, 18026–18035

[e] Huber, M., Rettler, I., Bernasconi, K., Frenk, E., Lavrijsen, S. P. M., Ponec, M., Bon, A., Lautenschlager, S., Schorderet, D. F., and Hohl, D. (1995) *Science* **267**, 525–528

[f] Candi, E., Melino, G., Lahm, A., Ceci, R., Rossi, A., Kim, I. G., Ciani, B., and Steinert, P. M. (1998) *J. Biol. Chem.* **273**, 13693–13702

[g] Yaffe, M. B., Beegen, H., and Eckert, R. L. (1992) *J. Biol. Chem.* **267**, 12233–12238

[h] Hohl, D., Mehrel, T., Lichti, U., Turner, M. L., Roop, D. R., and Steinert, P. M. (1991) *J. Biol. Chem.* **266**, 6626–6636

[i] Resing, K. A., Walsh, K. A., Haugen-Scofield, J., and Dale, B. A. (1989) *J. Biol. Chem.* **264**, 1837–1845

[j] Mack, J. W., Steven, A. C., and Steinert, P. M. (1993) *J. Mol. Biol.* **232**, 50–66

[k] ten Grotenhuis, E., Demel, R. A., Ponec, M., Boer, D. R., van Miltenburg, J. C., and Bouwstra, J. A. (1996) *Biophys. J.* **71**, 1389–1399

[l] Bouwstra, J. A., Thewalt, J., Gooris, G. S., and Kitson, N. (1997) *Biochemistry* **36**, 7717–7725

[m] Montagna, W., Prota, G., and Kenney, J. A., Jr. (1993) *Black Skin Structure and Function*, Academic Press, San Diego, California

[n] Potterf, S. B., Muller, J., Bernardini, I., Tietze, F., Kobayashi, T., Hearing, V. J., and Gahl, W. A. (1996) *J. Biol. Chem.* **271**, 4002–4008

[o] Roméro-Graillet, C., Aberdam, E., Biagoli, N., Massabni, W., Ortonne, J.-P., and Ballotti, R. (1996) *J. Biol. Chem.* **271**, 28052–28056

[p] Greaves, M. W., and Weinstein, G. D. (1995) *N. Engl. J. Med.* **332**, 581–588

[q] Rogers, S. L., Tint, I. S., Fanapour, P. C., and Gelfand, V. I. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3720–3725

[r] Martin, P. (1997) *Science* **276**, 75–81

[s] Byers, P. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4029–4077, McGraw-Hill, New York

[t] Christiano, A. M., Ryynänen, M., and Uitto, J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3549–3553

[u] Tomfohrde, J., Silverman, A., Barnes, R., Fernandez-Vina, M. A., Young, M., Lory, D., Morris, L., Wuepper, K. D., Stastny, P., Menter, A., and Bowcock, A. (1994) *Science* **264**, 1141–1145

[v] Boehncke, W.-H., Dressel, D., Zollner, T. M., and Kaufmann, R. (1996) *Nature (London)* **379**, 777

Fibrillin and Marfan's syndrome. Most connective tissues contain insoluble, beaded microfibrils 10–12 nm in diameter. A component of some of these microfibrils, which are often found in elastic tissue, was purified from media used to culture human fibroblasts in 1986. This protein, called fibrillin, is a single-chain 350-kDa glycoprotein which contains ~14% cysteine.[673–674] Using a DNA probe based on the partially cloned fibrillin gene, the location of the gene was established on the long arm of chromosome 15 at a site previously identified as that of a gene defective in Marfan's syndrome. This disorder often causes dislocation of lenses of the eyes and aortic aneurysm as well as elongated limbs and fingers. Point mutations in the fibrillin gene have been identified in both Marfan's patients and family members that carry the defective gene.[675,676]

The cuticles of invertebrates. The tough elastic cuticle of the nematode *Caenorhabditis* is largely collagen. However, the molecules are smaller than in vertebrates and there are ~100 different genes whose transcription gives rise to a large variety of similar proteins.[652] Cuticles of some annelids have unusually long collagens.[653,677] In contrast, the epithelial cells of insects and other arthropods secrete chitin which serves as the framework for development of a thick and often hard cuticle or exoskeleton. The cuticle also contains a variety of proteins.[678,679] In some instances mineralization by calcium carbonate occurs. During the later phases of the cuticle development extensive crosslinking of the proteins takes place. This is largely by reactions between modified aromatic side chains and resembles the chemistry of formation of melanin and lignin (Chapter 25).[680]

Bones, teeth, and shells. Living organisms are able to induce the formation of over 60 inorganic compounds.[681] Most of these are formed by animals. Two forms of calcium carbonate, **calcite** and **aragonite**, predominate.[682] These minerals form shells, exoskeleton bones, bones, teeth, and other specialized structures.

While some organisms promote mineral deposition completely outside of their cell coats, the mineralization is usually controlled by the proteins and polysaccharides lying around and between cells.[683–687]

Bone deposition begins in the proteoglycan and collagen II matrix of cartilage. Later these polymers are largely replaced by collagen I which accounts for 90% of the organic material in mature bone. Bone collagen has a distinctive pattern of crosslinking.[639,688,689] Some borohydride-reducible crosslinks remain throughout adult life.[639] However, complex hydroxypyridine and pyrrole linkages are more characteristic of bone collagen. Embedded in the bone matrix are spidery cells, the **osteocytes**. Among these are **osteoblasts**,[689a] which secrete the collagen and other proteins that promote the laying down of calcium phosphate. Also present are large multinucleate **osteoclasts**[689b] which dissolve bone and reabsorb calcium and phosphate. Cells of both types remain active in mature bone, which is both a structural material and a store of calcium and phosphorus.

How do osteoblasts induce calcium phosphate deposition? The first crystals laid down within the cartilage matrix are of carbonate apatite only 2–3 nm thick and tens of nanometers in length. They are intimately associated with the collagen and other components of the matrix[690] and may be formed within matrix vesicles.[691] From observations with large biologically formed calcium carbonate crystals it is known that proteins or other organic initiators of crystallization can be found embedded in mature crystals. If the same is true in bone what are the initiators? The answer is uncertain but it is known that without osteoblasts the partially mineralized cartilage will not become bone.[692]

A key to the development of osteoblasts appears to be an **osteoblast-specific transcription factor OS/2** or **Cbfa1**.[693–695] Mutations in human Cbfa1 are linked to a series of skeletal defects.[696] A unique change accompanying conversion of a precursor cell into an osteoblast is the formation of a 49-residue γ-carboxyglutamate (Gla)-containing protein called **osteocalcin**.[697] (See also Box 15-F). It is the most abundant noncollagenous protein of bone. Its three Gla residues doubtless help to bind calcium ions and osteocalcin may be an initiator of crystallization. Osteocalcin has also been found in fish scales.[698] Also present in bone is a 74-residue **matrix Gla protein** which has 5 Gla residues.[699]

Other phosphoproteins, glycoproteins, and proteoglycans may also be required for mineralization. For example, the 286-residue glycoprotein **osteonectin**[323,700,701] accounts for 3% of total bone protein. It contains two Ca^{2+}-binding motifs and inhibits growth of hydroxylapatite crystals, but its role in bone development is not clear. One of the phosphoproteins of developing bone has been identified as the 24 kDa propeptide cut from the N-terminal end of the α1 chain of type I procollagen.[702] Another bone protein, the sialic acid-containing **osteopontin** has the cell-binding sequence GRGDS identical to that in fibronectin.[703,704] Since it also binds to hydroxylapatite this protein may form a bridge between cell surfaces and the mineral matrix.

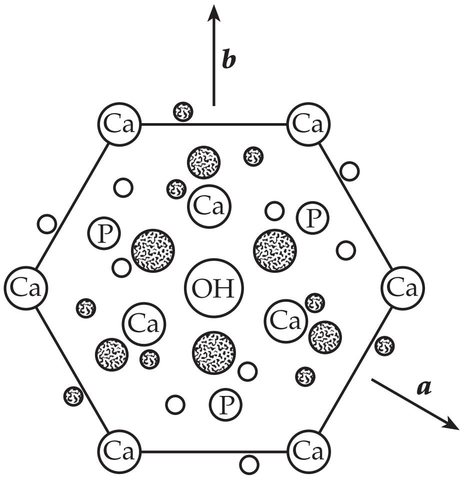


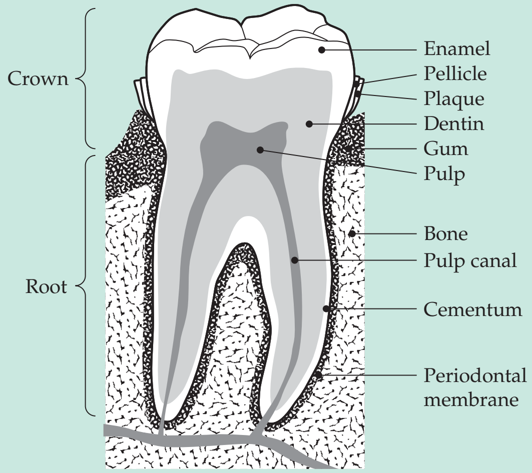
Figure 8-34 An end-on view of a crystallite of hydroxylapatite. The shaded atoms of Ca, P, and O represent an underlying layer. The OH^- groups form a longitudinal H-bonded array in the center. From J. A. Weatherell and C. Robinson,[705] p. 55. A small amount of Mg^{2+} is present in place of Ca^{2+} and a very small fraction of the OH^- is replaced by fluoride ion which has a bone strengthening effect.

The mineral phase of bone is largely hydroxylapatite, $Ca_{10}(PO_4)_6(OH)_2$ (Fig. 8-34), which is essentially in chemical equilibrium with the calcium and phosphate ions present in the blood serum. Thus, bone cells can easily promote either the deposition or dissolution of the mineral phase by localized change in pH, concentrations of Ca^{2+} or HPO_4^{2-}, or of chelating compounds such as ATP or inorganic pyrophosphate. Small 100 nm vesicles rich in acidic phospholipids and containing both Ca^{2+} and the enzymes alkaline phosphatase and pyrophosphatase, may play an essential role in calcification. Perhaps they release the calcium and enzymes that generate inorganic phosphate.[691]

Bone is noted for its continual "remodeling" caused by the action of both osteoblasts and osteoclasts.[704a,b] The latter are large multinuclear cells derived from the same precursors as give rise to macrophages.[689b,704] Osteoclasts generate a sealed, acidic compartment on a bone surface, using a vacuolar ATPase to pump protons from the cytosol of the osteoclast, using H_2CO_3 as a source of protons.[706] The HCO_3^- then leaves cells in exchange for Cl^- which then also enters the acidified compartment. Not only is the calcium phosphate dissolved but lysomal enzymes digest the organic materials. Each osteoclast creates a long tun-

BOX 8-G THE BIOCHEMISTRY OF TEETH

Like other bones, mammalian teeth are composed largely of collagen and hydroxylapatite. The much studied rat incisor is 65% mineral; about 85% of the organic material is type I collagen. A tooth consists of three mineralized tissues together with the internal blood vessels and nerves of the pulp.[a] The outer layer of **enamel** is formed by secretion from a thin layer of epithelial cells, the **ameloblasts**. The enamel matrix is devoid of collagen but contains two characteristic groups of proteins, the **amelogenins** and the **enamelins**.



This cross section of a human molar shows deposits of pellicle and plaque near the gum line[h]

The amelogenins are hydrophobic but are also rich in proline, histidine, and glutamic acid.[a,b] They account for 90% of the matrix proteins but are replaced by the initially less abundant enamelins[c] in fully mineralized teeth.[a,d] Another protein present in developing enamel is **ameloblastin**, which appears to be unique to amyloblasts.[a] As mineralization of the enamel progresses the matrix is lost and 98% of the enamel is hydroxylapatite.[a]

Dentin is of epithelial-mesenchymal origin. The **odontoblasts** secrete an extracellular matrix that is rich in collagen I and which also contains all of the other major bone proteins as well as a **dentin sialoprotein**[e] and poorly characterized phosphate carriers, the **phosphophoryns**.

The third hard tissue of teeth is **cementum**, which binds teeth to the periodental ligaments. The ligaments contain fibrillar collagen which inserts into the cementum.[f,g]

Except for the common cold, tooth decay (**caries**) is the most prevalent disease in the United States.[h,i] Caries is initiated by attack of acids produced by bacterial fermentations on the enamel. The saliva contains calcium and phosphate and is supersaturated with respect to these ions. As a result the enamel surface is continuously recalcified. The 43-residue salivary protein **statherin** retards precipitation of calcium phosphates from the saliva, preventing excessive build-up of calcified deposits on the teeth.[j] Other proline-rich proteins may play a role in recalcification repair.[k]

A freshly cleaned tooth surface quickly becomes coated with a thin **pellicle** of salivary proteins. This provides a surface for growth of **dental plaque**, which contains many bacteria and adhesive polysaccharides such as dextrans.[l] The latter are generated from dietary sucrose by such bacteria as *Streptococcus mutans*. (Chapter 20) and others.[m] Many factors affect the probability of tooth decay. A high sucrose diet promotes decay.[n] While most people have some trouble with tooth decay, 1 or 2 per thousand remain totally free of caries and seem to be immune. Many factors must affect resistance to caries. For example, individuals vary in the kinds and numbers of bacteria present on teeth and in the structure of tooth enamel.[o] Addition of fluoride ion to water supplies at a level of 1ppm (0.05 mM) is generally believed to reduce the incidence of tooth decay. However, caries has been declining in many developed countries at rates that are the same for water with or without fluoride.[p,q] If teeth escape caries **periodental disease**, caused by bacteria, is often a major problem for older people.[r]

---

[a] Krebsbach, P. H., Lee, S. K., Matsuki, Y., Kozak, C. A., Yamada, K. M., and Yamada, Y. (1996) *J. Biol. Chem.* **271**, 4431–4435

[b] Renugopalakrishnan, V., Strawich, E. S., Horowitz, P. M., and Glimcher, M. J. (1986) *Biochemistry* **25**, 4879–4887

[c] Deutsch, D., Palmon, A., Fisher, L. W., Kolodny, N., Termine, J. D., and Young, M. F. (1991) *J. Biol. Chem.* **266**, 16021–16028

[d] Robinson, C., Weatherell, J. A., and Höhling, H. J. (1983) *Trends Biochem. Sci.* **8**, 284–287

[e] Ritchie, H. H., Hou, H., Veis, A., and Butler, W. T. (1994) *J. Biol. Chem.* **269**, 3698–3702

[f] Arzate, H., Olson, S. W., Page, R. C., Gown, A. M., and Narayanan, A. S. (1992) *FASEB J.* **6**, 2990–2995

[g] Yamauchi, M., Katz, E. P., and Mechanic, G. L. (1986) *Biochemistry* **25**, 4907–4913

[h] Sanders, H. J. (1980) *Chem. Eng. News* **(Feb. 25)** **58(8)** 30–39

[i] Shaw, J. H. (1987) *N. Engl. J. Med.* **317**, 996–1004

[j] Schlesinger, D. H., and Hoy, D. I. (1977) *J. Biol. Chem.* **252**, 1689–1695

[k] Clements, S., Mehansho, H., and Carlson, D. M. (1985) *J. Biol. Chem.* **260**, 13471–13477

[l] Kolenbrander, P. E., Ganeshkumar, N., Cassels, F. J., and Hughes, C. V. (1993) *FASEB J.* **7**, 406–408

[m] Abeygunawardana, C., and Bush, C. A. (1990) *Biochemistry* **29**, 234–248

[n] Newbrun, E. (1982) *Science* **217**, 418–423

[o] Cevc, G., Cevc, P., Schara, M., and Skaleric, U. (1980) *Nature (London)* **286**, 425–426

[p] Diesendorf, M. (1986) *Nature (London)* **322**, 125–129

[q] Hileman, B. (1988) *Chem. Eng. News* **Aug. 1**, 26–42

[r] Williams, R. C. (1990) *N. Engl. J. Med.* **322**, 373–382

nel into the interior of the bone. However, the tunnel is soon lined with new osteoblasts and new bone is laid down.[707] In **osteoporosis**, a very common disease of older persons, the rate of resorption of bone by osteoclasts exceeds that of bone formation by osteoblasts. Women typically lose 50% and men ~30% of their calcium phosphate from vertebrae and ends of long bones as they age.[708] In osteopetrosis (marble bone disease) excessive calcification of bone and other tissues occurs as a result of a deficiency of carbonic anhydrase.[709] See also Chapter 13. A disease of progressive calcification of soft tissues[710] results from an excess of **bone morphogenic factor-4**, one of a group of protein factors acting on bone development (Chapter 32).[631] **Paget's disease** is another disorder of bone remodeling that leads to overproduction of bone of poor quality.[711] As discussed in Box 22-C, the circulating level of Ca^{2+} as well as the cellular uptake of this ion are controlled by vitamin D and its metabolites and by calcitonin and the parathyroid hormones.

Lack of a pyrophosphate ion pump in cartilage cells may cause a deficit in pyrophosphate in the surroundings of the forming bone. In mice with a defect in this pump bony spurs, similar to those in human osteoarthritis, are formed.[711a] Over half of the world's population of persons over 65 years of age are afflicted by arthritis, over 100 types being known.[711b]

Calcium carbonate in several different crystalline forms arises biologically.[705,712] Sometimes the mineral is deposited within vesicles inside the cell. Thus some species of the protozoan *Hymenomonas* use Golgi vesicles as sites for construction of segments of plantlike cell walls complete with crystalline calcium carbonate plates attached to the outer surface.[713] These wall segments are then transported to their final locations. The intricately sculptured spicules of sea urchins and sponges are single crystals of calcium carbonate which have grown within intracellular vesicles.[687,714] Diatoms also form their shells of nearly pure, hydrated SiO_2 (See Box 4-B) entirely within membrane-bound vesicles.[715] On the other hand, shells of molluscs are usually deposited outside the cells of the organism but again under the influence of a protein matrix.[716,717] The animals apparently actively pump bicarbonate outward where it reacts with Ca^{2+} (Eq. 8-11).[718]

$$Ca^{2+} + HCO_3^- \rightarrow CaCO_3 + H^+ \qquad (8\text{-}11)$$

The protons released then react with more bicarbonate to form carbonic acid which is converted to CO_2 and water (Eq. 8-12) by the enzyme carbonic anhydrase.

$$H_2CO_3 \xrightarrow{\substack{\text{Carbonic}\\ \text{anhydrase}}} CO_2 + H_2O \qquad (8\text{-}12)$$

A problem that arises within organisms is avoiding crystallization under supersaturating conditions. For example, normal urine is supersaturated with calcium oxalate. To prevent formation of renal calculi (stones)[719] an inhibitory glycoprotein is present and slows the formation and growth of crystals.[720] Under some disease conditions calcium carbonate stones may form in pancreatic ducts. A 17 kDa lectinlike glycoprotein called **lithostatine** has been proposed to inhibit stone formation by binding to certain planes on $CaCO_3$ microcrystals just as antifreeze proteins (Box 4-D) inhibit ice formation.[721] However, this proposed function for lithostatine is doubtful.[722,723] Pathological deposits of crystalline calcium pyrophosphate and basic calcium phosphates are sometimes present in joints,[724] even in Neanderthal skeletons.[725]

3. Cell Walls of Fungi and Green Plants

The cell walls of yeasts and other fungi are made up largely of glucans, chitin, and a mannan-protein complex. Yeasts contain predominately glucans but chitin is the major polysaccharide in many other fungi. The most abundant glucan is a β1,3 linked polymer with about 3% β1,6 linkages to the branches and a molecular mass of ~ 240 kDa. In addition there is about 15% of a highly branched 1,6 linked glucan containing 1,3 linked branches. At the outer surface of the wall are mannoproteins[726–728] which carry small serine- or threonine-linked oligomannans as well as large highly branched mannans linked to asparagine through the usual *N*-acetylglucosamine-containing core structure. Some of these highly branched mannan chains serve as species-specific antigens.[727] Like those of the bacterial and animal cell surfaces, the antigens vary in structure, a fact with important medical implications. Curiously, many fungi have surface fimbriae which are composed of **collagen**, which is usually regarded as exclusively an animal protein.[729]

Cell walls of higher plants (Fig. 1-7, 4-14) are composed largely of polysaccharides. They are discussed briefly in Chapter 20.

References

1. Christie, W. W. (1973) *Lipid Analysis; Isolation, Preparation, Identification and Structural Analysis of Lipids*, Pergamon, Oxford
2. Hitchcock, C., and Nichols, B. W. (1971) *Plant Lipid Biochemistry*, Academic Press, New York
3. Kates, M. (1972) in *Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 3, Part II (Work, T. S., and Work, E., eds), North-Holland Publ., Amsterdam
4. Weete, J. W. (1980) *Lipid Biochemistry of Fungi and Other Organisms*, Plenum, New York
5. Gurr, M. I., and James, A. T. (1975) *Lipid Biochemistry, an Introduction*, 2nd ed., Cornell Univ. Press, Ithaca, New York
6. Porter, K. R., and Bonneville, M. A. (1973) *Fine Structure of Cells and Tissues*, 4th ed., Lea & Febiger, Philadelphia, Pennsylvania
7. Aveldaño, M. I., and Sprecher, H. (1987) *J. Biol. Chem.* **262**, 1180 – 1186
8. Gunstone, F. D. (1967) *An Introduction to the Chemistry and Biochemistry of Fatty Acids and Their Glycerides*, 2nd ed., Chapman and Hall, London (Chapter 6)
9. Steck, P. A., Schwartz, B. A., Rosendahl, M. S., and Gray, G. R. (1978) *J. Biol. Chem.* **253**, 5625 – 5629
10. McNeil, M., Daffe, M., and Brennan, P. J. (1991) *J. Biol. Chem.* **266**, 13217 – 13223
11. Qureshi, N., Takayama, K., and Schnoes, H. K. (1980) *J. Biol. Chem.* **255**, 182 – 189

11a. Adlof, R. O., Duval, S., and Emken, E. A. (2000) *Lipids* **35**, 131 – 135

12. Cravatt, B. F., Prospero-Garcia, O., Siuzdak, G., Gilula, N. B., Henriksen, S. J., Boger, D. L., and Lerner, R. A. (1995) *Science* **268**, 1506 – 1509
13. Hummel, H. E., Gaston, L. K., Shorey, H. H., Kaae, R. S., Byrne, K. J., and Silverstein, R. M. (1973) *Science* **181**, 873 – 874
14. Klun, J. A., Chapman, O. L., Mattes, K. C., Wojtkowski, P. W., Beroza, M., and Sonnet, P. E. (1973) *Science* **181**, 661 – 663
15. Moolenaar, W. H. (1995) *J. Biol. Chem.* **270**, 12949 – 12952
16. Wells, W. W., and Eisenberg, F., Jr., eds. (1978) *Cyclitols and Phosphoinositides*, Academic Press, New York
17. DeBuch, H. (1978) *Trends Biochem. Sci.* **3**, 44 – 45
18. Chilton, F. H., and Connell, T. R. (1988) *J. Biol. Chem.* **263**, 5260 – 5265
19. Ford, D. A., and Gross, R. W. (1994) *Biochemistry* **33**, 1216 – 1222
20. Friedman, S. M., ed. (1978) *Biochemistry of Thermophily*, Academic Press, New York
21. Kushwaha, S. C., Kates, M., Sprott, G. D., and Smith, I. C. P. (1981) *Science* **211**, 1163 – 1164
22. Luzzati, V., Gambacorta, A., DeRosa, M., and Gulik, A. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 25 – 47
23. Goldfine, H., and Langworthy, T. A. (1988) *Trends Biochem. Sci.* **13**, 217 – 221
24. Ferrante, G., Ekiel, I., and Sprott, G. D. (1986) *J. Biol. Chem.* **261**, 17062 – 17066
25. Pond, J. L., Langworthy, T. A., and Holzer, G. (1986) *Science* **231**, 1134 – 1136
26. Florin-Christensen, J., Florin-Christensen, M., Knudson, J., and Rasmussen, L. (1986) *Trends Biochem. Sci.* **11**, 354 – 355
27. Cooney, R. V., Mumma, R. O., and Benson, A. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4262 – 4264
28. Renou, J.-P., Giziewicz, J. B., Smith, I. C. P., and Jarrell, H. C. (1989) *Biochemistry* **28**, 1804 – 1814
29. Gounaris, K., and Barber, J. (1983) *Trends Biochem. Sci.* **8**, 378 – 381
30. Knowles, F. C., and Benson, A. A. (1983) *Trends Biochem. Sci.* **8**, 178 – 180
31. Tupper, S., Wong, P. T. T., Kates, M., and Tanphaichitr, N. (1994) *Biochemistry* **33**, 13250 – 13258
32. Andersson, A.-S., Rilfors, L., Bergqvist, M., Persson, S., and Lindblom, G. (1996) *Biochemistry* **35**, 11119 – 11130

32a. Vikström, S., Li, L., and Wieslander, Å. (2000) *J. Biol. Chem.* **275**, 9296 – 9302

33. Nishihara, M., Utagawa, M., Akutsu, H., and Koga, Y. (1992) *J. Biol. Chem.* **267**, 12432 – 12435
34. Powles, T. J., Easty, D. M., Easty, G. C., Bondy, P. K., and Munro-Neville, A. (1973) *Nature (London), New Biol.* **245**, 83
35. Lynch, D. V. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 285 – 308, CRC Press, Boca Raton, Florida
36. Araki, S., Satake, M., Ando, S., Hayashi, A., and Fujii, N. (1986) *J. Biol. Chem.* **261**, 5138 – 5144
37. Araki, S., Abe, S., Odani, S., Ando, S., Fujii, N., and Satake, M. (1987) *J. Biol. Chem.* **262**, 14141 – 14145
38. Song, W., and Rintoul, D. A. (1989) *Biochemistry* **28**, 4194 – 4200
39. Hsieh, T. C.-Y., Lester, R. L., and Laine, R. A. (1981) *J. Biol. Chem.* **256**, 7747 – 7755
40. Barr, K., and Lester, R. L. (1984) *Biochemistry* **23**, 5581 – 5588
41. Dickson, R. C., Nagiec, E. E., Wells, G. B., Nagiec, M. M., and Lester, R. L. (1997) *J. Biol. Chem.* **272**, 29620 – 29625
42. Yamakawa, T., and Nagai, Y. (1978) *Trends Biochem. Sci.* **3**, 128 – 131
43. Hansson, G. C. (1983) *J. Biol. Chem.* **258**, 9612 – 9615
44. Clausen, H., Holmes, E., and Hakomori, S.-I. (1986) *J. Biol. Chem.* **261**, 1388 – 1392
45. Farooqui, A. A. (1981) *Adv. Lipid Res.* **18**, 159 – 202
46. Godchaux, W., III, and Leadbetter, E. R. (1984) *J. Biol. Chem.* **259**, 2982 – 2990
47. Abbanat, D. R., Leadbetter, E. R., Godchaux, W., III, and Escher, A. (1986) *Nature (London)* **324**, 367 – 369
48. Jones, M. N., and Chapman, D., eds. (1994) *Micelles, Monolayers and Biomembranes*, Wiley, New York
49. Jain, M. (1988) *Introduction to Biological Membranes*, 2nd ed., Wiley, New York
50. Lofan, R., and Nicolson, G. L. (1981) in *Advanced Cell Biology* (Schwartz, L. M., and Azar, M. M., eds), Van Nostrand Reinhold Co., New York
51. Storch, J., and Kleinfeld, A. M. (1985) *Trends Biochem. Sci.* **10**, 418 – 421
52. Robertson, R. N. (1983) *The Lively Membranes*, Cambridge Univ. Press, Cambridge
53. Starzak, M. (1984) *The Physical Chemistry of Membranes*, Academic Press, San Diego, California
54. Gennis, R. B. (1989) *Biomembranes: Molecular Structure and Function*, Springer-Verlag, New York
55. Morell, P., and Norton, W. T. (1980) *Sci. Am.* **242**(May), 88 – 116
56. Martenson, R. E., ed. (1992) *Myelin: Biology and Chemistry*, CRC Press, Boca Raton, Florida
57. Weimbs, T., and Stoffel, W. (1994) *Biochemistry* **33**, 10408 – 10415
58. Korn, E. D. (1966) *Science* **153**, 1491 – 1498
59. Tanford, C. (1980) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, 2nd ed., Wiley, New York
60. Danielli, J. F., and Davsen, H. (1985) *J. Cell. Comp. Physiol.* **5**, 495 – 508
61. Singer, S. J., and Nicolson, G. L. (1972) *Science* **175**, 720 – 731
62. Bretscher, M. S., and Raff, M. C. (1975) *Nature (London)* **258**, 43 – 49
63. Quinn, P. J., and Cherry, R. J., eds. (1992) *Structural and Dynamic Properties of Lipids and Membranes*, Portland Press, London
64. Disalvo, E. A., and Simon, S. A., eds. (1995) *Permeability and Stability of Lipid Bilayers*, CRC Press, Boca Raton, Florida
65. Jacobson, K., Sheets, E. D., and Simson, R. (1995) *Science* **268**, 1441 – 1442
66. Pascher, I., Lundmark, M., Nyholm, P.-G., and Sundell, S. (1992) *Biochim. Biophys. Acta.* **1113**, 339 – 373
67. Cameron, D. G., Gudgin, E. F., and Mantsch, H. H. (1981) *Biochemistry* **20**, 4496 – 4500
68. Zasadzinski, J. A. N., Helm, C. A., Longo, M. L., Weisenhorn, A. L., Gould, S. A. C., and Hansma, P. K. (1991) *Biophys. J.* **59**, 755 – 760
69. Wiener, M. C., and White, S. H. (1992) *Biophys. J.* **61**, 428 – 433
70. Smith, D. P. E., Bryant, A., Quate, C. F., Rabe, J. P., Gerber, C. H., and Swalen, J. D. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 969 – 972
71. Yang, J., Tamm, L. K., Tillack, T. W., and Shao, Z. (1993) *J. Mol. Biol.* **229**, 286 – 290
72. Nagle, J. F., Zhang, R., Tristram-Nagle, S., Sun, W., Petrache, H. I., and Suter, R. M. (1996) *Biophys. J.* **70**, 1419 – 1431
73. Mueller, P., Rudin, D. O., Tien, H. T., and Wescott, W. C. (1962) *Nature (London)* **194**, 979 – 980
74. Howard, K. P., and Prestegard, J. H. (1995) *J. Am. Chem. Soc.* **117**, 5031 – 5040
75. Ostro, M. J. (1987) *Sci. Am.* **256**(Jan), 103 – 111
76. Philippot, J. R., and Schuber, F., eds. (1995) *Liposomes as Tools in Basic Research and Industry*, CRC Press, Boca Raton, Florida
77. Harwood, J. L. (1992) *Trends Biochem. Sci.* **17**, 203 – 204

77a. Schröder, M. (1980) *Chem. Rev.* **80**, 187 – 213

77b. Mohapatra, S. K., and Behrman, E. J. (1982) *J. Inorg. Biochem* **16**, 85 – 89

77c. Deetz, J. S., and Behrman, E. J. (1981) *Intl. J. Peptide Prot. Res.* **17**, 495 – 500

78. Norrby, P.-O., Becker, H., and Sharpless, K. B. (1996) *J. Am. Chem. Soc.* **118**, 35 – 42
79. Montal, M., and Mueller, P. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3561 – 3566
80. Ladha, S., Mackie, A. R., Harvey, L. J., Clark, D. C., Lea, E. J. A., Brullemans, M., and Duclohier, H. (1996) *Biophys. J.* **71**, 1364 – 1373
81. Jones, M. N., and Chapman, D. (1995) *Micelles, Monolayers, and Biomembranes*, Wiley-Liss, New York

81a. Knobler, C. M. (1990) *Adv. Chem. Phys.* **77**, 397 – 449

82. Bryce, M. R., and Petty, M. C. (1995) *Nature (London)* **374**, 771 – 776
83. Zasadzinski, J. A., Viswanathan, R., Madsen, L., Garnaes, J., and Schwartz, D. K. (1994) *Science* **263**, 1726 – 1733
84. Fettiplace, R., Andrews, D. M., and Haydon, D. A. (1971) *J. Membr. Biol.* **5**, 277 – 296
85. White, S. W., and King, G. I. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6532 – 6536
86. Snyder, R. G., Maroncelli, M., Qi, S. P., and Strauss, H. L. (1981) *Science* **214**, 188 – 190
87. Di, L., and Small, D. M. (1995) *Biochemistry* **34**, 16672 – 16677
88. Tu, K., Tobias, D. J., Blasie, J. K., and Klein, M. L. (1996) *Biophys. J.* **70**, 595 – 608
89. Katsaras, J., Raghunathan, V. A., Dufourc, E. J., and Dufourcq, J. (1995) *Biochemistry* **34**, 4684 – 4688
90. Snyder, R. G., Liang, G. L., Strauss, H. L., and Mendelsohn, R. (1996) *Biophys. J.* **71**, 3186 – 3198
91. Chen, S. C., Sturtevant, J. M., and Gaffney, B. J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5060 – 5063

References

92. Sun, W.-J., Tristram-Nagle, S., Suter, R. M., and Nagle, J. F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7008–7012
93. Mou, J., Yang, J., and Shao, Z. (1994) *Biochemistry* **33**, 4439–4443
94. Wang, Z.-q, Lin, H.-n, Li, S., and Huang, C.-h. (1995) *J. Biol. Chem.* **270**, 2014–2023
95. Wang, G., Lin, H.-n, Li, S., and Huang, C.-h. (1995) *J. Biol. Chem.* **270**, 22738–22746
96. Wang, Z.-q, Lin, H.-n, Li, S., and Huang, C.-h. (1994) *J. Biol. Chem.* **269**, 23491–23499
97. Dufourc, E. J., Parish, E. J., Chitrakorn, S., and Smith, I. C. P. (1984) *Biochemistry* **23**, 6062–6071
98. Maulik, P. R., and Shipley, G. G. (1996) *Biochemistry* **35**, 8025–8034
99. McMullen, T. P. W., and McElhaney, R. N. (1997) *Biochemistry* **36**, 4979–4986
100. Finegold, L., ed. (1993) *Cholesterol in Membrane Models*, CRC Press, Boca Raton, Florida
101. Recktenwald, D. J., and McConnell, H. M. (1981) *Biochemistry* **20**, 4505–4510
102. Needham, D., McIntosh, T. J., and Evans, E. E. (1988) *Biochemistry* **27**, 4668–4673
103. Goldfine, H., Johnston, N. C., and Phillips, M. C. (1981) *Biochemistry* **20**, 2908–2916
104. Lewis, R. V. A. H., Pohle, W., and McElhaney, R. N. (1996) *Biophys. J.* **70**, 2736–2746
105. Shimshick, E. J., and McConnell, H. M. (1973) *Biochemistry* **12**, 2351–2360
106. Almeida, P. F. F., Vaz, W. L. C., and Thompson, T. E. (1992) *Biochemistry* **31**, 7198–7210
107. Träuble, H., and Eibl, H. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 214–219
108. Basánez, G., Nieva, J. L., Rivas, E., Alonso, A., and Goni, F. M. (1996) *Biophys. J.* **70**, 2299–2306
109. Smith, R. L., and Oldfield, E. (1984) *Science* **225**, 280–288

109a. Smith, I. C. P. (1983) in *NMR of Newly Accessible Nuclei*, Vol. 2 (Laslo, P., ed), pp. 1–26, Academic Press, New York

110. Marassi, F. M., and Macdonald, P. M. (1992) *Biochemistry* **31**, 10031–10036
111. Oldfield, E., Bowers, J. L., and Forbes, J. (1987) *Biochemistry* **26**, 6919–6923
112. Bruzik, K. S., and Nyholm, P.-G. (1997) *Biochemistry* **36**, 566–575
113. Hong, M., Schmidt-Rohr, K., and Nanz, D. (1995) *Biophys. J.* **69**, 1939–1950
114. Thurmond, R. L., Niemi, A. R., Lindblom, G., Wieslander, Å., and Rilfors, L. (1994) *Biochemistry* **33**, 13178–13188
115. Boden, N., Jones, S. A., and Sixl, F. (1991) *Biochemistry* **30**, 2146–2155
116. Davis, J. H. (1993) in *Cholesterol in Membrane Models* (Finegold, L., ed), pp. 67–157, CRC Press, Boca Raton, Florida
117. Baenziger, J. E., Jarrell, H. C., and Smith, I. C. P. (1992) *Biochemistry* **31**, 3377–3385
118. De Boeck, H., and Zidovetzki, R. (1992) *Biochemistry* **31**, 623–630
119. de Kruijft, B., Cullis, P. R., and Verkleij, A. J. (1980) *Trends Biochem. Sci.* **5**, 79–81
120. Peng, X., and Jonas, J. (1992) *Biochemistry* **31**, 6383–6390
121. Hong, M., Schmidt-Rohr, K., and Zimmermann, H. (1996) *Biochemistry* **35**, 8335–8341
122. Lu, D., Singh, D., Morrow, M. R., and Grant, C. W. M. (1993) *Biochemistry* **32**, 290–297
123. Salmon, A., and Hamilton, J. A. (1995) *Biochemistry* **34**, 16065–16073
124. Ruocco, M. J., Siminovitch, D. J., Long, J. R., Das Gupta, S. K., and Griffin, R. G. (1996) *Biophys. J.* **71**, 1776–1788
125. Bruzik, K. S., and Harwood, J. S. (1997) *J. Am. Chem. Soc.* **119**, 6629–6637
126. Casal, H. L., and McElhaney, R. N. (1990) *Biochemistry* **29**, 5423–5427
127. Hübner, W., Mantsch, H. H., Paltauf, F., and Hauser, H. (1994) *Biochemistry* **33**, 320–326
128. Léger, C. L., Daveloose, D., Christon, R., and Viret, J. (1990) *Biochemistry* **29**, 7269–7275
129. Jakobsson, E. (1997) *Trends Biochem. Sci.* **22**, 339–344
130. Berger, O., Edholm, O., and Jähnig, F. (1997) *Biophys. J.* **72**, 2002–2013
131. Chiu, S.-W., Clark, M., Balaji, V., Subramaniam, S., Scott, H. L., and Jakobsson, E. (1995) *Biophys. J.* **69**, 1230–1245
132. Browning, J. L. (1981) *Biochemistry* **20**, 7144–7151
133. Cevc, G., Watts, A., and Marsh, D. (1981) *Biochemistry* **20**, 4955–4965
134. Harlos, K., and Eibl, H. (1980) *Biochemistry* **19**, 895–899
135. Koynova, R. D., Kuttenreich, H. L., Tenchov, B. G., and Hinz, H.-J. (1988) *Biochemistry* **27**, 4612–4619
136. Hayakawa, E., Naganuma, M., Mukasa, K., Shimozawa, T., and Araiso, T. (1998) *Biophys. J.* **74**, 892–898
137. Katsaras, J., Jeffrey, K. R., Yang, D. S.-C., and Epand, R. M. (1993) *Biochemistry* **32**, 10700–10707
138. Srisiri, W., Sisson, T. M., O'Brien, D. F., McGrath, K. M., Han, Y., and Gruner, S. M. (1997) *J. Am. Chem. Soc.* **119**, 4866–4873
139. Siegel, D. P., and Epand, R. M. (1997) *Biophys. J.* **73**, 3089–3111
140. Thurmond, R. L., Lindblom, G., and Brown, M. F. (1993) *Biochemistry* **32**, 5394–5410
141. Siegel, D. P., Banschbach, J., and Yeagle, P. L. (1989) *Biochemistry* **28**, 5010–5019
142. Siegel, D. P., and Banschbach, J. L. (1990) *Biochemistry* **29**, 5975–5981
143. Landau, E. M., and Rosenbusch, J. P. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14532–14535
144. Delacroix, H., Gulik-Krzywicki, T., and Seddon, J. M. (1996) *J. Mol. Biol.* **258**, 88–103

144a. Tenchov, B., Koynova, R., and Rapp, G. (1998) *Biophys. J.* **75**, 853–866

145. Russell, N. J. (1984) *Trends Biochem. Sci.* **9**, 108–112
146. deMendoza, D., and Cronan, J. E., Jr. (1983) *Trends Biochem. Sci.* **8**, 49–52
147. Mantsch, H. H., Madec, C., Lewis, R. N. A. H., and McElhaney, R. N. (1987) *Biochemistry* **26**, 4045–4049
148. Bechtel, D. B., Mueller, D. D., Whaley, T. W., and Bulla, L. A., Jr. (1985) *J. Biol. Chem.* **260**, 9784–9792
149. Bérubé, L. R., and Hollingsworth, R. I. (1995) *Biochemistry* **34**, 12005–12011
150. Huster, D., Jin, A. J., Arnold, K., and Gawrisch, K. (1997) *Biophys. J.* **73**, 855–864
151. Träuble, H. (1971) *J. Membr. Biol.* **4**, 193–208
152. Wilson, M. A., and Pohorille, A. (1996) *J. Am. Chem. Soc.* **118**, 6580–6587
153. Ho, C., and Stubbs, C. D. (1997) *Biochemistry* **36**, 10630–10637
154. Bretscher, M. S. (1980) *Trends Biochem. Sci.* **5**, VI–VII
155. Devaux, P., and McConnell, H. M. (1972) *J. Am. Chem. Soc.* **94**, 4475–4481
156. Kell, D. B. (1984) *Trends Biochem. Sci.* **9**, 86–88
157. Picard, F., Paquet, M.-J., Dufourc, E. J., and Auger, M. (1998) *Biophys. J.* **74**, 857–868
158. Schmidt, T., Schütz, G. J., Baumgartner, W., Gruber, H. J., and Schindler, H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2926–2929
159. Sheets, E. D., Lee, G. M., Simson, R., and Jacobson, K. (1997) *Biochemistry* **36**, 12449–12458
160. Hopkins, C. R. (1992) *Trends Biochem. Sci.* **17**, 27–32
161. Kleinfeld, A. M., Chu, P., and Storch, J. (1997) *Biochemistry* **36**, 5702–5711
162. Zhao, J., Zhou, Q., Wiedmer, T., and Sims, P. J. (1998) *Biochemistry* **37**, 6361–6366
163. Wolff, D., Canessa-Fischer, M., Vargas, F., and Diaz, G. (1971) *J. Membr. Biol.* **6**, 304–314
164. Sherbet, G. V. (1978) *The Biophysical Characterisation of the Cell Surface*, Academic Press, London
165. Jones, M. N. (1975) *Biological Interfaces*, Elsevier, Amsterdam
166. McLaughlin, S. (1989) *Ann. Rev. Biophys. Biophys. Chem.* **18**, 113–136
167. Aidley, D. J. (1971) *The Physiology of Excitable Cells*, Cambridge Univ. Press, London and New York
168. Nystrom, R. A. (1973) *Membrane Physiology*, Prentice-Hall, Englewood Cliffs, New Jersey
169. Nicholls, D. G., and Ferguson, S. J. (1992) *Bioenergetics 2*, Academic Press, London
170. Wood, P. M. (1985) *Trends Biochem. Sci.* **10**, 106–107
171. Teissié, J., Prats, M., LeMassu, A., Stewart, L. C., and Kates, M. (1990) *Biochemistry* **29**, 59–65
172. Teissié, J., Gabriel, B., and Prats, M. (1993) *Trends Biochem. Sci.* **18**, 243–246
173. Heberle, J., Riesle, J., Thiedemann, G., Oesterhelt, D., and Dencher, N. A. (1994) *Nature (London)* **370**, 379–382
174. Gabriel, B., and Teissié, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14521–14525
175. Devaux, P. F. (1991) *Biochemistry* **30**, 1163–1173
176. Victorov, A. V., Janes, N., Taraschi, T. F., and Hoek, J. B. (1997) *Biophys. J.* **72**, 2588–2598
177. Shiffer, K. A., Rood, L., Emerson, R. K., and Kuypers, F. A. (1998) *Biochemistry* **37**, 3449–3458
178. Michel, H. (1983) *Trends Biochem. Sci.* **8**, 56–59
179. Cowan, S. W., and Rosenbusch, J. P. (1994) *Science* **264**, 914–916
180. von Heijne, G. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 167–192
181. Bloom, M. (1995) *Biophys. J.* **69**, 1631–1632
182. Weimbs, T., and Stoffel, W. (1992) *Biochemistry* **31**, 12289–12296
183. Yamaguchi, Y., Ikenaka, K., Niinobe, M., Yamada, H., and Mikoshiba, K. (1996) *J. Biol. Chem.* **271**, 27838–27846
184. Jennings, M. L. (1989) *Ann. Rev. Biochem.* **58**, 999–1027
185. Bowie, J. U. (1997) *J. Mol. Biol.* **272**, 780–789
186. Gafvelin, G., Sakaguchi, M., Andersson, H., and von Heijne, G. (1997) *J. Biol. Chem.* **272**, 6119–6127
187. Jähnig, F. (1990) *Trends Biochem. Sci.* **15**, 93–95
188. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1994) *Biochemistry* **33**, 3038–3049
189. Persson, B., and Argos, P. (1996) *Protein Sci.* **5**, 363–371
190. Fasman, G. D., and Gilbert, W. A. (1990) *Trends Biochem. Sci.* **15**, 89–92
191. Milpetz, F., Argos, P., and Persson, B. (1995) *Trends Biochem. Sci.* **20**, 204–205
192. Shai, Y. (1995) *Trends Biochem. Sci.* **20**, 460–464

192a. Killian, J. A., and von Heijne, G. (2000) *Trends Biochem. Sci.* **25**, 429–434

192b. Lew, S., Ren, J., and London, E. (2000) *Biochemistry* **39**, 9632–9640

193. Popov, M., Tam, L. Y., Li, J., and Reithmeier, R. A. F. (1997) *J. Biol. Chem.* **272**, 18325–18332
194. Lee, G. F., Burrows, G. G., Lebert, M. R., Dutton, D. P., and Hazelbauer, G. L. (1994) *J. Biol. Chem.* **269**, 29920–29927
195. Lee, G. F., Dutton, D. P., and Hazelbauer, G. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5416–5420
196. Yu, H., Kono, M., McKee, T. D., and Oprian, D. D. (1995) *Biochemistry* **34**, 14963–14969
197. Sansom, M. S. P., Son, H. S., Sankararamakrishnan, R., Kerr, I. D., and Breed, J. (1995) *Biophys. J.* **68**, 1295–1310

References

198. Stopar, D., Jansen, K. A. J., Páli, T., Marsh, D., and Hemminga, M. A. (1997) *Biochemistry* **36**, 8261–8268
199. Nikaido, H. (1994) *J. Biol. Chem.* **269**, 3905–3908
200. Mishra, V. K., and Palgunachari, M. N. (1996) *Biochemistry* **35**, 11210–11220
201. Mishra, V. K., Palgunachari, M. N., Lund-Katz, S., Phillips, M. C., Segrest, J. P., and Anantharamaiah, G. M. (1995) *J. Biol. Chem.* **270**, 1602–1611
202. Casey, P. J. (1995) *Science* **268**, 221–225
203. McIlhinney, R. A. J. (1990) *Trends Biochem. Sci.* **15**, 387–391
204. Buss, J. E., Mumby, S. M., Casey, P. J., Gilman, A. G., and Sefton, B. M. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7493–7497
205. Johnson, D. R., Bhatnagar, R. S., Knoll, L. J., and Gordon, J. I. (1994) *Ann. Rev. Biochem.* **63**, 869–914
206. Peseckis, S. M., and Resh, M. D. (1994) *J. Biol. Chem.* **269**, 30888–30892
207. McLaughlin, S., and Aderem, A. (1995) *Trends Biochem. Sci.* **20**, 272–276
208. Milligan, G., Parenti, M., and Magee, A. I. (1995) *Trends Biochem. Sci.* **20**, 181–186
209. Wilson, P. T., and Bourne, H. R. (1995) *J. Biol. Chem.* **270**, 9667–9675
210. Duncan, J. A., and Gilman, A. G. (1996) *J. Biol. Chem.* **271**, 23594–23600
211. Marshall, C. J. (1993) *Science* **259**, 1865–1866
212. Glomset, J. A., Gelb, M. H., and Farnsworth, C. C. (1990) *Trends Biochem. Sci.* **15**, 139–142
213. Casey, P. J., and Seabra, M. C. (1996) *J. Biol. Chem.* **271**, 5289–5292
214. Shipton, C. A., Parmryd, I., Swiezewska, E., Andersson, B., and Dallner, G. (1995) *J. Biol. Chem.* **270**, 566–572
215. Wedegaertner, P. B., Wilson, P. T., and Bourne, H. R. (1995) *J. Biol. Chem.* **270**, 503–506
216. Kikuchi, A., Sagami, H., and Ogura, K. (1999) *J. Biol. Chem.* **274**, 18011–18016
217. Turner, A. J. (1994) in *Essays in Biochemistry*, Vol. 28 (Tipton, K. F., ed), pp. 113–128, Portland Press, London
218. Low, M. G. (1987) *Biochem. J.* **244**, 1–13
219. Lisanti, M. P., and Rodriguez-Boulan, E. (1990) *Trends Biochem. Sci.* **15**, 113–118
220. Thomas, J. R., Dwek, R. A., and Rademacher, T. W. (1990) *Biochemistry* **29**, 5413–5422
221. Brewis, I. A., Ferguson, M. A. J., Mehlert, A., Turner, A. J., and Hooper, N. M. (1995) *J. Biol. Chem.* **270**, 22946–22956
222. Previato, J. O., Jones, C., Xavier, M. T., Wait, R., Travassos, L. R., Parodi, A. J., and Mendonça-Previato, L. (1995) *J. Biol. Chem.* **270, No 13**, 7241–7250
223. Heinz, D. W., Ryan, M., Smith, M. P., Weaver, L. H., Keana, J. F. W., and Griffith, O. H. (1996) *Biochemistry* **35**, 9496–9504
224. Williams, A. F., and Gagnon, J. (1982) *Science* **216**, 696–703
225. Mehlert, A., Richardson, J. M., and Ferguson, M. A. J. (1998) *J. Mol. Biol.* **277**, 379–392
225a. Eisenhaber, B., Bork, P., Yuan, Y., Löffler, G., and Eisenhaber, F. (2000) *Trends Biochem. Sci.* **25**, 340–341
226. Reynolds, J. A., and Trayer, H. (1971) *J. Biol. Chem.* **246**, 7337–7342
227. Pryde, J. G. (1986) *Trends Biochem. Sci.* **11**, 160–163
228. Bennett, V. (1985) *Ann. Rev. Biochem.* **54**, 273–304
229. Justice, J. M., Murtagh, J. J., Jr., Moss, J., and Vaughan, M. (1995) *J. Biol. Chem.* **270**, 17970–17976
230. Neville, D. M., Jr., and Glossmann, H. (1971) *J. Biol. Chem.* **246**, 6335–6338
231. Inaba, M., and Maede, Y. (1988) *J. Biol. Chem.* **263**, 17763–17771
232. Steck, T. L. (1974) *J. Cell Biol.* **62**, 1–19
233. Luna, E. J., Kidd, G. H., and Branton, D. (1979) *J. Biol. Chem.* **254**, 2526–2532
234. Viitala, J., and Järnefelt, J. (1985) *Trends Biochem. Sci.* **10**, 392–395
235. Marchesi, V. T., Furthmayr, H., and Tomita, M. (1976) *Ann. Rev. Biochem.* **45**, 667–698
236. Jay, D., and Cantley, L. (1986) *Ann. Rev. Biochem.* **55**, 511–538
237. Murthy, S. N. P., Wilson, J., Zhang, Y., and Lorand, L. (1994) *J. Biol. Chem.* **269**, 22907–22911
238. Tanner, M. J. A. (1996) *Nature (London)* **382**, 209–210
239. Müller-Berger, S., Karbach, D., König, J., Lepke, S., Wood, P. G., Appelhans, H., and Passow, H. (1995) *Biochemistry* **34**, 9315–9324
240. Erickson, H. K. (1997) *Biochemistry* **36**, 9958–9967
241. Fukuda, M., Lauffenburger, M., Sasaki, H., Rogers, M. E., and Dell, A. (1987) *J. Biol. Chem.* **262**, 11952–11957
242. MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997) *Science* **276**, 131–133
243. Challou, N., Goormaghtigh, E., Cabiaux, V., Conrath, K., and Ruysschaert, J.-M. (1994) *Biochemistry* **33**, 6902–6910
244. Huang, C.-H., Reid, M., Daniels, G., and Blumenfeld, O. O. (1993) *J. Biol. Chem.* **268**, 25902–25908
245. Marchesi, V. T. (1985) *Ann. Rev. Cell Biol.* **1**, 531–561
246. Fujita, T., Ralston, G. B., and Morris, M. B. (1998) *Biochemistry* **37**, 264–271
247. Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S. C., and Branton, D. (1993) *Science* **262**, 2027–2030
248. Cherry, L., Menhart, N., and Fung, L. W.-M. (1999) *J. Biol. Chem.* **274**, 2077–2084
249. Begg, G. E., Harper, S. L., Morris, M. B., and Speicher, D. W. (2000) *J. Biol. Chem.* **275**, 3279–3287
250. Fujita, T., Ralston, G. B., and Morris, M. B. (1998) *Biochemistry* **37**, 272–280
250a. Rief, M., Pascual, J., Saraste, M., and Gaub, H. E. (1999) *J. Mol. Biol.* **286**, 553–561
251. Fowler, V. M. (1986) *Nature (London)* **322**, 777–778
252. Carraway, K. L. (1975) *Biochim. Biophys. Acta.* **415**, 379–410
253. Byers, T. J., and Branton, D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6153–6157
254. Elgsaeter, A., Stokke, B. T., Mikkelsen, A., and Branton, D. (1986) *Science* **234**, 1217–1223
255. Pantazatos, D. P., and MacDonald, R. I. (1997) *J. Biol. Chem.* **272**, 21052–21059
256. Luna, E. J., and Hitt, A. L. (1992) *Science* **258**, 955–964
257. Reichstein, E., and Blostein, R. (1975) *J. Biol. Chem.* **250**, 6256–6263
258. Bennett, V. (1992) *J. Biol. Chem.* **267**, 8703–8706
259. Michaely, P., and Bennett, V. (1993) *J. Biol. Chem.* **268**, 22703–22709
260. Kordeli, E., Lambert, S., and Bennett, V. (1995) *J. Biol. Chem.* **270**, 2352–2359
261. An, X.-L., Takakuwa, Y., Nunomura, W., Manno, S., and Mohandas, N. (1996) *J. Biol. Chem.* **271**, 33187–33191
262. Workman, R. F., and Low, P. S. (1998) *J. Biol. Chem.* **273**, 6171–6176
263. Malik, S., Sami, M., and Watts, A. (1993) *Biochemistry* **32**, 10078–10084
264. Azim, A. C., Knoll, J. H. M., Beggs, A. H., and Chishti, A. H. (1995) *J. Biol. Chem.* **270**, 17407–17413
265. Wolfe, L. C., John, K. M., Falcone, J. C., Byrne, A. M., and Lux, S. E. (1982) *N. Engl. J. Med.* **307**, 1367–1374
266. Diakowski, W., and Sikorski, A. F. (1995) *Biochemistry* **34**, 13252–13258
267. Marfatia, S. M., Lue, R. A., Branton, D., and Chishti, A. H. (1995) *J. Biol. Chem.* **270**, 715–719
268. Krauss, S. W., Chasis, J. A., Rogers, C., Mohandas, N., Krockmalnic, G., and Penman, S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7297–7302
269. Khodadad, J. K., Waugh, R. E., Podolski, J. L., Josephs, R., and Steck, T. L. (1996) *Biophys. J.* **70**, 1036–1044
270. Hansen, J. C., Skalak, R., Chien, S., and Hoger, A. (1996) *Biophys. J.* **70**, 146–166
271. Rice-Evans, C. A., and Dunn, M. J. (1982) *Trends Biochem. Sci.* **7**, 282–286
272. Mohandas, N., and Evans, E. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 787–818
273. Agre, P., Casella, J. F., Zinkham, W. H., McMillan, C., and Bennett, V. (1985) *Nature (London)* **314**, 380–383
274. Baines, A. J., and Bennett, V. (1985) *Nature (London)* **315**, 410–413
275. Clark, E. A., and Brugge, J. S. (1995) *Science* **268**, 233–239
276. Dejana, E., Corada, M., and Lampugnani, M. G. (1995) *FASEB J.* **9**, 910–918
277. Malik, R. K., and Parsons, J. T. (1996) *J. Biol. Chem.* **271**, 29785–29791
278. Lo, S. H., An, Q., Bao, S., Wong, W.-K., Liu, Y., Janmey, P. A., Hartwig, J. H., and Chen, L. B. (1994) *J. Biol. Chem.* **269**, 22310–22319
279. Horwitz, A. F. (1997) *Sci. Am.* **276**(May), 68–75
280. de Pereda, J. M., Wiche, G., and Liddington, R. C. (1999) *EMBO J.* **18**, 4087–4095
281. Springer, T. A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 65–72
282. Bazzoni, G., and Hemler, M. E. (1998) *Trends Biochem. Sci.* **23**, 30–34
283. Emsley, J., King, S. L., Bergelson, J. M., and Liddington, R. C. (1997) *J. Biol. Chem.* **272**, 28512–28517
284. Luscinskas, F. W., and Lawler, J. (1994) *FASEB J.* **8**, 929–938
285. Simon, K. O., Otey, C. A., Pavelko, F. M., and Burridge, K. (1991) *Curr. Top. Membr. Transp.* **38**, 57–64
286. Edelman, G. M. (1983) *Science* **219**, 450–457
287. Becker, J. W., Erickson, H. P., Hoffman, S., Cunningham, B. A., and Edelman, G. M. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1088–1092
288. Cunningham, B. A., Hemperly, J. J., Murray, B. A., Prediger, E. A., Brackenbury, R., and Edelman, G. M. (1987) *Science* **236**, 799–806
289. Kudo, M., Kitajima, K., Inoue, S., Shiokawa, K., Morris, H. R., Dell, A., and Inoue, Y. (1996) *J. Biol. Chem.* **271**, 32667–32677
290. Newman, P. J., Berndt, M. C., Gorski, J., White, G. C., II, Lyman, S., Paddock, C., and Muller, W. A. (1990) *Science* **247**, 1219–1222
291. DeLisser, H. M., Yan, H. C., Newman, P. J., Muller, W. A., Buck, C. A., and Albelda, S. M. (1993) *J. Biol. Chem.* **268**, 16037–16046
291a. Chen, H. A., Pfuhl, M., McAlister, M. S. B., and Driscoll, P. C. (2000) *Biochemistry* **39**, 6814–6824
292. Casasnovas, J. M., Stehle, T., Liu, J.-h, Wang, J.-h, and Springer, T. A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4134–4139
293. Wang, J.-H., Pepinsky, R. B., Stehle, T., Liu, J.-H., Karpusas, M., Browning, B., and Osborn, L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5714–5718
294. Springer, T. A. (1990) *Nature (London)* **346**, 425–434

References

295. Wagner, G. (1995) *Science* **267**, 342

296. Nagar, B., Overduin, M., Ikura, M., and Rini, J. M. (1996) *Nature (London)* **380**, 360–364

297. Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grübel, G., Legrand, J.-F., Als-Nielsen, J., Colman, D. R., and Hendrickson, W. A. (1995) *Nature (London)* **374**, 327–336

298. Overduin, M., Harvey, T. S., Bagby, S., Tong, K. I., Yau, P., Takeichi, M., and Ikura, M. (1995) *Science* **267**, 386–389

298a. Nollet, F., Kools, P., and van Roy, F. (2000) *J. Mol. Biol.* **299**, 551–572

298b. Pertz, O., Bozic, D., Koch, A. W., Fauser, C., Brancaccio, A., and Engel, J. (1999) *EMBO J.* **18**, 1738–1747

299. Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M., and Rossmann, M. G. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4140–4145

300. Patel, D. J., and Gumbiner, B. M. (1995) *Nature (London)* **374**, 306–307

301. Berx, G., Cleton-Jansen, A.-M., Nollet, F., de Leeuw, W. J. F., van de Vijver, M. J., Cornelisse, C., and van Roy, F. (1995) *EMBO J.* **14**, 6107–6115

302. Vartio, T., and Vaheri, A. (1983) *Trends Biochem. Sci.* **8**, 442–444

303. Hynes, R. O. (1986) *Sci. Am.* **254**(Jun), 42–51

304. Ruoslahti, E. (1988) *Ann. Rev. Biochem.* **57**, 375–413

305. Engel, J., Odermatt, E., Engel, A., Madri, J. A., Furthmayr, H., Rohde, H., and Timpl, R. (1981) *J. Mol. Biol.* **150**, 97–120

306. Ingham, K. C., Brew, S. A., Huff, S., and Litvinovich, S. V. (1997) *J. Biol. Chem.* **272**, 1718–1724

307. Johnson, K. J., Sage, H., Briscoe, G., and Erickson, H. P. (1999) *J. Biol. Chem.* **274**, 15473–15479

308. Grant, R. P., Spitzfaden, C., Altroff, H., Campbell, I. D., and Mardon, H. J. (1997) *J. Biol. Chem.* **272**, 6159–6166

309. Hynes, R. O. (1990) *Fibronectins*, Springer-Verlag, New York

310. Moyano, J. V., Carnemolla, B., Domínguez-Jiménez, C., García-Gila, M., Albar, J. P., Sánchez-Aparicio, P., Leprini, A., Querzé, G., Zardi, L., and Garcia-Pardo, A. (1997) *J. Biol. Chem.* **272**, 24832–24836

311. Greenberg, C. S., Birckbichler, P. J., and Rice, R. H. (1991) *FASEB J.* **5**, 3071–3077

312. Jeong, J.-M., Murthy, S. N. P., Radek, J. T., and Lorand, L. (1995) *J. Biol. Chem.* **270**, 5654–5658

313. An, S. S. A., Jiménez-Barbero, J., Petersen, T. E., and Llinás, M. (1992) *Biochemistry* **31**, 9927–9933

314. Spitzfaden, C., Grant, R. P., Mardon, H. J., and Campbell, L. D. (1997) *J. Mol. Biol.* **265**, 565–579

315. Kornblihtt, A. R., Pesce, C. G., Alonso, C. R., Cramer, P., Srebrow, A., Werbajh, S., and Muro, A. F. (1996) *FASEB J.* **10**, 248–257

316. Yoneda, A., Ogawa, H., Kojima, K., and Matsumoto, I. (1998) *Biochemistry* **37**, 6351–6360

317. Fischer, D., Chiquet-Ehrismann, R., Bernasconi, C., and Chiquet, M. (1995) *J. Biol. Chem.* **270**, 3378–3384

318. Yokosaki, Y., Matsuura, N., Higashiyama, S., Murakami, I., Obara, M., Yamakido, M., Shigeto, N., Chen, J., and Sheppard, D. (1998) *J. Biol. Chem.* **273**, 11423–11428

319. Denda, S., Müller, U., Crossin, K. L., Erickson, H. P., and Reichardt, L. F. (1998) *Biochemistry* **37**, 5464–5474

320. Oberhauser, A. F., Marszalek, P. E., Erickson, H. P., and Fernandez, J. M. (1998) *Nature (London)* **393**, 181–185

321. Stetefeld, J., Mayer, U., Timpl, R., and Huber, R. (1996) *J. Mol. Biol.* **257**, 644–657

322. Colognato, H., MacCarrick, M., O'Rear, J. J., and Yurchenco, P. D. (1997) *J. Biol. Chem.* **272**, 29330–29336

323. Lane, T. F., and Sage, E. H. (1994) *FASEB J.* **8**, 163–173

324. Venstrom, K. A., and Reichardt, L. F. (1993) *FASEB J.* **7**, 996–1003

325. Gherzi, R., Carnemolla, B., Siri, A., Ponassi, M., Balza, E., and Zardi, L. (1995) *J. Biol. Chem.* **270**, 3429–3434

326. Ruoslahti, E., and Pierschbacher, M. D. (1987) *Science* **238**, 491–497

327. D'Souza, S. E., Ginsberg, M. H., and Plow, E. F. (1991) *Trends Biochem. Sci.* **16**, 246–250

328. Leahy, D. J., Hendrickson, W. A., Aukhil, I., and Erickson, H. P. (1992) *Science* **258**, 987–991

329. Stanley, K. K., Newby, A. C., and Luzio, J. P. (1982) *Trends Biochem. Sci.* **7**, 145–147

330. Roberts, M. F. (1996) *FASEB J.* **10**, 1159–1172

331. Han, S. K., Yoon, E. T., Scott, D. L., Sigler, P. B., and Cho, W. (1997) *J. Biol. Chem.* **272**, 3573–3582

332. Caruthers, A., and Melchior, D. L. (1986) *Trends Biochem. Sci.* **11**, 331–335

333. Gutknecht, J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6443–6446

334. Paula, S., Volkov, A. G., and Deamer, D. W. (1998) *Biophys. J.* **74**, 319–327

335. Stein, W. D. (1990) *Channels, Carriers and Pumps: an Introduction to Membrane Transport*, Academic Press, San Diego, California

335a. Paulsen, I. T., Nguyen, L., Sliwinski, M. K., Rabus, R., and Saier, M. H., Jr. (2000) *J. Mol. Biol.* **301**, 75–100

336. Fox, R. O., Jr., and Richards, F. M. (1982) *Nature (London)* **300**, 325–330

337. Kempf, C., Klausner, R. D., Weinstein, J. N., Van Renswoude, J., Pincus, M., and Blumenthal, R. (1982) *J. Biol. Chem.* **257**, 2469–2476

338. Klebba, P. E., Hofnung, M., and Charbit, A. (1994) *EMBO J.* **13**, 4670–4675

339. Schirmer, T., Keller, T. A., Wang, Y.-F., and Rosenbusch, J. P. (1995) *Science* **267**, 512–514

340. Meyer, J. E. W., Hofnung, M., and Schulz, G. E. (1997) *J. Mol. Biol.* **266**, 761–775

341. Jordy, M., Andersen, C., Schülein, K., Ferenci, T., and Benz, R. (1996) *J. Mol. Biol.* **259**, 666–678

341a. Dumas, F., Koebnik, R., Winterhalter, M., and Van Gelder, P. (2000) *J. Biol. Chem.* **275**, 19747–19751

342. Samartzidou, H., and Delcour, A. H. (1998) *EMBO J.* **17**, 93–100

343. Pauptit, R. A., Schirmer, T., Jansonius, J. N., Rosenbusch, J. P., Parker, M. W., Tucker, A. D., Tsernoglou, D., Weiss, M. S., and Schulz, G. E. (1991) *J. Structural Biol.* **107**, 136–145

344. Kreusch, A., Neubüser, A., Schiltz, E., Weckesser, J., and Schulz, G. E. (1994) *Protein Sci.* **3**, 58–63

344a. Müller, D. J., and Engel, A. (1999) *J. Mol. Biol.* **285**, 1347–1351

345. Bainbridge, G., Mobasheri, H., Armstrong, G. A., Lea, E. J. A., and Lakey, J. H. (1998) *J. Mol. Biol.* **275**, 171–176

346. Van Gelder, P., Saint, N., Phale, P., Eppens, E. F., Prilipov, A., van Boxtel, R., Rosenbusch, J. P., and Tommassen, J. (1997) *J. Mol. Biol.* **269**, 468–472

347. Lyer, R., Wu, Z., Woster, P. M., and Delcour, A. H. (2000) *J. Mol. Biol.* **297**, 933–945

348. Watanabe, M., Rosenbusch, J., Schirmer, T., and Karplus, M. (1997) *Biophys. J.* **72**, 2094–2102

349. Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N., and Rosenbusch, J. P. (1992) *Nature (London)* **358**, 727–733

350. Stathopoulos, C. (1996) *Protein Sci.* **5**, 170–173

351. Mannella, C. A. (1992) *Trends Biochem. Sci.* **17**, 315–320

352. Li, A. Z., Huang, H., Re, X., Qi, L. J., and Marx, K. A. (1998) *Biophys. J.* **74**, 964–973

353. Reumann, S., Maier, E., Benz, R., and Heldt, H. W. (1995) *J. Biol. Chem.* **270**, 17559–17565

354. Chrispeels, M. J., and Agre, P. (1994) *Trends Biochem. Sci.* **19**, 421–425

355. Knepper, M. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6255–6258

355a. Borgnia, M., Nielsen, S., Engel, A., and Agre, P. (1999) *Ann. Rev. Biochem.* **68**, 425–458

356. Ringler, P., Borgnia, M. J., Stahlberg, H., Maloney, P. C., Agre, P., and Engel, A. (1999) *J. Mol. Biol.* **291**, 1181–1190

357. de Groot, B. L., Heymann, J. B., Engel, A., Mitsuoka, K., Fujiyoshi, Y., and Grubmüller, H. (2000) *J. Mol. Biol.* **300**, 987–994

358. Cheng, A., van Hoek, A. N., Yeager, M., Verkman, A. S., and Mitra, A. K. (1997) *Nature (London)* **387**, 627–630

359. Ma, T., Song, Y., Gillespie, A., Carlson, E. J., Epstein, C. J., and Verkman, A. S. (1999) *J. Biol. Chem.* **274**, 20071–20074

359a. Zeuthen, T., and Klaerke, D. A. (1999) *J. Biol. Chem.* **274**, 21631–21636

360. Barone, L. M., Shih, C., and Wasserman, B. P. (1997) *J. Biol. Chem.* **272**, 30672–30677

361. Calamita, G., Bishai, W. R., Preston, G. M., Guggino, W. B., and Agre, P. (1995) *J. Biol. Chem.* **270**, 29063–29066

362. Deen, P. M. T., Verdijk, M. A. J., Knoers, N. V. A. M., Wieringa, B., Monnens, L. A. H., van OS, C. H., and van Oost, B. A. (1994) *Science* **264**, 92–95

363. Bai, L., Fushimi, K., Sasaki, S., and Marumo, F. (1996) *J. Biol. Chem.* **271**, 5171–5176

364. Jan, L. Y., and Jan, Y. N. (1994) *Nature (London)* **371**, 119–122

365. Armstrong, C. (1998) *Science* **280**, 56–57

366. Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., and MacKinnon, R. (1998) *Science* **280**, 69–77

367. Heginbothan, L., Odessey, E., and Miller, C. (1997) *Biochemistry* **36**, 10335–10342

368. MacKinnon, R., Cohen, S. L., Kuo, A., Lee, A., and Chait, B. T. (1998) *Science* **280**, 106–109

369. Aidley, D. J., and Stanfield, P. R. (1996) *Ion Channels, Molecules in Action*, Cambridge Univ. Press, New York

370. Kreusch, A., Pfaffinger, P. J., Stevens, C. F., and Choe, S. (1998) *Nature (London)* **392**, 945–948

370a. Gulbis, J. M., Zhou, M., Mann, S., and MacKinnon, R. (2000) *Science* **289**, 123–127

370b. Kobertz, W. R., Williams, C., and Miller, C (2000) *Biochemistry* **39**, 10347–10352

370c. Roux, B., Bernèche, S., and Im, W. (2000) *Biochemistry* **39**,

371. Klingenberg, M. (1981) *Nature (London)* **290**, 449–454

372. Prince, R. C., Gunson, D. E., and Scarpa, A. (1985) *Trends Biochem. Sci.* **10**, 99

373. Smith, R., Separovic, F., Milne, T. J., Whittaker, A., Bennett, F. M., Cornell, B. A., and Makriyannis, A. (1994) *J. Mol. Biol.* **241**, 456–466

374. Dathe, M., Schümann, M., Wieprecht, T., Winkler, A., Beyermann, M., Krause, E., Matsuzaki, K., Murase, O., and Bienert, M. (1996) *Biochemistry* **35**, 12612–12622

375. Kiyota, T., Lee, S., and Sugihara, G. (1996) *Biochemistry* **35**, 13196–13204

376. Monette, M., and Lafleur, M. (1996) *Biophys. J.* **70**, 2195–2202

377. Ojcius, D. M., and Young, J. D.-E. (1991) *Trends Biochem. Sci.* **16**, 225–229

378. Langs, D. A. (1988) *Science* **241**, 188–191

References

379. Doyle, D. A., and Wallace, B. A. (1997) *J. Mol. Biol.* **266**, 963–977
380. Iqbal, M., and Balaram, P. (1981) *Biochemistry* **20**, 7278–7284
381. Malashkevich, V. N., Kammerer, R. A., Efimov, V. P., Schulthess, T., and Engel, J. (1996) *Science* **274**, 761–765
382. Unwin, N. (1995) *Nature (London)* **373**, 37–43
383. Harrold, F. M., and Baarda, J. R. (1967) *J. Bacteriol.* **94**, 53–60
384. Harold, F. M., and Van Brunt, J. (1977) *Science* **197**, 372–373
385. Pressman, B. C. (1976) *Ann. Rev. Biochem.* **45**, 501–530
386. Hamilton, J. A., Sabesan, M. N., and Steinrauf, L. K. (1981) *J. Am. Chem. Soc.* **103**, 5880–5885
387. Duax, W. L., Hauptman, M., Weeks, C. M., and Norton, D. A. (1972) *Science* **176**, 911–914
388. Karle, I. L., Karle, J., Wieland, T., Burgemeister, W., Faulstich, H., and Witkop, B. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1836–1840
389. Pangborn, W., Duax, W., and Langs, D. (1987) *J. Am. Chem. Soc.* **109**, 2163–2165
390. Karle, I. L. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7155–7159
391. Inabayashi, M., Miyauchi, S., Kamo, N., and Jin, T. (1995) *Biochemistry* **34**, 3455–3460
392. Ledger, P. W., and Tanzer, M. L. (1984) *Trends Biochem. Sci.* **9**, 313–314
393. Marger, M. D., and Saier, M. H., Jr. (1993) *Trends Biochem. Sci.* **18**, 13–20
394. Goswitz, V. C., and Brooker, R. J. (1995) *Protein Sci.* **4**, 534–537
395. Seidel, H. M., and Knowles, J. R. (1994) *Biochemistry* **33**, 5641–5646
396. Cloherty, E. K., Sulzman, L. A., Zottola, R. J., and Carruthers, A. (1995) *Biochemistry* **34**, 15395–15406
397. Shanahan, M. F., and D'Artel-Ellis, J. (1984) *J. Biol. Chem.* **259**, 13878–13884
398. Bell, G. I., Burant, C. F., Takeda, J., and Gould, G. W. (1993) *J. Biol. Chem.* **268**, 19161–19164
399. Silverman, M. (1991) *Ann. Rev. Biochem.* **60**, 757–794
400. Mueckler, M. (1994) *Eur. J. Biochem.* **219**, 713–725
401. Maher, F., Vannucci, S. J., and Simpson, I. A. (1994) *FASEB J.* **8**, 1003–1011
402. Burant, C. F., Takeda, J., Brot-Laroche, E., Bell, G. I., and Davidson, N. O. (1992) *J. Biol. Chem.* **267**, 14523–14526
403. Zamora-León, S. P., Golde, D. W., Concha, I. I., Rivas, C. I., Delgado-López, F., Baselga, J., Nualart, F., and Vera, J. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1847–1852
404. Coderre, L., Kandror, K. V., Vallega, G., and Pilch, P. F. (1995) *J. Biol. Chem.* **270**, 27584–27588
405. Fischer, Y., Thomas, J., Sevilla, L., Munoz, P., Becker, C., Holman, G., Kozka, I. J., Palacín, M., Testar, X., Kammermeier, H., and Zorzano, A. (1997) *J. Biol. Chem.* **272**, 7085–7092

405a. Shepherd, P. R., and Kahn, B. B. (1999) *N. Engl. J. Med.* **341**, 248–257

406. Mueckler, M., and Makepeace, C. (1997) *J. Biol. Chem.* **272**, 30141–30146
407. Cloherty, E. K., Diamond, D. L., Heard, K. S., and Carruthers, A. (1996) *Biochemistry* **35**, 13231–13239
408. Seatter, M. J., De La Rue, S. A., Porter, L. M., and Gould, G. W. (1998) *Biochemistry* **37**, 1322–1326
409. Rumsey, S. C., Kwon, O., Xu, G. W., Burant, C. F., Simpson, I., and Levine, M. (1997) *J. Biol. Chem.* **272**, 18982–18989
410. Wiese, T. J., Dunlap, J. A., and Yorek, M. A. (1994) *J. Biol. Chem.* **269**, 22705–22711
411. Kasahara, M., Shimoda, E., and Maeda, M. (1997) *J. Biol. Chem.* **272**, 16721–16724
412. Turk, E., Kerner, C. J., Lostao, M. P., and Wright, E. M. (1996) *J. Biol. Chem.* **271**, 1925–1934
413. Chen, X.-Z., Coady, M. J., Jackson, F., Berteloot, A., and Lapointe, J.-Y. (1995) *Biophys. J.* **69**, 2405–2414
414. Panayotova-Heiermann, M., Eskandari, S., Turk, E., Zampighi, G. A., and Wright, E. M. (1997) *J. Biol. Chem.* **272**, 20324–20327
415. Prasad, P. D., Wang, H., Kekuda, R., Fujita, T., Fei, Y.-J., Devoe, L. D., Leibach, F. H., and Ganapathy, V. (1998) *J. Biol. Chem.* **273**, 7501–7506
416. Malandro, M. S., and Kilberg, M. S. (1996) *Ann. Rev. Biochem.* **65**, 305–336
417. Utsunomiya-Tate, N., Endou, H., and Kanai, Y. (1996) *J. Biol. Chem.* **271**, 14883–14890
418. Kong, C.-T., Yet, S.-F., and Lever, J. E. (1993) *J. Biol. Chem.* **268**, 1509–1512
419. Su, T.-Z., Wang, M., Syu, L.-J., Saltiel, A. R., and Oxender, D. L. (1998) *J. Biol. Chem.* **273**, 3173–3179
420. Yao, D., Mackenzie, B., Ming, H., Varoqui, H., Zhu, H., Hediger, M. A., and Erickson, J. D. (2000) *J. Biol. Chem.* **275**, 22790–22797

420a. Reimer, R. J., Chaudhry, F. A., Gray, A. T., and Edwards, R. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7715–7720

421. Mosckovitz, R., Udenfriend, S., Felix, A., Heimer, E., and Tate, S. S. (1994) *FASEB J.* **8**, 1069–1074
422. Chillarón, J., Estévez, R., Mora, C., Wagner, C. A., Suessbrich, H., Lang, F., Gelpí, J. L., Testar, X., Busch, A. E., Zorzano, A., and Palacín, M. (1996) *J. Biol. Chem.* **271**, 17761–17770
423. Trotti, D., Rossi, D., Gjesdal, O., Levy, L. M., Racagni, G., Danbolt, N. C., and Volterra, A. (1996) *J. Biol. Chem.* **271**, 5976–5979
424. Peghini, P., Janzen, J., and Stoffel, W. (1997) *EMBO J.* **16**, 3822–3832
425. Lebrun, B., Sakaitani, M., Shimamoto, K., Yasuda-Kamatani, Y., and Nakajima, T. (1997) *J. Biol. Chem.* **272**, 20336–20339
426. Ponce, J., Biton, B., Benavides, J., Avenet, P., and Aragón, C. (2000) *J. Biol. Chem.* **275**, 13856–13862
427. Pantanowitz, S., Bendahan, A., and Kanner, B. I. (1993) *J. Biol. Chem.* **268**, 3222–3225
428. Bennett, E. R., and Kanner, B. I. (1997) *J. Biol. Chem.* **272**, 1203–1210
429. Velaz-Faircloth, M., Guadano-Ferraz, A., Henzi, V. A., and Fremeau, R. T., Jr. (1995) *J. Biol. Chem.* **270**, 15755–15761
430. Chen, J.-G., Liu-Chen, S., and Rudnick, G. (1998) *J. Biol. Chem.* **273**, 12675–12681
431. Nikaido, H., and Saier, M. H., Jr. (1992) *Science* **258**, 936–942
432. Will, A., Grassl, R., Erdmenger, J., Caspari, T., and Tanner, W. (1998) *J. Biol. Chem.* **273**, 11456–11462
433. Kaback, H. R., Bibi, E., and Roepe, P. D. (1990) *Trends Biochem. Sci.* **15**, 309–314
434. Frillingos, S., Wu, J., Venkatesan, P., and Kaback, H. R. (1997) *Biochemistry* **36**, 6408–6414
435. Green, A. L., Anderson, E. J., and Brooker, R. J. (2000) *J. Biol. Chem.* **275**, 23240–23246
436. Zhao, M., Zen, K.-C., Hubbell, W. L., and Kaback, H. R. (1999) *Biochemistry* **38**, 7407–7412
437. Zani, M. L., Pourcher, T., and Leblanc, G. (1994) *J. Biol. Chem.* **269**, 24883–24889
438. Pourcher, T., Bibi, E., Kaback, H. R., and Leblanc, G. (1996) *Biochemistry* **35**, 4161–4168
439. McDonald, T. P., Walmsley, A. R., and Henderson, P. J. F. (1997) *J. Biol. Chem.* **272**, 15189–15199
440. Fei, Y.-J., Liu, W., Prasad, P. D., Kekuda, R., Oblak, T. G., Ganapathy, V., and Leibach, F. H. (1997) *Biochemistry* **36**, 452–460
441. Brandsch, M., Thunecke, F., Küllertz, G., Schutkowski, M., Fischer, G., and Neubert, K. (1998) *J. Biol. Chem.* **273**, 3861–3864
442. Boorer, K. J., Frommer, W. B., Bush, D. R., Kreman, M., Loo, D. D. F., and Wright, E. M. (1996) *J. Biol. Chem.* **271**, 2213–2220
443. Boorer, K. J., and Fischer, W.-N. (1997) *J. Biol. Chem.* **272**, 13040–13046
444. Someya, Y., Niwa, A., Sawai, T., and Yamaguchi, A. (1995) *Biochemistry* **34**, 7–12
445. Nikaido, K., Liu, P.-Q., and Ames, G. F.-L. (1997) *J. Biol. Chem.* **272**, 27745–27752
446. Létoffé, S., Delepelaire, P., and Wandersman, C. (1996) *EMBO J.* **15**, 5804–5811
447. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277**, 1453–1462
448. Decottignies, A., Lambert, L., Catty, P., Degand, H., Epping, E. A., Moye-Rowley, W. S., Balzi, E., and Goffeau, A. (1995) *J. Biol. Chem.* **270**, 18150–18157

448a. Driessen, A. J. M., Rosen, B. P., and Konings, W. N. (2000) *Trends Biochem. Sci.* **25**, 397–401

449. Sharom, F. J., DiDiodato, G., Yu, X., and Ashbourne, K. J. D. (1995) *J. Biol. Chem.* **270**, 10334–10341
450. Ko, Y. H., and Pedersen, P. L. (1995) *J. Biol. Chem.* **270**, 22093–22096
451. Nossal, N. G., and Heppel, L. A. (1966) *J. Biol. Chem.* **241**, 3055–3062
452. Ames, G. F.-L. (1986) *Ann. Rev. Biochem.* **55**, 397–425
453. Ames, G. F.-L., and Higgins, C. F. (1983) *Trends Biochem. Sci.* **8**, 97–100
454. Yao, N., Ledvina, P. S., Choudhary, A., and Quiocho, F. A. (1996) *Biochemistry* **35**, 2079–2085
455. Quiocho, F. A., and Vyas, N. K. (1984) *Nature (London)* **310**, 381–386
456. Gilliland, G. L., and Quiocho, F. A. (1981) *J. Mol. Biol.* **146**, 341–362
457. Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1988) *Science* **242**, 1290–1295
458. Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1991) *J. Biol. Chem.* **266**, 5226–5237
459. Shilton, B. H., Shuman, H. A., and Mowbray, S. L. (1996) *J. Mol. Biol.* **264**, 364–376
460. Oh, B.-H., Kang, C.-H., De Bondt, H., Kim, S.-H., Nikaido, K., Joshi, A. K., and Ames, G. F.-L. (1994) *J. Biol. Chem.* **269**, 4135–4143
461. Oh, B.-H., Ames, G. F.-L., and Kim, S.-H. (1994) *J. Biol. Chem.* **269**, 26323–26330
462. Sugiyama, S., Vassylyev, D. G., Matsushima, M., Kashiwagi, K., Igarashi, K., and Morikawa, K. (1996) *J. Biol. Chem.* **271**, 9519–9525
463. Pflugrath, J. W., and Quiocho, F. A. (1988) *J. Mol. Biol.* **200**, 163–180
464. Manson, M. D., Boos, W., Bassford, P. J., Jr., and Rasmussen, B. A. (1985) *J. Biol. Chem.* **260**, 9727–9733
465. Reizer, J., Michotey, V., Reizer, A., and Saier, M. H., Jr. (1994) *Protein Sci.* **3**, 440–450
466. Garrett, D. S., Seok, Y.-J., Liao, D.-I., Peterkofsky, A., Gronenborn, A. M., and Clore, G. M. (1997) *Biochemistry* **36**, 2517–2530
467. Chauvin, F., Brand, L., and Roseman, S. (1994) *J. Biol. Chem.* **269**, 20263–20269

467a. Zhu, P.-P., Szczepanowski, R. H., Nosworthy, N. J., Ginsburg, A., and Peterkofsky, A. (1999) *Biochemistry* **38**, 15470–15479

468. Jones, B. E., Rajagopal, P., and Klevit, R. E. (1997) *Protein Sci.* **6**, 2107–2119
469. El-Kabbani, O. A. L., Waygood, E. B., and Delbaera, L. T. J. (1987) *J. Biol. Chem.* **262**, 12926–12929

References

470. Reizer, J., Hoischen, C., Reizer, A., Pham, T. N., and Saier, M. H., Jr. (1993) *Protein Sci.* **2**, 506–521

471. Eberstadt, M., Grdadolnik, S. G., Gemmecker, G., Kessler, H., Buhr, A., and Erni, B. (1996) *Biochemistry* **35**, 11286–11292

472. Van Dijk, A. A., Scheek, R. M., Dijkstra, K., Wolters, G. K., and Robillard, G. T. (1992) *Biochemistry* **31**, 9063–9072

473. Gemmecker, G., Eberstadt, M., Buhr, A., Lanz, R., Golic Grdadolnik, S., Kessler, H., and Erni, B. (1997) *Biochemistry* **36**, 7408–7417

474. Pelton, J. G., Torchia, D. A., Remington, S. J., Murphy, K. P., Meadow, N. D., and Roseman, S. (1996) *J. Biol. Chem.* **271**, 33446–33456

475. Lanz, R., and Erni, B. (1998) *J. Biol. Chem.* **273**, 12239–12243

476. Nunn, R. S., Markovic-Housley, Z., Génovésio-Taverne, J.-C., Flükiger, K., Rizkallah, P. J., Jansonius, J. N., Schirmer, T., and Erni, B. (1996) *J. Mol. Biol.* **259**, 502–511

477. Roseman, S., and Meadow, N. D. (1990) *J. Biol. Chem.* **265**, 2993–2996

478. Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W., and Remington, S. J. (1993) *Science* **259**, 673–677

479. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, Freeman, San Francisco, California

480. Masuda, T., Dobson, G. P., and Veech, R. L. (1990) *J. Biol. Chem.* **265**, 20321–20334

481. Catterall, W. A. (1995) *Ann. Rev. Biochem.* **64**, 493–531

482. Valverde, M. A., Hardy, S. P., and Sepúlveda, F. V. (1995) *FASEB J.* **9**, 509–515

483. Schmidt-Rose, T., and Jentsch, T. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7633–7638

484. Jentsch, T. J., Steinmeyer, K., and Schwarz, G. (1990) *Nature (London)* **348**, 510–514

485. Middleton, R. E., Pheasant, D. J., and Miller, C. (1994) *Biochemistry* **33**, 13189–13198

486. Schwappach, B., Stobrawa, S., Hechenberger, M., Steinmeyer, K., and Jentsch, T. J. (1998) *J. Biol. Chem.* **273**, 15110–15118

487. Gronemeier, M., Condie, A., Prosser, J., Steinmeyer, K., Jentsch, T. J., and Jockusch, H. (1994) *J. Biol. Chem.* **269**, 5963–5967

488. Pucéat, M., Korichneva, I., Cassoly, R., and Vassort, G. (1995) *J. Biol. Chem.* **270**, 1315–1322

489. Jarolim, P., Shayakul, C., Prabakaran, D., Jiang, L., Stuart-Tilley, A., Rubin, H. L., Simova, S., Zavadil, J., Herrin, J. T., Brouillette, J., Somers, M. J. G., Seemanova, E., Brugnara, C., Guay-Woodford, L. M., and Alper, S. L. (1998) *J. Biol. Chem.* **273**, 6380–6388

489a. Fujinaga, J., Tang, X.-B., and Casey, J. R. (1999) *J. Biol. Chem.* **274**, 6626–6633

490. Welsh, M. J., and Smith, A. E. (1995) *Sci. Am.* **273**(Dec), 52–59

491. Seibert, F. S., Linsdell, P., Loo, T. W., Hanrahan, J. W., Riordan, J. R., and Clarke, D. M. (1996) *J. Biol. Chem.* **271**, 27493–27499

492. Romero, M. F., Hediger, M. A., Boulpaep, E. L., and Boron, W. F. (1997) *Nature (London)* **387**, 409–413

493. Burnham, C. E., Amlal, H., Wang, Z., Shull, G. E., and Soleimani, M. (1997) *J. Biol. Chem.* **272**, 19111–19114

494. Muchhal, U. S., Pardo, J. M., and Raghothama, K. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10519–10523

495. Smith, F. W., Ealing, P. M., Hawkesford, M. J., and Clarkson, D. T. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9373–9377

496. Zhou, J.-J., Theodoulou, F. L., Muldin, I., Ingemarsson, B., and Miller, A. J. (1998) *J. Biol. Chem.* **273**, 12017–12023

497. Schroers, A., Burkovski, A., Wohlrab, H., and Krämer, R. (1998) *J. Biol. Chem.* **273**, 14269–14276

498. Phelps, A., Briggs, C., Mincone, L., and Wohlrab, H. (1996) *Biochemistry* **35**, 10757–10762

499. Collins, J. F., and Ghishan, F. K. (1996) *FASEB J.* **10**, 751–759

500. Garcia, C. K., Brown, M. S., Pathak, R. K., and Goldstein, J. L. (1995) *J. Biol. Chem.* **270**, 1843–1849

501. Kuo, S.-M., and Aronson, P. S. (1996) *J. Biol. Chem.* **271**, 15491–15497

502. Yellen, G. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 227–246

503. Miller, C. (1991) *Science* **252**, 1092–1096

504. Philipson, L. H., and Steiner, D. F. (1995) *Science* **268**, 372–373

505. Inagaki, N., Gonoi, T., Clement, J. P., IV, Namba, N., Inazawa, J., Gonzalez, G., Aguilar-Bryan, L., Seino, S., and Bryan, J. (1995) *Science* **270**, 1166–1170

506. Bränström, R., Efendic, S., Berggren, P.-O., and Larsson, O. (1998) *J. Biol. Chem.* **273**, 14113–14118

507. Ruknudin, A., Schulze, D. H., Sullivan, S. K., Lederer, W. J., and Welling, P. A. (1998) *J. Biol. Chem.* **273**, 14165–14171

508. Day, M. L., Johnson, M. H., and Cook, D. I. (1998) *EMBO J.* **17**, 1952–1960

509. McCormick, K. A., Isom, L. L., Ragsdale, D., Smith, D., Scheuer, T., and Catterall, W. A. (1998) *J. Biol. Chem.* **273**, 3954–3962

510. Favre, I., Moczydlowski, E., and Schild, L. (1996) *Biophys. J.* **71**, 3110–3125

510a. Lipkind, G. M., and Fozzard, H. A. (2000) *Biochemistry* **39**, 8161–8170

511. Pérez-García, M. T., Chiamvimonvat, N., Marban, E., and Tomaselli, G. F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 300–304

512. Kosari, F., Sheng, S., Li, J., Mak, D.-O. D., Foskett, J. K., and Kleyman, T. R. (1998) *J. Biol. Chem.* **273**, 13469–13474

513. Garty, H. (1994) *FASEB J.* **8**, 522–528

514. Miller, R. J. (1992) *J. Biol. Chem.* **267**, 1403–1406

515. Mitterdorfer, J., Sinnegger, M. J., Grabner, M., Striessnig, J., and Glossmann, H. (1995) *Biochemistry* **34**, 9350–9355

516. McPherson, P. S., and Campbell, K. P. (1993) *J. Biol. Chem.* **268**, 13765–13768

517. Jeyakumar, L. H., Copello, J. A., O'Malley, A. M., Wu, G.-M., Grassucci, R., Wagenknecht, T., and Fleischer, S. (1998) *J. Biol. Chem.* **273**, 16011–16020

518. Kraus, R. L., Sinnegger, M. J., Glossmann, H., Hering, S., and Striessnig, J. (1998) *J. Biol. Chem.* **273**, 5586–5590

519. Kim, S., Lakhani, V., Costa, D. J., Sharara, A. I., Fitz, J. G., Huang, L.-W., Peters, K. G., and Kindman, L. A. (1995) *J. Biol. Chem.* **270**, 5266–5269

520. Mezna, M., and Michelangeli, F. (1996) *J. Biol. Chem.* **271**, 31818–31823

521. Nicoll, D. A., Hryshko, L. V., Matsuoka, S., Frank, J. S., and Philipson, K. D. (1996) *J. Biol. Chem.* **271**, 13385–13391

522. Crumpton, M. J., and Dedman, J. R. (1990) *Nature (London)* **345**, 212

523. Swairjo, M. A., and Seaton, B. A. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 193–213

524. Benz, J., Bergner, A., Hofmann, A., Demange, P., Göttig, P., Liemann, S., Huber, R., and Voges, D. (1996) *J. Mol. Biol.* **260**, 638–643

525. Burger, A., Berendes, R., Liemann, S., Benz, J., Hofmann, A., Göttig, P., Huber, R., Gerke, V., Thiel, C., Römisch, J., and Weber, K. (1996) *J. Mol. Biol.* **257**, 839–847

526. Demange, P., Voges, D., Benz, J., Liemann, S., Göttig, P., Berendes, R., Burger, A., and Huber, R. (1994) *Trends Biochem. Sci.* **19**, 272–276

527. Campos, B., Mo, Y. D., Mealy, T. R., Li, C. W., Swairjo, M. A., Balch, C., Head, J. F., Retzinger, G., Dedman, J. R., and Seaton, B. A. (1998) *Biochemistry* **37**, 8004–8010

528. Ling, G. N. (1984) *In Search of the Physical Basis of Life*, Plenum, New York

529. Baker, P. F., and Willis, J. S. (1970) *Nature (London)* **226**, 521–523

530. Kyte, J. (1972) *J. Biol. Chem.* **247**, 7642–7649

531. Lingrel, J. B., and Kuntzweiler, T. (1994) *J. Biol. Chem.* **269**, 19659–19662

532. Lutsenko, S., Daoud, S., and Kaplan, J. H. (1997) *J. Biol. Chem.* **272**, 5249–5255

533. Scheiner-Bobis, G., and Schreiber, S. (1999) *Biochemistry* **38**, 9198–9208

534. Gatto, C., Wang, A. X., and Kaplan, J. H. (1998) *J. Biol. Chem.* **273**, 10578–10585

535. Rossier, B. C., Geering, K., and Kraehenbuhl, J. P. (1987) *Trends Biochem. Sci.* **12**, 483–487

536. Minor, N. T., Sha, Q., Nichols, C. G., and Mercer, R. W. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6521–6525

537. Lutsenko, S., and Kaplan, J. H. (1995) *Biochemistry* **34**, 15607–15613

538. Nelson, N., and Taiz, L. (1989) *Trends Biochem. Sci.* **14**, 113–116

539. Solioz, M., and Vulpe, C. (1996) *Trends Biochem. Sci.* **21**, 237–241

539a. Ueno, S., Kaieda, N., Koyama, N. (2000) *J. Biol. Chem.* **275**, 14537–14540

540. Carafoli, E. (1994) *FASEB J.* **8**, 993–1002

541. MacLennan, D. H., Rice, W. J., and Green, N. M. (1997) *J. Biol. Chem.* **272**, 28815–28818

542. Myung, J., and Jencks, W. P. (1995) *Biochemistry* **34**, 3077–3083

543. Dode, L., De Greef, C., Mountian, I., Attard, M., Town, M. M., Casteels, R., and Wuytack, F. (1998) *J. Biol. Chem.* **273**, 13982–13994

543a. Menguy, T., Corre, F., Bouneau, L., Deschamps, S., Vuust Moller, J., Champeil, P., le Maire, M., and Falson, P. (1998) *J. Biol. Chem.* **273**, 20134–20143

544. Melle-Milovanovic, D., Milovanovic, M., Nagpal, S., Sachs, G., and Shin, J. M. (1998) *J. Biol. Chem.* **273**, 11075–11081

545. Rulli, S. J., Horiba, M. N., Skripnikova, E., and Rabon, E. C. (1999) *J. Biol. Chem.* **274**, 15245–15250

546. Asano, S., Inoie, M., and Takeguchi, N. (1987) *J. Biol. Chem.* **262**, 13263–13268

547. Seto-Young, D., Hall, M. J., Na, S., Haber, J. E., and Perlin, D. S. (1996) *J. Biol. Chem.* **271**, 581–587

548. Auer, M., Scarborough, G. A., and Kühlbrandt, W. (1998) *Nature (London)* **392**, 840–843 549. Orlowski, J., and Grinstein, S. (1997) *J. Biol. Chem.* **272**, 22373–22376

549. Orlowski, J. and Grinstein, S. (1997) *J. Biol. Chem.* **272**, 22373–22376

549a. Wakabayashi, S., Pang, T., Su, X., and Shigekawa, M. (2000) *J. Biol. Chem.* **275**, 7942–7949

549b. Williams, K. A. (2000) *Nature (London)* **403**, 112–115

550. Kasho, V. N., Stengelin, M., Smirnova, I. N., and Faller, L. D. (1997) *Biochemistry* **36**, 8045–8052

551. Campos, M., and Beaugé, L. (1997) *Biochemistry* **36**, 14228–14237

552. Vilsen, B. (1997) *Biochemistry* **36**, 13312–13324

552a. Boxenbaum, N., Daly, S. E., Javaid, Z. Z., Lane, L. K., and Blostein, R. (1998) *J. Biol. Chem.* **273**, 23086–23092

553. Zhang, P., Toyoshima, C., Yonekura, K., Green, N. M., and Stokes, D. L. (1998) *Nature (London)* **392**, 835–839

553a. Toyoshima, C., Nakasako, M., Nomura, H., and Ogawa, H. (2000) *Nature (London)* **405**, 647–655

References

553b. MacLennan, D. H., and Green, N. M. (2000) *Nature (London)* **405**, 633–634

553c. Zhang, Z., Lewis, D., Strock, C., and Inesi, G. (2000) *Biochemistry* **39**, 8758–8767

554. Crider, B. P., Andersen, P., White, A. E., Zhou, Z., Li, X., Mattsson, J. P., Lundberg, L., Keeling, D. J., Xie, X.-S., Stone, D. K., and Peng, S.-B. (1997) *J. Biol. Chem.* **272**, 10721–10728

555. Nanda, A., Brumell, J. H., Nordström, T., Kjeldsen, L., Sengelov, H., Borregaard, N., Rotstein, O. D., and Grinstein, S. (1996) *J. Biol. Chem.* **271**, 15963–15970

555a. Kelly, R. B. (1999) *Trends Biochem. Sci.* **24**, M29–M33

556. Desjardins, M., Celis, J. E., van Meer, G., Dieplinger, H., Jahraus, A., Griffiths, G., and Huber, L. A. (1994) *J. Biol. Chem.* **269**, 32194–32200

557. Peng, S.-B., Crider, B. P., Tsai, S. J., Xie, X.-S., and Stone, D. K. (1996) *J. Biol. Chem.* **271**, 3324–3327

558. Pearse, B. M. F., and Crowther, R. A. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 49–68

559. Schmid, S. L., and Damke, H. (1995) *FASEB J.* **9**, 1445–1453

559a. Marsh, M., and McMahon, H. T. (1999) *Science* **285**, 215–220

559b. Heuser, J. (1981) *Trends Biochem. Sci.* **6**, 64–68

560. Pearse, B. M. F., and Bretscher, M. S. (1981) *Ann. Rev. Biochem.* **50**, 85–101

561. Pishvaee, B., Munn, A., and Payne, G. (1997) *EMBO J.* **16**, 2227–2239

562. Rothman, J. E. (1996) *Protein Sci.* **5**, 185–194

563. Brodsky, F. M., Hill, B. L., Acton, S. L., Näthke, I., Wong, D. H., Ponnambalam, S., and Parham, P. (1991) *Trends Biochem. Sci.* **16**, 208–213

564. Lemmon, S. K., and Jones, E. W. (1987) *Science* **238**, 504–509

564a. Anderson, R. G. W. (1998) *Ann. Rev. Biochem.* **67**, 199–225

564b. Shin, J.-S., Gao, Z., and Abraham, S. N. (2000) *Science* **289**, 785–788

564c. Mulvey, M. A., and Hultgren, S. J. (2000) *Science* **289**, 732–733

565. Lowe, M., and Kreis, T. E. (1995) *J. Biol. Chem.* **270**, 31364–31371

566. Schekman, R., and Orci, L. (1996) *Science* **271**, 1526–1533

566a. Eugster, A., Frigerio, G., Dale, M., and Duden, R. (2000) *EMBO J.* **19**, 3905–3917

567. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R., and Klug, A. (1978) *Nature (London)* **276**, 362–368

567a. Rapoport, I., Chen, Y. C., Cupers, P., Shoelson, S. E., and Kirchhausen, T. (1998) *EMBO J.* **17**, 2148–2155

567b. Huang, K. M., D'Hondt, K., Riezman, H., and Lemmon, S. K. (1999) *EMBO J.* **18**, 3897–3908

567c. Owen, D. J., Vallis, Y., Pearse, B. M. F., McMahon, H. T., and Evans, P. R. (2000) *EMBO J.* **19**, 4216–4227

567d. Wendland, B., Steece, K. E., and Emr, S. D. (1999) *EMBO J.* **18**, 4383–4393

567e. Scales, S. J., and Scheller, R. H. (1999) *Nature (London)* **401**, 123–124

568. Ungewickell, E., Ungewickell, H., Holstein, S. E. H., Lindner, R., Prasad, K., Barouch, W., Martin, B., Greene, L. E., and Eisenberg, E. (1995) *Nature (London)* **378**, 632–635

569. Mellman, I., Fuchs, R., and Helenius, A. (1986) *Ann. Rev. Biochem.* **55**, 663–700

570. Brown, M. S., and Goldstein, J. L. (1974) *Sci. Am.* **251**(Nov), 58–66

571. Brown, M. S., and Goldstein, J. L. (1986) *Science* **232**, 34–47

572. Holstein, T. W., Benoit, M., Herder, G., Wanner, G., David, C. N., and Gaub, H. E. (1994) *Science* **265**, 402–404

573. Ferro-Novick, S., and Jahn, R. (1994) *Nature (London)* **370**, 191–193

574. Mayorga, L. S., Berón, W., Sarrouf, M. N., Colombo, M. I., Creutz, C., and Stahl, P. D. (1994) *J. Biol. Chem.* **269**, 30927–30934

575. Burgoyne, R. D. (1988) *Nature (London)* **331**, 20

576. Wilson, D. W., Wilcox, C. A., Flynn, G. C., Chen, E., Kuang, W.-J., Henzel, W. J., Block, M. R., Ullrich, A., and Rothman, J. E. (1989) *Nature (London)* **339**, 355–359

577. Colombo, M. I., Taddese, M., Whiteheart, S. W., and Stahl, P. D. (1996) *J. Biol. Chem.* **271**, 18810–18816

578. Goda, Y. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 769–772

579. Lugtenberg, B. (1981) *Trends Biochem. Sci.* **6**, 262–266

580. Graham, L. L., Beveridge, T. J., and Nanninga, N. (1991) *Trends Biochem. Sci.* **16**, 328–329

581. Oliver, D. B. (1987) in Escherichia coli *and* Salmonella typhi-murium (Neidhardt, F. C., ed), pp. 56–70, Am. Soc. for Micro-biology, Washington, DC

582. Glauner, B., Höltje, J.-V., and Schwarz, U. (1988) *J. Biol. Chem.* **263**, 10088–10095

583. Park, J. T. (1987) in Escherichia coli *and* Salmonella typhimurium (Neidhardt, F. C., ed), pp. 23–32, Am. Soc. for Microbiology, Washington, DC

584. Scherrer, R. (1984) *Trends Biochem. Sci.* **9**, 242–245

585. Nikaido, H., and Vaara, M. (1987) in Escherichia coli *and* Salmonella typhimurium (Neidhardt, F. C., ed), pp. 7–22, Am. Soc. for Microbiology, Washington, DC

586. Hancock, R. E. W., Karunaratne, D. N., and Bernegger-Egli, C. (1994) in *Bacterial Cell Wall (New Comprehensive Biochemistry)*, Vol. 27 (Ghuysen, J.-M., and Hakenbeck, R., eds), pp. 263–279, Elsevier, Amsterdam

587. Raetz, C. R. H., and Dowhan, W. (1990) *J. Biol. Chem.* **265**, 1235–1238

588. Osborn, M. J. (1969) *Ann. Rev. Biochem.* **38**, 501–538

589. Choi, D.-S., Yamada, H., Mizuno, T., and Mizushima, S. (1986) *J. Biol. Chem.* **261**, 8953–8957

590. Diedrich, D. L., and Schnaitman, C. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3708–3712

591. Raetz, C. R. H., Ulevitch, R. J., Wright, S. D., Sibley, C. H., Ding, A., and Nathan, C. F. (1991) *FASEB J.* **5**, 2652–2660

592. Hollingsworth, R. I., and Carlson, R. W. (1989) *J. Biol. Chem.* **264**, 9300–9303

593. Ogawa, T. (1994) *Eur. J. Biochem.* **219**, 737–742

594. Rietschel, E. T., Kirikae, T., Schade, F. U., Mamat, U., Schmidt, G., Loppnow, H., Ulmer, A. J., Zähringer, U., Seydel, U., Di Padova, F., Schreier, M., and Brade, H. (1994) *FASEB J.* **8**, 217–225

595. Masoud, H., Moxon, E. R., Martin, A., Krajcarski, D., and Richards, J. C. (1997) *Biochemistry* **36**, 2091–2103

596. Aspinall, G. O., and Monteiro, M. A. (1996) *Biochemistry* **35**, 2498–2504

597. Rietschel, E. T., and Brade, H. (1992) *Sci. Am.* **267**(Aug), 54–61

598. Costerton, J. W., Irwin, R. T., and Cheng, K. J. (1981) *Crit. Rev. Microbiol* **8**, 303–338

599. Fischer, W. (1994) in *Bacterial Cell Wall (New Comprehensive Biochemistry)*, Vol. 27 (Ghuysen, J.-M., and Hakenbeck, R., eds), pp. 199–215, Elsevier, Amsterdam

600. Ganfield, M.-C. W., and Pieringer, R. A. (1980) *J. Biol. Chem.* **255**, 5164–5169

601. Johnson, S. D., Lacher, K. P., and Anderson, J. S. (1981) *Biochemistry* **20**, 4781–4785

602. Ivatt, R. J., and Gilvarg, C. (1979) *J. Biol. Chem.* **254**, 2759–2765

603. Venisse, A., Rivière, M., Vercauteren, J., and Puzo, G. (1995) *J. Biol. Chem.* **270**, 15012–15021

604. Brennan, P. J., and Nikaido, H. (1995) *Ann. Rev. Biochem.* **64**, 29–63

605. Liu, J., Barry, C. E., III, Besra, G. S., and Nikaido, H. (1996) *J. Biol. Chem.* **271**, 29545–29551

606. Mikusová, K., Mikus, M., Besra, G. S., Hancock, I., and Brennan, P. J. (1996) *J. Biol. Chem.* **271**, 7820–7828

607. Lopez-Marin, L. M., Quesada, D., Lakhdar-Ghazal, F., Tocanne, J.-F., and Lanéelle, G. (1994) *Biochemistry* **33**, 7056–7061

608. Sprott, G. D., Ekiel, I., and Dicaire, C. (1990) *J. Biol. Chem.* **265**, 13735–13740

609. Peters, J., Nitsch, M., Kühlmorgen, B., Golbik, R., Lupas, A., Kellermann, J., Engelhardt, H., Pfander, J.-P., Müller, S., Goldie, K., Engel, A., Stetter, K.-O., and Baumeister, W. (1995) *J. Mol. Biol.* **245**, 385–401

610. Thomas, S. R., and Trust, T. J. (1995) *J. Mol. Biol.* **245**, 568–581

611. Peters, J., Baumeister, W., and Lupas, A. (1996) *J. Mol. Biol.* **257**, 1031–1041

612. Fischetti, V. A. (1991) *Sci. Am.* **264**(Jun), 58–65

613. Goward, C. R., Scawen, M. D., Murphy, J. P., and Atkinson, T. (1993) *Trends Biochem. Sci.* **18**, 136–139

614. Schneewind, O., Fowler, A., and Faull, K. F. (1995) *Science* **268**, 103–106

615. Knörle, R., and Hübner, W. (1995) *Biochemistry* **34**, 10970–10975

616. Shaper, N. L., Hollis, G. F., Douglas, J. G., Kirsch, I. R., and Shaper, J. H. (1988) *J. Biol. Chem.* **263**, 10420–10428

617. Lin, C. Q., and Bissell, M. J. (1993) *FASEB J.* **7**, 737–743

618. Har-El, R., and Tanzer, M. L. (1993) *FASEB J.* **7**, 1115–1123

619. Mayne, R., and Burgeson, R. E., eds. (1987) *Structure and Function of Collagen Types*, Academic Press, New York, Orlando, San Diego

620. Nimni, M. E., ed. (1988) *Collagen: Biochemistry, Biomechanics, Biotechnology*, CRC Press, Boca Raton, Florida

621. Ramachandran, G. N. (1988) *Intl. J. Peptide Prot. Res.* **31**, 1–16

622. Martin, G. R., Timpl, R., Müller, P. K., and Kühn, K. (1985) *Trends Biochem. Sci.* **10**, 285–287

623. Prockop, D. J., and Kivirikko, K. I. (1995) *Ann. Rev. Biochem.* **64**, 403–434

624. van der Rest, M., and Garrone, R. (1991) *FASEB J.* **5**, 2814–2823

625. Byers, P. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4029–4077, McGraw-Hill, New York

626. Brodsky, B., and Shah, N. K. (1995) *FASEB J.* **9**, 1537–1546

627. Koivu, J., and Myllylä, R. (1987) *J. Biol. Chem.* **262**, 6159–6164

628. Prockop, D. J. (1990) *J. Biol. Chem.* **265**, 15349–15352

629. Annunen, P., Autio-Harmainen, H., and Kivirikko, K. I. (1998) *J. Biol. Chem.* **273**, 5989–5992

630. Reiser, K., McCormick, R. J., and Rucker, R. B. (1992) *FASEB J.* **6**, 2439–2449

631. Li, S.-W., Sieron, A. L., Fertala, A., Hojima, Y., Arnold, W. V., and Prockop, D. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5127–5130

632. Meek, K. M., Chapman, J. A., and Hardcastle, R. A. (1979) *J. Biol. Chem.* **254**, 10710–10714

633. Chan, V. C., Ramshaw, J. A. M., Kirkpatrick, A., Beck, K., and Brodsky, B. (1997) *J. Biol. Chem.* **272**, 31441–31446

References

634. Wess, T. J., Hammersley, A. P., Wess, L., and Miller, A. (1998) *J. Mol. Biol.* **275**, 255–267
634a. Kramer, R. Z., Venugopal, M. G., Bella, J., Mayville, P., Brodsky, B., and Berman, H. M. (2000) *J. Mol. Biol.* **301**, 1191–1205
635. Eyre, D. R., Paz, M. A., and Gallop, P. M. (1984) *Ann. Rev. Biochem.* **53**, 717–748
636. Miller, A. (1982) *Trends Biochem. Sci.* **7**, 13–18
637. Eyre, D. R. (1980) *Science* **207**, 1315–1322
638. Wu, J.-J., and Eyre, D. R. (1984) *Biochemistry* **23**, 1850–1857
639. Hanson, D. A., and Eyre, D. R. (1996) *J. Biol. Chem.* **271**, 26508–26516
640. Kiss, I., Deák, F., Holloway, R. G., Jr., Delius, H., Mebust, K. A., Frinberger, E., Argraves, W. W., Tsonis, P. A., Winterbottom, N., and Goetink, P. F. (1989) *J. Biol. Chem.* **264**, 8126–8134
641. Tanaka, S., Avigad, G., Eikenberry, E. F., and Brodsky, B. (1988) *J. Biol. Chem.* **263**, 17650–17657
642. Kleman, J.-P., Aeschlimann, D., Paulsson, M., and van der Rest, M. (1995) *Biochemistry* **34**, 13768–13775
643. Mazzorana, M., Gruffat, H., Sergeant, A., and van der Rest, M. (1993) *J. Biol. Chem.* **268**, 3029–3032
644. Martin, G. R., Timpl, R., Miller, P. K., and Kuhn, K. (1985) *Trends Biochem. Sci.* **10**, 285–287
645. Gunwar, S., Ballester, F., Noelken, M. E., Sado, Y., Ninomiya, Y., and Hudson, B. G. (1998) *J. Biol. Chem.* **273**, 8767–8775
646. van der Rest, M., and Mayne, R. (1988) *J. Biol. Chem.* **263**, 1615–1618
647. de Crombrugghe, B., and Pastan, I. (1982) *Trends Biochem. Sci.* **7**, 11–13
648. Myers, J. C., Yang, H., D'Ippolito, J. A., Presente, A., Miller, M. K., and Dion, A. S. (1994) *J. Biol. Chem.* **269**, 18549–18557
649. Lui, V. C. H., Ng, L. J., Sat, E. W. Y., Nicholls, J., and Cheah, K. S. E. (1996) *J. Biol. Chem.* **271**, 16945–16951
650. de Wet, W., Bernard, M., Benson-Chanda, V., Chu, M.-L., Dickson, L., Weil, D., and Ramirez, F. (1987) *J. Biol. Chem.* **262**, 16032–16036
651. Chu, M.-L., de Wet, W., Bernhard, M., and Ramirez, F. (1985) *J. Biol. Chem.* **260**, 2315–2370
652. Kramer, J. M. (1994) *FASEB J.* **8**, 329–336
653. Engel, J. (1997) *Science* **277**, 1785–1786
654. Rosenbloom, J., Abrams, W. R., and Mecham, R. (1993) *FASEB J.* **7**, 1208–1218
655. Hukins, D. W. L. (1984) *Connective Tissue Matrix*, MacMillan, New York
656. Fransson, L.-Å. (1987) *Trends Biochem. Sci.* **12**, 406–411
657. Burg, M. A., Tillet, E., Timpl, R., and Stallcup, W. B. (1996) *J. Biol. Chem.* **271**, 26110–26116
658. Caplan, A. I. (1984) *Sci. Am.* **251**(Oct), 84–94
659. Heinegård, D., and Oldberg, Å. (1989) *FASEB J.* **3**, 2042–2051
660. Cheng, F., Heinegård, D., Fransson, L.-Å., Bayliss, M., Bielicki, J., Hopwood, J., and Yoshida, K. (1996) *J. Biol. Chem.* **271**, 28572–28580
661. Hauser, N., and Paulsson, M. (1994) *J. Biol. Chem.* **269**, 25747–25753
662. Beck, K., Gambee, J. E., Bohan, C. A., and Bächinger, H. P. (1996) *J. Mol. Biol.* **256**, 909–923
663. Yurchenco, P. D., and Schittny, J. C. (1990) *FASEB J.* **4**, 1577–1590
663a. Yurchenco, P. D., Birk, D. E., and Mecham, R. P., eds. (1994) *Extracellular Matrix Assembly and Structure*, Academic Press, San Diego, California
664. Utani, A., Nomizu, M., Timpl, R., Roller, P. P., and Yamada, Y. (1994) *J. Biol. Chem.* **269**, 19167–19175
665. Pöschl, E., Mayer, U., Stetefeld, J., Baumgartner, R., Holak, T. A., Huber, R., and Timpl, R. (1996) *EMBO J.* **15**, 5154–5159
666. Durkin, M. E., Gautam, M., Loechel, F., Sanes, J. R., Merlie, J. P., Albrechtsen, R., and Wewer, U. M. (1996) *J. Biol. Chem.* **271**, 13407–13416
667. Beckmann, G., Hanke, J., Bork, P., and Reich, J. G. (1998) *J. Mol. Biol.* **275**, 725–730
668. Colognato-Pyke, H., O'Rear, J. J., Yamada, Y., Carbonetto, S., Cheng, Y.-S., and Yurchenco, P. D. (1995) *J. Biol. Chem.* **270**, 9398–9406
669. Kramer, J. M., Cox, G. N., and Hirsch, D. (1985) *J. Biol. Chem.* **260**, 1945–1951
670. Beck, K., Hunter, I., and Engel, J. (1990) *FASEB J.* **4**, 148–160
671. Sasaki, T., Kostka, G., Göhring, W., Wiedemann, H., Mann, K., Chu, M.-L., and Timpl, R. (1995) *J. Mol. Biol.* **245**, 241–250
672. Aeschlimann, D., and Paulsson, M. (1991) *J. Biol. Chem.* **266**, 15308–15317
673. Reinhardt, D. P., Ono, R. N., and Sakai, L. Y. (1997) *J. Biol. Chem.* **272**, 1231–1236
673a. Kettle, S., Yuan, X., Grundy, G., Knott, V., Downing, A. K., and Handford, P. A. (1999) *J. Mol. Biol.* **285**, 1277–1287
673b. Ritty, T. M., Broekelmann, T., Tisdale, C., Milewicz, D. M., and Mecham, R. P. (1999) *J. Biol. Chem.* **274**, 8933–8940
674. Reinhardt, D. P., Keene, D. R., Corson, G. M., Pöschl, E., Bächinger, H. P., Gambee, J. E., and Sakai, L. Y. (1996) *J. Mol. Biol.* **258**, 104–116
675. McKusick, V. A. (1991) *Nature (London)* **352**, 279–281
676. Francke, U., and Furthmayr, H. (1994) *N. Engl. J. Med.* **330**, 1384–1385
677. Mann, K., Mechling, D. E., Bächinger, H. P., Eckerskorn, C., Gaill, F., and Timpl, R. (1996) *J. Mol. Biol.* **261**, 255–266
678. Durkin, M. E., Carlin, B. E., Vergnes, J., Bartos, B., and Merlie, J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1570–1574
679. Richards, A. G. (1978) in *Biochemistry of Insects* (Rockstein, M., ed), Academic Press, New York (p. 205)
680. Schaefer, J., Kramer, K. J., Garbow, J. R., Jacob, G. S., Stejskal, E. O., Hopkins, T. L., and Speirs, R. D. (1987) *Science* **235**, 1200–1204
681. Simkiss, K., and Wilbur, K. M. (1989) *Biomineralization*, Academic Press, San Diego, California
682. DeOliveira, D. B., and Laursen, R. A. (1997) *J. Am. Chem. Soc.* **119**, 10627–10631
683. Weiner, S., and Addadi, L. (1991) *Trends Biochem. Sci.* **16**, 252–256
684. Albeck, S., Aizenberg, J., Addadi, L., and Weiner, S. (1993) *J. Am. Chem. Soc.* **115**, 11691–11697
685. Bonucci, E., ed. (1992) *Calcification in Biological Systems*, CRC Press, Boca Raton, Florida
686. Mann, S., Archibald, D. D., Didymus, J. M., Douglas, T., Heywood, B. R., Meldrum, F. C., and Reeves, N. J. (1993) *Science* **261**, 1286–1292
687. Aizenberg, J., Hanson, J., Ilan, M., Leiserowitz, L., Koetzle, T. F., Addadi, L., and Weiner, S. (1995) *FASEB J.* **9**, 262–268
688. Otsubo, K., Katz, E. P., Mechanic, G. L., and Yamauchi, M. (1992) *Biochemistry* **31**, 396–402
689. Fledelius, C., Johnsen, A. H., Cloos, P. A. C., Bonde, M., and Qvist, P. (1997) *J. Biol. Chem.* **272**, 9755–9763
689a. Ducy, P., Schinke, T., and Karsenty, G. (2000) *Science* **289**, 1501–1504
689b. Teitelbaum, S. L. (2000) *Science* **289**, 1504–1508
690. Weiner, S., and Traub, W. (1992) *FASEB J.* **6**, 879–885
691. McLean, F. M., Keller, P. J., Genge, B. R., Walters, S. A., and Wuthier, R. E. (1987) *J. Biol. Chem.* **262**, 10481–10488
692. Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., Shimizu, Y., Bronson, R. T., Gao, Y.-H., Inada, M., Sato, M., Okamoto, R., Kitamura, Y., Yoshiki, S., and Kishimoto, T. (1997) *Cell* **89**, 755–764
693. Ducy, P., Zhang, R., Geoffroy, V., Ridall, A. L., and Karsenty, G. (1997) *Cell* **89**, 747–754
694. Dickman, S. (1997) *Science* **276**, 1502
695. Rodan, G. A., and Harada, S.-i. (1997) *Cell* **89**, 677–680
696. Mundlos, S., Otto, F., Mundlos, C., Mulliken, J. B., Aylsworth, A. S., Albright, S., Lindhout, D., Cole, W. G., Henn, W., Knoll, J. H. M., Owen, M. J., Mertelsmann, R., Zabel, B. U., and Olsen, B. R. (1997) *Cell* **89**, 773–779
697. Price, P. A., and Williamson, M. K. (1985) *J. Biol. Chem.* **260**, 14971–14975
698. Nishimoto, S. K., Araki, N., Robinson, F. D., and Waite, J. H. (1992) *J. Biol. Chem.* **267**, 11600–11605
699. Price, P. A., Rice, J. S., and Williamson, M. K. (1994) *Protein Sci.* **3**, 822–830
700. Hohenester, E., Maurer, P., and Timpl, R. (1997) *EMBO J.* **16**, 3778–3786
701. Kelm, R. J., Jr., Swords, N. A., Orfeo, T., and Mann, K. G. (1994) *J. Biol. Chem.* **269**, 30147–30153
702. Fisher, L. W., Robey, P. G., Tuross, N., Otsuka, A. S., Tepen, D. A., Esch, F. S., Shimasaki, S., and Termine, J. D. (1987) *J. Biol. Chem.* **262**, 13457–13463
703. Denhardt, D. T., and Guo, X. (1993) *FASEB J.* **7**, 1475–1482
704. Shanmugam, V., Chackalaparampil, I., Kundu, G. C., Mukherjee, A. B., and Mukherjee, B. B. (1997) *Biochemistry* **36**, 5729–5738
704a. Rodan, G. A., and Martin, T. J. (2000) *Science* **289**, 1508–1514
704b. Service, R. F. (2000) *Science* **289**, 1498–1500
705. Zipkin, I., ed. (1973) *Biological Mineralization*, Wiley, New York
706. Schlesinger, P. H., Blair, H. C., Teitelbaum, S. L., and Edwards, J. C. (1997) *J. Biol. Chem.* **272**, 18636–18643
707. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) *Molecular Biology of the Cell*, 3rd ed., Garland, New York (p. 1182–1186)
708. Riggs, B. L., and Melton, L. J. (1992) *N. Engl. J. Med.* **327**, 620–627
709. Sly, W. S., and Hu, P. Y. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4113–4124, McGraw-Hill, New York
710. Roush, W. (1996) *Science* **273**, 1170
711. Delmas, P. D., and Meunier, P. J. (1997) *N. Engl. J. Med.* **336**, 558–566
711a. Ho, A. M., Johnson, M. D., and Kingsley, D. M. (2000) *Science* **289**, 265–270
711b. Hagmann, M. (2000) *Science* **289**, 225–226
712. Hohling, H. J., Barckhaus, R. H., and Kreftin, E. R. (1980) *Trends Biochem. Sci.* **5**, 8–11
713. Outka, D. E., and Williams, D. C. (1971) *Journal of Protozoology* **18**, 285–297
714. Inoue, S., and Okazaki, K. (1977) *Sci. Am.* **236**(Apr), 83–92
715. Volcani, B. E. (1981) in *Silicon and Silicious Structures in Biological Systems* (Simpson, T. L., and Volcani, B. E., eds), Springer-Verlag, New York (p. 157)
716. Mann, S. (1988) *Nature (London)* **332**, 119–124
717. Calvert, P. (1988) *Nature (London)* **334**, 651–652
718. Watanabe, N., and Wilbur, K. M., eds. (1976) *The Mechanisms of Mineralization in Invertebrates and Plants*, Univ. N. Carolina Press, Columbia, South Carolina

References

719. Lemann, J., Jr. (1993) *N. Engl. J. Med.* **328**, 880–881
720. Nakagawa, Y., Abram, V., Kézdy, F. J., Kaiser, E. T., and Coe, F. L. (1983) *J. Biol. Chem.* **258**, 12594–12600
721. Bertrand, J. A., Pignol, D., Bernard, J.-P., Verdier, J.-M., Dagorn, J.-C., and Fontecilla-Camps, J. C. (1996) *EMBO J.* **15**, 2678–2684
722. Bimmler, D., Graf, R., Scheele, G. A., and Frick, T. W. (1997) *J. Biol. Chem.* **272**, 3073–3082
723. De Reggi, M., Gharib, B., Patard, L., and Stoven, V. (1998) *J. Biol. Chem.* **273**, 4967–4971
724. Cheung, H. S., Kurup, I. V., Sallis, J. D., and Ryan, L. M. (1996) *J. Biol. Chem.* **271**, 28082–28085
725. Rothschild, B. M., and Thillaud, P. L. (1991) *Nature (London)* **349**, 288
726. Cabib, E., Roberts, R., and Bowers, B. (1982) *Ann. Rev. Biochem.* **51**, 763–793
727. Ballou, C. E., and Raschke, W. C. (1974) *Science* **184**, 127–134
728. Frevert, J., and Ballou, C. E. (1985) *Biochemistry* **24**, 753–759
729. Celerin, M., Ray, J. M., Schisler, N. J., Day, A. W., Stetler-Stevenson, W. G., and Laudenbach, D. E. (1996) *EMBO J.* **15**, 4445–4453

Study Questions

1. Compare the chemical makeup of the extracellar “coat” or “matrix” materials secreted by the following cells: bacteria, fibroblasts, osteoblasts, plant cells, fungi.

2. The iodine number of a compound is defined as the number of grams of I_2 absorbed (through addition to C=C bonds to give a diiodo derivative) per 100 g of fat. NOTE: Iodine monochloride or iodine monobromide are the usual halogenating reagents but the iodine number is expressed in terms of grams of I_2. The saponification number is the number of milligrams of KOH needed to completely saponify (hydrolyze and neutralize the resulting fatty acids) 1 g of fat. A pure triglyceride has a saponification number of 198 and an iodine number of 59.7.

   a) What is the relative molecular mass?
   b) What is the average chain length of the fatty acids?
   c) What is the number of double bonds in the molecule?

3. Spermaceti (a wax from the head of the sperm whale) resembles high molecular mass hydrocarbons in physical properties and inertness toward $Br_2 / CHCl_3$ and $KMNO_4$; on qualitative analysis, it gives positive tests only for carbon and hydrogen. However, its IR spectrum shows the presence of an ester linkage, and quantitative analysis gives the empirical formula $C_{16}H_{32}O$. A solution of the wax in alcoholic KOH is refluxed for a long time. Titration of an aliquot shows that one equivalent of base is consumed for every 475 g of wax. Water and ether are added to the cooled reflux mixture, and the aqueous and etheral layers are separated. Acidification of the aqueous layer yields a solid A with a neutralization equivalent of 260 ± 5. Evaporation of the ether layer gives solid B, which could not be titrated. Reduction of either spermaceti or A by lithium aluminum hydride gave B as the only product.

   What is the likely structure of spermaceti?

4. Stearic acid (1.16 g) was dissolved in 100 ml of ethanol. A 10 µl portion of the resulting solution was pipetted onto a clean surface of a dilute HCl solution (in a shallow tray) where it spread to form a monolayer of stearic acid. The layer was compressed (by moving a Teflon barrier across the tray) until the surface pressure π started to rise sharply and reached ~20 dyn / cm. Note that $\pi = \gamma_o - \gamma$ where γ is the measured surface tension with the film present and γ_o is the higher surface tension of water alone. The compressed film occupied a 20 x 24 cm area. Calculate the cross-sectional area of an alkyl chain in stearic acid. [See J. B. Davenport, *in Biochemistry and Methodology of Lipids* (A. R. Johnson and J. B. Davenport, eds.), pp. 47-83. Wiley-Intersience, New York, 1971; and M. C. Phillips, *in Progress in Surface and Membrane Science* (J. F. Danielli, D. M. Rosenberg, and D. A. Cadenhead, eds.) Vol. 5, pp. 139-221. Academic Press, New York, 1972.]

5. In 1925, E. Gorter and F. Grendel (*J. Exp. Med.* **41**, 439) reported measurements in which they extracted lipid from red blood cell membranes with acetone, spread the lipids as a monolayer, and measured the area of the compressed monolayer. They then estimated the surface area of an erythrocyte and calculated that the ratio of the lipids (as a monolayer) to the surface area of the red blood cell was 1.9–2.0. More modern experiments gave the following: each erythrocyte membrane contains 4.5×10^{-16} mol of phospholipid and 3.1×10^{-16} mol of cholesterol.

   a) If the cross-sectional areas of phospholipid and cholesterol molecules in a membrane are taken as 0.70 and 0.38 nm^2, respectively, what surface area would be occupied in a monolayer?
   b) If the measured surface area of an erythrocyte is 167 μm^2, what is the ratio of the area calculated in (a) to the area of the cell surface?
   c) How might you explain the difference between this answer and that of Gorter and Grendel? See E. D. Korn (1966) *Science* **153**, 1491–1498.

Study Questions

6. The following experimental observations are related to biological membrane structure and function. Discuss the implications of each observation with respect to membrane structure.

   a) Many macrocyclic antibiotics (nonactin, valinomycin, and others) form 1: 1 complexes with alkali metal cations in a highly selective manner. The complexes are readily soluble in nonpolar organic solvents. These antibiotics increase the electrical conductance and permeability to alkali metal cations of synthetic phospholipid membranes. Valinomycin increases the electrical conductance of thylakoid membranes of chloroplasts in the presence of K^+ but not in the presence of Na^+; it also uncouples oxidative phoshorylation in mitochondria. (See Chapter 18.)
   b) Treatment of intact chloroplasts with a galactolipase releases glactose from galactosyl diglycerides. Treatment of red blood cell ghosts with phopholipase C releases about 75% of the lipid P in water-soluble form. In neither case is the structural integrity of the membrane destroyed.
   c) Using sphingomyelin as a hapten, antibodies specific for this lipid can be produced. When red blood cell ghosts are exposed to these antibodies, it can be shown that the antibodies react, but only on one side of the membrane.

7. Describe the structure of biological membranes and the characteristic functions of lipid-, protein-, and carbohydrate-containing components. Describe the differences between inner and outer membrane surfaces.

8. Compare the distribution of triglycerides, phosphatidylcholine, phosphatidylethanolamine, sphingomyelins, glycolipids, and cholesterol within cells. Consider differences between the two sides of membranes.

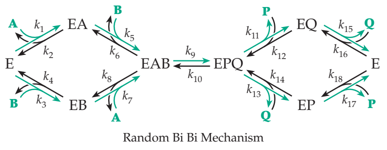
9. Consider the chemistry underlying the labeling of cell surfaces with each of the following:
   a) Lactoperoxidase,
   b) galactose oxidase,
   c) formylmethionylsulfone methyl phosphate,
   d) the diazonium salt of diiodosulfanilic acid,
   e) fluorescent antibodies,
   f) antibodies conjugated with ferritin.

   Write the equations for the chemical reactions involved. State what surface groups will be labeled by each reagent. List special advantages of each of these reagents.

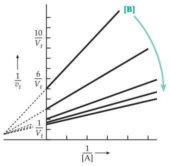
10. Which would be the more effective detergent in the pH range 2 to 3, sodium lauryl sulfonate or sodium laurate? Why?

11. Suppose that a cell contains 10 mM Na^+ and 100 mM K^+ and that it is bathed in extracellular fluid containing 100 mM Na^+ and 5 mM K^+. How much energy will be required to transport three equivalents of Na^+ out and two equivalents of K^+ in? Compare this with $\Delta G'$ for hydrolysis of ATP at pH 7. Assume that the membrane is permeable to Cl^-.

12. An *E. coli* cell is said to contain about 10^5 molecules of an envelope protein of MW = 36,500. If the latter is spherical and the spheres are closely packed in a hexogonal lattice, how much of the surface area of the bacterium would be covered? What would the diameter of the protein be? What spacing would be required if 10^5 molecules covered the surface completely? Suggest a shape for the protein molecule that is consistent with the requirement.



Random Bi Bi Mechanism



Bimolecular reactions of two molecules, A and B, to give two products, P and Q, are catalyzed by many enzymes. For some enzymes the substrates A and B bind into the active site in an ordered sequence while for others, binding may be in a random order. The scheme shown here is described as random Bi Bi in a classification introduced by Cleland. Eighteen rate constants, some second order and some first order, describe the reversible system. Determination of these kinetic parameters is often accomplished using a series of double reciprocal plots (Lineweaver-Burk plots), such as those at the right.

Contents

455 .... **A. Information from Kinetics**
455 .... 1. Measuring the Speed of an Enzymatic Reaction
457 .... *First-order reactions*
457 .... *Turnover numbers and units of activity*
458 .... *Second-order reactions*
458 .... *Reversible chemical reactions*
458 .... 2. Formation and Reaction of Enzyme–Substrate Complexes
460 .... *Linear forms for rate equations*
460 .... *Nonlinear equations and integrated rate equations*
461 .... *Kinetics with high enzyme concentrations*
461 .... 3. Diffusion and the Rate of Encounter of an Enzyme with Substrate
462 .... *The "cage effect" and rotation of molecules*
463 .... *The rate of substrate binding*
463 .... *The displacement of bound ligands by substrate*
464 .... 4. Reversible Enzymatic Reactions
464 .... *Reactions of two or more substrates*
465 .... *"Ping-pong" mechanisms*
466 .... *Isomechanisms*
466 .... *Dead-end complexes*
466 .... *Handling rate equations for complex mechanisms*
467 .... *The rapid equilibrium assumption*
467 .... *Isotope exchange at equilibrium*
468 .... 5. Kinetics of Rapid Reactions
468 .... *Flowing substrates together*
468 .... *Observing relaxation*
468 .... *Rapid photometric methods*
468 .... *Some results*
469 .... 6. Cryoenzymology
469 .... 7. The Effect of pH on Enzymatic Action
471 .... **B. Inhibition and Activation of Enzymes**
471 .... 1. Competitive Inhibitors
473 .... 2. Noncompetitive Inhibition and Activation; Allosteric Sites
475 .... 3. Inhibitors in the Study of Mechanisms
475 .... 4. Allosteric Effectors in the Regulation of Enzyme Activity
477 .... 5. Irreversible Inhibition of Enzymes
478 .... **C. The Specificity of Enzymatic Action**
478 .... 1. Complementarity of Substrate and Enzyme Surfaces
478 .... 2. Stereospecificity and Prochiral Centers
480 .... *Stereochemical numbering*
480 .... *Trigonal prochiral centers*
481 .... 3. Induced Fit and Conformational Changes
481 .... 4. Specificity and k_{cat}
482 .... 5. Proofreading
482 .... **D. Mechanisms of Catalysis**
482 .... 1. The "Transition State"
483 .... *Quantitative transition state theory*
484 .... *Transition state inhibitors*
484 .... *Describing the transition state*
486 .... *Getting to the transition state*
486 .... 2. Microscopic Reversibility
486 .... 3. Acid and Base Catalysis
486 .... *Acidic and basic groups in enzymes*
487 .... *Acid–base catalysis of mutarotation*
487 .... *General base and general acid catalysis*
490 .... *The Brönsted relationships*
490 .... *Concerted acid–base or tautomeric catalysis*
491 .... 4. Hydrogen Bonding and the Transfer of Protons within Active Sites
491 .... *Ultra–fast proton transfer*
492 .... *Proton transfer rates*
493 .... *Marcus theory*
493 .... *Diffusion-controlled dissociation of protons*
493 .... *Coupled proton transfers*
493 .... *Unusually strong hydrogen bonds*
494 .... 5. Covalent Catalysis
494 .... 6. Proximity and Orientation
495 .... 7. The Microenvironment
496 .... 8. Strain and Distortion
497 .... 9. Why Oligomeric Enzymes?
497 .... 10. Summary
497 .... **E. Classification of Enzymes**

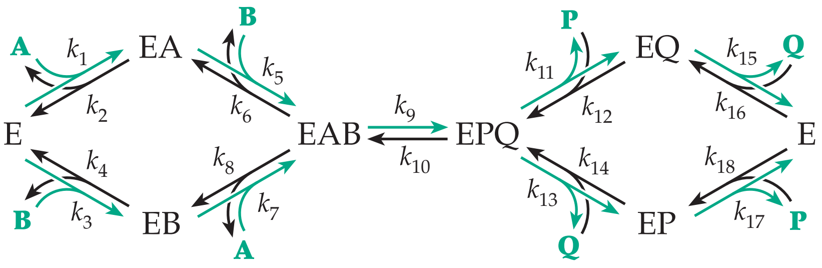
498 .... **References**
501 .... **Study Questions**

Boxes

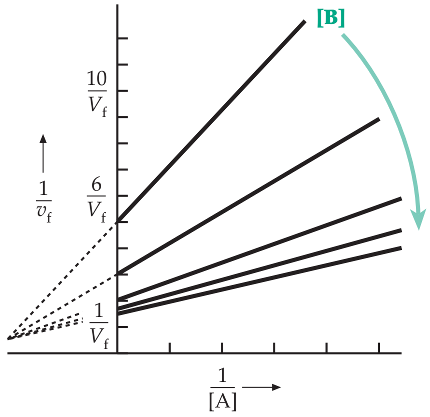
456 .... Box 9-A A Historical Note on Enzymes
470 .... Box 9-B Growth Rates of Cells
473 .... Box 9-C The Sulfonamides as Antimetabolites
479 .... Box 9-D Receptors, Agonists, and Antagonists
485 .... Box 9-E Chorismate Mutase
488 .... Box 9-F Immunophilins as Rotamases

Enzymes: The Catalysts of Cells

9



Random Bi Bi Mechanism



Most of the machinery of living cells is made of enzymes. Thousands of them have been extracted from cells and have been purified and crystallized. Many others are recognized only by their catalytic action and have not yet been isolated in pure form. Most enzymes are soluble globular proteins but an increasing number of RNA molecules are also being recognized as enzymes. Many structural proteins of the cell also act as catalysts. For example, the muscle proteins actin and myosin together catalyze the hydrolysis of ATP and link the hydrolysis to movement (Chapter 19). Catalysis is one of the most fundamental characteristics of life.

How do we recognize that a protein or RNA molecule *is* an enzyme? The answer is that *enzymes are recognized primarily by their ability to catalyze a chemical reaction.* For this reason, an everyday operation for many biochemists is the measurement of the catalytic activity of enzymes. Only by measuring the rates of the catalyzed reactions carefully and quantitatively has it been possible to isolate and purify these remarkable molecules.

Since the beginning of biochemical investigation enzymes have held a special fascination for chemists and biologists. How can these easily destroyed substances catalyze reactions with such speed and without formation of significant quantities of side products? Some enzymes increase the velocity of a single chemical reaction of a specific compound by a factor of as much as 10^{10}. How can a protein do this? In this chapter we'll consider both ways of measuring enzymatic activity and basic mechanisms of catalysis.

A. Information from Kinetics

The quantitative study of catalysis by enzymes, i.e., the study of enzyme kinetics, is a highly developed branch of biochemistry. It is one of our most important means of learning about the mechanisms of catalysis at the active sites of enzymes.[1–13a] By determining **rate constants *k*** under a variety of conditions we can learn just how fast an enzyme can act, how tightly it binds its substrates to form the enzyme-substrate (ES) complexes essential to catalysis, how specific it is with respect to substrate structure, and how it is affected by compounds that inhibit or activate the catalysis.

1. Measuring the Speed of an Enzymatic Reaction

A major goal in kinetic studies is to establish a **rate equation** which describes the velocity of a reaction in terms of **kinetic constants** and other experimentally measurable parameters. To measure the velocity of any chemical reaction precisely we must start the reaction at a definite time by rapidly mixing together the two or more reactants. Then, while keeping the mixture at an accurately constant temperature and pH, we must measure the concentration of a reactant or product after a fixed time interval, or at various times. No end of ingenuity has gone into devising ways of doing this for particular enzymes. Whatever the procedure, the information we must obtain is the **rate** at which some concentration changes with time. We can then construct a **progress curve** showing the decrease in concentration [S] of the reactant (**substrate**) or the increase in the concentration of **product** [P] with time (Fig. 9-1).

BOX 9-A A HISTORICAL NOTE ON ENZYMES

While the earliest physiologists postulated a “vital force” to explain the chemical reactions of cells, the existence of biochemical substances promoting reactions outside of the body was recognized at least by the early 1600s. However, the role of yeast in fermentation was still unknown and it was thought that both alcoholic fermentation and animal digestion were caused by unknown substances called **ferments**. In 1752, Reamur demonstrated the solvent power of the gastric juice of birds and by 1783 Spallanzani had extended the studies to humans and other species.[a] In 1836 Schwann isolated the enzyme **pepsin**[b] from gastric juice.

In the same year Jacob Berzelius introduced the concept of catalysis, which he developed as a result of studies of the effects of acids and bases in promoting the hydrolysis of starch and of the effects of metals on the decomposition of hydrogen peroxide. Berzelius proposed the term **catalyst** from the Greek “katalysis,” meaning “dissolution.” Although he had been concerned primarily with inorganic catalysts, Berzelius recognized that a natural catalyst, an amylase that causes the hydrolysis of starch, had already been isolated from germinating barley in 1833 and that “in living plants and animals thousands of catalytic processes take place.”[c] Some chemists and physiologists accepted Berzelius’ concept of biochemical catalysis immediately, but many did not. The matter was complicated by Pasteur’s discovery that yeast cells were the causative agent of alcoholic fermentation. Only after Edward Buchner’s reports in 1897 that a juice formed by grinding yeast with sand and filtering could still ferment sugar (see Chapters 15 and 17) was the reality of enzymes in metabolism[d] generally accepted. The word **enzyme** (from the Greek “in yeast”) was introduced earlier by F. W. Kuhne, professor of physiology in Heidelberg and a person who had accepted Berzelius’ concept.[c]

---

[a] Richmond, C. (1986) *Trends Biochem. Sci.* **11**, 528–530
[b] Schwann, T. (1836) *Arch. Anat. Physiol.*, 90–138
[c] Hoffmann-Ostenhof, O. (1978) *Trends Biochem. Sci.* **3**, 186–187
[d] Buchner, E. (1897) *Ber. Deut. Chem. Ges.* **30**, 117–124

The velocity v of an enzymatic reaction is defined as the rate at which a substrate disappears or at which a product is formed, the two being identical:

$$v = -d[\mathrm{S}] / dt = d[\mathrm{P}] / dt \qquad (9\text{-}1)$$

Under the steady-state conditions that usually apply (see p. 449), *the rate of increase of product will be the same as the rate of decrease of substrate.* The units of velocity are moles per liter per second (M s^{-1}) or more traditionally in enzymology moles per liter per minute. We are interested in the **instantaneous velocity,** which at any time is given by the slope of the progress curve (Fig. 9-1). We usually want to measure the velocity immediately after the reaction is started to avoid the decrease in rate that comes from the depletion of substrate or accumulation of products. In some cases, as in the example given in Fig. 9-1, this is difficult to do with accuracy.

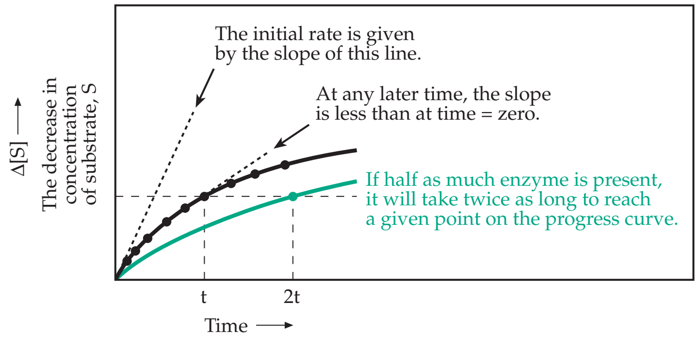


Figure 9-1 The progress curve for an enzymatic reaction in which the substrate S is converted into products.

In some chemical reactions, which involve first-order processes, a logarithmic plot of the progress curve (log[S] vs t) gives a straight line so that the initial slope need not be determined. However, most **enzyme assays** give progress curves that remain nearly linear for only short periods of time. In these cases we need very sensitive methods for detecting products. These may involve colorimetric or fluorimetric measurements or the use of radioactive substrates. One of the most sensitive approaches is to arrange the assay so that a product of the reaction serves as a catalyst for another enzymatic process, thus amplifying the amount of final product to be measured.[14,15] Another approach is to measure velocities of reactions in a very small volume, e.g., 200 nanoliters, and to continuously separate and measure products using capillary electrophoresis and fluorescence detection.[16]

If the progress curve is not a straight line at the beginning of the reaction (0 time) and if the amount of compound that reacts in a fixed interval of time is taken as the rate, an erroneous answer will be obtained. Sometimes an **integrated rate expression** that describes the time course of product formation can be used (see Eq. 9-22). Even when the progress curve is nonlinear as shown in Fig. 9-1, it is possible to estimate relative rates by noting that in most cases if one unit of enzyme yields a certain amount of product in time t_1, the same amount of product will be formed by n units of enzyme in time t_1/n.

First-order reactions. In many chemical reactions the rate of decrease of the concentration of a given reactant [A] is found experimentally to be directly proportional to the concentration of that reactant at any given time:

$$v = -d[\mathrm{A}]/dt = k[\mathrm{A}] \tag{9-2}$$

Such a reaction is described as **first order** and the proportionality constant ***k*** is known as the **rate constant.** Such first-order kinetics is observed for **unimolecular processes** in which a molecule of A is converted into product P in a given time interval with a probability that does not depend on interaction with another molecule. An example is radioactive decay. Enzyme–substrate complexes often react by unimolecular processes. In other cases, a reaction is **pseudo-first order**; compound A actually reacts with a second molecule such as water, which is present in such excess that its concentration does not change during the experiment. Consequently, the velocity is apparently proportional only to [A].

A first-order rate constant k has units of s^{-1}. When [A] = 1, $v = k$. Thus, k is a measure of the speed in mol l^{-1} s^{-1} of the reaction of a substance at unit activity. As a first-order reaction proceeds, [A] decreases and at time t is given by any one of the following three equivalent expressions (Eq. 9-3). These are obtained by integration of Eq. 9-2 in which t_0 is the time at which the reaction was started.

$$\begin{gathered}[\mathrm{A}] = A_0 e^{-kt} \\ \ln([\mathrm{A}_0]/[\mathrm{A}]) = kt \\ \log[\mathrm{A}_0] - \log[\mathrm{A}] = kt/2.303\end{gathered} \tag{9-3}$$

Equation 9-3 is the equation of exponential decay. A characteristic of exponential decay is that [A] is halved in a time that is independent of concentration. The **half-life** is $t_{1/2}$:

$$t_{1/2} = \ln 2/k = 0.693/k \tag{9-4}$$

The **relaxation time** τ for A is defined by Eq. 9-5 and represents the time required for the concentration [A] to fall to $1/e$ (or ~0.37) of its initial value.

$$\tau = 1/k = t_{1/2}/\ln 2 \tag{9-5}$$

Turnover numbers and units of activity. When an enzyme is catalyzing product formation at the maximum possible rate, V_{max}, we can usually assume that the active site of every molecule of enzyme contains a substrate in the form of an intermediate enzyme–substrate complex ES and that this complex is being converted to products according to Eq. 9-6:

$$\begin{gathered}d[\mathrm{P}]/dt = V_{\mathrm{max}} = k[\mathrm{ES}] = k[\mathrm{E}]_t \\ \text{or} \qquad V_{\mathrm{max}} = k_{\mathrm{cat}}[\mathrm{E}]_t\end{gathered} \tag{9-6}$$

Here E_t is the total enzyme, namely, the free enzyme E plus enzyme–substrate complex ES. The equation holds only at **substrate saturation,** that is, when the substrate concentration is high enough that essentially all of the enzyme has been converted into the intermediate ES. The process is first order in enzyme but is zero order in substrate. The rate constant k is a measure of the speed at which the enzyme operates. When the concentration $[\mathrm{E}]_t$ is given in moles per liter of *active sites* (actual molar concentration multiplied by the number of active sites per mole) the constant k is known as the **turnover number**, the **molecular activity**, or $\boldsymbol{k_{\mathrm{cat}}}$. The symbol k_{cat} is also used in place of k in Eq. 9-6 for complex rate expressions in which k_{cat} cannot represent a single rate constant but is an algebraic expression that contains a number of different constants.

Turnover numbers can be measured only when the concentration of the enzyme is known. Partly for this reason the activity of an enzyme is usually given as **specific activity,** the *units of activity per milligram of protein*. One **international unit** is the amount of enzyme that produces 1 µmol of product per minute under standard (usually optimal) conditions. The International Union of Biochemistry[17] has recommended a larger unit, the **katal** (kat), the amount of enzyme that converts one mol s^{-1} of substrate to product.

1 kat = 6 x 10^7 international units
1 international unit = 16.67 nkat (nanokatals)

If the enzyme is pure and is saturated with substrate under the standard assay conditions, the following relationships hold.

$$\begin{aligned}\text{Turnover No.} &= \text{katals/mol of active sites} \\ &= [\text{nkat/mg}] \times M_r \times 10^{-6}/n \\ &= [\text{international units/mg}] \times M_r \times 10^{-3}/60n\end{aligned}$$

Here M_r is the relative molecular mass of the enzyme and n is the number of active sites per molecule. Since the activity of an enzyme is dependent on both temperature and pH, these variables must be specified.

Turnover numbers of enzymes vary from <1 to ~ 10^6 s^{-1}. Trypsin, chymotrypsin, and many intracellular enzymes have turnover numbers of ~10^2 s^{-1}. The fastest enzymes, which include **catalase** (Chapter 16), **carbonic anhydrase** (Chapter 13), and **Δ^5-3-oxosteroid isomerase** (Chapter 13), have maximum turnover numbers of 2 x 10^5 s^{-1} or more. Compare these reaction rates with those of a typical organic synthesis in the laboratory. A reaction mixture must often be heated for hours ($k < 10^{-3}$ s^{-1}). Many enzymes accelerate rates by factors of greater than 10^6 over those observed in the absence of an enzyme at a comparable temperature and pH. Enzymes often bring two or more substrates together, binding them at a specific location in their active sites. Because of this, rapid reactions can be catalyzed even when the reactants are present in low concentrations.

Second-order reactions. For a chemical reaction to occur between two molecules, A and B (Eq. 9-7), they must meet and collide.

$$A + B \xrightarrow{k_2} P \qquad (9\text{-}7)$$

The velocity of such a **second-order** process is characterized by a **bimolecular rate constant** k_2 and is proportional to the product of the concentrations of A *and* B:

$$v = k_2[A]\,[B] \qquad (9\text{-}8)$$

The units of k_2 are M^{-1} s^{-1}. If [B] is present at unit activity, the rate is k_2[A], a quantity with units of s^{-1}. We can see that the bimolecular, or second-order, rate constant for reaction of A with B may be compared with first-order constants when the second reactant B is present at unit activity. In many real situations, reactant B is present in large excess and in a virtually constant concentration. The reaction is pseudo-first order and the experimentally observed rate constant k_2[B] is an *apparent first-order rate constant*. The bimolecular rate constant k_2 can be obtained by dividing the apparent constant by [B].

Reversible chemical reactions. In any reversible process, we must consider rate constants for both the forward and the reverse reactions. At equilibrium a reaction proceeds in the forward direction at exactly the same velocity as in the reverse reaction so that no change occurs. For this reason there is always a relationship between the equilibrium constant and the rate constants. For Eq. 9-9, k_1 is the bimolecular rate constant for the forward reaction and k_2 the unimolecular rate constant for the reverse reaction. The equilibrium constant *K* can easily be shown, from Eq. 9-2 and 9-8, to equal k_1 / k_2 for the reaction of Eq. 9-9.

$$A + B \underset{k_2}{\overset{k_1}{\rightleftarrows}} P$$

$$\text{Equilibrium constant} = K = \frac{[P]}{[A][B]} = \frac{k_1}{k_2} \qquad (9\text{-}9)$$

The student should be aware that in kinetic equations rate constants are usually numbered consecutively via subscripts and that the subscripts do not imply anything about the molecularity. The system which is used here employs odd-numbered constants for steps in the forward direction and even-numbered constants for steps in the reverse direction. However, *many authors number the steps in the forward direction consecutively and those in the reverse direction with corresponding negative subscripts.*

What relationships exist between experimentally observable rates and k_1 and k_2 for a reversible reaction? Consider first the simplest case (Eq. 9-10):

$$A \underset{k_2}{\overset{k_1}{\rightleftarrows}} B \qquad (9\text{-}10)$$

If pure A is placed in a solution, its concentration will decrease until it reaches an equilibrium with the B which has been formed. It is easy to show that in this case [A] does not decay exponentially but $[A] - [A]_{equil}$ does. If $\log([A] - [A]_{equil})$ is plotted against time a first-order rate constant *k*, characteristic of *the rate of approach to equilibrium*, will be obtained. Its relationship to $k_1$and k_2 is given by Eq. 9-11.

$$k = k_1 + k_2 \qquad (9\text{-}11)$$

The relaxation time τ for approach to equilibrium can be expressed as follows:

$$\tau^{-1} = k = \tau_1^{-1} + \tau_2^{-1} = k_1 + k_2 \qquad (9\text{-}12)$$

and for the more complex case of Eq. 9-9 as

$$k = \tau^{-1} = k_1([A]_e + [B]_e) + k_2 \qquad (9\text{-}13)$$

where $[A]_e$ and $[B]_e$ are the equilibrium concentrations of A and B.[18]

2. Formation and Reaction of Enzyme–Substrate Complexes

An abundance of evidence indicates that the first step in enzymatic catalysis is the combining of the enzyme and substrate reversibly to form a complex, ES (Eq. 9-14). Formation of the complex is normally reversible. ES can either break up to form enzyme and substrate again or it can undergo conversion to a product or products, often by a unimolecular process. Three rate constants are needed to describe this system for a reaction that is irreversible overall. A complete description of the kinetic behavior is fairly involved.

$$E + S \underset{k_2}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_3} E + P \tag{9-14}$$

For example, the kinetics may be different within cells, where molar concentrations of enzymes often exceed those of substrate, than in the laboratory. In most laboratory experiments the enzyme is present at an extremely low concentration (e.g., 10^{-8} M) while the substrate is present in large excess. Under these circumstances the **steady-state approximation** can be used. For this approximation *the rate of formation of ES from free enzyme and substrate is assumed to be exactly balanced by the rate of conversion of ES on to P*. That is, for a relatively short time during the duration of the experimental measurement of velocity, the concentration of ES remains essentially constant. To be more precise, the steady-state criterion is met if the absolute rate of change of a concentration of a transient intermediate is very small compared to that of the reactants and products.[19]

The **Michaelis–Menten equation** (Eq. 9-15) describes the initial reaction rate of a single substrate with an enzyme under steady-state conditions.

$$v = \frac{V_{\max}}{1 + K_m/[S]} = \frac{V_{\max}[S]}{K_m + [S]}$$
$$\text{where } K_m = (k_2 + k_3)/k_1 \tag{9-15}$$

This can be rearranged as follows:

$$\frac{V_{\max}}{v} = 1 + K_m/[S] \tag{9-16}$$

Equation 9-15 provides a relationship between the velocity observed at a particular substrate concentration and the maximum velocity that would be achieved at infinite substrate concentration. The quantities $V_{\max}$ and K_m are often referred to as the **kinetic parameters** of an enzyme and their determination is an important part of the characterization of an enzyme. Equation 9-15 can be derived by setting the rate of formation of the ES complex ($k1$[E][S]) in the steady state equal to its rate of breakdown, ($[k_2 + k_3]$[ES]). Rearranging and substituting K_m, as defined in Eq. 9-15, we obtain Eq. 9-17.

$$[E][S] = \frac{(k_2 + k_3)}{k_1}[ES] = K_m[ES] \tag{9-17}$$

Using this equation, together with a mass balance relationship ($[E] = [E]_t - [ES]$), we can solve for $[ES]/[E]_t$, the fraction of enzyme combined as enzyme–substrate complex (Eq. 9-18).

$$\frac{[ES]}{[E]_t} = \frac{[S]}{K_m + [S]} \tag{9-18}$$

The maximum velocity, $V_{\max} = k_3[E]_t$, is attained only when all of the enzyme is converted into ES. Under other conditions $v = k_3$[ES] and Eq. 9-19 holds.

$$[ES] / [E]_t = v / V_{\max} \tag{9-19}$$

Substituting from Eq. 9-19 into Eq. 9-18 gives the Michaelis–Menten equation (Eq. 9-15).

In many cases, the rate at which ES is converted back to free E and S is much greater than the rate of conversion of ES to products ($k_2 \gg k_3$). In such cases K_m equals k_2/k_1, the dissociation constant for breakdown of ES to free enzyme and substrate (sometimes called K_s). Thus, *K_m sometimes has a close inverse relationship to the strength of binding of substrate to enzyme.* In such cases, $1/K_m$ is a measure of the affinity of the substrate for its binding site on the enzyme and for a series of different substrates acted on by the same enzyme. The more tightly bound substrates have the lower values of K_m. But beware! The condition that k_3 is negligible compared to k_2 may be met with some (poorer) substrates but may not always be met by others. From Eq. 9-15 we see that K_m is always greater than or equal to K_s, but it may be less than K_s for more complex mechanisms.

Figure 9-2 shows a plot of velocity against substrate concentration as given by Eq. 9-15. The position of $V_{\max}$ on the ordinate is marked, but it should be clear that the experimental velocity (v) can never attain $V_{\max}$ unless [S] is very high relative to K_m. The value of v approaches $V_{\max}$ asymptotically. Since K_m is defined as the value of [S] at which $v = V_{\max}/2$, its value can be estimated from Fig. 9-2. However, K_m cannot be determined reliably because of the difficulty of establishing the value of $V_{\max}$ from a plot of this type. Notice that the curve of Figure 9-2 is identical in form to the saturation curve for reversible binding shown in Fig. 7-1.

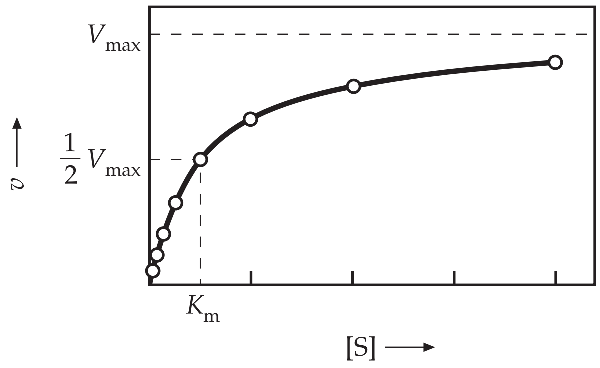


Figure 9-2 Plot of observed velocity v vs substrate concentration [S] for an enzyme-catalyzed reaction.

Linear forms for rate equations. To obtain K_m and V_{max} from experimental rate data, Eq. 9-15 can be transformed by algebraic rearrangement into one of several linear forms. The popular **double-reciprocal** or **Lineweaver–Burk** plot of $1/v$ against $1/[\text{S}]$ (Fig. 9-3) is described by Eq. 9-20. The values of K_m/V_{max} and $1/V_{\text{max}}$ can be evaluated from the slope and intercept, respectively, of this straight line plot.

$$\frac{1}{v} = \frac{1}{V_{\text{max}}} + \frac{K_m}{V_{\text{max}}}\frac{1}{[\text{S}]} \qquad (9\text{-}20)$$

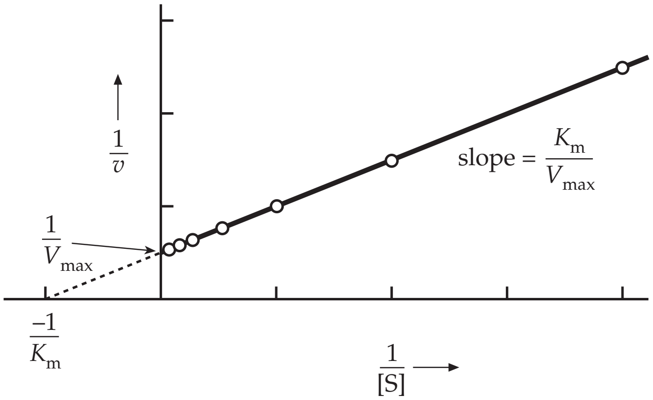


Figure 9-3 Double-reciprocal or Lineweaver–Burk plot of $1/v$ vs $1/[\text{S}]$. The intercept on the vertical axis gives $1/V_{\text{max}}$ and the slope gives K_m/V_{max}. The intercept on the horizontal axis equals $-1/K_m$.

Another linear plot, the **Eadie–Hofstee plot**, is that of $v/[\text{S}]$ vs v (Fig. 9-4). It is related to the Scatchard plot (Fig. 7-3) and is fitted by Eq. 9-21.

$$\frac{v}{[\text{S}]} = \frac{V_{\text{max}}}{K_m} - v\frac{1}{K_m} \qquad (9\text{-}21)$$

While the point for $[\text{S}] = 0$ and $v = 0$ cannot be plotted, the ratio $v/[\text{S}]$ approaches V_{max}/K_m as v approaches zero. Notice the distribution of the points in Fig. 9-4. Substrate concentrations were chosen such that the increase in velocity from point to point is more or less constant, a desirable experimental situation. The points on the Eadie–Hofstee plot are also nearly evenly distributed, but those of the Lineweaver–Burk plot are compressed at one end. (However, if the substrate concentrations for successive points are selected in the ratios 1, 1/3, 1/5, 1/7, and 1/9, the spacing will be uniform on the Lineweaver–Burk plot.) A second advantage of the Eadie–Hofstee plot is that the entire range of possible substrate concentration from near zero to infinity can be fitted onto a single plot.

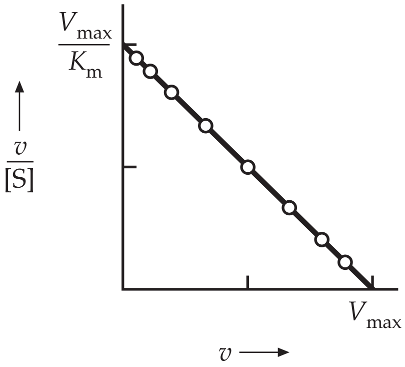
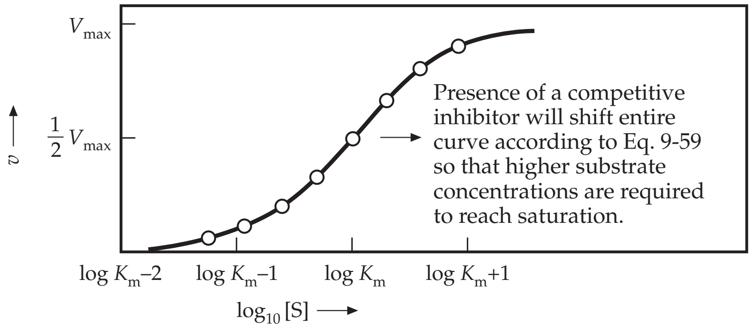


Figure 9-4 The Eadie–Hofstee plot of $v/[\text{S}]$ against v. The slope is $-1/K_m$; the intercept on the vertical axis is V_{max}/K_m and that on the horizontal axis is V_{max}.

As pointed out at the beginning of Section A, depletion of substrate with time lowers [S] from its initial value to some extent. This can have an especially adverse effect on points obtained at low [S], e.g., points on the right side of Fig. 9-3 or on the left side of Fig. 9-4. Equations 9-20 and 9-21 will be more precise if the average substrate concentration over the time period of the assay, rather than that at zero time,[20] is used.

Nonlinear equations and integrated rate equations. In discussions of the *control* of metabolism through regulation of enzymatic activity, it is often better to plot v against log [S] as in Fig. 9-5. This plot also has the virture that the entire range of attainable substrate concentrations can be plotted on one piece of paper if the point for $[\text{S}] = 0$ (at minus infinity) is omitted. The same scale can be used for all enzymes. The plot is S-shaped, both for simple cases that are represented by Eq. 9-15 and for enzymes that bind substrate cooperatively (Section B,5). Thus, the classification or "hyperbolic" vs "sigmoidal" is lost. However, the degree of cooperativity can be directly measured from the midpoint slope of the curve. A disadvantage is that it is awkward to measure V_{max} from a plot of this type, and it may be preferable to obtain it from a linear plot (Figs. 9-3 or 9-4). Alternatively, computer-assisted methods can be used to obtain both K_m and V_{max} and to fit a curve to the experimental points as in Fig. 9-5.

An attractive alternative to the use of initial velocities and linear plots is to measure the kinetic parameters V_{max} and K_m using points all along the progress curve (Fig. 9-1). Various procedures for doing this have been devised.[21,22] For example, the integrated form of Eq. 9-15 can be given as Eq. 9-22 or 9-23.



Presence of a competitive inhibitor will shift entire curve according to Eq. 9-59 so that higher substrate concentrations are required to reach saturation.

Figure 9-5 Plot of v against log [S] for an enzyme-catalyzed reaction.

$$V_{max}t = [P] + K_m \ln \frac{[S]_0}{[S]_0 - [P]} \qquad (9\text{-}22)$$

$$V_{max}\, t = [S]_0 - [S] + K_m \ln ([S]_0 / [S]) \qquad (9\text{-}23)$$

Here $[S]_0$ is the initial substrate concentration at $t = 0$ and [S] and [P] are the substrate and product concentrations at any later time. In one type of linear plot,[21] based on Eq. 9-23, $([S]_0 - [S]) / t$ is plotted against $1/t \ln ([S]_0 / [S])$, the slope of the plot being $-K_m$ and the intercepts being V_{max} and V_{max} / K_m. Computer programs are available for analysis of enzymatic progress curves, even for complex mechanisms.[23,24]

Kinetics with high enzyme concentrations. Laboratory studies of the kinetics of purified enzymes are usually conducted with enzyme concentrations of 10^{-7} to 10^{-10} M, but within cells enzyme concentrations are probably often in the range 10^{-6} to 10^{-5} M,[25] which may be higher than the concentrations of the substrates upon which the enzymes act. Be cautious in drawing conclusions about kinetics under such circumstances! Methods have been devised for handling kinetic data when the concentration of enzyme is greater than K_m, a condition that can lead to intolerably high errors if the usual equations are applied.[26,27]

3. Diffusion and the Rate of Encounter of an Enzyme with Substrate

What determines the value of k_1 in Eq. 9-14? This rate constant represents the process by which the substrate and enzyme find each other, become mutually oriented, and bind to form ES. If orientation and binding are rapid enough, the rate will be determined by the speed with which the molecules can come together by diffusion. Large molecules in solution are free to travel for only a tiny fraction of their diameter as a result of their frequent collision with solvent molecules. The result is visible in the **Brownian movement** of microscopic particles suspended in a fluid. If an individual particle is followed, it is seen to undergo a "random walk," moving in first one direction then another. Albert Einstein showed that if the distances transversed by such particles in a given time Δt are measured, the mean square of these Δx values Δ^2 can be related by Eq. 9-24 to the diffusion constant D (which is usually given in units of $cm^2 s^{-1}$).

$$\Delta^2 = 2 D \,\Delta t \qquad (9\text{-}24)$$

For molecules, the Brownian movement cannot be observed directly but the diffusion constant can be measured, for example, by observing the rate of spreading of a boundary between two different concentrations of the substance.[28] The diffusion constant for $^1H^2HO$ (HDO) in H_2O at 25°C is 2.27×10^{-5} cm^2 s^{-1}, and the values for the ions K^+ and Cl^- are about the same.[29] For many small molecules D is approximately 10^{-5} cm^2 s^{-1}, and the value decreases as the size of the molecule increases. Thus, for the 13.7-kDa ribonuclease, $D = 1.1 \times 10^{-6}$ cm^2 s^{-1}, and for the 500-kDa myosin, D is ~ 1×10^{-7} cm^2 s^{-1}. For the spherical particle the **Stokes–Einstein equation** (Eq. 9-25) can be used to relate the diffusion coefficient to the radius, the coefficient of viscosity η, the Boltzmann constant k_B, and the temperature (Kelvin) T.

$$D = \frac{k_B T}{6\pi\eta r} \qquad (9\text{-}25)$$

To estimate the rate constant for a reaction that is controlled strictly by the frequency of collisions of particles, we must ask how many times per second one of a number n of particles will be hit by another of the particles as a result of Brownian movement. The problem was analyzed in 1917 by Smoluchowski,[30,31] who considered the rate at which a particle B diffuses toward a second particle A and disappears when the two collide. Using Fick's law of diffusion, he concluded that the number of **encounters** per milliliter per second was given by Eq. 9-26.

$$\text{Number of encounters/ml s}^{-1} = 4\pi (D_A + D_B)(r_A + r_B)\, n_A n_B \qquad (9\text{-}26)$$

Here D_A and D_B are the two diffusion coefficients, r_A and r_B are the radii (in Å) of the two particles, and n_A and n_B are the numbers of particles per milliliter. The number of encounters per liter per second is $(4\pi / 1000)(D_A + D_B)(r_A + r_B) N^2 [A][B]$, where N is Avogadro's number. Dividing this frequency by N gives the rate of collision v in M s^{-1}, a velocity that can be equated (Eq. 9-27) with k_D [A] [B], where k_D is a second-order rate constant which determines the **diffusion controlled limit**, that is, the maximum possible rate of a reaction.

$$v = \frac{4\pi}{1000}(D_A + D_B)(r_A + r_B)[A][B]N = k_D[A][B] \tag{9-27}$$

This equation can be rearranged to Eq. 9-28.

$$k_D = \frac{4\pi N}{1000}(D_A + D_B)(r_A + r_B)\text{M}^{-1}\,\text{s}^{-1} \tag{9-28}$$

While this equation is thought to overestimate the diffusion-limited rate constant slightly, it is a good approximation. If the diffusing particles are approximately spherical, diffusion constants D_A and D_B can be calculated from Eq. 9-25, and Eq. 9-28 becomes Eq. 9-29.

$$k_D = \frac{2RT}{3000\eta}\left(2 + \frac{r_A}{r_B} + \frac{r_B}{r_A}\right) \tag{9-29}$$

The value of k_D does not vary greatly as the ratio of radii r_A / r_B is changed, and in most cases it may be assumed that $r_A / r_B \approx 1$, in which case the equation simplifies (Eq. 9-30).

$$k_D \approx \frac{8RT}{3000\eta}\,\text{M}^{-1}\,\text{s}^{-1} \tag{9-30}$$

For water at 25°C, the coefficient of viscosity η is ~0.01 poise (1 poise = 10^{-5} newton cm^{-2}) which leads, according to Eq. 9-30, to $k_D \sim 0.7 \times 10^{10}\ \text{M}^{-1}\,\text{s}^{-1}$. It was shown by Debye[32] that this rate constant must be multiplied by a correction factor when charged particles rather than uncharged spheres diffuse together. This factor may be of the order of 5–10 for a substrate and enzyme carrying two or three charges and may act to either increase or decrease reaction rates. Since the viscosity of cytoplasm is 0.02–0.03 poise,[33] k_D will be reduced correspondingly.

A simple alternative derivation of Eq. 9-30 was developed by Dexter French,[34] who also provided the author with most of this discussion of encounter theory. Consider a small element of volume ΔV swept out by a particle as it moves through the solution for a distance equal to its own radius. This element of volume will equal πr^3:

$$\Delta V = \pi r^2 \times r = \pi r^3 \tag{9-31}$$

It will be swept out in a time Δt which can be calculated from Eq. 9-24 as $r^2 / 2D$. Substituting the value of D given by Eq. 9-25, we obtain for Δt:

$$\Delta t = \frac{3\pi\eta r^3}{k_B T} \tag{9-32}$$

Division of Eq. 9-31 by Eq. 9-32 gives the volume swept out per second by one particle:

$$\Delta V / \Delta t = k_B T / 3\eta\ \text{cm}^3\,\text{s}^{-1} \tag{9-33}$$

Since the collision radius for two particles of equal size is two times the particle radius, the effective volume swept out will be four times that given by Eq. 9-33. Since both particles are diffusing, the effective diffusion constant will be twice that used in obtaining Eq. 9-28. Thus, the effective volume swept out by the particle in a second will be eight times that given by Eq. 9-33. The volume swept out by one mole of particles is equal to k_D (recall that the second-order rate constant has dimensions of liter mol^{-1} s^{-1}). Thus, when converted to a moles per liter basis and multiplied by 8, Eq. 9-33 should (and does) become identical with the Smoluchowski equation (Eq. 9-30).

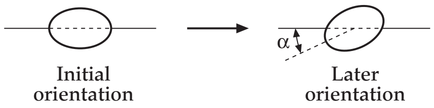
The volume given by Eq. 9-33 is about 1.4×10^{-11} cm^3, which could be represented approximately by a cube 2.4 µm on a side. If we compare this volume with that of a cell (Table 1-2) or of an organelle, we see that in one second an enzyme molecule will sweep out a large fraction of the volume of a small cell, mitochondrion, chloroplast, etc.

The "cage effect" and rotation of molecules. It is of interest to compare the bimolecular rate constant for encounters calculated by the Smoluchowski theory (Eq. 9-29) with the corresponding bimolecular rate constant for molecular collisions given by the kinetic theory of gases (Eq. 9-34).

$$k(\text{collision}) = \frac{N(r_A + r_B)^2}{1000}\left[8\pi k_B T\left(\frac{1}{m_A} + \frac{1}{m_B}\right)\right]^{1/2} \tag{9-34}$$

Here m_A and m_B are the masses of the two particles. This rate constant is also relatively independent of molecular size and for spheres varies from (4 to 11) x 10^{11} M^{-1} s^{-1}, over an order of magnitude greater than the rate constant for encounters. In a solution, molecules still collide at about the same rate as in a gas so that *100 to 200 collisions occur between two particles for each encounter*. However, during the time of the single encounter, the particles are together in a solvent "cage." While in this solvent cage, both substrate and enzyme molecules undergo random rotational motions. Successive collisions bring them together in different orientations, one of which is likely to lead to a sufficiently close match of complementary surfaces (of substrate and binding site) so that formation of a "productive" ES complex takes place.

Molecular rotation in a solution is described quantitatively by diffusion laws (analogous to Fick's laws) for which a **rotary diffusion constant** θ is defined.[35,36] Consider a group of molecules all oriented the same way initially, then undergoing rotary diffusion until their orientations become random. If we measure the orientation of each molecule by an angle we see that initially the value of cos α is 1 but that when the angles become random the mean value of cos α averaged over all molecules is zero. The rotary relaxation time τ is the time required for the mean value of cos α to fall to



1/*e* (which occurs when $\alpha = 68.5°$). For a sphere θ is given by Eq. 9-35.

$$\theta = \tfrac{1}{2}\tau = \frac{k_B T}{8\pi\eta r^3} \tag{9-35}$$

Ellipsoidal or rod-shaped molecules have two different rotary diffusion constants while, if the dimensions of the molecules are different along all three axes, three constants must be specified.[36]

From Eq. 9-35 we can calculate (if $\eta = 0.01$ poise) the following values for a small spherical molecule (substrate) of ~1 nm length and for a spherical enzyme of 5 nm diameter:

$r = 0.5$ nm $\theta \approx 1.3 \times 10^9\ \text{s}^{-1}$
$r = 2.5$ nm $\theta \approx 1.0 \times 10^7\ \text{s}^{-1}$

We see that smaller molecules rotate much faster than large ones and that rotational relaxation times for small proteins are of the same order of magnitude as k_D for diffusion-limited encounter. However, for very large molecules, especially long rods, the rotary relaxation time about a short axis may be a large fraction of a second.

The rate of substrate binding. At very low substrate concentrations the Michaelis–Menten equation (Eq. 9-15) simplifies as follows:

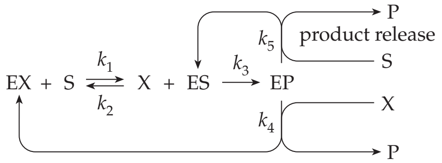
$$v = (V_{\text{max}} / K_m)[\text{S}] \tag{9-36}$$

Since $V_{\text{max}} = k_{\text{cat}}\,[\text{E}]_t$ (Eq. 9-6) and at low [S] most of $[\text{E}]_t$ is free E, we obtain the following equation:[2]

$$v = (k_{\text{cat}} / K_m)[\text{E}][\text{S}] \tag{9-37}$$

From this we see that k_{cat}/K_m *is the apparent second-order rate constant for the reaction of free enzyme with substrate*. As such it cannot exceed the diffusion controlled limit k_D of Eqs. 9-28 to 9-30 which falls in the range of $10^9 - 10^{11}\ \text{M}^{-1}\ \text{s}^{-1}$. Experimentally observed values of k_{cat}/K_m are always less than this limit, indicating that a certain time is required for a substrate molecule to become oriented and seated in the active site.[37] However, for several real enzymes values of k_{cat}/K_m of 10^7 to $3 \times 10^8\ \text{M}^{-1}\ \text{s}^{-1}$ have been observed. For triose phosphate isomerase Albery and Knowles obtained a value of $4 \times 10^8\ \text{M}^{-1}\ \text{s}^{-1}$, so close to the diffusion controlled limit that these authors regard this enzyme as a *nearly perfect catalyst*, one that could not have evolved further because it is already catalyzing the reaction with substrate at almost the maximum velocity that is possible.[37–39]

The displacement of bound ligands by substrate. Jenkins pointed out that in many instances a substrate must displace another ligand from the active site to form the ES complex.[40] For example, a binding site for an ionic substrate often already contains an ion, either of the product P or some other ion X which Jenkins calls a **substrate surrogate**. For this situation Eq. 9-14 must be replaced by the following set of equations:



(9-38)

Here k_2 and also k_4 and k_5, are second-order rate constants. The release of product, as determined by k_4 and k_5, may be rate-limiting. At zero time the reverse reactions may be ignored, and steady-state analysis shows that the Michaelis–Menten equation (Eq. 9-16b) will be replaced by Eq. 9-39. Here, D is a constant and A is also constant if X is present at a fixed concentration.

$$\begin{aligned} \frac{V_{\text{max}}}{v} &= 1 + \frac{k_2[\text{X}] + k_3}{k_1[\text{S}]} + \frac{k_3(1/k_5 - 1/k_1)}{k_4[\text{X}]/k_5 + [\text{S}]} \\ &= 1 + K_m/[\text{S}] + \text{D}/(\text{A} + [\text{S}]) \end{aligned} \tag{9-39}$$

When k_5, the rate constant for displacement of product by substrate, is very small, this equation simplifies to Eq. 9-40.

$$\begin{aligned} \frac{V_{\text{max}}}{v} &= 1 + k_3/k_4[\text{X}] + (k_2[\text{X}] + k_3)/k_1[\text{S}] \\ &= 1 + k_3/k_4[\text{X}] + K_m'/[\text{S}] \end{aligned} \tag{9-40}$$

At high concentrations of X the second term becomes negligible and the equation becomes identical to the Michaelis–Menten equation (Eq. 9-16) except that K_m', the apparent Michaelis constant, now increases as [X] increases (Eq. 9-41). This is exactly what is observed for competitive inhibitors where

$$K_m{}' = \frac{k_2[\text{X}] + k_3}{k_1} \tag{9-41}$$

Thus substrate surrogates at high concentrations are often competitive inhibitors. However, at low [X] the second term in Eq. 9-41 may be important and *X then serves as an activator*. Under many circumstances, Eq. 9-39 does not simplify further and a nonlinear relationship between $1/v$ and $1/[\text{S}]$ is observed, with the shape of the curves being influenced by the values of constants A and D (see also Section B,4).[40]

4. Reversible Enzymatic Reactions

For some enzyme-catalyzed reactions the equilibrium lies far to one side. However, many other reactions are freely reversible. Since a catalyst promotes reactions in both directions, we must consider the action of an enzyme on the reverse reaction. Let us designate the maximum velocity in the forward direction as V_f and that in the reverse direction as V_r. There will be a Michaelis constant for reaction of enzyme with product $K_{m\text{P}}$, while $K_{m\text{S}}$ will refer to the reaction with substrate.

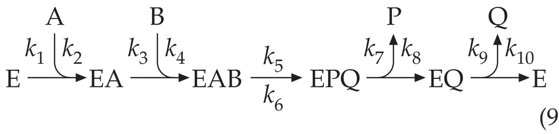
As in any other chemical reaction, there is a relationship between the rate constants for forward and reverse enzyme-catalyzed reactions and the equilibrium constant. This relationship, first derived by the British kineticist J. B. S. Haldane and proposed in his book *Enzymes*[41] in 1930, is known as the **Haldane relationship**. It is obtained by setting $v_\text{f} = v_\text{r}$ for the condition that product and substrate concentrations are those at equilibrium. For a single substrate–single product system it is given by Eq. 9-42.

$$K_\text{eq} = V_\text{f} K_{m\text{P}} / V_\text{r} K_{m\text{S}} \tag{9-42}$$

Because the Haldane relationship imposes constraints on the values of the velocity constants and Michaelis constants, it is of some value in understanding regulation of metabolism. Consider the case that the maximum forward velocity V_f is high and that $K_{m\text{S}}$ has a moderately low value (fairly strong binding of substrate). If the reaction is freely reversible ($K_\text{eq} \sim 1$) and the velocity of the reverse reaction V_r is about the same as that of the forward reaction, it will necessarily be true (see Eq. 9-42) that the product P will also be fairly tightly bound. If $V_\text{r} << V_\text{f}$, the value of $K_{m\text{P}}$ will have to be much *lower* than that of $K_{m\text{A}}$. In such a situation P will remain tightly bound to the enzyme and since V_r is low, it will tend to clog the enzyme. Such **product inhibition** may sometimes slow down a whole pathway of metabolism. In such a case, the only way that an enzymatic sequence can keep going in the forward direction is for product P to be removed rapidly by a subsequent reaction with a second enzyme. The first enzyme may be thought of as possessing a kind of "one-way valve" that turns off the flow in a metabolic pathway when the concentration of its product rises.

Reactions of two or more substrates. Enzymes frequently catalyze the reaction of two, three, or even more different molecules to give one, two, three, or more products. Sometimes all of the substrate molecules must be bound to an active site at the same time and are presumably lined up on the enzyme molecule in such a way that they can react in proper sequence. In other cases, the enzyme may transform molecule A to a product, and then cause the product to react with molecule B. The number of variations is enormous.[1,10,12]

The order in which two molecules A and B bind to an enzyme to form a complex EAB may be completely *random* or it may be *obligatorily ordered*. Both situations occur with real enzymes. Cleland introduced a widely used method of depicting the possibilities.[42] For example, Eq. 9-43 shows the reaction of A and B in an *ordered sequence* to form the complex EAB which is then isomerized to EPQ, the complex formed by binding the two products P and Q to the enzyme. The rate constant to the left of each vertical arrow or above each horizontal arrow refers to the reaction in the forward direction as indicated by the arrow while the other constants (to the right or below the arrows) refer to the reverse reactions. The velocity in the forward direction for an enzyme with ordered binding is given by Eq. 9-44a,



(9-43)

$$v_f = \frac{V_\text{f}[\text{A}][\text{B}]}{K_{\text{eqA}} K_{m\text{B}} + K_{m\text{B}}[\text{A}] + K_{m\text{A}}[\text{B}] + [\text{A}][\text{B}]} \tag{9-44a}$$

which may also be written in the reciprocal form (Eq. 9-44b):

$$\frac{1}{v_\text{f}} = \frac{1}{V_\text{f}}\left(1 + \frac{K_{m\text{A}}}{[\text{A}]} + \frac{K_{m\text{B}}}{[\text{B}]} + \frac{K_{\text{eqA}} K_{m\text{B}}}{[\text{A}][\text{B}]}\right) \tag{9-44b}$$

An alternative form of this equation (Eq. 9-45), proposed by Dalziel,[43] is sometimes used.

$$\frac{[\mathrm{E}]_t}{v_f} = \phi_0 + \frac{\phi_1}{[\mathrm{A}]} + \frac{\phi_2}{[\mathrm{B}]} + \frac{\phi_{12}}{[\mathrm{A}][\mathrm{B}]} \qquad (9\text{-}45)$$

The kinetic parameters of Eq. 9-44 are V_f, the maximum velocity in the forward direction, the two Michaelis constants, $K_{m\mathrm{B}}$ and $K_{m\mathrm{A}}$, and the equilibrium constant K_{eqA}, for reversible dissociation of the complex EA and which is equal to k_2 / k_1. The relationship between the parameters of Eq. 9-44 (K_m's, V's, and K_{eqA}'s) and the rate constants $k_1 - k_{10}$ is not obvious. However, remember that the parameters are experimental quantities determined by measurements on the enzyme. Sometimes, but not always, it is possible to deduce some of the values of individual rate constants from the experimental parameters.

An equation similar to 9-44a can be written for the velocity v_r of the reaction of P and Q. Also an equation can be written in similar form for $v_\mathrm{f} - v_\mathrm{r}$, i.e., the instantaneous velocity of the reaction in any mixture of all four components A, B, P, and Q.

The kinetic parameters of Eq. 9-44b are often obtained from experimental data through the use of reciprocal plots (Fig. 9-6). However, Eq. 9-44b is linear only if the concentration of one or the other of the substrates A and B is kept constant. For this reason a series of experiments is usually performed in which [A] is varied while [B] is held constant. Then [A] is held constant and [B] is varied. Each of these experiments leads to a family of lines (Fig. 9-6A) whose slopes and intercepts are measured. The slopes and intercepts of this family of curves are then plotted against the reciprocal of the second concentration, i.e., the one that was held fixed.

From a set of these **secondary plots**, V_f and one of the Michaelis constants can be determined (Fig. 9-6B and C). Using two sets of secondary plots, all of the constants of Eq. 9-44 may be established. Alternatively, a computer can be used to examine all of the data at once and to obtain the best values of the parameters. The latter approach is desirable because estimates of the standard deviations of the parameters can be obtained. However, the user must take care to ensure that the experimental errors are correctly estimated and are not simply estimates of how well the computer has fitted the points on the assumption that they contain no error.[23]

The meaning of the kinetic parameters may be slightly difficult to grasp. V_f is the velocity that would be obtained if both [A] and [B] were high enough to saturate the enzyme. Each K_m corresponds to that for a simple system in which the concentration of the second substrate is at a high, saturating value.

For the bimolecular reaction that we have considered, there are two Haldane relationships:

$$K_{\mathrm{eq}} = \frac{V_\mathrm{f} K_{m\mathrm{P}} K_{\mathrm{dQ}}}{V_\mathrm{r} K_{\mathrm{dA}} K_{m\mathrm{B}}} = \left(\frac{V_\mathrm{f}}{V_\mathrm{r}}\right)^2 \frac{K_{\mathrm{dP}} K_{m\mathrm{Q}}}{K_{m\mathrm{A}} K_{\mathrm{dB}}} \qquad (9\text{-}46)$$

Of these, only the first is ordinarily used.

"Ping-pong" mechanisms. A common type of mechanism that is especially prevalent for enzymes with tightly bound cofactors has been dubbed **ping-pong** because the enzyme alternates between

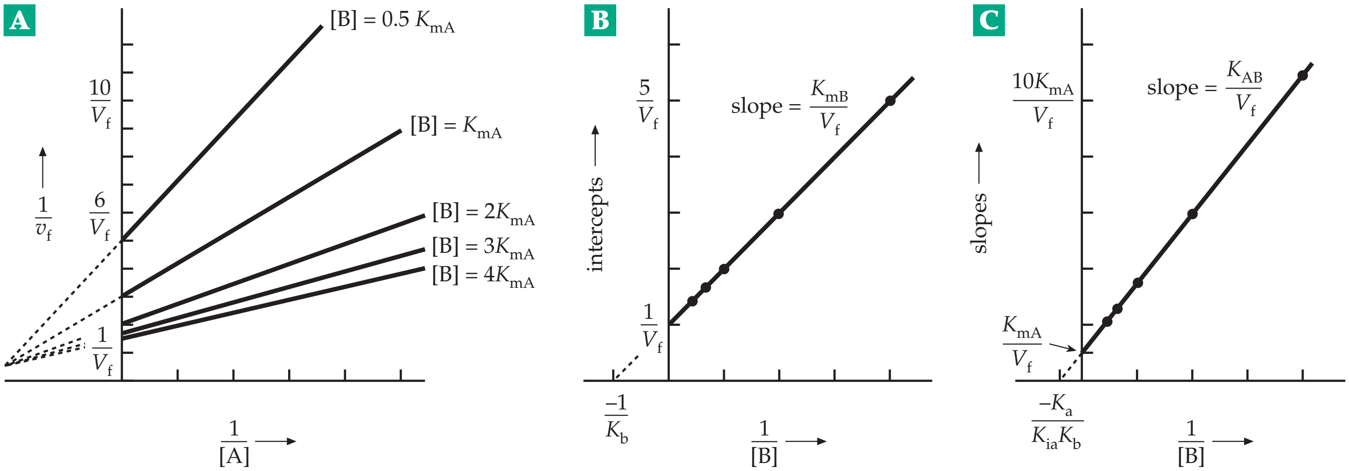
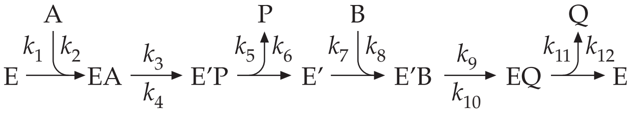


Figure 9-6 Reciprocal plots used to analyze kinetics of two-substrate enzymes. (A) Plot of $1 / v_\mathrm{f}$ against $1 / [\mathrm{A}]$ for a series of different concentrations of the second substrate B. (B) A secondary plot in which the intercepts from graph A are plotted against $1 / [\mathrm{B}]$. (C) Secondary plot in which the slopes from graph A have been plotted against $1 / [\mathrm{B}]$. The figures have been drawn for the case that $K_{m\mathrm{A}} = 10^{-3}$ M, $K_{m\mathrm{B}} = 2\,K_{m\mathrm{A}}$, and $K_{\mathrm{AB}} = K_{\mathrm{eqA}} K_{m\mathrm{B}}$ (Eq. 9-46) = $K_{m\mathrm{A}} / 200$ and [A] and [B] are in units of moles per liter. Eadie–Hofstee plots of $v_\mathrm{f} / [\mathrm{A}]$ vs v_f at constant [B] can also be used as the primary plots. The student can easily convert Eq. 9-44 to the proper form analogous to Eq. 9-21.

two forms E and E′ (Eq. 9-47). Substrate A reacts via complex EA to form E′, a modified enzyme that often



Ping-pong mechanism for reaction of A and B to form P and Q (9-47)

contains an altered coenzyme. At the same time A is changed to product P still bound to the enzyme. P dissociates leaving E′ which is then able to react with the second substrate B and to go through the second half of the cycle during which E′ is converted back to E. An example is provided by the aminotransferases (Eq. 14-25) in which the coenzyme pyridoxal phosphate is interconverted with pyridoxamine phosphate.

The rate equations of the ping-pong mechanism resemble that for the ordered bimolecular reaction (Eq. 9-44), but each has one less term (Eq. 9-48):

$$v_f = \frac{V_f[A][B]}{K_{mB}[A] + K_{mA}[B] + [A][B]} \tag{9-48a}$$

or

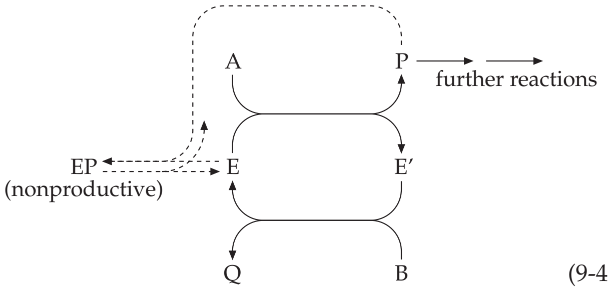
$$\frac{1}{v_f} = \frac{1}{V_f}\left(1 + \frac{K_{mA}}{[A]} + \frac{K_{mB}}{[B]}\right) \tag{9-48b}$$

A diagnostic feature of the ping-pong mechanism is that the families of lines (in double reciprocal plots) obtained when one substrate is held constant while the other is varied no longer intersect as they do in Fig. 9-6A but are parallel. They must be truly parallel and the experimentalist must be aware that nearly parallel lines may sometimes be observed for sequential reactions. Thus, if K_{eqA} of Eq. 9-44b is small enough the last term of that equation will be small compared to the other terms and the equation will be approximated by Eq. 9-48b. The reaction will appear to be ping-pong even though it is sequential and the reaction proceeds through the ternary complex EAB.

One less kinetic parameter can be obtained from an analysis of the data for a ping-pong mechanism than can be obtained for ordered reactions. Nevertheless, in Eq. 9-47, twelve rate constants are indicated. At least this many steps must be considered to describe the behavior of the enzyme. Not all of these constants can be determined from a study of steady-state kinetics, but they may be obtained in other ways.

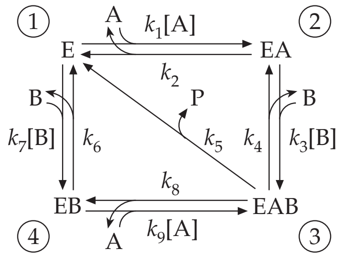
Isomechanisms. A catalyst functions over and over again without being altered. However, during a single turnover it is likely to undergo a temporary change. For example, if an enzyme assists in removing a proton from a substrate, some functional group of the enzyme will be protonated during part of the catalytic cycle. This proton must dissociate before another catalytic cycle begins, and in some instances this dissociation may be a rate-limiting step. Northrup and coworkers developed the use of product inhibition to identify such isomechanisms.[44]

Dead-end complexes. In the steady state of action of an enzyme with ping-pong kinetics, part of the enzyme is in form E and part in form E′. Ideally, E would have affinity only for A and Q, while E′ would have affinity only for B and P. However, in many real situations E also binds B and P weakly; similarly, E′ binds A and Q. This is because the products and reactants usually have structural features in common. The propensity of enzymes with ping-pong kinetics to form **dead-end complexes** (also called abortive complexes) may sometimes have a regulatory function. In Eq. 9-49 the reactions of Eq. 9-47 have been rearranged to depict a situation in which product P normally undergoes a sequence of further reactions. However, if P accumulates to a high enough concentration it can react reversibly to form a dead-end complex EP. This is an effective form of product inhibition which can be relieved only by a lowering of the concentration of P through its further metabolism.



(9-49)

Handling rate equations for complex mechanisms. While steady-state rate equations can be derived easily for the simple cases discussed in the preceding sections, enzymes are often considerably more complex and the derivation of the correct rate equations can be extremely tedious. The **topological theory of graphs**, widely used in analysis of electrical networks, has been applied to both steady-state and nonsteady-state enzyme kinetics.[45–50] The method employs diagrams of the type shown in Eq. 9-50. Here



(9-50)

the reaction of an enzyme with two substrates A and B with a **random** order of binding is depicted. (In contrast, Eq. 9-43 shows the case of ordered binding of two substrates.) When complex EAB is formed, it can decompose to free enzyme and to the single product P. Each one of the **nodes**, which are numbered 1–4 in the diagram, corresponds to a single form of the enzyme. The appropriate first-order rate constant or apparent first-order constant is placed by each arrow. The methods provide easy rules for deriving from such a scheme the steady-state rate equation.

The importance of the simplified schematic methods is apparent when one considers that the steady-state rate equation for Eq. 9-50 would have 6 terms in the numerator and 12 terms in the denominator.[51] In the more complex case in which EAB breaks down to two products P and Q with a random order of release, the rate equation contains 672 terms in the denominator. In such cases it is worthwhile to enlist the help of a computer in deriving the equation.[24,52–54]

The rapid equilibrium assumption. Rate equations for enzymes are often simplified if a single step, e.g., that of reaction of complex EAB to product in Eq. 9-50, is **rate limiting**.[54a] If it is assumed that all reaction steps preceding or following the rate-limiting step are at equilibrium, the equation for random binding with a two-substrate and two-product reaction simplifies to one whose form is similar to that obtained for ordered binding (Eq. 9-44). In the absence of products P and Q Eq. 9-44 will correctly represent the steady-state rate equation corresponding to Eq. 9-50. However, this simplification may not be valid for a very rapidly acting enzyme.

Isotope exchange at equilibrium. Consider the reaction of substrates A and B to form P and Q (Eq. 9-51). If both reactants and both products are present with the enzyme and in the ratio found at equilibrium no net reaction will take place. However, the reactants and products will be continually interconverted under the action of the enzyme. Now if a small amount of

$$A + B \rightleftarrows P + Q \qquad (9\text{-}51)$$

highly labeled reactant (A* or B*) is added, the rate at which isotope is transferred from the labeled reactant into one or the other of the products can be measured. In general, a label in one of the substrates will appear in only one of the products.

Figure 9-7A shows the rate of exchange of isotopically labeled glucose (glucose*) with glucose 6-phosphate as catalyzed by the enzyme hexokinase (Chapter 12). The exchange rate is plotted against the concentration of glucose 6-phosphate with the ratio [glucose] / [glucose 6-phosphate] constant at 1/19, such that an equilibrium ratio for reactants and products is always maintained. As can be seen from the graph, this exchange rate increases monotonically as substrate concentrations are increased. This is also true for the rate of ATP–ADP exchange. The fact that both exchange rates increase continuously indicates random binding of substrates.[55] The inequality of the two maximal exchange rates suggests that release of glucose 6-phosphate may be slower than that of ADP.

Figure 9-7B shows similar plots for lactate dehydrogenase.[53] In this case, after an initial rise (that is not regarded as significant), the pyruvate*–lactate exchange reaches a high constant value as the amount of pyruvate is increased (with a constant [pyruvate] / [lactate] ratio of 1/35). However, the NAD*–NADH exchange increases rapidly at first but then drops abruptly as the pyruvate and lactate concentrations continue to increase. This suggests an ordered mechanism (Eq. 9-43) in which NAD⁺ and NADH represent A and Q, respectively, and pyruvate and lactate represent B and P. As the concentrations of B and P become very high, the

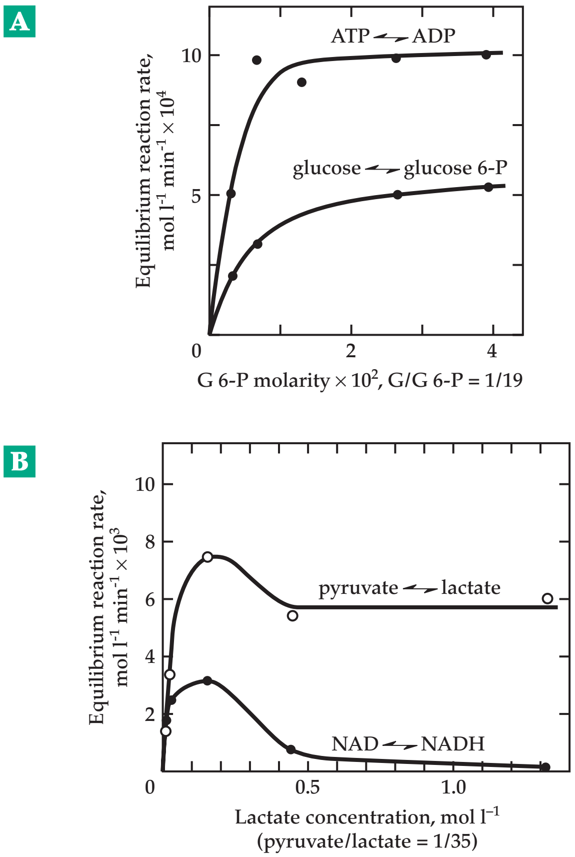


Figure 9-7 (A) Effect of glucose and glucose 6-phosphate concentrations on reaction rate of yeast hexokinase at equilibrium. Reaction mixtures contain 1–2.2 mM ATP, and 25.6 mM ADP at pH 6.5. From Fromm *et al*.[51] (B) Effect of lactate and pyruvate concentrations on equilibrium reaction rates of rabbit muscle lactate dehydrogenase. Reaction mixtures contained 1.7 mM NAD⁺, and 30 – 46 µM NADH in Tris-nitrate buffer, pH 7.9, 25°C. From Silverstein and Boyer.[53]

enzyme shuttles back and forth between EA and EQ, but these two complexes rarely dissociate to give free enzyme and A or Q. Hence, the A*–Q exchange rate drops.

In other cases a label may be transferred from A into P or from B into Q. Information on such exchanges has provided a valuable criterion of mechanism which is considered in Chapter 12, Section B,4.

5. Kinetics of Rapid Reactions

The fastest steps in an enzymatic process cannot be observed by conventional steady-state kinetic methods because the latter cannot be applied to reactions with half-times of less than about 10 s. Consequently, a variety of methods have been developed[18,56–59a] to measure rates in the range of 1 to 10^{13} s^{-1}.

Flowing substrates together. One of the first rapid kinetic methods to be devised consists of rapidly mixing two flowing solutions together in a special mixing device and allowing the resulting reaction mixture to move at a rate of several meters per second down a straight tube. At a flow velocity of 10 m s^{-1} a solution will move 1 cm in 10^{-3} s. Observations of the mixture are made at a suitable distance, e.g., 1 cm, and with various flow rates. Using spectrophotometry or other observation techniques, the formation or disappearance of a product or reactant can be followed. The special advantage of this technique is that observation can be made slowly. However, it may require large amounts of precious reactant solutions, e.g., those of purified enzymes.

In the **stopped flow** technique two solutions are mixed rapidly by the flow technique during a period of only 1–2 (or a few) milliseconds. A ram drives the solutions from syringes through a mixing chamber into an observation chamber. After the flow stops light absorption, fluorescence, conductivity, or other property, is measured. A means of rapid observation of changes during the reaction is essential. For example, light absorption may be measured by a photomultiplier with data being collected by a computer. Relaxation times as short as a few milliseconds or less can be observed in this way.[59a,b]

Observing relaxation. Kinetic measurements over periods of tens of microseconds or less can be made by rapidly inducing a small displacement from the equilibrium position of a reaction (or series of reactions) and observing the rate of return (relaxation) of the system to equilibrium. Best known is the **temperature jump** method devised by Eigen and associates. Over a period of about 10^{-6} s a potential difference of ~100 kV is applied across the experimental solution. A rapid electrical discharge from a bank of condensers passes through the solution (without any sparking) raising the temperature 2–10 degrees. All the chemical equilibria for which $\Delta H \neq 0$ are perturbed. If some property, such as the absorbance at a particular wavelength or the conductivity of the solution, is measured, very small relaxation times can be determined.

While it may not be intuitively obvious, if the displacement from equilibrium is small, the rate of return to equilibrium can always be expressed as a first-order process (e.g., see Eq. 9-13). In the event that there is more than one chemical reaction required to reequilibrate the system, each reaction has its own characteristic relaxation time. If these relaxation times are close together, it is difficult to distinguish them; however, they often differ by an order of magnitude or more. Therefore, two or more relaxation times can often be evaluated for a given solution. In favorable circumstances these relaxation times can be related directly to rate constants for particular steps. For example, Eigen measured the conductivity of water following a temperature jump[18] and observed the rate of combination of H^+ and OH^- for which τ at 23°C equals 37 x 10^{-6} s. From this, the rate constant for combination of OH^- and H^+ (Eq. 9-52) was calculated as follows (Eq. 9-53):

$$H^+ + OH^- \longrightarrow H_2O \tag{9-52}$$

$$k = 1/\{\tau([OH^-] + [H^+])\} = 1.3 \times 10^{11}\ M^{-1}\ s^{-1} \tag{9-53}$$

Pressure jump and electric field jump methods have also been used, as have methods depending upon periodic changes in some property. For example, absorption of ultrasonic sound causes a periodic change in the pressure of the system.

Rapid photometric methods. Another useful method has been to discharge a condenser through a flash tube over a period of 10^{-12} to 10^{-4} s, causing a rapid light absorption in a sample in an adjacent parallel tube. Following the flash, changes in absorption spectrum or fluorescence of the sample can be followed. The availability of intense lasers as light sources has made it possible to follow the results of light flashes of 5–10 picosecond duration and to measure extremely short relaxation times (Chapter 23).[58,59]

Some results. Rapid kinetic methods have revealed that enzymes often combine with substrates extremely quickly,[60] with values of k_1 in Eq. 9-14 falling in the range of 10^6 to 10^8 M^{-1} s^{-1}. Helix–coil transitions of polypeptides have relaxation times of about 10^{-8} s, but renaturation of a denatured protein may be much slower. The first detectable structural change in the vitamin A-based chromophore of the light-operated proton pump bacteriorhodopsin occurs in ~5 x 10^{-8} s, while a proton is pumped through the membrane in

~10^{-4} s.[61] Interconversion between chair and boat forms of cyclohexane derivatives may have $\tau \sim 10^{-5}$ s at room temperature, while rotation about a C–N bond in an amide linkage may be very slow with $\tau \sim 0.1$ s. The nonenzymatic hydration of the aldehyde pyridoxal phosphate via Eq. 13-1 occurs with $\tau = 0.01 - 0.1$ s, depending upon the pH.[59]

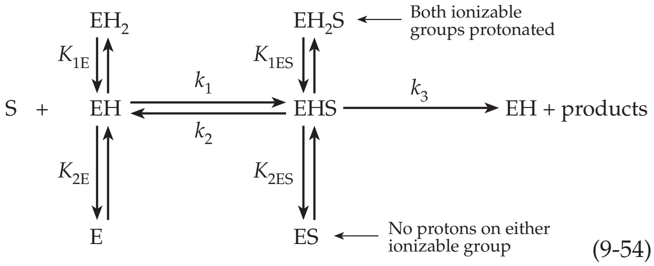
6. Cryoenzymology

An alternative to studying rapid reactions is to cool enzymes to a subzero temperature (down to –100ºC) where reactions proceed more slowly.[60,62] The enzyme must be dissolved in a suitable "cryosolvent," often containing 50–80% by volume of an organic solvent or solvents such as methanol, ethanol, dimethyl sulfoxide, dimethyl formamide, or ethylene glycol. Those containing methanol are especially desirable because of their low viscosities. Kinetics can be studied by various spectroscopic methods and stopped-flow, temperature-jump, and other rapid-reaction techniques can be applied. One goal of cryoenzymology is to stabilize otherwise unstable intermediates. X-ray crystallographic measurements can also be made at these low temperatures and it should be possible to observe structures of intermediate ES complexes. A problem is that the forms of the complexes stabilized may be side-products rather than true intermediates. However, as discussed in Chapter 3, a combination of low-temperature Laue X-ray diffraction and a laser-induced temperature jump may be feasible.

7. The Effect of pH on Enzymatic Action

Because proteins contain many acidic and basic groups it is not surprising that the activity of enzymes often varies strongly with pH. However, it is usually found that the state of protonation of only a few groups has a strong effect on activity. This is understandable because most ionized groups are on the outer surface of protein molecules and most are not close to an active site. Protonation or deprotonation of those groups will hardly ever have a major influence on events in the active site. Often only one or two ionizable groups have highly significant effects.

For many enzymes a plot of $V_{\max}$ against pH is a bell-shaped curve (Fig. 9-8). The **optimum** rate is observed at some intermediate pH, which is often, but not always, in the range of pH 6–9. This type of curve can be interpreted most simply by assuming the presence in the active site of two ionizable groups and three forms of the enzyme with different degrees of protonation: E, EH, and EH_2.


EH2
EH2S
Both ionizable groups protonated
K1E
K1ES
k1
k3
S + EH
EHS
EH + products
k2
K2E
K2ES
E
ES
No protons on either ionizable group

(9-54)

Let us designate the acid dissociation constants (K_a) for the two groups in the enzyme as K_{1E} and K_{2E} and those of the ES complex K_{1ES} and K_{2ES}. However, as discussed in Chapter 6, Section E,2 these consecutive K_a values cannot necessarily be assigned to single groups. They may belong to a system of interacting groups. Each may be the sum of more than one microscopic K_a and there may be extensive tautomerism within the active site. The rate constants k_1, k_2, and k_3 define the rates of formation and breakdown of the ES complex.

If it is assumed that only form EHS reacts to give products, the bell-shaped curves of Fig. 9-8 are obtained. The frequent observation of such curves supports the model of Eq. 9-54. It also suggests that the two ionizable groups may be intimately involved in catalysis, one as an acid and the other as a conjugate base (see Section E,5).

For the simple case illustrated in Eq. 9-54, the pH dependence of the initial maximum velocity, the apparent Michaelis constant, and $V_{\max} / K_m$ are given by Eqs. 9-55 to 9-57.

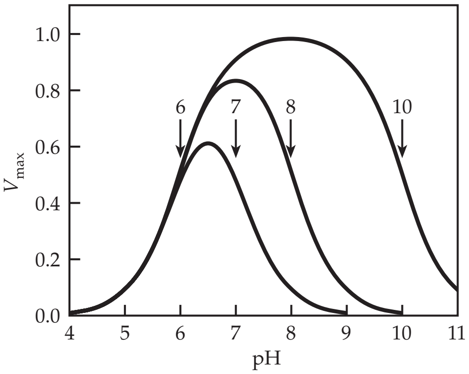


Figure 9-8 Expected dependence of $V_{\max}$ on the pH according to Eq. 9-54 with $k_3[E_t] = 1$, $pK_{1ES} = 6$, and pK_{2ES} as given on the graph. After Alberty.[63] Computer-drawn graph courtesy of Carol M. Metzler.

$$V_{\max} = \frac{k_3[\mathrm{E}]_t}{(1 + [\mathrm{H}^+]/K_{1\mathrm{ES}} + K_{2\mathrm{ES}}/[\mathrm{H}^+])} = \frac{k_3[\mathrm{E}]_t}{\mathrm{F}_{2\mathrm{ES}}} \tag{9-55}$$

$$K_m = \frac{(k_2 + k_3)}{k_1} \cdot \frac{(1 + [\mathrm{H}^+]/K_{1\mathrm{E}} + K_{2\mathrm{E}}/[\mathrm{H}^+])}{(1 + [\mathrm{H}^+]/K_{1\mathrm{ES}} + K_{2\mathrm{ES}}/[\mathrm{H}^+])} = K_m F_{2E}/F_{2ES} \tag{9-56}$$

$$\frac{V_{\max}}{K_m} = \frac{k_1 k_3}{(k_2 + k_3)} \cdot \frac{1}{(1 + [\mathrm{H}^+]/K_{1\mathrm{E}} + K_{2\mathrm{E}}/[\mathrm{H}^+])} = \frac{k_3[\mathrm{E}]_t}{K_m F_{2E}} \tag{9-57}$$

The denominators (which are the Michaelis pH functions given by the first three terms of Eq. 3-6) represent the fraction of enzyme or of ES complex in the monoprotonated state. The pH dependence of enzymatic action is often more complex than that shown in Fig. 9-8 and given by the foregoing equations. However, it is easy to write Michaelis pH functions (see Chapter 3) for enzymes with any number of dissociable groups in both E and ES and to write appropriate equations of the type of

BOX 9-B GROWTH RATES OF CELLS

How do we correctly describe the rate at which a cell grows? Consider bacteria in their rapidly growing "log phase."[a] Each cell divides after a fixed length of time, the **doubling time**. For *E. coli* this may be as short as 17 min in the early stages of growth but becomes somewhat longer as time goes on. A mean value of about 26 min. for *E. coli* at 37°C is typical. In contrast, the doubling time for mammalian cells in tissue culture is often about one day.

If a given volume of culture contains N_0 cells initially, the number N_n after n cell divisions will be:

$$N_n = 2^n N_0$$

From this equation we can calculate that a single bacterium with a generation time of 20 min can produce 2^{144} cells in 48 h of exponential growth. The **exponential growth rate constant *k*** is equal to *the number of doublings per unit time*. Thus, k is the reciprocal of the doubling time. It is easy to show that the number of bacteria present at time t will be given by the following equation.

$$N_t = 2^{kt} N_0$$

This can be rewritten in a form that can be used to determine k by counting the number of bacteria at zero time and at time t.

$$kt = \log_2(N_t/N_0) = \log_{10}(N_t/N_0)\,/\,0.301$$

Another way of expressing the growth is to equate the rate of increase of the number of bacteria with a growth rate constant μ multiplied by the number of bacteria present at that time.

$$dN/dt = \mu N$$

This is a general equation for an autocatalytic reaction and N could be replaced with a concentration, for example, the total content of cellular matter per liter in the medium. From the two preceding equations the following can be derived.[a]

$$\mu = k \ln 2 = 0.69k$$

When bacteria are transferred to new medium there is usually a lag before exponential growth begins. Exponential growth eventually stops and the culture enters **stationary phase**, which is usually followed by relatively rapid death of cells in the culture. It is often desirable to study cell growth under conditions of **continuous cultivation** in which a constant generation time is maintained but the density of cells in the medium does not increase. This can be done with a simple device known as the **chemostat**.[a,b] A culture vessel containing bacteria is stirred to ensure homogeneity. Fresh culture medium continuously enters the vessel from a reservoir and part of the content of the vessel, suspended bacteria included, is continuously removed through another tube. The bacterial population in the vessel builds up to a constant level and can be maintained at the same level for relatively long periods of time.

Here are some other statistics on cell growth. Bacteria growing exponentially expand their linear dimensions by 1.5 nm and synthesize 1.6 x 10^7 Da of new cell material in one second. This is equivalent to ~1000 small proteins and includes 23 ribosomes and 3000 base pairs of DNA at each growing point. The much larger HeLa human tumor cell grows by 0.13 nm / s but makes 4.6 x 10^8 Da / s of material.[c]

---

[a] Stanier, R. Y., Douderoff, M., and Adelberg, E. A. (1970) *The Microbial World*, 3rd ed., Prentice-Hall, Englewood Cliffs, New Jersey (pp. 298–324)

[b] Smith, H. L., and Waltman, P., eds. (1995) *The Theory of the Chemostat; Dynamics of Microbial Competition*, Cambridge University Press, London and New York

[c] Pollard, E. C. (1973) in *Cell Biology in Medicine* (Bittar, E. E., ed), pp. 357–377, Wiley, New York

Eqs. 9-55 to 9-57. Bear in mind that if the **free substrate** contains groups dissociating in the pH range of interest, a Michaelis pH function for the free substrates will also appear in the numerator of Eq. 9-56. If the pH dependence of the enyzme is regulated by a conformational change in the protein, there may be a cooperative gain or loss of more than one proton and the Michaelis pH function must reflect this fact. This can sometimes be accomplished by addition of a term related to Eq. 7-45. For more information see Dixon and Webb,[64] Cleland,[65,66] or Kyte.[67]

Plots of log V_{max} (or log-specific activity) and –log K_m or log (V_{max} / K_m) versus pH yield graphs of the type shown in Fig. 9-9. The curved segments of the graphs that extend for about 1.5 pH units on either side of each pK_a are asymptotic to straight lines of slope 1 when a single proton is involved or of a higher slope for multiple cooperative proton dissociation. The straight lines can be extrapolated and intersect at the pK_a values. (However, it is better to fit a complete curve of theoretically correct shape to the points.) Note that the curved line always passes below or above the intersection point at the value of log 2 = 0.30 except in the case of cooperative proton dissociation when it is closer.

Upward turns in the curve of log K_m vs pH correspond to pK_a values in the free enzyme or substrate and downward turns to pK_a values in ES. This approach to analysis of pH dependence has been adopted widely but often incorrectly. For example, many published curves have very sharp bends in which the curved portion covers less than 3 pH units and the curve is much closer than 0.30 to the extrapolated point. This suggests cooperative proton binding and an *apparent* pK_a that is related to $\bar{K}$ of Eq. 7-21. Cooperativity is always a possibility if a conformational change in the protein is involved.[67]

The simple treatment given above is based on the assumption that all proton dissociations are rapid compared to k_{cat}, that enzyme in only one state of protonation binds substrate, and that ES in only one state of protonation yields products. These assumptions are not always valid. It also assumes that both binding and dissociation of substrate are rapid, that is, to use Cleland's terminology the substrate is not "sticky." For a sticky substrate that dissociates more slowly than it reacts to form products ($k_3 > k_2$; Eq. 9-54), the values of pK_{1E} will be lowered and pK_{1E} of Eq. 9-53 will be raised by log ($1 + k_3 / k_2$).[65,66] In addition to the articles by Cleland, other detailed treatments of pH effects have been prepared by Brocklehurst and Dixon[69] and Tipton and Dixon.[70]

Fumarate hydratase (fumarase), which is discussed in Chapter 13, catalyzes the reversible hydration of fumaric acid to malic acid (Eq. 13-11). It was one of the first enzymes whose pH dependence was studied intensively. A bell-shaped pH dependence for both forward and reverse reactions was observed by Alberty and coworkers[63] and, using Eqs. 9-54 and 9-55, the two apparent pK_a values were measured.

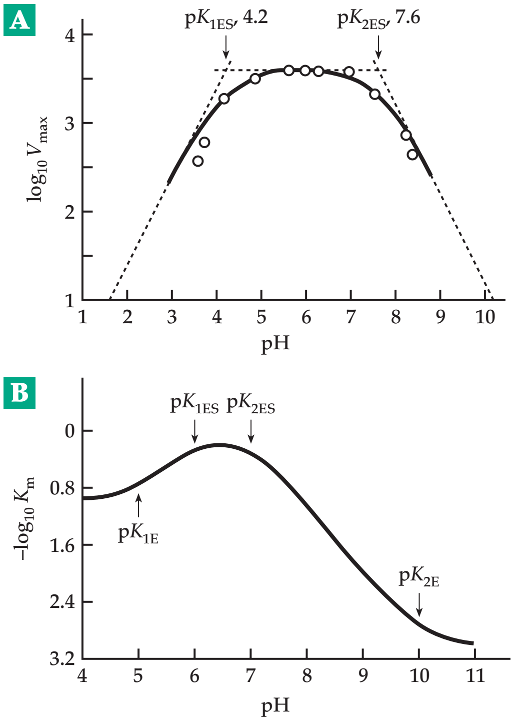


Figure 9-9 (A) Plot of log V_{max} vs pH for a crystalline bacterial α-amylase. From Ono, et al.[68] (B) Theoretical curve of log K_m vs pH for Eq. 9-56 with pK_{1E} = 5, pK_{2E} = 10, pK_{1ES} = 6, and pK_{2ES} = 7. Courtesy of C. Metzler.

B. Inhibition and Activation of Enzymes

The action of most enzymes is inhibited by many substances. Inhibition is often specific, and studies of the relationship between inhibitor structure and activity have been important to the development of our concepts of active sites and of complementarity of surfaces of biomolecules. Inhibition of enzymes is also the basis of the action of a very large fraction of important drugs. Inhibition may be **reversible** or **irreversible**, the latter leading to permanent inactivation of the enzyme. Often, but not always, irreversible inhibition is preceded by reversible binding of the inhibitor at a complementary site on the enzyme surface.

1. Competitive Inhibitors

Inhibitors with close structural similarities to a substrate tend to bind to the substrate site. In truly competitive inhibition, substrate and inhibitor not only compete for the same site but also their binding is reversible and mutually exclusive. The affinity of the inhibitor for the enzyme is expressed quantitatively through the **inhibition constant** K_i which is the *dissociation constant of the enzyme inhibitor complex EI* :

$$K_i = [E][I] / [EI] \tag{9-58}$$

Using the steady state assumption for the mechanism shown in Eq. 9-14, and writing a mass balance equation that includes not only free enzyme and ES but also EI we obtain an equation relating rate to substrate concentration. It is entirely analogous to Eq. 9-15 but K_m is replaced by an apparent Michaelis constant, K'_m:

$$K'_m = K_m \left(1 + \frac{[I]}{K_i}\right) \tag{9-59}$$

The relationships between v and [S] and between $1/v$ and $1/[S]$ are as follows:

$$\frac{v}{[S]} = \frac{V_{max}}{K_m'} - \frac{v}{K_m'} \tag{9-60}$$

$$\frac{1}{v} = \frac{1}{V_{max}} + \frac{K_m}{V_{max}[S]}\left(1 + \frac{[I]}{K_i}\right) \tag{9-61}$$

A commonly used test for competitive inhibition is to plot $1/v$ vs $1/[S]$ (Eq. 9-61), both in the absence of inhibitor and in the presence of one or more fixed concentrations of I. The result, in each case, is a family of lines of varying slope (Fig. 9-10) that converge on one of the axes at the value $1/V_{max}$. We see that the maximum velocity is unchanged by the presence of inhibitor. If sufficient substrate is added, the enzyme will be saturated with substrate and the inhibitor cannot bind. The value of K_i can be calculated using Eq. 9-61 from the change in slope caused by addition of inhibitor.

The effect of a fixed concentration of a competitive inhibitor on a plot of v against log [S] (Fig. 9-11) is to shift the curve to the right, i.e., toward higher values of [S], but without any change in shape (or in the value of V_{max}).

Another plot, introduced by Dixon,[71] is that of $1/v$ versus [I] at two or more fixed substrate concentrations. The student should be able to demonstrate that this plot contains a family of straight lines that intersect at a point to the left of the origin with coordinates $[I] = -K_i$ and $1/v = 1/V_{max}$. This plot may fail to distinguish certain types of inhibition discussed in the next section.[72]

Competitive inhibition is extremely common and has great significance for metabolic control and for the effects of drugs and of poisons. Simple ions are often competitive inhibitors. Since many biochemical substances carry negative charges, anions such as Cl^-, HCO_3^-, HPO_4^{2-}, and acetate$^-$ frequently act as competitive inhibitors.

Two special classes of competitive inhibitors are characterized by **slow binding** and **slow, tight binding** to active sites.[73–77] Among the very tight-binding inhibitors, which may also be slow to dissociate from active sites, are transition state inhibitors discussed in Section D,1.

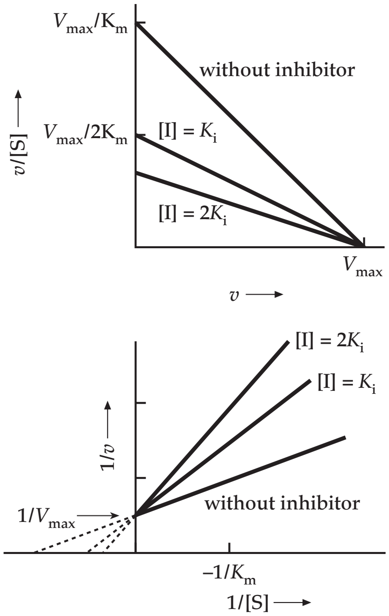


Figure 9-10 Effect of a competitive inhibitor on the Eadie–Hofstee plot (top) and on a double reciprocal plot (bottom). The apparent K_m (Eq. 9-59) is increased by increasing [I], but V_{max} is unchanged.

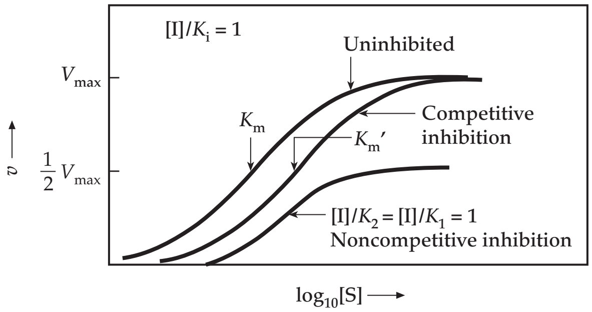


Figure 9-11 Plots of v vs. log [S] for competitive and noncompetitive inhibition.

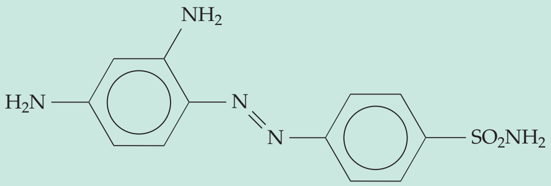
2. Noncompetitive Inhibition and Activation; Allosteric Sites

If an inhibitor binds not only to free enzyme but also to the enzyme substrate complex ES, inhibition is **noncompetitive**. In this case, S and I do not mutually exclude each other and both can be bound to the enzyme at the same time. Why does such an inhibitor slow an enzymatic reaction? In most instances, the structure of the inhibitor does not show a close similarity to that of substrate, which suggests that the binding of inhibitors is at an **allosteric site**, that is, at a site other than that of the substrate. The inhibition of the enzyme may result from a distortion of the three-dimensional structure of the enzyme which is caused by the binding of the inhibitor. This distortion may be transmitted to the active site even though the inhibitor binds far from that site. In some cases two distinctly different conformers of the protein may exist, one binding substrate well and the other binding inhibitor well (Fig. 9-12). In other instances the bound inhibitor may interfere with the catalytic action by partially overlapping the active site. In either case the ES complex reacts to give product in a normal way, but the ESI complex reacts more slowly or not at all.

Binding of a substance to an allosteric site sometimes has the effect of *increasing* the activity of an enzyme rather than inhibiting it. This may occur because the **activator** stabilizes the conformation that binds substrate best (Fig. 9-12). The quantitative treatment of such activation is similar to that of inhibition; allosteric inhibitors and activators are often considered together and are referred to as **modifiers** or

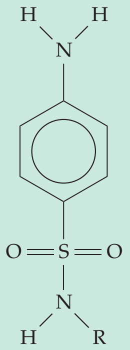
BOX 9-C THE SULFONAMIDES AS ANTIMETABOLITES

The development of the "sulfa drugs,"[a–c] derivatives of **sulfanilamide**, originated with studies of the staining of protozoal parasites by Paul Erhlich. In 1932 it was shown that the red dye 2,4-diaminoazobenzene-4'-sulfonamide (Prontosil) dramatically cured systemic infections by gram-positive bacteria. Subsequent studies revealed that bacteria converted the azo dye to sulfanilamide, a compound with strong bacteriostatic activity; that is, it inhibited bacterial growth without killing the bacteria. Although sulfanilamide had been used in large quantities since 1908 as an intermediate in synthesis of dyes, its potential as an antibacterial agent had not been recognized.

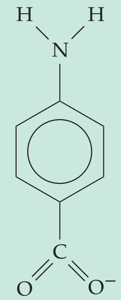


Prontosil

In 1935, D. D. Woods found that the growth inhibition of sulfanilamide was reversed by yeast extract.[d] From this source, in 1940 he isolated ***p*-aminobenzoic acid** and demonstrated that the inhibitory effect of 3×10^{-4} M sulfanilamide was overcome by 6×10^{-8} M *p*-aminobenzoate. The relationship between the two compounds was shown to be strictly competitive. If the sulfanilamide concentration was doubled, twice as much *p*-aminobenzoate was required to reverse the inhibition as before. These facts led to the formulation by Woods and by P. Fildes (in 1940) of the **antimetabolite theory**.[d,e] It was proposed that *p*-aminobenzoate was needed by bacteria and that sulfanilamide competed for a site on an enzyme designed to act on *p*-aminobenzoate. We now know that the idea was correct and that the enzyme on which the competition occurs catalyzes the synthesis of dihydropteroic acid (Fig. 25-19), a precursor to folic acid.



Sulfonamides
R=H in sulfanilamide



p-Aminobenzoate

While sulfanilamide itself is somewhat toxic, a variety of related drugs of outstanding value have been developed. Over 10,000 sulfonamides and related compounds have been tested for antibacterial action.

---

[a] Bardos, T. J. (1974) *Top. Curr. Chem.* **52**, 63–98

[b] Gale, E. F., Cundliffe, E., Reynolds, P. E., Richmond, M. H., and Waring, M. J. (1972) *The Molecular Basis of Antibiotic Action*, Wiley, New York

[c] Shepherd, R. G. (1970) in *Medicinal Chemistry*, 3rd ed. (Burger, A., ed), pp. 255–304, Wiley (Interscience), New York

[d] Woods, D. D. (1940) *Brit. J. Exp. Pathol.* **21**, 74–90

[e] Fildes, P. (1940) *Lancet* **1**, 955–957

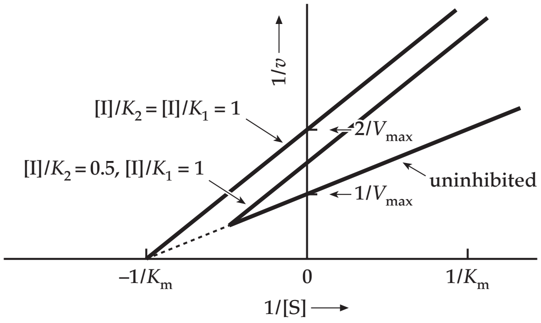


Figure 9-12 Double reciprocal plots for two cases of noncompetitive inhibition.

effectors. A general scheme[78,79] is given by Eq. 9-62.

$$\begin{tikzcd}
\mathrm{E + S} \arrow[r, shift left, "K_{ds}"] & \mathrm{ES} \arrow[l, shift left] \arrow[r, "k_3"] & \text{product} \\
+ & + & \\
\text{M (modifier)} \arrow[d, shift right] & \mathrm{M} \arrow[d, shift right] & \\
\mathrm{EM + S} \arrow[u, shift right, "K_1"'] \arrow[r, shift left, "K'_{ds}"] & \mathrm{ESM} \arrow[u, shift right, "K_2"'] \arrow[l, shift left] \arrow[r, "k_4"] & \text{product}
\end{tikzcd} \qquad (9\text{-}62)$$

Here K_1 and K_2 are equilibrium constants for dissociation of M from EM and ESM, respectively, while K_{ds} and K'_{ds} are the dissociation constant of ES and of ESM (to S and EM), respectively. Notice that, K'_{ds} is not independent of the others (Eq. 9-63):

$$K'_{ds} = K_{ds}K_2 / K_1 \qquad (9\text{-}63)$$

We have already considered competitive inhibition which is obtained when $K_2 = 0$ (and therefore $K'_{ds} = 0$). For this case, M is always an inhibitor and no activation is possible. Notice that the inhibition will appear competitive even if M binds at an allosteric site as in Fig. 9-13 or if the inhibited form does not react at all with substrate. Noncompetitive inhibition will be observed if ESM is formed but does not react, i.e., if $k_4 = 0$. Then the rate equation in reciprocal form will be given by Eq. 9-64.

$$\frac{1}{v} = \frac{1}{V_{\max}}\left(1 + \frac{[\mathrm{I}]}{K_2}\right) + \frac{K_m}{V_{\max}[\mathrm{S}]}\left(1 + \frac{[\mathrm{I}]}{K_1}\right) \qquad (9\text{-}64)$$

We see that $1/V_{\max}$ is multiplied by a term containing [I] and K_2. It is a characteristic of noncompetitive inhibition that the maximum velocity is decreased from that observed in the absence of the inhibitor. Also, no matter how high the substrate concentration, inhibition cannot be reversed completely.

Figure 9-12 shows a plot of $1/v$ against 1/[S] at a series of fixed values of [I]. For the case that $K_1 = K_2$ **(classical noncompetitive inhibition)**, a family of reciprocal plots that intersect on the horizontal axis at a value of $-1/K_m$ is obtained. On the other hand, if K_1 and K_2 differ (the general case of noncompetitive inhibition), the family of curves intersect at some other point to the left of the vertical axis and, depending upon the relative values of K_1 and K_2, either above or below the horizontal axis. The example illustrated is for $K_2 = 0.5K_1$; that is, for the binding of M to ES being twice as strong as that to E.

Figure 9-11 shows inhibition data for both the noncompetitive and the competitive cases plotted vs log [S]. The shift of the midpoint to the right in each case reflects the tendency of the inhibitor to exclude the substrate from binding, while the lowered value of the maximum velocity in the case of noncompetitive inhibition results from the failure of the substrate to completely displace the inhibitor from the enzyme

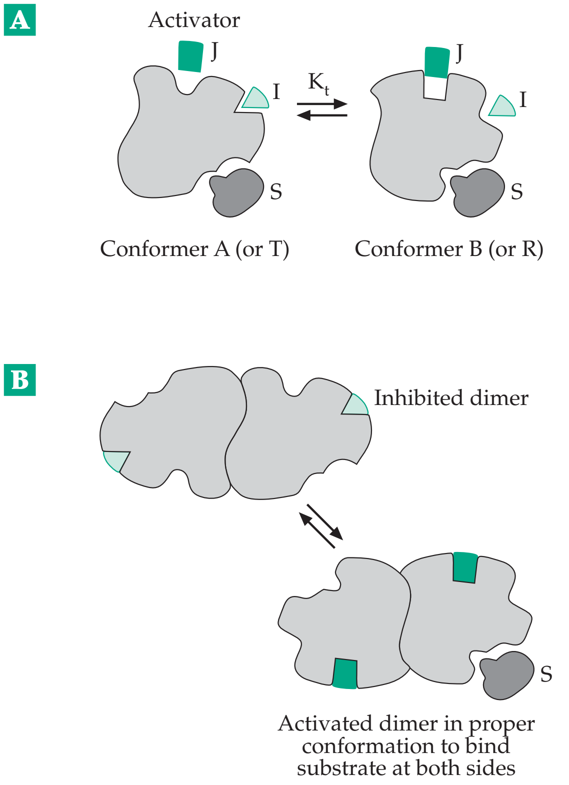


Figure 9-13 (A) An enzyme with binding sites for allosteric inhibitor I and activator J. Conformer A binds inhibitor I strongly but has little affinity for activator J or for substrate S. Conformer B binds S and catalyzes its reaction. It also binds activator J whose presence tends to lock the enzyme in the "on" conformation B. Conformers A and B are designated T and R in the MWC model of Monod, Wyman, and Changeux.[80] (B) Inhibited and activated dimeric enzymes.

even as [S] becomes very high. If an inhibitor binds only to ES and not to E, i.e., $K_1 = 0$, a family of parallel double-reciprocal plots of $1/v$ vs $1/[S]$ will be obtained. This case is referred to as **uncompetitive inhibition**. Multisubstrate enzymes with either ordered sequential or ping-pong mechanisms also often give parallel line plots with inhibitors.

Uncompetitive inhibitors of liver alcohol dehydrogenase (Chapter 15) could be used to treat cases of poisoning by methanol or ethylene glycol.[81–83] The aim is to prevent rapid oxidation to the toxic acids HCOOH and $HOCH_2COOH$, which lower blood pH, while the alcohols are excreted. Uncompetitive inhibitors have an advantage over competitive inhibitors as therapeutic agents in that the inhibition is not overcome when the substrate concentration is saturating.[84]

If rate constant k_4 of Eq. 9-62 is not zero we may observe either inhibition or activation. If $k_3 = k_4$ the effect of a modifier will be to alter the apparent K_m by either increasing it (inhibition) or decreasing it (activation). The maximum velocity will be unchanged. Monod *et al.*[80] referred to enzymes showing such behavior as **K systems**. On the other hand, if $K_1 = K_2$ and k_3 differs from k_4 we have a purely **V system**. In the general case a modifier affects both the apparent K_m and $V_{\max}$. Although in the foregoing discussion it has been assumed that activators bind at allosteric sites, as was pointed out in Section A,11, ions and other small molecules (substrate surrogates) that act as competitive inhibitors at high concentrations may be activators at low concentrations.[40]

Activation of enzymes by specific metallic ions is often observed. In many instances the metallic ion is properly regarded as a *second substrate* which must bind along with the first substrate before reaction can occur. Alternatively, the complex of the organic substrate with the metal ion can be considered the "true substrate." Thus, many enzymes act upon the magnesium complex of ATP (Chapter 12). The enzymes can either be regarded as three-substrate enzymes requiring $Mg^{2+} + ATP^{4-}$ and another substrate or as two-substrate enzymes acting upon $ATPMg^{2-}$ and a second substrate.

3. Inhibitors in the Study of Mechanisms

A substrate analog will frequently inhibit only one of the two forms of a multisubstrate enzyme with a ping-pong mechanism.[1,72] Reciprocal plots made for various inhibitor concentrations consist of a family of parallel lines reminiscent of uncompetitive inhibition. Observation of such parallel line plots can support a ping-pong mechanism for an enzyme but cannot prove it because in some cases parallel lines are observed for inhibition of enzymes acting by an ordered sequential mechanism. The following question arises naturally for any ordered bimolecular reaction (Eq. 9-43): Of the two substrates required by the enzymes, which one binds to the enzyme first? If the concentration of one substrate is kept constant while varying concentrations of an inhibitory analog of that substrate are added and $1/v$ is plotted against the reciprocal of the concentration of the other substrate, parallel lines are obtained if, and only if, the substrate of fixed concentration is B, the substrate is adding second in the binding sequence, *and* if I is its analog. The substrate binding first (A) is the one whose concentration was varied in the experiment.

Product inhibition (Section A,12) can also provide information about mechanisms. For example, if $1/v$ is plotted against $1/[A]$ in the presence and absence of the product Q, the product will be found to compete with A and to give a typical family of lines for competitive inhibition. On the other hand, a plot of $1/v$ vs $1/[B]$ in the presence and absence of Q will indicate noncompetitive inhibition if the binding of substrates is ordered (Eq. 9-43). In other words, only the A–Q pair of substrates are competitive. Product inhibition is also observed with enzymes having ping-pong kinetics (Eq. 9-47) as a result of formation of nonproductive complexes.

4. Allosteric Effectors in the Regulation of Enzyme Activity

The binding of a substance at an allosteric site with the induction of a conformational change forms the basis for many aspects of regulation. The term **allostery** (allosterism) usually refers to the effects of allosteric modifiers, which may be either inhibitors or activators, on oligomeric enzymes. However, as we have already seen (Eq. 9-62), monomeric enzymes may also be subject to allosteric regulation by modifiers. Consider a monomer that contains binding sites for substrate, inhibitor, and activator and which exists in conformations A and B as in Fig. 9-13. Let us assume (see Eq. 7-31) that conformer B binds both substrate and activator well but that it binds inhibitor poorly or not at all. On the other hand, A binds inhibitor well but binds substrate and activator poorly. This simple combination of two conformers with different binding properties provides a means by which enzymes can be turned "on" or "off" in response to changing conditions.

If an inhibitory substance builds up to a high concentration within a cell, it binds to conformer A; if the inhibitor concentration is high enough, virtually all of the enzyme will be locked in the inactive conformation A. The enzyme will be turned off or at least reduced to a low activity. On the other hand, in the presence of a high concentration of activator the enzyme will be turned on because it is locked in the B conformation. The relative concentrations of inhibitor, activator, and substrate within a cell at any given time will determine what fraction of the enzyme is in active conformation B. It is this interplay of inhibitory and activating effects that provides the basis for much of the regulation of cell chemistry.[80,85]

The effects of inhibitors or activators on the kinetics of the monomeric enzyme of Fig. 9-13 can be described by Eq. 9-62 to 9-64. Separate terms for both inhibition and activation can be included. The equilibrium between the two conformers can also be indicated explicitly according to Eq. 7-30. However, for monomeric enzymes it is usually not profitable to try to separate the two constants K_t and K_{BX} which describe the conformational change and binding of substrate or activator, respectively, in Eq. 7-30.

Most intracellular enzymes are oligomeric, and the binding of allosteric effectors leads to additional interesting effects. Binding constants or dissociation constants must be defined for both inhibitor and activator to both conformers A and B. Since all species must be taken into account in the mass balance, the equations are complex. However, the Monod–Wyman–Changeux (MWC) model (Chapter 7) gives a relatively simple picture. The saturation curve for an oligomeric enzyme following this model may be derived from Eq. 7-39 and is given by the following expression:

$$\mathrm{Y} = \frac{Lc\alpha(1 + c\alpha)^{n-1} + \alpha(1 + \alpha)^{n-1}}{L(1 + c\alpha)^n + (1 + \alpha)^n} \tag{9-65}$$

Here, L is the **allosteric constant** which is given (for a dimer) by Eq. 7-36. The constant c is the ratio of dissociation constants K_{BS} and K_{AS} for the two conformers:

$$c = K_{BS} / K_{AS} \tag{9-66}$$

Notice that in this chapter *dissociation* constants (K_d) of ES complexes are being used, whereas the equations of Chapter 7 are all written in terms of association constant (K_f). The parameter α is defined as follows, where K_{BS} equals $1/K_{BX}$ of Chapter 7.

$$\alpha = [\mathrm{S}] / K_{BS}) \tag{9-67}$$

To take account of effects of inhibitor and activator the *ratios* of dissociation constants of I from BI and AI and of activator J from BJ and AJ are defined as in Eq. 9-68. Likewise, "normalized concentrations" of I and J are defined (Eq. 9-69) as β and γ, respectively.

$$d = \frac{K_{\mathrm{BI}}}{K_{\mathrm{AI}}} > 1 \qquad e = \frac{K_{\mathrm{BJ}}}{K_{\mathrm{AJ}}} < 1 \tag{9-68}$$

$$\beta = [\mathrm{I}]/K_{\mathrm{AI}} \qquad \gamma = [\mathrm{J}]K_{\mathrm{BJ}} \tag{9-69}$$

According to the MWC model, in the presence of inhibitor and activator at normalized concentrations β and γ an enzyme will still follow Eq. 9-65, but the allosteric constant L will be replaced by an apparent allosteric constant L' (Eq. 9-70).[86] Figure 9-14 shows plots of Y vs. log α for two different values of L' for a tetramer with a specific value assumed for c. In both

$$L' = L\left[\frac{(1 + \beta d)(1 + \gamma e)}{(1 + \beta)(1 + \gamma)}\right]^n \tag{9-70}$$

cases, Y approaches 1 as log α increases, but since we are dealing with noncompetitive inhibition at high values of L', much of the enzyme will be in the T(A) conformation at saturation.

Noncompetitive inhibition cannot be completely reversed by very high substrate concentrations. Monod *et al*. defined for an allosteric enzyme a **function of state $\bar{\mathbf{R}}$** (Eq. 9-71) which is the fraction of total enzyme in the $\bar{R}$ (B) conformation:

$$\begin{aligned}\bar{R} &= \frac{(1 + \alpha)^n}{L(1 + c\alpha)^n + (1 + \alpha)^n} \\ &= \text{function of state}\end{aligned} \tag{9-71}$$

In a K system (Section B,2) it is the value of $\bar{R}$ that determines the velocity with which an enzyme reacts. Figure 9-14 also shows $\bar{R}$ as a function of log [α]. Note that when L' is low $\bar{R}$ does not approach zero even when [S] → 0. In other words, the enzyme is never completely turned off, just as when L' is high the enzyme is never completely turned on.

Figure 9-14 may be compared with Fig. 9-11 which shows similar curves for noncompetitive inhibition of a monomeric enzyme. The significant difference between the two figures is that saturation of the oligomeric enzyme occurs over a narrower concentration range than does that of the monomer, i.e., saturation of the oligomeric enzyme, especially in the presence of inhibitor, is *cooperative*. Note that cooperative binding of substrate requires that the free enzyme be largely in conformation T (A), as it is in the presence of an inhibitor. Allosteric interactions between two identical molecules, whether of substrate or of effector, are described by Monod *et al*.[80] as **homotropic interactions**. Such interactions lead to cooperativity or anticooperativity in binding. Allosteric interactions between two different molecules, e.g., a substrate and an activator are designated **heterotropic**.

For many enzymes the MWC model is unrealistically simple. The more general treatment of binding equilibria given in Chapter 7 may be applicable. However, in addition to K systems there are V systems in which a conformational change alters the maximum velocity (see Eq. 9-62)[87] and sometimes both substrate

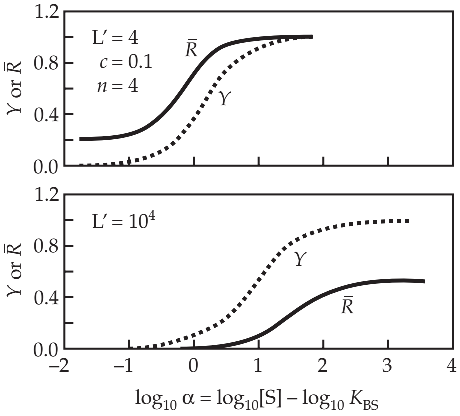
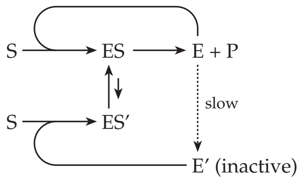


Figure 9-14 Fractional saturation Y and "function of state" $\bar{R}$ for hypothetical tetrameric enzymes following the MWC model. Curves are calculated for two different values of the apparent allosteric constant *L*′ (Eq. 9-70) and for *c* = 0.1 (Eq. 9-66). After Rubin and Changeux.[86]

affinity and maximum velocity.

The fact that data can be fitted to an equation is not proof that a mathematical model is correct; other models may predict the results just as well. For example, Jenkins has shown that the presence of ions that act as substrate surrogates (Section A,11) can produce cooperative or anticooperative binding curves for substrates.[40,88] Rabin[89] suggested the following explanation of cooperativity for a monomomeric enzyme with a single substrate-binding site. In its active conformation E the enzyme reacts with substrate rapidly and the ES complex yields product rapidly as in the upper loop of Eq. 9-72. However, a slow conformational change interconverts E and E′, a less active form with much lower affinity for substrate.



(9-72)

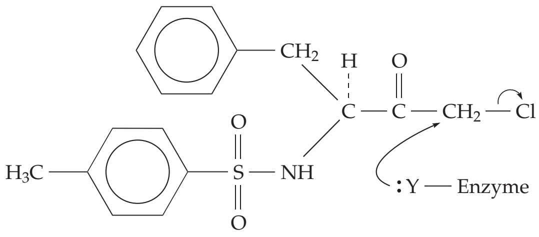
At the same time, the complex ES′, if formed, can equilibriate with ES and thereby alter the conformational state of the protein. At low substrate concentrations E′ will predominate and the enzymatic activity will be low. At high substrate concentrations most E will exist as ES and after release of product will tend to remain in the active conformation long enough to bind and act on another substrate molecule. This and other kinetic models of cooperativity are discussed by Newsholme and Start.[90]

The physiological significance of cooperative binding of substrates to enzymes is analogous to that of cooperative binding of oxygen by hemoglobin which provides for more efficient release of oxygen to tissues. However, in the presence of excess activator an enzyme is locked in the *R* (B) conformation and no cooperativity is seen in the binding of substrate. In this case, each binding site behaves independently. On the other hand, *there will be strong cooperativity in the binding of the activator*. The result is that control of the enzyme is sensitive to a higher power than the first of the activator concentration. Likewise, the turning off of the enzyme is more sensitive to inhibitor concentration as a result of cooperative binding of the latter. It seems likely that the evolution of oligomeric enzymes is at least partly a result of the greater efficiency of control mechanisms based on cooperative binding of effectors.

5. Irreversible Inhibition of Enzymes

Some of our most effective drugs, such as penicillin (Box 20-G) are irreversible inhibitors of specific enzymes. Such inhibitors are also of practical importance to biochemists who wish to inhibit specific enzymes such as proteases that might otherwise destroy proteins they are studying. Transition state analogs (Section E,1) and slow-binding inhibitors (Section B,1) may appear to inactivate an enzyme irreversibly because they bind tightly and dissociate slowly. However, true irreversible inhibitors, such as oxidizing agents or alkylating agents, cause a more permanent chemical modification, usually at or near the active site. Two groups of irreversible inhibitors are of special interest: affinity labeling agents and enzyme-activated inhibitors.

Affinity labeling agents (active-site directed inactivating reagents) have two essential properties: *a high affinity for the active site* of the specific enzyme to be inhibited and *the presence of a chemically highly reactive group* which can attack a functional group in a protein. A good example is provided by derivatives of chloromethylketones, which are reactive alkylating agents. One of these, *N*-tosylphenylalanine chloromethyl ketone (TPCK), is a potent inhibitor of chymotrypsin. In addition to the chloromethyl ketone group it contains

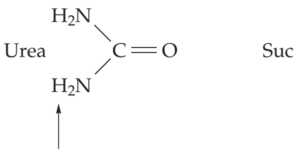


L-*N*-tosylphenylalanine chloromethylketone (TPCK) the phenylalanine side chain, which helps direct it into the substrate binding site of chymotrypsin, and a sulfonamide linkage that mimics the normal amide linkages of the substrate. The group Y, which becomes alkylated in an S_N2-like displacement of a chloride ion (see Eq. 9-76) is thought to be His 57 of the active site (see Fig. 12-10). For additional discussion see Walsh,[91] Kyte,[67] and Plapp.[91a]

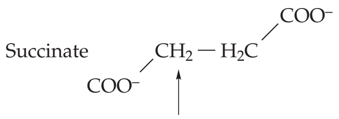
Enzyme-activated inhibitors (also called suicide substrates, k_{cat} inhibitors, or mechanism-based inhibitors) are chemically inert until they enter the active site of their target enzymes. Then, by passing through at least some of the normal stages of the catalytic action of the enzyme, they are converted to reactive intermediates that can become irreversibly bound to an enzyme.[92,93] For example, a halogen atom (F, Cl, Br) together with a proton may be eliminated from an intermediate to give an unsaturated compound to which a nucleophilic side chain of the protein adds. Several of these inhibitors are discussed in later chapters. Because of their high specificity many enzyme-activated inhibitors are potential drugs. One of them, α-difluoromethylornithine (Box 14-C), is said to bind covalently to only one protein in the body of a rat, namely, ornithine decarboxylase, the target enzyme for the drug.

C. The Specificity of Enzymatic Action

Enzymes are usually impressively specific in their action. The specificity toward substrate is sometimes almost absolute. For many years urea was believed to be the *only* substrate for the enzyme **urease** and succinate the only substrate for **succinate dehydrogenase**. Even after much searching for other substrates, only one or two closely related compounds could be found that were acted on at all. In other cases enzymes can use a class of compounds as substrates. For example, the **D-amino acid oxidase** of kidney oxidizes a variety of D-amino acids but does not touch L-amino acids.



One H can be replaced by —OH and some substrate activity remains



One H can be replaced by Cl with retention of substrate activity and by —CH_3 with partial retention of activity

Almost as impressive as the substrate specificity of enzymes is the specificity for a given type of reaction. Many substrates are capable of undergoing a variety of different chemical reactions, either unimolecular or with water or some other compound present in the cell. The enzyme catalyzes only one of these reactions. Although side reactions may occur to a small extent, the most impressive thing in comparing an enzyme-catalyzed reaction with an uncatalyzed organic reaction is that the latter often produces large amounts of side reaction products, but the enzymatic reaction does not.

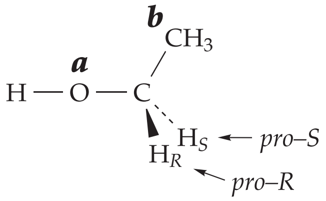
1. Complementarity of Substrate and Enzyme Surfaces

Impressed by the specificity of enzymatic action, biochemists early adopted a "lock-and-key" theory which stated that for a reaction to occur the substrate must fit into an active site precisely. Modern experiments have amply verified the idea. A vast amount of kinetic data on families of substrates and related competitive inhibitors support the idea and numerous X-ray structures of enzymes with bound inhibitors or with very slow substrates have given visual evidence of the reality of the lock-and-key concept. Directed mutation of genes of many enzymes of known three-dimensional structure has provided additional proof.

As anticipated, hydrophobic groups of substrates or inhibitors usually contact hydrophobic regions of the protein and the fit is tight. For example, the active site of chymotrypsin contains a "hydrophobic pocket" designed to hold a large hydrophobic side chain, thus providing the specificity observed with this enzyme (Table 3-2). Likewise, polar groups of substrates contact polar groups of the enzyme. The interactions are complementary, positive charges fitting against negative and with correctly formed hydrogen bonds. Trypsin, whose structure is similar to that of chymotrypsin, exerts its specificity for a positively charged side chain next to the bond cleaved (Table 3-2) by virtue of the presence of a negatively charged carboxylate at the bottom of the hydrophobic pocket. Several C=O and N–H groups of the peptide linkages in the substrate form hydrogen bonds to the edge of a β sheet in the protein, in effect making the substrate an added β strand in the sheet (Fig. 12-10). Aspartate aminotransferase acts on the dicarboxylic amino acids glutamate and aspartate. A pair of arginine side chains bind the two carboxylates of the substrate while the $-NH_3^+$ of the substrate is attracted to a negatively charged group in the coenzyme pyridoxal phosphate, which is also present in the active site (Eq. 14-39, Fig. 14-10).

2. Stereospecificity and Prochiral Centers

Most enzymes possess an infallible ability to recognize the difference between the right side and the left side of an organic substrate even when the latter has perfect bilateral symmetry. In fact, this ability is limited to **prochiral centers** of molecules and is a natural consequence of their reaction with the chiral enzyme.[94–98] Consider carbon atom number 1 of ethanol. The two attached hydrogen atoms are chemically identical and would react identically with a nonchiral reagent. Nevertheless, these atoms are no more equivalent stereochemically than are your right and your left arms. We say that the molecule has a prochiral center at C-1. A prochiral center on a tetrahedrally bonded carbon atom always contains two identical atoms or groups but a total of three different kinds of atoms or groups. If the priority of this H is elevated, e.g., by substitution of ^2H for ^1H the configuration would be *R* (the priority of this ^2H would become **c**).



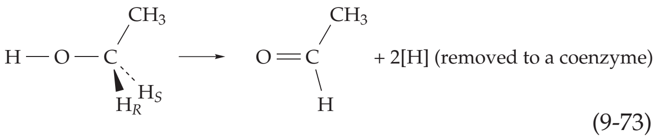
BOX 9-D RECEPTORS, AGONISTS, AND ANTAGONISTS

Inhibition of enzymes provides the basis for many of the effects of antibiotics and other chemotherapeutic substances (e.g., see Box 9-B). However, some drugs act on cell surface **receptors** which have ordinarily not been regarded as enzymes. According to receptor theory, developed around 1937, structurally similar drugs often elicit similar responses because they bind to the same receptor. A receptor may normally bind a hormone, neurotransmitter, or other metabolite whose geometry is partially shared by a drug. Binding of drugs of one class, termed **agonists** in the pharmacological literature, to an appropriate receptor triggers a response in a cell, similar to that of a hormone. On the other hand, compounds of related structure often act as **antagonists**, binding to receptor but failing to elicit a response. Agonist and antagonist often act in a strictly competitive fashion as in competitive inhibition of enzyme action.

We know now that many receptors *are* enzymes, some of which may act quite slowly. The active site may sometimes be far from the receptor binding site and sometimes in a separate subunit. The receptor can be viewed as an allosteric effector which binds at a distant site or as a ligand for a regulatory subunit of the enzyme complex. Alternatively, the active site may be viewed as the site for relaying a signal received from the hormone or other agonist.

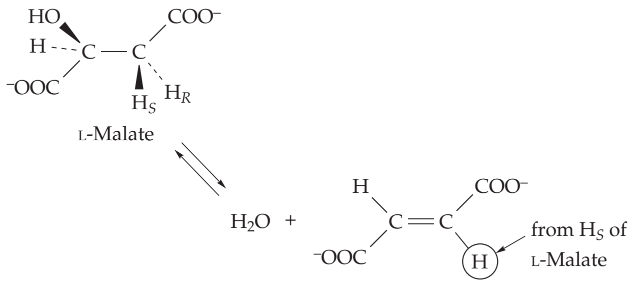
The two hydrogen atoms on C-1 of ethanol are called **enantiotopic** because replacement of one or the other by a fourth different kind of atom or group would produce a pair of enantiomers (page 41). This fact suggests a way of naming the positions occupied by two enantiotopic atoms or groups. We first assign priorities to all of the atoms or groups attached to the central carbon atom according to the *RS* system. Now, we ask whether the configuration will be *R* or *S* when the priority of one of the two identical groups is raised, e.g., by substitution of one of the hydrogen atoms by deuterium. If the configuration becomes *R*, that group occupies a *pro-R* position; if the configuration becomes *S*, it occupies a *pro-S* position. Referring to the preceding diagram, it is easy to see (by viewing down the bond to the group of lowest priority and applying the usual rule for determining configuration) that if the *pro-R* hydrogen (H_R) is replaced by deuterium (^2H), the configuration will be *R*. Conversely, replacement of H_S by deuterium will lead to the *S* configuration.

When ethanol is oxidized by the action of **alcohol dehydrogenase** (Eq. 9-73), only the *pro-R* hydrogen atom is removed. If the reaction is reversed in such a way that deuterium is introduced into ethanol from the reduced coenzyme the optically active *R*-2-deuterioethanol is formed. The ability of an enzyme to recognize a single hydrogen of a pair of hydrogens on a CH_2 group was at first a surprise to many biochemists.[99] However, it is a natural result of the complementarity of enzyme and substrate surfaces, just as the fit of a shoe is determined by the complementarity of surfaces of foot and shoe. Only a chiral catalyst can have this ability.



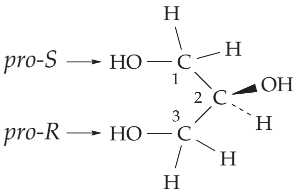
(9-73)

Another example is provided by malic acid, a chiral molecule which also contains a prochiral center (see Eq. 9-74). In this case replacement of the *pro-R* or *pro-S* hydrogen atom by another atom or group would yield a pair of diastereoisomers rather than enantiomers. Therefore, these hydrogen atoms are **diastereotopic**. When L-malic acid is dehydrated by fumarate hydratase (Chapter 13) the hydrogen in the *pro-R* position is removed but that in the *pro-S* position is not touched. This can be demonstrated by allowing the dehydration product, fumarate, to be hydrated to malate in ^2H_2O (Eq. 9-74). The malate formed contains deuterium in the *pro-R* position. If this malate is now isolated and placed with another portion of enzyme in H_2O, the deuterium is removed cleanly. The fumarate produced contains no deuterium.



(9-74)

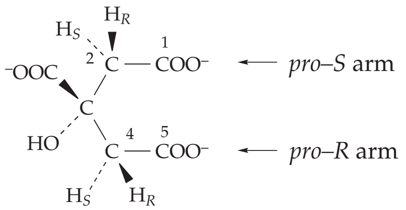
Stereochemical numbering. In some prochiral molecules, such as glycerol, the two ends of the main carbon chain form identical groups. Since the two ends are distinguishable to an enzyme, it is important to decide which should be labeled C-1 and which C-3. Hirschmann proposed a stereochemical numbering system[100] in which the carbons are numbered beginning with the end of the chain that occupies the *pro-S* position.



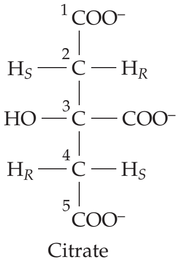
Stereochemical numbering for a symmetric prochiral molecule. Carbons are numbered beginning with the end of the chain that occupies the *pro-S* position.

In this numbering system derivatives of the parent prochiral compound are given the prefix *sn*-. Thus, glycerol phosphate, used by cells to construct phospholipids, usually bears a phosphate group on the $-CH_2OH$ in the *pro-R* position of glycerol and is therefore *sn*-3-glycerol phosphate.

The **citrate ion**, a very important prochiral metabolic intermediate, has three prochiral centers at C-2, C-3, and C-4, respectively. That at C-3 distinguishes the *pro-R* and *pro-S* arms and determines the stereochemical numbering. Citrate containing ^{14}C in the *sn*-1 position is called *sn*-citrate[1-^{14}C] and is the form of labeled citrate that is synthesized in living cells from oxaloacetate and [1-^{14}C]acetyl coenzyme A (see Fig. 10-6). The first step in the further metabolism of citrate is the elimination of the –OH group from C-3 together with the H_R proton from C-4 through the action of the enzyme **aconitate hydratase** (aconitase). In this case the proton at C-4 (in the *pro-R* arm) is selected rather than that at C-2.

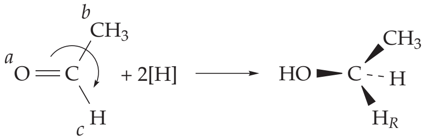


Citrate ion numbered according to the stereochemical system



Drawn using the Fisher convention

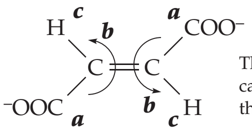
Trigonal prochiral centers. Planar trigonal atoms, such as those of aldehydes, ketones, and alkenes, are also prochiral if they are attached to three different kinds of atoms or groups. The two faces are enantiotopic or diastereotopic, if another chiral center is present. Hanson[94,94a] proposed that the faces be named as indicated in Eq. 9-75. The trigonal atom is viewed from one side and the three groups surrounding the carbon atom are given priorities, *a*, *b*, and *c*, as in the *RS* system (Chapter 2). If the sequence *a*, *b*, and *c* of priorites is clockwise, the face toward the reader is *re* (*rectus*); if counterclockwise, it is *si* (*sinister*). Priorities can often be assigned on the basis of the atomic numbers of the first atoms in the three groups (as in the example shown)



re (or *Re*) face is toward the reader
si (or *Si*) face is behind the page (9-75)

but, if necessary, replica atoms must be added as described in Chapter 2. Replica atoms attached to an originally trigonal carbon are ignored completely, but those attached to O or N of C=O or C=N may be required to establish the priorities of the groups around the carbon atom. For C=O and C=N the faces of the O and N atoms are taken to be the same as that of the attached trigonal carbon atom. In C=C the two carbons may have their *re* faces either on the same side or on opposite sides of the group.

Notice that addition of an atom of ^{2}H to the *re* face of acetaldehyde would give *R*-deuteroethanol (Eq. 9-75; reverse of Eq. 9-73). This is the reaction catalyzed by alcohol dehydrogenase. Addition of ^{1}H to the *re* face places the entering hydrogen in the *pro-R* position. Addition to the *si* face would place it in the *pro-S* position. Fumarate has two trigonal carbon atoms. The *si* faces of both are toward the viewer in the following structure (also shown in Eq. 9-74). Referring to Eq. 9-74, we see that an HO^- ion from water adds to the *si* face of one carbon atom of fumarate to give *s*-malate (L-malate). At the same time, a proton combines with the *re* face of the adjacent unsaturated carbon atom to enter the *pro-R* position.



Fumarate

The *si* faces of both trigonal carbon atoms are toward the viewer.

3. Induced Fit and Conformational Changes

Results of many X-ray studies indicate that the lock-and-key picture of enzyme action must be modified. If an enzyme is a "lock" and the substrate the "key," *the entrance of the key into the lock often induces a conformational change in the protein*. Binding of substrates may be imperfect in the initially formed ES complex but may be more nearly perfect a few nanoseconds or microseconds later as the protein readjusts its structure to accommodate the substrate. The substrate has induced a fit. For example, when an amino acid substrate binds to aspartate aminotransferase one whole domain of the enzyme moves inward, packing hydrophobic side chains of the protein against the substrate (Chapter 14). This strengthens the electrostatic interactions between the ion pairs that orient the substrate and align it for reaction. Similar conformational changes have been observed for citrate synthase, glycogen phosphorylase, various kinases (Chapter 12), alcohol dehydrogenase (Chapter 15), and a growing list of other enzymes. Accompanying changes in circular dichroism, ultraviolet spectra, and sedimentation constants are often observed.

For many other enzymes the observed conformational changes are subtle. A single loop of polypeptide chain or even of a side chain moves to cover the bound substrate. With trypsin and other serine proteases a flexible segment of the peptide chain becomes immobilized and forms more hydrogen bonds after substrate binds than before. It is probably quite unusual for a major unfolding and refolding of parts of a protein to take place. The term induced fit usually refers to substrate binding, but as substrates are interconverted within active sites successive changes in geometry and in charge distribution occur. Small conformational changes may be required at several stages of an enzymatic reaction to ensure that complementarity of substrate and enzyme is preserved.[101] In line with this idea are the facts that proteins are less tightly packed at active sites than in other parts of the molecules, and that active sites often lie between domains and usually are formed by several loops of the peptide chain (Chapter 2).[102]

Although conformational changes allow proteins to maintain good complementarity with substrates, it does not follow that substrates are therefore bound very tightly. This is easiest to understand for reversible reactions in which a substrate is the product of the reverse reaction. Not only binding but also the rate of release of product must be rapid. Very tight binding would retard release and cause product inhibition. Furthermore, as we can see from Eq. 9-37, the velocity of catalysis is the product of [S][E] x k_{cat} / K_m, where [E] is, *free enzyme*. Catalysis can be made more efficient in two ways: by increasing k_{cat} or by decreasing K_m, that is, by tighter binding. However, tighter binding will be advantageous only if K_m is not so low that the enzyme is approaching saturation with substrate. Otherwise, [E] will fall as K_m is lowered and little increase in rate will be observed.[2]

In fact, it appears that enzymes have evolved to have high values of k_{cat} and *high* values of K_m, that is, *weak binding* of substrates.[2] Frequently, in one conformational state of a protein the active site is open, with solvent molecules or substrate surrogates present, while in a second state of nearly equal energy it is closed around the substrate. The functional groups of the enzyme's active site can be bound either to external ions or to ionic groups of substrates and either to water or to hydrogen-bonding groups of substrates. In some cases two ionic groups in the protein may pair with each other in the open conformation and with ionic groups of the substrate in the closed conformation. Thus, the energy changes accompanying the conformational changes can be small but very good complementarity can exist in the ES complex, an important factor in establishing specificity.

4. Specificity and k_{cat}

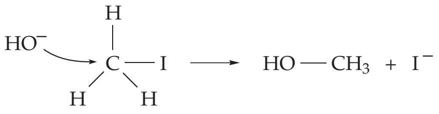
Enzyme specificity is often observed not only in binding but also in the rate at which ES is converted to products. Thus, it is the values of k_{cat} / K_m that determine specificity. Good examples are provided by chymotrypsin and related serine proteases (Chapter 12),[2] for which substrates with the shortest chains are often bound as well as those with longer chains but react more slowly. For example, *N*-acetylphenylalanine amide binds to chymotrypsin about as well as does the longer *N*-acetylphenylalanylalanine amide but reacts only 1/47 as fast.[2] One might anticipate that increasing the length of the substrate would make it bind more tightly because of the greater number of contacts between substrate and enzyme. It has often been suggested that the reason that this does *not* happen is that the binding of the longer, more specific substrates distorts the enzyme and that *the binding energy is now stored in the enzyme. It is as if the enzyme contained an internal spring which would be compressed when the substrate binds.* This would keep the binding weak but the energy in the spring might then be used to increase the velocity.

5. Proofreading

Although enzymes may be very specific they do make mistakes. This is of particular concern for processes such as protein synthesis in which the correct amino acid is placed at each position in the sequence with an error rate that has been estimated for *E. coli* as only 1 in 10^4. It would be impossible for an enzyme designed to attach valine to its specific transfer RNA to avoid attaching the smaller alanine if discrimination between the two were based solely on the Gibbs energy differences of binding.[103] However, it would be easier for an enzyme to exclude *larger* amino acids. This problem may be resolved by use of multistep screening.[104,105] For example, isoleucyl-tRNA synthetase (Chapter 29) does occasionally attach the smaller valine to the specific $\text{tRNA},^{\text{Ile}}$ but when it does the enzyme in a "proofreading and editing" step hydrolyzes off the incorrect amino acid. The active site for this hydrolyzing activity, whether at a different place on the enzyme surface or created by a conformational change, may be able to exclude sterically the larger isoleucyl residue while acting on the valyl-tRNA. This editing mechanism for isoleucyl-tRNA synthetase was demonstrated directly in 1998 by X-ray crystallography on complexes of the enzyme with L-isoleucine and L-valine. Both substrates fit into the ATP-requiring synthetic site but neither isoleucine nor isoleucyl-tRNA will fit into the editing site which is located in an adjacent β-barrel domain.[104,105] Proofreading steps based on differing chemical properties as well as size can also be visualized.[103,106]

D. Mechanisms of Catalysis

Kinetic studies tell how fast enzymes act but by themselves say nothing about *how* enzymes catalyze reactions. They do not give the **chemical mechanism** of catalysis, the step-by-step process by which a reaction takes place. Most of the individual steps involve the simultaneous breaking of a chemical bond and formation of a new bond. Consider a simple **displacement reaction**, that of a hydroxyl ion reacting with methyl iodide to give the products methanol and iodide ions.

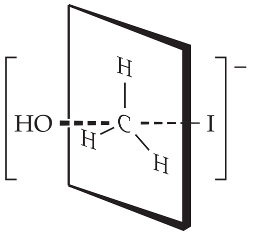


$$HO^- + CH_3{-}I \longrightarrow HO{-}CH_3 + I^- \qquad (9\text{-}76)$$

The reaction can be thought of as an "attack" by the OH^- ion on the "back side" of the carbon atom of the methyl groups with a simultaneous displacement of the I^-.

1. The "Transition State"

A reaction such as that of Eq. 9-76 is not instantaneous, and at some time between that at which the reactants exist and that at which products have been formed the C–I bond will be stretched and partially broken and the new C–O bond will be partially formed. The structure at this point is not that of an ordinary compound and is energetically unstable with respect to both the reactant and the products. The intermediate structure of the least stability is known as the **transition state**. Although no one has actually seen a transition state structure, we might represent that for Eq. 9-76 as follows:



Transition state structure

The negative charge is distributed between the attacking HO^- group and the departing iodide. The bonds to the central carbon atom are no longer tetrahedrally arranged but the C–H bonds lie in a single plane and the partial bonds to the OH and iodine atom lie at right angles to that plane.

It is useful in discussing a reaction mechanism to construct a **transition state diagram** in which Gibbs energy *G* is plotted against **reaction coordinate** (Fig. 9-15A). Energy *E* or enthalpy *H* may be plotted in the same way and authors frequently do not state whether *G*, *E*, or *H* is meant. The reaction coordinate is usually not assigned an exact physical meaning but represents the progress from reactants toward products. It is directly related to the extent to which an existing bond has been stretched and broken or a new one formed. The high energy point is the transition state. A somewhat more detailed idea of a transition state is obtained from a contour diagram such as that of Fig. 9-15B. Here, energy is plotted as a function of two distances, e.g., the lengthening C–I bond distance and the shortening C–O distance for Eq. 9-76. The path of minimum energy across the "saddle point" representing the transition state is indicated by the dashed line.

In a reaction coordinate diagram the difference in value of *G* between reactants and products is the overall Gibbs energy change ΔG for the reaction, while the difference in G between the transition state and reactants is $\Delta G^‡$, the **Gibbs energy of activation**. The magnitude of $\Delta G^‡$ represents the "energy barrier" to a reaction and largely determines the rate constant.

The diagrams in Fig. 9-15 are too simple because enzymatic reactions usually occur in several intermediate steps. There will be transition states for each step with valleys in between. The valleys correspond to intermediate species, which are sometimes very unstable. The passing from reactants to products in an enzymatic reaction can be likened to wandering through a series of mountain ranges of various heights and finally reaching the other side.

Quantitative transition state theory.[107–113] In the 1880s Arrhenius observed that the rate of chemical reactions varies with temperature according to Eq. 9-77 in a manner similar to the variation of an equilibrium constant with temperature (integrate Eq. 6-37 and compare). Here, k is a first-order rate constant, the quantity E_a is known as the **Arrhenius activation energy**, and the constant A is referred to as the "pre-exponential factor" or the "frequency factor".

$$k = A\,e^{-E_a / RT} \tag{9-77}$$

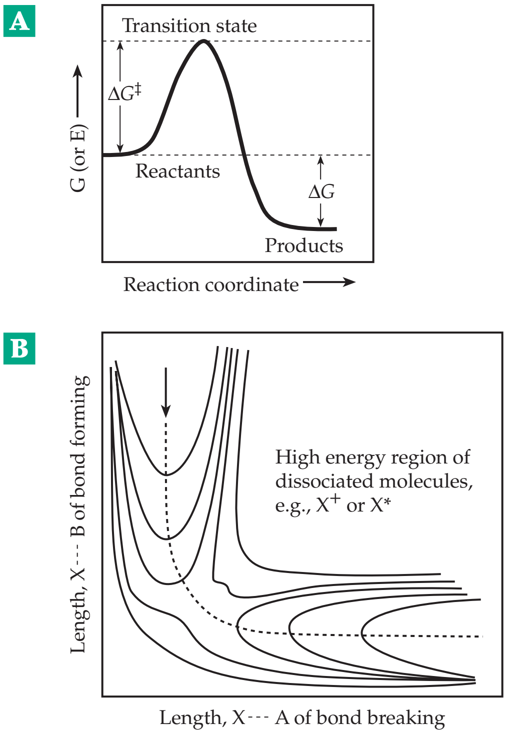

A
Transition state
$\Delta G^‡$
G (or E)
Reactants
ΔG
Products
Reaction coordinate
B
Length, X--- B of bond forming
High energy region of dissociated molecules, e.g., X^+ or X*
Length, X--- A of bond breaking

Figure 9-15 (A) Transition state diagram illustrating Gibbs energy vs reaction coordinate for conversion of reactants to products in a chemical reaction. (B) Contour map of Gibbs energy vs interatomic bond distances for reaction $B + X – A \rightarrow B – X + A$.

The Arrhenius equation, together with studies of the effects of salts on reaction rates and observation of quantitative correlations between rates and equilibrium constants, suggested that a rate constant for a reaction might be a product of a constant term which is nearly independent of temperature and a constant $K^‡$ which has the properties of an equilibrium constant for formation of the transition state. Eyring made this quantitative in 1935 with his "absolute rate theory"[112] according to which all transition states break down with a rate constant $\kappa k_B T / h$. Eyring reached this conclusion by assuming that *the rate of a chemical reaction is determined by the frequency of stretching of the bond that is being broken in the transition state*. To be more precise, it is the "normal-mode" oscillation of the transition state complex along the reaction coordinate.[109] This frequency ν was deduced by describing the vibrational energy as $h\nu$ (from quantum mechanics) and as k_B/T (from classical mechanics) and setting them equal.

$$h\nu = k_B T \text{ and}$$
$$\nu = k_B T / h$$

k_B is the Boltzmann constant
h is Planck's constant. (9-78)

The right side of Eq. 9-78 is usually multiplied by a **transmission coefficient κ**, which may vary from 1 to 0.1 or even much less. However, for lack of any better value, κ is usually assumed to be 1. From Eq. 9-78, at 25°C $\nu = 6.2 \times 10^{12}\ s^{-1}$. This is the maximum rate for a chemical reaction of molecules in the transition state. This is the rate for a single molecule and must be multiplied by the concentration of the reacting substance X in the transition state. This concentration $[X]^‡$ is determined by the equilibrium constant $K^‡ = [X]^‡ / [X]$. The velocity of the reaction becomes

$$\frac{-d[X]}{dt} = v = \frac{k_B T}{h} \cdot K^‡[X] \tag{9-79}$$

The first-order rate constant is

$$k_1 = \frac{-d[X]/dt}{[X]} = K\frac{k_B T}{h} \cdot K^‡ \quad s^{-1} \tag{9-80}$$

Since $\Delta G^‡ = -RT \ln K^‡$, Eq. 9-80 can be rewritten as Eq. 9-81, in which $\Delta G^‡$ is the Gibbs energy of activation.

$$k_1 \approx \frac{k_B T}{h} \cdot e^{-\Delta G^‡/RT} \quad s^{-1} \tag{9-81}$$

At 25°C, $\Delta G^\ddagger$ in kJ / mol, the following "practical" form of Eq. 9-81 can be written.

$$k_1 = 6.2 \times 10^{12}\, e^{-\Delta G^\ddagger/2.48}\ \mathrm{s}^{-1}$$
$$\log k_1 = 12.79 - \Delta G^\ddagger/5.71$$
$$\text{and } \Delta G^\ddagger = 73.0 - 5.71 \log k_1$$

Equation 9-81 is approximate and a more correct statistical mechanical treatment is available.[113,114] See also comment on p. 288 about log k (and log k_1) being unitless. Employing Eq. 6-14, we may expand Eq. 9-81 as follows:

$$k \approx \left(\frac{k_B T}{h} e^{\Delta S\ /R} \right) e^{-\Delta H^\ddagger/RT}\ \mathrm{s}^{-1} \tag{9-82}$$

From this it appears that $\Delta H^\ddagger\quad E_a$, the Arrhenius activation energy. A more correct treatment gives $\Delta H^\ddagger = E_a - RT$ for reactions in solution. However, since RT at 25°C is only 2.5 kJ mol^{-1}, the approximation that $\Delta H^\ddagger = E_a$ is often used. The preexponential term, in parentheses in Eq. 9-82, depends principally on $\Delta S^\ddagger$, the entropy change accompanying formation of the transition state. The quantities $\Delta G^\ddagger$, $\Delta H^\ddagger$, and $\Delta S^\ddagger$ are sometimes measured for enzymatic reactions but useful interpretations are difficult.

Equations 9-81 and 9-82 also show that a small decrease in transition state energy will give a large increase in rate; stabilization of the transition state by 5.7 kJ / mol (1.4 kcal / mol) will increase the rate 10-fold. If $\Delta G^\ddagger = 400$ kJ/mol, as for a strong covalent single bond, $k_1 = 5 \times 10^{-58}\ \mathrm{s}^{-1}$ and the half-life $t_{1/2} = .693 / k_1 = 1.3 \times 10^{57}$ s (4×10^{49} years, greater than the ~10^{10} years estimated age of the universe). If $\Delta G^\ddagger = 100$ kJ / mol, $\log k_1 \approx -4.7\ \mathrm{s}^{-1}$, $k_1 = 1.9 \times 10^{-5}$, and $t_{1/2} \approx 10$ h. This is about the rate of a typical nonenzymatic "model" reaction at 25°C. If $\Delta G^\ddagger = 50$ kJ / mol, $\log k_1 \approx 4$, as for a *fast* enzyme.

We can conclude that *enzymes make use of relatively small energy differences in catalyzing reactions*. The energies of numerous van der Waals interactions of hydrogen bonds and electrostatic attraction or repulsion of charges are sufficient. Nevertheless, we can see that anything that stabilizes the transition state (decreases $\Delta G^\ddagger$) will increase the rate of reaction. The role of a catalyst is to permit the formation of a transition state of lower energy (higher stability) than that for the uncatalyzed reaction. Stabilization of the transition state of a reaction by an enzyme suggests that the enzyme has a higher affinity for the transition state than it does for substrate or products, an idea that appears to have been expressed first by Haldane[41] and popularized by Pauling.[115]

> *I think that enzymes are molecules that are complementary in structure to the activated complexes of the reactions that they catalyze. The attraction of the enzyme molecule for the activated complex would thus lead to a decrease in its energy, and hence to a decrease in energy of activation of the reaction and to an increase in the rate of the reaction.*

Transition state inhibitors. Suppose that a chemical reaction of a compound S takes place with rate constant k_N through transition state T. Let the equilibrium constant for formation of T be $K_N^\ddagger$. Assume that an enzyme E can combine either with S with dissociation constant K_{dS} or with the compound in its transition state structure T with dissociation constant K_{dT} (Eq. 9-83).

$$\begin{tikzcd}
\begin{matrix}\mathrm{S}\\+\\\mathrm{E}\end{matrix} \arrow[r, shift left, "K_N^\ddagger"] \arrow[d, shift right, "K_{dS}"'] & \begin{matrix}\mathrm{T}\\+\\\mathrm{E}\end{matrix} \arrow[l, shift left] \arrow[r] \arrow[d, shift right, "K_{dT}"'] & \text{product} \\
\mathrm{ES} \arrow[u, shift right] \arrow[r, shift left, "K_E^\ddagger"] & \mathrm{ET} \arrow[l, shift left] \arrow[u, shift right] \arrow[r] & \text{product}
\end{tikzcd} \tag{9-83}$$

If equilibrium is assumed for all four sets of double arrows it is easy to show that Eq. 9-84 holds.

$$K_E^\ddagger / K_N^\ddagger = K_{dS} / K_{dT} \tag{9-84}$$

According to transition state theory, if the transmission coefficient $\kappa = 1$, T and ET will be transformed to products at the same rate. Thus, if the mechanisms of the nonenzymatic and enzymatic reactions are assumed the same, the ratio of maximum velocities for first-order transformation of ES and S will be given by Eq. 9-85. For some enzymes the ratio

$$k_E / k_N = K_E^\ddagger / K_N^\ddagger = K_{dS} / K_{dT} \tag{9-85}$$

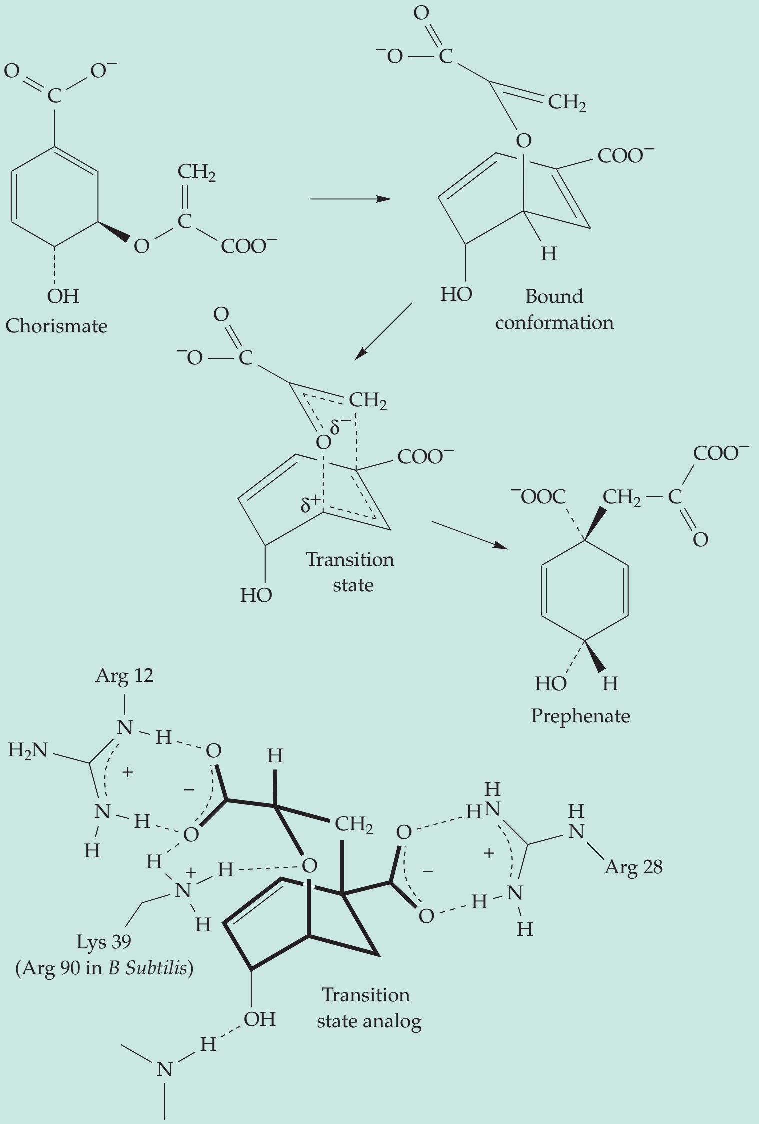
k_E / k_N may be 10^8 or more. Thus, if $K_{dS} \approx 10^{-3}$ the constant K_{dT} would be $\approx 10^{-11}$. The enzyme would be expected to bind the transition state structure (T) 10^8 times more tightly than it binds S.

The foregoing reasoning suggests that if structural analogs of T could be found for a particular reaction, they too might be very tightly bound—more so than ordinary substrate analogs.[116] Wolfenden[117] listed a series of compounds which may be **transition state inhibitors** of this type and many others have been described.[117a] Nevertheless, very tight, reversible binding is no *proof* that an inhibitor is actually binding to the transition state structure of the enzyme. For example, an inhibitor may bind to a conformational form of the enzyme that lies off to a side rather than directly on the reaction pathway for catalysis.

Describing the transition state. How can we describe the structure of the transition state when we have no direct knowledge of its structure? We can try

BOX 9-E CHORISMATE MUTASE

Bacteria, fungi, and plants convert products of glucose metabolism into the aromatic amino acids phenylalanine, tyrosine, and tryptophan by a complex sequence of reactions that is described in Chapter 25. One of these reactions, the conversion of **chorismate** into **prephenate,** is unique among enzyme-catalyzed reactions. It is a Claisen rearrangement that occurs readily *without catalysis* in an aqueous solution. However, the rate is increased over one million-fold by the enzyme **chorismate mutase**. The enzyme from *Bacillus subtilis* is a trimer of identical, small 127-residue subunits. Many studies, both of the nonenzymatic and enzymatic reactions, have suggested a “pericyclic” mechanism involving a polar chairlike transition state whose presumed structure is shown in the accompanying equation.[a–f] The transition state analog, whose structure is also depicted, is a powerful inhibitor. X-ray structural studies show that this inhibitor binds in the conformation shown with a number of interactions with polar side chain groups.[d,g] Some of these interactions are shown for the *E. coli* enzyme in the figure;[g] they are different in enzymes from *B. subtilis*[d] and yeast.[h] However, in no case are there groups in obvious positions to serve as acid–base catalysts and the maximum velocity V_{max} is independent of pH.[b] NMR studies show that the bound product prephenate displays large shifts in some ^{13}C resonances, indicating strong interaction with the polar side chains of the protein.[c] Theoretical calculations also suggested that these interactions, especially those with Arg 90 and Glu 78 of the *B. subtilis* enzymes,[f–l] help to stabilize the partial charge separation in the polar transition state. Studies of mutant enzymes have confirmed that the charged Arg 90 (Lys 39 for *E. coli*) side chain is essential.[k]



---

[a] Gray, J. V., Eren, D., and Knowles, J. R. (1990) *Biochemistry* **29**, 8872–8878

[b] Turnbull, J., Cleland, W. W., and Morrison, J. F. (1991) *Biochemistry* **30**, 7777–7782

[c] Rajagopalan, J. S., Taylor, K. M., and Jaffe, E. K. (1993) *Biochemistry* **32**, 3965–3972

[d] Chook, Y. M., Gray, J. V., Ke, H., and Lipscomb, W. N. (1994) *J. Mol. Biol.* **240**, 476–500

[e] Gray, J. V., and Knowles, J. R. (1994) *Biochemistry* **33**, 9953–9959

[f] Lyne, P. D., Mulholland, A. J., and Richards, W. G. (1995) *J. Am. Chem. Soc.* **117**, 11345–11350

[g] Lee, A. Y., Karplus, P. A., Ganem, B., and Clardy, J. (1995) *J. Am. Chem. Soc.* **117**, 3627–3628

[h] Xue, Y., Lipscomb, W. N., Graf, R., Schnappauf, G., and Braus, G. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10814–10818

[i] Wiest, O., and Houk, K. N. (1995) *J. Am. Chem. Soc.* **117**, 11628–11639

[j] Kast, P., Hartgerink, J. D., Asif-Ullah, M., and Hilvert, D. (1996) *J. Am. Chem. Soc.* **118**, 3069–3070

[k] Lin, S. L., Xu, D., Li, A., Rosen, M., Wolfson, H. J., and Nussinov, R. (1997) *J. Mol. Biol.* **271**, 838–845

[l] Ma, J., Zheng, X., Schnappauf, G., Braus, G., Karplus, M., and Lipscomb, W. N. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14640–14645

to predict the atomic coordinates of all of the atoms of the reacting substrates and of the protein. The transition state structure of the substrate must be between that of the last ES complex and the first EP complex formed. It should be similar to that of a transition state inhibitor. We can hope to obtain suitable X-ray or NMR structures and deduce an appropriate enzyme structure by molecular modeling. Theoretical calculations using quantum chemical methods are of value in this effort.[118,119]

Lengths of the breaking and forming bonds in a transition state are often estimated using **kinetic isotope effects** (KIEs) on the velocity of the reaction.[118–122] This is somewhat indirect. Nevertheless, by measuring these effects for a large number of substrates and several isotopes a picture of the transition state structure that is good enough to use as a model for design of an enzyme inhibitor may be obtained.[117a,121–124] Other physical measurements such as Raman difference spectroscopy of isotopically labeled inhibitors can also be of value.[125]

Getting to the transition state. Since some bonds must lengthen and some shorten and bond angles must also change, it may not be simple for an enzyme to catalyze necessary steps prior to the rate-limiting breakdown of the transition state. Hydrogen bonds within the protein may need to be broken and reformed with new bonding partners. Nonpolar side chains may need to shift to give a packing arrangement that will allow the movement of nuclei that must occur in forming the transition state structure. The enzyme must not only stabilize the transition state but also hold the substrate and escort it into the transition state configuration.

The attacking hydroxyl ion in Eq. 9-76 carries a negative charge which becomes distributed more or less equally between the hydroxyl and the iodine atom in the transition state. In this state the CH_3 group carries a partial positive charge as well. To provide good complementarity between an enzyme and this transition state, a change in the initial charge distribution within the enzyme will also be required. If a positively charged group initially binds the OH^- ion it may have to lose part of its charge and a group next to the iodide atom may have to gain positive charge. In the nonenzymatic reaction of Eq. 9-76 the ionic atmosphere provided by positive counterions in solution can continuously readjust to keep the negative charge effectively balanced at every step along the reaction coordinate and through the transition state. Within enzymes this adjustment may occur via redistribution of electrical charges within the polarizable network of internal hydrogen bonds. The enzyme structure must allow this. Because of the complexity of an enzymatic transition state it may be hard to compare it with the transition state of the corresponding nonenzymatic reaction. Effects of pressure[125a] and of temperature[125b] also raise doubts about the simple picture of Eq. 9-83. It also follows that tight binding of substrates in the ground state does not necessarily interfere with transition state stabilization.[126–129a]

2. Microscopic Reversibility

The important statistical mechanical principle of microscopic reversibility asserts that *the mechanism of any chemical reaction considered in the reverse direction must be exactly the inverse of the mechanism of the forward reaction*. A consequence of this principle is that if the mechanism of a reaction is known, that of the reverse reaction is also known. Furthermore, it follows that the *forward and reverse reactions catalyzed by an enzyme must occur at the same active site* on the enzyme and the transition state must be the same in both directions. The principle of microscopic reversibility is often useful when the likelihood of a given mechanism is being considered. If a mechanism is proposed for a reversible reaction in one direction the principle of microscopic reversibility will give an unambiguous mechanism for the reverse reaction. Sometimes this reverse mechanism will be chemically untenable and, recognizing this, the enzymologist can search for a better one.

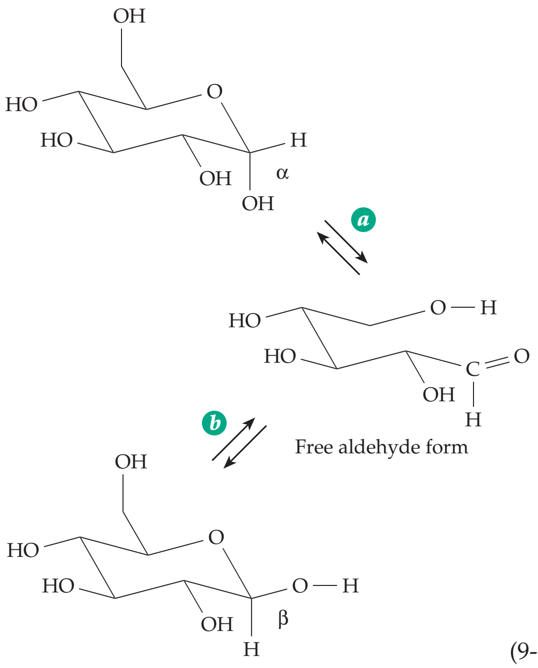
3. Acid and Base Catalysis

Many reactions that are promoted by enzymes can also be catalyzed by acids or bases or by both. An example is **mutarotation**, the reversible interconversion of the α- and β-anomeric forms of sugars (Eqs. 4-1 and 9-86). This reaction is catalyzed by a specific **mutarotase** and also by inorganic acids and bases. The frequently observed bell-shaped curve for the dependence of rate of catalysis on pH (e.g., Eqs. 9-55 to 9-57 and Fig. 9-8) also suggests participation of both protonated and unprotonated acid–base groups present in enzymes.

Acidic and basic groups in enzymes. In this discussion the symbols HB or H^+B will be used for acids and B^- or B for their conjugate bases. Remember that *strong acids* have *low* pK_a *values* and that the *conjugate bases formed from them are weak bases*. Likewise, *very weak acids have high* pK_a *values and their conjugate bases are strong*. The following are among the pK_a values that may be important in considering an enzyme mechanism. However, depending upon the environment of the side chain in a protein, these pK_a values may fall substantially outside of the indicated ranges.

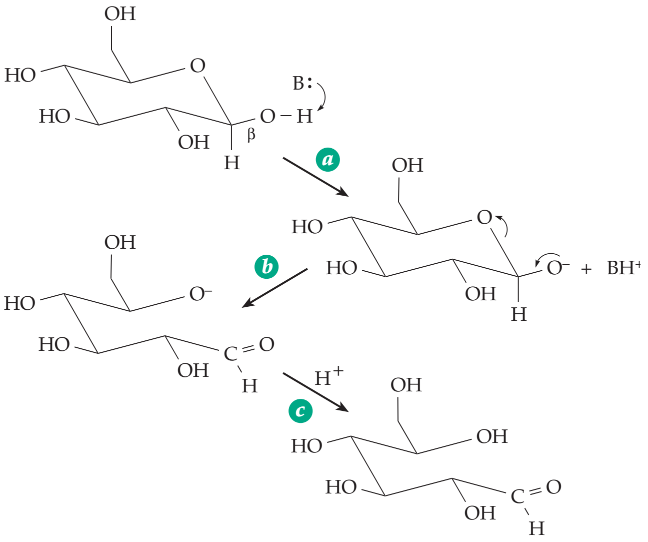
–1.74	H_3O^+ in 55.5 M water
~3.6	–COOH, terminal in peptide
4.3–4.7	–COOH in a glutamic or aspartic acid side chain
6.4–7.0	Imidazole (histidine) and phosphate ($–OPO_3H^–$)
7.5–8.0	$–NH_3^+$, terminal in peptide
8.5–10	–SH of cysteine side chains
9.5–11	Phenolic –OH of tyrosine
~10.5	$–NH_3^+$ of lysine side chain
~13.6	$–CH_2OH$ of serine side chain
15.7	free $HO^–$ in 55.5 M H_2O, 25°C

Acid–base catalysis of mutarotation. The mutarotation of glucose proceeds through the free aldehyde form as an intermediate (Eq. 9-86). A hydrogen atom is removed as H^+ from the oxygen at carbon 1 in step *a* and a proton (probably a different one) is transferred to the oxygen of the ring with cleavage of the O – C bond to the anomeric carbon atom. A similar process in reverse is required for step *b*. Transfers of hydrogen ions between atoms of oxygen, nitrogen, and sulfur atoms are a common feature of biochemical reactions. The bonds between hydrogen and O, N, and S atoms tend to be polarized strongly, leaving a partial positive charge on the hydrogen atoms. Consequently, the groups are weakly acidic and protons can be transferred from them relatively easily. It is reasonable to suppose that acid and base catalysis is related to these proton transfers.



(9-86)

General base and general acid catalysis. Base-catalyzed mutarotation might be formulated as follows: A hydroxyl ion or some other base attacks the proton on the anomeric –OH group of the sugar, removing it to form an anion and the conjugate acid BH^+ (Eq. 9-87, step *a*). The anion is isomerized to a second anion with the ring opened (step *b*). Addition of a proton (transfer of a proton from H_3O^+) produces the free aldehyde form of the sugar (step *c*).



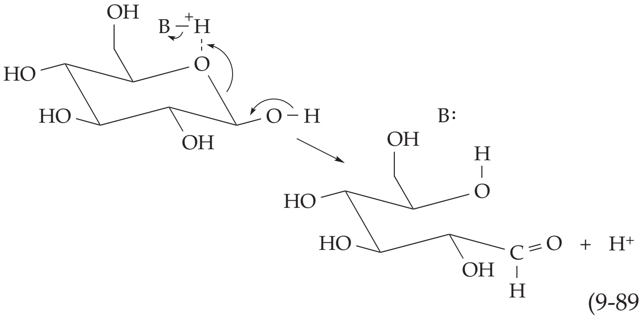
(9-87)

The catalytic base B: might be $HO^–$ or a weaker base such as ammonia or even water. For reactions the rate is proportional only to the concentration of $OH^–$ and the presence of other weaker bases has no effect.[129b] Such catalysis is referred to as **specific hydroxyl ion catalysis**.[19] More commonly, the rate is found to depend both on $[OH^–]$ and on the concentration of other weaker bases. In such cases the apparent first-order rate constant (k_{obs}) for the process is represented by a sum of terms (Eq. 9-88). The term k_{H_2O} is the rate in

$$k_{obs} = k_{H_2O} + K_{OH^–}[OH^–] + k_B[B] \qquad (9\text{-}88)$$

pure water and represents catalysis occurring by the action of water alone as either an acid or a base. The last two terms represent the contributions to the catalysis by $OH^–$ and by the other base, respectively. The term $k_B[B]$ represents **general base catalysis**.

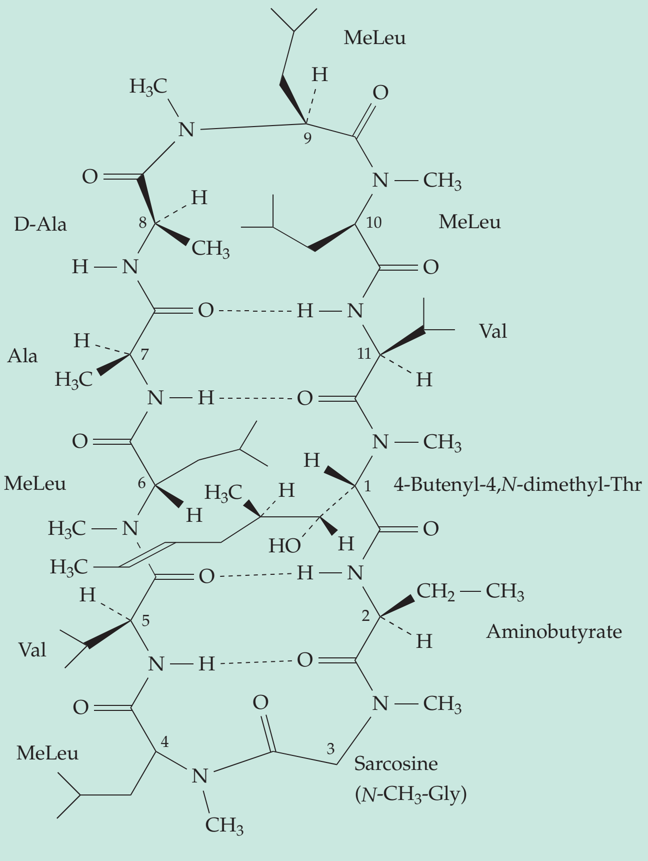
Catalysis of mutarotation by acids occurs if an acid donates a proton to the oxygen in the sugar ring as



(9-89)

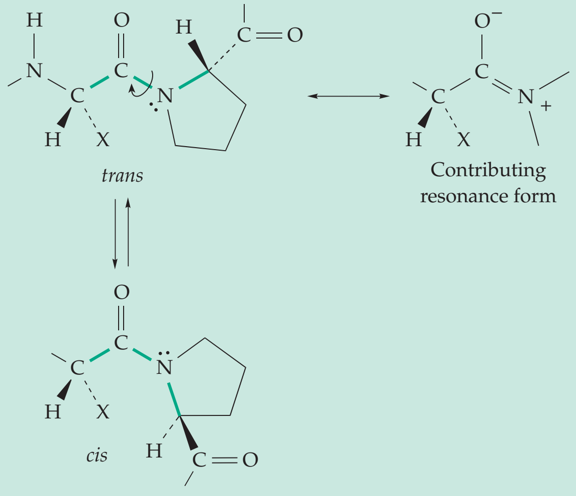
BOX 9-F IMMUNOPHILINS AS ROTAMASES

Since its introduction into clinical use in about 1979 the immunosuppresant **cyclosporin** has been responsible for a revolution in human organ transplantation.[a–c] The exact mechanism of action in suppressing T-lymphocyte-mediated autoimmune responses is still not completely clear, but cyclosporin, a cyclic lipophilic peptide from a fungus, was found to bind to specific proteins that were named **cyclophilins**.[d] Human cyclophilin A is a 165-residue protein which associates, in the crystal form, as a decamer with five-fold rotational and dihedral symmetry.[e] This protein is also found in almost all other organisms and has a highly conserved sequence.[f] Another immunosuppresant, a synthetic molecule known only as **FK506**, has an action similar to that of cyclosporin but binds to distinctly different proteins, the FK506-binding proteins (FKBPs), of which the 107-residue FKBP-12 is the best known.[g–j] These proteins also bind another immunosuppresant, **rapamycin**, which, however, has different biological properties than does FK506. Both FK506 and rapamycin are macrocyclic compounds but with structures very different from that of cyclosporin.

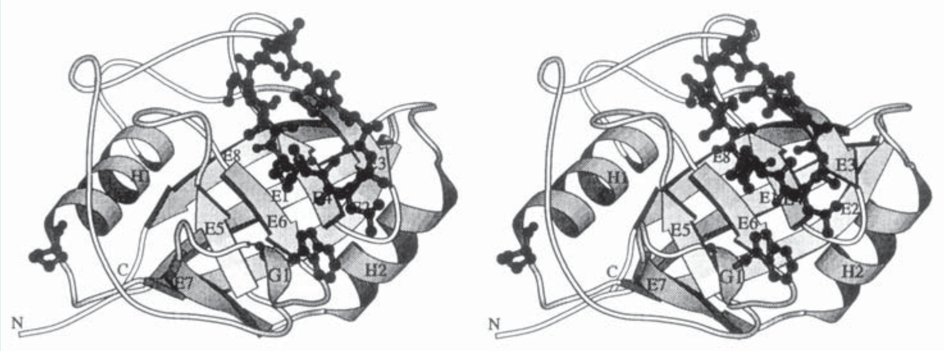

MeLeu
D-Ala
MeLeu
Val
Ala
MeLeu
4-Butenyl-4,N-dimethyl-Thr
Aminobutyrate
Val
MeLeu
Sarcosine
(N-CH3-Gly)

Cyclosporin A

It was a surprise to discover that all of the cyclophilins and FK506-binding proteins are **peptidyl prolyl *cis–trans* isomerases** or **rotamases**. They all catalyze the following simple and reversible reaction of a prolyl peptide linkage:


trans
cis
Contributing
resonance form

The transition state energy for the reaction is lowered by 33 kJ mol^{-1} by cyclophilin A[k] and by 27 kJ mol^{-1} by FKBP-12[g] with corresponding rate increases of ~ – 6 x 10^5- and 5 x 10^4-fold, respectively.


H1
E8
E3
E1
E4
E2
E5
E6
C
E7
G1
H2
N
H1
E8
E3
E1
E2
E5
E6
C
E7
G1
H2
N

Cyclosporin (black) bound to human cyclophilin. From Pflügl *et al.*[e] Courtesy of J. N. Jansonius.

BOX 9-F (continued)

For both enzymes the maximum velocity is independent of pH over a wide region. For cyclophilin the rate is nearly independent of the nature of residue X in the foregoing structure but FKBP prefers a hydrophobic residue.

How do these enzymes work? One possibility would be for the protein to transfer a proton to the nitrogen atom of the proline ring. This would destroy the partial double-bond character of the amide linkage and allow free rotation about a single bond. The same thing could be accomplished if a nucleophilic group from the enzyme formed a covalent adduct with the C=O of the substrate. However, this seems unlikely because there is no suitably placed nucleophile. A third possibility is that the enzyme distorts the substrate and stabilizes the transition state using only noncovalent interactions.[g,h,l] This would account for the lack of pH dependence and observable solvent isotope effects. However, in human cyclophilin the guanidinium group of arginine 55 hydrogen bonds to the peptide C=O of the substrate's proline residue and could easily shift to place a guanidinium proton against the proline ring nitrogen effectively protonating it and permitting rotation about the peptide linkage.[m]

Do cyclosporin and FK506 act as transition state inhibitors? If so, we might learn something about the mechanism from the three-dimensional structures of the inhibited rotamases.[i,l,n,o] From these structures as well as those with simpler substrate analogs, it is seen that the substrate is desolvated and that many nonbonded (van der Waals) interactions stabilize the binding. Hydrogen bonding is also important.[p,q] The presence of distinct hydrogen bonds to the peptide NH and C=O groups on the N-terminal side of the substrate X-Pro linkage and less well-defined bonding on the terminal side suggests that the C-terminal part may be rotated while the enzyme holds the N-terminal part.[p] It can also be concluded that mechanisms are probably not identical for all rotamases.

Cyclophilins and FKBPs are large families of proteins. In *E. coli* there are two cyclophilin genes,[r,s] three encoding FK506-binding proteins, and one of a third family of bacterial rotamases (parvulins) that lack sequence similarity to the other two families.[t,u] Yeast contains at least five cyclophilins[v] and even more eukaryotic FKBPs are known.[j] There is abundant evidence that these proteins play an important role in protein folding *in vivo* (Chapter 29).

---

[a] Kahan, B. D. (1989) *N. Engl. J. Med.* **321**, 1725–1738

[b] Schreiber, S. L. (1991) *Science* **251**, 283–287

[c] High, K. P., Joiner, K. A., and Handschumacher, R. E. (1994) *J. Biol. Chem.* **269**, 9105–9112

[d] Fruman, D. A., Burakoff, S. J., and Bierer, B. E. (1994) *FASEB J.* **8**, 391–400

[e] Pflügl, G. M., Kallen, J., Jansonius, J. N., and Walkinshaw, M. D. (1994) *J. Mol. Biol.* **244**, 385–409

[f] Lippuner, V., Chou, I. T., Scott, S. V., Ettinger, W. F., Theg, S. M., and Gasser, C. S. (1994) *J. Biol. Chem.* **269**, 7863–7868

[g] Fischer, S., Michnick, S., and Karplus, M. (1993) *Biochemistry* **32**, 13830–13837

[h] Orozco, M., Tirado-Rives, J., and Jorgensen, W. L. (1993) *Biochemistry* **32**, 12864–12874

[i] Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) *J. Mol. Biol.* **229**, 105–124

[j] Lam, E., Martin, M. M., Timerman, A. P., Sabers, C., Fleischer, S., Lukas, T., Abraham, R. T., O'Keefe, S. J., O'Neill, E. A., and Wiederrecht, G. J. (1995) *J. Biol. Chem.* **270**, 26511–26522

[k] Eberhardt, E. S., Loh, S. N., Hinck, A. P., and Raines, R. T. (1992) *J. Am. Chem. Soc.* **114**, 5437–5439

[l] Kakalis, L. T., and Armitage, I. M. (1994) *Biochemistry* **33**, 1495–1501

[m] Zhao, Y., and Ke, H. (1996) *Biochemistry* **35**, 7362–7368

[n] Thériault, Y., Logan, T. M., Meadows, R., Yu, L., Olejniczak, E. T., Holzman, T. F., Simmer, R. L., and Fesik, S. W. (1993) *Nature (London)* **361**, 88–91

[o] Konno, M., Ito, M., Hayano, T., and Takahashi, N. (1996) *J. Mol. Biol.* **256**, 897–908

[p] Kern, D., Kern, G., Scherer, G., Fischer, G., and Drakenberg, T. (1995) *Biochemistry* **34**, 13594–13602

[q] Göthel, S. F., Herrler, M., and Marahiel, M. A. (1996) *Biochemistry* **35**, 3636–3640

[r] Clubb, R. T., Ferguson, S. B., Walsh, C. T., and Wagner, G. (1994) *Biochemistry* **33**, 2761–2772

[s] Edwards, K. J., Ollis, D. L., and Dixon, N. E. (1997) *J. Mol. Biol.* **271**, 258–265

[t] Rudd, K. E., Sofia, H. J., Koonin, E. V., Plunkett, G., III, Lazar, S., and Rouviere, P. E. (1995) *Trends Biochem. Sci.* **20**, 12–14

[u] Rahfeld, J.-U., Rücknagel, K. P., Stoller, G., Horne, S. M., Schierhorn, A., Young, K. D., and Fischer, G. (1996) *J. Biol. Chem.* **271**, 22130–22138

[v] Matouschek, A., Rospert, S., Schmid, K., Glick, B. S., and Schatz, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6319–6323

---

shown in Eq. 9-89. Again, either **specific acid catalysis** (by H_3O^+) or **general acid catalysis** is possible.

Enzymes are not able to concentrate protons or hydroxyl ions to the point that *specific* base or acid catalysis would be effective. However, either general acid or general base catalysis can be accomplished by groups present in an enzyme in their normal states of protonation at the pH of the cell. Thus, if one of the enzymes must be dissociated to its conjugate base and a second must be protonated for reaction to take place (as in Eq. 9-54), it is reasonable to suppose that these two groups participate in acid–base catalysis. Of the acidic and basic groups present in proteins, the imidazole group of histidine would appear to be the most ideal both for general base catalysis and, because both forms may exist in nearly equal amounts at pH 7, for general acid catalysis. However, carboxyl, thiol, lysyl, tyrosyl, and N-terminal amino groups are all thought to function in various enzymes.

The mutarotase of *E. coli* has a turnover number of $10^4\ s^{-1}$. The plot of $-\log K_m$ vs pH indicates two pK_a values in the free enzyme at 5.5 and 7.6, while the plot of $\log V_{max}$ yields a single pK_a of 4.75 for the ES complex.[130] The latter might represent a carboxyl group on an imidazole group in its conjugate base form. Why doesn't the group having pK_a 7.6 in the free enzyme also show a pK_a in the ES complex? Either the group has no catalytic function so that EH_2S of Eq. 9-54 reacts to form products just as fast as does HES or the pK_a is so strongly shifted by substrate binding that it is not detected in the $\log V_{max}$ plot.

The Brönsted relationships. The effectiveness of a specific base as a general base catalyst can usually be related to its basicity (pK_a) via the **Brönsted equation** (Eq. 9-90). Here, k_B is defined by Eq. 9-88 and G_B is a

$$\log k_B = \log G_B + \beta\,(pK_a) \tag{9-90}$$

constant for a particular reaction. A similar equation (9-91) relates the constant k_A for general acid catalysis

$$\log k_{HB} = \log G_A - \alpha\,(pK_a) \tag{9-91}$$

to pK_a. These equations are linear Gibbs energy relationships similar to the ones discussed in Box 6-C. For the Brönsted equations to hold, the Gibbs energy of activation for the reaction must be directly related to the basicity or acidity of the catalyst.

The exponents β and α of Eqs. 9-90 and 9-91 measure the *sensitivity* of a reaction toward the basicity or acidity of the catalyst. It is easy to show that as β and α approach 1.0 general base or general acid catalysis is lost and that the rate becomes exactly that of specific hydroxyl ion or specific hydrogen ion catalysis.[131] As β or α approach zero basic or acidic catalysis is indetectable. Thus, general base or general acid catalysis is most significant when β or α is in the neighborhood of 0.5. Under these circumstances it is easy to see how a moderately weak basic group, such as the imidazole group of histidine, can be an unusually effective catalyst at pH 7.

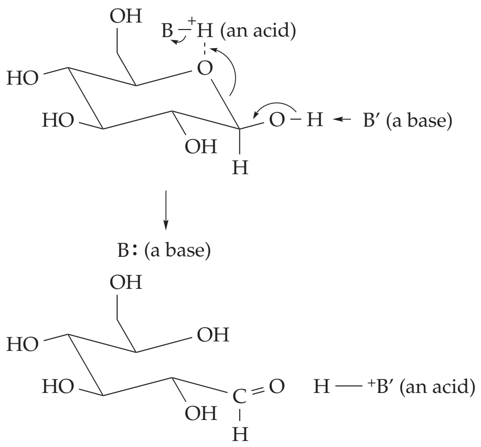
To determine α or β experimentally a plot of $\log k_B$ or $\log k_{HB}$ vs pK_a (a **Brönsted plot**) is made and the slope is measured. Statistical corrections (Chapter 7) should be applied for dicarboxylic acids and for ammonium ions from which one of three protons may be lost from the nitrogen atom. General base or general acid catalysis implies an important feature of any mechanism for which it is observed, namely, that removal of a proton or addition of a proton is involved in the rate-determining step of a reaction.

Concerted acid–base or tautomeric catalysis. A third possible type of catalysis requires that a base and an acid act *synchronously* to effect the breaking and formation of bonds in a single step. Thus, tetramethylglucose mutarotates very slowly in benzene containing either pyridine (a base) or phenol (an acid). However, when both pyridine and phenol are present, mutarotation is rapid. This suggested to Swain and Brown[132] a **concerted mechanism** (Eq. 9-92) in which both an acid and a base participate.

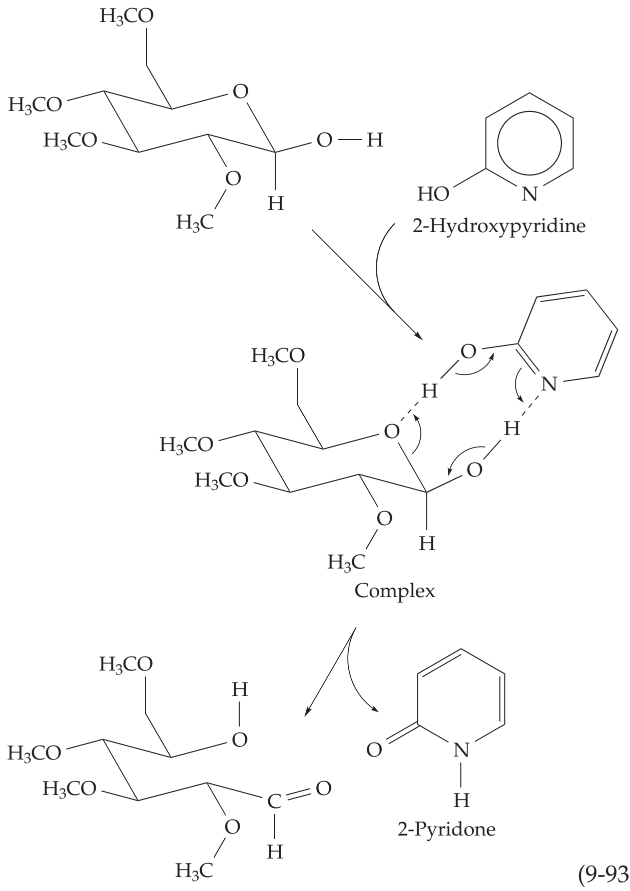
During the reaction shown, the acid BH^+ is converted to its conjugate base B and the base B′ to its conjugate acid H^+B'. It might seem that these agents, having been altered by the reaction, are not serving as true catalysts. However, a simple proton exchange will restore the original forms and complete the catalytic cycle. In aqueous solutions, water itself might act as the acid or the base or even both in concerted catalysis.

The original experimental evidence for concerted acid–base catalysis of the mutarotation in benzene is now considered unsound[133,134] and concerted acid–base catalysis has been difficult to prove for nonenzymatic reactions in aqueous solution. However, measurements of kinetic isotope effects seem to support Swain and Brown's interpretation.[135] Concerted acid–base catalysis by acetic acid and acetate ions may have been observed for the enolization of acetone[136] and it may be employed by enzymes.[136a]

Swain and Brown showed that a more effective catalyst for the mutarotation of sugars than a mixture of an acid and a base can be designed by incorporating the acidic and basic groups into *the same molecule*.[132,135] Thus, with 0.1 M tetramethylglucose in benzene solution, 0.001 M α-hydroxypyridine is 7000 times as effective a catalyst as a mixture of 0.001 M pyridine + 0.001 M phenol. Swain and Brown suggested the following completely concerted reaction mechanism for the **polyfunctional catalyst α-hydroxypyridine** in which the hydrogen-bonded complex formed (Eq. 9-93) is analogous to an enzyme–substrate complex. The product of the catalyst is 2-pyridone, a tautomer of



(9-92)



(9-93)

2-hydroxypyridine, with which it is in a rapid reversible equilibrium.

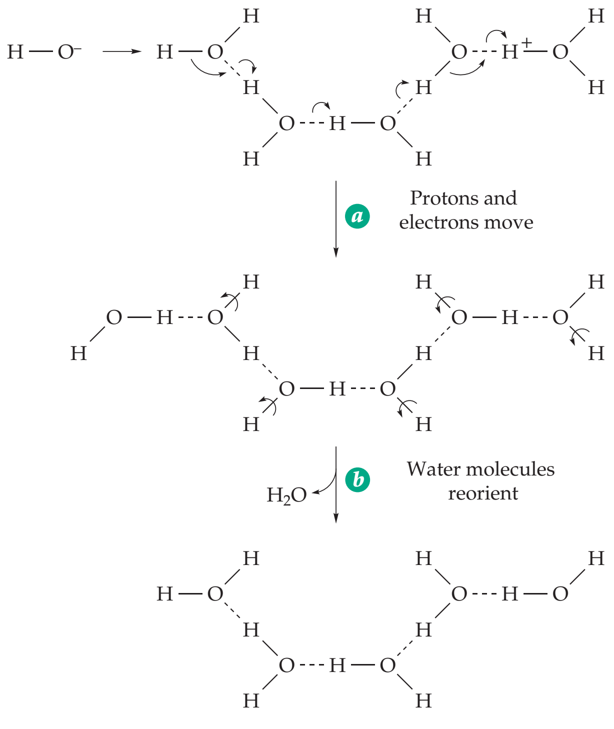
Rony called catalysis of the type illustrated in Eq. 9-93 tautomeric catalysis and suggested that its efficiency lies not simply in the close proximity of acidic and basic groups in the same molecule but also in the ability of the catalyst to repeatedly cycle between the two tautomeric states.[133] For an enzyme the tautomerization of the free catalyst could sometimes be rate determining (see Section A,4 on isomechanisms).

4. Hydrogen Bonding and the Transfer of Protons within Active Sites

An extensive network of hydrogen bonds runs throughout most proteins and may be especially complex within active sites. The network often runs through bound substrate molecules and immobilized water molecules in the active site cavity. This network arises in part because of the frequent occurrence of acidic and basic catalytic groups in active sites and by the fact that many substrates contain polar groups. The structure of an enzyme often seems to be more rigid in a complex with substrates or inhibitors than in the free state. Does this network of linked hydrogen bonds play a role in catalysis? If so, what? Wang suggested that rapid transfer of protons along rigidly and accurately held hydrogen bonds in the ES complex may be an essential feature of enzymatic catalysis.[137] This conclusion is supported by many more recent observations. A remaining question is whether this proton transfer can take place in such a way that the transition state barrier is lowered.

Ultrafast proton transfer. The diffusion-controlled limit for second-order rate constants (Section A3) is ~ 10^{10} M^{-1} s^{-1}. In 1956, Eigen, who had developed new methods for studying very fast reactions, discovered that protons and hydroxide ions react much more rapidly when present in a lattice of ice than when in solution.[138] He observed second-order rate constants of 10^{13} to 10^{14} M^{-1} s^{-1}. These represent rates almost as great as those of molecular vibration. For example, the frequency of vibration of the OH bond in water is about 10^{14} s^{-1}. The latter can be deduced directly from the frequency of infrared light absorbed in exciting this vibration: Frequency ν equals wave number (3710 cm^{-1} for –OH stretching) times *c*, the velocity of light (3 x 10^{10} cm s^{-1}).

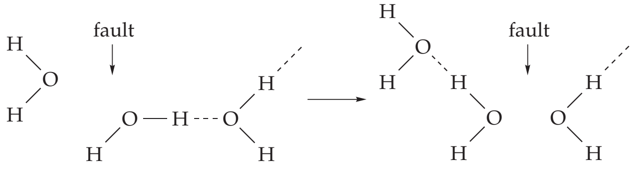
This ultrafast transfer of protons can be explained as follows: The OH^- ion and the proton, which is combined with a water molecule to form H_3O^+, are both hydrogen-bonded to adjacent water molecules. In ice a chain of hydrogen-bonded water molecules links the hydroxide and the hydrogen ions (Eq. 9-94). By synchronous movement of electron pairs from the OH^- ion and from each of the water molecules in the



A second pair of HO^- and H^+ can now use the same pathway

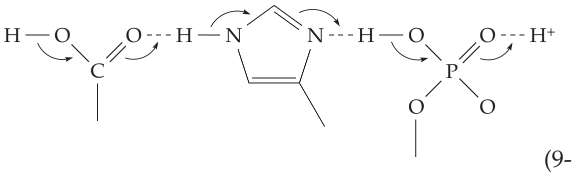
(9-94)

chain (as indicated by the little arrows) the neutralization can take place during the time of one molecular vibration (step *a*, Eq. 9-94). The positions of the oxygen atoms remain unchanged at the end of the reaction but the protons that were engaged in hydrogen bond formation have moved toward the left and are now attached to different oxygen atoms. The central chain of water molecules can be restored to its original state if each molecule rotates around one of its single bonds, swinging a hydrogen from right back to left (Eq. 9-94, step *b*). This must be a slower process which may occur one molecule at a time. The rotation of the leftmost water molecule would leave an empty space between two O-atoms, a "fault" in the H-bonded chain of the ice structure. This fault can be corrected by rotation of the second molecule (Eq. 9-95); however, this would create a fault at the right of the second water molecule. This would induce rotation of the third molecule, etc., causing the fault to migrate from left to right across the entire chain of water molecules (Eq. 9-94, step *b*) and leaving the chain ready to function again in transferring another proton.



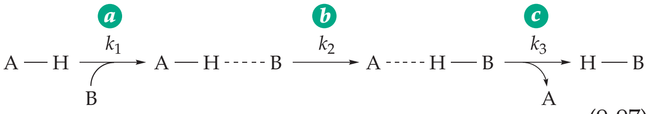
(9-95)

Because of the disorder present in ice, an array of water molecules such as that in Eq. 9-94 wouldn't revert to its exact original form in step *b*. However, active sites of enzymes are highly structured and proton transfers may occur with precision. For example, a synchronous shift of protons in an array of carboxylic acid, imidazolium, and phosphate groups can be envisioned readily (Eq. 9-96). The net effect of the process is to transfer a proton from one end of the chain to the other (as in Eq. 9-94) with facile tautomerization reactions providing the pathway. Such a pathway might be constructed by protein side chains to join the two sides of an active center promoting a concerted acid–base catalyzed reaction such as that of Eq. 9-92. Other tautomerization processes are possible within proteins if the existence of less stable minor tautomers of selected amide groups in the peptide backbone is allowed.[139] Nagle and Morowitz suggested a process similar to that of Eq. 9-94 but involving side chains of serine residues.[140]



(9-96)

Proton transfer rates. Consider the reversible reaction of a proton acceptor B with acid H-A (Eq. 9-97). Eigen pointed out that the reaction will be fastest if the two reactants form a hydrogen-bonded complex (Eq. 9-97, step *a*).[138] The hydrogen bonding shortens the distance from the proton to B and allows for very rapid transfer of the proton from A to B within the hydrogen-bonded complex (step *b*). The activation energy is close to zero. The complex dissociates in step *c* to form the products. The reactions are reversible, even though they have been indicated by unidirectional arrows.



(9-97)

The hydrogen-bonded complexes A–H---B and A ---H–B can be formed readily if A and B are oxygen or nitrogen bases such as $-COO^-$, $-OH$, $-NH_2$, or imidazole. In such cases, as in ice, the interconversion of the two complexes (step *b* in Eq. 9-97) is very fast. The overall rate of the proton transfer will then be limited by diffusion of B and H–A together or by diffusion of A and H–B apart. It might seem that these two processes should also both be very fast. However, the rates will be determined by the concentration ratio [A–H--- B]/[A--- H–B]. If [A–H--- B] equals [A--- H – B] the rate of dissociation of A–H–B will be half what it is if nearly all of the complex is A---H–B. If A–H---B predominates by 1000 to 1, the rate will be slowed much more.

The equilibrium constant for the reaction of Eq. 9-97 will be:

$$\begin{aligned} & K_{eq}^{AB} = K_a^{HA} / K_a^{HB} = k_f / k_r \\ \text{or} \quad & \log K_{eq} = pK_a^{HA} - pK_a^{HB} = \log k_f - \log k_r \end{aligned} \qquad (9\text{-}98)$$

Here k_f and k_r refer to the rates in the forward and reverse directions. If the pK_a's of HA and HB are equal k_f and k_r will be the same, but if they are very far from equal the reaction will be slowed in one direction. If proton transfer is a step in an enzymatic reaction it may be slowed enough to become rate limiting.

The difference in pK_a values between the proton donor and the proton acceptor in Eq. 9-97 can be expressed as the Gibbs energy change which at 25°C is equal to $\pm 5.71 \times \Delta pK_a$. This is often referred to as the **thermodynamic barrier** $\Delta G°$ to a reaction and $\Delta G^‡$ can be expressed as the sum of the thermodynamic barrier $\Delta G°$ plus an **intrinsic barrier** $\Delta G^‡_{int}$. For the proton transfer of Eq. 9-97 the intrinsic barrier (for step *b*) is thought to be near zero so that $\Delta G^‡ \approx 5.71\ \Delta pK_a$.

From this we can conclude that two pK_a values can be as much as eight units apart and $\Delta G^{\ddagger}$ will still be less than 50 kJ / mol, low enough to permit rapid enzymatic reactions. However, for transfer of a proton from a C–H bond to a catalytic group, for example, to form an enolate ion for an aldol condensation (Chapter 13), the intrinsic barrier is known to be about 50 kJ / mol.[141] In this case, to allow rapid enzymatic reaction either the thermodynamic barrier must be very low, as a result of closely matching pK_a values, or the enzyme must lower the intrinsic barrier. It may do both.

Marcus theory. Discussion of intrinsic barriers is often approached using a quantitative theory proposed by Marcus.[142–145] It was first applied to electron transfer (Chapter 16) but has been used for a great variety of nonenzymatic and enzymatic reactions. As used by Gerlt and Gassman,[141] the Marcus formalism describes the reaction coordinate (Fig. 9-15A) as an inverted parabola whose shape is determined by the overall Gibbs energy change $\Delta G°$ and the intrinsic barrier $\Delta G^{\ddagger}_{int}$. The value of the Gibbs energy at any point on the curve is designated G and the reaction coordinate x is taken as 0 for the reactants and 1 for the products.

$$G = -4\,\Delta G^{\ddagger}_{int}\,(x - 0.5)^2 \;+\; \Delta G°\,(x - 0.5)$$
$$\text{Limits: } \Delta G°/4 \le \Delta G^{\ddagger}_{int} \le 4\,\Delta G° \qquad (9\text{-}99)$$

Differentiation of Eq. 9-99 yields the position of the transition state coordinate $x^{\ddagger}$ as follows:

$$x^{\ddagger} = 0.5 + \Delta G°/8\,\Delta G^{\ddagger}_{int} \qquad (9\text{-}100)$$

and it follows that

$$\Delta G^{\ddagger} = \Delta G^{\ddagger}_{int}\,(1 + \Delta G°/4\,\Delta G^{\ddagger}_{int})^2 \qquad (9\text{-}101)$$

Diffusion-controlled dissociation of protons. The direct proton transfer between C-1 and C-2 during the action of sugar isomerases may seem puzzling. How can a highly mobile proton remain attached to a group in the enzyme for a millisecond or more instead of being transferred out to a solvent molecule? This can mean that the enzyme promotes the transfer of a hydride ion or of a hydrogen atom rather than a proton (see Chapter 13). If so, the observed proton exchange with solvent would be an unimportant side reaction. On the other hand, could the group in the enzyme that removes the proton be out of contact with the aqueous medium and thus able to hold onto the proton more tightly? In recent years, it has been recognized that neither of these explanations may be necessary. An imidazole group is a likely proton-carrying group at many active sites and it is thought that a proton cannot be expected to transfer out from an imidazolium group with a rate constant greater than ~10^3 s^{-1}.

The argument is as follows.[146] The rate of donation of a proton from H_3O^+ to imidazole (reverse of Eq. 9-102) is known to be diffusion controlled with a rate constant of 1.5 x 10^{10} $M^{-1}\,s^{-1}$.

$$\mathrm{ImH^+ + H_2O} \rightleftharpoons [\mathrm{ImH^+\text{----}OH_2}] \underset{\substack{\uparrow \\ \text{Very fast diffusion controlled} \\ k \sim 1.5 \times 10^{10}\ \mathrm{M^{-1}\,s^{-1}}}}{\rightleftharpoons} \mathrm{Im + OH_3^+} \qquad (9\text{-}102)$$

The equilibrium constant for Eq. 9-102, calculated from the pK_a of 7.0 for imidazole, is 10^{-7} M. Since K_{eq} is also the *ratio of the overall rate constants for the forward and reverse reactions*, we see that for the forward reaction $k_f = 10^{-7} \times 1.5 \times 10^{10} = 1.5 \times 10^3\ s^{-1}$. This slow rate results from the fact that in the intermediate complex (in brackets in Eq. 9-102) the proton is on the imidazole group most of the time. For a small fraction of the time it is on the coordinated molecule H_2O but reverts to being on the imidazole many times before the imidazole and OH_3^+ separate (see also Eqs. 9-97 and 9-98). Because of this unfavorable equilibrium within the complex, the diffusion-controlled rate of proton transfer from a protonated imidazole to water is far less than for proton transfer in the reverse direction.

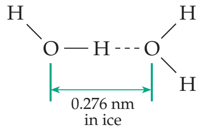
Coupled proton transfers. Enzymatic reactions often require the transfer of two or more protons. They may move individually, one proton at a time, or as in Eq. 9-93 they may move synchronously in a **coupled** or concerted process. Such coupled movement is generally not possible for heavier nuclei.[147] However, studies of solvent isotope effects using a **proton inventory** technique[111,148,149] have provided evidence favoring coupled proton transfers for a variety of enzymes. Movement of protons along hydrogen-bonded paths, as well as electron transfer, may take place with some participation of **quantum mechanical tunneling**.[150–152] Coupling to vibrational modes of the hydrogen-bonded protons may provide **vibration-assisted tunneling**.[153–154d] These reactions are associated with unusually large kinetic isotope effects.

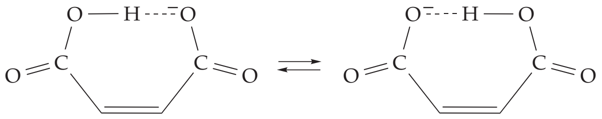
Unusually strong hydrogen bonds. The strength of a hydrogen bond is thought to be directly related to the length, which is ordinarily taken to be the distance between the two surrounding heavier atoms (see Chapter 2, Section B,3). Hydrogen bonds are sometimes classified on the basis of N–H---O distances[155,156] as:

<0.25 nm	very strong
0.25–0.265 nm	strong
>0.28 nm	weak
>0.37 nm	no van der Waals contact but electrostatic interaction still occurs

N–H---O distances may be a little longer than these.

The 0.276-nm hydrogen bonds in ice are regarded as moderately strong. However, if one of the oxygen atoms in an O–H---O hydrogen bond carries a negative charge, as in the maleate monoanion, it will be shorter.[156–158] Although the proton will be closer to one oxygen atom than to the other, it will be able to move between them by passing a very low transition state barrier. For this to occur the microscopic pK_a values for the two groups (when they are protonated) must be similar; for the maleate anion they are identical.

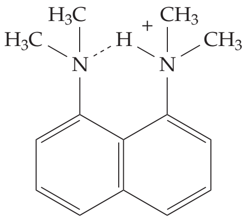




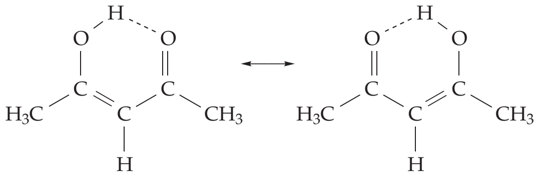
Monohydrogen maleate

The *strength* of a hydrogen bond can be measured for hydrogen-bonded complexes in the gas phase and range from 10–100 kJ / mol and even higher[155] for such complexes as $(FHF)^-$. It is more difficult to establish the strength in the liquid state or within the active site of an enzyme, but shorter hydrogen bonds are usually stronger than longer ones. Hydrogen bond distances in crystals of small molecules can be measured precisely by X-ray or neutron diffraction but greater uncertainty is present in distances within proteins. Lengths and probably strengths of hydrogen bonds in proteins can be measured by NMR methods.[158a,b] For example, solid-state NMR measurements on crystals of amino acids and other carboxylic acids have provided a plot of ^{1}H chemical shift vs hydrogen bond length.[159] For very short (0.24- to 0.25-nm) hydrogen bonds the ^{1}H chemical shift may be as great as 21 ppm. See also Fig. 3-30 and associated discussion. When dissolved in an ^1H_2O–^2H_2O mixture strongly hydrogen-bonded protons within a protein become enriched in ^{1}H. The ^{1}H / ^{2}H ratio of the hydrogen-bonded protons provides another measure of the hydrogen bond strength.[158c]

A short hydrogen bond is also present in such cations as the following:



The pK_a for dissociation of its proton is 12.3 and the hydrogen-bonded proton is probably located in the *center* of the bond with both amino groups sharing the charge.[160] Enols can also form unusually strong "resonance-assisted" hydrogen bonds:



The structure can be thought of as a resonance-stabilized enolate anion with a proton bound between the two oxygen atoms and equidistant from them.[156,157]

Do these short "**low-barrier hydrogen bonds**" have a special significance in enzymology? Proposals that they contribute to stabilization of transition states[161–166] have received some support[154,167] and aroused controversy.[158b,168–173] In later chapters we will examine specific enzymes in which low-barrier hydrogen bonds have been observed.

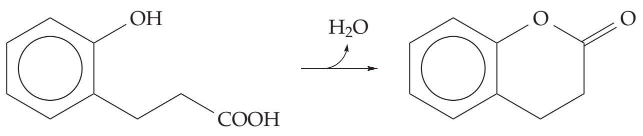
5. Covalent Catalysis

In addition to participating in acid–base catalysis, some amino acid side chains may enter into covalent bond formation with substrate molecules, a phenomenon that is often referred to as covalent catalysis.[174] When basic groups participate this may be called **nucleophilic catalysis**. Covalent catalysis occurs frequently with enzymes catalyzing nucleophilic displacement reactions and examples will be considered in Chapter 12. They include the formation of an acyl-enzyme intermediate by chymotrypsin (Fig. 12-11). Several of the coenzymes discussed in Chapters 14 and 15 also participate in covalent catalysis. These coenzymes combine with substrates to form reactive intermediate compounds whose structures allow them to be converted rapidly to the final products.

6. Proximity and Orientation

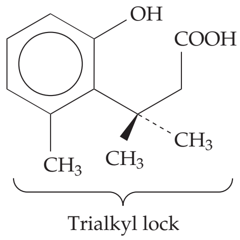
One of the earliest ideas about enzymes was that they simply brought reactants together and bound them side by side for a long enough time that the reactive groups might bump together and finally react. How important is this **proximity factor**? Page and Jencks estimated that rate enhancements by factors of 10^3 or more may be expected solely from the loss in the entropy of two reactants when they are bound in close proximity on an enzyme surface.[107,175,176] In view of the large entropy decrease involved, *the enthalpy of* *binding must be high*, and if this explanation is correct the *binding of the substrates* to the enzyme provides much of the driving force for catalysis. Westheimer described this by stating that enzymes use the substrate-binding force as an **entropy trap**.[108] The losses of translational and rotational entropy, which Page and Jencks estimated as up to −160 to −210 J /deg /mol, overcome the unfavorable entropy of activation that is usual in bimolecular reactions.

How precise must the orientation of substrates be for rapid reaction?[107,177,177a] Compounds such as the acid shown in Eq. 9-103 form an internal ester (lactone) spontaneously with elimination of water.



(9-103)

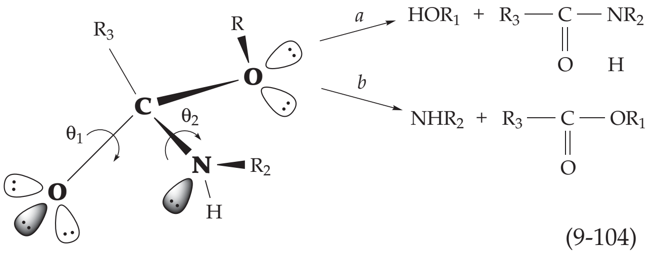
However, the following compound reacts at a rate over 10^{11} times that of the acid shown in Eq. 9-102. This is presumably because its conformation is highly restricted and the –COOH is constrained to frequently collide with the – OH group.[178,179] The three methyl



groups interdigitate and form a **trialkyl lock**. Orientation must also play a large role in enzymatic catalysis. As previously emphasized, enzymes often orient substrates precisely by formation of multiple hydrogen bonds. Small distortions by mutation or substitution of an essential metal by a different one can have very great effects. For example, because of an altered coordination pattern isocitrate dehydrogenases with Ca^{2+} in the active site has a maximum velocity of catalysis only 2.5 x 10^{-3} that with the normal Mg^{2+}-containing enzyme.[180] Bruice and associates concluded that enzymes must bring reacting groups into close proximity with orbitals of the reactants properly aligned in the ground state prior to moving to the transition state. They suggested that enzymes preorganize the enzyme-substrate complex into a **near attack conformation** in which the positions of reacting groups, the arrangements of hydrogen bonds, and the local dielectric constant in the active site, resemble those in the transition state. In this conformation the energy barrier to the transition state may be very low.[128-128b] Preorganization of the complex also acts to eliminate the slow components of solvent reorganization required for reaction in aqueous solutions.[128c,128d]

The necessity for reacting groups of substrates to collide with an orientation that allows productive interaction of electronic orbitals is often called a **stereoelectronic effect**. An example is the addition reaction of Eq. 9-74. The orbital of an unshared pair of electrons on the HO^- ion must be perpendicular to the plane of the double bond. Furthermore, if the proton becomes attached to the adjacent carbon in a synchronous or concerted manner it must enter from the opposite side, as it does in Eq. 9-74.

In many biochemical reactions an alcohol or amine is eliminated from a tetrahedrally bonded intermediate as in Eq. 9-104. Deslongchamps proposed a stereoelectronic theory[181,182] according to which elimination of either NH—R_2 or –O—R_1 from this intermediate will depend upon the values of the torsion angles θ_1 and θ_2 (Eq. 9-104). The theory asserts that for rapid elimination of OR_1 (Eq. 9-104*a*), unshared electron pairs on *both* the $-O^-$ and N atoms must be antiperiplanar to the bond being broken (see Figure 2-2). If rotation around the C–O bond (θ_1) occurs an orientation can be found in which an electron pair on each of the two oxygens will be antiperiplanar to the bond to NHR_2. From this orientation the latter will be eliminated (Eq. 9-104*b*). The theory has been supported by much experimental evidence involving reaction rates and product distribution among competing reactions of small conformationally restricted organic molecules.[181–184] Although more recent experiments[185] suggest that stereoelectronic factors may be of less significance than had been assumed, even a small decrease in transition state energy can be significant in an enzyme-catalyzed reaction. Enzymes may not only orient substrates in accord with stereoelectronic principles but also be able to promote conformational alterations in intermediates that allow them to take advantage of stereoelectronic factors. Examples are considered in Chapter 12.



(9-104)

7. The Microenvironment

A substrate bound at an active site may be in an environment quite different than that in an ordinary aqueous solution. In fact, the protein often surrounds the substrate to the extent that *the protein is the solvent*. What kind of **microenvironment** does the protein provide and can it assist in the catalytic process? Charged and dipolar groups of the protein provide an electrostatic field that provides part of the binding energy and which may also assist in catalysis, a concept expressed by Quastel as early as 1926.[183a] The ends of protein helices often seem to point at active centers.[186,187] In many cases more than one N terminus with its positive electrical potential (see Chapter 2) or more than one C terminus (negative potential) of a helix point to an active site. These helix dipoles may be important in stabilizing transition state structures[188] or in altering pK_a values of functional groups.[186] Fluctuations in charge distribution within a protein and in the hydrogen-bonding pattern of the protein with solvent and substrate may also be important.[189]

For some enzymatic reactions a transition state will be favored by a medium of very low dielectric constant and a correctly constructed active site can provide just such a surrounding.[189a] Hydrophobic groups may be packed around a site where an ion pair or other ionic interaction between enzyme and substrate occurs increasing the strength of that interaction. Conformational changes may enhance such effects. The dehydration of polar groups that must often occur upon binding of substrate may make these groups more reactive.[187,190,191] It has been suggested that the substantial volume changes ($\Delta V^‡$) that sometimes occur during formation of transition states for enzyme-catalyzed reactions may result largely from changes in hydration of groups on the enzyme surface and that these changes may play an important role in catalysis.

In the past most enzymologists tacitly assumed that the external medium in which enzymes act must be aqueous. However, many enzymes function well in media containing largely hydrocarbons. Enzymes in a dry, powdered form have been suspended directly in organic solvents.[192,193] Under these conditions, enzymes may contain only tightly bound “structural” water, together with less than one equivalent of a monolayer of water outside the protein. The enzymes often remain active and are able to catalyze reactions with an altered substrate specificity as well as different reactions overall.

8. Strain and Distortion

The fact that enzymes appear to bind their substrates in such a way as to surround and immobilize them means that something other than the kinetic energy of the substrate is needed to provide energy for the ES complex to pass over the transition state barrier. What is the source of this activation energy? As with nonenzymatic reactions, it must come ultimately from the translational energy of solvent or solute molecules bombarding the complex. Can enzymes act as “energy funnels” that effectively channel kinetic energy from spots on the enzyme surface to the active site?[194] This could either be through strictly mechanical movement or through induction by fluctuations of the electrical field.[195]

It is often suggested that enzymes assist in catalysis by distorting bond lengths or angles away from their normal values. If the distorted structure were closer to the transition state geometry than the undistorted structure, catalysis would be assisted (see *lysozyme*, Chapter 12). However, Levitt[196] concluded (see also Fersht[2]) that forces provided this way by a protein are small and that the protein would become distorted rather than a substrate. On the other hand, if binding of a substrate distorts the protein, could the resultant “stored energy” be used in some way to assist in catalysis? If binding of a substrate bends a “spring” (Section C,4), can the tension in the spring then be used to distort the enzyme to be more exactly complementary to the transition state? This concept has become popular.[2,113,197]

The amino acid side chains of proteins are usually well packed. However, neither the side chains nor the main chain are rigid and immobile. Some regions of the protein will contain empty spaces – packing defects. Lumry called these **mobile defects** because, as a result of fluctuations in side chain packing, they can move within a protein from one site to another for a considerable distance.[198] Much evidence, including X-ray studies at low temperatures,[199] supports the existence of many **conformational substates** in proteins. Some substates may bind substrates better than others and some may allow conversion to the transition state more readily than do others. Rotation of histidine rings, peptide linkages, – OH groups, or amide side chains[199a] may be required and has, in fact, been observed for some enzymes. Perhaps in one of the substates of an enzyme–substrate complex an especially favorable vibrational mode[200] leads the complex to the transition state. In the transition state the packing of side chains may be especially tight. A mobile defect present in the active site may have moved elsewhere. The binding of substrates to the protein in the transition state will also be tighter because of the conformational alterations that have occurred. The binding energy of the substrates is now being utilized to lower the transition state barrier. The substrate is literally squeezed into the transition state configuration.

Is it possible that the protein domains forming an active site act like **ferroelectric crystals**, which change their dipole moment in response to a change in electric field? The highly polarizable hydrogen-bonded network, the amide linkages, imidazole rings, guanidinium group, etc. of active sites may permit a flip-flop of the dipole moment as in domains of ferroelectric crystals.

In such crystals, e.g., those of the hydrogen-bonded KH_2PO_4, a 180° change of dipole-moment direction results from very small movements of heavy atoms together with larger movements of the hydrogen-bonded protons.[200a,b,c] Could a similar flip-flop in a protein domain be coupled to the passage of a substrate over the transition state barrier? I have not seen any discussion of this possibility, but the structure of protein domains would seem to allow it. One could also imagine that with a small change in the hydrogen-binding arrangement of protein groups an active site could become preorganized to favor a flip-flop along a hydrogen-bonded chain in a different direction for a subsequent step in a reaction sequence.

9. Why Oligomeric Enzymes?

As we have seen (Chapter 7), a large fraction of all proteins exist as dimers, trimers, and higher oligomers. Oligomeric proteins raise the osmotic pressure much less than would the same number of monomeric subunits and this may be crucial to a cell. Another advantage of oligomers may be reflected in the fact that active sites of enzymes are often at interfaces between two or more subunits. This may enhance the ability of enzymes to undergo conformational rearrangements that are required during their action, just as hemoglobin changes its oxygen affinity in concert with a change in inter-subunit contacts (Fig. 7-25).

A curious observation is that crystals of the dimeric pig heart malate dehydrogenase bind only one molecule of the substrate NAD^+ per dimer tightly; the second NAD^+ is bound weakly.[201,202] Similar *anticooperativity* has been reported for other crystalline dehydrogenases[203] and various other enzymes. An intriguing idea is that anticooperativity in binding might reflect a cooperative action between subunits during catalysis. Suppose that only conformation A of a dehydrogenase bound reduced substrate and NAD^+, while conformation B bound NADH and oxidized substrate. If reduced substrate and NAD^+ were present in excess and if oxidized substrate were efficiently removed from the scene by further oxidation, the following cooperative events could occur in the mixed AB dimer. The subunit of conformation A would bind substrates, react, and be converted to conformation B. At the same time, because of the strong AB interaction, the subunit that was originally in conformation B would be converted back to A and would be ready to initiate a new round of catalysis. Since conformation A has a low affinity for NADH, dissociation of the reduced coenzyme, which is often the slow step in dehydrogenase action, would be facilitated.[204–207] Such a **reciprocating** or **flip-flop** mechanism, suggested first by Harada and Wolfe,[204] is attractive because it provides a natural basis for the existence of the many known dimeric enzymes that do not exhibit evident allosteric properties. Attempts to verify this idea have largely failed. However, recent crystallographic studies of both glyceraldehyde phosphate dehydrogenase[206,207] and thymidylate synthase[208] are consistent with the proposal. There is still a possibility that coordinated reciprocal changes in distribution of electrical charges in the two subunits may also be important.

10. Summary

It appears that enzymes exert their catalytic powers by **first** bringing together substrates and binding them in proper orientations at the active site. **Second**, they often provide acidic and basic groups in the proper orientations to promote proton transfers within the substrate. **Third**, groups within the enzyme (especially nucleophilic groups) may enter into covalent interaction with the substrates to form structures that are more reactive than those originally present in the substrate. **Fourth**, the protein often closes around the substrate to immobilize it and to hold it in an environment, often lacking water, which could impede catalysis. The enzyme is also probably able to make small readjustments of its structure to provide good complementarity to the substrate at every stage and especially to the transition state structure. **Fifth**, the enzyme may be able to induce strain or distortion in the substrate perhaps accompanying a conformational change within the protein. The following question is often asked: “Why are enzymes such large molecules?”[209] At least part of the answer is that an enzyme usually interacts, sometimes via special domains, with other proteins.[206] Another part of the answer is evident when we consider that the formation of a surface complementary to that of the substrate and possessing reasonable rigidity requires a complex geometry in the peptide backbone. In addition, the enzyme must provide functional groups at the proper places to enter into catalysis. It may require a certain bulk to provide a low dielectric medium. Finally, if essential conformational changes occur during the course of the catalysis, we can only be surprised that nature has succeeded in packing so much machinery into such a small volume.

E. Classification of Enzymes

An official commission of the International Union of Biochemistry (IUB) has classified enzymes in the following six categories:[210]

1. *Oxidoreductases*. Enzymes catalyzing dehydrogenation or other oxidation and reduction reactions.

2. *Transferases*. These catalyze group transfer reactions.
3. *Hydrolases*. Enzymes catalyzing transfer of groups to the HO^- ion of H_2O.
4. *Lyases*. Enzymes promoting addition to double bonds or the reverse.
5. *Isomerases.* Enzymes catalyzing rearrangement reactions.
6. *Ligases* (synthetases). Enzymes that catalyze condensation with simultaneous cleavage of ATP and related reactions.

As an example, chymotrypsin is classified EC 3.4.4.5 according to the IUB system.[210] In this book a more mechanistically based classification is used. Because some official names are quite long, traditional trivial names for enzymes often have been retained. Remember that to be precise it is always necessary to mention the species from which an enzyme was isolated and, if possible, the strain. Also remember that almost every significant genetic difference is reflected in some change in some protein. It is possible that the enzyme you are working with is slightly different from the same enzyme prepared in a different laboratory.

References

1. Fromm, H. (1975) *Initial Rate Enzyme Kinetics*, Springer-Verlag, New York
2. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, Freeman, New York
3. Dixon, M., and Webb, E. C., eds. (1979) *Enzymes*, 3rd ed., Academic Press, New York
4. Wharton, C. W., and Eisenthal, R. (1981) *Molecular Enzymology*, Wiley, New York
5. Engel, P. C. (1982) *Enzyme Kinetics, the Steady-State Approach*, Chapman and Hall, London
6. Hammes, G. G. (1982) *Enzyme Catalysis and Regulation*, Academic Press, New York
7. Kull, F. J. (1994) *Principles of Biomolecular Kinetics and Binding*, CRC Press, Boca Raton, Florida
8. Suckling, C. J., ed. (1990) *Enzyme Chemistry*, 2nd ed., Chapman and Hall, New York
9. Price, N. C., and Stevens, L., eds. (1989) *Fundamentals of Enzymology*, 2nd ed., Oxford Univ. Press, Oxford,England
10. Kuby, S. A. (1990) *Enzymes: a Comprehensive Study*, CRC Press, Boca Raton, Florida
11. Schulz, A. R. (1994) *Enzyme Kinetics From Diastase to Multi-enzyme Systems*, Cambridge Univ. Press, New York
12. Segel, I. H. (1975) *Enzyme Kinetics:Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, Wiley, New York
13. Cornish-Bowden, A. (1995) *Fundamentals of Enzyme Kinetics*, 2nd ed., Portland Press, Brookfield, Vermont

13a. Purich, D. L., and Allison, R. D. (2000) *Trends Biochem. Sci.* **25**, 455

14. Lowry, O. H., and Passonneau, J. V. (1972) *A Flexible System of Enzymatic Analysis*, Academic Press, New York
15. Passonneau, J., and Lowry, O., eds. (1993) *Enzymatic Analysis*, 1st ed., Humana Press, Totowa, New Jersey
16. Liu, Y.-M., and Sweedler, J. V. (1995) *J. Am. Chem. Soc.* **117**, 8871–8872
17. Liébecq, C., ed. (1992) *Biochemical Nomenclature*, Portland Press, London and Chapel Hill, North Carolina (for the International Union of Biochemistry and Molecular Biology)
18. Caldin, E. F. (1964) *Fast Reactions in Solution*, Wiley, New York
19. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology*, Dover, Mineola, New York (p. 393)
20. Glick, N., Landman, A. D., and Roufogalis, B. D. (1979) *Trends Biochem. Sci.* **4**, N82 – N83

20a. Wilkinson. (1980) *Biochem. J.* **80**, 324 –

21. Orsi, B. A., and Tipton, K. E. (1979) *Methods Enzymol.* **63**, 159–183
22. Wharton, C. W., and Szawelski, R. J. (1982) *Biochem. J.* **203**, 351–360
23. Cleland, W. W. (1979) *Methods Enzymol.* **63**, 103–138
24. Frieden, C. (1994) *Methods Enzymol.* **240**, 311–322
25. Schechter, A. N. (1970) *Science* **170**, 273–280
26. Cha, S. (1970) *J. Biol. Chem.* **245**, 4814–4818
27. Martinez, M. B., Flickinger, M. C., and Nelsestuen, G. L. (1996) *Biochemistry* **35**, 1179–1186
28. Bull, H. B. (1971) *An Introduction to Physical Biochemistry*, 2nd ed., Davis Co., Philadelphia, Pennsylvania
29. Eisenstein, B. I. (1987) *Escherichia coli and Salmonella typhimurium*, Am. Soc. Microbiology, Washington, D.C., FC Niedhardt, ed. (pp.84–90)
30. Smoluchowski, M. (1917) *Z. Phys. Chem* **92**, 129–168
31. Caldin, E. F. (1964) *Fast Reactions in Solution*, Wiley, New York (pp. 10 and 279)
32. Debye, P. (1942) *Trans. Electrochem. Soc.* **82**, 265–272

32a. Elcock, A. H., Huber, G. A., and McCammon, J. A. (1997) *Biochemistry* **36**, 16049–16058

32b. Selzer, T. and Schreiber, G. (1999) *J. Mol. Biol.* **287**, 409–419

33. Mastro, A. M., Babich, M. A., Taylor, W. D., and Keith, A. D. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3414–3418
34. French, D. (1957) *Brewers Digest* **32**, 50–56
35. Cohn, E. J., and Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, Van Nostrand-Reinhold, Princeton, New Jersey (pp. 90–93)
36. Koenig, S. H. (1975) *Biopolymers* **14**, 2421–2423
37. Blacklow, S. C., Raines, R. T., Lim, W. A., Zamore, P. D., and Knowles, J. R. (1988) *Biochemistry* **27**, 1158–1167
38. Albery, W. J., and Knowles, J. R. (1976) *Biochemistry* **15**, 5627–5631
39. Albery, W. J., and Knowles, J. R. (1976) *Biochemistry* **15**, 5631–5640
40. Jenkins, W. T. (1982) *Adv. Enzymol.* **53**, 307–344
41. Haldane, J. B. S. (1930) *Enzymes*, Longmans, Green, New York
42. Cleland, W. W. (1970) in *The Enzymes*, 3rd ed., Vol. 2 (Boyer, P. D., ed), Academic Press, New York
43. Dalziel, K. (1957) *Acta Chem. Scand.* **11**, 1706–1723
44. Rebholz, K. L., and Northrop, D. B. (1994) *Arch. Biochem. Biophys.* **312**, 227–233
45. Volkenstein, M. V., and Goldstein, B. N. (1966) *Biokhim.* **31**, 541–547
46. Volkenstein, M. V., and Goldstein, B. N. (1966) *Biochim. Biophys. Acta.* **115**, 471–477
47. Fromm, H. J. (1970) *Biochem. Biophys. Res. Commun.* **40**, 692–697
48. Seshagiri, N. (1972) *J. Theor. Biol.* **34**, 469–486
49. Huang, C. Y. (1979) *Methods Enzymol.* **63**, 54–84
50. Chou, K., and Forsén, S. (1980) *Biochem. J.* **187**, 829–835
51. Fromm, H. J., Silverstein, E., and Boyer, P. D. (1964) *J. Biol. Chem.* **239**, 3645–3652
52. Fromm, S. J., and Fromm, H. J. (1999) *Biochem. Biophys. Res. Commun.* **265**, 448–452
53. Silverstein, E., and Boyer, P. D. (1964) *J. Biol. Chem.* **239**, 3901–3907
54. Hurst, R. O. (1969) *Can. J. Biochem. Physiol.* **47**, 941–944

54a. Cha, S. (1968) *J. Biol. Chem.* **243**, 820–825

55. Srere, P. A. (1967) *Science* **158**, 936–937
56. Hiromi, K., ed. (1979) *Kinetics of Fast Enzyme Reaction-Theory and Practice*, John Wiley, New York
57. Rentzepis, P. M. (1978) *Science* **202**, 174–182
58. Hammes, G. G., and Schimmel, P. R. (1970) in *The Enzymes*, 3rd ed., Vol. 2 (Boyer, P. D., ed), pp. 67–114, Academic Press, New York
59. Ahrens, M.-L., Maass, G., Schuster, P., and Winkler, H. (1970) *J. Am. Chem. Soc.* **92**, 6134–6139

59a. Gutfreund, H. (1999) *Trends Biochem. Sci.* **24**, 457–460

59b. Beechem, J. M. (1998) *Biophys. J.* **74**, 2141

60. Fink, A. L. (1979) *Trends Biochem. Sci.* **4**, 8–10
61. Balashov, S. P., Imasheva, E. S., Ebrey, T. G., Chen, N., Menick, D. R., and Crouch, R. K. (1997) *Biochemistry* **36**, 8671–8676
62. Fink, A. L., and Geeves, M. A. (1979) *Methods Enzymol.* **63**, 336–370
63. Alberty, R. A. (1956) *Adv. Enzymol.* **17**, 1–64
64. Dixon, M., and Webb, E. C., eds. (1979) *Enzymes*, 3rd ed., Academic Press, New York (pp.138–163)
65. Cleland, W. W. (1977) *Adv. Enzymol.* **45**, 273–387
66. Cleland, W. W. (1982) *Methods Enzymol.* **87**, 390–405
67. Kyte, J. (1995) *Mechanism in Protein Chemistry*, Garland Publ., New York
68. Ono, S., Hiromi, K., and Yashikawa, Y. (1958) *Bull. Chem. Soc. Jap.* **31**, 957–962
69. Brocklehurst, K., and Dixon, H. B. F. (1976) *Biochem. J.* **155**, 61–70
70. Tipton, K. F., and Dixon, H. B. F. (1979) *Methods Enzymol.* **63**, 183–234
71. Dixon, M. (1953) *Biochem. J.* **55**, 170–171
72. Purich, L. D., and Fromm, H. J. (1972) *Biochim. Biophys. Acta.* **268**, 1–3
73. Williams, J. W., and Morrison, J. F. (1979) *Methods Enzymol.* **63**, 437–467
74. Morrison, J. F. (1982) *Trends Biochem. Sci.* **7**, 102–105
75. Frieden, C. (1964) *J. Biol. Chem.* **239**, 3522–3531

References

76. Morrison, J. F., and Walsh, C. T. (1988) *Adv. Enzymol.* **61**, 201–301
77. Gutheil, W. G., and Bachovin, W. W. (1993) *Biochemistry* **32**, 8723–8731
78. Di Cera, E., Hopfner, K.-P., and Dang, Q. D. (1996) *Biophys. J.* **70**, 174–181
79. Botts, J., and Morales, M. (1953) *Trans. Faraday Soc.* **49**, 696–707
80. Monod, J., Wyman, J., and Changeux, J. D. (1965) *J. Mol. Biol.* **12**, 88–118
81. Plapp, B. V., Leidal, K. G., Smith, R. K., and Murch, B. P. (1984) *Arch. Biochem. Biophys.* **230**, 30–38
82. Chadha, V. K., Leidal, K. G., and Plapp, B. V. (1983) *Journal of Medicinal Chemistry* **26**, 916–922
83. Cho, H., Ramaswamy, S., and Plapp, B. V. (1997) *Biochemistry* **36**, 382–389
84. Westley, A. M., and Westley, J. (1996) *J. Biol. Chem.* **271**, 5347–5352
85. Hervé, G., ed. (1989) *Allosteric Enzymes*, CRC Press, Boca Raton, Florida
86. Rubin, M. M., and Changeux, J. D. (1966) *J. Mol. Biol.* **21**, 265–274
87. Grant, G. A., Schuller, D. J., and Banaszak, L. J. (1996) *Protein Sci.* **5**, 34–41
88. Williams, R. O., Young, J. R., and Majiwa, P. A. O. (1979) *Nature (London)* **282**, 847–849
89. Rabin, B. R. (1967) *Biochem. J.* **102**, 22c
90. Newsholme, E. A., and Start, C. (1973) *Regulation in Metabolism*, Wiley, New York
91. Walsh, C. (1979) *Enzymatic Reaction Mechanism*, Freeman, San Francisco, California
91a. Plapp, B. V. (1982) *Methods Enzymol.* **87**, 469–499
92. Sandler, M., ed. (1980) *Enzyme Inhibitors as Drugs*, Univ. Park Press, Baltimore, Maryland
93. Bey, P. (1981) *Chem. Ind. (London)*, 139–144
94. Hanson, K. R. (1966) *J. Am. Chem. Soc.* **88**, 2731–2742
94a. Prelog, V., and Helmchen, G. (1982) *Angew. Chem. Int. Ed. Engl.* **21**, 567–583
95. Bentley, R. (1969) *Molecular Asymmetry in Biology*, Vol. 1, Academic Press, New York (pp. 49–56)
96. Bentley, R. (1970) *Molecular Asymmetry in Biochemistry*, Vol. 2, Academic Press, New York
97. Alworth, W. L. (1972) *Stereochemistry and its Application in Biochemistry*, Wiley-Interscience, New York
98. Bentley, R. (1978) *Nature (London)* **276**, 673–676
99. Barry, J. M. (1997) *Trends Biochem. Sci.* **22**, 228–230
100. Hirschmann, H. (1960) *J. Biol. Chem.* **235**, 2762–2767
101. Post, C. B., and Ray, W. J. J. (1995) *Biochemistry* **34**, 15881–15885
102. Tsou, C.-L. (1986) *Trends Biochem. Sci.* **11**, 427–429
103. Ferscht, A. R. (1980) *Trends Biochem. Sci.* **5**, 262–265
104. Fersht, A. R. (1998) *Science* **280**, 541
105. Nureki, O., Vassylyev, D. G., Tateno, M., Shimada, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T. L., Schimmel, P., and Yokoyama, S. (1998) *Science* **280**, 578–582
106. Hopfield, J. J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5248–5252
107. Jencks, W. P. (1975) *Adv. Enzymol.* **43**, 219–410
108. Westheimer, F. H. (1962) *Adv. Enzymol.* **24**, 441–482
109. Kraut, J. (1988) *Science* **242**, 533–539
110. Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., Freeman, New York
111. Gandour, R. D., and Schowen, R. L., eds. (1978) *Transition States of Biochemical Processes*, Plenum, New York
112. Laidler, K. J. (1969) *Theories of Chemical Reaction Rates*, McGraw-Hill, New York
113. Kyte, J. (1995) *Mechanism in Protein Chemistry*, Garland Publ., New York (pp. 199–213)
114. Marcus, R. A. (1992) *Science* **256**, 1523–1524
115. Pauling, L. (1948) *Nature (London)* **161**, 707–713
116. Linehard, G. E. (1973) *Science* **180**, 149–154
117. Wolfenden, R. (1976) *Annu Rev Biophys Bioeng.* **5**, 271–306
117a. Schramm, V. L. (1998) *Ann. Rev. Biochem.* **67**, 693–720
118. Nielsen, P. A., Glad, S. S., and Jensen, F. (1996) *J. Am. Chem. Soc.* **118**, 10577–10583
119. Zheng, Y.-J., and Bruice, T. C. (1997) *J. Am. Chem. Soc.* **119**, 8137–8145
120. Merkler, D. J., Kline, P. C., Weiss, P., and Schramm, V. L. (1993) *Biochemistry* **32**, 12993–13001
121. Degano, M., Almo, S. C., Sacchettini, J. C., and Schramm, V. L. (1998) *Biochemistry* **37**, 6277–6285
122. Glad, S. S., and Jensen, F. (1997) *J. Am. Chem. Soc.* **119**, 227–232
123. Schramm, V. L., Horenstein, B. A., and Kline, P. C. (1994) *J. Biol. Chem.* **269**, 18259–18262
124. Kline, P. C., and Schramm, V. L. (1994) *J. Biol. Chem.* **269**, 22385–22390
125. Deng, H., Kurz, L. C., Rudolph, F. B., and Callender, R. (1998) *Biochemistry* **37**, 4968–4976
125a. Cho, Y.-K., and Northrop, D. B. (1999) *Biochemistry* **38**, 7470–7475
125b. Wolfenden, R., Snider, M., Ridgway, C., and Miller, B. (1999) *J. Am. Chem. Soc.* **121**, 7419–7420
126. Menger, F. M. (1992) *Biochemistry* **31**, 5368–5373
127. Goldsmith, J. O., and Kuo, L. C. (1993) *J. Biol. Chem.* **268**, 18481–18484
128. Lightstone, F. C., and Bruice, T. C. (1996) *J. Am. Chem. Soc.* **118**, 2595–2605
128a. Bruice, T. C., and Lightstone, F. C. (1999) *Acc. Chem. Res.* **32**, 127–136
128b. Bruice, T. C., and Benkovic, S. J. (2000) *Biochemistry* **39**, 6267–6274
128c. Cannon, W. R., Singleton, S. F., and Benkovic, S. J. (1996) *Nature Struct. Biol.* **3**, 821–833
128d. Cannon, W. R., and Benkovic, S. J. (1998) *J. Biol. Chem.* **273**, 26257–26260
129. Murphy, D. J. (1995) *Biochemistry* **34**, 4507–4510
129a. Warshel, A. (1998) *J. Biol. Chem.* **273**, 27035–27038
129b. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology*, Dover, Mineola, New York (pp. 210–213)
130. Hucho, F., and Wallenfels, K. (1971) *Eur. J. Biochem.* **23**, 489–496
131. Jencks, W. P. (1969) *Catalysis in Chemistry and Enzymology*, McGraw-Hill, New York (pp. 170–199)
132. Swain, C. G., and Brown, J. F., Jr. (1952) *J. Am. Chem. Soc.* **74**, 2534–2537 and 2538–2543
133. Rony, P. R. (1969) *J. Am. Chem. Soc.* **91**, 6090–6096
134. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology*, Dover, Mineola, New York (p. 199)
135. Engdahl, K.-Å., Bivehed, H., Ahlberg, P., and Saunders, W. H., Jr. (1983) *J. Am. Chem. Soc.* **105**, 4767–4774
136. Hegarty, A. F., and Jencks, W. P. (1975) *J. Am. Chem. Soc.* **97**, 7188–7189
136a. Williams, A. (1999) *Concerted Organic and Bio-Organic Mechanisms*, CRC Press, Boca Raton, Florida
137. Wang, J. H. (1968) *Science* **161**, 328–334
138. Eigen, M. (1964) *Angew. Chem. Int. Ed. Engl.* **3**, 1–19
139. Metzler, D. E. (1979) *Adv. Enzymol.* **50**, 1–40
140. Nagle, J. F., and Morowitz, H. J. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 298–302
141. Gerlt, J. A., and Gassman, P. G. (1993) *J. Am. Chem. Soc.* **115**, 11552–11568
142. Cohen, A. O., and Marcus, R. A. (1968) *J. Phys. Chem.* **72**, 4249–4256
143. Marcus, R. A. (1969) *J. Am. Chem. Soc.* **91**, 7224–7225
144. Guthrie, J. P. (1996) *J. Am. Chem. Soc.* **118**, 12878–12885
145. Guthrie, J. P. (1996) *J. Am. Chem. Soc.* **118**, 12886–12890
146. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology*, Dover, Mineola, New York (pp. 207–213)
147. Dewar, M. J. S. (1984) *J. Am. Chem. Soc.* **106**, 209–219
148. Gandour, R. D., Maggiora, G. M., and Schowen, R. L. (1974) *J. Am. Chem. Soc.* **96**, 6967–6979
149. Quinn, D. M. (1987) *Chem. Rev.* **87**, 955–979
150. Klinman, J. P. (1989) *Trends Biochem. Sci.* **14**, 368–373
151. Rucker, J., Cha, Y., Jonsson, T., Grant, K. L., and Klinman, J. P. (1992) *Biochemistry* **31**, 11489–11499
152. Hwang, J.-K., and Warshel, A. (1996) *J. Am. Chem. Soc.* **118**, 11745–11751
153. Barbara, P. F., Walker, G. C., and Smith, T. P. (1992) *Science* **256**, 975–981
154. Kearley, G. J., Fillaux, F., Baron, M.-H., Bennington, S., and Tomkinson, J. (1994) *Science* **264**, 1285–1289
154a. Antoniou, D., and Schwartz, S. D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12360–12365
154b. Basran, J., Sutcliffe, M. J., and Scrutton, N. S. (1999) *Biochemistry* **38**, 3218–3222
154c. Sutcliffe, M. J., and Scrutton, N. S. (2000) *Trends Biochem. Sci.* **25**, 405–408
154d. Huang, Y., Rettner, C. T., Auerbach, D. J., and Wodtke, A. M. (2000) *Science* **290**, 111–114
155. Hibbert, F., and Emsley, J. (1990) *Adv. Phys. Org. Chem.* **26**, 255–379
156. Gilli, P., Bertolasi, V., Ferretti, V., and Gilli, G. (1994) *J. Am. Chem. Soc.* **116**, 909–915
157. Perrin, C. L. (1994) *Science* **266**, 1665–1668
158. Garcia-Viloca, M., González-Lafont, A., and Lluch, J. M. (1997) *J. Am. Chem. Soc.* **119**, 1081–1086
158a. Bao, D., Huskey, W. P., Kettner, C. A., and Jordan, F. (1999) *J. Am. Chem. Soc.* **121**, 4684–4689
158b. Garcia-Viloca, M., Gelabert, R., González-Lafont, A., Moreno, M., and Lluch, J. M. (1998) *J. Am. Chem. Soc.* **120**, 10203–10209
158c. Bowers, P. M., and Klevit, R. E. (2000) *J. Am. Chem. Soc.* **122**, 1030–1033
159. McDermott, A., and Ridenour, C. F. (1996) in *Encyclopedia of NMR*, (Emsley, J. W., ed.), pp. 3820–3824, Wiley, Sussex, England
160. Frey, P. A. (1995) *Science* **269**, 104–106
161. Cleland, W. W., Frey, P. A., and Gerlt, J. A. (1998) *J. Biol. Chem.* **273**, 25529–25532
162. Cleland, W. W., and Kreevoy, M. M. (1994) *Science* **264**, 1887–1890
163. Frey, P. A., Whitt, S. A., and Tobin, J. B. (1994) *Science* **264**, 1927–1930
164. Cassidy, C. S., Lin, J., and Frey, P. A. (1997) *Biochemistry* **36**, 4576–4584
165. Halkides, C. J., Wu, Y. Q., and Murray, C. J. (1996) *Biochemistry* **35**, 15941–15948
166. Zheng, Y.-J., and Bruice, T. C. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4285–4288
167. Richard, J. P. (1998) *Biochemistry* **37**, 4305–4309
168. Scheiner, S., and Kar, T. (1995) *J. Am. Chem. Soc.* **117**, 6970–6975
169. Warshel, A., Papazyan, A., Kollman, P. A., Cleland, W. W., Kreevoy, M. M., and Frey, P. A. (1995) *Science* **269**, 102–106
170. Shan, S.-o, Loh, S., and Herschlag, D. (1996) *Science* **272**, 97–101
171. Ash, E. L., Sudmeier, J. L., De Fabo, E. C., and Bachovchin, W. W. (1997) *Science* **278**, 1128–1132

References

172. Shan, S.-o, and Herschlag, D. (1996) *J. Am. Chem. Soc.* **118**, 5515–5518
173. Warshel, A., and Papazyan, A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13665–13670
174. Spector, L. B. (1982) *Covalent Catalysis by Enzymes*, Springer-Verlag, New York
175. Page, M. I., and Jencks, W. P. (1971) *Proc. Natl. Acad. Sci. U.S.A.* **68**, 1678–1683
176. Kirsch, J. F. (1973) *Ann. Rev. Biochem.* **42**, 205–234
177. Koshland, D. E., Jr., Carraway, K. W., Dafforn, G. A., Goss, J. D., and Storm, D. R. (1971) *Cold Spring Harbor Symp. Quant. Biol.* **36**, 13–20

177a. Villà, J., Strajbl, M., Glennon, T. M., Sham, Y. Y., Chu, Z. T., and Warshel, A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11899–11904

178. Milstien, S., and Cohen, L. A. (1972) *J. Am. Chem. Soc.* **94**, 9158–9165
179. Karle, J. M., and Karle, I. L. (1972) *J. Am. Chem. Soc.* **94**, 9182–9189
180. Mesecar, A. D., Stoddard, B. L., and Koshland, D. E., Jr. (1997) *Science* **277**, 202–206
181. Deslongchamps, P., Lebreux, C., and Taillefer, R. (1973) *Can. J. Chem.* **51**, 1665–1669
182. Deslongchamps, P. (1975) *Tetrahedron* **31**, 2463–2490
183. Perrin, C. L., and Arrhenius, M. L. (1982) *J. Am. Chem. Soc.* **104**, 2839–2842
184. Evans, C. M., Glenn, R., and Kirby, A. J. (1982) *J. Am. Chem. Soc.* **104**, 4706–4707
185. Kuo, L. C., Fukuyama, J. M., and Makinen, M. W. (1983) *J. Mol. Biol.* **163**, 63–105
186. Plogman, J. H., Drenth, G., Kalk, K. H., and Hol, W. G. J. (1979) *J. Mol. Biol.* **127**, 149–162
187. Warshel, A. (1981) *Biochemistry* **20**, 3167–3177
188. Hol, W. G. J., van Duijen, P. T., and Berendsen, H. J. C. (1978) *Nature (London)* **273**, 443–446
189. Welch, G. R., ed. (1986) *The Fluctuating Enzyme*, Wiley, New York

189a. Mertz, E. L., and Krishtalik, L. I. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2081–2086

190. Dewar, M. J. S., and Storch, D. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2225–2229
191. Dewar, M. J. S., and Dieter, K. M. (1988) *Biochemistry* **27**, 3302–3308
192. Yennawar, H. P., Yennawar, N. H., and Farber, G. K. (1995) *J. Am. Chem. Soc.* **117**, 583–585
193. Schmitke, J. L., Wescott, C. R., and Klibanov, A. M. (1996) *J. Am. Chem. Soc.* **118**, 3360–3365
194. Kell, D. B. (1982) *Trends Biochem. Sci.* **7**, 1
195. Robertson, B., and Astumian, R. D. (1992) *Biochemistry* **31**, 138–141
196. Levitt, M. (1974) in *Peptides, Polypeptides and Proteins* (Blout, E. R., Bouey, F. A., Goodman, M., and Lotan, N., eds), p. 99, Wiley, New York
197. Avis, J. M., and Fersht, A. R. (1993) *Biochemistry* **32**, 5321–5326
198. Lumry, R., and Rosenberg, A. (1974) *Colloques Internationaux du C.N.R.S.* **246**, 53–62
199. Rader, S. D., and Agard, D. A. (1997) *Protein Sci.* **6**, 1375–1386

199a. Word, J. M., Lovell, S. C., Richardson, J. S., and Richardson, D. C. (1999) *J. Mol. Biol.* **285**, 1735–1747

200. Bialek, W., and Onuchic, J. N. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5908–5912

200a. Bystrov, D. S., and Popova, E. A. (1987) *Ferroelectrics* **72**, 147–155

200b. Emsley, J., Reza, N. M., Dowes, H. M., Hursthouse, M. B., and Kurodo, R. (1988) *Phosphorus and Sulfur* **35**, 141–149

200c. Rakvin, and Dalal. (1992) *Ferroelectrics* **135**, 227–236

201. Banaszak, L. A., and Bradshaw, R. A. (1975) in *The Enzymes*, 3rd ed., Vol. 11 (Boyer, P. D., ed), Academic Press, New York (pp. 369–396)
202. Zimmerle, C. T., and Alter, G. M. (1993) *Biochemistry* **32**, 12743–12748
203. Niefind, K., Hecht, H.-J., and Schomburg, D. (1995) *J. Mol. Biol.* **251**, 256–281
204. Harada, K., and Wolfe, R. G. (1968) *J. Biol. Chem.* **243**, 4131–4137
205. Lazdunski, M., Petitclerc, C., Chappelet, D., and Lazdunski, C. (1971) *Eur. J. Biochem.* **20**, 124–139
206. Skarzynski, T., Moody, P. C. E., and Wonacott, A. J. (1987) *J. Mol. Biol.* **193**, 171–187
207. Yun, M., Park, C.-G., Kim, J.-Y., and Park, H.-W. (2000) *Biochemistry* **39**, 10702–10710
208. Anderson, A. C., O'Neil, R. H., DeLano, W. L., and Stroud, R. M. (1999) *Biochemistry* **38**, 13829–13836
209. Srere, P. A. (1984) *Trends Biochem. Sci.* **9**, 387–390
210. Nomenclature Committee of the International Union of Biochemistry and Molecular Biology. (1992) *Enzyme Nomenclature*, Academic Press, San Diego, California

Study Questions

1. Give two reasons why enzymes are important to living organisms.

2. If an enzyme catalyzes the reaction $A \rightarrow B + C$, will it also catalyze the reaction $B + C \rightarrow A$?

3. Can an enzyme use a site to catalyze a reverse reaction different from what it uses to catalyze the reaction in the forward direction? Explain your answer.

4. Define the "steady state" of living cells and contrast with a state of chemical equilibrium.

5. How is the instantaneous initial velocity of an enzymatic reaction measured? What precautions must be taken to ensure that a true instantaneous velocity is being obtained?

6. How is the rate of an enzyme action influenced by changes in:
   a) temperature
   b) enzyme concentration
   c) substrate concentration
   d) pH

7. In what ways are enzymatic reactions typical of ordinary catalytic reactions in organic or inorganic chemistry, and in what ways are they distinct?

8. In general, for a reaction that can take place with or without catalysis by an enzyme, what is the effect, if any, of the enzyme on
   a) standard Gibbs energy change of the reaction
   b) energy of activation of the reaction
   c) initial velocity of the reaction
   d) temperature coefficient of the rate constant

9. What are the dimensions of the rate constant for zero-, first-, and second-order reactions? If a first-order reaction is half completed in 2 min, what is its rate constant?

10. A sample of hemin containing radioactive iron, ^{59}Fe, was assayed with a Geiger-Müller counter at intervals of time with the following results:

Time (days)	Radioactive counts per min
0	3981
2	3864
6	3648
10	3437
14	3238
20	2965

Determine the half-life of ^{59}Fe and the value for the decay constant.

11. The kinetics of the aerobic oxidation of enzymatically reduced nicotinamide adenine dinucleotide (NADH) have been investigated at pH 7.38 at 30°C. The reaction rate was followed spectrophotometrically by measuring the decrease in absorbance at 340 nm over a period of 30 min. The reaction may be represented as

$$\text{NADH} + \text{H}^+ + \text{riboflavin} \rightarrow \\ \text{NAD}^+ + \text{dihydroriboflavin}$$

Time (days)	A (at 340 nm)
1	0.347
2	0.339
5	0.327
9	0.302
16.5	0.275
23	0.254
27	0.239
30	0.229

Determine the rate constant and the order of the reaction.

12. You have isolated and purified a new enzyme (E) which converts a single substrate (S) into a single product (P). You have determined M_r by gel filtration as ~ 46,400. However, in SDS gel electrophoresis, a molecular mass of ~ 23 kDa was indicated for the single protein band observed. A solution of the enzyme was analyzed in the following way. The absorbance at 280 nm was found to be 0.512. A 1.00 ml portion of the same solution was subjected to amino acid analysis and was found to contain 71.3 nmol of tryptophan. N-terminal analysis on the same volume of enzyme revealed 23.8 nmol of N-terminal alanine.

   a) What is the approximate molecular mass of the enzyme? Discuss this answer. Be sure to use an appropriate number of significant figures in this and other calculations.
   b) What is the concentration of your enzyme in moles per liter of active sites?
   c) What is the molar extinction coefficient ε at 280 nm where $A = \varepsilon cl$; A = absorbance, c = mol / liter, and l = cell width in cm. Assume that all spectrophotometric measurements are made in 1.00 cm cuvettes.

   A second preparation of the enzyme had an absorbance at 280 nm of 0.485. This enzyme was diluted very carefully: 1.00 ml into 250 ml and this diluted enzyme was used for the following experiments (I to III).

Study Questions

Experiment I. A 1.00 ml portion of the diluted enzyme was added to 250 ml of buffered substrate at pH 7.0 and was mixed rapidly. The resulting initial substrate concentration $[S]_o$ was 1.000 mM. This reaction mixture was held at 25.0°C and portions were removed periodically at time t for analysis of the product P formed. The results follow. Plot [P] vs. time.

Time t (s)	[P] (mM)	Time t (s)	Remaining [S] (mM)
200	0.104	2800	0.070
400	0.208	3200	0.040
800	0.392	3600	0.022
1200	0.554	4400	0.0060
1600	0.695	6000	0.00048
2000	0.800		
2400	0.881		

The integrated rate equation corresponding to the Michaelis–Menten equation

$v = V_{max}\,[S]\,/\,(K_m + [S])$ is as follows:

$V_{max} \cdot t = [S]_o - [S] + K_m \ln([S]_o/[S])$

d) Plot $([S]_o - [S])/t$ vs. $1/t \times \ln([S]_o/[S])$. From this plot evaluate K_m and V_{max}. Make this and other plots to appropriate scale on good quality graph paper.

e) What is k_{cat}?

Experiment II. In a second experiment, a series of test tubes were set up, each containing a different amount of buffered substrate at pH 7 but each in a volume of exactly 4.00 ml. The same enzyme solution used in part d (absorbance at 280 nm = 0.485) was diluted 2.00 ml in 250 ml as in I, then again 2.00 ml in 200 ml. Portions of 1.00 ml of this twice diluted enzyme were added at $t = 0$ to each of the test tubes of buffered substrate. The reaction was stopped in just 10.0 minutes by adding perchloric acid; a suitable reagent was added to provide for a colorimetric determination of the product. The results were as follows:

[S] (mM)	Amount of product (μmol/tube)	[S] (mM)	Amount of product (μmol/tube)
10.0	2.29	0.600	1.31
5.00	2.18	0.400	1.07
2.50	2.00	0.200	0.686
1.200	1.69	0.100	0.400
0.800	1.48		

f) Plot $1/v$ vs. $1/[S]$ where v is in units of μmol per tube and [S] in millimoles / liter. Evaluate K_m, V_{max}, and k_{cat} from this plot. Fit your data by eye. It is generally agreed that a linear least squares fit is inappropriate unless suitable weighting factors are used.

g) Plot the same data as $v/[S]$ vs. v. Again evaluate K_m and V_{max}.

Experiment III. The preceding experiment was repeated but an inhibitor was present in each tube in a concentration equal to 5.00 mM, 10.0 mM, or 25.0 mM. Two different inhibitors were used, A and B. The following results were obtained.

μmol of product formed in 10 min/tube

[S] (mM)	[I] = 5.00 mM Inhibitor: A	5.00 mM B	25.0 mM Inhibitor: A	10.0 mM B
5.00	2.00	1.09	1.50	0.727
2.50	0.71	1.00	1.09	0.667
1.20	1.31	0.848	0.686	0.565
0.800	1.07	0.739	0.505	0.492
0.600	0.900	0.654	0.400	0.436
0.400	0.686	0.533	0.282	0.356
0.200	0.400	0.343	0.150	0.229
0.100	0.218	0.200	0.077	0.133

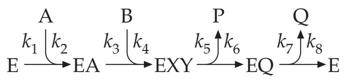
h) Plot $1/v$ vs. $1/[S]$ for each of these sets of data. For each case evaluate V_{max}, apparent K_m, and inhibitor constants $K_I = [I][E]/[EI]$. There may be two K_I values.

i) For uninhibited enzymes, when [S] = 0.4 mM what fraction of the enzyme is ES? Free E?

j) For enzyme in the presence of 10.00 mM inhibitor and 0.8 mM substrate, what fraction of the total enzyme is ES? EI? free E?

k) Recall that $v = [ES]\,k_{cat} = [E][S] \cdot k_{cat}/K_m$ in many cases. What will be the relative rates of product formation from your substrate S and another competing substrate S′ present at the same concentration which also reacts with your enzyme with $K_m = 0.01$ mM and $k_{cat} = 200$ / s if [S] = [S′]?

13. Using the method of graphs, write the initial rate equation for the following system with A, B, P, and Q present.



Odd-numbered rate constants are for forward reactions; even-numbered constants are for reverse.

a) Satisfy yourself that

$$\frac{-d[\mathrm{A}]}{dt} = k_1[\mathrm{A}][\mathrm{E}] - k_2[\mathrm{EA}]$$
$$= \frac{-d[\mathrm{B}]}{dt} = k_3[\mathrm{B}][\mathrm{EA}] - k_4[\mathrm{EXY}]$$
$$= \frac{d[\mathrm{P}]}{dt} = k_5[\mathrm{EXY}] - k_6[\mathrm{P}][\mathrm{EQ}]$$
$$= \frac{d[\mathrm{Q}]}{dt} = k_7[\mathrm{EQ}] - k_8[\mathrm{E}][\mathrm{Q}]$$

The determinants (given by the method of graphs) which provide the steady-state concentrations of [E] and of the various exzyme–substrate and enzyme–product complexes are

$$[\mathrm{E}] = k_2k_7[k_4 + k_5] + k_3k_5k_7[\mathrm{B}] + k_2k_4k_6[\mathrm{P}]$$

$$[\mathrm{EA}] = k_1k_7[k_4 + k_5][\mathrm{A}] + k_1k_4k_6[\mathrm{A}][\mathrm{P}] + k_4k_6k_8[\mathrm{P}][\mathrm{Q}]$$

$$[\mathrm{EXY}] = k_1k_3k_7[\mathrm{A}][\mathrm{B}] + k_2k_6k_8[\mathrm{P}][\mathrm{Q}] + k_1k_3k_6[\mathrm{A}][\mathrm{B}][\mathrm{P}] + k_3k_6k_8[\mathrm{B}][\mathrm{P}][\mathrm{Q}]$$

$$[\mathrm{EQ}] = k_2k_8[\ k_4 + k_5][\mathrm{Q}] + k_1k_3k_5[\mathrm{A}][\mathrm{B}] + k_3k_5k_8[\mathrm{B}][\mathrm{Q}]$$

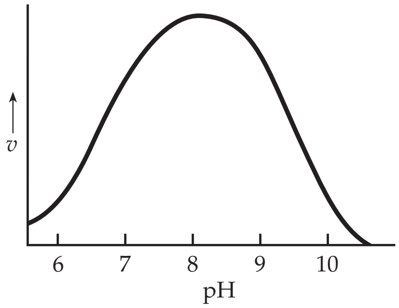
b) To obtain an expression for v_f the expression for $-d[\mathrm{A}]/dt$ may be multiplied by $[\mathrm{E}]_t$ and divided by $[\mathrm{E}] + [\mathrm{EA}] + [\mathrm{EXY}] + [\mathrm{EQ}] = [\mathrm{E}]_t$. Then the above determinants may be substituted.

c) What is $V_{\mathrm{max,forward}}$ and $V_{\mathrm{max,reverse}}$? HINT: $V_{\mathrm{max,forward}}$ is obtained when [P] = [Q] = 0 and [A] = [B] = ∞. What are K_{mA} and K_{mB}? K_{mA} is obtained when [P] = [Q] = 0, [B] = ∞, and $v = 1/2\ V_{\mathrm{max,forward}}$.

d) With a knowledge of the kinetic parameters, indicate how the eight rate constants may be obtained if the total concentration of enzyme, $[\mathrm{E}]_t$, is known.

14. Anticooperativity was observed in the plot of velocity vs. substrate concentration for an enzyme. Can this observation be explained by the Monod–Wyman–Changeux model for oligomeric enzymes? By the model of Koshland? Explain.

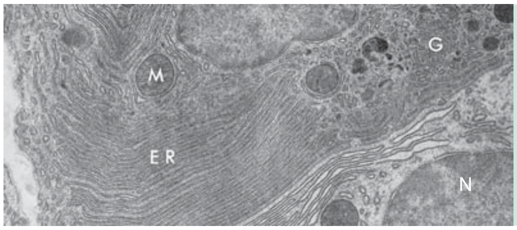
15. Interpret, for each of the following cases, the curve showing measured initial velocity at constant substrate concentration (*not* maximum velocity) against pH for an enzyme-catalyzed reaction.



a) The substrate is neutral and contains no acidic or basic groups. The Michaelis constant is found to be independent of pH over the range studied.

b) The substrate is neutral and contains no acidic or basic groups, and the maximum velocity is found to be independent of pH.

c) The substrate is an α-amino acid, and the maximum velocity is found to be independent of pH.



This electron micrograph of a thin section of pancreatic epithelial cells shows parts of two secretory cells that are synthesizing proenzymes (zymogens). Nuclei (N) are seen at the top center and lower right. Numerous ribosomes (barely seen here) line the many membranes of the endoplasmic reticulum (ER) and are most abundant near lateral and basal (left) surfaces of the cells. A few mitochondria (M) are present. The synthesized proteins move toward the apical surfaces (a small piece is seen at upper right), passing through the Golgi region (G) and accumulating in zymogen granules (Z) before secretion. From Porter, K. R. and Bonneville, M. A., *Fine Structure of Cells and Tissues*, Lea and Febiger, Philadelphia 1973. Courtesy of Mary Bonneville.

Contents

505 .... **A. The Starting Materials**
505 .... 1. Digestion
507 .... 2. Sources of Energy
507 .... **B. Catabolism and the Synthesis of ATP**
507 .... 1. Priming or Activation of Metabolites
508 .... 2. Interconversions of Sugar Phosphates
508 .... 3. Glycolysis and Fermentation
511 .... 4. Pyruvate Dehydrogenase
511 .... 5. Beta Oxidation
512 .... 6. The Electron Transport Chain, Oxidative Phosphorylation
515 .... 7. The Citric Acid Cycle
515 .... **C. Biosynthesis**
515 .... 1. Reversing Catabolic Pathways
517 .... 2. Photosynthesis
517 .... **D. Synthesis and Turnover of Macromolecules**
518 .... 1. Folding and Maturation of Proteins
519 .... 2. Transport of Proteins to Their Destinations within a Cell
519 .... *Signal sequences and translocation*
520 .... *Ticketing destinations*
521 .... *Vesicular transport and the Golgi system*
521 .... 3. Posttranslational Alterations
522 .... *Proteolytic processing*
522 .... *Altered ends*
523 .... 4. Intracellular Degradation of Proteins
527 .... 5. Turnover of Nucleic Acids
530 .... **E. Classifying Enzymatic Reactions**

531 .... **References**
533 .... **Study Questions**

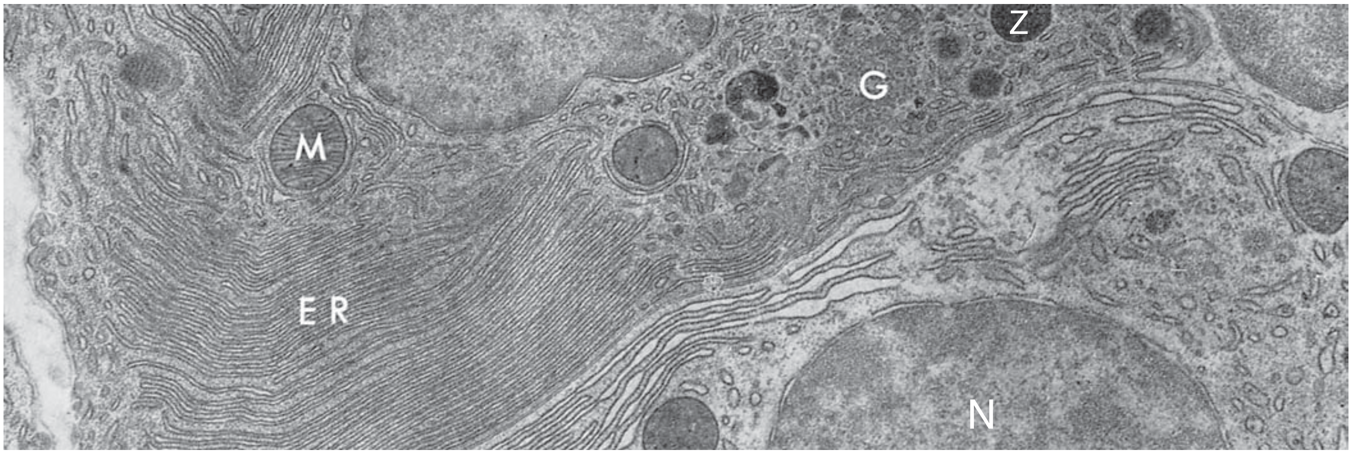
Boxes

513 .... Box 10-A An Early Labeling Experiment
517 .... Box 10-B Discovery of the Citric Acid Cycle
524 .... Box 10-C Ubiquitin
528 .... Box 10-D Drawing Those Little Arrows

Tables

526 .... Table 10-1 Types of Biochemical Reactions with Ionic Mechanisms

10 An Introduction to Metabolism



The first section of this book has dealt with the basic structures and with many of the complexities of the materials formed by living cells. In the next major section we will look at the chemical reactions that build and maintain a cell and that permit it to grow and to be responsive to external stimuli. These reactions are organized into **metabolic sequences** or **pathways** that form a complex, branched, and interconnected network. It would be pointless to try to memorize all of them. However, at this point it will be worthwhile to consider the significance of a few of the major sequences, which describe central pathways of metabolism.

Beginning in the nineteenth century, the investigation of these pathways and of the associated enzymes has provided much of our knowledge about the chemistry of living things. The protein products that are encoded in the structural genes that direct these pathways make up a large fraction of the material in a cell. It will be well worth the reader's effort to examine the chemistry of these metabolic sequences in detail. Regulatory mechanisms that are applicable to them, and to other pathways, are described in Chapter 11 and chemical details of enzyme action are considered in Chapters 12–16. Then, in Chapter 17, the chemical logic of the reaction sequences is considered in more detail.

A. The Starting Materials

All cells of all organisms take in chemical starting materials and give off chemical products. They all have a source of energy and generate heat during their metabolism. They all synthesize complex organic substances and maintain a high degree of organization, that is, a state of relatively low entropy. The materials taken up by cells are often organic compounds which not only supply material for the synthesis of cellular constituents but also may be degraded to provide energy. A characteristic of all cells is the ability simultaneously to synthesize thousands of complex proteins and other materials and, at the same time, to break down (catabolize) the same types of compounds. Since cells both synthesize and catabolize most cellular components there is a continuous turnover of the very structural components of which they are composed. Metabolism encompasses all of these processes.

The most rapid catabolic reactions are usually those that provide the cell's energy. Organisms vary greatly in the materials used for food. Human beings, as well as many other organisms, break down carbohydrates, lipids, and proteins to obtain energy and starting materials for biosynthesis. In contrast, some organisms use only one or a few simple organic compounds while autotrophs satisfy their needs entirely with inorganic materials. Looking at all of the species of living things we find an extraordinary range of specialization in metabolism as well as in structure. Nevertheless, there are many common features of metabolism. For example, most cells utilize glucose or a close relative as a source of biosynthetic intermediates. Pathways for synthesis of nucleic acids and proteins are similar in all species. Even the control of growth and development depends upon proteins whose structures are often conserved throughout the living world.

1. Digestion

We humans must digest most foods before we can utilize them. The same is true for most bacteria, which need amino acids generated by the breakdown of

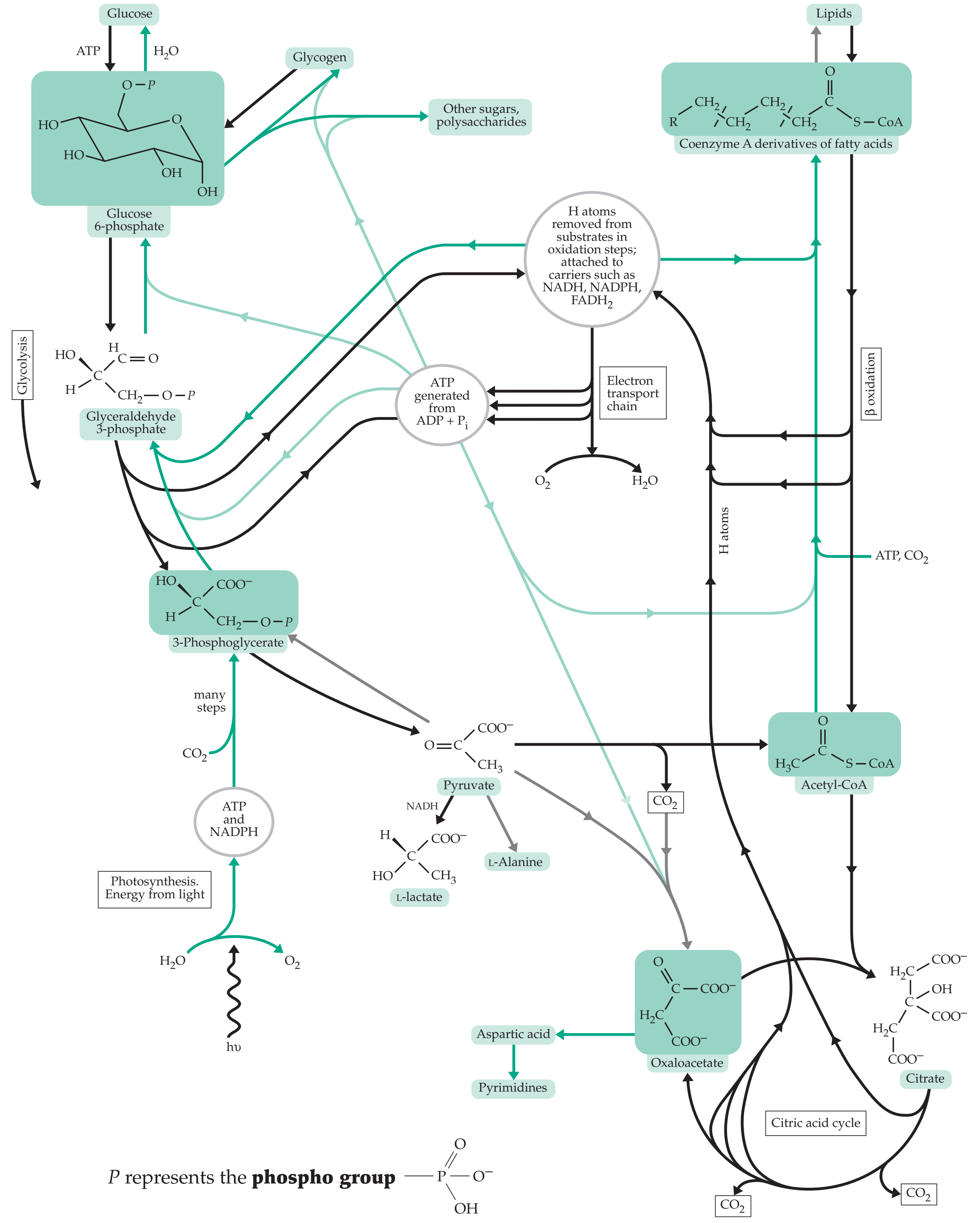
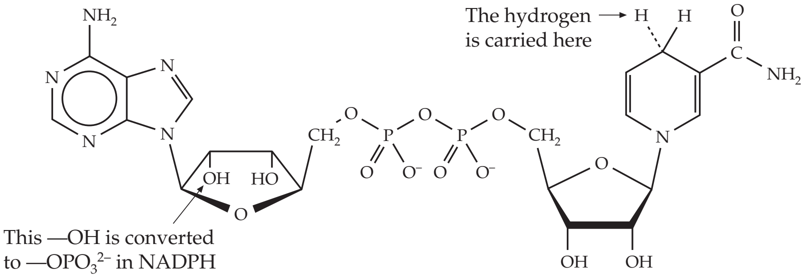

Glucose
ATP
H_2O
Glycogen
Other sugars, polysaccharides
Glucose 6-phosphate
Lipids
Coenzyme A derivatives of fatty acids
H atoms removed from substrates in oxidation steps; attached to carriers such as NADH, NADPH, $FADH_2$
Glycolysis
Glyceraldehyde 3-phosphate
ATP generated from ADP + P_i
Electron transport chain
O_2
H_2O
β oxidation
H atoms
ATP, CO_2
3-Phosphoglycerate
many steps
CO_2
Pyruvate
NADH
L-lactate
L-Alanine
Acetyl-CoA
CO_2
ATP and NADPH
Photosynthesis. Energy from light
H_2O
O_2
hυ
Aspartic acid
Pyrimidines
Oxaloacetate
Citrate
Citric acid cycle
CO_2
CO_2
P represents the **phospho group**

Figure 10-1 An overall view of some metabolic sequences. Several major pathways of catabolism are indicated by heavy lines. The glycolysis pathway, leading to pyruvate and lactate, starts at the top left, while the β oxidation pathway of fatty acids is on the right. Biosynthetic routes are shown in green lines. A few of the points of synthesis and utilization of ATP are indicated by the lighter green lines. Some of the oxidation–reduction reactions that produce or utilize the reduced hydrogen carriers NADH, NADPH, and $FADH_2$ are also indicated. The citric acid cycle, a major supplier of these molecules, is shown at the bottom right, while photosynthesis, the source of reduced hydrogen carriers in green plants, and the source of nearly all energy for life, is shown at the bottom left.

proteins. Our digestive enzymes, which act on carbohydrates, proteins, lipids, and nucleic acids, are synthesized by cells in the salivary glands, stomach, pancreas, and intestinal lining. The properties of these enzymes are described in Chapter 12. Bacteria secrete digestive enzymes into their surroundings. Although most green plants have no need to digest foods, they too have enzymes closely related to our digestive enzymes. These enzymes break bonds in proteins or carbohydrate polymers to allow expansion of leaves and stems, ripening of fruits, and other physical changes. All cells carry out **processing** or **maturation** of their newly synthesized polymers. This often involves the cutting off of one or more pieces of a protein, polysaccharide, or nucleic acid. The processing enzymes are also relatives of digestive enzymes.

2. Sources of Energy

All cells require a source of chemical energy. This is provided to a large extent by ATP (Chapter 6), whose hydrolysis can be coupled to reactions of synthesis, to transport across membranes, and to other endergonic processes. For this reason, all cells have active pathways for the synthesis of ATP. *Biosynthesis often also involves chemical reduction* of intermediates; therefore, cells must have a means of generating suitable reductants. The reagent of choice is often the hydrogen-carrying coenzyme **NADPH**. It is a phosphorylated form of NADH, a derivative of the vitamin **nicotinamide** (Box 15-A), whose structure follows. Reduction of metabolic intermediates by NADPH provides a second mechanism by which chemical energy is harnessed for biosynthesis. A third source of energy, which is considered in Chapter 18, is the gradient of hydrogen ions that is set up across cell membranes by oxidative or photochemical processes.


The hydrogen is carried here
This —OH is converted to —OPO$_3^{2-}$ in NADPH
NH$_2$
N
N
N
N
OH
HO
O
CH$_2$
O
P
O
P
O
O$^-$
CH$_2$
H
H
O
C
NH$_2$
N
OH
OH

The reduced coenzyme NADH

B. Catabolism and the Synthesis of ATP

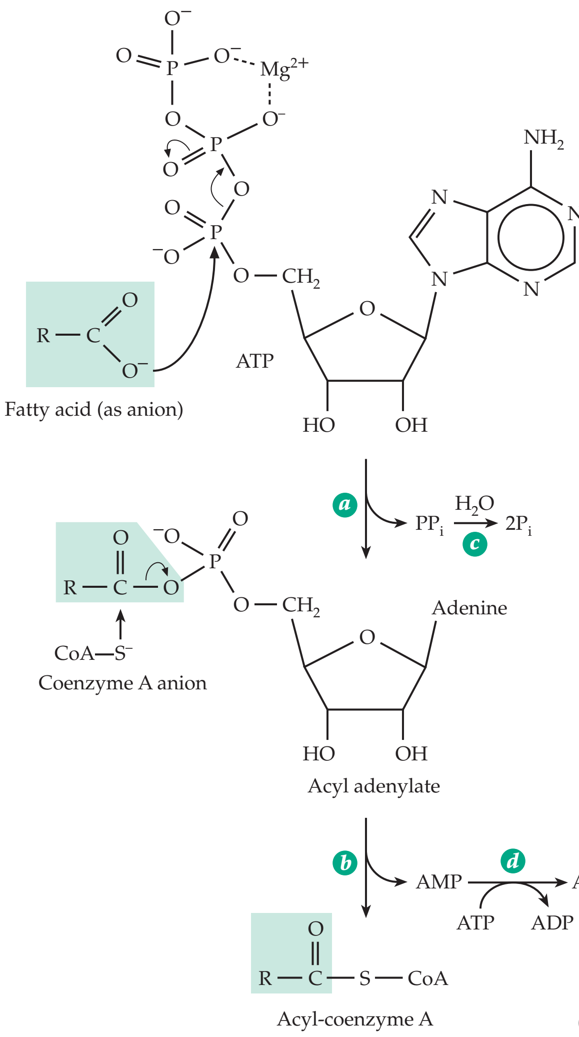
The catabolic sequences by which cells obtain energy often appear dominant. For animals, fungi, and nonphotosynthetic bacteria these pathways are used to metabolize large amounts of food and to produce large amounts of ATP. Because they are also related to biosynthetic pathways they have a central importance in virtually all organisms.

1. Priming or Activation of Metabolites

After a polymeric nutrient is digested and the monomeric products are absorbed into a cell, an energy-requiring **priming reaction** is usually required. For example, the hydrolysis of fats (whether in the gut or within cells) produces free fatty acids. Before undergoing further metabolism, these fatty acids are combined with the thiol compound **coenzyme A** to form **fatty acyl-CoA** derivatives. This requires a thermodynamically unfavorable reaction that necessitates the "expenditure" of ATP, that is, its hydrolysis to AMP and inorganic pyrophosphate (PP$_i$). Likewise, glucose, when taken into cells, is converted into the phosphate ester glucose 6-phosphate (glucose-6-*P*), again with an associated cleavage of ATP. *The major metabolic pathways often start with one of these two substances: a fatty acyl-CoA derivative or glucose 6-phosphate.* The structures of both are given at the top of Fig. 10-1, which is designed to provide an overall view of metabolism.

The phospho group of glucose-6-*P* is not very reactive but it provides a "handle" which helps enzymes to recognize and to hold onto this glucose derivative. Likewise, the coenzyme A molecule, whose complete structure is given in Chapter 14, provides a large and complex handle for enzymes. However, from a chemical viewpoint, the formation of the thioester linkage of a fatty acid with the –SH group of coenzyme A in the acyl-CoA is more important. This alters the electronic structure of the fatty acid molecule, "activating" it toward the reactions that it must undergo subsequently. Thus, priming reactions often provide chemical changes essential to the mechanisms of subsequent reactions.

Activation of a carboxylic acid by formation of an acyl-CoA derivative (Eq. 10-1) is of special importance because of its widespread use by all organisms. The basic reaction occurs by two steps. The first (Eq. 10-1, step *a*) is a displacement reaction on the phosphorus atom of ATP to form an **acyl adenylate**, a mixed anhydride of the carboxylic acid and a substituted phosphoric acid. Such mixed anhydride intermediates, or **acyl phosphates**, are central to much of cellular energy metabolism because *they preserve the high group-transfer potential of a phospho group of ATP (Chapter 6) while imparting a high group-transfer potential to the acyl group*. This allows the acyl group to be transferred in step *b*



(10-1)

of Eq. 10-1 onto the sulfur atom of coenzyme A to form an acyl-coenzyme A in which the high group transfer potential of the acyl group is conserved (see Table 6-6). Two additional steps are linked to this sequence. In step *c* the inorganic pyrophosphate that is displaced in step *a* is hydrolyzed to two molecules of HPO_4^{2-} (P_i), and in step *d* the AMP formed is phosphorylated by ATP to form ADP. The overall sequence of Eq. 10-1 leads to hydrolysis of two molecules of ATP to ADP + P_i. In other words, two molecules of ATP are spent in activating a carboxylic acid by conversion to an acyl-coenzyme A.

2. Interconversions of Sugar Phosphates

The strategy employed by most cells in the catabolism of several 6-carbon sugars is to convert them to glucose 6-phosphate and, in the several steps outlined in Fig. 10-2, to cleave this hexose phosphate to two equivalent molecules of **glyceraldehyde 3-phosphate**. This triose phosphate can then be metabolized further. Notice the chemical nature of the reactions involved in the formation of glyceraldehyde 3-phosphate. The first step (Fig. 10-2, step *a*) is the transfer of the phospho group from ATP to the 6 position of glucose. Step *b* is the reversible opening of the sugar ring and step *c* the isomerization of an aldose to a ketose. Step *d* is a second transfer of a phospho group from ATP, another priming reaction that provides the second of the phosphate handles for the two molecules of triose phosphate that are formed. Step *e* is an aldol cleavage which breaks the C–C bond in the center of the ketose chain, while step *f* is another sugar isomerization that is chemically similar to the one in step *c*.

We see that in the six steps shown in Fig. 10-2 for conversion of glucose to two molecules of glyceraldehyde 3-phosphate, there are only four kinds of reaction: (1) phospho transfer from ATP; (2) opening of a sugar ring; (3) aldose-ketose isomerization; and (4) aldol cleavage. These types of reactions, which are described in more detail in Chapters 12 and 13, occur in many places in metabolism. As is indicated in Fig. 10-2, glucose 6-phosphate can also be generated from glycogen or starch. The first step (Fig. 10-2, step *g*) is phosphorolysis, cleavage by an inorganic phosphate ion. This leads directly to glucose 1-phosphate. The latter is isomerized by a phospho transfer process (step *h*).

3. Glycolysis and Fermentation

A major route of breakdown of carbohydrates is the **Embden–Meyerhof–Parnas** pathway, often referred to simply as **glycolysis**. It is indicated on the left side of Fig. 10-1 and in more detail in Figs. 10-2 and 10-3. The pathway begins with the reactions of Fig. 10-2, with either free glucose or glycogen as starting materials. Its end products may be reduced materials such as **lactic acid** or **ethanol**, which are formed under *anaerobic conditions* (Fig. 10-3). However, under *aerobic conditions* the product is **acetyl-coenzyme A**, whose acetyl group can then be oxidized to carbon dioxide and water in the **citric acid cycle**.

After the sugar chain is cleaved in the glycolysis sequence, the two resultant molecules of glyceraldehyde 3-phosphate are oxidized to 3-phosphoglycerate (Fig. 10-3, steps *a,b*). The oxidant is the hydrogen carrier NAD^+, the oxidized form of NADH. Cells frequently use NAD^+ to dehydrogenate alcohols to aldehydes or ketones, with one atom of hydrogen and an electron from the alcohol becoming attached to the NAD^+ to give NADH and the other hydrogen being released as H^+ (Eq. 10-2).

The oxidation of glyceraldehyde 3-phosphate is considerably more complex. The oxidation of an aldehyde to a carboxylic acid is a strongly exergonic process and the oxidation of glyceraldehyde 3-phosphate by cells is almost always coupled to the synthesis

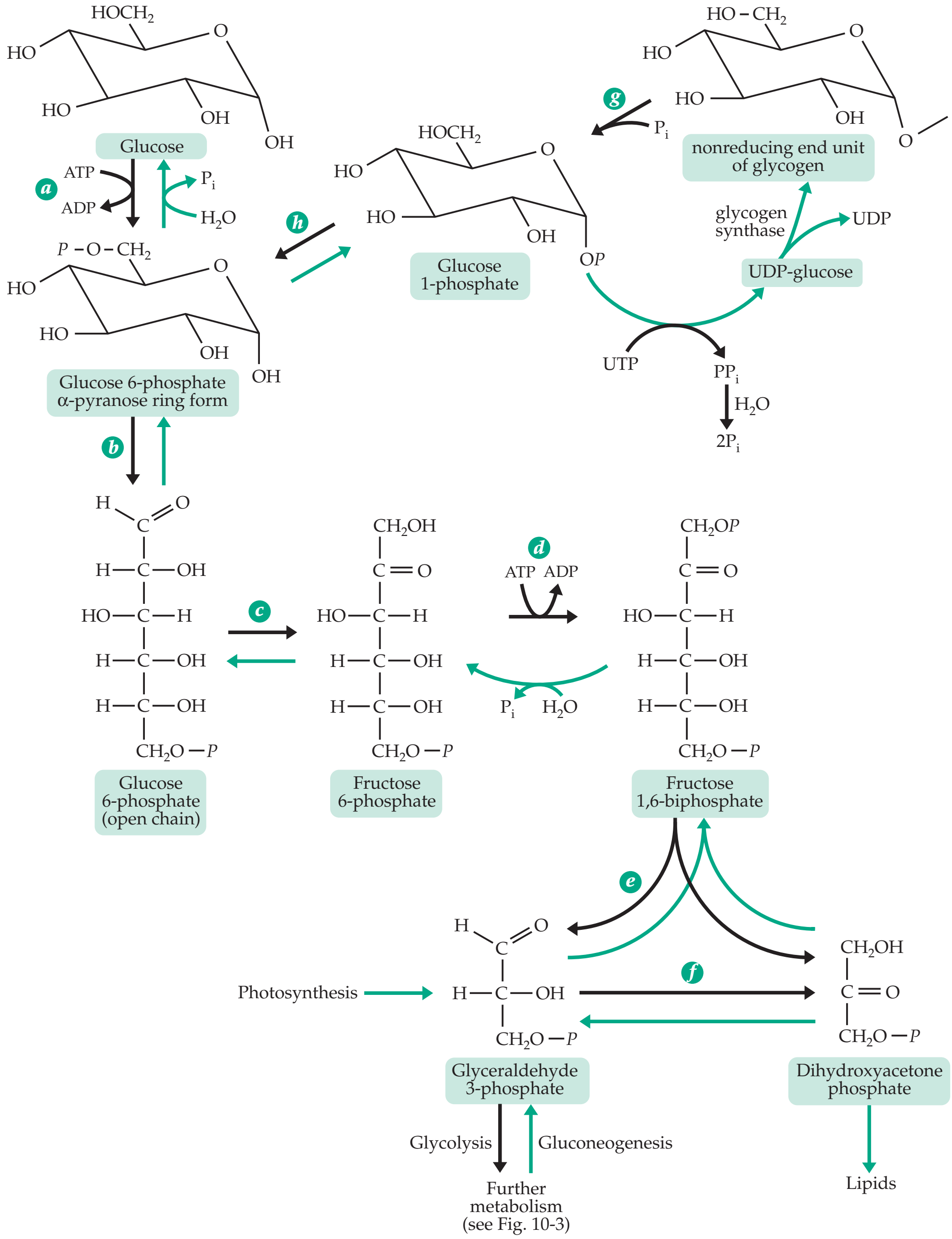
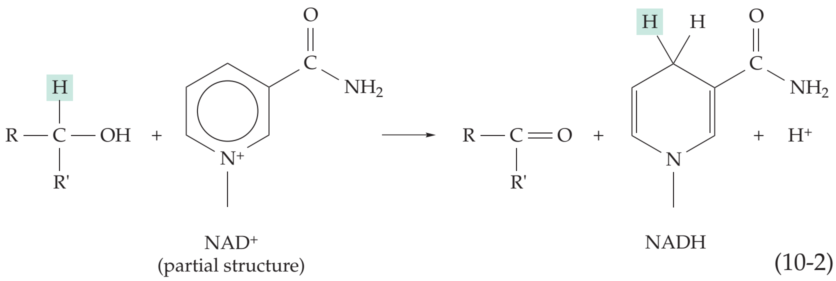

HOCH2
HO
O
HO
OH
OH
Glucose
ATP
a
ADP
Pi
H2O
P –O–CH2
O
HO
HO
OH
OH
Glucose 6-phosphate
α-pyranose ring form
h
HOCH2
HO
O
HO
OH
OP
Glucose
1-phosphate
g
Pi
HO–CH2
HO
O
HO
OH
O
nonreducing end unit
of glycogen
glycogen
synthase
UDP
UDP-glucose
UTP
PPi
H2O
2Pi
b
H
C
O
H—C—OH
HO—C—H
H—C—OH
H—C—OH
CH2O—P
Glucose
6-phosphate
(open chain)
c
CH2OH
C═O
HO—C—H
H—C—OH
H—C—OH
CH2O—P
Fructose
6-phosphate
d
ATP ADP
Pi
H2O
CH2OP
C═O
HO—C—H
H—C—OH
H—C—OH
CH2O—P
Fructose
1,6-biphosphate
e
H
C
O
Photosynthesis
H—C—OH
CH2O—P
f
CH2OH
C═O
CH2O—P
Glyceraldehyde
3-phosphate
Dihydroxyacetone
phosphate
Glycolysis
Gluconeogenesis
Further
metabolism
(see Fig. 10-3)
Lipids

Figure 10-2 The interconversion of glucose, glycogen, and glyceraldehyde-3-phosphate in the pathways of glycolysis, gluconeogenesis, and glycogen synthesis. Pathways of catabolism are indicated with black lines and those of biosynthesis with green lines.

of ATP. The importance of this coupling can be understood by the fact that *for many organisms living by anaerobic fermentation reactions, this one oxidation step provides the sole source of ATP*. It is the mechanism for coupling this oxidation to synthesis of ATP that accounts for the complexity of the glyceraldehyde 3-phosphate dehydrogenase reaction, whose chemistry is presented in detail in Chapter 15. As is indicated in the simplified version of Fig. 10-3 (step *a*) inorganic phosphate (P_i) is a reactant and the product released by the enzyme is 1,3-bisphosphoglycerate. This compound, which has been formed by the oxidative process, is an anhydride of phosphoric acid with 3-phosphoglyceric acid. The phospho group of such **acyl phosphates**, like that of ATP, has a high group transfer potential and can be transferred to ADP to generate ATP (Fig. 10-3, step *b*). Since each glucose molecule is cleaved to two molecules of triose phosphate, two molecules of ATP are generated in this process by the fermentation of one molecule of glucose. This is enough for a bacterium to live on if it ferments enough sugar.

Further metabolism of 3-phosphoglycerate involves an isomerization by means of a phosphotransferase (mutase reaction, step *c*) to form 2-phosphoglycerate which then loses water in an elimination reaction (step *d*). Elimination of a hydroxyl group in a β position relative to a carboxyl group together with an α-proton is another very frequently used type of metabolic reaction. In this case the product, **phosphoenolpyruvate** (PEP), is a unique and important metabolite. It is a key intermediate in the biosynthesis of aromatic amino acids and in numerous other biosynthetic sequences.


H
R—C—OH
R'
+
O
C
NH$_2$
N$^+$
NAD$^+$
(partial structure)
R—C=O
R'
+
H
H
O
C
NH$_2$
N
NADH
+ H$^+$

(10-2)

In the Embden–Meyerhof–Parnas pathway PEP transfers its phospho group to ADP (step *e*) to generate ATP and pyruvate. The latter is shown enclosed in brackets in Fig. 10-3 as the enolic form, which, tautomerizes to the oxo form (step *f*). The fact that the oxo form of pyruvate is much more stable than the enol gives the phospho group of phosphoenolpyruvate its high group transfer potential. The metabolic significance of this step is that the phosphate handles of the glycolytic intermediates, which were initially transferred on from ATP, are now returned to ADP with the regeneration of an equivalent amount of ATP.

The product of this metabolic sequence, **pyruvate**, is a metabolite of central importance. Its fate depends upon the conditions within a cell and upon the type of cell. When oxygen is plentiful pyruvate is usually converted to acetyl-coenzyme A, but under anaerobic conditions it may be reduced by NADH + H$^+$ to the alcohol **lactic acid** (Fig. 10-3, step *h*). This reduction exactly balances the previous oxidation step, that is, the oxidation of glyceraldehyde 3-phosphate to 3-phosphoglycerate (steps *a* and *b*). With a balanced sequence of an oxidation reaction, followed by a reduction reaction, glucose can be converted to lactate in the absence of oxygen, a **fermentation** process. The lactic acid fermentation occurs not only in certain bacteria but also in our own muscles under conditions of extremely vigorous exercise. It also occurs continuously in some tissues, e.g., the transparent lens and cornea of the eye.

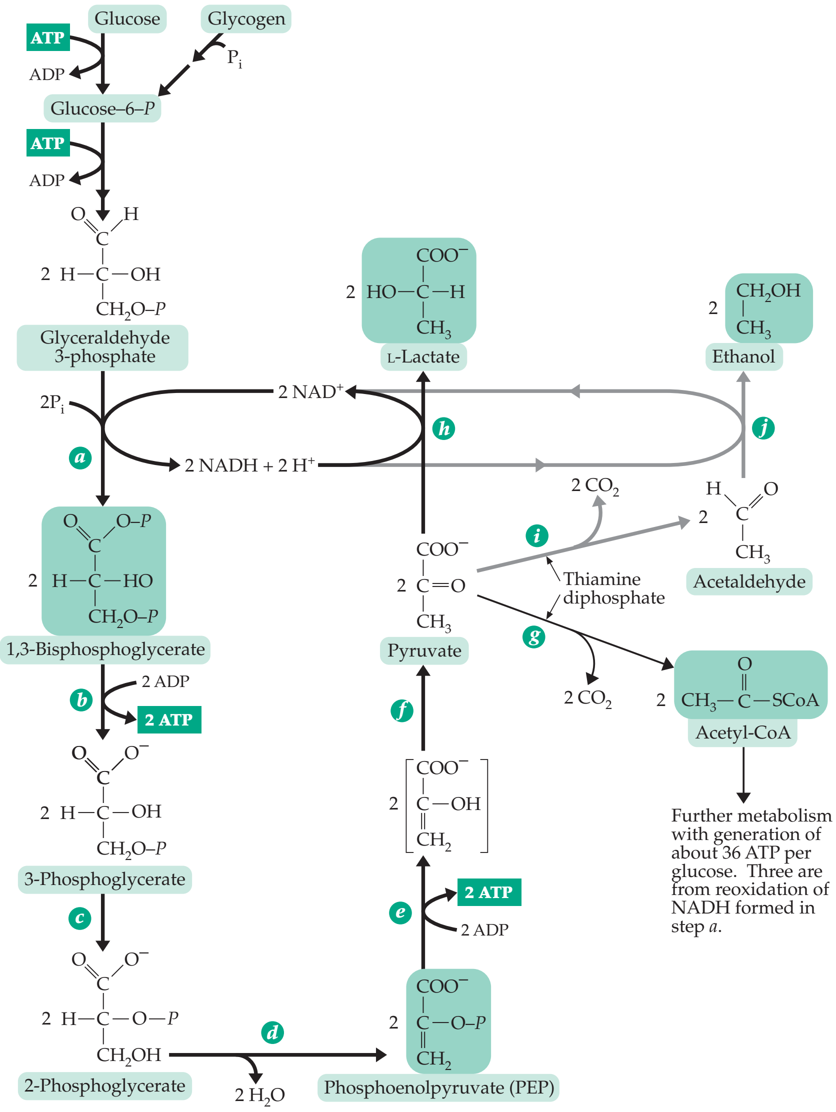

Glucose
Glycogen
ATP
ADP
P$_i$
Glucose–6–*P*
ATP
ADP
2 H—C—OH
CH$_2$O–*P*
Glyceraldehyde 3-phosphate
2P$_i$
a
2 NAD$^+$
2 NADH + 2 H$^+$
2 HO—C—H
CH$_3$
L-Lactate
h
2 CH$_2$OH CH$_3$
Ethanol
j
2 H—C—HO
CH$_2$O–*P*
1,3-Bisphosphoglycerate
b
2 ADP
2 ATP
2 H—C—OH
CH$_2$O–*P*
3-Phosphoglycerate
c
2 H—C—O—*P*
CH$_2$OH
2-Phosphoglycerate
d
2 H$_2$O
2 C—O–*P*
CH$_2$
Phosphoenolpyruvate (PEP)
e
2 ATP
2 ADP
2 C—OH
CH$_2$
f
COO$^-$
2 C=O
CH$_3$
Pyruvate
i
2 CO$_2$
Thiamine diphosphate
g
2 CO$_2$
2 H C O CH$_3$
Acetaldehyde
2 CH$_3$—C—SCoA
Acetyl-CoA
Further metabolism with generation of about 36 ATP per glucose. Three are from reoxidation of NADH formed in step *a*.

Figure 10-3 Coupling of the reactions of glycolysis with formation of lactic acid and ethanol in fermentations. Steps *a* to *g* describe the Embden–Meyerhof–Parnas pathway. Generation of 2 ATP in step *b* can provide all of the cell's energy.

In the well-known fermentation of sugar to ethanol by yeast, the pyruvate generated by glycolysis is first decarboxylated to acetaldehyde (Fig. 10-3, step *i*). This decarboxylation of a 2-oxo acid is chemically difficult and the enzyme catalyzing it makes use of a special reagent known as a **coenzyme**. For this type of reaction, the coenzyme is the diphosphate (pyrophosphate) ester of **thiamin** (**vitamin B_1**). Its mode of action is discussed in Chapter 14. It is usually needed when a 2-oxo acid is decarboxylated. The alcoholic fermentation by yeast is completed by reduction of acetaldehyde to ethanol (Fig. 10-3, step *j*), again using the NADH produced in the oxidation of glyceraldehyde 3-phosphate. The conversion of glucose to lactic acid or to ethanol and CO_2 are just two of many fermentation reactions, most of which are carried out by bacteria and which are dealt with further in Chapter 17. An important requirement is that the Gibbs energy change for the overall fermentation reaction be sufficiently negative that ATP synthesis can be coupled to it. Thus, using data from Table 6-4:

$$\begin{aligned}&C_6H_{12}O_6 \text{ (glucose)} \rightarrow 2\ C_3H_5O_3 \text{ (lactate)} + 2H^+\\&\Delta G' \text{ (pH 7)} = -196 \text{ kJ/mol } (-46.8 \text{ kcal/mol}) \qquad (10\text{-}3)\end{aligned}$$

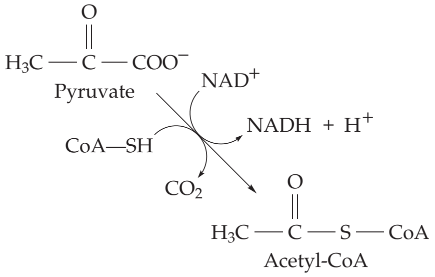
$$\begin{aligned}&2ADP^{3-} + 2P_i^{2-} + H^+ \rightarrow H_2O + ATP^{4-}\\&\Delta G' \text{ (pH 7)} = +69 \text{ kJ/mol} \qquad (10\text{-}4)\end{aligned}$$

$$\begin{aligned}&\text{Sum: } C_6H_{12}O_6 + 2ADP^{3-} + 2P_i^{2-} \rightarrow\\&\qquad H_2O + 2\ C_3H_5O_3 + 2H^+ + 2\ ATP^{4-}\\&\Delta G' \text{ (pH 7)} = -127 \text{ kJ/mol} \qquad (10\text{-}5)\end{aligned}$$

A requirement for all fermentations is the existence of a mechanism for coupling ATP synthesis to the fermentation reactions. In the lactic acid and ethanol fermentations this coupling mechanism consists of the formation of the intermediate 1,3-bisphosphoglycerate by the glyceraldehyde 3-phosphate dehydrogenase (Fig. 10-3, step *a*). This intermediate contains parts of both the products ATP and lactate or ethanol.

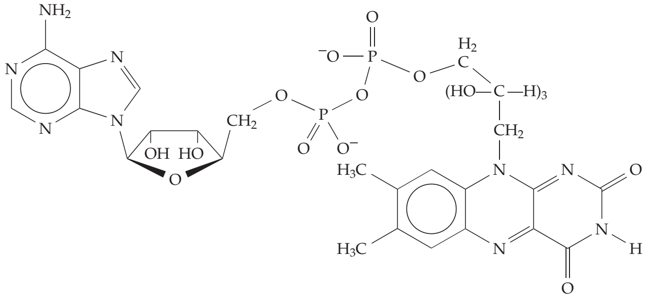
4. Pyruvate Dehydrogenase

In most organisms undergoing aerobic metabolism, pyruvate is oxidized to acetyl-CoA in a complex process involving its decarboxylation (Eq. 10-6). This **oxidative decarboxylation**, like the decarboxylation of pyruvate to acetaldehyde, requires thiamin diphosphate. In addition, an array of other catalysts participate in the process (see Fig. 15-15). Among these are the electron carrier **flavin adenine diphosphate** (**FAD**), which is derived from the vitamin **riboflavin**. Like NAD^+, this compound can accept two electrons (plus two protons) to form $FADH_2$. However, it sometimes serves as a one-electron carrier.



(10-6)

Acetyl-CoA is another major metabolic intermediate.



The coenzyme flavin adenine diphosphate (FAD)

It is an acyl-CoA of the type mentioned in Section 1 and can also be formed from acetate, ATP, and coenzyme A. Although the human diet contains some acetic acid, the two major sources of acetyl-CoA in our bodies are the oxidative decarboxylation of pyruvate (Eq. 10-6) and the breakdown of fatty acid chains. Let us consider the latter process before examining the further metabolism of acetyl-CoA.

5. Beta Oxidation

Whether fatty acids are oxidized to obtain energy or are utilized for biosynthesis, they are first converted to their acyl-CoA forms and are then cleaved to the two-carbon units represented by the acetyl groups of acetyl-CoA. The **beta oxidation** sequence, by which this occurs, is represented by the solid vertical arrow on the right side of Fig. 10-1 and is shown in greater detail in Fig. 10-4. We see from the latter figure that there are two dehydrogenation steps. The first (step *b*) removes hydrogen atoms from the α- and β-carbon atoms to produce a *trans* α, β unsaturated fatty acyl-CoA (Enoyl-CoA). The hydrogen acceptor is FAD. It is needed because this reaction requires a more powerful oxidant than does the dehydrogenation of an alcohol, for which NAD^+ is adequate. Addition of water to the double bond of the unsaturated acyl-CoA (step *c*) generates an alcohol which is then dehydrogenated (step *d*) by NAD^+ to form a β-oxoacyl-CoA. This is cleaved (step *e*) by a reaction (thiolysis) with another molecule of coenzyme A to form acetyl-CoA and a new acyl-CoA with a shortened chain. The latter is "recycled" by passage through the β oxidation sequence repeatedly until the chain is shortened to a 2- or 3-carbon fragment, acetyl-CoA, or propionyl-CoA. Since most dietary fatty acids contain an even number of carbon atoms, acetyl-CoA is the predominant product.

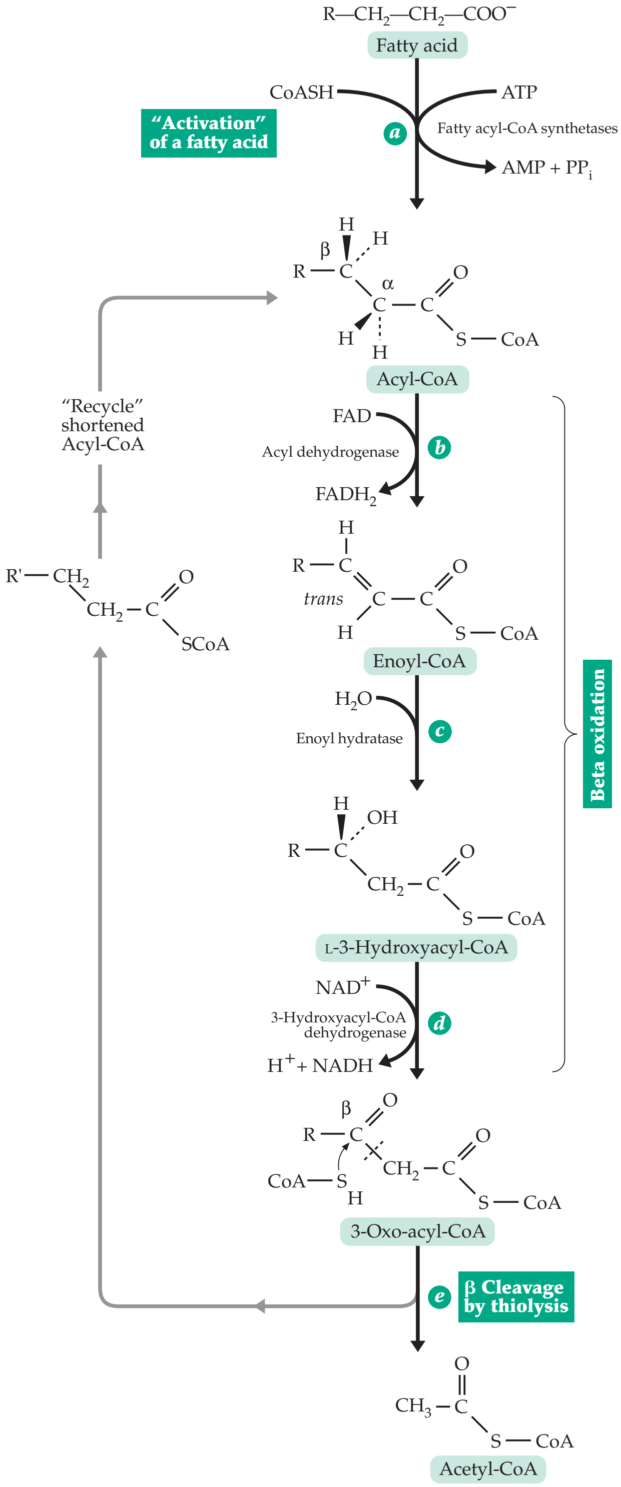

R—CH₂—CH₂—COO⁻
Fatty acid
CoASH
ATP
"Activation" of a fatty acid
a
Fatty acyl-CoA synthetases
AMP + PPᵢ
β
α
Acyl-CoA
"Recycle" shortened Acyl-CoA
FAD
Acyl dehydrogenase
b
FADH₂
trans
Enoyl-CoA
R'— CH₂
CH₂ — C
SCoA
H₂O
Enoyl hydratase
c
Beta oxidation
OH
L-3-Hydroxyacyl-CoA
NAD⁺
3-Hydroxyacyl-CoA dehydrogenase
d
H⁺ + NADH
CoA—S
3-Oxo-acyl-CoA
e
β Cleavage by thiolysis
CH₃ — C
S — CoA
Acetyl-CoA

Figure 10-4 Reactions of fatty acid activation and of breakdown by β oxidation.

The beta oxidation of fatty acids, like the dehydrogenation of pyruvate to acetyl-CoA, takes place within the inner **matrix** of the mitochondria in eukaryotic organisms. The reduced hydrogen carriers $FADH_2$ and NADH transfer their electrons to other carriers located within the **inner membrane** of the mitochondria. In bacteria the corresponding reactions occur with electron carriers present in the plasma membrane. Both FAD and NAD^+ are regenerated in this way and are able to again accept hydrogen from the beta oxidation reactions. This transfer of hydrogen atoms from substrates to hydrogen carriers is typical of biological oxidation processes.

6. The Electron Transport Chain, Oxidative Phosphorylation

Reoxidation of the reduced carriers NADH and $FADH_2$ actually involves a sequence of electron carriers, the **electron transport chain**, whose function is indicated below the circle near the center of Fig. 10-1. The oxidation of reduced NADH by O_2 (Eq. 10-7) is a highly exergonic process and is accompanied by the

$$\begin{aligned}&NADH + H^+ + 1/2\ O_2 \rightarrow NAD^+ + H_2O\\&\Delta G'\ (\mathrm{pH}\ 7) = -219\ \mathrm{kJ/mol}\end{aligned} \tag{10-7}$$

generation of about three molecules of ATP (from ADP and inorganic phosphate). This process, termed **oxidative phosphorylation**, is the principal source of usable energy (in the form of ATP) provided by breakdown of both carbohydrates and fats in the human body.

The mechanism of oxidation of NADH in the electron transport chain appears to occur by transfer of a hydrogen atom together with two electrons (a hydride ion H^-). Oxidation of $FADH_2$ to FAD might occur by transfer of two hydrogen atoms or by transfer of $H^- + H^+$. However, it is useful to talk about all of these compounds as **electron carriers** with the understanding that movement of one or both of the electrons may be accompanied by transfer of H^+. The electron transport complex is pictured in a very simplified form in Fig. 10-5.

The electrons donated from NADH or other reductants, upon entering this complex, travel from one carrier to the next, with each carrier being a somewhat more powerful oxidant than the previous one. The

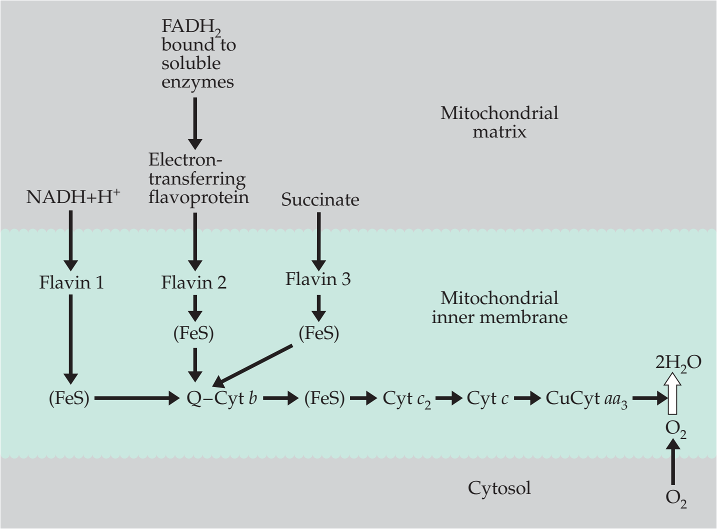

FADH2 bound to soluble enzymes
Mitochondrial matrix
Electron-transferring flavoprotein
NADH+H+
Succinate
Flavin 1
Flavin 2
Flavin 3
Mitochondrial inner membrane
(FeS)
(FeS)
2H2O
(FeS)
Q–Cyt b
(FeS)
Cyt c1
Cyt c
CuCyt aa3
O2
Cytosol
O2

Figure 10-5 An abbreviated version of the electron transport chain of mitochondria. Four electrons reduce O_2 to 2 H_2O. For details see Figs. 18-5 and 18-6.

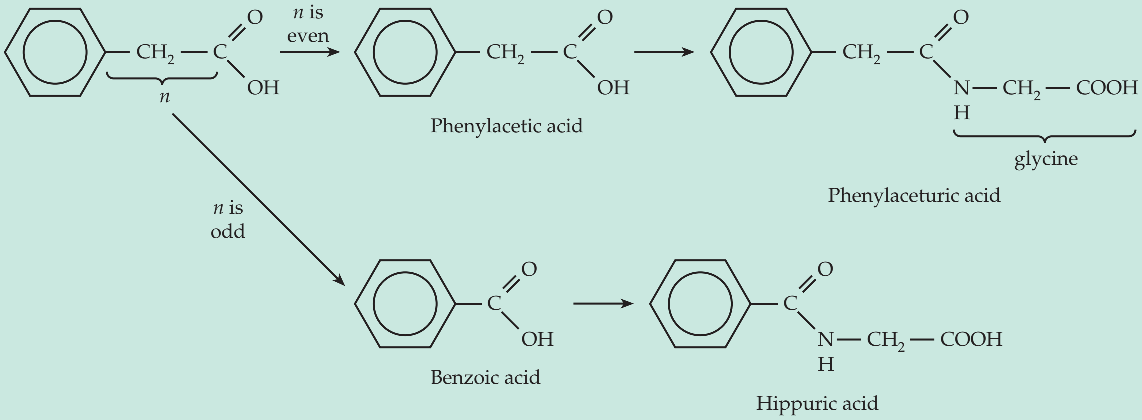
final carrier, known as cytochrome *aa*$_3$ or **cytochrome oxidase**, reacts with molecular oxygen. Each molecule of O_2, together with 4 H^+, is converted to two molecules of water. This stoichiometry requires that two molecules of NADH pass four electrons down the chain for each O_2 reduced.

The chemical structures of the components of the mitochondrial electron transport chain are varied but fall into several distinct classes. Most are proteins but these proteins always contain special coenzymes or **prosthetic groups** able to engage in oxidation–reduction reactions. At the substrate end of the chain NADH passes a hydrogen atom together with its bonding electron pair to a riboflavin-containing coenzyme **riboflavin 5'-phosphate** (FMN), which is tightly bound into a protein (designated Flavin 1 in Fig. 10-5). Similar **flavoproteins**, Flavin 2, and Flavin 3, act as oxidants for $FADH_2$ arising during beta

BOX 10-A AN EARLY LABELING EXPERIMENT

In 1904, long before the advent of radioactive tracers, Knoop synthesized fatty acids labeled by chemical attachment of a benzene ring at the end opposite the carboxyl group. He prepared these compounds with both odd and even numbers of carbon atoms in straight chains and fed them to dogs. From the dogs' urine he isolated **hippuric acid** and **phenylaceturic acid**, which are the amides of glycine with benzoic acid and phenylacetic acid, respectively. Knoop showed that the phenylacetic acid was produced from those fatty acids with an even number of carbon atoms, while the benzoic acid was formed from those with an odd number. From these experimental results, Knoop deduced that *fatty acid degradation occurs two carbon atoms at a time* and proposed his famous **β oxidation theory**.

Later experiments using isotopic labeling with ^{13}C verified Knoop's proposals,[a] but study of isolated enzymes was not possible until after the discovery of coenzyme A in 1950. Then, studies of fatty acid oxidation by extracts from isolated mitochondria established the details of the pathway.[b]


CH2
C
O
OH
n
n is even
n is odd
Phenylacetic acid
N
H
CH2
COOH
glycine
Phenylaceturic acid
Benzoic acid
Hippuric acid

---

[a] Weinhouse, S. (1995) *FASEB J.* **9**, 820–821

[b] Quastel, J. H. (1984) *Trends Biochem. Sci.* **9**, 117–118

oxidation and oxidation of succinate, whose significance to metabolism is discussed in the next section.

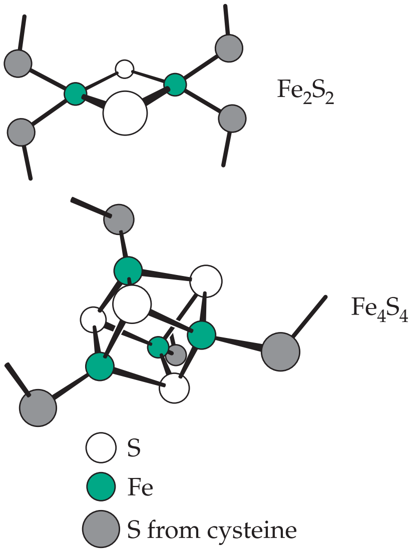
A significant difference between NADH and $FADH_2$ is that the former diffuses freely between the **dehydrogenases** that transfer hydrogen to it and the flavoprotein NADH dehydrogenase (Flavin 1) that reoxidizes it. However, FAD, whether in its oxidized state or as $FADH_2$, stays tightly bound to proteins at all times. Only hydrogen atoms or electrons are transferred into or out of these proteins. During beta oxidation the $FADH_2$ generated remains tied to the fatty acyl-CoA dehydrogenase protein. However, the two hydrogen atoms of this $FADH_2$ are transferred to another molecule of FAD, which is bound to an **electron-transferring flavoprotein.** This protein carries electrons one at a time, as FADH, by diffusion to the inner surface of the inner mitochondrial membrane. There it transfers the hydrogen atoms that it carries to the flavoprotein designated Flavin 2 in Fig. 10-5. The electron-transferring flavoprotein can also be viewed as a carrier of single electrons, each accompanied by H^+.

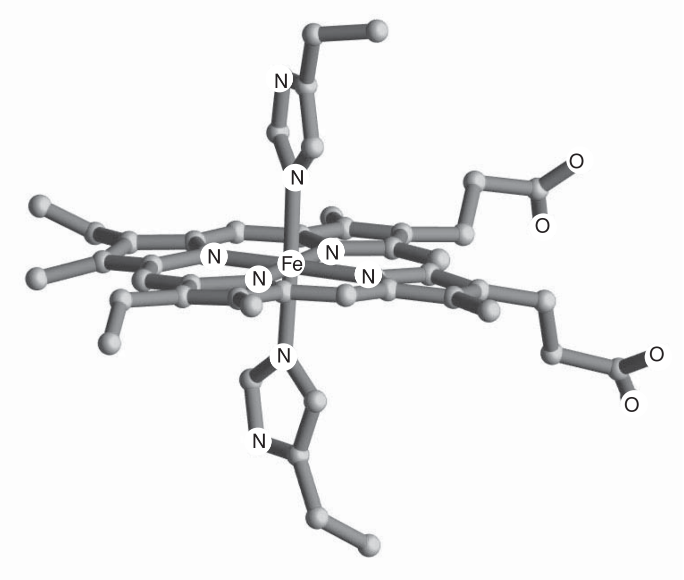
A second group of electron carriers in mitochondrial membranes are the **iron–sulfur [Fe–S] clusters** which are also bound to proteins. Iron–sulfur proteins release Fe^{3+} or Fe^{2+} plus H_2S when acidified. The "inorganic clusters" bound into the proteins have characteristic compositions such as Fe_2S_2 and Fe_4S_4. The sulfur atoms of the clusters can be regarded as sulfide ions bound to the iron ions. The iron atoms are also attached to other sulfur atoms from cysteine side chains from the proteins. The Fe–S proteins are often tightly associated with other components of the electron transport chain. For example, the flavoproteins Flavin 1, Flavin 2, and Flavin 3 shown in Fig. 10-5 all contain Fe–S clusters as does the Q-cytochrome *b* complex. All of these Fe–S clusters seem to be one-electron carriers.

A third hydrogen carrier of mitochondrial membranes, and the only one that is not unequivocally associated with a specific protein, is the isoprenoid quinone **ubiquinone** or **coenzyme Q** (Q in Fig. 10-5). Ubiquinone apparently serves as a common carrier, collecting electrons from three or more separate input ends of the chain and directing them along a single pathway to O_2.

The final group of mitochondrial redox components are one-electron carriers, small proteins (**cytochromes**) that contain iron in the form of the porphyrin complex known as **heme**. These carriers, which are discussed in Chapter 16, exist as several chemically distinct types: *a*, *b*, and *c*. Two or more components of each type are present in mitochondria. The complex cytochrome aa_3 deserves special comment. Although cytochromes are single-electron carriers, the cytochrome aa_3 complex must deliver four electrons to a single O_2 molecule. This may explain why the monomeric complex contains two hemes and two **copper** atoms which are also able to undergo redox reactions.[1,2]

Although the components of the electron transport chain have been studied intensively, there is still some mystery associated with the process by which ATP synthesis is coupled to electron transport (oxidative phosphorylation). A theory originally proposed by Mitchell[3] and now generally accepted[4,5] is that passage of electrons through the chain "pumps" protons from the inside to the outside of the tight inner mitochondrial membrane. As a result, protons accumulate along the outside of this membrane, as do negative counterions along the inside. The membrane becomes charged like a miniature electrical condenser. The synthesis of ATP takes place in the small knoblike **ATP synthase** which is partially embedded in the same membranes. The Gibbs energy for the formation of ATP from ADP and inorganic phosphate is apparently supplied by the flow of protons back to the inside of the membrane through the ATP synthase. Possible





Iron in a heme in a *b* type cytochrome

ways in which this may occur as well as many other aspects of electron transport are dealt with in Chapter 18.

7. The Citric Acid Cycle

The 2-carbon acetyl units removed from fatty acid chains or generated from sugars by glycolysis and the action of pyruvate dehydrogenase must be completely oxidized to carbon dioxide to provide cells with the maximum amount of energy. The oxidation of an acetyl group is a difficult chemical process, and probably for this reason nature has devised an elegant catalytic cycle, the **citric acid cycle**, which is indicated at the lower right in Fig. 10-1 and is shown in detail in Fig. 10-6. The cycle begins when the 4-carbon **oxaloacetate** (also spelled oxalacetate) is condensed with an acetyl group of acetyl-CoA to form the 6-carbon **citrate**. Then, in the remaining reactions of the cycle, two carbon atoms are removed as CO_2 and oxaloacetate is regenerated. Several oxidation steps are involved, each of which feeds additional **reducing equivalents** (i.e., hydrogen atoms removed from substrates) into the pool of hydrogen carriers and allows for more synthesis of ATP via the electron transport chain. The importance of this pathway to an organism can be understood from the fact that when glucose is oxidized completely to CO_2 and water via the citric acid cycle, about 38 molecules of ATP are formed, 19 times as much as by fermentation.

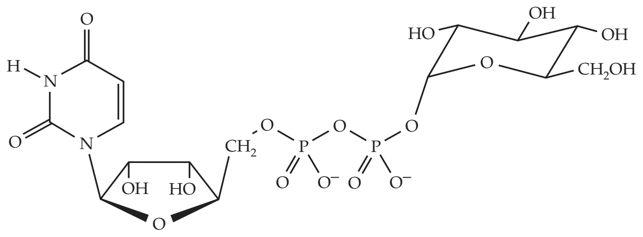
Although oxaloacetate is regenerated and therefore not consumed during the operation of the citric acid cycle, it also enters other metabolic pathways. To replace losses, oxaloacetate can be synthesized from pyruvate and CO_2 in a reaction that uses ATP as an energy source. This is indicated by the heavy gray line leading downward to the right from pyruvate in Fig. 10-1 and at the top center of Fig. 10-6. This reaction depends upon yet another coenzyme, a bound form of the vitamin **biotin**. Pyruvate is formed from breakdown of carbohydrates such as glucose, and the need for oxaloacetate in the citric acid cycle makes the oxidation of fats in the human body dependent on the concurrent metabolism of carbohydrates.

C. Biosynthesis

At the same time that cells break down foodstuffs to obtain energy, they are continuously creating new materials. The green lines in Fig. 10-1 indicate some pathways by which such biosynthesis takes place. Examining the left side of the figure, we see that either pyruvate or 3-phosphoglycerate can be converted back to glucose 6-phosphate and that the latter can be used to synthesize glycogen or other sugars or polysaccharides.

1. Reversing Catabolic Pathways

Breakdown of carbohydrates is thermodynamically spontaneous. Therefore, *cells cannot simply use catabolic pathways operating in reverse without finding ways of coupling the cleavage of ATP to synthesis.* In the formation of glucose from pyruvate in the liver, a process known as **gluconeogenesis**, there are three distinct points at which the enzymes used differ from those used in catabolism: (1) Pyruvate is converted to phosphoenolpyruvate by a mechanism utilizing more than one molecule of ATP, a pathway that is discussed in detail in Chapter 17; (2) as is shown in Fig. 10-2, fructose 1,6-bisphosphate is hydrolyzed to fructose 6-phosphate and inorganic phosphate by a **phosphatase** rather than through reversal of step *d*, which would form ATP; and (3) glucose-6-phosphate is hydrolyzed by a phosphatase rather than following the reverse of step *a* in Fig. 10-2. Furthermore, glycogen is synthesized from glucose 6-phosphate, not by reversal of the phosphorylase reaction (Fig. 10-2, step *g*), but via a new intermediate, **uridine diphosphate glucose** (UDPG), whose formation involves cleavage of UTP. The latter is generated by phospho transfer from ATP. Inorganic **pyrophosphate** (PP_i) is a product of UDPG formation and is removed by hydrolytic cleavage to two molecules of inorganic phosphate by the enzyme **pyrophosphatase**. This reaction helps to make the biosynthesis thermodynamically spontaneous by removing one of the reaction products.



Uridine diphosphate glucose (UDPG)

A similar situation exists in the case of fatty acid synthesis, which proceeds from acetyl-CoA and reverses fatty acid breakdown. However, both carbon dioxide and ATP, a source of energy, are needed in the synthetic pathway. Furthermore, while oxidation of fatty acids requires NAD^+ as one of the oxidants, and generates NADH, the biosynthetic process often requires the related NADPH. These patterns seen in biosynthesis of sugars and fatty acids are typical. *Synthetic reactions resemble the catabolic sequences in reverse, but distinct differences are evident.* These can usually be related to the requirement for energy and often also to control mechanisms.

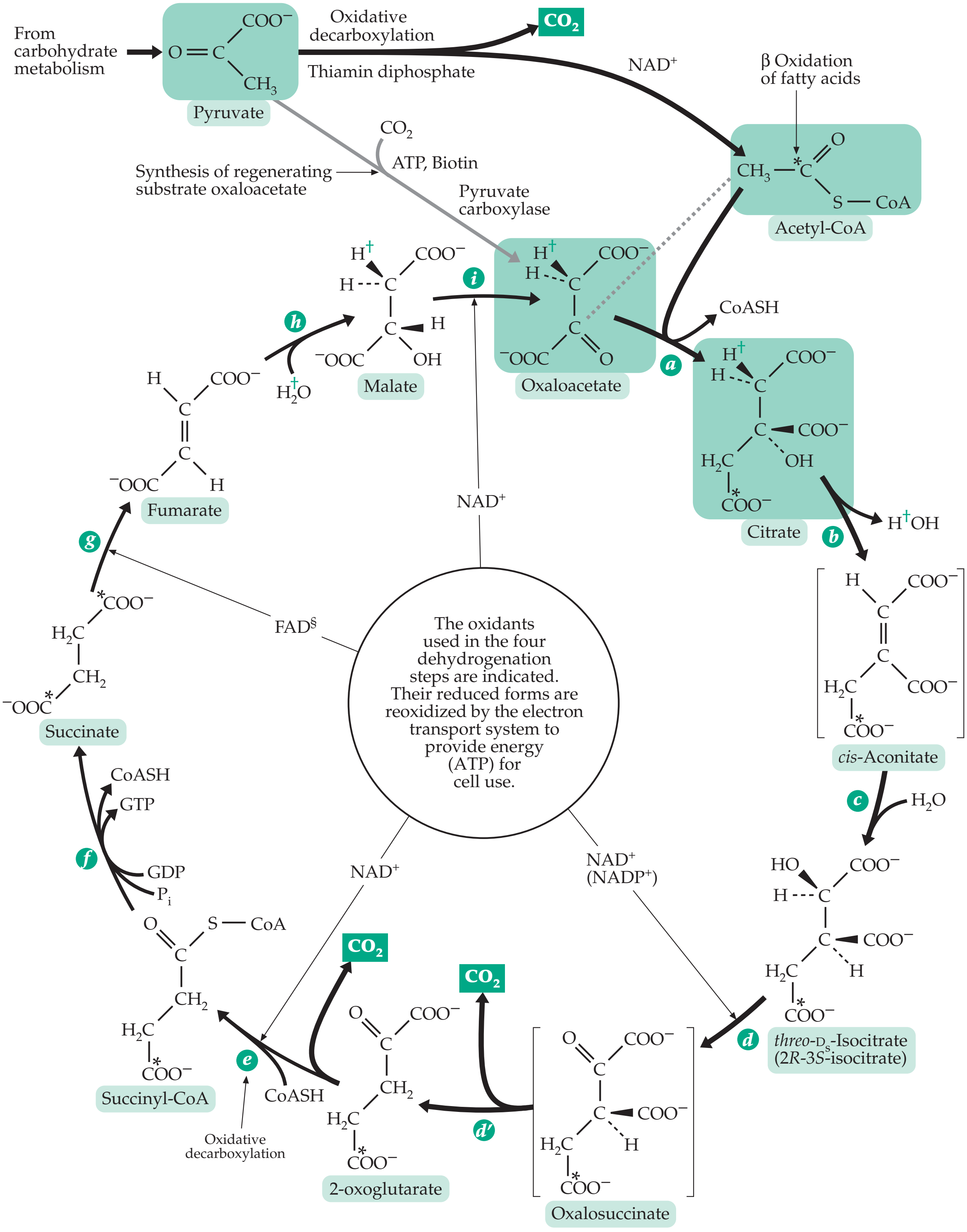

From carbohydrate metabolism
Pyruvate
Oxidative decarboxylation
Thiamin diphosphate
CO2
NAD+
β Oxidation of fatty acids
Acetyl-CoA
CO2
ATP, Biotin
Pyruvate carboxylase
Synthesis of regenerating substrate oxaloacetate
Oxaloacetate
CoASH
a
Citrate
b
H†OH
cis-Aconitate
c
H2O
threo-Ds-Isocitrate (2R-3S-isocitrate)
d
Oxalosuccinate
NAD+ (NADP+)
d′
CO2
2-oxoglutarate
CO2
e
CoASH
Oxidative decarboxylation
NAD+
Succinyl-CoA
f
CoASH
GTP
GDP
Pi
Succinate
g
FAD§
Fumarate
h
H2†O
Malate
i
NAD+
The oxidants used in the four dehydrogenation steps are indicated. Their reduced forms are reoxidized by the electron transport system to provide energy (ATP) for cell use.

Figure 10-6 Reactions of the citric acid cycle (Krebs' tricarboxylic acid cycle). Asterisks designate positions of isotopic label from entrance of carboxyl-labeled acetate into the cycle. Note that it is *not* the two carbon atoms from acetyl-CoA that are immediately removed as CO_2 but two atoms from oxaloacetate. Only after several turns of the cycle are the carbon atoms of the acetyl-CoA completely converted to CO_2. Nevertheless, the cycle can properly be regarded as a mechanism of oxidation of acetyl groups to CO_2. Green daggers (†) designate the position of ^2H introduced into malate as $^2H^+$ from the medium. FAD§ designates covalently bound 8-histidyl-FAD (see Chapter 15).

BOX 10-B DISCOVERY OF THE CITRIC ACID CYCLE (KREBS' TRICARBOXYLIC ACID CYCLE)

One of the first persons to study the oxidation of organic compounds by animal tissues was T. Thunberg, who between 1911 and 1920 discovered about 40 organic compounds that could be oxidized by animal tissues. Salts of succinate, fumarate, malate, and citrate were oxidized the fastest. Well aware of Knoop's β oxidation theory, Thunberg proposed a cyclic mechanism for oxidation of acetate. Two molecules of this two-carbon compound were supposed to condense (with reduction) to succinate, which was then oxidized as in the citric acid cycle to oxaloacetate. The latter was decarboxylated to pyruvate, which was oxidatively decarboxylated to acetate to complete the cycle. *One of the reactions essential for this cycle could not be verified experimentally.* It is left to the reader to recognize which one.

In 1935, A. Szent-Györgyi discovered that all of the carboxylic acids that we now recognize as members of the citric acid cycle stimulated respiration of animal tissues that were oxidizing other substrates such as glucose. Drawing on this knowledge, Krebs and Johnson in 1937 proposed the citric acid cycle.[a,b,c] Krebs provided further confirmation in 1940 by the observation that malonate, a close structural analog and competitive inhibitor of succinate, in concentrations as low as 0.01 M blocked the respiration of tissues by stopping the oxidation of succinate to fumarate.[c]

$$^{-}OOC{-}CH_2{-}COO^{-}$$

Malonate

$$^{-}OOC{-}CH_2{-}CH_2{-}COO^{-}$$

Succinate

In muscle, 90% of all respiration was inhibited and succinate was shown to accumulate, powerful proof of the importance of the citric acid cycle in the respiration of animal tissues.

---

[a] Krebs, H. A., and Johnson, W. A. (1937) *Enzymologia* **4**, 148–156

[b] Krebs, H. A. (1981) *Reminiscences and Reflections* Clarendon Press, Oxford

[c] Fruton, J. S. (1999) *Proteins, Enzymes, Genes: the Interplay of Chemistry and Biology* Yale Univ. Press, New Haven, CT

[d] Quastel, J. H. (1978) *Trends Biochem. Sci.* **3**, 68–69

2. Photosynthesis

The principal means of formation of glucose in nature is through photosynthesis in green plants (Fig. 10-1, lower left). Light energy is captured by chlorophyll and is used to transfer electrons from chlorophyll to other electron carriers, the most important of which is $NADP^+$. It is reduced to NADPH which is used to reduce carbon dioxide to sugar phosphates in a complex series of reactions known as the **reductive pentose phosphate pathway**, which is described in Chapter 17. ATP is also required for photosynthetic reduction of CO_2. It is generated by allowing some of the electrons to flow back through an electron transport chain in the membranes of the chloroplasts. This chain closely resembles that from Q to cytochrome *c* in mitochondria (Fig. 10-5), and the generation of ATP in this **photosynthetic phosphorylation** occurs in a manner analogous to that in the electron transport chain of mitochondria. In green plants the electrons removed from chlorophyll in one light-requiring reaction are replaced by electrons formed during the cleavage of water in a second light-dependent reaction, a reaction that also releases oxygen, O_2, and generates hydrogen ions (H^+). The first stable product from reduction of CO_2 in photosynthesis is 3-phosphoglycerate. It can be converted to sugars by pathways analogous to those employed by animals in gluconeogenesis. One interesting difference is the use in chloroplasts of NADPH + H^+ in reduction of 3-phosphoglycerate (reverse of step *a* in Fig. 10-3).

A small number of other biosynthetic pathways, which are used by both photosynthetic and nonphotosynthetic organisms, are indicated in Fig. 10-1. For example, pyruvate is converted readily to the amino acid **L-alanine** and oxaloacetate to **L-aspartic acid**; the latter, in turn, may be utilized in the biosynthesis of pyrimidines. Other amino acids, purines, and additional compounds needed for construction of cells are formed in pathways, most of which branch from some compound shown in Fig. 10-1 or from a point on one of the pathways shown in the figure. In virtually every instance biosynthesis is dependent upon a supply of energy furnished by the cleavage to ATP. In many cases it also requires one of the hydrogen carriers in a reduced form. While Fig. 10-1 outlines in briefest form a minute fraction of the metabolic pathways known, the ones shown are of central importance.

D. Synthesis and Turnover of Macromolecules

Proteins make up the bulk of the catalytic machinery of cells and together with other macromolecules most of the structure. Therefore, the synthesis and degradation of proteins and the control of those processes are of great importance to cells. Although the emphasis in this section is on proteins, similar considerations apply to nucleic acids, polysaccharides, and other macromolecules.

With the exception of some antibiotics and other short-chain molecules, all polypeptides are formed on ribosomes, which assemble proteins according to the sequences of nucleotides in the messenger RNA (mRNA) molecules. The basic chemistry is simple. The carboxyl groups of the amino acids are converted to reactive acyl adenylates by reaction with ATP, just as in Eq. 10-1. Each "activated" amino acid is carried on a molecule of transfer RNA (tRNA) and is placed in the reactive site of a ribosome when the appropriate codon of the mRNA has moved into the site. The growing peptide chain is then transferred by a displacement reaction onto the amino group of the activated amino acid that is being added to the peptide chain. In this manner, new amino acids are added one at a time to the carboxyl end of the chain, which always remains attached to a tRNA molecule. The process continues until a stop signal in the mRNA ends the process and the completed protein chain is released from the ribosome. Details are given in Chapter 29.

1. Folding and Maturation of Proteins

A newly synthesized peptide chain probably folds quickly. However, the cytosol provides an environment rich in other proteins and other macromolecules that can interact with the new peptide and may catalyze or inhibit folding. Among the most abundant proteins of bacteria or eukaryotes are proteins known as **chaperonins**. They apparently help polypeptide chains to fold correctly, partly by "chaperoning" them through the cytoplasm and across cell membranes, protecting them from becoming entangled with other proteins and macromolecules while they fold.[6–9] There are several classes of chaperonins. Most are oligomers made up of 10- to 90-kDa subunits. They have a variety of names, which may be somewhat confusing. The first chaperonins were identified only as **heat shock proteins** (Hsp). They are produced in large amounts by bacteria or other cells when the temperature is raised quickly, and are designated Hsp70, Hsp90, etc., where the number is the subunit mass in kDa. Other chaperonins were recognized as products of genes needed for replication of the DNA of bacteriophage λ in *E. coli.* Consequently, one major chaperonin of the Hsp70 class is designated **DnaK**. The Hsp70 protein of mitochondria was named "binding protein" or BiP.[10] Other workers have abbreviated chaperonins as Cpn70, Cpn60, etc.

The abbreviations Hsp70, Cpn70, DnaK, and BiP all refer to a group of similar 70-kDa proteins that are apparently found in all organisms. Their role seems to be to stabilize unfolded proteins prior to final folding in the cytosol or after translocation into the endoplasmic reticulum (ER) or into mitochondria or other organelles. Each of these proteins consists of two functional domains. A 52-kDa domain at the C terminus binds 7- or 8-residue segments of unfolded peptide chains in an elongated conformation.[9] At some point in its reaction cycle the N-terminal 40-kDa domain, which binds ATP tightly, causes the ATP to be hydrolyzed to ADP and inorganic phosphate: $ATP + H_2O \rightarrow ADP + P_i$. Binding and release of a polypeptide by the *E. coli* DnaK protein is coupled tightly to this exergonic **ATPase** reaction.[10a,10b] The reaction is dependent upon potassium ions[11] and is regulated by two **co-chaperones, DnaJ** (Hsp40)[11a-11c] and **GrpE**.[10a] Both ATP and extended polypeptides bind weakly to DnaK, and the ATP in the DnaK•polypeptide•ATP complex is hydrolyzed slowly to ADP and inorganic phosphate.[10a,10b,12] Co-chaperone DnaJ, which shares a largely α-helical structural motif with J-domains in various other proteins, binds to the complex.[11a, 11b, 11c] It probably induces a conformational change that leads to rapid hydrolysis of ATP. In the resulting complex both ADP and the polypeptide are bound tightly to the DnaK protein and dissociate from it very slowly. The *E. coli* co-chaperone GrpE acts as a **nucleotide exchange factor** that catalyzes rapid loss of ADP from the complex. If ATP binds to this DnaK•polypeptide complex the polypeptide is released.[10a] This cycle, which can be repeated, accomplishes the function of DnaK in protecting extended polypeptides and releasing them under appropriate conditions. There is also evidence that DnaK participates in refolding of misfolded proteins.[11d] It cooperates with a ribosome-associated prolylisomerase in bacteria.[11e] The three-dimensional structure of the ATPase domain of Hsp70 is strikingly similar to that of the enzyme **hexokinase** (Chapter 12) and to that of the muscle protein **actin**, Fig. 7-10.[13,14] Archaeal chaperonins lack the Cpn10 ring but have lid-like extensions at the cylinder ends.[15a]

The Cpn60 class of chaperonins are amazing cage-like structures. Each oligomer is composed of two rings, each made up of seven 60-kDa subunits stacked back-to-back. Cpn60 structures from archaeobacteria are similar but may have 8- or 9-subunit rings.[15,15a] The best known member of this group is the *E. coli* protein known as **GroEL** whose three-dimensional structure is depicted in Box 7-A.[16,17] This 14-subunit oligomer of GroEL is a cylinder which is capped by a smaller ring composed of seven 10-kDa subunits of **GroES**, a Cpn10 protein.[17,18] Like Hsp70, GroEL has ATPase activity and an ATP binding site in its large equatorial domain.[16] Archaeal chaperonins lack the Cpn10 ring but have lid-like extensions at the cylinder ends.[15a]

Within the GroEL–GroES cage polypeptide chains can fold without becoming entangled with other proteins or being cleaved by protein-hydrolyzing enzymes of the cytoplasm. An unfolded or partially folded protein may diffuse into the open end of the complex and bind temporarily via noncovalent interactions. The hydrophobic inner surface of the complex may favor the formation of helices within the folding protein. Binding of the polypeptide substrate and capping with GroES causes a major cooperative conformational change with doubling of the internal volume (Box 7-A). The character of the inner wall also changes as a result of exposure of hydrophilic groups.[18a,18b] Perhaps the expansion also stretches segments of the unfolded polypeptide. Subsequent hydrolysis of ATP may be coupled to release of the bound polypeptide, which may then leave the complex or bind again for another chance to fold correctly.[17,19] The two ends of the GroEL•GroES complex may function alternately, with each end in turn receiving the GroES cap.[17]

Hsp90 is one of the most abundant cytosolic proteins in eukaryotic cells, but it seems to chaperone only a few proteins, among which are steroid-hormone receptors.[20,20a] Another group of over 30 different chaperone proteins participate in assembly of external bacterial cell-surface structures such as pili (Fig. 7-9). The chaperone PapD consists of two domains, both having an immunoglobulin-like fold.[20b] PapD binds to the pilin subunits, escorting them to the site of pilus assembly. Folding of integral membrane proteins may be facilitated by a chaperone function of membrane lipids.[20c]

For most proteins the initial synthesis is followed by a sequence of **processing** or **maturation** reactions. These reactions sometimes involve alteration of amino acid side chains and very often include hydrolytic cleavages by which pieces of the peptide chain are cut off. The initially synthesized polypeptides are often not functional and gain biological activity only after one or more pieces have been removed. For example, digestive enzymes are usually secreted as **zymogens** or **proenzymes** which are activated by hydrolytic cleavages only after secretion into the digestive tract. Many proenzymes are components of the extremely complex cascades of activating reactions involved in **blood clotting** (Chapter 12) and in the defensive **complement system** (Chapter 31).

Most peptide hormones are cut out from larger proteins. For example, human insulin is synthesized as a 110-residue **preproinsulin** which is converted in stages to the active two-chain, 51-residue hormone (Eq. 10-8 and Fig. 10-7).[20d,20e]

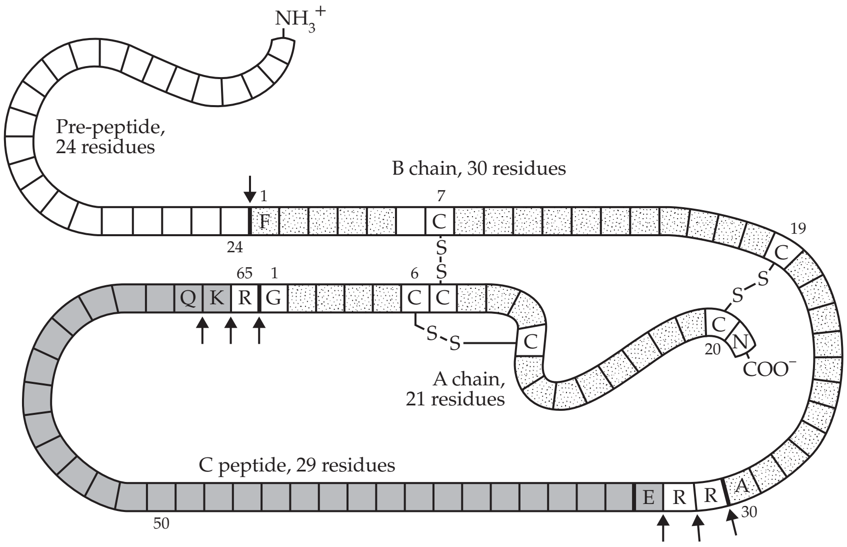


Figure 10-7 Schematic diagram of the structure of human preproinsulin. The 24-residue prepeptide, the 29-residue C-peptide and basic residues 31, 32, 64, and 65 are cut from the peptide upon conversion to insulin as indicated by the small arrows. Some amino acid residues are identified using the one-letter codes. See Fig. 7-17 for details of insulin structure.

$$\underset{\text{(110 residues)}}{\text{preproinsulin}} \rightarrow \underset{\text{(86 residues, one chain)}}{\text{proinsulin}} \rightarrow \underset{\text{(51 residues, two disulfide-linked chains)}}{\text{insulin}} \tag{10-8}$$

The 24-residue prepeptide is cleaved from preproinsulin within a few minutes of synthesis. Then, over a period of about an hour additional cleavages occur to give the final product.[21] This crystallizes as the zinc hexamer (Fig. 7-18) within the "dense cores" of storage vesicles from which it is released into the bloodstream as needed.[22] Other examples are described in Section 3.

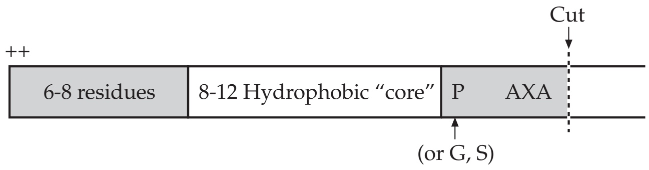
2. Transport of Proteins to Their Destinations within a Cell

The synthesis of peproinsulin on ribosomes in the cytosol and the release of mature insulin from secretion vesicles involve not only the chain cleavages but also transport of the polypeptide across membranes of the ER, Golgi compartments, and secretion vesicles.

Signal sequences and translocation. In 1971, Blobel and Sabatini postulated that for some proteins an N-terminal segment of newly synthesized polypeptide chains contains a signal sequence or leader peptide of 15–30 amino acids which carries information concerning the location of the mature protein in the cell.[23–28b] The rather nonpolar signal sequences of proteins destined for secretion from cells would interact with and pass through the membrane of the ER. On the other side of the membrane, a **signal peptidase** would cut off the signal peptide. In many cases, glycosyltransferases would add sugar residues near the N terminus to create a hydrophilic end which would help to pull the rest of the peptide chain through the membrane. A similar situation would hold for secretion of proteins by bacteria.

The signal hypothesis has been proven correct but with considerable added complexity. It now appears that when the N terminus of a new polypeptide chain carrying the proper signal sequence emerges from a ribosome it is intercepted by a **signal recognition particle (SRP)**, a 250-kDa ribonucleoprotein consisting of six polypeptide chains and a small 300-residue 7S RNA chain.[29–33] Binding to the ribosome, the SRP temporarily blocks further elongation of the peptide chain until it bumps against and binds to a 72-kDa membrane protein called the **SRP receptor** or **docking protein**. Then translation begins again and the protein moves into the ER, where it is cleaved by the signal peptidase and is modified further. Although the bacterial SRP is a single protein called **Ffh** (for 54 homolog),[33a] the basic machinery for signal recognition and secretion is remarkably conserved from *E. coli* to humans.[33b] Bacterial proteins that are destined for secretion or for a function in the periplasmic space or in external membranes also contain signal sequences which are cut off by a signal peptidase embedded in the plasma membrane.[33c]

Signal sequences vary in structure but usually have a net positive charge within the first 5–8 residues at the N terminus. This region is followed by a "hydrophobic core" made up of 8–10 residues with a strong tendency toward α helix formation. This sequence is often followed by one or a few proline, glycine, or serine residues and then a sequence AXA that immediately precedes the cleavage site. Here, A is usually alanine in prokaryotes but may also be glycine, serine, or threonine in eukaryotes. Residue X is any amino acid.[25,27,28a,28b]



It had often been assumed that a hydrophobic signal sequence, perhaps folded into a hairpin loop, spontaneously inserts itself into an ER membrane to initiate translocation. However, study of the genetics of protein transport suggests otherwise. Over 50 different genetic loci affect the translocation of proteins in yeast.[31,32,33d] Products of these **secretory genes**, which are named *SEC* 61, *SEC* 62, *SEC* 63, etc., are proteins with corresponding names, e.g., Sec 61 protein (or Sec 61p). Study of these proteins suggested that the Sec 61, Sec 62, and Sec 63 proteins are all directly involved in transport into the ER.[31,34,35] Additional proteins are needed for movement of vesicles out from the Golgi and for delivery of secretory vesicles to the plasma membrane.[33d] The protein Sec 61p has been identified as homologous to the α subunit of a similar trimeric αβγ Sec 61p complex of mammalian tissues.[32] It also appears to correspond to Sec Y, a protein required for secretion of proteins through the plasma membrane of *E. coli* into the periplasmic space. Proteins corresponding to the β and γ subunits of mammalian related proteins have been identified in yeast and *E. coli*[36] and in *Arabidopsis*, where it is necessary to bring proteins into the thylakoid membrane of the chloroplasts.[37]

In *E. coli* the products of genes *SecY*, *SecE*, and *SecG* are integral membrane proteins that, together with additional proteins, form a proteinaceous pore through which proteins pass.[38–40a] Additional proteins required for translocation of some bacterial proteins are the chaperonin **SecB**, a tetramer of 17.3-kDa subunits,[41] and **SecA**, a soluble ATPase that may participate in docking and serve as an engine for transport throught the pore.[38,40a,42] In eukaryotic cells a membrane-bound Ca^{2+}-dependent chaperonin called **calnexin** assists in bringing the protein into the ER in a properly folded state.[43,44]

Unlike transport across the membranes of the ER, transport across plasma membranes of bacteria often requires both hydrolysis of ATP and energy provided by the membrane electrical potential.[33,38,44–48] Secretion into the periplasmic space has been well characterized but less is known about transport of proteins into the external membranes of *E. coli*.[48] A 16 kDa periplasmic chaperone may be required.[48a]

Many bacteria have a second complete secretion system.[48b] This multi-gene type II system is present but usually inactive in *E. coli*. A third (type III) system is present in many pathogenic bacteria, and has evolved for delivery of specialized structural and regulatory proteins into host cells.[48c,48d]

Ticketing destinations. We have seen that proteins that contain suitable signal sequences are exported from the cytoplasm while other proteins remain. Some proteins are secreted while others take up residence as integral or peripheral membrane proteins or as soluble proteins within an organelle. All of the available evidence indicates that it is the sequence of a protein that determines its destination. Proteins targeted to pass through the inner mitochondrial or chloroplast membrane have 20- to 70-residue presequences that are rich in arginine and lysine and which are removed when the protein reaches its destination.[49–53] Proteins meant to go to the intermembrane space or the outer membrane of mitochondria have a sequence containing basic amino acids followed by a long stretch of uncharged residues.[54] Peroxisomal preproteins may have signal sequences such as SKL at the C terminus[54–56] and proteins destined to become attached to phosphatidylinositol glycan anchors (Fig. 8-13)[57] as well as some bacterial surface proteins[57a] have signal peptides at both N and C termini. Some lysosomal membrane glycoproteins have an LE pair in the N-terminal cytoplasmic tail[58] and proteins with suitable dileucine and related pairs are often taken up in lysosomes.[59] Soluble proteins that are resident in the cisternae of the ER often have the C-terminal sequence KDEL or HDEL.[29,60–64] It may serve both as a **retention signal** and as a **retrieval signal** for return of the protein if it passes on into the Golgi vesicles.[65,66]

Transport of proteins into mitochondria is dependent upon both cytosolic chaperonins and mitochondrial chaperonins of the Hsp70 and Hsp60 classes.[67–70] Entrance into mitochondria[71–73] resembles passage through membranes of the ER. However, entry to the nucleus through the nuclear pore complex requires other localization signals[73a] as well as specialized proteins.[74,74a] The sorting of proteins into six different compartments within a chloroplast requires a whole series of recognition signals as well as chaperonins and channel proteins.[75–79] The various signal sequences that determine a protein's interactions with chaperonins, docking proteins, and proteinaceous pore complexes may be complex and overlapping but evolution has selected sequences that allow cells to live and function. Computer programs for detecting sorting signals are being developed.[79a] Not only proteins but also other macromolecules are automatically sorted to their correct destination.

Vesicular transport and the Golgi system. Movement of secreted proteins through the cytosol from the ER to the external surface occurs through the formation and opening up of small vesicles about 70 nm in diameter. It occurs in steps that involve passage from the ER to the various Golgi membranes. Figure 10-8 provides a sketch of the system of ER, Golgi, secretion vesicles, and lysosomes.[63] Newly formed secretory proteins flow from the ER into an intermediate compartment where vesicles are formed and are carried to the *cis* Golgi network (CGN). They move step-by-step through the Golgi stack (GS) and into the *trans* Golgi network (TGN). New vesicles are formed to carry proteins between each pair of membranous compartments. In each compartment new glycosylation reactions or other modifications may occur. Finally, vesicles carry the mature proteins to the plasma membrane, lysosomes, or vacuoles.[79b] A related process is *uptake* of proteins by endocytosis to form, consecutively, early and late endosomes which fuse with lysosomes (Chapter 8, Section C,6, Fig. 10-8).

At every step in these processes vesicles are formed and are carried to the next destination where the vesicles fuse with the new membrane and discharge their contents.[80–85a] This remarkable process is complex and highly specific. Rothman proposed that the vesicles to be transported are "docked" on appropriate receptor molecules (called SNARES) on the destination membrane. This is accomplished with the aid of specific soluble marker proteins (called SNAPS) with surfaces complementary to those of the receptors.[80] Proteins are sorted in this way according to their destination signals. The Golgi system is considered further in Chapter 20.

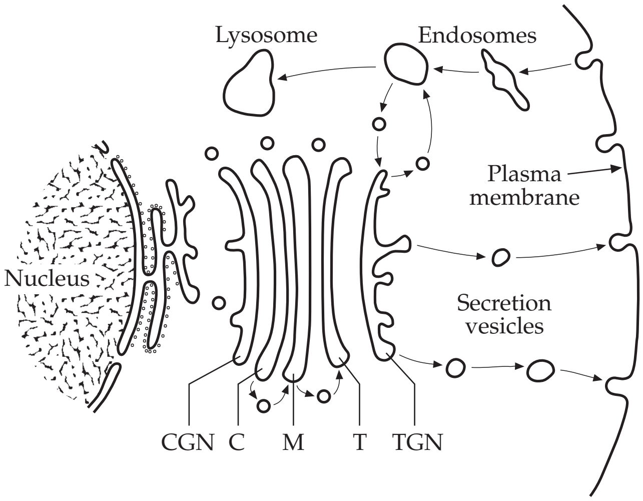
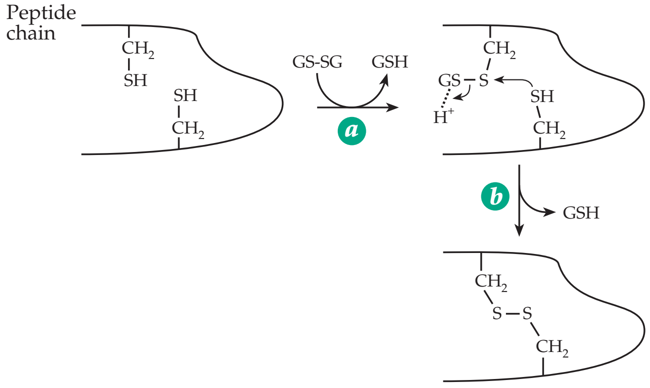


Figure 10-8 Current version of protein synthesis and processing via ER, Golgi, and secretory vesicles. CGN, *cis*-Golgi network: C, T, M are the *cis*, medial, and *trans* compartments of the Golgi stack; TGN, *trans* Golgi network. Arrows indicate some of the movements of transport vesicles.

3. Posttranslational Alterations

Chemical alterations to a protein begins as soon as the peptide chain is formed. Proteins translocated into the periplasmic space of bacteria or into the cisternae of the ER of eukaryotic cells meet several important enzymes. Peptidylprolyl isomerase (discussed in Box 9-F) assists the folding of the peptide chain and protein disulfide isomerases help to form disulfide linkages. The cytosol provides a reducing environment in which many proteins that have cysteine side chains carry free –SH groups. In contrast, the periplasmic space and the ER have more oxidizing environments in which disulfide bridges may form. As shown in Eq. 10-9, the oxidizing agent is often assumed to be the disulfide form of the tripeptide **glutathione** (see Box 11-B),



(10-9)

which is commonly abbreviated GS–SG. It can react with proteins by disulfide exchange reactions to form mixed disulfide groups with release of reduced glutathione, GSH (Eq. 10-9, step *a*). The mixed disulfides can be converted to disulfide bridges by a second exchange reaction (Eq. 10-9, step *b*). However, in the bacterial periplasm the dithiol protein DsbA appears to be the major oxidant (see Box 15-C). It functions together with other proteins, some of which carry electrons from the electron transport chains of the bacteria.[85a,85b] The rates of these reactions are greatly increased by the protein disulfide isomerases.[85a–85e,86–91] These enzymes tend to promote formation of disulfide linkages between the correct pairs of SH groups. If there are three or more –SH groups in a chain some incorrect pairing may, and often does, occur. The protein disulfide isomerases break these bonds and allow new ones to form.[92] The active sites of these isomerases contain pairs of –SH groups which can be oxidized to internal –S–S– bridges by NAD^+-dependent enzymes. These enzymes and their relatives **thioredoxin** and **glutaredoxin** are discussed further in Box 15-C. Glutathione and oxidation–reduction buffering are considered in Box 11-B.

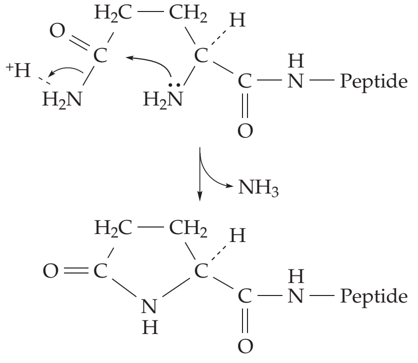
Proteolytic processing. For proteins exported from the cytosol, a signal peptidase (or leader peptidase) is usually waiting in the bacterial plasma membrane or in membranes of the ER. The *E. coli* leader peptidase is an integral membrane protein with its catalytic domain in the periplasm.[45,93] For insulin the 24-residue prepiece serves as the signal sequence which is hydrolyzed off in the ER. However, the other cleavages shown in Fig. 10-7 appear to take place in immature secretory granules.[21,22] For many proteins proteolytic processing begins earlier than this and may occur in more than one location. Eukaryotic processing proteases often resemble the bacterial protease subtilisin[20e] but cut peptide chains preferentially after dibasic amino acid pairs, e.g., Arg-Arg, Arg-Lys.[94–96] There are also other processing enzymes with differing specificities.[97]

Although most polypeptides leave ribosomes as single proteins, some are **polyproteins**, which give rise to two or more functional peptides. Polyproteins occur in all cells but are especially prevalent among virally encoded peptides. For example, the polio virus polyprotein is cut into at least ten pieces by proteases, some present in the host cell normally and some encoded by the virus.[98] Many neurohormones arise from polyproteins that undergo processing as they travel down an axon from the cell body before being secreted into a synapse (Chapter 30). Within our own brains the peptide **prepro-opiomelanocortin** undergoes numerous cleavages to give rise to at least seven different neurohormones (see Fig. 30-2).[96,99–101]

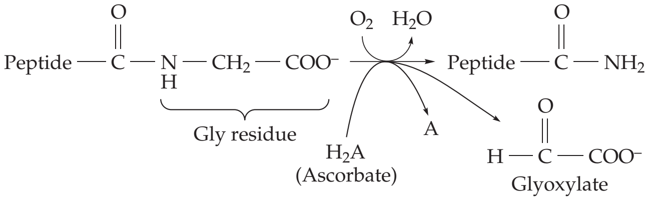
Altered ends. Proteins often contain "blocked" end groups in place of free $-NH_3^+$ or $-COO^-$. For many of the cytosolic proteins, which do not carry leader sequences, the N-terminal methionine that is always used to initiate the ribosomal synthesis of proteins in eukaryotes is removed. If the next residue after the N-terminal methionine is small and uncharged, the methionine is usually hydrolyzed off enzymatically after 30–40 residues have been added to the growing peptide chain.[102–105] An acetyl group is transferred onto the NH_2 termini of about 85% of all cytosolic proteins, whether or not the initiator methionine has been removed.[106] In other cases fatty acyl groups may be transferred to the terminal $-NH_2$ group or to side chain SH groups. For example, a **myristoyl** (tetradecanoyl) group is frequently combined in amide linkage at the N terminus of cellular and virally encoded proteins. A palmitoyl group is joined in thioester linkage to a cysteine side chain near the N terminus of the *E. coli* periplasmic lipoprotein (Fig. 8-28).[107] Together with the N-terminal myristoyl group it forms a membrane anchor for this protein. Polyprenyl groups are transferred onto cysteine –SH groups at or near the C termini of many eukaryotic proteins to form a different membrane anchor.[108–110] Some of the enzymes act on the sequence CAAX, where X = M, S, Q, C but not L or I. After transfer of a farnesyl or geranylgeranyl group onto the cysteine the AAX is removed proteolytically and the new terminal cysteinyl carboxyl group is methylated. Surface proteins of gram-positive bacteria are joined by amide linkage from their C termini to a pentaglycine chain of the peptidoglycan layer (Fig. 8-28)[57a]

A pyroglutamyl N terminus is found in the thyrotropin-releasing hormone (Fig. 2-4) and in many other peptide hormones and proteins. It presumably arises by attack of the α-NH_2 group of an N-terminal glutamine on the side chain amide group with release of NH_3 (Eq. 10-10).[111,112]

Another frequent modification at the C terminus of peptide hormones and of other proteins is **amidation**. In this reaction a C-terminal glycine is oxidatively removed as glyoxylate in an O_2^-, copper- and ascorbate



(10-10)



(10-11)

(vitamin C)-dependent process (Eq. 10-11).[113–117] See also Chapter 18, Section F,2.

An example of both of these modifications is formation of thyrotropin-releasing hormone shown in Fig. 2-4 from its immediate precursor:

$$\text{Gln–His–Pro–Gly} \rightarrow \text{pyroglutamyl–His–Pro–NH}_2 \tag{10-12}$$

Proteins with long C-terminal hydrophobic signal sequences may become attached to phosphatidylinositol-glycan anchors embedded in the plasma membrane (Fig. 8-13). An example is a human alkaline phosphatase in which the α carboxyl of the terminal aspartate residue forms an amide linkage with the ethanolamine part of the anchor. Attachment may occur by a direct attack of the $-NH_2$ group of the ethanolanine on a peptide linkage in a transacylation reaction that releases a 29-residue peptide from the C terminus.[118,119] (See Chapter 29).

Many other covalent modifications of proteins are dealt with in other sections of the book.[120–122] A few are described in Chapter 2, Eqs. 2-14 to 2-22. Reversible alterations used to regulate enzymes are considered in Chapter 11. Of these, the phosphorylation of –OH groups of serine, threonine, and tyrosine is the most important. A large fraction of all cellular proteins appear to be modified in this way. Protein glycosylation, the transfer of glycosyl groups onto –OH side chain groups of serine and threonine (Chapter 4, and Chapter 20) and nonenzymatic glycation (Eq. 4-8) also affect many proteins, often at turns in the peptide chain. Hydroxylation, glycosylation, and other modifications of collagen are described in Chapter 8. Another common reversible alteration is formation of sulfate esters of tyrosine –OH groups.[99,123] Reactions by which cofactors become attached to proteins[124] are described in Chapters 15–17.

Methylation,[125] hydroxylation, and other irreversible modifications often affect specific residues in a protein. Oxidative alterations occur during aging of proteins (Chapter 18).[126] A few proteins even undergo "splicing" that alters the amino acid sequence (Box 29-E).[127] All of these reactions not only affect the properties of the proteins but also participate in driving the turnover of these macromolecules.

4. Intracellular Degradation of Proteins

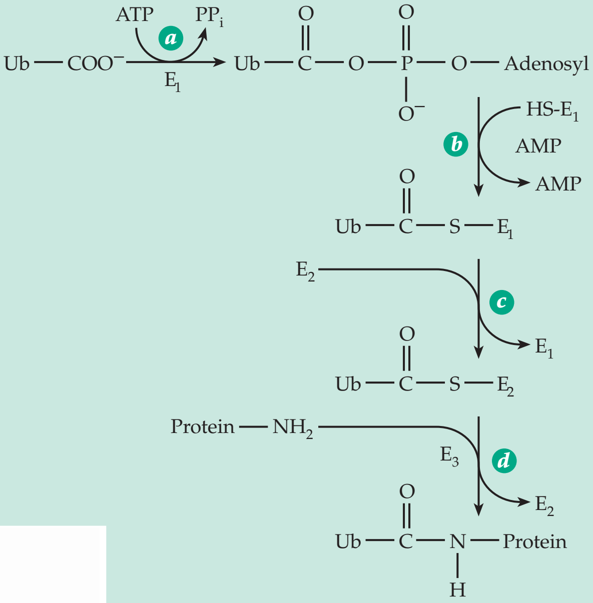
Once a protein has reached its correct location and has acquired its proper function, it usually has a limited lifetime, which may average only a few hours or a few days. The protein is then hydrolytically degraded back to its constituent amino acids.[128] Defective and damaged proteins are usually degraded much more rapidly than are intact proteins.[129,130] Under conditions of starvation, proteins are broken down more rapidly than usual to supply the cell with energy. Rapid degradation of proteins is often induced at certain stages of differentiation. For example, spore-forming bacteria contain a protease that becomes activated upon germination of the spore.[131] Within minutes this enzyme digests stored proteins to provide amino acids for the synthesis of new proteins during growth.

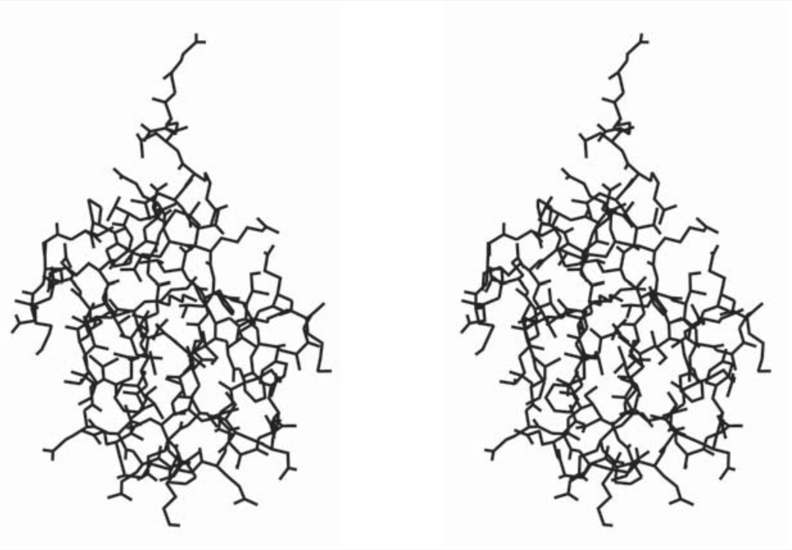
Eukaryotic cells degrade proteins within both the cytosol and lysosomes. Lysosomes apparently take up many proteins but have a preference for N-terminal KFERQ[132] and also for particular types of glycosylation (Chapter 20). Lysosomes act on many long-lived proteins.[133,133a] Once within the lysosomes, the proteins are broken down into amino acids with a half-life of ~8 minutes. During nutritional deprivation, the rate of uptake of proteins by lysosomes increases markedly. The same is true during certain developmental changes, for example, when a tadpole loses its tail.

Many short-lived proteins are degraded within the cytosol in ATP-dependent processes. A major process involves the small protein **ubiquitin** (Box 10-C).[134] Once "labeled" by formation of an isopeptide linkage to ubiquitin, a peptide is attacked by proteases in the **proteasome** complexes (Box 7-A, Chapter 12). There it is quickly degraded. Other proteases, most of which do not require ATP, are also present in the cytoplasm (Chapter 12). How do these enzymes as well as those within the lysosomes work together to produce a harmonious turnover of the very substance of our tissues? How is it possible that one protein has a long half life of many days while another lasts only an hour or two in the same cell? The answer seems to be that

BOX 10-C UBIQUITIN

The small 76-residue protein called ubiquitin[a-c] is probably present in all eukaryotic cells. Found in the nucleus, cytoplasm, cell surface membranes, and extracellular fluids, ubiquitin is often joined by isopeptide linkages from its C-terminal glycine to the ε-amino groups of lysine side chains of other proteins. Ubiquitin has one of the most conserved of all known amino acid sequences. No amino acid substitutions have been found among animal species and only three differences distinguish plant ubiquitins from that of humans. In its three-dimensional structure ubiquitin is compact, tightly hydrogen bonded, and roughly spherical. It contains an α helix, a mixed β sheet, and a distinct hydrophobic core.





Stereoscopic drawing of human ubiquitin. The C terminus is at the top. From Vijay-Kumar *et al.*[c] Courtesy of William J. Cook

Linkage of ubiquitin to other proteins occurs through the action of a ligase system which catalyzes four sequential reactions as shown in the accompanying equation. In step *a* a **ubiquitin activating enzyme** (E_1) forms a C-terminal acyl adenylate by reaction of ubiquitin (Ub) and ATP.[d,e] In step *b* a sulfhydryl group of the same enzyme then displaces the AMP part to form a thioester linkage to ubiquitin. The chemistry of the reaction is the same as that in Eq. 10-1. The activated ubiquitin is next transferred (step *c*) by transacylation to several **ubiquitin-conjugating enzymes** (E_2), also called ubiquitin-carrier proteins.[f-i] These in turn (step *d*) transfer the ubiquitin to amino groups of lysine side chains of target proteins (Prot-NH_2). A third protein, **ubiquitin-protein ligase** (E_3), is sometimes required for this last step. Most E_2 enzymes, acting without E_3, couple ubiquitin with amines or with small basic proteins whose cellular functions are still unclear. For example, an E_2 which is the product of the yeast gene CDC34 appears to function in the **cell cycle** (Fig. 11-15). Its absence from yeast is lethal[f,i,j] and it is clear that ubiquitin-mediated hydrolysis of the specialized proteins called **cyclins** is essential to operation of the cycle.[k,l]

The best understood function of ubiquitin is in nonlysosomal degradation of proteins.[j,m,n] Protein E_3 appears to select the proteins for degradation, binding them and catalyzing the formation of the isopeptide linkage to substrate. After one ubiquitin molecule has been attached, and while still held by E_3, a second activated ubiquitin is coupled to Lys 48 of the first ubiquitin. This process may continue until several molecules of ubiquitin are joined by Gly-Lys isopeptide linkages to form a **polyubiquitin** chain. Sometimes more than one lysine of the substrate becomes polyubiquitinated to form branched chains. A free α-NH_3^+ group on the protein being degraded is essential for rapid conjugation with ubiquitin and certain N-terminal residues such as arginine favor the conjugation and subsequent hydrolytic breakdown. Some proteins become attached to ubiquitin only after arginine is transferred

BOX 10-C (continued)

onto their N-termini from an aminoacyl-tRNA.[o]

Polyubiquitin chains serve as recognition markers that induce rapid hydrolysis of the marked proteins in the 2000 kDa 26S proteasome or **multicatalytic** protease.[p–s] This complex is discussed in Box 7-A and again in Chapter 12. During hydrolytic destruction of the protein the ubiquitin is released for reuse by **ubiquitin carboxyl-terminal hydrolases** or **isopeptidases** which cleave the thioester or isopeptide linkages that tie ubiquitin to proteins.[t,u]

While proteins may be modified to favor rapid ubiquitination, others may be altered to protect them from ubiquitination. For example, calmodulin produced from a cloned gene in bacteria is a good substrate for ubiquitination but within cells it appears to be protected by the posttranslational conversion of Lys 115 to trimethyllysine.[v]

About 10% of the histone H2A present in higher eukaryotes is ubiquitinated at Lys 119.[w] In the slime mold *Physarum* the content of ubiquitinated histones H2A and H2B changes rapidly during the various stages of mitosis. Apparently, ubiquitin must be cleaved from the histones to permit packaging of DNA into metaphase chromosomes and must become attached to the histones in some regions of the chromosomes to allow unfolding of the highly packed nucleosomes. A yeast enzyme that attaches ubiquitin to histones is encoded by the gene *RAD6*, which is required for DNA repair, sporulation, and other cellular processes. Terminal differentiation of reticulocytes to form erythrocytes involves loss of specific enzymes as well as of entire mitochondria. These processes also depend upon ubiquitin. Ubiquitin is one of the components of the paired helical filaments present in brains of persons with Alzheimer's disease.[x]

In most organisms there are two arrangements for ubiquitin genes. There is a cluster of up to 100 tandomly repeated genes whose transcription gives rise to polyubiquitin, a chain of ubiquitin molecules joined by Gly-Met linkages. These must be cleaved, perhaps by the same ubiquitin C-terminal hydrolase that releases ubiquitin from its conjugates.[y] Other ubiquitin genes are fused to genes encoding ribosomal proteins. The resulting polyproteins have the ribosomal peptides fused to the C termini of the ubiquitin sequences and must be proteolytically cleaved to give mature proteins.[z]

Recently a variety of modifiers of ubiquitin ligases have been discovered[aa,bb] as have ubiquitin-like domains in other proteins.[cc] These findings elucidate the complexity of the sorting of proteins and removal of improperly folded and otherwise defective proteins from the secretory pathway and return to the proteasomes in the cytosol.[dd,ee] They also suggest important roles for ubiquitination in a broad range of metabolic controls.

---

[a] Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D., Vierstra, R. D., Hatfield, P. M., and Cook, W. J. (1987) *J. Biol. Chem.* **262**, 6396–6399

[b] Vierstra, R. D., Langan, S. M., and Schaller, G. E. (1986) *Biochemistry* **25**, 3105–3108

[c] Vijay-Kumar, S., Bugg, C. E., and Cook, W. J. (1987) *J. Mol. Biol.* **194**, 531–544

[d] Hershko, A. (1991) *Trends Biochem. Sci.* **16**, 265–268

[e] Pickart, C. M., Kasperek, E. M., Beal, R., and Kim, A. (1994) *J. Biol. Chem.* **269**, 7115–7123

[f] Jentsch, S., Seufert, W., Sommer, T., and Reins, H.-A. (1990) *Trends Biochem. Sci.* **15**, 195–198

[g] Cook, W. J., Jeffrey, L. C., Sullivan, M. L., and Vierstra, R. D. (1992) *J. Biol. Chem.* **267**, 15116–15121

[h] Cook, W. J., Jeffrey, L. C., Xu, Y., and Chau, V. (1993) *Biochemistry* **32**, 13809–13817

[i] Blumenfeld, N., Gonen, H., Mayer, A., Smith, C. E., Siegel, N. R., Schwartz, A. L., and Ciechanover, A. (1994) *J. Biol. Chem.* **269**, 9574–9581

[j] Ciechanover, A., and Schwartz, A. L. (1994) *FASEB J.* **8**, 182–191

[k] Barinaga, M. (1995) *Science* **269**, 631–632

[l] Dorée, M., and Galas, S. (1994) *FASEB J.* **8**, 1114–1121

[m] Pickart, C. M. (2000) *Trends Biochem. Sci.* **25**, 544–548

[n] Johnson, E. S., Ma, P. C. M., Ota, I. M., and Varshavsky, A. (1995) *J. Biol. Chem.* **270**, 17442–17456

[o] Ciechanover, A., Ferber, S., Ganoth, D., Elias, S., Hershko, A., and Arfin, S. (1988) *J. Biol. Chem.* **263**, 11155–11167

[p] Goldberg, A. L. (1995) *Science* **268**, 522–523

[q] Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) *Science* **268**, 533–539

[r] Thrower, J. S., Hoffman, L., Rechsteiner, M., and Pickart, C. M. (2000) *EMBO J.* **19**, 94–102

[s] Peters, J.-M. (1994) *Trends Biochem. Sci.* **19**, 377–382

[t] Stein, R. L., Chen, Z., and Melandri, F. (1995) *Biochemistry* **34**, 12616–12623

[u] Wilkinson, K. D., Tashayev, V. L., O'Connor, L. B., Larsen, C. N., Kasperek, E., and Pickart, C. M. (1995) *Biochemistry* **34**, 14535–14546

[v] Johnston, S. C., Riddle, S. M., Cohen, R. E., and Hill, C. P. (1999) *EMBO J.* **18**, 3877–3887

[w] Davie, J. R., and Murphy, L. C. (1990) *Biochemistry* **29**, 4752–4757

[x] Mori, H., Kondo, J., and Ihara, Y. (1987) *Science* **235**, 1641–1644

[y] Finley, D., and Varshavsky, A. (1985) *Trends Biochem. Sci.* **10**, 343–347

[z] Baker, R. T., Tobias, J. W., and Varshavsky, A. (1992) *J. Biol. Chem.* **267**, 23364–23375

[aa] Tyers, M., and Willems, A. R. (1999) *Science* **284**, 601–604

[bb] Turner, G. C., Du, F., and Varshavsky, A. (2000) *Nature (London)* **405**, 579–583

[cc] Hochstrasser, M. (2000) *Science* **289**, 563–564

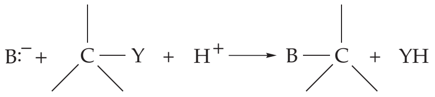
[dd] Ellgaard, L., Molinari, M., and Helenius, A. (1999) *Science* **286**, 1882–1888

[ee] Plemper, R. K., and Wolf, D. H. (1999) *Trends Biochem. Sci.* **24**, 266–270

TABLE 10-1
Types of Biochemical Reactions with Ionic Mechanisms

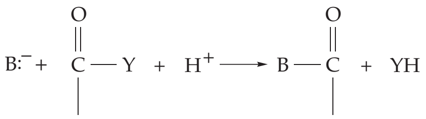
1. Nucleophilic displacement, often via an addition–elimination sequence

A.



(Examples: methyl-, glycosyl-transferases)

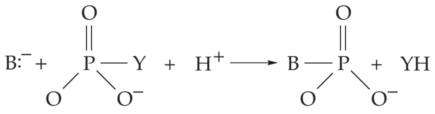
B.



$$\mathrm{B{:}^-} + \overset{\mathrm{O}}{\overset{\|}{\mathrm{C}}}\!-\!\mathrm{Y} + \mathrm{H^+} \longrightarrow \mathrm{B}\!-\!\overset{\mathrm{O}}{\overset{\|}{\mathrm{C}}} + \mathrm{YH}$$

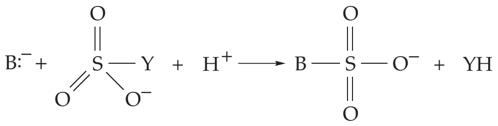
(Examples: esterases, peptidases)

C.



(Examples: phosphatases, kinases, phosphomutases, ribozymes)

D.



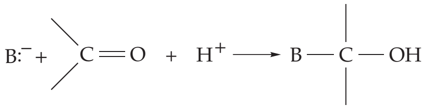
(Examples: sulfatases, sulfotransferases)

2. Addition

A. To polarized double bonds such as $>\!C{=}O$

and $>\!C{=}N-$

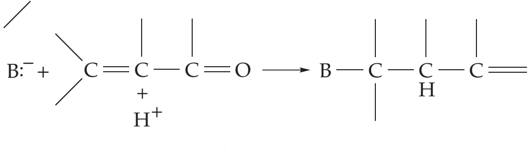
(This reaction most often occurs as a step in an enzymatic process e.g., formation of hemiacetals, hemiketals, hemimercaptals, carbinolamines)



(Example: mutarotase)

B. To double bond conjugated with $>\!C{=}O$ or

$>\!C{=}N-$



(Examples: acyl-CoA, fumarate hydratases, aspartase)

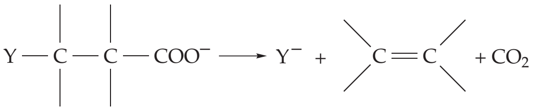
C. To isolated double bonds *(Example: oleate hydratase)* *(see p. 689)*

3. Elimination

A. and B. Precisely the opposite of *addition*

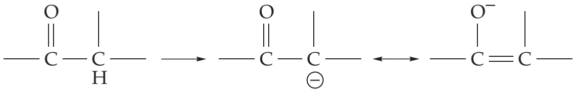
(Eliminations that are the reverse of type 2A are frequent steps in more complex enzyme mechanisms)

C. Decarboxylative elimination



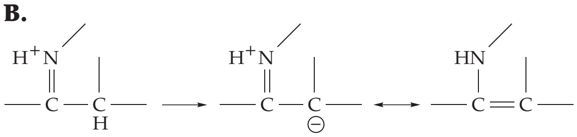
4. Formation of stabilized enolate anions and enamines

A.



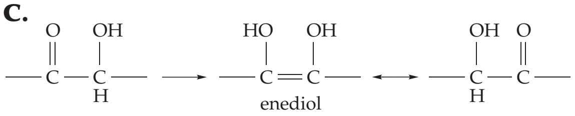
$$-\overset{\mathrm{O}}{\overset{\|}{\mathrm{C}}}-\underset{\mathrm{H}}{\overset{|}{\mathrm{C}}}- \longrightarrow -\overset{\mathrm{O}}{\overset{\|}{\mathrm{C}}}-\underset{\ominus}{\overset{|}{\mathrm{C}}}- \longleftrightarrow -\overset{\mathrm{O^-}}{\overset{|}{\mathrm{C}}}=\overset{|}{\mathrm{C}}-$$

B.



Isomerization reactions

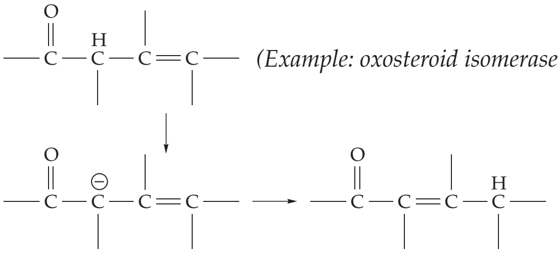
C.



$$-\overset{\mathrm{O}}{\overset{\|}{\mathrm{C}}}-\underset{\mathrm{H}}{\overset{\mathrm{OH}}{\overset{|}{\mathrm{C}}}}- \longrightarrow \underset{\text{enediol}}{-\overset{\mathrm{HO}}{\overset{|}{\mathrm{C}}}=\overset{\mathrm{OH}}{\overset{|}{\mathrm{C}}}-} \longleftrightarrow -\underset{\mathrm{H}}{\overset{\mathrm{OH}}{\overset{|}{\mathrm{C}}}}-\overset{\mathrm{O}}{\overset{\|}{\mathrm{C}}}-$$

(Example: sugar isomerases)

D.

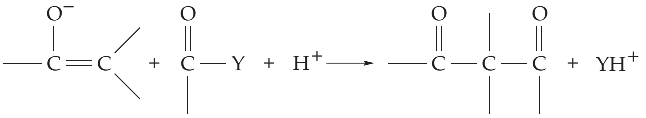


(Example: oxosteroid isomerases)

TABLE 10-1
(continued)

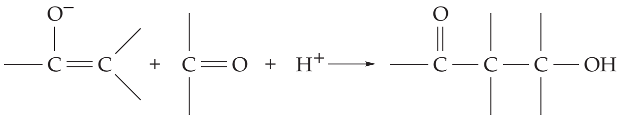
5. Stabilized enolate anions as nucleophiles: formation of carbon–carbon bonds (β condensation)

A. Displacement on a carbonyl group



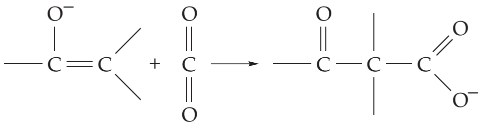
(Example: 3-Oxoacyl–CoA transferase (Y = —S—CoA)

B. Addition to a carbonyl group: aldol condensation



(Examples: aldolases, citrate synthases)

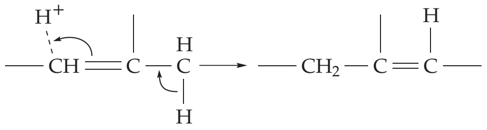
C. Addition to a carbon dioxide (β carboxylation); decarboxylation



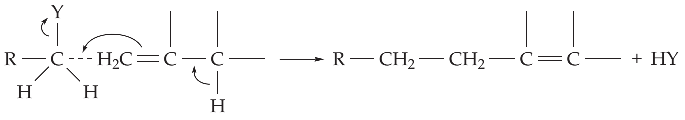
(Examples: phosphoenolpyruvate carboxylase, oxaloacetate decarboxylases)

6. Some isomerization and rearrangement reactions

A. Allylic rearrangement (1,3-proton shift)

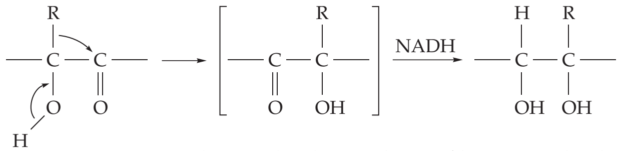


B. Allylic rearrangement with condensation



(Example: condensation of dimethylallyl pyrophosphate with isopentenyl pyrophosphate)

C. Rearrangements with alkyl or hydride ion shift



Glyoxalase

(Examples: biosynthesis of leucine and valine, xylose isomerase, glyoxalase)

the turnover rate of a protein is determined in large part by its sequence. Some proteins are tightly folded and have few bends on the outside that have sequences meeting the specificity requirements of intracellular proteolytic enzymes. These have long half lives. Other proteins may have external loops with sequences susceptible to attack or sequences that favor rapid reaction with the ubiquitin system or uptake into lysosomes. Prematurely terminated proteins and peptide fragments from partial degradation of proteins may tend to be unfolded at the N terminus and to be attacked rapidly.[128,135] Proteins that have undergone covalent modifications or oxidative damage also seem to be hydrolyzed rapidly.[136]

Regions rich in proline, glutamate, serine, and threonine (PEST regions) may be good substrates for Ca^{2+}-activated cytosolic proteases.[137] The ubiquitin system appears to act most slowly on a protein when the normal initiation amino acid methionine is present at the N terminus. For example, the half-life of β-galactosidase in yeast is over 20 hours. Replacement of the methionine with S, A, T, V, or G has little effect. However, replacement with other amino acids shortens the half-life as follows: I and E, 30 min.; Y and Q, 10 min.; F, L, D, and K, 3 min.; and R, 2 min.[138–140a]

5. Turnover of Nucleic Acids

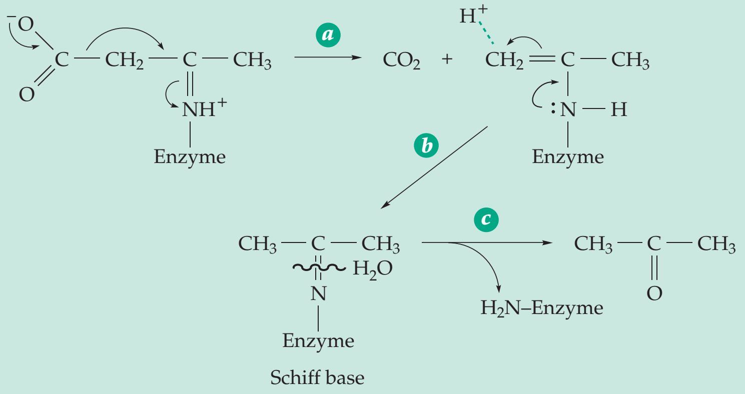
Because of its special role in carrying genetic information, DNA is relatively stable. An elaborate system of repair enzymes (Chapter 27) act to correct errors and to help DNA to preserve its genetic information. However, in some specialized cells such as those forming immunoglobulin, the DNA too undergoes major rearrangements (Chapter 31). RNA molecules are subject to extensive processing. This includes the conversion of RNA bases to modified forms, chain cleavages during maturation of ribosomal and transfer RNAs, cutting and splicing of gene transcripts to form mRNAs and finally degradation of the mRNA (Chapter 28). Proteins that serve as **RNA chaperones** assist in folding these molecules.[141]

BOX 10-D DRAWING THOSE LITTLE ARROWS

Organic mechanisms are often indicated by arrows that show the flow of electrons in individual steps of a reaction. Many errors are made by students on exams and even in published research papers. The arrows are often drawn backward, are too numerous, or do not clearly indicate electron flow. Here are some tips.

1

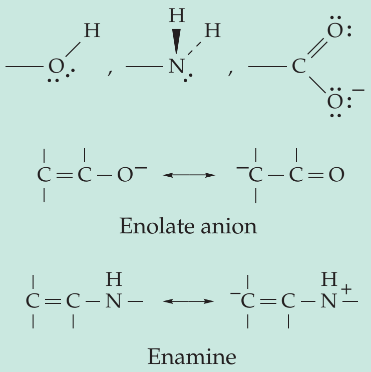
Always write a mechanism step-by-step. Never combine two steps (e.g., *a* and *b* at right) *in which electron flow occurs in opposite directions*. Notice that step *c*, the hydrolysis of a Schiff base, is also a two-step process. The reaction is a familiar one that is commonly indicated as shown here. However, this scheme does not show a detailed mechanism for step *c*.



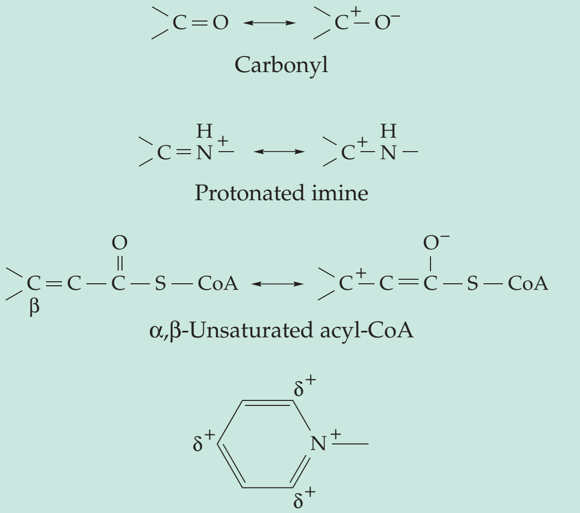
2

Identify the **nucleophilic** and **electrophilic** centers before starting to write a mechanism. Oxygen, sulfur, and nitrogen atoms are usually nucleophilic, e.g., those in —OH, —NH_2, —COO^-, —SH, —OPO_3^{2-}, and enolate anions. The weak nucleophiles —OH and —SH may be converted into strong nucleophiles —O^- —S^- by removal of a proton by a basic group of an enzyme. The nonnucleophilic —NH_3^+ becomes the good nucleophile —NH_2 by loss of a proton. Nucleophilic centers contain unshared electron pairs. Nucleophiles are basic but the basicity, as indicated by proton binding (by the pK_a), is not necessarily proportional to nucleophilic strength (nucleophilicity). Enolate anions and enamines provide nucleophilic centers on carbon atoms, important in formation and cleavage of C—C bonds.

Electrophilic centers include acidic hydrogen atoms, metal ions, the carbon atoms of carbonyl groups, and the β-carbon atoms of α, β unsaturated acids, ketones, or acyl-CoA derivatives. Highly polarized groups such as carbonyl and enamine generate electrophilic centers as indicated by the positive charges. They also affect more distant positions in conjugated systems, e.g., in α, β-unsaturated acyl-CoA derivatives, and in intermediates formed from thiamin diphosphate and pyridoxal phosphate.



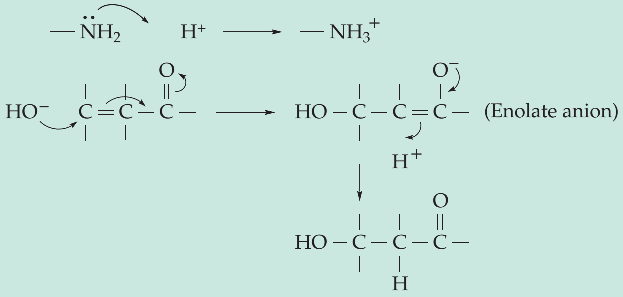
Groups with nucleophilic centers indicated by unshared electron pairs and/or negative charge.

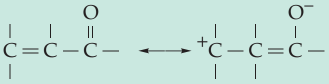


Some electrophilic centers are indicated by + or δ^+

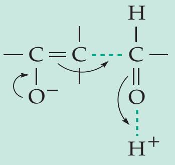
BOX 10-D (continued)

3 When an ionic organic reaction (the kind catalyzed by most enzymes) occurs a nucleophilic center joins with an electrophilic center. We use arrows to show the movement of pairs of electrons. The movement is always *away* from the nucleophile which can be thought of as "attacking" an electrophilic center. Notice the first step in the second example at right. The unsaturated ketone is polarized initially. However, this is not shown as a separate step. Rather, the flow of electrons from the double bond, between the α- and β-carbons into the electron-accepting C=O groups, is coordinated with the attack by the nucleophile. Dotted lines are often used to indicate bonds that will be formed in a reaction step, e.g., in an aldol condensation (right). Dashed or dotted lines are often used to indicate partially formed and partially broken bonds in a transition state, e.g., for the aldol condensation (with prior protonation of the aldehyde). However, *do not put arrows on transition state structures.*

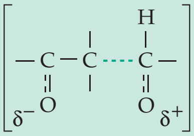




Polarized unsaturated ketone

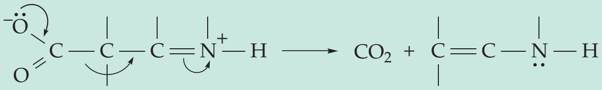


Aldol condensation



Transition state

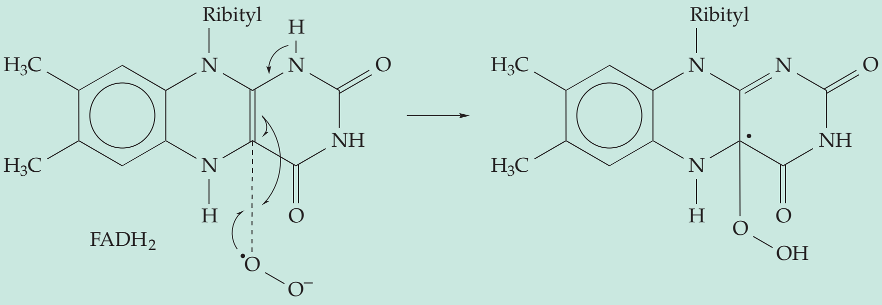
4 In a given reaction step all of the arrows must point the same way. *The arrows point into bonds that are forming* or toward atoms that will carry an unshared electron pair in the product. *Arrows originate from unshared electron pairs or from bonds that are breaking.*



All arrows point the same way

5 Never start an arrow from the same bond that another arrow is forming; i.e., electrons flow out of *alternate* bonds.

6 For reactions of radicals (homolytic reactions) arrows are used to indicate motion of single electrons rather than of electron pairs. It is desirable to use arrows with *half-heads*. For example, reaction of a superoxide radical (•O$_2^-$) with FADH$_2$ could occur as follows:



Half-heads on arrows indicate one-electron movement

E. Classifying Enzymatic Reactions

The majority of enzymes appear to contain, in their active centers, only the side chains of amino acids. Most of the reactions catalyzed by these enzymes can be classified into a small number of types as is indicated in Table 10-1: (Type 1) **displacement** or **substitution** reactions in which one base or nucleophile replaces another, (Type 2) **addition** reactions in which a reagent adds to a double bond, and (Type 3) **elimination** reactions by which groups are removed from a substrate to create double bonds. Note that the latter are the reverse of addition reactions. Two other groups of reactions depend upon formation of transient **enolic forms**. These include (Type 4) **isomerases** and (Type 5) **lyases**, a large and important group of reactions that form or cleave carbon–carbon bonds. Finally, there is a group (Type 6) of **isomerization** and **rearrangement** reactions that do not appear to fit any of the foregoing categories. Another quite different group of enzyme-catalyzed reactions, which are considered in Chapter 16, function with the participation of free radical intermediates.

Biochemical displacement reactions include all of the hydrolytic reactions by which biopolymers are broken down to monomers as well as most of the reactions by which the monomers are linked together to form polymers. Addition reactions are used to introduce oxygen, nitrogen, and sulfur atoms into biochemical compounds and elimination reactions often provide the driving force for biosynthetic sequences. Complex enzymatic processes are often combinations of several steps involving displacement, addition, or elimination. The reactions involving formation or cleavage of C–C bonds are essential to biosynthesis and degradation of the various carbon skeletons found in biomolecules, while the isomerization reactions provide connecting links between the other kinds of reactions in the establishment of metabolic pathways.

In Chapters 12 and 13 the individual reactions of metabolism are classified into these types and the enzymes that catalyze them are described in some detail. The chemistry of coenzymes and metalloenzymes are presented systematically in Chapters 14 to 16, and in Chapter 17 the logic of the combining of individual reactions into metabolic sequences is considered. It is not necessary to read Chapters 12–16 in their entirety since much of their content is reference material. In the later chapters on metabolism, cross-references point out the discussions of individual enzymes in Chapters 12–16.

The following are topics that may be especially valuable to the student and which might be read initially: in Chapter 12, lysozyme (Section B,5), chymotrypsin (Section C,1), kinases (Section D,9), multiple displacement, reactions (Section G); in Chapter 13, imines (Section A,2), addition to C=C bonds (Section A, 4,5), beta cleavage and condensation (Section C); in Chapter 14, thiamin diphosphate (Section D), pyridoxal phosphate (Section E); in Chapter 15, NAD (Section A).

References

1. Iwata, S., Ostermeier, C., Ludwig, B., and Michel, H. (1995) *Nature* **376**, 660–669
2. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., and Yoshikawa, S. (1995) *Science* **269**, 1069–1074
3. Mitchell, P. (1966) *Biol. Rev. Cambridge Philos. Soc.* **41**, 445–502
4. Mitchell, P. (1979) *Science* **206**, 1148–1159
5. Nicholls, D. G., and Ferguson, S. J. (1992) *Bioenergetics 2*, Academic Press, London
6. Gething, M.-J., and Sambrook, J. (1992) *Science* **355**, 33–45
7. Frydman, J., Nimmesgern, E., Ohtsuka, K., and Hartl, F. U. (1994) *Nature* **370**, 111–117
8. Buchner, J. (1994) *Trends Biochem. Sci.* **19**, 559
9. Hartl, F.-U., Hlodan, R., and Langer, T. (1994) *Trends Biochem. Sci.* **19**, 20–25
10. Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991) *Nature* **353**, 726–730

10a. Russell, R., Jordan, R., and McMacken, R. (1998) *Biochemistry* **37**, 596–607

10b. Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F.-U., and Kuriyan, J. (1997) *Science* **276**, 431–435

11. Wilbanks, S. M., and McKay, D. B. (1995) *J. Biol. Chem.* **270**, 2251–2257

11a. Kelley, W. L. (1998) *Trends Biochem. Sci.* **23**, 222–227

11b. Martinez-Yamout, M., Legge, G. B., Zhang, O., Wright, P. E., and Dyson, H. J. (2000) *J. Mol. Biol.* **300**, 805–818

11c. Westermann, B., and Neupert, W. (1997) *J. Mol. Biol.* **272**, 477–483

11d. Mogk, A., Tomoyasu, T., Goloubinoff, P., Rüdiger, S., Röder, D., Langen, H., and Bukau, B. (1999) *EMBO J.* **18**, 6934–6949

11e. Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999) *Nature (London)* **400**, 693–696

12. McCarty, J. S., Buchberger, A., Reinstein, J., and Bukau, B. (1995) *J. Mol. Biol.* **249**, 126–137

13. Flaherty, K. M., DeLuca-Flaherty, C., and McKay, D. B. (1990) *Nature* **346**, 623–628

14. Flaherty, K. M., Wilbanks, S. M., DeLuca-Flaherty, C., and McKay, D. B. (1994) *J. Biol. Chem.* **269**, 12899–12907

15. Phipps, B. M., Typke, D., Hegeri, R., Volker, S., Hoffmann, A., Stetter, K. O., and Baumeister, W. (1993) *Nature* **361**, 475–477

15a. Schoehn, G., Hayes, M., Cliff, M., Clarke, A. R., and Saibil, H. R. (2000) *J. Mol. Biol.* **301**, 323–332

16. Braig, D., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1994) *Nature* **371**, 578–586

17. Hayer-Hartl, M. K., Martin, J., and Hartl, F. U. (1995) *Science* **269**, 836–841

18. Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., Clarke, A. R., and Saibil, H. R. (1994) *Nature* **371**, 261–264

18a. Betancourt, M. R., and Thirumalai, D. (1999) *J. Mol. Biol.* **287**, 627–644

18b. Wang, Z., Feng, H.-p, Landry, S. J., Maxwell, J., and Gierasch, L. M. (1999) *Biochemistry* **38**, 12537–12546

19. Shtilerman, M., Lorimer, G. H., and Englander, S. W. (1999) *Science* **284**, 822–825

20. Weaver, A. J., Sullivan, W. P., Felts, S. J., Owen, B. A. L., and Toft, D. O. (2000) *J. Biol. Chem.* **275**, 23045–23052

20a. Buchner, J. (1999) *Trends Biochem. Sci.* **24**, 136–141

20b. Sauer, F. G., Fütterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J., and Waksman, G. (1999) *Science* **285**, 1058–1061

20c. Bogdanov, M., and Dowhan, W. (1999) *J. Biol. Chem.* **274**, 36827–36830

20d. Orci, L., Vassalli, J.-D., and Perrelet, A. (1988) *Sci. Am.* **259**(Sep), 85–94

20e. Zhou, A., Webb, G., Zhu, X., and Steiner, D. F. (1999) *J. Biol. Chem.* **274**, 20745–20748

21. Smeekens, S. P., Montag, A. G., Thomas, G., Albiges-Rizo, C., Carroll, R., Benig, M., Phillips, L. A., Martin, S., Ohagi, S., Gardner, P., Swift, H. H., and Steiner, D. F. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8822–8826

22. Huang, X. F., and Arvan, P. (1995) *J. Biol. Chem.* **270**, 20417–20423

23. Blobel, G., and Sabatini, D. D. (1971) in *Biomembranes*, Vol. 2 (Manson, L. A., ed), pp. 193–195, Plenum, New York

24. Blobel, G., and Dobberstein, B. (1975) *J. Cell Biol.* **67**, 852

References

25. Landry, S. J., and Gierasch, L. M. (1991) *Trends Biochem. Sci.* **16**, 159–163
26. Gierasch, L. M. (1989) *Biochemistry* **28**, 923–930
27. Jain, R. G., Rusch, S. L., and Kendall, D. A. (1994) *J. Biol. Chem.* **269**, 16305–16310
28. Andersson, H., and von Heijne, G. (1993) *J. Biol. Chem.* **268**, 21389–21393
28a. von Heijne, G. (1998) *Nature (London)* **396**, 111–113
28b. Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000) *J. Mol. Biol.* **300**, 1005–1016
29. Verner, K., and Schatz, G. (1988) *Science* **241**, 1307–1313
30. Rapoport, T. A. (1992) *Science* **258**, 931–935
30a. Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., and Doudna, J. A. (2000) *Science* **287**, 1232–1239
31. Sanders, S. L., and Schekman, R. (1992) *J. Biol. Chem.* **267**, 13791–13794
32. Hartmann, E., Sommer, T., Prehn, S., Görlich, D., Jentsch, S., and Rapoport, T. A. (1994) *Nature* **367**, 654–657
32a. Clemons, W. M., Jr., Gowda, K., Black, S. D., Zwieb, C., and Ramakrishnan, V. (1999) *J. Mol. Biol.* **292**, 697–705
33. Wickner, W. T. (1994) *Science* **266**, 1197–1198
33a. Moser, C., Mol, O., Goody, R. S., and Sinning, I. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11339–11344
33b. Samuelson, J. C., Chen, M., Jiang, F., Möller, I., Wiedmann, M., Kuhn, A., Phillips, G. J., and Dalbey, R. E. (2000) *Nature (London)* **406**, 637–640
33c. Paetzel, M., Strynadka, N. C. J., Tschantz, W. R., Casareno, R., Bullinger, P. R., and Dalbey, R. E. (1997) *J. Biol. Chem.* **272**, 9994–10003
33d. Guo, W., Grant, A., and Novick, P. (1999) *J. Biol. Chem.* **274**, 23558–23564
34. Noël, P. J., and Cartwright, I. L. (1994) *EMBO J.* **13**, 5253–5261
35. Brodsky, J. L., Goeckeler, J., and Schekman, R. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9643–9646
36. Dobberstein, B. (1994) *Nature* **367**, 599–600
37. Laidler, V., Chaddock, A. M., Knott, T. G., Walker, D., and Robinson, C. (1995) *J. Biol. Chem.* **270**, 17664–17667
38. Dalbey, R. E., and Robinson, C. (1999) *Trends Biochem. Sci.* **24**, 17–22
39. Douville, K., Price, A., Eichler, J., Economou, A., and Wickner, W. (1995) *J. Biol. Chem.* **270**, 20106–20111
39a. Eichler, J., Brunner, J., and Wickner, W. (1997) *EMBO J.* **16**, 2188–2196
39b. Matsumoto, G., Yoshihisa, T., and Ito, K. (1997) *EMBO J.* **16**, 6384–6393
39c. Scotti, P. A., Urbanus, M. L., Brunner, J., de Gier, J.-W. L., von Heijne, G., van der Does, C., Driessen, A. J. M., Oudega, B., and Luirink, J. (2000) *EMBO J.* **19**, 542–549
40. Meyer, T. H., Ménétret, J.-F., Breitling, R., Miller, K. R., Akey, C. W., and Rapoport, T. A. (1999) *J. Mol. Biol.* **285**, 1789–1800
40a. Yahr, T. L., and Wickner, W. T. (2000) *EMBO J.* **19**, 4393–4401
41. Randall, L. L., and Hardy, S. J. S. (1995) *Trends Biochem. Sci.* **20**, 65–69
42. Ulbrandt, N. D., London, E., and Oliver, D. B. (1992) *J. Biol. Chem.* **267**, 15184–15192
43. Bergeron, J. J. M., Brenner, M. B., Thomas, D. Y., and Williams, D. B. (1994) *Trends Biochem. Sci.* **19**, 124–128
44. Jungery, M., Pasvol, G., Newbold, C. I., and Weatherall, D. J. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1018–1022
45. Cao, G., Kuhn, A., and Dalbey, R. E. (1995) *EMBO J.* **14**, 866–875
46. Kawasaki, S., Mizushima, S., and Tokuda, H. (1993) *J. Biol. Chem.* **268**, 8193–8198
47. Wickner, W., Driessen, A. J. M., and Hartl, F.-U. (1991) *Ann. Rev. Biochem.* **60**, 101–124
48. Matsuyama, S.-i, Tajima, T., and Tokuda, H. (1995) *EMBO J.* **14**, 3365–3372
48a. Schäfer, U., Beck, K., and Müller, M. (1999) *J. Biol. Chem.* **274**, 24567–24574
48b. Russel, M. (1998) *J. Mol. Biol.* **279**, 485–499
48c. Galán, J. E., and Collmer, A. (1999) *Science* **284**, 1322–1328
48d. Tamano, K., Aizawa, S.-I., Katayama, E., Nonaka, T., Imajoh-Ohmi, S., Kuwae, A., Nagai, S., and Sasakawa, C. (2000) *EMBO J.* **19**, 3876–3887
49. Schatz, G. (1993) *Protein Sci.* **2**, 141–146
50. Pfeffer, S. R., and Rothman, J. E. (1987) *Ann. Rev. Biochem.* **56**, 829–852
51. Lodish, H. F. (1988) *J. Biol. Chem.* **263**, 2107–2110
52. Smeekens, S., Bauerle, C., Hageman, J., Keegstra, K., and Weisbeek, P. (1986) *Cell* **46**, 365–375
53. Roise, D., and Schatz, G. (1988) *J. Biol. Chem.* **263**, 4509–4511
54. Hurt, E. C., and van Loon, A. P. G. M. (1986) *Trends Biochem. Sci.* **11**, 204–207
55. Aitchison, J. D., Murray, W. W., and Rachubinski, R. A. (1991) *J. Biol. Chem.* **266**, 23197–23203
56. Wolins, N. E., and Donaldson, R. P. (1994) *J. Biol. Chem.* **269**, 1149–1153
57. Takeda, J., and Kinoshita, T. (1995) *Trends Biochem. Sci.* **20**, 367–371
57a. Mazmanian, S. K., Liu, G., Ton-That, H., and Schneewind, O. (1999) *Science* **285**, 760–762
58. Ogata, S., and Fukuda, M. (1994) *J. Biol. Chem.* **269**, 5210–5217
59. Pond, L., Kuhn, L. A., Teyton, L., Schutze, M.-P., Tainer, J. A., Jackson, M. R., and Peterson, P. A. (1995) *J. Biol. Chem.* **270**, 19989–19997
60. Pelham, H. R. B. (1990) *Trends Biochem. Sci.* **15**, 483–486
61. Wilson, D. W., Lewis, M. J., and Pelham, H. R. B. (1993) *J. Biol. Chem.* **268**, 7465–7468
62. Mallabiabarrena, A., Jiménez, M. A., Rico, M., and Alarcón, B. (1995) *EMBO J.* **14**, 2257–2268
63. Luzio, J. P., and Banting, G. (1993) *Trends Biochem. Sci.* **18**, 395–398
64. Peter, F., Van, P. N., and Soling, H.-D. (1992) *J. Biol. Chem.* **267**, 10631–10637
65. Beh, C. T., and Rose, M. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9820–9823
66. Hoe, M. H., Slusarewicz, P., Misteli, T., Watson, R., and Warren, G. (1995) *J. Biol. Chem.* **270**, 25057–25063
67. Schneider, H.-C., Berthold, J., Bauer, M. F., Dietmeier, K., Guiard, B., Brunner, M., and Neupert, W. (1994) *Nature* **371**, 768–773
68. Bhattacharyya, T., Karnezis, A. N., Murphy, S. P., Hoang, T., Freeman, B. C., Phillips, B., and Morimoto, R. I. (1995) *J. Biol. Chem.* **270**, 1705–1710
69. Schmitt, M., Neupert, W., and Langer, T. (1995) *EMBO J.* **14**, 3434–3444
70. Stuart, R. A., Cyr, D. M., Craig, E. A., and Neupert, W. (1994) *Trends Biochem. Sci.* **19**, 87–92
71. Lithgow, T., Glick, B. S., and Schatz, G. (1995) *Trends Biochem. Sci.* **20**, 98–101
72. Pfanner, N., Craig, E. A., and Meijer, M. (1994) *Trends Biochem. Sci.* **19**, 368–372
73. Mayer, A., Nargang, F. E., Neupert, W., and Lill, R. (1995) *EMBO J.* **14**, 4204–4211
73a. Heard, T. S., and Weiner, H. (1998) *J. Biol. Chem.* **273**, 29389–29393
74. Görlich, D., Vogel, F., Mills, A. D., Hartmann, E., and Laskey, R. A. (1995) *Nature* **377**, 246–248
74a. Koehler, C. M., Merchant, S., and Schatz, G. (1999) *Trends Biochem. Sci.* **24**, 428–432
75. Smeekens, S., Weisbeek, P., and Robinson, C. (1990) *Trends Biochem. Sci.* **15**, 73–76
76. Pilon, M., Wienk, H., Sips, W., de Swaaf, M., Talboom, I., van 't Hof, R., de Korte-Kool, G., Demel, R., Weisbeek, P., and de Kruijff, B. (1995) *J. Biol. Chem.* **270**, 3882–3893
77. Schnell, D. J., Kessler, F., and Blobel, G. (1994) *Science* **266**, 1007–1011
78. Voelker, R., and Barkan, A. (1995) *EMBO J.* **14**, 3905–3914
79. Viitanen, P. V., Schmidt, M., Buchner, J., Suzuki, T., Vierling, E., Dickson, R., Lorimer, G. H., Gatenby, A., and Soll, J. (1995) *J. Biol. Chem.* **270**, 18158–18164
79a. Nakai, K., and Horton, P. (1999) *Trends Biochem. Sci.* **24**, 34–35
79b. Klionsky, D. J. (1998) *J. Biol. Chem.* **273**, 10807–10810
80. Rothman, J. E. (1994) *Nature* **372**, 55–63
81. Pryer, N. K., Wuestehube, L. J., and Schekman, R. (1992) *Ann. Rev. Biochem.* **61**, 471–516
82. Bennett, M. K., and Scheller, R. H. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2559–2563
83. Duden, R., Hosobuchi, M., Hamamoto, S., Winey, M., Byers, B., and Schekman, R. (1994) *J. Biol. Chem.* **269**, 24486–24495
84. Edelmann, L., Hanson, P. I., Chapman, E. R., and Jahn, R. (1995) *EMBO J.* **14**, 224–231
84a. Dulubova, I., Sugita, S., Hill, S., Hosaka, M., Fernandez, I., Südhof, T. C., and Rizo, J. (1999) *EMBO J.* **18**, 4372–4382
84b. Katz, L., Hanson, P. I., Heuser, J. E., and Brennwald, P. (1998) *EMBO J.* **17**, 6200–6209
84c. Gerona, R. R. L., Larsen, E. C., Kowalchyk, J. A., and Martin, T. F. J. (2000) *J. Biol. Chem.* **275**, 6328–6336
85. Morgan, A., and Burgoyne, R. D. (1995) *EMBO J.* **14**, 232–239
85a. Glockshuber, R. (1999) *Nature (London)* **401**, 30–31
85b. Bolhuis, A., Venema, G., Quax, W. J., Bron, S., and van Dijl, J. M. (1999) *J. Biol. Chem.* **274**, 24531–24538
85c. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 905–920
85d. Berardi, M. J., and Bushweller, J. H. (1999) *J. Mol. Biol.* **292**, 151–161
85e. van den Berg, B., Chung, E. W., Robinson, C. V., Mateo, P. L., and Dobson, C. M. (1999) *EMBO J.* **18**, 4794–4803
86. Freedman, R. B., Hirst, T. R., and Tuite, M. F. (1994) *Trends Biochem. Sci.* **19**, 331–335
87. Martin, J. L., Bardwell, J. C. A., and Kuriyan, J. (1993) *Nature* **365**, 465–468
88. Kishigami, S., Kanaya, E., Kikuchi, M., and Ito, K. (1995) *J. Biol. Chem.* **270**, 17072–17074
89. Jander, G., Martin, N. L., and Beckwith, J. (1994) *EMBO J.* **13**, 5121–5127
90. Kanaya, E., Anaguchi, H., and Kikuchi, M. (1994) *J. Biol. Chem.* **269**, 4273–4278
91. Frech, C., and Schmid, F. X. (1995) *J. Biol. Chem.* **270**, 5367–5374
92. Chivers, P. T., Laboissiére, M. C. A., and Raines, R. T. (1996) *EMBO J.* **15**, 2659–2667
93. Tschantz, W. R., Paetzel, M., Cao, G., Suciu, D., Inouye, M., and Dalbey, R. E. (1995) *Biochemistry* **34**, 3935–3941
94. Steiner, D. F., Smeekens, S. P., Ohagi, S., and Chan, S. J. (1992) *J. Biol. Chem.* **267**, 23435–23438
95. De Bie, I., Savaria, D., Roebroek, A. J. M., Day, R., Lazure, C., Van de Ven, W. J. M., and Seidah, N. G. (1995) *J. Biol. Chem.* **270**, 1020–1028
96. Fisher, J. M., and Scheller, R. H. (1988) *J. Biol. Chem.* **263**, 16515–16518
97. Rehfeld, J. F., Hansen, C. P., and Johnsen, A. H. (1995) *EMBO J.* **14**, 389–396
98. Ypma-Wong, M. F., Filman, D. J., Hogle, J. M., and Semler, B. L. (1988) *J. Biol. Chem.* **263**, 17846–17856
99. Bateman, A., Solomon, S., and Bennett, H. P. J. (1990) *J. Biol. Chem.* **265**, 22130–22136
100. Richter, D. (1983) *Trends Biochem. Sci.* **8**, 278–280

References

101. Douglass, J., Civelli, O., and Herbert, E. (1984) *Ann. Rev. Biochem.* **53**, 665 – 715
102. Arfin, S. M., and Bradshaw, R. A. (1988) *Biochemistry* **27**, 7979 – 7984
103. Tolan, D. R., Amsden, A. B., Putney, S. D., Urdea, M. S., and Penhoet, E. E. (1984) *J. Biol. Chem.* **259**, 1127 – 1131
104. Rubenstein, P. A., and Martin, D. J. (1983) *J. Biol. Chem.* **258**, 3961 – 3966
105. Sheff, D. R., and Rubenstein, P. A. (1992) *J. Biol. Chem.* **267**, 20217 – 20224
106. Kulkarni, M. S., and Sherman, F. (1994) *J. Biol. Chem.* **269**, 13141 – 13147
107. Gan, K., Gupta, S. D., Sankaran, K., Schmidt, M. B., and Wu, H. C. (1993) *J. Biol. Chem.* **268**, 16544 – 16550
108. Vogt, A., Sun, J., Qian, Y., Tan-Chiu, E., Hamilton, A. D., and Sebti, S. M. (1995) *Biochemistry* **34**, 12398 – 12403
109. Parish, C. A., and Rando, R. R. (1994) *Biochemistry* **33**, 9986 – 9991
110. Pompliano, D. L., Rands, E., Schaber, M. D., Mosser, S. D., Anthony, N. J., and Gibbs, J. B. (1992) *Biochemistry* **31**, 3800 – 3807
111. Busby, W. H., Jr., Quackenbush, G. E., Humm, J., Youngblood, W. W., and Kizer, J. S. (1987) *J. Biol. Chem.* **262**, 8532 – 8536
112. Fischer, W. H., and Spiess, J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3628 – 3632
113. Katopodis, A. G., Ping, D., Smith, C. E., and May, S. W. (1991) *Biochemistry* **30**, 6189 – 6194
114. Merkler, D. J., Kulathila, R., Consalvo, A. P., Young, S. D., and Ash, D. E. (1992) *Biochemistry* **31**, 7282 – 7288
115. Ping, D., Katopodis, A. G., and May, S. W. (1992) *J. Am. Chem. Soc.* **114**, 3998 – 4000
116. Eipper, B. A., Milgram, S. L., Husten, E. J., Yun, H.-Y., and Mains, R. E. (1993) *Protein Sci.* **2**, 489 – 497
117. Bradbury, A. F., and Smyth, D. G. (1991) *Trends Biochem. Sci.* **16**, 112 – 115
118. Low, M. G., and Saltiel, A. R. (1988) *Science* **239**, 268 – 275
119. Lisanti, M. P., and Rodriguez-Boulan, E. (1990) *Trends Biochem. Sci.* **15**, 113 – 118
120. Graves, D. J., Martin, B. L., and Wang, J. H. (1994) *Co-and post-translational modification of proteins*, Oxford Univ. Press, New York
121. Tuboi, S., Taniguchi, N., and Katunuma, N., eds. (1992) *Post-Translation Modification of Proteins*, CRC Press, Boca Raton, Florida
122. Barrett, G. C., ed. (1985) *Chemistry and Biochemistry of the Amino Acids*, Chapman and Hall, London; New York
123. Niehrs, C., Kraft, M., Lee, R. W. H., and Huttner, W. B. (1990) *J. Biol. Chem.* **265**, 8525 – 8532
124. Rucker, R. B., and Wold, F. (1988) *FASEB J.* **2**, 2252 – 2261
125. Klotz, A. V., Leary, J. A., and Glazer, A. N. (1986) *J. Biol. Chem.* **261**, 15891 – 15894
126. Oliver, C. N., Ahn, B., Moerman, E. J., Goldstein, S., and Stadtman, E. R. (1987) *J. Biol. Chem.* **262**, 5488 – 5491
127. Cooper, A. A., and Stevens, T. H. (1995) *Trends Biochem. Sci.* **20**, 351 – 356
128. Rechsteiner, M., Rogers, S., and Rote, K. (1987) *Trends Biochem. Sci.* **12**, 390 – 394
129. Stadtman, E. R. (1986) *Trends Biochem. Sci.* **11**, 11 – 12
130. Rivett, A. J. (1985) *J. Biol. Chem.* **260**, 300 – 305
131. Loshon, C. A., Swerdlow, B. M., and Setlow, P. (1982) *J. Biol. Chem.* **257**, 10838 – 10845
132. Olden, K., Parent, J. B., and White, S. C. (1982) *Biochim. Biophys. Acta.* **650**, 209 – 232
133. Chiang, H.-L., and Dice, J. F. (1988) *J. Biol. Chem.* **263**, 6797 – 6805
133a. Dell'Angelica, E. C., Mullins, C., Caplan, S., and Bonifacino, J. S. (2000) *FASEB J.* **14**, 1265 – 1278
134. Hershko, A., and Ciechanover, A. (1992) *Ann. Rev. Biochem.* **61**, 761 – 807
135. Dice, J. F. (1987) *FASEB J.* **1**, 349 – 357
136. Stadtman, E. R. (1990) *Biochemistry* **29**, 6323 – 6331
137. Rogers, S., Wells, R., and Rechsteiner, M. (1986) *Science* **234**, 364 – 368
138. Bachmair, A., Finley, D., and Varshavsky, A. (1986) *Science* **234**, 179 – 186
139. Gonda, D. K., Bachmair, A., Wünning, I., Tobias, J. W., Lane, W. S., and Varshavsky, A. (1989) *J. Biol. Chem.* **264**, 16700 – 16712
140. Madura, K., and Varshavsky, A. (1994) *Science* **265**, 1454 – 1458
140a. Davydov, I. V., and Varchavsky, A. (2000) *J. Biol. Chem.* **275**, 22931 – 22941
141. Herschlag, D. (1995) *J. Biol. Chem.* **270**, 20871 – 20874

Study Questions

1. Outline in detail, using structural formulas, the enzyme-catalyzed reactions by which cells in the human body convert glyceraldehyde 3-phosphate into pyruvate.

2. Describe, using chemical structural formulas, the reactions involved in the breakdown of glycogen to glucose 1-phosphate and the synthesis of glycogen from glucose 1-phosphate.

3. Describe the reaction steps in gluconeogenesis by which pyruvate is converted into glyceraldehyde 3-phosphate.

4. Compare the reactions of pyruvate that give rise to the following three compounds. List coenzymes or electron-carriers involved in each case and indicate any intermediate compounds.
   a) Ethanol
   b) Lactic acid
   c) Acetyl-Coenzyme A

5. Mammalian sperm cells metabolize D-fructose preferentially as a source of energy. Fructose is formed in cells of the seminal vesicle from D-glucose via reduction to the sugar alcohol sorbitol using NADPH, followed by oxidation of sorbitol to fructose using NAD^+. The fructose concentration in human semen is about 12 mM, whereas the glucose concentration within cells is usually less than 1 mM. If the ratio $[NADPH] / [NADP^+]$ is 10^4 times higher than the ratio $[NADH] / [NAD^+]$, what is the minimum glucose concentration in cells that could allow formation of 12 mM fructose? The standard Gibbs energies of formation from the elements $G^o{}_f$ in kJ / mol at 25°C are: D-glucose –917, D-fructose –915.

6. Outline the reactions by which glyceraldehyde 3-phosphate is converted to 3-phosphoglycerate with coupled synthesis of ATP in the glycolysis pathway. Show important mechanistic details.

7. Why can't acetyl CoA be converted to glucose in animals?

8. Describe the parallel reaction sequences between the citric acid cycle and the β oxidation pathway.

9. Contrary to legend, camels do not store water in their humps, which actually consist of a large fat deposit. How can these fat deposits serve as a source of water? Calculate the amount of water (liters) that can be produced by the camel from 500 g of fat. Assume for simplicity that the fat consists entirely of tripalmitin.

10. A little-known microorganism carries out the following net reaction in a series of enzymatic steps:

$$\text{heptanoyl CoA} + \text{CoASH} + \text{NAD}^+ + \text{FAD} + H_2O \rightarrow \text{pentanoyl CoA} + \text{acetyl CoA} + \text{NADH} + \text{FADH}_2 + \text{H}^+$$

The net result is exactly the same as that of the β oxidation pathway, but for this microorganism, the pathway is demonstrably different in that pentanal, $CH_3(CH_2)_3CHO$, is an intermediate in the sequence. Which enzyme(s) of the β oxidation pathway does this microorganism lack? Propose an enzymatic reaction to account for the formation of this intermediate and a series for its conversion to the acyl CoA. ATP is not required.

11. Most natural fatty acids are even-numbered. What is the product of the *final* thiolase reaction to form an odd-chain fatty acid? Give a brief explanation.

12. It has been calculated that an average man takes in 1.5 kg of solid food and 1.4 kg of water per day. He gives off 3.5 kg of waste and 0.75 kg of sweat. It would thus appear that he should be losing over 1 kg per day through normal activites. How do you account for the fact that his weight remains relatively constant?

13. Suppose that fatty acids, instead of being broken down in two carbon fragments, were metabolized in three carbon units. What product(s) would Dr. Knoop have observed in the urine of his experimental dogs?

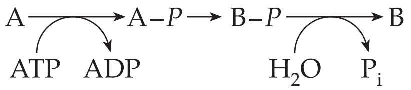
14. A renowned pharmacologist has announced the discovery of a new drug that specifically inhibits the fatty acid oxidation pathway within minutes of ingestion. The effects last for several hours only. The drug has no other effects on the subjects. The pharmacologist argues that this will increase athletic performance by shifting oxidative metabolism entirely to the more rapidly mobilized carbohydrate degradation pathway. Assume that the drug does work exactly as he suggests. Explain in a sentence or two how the drug would affect the performance of
a) a sprinter in the 100 meter dash
b) a long-distance runner in a marathon

15. Some bacteria use a "dicarboxylic acid cycle" to oxidize glyoxylate $OHC{-}COO^-$ to CO_2. The regenerating substrate for this cycle is acetyl-CoA. It is synthesized from glyoxylate by a complex pathway that begins with conversion of two molecules of glyoxylate to tartronic semialdehyde: $^-OOC{-}CHOH{-}CHO$. The latter is then dehydrogenated to D-glycerate. Write out a detailed scheme for the dicarboxylate cycle. Also indicate how glucose and other cell constituents can be formed from intermediates created in this biosynthetic pathway.

16. Write a balanced fermentation sequence by which glycogen can be converted rapidly to *sn*-glycerol 3-phosphate and pyruvate in insect flight muscle. How many molecules of ATP per glucose unit of glycogen will be formed?

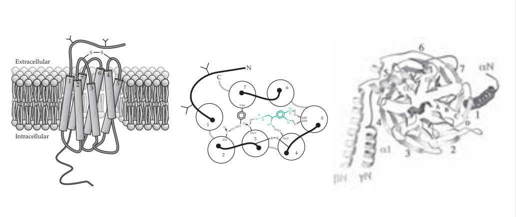
17. Show which parts (if any) of the citric acid cycle are utilized in each of the following reactions and what if any additional enzymes are needed in each case.
a) Oxidation of acetyl-CoA to CO_2
b) Catabolism of glutamate to CO_2
c) Biosynthesis of glutamate from pyruvate
d) Formation of propionate from pyruvate

18. a) If the Gibbs energy change $\Delta G'$ (pH 7) for the reaction $A \rightarrow B$ is +25 kJ/mol at 25°C, what would the ratio of [B]/[A] be at equilibrium?
b) Suppose that the reaction were coupled to the cleavage of ATP as follows:



Suppose further that the group transfer potential $(-\Delta G')$ for the phospho group of A–*P* at 25°C, pH 7 is 12 kJ/mol and that the equilibrium constant for conversion of A–*P* to B–*P* is the same as that for conversion of A to B. Calculate the concentrations of A, B, A–*P* and B–*P* at equilibrium if the phosphorylation state ratio R_p is $10^4\ M^{-1}$.

19. This problem refers to the 13 meter climb described in Chapter 6, Study Question 14.
a) Assuming that muscle accounts for 30% of total body mass, estimate the amounts (mmol/kg, total mmol, and total grams) of each of the following in their coenzyme forms: nicotinamide, pantothenic acid, thiamin. You may be able to obtain rough estimates from the vitamin content of pork.
b) Calculate how many times, on the average, each molecule of nicotinamide would undergo reduction and re-oxidation (i.e., turn over) during the climb. Do the same for pantothenic acid (cycling between acyl-CoA and free CoA forms) and for thiamin through its catalytic cycles.



Much of the control of cellular metabolism is accomplished by hormones and other molecules that bind to receptor proteins embedded in the plasma membrane. Left: Many receptors are seven-helix transmembrane proteins, the best known being the light receptors of the rhodopsin family (Chapter 23). Center: A related adrenergic receptor, viewed here from the extracellular side, has a molecule of the hormone adrenaline (green) bound in the center (Fig. 11-6). Right: These receptors interact with GTP-hydrolyzing "G proteins," which pass signals, often for short periods of time, to enzymes and other proteins. One structural domain of a G protein (Fig. 11-7C) shown here contains a β propeller domain which binds to the Ras-like domain of the α subunit. A few residues of the latter are visible in the center of this image.

Contents

535 .... **A. Pacemakers and the Control of Metabolic Flux**
538 .... **B. Genetic Control of Enzyme Synthesis**
538 .... 1. One Gene or Many?
538 .... 2. Repression, Induction, and Turnover
539 .... 3. Differences among Species
539 .... **C. Regulation of the Activity of Enzymes**
539 .... 1. Allosteric Control
540 .... *Regulatory subunits*
541 .... *Glycolysis and gluconeogenesis*
541 .... 2. Covalent Modification by Phosphorylation and Dephosphorylation
541 .... *Protein kinases and cyclic AMP*
544 .... *Protein phophatases*
545 .... *Phophorylation in bacteria*
545 .... 3. Other Modification Reactions Involving Group Transfer
549 .... 4. Thiol–disulfide Equilibria
549 .... 5. Regulatory Effects of H$^+$, Ca^{2+}, and Other Specific Ions
552 .... 6. Compartments and Organized Assemblies
553 .... **D. Hormones and Their Receptors**
553 .... 1. Beta Adrenergic Receptors and Related Seven-Helix Proteins
556 .... 2. Adenylate Cyclases (Adenylyl Cyclases)
557 .... 3. Guanine Nucleotide-Binding Proteins (G Proteins)
558 .... *Monomeric G proteins*
559 .... *Acylation and prenylation*
559 .... *Three-dimensional structures*
561 .... 4. Guanylate Cyclase (Guanylyl Cyclase), Nitric Oxide, and the Sensing of Light
561 .... 5. Bacterial Chemoreceptors
563 .... **E. Calcium, Inositol Polyphosphates, and Diacylglycerols**
563 .... 1. Alpha Adrenergic Receptors
563 .... 2. Phosphatidylinositol and the Release of Calcium Ions
566 .... **F. Regulatory Cascades**
566 .... 1. Advantages of Regulatory Cascades
567 .... 2. Substrate Cycles
567 .... **G. Insulin and Related Growth-Regulating Hormones**
568 .... 1. Metabolic Effects of Insulin
568 .... 2. Insulin Receptors
569 .... 3. A Second Messenger for Insulin?
571 .... **H. Growth Factors, Oncogenes, and the Cell Cycle**
571 .... 1. Oncogenes and Proto-oncogenes
576 .... *The ras oncogenes*
576 .... *Transcription factors*
578 .... 2. The MAP Kinase Cascade
580 .... 3. The Cell Cycle and Control of Growth

581 .... **References**
587 .... **Study Questions**

Boxes

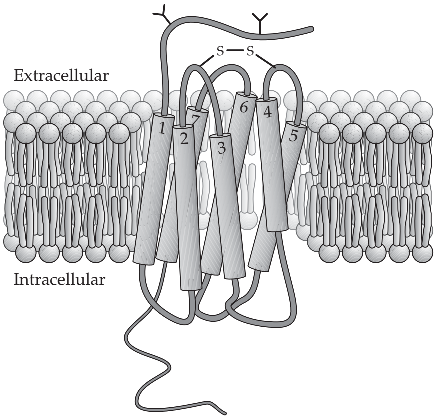
546 .... Box 11-A Cholera Toxin and Other Dangerous Proteins
550 .... Box 11-B Glutathione, Intracellular Oxidation–Reduction Buffer
557 .... Box 11-C The Attraction of *Dictyostelium* to Cyclic AMP
573 .... Box 11-D Cancer

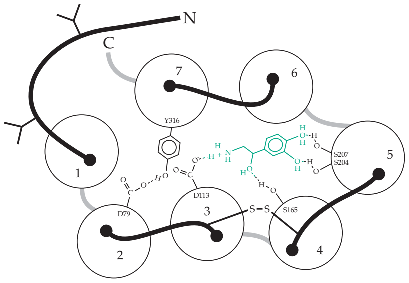
Tables

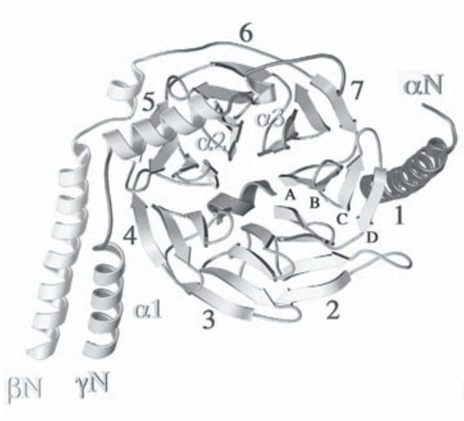
543 .... Table 11-1 A Few Covalent Modification Reactions Utilized to Control Metabolism
554 .... Table 11-2 Some Molecules (Second Messengers) That Carry Intracellular Signals
572 .... Table 11-3 A Few Oncogenes That Have Interested Biochemists
577 .... Table 11-4 A Few Abbreviations Used in Discussions of Cell Signaling

The Regulation of Enzymatic Activity and Metabolism

11







Living cells must operate with controls that provide a stable environment and a relatively constant supply of materials needed for biosynthesis and for meeting the energy needs of cells. They must also be responsive to changes in their environment and must be able to undergo mitosis and reproduce when appropriate. The necessary control of metabolism and of growth is accomplished largely through mechanisms that regulate the locations, the amounts, and the catalytic activities of enzymes. The purpose of this chapter is to summarize these control mechanisms and to introduce terminology and shorthand notations that will be used throughout this book. Many of the **control elements** considered are summarized in Fig. 11-1.

A. Pacemakers and the Control of Metabolic Flux

Metabolic control can be understood to some extent by focusing attention on those enzymes that catalyze rate-limiting steps in a reaction sequence. Such **pacemaker enzymes**[1–4] are often involved in reactions that determine the overall respiration rate of a cell, reactions that initiate major metabolic sequences, or reactions that initiate branch pathways in metabolism. Often the first step in a unique biosynthetic pathway for a compound acts as the pacemaker reaction. Such a reaction may be described as the **committed step** of the pathway. It usually proceeds with a large decrease in Gibbs energy and tends to be tightly controlled. Both the rate of synthesis of the enzyme protein and the activity of the enzyme, once it is formed, may be inhibited by **feedback inhibition** which occurs when an end product of a biosynthetic pathway accumulates and inhibits the enzyme. Enzyme activity may also be turned on or off by the effect of a hormone, by some other external stimulus, or by internal mechanisms that sense the metabolic state of the cell. Enzymes, other than the pacemaker, that catalyze reactions in a pathway may not be regulated and may operate at a steady state close to equilibrium.

If conditions within a cell change, a pacemaker reaction may cease to be rate limiting. A reactant plentiful in one circumstance may, in another, be depleted to the point that the rate of its formation from a preceding reaction determines the overall rate. Thus, as we have seen in Chapter 10, metabolism of glucose in our bodies occurs through the rapidly interconvertible phosphate esters glucose 6-phosphate and fructose 6-phosphate. The pacemaker enzyme in utilization of glucose or of glycogen is often **phosphofructokinase** (Fig. 11-2, step *b*) which catalyzes further metabolism of fructose 6-phosphate. However, if metabolism by this route is sufficiently rapid, the rate of formation of glucose 6-phosphate from glucose catalyzed by **hexokinase (**Fig. 11-2, step *a*) may become rate limiting.

Some catabolic reactions depend upon ADP, but under most conditions its concentration is very low because it is nearly all phosphorylated to ATP. Reactions utilizing ADP may then become the rate-limiting pacemakers in reaction sequences. Depletion of a reactant sometimes has the effect of changing the whole pattern of metabolism. Thus, if oxygen is unavailable to a yeast, the reduced coenzyme NADH accumulates and reduces pyruvate to ethanol plus CO_2 (Fig. 10-3). The result is a shift from oxidative metabolism to fermentation.

Pacemaker enzymes are often identified by the fact that the measured **mass action ratio**, e.g., for the reaction $A + B \rightarrow P + Q$,the ratio [P][Q]/[A][B], is far

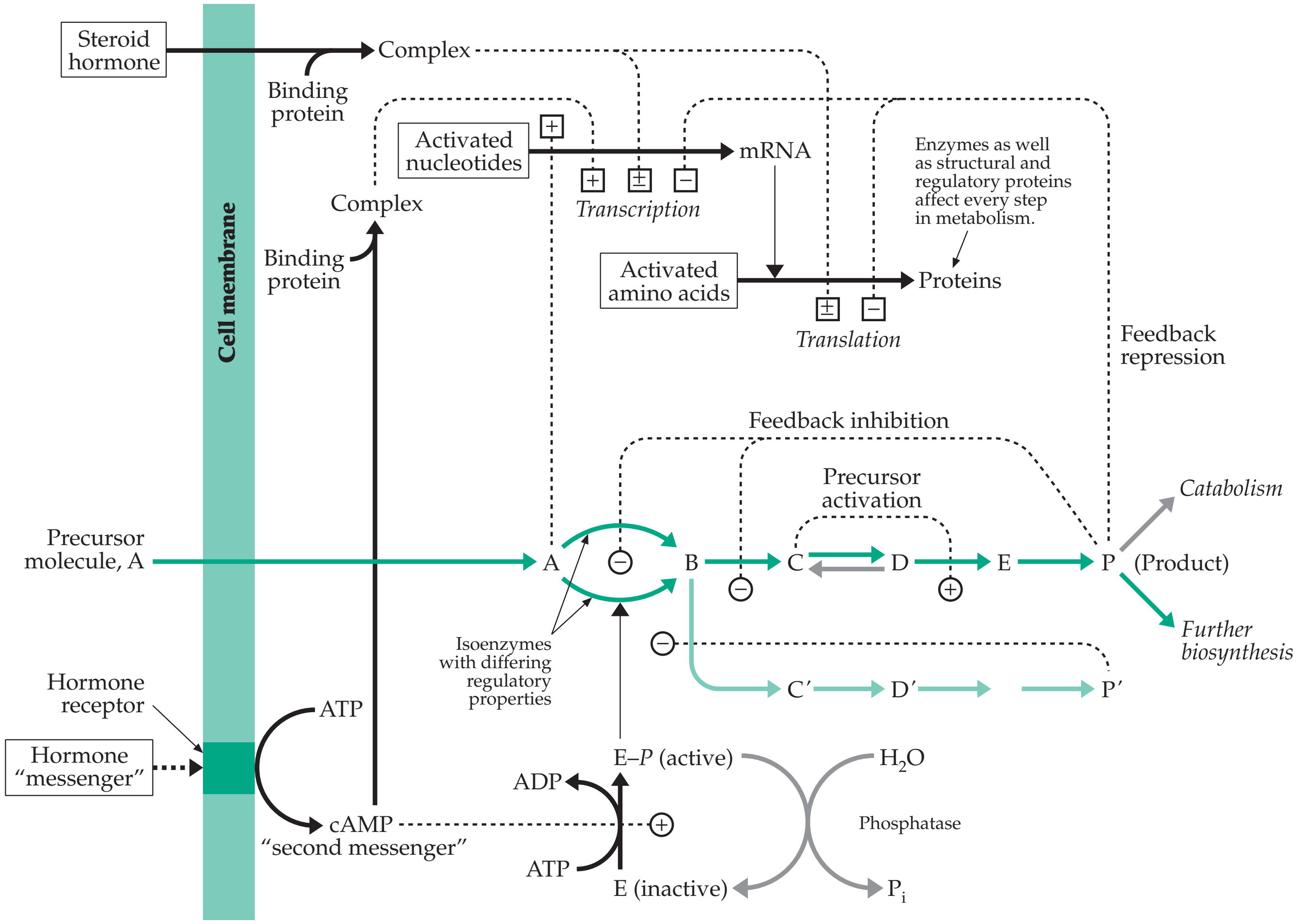

Steroid hormone
Complex
Binding protein
Cell membrane
Activated nucleotides
+
+
±
−
mRNA
Transcription
Enzymes as well as structural and regulatory proteins affect every step in metabolism.
Complex
Binding protein
Activated amino acids
Proteins
±
−
Translation
Feedback repression
Feedback inhibition
Precursor activation
Catabolism
Precursor molecule, A
A
−
B
−
C
D
+
E
P
(Product)
Further biosynthesis
Isoenzymes with differing regulatory properties
−
C′
D′
P′
Hormone receptor
Hormone "messenger"
ATP
cAMP "second messenger"
E–P (active)
ADP
ATP
+
E (inactive)
H2O
Phosphatase
Pi

Figure 11-1 Some control elements for metabolic reactions. *Throughout the book modulation of the activity of an enzyme by allosteric effectors or of transcription and translation of genes is indicated by dotted lines from the appropriate metabolite.* The lines terminate in a minus sign for inhibition or repression and in a plus sign for activation or derepression. Circles indicate direct effects on enzymes, while boxes indicate repression or induction of enzyme synthesis.

from that predicted by the known equilibrium constant for the reaction. Another approach to identifying a pacemaker enzyme is to inhibit it and observe resulting changes in steady-state metabolite concentrations.[5] If the pacemaker enzyme is inhibited, its substrate and other compounds preceding the step catalyzed by this enzyme will accumulate. At the same time the concentrations of products of the pacemaker reaction will drop as a result of their relatively more rapid removal by enzymes catalyzing subsequent steps. However, conclusions drawn by this approach may sometimes be erroneous.[6]

In spite of its usefulness, the pacemaker concept is oversimplified. It is often impossible to identify a specific pacemaker enzyme. When both catabolism and biosynthesis occur (e.g., as in the scheme in Fig. 11-2) it may be more useful to model the entire system with a computer than to try to identify pacemakers.[7,8] It is also important to realize that reaction rates may be determined by the rate of diffusion of a compound through a membrane. Thus, membrane transport processes can serve as pacemakers.

A general approach to analysis of complex metabolic pathways or to cell growth was introduced by Savageau,[9–11] and similar approaches have been followed by others.[3,12–19] They emphasize the **flux** of material through a pathway under steady-state conditions and recognize that every enzyme in a sequence can have some effect on the overall rate. Consider a chain of enzymes E_1 to E_n acting on substrate S_0, which is converted via substrates S_1, S_2 ... S_{n-1} into product S_n:

$$S_0 \xrightarrow{E_1} S_1 \xrightarrow{E_2} S_2 \xrightarrow{E_3} S_3 \cdots \xrightarrow{E_{n-1}} S_{n-1} \xrightarrow{E_n} S_n \qquad (11\text{-}1)$$

The flux $F = d[S_n]/dt$ is a constant. If the concentration of E_i were to change by an infinitesimal amount a corresponding change in F might be observed. The

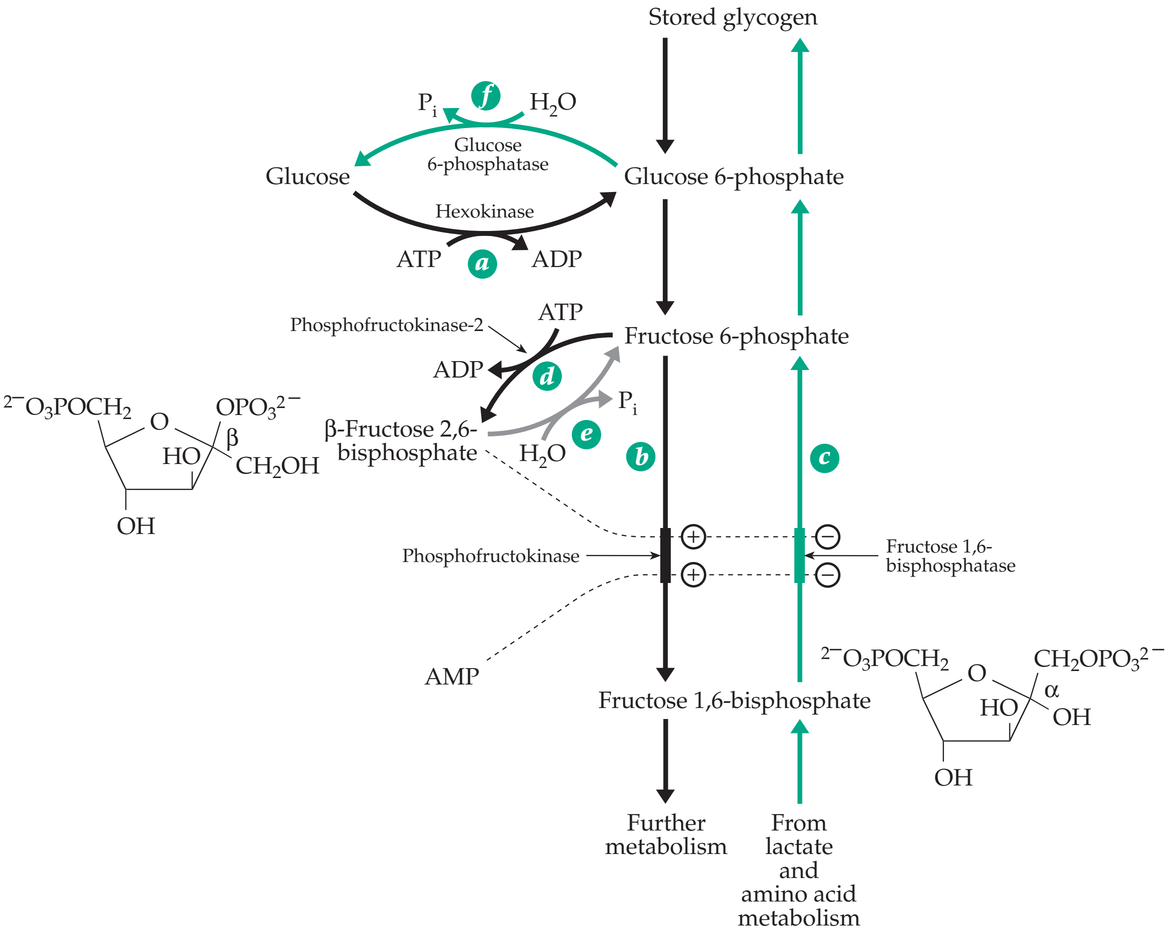

Stored glycogen
Glucose
Glucose 6-phosphatase
Hexokinase
Glucose 6-phosphate
Pi
H2O
ATP
ADP
a
f
Phosphofructokinase-2
ATP
ADP
Fructose 6-phosphate
d
e
Pi
H2O
β-Fructose 2,6-bisphosphate
b
c
Phosphofructokinase
Fructose 1,6-bisphosphatase
AMP
Fructose 1,6-bisphosphate
Further metabolism
From lactate and amino acid metabolism
OH
HO
O
β
CH2OH
OPO3 2−
2−O3POCH2
α
CH2OPO3 2−
OH

Figure 11-2 Roles of phosphofructose kinase and fructose 1,6-bisphosphatase in the control of the breakdown (➞) and storage (➞) of glycogen in muscle. The uptake of glucose from blood and its release from tissues is also illustrated. The allosteric effector fructose 2,6-bisphosphate (Fru-2,6-P_2) regulates both phosphofructokinase and fructose 2,6-bisphosphatase. These enzymes are also regulated by AMP if it accumulates. The activity of phosphofructokinase-2 (which synthesizes Fru-2,6-P_2) is controlled by a cyclic AMP-dependent kinase and by dephosphorylation by a phosphatase.

ratio of the change in flux to the change in [E_i] is the sensitivity coefficient Z_i:

$$Z_i = \frac{dF/F}{d[E_i]/[E_i]} = \frac{d \ln F}{d \ln[E_i]} \tag{11-2}$$

The sensitivity coefficient can range from 1 for a pacemaker enzyme to almost zero for a very active enzyme that does not significantly limit the flux. The sum of Z_1 to Z_n is equal to 1:

$$\sum_{n}^{i=1} Z_i = 1 \tag{11-3}$$

Not only can the concentration [E_i] change but also allosteric effectors can alter the activity. Kacser and Burns defined this in terms of a **controllability coefficient κ**, the ratio of the logarithmic change in velocity to the change in the concentration of an effector:

$$\kappa = d \ln v / d \ln [P] \tag{11-4}$$

It follows that the logarithmic change in flux $d \ln F$ can be related to the change in effector concentration [P] in the following way:

$$d \ln F = Z_i \kappa d \ln [P] \tag{11-5}$$

From this equation we see that for the flux to be sensitively dependent upon [P], both Z_i and κ must be reasonably large. This is exactly the case for many of the enzymes that have been identified as pacemakers. This approach can be applied to many aspects of metabolic control including cell growth.[20]

B. Genetic Control of Enzyme Synthesis

All cellular regulatory mechanisms depend upon the synthesis of proteins, that is, upon the expression of genetic information. Within a given cell many genes are transcribed continuously but others may remain unexpressed. The rates of both transcription and degradation of mRNA are regulated, as are the rates of synthesis of enzymes on the ribosomes in the cytoplasm and the rates of protein turnover. Although a single copy of a gene in each member of a pair of chromosomes is often adequate, there are situations in which extra copies of part of the DNA of a cell are formed. Such **gene amplification**, which is dealt with further in Chapter 27, provides the possibility for very rapid synthesis of an enzyme or other protein. It often happens in highly specialized cells such as those of the silkworm's silk glands which must make enormous amounts of a small number of proteins.[21]

1. One Gene or Many?

Many enzymes exist within a cell as two or more **isoenzymes**, enzymes that catalyze the same chemical reaction and have similar substrate specificities. They are not isomers but are distinctly different proteins which are usually encoded by different genes.[22,23] An example is provided by aspartate aminotransferase (Fig. 2-6) which occurs in eukaryotes as a pair of cytosolic and mitochondrial isoenzymes with different amino acid sequences and different isoelectric points. Although these isoenzymes share less than 50% sequence identity, their internal structures are nearly identical.[24–27] The two isoenzymes, which also share structural homology with that of *E. coli*,[28] may have evolved separately in the cytosol and mitochondria, respectively, from an ancient common precursor. The differences between them are concentrated on the external surface and may be important to as yet unknown interactions with other protein molecules.

Isoenzymes are designated in various ways. They are often *numbered in the order of decreasing electrophoretic mobility* at pH 7 to 9: Most enzymes are negatively charged in this pH range and the one that migrates most rapidly toward the anode is numbered one. This is the same convention used in the electrophoresis of blood proteins (e.g. see Box 2-A).

Lactate dehydrogenase exists in the cytoplasm of humans and most animals as *five forms* which are easily separable by electrophoresis and are evenly spaced on electropherograms.[8] This enzyme is a tetramer made of two kinds of subunits. Isoenzyme 1, which has the highest electrophoretic mobility, consists of four identical type B subunits. The slowest moving tetramer (isoenzyme 5) consists of four type A subunits, while the other three forms, AB_3, A_2B_2, and A_3B, contain both subunits in different proportions. The two subunits are encoded by separate genes which are expressed to different extents in different tissues. Thus, heart muscle and liver produce mainly subunit B, while skeletal muscle produces principally subunit A. A third subunit type (C) is found only in the testes.[29]

Why do cells produce isoenzymes? One reason may be that enzymes with differing kinetic parameters are needed.[30] Substrate concentrations may vary greatly between different tissues; between different subcellular compartments; and at different developmental stages of an organism. While the need for various isoenzymes of lactate dehydrogenase is not well understood,[31] it is easier to understand the roles of the multiple forms of hexokinase, the enzyme that catalyzes the reaction of step *a* in Fig. 11-2. The brain enzyme has a high affinity for glucose (K_m = 0.05 mM). Thus, it is able to phosphorylate glucose and to make that substrate available to the brain for metabolism, even if the glucose concentration in the tissues falls to low values.[31a] On the other hand, **glucokinase**, the hexokinase isozyme found in liver, has a much higher K_m of ~10 mM. It functions to remove excess glucose from blood, whose normal glucose content is ~5.5 mM. Glucokinase reaches its maximal activity only when the glucose concentration becomes much higher.[32] This happens after a meal when the absorbed glucose passes through the portal circulation directly to the liver.

Another important source of variation in enzymes as well as in other proteins is **alternative splicing** of mRNA.[33] For example, transcription of the mouse α-amylase gene in the salivary gland starts at a different site (promoter) than does transcription in the liver. The two common isoforms of the insulin receptor (Fig. 11-11) arise because a 36-nucleotide (12-amino acid) exon is spliced out of the mRNA for the shorter protein. Isoenzymes of aldolase[34] and of many other proteins are formed in a similar manner. **Frame-shifting** during protein synthesis (Chapter 29) and also post-translational alterations may give rise to additional modified forms. They are often synthesized in relatively small amounts but may be essential to the life of the cell. In addition, genetic variants of almost any protein will be found in any population. These often differ in sequence by a single amino acid.

2. Repression, Induction, and Turnover

The synthesis of some enzymes is referred to as **constitutive**, implying that the enzyme is formed no matter what the environmental conditions of the cell. For example, many bacteria synthesize the enzymes required to catabolize glucose under all conditions of growth. Other enzymes, known as **inducible**, are often produced only in small amounts. However, if cells are grown in specific inducing conditions for these enzymes, they are synthesized in larger quantities. For example, when *E. coli* is cultivated in the presence of lactose, several of the enzymes required for the catabolism of that disaccharide are formed. Synthesis of these enzymes is normally **repressed**. The genes which code for them are kept turned off through the action of protein **repressors** which bind to specific sites on the DNA and block transcription of the genes that they control (Fig. 11-1). Repressors have allosteric properties; in one conformation they bind tightly to DNA but in another they do not bind. For example, the free tryptophan repressor binds to DNA only weakly, but if a high concentration of tryptophan develops within the cell the tryptophan binds to an allosteric site on the repressor protein (Fig. 5-35). This changes the conformation to one that binds tightly to the appropriate control sequence in the DNA. In the case of lactose catabolism the free repressor binds to control sequences in the DNA until the **inducer**, allolactose (Chapters 4 and 28), binds at an allosteric site. This decreases the affinity of the repressor for DNA and the controlled genes are **derepressed**. There are also many protein **transcription factors** that have a positive effect, binding to DNA and promoting transcription of specific genes.

Synthesis of many enzymes is repressed most of the time. The appearance of an enzyme at a particular stage in the life of an organism as well as the differing distributions of isoenzymes within differentiated tissue result from derepression. The control of enzyme synthesis may also be exerted during the splicing of transcripts and at the translational level as well. These control mechanisms are often relatively slow, with response times of hours or even days. However, effects on the synthesis of some hormones, such as insulin (Section G), may be observed within a few minutes.

Genetic factors influence the rate of not only synthesis of proteins but also their breakdown, i.e., the rate of turnover. As we have seen in Chapter 10, some enzymes are synthesized as inactive proenzymes which are later modified to active forms, and active enzymes are destroyed, both by accident and via deliberate hydrolytic pathways. Protein **antienzymes** may not only inhibit enzymes but may promote their breakdown.[35] An example is the antienzyme that controls ornithine decarboxylase, a key enzyme in the synthesis of the polyamines that are essential to growth.[36,37] As with all cell constituents, the synthesis of enzymes and other proteins is balanced by degradation.

3. Differences among Species

Catalytic mechanisms of enzymes have usually been conserved throughout evolution, and certain residues in an enzyme may be invariant among many species. However, there are usually many differences in the distribution of amino acid residues on the surface of the proteins. Since changes in the surface shape of a protein molecule may alter the sensitivity to a potential allosteric effector, very different regulatory properties are sometimes found between species.

C. Regulation of the Activity of Enzymes

Some regulation of metabolism is provided by the kinetic properties of the enzymes. Thus, the value (mol/l) of the Michaelis constant K_m for an enzyme is usually low if the substrate normally occurs in very low concentration. It is likely to be higher if the enzyme acts on an abundant substrate. A tightly bound product of an enzymatic reaction will be released slowly from an enzyme if the product concentration within the cell is too high. However, while some regulation of metabolism is provided in such simple ways, the rapid changes that result from stimulation by hormones or by nerve impulses depend upon additional specific regulatory mechanisms that are discussed in the following sections.

1. Allosteric Control

Probably the most common and widespread control mechanisms in cells are **allosteric inhibition** and **allosteric activation**. These mechanisms are incorporated into metabolic pathways in many ways, the most frequent being **feedback inhibition**. This occurs when an end product of a metabolic sequence accumulates and turns off one or more enzymes needed for its own formation. It is often *the first enzyme unique to the specific biosynthetic pathway for the product that is inhibited.* When a cell makes two or more isoenzymes, only one of them may be inhibited by a particular product. For example, in Fig. 11-1 product P inhibits just one of the two isoenzymes that catalyzes conversion of A to B; the other is controlled by an enzyme modification reaction. In bacteria such as *E. coli*, three isoenzymes, which are labeled I, II, and III in Fig. 11-3, convert aspartate to β-aspartyl phosphate, the precursor to the end products threonine, isoleucine, methionine, and lysine. Each product inhibits only one of the isoenzymes as shown in the figure.

Feedback can also be positive. Since AMP is a product of the hydrolysis of ATP, its accumulation is a signal to *activate* key enzymes in metabolic pathways that generate ATP. For example, AMP causes allosteric activation of glycogen phosphorylase, which catalyzes the first step in the catabolism of glycogen. As is shown in Fig. 11-5, the allosteric site for AMP or IMP binding is more than 3 nm away from the active site. Only a subtle conformational change accompanies binding of IMP.[38] However, ^{31}P NMR studies indicate that binding of the activator may induce a change in the state of protonation of the phosphate group of the coenzyme pyridoxal phosphate at the active site as explained in Chapter 12.

A second pattern of allosteric control may be referred to as **precursor activation** or feed**-**forward control. A metabolite acting as an allosteric effector turns on an enzyme that either acts directly on that metabolite or acts on a product that lies further ahead in the sequence. For example, in Fig. 11-1, metabolite C activates the enzyme that catalyzes an essentially irreversible reaction of compound D. An actual example is provided by glycogen synthase, whose inactive "dependent" or D form is activated allosterically by the glycogen precursor glucose 6-phosphate.[39] See also phosphorylase kinase (Section 2).

Regulatory subunits. Some enzymes consist not only of catalytic subunits, which contain the active sites, but also of **regulatory subunits**. The latter bind to the catalytic subunits and serve as allosteric modifiers. The binding of inhibitors or activators to specific sites on the regulatory subunits induces conformational changes in these subunits, altering their interaction with the catalytic subunits. A well-known example is aspartate carbamoyltransferase (ACTase) from *E. coli* (Fig. 7-20).[40–42] Its regulatory subunits carry binding sites for cytidine triphosphate (CTP), which acts as a specific allosteric inhibitor of the enzyme. The significance of this fact is that ACTase catalyzes the first reaction specific to the pathway of synthesis of pyrimidine nucleotides (Chapter 25). CTP is an end product of that pathway and exerts feedback inhibition on the enzyme that initiates its synthesis. ATP also binds to the regulatory subunits and causes inhibition of the enzyme.

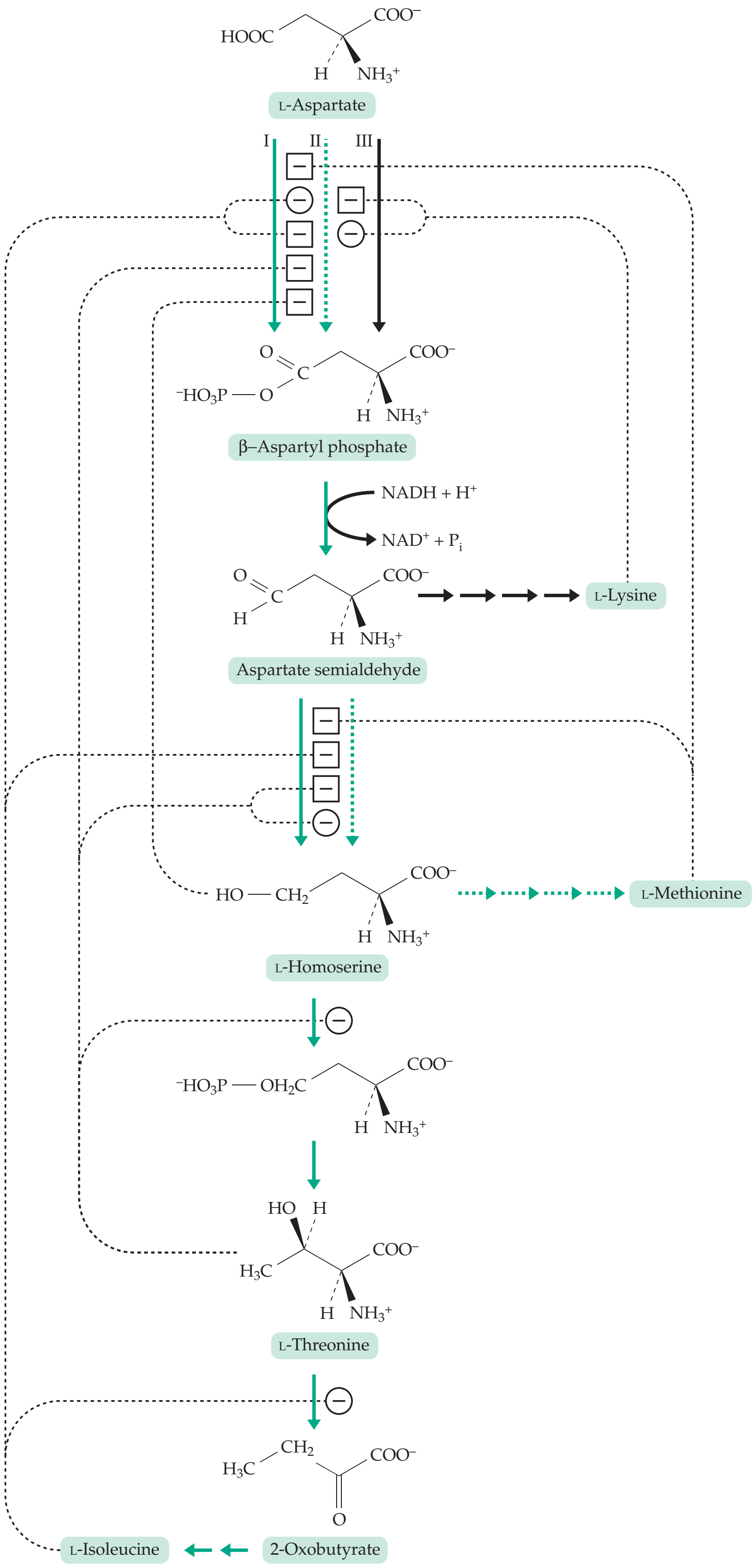
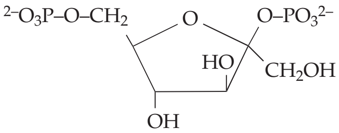


Figure 11-3 Feedback inhibition of enzymes involved in the biosynthesis of threonine, isoleucine, methionine, and lysine in *E. coli*. These amino acids all arise from L-aspartate, which is formed from oxaloacetate generated by the biosynthetic reactions of the citric acid cycle (Fig. 10-6). ⊖ Allosteric inhibition. ⊟ Repression of transcription of the enzyme or of its synthesis on ribosomes.

In the presence of CTP the binding of the substrates is cooperative, as would be anticipated if there is a two-state conformational equilibrium involving the catalytic subunits. This would be similar to the case depicted in Fig. 9-13 except that trimers rather than dimers are involved and the inhibitor is a part of the regulatory subunit and is controlled by binding of CTP.

Binding of ATP to ACTase decreases the cooperativity in substrate binding, again as predicted for a two-state model. However, as suggested by its structure, this enzyme system is more complex, as also indicated by the observed anticooperative binding of the activator CTP. X-ray studies show that binding of ligands causes a movement along subunit interfaces as well as localized conformational changes within the subunits. These are reminiscent of the changes seen upon binding of oxygen to hemoglobin (Fig. 7-25).

Glycolysis and gluconeogenesis. The highly regulated enzymes phosphofructokinase and fructose 1,6-bisphosphatase catalyze steps *b* and *c* of the reactions in Fig. 11-2, reactions that control glucose metabolism in cells. These enzymes have been studied for many years but the important allosteric effector **β-fructose 2,6-bisphosphate** was not discovered until 1980.[43–46]



β-Fructose
2,6-bisphosphate

This compound, which is formed from fructose 6-phosphate by a new enzyme, **phosphofructokinase-2** (also called fructose 6-phosphate 2-kinase) in step *d*, Fig. 11-2 activates phosphofructokinase allosterically. At the same time it inhibits fructose 1,6-bisphosphatase, an enzyme required for reversal of glycogen breakdown, that is, for the conversion of various metabolites arising from amino acids into glycogen (Fig. 11-2, step *c*).[46a] These same two key regulated enzymes are also affected by many other metabolites. For example, ATP in excess inhibits phosphofructokinase, decreasing the overall rate of glucose metabolism and consequently of ATP production. Citrate, which is exported from mitochondria when carbohydrate metabolism is excessive, inhibits the same enzyme. On the other hand, AMP acts together with fructose 2,6-bisphosphate to activate the pathway for glycogen breakdown and to inhibit that for its synthesis (Fig. 11-2). The concentration of the regulator fructose 2,6-bisphosphate is controlled by mechanisms that are discussed in the following section.

2. Covalent Modification by Phosphorylation and Dephosphorylation

Rapid alteration of the activities of enzymes is often accomplished by **reversible covalent modification**.[39,47] Many different modification reactions are known (Table 11-1) and there are doubtless many more to be discovered. Probably the most widespread and certainly the most studied is **phosphorylation**, the transfer of a phospho group from ATP or other suitable donor to a side chain group on the enzyme. An example is the phosphorylation by ATP of hydroxyl groups of specific serine residues in the two enzymes glycogen phosphorylase and glycogen synthase. These modifications are accomplished through a series or **cascade** of reactions initiated by the binding of hormones to cell surface receptors or by nerve impulses as is shown in Fig. 11-4. In the absence of such a stimulus, glycogen phosphorylase is present in its unphosphorylated or *b* form. Although this form can be activated allosterically by AMP, it is normally nearly inactive. When an appropriate hormone binds to the cell surface a cascade of reactions, as described in Section D, leads to activation of an enzyme called **phosphorylase kinase**.[47a] This enzyme transfers a phospho group from ATP to the –OH group of the side chain of Ser 14 in each subunit (Fig. 11-4, left center), converting the enzyme into the active glycogen-degrading **phosphorylase *a***. This switches the cellular metabolism from that designed to deposit the storage polysaccharide glycogen to one that degrades glycogen to provide the cell with energy. Serine 14 of glycogen phosphorylase is located adjacent to the allosteric AMP binding site (Fig. 11-5) and is surrounded by positively charged arginine and lysine side chains. Phosphorylation induces a rearrangement of hydrogen bonds involving these residues and in some way sends an appropriate signal to the active site. It also increases the affinity for AMP in the allosteric sites.[48,48a] Phosphorylase kinase is allosterically activated by AMP, a product of its action – a feed-forward activator.

The control of glycogen phosphorylase by the phosphorylation–dephosphorylation cycle was discovered in 1955 by Edmond Fischer and Edwin Krebs[50] and was at first regarded as peculiar to glycogen breakdown. However, it is now abundantly clear that similar reactions control most aspects of metabolism.[51] Phosphorylation of proteins is involved in control of carbohydrate, lipid, and amino acid metabolism; in control of muscular contraction, regulation of photosynthesis in plants,[52] transcription of genes,[51] protein syntheses,[53] and cell division; and in mediating most effects of hormones.

Protein kinases and cyclic AMP. Phosphorylase kinase is one of hundreds of different protein kinases which differ in specificity toward their substrates, in

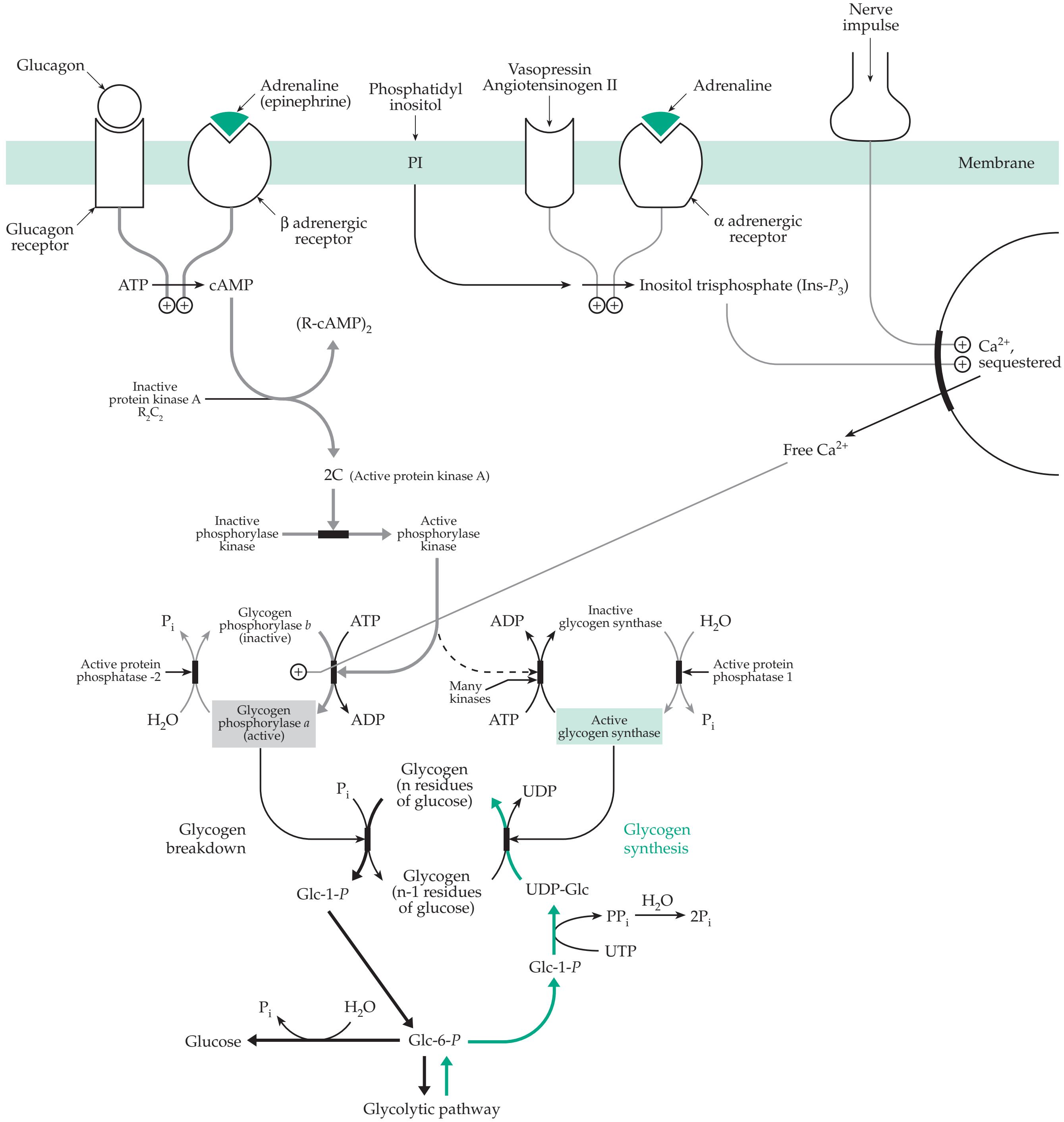

Nerve impulse
Glucagon
Adrenaline (epinephrine)
Phosphatidyl inositol
Vasopressin
Angiotensinogen II
Adrenaline
PI
Membrane
Glucagon receptor
β adrenergic receptor
α adrenergic receptor
ATP
cAMP
Inositol trisphosphate (Ins-P_3)
(R-cAMP)$_2$
Inactive protein kinase A R_2C_2
Ca^{2+}, sequestered
Free Ca^{2+}
2C (Active protein kinase A)
Inactive phosphorylase kinase
Active phosphorylase kinase
P_i
Glycogen phosphorylase *b* (inactive)
ATP
ADP
Inactive glycogen synthase
H_2O
Active protein phosphatase -2
Active protein phosphatase 1
Many kinases
H_2O
Glycogen phosphorylase *a* (active)
ADP
ATP
Active glycogen synthase
P_i
Glycogen (n residues of glucose)
P_i
UDP
Glycogen breakdown
Glycogen synthesis
Glc-1-*P*
Glycogen (n-1 residues of glucose)
UDP-Glc
H_2O
PP_i
$2P_i$
UTP
Glc-1-*P*
P_i
H_2O
Glucose
Glc-6-*P*
Glycolytic pathway

Figure 11-4 Cascades of phosphorylation and dephosphorylation reactions involved in the control of the metabolism of glycogen. Heavy arrows (➡) show pathways by which glucosyl units of glycogen are converted into free glucose or enter the glycolytic pathway. Green arrows (➡) trace the corresponding biosynthetic pathways. Gray arrows (➡) trace the pathway of activation of glycogen phosphorylase by a hormone such as adrenaline (epinephrine) or glucagon and by the action of protein kinases. A few of the related pathways in the network of reactions that affect glycogen metabolism are also shown. This includes protein phosphatases, which remove phospho groups from proteins and allow cells to relax to the state that preceded activation. One of these (protein phosphatase 1) is activated by phosphorylation by an insulin-stimulated protein kinase.[49] However, the significance is uncertain; control of glycogen synthase is complex.[49a]

TABLE 11-1
A Few Covalent Modification Reactions Utilized to Control Metabolism

Reaction	Example	Location of discussion
A. Phosphorylation–dephosphorylation		
Phosphorylation of Ser, Thr	Glycogen phosphorylase	This section
Phosphorylation of Tyr	Insulin receptor	Section G
Adenylylation, Uridylylation	Glutamine synthetase	Chapter 25
ADP-ribosylation		This section
B. Methylation of carboxyl groups	Bacterial glutamyl	Section D,5
	Aspartyl	Box 12-A
	Protein phosphatase 2A	Section C,3
	Ras	Section D,3
C. Formation of carbamino groups		
	In hemoglobin	Eq. 7-23
	In ribulose bisphosphate carboxylase	Chapters 13,23
D. Acylation		
Acetylation	Histones	Chapter 27
Palmitoylation	Ras	Section D,3
Prenylation	Ras	Chapter 22
E. Disulfide formation and cleavage		Chapters 10,15 Section C,4

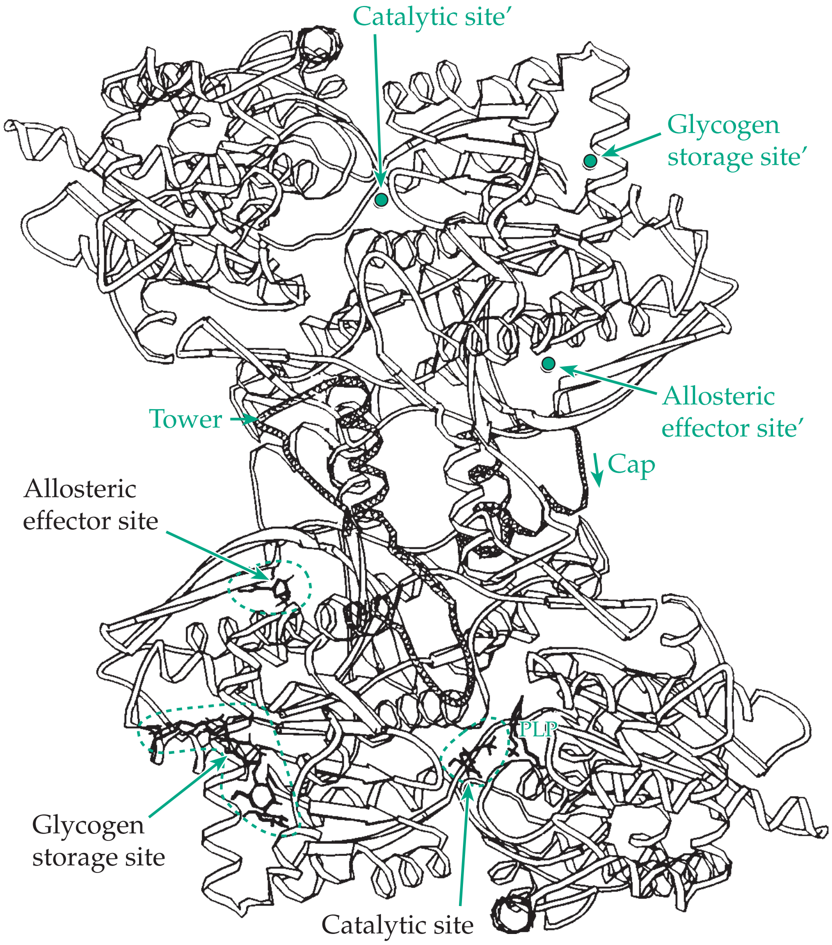
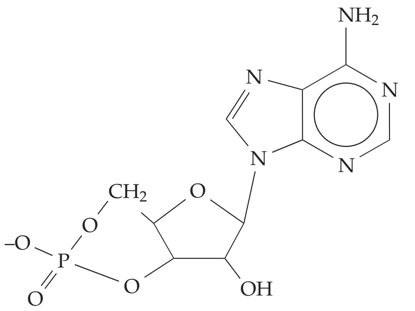


Figure 11-5 Schematic diagram of the glycogen phosphorylase dimer. The view is down the twofold axis of the dimer with the allosteric and Ser-*P* sites toward the viewer. Access to the catalytic site is from the far side of the molecule. The diagram shows the major change in conformation of the amino-terminal residues on phosphorylation. Residues 10–23 of glycogen phosphorylase *b* are shown as a thick solid line. These residues are not well ordered and make intrasubunit contacts. Upon phosphorylation, residues 10–23 change their conformation and are shown dark crosshatched with the position of Ser14-*P* indicated at the intersubunit interface. The fold of residues 24–80 through the α1 helix, the cap, and the α2 helix is shown lightly cross-hatched. The AMP allosteric effector site is located between the α2 helix and the cap region of the other subunit. The glycogen storage site is located on the surface of the subunit and is associated with a long α helix. The catalytic site is at the center of the subunit where the two domains come together. Courtesy of Louise N. Johnson.[48]

the functional groups phosphorylated, and in their allosteric activators.[39,51,54–56] There may be more than 1000 different protein kinase genes in vertebrate animals, accounting for ~2% of the genome.[57] Many cytosolic protein kinases transfer a phospho group from ATP to either a serine or threonine side chain at a β bend or other surface feature of the substrate protein. Some sites of phosphorylation in the substrate proteins contain lysine or arginine residues, separated from the serine or threonine by only one residue but many other sequences may also surround phosphorylation sites.[51,58] In the case of the 750-residue glycogen synthase, seven serine residues in different parts of the chain are phosphorylated by the action of at least five different kinases.[49,49a,59,60] The various kinases phosphorylate groups at different sites and their effects are roughly additive.

Some of the best known protein kinases (designated PKA or cAPK) are those that depend upon **3', 5'-cyclic AMP (cAMP)** as an allosteric effector. They are oligomeric proteins of composition R_2C_2, where R is a regulatory subunit and C is a catalytic subunit. Unless cAMP is present, the regulatory subunits interact with the catalytic subunits, keeping them in an inactive inhibited form. However, when cAMP is present it binds to the regulatory subunit dimer, releasing the two active catalytic units (Eq. 11-6). This reaction is



3',5'-cyclic AMP

$$R_2C_2 + 4\ cAMP \rightarrow R_2\ (cAMP)_4 + 2C \qquad (11\text{-}6)$$

reversible and as the concentration of cAMP is reduced by hydrolysis (see Eq. 11-8) the regulatory units recombine with the catalytic subunits and again inhibit them.

There are two prominent types of mammalian cAMP-dependent protein kinases.[51,61] The catalytic subunit is identical for both; the 41-kDa peptide as isolated from beef heart has 350 residues and an N terminus blocked by a myristoyl (tetradecanoyl) group.[62] One phosphoserine and one phosphothreonine are also present.[51] The ~50-kDa regulatory subunits vary in size and may also be subject to additional regulation by phosphorylation.[63] Three-dimensional structures are known for both the catalytic[62,64,65] and the regulatory[66] subunits. A **cyclic GMP** (cGMP)-activated protein kinase is present in some mammalian tissues[67,68] and is widespread in invertebrates.[51,67] The cyclic nucleotide-binding domains of these kinases have structures similar to that of the *E. coli* catabolite activator protein (Fig. 28-6)[51,68] and the catalytic domains are structurally similar to those of many other kinases.

Among the kinases that phosphorylate glycogen synthase is a **casein kinase**, named for the fact that the milk protein casein is also a good substrate. A family of casein kinases are found in the cytoplasm and nuclei of all eukaryotic cells. They phosphorylate serine and threonine side chains but have structures distinct from those of the cAMP-dependent kinases.[69–72] Casein kinase-2 (CK2) phosphorylates many proteins, including several that function in gene replication, transcription, and cell growth and division.[72a] **Theileriosis**, a parasitic disease of cattle in Africa, is caused by the tick-borne protist *Theileria parva*. The condition is often fatal as a result of a leukemia-like condition resulting from overexpression of the casein kinase-2 gene.[72]

Phosphorylase kinase is one of a large group of *specialized* protein kinases, each of which acts on a small number of proteins. It is regulated both by covalent modification and allosterically by **calcium ions**.[73–78] It contains four different kinds of subunits ranging in size from 17 kDa to about 145 kDa and has the composition $(\alpha\beta\gamma\delta)_4$. Phosphorylation of one serine on each of the 145-kDa α and 120-kDa β regulatory subunits is catalyzed by a cAMP-dependent protein kinase. The δ subunit is the Ca^{2+}-binding protein **calmodulin** (Fig. 6-8) and serves as a regulatory subunit sensitive to Ca^{2+}. The 45-kDa γ subunit contains the catalytic domain as well as a calmodulin-binding domain.[79] Other Ca^{2+}-dependent protein kinases include the **protein kinase C** family, which is discussed further in Section E, Ca^{2+}/calmodulin-dependent protein kinases,[80,81] and a plant kinase with a regulatory domain similiar to calmodulin.[82]

Protein tyrosine kinases (PTKs) place phospho groups on the phenolic oxygen atoms of tyrosyl residues of some proteins.[83–85] The resulting phosphotyrosine accounts for only about 1/3000 of the phospho groups in proteins but has aroused interest for two reasons. First, binding of growth hormones, such as **epidermal growth factor** (EGF), **platelet-derived growth factor** (PDGF), and **insulin**, to their receptors stimulates tyrosine-specific protein kinase activity of the receptor proteins (see Figs. 11-11 to 11-13).[69,83,86] Second, a number of cancer-causing **oncogenes** encode similar kinases (Section H). Tyrosine protein kinases are essential components of the cell division cycle (see Fig. 11-15).

Protein phosphatases. Most regulatory alterations in enzymatic activity are spontaneously reversible. The concentration of the allosteric effector soon drops and the covalent modifications are reversed so that the system relaxes to a state approximating the original one. Among the enzymes required for this reversal are the protein phosphatases that remove the phospho groups placed on amino acid side chains by kinases.[54,86–91] For example, four different phosphatases act on the (inactive) phosphorylated glycogen synthase and on the phosphorylated forms of glycogen phosphorylase and phosphorylase kinase. At least one of these (a protein phosphatase 1; see Fig. 11-4) is regulated by phosphorylation that is stimulated by insulin.[49,49a] The matter is made yet more complex by the presence in tissues of small protein inhibitors which can prevent the action of these phosphatases.[88,92–94] Another naturally occurring inhibitor is the polyether fatty acid **okadaic acid**, a "shellfish poison," actually produced by dinoflagellates but which accumulates in sponges and in mussels and other bivalves.[94] See also Chapter 12, D,4.

Most protein phosphatases are specific toward phosphoserine and phosphothreonine[89,95] or toward phosphotyrosine residues.[90,96–98] Some have a dual specificity.[99] There are Ca^{2+}-activated phosphatases[100] and phosphatases with transmembrane segments connected to receptors on the external cell surface. Some phosphorylation–dephosphorylation regulatory systems utilize bifunctional enzymes consisting of a protein kinase fused to a protein phosphatase. For example, the same *E. coli* enzyme that phosphorylates isocitrate dehydrogenase also hydrolyzes off the phospho group that it has put on.[101] Another bifunctional kinase–phosphatase catalyzes both the formation (reaction *d* of Fig. 11-2) and breakdown (reaction *e*) of fructose 2,6–bisphosphate in eukaryotic tissues. This kinase–phosphatase enzyme is, in turn, controlled by phosphorylation by a cAMP-dependent protein kinase. In the liver, this latter phosphorylation strongly inhibits the kinase activity preventing buildup of fructose 2,6-bisphosphate, the allosteric activator of phosphofructokinase-1, which catalyzes reaction *b* of Fig. 11-2. Consequently, when a pulse of cAMP is generated within a hepatocyte, glycolysis is blocked but **glycogenolysis**, the breakdown of glycogen by phosphorylase, is stimulated. The glucose 6-phosphate that is formed is hydrolyzed by glucose-6-phosphatase releasing glucose into the bloodstream. On the other hand, in heart and probably in most other tissues, a different isoenzyme form of the bifunctional enzyme is present. Phosphorylation by cAMP activates the kinase and inhibits the phosphatase. Consequently, in these tissues cAMP induces the activation of glycogenolysis and glycolysis coordinately.[102–105] In brain a different allosteric activator, **ribose 1, 5-bisphosphate**, may regulate phosphofructokinase-1.[106]

Phosphorylation in bacteria. A bacterial enzyme whose activity is controlled by phosphorylation is **isocitrate dehydrogenase**. Transfer of a phospho group to the –OH of Ser 113 completely inactivates the *E. coli* enzyme,[107,108] causing isocitrate to build up in the citric acid cycle (Fig. 10-6) and to be diverted into the glyoxylate pathway, which is depicted in Fig. 17-16. In this instance, it appears likely that the negative charge of the added phospho group causes electrostatic repulsion of the substrate isocitrate. In agreement with this concept, mutation of Ser 113 of this enzyme to Asp (mutant S113 D) also inactivates the enzyme.[109]

In addition to kinase–phosphatase cycles, bacteria use at least two other ATP-dependent regulatory mechanisms.[110] In the **sensor kinase/response regulator** (or "two-component") systems[110-111] a sensor protein, upon being allosterically activated, phosphorylates itself (autophosphorylation) on a specific histidyl residue to form an *N*-phosphohistidine derivative. The phosphohistidine then transfers its phospho group onto a specific aspartyl group in the response regulator causing the regulator to bind to its target protein and to exert its regulatory effect. The best known example involves the control of the motion of bacterial flagella which is discussed briefly in Section D,5 and further in Chapter 19.

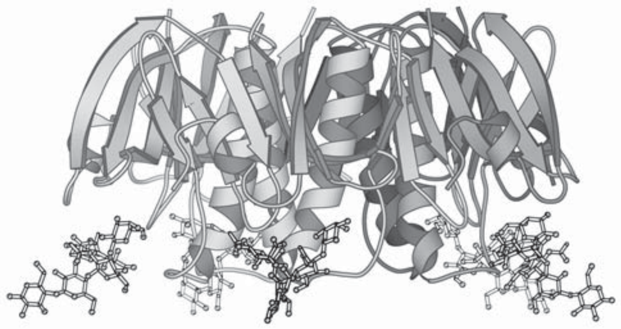
The third bacterial regulatory device is the phosphoenolpyruvate:sugar phosphotransferase system (Eq. 8-4). It is involved not only in transport but also in controlling a variety of physiological processes.[110,112,113]

3. Other Modification Reactions Involving Group Transfer

Nucleotidylation, the transfer of an entire nucleotidyl unit, rather than just a phospho group, to a protein is sometimes utilized for regulation. For example, glutamine synthetase of *E. coli* is modified by **adenylylation**, transfer of an adenylyl group from ATP to a tyrosine side chain.[114 – 116] Relaxation to the unmodified enzyme is catalyzed by a deadenylylating enzyme. A regulatory protein that undergoes reversible uridylylation[117] also functions in this system (see Fig. 24-7). Several mitochondrial and cytoplasmic proteins are modified by attachment of ADP-ribosyl groups to specific guanidino groups of arginine and other side chain groups.[114,118–123] This **ADP–ribosylation** requires the coenzyme NAD^+ as the ADP-ribosyl donor and is also catalyzed by bacterial toxins such as cholera toxin (Box 11-A) and by diphtheria toxin (Box 29-A). A second kind of ADP–ribosylation occurs within nuclei where the enzyme **poly(ADP–ribose) synthase** catalyzes both an initial ADP–ribosylation of an amino acid side chain group and also addition of more ADP-ribosyl units to form a polymer[124] (see Chapter 27).

Like inorganic phosphate, inorganic sulfate can be converted into an activated form in which the sulfate resembles the terminal phospho group of ATP (Eq. 17-38). The resulting **activated sulfo group** can be transferred to other compounds including enzymes. Formation of

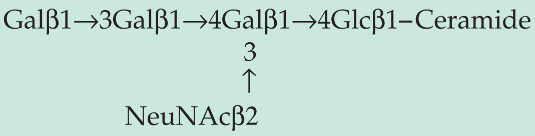
BOX 11-A CHOLERA TOXIN AND OTHER DANGEROUS PROTEINS



The three-dimensional structure of cholera toxin.[e] Side view of the β subunit pentamer as a ribbon drawing. Bound noncovalently to it are five molecules of the ganglioside G_{M1} (compare with the structure in Fig. 7-5). The diacyl glycerol parts of the gangliosides are buried in the membrane that lies below the toxin molecule. Courtesy of W. G. J. Hol.

As recently as 1980 it was estimated that there were ~ 100 million cases of acute diarrhea in Asia, Africa, and Latin America;[a] in 1991 there were four million deaths among children under five years of age.[b] The causative agents are bacteria and one of the most dangerous is *Vibrio cholerae*, which multiplies in the small intestine and secretes an **exotoxin**. Cholera toxin causes such a rapid loss of fluid and salts from the body that death occurs very quickly, even in adults. There is little cellular damage and almost all deaths can be prevented by intravenous administration of water, salts, and the antibiotic tetracycline. Fluids can also be given orally if glucose, which promotes intestinal absorption, is included with the Na^+, K^+, Cl^-, and HCO_3^- salts.[b]

Cholera toxin consists of one 240-residue A subunit and a ring of five 103-residue B subunits.[c–e] The latter is a “targeting complex,” a bacterial lectin each subunit of which binds to the galactose and sialic acid termini of a single molecule of ganglioside G_{M1}.

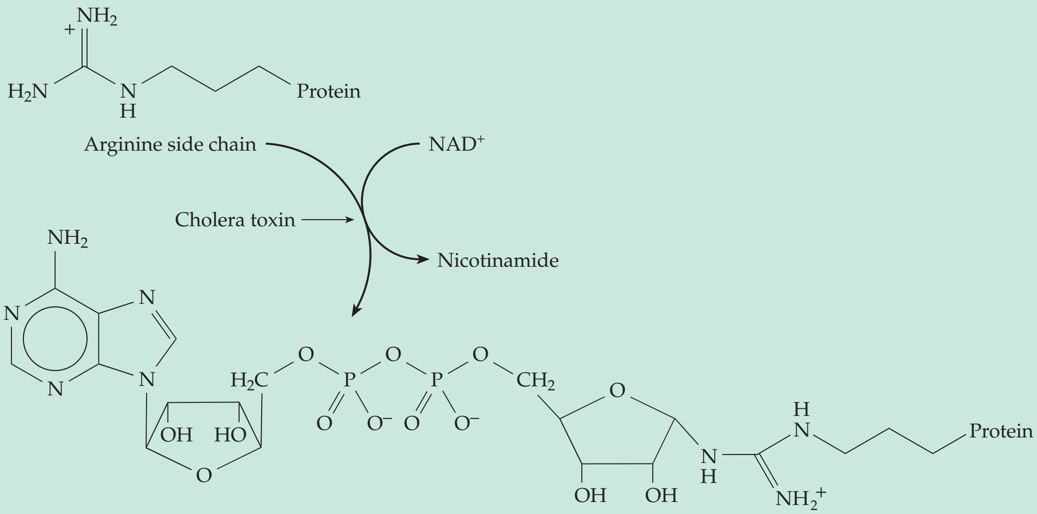


$$\text{Gal}\beta1\rightarrow3\underset{\substack{3\\ \uparrow\\ \text{NeuNAc}\beta2}}{\text{Gal}}\beta1\rightarrow4\text{Gal}\beta1\rightarrow4\text{Glc}\beta1\text{–Ceramide}$$

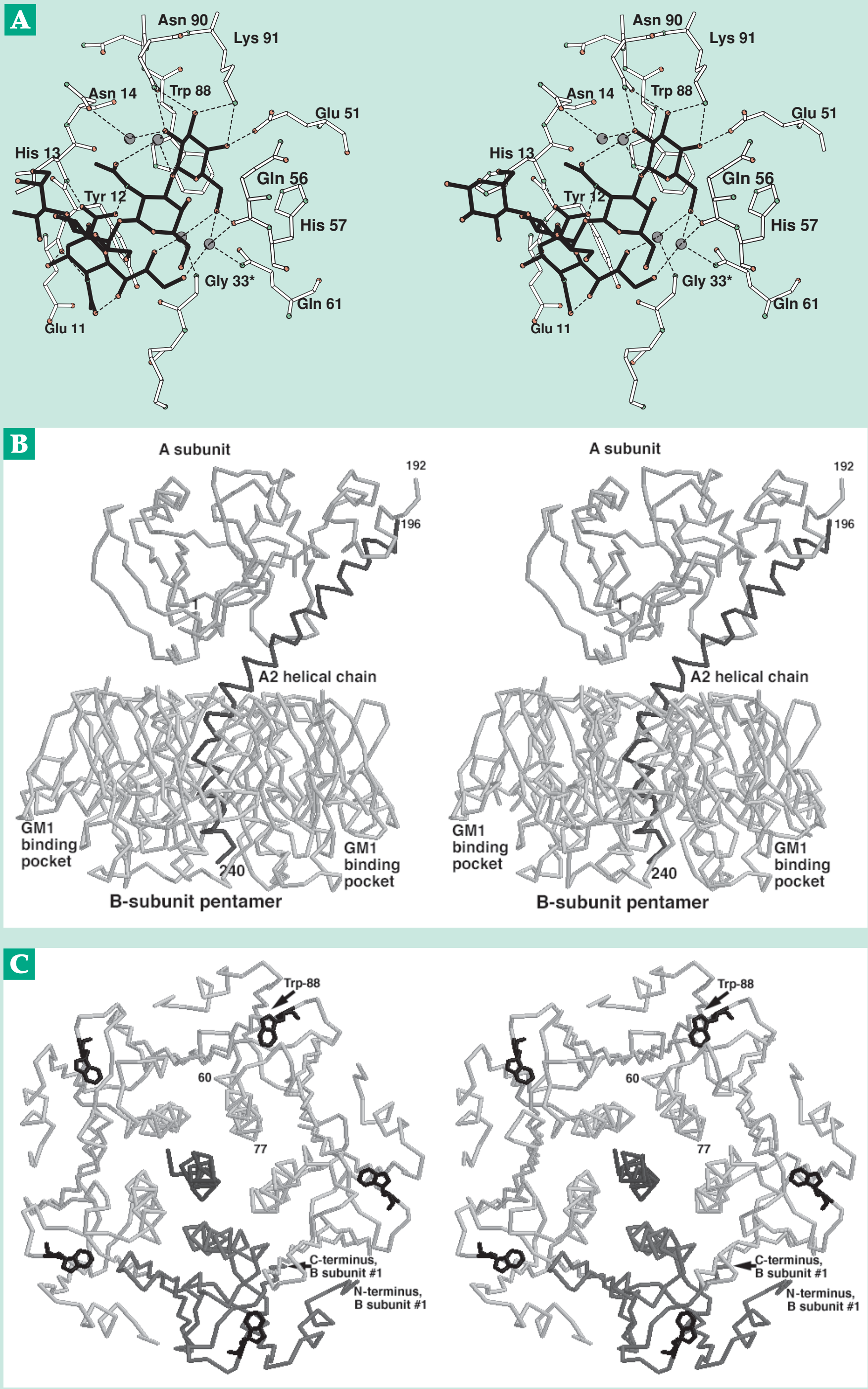
Ganglioside G_{M1} (See also Fig. 20-11)

The nonpolar ceramide chains of the ganglioside are buried in the membrane bilayer. The initial tight binding ($K_f \sim 2 \times 10^7\ M^{-1}$) of the toxin is followed by a poorly understood process in which the A subunit enters the cell. The peptide chain of this subunit is cut by bacterial proteases to give A_1 and A_2 fragments joined by an S–S bridge. Within the victim’s cells the disulfide is reduced to generate the active toxin, the 192- to 194- residue A_1 fragment.[f,g]

This toxin subunit is an enzyme, an ADP-ribosyltransferase which catalyzes transfer of ADP-ribosyl units from the coenzyme NAD^+ to specific arginine side chains to form *N*-ADP-ribosyl derivatives of various proteins. Of the proteins modified by cholera toxin, the most significant is the guanyl nucleotide regulatory protein G_s of the adenylate cyclase system.[c,f,h] ADP ribosylation of arginine 201 of the α subunit of protein G_s inhibits the GTP hydrolysis that normally allows the protein to relax to an unactivated form.[e] The ADP-ribosylated G_s keeps adenylate cyclase activated continuously and



BOX 11-A (continued)



(A) Stereoscopic view of the G_{M1} pentasaccharide binding site, showing both direct and solvent-mediated hydrogen bonding interactions between sugar and protein residues. The viewpoint is from "underneath" the membrane binding surface of the B pentamer. Starred residues are from an adjacent monomer. The terminal galactose residue of the pentasaccharide (upper right) is most deeply inserted into the binding site and is involved in the greatest number of identifiable binding interactions. The sialic acid residue, near the bottom of the figure, is also involved in several hydrogen-bonding interactions. Hydrophobic interactions include the approach of the sialic acid acetyl group to the edge of the Tyr 12 phenyl ring and the positioning of the terminal galactose sugar ring parallel to the indole ring of Trp 88. Figures generated using program MolScript (Kraulis, 1991). From Merritt *et al*.[e] (B) Side view of the intact AB_5 toxin as an α-carbon trace. The nicked A chain is at the top. The nick can be seen in the upper right corner, as can a disulfide bridge connecting the A_1 and A_2 segments. A single α helix extends into the "pore" in the center of the B pentamer. The side chains of Trp 88 of the B subunits have been added to mark the ganglioside binding sites. These side chains are also seen in (A). (C) View of the AB_5 molecule from the bottom showing the helix from the A_2 fragment surrounded by a tight cage of five long helices from the B subunits. Residues 237–240 of the A_2 fragment have the KDEL sequence and may extend from the pore to contact the membrane. (B) and (C) courtesy of Edwin M. Westbrook. See Zhang *et al*.[c]

generating abnormally high concentrations of cAMP. It is the effect of the cAMP on proteins of the intestinal mucosal cell membranes that causes the disastrous excessive secretion of water and salts that are characteristic of cholera.

Most strains of *V. cholerae* are relatively harmless but they may suddenly be transformed into a virulent toxin-producing form by infection with a bacterial virus similar to M13 (Chapter 5), which carries the toxin gene. Entrance into the *Vibrio* cell occurs with the help of pili which are present in many strains.[i,j]

Many other bacterial toxins have AB_n structures similar to that of cholera toxin. For example, toxin-producing strains of *E. coli* are also important causes of diarrhea in humans and in domestic animals.[k–m]

BOX 11-A CHOLERA TOXIN AND OTHER DANGEROUS PROTEINS (continued)

The heat-labile *E. coli* enterotoxin, whose gene is carried on a plasmid, is a close relative of cholera toxin[n,o] and also catalyzes ADP ribosylation of arginine 201 of the $G_{s\alpha}$ subunit.[m] *Bordetella pertussis*, which causes whooping cough, forms a similar toxin that attacks the inhibitory regulatory protein G_i[p,q] as well as transducin and inactivates them by ADP ribosylation. Diphtheria toxin (Box 29-A), the exotoxin from *Pseudomonas aeruginosa*, and the toxin from *Clostridium botulinum* also catalyze ADP-ribosylation reactions.[k,r]

Some bacteria produce effects similar to those of cholera toxin in different ways. For example, among a variety of toxic proteins produced by *Bacillus anthracis*, the causative agent of anthrax, is an adenylate cyclase that is able to enter the host's cells.[s] Similarly, *B. pertussis*, in addition to its ADP-ribosylating toxin, produces a calmodulin-stimulated adenylate cyclase which, when taken up by phagocytic cells, disrupts their function in the body's defense system.[t–w] In addition to the heat-labile enterotoxin, some strains of *E. coli* produce *heat-stable* toxins, small 18-residue peptides related in structure to the intestinal peptide **guanylin**.[l,x] These peptides bind to a membrane-bound guanylate cyclase activating fluid secretion into the intestine.

Why do our cells obligingly provide both initial receptors and means of uptake for these dangerous toxic proteins? Some ADP-ribosylation reactions are a natural part of cell function[y] and some hormones, for example thyrotropin, seem to stimulate their activity. It may be that the toxins use mechanisms designed to respond to normal hormonal stimulation.

---

[a] Holmgren, J. (1981) *Nature (London)* **292**, 413–417

[b] Hirschhorn, N., and Greenough, W. B., III (1991) *Sci. Am.* **264**(May), 50–56

[c] Zhang, R.-G., Scott, D. L., Westbrook, M. L., Nance, S., Spangler, B. D., Shipley, G. G., and Westbrook, E. M. (1995) *J. Mol. Biol.* **251**, 563–573

[d] Zhang, R.-G., Westbrook, M. L., Westbrook, E. M., Scott, D. L., Otwinowski, Z., Maulik, P. R., Reed, R. A., and Shipley, G. G. (1995) *J. Mol. Biol.* **251**, 550–562

[e] Merritt, E. A., Sarfaty, S., van den Akker, F., L'Hoir, C., Martial, J. A., and Hol, W. G. J. (1994) *Protein Sci.* **3**, 166–175

[f] Moss, J., Stanley, S. J., Morin, J. E., and Dixon, J. E. (1980) *J. Biol. Chem.* **255**, 11085–11087

[g] Janicot, M., Fouque, F., and Desbuquois, B. (1991) *J. Biol. Chem.* **266**, 12858–12865

[h] Tsai, S.-C., Adamik, R., Moss, J., and Vaughan, M. (1991) *Biochemistry* **30**, 3697–3703

[i] Williams, N. (1996) *Science* **272**, 1869–1870

[j] Waldor, M. K., and Mekalanos, J. J. (1996) *Science* **272**, 1910–1914

[k] Moss, J., and Vaugham, M., eds. (1990) *ADP–Ribosylating Toxins and G Proteins: Insights into Signal Transduction*, American Society for Microbiology, Washington, DC

[l] Sato, T., Ozaki, H., Hata, Y., Kitagawa, Y., Katsube, Y., and Shimonishi, Y. (1994) *Biochemistry* **33**, 8641–8650

[m] van den Akker, F., Merritt, E. A., Pizza, M. G., Domenighini, M., Rappuoli, R., and Hol, W. G. J. (1995) *Biochemistry* **34**, 10996–11004

[n] Sixma, T. K., Pronk, S. E., Kalk, K. H., van Zanten, B. A. M., Berghuis, A. M., and Hol, W. G. J. (1992) *Nature (London)* **355**, 561–564

[o] Sixma, T. K., Kalk, K. H., van Zanten, B. A. M., Dauter, Z., Kingma, J., Witholt, B., and Hol, W. G. J. (1993) *J. Mol. Biol.* **230**, 890–918

[p] Antoine, R., and Locht, C. (1994) *J. Biol. Chem.* **269**, 6450–6457

[q] Goodemote, K. A., Mattie, M. E., Berger, A., and Spiegel, S. (1995) *J. Biol. Chem.* **270**, 10270–10277

[r] Ohtsuka, T., Nagata, K.-i, Iiri, T., Nozawa, Y., Ueno, K., Ui, M., and Katada, T. (1989) *J. Biol. Chem.* **264**, 15000–15005

[s] Arora, N., Klimpel, K. R., Singh, Y., and Leppla, S. H. (1992) *J. Biol. Chem.* **267**, 15542–15548

[t] Benz, R., Maier, E., Ladant, D., Ullmann, A., and Sebo, P. (1994) *J. Biol. Chem.* **269**, 27231–27239

[u] Heveker, N., Bonnaffé, D., and Ullmann, A. (1994) *J. Biol. Chem.* **269**, 32844–32847

[v] Otero, A. S., Yi, X. B., Gray, M. C., Szabo, G., and Hewlett, E. L. (1995) *J. Biol. Chem.* **270**, 9695–9697

[w] Hackett, M., Guo, L., Shabanowitz, J., Hunt, D. F., and Hewlett, E. L. (1994) *Science* **266**, 433–435

[x] Ozaki, H., Sato, T., Kubota, H., Hata, Y., Katsube, Y., and Shimonishi, Y. (1991) *J. Biol. Chem.* **266**, 5934–5941

[y] Moss, J., and Vaughan, M. (1995) *J. Biol. Chem.* **270**, 12327–12330

---

tyrosine-*O*-sulfate residues may be a widespread regulatory mechanism.[125–127] One of the proteins known to contain a tyrosine sulfate residue is the blood protein fibrinogen. Many polysaccharides and oligosaccharides on glycoproteins exist in part as sulfate esters (Fig. 4-11).[128,129]

Carboxyl groups of certain glutamate side chains in proteins that control bacterial **chemotaxis** are methylated reversibly to form methyl esters.[130] This **carboxylmethylation** occurs as part of a reaction sequence by which the bacteria sense compounds that can serve as food or that indicate the presence of food (Section D,5, Fig. 11-8, and Chapter 19). Demethylation occurs by hydrolysis, which is catalyzed by esterases. Carboxylmethylation also occurs in eukaryotic cells but is often substoichiometric and part of a mechanism for repair of isomerized or racemized aspartyl residues in aged proteins (Box 12-A). However, the major eukaryotic protein phosphatase 2A is carboxylmethylated at its C terminus,[131] as are the Ras proteins discussed in Section D,3.

It was pointed out in Chapter 4 that many proteins, especially those secreted from cells or taking up residence within membranes, are **glycosylated**. Specific glycosyl groups may be removed or added to a protein. Such alterations may all be regarded as parts of control mechanisms that direct the proteins to their proper locations or determine the length of time that they remain active. The nuclear histones undergo extensive and reversible **acetylation**[132,133] which is thought to be important to replication of DNA and to transcription (Chapters 27 and 28).

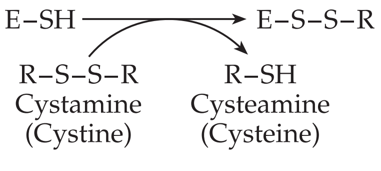
4. Thiol-Disulfide Equilibria

The activity of enzymes is sometimes controlled by the formation or the reductive cleavage of disulfide linkages between cysteine residues within the protein. An example is provided by the effects of light on enzymes of chloroplasts. Light absorbed by the photosynthetic reaction centers generates NADPH, which in turn reduces **thioredoxin**. This is a small protein containing a readily accessible S–S bridge that can be reduced to a pair of SH groups (Box 15-C). Within illuminated chloroplasts, these newly formed SH groups reduce disulfide bridges to activate a series of enzymes including fructose 1,6-bisphosphatase that participate in the photosynthetic incorporation of CO_2 into sugars[134,135] (see Chapter 23). Thioredoxins also function within bacteria, fungi, and animals, serving as electron carriers for processes, some of which are involved in metabolic control. A number of known enzymes and other proteins, including insulin, contain reducible S–S bridges within peptide loops as in the thioredoxins. Another possible control mechanism for SH-containing enzymes depends upon formation of **mixed disulfides** with small SH-containing metabolites. For example, **cysteamine disulfide**, a minor constituent of cells, cysteine, or some small disulfide, could be released

$$H_3{}^+N{-}CH_2{-}CH_2{-}S{-}S{-}CH_2{-}CH_2{-}NH_3{}^+$$

Cysteamine disulfide (cystamine)

following hormonal or other stimulation and could react by disulfide exchange (Eq. 11-7) to either inactivate or activate an enzyme.



(11-7)

Almost all cells contain a high concentration (3–9 mM) of the thiol-containing tripeptide **glutathione** (G–SH, Box 11-B). In its disulfide form it participates in forming disulfide bridges in secreted extracellular proteins (Eq. 10-9) via intermediate mixed disulfides. Mixed disulfides with glutathione as well as with other thiols can also be formed within cells by oxidative reactions. Reduction of these disulfides by reduced glutathione will return the enzymes to their reduced states. The small protein **glutaredoxin**, whose eukaryotic forms are also called **thioltransferase**, resembles thioredoxins. It undergoes reduction by glutathione and, in turn, reduces S–S linkages in a different set of proteins than those reduced by thioredoxin.

Low-molecular-mass thiols such as coenzyme A and protein-bound thiol cofactors such as phosphopantetheine are present in all cells. Their SH groups can also be oxidized to disulfides and it is of interest that in resting bacterial spores these compounds exist largely as disulfides or mixed disulfides. Upon germination of the spores special enzymes reduce the disulfides.[136] Some proteins involved in control of protein synthesis contain SH groups that add covalently to C-6 atoms of a uracil ring in specific mRNA molecules. Control of their state of reduction may also be important.[137]

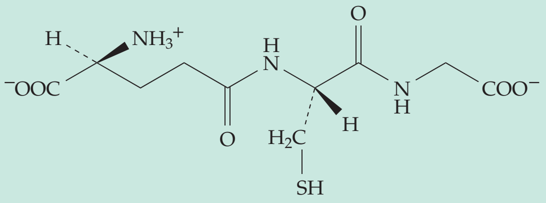
5. Regulatory Effects of H^+, Ca^{2+}, and Other Specific Ions

The pH of the cytoplasm is controlled tightly. Yet transient changes can occur and may affect many aspects of metabolism. For example, rapid glycolysis leads to lactic acid formation with an associated drop in internal pH. Both increases and decreases in pH have been associated with successive stages in embryonic or larval development.[138] Cytoplasmic pH changes may serve as regulatory signals. A well-understood example is the Bohr effect on oxygenation of hemoglobin (Chapter 7). Another is the protein kinase C-stimulated H^+/Na^+ exchange through membranes. Because the Na^+ concentration is high outside cells and low inside, the exchange leads to an increase in cytosolic pH with many resultant effects on metabolism.[139,140] Exchange of external Cl^- for internal $HCO_3{}^-$ also affects pH.[141]

Uptake of Ca^{2+} into cells, or release of this ion from intracellular stores, is a major regulatory mechanism in many if not all cells (see Section E). Mn^{2+} activates **phosphoenolpyruvate carboxykinase** (Eq. 13-46) and may be a regulator of gluconeogenesis.[142] Iron controls the synthesis of ferritin and of transferrin receptors[137] (Chapter 16). The specific metal ions present in many biological macromolecules are likely to participate in additional regulatory processes.

Phosphate and bicarbonate ions are important substrates for many enzymatic processes and as such have regulatory functions. Bicarbonate controls the key enzyme of photosynthesis, **ribulose bisphosphate carboxylase**, by carbamate formation (Fig. 13-12). Chloride ions activate amylases and may affect the action of “G proteins” that mediate hormone actions. Other observed effects of ions are too numerous to mention.

BOX 11-B GLUTATHIONE, INTRACELLULAR OXIDATION–REDUCTION BUFFER



Glutathione (γ-L-Glutamyl-L-cysteinylglycine)

In 1888, de Rey-Pailhade discovered the sulfur-containing tripeptide that we now know as glutathione (G–SH).[a] By 1929 it had been characterized by F. G. Hopkins and others as an unusual tripeptide present in most, if not all, eukaryotic cells. Within animal cells the concentration is typically 1–5 mM. Lower levels are found in many bacteria.[b–g]

The most interesting chemical characteristics of glutathione are the γ-glutamyl linkage and the presence of a free SH group. The latter can be oxidized to form a disulfide bridge linking two glutathione molecules.

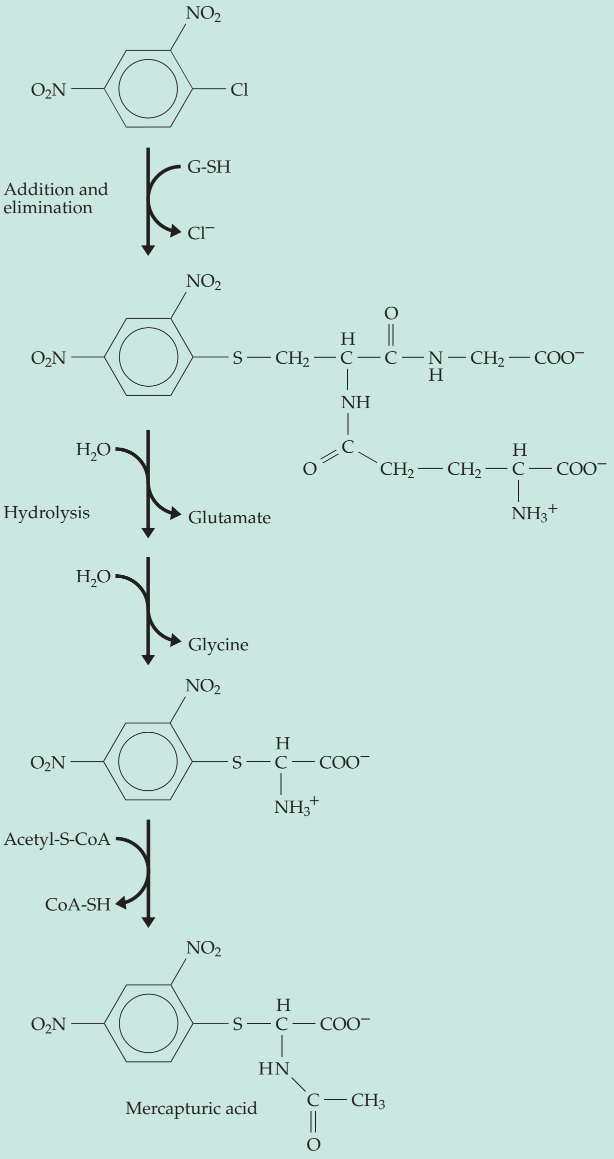
$$2G—SH - 2[H] \rightarrow G—S—S—G$$
$$E°' \text{ (pH 7)} = -0.25 \text{ V}$$

It is this redox reaction that has focused attention on glutathione as an intracellular reducing agent whose primary function may be to keep the SH groups of proteins reduced (see Section C). Glutathione is usually maintained in its reduced form by the flavoprotein glutathione reductase (Fig. 15-12).

Although it is primarily an intracellular compound, glutathione is secreted by epithelial and other cells. It may regulate the redox state of proteins in plasma and other extracellular fluids as well as within cells. In addition, glutathione released from the liver may be an important source of cysteine for other tissues. In the endoplasmic reticulum and the periplasm of bacteria glutathione functions in crosslinking thiol groups in newly formed proteins (Eq. 10-9).

Glutathione also has a series of **protective functions**. It reduces peroxides via the selenium-containing **glutathione peroxidase** (Box 15-H) and is part of a system for trapping and detoxifying harmful free radicals. The importance of this function is suggested by the fact that *Entamoeba histolytica*, an organism that lacks both mitochondria and aerobic respiration, produces no glutathione.[h] It may be that the primary function of glutathione in eukaryotes is to protect cells against oxygen toxicity associated with their mitochondria.[i] The renewal of free radicals and regeneration of protein –SH groups may involve the cooperation of glutathione, glutaredoxin (Box 15 - C), and ascorbic acid.[j]

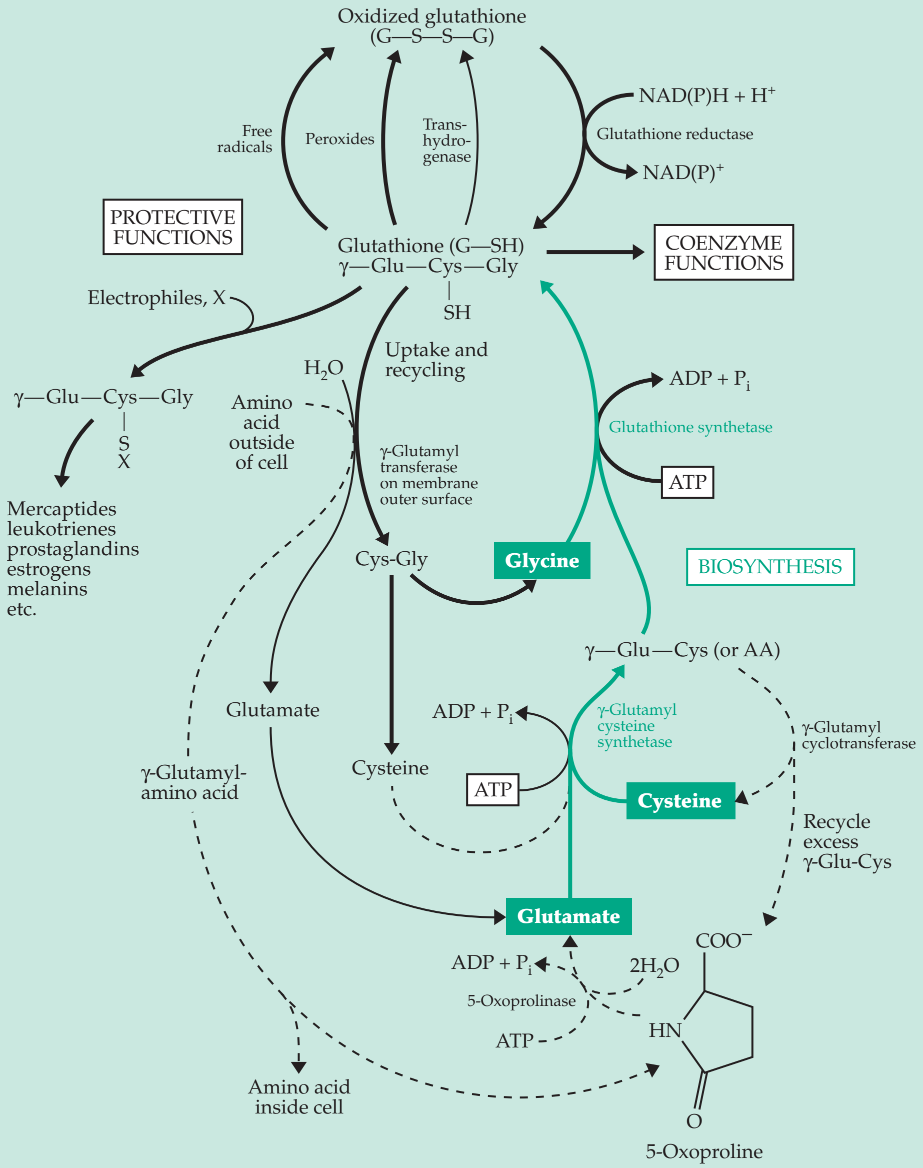
Another protective function is fulfilled by the formation of soluble **mercaptides** and other "conjugates" of glutathione with many foreign substances (**xenobiotics**). These conjugates are made by the action of a large family of **glutathione transferases**[e,k,l] (see also Chapter 13) which catalyze addition reactions of the thiolate group of glutathione with epoxides, alkylating compounds, and chlorinated aromatic hydrocarbons. The addition step is often followed by elimination, e.g., of chloride ion, as in the following example. Two steps of hydrolysis and an acetylation by acetyl-CoA form a mercapturic acid:



BOX 11-B (continued)

The mercapturic acids and related compounds can then be exported from cells by an ATP-dependent export pump.[m] Glutathione is a coenzyme for **glyoxalase** (Eq. 13-33), **maleylacetoacetate isomerase** (Eq. 13-20), and **DDT dehydrochlorinase**. The latter enzyme catalyzes elimination of HCl from molecules of the insecticide and is especially active in DDT-resistant flies.[a] Glutathione is said to be the specific factor eliciting the feeding reaction of *Hydra*; that is, the release of glutathione from injured cells causes the little animal to engulf food.

Synthesis of glutathione occurs within cells via the ATP-dependent reactions in the following scheme.[n,o] Much more γ-glutamylcysteine is synthesized than is converted to glutathione and the excess is degraded by γ-glutamyl cyclotransferase to form the cyclic amide 5-oxoproline. Cleavage of ATP is required to reopen the ring to form glutamate. Although biosynthesis is exclusively intracellular, most glutathione appears to be secreted and degraded by extracellular enzymes. The membrane-bound γ-glutamyl transferase catalyzes hydrolysis of


Oxidized glutathione
(G—S—S—G)
Free radicals
Peroxides
Trans-hydro-genase
NAD(P)H + H⁺
Glutathione reductase
NAD(P)⁺
PROTECTIVE FUNCTIONS
Glutathione (G—SH)
γ—Glu—Cys—Gly
SH
COENZYME FUNCTIONS
Electrophiles, X
γ—Glu—Cys—Gly
S
X
Mercaptides
leukotrienes
prostaglandins
estrogens
melanins
etc.
H₂O
Uptake and recycling
Amino acid outside of cell
γ-Glutamyl transferase on membrane outer surface
ADP + Pᵢ
Glutathione synthetase
ATP
Cys-Gly
Glycine
BIOSYNTHESIS
γ—Glu—Cys (or AA)
Glutamate
ADP + Pᵢ
γ-Glutamyl cysteine synthetase
γ-Glutamyl cyclotransferase
γ-Glutamyl-amino acid
Cysteine
ATP
Cysteine
Recycle excess γ-Glu-Cys
Glutamate
COO⁻
ADP + Pᵢ
2H₂O
5-Oxoprolinase
HN
ATP
Amino acid inside cell
O
5-Oxoproline

Scheme illustrating interrelationships of the biosynthesis and protective, coenzymatic, and transport functions of glutathione. See also Meister.[c]

BOX 11-B GLUTATHIONE, INTRACELLULAR OXIDATION–REDUCTION BUFFER (continued)

glutathione to cysteinylglycine which is further cleaved by a peptidase. The activity of γ-glutamyltransferase varies among tissues and is especially high in cells of the kidney tubules. The cysteine and glycine released by the peptidase may reenter the cells by a Na^+-dependent process. Meister proposed that the γ-glutamyltransferase, acting on amino acids, forms γ-glutamyl amino acids (see scheme) which are released *within* cells and are cleaved by γ-glutamyl cyclotransferase.[c] The cleavage of glutathione would provide the driving force for amino acid uptake. However, this is probably a minor pathway.[p]

Trypanosomes contain little glutathione but a large amount of **trypanothione** [N^1,N^8-bis (glutathionyl) spermidine].[q] This diamide of spermidine (Chapter 24) is in equilibrium with its disulfide, a 24-membered macrocyclic structure, and appears to have functions similar to those of glutathione. Trypanothione reductase, which is unique to trypanosomes, is a potential target for antitrypanosomal drugs,[q–s] as is trypanothione synthetase.[t] Bacteria that do not synthesize glutathione usually accumulate some other thiol, e.g., α-glutamylcysteine or coenzyme A, in high concentrations.[u,v]

---

[a] Meister, A. (1988) *Trends Biochem. Sci.* **13**, 185–188

[b] Bernofsky, C., and Wanda, S.-Y. C. (1982) *J. Biol. Chem.* **257**, 6809–6817

[c] Meister, A. (1988) *J. Biol. Chem.* **263**, 17205–17208

[d] Inoue, M. (1985) in *Renal Biochemistry* (Kinne, R. K. H., ed), pp. 225–269, Elsevier, Amsterdam

[e] Orrenius, S., Ormstad, K., Thor, H., and Jewell, S. A. (1983) *Fed. Proc.* **42**, 3177–3188

[f] Viña, J., ed. (1990) *Glutathione: Metabolism and Physiological Functions*, CRC Press, Inc., Boca Raton, Florida

[g] Dolphin, D., Poulson, R., and Avramovic, O., eds. (1989) *Glutathione - Chemical, Biochemical and Medical Aspects (Coenzymes and Cofactors)*, Vol. 3, Wiley, New York (Parts A & B)

[h] Fahey, R. C., Newton, G. L., Arrick, B., Overdank-Bogart, T., and Aley, S. B. (1984) *Science* **224**, 70–72

[i] Garcia de la Asuncion, J., Millan, A., Pla, R., Bruseghini, L., Esteras, A., Pallardo, F. V., Sastre, J., and Viña, J. (1996) *FASEB J.* **10**, 333–338

[j] Meister, A. (1994) *J. Biol. Chem.* **269**, 9397–9400

[k] Ji, X., Johnson, W. W., Sesay, M. A., Dickert, L., Prasad, S. M., Ammon, H. L., Armstrong, R. N., and Gilliland, G. L. (1994) *Biochemistry* **33**, 1043–1052

[l] Hebert, H., Schmidt-Krey, I., and Morgenstern, R. (1995) *EMBO J.* **14**, 3864–3869

[m] Ishikawa, T. (1992) *Trends Biochem. Sci.* **17**, 463–468

[n] Polekhina, G., Board, P. G., Gali, R. R., Rossjohn, J., and Parker, M. W. (1999) *EMBO J.* **18**, 3204–3213

[o] Lu, S. C. (1999) *FASEB J.* **13**, 1169–1183

[p] Lee, W., Hawkins, R., Peterson, D., and Viña, J. (1996) *J. Biol. Chem.* **271**, 19129–19133

[q] Henderson, G. B., Ulrich, P., Fairlamb, A. H., Rosenberg, I., Pereira, M., Sela, M., and Cerami, A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5374–5378

[r] Bollinger, J. M., Jr., Kwon, D. S., Huisman, G. W., Kolter, R., and Walsh, C. T. (1995) *J. Biol. Chem.* **270**, 14031–14041

[s] Sullivan, F. X., Sobolov, S. B., Bradley, M., and Walsh, C. T. (1991) *Biochemistry* **30**, 2761–2767

[t] Henderson, G. B., Yamaguchi, M., Novoa, L., Fairlamb, A. H., and Cerami, A. (1990) *Biochemistry* **29**, 3924–3929

[u] Sundquist, A. R., and Fahey, R. C. (1989) *J. Biol. Chem.* **264**, 719–725

[v] Swerdlow, R. D., Green, C. L., Setlow, B., and Setlow, P. (1979) *J. Biol. Chem.* **254**, 6835–6837

6. Compartments and Organized Assemblies

The geometry of cell construction provides another important aspect of metabolic control. In a bacterium, the periplasmic space (Fig. 8-28) provides a compartment that is separate from the cytosol. Some enzymes are localized in this space and do not mix with those within the cell. Other enzymes are fixed within or attached to the membrane. Eukaryotic cells have more compartments: nuclei, mitochondria (containing both matrix and intermembrane spaces), lysosomes, microbodies, plastids, and vacuoles. Within the cytosol the tubules and vesicles of the endoplasmic reticulum (ER) separate off other membrane-bounded compartments. The rate of transport of metabolites through the membranes between compartments is limited and often is controlled tightly.

While many enzymes appear to be dissolved in the cytosol and have no long-term association with other proteins, enzymes that catalyze a series of consecutive reactions may form complexes within which substrates are **channeled**.[143–147] Many enzymes are attached to membranes where they may be held close together as organized assemblies.[146] This appears to be the case for oxidative enzymes of mitochondria (Chapter 18) and for the cytoplasmic fatty acid synthetases (Chapter 21). In bacterial fatty acid synthesis, the product of the first enzyme is covalently attached to a “carrier” and, while so attached, is subjected to the action of a series of other enzymes. In eukaryotes several enzymes form domains of a single fatty acid synthase. Efficient substrate channeling results. Tryptophan synthase (Fig. 25-3) passes indole through a tunnel between subunits.[146a] Both NH_3 and carbamate ions pass through tunnels between subunits of *E. coli* carbamoyl phosphate synthetase (Eq. 24-22). The product carbamoyl phosphate may then be passed directly to aspartate carbamoyltransferase for synthesis of carbamoylaspartate in the pyrimidine biosynthetic pathway.[146b] Channeling is sometimes difficult to prove. Geck and Kirsch have provided a generally useful technique for testing. A large amount of a genetically modified, inactive form of an enzyme is added. Unless channeling occurs, this will decrease the rate of a reaction whose rate is limited by diffusion or by instability of an intermediate.[146c]

Proteasomes (Box 7-A) have enzymatic sites within a protected box which limits the escape of long peptide fragments. Membrane anchors (Chapter 8), often consisting of acyl or polyprenyl groups, hold many proteins to cell surfaces and strong protein–protein bonds hold many others.[146d] Enzymes involved in cell signaling, and discussed in the following section, are often anchored close together on membrane surfaces.[148,149]

D. Hormones and Their Receptors

A major element in the control of the metabolism of a cell is provided by chemical messages sent from *other cells* and sensed by receptors on the cell membrane or in the cytoplasm. Hormones such as insulin, adrenaline, and the sex hormones are released from an organ and travel through the blood, affecting tissues throughout the body. There are many such hormones. A large number of other hormones have more local effects, influencing mostly adjacent cells. When released at nerve endings these substances are called **neurotransmitters**. Chemical messages are probably sent from virtually all cells to their immediate neighbors, affecting both their growth and their behavior. In recent years the very reactive and toxic compounds **nitric oxide (NO)** and **carbon monoxide (CO)** have been identified as important hormones. These compounds can diffuse rapidly and react with many compounds within cells. Their actions are more rapid, more rapidly ended, and probably less specific than those of most hormones.

Hormones bind to **receptors** on their "target" cells. The receptors are often integral membrane proteins on cell surfaces. Binding of the hormone often causes a conformational alteration that "activates" the receptor. In some cases the receptor is an enzyme and the hormone an allosteric activator. In others, the activated receptor interacts with an enzyme in the cytoplasm or on the membrane facing the cytoplasm. This enzyme may generate a **second messenger**, a substance that can diffuse throughout the cell and alter metabolism by exerting allosteric effects on various other enzymes. The best known second messenger is cyclic AMP but there are many others, a few of which are listed in Table 11-2.

Some hormone–receptor complexes enter the cell via endocytosis in coated vesicles. Within the cell both the receptor and the bound hormone, if a peptide, may be degraded by proteases. The initial binding of the hormone may induce the release of a second messenger, while the degradation of the receptor complex at a later time releases peptides that may be additional second messengers. This is one way in which hormones may elicit a rapid response followed by delayed responses.

While many hormones bind to surface receptors the steroid hormones, which are lipid in nature, pass through the cell membrane and bind to receptor proteins in the nucleus. The resulting hormone–protein complexes induce changes in gene expression through regulation of transcription (Fig. 11-1, top). These receptors are considered in Chapter 22 and hormones are considered further in Chapter 30.

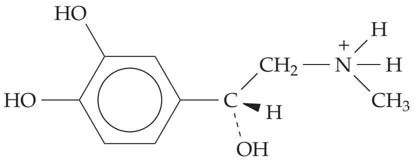
1. Beta Adrenergic Receptors and Related Seven-Helix Proteins

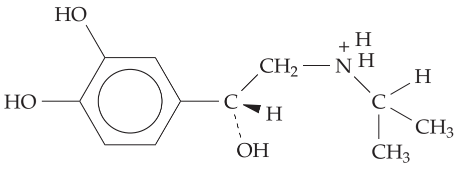
Sites that bind **adrenaline** (epinephrine), **noradrenaline** (norepinephrine), and related **catecholamines** (see Chapter 30) to almost all cell surfaces are classified as either **α adrenergic** or **β adrenergic receptors**. The β receptors, which have been studied the most,[150] occur as two major types. The β_1 receptors have approximately equal affinity for adrenaline and noradrenaline, whereas the β_2 receptors, the most common type, are more nearly specific for adrenaline. Binding of the hormone or other agonist such as **isoproterenol** to any of the β receptors stimulates cAMP formation within the cell.[153] The receptors belong to a family of integral membrane proteins related in sequence to the **opsins** of the retina and the light-operated proton pump bacteriorhodopsin (Fig. 23-45) and probably have a very similar three-dimensional structure. Based upon the known three-dimensional structures of rhodopsin and bacteriorhodopsin (Chapter 23), sequence comparisons, and much other evidence, all of these receptors are probably folded into groups of seven hydrophobic membrane-spanning helices arranged as closely packed bundles with folded loops protruding into the cytoplasm and into the extracellular space. The N termini of the proteins are thought to be in the extracellular space and the C termini in the cytoplasm as shown in Fig. 11-6.[153a,b,c] Defects in structure or functioning of human β receptors have been associated with both asthma[154] and heart failure.[155,155a,b]

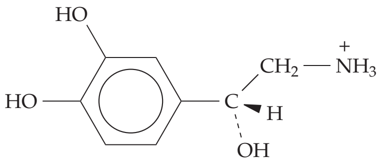
Among the proteins phosphorylated in response to formation of cAMP are the β adrenergic receptors themselves. Phosphorylation occurs within a cluster of serine and threonine residues near the C terminus as is indicated in Fig. 11-6A by the action of various kinases including a **receptor kinase**[156,156a] which may be anchored nearby.[157] The effect of this C-terminal phosphorylation is to decrease the sensitivity of the receptor so that after a few minutes it conveys a diminished response. Thus, cAMP exerts feedback inhibition of its own synthesis. *Desensitization upon continuous occupancy*

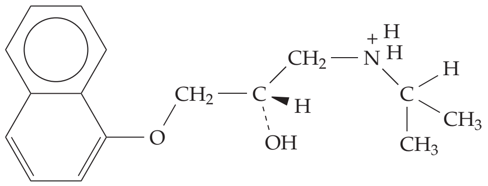
TABLE 11-2
Some Molecules (Second Messengers) That Carry Intracellular Signals

Compound	Metabolic state	Response	Location of discussion
Cyclic AMP	Stimulation of β adrenergic receptors	Increased glycogenolysis, glycolysis	Sections C,2 and D,2;
Cyclic GMP	Visual stimulation	Neuronal signal	Chapter 23, Section D
Ca^{2+}	Ca^{2+} channels open	Muscular contraction, others	Box 6-D, Section E Chapter 19, Section B,4
Inositol-1,4,5-trisphosphate and related compounds	Stimulation of α adrenergic receptors, various other stimuli	Opens Ca^{2+} channels in ER	Section E
Diadenosine 5'-tetraphosphate (Ap_4A)	Oxidative or heat stress		Chapters 28, 29
Guanosine 5'-diphosphate, 3-diphosphate (ppGpp)	Nutritional stress		Chapter 29
Mn^{2+}	Low glucose	Increased PECK activity	Chapter 17, Section L
		Increased arginase activity	Chapter 16, Section E
Cyclic ADP-ribose (cADPR) and 2'-P-cADPR			Chapter 15, Section E


Adrenaline (epinephrine)


Isoproterenol


Noradrenaline


Propranolol

by an agonist is a property of many other receptors as well. When a receptor is no longer occupied, phosphatases remove the phospho groups added by the receptor kinase permitting the sensitivity to rise again. This phosphorylation–dephosphorylation cycle appears to be only one of several mechanisms by which cells regulate receptor sensitivity.[155b,158,158a,158b]

Many other receptors also have a seven-helix structure similar to that of the adrenergic receptors. These include receptors for the following: **glucagon**, one of the pancreatic hormones regulating glucose metabolism;[159,160] **vasopressin** (Fig. 2-4);[161,162] **lutropin**, another pituitary hormone;[163] other gonadotropins;[164] the **thyrotropin-releasing factor** (TRF; Fig. 2-4);[165] a receptor for **KDEL** peptide sequences. The KDEL receptor functions in the return of soluble proteins containing the KDEL motif from the Golgi to the endoplasmic reticulum.[152]

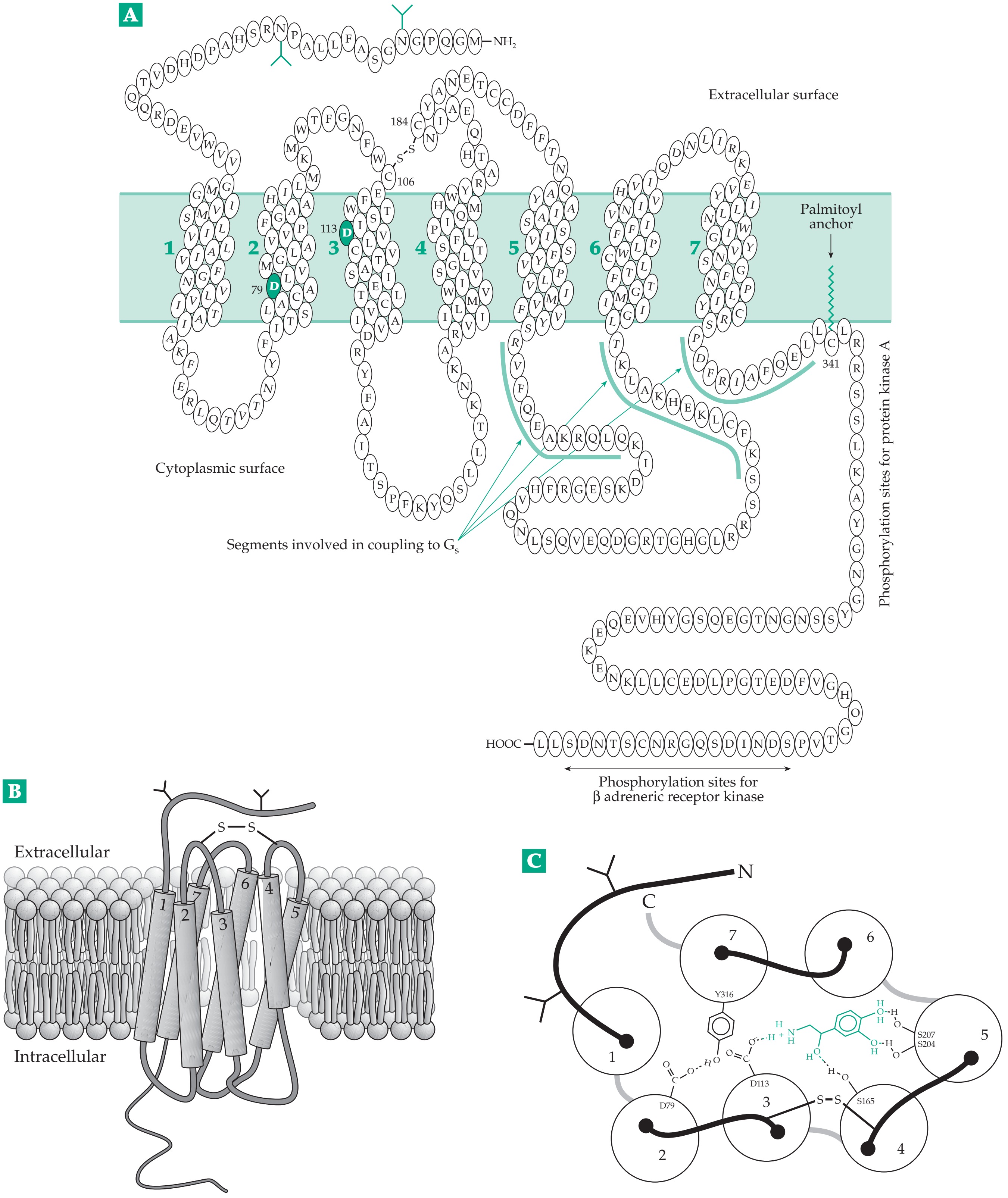
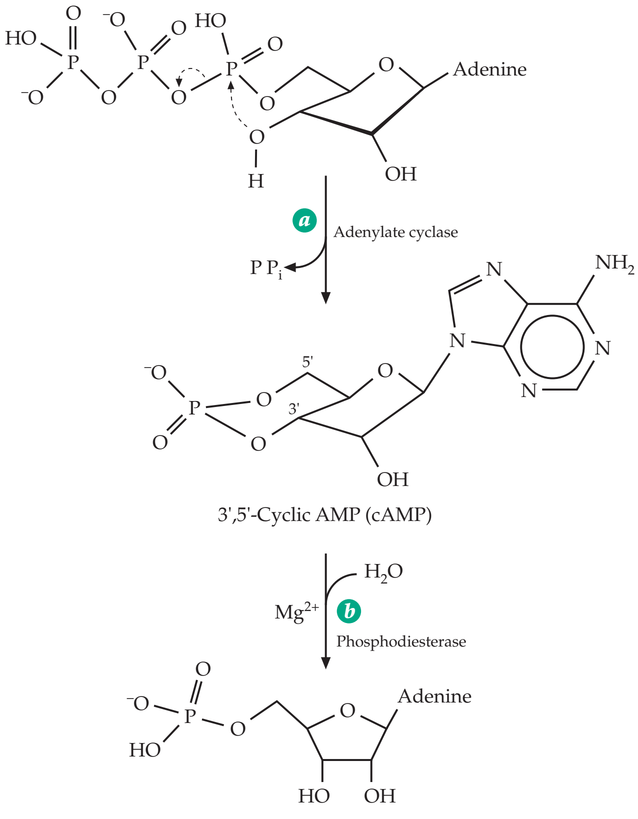

A
NH$_2$
Extracellular surface
184
106
113
79
1
2
3
4
5
6
7
Palmitoyl anchor
341
Cytoplasmic surface
Segments involved in coupling to G$_s$
Phosphorylation sites for protein kinase A
HOOC
Phosphorylation sites for β adreneric receptor kinase
B
S—S
Extracellular
Intracellular
C
N
C
Y316
D79
D113
S207
S204
S165
S–S

Figure 11-6 (A) Proposed organization of the human β_2 adrenergic receptor. The 413-residue polypeptide chain is arranged according to the model for rhodopsin as seven transmembrane helices. The two N-glycosylation sites, which may carry very large oligosaccharides, are indicated by Y. The palmitoylated cysteine 341 is shown with its alkyl side chain embedded in the membrane. Aspartates 79 and 113, also shown in (C), are shaded. After Strosberg[151] and Scheel and Pelham.[152] (B) Arrangement of the seven helices suggested by the rhodopsin structure. (C) Hypothetical view of the receptor from the external membrane surface showing a molecule of noradrenaline bound to hydrophilic residues deep in the cleft between the helices. From Strosberg.[151]

2. Adenylate Cyclases (Adenylyl Cyclases)

All of the effects of the catecholamines bound to β adrenergic receptors and of **glucagon**, **ACTH**, and many other hormones appear to be mediated by **adenylate cyclase**. This integral membrane protein catalyzes the formation of cAMP from ATP (Eq. 11-8, step *a*). The reaction, whose mechanism is considered in Chapter 12, also produces inorganic pyrophosphate. The released cAMP acts as the second messenger and diffuses rapidly throughout the cell to activate the cAMP-dependent protein kinases and thereby to stimulate phosphorylation of a selected group of proteins (Fig. 11-4). Subsequent relaxation to a low level of cytosolic cAMP is accomplished by hydrolysis of the cAMP by a phosphodiesterase (Eq. 11-8, step *b*).[166,167] In the absence of phosphodiesterase cAMP is extremely stable kinetically. However, it is thermodynamically unstable with respect to hydrolysis.

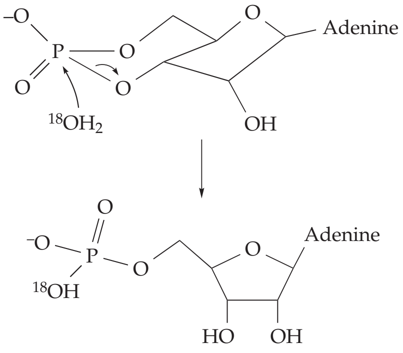
The existence of cAMP as a compound mediating the action of adrenaline and glucagon on glycogen phosphorylase was first recognized in 1956 by Sutherland.[168,169] However, for many years most biochemists regarded cAMP as a curiosity and the regulatory chemistry of phosphorylase as an unusual specialization. That view was altered drastically when cAMP was found to function as a second messenger in the action of over 20 different hormones. Phosphorylation by cAMP-dependent protein kinases regulates several enzymes concerned with energy metabolism or with the control of cell division. Phosphorylation reactions can lead to different responses in different specialized cells. Proteins of membranes, microtubules, and ribosomes are phosphorylated, as are nuclear histones. Cyclic AMP mediates the action of some neurotransmitters released at synapses and functions as a signal between cells of some slime molds (Box 11-C). Transcription of genes can be either stimulated or inhibited by cAMP.[170] Even in *E. coli* cAMP is generated and acts as a positive effector for transcription of some genes (Chapter 28). Cyclic AMP may function in signaling *stress* in plants.[171]



(11-8)

The isolation and characterization of adenylate cyclases have been difficult because of their location within membranes and because there is so little (usually only 0.001– 0.01%) of the membrane protein.[172] The 1060- to 1250-residue, 120 kDa proteins are also easily denatured. However, the sequences of cloned adenylate cyclase genes and the observed patterns of synthesis of the corresponding mRNAs have been revealing. There are at least eight mammalian adenylate cyclase genes with complex regulatory properties.[173 – 177] The sequences suggest that most of the isoenzymes are integral membrane proteins, *each with 12 transmembrane helices* organized as two sets of six with a large ~40-kDa cytoplasmic domain between the sets. This is similar to the organization of the cystic fibrosis transmembrane conductance regulator (Box 26-A) and some other membrane transporters. However, there is no firm evidence that adenylate cyclases contain ion channels. These enzymes are also discussed in Chapter 12, Section D,9.

A quantitative indication of the importance of the cAMP system within cells can be derived from measurement of the kinetics of the incorporation of ^{18}O from water into the α-phospho groups of AMP, ADP, and ATP. This incorporation will result from hydrolysis of cAMP by the phosphodiesterases that allow relaxation to a low cAMP level (Eq. 11-8). It is thought that in human blood platelets this represents the major pathway of this labeling (Eq. 11-9), which occurs at a rate of about 1.1 µmol of ^{18}O kg^{-1}s^{-1}.



(11-9)

Stimulation of the platelets by prostacyclin (prostaglandin I_2; see Chapter 22) leads to a 10- to 40- fold increase in cAMP concentration and a 4- to 5-fold increase in the rate of ^{18}O incorporation.[178] A quite different *soluble* adenylate cyclase is present in spermatozoa. It is stimulated directly by bicarbonate ions. It may be a **bicarbonate sensor** in sperm cells and in some other bicarbonate-responsive tissues as well as in cyanobacteria.[178a,178b]

3. Guanine Nucleotide-Binding Proteins (G Proteins)

The β adrenergic receptors are not coupled directly to adenylate cyclase but interact through an intermediary stimulatory protein **G_s**, which contains three subunits, α, β, and γ.[179–182] We know that the G_s protein associated with β adrenergic stimulation is only one of a very large number of related **G proteins**, so named because of their property of binding and hydrolyzing GTP. In its unactivated state the α subunit of a $G_{\alpha\beta\gamma}$ heterotrimer carries a molecule of bound GDP. Apparently, the G_s proteins and the hormone receptors, which are activated by the binding of hormones, diffuse within the membrane until they make contact and form molecules of the G_s•GDP complex (Eq. 11-10, step *b*). The complex then undergoes a rapid exchange of the bound GDP for GTP, after which the hormone and receptor dissociate. The G_α•GTP complex may also dissociate from the complex, perhaps entering the cytosol as a soluble protein (Eq. 11-10, step *c*). The G_α•GTP complex combines with adenylate cyclase (step *d*) and activates it to generate cAMP. However, the activation is transient. G_α also contains GTP-hydrolyzing (GTPase) activity and within a few minutes G_α•GTP is completely converted to G_α•GDP and the adenylate cyclase dissociates (Eq. 11-10, step *e*). The G_α•GDP recombines with the βγ complex which may serve as a membrane anchor, to complete the regulatory cycle. The overall effect is for hormone binding to cause a rapid release of cAMP in a short burst that may last only about 15 s.

It is not the binding of GTP but a slow subsequent conformational alteration that activates the G_α•GTP adenylate cyclase complex. This was deduced by study of analogs of GTP, such as guanosine 5'-(β, γ-imido) triphosphate (GMP-*P*-(NH)-*P* or GppNp), which are

BOX 11-C THE ATTRACTION OF *DICTYOSTELIUM* TO CYCLIC AMP

In higher animals cyclic AMP (cAMP) is an intracellular second messenger, but in the cellular slime mold *Dictyostelium discoideum* it serves to convey signals *between* cells.[a–c] As was mentioned on p. 20, the organism exists as individual amebas until the food supply is exhausted. Then the cells begin to signal their lack of food by secreting pulses of cAMP. Because they also secrete a phosphodiesterase, the cAMP is short-lived.[d] However, it is present long enough for any other nearby ameba to sense the gradient of cAMP concentration from one end of the cell to the other. As little as a 2% difference in concentration can induce chemotaxis.[e] The amebas move toward the source of cAMP and emit pulses of the compound. This results in the formation of aggregation centers in which the concentration of cAMP oscillates spontaneously as the cAMP moves outward in waves. The cells move up the concentration gradient until the peak of a wave reaches them. Then they move in a random direction until the next wave reaches them and again orients their motion.

After ~50 movement steps an aggregation center contains ~10^5 cells which now follow a "development program." The amebas adhere to each other to form 1- to 2 mm-long multicellular "slugs" in which all of the cells move forward together. About 30 h after aggregation begins the slugs stop and form stalks with spore-containing fruiting bodies on top. The lead cells in a slug become the stalk.[a,c]

Two types of G protein-coupled cAMP receptors in the cell membranes of the amebas have been identified. One leads to activation of adenylate cyclase and the release of new pulses of cAMP and the other to activation of guanylate cyclase.[e] This enzyme causes a rapid 7- to 10-fold increase in intracellular cGMP which plays an important role in controlling chemotaxis.

Some species of cellular slime molds use other chemical attractants (**acrasins**). For example, *D. minutum* secretes an analog of folic acid[f] and cells of *Polysphondylium violaceum* are attracted by the ethyl ester of *N*-propionyl-γ-L-glutamyl-L-ornithine-δ-lactam.[g]

---

[a] Gerisch, G. (1987) *Ann. Rev. Biochem.* **56**, 853–879

[b] Devreotes, P. (1989) *Science* **245**, 1054–1058

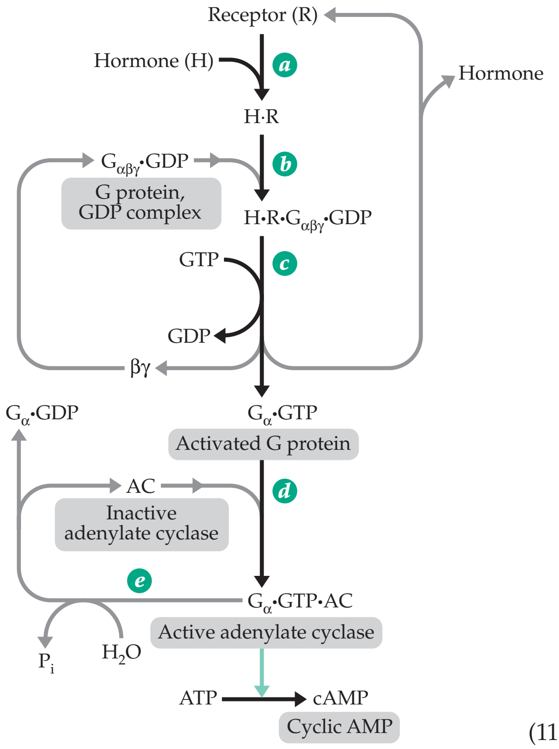
[c] Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) *Molecular Biology of the Cell*, 3rd ed., Garland, New York

[d] Levine, H., Aranson, I., Tsimring, L., and Truong, T. V. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6382–6386

[e] Kuwayama, H., and Van Haastert, P. J. M. (1996) *J. Biol. Chem.* **271**, 23718–23724

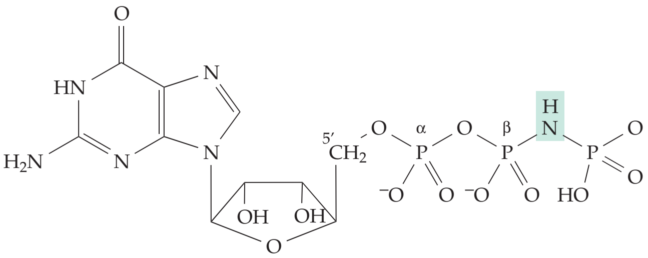
[f] Schapp, P., Konijn, T. M., and van Haastert, P. J. M. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2122–2126

[g] Shimomura, O., Suthers, H. L. B., and Bonner, J. T. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7376–7379



(11-10)

not hydrolyzed to GDP. These compounds may require many minutes to activate the complex, but because the GppNp is not hydrolyzed, the complex remains activated and able to catalyze cAMP formation continuously.



Guanosine 5'-(β,γ-imido) triphosphate (GppNp);
in place of the NH, O is present in GTP, S is present in GTPγS

Occupied receptors for adrenaline, glucagon, ACTH, and histamine activate adenylate cyclase via G_s proteins. Other G_s proteins, which contain subunits designated α_{olf} and which exist as a number of subtypes, mediate **olfactory responses**. Subunit α_o is another specialized polypeptide which is located primarily in neural tissues. A variety of additional G proteins have been discovered in organisms ranging from bacteria to mammals.[179,183–186] All have similar structures with 39- to 45-kDa α subunits, 35- to 36-kDa β subunits and 5- to 8-kDa γ subunits. Whereas the α subunits are unique to each G protein, β and γ subunits may be shared among several G proteins. These proteins appear to function with many kinds of hormone receptors and in nerve transmission, secretion, and endocytosis.[187] Mammals can form at least 23 different α_s subunits, which are coded for by 16 genes. Some of the forms arise by alternative splicing of mRNA transcribed from a single gene.[180] There are at least 6 β and 12 γ subunits.[179,180,182,188,189] Other G proteins, found in heart muscle, induce the opening of K^+ or Ca^{2+} channels in cell membranes.[185,190] It has often been assumed that all of the effects of G proteins are mediated by activated α subunits. However, the βγ complex, which sticks together very tightly, can also have signaling functions.[191] For example, an auxiliary protein called **phosducin** binds to the βγ complex of G_s, slowing the reconversion to the αβγ trimer.[192] Phosducin was discovered in rod cells of the retina but has been found to be distributed broadly in other tissues as well.[192,193] In yeast the βγ subunit of a G protein linked to the receptor for the yeast mating factor initiates the pheromone response of a cell.[193a]

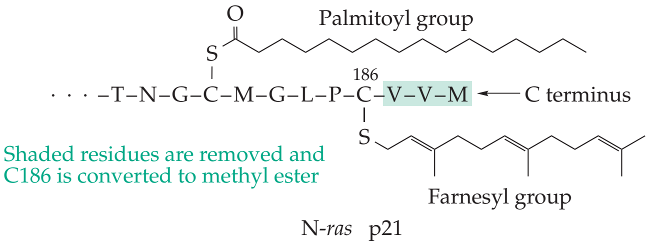
There is another large class of receptors whose occupancy by an agonist leads to *inhibition* of adenylate cyclase. These include the α_2 adrenergic receptors, receptors for acetylcholine, adenosine, prostaglandin E_2 (Chapter 21), somatostatin, and some receptors for dopamine. Their responses are mediated by inhibitory proteins G_i, which closely resemble G_s in their sizes, amino acid sequences, and heterotrimeric structures, but which inhibit adenylate cyclase when activated.[180] A clear distinction between the G_s and G_i proteins is evident in the fact that G_s is irreversibly activated by the action of cholera toxin, while G_i loses its ability to respond to occupied receptors when modified by the action of *Pertussis* toxin (Box 11-A). A specialized heterotrimeric G_i protein known as **transducin** mediates the light-induced activation of a **cyclic GMP phosphodiesterase** in the retina[194,195] (see Chapter 23). Its α subunit is designated α_t. The related **gustducin** is found in taste buds.[196]

Monomeric G proteins. An entirely different class of G proteins was discovered when it was found that the small 189-residue, 21-kDa protein products of the human oncogenes and proto-oncogenes known as **ras** are monomeric G proteins.[186,197–200] There are over 80 related proteins of this group of nine families.[200a] They include the much larger **elongation factor** EF-Tu, which functions in protein synthesis (Chapter 29). Even though there is a large difference in size, the three-dimensional structures of the GTP-binding domains of EF-Tu[201–204] and of the 21-kDa Ras p21 proteins[205–209] are very similar (Fig. 11-7). They also resemble that of adenylate kinase (Fig. 12-30).[205] The human genome contains four "true" *ras* genes: H-*ras*, N-*ras*, K-*ras*, and K-*rasB*. The occurrence of mutations in these genes in human cancer is discussed in Section H, as are the complex signaling functions of *ras* genes, (see also Fig. 11-13).[210]

Other proteins in the Ras superfamily include some members of the **Rho** family, which function in the cytoskeleton.[200a,211,212,212a,b,c,213] These include at least 14 mammalian proteins, among them RhoA, RhoB, etc., Rac1, Rac2, etc., and Cdc42. The latter is associated with the formation of fingerlike extensions of the cell membrane containing actin bundles.[200a,212d,212e] Genes *rac1* and *rac2* encode proteins involved in the initiation of superoxide radical formation by activated phagocytes (Chapter 18).[214] The **Rab** family of proteins are involved in exocytosis and endocytosis and vesicular transport. There are 24 or more human *rab* genes.[212,215,216] Yeast cells also contain Ras proteins. One *RAS* gene must be present for yeast to form spores.[217] **ADP-ribosylation factors** (ARFs), which stimulate the action of cholera toxin A subunit (Box 11-A) are also members of the Ras superfamily.[120,217a]

Like the trimeric G proteins, the monomeric proteins also serve as “timed switches” that are turned on by GDP–GTP exchange and are turned off by the hydrolysis of the bound GTP. In the absence of other regulatory proteins both the exchange of bound GDP for GTP and the hydrolysis of bound GTP are slow. While the receptor-associated trimeric G proteins are activated by the hormone receptors, the monomeric G proteins are activated toward GDP–GTP exchange by binding to proteins called **guanine nucleotide dissociation stimulators** (**GDSs**).[217b] They speed the release of GDP from the G protein allowing the active GTP complex to be formed. The velocity of hydrolysis of the bound GTP in the activated G protein is greatly increased by **GTPase activating proteins** (**GAPs**). These auxiliary proteins can be thought of as signaling molecules that pass their messages to the G protein, controlling the extent of its activation and therefore the strength of a signal that it sends on to the processes that it is controlling.[200,210,218] One group of GAP proteins are known as **regulators of G-protein signaling** (RGS).[218-218b] These proteins, which contain a characteristic 12-residue core, were first recognized as negative regulators of signaling in yeast cells called **GDP-dissociation inhibitors** (GDIs).[219] They can also be regarded as a family of GAPs.

Acylation and prenylation. The amino terminus (usually glycine) of the α subunit of any G protein is nearly always converted to an ***N*-myristoyl group**.[220–223] This modification occurs in a **cotranslational process** after removal of the initiating methionine (Chapter 29) and can be described as an acyl transfer from coenzyme A.[224] The C termini of the γ subunits of heterotrimeric G proteins, and also of the monomeric proteins of the Ras family, also undergo processing. For example, the C-terminal end of an intact Ras protein contains 18 residues which probably assume a largely α-helical conformation. A cysteine side chain near the terminus and having the sequence CAAX is converted within cells to a **thioether** with all-*trans*-farnesol or longer polyprenyl alcohols.[225] Another nearby cysteine may form a **thioester** with palmitic acid,[221,226,227] for example:



Shaded residues are removed and C186 is converted to methyl ester

N-*ras* p21

The farnesyl (or longer) polyprenyl and palmitoyl groups, together with nearby nonpolar side chains, serve as a membrane anchor. After the prenylation the C-terminal three amino acids (VVM in N-*ras* p21) are removed proteolytically and the new terminal carboxyl group of the farnesylated Cys 186 is converted to a methyl ester.[212c,225,228] N-terminal myristoylation is usually regarded as irreversible. However, the thioester linkages by which palmitoyl groups are attached to proteins are labile and may be cleaved rapidly by hydrolases. It follows that palmitoylation can be a rapid, regulated modification of proteins,[229–230a,b] strongly affecting the location and extent of adherance to membrane surfaces.

Three-dimensional structures. The structure of the GTP-binding domain of elongation factor EF-Tu was determined by Jurnak in 1985[201] and that of the complete three-domain structure later.[202,203] When the structure of the *catalytic domain* of the first Ras protein was determined (Fig. 11-7A) it was clear that it was similar to that of EF-Tu.[205,207] The same was true for the transducin α_t,[194,231,232] for the inhibitory $G_{i\alpha1}$,[218b,233,234] and for other G proteins.[235] In every case the differences in structure of the enzyme with GDP or with analogs of GTP were small and limited to a region close to the γ-phospho group of the bound GTP. This group can be seen clearly in Fig. 11-7A adjacent to residue 60 of the protein. See also Fig. 12-36.

The β strands and helices in the GTP-binding domain of the G proteins are connected by a series of loops. The first, second, and fourth loops (residues 10–17, 32–40, and 57–65, respectively) in the Ras structure of Fig. 11-7A form the catalytic site in which the hydrolysis to GDP takes place. The first loop, also called the P loop, is conserved in all GTP-binding proteins. In the Ras proteins it has the following sequence:

10 17	
G–A–G–G–V–G–K–S	Ras
G–X–X–X–X–G–K–S(T)	Consensus

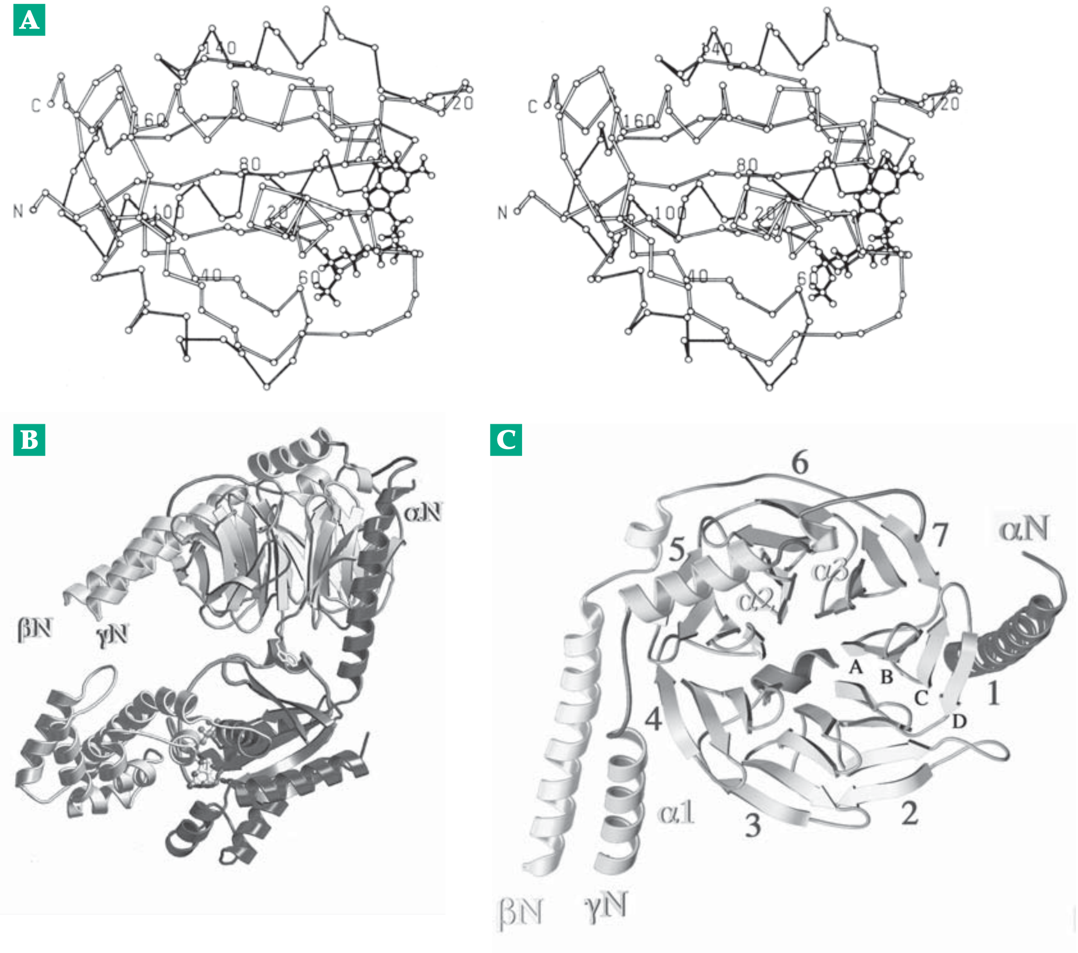
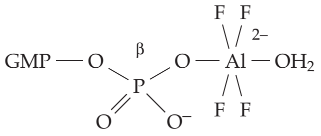


Figure 11-7 (A) Stereoscopic drawing showing α-carbon positions in the 166-residue "catalytic domain" of the human c-H-*ras* gene product and the bound GTP analog GppNp (solid bond lines). The intact protein contains an additional 23-residue C-terminal extension. From Pai *et al.*[241] (B), (C) Ribbon drawings showing two nearly orthogonal views of a hetero-trimeric inhibitory G_i brain protein produced using a cloned bovine gene. (B) The amino termini (N) of the three subunits are seen in the left-to-right order: γ, β, α. A side view of the β propeller domain of the β subunit is seen at top center. The Ras-like domain and the additional large helical domain of the α subunit are marked. (C) View from the flared end of the β propeller looking toward the α subunit. The strands of each propeller blade are labeled A, B, C, and D and the seven blades are numbered around the periphery of the propeller. From Wall *et al.*[242] (B) and (C) courtesy of Stephen Sprang.

Also shown here is the "consensus" sequence for all GTP-binding proteins. Cancer cells frequently have mutations in Ras at Gly12, which may be substituted by Asp, Lys, Val, or Arg, and at Gly 13 and Gln 61. The latter is in loop 4, a part of the sequence that appears especially mobile but is highly conserved. These mutations all activate the protein, i.e., they decrease the GTPase activity, allowing the G protein to remain in its active conformation longer than normal.

Ras proteins fulfill their functions by interacting closely with two or more proteins in signaling pathways as described in Section H. Other G proteins have additional domains. The 405-residue EF-Tu from *Thermus thermophilus* has three domains: the C-terminal nucleotide-binding domain and two β-barrel domains following it. A major difference in conformation is observed between forms of the protein with bound GTP or bound GDP, the GDP form being opened up by a hinging motion of ~ 90° between domains I and II.[202,203] The conformational change is triggered by the GDP–GTP exchange. An enhanced rate of GTP hydrolysis, which results from binding of a transfer RNA and binding to a ribosome, is apparently accomplished by interaction with loop 2 of the catalytic domain (corresponding to residues 32 – 40 in Fig. 11-7A). The function of this protein is considered further in Chapter 29.

Transducin has a 113-residue domain inserted into loop 2 of the catalytic domain. In this case, too, a large hinging movement is associated with the GDP–GTP exchange. The light-activated receptor rhodopsin induces the conformational change.[194,231] Again, structural changes resulting from the presence or absence of interactions of the protein with the γ-phospho group of GTP are involved. The GTPase activity of heterotrimeric G proteins can also be activated by aluminum fluoride AlF_4.[232,236] From the X-ray structure of the transducin α•GDP•AlF_4^- complex it is seen that hydrated AlF_4^- is covalently bonded in the position of the γ-phospho group and makes hydrogen bonds to active site groups, mimicking a possible transition state for the GTPase reaction[232] (see Chapter 12, Section D).



The β subunit of the βγ complex of transducin is a **seven-bladed β propeller** (Fig.11-7C). It is composed of seven GH–WD repeat units: $-[GH - X_n - WD]_{4-8}-$ where GH = Gly – His, WD = Trp–Asp, and X_n is a core repeating sequence, usually 32 –42 residues in length. This motif is also found in at least 40 other eukaryotic proteins.[188,237–240] In the βγ complex (Fig. 11-7B) the γ subunit assumes an elongated, largely α-helical structure. It is often anchored at its C terminus by a farnesyl or geranylgeranyl chain, while G_α may be myristoylated or palmitoylated.[195,230,243]

The three-dimensional structure of the GDP complex of the intact transducin heterotrimer[195] also shows a tight interaction between α and β subunits. The major interaction is probably disrupted by replacement of the bound GDP by GTP and the conformational change that occurs around the γ - phospho group. This explains the dissociation of the α subunit from βγ upon activation. An entirely similar picture has been obtained for the action of the inhibitory G protein, G_{i2}, for which structures of the α subunit and of the αβγ heterotrimer (Fig. 11-7,B,C) have been determined.[188,233,234,242] The structures resemble those of transducin, but differ in details.

4. Guanylate Cyclase (Guanylyl Cyclase), Nitric Oxide, and the Sensing of Light

Formation of the less abundant cyclic guanosine monophosphate (cGMP) is catalyzed by guanylate cyclases found in both soluble and particulate fractions of tissue homogenates.[244–247] However, its significance in metabolism is only now becoming well understood. There are cGMP-dependent protein kinases,[247] but, until recently cGMP could not be regarded as a second messenger for any mammalian hormone. Now we know that cGMP has several essential functions. It mediates the effects of the **atrial natriuretic factor**, a peptide hormone causing dilation of blood vessels (Box 23-D),[245,248–250] and also of a peptide called **guanylin**, which is formed in the intestinal epithelium.[245,251] The receptors for both of these peptides are *transmembrane proteins with cytoplasmic domains that have guanylate cyclase activity*. The released cGMP appears to induce relaxation of smooth muscles of the blood vessel walls. Another mediator of smooth muscle relaxation is the **endothelial cell-derived relaxing factor**, which is evidently **nitric oxide**, **NO**[252] (see Chapter 18). A soluble guanylate cyclase appears to be an NO receptor and is activated by binding of the NO to a heme group in one of the two subunits of the cyclase.[253 – 255] The cyclase can also be activated by carbon monoxide, CO, in a similar fashion.[256,257] Carbon monoxide is produced in the body by degradation of heme (Fig. 24-24) and is thought to be a neurohormone.[258]

Cyclic GMP also plays an important role in vision. Specific phosphodiesterases hydrolyze cGMP to GMP. The cGMP phosphodiesterases of the rod and cone cells of the retina are activated by the G protein transducin, which has been activated by light absorbed by the visual pigments.[194] The resulting decrease in the cGMP concentration is thought to cause the closing of cation channels and thereby to initiate a nerve impulse (see Chapter 23). Cyclic nucleotide forms of the pyrimidine nucleotides are present only in very small amounts in cells.

5. Bacterial Chemoreceptors

Bacteria are attracted to foods with the aid of chemoreceptors that bind certain amino acids such as aspartate. The receptors send signals to the mechanisms that ensure that the bacterium is swimming toward the food[113,259, 259a] (more details are given in Chapter 19). The aspartate and serine chemoreceptors found in membranes of *E. coli* and *Salmonella typhimurium* are among the most carefully studied of all receptors. Like the seven-helix receptors, they have an extracellular sensory domain to which the signaling molecule, aspartate or serine, binds, and a transmembrane helical bundle. In these receptors the transmembrane part consists of four helical segments coming from the two subunits of the dimeric protein. For the aspartate receptor there is a high-resolution X-ray structure for the cytoplasmic domain[260,261] including the bound aspartate[262] as is shown in Fig. 11-8A,B.[260,263] The structures of the receptors for aspartate, serine, and many other signaling molecules are evidently very similar. The structure of the cytosolic domain of the serine receptor has been determined and a model for the complete receptor has been proposed.[259a] The molecular mass of the 188-residue extracellular domain is ~18 kDa, while that of the larger cytoplasmic domains is ~36 kDa. The latter includes a linker region as well as a long four-helical bundle domain which can be divided into a methylation domain and a signaling domain as shown in (Fig. 11-8C).

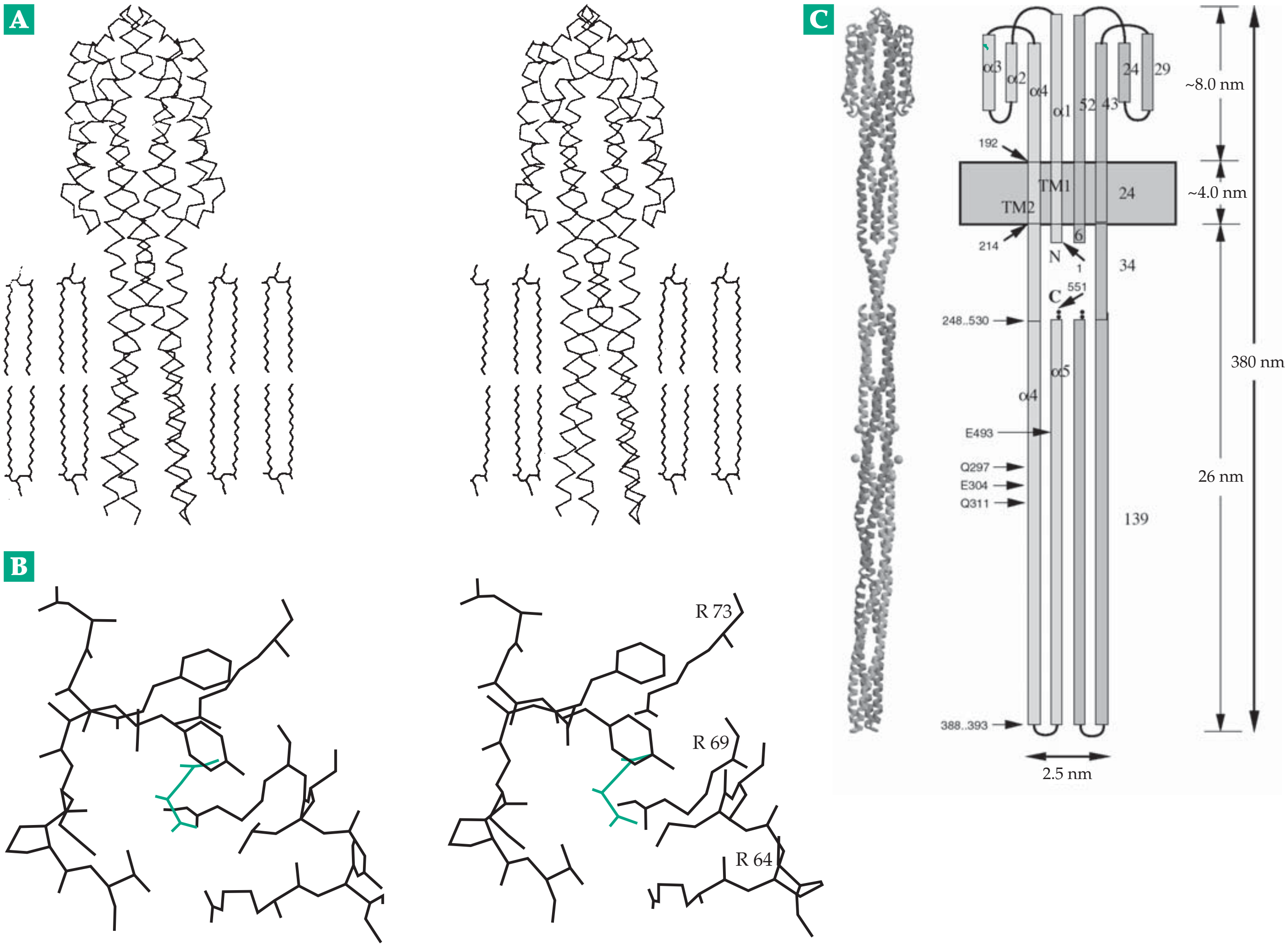


Figure 11-8 (A) Stereoscopic view of the sensory domain of the dimer with modeled transmembrane region. Each monomer contains a four α-helix bundle with two helices continuing through the membrane. From Scott *et al.*[260] (B) Stereoscopic image of aspartate in the major binding site of the receptor. The atomic model of aspartate (green) has been fitted into the observed difference electron density map. From Yeh *et al.*[262] (C) Model of an intact *E. coli* serine chemoreceptor. Left. Ribbon drawing viewed perpendicular to the molecular twofold axis. Methylation sites are represented by the dark balls by the cytoplasmic domain. The bound serine is drawn as a partially hidden green ball at the upper left in the extracellular domain. Right. Diagram of the receptor. The presumed membrane bilayer is represented by the gray band. Positions of some residues are marked on the left side. On the right, the numbers of residues in various peptide segments are indicated. From Kim *et al.*[259a] Courtesy of Sung-Hou Kim.

What happens chemically when aspartate binds into a deep pocket in the center of the sensory domain? What kind of signal can be passed from the aspartate to the "signaling domain"? Several possibilities were discussed by Kim,[264] who suggested that the binding of aspartate causes the rotation of the sensory domain of one subunit relative to that of the other and that this rotation is transmitted through the membrane to the signaling domain. The signaling domain of the receptor forms a complex with a multimeric **protein histidine kinase** called **CheA** (chemotaxis protein A) together with an auxiliary protein **CheW**. A more likely possibility is that some kind of "piston" action occurs between the transmembrane helices of the two subunits.[264a] Binding of aspartate to the receptor, and the associated alterations in the receptor•CheA•CheW complex, causes a strong *decrease* in the catalytic activity of CheA.

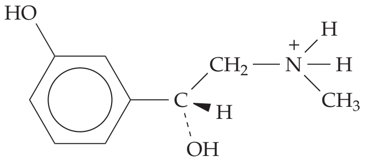
In its active form CheA undergoes **autophosphorylation**, that is, the phosphorylation of a histidine imidazole group in one of its subunits by the protein kinase active site of an adjacent subunit. The phospho group is then transferred from phospho-CheA to another protein, **CheY**. Phospho-CheY interacts with the flagellar motor proteins (Chapter 19) periodically causing a reversal of direction of the bacterial flagella. As a result the bacteria tumble and then usually move in a new direction. If the attractant molecule aspartate is present on the receptor the signal is weakened and the bacteria continue to swim in the same direction–toward food. Phospho-CheY is spontaneously hydrolyzed to remove the phospho group, but the dephosphorylation is promoted by an additional protein **CheZ**. At least two other auxiliary proteins also function in this control system. **CheR** is a **methyltransferase** that methylates the glutamate side chain carboxylates in the methylation domain of the coiled coil region and **CheB** is a **methylesterase** that removes the methyl groups. These two proteins control the methylation level in the coiled coil region of the receptor (Fig. 11-8A). Increased methylation increases the strength of the signal sent from the receptor, perhaps because methylation removes negative carboxylate charges that repell each other and weaken the structure, interfering with signaling.[259] See also Fig. 19-3.

E. Calcium, Inositol Polyphosphates, and Diacylglycerols

Calcium ions entering cells from the outside or released from internal stores trigger many biological responses (see Box 6-D). Within cells Ca^{2+} often accumulates in mitochondria, in the ER, or in vesicles called **calciosomes**.[265] Release of the stored Ca^{2+} is induced by hormones or by nerve impulses. For example, impulses flow from the nerve endings into the muscle fibers and along the invaginations of the plasma membrane called transverse tubules (Chapter 19). There they induce release of Ca^{2+} from the ER. The released ions activate enzymes[266] and induce contraction of the muscle fibers. In many cells, Ca^{2+} causes release of secreted materials, for example, neurotransmitters in the brain.[267,268]

1. Alpha Adrenergic Receptors

The release of stored Ca^{2+} is often triggered by α adrenergic receptors.[269] Like β adrenergic receptors, the α receptors are activated by adrenaline. Specific inhibitors distinguish them. For example, the β receptors are inhibited by propranolol, while the α receptors are blocked by phenoxybenzamine. The synthetic agonist **phenylephrine** activates only α receptors and no increases in cAMP or in protein kinase activity are observed.


Phenylephrine

There are two major α adrenergic receptor subtypes.[151,269] Activation of the α_2 receptors, which are present in various tissues including blood platelets, causes *inhibition* of adenylate cyclase. This inhibition is evidently mediated by the G_i protein considered in Section D,3. The nucleotide sequences of cloned α_2 receptor genes and other properties suggest close structural similarity to the β receptors.[270–273] Subtle differences in hydrogen bonding to the serine side chains shown in Fig. 11-6C may distinguish α_2 from β_2 receptors.[273] A characteristic of α_2 receptors is that pertussis toxin (Box 11-A) abolishes the inhibition of adenylate cyclase, which they mediate. In contrast, the action of α_1 adrenergic receptors is not affected by pertussis toxin.

The α_1 receptors are activated not only by catecholamines but also by the hormones **vasopressin** and **angiotensin II**. Binding of these hormones to α_1 receptors induces a complex response that involves rapid hydrolysis of **phosphatidylinositol** derivatives and release of Ca^{2+} into the cytoplasm, and of diacylglycerols into the lipid bilayer of the membrane. The response is mediated by another G protein called G_q.[274,275] When this G protein is activated it induces the hydrolysis of **phosphatidylinositol 4,5-bisphosphate** (PtdInsP_2), a normal minor component of the lipid bilayer, by **phospholipase C** (phosphoinositidase C).[265,276–281]

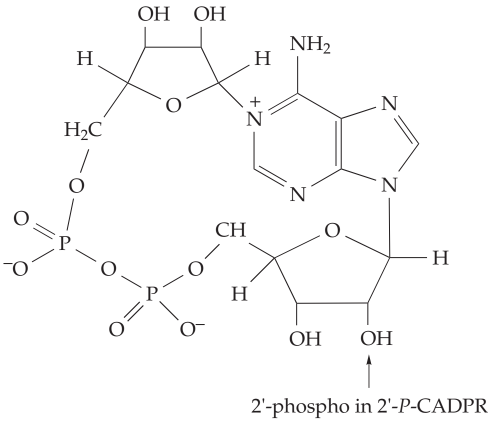
$$\text{Receptor} \rightarrow G_q \rightarrow \text{phospholipase C} \rightarrow \left\{\begin{array}{l}\text{inositol phosphates}\\ + \text{ diacylglycerol (DAG)} \rightarrow \text{increased } [Ca^{2+}]\end{array}\right. \quad (11\text{-}11)$$

The products, **inositol 1,4,5-trisphosphate** [abbreviated InsP_3 or Ins(1,4,5)P_3], and **diacylglycerol** (DAG) are both regarded as second messengers for the catecholamines acting on α_1 receptors and for about 20 other hormones, neurotransmitters, and growth factors upon binding to their specific receptors.[269,282] In addition to vasopressin, the gonadotropin-releasing hormones, histamine, thrombin (upon binding to platelet surfaces), and acetylcholine (upon binding to its "muscarinic" receptors; Chapter 30) all stimulate inositol phosphate release.

2. Phosphatidylinositol and the Release of Calcium Ions

The phosphoinositides constitute ~2–8% of the lipid of eukaryotic cell membranes but are metabolized more rapidly than are other lipids.[265,278,279,283–285] A simplified picture of this metabolism is presented in Fig. 11-9. Phosphatidylinositol is converted by the consecutive action of two kinases into phosphatidylinositol 4,5-bisphosphate.[286,287] The InsP_3 released from this precursor molecule by receptor-stimulated phospholipase C is thought to mobilize calcium ions by opening "gates" of calcium channels in membranes of the ER or of calciosomes.[282,288–289a] It diffuses across the peripheral cytoplasm to **InsP_3 receptors** which are embedded in the membranes of the ER.[282,290] One of the several isotypes of InsP_3 receptors is a 2749-residue protein thought to contain a calcium ion channel.[290] Similar receptors are also found in inner membranes of the nucleus.[291,292]

Several uncertainties have complicated our understanding of the role of Ca^{2+} in signaling. What is the source of Ca^{2+}? How much of it enters cells from the outside and how much is released from internal stores? Where are the internal stores? What other kinds of ion channels are present and what second messengers regulate them? The sarcoplasmic reticulum of skeletal muscle and also membranes in many other cells contain **ryanodine receptors** as well as InsP_3 receptors.[282,293] Both of these receptors have similar structures and contain Ca^{2+} channels. However, the ryanodine receptors are activated by **cyclic ADP ribose** (cADPR),[294,295] which was first discovered as a compound inducing the release of Ca^{2+} in sea urchin eggs.[296] The 2-phospho derivative of cADPR may also have a similar function.[297]



Cyclic ADP-ribose (cADPR)

Phospholipase C, which initiates the release of phosphatidylinositol derivatives, also requires Ca^{2+} for activity. It is difficult to determine whether release of Ca^{2+} is a primary or secondary response. There are three isoenzyme types of phospholipase C−β, γ, and δ− and several subforms of each with a variety of regulatory mechanisms.[298–300a] For example, the γ isoenzymes are activated by binding to the tyrosine kinase domain of receptors such as that for epidermal growth factor (see Fig. 11-13). In contrast, the β forms are often activated by inhibitory G_i proteins and also by G_q, which is specific for inositol phosphate release.

Calcium ions are usually released in distinct pulses or "quanta." The kinetic characteristics of the system of receptors, diffusing InsP_3, calcium buffers, and calcium pumps in the cell membrane, and the membrane potential may account for this behavior.[174,301–306] Since the phosphoinositols are chelators of Ca^{2+}, the equilibria involved in this control system are complex.[307]

In addition to InsP_3, several other inositol derivatives are released by adrenergic stimulation. **Inositol 1,3,4,5-tetrakisphosphate** (InsP_4) is formed from InsP_3 by action of a soluble kinase.[308,309,309a,b] A controversial suggestion is that InsP_4 may induce opening of Ca^{2+} channels through the outer membrane of the cell[310–312] and may also function to promote storage of Ca^{2+}. It also acts as a transcriptional regulator. Other possible second messengers are inositol 1,2-cyclic-3,4-trisphosphate and related metabolites that arise by the action of phospholipase C on PtdIns 4-P followed by additional actions of kinases and phosphatases.[278,309,313–315] A different inositol tetrakisphosphate, Ins(3,4,5,6)P_4, may control chloride ion channels.[316,317] Both inositol 1,3,4,5,6-pentakisphosphate (**InsP_5**) and inositol hexakisphosphate (**InsP_6**) are also found in plants[318] and animals.[319] InsP_5 serves as an allosteric effector regulating hemoglobin in avian erythrocytes (Chapter 7). InsP_6, also known as **phytate**, is present in large amounts in cereals and recently has been found to play a role in regulating the export of mRNA from the nucleus.[319a] Additional phospho groups can be added to InsP_6 to form pyrophosphates[320,320a] and other complex polyphosphates.[321] Both InsP_5 and InsP_6 can also serve as precursors to Ins(1,3,4,5)P_4 and Ins(1,4,5)P_3 by hydrolytic dephosphorylation.[322]

Following stimulation of a cell the induced metabolism of phosphoinositides decays rapidly. The diacylglycerols are converted into phosphatidic acid and resynthesized into phospholipids (Chapter 21). The InsP_4, InsP_3, and other inositol metabolites are hydrolyzed by phosphatases.[278,309,323–327] Two of these phosphatases are inhibited by Li^+ as indicated in Fig. 11-9. They may represent one site of action of **lithium ions** in the brain.[324,327,328] Lithium salts are one of the most important drugs for treatment of **bipolar (manic-depressive) illness**. By blocking the release of free inositol, which can be resynthesized into PtdInsP_2, Li^+ may prevent neuronal receptors from becoming too active. However, the basis for the therapeutic effect of lithium ions remains uncertain.

The diacylglycerols released by phospholipase C diffuse laterally through the bilayer and, together with the incoming Ca^{2+}, activate **protein kinases C**. These kinases also require **phosphatidylserine** for their activity and phosphorylate serine and threonine side chains in a variety of proteins.[329–330b] They are stimulated by the released unsaturated diacylglycerols. In addition protein kinases C can be activated by **phorbol esters**, which are the best known tumor promoters (Box 11-D). The diacylglycerol requirement favors a function for these protein kinases in membranes. They also appear to cooperate with calmodulin to activate the Ca^{2+}-dependent contraction of smooth muscle.[330]

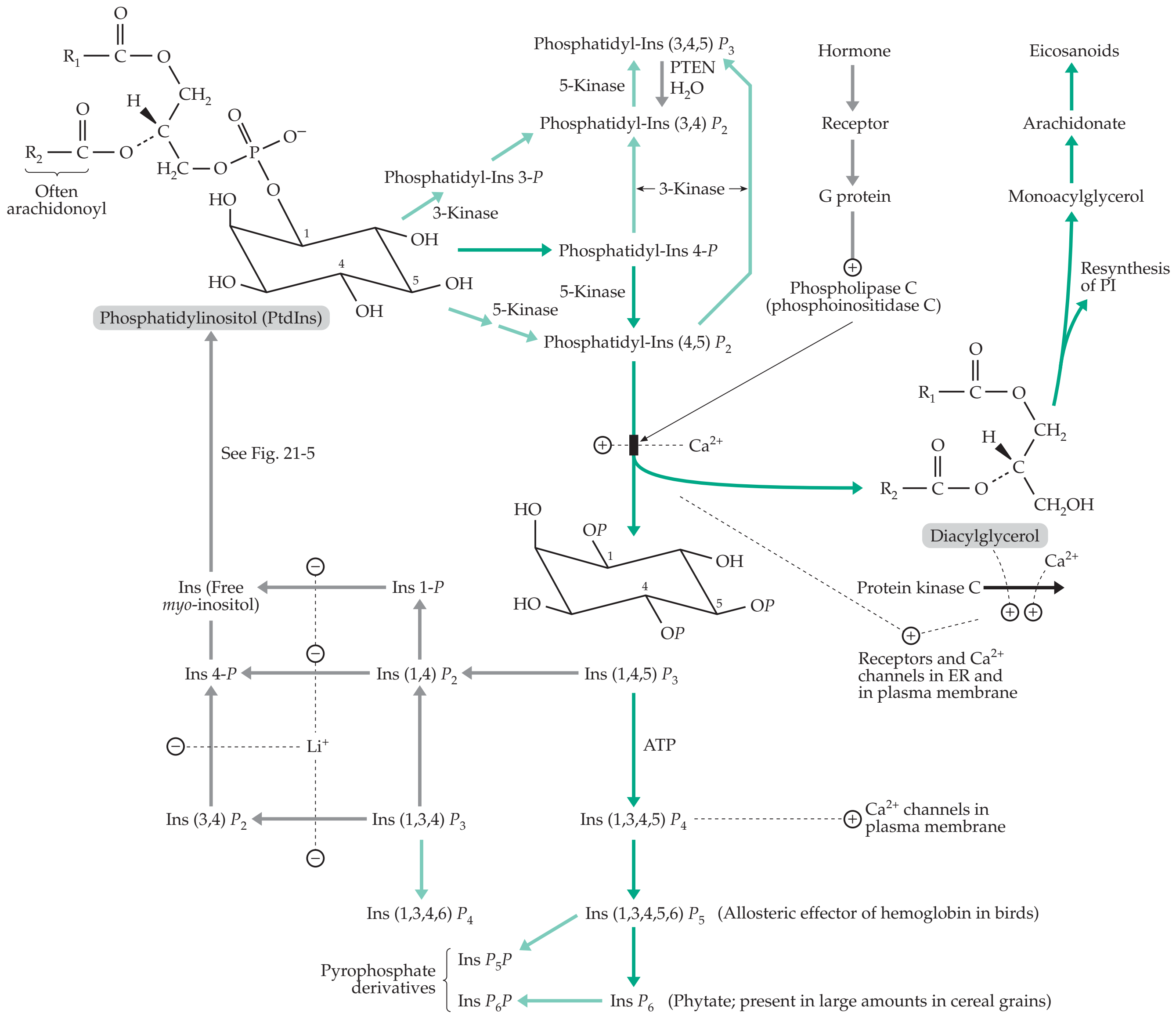
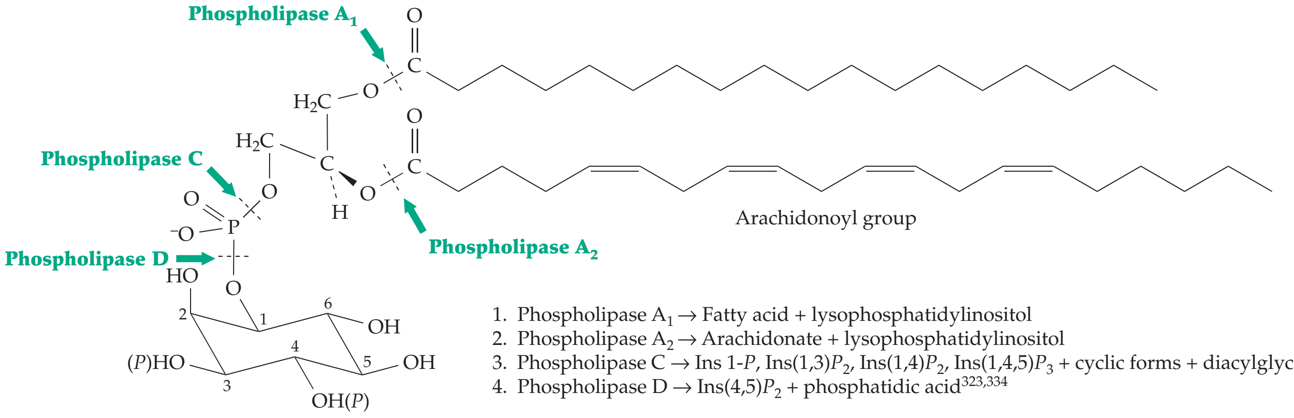

Often arachidonoyl
Phosphatidylinositol (PtdIns)
Phosphatidyl-Ins (3,4,5) P_3
5-Kinase
PTEN
H_2O
Phosphatidyl-Ins (3,4) P_2
Phosphatidyl-Ins 3-P
3-Kinase
3-Kinase
Phosphatidyl-Ins 4-P
5-Kinase
5-Kinase
Phosphatidyl-Ins (4,5) P_2
Hormone
Receptor
G protein
Phospholipase C (phosphoinositidase C)
Eicosanoids
Arachidonate
Monoacylglycerol
Resynthesis of PI
See Fig. 21-5
Ca^{2+}
Diacylglycerol
Ca^{2+}
Protein kinase C
Receptors and Ca^{2+} channels in ER and in plasma membrane
Ins (Free myo-inositol)
Ins 1-P
Ins 4-P
Ins (1,4) P_2
Ins (1,4,5) P_3
Li^+
ATP
Ins (3,4) P_2
Ins (1,3,4) P_3
Ins (1,3,4,5) P_4
Ca^{2+} channels in plasma membrane
Ins (1,3,4,6) P_4
Ins (1,3,4,5,6) P_5 (Allosteric effector of hemoglobin in birds)
Pyrophosphate derivatives
Ins P_5P
Ins P_6P
Ins P_6 (Phytate; present in large amounts in cereal grains)

Figure 11-9 Scheme showing synthesis and release of diacylglycerol and inositol phosphates and their regulation of calcium concentration in response to hormonal stimulation.

Diacylglycerols released by phospholipase C usually contain **arachidonic acid** in the 2 position. Hydrolytic cleavage of this linkage[284,331] is a major source of arachidonate for synthesis of **eicosanoids** such as the **prostaglandins** whose functions are discussed in Chapter 21. Arachidonate can also be formed directly by the action of phospholipase A_2 on membrane phospholipids[332,332a] (see the structure on p. 566). It has been suggested that protein G_q, which activates the hydrolysis of phosphoinositides, may also regulate phospholipase A_2 directly. Released diacylglycerols may not only activate protein kinases C but also have a direct role in promoting the membrane fusion required in exocytosis and endocytosis.[332] Other breakdown pathways of phosphatidylinositols are also indicated in the following structure.[298,333]

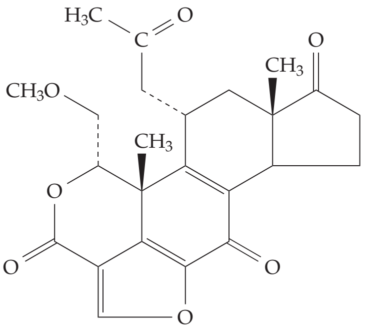
Initially most attention was paid to the water-soluble inositol phosphates that are released from phosphoinositides. However, phosphoinositol derivatives that retain the diacylglycerol part of the molecule have regulatory functions while remaining in membranes. A phosphatidylinositol 3-kinase phosphorylates the 3-OH of inositol in PtdIns, PtdIns 4-P and PtdIns (4,5)P_2 to give PtdIns 3-P, PtdIns(3,4)P_2, and PtdIns(3,4,5)P_2, respectively.[335–337d] However, they are ideally suited to function as spatially restricted membrane signals.[337a] They affect protein kinase C as well as the Ser / Thr protein kinase Akt (discussed in Section F,3),[326b] and have several functions in vesicular membrane transport and in regulation of the cytoskeleton.[337e,f]



1. Phospholipase $A_1 \rightarrow$ Fatty acid + lysophosphatidylinositol
2. Phospholipase $A_2 \rightarrow$ Arachidonate + lysophosphatidylinositol
3. Phospholipase C $\rightarrow$ Ins 1-P, Ins(1,3)P_2, Ins(1,4)P_2, Ins(1,4,5)P_3 + cyclic forms + diacylglycerol
4. Phospholipase D $\rightarrow$ Ins(4,5)P_2 + phosphatidic acid[323,334]

Phosphatidylinositol

Characteristic of many of the phosphoinositide-regulated proteins is a proline-rich PH domain (Table 7-3) which can transmit regulatory signals to additional proteins in a cascade. Their importance is emphasized by the finding that a 405-residue phosphatase **PTEN** (named after its gene symbol) which catalyses hydrolytic removal of one phospho group from PtdIns(3,4,5) P_3 to form PtdIns(3,4) P_2, (Fig. 11-9) is a major human tumor suppressor (see Box 11-D).[338,338a,b] These lipids are also important in insulin action (Section G,3). The steroid-like fungal metabolite wortmannin is a specific inhibitor of the 3-kinase.[339,339a]



Wortmannin

F. Regulatory Cascades

The effect of a regulated change in the activity of an enzyme is often amplified through a cascade mechanism. The first enzyme acts on a second enzyme, the second on a third, etc. The effect is to rapidly create a large amount of the active form of the last enzyme in the series.

We have already considered regulatory cascades initiated respectively by the β and α_2 adrenergic receptors. The effect of these cascades on glycogen phosphorylase is outlined on the left side of Fig. 11-4. One branch of the cascade sequence begins with release of adrenaline under control of the autonomic nervous system. In muscle, binding of this hormone to the β receptors on the cell membrane releases cAMP which activates a protein kinase. The kinase then phosphorylates phosphorylase kinase. At this point the muscles are prepared for the rapid breakdown of glycogen. However, an additional initiating signal is the release of Ca^{2+} into the cytoplasm in response to impulses to specific muscles via the motor neurons. Calcium ions can also be released in the liver by α adrenergic stimulation. Phosphorylase kinase is activated by the calcium ions, and in their presence it converts inactive phosphorylase *b* to the active phosphorylase *a*. Both protein kinases and phosphorylase kinase also act on glycogen synthase, phosphorylating it and converting it to an inactive form. This turns off the biosynthetic pathway at the same time that glycogenolysis (glycogen breakdown) is turned on. Spontaneous reversion of the enzymes to their resting states occurs through the action of phosphatases that cut off the phospho groups placed on the protein by the kinases. Also essential are phosphodiesterases, which destroy the cAMP, and the calcium ion pump, which reduces the concentration of the activating calcium ion to a low level.

Elaborate cascades initiate the clotting of blood (Chapter 12) and the action of the protective complement system (Chapter 31). Cascades considered later in the book are involved in controlling transcription (Fig. 11-13) and in the regulation of mammalian pyruvate dehydrogenase (Eq. 17-9), 3-hydroxy-3-methylglutaryl-CoA reductase and eicosanoids (Chapter 21), and glutamine synthetase (Chapter 24).

1. Advantages of Regulatory Cascades

Computer simulations as well as studies of experimental models have led to the following conclusions.[340,341] Even simple cascade mechanisms, such as the one shown in Fig. 11-10, can provide a more flexible response to allosteric effectors (such as e_1 in the figure) than if the effector acted directly on the enzyme rather than on the protein kinase. The cascade also provides **amplification**. This is especially true if additional cycles are added. A response can result from the binding of only a small number of hormone molecules to a receptor in a cell membrane or from activation of only a few molecules of a protease in the initiation of blood clotting. A striking amplification occurs in visual responses. Under appropriate conditions a single quantum of light falling on a receptor cell in the retina of an eye can initiate a nerve impulse (Chapter 23). The latter requires the flow of a large number of Na^+ ions across the plasma membrane. It would be hard to imagine how absorption of one quantum could initiate a photochemical reaction leading to that much sodium transport without intermediate amplification stages. Another advantage of cascades is that they may provide **ultrasensitive responses**. Not only can a response be sensitive to a higher power than the first of the concentration of a signaling molecule but also the amplification provided by the cascade confers a high sensitivity to the response.[342]

Cascade systems also provide for response to more than one allosteric stimulus in a single pathway. Thus, as shown in Fig. 11-4, glycogen catabolism can be initiated in more than one way. Two pathways are known for initiation of both blood clotting and activation of the complement system. Many pathways activate the MAP kinase pathway shown in Fig. 11-13.

2. Substrate Cycles

Although cascade systems offer advantage to cells, there is a distinct energy expenditure associated with the controls.[343] As can be seen in Fig. 11-10, in addition to providing for turning on and turning off the regulated enzyme, the kinase and the phosphatase together catalyze hydrolysis of ATP to ADP and inorganic phosphate. Thus, the regulated enzyme is continually cycling between active and inactive forms. The relative amounts of each are determined by the amounts of the effectors e_1 and e_2, the concentrations of modifying enzymes, and the kinetic constants. Such cycles have sometimes been called "futile cycles" because they seemingly waste ATP. This is particularly true of cycles that involve major metabolites, sometimes called **substrate cycles**. An example is the conversion of fructose 6-*P* to fructose 1,6-bisphosphate by phosphofructokinase and hydrolysis of the bisphosphate back to fructose 6-*P* by a phosphatase (Fig. 11-2). However, *the cycles are not futile* because the hydrolysis of the ATP provides the energy required to maintain the concentrations of modified (phosphorylated) enzyme or enzymes at steady-state levels that are required for efficient catalysis.[341,344] The energy utilized in this way in regulation of enzymes is often small.[345] Substrate cycles are not only an unavoidable consequence of the need to regulate enzymes but also provide for improved sensitivity in metabolic control. Consider Fig. 11-2 again. If the enzymes are set for conversion of fructose 1,6-P_2 to glycogen (backwards direction), the flux of materials flowing in the forward direction (glycolysis) will be efficiently curtailed because any fructose 1,6-P_2 formed will be hydrolyzed rapidly by the active phosphatase. Under these circumstances the flux in the forward direction will be ultrasensitive to the activation of phosphofructokinase and to the inhibition of fructose 1,6-bisphosphatase.[346,347]

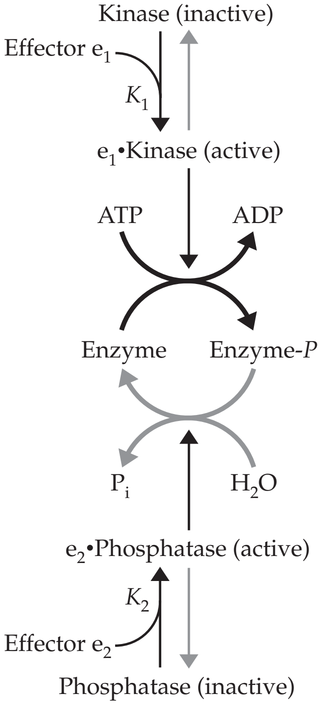


Figure 11-10 A "monocyclic" regulatory cascade involving phosphorylation and dephosphorylation of an enzyme. After Stadtman and Chock.[341]

G. Insulin and Related Growth-Regulating Hormones

Since its isolation in 1921, insulin has been the object of an enormous amount of experimentation aimed at clarifying its mode of action. It is produced by the β cells of the pancreatic islets of Langerhans and released into the bloodstream in response to elevated glucose levels. The absence of insulin or of a normal response to insulin results in the condition of **diabetes mellitus**, which is the most prevalent human metabolic disorder (see Box 17-G).[348]

Insulin or insulin-like material is also produced in ciliated protozoa, vertebrate and invertebrate animals, fungi, green plants,[349,350] and even in *E. coli*.[351]

1. Metabolic Effects of Insulin

Insulin has many effects on metabolism.[348,352,353] (Some of these are listed in Table 17-3). They can be summarized by saying that: (1) In most tissues insulin stimulates synthesis of proteins, glycogen, and lipids. (2) It affects the permeability of membranes, promoting the uptake and utilization of glucose and amino acids and of various ions from the blood. (3) It promotes both synthesis of glycogen and the breakdown of glucose by glycolysis. At the same time, it inhibits synthesis of glucose from amino acids by the gluconeogenesis pathway. The effects are not uniformly the same in all tissues. Some can be observed within a few minutes after administration, presumably as a result of regulation of enzymes. Effects on mRNA metabolism, protein synthesis, and cell growth are seen at later times.

The uptake of glucose by brain, liver, kidneys, erythrocytes, and the islets of Langerhans is unaffected by insulin. However, in muscle and adipose tissues insulin stimulates glucose uptake. Part of this effect results from insulin-induced translocation of molecules of the 509-residue glucose transport protein GLUT4 (Chapter 8) from the cytosol into the plasma membrane where it can function.[354–356a] Insulin apparently also increases the rate of synthesis of the transporters.

Insulin stimulates the phosphorylation of serine side chains of many proteins, including ATP citratelyase, acetyl-CoA carboxylase, and ribosomal subunit S6. At the same time it stimulates the dephosphorylation of other proteins, including acetyl-CoA carboxylase, glycogen synthase in skeletal muscle (Fig. 11-4),[49] pyruvate dehydrogenase, and hormone-sensitive lipase in adipose tissue. Yet another effect of insulin is to alter the amounts of specific messenger RNA molecules. For example, the transcription of the gene for phosphoenolpyruvate carboxykinase (PEPCK; Eq. 13-46) a key enzyme in gluconeogenesis, is inhibited by insulin within seconds after binding. The hexokinase isoenzyme called **glucokinase** phosphorylates glucose to glucose-6-*P* in liver and in pancreatic β cells. Its synthesis in liver is induced by insulin. (However, in the β cells the synthesis of glucokinase is induced by glucose.)

2. Insulin Receptors

All of the effects of insulin appear to result from its binding to insulin receptors, of which ~10^2 to 10^5 are present in the plasma membranes of most animal cells. First isolated in 1972,[357] insulin receptors and their cloned genes have been studied intensively. The receptors are $\alpha_2\beta_2$ disulfide crosslinked oligomers composed of pairs of identical 120- to 135-kDa α subunits and 95-kDa β subunits. An α and a β subunit are cut from a single precursor chain. The human insulin receptor precursor exists as two isoforms, A and B, which arise as a result of a difference in splicing of the mRNA.[358] Twenty-one introns are removed by splicing to form the mRNA for the 1355-residue B precursor.[348,358a] The mRNA for the A form lacks a 36-nucleotide segment (exon 11) that is discarded during splicing. The α chains come from the N-terminal part of the precursor and the β chains from the C-terminal part. The receptor sequence is numbered as in the longer B form precursor. (However, many authors number the chains of the A form receptor as in its precursor, 12 less than the numbers given here for residues 719 or higher[359]). The B form receptor has 731-residue α chains while the A form has 719-residue α chains as a result of the missing sequence from exon 11 (see Fig. 11-11A). Four residues (732–735 in the B form) are cut out and discarded, leaving 620-residue β chains for both isoforms. The chains become linked by three or more disulfide crossbridges.[360]

The two α subunits, which contain the insulin binding sites, are apparently present entirely on the outer surface of the plasma membrane. The β subunits pass through the membrane with their C termini in the cytoplasm.[361–363] Both α and β subunits are glycoproteins. Study of the amino acid sequences suggests that only one 23-residue segment (residues 930–952 of the B isoform) of hydrophobic amino acid residues in each β subunit is likely to exist as an α helix that spans the membrane. This raises several questions. How can a signal be sent from the cell surface into the cytoplasm through a single pair of α helices? Are there additional parts of the receptor that span the bilayer of the membrane? To make it easier to visualize these questions refer to Fig. 11-11B, which is a more realistic, although fanciful, drawing than that in Fig.11-11A. Electron microscopy and crystallography are now providing the first direct images of the receptor.[363a, b, c]

One possible way in which a signal could be sent through the membrane is for binding of insulin to promote aggregation of two or more receptors.[364] If insulin induces the receptors to stick together on the outside of the cell, the parts protruding into the cytoplasm would also tend to aggregate. This could induce a response. A second possibility is that a conformational change in the α subunits pulls or pushes on the β subunit allowing the latter to be exposed less or more on the cytoplasmic side. This difference might be sufficient to cause a conformational change in the cytoplasmic domain of the subunit. A third possibility involves activation by a twisting mechanism as proposed for the aspartate receptor (Fig. 11-8).

A large part of the cytoplasmic domain of the insulin

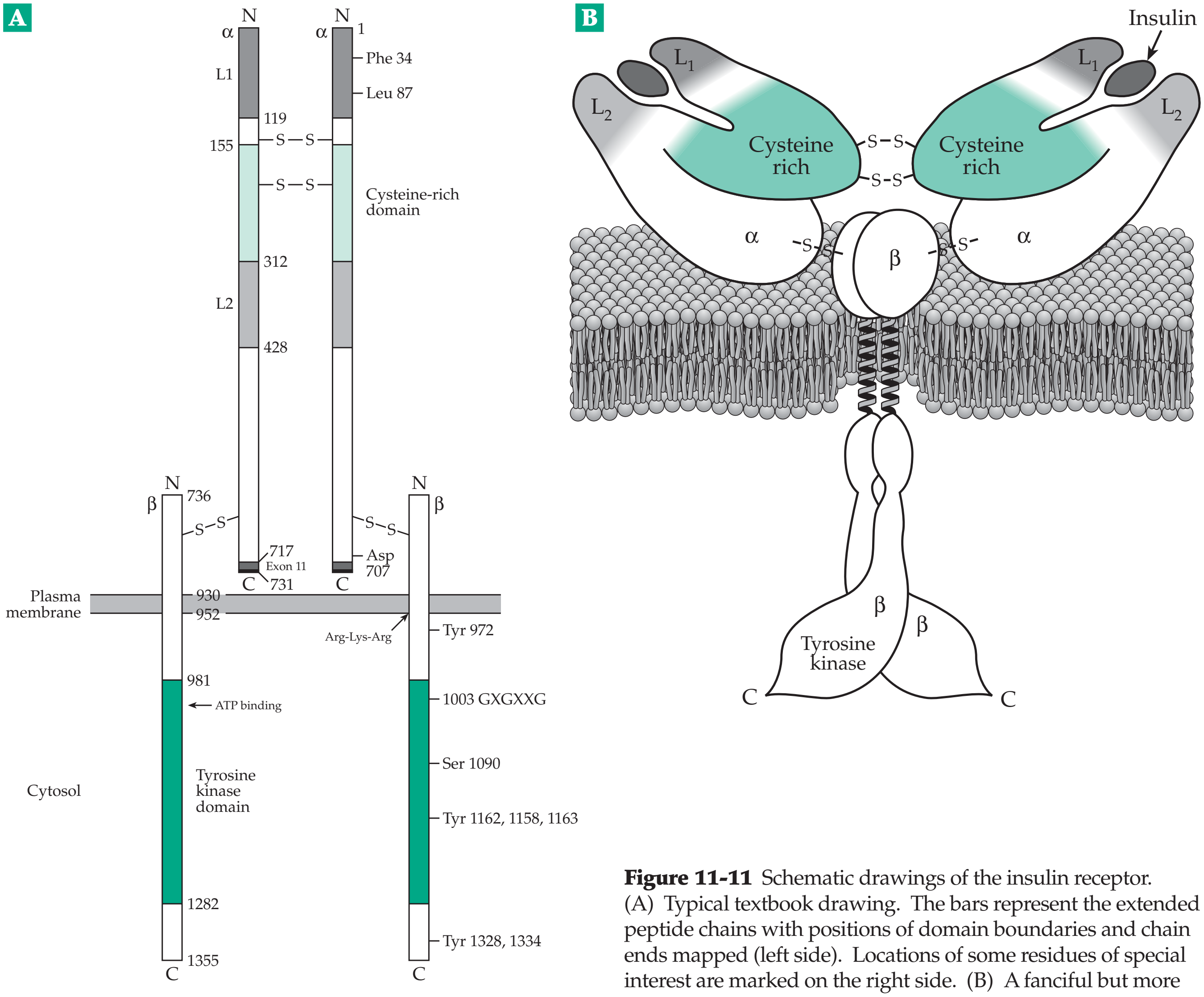

A
N
α
L1
119
155
S—S
S—S
312
L2
428
N
1
α
Phe 34
Leu 87
Cysteine-rich domain
N
β
736
S–S
717
Exon 11
731
C
S–S
Asp 707
C
N
β
Plasma membrane
930
952
Arg-Lys-Arg
Tyr 972
981
ATP binding
1003 GXGXXG
Ser 1090
Cytosol
Tyrosine kinase domain
Tyr 1162, 1158, 1163
1282
Tyr 1328, 1334
1355
C
C
B
Insulin
L_1
L_1
L_2
L_2
Cysteine rich
Cysteine rich
S–S
S–S
α
S–S
β
S–S
α
β
β
Tyrosine kinase
C
C

Figure 11-11 Schematic drawings of the insulin receptor. (A) Typical textbook drawing. The bars represent the extended peptide chains with positions of domain boundaries and chain ends mapped (left side). Locations of some residues of special interest are marked on the right side. (B) A fanciful but more realistic picture of the receptor. See Fabry *et al.*[367]

receptor consists of a tyrosine-specific protein kinase whose three-dimensional structure is depicted in Fig. 11-12.[365,365a] Not only can it phosphorylate –OH groups on tyrosine side chains of other proteins[362] but also it catalyzes ATP-dependent **autophosphorylation** of several residues in the C-terminal region. The most readily phosphorylated residue is Tyr 1158.[366] Some diabetic individuals have receptors with impaired tyrosine kinase activity.[362]

Using directed mutation of the cloned receptor gene, Lys 1018 in the ATP-binding part of the tyrosine kinase domain was replaced by alanine. This caused a loss both of kinase activity and of biologic response to insulin.[362] Thus, both the tyrosine kinase activity and autophosphorylation appear essential. If so, aggregation of two or more receptors may increase the extent of autophosphorylation and initiate a response.

Studies of many mutant proteins indicate that insulin binds to the α chains of the receptor between the two domains labeled L1 and L2 in Fig. 11-11.[368] Among essential residues is Phe 39 (marked).[369] Insulin contains three disulfide linkages (Fig. 7-16) and might undergo a thiol–disulfide exchange reaction (Eq. 11-7) with SH groups present in a cytoplasmic cysteine-rich domain of the α subunit of the receptor or with an external thiol compound.[369a,b] Such an exchange may also be essential for activation of the receptor.[370]

3. A Second Messenger for Insulin?

The known regulatory effects of insulin (Table 17-3) often involve phosphorylation of serine or threonine side chains on specific proteins. The tyrosine kinase of the activated insulin receptors does not catalyze such phosphorylation. Therefore, it seems likely that one or more second messengers or mediator substances are needed. Much effort has gone into searching for

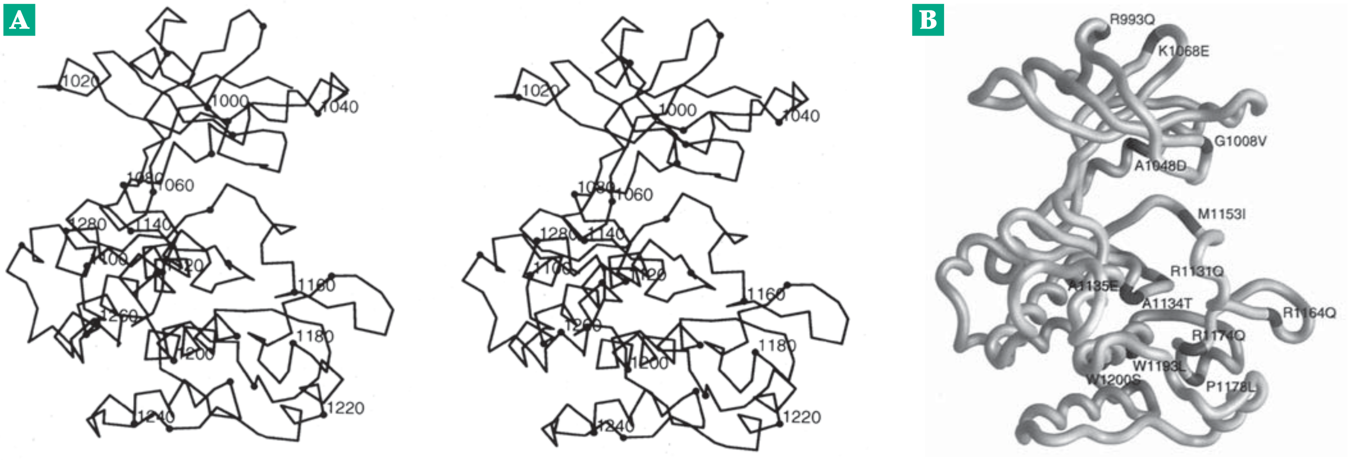


Figure 11-12 (A) Stereoscopic view of an α-carbon trace of the insulin receptor kinase domain. Every tenth residue is marked with a filled circle and every twentieth residue is labeled. (B) Locations of missense mutations in noninsulin-dependent diabetes mellitus patients mapped onto the receptor kinase structure. The mutations are R993Q, G1008V, A1048D, R1164Q, R1174Q, P1178L, W1193L, and W1200S. Here, R993Q is the mutant in which arginine 993 is replaced with glutamine, etc. From Hubbard *et al*.[365]

substrates for the tyrosine kinase,[353,362,371,372] whose activity on external protein substrates reaches a maximum after receptor tyrosines 1158, 1162, and 1163 have been autophosphorylated. Phosphorylation of such **insulin receptor substrates (IRSs)** as well as of some smaller "adapter" proteins is thought to initiate a series of complex cascades that serve to pass the insulin signal on to a variety of sites of action within a cell.[373–376] The large 185-kDA **insulin receptor substrate-1 (IRS-1)**, which is present in most cells, is phosphorylated on several tyrosines, usually within the sequences YXXM or YMXM. Phosphotyrosine within such sequences is known to be a ligand for the recognition domain SH2, which is present in many proteins. Binding of SH2 domains of other proteins to IRS-1 allows the insulin signal to be passed to several different proteins in the branched signaling pathways. Other insulin receptor substrates include **IRS-2**,[376a,b] protein Gab-1[376c], and the smaller protein called **Shc**, which occurs as 46-, 52-, and 66-kDa isoforms and is discussed further on p. 568.[376,377] Shc is an adapter protein which forms a complex that triggers the activation of the G protein Ras and the MAP kinase cascade shown in Fig. 11-13.[377,378] This pathway (see Section H,2) is thought to mediate the mitogenic (growth promoting) effects of insulin and also to promote phosphorylation of serine and threonine side chains of many proteins in the cytoplasm, the cytoskeleton, ribosomes, membranes, and the nucleus.[379,380] One serine kinase specifically phosphorylates the insulin receptor on serine 1078.[381]

In addition to IRS-1, IRS-2, and Shc, there are additional adapter proteins that interact with the phosphorylated tyrosine kinase domain of the insulin receptor.[382] Furthermore, the receptor tyrosine kinase may catalyze *direct* phosphorylation of some proteins, e.g. a cytoplasmic loop of the β_2 adrenergic receptor (Fig. 11-6), without intervention of an adapter.[383]

Phosphorylated IRS-1 activates a second signaling pathway by interacting with an 85-kDa SH2-containing protein that is a subunit of phophatidylinositol 3-kinase.[384–386] This activates the 110-kDa catalytic subunit of the 3-kinase, which catalyzes formation of phosphatidylinositol 3-phosphate as well as PtdIns (3,4)P_2 and PtdIns (3,4,5)P_3.[387,387a] These compounds, which remain within membranes, activate other branches of the signaling cascade, some of which may converge with those of the MAP kinase cascade. However, there appears to be specific activation of a ribosomal Ser/Thr kinase that, among other activities, phosphorylates ribosomal protein S6, a component of the small ribosomal subunit.[388] It also phosphorylates some isoforms of protein kinase C and other enzymes. PtdIns 3-kinase may also activate 6-phosphofructo-2-kinase (Fig. 11-2, step *d*).[384,388]

One of the most important effects of insulin is to increase glucose uptake by cells.[373,389] The mechanism is thought to depend upon the transporter protein GLUT4, which is stored within the membranes of small cytoplasmic vesicles. Binding of insulin to its receptors induces movement of these vesicles to the plasma membrane where fusion with the plasma membrane makes the GLUT4 molecules available for glucose transport.[390] Phosphatidylinositol 3-kinase also plays an important role. The PtdIns(3,4)P_2 and PtdIns(3,4,5)P_3 generated by this enzyme (Fig. 11-9) remain in the membrane and provide sites to which various proteins, e.g., those containing PH domains (Chapter 7) may bind. Among the proteins "recruited" to the inner membrane surface in this way is a Ser/Thr protein kinase known as PKB/Akt. The Akt abbreviation refers to its relationship to a particular viral oncogene.[390a–390e] and the phosphatidyl-inositol derivatives generated by the 3-kinase may have a direct effect upon the exocytosis of the GLUT-4 containing vesicles as may Gq-coupled receptors.[391]

Although insulin unquestionably stimulates phosphorylation of many proteins, the major metabolic effect appears to result from *dephosphorylation* by phosphatases of phosphorylated forms of several enzymes.[379,392] These include **glycogen synthase** and **pyruvate dehydrogenase**, which are both activated by dephosphorylation, and **glycogen phosphorylase**[393] and **hormone-sensitive lipase**, which are deactivated by dephosphorylation. These changes result in increased synthesis and storage of both glycogen and triacyglycerols. Insulin also stimulates a membrane-bound cAMP phosphodiesterase causing a reduction in cAMP concentrations.[394]

Insulin injected into the human body has a half-life of only 5–10 minutes. Much of the hormone is destroyed by hydrolytic cleavage of the peptide chain by a non-lysosomal protease.[395–397] The receptors together with the hormone are taken into cells by endocytosis from clathrin-coated pits and both may be degraded in the lysosomes.[396,398] Some signaling may arise from insulin receptors in endosomes.[399] Some of the hormone enters the nucleus, perhaps still bound to the receptors,[400] and may have a direct effect upon gene transcription. Some receptors are recycled to the cell surface.[398] This latter process, together with that of new receptor synthesis, controls the number of receptors on the surface and, in turn, the sensitivity of the cell to insulin. This is only one mechanism used for desensitization or **downregulation** of receptor sensitivity and numbers, something that happens normally for any hormone when its concentration is high.[401,402]

H. Growth Factors, Oncogenes, and the Cell Cycle

Insulin is just one of a large number of proteins that are secreted by cells and which influence the growth of nearby cells.[403–405] For example, blood platelets, which aggregate at the site of an injury to a blood vessel, contain granules (α granules) which, when the platelet is "activated" (see Box 8-A) release **platelet-derived growth factor** (PDGF), a 28- to 31-kDa glycosylated peptide which stimulates growth and tissue repair in the injured region.[406,407] There are many other protein growth factors, a few of which are considered here. Others are discussed in Chapters 30–32.

1. Oncogenes and Proto-oncogenes

Studies of cancer-causing cellular oncogenes and **proto-oncogenes** (**c-*onc***) identified in the human genome[408–411] have contributed greatly to our understanding of the action of growth factors. Proto-oncogenes are segments of DNA that code for proteins that have a normal function but which may become "activated" by a mutation or by chromosomal rearrangement to become cancer-causing oncogenes (Box 11-D). Oncogenes were discovered first in oncogenic (cancer-causing) retroviruses,[412] where they are designated v-*onc*. The v-*onc* genes are usually very similar but not identical to the corresponding c-*onc* genes and are thought to have arisen from them. The v-*onc* genes are often incomplete or have become fused to other genes by genetic recombination. Considerable excitement has attended the discovery that *many solid human tumors contain activated oncogenes*,[413] several of which are listed in Table 11-3.

In 1983 it was found that the oncogene *sis*, which is carried by the simian sarcoma virus, has a 104-residue sequence nearly identical to that of human PDGF.[404,414] The PDGF receptor has a large extracellular domain consisting of five immunoglobulin-like domains. Like the insulin receptor, it has tyrosine kinase activity which resides in a C-terminal cytoplasmic domain.[406,407,415] This suggests that the malignant transformation of cells by the viral v-*sis* gene leads to an excessive production of normal PDGF and, consequently, to excessive growth. The sequences related to v-*sis* are often found in human tumors and are located on chromosome 22.[416] Rearrangements by which part of this chromosome is moved to another location are well known and sometimes lead to conversion of the proto-oncogene to an active cancer-inducing oncogene.

The avian v-*erbB* oncogene and oncogene *neu*[417-419] both have sequences homologous to that of the gene for the receptor for the 53-residue **epidermal growth factor** EGF.[420–423a] The corresponding cellular c-*erbB* is the gene for the EGF receptor, a large 170-kDa 1186-residue protein. It also resembles the insulin receptor in having an N-terminal domain outside the cell, a single hydrophobic helix that spans the membrane, and a cytoplasmic domain with tyrosine-specific protein kinase activity.[403,419,424,425] It differs from the insulin receptor in being a single peptide chain. However, when EGF binds, the receptor dimerizes and the protein kinase is activated.[426] It phosphorylates its own Tyr 1173 near the C terminus, as well as tyrosines in the **lipocortins**, 36- to 38-kDa calcium-binding proteins located on the cytosolic face of plasma membranes.[427] The activated receptor also stimulates phospholipase C with a resultant increase in concentrations of inositol trisphosphate and Ca^{2+} (see Fig. 11-13).[428] The active oncogene v-*erbB* encodes a major fragment of the EGF receptor. However, most of the N-terminal domain

TABLE 11-3
A Few Oncogenes That Have Interested Biochemists

Oncogene Symbol	Source	Properties
v-*sis*	Simian sarcoma virus	Gene product is closely related to the B-chain of platelet-derived growth factor
v-*erbB*	Avian erythroblastosis virus	Gene product is shortened version of the EGF receptor, a tyrosine kinase
neu	(Her-2) Rat neuroblastomas A similar gene is found in human breast cancers and adenocarcinomas	Homologous to EGF receptor; has tyrosine kinase activity; may control phosphotidylinositol 3-kinase
v-*src*	Rous avian sarcoma virus	Gene product is another tyrosine-specific protein kinase
abl	Chronic myelogenous leukemia	Similar to *src;* characteristic chromosomal translocations yield cancer
ras, *has*, v-H-*ras* v-K-*ras* N-*ras*	Human bladder, colon, lung carcinomas	Homology with G_s and G_i regulatory proteins; several human proto-oncogenes exist; a single mutation in H-*ras* may lead to cancer
bas, *kis*	Murine sarcoma viruses	
v-*myc*	Burkitt's lymphoma, mouse plasmacytomas, avian retrovirus MC29	Nuclear location for gene product; which forms complex with protein Jun
mos	Burkitt's lymphoma, mouse plasmacytomas	
v-*fos*	Osteosarcoma virus (mouse)	Nuclear location for gene product, which forms complex with protein Jun
v-*jun*		Protein product is subunit of transcription factor AP-1
*bcl*1		Cyclin D1

including the EGF binding site is missing and the C-terminal end has been shortened. The tyrosine kinase domain is intact but Tyr 1173 is missing.

The rat *neu* oncogene (also called *erbB*-2 and *HER*-2), which apparently is derived from the gene for another growth factor receptor, differs from normal *neu* by a single nucleotide. This change causes valine to be substituted for glutamic acid at position 664 in the membrane-spanning domain of the 185-kDa protein.[418] This evidently gives an overactive and perhaps uncontrolled tyrosine-specific protein kinase. The c-*erbA* proto-oncogene appears to be the nuclear receptor for the thyroid hromone **triiodothyronine** (Chapter 25).

The Rous sarcoma virus oncogene v-*src* and a family of related oncogenes are derived from protein tyrosine kinases that are attached with the aid of a myristoyl anchor to the inner, cytoplasmic surfaces of membranes.[429–431] They may be activated by interaction with an occupied surface receptor. It has been difficult to understand the functioning of the normal Src protein. However, inactivation of c-*src* in mice caused the serious bone disease **osteopetrosis** in which the osteoclasts fail to function properly in resorbing the bone matrix, thereby allowing excessive accumulation of calcium phosphate.[432,433] A c-*src* deficiency also decreases formation of the bone adhesion protein **osteopontin**, an RGD protein.[434] There is actually a *family* of src proteins, some of which have important functions in lymphocytes.[435–438] The gene for one of these is mutated in the β cell disorder **agammaglobulinemia**.[437]

Another oncogene derived from a tyrosine kinase

BOX 11-D CANCER

Although cancer occurs in about 200 clinically distinct types, most cancers can be classified into four categories. In **leukemias**, which account for 3% of the ~700,000 cases of cancer diagnosed per year in the United States, an abnormal number of leukocytes are produced by the bone marrow. **Lymphomas**, such as Hodgkin's disease and Burkitt's lymphoma, arise from lymphocytes. They account for ~5% of human cancers. In these diseases malignant cells are produced in the spleen and lymph nodes and sometimes aggregate in lymphoid tissues. **Sarcomas**, solid tumors of bone or other connective tissue, contribute ~2% to the total of human cancers, while **carcinomas**, cancers of epithelial tissue, account for 85%. Carcinomas may develop from either the external or the internal epithelia, including the glands, lungs, and nerves.[a–d] More than one-third of all cancers in the United States are nonmelanoma carcinomas of the skin;[e] for these the mortality rate is low. Lung, colorectal, breast, and prostate tumors account for 55% of cancer deaths.[c] Carcinomas predominate in humans, but lymphomas, leukemias, and sarcomas are much more prevalent in laboratory animals and fowl.

An important characteristic of cancer cells is their uncontrolled proliferation. They don't respond to the normal signals from adjacent cells that indicate that cell division should stop. Cancer cells also differ dramatically from those present in warts and other benign tumors and in psoriasis. These conditions also result in excessive proliferation of cells and partial derangement of normal regulatory processes.

A second characteristic of cancer cells is that they usually appear less differentiated than the tissues from which they arise and are more like embryonic cells. Many cancers produce **ectopic proteins**, proteins inappropriate to the tissue involved and often identical to proteins synthesized by embryonic or fetal cells. A well-known example is **α-fetoprotein**, a 72-kDa glycoprotein normally present in serum in almost undetectable amounts but present in large amounts when some types of cancer are present.[f] A third property of cancers is the tendency toward **metastasis**, the detachment of cells from the cancer and their development in distant parts of the body.[g–i]

Cancer cells don't grow any faster than normal cells but they continue to divide when normal cells would not. For this reason, a cancer can grow rapidly and its demands for nutrients can literally starve the host. Cancer tends to weaken the immune system, making the host more susceptible to infections. In addition, cancers often interfere directly with the functioning of various organs and may cause death in this way.

What initiates a cancer? We know that cancers can arise from only one or a very small number of cells and that cancer can be induced by carcinogenic chemical compounds, by certain viruses, and by radiation. Use of tobacco appears to be responsible for about 30% of all human tumors.[c,j] Diet also affects the likelihood of developing cancer. For example, diets containing less animal fat and more fruits and vegetables are associated with lower levels of colon cancer.[c,k] Many carcinogenic compounds are naturally present in foods. Genetic factors help to determine susceptibility to cancer[b,l] and characteristic chromosomal aberrations are usually associated with cancer.[m–o] The incidence of cancer increases markedly with age.

A common feature of all the agents that induce cancer is the production of mutations and cancer probably always involves some alteration in the cell's DNA. The long lag between exposure to carcinogenic materials and development of cancers, often 20 years or more, suggested that more than one mutation or chromosomal rearrangement is required for production of a cancer. Recent evidence confirms that several mutations are required.[b,p] Relevant to this conclusion is the fact that carcinogenic compounds can be applied to the skin in amounts sufficient to cause a number of mutations in the epithelial cells but insufficient to actually induce cancer. Then, even many years later, irritant compounds known as **cancer promoters** can be applied and cancer will develop promptly. The promoters apparently induce cell proliferation which leads to more errors in DNA replication, converting an initially mutated cell to a cancerous cell. The most studied promoters are the phorbol esters, which are known to activate the protein kinase C isoenzymes (Section E,2). Any factors that increase rates of cell division such as some hormones, excess calories, or chronic inflammation cause increased cancer.[c] Cell divisions are accompanied by errors in replication of DNA and sometimes by translocation or deletion of parts of chromosomes. Chronic infection by bacteria, viruses, or other organisms may cause cancer as a result of continuing inflammation. Other cancers arise from integration of viral DNA into the host's DNA.

Virally induced cancers are often epidemic among poultry and rodents. When infected with cancer-causing viruses from these animals, cells in culture often become **transformed**. Whereas normal cells tend to respond to **contact inhibition** and grow as a monolayer, transformed cells continue to divide after the monolayer is complete. In laboratory studies, transformation of cells is often taken as the equivalent to an early step in cancer production in an animal. Studies of transformation led to the identification and characterization of several **viral oncogenes** which cause the transformation. These are designated by abbreviations such as v-*src* (the oncogene of Rous sarcoma virus, which induces cancer in chickens) and v-*sis* (the oncogene of simian sarcoma virus, which causes cancers in monkeys). Some other oncogenes are described in the main text. Naturally occurring gene

BOX 11-D CANCER (continued)

sequences closely homologous to those of the viral oncogenes have been found in many solid human tumors. Study of oncogenes has shown that they are related to and derived from **proto-oncogenes**, normal cellular genes that are involved in control of growth and differentiation.

Oncogenes are often "amplified" in tumor cells so that their copy number is greater than that of the corresponding genes in normal cells. For example an oncogene related to the viral oncogene *neu* is amplified in many human breast and ovarian cancers[q] and oncogene *src* in many colon cancers.[r] Mutated *ras* genes have been found in over one-third of human colorectal cancers.[s] Amplified oncogenes *ras*, *myc,* and *myb* have been observed in other cancers.[t] Oncogenes are often overexpressed or are responsible for over-expression of other genes. One idea that developed from these observations is that cancer cells may secrete new or mutated growth factors that stimulate their own receptors (**autocrine** stimulation) in a way that promotes uncontrolled growth.[u]

Cancer develops in stages and in many cases defective proto-oncogenes appear before truly malignant cells appear. The latter must result from additional mutations that often involve *loss* of parts of chromosomes. A major breakthrough in our understanding of cancer and how it is induced by loss of genes has come from studies of some rare cancers that are inherited in a Medelian fashion. One of these is **retinoblastoma**, an intraocular tumor which affects 1 child in 20,000. Homozygotes always develop the disease between the ages of 1 and 5. The hereditary defect has been traced to the absence of a functional retinoblastoma gene *RB1*, which is found in band q14 of chromosome 13.[v] Additional mutational events are required to induce cancer. The *RB1* gene was the first **tumor-suppressor** gene (anti-oncogene) identified. These suppressor genes encode proteins that inhibit growth.[w – z] The retinoblastoma gene encodes a 105-kDa DNA-binding phosphoprotein (Rb-*P*).[aa – bb] The Rb protein is phosphorylated and dephosphorylated in a cyclic fashion that is synchronized with the cell replication cycle (Fig. 11-15). It forms a complex with a transcription factor E2F that functions in transcription of the adenovirus genes and is also involved in control of the cell replication cycle.[cc] Deletion of the Rb gene from mice leads to death of embryos homozygous for the mutation.[dd]

Study of other rare hereditary cancers has led to the location of 20 or more additional probable tumor-suppressor genes. One of these, ***p53***, is inactive in over 50% of all human cancers and over 90% of sqamous cell carcinomas of the skin.[ee] In small-cell lung cancers and osteosarcoma *both RB* and *p53* are inactive.[z] Protein p53 is a stronger tumor suppressor than protein Rb. Results of a variety of experiments have suggested that p53, a DNA-binding protein of known structure,[ff] plays a key role in checking DNA for damage at the G_1 to S-phase checkpoint in the cell cycle. If the DNA has too many defects the cycle is stopped in the G_1 stage and the cell may be killed by the process known as **apoptosis**.[gg – jj] Protein p53 has been called the "guardian of the genome." The mechanisms by which it functions are complex and poorly understood. It may act with the assistance of Rb and many other proteins. Mutations in DNA and their repair are discussed in Chapter 27 and the cell cycle is discussed in this chapter and further in Chapters 26 and 32.

Many other cancer susceptibility genes also encode suppressors. Mutation in genes *BRACA1*[kk] and *BRACA2*[ll] are responsible for early onset ovarian and breast cancer. Gene *DPC4* may be a suppressor of pancreatic cancer.[mm] The gene *ptc* (patched), first studied as a developmental gene in *Drosophila,* may encode a suppressor of basal cell carcinoma, the commonest form of human skin cancer.[nn] Gene *p16* (also called *CDKN2*) may be a major suppressor that is mutated in many cancers including the dangerous skin melanoma.[oo] Mutations in the *NF* gene, which may be a cytoskeletal protein, are associated with **neurofibromatosis**,[pp] a relatively common hereditary disease causing tumors of the nervous system. The tumors are usually not malignant but are numerous and disfiguring. Several cancer susceptibility genes are associated with faulty mismatch repair of DNA. Among these is the *APC* gene, whose malfunction is associated with human **familial adenomatous polyposis** which causes thousands of benign tumors in the lining of the large intestine and often colorectal cancer.[qq,rr,ss] In the much more common nonpolyposis colon cancer a complex pattern of instability in several genes is associated with DNA repair,[tt] a topic dealt with further in Chapter 27. The *ATM* gene defective in **ataxia telangiectasia** (Chapter 27) may encode a phosphatidylinositol kinase that is in some way involved in repair of DNA.[ss,uu,vv] A transcription factor gene *nm23* may be a suppressor gene for metastasis[ww] and the cell–cell adhesion molecule **E-cadherin** may suppress tumor invasion in some kinds of breast cancer.[xx] The **VHL** (van Hippel-Lindau cancer syndrome) suppressor protein is defective in the majority of kidney cancers. It normally binds to the **elongin complex**, a DNA-binding complex that functions in control of transcription.[zz]

In addition to treatment by surgery there are numerous chemical approaches to combating cancer.[xy] They usually exploit the tendency of cancers to grow continuously. For example, a toxic analog of a metabolite needed for growth, such as methotrexate (Chapter 16) and 5-fluorouridine (Box 28-C) may be taken up more rapidly by tumor cells than by normal cells. A variety of DNA-binding compounds are useful in chemotherapy (Box 5-B). Alkylating agents such as the nitrogen mustards (Eq. 5-20) and certain antibiotic compounds are also used widely, as are intercalating compounds

BOX 11-D (continued)

such as adriamycin (14-hydroxydaunomycin; Fig. 5-22) and cisplatin (Box 5-B). A disadvantage to most present-day chemotherapy is that normal proliferation of cells, especially of glandular tissues, intestinal epithelium, hair, etc., is severely damaged. A possibility for circumventing this problem is to put normal growth "on hold" temporarily while cancer growth is being inhibited. Some inhibitors of **topoisomerase I** (see Chapter 27) have low toxicity and could be useful. Another therapeutic approach is based on the fact that one function of the immune system is to destroy cancerous or precancerous cells. The immune system tends to weaken with age, which may be one reason that the incidence of cancers rises rapidly in older age Are there ways of stimulating the immune system into increased activity against cancer cells? Another approach is to find ways of increasing the activity of tumor-suppressor genes.

Can drugs be developed to *prevent* cancer? Such "chemoprevention"[yz] may be appropriate for persons carrying genes that make them highly susceptible to cancer. The antiestrogenic drug tamoxifen (Chapter 22) is currently being tested on women with a high risk for breast cancer. Use of oral contraceptives appears to have cut the risk of endometrial cancer substantially.[yy] Newer approaches to contraception may help prevent breast cancer and chemoprevention may also be possible for prostate cancer.[yy]

---

[a] Cairns, J. (1978) *Cancer: Science and Society*, Freeman, San Francisco, California

[b] Cavenee, W. K., and White, R. L. (1995) *Sci. Am.* **272**(Mar), 72–79

[c] Ames, B. N., Gold, L. S., and Willett, W. C. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5258–5265

[d] Ruddon, R. W. (1987) *Cancer Biology*, 2nd ed., Oxford Univ. Press, London

[e] Preston, D. S., and Stern, R. S. (1992) *N. Engl. J. Med.* **327**, 1649–1662

[f] Zhang, D., Hoyt, P. R., and Papaconstantinou, J. (1990) *J. Biol. Chem.* **265**, 3382–3391

[g] Packard, B. (1986) *Trends Biochem. Sci.* **11**, 490–491

[h] Feldman, M., and Eisenbach, L. (1988) *Sci. Am.* **259**(Nov), 60–85

[i] Marx, J. (1993) *Science* **259**, 626–629

[j] zur Hausen, H. (1991) *Science* **254**, 1167–1173

[k] Willett, W. (1989) *Nature (London)* **338**, 389–394

[l] Dragani, T. A., Canzian, F., and Pierotti, M. A. (1996) *FASEB J.* **10**, 865–870

[m] Solomon, E., Borrow, J., and Goddard, A. D. (1991) *Science* **254**, 1153–1160

[n] Nowell, P. C. (1994) *FASEB J.* **8**, 408–413

[o] Pennisi, E. (1996) *Science* **272**, 649

[p] Marx, J. (1989) *Science* **246**, 1386–1388

[q] Slamon, D. J., Godolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., Ullrich, A., and Press, M. F. (1989) *Science* **244**, 707–712

[r] Cartwright, C. A., Meisler, A. I., and Eckhart, W. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 558–562

[s] Bos, J. L., Fearon, E. R., Hamilton, S. R., Verlaan-de Vries, M., van Boom, J. H., van der Eb, A. J., and Vogelstein, B. (1987) *Nature (London)* **327**, 293–297

[t] Yokota, J., Tsunetsugu-Yokota, Y., Battifora, H., Le Fevre, C., and Cline, M. J. (1986) *Science* **231**, 261–265

[u] Sporn, M. B., and Roberts, A. B. (1985) *Nature (London)* **313**, 745–747

[v] Weinberg, R. A. (1990) *Trends Biochem. Sci.* **15**, 199–202

[w] Stanbridge, E. J. (1990) *Science* **247**, 12–13

[x] Weinberg, R. A. (1991) *Science* **254**, 1138–1146

[y] Knudson, A. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10914–10921

[z] Yokota, J., and Sugimura, T. (1993) *FASEB J.* **7**, 920–925

[aa] Wiman, K. G. (1993) *FASEB J.* **7**, 841–845

[bb] Cobrinik, D., Dowdy, S. F., Hinds, P. W., Mittnacht, S., and Weinberg, R. A. (1992) *Trends Biochem. Sci.* **17**, 312–315

[cc] Nevins, J. R. (1992) *Science* **258**, 424–429

[dd] Jacks, T., Fazeli, A., Schmitt, E. M., Bronson, R. T., Goodell, M. A., and Weinberg, R. A. (1992) *Nature (London)* **359**, 295–300

[ee] Ziegler, A., Jonason, A. S., Leffell, D. J., Simon, J. A., Sharma, H. W., Kimmelman, J., Remington, L., Jacks, T., and Brash, D. E. (1994) *Nature (London)* **372**, 773–776

[ff] Cho, Y., Gorina, S., Jeffrey, P. D., and Pavletich, N. P. (1994) *Science* **265**, 346–355

[gg] Marx, J. (1993) *Science* **262**, 1644–1645

[hh] Enoch, T., and Norbury, C. (1995) *Trends Biochem. Sci.* **20**, 426–430

[ii] Kaufmann, W. K., and Paules, R. S. (1996) *FASEB J.* **10**, 238–247

[jj] Hartwell, L. H., and Kastan, M. B. (1994) *Science* **266**, 1821–1828

[kk] Futreal, P. A., and 26 other authors (1994) *Science* **266**, 120–122

[ll] Wooster, R., and 30 other authors (1994) *Science* **265**, 2088–2090

[mm] Hahn, S. A., Schutte, M., Hoque, A. T. M. S., Moskaluk, C. A., da Costa, L. T., Rozenblum, E., Weinstein, C. L., Fischer, A., Yeo, C. J., Hruban, R. H., and Kern, S. (1996) *Science* **271**, 350–353

[nn] Pennisi, E. (1996) *Science* **272**, 1583–1584

[oo] Marx, J. (1994) *Science* **265**, 1364–1365

[pp] Rouleau, G. A., and 20 other authors (1993) *Nature (London)* **363**, 515–521

[qq] Peltomäki, P., Aaltonen, L. A., Sistonen, P., Pylkkänen, L., Mecklin, J.-P., Järvinen, H., Green, J. S., Jass, J. R., Weber, J. L., Leach, F. S., Petersen, G. M., Hamilton, S. R., de la Chapelle,A., and Vogelstein, B. (1993) *Science* **260**, 810–819

[rr] Huang, J., Papadopoulos, N., McKinley, A. J., Farrington, S. M., Curtis, L. J., Wyllie, A. H., Zheng, S., Willson, J. K. V., Markowitz, S. D., Morin, P., Kinzler, K. W., Vogelstein, B., and Dunlop, M. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9049–9054

[ss] Kolodner, R. D. (1995) *Trends Biochem. Sci.* **20**, 397–401

[tt] Papadopoulos, N., and 19 other authors (1994) *Science* **18**, 1625–1629

[uu] Keith, C. T., and Schreiber, S. L. (1995) *Science* **270**, 50–51

[vv] Sanchez, Y., Desany, B. A., Jones, W. J., Liu, Q., Wang, B., and Elledge, S. J. (1996) *Science* **271**, 357–360

[ww] Marx, J. (1993) *Science* **261**, 428–429

[xx] Berx, G., Cleton-Jansen, A.-M., Nollet, F., de Leeuw, W. J. F., van de Vijver, M. J., Cornelisse, C., and van Roy, F. (1995) *EMBO J.* **14**, 6107–6115

[xy] Hoffman, E. J. (1999) *Cancer and the Search for Selective Biochemical Inhibitors*, CRC Press, Boca Raton, Florida

[yy] Henderson, B. E., Ross, R. K., and Pike, M. C. (1993) *Science* **259**, 633–638

[yz] Young, R. C. (2000) *Nature (London)* **408**, 141

[zz] Stebbins, C. E., Kaelin, W. G., Jr., and Pavletich, N. P. (1999) *Science* **284**, 455–461

gene is the Abelson murine leukemia virus v-*abl*.[438a] In the mouse the c-*abl* gene is split into at least ten exons. However, v-*abl* contains all of these as a correctly spliced sequence suggesting that v-*abl* was derived from a c-*abl* messenger RNA. Part of one exon of the cellular gene is missing in v-*abl* and there is at least one base substitution mutation as well. Human c-*abl* is located on the long arm q of chromosome 9. This arm has long been known to be translocated to chromosome 22 (the "Philadelphia translocation") in patients with chronic myelogenous leukemia. There the c-*abl* gene is fused with another gene.[439] The human c-*sis* gene is also located in the region of chromosome 22 that is translocated to chromosome 9 in the same patients.

A group of human pituitary tumors have been shown to contain an oncogene that is apparently a mutated gene for the stimulatory protein G_s.[440] This is the protein that activates adenylate cyclase in response to hormonal activation. The oncogenic mutations inhibit the GTPase activity that normally turns off this activation. These tumors secrete growth hormone which binds to receptors on the tumor cells activating the defective G_s proteins and causing excessive synthesis of cAMP. This in turn promotes growth of the tumor.

The ras oncogenes. Activated *ras*-oncogenes have been found in at least 25% of all human tumors. The proto-oncogenes, which are designated c-H-*ras*, c-K-*ras,* and c-N-*ras*, are found on the short arms of chromosomes 11, 12, and 1, respectively. A single base substitution (G → T) at position 35 of any of the genes, resulting in a Ras protein containing valine instead of glycine at position 12 of the 21 kDa protein product (usually designated p21) produces an active oncogene. Substitutions at position 13 or at positions 59 and 61, which are adjacent to Gly 12 in the three-dimensional structure, can also activate the oncogenes.[441] From the drawing in Fig. 11-7A the locations of glycines 12 and 13 and of residues 59/61 are seen to be close to the β phospho group of bound GTP. Proteins encoded by activated *ras* oncogenes are less active in catalyzing GTP hydrolysis than are the corresponding normal proteins. In addition, the GTPase-activating protein (GAP) that binds to normal *ras* proteins and stimulates their GTPase activity does not affect the mutant oncogenic proteins.

A single-base alteration is capable of activating a *ras* gene with respect to cell transformation. However, initiation of a malignant tumor requires the additional presence of at least a second activated oncogene such as *myc*[442,443] or *fos*[444] or previous transformation of a fibroblast into a nonmalignant but "immortalized" form by treatment with carcinogens. The c-*myc* gene, which is found in active form in plasmacytomas (tumors of B lymphocytes) of mice as well as in the human Burkitt's lymphoma, is normally located on human chromosome 8. In most Burkitt's lymphomas a translocation has brought the c-*myc* gene into the locus of the immunoglobulin heavy chains on chromosome 14. There its transcription may be subject to different controls than in its original location.[445,446] During the translocation process the c-*myc* gene is often broken within the first intron. Thus, the activated gene lacks the first exon and is placed after a new controlling sequence that may drastically alter its transcription rate. Viral *myc* genes are found in at least one human virus, **cytomegalovirus**, which has been associated with carcinomas.

Transcription factors. The proto-oncogenes c-*myc*[447–451a], c-*myb*,[452–454] c-*fos,* c-*jun* and c-*ets*[455] all encode nuclear proteins involved in regulation of transcription. The 39 kDa protein Jun, which is encoded by c-*jun*, is a major component of the **transcriptional activator** called **AP-1**.[456–459] It binds to palindromic **enhancer** sites (Chapter 28) in DNA promoters to increase the transcription rate for a group of genes. Jun is actually a multigene family whose encoded proteins bind to DNA as complexes formed with the 62 kDa phosphoprotein Fos, the product of the c-*fos* gene.[460] The heterodimeric Fos/Jun complex is held together, at least in part, by interactions between leucine side chains lying along a pair of parallel α-helices in a "leucine zipper" (Fig. 5-36; Fig. 2-21).[461,462]

Regulation of the synthesis of Jun is complex, but growth factors such as Neu, EGF,[463] and PDGF stimulate transcription of c-*jun* in cultured cells. Messenger RNA for synthesis of Fos appears within a few minutes of stimulation of PDGF receptors.[464,465] This is one of the earliest known nuclear reactions to a mitogenic stimulus and suggested that a Ras p21 protein is involved in stimulating transcription of c-*fos* in the signaling pathway from PDGF.[466] Synthesis of Fos is also induced by a variety of other stimuli.[456] Upon translocation into the nucleus Fos combines with preexisting Jun to form AP-1, which binds to sites on DNA and induces the transcription of a large number of proteins (Chapter 28). Deletion of the c-*fos* gene in mice leads to defects in developing bone, teeth, and blood cells,[467] while excessive synthesis of Fos has been associated with the human bone disease **fibrous dysplasia**.[468]

2. The MAP Kinase Cascade

Insulin, platelet-derived growth factor (PDGF), epidermal growth factor (EGF), and many other proteins have **mitogenic activity**, that is, they induce cells to transcribe genes, to grow, and to divide. How is this accomplished? Study of such oncogenes as *src* and *ras* suggested that the proteins that they encode also participate in the process, as do transcription factors, including those encoded by the proto-oncogenes

Table 11-4
A Few Abbreviations Used in Discussions of Cell Signaling

AP-1	A major transcriptional activator protein
βARK	Beta adrenergic receptor kinase
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ERK-1, ERK-2	Extracellular signal-regulated protein kinases (proline-directed protein kinases)
Fos	Protein encoded by proto-oncogene *fos* (Table 11-3)
Grb2	Adapter protein containing one SH2 and two SH3 domains
IGF-1	Insulin-like growth factor 1
InsRTK	Insulin receptor tyrosine kinase
InsP_3 or IP_3	Inositol (1,3,5)-trisphosphate
Jun	Protein encoded by proto-oncogene *jun*, an AP-1 gene
MAPK	Mitogen-activated protein kinase (also designated ERK)
MAPKK	A kinase acting on MAPK
MEK-1,MEK-2	Mitogen-activated ERK-activating kinases; dual function Ser/Thr and Tyr protein kinases; also called MAP kinase kinases
c-*myc*	A cellular proto-oncogene
PDGF	Platelet-derived growth factor
PtdIns (or PI) 3-kinase	Phosphatidylinositol 3-kinase
PKA (cAPK)	cyclic AMP-dependent protein kinase
PKC, PKCα, β, γ	Protein kinase C,α, β, γ subforms
PKR	dsRNA-activated protein kinase
PLA	Phospholipase A
PLC	Phospholipase C
PP2A	Protein phosphatase, type 2A
PTK	Protein tyrosine kinase
PTP (PTPase, PTP 1B, etc.)	Protein tyrosine phosphatases
pY	Phosphotyrosine residue
Raf-1	A cytoplasmic serine/threonine protein kinase; also called MAP kinase kinase kinase (MAP3K)
Ras	A monomeric G protein encoded by the proto-oncogene *ras*
RSK	Ribosomal protein S6 kinase
RTK	Receptor tyrosine kinase
RTK-*P*	Phosphorylated receptor tyrosine kinase
SH2, SH3	Src homology domains 2 and 3
Sos	"son of sevenless," a GDP – GTP exchange factor named for a similarity to the protein encoded by the *Drosophila* sevenless gene
Src	Protein tyrosine kinase encoded by oncogene *src* (Table 11-3); contains recognition domains SH2 and SH3
V2R	Type 2 vasopressin receptor
Shc	A proline-rich adapter protein containing SH2 and PH domains

myc, *fos*, and *jun*. The Src and Ras proteins are anchored on the inner surfaces of cytoplasmic membranes. Although the exact functional relationships of all of the components, one to another, has not been completely established, a general picture, usually described as the **mitogen-activated protein kinase** (MAP kinase) **cascade**, has emerged.[69,380,469-470b] This is sketched in simplified form in Fig. 11-13. The many amplification steps provide for ultrasensitive responses.[342]

The binding of a hormone or growth factor (a ligand) to a dimeric receptor activates the protein kinase domain of the receptor which phosphorylates a number of tyrosine hydroxyl groups of the receptor itself. This autophosphorylation is followed by a variety of events, which include phosphorylation of tyrosine side chains of various other proteins.[426] Another major event is the *binding of a variety of different protein molecules containing recognition domains to the phosphotyrosyl groups of the activated receptors.*[471] The major recognition motif is the SH2 domain. See Figs. 7-30 and 11-14.[472 – 475] Proteins containing SH2 domains can bind to the phosphotyrosyl groups of the activated receptors and while bound become phosphorylated by the receptor tyrosine kinase action and/or be activated allosterically.

Other proteins interact with the receptors indirectly through adapter molecules which have no catalytic activity. Two well-known adapter proteins are **Grb2**[476] and **Shc**.[477,478] The 25-kDa protein Grb2 consists entirely of recognition domains, one SH2 and two SH3 domains (Fig. 11-14). The larger Shc, which is found in all mammalian tissues, contains a 200-residue phosphotyrosyl-binding PH domain (Chapter 7) at the N terminus, a collagen-like domain that binds to Grb2 and an SH2 domain at the C terminus.[376,377,477]

Adapter Grb2 binds to a phosphotyrosine side chain of an activated receptor, such as that for EGF, and simultaneously binds to the GDP-GTP exchange protein called **Sos** (Table 11-4). This signals Sos to activate the membrane-bound Ras protein by converting it into the GTP form. The second adapter Shc may also participate in formation of the receptor kinase - Grb2–Sos complex,[479] perhaps permitting formation of a more robust complex that may receive signals from more than one kind of receptor. Functions of other members of the Grb adapter family are being discovered.[479a]

Activated Ras binds to and activates the cytoplasmic serine/threonine protein kinase called **Raf-1**.[380,480,481] This kinase becomes transiently activated within 2–3 min of the binding of a mitogen to a receptor. Raf-1 initiates a cascade of other protein kinases by acting on the dual-function Ser/Thr and tyrosine protein kinases called **MEK-1** and **MEK-2**. The phosphorylated, active MEK proteins phosphorylate the mitogen-activated protein kinases **MAPK** which act on a variety of other proteins. Two of the best known MAPK proteins are designated **ERK-1** and **ERK-2**. These are *proline-directed* kinases which phosphorylate serines and threonines that are neighbors to prolines, e.g. in the sequence PLS/TP.[380] The activated ERKs are able to phosphorylate a large number of different proteins including nuclear proteins that control the transcription of such protein transcription factors as AP-1 and Myc. A protein known as the **serum response factor** binds to nucleotide sequences CC(A/T)$_6$GC in the DNA to locate initiation sites for transcription. Other proteins that have been phosphorylated by the MAP kinase cascade then induce transcription.[482] The induction of c-*fos* mRNA is one of the earliest identified responses to growth factors.[375,456,483] Protein Jun, whose synthesis is induced independently by almost all growth factors, is usually present in excess.

The Fos/Jun complex is transcription factor AP-1, which induces transcription of many genes needed for cell growth. However, transcription of specific genes often depends upon additional nucleotide sequences. For example, the sequence CGGAAA is present in an **insulin response element** found in the promoter sequences of genes encoding such proteins as phosphoenolpyruvate carboxykinase, glyceraldehyde phosphate dehydrogenase, and prolactin–proteins whose synthesis is induced by insulin.[484]

The MAPK cascade also has direct effects upon protein synthesis, i.e., on the translation of mRNA messages. For example, insulin stimulates phosphorylation of proteins that regulate a translation initiation factor, a protein called eIF-4E (see Chapter 29). Phosphorylation of inhibitory proteins allows them to dissociate from the initiation factor so that protein synthesis can proceed.[485,486]

The scheme in Fig. 11-13 is complex, but in reality it is *much* more complex than is shown. Each protein kinase (receptor kinase, Raf-1, MEK, and MAPK) will phosphorylate not only the proteins indicated in this scheme but also any others that meet the specificity requirements of the kinases. Thus, there will be branches diverging from the pathways shown.[486a,b,c] There are isoenzymes that provide further divergence and interaction.[487] Not only do pathways diverge but also others *converge*. Thus, binding of many different ligands to their receptors activates the same MAPK cascade. For example, seven-helix G protein-coupled receptors release their βγ subunits which may also activate Ras as indicated on the right edge of Fig. 11-13A. At the same time the α subunits of the heterotrimer G proteins can affect not only adenylate cyclase but also phospholipases C which can, in some cases, also activate the MAP kinase pathway.[488,489] Sphingosine 1-phosphate may be released from membrane spingolipids and activate the same cascade.[490] However, hormones do not all affect cells in the same way. In view of all the converging pathways, how is this possible? Part of the answer lies in the proximity or spatial separation of components of the pathway. The kinases exist in complexes with other signaling proteins and may phosphorylate them, sending a signal back, as well as forward via other protein substrates. There are also unknown kinetic considerations. Hormones, neurotransmitters, and calcium ions are often released in pulses. The signaling system must integrate effects of all the stimuli that arise from different parts of the cell, at different times, and from differing receptors. At the same time all of the phosphorylated proteins are acted upon by phosphatases that may either activate or deactivate the proteins and by proteases that process newly formed peptides and modify or destroy mature proteins. The various modifying enzymes act on cytosolic proteins, proteins of membranes, of the cytoskeleton, and of the nucleus. The regulatory processes that we discuss in such minute detail involve the very substance of living cytoplasm which is ever-changing and responding to its surroundings. The flow of energy, provided by synthesis of ATP and by the use of ATP by kinases, phosphatases, and the protein synthetic machinery, goes along with the flow of information and drives the signaling network. Evolution has shaped this system to allow it to respond appropriately for every species.

3. The Cell Cycle and Control of Growth

When a cell divides it is of utmost importance that the DNA be replicated reliably. This requires that the dividing cell be large enough and contain enough biosynthetic precursor materials to complete the elaborate process of DNA synthesis and of mitosis. The **cell**

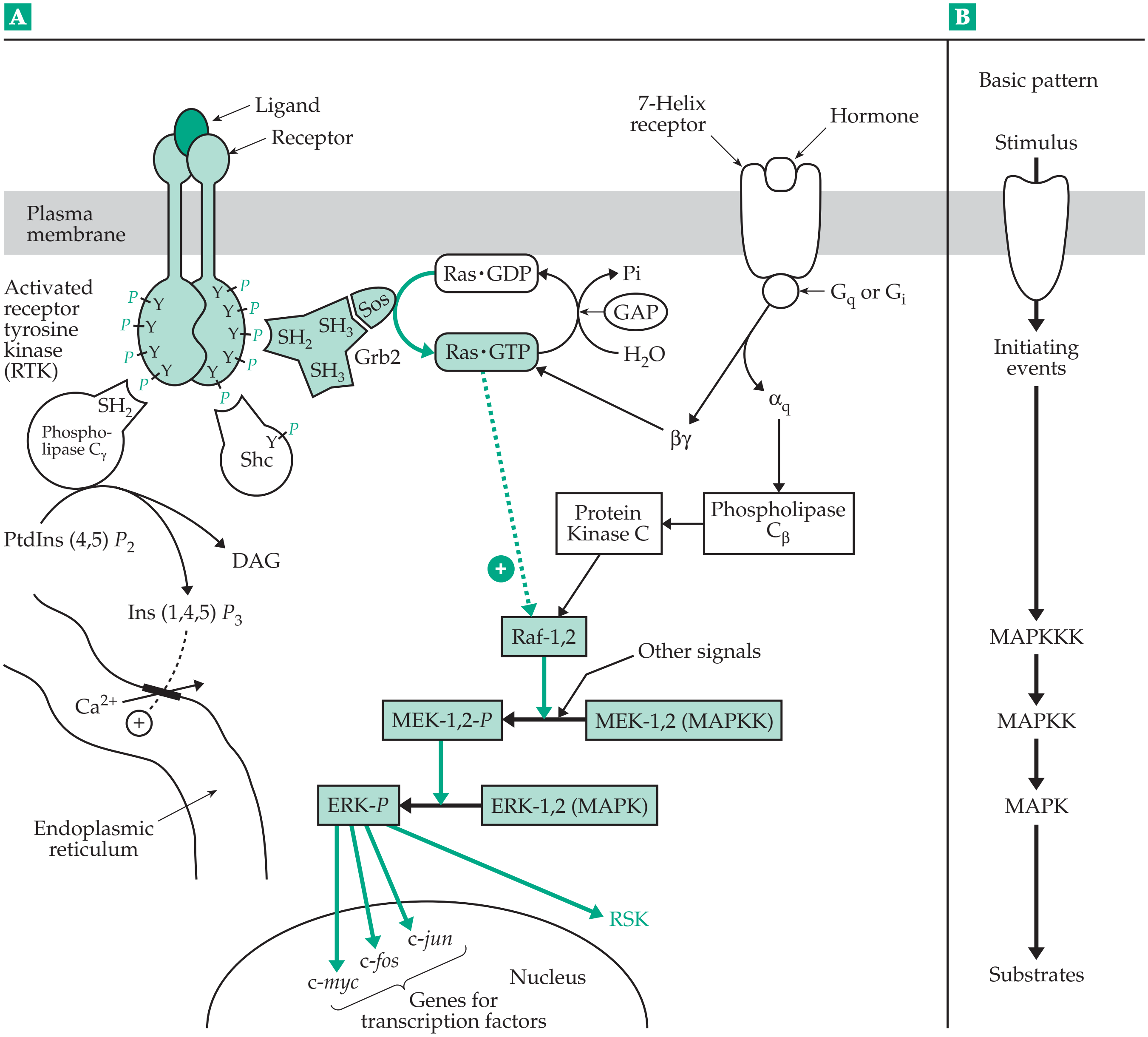

A
B
Ligand
Receptor
Plasma membrane
Activated receptor tyrosine kinase (RTK)
Y
P
SH2
Sos
SH3
Grb2
Phospholipase Cγ
Shc
PtdIns (4,5) P2
DAG
Ins (1,4,5) P3
Ca2+
+
Endoplasmic reticulum
Ras·GDP
Ras·GTP
Pi
GAP
H2O
7-Helix receptor
Hormone
Gq or Gi
αq
βγ
Protein Kinase C
Phospholipase Cβ
+
Raf-1,2
Other signals
MEK-1,2-P
MEK-1,2 (MAPKK)
ERK-P
ERK-1,2 (MAPK)
RSK
c-myc
c-fos
c-jun
Genes for transcription factors
Nucleus
Basic pattern
Stimulus
Initiating events
MAPKKK
MAPKK
MAPK
Substrates

Figure 11-13 (A) A simplified version of the mitogen-activated kinase (MAPK) signaling cascade. At left is shown a hormone receptor, e.g., that for the epidermal growth factor (EGF). The receptor tyrosine kinase undergoes autophosphorylation on numerous tyrosines. The resulting phosphotyrosyl (Y-P) groups bind to SH2 domains of adapters such as Grb2 and Shc. Two pathways from the activated receptor are shown. At the left is activation of phospholipase Cγ and formation, at a membrane-bound site, of inositol trisphosphate and diacylglycerol (DAG). The main pathway, in the center, activates Ras with the aid of the G protein Sos. Activated Ras, in turn, activates Raf and successive components of the MAPK cascade. At the right a seven-helix receptor activates both phospholipase Cβ and Ras via interaction with a βγ subunit. (B) A generalized scheme for the MAP kinase pathway. See Seger and Krebs.[380]

replication cycle (or simply cell cycle) is commonly shown as a circle in which the time from one cell division to the next, in a rapidly growing organism or tissue, is represented by the circumference. The time required for DNA synthesis is the **S-phase** and the time required for mitosis the mitotic or **M-phase** (Fig. 11-15). After metaphase there is a **gap** in time denoted **G_1**. A second gap **G_2** separates the synthetic S-phase and the M-phase. The total time required for one cycle varies with conditions. It may be as short as 8–60 min in an early embryo but is usually two hours or more. A slowly growing cell may pause before the G_1 phase in a nongrowing **G_0 phase**.[491,492]

What controls the cell replication cycle? Signals

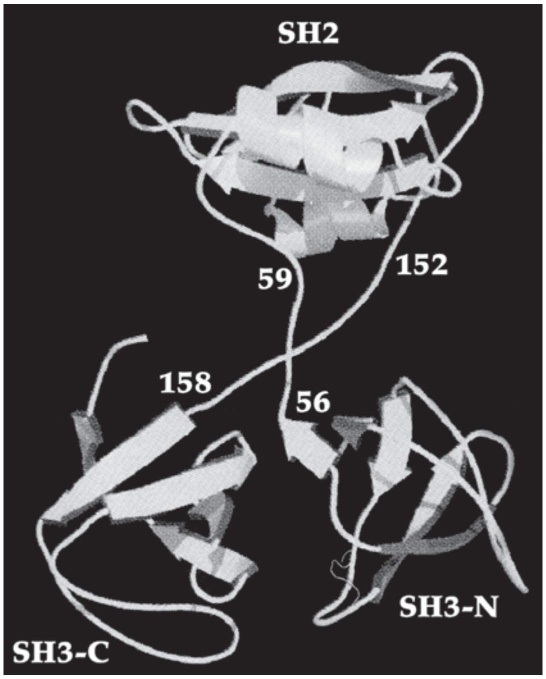

SH2
59
152
158
56
SH3-N
SH3-C

Figure 11-14 Ribbon drawing of the three-dimensional structure of adapter protein Grb2. The two SH3 domains at the N and C termini are labeled, as is the central SH2 domain. Produced with programs MolScript and Raster3D. From Maignan *et al.*[476] Courtesy of Arnaud Ducruix.

from growth factors can stimulate a cell to leave G_0 and enter G_1. However, to go further the cell depends upon a large number of proteins of which two types are prominent. The **cyclins** are labile 45- to 60-kDa proteins which are degraded by the ubiquitin system as part of the cycle. The cyclins associate with a series of protein Ser/Thr kinases known as **cyclin-dependent kinases** (**CDK**s). It is the cyclin–CDK complexes that signal the start of the next step in the cycle.[493–495]

The number of different cyclins and CDK enzymes needed varies with the organism, often being greater for more complex species. As a rule, there are at least two types of cyclins. The **G_1** or **start cyclins** initiate the passage through the start (G_1 checkpoint) into the S-phase, while **mitotic cyclins** initiate the passage from the S-phase into the M-phase. Both vertebrates and *Drosophila* utilize at least four different types of cyclin (A, B, D, and E) and also four or more CDKs (CDK1, CDK2, CDK4, and CDK6).[492] A somewhat different set of these proteins are found in yeast.

Control of cell growth is directly related to the cell cycle. Mitogenic signals from growth factors act to initiate progression through G_1, apparently by stimulating transcription of D-type cyclins.[496] They may affect other steps as well. Oncogenic signals can arise from such oncogenes as *ras* and *abl*.[497,498] As pointed out in Box 11-D, the tumor-suppressor protein Rb becomes phosphorylated in synchrony with the cell

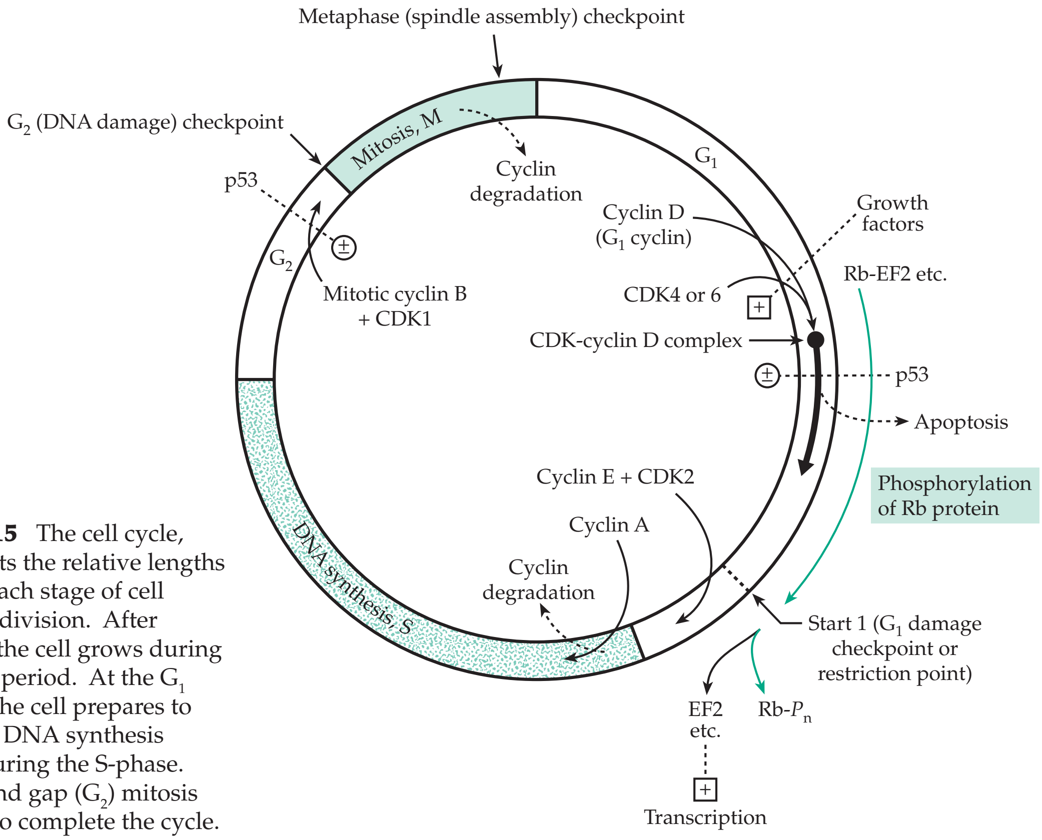

Metaphase (spindle assembly) checkpoint
G$_2$ (DNA damage) checkpoint
Mitosis, M
p53
±
G$_2$
Mitotic cyclin B + CDK1
Cyclin degradation
G$_1$
Cyclin D (G$_1$ cyclin)
Growth factors
Rb-EF2 etc.
CDK4 or 6
+
CDK-cyclin D complex
±
p53
Apoptosis
Phosphorylation of Rb protein
Cyclin E + CDK2
Cyclin A
Cyclin degradation
DNA synthesis, S
Start 1 (G$_1$ damage checkpoint or restriction point)
EF2 etc.
Rb-P_n
+
Transcription

Figure 11-15 The cell cycle, which depicts the relative lengths of time for each stage of cell growth and division. After mitosis (M) the cell grows during the G_1 (gap) period. At the G_1 checkpoint the cell prepares to divide, with DNA synthesis occurring during the S-phase. After a second gap (G_2) mitosis takes place to complete the cycle.

cycle. This seems to be an essential process for passage through the G_1 "checkpoint." The checkpoints in the cycle should be viewed not as points but as essential interlocking processes that must be accomplished before transcription of genes essential for the next step in the cycle can take place.[499] A possibility is that the cyclin-dependent kinase in the cyclin D•CDK4 (or CDK6) complex phosphorylates the Rb protein which in its unphosphorylated state forms a complex with transcriptional regulators of the E2F family. Phosphorylation of Rb allows this complex to dissociate and permits E2F to induce transcription of genes encoding essential proteins for the next step in the cycle.[499a,b,c] Each CDK may also phosphorylate its associated cyclin, a modification that may be essential to the proteolysis of the cyclin and progression to the next stage of the cycle.

The powerful cancer suppressor, p53, which is also described in Box 11-D, in some way senses DNA damage. It prevents passage through the G_1 checkpoint and also through the G_2 checkpoint if the DNA has not been adequately repaired.[496,497] Protein p53, whose three-dimensional structure is known,[500,501] binds to DNA and also induces transcription of genes that cause arrest of the cell cycle. It may also induce cell death (apoptosis), a process that may also require the protein product of protooncogene c-*myc*.[502-505] A variety of protein kinases and phosphatases act on p53 and influence its activity.[506]

References

1. Krebs, H. A. (1957) *Endeavour* **16**, 125 – 132
2. Newsholme, E. A., and Start, C. (1973) *Regulation in Metabolism*, Wiley, New York
3. Heinrich, R., Rapoport, S. M., and Rapoport, T. A. (1977) *Prog. Biophys. and Mol. Biol.* **32**, 1 – 82
4. Martin, B. R. (1987) *Metabolic Regulation*, Blackwell Scientific Publ., Oxford
5. Chance, B., and Williams, G. R. (1955) *J. Biol. Chem.* **217**, 477 – 488
6. Heinrich, R., and Rapoport, T. A. (1974) *Eur. J. Biochem.* **42**, 89 – 95; 97 – 105
7. Mendes, P. (1997) *Trends Biochem. Sci.* **22**, 361 – 363
8. Wright, B. E. (1986) *Trends Biochem. Sci.* **11**, 164 – 165
9. Savageau, M. A. (1987) *Trends Biochem. Sci.* **12**, 219 – 220
10. Voit, E. O., and Savageau, M. A. (1987) *Biochemistry* **26**, 6869 – 6880
11. Hlavacek, W. S., and Savageau, M. A. (1996) *J. Mol. Biol.* **255**, 121 – 139
12. Voit, E. O. (1987) *Trends Biochem. Sci.* **12**, 221
13. Kacser, H., and Burns, J. A. (1979) *Biochem. Soc. Trans.* **7**, 1149 – 1166
14. Crabtree, B., and Newsholme, E. A. (1987) *Trends Biochem. Sci.* **12**, 4 – 12
15. Kholodenko, B. N., and Westerhoff, H. V. (1995) *Trends Biochem. Sci.* **20**, 52 – 54
16. Fell, D. (1997) *Understanding the Control of Metabolism*, Portland Press, London
17. Quant, P. A. (1993) *Trends Biochem. Sci.* **18**, 26 – 30
18. Salter, M., Knowles, R. G., and Pogson, C. I. (1994) in *Essays in Biochemistry*, Vol. 28 (Tipton, K. F., ed), Portland Press, London
19. Schulze, E.-D. (1994) *Flux Control in Biological Systems*, Academic Press, San Diego, California
20. Kacser, H., and Burns, J. A. (1973) *Symp. Soc. Exp. Biol* **XXVII**, 65 – 104
21. Stark, G. R., and Wahl, G. M. (1984) *Ann. Rev. Biochem.* **53**, 447 – 491
22. Wilkinson, J. H. (1970) *Isoenzymes*, 2nd ed., Lippincott, Philadelphia, Pennsylvania
23. Markert, C. L., ed. (1975) *Isozymes*, Vol. 4, Academic Press, New York
24. Christen, P., and Metzler, D. E., eds. (1985) *Transaminases*, Wiley, New York
25. Salerno, C., Giartosio, A., and Fasella, P. (1986) in *Vitamin B_6 Pyridoxal Phosphate*, Vol. 1b (Dolphin, D., Poulson, R., and Avramovic, O., eds), Wiley, New York
26. McPhalen, C. A., Vincent, M. G., and Jansonius, J. N. (1992) *J. Mol. Biol.* **225**, 495 – 517
27. Mehta, P. K., Hale, T. I., and Christen, P. (1989) *Eur. J. Biochem.* **186**, 249 – 253
28. Kamitori, S., Okamoto, A., Hirotsu, K., Higuchi, T., Kuramitsu, S., Kagamiyama, H., Matsuura, Y., and Katsube, Y. (1990) *J. Biochem.* **108**, 175 – 184
29. Hogrefe, H. H., Griffith, J. P., Rossman, M. G., and Goldberg, E. (1987) *J. Biol. Chem.* **262**, 13155 – 13162
30. Purich, D. L., and Fromm, H. J. (1972) *Curr. Top. Cell. Regul.* **6**, 131 – 167
31. Huijing, F. (1979) *Trends Biochem. Sci.* **4**, N132 – N134

31a. Aleshin, A. E., Kirby, C., Liu, X., Bourenkov, G. P., Bartunik, H. D., Fromm, H. J., and Honzatko, R. B. (2000) *J. Mol. Biol.* **296**, 1001 – 1015

32. Purich, D. L., Fromm, H. J., and Rudolph, F. R. (1973) *Adv. Enzymol.* **39**, 249 – 326
33. Breitbar, R. E., Andreadis, A., and Nadal - Ginard, B. (1987) *Ann. Rev. Biochem.* **56**, 467 – 495
34. Joh, K., Arai, Y., Mukai, T., and Hori, K. (1986) *J. Mol. Biol.* **190**, 401 – 410
35. Haldane, J. B. S. (1930) *Enzymes*, Longmans, Green, New York
36. Hayashi, S.-i, Murakami, Y., and Matsufuji, S. (1996) *Trends Biochem. Sci.* **21**, 27 – 30
37. Li, X., Stebbins, B., Hoffman, L., Pratt, G., Rechsteiner, M., and Coffino, P. (1996) *J. Biol. Chem.* **271**, 4441 – 4446
38. Sprang, S. R., Acharya, K. R., Goldsmith, E. J., Stuart, D. I., Varvill, K., Fletterick, R. J., Madsen, N. B., and Johnson, L. N. (1988) *Nature (London)* **336**, 215 – 221
39. Cohen, P. (1983) *Control of Enzyme Activity*, 2nd ed., Chapman and Hall, London
40. Kantrowitz, E. R., and Lipscomb, W. N. (1990) *Trends Biochem. Sci.* **15**, 53 – 59
41. Zhou, B.-B., and Schachman, H. K. (1993) *Protein Sci.* **2**, 103 – 112
42. Xi, X. G., De Staercke, C., Van Vliet, F., Triniolles, F., Jacobs, A., Stas, P. P., Ladjimi, M. M., Simon, V., Cunin, R., and Hervé, G. (1994) *J. Mol. Biol.* **242**, 139 – 149
43. Uyeda, K., Furuya, E., and Sherry, A. D. (1981) *J. Biol. Chem.* **256**, 8679 – 8684
44. Hers, H., Hue, L., and Schaftingen, E. (1982) *Trends Biochem. Sci.* **7**, 329 – 331
45. Lively, M. O., El-Maghrabi, M. R., Pilkis, J., D'Angelo, G., Colosia, A. D., Ciavola, J.-A., Fraser, B. A., and Pilkis, S. J. (1988) *J. Biol. Chem.* **263**, 839 – 849
46. Kitamura, K., and Uyeda, K. (1988) *J. Biol. Chem.* **263**, 9027 – 9033

46a. Choe, J.-Y., Fromm, H. J., and Honzatko, R. B. (2000) *Biochemistry* **39**, 8565 – 8574

47. Graves, D. J., Martin, B. L., and Wang, J. H. (1994) *Co – and Post – Translational Modification of Proteins*, Oxford Univ. Press, New York

47a. Skamnaki, V. T., Owen, D. J., Noble, M. E. M., Lowe, E. D., Lowe, G., Oikonomakos, N. G., and Johnson, L. N. (1999) *Biochemistry* **38**, 14718 – 14730

48. Johnson, L. N. (1992) *FASEB J.* **6**, 2274 – 2282

48a. Cheng, A., Fitzgerald, T. J., Bhatnagar, D., Roskoski, R., Jr., and Carlson, G. M. (1988) *J. Biol. Chem.* **263**, 5534 – 5542

49. Dent, P., Lavoinne, A., Nakielny, S., Caudwell, F. B., Watt, P., and Cohen, P. (1990) *Nature (London)* **348**, 302 – 308

49a. Brady, M. J., Bourbonais, F. J., and Saltiel, A. R. (1998) *J. Biol. Chem.* **273**, 14063 – 14066

50. Krebs, E. G., and Fischer, E. H. (1956) *Biochim. Biophys. Acta.* **20**, 150 – 157
51. Edelman, A. M., Blumenthal, D. K., and Krebs, E. G. (1987) *Ann. Rev. Biochem.* **56**, 567 – 613
52. Allen, J. F. (1992) *Trends Biochem. Sci.* **17**, 12 – 17
53. Hershey, J. W. B. (1989) *J. Biol. Chem.* **264**, 20823 – 20826
54. Kennelly, P. J., and Krebs, E. G. (1991) *J. Biol. Chem.* **266**, 15555 – 15558
55. Hanks, S. K., and Hunter, T. (1995) *FASEB J.* **9**, 576 – 596
56. Woodgett, J. R., ed. (1994) *Protein Kinases*, IRL Press, Oxford
57. Hunter, T. (1994) *Sem. Cell Biol.* **5**, 367 – 376
58. Kemp, B. E., Parker, M. W., Hu, S., Tiganis, T., and House, C. (1994) *Trends Biochem. Sci.* **19**, 440 – 444
59. Imazu, M., Stricklund, W. G., Chrisman, T. D., and Exton, J. H. (1984) *J. Biol. Chem.* **259**, 1813 – 1821
60. He, X., Saint-Jeannet, J.-P., Woodgett, J. R., Varmus, H. E., and David, I. B. (1995) *Nature (London)* **374**, 617 – 622
61. Hartl, F. T., and Roskoski, R., Jr. (1983) *J. Biol. Chem.* **258**, 3950 – 3955
62. Zheng, J., Knighton, D. R., Xuong, N.-H., Taylor, S. S., Sowadski, J. M., and Eyck, L. F. T. (1993) *Protein Sci.* **2**, 1559 – 1573
63. Beebe, S. J., Reimann, E. M., and Schlender, K. K. (1984) *J. Biol. Chem.* **259**, 1415 – 1422
64. Taylor, S. S., Knighton, D. R., Zheng, J., Sowadski, J. M., Gibbs, C. S., and Zoller, M. J. (1993) *Trends Biochem. Sci.* **18**, 84 – 89
65. Narayana, N., Cox, S., Shaltiel, S., Taylor, S. S., and Xuong, N.-h. (1997) *Biochemistry* **36**, 4438 – 4448

References

66. Su, Y., Dostmann, W. R. G., Herberg, F. W., Durick, K., Xuong, N.-h, Eyck, L. T., Taylor, S. S., and Varughese, K. I. (1995) *Science* **269**, 807–813

67. Lincoln, T. M., Thompson, M., and Cornwell, T. L. (1988) *J. Biol. Chem.* **263**, 17632–17637

68. Weber, I. T., Shabb, J. B., and Corbin, J. D. (1989) *Biochemistry* **28**, 6122–6127

69. Fry, M. J., Panayotou, G., Booker, G. W., and Waterfield, M. D. (1993) *Protein Sci.* **2**, 1785–1797

70. Xu, R.-M., Carmel, G., Sweet, R. M., Kuret, J., and Cheng, X. (1995) *EMBO J.* **14**, 1015–1023

71. Allende, J. E., and Allende, C. C. (1995) *FASEB J.* **9**, 313–323

72. Seldin, D. C., and Leder, P. (1995) *Science* **267**, 894–897

72a. Niefind, K., Guerra, B., Pinna, L. A., Issinger, O.-G., and Schomburg, D. (1998) *EMBO J.* **17**, 2451–2462

73. Chan, K.- F., and Graves, D. J. (1984) in *Calcium and Cell Function*, Vol. 5 (Cheung, W. Y., ed), pp. 1–32, Academic Press, New York

74. Heilmeyer, L. M. G., Jr. (1991) *Biochim. Biophys. Acta.* **1094**, 168–174

75. Norcum, M. T., Wilkinson, D. A., Carlson, M. C., Hainfeld, J. F., and Carlson, G. M. (1994) *J. Mol. Biol.* **241**, 94–102

76. Kee, S. M., and Graves, D. J. (1987) *J. Biol. Chem.* **262**, 9448–9453

77. Paudel, H. K., and Carlson, G. M. (1987) *J. Biol. Chem.* **262**, 11912–11915

78. Cheng, A., Fitzgerald, T. J., Bhatnagar, D., Roskoski, R., Jr., and Carlson, G. M. (1988) *J. Biol. Chem.* **263**, 5534–5542

79. Owen, D. J., Papageorgiou, A. C., Garman, E. F., Noble, M. E. M., and Johnson, L. N. (1995) *J. Mol. Biol.* **246**, 374–381

80. Ohmstede, C.-A., Bland, M. M., Merrill, B. M., and Sahyoun, N. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5784–5788

81. Hanson, P. I., and Schulman, H. (1992) *Ann. Rev. Biochem.* **61**, 559–601

82. Harper, J. F., Sussman, M. R., Schaller, G. E., Putnam-Evans, C., Charbonneau, H., and Harmon, A. C. (1991) *Science* **252**, 951–954

83. Ullrich, A., Bell, J. R., Chen, E. Y., Herrera, R., Petruzzelli, L. M., Dull, T. J., Gray, A., Coussens, L., Liao, Y.-C., Tsubokawa, M., Mason, A., Seeburg, P. H., Grunfield, C., Roden, O. M., and Ramachandran, J. (1985) *Nature (London)* **313**, 756–761

84. Wang, J. Y. J. (1994) *Trends Biochem. Sci.* **19**, 373–376

85. Clark, E. A., Shattil, S. J., and Brugge, J. S. (1994) *Trends Biochem. Sci.* **19**, 464–469

86. Sun, H., and Tonks, N. K. (1994) *Trends Biochem. Sci.* **19**, 480–485

87. Ingebritsen, T. S., and Cohen, P. (1983) *Science* **221**, 331–338

88. Cohen, P. (1989) *Ann. Rev. Biochem.* **58**, 453–508

89. Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A. C., and Kuriyan, J. (1995) *Nature (London)* **376**, 745–753

90. Fischer, E. H., Charbonneau, H., and Tonks, N. K. (1991) *Science* **253**, 401–406

91. Yuvaniyama, J., Denu, J. M., Dixon, J. E., and Saper, M. A. (1996) *Science* **272**, 1328–1331

92. MacKintosh, C., and MacKintosh, R. W. (1994) *Trends Biochem. Sci.* **19**, 444–448

93. Hidaka, H., and Kobayashi, R. (1994) in *Essays in Biochemistry*, Vol. 28 (Tipton, K. F., ed), pp. 73–98, Portland Press, London

94. Cohen, P., Holmes, C. F. B., and Tsukitani, Y. (1990) *Trends Biochem. Sci.* **15**, 98–102

95. Price, N. E., and Mumby, M. C. (2000) *Biochemistry* **39**, 11312–11318

96. Zhang, M., Van Etten, R. L., and Stauffacher, C. V. (1994) *Biochemistry* **33**, 11097–11105

97. Barford, D., Flint, A. J., and Tonks, N. K. (1994) *Science* **263**, 1397–1403

98. Stone, R. L., and Dixon, J. E. (1994) *J. Biol. Chem.* **269**, 31323–31326

99. Chen, W., Wilborn, M., and Rudolph, J. (2000) *Biochemistry* **39**, 10781–10789

100. Kissinger, C. R., Parge, H. E., Knighton, D. R., Lewis, C. T., Pelletier, L. A., Tempczyk, A., Kalish, V. J., Tucker, K. D., Showalter, R. E., Moomaw, E. W., Gastinel, L. N., Habuka, N., Chen, X., Maldonado, F., Barker, J. E., Bacquet, R., and Villafranca, J. E. (1995) *Nature (London)* **378**, 641–644

101. LaPorte, D. C., and Chung, T. (1985) *J. Biol. Chem.* **260**, 15291–15297

102. Pilkis, S. J., Claus, T. H., Kurland, I. J., and Lange, A. J. (1995) *Ann. Rev. Biochem.* **64**, 799–835

103. Istvan, E. S., Hasemann, C. A., Kurumbail, R. G., Uyeda, K., and Deisenhofer, J. (1995) *Protein Sci.* **4**, 2439–2441

104. Vertommen, D., Bertrand, L., Sontag, B., Di Pietro, A., Louckx, M. P., Vidal, H., Hue, L., and Rider, M. H. (1996) *J. Biol. Chem.* **271**, 17875–17880

105. Abe, Y., Minami, Y., Li, Y., Nguyen, C., and Uyeda, K. (1995) *Biochemistry* **34**, 2553–2559

106. Ogushi, S., Lawson, J. W. R., Dobson, G. P., Veech, R. L., and Uyeda, K. (1990) *J. Biol. Chem.* **265**, 10943–10949

107. Nimmo, H. G. (1984) *Trends Biochem. Sci.* **9**, 475–478

108. Dean, A. M., and Koshland, D. E., Jr. (1993) *Biochemistry* **32**, 9302–9309

109. Dean, A. M., and Koshland, D. E., Jr. (1990) *Science* **249**, 1044–1046

110. Saier, M. H., Jr., Wu, L.-F., and Reizer, J. (1990) *Trends Biochem. Sci.* **15**, 391–395

110a. Song, H. K., Lee, J. Y., Lee, M. G., Moon, J., Min, K., Yang, J. K., and Suh, S. W. (1999) *J. Mol. Biol.* **293**, 753–761

111. Swanson, R. V., Alex, L. A., and Simon, M. I. (1994) *Trends Biochem. Sci.* **19**, 485–490

112. Saier, M. H., Jr. (1989) *Microbiol. Rev.* **53**, 109–120

113. Stock, J. B., Stock, A. M., and Mottonen, J. M. (1990) *Nature (London)* **344**, 395–400

114. Boon, P., Chock, P. B., Rhee, S. G., and Stadtman, E. R. (1980) *Ann. Rev. Biochem.* **49**, 813–843

115. Mura, U., Chock, P. B., and Stadtman, E. R. (1981) *J. Biol. Chem.* **256**, 13022–13029

116. Almassy, R. J., Janson, C. A., Hamlin, R., Xuong, N.-H., and Eisenberg, D. (1986) *Nature (London)* **323**, 304–309

117. Garcia, E., and Rhee, S. G. (1983) *J. Biol. Chem.* **258**, 2246–2253

118. Moss, J., and Vaughan, M. (1995) *J. Biol. Chem.* **270**, 12327–12330

119. Amor, J. C., Harrison, D. H., Kahn, R. A., and Ringe, D. (1994) *Nature (London)* **372**, 704–708

120. Boman, A. L., and Kahn, R. A. (1995) *Trends Biochem. Sci.* **20**, 147–150

121. Lee, F.-J. S., Stevens, L. A., Hall, L. M., Murtagh, J. J., Jr., Kao, Y. L., Moss, J., and Vaughan, M. (1994) *J. Biol. Chem.* **269**, 21555–21560

122. Moss, J., Stanley, S. J., Nightingale, M. S., Murtagh, J. J., Jr., Monaco, L., Mishima, K., Chen, H.-C., Williamson, K.-C., and Tsai, S.-C. (1992) *J. Biol. Chem.* **267**, 10481–10488

123. Guse, A. H., da Silva, C. P., Weber, K., Ashamu, G. A., Potter, B. V. L., and Mayr, G. W. (1996) *J. Biol. Chem.* **271**, 23946–23953

124. Gaal, J. C., Smith, K. R., and Pearson, C. K. (1987) *Trends Biochem. Sci.* **12**, 129–130

125. Huttner, W. B. (1987) *Trends Biochem. Sci.* **12**, 361–363

126. Sakakibara, Y., Takami, Y., Zwieb, C., Nakayama, T., Suiko, M., Nakajima, H., and Liu, M.-C. (1995) *J. Biol. Chem.* **270**, 30470–30478

127. Rosenquist, G. L., and Nicholas, H. B., Jr. (1993) *Protein Sci.* **2**, 215–222

128. Sundaram, K. S., and Lev, M. (1992) *J. Biol. Chem.* **267**, 24041–24044

129. Hooper, L. V., Manzella, S. M., and Baenziger, J. U. (1996) *FASEB J.* **10**, 1137–1146

130. Terwilliger, T. C., and Koshland, D. E., Jr. (1984) *J. Biol. Chem.* **259**, 7719–7725

131. Favre, B., Zolnierowicz, S., Turowski, P., and Hemmings, B. A. (1994) *J. Biol. Chem.* **269**, 16311–16317

132. Edmondson, D. G., and Roth, S. Y. (1996) *FASEB J.* **10**, 1173–1182

133. Kuo, M.-H., Brownell, J. E., Sobel, R. E., Ranalli, T. A., Cook, R. G., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996) *Nature (London)* **383**, 269–272

134. Brandes, H. K., Larimer, F. W., and Hartman, F. C. (1996) *J. Biol. Chem.* **271**, 3333–3335

135. Brandes, H. K., Hartman, F. C., Lu, T. -Y. S., and Larimer, F. W. (1996) *J. Biol. Chem.* **271**, 6490–6496

136. Swerdlow, R. D., Green, C. L., Setlow, B., and Setlow, D. (1979) *J. Biol. Chem.* **254**, 6835–6837

137. Hentze, M. W., Rouault, T. A., Harford, J. B., and Klausner, R. D. (1989) *Science* **244**, 357–359

138. Wadsworth, W. G., and Riddle, D. L. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8435–8438

139. Boron, W. F. (1984) *Nature (London)* **312**, 312

140. Siffert, W., and Akkerman, J. W. N. (1988) *Trends Biochem. Sci.* **13**, 148–151

141. Reinertsen, K. V., Tonnessen, T. I., Jacobsen, J., Sandvig, K., and Olsnes, S. (1988) *J. Biol. Chem.* **263**, 11117–11125

142. Schramm, V. L. (1982) *Trends Biochem. Sci.* **7**, 369–371

143. Srere, P. A. (1987) *Ann. Rev. Biochem.* **56**, 89–124

144. Srivastava, D. K., and Bernhard, S. A. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 175–204

145. Ovádi, J. (1988) *Trends Biochem. Sci.* **13**, 486–490

146. Kaprelyants, A. S. (1988) *Trends Biochem. Sci.* **13**, 43–46

146a. Miles, E. W., Rhee, S., and Davies, D. R. (1999) *J. Biol. Chem.* **274**, 12193–12196

146b. Serre, V., Guy, H., Penverne, B., Lux, M., Rotgeri, A., Evans, D., and Hervé, G. (1999) *J. Biol. Chem.* **274**, 23794–23801

146c. Geck, M. K., and Kirsch, J. F. (1999) *Biochemistry* **38**, 8032–8037

146d. Dell'Acqua, M. L., and Scott, J. D. (1997) *J. Biol. Chem.* **272**, 12881–12884

147. Ovádi, J., and Srere, P. A. (1992) *Trends Biochem. Sci.* **17**, 445–447

148. Inagaki, N., Ito, M., Nakano, T., and Inagaki, M. (1994) *Trends Biochem. Sci.* **19**, 448–452

149. Faux, M. C., and Scott, J. D. (1996) *Trends Biochem. Sci.* **21**, 312–315

150. Fraser, C. M. (1989) *J. Biol. Chem.* **264**, 9266–9270

151. Strosberg, A. D. (1993) *Protein Sci.* **2**, 1198–1209

152. Scheel, A. A., and Pelham, H. R. B. (1996) *Biochemistry* **35**, 10203–10209

153. Levitzki, A. (1988) *Science* **241**, 800–806

153a. Stacey, M., Lin, H.-H., Gordon, S., and McKnight, A. J. (2000) *Trends Biochem. Sci.* **25**, 284–289

153b. Leurs, R., Smit, M. J., Alewijnse, A. E., and Timmerman, H. (1998) *Trends Biochem. Sci.* **23**, 418–422

153c. Bockaert, J., and Pin, J. P. (1999) *EMBO J.* **18**, 1723–1729

154. Insel, P. A., and Wasserman, S. I. (1990) *FASEB J.* **4**, 2732–2736

155. Turki, J., Lorenz, J. N., Green, S. A., Donnelly, E. T., Jacinto, M., and Liggett, S. B. (1996) *Proc. Natl. Acad. Sci.* U.S.A. **93**, 10483–10488

155a. Rohrer, D. K., Chruscinski, A., Schauble, E. H., Bernstein, D., and Kobilka, B. K. (1999) *J. Biol. Chem.* **274**, 16701 – 16708

155b. Limbird, L. E., and Vaughan, D. E. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7125 – 7127

156. Premont, R. T., Inglese, J., and Lefkowitz, R. J. (1995) *FASEB J.* **9**, 175 – 182

156a. Pitcher, J. A., Freedman, N. J., and Lefkowitz, R. J. (1998) *Ann. Rev. Biochem.* **67**, 653 – 692

157. Mochly-Rosen, D. (1995) *Science* **268**, 247 – 251

158. Benovic, J. L., DeBlasi, A., Stone, W. C., Caron, M. G., and Lefkowitz, R. J. (1989) *Science* **246**, 235 – 240

158a. Danner, S., Frank, M., and Lohse, M. J. (1998) *J. Biol. Chem.* **273**, 3223 – 3229

158b. Gabilondo, A. M., Hegler, J., Krasel, C., Boivin-Jahns, V., Hein, L., and Lohse, M. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12285 – 12290

159. Rodbell, M., Birnbaumer, L., Pohl, S. L., and Krans, H. M. J. (1971) *J. Biol. Chem.* **246**, 1877 – 1882

160. Unson, C. G., Cypess, A. M., Wu, C.-R., Goldsmith, P. K., Merrifield, R. B., and Sakmar, T. P. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 310 – 315

161. Rosenthal, W., Antaramian, A., Gilbert, S., and Birnbaumer, M. (1993) *J. Biol. Chem.* **268**, 13030 – 13033

162. Sharif, M., and Hanley, M. R. (1992) *Nature (London)* **357**, 279 – 280

163. McFarland, K. C., Sprengel, R., Phillips, H. S., Köhler, M., Rosemblit, N., Nikolics, K., Segaloff, D. L., and Seeburg, P. H. (1989) *Science* **245**, 494 – 499

164. Kudo, M., Osuga, Y., Kobilka, B. K., and Hsueh, A. J. W. (1996) *J. Biol. Chem.* **271**, 22470 – 22478

165. Perlman, J. H., Laakkonen, L., Osman, R., and Gershengorn, M. C. (1994) *J. Biol. Chem.* **269**, 23383 – 23386

166. Conti, M., Iona, S., Cuomo, M., Swinnen, J. V., Odeh, J., and Svoboda, M. E. (1995) *Biochemistry* **34**, 7979 – 7987

167. Smith, K. J., Scotland, G., Beattie, J., Trayer, I. P., and Houslay, M. D. (1996) *J. Biol. Chem.* **271**, 16703 – 16711

168. Sutherland, E. W. (1972) *Science* **177**, 401 – 408

169. Pastan, I. (1972) *Sci. Am.* **227**(Aug), 97 – 105

170. Lalli, E., and Sassone-Corsi, P. (1994) *J. Biol. Chem.* **269**, 17359 – 17362

171. Bolwell, G. P. (1995) *Trends Biochem. Sci.* **20**, 492 – 495

172. Taussig, R., and Gilman, A. G. (1995) *J. Biol. Chem.* **270**, 1 – 4

173. Dessauer, C. W., Tesmer, J. J. G., Sprang, S. R., and Gilman, A. G. (1998) *J. Biol. Chem.* **273**, 25831 – 25839

174. Cooper, D. M. F., Mons, N., and Karpen, J. W. (1995) *Nature (London)* **374**, 421 – 424

175. Désaubry, L., Shoshani, I., and Johnson, R. A. (1996) *J. Biol. Chem.* **271**, 14028 – 14034

176. Hurley, J. H. (1999) *J. Biol. Chem.* **274**, 7599 – 7602

177. Hellevuo, K., Yoshimura, M., Mons, N., Hoffman, P. L., Cooper, D. M. F., and Tabakoff, B. (1995) *J. Biol. Chem.* **270**, 11581 – 11589

178. Walseth, T. F., Gander, J. E., Eide, S. J., Krick, T. P., and Goldberg, N. D. (1983) *J. Biol. Chem.* **258**, 1544 – 1558

178a. Chen, Y., Cann, M. J., Litvin, T. N., Lourgenko, V., Sinclair, M. L., Levin, L. R., and Buck, J. (2000) *Science* **289**, 625 – 628

178b. Kaupp, U. B., and Weyand, I. (2000) *Science* **289**, 559 – 560

179. Hamm, H. E. (1998) *J. Biol. Chem.* **273**, 669 – 672

180. Neer, E. J. (1994) *Protein Sci.* **3**, 3 – 14

181. Collins, S., Caron, M. G., and Lefkowitz, R. J. (1992) *Trends Biochem. Sci.* **17**, 37 – 39

182. Hepler, J. R., and Gilman, A. G. (1992) *Trends Biochem. Sci.* **17**, 383 – 387

183. Levitzki, A. (1988) *Trends Biochem. Sci.* **13**, 298 – 301

184. Rawls, R. L. (1987) *Chem. Eng. News* **65 December 21**, 26 – 39

185. Wang, N., Yan, K., and Rasenick, M. M. (1990) *J. Biol. Chem.* **265**, 1239 – 1242

186. Wang, Q., Mullah, B. K., and Robishaw, J. D. (1999) *J. Biol. Chem.* **274**, 17365 – 17371

187. Mayorga, L. S., Diaz, R., and Stahl, P. D. (1989) *Science* **244**, 1475 – 1477

188. Coleman, D. E., and Sprang, S. R. (1996) *Trends Biochem. Sci.* **21**, 41 – 44

189. Ray, K., Kunsch, C., Bonner, L. M., and Robishaw, J. D. (1995) *J. Biol. Chem.* **270**, 21765 – 21771

190. Brown, A. M. (1991) *FASEB J.* **5**, 2175 – 2179

191. Chen, J., DeVivo, M., Dingus, J., Harry, A., Li, J., Sui, J., Carty, D. J., Blank, J. L., Exton, J. H., Stoffel, R. H., Inglese, J., Lefkowitz, R. J., Logothetis, D. E., Hildebrandt, J. D., and Iyengar, R. (1995) *Science* **268**, 1166 – 1169

192. Schulz, K., Danner, S., Bauer, P., Schröder, S., and Lohse, M. J. (1996) *J. Biol. Chem.* **271**, 22546 – 22551

193. Yoshida, T., Willardson, B. M., Wilkins, J. F., Jensen, G. J., Thornton, B. D., and Bitensky, M. W. (1994) *J. Biol. Chem.* **269**, 24050 – 24057

193a. Whiteway, M., Hougan, L., Dignard, D., Thomas, D. Y., Bell, L., Saari, G. C., Grant, F. J., O'Hara, P., and MacKay, V. L. (1989) *Cell* **56**, 467 – 477

194. Noel, J. P., Hamm, H. E., and Sigler, P. B. (1993) *Nature (London)* **366**, 654 – 663

195. Lambright, D. G., Sondek, J., Bohm, A., Skiba, N. P., Hamm, H. E., and Sigler, P. B. (1996) *Nature (London)* **379**, 311 – 319

196. McLaughlin, S. K., McKinnon, P. J., and Margolskee, R. F. (1992) *Nature (London)* **357**, 563 – 569

197. Hall, A. (1990) *Science* **249**, 635 – 640

198. Valencia, A., Chardin, P., Wittinghofer, A., and Sander, C. (1991) *Biochemistry* **30**, 4637 – 4648

199. Macara, I. G., Lounsbury, K. M., Richards, S. A., McKiernan, C., and Bar-Sagi, D. (1996) *FASEB J.* **10**, 625 – 630

200. Boguski, M. S., and McCormick, F. (1993) *Nature (London)* **366**, 643 – 654

200a. Scita, G., Tenca, P., Frittoli, E., Tocchetti, A., Innocenti, M., Giardina, G., and Di Fiore, P. P. (2000) *EMBO J.* **19**, 2393 – 2398

201. Jurnak, F. (1985) *Science* **230**, 32 – 36

202. Sprinzl, M. (1994) *Trends Biochem. Sci.* **19**, 245 – 250

203. Berchtold, H., Reshetnikova, L., Reiser, C. O. A., Schirmer, N. K., Sprinzl, M., and Hilgenfeld, R. (1993) *Nature (London)* **365**, 126 – 132

204. Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S., and Leberman, R. (1996) *Nature (London)* **379**, 511 – 518

205. Jurnak, F. (1988) *Trends Biochem. Sci.* **13**, 195 – 198

206. Downward, J. (1990) *Trends Biochem. Sci.* **15**, 469 – 472

207. Wittinghofer, A., and Pai, E. F. (1991) *Trends Biochem. Sci.* **16**, 382 – 387

208. Schlichting, I., Almo, S. C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E. F., Petsko, G. A., and Goody, R. S. (1990) *Nature (London)* **345**, 309 – 315

209. Kraulis, P. J., Domaille, P. J., Campbell-Burk, S. L., Van Aken, T., and Laue, E. D. (1994) *Biochemistry* **33**, 3515 – 3531

210. Bokoch, G. M., and Der, C. J. (1993) *FASEB J.* **7**, 750 – 759

211. Takai, Y., Sasaki, T., Tanaka, K., and Nakanishi, H. (1995) *Trends Biochem. Sci.* **20**, 227 – 231

212. Fischer von Mollard, G., Stahl, B., Li, C., Südhof, T. C., and Jahn, R. (1994) *Trends Biochem. Sci.* **19**, 164 – 168

212a. Burridge, K. (1999) *Science* **283**, 2028 – 2029

212b. Lin, R., Cerione, R. A., and Manor, D. (1999) *J. Biol. Chem.* **274**, 23633 – 23641

212c. Desrosiers, R. R., Gauthier, F., Lanthier, J., and Béliveau, R. (2000) *J. Biol. Chem.* **275**, 14949 – 14957

212d. Schwartz, M. A., and Shattil, S. J. (2000) *Trends Biochem. Sci.* **25**, 388 – 391

212e. Zhang, B., Zhang, Y., Wang, Z.-x, and Zheng, Y. (2000) *J. Biol. Chem.* **275**, 25299 – 25307

213. Symons, M. (1996) *Trends Biochem. Sci.* **21**, 178 – 181

214. Ando, S., Kaibuchi, K., Sasaki, T., Hiraoka, K., Nishiyama, T., Mizuno, T., Asada, M., Nunoi, H., Matsuda, I., Matsuura, Y., Polakis, P., McCormick, F., and Takai, Y. (1992) *J. Biol. Chem.* **267**, 25709 – 25713

215. Nuoffer, C., and Balch, W. E. (1994) *Ann. Rev. Biochem.* **63**, 949 – 990

216. Rybin, V., Ullrich, O., Rubino, M., Alexandrov, K., Simon, I., Seabra, M. C., Goody, R., and Zerial, M. (1996) *Nature (London)* **383**, 266 – 269

217. Garrett, M. D., Self, A. J., van Oers, C., and Hall, A. (1989) *J. Biol. Chem.* **264**, 10 – 13

217a. Moss, J., and Vaughan, M. (1998) *J. Biol. Chem.* **273**, 21431 – 21434

217b. Hutchinson, J. P., and Eccleston, J. F. (2000) *Biochemistry* **39**, 11348 – 11359

218. Scheffzek, K., Ahmadian, M. R., and Wittinghofer, A. (1998) *Trends Biochem. Sci.* **23**, 257 – 262

218a. Berman, D. M., and Gilman, A. G. (1998) *J. Biol. Chem.* **273**, 1269 – 1272

218b. Coleman, D. E., and Sprang, S. R. (1999) *J. Biol. Chem.* **274**, 16669 – 16672

219. Schalk, I., Zeng, K., Wu, S.-K., Stura, E. A., Matteson, J., Huang, M., Tandon, A., Wilson, I. A., and Balch, W. E. (1996) *Nature (London)* **381**, 42 – 48

220. Buss, J. E., Mumby, S. M., Casey, P. J., Gilman, A. G., and Sefton, B. M. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7493 – 7497

221. Wedegaertner, P. B., Wilson, P. T., and Bourne, H. R. (1995) *J. Biol. Chem.* **270**, 503 – 506

222. Song, J., Hirschman, J., Gunn, K., and Dohlman, H. G. (1996) *J. Biol. Chem.* **271**, 20273 – 20283

223. Hepler, J. R., Biddlecome, G. H., Kleuss, C., Camp, L. A., Hofmann, S. L., Ross, E. M., and Gilman, A. G. (1996) *J. Biol. Chem.* **271**, 496 – 504

224. Peseckis, S. M., and Resh, M. D. (1994) *J. Biol. Chem.* **269**, 30888 – 30892

225. Giner, J.-L., and Rando, R. R. (1994) *Biochemistry* **33**, 15116 – 15123

226. Liu, L., Dudler, T., and Gelb, M. H. (1996) *J. Biol. Chem.* **271**, 23269 – 23276

227. Casey, P. J. (1995) *Science* **268**, 221 – 225

228. Maltese, W. A., Sheridan, K. M., Repko, E. M., and Erdman, R. A. (1990) *J. Biol. Chem.* **265**, 2148 – 2155

229. Milligan, G., Parenti, M., and Magee, A. I. (1995) *Trends Biochem. Sci.* **20**, 181 – 186

230. Duncan, J. A., and Gilman, A. G. (1996) *J. Biol. Chem.* **271**, 23594 – 23600

230a. Iiri, T., Backlund, P. S., Jr., Jones, T. L. Z., Wedegaertner, P. B., and Bourne, H. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14592 – 14597

230b. Lee, T. W., Seifert, R., Guan, X., and Kobilka, B. K. (1999) *Biochemistry* **38**, 13801 – 13809

231. Lambright, D. G., Noel, J. P., Hamm, H. E., and Sigler, P. B. (1994) *Nature (London)* **369**, 621 – 628

232. Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E., and Sigler, P. B. (1994) *Nature (London)* **372**, 276 – 279

References

233. Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G., and Sprang, S. R. (1994) *Science* **265**, 1405–1412
234. Mixon, M. B., Lee, E., Coleman, D. E., Berghuis, A. M., Gilman, A. G., and Sprang, S. R. (1995) *Science* **270**, 954–960
235. Kjeldgaard, M., Nyborg, J., and Clark, B. F. C. (1996) *FASEB J.* **10**, 1347–1368
236. Higashijima, T., Graziano, M. P., Suga, H., Kainosho, M., and Gilman, A. G. (1991) *J. Biol. Chem.* **266**, 3396–3401
237. Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E., and Sigler, P. B. (1996) *Nature (London)* **379**, 369–374
238. Clapham, D. E. (1996) *Nature (London)* **379**, 297–299
239. Neer, E. J., Schmidt, C. J., Nambudripad, R., and Smith, T. F. (1994) *Nature (London)* **371**, 297–300
240. Faber, X. X. (1994) *Structure* **3**, 551–559
241. Pai, E. F., Kabsch, W., Krengel, U., Holmes, K. C., John, J., and Wittinghofer, A. (1989) *Nature (London)* **341**, 209–214
242. Wall, M. A., Posner, B. A., and Sprang, S. R. (1998) *Structure* **6**, 1169–1183
243. Wilson, P. T., and Bourne, H. R. (1995) *J. Biol. Chem.* **270**, 9667–9675
244. Zhao, Y., Brandish, P. E., DiValentin, M., Schelvis, J. P. M., Babcock, G. T., and Marletta, M. A. (2000) *Biochemistry* **39**, 10848–10854
245. Garbers, D. L., and Lowe, D. G. (1994) *J. Biol. Chem.* **269**, 30741–30744
246. Subbaraya, I., Ruiz, C. C., Helekar, B. S., Zhao, X., Gorczyca, W. A., Pettenati, M. J., Rao, P. N., Palczewski, K., and Baehr, W. (1994) *J. Biol. Chem.* **269**, 31080–31089
247. Lincoln, T. M., and Cornwell, T. L. (1993) *FASEB J.* **7**, 328–338
248. Houslay, M. D. (1985) *Trends Biochem. Sci.* **10**, 465–466
249. Garbers, D. L. (1989) *J. Biol. Chem.* **264**, 9103–9106
250. Lowe, D. G., Chang, M.-S., Hellmiss, R., Chen, E., Singh, S., Garbers, D. L., and Goeddel, D. V. (1989) *EMBO J.* **8**, 1377–1384
251. Hamra, F. K., Forte, L. R., Eber, S. L., Pidhorodeckyj, N. V., Krause, W. J., Freeman, R. H., Chin, D. T., Tompkins, J. A., Fok, K. F., Smith, C. E., Duffin, K. L., Siegel, N. R., and Currie, M. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10464–10468
252. Stamler, J. S., Singel, D. J., and Loscalzo, J. (1992) *Science* **258**, 1898–1902
253. Yuen, P. S. T., Doolittle, L. K., and Garbers, D. L. (1994) *J. Biol. Chem.* **269**, 791–793
254. Stone, J. R., and Marletta, M. A. (1996) *Biochemistry* **35**, 1093–1099
255. Mayer, B., Schrammel, A., Klatt, P., Koesling, D., and Schmidt, K. (1995) *J. Biol. Chem.* **270**, 17355–17360
256. Stone, J. R., and Marletta, M. A. (1995) *Biochemistry* **34**, 16397–16403
257. Deinum, G., Stone, J. R., Babcock, G. T., and Marletta, M. A. (1996) *Biochemistry* **35**, 1540–1547
258. Verma, A., Hirsch, D. J., Glatt, C. E., Ronnett, G. V., and Snyder, S. H. (1993) *Science* **259**, 381–384
259. Stock, J. B., Surette, M. G., McCleary, W. R., and Stock, A. M. (1992) *J. Biol. Chem.* **267**, 19753–19756
259a. Kim, K. K., Yokota, H., and Kim, S.-H. (1999) *Nature* (*London*) **400**, 787–792
260. Scott, W. G., Milligan, D. L., Milburn, M. V., Privé, G. G., Yeh, J., Koshland, D. E., Jr., and Kim, S.-H. (1993) *J. Mol. Biol.* **232**, 555–573
261. Danielson, M. A., Biemann, H.-P., Koshland, D. E., Jr., and Falke, J. J. (1994) *Biochemistry* **33**, 6100–6109
262. Yeh, J. I., Biemann, H.-P., Privé, G. G., Pandit, J., Koshland, D. E., Jr., and Kim, S.-H. (1996) *J. Mol. Biol.* **262**, 186–201
263. Danielson, M. A., Bass, R. B., and Falke, J. J. (1997) *J. Biol. Chem.* **272**, 32878–32888
264. Kim, S.-H. (1994) *Protein Sci.* **3**, 159–165
264a. Ottemann, K. M., Xiao, W., Shin, Y.-K., and Koshland, D. E., Jr. (1999) *Science* **285**, 1751–1754
265. Berridge, M. J., and Irvine, R. F. (1989) *Nature (London)* **341**, 197–205
266. Schulman, H., and Lou, L. L. (1989) *Trends Biochem. Sci.* **14**, 62–66
267. Carafoli, E. (1987) *Ann. Rev. Biochem.* **56**, 395–433
268. Rasmussen, H. (1989) *Sci. Am.* **261**(Oct), 66–73
269. Kobilka, B. K., Matsui, H., Kobilka, T. S., Yang-Feng, T. L., Francke, U., Caron, M. G., Lefkowitz, R. J., and Regan, J. W. (1987) *Science* **238**, 650–656
270. Cotecchia, S., Kobilka, B. K., Daniel, K. W., Nolan, R. D., Lapetina, E. Y., Caron, M. G., Lefkowitz, R. J., and Regan, J. W. (1990) *J. Biol. Chem.* **265**, 63–69
271. Ceresa, B. P., and Limbird, L. E. (1994) *J. Biol. Chem.* **269**, 29557–29564
272. Scheer, A., Fanelli, F., Costa, T., De Benedetti, P. G., and Cotecchia, S. (1996) *EMBO J.* **15**, 3566–3578
273. Hwa, J., and Perez, D. M. (1996) *J. Biol. Chem.* **271**, 6322–6327
274. Boyer, J. L., Waldo, G. L., Evans, T., Northup, J. K., Downes, C. P., and Harden, T. K. (1989) *J. Biol. Chem.* **264**, 13917–13922
275. Wu, D., Jiang, H., and Simon, M. I. (1995) *J. Biol. Chem.* **270**, 9828–9832
276. Nishizuka, Y. (1995) *FASEB J.* **9**, 484–496
277. Majerus, P. W. (1992) *Ann. Rev. Biochem.* **61**, 225–250
278. Majerus, P. W., Conolly, T. M., Bansal, V. S., Inhorn, R. C., Ross, T. S., and Lips, D. L. (1988) *J. Biol. Chem.* **263**, 3051–3054
279. Rhee, S. G., Suh, P.-G., Ryu, S.-H., and Lee, S. Y. (1989) *Science* **244**, 546–550
280. Hough, E., Hansen, L. K., Birknes, B., Jynge, K., Hansen, S., Hordvik, A., Little, C., Dodson, E., and Derewenda, Z. (1989) *Nature (London)* **338**, 357–360
281. Essen, L.-O., Perisic, O., Cheung, R., Katan, M., and Williams, R. L. (1996) *Nature (London)* **380**, 595–602
282. Berridge, M. J. (1993) *Nature (London)* **361**, 315–325
283. Parthasarathy, R., and Eisenberg, F. J. (1986) *Biochem. J.* **235**, 313–322
284. Majerus, P. W., Connolly, T. M., Deckmyn, H., Ross, T. S., Bross, T. E., Ishii, H., Bansal, V. S., and Wilson, D. B. (1986) *Science* **234**, 1519–1526
285. Houslay, M. D. (1987) *Trends Biochem. Sci.* **12**, 1–2
286. Nickels, J. T., Jr., Buxeda, R. J., and Carman, G. M. (1994) *J. Biol. Chem.* **269**, 11018–11024
287. Boronenkov, I. V., and Anderson, R. A. (1995) *J. Biol. Chem.* **270**, 2881–2884
288. Hughes, A. R., Takemura, H., and Putney, J. W., Jr. (1988) *J. Biol. Chem.* **263**, 10314–10319
289. Furuichi, T., Yoshikawa, S., Miyawaki, A., Wada, K., Maeda, N., and Mikoshiba, K. (1989) *Nature (London)* **342**, 32–38
289a. Zimmermann, B., and Walz, B. (1999) *EMBO J.* **18**, 3222–3231
290. Yoshikawa, F., Morita, M., Monkawa, T., Michikawa, T., Furuichi, T., and Mikoshiba, K. (1996) *J. Biol. Chem.* **271**, 18277–18284
291. Humbert, J.-P., Matter, N., Artault, J.-C., Köppler, P., and Malviya, A. N. (1996) *J. Biol. Chem.* **271**, 478–485
292. Hennager, D. J., Welsh, M. J., and DeLisle, S. (1995) *J. Biol. Chem.* **270**, 4959–4962
293. Taylor, C. W., and Marshall, I. C. B. (1992) *Trends Biochem. Sci.* **17**, 403–407
294. Gu, Q.-M., and Sih, C. J. (1994) *J. Am. Chem. Soc.* **116**, 7481–7486
295. Berridge, M. J. (1993) *Nature (London)* **365**, 388–389
296. Lee, H. C., Walseth, T. F., Bratt, G. T., Hayes, R. N., and Clapper, D. L. (1989) *J. Biol. Chem.* **264**, 1608–1615
297. Vu, C. Q., Lu, P.-J., Chen, C.-S., and Jacobson, M. K. (1996) *J. Biol. Chem.* **271**, 4747–4754
298. Dennis, E. A., Rhee, S. G., Billah, M. M., and Hannun, Y. A. (1991) *FASEB J.* **5**, 2068–2077
299. Sternweis, P. C., and Smrcka, A. V. (1992) *Trends Biochem. Sci.* **17**, 502–506
300. Rhee, S. G., and Choi, K. D. (1992) *J. Biol. Chem.* **267**, 12393–12396
300a. Kim, M. J., Chang, J.-S., Park, S. K., Hwang, J.-I., Ryu, S. H., and Suh, P.-G. (2000) *Biochemistry* **39**, 8674–8682
301. Jafri, M. S., and Keizer, J. (1995) *Biophys. J.* **69**, 2139–2153
302. Combettes, L., Cheek, T. R., and Taylor, C. W. (1996) *EMBO J.* **15**, 2086–2093
303. van de Put, F. H. M. M., De Pont, J. J. H. H. M., and Willems, P. H. G. M. (1994) *J. Biol. Chem.* **269**, 12438–12443
304. Kukuljan, M., Rojas, E., Catt, K. J., and Stojilkovic, S. S. (1994) *J. Biol. Chem.* **269**, 4860–4865
305. Berridge, M. J., and Galione, A. (1988) *FASEB J.* **2**, 3074–3082
306. Meyer, T., Wensel, T., and Stryer, L. (1990) *Biochemistry* **29**, 32–37
307. Luttrell, B. M. (1993) *J. Biol. Chem.* **268**, 1521–1524
308. Wilson, M. P., and Majerus, P. W. (1996) *J. Biol. Chem.* **271**, 11904–11910
309. Shears, S. B. (1989) *J. Biol. Chem.* **264**, 19879–19886
309a. Zhu, D.-M., Tekle, E., Huang, C. Y., and Chock, P. B. (2000) *J. Biol. Chem.* **275**, 6063–6066
309b. Odom, A. R., Stahlberg, A., Wente, S. R., and York, J. D. (2000) *Science* **287**, 2026–2029
310. Ryu, S. H., Lee, S. Y., Lee, K.-Y., and Rhee, S. G. (1987) *FASEB J.* **1**, 388–393
311. Fukuda, M., and Mikoshiba, K. (1996) *J. Biol. Chem.* **271**, 18838–18842
312. Bird, G. St. J., and Putney, J. W., Jr. (1996) *J. Biol. Chem.* **271**, 6766–6770
313. Dixon, J. F., and Hokin, L. E. (1987) *J. Biol. Chem.* **262**, 13892–13895
314. Heinz, D. W., Ryan, M., Bullock, T. L., and Griffith, O. H. (1995) *EMBO J.* **14**, 3855–3863
315. Ali, N., Craxton, A., and Shears, S. B. (1993) *J. Biol. Chem.* **268**, 6161–6167
316. Xie, W., Kaetzel, M. A., Bruzik, K. S., Dedman, J. R., Shears, S. B., and Nelson, D. J. (1996) *J. Biol. Chem.* **271**, 14092–14097
317. Ismailov, I. I., Fuller, C. M., Berdiev, B. K., Shlyonsky, V. G., Benos, D. J., and Barrett, K. E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10505–10509
318. Dasgupta, S., Dasgupta, D., Sen, M., Biswas, S., and Biswas, B. B. (1996) *Biochemistry* **35**, 4994–5001
319. Voglmaier, S. M., Bembenek, M. E., Kaplin, A. I., Dormán, G., Olszewski, J. D., Prestwich, G. D., and Snyder, S. H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4305–4310
319a. Chi, T. H., and Crabtree, G. R. (2000) *Science* **287**, 1937–1939
320. Menniti, F. S., Oliver, K. G., Putney, J. W., Jr., and Shears, S. B. (1993) *Trends Biochem. Sci.* **18**, 53–56
320a. Yang, X., Safrany, S. T., and Shears, S. B. (1999) *J. Biol. Chem.* **274**, 35434–35440

321. Stephens, L., Radenberg, T., Thiel.U, Vogel, G., Khoo, K.-H., Dell, A., Jackson, T. R., Hawkins, P. T., and Mayr, G. W. (1993) *J. Biol. Chem.* **268**, 4009 – 4015

322. Van Dijken, P., de Haas, J.-R., Craxton, A., Erneux, C., Shears, S. B., and Van Haastert, P. J. M. (1995) *J. Biol. Chem.* **270**, 29724 – 29731

323. Bansal, V. S., Caldwell, K. K., and Majerus, P. W. (1990) *J. Biol. Chem.* **265**, 1806 – 1811

324. York, J. D., Ponder, J. W., Chen, Z.-w, Mathews, F. S., and Majerus, P. W. (1994) *Biochemistry* **33**, 13164 – 13171

325. Norris, F. A., Auethavekiat, V., and Majerus, P. W. (1995) *J. Biol. Chem.* **270**, 16128 – 16133

326. Jefferson, A. B., and Majerus, P. W. (1995) *J. Biol. Chem.* **270**, 9370 – 9377

326a. Norris, F. A., Wilson, M. P., Wallis, T. S., Galyov, E. E., and Majerus, P. W. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14057 – 14059

326b. Munday, A. D., Norris, F. A., Caldwell, K. K., Brown, S., Majerus, P. W., and Mitchell, C. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3640 – 3645

327. Bone, R., Springer, J. P., and Atack, J. R. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10031 – 10035

328. Ganzhorn, A. J., Lepage, P., Pelton, P. D., Strasser, F., Vincendon, P., and Rondeau, J.-M. (1996) *Biochemistry* **35**, 10957 – 10966

329. Burgoyne, R. D. (1989) *Trends Biochem. Sci.* **14**, 87 – 88

330. Huang, K.-P., Huang, F. L., Nakabayashi, H., and Yoshida, Y. (1988) *J. Biol. Chem.* **263**, 14839 – 14845

330a. Ron, D., and Kazanietz, M. G. (1999) *FASEB J.* **13**, 1658 – 1676

330b. Ward, N. E., Stewart, J. R., Ionnides, C. G., and O'Brian, C. A. (2000) *Biochemistry* **39**, 10319 – 10329

331. Pessin, M. S., and Raben, D. M. (1989) *J. Biol. Chem.* **264**, 8729 – 8738

332. Exton, J. H. (1990) *J. Biol. Chem.* **265**, 1 – 4

332a. Perisic, O., Paterson, H. F., Mosedale, G., Lara-González, S., and Williams, R. L. (1999) *J. Biol. Chem.* **274**, 14979 – 14987

333. Roberts, M. F. (1996) *FASEB J.* **10**, 1159 – 1172

334. Liscovitch, M., Chalifa, V., Pertile, P., Chen, C.-S., and Cantley, L. C. (1994) *J. Biol. Chem.* **269**, 21403 – 21406

335. Stein, R. L., Melandri, F., and Dick, L. (1935) *Biochemistry* **35**, 3899 – 3908

336. Ward, S. G., Mills, S. J., Liu, C., Westwick, J., and Potter, B. V. L. (1995) *J. Biol. Chem.* **270**, 12075 – 12084

337. Woscholski, R., Kodaki, T., Palmer, R. H., Waterfield, M. D., and Parker, P. J. (1995) *Biochemistry* **34**, 11489 – 11493

337a. Odorizzi, G., Babst, M., and Emr, S. D. (2000) *Trends Biochem. Sci.* **25**, 229 – 235

337b. Walker, E. H., Perisic, O., Ried, C., Stephens, L., and Williams, R. L. (1999) *Nature (London)* **402**, 313 – 320

337c. Ching, T.-T., Wang, D.-S., Hsu, A.-L., Lu, P.-J., and Chen, C.-S. (1999) *J. Biol. Chem.* **274**, 8611 – 8617

337d. Bertsch, U., Deschermeier, C., Fanick, W., Girkontaite, I., Hillemeier, K., Johnen, H., Weglöhner, W., Emmrich, F., and Mayr, G. W. (2000) *J. Biol. Chem.* **275**, 1557 – 1564

337e. Jackson, T. R., Kearns, B. G., and Theibert, A. B. (2000) *Trends Biochem. Sci.* **25**, 489 – 495

337f. Gillooly, D. J., Morrow, I. C., Lindsay, M., Gould, R., Bryant, N. J., Gaullier, J.-M., Parton, R. G., and Stenmark, H. (2000) *EMBO J.* **19**, 4577 – 4588

338. Pennisi, E. (1997) *Science* **275**, 1876 – 1878

338a. Stambolic, V., Suzuki, A., de la Pompa, J. L., Brothers, G. M., Mirtsos, C., Sasaki, T., Ruland, J., Penninger, J. M., Siderovski, D. P., and Mak, T. W. (1998) *Cell* **95**, 29 – 39

338b. Toker, A., and Cantley, L. C. (1997) *Nature (London)* **387**, 673 – 675

339. Ui, M., Okada, T., Hazeki, K., and Hazeki, O. (1995) *Trends Biochem. Sci.* **20**, 303 – 307

339a. Takasuga, S., Katada, T., Ui, M., and Hazeki, O. (1999) *J. Biol. Chem.* **274**, 19545 – 19550

340. Chock, P. B., Rhee, S. G., and Stadtman, E. R. (1980) *Ann. Rev. Biochem.* **49**, 813 – 843

341. Stadtman, E. R., and Chock, P. B. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2761 – 2765

342. Huang, C.-Y. F., and Ferrell, J. E., Jr. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10078 – 10083

343. Goldbeter, A., and Koshland, D. E., Jr. (1987) *J. Biol. Chem.* **262**, 4460 – 4471

344. Schacter-Noiman, E., Chock, P. B., and Stadtman, E. R. (1983) *Philos. Trans. R. Soc. London B* **302**, 157 – 166

345. Schacter, E., Chock, P. B., and Stadtman, E. R. (1984) *J. Biol. Chem.* **259**, 12260 – 12264

346. Newsholme, E. A., Challiss, R. A. J., and Crabtree, B. (1984) *Trends Biochem. Sci.* **9**, 277 – 280

347. Koshland, D. E., Jr. (1987) *Trends Biochem. Sci.* **12**, 225 – 228

348. Taylor, S. I. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 843 – 896, McGraw-Hill, New York

349. Müller, G., Rouveyre, N., Crecelius, A., and Bandlow, W. (1998) *Biochemistry* **37**, 8683 – 8695

350. Smit, A. B., Vreugdenhil, E., Ebberink, R. H. M., Geraerts, W. P. M., Klootwijk, J., and Joosse, J. (1988) *Nature (London)* **331**, 535 – 538

351. LeRoith, D., Shiloach, J., Roth, J., and Lesniak, M. A. (1983) *J. Biol. Chem.* **256**, 6533 – 6536

352. Rosen, O. M. (1987) *Science* **237**, 1452 – 1458

353. White, M. F., and Kahn, C. R. (1994) *J. Biol. Chem.* **269**, 1 – 4

354. Barnard, R. J., and Youngren, J. F. (1992) *FASEB J.* **6**, 3238 – 3244

355. Bell, G. I., Burant, C. F., Takeda, J., and Gould, G. W. (1993) *J. Biol. Chem.* **268**, 19161 – 19164

356. Mueckler, M. (1994) *Eur. J. Biochem.* **219**, 713 – 725

356a. Li, J., Houseknecht, K. L., Stenbit, A. E., Katz, E. B., and Charron, M. J. (2000) *FASEB J.* **14**, 1117 – 1125

357. Cuatrecasas, P. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1277 – 1281

358. Pashmforoush, M., Yoshimasa, Y., and Steiner, D. F. (1994) *J. Biol. Chem.* **269**, 32639 – 32648

359. Ullrich, A., Bell, J. R., Chen, E. Y., Herrera, R., Petruzzelli, L. M., Dull, T. J., Gray, A., Coussens, L., Liao, Y.-C., Tsubokawa, M., Mason, A., Seeburg, P. H., Grunfeld, C., Rosen, O. M., and Ramachandran, J. (1985) *Nature (London)* **313**, 756 – 761

360. Xu, Q.-Y., Paxton, R. J., and Fujita-Yamaguchi, Y. (1990) *J. Biol. Chem.* **265**, 18673 – 18681

361. Herrera, R., Petruzzelli, L., Thomas, N., Bramson, H. N., Kaiser, E. T., and Rosen, O. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7899 – 7903

362. McClain, D. A., Maegawa, H., Lee, J., Dull, T. J., Ullrich, A., and Olefsky, J. M. (1987) *J. Biol. Chem.* **262**, 14663 – 14671

363. Perlman, R., Bottaro, D. P., White, M. F., and Kahn, C. R. (1989) *J. Biol. Chem.* **264**, 8946 – 8950

363a. Rouard, M., Bass, J., Grigorescu, F., Garrett, T. P. J., Ward, C. W., Lipkind, G., Jaffiole, C., Steiner, D. F., and Bell, G. I. (1999) *J. Biol. Chem.* **274**, 18487 – 18491

363b. Luo, R. Z.-T., Beniac, D. R., Fernandes, A., Yip, C. C., and Ottensmeyer, F. P. (1999) *Science* **285**, 1077 – 1080

363c. Woldin, C. N., Hing, F. S., Lee, J., Pilch, P. F., and Shipley, G. G. (1999) *J. Biol. Chem.* **274**, 34981 – 34992

364. Lemmon, M. A., and Schlessinger, J. (1994) *Trends Biochem. Sci.* **19**, 459 – 563

365. Hubbard, S. R., Wei, L., Ellis, L., and Hendrickson, W. A. (1994) *Nature (London)* **372**, 746 – 754

365a. Hubbard, S. R. (1997) *EMBO J.* **16**, 5572 – 5581

366. White, M. F., Shoelson, S. E., Keutmann, H., and Kahn, C. R. (1988) *J. Biol. Chem.* **263**, 2969 – 2980

367. Fabry, M., Schaefer, E., Ellis, L., Kojro, E., Fahrenholz, F., and Brandenburg, D. (1992) *J. Biol. Chem.* **267**, 8950 – 8956

368. Williams, P. F., Mynarcik, D. C., Yu, G. Q., and Whittaker, J. (1995) *J. Biol. Chem.* **270**, 3012 – 3016

369. Kjeldsen, T., Wiberg, F. C., and Andersen, A. S. (1994) *J. Biol. Chem.* **269**, 32942 – 32946

369a. Sparrow, L. G., McKern, N. M., Gorman, J. J., Strike, P. M., Robinson, C. P., Bentley, J. D., and Ward, C. W. (1997) *J. Biol. Chem.* **272**, 29460 – 29467

369b. Garant, M. J., Kole, S., Maksimova, E. M., and Bernier, M. (1999) *Biochemistry* **38**, 5896 – 5904

370. Clark, S., and Harrison, L. C. (1983) *J. Biol. Chem.* **258**, 11434 – 11437

371. Saltiel, A. R. (1994) *FASEB J.* **8**, 1034 – 1040

372. Carter, W. G., Asamoah, K. A., and Sale, G. J. (1995) *Biochemistry* **34**, 9488 – 9499

373. Myers, M. G., Jr., Sun, X. J., and White, M. F. (1994) *Trends Biochem. Sci.* **19**, 289 – 293

374. Yenush, L., Makati, K. J., Smith-Hall, J., Ishibashi, O., Myers, M. G. J., and White, M. F. (1996) *J. Biol. Chem.* **271**, 24300 – 24306

375. Kowalski-Chauvel, A., Pradayrol, L., Vaysse, N., and Seva, C. (1996) *J. Biol. Chem.* **271**, 26356 – 26361

376. Paz, K., Voliovitch, H., Hadari, Y. R., Roberts, C. T., Jr., LeRoith, D., and Zick, Y. (1996) *J. Biol. Chem.* **271**, 6998 – 7003

376a. Rother, K. I., Imai, Y., Caruso, M., Beguinot, F., Formisano, P., and Accili, D. (1998) *J. Biol. Chem.* **273**, 17491 – 17497

376b. Clark, S. F., Molero, J.-C., and James, D. E. (2000) *J. Biol. Chem.* **275**, 3819 – 3826

376c. Lehr, S., Kotzka, J., Herkner, A., Sikmann, A., Meyer, H. E., Krone, W., and Müller-Wieland, D. (2000) *Biochemistry* **39**, 10898 – 10907

377. Ricketts, W. A., Rose, D. W., Shoelson, S., and Olefsky, J. M. (1996) *J. Biol. Chem.* **271**, 26165 – 26169

378. Okada, S., and Pessin, J. E. (1996) *J. Biol. Chem.* **271**, 25533 – 25538

379. Czech, M. P., Klarlund, J. K., Yagaloff, K. A., Bradford, A. P., and Lewis, R. E. (1988) *J. Biol. Chem.* **263**, 11017 – 11020

380. Seger, R., and Krebs, E. G. (1995) *FASEB J.* **9**, 726 – 735

381. Carter, W. G., Sullivan, A. C., Asamoah, K. A., and Sale, G. J. (1996) *Biochemistry* **35**, 14340 – 14351

382. Kasus-Jacobi, A., Perdereau, D., Auzan, C., Clauser, E., Van Obberghen, E., Mauvais-Jarvis, F., Girard, J., and Burnol, A.-F. (1998) *J. Biol. Chem.* **273**, 26026 – 26035

383. Baltensperger, K., Karoor, V., Paul, H., Ruoho, A., Czech, M. P., and Malbon, C. C. (1996) *J. Biol. Chem.* **271**, 1061 – 1064

384. Hara, K., Yonezawa, K., Sakaue, H., Ando, A., Kotani, K., Kitamura, T., Kitamura, Y., Ueda, H., Stephens, L., Jackson, T. R., Hawkins, P. T., Dhand, R., Clark, A. E., Holman, G. D., Waterfield, M. D., and Kasuga, M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7415 – 7419

385. Lam, K., Carpenter, C. L., Ruderman, N. B., Friel, J. C., and Kelly, K. L. (1994) *J. Biol. Chem.* **269**, 20648 – 20652

386. Sutherland, C., O'Brien, R. M., and Granner, D. K. (1995) *J. Biol. Chem.* **270**, 15501 – 15506

387. Domin, J., Dhand, R., and Waterfield, M. D. (1996) *J. Biol. Chem.* **271**, 21614 – 21621

387a. Kosaki, A., Yamada, K., Suga, J., Otaka, A., and Kuzuya, H. (1998) *J. Biol. Chem.* **273**, 940 – 944

References

388. Lefebvre, V., Méchin, M.-C., Louckx, M. P., Rider, M. H., and Hue, L. (1996) *J. Biol. Chem.* **271**, 22289 – 22292
389. Quon, M. J., Butte, A. J., Zarnowski, M. J., Sesti, G., Cushman, S. W., and Taylor, S. I. (1994) *J. Biol. Chem.* **269**, 27920 – 27924
390. Kandror, K. V., and Pilch, P. F. (1996) *J. Biol. Chem.* **271**, 21703 – 21708
390a. Marte, B. M., and Downward, J. (1997) *Trends Biochem. Sci.* **22**, 355 – 358
390b. Hemmings, B. A. (1997) *Science* **275**, 628 – 630
390c. Chan, T. O., Rittenhouse, S. E., and Tsichlis, P. N. (1999) *Ann. Rev. Biochem.* **68**, 965 – 1014
390d. Aguirre, V., Uchida, T., Yenush, L., Davis, R., and White, M. F. (2000) *J. Biol. Chem.* **275**, 9047 – 9054
390e. Nakae, J., Park, B.-C., and Accili, D. (1999) *J. Biol. Chem.* **274**, 15982 – 15985
391. Kishi, K., Hayashi, H., Wang, L., Kamohara, S., Tamaoka, K., Shimizu, T., Ushikubi, F., Narumiya, S., and Ebina, Y. (1996) *J. Biol. Chem.* **271**, 26561 – 26568
392. Yamauchi, K., Ribon, V., Saltiel, A. R., and Pessin, J. E. (1995) *J. Biol. Chem.* **270**, 17716 – 17722
393. Zhang, J., Hiken, J., Davis, A. E., and Lawrence, J. C., Jr. (1989) *J. Biol. Chem.* **264**, 17513 – 17523
394. Robinson, F. W., Smith, C. J., Flanagan, J. E., Shibata, H., and Kono, T. (1989) *J. Biol. Chem.* **264**, 16458 – 16464
395. Najjar, S. M., Choice, C. V., Soni, P., Whitman, C. M., and Poy, M. N. (1998) *J. Biol. Chem.* **273**, 12923 – 12928
396. Haft, C. R., Klausner, R. D., and Taylor, S. I. (1994) *J. Biol. Chem.* **269**, 26286 – 26294
397. Safavi, A., Miller, B. C., Cottam, L., and Hersh, L. B. (1996) *Biochemistry* **35**, 14318 – 14325
398. Kublaoui, B., Lee, J., and Pilch, P. F. (1995) *J. Biol. Chem.* **270**, 59 – 65
399. Bevan, A. P., Burgess, J. W., Drake, P. G., Shaver, A., Bergeron, J. J. M., and Posner, B. I. (1995) *J. Biol. Chem.* **270**, 10784 – 10791
400. Podlecki, D. A., Smith, R. M., Kao, M., Tsai, P., Huecksteadt, T., Brandenburg, D., Lasher, R. S., Jarett, L., and Olefsky, J. M. (1987) *J. Biol. Chem.* **262**, 3362 – 3368
401. Cherniack, A. D., Klarlund, J. K., Conway, B. R., and Czech, M. P. (1995) *J. Biol. Chem.* **270**, 1485 – 1488
402. Unoue, G., Cheatham, B., and Kahn, C. R. (1996) *J. Biol. Chem.* **271**, 28206 – 28211
403. Yarden, Y., and Ullrich, A. (1988) *Biochemistry* **27**, 3113 – 3119
404. James, R. (1984) *Ann. Rev. Biochem.* **53**, 259 – 292
405. Nilsen-Hamilton, M., ed. (1994) *Growth Factors and Signal Transduction in Development*, Wiley-Liss, New York
406. Claesson-Welsh, L. (1994) *J. Biol. Chem.* **269**, 32023 – 32026
407. Williams, L. T. (1989) *Science* **243**, 1564 – 1570
408. Hesketh, R. (1995) *The Oncogene Facts Book*, Academic Press, San Diego, California
409. Hunter, T. (1984) *Sci. Am.* **251**(Aug), 70 – 79
410. Marx, J. (1994) *Science* **266**, 1942 – 1944
411. Cooper, G. M., ed. (1990) *Oncogenes*, Jones & Bartlett, Boston, Massachusetts
412. Bishop, J. M. (1996) *FASEB J.* **10**, 362 – 364
413. Weinberg, R. A. (1983) *Sci. Am.* **249**(Nov), 126 – 142
414. Hickman, C. P. (1973) *Biology of the Invertebrates*, Mosby, St. Louis, Missouri
415. Yarden, Y., Escobedo, J. A., Kuang, W. J., Yang-Feng, T. L., Daniel, T. O., Tremble, P. M., Chen, E. Y., Ando, M. E., Harkins, R. N., Francke, V., Fried, V. A., Ullrich, A., and Williams, L. T. (1986) *Nature (London)* **323**, 226 – 232
416. Robbins, K. C., Antoniades, H. N., Devare, S. G., Hunkapillar, M. W., and Aaronson, S. A. (1983) *Nature (London)* **305**, 605 – 608
417. Yarden, Y., and Peles, E. (1991) *Biochemistry* **30**, 3543 – 3550
418. Peles, E., Lamprecht, R., Ben-Levy, R., Tzahar, E., and Yarden, Y. (1992) *J. Biol. Chem.* **267**, 12266 – 12274
419. Carpenter, G., and Cohen, S. (1990) *J. Biol. Chem.* **265**, 7709 – 7712
420. Suen, T.-C., and Goss, P. E. (2000) *J. Biol. Chem.* **275**, 6600 – 6607
421. Stover, D. R., Becker, M., Liebetanz, J., and Lydon, N. B. (1995) *J. Biol. Chem.* **270**, 15591 – 15597
422. Kauffmann-Zeh, A., Thomas, G. M. H., Ball, A., Prosser, S., Cunningham, E., Cockcroft, S., and Hsuan, J. J. (1995) *Science* **268**, 1188 – 1190
422a. Jones, J. T., Ballinger, M. D., Pisacane, P. I., Lofgren, J. A., Fitzpatrick, V. D., Fairbrother, W. J., Wells, J. A., and Sliwkowski, M. X. (1998) *J. Biol. Chem.* **273**, 11667 – 11674
423. Montelione, G. T., Wüthrich, K., and Scheraga, H. A. (1988) *Biochemistry* **27**, 2235 – 2243
424. Schlessinger, J. (1988) *Biochemistry* **27**, 3119 – 3123
425. Montelione, G. T., Wüthrich, K., and Scheraga, H. A. (1988) *Biochemistry* **27**, 2235 – 2243
426. Chantry, A. (1995) *J. Biol. Chem.* **270**, 3068 – 3073
427. Karasik, A., Pepinsky, R. B., Shoelson, S. E., and Kahn, C. R. (1988) *J. Biol. Chem.* **263**, 11862 – 11867
428. Kauffmann-Zeh, A., Klinger, R., Endemann, G., Waterfield, M. D., Wetzker, R., and Hsuan, J. J. (1994) *J. Biol. Chem.* **269**, 31243 – 31251
429. Kaplan, K. B., Bibbins, K. B., Swedlow, J. R., Arnaud, M., Morgan, D. O., and Varmus, H. E. (1994) *EMBO J.* **13**, 4745 – 4756
429a. Williams, J. C., Weijland, A., Gonfloni, S., Thompson, A., Courtneidge, S. A., Superti-Furga, G., and Wierenga, R. K. (1997) *J. Mol. Biol.* **274**, 757 – 775
430. Okada, M., and Nakagawa, H. (1989) *J. Biol. Chem.* **264**, 20886 – 20893
431. Buss, J. E., and Sefton, B. M. (1985) *Journal of Virology* **53**, 7 – 12
432. Soriano, P., Montgomery, C., Geske, R., and Bradley, A. (1991) *Cell* **64**, 693 – 702
433. Lowe, C., Yoneda, T., Boyce, B. F., Chen, H., Mundy, G. R., and Soriano, P. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4485 – 4489
434. Tezuka, K.-i, Denhardt, D. T., Rodan, G. A., and Harada, S.-i. (1996) *J. Biol. Chem.* **271**, 22713 – 22717
435. Bolen, J. B., and Veillette, A. (1989) *Trends Biochem. Sci.* **14**, 404 – 407
436. Mustelin, T., and Burn, P. (1993) *Trends Biochem. Sci.* **18**, 215 – 220
437. Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarström, L., Kinnon, C., Levinsky, R., Bobrow, M., Smith, C. I. E., and Bentley, D. R. (1993) *Nature (London)* **361**, 226 – 233
438. Hatakeyama, M., Kono, T., Kobayashi, N., Kawahara, A., Levin, S. D., Perlmutter, R. M., and Taniguchi, T. (1991) *Science* **252**, 1523 – 1528
438a. Zou, X., and Calame, K. (1999) *J. Biol. Chem.* **274**, 18141 – 18144
439. Daley, G. Q., Van Etten, R. A., and Baltimore, D. (1990) *Science* **247**, 824 – 830
440. Landis, C. A., Masters, S. B., Spada, A., Pace, A. M., Bourne, H. R., and Vallar, L. (1989) *Nature (London)* **340**, 692 – 696
441. Barbacid, M. (1987) *Ann. Rev. Biochem.* **56**, 779 – 827
442. Kato, G. J., and Dang, C. V. (1992) *FASEB J.* **6**, 3065 – 3072
443. Marcu, K. B., Bossone, S. A., and Patel, A. J. (1992) *Ann. Rev. Biochem.* **61**, 809 – 860
444. Greenhalgh, D. A., Welty, D. J., Player, A., and Yuspa, S. H. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 643 – 647
445. Rabbitts, T. H., Forster, A., Hamlyn, P., and Baer, R. (1984) *Nature (London)* **309**, 592 – 597
446. Shima, E. A., Le Beau, M. M., McKeithan, T. W., Minowada, J., Showe, L. C., Mak, T. W., Minden, M. D., Rowley, J. D., and Diaz, M. O. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3439 – 3443
447. Cramer, C. J., and Truhlar, D. G. (1993) *J. Am. Chem. Soc.* **115**, 5745 – 5753
448. Perrin, C. L., Armstrong, K. B., and Fabian, M. A. (1994) *J. Am. Chem. Soc.* **116**, 715 – 722
449. Gabb, H. A., and Harvey, S. C. (1993) *J. Am. Chem. Soc.* **115**, 4218 – 4227
450. Plavec, J., Tong, W., and Chattopadhyaya, J. (1993) *J. Am. Chem. Soc.* **115**, 9734 – 9746
451. Ellervik, U., and Magnusson, G. (1994) *J. Am. Chem. Soc.* **116**, 2340 – 2347
451a. Hermeking, H., Rago, C., Schuhmacher, M., Li, Q., Barrett, J. F., Obaya, A. J., O'Connell, B. C., Mateyak, M. K., Tam, W., Kohlhuber, F., Dang, C. V., Sedivy, J. M., Eick, D., Vogelstein, B., and Kinzler, K. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2229 – 2234
452. Eden, A., Simchen, G., and Benvenisty, N. (1996) *J. Biol. Chem.* **271**, 20242 – 20245
453. Postel, E. H., Berberich, S. J., Flint, S. J., and Ferrone, C. A. (1993) *Science* **261**, 478 – 480
454. Ying, G.-G., Proost, P., van Damme, J., Bruschi, M., Introna, M., and Golay, J. (2000) *J. Biol. Chem.* **275**, 4152 – 4158
455. Macleod, K., Leprince, D., and Stehelin, D. (1992) *Trends Biochem. Sci.* **17**, 251 – 256
456. Karin, M. (1995) *J. Biol. Chem.* **270**, 16483 – 16486
457. Turner, R., and Tjian, R. (1989) *Science* **243**, 1689 – 1694
458. Vogt, P. K., and Bos, T. J. (1989) *Trends Biochem. Sci.* **14**, 172 – 175
459. Neuberg, M., Schuermann, M., Hunter, J. B., and Müller, R. (1989) *Nature (London)* **338**, 589 – 590
460. Treisman, R. (1995) *EMBO J.* **14**, 4905 – 4913
461. Abel, T., and Maniatis, T. (1989) *Nature (London)* **341**, 24 – 25
462. Glover, J. N. M., and Harrison, S. C. (1995) *Nature (London)* **373**, 257 – 261
463. Quantin, B., and Breathnach, R. (1988) *Nature (London)* **334**, 538 – 539
464. Greenberg, M. E., and Ziff, E. B. (1984) *Nature (London)* **311**, 433 – 438
465. Müller, R., Bravo, R., Burckhardt, J., and Curran, T. (1984) *Nature (London)* **312**, 716 – 720
466. Fukumoto, Y., Kaibuchi, K., Oku, N., Hori, Y., and Takai, Y. (1990) *J. Biol. Chem.* **265**, 774 – 780
467. Wang, Z.-Q., Ovitt, C., Grigoriadis, A. E., Möhle-Steinlein, U., Rüther, U., and Wagner, E. F. (1992) *Nature (London)* **360**, 741 – 745
468. Candeliere, G. A., Glorieux, F. H., Prud'homme, J., and St.-Arnaud, R. (1995) *N. Engl. J. Med.* **332**, 1546 – 1551
469. Egan, S. E., and Weinberg, R. A. (1993) *Nature (London)* **365**, 781 – 783
470. Davis, R. J. (1994) *Trends Biochem. Sci.* **19**, 470 – 473
470a. Gutkind, J. S. (1998) *J. Biol. Chem.* **273**, 1839 – 1842
470b. Whitmarsh, A. J., and Davis, R. J. (1998) *Trends Biochem. Sci.* **23**, 481 – 485
471. Pawson, T. (1995) *Nature (London)* **373**, 573 – 580
472. Daly, R. J., Sanderson, G. M., Janes, P. W., and Sutherland, R. L. (1996) *J. Biol. Chem.* **271**, 12502 – 12510
473. Feller, S. M., Ren, R., Hanafusa, H., and Baltimore, D. (1994) *Trends Biochem. Sci.* **19**, 453 – 458
474. Pei, D., Lorenz, U., Klingmüller, U., Neel, B. G., and Walsh, C. T. (1994) *Biochemistry* **33**, 15483 – 15493

References

475. Ladbury, J. E., Hensmann, M., Panayotou, G., and Campbell, I. D. (1996) *Biochemistry* **35**, 11062–11069
476. Maignan, S., Guilloteau, J.-P., Fromage, N., Arnoux, B., Becquart, J., and Ducruix, A. (1995) *Science* **268**, 291–293
477. Mikol, V., Baumann, G., Zurini, M. G. M., and Hommel, U. (1995) *J. Mol. Biol.* **254**, 86–95
478. Zhou, M.-M., Ravichandran, K. S., Olejniczak, E. T., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J., and Fesik, S. W. (1995) *Nature (London)* **378**, 584–592
479. Bonfini, L., Migliaccio, E., Pelicci, G., Lanfrancone, L., and Pelicci, P. G. (1996) *Trends Biochem. Sci.* **21**, 257–261

479a. Nantel, A., Mohammad-Ali, K., Sherk, J., Posner, B. I., and Thomas, D. Y. (1998) *J. Biol. Chem.* **273**, 10475–10484

480. Burgering, B. M. T., and Bos, J. L. (1995) *Trends Biochem. Sci.* **20**, 18–22
481. Pumiglia, K. M., LeVine, H., Haske, T., Habib, T., Jove, R., and Decker, S. J. (1995) *J. Biol. Chem.* **270**, 14251–14254
482. Thompson, M. J., Roe, M. W., Malik, R. K., and Blackshear, P. J. (1994) *J. Biol. Chem.* **269**, 21127–21135
483. Jhun, B. H., Haruta, T., Meinkoth, J. L., Leitner, J. W., Draznin, B., Saltiel, A. R., Pang, L., Sasaoka, T., and Olefsky, J. M. (1995) *Biochemistry* **34**, 7996–8004
484. Jacob, K. K., Ouyang, L., and Stanley, F. M. (1995) *J. Biol. Chem.* **270**, 27773–27779
485. Pause, A., Belsham, G. J., Gingras, A.-C., Donzé, O., Lin, T.-A., Lawrence, J. C., Jr., and Sonenberg, N. (1994) *Nature (London)* **371**, 762–767
486. Proud, C. G. (1994) *Nature (London)* **371**, 747–748

486a. Wartmann, M., Hofer, P., Turowski, P., Saltiel, A. R., and Hynes, N. E. (1997) *J. Biol. Chem.* **272**, 3915–3923

486b. Weng, G., Bhalla, U. S., and Lyengar, R. (1999) *Science* **284**, 92–96

486c. Petosa, C., Masters, S. C., Bankston, L. A., Pohl, J., Wang, B., Fu, H., and Liddington, R. C. (1998) *J. Biol. Chem.* **273**, 16305–16310

487. Cobb, M. H., and Goldsmith, E. J. (1995) *J. Biol. Chem.* **270**, 14843–14846
488. Inglese, J., Koch, W. J., Touhara, K., and Lefkowitz, R. J. (1995) *Trends Biochem. Sci.* **20**, 151–156
489. van Biesen, T., Hawes, B. E., Raymond, J. R., Luttrell, L. M., Koch, W. J., and Lefkowitz, R. J. (1996) *J. Biol. Chem.* **271**, 1266–1269
490. Wu, J., Spiegel, S., and Sturgill, T. W. (1995) *J. Biol. Chem.* **270**, 11484–11488
491. Nurse, P. (2000) *Science* **289**, 1711–1716
492. Edgar, B. A., and Lehner, C. F. (1996) *Science* **274**, 1646–1652
493. Kirschner, M. (1992) *Trends Biochem. Sci.* **17**, 281–285
494. Horne, M. C., Goolsby, G. L., Donaldson, K. L., Tran, D., Neubauer, M., and Wahl, A. F. (1996) *J. Biol. Chem.* **271**, 6050–6061
495. Tyson, J. J., Novak, B., Odell, G. M., Chen, K., and Thron, C. D. (1996) *Trends Biochem. Sci.* **21**, 89–96
496. Sherr, C. J. (1996) *Science* **274**, 1672–1677
497. Hartwell, L. H., and Kastan, M. B. (1994) *Science* **266**, 1821–1828
498. Hennig, M., Jansonius, J. N., Terwisscha van Scheltinga, A. C., Dijkstra, B. W., and Schlesier, B. (1995) *J. Mol. Biol.* **254**, 237–246
499. Elledge, S. J. (1996) *Science* **274**, 1664–1672

499a. Chin, L., Pomerantz, J., and DePinho, R. A. (1998) *Trends Biochem. Sci.* **23**, 291–296

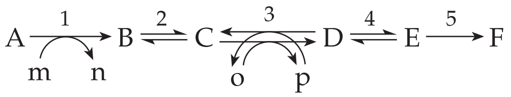
499b. Nead, M. A., Baglia, L. A., Antinore, M. J., Ludlow, J. W., and McCance, D. J. (1998) *EMBO J.* **17**, 2342–2352

499c. Bakiri, L., Lallemand, D., Bossy-Wetzel, E., and Yaniv, M. (2000) *EMBO J.* **19**, 2056–2068

500. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., and Pavletich, N. P. (1996) *Science* **274**, 948–953
501. Gorina, S., and Pavletich, N. P. (1996) *Science* **274**, 1001–1005
502. Galaktionov, K., Chen, X., and Beach, D. (1996) *Nature (London)* **382**, 511–517
503. Agarwal, M. L., Taylor, W. R., Chernov, M. V., Chernova, O. B., and Stark, G. R. (1998) *J. Biol. Chem.* **273**, 1–4
504. Carr, A. M. (2000) *Science* **287**, 1765–1766
505. Hirao, A., Kong, Y.-Y., Matsuoka, S., Wakeham, A., Ruland, J., Yoshida, H., Liu, D., Elledge, S. J., and Mak, T. W. (2000) *Science* **287**, 1824–1827
506. Gatti, A., Li, H.-H., Traugh, J. A., and Liu, X. (2000) *Biochemistry* **39**, 9837–9842

Study Questions

1. Illustrated is a generalized metabolic pathway in which capital letters indicate major metabolites in the pathway, small letters indicate cofactors and numbers indicate enzymes catalyzing the reactions.



List and describe four different ways in which the pathway might be regulated, referring to the specific enzymes, reactants, and cofactors indicated in the diagram. NOTE: Do not just refer to four different reactions as possible sites of regulation, but give four different general methods for regulation.

2. Describe the role of β-fructose-2,6-bisphosphate in the control of the further breakdown and resynthesis of glucose 1-phosphate.

3. It has been proposed that the substrate cycle involving phosphofructokinase and fructose bisphosphatase is used by bumblebees to warm their flight muscles to 30°C before flight begins. Clark *et al.* (1973) *Biochem. J.* **134**, 589–597 found maximal rates of catalytic activity for both enzymes to be about 44 μ mol / min / g of fresh tissue. In flying bees glycolysis occurred at a rate of about 20 μ mol / min / g of tissue with no substrate cycling. In resting bees at 27°C no cycling was detected, but at 5°C substrate cycling occurred at the rate of 10.4 μ mol / min / g while glycolysis had slowed to 5.8 μ mol / min / g. If the cycling provides heat to warm the insect, estimate how long it would take to reach 30°C if a cold (5°C) bee could carry out cycling at the maximum rate of 40 μ mol / min / g and if no heat were lost to the surroundings.

4. a) Compare the reaction cycle of a small GTPase (G protein) that is regulated by the actions of a GTPase-activating protein (GAP) and a guanine-nucleotide exchange factor (GEF) with that of a G protein linked to a receptor.
   b) Some G proteins have been described as "timed switches" and others as "triggered switches." In what ways might these two groups differ? See Kjeldgaard *et al.* (1996) *FASEB J.* **10**, 1347–1368.

View of a modified bovine fibrinogen molecule. The 45-nm-long disulfide-linked dimer is composed of three nonidentical polypeptide chains. The N termini of the six chains from the two halves come together in the center in a small globular "disulfide knot." The C termini form globular domains at the ends. The 340-kDa molecule has been treated with a lysine-specific protease which has removed portions of two chains to give the ~285-kDa molecule whose crystal structure is shown. Arrows point to attached oligosaccharides. From Brown *et al.* (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 85–90. Courtesy of Carolyn Cohen.

Contents

589 .... **A. Factors Affecting Rates of a Displacement Reaction**
590 .... **B. Nucleophilic Displacements on Singly Bonded Carbon Atoms**
590 .... 1. Inversion as a Criterion of Mechanism
591 .... 2. Transmethylation
592 .... 3. Kinetic Isotope Effects
593 .... 4. Glycosyltransferases
593 .... *Inversion or retention?*
595 .... *Double-displacement mechanisms*
598 .... *Carbocationic intermediates*
599 .... 5. Lysozymes and Chitinases
599 .... *Catalytic side chain groups*
600 .... *Kinetic isotope effect for lysozyme*
601 .... *Does lysozyme distort its substrate?*
601 .... *Help from a neighboring group*
602 .... 6. Cellulases and Other β-Glycosidases
604 .... 7. Glycogen Phosphorylase
605 .... 8. Starch-Hydrolyzing Enzymes
608 .... **C. Displacement Reactions on Carbonyl Groups**
609 .... 1. The Serine Proteases
609 .... *Serine as a nucleophile*
610 .... *Acyl-enzyme intermediates*
611 .... *Three-dimensional structures*
613 .... *The catalytic cycle*
614 .... *The catalytic triad*
614 .... *The "oxyanion hole"*
616 .... *Stereoelectronic considerations*
616 .... *pH dependence*
616 .... *Substrate specificity*
618 .... 2. The Cysteine Proteases (Thiol Proteases)
620 .... 3. N-Terminal Nucleophile Hydrolases and Related Enzymes
621 .... 4. The Aspartic Proteases
625 .... 5. Metalloproteases
627 .... 6. ATP-Dependent Proteases
628 .... 7. The Many Functions of Proteases
629 .... 8. Protease Inhibitors of Animals and Plants
631 .... 9. Coagulation of Blood
634 .... 10. Esterases and Lipases
637 .... 11. Other Acyltransferases
637 .... **D. Displacement on a Phosphorus Atom**
638 .... 1. Questions about Mechanisms
638 .... *Geometric complexities*
639 .... *Metaphosphate ions*
639 .... *Coping with negative charges*
639 .... 2. Magnetic Resonance Studies
642 .... 3. Stereochemistry
645 .... 4. Phosphatases
647 .... 5. Ribonucleases
649 .... 6. Ribonuclease P, Ribozymes, and Peptidyl Transferase
652 .... 7. Deoxyribonucleases
653 .... 8. Mutases
654 .... 9. Molecular Properties of Kinases
657 .... 10. Nucleotidyl Transferases
658 .... **E. The Adenylate Kinase Fold, the P Loop, and ATPases and GTPases**
659 .... **F. Displacements on Sulfur Atoms**
660 .... **G. Multiple Displacement Reactions and the Coupling of ATP Cleavage to Endergonic Processes**
660 .... 1. Transfer of Phospho, Pyrophospho, and Adenylyl Groups from ATP
660 .... 2. Acyl Phosphates
661 .... 3. General Mechanism of Formation of Thioesters, Esters, and Amides
662 .... 4. Coenzyme A Transferases

663 .... **References**
674 .... **Study Questions**

Boxes

594 .... Box 12-A Carboxymethylation of Proteins
596 .... Box 12-B Arsenic
622 .... Box 12-C Synthetic Protease Inhibitors
630 .... Box 12-D Molecular Mousetraps
636 .... Box 12-E Insecticides
658 .... Box 12-F The Toxicity of Aluminum

Tables

603 .... Table 12-1 Acidic and Basic Catalytic Groups in a Few Glycosyltransferases

Transferring Groups by Displacement Reactions

12

The majority of enzymes that are apt to be mentioned in any discussion of metabolism catalyze nucleophilic displacement reactions (Type 1, Table 10-1). These include most of the reactions by which the energy of ATP cleavage is harnessed and by which polymers are assembled from monomers. They include reactions by which pieces, large or small, are transferred onto or off of polymers as well as the reactions by which polymers are cleaved into pieces.[1–3d]

In these reactions a **nucleophilic group** (a base), designated B^- in Table 10-1, approaches an **electrophilic center,** often an electron-deficient carbon or phosphorus atom. It forms a bond with this atom, at the same time displacing some other atom, usually O, N, S, or C. The displaced atom leaves with its bonding electron pair and with whatever other chemical group is attached, the entire unit being called the **leaving group**. This is designated YH in Table 10-1. Simultaneous or subsequent donation of a proton from an acidic group of the enzyme, or from water, to the O, N, or S atom of the leaving group is usually required to complete the reaction. The entering base B, which may or may not carry a negative charge, must often be generated by enzyme-catalyzed removal of a proton from the conjugate acid BH. If BH is water, the enzyme is a **hydrolase**; otherwise, it is called a **transferase**.

For purposes of classifying the reactions of metabolism, in this book the nucleophilic displacements are grouped into four subtypes (Table 10-1). These are displacements on (A) a saturated carbon atom, often from a methyl group or a glycosyl group; (B) a carbonyl group of an ester, thioester, or amide; (C) a phospho group; or (D) a sulfur atom. In addition, many enzymes employ in sequence a displacement on a carbon atom followed by a second displacement on a phosphorus atom (or vice versa).

A. Factors Affecting Rates of a Displacement Reaction

In Chapter 9, the displacement of an iodide ion from methyl iodide by a hydroxide ion (Eq. 9-76) was considered. Can we similarly displace a methyl group from ethane, $CH_3{-}CH_3$, to break the C–C bond and form CH_3OH? The answer is no. Ethane is perfectly stable in sodium hydroxide and is not cleaved by a simple displacement process within our bodies. Likewise, long hydrocarbon chains such as those in the fatty acids cannot be broken by a corresponding process during metabolism of fatty acids. Not every structure will allow a nucleophilic displacement reaction to occur and not every anion or neutral base can act to displace another group.

At least four factors affect the likelihood of a displacement reaction:[1–3d]

(1) *The position of the equilibrium in the overall reaction*. An example is provided by the hydrolases that catalyze cleavage of amide, ester, and phosphodiester linkages using water as the entering nucleophile. Because enzymes usually act in an environment of high water content, the equilibrium almost always favors hydrolysis rather than the reverse reactions of synthesis. However, in a nonaqueous solvent the same enzyme will catalyze synthetic reactions.

(2) *The reactivity of the entering nucleophile*. Nucleophilic reactivity or **nucleophilicity** is partly determined by basicity. Compounds that are strong bases tend to react more rapidly in nonenzymatic displacement reactions than do weaker bases; the hydroxyl ion, HO^-, is a better nucleophile than $-COO^-$. However, enzymes usually act optimally near a neutral pH. Under these conditions the $-COO^-$ group may be

more reactive than strongly basic groups such as $-NH_2$ or HO^- because at pH 7 the latter groups will be almost completely protonated and the active nucleophile will be present in low concentrations.

A second factor affecting nucleophilic reactivity is **polarizability**, which is the ease with which the electronic distribution around an atom or within a chemical group can be distorted. A large atomic radius and the presence of double bonds in a group both tend to increase polarizability. In most cases, the more highly polarizable a group, the more rapidly it will react in a nucleophilic displacement, apparently because a polarizable group is able to form a partial bond at a greater distance than can a nonpolarizable group. Thus, I^- is more reactive than Br^-, which is more reactive than Cl^-. Polarizable bases such as imidazole are often much more reactive than nonpolarizable ones such as $-NH_2$. Sulfur compounds also tend to have a high nucleophilic reactivity. In displacement reactions on carbonyl groups (reaction type 1B), the less polarizable "hard" nucleophiles are more reactive than polarizable ones such as I^-. Attempts have been made to relate quantitatively nucleophilic reactivity to basicity plus polarizability.[1]

Certain chemical groups, e.g., those in which an atom with unpaired electrons is directly bonded to the nucleophilic center undergoing reaction, are more reactive than others of similar basicity. This **α effect** has been invoked to explain the high reactivity of the poisons hydroxylamine (NH_2OH) and cyanide ion[4] and other puzzling results.[1]

(3) *The chemical nature of the leaving group that is displaced*. The chemistry of the leaving group affects both the rate and the equilibrium position in nucleophilic displacements. The leaving group must accommodate a pair of electrons and often must bear a negative charge. A methyl group displaced from ethane or methane as CH_3^- would be an extremely poor leaving group; the pK_a of methane as an acid[5] has been estimated as 47. Iodide ion is a good leaving group, but F^- is over 10^4 times poorer.[5] In an aqueous medium, phosphate is a much better leaving group than OH^-, and pyrophosphate and tripolyphosphate are even better.

(4) *Special structural features present in the substrate*. Enzymes are usually constructed so that they recognize unique features in substrates. As a consequence, they have many ways of lowering the energy of the transition state and increasing the apparent nucleophilic reactivity.

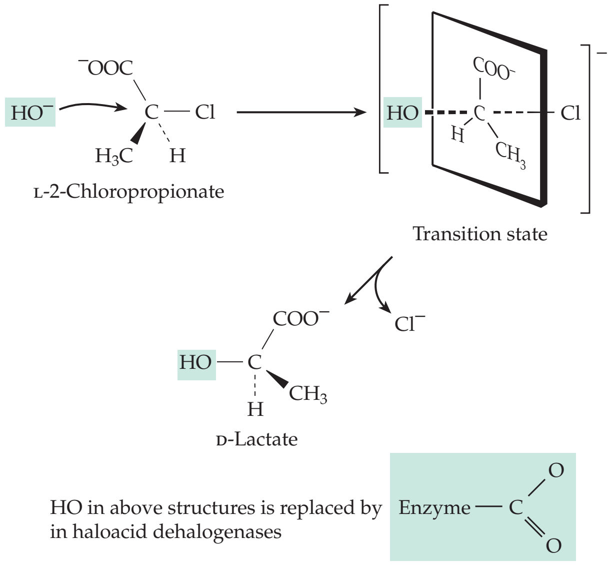
B. Nucleophilic Displacements on Singly Bonded Carbon Atoms

The enzymatic counterpart of Eq. 9-76, the displacement of I^- from a saturated carbon atom by HO^-, is catalyzed by the **haloacid dehalogenases** of soil pseudomonads (Eq. 12-1).[6–8] Even the poor leaving group F^- can be displaced by OH^- in the active site of these enzymes.

$$HO^- + F{-}CH_2COO^- \rightarrow HO{-}CH_2{-}COO^- + F^- \qquad (12\text{-}1)$$

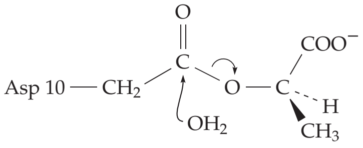
1. Inversion as a Criterion of Mechanism

An interesting result was obtained when a haloacid dehalogenase was tested with a substrate containing a chiral center.[6] Reaction of L-2-chloropropionate with hydroxyl ion gave only D-lactate, a compound with a chirality opposite to that of the reactant. A plausible explanation is that the hydroxyl ion attacks the central carbon from behind the chlorine atom (Eq. 12-2). The resulting five-bonded transition state (center) loses a chloride ion to form the product D-lactate in which inversion of configuration has occurred. *Inversion always accompanies single displacement reactions in which bond breaking and bond formation occur synchronously, as in Eq. 12-2*. This is true for both enzymatic and nonenzymatic reactions. However, the occurrence of inversion does not rule out more complex mechanisms. Indeed, in this case the displacing group is evidently not HO^- but a carboxylate group from the enzyme.



(12-2)

In one of the haloacid dehalogenases, a 232-residue protein for which the three-dimensional structure is known,[8,9] Asp 10 is in a position to carry out the initial attack which would give an enzyme-bound intermediate with an ester linkage:



Hydrolytic cleavage, as indicated in the diagram, yields the product D-lactate.[7] Thus, we have a direct displacement with inversion followed by an additional hydrolytic step. This is an example of **covalent catalysis**, the enzyme providing a well-oriented reactive group instead of generating a hydroxyl group from a bound water molecule.

Related **haloalkane dehalogenases** as well as **epoxide hydrolases** and a large superfamily of other enzymes utilize similar mechanisms.[10,11,11a] In the active site of a haloalkane dehalogenase from *Xanthobacter* (Fig. 12-1) the carboxylate of Asp 124 acts as the attacking nucleophile that displaces Cl^- from the substrate dichloroethane, which is held in a small cavity with the aid of two tryptophan indole rings.[12–14] However, the histidine–aspartate pair and the bound water molecule shown in Fig. 12-1 are essential for the subsequent hydrolysis of the intermediate ester. (see also Fig. 12-11).[15] The substrate shown in Fig. 12-1 is 1,2-dichloroethane, a widespread environmental pollutant that is not known to occur naturally. An interesting question considered by Pries *et al.*[16] is how this dehalogenase has evolved in the years since 1922 when industrial production of dichloroethane began.

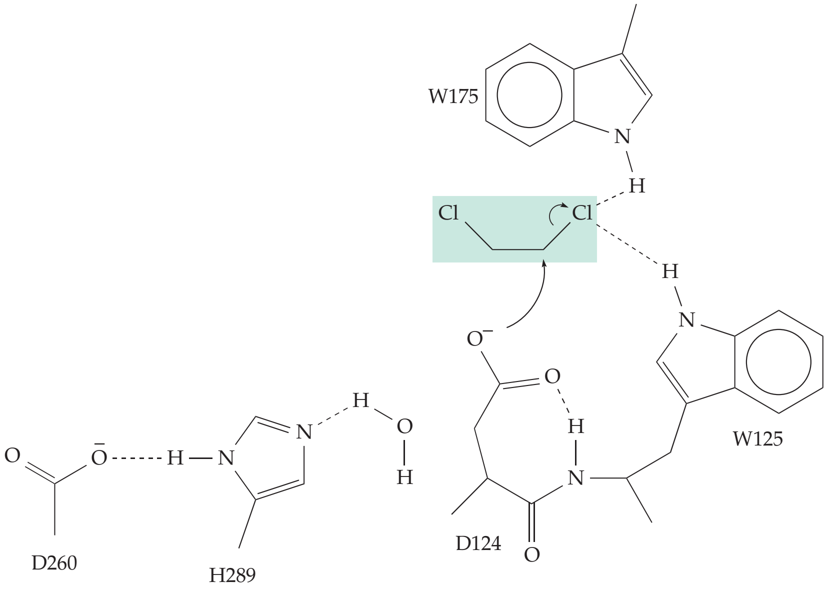


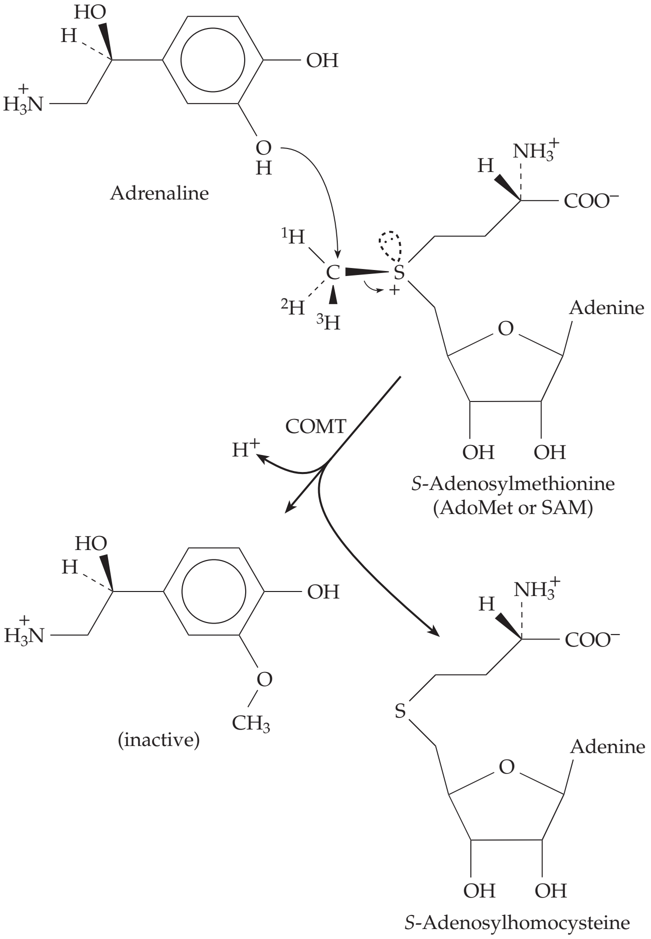
Figure 12-1 The active site structure of haloalkane dehalogenase from *Xanthobacter autotrophicus* with a molecule of bound dichloroethane. See Pries *et al.*[13] The arrows illustrate the initial nucleophilic displacement. The D260 – H289 pair is essential for the subsequent hydrolysis of the intermediate ester formed in the initial step.

2. Transmethylation

Nucleophilic attack on a methyl group or other alkyl group occurs most readily if the carbon is attached to an atom bearing a positive charge, for example, the sulfur atom of a sulfonium ion such as that present in ***S*-adenosylmethionine** (abbreviated **AdoMet** or SAM). The ensuing **transmethylation** reaction results in the transfer of the methyl group from the sulfur to the attacking nucleophile (Eq. 12-3). Transmethylation is an important metabolic process by which various oxygen, nitrogen, and other atoms at precise positions in proteins, nucleic acids, phospholipids, and other small molecules are methylated.[17] The methyl group donor is usually *S*-adenosylmethionine. This compound has two chiral centers, one at the α-carbon of the amino acid and one at the sulfur atom, with an unshared electron pair serving as the fourth group around the S atom. In the naturally occurring AdoMet both centers have the *S* configuration.[18] The reaction of Eq. 12-3, which is catalyzed by **catechol *O*-methyltransferase** (COMT), inactivates the neurotransmitters adrenaline, dopamine, and related compounds by methylation. When the substrate contains a chiral methyl group ($-C^1H^2H^3H$; see also Chapter 13), the inversion of the methyl group expected for a simple S_N2-like reaction is observed.[19]

Structural features similar to those of COMT[20] are found in glycine *N*-methyltransferase[21] and guanidinoacetate methyltransferase[22] from liver, and transferases that place methyl groups on N-6 of adenine[23–25a] and N-4 of cytosine or on C-5 of cytosine[26] in nucleic acids.[27] A stereoscopic view of the active site of glycine *N*-methyltransferase is shown in Fig. 12-2. An acetate ion present in the active site in the crystal has been converted into glycine by computer-assisted modeling. The amino group of the bound glycine is adjacent to the methyl group of AdoMet. A nearby glutamic acid side chain (E15) may have removed a proton from the glycine dipolar ion to create the free $-NH_2$ group required for the reaction. In COMT a magnesium Mg^{2+} ion binds to the two aromatic hydroxyl groups of the substrate and helps to hold it correctly in the active site.[28–29a]

In the case of the more difficult C-methylation of uracil 54 in transfer RNA, an –SH group from Cys 324 of the methyltransferase adds to the C-6 position of U54 of the substrate, which may be any of the *E. coli* tRNAs (Eq. 12-4, step *a*) .[30] In the



(12-3)

adduct carbon atom C-5 acquires substantial nucleophilic character which permits the transfer of the methyl group in step *b*. The adduct breaks up (step *c*) and the product is released.

Methylation of nucleic acids is considered further in Chapers 27 and 28. Methylation of carboxyl groups of certain proteins regulates motion of bacterial flagella (Chapter 19) and other aspects of metabolism[31] while methylation of isomerized aspartyl residues is part of a protein repair process (Box 12-A). Methyl groups are usually transferred from *S*-adenosylmethionine, but sometimes from a folic acid derivative (Chapter 15) or from a cobalt atom of a corrin ring (Chapter 16).

3. Kinetic Isotope Effects

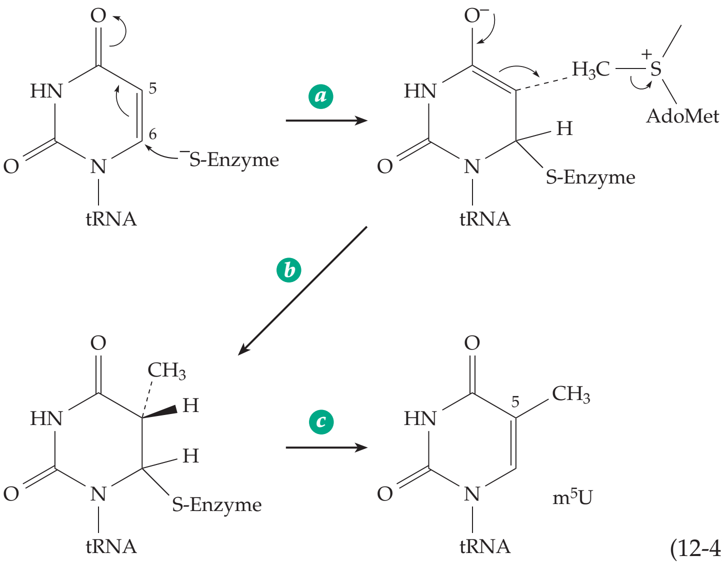
An S–^{12}C bond breaks a little faster than an S–^{13}C bond in a nucleophilic displacement reaction. This **primary kinetic isotope effect** (KIE)[3a,3b,31a–d] is usually discussed first for breakage of a C–H bond. In a linear transition state, in which the C–H bond is being stretched, then cleaved, the difference between the transition state barrier for a C–^{1}H bond and a C–^{2}H bond is thought to arise principally from a difference in the energies of the C–H stretching vibration.[3a] This vibrational energy, in the ground state of a molecule (the zero-point energy) is equal to $1/2\ h\nu_0$ where ν_0 may be observed in the infrared absorption spectrum (see Fig. 23-2). For a C–^{2}H bond, with a stretching vibration at a wave number of ~2900 cm^{-1}, the zero point energy is about + 17.4 kJ mol^{-1}. For a C–^{2}H bond, with a stretching wave number of ~2200 cm^{-1}, it is about +17.4 kJ mol^{-1}. This difference is a result of the difference in mass of ^{1}H (1.67×10^{-24} g) and ^{2}H (2.34×10^{-24} g). The isotope effect arises because the vibration occurs along the axis of the bond being broken, and the vibrational energy is converted into translational motion along the reaction coordinate, in effect lowering the transition state barrier by $\Delta G^{0\ddagger}$. The difference in this effect between ^{1}H and ^{2}H ($\Delta\Delta G^{0\ddagger}$) gives the ratio of expected rate constants as follows:

$$\frac{k^{C-H}}{k^{C-D}} = e^{-\Delta\Delta G^{0\pm}/RT} = 7$$

For heavier nuclei, such as C, N, O the kinetic isotope effect is much smaller.

When ^{13}C was introduced into the methyl group of *S*-adenosylmethionine and its rate of reaction with catechol *O*-methyltransferase was compared with that of the normal ^{12}C-containing substrate, the expected effect on V_{max}, expressed as a first-order rate constant, was seen: $k_{12}/k_{13} = 1.09 \pm 0.05$. This effect is small but it can be measured reliably and establishes that the methyl transfer step rather than substrate binding, product release, or a conformational change in the protein is rate limiting.[32]

Substitution of ^{1}H by ^{2}H in the CH_3 group has a larger effect. This **secondary kinetic isotope effect** (or α-deuterium



(12-4)

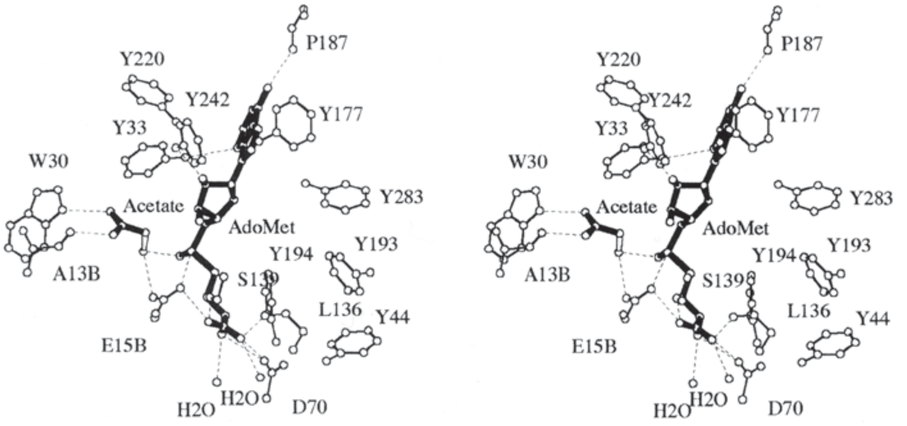


Fig. 12-2 *S*-Adenosylmethionine (AdoMet, solid bonds) bound in the active site of glycine *N*-methyltransferase together with an acetate ion bound in the glycine site. Glycine was built by attaching an amino group (open bond) to the acetate (solid bonds). Possible polar interactions (O–O and O–N < 0.31 nm and O–S < 0.4 nm) are indicated by dotted lines. Tyrosine residues located at the inner surface of the active site are also shown. From Fu *et al.*[21]

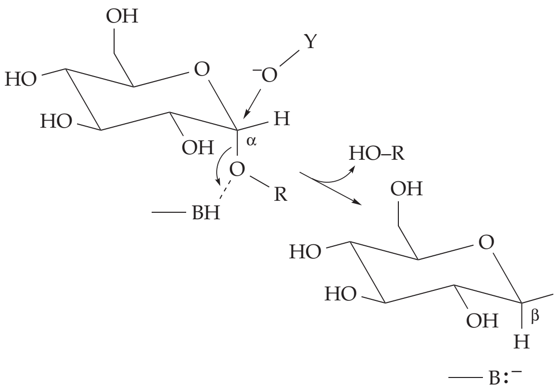
effect) arises because of small differences in the vibrational energies of the methyl group resulting from the differences in mass of ^{1}H and ^{2}H. It often leads to a more rapid reaction for the ^{1}H-containing substrates. In a model nonenzymatic displacement reaction of a similar type[32] the ratio of rate constants was $k_{1_H} / k_{2_H} = 1.17 \pm 0.02$. However, for COMT an inverse α-deuterium effect was seen: $k_{(\mathrm{C}^1\mathrm{H}_3)} / k_{(\mathrm{C}^2\mathrm{H}_3)} = 0.83 \pm 0.05$. Theoretical calculations suggested that such an effect might be observed if the enzyme compresses an S_N2-like transition state, shortening the bonds from the central carbon atom to both the oxygen atom of the entering nucleophile and the sulfur atom of the leaving group.[33,34] However, more recent calculations suggest that it is difficult to draw such conclusions from secondary isotope effects.[35,36] Calculations by Zheng, Kahn, and Bruice predict that in the gas phase the two substrates, upon collision, will react with a very low energy barrier.[29,29a] For this enzyme, as for the haloalkane dehalogenase (Fig. 12-1), it has been concluded that the enzyme excludes water from the active site and binds the two substrates very close together and in a correct orientation for reaction. The computations predicted secondary isotope effects for the enzymatic reaction similar to those measured by Hegazi *et al.*[32] suggesting that the transition states for the enzymatic reactions closely resemble those for the gas phase.

4. Glycosyltransferases

The polarization of a single C–O bond in an ether is quantitatively much less than that of the C–S$^+$ bond of *S*-adenosylmethionine, and simple ethers are not readily cleaved by displacement reactions. However, glycosides, which contain a carbon atom attached to *two* oxygen atoms, do undergo displacement reactions which lead to **hydrolysis**, **phosphorolysis**, or **glycosyl exchange**. The corresponding enzymes are glycosylhydrolases, phosphorylases, and glycosyltransferases. Two characteristics are commonly used to classify glycosylhydrolases that act on polysaccharides: **Endoglycanases** cut at random locations within the chains of sugar units, while **exoglycanases** cut only at one end or another, usually at the nonreducing end. **Inverting enzymes** invert the configuration at the anomeric carbon atom which they attack, while **retaining enzymes** do not. Although simple glycosides undergo acid-catalyzed hydrolysis readily, uncatalyzed hydrolysis is extremely slow, the estimated first order rate constant[37] at 25°C being about 2×10^{-15} s^{-1}. Polysaccharides may be the most stable of all bioploymers.

Glycosyltransferases are numerous. For example, the amino acid sequences for about 500 glucosidases, which hydrolyze linkages between glucose residues, have been determined and this one group of enzymes has been classified into over 60 families[38,39] and eight larger groups.[40] Many three-dimensional structures have been reported and numerous studies of the reaction mechanism for both enzymatic and nonenzymatic hydrolysis of glycosides have been conducted. Most of the experimental results have been carefully verified. Nevertheless, they serve to illustrate how difficult it is to understand how enzymes catalyze this simple type of reaction.[41]

Inversion or retention? Equation 12-5 pictures the reaction of a glycoside (such as a glucose unit at

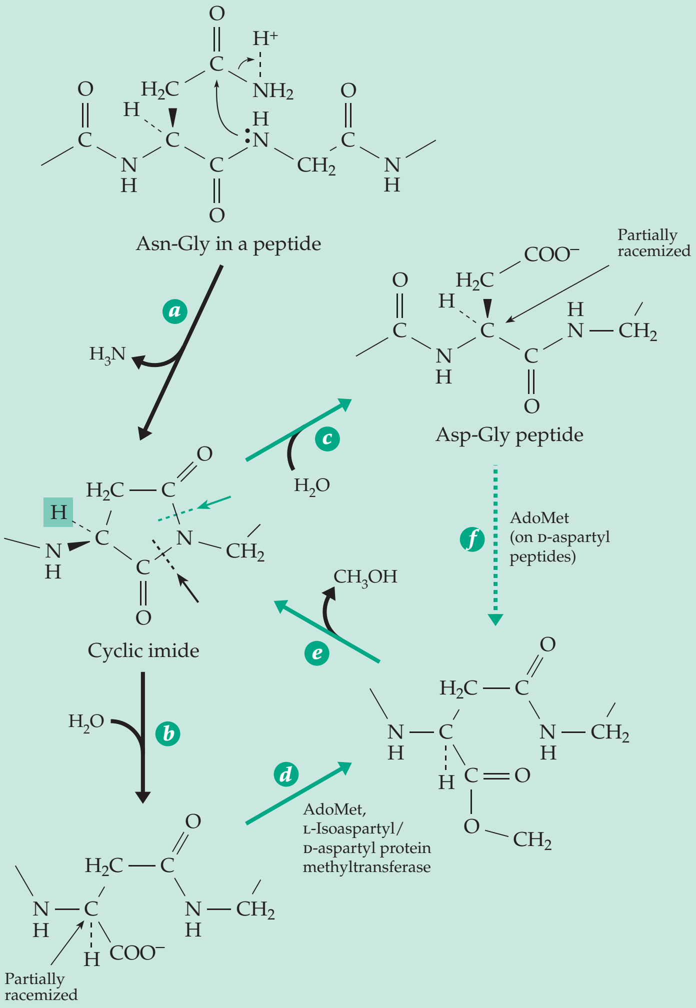


(12-5)

BOX 12-A CARBOXYMETHYLATION OF PROTEINS

Methylation and demethylation of the carboxyl groups of the side chains of specific glutamyl residues play a role in bacterial chemotaxis (Chapter 19, Section A). However, similar functions in eukaryotes have not been found. There is a specific methyltransferase that methylates C-terminal carboxyl groups on prenylated and sometimes palmitoylated peptides.[a] Another type of carboxylmethyltransferase acts on only a small percentage of many proteins and forms labile methyl esters. These methyltransferases have a dual specificity, acting on **L-isospartyl** residues and usually also on **D-aspartyl** residues. Both of these amino acids can arise from deamidation of asparagine, especially in Asn-Gly sequences (Eq. 2-24 and steps *a* and *b* of following scheme).[b–g] A similar isomerization of aspartyl residues occurs more slowly. Asn-Gly sequences are present in many proteins and provide weak linkages whose isomerization is an inevitable aspect of aging.[e,h] Furthermore, the α proton of the cyclic imide (green box) is more readily dissociated than in a standard peptide, leading to racemization (step *c*). The isoaspartyl and D-aspartyl residues provide kinks in the peptide chain which may interfere with normal function and turnover.

The L-isoaspartyl / D-aspartyl methyltransferase (step *d*), which is especially abundant in the brain,[i,j] provides for partial repair of these defects.[k–o] The methyl esters of isoaspartyl residues readily undergo demethylation with a return to a cyclic imide (step *e*). The cyclic imide is opened hydrolytically (step *c*) in part to an isoaspartyl residue, but in part to a normal aspartyl form. The combined action of the carboxylmethyltransferase and the demethylation reaction tends to repair the isomerized linkages. Methylation of D- aspartyl residues returns them to the cyclic imide (steps *f*, *e*), allowing them to also return to the normal L configuration. Nevertheless, the protein will have a different net charge than it did originally and must be considered a new modified form.



---

[a] Dai, Q., Choy, E., Chiu, V., Romano, J., Slivka, S. R., Steitz, S. A., Michaelis, S., and Philips, M. R. (1998) *J. Biol. Chem.* **273**, 15030–15034

[b] Geiger, T., and Clarke, S. (1987) *J. Biol. Chem.* **262**, 785–794

[c] Tyler-Cross, R., and Schirch, V. (1991) *J. Biol. Chem.* **266**, 22549–22556

[d] Brennan, T. V., and Clarke, S. (1993) *Protein Sci.* **2**, 331–338

[e] Paranandi, M. V., Guzzetta, A. W., Hancock, W. S., and Aswad, D. W. (1994) *J. Biol. Chem.* **269**, 243–253

[f] Donato, A. D., Ciardiello, M. A., de Nigris, M., Piccoli, R., Mazzarella, L., and D'Alessio, G. (1993) *J. Biol. Chem.* **268**, 4745–4751

[g] Tomizawa, H., Yamada, H., Ueda, T., and Imoto, T. (1994) *Biochemistry* **33**, 8770–8774

[h] Man, E. H., Sandhouse, M. E., Burg, J., and Fisher, G. H. (1983) *Science* **220**, 1407–1408

[i] Orpiszewski, J., and Aswad, D. W. (1996) *J. Biol. Chem.* **271**, 22965–22968

[j] Najbauer, J., Orpiszewski, J., and Aswad, D. W. (1996) *Biochemistry* **35**, 5183–5190

[k] Johnson, B. A., Murray, E. D. J., Clarke, S., Glass, D. B., and Aswad, D. W. (1987) *J. Biol. Chem.* **262**, 5622–5629

[l] Brennan, T. V., Anderson, J. W., Jia, Z., Waygood, E. B., and Clarke, S. (1994) *J. Biol. Chem.* **269**, 24586–24595

[m] Mudgett, M. B., and Clarke, S. (1993) *Biochemistry* **32**, 11100–11111

[n] Mudgett, M. B., and Clarke, S. (1994) *J. Biol. Chem.* **269**, 25605–25612

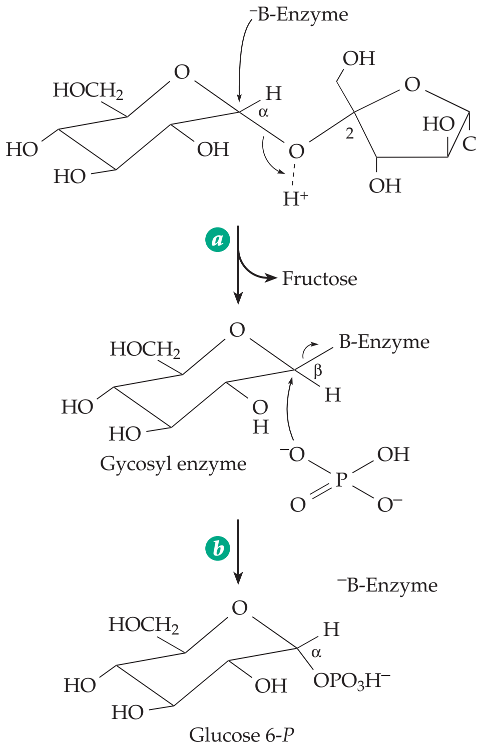
[o] Aswad, D. W., ed. (1995) *Deamidation and Isoaspartate Formation in Peptides and Proteins*, CRC Press, Boca Raton, Florida

the end of a starch chain) with a nucleophile Y–O⁻ as the displacing group. An enzyme-bound acidic group –BH is shown assisting the process. Inversion of configuration with formation of a product of the β configuration at the anomeric carbon atom would be predicted and is observed for many of these enzymes. However, many others do *not* cause inversion.[41,42] An example is the reaction catalyzed by **sucrose phosphorylase** from *Pseudomonas saccharophila*:

$$\text{Sucrose (an } \alpha\text{-glucoside of fructose) + inorganic phosphate } (P_i) \rightarrow \alpha\text{-D-glucose 1-phosphate + fructose} \qquad (12\text{-}6)$$

Two possible explanations for the lack of inversion during this reaction are that the enzyme acts by a **double-displacement** reaction or through a stabilized **carbocationic intermediate**. Let us consider these possibilities in turn.

Double-displacement mechanisms. In a double-displacement mechanism sucrose phosphorylase would catalyze two consecutive single displacements, each with inversion. A nucleophilic group of the enzyme would react in Eq. 12-7, step *a*. In step *b*, a phosphate would react to regenerate the enzyme with its free nucleophilic group –B⁻.



(12-7)

Four kinds of experiments were used to identify this double-displacement mechanism.

(1) *Kinetics***.** In a double-displacement mechanism the enzyme shuttles between free enzyme and the intermediate carrying the substrate fragment (here, the glycosyl enzyme). With sucrose phosphorylase the maximum velocity varies with the concentrations of sucrose and HPO_4^{2-} in the characteristic fashion expected for this "ping-pong" mechanism (Eq. 9-47).[43]

(2) *Exchange reactions***.** In a double-displacement mechanism sucrose containing ^{14}C in the fructose portion of the molecule should react with free enzyme E to form glycosyl enzyme and free radioactive fructose (Eq. 12-8). The ^{14}C-containing groups are designated here by the asterisks.

$$\underset{\text{(Sucrose)}}{\text{D-Glucosyl-fructose*}} + \text{E} \rightleftarrows \text{glucosyl-E + fructose*} \qquad (12\text{-}8)$$

If a very low molar concentration of enzyme is present, and a large excess of nonradioactive fructose is added, the enzyme will catalyze no net reaction but will change back and forth repeatedly between the free enzyme and glucosyl enzyme. Each time, in the reverse reaction, it will make use primarily of unlabeled fructose. The net effect will be catalysis of a sucrose–fructose exchange:

$$\text{Sucrose* + fructose} \rightleftarrows \text{sucrose + fructose*} \qquad (12\text{-}9)$$

This exchange reaction, as well as other predicted exchanges, has been observed.[44] Although the exchange criterion of the mechanism is often applied to enzymatic processes, the observation of exchange reactions does not *prove* the existence of a covalently bound intermediate. Furthermore, enzymes using double-displacement mechanisms may not always catalyze the expected exchanges.

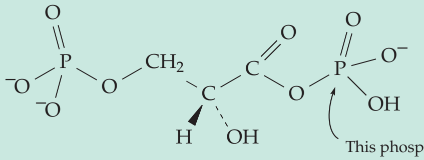
(3) *Arsenolysis***.** Sucrose phosphorylase also catalyzes the cleavage of sucrose by arsenate and promotes a rapid cleavage (**arsenolysis**) of glucose 1-phosphate to free glucose. This reaction is evidently a result of a displacement by arsenate on a glycosyl enzyme intermediate. The resulting unstable glucose 1-arsenate (see Box 12-B) is hydrolyzed rapidly. Arsenolysis is a general way of trapping reactive enzyme-bound intermediates that normally react with phosphate groups. Arsenate is one of many substrate analogs that can be used to siphon off reactive enzyme-bound intermediates into nonproductive side paths.

(4) *Identification of glycosyl-enzyme intermediates***.** Studies with pure enzymes often make it possible to confirm directly the existence of enzyme-bound intermediates. The intermediates detected are frequently **glycosyl esters** of glutamate or aspartate side chain

BOX 12-B ARSENIC

Arsenate, AsO_4^{3-}, is chemically similar to phosphate in size and geometry and in its ability to enter into biochemical reactions. However, arsenate esters are far less stable than phosphate esters. If formed in the active site of an enzyme, they are quickly hydrolyzed upon dissociation from the enzyme. This fact accounts for some of the toxicity of arsenic compounds.[a]

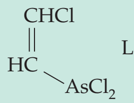
Arsenate will replace phosphate in all phosphorolytic reactions, e.g., in the cleavage of glycogen by glycogen phosphorylase, of sucrose by sucrose phosphorylase, and in the action of purine nucleoside phosphorylase.[b] Glucose 1-arsenate or ribose-1-arsenate is presumably a transient intermediate which is hydrolyzed to glucose. The overall process is called **arsenolysis**. Another reaction in which arsenate can replace phosphate is the oxidation of glyceraldehyde 3-phosphate in the presence of P_i to form 1,3-bisphosphoglycerate:



1,3-Biphosphoglycerate

This phospho group is normally transferred to ADP to form ATP

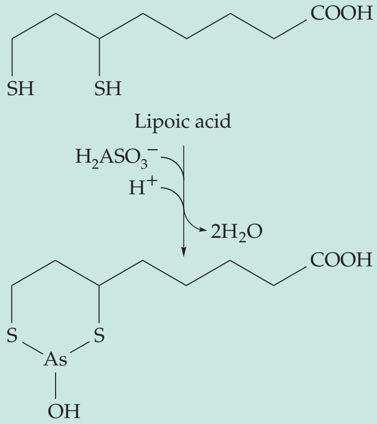
The subsequent transfer of the 1-phospho group to ADP is an important energy-yielding step in metabolism (Chapter 12). When arsenate substitutes for phosphate the acyl arsenate (1-arseno-3-phosphoglycerate) is hydrolyzed to 3-phosphoglycerate. As a consequence, in the presence of arsenate oxidation of 3-phosphoglyceraldehyde continues but ATP synthesis ceases. Arsenate is said to **uncouple** phosphorylation from oxidation. Arsenate can also partially replace phosphate in stimulating the respiration of mitochondria and is an uncoupler of oxidative phosphorylation (Chapter 18). Enzymes that normally act on a phosphorylated substrate will usually catalyze a slow reaction of the corresponding unphosphorylated substrate in the presence of arsenate. Apparently, the arsenate ester of the substrate forms transiently on the enzyme surface, permitting the reaction to occur.



Lewisite

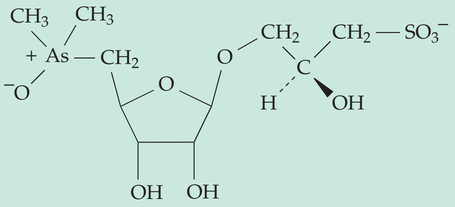
Arsenite is noted for its tendency to react rapidly with thiol groups,[c] especially with pairs of adjacent (vicinal) or closely spaced –SH groups[d] as in lipoic acid. By blocking oxidative enzymes requiring lipoic acid (Chapter 15), arsenite causes the accumulation of pyruvate and of other 2-oxo acids. Similar chemistry underlies the action of the mustard gas Lewisite, which also attacks lipoic acid.[de]

Many, perhaps all, organisms have enzyme systems for protection against arsenic compounds.



In *E. coli* arsenate is reduced to arsenite by a glutaredoxin- and NADH-dependent system.[e–g] The arsenite as well as antimonite and tellurite are pumped out by an ATP-dependent transporter. The genes for reductase, periplasmic-binding protein, and transporter components are encoded in a conjugative plasmid.[h,i] A quite similar system functions in yeast.[j]

Marine algae as well as some higher aquatic plants detoxify and excrete arsenate by conversion to various water-soluble organic forms such as trimethyarsonium lactic acid (Chapter 8) and the following ribofuranoside.[a]



2-Hydroxy-3-sulfopropyl-5-deoxy-5-(dimethylarseno)-ribofuranoside

Arsenic-containing phospholipids are also formed (Chapter 8).[k]

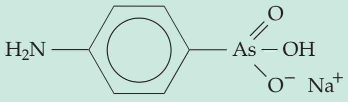
Arsenic is present at high levels in some soils and contamination of drinking water with arsenic is a major problem is some areas of the world. In West Bengal, India, millions of people drink contaminated well water. As a result hundreds of thousands have developed debilitating nodular keratoses on their feet.[l,m] The problem is made worse by the increasing

BOX 12-B (continued)

use of fresh water for irrigation and the difficulty of removing arsenic contamination.

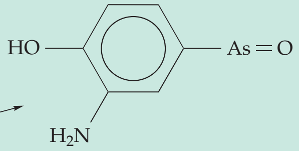
Although it is most known for its toxicity, arsenic may be an essential nutrient. Data from feeding of chicks, goats, rats, and miniature pigs suggest a probable human need for arsenic of ~12 μg / day.[n]

Compounds of arsenic have been used in medicine for over 2000 years, but only in the past century have specific arsenicals been created as drugs. In 1905, it was discovered that sodium arsanilate is toxic to trypanosomes. The development by P. Ehrlich of **arsenicals** for the treatment of syphilis (in 1909) first focused attention on the possibility of effective chemotherapy against bacterial infections.



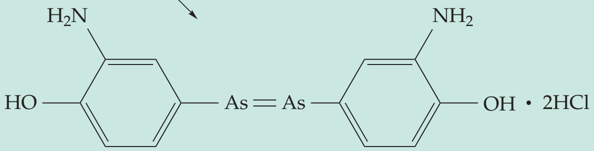
Substitution by —CH$_2$CONH$_2$ yields "tryparsamide," a more effective anti-trypanosome drug

Sodium arsanilate



Oxophenarsine

Two of Ehrlich's anti-syphilis drugs



Salvarsan

---

[a] Knowles, F. C., and Benson, A. A. (1983) *Trends Biochem. Sci.* **8**, 178–180

[b] Kline, P. C., and Schramm, V. L. (1995) *Biochemistry* **34**, 1153–1162

[c] Lam, W.-C., Tsao, D. H. H., Maki, A. H., Maegley, K. A., and Reich, N. O. (1992) *Biochemistry* **31**, 10438–10442

[d] Li, J., and Pickart, C. M. (1995) *Biochemistry* **34**, 15829–15837

[de] Ord, M. G., and Stocken, L. A. (2000) *Trends Biochem. Sci.* **25**, 253–256

[e] Silver, S., Nucifora, G., Chu, L., and Misra, T. K. (1989) *Trends Biochem. Sci.* **14**, 76–80

[f] Ji, G., Garber, E. A. E., Armes, L. G., Chen, C.-M., Fuchs, J. A., and Silver, S. (1994) *Biochemistry* **33**, 7294–7299

[g] Rosen, B. P., Weigel, U., Karkaria, C., and Gangola, P. (1988) *J. Biol. Chem.* **263**, 3067–3070

[h] Gladysheva, T. B., Oden, K. L., and Rosen, B. P. (1994) *Biochemistry* **33**, 7288–7293

[i] Chen, Y., and Rosen, B. P. (1997) *J. Biol. Chem.* **272**, 14257–14262

[j] Wysocki, R., Bobrowicz, P., and Ulaszewski, S. (1997) *J. Biol. Chem.* **272**, 30061–30066

[k] Cooney, R. V., Mumma, R. O., and Benson, A. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4262–4264

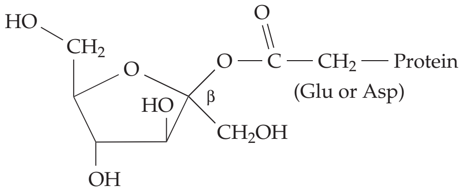
[l] Bagla, P., and Kaiser, J. (1996) *Science* **274**, 174–175

[m] Saha, D. P., and Subramanian, K. S. (1996) *Science* **274**, 1287–1288

[n] Nielsen, F. H. (1991) *FASEB J.* **5**, 2661–2667

---

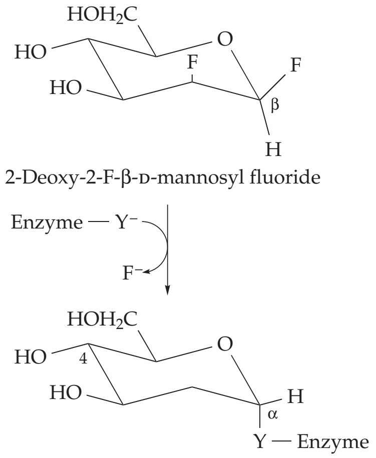
carboxyl groups. Such an intermediate, fructofuranosyl-enzyme, was identified tentatively for sucrose phosphorylase and also for a related levan sucrase.[45] Recently, identification of glycosyl-enzyme intermediates has been accomplished for many other glycosyltransferases. Among these are human pancreatic and salivary α-amylases,[42,46] α-glucosidases, and some cellulases and xylanases.



Fructofuranosyl - enzyme

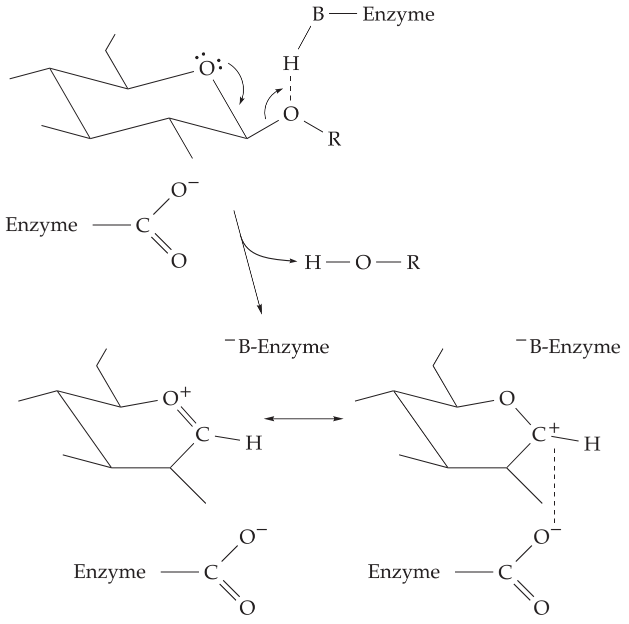
For these hydrolytic enzymes the glycosyl-enzyme would be attacked by a hydroxyl ion derived from H_2O, whose deprotonation would presumably be assisted by the conjugate base (—B: in Eq. 12-5) of the catalytic acidic group.

A convenient way to form and identify covalently linked glucosyl-enzyme intermediates, developed by Withers and coworkers, employs enzyme-activated inhibitors such as 2-deoxy-2-fluoroglycosyl fluorides (Eq. 12-10).[42,47,48] The 2-F substituent greatly decreases both the rate of formation of a glycosyl-enzyme and its rate of breakdown by hydrolysis or transglycosylation. This may be in part because these compounds lack the 2–OH group which helps to stabilize, by hydrogen bonding, the complexes of normal substrates. A second factor is the high electronegativity of –F versus –OH, which leads to significant loss of stability for a transition state in which the anomeric carbon atom carries a significant amount of positive charge (see next section). Having a good leaving group, such as fluorine or 2,4-dinitrophenyl, the compounds react rapidly to give stable glucosyl-enzymes which can be characterized. In the example shown in Eq. 12-10, the mannosyl-enzyme was identified by the chemical shifts and line-widths of the ^{19}F resonances of the intermediates and the anomeric configuration was established. More recently, mass spectrometry has been used.[49] For example, a maltotriosyl-enzyme intermediate in the action of glycogen-debranching enzyme was identified by separation of an active site peptide by HPLC followed by mass spectrometry (Chapter 3).[50]



(12-10)

Carbocationic intermediates. In a second mechanism of nucleophilic displacement the leaving group departs (often in a protonated form) before the entering nucleophile reacts.

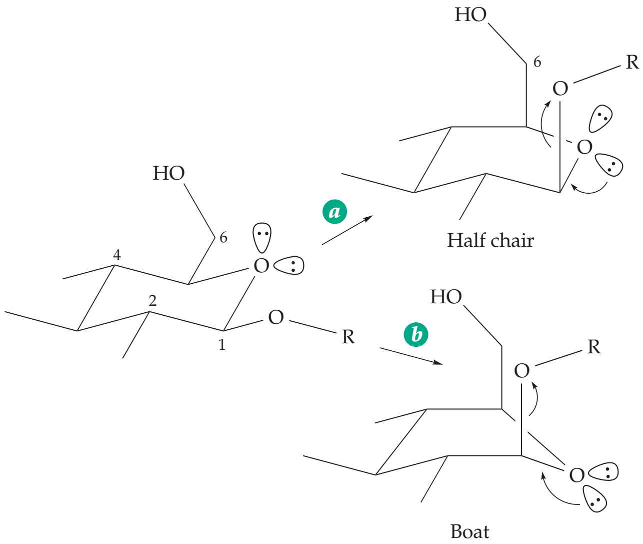


Oxocarbenium ion, a stabilized carbocation

(12-11)

A carbocation is formed as shown in Eq. 12-11, which represents just one-half of the overall displacement reaction. In the common terminology of physical organic chemistry this is an S_N1 reaction rather than an S_N2 reaction of the kind shown in Eqs. 12-3 and 12-5. This terminology is not quite appropriate for enzymes because the breakdown of ES complexes to product is usually a zero-order process and the numbers 1 and 2 in the symbols S_N1 and S_N2 refer to the order or molecularity of the reaction. It is better to speak of S_N1-like or S_N2-like reactions.

The carbocation in Eq. 12-11 is depicted as a resonance hybrid of two forms. One of these is an **oxocarbenium** (or oxonium) ion which contains a double bond between carbon and oxygen.[50a] This double-bonded structure can be visualized as arising from the original structure by an internal displacement or elimination by the unshared electrons on oxygen, as indicated by the small arrows. In the oxocarbenium ion, the ring atoms C-2, C-1, and C-5 and the oxygen atom are almost coplanar and the ring conformation is described as **half-chair** or **sofa**. As pointed out in Chapter 9, the theory of stereoelectronic control predicts that elimination of the group –OR of Eq. 12-11 should occur most easily if a lone pair of electrons on the ring oxygen atom are antiperiplanar to the O–R bond. This is impossible for a β-glycoside with a ring in the chair conformation shown in Eq. 12-11. This fact suggested that enzymes cleaving β-glycosides may preferentially bind substrate with the appropriate sugar ring in a less stable half-chair or flexible boat conformation prior to bond cleavage (Eq. 12-12).[51,52] This would allow an unshared pair of electrons on oxygen antiperiplanar to the C-1 to O – R bond to participate in the elimination reaction as is indicated in the following diagrams.



(12-12)

On the other hand, Bennet and Sinnott provided evidence that an antiperiplanar lone electron pair is *not* needed in acid-catalyzed cleavage of glycosides via a carbocationic intermediate.[53,54] Theories of stereoelectronic control must be applied to enzymes with caution!

5. Lysozymes and Chitinases

Polysaccharide chains in the peptidoglycan layer (Fig. 8-29) of the cell walls of bacteria are attacked and cleaved by lysozymes,[55] enzymes that occur in tears and other body secretions and in large amounts in egg white. Some bacteria and fungi, and even viruses, contain lysozymes.[56] Their function is usually to protect against bacteria, but lysozyme of phage T4 is a component of the baseplate of the virus tail (Box 7-C). Its role is to cut a hole in the bacterial cell wall to permit injection of the virus' own DNA. Egg white lysozyme, the first enzyme for which a complete three-dimensional structure was determined by X-ray diffraction,[55] is a 129-residue protein. The active site is in a cleft between a large domain with a nonpolar core and a smaller β-sheet domain that contains many hydrogen-bonded polar side chains (Figs. 12-3, 12-4). Human lysozyme has a similar structure and properties.[57–59] The T4 lysozyme has an additional C-terminal domain whose function may be to bind the crosslinking peptide of the *E. coli* peptidoglycan. Goose lysozyme is similar in part to both hen lysozyme and T4 lysozyme. All three enzymes, as well as that of our own tears, may have evolved from a common ancestral protein.[60] On the other hand, *Streptomyces erythaeus* has developed its own lysozyme with a completely different structure.[61] An extensive series of T4 lysozyme mutants have been studied in efforts to understand protein folding and stability.[61–63]

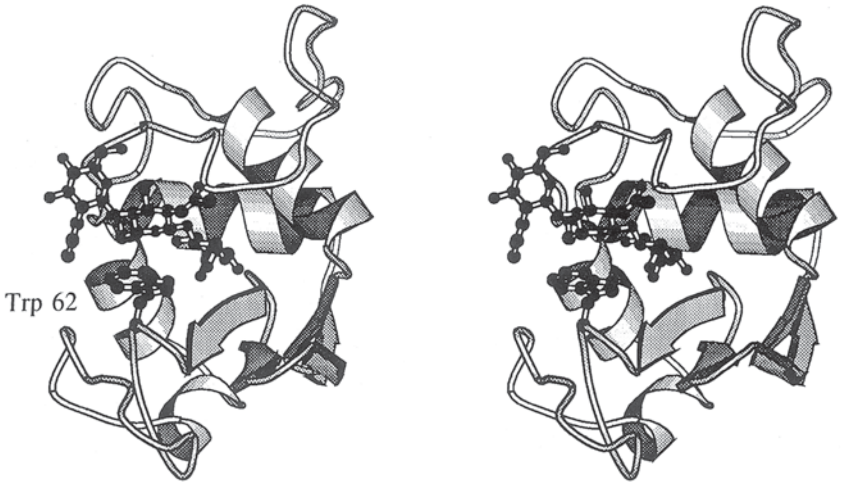


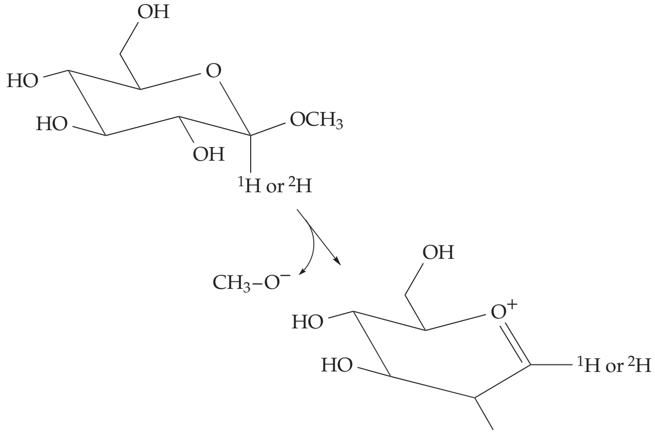
Figure 12-3 Stereoscopic MolScript view of hen egg white lysozyme complexed with a trisaccharide of *N*-acetylglucosamine $(GlcNAc)_3$ in binding subsites A, B, and C. The side chain of Trp 62 and the trisaccharide above it are shown in ball-and-stick form. From Maenaka *et al.*[65] Courtesy of Izumi Kumagai.

Catalytic side chain groups. Six *N*-acetylglucosamine (GlcNAc) or *N*-acetylmuramic acid (MurNAc) rings of the polysaccharide substrate are able to fit precisely into six subsites (designated A to F) in a groove between the two structural domains of egg white lysozyme (Fig. 12-4). The bond between the fourth and fifth rings (subsites D and E) is then cleaved. At the active site, the side chain carboxyl group of Glu 35 is positioned correctly to serve as the proton donor BH of Eq. 12-11, while the carboxyl of Asp 52 lies on the opposite side of the groove. Both Glu 35 and Asp 52 have relatively high pK_a values for carboxyl (microscopic pK_a's are ~5.3 and 4.6, respectively, in the fully protonated active site when the ionic strength is ~ 0.2)[64] as a result of the hydrophobic environment and hydrogen bonding to other groups. If the pH is raised from an initially low value, a proton from Asp 52 usually dissociates first and the electrostatic field of the resulting anion keeps Glu 35 largely protonated until the pH approaches 6. Positively charged basic groups nearby affect the pK_a values; hence, the behavior of the enzyme is sensitive to the ionic strength of the medium.[63] The Asp 52 anion lies only ~ 0.3 nm from the center of positive charge expected for an oxocarbenium ion intermediate and presumably stabilizes it. Replacement of Glu 35 by Gln destroys all catalytic activity and replacement of Asp 52 by Ala or Asn decreases activity to ~4–5% of the original.[66–68] Less than 1% activity remained for the D52S mutant.[69]

Nevertheless, Asp 52 is not absolutely essential for lysozyme activity. Goose lysozyme lacks this catalytic aspartate.[70] Matsumura and Kirsch suggested that carboxyl groups of glycine residues covalently attached to *N*-acetylmuramic acid rings in the natural substrates may participate in catalysis.[66] Many mutant forms of lysozyme have been studied. Of special interest is the D52E mutant. In this enzyme the carboxylate of the longer glutamate side chain reacts with the oxocarbenium ion intermediate to form a covalent adduct and apparently alter the basic mechanism.[71] Replacement of asparagine 46, which can hydrogen bond to Asp 52, also decreases k_{cat} greatly, suggesting a role in catalysis.[72]

The lysozyme-catalyzed reaction is completed by stereospecific addition of a hydroxyl ion to the oxocarbenium ion with the original β configuration being retained in the product. Such stereospecificity for reactions of enzyme-bound carbocations is not surprising because the enzyme probably assists in generation, on the appropriate side of the sugar ring, of the attacking hydroxyl ion.

Kinetic isotope effect for lysozyme. A secondary kinetic isotope effect is expected because a molecule with ^{1}H in the number 1 position can be converted to the corresponding oxocarbenium ion somewhat more easily than the molecule with ^{2}H in the same position (Eq. 12-13). For example, in the nonenzymatic acid-catalyzed hydrolysis of a methyl- glucoside, a reaction also believed to proceed through a carbocation intermediate,[41,75] the ratio k_{1_H} / k_{2_H} is 1.14 for the α anomer and 1.09 for the β anomer.[53] In the base-catalyzed hydrolysis of the same compound, which is believed to occur by a double-displacement reaction involving participation of the neighboring OH group on C-2, the ratio k_{1_H} / k_{2_H} is 1.03. The corresponding ratio measured



(12-13)

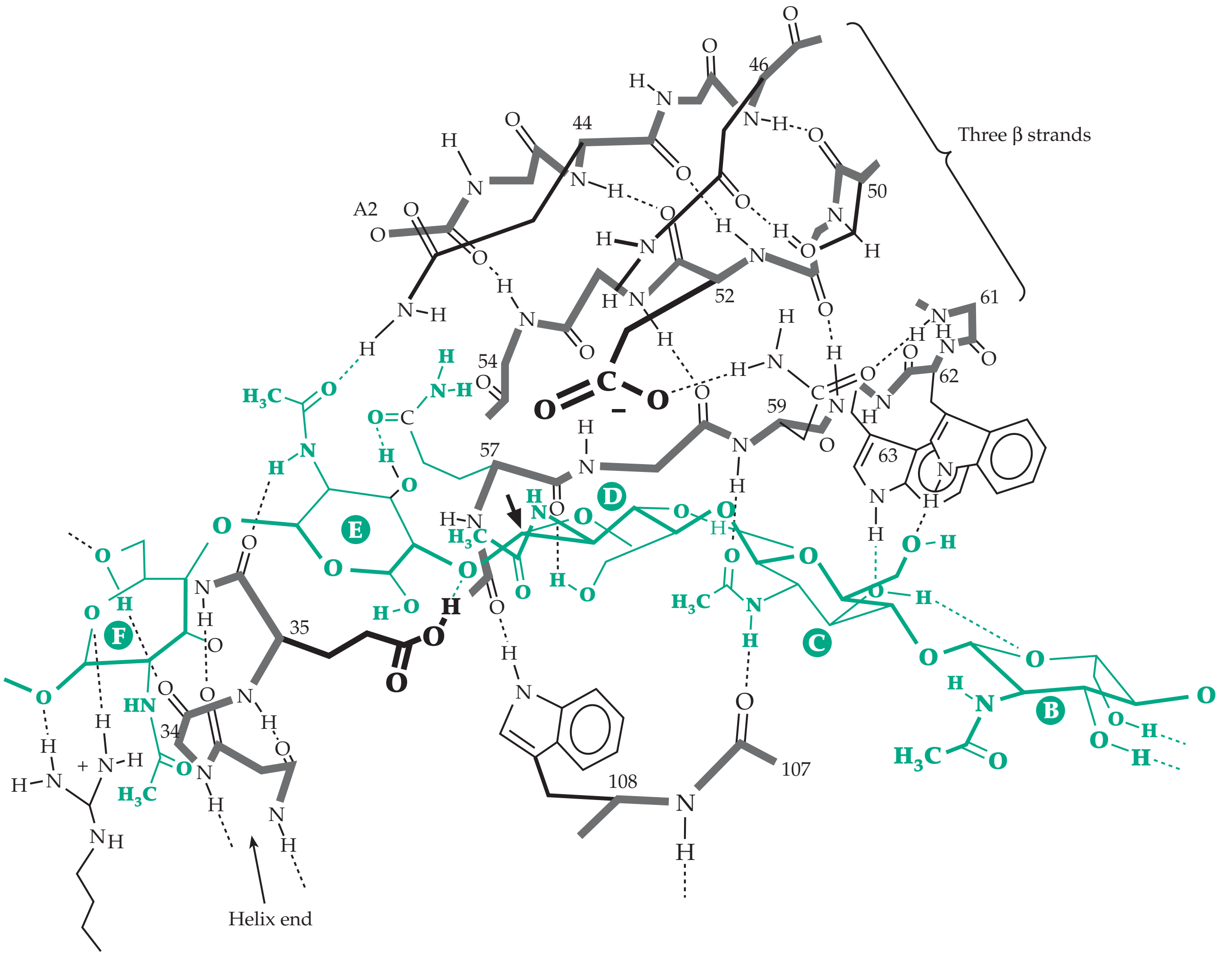


Figure 12-4 Schematic drawing of the active site of egg white lysozyme with substrate in place and about to be cleaved. Three strands of the small β-sheet domain, which contains an extensive hydrogen bond network, are seen at the top. Notice that this view into the active site is different from that in Fig. 12-3. A 180° rotation of either figure will make the views more similar. The three-strand β sheet can be seen (in stereo) at the lower right in Fig. 12-3 and the helix end carrying Glu 35 in the right center foreground. A segment of a chitin oligosaccharide (green), whose reducing end is to the left, is bound into subsites B –F. Cleavage occurs between rings D and E as indicated by the heavy arrow which points to the anomeric carbon atom of ring D. The side chain of Glu 35 is shown protonating the bridge oxygen. Ring D has been distorted into a twist conformation to facilitate cleavage. A larger domain, which contains one long α helix, is at the bottom of the drawing. The active site lies in a cleft between the domains. Notice the chain of hydrogen bonds between the carbonyl of residue 107 in the large domain and the peptide NH of residue 59 in the small domain. Also notice the aromatic side chains which are usually present in carbohydrase active sites. Drawing is based in part on those of Irving Geis[73] and Levitt[74] and on a sketch by author from a three-dimensional model.

for the action of lysozyme is 1.11, much closer to that of the oxocarbenium ion mechanism than to that of the double-displacement mechanism. Similar observations have been made with amylases.[76] Kinetic isotope effects have also been measured for ^{12}C vs ^{13}C in the anomeric position, ^{16}O vs ^{18}O in the leaving group ($-OCH_3$) of the methyl glucosides, and for other locations[53] as well as for hydrolysis of glucosyl fluorides.[77–79] The results are generally supportive of the oxocarbenium ion mechanism for lysozyme. However, as mentioned in Section B,3, the interpretation of secondary isotope effects is difficult. Such effects cannot reliably identify a carbohydrase mechanism.[80,81]

For acid-catalyzed hydrolysis of methyl glucosides[53] the kinetic isotope effect observed for the oxygen of the leaving group was $k_{^{16}O} / k_{^{18}O} = 1.024-1.026$. Observation of similar effects for enzymes supports the participation of an acidic group of the protein (Glu 35 of lysozyme) in catalysis but does not eliminate the possibility of concerted involvement of a nucleophilic group, e.g., Asp 52 in lysozyme.[81,82]

Does lysozyme distort its substrate? An early study of models indicated that for six sugar rings of a substrate to bind tightly into the active site of lysozyme, the ring in subsite D, which contains the carbon atom on which the displacement occurs, had to be distorted from its normal chair conformation into the half-chair conformation.[83] This is illustrated in Fig. 12-4. It was suggested that by binding the substrate chain at six different sites the enzyme provides leverage to distort the ring in subsite D into a conformation similar to that of the transition state. This idea was criticized on the basis that an enzyme would be too flexible to act in this manner.[74] Furthermore, the non-hydrolyzed trisaccharide MurNAc–GlcNac–MurNAc was shown to fit into subsites B, C, and D of the active site groove without distortion.[84] However, it is bound weakly.

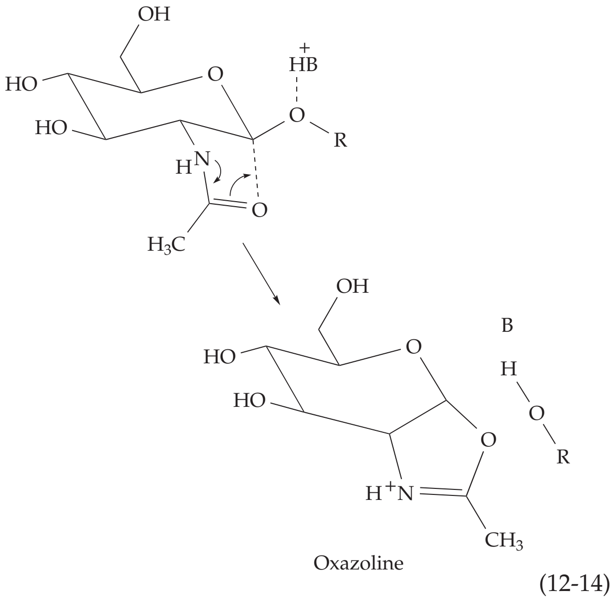
Can electrical forces acting within the active site help to distort the substrate and assist in formation of the carbocation intermediate? Levitt and Warshel suggested this possibility and proposed that the necessary electrostatic force arises from the arrangement of dipoles within the peptide backbone of the protein and in the amino acid side chains.[74] As can be seen from Fig. 12-4, the enzyme forms many hydrogen bonds with the substrates. Of special interest is a chain of hydrogen bonds that passes from the backbone carbonyl of Ala 107 through the 2-acetamido group of the substrate in subsite C and into the edge of the β sheet at the backbone NH of Asn 59. This interaction provides specificity toward GlcNac-containing substrates. Perhaps oscillation of charge within such polarizable chains of hydrogen bonds can also help a substrate to move toward its transition state structure.[85,86]

From simulations of lysozyme action by the methods of molecular dynamics, Post and Karplus observed that Glu 35 tended not to hydrogen bond to the exocyclic bridge oxygen, even though that would be a logical step in the protonation (by –BH) shown in Eq. 12-11. They suggested that Glu 35 may protonate the ring oxygen.[87] The ring could then open to form an exocyclic oxonium ion, which could hydrolyze and cyclize to the final product. The initial elimination could receive stereoelectronic assistance from the antiperiplanar lone pair of electrons on the bridge oxygen. This proposal has been criticized.[81,88] For example, it is hard to explain the rapid cleavage of glycosyl fluorides by this exocyclic oxonium mechanism.

As attractive as the arguments for a oxocarbenium mechanism in lysozyme are, is it still possible to explain the experimental observations by a double displacement in which Asp 52 serves as the nucleophile to form a glycosyl enzyme?[38,82,82a] Sucrose phosphorylase and other glucosyl transferases, like lysozyme, apparently have a carboxylate ion at the active site. In some enzymes, the carboxylate ion forms a covalent glucosyl enzyme, whereas in lysozyme it apparently only stabilizes a carbocation. Is there really a difference in mechanism? Do glucosyl enzymes form only with certain substrates or upon denaturation of the enzymes? Nature hides her secrets well. The difficulty in pinning down the fine mechanistic details of enzymatic action makes it essential to be skeptical, to examine data critically, and to try to imagine all possible alternatives –even when things seem to be proven beautifully.

Help from a neighboring group. In the 1960s it was suggested that the acetamido group of *N*-acetylglucosamine residues might participate as a nucleophile, either stabilizing an oxocarbenium ion or forming an oxazoline intermediate.[89–91] The proposal received little support, as applied to lysozyme, until sequences and structures of many larger carbohydrases were determined. Among these are **chitinases**, enzymes that act on the same substrates as the small lysozymes. One group of plant chitinases have a structure similar to that of egg white lysozyme. The 243-residue enzyme from barley seeds apparently has Glu 67 as a proton donor and Glu 89 as a possible stabilizing nucleophile.[92] Another group of chitinases from both plants and bacteria have active sites at ends of $(\alpha\beta)_6$ barrels.[93–95] They all have a proton-donating glutamate in a conserved position but *no aspartate* that could serve as a nucleophile. Study of complexes with substrates and inhibitors with these enzymes has provided direct evidence of ring distortion and of the probable role of the acetamido group as indicated in Eq. 12-14.[94,95] The distortion is beyond that in a sofa conformation and allows for maximum stereoelectronic assistance as well as participation of the acetamido group.

In these relatively large enzymes the substrate is deeply buried and cannot reach a conformation



(12-14)

approaching that of the transition state without having a strained ring conformation. These enzymes may use binding forces exerted on many parts of the substrate to stabilize the transition state structure.

6. Cellulases and Other β-Glycosidases

Cellulose, the most abundant of all biopolymers, is extremely stable but is attacked by a host of bacterial and fungal β-glycanases.[96] Animals do not ordinarily produce cellulases but some termites do.[97] Cellulase structures are varied, being represented by 10 of 57 different glycosylhydrolase families.[98] Most, like lysozyme, retain the β configuration in their products but some invert.[98–100]

Among the cellulases are **endoglycanases**, which cleave chains at random positions, and **exoglycanases** (also called **cellobiohydrolases**) that cleave cellobiose units from ends of chains. Some act on the nonreducing ends and some on reducing ends; a mixture of enzymes is most effective.[99,101,102] In some bacteria a whole series of different cellulases, together with a large (197-kDa) organizing protein, form a **cellusome**, a complex with high catalytic activity for crystalline cellulose.[103–105] Cellulases usually have tightly packed catalytic domains which may vary in size from ~200 residues[98] to over 400.[106] In most cases the catalytic domain is connected by a flexible linker to one or more small, globular **cellulose-binding domains**. These vary in size from 36 to 200 residues and often have a β-barrel fold.[104,107–108a] Their function is to hold the enzymes to the cellulose surfaces. They may also facilitate disruption of the tightly hydrogen bonded cellulose structure (Fig. 4-5). As with other carbohydrases, carboxyl groups of amino acid side chains provide the major catalytic groups (Table 12-1). An extensive hydrogen-bond network which often includes imidazole groups may influence activity.[109,110]

A striking feature of a 411-residue endoglucanase from *Fusarium* was revealed by the binding of a nonhydrolyzable thiooligosaccharide substrate analog. The pyranose ring at the cleavage site was distorted in an identical manner to that mentioned in the preceding section for a chitinase thought to use the substrate's acetamido group as a nucleophile (Eq. 12-14).[106] The distortion observed is beyond that required for a sofa conformation and allows for the maximum stereoelectronic assistance (Eq. 12-12*b*). An oxygen atom of the E197 carboxylate, the catalytic nucleophile, occupies a position in the complex that is coincident with that of the C2 acetamido oxygen in the catalytic site of the chitinase discussed in the previous section.

The *inverting* β-glucanases also have two catalytic acid base groups but they are ~0 .9 – 1.0 nm apart rather than ~ 0.6 nm for retaining enzymes. This allows space for a water molecule whose ^-OH is the nucleophilic reactant (^-OY in Eq. 12-5) and in which a carboxylate group assists in dissociating the water molecule.[98] (This mechanism is illustrated for glucoamylase in Fig. 12-7).

Structural features met in some cellulases include an α,α barrel[111] similar to that of glucoamylase (Fig. 2-29) and, in a cellobiohydrolase,[101] a 5-nm-long tunnel into which the cellulose chains must enter. Ten well-defined subsites for glycosyl units are present in the tunnel.[101] A feature associated with this tunnel is **processive** action, movement of the enzyme along the chain without dissociation,[105] a phenomenon observed long ago for amylases (see Section 9) and often observed for enzymes acting on nucleic acids.

Another group of **β-glucanases**, found in plants and their seeds, hydrolyze β-1,3-linked glucans[112] and, in some cases, also mixed 1,3- and 1,4-linked polysaccharides. A characteristic enzyme is found in barley (Table 12-1).[113,114] Similar enzymes are produced by some bacteria[113] and also by molluscs.[115]

Xylanases act on the β-1,4-linked xylan, the most abundant of the **hemicelluloses** that constitutes over 30% of the dry weight of terrestrial plants.[116] They resemble cellulases and cooperate with cellulases and **xylosidases**[117] in digestion of plant cell walls.[110,116,118–121] **Galactanase** digests the β-1,4-linked component of pectins.[122]

The final step in degradation of cellulose is hydrolysis of cellobiose to glucose. This is accomplished by **β-glucosidases**, enzymes that also hydrolyze lactose, phosphorylated disaccharides, and cyanogenic glycosides.[123 – 126] Lactose is also cleaved by **β-galactosidases**.[127] The large 1023-residue β-galactosidase from *E. coli*[128] is famous in the history of molecular biology as a component of the *lac* operon.[129 – 132] Its properties are also employed to assist in the cloning of genes (Chapter 26).

TABLE 12-1
Acidic and Basic Catalytic Groups in a Few Glycosyltransferases[a]

Enzyme	Number of residues	Inverting?	Glycosyl-enzyme identified?	Nucleophile $(-COO^-)$	Electrophile (proton donor) * Assisting group
Lysozyme					
Human, hen[b]	130	No		E35	D53(52)
Bacteriophage T4[c]		No		E11	D20
Chitinase					
Rubber plant[d]	273	No			E127
Cellulases					
E1 endocellulase,					
Acidothermus[e]	521				
catalytic domain	358	No		E282	E162, D252*
Endoglucanase I					
Fusarium[f]	411	No		E197	E202
Endoglucanase Cen A					
Cellulomonas[g]	>351	Yes		D392	D78
1,3-β-D-Glucanase					
Barley[h]		No		E231	E288
Xylanase					
B. circulans[i]		No	Yes	E78	E172
β-Glucosidase					
Agrobacterium[j]	458	No	Yes	E358	E170, Y298*
β-Galactosidase					
E. coli[k]	1023	No	Yes	E537	E461, Y503*
α-Amylase					
Human and pig[l]	496	No		D197	E233, D300*
Barley[m]		No		D179	E204, D289*
Aspergillus[n]		No		D206	E230
Cyclodextrin glucosyltransferase					
B. circulans[o]				D229	E257, D328
α-Glucosidase					
Saccharomyces[p]	No	Yes		D214	E233, D300
Glycogen debranching enzyme, rabbit[s]	No	Yes		D549	
Glucoamylase					
Aspergillus[q]	616	Yes		H_2O, E400	E179
β-Amylase					
Soybean[r]		Yes		E186	E380
Glucocerebrosidase					
Human[t]				E340	

[a] For classification of glycosyl hydrolases into families, see Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.-P., and Davies, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7090–7094

[b] Matsumura, I., and Kirsch, J. F. (1996) *Biochemistry* **35**, 1881–1889; Harata, K., Muraki, M., Hayashi, Y., and Jigami, Y. (1992) *Protein Sci.* **1**, 1447–1453

[c] Brzozowski, A. M., and Davies, G. J. (1997) *Biochemistry* **36**, 10837–10845; Hardy, L. W., and Poteete, A. R. (1991) *Biochemistry* **30**, 9457–9463 Kuroki, R., Weaver, L. H., and Matthews, B. W. (1993) *Science* **262**, 2030–2033

[d] Tews, I., Terwisscha van Scheltinga, A. C., Perrakis, A., Wilson, K. S., and Dijkstra, B. W. (1997) *J. Am. Chem. Soc.* **119**, 7954–7959

[e] Sakon, J., Adney, W. S., Himmel, M. E., Thomas, S. R., and Karplus, P. A. (1996) *Biochemistry* **35**, 10648–10660; Ghidoni, R., Sonnino, S., Tettamanti, G., Baumann, N., Reuter, G., and Schauer, R. (1980) *J. Biol. Chem.* **255**, 6990–6995

[f] Sulzenbacher, G., Schülein, M., and Davies, G. J. (1997) *Biochemistry* **36**, 5902–5911; Sulzenbacher, G., Driguez, H., Henrissat, B., Schülein, M., and Davies, G. J. (1996) *Biochemistry* **35**, 15280–15287; Mackenzie, L. F., Davies, G. J., Schülein, M., and Withers, S. G. (1997) *Biochemistry* **36**, 5893–5901

[g] Damude, H. G., Withers, S. G., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. J. (1995) *Biochemistry* **34**, 2220–2224

[h] Chen, L., Garrett, T. P. J., Fincher, G. B., and Hoj, P. B. (1995) *J. Biol. Chem.* **270**, 8093–8101

[i] Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J., and Yaguchi, M. (1994) *Protein Sci.* **3**, 467–475; Lawson, S. L., Wakarchuk, W. W., and Withers, S. G. (1997) *Biochemistry* **36**, 2257–2265; Sidhu, G., Withers, S. G., Nguyen, N. T., McIntosh, L. P., Ziser, L., and Brayer, G. D. (1999) *Biochemistry* **38**, 5346–5354

[j] Wang, Q., Trimbur, D., Graham, R., Warren, R. A. J., and Withers, S. G. (1995) *Biochemistry* **34**, 14554–14562

[k] Gebler, J. C., Aebersold, R., and Withers, S. G. (1992) *J. Biol. Chem.* **267**, 11126–11130; Jacobson, R. H., Zhang, X.-J., DuBose, R. F., and Matthews, B. W. (1994) *Nature (London)* **369**, 761–766; Richard, J. P., Huber, R. E., Lin, S., Heo, C., and Amyes, T. L. (1996) *Biochemistry* **35**, 12377–12386

[l] Brayer, G. D., Luo, Y., and Withers, S. G. (1995) *Protein Sci.* **4**, 1730–1742; Qian, M., Haser, R., Buisson, G., Duée, E., and Payan, F. (1994) *Biochemistry* **33**, 6284–6294

[m] Kadziola, A., Sogaard, M., Svensson, B., and Haser, R. (1998) *J. Mol. Biol.* **278**, 205–217

[n] Brzozowski, A. M., and Davies, G. J. (1997) *Biochemistry* **36**, 10837–10845; Matsuura, Y., Kusunoki, M., Harada, W., and Kakudo, M. (1984) *J. Biochem.* **95**, 697–702

[o] Knegtel, R. M. A., Strokopytov, B., Penninga, D., Faber, O. G., Rozeboom, H. J., Kalk, K. H., Dijkhuizen, L., and Dijkstra, B. W. (1995) *J. Biol. Chem.* **270**, 29256–29264

[p] McCarter, J. D., and Withers, S. G. (1996) *J. Biol. Chem.* **271**, 6889–6894

[q] Christensen, U., Olsen, K., Stoffer, B. B., and Svensson, B. (1996) *Biochemistry* **35**, 15009–15018

[r] Mikami, B., Degano, M., Hehre, E. J., and Sacchettini, J. C. (1994) *Biochemistry* **33**, 7779–7787; Adachi, M., Mikami, B., Katsube, T., and Utsumi, S. (1998) *J. Biol. Chem.* **273**, 19859–19865

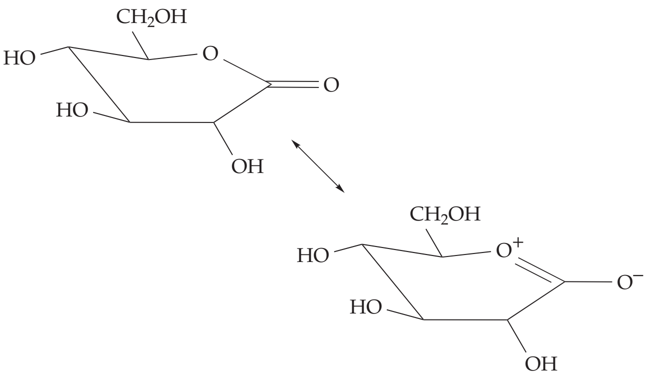
[s] Braun, C., Lindhorst, T., Madsen, N. B., and Withers, S. G. (1996) *Biochemistry* **35**, 5458–5463

[t] Withers, S. G., and Aebersold, R. (1995) *Protein Sci.* **4**, 361–372

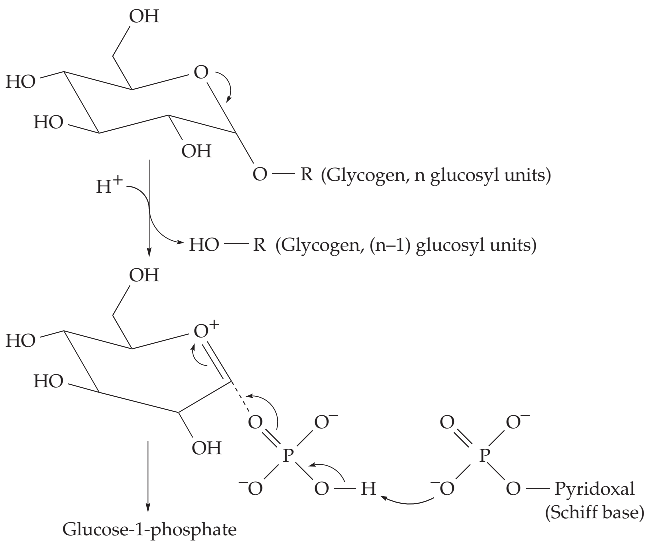
7. Glycogen Phosphorylase

A large number of different glycosyltransferases act on the α1,4 linkages of glycogen, starch, and related polysaccharides. Among these, one of the most studied is glycogen phosphorylase. It is not a hydrolase, but it catalyzes cleavage of α1,4 linkages at the nonreducing ends of glycogen molecules by displacement with inorganic phosphate to give α-D-glucose-1-phosphate. It is a very large enzyme (841 residues in rabbit muscle) whose structure is shown in Fig. 11-5. Its complex regulatory mechanisms were discussed briefly in Chapter 11.

Neither partial exchange reactions nor inversion of configuration occur when glycogen phosphorylase acts on its substrates. The enzyme apparently does nothing until both substrates are present. These are glycogen + inorganic phosphate or, for the reverse reaction, glycogen (shifted over by one sugar binding subsite) + glucose-1-phosphate. The active site is in a deep groove in the enzyme. Evidently, the protein must close and fold around the substrates before it becomes active.[133] An oxocarbenium ion mechanism has also been proposed for this enzyme, partly on the basis of strong inhibition of phosphorylase by **5-gluconolactone**,[134,135] a compound having a half-chair conformation and perhaps acting as a transition state inhibitor (Chapter 9). This gluconolactone also inhibits many other carbohydrases.[123,136]



5-Gluconolactone



(12-15)

An unexpected discovery was that glycogen phosphorylase contains a molecule of the coenzyme pyridoxal 5'-phosphate (PLP) bound into the center of the protein behind the active site[134,135,137–139] with its phosphate group adjacent to the binding site of the phosphate of glucose-1-phosphate. It probably serves as a general base catalyst, e.g., assisting the attack of a phosphate ion on the oxocarbenium formed by cleavage of the glycogen chain (Eq. 12-15).[137,140–142] A key observation by Graves and associates was that pyridoxal alone does not activate apo-phosphorylase but that pyridoxal plus phosphite, phosphate, or fluorophosphate does provide up to 20% of full activity.[143] X-ray crystallography confirmed that these activating anions are bound into the active site at the approximate position of the phosphate group of PLP.[139,143a] Glycogen phosphorylase is being studied by X-ray diffraction techniques that allow observation of structural changes in as short a time as a few milliseconds (see Chapter 3, Section H).[144]

The regulation of glycogen phosphorylase, like that of many other allosteric proteins, depends upon the existence of two distinct conformational states, whose structures have been established by crystallography.[137] It is not immediately evident how they can affect the active site. In the R-state the enzyme has a low affinity for both substrates and activators such as AMP. In the T-state the affinities are much higher. For example, that of inorganic phosphate is raised by a factor of fifteen.[137] As we have already seen (Figs. 11-4, 11-5) the relatively inactive phosphorylase *b* is converted to the active phosphorylase *a* by phosphorylation of the side chain of serine 14, a structural change which favors the R-state. In the inactive T-state of phosphorylase *b*, an 18-residue N-terminal segment of the polypeptide is not seen by X-ray diffraction, presumably because it does not assume a fixed conformation but projects into a solvent channel in the crystal and moves freely within it. However, in the active R-state of phosphorylase *a*, in which Ser 14 has been phosphorylated, the N-terminal segment is rigid. The phospho group on Ser 14 binds to Arg 69 and other arginine, lysine, and histidine side chains from both subunits.[137,145] The phosphorylation of Ser 14 occurs 1.5 nm from the active sites. The conformational changes induced by phosphorylation of Ser 14 cause a rotation of about 10° between subunits, somewhat reminiscent of the changes accompanying oxygenation of hemoglobin (Fig. 7-25).

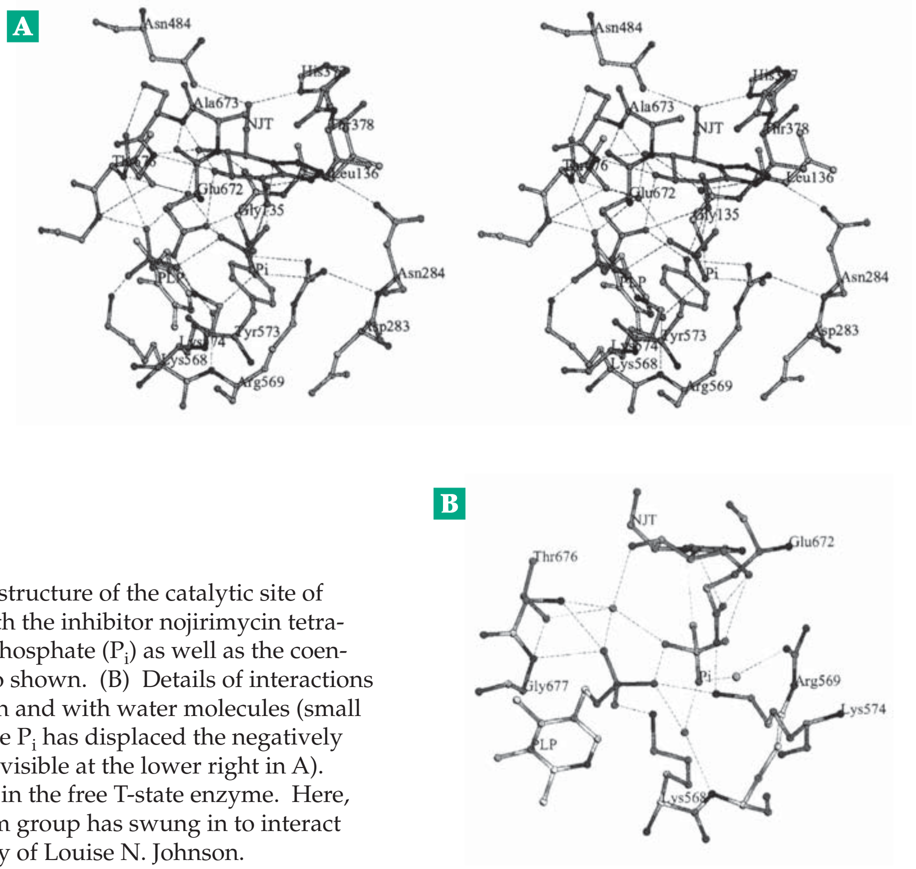
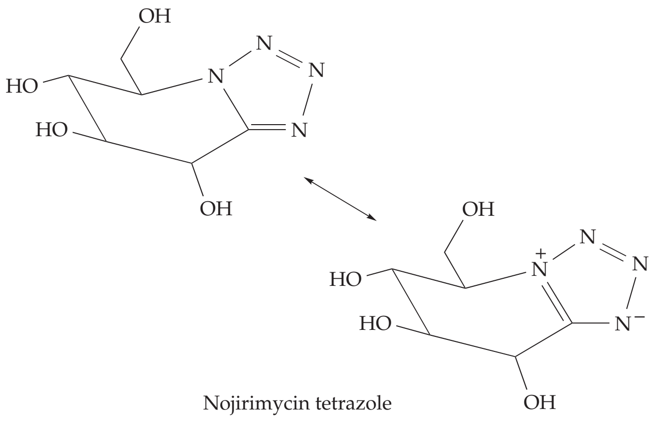


Figure 12-5 (A) Stereoscopic view of the structure of the catalytic site of phosphorylase *b* in the inhibited T-state with the inhibitor nojirimycin tetrazole bound into the active site. Inorganic phosphate (P_i) as well as the coenzyme pyridoxal 5'-phosphate (PLP) are also shown. (B) Details of interactions of the inhibitor, P_i, and PLP with the protein and with water molecules (small circles). This is a weak-binding state but the P_i has displaced the negatively charged side chain carboxylate of Asp 283 (visible at the lower right in A). This carboxylate blocks access to the P_i site in the free T-state enzyme. Here, the positively charged Arg 569 guanidinium group has swung in to interact with the P_i. From Mitchell *et al.*[133] Courtesy of Louise N. Johnson.

A related structural change accompanies binding of caffeine, adenosine, or AMP at high concentrations. These substances inhibit the enzyme by binding at a site next to the catalytic site (Fig. 11-5), stabilizing the T-state and causing a loop of protein to move into the catalytic site and to block it.[137] The catalytic site lies deeply buried in the protein between the two large structural domains (Fig. 11-5). The substrate is held by a network of hydrogen bonds, some of which are shown in Fig. 12-5, in which the active site contains an inhibitory substrate analog **nojirimycin tetrazole**, which is viewed toward the edges of the two rings.



Nojirimycin tetrazole

The enzyme is in the weak-binding T-state but inorganic phosphate (P_i) is bound below it and next to the phosphate group of PLP, as required by the mechanism of Eq. 12-15.

The difference in binding affinities for the T- and R-states lies in a flexible loop of residues 280–288, which in the T-state blocks access to the substrate-binding cleft. The universally conserved Asp 238 behaves as a substrate mimic, occupying the P_i site.[137,146,147] In the R-state this residue moves, allowing P_i to enter and bind (Fig. 12-5).

Most of the structure of mammalian phosphorylases is conserved in species as diverse as *E. coli* and the potato.[148] However, *E. coli* maltodextrin phosphorylase[143a,149–151] and potato phosphorylases have less sophisticated regulatory mechanisms than do the animal enzymes. Another feature of glycogen phosphorylase is a "glycogen storage site" about 2.5 nm from the active site (Fig. 11-5).[152] This provides a means for the enzyme to hold onto the giant glycogen molecule while "nibbling off" the outside ends of nearby branches.

8. Starch-Hydrolyzing Enzymes

Among the hydrolases are the widely distributed **α-amylases**, *endo*-glycosidases which hydrolyze starch chains by random attack at points far from chain ends to form short polysaccharide chains known as **dextrins** as well as simpler sugars.[153] The catalyzed reactions proceed with retention of the original α configuration. Alpha-amylases are found in fungi, plants, and animals. One powerful enzyme of this class is present in the saliva of most humans and other isoenzymes are formed by the pancreas.[154] They are encoded by a family of genes, a fact that accounts for the existence of some healthy individuals completely lacking salivary amylase.[155]

Structures are known for human[154] and porcine[156–159] α-amylases as well as for corresponding enzymes from barley,[160] mealworms,[161] fungi, and bacteria.[162–164] The α-amylase from *Aspergillus oryzae* (Taka-amylase), widely used in laboratory work, was the first for which a structure was determined.[165,166] The α-amylases fold into three domains (Fig. 12-6), with the active site in the center of an $(\alpha/\beta)_8$ barrel. All of the α-amylases contain one or more bound Ca^{2+} ions. Some, including the human α-amylases, also require a **chloride ion**. The Cl^- is held by a pair of arginine guanidinium groups[154] and interacts with adjacent carboxyl groups, inducing pK_a shifts and allosteric activation.[163]

The sequences of α-amylases vary widely but a conserved cluster of one glutamate and two aspartates is usually present (Table 12-1).

Studies of the **action patterns**, i.e., the distribution of products formed by α-amylases when acting on a variety of α 1,4-linked oligosaccharides, suggested that the substrate binding region of the porcine pancreatic enzyme has five subsites, each binding one glucose residue.[153] The α-glucan chain is cleaved between the residues bound at the second and third subsites (numbered from the reducing end of the substrate) by a lysozyme-like mechanism. Endolytic enzymes, which cleave biopolymer chains internally, are usually thought to carry out random attack. However, after the initial catalytic reaction one of the polysaccharide products of porcine pancreatic amylase action often does not leave the enzyme. The polysaccharide simply "slides over" until it fully occupies all of the subsites of the substrate binding site and a second "attack" occurs. An average of seven catalytic events occur each time it forms a complex with amylase.[167] Is there a mechanism by which the enzyme deliberately promotes the "sliding" of the substrate in this **multiple attack** or **processive** mechanism[168,169] or does the dissociated product simply diffuse a short distance while enclosed in a solvent cage? The latter explanation may be adequate for some enzymes. In contrast to pancreatic α-amylase a bacterial "maltogenic" α-amylase produces principally maltose as a product.[169a,b]

Digestion of starch and glycogen by α-amylases produces a mixture of glucose, oligosaccharides, and dextrins, which in the human body are further degraded by **α-glucosidases** of the brush border membrane of the small intestine.[154] A lysosomal form of the enzyme is missing in Pompe's disease (Box 20-D).[170,171] The α-glucosidases are also members of the α-amylase family.[46,172] For digestion of the branched amylopectin and glycogen a **debranching enzyme** and **α-1,6-glucosidase** activity are required. In mammals these activities are found in a single polypeptide chain with separate but adjacent active sites.[50,173,174] The debranching enzyme catalyzes transfer of oligosaccharide chains from α1,6-linked branch positions to new locations at ends of chains with α1,4 linkages. A bacterial oligo-1,6-glucosidase has a catalytic site formed from

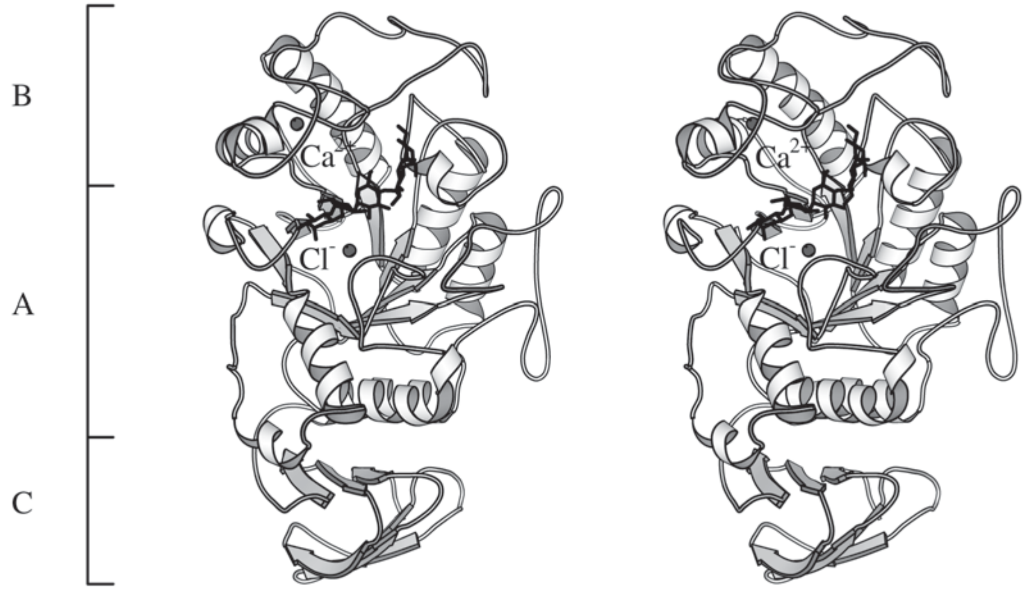
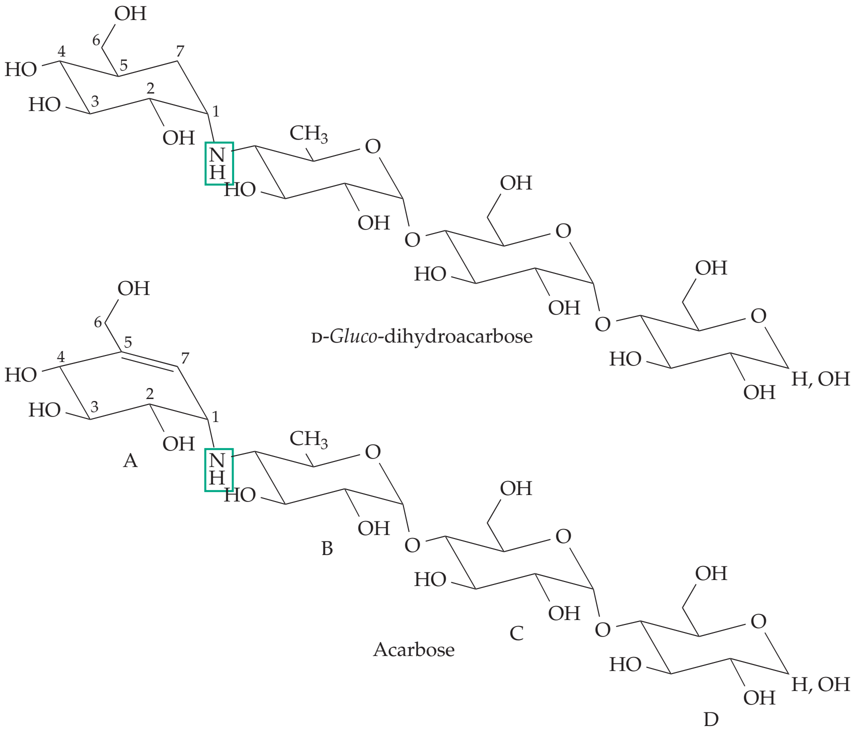


Figure 12-6 Drawing showing the overall polypeptide chain fold and relative positioning of the three structural domains of human pancreatic α-amylase. Also drawn are the locations of the calcium and chloride binding sites. Overlaid is the placement of a modified form of the inhibitor acarbose (p. 607) that binds in the active site cleft. MolScript drawing courtesy of G. Sidhu and G. Brayer.

D199, E255, and D329 similar to that of α-amylases.[175]

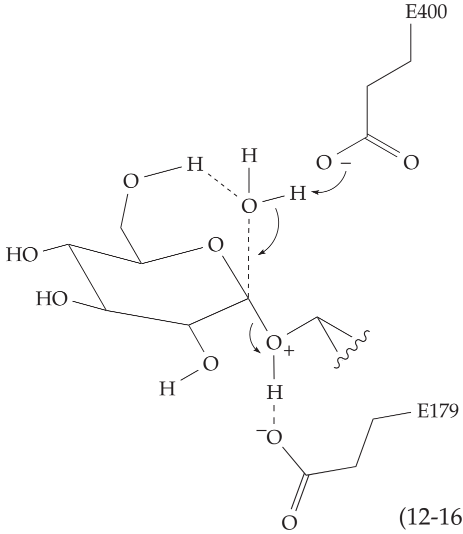
Another member of the α-amylase family has principally a glycosyltransferase activity. **Cyclodextrin glucanotransferase** forms cyclodextrins (Box 4-A) by a transferase reaction in which a 6- to 8- member oligo-saccharyl group is transferred from a straight amylase chain onto a protein side chain (Glu 257) and then joins the ends of the oligosaccharide to form the cyclo-dextrin rings (an overall double-displacement process).[176–177b] However, large circular dextrins are produced initially and are then converted into the smaller cyclodextrins.[178]

A starch-digesting enzyme of great industrial importance is **glucoamylase**, whose $(\alpha,\alpha)_6$-barrel structure is shown in Fig. 2-29. That figure also shows the tetrasaccharide inhibitor **acarbose** in the active site. The ring at the nonreducing end is deeply embedded in the protein and, as shown in Fig. 12-7 for the related D-*gluco*-dihydroacarbose, is held by many hydrogen bonds. This slow enzyme (k_{cat} ~ 50 s^{-1} at 45°C)[179] cuts off a single glucose unit, then releases itself from the starch before releasing the glucose and rebinding to the starch. The catalytic acid has been identified as the carboxyl group of Glu 179. In Fig. 12-7 it is seen, presumably as a carboxylate group, tightly hydrogen bonded to what is doubtless the bridge $-NH_2^+-$ between rings A and B of the inhibitor D-*gluco*-dihydroacarbose. Thus, the complex mimics that of a true substrate protonated on the bridge oxygen, a possible first step in normal catalysis. In accord with this mechanism (Eq. 12-16), glucoamylase is an inverting enzyme.



Notice in Fig. 12-7 the abundance of polar groups, many of which are charged. Four carboxyl groups, from D55, E179, E180, and E400, are present in the active site region and participate in the hydrogen bonded network. Active sites of other carbohydrases differ from that of glucoamylase. However, the presence of many hydrogen-bonded polar groups, including two or more carboxyl groups, is characteristic. As illustrated in Figs. 9-8 and 9-9, the pH dependence of the maximum velocity v_{max} is often determined by these carboxyl groups. At least one of them must be protonated for maximum catalytic activity and if two are protonated activity may fall again. Under these circumstances a single proton bound to a molecule of enzyme may spend a fraction of its time on each of several different carboxylate, imidazole, or other groups.[181] In the case of glucoamylase it was concluded from the X-ray structure that the catalytic acid group, E179, is probably *unprotonated* most of the time.[181] However, it can still bind to starch. Then, after a fraction of a second a proton may be transferred onto the Glu 179 carboxylate and from there to the bridge oxygen of the substrate, inducing reaction according to Eq. 12-16. It may even be necessary to have enzyme protonated initially on a group other than E179 to allow small conformational changes to occur prior to formation of the final activated complex. Such essential conformational changes have often been invoked for glucoamylase.[182]

Just as most cellulases have special cellulose-binding domains, glucoamylase has a compact C-terminal starch-binding domain (residues 509–616) similar to the



(12-16)

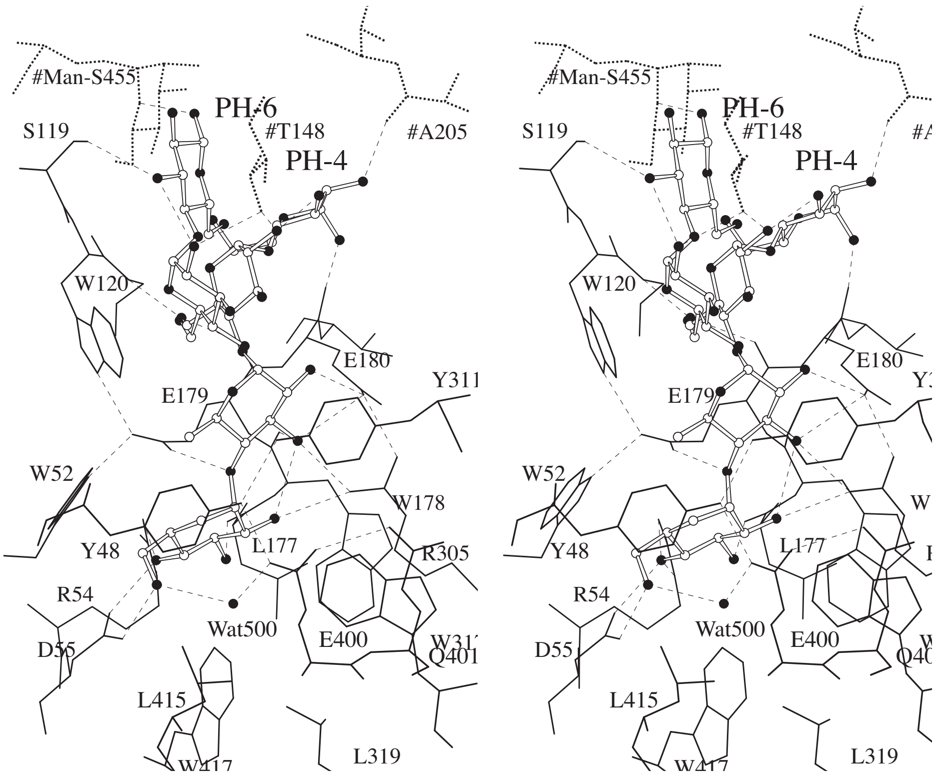


Figure 12-7 Stereoscopic view of the inhibitor D-*gluco*-dihydroacarbose in its complex with glucoamylase of *Aspergillus*. Residues from symmetry-related molecules of the enzyme are shown as dotted lines. Ring A at the nonreducing end of an amylase chain is thought to bind in a similar way (with ring A at the bottom in this figure). Cleavage is between the A and B rings E178 acting as a proton donor to the bridge oxygen (or NH for the inhibitor). The attacking nucleophile is thought to be a water molecule, which is labeled Wat 500 and is held by the assisting carboxylate of E400.[180] Courtesy of Alexander Aleshin.

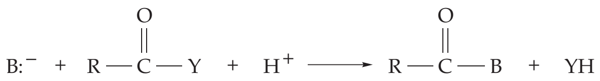
corresponding domain of cellulases.[183] It is connected by a glycosylated linker to the 470-residue catalytic subunit. Cyclodextrin glucanotransferase also has a starch-binding domain.[184]

Beta-amylases, characteristic plant and bacterial enzymes[184a], have an *exo* action, cutting off chain ends two sugar units at a time as *maltose*. The original α linkage is inverted, the product being β-maltose. The $(\alpha/\beta)_8$-barrel structure is unlike that of glucoamylase, but the spacing of active site carboxyl groups suggests that a water molecule is held and activated as in glucoamylase.[185,186] The multiple subsite structure of the active site may permit the substrate amylase to slip forward after a maltose product molecule leaves the site. This may account for the observed processive action.[186]

Many specialized glycosyl transferases synthesize glycogen, starch, cellulose, and other polysaccharides and add glycosyl groups to glycoproteins and glycolipids.[187,188] Often the glycosyl group is transferred from a carrier such as uridine diphosphate (UDP). An example is glycogen synthase (Fig. 11-4), which transfers glucosyl groups of UDP-glucose to the 4'–OH groups at the nonreducing ends of the bushlike glycogen molecules. Other similar synthetic reactions are considered in Chapter 20.

C. Displacement Reactions on Carbonyl Groups

A second major class of nucleophilic displacement reactions (Type 1B in Table 10-1) involve replacement of a group Y attached to a carbonyl carbon:

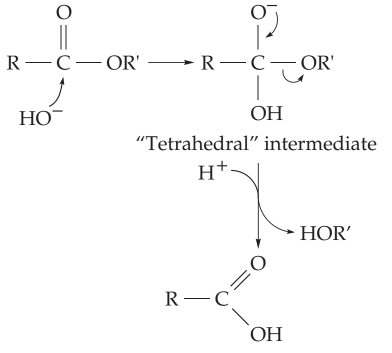


$$B:^- + R-\overset{\overset{\displaystyle O}{\|}}{C}-Y + H^+ \longrightarrow R-\overset{\overset{\displaystyle O}{\|}}{C}-B + YH$$

(12-17)

Group Y may be –OR (esters), –SR (thioesters) or –NHR (amides or peptides). If B^- is a hydroxyl ion formed from H_2O, the reaction is **hydrolysis**. If B^- is the anion of an alcohol, thiol, or amine, the reaction is **transacylation**. Transacylation is an essential process in biosynthesis of proteins and lipids, but it is the digestive enzymes, which catalyze hydrolysis, that have been studied most intensively.

Uncatalyzed hydrolysis of a peptide linkage is very slow with $t_{1/2}$ at neutral pH and 25°C of ~300–600 years.[189] Both acids and bases catalyze hydrolysis, but enzymes are needed for rapid digestion. The carbonyl group C=O is highly polarized, with the resonance form $C^+–O^-$ contributing substantially to its structure. An attack by a base will take place readily on the electrophilic carbon atom. While the reactivity of the carbonyl group in esters and amides is relatively low because of the resonance stabilization of these groups, the carbon atom still maintains an electrophilic character and may combine with basic groups. Thus, in the base-catalyzed hydrolysis of esters a hydroxyl ion adds to the carbonyl group to form a transient single-bonded "tetrahedral" intermediate (Eq. 12-18). Similar intermediates are believed to form during the action of many enzymes. However, for purposes of classification, we can regard them as simple displacement reactions on a carbon atom with the understanding that there are probably transient single-bonded intermediates.



(12-18)

A large number of hydrolytic enzymes, the **proteases**, and **peptidases** act on peptide linkages of proteins.[190] At present the traditional name protease, which implies proteolysis, is most often used. However, the IUB encourages use of the name **proteinase**. Although this seems less specific in meaning, its use will probably increase. Some proteases trim newly formed peptide chains, others convert proteins from precursor forms into biologically active molecules, and others digest proteins. In addition to endopeptidases, such as trypsin and chymotrypsin, which cleave at positions *within* a long peptide chain, there are many enzymes that cleave amino acids from the *ends* of chains. These are usually called peptidases and are designated **aminopeptidases** if they cleave from the N terminus and **carboxypeptidases** if they cleave from the C terminus. Most of these enzymes can be classified into **serine** proteases, **cysteine** proteases, **aspartic** proteases, **metallo** proteases, or **N-terminal nucleophile** hydrolases, depending upon the chemical nature of the active site. These groups are further divided into families or "clans."[190]

1. The Serine Proteases

The digestive enzymes **trypsin**, **chymotrypsin**, **elastase**, and **proteinase E** are related serine proteases. All three are synthesized in the pancreas which secretes 5–10 g per day of proteins, mostly the inactive proenzymes (zymogens) of digestive enzymes.[191,192] These proenzymes are synthesized and "packaged" as **zymogen granules** which travel to the surfaces of the secretory cells. The contents of the granules are secreted into the extracellular medium and are discharged via the pancreatic duct into the small intestine. At their sites of action, the zymogens are converted into active enzymes by the cutting out of one or more pieces from the precursor. This occurs in a cascade-type process triggered by **enteropeptidase** (historically enterokinase), another serine protease which is secreted by the intestinal lining.[192a] Human enteropeptidase consists of a 235-residue catalytic subunit bonded through a disulfide bridge to a larger 784-residue membrane-anchoring subunit.[193–195] It attacks specifically **trypsinogen**, converting it to active trypsin.[196,197] Trypsin in turn activates the other zymogens, as is indicated in Fig. 12-8. Trypsin can also activate its own zymogen, trypsinogen, in an autocatalytic process.

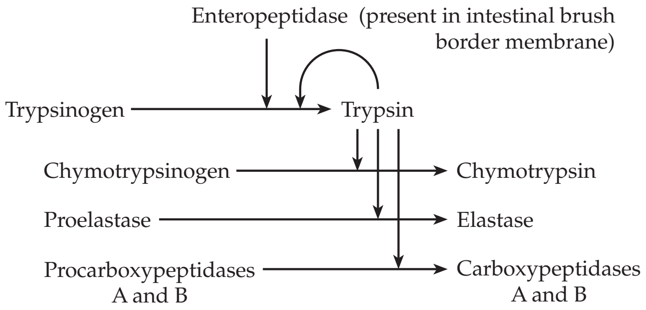
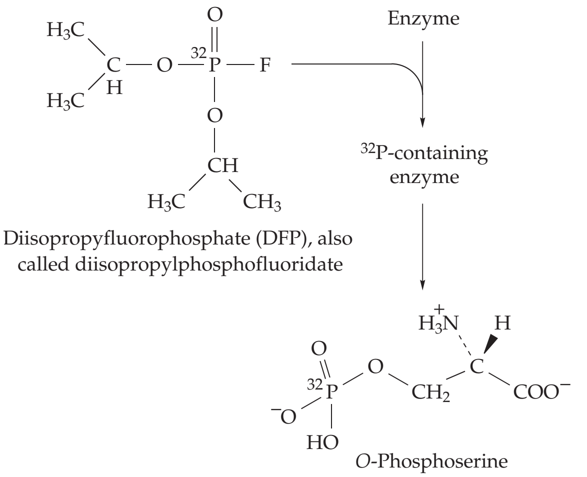

Enteropeptidase (present in intestinal brush border membrane)
Trypsinogen
Trypsin
Chymotrypsinogen
Chymotrypsin
Proelastase
Elastase
Procarboxypeptidases A and B
Carboxypeptidases A and B

Figure 12-8 Cascade of reactions that activate pancreatic proteases. Enteropeptidase, or trypsin, cleaves the proenzyme (zymogen) at specific sites.

Chymotrypsinogen consists of a single 245-residue chain. The amino acid residues in chymotrypsin, trypsin, and elastase are usually all numbered according to their position in this zymogen. Inactive proenzymes are formed as precursors to enzymes of many different classes and are activated in a variety of ways. A part of the polypeptide chain of the proenzymes is often folded over the active site, interacting in a nonsubstrate-like fashion and blocking the site.[197a]

Serine as a nucleophile. An early clue to the mechanism of action of chymotrypsin came from investigation of the related **acetylcholinesterase**. This key enzyme of the nervous system is inactivated irreversibly by powerful phosphorus-containing poisons that had been developed as insecticides and as war gases (nerve gases, Box 12-C). Around 1949, the nerve gas diisopropylfluorophosphate (DFP) was shown also to inactivate chymotrypsin. When radioactive ^{32}P-containing DFP was allowed to react the ^{32}P became covalently attached to the enzyme. When the labeled enzyme was denatured and subjected to acid hydrolysis the phosphorus stuck tightly; the radioactive fragment was identified as *O*-phosphoserine. It was evident that this product could be formed by an attack of the hydroxyl group of the serine side chain on the phosphorus with displacment of the fluoride ion. This is a nucleophilic displacement on phosphorus, occurring on an enzyme that normally catalyzes displacement on C=O. The DFP molecule acts as a pseudosubstrate which reacts with the enzyme in a manner analogous to that of a true substrate but which does not complete the reaction sequence normally.



Diisopropylfluorophosphate (DFP), also called diisopropylphosphofluoridate

O-Phosphoserine

(12-19)

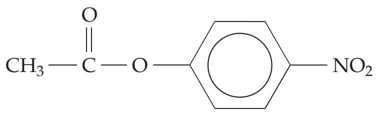
From study of peptides formed by partial hydrolysis of the ^{32}P-labeled chymotrypsin, the sequence of amino acids surrounding the reactive serine was established and serine 195 was identified as the residue whose side chain hydroxyl group became phosphorylated. The same sequence Gly-Asp-Ser-Gly was soon discovered around reactive serine residues in trypsin, thrombin, elastase, and in the trypsin-like **cocoonase** used by silkmoths to escape from their cocoons.[198] We know now that these are only a few of the enzymes in a very large family of serine proteases, most of which have related active site sequences.[199,200] Among these are **thrombin** and other enzymes of the blood-clotting cascade (Fig. 12-17), proteases of lysosomes, and secreted proteases.

Numerous serine proteases, including trypsin-like enzymes called **tryptases**[201–204a] and chymotrypsin-like chymases,[205–207] are found within tissues in which they are stored in granules of mast cells,[208] neutrophils, lymphocytes, and cytotoxic T cells.[205] Secretory granules of mast cells present in skin and other tissues contain high concentrations of tryptase and chymase precursors[202,206] which may be released as part of an inflammatory response. Tryptase may be involved in asthma and other allergic resposes.[201] **Cathepsin G**[209] (proteinase II), **neutrophil elastase**, and **proteinase III**[210] are found in granules of neutrophils and monocytes as well as in mast cells.[209] Cytoplasmic granules of cytotoxic T cells contain at least seven proteases called **granzymes** that can be released to attack target cells.[211–213a]

Many secreted proteins, as well as smaller peptide hormones, are acted upon in the endoplasmic reticulum by tryptases and other serine proteases. They often cut between pairs of basic residues such as KK, KR, or RR.[214–216] A substilisin-like protease cleaves adjacent to methionine.[217] Other classes of proteases (e.g., zinc-dependent carboxypeptidases) also participate in this processing. **Serine carboxypeptidases** are involved in processing human prohormones.[218] Among the serine carboxypeptidases of known structure is one from wheat[219] and **carboxypeptidase Y**, a vacuolar enzyme from yeast.[220] Like the pancreatic metallocarboxypeptidases discussed in Section 4, these enzymes remove one amino acid at a time, a property that has made carboxypeptidases valuable reagents for determination of amino acid sequences. Carboxypeptidases may also be used for modification of proteins by removal of one or a few amino acids from the ends.

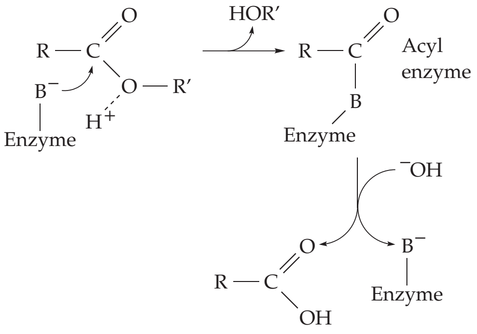
The variety of bacterial serine proteases known include the 198-residue **α-lytic protease** of *Myxobacter*,[221] a family of at least 80 **subtilisins** which are produced by various species of *Bacillus*[222–225a] as well as by many other organisms,[226] and a trypsinlike enzyme from *Streptomyces griseus*.[227,228] **Tripeptidyl peptidases**, subtilisin-like enzymes, cut tripeptides from the N-termini of proteins.[228a,b,c] One participates in lysosomal protein degradation and the other, an oligomer of 138 kDa subunits, cuts precursor proteins to form neuropeptides and other hormones (Chapter 30).

Acyl-enzyme intermediates. Serine proteases are probably the most studied of any group of enzymes.[229] Early work was focused on the digestive enzymes. The pseudosubstrate, *p*-nitrophenyl acetate, reacts with chymotrypsin at pH 4 (far below the optimum pH for hydrolysis) with rapid release of *p*-nitrophenol and formation of acetyl derivative of the enzyme.



p-Nitrophenyl acetate

This acetyl enzyme hydrolyzes very slowly at pH 4 but rapidly at higher pH. These experiments suggested a double displacement mechanism:



(12-20)

Although the experiments with DFP suggested that the $-O^-$ group from Ser 195 might be the base $-B^-$ in this equation, there was reluctance to accept this deduction because of the very weak acidity of the $-CH_2OH$ group. Furthermore, the pH dependence of catalysis suggested an imidazole group of a histidine side chain as the attacking nucleophile B. Indeed, imidazoles catalyze the nonenzymatic hydrolysis of *p*-nitrophenyl acetate with formation of unstable *N*-acetyl imidazoles as intermediates. Thus, while the stable end products of reactions with pseudosubstrates were unquestionably derivatives of serine, the possibility remained that these were side products and that histidine was involved in transient, rapidly forming, and reacting intermediates. It remained for the results of the then newly developed science of X-ray crystallography to clarify this question.[230]

Three-dimensional structures. The structures of chymotrypsin,[199,230,231] trypsin,[232,233] elastase,[234,235] thrombin,[236] kallikrein,[237,238] and many other enzymes are similar, with the basic fold shown in Fig. 12-9.[199,228] Both Ser 195 and His 57 (or corresponding residues) are present in the active site (Figs. 12-9, 12-10). From the observed positions of competitive inhibitors occupying the active site, the modes of binding depicted in Fig. 12-10A for the chymotrypsin family and in Fig. 12-10B for the subtilisin family have been deduced. Bear in mind that the X-ray diffraction results do not

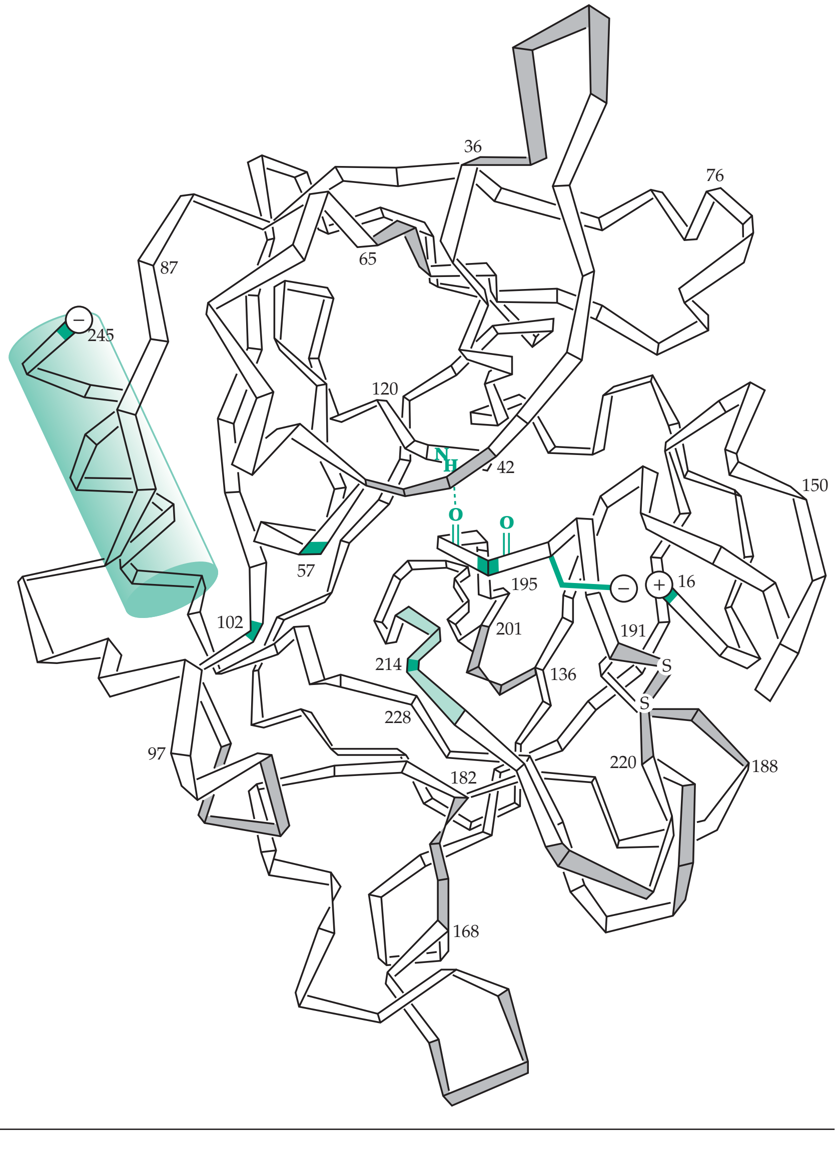


Figure 12-9 Alpha-carbon diagram of the three-dimensional structure of pancreatic elastase. A principal structural feature is a pair of β cylinders. One of these, at the top of the figure, is viewed end-on, while the other, at the bottom of the figure, is viewed from a side. The prominent interface between them is seen in the center. The α-carbon positions of the catalytic triad serine 195, histidine 57, and aspartate 102 are marked. As shown in Fig. 12-10 the catalytic triad is located across the interface between β cylinders, which may allow for easier conformational alterations during the action of the enzyme. The four disulfide bridges and the long C-terminal helix are also emphasized. The cross-hatched loop regions are residues present in elastase but not present in chymotrypsin. The green shaded strand (residues 214–216) is a segment that joins residues P_1 and P_2 of the substrate to the edge of the lower β cylinder in an extended β structure. The ion pair formed by aspartate 194 and isoleucine 16 during zymogen activation is also shown, as are two peptide carbonyl groups that protrude into the interface area, one forming a hydrogen bond across the interface. Modified from a drawing of Sawyer *et al.*[234] which was in turn based on the original chymotrypsin drawing of Annette Snazle.[230]

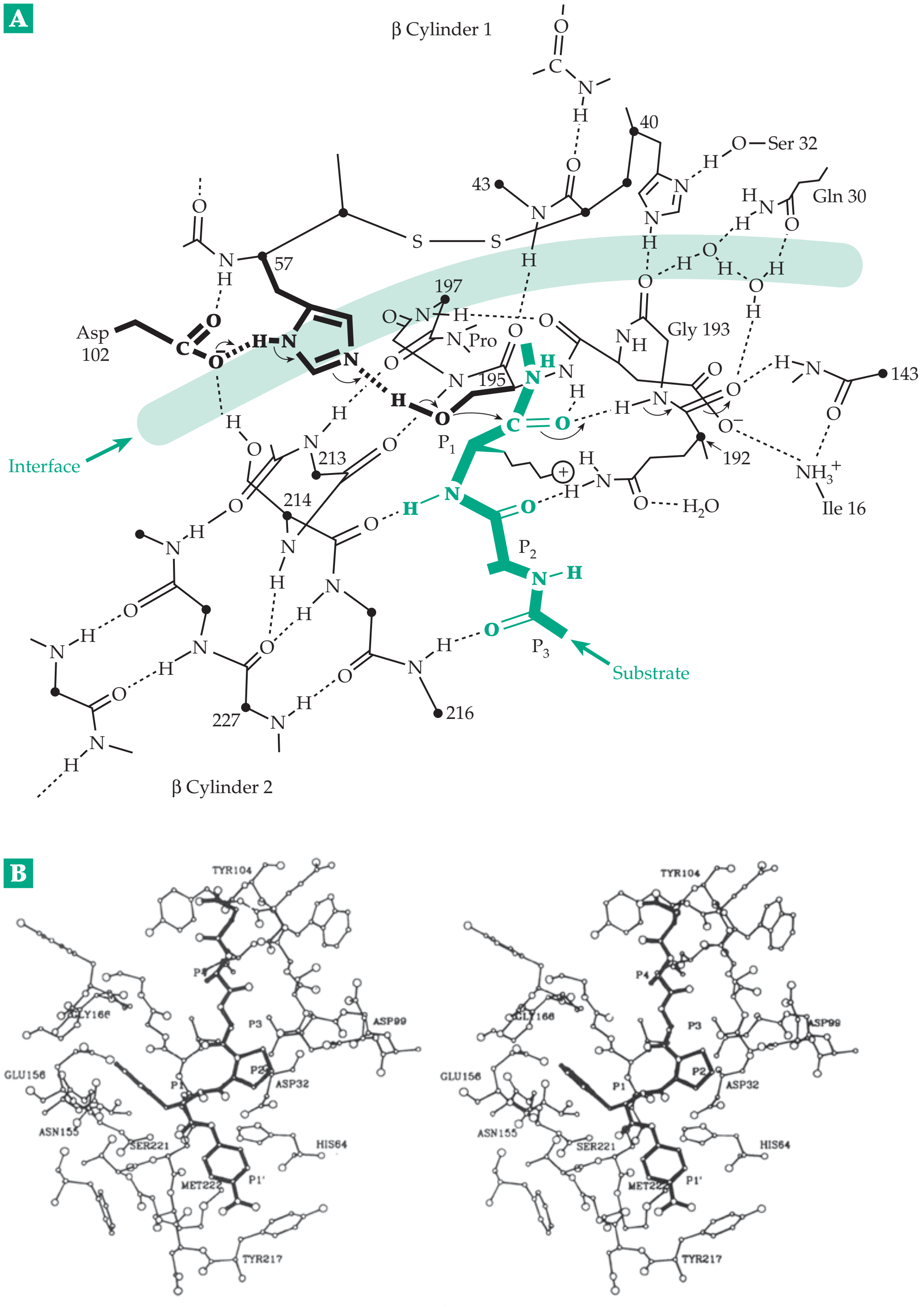
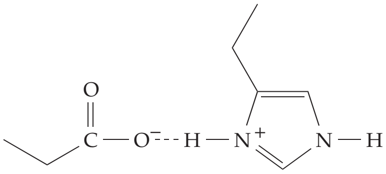


Figure 12-10 (A) Part of the hydrogen-bonding network of trypsin and other serine proteases with a bound trypsin substrate (green). Residues P_1 to P_3, as defined in Fig. 12-14, have been marked. The view is similar to that in Fig. 12-9 but some background lines have been omitted. Note the two competing hydrogen-bonded chains passing through C=O of residue 227 in β cylinder 2. One of these chains passes through the backbone of the substrate and the other through residues 214 and 195 across the interface between domains into β cylinder 2. Asp 102, His 57, and Ser 195 of the catalytic triad are emphasized. Arrows indicate probable movement of electrons from the negative charge of Asp 102 into the "oxyanion hole." After Metzler.[85] Based on papers of Sawyer *et al.*[234] and Huber and Bode.[240] (B) Stereoscopic view showing the model substrate *N*-succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide bound into the active site of subtilisin BPN′. Residues of the catalytic triad (Ser 221, His 64, Asp 32) and some others are labeled. Subsites P_1', P_1, P_2, P_3, and P_4 (see Fig. 12-14) are also labeled. Notice that site P_3, which is near the top in this drawing, is at the bottom in (A). Based on X-ray data of R. Bott and M. Ultsch at 0.2 nm resolution. From Wells and Estell.[223]

show where the hydrogen atoms are and that these have been added in Fig. 12-10. The imidazole group of His 57 is located next to the side chain –OH group of Ser 195 and is able to form an N---H–O hydrogen bond to it. The obvious conclusion is that His 57 acts as a general base catalyst that assists in removing the proton from the –OH of Ser 195, making that hydroxyl group more nucleophilic than it would be otherwise. This may happen after the –OH group has started to add to the substrate carbonyl.

The second nitrogen atom of His 57 is hydrogen bonded to the carboxylate group of Asp 102, which is in turn hydrogen bonded to two other groups. Aspartate 102 has one of the few carboxylate side chains that is buried inside the protein. To Blow,[230,231,239] the structure suggested a **charge-relay system** by which negative charges might move synchronously from Asp 102 to the imidazole which could then deprotonate the hydroxyl group of Ser 195, allowing the serine oxygen to add to the substrate carbonyl to form the tetrahedral or **oxyanion** intermediate, which is depicted in step *b* of Fig. 12-11. The small arrows in Fig. 12-10 also indicate the movement of charge. Blow suggested that in the extreme case a *proton* might be transferred also from the His 57 imidazole to the Asp 102 carboxylate. However, a variety of experiments, including studies by ^{15}N NMR[241,242] and by neutron diffraction,[243] suggest that the imidazole does not transfer its proton to the carboxylate of Asp 102,[244,245] unless it does so transiently, e.g., in a transition state complex. Instead, the carboxylate and the imidazolium ions formed by protonation of His 57 probably exist as a tight ion pair:



Support for this concept is provided by ^1H NMR studies which have identified a downfield resonance of the hydrogen-bonded proton in this pair at ~18 ppm in chymotrypsinogen and chymotrypsin at low pH and at ~14.9–15.5 ppm at high pH values.[246,247] Similar resonances are seen in the α-lytic protease,[248] in subtilisin,[249] in adducts of serine proteases with boronic acids[250,251] or peptidyl trifluoromethyl ketones,[252] in alkylated derivative of the active site histidine,[253] and in molecular complexes that mimic the Asp-His pair in the active sites of serine proteases.[254]

The catalytic cycle. Figure 12-11 depicts the generally accepted sequence of reactions for a serine protease. If we consider both the formation and the subsequent hydrolysis of the acyl-enzyme intermediate with appropriate oxyanion intermediates, there are at least seven distinct steps. As indicated in this figure, His 57 not only accepts a proton from the hydroxyl

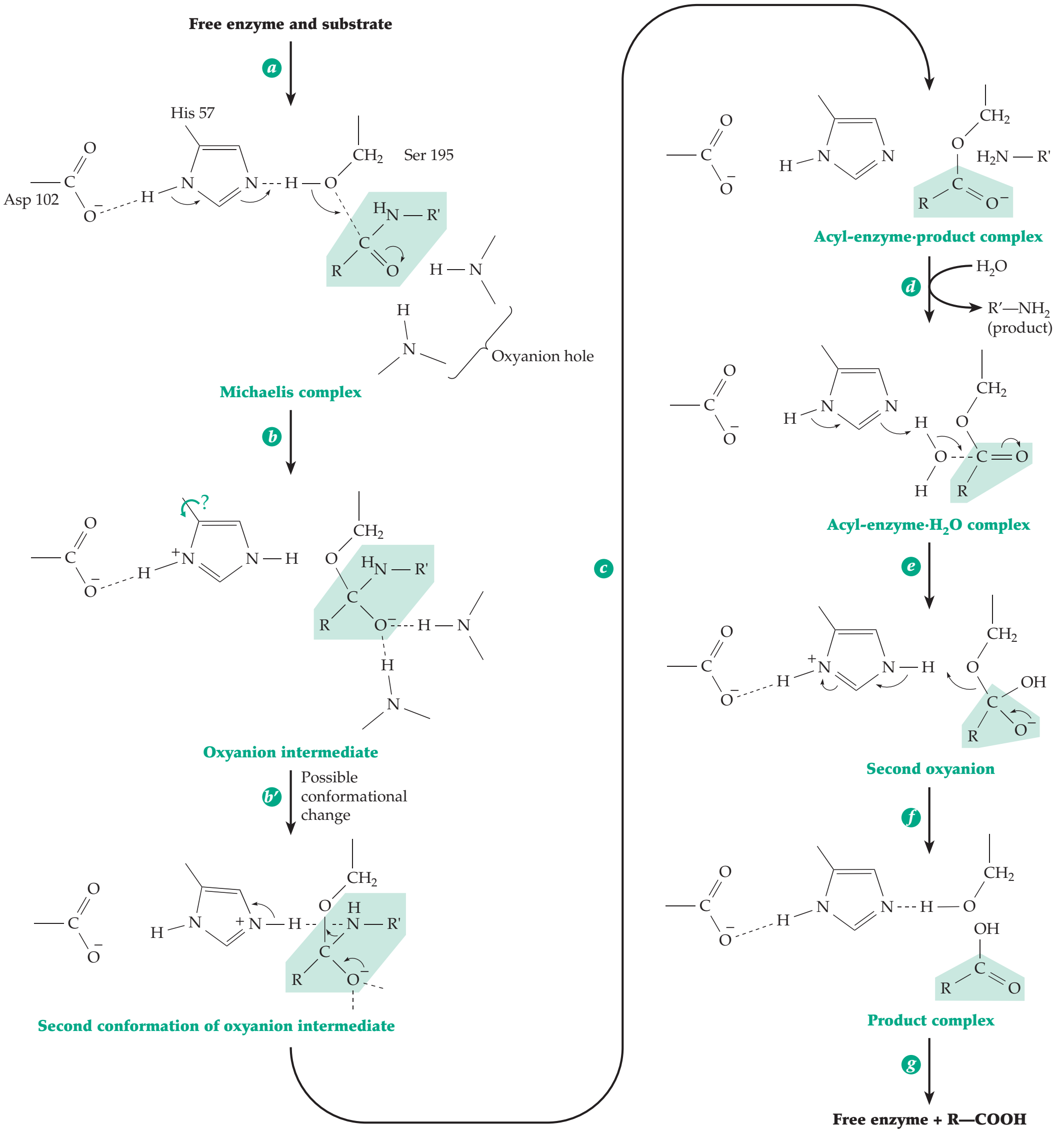

Free enzyme and substrate
a
His 57
Ser 195
Asp 102
Oxyanion hole
Michaelis complex
b
Oxyanion intermediate
b'
Possible conformational change
Second conformation of oxyanion intermediate
c
Acyl-enzyme·product complex
d
H2O
R'—NH2 (product)
Acyl-enzyme·H2O complex
e
Second oxyanion
f
Product complex
g
Free enzyme + R—COOH

Figure 12-11 Sequence of chemical reactions involved in the action of a serine protease. The oxyanion hole structure has been omitted in the right-hand column. The imidazole ring may rotate, as indicated by the green arrow, to provide two different conformations.

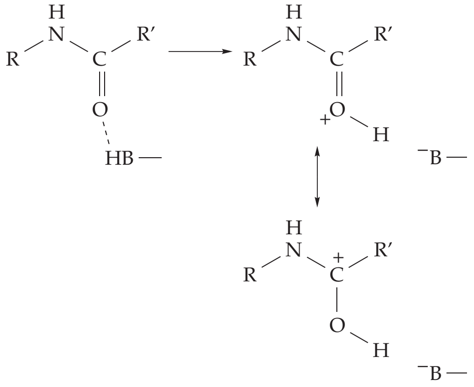
group of Ser 195 (step *b*) but probably also functions in protonation of the –NH–R′ leaving group (step *c*). An unprotonated –NH–R′ would be such a poor leaving group that the oxyanion intermediate would not go on to acyl-enzyme. However, donation of a proton to that leaving group from the protonated His 57 (general acid catalysis) permits elimination of H_2N-R' (step *c*). The product of this step is the acyl-enzyme intermediate which must be hydrolyzed to complete the catalytic cycle. This is accomplished through steps *d–f* of Fig. 12-11. The product $R'-NH_2$ is replaced by H_2O, which, in steps paralleling steps *b* and *c*, is converted to an HO^- ion that serves as the attacking nucleophile to form (in step *e*) a second oxyanion intermediate which is cleaved to the second product R–COOH. The water molecule that enters in step *d*, and which participates in hydrolysis of the acyl enzyme, has apparently been observed directly by time-resolved Laue crystallography at low temperature.[255]

The catalytic triad. The significance of the charge-relay effect may not be fully understood but the importance of the Ser•His•Asp cluster, which has become known as the catalytic triad, cannot be doubted. It has evolved independently in several subfamilies of bacterial and plant proteases, a well-known example of "convergent evolution." This triad is also found throughout a broad range of many different kinds of enzymes and other proteins. In pancreatic enzymes the catalytic triad consists of Ser 195•His 57•Asp 102 but in the bacterial subtilisin it consists of Ser 221•His 64•Asp 32 and in a wheat serine carboxypeptidase it is Ser 146•His 397•Asp 338.[219] In all three cases the folding patterns of the polypeptides are entirely different but the geometry of the triad is the same. Another folding pattern is seen in a protease encoded by cytomegalovirus, which contains an active site Ser 132•His 63 pair.[256,257]

Investigation of a host of mutant proteins also demonstrates the importance of the catalytic triad. For example, if either the histidine or the serine of the triad of subtilisin was replaced by alanine the catalytic activity decreased by a factor of 2×10^6 and replacement of the aspartate of the triad by alanine decreased activity by a factor of 3×10^4.[229,258] When Asp 102 of trypsin is replaced by asparagine the catalytic activity falls by four orders of magnitude.[259] This may be in part because the histidine in this mutant is hydrogen bonded to Asn 102 as the tautomer with a proton on N^{ε}, the nitrogen that should serve as the catalytic base in step *b* (Fig. 12-11).[260] A mutant in which Ser 214 (see Fig. 12-10) was replaced with alanine is fully active but charged residues in this position interfere with catalysis.[261]

Does the "low-barrier hydrogen bond" in the catalytic triad play any special role in catalysis? Blow's suggestion of a charge relay from Asp 102 to Ser 195 of chymotrypsin is probably correct. Some theoretical calculations have indicated the possibility of synchronous movement of the two protons in the system during step *a* of the sequence shown in Fig. 12-11, with the proton in the strong hydrogen bond to Asp 102 moving away from His 57 and toward the midpoint distance of the hydrogen bond. Could the presumed high energy of the short hydrogen bond be harnessed to lower the transition state energy? A realistic possibility is that the hydrogen bond increases the polarizability of the catalytic triad, facilitating movement of the substrate to the transition state. Cassidy *et al.* suggested that the strong hydrogen bond may be formed by compression of the triad resulting from binding of substrate in the S_1 and S_2 subsites.[252] They suggested that this would raise the pK_a of His 57 from ~7 in the free enzyme to 10–12, high enough to enable it to remove the proton from the Ser 195 –OH group and low enough to allow the protonated form to be the proton donor to the leaving group (step *c*, Fig. 12-11).

The "oxyanion hole." A third mechanism by which an enzyme can assist in a displacement reaction on a carbonyl group is through protonation of the carbonyl oxygen atom by an acidic group of the enzyme (Eq. 12-21). This will greatly increase the positive charge on the carbon atom making attack by a nucleophile easier and will also stabilize the tetrahedral intermediate. Although the carbonyl oxygen is very weakly basic, it can interact with a suitably oriented acidic group of the enzyme. In many serine proteases this acidic function is apparently fulfilled by NH groups of two amide linkages. In chymotrypsin these are the backbone NH groups of Ser 195 and Gly 193 (Figs. 12-10,12-12). No actual transfer of a proton to the carbonyl oxygen of the substrate is expected. However, the NH groups are positive ends of amide dipoles and can interact electrostatically with the negative charge that develops on the oxyanion. The fit of substrate into the **oxyanion hole** between the two



(12-21)

NH groups is apparently good only for the tetrahedrally bonded oxyanion intermediate,[262] a structure thought to be close to that of the transition state. The importance of the oxyanion hole for catalysis has been supported by theoretical calculations[245] and by the study of mutant enzymes.[263] In subtilisin, in which a side chain of asparagine forms part of the oxyanion hole, replacement of Asn with the isosteric Leu causes k_{cat} to fall by a factor of about 200 while K_m is unaffected.[264] It is also of interest that thiono ester substrates, in which the C=O of an oxygen ester has been replaced by C=S, bind with normal affinity to chymotrypsin but are not hydrolyzed at significant rates.[265]

The formation of the oxyanion intermediate during serine protease action is also supported by the existence of tetrahedral forms of enzymes inhibited by substrate-like aldehydes. The –OH group of Ser 195 can add to the carbonyl group to form hemiacetals. For example, a ^{13}C-enriched aldehyde whose carbonyl carbon had a chemical shift of 204 ppm gave a 94 ppm resonance as it formed the tetrahedral hemiacetal with one of the inhibitory aldehydes, *N*-acetyl-L-Leu-L-Leu-L-arginal (**leupeptin**; Box 12-C). A natural product from a species of *Streptomyces*, this aldehyde inhibits trypsin and several other enzymes strongly.[266] Adducts of wheat serine carboxypeptidases with aldehyde inhibitors have also been observed.[267] While the carbonyl group of the substrate amide linkage that is to be cleaved apparently can't form strong hydrogen bonds to the NH groups of the oxyanion hole, that of the acyl-enzyme intermediate can, as judged by resonance Raman spectroscopy.[268,269] The strength of the hydrogen bonding, as judged by the stretching frequency of the C=O group, is correlated with the reactivity of the acyl group.[269]

Chymotrypsinogen and related proenzymes have extremely low catalytic activity even though a major part of the substrate binding site as well as the catalytic triad system are already in place. However, the oxyanion hole is created during activation of the proenzyme by a subtle conformational change[197,262,271] that involves the chain segment containing Gly 193 (Fig. 12-12). This is further evidence of the importance of this part of the active site structure.

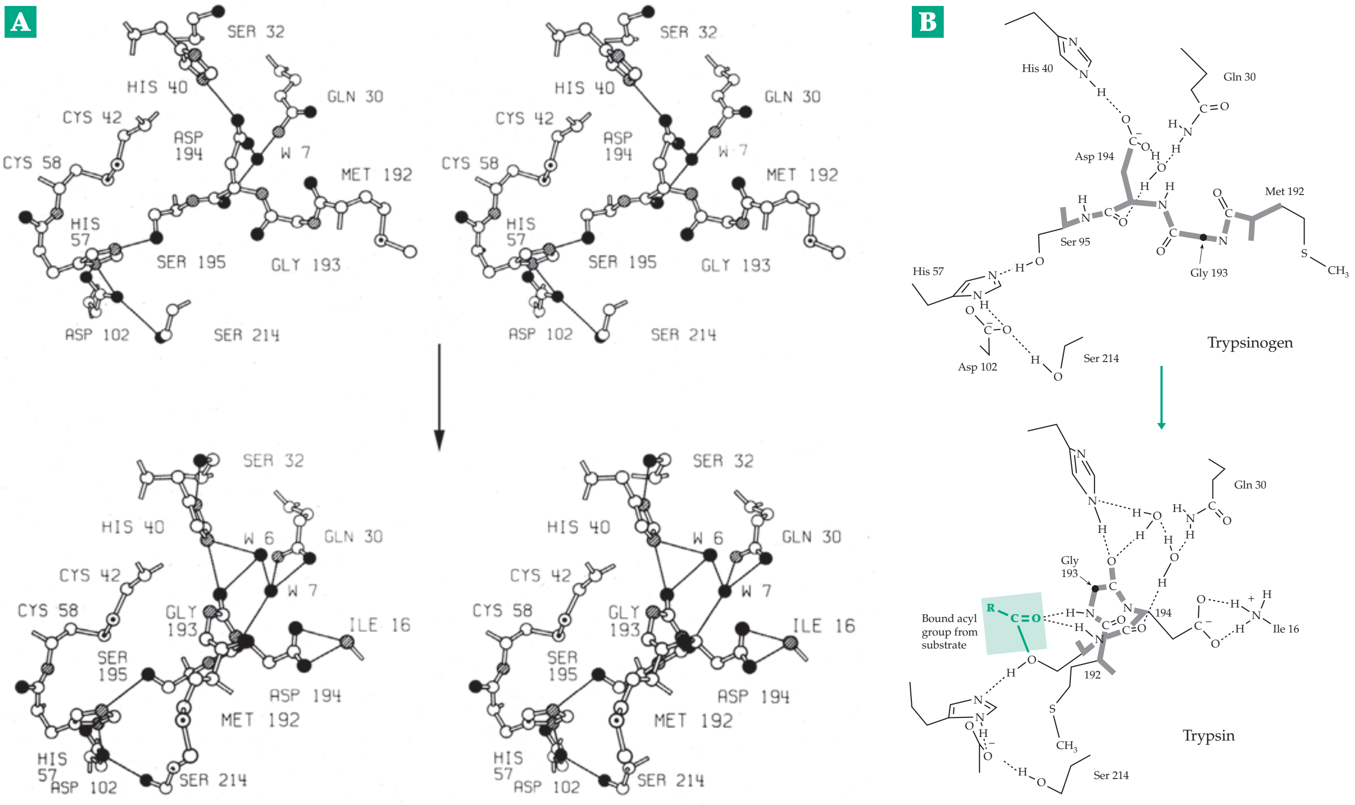

A
SER 32
HIS 40
GLN 30
CYS 42
ASP 194
W 7
CYS 58
MET 192
HIS 57
SER 195
GLY 193
ASP 102
SER 214
SER 32
HIS 40
W 6
GLN 30
CYS 42
W 7
CYS 58
GLY 193
ILE 16
SER 195
ASP 194
MET 192
HIS 57
ASP 102
SER 214
B
His 40
Gln 30
Asp 194
Met 192
Ser 95
Gly 193
His 57
Asp 102
Ser 214
Trypsinogen
Gln 30
Gly 193
Bound acyl group from substrate
R
C=O
194
Ile 16
192
Trypsin
Ser 214

Figure 12-12 Formation of the oxyanion hole following cleavage of trypsinogen between Lys 15 and Ile 16. (A) Stereoscopic view. (B) Schematic representation. The newly created terminal –NH_3^+ of Ile 16 forms a hydrogen-bonded ion pair with the carboxylate of Asp 194. This breaks the hydrogen bond between Asp 194 and His 40 in trypsinogen, inducing the peptide segment 192–194 to shift from an extended conformation to a helical form in which the NH groups of Gly 193 and Ser 195 form the oxyanion hole. Notice that the positions and interactions of Asp 102, His 57, and Ser 195, the catalytic triad, are very little changed. From Birktoft *et al.*[270]

Stereoelectronic considerations. The amide group that is cleaved by a protease is a resonance hybrid of structures A and B of Fig. 12-13. The unshared pair of electrons on the nitrogen atom of structure A and the third unshared pair on the oxygen atoms of structure B (shaded orbitals) have been drawn in this figure in such a way that they are anti-periplanar to the entering serine oxygen. This is required in the transition state according to stereoelectronic theory. The newly created (shaded) electron pair on oxygen is one of those that hydrogen bonds to an NH group of the oxyanion hole. The tetrahedral intermediate has another unshared pair, which is not hydrogen bonded, and is antiperiplanar to the HN-R leaving group. Thus, it appears to be set up for easy elimination. Nevertheless, there is some doubt about the need for adherence to the stereoelectronic "rule" that two antiperiplanar lone pairs are necessary for elimination.

There is another complication. The leaving group –NH–R′ cannot be eliminated from the oxyanion until it is protonated, presumably by the imidazolium group of His 57. However, the proton on His 57 will be adjacent to the proton that is already on this nitrogen rather than to the unshared pair of electrons on the same nitrogen atom. A conformational change in which the tilt of the catalytic imidazole ring is altered (step *b*′, Fig. 12-11) or in which the ring rotates, may have to precede the protonation of the leaving group (step *c*, Fig. 12-11).[272–274b] This change may be assisted by the presence of positive charge on the protonated imidazole, but it still does not solve the problem.

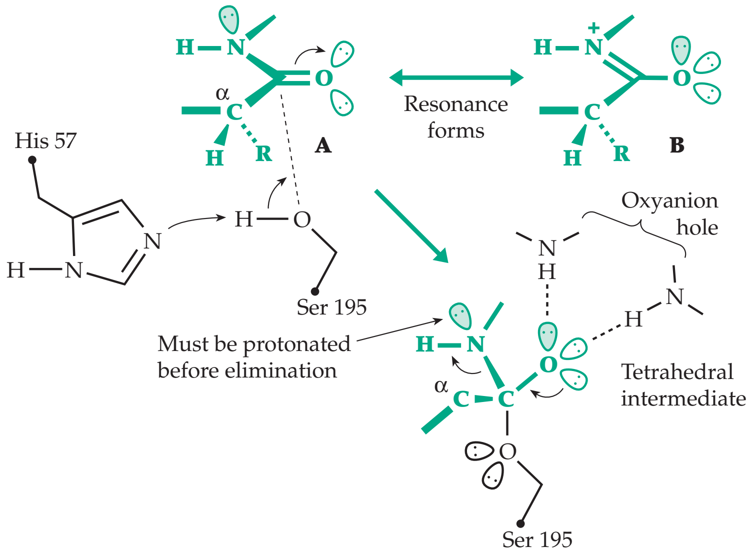


Figure 12-13 The stereochemistry of formation of a tetrahedral intermediate by a serine protease. The most probable orientation of groups as deduced by model building is shown. The shaded orbitals in A and B are antiperiplanar to the entering oxygen of Ser 195. See Polgár and Halász.[272]

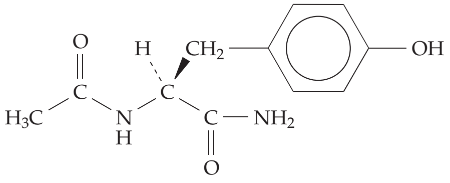
Bizzozero and Butler suggested that a rapid inversion of this chiral center on the –NH–R′ group may be required prior to protonation and elimination.[273] This would explain their observation that *N*-alkylated peptide linkages of otherwise fast substrates are not cleaved by chymotrypsin. A high-resolution structure determination demonstrated that a very slow elastase substrate that contains *N*-methylleucine at the cleavage site forms a normal ES complex.[275] However, inversion would be hindered by the *N*-methyl group. An alternative in inversion would be a torsional rearrangement of the substrate during binding.[276] The structure of an acyl-enzyme intermediate with elastase has been determined by crystallographic cryoenzymology. The intermediate was allowed to accumulate at –26°C, after which the temperature was lowered to –50°C and the structure determined. The structure shows the carbonyl group in the oxyanion hole as anticipated.[277]

pH dependence. A plot of k_{cat} / K_m for chymotrypsin is bell shaped with a maximum around pH 7.8 and pK_a values of 6.8 and 8.8. These represent pK_as of the free enzyme (Eq. 9-57). That of 6.8 has been shown to represent His 57. As the pH is lowered from the optimum the affinity for substrate falls off as His 57 becomes protonated. The high pK_a of 8.8 is thought to belong to the N-terminal amino group of Ile 16, a group that is generated during the conversion of the proenzyme to active enzyme.[278] The Ile 16 amino group forms an ion pair with Asp 194 (Fig. 12-12B) which is next to the serine at the active center. This salt linkage helps to hold the enzyme in the required conformation for reaction and its deprotonation at high pH causes a decrease in substrate affinity.[279]

The value of k_{cat} is also pH dependent and falls off at low pH around a pK_a value of from 6–7 depending upon the substrate. However, no higher pK_a affects k_{cat} in the experimentally accessible range. These results provide the basis for believing that an unprotonated His 57 is needed in the ES complex for catalysis to occur. Similar conclusions have been reached for other serine proteases.

Substrate specificity. Like most other enzymes, proteases display distinct preferences for certain substrates. These are often discussed using the nomenclature of Fig. 12-14. The substrate residue contributing the carbonyl of the amide group to be cleaved is designated P_1 and residues toward the N terminus as P_2, P_3, etc., as is shown in Fig. 12-14. Residues toward the C terminus from the peptide linkage to be cleaved are designated P_1', P_2', etc. Chymotrypsin acts most rapidly if the P_1 residue is one of the aromatic amino acids.[280] Thus, the S_1 part of the substrate binding site must bind preferentially to large, flat aromatic rings. The crystal structure of chymotrypsin showed this site to be composed of nonpolar side chain groups. On the other hand, in trypsin the specificity portion of the S_1 site is a deep "specificity pocket" containing a fixed negative charge provided by the carboxylate side chain group of Asp 189. This explains why trypsin acts only upon peptide linkages containing the positively charged arginine or lysine residues in the P_1 position. In elastase the specificity pocket is partly filled by nonpolar side chains and the enzyme can accommodate only small P_1 side chains such as the methyl group of alanine. Replacement of Asp 189 of trypsin by lysine led to the predicted loss of specificity for basic side chains. However, the mutant enzyme did not become specific for negatively charged side chains.[281] This is presumably because the lysine $-NH_3^+$ group was not located in the same position as the Asp 189 carboxylate as a result of a different packing of side chains between native and mutant enzymes. Residues P_1' and P_2' also have major effects on substrate binding by serine proteases.[282–284] Now many mutant forms of subtilisin[225,285] and other serine proteases are being made and are yielding a more sophisticated understanding of the basis of the specificity of these enzymes. An important factor that has emerged is flexibility of surface loops in allowing an enzyme to adjust its structure to give a better fit to some substrates.[286,287] The specificity of the serine proteases is also being exploited in the design of specific inhibitors (Box 12-D).

Many serine proteases react with *p*-nitrophenylacetate to give acetyl enzymes. However, its rate of hydrolysis to give acetate is orders of magnitude slower than that of acyl-enzymes derived from small substrates such as the chymotrypsin substrate *N*-acetyltyrosine amide.



N-Acetyltyrosine amide

In addition to the specificity-determining P_1 aromatic side chain, the amide groups of this substrate can form specific hydrogen bonds to the protein (Fig. 12-10). These hydrogen bonds presumably help the enzyme to recognize the compound, which is bound with K_m of ~0.03 M and is hydrolyzed (with liberation of NH_3)[288,289] with k_{cat} ~0.17 s^{-1}.

What happens when the length of the substrate is extended in the direction of the N terminus? The tyrosyl residue in the foregoing compound may be designated P_1. For extended substrates which contain P_2, P_3, and additional residues (this includes most natural substrates) the K_m values decrease very little from that of short substrates despite the larger number of "subsites" to which the substrate is bound. However, the maximum velocity is often much greater for the extended substrates than for short ones. Thus, for *N*-acetyltyrosyl-glycine amide K_m is 0.017 M, only a little less than for *N*-acetyltyrosine amide, but k_{cat} is 7.5 s^{-1}, 440 times greater than for the shorter substrate.[229,288,289] Other examples have been tabulated by Fersht.[279]

These observations suggest that *the binding energy that would be expected to increase the tightness of binding is, instead, causing an increase in* V_{max}[279,290] that is, it is reducing the Gibbs energy of activation. How can this be? Imagine that as the extended substrate binds, for example, into the subsite S_2 (which binds residue P_2), it must compress a spring in the enzyme. Could not the compressed spring now provide a source of energy for assisting in peptide bond cleavage? If this is the case, we must ask "what are the springs?" Are there amide linkages of the peptide backbone that are distorted when the substrate hydrogen bonds into the site? How is the distortion transmitted to the active site and how does it stabilize the transition

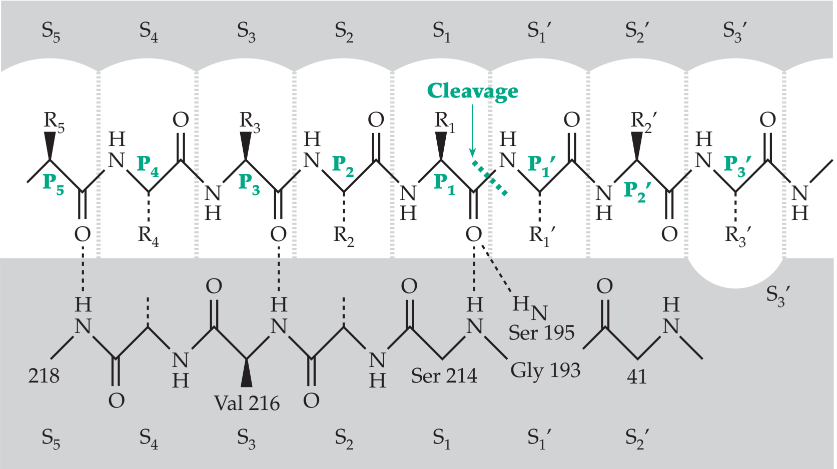


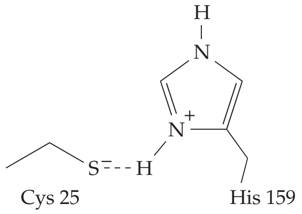
Figure 12-14 Standard nomenclature used to define the residues P_1, P_2 . . . toward the N terminus and P_1', P_2' . . . toward the C terminus of a peptide substrate for a protease. The corresponding subsites of the protease are designated S_1, S_2 . . . S_1', S_2'.

state? We are far from understanding the answers to these questions. However, it is of interest that one chain of H-bonds that passes through the amide between P_1 and P_2 also passes through a carbonyl group of a β bulge (Fig. 12-10). A second chain of H-bonds that connects to the same β bulge passes through the backbone amide of the active site Ser 195, across the interface between the two domains, and through the backbone of the second β cylinder.[85] Perhaps when substrate binds, the strengthening of the first H-bond chain weakens the second through competition at the β bulge, allowing some subtle rearrangement in the active site structure. We are talking here about the P_1–P_2 amide. What happens at P_2–P_3, and further along the chain in both directions? It is important to understand these effects; the phenomenon of increased reaction rate for longer, more specific substrates is observed with many proteases and other enzymes as well.

Another difference between small substrates and longer, more specific substrates has been found in studies of the effects of changing the solvent from pure H_2O to mixtures containing an increasing mole fraction of ^2H_2O. This is called a **proton inventory**.[291] For the serine proteases the rates are decreased as the ^2H content of the solvent increases, a fact that suggests that some step involving a proton transfer, for example, the deprotonation of Ser 195, is rate limiting. For simple substrates, the effect is directly proportional to the mole fraction of ^2H. However, for extended substrates, a quadratic dependence on the mole fraction of ^2H is observed. This suggests that a process involving synchronous transfer of two protons, as in the postulated charge relay system, may be more important in extended substrates than in simple ones and may account for the more rapid action on these substrates.[290,292] However, other interpretations of the data are possible, leaving this conclusion uncertain.[229,291]

2. The Cysteine Proteases (Thiol Proteases)

Papain from the papaya is one of a family of enzymes that includes **bromelain** of the pineapple, **ficin** of the fig, and **actinidin** of the kiwifruit.[293] Additional cysteine proteases[294] from the latex of the papaya tree *Carica papaya* are known as **caricain**, **chymopapain**,[294] and **glycyl endopeptidase**.[295] All are members of a large superfamily which includes at least 12 mammalian enzymes and many others from both eukaryotic and prokaryotic organisms.[296] All share with papain a characteristic structure which was determined by Drenth and coworkers in 1968.[297] The participating nucleophile in the active sites of these enzymes is an SH group, that of Cys 25 in papain. An adjacent imidazole from His 159 removes the proton from the SH group to form a thiolate-imidazolium pair.[298–300]



The close proximity of the imidazolium group greatly lowers the microscopic pK_a of the Cys 25 thiol group and the proximity of the resulting hydrogen-bonded S^- group greatly raises the microscopic pK_a of the imidazolium group.

Studies of the pH dependence of V_{max}/k_{cat} (Eq. 9-57) reveal a bell-shaped dependence on pH with pK_a values[301,302] of ~4 and ~9. However, the ion pair is formed at a pH below four with apparent pK_a values of 2.5, 2.9, and 3.3 for ficin, caricain, and papain, respectively.[302] These low values can be assigned principally to Cys 25 with only very small contributions from His 159 (see Eq. 6-75). A third pK_a, perhaps of a nearby carboxylate from Glu 50, affects the rate. For caricain the nearby Asp 158 (Fig. 12-15) has been implicated.[303,304]

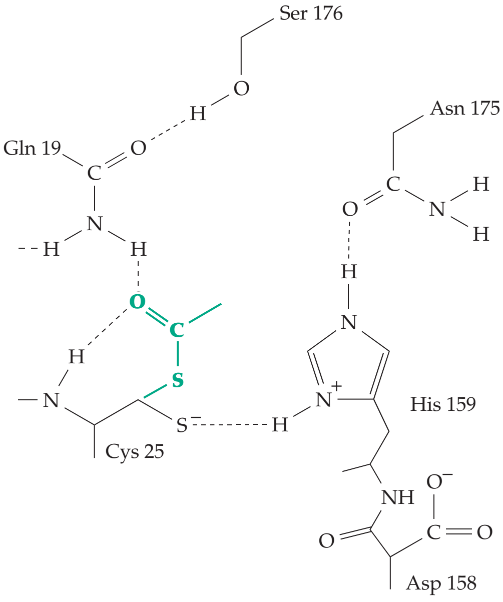


Figure 12-15 Schematic drawing of the active site of a cysteine protease of the papain family with a partial structure of an acyl-enzyme intermediate in green. The thiolate-imidazolium pair of Cys 25•His 159 lies deep in the substrate-binding cleft and bridges an interface between two major structural domains, just as the Ser•His pair does in serine proteases (Fig. 12-10). This may facilitate small conformational changes during the catalytic cycle. Asn 175 provides a polarizable acceptor for positive charge, helping to stabilize the preformed ion pair, and allows easy transfer of an imidazolium proton to the product of substrate cleavage. The peptide NH of Cys 25 and the side chain of Gln 19 form an oxyanion hole.

As shown in Fig. 12-15, the side chain of Asn 175 provides papain with a third member of a catalytic triad analogous to that of serine proteases.[305] Glutamine 19, together with the peptide backbone NH of Cys 25, provides an oxyanion hole.[306–308] Many studies, including structure determinations on bound aldehyde and other inhibitors, and observation by ^{13}C NMR[309] indicate that thiol proteases act by addition of the thiolate anion to the peptide carbonyl of the P_1 residue, just as in step *b* of Fig. 12-11 (see also Fig. 12-12) for serine proteases.[308,310] However, alternative sequences of proton transfer have been suggested.[311] A possible role for a strong hydrogen bond has also been proposed.[310] For a detailed discussion see Brocklehurst *et al.*[312]

Most of the lysosomal proteases called **cathepsins** are small 20- to 40-kDa glycoproteins found in all animal tissues.[313] Most are cysteine proteases which function best and are most stable in the low pH reducing environment of lysosomes. They resemble papain in size, amino acid sequence, and active site structures. Papain is nonspecific but most cathepsins have definite substrate preferences. Cathepsin B is the most abundant. There are smaller amounts of related cathepsins H (an aminopeptidase)[314] and L[315] and still less of cathepsins C, K, and others. Cathepsin B is both an endopeptidase and an exopeptidase.[316] It acts on peptides with arginine at either P_1 or P_2 but also accepts bulky hydrophobic residues in P_1 and prefers tyrosine at P_3.[317] Cathepsin S is less stable at higher pH than other cathepsins and has a more limited tissue distribution, being especially active in the immune system.[318,319]

Cathepsin K is especially abundant in the bone resorbing osteoclasts (Chapter 8). It is essential to normal bone structure and its absence is associated with the rare hereditary disease pycnodysostosis (pycno) which causes short stature, fragile bones, and skull and skeletal deformities.[320] It may also play a role in the very common bone condition osteoporosis. Cathepsin C is also called **dipeptidyl peptidase**. It removes N-terminal dipeptides from many intracellular proteins activating many enzymes, including some other cathepsins.[321,322] A **prohormone thiol protease** cleaves peptide chains between pairs of basic residues, e.g., in the brain peptide precursor proenkephalin (Chapter 30), and also on the N-terminal side of arginine residues.[323] **Pyroglutamate aminopeptidase** removes pyroglutamyl (5-oxoprolyl) groups from amino termini of some peptides and proteins (see Fig. 2-4).[324] Another cysteine protease cleaves **isopeptide** linkages such as those formed by transglutaminase or those involving ubiquitin (Box 10-C).[325] Another cysteine protease present in animal tissues was recognized by its ability to hydrolyze the anticancer drug bleomycin. This **bleomycin hydrolase** is a hexamer with a central channel lined with papainlike active sites as in the proteasome structure (Box 7-A)[326] The enzyme also binds to DNA.[327] Unfortunately, cancer tissues often contain high levels of the enzyme, whereas it is low in skin and lung tissues, which are damaged by bleomycin.

The Ca^{2+}-dependent neutral proteases called **calpains** are found within the cells of higher animals. The 705-residue multidomain peptide chain of a chicken calpain contains a papain-like domain as well as a calmodulin-like domain.[328] It presumably arose from fusion of the genes of these proteins. At least six calpains with similar properties are known.[329] Some have a preference for myofibrillar proteins or neurofilaments.[330] They presumably function in normal turnover of these proteins and may play a role in numerous calcium-activated cellular processes.[331–332a]

A group of **cysteinyl aspartate-specific proteases (caspases)** play an essential role in programmed cell death (apoptosis).[333 – 335] Recall that nematodes are cell-constant organisms. For maturation of *Caenorhabditis elegans* 131 programmed cell deaths must occur at specific stages of development. An essential gene for this process was identified and named ***CED-3***. Deletion of this gene completely blocked the death of these cells. *CED*-3 encodes a cysteine protease that is highly homologous to the mammalian **interleukin-1β-converting enzyme** (**ICE** or caspase 1) which cleaves the 31-kDa **pro-interleukin-1β** to form the active 175 kDa species of this cytokine (Chapter 30). At least ten caspases are known and many observations have confirmed their role in apoptosis (see also Chapter 32). In caspases 1 and 2 the side chains of Cys 285 and His 237 form the catalytic dyad and peptide NH groups of Cys 285 and Gly 238 form the oxyanion hole.[333]

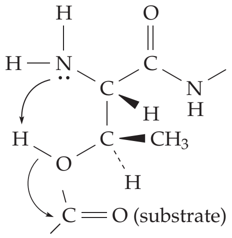
Parasites often use proteases in attacks on their hosts. The cysteine protease **cruzain** is secreted by the trypanosomes that cause Chagas' disease and is essential to their survival within the human body.[336,337] Cruzain is consequently an attractive target for development of drugs for treating this major disease.

Among cysteine proteases of bacteria is a papain-like enzyme from *Clostridium histolyticum* with a specificity similar to that of trypsin.[338] The anaerobic *Porphyromonas gingivalis*, which is implicated in periodental disease, produces both arginine- and lysine-specific cysteine proteases designated **gingipains**.[339,339a] Some virally encoded cysteine proteases, including one from the polio virus, have trypsin-like sequences with the serine of the catalytic triad replaced by cysteine.[340,341] A human adenovirus protease also has a Cys•His•Glu triad but a totally different protein fold.[342]

Zymogens of cysteine proteases usually have a long terminal extension which is removed, sometimes by autoactivation. Propapain has a 107-residue extension.[343] The 322-residue cathepsin B carries an unusually short 62-residue extension in its proenzyme form.[315,343,344] In every case the N-terminal extension folds into a domain, one of whose functions is to block the active site cleft.

3. N-terminal Nucleophile Hydrolases and Related Enzymes

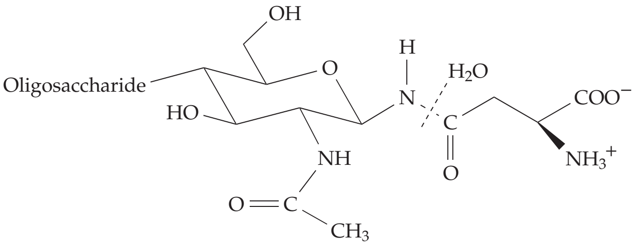
The most recently discovered group of proteases are the N-terminal threonine hydrolases of the **multicatalytic protease complex** (MPC) of proteasomes. The enzymes are arranged in a regular array inside proteasomal compartments as shown in Box 7-A. The active site is a catalytic dyad formed from the amino group at the N terminus of the β subunits.[345–346a]


H
H—N
O
C
Enzyme
C
N
H
H
H
C
CH3
O
H
C=O (substrate)

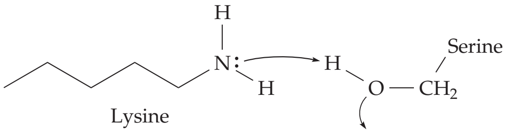
Hydroxyl • α-amine catalytic dyad

Proteasomes of *Thermoplasma* contain a single type of β subunit but eukaryotic proteasomes contain subunits with at least three distinct substrate preferences.[347–349c] They all appear to use the same hydrolytic mechanism but in their substrate specificities they are chymotrypsin-like, peptidylglutamyl-peptide hydrolyzing, branched chain amino acid preferring, and small neutral amino acid preferring based on the P_1 amino acid residue. In the spleen some of the β subunits of the proteasomes appear to have been replaced by proteins encoded by the major histocompatibility complex of the immune system (Chapter 31).[347] This may alter the properties of the proteasome to favor their function in antigen processing. Proteasomes are also ATP- and ubiquitin-dependent, as discussed in Section 6.

The enzyme glucosylasparaginase (aspartylglucosaminidase) is one of a group of other enzymes that use N-terminal threonyl groups as catalytic dyads.[346,350–353] It removes N-linked glycosyl groups from asparagine side chains of proteolytically degraded proteins and, as indicated on the accompanying structural formula, releases free aspartate and a 1-amino-*N*- acetylglucosamine-containing oligosaccharide. The amino group is then released as NH_3 from the product by acid catalysis in the lysosome.[350]

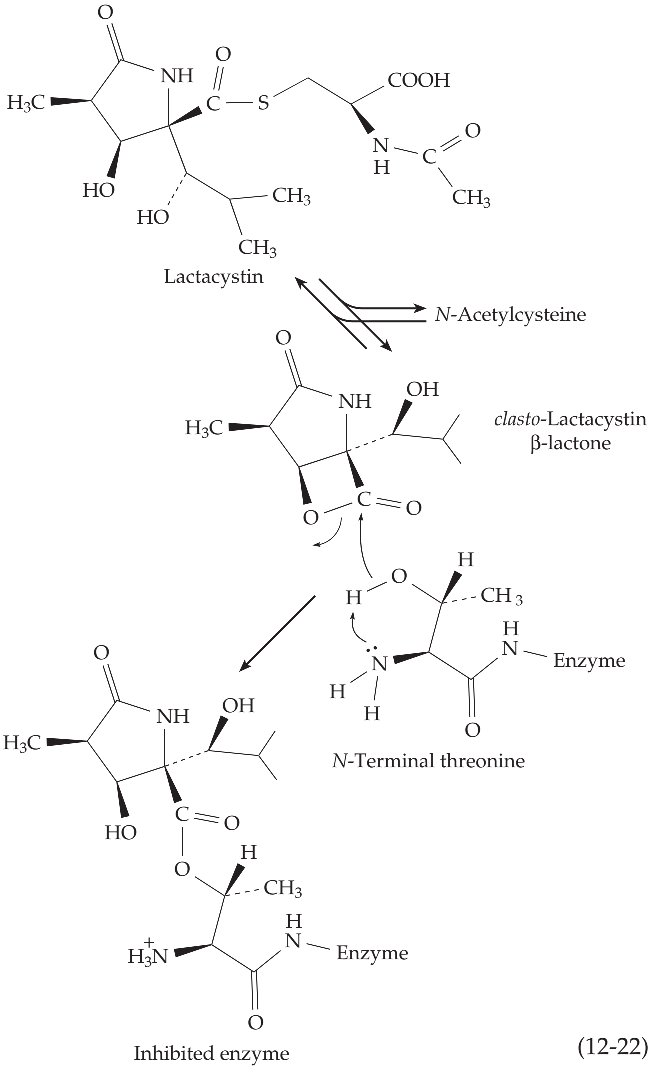

OH
Oligosaccharide
O
HO
NH
O=C
CH3
H
N
H2O
C
O
COO−
NH3+

Glutamine PRPP amidotransferase (Fig. 25-15) and a penicillin acylase have similar active sites and overall structures.[354] The $-NH_2$ group is basic enough in the environment of the protein to remove the proton from the threonine –OH group, activating it as indicated below. Several serine proteases use an $-NH_2$ group of a lysine to form a serine•lysine dyad.[346]


H
N
H
Lysine
H
O
CH2
Serine

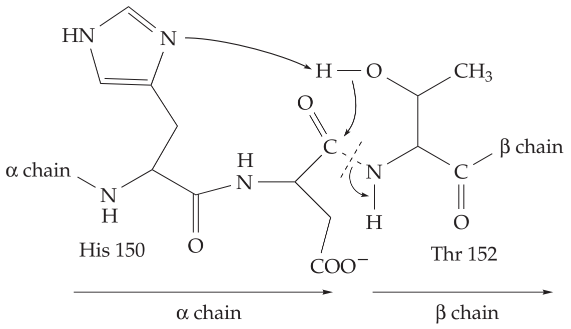
Among these are the well-known *E. coli* **leader peptidase**[355,356] and other signal peptidases.[357] These are integral membrane proteins that cleave N-terminal signal sequences from proteins incorporated into plasma membranes. Another enzyme of this class is the **lexA** repressor and protease discussed in Chapter 28.

A specific inhibitor of the major proteasomal activities is **lactacystin**, a compound formed by *Streptomyces*. Lactacystin is converted reversibly, by loss of *N*-acetylcysteine, into a β-lactone known as *clasto*-lactacystin. The N-terminal amino group attacks the reactive four-membered ring of the lactone (Eq. 12-22).[358,359]


O
NH
O
C
S
COOH
H3C
N
H
C
O
CH3
HO
HO
CH3
CH3
Lactacystin
N-Acetylcysteine
O
NH
OH
H3C
clasto-Lactacystin
β-lactone
O
C=O
H
O
CH3
H
H
N
Enzyme
N
H
H
O
N-Terminal threonine
O
NH
OH
H3C
HO
C=O
O
H
CH3
H
N
Enzyme
H3N+
O
Inhibited enzyme

(12-22)

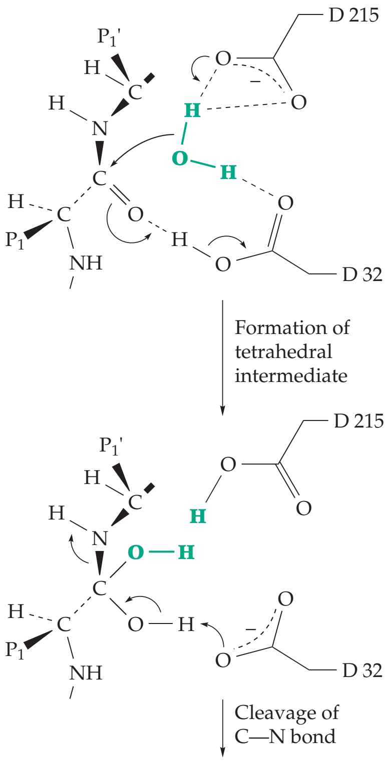
The active sites of the N-terminal nucleophile hydrolases are generated *autocatalytically*.[360–362] A single peptide chain is cleaved to form α and β chains as in the following diagram. An activating nucleophile such as histidine removes a proton from an adjacent threonine –OH and the resulting alkoxide ion attacks the adjacent peptide linkage, presumably via a tetrahedral intermediate, to form an ester linkage. Compare this sequence with reactions in Box 12-A and Eq. 14-41. Hydrolysis generates the N-terminal threonine as indicated:



4. The Aspartic Proteases

A fourth large group of protein-hydrolyzing enzymes consists of **pepsin** of the stomach[363] and related enzymes.[364] Each of these ~320-residue proteins is folded into two domains which associate with a pseudotwofold axis of symmetry that passes through the active site.[365] A second human gastric aspartic proteinase is **gastricsin**.[365] The related **chymosin** (rennin),[366] which is obtained from the fourth stomach of the calf, causes a rapid clotting of milk and is widely used in manufacture of cheese. Pepsin has a broad specificity but cleaves preferentially between pairs of hydrophobic residues, converting proteins into soluble fragments. It is unusual in being able to cleave X-Pro peptide bonds.[367] Chymosin has a more restricted specificity, cutting the κ-casein of milk between a Phe–Met bond. This decreases the stability of the milk micelles, inducing clotting.[366] The serum protein **renin** (distinct from rennin), the lysosomal **cathepsins D** and **E**,[368–370] an **aspartyl aminopeptidase**,[371] and various fungal proteases are also closely related.[372–375] Renin,[376,377] which is synthesized largely in the kidneys, is involved in blood pressure regulation (Box 22-D). More distantly related aspartic proteases are encoded by retroviruses.

The pepsin family is most active in the low pH range 1–5. All of the enzymes contain two especially reactive aspartate carboxyl groups.[378] One of them (Asp 215 in pepsin) reacts with site-directed diazonium compounds and the other (Asp 32) with site-directed epoxides.[379] It is attractive to think that one of these carboxyl groups might be the nucleophile in a double displacement mechanism. The second carboxyl could then be the proton donor to the cleaving group. The acyl-enzyme would be an acid anhydride. However, X-ray studies on pepsin[363,380] and on related fungal enzymes such as **penicillopepsin**[381] and others[373,375] suggested a different possibility: A water molecule, hydrogen bonded to one of the active site Asp carboxylates or bridging between them, becomes the nucleophile. A proton, held by the carboxylate pair, protonates the substrate carbonyl to facilitate nucleophilic attack.[381] This is illustrated in Eq. 12-23 but without detail. Several intermediate sequences of reaction steps are possible.[382] At the beginning of the sequence one of the two symmetrically placed carboxylates is protonated while the other is not. Also notice that although the active site is symmetric, the substrate is bound asymmetrically as determined by its hydrogen bonding into an extended binding site.



(12-23)

A characteristic feature of catalysis by the aspartic proteases is a tendency, with certain substrates, to catalyze transpeptidation reactions of the following type.

$$\text{Acetyl-Phe-Tyr} + \text{Acetyl-Phe*} \rightarrow \text{Acetyl-Phe*-Tyr} \qquad (12\text{-}24)$$

Here Phe* is an isotopically labeled residue. Although such reactions suggested the possibility of some kind of activated amino group on the tyrosine that is cut off in the initial cleavage of the unlabeled substrate, it is more likely that the released tyrosine stays in the active site,[383] while the acetyl-Phe fragment exchanges with acetyl*-Phe.

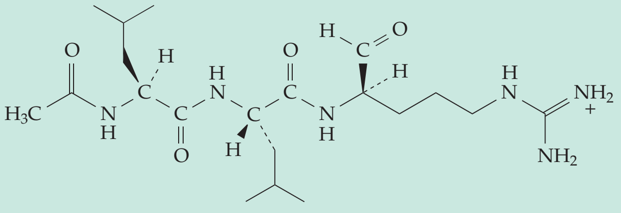
BOX 12-C SYNTHETIC PROTEASE INHIBITORS

One of the goals of synthetic medicinal chemistry is to design potent inhibitors of clinically important proteases. Elastase inhibitors may be useful for treatment of emphysema, pancreatitis, and arthritis,[a,b] while inhibitors of the angiotensinogen-converting enzyme or of renin (Box 22-D) can help control blood pressure. Inhibition of thrombin, factor Xa, or other blood clotting factors (Fig. 12-17) may prevent blood clots and inhibition of the cytosolic tryptase may provide a new treatment for asthma. Inhibition of the cysteine protease cathepsin K may help combat osteoporosis and inhibition of cysteine proteases of corona viruses may fight the common cold. Cysteine proteases of schistosomes are also targets for protease inhibitors.[c]

Many chemical approaches are used in designing inhibitors. Often, a naturally occurring inhibitor provides a starting point. The availability of high-resolution structures of the target enzymes and of various enzyme-inhibitor complexes assists in the rational design of tight-binding inhibitors. Use of combinatorial chemistry (Chapter 3)[d] is another source of potential inhibitors. To be of practical use in medicine many criteria of stability, solubility, and low toxicity must be met. While most inhibitors are disappointing as drugs, their use in laboratory experimentation has clarified a great deal of biochemistry.

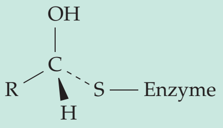
A straightforward approach is to hunt for short polypeptides that meet the specificity requirement of an enzyme but which, because of peculiarities of the sequence, are acted upon very slowly. Such a peptide may contain unusual or chemically modified amino acids. For example, the peptide **Thr-Pro-nVal-NMeLeu-Tyr-Thr** (nVal = norvaline; NMeLeu = *N*-methylleucine) is a very slow elastase substrate whose binding can be studied by X-ray diffraction and NMR spectroscopy.[e] Thiol proteases are inhibited by **succinyl-Gln-Val-Val-Ala-Ala-*p*-nitroanilide**, which includes a sequence common to a number of naturally occurring peptide inhibitors called **cystatins**.[f] They are found in various animal tissues where they inhibit cysteine proteases.

A group of inhibitors such as **leupeptin** have C-terminal aldehyde groups. Small oligopeptides with this structure and with appropriate specificity-determining side chains form tetrahedral hemiacetals, which may mimic transition state structures, at the active sites of the target enzymes.[g,h]

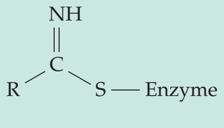


Leupeptin (*Actinomyces*)

Leupeptin is a slow, tight-binding inhibitor of trypsin. Some peptide aldehydes are potent, reversible inhibitors of cysteine proteases forming hemithioacetals with the active site cysteine.[i] Similarly, peptide nitriles form thioimidate adducts.[h]

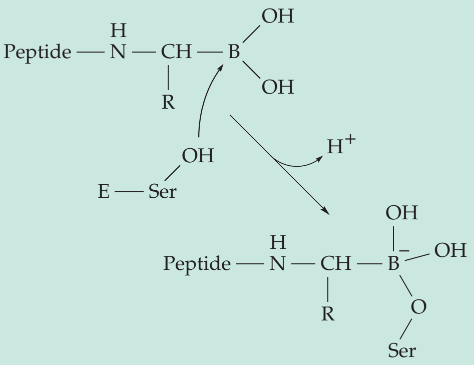


Hemithioacetal



Thioimidate

The peptide **boronic acids** form adducts with the active site serine of serine proteases.[j–l]



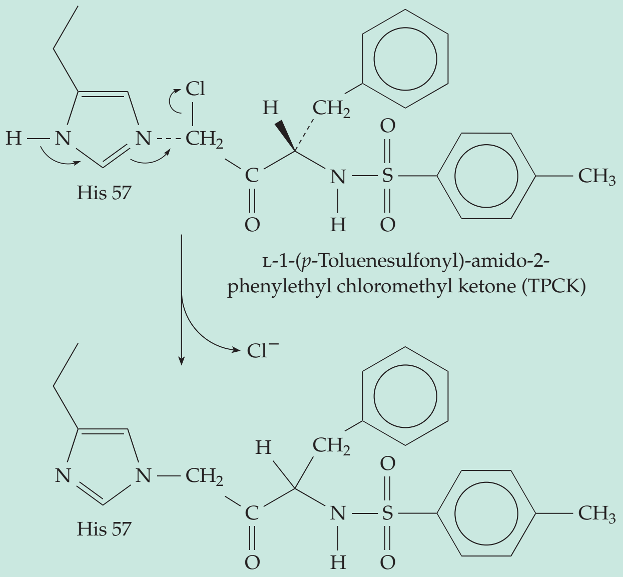
Adduct

However, both X-ray crystallography[m] and ^{11}B NMR[n] have shown that imidazole of the catalytic triad may also add to the boronic acid and that a tetrahedral adduct with both the serine oxygen and histidine nitrogen covalently bonded to boron can also be formed.[m] Thus, in reversibly inhibited enzymes a mixture of different chemical species may exist. Inhibitors can be designed to bind more tightly by providing additional bonding opportunities. For example, a suitably placed cyano group on a phenylalanine ring in the P_1 position of thrombin

BOX 12-C (continued)

or other serine proteases can form a hydrogen bond to the peptide NH of Gly 219 (see Fig. 12-9).[o]

Numerous synthetic active-site directed or enzyme-activated irreversible inhibitors have been designed.[p] For example, the following chloroketone inhibits chymotrypsin but does not act on trypsin. The corresponding structure with a lysine side chain (TLCK) inhibits only trypsin. These **affinity labeling compounds** initially bind noncovalently at the active site.



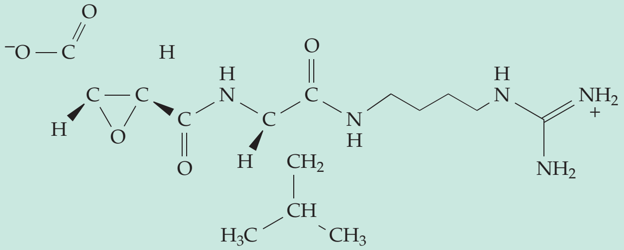
L-1-(*p*-Toluenesulfonyl)-amido-2-phenylethyl chloromethyl ketone (TPCK)

However, α-chloroketones are powerful alkylating agents and the bound inhibitor attacks His 57 of the catalytic triad system. The reaction is probably more complex than is indicated in the foregoing equation and may involve an epoxy ether intermediate.[q] Many other peptide chloromethyl ketone inhibitors have been devised.[q,r]

Isocoumarins inactivate many serine proteases. For example, 7-amino-4-chloro-3-methoxyisocoumarin acylates serine 195 of elastases as follows.[s]

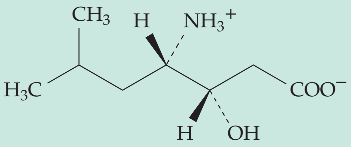
Many other enzyme-activated inhibitors are being developed.[c,d,t]

Epoxy groups, such as that of E-64, a compound isolated from the culture medium of a species of *Aspergillus*, react irreversibly with the active site thiolate group of cysteine proteases.[i,u,v] Related epoxides may become useful medications against abnormal cathepsin levels.



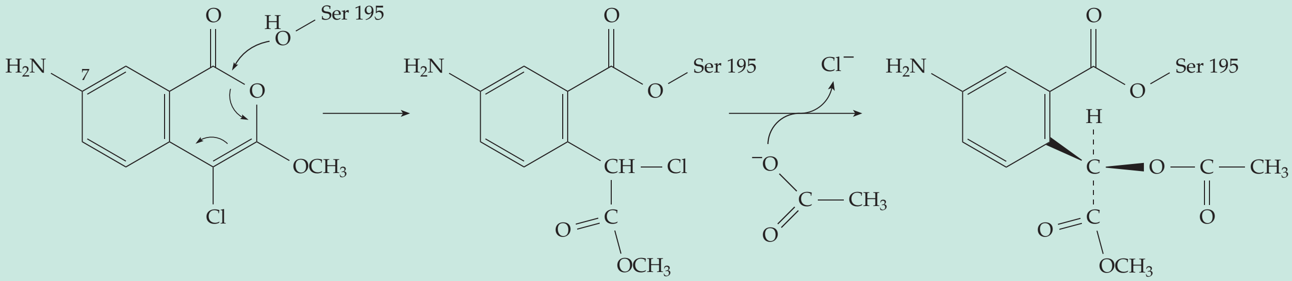
E-64, from *Aspergillus japonicum*

All of the aspartic proteases are inhibited by **pepstatin**, a peptide produced by some species of *Actinomyces* and which contains two residues of the unusual amino acid **statine** (**sta**).[w] Pepstatin has the sequence Isovaleryl-L-Val-L-Val-Sta-L-Ala-Sta.



Statine (Sta)

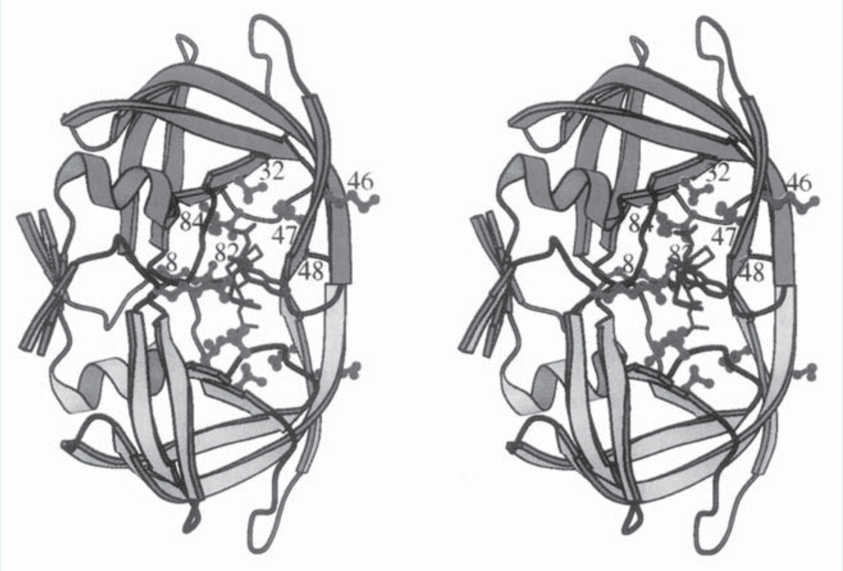
The statine residue mimics the noncovalently bonded tetrahedral intermediate, permitting formation of a very tight complex. Pepstatin is a poor inhibitor of human renin but its existence has inspired the synthesis of numerous related compounds, some of which are effective renin inhibitors.[x,y] Some of these inhibitors use the human angiotensinogen sequence with a secondary alcohol group mimicking the tetrahedral intermediate.[x,z]



BOX 12-C SYNTHETIC PROTEASE INHIBITORS (continued)

These aspartic protease inhibitors are also “lead compounds” in the development of inhibitors of HIV protease.[aa–cc] As in statine-based inhibitors, the site of occupancy by the catalytic H_2O (green in Eq. 12-23) is occupied in the inhibitor by something that mimics a tetrahedral intermediate with –CHOH–, $-PO_2H-$, etc.[bb] Tremendous efforts are being expended in designing these inhibitors and considerable success has been achieved. However, the rapid development of mutant strains of the virus with drug-resistant proteases presents a major challenge.[cc,dd]

Mercaptans of suitable structure bind tightly to Zn^{2+} in the active sites of metalloproteases. For example, **captopril**[ee] is a tight-binding competitive inhibitor of the angiotensinogen-converting enzyme that is effective in lowering blood pressure and the first of many related inhibitors.[ff] As mentioned in the text, a variety of inhibitors that mimic the geometry of a tetrahedral intermediate or transition state are also potent inhibitors of metalloproteases.



Stereoscopic ribbon structure of the HIV-1 protease with the synthetic inhibitor Sequinivir[cc] bound in the active site. One of the two identical subunits (top) is shaded darker than the second (bottom). When mutated, the amino acid side chains shown in ball-and-stick form with residue numbers shown for the top subunit led to drug-resistant viruses. Courtesy of Alex Wlodawer, National Cancer Institute.[cc]

---

[a] Powers, J. C., Oleksyszyn, J., Narasimhan, S. L., Kam, C.-M., Radhakrishnan, R., and Meyer, E. F., Jr. (1990) *Biochemistry* **29**, 3108–3118

[b] Mattos, C., Giammona, D. A., Petsko, G. A., and Ringe, D. (1995) *Biochemistry* **34**, 3193–3203

[c] Seife, C. (1997) *Science* **277**, 1602–1603

[d] Peisach, E., Casebier, D., Gallion, S. L., Furth, P., Petsko, G. A., Hogan, J. C., Jr., and Ringe, D. (1995) *Science* **269**, 66–69

[e] Meyer, E. F., Jr., Clore, G. M., Gronenborn, A. M., and Hansen, H. A. S. (1988) *Biochemistry* **27**, 725–730

[f] Yamamoto, A., Tomoo, K., Doi, M., Ohishi, H., Inoue, M., Ishida, T., Yamamoto, D., Tsuboi, S., Okamoto, H., and Okada, Y. (1992) *Biochemistry* **31**, 11305–11309

[g] Ortiz, C., Tellier, C., Williams, H., Stolowich, N. J., and Scott, A. I. (1991) *Biochemistry* **30**, 10026–10034

[h] Dufour, E., Storer, A. C., and Ménard, R. (1995) *Biochemistry* **34**, 9136–9143

[i] Mehdi, S. (1991) *Trends Biochem. Sci.* **16**, 150–153

[j] Bone, R., Shenvi, A. B., Kettner, C. A., and Agard, D. A. (1987) *Biochemistry* **26**, 7609–7614

[k] Takahashi, L. H., Radhakrishnan, R., Rosenfield, R. E., Jr., and Meyer, E. F., Jr. (1989) *Biochemistry* **28**, 7610–7617

[l] Nienaber, V. L., Mersinger, L. J., and Kettner, C. A. (1996) *Biochemistry* **35**, 9690–9699

[m] Stoll, V. S., Eger, B. T., Hynes, R. C., Martichonok, V., Jones, J. B., and Pai, E. F. (1998) *Biochemistry* **37**, 451–462

[n] Zhong, S., Jordan, F., Kettner, C., and Polgar, L. (1991) *J. Am. Chem. Soc.* **113**, 9429–9435

[o] Lee, S.-L., Alexander, R. S., Smallwood, A., Trievel, R., Mersinger, L., Weber, P. C., and Kettner, C. (1997) *Biochemistry* **36**, 13180–13186

[p] Bode, W., Meyer, E., Jr., and Powers, J. C. (1989) *Biochemistry* **28**, 1951–1963

[q] Kreutter, K., Steinmetz, A. C. U., Liang, T.-C., Prorok, M., Abeles, R. H., and Ringe, D. (1994) *Biochemistry* **33**, 13792–13800

[r] Wolf, W. M., Bajorath, J., Müller, A., Raghunathan, S., Singh, T. P., Hinrichs, W., and Saenger, W. (1991) *J. Biol. Chem.* **266**, 17695–17699

[s] Meyer, E. F., Jr., Presta, L. G., and Radhakrishnan, R. (1985) *J. Am. Chem. Soc.* **107**, 4091–4094

[t] Groutas, W. C., Kuang, R., Venkataraman, R., Epp, J. B., Ruan, S., and Prakash, O. (1997) *Biochemistry* **36**, 4739–4750

[u] Yamamoto, D., Matsumoto, K., Ohishi, H., Ishida, T., Inoue, M., Kitamura, K., and Mizuno, H. (1991) *J. Biol. Chem.* **266**, 14771–14777

[v] Varughese, K. I., Su, Y., Cromwell, D., Hasnain, S., and Xuong, N.-h. (1992) *Biochemistry* **31**, 5172–5176

[w] Gómez, J., and Freire, E. (1995) *J. Mol. Biol.* **252**, 337–350

[x] Cooper, J., Quail, W., Frazao, C., Foundling, S. I., Blundell, T. L., Humblet, C., Lunney, E. A., Lowther, W. T., and Dunn, B. M. (1992) *Biochemistry* **31**, 8142–8150

[y] Tong, L., Pav, S., Lamarre, D., Pilote, L., LaPlante, S., Anderson, P. C., and Jung, G. (1995) *J. Mol. Biol.* **250**, 211–222

[z] Cooper, J. B., Foundling, S. I., Blundell, T. L., Boger, J., Jupp, R. A., and Kay, J. (1989) *Biochemistry* **28**, 8596–8603

[aa] Hui, K. Y., Manetta, J. V., Gygi, T., Bowdon, B. J., Keith, K. A., Shannon, W. M., and Lai, M.-H. T. (1991) *FASEB J.* **5**, 2606–2610

[bb] Abdel-Meguid, S. S., Zhao, B., Murthy, K. H. M., Winborne, E., Choi, J.-K., DesJarlais, R. L., Minnich, M. D., Culp, J. S., Debouck, C., Tomaszek, T. A., Meek, T. D., and Dreyer, G. B. (1993) *Biochemistry* **32**, 7972–7980

[cc] Ridky, T., and Leis, J. (1995) *J. Biol. Chem.* **270**, 29621–29623

[dd] Chen, Z., Li, Y., Schock, H. B., Hall, D., Chen, E., and Kuo, L. C. (1995) *J. Biol. Chem.* **270**, 21433–21436

[ee] Vidt, D. G., Bravo, E. L., and Fouad, F. M. (1982) *N. Engl. J. Med.* **306**, 214–219

[ff] Gros, C., Noël, N., Souque, A., Schwartz, J.-C., Danvy, D., Plaquevent, J.-C., Duhamel, L., Duhamel, P., Lecomte, J.-M., and Bralet, J. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4210–4214

Pepsin is secreted as the inactive pepsinogen, which is activated by H^+ ions at a pH below 5. Determination of its crystal structure revealed that in the proenzyme the N-terminal 44-residue peptide segment lies across the active site, blocking it.[384] At low pH the salt bridges that stabilize the proenzyme are disrupted and the active site is opened up to substrates.

While the cellular aspartate proteases are over 300 residues in length, the retroviral proteases are less than one-half this size.[385–388] That of the human HIV-I protease contains only 99 residues. These enzymes are cut from a polyprotein (encoded by the viral ***gag*** and ***pol*** genes (Fig. 28-26). The *pol* gene encodes four other essential enzymes as well, and these are cut apart at eight different sites by action of the protease.[388] Despite its small size, it displays sequence homologies with the larger cellular aspartate proteases and has a related three-dimensional structure. Each chain has only one active site aspartate; the functional enzymes are dimers and the catalytic mechanism appears to be similar to that of pepsin.[388–392] A great deal of effort is being devoted to designing synthetic HIV protease inhibitors, which are used in the treatment of AIDS (Box 12-C).

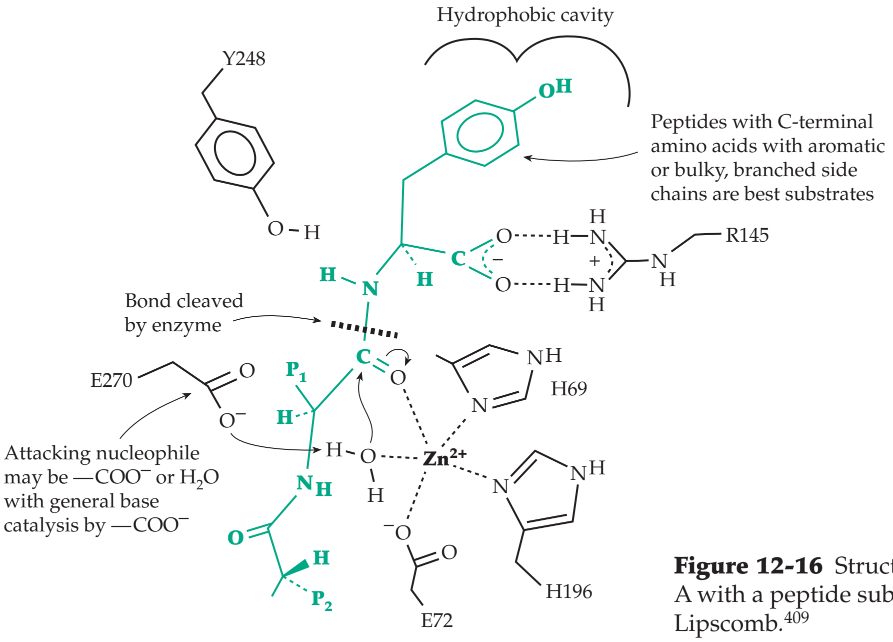
5. Metalloproteases

The pancreatic **carboxypeptidases** are characterized by the presence of one firmly bound **zinc ion** in each molecule. The Zn^{2+} can be removed and can be replaced by other metal ions such as Co^{2+} and Ni^{2+}, in some cases with reconstitution of catalytic activity. The human pancreas synthesizes and secretes proenzymes for two forms of carboxypeptidase A, with a preference for C-terminal hydrophobic residues, as well as carboxypeptidases B, which prefer C-terminal basic residues. Additional A and B forms[393] as well as more specialized nondigestive carboxypeptidases are also known. In eukaryotic cells the latter participate in processing of proteins. Following removal of N-terminal signal sequences, processing often continues with removal of basic residues from the C termini by **carboxypeptidases N**, **H** (also called E or enkephalin convertase),[394] and **M**, all of which are metalloenzymes.[395,396] Carboxypeptidase N removes C-terminal arginines from many biologically important peptides. It also circulates in the plasma and protects the body by inactivating such potent inflammatory peptides as the kinins and anaphylotoxins.[397] Carboxypeptidase H is located in secretory granules, while carboxypeptidase M is membrane associated.[396] Dipeptidyl carboxypeptidase (**angiotensin-converting enzyme**) removes the C-terminal Pro–Phe dipeptide from angiotensinogen to generate the potent pressor agent angiotensin I (Box 22-D) and cleaves dipeptides from many other substrates as well.[398] A **D-alanyl-D-alanyl carboxypeptidase** cleaves D-alanine from the ends of cell wall peptides (Chapter 20).[399]

Another well-known zinc-containing enzyme is **thermolysin**, a nonspecific *endopeptidase* widely used in laboratories. Produced by a themophilic bacterium, it is unusually resistant to heat. It contains four bound calcium ions in addition to the active site zinc.[400–402] The active site structure resembles that of pancreatic carboxypeptidase A and the two enzymes appear to act by similar mechanisms.[401,403,404] The mammalian zinc endopeptidase **neprilysin**, an integral membrane protein involved in inactivation of enkephalins and other signaling peptides, also resembles thermolysin.[405] A related neutral endopeptidase is the product of a gene called *PEX* (phosphate-regulating gene with homologies to endopeptidases on the X chromosome). The absence of the *PEX* gene product causes **X-linked hypophosphatemic rickets** which leads to excessive loss of phosphate from the body with defective mineralization of bone.[406] The **mitochondrial processing peptidase**, which removes signal sequences from the N termini of mitochondrial proteins, also contains Zn^{2+} at its active site.[407,408] However, it is an αβ heterodimer and a member of an additional family of enzymes, one of which includes human insulin-degrading enzyme.

In both carboxypeptidase A and thermolysin the active site Zn^{2+} is chelated by two imidazole groups and a glutamate side chain (Fig. 12-16). In carboxypeptidase A, Arg 145, Tyr 248, and perhaps Arg 127 form hydrogen bonds to the substrate. A water molecule is also bound to the Zn^{2+} ion. The presence of the positively charged side chain of Arg 145 and of a hydrophobic pocket accounts for the preference of the enzyme for C-terminal amino acids with bulky, nonpolar side chains. The Zn^{2+} in thermolysin is also bound to two imidazole groups and that in D-alanyl-D-alanyl carboxypeptidase to three.

The presence of a zinc ion in the metalloproteases immediately suggested a role in catalysis. Unlike protons, which have a weak affinity for the oxygen of an amide carbonyl group, a metal ion can form a strong complex. If held in position by other ligands from the protein, a properly placed zinc ion might be expected to greatly enhance the electrophilic nature of the carbon atom of the C=O group. However, it has been difficult to establish the exact mechanism of action.[410] Carboxypeptidase A cleaves both peptides and ester substrates. For peptides, K_m is the same for various metals while k_{cat} changes, but the converse is true for ester substrates.[411] From its position Glu 270 (Glu 143 in thermolysin) seems to be the logical nucleophile to attack the substrate to form an acyl-enzyme intermediate, an anhydride. Using the following specific *ester* substrate, Makinen *et al.* showed that at very low temperatures of −40° to −60°C, in solvents such as 50:50 ethylene glycol:water, an acyl-enzyme intermediate can be detected spectroscopically. It could even be separated from free enzyme[412] by gel



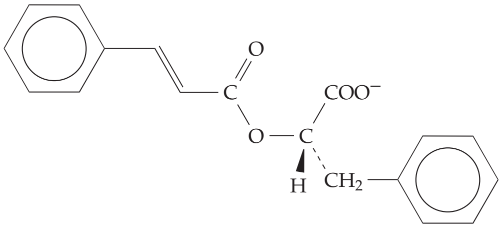
Peptides with C-terminal amino acids with aromatic or bulky, branched side chains are best substrates

Bond cleaved by enzyme

Attacking nucleophile may be —COO⁻ or H_2O with general base catalysis by —COO⁻

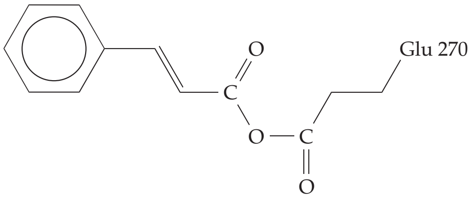
Figure 12-16 Structure of the active site of carboxypeptidase A with a peptide substrate present. See Christianson and Lipscomb.[409]

filtration at –60°C and its cyanoborohydride reduction product was characterized.[413]



O-(*trans*-*p*-cinnamoyl)-L-β-phenyllactate

The intermediate appears to be the acid anhydride formed by Glu-270.[414,415] The conformation of the intermediate was deduced by ENDOR spectroscopy and its formation and reaction interpreted according to stereoelectronic principles.[416]



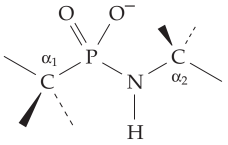
If the anhydride mechanism is correct, the water molecule bound to the Zn^{2+} probably provides an HO^- ion necessary for the hydrolysis of the intermediate anhydride.[417]

Although these results seem convincing there are objections.[410] The mechanism deduced for hydrolysis of an ester may be different than that for a peptide.[418] Furthermore, many observations favor an alternative mechanism. A hydroxide ion derived from a water molecule bound to the zinc ion may be the initial attacking nucleophile.[404,419–422] Both X-ray crystallographic studies and EPR investigations[403,414] show that the zinc in carboxypeptidases can coordinate at least five surrounding atoms. As shown in Fig. 12-16, the Zn^{2+} could hold the attacking water molecule and, at the same time, provide the positive charge for stabilizing the resulting tetrahedral intermediate. Glu 270 (or in thermolysin His 231) acts as a base to deprotonate the bound H_2O. Mock and Stanford argue that the H_2O is probably displaced when the peptide carbonyl binds and that the H_2O is then deprotonated and adds to the polarized carbonyl.[404]

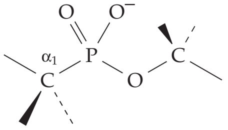
The pH dependence of the action of carboxypeptidase A is determined by pK_a values of ~6 and ~9.5 for the free enzyme[422] and of ~6.4 and ~9 for k_{cat}.[420] For thermolysin the values for k_{cat} are ~5 and ~8.[404] Assignment of pK_a values has been controversial. They may all be composites of two or more microscopic constants but probably, at least for carboxypeptidase, the low pK_a is largely that of Glu 270 while the high one represents largely the dissociation of a proton from the zinc-bound H_2O.

Earlier studies of carboxypeptidase had indicated that Tyr 248 moves its position dramatically upon substrate binding, and it was suggested that its phenolic –OH group protonates the leaving group in the acylation step. However, a mutant in which Tyr 248 was replaced by phenylalanine still functions well.[423]

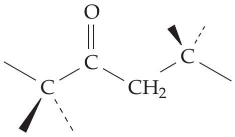
Phosphonamidates,[424,425] phosphonate esters, and oxo-methylene substrate analogs, which presumably mimic the geometry of the tetrahedral intermediate or transition state of the intermediate or transition state of the catalytic cycle, are effective inhibitors of zinc proteases. X-ray studies of complexes of such inhibitors with both thermolysin and carboxypeptidase A support the suggestion that the Zn^{2+} binds the carbonyl oxygen of the amide bond that is to be cleaved in a substrate and that a glutamate side chain could activate a Zn^{2+}-bound water to form an attacking nucleophile. A similar conclusion was reached from the structure of a thermolysin–product complex.[426]



Phosphonamidate



Phosphonate ester

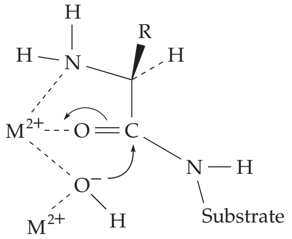


Oxo-methylene analog

Another large group of zinc proteases are the **matrix metalloproteases** which act on proteins of the extracellular matrix, such as collagen, proteoglycans, and fibronectin.[427–429c] These enzymes have been classified as **collagenases**, **gelatinases** (which act on denatured collagen), **stromelysins** (which are activated in response to inflammatory stimuli),[430–432] and a membrane type. The group also includes the digestive enzyme **astacin**, from the crayfish,[433] and metalloproteases from snake venoms.[429,434] The matrix metalloproteases are essential to the remodeling of the extracellular matrix that occurs during wound healing, tissue growth, differentiation, and cell death. An example of tissue remodeling is the development of dental enamel (Box 8-G). The proteinaceous matrix formed initially must be digested away, perhaps by the metalloprotease **enamelysin**, and replaced by the dense mineral of enamel.[435] Excessive secretion of collagenase by fibroblasts is observed in **rheumatoid arthritis** and other inflammatory conditions.

At least 64 different matrix metalloproteins are known.[427] Each enzyme consists of three domains. An 80- to 90-residue N-terminal propeptide domain contains a cysteine whose $-S^-$ group binds to the active site zinc, screening it from potential substrates. The central catalytic domain is followed by a hinge region and a C-terminal domain that resembles the serum iron binding and transporting **hemopexin**.[427,436] The mechanism of action is probably similar to that of thermolysin.[430]

Many aminopeptidases are metalloenzymes.[437] Most studied is the cytosolic **leucine aminopeptidase** which acts rapidly on N-terminal leucine and removes other amino acids more slowly. Each of the subunits of the hexameric enzyme contains *two* divalent metal ions, one of which must be Zn^{2+} or Co^{2+}.[438,439] A methionine aminopeptidase from *E. coli* contains two Co^{2+} ions[440,441] and a proline-specific aminopeptidase from the same bacterium two Mn^{2+}.[442] In all of these enzymes the metal ions are present as dimetal pairs similar to those observed in phosphatases and discussed in Section D,4 and to the Fe–Fe pairs of hemerythrin and other diiron proteins (Fig. 16-20). A hydroxide ion that bridges the metal ions may serve as the nucleophile in the aminopeptidases.[438] A bound bicarbonate ion may assist.[438a]



A metalloenzyme **peptide deformylase** removes the formyl groups from the N termini of bacterial proteins. Although the active site is similar to that of thermolysin,[443] the Zn^{2+} form of peptide deformylase is unstable. Both Ni^{2+} and Fe^{2+} form active, stable enzymes.[444,445]

6. ATP-Dependent Proteases

Much of metabolism is driven by the Gibbs energy of hydrolysis of ATP so it shouldn't be surprising that ATP is sometimes rather directly involved in hydrolytic degradation of polypeptide chains. Much of protein processing occurs in the endoplasmic reticulum,[446] which also assists in sorting unneeded and defective proteins for degradation in the ubiquitin–proteasome system. Proteasomal degradation occurs in the cytosol, in the nucleus, and along the ER. However, the predominant location is the *centrosome*.[446a]

Polyubiquitination of proteins requires ATP (Box 10-C) and additional ATP is utilized in the proteasomes (Box 7-A) during the selection of polyubiquitinated proteins for hydrolysis.[447,448] With 28 subunits the 26S proteasome is complex and not fully understood.[448] The ATP-hydrolyzing subunits appear to all be in the cap regions. Is the ATP used to open and close the entry pores? To induce conformational changes in all subunits as part of a catalytic cycle? Or to unfold folded proteins to help them enter the proteasomes?[449]

Some answers may be obtained from smaller bacterial, mitochondrial, and chloroplast ATP-dependent proteases. Cells of *E. coli* contain at least nine proteases, which have been named after the musical scale as Do, Re, Mi, Fa, So, La, Ti, Di, and Ci.[450,451] Most are serine proteases but two, Ci and Pi, are metalloproteins. Protease **La** (**Lon protease**, encoded by gene ***lon***) has attracted particular attention because the hydrolysis of two molecules of ATP occurs synchronously with cleavage of a peptide linkage in the protein chain.[452] This enzyme, as well as protease **Ti** (more often called **Clp**, for caseinolytic protease),[451,453,454] is ATP-dependent.[451]

The Lon protease of *E. coli* is a large 88-kDa serine protease with the catalytic domain, containing active site Ser 679, in the C-terminal half. The N-terminal portion contains two ATP-binding motifs and a linker region.[455,456] A homologous protein known either as Lon or as PIM1 is present in mitochondria.[457,458] The *E. coli* proteases Clp (Ti) include ClpAP and ClpXP, which are heterodimers, each containing the catalytic subunit ClpP and either ATPase ClpA or ClpX. The active enzyme may be designated **ClpAP**.[459–461] A homolog of ClpP has been found in human mitochondria.[461a] The ATPase Clp forms seven-subunit rings resembling the rings of proteasomes (Box 7-A), while the catalytic subunit ClpA forms six-membered rings.[451,454] Despite the mismatch in symmetry and the fact that they are serine proteases, these enzymes appear similar to proteasomes. However, each catalytic subunit has an ATPase neighbor. Why is it needed? The related **ClpB** is both an ATPase and a chaperonin. It is essential for survival of *E. coli* at high temperatures.[461b]

Another ATP-dependent protease identified among heat shock proteins of *E. coli* is known as **Hs1V–Hs1U** or (ClpQ–ClpY). It has a threonine protease active site and is even more closely remeniscient of eukaryotic proteasomes.[462–463a] Also active in *E. coli* is another ring-like protease, a membrane-bound zinc endopeptidase **FtsH** (or HflB).[463b] Similar eukaryotic proteases also exist.[460]

7. The Many Functions of Proteases

Many of the enzymes considered in the preceding sections function within the *digestive* tract. Others function in the *processing* of newly formed peptide chains, while others act in the *intracellular degradation* of proteins. Yet others are secreted from cells and function in the external surroundings. Both processing and intracellular degradation can be viewed as parts of *biosynthetic loops* (Chapter 17) that synthesize mature proteins, and then degrade them when they have served their function or have become damaged. The pathways involved are varied and complex, a natural result of thousands of enzymes and other compounds acting on accessible and chemically appropriate parts of the proteins. Evolutionary selection has evidently led to the particular set of pathways that we observe for any organisms.

Because they must often cleave large **polyproteins**, many viruses encode processing proteases.[464–466b] For example, the entire RNA genome of the poliovirus encodes a large polyprotein which is cut by two virally encoded chymotrypsin-like cysteine proteases within Tyr-Gly and Gln-Gly sequences;[341,465] Asn-Ser sequences are also cut, apparently autocatalytically. As we have already seen, retroviruses encode their own aspartic protease. Most cellular and secreted proteins of bacteria or eukaryotes also undergo processing. This ranges from removal of *N*-formyl groups to cutting off signal peptides, addition or formation of prosthetic groups, internal cleavages, and modifications at the C termini. A variety of peptidases are required.

Degradation of proteins, which converts them back to amino acids as well as other products, takes place in part in the cytosol via the ubiquitin-protease system. Proteolysis also occurs in ER and external cell spaces through the action of membrane-bound and secreted proteases. Loosely folded proteins, which can arise in various ways, are subject to rapid degradation. For example, synthesis of a polypeptide chain on a ribosome may be accidentally disrupted with formation of a protein with a shorted chain. Mature proteins may become damaged. For example, certain histidine residues are readily attacked by oxidizing reagents. Oxidative damage may mark proteins for rapid degeneration.[450] In *E. coli,* proteases So and Re attack oxidized glutamine synthetase much more rapidly than the intact native enzyme.[450,467] In mammalian cells the proteasomes degrade oxidized as well as poorly folded proteins.[468] However, proteasomal digestion is not complete and peptide fragments may be secreted. Peptide fragments of proteins are also formed in the ER and may be secreted. Thus, a variety of small peptide hormones and other biologically active peptides are both generated and inactivated in or around cells.[469–471] Some receptors are activated by proteases.[471a]

Circulating proenzymes of the blood clotting factors, of the complement system (Chapter 31), represent a specialized group of secreted **signaling proteins** that are able to initiate important defensive cascades. Proteases also act more directly in defense systems of the body. For example, serine proteases cause lysis of the target cells of cytotoxic T lymphocytes[472] (Chapter 31) and activated neutrophils[473] (Chapter 18). At the same time, pathogenic bacteria often secrete proteases that assist in attack on their hosts[474] and schistosomes secrete an elastase that helps them penetrate skin and invade their hosts.[475]

8. Protease Inhibitors of Animals and Plants

Premature conversion of proenzymes such as trypsinogen into active proteases in the pancreas would be disastrous. To prevent this, the pancreas also produces inhibitors. The complex cascades that control blood clotting would also be unstable were it not for the presence in blood of numerous inhibitory proteins. Indeed, inhibitors of proteases are found everywhere in animals, plants, and microorganisms.[476–478] They are usually proteins but small antibiotic protease inhibitors are also produced by microorganisms. Protein inhibitors are usually specific for a given type of enzyme: The **serpins** inhibit serine proteases, and the **cystatins** inhibit cysteine proteases.[479–483] Other inhibitors block the action of metalloproteases[484–486] and aspartic proteases. Inhibitors help not only host organisms but also parasites. For example, *Ascaris suum*, a very large nematode that is thought to infect 1/4 of the human population of the earth and nearly all of the pigs, secretes a pepsin inhibitor.[487] However, the large 720-kDa serum **α-macroglobulins** (Box 12-D) inhibit all classes of proteases.

About 20 families of protein inhibitors of proteases have been described.[488] The egg white **ovomucoids** comprise one family. Turkey ovomucoid is a three-domain protein whose 56-residue third domain is a potent inhibitor of most serine proteases.[488,489] The 58-residue **pancreatic trypsin inhibitor**[490] is a member of another family of small proteins. A 36-residue insect (locust) protease inhibitor is even smaller.[491]

Inhibitors that block the action of trypsin and other proteases are found in many plants. The inhibitor activity is usually highest in seeds and tubers, but synthesis of inhibitors can be induced in other parts of plants e.g. tomato and potato by wounding.[492] See also Chapter 31. Legumes contain small 60- to 76-residue inhibitors (Bowman–Birk inhibitors) each containing seven disulfide linkages, and they are relatively stable toward cooking and toward acid denaturation in the stomach.[493,494] Although they interfere with protein digestion they seem to have an anticarcinogen effect.[493] Soybeans also contain a larger 181-residue trypsin inhibitor (Kunitz type).[495]

Several inhibitor–protease complexes have been crystallized and details of their interactions are known. For example, the pancreatic trypsin inhibitor binds at the active site of trypsin with $K_f > 10^{13}\ M^{-1}$ at neutral pH.[496] The two molecules fit snugly together,[490,497] the inhibitor being bound as if it were a peptide substrate with one edge of the inhibitor molecule forming an antiparallel β structure with a peptide chain in the enzyme. Lysine 15, which forms part of this β structure, enters the specific P_1 binding site for a basic amino acid in a substrate. Thus, the protease inhibitor is a modified substrate which may actually undergo attack at the active site. However, the fit between the two molecules is so tight that it is hard for a water molecule to enter and complete the catalytic act. The complex reacts very slowly, keeping the enzyme inactive. There is not enough inhibitor to interfere with the large amounts of trypsin formed from trypsinogen in the small intestine so that trypsin can function there.

Most protease inhibitors act by mechanisms similar to that of the pancreatic trypsin inhibitor. They are very slow substrates with a **reactive loop** that carries suitable P_1, P_2, and P_1' residues that meet the specificity requirements of the enzyme. Additional noncovalent interactions prevent dissociation and make the energy barrier for hydrolysis so high that the reaction is extremely slow.[488,494,495]

The serpins are larger ~400-residue proteins.[498–500] They also form complexes in which hydrolysis and release of the serpin occurs very slowly. However, structural analysis before reaction and after release showed that a major rearrangement occurs in the serpin structure. The P_1 through ~P_{15} residues of the cleaved reactive loop are inserted into the center of the main β sheet of the serpin, leaving the P_1 and P_1' residues ~7 nm apart.[499,500] It seems likely that the rearrangement begins during formation of the tight inhibited complex, which cannot be dissociated by boiling in a sodium dodecyl sulfate (SDS) solution and which may be an acyl-enzyme.[501,502] A serpin molecule can act only once.

Blood contains several serpins. They are abundant, accounting for about 10% of the total protein of human plasma.[478,500] The most abundant is the **α1-protease inhibitor** or α_1-antitrypsin, a 394-residue glycoprotein component of the α-globulin fraction of blood serum.[500a] There is no trypsin in tissues, but this inhibitor blocks the action of other serine proteases, including cathepsin G and **leukocyte elastase**.[503] Hereditary absence of α_1-protease inhibitor often leads to severe **pulmonary emphysema** at an early age. Elastase released by neutrophils at sites of inflammation degrades many components of connective tissue including elastin, collagen, and proteoglycans. Without the presence of protease inhibitor too much damage is done to surrounding tissue. A lack of this inhibitor is one of the commonest genetic defects among persons of European ancestry, affecting 1 in 750 persons born.[504–506] The most serious known mutation is a replacement of guanine by adenine at a specific point in the DNA and a resultant replacement of Glu 342 with lysine in the protease inhibitor. This in some manner adversely affected the processing and secretion of the protein.

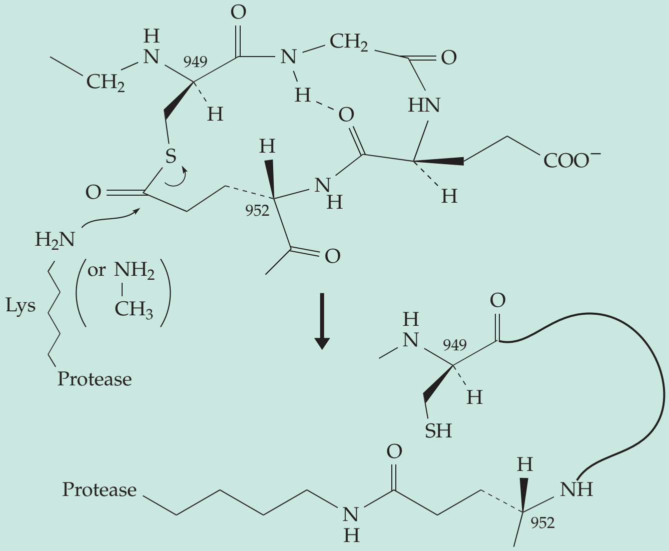
At its reactive site the α_1-protease inhibitor has the sequence

Ala-Ile-Pro-Met*-Ser-Ile-Pro-Pro,

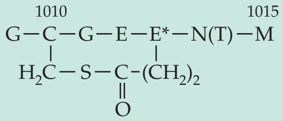
the Met-Ser pair marked by the asterisk fitting into the P_1-P_1' sites (Fig. 12-14). The methionine (Met 358) in

BOX 12-D MOLECULAR MOUSETRAPS

The large 720-kD α_2-macroglobulin of human serum, as well as related proteins of vertebrate and invertebrate circulatory systems and of egg whites of birds and reptiles, is a trap for proteases.[a–f] Human macroglobulin is a homotetramer consisting of two pairs of identical 180-kDa subunits, each pair being held together in an antiparallel configuration by two disulfide bridges. Each subunit contains a "bait region" with cleavage sites appropriate for nearly all known endoproteases[f] and also a thioester linkage as explained later. Electron microscopic reconstructions of the native protein and its complexes with proteases show that a major structural transformation occurs.[e,f] The macroglobulin traps two protease molecules of the size of trypsin, or one larger one such as plasmin, in an internal cavity. The internal thioesters, which are formed between Cys 949 and Gln 952 (with loss of NH_3) in each subunit, become reactive[g] and form covalent bonds with ε-amino groups of various lysine side chains of the trapped proteases.



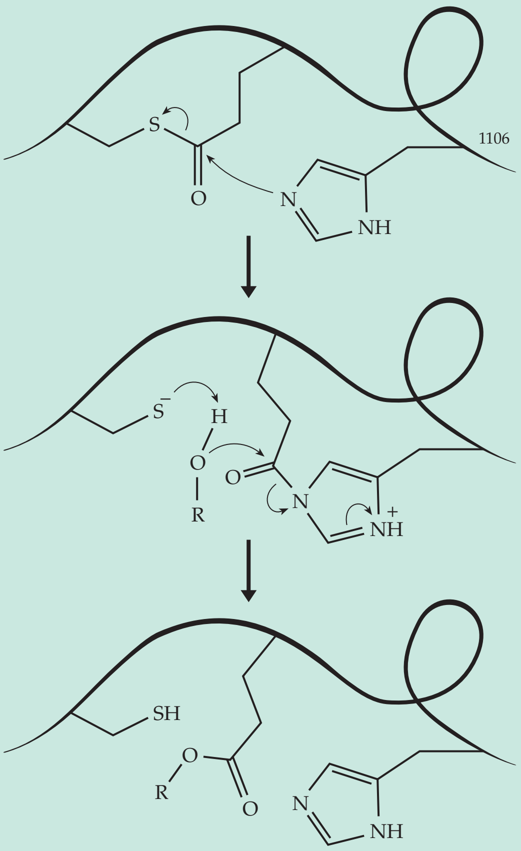
The serum proteins C3 and C4, members of the complement system (Chapter 31), are converted into their active forms, C3b and C4b, by proteolytic removal of short N-terminal peptide fragments. Both C3b and C4b bind tightly to cell surfaces, a feature that helps to direct the complement system's attack to the surfaces of invading organisms. This tight binding also involves covalent attachment of macromolecules by reaction with a preformed thioester just as with α_2-macroglobulin.[h–j] In fact, the thioester linkage was first discovered in the complement proteins. Both C3 and C4 contain the thioester within the following sequence, which is the same as that in α_2-macroglobulin:



$$\begin{array}{l} \overset{1010}{\text{G}}-\text{C}-\text{G}-\text{E}-\text{E*}-\text{N(T)}-\overset{1015}{\text{M}} \\ \quad\;\, | \qquad\qquad\quad\;\; | \\ \text{H}_2\text{C}-\text{S}-\underset{\underset{\text{O}}{\|}}{\text{C}}\text{-(CH}_2)_2 \end{array}$$

E*, encoded as Q

Here the side chains have been added for the thioester-forming cysteine and glutamate and the sequence numbers are for C3. The thioester-forming glutamate is labeled E* because it is not encoded as glutamate but as glutamine, suggesting a mechanism of thioester formation. Protein C4 exists as two subforms, C4A and C4B. Both C3 and C4A react predominately with lysine amino groups, but C4B reacts with –OH groups of cell surface polysaccharides.[h] It has a histidine at position 1106. There is good evidence that it is adjacent to the C=O group of the thioester and reacts to form an acyl-imidazole which is more reactive with hydroxyl groups than is a thioester:



Activation of C3 and C4 apparently allows the preformed thioester, which is buried in the interior

BOX 12-D (continued)

of these proteins, to become exposed on the external surface where it can react. The high group transfer potential of the thioester ensures that the reaction will go to completion.

---

[a] Sottrup-Jensen, L. (1989) *J. Biol. Chem.* **264**, 11539–11542

[b] Fothergill, J. (1982) *Nature (London)* **298**, 705–706

[c] Jacobsen, L., and Sottrup-Jensen, L. (1993) *Biochemistry* **32**, 120–126

[d] Andersen, G. R., Koch, T. J., Dolmer, K., Sottrup-Jensen, L., and Nyborg, J. (1995) *J. Biol. Chem.* **270**, 25133–25141

[e] Boisset, N., Taveau, J.-C., Pochon, F., and Lamy, J. (1996) *J. Biol. Chem.* **271**, 25762–25769

[f] Qazi, U., Gettins, P. G. W., Strickland, D. K., and Stoops, J. K. (1999) *J. Biol. Chem.* **274**, 8137–8142

[g] Gettins, P. G. W. (1995) *Biochemistry* **34**, 12233–12240

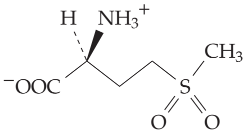
[h] Law, S. K. A., and Dodds, A. W. (1997) *Protein Sci.* **6**, 263–274

[i] Khan, S. A., Sekulski, J. M., and Erickson, B. W. (1986) *Biochemistry* **25**, 5165–5171

[j] Dodds, A. W., Ren, X.-D., Willis, A. C., and Law, S. K. A. (1996) *Nature (London)* **379**, 177–179

---

this sequence, as well as another nearby methionine residue, is very susceptible to oxidation to sulfoxides:



Methionine sulfoxide

The oxidation, which may be caused by such agents as myeloperoxidase (Chapter 16) released from leukocytes,[507] inactivates the inhibitor. This may be physiologically important in permitting the proteases to be *uninhibited* in the immediate vicinity of the leukocyte. Cigarette smoke also inactivates α_1-protease inhibitor by oxidation of the same methionine residues and the lungs of smokers contain the oxidized inhibitor.[508] However, the major cause of emphysema among smokers appears to be an increase in released neutrophil elastase.[509] One approach to the treatment of emphysema involves weekly intravenous injection of α_1-antitrypsin.[510] This treatment may be improved by use of genetically engineered oxidation-resistant variants of the antitrypsin such as Met 385→Val.[504,511] Efforts are also being made to introduce an α_1-antitrypsin gene into lung epithelial cells.[510,512]

Blood plasma also contains at least nine other protease inhibitors. One of these, the thrombin inhibitor **antithrombin III** (Section 9), contains the sequence Arg-Ser-Leu at the P_1, P_1', and P_2' sites. A tragic case of a person born with a Met 385 → Arg mutation in α_1-antitrypsin has been reported.[513] This converted the antitrypsin to an antithrombin causing a bleeding disorder that was eventually fatal.

9. Coagulation of Blood

The clotting of blood following injury and the subsequent dissolving of the clot are familiar phenomena that involve several cascades of proteolytic enzymes together with a number of accessory **cofactors**.[514–516] The first step in clotting results from "activation" of blood platelets which aggregate to form a platelet plug that slows bleeding.[514,517] The clot, which is formed by the insoluble **fibrin**, grows on the platelet surfaces and strengthens the plug. The initial rapid formation of a clot occurs via the **tissue factor pathway** (or extrinsic pathway; right side of Fig. 12-17) which is triggered by the exposure in injured tissues of **tissue factor** (TF), a transmembrane glycoprotein[518–522] and a member of the cytokine receptor superfamily.[518] Human TF is a 263-residue protein with a single membrane-spanning region and a small 20-residue C-terminal cytoplasmic domain. The 219-residue extracellular domain consists largely of two IgG-like domains (Fig. 12-18).[519,523] This protein stimulates the conversion of fibrinogen[524–526] into the insoluble fibrin through the action of three proteases–factor VIIa, factor Xa, and thrombin. These enzymes are generated from proenzymes VII, X, and prothrombin, respectively in a cascade mechanism.

Factor VII binds tightly to TF,[527,527a,b] which also binds Ca^{2+} and phospholipid of the cell membranes. Within this complex a plasma protease, such as thrombin or factor VIIa or Xa, cleaves a single Arg-Ile bond in VII to form active VIIa.[528–530] The TF•VIIa complex is a very active protease which cleaves a specific peptide bond in factor X to form Xa[531–533] which continues the cascade. Notice that there are autocatalytic features: VII can be converted to the active VIIa by the action of Xa and the accessory factor Va is generated from the precursor, factor V, in part through the action of thrombin.[514,534–535] Factors Xa and Va together with Ca^{2+} and phospholipid form the active **prothrombinase** complex which attacks prothrombin to form the active enzyme thrombin.[536,537] The roles of factor Va and Ca^{2+} appear to be to hold the prothrombin and the activated protease Xa together on the phospholipid surface.[538] This localizes the clotting. Factor Xa is unusually specific, cleaving only after arginine in the sequence. Its activation of prothrombin results from cleavage of two

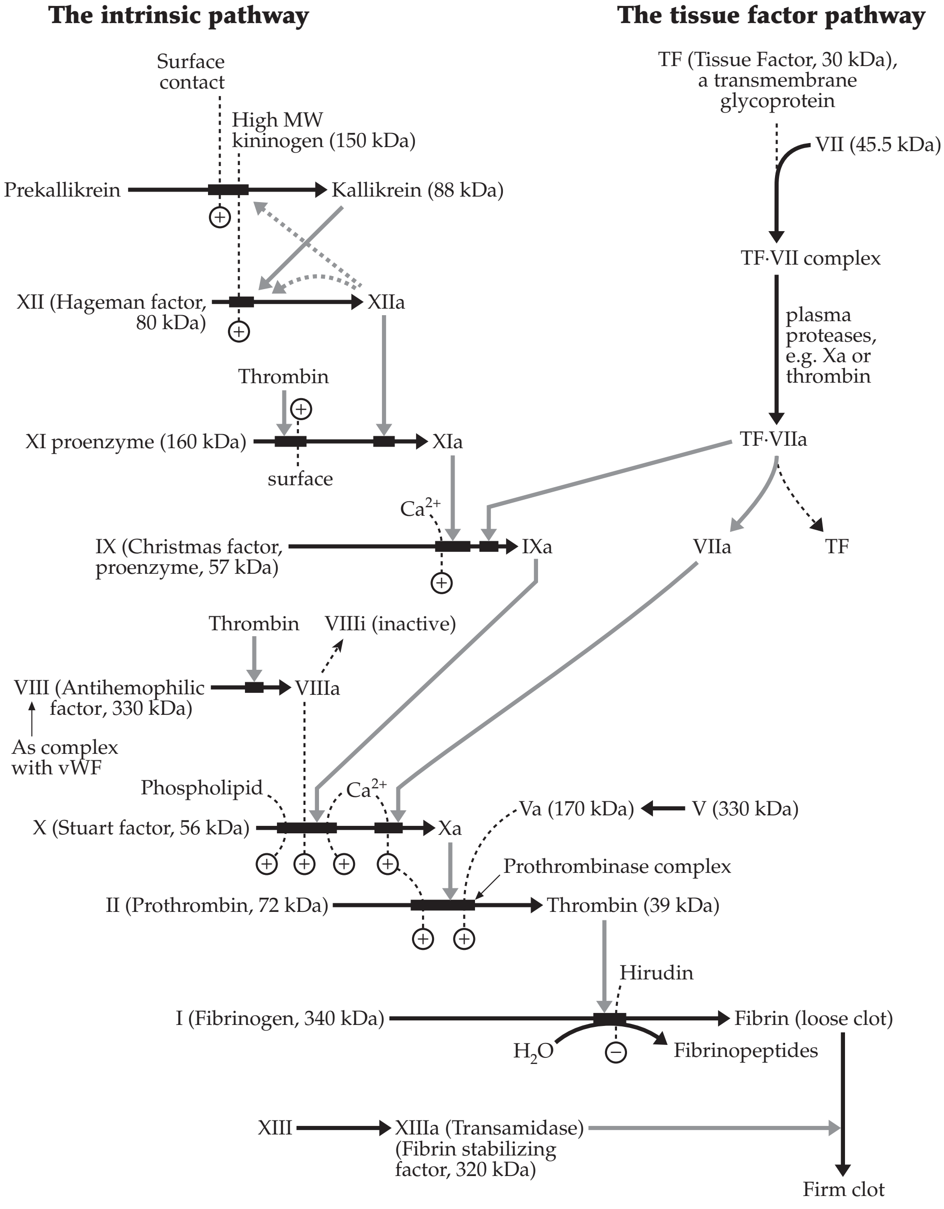

The intrinsic pathway
The tissue factor pathway
Surface contact
High MW kininogen (150 kDa)
Prekallikrein
Kallikrein (88 kDa)
XII (Hageman factor, 80 kDa)
XIIa
Thrombin
XI proenzyme (160 kDa)
XIa
surface
Ca2+
IX (Christmas factor, proenzyme, 57 kDa)
IXa
Thrombin
VIIIi (inactive)
VIII (Antihemophilic factor, 330 kDa)
VIIIa
As complex with vWF
Phospholipid
Ca2+
X (Stuart factor, 56 kDa)
Xa
Va (170 kDa)
V (330 kDa)
Prothrombinase complex
II (Prothrombin, 72 kDa)
Thrombin (39 kDa)
Hirudin
I (Fibrinogen, 340 kDa)
Fibrin (loose clot)
H2O
Fibrinopeptides
XIII
XIIIa (Transamidase) (Fibrin stabilizing factor, 320 kDa)
Firm clot
TF (Tissue Factor, 30 kDa), a transmembrane glycoprotein
VII (45.5 kDa)
TF·VII complex
plasma proteases, e.g. Xa or thrombin
TF·VIIa
VIIa
TF

Figure 12-17 Major components of the human blood clotting cascades. The site of action of the leech anticoagulant protein hirudin is also indicated.

peptide bonds, which releases the 39-kDa thrombin from the much larger, immobilized 72-kDa proenzyme.[236,531]

Thrombin, like most other clotting factors, is also a serine protease. However, the clotting factors are multidomain proteins that are more elaborate and more specific[539] than the digestive enzymes. Prothrombin, as well as factors VII, IX, and X and the anticoagulant protein C, contain at their N-terminal ends several residues of γ-carboxyglutamate (Gla), an amino acid generated in a posttranslational modification that depends upon vitamin K (Chapter 15). In human prothrombin there are ten of these in the following N-terminal sequence, where E* represents Gla. Since many of these enzymes are also dependent upon activation by calcium ions (see Fig. 12-17), it is thought that the function of Gla is to assist in the binding of Ca^{2+} which helps to tie these proteins to the phospholipids of platelet surfaces. In factors VII, IX, X, and protein C this Ca^{2+}-binding domain is followed by two epidermal growth factor (EGF)-like domains, each containing one residue of *erythro*-β-hydroxyaspartate or hydroxyasparagine formed by hydroxylation of an aspartate or asparagine residue in the first EGF-like domain.[540,540a,b] The C-terminal catalytic domain of each enzyme contains the protease active site.

```
          10           20           30
           •            •            •
ANTFLE*E*VRKGNLE*RE*CVE*E*TCSYE*E*AAFE*ALE*SS
```

Fibrinogen is an elongated molecule with an $(\alpha\beta\gamma)_2$ structure.[524,541,541a] Thrombin cleaves specific Arg-Gly bonds in the α and β chains releasing short (14- to16-residue) “fibrinopeptides” from the N termini of the peptide chains. This leaves Gly-Pro-Arg “knobs” at

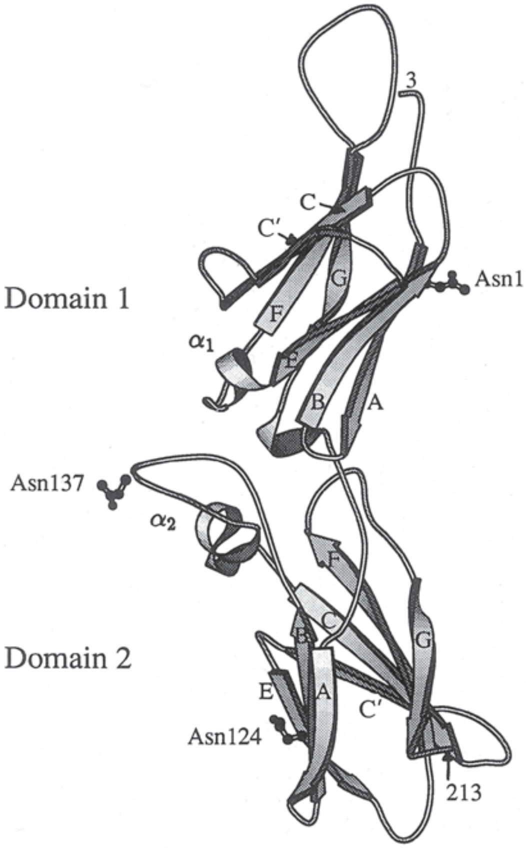


Figure 12-18 A ribbon drawing of the three-dimensional structure of the extracellular region (residues 3 to 213) of human tissue factor. Strands of domain 1 and 2 are labeled A to G. Two helices are labeled α_1, and α_2 and the three asparagines that provide the glycosylation sites are also marked. MolScript drawing from Harlos *et al*.[520]

the N termini of the α-chains and these fit into complementary "holes" in the γ chains to form noncovalently linked aggregates.[542–544] The clot that forms is unstable, but it is soon crosslinked by the action of the **transamidase** (transglutaminase or **factor XIIIa**; Eq. 2-23).[545,546] The fibrin monomers, von Willebrand factor (discussed in following paragraphs), fibronectin, collagen, and other proteins all become crosslinked.

The slower **intrinsic mechanism** first described in 1964[547,548] consists of a cascade involving six proteases (Fig. 12-17, left side). Again, autocatalytic cycles are present in the activation by XIIa of both prekallikrein[549] and XII and in activation by thrombin of factors XI and VIII. This intrinsic pathway is initiated by the serine protease proenzymes prekallikrein and factor XII together with the accessory protein high-molecular-mass-H- kininigen.[549] Activation occurs when blood contacts surfaces such as glass or kaolin (a clay).[547,548,550] Factor XI can also be activated by thrombin. Heredity absence of factor XI leads to bleeding problems, especially after surgery, but absence of XII does not. This fact suggests that direct activation of XI by thrombin is important and that the kallikrein–factor XII pathway is usually less important.

Factor IX (Christmas factor) is next in the intrinsic mechanism cascade. It can be activated either by XIa or by VIIa of the tissue factor pathway. The absence of a functional factor IX leads to the inherited X-linked bleeding disorder **hemophilia B** which affects 1 in 30,000 males. The condition can be mild or very serious[551,552] and may be caused by a variety of mutations or by incorrect splicing of the messenger RNA for the 416-residue factor IX. The level of factor IX in blood increases with age, almost doubling by old age.[552a]

Factor IXa causes a rapid activation of factor X only if Ca^{2+}, phospholipid,[553,554] and the accessory factor VIIIa[555] are present. The IXa•VIIIa complex acts on X about 2×10^5 times faster than does IXa alone. This complex cleaves the same bonds in X as does the VIIa•Va complex formed in the tissue factor pathway.[514] The 2332-residue factor VIII and factor V have similar structures that include three repeats of a domain homologous to the blue copper-containing plasma protein **ceruloplasmin** (Chapter 16).[556–559] Tyrosine 1680 of VIII apparently must be converted to a sulfate ester for full activity.[560]

The absence of factor VIII in about 1/10,000 males born, leads to the severe X-linked bleeding disorder **hemophilia A**. Human factor VIII is encoded by a 186-kb gene containing 26 coding exons. Severe cases of hemophilia are usually a result of point mutations that produce stop codons in this gene and therefore a shortened protein. Milder cases may result from an amino acid substitution.[561] Factor VIII circulates in the plasma as a complex with the **von Willebrand factor**, (vWF), a large multimeric protein derived by proteolytic cleavage, glycosylation, and sulfation of a 2813-residue precursor.[562–566] The mature 2050-residue vWF stabilizes factor VIII in the blood. Another important function of vWF is to bind platelets to damaged endothelial surfaces.[567,568] Like fibrinogen, vWF contains RGD sequences specific for binding to adhesion receptors.[569,570] It also binds to collagen. A serious inherited bleeding disorder caused by deficiency of vWF was first discovered among inhabitants of islands in the Gulf of Bothnia, Finland, in 1926 by von Willebrand.[563,571] If mild forms of the disease are included, vWD deficiency is the commonest bleeding problem. However, abnormalities have been identified in almost every one of the proteins in the coagulation cascades.[514]

Until recently, these bleeding diseases were treated by regular injections of the appropriate factors isolated from human plasma. This was both costly and carried a high risk of infection by HIV, hepatitis, or other contaminating viruses. Now, cloned genes are used for commercial production of the factors.[572,573] Experiments in animals, designed to lead to eventual gene replacement therapy, have been somewhat successful.[573,574] It is of interest to compare the relative concentrations of several of the clotting factors in plasma and to compare these with their positions in the cascade of the tissue factor (extrinsic) pathway.[515]

	mg / L	Mass kDa	μM
VII	0.5	45.5	0.01
V	10.0	330	0.03
Prothrombin	150	72	2.1
Fibrinogen	3,000	340	8.8

Why do we have the intrinsic pathway when the tissue factor pathway provides rapid clot formation? The answer seems to be that the tissue factor pathway is needed immediately after injury but that it is turned off quickly by the **anticoagulation systems** of the body. As a result the protease **plasmin** begins to dissolve (lyse) the clot within a few hours. The intrinsic pathway is apparently needed to maintain the clot for a longer period.[514]

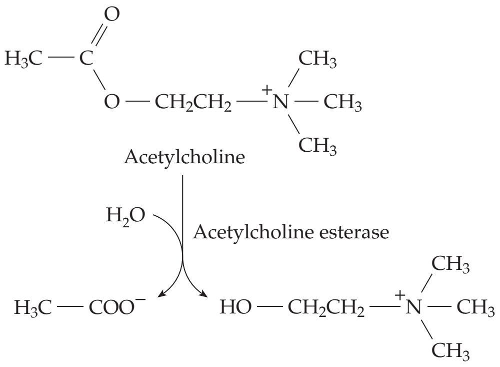
What prevents the clotting mechanism, with its autocatalytic cycles, from running out of control? In part the answer lies in the localization of the enzymatic activation to tissue surfaces near a wound. The flow of blood rapidly dilutes components that escape into the general circulation and liver cells take up and destroy the active proteases. A variety of circulating antiproteases including the **tissue factor pathway inhibitor**[575] act on these escaped enzymes. Two anticoagulation systems that are localized on vessel walls[576] also come into action very quickly. The circulating polysaccharide heparin (Chapter 4) forms a complex with the serpin **antithrombin**. Antithrombin traps thrombin as an inactive complex or compound[577,578a] and heparin greatly accelerates the inactivation.[578] Kallikrein and factors IXa, Xa, XIa, and XIIa are also inhibited.

Thrombin is an allosteric protein which exists as a mixture, in nearly equal amounts, of a fast-acting form that cleaves fibrinogen and is stabilized by Na$^+$, and a slow-acting form that initiates the second anticoagulant cascade.[579] The slow acting form, bound to **thrombomodulin**,[527a,540b,580,580a,b,581] an endothelial cell surface protein, attacks the proenzyme of another serum protease called **protein C**.[582,582a] Activated protein C (APC) inactivates the accessory clotting factors Va and VIIIa. The accessory factor **protein S**[583] is also needed for rapid inactivation. A blood clot is temporary and its dissolution begins as soon as it is formed, largely through the action of plasmin, a protease derived from the circulating 791-residue **plasminogen**[584–588] through the action of yet other proteases. Plasminogen often becomes crosslinked to cell surface proteins by transglutaminase.[586] Plasminogen activators include **urokinase**,[589] a protease present in kidney tissue and urine, and **tissue plasminogen activator** (tPA),[590,591] a 527-residue protease which is now produced by recombinant DNA technology. It is sometimes used to dissolve blood clots in emergency situations, such as myocardial infarction and pulmonary embolism,[592] but can also cause serious bleeding problems. Plasmin is inhibited by the plasma **α_2 antiplasmin** and is inactivated by action of clotting factor XIIa. Likewise, tPA is inhibited by several protease inhibitors present in tissues and in plasma. One is a fast-acting serpin called plasminogen activator inhibitor.[593] Another anticoagulant compound of medical interest is **hirudin**, a 65-residue peptide from the leech. It binds very tightly to thrombin ($K_d = 1$ pM) preventing its action.[594–596] Insects also produce antithrombin.[597,598] Ticks[599,600] and some insects[598] inject proteins that inhibit Xa selectively.[599] Anticoagulants are of great medical importance and much effort is being devoted to the design of better inhibitors of thrombin,[601–603] factor Xa,[604] and other components of the blood coagulation cascade.

Inherited deficiencies of the anticoagulant pathways with associated problems of thrombosis are known. These include problems with protein C,[576,605] plasminogen,[606] and antithrombin.[607,608]

10. Esterases and Lipases

A group of esterases hydrolyze simple oxygen esters. Some of these are designed to hydrolyze a particular ester or small group of esters, while others have a more nonspecific action. **Acetylcholinesterase**[609–611a] is specific for acetylcholine (Eq. 12-25), a neurotransmitter that is released at many nerve synapses and neuromuscular junctions (Chapter 30). The acetylcholine, which is very toxic in excess, must be destroyed rapidly to prepare the synapse for transmission of another impulse:



(12-25)

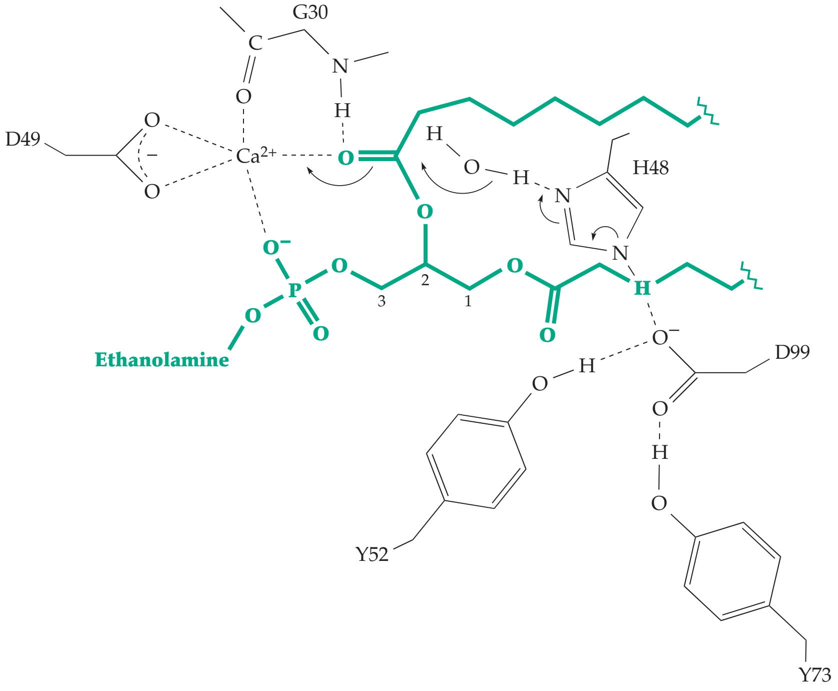
The more widely distributed **butyrylcholinesterase**[612] is less specific but prefers butyrylcholine. Acetylcholinesterase is a very efficient catalyst:[613–615] $k_{cat} = 1.6 \times 10^4$ s^{-1}

and $k_{cat} / K_m = 2 \times 10^8\ M^{-1}s^{-1}$ at 25°C. It exists as a series of molecular forms containing varying numbers of 68-kDa catalytic subunits, 100-kDa structural subunits and subunits with triple-helical ~120-kDa collagen-like "tails."[609,616] The subunits are joined by disulfide bridges to give aggregates that range from simple dimers of catalytic subunits to tailed forms containing 8–12 catalytic subunits and non-collagenous structural subunits as well. The tailed forms are secreted from cells and may be designed to take up residence in the basal lamina of synapses, whereas the dimers are apparently attached to phosphatidylinositol anchors in the membranes (Fig. 8-13).[617] Human liver **carboxylesterases** are relatively nonspecific enzymes that hydrolyze ester groups of various drugs and toxins including cocaine and heroin. Products are often excreted in the urine.[618] **Thioesterases** function in biosynthesis of fatty acids, polyketides (Chapter 21), and many other substances.[619]

Lipases[620,621] hydrolyze triacylglycerols. The pancreatic digestive lipase[622,623] acts faster on emulsified fats than on glycerol esters in true solution but requires the cooperation of a small 10-kDa **colipase**.[624] **Cholesterol esterase** hydrolyzes not only cholesterol esters but also esters of fat-soluble vitamins, phospholipids, and triacylglycerols.[625,626] Other lipases include gastric[627] and hepatic lipases and a **lysosomal acid lipase**[628] which also attacks neutral lipids. Plasma **lipoprotein lipase**[629,630] digests fats in the chylomicrons and from the very low-density lipoproteins of blood. **Hormone-sensitive lipase** hydrolyzes stored triacylglycerols in the cytosol in response to catecholamines, ACTH, and other hormones.[631] The **phospholipases** attack phospholipids, while **cutinase**,[632] produced by some fungi, cleaves the ester linkages in the cutin (Chapter 21) of plant surfaces. Fungal lipases are important industrial commodities. Numerous structural and mechanistic studies have been made with them.[620,629,633–635] The gene for a lipase from *Candida rugosa* has been synthesized using codons that maximize its expression in *Saccharomyces cerevisiae* and which allow for further genetic engineering of the lipase.[636]

All of these esterases appear to act by mechanisms closely related to those of proteases. Acetylcholinesterase contains an active site serine that reacts with organophosphorus compounds (Box 12-E) and is part of an Asp-His-Ser catalytic triad which lies in a deep "gorge" as well as an oxyanion hole.[637] A surprise is the absence of an essential carboxylate group that might bind the positively charged trimethylammonium group of acetylcholine. Instead, the lining of the gourge is rich in aromatic side chains which may interact with the methyl groups of the substrate and by their polarizability stabilize the charge.[611,638] Most lipases, including cutinase, also have an Asp-His-Ser or Glu-His-Ser triad as well as some form of oxyanion hole.[620,632,639] Like the serine proteases, the lipases have bell-shaped pH optima.

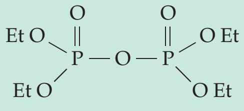
Phospholipase A_2 cleaves the ester linkage at the 2 position in phospholipids.[640–642] One isoenzyme form is secreted by the pancreas as a proenzyme whose N-terminal seven residues are removed by trypsin to give an active 125-residue enzyme. Phospholipase A_2 of a similar type is abundant in venoms of reptiles and bees. The venom and pancreatic enzymes have closely similar three-dimensional structures.[643–645] Although the folding pattern is different from that of chymotrypsin, imidazole (His 48) and carboxylate (Asp 99) groups are present in the active site in an orientation resembling that of the catalytic triad of serine protease. The enzyme requires calcium ions, one of which binds at an appropriate point for complexation with the substrate carbonyl as it is converted to an oxyanion intermediate. The backbone NH of Gly 30 may also serve as an oxyanion ligand. However, in most phospholipases there is no active site serine. Instead, a water molecule is positioned to serve as the attacking nucleophile in formation of the oxyanion as is indicated in the following scheme, which shows a truncated phosphatidylethanolamine as the substrate.[643] Phospholipase A_2 is up to 1000 or more times as active on phospholipids in micelles as on dissolved substrates.[621,646,647] Apparently, the



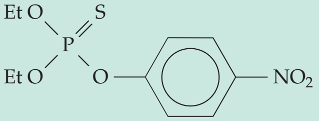
BOX 12-E INSECTICIDES

Over 200 organic insecticides, designed to kill insects without excessive danger to humans and animals, are presently in use.[a–e] Many of these compounds act by inhibiting cell respiration; others uncouple ATP synthesis from electron transport. The chlorinated hydrocarbons such as DDT act on nerves in a manner that is still not fully understood. One of the largest classes of insecticides acts on the enzyme acetylcholinesterase of nerve synapses. Like chymotrypsin, it contains an active site serine residue that reacts with organophosphorus compounds. The extreme toxicity of esters of pyrophosphate and of dialkylphosphonofluoridates was recognized in the 1930s and led to their development in Germany and in England as insecticides and as nerve gases. Among the most notorious is diisopropylphosphonofluoridate (diisopropylfluorophosphate; DFP), for which the LD_{50} (dose lethal to 50% of the animals tested) is only 0.5 mg kg^{-1} intravenously. This exceedingly dangerous compound can cause rapid death by absorption through the skin.

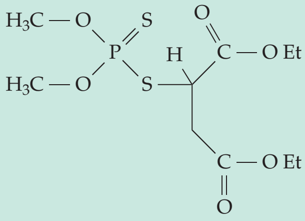
The following structures are a few of the organophosphorus compounds and other acetylcholinesterase inhibitors that are selectively toxic to insects.



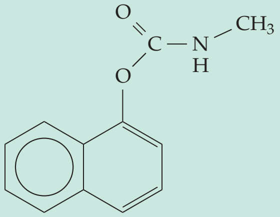
Tetraethyl pyrophosphate
(TEPP)



Parathion



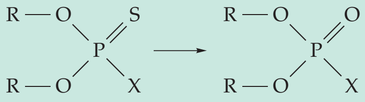
Malathion



Carbaryl

The characteristic high group transfer potential of a phospho group in pyrophosphate linkage, which makes ATP so useful in cells, also permits tetraethyl pyrophosphate to phosphorylate active sites of acetylcholinesterases. While TEPP is very toxic, it is rapidly hydrolyzed; all harmful residues are gone within a few hours after use.

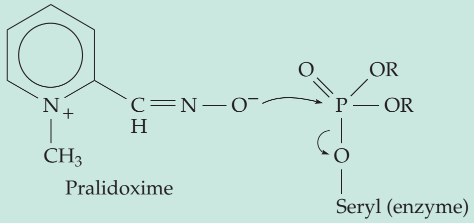
Two insecticides that have been used widely are **parathion** and **malathion**. They are less toxic than DFP or TEPP and do not become effective insecticides until they undergo bioactivation during which conversion from a P = S to a P = O compound occurs:



Highly toxic

The desulfuration reaction involves microsomal oxidases of the liver, the sulfur being oxidized ultimately to sulfate.[f] The reactivity of parathion with cholinesterases depends upon the high group transfer potential imparted by the presence of the excellent leaving group, the *p*-nitrophenolate anion. If the P– O linkage to this group is hydrolyzed before the desulfuration takes place, the phosphorus compound is rendered harmless. Thus, the design of an effective insecticide involves finding a compound which insects activate rapidly but which is quickly degraded by higher animals. Other factors, such as rate of penetration of the insect cuticle and rate of excretion from the organism, are also important.

The phosphorylated esterases formed by the action of organophosphorus inhibitors are very stable, but some antidotes can reverse the inhibition. The oxime of 2-formyl-1-methylpyridinium ion (pralidoxime) is very effective.[g] Its positive charge permits it to bind to the site normally occupied by the quarternary nitrogen of acetylcholine and to displace the dialkylphospho group:



Carbaryl, a widely used methyl carbamate, is a pseudosubstrate of acetylcholinesterase that reacts 10^5 to 10^6 times more slowly than do normal substrates. The carbamoylated enzyme formed is not as stable as the phosphorylated enzymes and the inhibition is reversible.

A basic problem is that most insecticides are designed to attack the central nervous system of the insect, the system that depends heavily upon acetylcholine. However, in human beings the readily accessible peripheral nervous system also depends upon acetylcholine, e.g., in neuronuscular junctions. The danger of poisoning is great. Another approach is to attack the glutamate-dependent peripheral system in insects, e.g., with inhibitors of glutamate decarboxylase. Glutamate functions as a neurotransmitter in the human body but only in the well-

BOX 12-E (continued)

protected central nervous system.

Another important problem is the development of insects resistant to insecticides. This often arises as a result of increased levels of carboxylesterases which hydrolyze both organophosphates and carbaryl.[h,i] A mutation that changed a single active site glycine to aspartate in a carboxylesterase of a blowfly changed the esterase to an organophosphorus hydrolase which protected the fly against insecticides.[j]

---

[a] Büchel, K. H., ed. (1983) *Chemistry of Pesticides*, Wiley, New York

[b] Hassall, K. A. (1990) *The Biochemistry and Uses of Pesticides*, 2nd ed., VCH Publ., Weinheim

[c] Kamrin, M. A., ed. (1997) *Pesticide Profiles*, CRC Press, Boca Raton, Florida

[d] Casida, J. E. (1973) *Ann. Rev. Biochem.* **42**, 259–278

[e] Wilkinson, C. F., ed. (1976) *Insect Biochemistry and Physiology*, Plenum, New York

[f] Nakatsugawa, T., Tolman, N. M., and Dahm, P. A. (1969) *Biochem. Pharmacol.* **18**, 1103–1114

[g] Wilson, I. B., and Ginsburg, S. (1955) *Biochim. Biophys. Acta.* **18**, 168–170

[h] Raymond, M., Callaghan, A., Fort, P., and Pasteur, N. (1991) *Nature (London)* **350**, 151–153

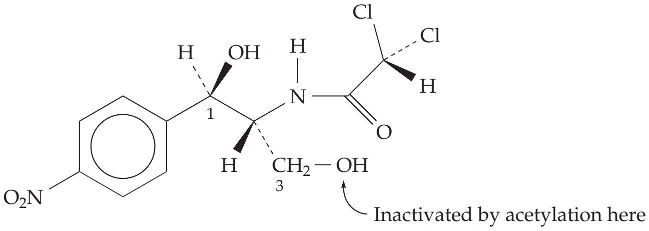
[i] Karunaratne, S. H. P. P., Hemingway, J., Jayawardena, K. G. I., Dassanayaka, V., and Vaughan, A. (1995) *J. Biol. Chem.* **270**, 31124–31128

[j] Newcomb, R. D., Campbell, P. M., Ollis, D. L., Cheah, E., Russell, R. J., and Oakeshott, J. G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7464–7468

substrate-binding cavity of the protein is designed to accommodate phospholipid molecules in the preferred conformation found in the micelles.[643] Most of the other lipases have lids which close over the active sites and impede access of substrates. Binding to a phospholipid surface apparently induces a conformational change that opens the lid. This allows substrate to enter from the lipid surface.[620,621,646] Cutinases, which do not display interfacial activation, do not have lids. Phospholipase A_2 activity is stimulated by an applied electrical field, a result that suggests that its activity *in vivo* may be regulated in part by the membrane potential.[648]

11. Other Acyltransferases

Acyl groups are frequently transferred from amides or esters to various acceptors in biosynthetic reactions. Among the many known acyltransferases are the ribosomal **peptidyl transferases** (Chapter 29), a transacylase involved in bacterial peptidoglycan synthesis (Chapter 20), transglutaminase (Eq. 2-23),[649,650] γ-glutamylcyclotransferase (Box 11-B), and transacylation reactions involving acyl-CoA derivatives. Examples of the latter are *N*-acetylation[651] or myristoylation (Chapter 8) of proteins, the formation of acetylcholine from choline[652] and of acetylcarnitine from carnitine (Eq. 17-4), and acetylation of the antibiotic chloramphenicol. The high group transfer potential of thioesters ensures that these reactions proceed to completion. **Chloramphenicol acetyltransferase** (**CAT**),[653] the enzyme that catalyzes acetylation and inactivation of the antibiotic by bacteria, is much used in studies of gene expression (Chapter 28). A catalytic histidine removes the proton from the 3-OH group and a serine hydroxyl provides an oxyanion hole to accommodate the anticipated tetrahedral intermediate.[652,654] The steroidal antibiotic fusidic acid (Chapter 22, Section G) is a competitive inhibitor.[655] A related transferase is aspartate carbamoyltransferase (Fig. 7-20; Chapter 24).



Chloramphenicol, a broad-spectrum *Streptomyces* antibiotic

Penicillin and related antibiotics are inactivated by β-lactamases (Box 20-G), some of which resemble serine proteases in forming acyl enzymes with active site serine side chains.[656,657] Others are zinc metalloenzymes.[658,659] **Amidohydrolases** such as **asparaginase** and **glutaminase,**[660,661] deacetylases,[662] and many other hydrolases can also be described as acyltransferases.

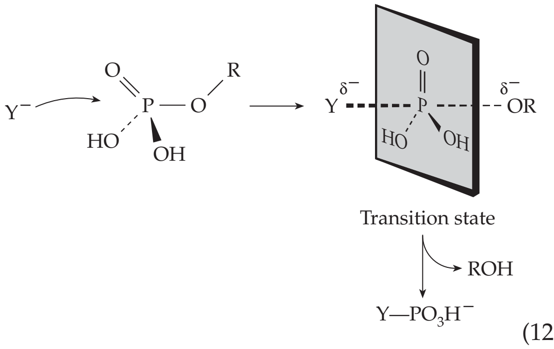
D. Displacement on a Phosphorus Atom

Nucleophilic displacements on phosphorus (Table 10-1, reaction type 1C) are involved in virtually every aspect of cellular energetics and in many aspects of biosynthesis. One large group of such enzymes are **phosphotransferases**, which transfer **phospho** (also called **phosphono** or, traditionally in biochemistry, **phosphoryl**) groups from one nucleophilic center to another. When transfer is to water the enzymes are called **phosphatases**, and when from one group in a molecule to another in the same molecule, **mutases**.

Enzymes that transfer a phospho group from ATP to water are **ATPases** and those that transfer the phospho group from ATP to some other nucleophile are **kinases**. Substituted phospho groups can also be transferred. Thus, **nucleases**, members of a larger class of **phosphodiesterases**, hydrolyze nucleic acids by transfer of nucleotidyl groups to a hydroxyl group of water. **Polynucleotide polymerases** transfer nucleotidyl groups to growing polynucleotide chains. An intramolecular nucleotidyl transferase is adenylate cyclase. **Topoisomerases** carry out a sequence of phosphotransferase reactions.

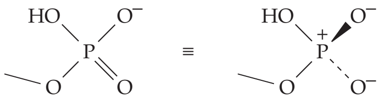
1. Questions about Mechanisms

Consider the following general equation for transfer to nucleophile Y^- of a phospho group attached in an ester or an anhydride linkage to form ROH, which could be an alcohol, carboxylic acid, or a phosphoric acid such as ATP.



(12-26)

This equation could also represent a half-reaction in a double-displacement process. As with displacements on saturated carbon atoms, two basic mechanisms can be imagined. The first is S_N2-like or **associative**.[663–667] The transition state might be represented where the bonds from the phosphorous atom to Y and to O are approximately equally formed. In an **in-line** displacement, where Y, P, and –OR are colinear, this mechanism leads to inversion if a chiral phospho group is used. There is another possibility for the associative mechanism. Whereas a carbon atom can form only four stable covalent bonds, phosphorus is able to form five. While nucleophilic attack on carbon leads to a *transient* five-bonded transition state (Eq. 12-2), attack on phosphorus could produce a relatively long-lived pentacovalent intermediate (Eq. 12-27). Notice that two transition states are involved. Remember that our conventional way of drawing phosphate esters with a double bond from phosphorus to one of the oxygens is misleading. All of the P–O bonds share some of the double-bond character and the phosphate group has many characteristics of a completely single-bonded structure:

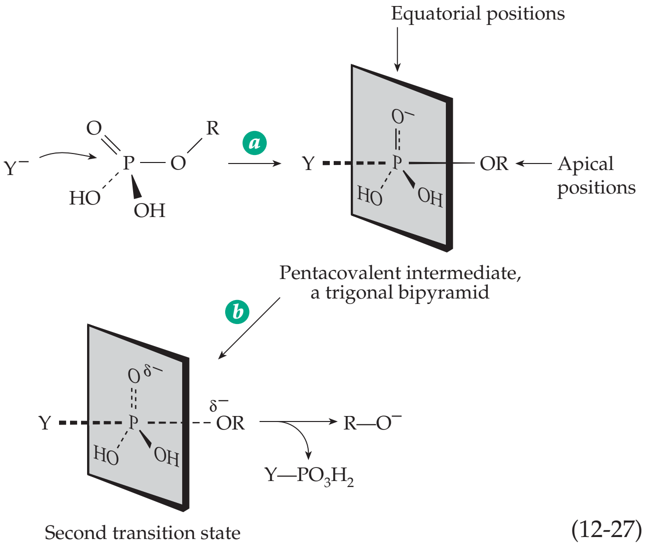


Likewise, in the transition state structures of Eq. 12-27, step *b,* the P–O bonds, except for that to –OR, are equivalent and all have partial double-bond character.

Geometric complexities. The geometry of the pentacovalent intermediate in Eq. 12-27 is that of a **trigonal bipyramid**. In this structure the bond angles in the **equatorial** plane are 120°, whereas all of the angles between any of those bonds and the two that attach to the groups in the **apical** positions are 90°. This disparity arises naturally from the fact that it is impossible to place five points on the surface of a sphere all equidistant one from the other. The attack of Y^- from the side opposite O–R (an in-line attack) leads to a trigonal bipyramid in which –O–R and –Y occupy the two apical positions. However, if Y^- attacks a face opposite one of the other oxygens (**adjacent** attack), –O–R will take an equatorial position.

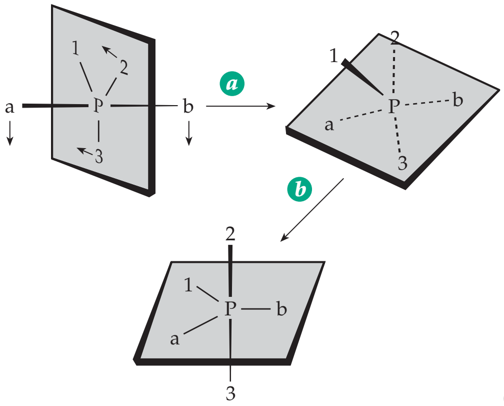
The chemical reactivities of groups in the apical and equatorial positions of pentacovalent intermediates are different.[664] In particular, elimination of a nucleophilic group to form a tetrahedral phosphate is easier from an apical position than from an equatorial position. For the in-line displacement of Eq. 12-27 elimination of RO^- should be easy. However an adjacent attack would leave –OR in an equatorial position. Before it could be eliminated, the intermediate would probably have to undergo a **permutational rearrangement** by which –OR would be transferred from an equatorial to an apical position.

One type of permutational rearrangement, known as **pseudorotation**, can be visualized as in Eq. 12-28.[668,669] The axial groups a and b move back while equatorial groups 2 and 3 move forward, still in the same equato-



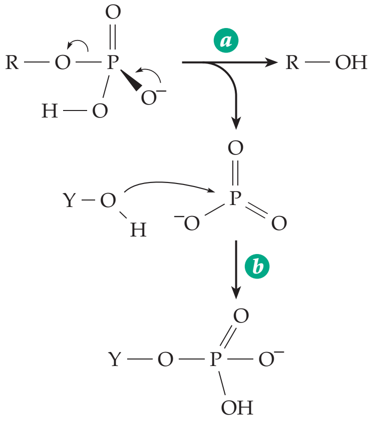
(12-27)

rial plane. Equatorial group 1 does not move. This decreases the 120° bond angles between the equatorial groups and increases the bond angles between group 1 and the axial groups until all four bond angles to group 1 are equivalent. The resulting **square pyramid** is a high-energy transition state structure in the pseudorotation process and can either revert to the original structure or, by continued motion of the groups in the same directions as before, to the structures shown at the bottom of Eq. 12-28. In the final structure, groups 2 and 3 are axial and the original axial groups a and b are equatorial. Pseudorotation is slow enough that it could be rate limiting in enzymatic reactions. Stereoelectronic effects could also affect these rates.[670,671] However, much evidence indicates that enzymes almost always avoid these complexities by using in-line mechanisms.



(12-28)

Metaphosphate ions. An alternative to an associative mechanism is an S_N1-like or **dissociative** mechanism which can occur by elimination of a **metaphosphate ion** (Eq. 12-29, step *a*). A nucleophilic reagent can then add to the eliminated metaphosphate in step *b*. The formation of metaphosphate is analogous to formation of carbocation during the action of lysozyme (Eq. 12-11). This dissociative mechanism could lead either to racemization or to inversion of a chiral phosphate. Metaphosphate ions have been shown to exist. They are generated in certain nonenzymatic elimination reactions in aprotic solvents[672–674] and they are reactive electrophiles able to react as in Eq. 12-29. However, there is some doubt that they can exist free in aqueous media.[675] Jencks and associates concluded from studies of linear Gibbs energy relationships that in the transition state for nonenzymatic phospho transfer reactions there is a large amount of bond breaking and a small amount of new bond formation in the transition state[676] but no free metaphosphate. There is a large dependence of the rate on the pK_a of the leaving group, but there is still a small dependence on the pK_a of the entering nucleophile. Thus, we have *a dissociative mechanism but without free metaphosphate.* The concept is supported by studies of kinetic isotope effects.[677,678]



(12-29)

Coping with negative charges. We visualize phospho group transfers as involving attack by a nucleophile bearing at least a partial negative charge. However, phospho groups also carry one or two negative charges and that in ATP even more. Therefore, it does not seem surprising that many phosphotransferases are metalloenzymes, sometimes containing bound Zn^{2+} which neutralizes some of the negative charge. Furthermore, enzymes usually accept ATP as a substrate only when it is accompanied by a divalent metal ion,[679] usually Mg^{2+}. Another way in which enzymes deal with the negative charges on phospho groups is to have arginine side chains in appropriate positions to interact by forming strong ion paired-hydrogen bonds. It is often assumed that it is essential to neutralize the charge on the phospho group to avoid electrostatic repulsion from a partial or complete negative charge on the attacking nucleophilic center. However, with a dissociative mechanism and a metaphosphate-like intermediate the transition state may be reached without charge neutralization.[679,680] Interactions of Mg^{2+} with phosphorus-containing substrates, like those of fixed positive charges in the protein may also be essential for binding the substrate correctly to an enzyme.

2. Magnetic Resonance Studies

There have been many investigations of phosphotransferases by NMR and EPR methods.[681,682] One approach is to use paramagnetic ions such as Mn^{2+}, Cu^{2+}, or Cr^{3+} to induce nuclear relaxation in substrate and coenzyme molecules at active sites of enzymes. Flavin radicals and specifically introduced nitroxide spin labels can serve as well. Paramagnetic ions greatly increase the rate of magnetic relaxation of nearby nuclei (Chapter 3, Section I). Thus, small amounts of

Mn^{2+} in a sample lead to broadening of lines in ordinary ^1H, ^{13}C, and ^{31}P NMR spectra.

Useful information about enzymes can sometimes be obtained by observing effects of paramagnetic ions on the NMR signal of protons in the solvent water. The relaxation time of solvent protons is usually greater than 1 s. However, in the ion $Mn(H_2O)_6^{2+}$ the protons of the coordinated water molecules relax much more rapidly, both T_1 and T_2 being ~10^{-1} s. Since the coordinated water molecules usually exchange rapidly with the bulk solvent, a small number of manganese ions can cause a significant increase in relaxation rate for all of the water protons. Broadening of the proton band is observed and differences in T_1 and T_2 can be measured by appropriate methods. Paramagnetic relaxation effects usually increase as the inverse sixth power of the internuclear distance. Knowing the Mn^{2+}–H distance to be 0.287 ± 0.005 nm for hydrated Mn^{2+}, it has been possible to relate the effects on T_1 and T_2 to the number of H_2O molecules coordinated at any one time to a protein-bound metal ion and to their rate of exchange with the bulk solvent.

Relaxation effects on ^1H, ^{13}C, and ^{31}P, while more difficult to observe, can provide geometric information about active sites. The theory is complex, but under some conditions the paramagnetically induced relaxation can be described adequately by Eq. 12-30.

$$r = C\,[T_{1M} f(\tau_c)]^{1/6} \qquad (12\text{-}30)$$

Here r is the internuclear distance, C is a combination of physical constants, and T_{1M} is the longitudinal relaxation time. The complex function $f(\tau_c)$ depends upon the correlation time τ_c, the resonance frequency of the nucleus being observed, and the frequency of precession of the electron spins at the paramagnetic centers. The value of τ_c can be estimated (Chapter 3) and, in turn, the distance r according to Eq. 12-30.

Such studies on creatine kinase (Eq. 12-31) utilized both a bound Mn^{2+} ion and a nitroxide spin label to estimate distances of various protons from the nitroxide.[683] Together with EPR measurements (Box 8-C), which gave the Mn^{2+}–nitroxide distance, a model of the ATP•Mn^{2+} complex in the active site was constructed. Additional EPR experiments on Mn^{2+} complexes with ATP and ADP containing ^{17}O in the α, β, or γ phospho groups showed that in the enzyme•ATP• creatine complex the metal ion is bound to all three phospho groups of ATP. It remained coordinated with the two phospho groups of ADP and also that of the phosphocreatine product in the enzyme•ADP•creatine-*P* complex as well as in the transition state, which is pictured occurring via a metaphosphate ion.[684]

$$\text{Creatine-}P + \text{ADP} \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \text{ATP} + \text{creatine} \qquad (12\text{-}31)$$

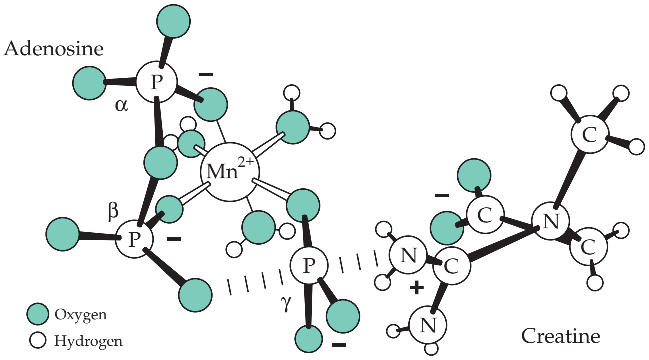


Figure 12-19 Proposed transition state structure formed from Mn^{2+}, ATP, and creatine bound in the active site of muscle creatine kinase. Based on EPR spectroscopy with regiospecifically ^{17}O-labeled substrates. The electrical charges have been added in one possible constellation. However, hydrogen atoms bound to phospho groups are not shown. After Leyh *et al.*[684]

Phosphorus-31 NMR has been used to measure **internal equilibrium constants** within enzyme-substrate (ES) complexes.[663,685–687] By having both substrate and product concentrations high enough to saturate the enzyme, all of the enzyme exists as ES and enzyme–product (EP) complexes in equilibrium with each other. For a phosphotransferase at least one substrate and one product contain phosphorus. Although the NMR resonances are broadened by binding to the large, slowly tumbling protein, their areas can be measured satisfactorily and can be used to calculate an equilibrium constant such as that for Eq. 12-32:

$$\text{ES} \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \text{EP}$$

$$K' = [\text{EP}]/[\text{ES}] \qquad (12\text{-}32)$$

An example is illustrated in Fig. 12-20. In this experiment[685] the relative areas of the ^{31}P signals of ADP (one for free ADP and one, slightly more intense, for MgADP) and of the signal for phosphoenolpyruvate (PEP) were measured in the absence of enzyme and in the presence of a catalytic amount of pyruvate kinase (Fig. 12-20A). The results verified that the equilibrium constant for the overall reaction (Eq. 12-33) is very high (3300).

$$\text{Phosphoenolpyruvate (PEP)} + \text{ADP} \rightarrow \text{ATP} + \text{pyruvate} \qquad (12\text{-}33)$$

However, with an excess of enzyme (Fig. 12-20B) the internal constant was estimated as

$$K' = \frac{[\text{E} \bullet \text{MgATP} \bullet \text{pyruvate}]}{[\text{E} \bullet \text{MgADP} \bullet \text{PEP}]} = 0.5 - 1.0 \tag{12-34}$$

This is a difficult measurement and reinvestigation by another group[686] indicated that the amount of the PEP-containing complex had been overestimated and that K'=10.

If the rates of the forward and backward reactions in Eq. 12-34 are of the same order as the spin-lattice relaxation times (T_1) of the ^{31}P in the bound substrate and product, the rate constants k_f and k_r can be evaluated by **saturation-transfer NMR**.[688] This is done by irradiating one resonance, e.g., that of the γ-P of ATP, and observing whether this causes a loss of intensity of the resonance for the product which is receiving its phospho group from ATP. This technique was used to observe the creatine kinase reaction (Eq. 12-31) in living muscle in both relaxed and contracting states. For resting muscle the observed forward flux was 1.7 x 10^{-3} M s^{-1} and the backward flux 1.2 x 10^{-3} M s^{-1}. Thus, this reaction, which supplies ATP for contraction of the muscle from stored phosphocreatine, appears to be operating at or near equilibrium. This had been assumed but had previously been difficult to prove. Two-dimensional NMR techniques can now be used for this kind of measurement.[689]

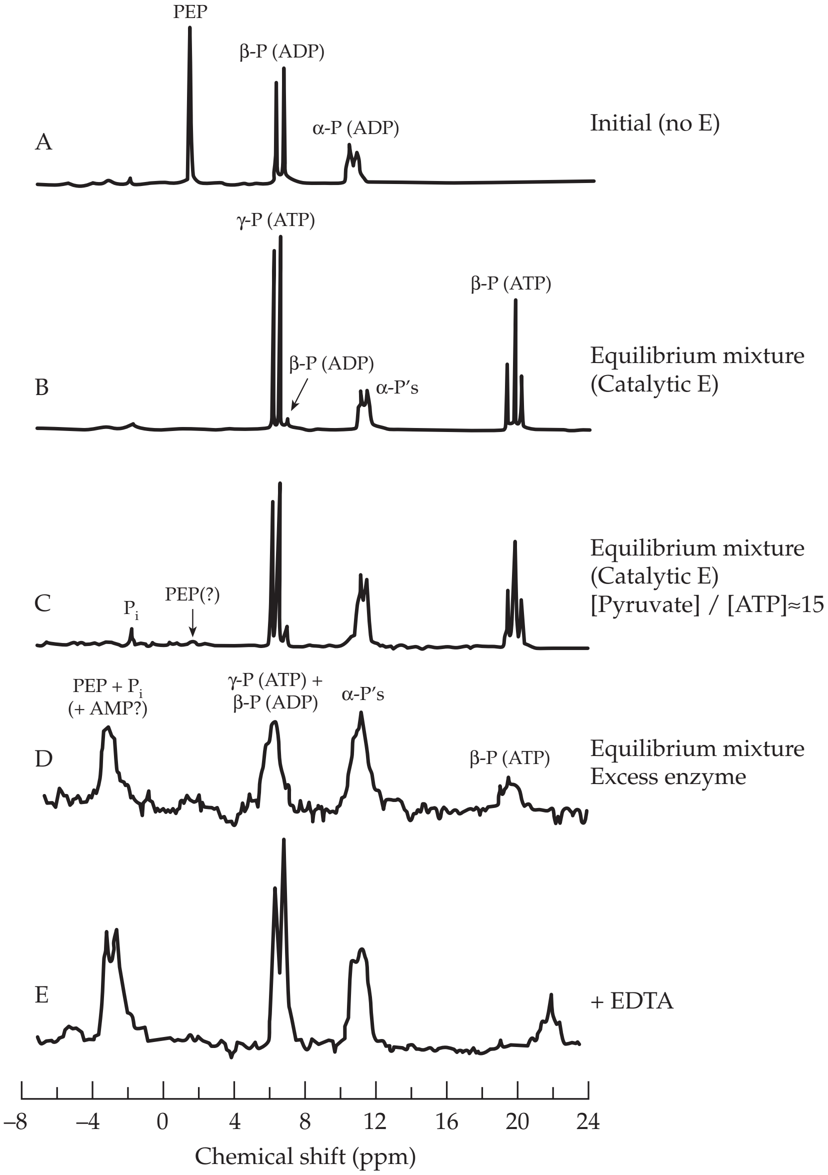
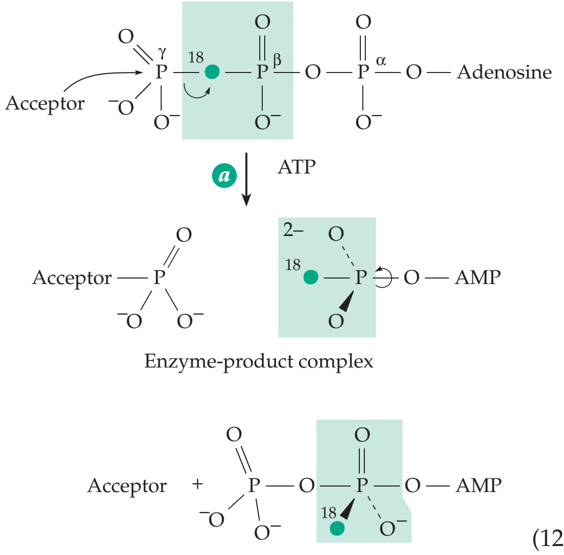


Figure 12-20 Equilibria in pyruvate kinase reaction as studied by ^{31}P NMR at 40.3 MHz, pH 8.0, 15°C. (A–C) Equilibria with low enzyme in levels ~15% ^2H_2O. (A) ^{31}P NMR spectrum of 1.5 ml of reaction mixture; PEP, 13.3 mM; ADP, 14.1 mM; $MgCl_2$, 20 mM; potassium Hepes buffer, 100 mM; KC, 50 mM without enzyme. (B) Equilibrium mixture after the addition of ~1 mg of pyruvate kinase to the reaction mixture. (C) Equilibrium after the addition of potassium pyruvate (final concentration of 200 mM) to the sample of the spectrum in (B). (D,E) Equilibrium with enzyme concentrations in excess of the substrates. Sample volumes ~1.1 ml with 10% ^2H_2O. (D) Equilibrium mixture set up with enzyme (2.8 mM active sites); 2.8 mM PEP; 2.4 mM ADP; 5.7 mM $MgCl_2$; 100 mM potassium Hepes; 100 mM KCl. (E) Spectrum after the addition of 50 µl of 400 mM EDTA (pH readjusted to 8.0) to the sample of spectrum D. The EDTA removes metal ions, stopping the catalytic reactions and sharpening the resonances. From Nageswara Rao *et al.*[685]

When ^{31}P is bonded to ^{18}O the chemical shift of the ^{31}P is altered by 0.0206 ppm from that when the phosphorus is bonded to ^{16}O. This allows ^{18}O labels introduced into phospho groups to serve as tracers which can be followed continuously during reactions.[683] The technique is useful in studies of stereochemistry (see Section 2) and for examination of **positional isotope exchange**.[690] This latter technique is often used with ATP containing ^{18}O in the β,γ-bridge position. If an enzyme transfers the terminal (γ) phospho group to an acceptor via a phosphoenzyme but without loss of the ADP, we may expect positional isomerization. The ^{18}O will move between the β,γ-bridge position and a nonbonding position as the phospho group is repeatedly transferred back and forth between ATP and the acceptor and as the phospho group rotates.[682,690]

Equation 12-35 shows one part of this isomerization. The negative oxygens have been omitted here to avoid implying a known state of protonation or a localization of charge.

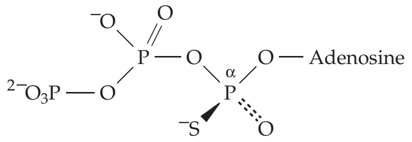


(12-35)

Figure 12-21 illustrates the use of the technique in investigating the possible participation of metaphosphate in a nonenzymatic reaction.

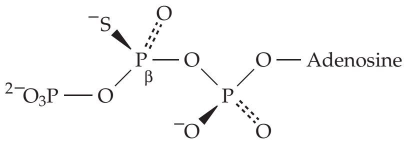
3. Stereochemistry

Evidence for an in-line S_N2-like mechanism of most enzymatic phospho group transfer reactions comes largely from study of chiral phospho groups.[663,692–695] A chiral phosphate can be introduced at either the α or β phosphorus of ATP by substitution of one of the oxygen atoms by sulfur. A chiral phospho group in the β position can be formed by substituting one oxygen by S and a second by ^{18}O.

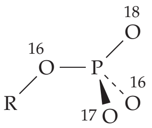


ATP αS (*R* configuration at P)

Notice that the negative charge is largely localized on sulfur in these phosphorothioate compounds.[696] More general is the use of ^{17}O and ^{18}O to form a chiral phospho group:



ATP βS (*S* configuration at P)



An ester chiral at the phosphorus atom; *R* configuration

Considerable ingenuity was required in both the synthesis of these chiral compounds[695,697] and the stereochemical analysis of the products formed from them by enzymes.[698–700] In one experiment the phospho group was transferred from chiral phenyl phosphate to a diol acceptor using *E. coli* alkaline phosphatase as a catalyst (Eq. 12-36). In this reaction transfer of the phospho group occurred without inversion. The chirality of the product was determined as follows. It was cyclized by a nonenzymatic in-line displacement to give equimolar ratios of three isomeric cyclic diesters. These were methylated with diazomethane to a mixture of three pairs of diastereoisomers triesters. These diastereoisomers were separated and the chirality was determined by a sophisticated mass spectrometric analysis.[692] A simpler analysis employs ^{31}P NMR spectroscopy and is illustrated in Fig. 12-22. Since alkaline phosphatase is relatively nonspecific, most phosphate esters produced by the action of phosphotransferases can have their phospho groups transferred without inversion to 1,2-propanediol and the chirality can be determined by this method.

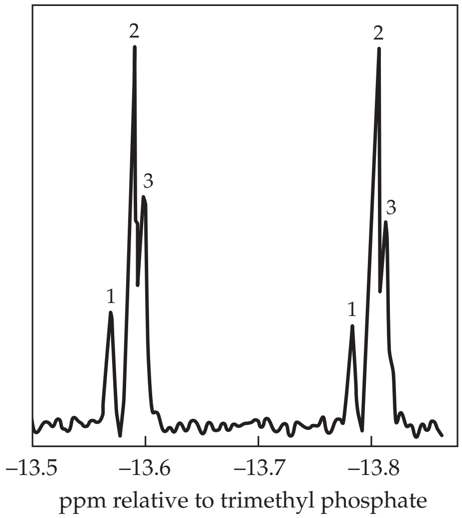
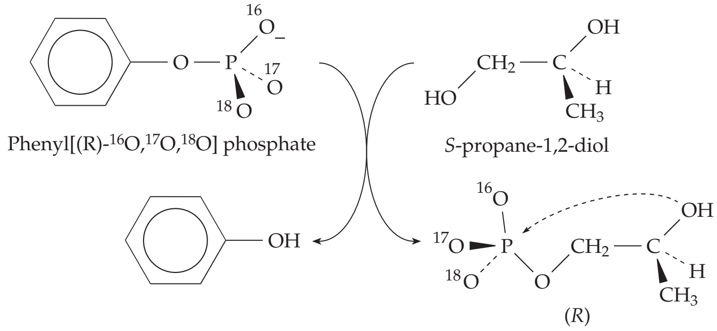


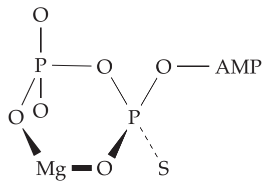
Figure 12-21 The ^{31}P NMR spectrum at 101.2 MHz of P_α of isotopically labeled ADP. This was recovered from an experiment in which ADP containing 87 atom % of ^{18}O in all four oxygens around P_β was allowed to undergo partial (~20%) nonenzymatic hydrolysis to AMP and P_i. Peaks 1 represent the species containing no ^{18}O bonded to P_α. Peaks 2 represent all species with ^{18}O in the P_α–O–P_β bridge, and peaks 3 represent species with ^{18}O in nonbridging positions at P_α. These have undergone positional isotope exchange. From Lowe and Tuck.[691]



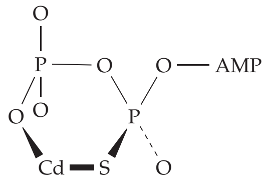
(12-36)

Although inversion was not observed with the *E. coli* alkaline phosphatase, it has been observed for ribonucleases and many other hydrolytic enzymes and for most kinases transferring phospho groups from ATP. The difference lies in the existence of a phosphoenzyme intermediate in the action of alkaline phosphatase (see Eq. 12-38). Each of the two phosphotransferase steps in the phosphatase action apparently occurs with inversion. The simplest interpretation of all the experimental results is that *phosphotransferases usually act by in-line S_N2-like mechanisms which may involve metaphosphate-ion-like transition states that are constrained to react with an incoming nucleophile to give inversion*. An adjacent attack with pseudorotation would probably retain the original configuration and is therefore excluded.

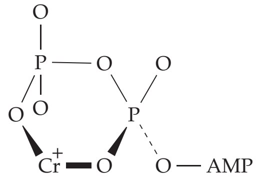
The substrate for many phosphotransferases is MgATP. Which of the possible isomers of this chelate complex is utilized by these enzymes? Since Mg^{2+} associates and dissociates rapidly from the complexes there are several possibilities: a tridentate complex with oxygens from α, β, and γ phospho groups coordinated with the metal ion, an α,β-bidentate, a β,γ-bidentate, or a monodentate complex. Most evidence suggests that β,γ-bidentate complexes of the following types are the true substrates.



Mg–ATP βS complex
with Λ screw sense,
R configuration at P



Cd–AT βS complex
with Λ screw sense,
S configuration at P

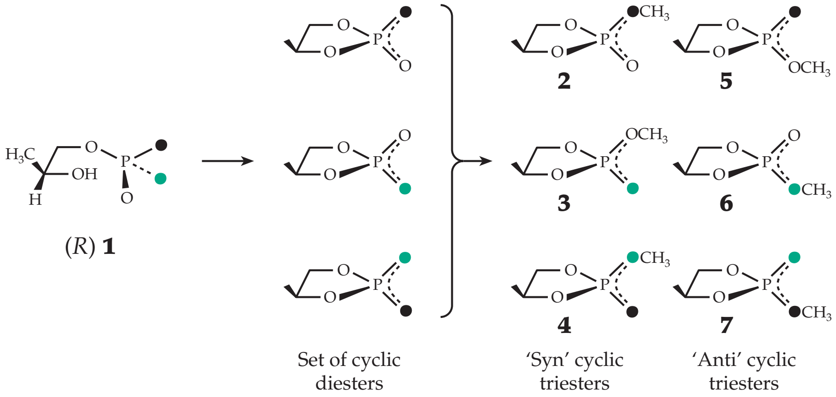


Cr–ATP complex
with Δ screw sense

The first structure drawn is for a Mg–ATP *S* complex with a chiral β-phosphothioate group. The Mg^{2+} is expected to bond to oxygen. However, in the second complex, in which Cd^{2+} has been substituted for Mg^{2+}, it is expected that the Cd^{2+} will bond to sulfur. Therefore, the stereochemically equivalent structure will be obtained only when the compound has the *S* configuration at the phosphorus. In the third of the foregoing complexes, Mg^{2+} has been replaced by Cr^{3+} to give an **exchange-inert** complex of ATP in which the Cr^{3+} will remain attached firmly to the same two oxygen atoms under most experimental conditions. Notice also that in this complex, the AMP portion occupies a different position than in the first two complexes. It is called the Δ **screw-sense isomer**,[702] while the upper two complexes are Λ screw-sense isomers. We may reasonably expect that an enzyme, which may provide additional ligands to the metal ions, will prefer one or the other of these screw-sense isomers. For α,β-bidentate ATP complexes both α and β phosphorous atoms become chiral centers and even more isomers are possible.

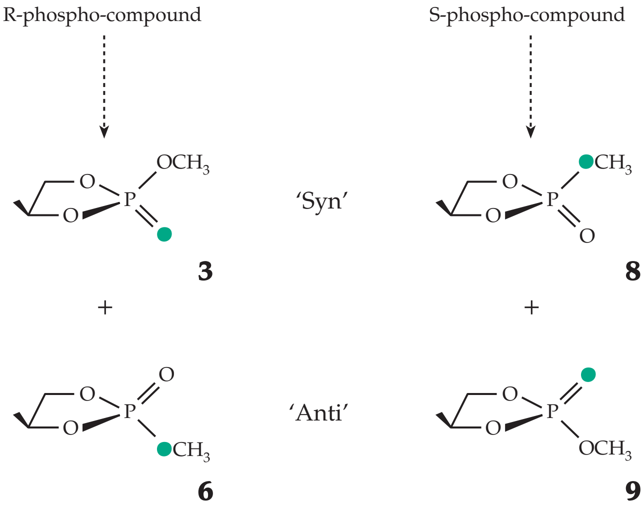
The use of exchange inert Cr^{3+} and Co^{3+} complexes of ATP has been developed by Cleland and associates.[702–705] The β,γ-bidentate chromium complexes were separated into the Λ and Δ isomers.[704] Each was separated further into a pair of "ring-puckering isomers". These metal complexes are all competitive inhibitors of MgATP and both ring puckering isomers of the Δ screw sense are very slow substrates of various kinases. The Λ isomers serve as very slow substrates for pyruvate kinase, adenylate kinase, and fructose-6-phosphate kinase.[703] The Δ exo isomer is the strongest inhibitor of creatine kinase, suggesting this same conformation for the MgATP substrate. The disastereoisomers of ATPαS and ATPβS have also been tested with kinases in the presence of either Mg^{2+} or Cd^{2+}. With creatine kinase[695,706] the isomers with the R configuration at phosphorus are preferred in the presence of Mg^{2+} but those with the *S* configuration are preferred in the presence of Cd^{2+}. This might be related to the previously mentioned preferences of the metals for O vs. S ligands, and would suggest that this enzyme prefers the Λ isomer (see foregoing structures). However, EPR studies with the corresponding Mn^{2+} complexes suggest that the Δ isomer is preferred by the enzyme.[693] There are 12 isomers of monoammine Cr(III)ATP. Their use has provided information about the location of water molecules in metal complexes of kinases.[707]

In the following sections we will consider several individual phosphotransferases.

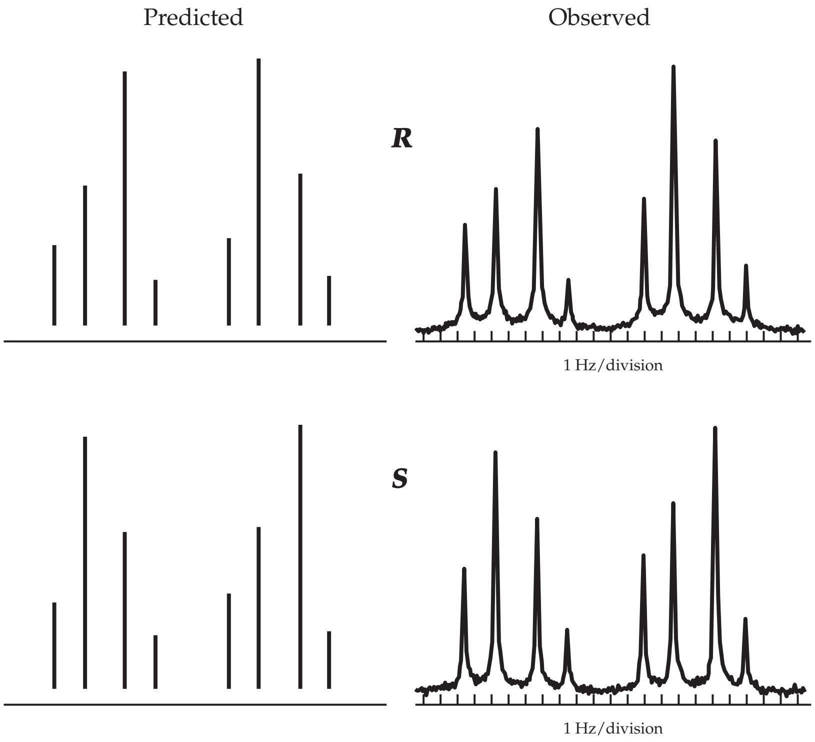

H$_3$C
OH
H
O
P
(R) 1
2
3
4
5
6
7
OCH$_3$
CH$_3$
Set of cyclic diesters
'Syn' cyclic triesters
'Anti' cyclic triesters

The product 1-phospho-(*S*)-propane-1,2-diol **1** (here shown as *R* at phosphorus) is converted by in-line ring closure to an equimolar mixture of three cyclic diesters. These are methylated to give six cyclic triesters. Of these, only **3** and **6** give sharp ^{31}P resonance because the ^{17}O in the others broadens the lines.

O = ^{16}O
● = ^{17}O
🟢 = ^{18}O


R-phospho-compound
S-phospho-compound
OCH$_3$
CH$_3$
O
P
3
8
+
'Syn'
'Anti'
6
9

Two cyclic triesters **3** and **6** from an *R*-phospho compound and two others **8** and **9** from an *S*-phospho compound will give sharp ^{31}P NMR resonances.

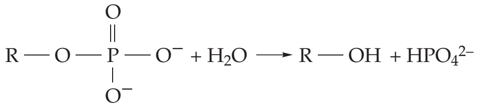


Predicted and observed ^{31}P NMR spectra of the mixtures of syn and anti cyclic triesters derived from labeled samples of 1 phospho-(*S*)-propane-1,2-diols that are *R* and *S* at phosphorus.

Figure 12-22 Method for determining chirality of phospho groups containing ^{16}O, ^{17}O, and ^{18}O. From Buchwald and Knowles.[701]

4. Phosphatases

The phosphatases catalyze hydrolysis of phosphate esters to produce inorganic phosphate:[667]

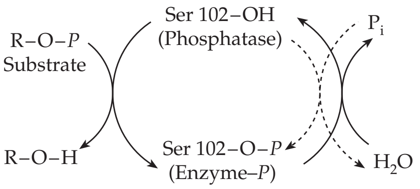


$$R—O—\overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle O^-}{|}}{P}}—O^- + H_2O \longrightarrow R—OH + HPO_4^{2-} \qquad (12\text{-}37)$$

The **acid phosphatases** and **alkaline phosphatases** are nonspecific and cleave many different phosphate esters, whereas **glucose-6-phosphatase**, **fructose-1,6-bisphosphatase**, and many others are specific for single substrates. The nonspecific phosphatases may provide inorganic phosphate ions in places where they are needed, e.g., in mineralizing bone. All phosphatases help to drive metabolic cycles (Chapter 17).

The alkaline phosphatases are found in bacteria, fungi, and higher animals but not in higher plants. In *E. coli* alkaline phosphatase is concentrated in the periplasmic space. In animals it is found in the brush border of kidney cells, in cells of the intestinal mucosa, and in the osteocytes and osteoblasts of bone. It is almost absent from red blood cells, muscle, and other tissues which are not involved extensively in transport of nutrients.

The alkaline phosphatase of *E. coli* is a dimer of 449-residue subunits which requires Zn^{2+}, is allosterically activated by Mg^{2+}, and has a pH optimum above 8.[667,708–711] At a pH of ~4, incubation of the enzyme with inorganic phosphate leads to formation of a phosphoenzyme. Using ^{32}P-labeled phosphate, it was established that the phosphate becomes attached in ester linkages to serine 102. The same active site sequence Asp-Ser-Ala is found in mammalian alkaline phosphatases. These results, as well as the stereochemical arguments given in Section 2, suggest a double-displacement mechanism of Eq. 12-38:



(12-38)

The active site contains two Zn^{2+} ions and one Mg^{2+} ion which are held by imidazole and carboxylate groups. The inorganic phosphate in an enzyme–product complex is bound to both zinc ions (Fig. 12-23). The Ser 102 side chain is above one Zn. In the enzyme-*P* intermediate it would be linked to the phospho group as an ester which would then be hydrolyzed, reversibly, by a water molecule bound to Zn.[712–713a] This water presumably dissociates to Zn^+–OH and its bound hydroxyl ion carries out the displacement. This reaction may be preceded by a proton transfer to an oxygen atom of the phospho group.[714]

Acid phosphatases, which have pH optima of ~5 and are inhibited by fluoride ion, occur in bacteria, fungi, plants, and animals. In bone, the acid phosphatase content is high in the osteoclasts which function in the resorption of calcium from bone. The highest content of acid phosphatase in humans is in the prostate and an elevated serum level has long been used as a diagnostic indicator of prostatic cancer.[715] Acid phosphatase is a periplasmic enzyme in *E. coli*.[716] Phosphoenzymes have been trapped from both plant and animal acid phosphatases.[717] For example, a brief incubation with ^{32}P-labeled *p*-nitrophenyl phosphate followed by rapid denaturation in an alkaline medium gave a covalently labeled protein from which ^{32}P-containing N δ-phosphohistidine was isolated. In agreement with kinetic evidence that a phosphoenzyme is a true intermediate, there is no inversion of chiral phospho groups by liver or prostate acid phosphatases.[718]

Some acid phosphatases from animals and plants are violet in color and contain iron (Chapter 16) and an Mn^{3+}-containing acid phosphatase has been isolated from sweet potatoes.[720] These enzymes have dimetal centers, often containing one Zn^{2+} and one Fe^{3+} with bridging carboxylate and hydroxide ions between the metals. Imidazole, tyrosinate, and carboxylate side chains hold the metals as in Fig. 16-20. A water molecule bound to the Fe^{3+} is thought to dissociate with a low pK_a of 4.8 to give an $Fe^{3+} \cdot {}^-OH$ complex. The hydroxyl ion can then attack the phospho groups, one

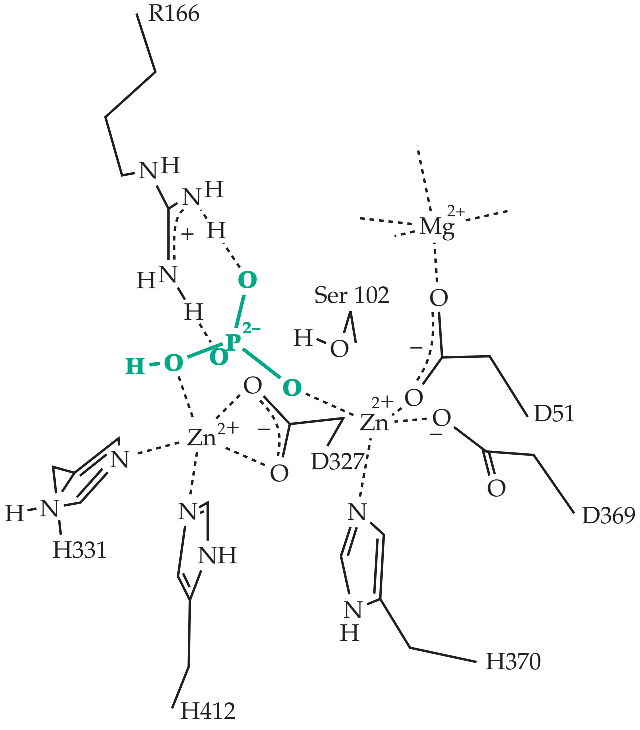


Figure 12-23 Schematic drawing of the product inorganic phosphate bound in the active site of *E. coli* alkaline phosphatase. See Ma and Kantrowitz.[719]

of whose oxygen atoms is coordinated with the Zn^{2+}.

The mechanism resembles that proposed for a **phosphotriesterase** (Fig. 12-24). The triesterase catalyzes detoxification of organophosphorus toxins such as parathion (Box 12-E) and seems to have evolved rapidly from a homologous protein of unknown function.[721] The phosphotriesterase contains two Zn^{2+} ions in a dimetal center. An unusual structural feature is a carbamate group, formed from Lys 169 and CO_2, which provides a bridging ligand for the metal pair.[721–725] A carbamylated lysine also functions in ribulose bisphosphate carboxylase (Fig. 13-11).

Phosphatases specific for such substrates as glucose-6-phosphate, fructose-1,6-bisphosphate, and phosphoglycolate help to drive metabolic cycles (Chapter 17). The 335-residue **fructose-1,6-bisphosphatase** associates to form a tetramer with D_2 symmetry.[726–730] The allosteric enzyme exists in two conformational states (see Chapter 11). Activity is dependent upon Mg^{2+} or other suitable divalent cation, e.g., Mn^{2+} or Zn^{2+}, and is further enhanced by K^+ or NH_3^+. While the dimetal sites depicted in Figs. 12-23 and 12-24 are quite rigid and undergo little change upon formation of complexes with substrates or products, the active site of fructose-1,6-bisphosphatase is more flexible. There are three metal-binding sites but they contain no histidine side chains and have been seen clearly only in a product complex.[727,728] Perhaps because of the need for flexibility involved in allosteric changes, the active site is not fully formed until the substrate binds. **Fructose-2,6-bisphosphatase** forms one domain of a bifunctional kinase-phosphatase (Chapter 11). It has two imidazole rings, as well as side chains from a glutamate and two arginine residues at the catalytic and substrate-binding site.[728a]

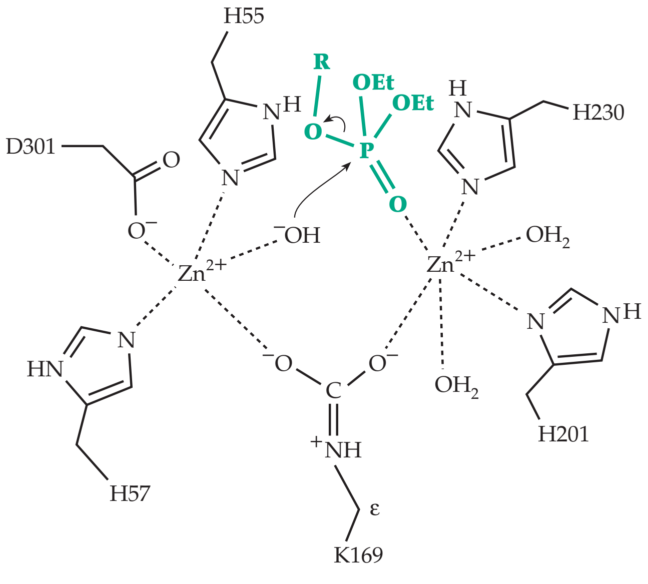


Figure 12-24 Hypothetical event in the action of a phosphotriesterase. A carbamylated lysine (lower center), as well as a water molecule, bridge the two Zn^{2+} ions, which are held by imidazole and aspartate carboxylate groups. The bound H_2O can be deprotonated to give the HO^- complex shown. The substrate may displace the HO^- ion from the right-hand zinc and thereby move close to the bound HO^- which attacks as indicated. Based on Cd^{2+}-containing structure and discussion by Benning *et al.*[722]

The 357-residue mammalian glucose-6-phosphatase plays an important role in metabolism (Chapter 17). Defects in the enzyme cause a glycogen storage disease (Box 20-D) and severe disruption of metabolism.[731] However, the molecular basis of its action is not well-known. Furthermore, the active site of the enzyme is located in the lumen of the endoplasmic reticulum[732] and glucose-6-phosphate must pass in through the plasma membrane. An additional glucose-6-phosphate transporter subunit may be required to allow the substrate to leave the cytoplasm.[73]

Pyrophosphatases, which are present in all cells, and catalyze hydrolysis of inorganic pyrophosphate (PP_1) to orthophosphate (P_i) (see Chapter 6, Section D), also drive metabolic sequences. The very active pyrophosphatase of *E. coli* has a turnover number of over 2×10^4 s^{-1} at 37°C. The 1000 molecules per cell are sufficient to immediately hydrolyze any pyrophosphate produced by bacterial metabolism.[733] The much studied soluble pyrophosphatases of *E. coli*,[734,735] yeast,[736] and other organisms[736a,b] are metalloenzymes that are most active with Mg^{2+}. Two Mg^{2+} ions are held, mostly by carboxylate side chains, while a third apparently enters the active site as magnesium pyrophosphate, perhaps $MgP_2O_7^-$. As with other metallohydrolases, a metal-bound hydroxyl ion may serve as the attacking nucleophile.

At least three distinct families of **protein phosphatases** remove phosphate groups from serine, threonine, and tyrosine side chains in proteins.[737] Their role in control of numerous biochemical processes has been discussed in Chapter 11, Section C,2. The catalytic domains or subunits of the protein phosphatases act together with regulatory domains or separate regulatory subunits to control thousands of reactions. For example, protein phosphatase 1 (PP1) together with a glycogen-targeting subunit dephosphorylates inactive glycogen kinase (see Fig. 11-4). Belonging to the same family is **calcineurin** (PP2B), a phosphatase activated by Ca^{2+} through binding to calmodulin (Box 6-D). There are two families of Ser/Thr phosphatases. Their polypeptide folding patterns differ, but the active sites have similar dimetal centers resembling those in Figs. 12-24 and 16-20 with $Mn^{2+} + Fe^{2+}$, $Zn^{2+} + Fe^{3+}$, and probably other pairs of metals.[737–740] The family containing PP1 has weak sequence homology with the purple acid phosphatases.[740] Another common feature of these enzymes is a conserved His-Asp dyad (His 125 and Asp 95 in PP1) which is thought to be a proton donor, protonating the leaving group ($-O^-$) in a manner reminiscent of the serine proteases. The reaction is evidently initiated by attack of an $^-$OH ion held by the dimetal center. This would resemble the mechanism pictured in Fig. 12-24 except that the phospho group would carry two negative charges.

The protein tyrosine phosphatases also exist as several families with numerous functions in control of transcription, growth, differentiation, and metabolism.[741–743] These enzymes function by a double-displacement mechanism, as in Eq. 12-38, but with a cysteine side chain rather than serine. The cysteine is present in the conserved sequence (H/V)**C**X_5**R**(S/T). The arginine binds the phospho group and helps to stabilize the transition state, which probably is metaphosphate-like.[742]

5. Ribonucleases (RNases)

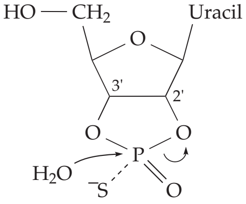
Many hydrolases act on phosphodiester linkages, which abound in nature.[725,744] Some are digestive enzymes but others serve more specific metabolic functions. **Ribonuclease A** (**RNase A**), the pancreatic digestive enzyme responsible for breakdown of RNA, was one of the first enzymes for which a structure was deduced. By 1963 Moore and Stein and their associates, who had earlier developed ion exchange methods for separating amino acids and peptides (Fig. 3-6), had determined the sequence of the 124-residue bovine enzyme.[745] They observed that Lys 41 was unusually reactive with dinitrofluorobenzene and that photooxidation of His 12 and His 119, which are almost at opposite ends of the peptide chain, inactivates the enzyme. They concluded that both histidines are at the active site, a conclusion that was later substantiated by X-ray crystallography.[746–748] A segment 12 nucleotides in length can fit into the cleft in the enzyme that contains the active site. The negatively charged phosphates of the RNA backbone form 8 – 9 electrostatic bonds to lysine and arginine side chains of the enzyme.[749,750] However, the only close interactions of the nucleic acid bases with the enzyme occur at the site of cleavage as shown in Fig. 12-25. The four residues His 12, Lys 41, Thr 45, and His 119 are strictly conserved in the RNAse A superfamily.[749] The carboxylate of Asp 121 apparently helps to orient the proper tautomer of His 119 for catalysis.[751]

Ribonuclease A was the first enzyme to be synthesized in the laboratory. Fully active ribonuclease has been synthesized,[752] as have new modified enzymes. For example a 63-residue peptide made up of five segments of the native RNase sequence retained measurable catalytic activity.[753] Using total synthesis, unnatural amino acids, such as 4-fluorohistidine, have been incorporated at specific positions in RNAse.[752]

Cleavage of a phosphodiester linkage in the substrate chain occurs in two steps. In the first or *transesterification* step, the hydroxyl group on the 2' position of the ribose ring is thought to be deprotonated by attack of either the imidazole of His 12, as shown in Fig. 12-25, or by the adjacent amino group of Lys 41. (In the latter case His 12 would have to remove a proton from the $-NH_3^+$ group of Lys 41 before it could attack.) In either case the positive charge of Lys 41 would help to neutralize the negative charge on the phosphate. The deprotonation of the 2'–OH may occur synchronously with its attack on the adjacent 3' phospho group. An in-line displacement of the oxygen attached to the 5' carbon of the next nucleotide unit is thought to be assisted by His 119. Its protonated imidazolium group may transfer a proton to a phosphate oxygen atom prior to or synchronously with formation of the new P–O bond in step *a* (Fig. 12-25).[754] The intermediate formed in step *a* is a cyclic 2', 3'-diphosphate which then undergoes hydrolysis by attack of a water molecule in step *b* to give the free nucleoside 3'-phosphate. The overall reaction is a two-step double-displacement, analogous to that with chymotrypsin, except that a neighboring group in the substrate rather than an amino acid side chain is the nucleophilic catalyst. The pH dependence of the enzyme is in agreement with this mechanism because there are two pK_a values of ~5.4 and ~6.4 which regulate the catalytic activity. Microscopic pK_a values of His 12 and His 119 have been measured by NMR spectroscopy as ~ 6.1 and ~ 6.3 and are shifted somewhat by binding of nucleotides.

A bacterial peptidase splits a 20-residue fragment containing His 12 from the N-terminal end of RNase A. This "S-peptide" can be recombined with the rest of the molecule, which is inactive, to give a functional enzyme called ribonuclease S. In a similar way, residues 119–124 of RNase A can be removed by digestion with carboxypeptidase to give an inactive protein which lacks His 119. In this case, a synthetic peptide with the sequence of residues 111 – 124 of RNase A forms a complex with the shortened enzyme restoring full activity.[755]

Stereochemical studies support in-line mechanisms for both the transesterification and hydrolysis steps of ribonuclease catalysis. For example, chiral uridine 2',3'-cyclic phosphorothioates are hydrolyzed with inversion of configuration, with the diastereoisomer shown yielding a 2'-monophosphothioate of the *R* configuration at phosphorus.



Uridine 2',3'-cyclic phosphorothioate

Transesterification step is also in-line[757] as it is shown in Fig. 12-25. Study of kinetic isotope effects in H_2O–D_2O mixtures suggested that two protons may move synchronously as the enzyme–substrate complex passes through the transition state.[758] Although RNase A is one of the most studied of all enzymes, there are still uncertainties about the mechanism. Is a proton removed first by His 12 (Fig. 12-25), as has long been assumed, or is a proton transferred first from His 119 to the oxygen of the phospho group?[759,760] Is the reaction concerted, as suggested by kinetic isotope effects,[761] or is there a pentacovalent intermediate?

The specificity of RNase A for a pyrimidine on the 3' side of the phosphodiester bond that is cleaved is evidently ensured by the pair of hydrogen bonds from O-2' of the pyrimidine to the backbone NH of Thr 45 and a second from the N-4' proton to the side chain OH of the same threonine (Fig. 12-25). Other nucleases, such as ribonuclease T_2,[762] with different specificities also make use of hydrogen bonding of the base at the 3' side of the cleavage point with backbone amide groupings.

Various bacterial ribonucleases as well as the fungal ribonucleases T_1, U_1, and U_2 (see also Fig. 5-43) have amino acid sequences related to that of RNase A[763,764,764a] but with distinctly different three-dimensional structures. The active sites contain Glu, His, and Arg side chains. For RNase T_1, Glu 58 and His 92 appear to provide acid–base catalysis with assistance from Tyr 38, Arg 77, and His 40.[763,765] A glutamate carboxylate also appears to be the catalytic base in the related RNase, called **barnase**, from *Bacillus amyloliquefaciens*.[766]

In addition to extracellular digestive enzymes, the RNase family contains many intracellular enzymes that are involved in turnover of RNA.[767,768] RNase H digests away RNA primers during DNA synthesis (Chapter 27). RNase H activity is also present in a domain of viral reverse transcriptases and is absolutely essential for the replication of HIV and other retroviruses.[769] The structures of the reverse transcriptase RNase H domain and of the *E. coli* enzyme are similar.[769–771] Unlike RNase A, the RNases H are metalloenzymes which apparently contain two Mg^{2+} ions held by carboxylate groups and utilize a metal bound HO^- ion as in previously discussed phosphatases. Secreted RNases sometimes have specific functions. For example, the 123-residue **angiogenin** is homologous to pancreatic RNases but acts to induce formation of new blood vessels (angiogenesis).[772–774] This is essential to growth of solid cancers as well as for normal growth. The enzyme is a very poor catalyst but its RNase activity appears essential for its biological function. Mutation of any of the catalytic residues His 13, Lys 40, or His 114 abolishes all angiogenic activity. A neurotoxin secreted by eosinophils[775] is one of a group of selectively toxic RNases.[776,777] Intracellular RNases are often found as complexes of specific inhibitor proteins.[778]

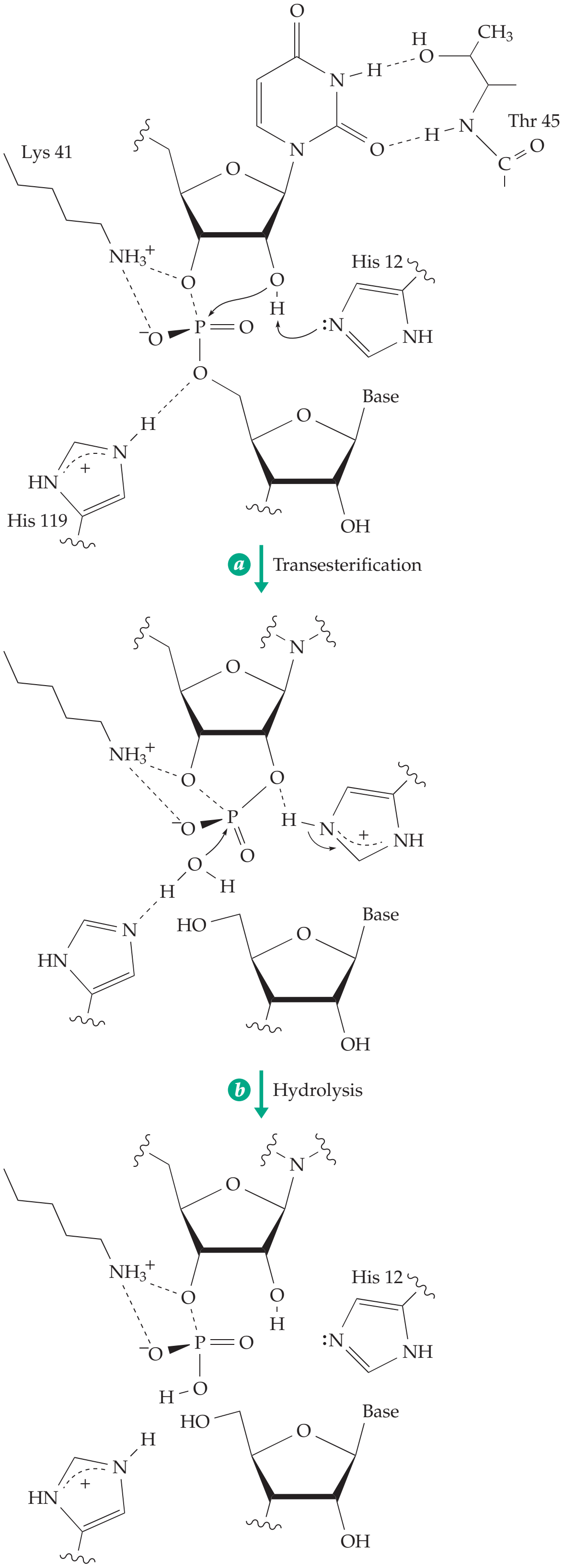


Figure 12-25 Proposed two-step in-line reaction mechanism for ribonuclease A. The hydrogen bonding that provides recognition of the pyrimidine base at the 3' end created by the cleavage is also shown. See Wladkowski *et al.*[756]

6. Ribonuclease P, Ribozymes, and Peptidyl Transferase

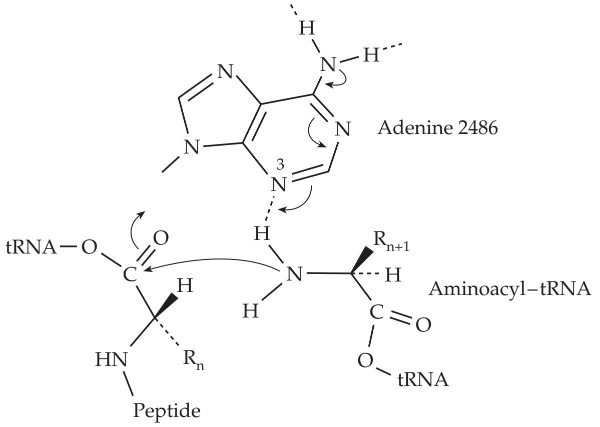
A very different ribonuclease participates in the biosynthesis of all of the transfer RNAs of *E. coli*. **Ribonuclease P** cuts a 5' leader sequence from precursor RNAs to form the final 5' termini of the tRNAs. Sidney Altman and coworkers in 1980 showed that the enzyme consists of a 13.7-kDa protein together with a specific 377-nucleotide RNA component (designated M1 RNA) that is about five times more massive than the protein.[779] Amazingly, the M1 RNA alone is able to catalyze the ribonuclease reaction with the proper substrate specificity.[780–782a] The protein apparently accelerates the reaction only about twofold for some substrates but much more for certain natural substrates. The catalytic center is in the RNA, which functions well only in a high salt concentration. A major role of the small protein subunit may be to provide counterions to screen the negative charges on the RNA and permit rapid binding of substrate and release of products.[783] Eukaryotes, as well as other prokaryotes, have enzymes similar to the *E. coli* RNase P. However, the eukaryotic enzymes require the protein part as well as the RNA for activity.[784]

Thomas Cech and associates independently discovered another class of catalytic RNA molecules. These are **self-splicing RNAs** that cut out intervening sequences from themselves to generate ribosomal RNA precursors (see Chapter 28).[785–787] They act only once and are therefore not enzymes. However, the introns that are cut out during self-splicing are **ribozymes** which, like the RNA from ribonuclease P, can act catalytically and have properties similar to those of protein enzymes. They exhibit the kinetic properties of enzymes and are denatured by heat. The RNAs are folded into compact structures resembling those of globular proteins. Like tRNA, they contain loops and hydrogen-bonded stems. Phylogenetic comparisons (Chapter 29) of the M1 RNA of ribonuclease P isolated from various species have allowed prediction of precise secondary structures.[783,788,789] A simplified M1 RNA consisting of 263 nucleotides from conserved regions of the molecule is catalytically effective.[783] Tetrahymena ribozyme also has a complex structure with a 247-nucleotide catalytic core formed by two structural domains (Fig. 12-26).[790,791] The crystal structure of a third ribozyme, one found in the RNA of the human pathogen hepatitis delta virus (HDV), has also been determined.[792] It is a smaller 72-nucleotide self-cleaved molecule with a very different structure from that in Fig. 12-26. It makes use of a double pseudoknot (see Fig. 5-29) to bind the RNA into a compact, tightly hydrogen-bonded structure with a deep active site cleft. It is the fastest known naturally occurring self-cleaving RNA and is able to react at a rate of more than 1 s^{-1} at its optimum temperature of 65°C.

Smaller self-cleaving RNAs have been found among plant viruses and viroids. Many of them have a common catalytic core which can be converted into 30- to 40-nucleotide ribozymes. Only 17 nucleotides and three hydrogen-bonded helical stems are required to form the self-cleaving "hammerhead" domain, which has a structural similarity to the catalytic core of the *Tetrahymena* ribozyme. The **hammerhead ribozymes** (Fig. 12-27) represent one form of small ribozyme.[793–797] Another is the **hairpin ribozyme** shown in Fig. 12-28,[798,801] which also shows the even smaller lead-dependent "**leadzyme**," a ribozyme that doesn't occur in nature.

In an intact viral self-cleaving RNA the entire catalytic center is formed from a single strand. Stems II and III of the hammerhead ribozyme (Fig. 12-27C) are closed by large loops. In the ribozyme shown, loop III has been cut off and stem II has been closed by a tight loop to form a compact catalytic RNA that will cut a substrate having a suitable nucleotide sequence for binding to the ribozyme. Only 12 bases in this ribozyme are highly conserved. By varying the sequences in the ribozyme half of stems I and III, catalysts that cleave after any sequence GUX, where X=A, C, or V, can be designed. Such catalysts are useful in the laboratory and potentially also in medicine.

What groups of a ribozyme bind to substrates and what groups participate in catalysis? Like peptides, RNAs have amide groups that can hydrogen bond to substrates. Adenine and cytosine can supply protonated amino groups which could participate in acid–base catalysis. This is evidently the case in the **peptidyl transferase** centers of ribosomes. The RNA in these centers catalyzes a transesterification in which an aminoacyl group is transferred from an aminoacyl-tRNA anto the growing polypeptide chain attached to a second tRNA molecule. The reaction is evidently catalyzed by a universally conserved adenine ring located at position 2486 in the *Haloarcula marismortii* 23S RNA (position 2451 in *E. coli*). There are no protein groups within 1.8 nm of the location of peptide bond synthesis.[798a] The active site adenine appears to be much more basic than normal. A high pK_a of 7.6 controls the peptidyl transferase, and also controls the methylation of the active site adenine by dimethylsulfate.[798b] The site of protonation is thought to be largely N3 of adenine 2486, which is probably the basic center involved in catalysis. The peptidyltransferase reaction may be initiated as follows.[798a]



A similar catalytic mechanism is probably used by the small 85-nucleotide hepatitis delta virus ribozyme whose catalytic base is thought to be N3 of cytosine 75, which is associated with a pK_a of ~6.1.[798c] Both this HDV ribozyme and the ribosomal RNA resemble serine proteases with histidine as the catalytic base. However, the self-splicing RNA of *Tetrahymena* initiates a nucleophilic attack with the 3'–OH group of a guanosine molecule that is bound to a site in the P7 region (Fig. 12-16A) and which acts as a cofactor (see Fig. 28-18). Ribonuclease P and all group I and II self-cleaving introns also use an external nucleophile such as a guanosine –OH and form 3'-OH and *O*-phosphate or

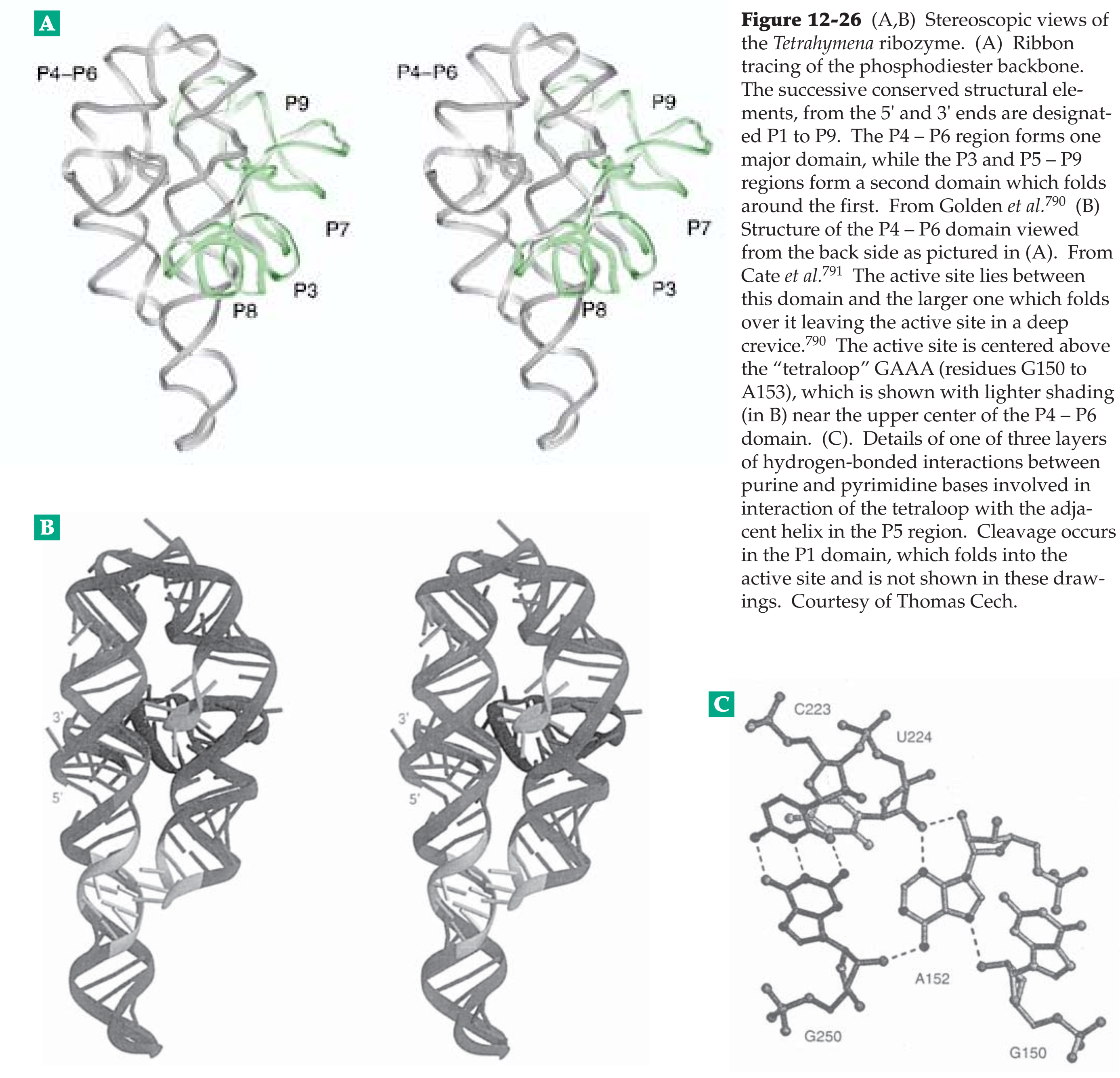


Figure 12-26 (A,B) Stereoscopic views of the *Tetrahymena* ribozyme. (A) Ribbon tracing of the phosphodiester backbone. The successive conserved structural elements, from the 5' and 3' ends are designated P1 to P9. The P4 – P6 region forms one major domain, while the P3 and P5 – P9 regions form a second domain which folds around the first. From Golden *et al.*[790] (B) Structure of the P4 – P6 domain viewed from the back side as pictured in (A). From Cate *et al.*[791] The active site lies between this domain and the larger one which folds over it leaving the active site in a deep crevice.[790] The active site is centered above the "tetraloop" GAAA (residues G150 to A153), which is shown with lighter shading (in B) near the upper center of the P4 – P6 domain. (C). Details of one of three layers of hydrogen-bonded interactions between purine and pyrimidine bases involved in interaction of the tetraloop with the adjacent helix in the P5 region. Cleavage occurs in the P1 domain, which folds into the active site and is not shown in these drawings. Courtesy of Thomas Cech.

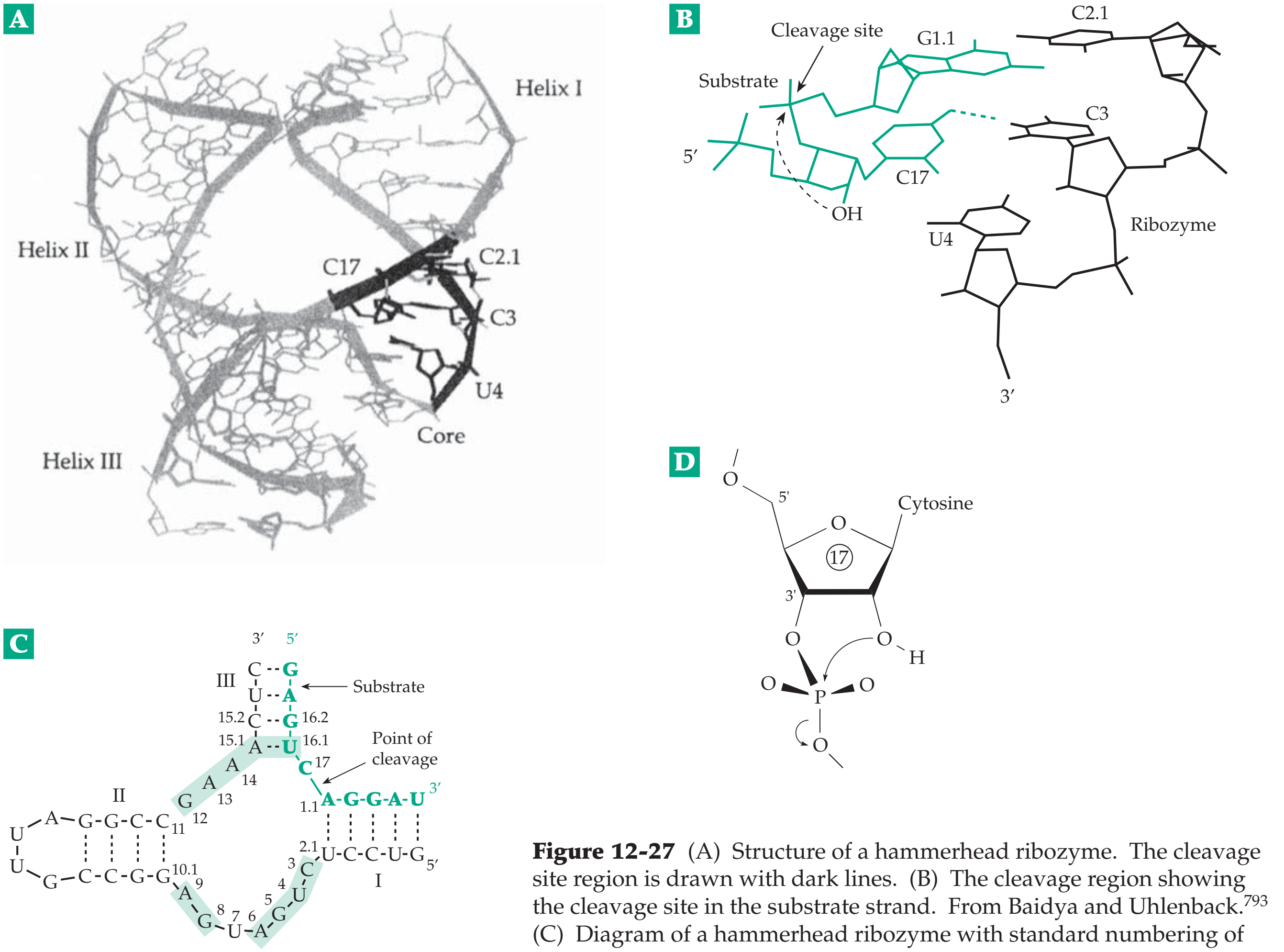

A
Helix I
Helix II
C17
C2.1
C3
U4
Core
Helix III
B
Cleavage site
G1.1
C2.1
Substrate
5′
C3
C17
OH
Ribozyme
U4
3′
D
O
5′
O
Cytosine
17
3′
O
O
H
O
P
O
O
C
3′ 5′
III
C - - G
U - - A
Substrate
15.2 C - - G 16.2
15.1 A - - U 16.1
Point of cleavage
C 17
A 14
A
G 13
II
A-G-G-A-U 3′
1.1
U A G-G-C-C 11 12
U G C-C-G-G 10.1 9
2.1 U-C-C-U-G 5′
3 C
I
4 U
5
A 8 7 6 G
G U A

Figure 12-27 (A) Structure of a hammerhead ribozyme. The cleavage site region is drawn with dark lines. (B) The cleavage region showing the cleavage site in the substrate strand. From Baidya and Uhlenback.[793] (C) Diagram of a hammerhead ribozyme with standard numbering of nucleotides. The three helical stems are labeled I, II, and III. From Bevers *et al.*[796] (D) Simplified cleavage mechanism which resembles step *a* of Fig. 12-25 and ignores the known participation of metal ions.

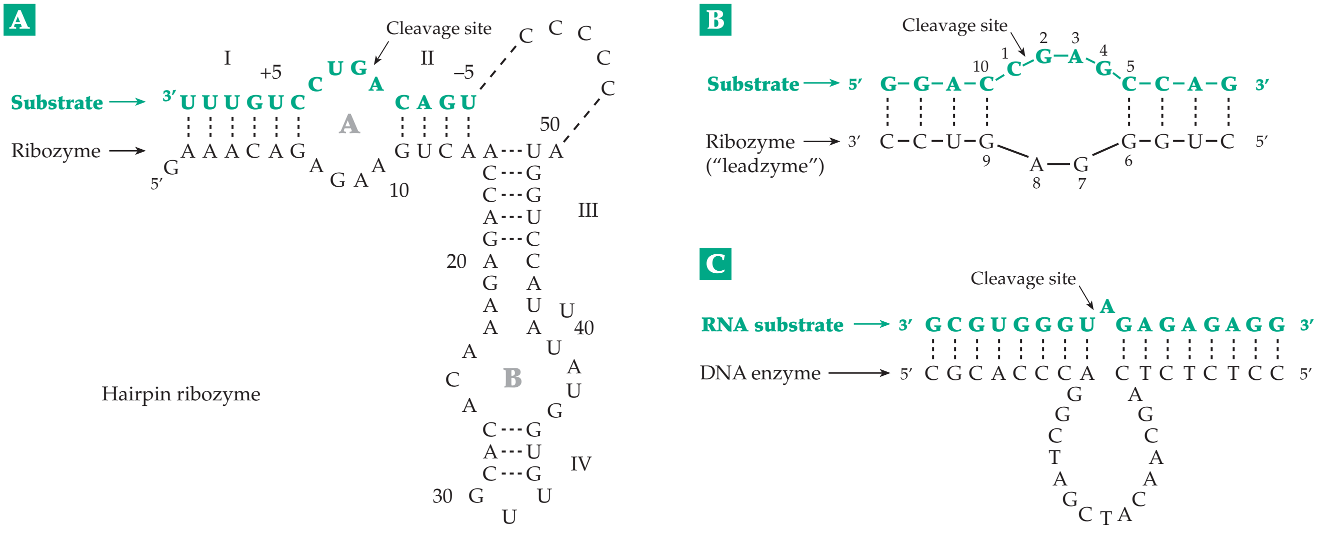
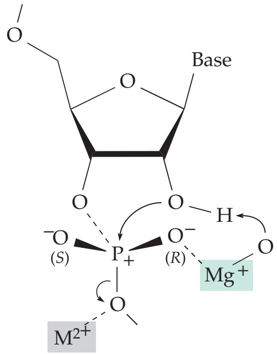

A
Cleavage site
I
+5
U G
C A
II
−5
C C C
C
C
Substrate
3′ U U U G U C
C A G U
A
50
Ribozyme
5′ G A A A C A G
A G A A
10
G U C A A - - - U A
C - - - G
C - - - G
A - - - U
III
G - - - C
20 A C
G A
A U U
A A 40
A U
C B A
A U
G
C - - - G
A - - - U
IV
C - - - G
30 G U
U
Hairpin ribozyme
B
Cleavage site
2 3
1 G-A 4
10 C G 5
Substrate 5′ G-G-A-C C-C-A-G 3′
Ribozyme 3′ C-C-U-G G-G-U-C 5′
(“leadzyme”)
9 6
A-G
8 7
C
Cleavage site
A
RNA substrate 3′ G C G U G G G U G A G A G A G G 3′
DNA enzyme 5′ C G C A C C C A C T C T C T C C 5′
G A
G G
C C
T A
A A
G C
C T A

Figure 12-28 (A) A hairpin ribozyme formed from the minus strand of a satellite RNA associated with tobacco ringspot virus. On the basis of hydroxyl radical footprinting (see Fig. 5-50), to identify protected areas a folding pattern that brings domains A and B together to form a compact catalytic core has been proposed.[798] (B) A “leadzyme,” a ribozyme dependent upon Pb^{2+} for cleavage of RNA.[802,803] (C) An RNA-cleaving DNA enzyme.[804]

phosphodiester ends at the cleavage points. Hammerhead ribozymes utilize the 2'-OH of the ribose at the cleavage site to form a 2',3'-cyclic phosphate ester as in step *a* of the ribonuclease A reaction (Fig. 12-25) and as indicated in Fig. 12-27D. Ribozymes act by in-line mechanisms, causing inversion of the configuration at the phosphorus.[799,800]

Most ribozymes, as well as RNase P,[805,806] require one or two metal ions for activity.[807] Magnesium ions predominate and many Mg^{2+} ions are bound at distinct sites in crystalline ribozymes. Hammerhead and hairpin ribozymes work reasonably well with monovalent ions. One proposed mechanism is for an HO^- ion bound to the Mg^{2+} to remove H^+ from the 2'–OH of the ribose ring as follows:



When the *pro-R* oxygen of the phospho group (labeled *R* in the accompanying structure) was replaced by sulfur the rate of cleavage decreased 10^3-fold. However, the rate was restored fully when Cd^{2+} was added, presumably because of the high affinity of cadmium ions for sulfur ligands.[808] This supports the possibility of a dual role for Mg^{2+} in activating a water molecule to provide HO^- and in stabilizing negative charges on the phospho group by interaction with the *pro-R* oxygen in the transition state. Several investigators have suggested that *two* metal ions may be needed. One possibility is for a mechanism similar to that proposed for alkaline phosphatase (Fig. 12-25) and other phosphotransferases.[809,810] The *Tetrahymena* ribozyme functions best if both Mg^{2+} and Mn^{2+} are present.[811] A second metal may act as a Lewis acid facilitating loss of the 5'-OH as indicated by M^{2+} (in gray) in the preceding diagram. Metal ions may also participate in conformational changes as well as have structural functions in ribozymes.[812,813] Studies of the kinetics of action of hammerhead ribozymes have suggested that the intrinsic ribozyme-substrate binding energy is utilized for catalysis.[814] This may be possible because the ribozyme is only partially folded in the ground state but it folds into a tighter conformation in the transition state.[815]

Methods have been devised for generating enormous numbers of RNA molecules with random sequences and for selecting those with unusual catalytic activities.[816–819] Among the new catalysts produced in this way are very small ribozymes that cleave RNA specifically in the presence of Pb^{2+} (Fig. 12-28B).[802,803] The leadzyme is more active with neodymium (Nd^{3+}) + Pb^{2+} than with lead alone, suggesting a two-metal mechanism.[820] Other artificial ribozymes include RNA ligases,[817] acyltransferases,[821] and DNA hydrolases.[822,823] Is it possible to find a DNA enzyme? Without the 2'–OH of ribose to form hydrogen bonds it seemed doubtful, but an RNA-cleaving DNA enzyme has been selected from a population of ~10^{14} different small DNA molecules. The DNA enzyme (Fig. 12-28C) will cleave RNA, whose sequence fulfills the base pairing requirements of two 8-deoxynucleotide recognition domains. Cleavage occurs between an unpaired purine and a paired pyrimidine using a metal-dependent mechanism that gives a 2',3'-cyclic phosphate as in ribonuclease A cleavage.[804]

7. Deoxyribonucleases (DNases)

A multitude of nucleases cleave DNA, single- or double-stranded. They range from the pancreatic digestive enzyme DNase I through specialized nucleases that function during DNA repair and the hundreds of restriction endonucleases that have become so valuable in modern laboratory work. Some nucleases leave a 3'-phosphate ester at a cut end in a DNA chain, while others leave a 5'-phosphate end.[824] Many nucleases are dealt with in later chapters. Only a few will be mentioned here.

One of the most studied enzymes of this group is the 149-residue micrococcal (staphylcoccal) nuclease from *Micrococcus* which cleaves either RNA or single- or double-stranded DNA. The relatively nonspecific enzyme cuts nearly randomly at the 5' side of the phosphodiester linkages, leaving 3'-phosphate groups. It enhances the uncatalyzed hydrolysis rate at least 10^{16}-fold.[825–827] The crystal structure showed that the majority of the acidic and basic side chains of the protein interact with each other through clusters of hydrogen bonds. At the active site the side chains of both Arg 35 and Arg 87 form pairs of hydrogen bonds to the 5'-phosphate group of the specific inhibitor deoxythymidine 3',5'-diphosphate (Fig. 12-29). While Arg 87 appears to be in a position to protonate the leaving group –O$^-$, ^{13}C NMR experiments showed that all of the arginine side chains had pK_a values above 11.6. However, Tyr 85 has a pK_a of 9.5, which appears to control k_{cat} / K_m.[828] A Y85F mutant lacks this pK_a. The X-ray structure also suggests that Glu 43 may be the attacking nucleophile and that it may deprotonate a water molecule bound to the Glu 43 carboxylate. The resulting HO^- probably carries out a direct in-line attack as shown in Fig. 12-29. Mutants such as E43D, E43Q, and E43S have greatly decreased activity,[829] in

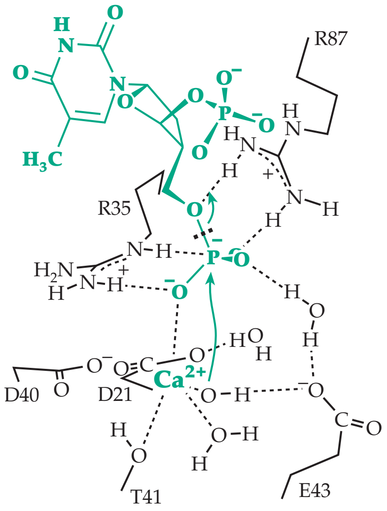


Figure 12-29 Drawing showing the hydrogen-bonding interactions between the guanidinium ions of arginines 35 and 87 of the micrococcal (staphylococcal) nuclease with the 5'-phosphate of the inhibitor thymidine 3',5'-diphosphate in the complex of E + I + Ca^{2+}. A possible mechanism is illustrated. A hydroxyl ion bound to Ca^{2+} carries out an in-line attack on the phosphorus. See Libson *et al.*[826]

agreement with this mechanism. The nearby Ca^{2+} is essential. For this nonspecific nuclease there is no hydrogen bonding of a purine or pyrimidine base of the substrate to the enzyme.

The digestive enzyme, pancreatic DNAse I, makes single-stranded cuts in double-stranded (ds) DNA. An exposed strand of peptide chain from the enzyme binds into the minor groove of B-type DNA.[830] Because this groove becomes too narrow in long (A+T) rich sequences, they are cleaved slowly. Certain hypersensitive sites are cleaved very rapidly, perhaps because the DNA at these regions is bent or is able to bend to give a very good fit to the enzyme active site. A histidine which is hydrogen bonded to a nearby carboxylate of a glutamate side chain appears to be a catalytic base that acts upon a water molecule as in phospholipase A (Section D,10), displacing the 3' oxygen of the phosphodiester linkage. An imidazolium group from a second histidine is hydrogen bonded to an aspartate carboxylate and a tyrosine –OH to form a catalytic triad that can protonate the 3' $-O^-$ as it is displaced.[831] Two Mg^{2+} ions are also required. Both are held by different carboxylate side chains and may also interact with oxygen atoms of the phospho group to neutralize charge and stabilize the transition state.

In contrast to DNAse I, the **restriction endonucleases**, which are discussed in Chapter 5, Section H,2 and in Chapter 26, have precise substrate sequence specificities. Three of the best known restriction endonucleases are called *Eco*RI,[832–834] an enzyme which binds to and cuts both strands of the palindromic sequence 5'-GAATTC; *Eco*RV,[835] which cuts both strands in the center of the sequence 5'-GATATC; and *Bam*H1, which binds to the sequence 5'-GGATCC and cleaves after the 5' G on each strand.[836] A high-resolution structure is also known for *Cfr*10I, which recognizes the less strict sequence 5'-PuCCGGPy and cleaves both strands after the 5' Pu.[837] All of these enzymes require Mg^{2+} and have active sites containing carboxylate groups. Two-metal mechanisms have been suggested.

Restriction endonuclease *Eco*RI is able to cut a chain in dsDNA which has a chiral phophorothioate group at the specific cleavage site.[838] The reaction occurs with inversion of configuration at phosphorus, suggesting direct in-line attack by a hydroxyl ion generated from H_2O.

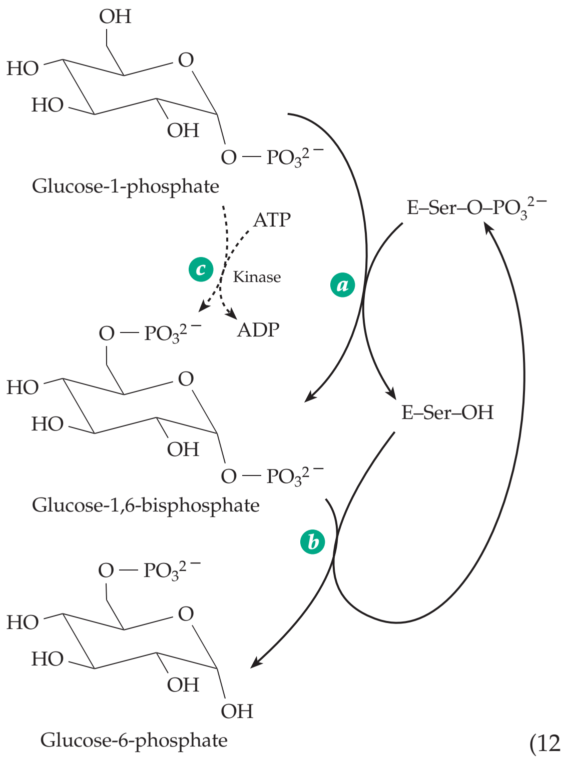
Attempts are being made to design semisynthetic restriction endonucleases specific for single-stranded DNA or RNA. For example, an oligonucleotide with a sequence complementary to a sequence adjacent the linkage that is to be cut can be covalently linked to a relatively nonspecific nuclease. Such an enzyme derived from micrococal nuclease cuts a single-stranded chain of either DNA or RNA adjacent to the double-stranded region of the ES complex.[839]

8. Mutases

Phosphotransferases that shift phospho groups from one position within a substrate to another are often called **mutases**. For example, **phosphoglucomutase** catalyzes the interconversion of glucose 1-phosphate and glucose 6-phosphate, an important reaction that bridges glycogen metabolism and glycolysis (Fig. 11-2). This 561-residue protein operates through formation of an intermediate **phosphoenzyme**.[840–842] The phospho group becomes attached to the OH of Ser 116 and can be transferred either to the 6 or the 1 position of a glucose phosphate (step *a* and reverse of step *b* in Eq. 12-39).

The two-step reaction accomplishes the reversible isomerization of glucose 1-phosphate and glucose 6-phosphate via **glucose 1,6-bisphosphate**. Evidently, the glucose bisphosphate, without leaving the enzyme is reoriented to allow transfer of the phospho group to either the 1- or 6-position.[841] The phospho enzyme is relatively unstable and can undergo hydrolysis to free enzyme and P_i. To prevent loss of active enzyme in this way, a separate reaction (catalyzed by a kinase; Eq. 12-39, step *c*)[843] generates glucose 1,6-bisphosphate, which rephosphorylates any free enzyme formed by hydrolysis of the phospho enzyme. The glucose 1,6-bisphosphate can be regarded as a cofactor or **cosubstrate** for the reaction.

Phosphoglycerate mutase, which interconverts 2-phosphoglycerate and 3-phosphoglycerate in glycolysis (Fig. 10-3, step *c*), functions by a similar mecha-



(12-39)

nism.[843–845] However, the enzyme-bound phospho group is carried on an imidazole group. The essential cosubstrate required by some phosphoglycerate mutases is 2,3-bisphosphoglycerate.[845] It is formed from the glycolytic intermediate 1,3-bisphosphoglycerate (Fig. 10-3) by action of another mutase, **bisphosphoglycerate mutase**.[843,846]

$$\text{1,3-bisphosphoglycerate} \rightarrow \text{2,3-bisphosphoglycerate} \tag{12-40}$$

This is also the pathway for synthesis of 2,3-biphosphoglycerate in red blood cells where it serves as an important allosteric regulator (Chapter 7). Two human phosphoglycerate mutase isoenzymes are known. One is found in muscle and the other in brain and other tissues.[844] A hereditary lack of the muscle type enzyme is one of the known types of glycogen storage diseases (Box 20-D).

Human **phosphomannomutase**, which catalyzes the interconversion of mannose 1- and 6- phosphates, appears to carry the phospho group on an aspartate side chain in the sequence **D**XDX (T/V), which is conserved in a family of phosphomutases and phosphatases.[847] The first aspartate in the sequence is phosphorylated during the enzymatic reaction.

9. Molecular Properties of Kinases

Kinases transfer phospho groups from polyphosphates such as ATP to oxygen, nitrogen, or sulfur atoms of a second substrate.[848] Examples include **hexokinase**, the enzyme responsible for synthesis of glucose 6-phosphate from free glucose and ATP (Fig. 11-2, step *a*); **phosphofructokinase**, which forms fructose 1,6 bisphosphate in the glycolysis pathway (Fig. 11-2, step *b*); and **phosphoglycerate kinase**, and **pyruvate kinase**, both of which form ATP from ADP in the glycolysis pathway (Fig. 10-3, steps *b*, *c*, and *f*). There are many others. Kinases vary greatly in size and in three-dimensional structure. For example, a small **adenylate kinase**, which phosphorylates AMP to ADP (Eq. 6-65), is a 22-kDa monomer of 194 residues. Pyruvate kinase is a tetramer of 60-kDa subunits and muscle phosphofructokinases are tetramers of 75- to 85-kDa subunits. The three-dimensional structures also vary. While all kinases consist of two domains built around central β sheets (Fig. 12-30), there are several different folding patterns.[849] The two-domain structures all have deep clefts which contain the active sites. Both adenylate kinase (Fig. 12-30) and hexokinase crystallize in two or more forms with differing conformations.[850] This and other evidence suggests that as a kinase binds and recognizes its two correct substrates, the active site cleft closes by a hinging action that brings together the reacting molecules in the correct orientation.[851–853] In the crystal structure shown in Fig. 12-30 both ADP and AMP are bound in a nonproductive complex. If the ADP were replaced with ATP (or the AMP with a second ADP) to form a productive complex the two reacting phospho groups would be ~0.8 nm apart. A reaction could not occur without further closing of the active site cleft.[854] Evidence for domain closure has been obtained for many other kinases. For example, substrate complexes of phosphoglycerate kinase have been crystallized in both "open" forms and "closed" forms in which a 30° hinge-bending movement has brought the ligands together for an in-line phospho group transfer.[855]

For many enzymes an ATP binding site has been revealed by study of nonhydrolyzable analogs of ATP such as "AMP-PNP" whose structure is shown in Fig. 12-31.[856] AMP-PNP has been used in thousands of investigations of ATP-dependent processes.[857] For example, the structure of a phosphoglycerate • Mg^{2+} • AMP-PNP complex in the active site of phosphoglycerate kinase has been determined.[858] Modeling of a transition state complex indicates that all three negatively charged oxygens of the ATP portion of structure are stabilized by hydrogen bonding.[852] Related analogs such as AMP-PCH_2P (Fig. 12-3) have been used in similar ways.[859] Another analog Mg^{2+} • Ap_5A (Fig. 12-31) is a bisubstrate inhibitor which binds to adenylate kinases, fixing the enzymes in a closed conformation that is thought to resemble the transition state.[859,860]

The Mg^{2+} complex of ATP is regarded as the true substrate for kinases. The metal usually also binds both to the phospho groups of ATP and to groups on

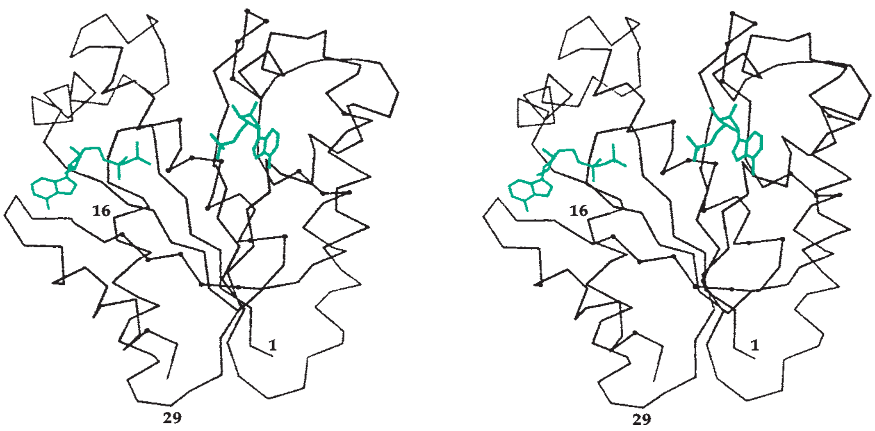
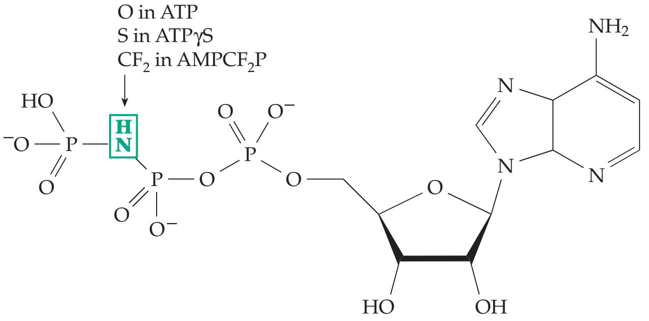
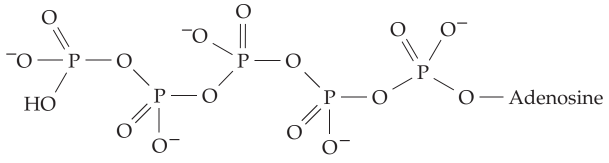


Figure 12-30 Stereoscopic α-carbon plots of a 194-residue subunit of adenylate kinase from the archaebacterium *Sulfolobus acidocaldarius* with ADP (left side) and AMP (right side) bound into the active site. From Vonrheim *et al.*[854] Courtesy of G. E. Schulz.



AMP-PNP



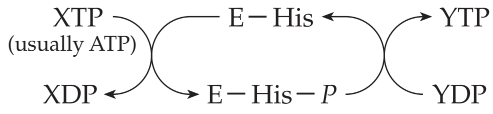
Ap_5A, a bisubstrate inhibitor of adenylate kinase

Figure 12-31 Some useful analogs of ATP.

the enzyme. Crystallographic investigations as well as studies with exchange-inert ATP complexes (Section 2) have suggested that the metal ion is bound initially to the terminal (γ) and adjacent (β) phospho groups of ATP. However, this metal bridge could prevent rather than assist the reaction. For some phosphotransferase reactions with exchange-inert complexes the product appears to be the α,β-bidentate complex of ADP (see Section 2) suggesting that movement of the metal ion from the β,γ- to the α,β-bidentate ATP complex may occur either prior to or during transfer of the phospho group. Theoretical calculations also suggest a movement in the metal complexation as the reaction progresses.[861] EPR studies that involved observation of hyperfine coupling between ^{17}O in substrates and Mn^{2+} in the active site led Reed and Leyh to conclude that the activating metal ion is bound to all three phospho groups in the transition state.[849]

Adenylate kinase performs the essential function of recovering AMP formed by many enzymatic processes and converting it to ADP (Eq. 6-65) which can be reconverted to ATP by oxidative or substrate level phosphorylation. The enzyme is present in all organisms. In vertebrates different isoenzymes function in the cytosol, mitochondrial intermembrane space, and mitochondrial matrix.[862,863] A group of other **nucleotide** and **deoxynucleotide kinases** convert nucleoside monophosphates into diphosphates.[864,865] Some of them, e.g., **uridylate kinase** are similar in structure and properties to adenylate kinase.[866,867] Another member of the adenylate kinase family is phosphoribulokinase, an important photosynthetic enzyme (see Fig. 17-14, step *a*).[868]

Most kinases transfer chiral phospho groups with inversion and fail to catalyze partial exchange reactions that would indicate phosphoenzyme intermediates. However, **nucleoside diphosphate kinase** contains an active site histidine which is phosphorylated to form a phosphoenzyme.[869] The enzyme catalyzes phosphorylation of nucleoside diphosphates other than ADP by a nucleotide triphosphate, usually ATP.



(12-41)

Here, X is usually adenosine and Y is any ribonucleoside or deoxyribonucleoside. This enzyme supplies all of the nucleotide triphosphates except ATP for use in the many cellular processes that require them.[870–872] The enzyme aligns the substrate, holding the phospho group with a pair of arginines and a magnesium ion. The phospho group is aligned for an in-line displacement by N^{δ} of His 122, part of a hydrogen-bonded His–Glu dyad. Formation of the phosphoenzyme occurs in less than 1 ms without significant conformational change other than a 30° rotation of the histidine ring.[870,873]

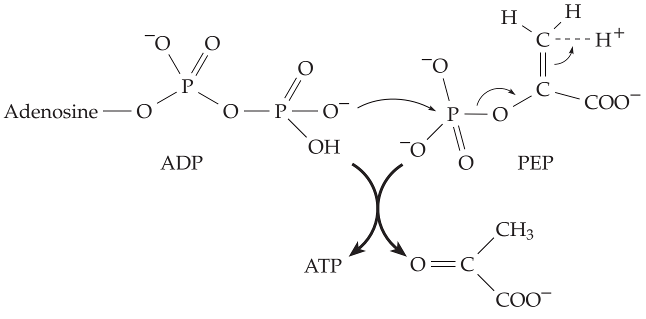
Hexokinase, the enzyme that phosphorylates glucose to glucose 6-phosphate, exists as four isoenzymes in mammals. Hexokinases I, II, and III are large ~ 100-kDa monomers with similar amino acid sequences, a single active site, and complex allosteric regulatory properties. The three-dimensional structure is known for hexokinase I, which is the pacemaker of glycolysis in the brain.[874,875] Hexokinase IV (glucokinase) is a 50-kDa protein that is found in liver and in the β cells of the pancreas. It has a low K_m and is not inhibited by glucose 6-phosphate, properties that allow for rapid uptake of glucose after a meal.[876] Its properties are similar to those of a major isoenzyme of yeast,[877] a dimer of identical 50-kDa subunits whose structure is also known.[878] The glucose-binding residues are conserved in yeast hexokinase, in glucokinase, and in brain hexokinase. The sequence of the latter suggests that it may have arisen by a doubling of a shorter hexokinase gene.[879]

Fructose 6-phosphate kinase (phosphofructokinase) has attracted much attention because of its regulatory properties (Chapter 11).[880–882] Prokaryotic forms are somewhat simpler.[883] The related fructose 6-phosphate 2-kinase is a component of a bifunctional kinase-phosphatase (Fig. 11-2, steps *d* and *e*) and has a structure similar to that of adenylate kinase.[884]

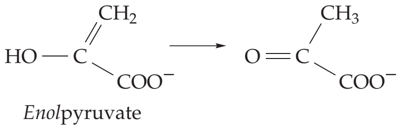
Phosphoglycerate kinase is encoded by the mammalian X chromosome. Several mutant forms are associated with hemolytic anemia and mental disorders.[885,886] Mutation of Glu 190, which is in the hinge region far from the active site, to Gln or Asn markedly reduces enzymatic activity.[887]

Creatine kinase transfers a phospho group to a nitrogen atom of the quanidinium group of creatine (Eq. 12-31 and Fig. 12-19). Several isoenzymes participate in its function of buffering the ATP level in tissues such as muscle fibers, neurons, photoreceptors, and spermatozoa which experience high and fluctuating energy needs.[888] A form from the mitochondrial intermembrane space of chicken heart is an octomer of 380-residue subunits.[889] The structurally and mechanistically similar **arginine kinase** has an analogous function in many invertebrates, e.g., in the horseshoe crab, which provided enzyme for a structure determination.[890]

The ~500-residue subunits of **pyruvate kinase** consist of four domains,[891] the largest of which contains an 8-stranded barrel similar to that present in triose phosphate isomerase (Fig. 2-28). Although these two enzymes catalyze different types of reactions, a common feature is an enolic intermediate. One could imagine that pyruvate kinase protonates its substrate phosphoenolpyruvate (PEP) synchronously with the phospho group transfer (Eq. 12-42). However, the enzyme catalyzes the rapid conversion of the enolic form of pyruvate to the oxo form (Eq. 12-43) adding the proton sterospecifically to the *si* face. This and other evidence favors the enol as a true intermediate and a product of the phosphotransfer step.[891] Pyruvate kinase requires not only two equivalents of a divalent cation such as Mg^{2+} or Mn^{2+} but also a monovalent cation, usually K^+. However, Li^+, Na^+, NH_4^+, Rb^+, Tl^+, and others can substitute. The monovalent cation induces an essential conformational change. Using ^{205}Tl NMR it was found that the thallium ion binds about 0.6 nm from Mn^{2+} that is also present in the active site.[892] All three metals interact directly with the γ phospho group of ATP.[891] Pyruvate kinase is a regulated allosteric enzyme present in four isoenzymic forms in mammals.[893–894a]



(12-42)



(12-43)

Protein kinases, which were discussed in Chapter 11, phosphorylate selected – OH groups of serine, threonine, and tyrosine side chains in proteins. Examination of the sequences in the complete genome of yeast (*Saccharomyces cerevisiae*) indicates the presence of at least 113 protein kinase genes, which account for ~ 2% of the total DNA.[895] Higher eukaryotes have more. While structures of these enzymes vary widely, they share a common two-domain catalytic core structure.[896,897] The best known, and one of the simplest of them, is the catalytic subunit of cyclic AMP-dependent protein kinase.[897,898] The substrate MgATP binds into the active site cleft with the γ-phospho group protruding to meet the appropriate site of a bound protein substrate (Fig. 12-32). Most protein kinases are regulated by an "activation loop" that must be correctly placed before the ES complex can be formed. As discussed in Chapter 11, regulation of the cAMP-dependent kinase depends upon inhibition by a regulatory subunit. In the cAMP-dependent kinases and many other protein kinases the activation loop (not shown in Fig. 12-32), which helps to form the substrate site, contains a phosphothreonine residue which is essential for activity. It is a stable feature of the cAMP-activated kinase, incorporated at position 197, but for some tyrosine kinases it is generated by autophosphorylation.[899]

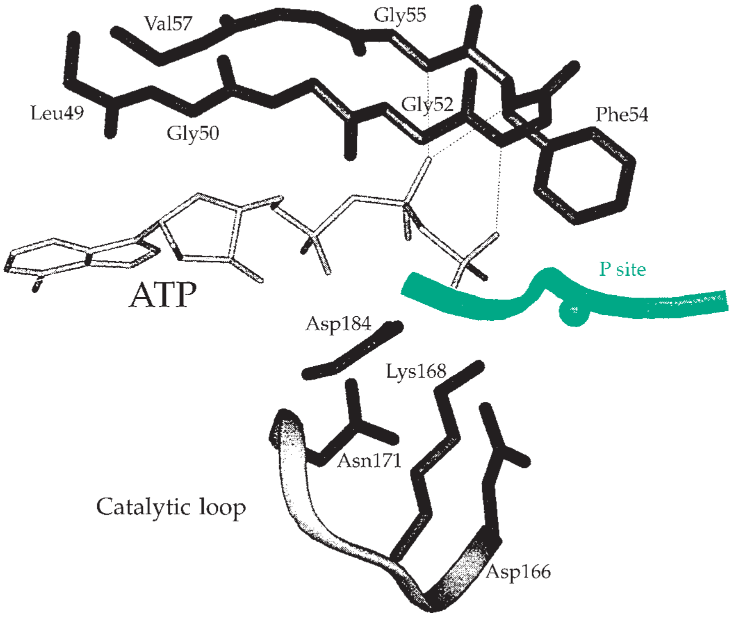


Figure 12-32 The active site of cyclic AMP-dependent protein kinase with bound ATP and a segment of an inhibitory peptide (green) blocking the substrate site. The small ball is an alanine side chain; it would be Ser or Thr in a substrate. The ATP is clamped by the glycine-rich loop at the top of the figure. Hydrogen bonds from peptide NH groups bind the β and γ phospho groups of ATP. A magnesium ion (not shown) also binds to the β and γ phosphate oxygens and to the invariant Asp 184. The Asp 166 carboxylate is probably the catalytic base for deprotonation of the substrate –OH. From Grant *et al.*[896] Courtesy of Susan S. Taylor.

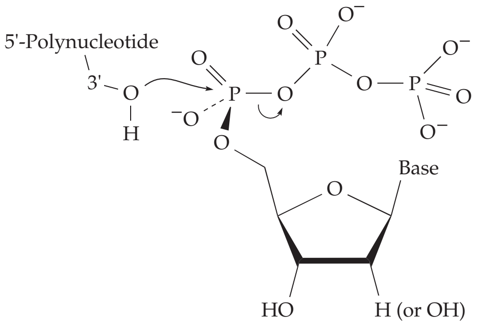
For the enzyme to be activated the phosphothreonine must form a hydrogen-bonded ion pair with Lys 189 and be hydrogen bonded to His 87, tying together critical regions of the catalytic domain. The regulatory subunits, unless occupied by cAMP (Chapter 11) are competitive inhibitors of the substrates.[900] Cyclic GMP-activated kinases also have distinct functions.[901] In **tryosine kinases**[902–904] C-terminal src-homology domains (Fig. 7-30 and Chapter 11) fold over and interact in an inhibitory fashion until an appropriate activating signal is received (see Fig. 11-13).

A relative of the kinases is **adenylate cyclase**, whose role in forming the allosteric effector 3',5'-cyclic AMP (cAMP) was considered in Chapter 11. This enzyme catalyzes a displacement on P_α of ATP by the 3'-hydroxyl group of its ribose ring (see Eq. 11-8 , step *a*). The structure of the active site is known.[905] Studies with ATPαS suggest an in-line mechanism resembling that of ribonuclease (step *a*, Eq. 12-25). However, it is Mg^{2+} dependent, does not utilize the two-histidine mechanism of ribonuclease A, and involves an aspartate carboxylate as catalytic base.[906] All isoforms of adenylate cyclase are activated by the α subunits of some G proteins (Chapter 11). The structures[907] of $G_{s\alpha}$ and of its complex with adenylate kinase[905] have been determined. The $G_{s\alpha}$ activator appears to serve as an allosteric effector.

Guanylate cyclases, which form cyclic GMP, occur in particulate and soluble forms.[908] The latter have been of great interest because they are activated by nitric oxide (NO). The soluble guanylate cyclases are αβ heterodimers. The C-terminal regions of both α and β subunits are homologous to the catalytic domain of adenylate cyclase. The N-terminal domain of the α subunits contains heme whose Fe atom is coordinated by a histidine imidazole.[908,908a] This iron atom is apparently the receptor for NO, a major gaseous hormone, which is discussed in Chapter 18.

10. Nucleotidyl Transferases

An important group of enzymes transfer substituted phospho groups, most often nucleotidyl groups. The nucleases, ATPases, and GTPases, which have already been discussed, belong to this group as do the nucleic acid synthesizing enzymes, the **DNA** and **RNA polymerases**,[909–911] **reverse transcriptase**,[912] and **topoisomerases**. As with other phosphotransferases, the nucleotidyl transfers occur with inversion[913,913a] and crystallographic investigations also support in-line mechanisms as illustrated in the following scheme. Two metal ions assist.



The inorganic pyrophosphate formed is hydrolyzed to inorganic phosphate by pyrophosphatase. Specific information about the polymerases and topoisomerases is given in Chapters 27 and 28.

E. The Adenylate Kinase Fold, the P Loop, and ATPases and GTPases

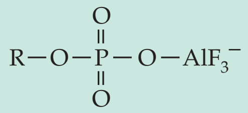
A magnesium–ATP-binding fragment consisting of ~40 residues at the N terminus of adenylate kinase contains sequences homologous to those in the GTP-binding "G proteins" such as the protooncogenes *ras* (Fig. 11-7A) and also to sequences in myosin and in mitochondrial ATP synthase.[914] This includes the glycine-rich "P loop" which extends from Gly 15 to Gly 22 in the porcine cytosolic enzyme and contains a highly conserved lysine [Lys 21 of porcine adenylate

BOX 12-F THE TOXICITY OF ALUMINUM

Aluminum, in the form of oxides and silicate minerals, is the most abundant element in the earth's crust. Yet it appears to be actively excluded from most living organisms. It seems surprising that a naturally occuring Al^{3+}-dependent enzyme hasn't been found, but there is no evidence that Al is an essential element. Until recently, it was usually regarded as harmless. The aluminum salts known as alums are used in baking powders and have been added to pickles and other foods. Aluminum sulfate is often used as a coagulant to clarify turbid drinking water. Because of the insolubility of $Al(OH)_3$ and of aluminum phosphate the concentration of Al^{3+} is very low at neutral pH. However, small amounts of $AlOH^{2+}$, $Al(OH)_2^+$, and $Al(OH)_4^-$ are present in water. Fluoride complexes such as AlF^{2+} may also be present. Soluble complexes, such as that with citrate, may permit some Al^{3+} to be absorbed by the body.[a,b]

The toxicity of aluminum has been recognized most clearly by the development of bone disease caused by deposition of Al in bones of patients on hemodialysis[a–c] and in infants on intravenous therapy.[d,e] Excessive Al in the water used for dialysis may cause brain damage. Dietary aluminum may be one cause of Alzheimer's disease,[f–h] but this is controversial as is a possible role of aluminum in vaccines in causing inflammation in muscle.[i,j] Solubilization of soil aluminum by acid rain has been blamed for the decline of forests in Europe and North America,[j] for the death of fish in acid waters,[k] and for very large reductions in yield for many crops.[l,m] An aluminum-resistant strain of buckwheat makes and secretes from its roots large amounts of oxalate which binds and detoxifies the Al^{3+} ions.[m]

Al^{3+} has a radius somewhat less than that of Fe^{3+} (Table 6-10) and it may sometimes occupy empty Fe^{3+} binding sites. Thus, the transferrin[n] in blood carries some Al^{3+}, although citrate is probably a more important carrier.[o] Al^{3+} binds preferentially to oxygen ligands and can compete with Mg^{2+}. However, the slower rate of ligand exchange reactions with Al^{3+} may interfere with the proper functioning of the metal. Brain hexokinase is strongly inhibited[p] by Al^{3+} and the binding of Al^{3+} to tubulin decreases the rates of GTP hydrolysis and of Ca^{2+}-induced depolymerization of microtubules.[q] Aluminofluoride ions, such as AlF_4^-, react with phosphates to form ions such as,



$$R-O-\underset{\underset{O}{\|}}{\overset{\overset{O}{\|}}{P}}-O-AlF_3^-$$

which may be potent competitive inhibitors of enzymes acting on ATP, GTP, or other phosphate-containing substrate.[r,s] However, Fe^{3+} can be replaced by Al^{3+} in a purple acid phosphatase (Chapter 16) with retention of good catalytic activity.[t]

---

[a] Martin, R. B. (1986) *Clinical Chemistry* **32**, 1797–1806

[b] Macdonald, T. L., and Martin, R. B. (1988) *Trends Biochem. Sci.* **13**, 15–19

[c] Andress, D. L., Kopp, J. B., Maloney, N. A., Coburn, J. W., and Sherrard, D. J. (1987) *N. Engl. J. Med.* **316**, 292–296

[d] Sedman, A. B., Klein, G. L., Merritt, R. J., Miller, N. L., Weber, K. O., Gill, W. L., Anand, H., and Alfrey, A. C. (1985) *N. Engl. J. Med.* **312**, 1337–1342

[e] Bishop, N. J., Morley, R., Chir, B., Day, J. P., and Lucas, A. (1997) *N. Engl. J. Med.* **336**, 1557–1561

[f] Good, P. F., and Perl, D. P. (1993) *Nature (London)* **362**, 418

[g] Shen, Z. M., Perczel, A., Hollósi, M., Nagypál, I., and Fasman, G. D. (1994) *Biochemistry* **33**, 9627–9636

[h] Walker, P. R., LeBlanc, J., and Sikorska, M. (1989) *Biochemistry* **28**, 3911–3915

[i] Landsberg, J. P., McDonald, B., and Watt, F. (1992) *Nature (London)* **360**, 65–68

[ij] Malakoff, D. (2000) *Science* **288**, 1323–1324

[j] Godbold, D. L., Fritz, E., and Hüttermann, A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3888–3892

[k] Birchall, J. D., Exley, C., Chappell, J. S., and Phillips, M. J. (1989) *Nature (London)* **338**, 146–148

[l] Barinaga, M. (1997) *Science* **276**, 1497

[m] Ma, J. F., Zheng, S. J., Matsumoto, H., and Hiradate, S. (1997) *Nature (London)* **390**, 569–570

[n] Roskams, A. J., and Connor, J. R. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9024–9027

[o] Martin, R. B., Savory, J., Brown, S., Bertholf, R. L., and Wills, M. R. (1987) *Clinical Chemistry* **33**, 405–407

[p] Viola, R. E., Morrison, J. F., and Cleland, W. W. (1987) *Biochemistry* **19**, 3131–3137

[q] Macdonald, T. L., Humphreys, W. G., and Martin, R. B. (1987) *Science* **236**, 183–186

[r] Troullier, A., Girardet, J.-L., and Dupont, Y. (1992) *J. Biol. Chem.* **267**, 22821–22829

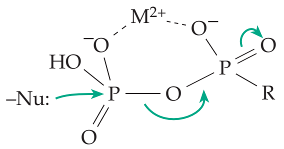
[s] Chabre, M. (1990) *Trends Biochem. Sci.* **15**, 6–10

[t] Merkx, M., and Averill, B. A. (1999) *J. Am. Chem. Soc.* **121**, 6683–6689

kinase and Lys 17 of the archaeal enzyme (Fig. 12-30)] and of the *ras* oncogene product).[915] The lysine side chain appears to interact with the β and / or γ phospho group of ATP or GTP. The peptide chain of the P loop wraps around the β phospho group as the chain turns from the first β strand into the first helix. This can be seen in Fig. 12-30, in which the P loop surrounds the β phospho group of ADP. Two peptide NH groups also bind to the β phospho group on a side opposite to that occupied by Mg^{2+}. A similar loop in the dinucleotide-binding domain of a dehydrogenase can be seen hydrogen bonded to the P_α and P_β phospho groups of NAD^+ in Fig. 2-13. Consensus sequences for three groups of glycine-rich loops are:[896]

Dinucleotide-binding	G X G X X G
P loop	G X X X X G K (S/T)
Protein kinase	G X G X X G X V

The protein kinase loop is seen at the top of Fig. 12-32 and extends from Gly 50 to Val 57. These conserved loops help to hold the ATP in place and to orient it correctly. Do they have any other significance? The answer is not clear. These glycine-rich loops fold across the β–γ diphosphate linkage that is broken when ATP or GTP is hydrolyzed. Cleavage of both of these molecules is associated with movement. Kinases close around ATP, and parts of G proteins (Fig. 11-7) move when these regulatory devices function.[915,916] Cleavage of ATP causes movement in the ATPase heads of muscle myosin, providing the force for muscular contraction (Chapter 19). Somewhat the opposite occurs in mitochondrial ATP synthase (Chapter 18) when movement in the synthase heads snaps ADP and inorganic phosphate together to form ATP. Common to all of these processes is a movement of charge within the ATP (or its cleavage products) from the attack nucleophile into the neighboring phospho group, as is indicated by the arrows in the following diagram.



Not shown in this diagram is an accompanying flow of positive charge which may include movement of a metal ion, addition and loss of protons, and which may induce conformational changes.[914] The latter are essential not only to muscle contraction and ATP synthesis but also to many other processes that depend upon the Gibbs energy of cleavage of ATP and related compounds. This includes the pumping of ions against concentration gradients (Chapter 8), the action of **topoisomerases** (Chapter 27) which function to alter the supercoiling in DNA, and the functioning of the phosphotransferase system by which sugars and other compounds are brought into bacterial cells (Chapter 8).

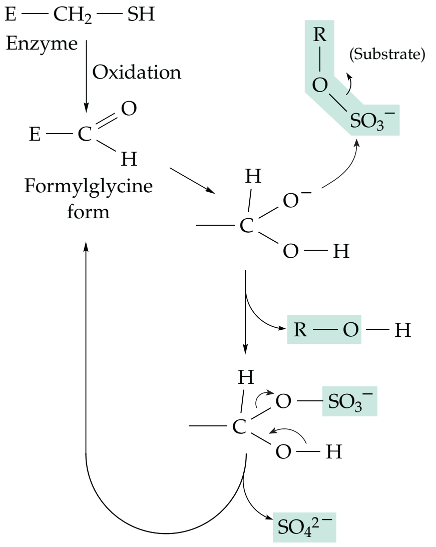
F. Displacements on Sulfur Atoms

Nucleophilic displacement reactions occur on sulfur atoms in various oxidation states. A common reaction is **thiol–disulfide** exchange (Eqs. 10-9, 11-7), a reaction in which a nucleophilic thiolate anion attacks one of the sulfur atoms of a disulfide. Proteins such as **thioredoxin** of *E. coli* and **thioltransferases** (Box 15-C), which contain internal disulfide bridges, can be reduced by disulfide exchange with thiols such as glutathione (Box 11-B). The reduced proteins may then undergo similar exchange reactions that cleave disulfide linkages in other molecules. An example is glutathione reductase (Fig. 15-12). Thioltransferases may also serve as **protein disulfide isomerases** (Chapter 10, Section D,3). Nucleophilic displacements on sulfur or on selenium atoms are steps in a variety of enzymatic reactions. Among these are glutathione peroxidase (Eq. 15-59) and thiosulfate: cyanide sulfotransferase (Eq. 24-45).

While esters of sulfuric acid do not play as central a role in metabolism as do phosphate esters, they occur widely. Both oxygen esters ($R–O–SO_3^-$, often referred to as **O-sulfates**) and derivatives of sulfamic acid ($R–NH–SO_3^-$, **N-sulfates**) are found, the latter occurring in mucopolysaccharides such as heparin. Sulfate esters of mucopolysaccharides and of steroids are ubiquitous and sulfation is the most abundant known modification of tyrosine side chains. Choline sulfate and ascorbic acid 2-sulfate are also found in cells. Sulfate esters of phenols and many other organic sulfates are present in urine.

Sulfotransferases[917–920a] transfer sulfo groups to O and N atoms of suitable acceptors (reaction type 1D, Table 10-1). Usually, transfer is from the "active sulfate," **3'-phosphoadenosine 5'-phosphosulfate** (**PAPS**),[921] whose formation is depicted in Eq. 17-38. **Sulfatases** catalyze hydrolysis of sulfate esters. The importance of such enzymes is demonstrated by the genetic **mucopolysaccharidoses**. In four of these disease-specific sulfatases that act on iduronate sulfate, heparan *N*-sulfate, galactose-6-sulfate, or *N*-acetylglucosamine-4-sulfate are absent. Some of these, such as heparan *N*-sulfatase deficiency, lead to severe mental retardation, some cause serious skeletal abnormalities, while others are mild in their effects.[922]

Sulfatases are unusual in having a residue of **formylglycine** at the active site. This is generated oxidatively from cysteine in human enzymes[923,923a] and from serine in some bacterial sulfatases.[924,924a] Absence of this modification results in a multiple sulfatase deficiency disease. A probable mechanism of sulfatase based on a crystal structure determination is given by Eq. 12-44.[923]



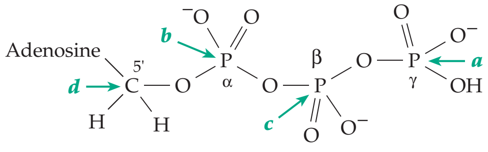
(12-44)

G. Multiple Displacement Reactions and the Coupling of ATP Cleavage to Endergonic Processes

A combination of successive displacement reactions of two types is required in many enzymatic reactions, including most of those by which the cleavage of ATP is coupled to biosynthesis. To harness the group transfer potential of ATP to drive an endergonic metabolic process there must be a mechanism of **coupling**. Otherwise, hydrolysis of ATP within a cell would simply generate heat. *An essential part of the coupling mechanism usually consists of a nucleophilic displacement on phosphorus followed by displacement on carbon.* Likewise, the synthesis of ATP and related compounds often begins with a displacement on carbon followed by one on phosphorus.

1. Transfer of Phospho, Pyrophospho, and Adenylyl Groups from ATP

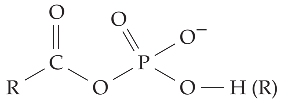
The first step in coupling ATP cleavage to any process is transfer of part of the ATP molecule to a nucleophile Y, usually by displacement on one of the three phosphorus atoms. The nucleophilic attack may be (*a*) on the terminal phosphorus (P_γ) with displacement of ADP or (*b*) on the internal phosphorus (P_α) with displacement of inorganic pyrophosphate. In the first case, $Y–PO_3H^−$ is formed; in the latter, **Y-adenylyl** (sometimes shortened to Y-adenyl) is formed. More rarely, displacement occurs (*c*) on the central phosphorus (P_β) with transfer of a pyrophospho group to the nucleophile. Still less frequent (*d*) is a displacement on C-5' as shown in Eq. 17-37. If the nucleophile Y in any of these displacement reactions is H_2O, the resulting hydrolysis tends to go to completion, i.e., the phospho, adenylyl, and pyrophospho groups of ATP all have high group transfer potentials (Table 6-6). If Y is an –OH group in an ordinary alcohol the transfer reaction also tends to go to completion because the group transfer potential of a simple phosphate ester is relatively low. Consequently, phosphorylation by ATP is often used as a means of introducing an essentially irreversible step in a metabolic pathway.



ATP

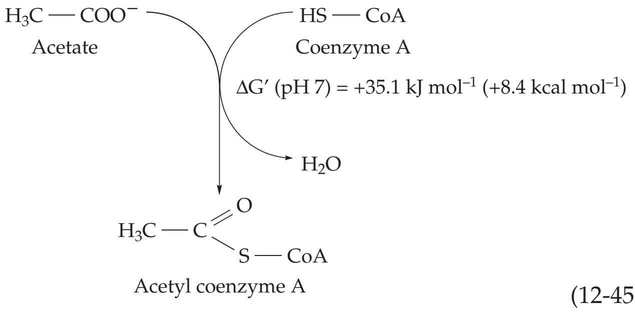
2. Acyl Phosphates

Transfer of a phospho or adenylyl group from ATP to the oxygen atom of a carboxylate group yields an **acyl phosphate**, a type of metabolic intermediate of special significance. Acyl phosphates are mixed anhydrides of carboxylic and phosphoric acids in which *both the acyl group and the phospho group have high group transfer potentials.* As a consequence, acyl phosphates can serve as metabolic intermediates through which the group transfer potential of ATP is transferred into other molecules and is harnessed to do chemical work.



Acyl phosphate

A typical example is the synthesis of acetyl coenzyme A (Eq. 12-45). See Fig. 14-1 for the complete structure of –SH group-containing coenzyme A.

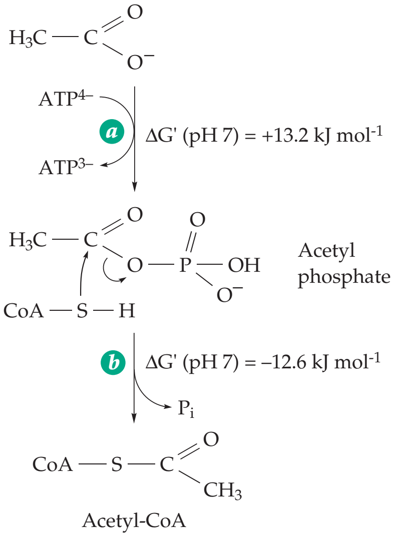


(12-45)

Because the acetyl group in the product also has a high group transfer potential, $\Delta G'$ for this reaction is highly positive and the formation of acetyl-CoA will not occur spontaneously. However, the sum of $\Delta G'$ for the reaction in Eq.12-45 plus that for the hydrolysis of ATP (Eq. 12-46) is nearly zero (+ 0.6kJ mol^{-1}).

$$\begin{aligned} &ATP^{4-} + H_2O \rightarrow ADP^{3-} + HPO_4^{2-} + H^+ \\ &\Delta G' = -34.5\ kJ\ mol^{-1}\ (-8.3\ kcal\ mol^{-1}) \end{aligned} \qquad (12\text{-}46)$$

Coupling of the two reactions is accomplished by first letting an oxygen atom of the nucleophilic carboxylate group attack P_γ of ATP to form acetyl phosphate (Eq. 12-47, step *a*). In the second step (step *b*) the sulfur atom of the –SH group of coenzyme A (often abbreviated CoA–SH) attacks the carbon atom of the acetyl phosphate with displacement of the good leaving group P_i. While $\Delta G'$ for Eq. 12-47, step *a,* is moderately positive (meaning that a relatively low concentration of acetyl phosphate will accumulate unless the [ATP] / [ADP] ratio is high), the equilibrium in Eq. 12-47, step *b,* favors the products.

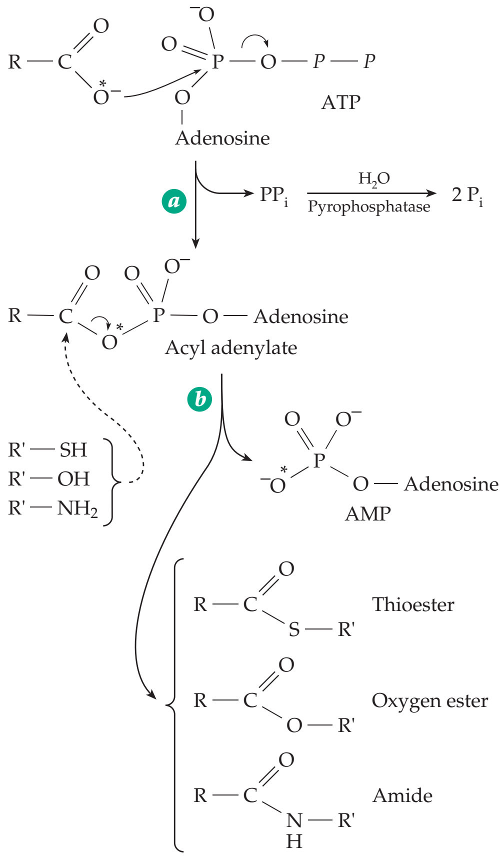


(12-47)

The two reactions of Eq. 12-47 are catalyzed by **acetate kinase**[925] and an ***S*-acetyltransferase**, respectively. The sequence represents an essential first stage in bacterial utilization of acetate for growth. It is also used in a few bacteria in reverse as a way of *generating* ATP in fermentation reactions. On the other hand, most eukaryotic cells make acetyl-CoA from acetate by coupling the synthesis to cleavage of ATP to AMP and P_i. A single enzyme **acetyl-CoA synthetase** (acetate thiokinase) catalyzes both steps in the reaction (Eq. 10-1). The sequence parallels that of Eq. 12-47, but the initial displacement is on P_α of ATP to form **acetyl adenylate**. This intermediate remains tightly bound to the enzyme until the second step in the sequence takes place. When ^{18}O is present in the acetate (designated by the asterisks in Eq. 12-48) it appears in the phospho group of AMP as expected for the indicated mechanism.

3. General Mechanism of Formation of Thioesters, Esters, and Amides

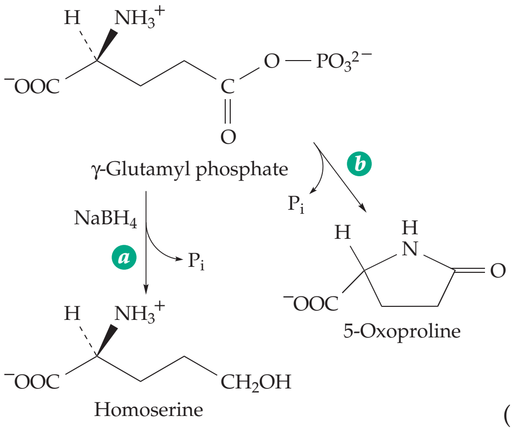
The sequences of Eqs. 12-47 and 12-48 are general ones used by cells for linking carboxylic acids to –OH, –SH, and –NH_2 groups to form oxygen esters, thioesters, or amides, respectively. ATP can be cleaved at either P_α or P_γ. If cleavage is at P_α (Eq. 12-48) the hydrolysis of inorganic pyrophosphate (PP_i) to P_i provides an additional coupling of ATP cleavage to biosynthesis as is discussed in Chapter 17, Section H.



(12-48)

An enzyme catalyzing a reaction similar to that of acetyl-CoA synthetase is **succinyl-CoA synthetase** (succinate thiokinase).[926,927] The enzyme from *E. coli* has been studied most. The first step is formation of a phosphoenzyme by transfer of the γ phospho group from ATP to N^ε of histidine 246 in the α subunit of the 140-kDa $\alpha_2\beta_2$ tetramer.[926] The phospho group is then transferred to succinate to form succinyl phosphate, which reacts with coenzyme A, as in step *b*, Eq. 12-48, to form succinyl-CoA. Crystallographic studies suggest that for this step the succinyl phosphate and the coenzyme A may be bound to opposite α subunits in the tetramer.[926] However, in some bacteria and in eukaryotes the enzyme appears to operate as an αβ heterodimer.[927]

Glutamine synthetase,[928,929] a large enzyme containing 12 identical 468-residue subunits with 622 symmetry (as in Fig. 7-12), has a major regulatory function in nitrogen metabolism, which is discussed in Chapter 24. Apparently, the intermediate acyl phosphate (γ-glutamyl phosphate) has a transient existence and all three reactants–glutamate, NH_4^+, and ATP–must be bound to the enzyme concurrently before the active site becomes functional. Early evidence for the acyl phosphate[930,931] included reduction by sodium borohydride to an alcohol (Eq. 12-49*a*) and isolation of the internal amide of glutamic acid **5-oxoproline** (Eq. 12-49*b*).



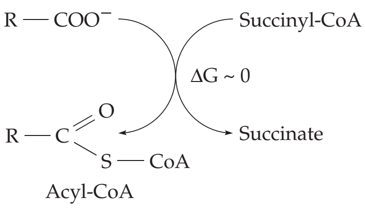
(12-49)

Ammonium ions appear to be bound at a specific site and to be deprotonated by the Asp 50 carboxylate.[929]

The **aminoacyl-tRNA synthetases** join amino acids to their appropriate tRNA molecules for protein synthesis. They have the very important task of selecting both a specific amino acid and a specific tRNA and joining them. The enzymes differ in size and other properties. However, they all appear to function by a common basic chemistry that makes use of cleavage of ATP at P_α (Eq. 12-48) via an intermediate aminoacyl adenylate and that is outlined also in Eq. 17-36. These enzymes are discussed in Chapter 29 .

4. Coenzyme A Transferases

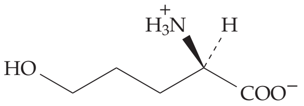
The following problem in energy transfer arises occasionally: A thioester, such as succinyl-CoA, is available to a cell and the energy available in its unstable linkage is needed for synthesis of a different thioester. It would be possible for a cell to first form ATP or GTP, using a synthetase reaction in reverse; then the ATP or GTP formed could be used to make the new linkage by the action of another acyl-CoA synthetase. However, special enzymes, the **CoA transferases**, function more directly (Eq. 12-50). The mechanism is not obvious.



(12-50)

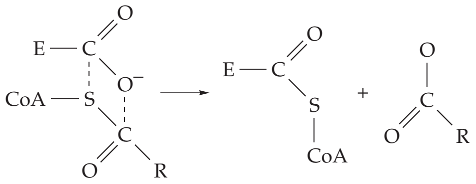
How can the CoA be transferred from one acyl group to another while still retaining the high group transfer potential of the acyl group?

The following experiments shed some light. Kinetic studies of **succinyl-CoA–acetoacetate CoA transferase** indicate a ping-pong mechanism. The enzyme alternates between two distinct forms, one of which has been shown to contain bound CoA.[932–934] The E-CoA intermediate formed from enzyme plus acetoacetyl-CoA was reduced with ^{3}H-containing sodium borohydride and the protein was completely hydrolyzed with HCl. Tritium-containing α-amino-δ-hydroxyvaleric acid was isolated. Since thioesters (as well as oxygen esters) are cleaved in a two-step process to alcohols by reduction with borohydride, it was concluded that the intermediate E-CoA is a thioester of the Glu 344 side chain. In exchange reactions ^{18}O from labeled succinate entered both the E-CoA intermediate and the carboxyl group of acetoacetate.



α-Amino-δ-hydroxyvaleric acid

A mechanism involving formation of a transient anhydride is similar to reactions discussed in preceding sections.[717,720] The student should be able to write out the step-by-step detail. Does this mechanism explain the ^{18}O exchange data? A second possibility is a **4-center reaction** (Eq. 12-51). However, mechanisms of this type have not been demonstrated for enzymatic reactions.



(12-51)

References

1. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology*, Dover, Mineola, New York (pp. 78–110)
2. Lowe, J. N., and Ingraham, L. L. (1974) *An Introduction to Biochemical Reaction Mechanisms*, Prentice-Hall, Englewood Cliffs, New Jersey
3. Bruice, T. C., and Benkovic, S. J. (1966) *Bioorganic Mechanisms*, Benjamin, New York (2 vols.)
3a. Kyte, J. (1995) *Mechanism in Protein Chemistry*, Garland Publ., New York
3b. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, Freeman, New York
3c. Silverman, R. B. (1999) *The Organic Chemistry of Enzyme-Catalyzed Reactions*, Academic Press, San Diego, California
3d. Sinnott, M., ed. (1998) *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I, Academic Press, San Diego, California
4. Dixon, J. E., and Bruice, T. C. (1971) *J. Am. Chem. Soc.* **93**, 6592–6597
5. Kosower, E. M. (1968) *An Introduction to Physical Organic Chemistry*, Wiley, New York (pp. 28 and 81)
6. Goldman, P., Milne, G. W. A., and Keister, D. B. (1968) *J. Biol. Chem.* **243**, 428–434
7. Liu, J.-Q., Kurihara, T., Miyagi, M., Esaki, N., and Soda, K. (1995) *J. Biol. Chem.* **270**, 18309–18312
8. Hisano, T., Hata, Y., Fujii, T., Liu, J.-Q., Kurihara, T., Esaki, N., and Soda, K. (1996) *J. Biol. Chem.* **271**, 20322–20330
9. Li, Y.-F., Hata, Y., Fujii, T., Hisano, T., Nishihara, M., Kurihara, T., and Esaki, N. (1998) *J. Biol. Chem.* **273**, 15035–15044
10. Grimmelikhuijzen, C. J. P., and Schaller, H. C. (1979) *Trends Biochem. Sci.* **4**, 265–267
11. Koonin, E. V., and Tatusov, R. L. (1994) *J. Mol. Biol.* **244**, 125–132
11a. Nardini, M., Ridder, I. S., Rozeboom, H. J., Kalk, K. H., Rink, R., Janssen, D. B., and Dijkstra, B. W. (1999) *J. Biol. Chem.* **274**, 14579–14586
12. Verschueren, K. H. G., Seljée, F., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1993) *Nature (London)* **363**, 693–698
12a. Ridder, I. S., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1999) *J. Biol. Chem.* **274**, 30672–30678
13. Pries, F., Kingma, J., Krooshof, G. H., Jeronimus-Stratingh, C. M., Bruins, A. P., and Janssen, D. B. (1995) *J. Biol. Chem.* **270**, 10405–10411
13a. Schindler, J. F., Naranjo, P. A., Honaberger, D. A., Chang, C.-H., Brainard, J. R., Vanderberg, L. A., and Unkefer, C. J. (1999) *Biochemistry* **38**, 5772–5778
14. Lightstone, F. C., Zheng, Y.-J., Maulitz, A. H., and Bruice, T. C. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8417–8420
15. Lightstone, F. C., Zheng, Y.-J., and Bruice, T. C. (1998) *J. Am. Chem. Soc.* **120**, 5611–5621
16. Pries, F., van den Wijngaard, A. J., Bos, R., Pentenga, M., and Janssen, D. B. (1994) *J. Biol. Chem.* **269**, 17490–17494
17. Chiang, P. K., Gordon, R. K., Tal, J., Zeng, G. C., Doctor, B. P., Pardhasaradhi, K., and McCann, P. P. (1996) *FASEB J.* **10**, 471–480
18. Cornforth, J. W., Reichard, S. A., Talalay, P., Carrell, H. L., and Glusker, J. P. (1977) *J. Am. Chem. Soc.* **99**, 7292–7301
19. Woodard, R. W., Tsai, M.-D., Floss, H. G., Crooks, P. A., and Coward, J. K. (1980) *J. Biol. Chem.* **255**, 9124–9127
20. Vidgren, J., Svensson, L. A., and Liljas, A. (1994) *Nature (London)* **368**, 354–357
21. Fu, Z., Hu, Y., Konishi, K., Takata, Y., Ogawa, H., Gomi, T., Fujioka, M., and Takusagawa, F. (1996) *Biochemistry* **35**, 11985–11993
22. Takata, Y., Konishi, K., Gomi, T., and Fujioka, M. (1994) *J. Biol. Chem.* **269**, 5537–5542
23. Schluckebier, G., O'Gara, M., Saenger, W., and Cheng, X. (1995) *J. Mol. Biol.* **247**, 16–20
24. Szczelkun, M. D., Jones, H., and Connolly, B. A. (1995) *Biochemistry* **34**, 10734–10743
25. O'Gara, M., Roberts, R. J., and Cheng, X. (1996) *J. Mol. Biol.* **263**, 597–606
25a. Reddy, Y. V. R., and Rao, D. N. (2000) *J. Mol. Biol.* **298**, 597–610
26. Patel, D. J. (1994) *Nature (London)* **367**, 688–690
27. Jeltsch, A., Christ, F., Fatemi, M., and Roth, M. (1999) *J. Biol. Chem.* **274**, 19538–19544
28. Lotta, T., Vidgren, J., Tilgmann, C., Ulmanen, I., Melén, K., Julkunen, I., and Taskinen, J. (1995) *Biochemistry* **34**, 4202–4210
29. Zheng, Y.-J., and Bruice, T. C. (1997) *J. Am. Chem. Soc.* **119**, 8137–8145
29a. Kahn, K., and Bruice, T. C. (2000) *J. Am. Chem. Soc.* **122**, 46–51
30. Kealey, J. T., and Santi, D. V. (1995) *Biochemistry* **34**, 2441–2446
31. Hrycyna, C. A., Yang, M. C., and Clarke, S. (1994) *Biochemistry* **33**, 9806–9812
31a. Cleland, W. W., O'Leary, M. H., and Northrop, D. B., eds. (1977) *Isotope Effects on Enzyme-Catalyzed Reactions*, Univ. Park Press, Baltimore, Maryland
31b. Klinman, J. P. (1978) in *Transition States of Biochemical Processes* (Gandour, R. D., and Schowen, R. L., eds), pp. 165–200, Plenum, New York
31c. Gandour, R. D., and Schowen, R. L., eds. (1978) *Transition States of Biochemical Processes*, Plenum, New York
31d. Schramm, V. L., Horenstein, B. A., and Kline, P. C. (1994) *J. Biol. Chem.* **269**, 18259–18262
32. Hegazi, M. F., Borchardt, R. T., and Schowen, R. L. (1979) *J. Am. Chem. Soc.* **101**, 4359–4365
33. Mihel, I., Knipe, J. O., Coward, J. K., and Schowen, R. L. (1979) *J. Am. Chem. Soc.* **101**, 4349–4351
34. Rodgers, J., Femec, D. A., and Schowen, R. L. (1982) *J. Am. Chem. Soc.* **104**, 3263–3268
35. Boyd, R. J., Kim, C.-K., Shi, Z., Weinberg, N., and Wolfe, S. (1993) *J. Am. Chem. Soc.* **115**, 10147–10152
36. Glad, S. S., and Jensen, F. (1997) *J. Am. Chem. Soc.* **119**, 227–232
37. Wolfenden, R., Lu, X., and Young, G. (1998) *J. Am. Chem. Soc.* **120**, 6814–6815
38. Henrissat, B., and Bairoch, A. (1993) *Biochem. J.* **293**, 781–788
39. Henrissat, B., and Bairoch, A. (1996) *Biochem. J.* **316**, 695–696
40. Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.-P., and Davies, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7090–7094
41. Sinnot, M. L. (1990) *Chem. Rev.* **90**, 1171– 1202
42. Withers, S. G., and Aebersold, R. (1995) *Protein Sci.* **4**, 361–372
43. Silverstein, R., Voet, J., Reed, D., and Abeles, R. H. (1967) *J. Biol. Chem.* **242**, 1338–1346
44. Doudoroff, M., Barker, H. A., and Hassid, W. Z. (1947) *J. Biol. Chem.* **168**, 725–732 and 733–746
45. LeBrun, E., and van Rapenbusch, R. (1980) *J. Biol. Chem.* **255**, 12034–12036
46. McCarter, J. D., and Withers, S. G. (1996) *J. Biol. Chem.* **271**, 6889–6894
47. Withers, S. G., Rupitz, K., and Street, I. P. (1988) *J. Biol. Chem.* **17**, 7929–7932
48. Withers, S. G., and Street, I. P. (1988) *J. Am. Chem. Soc.* **110**, 8551–8553
49. Miao, S., Ziser, L., Aebersold, R., and Withers, S. G. (1994) *Biochemistry* **33**, 7027–7032
50. Braun, C., Lindhorst, T., Madsen, N. B., and Withers, S. G. (1996) *Biochemistry* **35**, 5458–5463
50a. Namchuk, M. N., McCarter, J. D., Becalski, A., Andrews, T., and Withers, S. G. (2000) *J. Am. Chem. Soc.* **122**, 1270–1277
51. Gorenstein, D. G., Findlay, J. B., Luxon, B. A., and Kar, D. (1977) *J. Am. Chem. Soc.* **99**, 3473–3479
52. Kirby, A. J. (1983) *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, Springer, Berlin
53. Bennet, A. J., and Sinnott, M. L. (1986) *J. Am. Chem. Soc.* **108**, 7287–7294
54. Sinnott, M. L. (1986) in *Mechanisms of Enzymatic Reactions: Stereochemistry A. Steenbock Symposium* (Frey, P. A., ed), pp. 293–305, Elsevier, New York
55. Phillips, D. C. (1966) *Sci. Am.* **215** (Nov), 78–90
56. Sanz, J. M., García, P., and García, J. L. (1992) *Biochemistry* **31**, 8495–8499
57. Artymicek, P. J., and Blake, C. C. F. (1981) *J. Mol. Biol.* **152**, 737–762
58. Harata, K., Muraki, M., and Jigami, Y. (1993) *J. Mol. Biol.* **233**, 524–535
59. Harata, K., Muraki, M., Hayashi, Y., and Jigami, Y. (1992) *Protein Sci.* **1**, 1447–1453
60. Grütter, M. G., Weaver, L. H., and Matthews, B. W. (1983) *Nature (London)* **303**, 828–831
61. Zhang, X.-J., Baase, W. A., and Matthews, B. W. (1991) *Biochemistry* **30**, 2012–2017
62. Dao-pin, S., Sauer, U., Nicholson, H., and Matthews, B. W. (1991) *Biochemistry* **30**, 7142–7153
63. Matsumura, M., Becktel, W. J., Levitt, M., and Matthews, B. W. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6562–6566
64. Parsons, S. M., and Raftery, M. A. (1972) *Biochemistry* **11**, 1623–1628
65. Maenaka, K., Matsushima, M., Song, H., Sunada, F., Watanabe, K., and Kumagai, I. (1995) *J. Mol. Biol.* **247**, 281–293
66. Matsumura, I., and Kirsch, J. F. (1996) *Biochemistry* **35**, 1881–1889
67. Kuroki, R., Yamada, H., Moriyama, T., and Imoto, T. (1986) *J. Biol. Chem.* **261**, 13571–13574
68. Malcolm, B. A., Rosenberg, S., Corey, M. J., Allen, J. S., de Baetselier, A., and Kirsch, J. F. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 133–137
69. Hadfield, A. T., Harvey, D. J., Archer, D. B., MacKenzie, da, Jeenes, D. J., Radford, S. E., Lowe, G., Dobson, C. M., and Jonhson, L. N. (1994) *J. Mol. Biol.* **243**, 856–872
70. Weaver, L. H., Grütter, M. G., and Matthews, B. W. (1995) *J. Mol. Biol.* **245**, 54–68
71. Kuroki, R., Ito, Y., Kato, Y., and Imoto, T. (1997) *J. Biol. Chem.* **272**, 19976–19981
72. Matsumura, I., and Kirsch, J. F. (1996) *Biochemistry* **35**, 1890–1896
73. Dickerson, R. E., and Geis, I. (1969) *The Structure and Action of Proteins*, Harper and Row, New York (p. 71)
74. Levitt, M. (1974) in *Peptides, Polypeptides and Proteins* (Blout, E. R., Bouey, F. A., Goodman, M., and Lotan, N., eds), pp. 99–113, Wiley, New York
75. Huang, X., Surry, C., Hiebert, T., and Bennet, A. J. (1995) *J. Am. Chem. Soc.* **117**, 10614–10621
76. Matsui, H., Blanchard, J. S., Brewer, C. F., and Hehre, E. J. (1989) *J. Biol. Chem.* **264**, 8714–8716
77. Banait, N. S., and Jencks, W. P. (1991) *J. Am. Chem. Soc.* **113**, 7951–7958
78. Zhang, Y., Bommuswamy, J., and Sinnott, M. L. (1994) *J. Am. Chem. Soc.* **116**, 7557–7563
79. Tanaka, Y., Tao, W., Blanchard, J. S., and Hehre, E. J. (1994) *J. Biol. Chem.* **269**, 32306–32312
80. Knier, B. L., and Jencks, W. P. (1980) *J. Am. Chem. Soc.* **102**, 6789–6798
81. Huang, X., Tanaka, K. S. E., and Bennet, A. J. (1997) *J. Am. Chem. Soc.* **119**, 11147–11154

References

82. Hardy, L. W., and Poteete, A. R. (1991) *Biochemistry* **30**, 9457–9463
82a. Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schülein, M., and Withers, S. G. (1998) *Biochemistry* **37**, 11707–11713
82b. Vocadlo, D. J., Mayer, C., He, S., and Withers, S. G. (2000) *Biochemistry* **39**, 117–126
83. Banerjee, S. K., and Rupley, J. A. (1975) *J. Biol. Chem.* **250**, 8267–8274
84. Kelly, J. A., Sielecki, A. R., Sykes, B. D., James, M. N. G., and Phillips, D. C. (1979) *Nature (London)* **282**, 875–878
85. Metzler, D. E. (1979) *Adv. Enzymol.* **50**, 1–40
86. Dao-pin, S., Liao, D.-I., and Remington, S. J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5361–5365
87. Post, C. B., and Karplus, M. (1986) *J. Am. Chem. Soc.* **108**, 1317–1319
88. Fife, T. H., Jaffe, S. H., and Natarajan, R. (1991) *J. Am. Chem. Soc.* **113**, 7646–7653
89. Piszkiewicz, D., and Bruice, T. C. (1968) *J. Am. Chem. Soc.* **90**, 2156–2163
90. Lowe, G., and Sheppard, G. (1968) *J. Chem. Soc. Chem. Commun.*, 529–530
91. Capon, B. (1969) *Chem. Rev.* **69**, 407–498
92. Hart, P. J., Pfluger, H. D., Monzingo, A. F., Hollis, T., and Robertus, J. D. (1995) *J. Mol. Biol.* **248**, 402–413
93. Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B., and Dijkstra, B. W. (1995) *Biochemistry* **34**, 15619–15623
94. Tews, I., Terwisscha van Scheltinga, A. C., Perrakis, A., Wilson, K. S., and Dijkstra, B. W. (1997) *J. Am. Chem. Soc.* **119**, 7954–7959
95. Brameld, K. A., and Goddard, W. A., III (1998) *J. Am. Chem. Soc.* **120**, 3571–3580
96. Klyosov, A. A. (1990) *Biochemistry* **29**, 10577–10585
97. Watanabe, H., Noda, H., Tokuda, G., and Lo, N. (1998) *Nature (London)* **394**, 330–331
98. Davies, G. J., Tolley, S. P., Henrissat, B., Hjort, C., and Schülein, M. (1995) *Biochemistry* **34**, 16210–16220
99. Damude, H. G., Withers, S. G., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. J. (1995) *Biochemistry* **34**, 2220–2224
100. Barr, B. K., Wolfgang, D. E., Piens, K., Claeyssens, M., and Wilson, D. B. (1998) *Biochemistry* **37**, 9220–9229
101. Divne, C., Ståhlberg, J., Teeri, T. T., and Jones, T. A. (1998) *J. Mol. Biol.* **275**, 309–325
102. Barr, B. K., Hsieh, Y.-L., Ganem, B., and Wilson, D. B. (1996) *Biochemistry* **35**, 586–592
103. Choi, S. K., and Ljungdahl, L. G. (1996) *Biochemistry* **35**, 4906–4910
104. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J.* **15**, 5739–5751
105. Sulzenbacher, G., Schülein, M., and Davies, G. J. (1997) *Biochemistry* **36**, 5902–5911
106. Sulzenbacher, G., Driguez, H., Henrissat, B., Schülein, M., and Davies, G. J. (1996) *Biochemistry* **35**, 15280–15287
107. Johnson, P. E., Joshi, M. D., Tomme, P., Kilburn, D. G., and McIntosh, L. P. (1996) *Biochemistry* **35**, 14381–14394
108. Fukuda, M., Spooncer, E., Oates, J. E., Dell, A., and Klock, J. C. (1984) *J. Biol. Chem.* **259**, 10925–10935
108a. Ponyi, T., Szabó, L., Nagy, T., Orosz, L., Simpson, P. J., Williamson, M. P., and Gilbert, H. J. (2000) *Biochemistry* **39**, 985–991
109. Domínguez, R., Souchon, H., Lascombe, M.-B., and Alzari, P. M. (1996) *J. Mol. Biol.* **257**, 1042–1051
110. Roberge, M., Shareck, F., Morosoli, R., Kluepfel, D., and Dupont, C. (1997) *Biochemistry* **36**, 7769–7775
111. Juy, M., Amit, A. G., Alzari, P. M., Poljak, R. J., Claeyssens, M., Béguin, P., and Aubert, J.-P. (1992) *Nature (London)* **357**, 89–91
112. Chen, L., Fincher, G. B., and Hoj, P. B. (1993) *J. Biol. Chem.* **268**, 13318–13326
113. Chen, L., Garrett, T. P. J., Fincher, G. B., and Hoj, P. B. (1995) *J. Biol. Chem.* **270**, 8093–8101
114. Müller, J. J., Thomsen, K. K., and Heinemann, U. (1998) *J. Biol. Chem.* **273**, 3438–3446
115. Lindley, M. G., Shallenberger, R. S., and Herbert, S. M. (1976) *Food Chemistry* **1**, 149–159
116. Krengel, U., and Dijkstra, B. W. (1996) *J. Mol. Biol.* **263**, 70–78
117. Kasumi, T., Tsumuraya, Y., Brewer, C. F., Kersters-Hilderson, H., Claeyssens, M., and Hehre, E. J. (1987) *Biochemistry* **26**, 3010–3016
118. Lawson, S. L., Wakarchuk, W. W., and Withers, S. G. (1996) *Biochemistry* **35**, 10110–10118
119. Derewenda, U., Swenson, L., Green, R., Wei, Y., Morosoli, R., Shareck, F., Kluepfel, D., and Derewenda, Z. S. (1994) *J. Biol. Chem.* **269**, 20811–20814
120. McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Körner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) *Biochemistry* **35**, 9958–9966
121. He, X. M., and Carter, D. C. (1992) *Nature (London)* **358**, 209–215
122. Braithwaite, K. L., Barna, T., Spurway, T. D., Charnock, S. J., Black, G. W., Hughes, N., Lakey, J. H., Virden, R., Hazlewood, G. P., Henrissat, B., and Gilbert, H. J. (1997) *Biochemistry* **36**, 15489–15500
123. Sanz-Aparicio, J., Hermoso, J. A., Martínez-Ripoll, M., Lequerica, J. L., and Polaina, J. (1998) *J. Mol. Biol.* **275**, 491–502
124. Wang, Q., Trimbur, D., Graham, R., Warren, R. A. J., and Withers, S. G. (1995) *Biochemistry* **34**, 14554–14562
125. Namchuk, M. N., and Withers, S. G. (1995) *Biochemistry* **34**, 16194–16202
126. Hrmova, M., MacGregor, E. A., Biely, P., Stewart, R. J., and Fincher, G. B. (1998) *J. Biol. Chem.* **273**, 11134–11143
127. Febbraio, F., Barone, R., D'Auria, S., Rossi, M., Nucci, R., Piccialli, G., De Napoli, L., Orrù, S., and Pucci, P. (1997) *Biochemistry* **36**, 3068–3075
128. Jacobson, R. H., Zhang, X.-J., DuBose, R. F., and Matthews, B. W. (1994) *Nature (London)* **369**, 761–766
129. Gebler, J. C., Aebersold, R., and Withers, S. G. (1992) *J. Biol. Chem.* **267**, 11126–11130
130. Richard, J. P., Westerfeld, J. G., Lin, S., and Beard, J. (1995) *Biochemistry* **34**, 11713–11724
131. Richard, J. P., Huber, R. E., Lin, S., Heo, C., and Amyes, T. L. (1996) *Biochemistry* **35**, 12377–12386
132. Roth, N. J., Rob, B., and Huber, R. E. (1998) *Biochemistry* **37**, 10099–10107
133. Mitchell, E. P., Withers, S. G., Ermert, P., Vasella, A. T., Garman, E. F., Oikonomakos, N. G., and Johnson, L. N. (1996) *Biochemistry* **35**, 7341–7355
134. Tu, J.-I., Jacobson, G. R., and Graves, D. J. (1971) *Biochemistry* **10**, 1229–1236
135. Gold, A. M., Legrand, E., and Sanchez, G. R. (1971) *J. Biol. Chem.* **246**, 5700–5706
136. Levvy, G. A., and Snaith, S. M. (1972) *Adv. Enzymol.* **36**, 151–181
137. Barford, D., and Johnson, L. N. (1989) *Nature (London)* **340**, 609–616
138. Oikonomakos, N. G., Johnson, L. N., Acharya, K. R., Stuart, D. I., Barford, D., Hajdu, J., Varvill, K. M., Melpidou, A. E., Papageorgiou, T., Graves, D. J., and Palm, D. (1987) *Biochemistry* **26**, 8381–8389
139. Oikonomakos, N. G., Zographos, S. E., Tsitsanou, K. E., Johnson, L. N., and Acharya, K. R. (1996) *Protein Sci.* **5**, 2416–2428
140. Helmreich, E. J. M., and Klein, H. W. (1980) *Angew. Chem. Int. Ed. Engl.* **19**, 441–455
141. Klein, H. W., Im, M. J., Palm, D., and Helmreich, E. J. M. (1984) *Biochemistry* **23**, 5853–5861
142. Street, I. P., Rupitz, K., and Withers, S. G. (1989) *Biochemistry* **28**, 1581–1587
143. Parrish, R. F., Uhing, R. J., and Graves, D. J. (1977) *Biochemistry* **16**, 4824–4831
143a. Watson, K. A., McCleverty, C., Geremia, S., Cottaz, S., Driguez, H., and Johnson, L. N. (1999) *EMBO J.* **18**, 4619–4632
144. Duke, E. M. H., Wakatsuki, S., Hadfield, A., and Johnson, L. N. (1994) *Protein Sci.* **3**, 1178–1196
145. Buchbinder, J. L., Luong, C. B. H., Browner, M. F., and Fletterick, R. J. (1997) *Biochemistry* **36**, 8039–8044
146. Buchbinder, J. L., and Fletterick, R. J. (1996) *J. Biol. Chem.* **271**, 22305–22309
147. Gregoriou, M., Noble, M. E. M., Watson, K. A., Garman, E. F., Krulle, T. M., De La Fuente, C., Fleet, G. W. J., Oikonomakos, N. G., and Johnson, L. N. (1998) *Protein Sci.* **7**, 915–927
148. Nakano, K., and Fukui, T. (1986) *J. Biol. Chem.* **261**, 8230–8236
149. Watson, K. A., Schinzel, R., Palm, D., and Johnson, L. N. (1997) *EMBO J.* **16**, 1–14
150. Becker, S., Palm, D., and Schinzel, R. (1994) *J. Biol. Chem.* **269**, 2485–2490
151. Lin, K., Hwang, P. K., and Fletterick, R. J. (1995) *J. Biol. Chem.* **270**, 26833–26839
152. Goldsmith, E. J., Fletterick, R. J., and Withers, S. G. (1987) *J. Biol. Chem.* **262**, 1449–1455
153. Robyt, J. F., and French, D. (1970) *J. Biol. Chem.* **245**, 3917–3927
154. Brayer, G. D., Luo, Y., and Withers, S. G. (1995) *Protein Sci.* **4**, 1730–1742
155. MacDonald, R. J., Crerar, M. M., Swain, W. F., Pictet, R. L., Thomas, G., and Rutter, W. J. (1980) *Nature (London)* **287**, 117–122
156. Qian, M., Haser, R., Buisson, G., Duée, E., and Payan, F. (1994) *Biochemistry* **33**, 6284–6294
157. Qian, M., Haser, R., and Payan, F. (1995) *Protein Sci.* **4**, 747–755
158. Larson, S. B., Greenwood, A., Cascio, D., Day, J., and McPherson, A. (1994) *J. Mol. Biol.* **235**, 1560–1584
159. Machius, M., Vértesy, L., Huber, R., and Wiegand, G. (1996) *J. Mol. Biol.* **260**, 409–421
160. Kadziola, A., Sogaard, M., Svensson, B., and Haser, R. (1998) *J. Mol. Biol.* **278**, 205–217
161. Strobl, S., Maskos, K., Betz, M., Wiegand, G., Huber, R., Gomis-Rüth, F. X., and Glockshuber, R. (1998) *J. Mol. Biol.* **278**, 617–628
162. Fujimoto, Z., Takase, K., Doui, N., Momma, M., Matsumoto, T., and Mizuno, H. (1998) *J. Mol. Biol.* **277**, 393–407
163. Feller, G., le Bussy, O., Houssier, C., and Gerday, C. (1996) *J. Biol. Chem.* **271**, 23836–23841
164. Aghajari, N., Feller, G., Gerday, C., and Haser, R. (1998) *Protein Sci.* **7**, 564–572
165. Matsuura, Y., Kusunoki, M., Harada, W., and Kakudo, M. (1984) *J. Biochem.* **95**, 697–702
166. Brzozowski, A. M., and Davies, G. J. (1997) *Biochemistry* **36**, 10837–10845
167. Robyt, J. F., and French, D. (1970) *Arch. Biochem. Biophys.* **138**, 662–670
168. Robyt, J. F., and French, D. (1967) *Arch. Biochem. Biophys.* **122**, 8–16
169. delCardayré, S. B., and Raines, R. T. (1994) *Biochemistry* **33**, 6031–6037
169a. Dauter, Z., Dauter, M., Brzozowski, A. M., Christensen, S., Borchert, T. V., Beier, L., Wilson, K. S., and Davies, G. J. (1999) *Biochemistry* **38**, 8385–8392
169b. Park, K. H., Kim, M. J., Lee, H. S., Han, N. S., Kim, D., and Robyt, J. F. (1998) *Carbohydr. Res.* **313**, 235–246

References

170. Hermans, M. M. P., Kroos, M. A., van Beeuman, J., Oostra, B. A., and Reuser, A. J. J. (1991) *J. Biol. Chem.* **266**, 13507–13512
171. Raben, N., Nagaraju, K., Lee, E., Kessler, P., Byrne, B., Lee, L., LaMarca, M., King, C., Ward, J., Sauer, B., and Plotz, P. (1998) *J. Biol. Chem.* **273**, 19086–19092
172. Howard, S., and Withers, S. G. (1998) *Biochemistry* **37**, 3858–3864
173. Jespersen, H. M., MacGregor, E. A., Henrissat, B., Sierks, M. R., and Svensson, B. (1993) *Journal of Protein Chemistry* **12**, 791–805
174. Liu, W., Madsen, N. B., Fan, B., Zucker, K. A., Glew, R. H., and Fry, D. E. (1995) *Biochemistry* **34**, 7056–7061
175. Watanabe, K., Hata, Y., Kizaki, H., Katsube, Y., and Suzuki, Y. (1997) *J. Mol. Biol.* **269**, 142–153
176. Knegtel, R. M. A., Strokopytov, B., Penninga, D., Faber, O. G., Rozeboom, H. J., Kalk, K. H., Dijkhuizen, L., and Dijkstra, B. W. (1995) *J. Biol. Chem.* **270**, 29256–29264
177. Schmidt, A. K., Cottaz, S., Driguez, H., and Schulz, G. E. (1998) *Biochemistry* **37**, 5909–5915

177a. Uitdehaag, J. C. M., Mosi, R., Kalk, K. H., van der Veen, B. A., Dijkhuizen, L., Withers, S. G., and Dijkstra, B. W. (1999) *Nature Struct. Biol.* **6**, 432–436

177b. van der Veen, B. A., Uitdehaag, J. C. M., Penninga, D., van Alebeek, G.-J. W. M., Smith, L. M., Dijkstra, B. W., and Dijkhuizen, L. (2000) *J. Mol. Biol.* **296**, 1027–1038

178. Terada, Y., Yanase, M., Takata, H., Takaha, T., and Okada, S. (1997) *J. Biol. Chem.* **272**, 15729–15733
179. Fierobe, H.-P., Clarke, A. J., Tull, D., and Svensson, B. (1998) *Biochemistry* **37**, 3753–3759
180. Aleshin, A. E., Stoffer, B., Firsov, L. M., Svensson, B., and Honzatko, R. B. (1996) *Biochemistry* **35**, 8319–8328
181. Firsov, L. M., Neustroev, K. N., Aleshin, A. E., Metzler, C. M., Metzler, D. E., Scott, R. D., Stoffer, B., Christensen, T., and Svensson, B. (1994) *Eur. J. Biochem.* **223**, 293–302
182. Natarajan, S. K., and Sierks, M. R. (1997) *Biochemistry* **36**, 14946–14955
183. Williamson, M. P., Le Gal-Coëffet, M.-F., Sorimachi, K., Furniss, C. S. M., Archer, D. B., and Williamson, G. (1997) *Biochemistry* **36**, 7535–7539
184. Penninga, D., van der Veen, B. A., Knegtel, R. M. A., van Hijum, S. A. F. T., Rozeboom, H. J., Kalk, K. H., Dijkstra, B. W., and Dijkhuizen, L. (1996) *J. Biol. Chem.* **271**, 32777–32784

184a. Mikami, B., Adachi, M., Kage, T., Sarikaya, E., Nanmori, T., Shinke, R., and Utsumi, S. (1999) *Biochemistry* **38**, 7050–7061

185. Mikami, B., Degano, M., Hehre, E. J., and Sacchettini, J. C. (1994) *Biochemistry* **33**, 7779–7787
186. Adachi, M., Mikami, B., Katsube, T., and Utsumi, S. (1998) *J. Biol. Chem.* **273**, 19859–19865
187. Amado, M., Almeida, R., Carneiro, F., Levery, S. B., Holmes, E. H., Nomoto, M., Hollingsworth, M. A., Hassan, H., Schwientek, T., Nielsen, P. A., Bennett, E. P., and Clausen, H. (1998) *J. Biol. Chem.* **273**, 12770–12778
188. Baenziger, J. U. (1994) *FASEB J.* **8**, 1019–1025
189. Radzicka, A., and Wolfenden, R. (1996) *J. Am. Chem. Soc.* **118**, 6105–6109
190. Barrett, A. J., Rawlings, N. D., and Woessner, J. F., eds. (1998) *Handbook of Proteolytic Enzymes*, Academic Press, San Diego, California
191. Fushiki, T., and Iwai, K. (1989) *FASEB J.* **3**, 121–126
192. Pignol, D., Granon, S., Chapus, C., and Carlos, J. (1995) *J. Mol. Biol.* **252**, 20–24

192a. Lu, D., Fütterer, K., Korolev, S., Zheng, X., Tan, K., Waksman, G., and Sadler, J. E. (1999) *J. Mol. Biol.* **292**, 361–373

193. Light, A., and Janska, H. (1989) *Trends Biochem. Sci.* **14**, 110–112
194. Matsushima, M., Ichinose, M., Yahagi, N., Kakei, N., Tsukada, S., Miki, K., Kurokawa, K., Tashiro, K., Shiokawa, K., Shinomiya, K., Umeyama, H., Inoue, H., Takahashi, T., and Takahashi, K. (1994) *J. Biol. Chem.* **269**, 19976–19982
195. Kitamoto, Y., Veile, R. A., Donis–Keller, H., and Sadler, J. E. (1995) *Biochemistry* **34**, 4562–4568
196. Khan, A. R., and James, M. N. G. (1998) *Protein Sci.* **7**, 815–836
197. Hedstrom, L., Lin, T.-Y., and Fast, W. (1996) *Biochemistry* **35**, 4515–4523

197a. Khan, A. R., Khazanovich-Bernstein, N., Bergman, E. M., and James, M. N. G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10968–10975

198. Kramer, K. J., Felsted, R. L., and Law, J. H. (1973) *J. Biol. Chem.* **248**, 3021–3028
199. Lesk, A. M., and Fordham, W. D. (1996) *J. Mol. Biol.* **258**, 501–537
200. Cygler, M., Schrag, J. D., Sussman, J. L., Harel, M., Silman, I., Gentry, M. K., and Doctor, B. P. (1993) *Protein Sci.* **2**, 366–382
201. Pereira, P. J. B., Bergner, A., Macedo, R., S, Huber, R., Matschiner, G., Fritz, H., Sommerhoff, C. P., and Bode, W. (1998) *Nature (London)* **392**, 306–311

201a. Sommerhoff, C. P., Bode, W., Pereira, P. J. B., Stubbs, M. T., Stürzebecher, J., Piechottka, G. P., Matschiner, G., and Bergner, A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10984–10991

202. Schechter, N. M., Eng, G. Y., Selwood, T., and McCaslin, D. R. (1995) *Biochemistry* **34**, 10628–10638
203. Johnson, D. A., and Barton, G. J. (1992) *Protein Sci.* **1**, 370–377
204. Kido, H., Yokogoshi, Y., Sakai, K., Tashiro, M., Kishino, Y., Fukutomi, A., and Katunuma, N. (1992) *J. Biol. Chem.* **267**, 13573–13579

204a. Wong, G. W., Tang, Y., Feyfant, E., Sali, A., Li, L., Li, Y., Huang, C., Friend, D. S., Krilis, S. A., and Stevens, R. L. (1999) *J. Biol. Chem.* **274**, 30784–30793

205. Remington, S. J., Woodbury, R. G., Reynolds, R. A., Matthews, B. W., and Neurath, H. (1988) *Biochemistry* **27**, 8097–8105
206. McGrath, M. E., Mirzadegan, T., and Schmidt, B. F. (1997) *Biochemistry* **36**, 14318–14324
207. Sali, A., Matsumoto, R., McNeil, H. P., Karplus, M., and Stevens, R. L. (1993) *J. Biol. Chem.* **268**, 9023–9034
208. Springman, E. B., Dikov, M. M., and Serafin, W. E. (1995) *J. Biol. Chem.* **270**, 1300–1307
209. Hof, P., Mayr, I., Huber, R., Korzus, E., Potempa, J., Travis, J., Powers, J. C., and Bode, W. (1996) *EMBO J.* **15**, 5481–5491
210. Rao, N. V., Wehner, N. G., Marshall, B. C., Gray, W. R., Gray, B. H., and Hoidal, J. R. (1991) *J. Biol. Chem.* **266**, 9540–9548
211. Hershberger, R. J., Gershenfeld, H. K., Weissman, I. L., and Su, L. (1992) *J. Biol. Chem.* **267**, 25488–25493
212. Beresford, P. J., Kam, C.-M., Powers, J. C., and Lieberman, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9285–9290
213. Caputo, A., Garner, R. S., Winkler, U., Hudig, D., and Bleackley, R. C. (1993) *J. Biol. Chem.* **268**, 17672–17675

213a. Harris, J. L., Peterson, E. P., Hudig, D., Thornberry, N. A., and Craik, C. S. (1998) *J. Biol. Chem.* **273**, 27364–27373

214. Fuller, R. S., Brake, A. J., and Thorner, J. (1989) *Science* **246**, 482–486
215. Cromlish, J. A., Seidah, N. G., Marcinkiewicz, M., Hamelin, J., Johnson, D. A., and Chrétien, M. (1987) *J. Biol. Chem.* **262**, 1363–1373
216. Schaner, P., Todd, R. B., Seidah, N. G., and Nillni, E. A. (1997) *J. Biol. Chem.* **272**, 19958–19968
217. Rose, C., Vargas, F., Facchinetti, P., Bourgeat, P., Bambal, R. B., Bishop, P. B., Chan, S. M. T., Moore, A. N. J., Ganellin, C. R., and Schwartz, J.-C. (1996) *Nature (London)* **380**, 403–409
218. Shilton, B. H., Thomas, D. Y., and Cygler, M. (1997) *Biochemistry* **36**, 9002–9012
219. Bullock, T. L., Branchaud, B., and Remington, S. J. (1994) *Biochemistry* **33**, 11127–11134
220. Mortensen, U. H., Remington, S. J., and Breddam, K. (1994) *Biochemistry* **33**, 508–517
221. Brayer, G. D., Delbaere, L. T. J., and James, M. N. G. (1979) *J. Mol. Biol.* **131**, 743–775
222. Plou, F. J., Kowlessur, D., Malthouse, J. P. G., Mellor, G. W., Hartshorn, M. J., Pinitglang, S., Patel, H., Topham, C. M., Thomas, E. W., Verma, C., and Brocklehurst, K. (1996) *J. Mol. Biol.* **257**, 1088–1111
223. Wells, J. A., and Estell, D. A. (1988) *Trends Biochem. Sci.* **13**, 291–297
224. Kossiakoff, A. A., Ultsch, M., White, S., and Eigenbrot, C. (1991) *Biochemistry* **30**, 1211–1221
225. Sorensen, S. B., Bech, L. M., Meldal, M., and Breddam, K. (1993) *Biochemistry* **32**, 8994–8999

225a. Smith, C. A., Toogood, H. S., Baker, H. M., Daniel, R. M., and Baker, E. N. (1999) *J. Mol. Biol.* **294**, 1027–1040

226. Koszelak, S., Ng, J. D., Day, J., Ko, T. P., Greenwood, A., and McPherson, A. (1997) *Biochemistry* **36**, 6597–6604
227. Delbaere, L. T. J., and Brayer, G. D. (1980) *J. Mol. Biol.* **139**, 45–51
228. Blanchard, H., and James, M. N. G. (1994) *J. Mol. Biol.* **241**, 574–587

228a. Tomkinson, B. (1999) *Trends Biochem. Sci.* **24**, 355–359

228b. Rose, C., Vargas, F., Facchinetti, P., Bourgeat, P., Bambal, R. B., Bishop, P. B., Chan, S. M. T., Moore, A. N. J., Ganellin, C. R., and Schwartz, J.-C. (1996) *Nature (London)* **380**, 403–409

228c. Renn, S. C. P., Tomkinson, B., and Taghert, P. H. (1998) *J. Biol. Chem.* **273**, 19173–19182

229. Wharton, C. W. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), pp. 345–379, Academic Press, San Diego, California
230. Blow, D. M. (1997) *Trends Biochem. Sci.* **22**, 405–408
231. Tsukada, H., and Blow, D. M. (1985) *J. Mol. Biol.* **184**, 703–711
232. Bode, W., and Schwager, P. (1975) *J. Mol. Biol.* **98**, 693–717
233. Gaboriaud, C., Serre, L., Guy-Crotte, O., Forest, E., and Fontecilla-Camps, J.-C. (1996) *J. Mol. Biol.* **259**, 995–1010
234. Sawyer, L., Shotten, D. M., Campbell, J. W., Wendell, P. L., Muirhead, H., and Watson, H. C. (1978) *J. Mol. Biol.* **118**, 137–208
235. Bode, W., Meyer, E., Jr., and Powers, J. C. (1989) *Biochemistry* **28**, 1951–1963
236. Stubbs, M. T., and Bode, W. (1995) *Trends Biochem. Sci.* **20**, 23–28
237. Katz, B. A., Liu, B., Barnes, M., and Springman, E. B. (1998) *Protein Sci.* **7**, 875–885
238. Timm, D. E. (1997) *Protein Sci.* **6**, 1418–1425
239. Blow, D. M. (1971) in *The Enzymes*, 3rd ed., Vol. 3 (Boyer, P. D., ed), pp. 185–212, Academic Press, New York
240. Huber, R., and Bode, W. (1977) in *NMR in Biology* (Dwek, R. A., Campbell, I. D., Richards, R. E., and Williams, R. J. P., eds), pp. 1–31 , Academic Press, New York
241. Bachovchin, W. W., and Roberts, J. D. (1978) *J. Am. Chem. Soc.* **100**, 8041–8047

References

242. Smith, S. O., Farr-Jones, S., Griffin, R. G., and Bachovchin, W. W. (1989) *Science* **244**, 961–964
243. Kossiakoff, A. A., and Spencer, S. A. (1981) *Biochemistry* **20**, 6462–6474
244. Polgár, L. (1989) *Mechanisms of Protease Action*, CRC Press, Boca Raton, Florida
245. Warshel, A., Naray-Szabo, G., Sussman, F., and Hwang, J.-K. (1989) *Biochemistry* **28**, 3630–3637
246. Robillard, G., and Schulman, R. G. (1972) *J. Mol. Biol.* **71**, 507–511
247. Markley, J. L., and Westler, W. M. (1996) *Biochemistry* **35**, 11092–11097
248. Ash, E. L., Sudmeier, J. L., De Fabo, E. C., and Bachovchin, W. W. (1997) *Science* **278**, 1128–1132
249. Halkides, C. J., Wu, Y. Q., and Murray, C. J. (1996) *Biochemistry* **35**, 15941–15948
250. Zhong, S., Haghjoo, K., Kettner, C., and Jordan, F. (1995) *J. Am. Chem. Soc.* **117**, 7048–7055
251. Bachovchin, W. W., Wong, W. Y. L., Farr-Jones, S., Shenvi, A. B., and Kettner, C. A. (1988) *Biochemistry* **27**, 7689–7697
252. Cassidy, C. S., Lin, J., and Frey, P. A. (1997) *Biochemistry* **36**, 4576–4584
253. Tsilikounas, E., Rao, T., Gutheil, W. G., and Bachovchin, W. W. (1996) *Biochemistry* **35**, 2437–2444
254. Tobin, J. B., Whitt, S. A., Cassidy, C. S., and Frey, P. A. (1995) *Biochemistry* **34**, 6919–6924
255. Singer, P. T., Smalås, A., Carty, R. P., Mangel, W. F., and Sweet, R. M. (1993) *Science* **259**, 669–673
256. Shieh, H.-S., Kurumbail, R. G., Stevens, A. M., Stegeman, R. A., Sturman, E. J., Pak, J. Y., Wittwer, A. J., Palmier, M. O., Wiegand, R. C., Holwerda, B. C., and Stallings, W. C. (1996) *Nature (London)* **383**, 279–282
257. Tong, L., Quin, C., Massariol, M.-J., Bonneau, P. R., Cordingley, M. G., and Lagacé, L. (1996) *Nature (London)* **383**, 272–275
258. Carter, P., and Wells, J. A. (1988) *Nature (London)* **332**, 564–568
259. Corey, D. R., and Craik, C. S. (1992) *J. Am. Chem. Soc.* **114**, 1784–1790
260. Sprang, S., Standing, T., Fletterick, R. J., Stroud, R. M., Finer-Moore, J., Xuong, N.-H., Hamlin, R., Rutter, W. J., and Craik, C. S. (1987) *Science* **237**, 905–909
261. McGrath, M. E., Vásquez, J. R., Craik, C. S., Yang, A. S., Honig, B., and Fletterick, R. J. (1992) *Biochemistry* **31**, 3059–3064
262. Robertus, J. D., Kraut, J., Alden, R. A., and Birktoft, J. J. (1972) *Biochemistry* **11**, 4293–4303
263. Whiting, A. K., and Peticolas, W. L. (1994) *Biochemistry* **33**, 552–561
264. Bryan, P., Pantoliano, M. W., Quill, S. G., Hsiao, H.-Y., and Poulos, T. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3743–3745
265. Asbóth, B., and Polgár, L. (1983) *Biochemistry* **22**, 117–122
266. Ortiz, C., Tellier, C., Williams, H., Stolowich, N. J., and Scott, A. I. (1991) *Biochemistry* **30**, 10026–10034
267. Bullock, T. L., Breddam, K., and Remington, S. J. (1996) *J. Mol. Biol.* **255**, 714–725
268. Tonge, P. J., and Carey, P. R. (1989) *Biochemistry* **28**, 6701–6709
269. Tonge, P. J., and Carey, P. R. (1992) *Biochemistry* **31**, 9122–9125
270. Birktoft, J. J., Kraut, J., and Freer, S. T. (1976) *Biochemistry* **15**, 4481–4485
271. Brünger, A. T., Huber, R., and Karplus, M. (1987) *Biochemistry* **26**, 5153 –5162
272. Polgár, L., and Halász, R. (1982) *Biochem. J.* **207**, 1–10
273. Bizzozero, S. A., and Dutler, H. (1981) *Bioorg. Chem.* **10**, 46–62
274. Cruickshank, W. H., and Kaplan, H. (1974) *J. Mol. Biol.* **83**, 267–274
274a. Ash, E. L., Sudmeier, J. L., Day, R. M., Vincent, M., Torchilin, E. V., Haddad, K. C., Bradshaw, E. M., Sanford, D. G., and Bachovchin, W. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10371–10376
274b. Kidd, R. D., Sears, P., Huang, D.-H., Witte, K., Wong, C.-H., and Farber, G. K. (1999) *Protein Sci.* **8**, 410–417
275. Meyer, E. F., Jr., Clore, G. M., Gronenborn, A. M., and Hansen, H. A. S. (1988) *Biochemistry* **27**, 725–730
276. Wells, G. B., Mustafi, D., and Makinen, M. W. (1994) *J. Biol. Chem.* **269**, 4577–4586
277. Ding, X., Rasmussen, B. F., Petsko, G. A., and Ringe, D. (1994) *Biochemistry* **33**, 9285–9293
278. Oppenheimer, H. L., Labouesse, B., and Hess, G. P. (1966) *J. Biol. Chem.* **241**, 2720–2730
279. Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., Freeman, San Francisco, California
280. Perona, J. J., Hedstrom, L., Rutter, W. J., and Fletterick, R. J. (1995) *Biochemistry* **34**, 1489–1499
281. Graf, L., Craik, C. S., Patthy, A., Roczniak, S., Fletterick, R. J., and Rutter, W. J. (1987) *Biochemistry* **26**, 2616–2623
282. Schellenberger, V., Turck, C. W., and Rutter, W. J. (1994) *Biochemistry* **33**, 4251–4257
283. Kurth, T., Ullmann, D., Jakubke, H.-D., and Hedstrom, L. (1997) *Biochemistry* **36**, 10098–10104
284. Le Bonniec, B. F., Myles, T., Johnson, T., Knight, C. G., Tapparelli, C., and Stone, S. R. (1996) *Biochemistry* **35**, 7114–7122
285. Wangikar, P. P., Rich, J. O., Clark, D. S., and Dordick, J. S. (1995) *Biochemistry* **34**, 12302–12310
286. Perona, J. J., and Craik, C. S. (1995) *Protein Sci.* **4**, 337–360
287. Davis, J. H., and Agard, D. A. (1998) *Biochemistry* **37**, 7696–7707
288. Baumann, W. K., Bizzozero, S. A., and Dutler, H. (1970) *FEBS Letters* **8**, 257–260
289. Baumann, W. K., Bizzozero, S. A., and Dutler, H. (1973) *Eur. J. Biochem.* **39**, 381–387
290. Stein, R. L., Elrod, J. P., and Schowen, R. L. (1983) *J. Am. Chem. Soc.* **105**, 2446–2452
291. Chang, T. K., Chiang, Y., Guo, H.-X., Kresge, A. J., Mathew, L., Powell, M. F., and Wells, J. A. (1996) *J. Am. Chem. Soc.* **118**, 8802–8807
292. Stein, R. L., and Stimpler, A. M. (1987) *J. Am. Chem. Soc.* **109**, 4387–4390
293. Baker, E. N. (1980) *J. Mol. Biol.* **141**, 441–484
294. Maes, D., Bouckaert, J., Poortmans, F., Wyns, L., and Looze, Y. (1996) *Biochemistry* **35**, 16292–16298
295. O'Hara, B. P., Hemmings, A. M., Buttle, D. J., and Pearl, L. H. (1995) *Biochemistry* **34**, 13190–13195
296. Berti, P. J., and Storer, A. C. (1995) *J. Mol. Biol.* **246**, 273–283
297. Drenth, J., Jansonius, J. N., Koekoek, R., and Wolthers, B. G. (1971) *Adv. Prot. Chem.* **25**, 79–115
298. Lewis, S. D., Johnson, F. A., and Shafer, J. A. (1981) *Biochemistry* **20**, 48–51
299. Roberts, D. D., Lewis, S. D., Ballou, D. P., Olson, S. T., and Shafer, J. A. (1986) *Biochemistry* **25**, 5595–5601
300. Johnson, F. A., Lewis, S. D., and Shafer, J. A. (1981) *Biochemistry* **20**, 44–48
301. Keillor, J. W., and Brown, R. S. (1992) *J. Am. Chem. Soc.* **114**, 7983–7989
302. Pinitglang, S., Watts, A. B., Patel, M., Reid, J. D., Noble, M. A., Gul, S., Bokth, A., Naeem, A., Patel, H., Thomas, E. W., Sreedharan, S. K., Verma, C., and Brocklehurst, K. (1997) *Biochemistry* **36**, 9968–9982
303. Ménard, R., Khouri, H. E., Plouffe, C., Laflamme, P., Dupras, R., Vernet, T., Tessier, D. C., Thomas, D. Y., and Storer, A. C. (1991) *Biochemistry* **30**, 5531–5538
304. Katerelos, N. A., and Goodenough, P. W. (1996) *Biochemistry* **35**, 14763–14772
305. Vernet, T., Tessier, D. C., Chatellier, J., Plouffe, C., Lee, T. S., Thomas, D. Y., Storer, A. C., and Ménard, R. (1995) *J. Biol. Chem.* **270**, 16645–16652
306. Ménard, R., Carrière, J., Laflamme, P., Plouffe, C., Khouri, H. E., Vernet, T., Tessier, D. C., Thomas, D. Y., and Storer, A. C. (1991) *Biochemistry* **30**, 8924–8928
307. Ménard, R., Laflamme, P., Plouffe, C., Vernet, T., Tessier, D. C., Thomas, D. Y., and Storer, A. C. (1995) *Biochemistry* **34**, 464–471
308. Doran, J. D., Tonge, P. J., Mort, J. S., and Carey, P. R. (1996) *Biochemistry* **35**, 12487–12494
309. Gamcsik, M. P., Malthouse, J. P. G., Primrose, W. U., Mackenzie, N. E., Boyd, A. S. F., Russell, R. A., and Scott, A. I. (1983) *J. Am. Chem. Soc.* **105**, 6324–6325
310. Zheng, Y.-J., and Bruice, T. C. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4285–4288
311. Arad, D., Langridge, R., and Kollman, P. A. (1990) *J. Am. Chem. Soc.* **112**, 491–502
312. Brocklehurst, K., Watts, A. B., Patel, M., Verma, C., and Thomas, E. W. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), Academic Press, San Diego, California
313. Bond, J. S., and Butler, P. E. (1987) *Ann. Rev. Biochem.* **56**, 333–364
314. Takahashi, T., Dehdarani, A. H., and Tang, J. (1988) *J. Biol. Chem.* **263**, 10952–10957
315. Coulombe, R., Grochulski, P., Sivaraman, J., Ménard, R., Mort, J. S., and Cygler, M. (1996) *EMBO J.* **15**, 5492–5503
316. Jia, Z., Hasnain, S., Hirama, T., Lee, X., Mort, J. S., To, R., and Huber, C. P. (1995) *J. Biol. Chem.* **270**, 5527–5533
317. Taralp, A., Kaplan, H., Sytwu, I.-I., Vlattas, I., Bohacek, R., Knap, A. K., Hirama, T., Huber, C. P., and Hasnain, S. (1995) *J. Biol. Chem.* **270**, 18036–18043
318. McGrath, M. E., Palmer, J. T., Brömme, D., and Somoza, J. R. (1998) *Protein Sci.* **7**, 1294–1302
319. Shi, G.-P., Webb, A. C., Foster, K. E., Knoll, J. H. M., Lemere, C. A., Munger, J. S., and Chapman, H. A. (1994) *J. Biol. Chem.* **269**, 11530–11536
320. Gelb, B. D., Shi, G.-P., Chapman, H. A., and Desnick, R. J. (1996) *Science* **273**, 1236–1238
321. Ishidoh, K., Muno, D., Sato, N., and Kominami, E. (1991) *J. Biol. Chem.* **266**, 16312–16317
322. Wolters, P. J., Raymond, W. W., Blount, J. L., and Caughey, G. H. (1998) *J. Biol. Chem.* **273**, 15514–15520
323. Schiller, M. R., Mende-Mueller, L., Moran, K., Meng, M., Miller, K. W., and Hook, V. Y. H. (1995) *Biochemistry* **34**, 7988–7995
324. Friedman, T. C., Kline, T. B., and Wilk, S. (1985) *Biochemistry* **24**, 3907–3913
325. Matsui, S.-I., Sandberg, A. A., Negoro, S., Seon, B. K., and Goldstein, G. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1535–1539
326. Joshua-Tor, L., Xu, H. E., Johnston, S. A., and Rees, D. C. (1995) *Science* **269**, 945–950
327. Richards, F. M. (1991) *Sci. Am.* **264** (Jan), 54–63
328. Ohno, S., Emori, Y., Imajoh, S., Kawasaki, H., Kisaragi, M., and Suzuki, K. (1984) *Nature (London)* **312**, 566–570
329. Saido, T. C., Sorimachi, H., and Suzuki, K. (1994) *FASEB J.* **8**, 814–822
330. Chian, H.-L., and Dice, J. F. (1988) *J. Biol. Chem.* **263**, 6797–6805
331. Tompa, P., Baki, A., Schád, É., and Friedrich, P. (1996) *J. Biol. Chem.* **271**, 33161–33164

References

332. Stabach, P. R., Cianci, C. D., Glantz, S. B., Zhang, Z., and Morrow, J. S. (1997) *Biochemistry* **36**, 57 – 65

332a. Hosfield, C. M., Elce, J. S., Davies, P. L., and Jia, Z. (1999) *EMBO J.* **18**, 6880 – 6889

333. Nicholson, D. W., and Thornberry, N. A. (1997) *Trends Biochem. Sci.* **22**, 299 – 306

334. Villa, P., Kaufmann, S. H., and Earnshaw, W. C. (1997) *Trends Biochem. Sci.* **22**, 388 – 393

335. Munday, N. A., Vaillancourt, J. P., Ali, A., Casano, F. J., Miller, D. K., Molineaux, S. M., Yamin, T.-T., Yu, V. L., and Nicholson, D. W. (1995) *J. Biol. Chem.* **270**, 15870 – 15876

336. McGrath, M. E., Eakin, A. E., Engel, J. C., McKerrow, J. H., Craik, C. S., and Fletterick, R. J. (1995) *J. Mol. Biol.* **247**, 251 – 259

337. Gillmor, S. A., Craik, C. S., and Fletterick, R. J. (1997) *Protein Sci.* **6**, 1603 – 1611

338. Porter, W. H., Cunningham, L. W., and Mitchell, W. M. (1971) *J. Biol. Chem.* **246**, 7675 – 7682

339. Pavloff, N., Pemberton, P. A., Potempa, J., Chen, W.-C. A., Pike, R. N., Prochazka, V., Kiefer, M. C., Travis, J., and Barr, P. J. (1997) *J. Biol. Chem.* **272**, 1595 – 1600

339a. Nelson, D., Potempa, J., Kordula, T., and Travis, J. (1999) *J. Biol. Chem.* **274**, 12245 – 12251

340. Malcolm, B. A. (1995) *Protein Sci.* **4**, 1439 – 1445

341. Mosimann, S. C., Cherney, M. M., Sia, S., Plotch, S., and James, M. N. G. (1997) *J. Mol. Biol.* **273**, 1032 – 1047

342. Ding, J., McGrath, W. J., Sweet, R. M., and Mangel, W. F. (1996) *EMBO J.* **15**, 1778 – 1783

343. Fox, T., de Miguel, E., Mort, J. S., and Storer, A. C. (1992) *Biochemistry* **31**, 12571 – 12576

344. Podobnik, M., Kuhelj, R., Turk, V., and Turk, D. (1997) *J. Mol. Biol.* **271**, 774 – 788

345. Seemüller, E., Lupas, A., Stock, D., Löwe, J., Huber, R., and Baumeister, W. (1995) *Science* **268**, 579 – 582

346. Paetzel, M., and Dalbey, R. E. (1997) *Trends Biochem. Sci.* **22**, 28 – 31

346a. Kisselev, A. F., Songyang, Z., and Goldberg, A. L. (2000) *J. Biol. Chem.* **275**, 14831 – 14837

347. Eleuteri, A. M., Kohanski, R. A., Cardozo, C., and Orlowski, M. (1997) *J. Biol. Chem.* **272**, 11824 – 11831

348. Kopp, F., Kristensen, P., Hendil, K. B., Johnsen, A., Sobek, A., and Dahlmann, B. (1995) *J. Mol. Biol.* **248**, 264 – 272

349. Cardozo, C., Michaud, C., and Orlowski, M. (1999) *Biochemistry* **38**, 9768 – 9777

349a. Groll, M., Heinemeyer, W., Jäger, S., Ullrich, T., Bochtler, M., Wolf, D. H., and Huber, R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10976 – 10983

349b. Witt, E., Zantopf, D., Schmidt, M., Kraft, R., Kloetzel, P.-M., and Krüger, E. (2000) *J. Mol. Biol.* **301**, 1 – 9

349c. Orlowski, M., and Wilk, S. (2000) *Arch. Biochem. Biophys.* **383**, 1 – 16

350. Tikkanen, R., Riikonen, A., Oinonen, C., Rouvinen, J., and Peltonen, L. (1996) *EMBO J.* **15**, 2954 – 2960

351. Xuan, J., Tarentino, A. L., Grimwood, B. G., Plummer, T. H., Jr., Cui, T., Guan, C., and Van Roey, P. (1998) *Protein Sci.* **7**, 774 – 781

352. Liu, Y., Guan, C., and Aronson, N. N., Jr. (1998) *J. Biol. Chem.* **273**, 9688 – 9694

353. Guan, C., Liu, Y., Shao, Y., Cui, T., Liao, W., Ewel, A., and Whitaker, R. (1998) *J. Biol. Chem.* **273**, 9695 – 9702

354. Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C. E., Smith, J. L., Tomchick, D. R., and Murzin, A. G. (1995) *Nature (London)* **378**, 416 – 419

355. Dalbey, R. E., and Wickner, W. (1988) *J. Biol. Chem.* **263**, 404 – 408

356. Tschantz, W. R., Sung, M., Delgado-Partin, V. M., and Dalbey, R. E. (1993) *J. Biol. Chem.* **268**, 27349 – 27354

357. Dalbey, R. E., Lively, M. O., Bron, S., and Van Dijl, J. M. (1997) *Protein Sci.* **6**, 1129 – 1138

358. Fenteany, G., and Schreiber, S. L. (1998) *J. Biol. Chem.* **273**, 8545 – 8548

359. Bogyo, M., McMaster, J. S., Gaczynska, M., Tortorella, D., Goldberg, A. L., and Ploegh, H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6629 – 6634

360. Schmidtke, G., Kraft, R., Kostka, S., Henklein, P., Frömmel, C., Löwe, J., Huber, R., Kloetzel, P. M., and Schmidt, M. (1996) *EMBO J.* **15**, 6887 – 6898

361. Heinemeyer, W., Fischer, M., Krimmer, T., Stachon, U., and Wolf, D. H. (1997) *J. Biol. Chem.* **272**, 25200 – 25209

362. Ditzel, L., Huber, R., Mann, K., Heinemeyer, W., Wolf, D. H., and Groll, M. (1998) *J. Mol. Biol.* **279**, 1187 – 1191

363. Andreeva, N. S., Zdanov, A. S., Gustchina, A. E., and Fedorov, A. A. (1984) *J. Biol. Chem.* **259**, 11353 – 11365

364. Fusek, M., and Vetvicka, V., eds. (1995) *Aspartic Proteinases: Physiology and Pathology*, CRC Press, Boca Raton, Florida

365. Moore, S. A., Sielecki, A. R., Chernaia, M. M., Tarasova, N. I., and James, M. N. G. (1995) *J. Mol. Biol.* **247**, 466 – 485

366. Strop, P., Sedlacek, J., Stys, J., Kaderabkova, Z., Blaha, I., Pavlickova, L., Pohl, J., Fabry, M., Kostka, V., Newman, M., Frazao, C., Shearer, A., Tickle, I. J., and Blundell, T. L. (1990) *Biochemistry* **29**, 9863 – 9871

367. Vance, J. E., LeBlanc, D. A., and London, R. E. (1997) *Biochemistry* **36**, 13232 – 13240

368. Majer, P., Collins, J. R., Gulnik, S. V., and Erickson, J. W. (1997) *Protein Sci.* **6**, 1458 – 1466

369. Krieger, T. J., and Hook, V. Y. H. (1992) *Biochemistry* **31**, 4223 – 4231

370. Fineschi, B., and Miller, J. (1997) *Trends Biochem. Sci.* **22**, 377 – 382

371. Wilk, S., Wilk, E., and Magnusson, R. P. (1998) *J. Biol. Chem.* **273**, 15961 – 15970

372. Tong, L., Pav, S., Lamarre, D., Pilote, L., LaPlante, S., Anderson, P. C., and Jung, G. (1995) *J. Mol. Biol.* **250**, 211 – 222

373. Yang, J., Teplyakov, A., and Quail, J. W. (1997) *J. Mol. Biol.* **268**, 449 – 459

374. Gómez, J., and Freire, E. (1995) *J. Mol. Biol.* **252**, 337 – 350

375. Symersky, J., Monod, M., and Foundling, S. I. (1997) *Biochemistry* **36**, 12700 – 12710

376. Green, D. W., Aykent, S., Gierse, J. K., and Zupec, M. E. (1990) *Biochemistry* **29**, 3126 – 3133

377. Baldwin, E. T., Bhat, T. N., Gulnik, S., Hosur, M. V., Sowder, R. C., Cachau, R. E., Collins, J., Silva, A. M., and Erickson, J. W. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6796 – 6800

378. Goldblum, A. (1988) *Biochemistry* **27**, 1653 – 1658

379. Subramanian, E. (1978) *Trends Biochem. Sci.* **3**, 1 – 3

380. Andreeva, N. S., Zdanov, A. S., Gustchina, A. E., and Fedorov, A. A. (1984) *J. Biol. Chem.* **259**, 11353 – 11365

381. James, M. N. G., Sielecki, A. R., Hayakawa, K., and Gelb, M. H. (1992) *Biochemistry* **31**, 3872 – 3886

382. Meek, T. D. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), pp. 327 – 344, Academic Press, San Diego, California

383. Blum, M., Cunningham, A., Pang, H., and Hofmann, T. (1991) *J. Biol. Chem.* **266**, 9501 – 9507

384. James, M. N. G., and Sielecki, A. R. (1986) *Nature (London)* **319**, 33 – 38

385. Blundell, T. L., Lapatto, R., Wilderspin, A. F., Hemmings, A. M., Hobart, P. M., Danley, D. E., and Whittle, P. J. (1990) *Trends Biochem. Sci.* **15**, 425 – 430

386. Katz, R. A., and Skalka, A. M. (1994) *Ann. Rev. Biochem.* **63**, 133 – 173

387. Wlodawer, A., Miller, M., Jaskólski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. H. (1989) *Science* **245**, 616 – 620

388. Rose, R. B., Craik, C. S., Douglas, N. L., and Stroud, R. M. (1996) *Biochemistry* **35**, 12933 – 12944

389. Chatfield, D. C., and Brooks, B. R. (1995) *J. Am. Chem. Soc.* **117**, 5561 – 5572

390. Rodriguez, E. J., Angeles, T. S., and Meek, T. D. (1993) *Biochemistry* **32**, 12380 – 12385

391. Liu, H., Müller-Plathe, F., and van Gunsteren, W. F. (1996) *J. Mol. Biol.* **261**, 454 – 469

392. Silva, A. M., Cachau, R. E., Sham, H. L., and Erickson, J. W. (1996) *J. Mol. Biol.* **255**, 321 – 346

393. Reverter, D., Ventura, S., Villegas, V., Vendrell, J., and Avilés, F. X. (1998) *J. Biol. Chem.* **273**, 3535 – 3541

394. Varlamov, O., Leiter, E. H., and Fricker, L. (1996) *J. Biol. Chem.* **271**, 13981 – 13986

395. Rodríguez, C., Brayton, K. A., Brownstein, M., and Dixon, J. E. (1989) *J. Biol. Chem.* **264**, 5988 – 5995

396. Tan, F., Chan, S. J., Steiner, D. F., Schilling, J. W., and Skidgel, R. A. (1989) *J. Biol. Chem.* **264**, 13165 – 13170

397. Tan, F., Weerasinghe, D. K., Skidgel, R. A., Tamei, H., Kaul, R. K., Roninson, I. B., Schilling, J. W., and Erdös, E. G. (1990) *J. Biol. Chem.* **265**, 13 – 19

398. Ehlers, M. R. W., and Riordan, J. F. (1989) *Biochemistry* **28**, 5311 – 5318

399. Dideberg, O., Charlier, P., Dive, G., Joris, B., Frere, J. M., and Ghuysen, J. M. (1982) *Nature (London)* **299**, 469 – 470

400. Weaver, L. H., Kester, W. R., and Matthews, B. W. (1977) *J. Mol. Biol.* **114**, 119 – 132

401. Hausrath, A. C., and Matthews, B. W. (1994) *J. Biol. Chem.* **269**, 18839 – 18842

402. Holland, D. R., Hausrath, A. C., Juers, D., and Matthews, B. W. (1995) *Protein Sci.* **4**, 1955 – 1965

403. Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H., and Matthews, B. W. (1987) *Biochemistry* **26**, 8542 – 8553

404. Mock, W. L., and Stanford, D. J. (1996) *Biochemistry* **35**, 7369 – 7377

405. Marie-Claire, C., Ruffet, E., Antonczak, S., Beaumont, A., O'Donohue, M., Roques, B. P., and Fournié-Zaluski, M.-C. (1997) *Biochemistry* **36**, 13938 – 13945

406. Lipman, M. L., Panda, D., Bennett, H. P. J., Henderson, J. E., Shane, E., Shen, Y., Goltzman, D., and Karaplis, A. C. (1998) *J. Biol. Chem.* **273**, 13729 – 13737

407. Ogishima, T., Niidome, T., Shimokata, K., Kitada, S., and Ito, A. (1995) *J. Biol. Chem.* **270**, 30322 – 30326

408. Luciano, P., Tokatlidis, K., Chambre, I., Germanique, J.-C., and Géli, V. (1998) *J. Mol. Biol.* **280**, 193 – 199

409. Christianson, D. W., and Lipscomb, W. N. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7568 – 7572

410. Mock, W. L. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), pp. 425 – 453, Academic Press, San Diego, California

411. Campbell, P., and Nashed, N. T. (1982) *J. Am. Chem. Soc.* **104**, 5221 – 5226

412. Makinen, M. W., Fukuyama, J. M., and Kuo, L. C. (1982) *J. Am. Chem. Soc.* **104**, 2667 – 2669

413. Sander, M. E., and Witzel, H. (1985) *Biochem. Biophys. Res. Commun.* **132**, 681 – 687

414. Kuo, L. C., Fukuyama, J. M., and Makinen, M. W. (1983) *J. Mol. Biol.* **163**, 63 – 105

References

415. Suh, J., Hong, S.-B., and Chung, S. (1986) *J. Biol. Chem.* **261**, 7112–7114
416. Mustafi, D., and Makinen, M. W. (1994) *J. Biol. Chem.* **269**, 4587–4595
417. Kuo, L. C., and Makinen, M. W. (1985) *J. Am. Chem. Soc.* **107**, 5255–5261
418. Cleland, W. W. (1977) *Adv. Enzymol.* **45**, 273–286
419. Kim, H., and Lipscomb, W. N. (1990) *Biochemistry* **29**, 5546–5555
420. Osumi, A., Rahmo, A., King, S. W., Przystas, T. J., and Fife, T. H. (1994) *Biochemistry* **33**, 14750–14757
421. Mock, W. L., and Zhang, J. Z. (1991) *J. Biol. Chem.* **266**, 6393–6400
422. Zhang, K., and Auld, D. S. (1995) *Biochemistry* **34**, 16306–16312
423. Hilver, D., Gardell, S. J., Rutter, W. J., and Kaiser, E. T. (1986) *J. Am. Chem. Soc.* **108**, 5298–5304
424. Kam, C.-M., Nishino, N., and Powers, J. C. (1979) *Biochemistry* **18**, 3032–3038
425. Bartlett, P. A., and Marlowe, C. K. (1983) *Biochemistry* **22**, 4618–4624
426. Holden, H. M., and Matthews, B. W. (1988) *J. Biol. Chem.* **263**, 3256–3260
427. Massova, I., Kotra, L. P., Fridman, R., and Mobashery, S. (1998) *FASEB J.* **12**, 1075–1095
428. Bond, J. S., and Beynon, R. J. (1995) *Protein Sci.* **4**, 1247–1261
429. Stöcker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Rüth, F.-X., McKay, D. B., and Bode, W. (1995) *Protein Sci.* **4**, 823–840

429a. Morgunova, E., Tuuttila, A., Bergmann, U., Isupov, M., Lindqvist, Y., Schneider, G., and Tryggvason, K. (1999) *Science* **284**, 1667–1670

429b. Tortorella, M. D., and 27 other authors. (1999) *Science* **284**, 1664–1666

429c. Nagase, H., and Woessner, J. F., Jr. (1999) *J. Biol. Chem.* **274**, 21491–21494

430. Becker, J. W., Marcy, A. I., Rokosz, L. L., Axel, M. G., Burbaum, J. J., Fitzgerald, P. M. D., Cameron, P. M., Esser, C. K., Hagmann, W. K., Hermes, J. D., and Springer, J. P. (1995) *Protein Sci.* **4**, 1966–1976
431. Wetmore, D. R., and Hardman, K. D. (1996) *Biochemistry* **35**, 6549–6558
432. Arumugam, S., Hemme, C. L., Yoshida, N., Suzuki, K., Nagase, H., Berjanskii, M., Wu, B., and Van Doren, S. R. (1998) *Biochemistry* **37**, 9650–9657
433. Gomis-Rüth, F. X., Stöcker, W., Huber, R., Zwilling, R., and Bode, W. (1993) *J. Mol. Biol.* **229**, 945–968
434. Gomis-Rüth, F. X., Kress, L. F., Kellermann, J., Mayr, I., Lee, X., Huber, R., and Bode, W. (1994) *J. Mol. Biol.* **239**, 513–544
435. Llano, E., Pendás, A. M., Knäuper, V., Sorsa, T., Salo, T., Salido, E., Murphy, G., Simmer, J. P., Bartlett, J. D., and López-Otín, C. (1997) *Biochemistry* **36**, 15101–15108
436. Gomis-Rüth, F. X., Gohlke, U., Betz, M., Knäuper, V., Murphy, G., López-Otín, C., and Bode, W. (1996) *J. Mol. Biol.* **264**, 556–566
437. Taylor, A. (1993) *FASEB J.* **7**, 290–298
438. Sträter, N., and Lipscomb, W. N. (1995) *Biochemistry* **34**, 14792–14800

438a. Sträter, N., Sun, L., Kantrowitz, E. R., and Lipscomb, W. N. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11151–11155

439. Chen, G., Edwards, T., D'souza, V. M., and Holz, R. C. (1997) *Biochemistry* **36**, 4278–4286
440. Lowther, W. T., Orville, A. M., Madden, D. T., Lim, S., Rich, D. H., and Matthews, B. W. (1999) *Biochemistry* **38**, 7678–7688
441. Bradshaw, R. A., Brickey, W. W., and Walker, K. W. (1998) *Trends Biochem. Sci.* **23**, 263–267
442. Wilce, M. C. J., Bond, C. S., Dixon, N. E., Freeman, H. C., Guss, J. M., Lilley, P. E., and Wilce, J. A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3472–3477
443. Chan, M. K., Gong, W., Rajagopalan, P. T. R., Hao, B., Tsai, C. M., and Pei, D. (1997) *Biochemistry* **36**, 13904–13909
444. Dardel, F., Ragusa, S., Lazennec, C., Blanquet, S., and Meinnel, T. (1998) *J. Mol. Biol.* **280**, 501–513
445. Becker, A., Schlichting, I., Kabsch, W., Schultz, S., and Wagner, A. F. V. (1998) *J. Biol. Chem.* **273**, 11413–11416
446. Andrews, D. W., and Johnson, A. E. (1996) *Trends Biochem. Sci.* **21**, 365–369

446a. Fabunmi, R. P., Wigley, W. C., Thomas, P. J., and DeMartino, G. N. (2000) *J. Biol. Chem.* **275**, 409–413

447. Sommer, T., and Wolf, D. H. (1997) *FASEB J.* **11**, 1227–1233
448. Ciechanover, A., and Schwartz, A. L. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2727–2730
449. Rubin, D. M., Glickman, M. H., Larsen, C. N., Dhruvakumar, S., and Finley, D. (1998) *EMBO J.* **17**, 4909–4919
450. Lee, Y. S., Park, S. C., Goldberg, A. L., and Chung, C. H. (1988) *J. Biol. Chem.* **263**, 6643–6646
451. Shin, D. H., Lee, C. S., Chung, C. H., and Suh, S. W. (1996) *J. Mol. Biol.* **262**, 71–76
452. Chin, D. T., Goff, S. A., Webster, T., Smith, T., and Goldberg, A. L. (1988) *J. Biol. Chem.* **263**, 11718–11728
453. Flanagan, J. M., Wall, J. S., Capel, M. S., Schneider, D. K., and Shanklin, J. (1995) *Biochemistry* **34**, 10910–10917
454. Kessel, M., Maurizi, M. R., Kim, B., Kocsis, E., Trus, B. L., Singh, S. K., and Steven, A. C. (1995) *J. Mol. Biol.* **250**, 587–594
455. Suzuki, C. K., Rep, M., Maarten van Dijl, J., Suda, K., Grivell, L. A., and Schatz, G. (1997) *Trends Biochem. Sci.* **22**, 118–123
456. Roudiak, S. G., and Shrader, T. E. (1998) *Biochemistry* **37**, 11255–11263
457. Wang, N., Maurizi, M. R., Emmert-Buck, L., and Gottesman, M. M. (1994) *J. Biol. Chem.* **269**, 29308–29313
458. Savel'ev, A. S., Novikova, L. A., Kovaleva, I. E., Luzikov, V. N., Neupert, W., and Langer, T. (1998) *J. Biol. Chem.* **273**, 20596–20602
459. Singh, S. K., Guo, F., and Maurizi, M. R. (1999) *Biochemistry* **38**, 14906–14915
460. Kihara, A., Akiyama, Y., and Ito, K. (1998) *J. Mol. Biol.* **279**, 175–188
461. Lupas, A., Flanagan, J. M., Tamura, T., and Baumeister, W. (1997) *Trends Biochem. Sci.* **22**, 399–404

461a. de Sagarra, M. R., Mayo, I., Marco, S., Rodríguez-Vilarino, S., Oliva, J., Carrascosa, J. L., and Castano, J. G. (1999) *J. Mol. Biol.* **292**, 819–825

461b. Kim, K. I., Cheong, G.-W., Park, S.-C., Ha, J.-S., Woo, K. M., Choi, S. J., and Chung, C. H. (2000) *J. Mol. Biol.* **303**, 655-666

462. Rohrwild, M., Coux, O., Huang, H.-C., Moerschell, R. P., Yoo, S. J., Seol, J. H., Chung, C. H., and Goldberg, A. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5808–5813
463. Missiakas, D., Schwager, F., Betton, J.-M., Georgopoulos, C., and Raina, S. (1996) *EMBO J.* **15**, 6899–6909

463a. Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D., and Huber, R. (2000) *Nature (London)* **403**, 800–805

463b. Shotland, Y., Teff, D., Koby, S., Kobiler, O., and Oppenheim, A. B. (2000) *J. Mol. Biol.* **299**, 953–964

464. Kräusslich, H.-G., and Wimmer, E. (1988) *Ann. Rev. Biochem.* **57**, 701–754
465. Ypma-Wong, M. F., Filman, D. J., Hogle, J. M., and Semler, B. L. (1988) *J. Biol. Chem.* **263**, 17846–17856
466. López-Otín, C., Simón-Mateo, C., Martínez, L., and Vinuela, E. (1989) *J. Biol. Chem.* **264**, 9107–9110

466a. Wu, Z., Yao, N., Le, H. V., and Weber, P. C. (1998) *Trends Biochem. Sci.* **23**, 92–93

466b. Barbato, G., Cicero, D. O., Cordier, F., Narjes, F., Gerlach, B., Sambucini, S., Grzesiek, S., Matassa, V. G., De Francesco, R., and Bazzo, R. (2000) *EMBO J.* **19**, 1195–1206

467. Roseman, J. E., and Levine, R. L. (1987) *J. Biol. Chem.* **262**, 2101–2110
468. Grune, T., Reinheckel, T., and Davies, K. J. A. (1997) *FASEB J.* **11**, 526–534
469. Erdös, E. G., and Skidgel, R. A. (1989) *FASEB J.* **3**, 145–151
470. Hui, K.-S. (1988) *J. Biol. Chem.* **263**, 6613–6618
471. Zisfein, J. B., Graham, R. M., Dreskin, S. V., Wildey, G. M., Fischman, A. J., and Honey, C. J. (1987) *Biochemistry* **26**, 8690–8697

471a. Loew, D., Perrault, C., Morales, M., Moog, S., Ravanat, C., Schuhler, S., Arcone, R., Pietropaolo, C., Cazenave, J.-P., van Dorsselaer, A., and Lanza, F. (2000) *Biochemistry* **39**, 10812–10822

472. Lobe, C. G., Finlay, B. B., Paranchych, W., Paetkau, V. H., and Bleackley, R. C. (1986) *Science* **232**, 858–861
473. Melloni, E., Pontremoki, S., Salamino, F., Sparatore, B., Michetti, M., Sacco, O., and Horecker, B. L. (1986) *J. Biol. Chem.* **261**, 11437–11439
474. Pohlner, J., Halter, R., Beyreuther, K., and Meyer, T. F. (1987) *Nature (London)* **325**, 458–462
475. Newport, G. R., McKerrow, J. H., Hedstrom, R., Petitt, M., McGarrigle, L., Barr, D. J., and Agabian, N. (1988) *J. Biol. Chem.* **263**, 13179–13184
476. Folk, J. E. (1980) *Ann. Rev. Biochem.* **49**, 517–531
477. McGrath, M. E., Gillmor, S. A., and Fletterick, R. J. (1995) *Protein Sci.* **4**, 141–148
478. Salier, J.-P. (1990) *Trends Biochem. Sci.* **15**, 435–439
479. Laskowski, M., Jr., and Kato, I. (1980) *Ann. Rev. Biochem.* **49**, 593–626
480. Barrett, A. J. (1987) *Trends Biochem. Sci.* **12**, 193–196
481. Freije, J. P., Balbín, M., Abrahamson, M., Velasco, G., Dalboge, H., Grubb, A., and López-Otín, C. (1993) *J. Biol. Chem.* **268**, 15737–15744
482. Martin, J. R., Craven, C. J., Jerala, R., Kroon-Zitko, L., Zerovnik, E., Turk, V., and Waltho, J. P. (1995) *J. Mol. Biol.* **246**, 331–343
483. Brown, W. M., and Dziegielewska, K. M. (1997) *Protein Sci.* **6**, 5–12
484. Williamson, R. A., Carr, M. D., Frenkiel, T. A., Feeney, J., and Freedman, R. B. (1997) *Biochemistry* **36**, 13882–13889
485. Apte, S. S., Olsen, B. R., and Murphy, G. (1995) *J. Biol. Chem.* **270**, 14313–14318
486. Baumann, U., Bauer, M., Létoffé, S., Delepelaire, P., and Wandersman, C. (1995) *J. Mol. Biol.* **248**, 653–661
487. Martzen, M. R., McMullen, B. A., Smith, N. E., Fujikawa, K., and Peanasky, R. J. (1990) *Biochemistry* **29**, 7366–7372
488. Huang, K., Lu, W., Anderson, S., Laskowski, M., Jr., and James, M. N. G. (1995) *Protein Sci.* **4**, 1985–1997
489. Lu, W., Qasim, M. A., and Kent, S. B. H. (1996) *J. Am. Chem. Soc.* **118**, 8518–8523
490. Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Diesenhofer, J., and Steigemann, W. (1974) *J. Mol. Biol.* **89**, 73–101

References

491. Mer, G., Hietter, H., Kellenberger, C., Renatus, M., Luu, B., and Lefèvre, J.-F. (1996) *J. Mol. Biol.* **258**, 158–171
492. Conconi, A., and Ryan, C. A. (1993) *J. Biol. Chem.* **268**, 430–435
493. Werner, M. H., and Wemmer, D. E. (1991) *Biochemistry* **30**, 3356–3364
494. McBride, J. D., Brauer, A. B. E., Nievo, M., and Leatherbarrow, R. J. (1998) *J. Mol. Biol.* **282**, 447–457
495. Song, H. K., and Suh, S. W. (1998) *J. Mol. Biol.* **275**, 347–363
496. Brandt, P., and Woodward, C. (1987) *Biochemistry* **26**, 3156–3162
497. Huber, R., and Bode, W. (1978) *Acc. Chem. Res.* **11**, 114–122
498. Potempa, J., Korzus, E., and Travis, J. (1994) *J. Biol. Chem.* **269**, 15957–15960
499. Whisstock, J., Skinner, R., and Lesk, A. M. (1998) *Trends Biochem. Sci.* **23**, 63–67
500. Lukacs, C. M., Rubin, H., and Christianson, D. W. (1998) *Biochemistry* **37**, 3297–3304

500a. Whisstock, J. C., Skinner, R., Carrell, R. W., and Lesk, A. M. (2000) *J. Mol. Biol.* **296**, 685–699

501. Lawrence, D. A., Ginsburg, D., Day, D. E., Berkenpas, M. B., Verhamme, I. M., Kvassman, J.-O., and Shore, J. D. (1995) *J. Biol. Chem.* **270**, 25309–25312
502. Kaslik, G., Kardos, J., Szabó, E., Szilágyi, L., Závodszky, P., Westler, W. M., Markley, J. L., and Gráf, L. (1997) *Biochemistry* **36**, 5455–5464
503. Bode, W., Wei, A.-Z., Huber, R., Meyer, E., Travis, J., and Neumann, S. (1986) *EMBO J.* **5**, 2453–2458
504. Boswell, D. R., and Carrell, R. (1986) *Trends Biochem. Sci.* **11**, 102–103
505. Curiel, D. T., Holmes, M. D., Okayama, H., Brantly, M. L., Vogelmeier, C., Travis, W. D., Stier, L. E., Perks, W. H., and Crystal, R. G. (1989) *J. Biol. Chem.* **264**, 13938–13945
506. Lomas, D. A., Finch, J. T., Seyama, K., Nukiwa, T., and Carrell, R. W. (1993) *J. Biol. Chem.* **268**, 15333–15335
507. Carrell, R. W., Jeppsson, J.-O., Laurell, C.-B., Brennan, S. O., Owen, M. C., Vaughan, L., and Boswell, D. R. (1982) *Nature (London)* **298**, 329–334
508. Carp, H., Miller, F., Hoidal, J. R., and Janoff, A. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2041–2045
509. Radinsky, L. (1983) *Science* **221**, 1187–1191
510. Culliton, B. J. (1989) *Science* **246**, 750–751
511. Matheson, N. R., Gibson, H. E., Hallewell, R. A., Barr, P. J., and Travis, J. (1986) *J. Biol. Chem.* **261**, 10404–10409
512. Rosenfeld, M. A., Siegfried, W., Yoshimura, K., Yoneyama, K., Fukayama, M., Stier, L. E., Pääkkö, P. K., Gilardi, P., Stratford-Perricaudet, L. D., Perricaudet, M., Jallat, S., Pavirani, A., Lecocq, J.-P., and Crystal, R. G. (1991) *Science* **252**, 431–434
513. Owen, M. C., Brennan, S. O., Lewis, J. H., and Carrell, R. W. (1983) *N. Engl. J. Med.* **309**, 694–698
514. Davie, E. W., Fujikawa, K., and Kisiel, W. (1991) *Biochemistry* **30**, 10364–10370
515. Smith, E. L., Hill, R. L., Lehman, I. R., Lefkowitz, R. J., Handler, P., and White, A. (1983) in *Principles of Biochemistry, Mammalian Biochemistry*, 7th ed., pp. 17–37, McGraw-Hill, New York
516. Mann, K. G., Jenny, R. J., and Krishnaswamy, S. (1988) *Ann. Rev. Biochem.* **57**, 915–956
517. Zucker, M. B. (1980) *Sci. Am.* **242** (Jun), 86–103
518. Gibbs, C. S., McCurdy, S. N., Leung, L. L. K., and Paborsky, L. R. (1994) *Biochemistry* **33**, 14003–14010
519. Muller, Y. A., Ultsch, M. H., Kelley, R. F., and de Vos, A. M. (1994) *Biochemistry* **33**, 10864–10870
520. Harlos, K., Martin, D. M. A., O'Brien, D. P., Jones, E. Y., Stuart, D. I., Polikarpov, I., Miller, A., Tuddenham, E. G. D., and Boys, C. W. G. (1994) *Nature (London)* **370**, 662–666
521. Martin, D. M. A., Boys, C. W. G., and Ruf, W. (1995) *FASEB J.* **9**, 852–859
522. Ruf, W., Kelly, C. R., Schullek, J. R., Martin, D. M. A., Polikarpov, I., Boys, C. W. G., Tuddenham, E. G. D., and Edgington, T. S. (1995) *Biochemistry* **34**, 6310–6315
523. Muller, Y. A., Ultsch, M. H., and de Vos, A. M. (1996) *J. Mol. Biol.* **256**, 144–159
524. Doolittle, R. F. (1981) *Sci. Am.* **245** (Dec), 126–135
525. Fu, Y., Weissbach, L., Plant, P. W., Oddoux, C., Cao, Y., Liang, T. J., Roy, S. N., Redman, C. M., and Grieninger, G. (1992) *Biochemistry* **31**, 11968–11972
526. Hunziker, E. B., Straub, P. W., and Haeberli, A. (1990) *J. Biol. Chem.* **265**, 7455–7463
527. Banner, D. W., D'Arcy, A., Chène, C., Winkler, F. K., Guha, A., Konigsberg, W. H., Nemerson, Y., and Kirchhofer, D. (1996) *Nature (London)* **380**, 41–46

527a. Banner, D. W. (2000) *Nature (London)* **404**, 449–450

527b. Chang, Y.-J., Hamaguchi, N., Chang, S.-C., Ruf, W., Shen, M.-C., and Lin, S.-W. (1999) *Biochemistry* **38**, 10940–10948

528. Muranyi, A., Finn, B. E., Gippert, G. P., Forsén, S., Stenflo, J., and Drakenberg, T. (1998) *Biochemistry* **37**, 10605–10615
529. Andrews, B. S. (1991) *Trends Biochem. Sci.* **16**, 31–36
530. Neuenschwander, P. F., and Morrissey, J. H. (1992) *J. Biol. Chem.* **267**, 14477–14482
531. Altieri, D. C. (1995) *FASEB J.* **9**, 860–865
532. Brandstetter, H., Kühne, A., Bode, W., Huber, R., von der Saal, W., Wirthensohn, K., and Engh, R. A. (1996) *J. Biol. Chem.* **271**, 29988–29992
533. Sabharwal, A. K., Padmanabhan, K., Tulinsky, A., Mathur, A., Gorka, J., and Bajaj, S. P. (1997) *J. Biol. Chem.* **272**, 22037–22045
534. Dharmawardana, K. R., and Bock, P. E. (1998) *Biochemistry* **37**, 13143–13152

534a. Macedo-Ribeiro, S., Bode, W., Huber, R., Quinn-Allen, M. A., Kim, S. W., Ortel, T. L., Bourenkov, G. P., Bartunik, H. D., Stubbs, M. T., Kane, W. H., and Fuentes-Prior, P. (1999) *Nature (London)* **402**, 434–439

535. Xue, J., Kalafatis, M., Silveira, J. R., Kung, C., and Mann, K. G. (1994) *Biochemistry* **33**, 13109–13116
536. Comfurius, P., Smeets, E. F., Willems, G. M., Bevers, E. M., and Zwaal, R. F. A. (1994) *Biochemistry* **33**, 10319–10324
537. Betz, A., and Krishnaswamy, S. (1998) *J. Biol. Chem.* **273**, 10709–10718
538. Bottenus, R. E., Ichinose, A., and Davie, E. W. (1990) *Biochemistry* **29**, 11195–11209
539. Slon-Usakiewicz, J. J., Purisima, E., Tsuda, Y., Sulea, T., Pedyczak, A., Féthière, J., Cygler, M., and Konishi, Y. (1997) *Biochemistry* **36**, 13494–13502
540. Hughes, P. E., Morgan, G., Rooney, E. K., Brownlee, G. G., and Handford, P. (1993) *J. Biol. Chem.* **268**, 17727–17733

540a. Mathur, A., and Bajaj, S. P. (1999) *J. Biol. Chem.* **274**, 18477– 18486

540b. Tolkatchev, D., Ng, A., Zhu, B., and Ni, F. (2000) *Biochemistry* **39**, 10365–10372

541. Xu, W.-f, Chung, D. W., and Davie, E. W. (1996) *J. Biol. Chem.* **271**, 27948–27953

541a. Brown, J. H., Volkmann, N., Jun, G., Henschen-Edman, A. H., and Cohen, C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 85–90

542. Doolittle, R. F., Everse, S. J., and Spraggon, G. (1996) *FASEB J.* **10**, 1464–1470
543. Everse, S. J., Spraggon, G., Veerapandian, L., Riley, M., and Doolittle, R. F. (1998) *Biochemistry* **37**, 8637–8642
544. Spraggon, G., Everse, S. J., and Doolittle, R. F. (1997) *Nature (London)* **389**, 455–462
545. Murthy, S. N. P., Wilson, J. H., Lukas, T. J., Veklich, Y., Weisel, J. W., and Lorand, L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 44–48
546. Pedersen, L. C., Yee, V. C., Bishop, P. D., Le Trong, I., Teller, D. C., and Stenkamp, R. E. (1994) *Protein Sci.* **3**, 1131–1135
547. Bernardo, M. M., Day, D. E., Olson, S. T., and Shore, J. D. (1993) *J. Biol. Chem.* **268**, 12468–12476
548. Beaubien, G., Rosinski-Chupin, I., Mattei, M. G., Mbikay, M., Chrétien, M., and Seidah, N. G. (1991) *Biochemistry* **30**, 1628–1635
549. Herwald, H., Renné, T., Meijers, J. C. M., Chung, D. W., Page, J. D., Colman, R. W., and Müller-Esterl, W. (1996) *J. Biol. Chem.* **271**, 13061–13067
550. Baglia, F. A., Jameson, B. A., and Walsh, P. N. (1993) *J. Biol. Chem.* **268**, 3838–3844
551. Hamaguchi, N., Charifson, P. S., Pedersen, L. G., Brayer, G. D., Smith, K. J., and Stafford, D. W. (1991) *J. Biol. Chem.* **266**, 15213–15220
552. Kurachi, S., Hitomi, Y., Furukawa, M., and Kurachi, K. (1995) *J. Biol. Chem.* **270**, 5276–5281

552a. Kurachi, S., Deyashiki, Y., Takeshita, J., and Kurachi, K. (1999) *Science* **285**, 739–743

553. Freedman, S. J., Furie, B. C., Furie, B., and Baleja, J. D. (1995) *Biochemistry* **34**, 12126–12137
554. Freedman, S. J., Blostein, M. D., Baleja, J. D., Jacobs, M., Furie, B. C., and Furie, B. (1996) *J. Biol. Chem.* **271**, 16227–16236
555. Lenting, P. J., Christophe, O. D., ter Maat, H., Rees, D. J. G., and Mertens, K. (1996) *J. Biol. Chem.* **271**, 25332–25337
556. Tagliavacca, L., Moon, N., Dunham, W. R., and Kaufman, R. J. (1997) *J. Biol. Chem.* **272**, 27428–27434
557. Pittman, D. D., Wang, J. H., and Kaufman, R. J. (1992) *Biochemistry* **31**, 3315–3325
558. Gilbert, G. E., and Baleja, J. D. (1995) *Biochemistry* **34**, 3022–3031
559. Gilbert, G. E., and Drinkwater, D. (1993) *Biochemistry* **32**, 9577–9585
560. Leyte, A., van Schijndel, H. B., Niehrs, C., Huttner, W. B., Verbeet, M. P., Mertens, K., and van Mourik, J. A. (1991) *J. Biol. Chem.* **266**, 740–746
561. Gitschier, J., Wood, W. I., Shuman, M. A., and Lawn, R. M. (1986) *Science* **232**, 1415 – 1416
562. Titani, K., and Walsh, K. A. (1988) *Trends Biochem. Sci.* **13**, 94–97
563. Sadler, J. E. (1998) *Ann. Rev. Biochem.* **67**, 395–424
564. Ruggeri, Z. M., and Ware, J. (1993) *FASEB J.* **7**, 308–316
565. Huizinga, E. G., van der Plas, R. M., Kroon, J., Sixma, J. J., and Gros, P. (1997) *Structure* **5**, 1147–1156
566. Emsley, J., Cruz, M., Handin, R., and Liddington, R. (1998) *J. Biol. Chem.* **273**, 10396–10401
567. Sugimoto, M., Dent, J., McClintock, R., Ware, J., and Ruggeri, Z. M. (1993) *J. Biol. Chem.* **268**, 12185–12192
568. George, J. N., Nurden, A. T., and Phillips, D. R. (1984) *N. Engl. J. Med.* **311**, 1084–1098
569. Cruz, M. A., Yuan, H., Lee, J. R., Wise, R. J., and Handin, R. I. (1995) *J. Biol. Chem.* **270**, 10822–10827
570. Beacham, D. A., Wise, R. J., Turci, S. M., and Handin, R. I. (1992) *J. Biol. Chem.* **267**, 3409–3415
571. Zhang, Z. P., Blombäck, M., Nyman, D., and Anvret, M. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7937–7940

References

572. Kaufman, R. J. (1989) *Nature (London)* **342**, 207–208
573. Dwarki, V. J., Belloni, P., Nijjar, T., Smith, J., Couto, L., Rabier, M., Clift, S., Berns, A., and Cohen, L. K. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1023–1027
574. Kay, M. A., Rothenberg, S., Landen, C. N., Bellinger, D. A., Leland, F., Toman, C., Finegold, M., Thompson, A. R., Read, M. S., Brinkhous, K. M., and Woo, S. L. C. (1993) *Science* **262**, 117–119
575. Jesty, J., Wun, T.-C., and Lorenz, A. (1994) *Biochemistry* **33**, 12686–12694
576. Hajjar, K. A. (1994) *N. Engl. J. Med.* **331**, 1585–1587
577. Björk, I., Nordling, K., and Olson, S. T. (1993) *Biochemistry* **32**, 6501–6505
578. Streusand, V. J., Björk, I., Gettins, P. G. W., Petitou, M., and Olson, S. T. (1995) *J. Biol. Chem.* **270**, 9043–9051
578a. Whisstock, J. C., Pike, R. N., Jin, L., Skinner, R., Pei, X. Y., Carrell, R. W., and Lesk, A. M. (2000) *J. Mol. Biol.* **301**, 1287–1305
579. Di Cera, E., Guinto, E. R., Vindigni, A., Dang, Q. D., Ayala, Y. M., Wuyi, M., and Tulinsky, A. (1995) *J. Biol. Chem.* **270**, 22089–22092
580. Weisel, J. W., Nagaswami, C., Young, T. A., and Light, D. R. (1996) *J. Biol. Chem.* **271**, 31485–31490
580a. Fuentes-Prior, P., Iwanaga, Y., Huber, R., Pagila, R., Rumennik, G., Seto, M., Morser, J., Light, D. R., and Bode, W. (2000) *Nature (London)* **404**, 518–525
580b. Baerga-Ortiz, A., Rezaie, A. R., and Komives, E. A. (2000) *J. Mol. Biol.* **296**, 651–658
581. Esmon, C. T. (1995) *FASEB J.* **9**, 946 – 955
582. Colpitts, T. L., Prorok, M., and Castellino, F. J. (1995) *Biochemistry* **34**, 2424–2430
582a. Shen, L., Dahlbäck, B., and Villoutreix, B. O. (2000) *Biochemistry* **39**, 2853–2860
583. Öhlin, A., Landes, G., Bourden, P., Oppenheimer, C., Wydro, R., and Stenflo, J. (1988) *J. Biol. Chem.* **263**, 19240–19248
584. Plow, E. F., Herren, T., Redlitz, A., Miles, L. A., and Hoover-Plow, J. L. (1995) *FASEB J.* **9**, 939–945
585. Menhart, N., Hoover, G. J., McCance, S. G., and Castellino, F. J. (1995) *Biochemistry* **34**, 1482–1488
586. Bendixen, E., Harpel, P. C., and Sottrup-Jensen, L. (1995) *J. Biol. Chem.* **270**, 17929–17933
587. Pirie-Shepherd, S. R., Jett, E. A., Andon, N. L., and Pizzo, S. V. (1995) *J. Biol. Chem.* **270**, 5877–5881
588. Marti, D. N., Hu, C.-K., An, S. S. A., von Haller, P., Schaller, J., and Llinás, M. (1997) *Biochemistry* **36**, 11591–11604
589. Bogusky, M. J., Dobson, C. M., and Smith, R. A. G. (1989) *Biochemistry* **28**, 6728–6735
590. Byeon, I.-J. L., Kelley, R. F., Mulkerrin, M. G., An, S. S. A., and Llinás, M. (1995) *Biochemistry* **34**, 2739–2750
591. Parry, M. A. A., Zhang, X. C., and Bode, W. (2000) *Trends Biochem. Sci.* **25**, 53–59
592. Oates, J. A., Wood, A. J. J., Loscalzo, J., and Braunwald, E. (1988) *N. Engl. J. Med.* **319**, 925–931
593. Mottonen, J., Strand, A., Symersky, J., Sweet, R. M., Danley, D. E., Geoghegan, K. F., Gerard, R. D., and Goldsmith, E. J. (1992) *Nature (London)* **355**, 270–273
594. Szyperski, T., Antuch, W., Schick, M., Betz, A., Stone, S. R., and Wüthrich, K. (1994) *Biochemistry* **33**, 9303–9310
595. Betz, A., Hofsteenge, J., and Stone, S. R. (1992) *Biochemistry* **31**, 4557–4562
596. Vitali, J., Martin, P. D., Malkowski, M. G., Robertson, W. D., Lazar, J. B., Winant, R. C., Johnson, P. H., and Edwards, B. F. P. (1992) *J. Biol. Chem.* **267**, 17670–17678
597. van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kröger, B., Höffken, W., and Bode, W. (1995) *EMBO J.* **14**, 5149–5157
598. Stark, K. R., and James, A. A. (1998) *J. Biol. Chem.* **273**, 20802–20809
599. Lim-Wilby, M. S. L., Hallenga, K., De Maeyer, M., Lasters, I., Vlasuk, G. P., and Brunck, T. K. (1995) *Protein Sci.* **4**, 178–186
600. van de Locht, A., Stubbs, M. T., Bode, W., Friedrich, T., Bollschweiler, C., Höffken, W., and Huber, R. (1996) *EMBO J.* **15**, 6011–6017
601. Tabernero, L., Chang, C. Y. Y., Ohringer, S. L., Lau, W. F., Iwanowicz, E. J., Han, W.-C., Wang, T. C., Seiler, S. M., Roberts, D. G. M., and Sack, J. S. (1995) *J. Mol. Biol.* **246**, 14–20
602. Cheng, Y., Slon-Usakiewicz, J. J., Wang, J., Purisima, E. O., and Konishi, Y. (1996) *Biochemistry* **35**, 13021–13029
603. Krishnan, R., Tulinsky, A., Vlasuk, G. P., Pearson, D., Vallar, P., Bergum, P., Brunck, T. K., and Ripka, W. C. (1996) *Protein Sci.* **5**, 422–433
604. Kamata, K., Kawamoto, H., Honma, T., Iwama, T., and Kim, S.-H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6630–6635
605. Lu, D., Bovill, E. G., and Long, G. L. (1994) *J. Biol. Chem.* **269**, 29032–29038
606. Ichinose, A. (1992) *Biochemistry* **31**, 3113–3118
607. Olds, R. J., Lane, D. A., Chowdhury, V., De Stefano, V., Leone, G., and Thein, S. L. (1993) *Biochemistry* **32**, 4216–4224
608. Watton, J., Longstaff, C., Lane, D. A., and Barrowcliffe, T. W. (1993) *Biochemistry* **32**, 7286–7293
609. Taylor, P. (1991) *J. Biol. Chem.* **266**, 4025–4028
610. Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. (1991) *Science* **253**, 872–879
611. Maelicke, A. (1991) *Trends Biochem. Sci.* **16**, 355–356
611a. Koellner, G., Kryger, G., Millard, C. B., Silman, I., Sussman, J. L., and Steiner, T. (2000) *J. Mol. Biol.* **296**, 713–735
612. Vellom, D. C., Radic', Z., Li, Y., Pickering, N. A., Camp, S., and Taylor, P. (1993) *Biochemistry* **32**, 12–17
613. Kovach, I. M., Huber, J. H.-A., and Schowen, R. L. (1988) *J. Am. Chem. Soc.* **110**, 590–593
614. Radic, Z., Kirchhoff, P. D., Quinn, D. M., McCammon, J. A., and Taylor, P. (1997) *J. Biol. Chem.* **272**, 23265–23277
615. Wlodek, S. T., Antosiewicz, J., and Briggs, J. M. (1997) *J. Am. Chem. Soc.* **119**, 8159–8165
616. Haas, R., Marshall, T. L., and Rosenberry, T. L. (1988) *Biochemistry* **27**, 6453–6457
617. Krejci, E., Thomine, S., Boschetti, N., Legay, C., Sketelj, J., and Massoulié, J. (1997) *J. Biol. Chem.* **272**, 22840–22847
618. Pindel, E. V., Kedishvili, N. Y., Abraham, T. L., Brzezinski, M. R., Zhang, J., Dean, R. A., and Bosron, W. F. (1997) *J. Biol. Chem.* **272**, 14769–14775
619. Li, J., Szittner, R., Derewenda, Z. S., and Meighen, E. A. (1996) *Biochemistry* **35**, 9967–9973
620. Derewenda, Z. S., and Sharp, A. M. (1993) *Trends Biochem. Sci.* **18**, 20–25
621. Quinn, D. M., and Feaster, S. R. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), pp. 455–482, Academic Press, San Diego, California
622. Bourne, Y., Martinez, C., Kerfelec, B., Lombardo, D., Chapus, C., and Cambillau, C. (1994) *J. Mol. Biol.* **238**, 709–732
623. Egloff, M.-P., Marguet, F., Buono, G., Verger, R., Cambillau, C., and van Tilbeurgh, H. (1995) *Biochemistry* **34**, 2751–2762
624. Dahim, M., and Brockman, H. (1998) *Biochemistry* **37**, 8369–8377
625. Jentoft, N. (1990) *Trends Biochem. Sci.* **15**, 291–294
626. Chen, J. C.-H., Miercke, L. J. W., Krucinski, J., Starr, J. R., Saenz, G., Wang, X., Spilburg, C. A., Lange, L. G., Ellsworth, J. L., and Stroud, R. M. (1998) *Biochemistry* **37**, 5107–5117
627. Gargouri, Y., Moreau, H., Pieroni, G., and Verger, R. (1988) *J. Biol. Chem.* **263**, 2159–2162
628. Warner, T. G., Dambach, L. M., Shin, J. H., and O'Brien, J. S. (1981) *J. Biol. Chem.* **256**, 2952–2957
629. Emmerich, J., Beg, O. U., Peterson, J., Previato, L., Brunzell, J. D., Brewer, J., HB, and Santamarina-Fojo, S. (1992) *J. Biol. Chem.* **267**, 4161–4165
630. Kobayashi, J., Applebaum-Bowden, D., Dugi, K. A., Brown, D. R., Kashyap, V. S., Parrott, C., Duarte, C., Maeda, N., and Santamarina-Fojo, S. (1996) *J. Biol. Chem.* **271**, 26296–26301
631. Shen, W.-J., Patel, S., Natu, V., and Kraemer, F. B. (1998) *Biochemistry* **37**, 8973–8979
632. Nicolas, A., Egmond, M., Verrips, C. T., de Vlieg, J., Longhi, S., Cambillau, C., and Martinez, C. (1996) *Biochemistry* **35**, 398–410
633. Grochulski, P., Bouthillier, F., Kazlauskas, R. J., Serreqi, A. N., Schrag, J. D., Ziomek, E., and Cygler, M. (1994) *Biochemistry* **33**, 3494–3500
634. Lang, D., Hofmann, B., Haalck, L., Hecht, H.-J., Spener, F., Schmid, R. D., and Schomburg, D. (1996) *J. Mol. Biol.* **259**, 704–717
635. Uppenberg, J., Öhrner, N., Norin, M., Hult, K., Kleywegt, G. J., Patkar, S., Waagen, V., Anthonsen, T., and Jones, T. A. (1995) *Biochemistry* **34**, 16838–16851
636. Brocca, S., Schmidt-Dannert, C., Lotti, M., Alberghina, L., and Schmid, R. D. (1998) *Protein Sci.* **7**, 1415–1422
637. Ordentlich, A., Barak, D., Kronman, C., Ariel, N., Segall, Y., Velan, B., and Shafferman, A. (1998) *J. Biol. Chem.* **273**, 19509–19517
638. Axelsen, P. H., Harel, M., Silman, I., and Sussman, J. L. (1994) *Protein Sci.* **3**, 188–197
639. Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E., Dodson, G., Tolley, S., Turkenburg, J. P., Christiansen, L., Huge-Jensen, B., Norskov, L., Thim, L., and Menge, U. (1990) *Nature (London)* **343**, 767–770
640. Dennis, E. A. (1997) *Trends Biochem. Sci.* **22**, 1–2
641. Balsinde, J., and Dennis, E. A. (1997) *J. Biol. Chem.* **272**, 16069–16072
642. Leslie, C. C. (1997) *J. Biol. Chem.* **272**, 16709–16712
643. Scott, D. L., White, S. P., Otwinowski, Z., Yuan, W., Gelb, M. H., and Sigler, P. B. (1990) *Science* **250**, 1541–1546
644. Brunie, S., Bolin, J., Gewirth, D., and Sigler, P. B. (1985) *J. Biol. Chem.* **260**, 9742–9749
645. Scott, D. L., Achari, A., Vidal, J. C., and Sigler, P. B. (1992) *J. Biol. Chem.* **267**, 22645–22657
646. van den Berg, B., Tessari, M., de Haas, G. H., Verheij, H. M., Boelens, R., and Kaptein, R. (1995) *EMBO J.* **14**, 4123–4131
647. Gelb, M. H., Jain, M. K., Hanel, A. M., and Berg, O. G. (1995) *Ann. Rev. Biochem.* **64**, 653–688
648. Thuren, T., Tulkki, A.-P., Virtanen, J. A., and Kinnunen, P. K. J. (1987) *Biochemistry* **26**, 4907–4910
649. Huber, M., Yee, V. C., Burri, N., Vikerfors, E., Lavrijsen, A. P. M., Paller, A. S., and Hohl, D. (1997) *J. Biol. Chem.* **272**, 21018–21026
650. Yee, V. C., Pedersen, L. C., Le Trong, I., Bishop, P. D., Stenkamp, R. E., and Teller, D. C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7296–7300

651. Kulkarni, M. S., and Sherman, F. (1994) *J. Biol. Chem.* **269**, 13141–13147

652. Wu, D., and Hersh, L. B. (1995) *J. Biol. Chem.* **270**, 29111–29116

653. Schwartz, B., and Drueckhammer, D. G. (1996) *J. Am. Chem. Soc.* **118**, 9826–9830

654. Ellis, J., Bagshaw, C. R., and Shaw, W. V. (1995) *Biochemistry* **34**, 16852–16859

655. Murray, I. A., Cann, P. A., Day, P. J., Derrick, J. P., Sutcliffe, M. J., Shaw, W. V., and Leslie, A. G. W. (1995) *J. Mol. Biol.* **254**, 993–1005

656. Guillaume, G., Vanhove, M., Lamotte-Brasseur, J., Ledent, P., Jamin, M., Joris, B., and Frère, J.-M. (1997) *J. Biol. Chem.* **272**, 5438–5444

657. Maveyraud, L., Pratt, R. F., and Samama, J.-P. (1998) *Biochemistry* **37**, 2622–2628

658. Crowder, M. W., Wang, Z., Franklin, S. L., Zovinka, E. P., and Benkovic, S. J. (1996) *Biochemistry* **35**, 12126–12132

659. Orellano, E. G., Girardini, J. E., Cricco, J. A., Ceccarelli, E. A., and Vila, A. J. (1998) *Biochemistry* **37**, 10173–10180

660. Lubkowski, J., Wlodawer, A., Ammon, H. L., Copeland, T. D., and Swain, A. L. (1994) *Biochemistry* **33**, 10257–10265

661. Stewart, A. E., Arfin, S. M., and Bradshaw, R. A. (1995) *J. Biol. Chem.* **270**, 25–28

662. Nakamura, N., Inoue, N., Watanabe, R., Takahashi, M., Takeda, J., Stevens, V. L., and Kinoshita, T. (1997) *J. Biol. Chem.* **272**, 15834–15840

663. Knowles, J. R. (1980) *Ann. Rev. Biochem.* **49**, 877–919

664. Westheimer, F. H. (1968) *Acc. Chem. Res.* **1**, 70–78

665. Westheimer, F. H. (1987) *Science* **235**, 1173–1178

666. Hengge, A. C. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), pp. 517–542, Academic Press, San Diego, California

667. Vincent, J. B., Crowder, M. W., and Averill, B. A. (1992) *Trends Biochem. Sci.* **17**, 105–110

668. Mislow, K. (1970) *Acc. Chem. Res.* **3**, 321–331

669. Bunton, C. A. (1970) *Acc. Chem. Res.* **3**, 257–265

670. Gorenstein, D. G., Luxon, B. A., and Findlay, J. B. (1979) *J. Am. Chem. Soc.* **101**, 5869–5875

671. Taira, K., Mock, W. L., and Gorenstein, D. G. (1984) *J. Am. Chem. Soc.* **106**, 7831–7835

672. Calvo, K. C., and Westheimer, F. H. (1984) *J. Am. Chem. Soc.* **106**, 4205–4210

673. Friedman, J. M., Freeman, S., and Knowles, J. R. (1988) *J. Am. Chem. Soc.* **110**, 1268–1275

674. Wu, Y.-D., and Houk, K. N. (1993) *J. Am. Chem. Soc.* **115**, 11997–12002

675. Buchwald, S. L., Friedman, J. M., and Knowles, J. R. (1984) *J. Am. Chem. Soc.* **106**, 4911–4916

676. Herschlag, D., and Jencks, W. P. (1989) *J. Am. Chem. Soc.* **111**, 7579–7586

677. Hengge, A. C., Edens, W. A., and Elsing, H. (1994) *J. Am. Chem. Soc.* **116**, 5045–5049

678. Jankowski, S., Quin, L. D., Paneth, P., and O'Leary, M. H. (1994) *J. Am. Chem. Soc.* **116**, 11675–11677

679. Herschlag, D., and Jencks, W. P. (1990) *Biochemistry* **29**, 5172–5179

680. Admiraal, S. J., and Herschlag, D. (1999) *J. Am. Chem. Soc.* **121**, 5837–5845

681. Cohn, M., and Reed, G. H. (1982) *Ann. Rev. Biochem.* **51**, 365–394

682. Villafranca, J. J., and Raushel, F. M. (1980) *Annu Rev Biophys Bioeng.* **9**, 363–392

683. McLaughlin, A. C., Leigh, J. S., Jr., and Cohn, M. (1976) *J. Biol. Chem.* **251**, 2777–2787

684. Leyh, T. S., Goodhart, P. J., Nguyen, A. C., Kenyon, G. L., and Reed, G. H. (1985) *Biochemistry* **24**, 308–316

685. Nageswara Rao, B. D., Kayne, F. J., and Cohn, M. (1979) *J. Biol. Chem.* **254**, 2689–2696

686. Stackhouse, J., Nambiar, K. P., Burbaum, J. J., Stauffer, D. M., and Benner, S. A. (1985) *J. Am. Chem. Soc.* **107**, 2757–2763

687. Rao, B. D. N., and Cohn, M. (1981) *J. Biol. Chem.* **256**, 1716–1721

688. Gadian, D. G., Radda, G. K., Brown, T. R., Chance, E. M., Dawson, M. J., and Wilke, D. R. (1981) *Biochem. J.* **194**, 215–228

689. Mendz, G. L., Robinson, G., and Kuchel, P. W. (1986) *J. Am. Chem. Soc.* **108**, 169–173

690. Midelfort, C. F., and Rose, I. A. (1976) *J. Biol. Chem.* **251**, 5881–5887

691. Lowe, G., and Tuck, S. P. (1986) *J. Am. Chem. Soc.* **108**, 1300–1301

692. Abbott, S. J., Jones, S. R., Weinman, S. A., Bockhoff, F. M., McLafferty, F. W., and Knowles, J. R. (1979) *J. Am. Chem. Soc.* **101**, 4323–4332

693. Pliura, D. H., Schomburg, D., Richard, J. P., Frey, P. A., and Knowles, J. R. (1980) *Biochemistry* **19**, 325–329

694. Sammons, R. D., and Frey, P. A. (1982) *J. Biol. Chem.* **257**, 1138–1141

695. Eckstein, F. (1980) *Trends Biochem. Sci.* **5**, 157–159

696. Baraniak, J., and Frey, P. A. (1988) *J. Am. Chem. Soc.* **110**, 4059–4060

697. Richard, J. P., and Frey, P. A. (1982) *J. Am. Chem. Soc.* **104**, 3476–3481

698. Hassett, A., Blättler, W., and Knowles, J. R. (1982) *Biochemistry* **21**, 6335–6340

699. Mehdi, S., and Gerlt, J. A. (1984) *Biochemistry* **23**, 4844–4852

700. Sammons, R. D., Frey, P. A., Bruzik, K., and Tsai, M.-D. (1983) *J. Am. Chem. Soc.* **105**, 5455–5461

701. Buchwald, S. L., and Knowles, J. R. (1980) *J. Am. Chem. Soc.* **102**, 6602–6604

702. Cornelius, R. D., and Cleland, W. W. (1978) *Biochemistry* **17**, 3279–3286

703. Dunaway-Mariano, D., and Cleland, W. W. (1980) *Biochemistry* **19**, 1506–1515

704. Speckhard, D. C., Pecoraro, V. L., Knight, W. B., and Cleland, W. W. (1986) *J. Am. Chem. Soc.* **108**, 4167–4171

705. Lin, I., and Dunaway-Mariano, D. (1988) *J. Am. Chem. Soc.* **110**, 950–956

706. Burgers, P. M. J., and Eckstein, F. (1980) *J. Biol. Chem.* **255**, 8229–8233

707. Lester, L. M., Rusch, L. A., Robinson, G. J., and Speckhard, D. C. (1998) *Biochemistry* **37**, 5349–5355

708. Matlin, A. R., Kendall, D. A., Carano, K. S., Banzon, J. A., Klecka, S. B., and Solomon, N. M. (1992) *Biochemistry* **31**, 8196–8200

709. Simopoulos, T. T., and Jencks, W. P. (1994) *Biochemistry* **33**, 10375–10380

710. Han, R., and Coleman, J. E. (1995) *Biochemistry* **34**, 4238–4245

711. Craig, D. B., Arriaga, E. A., Wong, J. C. Y., Lu, H., and Dovichi, N. J. (1996) *J. Am. Chem. Soc.* **118**, 5245–5253

712. Kimura, E., Kodama, Y., Koike, T., and Shiro, M. (1995) *J. Am. Chem. Soc.* **117**, 8304–8311

713. Stec, B., Hehir, M. J., Brennan, C., Nolte, M., and Kantrowitz, E. R. (1998) *J. Mol. Biol.* **277**, 647–662

713a. Stec, B., Holtz, K. M., and Kantrowitz, E. R. (2000) *J. Mol. Biol.* **299**, 1303–1311

714. Florián, J., and Warshel, A. (1997) *J. Am. Chem. Soc.* **119**, 5473–5474

715. Herschman, H. R. (1980) *Trends Biochem. Sci.* **5**, 82–84

716. Ostanin, K., and Van Etten, R. L. (1993) *J. Biol. Chem.* **268**, 20778–20784

717. VanEtten, R. L., Waymack, P. P., and Rehkop, D. M. (1974) *J. Am. Chem. Soc.* **96**, 6783–6785

718. Buchwald, S. L., Saini, M. S., Knowles, J. R., and Van Etten, R. L. (1984) *J. Biol. Chem.* **259**, 2208–2213

719. Ma, L., and Kantrowitz, E. R. (1996) *Biochemistry* **35**, 2394–2402

720. Sugiura, Y., Kawabe, H., Tanaka, H., Fujimoto, S., and Ohara, A. (1981) *J. Biol. Chem.* **256**, 10664–10670

721. Hong, S.-B., Kuo, J. M., Mullins, L. S., and Raushel, F. M. (1995) *J. Am. Chem. Soc.* **117**, 7580–7581

722. Benning, M. M., Kuo, J. M., Raushel, F. M., and Holden, H. M. (1995) *Biochemistry* **34**, 7973–7978

723. Vanhooke, J. L., Benning, M. M., Raushel, F. M., and Holden, H. M. (1996) *Biochemistry* **35**, 6020–6025

723a. Benning, M. M., Hong, S.-B., Raushel, F. M., and Holden, H. M. (2000) *J. Biol. Chem.* **275**, 30556–30560

724. Kuo, J. M., Chae, M. Y., and Raushel, F. M. (1997) *Biochemistry* **36**, 1982–1988

724a. Shim, H., and Raushel, F. M. (2000) *Biochemistry* **39**, 7357–7364

725. Williams, N. H. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), pp. 543–561, Academic Press, San Diego, California

726. Villeret, V., Huang, S., Zhang, Y., and Lipscomb, W. N. (1995) *Biochemistry* **34**, 4307–4315

727. Kurbanov, F. T., Choe, J.-y, Honzatko, R. B., and Fromm, H. J. (1998) *J. Biol. Chem.* **273**, 17511–17516

728. Choe, J.-Y., Fromm, H. J., and Honzatko, R. B. (2000) *Biochemistry* **39**, 8565–8574

728a. Lee, Y.-H., Ogata, C., Pflugrath, J. W., Levitt, D. G., Sarma, R., Banaszak, L. J., and Pilkis, S. J. (1996) *Biochemistry* **35**, 6010–6019

729. Lu, G., Giroux, E. L., and Kantrowitz, E. R. (1997) *J. Biol. Chem.* **272**, 5076–5081

730. Villeret, V., Huang, S., Zhang, Y., Xue, Y., and Lipscomb, W. N. (1995) *Biochemistry* **34**, 4299–4306

731. Lei, K.-J., Pan, C.-J., Liu, J.-L., Shelly, L. L., and Chou, J. Y. (1995) *J. Biol. Chem.* **270**, 11882–11886

732. Clottes, E., and Burchell, A. (1998) *J. Biol. Chem.* **273**, 19391–19397

733. Josse, J., and Wong, S. C. K. (1971) in *The Enzymes*, 3rd ed., Vol. 4 (Boyer, P. D., ed), pp. 499–541, Academic Press, New York

734. Harutyunyan, E. H., Oganessyan, V. Y., Oganessyan, N. N., Avaeva, S. M., Nazarova, T. I., Vorobyeva, N. N., Kurilova, S. A., Huber, R., and Mather, T. (1997) *Biochemistry* **36**, 7754–7760

735. Salminen, A., Efimova, I. S., Parfenyev, A. N., Magretova, N. N., Mikalahti, K., Goldman, A., Baykov, A. A., and Lahti, R. (1999) *J. Biol. Chem.* **274**, 33898–33904

736. Pohjanjoki, P., Lahti, R., Goldman, A., and Cooperman, B. S. (1998) *Biochemistry* **37**, 1754–1761

736a. Lappänen, V.-M., Nummelin, H., Hansen, T., Lahti, R., Schäfer, G., and Goldman, A. (1999) *Protein Sci.* **8**, 1218–1231

736b. Baykov, A. A., Fabrichniy, I. P., Pohjanjoki, P., Zyryanov, A. B., and Lahti, R. (2000) *Biochemistry* **39**, 11939–11947

737. Barford, D. (1996) *Trends Biochem. Sci.* **21**, 407–412

738. Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A. C., and Kuriyan, J. (1995) *Nature (London)* **376**, 745–753

739. Egloff, M.-P., Cohen, P. T. W., Reinemer, P., and Barford, D. (1995) *J. Mol. Biol.* **254**, 942–959

740. Zhang, J., Zhang, Z., Brew, K., and Lee, E. Y. C. (1996) *Biochemistry* **35**, 6276–6282

References

741. Zhang, M., Zhou, M., Van Etten, R. L., and Stauffacher, C. V. (1997) *Biochemistry* **36**, 15–23
742. Hengge, A. C., Sowa, G. A., Wu, L., and Zhang, Z.-Y. (1995) *Biochemistry* **34**, 13982–13987
743. Yuvaniyama, J., Denu, J. M., Dixon, J. E., and Saper, M. A. (1996) *Science* **272**, 1328–1331
744. D'Alessio, G., and Riordan, J. F., eds. (1997) *Ribonuclease Structure and Functions*, Academic Press, San Diego, California
745. Moore, S., and Stein, W. H. (1973) *Science* **180**, 458–464
746. Wlodawer, A., Svensson, L. A., Sjölin, L., and Gilliland, G. L. (1988) *Biochemistry* **27**, 2705–2717
747. Santoro, J., González, C., Bruix, M., Neira, J. L., Nieto, J. L., Herranz, J., and Rico, M. (1993) *J. Mol. Biol.* **229**, 722–734
748. Tilton, R. F., Jr., Dewan, J. C., and Petsko, G. A. (1992) *Biochemistry* **31**, 2469–2481
749. Fedorov, A. A., Joseph-McCarthy, D., Fedorov, E., Sirakova, D., Graf, I., and Almo, S. C. (1996) *Biochemistry* **35**, 15962–15979
750. Boix, E., Nogués, M. V., Schein, C. H., Benner, S. A., and Cuchillo, C. M. (1994) *J. Biol. Chem.* **269**, 2529–2534
751. Schultz, L. W., Quirk, D. J., and Raines, R. T. (1998) *Biochemistry* **37**, 8886–8898
752. Jackson, D. Y., Burnier, J., Quan, C., Stanley, M., Tom, J., and Wells, J. A. (1994) *Science* **266**, 243–247
753. Gutte, B. (1977) *J. Biol. Chem.* **252**, 663 – 670
754. Breslow, R., Dong, S. D., Webb, Y., and Xu, R. (1996) *J. Am. Chem. Soc.* **118**, 6588–6600
755. Martin, P. D., Coscha, M. S., and Edwards, B. F. P. (1987) *J. Biol. Chem.* **262**, 15930–15938
756. Wladkowski, B. D., Svensson, L. A., Sjolin, L., Ladner, J. E., and Gilliland, G. L. (1998) *J. Am. Chem. Soc.* **120**, 5488–5498
757. Eckstein, F. (1979) *Acc. Chem. Res.* **12**, 204–210
758. Matta, M. S., and Vo, D. T. (1986) *J. Am. Chem. Soc.* **108**, 5316–5318
759. Herschlag, D. (1994) *J. Am. Chem. Soc.* **116**, 11631–11635
760. Breslow, R., and Chapman, W. H., Jr. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10018–10021
761. Sowa, G. A., Hengge, A. C., and Cleland, W. W. (1997) *J. Am. Chem. Soc.* **119**, 2319–2320
762. Kawata, Y., Sakiyama, F., Hayashi, F., and Kyogoku, Y. (1990) *Eur. J. Biochem.* **187**, 255–262
763. Pletinckx, J., Steyaert, J., Zegers, I., Choe, H.-W., Heinemann, U., and Wyns, L. (1994) *Biochemistry* **33**, 1654–1662
764. Kurihara, H., Nonaka, T., Mitsui, Y., Ohgi, K., Irie, M., and Nakamura, K. T. (1996) *J. Mol. Biol.* **255**, 310–320
764a. Noguchi, S., Satow, Y., Uchida, T., Sasaki, C., and Matsuzaki, T. (1995) *Biochemistry* **34**, 15583–15591
765. Cordes, F., Starikov, E. B., and Saenger, W. (1995) *J. Am. Chem. Soc.* **117**, 10365–10372
766. Buckle, A. M., and Fersht, A. R. (1994) *Biochemistry* **33**, 1644–1653
767. Deutscher, M. P. (1988) *Trends Biochem. Sci.* **13**, 136–139
768. Benner, S. A., and Alleman, R. K. (1989) *Trends Biochem. Sci.* **14**, 396–397
769. Davies, J. F., II, Hostomska, Z., Hostomsky, Z., Jordan, S. R., and Matthews, D. A. (1991) *Science* **252**, 88–95
770. Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Ikehara, M., Matsuzaki, T., and Morikawa, K. (1990) *Nature (London)* **347**, 306–309
771. Yang, W., Hendrickson, W. A., Crouch, R. J., and Satow, Y. (1990) *Science* **249**, 1398–1405
772. Acharya, K. R., Shapiro, R., Riordan, J. F., and Vallee, B. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2949–2953
773. Lequin, O., Albaret, C., Bontems, F., Spik, G., and Lallemand, J.-Y. (1996) *Biochemistry* **35**, 8870–8880
774. Shapiro, R. (1998) *Biochemistry* **37**, 6847–6856
775. Rosenberg, H. F., Tenen, D. G., and Ackerman, S. J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4460–4464
776. Leland, P. A., Schultz, L. W., Kim, B.-M., and Raines, R. T. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10407–10412
777. D'Alessio, G., Di Donato, A., Parente, A., and Piccoli, R. (1991) *Trends Biochem. Sci.* **16**, 104–106
778. Hofsteenge, J., Kieffer, B., Matthies, R., Hemmings, B. A., and Stone, S. R. (1988) *Biochemistry* **27**, 8537–8544
779. Kole, R., and Altman, S. (1981) *Biochemistry* **20**, 1902–1906
780. Altman, S., Baer, M., Guerrier-Takada, C., and Vioque, A. (1986) *Trends Biochem. Sci.* **11**, 515–518
781. Altman, S., Kirsebom, L., and Talbot, S. (1993) *FASEB J.* **7**, 7–14
782. Yuan, Y., and Altman, S. (1995) *EMBO J.* **14**, 159–168
782a. Massire, C., Jaeger, L., and Westhof, E. (1998) *J. Mol. Biol.* **279**, 773–793
783. Waugh, D. S., Green, C. J., and Pace, N. R. (1989) *Science* **244**, 1569–1571
784. True, H. L., and Celander, D. W. (1998) *J. Biol. Chem.* **273**, 7193–7196
785. Cech, T. R. (1987) *Science* **236**, 1532–1539
786. Zaug, A. J., and Cech, T. R. (1986) *Science* **231**, 470–475
787. Cech, T. R. (1986) *Sci. Am.* **255** (Nov), 64–75
788. Brown, J. W., Haas, E. S., Gilbert, D. G., and Pace, N. R. (1994) *Nucleic Acids Res.* **22**, 3660–3662
789. Pan, T. (1995) *Biochemistry* **34**, 902–909
790. Golden, B. L., Gooding, A. R., Podell, E. R., and Cech, T. R. (1998) *Science* **282**, 259–264
791. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science* **273**, 1678–1685
792. Ferré-D'Amaré, A. R., Zhou, K., and Doudna, J. A. (1998) *Nature (London)* **395**, 567–574
793. Baidya, N., and Uhlenbeck, O. C. (1997) *Biochemistry* **36**, 1108–1114
794. Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) *Nature (London)* **372**, 68–74
795. Scott, W. G., Murray, J. B., Arnold, J. R. P., Stoddard, B. L., and Klug, A. (1996) *Science* **274**, 2065–2069
796. Bevers, S., Xiang, G., and McLaughlin, L. W. (1996) *Biochemistry* **35**, 6483–6490
796a. Bevers, S., Ha, S. B., and McLaughlin, L. W. (1999) *Biochemistry* **38**, 7710–7718
796b. Kore, A. R., and Eckstein, F. (1999) *Biochemistry* **38**, 10915–10918
796c. Lyne, P. D., and Karplus, M. (2000) *J. Am. Chem. Soc.* **122**, 166–167
796d. Torres, R. A., and Bruice, T. C. (2000) *J. Am. Chem. Soc.* **122**, 781–791
797. Simorre, J.-P., Legault, P., Baidya, N., Uhlenbeck, O. C., Maloney, L., Wincott, F., Usman, N., Beigelman, L., and Pardi, A. (1998) *Biochemistry* **37**, 4034–4044
798. Sargueil, B., Pecchia, D. B., and Burke, J. M. (1995) *Biochemistry* **34**, 7739–7748
798a. Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 920–930
798b. Muth, G. W., Ortoleva-Donnelly, L., and Strobel, S. A. (2000) *Science* **289**, 947–950
798c. Nakano, S.-i, Chadalavada, D. M., and Bevilacqua, P. C. (2000) *Science* **287**, 1493–1497
799. Scott, W. G., and Klug, A. (1996) *Trends Biochem. Sci.* **21**, 220–224
800. Grasby, J. A. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), pp. 563–571, Academic Press, San Diego, California
801. Hampel, K. J., Walter, N. G., and Burke, J. M. (1998) *Biochemistry* **37**, 14672–14682
802. Pan, T., and Uhlenbeck, O. C. (1992) *Nature (London)* **358**, 560–563
803. Chartrand, P., Usman, N., and Cedergren, R. (1997) *Biochemistry* **36**, 3145–3150
804. Santoro, S. W., and Joyce, G. F. (1998) *Biochemistry* **37**, 13330–13342
805. Beebe, J. A., Kurz, J. C., and Fierke, C. A. (1996) *Biochemistry* **35**, 10493–10505
806. Torres, R. A., and Bruice, T. C. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11077–11082
807. Pyle, A. M. (1993) *Science* **261**, 709–714
808. Peracchi, A., Beigelman, L., Scott, E. C., Uhlenbeck, O. C., and Herschlag, D. (1997) *J. Biol. Chem.* **272**, 26822–26826
809. Pontius, B. W., Lott, W. B., and von Hippel, P. H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2290–2294
810. Bruice, T. C., Tsubouchi, A., Dempcy, R. O., and Olson, L. P. (1996) *J. Am. Chem. Soc.* **118**, 9867–9875
811. Weinstein, L. B., Jones, B. C. N. M., Cosstick, R., and Cech, T. R. (1997) *Nature (London)* **388**, 805–808
812. Christian, E. L., and Yarus, M. (1993) *Biochemistry* **32**, 4475–4480
813. Orita, M., Vinayak, R., Andrus, A., Warashina, M., Chiba, A., Kaniwa, H., Nishikawa, F., Nishikawa, S., and Taira, K. (1996) *J. Biol. Chem.* **271**, 9447–9454
814. Hertel, K. J., Peracchi, A., Uhlenbeck, O. C., and Herschlag, D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8497–8502
815. Peracchi, A., Karpeisky, A., Maloney, L., Beigelman, L., and Herschlag, D. (1998) *Biochemistry* **37**, 14765–14775
816. Joyce, G. F. (1992) *Sci. Am.* **267** (Dec), 90–97
817. Ekland, E. H., Szostak, J. W., and Bartel, D. P. (1995) *Science* **269**, 364–370
818. Wilson, C., and Szostak, J. W. (1995) *Nature (London)* **374**, 777–782
819. Wright, M. C., and Joyce, G. F. (1997) *Science* **276**, 614–617
820. Ohmichi, T., and Sugimoto, N. (1997) *Biochemistry* **36**, 3514–3521
821. Suga, H., Cowan, J. A., and Szostak, J. W. (1998) *Biochemistry* **37**, 10118–10125
822. Tsang, J., and Joyce, G. F. (1996) *J. Mol. Biol.* **262**, 31–42
823. Geyer, C. R., and Sen, D. (1998) *J. Mol. Biol.* **275**, 483–489
824. Ceska, T. A., and Sayers, J. R. (1998) *Trends Biochem. Sci.* **23**, 331–336
825. Weber, D. J., Libson, A. M., Gittis, A. G., Lebowitz, M. S., and Mildvan, A. S. (1994) *Biochemistry* **33**, 8017–8028
826. Libson, A. M., Gittis, A. G., and Lattman, E. E. (1994) *Biochemistry* **33**, 8007–8016
827. Hale, S. P., Poole, L. B., and Gerlt, J. A. (1993) *Biochemistry* **32**, 7479–7487
828. Grissom, C. B., and Markley, J. L. (1989) *Biochemistry* **28**, 2116–2124
829. Serpersu, E. H., Hibler, D. W., Gerlt, J. A., and Mildvan, A. S. (1989) *Biochemistry* **28**, 1539–1548
830. Suck, D., Lahm, A., and Oefner, C. (1988) *Nature (London)* **332**, 464–468
831. Jones, S. J., Worrall, A. F., and Connolly, B. A. (1996) *J. Mol. Biol.* **264**, 1154–1163
832. Frederick, C. A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.-C., Greene, P., Boyer, H. W., and Rosenberg, J. M. (1984) *Nature (London)* **309**, 327–331

833. McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J., and Rosenberg, J. M. (1986) *Science* **234**, 1526–1541

834. McCarthy, A. D., and Hardie, D. G. (1984) *Trends Biochem. Sci.* **9**, 60–63

835. Kostrewa, D., and Winkler, F. K. (1995) *Biochemistry* **34**, 683–696

836. Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1995) *Science* **269**, 656–663

837. Bozic, D., Grazulis, S., Siksnys, V., and Huber, R. (1996) *J. Mol. Biol.* **255**, 176–186

838. Connolly, B. A., Echstein, F., and Pingoud, A. (1984) *J. Biol. Chem.* **259**, 10760–10763

839. Corey, D. R., Pei, D., and Schultz, P. G. (1989) *Biochemistry* **28**, 8277–8786

840. Ray, W. J., Jr., Hermodsen, M. A., Puvathingal, J. M., and Mahoney, W. C. (1983) *J. Biol. Chem.* **258**, 9166–9174

841. Percival, M. D., and Withers, S. G. (1992) *Biochemistry* **31**, 505–512

842. Dai, J.-B., Liu, Y., Ray, W. J., Jr., and Konno, M. (1992) *J. Biol. Chem.* **267**, 6322–6337

843. Rose, Z. B. (1986) *Trends Biochem. Sci.* **11**, 253–255

844. Shanske, S., Sakoda, S., Hermodson, M. A., DiMauro, S., and Schon, E. A. (1987) *J. Biol. Chem.* **262**, 14612–14617

845. Rigden, D. J., Alexeev, D., Phillips, S. E. V., and Fothergill-Gilmore, L. A. (1998) *J. Mol. Biol.* **276**, 449–459

846. Ravel, P., Craescu, C. T., Arous, N., Rosa, J., and Garel, M. C. (1997) *J. Biol. Chem.* **272**, 14045–14050

847. Collet, J.-F., Stroobant, V., Pirard, M., Delpierre, G., and Van Schaftingen, E. (1998) *J. Biol. Chem.* **273**, 14107–14112

848. Kalckar, H. M. (1985) *Trends Biochem. Sci.* **10**, 291–293

849. Reed, G. H., and Leyh, T. S. (1980) *Biochemistry* **19**, 5472–5480

850. Schulz, G. E., and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York (pp. 222–226)

851. Gerstein, M., Schulz, G., and Chothia, C. (1993) *J. Mol. Biol.* **229**, 494–501

852. Bernstein, B. E., and Hol, W. G. J. (1998) *Biochemistry* **37**, 4429–4436

853. Matte, A., Tari, L. W., and Delbaere, L. T. J. (1998) *Structure* **6**, 413–419

854. Vonrhein, C., Bönisch, H., Schäfer, G., and Schulz, G. E. (1998) *J. Mol. Biol.* **282**, 167–179

855. Bernstein, B. E., Michels, P. A. M., and Hol, W. G. J. (1997) *Nature (London)* **385**, 275–278

856. Yount, R. G., Babcock, D., Ballantyne, W., and Ojala, D. (1971) *Biochemistry* **10**, 2484–2489

857. Yount, R. G. (1975) *Adv. Enzymol.* **43**, 1–56

858. McPhillips, T. M., Hsu, B. T., Sherman, M. A., Mas, M. T., and Rees, D. C. (1996) *Biochemistry* **35**, 4118–4127

859. Schlauderer, G. J., Proba, K., and Schulz, G. E. (1996) *J. Mol. Biol.* **256**, 223–227

860. Byeon, I.-J. L., Shi, Z., and Tsai, M.-D. (1995) *Biochemistry* **34**, 3172–3182

861. Ray, B. D., Chau, M. H., Fife, W. K., Jarori, G. K., and Nageswara Rao, B. D. (1996) *Biochemistry* **35**, 7239–7246

862. Schlauderer, G. J., and Schulz, G. E. (1996) *Protein Sci.* **5**, 434–441

863. Schricker, R., Magdolen, V., Strobel, G., Bogengruber, E., Breitenbach, M., and Bandlow, W. (1995) *J. Biol. Chem.* **270**, 31103–31110

864. Teplyakov, A., Sebastiao, P., Obmolova, G., Perrakis, A., Brush, G. S., Bessman, M. J., and Wilson, K. S. (1996) *EMBO J.* **15**, 3487–3497

865. Zhang, Y., Li, Y., Wu, Y., and Yan, H. (1997) *J. Biol. Chem.* **272**, 29343–29350

866. Müller-Dieckmann, H.-J., and Schulz, G. E. (1995) *J. Mol. Biol.* **246**, 522–530

867. Schlichting, I., and Reinstein, J. (1997) *Biochemistry* **36**, 9290–9296

868. Runquist, J. A., Harrison, D. H. T., and Miziorko, H. M. (1998) *Biochemistry* **37**, 1221–1226

869. Moréra, S., Chiadmi, M., LeBras, G., Lascu, I., and Janin, J. (1995) *Biochemistry* **34**, 11062–11070

870. Webb, P. A., Perisic, O., Mendola, C. E., Backer, J. M., and Williams, R. L. (1995) *J. Mol. Biol.* **251**, 574–587

871. Abdulaev, N. G., Karaschuk, G. N., Ladner, J. E., Kakuev, D. L., Yakhyaev, A. V., Tordova, M., Gaidarov, I. O., Popov, V. I., Fujiwara, J. H., Chinchilla, D., Eisenstrein, E., Gilliland, G. L., and Ridge, K. D. (1998) *Biochemistry* **37**, 13958–13967

872. Mesnildrey, S., Agou, F., Karlsson, A., Bonne, D. D., and Véron, M. (1998) *J. Biol. Chem.* **273**, 4436–4442

873. Turano, A., Furey, W., Pletcher, J., Sax, M., Pike, D., and Kluger, R. (1982) *J. Am. Chem. Soc.* **104**, 3089–3095

874. Aleshin, A. E., Zeng, C., Bourenkov, G. P., Bartunik, H. D., Fromm, H. J., and Honzatko, R. B. (1998) *Structure* **6**, 39–50

875. Aleshin, A. E., Zeng, C., Bartunik, H. D., Fromm, H. J., and Honzatko, R. B. (1998) *J. Mol. Biol.* **282**, 345–357

876. Lundblad, V. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8415–8416

877. Behlke, J., Heidrich, K., Naumann, M., Müller, E.-C., Otto, A., Reuter, R., and Kriegel, T. (1998) *Biochemistry* **37**, 11989–11995

878. Anderson, C. M., Stenkamp, R. E., and Steitz, T. A. (1978) *J. Mol. Biol.* **123**, 15–23

879. Schwab, D. A., and Wilson, J. E. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2563–2567

880. Schirmer, T., and Evans, P. R. (1990) *Nature (London)* **343**, 140–145

881. Auzat, I., Le Bras, G., and Garel, J.-R. (1995) *J. Mol. Biol.* **246**, 248–253

882. Blake, C. C. F., and Evans, P. R. (1974) *J. Mol. Biol.* **484**, 585–601

883. Berger, S. A., and Evans, P. R. (1992) *Biochemistry* **31**, 9237–9242

884. Mizuguchi, H., Cook, P. F., Hasemann, C. A., and Uyeda, K. (1997) *Biochemistry* **36**, 8775–8784

885. Fujii, H., Krietsch, W. K. G., and Yoshida, A. (1980) *J. Biol. Chem.* **255**, 6421–6423

886. Tanaka, K. R., and Paglia, D. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3485–3511, McGraw-Hill, New York

887. Mas, M. T., Bailey, J. M., and Resplandor, Z. E. (1988) *Biochemistry* **27**, 1168–1172

888. Forstner, M., Müller, A., Stolz, M., and Wallimann, T. (1997) *Protein Sci.* **6**, 331–339

889. Fritz-Wolf, K., Schnyder, T., Wallimann, T., and Kabsch, W. (1996) *Nature (London)* **381**, 341–345

890. Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R., and Chapman, M. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8449–8454

891. Larsen, T. M., Benning, M. M., Rayment, I., and Reed, G. H. (1998) *Biochemistry* **37**, 6247–6255

892. Loria, J. P., and Nowak, T. (1998) *Biochemistry* **37**, 6967–6974

893. Lovell, S. C., Mullick, A. H., and Muirhead, H. (1998) *J. Mol. Biol.* **276**, 839–851

894. Cheng, X., Friesen, R. H. E., and Lee, J. C. (1996) *J. Biol. Chem.* **271**, 6313–6321

894a. Valentini, G., Chiarelli, L., Fortin, R., Speranza, M. L., Galizzi, A., and Mattevi, A. (2000) *J. Biol. Chem.* **275**, 18145–18152

895. Hunter, T., and Plowman, G. D. (1997) *Trends Biochem. Sci.* **22**, 18–22

896. Grant, B. D., Hemmer, W., Tsigelny, I., Adams, J. A., and Taylor, S. S. (1998) *Biochemistry* **37**, 7708–7715

897. Cheng, X., Shaltiel, S., and Taylor, S. S. (1998) *Biochemistry* **37**, 14005–14013

898. Engh, R. A., Girod, A., Kinzel, V., Huber, R., and Bossemeyer, D. (1996) *J. Biol. Chem.* **271**, 26157–26164

899. Herberg, F. W., Zimmermann, B., McGlone, M., and Taylor, S. S. (1997) *Protein Sci.* **6**, 569–579

900. Su, Y., Dostmann, W. R. G., Herberg, F. W., Durick, K., Xuong, N.-h, Eyck, L. T., Taylor, S. S., and Varughese, K. I. (1995) *Science* **269**, 807–813

901. Lohmann, S. M., Vaandrager, A. B., Smolenski, A., Walter, U., and DeJonge, H. R. (1997) *Trends Biochem. Sci.* **22**, 307–312

902. Xu, W., Harrison, S. C., and Eck, M. J. (1997) *Nature (London)* **385**, 595–602

903. Sicheri, F., Moarefi, I., and Kuriyan, J. (1997) *Nature (London)* **385**, 602–609

904. Adams, J. A. (1996) *Biochemistry* **35**, 10949–10956

905. Tesmer, J. J. G., Sunahara, R. K., Gilman, A. G., and Sprang, S. R. (1997) *Science* **278**, 1907–1916

906. Tesmer, J. J. G., Sunahara, R. K., Johnson, R. A., Gosselin, G., Gilman, A. G., and Sprang, S. R. (1999) *Science* **285**, 756–760

907. Sunahara, R. K., Tesmer, J. J. G., Gilman, A. G., and Sprang, S. R. (1997) *Science* **278**, 1943–1947

908. Zhao, Y., Schelvis, J. P. M., Babcock, G. T., and Marletta, M. A. (1998) *Biochemistry* **37**, 4502–4509

908a. Zhao, Y., Brandish, P. E., DiValentin, M., Schelvis, J. P. M., Babcock, G. T., and Marletta, M. A. (2000) *Biochemistry* **39**, 10848–10854

909. Joyce, C. M., and Steitz, T. A. (1994) *Ann. Rev. Biochem.* **63**, 777–822

910. Kim, Y., Eom, S. H., Wang, J., Lee, D.-S., Suh, S. W., and Steitz, T. A. (1995) *Nature (London)* **376**, 612–616

911. Doublié, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) *Nature (London)* **391**, 251–258

912. Patel, P. H., Jacobo-Molina, A., Ding, J., Tantillo, C., Clark, A. D., Jr., Raag, R., Nanni, R. G., Hughes, S. H., and Arnold, E. (1995) *Biochemistry* **34**, 5351–5363

913. Bartlett, P. A., and Eckstein, F. (1988) *J. Biol. Chem.* **257**, 8879–8884

913a. Stivers, J. T., Nawrot, B., Jagadeesh, G. J., Stec, W. J., and Shuman, S. (2000) *Biochemistry* **39**, 5561–5572

914. Smith, C. A., and Rayment, I. (1996) *Biophys. J.* **70**, 1590–1602

915. Wittinghofer, A., and Pai, E. F. (1991) *Trends Biochem. Sci.* **16**, 382–387

916. Coleman, D. E., and Sprang, S. R. (1998) *Biochemistry* **37**, 14376–14385

917. Falany, C. N. (1997) *FASEB J.* **11**, 206–216

918. Varin, L., Marsolais, F., Richard, M., and Rouleau, M. (1997) *FASEB J.* **11**, 517–525

919. Kakuta, Y., Petrotchenko, E. V., Pedersen, L. C., and Negishi, M. (1998) *J. Biol. Chem.* **273**, 27325–27330

920. Zhang, H., Varmalova, O., Vargas, F. M., Falany, C. N., and Leyh, T. S. (1998) *J. Biol. Chem.* **273**, 10888–10892

920a. Hiraoka, N., Nakagawa, H., Ong, E., Akama, T. O., Fukuda, M. N., and Fukuda, M. (2000) *J. Biol. Chem.* **275**, 20188 –20196

921. Klaassen, C. D., and Boles, J. W. (1997) *FASEB J.* **11**, 404–418

References

922. Neufeld, E. F., and Muenzer, J. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2465–2494, McGraw-Hill, New York
923. Lukatela, G., Krauss, N., Theis, K., Selmer, T., Gieselmann, V., von Figura, K., and Saenger, W. (1998) *Biochemistry* **37**, 3654–3664
923a. Dierks, T., Lecca, M. R., Schlotterhose, P., Schmidt, B., and von Figura, K. (1999) *EMBO J.* **18**, 2084–2091
924. Waldow, A., Schmidt, B., Dierks, T., von Bülow, R., and von Figura, K. (1999) *J. Biol. Chem.* **274**, 12284–12288
924a. Szameit, C., Miech, C., Balleininger, M., Schmidt, B., von Figura, K., and Dierks, T. (1999) *J. Biol. Chem.* **274**, 15375–15381
925. Fox, D. K., and Roseman, S. (1986) *J. Biol. Chem.* **261**, 13487–13497
926. Joyce, M. A., Fraser, M. E., James, M. N. G., Bridger, W. A., and Wolodko, W. T. (2000) *Biochemistry* **39**, 17–25
927. Johnson, J. D., Muhonen, W. W., and Lambeth, D. O. (1998) *J. Biol. Chem.* **273**, 27573–27579
928. Liaw, S.-H., and Eisenberg, D. (1994) *Biochemistry* **33**, 675–681
929. Alibhai, M., and Villafranca, J. J. (1994) *Biochemistry* **33**, 682–686
930. Meister, A. (1968) *Adv. Enzymol.* **31**, 183 – 218
931. Meister, A. (1974) in *The Enzymes*, 3rd ed., Vol. 10 (Boyer, P. D., ed), pp. 669–754, Academic Press, New York
932. Whitty, A., Fierke, C. A., and Jencks, W. P. (1995) *Biochemistry* **34**, 11678–11689
933. Rochet, J.-C., and Bridger, W. A. (1994) *Protein Sci.* **3**, 975–981
934. Selmer, T., and Buckel, W. (1999) *J. Biol. Chem.* **274**, 20772–20778

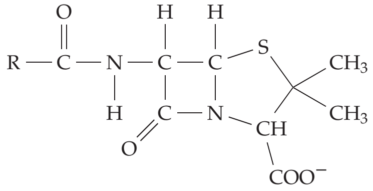
Study Questions

1. Outline the reactions by which glyceraldehyde 3-phosphate is converted to 3-phosphoglycerate with coupled synthesis of ATP in the glycolysis pathway. Show important mechanistic details.

2. Papain is a protein-hydrolyzing (proteolytic) enzyme with an –SH group and an imidazole group at the active site. Write a reasonable structure for a "tetrahedral intermediate" that would be expected to arise during formation of an acyl enzyme intermediate.

3. Adenylate kinase catalyzes the interconversion of ATP, AMP, and ADP.
   a) Draw a reasonable structure for a pentacovalent intermediate derived from ATP and AMP.
   b) Draw a reasonable structure for the transition state leading from ATP + AMP to two molecules of ADP in an S_N2-like reaction.

4. Penicillin inhibits a D-alanyl-D-alanine transpeptidase that catalyzes the reaction



Penicillin

R′-D-alanyl-D-alanine + H_2N-R →
R′-D-alanyl-NH-R + D-alanine

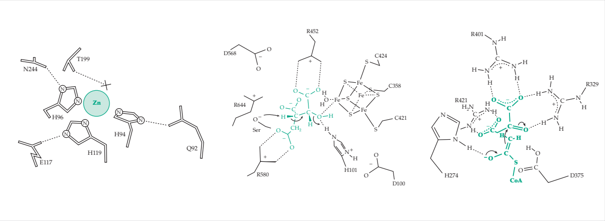
where R and R′ are different parts of a bacterial peptidoglycan. Write a step-by-step mechanism for this reaction and indicate how penicillin may inhibit the enzyme by combining with it irreversibly. See Strynodka *et al., Nature* **359**, 700–705, 1992, for related reaction of penicillin with a penicillinase.

5. Write out step-by-step chemical mechanisms for the following enzymatic reaction. Use small arrows to indicate directions of electron flow. Remember to have all electrons move in the same direction in any single structure.

Amino acid + tRNA + ATP →
Aminoacyl-tRNA + AMP + PP_i

6. Trypsin in which Asp 102 has been replaced by Asn has 10^4 times less catalytic activity than natural trypsin at neutral pH. From the crystal structure of the mutant enzyme it appears that the imidazole group of His 57 is held by the Asn side chain in the wrong tautomeric form for catalysis. Explain. Compare this incorrect tautomeric form with that in the initial structure shown in Fig. 12-11.

7. A recent discovery in biochemistry is that RNA can act as an enzyme in chemical reactions, usually reactions involving RNA hydrolysis. Discuss the features of RNA structure that might favor evolution of enzymes composed entirely of a single polyribonucleotide chain, and describe a proposed mechanism for RNA-catalyzed hydrolysis of RNA molecules.



One of the simplest biochemical addition reactions is the hydration of carbon dioxide to form carbonic acid, which is released from the zinc-containing **carbonic anhydrase** (left, Fig. 13-1) as HCO_3^-. **Aconitase** (center, Fig. 13-4) is shown here removing a water molecule from isocitrate, an intermediate compound in the citric acid cycle. The H_2O that is removed will become bonded to an iron atom of the Fe_4S_4 cluster at the active site as indicated by the black H_2O. An enolate anion derived from acetyl-CoA adds to the carbonyl group of oxaloacetate to form citrate in the active site of **citrate synthase** (right, Fig. 13-9) to initiate the citric acid cycle.

Contents

677 .... **A. Addition of R–OH, R–NH$_2$, and R–SH to Polarized Double Bonds**
677 .... 1. Carbonic Anhydrase
679 .... 2. Imines (Schiff Bases)
680 .... 3. Stereochemistry of Addition to Trigonal Carbon Atoms
681 .... 4. Addition of Carbon–Carbon Double Bonds, Often Reversible Reactions
681 .... *Enoyl-CoA hydratase*
682 .... *Glutathione S-transferases*
682 .... *Chlorobenzoyl-CoA dehalogenase*
683 .... 5. Addition to Double Bonds Adjacent to Carboxylate Groups
683 .... *Fumarate hydratase*
685 .... *Some other fumarate-forming reactions*
685 .... *Enolase*
686 .... *Pectate lyase and related enzymes*
686 .... 6. Aconitase and Related Iron–Sulfur Proteins
688 .... 7. Addition to or Formation of Isolated Double Bonds
689 .... 8. Conjugative and Decarboxylative Elimination Reactions
690 .... 9. Isomerization Assisted by Addition
690 ....10. Reversibility of Addition and Elimination Reactions
691 .... **B. Enolic Intermediates in Enzymatic Reactions**
691 .... 1. Mandelate Racemase and Related Enzymes
692 .... 2. Isomerases
692 .... *Aldose–ketose interconversions*
693 .... *Catalysis of ring opening by isomerases*
693 .... *Triose phosphate isomerase*
695 .... *Xylose isomerase and the hydride shift mechanism*
696 .... *3-Oxosteroid isomerases*
697 .... *4-Oxalocrotonate tautomerase*
697 .... 3. Internal Oxidation–Reduction by Dehydration of Dihydroxyacids
697 .... 4. Formation and Metabolism of Methylglyoxal (Pyruvaldehyde)
698 .... **C. Beta Cleavage and Condensation**
698 .... 1. Displacement on a Carbonyl Group
699 .... 2. Addition of an Enolate Anion to a Carbonyl Group or an Imine; Aldolases
700 .... *Polycarboxylic acid synthases*
703 .... *Citrate cleaving enzymes*
704 .... 3. Chiral Acetates and Their Use in Stereochemical Studies
705 .... 4. Addition of an Enolate Ion to Carbon Dioxide and Decarboxylation
705 .... *Decarboxylation of β-oxoacids*
705 .... *Linked oxidation and decarboxylation*
705 .... *Phosphoenolpyruvate, a key metabolic intermediate*
706 .... *Ribulose bisphosphate carboxylase*
710 .... *Carbon dioxide or bicarbonate ion?*
711 .... 5. Incorporation of Bicarbonate into Carboxyl Groups
711 .... *PEP mutase and the synthesis of phosphonates*
712 .... **D. Some Isomerization and Rearrangement Reactions**

713 .... **References**
717 .... **Study Questions**

Boxes

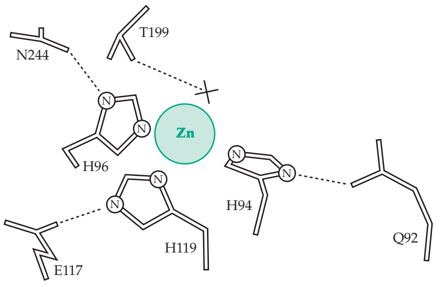
680 .... Box 13-A Zinc
687 .... Box 13-B EPSP Synthase and the Herbicide Glyphosate

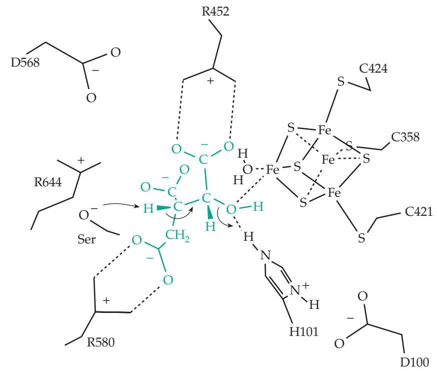
Tables

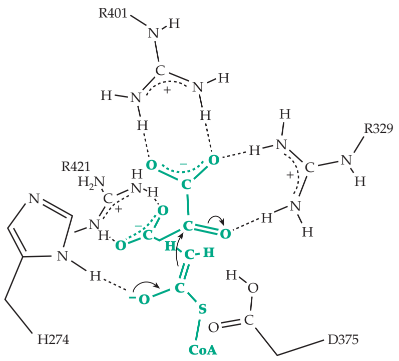
701 .... Table 13-1 Products Arising from Reactions of Acetyl-CoA with a Second Substrate with Catalysis by a Polycarboxylate Synthase

Enzymatic Addition, Elimination, Condensation, and Isomerization: Roles for Enolate and Carbocation Intermediates

13





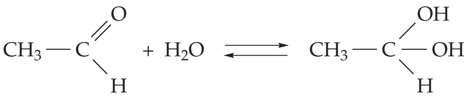


In Chapter 12 we considered reactions by which living cells are able to transfer groups from one molecule to another using nucleophilic displacements. We also showed how transfer reactions can be utilized by ligases to join two molecules together with the Gibbs energy of cleavage of ATP or of a related molecule driving the reaction. In this chapter we will examine addition reactions, which provide a simple way of joining two molecules by means of C–O, C–N, C–S, or C–C bonds. Among these are the aldol and Claisen-type condensations by which C–C bonds are formed. We will also consider elimination reactions and decarboxylations, which are the reverse of addition and condensation reactions, as well as mechanistically related isomerizations. Many reactions of these types occur in the major pathways of metabolism.

A. Addition of R–OH, R–NH_2, and R–SH to Polarized Double Bonds

Next to nucleophilic displacement, the commonest mechanistic processes in enzymatic catalysis are addition to double bonds and elimination to form double bonds. These often involve addition of a nucleophile together with a proton to a highly polarized double bond such as C=O or C=N–. In other reactions, which are discussed in Section C,2, the nucleophile attacks one end of a C=C bond that is polarized by conjugation with C=O or C=N.

Alcohols, amines, and thiols add readily to the electrophilic carbon of the carbonyl group to form **hemiacetals**, **carbinolamines, hemiketals,** and **hemimercaptals**. An example is the formation of ring structures of sugars (Eq. 4-1). Water can also add to carbonyl groups and most aliphatic carbonyl compounds exist in water as an equilibrium mixture with a covalent hydrate (Eq. 13-1). For example, acetaldehyde in aqueous solution consists of a mixture of about 50% free aldehyde and 50% hydrate in rapidly reversible equilibrium[1,2] and formaldehyde is over 99.9% hydrated.[3]

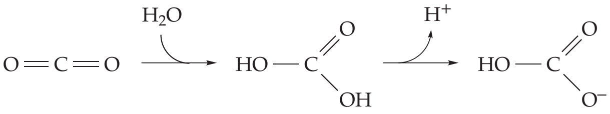


$$CH_3-\overset{\displaystyle O}{\overset{\|}{C}}-H + H_2O \rightleftharpoons CH_3-\underset{\displaystyle H}{\overset{\displaystyle OH}{C}}-OH \qquad (13\text{-}1)$$

Addition reactions often occur as parts of more complex reactions. For example, a thiol group of glyceraldehyde-3-phosphate dehydrogenase reacts with the aldehyde substrate to form a hemimercaptal, which is subsequently oxidized to a thioester (see Fig. 15-6).

1. Carbonic Anhydrase

Another simple addition reaction is the hydration of CO_2 to form the bicarbonate ion. Without catalysis the reaction may require several seconds,[4,5] the apparent first-order rate constant being ~0.03 s^{-1} at 25°C. Cells must often hasten the process. The specific catalyst carbonic anhydrase is widespread in its distribution



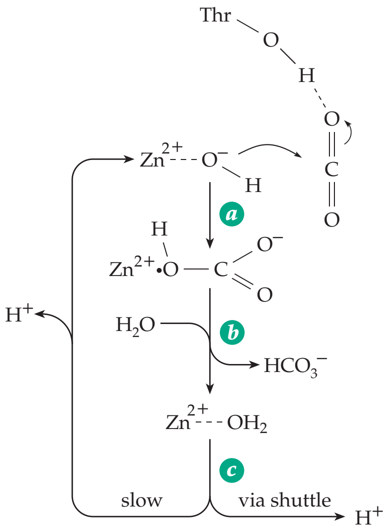
$$O=C=O \xrightarrow{H_2O} HO-C(=O)-OH \xrightarrow{-H^+} HO-C(=O)-O^- \qquad (13\text{-}2)$$

and is especially active in tissues (e.g., red blood cells, and lungs) that are involved in respiration. One liter of mammalian blood contains 1–2 g of this enzyme, a monomeric 30-kDa protein containing ~260 amino acids and one tightly bound ion of Zn^{2+}. Erythrocytes contain two isoenzymes (I and II) of carbonic anhydrase and the human body contains at least eight distinct isoenzymes (I – VIII).[5–8] They are found wherever there is a high demand for CO_2 or bicarbonate. Isoenzyme I, II, III, and VII are cytosolic. Carbonic anhydrase I is specific to erythrocytes, while isoenzyme II is present in most cells. Hereditary lack of carbonic anhydrase II is associated with **osteopetrosis** (marble bone disease), a condition involving failure of bone resorption and the calcification of other tissues.[9] The generation of acidity according to Eq. 13-2 is presumably required for dissolution of bone by the osteoclasts.

Isoenzymes III and VII have a more specialized distribution. Carbonic anhydrase III is abundant in adipocytes which use bicarbonate in fatty acid synthesis.[7] Isoenzyme V is present in the mitochondrial matrix and is also abundant in both adipocytes and liver.[7,8] Isoenzyme IV is a larger membrane-associated form, while VI is secreted into the saliva.[10] Carbonic anhydrase has also been identified in *E. coli*.,[11] in a methanobacterium,[12] and in green plants.[13,13a] A 60-kDa carbonic anhydrase called **nacrein** is found in the organic matrix of the nacreous layer of the pearl oyster, the layer that forms aragonite (orthorhombic calcium carbonate) in the shell and in pearls.[14]

X-ray studies of carbonic anhydrases I and II, from human blood, revealed that both have an ellipsoidal shape of dimensions ~4.1 x 4.1 x 4.7 nm.[15,16] The zinc atom in each molecule lies in a deep pocket ~1.2 nm from the surface and is surrounded by three histidine side chains and one H_2O or OH^- ion, the four ligands forming a distorted tetrahedron (Fig. 13-1). The coordinating imidazole group from His 119 is hydrogen bonded to a carboxylate group of Glu 117, a feature reminiscent of the charge-relay system of serine protease. This carboxylate group is also bound into a more extended hydrogen-bonded network, part of which is indicated in Fig. 13-1. The other imidazole groups also form hydrogen bonds to protein groups and the zinc-bound H_2O is involved in an extensive hydrogen-bonded network with several other bound water molecules and protein side chains.[17a] Most of these structural features are conserved in the other mammalian isoenzymes.[5] However, from X-ray absorption spectroscopy (EXAFS) it appears that spinach carbonic anhydrase contains one or more sulfur ligands to the zinc,[13] while the enzymes from the archaeon *Methanosarcina thermophila* have a left-handed β helix structure (see Figs. 2-17 and 13-3).[12] Nevertheless, this enzyme has the same three-histidine Zn–OH structure found in the mammalian enzymes.

Carbonic anhydrase II is among the most rapid enzymes known, with the turnover number at 25°C for hydration of CO_2 being ~10^6 s^{-1}. The same enzyme catalyzes hydration of acetaldehyde (Eq. 13-1) but at a 1000-fold slower rate. A pK_a of ~7 controls the activity. This appears to represent the loss of H^+ from the Zn^{2+}–OH_2 complex[18] to give Zn^+–OH. The latter is in effect a stabilized hydroxide ion existing at a pH at which OH^- is normally not present in quantity. It is this hydroxide ion that adds to the CO_2 or to the aldehyde substrate (Eq. 13-3, step *a*).[18–20] In step *b* a water molecule replaces the departing bicarbonate.[17] A variety of data indicate that proton transfers mediated by the enzyme are essential parts of the carbonic anhydrase mechanism.[18–19b,21–23] One proposal is that the nearby imidazole group of His 64 (not shown in Fig. 13-1) deprotonates the bound H_2O via a hydrogen-bonded network of bound water molecules (Eq. 13-3, step *c*). The side chain of Thr 199 may function in a cyclic proton transfer in step *c*.[19] The proton generated in step *c* is released to the solvent in a process that is catalyzed by buffer anions or by amines such as histamine. The latter binds at the edge of the active site and forms an additional hydrogen-bonded pathway to the zinc-bound H_2O.[24] A different proton shuttle pathway has been proposed for the slower carbonic anhydrase III.[25]



(13-3)

Related to the reaction catalyzed by carbonic anhydrase is the addition of an amino group to CO_2 (Eq. 2-21) to form a carbamino group (–NH–COO^-). This reaction is essential to the functioning of hemoglobin, which must carry large amounts of CO_2, as carbamino groups, from tissues to the lungs, (Eq. 7-47) and to some enzymes such as ribulose bisphosphate carboxylase (Figs. 13-10, 13-11).

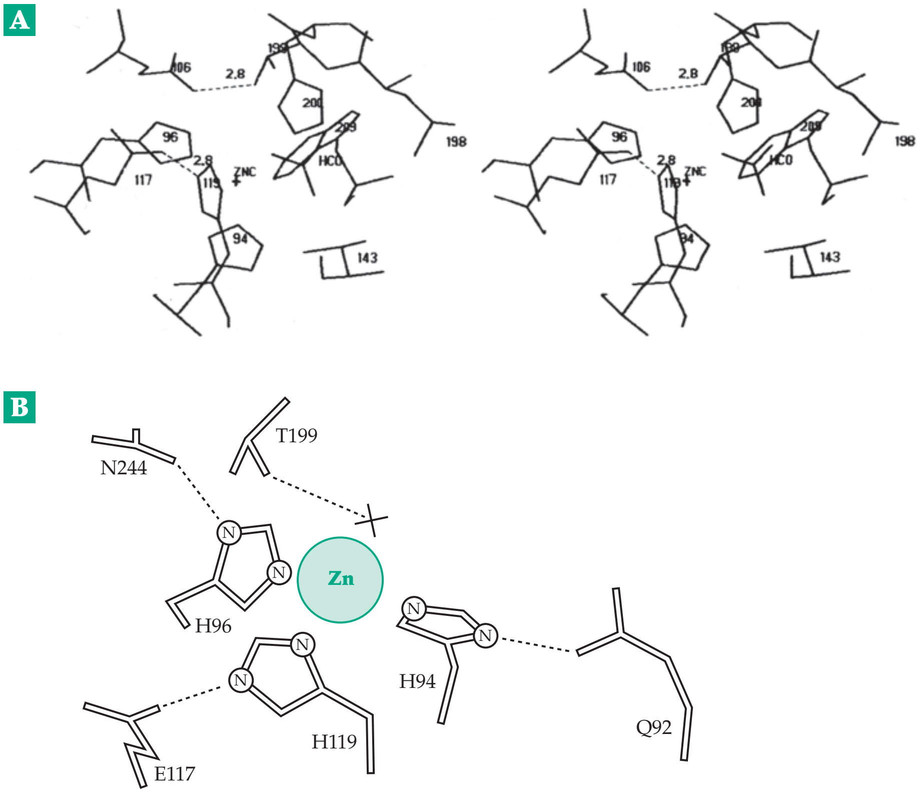
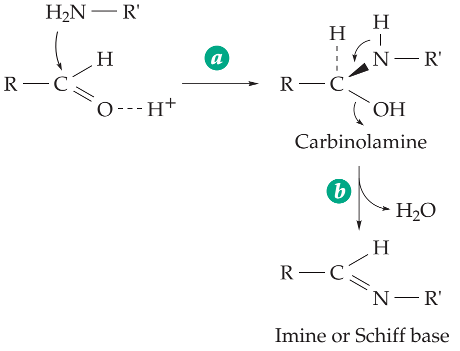


Figure 13-1 (A) Stereoscopic view showing the binding pocket for HCO_3^- in the active site of human carbonic acid anhydrase I. Also shown are the hydrogen-bonded pairs E117–H119 and E106–T199. ZNC, zinc ion; HCO, bicarbonate. From Kumar and Kannan.[16] (B) View of the active site of carbonic anhydrase II. The orientation is a little different than in (A). The location of H_2O or ^-OH bound to the zinc ion is marked X. Hydrogen bonds from the three histidines that form coordinate bonds to the zinc, Q92, E117, the backbone carbonyl of N244, and the hydrogen bond from the zinc-bound hydroxyl to the T199 side chain are shown as dashed lines. From Kiefer *et al.*[17]

2. Imines (Schiff Bases)

As we have seen already, many enzymatic reactions depend upon formation of **imines**, which are commonly called Schiff bases. The two-step formation of Schiff bases consists of addition of an amino group to a carbonyl group to form a carbinolamine followed by elimination of water (Eq. 13-4).[26] One group of **aldolases** (Section D) have, at their active centers, lysine side chains which form Schiff bases with the ketone substrates prior to the principal reaction of breaking or forming a C – C bond. Similarly, the initial reaction of the aldehyde coenzyme **pyridoxal phosphate** with amino acid substrates is the formation of Schiff bases (Fig. 14-4). Indeed, *the groups C=O and H$_2$N– are inherently complementary* and their interaction through imine formation is extremely common.



(13-4)

Schiff bases often form within a fraction of a second, but one or both steps of Eq. 13-4 may require catalysis to achieve enzymatic velocities.[27] The reaction is usually completely reversible and formation constants are often low enough that a carbonyl compound present in small amounts will not react extensively with an amine unless the two are brought together on an enzyme surface. If the amino group is hydrogen bonded to the enzyme, the proton may remain on the Schiff base nitrogen, enhancing the electron-accepting properties of the C=N group. The pK_a values of Schiff bases with aliphatic aldehydes are usually 1– 2 units lower than those of the corresponding primary amines. However, the local environment may sometimes cause the pK_a to be increased. Most Schiff bases are reduced readily by sodium borohydride or sodium cyanoborohydride to form secondary amines in which the original aldehydes are bound covalently to the original amino groups (Eq. 3-34). This provides a method for locating sites of Schiff base formation in enzymes. An isotopically substituted aldehyde or amine can be employed, or an isotopic label can be introduced, by using ^{2}H- or ^{3}H-containing sodium borohydride in the reduction.

3. Stereochemistry of Addition to Trigonal Carbon Atoms

Adducts formed by enzymatic addition of nucleophiles to carbonyl groups are usually chiral. For example, addition of an amino group to the "front" side of the carbonyl carbon in Eq. 13-5 (the *si* face as defined in Chapter 9) creates a carbinolamine of the *S* configuration. Addition of the amino group from behind the plane of the paper (the *re* face) would lead to a carbinolamine of the *R* configuration.

BOX 13-A ZINC

The average human ingests 10–15 mg of zinc a day.[a] Although it is poorly absorbed, the concentrations of zinc in tissues are relatively high and the metal plays an essential role in a multitude of enzymes. The total zinc content of a 70-kg person is 1.4–2.3 g. A typical tissue concentration of Zn^{2+} is 0.3–0.5 mM; an unusually high content of ~15 mM is found in the prostate gland.

Zinc ion is much more tightly bound to most organic ligands than is Mg^{2+} (Table 6-9). It has a filled 3*d* shell and tends to form four ionic bonds with a tetrahedral geometry, often with nitrogen- or sulfur-containing ligands.[b] Unlike Mg^{2+}, which interacts rapidly and reversibly with enzymes, Zn^{2+} tends to be tightly bound within over 300 **metalloenzymes**.[c–e] A common feature is the surrounding of the Zn^{2+} at the active center by three imidazole groups, the fourth coordination position being free for interaction with substrate. The second nitrogen of the imidazole ring in many instances is hydrogen bonded to a main chain carbonyl group of the peptide, a feature that is also shared by histidines in other metalloproteins.

The most important chemical function of Zn^{2+} in enzymes is probably that of a Lewis acid providing a concentrated center of positive charge at a nucleophilic site on the substrate.[f] This role for Zn^{2+} is discussed for carboxypeptidases (Fig.12-16) and thermolysin,[g] alkaline phosphatase (Fig. 12-23),[h] RNA polymerases, DNA polymerases, carbonic anhydrase (Fig. 13-1),[i] class II aldolases (Fig. 13-7), some alcohol dehydrogenases (Fig. 15-5), and superoxide dismutases (Fig.16-22). Zinc ions in enzymes can often be replaced by Mn^{2+}, Co^{2+}, and other ions with substantial retention of catalytic activity.[f,j]

In addition to its function in catalysis, zinc often plays an important structural role, e.g., in the **zinc finger** transcriptional regulators (Fig. 5-38).[k] Zinc ions bind to insulin and stabilize its hexameric structure (Fig. 7-18).[l] Six Zn^{2+} ions are present in the hexagonal tail plate of the T-even bacteriophage (Box 7-C) and appear to be essential for invasion of bacteria.[m] In carnivores, the **tapetum**, the reflecting layer behind the retina of the eye of many animals, contains crystals of the Zn^{2+}–cysteine complex.

Since zinc ions have no color their presence has often been overlooked. Zinc ions will doubtless be found in many more places within cells. Zinc is usually the major component of the bound metals in the **metallothioneins** (Box 6-E). These small 6.6-kDa proteins which contain ~33% cysteine and bind as many as six ions of Cd^{2+}, Hg^{2+}, Cu^{2+}, or Zn^{2+} per molecule are present in all animal tissues as well as in plants and some bacteria.[j]

From a nutritional viewpoint, Cu^{2+} competes with zinc ion, as does the very toxic Cd^{2+}. The latter accumulates in the cortex of the kidney. Dietary cadmium in concentrations less than those found in human kidneys shortens the lives of rats and mice. However, some marine diatoms contain a cadmium-dependent carbonic anhydrase.[n] Although zinc deficiency was once regarded as unlikely in humans, it is now recognized as occurring under a variety of circumstances[o,p] and is well-known in domestic animals.[q] Consumption of excessive amounts of protein as well as alcoholism, malabsorption, sickle cell anemia, and chronic kidney disease can all be accompanied by zinc deficiency.

---

[a] O'Dell, B. L., and Campbell, B. J. (1971) *Comprehensive Biochemistry* **21**, 179–216

[b] Bock, C. W., Katz, A. K., and Glusker, J. P. (1995) *J. Am. Chem. Soc.* **117**, 3754–3765

[c] Berg, J. M., and Shi, Y. (1996) *Science* **271**, 1081–1085

[d] Vallee, B. L., and Auld, D. S. (1993) *Biochemistry* **32**, 6493–6500

[e] Coleman, J. E. (1992) *Ann. Rev. Biochem.* **61**, 897–946

[f] Mildvan, A. S. (1974) *Ann. Rev. Biochem.* **43**, 357–399

[g] Holland, D. R., Hausrath, A. C., Juers, D., and Matthews, B. W. (1995) *Protein Sci.* **4**, 1955–1965

[h] Kimura, E., Kodama, Y., Koike, T., and Shiro, M. (1995) *J. Am. Chem. Soc.* **117**, 8304–8311

[i] Lesburg, C. A., and Christianson, D. W. (1995) *J. Am. Chem. Soc.* **117**, 6838–6844

[j] Kagi, J. H. R., Himmelhoch, S. R., Whanger, P. D., Bethune, J. L., and Vallee, B. L. (1974) *J. Biol. Chem.* **249**, 3537–3542

[k] Berg, J. M. (1990) *J. Biol. Chem.* **265**, 6513–6516

[l] Hill, C. P., Dauter, Z., Dodson, E. J., Dodson, G. G., and Dunn, M. F. (1991) *Biochemistry* **30**, 917–924

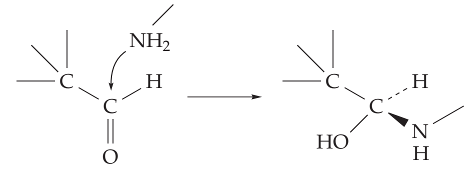
[m] Kozloff, L. M., and Lute, M. (1977) *J. Biol. Chem.* **252**, 7715–7724

[n] Lane, T. W., and Morel, F. M. M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4627–4631

[o] Prasad, A. S. (1984) *Fed. Proc.* **43**, 2829–2834

[p] Day, H. G. (1991) *FASEB J.* **5**, 2315–2316

[q] Luecke, R. W. (1984) *Fed. Proc.* **43**, 2823–2828



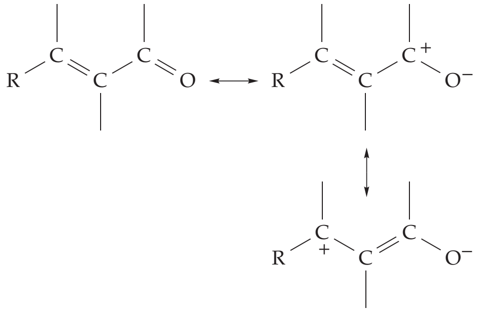
Addition from "front"
(*re* face of the carbonyl carbon)

S configuration

(13-5)

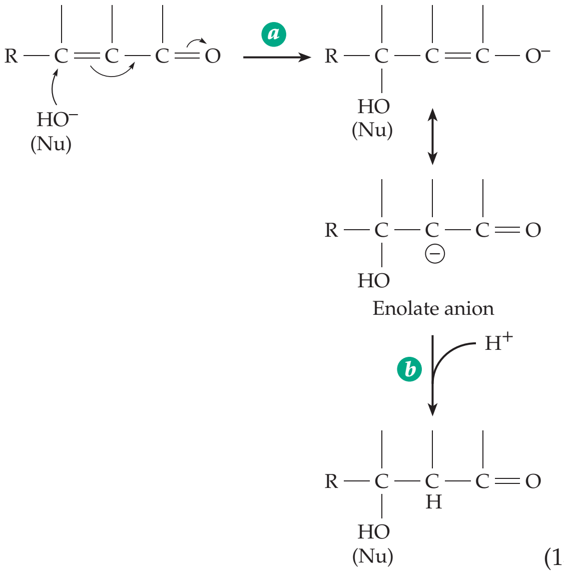
4. Addition to Carbon–Carbon Double Bonds, Often Reversible Reactions

Most of the reactions that we will consider in this chapter involve addition of a proton to a carbon atom or removal of a proton attached to a carbon atom. A frequent metabolic reaction is addition of water to a carbon–carbon double bond that is conjugated with a carbonyl group. This transmits the polarization of the carbonyl group to a position located two carbon atoms further along the chain.



Because of this effect bases can add to a carbon–carbon double bond in a position β to the carbonyl group as in Eq. 13-6, step *a*. The product of addition of HO⁻ is the anion of an **enol** in which the negative charge is distributed by resonance between the oxygen of the carbonyl group and the carbon adjacent to the carbonyl. A stable product results if a proton adds to the latter position (Eq. 13-6, step *b*). Many nucleophilic groups (Nu in Eq. 13-6) other than HO⁻, may add, the reaction being known as a **Michael addition**. If a neutral nucleophile Nu adds, the product may lose a proton, or transfer it to the adjacent carbon in Eq. 13-6, step *b*, to give a neutral end product. The reverse of an addition reaction of this type is also known as an **elimination reaction**.

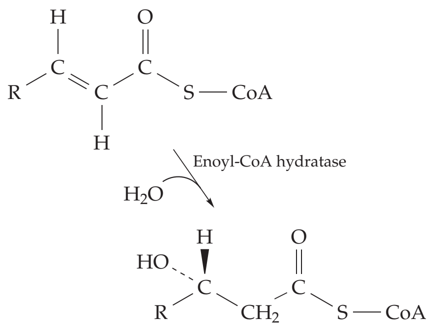
Enzymatic reactions involving addition to a C=C bond adjacent to a carbonyl group (or in which elimination occurs α,β to a carbonyl) are numerous. Except for some enzymes acting by free radical mechanisms, the nucleophilic group always adds at the β position suggesting that the mechanism portrayed by Eq. 13-7 is probable. It is noteworthy that *frequently in a metabolic sequence a carbonyl group is deliberately introduced to*



(13-6)

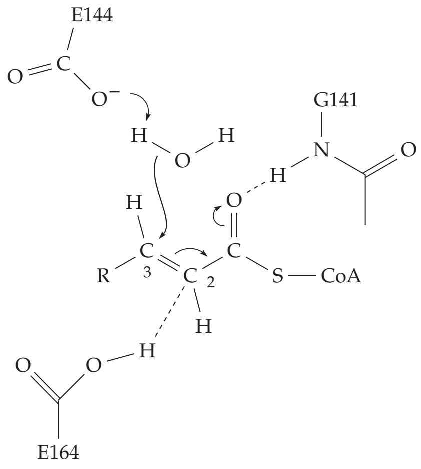
facilitate elimination or addition at adjacent carbon atoms. The carbonyl may be formed by oxidation of a hydroxyl group or it may be provided by a thioester formed with coenzyme A or with an **acyl carrier protein** (Chapter 14).

Enoyl-CoA hydratase. A specific example of the reaction in Eq. 13-6 is the addition of water to *trans*-α,β-unsaturated CoA derivatives (Eq. 13-7). It is catalyzed by enoyl–CoA hydratase (crotonase) from mitochondria and is a step in the β oxidation of fatty acids (Fig. 10-4).



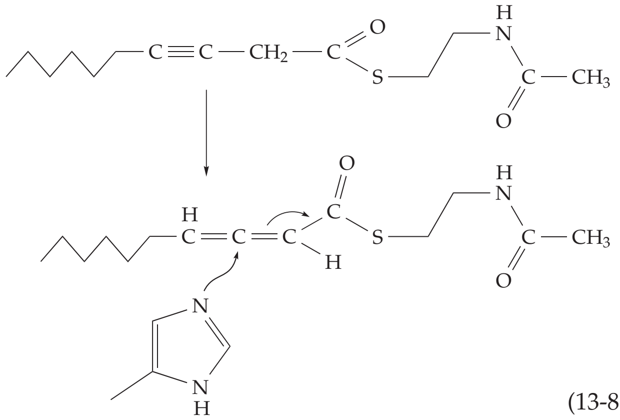
L-β-Hydroxyacyl-CoA (13-7)

The enzyme is a hexamer, actually a dimer of trimers made up of 291-residue polypeptide chains.[28] Acetoacetyl-CoA is a competitive inhibitor which binds into the active site and locates it. From the X-ray structure of the enzyme–inhibitor complex it can be deduced that the carboxylate group of E144 abstracts a proton from a water molecule to provide the hydroxyl ion that binds to the β position (Eq. 13-6, step *a*) and that the E164 carboxyl group donates a proton to the intermediate enolate anion in step *b*.[28] The hydroxyl group and the proton enter from the same side of the double bond in a *syn* addition. The X-ray structure shows that the E144 and E164 side chains are on the same side of the double bond, accounting for this stereochemistry.[28,28a]

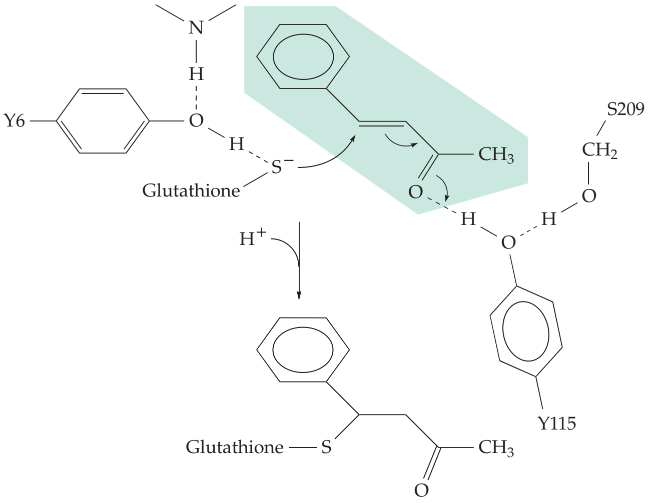


Cinnamoyl-CoA thiol esters (R = phenyl in the foregoing structure) contain a good light-absorbing group (chromophore). Binding to the protein induces distinct shifts in the ultraviolet absorption spectrum and also in ^{13}C NMR and Raman spectra. These shifts suggest that binding induces an enhanced positive charge at C-3 and formation of a strong hydrogen bond to the carbonyl group.[29] The X-ray studies indicate that this bond is to the G141 peptide nitrogen as shown. These results favor an enolate anion intermediate as in Eq. 13-6. However, kinetic isotope effects[30] as well as studies of proton exchange have suggested a concerted mechanism with water adding at the same time that a proton binds at C-2.[30–31a]

A closely related *E. coli* protein is a 79-kDa multifunctional enzyme that catalyzes four different reactions of fatty acid oxidation (Chapter 17). The amino-terminal region contains the enoyl hydratase activity.[32] A quite different enzyme catalyzes dehydration of thioesters of β-hydroxyacids such as 3-hydroxydecanoyl-acyl carrier protein (see Eq. 21-2) to both form and isomerize enoyl-ACP derivatives during synthesis of unsaturated fatty acids by *E. coli*. Again, a glutamate side chain is the catalytic base but an imidazole group of histidine has also been implicated.[33] This enzyme is inhibited irreversibly by the *N*-acetylcysteamine thioester of 3-decynoic acids (Eq. 13-8). This was one of the first enzyme-activated inhibitors to be studied.[34]



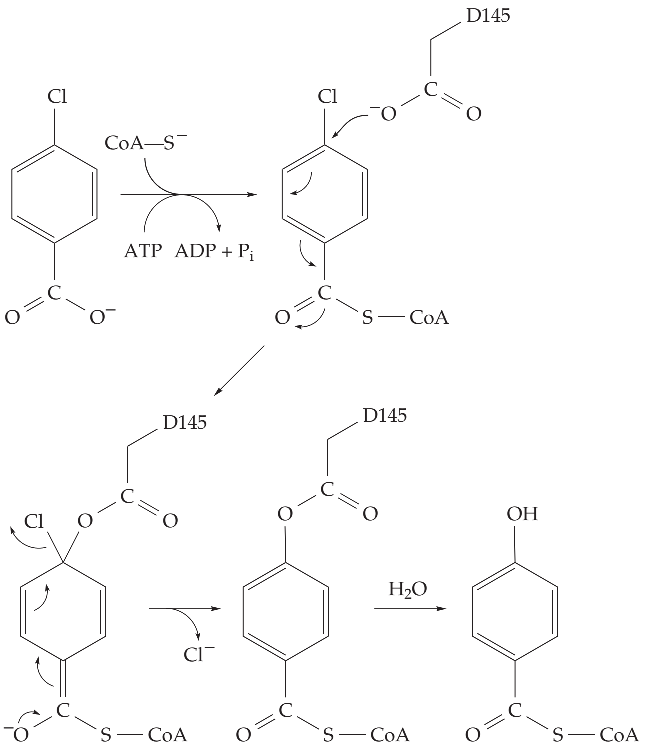
(13-8)



(13-9)

Glutathione S-transferases. Addition of glutathione (Box 11-B) to a large variety of different substrates containing electrophilic centers, such as that at the β position in an α,β-unsaturated ketone (Eq. 13-9), is catalyzed by the ubiquitous group of enzymes called glutathione *S*-transferases. There are six classes of eukayotic glutathione *S*-transferases–one membrane-associated microsomal[35] and five cytosolic.[36–40] They play an important role in detoxifying many xenobiotics and other relatively hydrophobic compounds by converting them to more soluble compounds that can be degraded and excreted easily (see Box 11-B). A pair of tyrosines appear to participate in catalysis, as is indicated in Eq. 13-9.[38,41] Glutathione *S*-transferases have attracted attention because of their role in detoxifying anticancer drugs. Cancer cells can become resistant to drugs as a result of excessive synthesis of these detoxifying enzymes. At the same time glutathione transferases protect human patients from drugs. In plants these enzymes may provide protection from insecticides and herbicides.[41a] Glutathione transferases of pathogenic organisms such as schistosomes are appropriate targets for new drugs as well as for vaccines.[42] Nonenzymatic Michael addition of glutathione to such compounds as **4-hydroxynonenal**, a product of peroxidation of the polyunsaturated arachidonate (see Eq. 21-15) are also biochemically important.

Chlorobenzoyl-CoA dehalogenase. The enzymatic release of chloride ion from 4-chloroxybenzoyl-CoA would be hard to explain by a simple nucleophilic displacement by HO$^-$. However, *addition* of HO$^-$, followed by elimination of Cl$^-$, would be chemically reasonable. Nevertheless, single-turnover studies revealed a more complex mechanism involving formation of a covalent adduct with the enzyme.[43,44] A carboxylate group adds and Cl$^-$ is eliminated to form of an oxygen ester, which is then hydrolyzed to the final product (Eq. 13-10). Studies of mutant enzymes[45] together with X-ray crystallography[46] support this mechanism. As with enoyl-CoA hydratase, binding of substrate to chlorobenzoyl-CoA dehalogenase causes alterations in ultraviolet, NMR, and Raman spectra that can be interpreted as indicating enhanced polarization of the benzoyl group.[47,48] For example, changes in the C = O stretching frequency suggest that the bond is elongated, presumably as a result of hydrogen bonding to the N – H of G114, which lies at the N terminus of an α helix and experiences the additional polarizing effect of the helix dipole (Fig. 2-20).[48,49] This effect can be compared with that mentioned previously for enoyl-CoA hydratase.

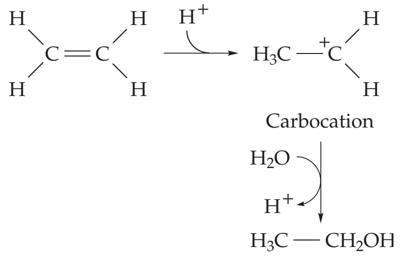


(13-10)

In fact, close sequence and structural homologies show that chlorobenzoyl-CoA dehalogenase, enoyl-CoA hydratase, and a variety of other hydratases, isomerases, synthases, lyases, and hydrolases belong to a large family of related proteins.[49a,49b,49c]

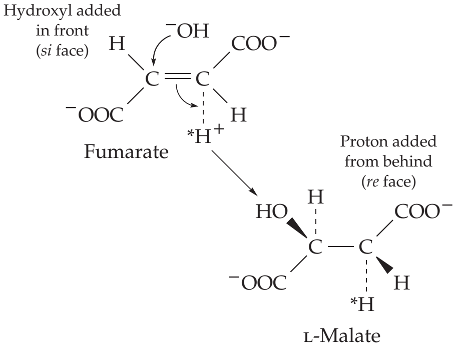
5. Addition to Double Bonds Adjacent to Carboxylate Groups

Biochemical reactions often involve addition to C = C bonds that are not conjugated with a true carbonyl group but with the poorer electron acceptor – COO$^-$. While held on an enzyme a carboxylate group may be protonated, making it a better electron acceptor. Nevertheless, there has been some doubt as to whether the carbanion mechanism of Eq. 13-6 holds for these enzymes. Some experimental data suggested a quite different mechanism, one that has been established for the nonenzymatic hydration of alkenes. An example is the hydration of ethylene by hot water with dilute sulfuric acid as a catalyst (Eq. 13-11), an industrial method of preparation of ethanol. The electrons of the double bond form the point of attack by a proton, and the resulting carbocation readily abstracts a hydroxyl ion from water. Direct addition of OH$^-$ to ethylene to form a carbanion is not favored, because there is no adjacent carbonyl group to stabilize the negative charge.



(13-11)

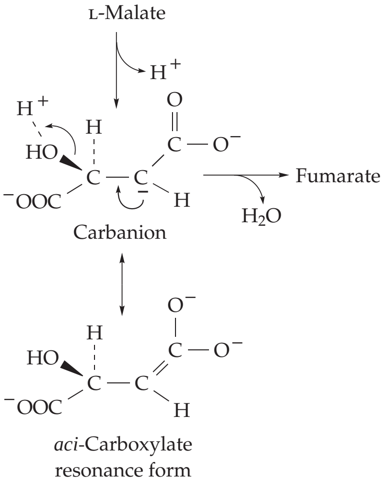
Fumarate hydratase. The most studied enzyme of this group is probably the porcine mitochondrial fumarate hydratase (fumarase; see also Chapter 9), a tetramer of 48.5-kDa subunits[50] with a turnover number of ~2 x 10^3 s^{-1}. It accelerates the hydration reaction more than 10^{15}-fold.[51] A similar enzyme, the 467-residue **fumarase C** whose three-dimensional structure is known,[52,53] is found in cells of *E. coli* when grown aerobically. The product of the fumarate hydratase reaction is L-malate (*S*-malate). The stereospecificity is extremely high. If the reaction is carried out in ^2H_2O an atom of ^2H is incorporated into the *pro-R* position, i.e., the proton is added strictly from the *re* face of the trigonal carbon (Eq. 13-12). To obtain L-malate the hydroxyl must have been added from the opposite side of the double bond. Such *anti* (*trans*) addition is much more common in both nonenzymatic and enzymatic reactions than is addition of both H and OH (or – Y) from the same side (*syn*, *cis*, or adjacent addition).[54] For concerted addition it is a natural result of stereoelectronic control. Almost *all* enzymatic addition and elimination reactions involving free carboxylic acids are anti with the proton entering from the *re* face,


Hydroxyl added
in front
(si face)
⁻OH
H
COO⁻
C=C
⁻OOC
H
*H⁺
Fumarate
Proton added
from behind
(re face)
HO
H
COO⁻
C—C
⁻OOC
H
*H
L-Malate

(13-12)

suggesting that there has been conservation throughout evolution of a single basically similar mechanism.[54,55]

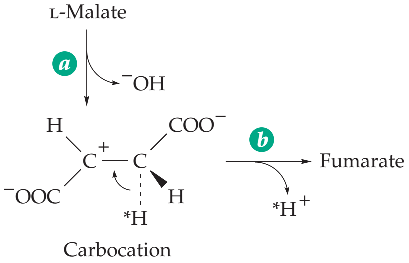
The pH dependence of the action of fumarate hydratase indicates participation of both an acidic and a basic group with pK_a values of 5.8 and 7.1.[56] See Chapter 9 for additional information. However, either anion or carbocation mechanisms might be possible. That the cleavage of the C–H bond is not rate limiting is suggested by the observation that malate containing ^{2}H in the *pro-R* position is dehydrated at the same rate as ordinary malate. If the anion mechanism (Eq. 13-13) is correct, the ^{2}H from the *pro-R* position of specifically labeled malate might be removed rapidly, while the loss of OH⁻ could be slower. If so, the ^{2}H would be "washed out" of L-malate faster than could happen by conversion to fumarate followed by rehydration to malate. In fact, the opposite was observed.


L-Malate
H⁺
H⁺
H
O
HO
C—O⁻
C—C
⁻OOC
H
Fumarate
H₂O
Carbanion
O⁻
H
C—O⁻
HO
C—C
⁻OOC
H
aci-Carboxylate
resonance form

(13-13)

The hydroxyl group was lost rapidly and the ^{2}H more slowly. This result suggested the carbocation ion mechanism of Eq. 13-14.

While no *primary* isotope rate effect was observed, when the hydrogen at C-2 or the *pro-S* hydrogen at C-3 of malate was replaced by ^{2}H or ^{3}H distinct secondary isotope effects were seen. Thus, $k(^1H) / k(^2H) = 1.09$ for both the *pro-S* and the C-3 hydrogen atoms.[57] These findings appeared to support the carbocation

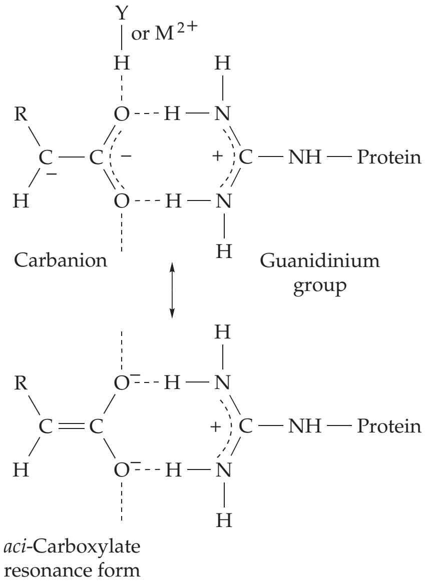

L-Malate
a
⁻OH
H
COO⁻
C⁺—C
⁻OOC
H
*H
b
Fumarate
*H⁺
Carbocation

(13-14)

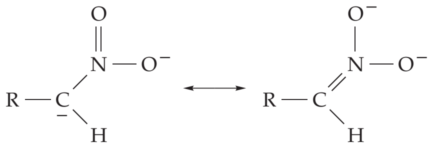
mechanism for the reasons considered in Chapter 12 and suggested that step *b* of Eq. 13-14 is rate limiting. The relative values of V_{max} for hydration of fumarate, fluorofumarate, and difluorofumarate (104, 410, and 86 mol ml^{-1} min^{-1} mg^{-1}) also seemed to support the carbocation mechanism.[58]

However, it has become clear that protons removed from a substrate to a basic group in a protein need not exchange rapidly with solvent (see Eq. 9-102). In fact, the proton removed by fumarate hydratase from malate is held by the enzyme for relatively long periods of time. Its rate of exchange between malate and solvent is slower than the exchange of a bound fumarate ion on the enzyme surface with another substrate molecule from the medium.[59] Thus, the overall rate is determined by the speed of dissociation of products from the enzyme and we cannot yet decide whether removal of a proton precedes or follows loss of OH⁻.

Two new lines of evidence suggest that proton abstraction comes first. A careful study of both ^{18}O and ^{2}H isotope effects[60] supports the carbanion intermediate, as does the strong inhibition by anions of 3-nitropropionate and 3-nitro-2-hydroxypropionate.[61] To provide a good electron sink the carboxylate group adjacent to the proton that is removed by fumarate hydratase must either be actually protonated in the enzyme–substrate (ES) complex or paired with and hydrogen bonded to a positively charged group.


Y
or M²⁺
H
H
R
O---H—N
C—C
+ C—NH—Protein
H
O---H—N
H
Carbanion
Guanidinium
group
H
R
O⁻--H—N
C=C
+ C—NH—Protein
H
O⁻--H—N
H
aci-Carboxylate
resonance form

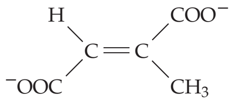
The anion formed by removal of the 3-H is analogous to the enolate anion of Eq. 13-6 and has a strong structural similarity to the readily formed anions of organic nitro compounds. The nitronate anions may, perhaps, be regarded as transition state inhibitors.



Nitronate anion

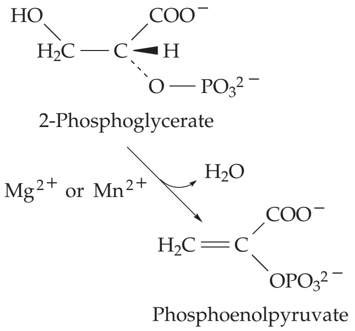
There is still a third possible mechanism for the fumarate hydratase reaction. The proton and hydroxyl groups may be added *simultaneously* in a concerted reaction.[62] However, observed kinetic isotope effects are not consistent with this mechanism.[60] In 1997 the structure of fumarase C of *E. coli* was reported.[52,53] Each active site of the tetrameric enzyme is formed using side chains from three different subunits. The H188 imidazole is hydrogen bonded to an active site water molecule and is backed up by the E331 carboxylate which forms a familiar catalytic pair. However, these results have not clarified the exact mode of substrate binding nor the details of the catalytic mechanism. Structural studies of fumarate hydratase from yeast[53a] and the pig[53b] are also in progress.

Some other fumarate-forming reactions. Other enzymes that catalyze elimination reactions that produce fumarate are **aspartate ammonia-lyase** (aspartase),[63] **argininosuccinate lyase** (Fig. 24-10, reaction *g*),[64,65] and **adenylosuccinate lyase** (Fig. 25-15). In every case it is NH_3 or an amine, rather than an OH group, that is eliminated. However, the mechanisms probably resemble that of fumarate hydratase. Sequence analysis indicated that all of these enzymes belong to a single **fumarase-aspartase family**.[64,65] The three-dimensional structure of aspartate ammonia-lyase resembles that of fumarate hydratase, but the catalytic site lacks the essential H188 of fumarate hydratase. However, the pK_a values deduced from the pH dependence of V_{max} are similar to those for fumarase.[64] **3-Methylaspartate lyase** catalyzes the same kind of reaction to produce ammonia plus *cis*-mesaconate.[63] Its sequence is not related to that of fumarase and it may contain a dehydroalanine residue (Chapter 14).[66]



cis-Mesaconate

Enolase. A key reaction in the metabolism of sugars is the dehydration of 2-phosphoglycerate to form **phosphoenolpyruvate** (PEP), the phospho derivative of the enolic form of pyruvic acid:



(13-15)

A carbanionic intermediate has often been suggested for this enzyme.[67,68] However, despite measurements of kinetic isotope effects and many other experiments it has been difficult to establish a detailed mechanism.[68]

Enolase has a complex metal ion requirement,[68a,69] usually met by Mg^{2+} and Mn^{2+}. From NMR studies of the relaxation of water protons, it was concluded that a Mn^{2+} ion coordinates two rapidly exchangeable water molecules in the free enzyme. When substrate binds, one of these water molecules may be immobilized and may participate in an addition reaction that forms phosphoenolpyruvate (reverse of reaction 13-15). A tightly bound "conformational" metal ion is located in the known three-dimensional structure in such a

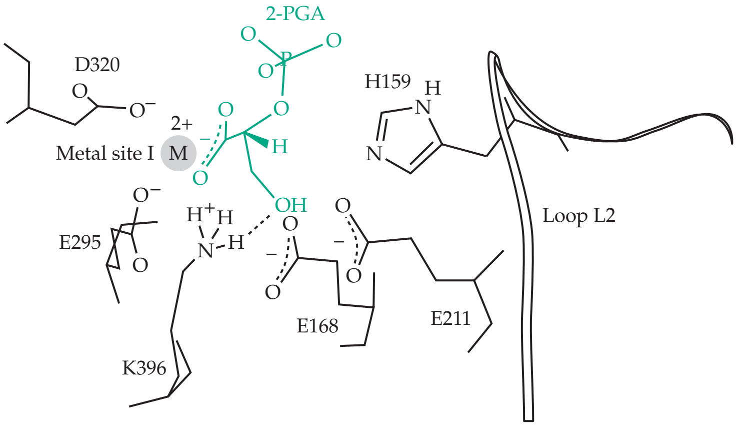


Figure 13-2 View of the active site of yeast enolase containing a bound molecule of 2-phospho-D-glycerate. The catalytic magnesium ion is at the left but the "conformational" metal is not visible here. The imidazole group of His 159 serves as the catalytic base and the $-NH_3{}^+$ of Lys 396 or Lys 345[73b] as the catalytic acid. From Vinarov and Nowak.[69]

way that it might function in this manner. A more loosely bound "catalytic" metal ion is also essential.[69–72]

Figure 13-2 shows a view of the active site of yeast enolase occupied by a molecule of bound 2-phosphoglycerate. Histidine 159 is probably the catalytic base that removes the α-proton to form an *aci*-anion which is stabilized by interaction with the catalytic Mg^{2+} ion. A protonated lysine 396 amino group may be the catalytic acid.[68a,69] The active site is surrounded by a complex hydrogen-bonded network.[73] As is discussed in Section B, a large number of other enzymes belong to an **enolase superfamily** of enzymes. Among them is **glucarate dehydratase**, which initiates a pathway for catabolism of D-glucarate and galactarate in *E. coli*.[73a]

Pectate lyase and related enzymes. A group of polysaccharide lyases cleave the chains of polymers of uronic acids with 1,4 linkages such as pectins, hyaluronan, heparin,[74,75] and dermatan sulfate (Fig. 4-11). These bacterial enzymes also employ an elimination mechanism.[76] The geometry of the β-linked galacturonic acid units of pectin is favorable for *anti* elimination of the 5-H and the *O*-glycosyl group in the 4 position (Eq. 13-16). However, the corresponding **hyaluronate lyase** (hyaluronidase) acting on glucuronic acid residues causes a *syn* elimination. These results suggest the formation of anionic intermediates which can eliminate a substituent from either the equatorial or axial position of the sugar ring. Hyaluronate lyase from the pathogenic *Streptococcus pneumoniae* apparently utilizes an imidazole group as the catalytic base and a tyrosine side chain as the proton donor in the reaction.[76a,76b]

Pectate lyase C from the plant pathogen *Erwinia*, which causes soft-rot in many different plants, has a parallel β barrel structure (Fig. 13-3)[77] which is similar to that of the tailspike protein shown in Fig. 2-17 and represents what may be a very large structural family of proteins.[78] The location of the active site is not obvious from the structure but on the basis of the effects of a large number of mutations, it is thought to be near the Ca^{2+}- binding site labeled in Fig. 13-3. Replacement of any of three Asp or Glu and three Lys residues in this region leads to loss of or reduction in catalytic activity.[79]

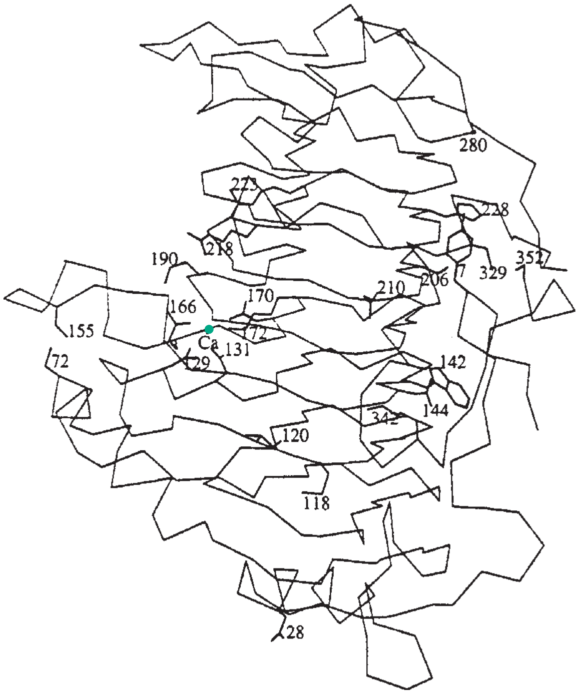
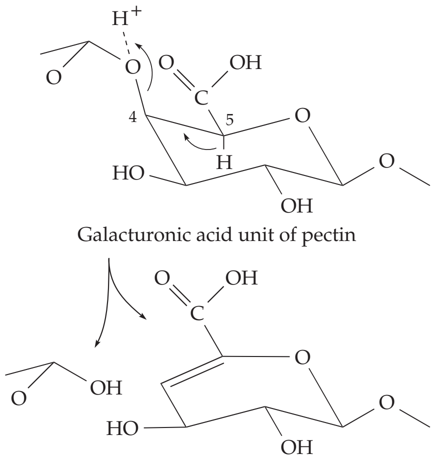


Figure 13-3 The three-dimensional structure of pectate lyase C showing locations of amino acids substituted by oligonucleotide-directed mutation of the cloned gene. The green dot labeled Ca is the Ca^{2+}-binding site.[79] Courtesy of Frances Jurnak.



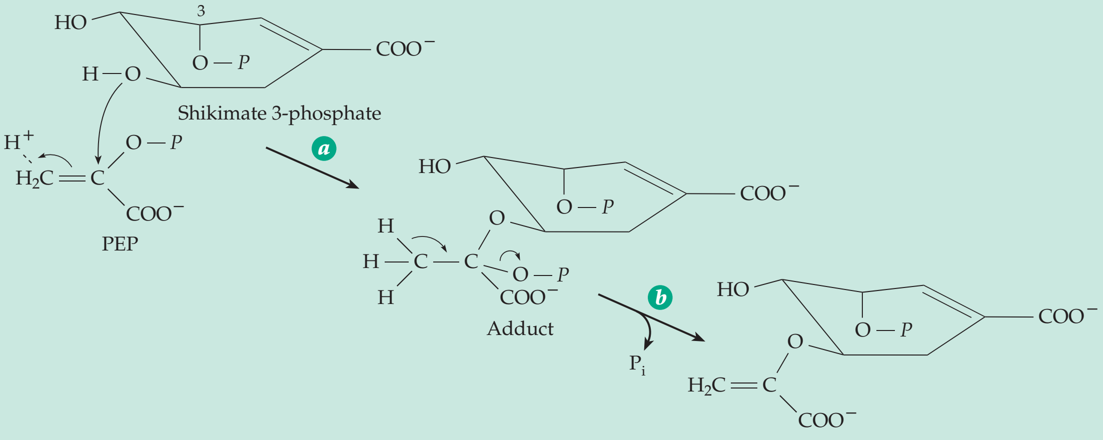
(13-16)

6. Aconitase and Related Iron – Sulfur Proteins

Two consecutive reactions of the citric acid cycle (Fig. 10-6), the dehydration of citrate to form *cis*-aconitate and the rehydration in a different way to form isocitrate (Eq. 13-17), are catalyzed by aconitase (aconitate hydratase). Both reactions are completely stereospecific. In the first (Eq. 13-17, step *a*), the *pro-R* proton from C-4 (stereochemical numbering) of citrate is removed and in step *c* isocitrate is formed. Proton addition is to the *re* face in both cases.

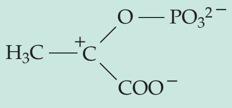
As with fumarate hydratase, the enzyme holds the abstracted proton (for up to 7×10^{-5} s) long enough so that a *cis*-aconitate molecule sometimes diffuses from the enzyme and (if excess *cis*-aconitate is present) is replaced by another. The result is that the new *cis*-aconitate molecule sometimes receives the proton (intermolecular proton transfer). The proton removed

BOX 13-B EPSP SYNTHASE AND THE HERBICIDE GLYPHOSATE

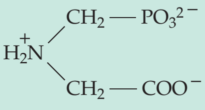


5-Enolpyruvoylshikimate-3-phosphate (EPSP)

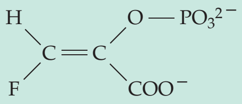
The reversible reaction of phosphoenolpyruvate (PEP) with shikimate 3-phosphate is a step in the synthesis of the aromatic amino acids (see Fig. 25-1). The chemical mechanism indicated was proposed by Leaven and Sprinson in 1964.[a] Step *a* is unusual[b] because it involves protonation on the methylene carbon of PEP and addition of a nucleophile at C-2, the opposite of the addition in the enolase reaction (Eq. 13-15, reverse). It is likely that formation of a cationic intermediate precedes that of the adduct shown. However, the structure of the adduct has been confirmed by isolation from the active site and by synthesis[c] and it has been observed in the active site by NMR spectroscopy.[d] The three-dimensional structure of EPSP synthase is known.[e,m] It is inhibited by the commercial herbicide **glyphosate** [*N*-(phosphonomethyl) glycine] whose structure is somewhat similar to that of the proposed carbocation that arises from PEP. The inhibitor (Z)-3-fluorophosphoenolpyruvate is a pseudosubstrate that reacts in step *a* to give a stable adduct unable to go through step *b* to form a product.[f–h] Glyphosate was for many years viewed as a transition-state analog but more recently has been shown to be a tight-binding noncompetitive inhibitor.[i]



Carbocation



Glyphosate



(*Z*)-3-Fluorophosphonoenolypyruvate

A related mechanism is utilized in the biosynthesis of UDP-muramic acid (Eq. 20-6).[j] There is an enolpyruvoyl adduct analogous to that of EPSP synthase; a proposed enolpyruvoyl-enzyme adduct with Cys 115 is not on the major path.[k,l] However, this enzyme is not inhibited by glyphosate.[i]

---

[a] Levin, J. G., and Sprinson, D. B. (1964) *J. Biol. Chem.* **239**, 1142–1150

[b] Barlow, P. N., Appleyard, R. J., Wilson, B. J. O., and Evans, J. N. S. (1989) *Biochemistry* **28**, 7985–7991

[c] Anderson, K. S., Sikorski, J. A., Benesi, A. J., and Johnson, K. A. (1988) *J. Am. Chem. Soc.* **110**, 6577–6579

[d] Appleyard, R. J., Shuttleworth, W. A., and Evans, J. N. S. (1994) *Biochemistry* **33**, 6812–6821

[e] Stallings, W. C., Abdel-Meguid, S. S., Lim, L. W., Shieh, H.-S., Dayringer, H. E., Leimgruber, N. K., Stegeman, R. A., Anderson, K. S., Sikorski, J. A., Padgette, S. R., and Kishore, G. M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5046–5050

[f] Alberg, D. G., Lauhon, C. T., Nyfeler, R., Fässler, A., and Bartlett, P. A. (1992) *J. Am. Chem. Soc.* **114**, 3535–3546

[g] Walker, M. C., Jones, C. R., Somerville, R. L., and Sikorski, J. A. (1992) *J. Am. Chem. Soc.* **114**, 7601–7603

[h] Ream, J. E., Yuen, H. K., Frazier, R. B., and Sikorski, J. A. (1992) *Biochemistry* **31**, 5528–5534

[i] Sammons, R. D., Gruys, K. J., Anderson, K. S., Johnson, K. A., and Sikorski, J. A. (1995) *Biochemistry* **34**, 6433–6440

[j] Samland, A. K., Amrhein, N., and Macheroux, P. (1999) *Biochemistry* **38**, 13162–13169

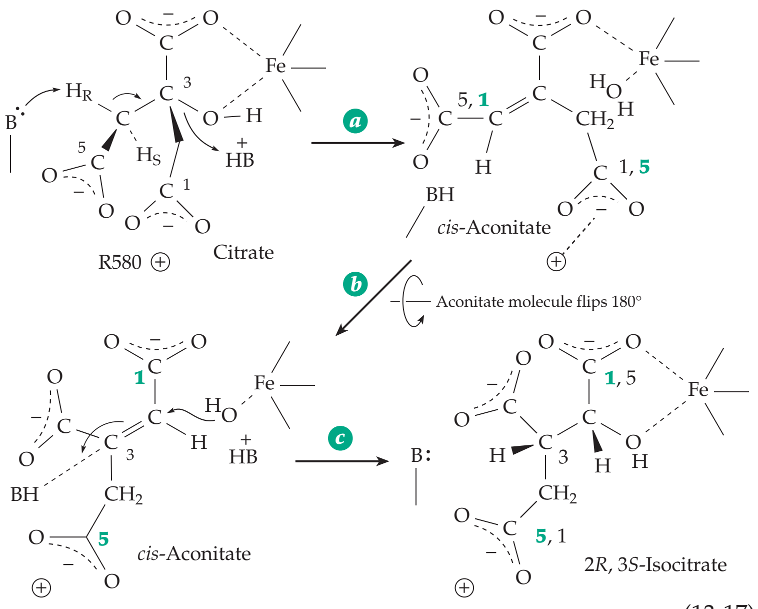
[k] Skarzynski, T., Kim, D. H., Lees, W. J., Walsh, C. T., and Duncan, K. (1998) *Biochemistry* **37**, 2572–2577

[l] Jia, Y., Lu, Z., Huang, K., Herzberg, O., and Dunaway-Mariano, D. (1999) *Biochemistry* **38**, 14165–14173

[m] Lewis, J., Johnson, K. A., and Anderson, K. S. (1999) *Biochemistry* **38**, 7372–7379

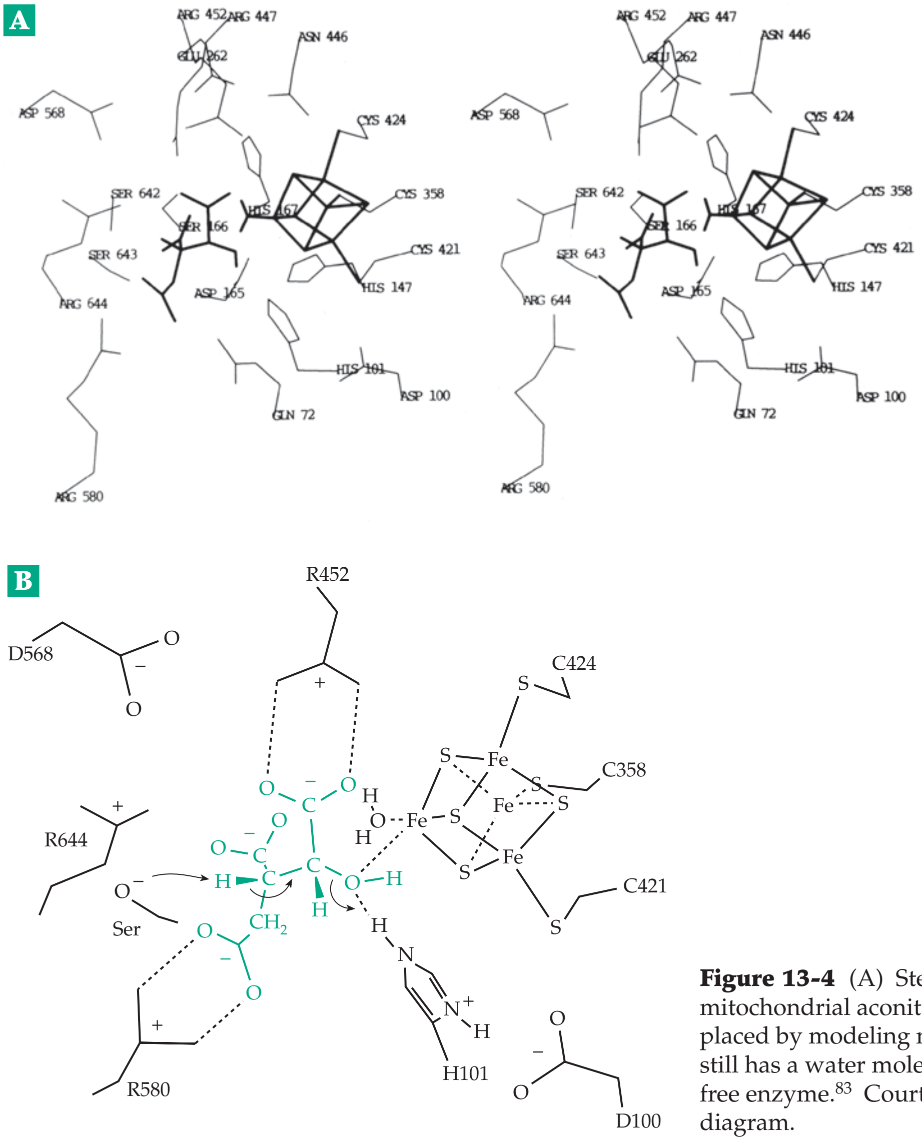
from citrate is often returned to the molecule in Eq. 13-17, step *b*, but the position of reentry is different from that of removal. Apparently, after the initial proton removal, the *cis*-aconitate that is formed "flips over" so that it can be rehydrated (Eq. 13-17, step *c*) with participation of the same groups involved in dehydration but with formation of the new product.[80,80a]

Aconitase contains iron in the form of an Fe_4S_4 iron-sulfur cluster (Fig. 13-4).[81–83] However, the enzyme is usually isolated in a form that does not show its maximum activity until it has been incubated with ferrous iron (Fe^{2+}). The inactive form of the enzyme is thought to contain an Fe_3S_4 cluster (Chapter 16) which is converted back to the Fe_4S_4 cluster by the incubation



B:
H_R
H_S
Fe
HB
R580
Citrate
a
cis-Aconitate
BH
b
Aconitate molecule flips 180°
cis-Aconitate
c
B:
2R, 3S-Isocitrate

(13-17)



A
ARG 452
ARG 447
ASN 446
GLU 262
ASP 568
CYS 424
SER 642
CYS 358
SER 166
HIS 167
SER 643
CYS 421
ARG 644
ASP 165
HIS 147
HIS 101
ASP 100
GLN 72
ARG 580
B
R452
D568
C424
C358
R644
Ser
C421
H101
R580
D100

Figure 13-4 (A) Stereoscopic view of the active site of mitochondrial aconitase with a molecule of L-isocitrate placed by modeling next to the Fe_4S_4 cluster. This cluster still has a water molecule bound to one iron atom as in the free enzyme.[83] Courtesy of C. D. Stout. (B) Interpretive diagram.

with Fe^{2+}.[84–87] From Mössbauer spectroscopy it was deduced that the iron atom that is lost upon inactivation is the binding site for the –OH group of citrate or isocitrate and of an adjacent carboxylate.[87] Since iron can engage in oxidation–reduction processes, aconitase could act by a mechanism different from others discussed here. However, an Fe–OH group could act much as does the Zn–OH of carbonic anhydrase (Eq. 13-4). Any redox chemistry may be involved in control of the enzyme.

X-ray studies have confirmed the proximity of the Fe_4S_4 cluster to bound substrate analogs.[83,88] The isocitrate shown in Fig. 13-4 was fitted into the active site by modeling. It shows the water molecule bound to the Fe_4S_4 cluster as observed for the free enzyme.[83] Notice that this cluster is held by three cysteine side chains (C358, C421, and C424) rather than the usual four (Chapter 16). The fourth iron atom is free to bind to water. When the enzyme acts on isocitrate this water must be displaced by the one generated from the substrate. The Fe probably acts as a polarizing electrophile that assists in the elimination but it is apparently the histidine–aspartate pair (H101–D100) that serves as the catalytic acid in generating the H_2O. What is the catalytic base? The only candidate seen in the structure is the Ser 642 –OH, dissociated to $–O^–$. Can this be correct? Mutational analysis supports the essential role of this side chain.[80a] The peptide NH and guanidinium groups of R644 appear to provide an "oxyanion hole" (Chapter 12) that stabilizes the negative charge.[86] Mutations also support the role of the His-Asp pair.[88,89] Another His–Asp pair (H167–D165) is located directly behind the bound substrate and Fe–OH_2 in Fig. 13-4A and is apparently also essential.[88]

Aconitase exists as both mitochondrial and cytosolic isoenzyme forms of similar structure. However, the cytosolic isoenzyme has a second function. In its apoenzyme form, which lacks the iron–sulfur cluster, it acts as the much-studied **iron regulatory factor**, or iron-responsive element binding protein (IRE-BP). This protein binds to a specific stem-loop structure in the messenger RNA for proteins involved in iron transport and storage (Chapter 28).[86,90]

Other enzymes in the aconitase family include **isopropylmalate isomerase** and **homoaconitase** enzymes functioning in the chain elongation pathways to leucine and lysine, both of which are pictured in Fig. 17-18.[90] There are also iron–sulfur dehydratases, some of which may function by a mechanism similar to that of aconitase. Among these are the two fumarate hydratases, **fumarases A and B,** which are formed in place of fumarase C by cells of *E. coli* growing anaerobically.[91,92] Also related may be bacterial L-serine and L-threonine dehydratases. These function without the coenzyme pyridoxal phosphate (Chapter 14) but contain iron–sulfur centers.[93–95] A **lactyl-CoA dehydratase** and some related iron-sulfur enzymes (pp. 861-862) may act via a mechanism related to that of vitamin B_{12}.

7. Addition to or Formation of Isolated Double Bonds

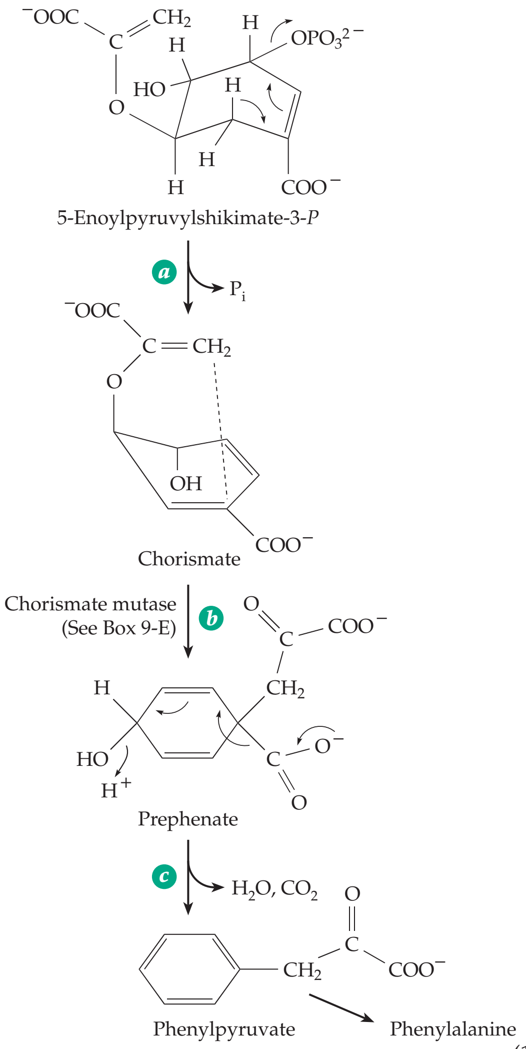
Only a few examples are known in which an enzyme induces addition to a double bond that is *not conjugated* with a carbonyl or carboxyl group. Pseudomonads have been observed to catalyze stereospecific hydration of oleic acid to D-10-hydroxystearate.[96] The addition is *anti* and the proton enters from the *re* face.

8. Conjugative and Decarboxylative Elimination Reactions

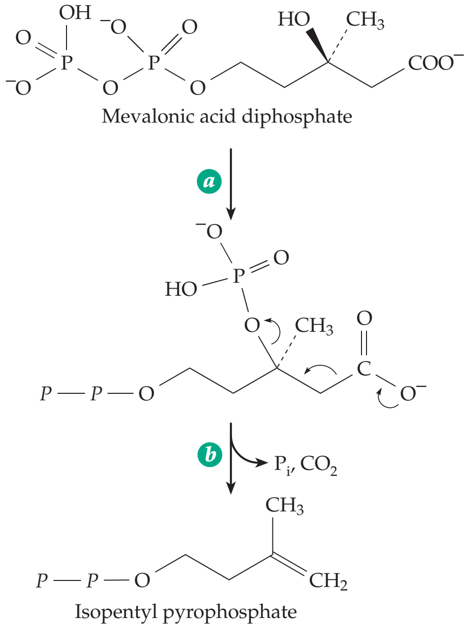
Elimination can occur if the electrophilic and nucleophilic groups to be removed are located not on adjacent carbon atoms but rather are separated from each other by a pair of atoms joined by a double bond. Such a conjugative elimination of phosphate is the last step in the biosynthesis of **chorismate** (Eq. 13-18, step *a*). Chorismate is converted to **prephenate** (see Box 9-E) which undergoes a conjugative and decarboxylative elimination (Eq. 13-18, step *c*) with loss of both water and CO_2 to form **phenylpyruvate**, the immediate biosynthetic precursor of phenylalanine. These reactions provide a good example of how *elimination reactions can be used to generate aromatic groups*. In fact, this is the usual method of synthesis of aromatic rings in nature.

Notice the stereochemistry of Eq. 13-18, step *a*. Orbital interaction rules predict that if the elimination is a concerted process it should be syn. The observed anti elimination suggests a more complex mechanism involving participation of a nucleophilic group of the enzyme.[97]

Elimination usually involves loss of a proton together with a nucleophilic group such as –OH, $–NH_3^+$, phosphate, or pyrophosphate. However, as in Eq. 13-18, step *c*, electrophilic groups such as $–COO^–$ can replace the proton. Another example is the conversion of **mevalonic acid-5-pyrophosphate** to **isopentenyl pyrophosphate** (Eq. 13-19): This is a key reaction in the biosynthesis of isoprenoid compounds such as cholesterol and vitamin A (Chapter 22). The phosphate ester formed in step *a* is a probable intermediate and the reaction probably involves a carbocationic intermediate generated by the loss of phosphate prior to the decarboxylation.


5-Enoylpyruvylshikimate-3-P
a
P_i
Chorismate
Chorismate mutase (See Box 9-E)
b
Prephenate
c
H_2O, CO_2
Phenylpyruvate
Phenylalanine

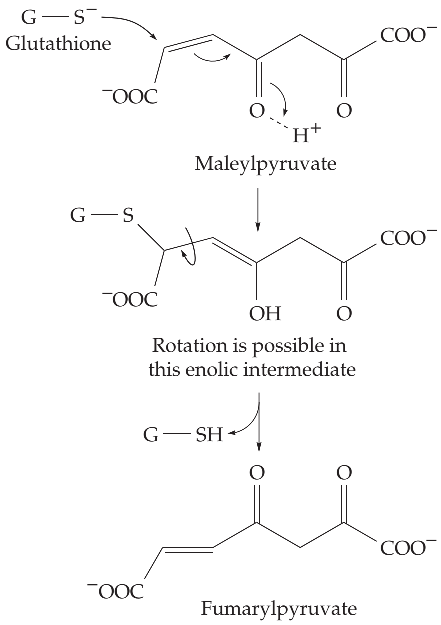
(13-18)


Mevalonic acid diphosphate
a
b
P_i, CO_2
Isopentyl pyrophosphate

(13-19)

9. Isomerization Assisted by Addition

An interesting use is made of addition to a double bond by glutathione-dependent ***cis–trans*** **isomerases**.[76] One of them converts **maleate** to fumarate with a turnover number of 300 s^{-1}. Similar enzymes, which participate in bacterial breakdown of aromatic compounds (Fig. 25-7), isomerize **maleylacetoacetate** and **maleylpyruvate** to the corresponding fumaryl derivatives (Eq. 13-20). The – SH group of bound glutathione is thought to add to the double bond. Rotation can then occur in the enolic intermediate. Thiocyanate ion catalyzes the isomerization of maleic acid nonenzymatically, presumably by a similar mechanism.


Glutathione
Maleylpyruvate
Rotation is possible in this enolic intermediate
G — SH
Fumarylpyruvate

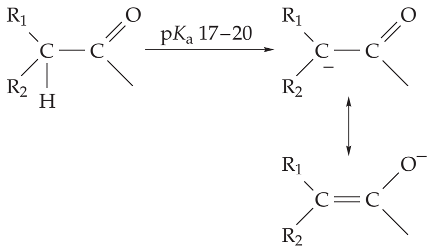
(13-20)

10. Reversibility of Addition and Elimination Reactions

Many addition and elimination reactions, e.g., the hydration of aldehydes and ketones, and reactions catalyzed by lyases such as fumarate hydratase are strictly reversible. However, biosynthetic sequences are often nearly irreversible because of the elimination of inorganic phosphate or pyrophosphate ions. Both of these ions occur in low concentrations within cells so that the reverse reaction does not tend to take place. In decarboxylative eliminations, carbon dioxide is produced and reversal becomes unlikely because of the high stability of CO_2. Further irreversibility is introduced when the major product is an aromatic ring, as in the formation of phenylpyruvate.

B. Enolic Intermediates in Enzymatic Reactions

Enzymologists have freely proposed enolate anions, enols, and enamines as intermediates for many years. Such intermediates have been demonstrated for some nonenzymatic acid- or base-catalyzed reactions, but how can enzymes form enolates at pH 7 without the use of strong acids or bases? The microscopic pK_a value of an α-hydrogen in a ketone or aldehyde is about 17–20.[72,98,99]



(13-21)

It will be similar for an acyl-CoA. However, for a carboxylic acid, protonated on oxygen, the pK_a will be much higher (~22–25) and for carboxylate anions even higher (~29–32).[100]

1. Mandelate Racemase and Related Enzymes

The degradation of mandelic acid by the bacterium *Pseudomonas putida* (Chapter 25) is initiated by mandalate racemase, another $(\alpha/\beta)_8$-barrel protein.[101] X-ray structures of bound inhibitors together with modeling suggest that the side chain of Lys 264 is the catalytic base that abstracts the α-H from *S*-mandelate (Fig. 13-5) and that the catalytic pair of His 297 and Asp 270 acts as proton donor, or, in the reverse direction, as catalytic base for deprotonation of *R*-mandelate.[102–104] The enzyme is structurally a member of the enolase superfamily[105] and requires Mg^{2+} for activity. The pH dependence of the reaction velocity k_{cat} reveals two pK_a values in the ES complex; ~6.4 and ~10.0. These are the same for *S*- and *R*- isomers. How can the pK_a values be assigned? Do they each belong in part to His 297 and in part to Lys 166? Do other adjacent groups also share? In the K166R mutant the lower pK_a is raised to 8.0. The D270N mutant has lost all but 0.01% of its catalytic activity.[104]

The carboxylate group of the mandelate interacts with side chains of E317, K264, and a Mg^{2+} ion (Fig. 13-5). These may serve both to protonate the carboxylate and also to help stabilize an *aci* anion formed upon dissociation of the α-hydrogen (Eq. 13-22). Both pK_a values for mandelic acid and its enolic form and the equilibrium constants for enolization in water are known[107] and are given beside the arrows in Eq. 13-22. It is difficult to imagine how a base with a pK_a near neutrality could remove an α-proton with a pK_a of 22.0, which is 14 pK units away from the pK_a of the catalytic base. This corresponds to a thermodynamic barrier (Eq. 9-97) of 14 x 5.7 = 80 kJ/mol, making the reaction impossibly slow. In addition to this thermodynamic barrier the *rates* of dissociation of carbon acids are known to be slow, presumably because of the lack of hydrogen bond formation between the C–H proton and the catalytic base (Eq. 9-97). This *intrinsic* barrier (Chapter 9) for simple ketones has been estimated as 45 kJ/mol for a total barrier of ~125 kJ. However, the observed $\Delta G^{\ddagger}$ as estimated from Eq. 9-81 is only ~57 kJ/mol. The enzyme must catalyze the reaction by lowering the very high thermodynamic barrier and perhaps also by lowering the intrinsic barrier.[72,108]

Protonation of the carboxylate greatly decreases the microscopic pK_a for loss of the α-hydrogen as a proton to form the enolic *aci* acid (Eq. 13-22). Double protonation, although depending upon a pK_a of –8,

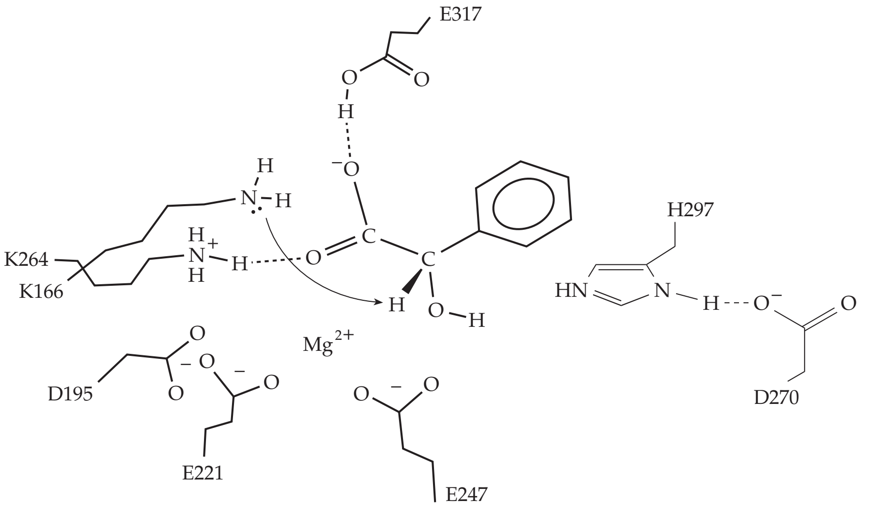
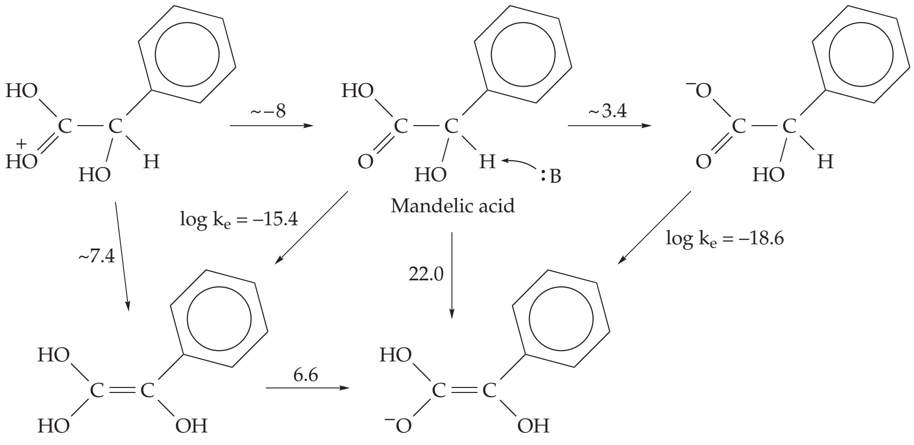


Figure 13-5 An *S*-mandelate ion in the active site of mandelate racemase. Only some of the polar groups surrounding the active site are shown. The enzyme has two catalytic acid–base groups. Lysine 166 is thought to deprotonate *S*-mandelate to form the *aci* anion, while His 297 deprotonates *R*-mandelate to form the same anion.[106]

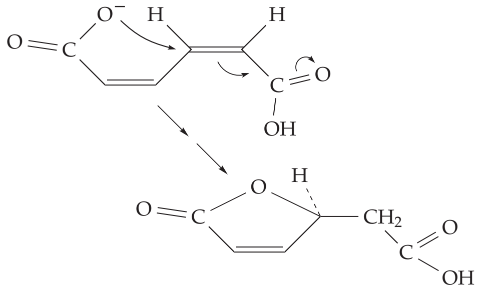


(13-22)

would reduce the pK_a of the α-hydrogen to ~ 7.4, making it very easy to enolize[109] and solving the thermodynamic problem. However, what forces could keep the substrate molecule in this unlikely state of protonation? Gerlt and Gassman proposed formation of a strong, short hydrogen bond to a carboxylate oxygen.[110] Formation of such a bond would lead to an increased positive charge on the alpha carbon, lowering the pK_a and therefore the thermodynamic barrier. It could also lower the intrinsic barrier,[108] for example, by permitting, to some extent, the formation of a C—H---N hydrogen bond in the transition state.[111] Polarization by the Mg^{2+} ion may also be involved. The very strong electrostatic forces within the active site may be sufficient to explain the formation of enolate anions or enols as intermediates in many different enzymes.[108,112]

A number of other racemases and epimerases may function by similar mechanisms. While some amino acid racemases depend upon pyridoxal phosphate (Chapter 14), several others function without this coenzyme. These include racemases for aspartate,[113] glutamate,[114–115a] proline, phenylalanine,[116] and diaminopimelate epimerase.[117] Some spiders are able to interconvert D and L forms of amino acid residues in intact polypeptide chains.[118,119]

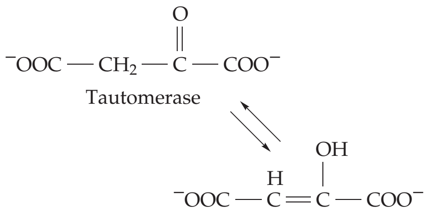
Another enzyme of the mandelate pathway of degradation of aromatic rings (Fig. 25-8) is the ***cis,cis*-muconate lactonizing enzyme** which catalyzes the reaction of Eq. 13-23. It has a three-dimensional structure almost identical to that of mandelate racemase but has incorporated an additional feature that allows formation, by addition, of the intermediate *aci*-acid.[101,105]



(13-23)

2. Isomerases

The isomerases that catalyze the simplest reactions are **tautomerases** that promote the oxo-enol (keto-enol) transformation. The widely distributed oxaloacetate tautomerase (Eq. 13-24) is especially active in animal tissues.[97,120] Oxaloacetate exists to a substantial extent in the enolic form: at 38°, ~ 6% enol, 13% oxo, and ~ 81% covalent hydrate.[120,121] A mammalian **phenylpyruvate tautomerase** has also been investigated.[122]

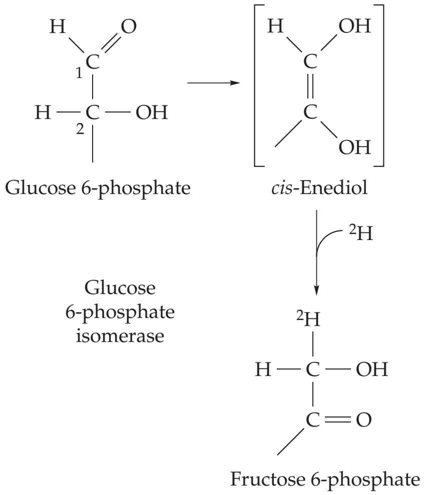


$$^-OOC-CH_2-\overset{\overset{\displaystyle O}{\|}}{C}-COO^- \rightleftharpoons {}^-OOC-\overset{H}{C}=\overset{\overset{\displaystyle OH}{|}}{C}-COO^-$$

(13-24)

The oxidation of one functional group of a molecule by an adjacent group in the same molecule is a feature of many metabolic sequences. In most cases an enolic intermediate, formed either from a ketone as in Eq. 13-25 or by a dehydration reaction (see Eq. 13-32), is postulated.

Aldose–ketose interconversions. A metabolically important group of enzymes catalyze the interconversion of aldose sugars with the corresponding 2-ketoses (Table 10-1, reaction 4C). Several sugar phosphates undergo rapid isomerization. **Glucose 6-phosphate isomerase** (Eq. 13-25) appears to function in all cells with a high efficiency.[123–125a] The 132 kDa dimeric protein from muscle converts glucose 6-phosphate to



(13-25)

fructose 6-phosphate with a turnover number of ~ 10^3 s^{-1}. Hereditary defects in this enzyme cause a variety of problems that range from mild to very severe.[126] **Mannose 6-phosphate isomerase** also forms fructose 6-phosphate, while **ribose 5-phosphate isomerase**[127] interconverts the 5-phosphates of D-ribose and D-xylulose in the pentose phosphate pathways (Chapter 17). Other enzymes, most often metalloenzymes, catalyze the isomerization of free sugars.

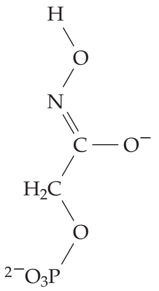
These enzymes vary widely in secondary and tertiary structure.[127a] Mannose-6-phosphate isomerase is a 45 kDa Zn^{2+}-containing monomer. The larger 65 kDa **L-fucose isomerase**, which also acts on D-arabinose, is a hexameric Mn^{2+}-dependent enzyme.[127a] **L-Arabinose isomerase** of *E. coli*, which interconverts arabinose and L-ribulose, is a hexamer of 60-kDa subunits[128] while the **D-xylose isomerase** of *Streptomyces* is a tetramer of 43-kDa subunits.[129] The nonenzymatic counterpart of the isomerization catalyzed by the enzyme is the base-catalyzed **Lobry deBruyn–Alberda van Ekenstein transformation** (Eq. 13-25).[130]

In 1895, Emil Ficher proposed an enediol intermediate for this isomerization. As would be expected, the enzyme-catalyzed isomerization of glucose-6-phosphate in ^2H_2O is accompanied by incorporation of deuterium into the product fructose 6-phosphate at C-1. In the reverse reaction ^{2}H-containing fructose 6-phosphate was found to react at only 45% of the rate of the ^{1}H-containing compound. Thus, the primary deuterium isotope effect expected for a rate-limiting cleavage of the C–H bond was observed (see Chapter 12, Section B,3).

When fructose 6-phosphate containing both ^{2}H and ^{14}C in the 1 position was isomerized in the presence of a large amount of nonlabeled fructose 6-phosphate, the product glucose 6-phosphate contained not only ^{14}C but also ^{2}H, and the distribution indicated that the ^{2}H had been transferred from the C-1 position into the C-2 position.[123] It was concluded that in over half the turnovers of the enzyme, ^{2}H removed from C-1 is put back on the same molecule at C-2. This intramolecular transfer of a proton suggests a syn transfer, the proton being removed and put back on the same side of the molecule. The carrier of the proton may be a histidine side chain.[131] This result, together with the known configuration of glucose at C-2, indicates that the intermediate is the *cis*-enediol and that addition of a proton at either C-1 or C-2 of the enediol is to the *re* face. However, mannose-6-phosphate isomerase catalyzes addition to the *si* face.

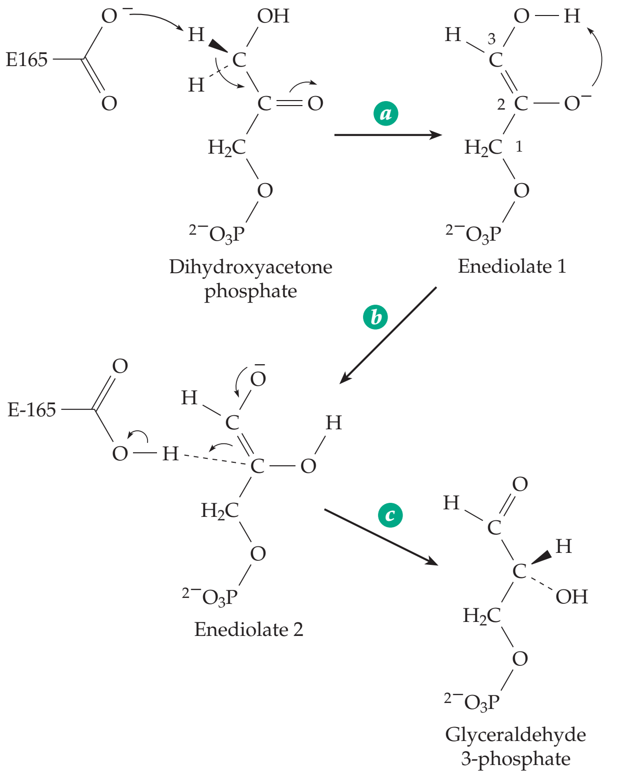
Catalysis of ring opening by isomerases. Glucose-6-phosphate isomerase catalyzes a second reaction, namely, the opening of the ring of the α-anomer of glucose 6-phosphate (one-half of the mutarotation reaction; Eq. 10-88). Noltmann suggested a concerted acid–base catalysis by two side chain groups as is indicated in Eq. 9-90. An NMR study showed that the isomerization of the α anomer occurs at least ten times faster than that of the β. Thus, β-glucose 6-phosphate is first converted to α-glucose 6-phosphate before it can be isomerized to fructose 6-phosphate.[132]

Triose phosphate isomerase. This dimeric 53-kDa enzyme interconverts the 3-phosphate esters of glyceraldehyde and dihydroxyacetone and is the fastest enzyme participating in glycolysis. Its molecular activity at 25° is ~2800 s^{-1} in the direction shown in Eq. 13-26 and ~250 s^{-1} in the reverse direction (the predominant direction in metabolism)[133] and is thought to operate at the diffusion-controlled limit (Chapter 9).[133] Each of the identical subunits consists of a striking $(\alpha/\beta)_8$ barrel (Fig. 13-6) with an active site at the carboxyl ends of the β strands.[134–137] Structures of the enzyme containing bound inhibitors such as phosphoglycolohydroxamate have also been determined (Fig. 13-6).[138] Triose phosphate isomerase is also one of the most investigated of all enzymes. Not only are its catalytic properties unusual but also there are known defects in the human enzyme[137] and it is also a potential target for antitrypanosomal drugs.[139]



Phosphoglycolohydroxamate (PGH), a tight-binding inhibitor (K_d = 15μM) that resembles enediolate 1 of Eq. 13-26

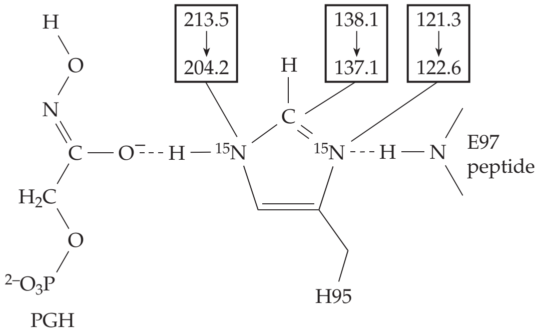
Although its high catalytic activity might favor intramolecular transfer of the proton removed by triose phosphate isomerase, little such transfer has



(13-26)

been observed.[140] This suggested that a relatively weak base such as a carboxylate group might serve as the proton acceptor at the C-3 position of dihydroxyacetone phosphate. Covalent labeling experiments and later X-ray studies have implicated Glu 165, whose carboxylate group is thought to remove the *pro-R* hydrogen atom from the hydroxymethyl group of dihydroxyacetone phosphate as indicated in Eq. 13-26. When this residue is replaced by Asp, most activity is lost.[141] Kinetic studies suggest that this carboxyl group has a pK_a of ~3.9 in the free enzyme. However, k_{cat} is independent of pH up to pH 10.[140] The pK_a of the phosphate also affects the binding (K_m), with only the phosphate dianion being a substrate. Study of the infrared spectrum of dihydroxyacetone bound to the enzyme revealed a shift of the carbonyl bands at 1733 cm^{-1} by about 20 cm^{-1} to ~1713 cm^{-1}. This might indicate a polarization and stretching of the carbonyl group by a positively charged histidine or lysine side chain of the protein.[142] The His 95 side chain is appropriately placed (Fig. 13-6) to function in this way[143] and its replacement by glutamine decreases catalytic activity by a factor of 400.[144] The –NH_3^+ group of lysine 12, another essential residue, is apparently needed for substrate binding.[136,145] After product is released, an isomerization within the enzyme is usually required to prepare it for acceptance of a new substrate. This may involve movement of protons between side chain groups or a conformational change or both. In the case of triose phosphate isomerase, a very rapid isomerization requiring about ~ 1 ms has been detected.[146]

To investigate further the function of His 95, Lodi and Knowles recorded the ^{13}C and ^{15}N NMR spectra of the three histidine rings[148] both in unligated enzyme and in the phosphoglycolohydroxamate (PGH) complex. The results were a surprise. The His 95 resonances did not change at all when the pH was changed from below 5 to 9.9, and the chemical shift values (shown above) indicate clearly that the ring is **unprotonated** in both free enzyme and in the complex. The key chemical shift values are shown on the diagram to the



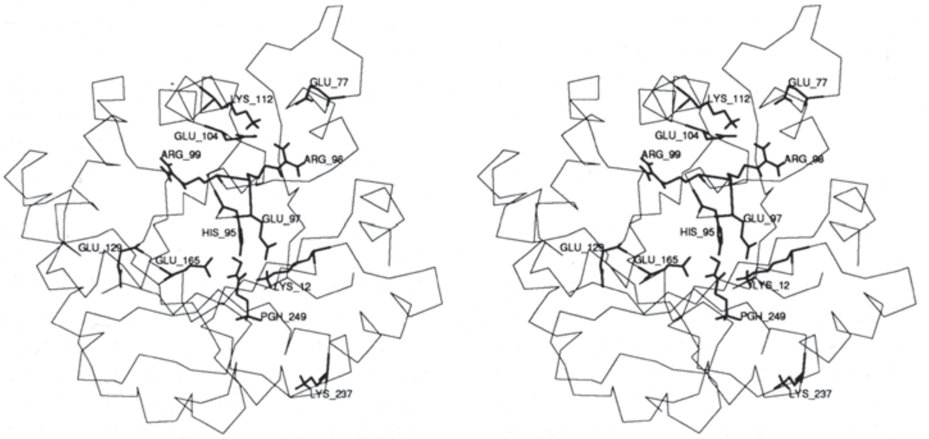
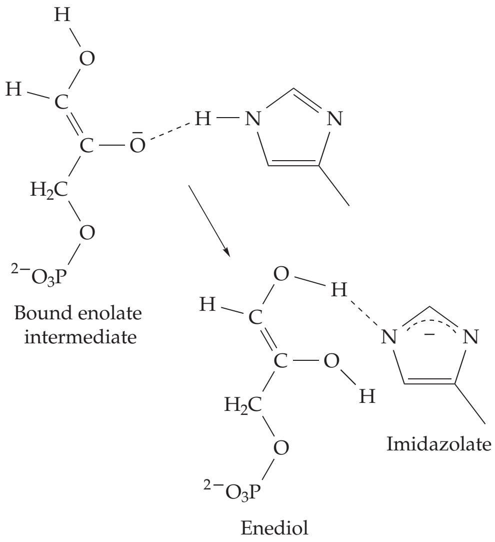


Figure 13-6 Stereoscopic view into the active site of triose phosphate isomerase showing side chains of some charged residues; PGH, a molecule of bound phosphoglycolohydroxamate, an analog of the substrate enolate.[138] The peptide backbone, as an alpha-carbon plot, is shown in light lines.[147] The $(\alpha / \beta)_8$-barrel structure is often called a TIM barrel because of its discovery in this enzyme. Courtesy of M. Karplus.

left; the arrows indicate the changes upon formation of the PGH complex. These results suggest that H95 is never protonated and that if it acts as the catalytic acid it does so by dissociating to an imidazolate ion (Eq. 13-27).[148]

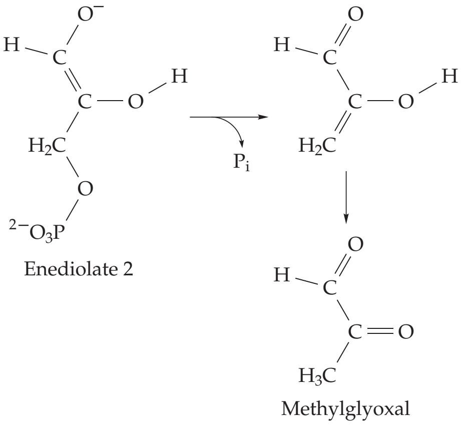


(13-27)

The imidazole would donate a proton to the enolate ion and then remove a proton from the resulting enediol. Although the proposed chemistry is unusual, it is argued that the high pK_a of the neutral imidazole acting as a proton donor would be matched with the high pK_a of the enol, permitting rapid reaction.[148] Nevertheless, two other possibilities exist. Theoretical calculations support the idea that the proton transfer between the two oxygens in the enediolate may occur *without catalysis by a proton donor*[149] as indicated in step *b* of Eq. 13-26. Another possibility is that a transient protonation of H95 by the adjacent E97 carboxyl group occurs and allows the histidine to participate in the proton transfer. Study of kinetic isotope effects has indicated coupled motion of protons and proton tunneling.[143]

Another detail should be mentioned. The active site of triose phosphate isomerase is formed by a series of loops connecting the α helices and β strands of the barrel. One of those loops, consisting of residues 167–176, folds over the active site after the substrate is bound to form a hinged lid that helps to hold the substrate in the correct orientation for reaction.[150–152] When the lid, which can be seen in Fig. 13-6, closes, the peptide NH of G171 forms a hydrogen bond to a phosphate oxygen atom of the substrate. This is only one of many known enzymes with deeply buried active sites that close in some similar fashion before a rapid reaction occurs.

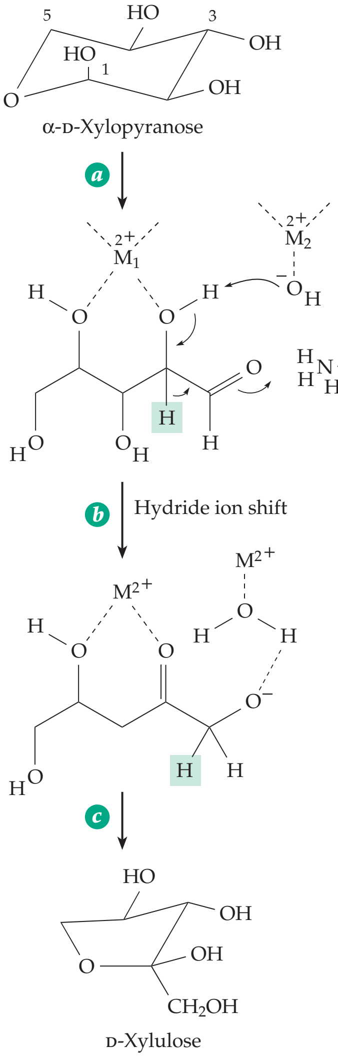
Although enzymes tend to be extremely specific they are not always completely able to avoid side reactions. Triose phosphate isomerase releases small amounts of **methylglyoxal** (Eq. 13-28), presumably as a result of elimination of phosphate from enediolate 2 of Eq. 13-26. Deletion of four residues from the hinged lid produced an enzyme in which this side reaction was increased 5.5-fold.[146]



(13-28)

Xylose isomerase and the hydride shift mechanism. This bacterial enzyme, which isomerizes D-xylose to D-xylulose, has an $(\alpha/\beta)_8$ barrel similar to that of triose phosphate isomerase. It is of industrial importance because it also catalyzes isomerization of D-glucose to the sweeter sugar D-fructose, a reaction used in preparation of high-fructose syrup. For this purpose the isomerase is often immobilized on an insoluble matrix such as diethylaminoethyl cellulose.[153] This enzyme, as well as other isomerases that act on free sugars, requires a metal ion such as Mg^{2+}, Co^{2+}, or Mn^{2+}. The three-dimensional structures, solved in several different laboratories,[129,154–158] show that there are two metal ions (ordinarily Mg^{2+}) about 0.5 nm apart and held by an array of glutamate and aspartate side chains. One glutamate carboxylate forms short ionic bonds to both metal ions and an essential histidine is also coordinated with metal ion 2. The three-dimensional structure is superficially similar to that of triose phosphate isomerase but there are major differences in properties. Xylose isomerase requires metal ions, acts on unphosphorylated non-ionic substrates, and does not catalyze detectable exchange of protons with the solvent. Furthermore, *X-ray structures do not show the presence of any catalytic base that could initiate formation of an intermediate enolate ion.* These facts suggested an alternative isomerization mechanism, one involving a hydride ion shift and well-known from studies of the Cannizarro reaction.[159] Because the non-ionic substrates and inhibitors bind weakly it has been difficult to obtain a clear picture of events in the active site. The substrate is the α-anomer of D-xylopyranose[155] and the enzyme catalyzes the opening of the sugar ring (step *a*, Eq. 13-29).[157] The details of this process are not clear but acid-base catalysis as in Eq. 9-87 is probable.

There is a consensus[156–158] that the open form of

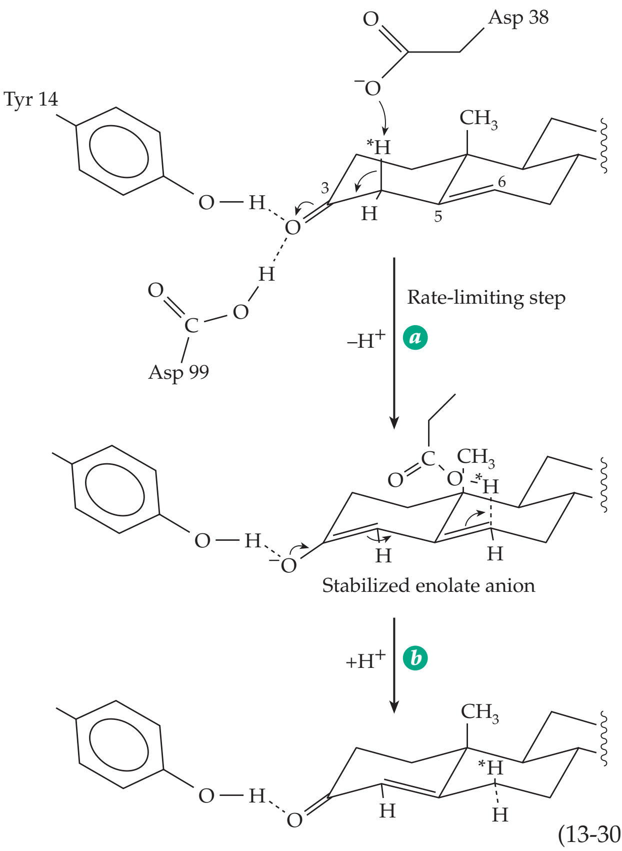


(13-29)

the sugar is bound to one of the two metal ions as shown in Eq. 13-29 and that a bound hydroxide ion on the second metal provides the catalytic base to remove a proton from the –OH at position 2. The isomerization occurs with the shaded hydrogen atom being shifted as H^- to the carbon atom of the carbonyl group. During the reaction metal ion 2 moves apart from metal ion 1 by ~0.09 nm,[129,157,158] another fact that has made analysis of X-ray data difficult. Protonation of the alkoxide ion at C-2 is followed by ring closure (step *c*, Eq. 13-29).

3-Oxosteroid isomerases. Cholesterol serves in the animal body as a precursor of all of the steroid hormones, including the 3-oxosteroids progesterone and testosterone (Chapter 22). The 3-hydroxyl group of cholesterol is first oxidized to an oxo group. This is followed by an essentially irreversible migration of the double bond in the 5,6-position into conjugation with the carbonyl group (Eq. 13-30) catalyzed by a 3-oxo-Δ^5-steroid isomerase. The small 125-residue enzyme from the bacterium *Pseudomonas testosteroni* has been studied by a great variety of methods.[160–164] The subunits associate as dimers or higher oligomers. The enzyme has a high content of nonpolar amino acid residues and is soluble in high concentrations of ethanol. This property is compatible with the location of the eukaryotic enzyme in the endoplasmic reticulum. The substrate binding site is a hydrophobic cavity.

Oxosteroid isomerase has a remarkably high molecular activity (~0.75 x 10^5 s^{-1}). Substrates containing ^{2}H in the 4 position react only one-fourth as rapidly as normal substrates. The large isotope effect suggested that cleavage of the C–H bond to form an enzyme-stabilized enolate anion is rate limiting (Eq. 13-30, step *a*).[160] The proton in the axial position at C-4 is removed preferentially but without complete stereospecificty.[165] The abstracted proton must be carried by a group on the enzyme and returned to the 6 position of the substrate, again in an axial orientation (Eq. 13-30, step *b*), a suprafacial transfer. Studies of the binding of spin-labeled substrate analogs, molecular modeling, and directed mutation indicate that Asp 38 is this proton carrier.[164,166] Little exchange of the proton with solvent is observed, presumably because of the extreme rapidity of the isomerase action. However, competitive inhibitors such as nortestosterone, whose double-bond arrangement is that of the product in Eq. 13-30, undergo exchange of one of the hydrogens at C-4 with the medium. The ultraviolet absorption band of the inhibitor at 248 nm is shifted to 258 nm upon combination with the enzyme, presumably as a result of formation of the

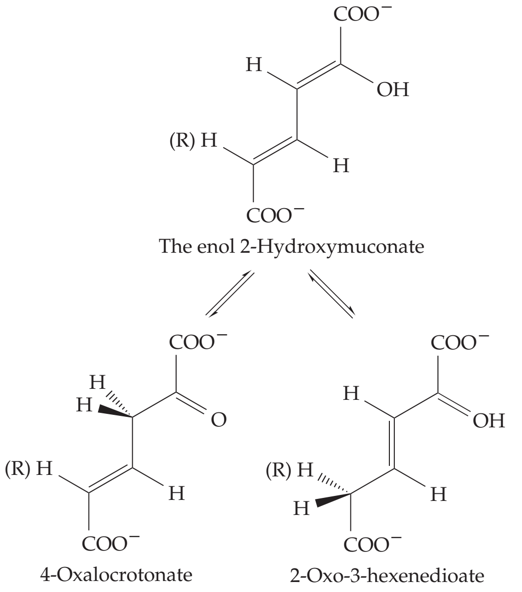


(13-30)

enolate anion. Tyrosine 14 has been identified as the catalytic acid, able to stabilize the enolate anion as depicted in Eq. 13-30.[161,162] A high-resolution NMR structure[167] and X-ray crystal structures[167a,b] have revealed an adjacent Asp 99 side chain that may assist.[167c]

Enzymatic isomerization of *cis*-aconitate to *trans*-aconitate apparently also involves proton abstraction,[165] with resonance in the anion extending into the carboxylic acid group. Its mechanism may be directly related to that of the oxosteroid isomerase. However, there are other 1,3-proton shifts in which neither a carbonyl nor a carboxyl group is present in the substrate (Eqs. 13-55, 13-56).

4-Oxalocrotonate tautomerase. This bacterial enzyme, which functions in the degradation of toluene (Chapter 25), is actually an isomerase. It catalyzes rapid interconversion of an unconjugated unsaturated α-oxoacid such as 4-oxalocrotonate with an intermediate enol (which may leave the enzyme) and the isomeric conjugated oxoacid (Eq. 13-31).[168–170] A related 5-carboxymethyl-2-hydroxymuconate isomerase catalyzes the same reaction when R = $-CH_2COO^-$ in Eq. 13-31. It is an unusually small enzyme consisting of only 62 amino acid residues. The pH dependence reveals pK_a values of 6.2 and 9.0 in the free enzyme and 7.7 and 8.5 in the ES complex.[171]



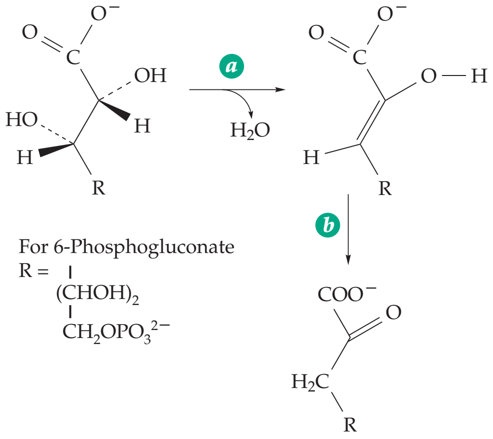
(13-31)

The pK_a of 6.2 has been associated with the amino-terminal proline 1.[172] 4-Oxalocrotonate tautomerase is one of a small group of enzymes that have been synthesized nonenzymatically with both L amino acids and as a mirror image constructed with D amino acids.[173]

3. Internal Oxidation–Reduction by Dehydration of Dihydroxyacids

When a carboxylic acid contains hydroxyl groups in both the α and β positions, dehydration leads to formation of an enol that can tautomerize to 2-oxo-3-deoxy derivatives of the original acid (Eq. 13-32). Thus, phosphogluconate dehydratase yields 2-oxo-3-deoxyphosphogluconate as the product. When the reaction is carried out in ^2H_2O the ^2H is incorporated with a random configuration at C-3, indicating that the enzyme catalyzes only the dehydration and that the tautomerization of the enol to the ketone is nonenzymatic.

The reaction of Eq. 13-32 initiates a unique pathway of sugar breakdown, the Entner–Douderoff pathway (Eq. 17-18), in certain organisms. The 6-phosphogluconate is formed by oxidation of the aldehyde group of glucose 6-phosphate. This pattern of oxidation of a sugar to an aldonic acid followed by dehydration according to Eq. 13-32 occurs frequently in metabolism. A related reaction is the dehydration of 2-phosphoglycerate by enolase (Eq. 13-15). In this case the product is phosphoenolpyruvate, a stabilized form of the enolic intermediate of Eq. 13-32 (when R = H).

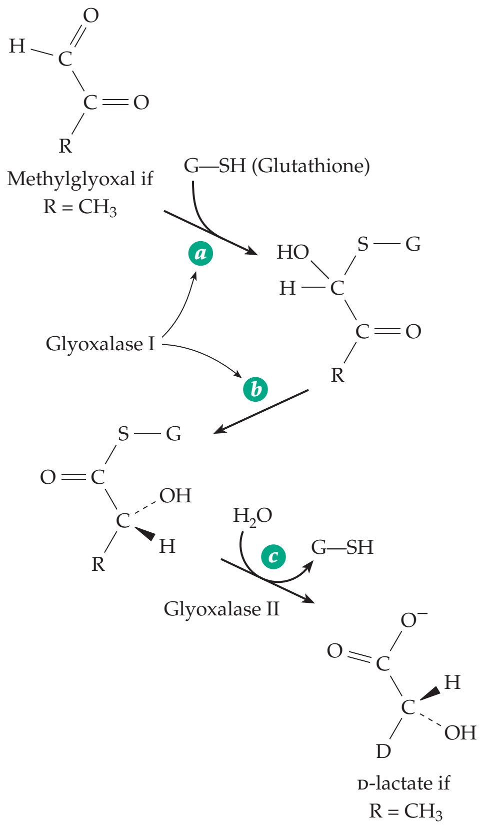


(13-32)

4. Formation and Metabolism of Methylglyoxal (Pyruvaldehyde)

The rather toxic methylglyoxal is formed in many organisms and within human tissues.[174] It arises in part as a side reaction of triose phosphate isomerase (Eq. 13-28) and also from oxidation of acetone (Eq. 17-7) or aminoacetone, a metabolite of threonine (Chapter 24).[175] In addition, yeast and some bacteria, including *E. coli*, have a **methylglyoxal synthase** that converts dihydroxyacetone to methylglyoxal, apparently using a mechanism similar to that of triose phosphate isomerase. It presumably forms enediolate 2 of Eq. 13-26, which eliminates inorganic phosphate to yield methylglyoxal as in Eq. 13-28.[176,176a] Methylglyoxal is converted to D-lactate by the two-enzyme **glyoxalase** system (Eq. 13-33).[177] The combined action of methylglyoxal synthase and the glyoxalases provides a bypass to the usual glycolysis pathway (Fig. 10-3). Although it does not provide energy to the cell, it releases inorganic phosphate from sugar phosphates that may accumulate under conditions of low phosphate because the free P_i concentration is too low to support the glyceraldehyde 3-phosphate dehydrogenase reaction (step *a* of Fig. 10-3).

Glyoxalase acts not only on methylglyoxal but also on other α-oxo-aldehydes. It is thought to be an important enzyme system that protects cells against these potentially dangerous metabolites. Glyoxalase consists of a pair of enzymes, **glyoxalase I** and **glyoxalase II**, which catalyze the reactions of Eq. 13-33. Each subunit of the 183-residue human glyoxalase I contains one tightly bound Zn^{2+} ion.[178] However, the *E.coli* enzyme[178a] is inactive with Zn^{2+} and maximally active with Ni^{2+}. The enzyme requires glutathione as a cofactor, and in step *a* of the reaction (Eq. 13-33) the glutathione adds to form a thiohemiacetal[179] which is then isomerized in step *b*. During this step retention of the abstracted proton is so complete that it was earlier thought to function by an intramolecular hydride ion shift as in xylose isomerase. More recent evidence favors an enolate anion intermediate. However, the three-dimensional structure is not related to that of other isomerases and the exact mechanism remains uncertain. Glyoxylase II, which catalyzes step *c* of Eq. 13-33, is an esterase.[175,180]

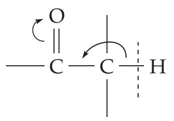


(13-33)

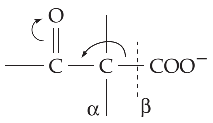
C. Beta Cleavage and Condensation

In Chapter 12 and in the preceding sections of this chapter we examined displacement and addition reactions involving nucleophilic centers on O, N, or S. Bonds from carbon to these atoms can usually be broken easily by acidic or basic catalysis. The breaking and making of C–C bonds does not occur as readily and the "carbon skeletons" of organic molecules often stick together tenaciously. Yet living cells must both form and destroy the many complex, branched carbon compounds found within them.

A major mechanistic problem in cleavage or formation of carbon–carbon bonds is the creation of a nucleophilic center on a carbon atom. The problem is most often solved by using the *activating influence of a carbonyl group to generate a resonance-stabilized enolate anion*. Just as the presence of a carbonyl group facilitates cleavage of an adjacent C–H bond, so it can also assist the cleavage of a C–C bond. The best known reactions of this type are the **aldol cleavage** and the **decarboxylation** of **β-oxo acids**. The latter has been referred to as β decarboxylation and its reverse as β carboxylation. In this book these terms have been extended to include other reactions by which bonds between the α and β carbon atoms of a carbonyl compound are broken or formed, and these will be referred to as **β cleavage and β condensation**.



Cleavage of C—H bond facilitated by a carbonyl group

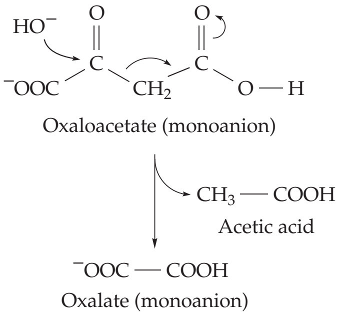


β Cleavage with release of CO_2

The β *condensation* reactions consist of displacement or addition reactions in which an enzyme-bound enolate anion acts as the nucleophile. We can group these condensation reactions into three categories as indicated by reaction types 5A, 5B, and 5C of Table 10-1.

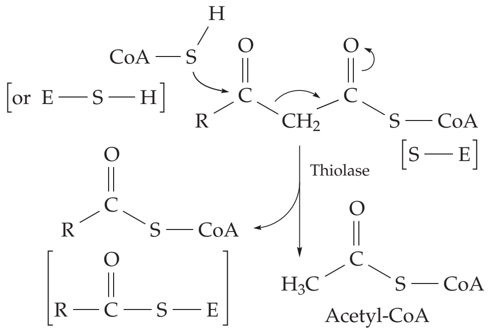
1. Displacement on a Carbonyl Group

A β-oxo acid, with proper catalysis, is susceptible to hydrolysis by attack of water on the carbonyl group. An example is the reaction catalyzed by **oxaloacetate acetylhydrolase** which has been isolated from *Aspergillus niger* (Eq. 13-34). A related cleavage is catalyzed by ribulose bisphosphate carboxylase (see Eq. 13-48).


HO⁻
O
O
C
C
⁻OOC
CH₂
O — H
Oxaloacetate (monoanion)
CH₃ — COOH
Acetic acid
⁻OOC — COOH
Oxalate (monoanion)

(13-34)

The **thiolases**[181] are lyases that cleave β-oxoacyl derivatives of CoA by displacement with a thiol group of another CoA molecule (Eq. 13-35). This is the chain cleavage step in the β oxidation sequence by which fatty acid chains are degraded (Fig. 10-4). Biosynthetic thiolases catalyze the condensation of two molecules of acetyl-CoA to form acetoacetyl-CoA, (see Eq. 17-5), a precursor to cholesterol and related compounds and to poly-β-hydroxybutyrate (Box 21-D). Because acetyl-CoA is a thioester, the reaction is usually described as a **Claisen condensation**. A related reaction, involving decarboxylation of malonyl-CoA, is a step in fatty acid synthesis. Since the thiolases are inhibited by –SH reagents it has been suggested that a thiol group in the enzyme reacts initially with the β-carbonyl group as in Eq. 13-35 to give an enzyme-bound *S*-acyl intermediate. The acyl group is then transferred to CoA in a second step.


H
CoA — S
[or E — S — H]
O
O
C
C
R
CH₂
S — CoA
[S — E]
Thiolase
O
C
R
S — CoA
[R — C — S — E]
O
C
H₃C
S — CoA
Acetyl-CoA

(13-35)

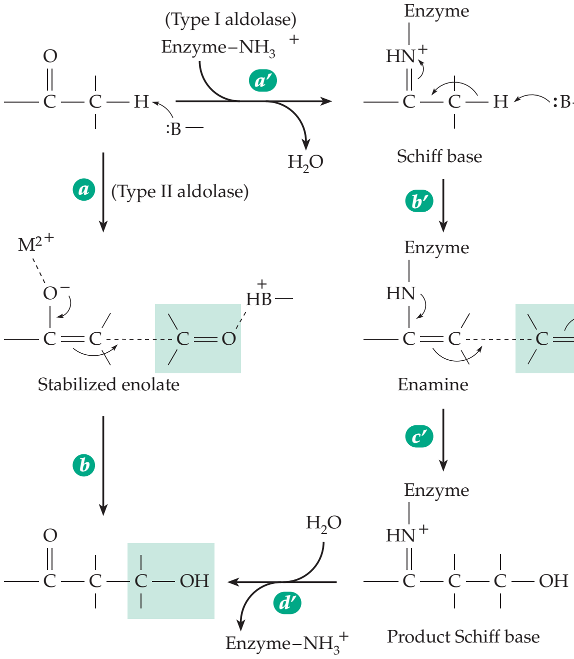
A very similar reaction is catalyzed by 3-hydroxy-3-methylglutaryl-CoA lyase (HMG-CoA lyase), which functions in the formation of acetoacetate in the human body (Eq. 17-5, step *c*) and also in the catabolism of leucine (Fig. 24-18)[182,183] and in the synthesis of **3-hydroxy-3-methylglutaryl-CoA**, the presursor of cholesterol (Eq. 17-5, step *b*)[183a]

2. Addition of an Enolate Anion to a Carbonyl Group or an Imine; Aldolases

The **aldol condensation** (Eq. 13-36), which is also illustrated in Box 10-D, is one of the more common reactions by which C–C bonds are formed[183b] and, in the reverse reaction, cleaved in metabolism. Aldolases are classified into two major types. The type II aldolases are metal-ion dependent, the metal ion stabilizing the intermediate enolate ion (Eq. 13-36, steps *a* and *b*). Type I aldolases, which include the most studied mammalian enzymes, have a more complex mechanism involving intermediate Schiff base forms (Eq. 13-36, steps *a′*, *b′*, *c′*, *d′*).[184] The best known members of this group are the **fructose bisphosphate aldolases** (often referred to simply as aldolases), which cleave fructose-1,6-P_2 during glycolysis (Fig. 10-2, step *e*).

These enzymes have been found in all plant and animal tissues examined and are absent only from a few specialized bacteria. Three closely related isoenzymes are found in vertebrates.[185,186] The much studied rabbit muscle aldolase A is a 158-kDa protein tetramer of identical peptide chains.[186,187] Aldolase B, which is lacking in hereditary fructose intolerance, predominates in liver and isoenzyme C in brain.[185]

Treatment with sodium borohydride of the enzyme-substrate complex of aldolase A and dihydroxyacetone phosphate leads to formation of a covalent linkage between the protein and substrate. This and other evidence suggested a Schiff base intermediate (Eq. 13-36). When ^{14}C-containing substrate was used, the borohydride reduction (Eq. 3-34) labeled a lysine side chain in the active site. The radioactive label was followed through the sequence determination and was found on Lys 229 in the chain of 363 amino acids.[186,188–188b] The enzyme is another $(\alpha / \beta)_8$-barrel protein and the side chain of Lys 229 projects into the interior of the barrel which opens at the C-terminal ends of the strands. The conjugate base form of another lysine, Lys 146, may represent the basic group B in Eq. 13-36,

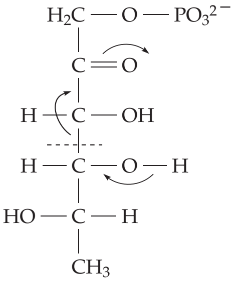

(Type I aldolase)
Enzyme–NH₃⁺
a′
H₂O
O
C — C — H
:B —
Enzyme
HN⁺
C — C — H
:B—
Schiff base
a
(Type II aldolase)
b′
M²⁺
O⁻
C = C
C = O
HB⁺—
Stabilized enolate
Enzyme
HN
C = C
C = O
HB⁺—
Enamine
b
c′
O
C — C — C — OH
H₂O
d′
Enzyme–NH₃⁺
Enzyme
HN⁺
C — C — C — OH
Product Schiff base

(13-36)

step *b*. Another possibility is that the adjacent phosphate group of the substrate acts as the acid–base catalyst for this step.[189] Aldolase A has been altered by mutations into a monomeric form that retains high catalytic activity,[190] something that has not often been accomplished for oligomeric enzymes.

The **type II aldolases** are not inactivated by sodium borohydride in the presence of substrate. A probable function of the essential metal ion is to polarize the carbonyl group as indicated in Eq. 13-36. In both yeast and *E. coli* aldolase the Zn^{2+} is held by 3 imidazole groups.[191,191a] An arginine side chain is a conserved residue involved in substrate binding in several class II aldolases.[192] Some blue-green algae contain both types of aldolase, as do the flagellates *Euglena* and *Chlamydomonas*.

The catabolism of L-fucose by *E. coli* requires cleavage of L-fuculose-1-phosphate to form dihydroxyacetone phosphate and D-lactaldehyde by a class II aldolase.[193]



L-Fuculose-1-phosphate

The mechanism of chain cleavage proposed on the basis of the structure and modeling[193] is illustrated in Fig. 13-7.

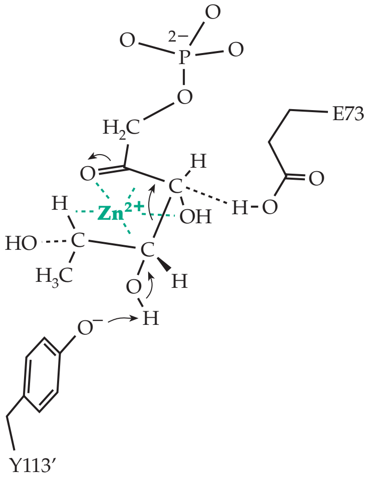
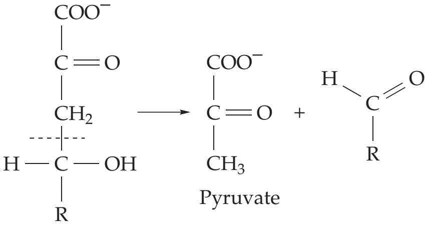


Figure 13-7 Interaction of the bound zinc ion of L-fuculose-1-phosphate aldolase and catalytic side chains with the substrate in the active site of the enzyme as revealed by X-ray crystallography and modeling. See Dreyer and Schulz.[193]

Special aldolases of both classes cleave and form C–C bonds throughout metabolism. Several of them act on 2-oxo-3-deoxy substrates forming pyruvate as one product (Eq. 13-37). The aldehyde product varies. In the Entner–Doudoroff pathway of carbohydrate metabolism (Chapter 21) 3-deoxy-2-oxo-6-phosphogluconate (KDPG) is cleaved to pyruvate and 3-phosphoglyceraldehyde.[194] The same products arise from the corresponding phosphogalactonate derivative.[195] The subunits of the trimeric KDPG aldolase have an $(\alpha / \beta)_8$-barrel structure similar to that of eukaryotic fructose 1,6-bisphosphate aldolase.[194] The 8-carbon sugar acid "KDO" of bacterial cell walls (Fig. 4-26) is cleaved by another aldolase. The catabolism of hydroxyproline leads to 4-hydroxy-2-oxoglutarate, which is cleaved to pyruvate and glyoxylate.[196] In the catabolism of deoxynucleotides, another aldolase converts 2-deoxyribose 5-phosphate to acetaldehyde and glyceraldehyde 3-phosphate.[197]

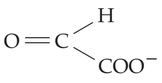
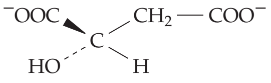
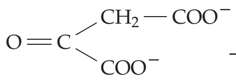
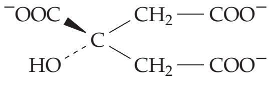
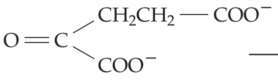
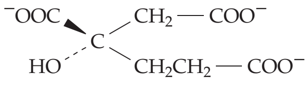
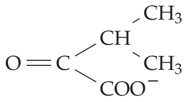
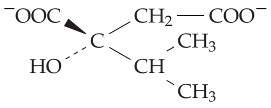
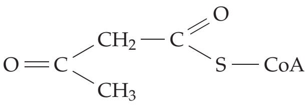
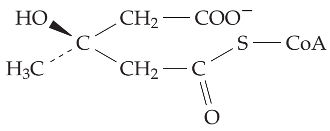


(13-37)

Closely related to aldolases is **transaldolase**, an important enzyme in the pentose phosphate pathways of sugar metabolism and in photosynthesis. The mechanism of the transaldolase reaction (Eq. 17-15) is similar to that used by fructose-1,6-bisphosphate aldolase with a lysine side chain forming a Schiff base and catalytic aspartate and glutamate side chains.[198]

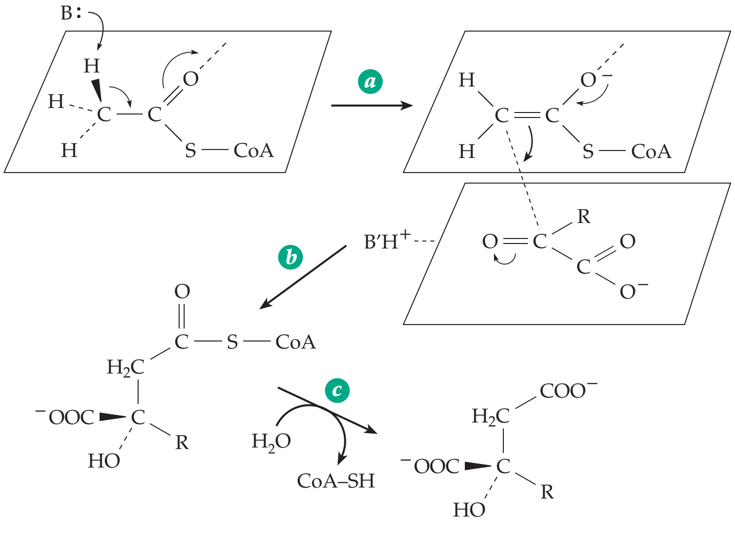
Polycarboxylic acid synthases. Several enzymes, including **citrate synthase,** the key enzyme which catalyzes the first step of the citric acid cycle, promote condensations of acetyl-CoA with ketones (Eq. 13-38). An α-oxo acid is most often the second substrate, and a thioester intermediate (Eq. 13-38) undergoes hydrolysis to release coenzyme A.[199] Because the substrate acetyl-CoA is a thioester, the reaction is often described as a Claisen condensation. The same enzyme that catalyzes the condensation of acetyl-CoA with a ketone also catalyzes the second step, the hydrolysis of the CoA thioester. These polycarboxylic acid synthases are important in biosynthesis. They carry out the initial steps in a *general chain elongation process* (Fig. 17-18). While one function of the thioester group in acetyl-CoA is to activate the methyl hydrogens toward the aldol condensation, the subsequent hydrolysis of the thioester linkage provides for overall irreversibility and "drives" the synthetic reaction.

TABLE 13-1
Products Arising from Reactions of Acetyl-CoA with a Second Substrate with Catalysis by a Polycarboxylate Synthase

Ketone substrate	Product	Further metabolites
 Glyoxylate	 L-Malate	Carbohydrates, etc., via glyoxylate pathway
 Oxaloacetate *Si*-Citrate synthase	 Citrate	2-Carbon unit from acetyl-CoA occupies *pro-S* position
 2-Oxoglutarate	 Homocitrate	Lysine via α-aminoadipic acid
 2-Oxoisovalerate	 α-Isopropylmalate	Leucine
 Acetoacetyl-CoA	 *S*-3-Hydroxy-3-methylglutaryl-CoA (HMG-CoA)	Free acetoacetate Isoprenoid compounds

The stereochemistry of the reaction is also illustrated in Eq. 13-38. These enzymes may be classified by designating the face of the carbonyl group to which the enolate anion adds. The *si* face is up in Eq. 13-38. The common citrate synthase of animal tissues[200] and that of *E. coli*[201] condense with the *si* face and are designated (*si*)-citrate synthases. A few anaerobic bacteria use citrate (*re*)-synthase having the opposite stereochemistry.[202] Many citrate synthases are ~100-kDa dimers[203,204] but some are hexamers[203,205] and are allosterically inhibited by NADH. The second substrates and products of several related reactions of acetyl-CoA are summarized in Table 13-1.

The *si*-citrate synthase of pigs is a dimer of 437-residue chains, each of which is organized into a large rigid domain and a smaller



Citrate if R = —CH$_2$—COO$^-$
L-Malate if R = H

(13-38)

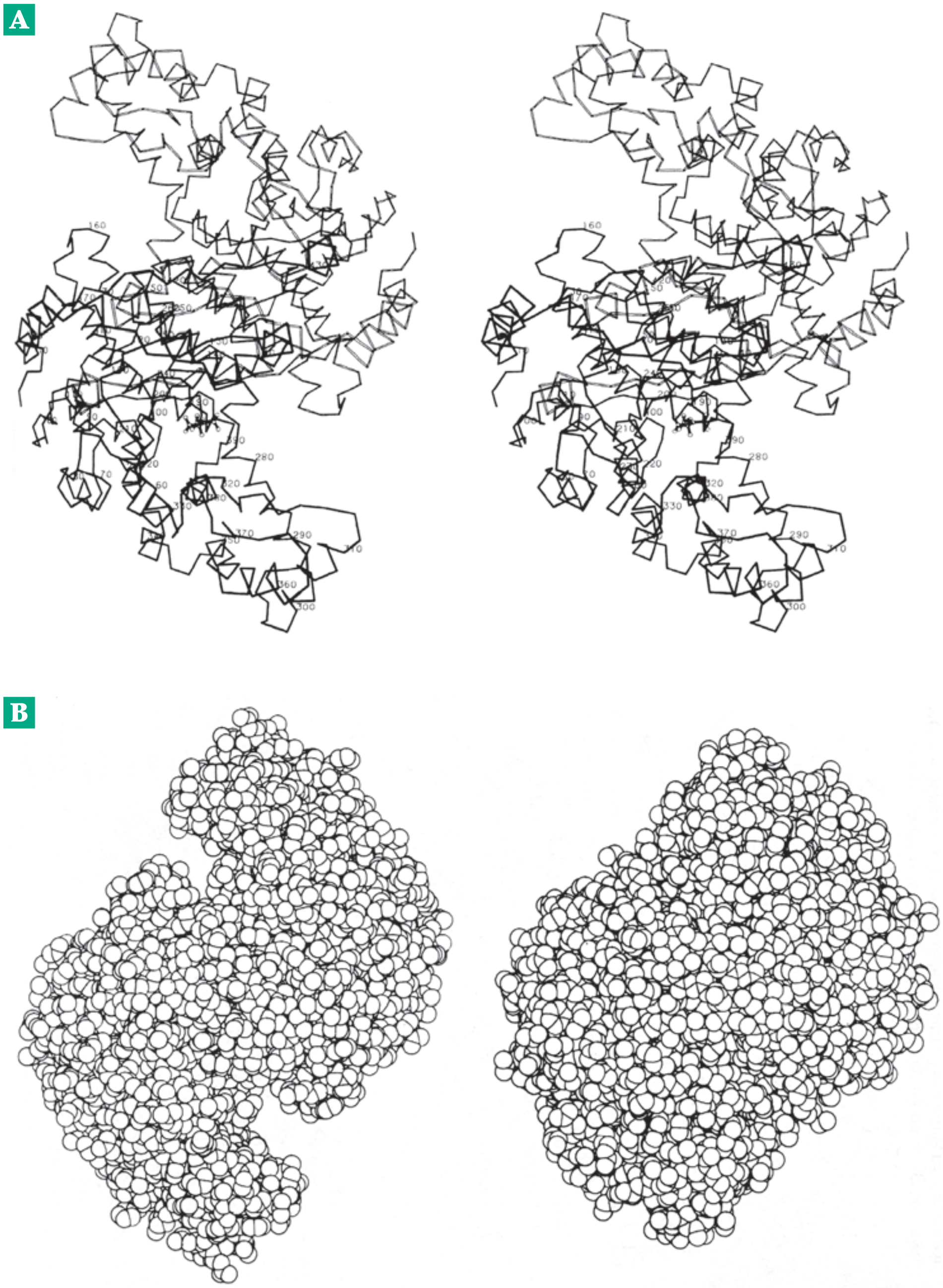
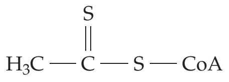


Figure 13-8 Three-dimensional structure of citrate synthase. (A) Stereoscopic alpha carbon trace of the dimeric pig enzyme in its open conformation.[209] A molecule of citrate is shown in the lower subunit. The view is down the two-fold axis. Courtesy of Robert Huber. (B) Space-filling representation of the open (left) and closed (right) forms of the same enzyme.[200] Courtesy of Stephen J. Remington.

flexible domain. The active site lies between the domains.[206] When oxaloacetate binds into the cleft the smaller domain undergoes a complex motion that closes the enzyme tightly around the substrate (Fig. 13-8).[200,207,208] The bound oxaloacetate is shown in Fig. 13-9. It is surrounded by a large number of polar side chains, including several from histidine and arginine residues. Of these, Arg 401 and Arg 421 of the second subunit bind the substrates' two carboxylate groups. In this tight complex the imidazole of His 320 is in the correct position to protonate the carbonyl oxygen of the oxaloacetate. The domain movement has also brought the groups that bind the acetyl-CoA into their proper position creating, by "induced fit," the acetyl-CoA binding site. This conformational change also accounts for the observation of an ordered kinetic mechanism with oxaloacetate binding before acetyl-CoA. The imidazole of His 274 is correctly oriented to abstract a proton from the methyl group of acetyl-CoA to generate the intermediate enolate anion. When oxaloacetate alone is bound into the active site the carbonyl stretching frequency, observed by infrared spectroscopy, is shifted from 1718 cm^{-1} for free oxaloacetate to 1697 cm^{-1}. This decrease of 21 cm^{-1} suggests a strong polarization of the C = O bond by its interaction with His 320 in the ground state.[211] This interaction is seen also in the ^{13}C NMR spectrum, the carbonyl resonance being shifted downfield by 6.8 ppm upon binding to the protein.[212] The enolate anion mechanism is also supported by kinetic isotope effects.[213]

The methyl protons of dithia-acetyl-CoA are much more reactive than those of acetyl-CoA (pK_a ~12.5 vs ~20). When citrate acts on this acetyl-CoA analog together with oxaloacetate, the expected enolate anion



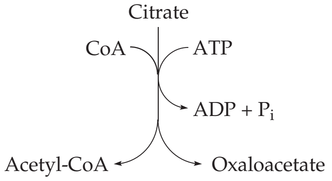
$$\mathrm{H_3C-\overset{\displaystyle S}{\overset{\|}{C}}-S-CoA}$$

Dithia-acetyl-CoA

forms rapidly as is indicated by the appearance of a 306 nm absorption band. However, it condenses very slowly with the oxaloacetate.[214] The binding equilibria, kinetics, and X-ray structures of complexes for a variety of other analogs of acetyl-CoA have also been studied.[210,215-216a] It appears that Asp 375 and His 274 work together to generate the enolate anion (Fig. 13-9). NMR measurements indicate formation of an unusually short hydrogen bond, but its significance is uncertain.[216a,216b] Malate synthase (Table 13-1) operates with a very similar mechanism but has an entirely different amino acid sequence and protein fold.[216c]

Other Claisen condensations are involved in synthesis of fatty acids and polyketides[217] (Chapter 21) and in formation of 3-hydroxy-3-methylglutaryl-CoA, the precursor to the polyprenyl family of compounds (Chapter 22). In these cases the acetyl group of acetyl-CoA is transferred by a simple displacement mechanism onto an –SH group at the active site of the synthase to form an acetyl-enzyme.[218,219] The acetyl-enzyme is the actual reactant in step *b* of Eq. 17-5 where this reaction, as well as that of HMG-CoA lyase, is illustrated.

Citrate cleaving enzymes. In eukaryotic organisms the synthesis of citrate takes place within the mitochondria, but under some circumstances citrate is exported into the cytoplasm. There it is cleaved by **ATP-citrate lyase**. To ensure that the reaction goes to completion, cleavage is coupled to the hydrolysis of ATP to ADP and inorganic phosphate (Eq. 13-39). The value of *G′* given here is extremely dependent upon the concentration of Mg^{2+} as a result of strong chelation of Mg^{2+} by citrate.[220] The reaction sequence is complex but can be understood in terms of an initial ATP-dependent synthesis of



$$\Delta G'\ (\text{pH } 7)^* = +0.9\ \text{kJ mol}^{-1} \qquad (13\text{-}39)$$

citryl-CoA using a mechanism similar to that in Eq. 12-47. There is evidence for both phosphoenzyme and citryl enzyme intermediates (Eq. 13-40).[221] Native ATP-citrate lyase is a tetramer of 110-kDa subunits. It usually contains some phosphoserine and phosphothreonine residues but they apparently have little effect on activity.[222] Phosphorylation is catalyzed by cAMP-dependent and by insulin-dependent protein kinases.[223,224] A related reaction is the ATP-dependent cleavage of malate to acetyl-CoA and glyoxylate. It requires two enzymes, malyl-CoA being an intermediate.[225]

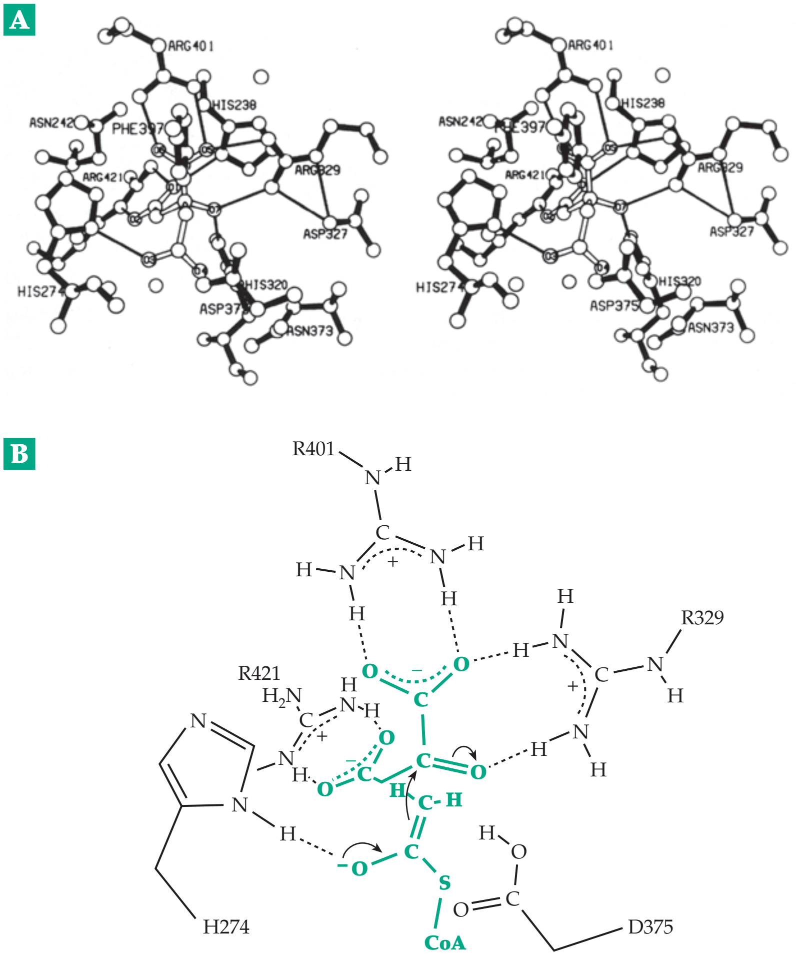
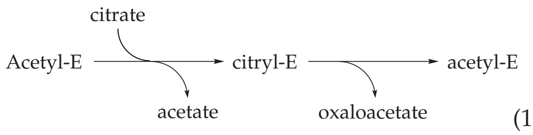


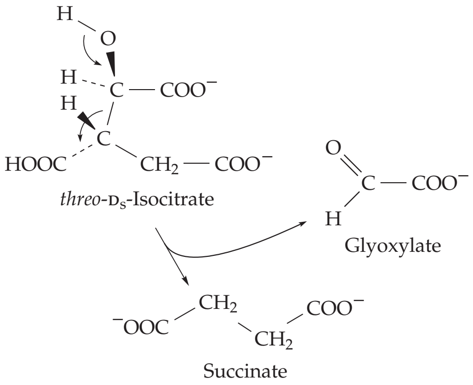
Figure 13-9 Active site of pig citrate synthase. (A) Stereoscopic view with a molecule of citrate in the active site.[200] Courtesy of Stephen J. Remington. (B) Interpretive view of the enolate anion of acetyl-CoA and oxaloacetate bound in the active site. Based on work by Kurz *et al.*[210]

A substrate-induced citrate lyase found in bacteria such as *E. coli* and *Klebsiella* promotes the anaerobic dissimilation of citrate splitting it to oxaloacetate and acetate.[226] The large ~585-kDa protein from *Klebsiella*[225] has the composition $(\alpha\beta\gamma)_6$, where the α, β, and γ subunits have masses of ~ 55-, 30-, and 10-kDa, respectively. The subunit carries an unusual covalently bound derivative of coenzyme A[227] (see also Chapter 14). The 10-kDa γ subunit serves as an **acyl-carrier protein**, the –SH of its prosthetic group being acetylated by a separate ATP-dependent ligase. The resulting acetyl-enzyme undergoes an acyl exchange (analogous to a CoA-transferase reaction; Eq. 12-50) to form a citryl enzyme before the aldol cleavage takes place (Eq. 13-40).



(13-40)

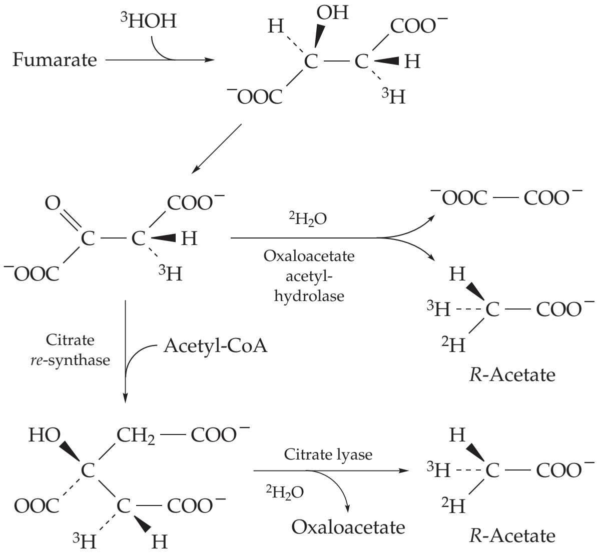
The first unique enzyme of the important glyoxylate pathway (Chapter 17), **isocitrate lyase**, cleaves isocitrate to succinate and glyoxylate (Eq. 13-41).[228] The carboxylate group that acts as electron acceptor would presumably be protonated by the enzyme.



ΔG' (pH 7) = 8.7 kJ mol^{-1} (13-41)

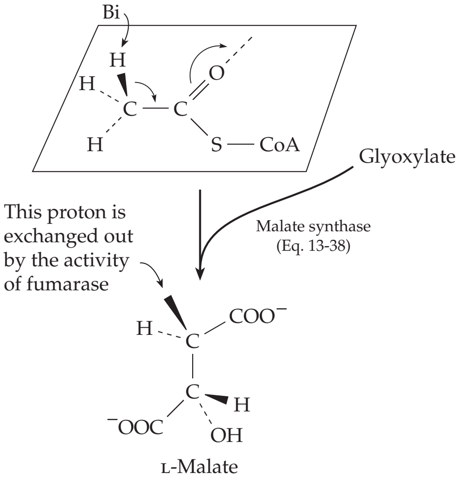
3. Chiral Acetates and Their Use in Stereochemical Studies

Consider the series of enzyme-catalyzed reactions shown in Eq. 13-42. Fumarate is hydrated in ^{3}H-containing water to malate which is oxidized to oxaloacetate. Hydrolysis of the latter with oxaloacetate acetyl-hydrolase (Eq. 13-34) in ^2H_2O gives oxalate and chiral (*R*) acetate. The identical product can be obtained by condensing oxaloacetate with acetyl-CoA using citrate (*re*)-synthetase. The resulting citrate is cleaved in ^2H_2O using a citrate lyase having the *si* specificity.[229,230] Acetate of the opposite chirality can be formed enzymatically beginning with [2,3-^{3}H]fumarate hydrated by fumarate hydratase in ordinary water. Chiral acetates have also been prepared nonenzymatically,[229] and their configuration has been established unequivocally.



(13-42)

During the action of both oxaloacetate acetyl-hydrolase and citrate lyase (Eq. 13-42) inversion of configuration occurs about the carbon atom that carries the negative charge in the departing enolate anion. That inversion also occurs during catalysis by citrate (*re*)-synthase and other related enzymes has been demonstrated through the use of chiral acetates.[229,230] The findings with malate synthase[199] are illustrated in Eq. 13-38. Presumably, a basic group B of the enzyme removes a proton to form the planar enolate anion. The second substrate glyoxylate approaches from the other side of the molecule and condenses as is shown. Since any one of the three protons in either *R* or *S* chiral acetyl-CoA might have been abstracted by base B, several possible combinations of isotopes are possible in the L-malate formed. One of the results of the experiment using chiral (*R*) acetyl-CoA is illustrated in Eq. 13-43. The reader can easily tabulate the results of removal of the ^{2}H or ^{3}H. However, notice that if the base –B: removes ^{2}H (D) or ^{3}H (T) the reaction will be much slower because of the kinetic isotope effects which are expected to be $^{H}k/^{D}k \approx 7$ and $^{H}k/^{T}k \approx 16$. A second important fact is that the *pro-R* hydrogen at C-3 in malate is specifically exchanged out into water by the action of fumarate hydratase. From the distribution of tritium in the malate and fumarate formed using the two chiral acetates, the inversion by malate synthase was established. See Kyte[231] for a detailed discussion.

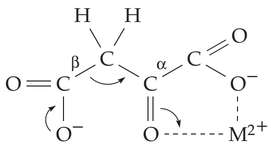

Bi
H
H
H
C
C
O
S — CoA
Glyoxylate
This proton is exchanged out by the activity of fumarase
Malate synthase (Eq. 13-38)
H
COO⁻
C
C
H
⁻OOC
OH
L-Malate

(13-43)

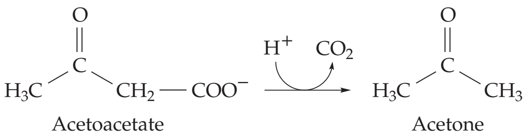
4. Addition of an Enolate Ion to Carbon Dioxide and Decarboxylation

The addition of an enolate anion to CO_2 to form a β-oxoacid represents one of the commonest means of incorporation of CO_2 into organic compounds. The reverse reaction of **decarboxylation** is a major mechanism of biochemical formation of CO_2. The equilibrium constants usually favor decarboxylation but the cleavage of ATP can be coupled to drive carboxylation when it is needed, e.g., in photosynthesis.

Decarboxylation of β-oxoacids. Beta-oxoacids such as oxaloacetic acid and acetoacetic acid are unstable, their decarboxylation being catalyzed by amines, metal ions, and other substances. Catalysis by amines depends upon Schiff base formation,[232] while metal ions form chelates in which the metal assists in electron withdrawal to form an enolate anion.[233–235]



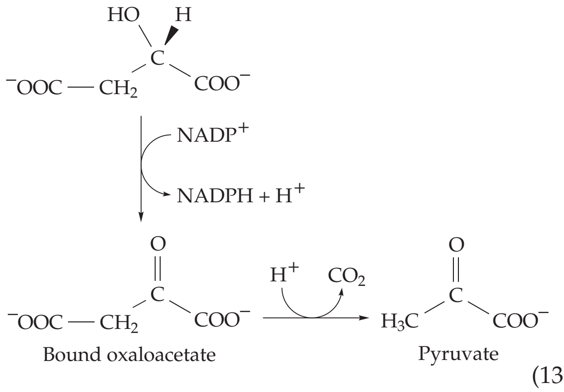
Can we apply any of this information from nonenzymatic catalysis to decarboxylating enzymes? Some decarboxylases do form Schiff bases with their substrates, and some are dependent on metal ions.[235] The acetone-forming fermentation of *Clostridium acetobutylicum* requires large amounts of acetoacetate decarboxylase (Eq. 13-44).



(13-44)

The enzyme is inactivated by borohydride in the presence of substrate, and acid hydrolysis of the inactivated enzyme yielded ε-*N*-isopropyllysine. Decarboxylation occurs from a Schiff base by a mechanism analogous to that of the aldol cleavage shown in Eq. 13-36.[236] Mechanistically related is 4-oxalocrotonate decarboxylase.[236a]

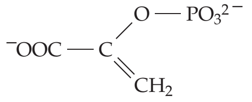
Linked oxidation and decarboxylation. Metabolic pathways often make use of oxidation of a β-hydroxy acid to a β-oxoacid followed by decarboxylation in the active site of the same enzyme. An example is conversion of L-malate to pyruvate (Eq. 13-45). The Mg^{2+} or Mn^{2+}-dependent decarboxylating malic dehydrogenase that catalyzes the reaction is usually called **the malic enzyme**. It is found in most organisms.[237–240] While a concerted decarboxylation and dehydrogenation may sometimes occur,[241–242] the enzymes of this group appear usually to operate with bound oxoacid intermediates as in Eq. 13-45.



(13-45)

Other reactions of this type are the oxidation of isocitrate to 2-oxoglutarate in the citric acid cycle (Fig. 17-4, steps *d* and *e*),[243] oxidation of 6-phosphogluconate to ribulose 5-phosphate (Eq. 17-12),[244] and corresponding reactions of isopropylmalate dehydrogenase[245,246] and tartrate dehydrogenase.[247,248] Crystallographic studies of isocitrate dehydrogenase using both photolabile "caged" isocitrate[247] and slow mutant forms[243] with polychromatic Laue crystallography (Chapter 3) have demonstrated the rapid formation of the anticipated intermediate **oxalosuccinate**.

Phosphoenolpyruvate, a key metabolic intermediate. A compound of central importance in metabolism is the phosphate ester of the enol form of pyruvate, commonly known simply as phosphoenolpyruvate (PEP).[249] It is formed in the glycolysis pathway by dehydration of 2-phosphoglycerate (Eq. 13-15) or by decarboxylation of oxaloacetate. Serving as a preformed enol from which a reactive enolate anion can be released for condensation reactions,[250,251] PEP is utilized in metabolism in many ways.

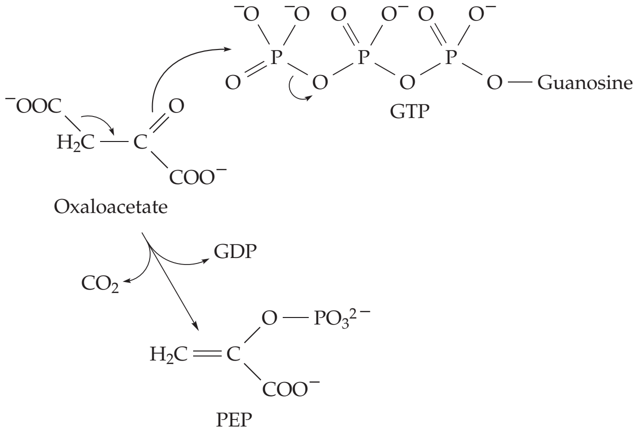


Phosphoenolpyruvate (PEP)

In animals and in many bacteria, PEP is formed by decarboxylation of oxaloacetate. In this reaction, which is catalyzed by **PEP carboxykinase** (PEPCK), a molecule of GTP, ATP, or inosine triphosphate captures and phosphorylates the enolate anion generated by the decarboxylation (Eq. 13-46).[252] The stereochemistry is such that CO_2 departs from the *si* face of the forming enol.[253] The phospho group is transferred from GTP with inversion at the phosphorus atom.[254] The enzyme requires a divalent metal ion, preferably Mn^{2+}. In fact, kinetic studies of the GTP-dependent avian mitochondrial enzyme indicate two metal-binding sites, one on the polyphosphate group of the bound GTP and one on carboxylate side chains of the protein.[252,255] The three-dimensional structure of the ATP-dependent *E. coli* enzyme reveals a nucleotide binding site similar to the ATP site of adenylate kinase (Fig. 12-30).[256] A definite binding site for CO_2 is also present in the enzyme.[257]

PEPCK is also activated by low concentrations of Fe^{2+} and this activation depends upon a protein that has been identified as glutathione peroxidase (Eq. 15-58). By destroying H_2O_2 the latter may allow the Fe^{2+} to prevent oxidation of an SH group on PEP carboxykinase.[258] Synthesis of PEPCK is stimulated by glucagon, evidently through a direct action of cAMP on transcription of the structural gene.[259] Transcription is also stimulated by glucocorticoids and thyroid hormone and is inhibited by insulin.

In some organisms, such as the parasitic *Ascaris suum*, PEPCK functions principally as a means of synthesis of oxaloacetate by reaction of PEP with CO_2 and GDP. These organisms lack pyruvate kinase, which allows for buildup of the high PEP concentration needed to drive this reaction.[260] In a similar way PEP can be converted to oxaloacetate by **PEP carboxytransphosphorylase**, an enzyme found only in propionic acid bacteria and in *Entamoeba*. The reaction (Eq. 13-47) is accompanied by synthesis of inorganic pyrophosphate which may be cleaved to "pull" the reaction in the indicated direction.

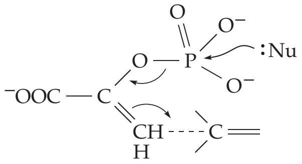


re Face is toward reader (13-46)

$$PEP + CO_2 + P_i \rightarrow PP_i + \text{oxaloacetate} \quad (13\text{-}47)$$

Oxaloacetate is also decarboxylated without phosphorylation of the enolate anion formed but with release of free pyruvate. Both pyruvate kinase and PEPCK can act as oxaloacetate decarboxylases.[261]

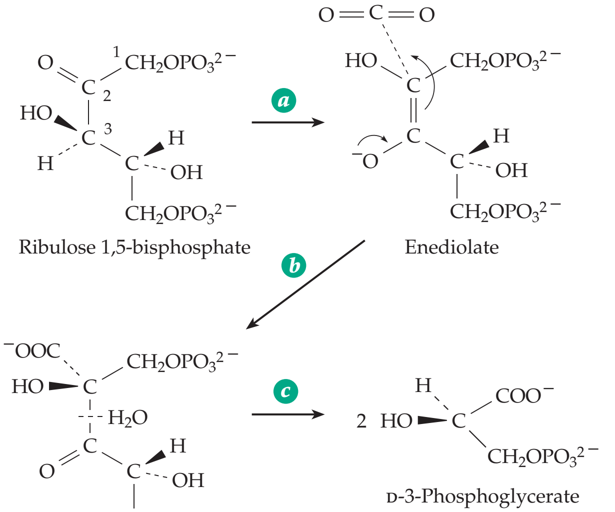
In the important reactions discussed in the following sections enolate ions are intermediates in carbon-carbon bond formation. Other examples are given in Eqs. 20-7 and 20-8, and Fig. 25-1, in which C–C bonds are formed by action of PEP as a carbon nucleophile, e.g., in aldol-like condensations.[242a] Yet another unusual type of reaction, involving formation of an enol ether linkage is illustrated in Box 13-B and in Eq. 25-3.



Ribulose bisphosphate carboxylase. The major route of incorporation of CO_2 into organic compounds is via photosynthesis. When ^{14}C-labeled CO_2 enters chloroplasts of green plants, the first organic ^{14}C-containing compound detected is 3-phosphoglycerate. Two molecules of this compound are formed through the action of **ribulose-1,5-bisphosphate carboxylase** (abbreviated **rubisco**), an enzyme present in chloroplasts and making up 16% of the protein of spinach leaves. This enzyme, whose structure is illustrated in Fig. 13-10, is thought to be the most abundant protein on earth. Because O_2 competes with CO_2 as a substrate, the enzyme also catalyzes an "oxygenase" reaction. It is therefore often called **ribulose-1,5-bisphosphate carboxylase/ oxygenase**.[262] The rubisco from most plants is a 500- to 560-kDa L_8S_8 oligomer as shown in Fig. 13-10. The large subunit is encoded in the chloroplast DNA. However, a family of nuclear genes encode the small subunits, which are synthesized as larger precursors, with N-terminal extensions being removed to give the mature subunits.[263,264] Two types of small subunits occupy different positions in the quasi-symmetric spinach rubisco.[264] In *Euglena* the small subunit is synthesized as a polyprotein precursor containing eight copies of the subunit.[265] Rubisco from the hydrogen-oxidizing bacterium *Alcaligenes eutrophus* has a similar quaternary structure,[266] but the enzyme from *Rhodospirillum rubrum* is a simple dimer.[267] In dinoflagellates the rubisco gene is present in nuclear DNA rather than in the chloroplasts.[268]

The carboxylation reaction catalyzed by rubisco differs from others that we have considered in that the carboxylated product is split by the same enzyme (Eq. 13-48). The mechanism shown in Eq. 13-48 was suggested by Bassham *et al*.[271] Ribulose bisphosphate, for which the enzyme is absolutely specific, is first converted to its enolic form, a 2,3-enediol. Loss of a proton from the 3-OH group forms the endiolate anion needed for the carboxylation in step *b*. The product of this step is a β-oxoacid which undergoes enzyme-catalyzed hydrolytic cleavage (Eq. 13-48, step *c*; see also Eq. 13-34). Support for this mechanism came from the observation that 2-carboxyarabinitol bisphosphate (Fig. 13-11B) is a potent inhibitor, possibly a transition state analog.[272] This inhibitor, bound into the active site of the enzyme from spinach, is seen in both Figs. 13-10B and 13-11A. An expanded version of Eq. 13-48 is given in Fig. 13-12, which is based upon modeling together with a variety of X-ray structures including that shown in Fig. 13-11.

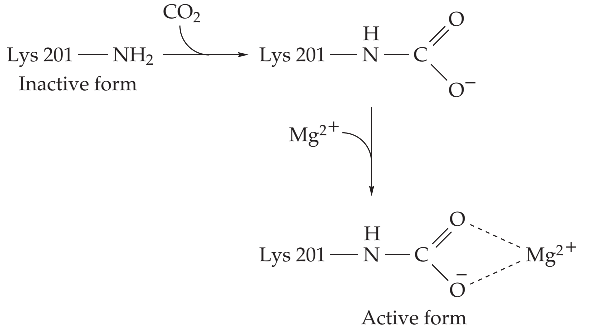
An essential Mg^{2+} ion is held by carboxylates of D203 and E204 and by modified K201. It also coordinates three molecules of H_2O in the free enzyme.[272a] Catalytic roles for the H294 and H327 imidazole groups are still being elucidated.[272b] Lysine 175 protonates C2 of the aci anion generated by C–C bond cleavage (step *e*, Fig. 13-12).[272c,276] Like many enzymes, rubisco exists in two major conformational states: open and closed.[272d]



(13-48)

Chemical studies also support the indicated mechanism. For example, the β-oxoacid intermediate formed in step *b* of Eq. 13-48 or Fig. 13-12 has been identified as a product released from the enzyme by acid denaturation during steady-state turnover.[273,274] Isotopic exchange with ^{3}H in the solvent[275] and measurement of ^{13}C isotope effects[277] have provided additional verification of the mechanism. The catalytic activity of the enzyme is determined by ionizable groups with pK_a values of 7.1 and 8.3 in the ES complex.[278]

The apparent value of K_m for total CO_2 (CO_2 + HCO_3^-) is high, 11–30 mM, but for the true substrate CO_2 it is only 0.45 mM. In intact chloroplasts the affinity for substrate is distinctly higher, with the K_m for total CO_2 dropping to ~0.6 mM. The difference appears to result largely from a regulatory reaction of CO_2 in which the side chain amino group of lysine 201 of the large subunit forms a carbamate (Eq. 13-49). Although carbamylation is spontanous, it is enhanced by an ATP-dependent process catalyzed by **rubisco activase.**[279] The carbamylation converts the side chain of Lys 201 into a negatively charged group that binds to an essential divalent metal, usually Mg^{2+}, in the active center[282] as is shown in Fig. 13-11, A and B. The nature of the reaction is uncertain. One possibility is that the activase is a chaperonin.[280] It appears to assist the enzyme in removing inhibitory sugars that arise by side reactions in the active site.[281] Rubisco is also regulated by the level of a natural inhibitor which has been identified as 2-carboxyarabinitol 1-phosphate. This is the same as the inhibitor shown in Fig. 13-11 but with one less phosphate group and consequent weaker binding.[273a]



(13-49)

In most plants photosynthesis is also strongly inhibited by O_2. This observation led to the discovery that O_2 competes directly for CO_2 at the active site of rubisco in a process called **photorespiration**. Chloroplasts inhibited by oxygen produce **glycolate** in large amounts[282a] as a result of the reaction of the intermediate enediolate ion formed in step *b* of Eq. 13-48 with O_2

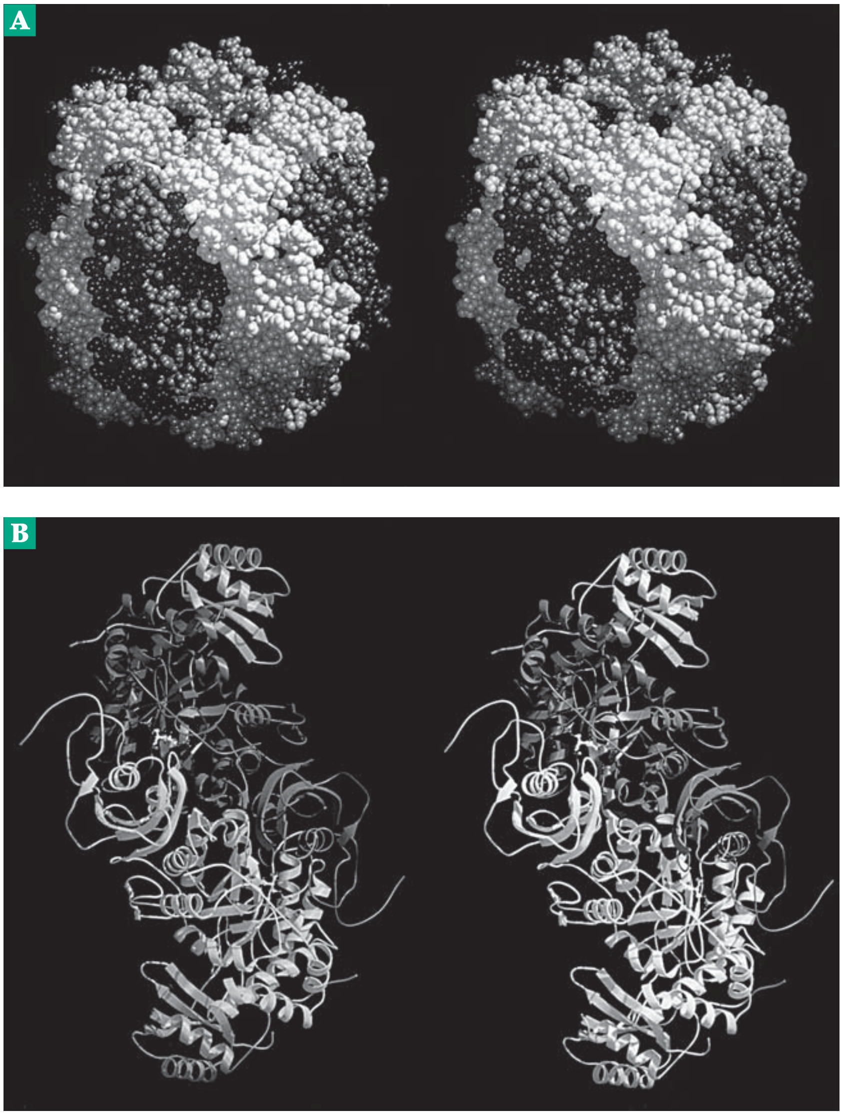


Figure 13-10 Stereoscopic view of ribulose bisphosphate carboxylase (rubisco) from spinach. (A) The symmetric L_8S_8 molecule contains eight 475-residue large subunits (in two shades of gray) and eight 123-residue small subunits (lighter gray in upper half of image). (B) One L_2S_2 substructure containing two active sites shared between adjacent large subunits with the bound inhibitor 2-carboxy-D-arabinitol 1,5-bisphosphate.[269] In the upper LS unit the S subunit (top) is light and the L subunit is dark. Courtesy of Inger Andersson. Similar structures have been determined for enzymes from tobacco[269a] and from the cyanobacterium *Synechoccus*.[270]

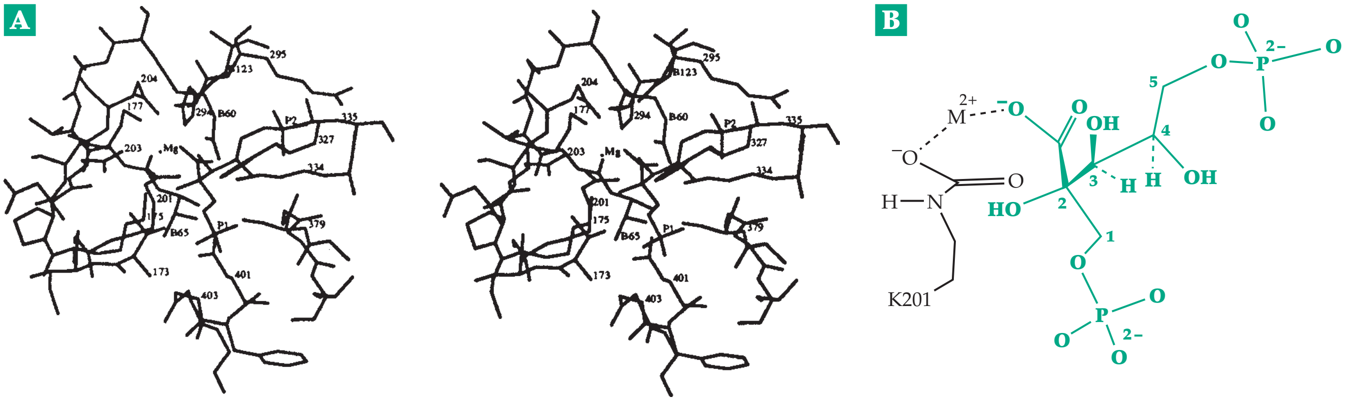


Figure 13-11 (A) Overview of the active site of spinach rubisco showing bound 2-carboxy-D-arabinitol 1,5-bisphosphate and Mg^{2+} and residues within hydrogen-bonding distance of these ligands. The hydroxyl groups at C2 and C3 of the inhibitor are in *cis* conformation. [269] Courtesy of Inger Andersson. (B) Structure of the inhibitor 2-carboxy-D-arabinitol 1,5-bisphosphate. A part of the carbamylated lysine 201 and the essential metal ion are also shown.

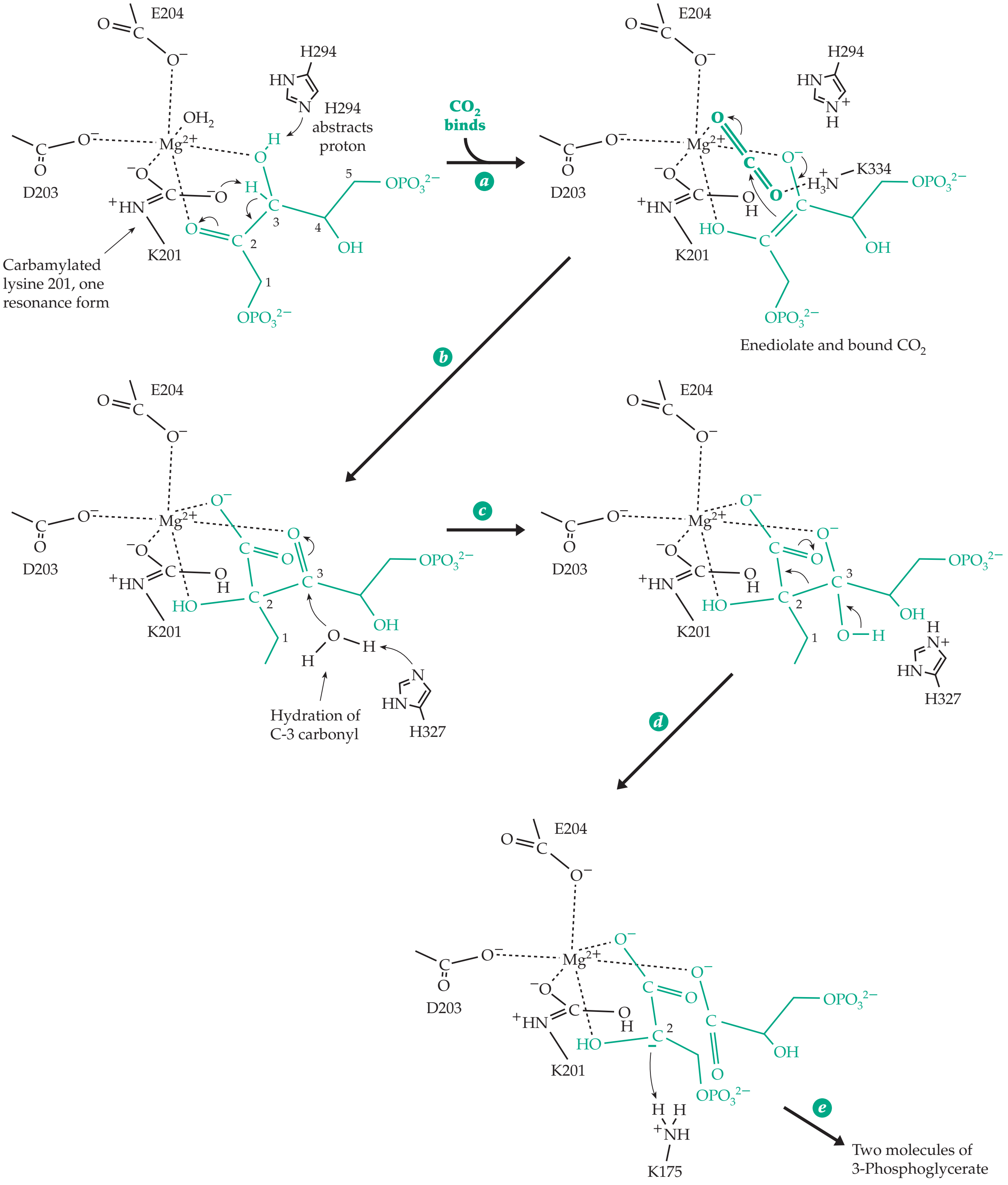
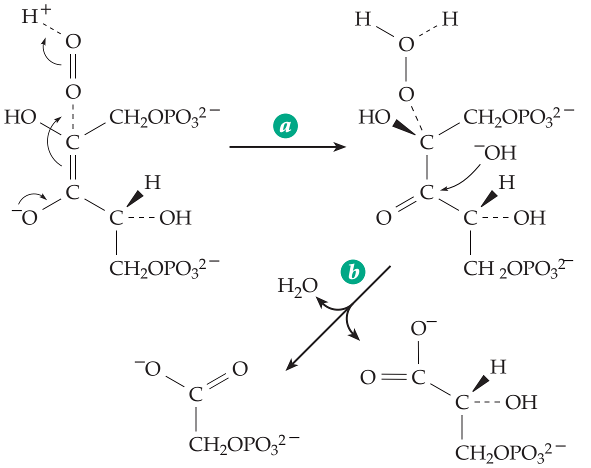

E204
H294
H294 abstracts proton
Mg2+
OH2
D203
K201
Carbamylated lysine 201, one resonance form
CO2 binds
a
K334
Enediolate and bound CO2
b
c
d
e
Hydration of C-3 carbonyl
H327
K175
Two molecules of 3-Phosphoglycerate

Figure 13-12 Proposed mechanism of action of ribulose bisphosphate carboxylase (rubisco). This is an abbreviated version of the mechanism as presented by Taylor and Andersson.[276] The binding of ribulose 1,5-bisphosphate occurs after carbamylation of lysine 201 and binding of a magnesium ion. Formation of an enediolate intermediate in step *a* is probably catalyzed by the carbamate group as indicated. Removal of a proton from the 3-OH, perhaps by H294, and addition of a proton to form an –OH group at C2 are also necessary. The CO_2 may bind to the Mg^{2+} and be polarized by interaction with other side chains prior to reaction. Carboxylation occurs by addition of the enediolate to CO_2 in step *b*. The hydration of the resulting 3-oxo group (step *c*) is necessary for cleavage of the C–C bond in step *d*. The participation of the newly formed carboxylate as an *aci* anion coordinated to Mg^{2+} is presumably involved. Protonation, with stereochemical inversion, is thought to involve K175, as shown. The two product molecules dissociate in step *e*.

(Eq. 13-50, step *a*). The peroxide formed in this way breaks up under the hydrolytic action of the enzyme to form phosphoglycolate and 3-phosphoglycerate (Eq. 13-50, step *b*).



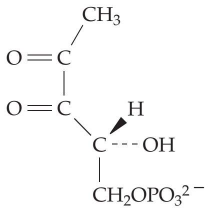
Phosphoglycolate 3-Phosphoglycerate

(13-50)

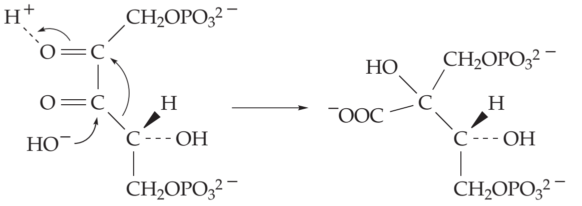
Molecular oxygen usually reacts rapidly with only those organic substrates, such as dihydroflavins, that are able to form stable free radicals. However, the endiolate anion of Eq. 13-50 may be able to donate a single electron to O_2 to form a superoxide–organic radical pair prior to formation of the peroxide (see also Eq. 15-30). Similar oxygenase side reactions have been observed for a variety of other enzymes that utilize carbanion mechanisms.[283] The reaction of rubisco with O_2 is of both theoretical and practical interest, the latter because of its significance in lowering the yield in photosynthesis (Chapter 23).

The simpler dimeric rubisco from *Rhodospirillum* is very inefficient in carboxylation and catalyzes much more oxygenation than do rubiscos of higher plants.[283a] Mutant enzymes that have impaired carboxylase and enhanced oxygenase are also known.[284,284a]

The small subunits of rubisco may help suppress undesirable side reactions.[285] For example, the following deoxypentodiulose phosphate can be formed by β elimination from the second intermediate of Eq. 13-48.



Some side products of the normal oxgenase reaction arise by elimination of a peroxide ion from the first intermediate of Eq. 13-50. The resulting dicarbonyl bisphosphate can rearrange to give a carboxytetritol bisphosphate (Eq. 13-51).[284]



Deoxypentodiulose-bisphosphate 2-Carboxytetritol-1,4-bisphosphate

(13-51)

Carbon dioxide or bicarbonate ion? An important question in the consideration of carboxylation and decarboxylation reactions is whether the reactant or the product is CO_2 or HCO_3^-. An approach to answering the question was suggested by Krebs and Roughton,[286]

$$CO_2 + H_2O \xrightarrow{a} H_2CO_3 \xrightarrow{b} H^+ + HCO_3^- \qquad (13\text{-}52)$$

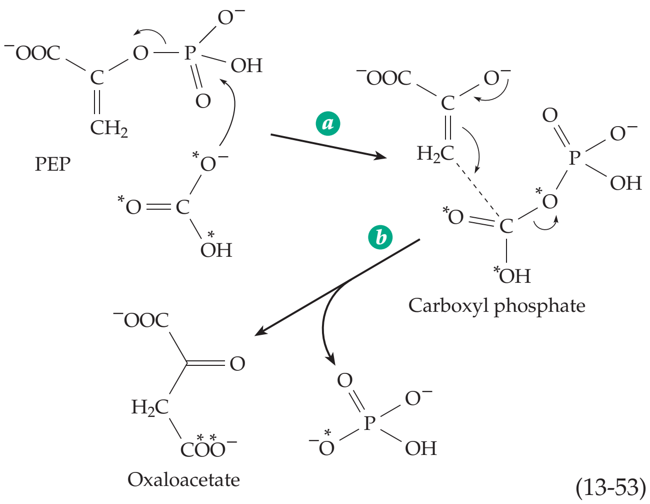
who pointed out that the attainment of equilibrium between free CO_2 and HCO_3^- (Eq. 13-52) may require several seconds. If an enzyme produces CO_2 as a product and progress of the reaction is followed manometrically, the pressure will rise higher than the equilibrium value as the CO_2 is evolved. Later when substrate is exhausted and the CO_2 equilibrates with bicarbonate, the pressure will fall again. The addition of carbonic anhydrase, which catalyzes Eq. 13-52, step *a*, abolishes the "overshoot." If bicarbonate is the primary product of a decarboxylation, there is a lag in the appearance of free CO_2.

A second approach is to use [^{18}O]bicarbonate and to follow the incorporation of ^{18}O into a carboxylated substrate. If CO_2 is the primary substrate only two labeled oxygen atoms enter the compound, whereas if HCO_3^- is the reactant three are incorporated.[287] A third technique is measurement of the rate of incorporation of CO_2 or bicarbonate in the carboxylated product. Over a short interval of time, e.g., 1 min, different kinetics will be observed for the incorporation of CO_2 and of bicarbonate.[288] Using these methods, it was established that the product formed in Eq. 13-46 and the reactant in Eq. 13-47 is CO_2. However, the carboxylation enzymes considered in the next section use bicarbonate as the substrate.

5. Incorporation of Bicarbonate into Carboxyl Groups

An important enzyme with a biosynthetic function in many bacteria and in all higher plants is **PEP carboxylase**,[289] which catalyzes the reaction of Eq. 13-53. This enzyme, in effect, accomplishes the reverse of Eq. 13-46 by converting the three-carbon PEP, by reaction with bicarbonate, into the four-carbon oxaloacetate. The latter is needed for "priming" of the citric acid cycle and for biosynthesis of such amino acids as aspartate and glutamate. That the enzyme functions in this way is indicated by the fact that mutants of *Salmonella* defective in the enzyme do not grow unless oxaloacetate or some other intermediate in the citric acid cycle is added to the medium. The enzyme from *S. typhimurium* is a 400-kDa tetramer with complex regulatory properties. The corresponding enzyme from spinach has 12 subunits and 12 bound Mn^{2+} ions. The enzyme also has a special function in the C_4 plants,[290] in which it is a component of a carbon dioxide concentrating system (Chapter 23).

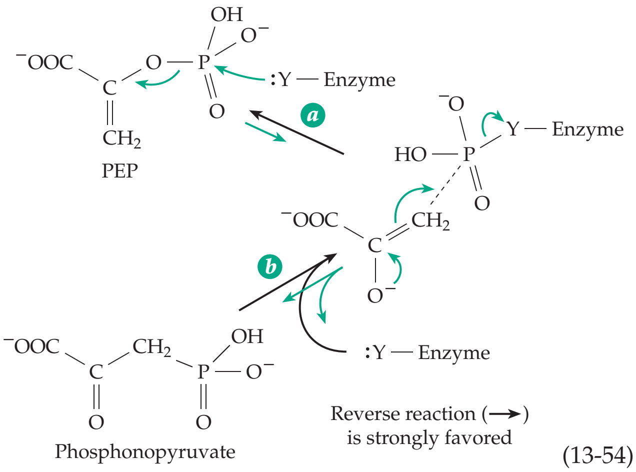
When [^{18}O]bicarbonate is a substrate, two labeled oxygen atoms enter the oxaloacetate, while the third appears in P_i. A concerted, cyclic mechanism could explain these results. However, study of kinetic isotope effects,[291] use of a substrate with a chiral thiophospho group,[292] and additional ^{18}O exchange studies[293] have ruled out this possibility. A transient carboxyl phosphate (Eq. 13-53) is evidently an intermediate.[294,295] The incorporation of the ^{18}O from bicarbonate into phosphate is indicated by the asterisks. The carboxyl group enters on the *si* face of PEP. However, there is another possibility.[295,296] The carboxyl phosphate, while in the active site adjacent to the enolate anion, may eliminate phosphate, the enolate ion adding to the resulting CO_2 to form the final product. According to this mechanism the group transfer potential of the phosopho group in PEP is utilized to concentrate the CO_2 and to localize it next to the enolate anion generated by the same process.



(13-53)

PEP carboxylase is lacking from animal tissues and fungi. In these creatures PEP is converted to pyruvate, which is then carboxylated to oxaloacetate with coupled cleavage of ATP by the action of **pyruvate carboxylase** (Eq. 14-3), an enzyme that not only utilizes bicarbonate ion but also contains **biotin**. However, there are mechanistic similarities between its action and that of PEP carboxylase.

PEP mutase and the synthesis of phosphonates. The lipids of some organisms, such as *Tetrahymena*, contain aminoethylphosphonate, a compound with a C–P bond (Chapter 8). There are also many other naturally occurring phosphono compounds and huge quantities of synthetic phosphonates, present in detergents, herbicides, and insecticides, are metabolized by bacteria.[297] Here we will consider only one step in the biosynthesis of phosphonates, the conversion of PEP into phosphonopyruvate (Eq. 13-54), a reaction catalyzed by **PEP mutase**. The phospho group is moved from the oxygen atom of PEP to the methylene carbon atom. When attached to carbon the phospho group is designated **phosphono**. The reaction occurs with retention of the configuration around phosphorus and a phosphoenzyme intermediate seems to be involved[298,299] as shown in Eq. 13-54. The equilibrium in this reaction strongly favors PEP. One further product formed from phosphonopyruvate is phosphonoethylamine, a component of phosphonolipids.[300] It is easy to imagine a synthetic route involving transamination followed by decarboxylation. Another mutase, similar to PEP mutase, shifts the carboxyphospho group $^-OOC{-}PO_2^-{-}$ exactly as in Eq. 13-54 as part of the biosynthetic pathway of a natural herbicide formed by *Streptomyces hygroscopicus*.[301,302] Soil bacteria contain specialized enzymes that catalyze the hydrolytic cleavage of P–C bonds.[302a]



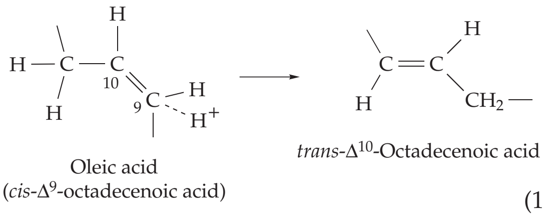
Reverse reaction (⟶) is strongly favored

(13-54)

D. Some Isomerization and Rearrangement Reactions

A few metabolic reactions do not fit into any of the categories discussed so far and apparently do not depend upon a coenzyme. Some of these involve transfer of alkyl groups or of hydrogen atoms from one carbon to another. The hydrogen atoms move by direct transfer without exchange with the medium. All of the reactions could involve carbocations but there is often more than one mechanistic possibility.

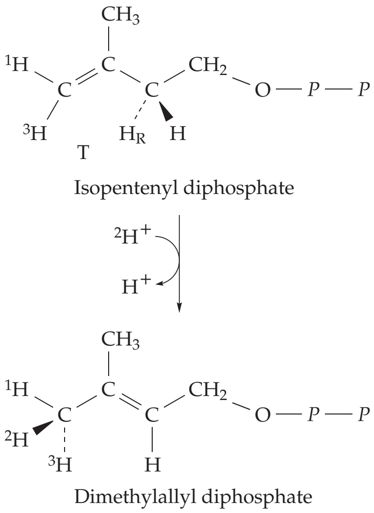
A simple 1,3-proton shift is shown in Table 10-1 as reaction type 6A. An example is the isomerization of oleic acid to *trans*-Δ^{10}-octadecenoic acid (Eq. 13-55) catalyzed by a soluble enzyme from a pseudomonad.[303]



H
H — C — C
10
9 C — H
H⁺
H
C = C
H
CH₂ —
Oleic acid
(*cis*-Δ⁹-octadecenoic acid)
trans-Δ¹⁰-Octadecenoic acid

(13-55)

A second example is isomerization of **isopentenyl diphosphate** to **dimethylallyl diphosphate** (Eq. 13-56).[304–307] The stereochemistry has been investigated using the ^{3}H-labeled compound shown in Eq. 13-56. The *pro-R* proton is lost from C-2 and a proton is added to the *re* face at C-4. When the reaction was carried out in ^2H_2O a chiral methyl group was produced as shown.[304] A concerted proton addition and abstraction is also possible, the observed *trans* stereochemistry being expected for such a mechanism. However, the fact that the enzyme is strongly inhibited by the cation of 2-(diethylamino) ethyl pyrophosphate, an analog of a probable carbocation intermediate. Cysteine and glutamate side chains are essential.[307]

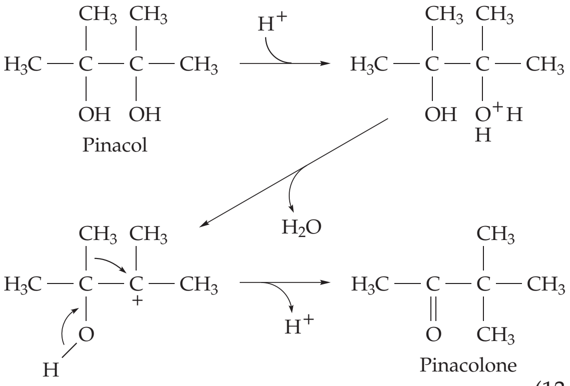


CH₃
¹H
C
C
CH₂
O — P — P
³H
H_R
H
T
Isopentenyl diphosphate
²H⁺
H⁺
CH₃
¹H
C
C
CH₂
O — P — P
²H
³H
H
Dimethylallyl diphosphate

(13-56)

Reaction type 6B of Table 10-1 is allylic rearrangement with simultaneous condensation with another molecule. The reaction, which is catalyzed by **prenyltransferases**,[307a] occurs during the polymerization of polyprenyl compounds (Fig. 22-1,Eqs. 22-2, 22-3). Experimental evidence favors a carbocation mechanism for all of these reactions.[308,309] See Chapter 22.

Reaction type 6C (Table 10-1) occurs during the biosynthesis of leucine and valine (Fig. 24-17). The rearrangement is often compared with the nonenzymatic acid-catalyzed pinacol–pinacolone rearrangement in which a similar shift of an alkyl group takes place (Eq. 13-57). The enzyme-catalyzed rearrangement presumably gives the structure drawn in brackets in Table 10-1. The same enzyme always catalyzes reduction with NADH to the diol, the Mg^{2+}-dependent enzyme being called **acetohydroxy acid isomeroreductase**. Rearrangement has never been observed without the accompanying reduction.[310–313] More complex rearrangements that occur during biosynthesis of sterols are described in Chapter 22.



CH₃ CH₃
H₃C — C — C — CH₃
OH OH
Pinacol
H⁺
CH₃ CH₃
H₃C — C — C — CH₃
OH O⁺H
H
H₂O
CH₃ CH₃
H₃C — C — C — CH₃
+
O
H
H⁺
CH₃
H₃C — C — C — CH₃
O CH₃
Pinacolone

(13-57)

References

1. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology*, Dover, Mineola, N. Y. (p. 465)

2. Sorensen, P. E., and Jencks, W. P. (1987) *J. Am. Chem. Soc.* **109**, 4675 – 4690

3. Kallen, R. G., and Jencks, W. P. (1966) *J. Biol. Chem.* **241**, 5845 – 5850, 5851 – 5863

4. Edsall, J. T., and Wyman, J. (1958) *Biophysical Chemistry*, Vol. I, Academic Press, New York (p. 550ff)

5. Fernley, R. T. (1988) *Trends Biochem. Sci.* **13**, 356 – 359

6. Sly, W. S., and Hu, P. Y. (1995) *Ann. Rev. Biochem.* **64**, 375 – 401

7. Hazen, S. A., Waheed, A., Sly, W. S., LaNoue, K. F., and Lynch, C. J. (1996) *FASEB J.* **10**, 481 – 490

8. Heck, R. W., Boriack-Sjodin, P. A., Qian, M., Tu, C., Christianson, D. W., Laipis, P. J., and Silverman, D. N. (1996) *Biochemistry* **35**, 11605 – 11611

9. Maren, T. H. (1985) *N. Engl. J. Med.* **313**, 179 – 181

10. Stams, T., Nair, S. K., Okuyama, T., Waheed, A., Sly, W. S., and Christianson, D. W. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13589 – 13594

11. Guilloton, M. B., Korte, J. J., Lamblin, A. F., Fuchs, J. A., and Anderson, P. M. (1992) *J. Biol. Chem.* **267**, 3731 – 3734

12. Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G., and Rees, D. C. (1996) *EMBO J.* **15**, 2323 – 2330

13. Rowlett, R. S., Chance, M. R., Wirt, M. D., Sidelinger, D. E., Royal, J. R., Woodroffe, M., Wang, Y.-F. A., Saha, R. P., and Lam, M. G. (1994) *Biochemistry* **33**, 13967 – 13976

13a. Kimber, M. S., and Pai, E. F. (2000) *EMBO J.* **19**, 1407 – 418

14. Miyamoto, H., Miyashita, T., Okushima, M., Nakano, S., Morita, T., and Matsushiro, A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9657 – 9660

15. Håkansson, K., Carlsson, M., Svensson, L. A., and Liljas, A. (1992) *J. Mol. Biol.* **227**, 1192 – 1204

16. Kumar, V., and Kannan, K. K. (1994) *J. Mol. Biol.* **241**, 226 – 232

17. Kiefer, L. L., Paterno, S. A., and Fierke, C. A. (1995) *J. Am. Chem. Soc.* **117**, 6831 – 6837

17a. Lesburg, C. A., Huang, C.-c, Christianson, D. W., and Fierke, C. A. (1997) *Biochemistry* **36**, 15780 – 15791

18. Liang, J.-Y., and Lipscomb, W. N. (1987) *Biochemistry* **26**, 5293 – 5301

19. Pocker, Y., and Janjic, N. (1989) *J. Am. Chem. Soc.* **111**, 731 – 733

19a. Earnhardt, J. N., Qian, M., Tu, C., Lakkis, M. M., Bergenhem, N. C. H., Laipis, P. J., Tashian, R. E., and Silverman, D. N. (1998) *Biochemistry* **37**, 10837 – 10845

19b. Denisov, V. P., Jonsson, B.-H., and Halle, B. (1999) *J. Am. Chem. Soc.* **121**, 2327 – 2328

20. Paneth, P., and O'Leary, M. H. (1985) *J. Am. Chem. Soc.* **107**, 7381 – 7384

21. Scolnick, L. R., and Christianson, D. W. (1996) *Biochemistry* **35**, 16429 – 16434

22. Toba, S., Colombo, G., and Merz, K. M., Jr. (1999) *J. Am. Chem. Soc.* **121**, 2290 – 2302

23. Merz, K. M., Jr., and Banci, L. (1997) *J. Am. Chem. Soc.* **119**, 863 – 871

24. Briganti, F., Mangani, S., Orioli, P., Scozzafava, A., Vernaglione, G., and Supuran, C. T. (1997) *Biochemistry* **36**, 10384 – 10392

25. Ren, X., Tu, C., Laipis, P. J., and Silverman, D. N. (1995) *Biochemistry* **34**, 8492 – 8498

26. Malatesta, V., and Cocivera, M. (1978) *Journal of Organic Chemistry* **43**, 1737 – 1742

27. Jencks, W. P. (1987) *Catalysis in Chemistry and Enzymology*, Dover, Mineola, New York

28. Engel, C. K., Mathieu, M., Zeelen, J. P., Hiltunen, J. K., and Wierenga, R. K. (1996) *EMBO J.* **15**, 5135 – 5145

28a. Engel, C. K., Kiema, T. R., Hiltunen, J. K., and Wierenga, R. K. (1998) *J. Mol. Biol.* **275**, 847 – 859

29. D'Ordine, R. L., Tonge, P. J., Carey, P. R., and Anderson, V. E. (1994) *Biochemistry* **33**, 12635 – 12643

30. Bahnson, B. J., and Anderson, V. E. (1991) *Biochemistry* **30**, 5894 – 5906

31. D'Ordine, R. L., Bahnson, B. J., Tonge, P. J., and Anderson, V. E. (1994) *Biochemistry* **33**, 14733 – 14742

31a. Hofstein, H. A., Feng, Y., Anderson, V. E., and Tonge, P. J. (1999) *Biochemistry* **38**, 9508 – 9516

32. Yang, S.-Y., He, X.-Y., and Schulz, H. (1995) *Biochemistry* **34**, 6441 – 6447

33. Annand, R. R., Kozlowski, J. F., Davisson, V. J., and Schwab, J. (1993) *J. Am. Chem. Soc.* **115**, 1088 – 1094

34. Brock, D. J. H., Kass, L. R., and Bloch, K. (1967) *J. Biol. Chem.* **242**, 4432 – 4440

35. Hebert, H., Schmidt-Krey, I., Morgenstern, R., Murata, K., Hirai, T., Mitsuoka, K., and Fujiyoshi, Y. (1997) *J. Mol. Biol.* **271**, 751 – 758

36. Hu, L., Borleske, B. L., and Colman, R. F. (1997) *Protein Sci.* **6**, 43 – 52

37. Rushmore, T. H., and Pickett, C. B. (1993) *J. Biol. Chem.* **268**, 11475 – 11478

38. Ji, X., Johnson, W. W., Sesay, M. A., Dickert, L., Prasad, S. M., Ammon, H. L., Armstrong, R. N., and Gilliland, G. L. (1994) *Biochemistry* **33**, 1043 – 1052

39. Oakley, A. J., Rossjohn, J., Lo Bello, M., Caccuri, A. M., Federici, G., and Parker, M. W. (1997) *Biochemistry* **36**, 576 – 585

40. Reinemer, P., Prade, L., Hof, P., Neuefeind, T., Huber, R., Zettl, R., Palme, K., Schell, J., Koelln, I., Bartunik, H. D., and Bieseler, B. (1996) *J. Mol. Biol.* **255**, 289 – 309

41. Xiao, G., Liu, S., Ji, X., Johnson, W. W., Chen, J., Parsons, J. F., Stevens, W. J., Gilliland, G. L., and Armstrong, R. N. (1996) *Biochemistry* **35**, 4753 – 4765

41a. Neuefeind, T., Huber, R., Reinemer, P., Knäblein, J., Prade, L., Mann, K., and Bieseler, B. (1997) *J. Mol. Biol.* **274**, 577 – 587

42. McTigue, M. A., Williams, D. R., and Tainer, J. A. (1995) *J. Mol. Biol.* **246**, 21 – 27

43. Yang, G., Liang, P.-H., and Dunaway-Mariano, D. (1994) *Biochemistry* **33**, 8327 – 8331

44. Crooks, G. P., Xu, L., Barkley, R. M., and Copley, S. D. (1995) *J. Am. Chem. Soc.* **117**, 10791 – 10798

45. Yang, G., Liu, R.-Q., Taylor, K. L., Xiang, H., Price, J., and Dunaway-Mariano, D. (1996) *Biochemistry* **36**, 10879 – 10885

46. Benning, M. M., Taylor, K. L., Liu, R.-Q., Yang, G., Xiang, H., Wesenberg, G., Dunaway-Mariano, D., and Holden, H. M. (1996) *Biochemistry* **35**, 8103 – 8109

47. Taylor, K. L., Liu, R.-Q., Liang, P.-H., Price, J., Dunaway-Mariano, D., Tonge, P. J., Clarkson, J., and Carey, P. R. (1995) *Biochemistry* **34**, 13881 – 13888

48. Clarkson, J., Tonge, P. J., Taylor, K. L., Dunaway-Mariano, D., and Carey, P. R. (1997) *Biochemistry* **36**, 10192 – 10199

49. Taylor, K. L., Xiang, H., Liu, R.-Q., Yang, G., and Dunaway-Mariano, D. (1997) *Biochemistry* **36**, 1349 – 1361

49a. Benning, M. M., Haller, T., Gerlt, J. A., and Holden, H. M. (2000) *Biochemistry* **39**, 4630 – 4639

49b. Xiang, H., Luo, L., Taylor, K. L., and Dunaway-Mariano, D. (1999) *Biochemistry* **38**, 7638 – 7652

49c. Babbitt, P. C., and Gerlt, J. A. (1997) *J. Biol. Chem.* **272**, 30591 – 30594

50. Sacchettini, J. C., Meininger, T., Rodrick, S., and Banaszak, L. J. (1986) *J. Biol. Chem.* **261**, 15183 – 15185

51. Bearne, S. L., and Wolfenden, R. (1995) *J. Am. Chem. Soc.* **117**, 9588 – 9589

52. Weaver, T., and Banaszak, L. (1996) *Biochemistry* **35**, 13955 – 13965

53. Weaver, T., Lees, M., and Banaszak, L. (1997) *Protein Sci.* **6**, 834 – 842

53a. Weaver, T., Lees, M., Zaitsev, V., Zaitseva, I., Duke, E., Lindley, P., McSweeny, S., Svensson, A., Keruchenko, J., Keruchenko, I., Gladilin, K., and Banaszak, L. (1998) *J. Mol. Biol.* **280**, 431 – 442

53b. Beeckmans, S., and Van Driessche, E. (1998) *J. Biol. Chem.* **273**, 31661 – 31669

54. Mohrig, J. R., Moerke, K. A., Cloutier, D. L., Lane, B. D., Person, E. C., and Onasch, T. B. (1995) *Science* **269**, 527 – 529

55. Rose, I. A. (1972) *CRC Critical Review of Biochemistry* **1**, 33 – 57

56. Brant, D. A., Barnett, L. B., and Alberty, R. A. (1963) *J. Am. Chem. Soc.* **85**, 2204 – 2209

57. Schmidt, D. E., Jr., Nigh, W. G., Tanzer, C., and Richards, J. H. (1969) *J. Am. Chem. Soc.* **91**, 5849 – 5854

58. Nigh, W. G., and Richards, J. H. (1969) *J. Am. Chem. Soc.* **91**, 5847 – 5848

59. Hansen, J. N., Dinovo, E. C., and Boyer, P. D. (1969) *J. Biol. Chem.* **244**, 6270 – 6279

60. Blanchard, J. S., and Cleland, W. W. (1980) *Biochemistry* **19**, 4506 – 4513

61. Porter, D. J. T., and Bright, H. J. (1980) *J. Biol. Chem.* **255**, 4772 – 4780

62. Jones, V. T., Lowe, G., and Potter, B. V. L. (1980) *Eur. J. Biochem.* **108**, 433 – 437

63. Botting, N. P., and Gani, D. (1992) *Biochemistry* **31**, 1509 – 1520

64. Shi, W., Dunbar, J., Jayasekera, M. M. K., Viola, R. E., and Farber, G. K. (1997) *Biochemistry* **36**, 9136 – 9144

65. Jayasekera, M. M. K., Shi, W., Farber, G. K., and Viola, R. E. (1997) *Biochemistry* **36**, 9145 – 9150

66. Goda, S. K., Minton, N. P., Botting, N. P., and Gani, D. (1992) *Biochemistry* **31**, 10747 – 10756

67. Weiss, P. M., Boerner, R. J., and Cleland, W. W. (1987) *J. Am. Chem. Soc.* **109**, 7201 – 7202

68. Anderson, S. R., Anderson, V. E., and Knowles, J. R. (1994) *Biochemistry* **33**, 10545 – 10555

68a. Vinarov, D. A., and Nowak, T. (1998) *Biochemistry* **37**, 15238 – 15246

69. Vinarov, D. A., and Nowak, T. (1999) *Biochemistry* **38**, 12138 – 12149

70. Zhang, E., Hatada, M., Brewer, J. M., and Lebioda, L. (1994) *Biochemistry* **33**, 6295 – 6300

71. Wedekind, J. E., Reed, G. H., and Rayment, I. (1995) *Biochemistry* **34**, 4325 – 4330

72. Duquerroy, S., Camus, C., and Janin, J. (1995) *Biochemistry* **34**, 12513 – 12523

73. Larsen, T. M., Wedekind, J. E., Rayment, I., and Reed, G. H. (1996) *Biochemistry* **35**, 4349 – 4358

73a. Hubbard, B. K., Koch, M., Palmer, D. R. J., Babbitt, P. C., and Gerlt, J. A. (1998) *Biochemistry* **37**, 14369 – 14375

73b. Thompson, T. B., Garrett, J. B., Taylor, E. A., Meganathan, R., Gerlt, J. A., and Rayment, I. (2000) *Biochemistry* **39**, 10662 – 10676

74. Linhardt, R. J., Turnbull, J. E., Wang, H. M., Loganathan, D., and Gallagher, J. T. (1990) *Biochemistry* **29**, 2611 – 2617

75. Desai, U. R., Wang, H.-M., and Linhardt, R. J. (1993) *Biochemistry* **32**, 8140 – 8145

76. Kiss, J. (1974) *Adv. Carbohydrate Chem. Biochem.* **29**, 229 – 303

76a. Li, S., Kelly, S. J., Lamani, E., Ferraroni, M., and Jedrzejas, M. J. (2000) *EMBO J.* **19**, 1228 – 1240

References

76b. Ponnuraj, K., and Jedrzejas, M. J. (2000) *J. Mol. Biol.* **299**, 885 – 895
77. Yoder, M. D., Keen, N. T., and Jurnak, F. (1993) *Science* **260**, 1503 – 1507
78. Cohen, F. E. (1993) *Science* **260**, 1444 – 1445
79. Kita, N., Boyd, C. M., Garrett, M. R., Jurnak, F., and Keen, N. T. (1996) *J. Biol. Chem.* **271**, 26529 – 26535
80. Kuo, D. J., and Rose, I. A. (1987) *Biochemistry* **26**, 7589 – 7596
80a. Lloyd, S. J., Lauble, H., Prasad, G. S., and Stout, C. D. (1999) *Protein Sci.* **8**, 2655 – 2662
81. Werst, M. M., Kennedy, M. C., Houseman, A. L. P., Beinert, H., and Hoffman, B. M. (1990) *Biochemistry* **29**, 10533 – 10540
82. Kilpatrick, L. K., Kennedy, M. C., Beinert, H., Czernuszewicz, R. S., Qiu, D., and Spiro, T. G. (1994) *J. Am. Chem. Soc.* **116**, 4053 – 4061
83. Lauble, H., Kennedy, M. C., Beinert, H., and Stout, C. D. (1992) *Biochemistry* **31**, 2735 – 2748
84. Kennedy, M. C., Emptage, M. H., and Beinert, H. (1984) *J. Biol. Chem.* **259**, 3145 – 3151
85. Kennedy, M. C., and Beinert, H. (1988) *J. Biol. Chem.* **263**, 8194 – 8198
86. Beinert, H., and Kennedy, M. C. (1993) *FASEB J.* **7**, 1442 – 1448
87. Kent, T. A., Emptage, M. H., Merkle, H., Kennedy, M. C., Beinert, H., and Münck, E. (1985) *J. Biol. Chem.* **260**, 6871 – 6881
88. Lauble, H., Kennedy, M. C., Beinert, H., and Stout, C. D. (1994) *J. Mol. Biol.* **237**, 437 – 451
89. Zheng, L., Kennedy, M. C., Beinert, H., and Zalkin, H. (1992) *J. Biol. Chem.* **267**, 7895 – 7903
90. Gruer, M. J., Artymiuk, P. J., and Guest, J. R. (1997) *Trends Biochem. Sci.* **22**, 3 – 6
91. Flint, D. H. (1993) *Biochemistry* **32**, 799 – 805
92. Flint, D. H., and McKay, R. G. (1994) *J. Am. Chem. Soc.* **116**, 5534 – 5539
93. Hofmeister, A. E. M., Grabowski, R., Linder, D., and Buckel, W. (1993) *Eur. J. Biochem.* **215**, 341 – 349
94. Grabowski, R., Hofmeister, A. E. M., and Buckel, W. (1993) *Trends Biochem. Sci.* **18**, 297 – 300
95. Hofmeister, A. E. M., Berger, S., and Buckel, W. (1992) *Eur. J. Biochem.* **205**, 743 – 749
96. Schroepfer, G. J., Jr. (1966) *J. Biol. Chem.* **241**, 5441 – 5447
97. Hill, R. K., and Newkome, G. R. (1969) *J. Am. Chem. Soc.* **91**, 5893 – 5894
98. Chiang, Y., and Kresge, A. J. (1991) *Science* **253**, 395 – 400
99. Vellom, D. C., Radic′, Z., Li, Y., Pickering, N. A., Camp, S., and Taylor, P. (1993) *Biochemistry* **32**, 12 – 17
100. Gerlt, J. A., and Gassman, P. G. (1993) *Biochemistry* **32**, 11943 – 11952
101. Petsko, G. A., Kenyon, G. L., Gerlt, J. A., Ringe, D., and Kozarich, J. W. (1993) *Trends Biochem. Sci.* **18**, 372 – 376
102. Neidhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R., Kenyon, G. L., and Gerlt, J. A. (1991) *Biochemistry* **30**, 9264 – 9273
103. St. Maurice, M., and Bearne, S. L. (2000) *Biochemistry* **39**, 13324 – 13335
104. Schafer, S. L., Barrett, W. C., Kallarakal, A. T., Mitra, B., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., and Kenyon, G. L. (1996) *Biochemistry* **35**, 5662 – 5669
105. Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R. J., Barrett, W. C., Reed, G. H., Rayment, I., Ringe, D., Kenyon, G. L., and Gerlt, J. A. (1996) *Biochemistry* **35**, 16489 – 16501
106. Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L., and Gerlt, J. A. (1995) *Science* **267**, 1159 – 1161
107. Chiang, Y., Kresge, A. J., Pruszynski, P., Schepp, N. P., and Wirz, J. (1990) *Angew. Chem. Int. Ed. Engl.* **29**, 792 – 794
108. Guthrie, J. P., and Kluger, R. (1993) *J. Am. Chem. Soc.* **115**, 11569 – 11572
109. Gerlt, J. A., Kozarich, J. W., Kenyon, G. L., and Gassman, P. G. (1991) *J. Am. Chem. Soc.* **113**, 9667 – 9669
110. Gerlt, J. A., and Gassman, P. G. (1993) *J. Am. Chem. Soc.* **115**, 11552 – 11568
111. Guthrie, J. P. (1997) *J. Am. Chem. Soc.* **119**, 1151 – 1152
112. Tobin, J. B., and Frey, P. A. (1996) *J. Am. Chem. Soc.* **118**, 12253 – 12260
113. Yamauchi, T., Choi, S.-Y., Okada, H., Yohda, M., Kumagai, H., Esaki, N., and Soda, K. (1992) *J. Biol. Chem.* **267**, 18361 – 18364
114. Glavas, S., and Tanner, M. E. (1999) *Biochemistry* **38**, 4106 – 4113
115. Ho, H.-T., Falk, P. J., Ervin, K. M., Krishnan, B. S., Discotto, L. F., Dougherty, T. J., and Pucci, M. J. (1995) *Biochemistry* **34**, 2464 – 2470
115a. Hwang, K. Y., Cho, C.-S., Kim, S. S., Sung, H.-C., Yu, Y. G., and Cho, Y. (1999) *Nature Struct. Biol.* **6**, 422 – 426
116. Stein, T., Kluge, B., Vater, J., Franke, P., Otto, A., and Wittmann-Liebold, B. (1995) *Biochemistry* **34**, 4633 – 4642
117. Wiseman, J. S., and Nichols, J. S. (1984) *J. Biol. Chem.* **259**, 8907 – 8914
118. Shikata, Y., Watanabe, T., Teramoto, T., Inoue, A., Kawakami, Y., Nishizawa, Y., Katayama, K., and Kuwada, M. (1995) *J. Biol. Chem.* **270**, 16719 – 16723
119. Heck, S. D., Faraci, W. S., Kelbaugh, P. R., Saccomano, N. A., Thadeio, P. F., and Volkmann, R. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4036 – 4039
120. Johnson, J. D., Creighton, D. J., and Lambert, M. R. (1986) *J. Biol. Chem.* **261**, 4535 – 4541
121. Cooper, A. J. L., Gines, J. Z., and Meister, A. (1983) *Chem. Rev.* **83**, 321 – 358
122. Pirrung, M. C., Chen, J., Rowley, E. G., and McPhail, A. T. (1993) *J. Am. Chem. Soc.* **115**, 7103 – 7110
123. Noltmann, E. A. (1972) in *The Enzymes*, 3rd ed., Vol. 6 (Boyer, P. D., ed), pp. 271 – 354, Academic Press, New York
124. McGee, D. M., Hathaway, G. H., Palmieri, R. H., and Noltmann, E. A. (1980) *J. Mol. Biol.* **142**, 29 – 42
125. Mushegian, A. R., and Koonin, E. V. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10268 – 10273
125a. Jeffery, C. J., Bahnson, B. J., Chien, W., Ringe, D., and Petsko, G. A. (2000) *Biochemistry* **39**, 955 – 964
126. Tanaka, K. R., and Paglia, D. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3485 – 3511, McGraw-Hill, New York
127. Woodruff, W. W., III, and Wolfenden, R. (1979) *J. Biol. Chem.* **254**, 5866 – 5867
127a. Seemann, J. E., and Schulz, G. E. (1997) *J. Mol. Biol.* **273**, 256 – 268
128. Wallace, L. J., Eiserling, F. A., and Wilcox, G. (1978) *J. Biol. Chem.* **253**, 3717 – 3720
129. Bogumil, R., Kappl, R., Hüttermann, J., and Witzel, H. (1997) *Biochemistry* **36**, 2345 – 2352
130. Speck, J. C., Jr. (1958) *Adv. Carbohydrate Chem.* **13**, 63 – 103
131. Gibson, D. R., Gracy, R. W., and Hartman, F. C. (1980) *J. Biol. Chem.* **255**, 9369 – 9374
132. Balaban, R. S., and Ferretti, J. A. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1241 – 1245
133. Blacklow, S. C., Raines, R. T., Lim, W. A., Zamore, P. D., and Knowles, J. R. (1988) *Biochemistry* **27**, 1158 – 1167
134. Muirhead, H. (1983) *Trends Biochem. Sci.* **8**, 326 – 330
135. Lolis, E., Alber, T., Davenport, R. C., Rose, D., Hartman, F. C., and Petsko, G. A. (1990) *Biochemistry* **29**, 6609 – 6618
136. Joseph-McCarthy, D., Lolis, E., Komives, E. A., and Petsko, G. A. (1994) *Biochemistry* **33**, 2815 – 2823
137. Mande, S. C., Mainfroid, V., Kalk, K. H., Goraj, K., Martial, J. A., and Hol, W. G. J. (1994) *Protein Sci.* **3**, 810 – 821
138. Davenport, R. C., Bash, P. A., Seaton, B. A., Karplus, M., Petsko, G. A., and Ringe, D. (1991) *Biochemistry* **30**, 5821 – 5826
139. Verlinde, C. L. M. J., Witmans, C. J., Pijning, T., Kalk, K. H., Hol, W. G. J., Callens, M., and Opperdoes, F. R. (1992) *Protein Sci.* **1**, 1578 – 1584
140. Rose, I. A., Fung, W.-J., and Warms, J. V. B. (1990) *Biochemistry* **29**, 4312 – 4317
140a. Harris, T. K., Cole, R. N., Comer, F. I., and Mildvan, A. S. (1998) *Biochemistry* **37**, 16828 – 16838
141. Joseph-McCarthy, D., Rost, L. E., Komives, E. A., and Petsko, G. A. (1994) *Biochemistry* **33**, 2824 – 2829
142. Belasco, J. G., and Knowles, J. R. (1980) *Biochemistry* **19**, 472 – 477
143. Alston, W. C., II, Kanska, M., and Murray, C. J. (1996) *Biochemistry* **35**, 12873 – 12881
144. Nickbarg, E. B., Davenport, R. C., Petsko, G. A., and Knowles, J. R. (1988) *Biochemistry* **27**, 5948 – 5960
145. Lodi, P. J., Chang, L. C., Knowles, J. R., and Komives, E. A. (1994) *Biochemistry* **33**, 2809 – 2814
146. Raines, R. T., and Knowles, J. R. (1987) *Biochemistry* **26**, 7014 – 7020
147. Bash, P. A., Field, M. J., Davenport, R. C., Petsko, G. A., Ringe, D., and Karplus, M. (1991) *Biochemistry* **30**, 5826 – 5832
148. Lodi, P. J., and Knowles, J. R. (1991) *Biochemistry* **30**, 6948 – 6956
149. Alagona, G., Ghio, C., and Kollman, P. A. (1995) *J. Am. Chem. Soc.* **117**, 9855 – 9862
150. Pompliano, D. L., Peyman, A., and Knowles, J. R. (1990) *Biochemistry* **29**, 3186 – 3194
151. Yüksel, K. Ü., Sun, A.-Q., Gracy, R. W., and Schnackerz, K. D. (1994) *J. Biol. Chem.* **269**, 5005 – 5008
152. Williams, J. C., and McDermott, A. E. (1995) *Biochemistry* **34**, 8309 – 8319
153. Carrell, H. L., Rubin, B. H., Hurley, T. J., and Glusker, J. P. (1984) *J. Biol. Chem.* **259**, 3230 – 3236
154. Farber, G. K., Glasfeld, A., Tiraby, G., Ringe, D., and Petsko, G. A. (1989) *Biochemistry* **28**, 7289 – 7297
155. Collyer, C. A., Goldberg, J. D., Viehmann, H., Blow, D. M., Ramsden, N. G., Fleet, G. W. J., Montgomery, F. J., and Grice, P. (1992) *Biochemistry* **31**, 12211 – 12218
156. Jenkins, J., Janin, J., Rey, F., Chiadmi, M., van Tilbeurgh, H., Lasters, I., De Maeyer, M., Van Belle, D., Wodak, S. J., Lauwereys, M., Stanssens, P., Mrabet, N. T., Snauwaert, J., Matthyssens, G., and Lambeir, A.-M. (1992) *Biochemistry* **31**, 5449 – 5458
157. Whitaker, R. D., Cho, Y., Cha, J., Carrell, H. L., Glusker, J. P., Karplus, P. A., and Batt, C. A. (1995) *J. Biol. Chem.* **270**, 22895 – 22906
158. Allen, K. N., Lavie, A., Petsko, G. A., and Ringe, D. (1995) *Biochemistry* **34**, 3742 – 3749
159. Hall, S. S., Doweyko, A. M., and Jordan, F. (1978) *J. Am. Chem. Soc.* **100**, 5934
160. Xue, L., Talalay, P., and Mildvan, A. S. (1991) *Biochemistry* **30**, 10858 – 10865
161. Brooks, B., and Benisek, W. F. (1994) *Biochemistry* **33**, 2682 – 2687
162. Austin, J. C., Zhao, Q., Jordan, T., Talalay, P., Mildvan, A. S., and Spiro, T. G. (1995) *Biochemistry* **34**, 4441 – 4447

References

163. Zhao, Q., Li, Y.-K., Mildvan, A. S., and Talalay, P. (1995) *Biochemistry* **34**, 6562–6572

164. Hawkinson, D. C., Pollack, R. M., and Ambulos, N. P., Jr. (1994) *Biochemistry* **33**, 12172–12183

165. Viger, A., Coustal, S., and Marquet, A. (1983) *J. Am. Chem. Soc.* **103**, 451–458

166. Holman, C. M., and Benisek, W. F. (1994) *Biochemistry* **33**, 2672–2681

167. Wu, Z. R., Ebrahimian, S., Zawrotny, M. E., Thornburg, L. D., Perez-Alvarado, G. C., Brothers, P., Pollack, R. M., and Summers, M. F. (1997) *Science* **276**, 415–418

167a. Kim, S. W., Cha, S.-S., Cho, H.-S., Kim, J.-S., Ha, N.-C., Cho, M.-J., Joo, S., Kim, K. K., Choi, K. Y., and Oh, B.-H. (1997) *Biochemistry* **36**, 14030–14036

167b. Choi, G., Ha, N.-C., Kim, S. W., Kim, D.-H., Park, S., Oh, B.-H., and Choi, K. Y. (2000) *Biochemistry* **39**, 903–909

167c. Thornburg, L. D., Hénot, F., Bash, D. P., Hawkinson, D. C., Bartel, S. D., and Pollack, R. M. (1998) *Biochemistry* **37**, 10499–10506

168. Chen, L. H., Kenyon, G. L., Curtin, F., Harayama, S., Bembenek, M. E., Hajipour, G., and Whitman, C. P. (1992) *J. Biol. Chem.* **267**, 17716–17721

169. Subramanya, H. S., Roper, D. I., Dauter, Z., Dodson, E. J., Davies, G. J., Wilson, K. S., and Wigley, D. B. (1996) *Biochemistry* **35**, 792–802

170. Taylor, A. B., Czerwinski, R. M., Johnson, W. H., Jr., Whitman, C. P., and Hackert, M. L. (1998) *Biochemistry* **37**, 14692–14700

171. Stivers, J. T., Abeygunawardana, C., Mildvan, A. S., Hajipour, G., and Whitman, C. P. (1996) *Biochemistry* **35**, 814–823

172. Czerwinski, R. M., Harris, T. K., Johnson, W. H., Jr., Legler, P. M., Stivers, J. T., Mildvan, A. S., and Whitman, C. P. (1999) *Biochemistry* **38**, 12358–12366

173. Fitzgerald, M. C., Chernushevich, I., Standing, K. G., Kent, S. B. H., and Whitman, C. P. (1995) *J. Am. Chem. Soc.* **117**, 11075–11080

174. Lo, T. W. C., Westwood, M. E., McLellan, A. C., Selwood, T., and Thornalley, P. J. (1994) *J. Biol. Chem.* **269**, 32299–32305

175. Ridderström, M., Saccucci, F., Hellman, U., Bergman, T., Principato, G., and Mannervik, B. (1996) *J. Biol. Chem.* **271**, 319–323

176. Richard, J. P. (1991) *Biochemistry* **30**, 4581–4585

176a. Saadat, D., and Harrison, D. H. T. (2000) *Biochemistry* **39**, 2950–2960

177. Lan, Y., Lu, T., Lovett, P. S., and Creighton, D. J. (1995) *J. Biol. Chem.* **270**, 12957–12960

178. Cameron, A. D., Ridderström, M., Olin, B., Kavarana, M. J., Creighton, D. J., and Mannervik, B. (1999) *Biochemistry* **38**, 13480–13490

178a. He, M. M., Clugston, S. L., Honek, J. F., and Matthews, B. W. (2000) *Biochemistry* **39**, 8719–8727

179. Rae, C., O'Donoghue, S. I., Bubb, W. A., and Kuchel, P. W. (1994) *Biochemistry* **33**, 3548–3559

180. Bito, A., Haider, M., Hadler, I., and Breitenbach, M. (1997) *J. Biol. Chem.* **272**, 21509–21519

181. Mathieu, M., Modis, Y., Zeelen, J. P., Engel, C. K., Abagyan, R. A., Ahlberg, A., Rasmussen, B., Lamzin, V. S., Kunau, W. H., and Wierenga, R. K. (1997) *J. Mol. Biol.* **273**, 714–728

181a. Modis, Y., and Wierenga, R. K. (2000) *J. Mol. Biol.* **297**, 1171–1182

182. Roberts, J. R., Narasimhan, C., and Miziorko, H. M. (1995) *J. Biol. Chem.* **270**, 17311–17316

183. Narasimhan, C., Roberts, J. R., and Miziorko, H. M. (1995) *Biochemistry* **34**, 9930–9935

183a. Vinarov, D. A., and Miziorko, H. M. (2000) *Biochemistry* **39**, 3360–3368

183b. Richard, J. P., and Nagorski, R. W. (1999) *J. Am. Chem. Soc.* **121**, 4763–4770

184. Marsh, J. J., and Lebherz, H. G. (1992) *Trends Biochem. Sci.* **17**, 110–113

185. Cox, T. M. (1994) *FASEB J.* **8**, 62–71

186. Morris, A. J., and Tolan, D. R. (1994) *Biochemistry* **33**, 12291–12297

187. Sygusch, J., Beaudry, D., and Allaire, M. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7846–7850

188. Lai, C. Y., and Oshima, T. (1971) *Arch. Biochem. Biophys.* **144**, 363–

188a. Choi, K. H., Mazurkie, A. S., Morris, A. J., Utheza, D., Tolan, D. R., and Allen, K. N. (1999) *Biochemistry* **38**, 12655–12664

188b. Dalby, A., Dauter, Z., and Littlechild, J. A. (1999) *Protein Sci.* **8**, 291–297

189. Periana, R. A., Motiu-DeGrood, R., Chiang, Y., and Hupe, D. J. (1980) *J. Am. Chem. Soc.* **102**, 3923–3927

190. Beernink, P. T., and Tolan, D. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5374–5379

191. Kadonaga, J. T., and Knowles, J. R. (1983) *Biochemistry* **22**, 130–136

191a. Hall, D. R., Leonard, G. A., Reed, C. D., Watt, C. I., Berry, A., and Hunter, W. N. (1999) *J. Mol. Biol.* **287**, 383–394

192. Qamar, S., Marsh, K., and Berry, A. (1996) *Protein Sci.* **5**, 154–161

193. Dreyer, M. K., and Schulz, G. E. (1996) *J. Mol. Biol.* **259**, 458–466

194. Lebioda, L., Hatada, M. H., Tulinsky, A., and Mavridis, I. M. (1982) *J. Mol. Biol.* **162**, 445–458

195. Meloche, H. P. (1981) *Trends Biochem. Sci.* **6**, 38–41

196. Dekker, E. E., and Kitson, R. P. (1992) *J. Biol. Chem.* **267**, 10507–10514

197. Wong, C.-H., Garcia-Junceda, E., Chen, L., Blanco, O., Gijsen, H. J. M., and Steensma, D. H. (1995) *J. Am. Chem. Soc.* **117**, 3333–3339

198. Jia, J., Schörken, U., Lindqvist, Y., Sprenger, G. A., and Schneider, G. (1997) *Protein Sci.* **6**, 119–124

199. Higgins, M. J. P., Kornblatt, J. A., and Rudney, H. (1972) in *The Enzymes*, 3rd ed., Vol. 7 (Boyer, P. D., ed), pp. 407–434, Academic Press, New York

200. Wiegand, G., and Remington, S. J. (1986) *Ann. Rev. Biophys. Biophys. Chem.* **15**, 97–117

201. Anderson, D. H., and Duckworth, H. W. (1988) *J. Biol. Chem.* **263**, 2163–2169

202. Wiegand, G., Remington, S., Deisenhofer, J., and Huber, R. (1984) *J. Mol. Biol.* **174**, 205–219

203. Rault-Leonardon, M., Atkinson, M. A. L., Slaughter, C. A., Moomaw, C. R., and Srere, P. A. (1995) *Biochemistry* **34**, 257–263

204. Russell, R. J. M., Ferguson, J. M. C., Hough, D. W., Danson, M. J., and Taylor, G. L. (1997) *Biochemistry* **36**, 9983–9994

205. Pereira, F. D., Donald, L. J., Hosfield, D. J., and Duckworth, H. W. (1994) *J. Biol. Chem.* **269**, 412–417

206. Karpusas, M., Holland, D., and Remington, S. J. (1991) *Biochemistry* **30**, 6024–6031

207. Chothia, C., and Lesk, A. M. (1985) *Trends Biochem. Sci.* **10**, 116–118

208. Evans, C. T., Kurz, L. C., Remington, S. J., and Srere, P. A. (1996) *Biochemistry* **35**, 10661–10672

209. Remington, S., Wiegand, G., and Huber, R. (1982) *J. Mol. Biol.* **158**, 111 – 152

210. Kurz, L. C., Roble, J. H., Nakra, T., Drysdale, G. R., Buzan, J. M., Schwartz, B., and Drueckhammer, D. G. (1997) *Biochemistry* **36**, 3981–3990

211. Kurz, L. C., and Drysdale, G. R. (1987) *Biochemistry* **26**, 2623–2627

212. Kurz, L. C., Ackerman, J. J. H., and Drysdale, G. R. (1985) *Biochemistry* **24**, 452–457

213. Clark, J. D., O'Keefe, S. J., and Knowles, J. R. (1988) *Biochemistry* **27**, 5961–5971

214. Wlassics, I. D., and Anderson, V. E. (1989) *Biochemistry* **28**, 1627–1633

215. Usher, K. C., Remington, S. J., Martin, D. P., and Drueckhammer, D. G. (1994) *Biochemistry* **33**, 7753–7759

215a. Kurz, L. C., Nakra, T., Stein, R., Plungkhen, W., Riley, M., Hsu, F., and Drysdale, G. R. (1998) *Biochemistry* **37**, 9724–9737

216. Schwartz, B., Drueckhammer, D. G., Usher, K. C., and Remington, S. J. (1995) *Biochemistry* **34**, 15459–15466

216a. Gu, Z., Drueckhammer, D. G., Kurz, L., Liu, K., Martin, D. P., and McDermott, A. (1999) *Biochemistry* **38**, 8022–8031

216b. Mulholland, A. J., Lyne, P. D., and Karplus, M. (2000) *J. Am. Chem. Soc.* **122**, 534–535

216c. Howard, B. R., Endrizzi, J. A., and Remington, S. J. (2000) *Biochemistry* **39**, 3156–3168

217. Smith, S. (1994) *FASEB J.* **8**, 1248–1259

218. Misra, I., Narasimhan, C., and Miziorko, H. M. (1993) *J. Biol. Chem.* **268**, 12129–12135

219. Misra, I., and Miziorko, H. M. (1996) *Biochemistry* **35**, 9610–9616

220. Guynn, R. W., and Veech, R. L. (1979) *J. Biol. Chem.* **254**, 1691–1698

221. Linn, T. C., and Srere, P. A. (1979) *J. Biol. Chem.* **254**, 1691–1698

222. Ranganathan, N. S., Linn, T. C., and Srere, P. A. (1982) *J. Biol. Chem.* **257**, 698–702

223. Elshourbagy, N. A., Near, J. C., Kmetz, P. J., Sathe, G. M., Southan, C., Strickler, J. E., Gross, M., Young, J. F., Wells, T. N. C., and Groot, P. H. E. (1990) *J. Biol. Chem.* **265**, 1430–1435

224. Pentyala, S. N., and Benjamin, W. B. (1995) *Biochemistry* **34**, 10961–10969

225. Hersch, L. B. (1973) *J. Biol. Chem.* **248**, 7295–7303

226. Nilekani, S., and SivaRaman, C. (1983) *Biochemistry* **22**, 4657–4663

227. Oppenheimer, N. J., Singh, M., Sweeley, C. C., Sung, S.-J., and Srere, P. A. (1979) *J. Biol. Chem.* **254**, 1000–1002

228. Ko, Y. H., Vanni, P., Munske, G. R., and McFadden, B. A. (1991) *Biochemistry* **30**, 7451–7456

229. Lenz, H., Buckel, W., Wunderwald, P., Biedermann, G., Buschmeier, V., Eggerer, H., Cornforth, J. W., Redmond, J. W., and Mallaby, R. (1971) *Eur. J. Biochem.* **24**, 207–215

230. Retey, J., Luthy, J., and Arigoni, D. (1970) *Nature (London)* **226**, 519–521

231. Kyte, J. (1995) *Mechanism in Protein Chemistry*, Garland Publ., New York (pp. 293–313)

232. Leussing, D. L., and Raghavan, N. V. (1980) *J. Am. Chem. Soc.* **102**, 5635–5643

233. Steinberger, R., and Westheimer, F. H. (1951) *J. Am. Chem. Soc.* **73**, 429–435

234. Kubala, G., and Martell, A. E. (1982) *J. Am. Chem. Soc.* **104**, 6602–6609

235. Waldrop, G. L., Braxton, B. F., Urbauer, J. L., Cleland, W. W., and Kiick, D. M. (1994) *Biochemistry* **33**, 5262–5267

236. Highbarger, L. A., Gerlt, J. A., and Kenyon, G. L. (1996) *Biochemistry* **35**, 41–46

236a. Stanley, T. M., Johnson, W. H., Jr., Burks, E. A., Whitman, C. P., Hwang, C.-C., and Cook, P. F. (2000) *Biochemistry* **39**, 718–726

237. Loeber, G., Infante, A. A., Maurer-Fogy, I., Krystek, E., and Dworkin, M. B. (1991) *J. Biol. Chem.* **266**, 3016–3021

238. Chou, W.-Y., Liu, M.-Y., Huang, S.-M., and Chang, G.-G. (1996) *Biochemistry* **35**, 9873–9879

239. Winning, B. M., Bourguignon, J., and Leaver, C. J. (1994) *J. Biol. Chem.* **269**, 4780–4786

240. Wei, C.-H., Chou, W.-Y., and Chang, G.-G. (1995) *Biochemistry* **34**, 7949–7954

241. Karsten, W. E., and Cook, P. F. (1994) *Biochemistry* **33**, 2096–2103

References

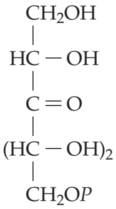
241a. Liu, D., Karsten, W. E., and Cook, P. F. (2000) *Biochemistry* **39**, 11955–11960
242. Edens, W. A., Urbauer, J. L., and Cleland, W. W. (1997) *Biochemistry* **36**, 1141–1147
242a. Gruys, K. J., and Sikorski, J. A. (1998) in *Comprehensive Biological Catalysis. A Mechanistic Reference*, Vol. I (Sinnott, M., ed), pp. 273–291, Academic Press, San Diego, California
243. Bolduc, J. M., Dyer, D. H., Scott, W. G., Singer, P., Sweet, R. M., Koshland, D. E., Jr., and Stoddard, B. L. (1995) *Science* **268**, 1312–1318
244. Berdis, A. J., and Cook, P. F. (1993) *Biochemistry* **32**, 2041–2046
245. Dean, A. M., and Dvorak, L. (1995) *Protein Sci.* **4**, 2156–2167
246. Wallon, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D., and Petsko, G. A. (1997) *J. Mol. Biol.* **266**, 1016–1031
247. Brubaker, M. J., Dyer, D. H., Stoddard, B., and Koshland, D. E., Jr. (1996) *Biochemistry* **35**, 2854–2864
248. Tipton, P. A. (1993) *Biochemistry* **32**, 2822–2827
249. Kalckar, H. M. (1985) *Trends Biochem. Sci.* **10**, 132–133
250. Peliska, J. A., and O'Leary, M. H. (1991) *J. Am. Chem. Soc.* **113**, 1841–1842
251. Seeholzer, S. H., Jaworowski, A., and Rose, I. A. (1991) *Biochemistry* **30**, 727–732
252. Matte, A., Tari, L. W., Goldie, H., and Delbaere, L. T. J. (1997) *J. Biol. Chem.* **272**, 8105–8108
253. Hwang, S. H., and Nowak, T. (1986) *Biochemistry* **25**, 5590–5595
254. Konopka, J. M., Lardy, H. A., and Frey, P. A. (1986) *Biochemistry* **25**, 5571–5575
255. Hlavaty, J. J., and Nowak, T. (1997) *Biochemistry* **36**, 3389–3403
256. Matte, A., Goldie, H., Sweet, R. M., and Delbaere, L. T. J. (1996) *J. Mol. Biol.* **256**, 126–143
257. Arnelle, D. R., and O'Leary, M. H. (1992) *Biochemistry* **31**, 4363–4368
258. Punekar, N. S., and Lardy, H. A. (1987) *J. Biol. Chem.* **262**, 6714–6719
259. Roseler, W. J., Vandenbark, G. R., and Hanson, R. H. (1989) *J. Biol. Chem.* **264**, 9657–9664
260. Rohrer, S. D., Saz, H. J., and Nowak, T. (1986) *J. Biol. Chem.* **261**, 13049–13055
261. Hebdo, C. A., and Nowak, T. (1982) *J. Biol. Chem.* **257**, 5503–5514
262. Hartman, F. C., and Harpel, M. R. (1994) *Ann. Rev. Biochem.* **63**, 197–234
263. Wasmann, C. C., Reiss, B., and Bohnert, H. J. (1988) *J. Biol. Chem.* **263**, 617–619
264. Shibata, N., Inoue, T., Fukuhara, K., Nagara, Y., Kitagawa, R., Harada, S., Kasai, N., Uemura, K., Kato, K., Yokota, A., and Kai, Y. (1996) *J. Biol. Chem.* **271**, 26449–26452
265. Tessier, L. H., Paulus, F., Keller, M., Vial, C., and Imbault, P. (1995) *J. Mol. Biol.* **245**, 22–33
266. Holzenburg, A., Mayer, F., Harauz, G., van Heel, M., Tokuoka, R., Ishida, T., Harata, K., Pal, G. P., and Saenger, W. (1987) *Nature (London)* **325**, 730–732
267. Erijman, L., Lorimer, G. H., and Weber, G. (1993) *Biochemistry* **32**, 5187–5195
268. Morse, D., Salois, P., Markovic, P., and Hastings, J. W. (1995) *Science* **268**, 1622–1624
269. Andersson, I. (1996) *J. Mol. Biol.* **259**, 160–174
269a. Schreuder, H. A., Knight, S., Curmi, P. M. G., Andersson, I., Cascio, D., Sweet, R. M., Brändén, C.-I., and Eisenberg, D. (1993) *Protein Sci.* **2**, 1136–1146
270. Newman, J., and Gutteridge, S. (1993) *J. Biol. Chem.* **268**, 25876–25886
271. Bassham, J. A., Benson, A. A., Kay, L. D., Harris, A. Z., Wilson, A. T., and Calvin, M. (1954) *J. Am. Chem. Soc.* **76**, 1760–1770
272. Siegel, M. I., and Lane, M. D. (1973) *J. Biol. Chem.* **248**, 5486–5498
272a. Taylor, T. C., and Andersson, I. (1996) *Nature Struct. Biol.* **3**, 95–101
272b. Harpel, M. R., Larimer, F. W., and Hartman, F. C. (1998) *Protein Sci.* **7**, 730–738
272c. Harpel, M. R., and Hartman, F. C. (1996) *Biochemistry* **35**, 13865–13870
272d. Duff, A. P., Andrews, T. J., and Curmi, R. M. G. (2000) *J. Mol. Biol.* **298**, 903–916
273. Pierce, J., Andrews, T. J., and Lorimer, G. H. (1986) *J. Biol. Chem.* **261**, 10248–10256
273a. Berry, J. A., Lorimer, G. H., Pierce, J., Seeman, J. R., Meeks, J., and Freas, S. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 734–738
274. Jaworowski, A., Hartman, F. C., and Rose, I. A. (1984) *J. Biol. Chem.* **259**, 6783–6789
275. Saver, B. G., and Knowles, J. R. (1982) *Biochemistry* **21**, 5398–5403
276. Taylor, T. C., and Andersson, I. (1997) *J. Mol. Biol.* **265**, 432–444
277. Roeske, C. A., and O'Leary, M. H. (1984) *Biochemistry* **23**, 6275–6284
278. Van Dyke, D. E., and Schloss, J. V. (1986) *Biochemistry* **25**, 5145–5156
279. van de Loo, F. J., and Salvucci, M. E. (1996) *Biochemistry* **35**, 8143–8148
280. de Jiménez, E. S., Medrano, L., and Martínez-Barajas, E. (1995) *Biochemistry* **34**, 2826–2831
281. Larson, E. M., O'Brien, C. M., Zhu, G., Spreitzer, R. J., and Portis, A. R., Jr. (1997) *J. Biol. Chem.* **272**, 17033–17037
282. Mueller, D. D., Schmidt, A., Pappan, K. L., McKay, R. A., and Schaefer, J. (1995) *Biochemistry* **34**, 5597–5603
282a. Ogren, W. L. (1984) *Ann. Rev. Plant Physiol.* **35**, 415–442
283. Tse, J. M.-T., and Schloss, J. V. (1993) *Biochemistry* **32**, 10398–10403
283a. Sugawara, H., Yamamoto, H., Shibata, N., Inoue, T., Okada, S., Miyake, C., Yokota, A., and Kai, Y. (1999) *J. Biol. Chem.* **274**, 15655–15661
284. Harpel, M. R., Serpersu, E. H., Lamerdin, J. A., Huang, Z.-H., Gage, D. A., and Hartman, F. C. (1995) *Biochemistry* **34**, 11296–11306
284a. Gutteridge, S., Rhoades, D. F., and Herrmann, C. (1993) *J. Biol. Chem.* **268**, 7818–7824
285. Morell, M. K., Wilkin, J.-M., Kane, H. J., and Andrews, T. J. (1997) *J. Biol. Chem.* **272**, 5445–5451
286. Krebs, H. A., and Roughton, F. J. W. (1948) *Biochem. J.* **43**, 550–555
287. Kaziro, Y., Hass, L. F., Boyer, P. D., and Ochoa, S. (1962) *J. Biol. Chem.* **237**, 1460–1468
288. Cooper, T. G., Tchen, T. T., Wood, H. G., and Benedict, C. R. (1968) *J. Biol. Chem.* **243**, 3857–3863
289. González, D. H., and Andreo, C. S. (1988) *Biochemistry* **27**, 177–183
290. Harpster, M. H., and Taylor, W. C. (1986) *J. Biol. Chem.* **261**, 6132–6136
291. O'Leary, M. H., Rife, J. E., and Slater, J. D. (1981) *Biochemistry* **20**, 7308–7314
292. Hansen, D. E., and Knowles, J. R. (1982) *J. Biol. Chem.* **257**, 14795–14798
293. Fujita, N., Izui, K., Nishino, I., and Katsuki, H. (1984) *Biochemistry* **23**, 1774–1779
294. González, D. H., and Andreo, C. S. (1989) *Trends Biochem. Sci.* **14**, 24–27
295. Knowles, J. R. (1989) *Ann. Rev. Biochem.* **58**, 195–221
296. Janc, J. W., Urbauer, J. L., O'Leary, M. H., and Cleland, W. W. (1992) *Biochemistry* **31**, 6432–6440
297. Lee, S.-L., Hepburn, T. W., Swartz, W. H., Ammon, H. L., Mariano, P. S., and Dunaway-Matiano, D. (1992) *J. Am. Chem. Soc.* **114**, 7346–7354
298. McQueney, M. S., Lee, S.-l, Bowman, E., Mariano, P. S., and Dunaway-Mariano, D. (1989) *J. Am. Chem. Soc.* **111**, 6885–6887
299. Seidel, H. M., and Knowles, J. R. (1994) *Biochemistry* **33**, 5641–5646
300. Bowman, E., McQueney, M., Barry, R. J., and Dunaway-Mariano, D. (1988) *J. Am. Chem. Soc.* **110**, 5575–5576
301. Hidaka, T., and Seto, H. (1989) *J. Am. Chem. Soc.* **111**, 8012–8013
302. Freeman, S., Pollack, S. J., and Knowles, J. R. (1992) *J. Am. Chem. Soc.* **114**, 377–378
302a. Morais, M. C., Zhang, W., Baker, A. S., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2000) *Biochemistry* **39**, 10385–10396
303. Mortimer, C. E., and Niehaus, W. G., Jr. (1974) *J. Biol. Chem.* **249**, 2833–2842
304. Clifford, K., Cornforth, J. W., Mallaby, R., and Phillips, G. T. (1971) *J. Chem. Soc. Chem. Commun.*, 1599–1600
305. Reardan, J. E., and Abeles, R. H. (1986) *Biochemistry* **25**, 5609–5616
306. Poulter, C. D., Muehlbacker, M., and Davis, D. R. (1989) *J. Am. Chem. Soc.* **111**, 3740–3742
307. Street, I. P., Coffman, H. R., Baker, J. A., and Poulter, C. D. (1994) *Biochemistry* **33**, 4212–4217
307a. Oh, S. K., Han, K. H., Ryu, S. B., and Kang, H. (2000) *J. Biol. Chem.* **275**, 18482–18488
308. Cane, D. E., Abell, C., Harrison, P. H., Hubbard, B. R., Kane, C. T., Lattman, R., Oliver, J. S., and Weiner, S. W. (1991) *Philos Trans R Soc Lond B Biol Sci* **332**, 123–129
309. Cane, D. E., Pawlak, J. L., and Horak, R. M. (1990) *Biochemistry* **29**, 5476–5490
310. Aulabaugh, A., and Schloss, J. V. (1990) *Biochemistry* **29**, 2824–2830
311. Biou, V., Dumas, R., Cohen-Addad, C., Douce, R., Job, D., and Pebay-Peyroula, E. (1997) *EMBO J.* **16**, 3405–3415
312. Halgand, F., Dumas, R., Biou, V., Andrieu, J.-P., Thomazeau, K., Gagnon, J., Douce, R., and Forest, E. (1999) *Biochemistry* **38**, 6025–6034
313. Proust-De Martin, F., Dumas, R., and Field, M. J. (2000) *J. Am. Chem. Soc.* **122**, 7688–7697

Study Questions

1. Discuss the role of carbonyl groups in facilitating reactions of metabolism.

2. Write a step-by-step sequence showing the chemical mechanisms involved in the action of a type I aldolase that catalyzes cleavage of fructose 1,6-bisphosphate. The enzyme is inactivated by sodium borohydride in the presence of the substrate. Explain this inactivation.

3. Some methylotrophic bacteria dehydrogenate methanol to formaldehyde. The latter undergoes an aldol condensation to form a hexulose-6-phosphate:



$$\begin{array}{c} CH_2OH \\ | \\ HC-OH \\ | \\ C=O \\ | \\ (HC-OH)_2 \\ | \\ CH_2OP \end{array}$$

   Write a reasonable cycle of biochemical reactions (utilizing this compound) by which three molecules of methanol can yield one molecule of glyceraldehyde 3-phosphate, a compound that can be either catabolized for energy or utilized for biosynthesis.

4. Malate is formed in the glyoxylate pathway by reaction of glyoxylate with Acetyl-CoA. Indicate the chemical mechanism of the reactions involved and structure of an intermediate species.

5. Acetyl-CoA is condensed with acetoacetyl-CoA to give an intermediate which is eventually converted to mevalonate.
   a) Show the sequence of reactions leading from acetyl-CoA to mevalonate with the structures of important intermediates.
   b) Mevalonate is an important intermediate on the pathway leading to what *class* of compounds?

6. a) Show the structures of the reactants for the 3-hydroxy-3-methylglutaryl-CoA synthase reaction.
   b) Free coenzyme A is liberated in the above reaction. From which of the two reactants did it come? Explain the metabolic significance of the liberation of free CoA.

7. Phosphoenolpyruvate carboxykinase catalyzes the following reaction:

   $$\text{Oxaloacetate} + \text{GTP} \rightarrow CO_2 + \text{phosphoenolpyruvate} + \text{GDP}$$

   Illustrate the probable mechanism of the reaction.

8. Write a complete step-by-step mechanism for the action of ATP-citrate lyase which catalyzes the following reaction (Eq. 13-39):

   $$\text{ATP} + \text{citrate} + \text{CoA}-\text{SH} \rightarrow \text{Acetyl}-\text{CoA} + \text{oxaloacetate} + \text{ADP} + P_i$$

9. Malonyl-CoA synthetase forms its product from free malonate and MgATP. Both phospho-enzyme and malonyl-enzyme intermediates have been detected. Suggest a sequential mechanism of action. See Kim, Y. S., and Lee, J. K. (1986) *J. Biol. Chem.* **261**, 16295–16297.

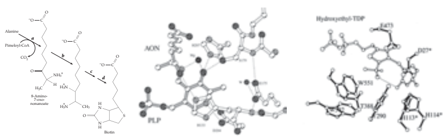
10. The carboxylation of acetone with HCO_3^- to form acetoacetate (the reverse of Eq. 13-44) is not a thermodynamically spontaneous process ($\Delta G°'$ ~17 kJ mol^{-1}). An anaerobic strain of *Xanthobacter* couples this reaction to the cleavage of ATP to AMP + 2 P_i. Suggest possible mechanisms. See Sluis, M. K., and Ensign, S. A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8456–8461.

11. The reaction of Eq. 10-11 begins with a hydroxylation, which can be viewed as insertion of an oxygen atom into a C–H bond. Draw the probable intermediate formed in this reaction and indicate how it can be converted to the final products.

12. Cyanase from *E. coli* decomposes cyanate (NCO^-) to ammonia and CO_2. Isotopic labeling studies show that the second oxygen atom in the product CO_2 comes not from H_2O but from a second substrate, bicarbonate ion. The correct equation for the enzymatic process is:

    $$NCO^- + H^+ + HCO_3^- \rightarrow NH_2-COO^- + CO_2$$

    Propose a sequence of steps for this enzymatic reaction, which has been investigated by Anderson and coworkers. Johnson, W. V., and Anderson, P. M. (1987) *J. Biol. Chem.* **262**, 9021–9025 and Anderson, P. M., Korte, J. J., Holcomb, T. A., Cho, Y.-g, Son, C.-m, and Sung, Y.-c. (1994) *J. Biol. Chem.* **269**, 15036–15045.



The vitamin **biotin** is formed in nature (left) by condensation of L-alanine with pimeloyl-CoA to form 8-amino-7-oxononanoate (AON). This compound is seen at the upper left of the center structure joined as a Schiff base with the coenzyme **pyridoxal phosphate** (PLP). This is a product complex of the enzyme AON synthase (see Webster *et al.*, *Biochemistry* **39**, 516-528, 2000) Courtesy of D. Alexeev, R. L. Baxter, and L. Sawyer. Biotin synthesis requires three other enzymes (steps *b*, *c*, *d*). Step *b* is catalyzed by a PLP-dependent transaminase. At the left is **thiamin diphosphate**, in the form of its 2-(1-hydroxyethyl) derivative, an intermediate in the enzyme pyruvate decarboxylase (Dobritzsch *et al.*, *J. Biol. Chem.* **273**, 20196-20204, 1998). Courtesy of Guoguang Lu. Thiamin diphosphate functions in all living organisms to cleave C–C bonds adjacent to C=O groups.

Contents

719 .... **A. ATP and the Nucleotide "Handles"**
720 .... **B. Coenzyme A and Phosphopantetheine**
723 .... **C. Biotin and the Formation of Carboxyl Groups from Bicarbonate**
724 .... 1. Biotin-Containing Enzymes
725 .... 2. The Mechanism of Biotin Action
725 .... *Carboxybiotin*
726 .... *Carboxyphosphate*
727 .... *The β carboxylation step*
729 .... 3. Control Mechanisms
729 .... 4. Pumping Ions with the Help of Biotin
730 .... **D. Thiamin Diphosphate**
730 .... 1. Chemical Properties of Thiamin
731 .... 2. Catalytic Mechanisms
733 .... 3. Structures of Thiamin-Dependent Enzymes
734 .... 4. The Variety of Enzymatic Reactions Involving Thiamin
736 .... 5. Oxidative Decarboxylation and 2-Acetylthiamin Diphosphate
736 .... 6. Thiamin Coenzymes in Nerve Action
737 .... **E. Pyridoxal Phosphate**
737 .... 1. Nonenzymatic Models
740 .... 2. A General Mechanism of Action of PLP
741 .... 3. The Variety of PLP-Dependent Reactions
741 .... *Loss of the α-Hydrogen* (Group *a*)
744 .... *Decarboxylation* (Group *b*)
745 .... *Side chain cleavage* (Group *c*)
745 .... *Ketimine intermediate as electron acceptor* (Group *d*)
746 .... *Glycogen phosphorylase*
747 .... 4. Pyridoxamine Phophate as a Coenzyme
747 .... 5. Stereochemistry of PLP-Requiring Enzymes
749 .... 6. Seeing Changes in the Optical Properties of the Coenzyme
750 .... *Imine groups in free enzymes*
750 .... *Absorption bands at 500 nm*
750 .... 7. Atomic Structures
751 .... 8. Constructing a Detailed Picture of the Action of a PLP Enzyme
753 .... **F. Pyruvoyl Groups and Other Unusual Electrophilic Centers**
754 .... 1. Decarboxylases
755 .... 2. Proline and Glycine Reductases
755 .... 3. Dehydroalanine and Histidine and Phenylalanine Ammonia-Lyases

758 .... **References**
763 .... **Study Questions**

Boxes

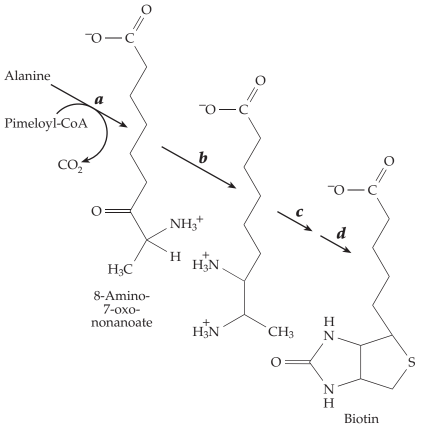
721 .... Box 14-A Discovery of the Vitamins
728 .... Box 14-B The Biotin-Binding Proteins Avidin and Streptavidin
738 .... Box 14-C The Vitamin B_6 Family: Pyridoxine, Pyridoxal, and Pyridoxamine
756 .... Box 14-D Dietary Requirements for B Vitamins

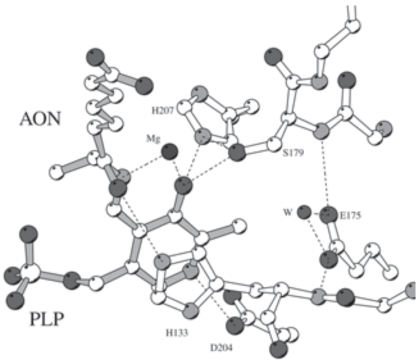
Tables

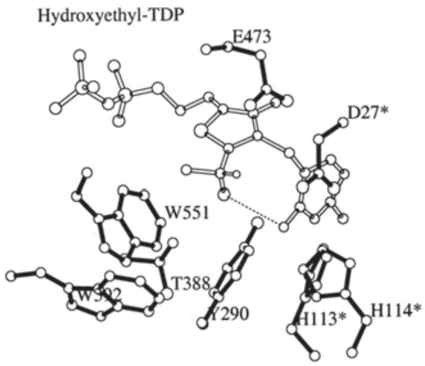
725 .... Table 14-1 Enzymes Containing Bound Biotin
735 .... Table 14-2 Enzymes Dependent upon Thiamin Diphosphate as a Coenzyme
743 .... Table 14-3 Some Enzymes That Require Pyridoxal Phosphate as a Coenzyme
753 .... Table 14-4 Some Pyruvoyl Enzymes

Coenzymes: Nature’s Special Reagents

14







Most of the reactions discussed in Chapters 12 and 13 are catalyzed by enzymes that contain only those functional groups found in the side chains of the constituent amino acids. **Coenzymes** are nonprotein molecules that function as essential parts of enzymes. Coenzymes often serve as “special reagents” needed for reactions that would be difficult or impossible using only simple acid–base catalysis. In many instances, they also serve as **carriers**, alternating catalysts that accept and donate chemical groups, hydrogen atoms, or electrons. Coenzymes will be considered here in three groups:

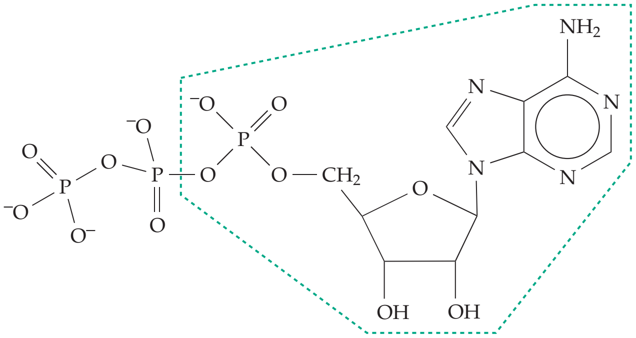
1. Compounds of high group transfer potential such as ATP and GTP that function in energy coupling within cells. Because it is cleaved and then dissociates from the enzyme to which it is bound, ATP may be regarded as a substrate rather than a coenzyme. However, as a phosphorylated form of ADP it may also be viewed as a carrier of high-energy phospho groups.

2. Compounds, often derivatives of **vitamins** that, while in the active site of the enzyme, alter the structure of a substrate in a way that permits it to react more readily. Coenzyme A, pyridoxal phosphate, thiamin diphosphate, and vitamin B_{12} coenzymes fall into this group.

3. Oxidative coenzymes with structures of precisely determined oxidation–reduction potential. Examples are NAD^+, $NADP^+$, FAD, and lipoic acid. They serve as carriers of hydrogen atoms or of electrons. Some of these coenzymes, such as NAD^+ and $NADP^+$, can usually dissociate rapidly and reversibly from the enzymes with which they function. Others, including FAD, are much more tightly bound and rarely if ever dissociate from the protein catalyst. Heme groups are covalently linked to proteins such as cytochrome *c* and cannot be dissociated without destroying the enzyme. Very tightly bound coenzyme groups are often called **prosthetic groups**, but there is no sharp line that divides prosthetic groups from the loosely bound coenzymes. For example, NAD^+ is bound weakly to some proteins but tightly to others. Oxidative coenzymes are discussed in Chapter 15.

A. ATP and the Nucleotide “Handles”

The role of ATP in “driving” biosynthetic reactions has been considered in Chapter 12, where attention was focused on the polyphosphate group which undergoes cleavage. What about the adenosine end? Here is a shapely structure borrowed from the nucleic acids. What is it doing as a carrier of phospho groups? At least part of the answer seems to be that the adenosine monophosphate (AMP) portion of the molecule is a “handle” which can be “grasped” by catalytic proteins. For some enzymes, such as acetyl-CoA synthetase (Eq. 12-45), the handle is important because the intermediate acyl adenylate must remain bound to the protein. Without the large adenosine group, there would be little for the protein to hold onto.

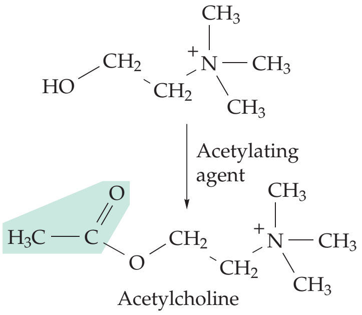

The AMP "handle" in ATP

AMP is only one of several handles to which nature attaches phospho groups to form di- and triphosphate derivatives. Like AMP, the other handles are nucleotides, the monomer units of nucleic acids. Thus, one enzyme requiring a polyphosphate as an energy source selects ATP, another CTP, or GTP. The nucleotide handles not only carry polyphosphate groups but also are present in other coenzymes, such as CoA, NAD$^+$, NADP$^+$, and FAD. In addition, they serve as carriers for small organic molecules. For example, **uridine diphosphate glucose** (UDP-Glc) Chapter 10, is a carrier of active glucosyl groups important in sugar metabolism and **cytidine diphosphate choline** is an intermediate in synthesis of phospholipids.

Recalling that acetyl adenylate (acetyl-AMP) is an intermediate in synthesis of acetyl-CoA, and comparing the biosynthesis of sugars, phospholipids, and acetyl-CoA, we see that in each case the enzyme involved requires a different nucleotide handle. The handle may provide a means of recognition which can help an enzyme to pick the right bit of raw material out of the sea of molecules surrounding it. Figure 5-6 shows the shapes of the four purine and pyrimidine bases forming the most common nucleotide handles. The distinctive differences both in shape and in hydrogen bond patterns are obvious. In binding to proteins, the hydrogen bond-forming groups in the purine and pyrimidine bases sometimes interact with precisely positioned groups in the protein. However, in some enzymes the "handle" is not precisely bound. This seems to be the case for adenine, which makes surprisingly few hydrogen bonds to proteins.[1] Hydroxyl groups of the ribose or deoxyribose ring also often form hydrogen bonds and the negatively charged oxygen atoms of the 5'-phosphate may interact with positively charged protein side chains.

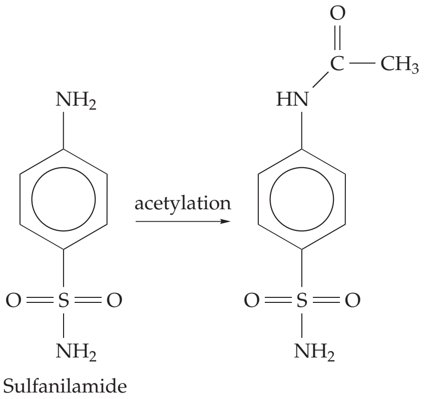
B. Coenzyme A and Phosphopantetheine

The existence of a special coenzyme required in biological acetylation was recognized by Fritz Lipmann in 1945.[2–3a] The joining of acetyl groups to other molecules is a commonplace reaction within living cells, one example being the formation of the neurotransmitter **acetylcholine**. In the laboratory acetylation is carried out with reactive compounds such as acetic anhydride or acetyl chloride.



(14-1)

Lipmann wondered what nature used in their place. His approach in seeking the biological "active acetate" is one that has been used successfully in solving many biochemical problems. He first set up a test system to examine the ability of extracts prepared from fresh liver tissue to catalyze the acetylation of sulfanilamide (Eq. 14-2). A specific color test was available for quantitative determination of very small amounts of the product. The rate of acetylation of sulfanilamide under standard conditions was taken as a measure of the activity of the biochemical acetylation system. Lipmann soon discovered that the reaction required ATP and that the ATP was cleaved to ADP concurrently with the formation of acetyl-sulfanilamide. He also found that dialysis or ultrafiltration rendered the liver extract almost inactive in acetylation. Apparently some essential material passed out through the semipermeable dialysis membrane. When the dialysate or ultrafiltrate was concentrated and added back, acetylation activity was restored. The unknown material was not destroyed by boiling, and Lipmann postulated that it was a new coenzyme which he called **coenzyme A** (CoA). Now the test system was used to estimate the amount of the



(14-2)

BOX 14-A DISCOVERY OF THE VITAMINS

Several mysterious and often fatal diseases which resulted from vitamin deficiencies were prevalent until the past century. Sailors on long sea voyages were often the victims. In the Orient, the disease **beriberi** was rampant and millions died of its strange paralysis called “polyneuritis.” In 1840, George Budd predicted that the disease was caused by the lack of some organic compound that would be discovered in “a not too distant future.”[a] In 1893, C. Eijkman, a Dutch physician working in Indonesia, observed paralysis in chicks fed white rice consumed by the local populace. He found that the paralysis could be relieved promptly by feeding an extract of rice polishings. This was one of the pieces of evidence that led the Polish biochemist Casimir Funk to formulate the “vitamine theory” in about 1912. Funk suggested that the diseases **beriberi**, **pellagra**, **rickets**, and **scurvy** resulted from lack in the diet of four different vital nutrients. He imagined them all to be amines, hence the name **vitamine**.

In the same year in England, F. G. Hopkins announced that he had fed rats on purified diets and discovered that amazingly small amounts of **accessory growth factors**, which could be obtained from milk, were necessary for normal growth.[b] By 1915, E. V. McCollum and M. Davis at the University of Wisconsin had recognized that rat growth depended on not one but at least two accessory factors. The first, soluble in fatty solvents, they called **A**, and the other, soluble in water, they designated **B**. Factor B cured beriberi in chicks. Later, when it was shown that vitamine A was not an amine, the “e” was dropped and **vitamin** became a general term.

Progress in isolation of the vitamins was slow, principally because of a lack of interest. According to R. R. Williams, when he started his work on isolation of the antiberiberi factor in 1910 most people were convinced that his efforts were doomed to failure, so ingrained was Pasteur’s idea that diseases were caused only by bacteria. In 1926, Jansen isolated a small amount of thiamin, but it was not until 1933 that Williams, working almost without financial support, succeeded in preparing a large amount of the crystalline compound from rice polishings. Characterization and synthesis followed rapidly.[c,d]

It was soon apparent that the new vitamin alone would not satisfy the dietary need of rats for the B factor. A second thermostable factor (B_2) was required in addition to thiamin (B_1), which was labile and easily destroyed by heating. When it became clear that factor B_2 contained more than one component, it was called **vitamin B complex**. There was some confusion until relatively specific animal tests for each one of the members had been devised. **Riboflavin** was found to be most responsible for the stimulation of rat growth, while **vitamin B_6** was needed to prevent a facial dermatitis or “rat pellagra.” **Pantothenic acid** was especially effective in curing a chick dermatitis, while **nicotinamide** was required to cure human pellagra. **Biotin** was required for growth of yeast.

The antiscurvy (antiscorbutic) activity was called **vitamin C**, and when its structure became known it was called **ascorbic acid**. The fat-soluble factor preventing rickets was designated **vitamin D**. By 1922, it was recognized that another fat-soluble factor, **vitamin E**, is essential for full-term pregnancy in the rat. In the early 1930s **vitamin K** and the **essential fatty acids** were added to the list of fat-soluble vitamins. Study of the human blood disorders “tropical macrocytic anemia” and “pernicious anemia” led to recognition of two more water-soluble vitamins, **folic acid** and **vitamin B_{12}**. The latter is required in minute amounts and was not isolated until 1948. Have all the vitamins been discovered? Rats can be reared on an almost completely synthetic diet. However, there is the possibility that for good health humans require some as yet undiscovered compounds in our diet. Furthermore, it is quite likely that we receive some essential nutrients that we cannot synthesize from bacteria in our intestinal tracts. An example may be the pyrroloquinoline quinone (PQQ).[e]

Why do we need vitamins? Early clues came in 1935 when nicotinamide was found in NAD^+ by H. von Euler and associates and in $NADP^+$ by Warburg and Christian. Two years later, K. Lohman and P. Schuster isolated pure **cocarboxylase**, a dialyzable material required for decarboxylation of pyruvate by an enzyme from yeast. It was shown to be thiamin diphosphate (Fig. 15-3). Most of the water-soluble vitamins are converted into coenzymes or are covalently bound into active sites of enzymes. Some lipid-soluble vitamins have similar functions but others, such as vitamin D and some metabolites of vitamin A, act more like hormones, binding to receptors that control gene expression or other aspects of metabolism.

---

[a] Hughes, R. E. (1973) *Medical History* **17**, 127–134

[b] Harris, L. D. (1937) *Vitamins in Theory and Practice*, Cambridge Univ. Press, London and New York

[c] Karlson, P. (1984) *Trends Biochem. Sci.* **9**, 536–537

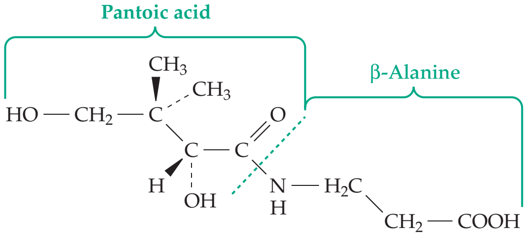
[d] Williams, R. R., and Spies, T. D. (1938) *Vitamin B_1 and Its Use in Medicine*, Macmillan, New York

[e] Killgore, J., Smidt, C., Duich, L., Romero-Chapman, N., Tinker, D., Reiser, K., Melko, M., Hyde, D., and Rucker, R. B. (1989) *Science* **245**, 850–852

coenzyme in a given volume of dialysate or in any other sample. When small amounts of CoA were supplied to the test system, only partial restoration of the acetylation activity was observed and the amount of restoration was proportional to the amount of CoA. With test system in hand to monitor various fractionation methods, Lipmann soon isolated the new coenzyme in pure form from yeast and liver.

Coenzyme A (Fig. 14-1) is a surprisingly complex molecule. The handle is AMP with an extra phospho group on its 3'-hydroxyl. The phosphate of the 5'-carbon is linked in anhydride (pyrophosphate) linkage to another phosphoric acid, which is in turn esterified with **pantoic acid**. Pantoic acid is linked to **β-alanine** and the latter to **β-mercaptoethylamine** through amide linkages, the reactive SH group being attached to a long (1.9-nm) semi-flexible chain. Coenzyme A can be cleaved by hydrolysis to **pantetheine** (Fig. 14-1), **pantetheine 4'-phosphate**, and **pantothenic acid**. These three compounds are all **growth factors**.

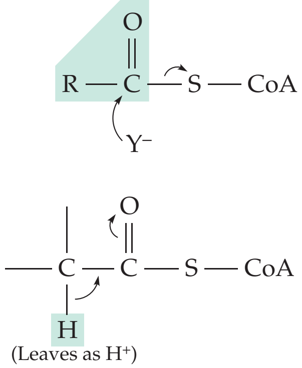
Pantothenic acid is a vitamin, which is essential to human life. Its name is derived from a Greek root that reflects its universal occurrence in living things. The bacterium *Lactobacillus bulgaricus*, which converts milk to yogurt, needs the more complex pantetheine for growth. It finds a ready supply of pantetheine in milk and has lost its ability to synthesize this compound. However, it can convert pantetheine to CoA. Pantetheine 4'-phosphate is required for growth of *Acetobacter suboxydans*. While it is not a dietary essential for most organisms, it is found in a covalently bound form in several enzymes.


Pantoic acid
β-Alanine
CH3
CH3
HO — CH2 — C
O
C — C
H
OH
N — H2C
H
CH2 — COOH

Pantothenic acid

While CoA was discovered as the "acetylation coenzyme," it has a far more general function. It is required, in the form of acetyl-CoA, to catalyze the synthesis of citrate in the citric acid cycle. It is essential to the β oxidation of fatty acids and carries propionyl and other acyl groups in a great variety of other metabolic reactions. About 4% of all known enzymes require CoA or one of its esters as a substrate.[4]

Coenzyme A has two distinctly different biochemical functions, which have already been considered briefly in Chapters 12 and 13 and can be summarized as follows:


O
R — C — S — CoA
Y⁻
O
— C — C — S — CoA
H
(Leaves as H+)

1. Activation of acyl group, $R-\overset{O}{\overset{\|}{C}}-$, toward transfer by nucleophilic displacment
2. Activation of a hydrogen atom adjacent to the carbonyl of a thioester

These functions depend, to a considerable extent, on the fact that the properties of the carbonyl group of a thioester are closely similar to those of an isolated carbonyl group in a ketone.[5]

Synthesis of fatty acids in bacteria requires a small **acyl carrier protein (ACP)** whose functions are similar to those of CoA. However, it contains pante-

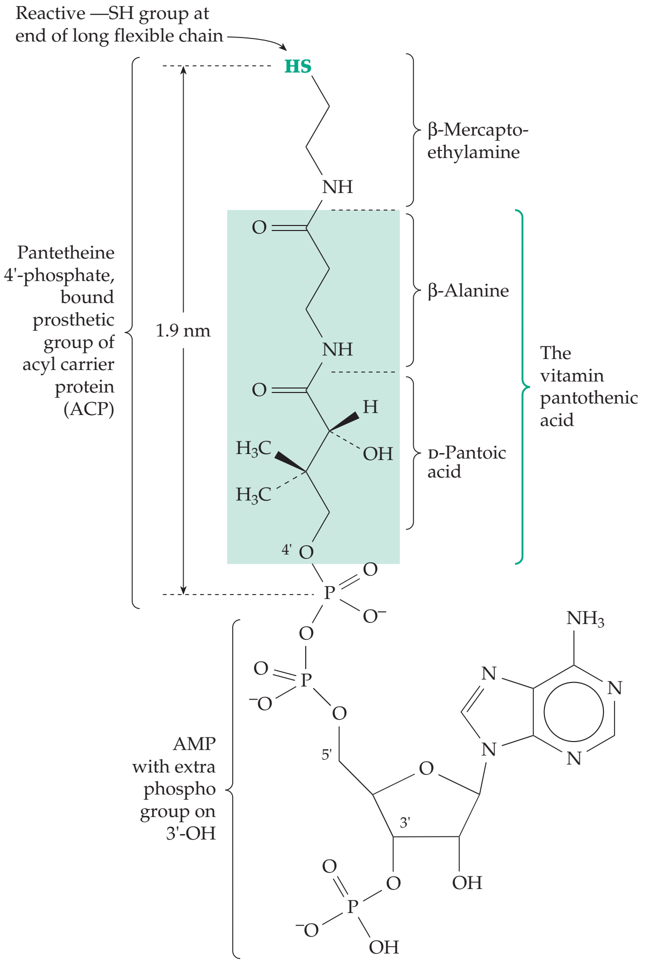
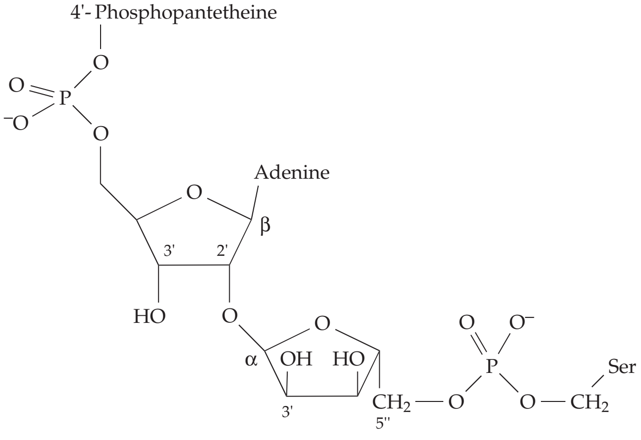

Reactive —SH group at end of long flexible chain
HS
β-Mercapto-ethylamine
NH
O
β-Alanine
NH
O
H
OH
H3C
H3C
D-Pantoic acid
4'
O
P
O
O⁻
The vitamin pantothenic acid
Pantetheine 4'-phosphate, bound prosthetic group of acyl carrier protein (ACP)
1.9 nm
O
O
P
⁻O
O
NH3
N
N
N
N
5'
O
3'
AMP with extra phospho group on 3'-OH
O
O
OH
P
⁻O
OH

Figure 14-1 Coenzyme A, an acyl-activating coenzyme containing the vitamin pantothenic acid.

theine 4'-phosphate covalently bonded through phosphodiester linkage to a serine side chain. In *E. coli* this is at position 36 in the 77-residue protein.[6,7] Here the nucleotide handle of CoA has been replaced with a larger and more complex protein which can interact in specific ways with the multiprotein fatty acid synthase complex described in Chapter 21. In higher organisms ACP is usually not a separate protein but a domain in a large synthase molecule. Bound phosphopantetheine is also found in enzymes involved in synthesis of peptide antibiotics (Chapter 29)[8,9] and polyketides (Chapter 21).[10,11] It is also present in subunits of cytochrome oxidase and of ATP synthase of *Neurospora* but appears to play only a structural role, being needed for proper assembly of these multimeric proteins.[12] In a citrate-cleaving enzyme, and in a bacterial malonate decarboxylase, phosphopantethine is attached to a serine side chain as 2′-(5′′-phosphoribosyl)-3′-dephospho-CoA.[13–14a]



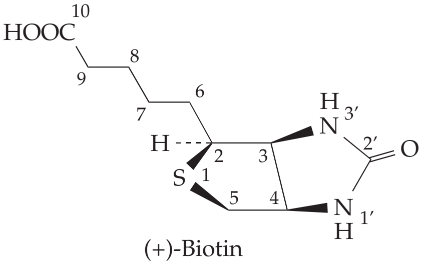
2'-(5''-Phophoribosyl)-3'-dephospho-CoA

Attachment of phosphopantetheine to proteins is catalyzed by a phosphotransferase that utilizes CoA as the donor. A phosphodiesterase removes the phosphopantetheine, providing a turnover cycle.[15–15b] A variety of synthetic analogs have been made.[4,16] The reactive center of CoA and phosphopantetheine is the SH group, which is carried on a flexible arm that consists in part of the β-alanine portion of pantothenic acid. A mystery is why pantoic acid, a small odd-shaped molecule that the human body cannot make, is so essential for life. The hydroxyl group is a potential reactive site and the two methyl groups may enter into formation of a “trialkyl lock” (p. 485), part of a sophisticated “elbow” or shoulder for the SH-bearing arm.

When it binds to citrate synthase acetyl-CoA appears to bind only after the enzyme has undergone a conformational change that closes the enzyme around its other substrate oxaloacetate (Fig. 13-8). The adenine ring of the long CoA handle is tightly bonded to the protein through hydrophobic interactions and hydrogen-bond recognition interactions. Additional specific hydrophobic interactions allow the enzyme to hold the acetyl group in a precise position where it can be acted upon by the catalytic groups. In the case of CoA transferase (Eq. 12-50) smaller thiols can replace CoA but with much lower catalytic rates. The acyl-CoA derivatives are weakly bound, the expected intrinsic binding energy of the pantetheine portion of the molecule apparently being used to increase k_{cat}.[17] From a study of kinetics and equilibria it was concluded that the binding energy of the interaction of the nucleotide portion of coenzyme A with the enzyme is utilized to increase the rate of formation and to stabilize the covalently linked E-CoA (see Eq. 12-50). On the other hand, binding of the pantoic acid part of the molecule decreases the stability of the transition state for breakup of the complex by ~40 kJ/mol, which corresponds to a 10^7-fold increase in the reaction rate.[16]

C. Biotin and the Formation of Carboxyl Groups from Bicarbonate

By 1901 it was recognized that yeast required for its growth an unknown material which was called **bios**. This was eventually found to be a mixture of pantothenic acid, inositol, and a third component which was named **biotin**. Biotin was also recognized as a factor promoting growth and respiration of the clover root nodule organism *Rhizobium trifolii* and as vitamin H, a material that prevented dermatitis and paralysis in rats that were fed large amounts of uncooked egg white. Isolation of the pure vitamin was a heroic task accomplished by Kögl in 1935. In one preparation 250 kg of dried egg yolk yielded only 1.1 mg of crystalline biotin.[18]

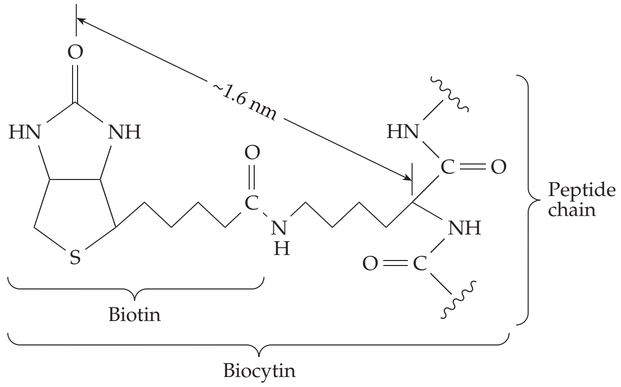


(+)-Biotin

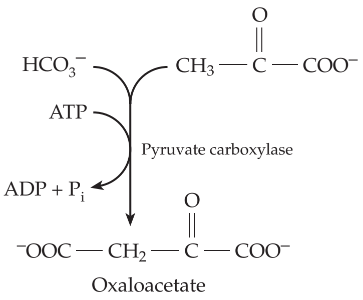
Biotin contains three chiral centers and therefore has eight stereoisomers.[18,19] Of these, only one, the dextrorotatory (+)-biotin, is biologically active.[19,20] The vitamin is readily oxidized to the sulfoxide and sulfone. The sulfoxide can be reduced back to biotin by a molybdenum-containing reductase in some bacteria (see also Chapter 16, Section H).[20a] Biotin is synthesized from pimeloyl-CoA (see chapter banner, p. 719 and Eq. 24-39). Four enzymes are required. Two of them, a synthase that catalyzes step *a* (banner) and a transaminase that catalyzes step *b*, contain the coenzyme pyridoxal phosphate (PLP).[20b] The final step is insertion of a sulfur atom from an iron-sulfur center.

1. Biotin-Containing Enzymes

Within cells the biotin is covalently bonded to proteins, its double ring being attached to a 1.6-nm flexible arm. The first clue to this fact was obtained from isolation of a biotin-containing material **biocytin**, ε-N-biotinyl-L-lysine, from autolysates (self-digests) of rapidly growing yeast.[21] It was subsequently shown that the lysine residue of the biocytin was originally present in proteins at the active sites of biotin-containing enzymes usually within the sequence AMKM[22] or VMKM.[23] Other conserved features also mark the attachment site.[24]

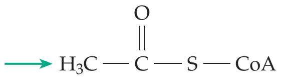


Biotin acts as a **carboxyl group carrier** in a series of carboxylation reactions, a function originally suggested by the fact that aspartate partially replaces biotin in promoting the growth of the yeast *Torula cremonis*. Aspartate was known to arise by transamination from oxaloacetate, which in turn could be formed by carboxylation of pyruvate. Subsequent studies showed that biotin was needed for an enzymatic ATP-dependent reaction of pyruvate with bicarbonate ion to form oxaloacetate (Eq. 14-3). This is a β carboxylation coupled to the hydrolysis of ATP.

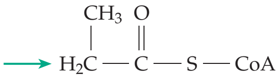


(14-3)

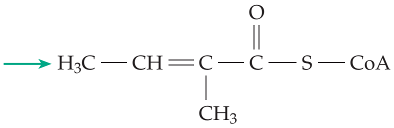
In addition to pyruvate carboxylase, other biotin-requiring enzymes act on **acetyl-CoA**, **propionyl-CoA**, and **β-methylcrotonyl-CoA**, using HCO_3^- to add carboxyl groups at the sites indicated by the arrows in the accompanying structures. Because of the presence of the C=C double bond conjugated with the carbonyl group, the carboxylation of β-methylcrotonyl-CoA is electronically analogous to β carboxylation.


$$\rightarrow H_3C-\overset{\overset{\displaystyle O}{\|}}{C}-S-CoA$$

Acetyl-CoA


$$\rightarrow H_2C-\overset{\overset{\displaystyle O}{\|}}{C}-S-CoA$$

Propionyl-CoA


$$\rightarrow H_3C-CH=\underset{\underset{\displaystyle CH_3}{|}}{C}-\overset{\overset{\displaystyle O}{\|}}{C}-S-CoA$$

β-Methylcrotonyl-CoA

Human cells, as well as those of other higher eukaryotes, carry genes for all four of these enzymes.[25–29] Acetyl-CoA carboxylase is a cytosolic enzyme needed for synthesis of fatty acids but the other three enzymes enter mitochondria when they function. The biotin-dependent carboxylases, which are listed in Table 14-1, have a variety of molecular sizes and subunits but show much evolutionary conservation in their sequences and chemical mechanisms.[22,25] In higher eukaryotes all of the catalytic apparatus of the enzymes is present in single large 190- to 200-kDa subunits. The 251-kDa subunit of yeast acetyl-CoA carboxylase consists of 2337 amino acid residues.[30] That of rats contains 2345[31] and that of the alga *Cyclotella* 2089.[32] Human cytosolic acetyl-Co carboxylase has 2347 residues while a mitochondria-associated form has an extra 136 residues, most of them in a hydrophobic N-terminal extension.[32a,b] Plants have two forms of the enzyme, one cytosolic and one located in plastids. In wheat, they have 2260 and 2311 residues, respectively.[32c,32d] Animal and fungal pyruvate carboxylases are also large ~500-kDa tetramers.[32e] The yeast enzyme consists of 1178-residue monomers.[33] In contrast, the 560-kDa human propionyl-CoA carboxylase is an $\alpha_4\beta_4$ tetramer.[34,35]

In bacteria and in at least some plant chloroplasts,[36,37] acetyl-CoA carboxylase consists of three different kinds of subunit and four different peptide chains. The much studied *E. coli* enzyme is composed of a 156-residue **biotin carboxyl carrier protein**,[38] a 449-residue **biotin carboxylase**, whose three-dimensional structure in known,[39,39a] and a **carboxyltransferase** subunit consisting of 304 (α)- and 319 (β)- residue chains. These all associate as a dimer of the three subunits (eight peptide chains).[40–42]

Biotin becomes attached to the proper ε-amino groups at the active centers of biotin enzymes by the action of **biotin holoenzyme synthetase** (biotinyl protein ligase), which utilizes ATP to form an intermediate biotinyl-AMP.[43–47a] Hereditary deficiency of this enzyme has been observed in a few children and has been treated by administration of extra biotin.[48] The *E. coli* biotin holoenzyme synthetase, whose three-dimensional structure is known, has a dual function. It is also a repressor of transcription of the biotin biosynthetic operon.[45] Intracellular degradation of biotin-containing proteins yields biotin-containing oligopeptides as well as biocytin. These are acted on by **biotinidase** to release free biotin.[49,50] The action of this enzyme in recycling biotin may be a controlling factor in the rate of formation of new biotin-dependent enzymes.

2. The Mechanism of Biotin Action

It may seem surprising that a coenzyme is needed for these carboxylation reactions. However, unless the cleavage of ATP were coupled to the reactions, the equilibria would lie far in the direction of decarboxylation. For example, the measured apparent equilibrium constant K' for conversion of propionyl-CoA to S methylmalonyl-CoA at pH 8.1 and 28°C[51] is given by Eq. 14-4.

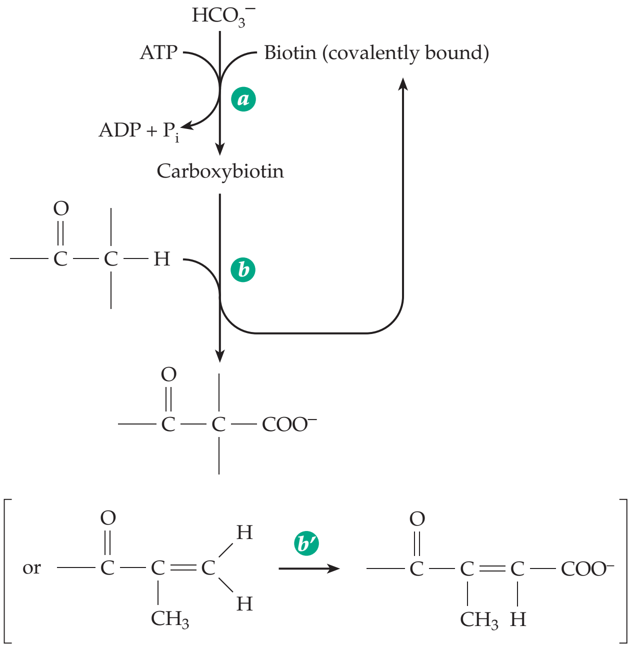
$$K' = \frac{[\mathrm{ADP}][\mathrm{P_i}][\mathrm{methylmalonyl{-}CoA}]}{[\mathrm{ATP}][\mathrm{HCO_3^-}][\mathrm{propionyl{-}CoA}]} = 5.7$$

$$\Delta G' = -\ 4.36\ \mathrm{kJ\ mol^{-1}} \qquad (14\text{-}4)$$

TABLE 14-1
Enzymes Containing Bound Biotin

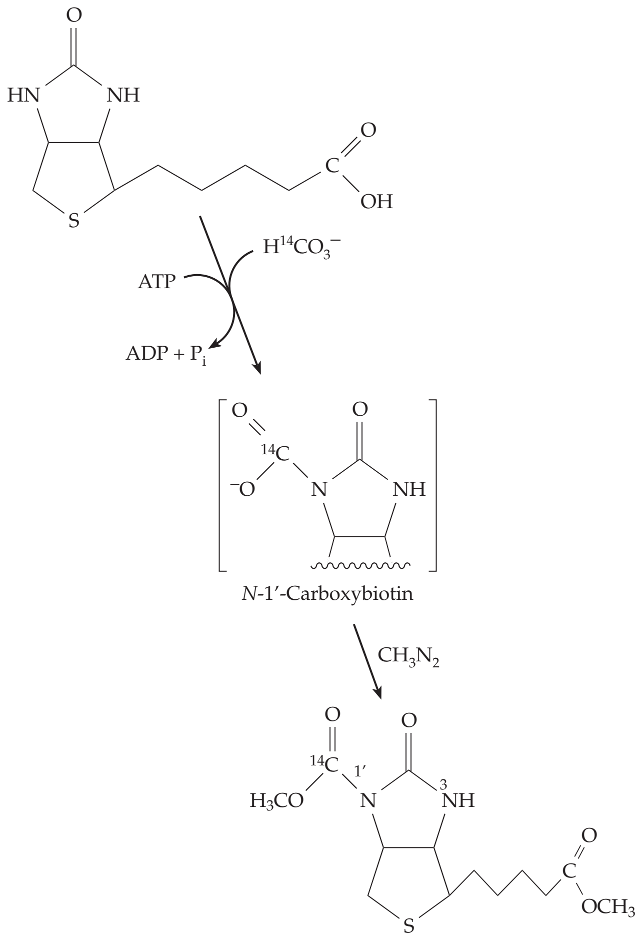
1. Catalyzing beta carboxylation using HCO_3^- with coupled cleavage of ATP to ADP + P_i
    - Acetyl-CoA carboxylase
    - Propionyl-CoA carboxylase
    - Pyruvate carboxylase
    - β-Methylcrotonyl-CoA carboxylase (δ carboxylation)
2. Carboxyl group transfer without cleavage of ATP
    - Carboxyltransferase of *Propionobacterium*
3. Biotin-dependent Na^+ pumps
    - Oxaloacetate decarboxylase
    - Methylmalonyl-CoA decarboxylase
    - Glutaconyl-CoA decarboxylase
4. Other
    - Malonate decarboxylase
    - Urea carboxylase

The function of biotin is to mediate the coupling of ATP cleavage to the carboxylation, making the overall reaction exergonic. This is accomplished by a two-stage process in which a **carboxybiotin** intermediate is formed (Eq. 14-5). There is *one known biotin-containing enzyme that does not utilize ATP*. Propionic acid bacteria contain a **carboxyltransferase** which transfers a carboxyl group reversibly from methylmalonyl-CoA to pyruvate to form oxaloacetate and propionyl-CoA (see Fig. 17-10). This huge enzyme consists of a central hexameric core of large 12S subunits to which six 5S dimeric subunits and twelve 123-residue biotinylated peptides are attached.[25,52] No ATP is needed because free HCO_3^- is not a substrate. However, biotin serves as the carboxyl group carrier in this enzyme too.



(14-5)

Carboxybiotin. The structure of biotin suggested that bicarbonate might be incorporated reversibly into its position 2'. However, this proved not to be true and it remained for F. Lynen and associates to obtain a clue from a "model reaction." They showed that purified β-methylcrotonyl-CoA carboxylase promoted the carboxylation of *free* biotin with bicarbonate ($H^{14}CO_3^-$) and ATP. While the carboxylated biotin was labile, treatment with diazomethane (Eq. 14-6) gave a stable dimethyl ester of ***N*-1'-carboxybiotin**.[53,54] The covalently bound biotin at active sites of enzymes was also successfully labeled with $^{14}CO_2$. Treatment of the labeled enzymes with diazomethane followed by hydrolysis with trypsin and pepsin gave authentic *N*-1'-carboxybiocytin. It was now clear that the cleavage of ATP is required to couple the CO_2 from HCO_3^- to the biotin to form carboxybiotin. The enzyme must

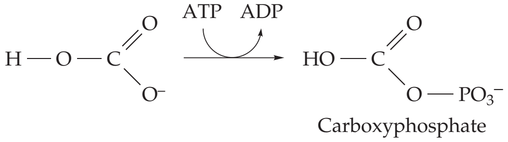


(14-6)

then transfer the carboxyl group from carboxybiotin to the substrate that is to be carboxylated. Enzymatic transfer of a carboxyl group from chemically synthesized carboxybiotin onto specific substrates confirmed the proposed mechanism.[55]

The biotin carboxyl carrier subunit of *E. coli* acetyl-CoA carboxylase contains the covalently bound biotin.[55a,b] The larger biotin carboxylase subunit catalyzes the ATP-dependent attachment of CO_2 to the biotin and the carboxyltransferase subunit catalyzes the final transcarboxylation step (Eq. 14-5, step *b*) by which acetyl-CoA is converted into malonyl-CoA. The biotin, which is attached to the carrier protein, is presumably able to move by means of its flexible arm from a site on the carboxylase to a site on the transcarboxylase.

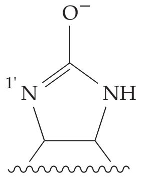
Carboxyphosphate. During the initial carboxylation step ^{18}O from labeled bicarbonate enters the P_i that is split from ATP. This suggested transient formation of **carboxyphosphate** by nucleophilic attack of HCO_3^- on ATP (Eq. 14-7). The carboxyl group of this reactive mixed anhydride[56] could then be transferred to biotin. This mechanism is supported by the fact that biotin carboxylase catalyzes the transfer of a



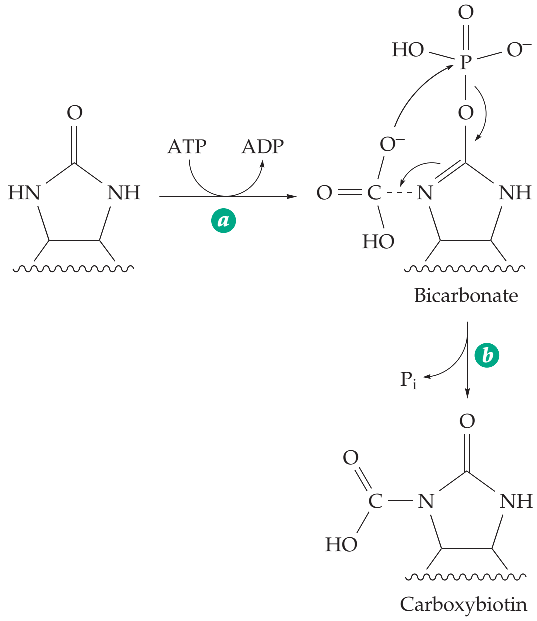
(14-7)

phospho group to ADP from carbamoyl phosphate, an analog of carboxyphosphate in a reaction that is analogous to the reverse of that in Eq. 14-7 and also by a slow bicarbonate-dependent ATPase activity that does not depend upon biotin.[57,58,58a]

The simplest mechanism for transfer of the carboxyl group of carboxyphosphate to biotin would appear to be nucleophilic displacement of the phosphate leaving group by N1' of biotin. The enzyme could presumably first catalyze removal of the N1' hydrogen to form a ureido anion.[59] Another reasonable possibility would be for the terminal phospho group of ATP to be transferred to biotin to form an *O*-phosphate[60,61] (Eq. 14-8, step *a*) which could react with bicarbonate as in Eq. 14-8, step *b*. Cleavage of the enol phosphate by attack of HCO_3^- would simultaneously create a nucleophilic center at N1' and carboxyphosphate ready to react with N1'. However, this could not easily explain the ATPase activity in the absence of biotin.

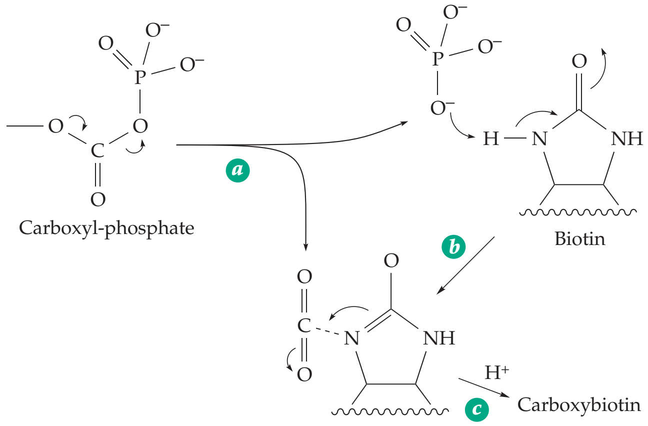


Ureido anion



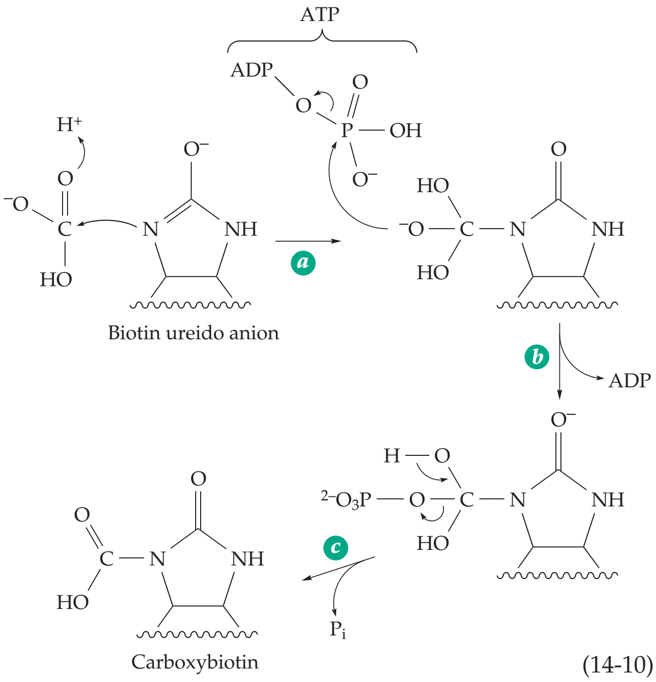
(14-8)

Either of the foregoing mechanisms requires that the ureido anion of biotin attack the rather unreactive carbon atom of carboxyphosphate. Another alternative, which is analogous to that suggested for PEP carboxylase (Eq. 13-53) is for carboxyphosphate to eliminate inorganic phosphate to give the more electrophilic CO_2 (Eq. 14-9, step *a*). The very basic inorganic phosphate trianion PO_4^{3-} that is eliminated could remove the proton from N1' of biotin to create the biotin ureido anion (step *b*), which could then add to CO_2 (step *c*).[22]



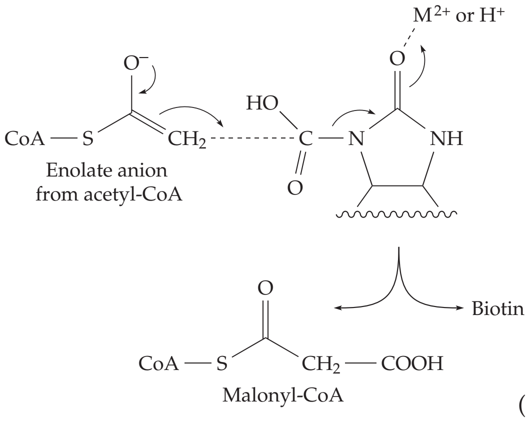
(14-9)

Have we checked all of the possibilities for the mechanism of biotin carboxylation? Kruger and associates[62,63] suggested that biotin, as a ureido anion, might add to bicarbonate to form a highly unstable intermediate which, however, could be phosphorylated by ATP (Eq. 14-10, steps *a* and *b*). This intermediate could undergo elimination of inorganic phosphate (step *c*), driving the reaction to completion. The observed transfer of isotope from ^{18}O-containing bicarbonate into ADP would be observed.



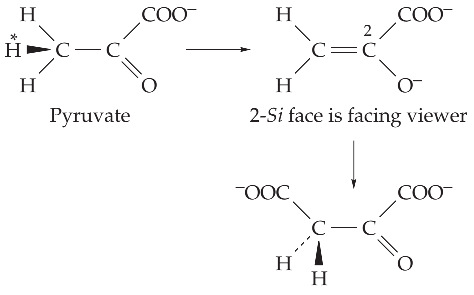
(14-10)

The β carboxylation step. Once formed, the carboxybiotin "head group" could swing to the carboxyltransferase site where transfer of the carboxyl group into the final product takes place. This might occur either by nucleophilic attack of an enolate anion on the carbonyl carbon (Eq. 14-11) or on CO_2 generated by reversal of the reactions of Eq. 14-9, steps *b* and *c*.



(14-11)

When pyruvate with a chiral methyl group is carboxylated by pyruvate carboxylase the configuration at C-3 is retained. The carboxyl enters from the 2-*si* side, the same side from which the proton (marked H*) was removed to form the enolate anion (Eq. 14-12). Comparable stereochemistry has been established for other biotin-dependent enzymes.[64,65]



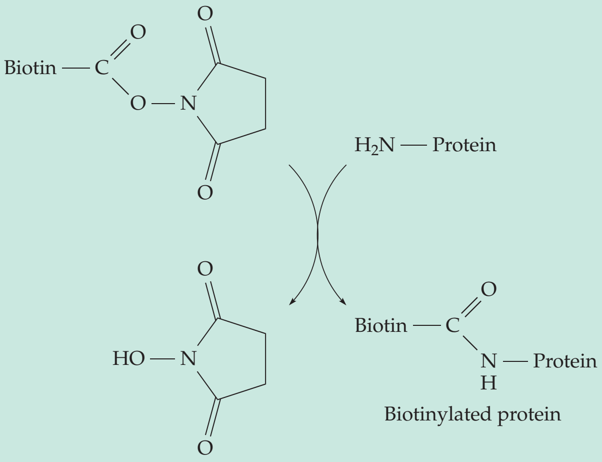
(14-12)

These enzymes do not catalyze any proton exchange at C-3 of pyruvate or at C-2 of an acyl-CoA unless the biotin is first carboxylated. This suggested that removal of the proton to the biotin oxygen and carboxylation might be synchronous. However, ^{13}C and ^{2}H kinetic isotope effects and studies of ^{3}H exchange[66] support the existence of a discrete enolate anion intermediate as shown in Eq. 14-11.[65,67] This mechanism is also consistent with the observation that propionyl-CoA

BOX 14-B THE BIOTIN-BINDING PROTEINS AVIDIN AND STREPTAVIDIN

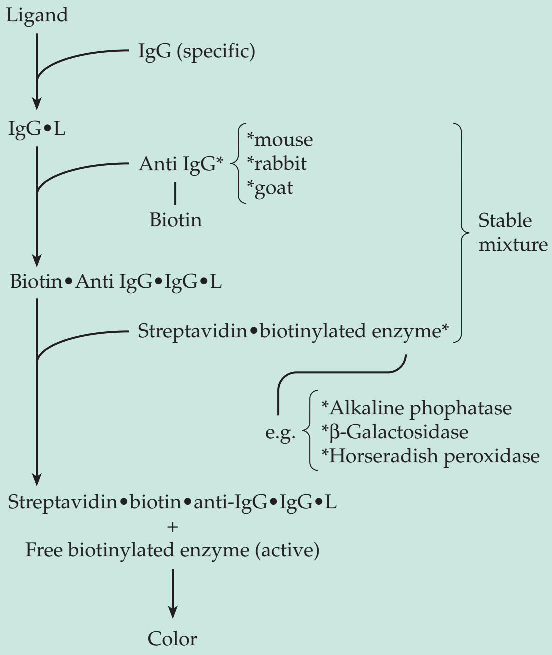
A biochemical curiosity is the presence in egg white of the glycoprotein **avidin**.[a,b] Each 68-kDa subunit of this tetrameric protein binds one molecule of biotin tenaciously with $K_f \sim 10^{15}\ M^{-1}$. Nature's purpose in placing this unusual protein in egg white is uncertain. Perhaps it is a storage form of biotin, but it is more likely an antibiotic that depletes the environment of biotin. A closely similar protein **streptavidin** is secreted into the culture medium by *Streptomyces avidinii*.[c] Its sequence is homologous to that of avidin. It has a similar binding constant for biotin and the two proteins have similar three-dimensional structures.[a,d–j] Biotin binds at one end of a β barrel formed from antiparallel strands and is held by multiple hydrogen bonds and a conformational alteration that allows a peptide loop to close over the bound vitamin.

Historically, avidin was important to the discovery of biotin. The bonding between avidin and biotin is so tight that inclusion of raw egg white in the diet of animals is sufficient to cause a severe biotin deficiency. Avidin has also been an important tool to enzymologists interested in biotin-containing enzymes. Avidin invariably inhibits these enzymes and inhibition by avidin is diagnostic of a biotin-containing protein. Recently, avidin and streptavidin have found widespread application in affinity chromatography, in immunoassays, and in the staining of cells and tissues.[d,k–p] These uses are all based on the ability to attach biotin covalently to side chain groups of proteins, polysaccharides, and other substances. The carboxyl group of the biotin "arm," which lies at the surface of the complex with avidin or streptovidin, can be converted to any of a series of reactive derivatives. For example, *p*-nitrophenyl or *N*-hydroxysuccinimide esters of biotin can be used to attach biotin to amino groups of proteins to form biotinylated proteins. Other reactive derivatives can be used to attach biotin to phenolic, thiol, or carbonyl groups.



Biotinylated protein

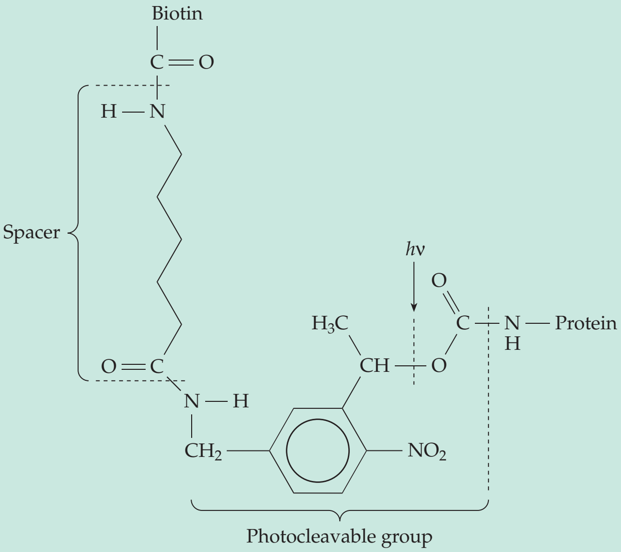
The affinity of avidin or streptavidin for the resulting biotinylated materials is still very high. This fact has permitted the application of this "biotin–avidin technology" to numerous aspects of research and diagnostic medicine. For example, a specific antibody (IgG) can be utilized for immunoassay of a hormone or other ligand. A second antibody, specific for the IgG–ligand complex, can be produced in a biotinylated form. A steptavidin complex of a biotinylated enzyme such as alkaline phosphatase, β-galactosidase, or horseradish peroxidase, for which a sensitive colorimetric assay is available, is allowed to react with the biotin•anti-IgG•IgG•ligand complex. The streptavidin now releases enzyme in proportion to the amount of ligand originally present. Since the biotinylated anti IgG and streptavidin•biotinylated enzyme can be stored as a stable mixture, the assay is simple and fast.



Avidin technology can also be applied to the isolation of proteins and other materials from cells. Because the irreversibility of the binding of biotin may be a problem, photocleavable biotin derivatives have been developed.[q] In the following structure, the biotin derivative has been joined to a protein (as in the first equation in this box) and is ready for separation, perhaps on a column containing immobilized avidin or streptavidin. After separation the biotin together with the linker and photocleavable group are cut off by a short irradiation with ultra-

BOX 14-B (continued)

violet light, leaving the protein in a free form.



[a] Livnah, O., Bayer, E. A., Wilshek, M., and Sussman, J. L. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5076–5080

[b] Pugliese, L., Coda, A., Malcovati, M., and Bolognesi, M. (1993) *J. Mol. Biol.* **231**, 698–710

[c] Meslar, H. W., Camper, S. A., and White, H. B., III. (1978) *J. Biol. Chem.* **253**, 6979–6982

[d] Wilchek, M., and Bayer, E. A. (1989) *Trends Biochem. Sci.* **14**, 408–412

[e] Punekar, N. S., and Lardy, H. A. (1987) *J. Biol. Chem.* **262**, 6714–6719

[f] Henderson, W. A., Pähler, A., Smith, J. L., Satow, Y., Merritt, E. A., and Phizackerley, R. P. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2190–2194

[g] Weber, P. C., Pantoliano, M. W., and Thompson, L. D. (1992) *Biochemistry* **31**, 9350–9354

[h] Schmidt, T. G. M., Koepke, J., Frank, R., and Skerra, A. (1996) *J. Mol. Biol.* **255**, 753–766

[i] Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J., and Salemme, F. R. (1989) *Science* **243**, 85–88

[j] Sano, T., and Cantor, C. R. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3180–3184

[k] Bayer, E. A., and Wilchek, M. (1980) *Meth. Biochem. Anal.* **26**, 1–46

[l] Childs, G. V., Naor, Z., Hazum, F., Tibolt, R., Westlund, K. N., and Hancock, M. B. (1983) *J. Histochem. Cytochem.* **31**, 1422–1425

[m] Wilchek, M., and Bayer, E. A., eds. (1990) *Methods in Enzymology*, Vol. 184, Academic Press, San Diego, California

[n] Savage, M. D., Mattson, G., Desai, S., Nielander, G. W., Morgensen, S., and Conklin, E. J. (1992) *Avidin-Biotin Chemistry: A Handbook*, Pierce, Rockford, Illinois

[o] Donnelson, J. E., and Wu, R. (1972) *J. Biol. Chem.* **247**, 4661–4668

[p] Laundon, C. H., and Griffith, J. D. (1987) *Biochemistry* **26**, 3759–3762

[q] Olejnik, J., Sonar, S., Krzymañska-Olejnik, E., and Rothschild, K. J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7590–7594

carboxylase and transcarboxylase both catalyze elimination of HF from β-fluoropropionyl-CoA to form the unsaturated acrylyl-CoA.[67] The elimination presumably occurs via an enolate anion intermediate as in Eq. 13-28.

A bound divalent metal ion, usually Mn^{2+}, is required in the transcarboxylation step. A possible function is to assist in enolization of the carboxyl acceptor. However, measurement of the effect of the bound Mn^{2+} on ^{13}C relaxation times in the substrate for pyruvate carboxylase indicated a distance of ~0.7 nm between the carbonyl carbon and the Mn^{2+}, too great for direct coordination of the metal to the carbonyl oxygen.[68] Another possibility is that the metal binds to the carbonyl of biotin as indicated in Eq. 14-11. Pyruvate carboxylase utilizes two divalent metal ions and at least one monovalent cation.[68a]

What is the role of the sulfur atom in biotin? Perhaps it interacts with CO_2, helping to hold it in a correct orientation for reaction. Perhaps it helps to keep the ureido ring of biotin planar, or perhaps it has no special function.[69]

3. Control Mechanisms

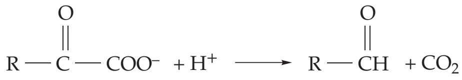
Most pyruvate carboxylases of animal and of yeast are allosterically activated by acetyl-CoA, but those of bacteria are usually not. The enzyme from chicken liver has almost no activity in the absence of acetyl-CoA, which appears to increase greatly the rate of formation of carboxyphosphate and to slow the side reaction by which carboxyphosphate is hydrolyzed to bicarbonate and phosphate.[70] The acetyl-CoA carboxylases of rat or chicken liver aggregate in the presence of citrate to form ~8000-kDa rods. Citrate is an allosteric activator for this enzyme but it acts only on a phosphorylated form and the primary control mechanism.[71,72] This enzyme in plants is a target for a group of herbicides that are selectively toxic to grasses.[73]

4. Pumping Ions with the Help of Biotin

Biotin-dependent **decarboxylases** act as sodium ion pumps in *Klebsiella*[74] and in various anaerobes.[22,75] For example, oxaloacetate is converted to pyruvate and bound carboxybiotin.[74,74a] The latter is decarboxylated to CO_2 at the same time that two Na^+ ions are transported from the inside to the outside of the cell. The function of this pump, like that of the Na^+, K^+-ATPase (Fig. 8-25) is to provide an electrochemical gradient that drives the transport of other ions and molecules through the membrane. Similar ion pumps are operated by decarboxylation of methylmalonyl-CoA[76] and glutaconyl-CoA.[77] Yeast (*Saccharomyces cerevisiae*) cannot make biotin and requires an unusually large amount of the vitamin when urea, allantoin, allantoic acid, and certain other compounds are supplied as the sole source of nitrogen for growth. The reason is that in this organism urea must first be carboxylated by the biotin-containing **urea carboxylase**[78] (see Eq. 24-25) before it can be hydrolyzed to NH_3 and CO_2.

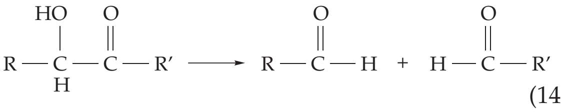
D. Thiamin Diphosphate

In Chapter 13 we considered the breaking of a bond between two carbon atoms, one of which is also bonded to a carbonyl group. These β cleavages are catalyzed by simple acidic and basic groups of the protein side chains. On the other hand, the decarboxylation of 2-oxo acids (Eq. 14-13) and the cleavage and formation of α-hydroxyketones (Eq. 14-14) depend upon thiamin diphosphate (**TDP**).[79–85] These reactions represent a second important method of making and breaking carbon–carbon bonds which we will designate **α condensation and α cleavage**. The common feature of all thiamin-catalyzed reactions is that the bond broken (or formed) is *immediately adjacent to the carbonyl group*, not one carbon removed, as in β cleavage reactions. No simple acid–base catalyzed mechanisms can be written; hence the need for a coenzyme.



$$R-\overset{O}{\overset{\|}{C}}-COO^- + H^+ \longrightarrow R-\overset{O}{\overset{\|}{C}}H + CO_2$$

(14-13)

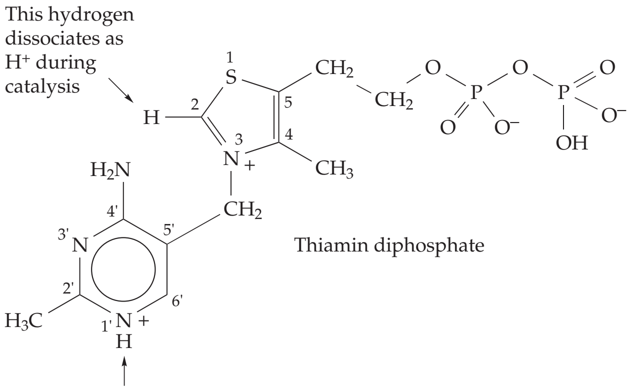


$$R-\underset{H}{\overset{HO}{\overset{|}{C}}}-\overset{O}{\overset{\|}{C}}-R' \longrightarrow R-\overset{O}{\overset{\|}{C}}-H + H-\overset{O}{\overset{\|}{C}}-R'$$

(14-14)

1. Chemical Properties of Thiamin

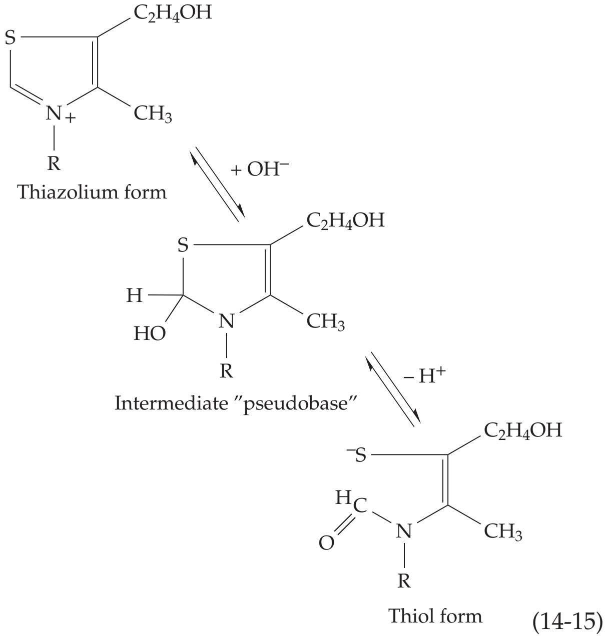
The weakly basic portion of thiamin or of its coenzyme forms is protonated at low pH, largely on N-1 of the pyrimidine ring.[86–88] The pK_a value is ~ 4.9. In basic solution, thiamin reacts in two steps with an opening of the thiazole ring (Eq. 14-15) to give the anion of a thiol form which may be crystallized as the sodium salt.[79,84] This reaction, like the competing reaction described in Eq. 7-19, and which leads to a yellow unstable form of the thiamin anion, is an example of a cooperative two-proton dissociation with linked structural changes. A very low concentration of the intermediate "pseudobase" is present during the titration. This property, which is unusual among small molecules, was instrumental in leading Williams *et al.* to the correct structure for the vitamin.[89] A still unanswered question is, What biological significance is associated with these reactions? Perhaps the thiol form depicted in Eq. 14-15 or the "yellow form" (Eq. 7-19) becomes attached to active sites of some proteins through disulfide linkages.



This hydrogen dissociates as H^+ during catalysis

Thiamin diphosphate

Protonation occurs here with pK_a ~ 4.9

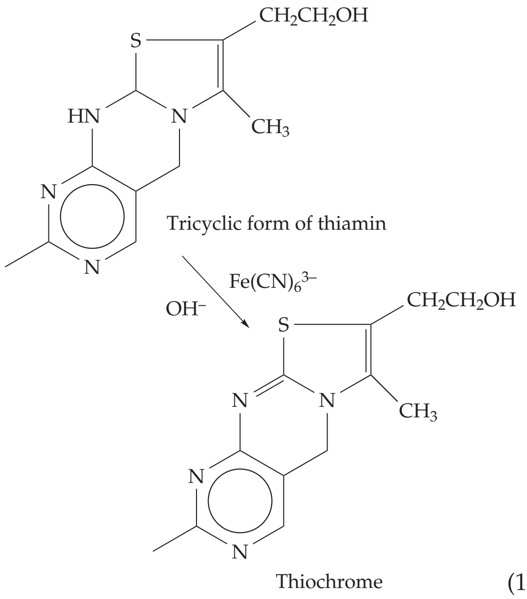


Thiazolium form

Intermediate "pseudobase"

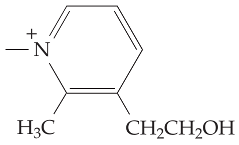
Thiol form (14-15)

Thiamin is unstable at high pH[90,91] and is destroyed by the cooking of foods under mildly basic conditions. The thiol form undergoes hydrolysis and oxidation by air to a disulfide. The tricyclic form (Eq. 7-19) is oxidized to **thiochrome**, a fluorescent compound whose formation from thiamin by treatment with alkaline hexacyanoferrate (III) (ferricyanide; Eq. 14-16) is the basis of a much used fluorimetric assay.

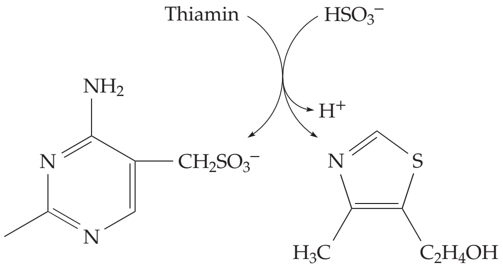


(14-16)

Treatment of thiamin with boiling 5N HCl deaminates it to the hydroxy analogue **oxythiamin**, a potent antagonist. **Pyrithiamin**, another competitor containing in place of the thiazolium ring, is very toxic especially to the nervous system.



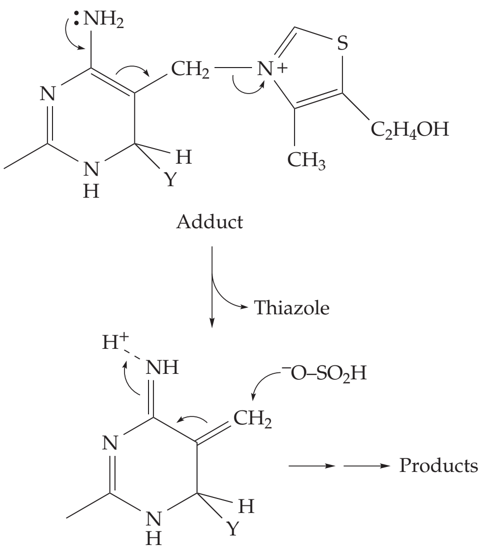
In a solution of sodium sulfite at pH 5, thiamin is cleaved by what appears to be a nucleophilic displacement reaction on the methylene group to give the free thiazole and a sulfonic acid.



(14-17)

In fact the mechanism of the reaction is more complex and is evidently initiated by addition of a nucleophile Y, such as $^-$OH or bisulfite, followed by elimination of the thiazole (Eq. 14-18).

A similar cleavage is catalyzed by thiamin-degrading enzymes known as thiaminases which are found in a number of bacteria, marine organisms, and plants. In a bacterial thiaminase, group Y of Eq. 14-18 is a cysteine –SH.[92,92a]

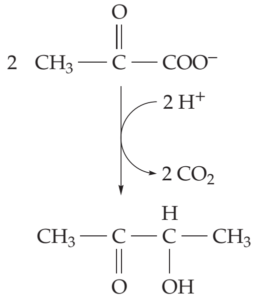


(14-18)

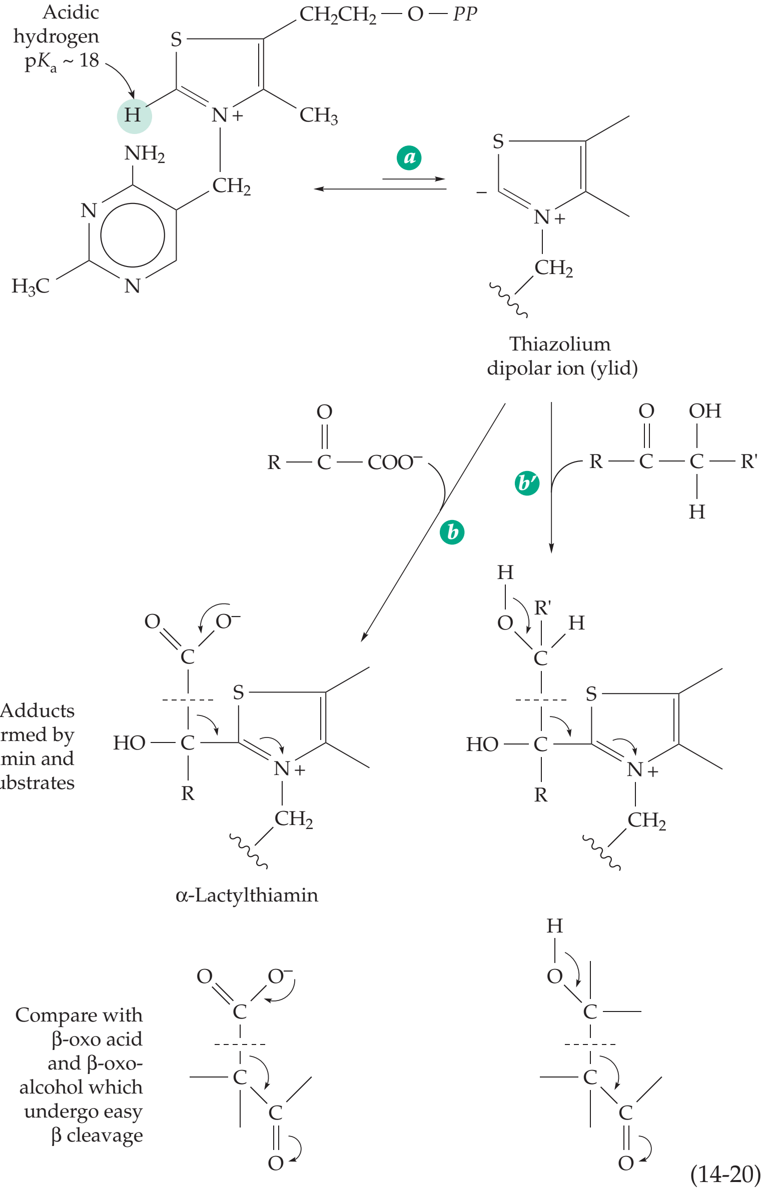
Thiamin is synthesized in bacteria, fungi, and plants from 1-deoxyxylulose 5-phosphate (Eq. 25-21), which is also an intermediate in the nonmevalonate pathway of polyprenyl synthesis. However, thiamin diphosphate is a coenzyme for synthesis of this intermediate (p. 736), suggesting that an alternative pathway must also exist. Each of the two rings of thiamin is formed separately as the esters 4-amino-5-hydroxymethylpyrimidine diphosphate and 4-methyl-5-(β-hydroxyethyl) thiazole monophosphate. These precursors are joined with displacement of pyrophosphate to form thiamin monophosphate.[92b] In eukaryotes this is hydrolyzed to thiamin, then converted to thiamin diphosphate by transfer of a diphospho group from ATP.[92b,c] In bacteria thiamin monophosphate is converted to the diphosphate by ATP and thiamin monophosphate kinase.[92b]

2. Catalytic Mechanisms

The first real clue to the mechanism of thiamin-dependent cleavage came in about 1950 when Mizuhara showed that at pH 8.4 thiamin catalyzes the nonenzymatic conversion of pyruvate into acetoin (Eq. 14-19).[93] Following Mizuhara's lead, Breslow investigated the same reaction using the then new NMR method.[94] He made the surprising discovery that the hydrogen atom in the 2 position of the thiazolium ring, between the sulfur and the nitrogen atoms, exchanged easily with deuterium of ^2H_2O. The pK_a of this proton has been estimated as ~18, low enough to permit rapid



(14-19)


Acidic
hydrogen
pK_a ~ 18
a
b
b′
Thiazolium
dipolar ion (ylid)
Adducts
formed by
thiamin and
substrates
α-Lactylthiamin
Compare with
β-oxo acid
and β-oxo-
alcohol which
undergo easy
β cleavage

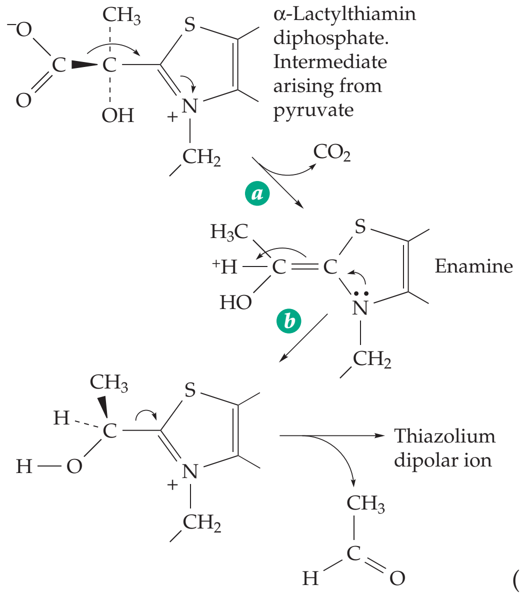
(14-20)

dissociation and replacement with ^{2}H.[84,95,96] The resulting **thiazolium dipolar ion** (or **ylid**) formed by this dissociation (Eq. 14-20, step *a*) is stabilized by the electrostatic interaction of the adjacent positive and negative charges. Breslow suggested that this dipolar ion is the key intermediate in reactions of thiamin-dependent enzymes.[94,97] The anionic center of the dipolar ion can react with a substrate such as an 2-oxo acid or 2-oxo alcohol by addition to the carbonyl group (Eq. 14-20, step *b* or 14-20, step *b*′). The resulting adducts are able to undergo cleavage readily, as indicated by the arrows showing the electron flow toward the =N$^+$– group.

Below the structures of the adducts in Eq. 14-20 are those of a 2-oxo acid and a β-ketol with arrows indicating the electron flow in decarboxylation and in the aldol cleavage. The similarities to the thiamin-dependent cleavage reaction are especially striking if one remembers that in some aldolases and decarboxylases the substrate carbonyl group is first converted to an N-protonated Schiff base before the bond cleavage.

We see that *the essence of the action of thiamin diphosphate as a coenzyme is to convert the substrate into a form in which electron flow can occur from the bond to be broken into the structure of the coenzyme*. Because of this alteration in structure, a bond breaking reaction that would not otherwise have been possible occurs readily. To complete the catalytic cycle, the electron flow has to be reversed again. The thiamin-bound cleavage product (an enamine) from either of the adducts in Eq. 14-20 can be reconverted to the thiazolium dipolar ion and an aldehyde as shown in step *b* of Eq. 14-21 for decarboxylation of pyruvate to acetaldehyde.

The adducts α-lactylthiamin and α-lactylthiamin diphosphate have both been synthesized.[84,98–100] As long as α-lactylthiamin is kept as a dry solid or at low pH, it is stable. However, it decarboxylates readily in neutral solution (Eq. 14-21). Decarboxylation is much

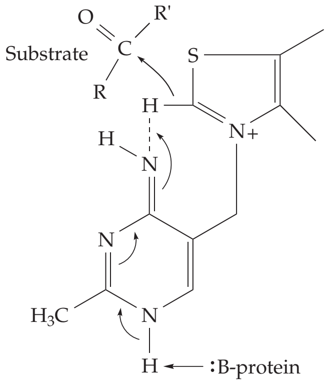

α-Lactylthiamin
diphosphate.
Intermediate
arising from
pyruvate
a
Enamine
b
Thiazolium
dipolar ion

(14-21)

more rapid in methanol, a fact that was predicted by Lienhard and associates.[101] They suggested that decarboxylation is easier in a solvent of low polarity because the transition state has a lower polarity than does lactylthiamin. An enzyme could assist the reaction by providing a relatively nonpolar environment.

The crystal structures of thiamin-dependent enzymes (see next section) as well as modeling[102,103] suggest that lactylthiamin pyrophosphate has the conformation shown in Eq. 14-21. If so, it would be formed by the addition of the ylid to the carbonyl of pyruvate in accord with stereoelectronic principles, and the carboxylate group would also be in the correct orientation for elimination to form the enamine in Eq. 14-21, step *b*.[82–83a] A transient 380- to 440-nm absorption band arising during the action of pyruvate decarboxylase has been attributed to the enamine.

What is the role of the pyrimidine portion of the coenzyme in these reactions? The pyrimidine ring has a large inductive effect on the basicity of the thiazolium nitrogen and may increase the rate of dissociation of the C-2 proton somewhat.[104] More significant is the fact that the $-NH_2$ group is properly placed to function as a basic catalyst in the generation of the thiazolium dipolar ion. However, the amino group of thiamin is not very basic ($pK_a \sim 4.9$) and the site of protonation at low pH is largely N-1 of the pyrimidine ring. In the protonated form the $-NH_2$ group is even less basic because of electron withdrawal into the ring. Studies of thiamin analogs suggested another possibility. Schellenberger[105] found pyruvate decarboxylase inactive when TDP was substituted by analogs with modified aminopyrimidine rings, e.g., with methylated or dimethylated amino groups or with N1 of the ring replaced by carbon (an aminopyridyl analog). More recently the experiment has been repeated with additional enzymes[106] and X-ray studies have shown that the analogs bind into the active site of transketolase in a normal way.[107] Of the compounds studied *only an aminopyridyl analog of TDP having a nitrogen atom at 1′* (but CH at 3') *had substantial catalytic activity.* Jordan and Mariam showed that *N*-1'-methylthiamin is a superior catalyst in nonenzymatic catalysis.[108] These results are consistent with the speculative scheme illustrated in the following drawing from the first edition of this book.[109] The $-NH_2$ group of the N1-protonated aminopyrimidine has lost a proton to form a normally *minor tautomer* in which the resulting imino group would be quite basic. Assisted by a basic group from the protein, it could abstract the proton from the thiazolium ring to form the ylid.



Crystallographic studies show that the catalytic base (:B-protein) is the carboxylate group of a conserved glutamate side chain (E59' in Fig. 14-2). Kern *et al.* used NMR spectroscopy of thiamin diphosphate present in native and mutant pyruvate decarboxylase and transketolase to monitor the exchange rates of the C2-H proton of the thiazolium ring.[109a] The results confirmed the importance of the conserved glutamate side chain for dissociation of the C2-H proton. Participation of other catalytic groups from the enzyme may also be important. However, these groups are not conserved in the whole family of enzymes. For example, glutamine 122, which is within hydrogen-bonding distance of both the substrate and thiamin amino group, is replaced by histidine in transketolase.[110] A variety of kinetic studies involving mutants,[111,111a] alternative substrates,[112,113] and isotope effects in substrates[114–116] and solvent[117,118] have not yet resolved the details of the proton transfers that occur within the active site.

3. Structures of Thiamin-Dependent Enzymes

By 1998, X-ray structures had been determined for four thiamin diphosphate-dependent enzymes: (1) a bacterial pyruvate oxidase,[119,120] (2) yeast and bacterial pyruvate decarboxylases,[121–122c] (3) transketolase,[110,123,124] and (4) benzoylformate decarboxylase.[124a] The reactions catalyzed by these enzymes are all quite different, as are the sequences of the proteins. However, the thiamin diphosphate is bound in a similar way in all of them. A conserved pattern of hydrogen bonds holds the diphosphate group to the protein and also provides ligands to a metal ion. This is normally Mg^{2+}, which is held in nearly perfect octahedral coordination by two phosphate oxygen atoms, a conserved aspartate carboxylate, a conserved asparagine amide, and a water molecule. The thiamin rings are in a less polar region. The amino group of the pyrimidine is adjacent to the 2–CH of the thiazole and N1' of the pyrimidine is apparently protonated and hydrogen bonded to the carboxylate group of a conserved glutamate side chain. This is shown in Fig. 14-2. Substitution of the corresponding glutamate 51 of yeast pyruvate decarboxylase by glutamine or alanine greatly reduced or eliminated catalytic activity.[125]

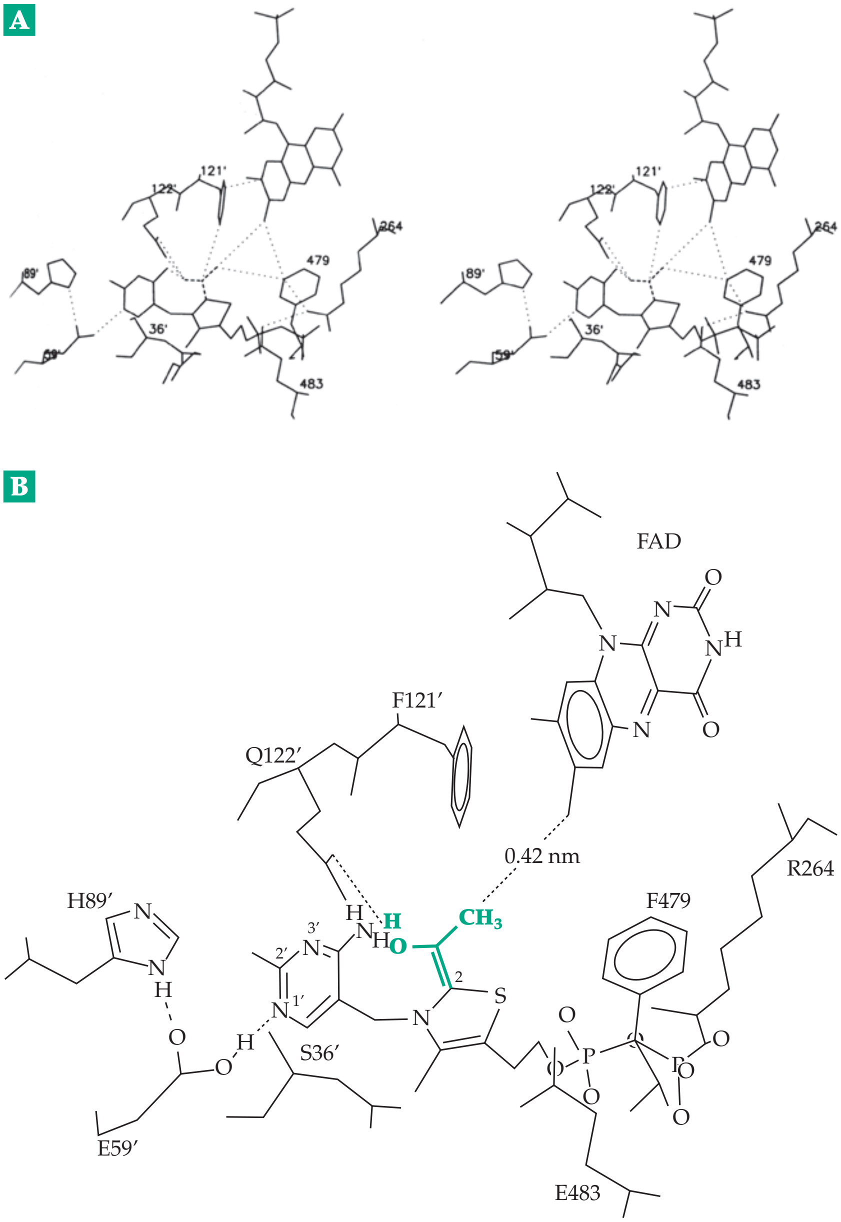
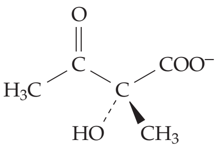


Figure 14-2 (A) Stereoscopic view of the active site of pyruvate oxidase from the bacterium *Lactobacillus plantarium* showing the thiamin diphosphate as well as the flavin part of the bound FAD. The planar structure of the part of the intermediate enamine that arises from pyruvate is shown by dotted lines. Only some residues that may be important for catalysis are displayed: G35', S36', E59', H89', F121', Q122', R264, F479, and E483. Courtesy of Georg E. Schulz.[119] (B) Simplified view with some atoms labeled and some side chains omitted. The atoms of the hypothetical enamine that are formed from pyruvate, by decarboxylation, are shown in green.

4. The Variety of Enzymatic Reactions Involving Thiamin

Most known thiamin diphosphate-dependent reactions (Table 14-2) can be derived from the five half-reactions, *a* through *e*, shown in Fig. 14-3. Each half-reaction is an α cleavage which leads to a thiamin- bound enamine (center, Fig. 14-3) The decarboxylation of an α-oxo acid to an aldehyde is represented by step *b* followed by *a* in reverse. The most studied enzyme catalyzing a reaction of this type is yeast **pyruvate decarboxylase,** an enzyme essential to alcoholic fermentation (Fig. 10-3). There are two ~250-kDa isoenzyme forms, one an α_4 tetramer and one with an $(\alpha\beta)_2$ quaternary structure. The isolation of α-hydroxyethylthiamin diphosphate from reaction mixtures of this enzyme with pyruvate[52] provided important verification of the mechanisms of Eqs. 14-14, 14-15. Other decarboxylases produce aldehydes in specialized metabolic pathways: indolepyruvate decarboxylase[126] in the biosynthesis of the plant hormone **indole-3-acetate** and benzoylformate decarboxylase in the mandelate pathway of bacterial metabolism (Chapter 25).[124a,127]

Formation of α-ketols from α-oxo acids also starts with step *b* of Fig. 14-3 but is followed by condensation with another carbonyl compound in step *c*, in reverse. An example is decarboxylation of pyruvate and condensation of the resulting active acetaldehyde with a second pyruvate molecule to give *R*-α-acetolactate, a reaction catalyzed by **acetohydroxy acid synthase** (acetolactate synthase).[128] Acetolactate is the precursor to valine and leucine. A similar ketol condensation, which is catalyzed by the same synthase, is required in the biosynthesis of isoleucine (Fig. 24-17). Since this



R-α-Acetolactate

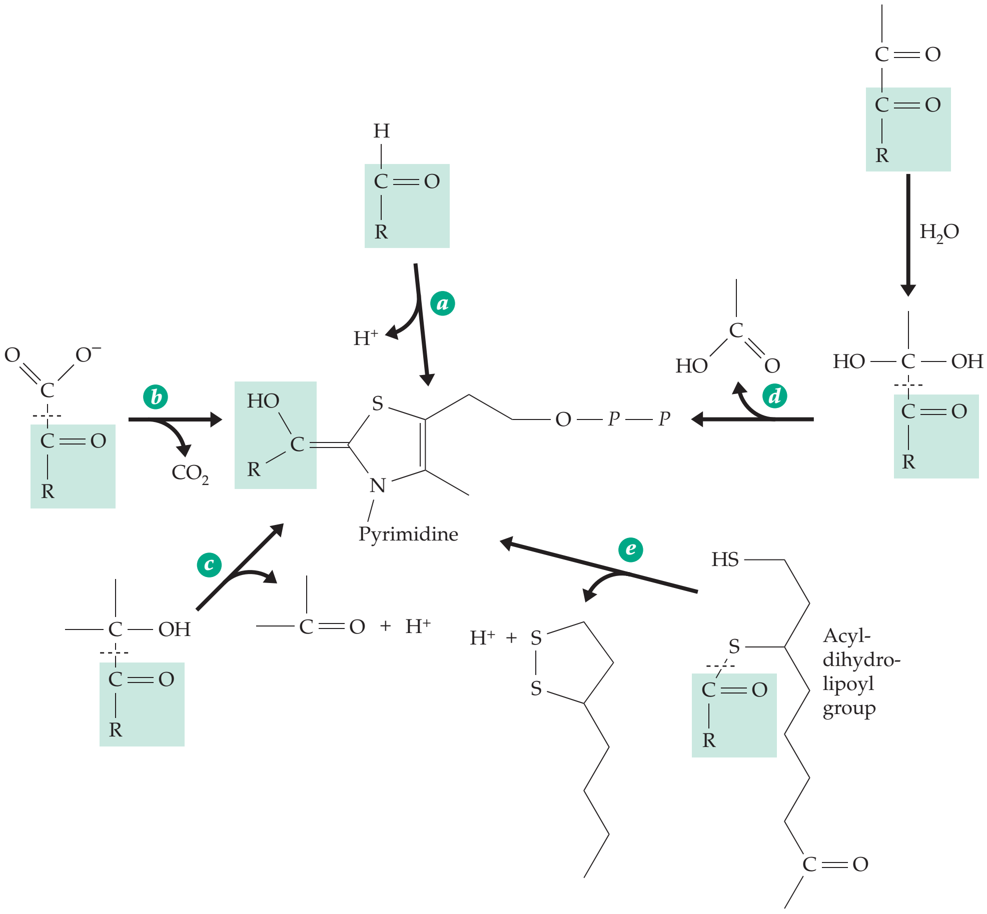

H
C=O
R
a
H+
C=O
C=O
R
H2O
O
O−
C
C=O
R
b
CO2
HO
C
R
S
N
Pyrimidine
O — P — P
C
HO
O
d
HO — C — OH
C=O
R
c
C — OH
C=O
R
C=O + H+
e
H+ + S
S
HS
S
C=O
R
Acyl-dihydro-lipoyl group
C=O

Figure 14-3 Half-reactions making up the thiamin-dependent α cleavage and α condensation reactions.

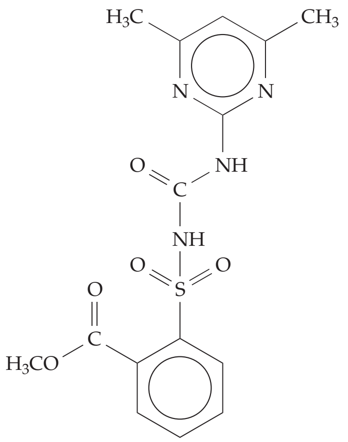
TABLE 14-2
Enzymes Dependent upon Thiamin Diphosphate as a Coenzyme

1. Nonoxidative
    - Pyruvate decarboxylase*
    - Indolepyruvate decarboxylase
    - Benzoylformate decarboxylase*
    - Glyoxylate carboligase
    - Acetohydroxy acid synthase (acetolactate synthase)
    - 1-Deoxy-D-xylulose 5-phosphate synthase
    - Transketolase*
    - Phosphoketolase
2. Oxidative decarboxylase
    - Pyruvate oxidase (FAD)*
    - Pyruvate dehydrogenase (Lipoyl, FAD, NAD+) multienzyme complex
    - Pyruvate:ferredoxin oxidoreductase
    - Indolepyruvate:ferredoxin oxidoreductase

---

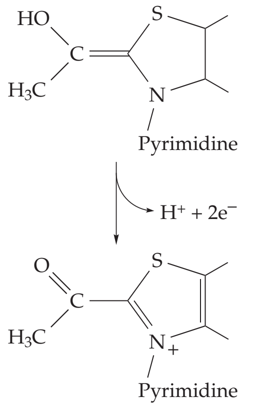
* Three-dimensional structures for these enzymes had been determined by 1998.

synthase is not present in mammals it is a popular target for herbicides.[129,130] It is inhibited by many of the most widely used herbicides including sulfometuron methyl, whose structure is shown here.


H3C
CH3
N
N
O
NH
C
NH
O
O
S
O
C
H3CO

Sulfometuron methyl

Acetolactate is a β-oxo acid and is readily decarboxylated to acetoin, a reaction of importance in bacterial fermentations (Eq. 17-26). Acetoin, of both *R* and *S*



2-Acetylthiamin diphosphate

(14-22)

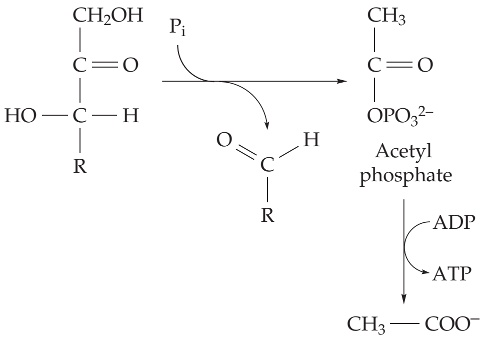
configurations, is also formed by pyruvate decarboxylases acting on acetaldehyde.[103,131] The ketol condensation of two molecules of glyoxylate with decarboxylation to form tartronic semialdehyde (see Fig. 17-6) is an important reaction in bacterial metabolism. It is catalyzed by **glyoxylate carboligase,**[132] another thiamin diphosphate-dependent enzyme. Formation of 1-deoxy-D-xylulose 5-phosphate, an intermediate in the nonmevalonate pathway of isoprenoid synthesis, is formed in a thiamin diphosphate-catalyzed condensation of pyruvate with glyceraldehyde 3-phosophate (Fig. 22-2)[132a] However, there is an unresolved problem. As previously mentioned, the same intermediate is thought to be a precursor to thiamin diphosphate (Eq. 25-21). This suggests the presence of an alternative pathway.

Ketols can also be formed enzymatically by cleavage of an aldehyde (step *a,* Fig. 14-3) followed by condensation with a second aldehyde (step *c*, in reverse). An enzyme utilizing these steps is **transketolase** (Eq. 17-15),[132b] which is essential in the pentose phosphate pathways of metabolism and in photosynthesis. α-Diketones can be cleaved (step *d*) to a carboxylic acid plus active aldehyde, which can react either via *a* or *c* in reverse. These and other combinations of steps are often observed as side reactions of such enzymes as pyruvate decarboxylase. A related thiamin-dependent reaction is that of pyruvate and acetyl-CoA to give the α-diketone, **diacetyl**, $CH_3COCOCH_3$.[133] The reaction can be viewed as a displacement of the CoA anion from acetyl-CoA by attack of thiamin-bound active acetaldehyde derived from pyruvate (reverse of step *d*, Fig. 14-3 with release of CoA).

5. Oxidative Decarboxylation and 2-Acetylthiamin Diphosphate

The oxidative decarboxylation of pyruvate to form acetyl-CoA or acetyl phosphate plays a central role in the metabolism of our bodies and of most other organisms. This reaction is usually formulated as the reverse of step *e* of Fig. 14-3, which shows the cleavage of an **acyl-dihydrolipoyl** derivative. However, there is a possibility that the lipoyl group functions not as shown in Fig. 14-3 but as an oxidant that converts the TDP enamine to 2-acetylthiamin diphosphate (Eq. 14-22)[134,135] and only after that as an acyl group carrier. A related reaction that is known to proceed through acetyl-TDP is the previously mentioned bacterial pyruvate oxidase. As seen in Fig. 14-2, this enzyme has its own oxidant, FAD, which is ready to accept the two electrons of Eq. 14-22 to produce bound acetyl-TDP. The electrons may be able to jump directly to the FAD, with thiamin and flavin radicals being formed at an intermediate stage.[135a] The electron transfers as well as other aspects of oxidative decarboxylation are discussed in Chapter 15, Section C.

A reaction that is related to that of transketolase but is likely to function via acetyl-TDP is **phosphoketolase**, whose action is required in the energy metabolism of some bacteria (Eq. 14-23). A product of phosphoketolase is acetyl phosphate, whose cleavage can be coupled to synthesis of ATP. Phosphoketolase presumably catalyzes an α cleavage to the thiamin-containing enamine shown in Fig. 14-3. A possible mechanism of formation of acetyl phosphate is elimination of H_2O from this enamine, tautomerization to 2-acetylthiamin, and reaction of the latter with inorganic phosphate.



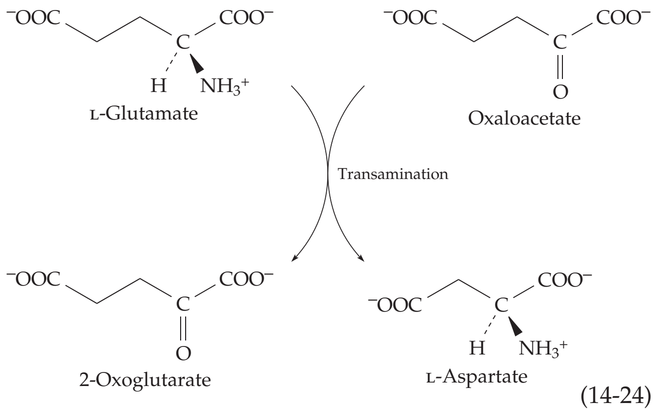
(14-23)

6. Thiamin Coenzymes in Nerve Action

The striking paralysis caused by thiamin deficiency together with studies of thiamin analogs as metabolites suggested a special action for this vitamin in nerves.[136] The thiamin analog **pyrithiamin** (p. 731) both induces paralytic symptoms and displaces thiamin from nerve preparations. The nerve poison **tetrodotoxin** (Chapter 30) blocks nerve conduction by inhibiting inward diffusion of sodium, but it also promotes release of thiamin from nerve membranes. Evidence for a metabolic significance of thiamin triphosphate comes from identification of soluble and membrane-associated thiamin triphosphatases[137] as well as a kinase that forms protein-bound thiamin triphosphate in the brain.[138] Mono-, tri-, and tetraphosphates also occur naturally in smaller amounts. One might speculate about a possible role for the rapid interconversion of cationic and yellow anionic forms of thiamin via the tricyclic form (Eq. 7-19) in some aspect of nerve conduction.

E. Pyridoxal Phosphate

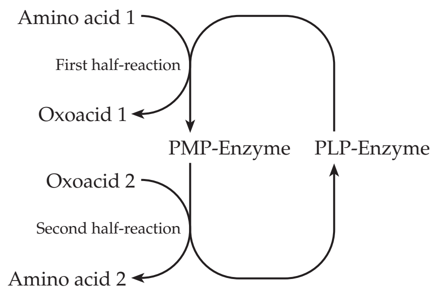
The phosphate ester of the aldehyde form of vitamin B_6, **pyridoxal phosphate** (pyridoxal-*P* or PLP), is required by many enzymes catalyzing reactions of amino acids and amines. The reactions are numerous, and pyridoxal phosphate is surely one of nature's most versatile catalysts. The story begins with biochemical **transamination**, a process of central importance in nitrogen metabolism. In 1937, Alexander Braunstein and Maria Kritzmann, in Moscow, described the transamination reaction by which amino groups can be transferred from one carbon skeleton to another.[139,140] For example, the amino group of glutamate can be transferred to the carbon skeleton of oxaloacetate to form aspartate and 2-oxoglutarate (Eq. 14-24).



(14-24)

This transamination reaction is a widespread process of importance in many aspects of the nitrogen metabolism of organisms. A large series of **transaminases** (**aminotransferases**), for which glutamate is most often one of the reactants, have been shown to catalyze the reactions of other oxoacids and amino acids.[141–143]

In 1944, Esmond Snell reported the nonenzymatic conversion of pyridoxal into pyridoxamine (Box 14-C) by heating with glutamate. He recognized that this was also transamination and proposed that pyridoxal might be a part of a coenzyme needed for aminotransferases and that these enzymes might act via two half-reactions that interconverted pyridoxal and pyridoxamine (Eq. 14-25). The hypothesis was soon verified and the coenzyme was identified as pyridoxal 5'-phosphate or pyridoxamine 5'-phosphate (Fig. 14-5).[144,145]



Sum: Amino acid 1 + oxoacid 2 ⟶
oxoacid 1 + amino acid 2 (14-25)

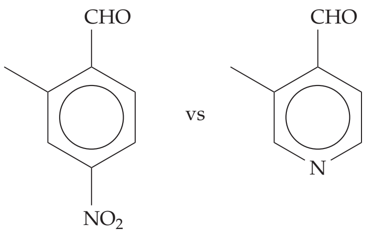
At about the same time, Gunsalus and coworkers noticed that the activity of **tyrosine decarboxylase** produced by lactic acid bacteria was unusually low when the medium was deficient in pyridoxine. Addition of pyridoxal plus ATP increased the decarboxylase activity of cell extracts.[146] PLP was synthesized and was found to be the essential coenzyme for this and a variety of other enzymes.[147]

1. Nonenzymatic Models

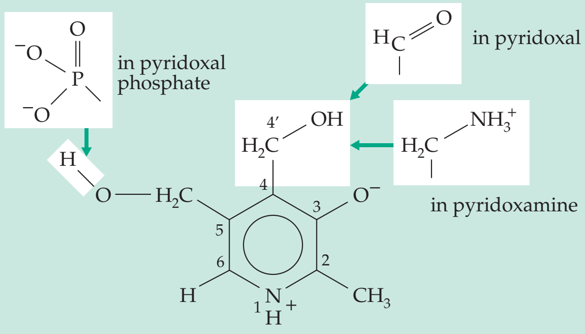
Pyridoxal or PLP, in the complete absence of enzymes, not only undergoes slow transamination with amino acids but also catalyzes many other reactions of amino acids that are identical to those catalyzed by PLP-dependent enzymes. Thus, *the coenzyme itself can be regarded as the active site of the enzymes* and can be studied in nonenzymatic reactions. The latter can be thought of as *models* for corresponding enzymatic reactions. From such studies Snell and associates drew the following conclusions.[148]

a. The aldehyde group of PLP reacts readily and reversibly with amino acids to form Schiff bases (Fig. 14-4) which react further to give products.

b. For an aldehyde to be a catalyst, a strong electron-attracting group, e.g., the ring nitrogen of pyridine (as in PLP), must be *ortho* or *para* to the –CHO group. A nitro group, also strongly electron attracting, can replace the pyridine nitrogen in model reactions.



BOX 14-C THE VITAMIN B_6 FAMILY: PYRIDOXINE, PYRIDOXAL, AND PYRIDOXAMINE



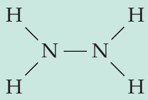
Pyridoxine (pyridoxol)
Vitamin B_6 alcohol

Pyridoxine, the usual commercial form of vitamin B_6, was isolated and synthesized in 1938. However, studies of bacterial nutrition soon indicated the existence of other naturally occurring forms of the new vitamin which were more active than pyridoxine in promoting growth of certain lactic acid bacteria. The amine **pyridoxamine** and the aldehyde **pyridoxal** were identified by Snell, who found that pyridoxal could be formed from pyridoxine by mild oxidation and that pyridoxamine could be formed from pyridoxal by heating in a solution with glutamic acid via a transamination reaction. These simple experiments also suggested the correct structures of the new forms of vitamin B_6. Animal tissues contain largely pyridoxal, pyridoxamine, and their phosphate esters. The lability of the aldehyde explains the ease of destruction of the vitamin by excessive heat or by light. On the other hand, plant tissues contain mostly pyridoxine, which is more stable. Kinases use ATP to form the phosphate esters, which are interconvertible within cells.[a–e] Pyridoxine 5'-phosphate can be oxidized to PLP[d,e,f] and the latter may undergo transamination to PMP. The acid–base chemistry and tautomerism of pyridoxine were discussed in Chapter 6, Section E,2.

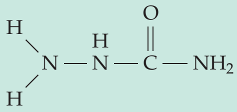
Many poisonous substances as well as useful drugs react with PLP-requiring enzymes. Thus, much of the toxic effect of the "carbonyl reagents" hydroxylamine, hydrazine, and semicarbazide stems from their formation of stable derivatives analogous to Schiff bases with PLP.

$H—O—NH_2$, $R—O—NH_2$

Hydroxylamines

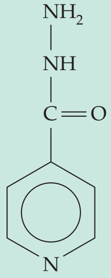


Hydrazine



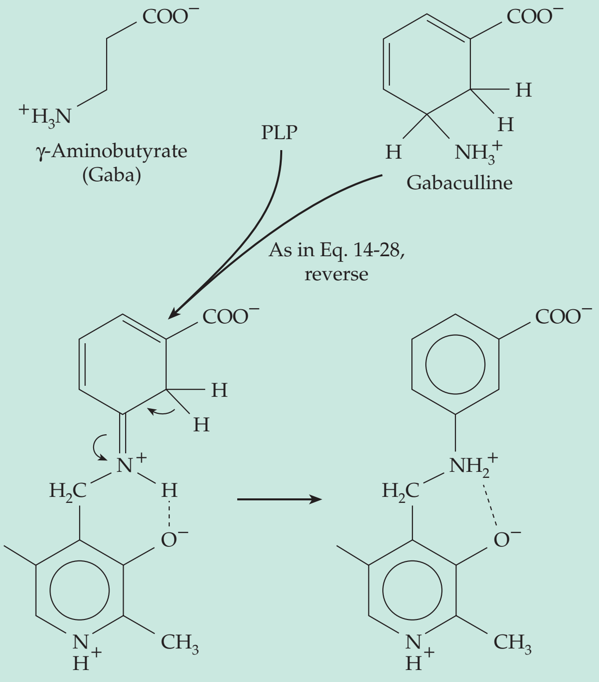
Semicarbazide

Isonicotinyl hydrazide (INH), one of the most effective drugs against tuberculosis, is inhibitory to pyridoxal kinase, the enzyme that converts pyridoxal to PLP.[c] Apparently, the drug reacts with pyridoxal to form a hydrazone which blocks the enzyme. Pyridoxal kinase is not the primary target of INH in mycobacteria. However, patients on long-term isonicotinyl hydrazide therapy sometimes suffer symptoms of vitamin B_6 deficiency.[g]



Isonicotinyl hydrazide
(isoniazid)

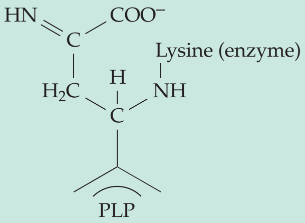
PLP-dependent enzymes are inhibited by a great variety of enzyme-activated inhibitors that react by several distinctly different chemical mechanisms.[h] Here are a few. The naturally occurring **gabaculline** mimics γ-aminobutyrate (Gaba) and inhibits γ-aminobutyrate aminotransferase as well as other PLP-dependent enzymes. The inhibitor follows the normal catalytic pathway as far as the ketimine. There, a proton is lost from the inhibitor permitting formation of a stable benzene ring and leaving the inhibitor stuck in the active site:



Stably bound
inhibitor

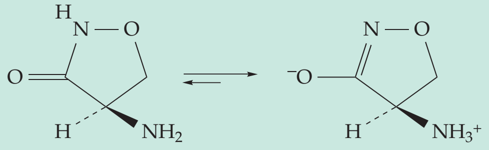
BOX 14-C (continued)

Beta-chloroalanine and serine *O*-sulfate can undergo β elimination (as in Eq. 14-29) in active sites of glutamate decarboxylase or aspartate aminotransferase. The enzymes then form free aminoacrylate, a reactive molecule that can undergo an aldol-type condensation with the external aldimine to give the following product.[i]



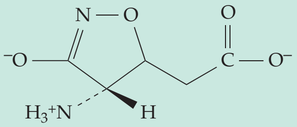
Nucleophilic groups from enzymes can add to double bonds, e.g., in an aminoacrylate Schiff base, or to multiple bonds present in the inhibitor. An example is γ-vinyl γ-aminobutyrate (4-amino-5-hexenoic acid), another inhibitor of brain γ-aminobutyrate aminotransferase which is a useful anticonvulsant drug.

Another enzyme-activated inhibitor is the streptomyces antibiotic **D-cycloserine** (oxamycin), an antitubercular drug that resembles D-alanine in structure. A potent inhibitor of alanine racemase, it also inhibits the non-PLP, ATP-dependent, **D-alanyl-D-alanine synthetase** which is needed in the biosynthesis of the peptidoglycan of bacterial cell walls.

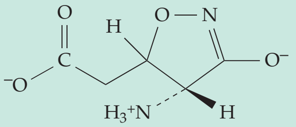


D-Cycloserine

L-Cycloserine inhibits many PLP enzymes and is toxic to humans. This observation led Khomutov *et al.* to synthesize the following more specific "cycloglutamates," structural analogs of glutamic acid with fixed conformations.[j,k] Nature apparently anticipated the synthetic chemist, because it has been reported that the mushroom *Tricholoma muscarium* contains one of the same compounds. It is said to impart two interesting properties to the mushroom: a delicious flavor and a lethal action on flies that alight on the mushroom's surface![l]

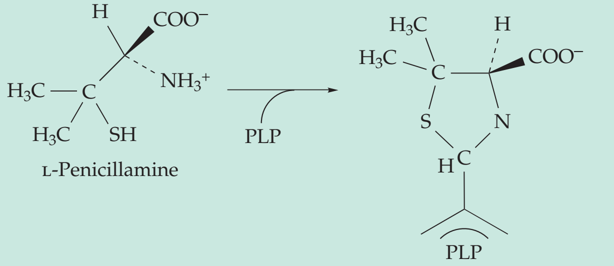


A cycloglutamate that inhibits aspartate aminotransferase



An isomeric cycloglutamate tricholomic acid found in certain mushrooms

The substituted cysteine derivative L-penicillamine causes convulsions and low glutamate decarboxylase levels in the brain, presumably because the Schiff base formed with PLP can then undergo cyclization, the SH group adding to the C=N to form a stable thiazolidine ring.



Toxopyrimidine, the alcohol derived from the pyrimidine portion of the thiamin molecule, is a structural analog of pyridoxal. When fed to rats or mice it induces running fits which can be stopped by administration of vitamin B_6. Phophorylation of toxopyrimidine by pyridoxal kinase may produce an antagonistic analog of PLP. In a similar fashion, 4-deoxypyridoxine, which was tested as a possible anticancer drug, caused convulsions and other symptons of vitamin B_6 deficiency in humans. A host of synthetic PLP derivatives have been made, some of which are effective in blocking PLP enzymes.[h]

---

[a] Lepkovsky, S. (1979) *Fed. Proc.* **38**, 2699–2700

[b] McCormick, D. B., Gregory, M. E., and Snell, E. E. (1961) *J. Biol. Chem.* **236**, 2076–2084

[c] Snell, E. E., and Haskell, B. E. (1970) *Comprehensive Biochemistry* **21**, 47–71

[d] McCormick, D. B. and Chen, H. (1999) *J. Nutr.* **129**, 325–327

[e] Hanna, M. L., Turner, A. J., and Kirkness, E. F. (1997) *J. Biol. Chem.* **272**, 10756–10760

[f] Ngo, E. O., LePage, G. R., Thanassi, J. W., Meisler, N., and Nutter, L. M. (1998) *Biochemistry* **37**, 7741–7748

[g] Lui, A., and Lumeng, L. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1B (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 601–674, Wiley, New York

[h] Walsh, C. T. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1B (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 43–70, Wiley, New York

[i] Likos, J. J., Ueno, H., Fedhaus, R. W., and Metzler, D. E. (1982) *Biochemistry* **21**, 4377–4386

[j] Khomutov, R. M., Koveleva, G. K., Severin, E. S., and Vdovina, L. V. (1967) *Biokhim.* **32**, 900–907

[k] Sastchenko, L. P., Severin, E. S., Metzler, D. E., and Khomutov, R. M. (1971) *Biochemistry* **10**, 4888–4894

[l] Iwasaki, H., Kamiya, T., Oka, O., and Veyanagi, J. (1969) *Chem. Pharm. Bull.* **17**, 866–872

c. The presence of an – OH group adjacent to the – CHO group greatly enhances the catalytic activity. Since certain metal ions, such as Cu^{2+} and Al^{3+}, increase the rates in model systems and are known to chelate with Schiff bases of the type formed with PLP, it was concluded that either a metal ion or a proton formed a chelate ring and helped to hold the Schiff base in a planar conformation (Fig. 14-6). *However, such a function for metal ions has not been found in PLP-dependent enzymes.*

d. In model systems the 5-hydroxymethyl and 2-methyl groups are not needed for catalysis. However, in enzymes the 5-CH_2OH group is essential for attachment of the phosphate handle. The 2-CH_3 group is usually not necessary for coenzymatic activity.

Many investigations of nonenzymatic reactions of PLP and related compounds have been and are still being conducted[149,150]

2. A General Mechanism of Action of PLP

Based upon consideration of the various known PLP-dependent enzymes of amino acid metabolism, Braunstein and Shemyakin in 1952 proposed a general mechanism of PLP action[151,152] which, in most details, was the same as the one proposed independently by Snell and associates on the basis of the nonenzymatic reactions.[148] The general mechanism, which has been verified by studies of many enzymes, can be stated as follows: *Pyridoxal phosphate reacts to convert the amino group of a substrate into a Schiff base that is electronically the equivalent of an adjacent carbonyl* (Fig. 14-4). However, a Schiff base of an amino acid with a simple aldehyde (for example, acetaldehyde) has the opposite polarity from that of C=O (see the following structures). Such an imine could not substitute for a carbonyl group in activating an α-hydrogen nor in facilitating C–C bond cleavage in the amino acid. It is necessary to have the strongly electron-attracting pyridine group conjugated with the C=N group in such a way that electrons can flow from the substrate into the coenzyme.

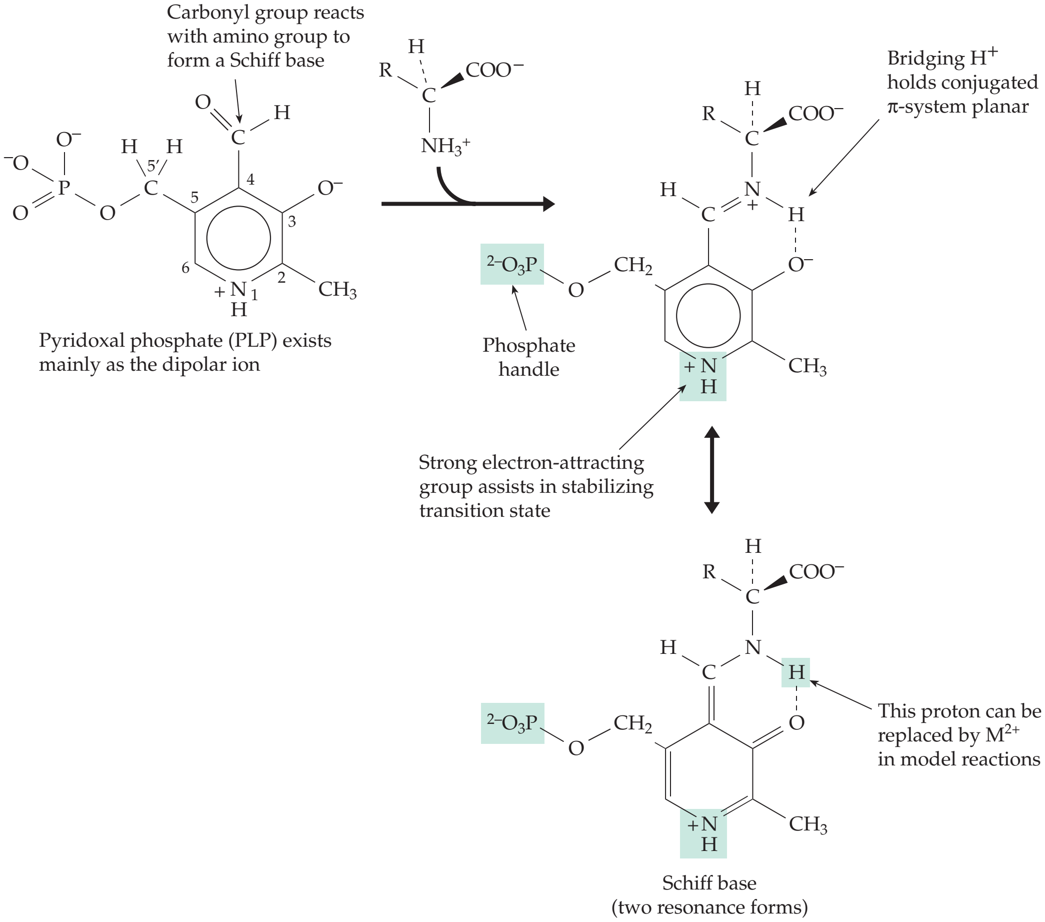
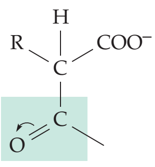
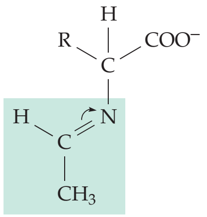

Carbonyl group reacts with amino group to form a Schiff base
Pyridoxal phosphate (PLP) exists mainly as the dipolar ion
Bridging H⁺ holds conjugated π-system planar
Phosphate handle
Strong electron-attracting group assists in stabilizing transition state
This proton can be replaced by M²⁺ in model reactions
Schiff base (two resonance forms)

Figure 14-4 Pyridoxal 5'-phosphate (PLP), a special coenzyme for reactions of amino acids.

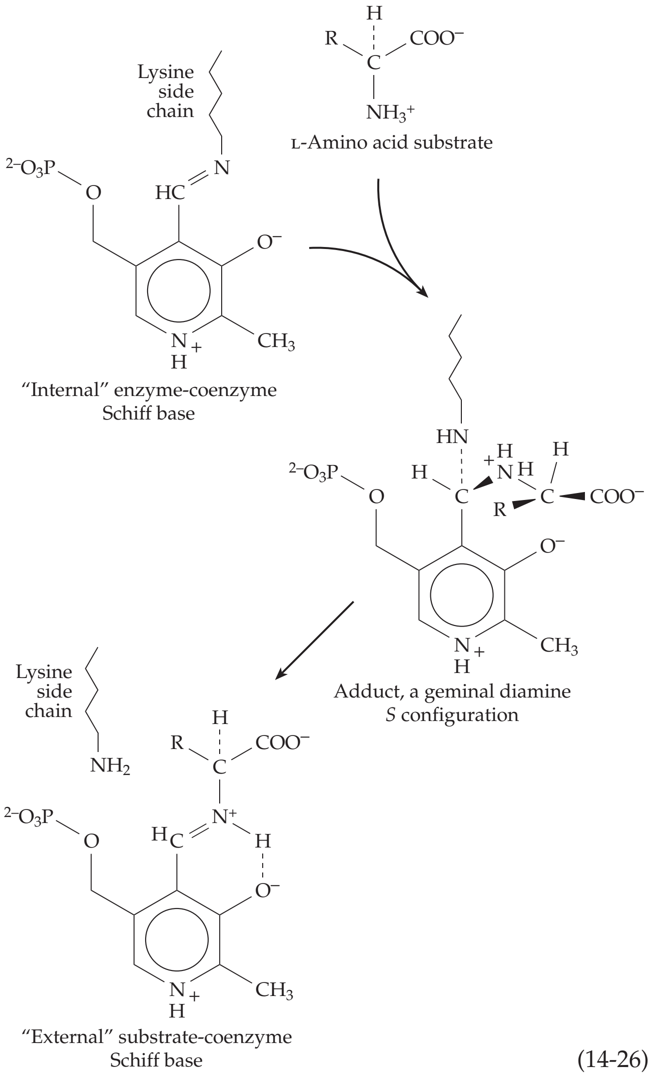


Carbonyl group



Schiff base with acetaldehyde

Before discussing the reactions of Schiff bases of PLP we should consider one fact that was not known in 1952. PLP is bound into an enzyme's active site as a Schiff base with a specific lysine side chain before a substate binds . This is often called the **internal aldimine**. When the substrate binds it reacts with the internal Schiff base by a two-step process called **transimination** (Eq. 14-26) to form the substrate Schiff base, which is also called the **external aldimine**.



(14-26)

3. The Variety of PLP-Dependent Reactions

In Fig. 14-5 the reactions of PLP-amino acid Schiff bases are compared with those of β-oxo-acids. Beta-hydroxy-α-oxo acids and Schiff bases of PLP with β-hydroxy-α-amino acids can react in similar ways. The reactions fall naturally into three groups (*a*,*b*,*c*) depending upon whether the bond cleaved is from the α-carbon of the substrate to the hydrogen atom, to the carboxyl group, or to the side chain. A fourth group of reactions of PLP-dependent enzymes (*d*) also involve removal of the α-hydrogen but are mechanistically more complex. Some of the many reactions catalyzed by these enzymes are listed in Table 14-3.

***Loss of the α-hydrogen* (Group *a*).** Dissociation of the α-hydrogen from the Schiff base leads to a **quinonoid–carbanionic intermediate** whose structure in depicted in Fig. 14-5. The name reflects the characteristics of the two resonance forms drawn. Like an enolate anion, this intermediate can react in several ways (1-4).

(1) **Racemization**. A proton can be added back to the original alpha position but without stereospecificity. A racemase which does this is important to bacteria. They must synthesize D-alanine and D-glutamic acid from the corresponding L-isomers for use in formation of their peptidoglycan envelopes.[153–154a] The combined actions of alanine racemase plus D-alanine aminotransferase, which produces D-glutamate as a product, provide bacteria with both D amino acids. A fungal alanine racemase is necessary for synthesis of the immunosuppresant cyclosporin (Box 9-F).[155,155a] High concentrations of free D-alanine are found in certain regions of the brain and also in various glands.[156]

The carboxyl group of an amino acid can also activate the α-hydrogen. This may be the basis for an aspartate racemase and other racemases that are *not* dependent upon PLP.[156–158] See also Chapter 13, Section B,4.

(2) **Cyclization**. A second kind of reaction is represented by the conversion of *S*-adenosylmethionine to **aminocyclopropanecarboxylic acid**, a precursor to the plant hormone **ethylene** (see Chapter 24).[159] The quinonoid intermediate cyclizes with elimination of methylthioadenosine to give a Schiff base of the product (Eq. 14-27).[160–161a] The cyclization step appears to be a simple S_N2-like reaction.[162]

(3) **Transamination**. A proton can add to the carbon attached to the 4 position of the PLP ring (Fig. 14-5) to form a second Schiff base, often referred to as a **ketimine** (Eq. 14-28). The latter can readily undergo hydrolysis to **pyridoxamine phosphate** (PMP) and an α-oxo acid. This sequence represents one of the

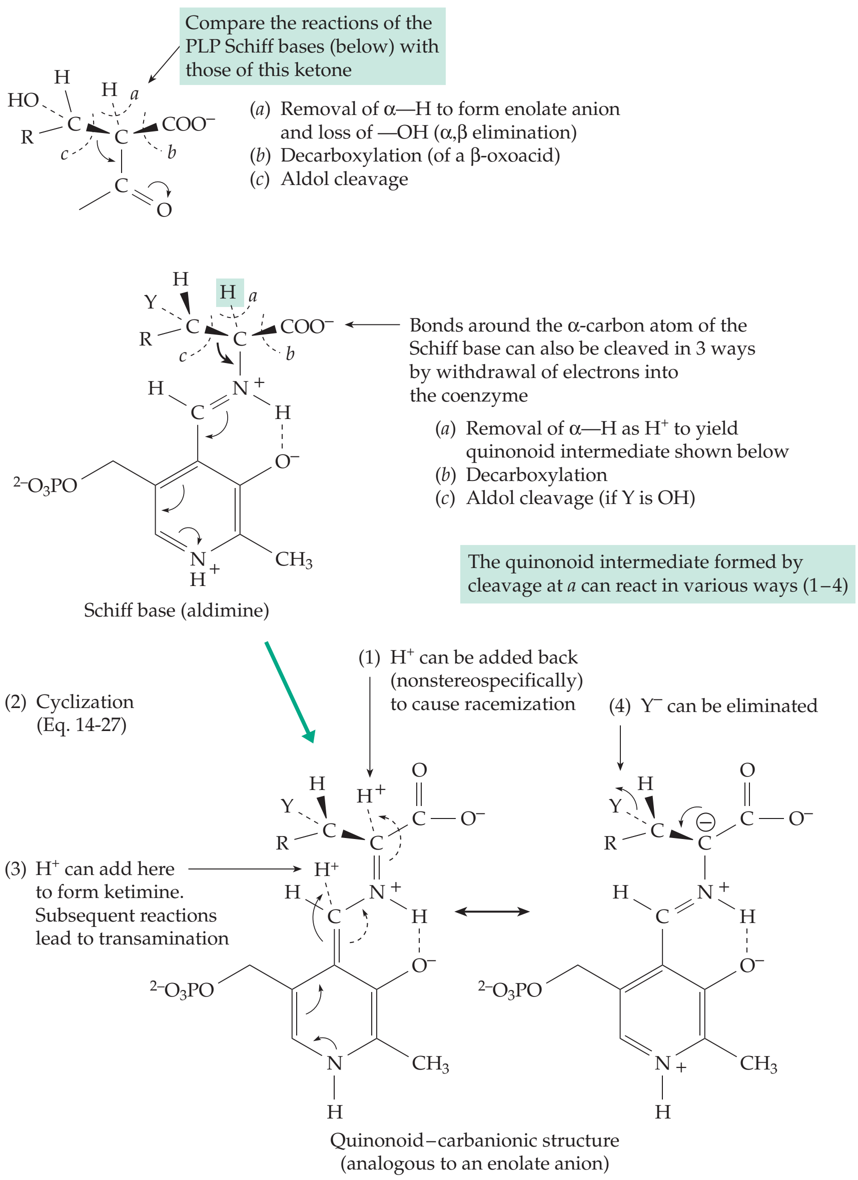

Compare the reactions of the PLP Schiff bases (below) with those of this ketone
HO
H
H
a
R
C
C
COO⁻
c
b
C
O
(a) Removal of α—H to form enolate anion and loss of —OH (α,β elimination)
(b) Decarboxylation (of a β-oxoacid)
(c) Aldol cleavage
H
H
a
Y
R
C
C
COO⁻
c
b
Bonds around the α-carbon atom of the Schiff base can also be cleaved in 3 ways by withdrawal of electrons into the coenzyme
(a) Removal of α—H as H⁺ to yield quinonoid intermediate shown below
(b) Decarboxylation
(c) Aldol cleavage (if Y is OH)
H
C
N⁺
H
O⁻
²⁻O₃PO
N⁺
H
CH₃
Schiff base (aldimine)
The quinonoid intermediate formed by cleavage at a can react in various ways (1–4)
(1) H⁺ can be added back (nonstereospecifically) to cause racemization
(2) Cyclization (Eq. 14-27)
(4) Y⁻ can be eliminated
(3) H⁺ can add here to form ketimine. Subsequent reactions lead to transamination
H
H⁺
Y
R
C
C
O
C—O⁻
H⁺
H
C
N⁺
H
O⁻
²⁻O₃PO
N
H
CH₃
Y
H
R
C
⊖
C
O
C—O⁻
H
C
N⁺
H
O⁻
²⁻O₃PO
N⁺
H
CH₃
Quinonoid–carbanionic structure (analogous to an enolate anion)

Figure 14-5 Some reactions of Schiff bases of pyridoxal phosphate. (*a*) Formation of the quinonoid intermediate, (*b*) elimination of a β substituent, and (*c*) transamination. The quinonoid–carbanionic intermediate can react in four ways (1–4) if enzyme specificity and substrate structure allow.

two half-reactions (Eqs. 14-24 and 14-25) required for enzymatic transamination.

Transaminases participate in metabolism of most of the amino acids, over 60 different enzymes have been identified.[142,163] Best studied are the **aspartate aminotransferases**, a pair of cytosolic and mitochondrial isoenzymes which can be isolated readily from animal hearts. Their presence in heart muscle and brain in high concentration is thought to be a result of their functioning in the malate–aspartate shuttle (Chapter 18). The sequences of the two proteins differ greatly, with only 50% of the residues being the same in both isoenzymes. However, these differences are largely on the outside surface, the folding pattern (Fig. 2-6) and internal structure (Fig. 14-6) are almost identical.[142,163–167a] Three-dimensional structures of aspartate aminotransferases of *E. coli*, yeast, chickens, and mammals are extremely similar, even though sequence identity may be as low as 20%.[167a] Most other transaminases also use the L-gutamate–oxoglutarate pair as one of the product–reactant pairs but a few prefer smaller substrates with uncharged side chains. An example is serine: pyruvate (or alanine:glyoxylate) aminotransferase, an important mitochondrial and peroxisomal enzyme in both animals and plants.[168] Other specialized aminotransferases act on aromatic amino acids,[168a] the branched chain amino acids valine, leucine, and isoleucine,[168b,c] and D-amino acids.[168d] Many of them are highly specific for individual amino acids such as phosphoserine,[168e] ornithine,[168f,g] *N*-acetylornithine,[168h,i] and 8-amino-7-oxononanoate (see banner, p. 719).[168j] An apparently internal transamination, which requires PMP and PLP, converts glutamate-1-semialdehyde into δ-aminolevulinate (see Eq. 24-44) in the pathway of porphyrin biosynthesis used by bacteria and plants.[168k]

(4) **Elimination and β replacement**. When a good leaving group is present in the β position of the amino acid it can be eliminated (Fig. 14-5, Eq. 14-29).[170] A large number of enzymes catalyze such reactions. Among them are **serine** and **threonine dehydratases,** which eliminate OH^- as H_2O;[171–173a] **tryptophan indole-lyase** (tryptophanase) of bacteria, which eliminates indole;[174–176c] **tyrosine phenol-lyase** (elimination of phenol);[177–178a] and **alliinase** of garlic (elimination of 1-propenylsulfenic acid).[179,180] Cystathionine, a precursor to methionine, eliminates L-homocysteine through the action of **cystathionine β lyase** (cystathionase).[181,182] Ammonia is eliminated from the β position of 2,3-diaminopropionate by a bacterial lyase.[183]

TABLE 14-3
Some Enzymes That Require Pyridoxal Phosphate as a Coenzyme

(*a*) Removing alpha hydrogen as H^+
- (1) Racemization
  - Alanine racemase*
- (2) Cyclization
  - Aminocyclopropane carboxylate synthase
- (3) Amino group transfer
  - Aspartate aminotransferase*
  - Alanine aminotransferase
  - D-Amino acid aminotransferase*
  - Branched chain aminotransferase
  - Gamma-aminobutyrate aminotransferase
  - ω-Amino acid:pyruvate aminotransferase*
  - Tyrosine aminotransferase
  - Serine:pyruvate aminotransferase
- (4) Beta elimination or replacement
  - D- and L- Serine dehydratases (deaminases)
  - Tryptophan indole-lyase (tryptophanase)*
  - Tyrosine phenol-lyase*
  - Alliinase
  - Cystathionine β-lyase (cystathionase)*
  - *O*-Acetylserine sulfhydrylase (cysteine synthase)
  - Cystathionine β-synthase
  - Tryptophan synthase*

(*b*) Removal of alpha carboxylate as CO_2
- Diaminopimelate decarboxylase
- Glycine decarboxylase (requires lipoyl group)
- Glutamate decarboxylase
- Histidine decarboxylase
- Dopa decarboxylase
- Ornithine decarboxylase*
- Tyrosine decarboxylase
- Dialkylglycine decarboxylase (a decarboxylating transaminase)*

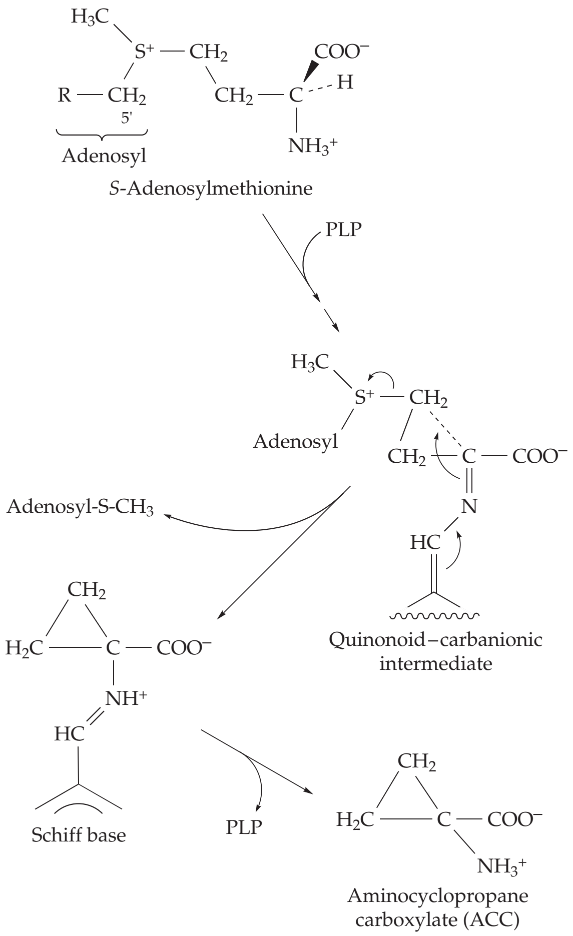
(*c*) Removal or replacement of side chain (or –H) by aldol cleavage
- Serine hydroxymethyltransferase
- Threonine aldolase
- δ-Aminolevulinate synthase
- Serine palmitoyltransferase
- 2-Amino-3-oxobutyrate-CoA ligase

(*d*) Reactions of ketimine intermediates
- Aspartate γ-decarboxylase
- Selenocysteine lyase
- *Nif S* protein of nitrogenase
- Gamma elimination and replacement
  - Cystathionine γ-synthase
  - Cystathionine γ-lyase
  - Threonine synthase

(*e*) Other enzymes
- Lysine 2,3-aminomutase
- Glycogen phosphorylase*
- Pyridoxamine phosphate (PMP) in synthesis of 3,6-dideoxy hexoses

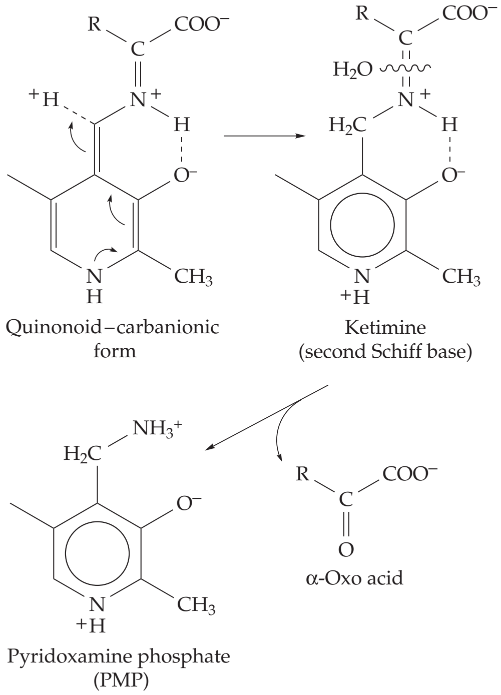
---

* The three-dimensional structures of these and other PLP-dependent enzymes were determined by 2000.



H3C
S+ — CH2
R — CH2
5'
CH2 — C
COO−
H
NH3+
Adenosyl
S-Adenosylmethionine
PLP
H3C
S+ — CH2
Adenosyl
CH2
C — COO−
N
HC
Adenosyl-S-CH3
Quinonoid–carbanionic intermediate
CH2
H2C
C — COO−
NH+
HC
Schiff base
PLP
CH2
H2C
C — COO−
NH3+
Aminocyclopropane carboxylate (ACC)

(14-27)



R
C
COO−
N+
+H
H
O−
N
H
CH3
Quinonoid–carbanionic form
R
C
COO−
H2O
N+
H2C
H
O−
N
+H
CH3
Ketimine (second Schiff base)
NH3+
H2C
O−
N
+H
CH3
Pyridoxamine phosphate (PMP)
R
C
COO−
O
α-Oxo acid

(14-28)

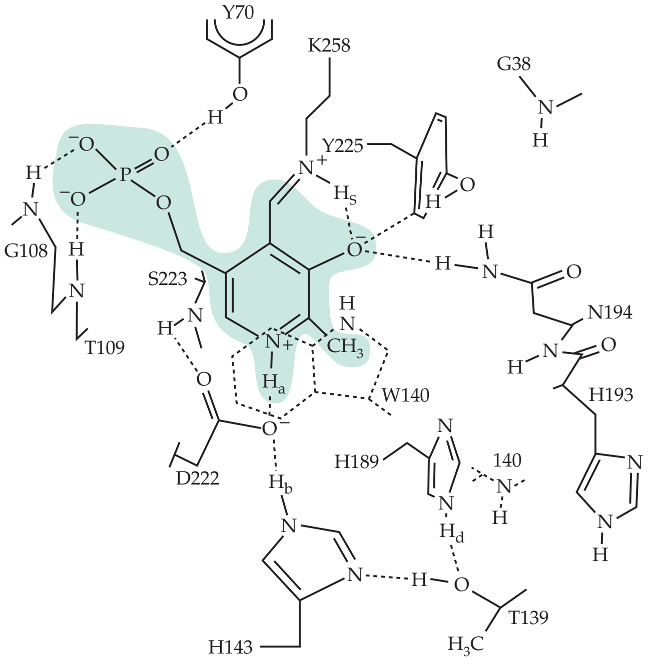
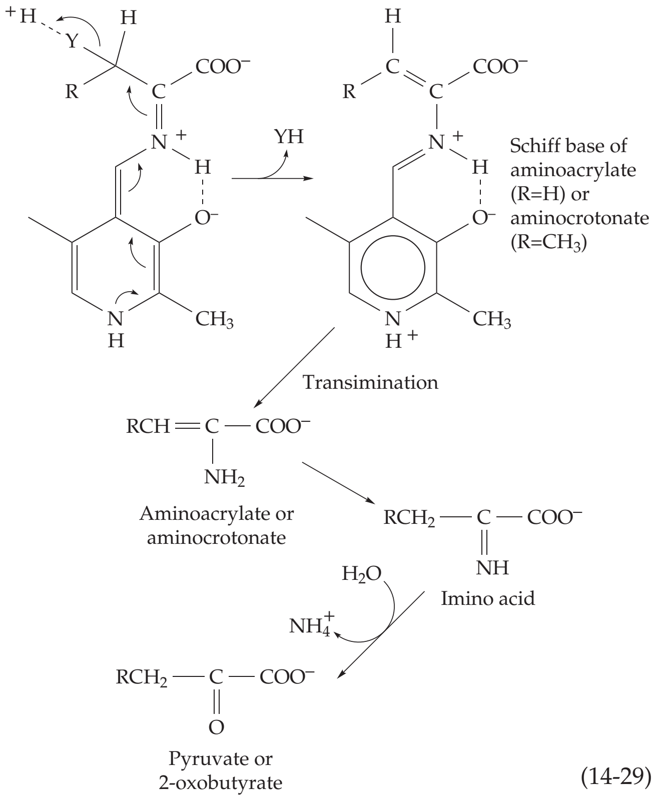
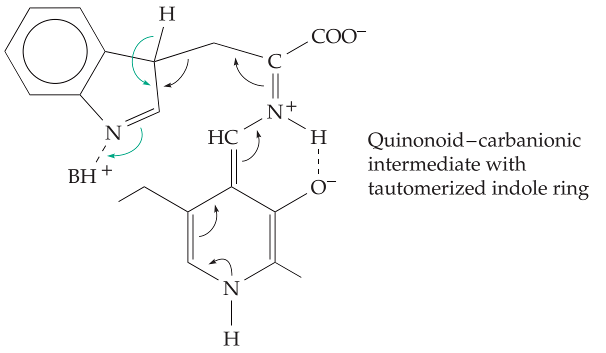

Y70
K258
G38
Y225
G108
S223
T109
N194
W140
H193
D222
H189
140
H143
T139
H_S
H_a
H_b
H_d
CH_3
H_3C

Figure 14-6 Drawing showing pyridoxal phosphate (shaded) and some surrounding protein structure in the active site of cytosolic aspartate aminotransferase. This is the low pH form of the enzyme with an *N*-protonated Schiff base linkage of lysine 258 to the PLP. The tryptophan 140 ring lies in front of the coenzyme. Several protons, labeled H_a, H_b, and H_d, are represented in ^1H NMR spectra by distinct resonances whose chemical shifts are sensitive to changes in the active site.[169]


YH
Schiff base of aminoacrylate (R=H) or aminocrotonate ($R=CH_3$)
Transimination
$RCH{=}C{-}COO^-$
NH_2
Aminoacrylate or aminocrotonate
$RCH_2{-}C{-}COO^-$
NH
Imino acid
H_2O
NH_4^+
$RCH_2{-}C{-}COO^-$
O
Pyruvate or 2-oxobutyrate

(14-29)

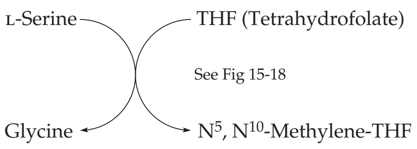
Beta replacement is catalyzed by such enzymes of amino acid biosynthesis as **tryptophan synthase** (Chapter 25),[184] ***O*-acetylserine sulfhydrylase** (cysteine synthase),[185–186a] and **cystathionine β-synthase** (Chapter 24).[187–188c] In both elimination and β replacement an unsaturated Schiff base, usually of aminoacrylate or aminocrotonate, is a probable intermediate (Eq. 14-29). Conversion to the final products is usually assumed to be via hydrolysis to free aminoacrylate, tautomerization to an imino acid, and hydrolysis of the latter, e.g., to pyruvate and ammonium ion (Eq. 14-29). However, the observed stereospecific addition of a proton at the β-C atom of 2-oxobutyrate[189] suggests that these steps may occur with the participation of groups from the enzyme. Before indole can be eliminated by tryptophan indolelyase the indole ring must presumably be tautomerized to the following form of the quinonoid intermediate. The same species may be created by tryptophan synthase upon addition of indole to the enzyme-bound aminoacrylate. The green arrows on the structure indicate the tautomerization that would occur to convert the indole ring to the structure found in tryptophan. The three-dimensional structure of tryptophan synthase is shown in Fig. 25-3. It is a complex of two enzymes with a remarkable tunnel through which the intermediate indole can pass.[184,190] An unusual PLP-dependent β replacement is used to synthesize a transfer RNA ester of **selenocysteine** prior to its insertion into special locations in a few proteins (Chapter 15, Section G, and Chapter 29).


COO^-
BH^+
Quinonoid–carbanionic intermediate with tautomerized indole ring

***Decarboxylation* (Group *b*).** The bond to the carboxyl group of an amino acid substrate is broken in reactions catalyzed by **amino acid decarboxylases**.[191,192] These also presumably lead to a transient quinonoid-carbanionic intermediate. Addition of a proton at the original site of decarboxylation followed by breakup of the Schiff base completes the sequence. Decarboxylation of amino acids is nearly irreversible and frequently appears as a final step in synthesis of amino compounds. For example, in the brain glutamic acid is decarboxylated to **γ-aminobutyric acid** (Gaba),[193–196b] while 3,4-dihydroxyphenylalanine (dopa) and 5-hydroxytryptophan are acted upon by an **aromatic amino acid decarboxylase** to form, respectively, the neurotransmitters **dopamine** and **serotonin**.[197–199b] Histidine is decarboxylated to **histamine**.[200–202] However, not all histidine decarboxylases use PLP as a coenzyme (Section F).

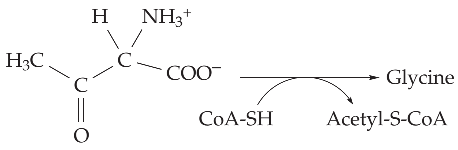
Arginine is converted by a PLP-dependent decarboxylase to agmatine (Fig. 24-12) which is hydrolyzed to **1,4-diaminopropane**.[191] This important cell constituent is also formed by hydrolysis of arginine to **ornithine** (Fig 24-10) and decarboxylation of the latter.[203–206c] **Lysine** is formed in bacteria by decarboxylation of *meso*-diamino-pimelic acid (Fig. 24-14). **Glycine** is decarboxylated oxidatively in mitochondria in a sequence requiring lipoic acid and tetrahydrofolate as well as PLP (Fig. 15-20).[207–209b] A **methionine** decarboxylase has been isolated in pure form from a fern.[210] The bacterial **dialkylglycine** decarboxylase is both a decarboxylase and an aminotransferase which uses pyruvate as its second substrate forming a ketone and L-alanine as products (See Eq. 14-37)[210a, 210b]

***Side chain cleavage* (Group *c*).** In a third type of reaction the side chain of the Schiff base of Fig. 14-5 undergoes aldol cleavage. Conversely, a side chain can be added by β condensation. The best known enzyme of this group is **serine hydroxymethyltransferase**, which converts serine to glycine and formaldehyde.[211–213b] The latter is not released in a free form but is transferred by the same enzyme specifically to **tetrahydrofolic acid** (Eq. 14-30), with which it forms a cyclic adduct.



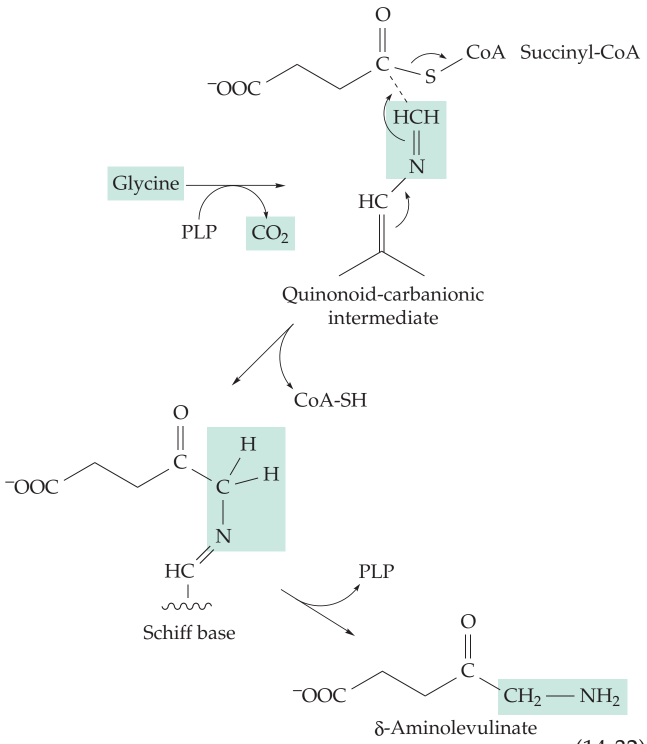
(14-30)

Threonine is cleaved to acetaldehyde by the same enzyme. A related reaction is indicated in Fig. 24-27 (top). In a more important pathway of degradation of threonine the hydroxyl group of its side chain is dehydrogenated to form 2-amino-3-oxobutyrate which is cleaved by a PLP-dependent enzyme to glycine and acetyl-CoA (Eq. 14-31).[214,215]



(14-31)

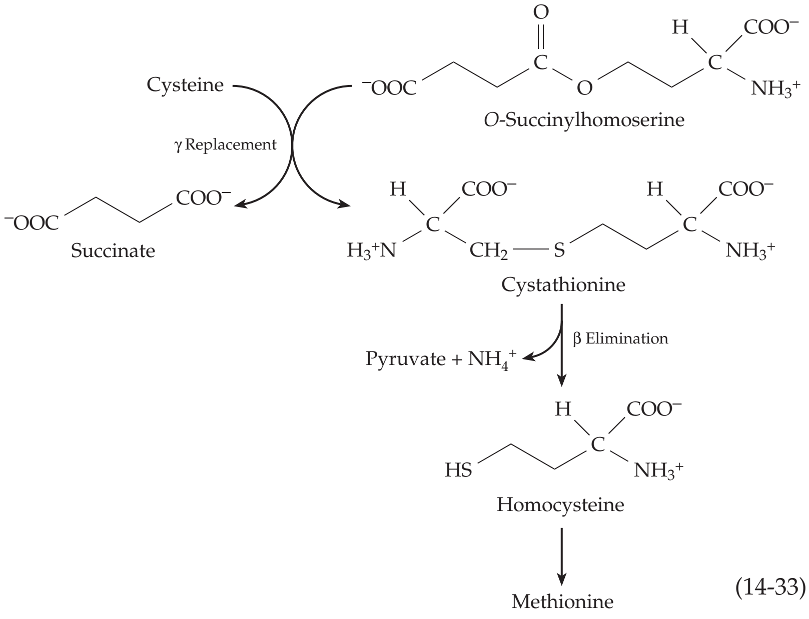
Conversely, ester condensation reactions join acyl groups from CoA derivatives to Schiff bases derived from glycine or serine. Succinyl-CoA is the acyl donor in Eq. 14-32 for the second known pathway for biosynthesis of **δ-aminolevulinic acid**, an intermediate in porphyrin synthesis (Chapter 24).[216 – 218b] The enzyme does not catalyze decarboxylation of glycine in the absence of succinyl-CoA, and the decarboxylation probably follows the condensation as indicated in Eq. 14-32.[219] In a similar reaction in the biosynthesis of **sphingosine** serine is condensed with palmitoyl-CoA and decarboxylated to form an aminoketone intermediate (Fig. 21-6).[219a] 8-Amino-7-oxonanonoate synthase forms a precursor of biotin (see banner, p. 719).[220]



(14-32)

***Ketimine intermediate as electron acceptor* (Group *d*).** The fourth group of PLP-dependent reactions are thought to depend upon formation of the ketimine intermediate of Eq. 14-28. In this form the original α-hydrogen of the amino acid has been removed and the $C=NH^+$ bond of the ketimine is polarized in a direction that favors electron withdrawal from the amino acid into the imine group. This permits another series of enzymatic reactions analogous to those of the β-oxo acid shown at the top of Fig. 14-5. Both elimination and C–C bond cleavage α,β to the C=N group of the ketimine can occur.

Enzymes of this group catalyze elimination of γ substituents from amino acids as illustrated in Fig. 14-7. Eliminated groups may be replaced by other substituents, either in the α or the β positions. The ketimine formed initially by such an enzyme (step *a*) undergoes elimination of the γ substituent (β with respect to the C=N group) along with a proton from the β position of the original amino acid to form an unsaturated intermediate which can react in one of three ways, depending upon the enzyme. Addition of HY' leads to γ replacement (step *c*), while addition of a proton at the α position leads, via reaction step *d*, to an α,β-unsaturated Schiff base. The latter can react by addition of HY' (β replacement, step *e*) or it can break down to an α-oxo acid and ammonium ion (step *f*), just as in the β elimination reactions of Eq. 14-29. An important γ replacement reaction is conversion of *O*-acetyl-, *O*-succinyl-, or *O*-phosphohomoserine to cystathionine (Eq. 14-33). This **cystathionine γ-synthase** reaction^{220a} lies on the pathway of biosynthesis of methionine by bacteria, fungi, and higher plants. Subsequent reactions include β elimination from cystathionine of **homocysteine**^{220b} which is then converted to methionine (Eq. 14-33). Threonine is formed from ***O*-phosphohomoserine** via γ elimination followed by β replacement with HO^-, a reaction catalyzed by **threonine synthase** (Fig. 24-13).^{220c}



(14-33)

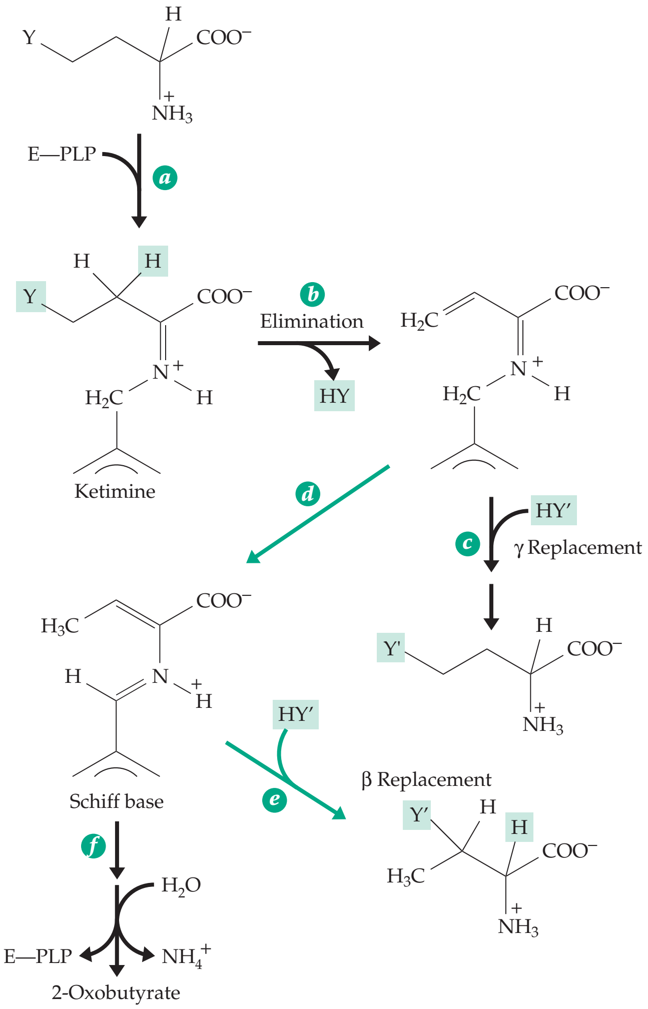
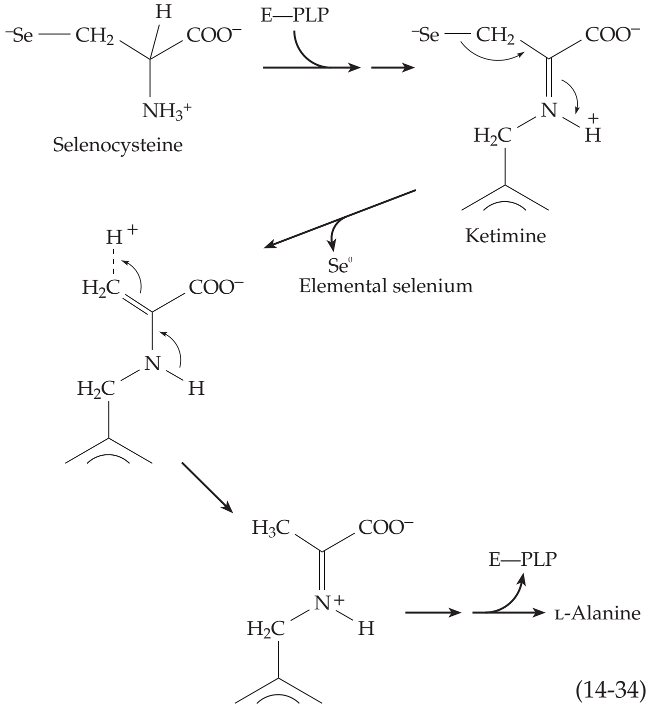


Figure 14-7 Some PLP-dependent reactions involving elimination of a γ substituent. Replacement by another γ substituent or by a substituent in the β position is possible, as is deamination to an α-oxo acid.

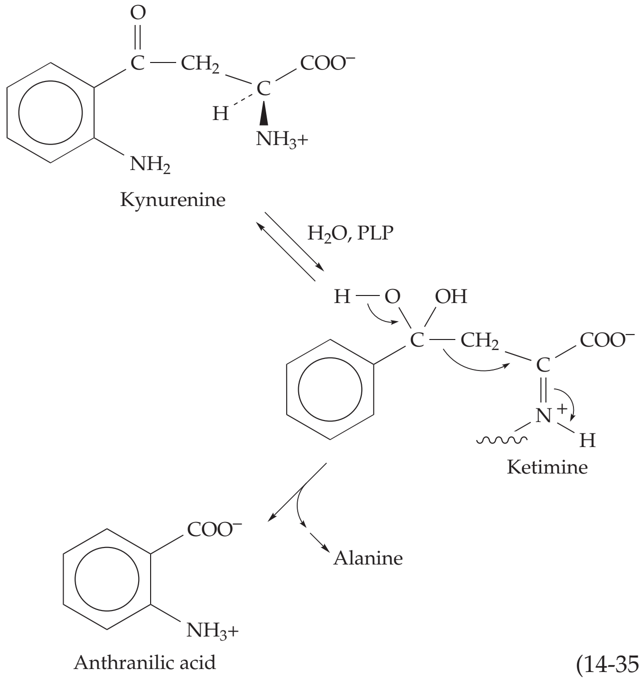
The loss of a β-carboxyl group as CO_2 can also occur through a ketimine or quinonoid intermediate. For example, the bacterial **aspartate β-decarboxylase**^{221} converts aspartate to alanine and CO_2. **Selenocysteine** is utilized to create the active sites of several enzymes (Chapter 15). Excess selenocysteine is degraded by the PLP-dependent **selenocysteine lyase**,$^{222–223b}$ which evidently eliminates elemental selenium from a ketimine or quinonoid state of an intermediate Schiff base (Eq. 14-34). A similar reaction may occur in the biosynthesis of iron–sulfur clusters. The **Nif S** protein is essential for formation of Fe_4S_4 clusters in the nitrogen-fixing enzyme nitrogenase. This enzyme is in some way involved in transferring the sulfur atom of cysteine into an iron–sulfur cluster.^{224} Alanine is the other product suggesting transfer of S^0 into the cluster using the sequence of Eq. 14-34.

Another related reaction that goes through a ketimine is the conversion of the amino acid **kynurenine** to alanine and anthranilic acid.^{225} It presumably depends upon hydration of the carbonyl group prior to β cleavage (Eq. 14-35). An analogous thiolytic cleavage utilizes CoA to convert 2-amino-4-ketopentanoate to acetyl-CoA and alanine.^{226}

Glycogen phosphorylase. While PLP is ideally designed to catalyze reactions of amino compounds it was surprising to find it as an essential cofactor for glycogen phosphorylase (Fig. 11-5). The PLP is linked as a Schiff base in the same way as in other PLP-depen-


⁻Se — CH₂
H
COO⁻
NH₃⁺
Selenocysteine
E—PLP
⁻Se — CH₂
COO⁻
N
H
H₂C
Ketimine
Se⁰
Elemental selenium
H⁺
H₂C
COO⁻
N
H
H₂C
H₃C
COO⁻
N⁺
H
H₂C
E—PLP
L-Alanine

(14-34)

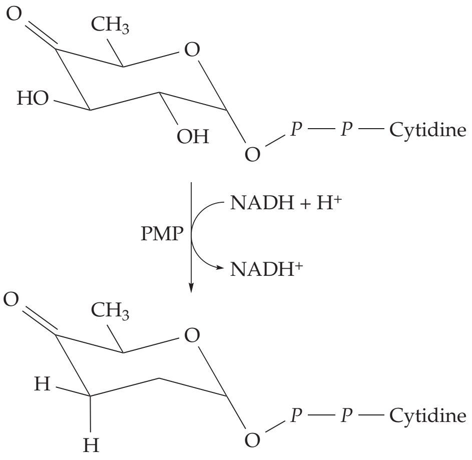

O
C — CH₂
C
COO⁻
H
NH₃+
NH₂
Kynurenine
H₂O, PLP
H — O
OH
C — CH₂
C
COO⁻
N⁺
H
Ketimine
COO⁻
NH₃+
Alanine
Anthranilic acid

(14-35)

dent enzymes, but there is no obvious function for the coenzyme ring. As suggested in Chapter 12, the phosphate group probably acts as an acid–base catalyst. It has been estimated that 50% of the vitamin B_6 in our body is present as PLP in muscle phosphorylase.[227] Studies of vitamin B_6-deficient rats suggest that PLP in phosphorylase serves as a reserve supply, much of which can be taken for other purposes during times of deficiency.

4. Pyridoxamine Phosphate as a Coenzyme

If PLP is a cofactor designed to react with amino groups of substrates, might not pyridoxamine phosphate (PMP) act as a coenzyme for reactions of carbonyl compounds? An example of this kind of function has been found[228–230] in the formation of 3,6-dideoxyhexoses needed for bacterial cell surface antigens (Fig. 4-15). Glucose (as cytidine diphosphate glucose; CDP-glucose) is first converted to 4-oxo-6-deoxy-CDP-glucose. The conversion of the latter to 3,6-dideoxy-CDP-glucose (Eq. 14-36) requires PMP as well as NADH or NADPH.


O
CH₃
O
HO
OH
O
P — P — Cytidine
PMP
NADH + H⁺
NADH⁺
O
CH₃
O
H
H
O
P — P — Cytidine

(14-36)

The student may find it of interest to propose a mechanism for this reaction, taking into account the expected direct transfer of a hydrogen from NADH as described in Chapter 15, before consulting published papers. Part of the reaction cycle appears to involve a free radical derived from the PMP. This is discussed further in Chapter 20 together with free radical-forming PLP enzymes

5. Stereochemistry of PLP-Requiring Enzymes

According to stereoelectronic principles, the bond in the substrate amino acid that is to be broken by a PLP-dependent enzyme should lie in a plane perpendicular to the plane of the cofactor–imine π system (Fig. 14-8). This would minimize the energy of the transition state by allowing maximum σ–π overlap between the breaking bond and the ring–imine π system. It also would provide the geometry closest to that of the planar quinonoid intermediate to be formed, thus minimizing molecular motion in the approach to the transition state. Figure 14-8 shows three orientations of an amino acid in which the α-hydrogen, the carboxyl group, and the side chain, respectively, are positioned for cleavage. For each orientation shown, another geometry suitable for cleavage of the same bond is

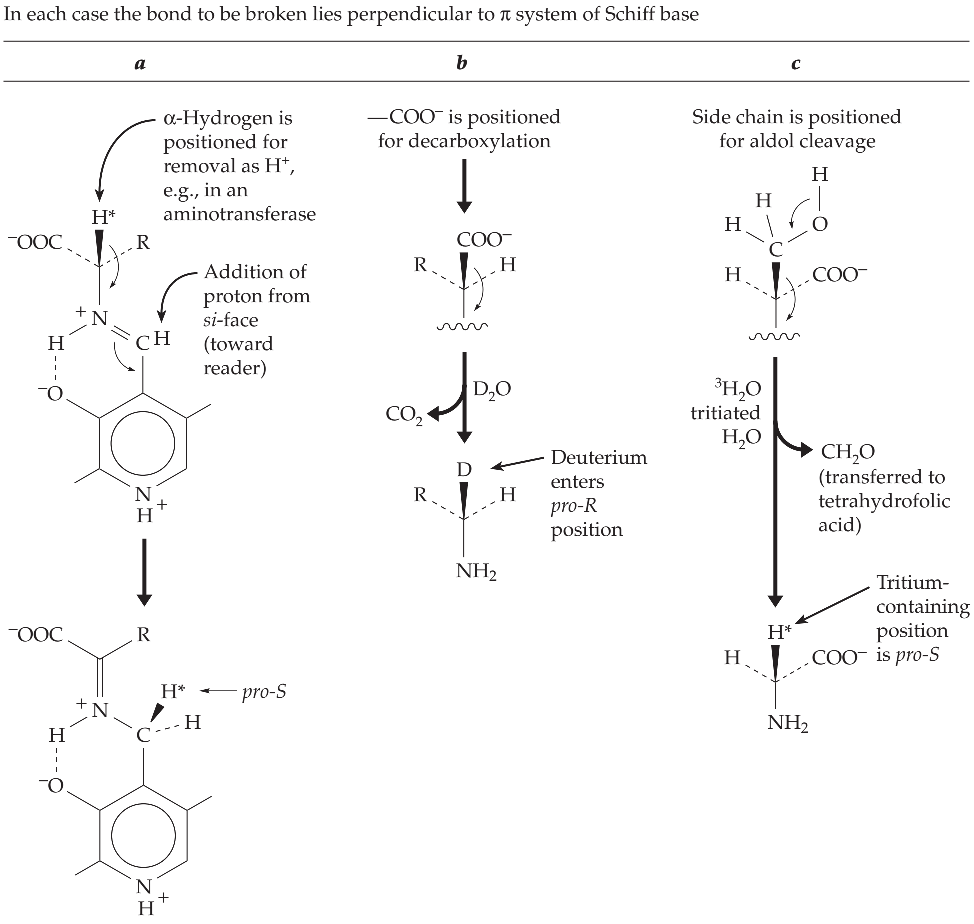
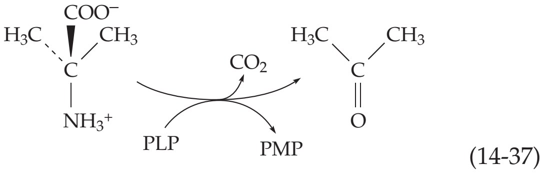

In each case the bond to be broken lies perpendicular to π system of Schiff base
a
b
c
α-Hydrogen is positioned for removal as H+, e.g., in an aminotransferase
H*
⁻OOC
R
Addition of proton from si-face (toward reader)
+N
H
C H
⁻O
N
H+
⁻OOC
R
H*
pro-S
+N
H
C
H
⁻O
N
H+
—COO⁻ is positioned for decarboxylation
COO⁻
R
H
D₂O
CO₂
D
R
H
NH₂
Deuterium enters pro-R position
Side chain is positioned for aldol cleavage
H
H
H
O
C
H
COO⁻
³H₂O tritiated H₂O
CH₂O (transferred to tetrahydrofolic acid)
Tritium-containing position is pro-S
H*
H
COO⁻
NH₂

Figure 14-8 Some stereochemical aspects of catalysis of PLP-requiring enzymes.

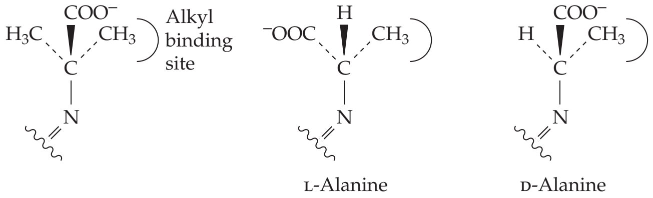
obtained by rotating the amino acid through 180°.

Dunathan suggested that this stereoelectronic requirement explains certain side reactions observed with PLP-requiring enzymes. The idea also received support from experiments with a bacterial α-dialkylglycinedecarboxylase.[231,232] The enzyme ordinarily catalyzes, as one half-reaction, the combination decarboxylation–transamination reaction shown in Eq. 14-37. It also acts on both D- and L-alanine, decarboxylating the former but catalyzing only removal of the α-H


COO⁻
H₃C
CH₃
C
NH₃⁺
CO₂
PLP
PMP
H₃C
CH₃
C
O

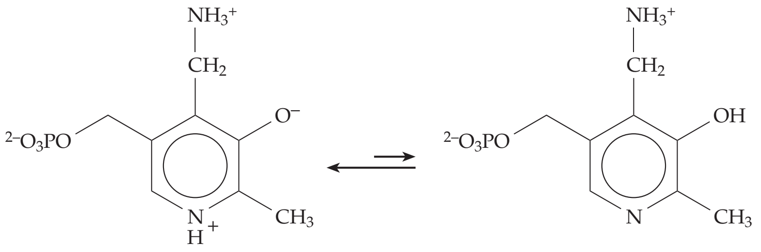
(14-37)

from L-alanine. The results can be rationalized by assuming that the enzyme possesses a definite site for one alkyl group but that the position of the second alkyl group can be occupied by –H or –COO⁻ and that the group labilized lies perpendicular to the π system:


COO⁻
H₃C
CH₃
Alkyl binding site
C
N
H
⁻OOC
CH₃
C
N
L-Alanine
COO⁻
H
CH₃
C
N
D-Alanine

Glycine is unreactive, suggesting that occupation of the alkyl binding site is required for catalysis.

According to Dunathan's postulate, there are only two possible orientations of the amino acid substrate in an aminotransferase. One is shown in Fig. 14-8. In the other, the amino acid is rotated 180° so that the α-hydrogen protrudes *behind* the plane of the paper. Dunathan studied **pyridoxamine:pyruvate aminotransferase**, an enzyme closely related to PLP-requiring aminotransferases and which catalyzes the transamination of pyridoxal with L-alanine to form pyridoxamine and pyruvate. The same reaction is catalyzed by the apoenzyme of aspartate aminotransferase. In both cases, when the alanine contained ^{2}H in the α position the ^{2}H was transferred stereospecifically into the *pro-S* position at C-4' of the pyridoxamine (indicated by asterisks in Fig. 14-8).[233] The results suggested that a group from the protein abstracts a proton from the α position and transfers it on the same side of the π system (**suprafacial transfer**), adding it to the *si* face of the C=N group as shown in Fig. 14-8. Later, the same stereospecific proton transfer was demonstrated for the PLP present in the holoenzyme.[234] Not surprisingly, the D-amino acid aminotransferase adds the proton to the *re* face of the C=N group.[235]



Dipolar ionic form
I 327 nm
II 253 nm
III 217 nm

Minor non-dipolar ionic tautomer
I 283 nm

(14-38)

When a decarboxylase acts on an amino acid in ^2H_2O, an atom of ^2H is incorporated in the *pro-R* position, the position originally occupied by the carboxyl group (Fig. 14-8). Cleavage of serine by serine hydroxymethyltransferase in ^3H-containing water leads to incorporation of ^3H in the *pro-S* position. Stereospecific introduction of ^2H or ^3H has been observed in the β position of 2-oxobutyrate formed in β or γ elimination reactions. Conversion of serine to tryptophan by tryptophan synthetase occurs without inversion at C-3.[236] These and many other observations on PLP-dependent enzymes[189,237,238] can be generalized by saying that enzymatic reactions of PLP Schiff bases usually take place on only one face of the relatively planar structure. This is the *si* face at C-4' of the coenzyme (see Fig. 14-8). This result is expected if a single acid–base group serves as proton acceptor in one step and as proton donor in a later step. This leads naturally to the observed retention of configuration in steps involving replacement and the suprafacial transfer of protons from one position on that face to another.

6. Seeing Changes in the Optical Properties of the Coenzyme

The absorption of light in the ultraviolet and visible regions is a striking characteristic of many coenzymes. It can be measured accurately and displayed as an absorption spectrum and may also give rise to circular dichroism and to fluorescence (Chapter 23). The optical properties of the vitamin B_6 coenzymes are sensitive to changes both in environment and in the state of protonation of groups in the molecule. For example, PMP in the neutral dipolar ionic form, which exists at pH 7, has three strong light absorption bands centered at 327, 253, and 217 nm.[239] The other ionic forms of PMP and other derivatives of vitamin B_6 also each have three absorption bands spaced at roughly similar intervals, but with varying positions and intensities. The minor tautomer of PMP containing an uncharged ring (Eq. 14-38) has its low-energy (long-wavelength) band at 283 nm. When both the ring nitrogen and phenolic oxygen are protonated, the band shifts again to 294 nm and if both groups are deprotonated the resulting anion absorbs at 312 nm. Thus, observation of the absorption spectrum of the coenzyme bound to an enzyme surface can tell us whether particular groups are protonated or unprotonated. The peak of bound PMP at 330 nm in aspartate aminotransferase (Fig. 14-9) is indicative of the dipolar ionic form (Eq. 14-38). However, the 5-nm shift from the position of free PMP suggests a distinct change in environment.

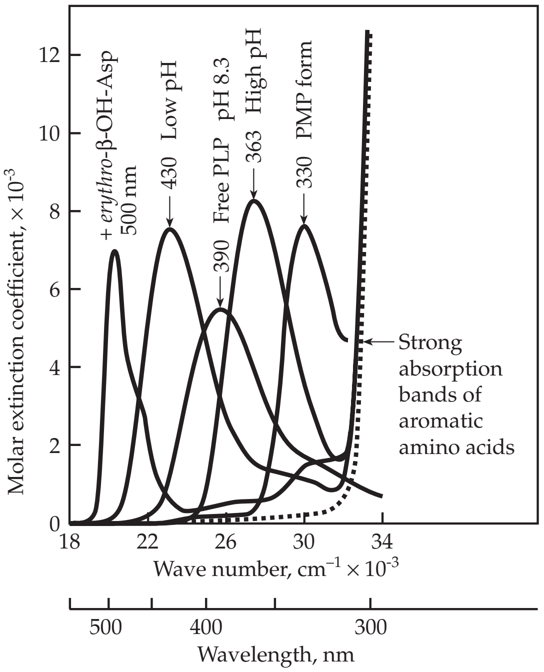
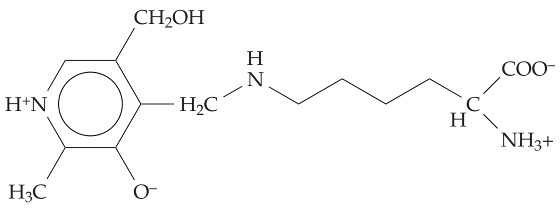


Figure 14-9 Absorption spectra of various forms of aspartate aminotransferase compared with that of free pyridoxal phosphate. The low pH form of the enzyme observed at pH <5 is converted to the high pH form with pK_a ~6.3. Addition of *erythro*-3-hydroxyaspartate produces a quinonoid form whose spectrum here is shown only 1/3 its true height. The spectrum of free PLP at pH 8.3 is also shown . The spectrum of the apoenzyme (- - -) contains a small amount of residual absorption of uncertain origin in the 300- to 400-nm region.

Pyridoxal phosphate exists in an equilibrium between the aldehyde and its covalent hydrate (as in Eq. 13-1). The aldehyde has a yellow color and absorbs at 390 nm (Fig. 14-9), while the hydrate absorbs at nearly the same position as does PMP. The absorption bands of Schiff bases of PLP are shifted even further to longer wavelengths, with N-protonated forms absorbing at 415–430 nm. Forms with an unprotonated C=N group absorb at shorter wavelengths.[149,240]

Imine groups in free enzymes. When, in 1957, W. T. Jenkins examined one of the first highly purified aspartate aminotransferase preparations he noted a surprising fact: The bound coenzyme, at pH 5, absorbed not at 390 nm, as does PLP, but at 430 nm, like a Schiff base (Fig. 14-9). When the pH was raised, the absorption band shifted to 363 nm. The result suggests dissociation of a proton (with a pK_a of ~ 6.3) from the hydrogen-bonded position in a Schiff base of the type shown in Fig. 14-6. It was quickly demonstrated for this enzyme and for many other PLP-dependent enzymes that reduction with sodium borohydride caused the spectrum to revert to one similar to that of PMP and fixed the coenzyme to the protein. After complete HCl digestion of such borohydride-reduced proteins a fluorescent amino acid containing the reduced pyridoxyl group was obtained and in every case was identified as ε-pyridoxyl-lysine. Thus, PLP-containing enzymes in the absence of substrates usually exist as Schiff bases with lysine side chains of the proteins. Even the PLP in glycogen phosphorylase is joined in this way. However, its absorption maximum at 330 nm shows that in phosphorylase it is present as the nondipolar ionic tautomer with a 3-OH group on the ring as in Eq. 14-38.



ε-Pyridoxyllysine

Absorption bands at 500 nm. With many PLP enzymes certain substrates and inhibitors cause the appearance of intense and unusually narrow bands at ~500 nm. Such a band is observed with aspartate aminotransferases acting on *erythro*-3-hydroxyaspartate (Fig. 14-9). This substrate undergoes transamination very slowly, and the 500-nm absorbing form which accumulates is probably an intermediate in the normal reaction sequence. A similar spectrum is produced by tryptophan indole-lyase acting on the competitive inhibitor L-alanine. Under the same conditions the enzyme promotes a rapid exchange of the α-hydrogen of the alanine with ^2H of ^2H_2O. Serine hydroxymethyltransferase gives a 495- to 500-nm band with both D-alanine and the normal product glycine.[241] Similar spectra have been produced in nonenzymatic model reactions[242] and probably represent the postulated quinonoid–carbanionic intermediates.

7. Atomic Structures

The three-dimensional structures of aspartate aminotransferases from *E. coli* to humans are very similar.[163–167a] The folding pattern (Fig. 2-6) and active site structure (Figs. 14-6 and 14-10) are completely conserved. The major domain of the protein contains a central β sheet surrounded by helices with coenzyme attached to a lysine at the C terminus of one of the β strands. The protein is a dimer with the two major domains held together by both polar and nonpolar interactions. The two active sites are located at the interface between the subunits and residues from both subunits participate in forming the active site (Fig. 14-10). The internal Schiff base is formed with Lys 258. The protonated ring nitrogen of the dipolar ionic PLP forms an ion pair with the carboxylate of Asp 222 which protrudes from a central seven-stranded β sheet. The phenolic $-O^-$ forms a hydrogen bond with the –OH of Tyr 225. The interactions of Asp 222 and Tyr 225 fix the ring as the dipolar ionic tautomer.

The phosphate group of the coenzyme, which ^{31}P NMR shows to be predominantly dianionic,[243] forms an ion pair with the side chain of Arg 266 and hydrogen bonds to a backbone N–H at the N terminus of a long helix where it can interact with the positive end of the helix dipole. In addition, the phosphate forms hydrogen bonds to four OH groups of Ser, Thr, and Tyr side chains. In front of the coenzyme ring are two guanidinium groups from the side chains of Arg 386 and of Arg 292* (from the second subunit). These have been shown by X-ray crystallography to bind the two carboxylate groups of a substrate such as glutamate, 2-oxoglutarate, aspartate, oxaloacetate, or cysteine sulfinate; of quasi-substrates such as 2-methylaspartate and *erythro*-β-hydroxyaspartate; or of dicarboxylic inhibitors (Fig. 14-10).

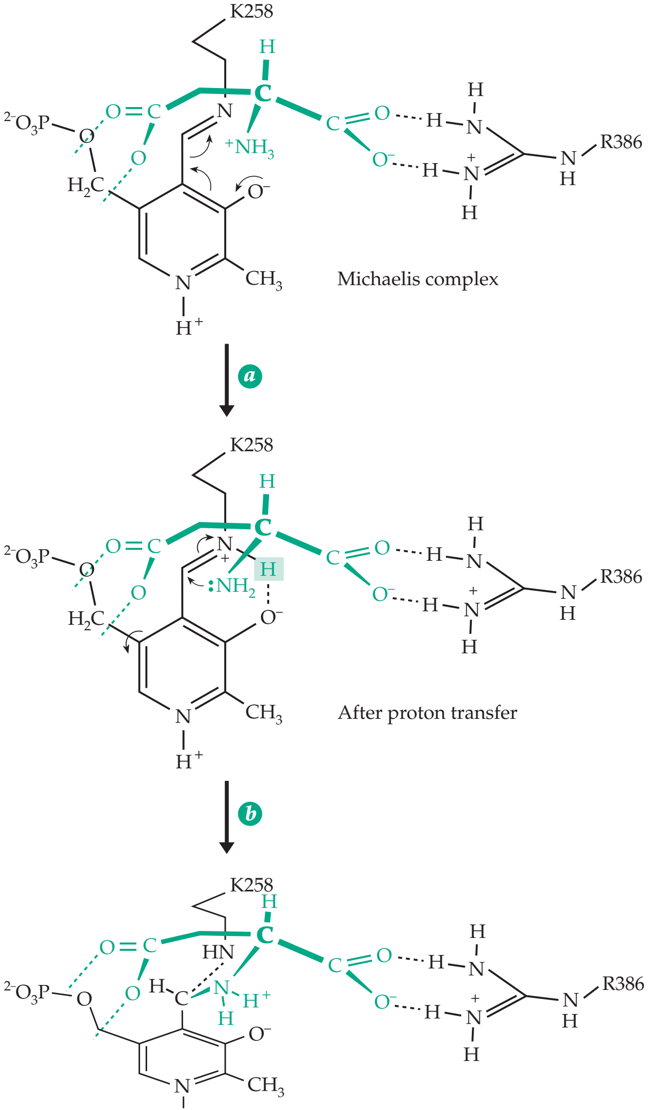
Comparison of amino acid sequences suggests that many other PLP-dependent enzymes have folding patterns similar to those of aspartate aminotransferase but that there are four or more additional different folding patterns.[205,243a,b] Among the enzymes resembling aspartate aminotransferase are ω-amino acid: pyruvate aminotransferase,[244] 2,3-dialkylglycine decarboxylase,[232] tyrosine phenol-lyase,[177] a bacterial ornithine decarboxylase,[182,204] and cystathionine β-lyase.[182] The tryptophan synthase β subunit has a second folding pattern,[184] while alanine racemase[153] and eukaryotic ornithine decarboxylases[205] have $(\alpha\beta)_8$-barrel structures resembling that in Fig. 2-28. A fourth structural pattern is that of D-amino acid aminotransferase.[235] It is anticipated that the branched-chain aminotransferase[245] will have a similar stucture. Glycogen phosphorylase (Fig 11-5) has a fifth folding pattern.

8. Constructing a Detailed Picture of the Action of a PLP Enzyme

Consider the number of different steps that must occur in about one-thousanths of a second during the action of an aminotransferase. First, the substrate binds to form the "Michaelis complex." Then the transimination (Eq. 14-26) takes place in two steps and is followed by the removal of the α-hydrogen to form the quinonoid intermediate. An additional four steps are needed to form the ketimine, to hydrolyze it, and to release the oxoacid product to give the PMP form of the enzyme. The reaction sequences in some of the other enzymes are even more complex. How can one enzyme do all this?

The first step in the sequence is the binding of the substrate to form the "Michaelis complex." The positive charges on Arg 386 and Arg 292 doubtless attract the carboxylate groups of the substrate and aid in guiding it toward a correct fit. In a similar manner the $-O^-$ of the coenzyme, which is distributed by resonance into the $-C=N$ of the Schiff base linkage, attracts the $-NH_3^+$ of the substrate. When a substrate or inhibitor binds to the two guanidinium groups a small structural domain of the enzyme moves and closes around the substrate which now has very little contact with the external solvent. In the initial "Michaelis complex" the $-NH_3^+$ group of the substrate lies directly in front of the C-4' carbon of the coenzyme (Fig. 14-10), where it can initiate the transimination reaction of Eq. 14-26. However, before this can happen a proton must be removed to convert the $-NH_3^+$ to $-NH_2$.

Long before the three-dimensional structures were known, Ivanov and Karpeisky[246] suggested that in the free enzyme the positively charged group (Arg 386) that binds the α-carboxylate of the substrate interacts electrostatically with the $-O^-$ of the coenzyme. This is one of the factors that keeps the pK_a of the $-CH=N^+H-$ that is conjugated with this $-O^-$ at a low value of ~6.3 (at 0.1 M anion concentration). However, in the Michaelis complex this interaction of the + charge of Arg 386 with the imine group must be weakened because of the pairing of the $\alpha-COO^-$ of the substrate with the + charge. This will increase the basicity of the imine nitrogen and will also cause a decrease in the pK_a of the substrate $-NH_3^+$, making is easier for a proton to jump from the $-NH_3^+$ to the imine group. This proton transfer is shown in Eq. 14-39, step *a*. Thus, the nucleophilic $-NH_2$ group is generated by a process that at the same time increases the electrophilic properties of the carbon atom of the imine group. This favors the immediate addition of $-NH_2$ to $-C=N^+H-$ to give the adduct a **geminal diamine** (Eq. 14-39, step *b*), which is shown in three dimensions in Fig. 14-10B.



(14-39)

Notice that each step in the overall sequence changes the electronic or steric characteristics of the complex in a way that facilitates the next step.[246] This is an important principle that is applicable throughout enzymology: *For an enzyme to be an efficient catalyst each step must lead to a change that sets the stage for the next*. These consecutive steps often require proton transfers, and each such transfer will influence the subsequent step in the sequence. Some steps also require alterations in the conformation of substrate, coenzyme, and enzyme. One of these is the transimination sequence (Eqs. 14-26, 14-39). On the basis of the observed loss of circular dichroism in the external aldimine, Ivanov and Karpeisky suggested that a

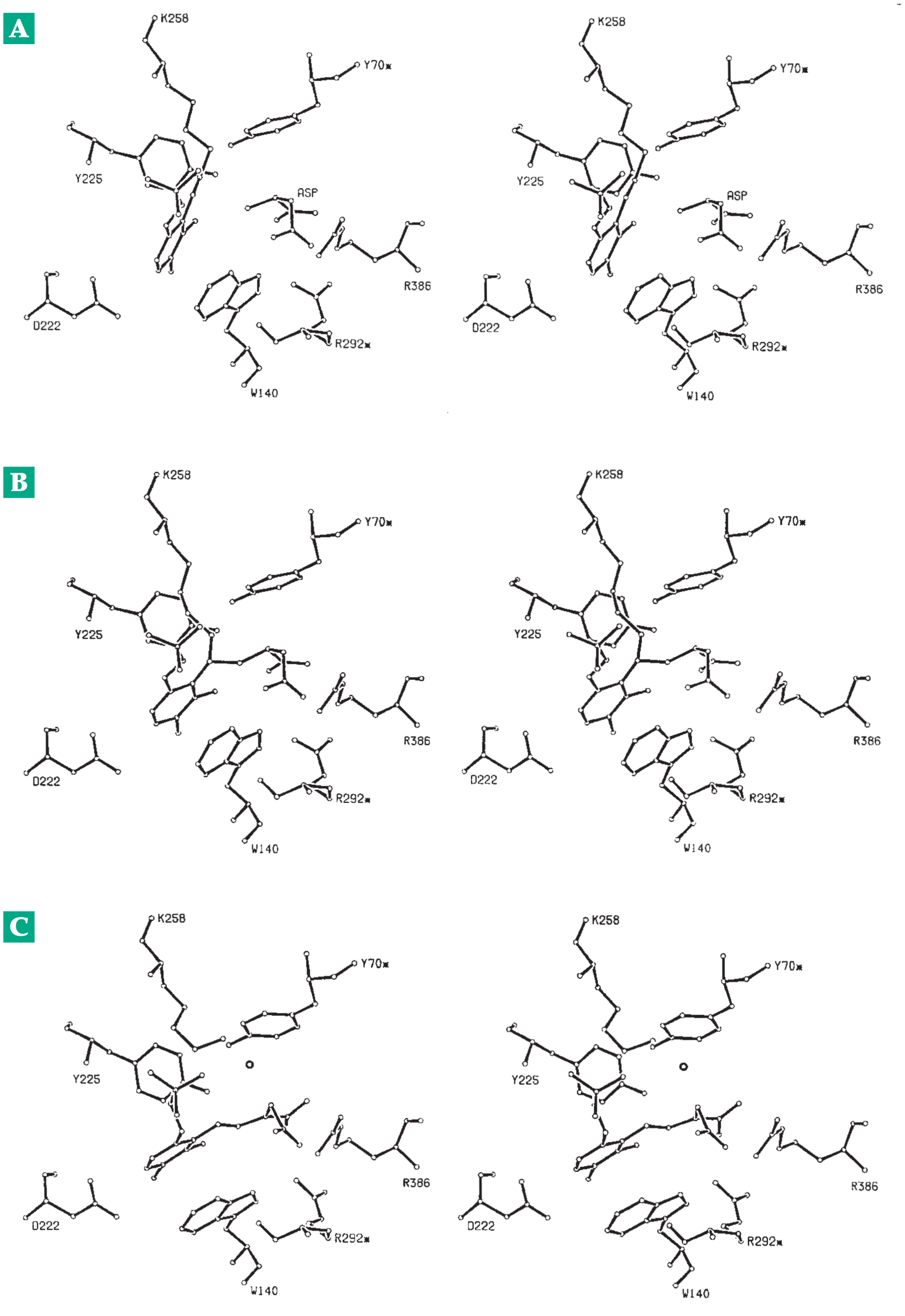


Figure 14-10 Models of catalytic intermediates for aspartate aminotransferase in a half-transamination reaction from aspartate to oxalocetate. For clarity, only a selection of the active site groups are shown. (A) Michaelis complex of PLP enzyme with aspartate. (B) Geminal diamine. (C) Ketimine intermediate. The circle indicates a bound water molecule. See Jansonius and Vincent in Jurnak and McPherson.[163] Courtesy of J.N. Jansonius.

rotation of the coenzyme occurs as the $-NH_2$ of the substrate adds to the C = N bond during transimination (Eq. 14-39, step *b*). This accomplishes the essential shortening by ~0.15 nm of the distance between C-4' and the N atom from a van der Waals contact distance to a covalent bond distance while the carboxylate groups of the substrate remain bound in their initial positions.

That the coenzyme really does change its orientation was suggested by a dramatic change in the absorption spectrum of a crystal recorded with plane polarized light (linear dichroism) when 2-methylaspartate was soaked into a crystal (Fig. 23-9).[247] X-ray crystallography confirmed rotation of the ring by ~30°.[163,242,247] To complete the transimination sequence, which is shown only partially by Eq. 14-39, another proton transfer is needed to move the positive charge on the substrate $-N^+H_2-$ to that of lysine 258, whose amino group is then eliminated. This requires additional tilting of the ring. The crystal structure of the external aldimine with α-methylaspartate has been determined[248] as have those of ketimines with glutamate and aspartate,[249,250] a carbinolamine,[251] and quinonoid complexes of related enzymes.[212a]

The ε-amino group, which is eliminated in Eq. 14-26, is basic and functions in the next several steps of catalysis, including the abstraction of the α-hydrogen and in its transfer to the 4'-carbon (Eq. 14-28). This amino group can be seen in Fig. 14-10C, where it is positioned beside a water molecule that is needed to hydrolyze the ketimine. From this figure it can be seen that the group is able to move from site to site on one side of the planar external aldimine, quinonoid–carbanionic, and nearly planar ketimine forms. If the Lys 258 amino group is the catalytic base for these reactions it must be present at pH 7. What is its pK_a? And why doesn't this pK_a show up in the plot of V_{max} vs pH as in Fig 9-9? In fact, the maximum velocity of aspartate aminotransferase and many other PLP-dependent enzymes is independent of pH over a broad range. The answer is probably that the pK_a's of both the $-NH_2$ groups of Lys 258 and the amino acid in the Michalis complex are high when there is no positive charge on the adjacent Schiff base $-C=N-$ but very low when the Schiff base is protonated (–C=NH$^+$–). This is a result of the large electrostatic effect of a closely adjacent charge in a medium of low dielectric constant (see Chapter 7, Section A). When the Schiff base is unprotonated, as in the first structure of Eq. 14-39, the adjacent amino group has a *high* pK_a and is mostly protonated. However, the proton can jump as in step *a* of that equation to give a protonated Schiff base. Now the unprotonated –NH$_2$ group has a very *low* pK_a, but it is still strongly nucleophilic and can readily add to the Schiff base double bond in step *b* of Eq. 14-39. In a similar fashion the amino group of Lys 258 will alternately be unprotonated and then protonated, its microscopic pK_a alternating between low and high, as it catalyzes the steps of abstracting the α-hydrogen and forming and hydrolyzing the ketimine. Only two microscopic pK_a values, one very high and one very low appear in the V_{max} vs pH profile. This alternation of microscopic pK_a values may be a common characteristic of enzymes that bind ionized substrates and can make good use of the strong electrostatic effects that arise in the active sites to facilitate essential proton transfers.

Aspartate aminotransferases are distinguished from most other PLP-dependent enzymes including transaminases by the relatively low pK_a of ~6.3 for the *free* enzyme. The unprotonated Schiff base in the free enzyme can then react with the protonated amino group of the substrate. How can other PLP enzymes react with amino acids when they have protonated Schiff bases even at relatively high pH? A logical answer is that some basic group with a low pK_a is close to the Schiff base and acts to deprotonate the substrate –NH$_3^+$ so that transimination can occur. Clausen *et al*. suggested that in cystathionine β-lyase this is probably tyrosine (Y111), which is adjacent to the Schiff base –C=NH$^+$– of the external aldimine and is thought to be ionized in the active enzyme.[182] NMR evidence suggests that protonation of an adjacent catalytic base occurs upon substrate binding in D-serine dehydratase as well.[252] Many other variations in active site environments are seen among PLP-dependent enzymes. As with aspartate aminotransferases, there is often an essential carboxylate group that holds a proton onto the pyridine ring of the coenzyme. However, in alanine racemase a guanidinium group from arginine is hydrogen bonded to the coenzyme ring.[153]

Below the active site of aspartate aminotransferase, as shown in Fig. 14-6, is a cluster of three buried histidine side chains in close contact with each other. The imidazole of H143 is hydrogen bonded to the D222 carboxylate, the same carboxylate that forms an ion pair with the coenzyme. This system looks somewhat like the catalytic triad of the serine proteases in reverse. As with the serine proteases, the proton-labeled H$_b$ in Fig. 14-6 can be “seen” by NMR spectroscopy (Fig. 3-30). So can the proton H$_a$ on the PLP ring. These protons act as built-in sensors able to detect small changes in the electronic environment. For example, when the Schiff base proton dissociates around the pK_a of ~6.2 the NMR resonance of H$_a$ shifts upfield from 17.2 to 15.2 ppm as a result of donation of electons into the ring from the –O$^-$ of the coenzyme.[169] This shift illustrates the reality of the strong electrostatic forces that operate across heterocyclic aromatic rings within active sites of proteins. In alanine racemase a different histidine cluster is present beneath the active site and constitutes part of the “solvent” in which the catalyzed reaction takes place. At least in the case of aspartate aminotransferase, none of the histidines are absolutely essential for activity but the hydrogen-bonded network, which can be altered in mutant forms, may be important.

The detailed description of a reaction sequence given here has to be altered for each specific enzyme. A vast amount of work, only a little of which is cited here, has been done on PLP enzymes.[253–254a] These studies involve calorimetry,[255] kinetics,[210,256–259] crystallography, optical spectroscopy,[188,260,261] NMR,[243,262–264] and genetic engineering and chemical modification.[178,185,186,265]

F. Pyruvoyl Groups and Other Unusual Electrophilic Centers

A few enzymes that might be expected to have PLP at their active sites have instead a prosthetic group consisting of pyruvic acid bound by an amide linkage, a **pyruvoyl group** (Table 14-4). These and several apparently related enzymes are the subject of this section.

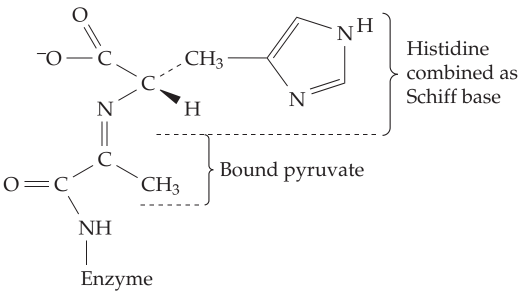
TABLE 14-4
Some Pyruvoyl Enzymes

Decarboxylases	Product
Histidine (bacterial)	
S-Adenosylmethionine	
Aspartate α- decarboxylase	β-Alanine
Phosphatidylserine	Phosphatidyl-ethanolamine
4' - Phosphopantothenylcysteine	4' - Phosphopan-tetheine
Reductases (clostridial)	
Proline	
Glycine	

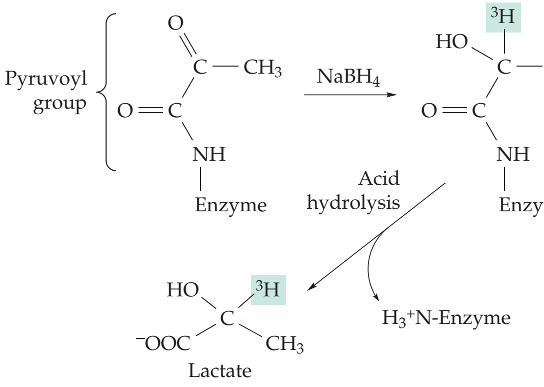
Adapted from van Peolje and Snell.[267]

1. Decarboxylases

Mammalian **histidine decarboxylase** contains PLP but the enzyme from many bacteria contains a pyruvoyl group, as do a few other decarboxylases both of bacterial and eukaryotic origin.[266,267] These enzymes are inhibited by carbonyl reagents and by borohydride. When ^3H-containing borohydride was used to reduce the histidine decarboxylase of *Lactobacillus*, ^3H was incorporated and was recovered in lactic acid following hydrolysis. This suggested the presence of a pyruvoyl group attached by an amide linkage and undergoing the chemical reactions that are shown in Eq. 14-40. Reduction in the presence of the substrate histidine resulted in covalent binding of the histidine to the bound pyruvate. Thus, as with the PLP-containing decarboxylases, a Schiff base is formed with the substrate. Decarboxylation is presumably accomplished by using the electron-attracting properties of the carbonyl group of the amide:



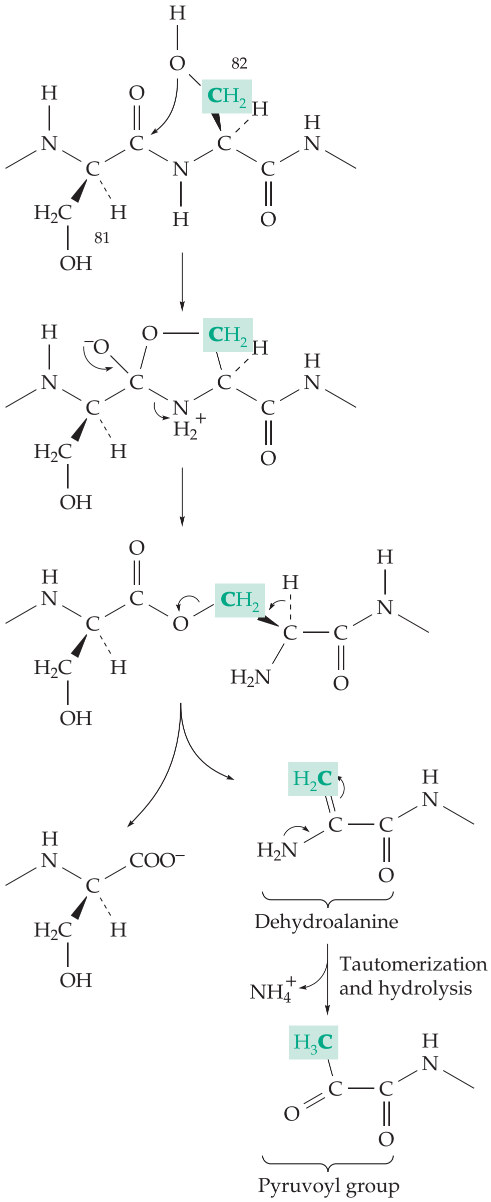
When ^{14}C-labeled serine was fed to organisms producing histidine decarboxylase, ^{14}C was incorporated into the bound pyruvoyl group (Fig. 14-11). Thus, serine is a precursor of the bound pyruvate. The enzyme is manufactured in the cell as a longer 307-residue proenzyme which associates as hexamers (designated π_6). The active enzyme was found to be formed by cleavage of the π chains between Ser 81 and Ser 82 to form 226-residue α chains and 81-residue β chains which associate as $(\alpha\beta)_6$.[270,271] The α chains carry the N-terminal pyruvoyl group, which was formed from Ser 82. The activation occurs spontaneously by incubations of the proenzyme for 24–48 h at pH ~7 in the presence of divalent metal ions.



(14-40)

Substituted serine residues under mildly alkaline conditions readily undergo α,β elimination to form **dehydroalanine** residues. When prohistidine decarboxylase containing ^{18}O in its serine side chains was activated ^{18}O was found in the carboxylate group of Ser 81 of the β chains. It was shown that it had been transferred from the side chain of Ser 82. This suggested the formation of an intermediate oxygen ester of Ser 81 during formation of the pyruvoyl group (Eq. 14-41).[272]

***S*-Adenosylmethionine decarboxylase** is the first enzyme in the biosynthetic pathway to spermidine (Chapter 24). Whether isolated from bacteria, yeast, animals, or other eukaryotes, this enzyme always contains a bound pyruvoyl group.[273–274b] Both the



(14-41)

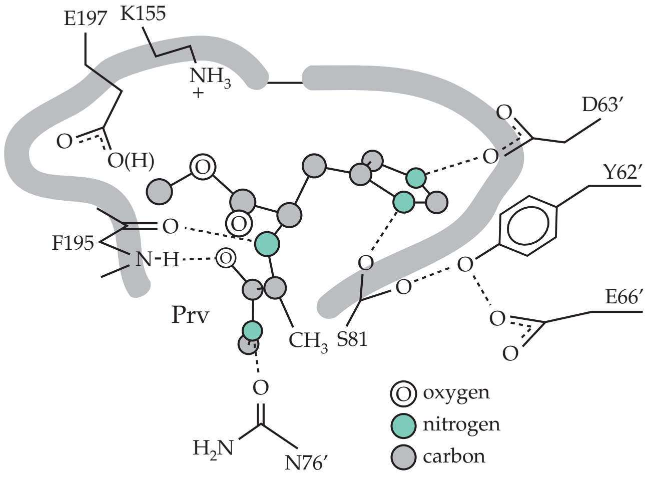
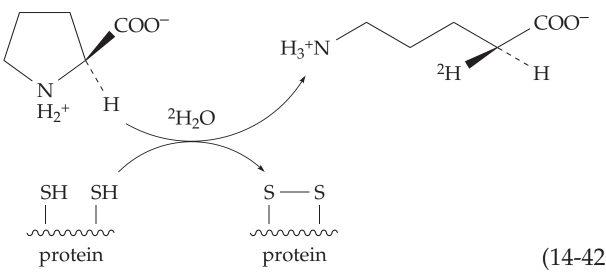


Figure 14-11 Schematic diagram of the active site of the pyruvoyl enzyme histidine decarboxylase showing key polar interactions between the pyruvoyl group and groups of the inhibitor *O*-methylhistidine and surrounding enzyme groups. Aspartate 63 appears to form an ion pair with the imidazolium group of the substrate.[268] Hydrogen bonds are indicated by dotted lines. See Gallagher *et al.*[269]

mammalian enzyme and that from *E. coli* are $(\alpha\beta)_4$ tetramers formed in a manner similar to that of bacterial histidine decarboxylase. Other pyruvoyl decarboxylases are **phosphatidylserine decarboxylase**,[275,276] an intrinsic membrane protein used to form phosphatidylethanolamine, **aspartate α-decarboxylase**,[277] which forms β alanine needed for biosynthesis of coenzyme A, and **4'-phosphopantethenoylcysteine decarboxylase**,[278] the second of two decarboxylases required in the synthesis of coenzyme A. Because of the lack of a primary amino group, its mechanism must be somewhat different from that of other enzymes in this group.[267]

2. Proline and Glycine Reductases

An enzyme required in the anaerobic breakdown of proline by clostridia utilizes a dithiol-containing protein to reductively open the ring (Eq. 14-42).[279,280]



(14-42)

This enzyme also contains an N-terminal pyruvoyl residue as does one subunit of a selenium-containing glycine reductase which utilizes a dithiol to convert glycine into acetate with coupled formation of ATP:

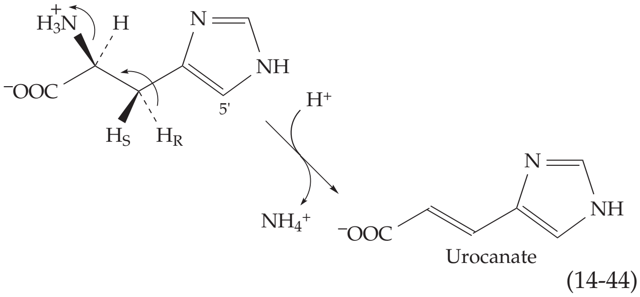
$$\begin{aligned}&\text{Glycine} + \text{ADP} + \text{P}_i + 2\,\text{e}^- \rightarrow \\ &\quad \text{acetate}^- + \text{NH}_4^+ + \text{ATP}\end{aligned} \tag{14-43}$$

In both cases it has been proposed that the pyruvoyl group forms a positively charged Schiff base with the substrate. The bond-breaking mechanisms are not obvious but some ideas have been proposed.[281–283a] The ATP may arise from reaction with an intermediate acetyl phosphate.[282] See Eq. 15-61.

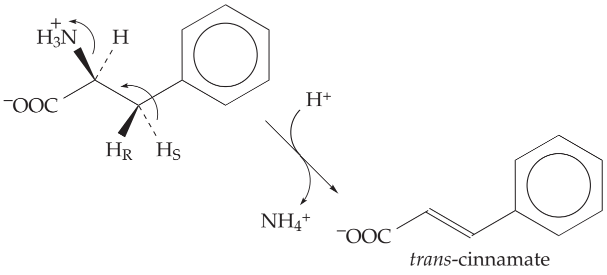
3. Dehydroalanine and Histidine and Phenylalanine Ammonia-Lyases

Catabolism of histidine in most organisms proceeds via an initial elimination of NH_3 to form **urocanic acid** (Eq. 14-44). The absence of the enzyme **L-histidine ammonia-lyase** (histidase) causes the genetic disease histidinemia.[284,285] A similar reaction is catalyzed by the important plant enzyme **L-phenylalanine ammonia-lyase**. It eliminates $-NH_3^+$ along with the *pro-S* hydrogen in the β position of phenylalanine to form *trans*-cinnamate (Eq. 14-45). Tyrosine is converted to *p*-coumarate by the same enzyme. Cinnamate and coumarate are formed in higher plants and are converted into a vast array of derivatives (Box 21-E, Fig. 25-8).

The reactions of Eqs. 14-44 and 14-45 are unusual because the nucleophilic substituent eliminated is on the α-carbon atom rather than the β. There is nothing in the structures of the substrates that would permit an



(14-44)



(14-45)

easy elimination of the α-amino group. Thus, it should not be surprising that both enzymes contain a special active center. When ^{2}H is introduced into the β position of phenylalanine, no isotope effect on the rate is observed. Rather, the rate-limiting step appears to be the release of ammonia from the coenzyme group. It appears that the enzyme must in some way make the amino group a much better leaving group than it would be otherwise.

Both enzymes are inhibited by sodium borohydride and also by nitromethane. After reduction with NaB^{3}H$_4$ and hydrolysis, ^{3}H-containing alanine was isolated. This suggested that they contain **dehydroalanine**, which could arise by dehydration of a specific serine residue.[286,287] For phenylalanine ammonia-lyase from *Pseudomonas putida* this active site residue has been identified as S143. Replacement by cysteine in the S143C mutant also gave active enzyme while S143A

BOX 14-D DIETARY REQUIREMENTS FOR B VITAMINS

It is difficult to determine the amounts of vitamins needed for good health and they may differ considerably from one individual to another. The quantities listed below are probably adequate for most young adults but must be increased during pregnancy and lactation and after very strenuous exercise.

Pantothenic acid: 10–15 mg/day. Deficiency causes apathy, depression, impaired adrenal function, and muscular weakness. ω -Methylpantothenic acid is a specific antagonist. The calcium salt, calcium pantothenate, is the usual commercial form.

Biotin: 0.15–0.3 mg/day. The discovery that biotin deficiency in young chickens can lead to sudden death resulted in a recommendation to supplement infant formulations with biotin.[a] **Desthiobiotin**, in which the sulfur has been removed and replaced by two hydrogen atoms, can replace biotin in some organisms and appears to lie on one pathway of biosynthesis.[b,c] **Oxybiotin,** in which the sulfur has been replaced by oxygen, is active for many organisms and partially active for others. No evidence for conversion to biotin itself has been reported, and oxybiotin may function satisfactorily in at least some enzymes.

Thiamin: 0.23 mg or more per 1000 kcal of food consumed and a minimum total of 0.8 mg/day. Replacement of the methyl group on the pyrimidine ring by ethyl, propyl, or isopropyl gives compounds with some vitamin activity, but replacement by hydrogen cuts activity to 5% of the original. The butyl analog is a competitive inhibitor.

Vitamin B$_6$: 1.5–2 mg/day; 0.4 mg/day for infants. Vitamin B$_6$ is widely distributed in foods, and symptoms of severe deficiency are seldom observed. However, a number of cases of convulsions have been attributed to partial destruction of vitamin B$_6$ in infant liquid milk formulas. Convulsions occurred when the vitamin B$_6$ content was reduced to about one-half that normally present in human milk.

Several cases of children with an abnormally high vitamin B$_6$ requirement (2–10 mg/day) have been reported, and rare metabolic diseases are known[d-f] in which specific enzymes, such as cystathionine synthase, have a reduced affinity for PLP. Patients with these diseases also benefit from a higher than normal intake of the vitamin. Excessive excretion of the vitamin may also occur, an example being provided by a strain of laboratory mice that require twice the normal amount of vitamin B$_6$ and which die in convulsions after a brief period of vitamin B$_6$ depletion.[d] Dietary supplementation with large amounts of vitamin B$_6$ for treatment of medical conditions such as carpal tunnel syndrome has been controversial.[g] Amounts of pyridoxine over 50 mg per day may damage peripheral nerves, probably because of the chemical reactivity of pyridoxal and PLP.

Nicotinamide: About 7.5 mg/day. Tryptophan can substitute to some extent. See also Box 15-A.

Riboflavin: About 1.5 mg/day. See also Box 15-B.

Folic acid: About 0.2–0.4 mg/day. See also Box 15-D.

Vitamin B$_{12}$: About 2 μg/day. See also Box 16-B.

Vitamin C (ascorbic acid): 50–200 mg/day. See also Box 18-D.

---

[a] Parry, R. J., and Kunitni, M. G. (1976) *J. Am. Chem. Soc.* **98**, 4024–4025

[b] Gibson, K. J., Lorimer, G. H., Rendina, A. R., Taylor, W. S., Cohen, G., Gatenby, A. A., Payne, W. G., Roe, D. C., Lockett, B. A., Nudelman, A., Marcovici, D., Nachum, A., Wexler, B. A., Marsilii, E. L., Turner, S., IM, Howe, L. D., Kalbach, C. E., and Chi, H. (1995) *Biochemistry* **34**, 10976–10984

[c] Huang, W., Jia, J., Gibson, K. J., Taylor, W. S., Rendina, A. R., Schneider, G., and Lindqvist, Y. (1995) *Biochemistry* **34**, 10985–10995

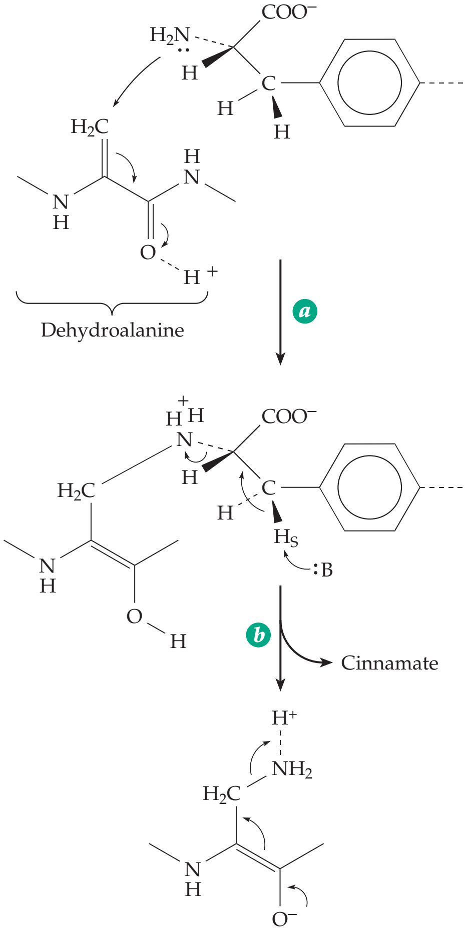
[d] Mudd, S. H., Levy, H. L., and Skovby, F. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1279–1327, McGraw-Hill, New York

[e] Bell, R. R., and Haskell, B. E. (1971) *Arch. Biochem. Biophys.* **147**, 588–601

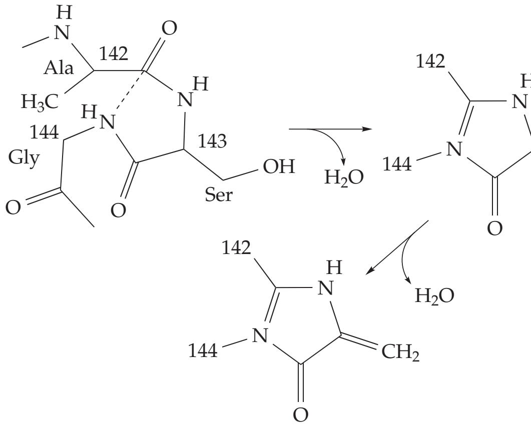
[f] Pascal, T. A., Gaull, G. E., Beratis, N. G., Gillam, B. M., Tallan, H. H., and Hirschhorn, K. (1975) *Science* **190**, 1209–1211

[g] Bender, D. A. (1999) *Br. J. Nutr.* **81**, 7–20

was inactive.[288,289] These results support the formation of dehydroalanine. One proposed mechanism of action of the enzymes involved addition of the substrate amino group to the C = C bond of the dehydroalanine followed by elimination (Eq. 14-46). Dissociation of the C – H bond in this step would be assisted by the electron-accepting properties of the adjacent aromatic ring. However, the proposed chemistry was not convincing. Noting that a 5-nitro substituent on the histidine greatly enhances the rate of reaction, Langer *et al.* proposed that the histidine reacts with the electrophilic carbon in the pyruvoyl center as in Eq. 14-47.[290,291] However, this chemistry, too, was unprecedented in enzymology. Determination of the three-dimensional structure of histidine ammonia-lyase[292] led to the discovery of a new prosthetic group and a solution to the problem. Two dehydration steps, (Eq. 14-47) convert an Ala-Gly-Ser sequence within the protein into **4-methylidine-imidazole-5-one (MIO)**, a modified dehydroalanine with enhanced electron-accepting properties. The proposed mechanism of action is portrayed in Eq. 14-48.

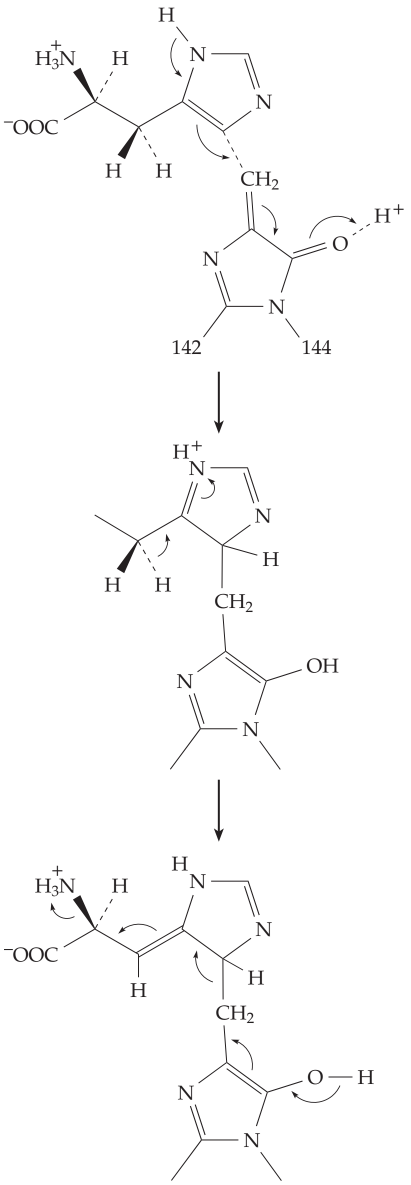


(14-46)



4-Methylidene-imidazole-5-one (MIO)

(14-47)



(14-48)

References

1. Moodie, S. L., Mitchell, J. B. O., and Thornton, J. M. (1996) *J. Mol. Biol.* **263**, 486–500
2. Lipmann, F. (1945) *J. Biol. Chem.* **160**, 173–190
3. Roskoski, R. N., Jr. (1987) *Trends Biochem. Sci.* **12**, 136–138

3a. Bentley, R. (2000) *Trends Biochem. Sci.* **25**, 302–305

4. Martin, D. P., and Drueckhammer, D. G. (1992) *J. Am. Chem. Soc.* **114**, 7287–7288
5. Baraniak, J., and Frey, P. A. (1988) *J. Am. Chem. Soc.* **110**, 4059–4060
6. Richter, D., and Hilz, H. (1979) *Trends Biochem. Sci.* **1**, N123–N124
7. Ohlrogge, J. B., Jaworski, J. G., and Post-Beittenmiller, D. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 3–32, CRC Press, Boca Raton, Florida
8. Pavela-Vrancic, M., Pfeifer, E., Schröder, W., von Döhren, H., and Kleinkauf, H. (1994) *J. Biol. Chem.* **269**, 14962–14966
9. Haese, A., Pieper, R., von Ostrowski, T., and Zocher, R. (1994) *J. Mol. Biol.* **243**, 116–122
10. Pieper, R., Ebert-Khosla, S., Cane, D., and Khosla, C. (1996) *Biochemistry* **35**, 2054 –2060
11. Cortes, J., Wiesmann, K. E. H., Roberts, G. A., Brown, M. J. B., Staunton, J., and Leadlay, P. F. (1995) *Science* **268**, 1487–1489
12. Brambl, R., and Plesofsky-Vig, N. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3644–3648
13. Berg, M., Hilbi, H., and Dimroth, P. (1996) *Biochemistry* **35**, 4689–4696
14. Dimroth, P. (1987) *Microbiol. Rev.* **51**, 320–340

14a. Hoenke, S., Wild, M. R., and Dimroth, P. (2000) *Biochemistry* **39**, 13223–13232

15. Jackowski, S., and Rock, C. O. (1984) *J. Biol. Chem.* **259**, 1891–1895

15a. McAllister, K. A., Peery, R. B., Meier, T. I., Fischl, A. S., and Zhao, G. (2000) *J. Biol. Chem.* **275**, 30864–30872

15b. Chirgadze, N. Y., Briggs, S. L., McAllister, K. A., Fischl, A. S., and Zhao, G. (2000) *EMBO J.* **19**, 5281–5287

16. Whitty, A., Fierke, C. A., and Jencks, W. P. (1995) *Biochemistry* **34**, 11678–11689
17. Fierke, C. A., and Jencks, W. P. (1986) *J. Biol. Chem.* **261**, 7603–7606
18. Langer, B. W., Jr., and Gyorgy, P. (1968) in *The Vitamins*, 2nd ed., Vol. 2 (Sebrell, W. H., Jr., and Harris, R. S., eds), pp. 294–322, Academic Press, New York
19. Bentley, R. (1985) *Trends Biochem. Sci.* **10**, 51–56
20. DeTitta, G. T., Edmonds, J. W., Stallings, W., and Donohue, J. (1976) *J. Am. Chem. Soc.* **98**, 1920–1926

20a. Garton, S. D., Temple, C. A., Dhawan, I. K., Barber, M. J., and Rajagopalan, K. V. (2000) *J. Biol. Chem.* **275**, 6798–6805

20b. McIver, L., Baxter, R. L., and Campopiano, D. J. (2000) *J. Biol. Chem.* **275**, 13888–13894

21. Wright, L. D., Cresson, E. L., Skeggs, H. R., Peck, R. L., Wolf, D. E., Wood, T. R., Valent, J., and Folkers, K. (1951) *Science* **114**, 635–636
22. Knowles, J. R. (1989) *Ann. Rev. Biochem.* **58**, 195–221
23. López-Casillas, F., Bai, D.-H., Luo, X., Kong, I. S., Hermodson, M. A., and Kim, K.-H. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5784–5788
24. Browner, M. F., Taroni, F., Sztul, E., and Rosenberg, L. E. (1989) *J. Biol. Chem.* **264**, 12680–12685
25. Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C., and Wood, H. G. (1988) *J. Biol. Chem.* **263**, 6461–6464
26. León-Del-Rio, A., Leclerc, D., Akerman, B., Wakamatsu, N., and Gravel, R. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4626–4630
27. Ahmad, P. M., and Ahmad, F. (1991) *FASEB J.* **5**, 2482–2485
28. Wang, X., Wurtele, E. S., Keller, G., McKean, A. L., and Nikolau, B. J. (1994) *J. Biol. Chem.* **269**, 11760–11769

28a. McKean, A. L., Ke, J., Song, J., Che, P., Achenbach, S., Nikolau, B. J., and Wurtele, E. S. (2000) *J. Biol. Chem.* **275**, 5582–5590

29. Weaver, L. M., Lebrun, L., Wurtele, E. S., and Nikolau, B. J. (1995) *Plant Physiol.* **107**, 1013–1014
30. Al-Feel, W., Chirala, S. S., and Wakil, S. J. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4534–4538
31. Lamzin, V. S., Dauter, Z., Popov, V. O., Harutyunyan, E. H., and Wilson, K. S. (1994) *J. Mol. Biol.* **236**, 759–785
32. Roessler, P. G., and Ohlrogge, J. B. (1993) *J. Biol. Chem.* **268**, 19254–19259

32a. Abu-Elheiga, L., Almarza-Ortega, D. B., Baldini, A., and Wakil, S. J. (1997) *J. Biol. Chem.* **272**, 10669–10677

32b. Abu-Elheiga, L., Brinkley, W. R., Zhong, L., Chirala, S. S., Woldegiorgis, G., and Wakil, S. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1444–1449

32c. Nikolskaya, T., Zagnitko, O., Tevzadze, G., Haselkorn, R., and Gornicki, P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14647–14651

32d. Schulte, W., Töpfer, R., Stracke, R., Schell, J., and Martini, N. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3465–3470

32e. Jitrapakdee, S., Booker, G. W., Cassady, A. I., and Wallace, J. C. (1997) *J. Biol. Chem.* **272**, 20522–20530

33. Lim, F., Morris, C. P., Occhiodoro, F., and Wallace, J. C. (1988) *J. Biol. Chem.* **263**, 11493–11497
34. Kalousek, F., Darigo, M. D., and Rosenberg, L. E. (1980) *J. Biol. Chem.* **255**, 60–65
35. Leon – Del-Rio, A., and Gravel, R. A. (1994) *J. Biol. Chem.* **269**, 22964–22968
36. Sasaki, Y., Hakamada, K., Suama, Y., Nagano, Y., Furusawa, I., and Matsuno, R. (1993) *J. Biol. Chem.* **268**, 25118–25123
37. Gornicki, P., Podkowinski, J., Scappino, L. A., DiMaio, J., Ward, E., and Haselkorn, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6860–6864
38. Nenortas, E., and Beckett, D. (1996) *J. Biol. Chem.* **271**, 7559–7567
39. Waldrop, G. L., Rayment, I., and Holden, H. M. (1994) *Biochemistry* **33**, 10249–10256

39a. Blanchard, C. Z., Lee, Y. M., Frantom, P. A., and Waldrop, G. L. (1999) *Biochemistry* **38**, 3393–3400

40. Li, S.-J., and Cronan, J. E., Jr. (1992) *J. Biol. Chem.* **267**, 855–863
41. Li, S.-J., and Cronan, J. E. J. (1992) *J. Biol. Chem.* **267**, 16841–16847
42. Ha, J., Daniel, S., Broyles, S. S., and Kim, K.-H. (1994) *J. Biol. Chem.* **269**, 22162–22168
43. Wood, H. G., Harmon, F. R., Wuhr, B., Hubner, K., and Lynen, F. (1980) *J. Biol. Chem.* **255**, 7397–7409
44. Shenoy, B. C., and Wood, H. G. (1988) *FASEB J.* **2**, 2396–2401
45. Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J., and Matthews, B. W. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9257–9261
46. Wolf, B. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3151–3177, McGraw-Hill, New York
47. Taroni, F., and Rosenberg, L. E. (1991) *J. Biol. Chem.* **266**, 13267– 13271

47a. Reche, P., and Perham, R. N. (1999) *EMBO J.* **18**, 2673–2682

48. Tanaka, K. (1981) *N. Engl. J. Med.* **304**, 839–840
49. Craft, D. V., Goss, N. H., Chandramouli, N., and Wood, H. G. (1985) *Biochemistry* **24**, 2471–2476
50. Chauhan, J., and Dakshinamurti, K. (1986) *J. Biol. Chem.* **261**, 4268–4275
51. Alberts, A. W., and Vagelos, P. R. (1972) in *The Enzymes*, 3rd ed., Vol. 6 (Boyer, P. D., ed), pp. 37–82, Academic Press, New York
52. Kumar, G. K., Beegen, H., and Wood, H. G. (1988) *Biochemistry* **27**, 5972–5978
53. Lynen, F., Knappe, J., Lorch, E., Jutting, G., and Ringelmann, E. (1959) *Angew. Chem. Int. Ed. Engl.* **71**, 481–486
54. Wood, H. G. (1979) *Trends Biochem. Sci.* **4**, N300–N302
55. Polakis, S. E., Guchhait, R. B., and Lane, M. D. (1972) *J. Biol. Chem.* **247**, 1235–1337

55a. Chapman-Smith, A., Forbes, B. E., Wallace, J. C., and Cronan, J. E., Jr. (1997) *J. Biol. Chem.* **272**, 26017–26022

55b. Reddy, D. V., Shenoy, B. C., Carey, P. R., and Sönnichsen, F. D. (2000) *Biochemistry* **39**, 2509–2516

56. Tipton, P. A., and Cleland, W. W. (1988) *J. Am. Chem. Soc.* **110**, 5866–5869
57. Tipton, P. A., and Cleland, W. W. (1988) *Biochemistry* **27**, 4317– 4325
58. Ogita, T., and Knowles, J. R. (1988) *Biochemistry* **27**, 8028–8033

58a. Branson, J. P., and Attwood, P. V. (2000) *Biochemistry* **39**, 7480–7491

59. Perrin, C. L., and Dwyer, T. J. (1987) *J. Am. Chem. Soc.* **109**, 5163–5167
60. Kluger, R., Davis, P. P., and Adawadkar, P. D. (1979) *J. Am. Chem. Soc.* **101**, 5995–6000
61. King, S. W., Natarajan, R., Bembi, R., and Fife, T. H. (1992) *J. Am. Chem. Soc.* **114**, 10715–10721
62. Kluger, R., and Taylor, S. D. (1991) *J. Am. Chem. Soc.* **113**, 996– 1001
63. Taylor, S. D., and Kluger, R. (1993) *J. Am. Chem. Soc.* **115**, 867 – 871
64. Hoving, H., Crysell, B., and Leadbay, P. F. (1985) *Biochemistry* **24**, 6163–6169
65. O'Keefe, S. J., and Knowles, J. R. (1986) *Biochemistry* **25**, 6077– 6084
66. Kuo, D. J., and Rose, I. A. (1993) *J. Am. Chem. Soc.* **115**, 387–390
67. Stubbe, J., Fish, S., and Abeles, R. H. (1980) *J. Biol. Chem.* **255**, 236–242
68. Reed, G. H., and Scrutton, M. G. (1974) *J. Biol. Chem.* **249**, 6156–6162

68a. Werneburg, B. G., and Ash, D. E. (1997) *Biochemistry* **36**, 14392–14402

69. DeTitta, G. T., Blessing, R. H., Moss, G. R., King, H. F., Sukumaran, D. K., and Roskwitalski, R. L. (1994) *J. Am. Chem. Soc.* **116**, 6485–6493
70. Legge, G. B., Branson, J. P., and Attwood, P. V. (1996) *Biochemistry* **35**, 3849–3856
71. Mabrouk, G. M., Helmy, I. M., Thampy, K. G., and Wakil, S. J. (1990) *J. Biol. Chem.* **265**, 6330–6338
72. Kim, K.-H., López-Casillas, F., Bai, D. H., Luo, X., and Pape, M. E. (1989) *FASEB J.* **3**, 2250–2256
73. Harwood, J. L. (1988) *Trends Biochem. Sci.* **13**, 330 – 331
74. LauBermair, E., Schwarz, E., Oesterhelt, D., Reinke, H., Beyreuther, K., and Dimroth, P. (1989) *J. Biol. Chem.* **264**, 14710–14715

74a. Jockel, P., Schmid, M., Choinowski, T., and Dimroth, P. (2000) *Biochemistry* **39**, 4320–4326

75. Laussermair, E., Schwarz, E., Oesterhelt, D., Reinke, H., Beyreuther, K., and Dimroth, P. (1989) *J. Biol. Chem.* **264**, 14710–14715
76. Huder, J. B., and Dimroth, P. (1993) *J. Biol. Chem.* **268**, 24564–24571

References

77. Bendrat, K., and Buckel, W. (1993) *Eur. J. Biochem.* **211**, 697 – 702

78. Ahmad, F., Lygre, D. G., Jacobson, B. E., and Wood, H. G. (1972) *J. Biol. Chem.* **247**, 6299 – 6305

79. Metzler, D. E. (1960) in *The Enzymes*, 2nd ed., Vol. 2 (Boyer, P. D., Lardy, H., and Myrback, K., eds), pp. 295 – 337, Academic Press, New York

80. Schellenberger, A., and Schowen, R. L., eds. (1988) *Thiamin Pyrophosphate Biochemistry*, CRC Press, Boca Raton, Florida (2 Vols.)

81. Krampitz, L. O. (1970) *Thiamin Diphosphate and Its Catalytic Functions*, Dekker, New York

82. Gubler, C. J., ed. (1976) *Thiamine*, Wiley, New York

83. Gallo, A., Mieyal, J. J., and Sable, H. Z. (1978) *Bioorg. Chem.* **4**, 147 – 177

83a. Jordan, F., Li, H., and Brown, A. (1999) *Biochemistry* **38**, 6369 – 6373

84. Kluger, R. (1987) *Chem. Rev.* **87**, 863 – 876

85. Sable, H. Z., ed., and Gubler, C. J., ed. (1982) *Ann. N.Y. Acad. Sci.* **378**

86. Karlson, P. (1984) *Trends Biochem. Sci.* **9**, 536 – 537

87. Panijpan, B. (1979) *J. Chem. Educ.* **56**, 805 – 806

88. Suchy, J., Mieyal, J. J., Bantle, G., and Sable, H. Z. (1972) *J. Biol. Chem.* **247**, 5905 – 5912

89. Williams, R. R., and Spies, T. D. (1938) *Vitamin B_1 and Its Use in Medicine*, Macmillan, New York

90. Kluger, R., Lam, J. F., Pezacki, J. P., and Yang, C.-M. (1995) *J. Am. Chem. Soc.* **117**, 11383 – 11389

91. Windheuser, J. J., and Higuchi, T. (1962) *J. Pharm. Sci.* **51**, 354 – 364

92. Costello, C. A., Kelleher, N. L., Abe, M., McLafferty, F. W., and Begley, T. P. (1996) *J. Biol. Chem.* **271**, 3445 – 3452

92a. Campobasso, N., Costello, C. A., Kinsland, C., Begley, T. P., and Ealick, S. E. (1998) *Biochemistry* **37**, 15981 – 15989

92b. Webb, E., and Downs, D. (1997) *J. Biol. Chem.* **272**, 15702 – 15707

92c. Nosaka, K., Onozuka, M., Nishino, H., Nishimura, H., Kawasaki, Y., and Ueyama, H. (1999) *J. Biol. Chem.* **274**, 34129 – 34133

93. Mizuhara, S., and Handler, P. (1954) *J. Am. Chem. Soc.* **76**, 571 – 573

94. Breslow, R. (1958) *J. Am. Chem. Soc.* **80**, 3719 – 3726

95. Hopmann, R. F. W., Brugnoni, G. P., and Fol, B. (1982) *J. Am. Chem. Soc.* **104**, 1341 – 1344

96. Washabaugh, M. W., and Jencks, W. P. (1989) *J. Am. Chem. Soc.* **111**, 683 – 692

97. Breslow, R., and McNelis, E. (1959) *J. Am. Chem. Soc.* **81**, 3080 – 3082

98. Kluger, R., Chin, J., and Smyth, T. (1981) *J. Am. Chem. Soc.* **103**, 884 – 888

99. Kluger, R., and Smyth, T. (1981) *J. Am. Chem. Soc.* **103**, 1216 – 1218

100. Kluger, R., Karimian, K., Gish, G., Pangborn, W. A., and DeTitta, G. T. (1987) *J. Am. Chem. Soc.* **109**, 618 – 620

101. Crosby, J., Stone, R., and Lienhard, G. E. (1970) *J. Am. Chem. Soc.* **92**, 2891 – 2900

102. Shin, W., Oh, D.-G., Chae, C.-H., and Yoon, T.-S. (1993) *J. Am. Chem. Soc.* **115**, 12238 – 12250

103. Lobell, M., and Crout, D. H. G. (1996) *J. Am. Chem. Soc.* **118**, 1867 – 1873

104. Gallo, A. A., and Sable, H. Z. (1974) *J. Biol. Chem.* **249**, 1382 – 1389

105. Schellenberger, A. (1967) *Angew. Chem. Int. Ed. Engl.* **6**, 1024 – 1035

106. Golbik, R., Neef, H., Hübner, G., König, S., Seliger, B., Meshalkina, L., Kochetov, G. A., and Schellenberger, A. (1991) *Bioorg. Chem.* **19**, 10 – 17

107. König, S., Schellenberger, A., Neef, H., and Schneider, G. (1994) *J. Biol. Chem.* **269**, 10879 – 10882

108. Jordan, F., and Mariam, Y. H. (1978) *J. Am. Chem. Soc.* **100**, 2534 – 2541

109. Metzler, D. E. (1977) *Biochemistry; The Chemical Reactions of Living Cells*, Academic Press, New York (440 – 441)

109a. Kern, D., Kern, G., Neef, H., Tittmann, K., Killenberg-Jabs, M., Wikner, C., Schneider, G., and Hübner, G. (1997) *Science* **275**, 67 – 70

110. Singleton, C. K., Wang, J. J.-L., Shan, L., and Martin, P. R. (1996) *Biochemistry* **35**, 15865 – 15869

111. Guo, F., Zhang, D., Kahyaoglu, A., Farid, R. S., and Jordan, F. (1998) *Biochemistry* **37**, 13379 – 13391

111a. Li, H., Furey, W., and Jordan, F. (1999) *Biochemistry* **38**, 9992 – 10003

112. Gish, G., Smyth, T., and Kluger, R. (1988) *J. Am. Chem. Soc.* **110**, 6230 – 6234

113. Annan, N., Paris, R., and Jordan, F. (1989) *J. Am. Chem. Soc.* **111**, 8895 – 8901

114. Alvarez, F. J., Ermer, J., Hübner, G., Schellenberger, A., and Schowen, R. L. (1991) *J. Am. Chem. Soc.* **113**, 8402 – 8409

115. Alvarez, F. J., Ermer, J., Hübner, G., Schellenberger, A., and Schowen, R. L. (1995) *J. Am. Chem. Soc.* **117**, 1678 – 1683

116. Sun, S., Duggleby, R. G., and Schowen, R. L. (1995) *J. Am. Chem. Soc.* **117**, 7317 – 7322

117. Harris, T. K., and Washabaugh, M. W. (1995) *Biochemistry* **34**, 13994 – 14000

118. Harris, T. K., and Washabaugh, M. W. (1995) *Biochemistry* **34**, 14001 – 14011

119. Muller, Y. A., and Schulz, G. E. (1993) *Science* **259**, 965 – 967

120. Muller, Y. A., Schumacher, G., Rudolph, R., and Schulz, G. E. (1994) *J. Mol. Biol.* **237**, 315 – 335

121. Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B., and Jordan, F. (1993) *Biochemistry* **32**, 6165 – 6170

122. Arjunan, P., Umland, T., Dyda, R., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D., and Jordan, F. (1996) *J. Mol. Biol.* **256**, 590 – 600

122a. Dobritzsch, D., König, S., Schneider, G., and Lu, G. (1998) *J. Biol. Chem.* **273**, 20196 – 20204

122b. Chang, A. K., Nixon, P. F., and Duggleby, R. G. (2000) *Biochemistry* **39**, 9430 – 9437

122c. Li, H., and Jordan, F. (1999) *Biochemistry* **38**, 10004 – 10012

123. Wikner, C., Meshalkina, L., Nilsson, U., Nikkola, M., Lindqvist, Y., Sundström, M., and Schneider, G. (1994) *J. Biol. Chem.* **269**, 32144 – 32150

124. Nikkola, M., Lindqvist, Y., and Schneider, G. (1994) *J. Mol. Biol.* **238**, 387 – 404

124a. Hasson, M. S., Muscate, A., McLeish, M. J., Polovnikova, L. S., Gerlt, J. A., Kenyon, G. L., Petsko, G. A., and Ringe, D. (1998) *Biochemistry* **37**, 9918 – 9930

125. Killenberg-Jabs, M., König, S., Eberhardt, I., Hohmann, S., and Hübner, G. (1997) *Biochemistry* **36**, 1900 – 1905

126. Koga, J., Adachi, T., and Hidaka, H. (1992) *J. Biol. Chem.* **267**, 15823 – 15828

127. Hasson, M. S., Muscate, A., Henehan, G. T. M., Guidinger, P. F., Petsko, G. A., Ringe, D., and Kenyon, G. L. (1995) *Protein Sci.* **4**, 955 – 959

128. Ibdah, M., Bar-Ilan, A., Livnah, O., Schloss, J. V., Barak, Z., and Chipman, D. M. (1996) *Biochemistry* **35**, 16282 – 16291

129. Bernasconi, P., Woodworth, A. R., Rosen, B. A., Subramanian, M. V., and Siehl, D. L. (1995) *J. Biol. Chem.* **270**, 17381 – 17385

130. Ott, K.-H., Kwagh, J.-G., Stockton, G. W., Sidorov, V., and Kakefuda, G. (1996) *J. Mol. Biol.* **263**, 359 – 368

131. Stivers, J. T., and Washabaugh, M. W. (1993) *Biochemistry* **32**, 13472 – 13482

132. Chang, Y.-Y., Wang, A.-Y., and Cronan, J. E., Jr. (1993) *J. Biol. Chem.* **268**, 3911 – 3919

132a. Kuzuyama, T., Takahashi, S., Takagi, M., and Seto, H. (2000) *J. Biol. Chem.* **275**, 19928 – 19932

132b. Wikner, C., Nilsson, U., Meshalkina, L., Udekwu, C., Lindqvist, Y., and Schneider, G. (1997) *Biochemistry* **36**, 15643 – 15649

133. Chuang, L. F., and Collins, E. B. (1968) *J. Bacteriol.* **95**, 2083 – 2089

134. Gruys, K. J., Datta, A., and Frey, P. A. (1989) *Biochemistry* **28**, 9071 – 9080

135. Flournoy, D. S., and Frey, P. A. (1989) *Biochemistry* **28**, 9594 – 9602

135a. Tittmann, K., Proske, D., Spinka, M., Ghisla, S., Rudolph, R., Hübner, G., and Kern, G. (1998) *J. Biol. Chem.* **273**, 12929 – 12934

136. Matsuda, T., and Cooper, J. R. (1983) *Biochemistry* **22**, 2209 – 2213

137. Barchi, R. L., and Viale, R. O. (1976) *J. Biol. Chem.* **251**, 193 – 197

138. Nishino, K., Itokawa, Y., Nishino, N., Piros, K., and Cooper, J. R. (1983) *J. Biol. Chem.* **258**, 1871 – 11878

139. Braunstein, A. E., and Kritzmann, M. G. (1937) *Enzymologia* **2**, 751 – 752

140. Braunstein, A. E., and Kritzmann, M. G. (1937) *Biokhim.* **2**, 242 – 262,859 – 874

141. Braunstein, A. E. (1973) in *The Enzymes*, 3rd ed., Vol. 9 (Boyer, P. D., ed), pp. 379 – 481, Academic Press, New York

142. Christen, P., and Metzler, D. E., eds. (1984) *Transaminases*, Wiley, New York

143. Torchinsky, Y. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1A (Dolphin, R., Poulson, R., and Avramovic, O., eds), pp. 169 – 221, Wiley, New York

144. Snell, E. E. (1944) *J. Biol. Chem.* **154**, 313 – 314

145. Snell, E. E. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1A (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 1 – 12, Wiley, New York

146. Gunsalus, I. C., Bellamy, W. D., and Umbreit, W. W. (1944) *J. Biol. Chem.* **155**, 685 – 686

147. Snell, E. E. (1981) in *Methods in Vitamin B_6 Nutrition* (Leklem, J. E., and Reynolds, R. D., eds), pp. 1 – 19, Plenum, New York

148. Metzler, D. E., Ikawa, M., and Snell, E. E. (1954) *J. Am. Chem. Soc.* **76**, 648 – 652

149. Kallen, R. G., Korpela, T., Martell, A. E., Matsushima, Y., Metzler, C. M., Metzler, D. E., Morozov, Y. V., Ralston, I. M., Savin, F. A., Torchinsky, Y. M., and Ueno, H. (1984) in *Transaminases* (Christen, P., and Metzler, D. E., eds), pp. 37 – 106, Wiley, New York

150. Leussing, D. L. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1A (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 69 – 115, Wiley, New York

151. Braunstein, A. E., and Kritzmann, M. G. (1952) *Dokl. Akad. Nauk. SSSR* **85**, 1115 – 1118

152. Braunstein, A. E., and Kritzmann, M. G. (1952) *Biokhim.* **18**, 393 – 411

153. Shaw, J. P., Petsko, G. A., and Ringe, D. (1997) *Biochemistry* **36**, 1329 – 1342

153a. Stamper, C. G. F., Morollo, A. A., and Ringe, D. (1998) *Biochemistry* **37**, 10438 – 10445

153b. Morollo, A. A., Petsko, G. A., and Ringe, D. (1999) *Biochemistry* **38**, 3293 – 3301

153c. Sun, S., and Toney, M. D. (1999) *Biochemistry* **38**, 4058 – 4065

154. Aswad, D. W., and Johnson, B. A. (1987) *Trends Biochem. Sci.* **12**, 155 – 158

154a. Watababe, A., Kurokawa, Y., Yoshimura, T., Kurihara, T., Soda, K., and Esaki, N. (1999) *J. Biol. Chem.* **274**, 4189 – 4194

References

155. Hoffmann, K., Schneider-Scherzer, E., Kleinkauf, H., and Zocher, R. (1994) *J. Biol. Chem.* **269**, 12710 – 12714
155a. Cheng, Y.-Q., and Walton, J. D. (2000) *J. Biol. Chem.* **275**, 4906 – 4911
156. Schnell, M. J., Cooper, O. B., and Snyder, S. H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2013 – 2018
157. Yamauchi, T., Choi, S.-Y., Okada, H., Yohda, M., Kumagai, H., Esaki, N., and Soda, K. (1992) *J. Biol. Chem.* **267**, 18361 – 18364
158. Yohda, M., Endo, I., Abe, Y., Ohta, T., Iida, T., Maruyama, T., and Kagawa, Y. (1996) *J. Biol. Chem.* **271**, 22017 – 22021
159. Li, N., and Mattoo, A. K. (1994) *J. Biol. Chem.* **269**, 6908 – 6917
160. Li, N., Jiang, X. N., Cai, G. P., and Yang, S. F. (1996) *J. Biol. Chem.* **271**, 25738 – 25741
161. Hohenester, E., White, M. F., Kirsch, J. F., and Jansonius, J. N. (1994) *J. Mol. Biol.* **243**, 947 – 949
161a. Capitani, G., Hohenester, E., Feng, L., Storici, P., Kirsch, J. F., and Jansonius, J. N. (1999) *J. Mol. Biol.* **294**, 745 – 756
162. Ramalingam, K., Lee, K.-M., Woodard, R. W., Bleecker, A. B., and Kende, H. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7820 – 7824
163. Rhee, S., Silva, M. M., Hyde, C. C., Rogers, P. H., Metzler, C. M., Metzler, D. E., and Arnone, A. (1997) *J. Biol. Chem.* **272**, 17293 – 17302
164. Malashkevich, V. N., Strokopytov, B. V., Borisov, V. V., Dauter, Z., Wilson, K. S., and Torchinsky, Y. M. (1995) *J. Mol. Biol.* **247**, 111 – 124
165. McPhalen, C. A., Vincent, M. G., and Jansonius, J. N. (1992) *J. Mol. Biol.* **225**, 495 – 517
166. Kamitori, S., Okamoto, A., Hirotsu, K., Higuchi, T., Kuramitsu, S., Kagamiyama, H., Matsuura, Y., and Katsube, Y. (1990) *J. Biochem.* **108**, 175 – 184
167. Okamoto, A., Higuchi, T., Hirotsu, K., Kuramitsu, S., and Kagamiyama, H. (1994) *J. Biochem.* **116**, 95 – 107
167a. Jeffery, C. J., Barry, T., Doonan, S., Petsko, G. A., and Ringe, D. (1998) *Protein Sci.* **7**, 1380 – 1387
168. Uchida, C., Funai, T., Oda, T., Ohbayashi, K., and Ichiyama, A. (1994) *J. Biol. Chem.* **269**, 8846 – 8856
168a. Matsui, I., Matsui, E., Sakai, Y., Kikuchi, H., Kawarabayasi, Y., Ura, H., Kawaguchi, S.-i, Kuramitsu, S., and Harata, K. (2000) *J. Biol. Chem.* **275**, 4871 – 4879
168b. Okada, K., Hirotsu, K., Sato, M., Hayashi, H., and Kagamiyama, H. (1997) *J. Biochem.* **121**, 637 – 641
168c. Davoodi, J., Drown, P. M., Bledsoe, R. K., Wallin, R., Reinhart, G. D., and Hutson, S. M. (1998) *J. Biol. Chem.* **273**, 4982 – 4989
168d. van Ophem, P. W., Peisach, D., Erickson, S. D., Soda, K., Ringe, D., and Manning, J. M. (1999) *Biochemistry* **38**, 1323 – 1331
168e. Hester, G., Stark, W., Moser, M., Kallen, J., Markovic'-Housley, Z., and Jansonius, J. N. (1999) *J. Mol. Biol.* **286**, 829 – 850
168f. Storici, P., Capitani, G., Müller, R., Schirmer, T., and Jansonius, J. N. (1999) *J. Mol. Biol.* **285**, 297 – 309
168g. Shah, S. A., Shen, B. W., and Brünger, A. T. (1997) *Structure* **5**, 1067 – 1075
168h. Ledwidge, R., and Blanchard, J. S. (1999) *Biochemistry* **38**, 3019 – 3024
168i. Smith, M. A., King, P. J., and Grimm, B. (1998) *Biochemistry* **37**, 319 – 329
168j. Käck, H., Sandmark, J., Gibson, K., Schneider, G., and Lindqvist, Y. (1999) *J. Mol. Biol.* **291**, 857 – 876
168k. Contestabile, R., Angelaccio, S., Maytum, R., Bossa, F., and John, R. A. (2000) *J. Biol. Chem.* **275**, 3879 – 3886
169. Kintanar, A., Metzler, C. M., Metzler, D. E., and Scott, R. D. (1991) *J. Biol. Chem.* **266**, 17222 – 17229
170. Miles, E. W. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1B (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 253 – 310, Wiley, New York
171. Marceau, M., Lewis, S. D., Kojiro, C. L., and Shafer, J. A. (1989) *J. Biol. Chem.* **264**, 2753 – 2757
172. Ogawa, H., Gomi, T., Konishi, K., Date, T., Nakashima, H., Nose, K., Matsuda, Y., Peraino, C., Pilot, H. C., and Fujioka, M. (1989) *J. Biol. Chem.* **264**, 15818 – 15823
173. Obmolova, G., Tepliakov, A., Harutyunyan, E., Wahler, G., and Schnackerz, K. D. (1990) *J. Mol. Biol.* **214**, 641 – 642
173a. Ogawa, H., Takusagawa, F., Wakaki, K., Kishi, H., Eskandarian, M. R., Kobayashi, M., Date, T., Huh, N.-H., and Pitot, H. C. (1999) *J. Biol. Chem.* **274**, 12855 – 12860
174. Snell, E. E. (1975) *Adv. Enzymol.* **42**, 287 – 333
175. Metzler, C. M., Viswanath, R., and Metzler, D. E. (1991) *J. Biol. Chem.* **266**, 9374 – 9381
176. Sloan, M. J., and Phillips, R. S. (1996) *Biochemistry* **35**, 16165 – 16173
176a. Isupov, M. N., Antson, A. A., Dodson, E. J., Dodson, G. G., Dementieva, I. S., Zakomirdina, L. N., Wilson, K. S., Dauter, Z., Lebedev, A. A., and Harutyunyan, E. H. (1998) *J. Mol. Biol.* **276**, 603 – 623
176b. Ikushiro, H., Hayashi, H., Kawata, Y., and Kagamiyama, H. (1998) *Biochemistry* **37**, 3043 – 3052
176c. Phillips, R. S., Sundararaju, B., and Faleev, N. G. (2000) *J. Am. Chem. Soc.* **122**, 1008 – 1014
177. Antson, A. A., Demidkina, T. V., Gollnick, P., Dauter, Z., Von Tersch, R. L., Long, J., Berezhnoy, S. N., Phillips, R. S., Harutyunyan, E. H., and Wilson, K. S. (1993) *Biochemistry* **32**, 4195 – 4206
177a. Sundararaju, B., Antson, A. A., Phillips, R. S., Demidkina, T. V., Barbolina, M. V., Gollnick, P., Dodson, G. G., and Wilson, K. S. (1997) *Biochemistry* **36**, 6502 – 6510
178. Chen, H. Y., Demidkina, T. V., and Phillips, R. S. (1995) *Biochemistry* **34**, 12276 – 12283
178a. Sundararaju, B., Chen, H., Shilcutt, S., and Phillips, R. S. (2000) *Biochemistry* **39**, 8546 – 8555
179. Block, E. (1985) *Sci. Am.* **252**(Mar), 114 – 119
180. Block, E., Gillies, J. Z., Gillies, C. W., Bazzi, A. A., Putman, D., Revelle, L. K., Wang, D., and Zhang, X. (1996) *J. Am. Chem. Soc.* **118**, 7492 – 7501
181. Gentry-Weeks, C. R., Spokes, J., and Thompson, J. (1995) *J. Biol. Chem.* **270**, 7695 – 7702
182. Clausen, T., Huber, R., Laber, B., Pohlenz, H.-D., and Messerschmidt, A. (1996) *J. Mol. Biol.* **262**, 202 – 224
183. Nagasawa, T., Tanizawa, K., Satoda, T., and Yamada, H. (1988) *J. Biol. Chem.* **263**, 958 – 964
184. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., and Davies, D. R. (1988) *J. Biol. Chem.* **263**, 17857 – 17871
184a. Rhee, S., Miles, E. W., and Davies, D. R. (1998) *J. Biol. Chem.* **273**, 8553 – 8555
184b. Weyand, M., and Schlichting, I. (1999) *Biochemistry* **38**, 16469 – 16480
184c. Sachpatzidis, A., Dealwis, C., Lubetsky, J. B., Liang, P.-H., Anderson, K. S., and Lolis, E. (1999) *Biochemistry* **38**, 12665 – 12674
184d. Rhee, S., Miles, E. W., Mozzarelli, A., and Davies, D. R. (1998) *Biochemistry* **37**, 10653 – 10659
185. Rege, V. D., Kredich, N. M., Tai, C.-H., Karsten, W. E., Schnackerz, K. D., and Cook, P. F. (1996) *Biochemistry* **35**, 13485 – 13493
185a. Burkhard, P., Tai, C.-H., Jansonius, J. N., and Cook, P. F. (2000) *J. Mol. Biol.* **303**, 279 – 286
186. Cook, P. F., Tai, C.-H., Hwang, C.-C., Woehl, E. U., Dunn, M. F., and Schnackerz, K. D. (1996) *J. Biol. Chem.* **271**, 25842 – 25849
186a. Benci, S., Vaccari, S., Mozzarelli, A., and Cook, P. F. (1997) *Biochemistry* **36**, 15419 – 15427
187. Kery, V., Bukovska, G., and Kraus, J. P. (1994) *J. Biol. Chem.* **269**, 25283 – 25288
188. Peracchi, A., Bettati, S., Mozzarelli, A., Rossi, G. L., Miles, E. W., and Dunn, M. F. (1996) *Biochemistry* **35**, 1872 – 1880
188a. Jhee, K.-H., McPhie, P., and Miles, E. W. (2000) *J. Biol. Chem.* **275**, 11541 – 11544
188b. Taoka, S., Widjaja, L., and Banerjee, R. (1999) *Biochemistry* **38**, 13155 – 13161
188c. Ojha, S., Hwang, J., Kabil, O., Penner-Hahn, J. E., and Banerjee, R. (2000) *Biochemistry* **39**, 10542 – 10547
189. Palcic, M. M., and Floss, H. G. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1A (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 25 – 68, Wiley, New York
190. Pan, P., Woehl, E., and Dunn, M. F. (1997) *Trends Biochem. Sci.* **22**, 22 – 27
191. Sukhareva, B. S. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1B (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 325 – 353, Wiley, New York
192. Sandmeier, E., Hale, T. I., and Christen, P. (1994) *Eur. J. Biochem.* **221**, 997 – 1002
193. Nathan, B., Hsu, C.-C., Bao, J., Wu, R., and Wu, J.-Y. (1994) *J. Biol. Chem.* **269**, 7249 – 7254
194. Porter, T. G., and Martin, D. L. (1988) *Biochim. Biophys. Acta.* **874**, 235 – 244
195. Bu, D.-F., Erlander, M. G., Hitz, B. C., Tillakaratne, N. J. K., Kaufman, D. L., Wagner-McPherson, C. B., Evans, G. A., and Tobin, A. J. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2115 – 2119
196. Dirkx, R., Jr., Thomas, A., Li, L., Lernmark, Å., Sherwin, R. S., DeCamilli, P., and Solimena, M. (1995) *J. Biol. Chem.* **270**, 2241 – 2246
196a. Qu, K., Martin, D. L., and Lawrence, C. E. (1998) *Protein Sci.* **7**, 1092 – 1105
196b. Kanaani, J., Lissin, D., Kash, S. F., and Baekkeskov, S. (1999) *J. Biol. Chem.* **274**, 37200 – 37209
197. Malashkevich, V. N., Filipponi, P., Sauder, U., Dominici, P., Jansonius, J. N., and Voltattorni, C. B. (1992) *J. Mol. Biol.* **224**, 1167 – 1170
198. Dominici, P., Filipponi, P., Schinninà, M. E., Barra, D., and Voltattorni, C. B. (1990) *Ann. N.Y. Acad. Sci.* **585**, 162 – 171
199. Reith, J., Benkelfat, C., Sherwin, A., Yasuhara, Y., Kuwabara, H., Andermann, F., Bachneff, S., Cumming, P., Diksic, M., Dyve, S. E., Etinne, P., Evans, A. C., Lal, S., Shevell, M., Savard, G., Wong, D. F., Chouinard, G., and Gjedde, A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11651 – 11654
199a. Hayashi, H., Tsukiyama, F., Ishii, S., Mizuguchi, H., and Kagamiyama, H. (1999) *Biochemistry* **38**, 15615 – 15622
199b. Nishino, J., Hayashi, H., Ishii, S., and Kagamiyama, H. (1997) *J. Biochem.* **121**, 604 – 611
200. Bhattacharjee, M. K., and Snell, E. E. (1990) *J. Biol. Chem.* **265**, 6664 – 6668
201. Chudomelka, P. J., Ramaley, R. F., and Murrin, L. C. (1990) *Neurochemical Research* **15**, 17 – 24
202. Yatsunami, K., Tsuchikawa, M., Kamada, M., Hori, K., and Higuchi, T. (1995) *J. Biol. Chem.* **270**, 30813 – 30817

203. Osterman, A. L., Kinch, L. N., Grishin, N. V., and Phillips, M. A. (1995) *J. Biol. Chem.* **270**, 11797–11802
204. Momany, C., Ernst, S., Ghosh, R., Chang, N.-L., and Hackert, M. L. (1995) *J. Mol. Biol.* **252**, 643–655
205. Osterman, A. L., Brooks, H. B., Jackson, L., Abbott, J. J., and Phillips, M. A. (1999) *Biochemistry* **38**, 11814–11826
206. Viguera, E., Trelles, O., Urdiales, J. L., Matés, J. M., and Sánchez-Jiménez, F. (1994) *Trends Biochem. Sci.* **19**, 318–319
206a. Almrud, J. J., Oliveira, M. A., Kern, A. D., Grishin, N. V., Phillips, M. A., and Hackert, M. L. (2000) *J. Mol. Biol.* **295**, 7–16
206b. Jackson, L. K., Brooks, H. B., Osterman, A. L., Goldsmith, E. J., and Phillips, M. A. (2000) *Biochemistry* **39**, 11247–11257
206c. Grishin, N. V., Osterman, A. L., Brooks, H. B., Phillips, M. A., and Goldsmith, E. J. (1999) *Biochemistry* **38**, 15174–15184
207. Kume, A., Koyata, H., Sakakibara, T., Ishiguro, Y., Kure, S., and Hiraga, K. (1991) *J. Biol. Chem.* **266**, 3323–3329
208. Pasternack, L. B., Laude, D. A., Jr., and Appling, D. R. (1992) *Biochemistry* **31**, 8713–8719
209. Kennard, O., and Salisbury, S. A. (1993) *J. Biol. Chem.* **268**, 10701–10704
209a. Okamura-Ikeda, K., Fujiwara, K., and Motokawa, Y. (1999) *J. Biol. Chem.* **274**, 17471–17477
209b. Guilhaudis, L., Simorre, J.-P., Blackledge, M., Marion, D., Gans, P., Neuburger, M., and Douce, R. (2000) *Biochemistry* **39**, 4259–4266
210. Akhtar, M., Stevenson, D. E., and Gani, D. (1990) *Biochemistry* **29**, 7648–7660
210a. Zhou, X., and Toney, M. D. (1999) *Biochemistry* **38**, 311–320
210b. Malashkevich, V. N., Strop, P., Keller, J. W., Jansonius, J. N., and Toney, M. D. (1999) *J. Mol. Biol.* **294**, 193–200
211. Scarsdale, J. N., Kazanina, G., Radaev, S., Schirch, V., and Wright, H. T. (1999) *Biochemistry* **38**, 8347–8358
212. Stover, P. J., Chen, L. H., Suh, J. R., Stover, D. M., Keyomarsi, K., and Shane, B. (1997) *J. Biol. Chem.* **272**, 1842–1848
212a. Szebenyi, D. M. E., Liu, X., Kriksunov, I. A., Stover, P. J., and Thiel, D. J. (2000) *Biochemistry* **39**, 13313–13323
213. Webb, H. K., and Matthews, R. G. (1995) *J. Biol. Chem.* **270**, 17204–17209
213a. Jagath, J. R., Sharma, B., Appaji Rao, N., and Savithri, H. S. (1997) *J. Biol. Chem.* **272**, 24355–24362
213b. Scarsdale, J. N., Radaev, S., Kazanina, G., Schirch, V., and Wright, H. T. (2000) *J. Mol. Biol.* **296**, 155–168
214. Tong, H., and Davis, L. (1994) *J. Biol. Chem.* **269**, 4057–4064
215. Tong, H., and Davis, L. (1995) *Biochemistry* **34**, 3362–3367
216. Cotter, P. D., Baumann, M., and Bishop, D. F. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4028–4032
217. Ferreira, G. C., Neame, P. J., and Dailey, H. A. (1993) *Protein Sci.* **2**, 1959–1965
218. Cox, T. C., Bottomley, S. S., Wiley, J. S., Bawden, M. J., Matthews, C. S., and May, B. K. (1994) *N. Engl. J. Med.* **330**, 675–679
218a. Hunter, G. A., and Ferreira, G. C. (1999) *J. Biol. Chem.* **274**, 12222–12228
218b. Hunter, G. A., and Ferreira, G. C. (1999) *Biochemistry* **38**, 3711–3718
219. Williams, R. D., Nixon, D. W., and Merrill, A. H., Jr. (1984) *Cancer Research* **44**, 1918–1923
220. Webster, S. P., Alexeev, D., Campopiano, D. J., Watt, R. M., Alexeeva, M., Sawyer, L., and Baxter, R. L. (2000) *Biochemistry* **39**, 516–528
220a. Steegborn, C., Messerschmidt, A., Laber, B., Streber, W., Huber, R., and Clausen, T. (1999) *J. Mol. Biol.* **290**, 983–996
220b. Steegborn, C., Clausen, T., Sondermann, P., Jacob, U., Worbs, M., Marinkovic, S., Huber, R., and Wahl, M. C. (1999) *J. Biol. Chem.* **274**, 12675–12684
220c. Laber, B., Gerbling, K.-P., Harde, C., Neff, K.-H., Nordhoff, E., and Pohlenz, H.-D. (1994) *Biochemistry* **33**, 3413–3423
221. Tate, S. E., and Meister, A. (1971) *Adv. Enzymol.* **35**, 503–543
222. Mihara, H., Kurihara, T., Watanabe, T., Yoshimura, T., and Esaki, N. (2000) *J. Biol. Chem.* **275**, 6195–6200
223. Esaki, N., Seraneeprakarn, V., Tanaka, H., and Soda, K. (1988) *J. Bacteriol.* **170**, 751–756
223a. Mihara, H., Maeda, M., Fujii, T., Kurihara, T., Hata, Y., and Esaki, N. (1999) *J. Biol. Chem.* **274**, 14768–14772
223b. Fujii, T., Maeda, M., Mihara, H., Kurihara, T., Esaki, N., and Hata, Y. (2000) *Biochemistry* **39**, 1263–1273
224. Zheng, L., and Dean, D. R. (1994) *J. Biol. Chem.* **269**, 18723–18726
225. Phillips, R. S., and Dua, R. K. (1991) *J. Am. Chem. Soc.* **113**, 7385–7388
226. Dua, R. K., Taylor, E. W., and Phillips, R. S. (1993) *J. Am. Chem. Soc.* **115**, 1264–1270
227. Krebs, E. G., and Fischer, E. H. (1964) *Vitam. Horm.(N. Y.)* **22**, 399–410
228. Rubenstein, P. A., and Strominger, J. L. (1974) *J. Biol. Chem.* **249**, 3776–3781
229. Johnson, D. A., Gassner, G. T., Bandarian, V., Ruzicka, F. J., Ballou, D. P., Reed, G. H., and Liu, H.-w. (1996) *Biochemistry* **35**, 15846–15856
230. Gonzalez-Porqué, P. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1B (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 391–419, Wiley, New York
231. Bailey, G. B., Chotamangsa, O., and Vuttivej, K. (1970) *Biochemistry* **9**, 3243–3248
232. Toney, M. D., Hohenester, E., Cowan, S. W., and Jansonius, J. N. (1993) *Science* **261**, 756–759
233. Dunathan, H. C. (1971) *Adv. Enzymol.* **35**, 79–134
234. Tobler, H. P., Gehring, H., and Christen, P. (1987) *J. Biol. Chem.* **262**, 8985–8989
235. Sugio, S., Petsko, G. A., Manning, J. M., Soda, K., and Ringe, D. (1995) *Biochemistry* **34**, 9661–9669
236. Syke, G. E., Potts, R., and Floss, H. G. (1974) *J. Am. Chem. Soc.* **96**, 1593–1595
237. Floss, H. G., and Vederas, J. C. (1982) in *New Comprehensive Biochemistry Stereochemistry*, Vol. 3 (Tamm, C., ed), pp. 161–199, Elsevier, Amsterdam
238. Palcic, M. M., and Floss, H. G. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1A (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 25–68, Wiley, New York
239. Metzler, D. E., Harris, C. M., Johnson, R. J., Siano, D. B., and Thomson, J. A. (1973) *Biochemistry* **12**, 5377–5392
240. Metzler, C. M., Cahill, A. E., and Metzler, D. E. (1980) *J. Am. Chem. Soc.* **102**, 6075–6082
241. Schirch, L., and Jenkins, W. T. (1964) *J. Biol. Chem.* **239**, 3801–3807
242. Metzler, C. M., Harris, A. G., and Metzler, D. E. (1988) *Biochemistry* **27**, 4923–4933
243. Schnackerz, K. D. (1986) in *Vitamin B_6, Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*, Vol. 1A (Dolphin, D., Poulson, R., and Avramovíc, O., eds), pp. 245–264, Wiley, New York
243a. Mehta, P. K., and Christen, P. (1998) in *Advances in Enzymology and Related Areas of Molecular Biology*, Vol. 74 (Purich, D. L., ed), Wiley, New York (pp. 129–184)
243b. Jansonius, J. N. (1998) *Current Opinion in Structural Biology* **8**, 759–769
244. Watanabe, N., Yonaha, K., Sakabe, K., Sakabe, N., Aibara, S., and Morita, Y. (1991) in *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors* (Fukui, T., Kagamiyama, H., Soda, K., and Wada, H., eds), pp. 121–124, Pergamon Press, Oxford
245. Huston, S. M., Bledsoe, R. K., Hall, T. R., and Dawson, P. A. (1995) *J. Biol. Chem.* **270**, 30344–30352
246. Ivanov, V. I., and Karpeisky, M. Y. (1969) *Adv. Enzymol.* **32**, 21–53
247. Arnone, A., Christen, P., Jansonius, J. N., and Metzler, D. E. (1985) in *Transaminases* (Christen, P., and Metzler, D. E., eds), pp. 326–370, Wiley, New York
248. Rhee, S., Silva, M. M., Hyde, C. C., Rogers, P. H., Metzler, C. M., Metzler, D. E., and Arnone, A. (1997) *J. Biol. Chem.* **272**, 17293–17302
249. Metzler, C. M., Mitra, J., Metzler, D. E., Makinen, M. W., Hyde, C. C., Rogers, P., and Arnone, A. (1988) *J. Mol. Biol.* **203**, 197–220
250. Malashkevich, V. N., Toney, M. D., and Jansonius, J. N. (1993) *Biochemistry* **32**, 13451–13462
251. von Stosch, A. G. (1996) *Biochemistry* **35**, 15260–15268
252. Metzler, C. M., Metzler, D. E., Kintanar, A., Scott, R. D., and Marceau, M. (1991) *Biochem. Biophys. Res. Commun.* **178**, 385–392
253. Fukui, T., Kagamiyama, H., Soda, K., and Wada, H., eds. (1991) *Enzymes Dependent on Pyridoxal Phosphate and Other Carbonyl Compounds as Cofactors*, Pergamon Press, Oxford
254. Marino, G., Sannia, G., and Bossa, F., eds. (1994) *Biochemistry of Vitamin B_6 and PQQ*, Birkhäuser, Basel
254a. Iriarte, A., Kagan, H. M., and Martinez-Carrion, M., eds. (2000) *Biochemistry and Molecular Biology of Vitamin B6 and PQQ-dependent Proteins*, Birkhäuser Verlag, Basel
255. Relimpio, A., Iriarte, A., Chlebowski, J. F., and Martinez-Carrion, M. (1981) *J. Biol. Chem.* **256**, 4478–4488
256. Gloss, L. M., and Kirsch, J. F. (1995) *Biochemistry* **34**, 12323–12332
257. Hwang, C.-C., Woehl, E. U., Minter, D. E., Dunn, M. F., and Cook, P. F. (1996) *Biochemistry* **35**, 6358–6365
258. Pan, P., and Dunn, M. F. (1996) *Biochemistry* **35**, 5002–5013
259. Anderson, K. S., Kim, A. Y., Quillen, J. M., Sayers, E., Yang, X.-J., and Miles, E. W. (1995) *J. Biol. Chem.* **270**, 29936–29944
260. Schnackerz, K. D., Tai, C.-H., Simmons, J. W., III, Jacobson, T. M., Rao, G. S. J., and Cook, P. F. (1995) *Biochemistry* **34**, 12152–12160
261. Hayashi, H., and Kagamiyama, H. (1995) *Biochemistry* **34**, 9413–9423
262. Mattingly, M. E., Mattingly, J. R., Jr., and Martinez-Carrion, M. (1982) *J. Biol. Chem.* **257**, 8872–8878
263. Metzler, D. E. (1997) *Methods Enzymol.* **280**, 30–40
264. Higaki, T., Tanase, S., Nagashima, F., Morino, Y., Scott, A. I., Williams, H. J., and Stolowich, N. J. (1991) *Biochemistry* **30**, 2519–2526
265. Goldberg, J. M., and Kirsch, J. F. (1996) *Biochemistry* **35**, 5280–5291
266. Recsei, P. A., and Snell, E. E. (1984) *Ann. Rev. Biochem.* **53**, 357–387
267. van Peolje, P. D., and Snell, E. E. (1990) *Ann. Rev. Biochem.* **59**, 29–59

References

268. Pishko, E. J., and Robertus, J. D. (1993) *Biochemistry* **32**, 4943– 4948
269. Gallagher, T., Snell, E. E., and Hackert, M. L. (1989) *J. Biol. Chem.* **264**, 12737–12743
270. Recsei, P. A., Huynh, Q. K., and Snell, E. E. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 973–977
271. Van Poelje, P. D., and Snell, E. E. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8449–8453
272. Huynh, Q. H., and Snell, E. E. (1986) *J. Biol. Chem.* **261**, 1521–1524
273. Shirahata, A., and Pegg, A. E. (1986) *J. Biol. Chem.* **261**, 13833–13837
274. Stanley, B. A., Shantz, L. M., and Pegg, A. E. (1994) *J. Biol. Chem.* **269**, 7901–7907

274a. Xiong, H., and Pegg, A. E. (1999) *J. Biol. Chem.* **274**, 35059–35066

275. Clancey, C. J., Chang, S.-C., and Dowhan, W. (1993) *J. Biol. Chem.* **268**, 24580–24590
276. Satre, M., and Kennedy, E. P. (1978) *J. Biol. Chem.* **253**, 479–483
277. Williamson, J. M., and Brown, G. M. (1979) *J. Biol. Chem.* **254**, 8074–8082
278. Yang, H., and Abeles, R. H. (1987) *Biochemistry* **26**, 4076 –4081
279. Seto, B., and Stadtman, T. C. (1976) *J. Biol. Chem.* **251**, 2435–2439
280. Seto, B. (1978) *J. Biol. Chem.* **253**, 4525–4529
281. Hodgins, D. S., and Abeles, R. H. (1969) *Arch. Biochem. Biophys.* **130**, 274–285
282. Arkowitz, R. A., and Abeles, R. H. (1989) *Biochemistry* **28**, 4639–4644
283. Stadtman, T. C., and Davis, J. N. (1991) *J. Biol. Chem.* **266**, 22147–22153

283a. Kabisch, U. C., Gräntzdörffer, A., Schierhorn, A., Rücknagel, K. P., Andreesen, J. R., and Pich, A. (1999) *J. Biol. Chem.* **274**, 8445–8454

284. Furuta, T., Takahashi, H., Shibasaki, H., and Kasuya, Y. (1992) *J. Biol. Chem.* **267**, 12600–12605
285. Levy, H. L., Taylor, R. G., and McInnes, R. R. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1107–1123, McGraw-Hill, New York
286. Havir, E. A., and Hanson, K. R. (1985) *Biochemistry* **24**, 2959–2967
287. Consevage, M. W., and Phillips, A. T. (1985) *Biochemistry* **24**, 301–308
288. Langer, M., Reck, G., Reed, J., and Rétey, J. (1994) *Biochemistry* **33**, 6462–6467
289. Langer, M., Lieber, A., and Rétey, J. (1994) *Biochemistry* **33**, 14034–14038
290. Langer, M., Pauling, A., and Rétey, J. (1995) *Angew. Chem. Int. Ed. Engl.* **34**, 1464–1465
291. Schuster, B., and Rétey, J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8433–8437
292. Schwede, T. F., Rétey, J., and Schulz, G. E. (1999) *Biochemistry* **38**, 5355–5361

Study Questions

1. Discuss the role of biotin in metabolism and the chemical mechanism of its action. Illustrate with examples.

2. Biotin has been shown to be an essential component of a bacterial oxaloacetate *decarboxylase* that pumps two sodium ions out of a cell for each oxaloacetate molecule decarboxylated. Propose a chemical mechanism for the functioning of biotin and also any ideas that you may have for the operation of the sodium pump.

3. Write out the equation by which **acetyl-CoA** and **pyruvate** can be converted into **L-glutamate**, which can be observed in living animals by ^{13}C NMR. If [2-^{13}C] sodium acetate were injected into the animal what labeling pattern could be anticipated in the L-glutamate?

4. Compare mechanisms and coenzyme requirements for biological decarboxylation of the following three groups of compounds:

   β-oxoacids
   α-oxoacids
   α-amino acids

5. Illustrate, using structural equations, the chemical mechanisms of the following biochemical reactions.

   a) Pyruvate → acetaldehyde + CO_2
   b) 2-Pyruvate → α-acetolactate + CO_2
   c) Pyruvate + NAD^+ + CoA-SH → acetyl-CoA + NADH + H^+ + CO_2
   d) Fructose 6-*P* + glyceraldehyde-3-*P* → xylulose-5-*P* + erythrose-4-*P*

6. 3-Fluoropyruvate is converted quantitatively by pyruvate decarboxylase from wheat germ into acetate, fluoride (F^-), and carbon dioxide. Propose a reaction mechanism. See Gish, G., Smyth, T., and Kluger, R. (1988) *J. Am. Chem. Soc.* **110**, 6230-6234.

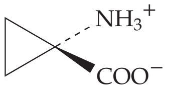
7. In *E. coli* **L-cysteine** is formed from L-serine and the sulfide ion S^{2-} in a reaction that also requires acetyl-CoA and is catalyzed by the consecutive action of an acyl transferase and cysteine synthase. Outline the mechanism of this conversion indicating participation of any essential coenzymes.

8. Illustrate, using structural equations, the chemical mechanisms of the following biochemical reactions.

   a) L-Glutamate + oxaloacetate → 2-oxoglutarate + L-aspartate
   b) L-Serine → pyruvate + NH_4^+
   c) L-Serine + indole (from cleavage of indole-3-glycerol phosphate) → L-tryptophan
   d) L-Serine + tetrahydrofolate → glycine + N^5, N^{10}-methylene-tetrahydrofolate
   e) L-Selenocysteine → L-alanine + Se^0

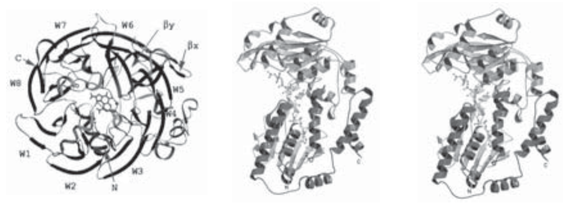
9. Threonine is formed by *E. coli* from homoserine via the intermediate γ-phosphohomoserine. Write out an abbreviated reaction sequence for its conversion to L-threonine by the action of threonine synthase.

10. Write out a plausible step-by-step mechanism by which 1-aminocyclopropane-1-carboxylate synthase (ACC synthase) of plant tissues can form ACC from *S*-adenosylmethionine. This reaction requires a specific cofactor



11. Tissues of the mammalian central nervous system contain a pyridoxal phosphate-dependent glutamate decarboxylase that catalyzes conversion of Glu to γ-aminobutyrate (GABA), an inhibitory synaptic transmitter. GABA is degraded by transimination with α-oxoglutarate as the acceptor to yield succinic semialdehyde, which then is oxidized to succinate by an NAD-linked dehydrogenase.

    a) Show how these reactions can operate as a shunt pathway that allows the citric acid cycle to function without the enzymes α-oxoglutarate dehydrogenase and succinate thiokinase.
    b) Is the shunt more or less efficient than the normal cycle from the standpoint of energy recovery? Explain.



Biological oxidation-reduction reactions rely upon many organic coenzymes and transition metal ions. Left, a molecule of the hydrogen carrier pyrroloquinoline quinone (**PQQ**, Fig. 15-23) is seen at the bottom of a 7-bladed β propeller of a bacterial methanol dehydrogenase. In adrenodoxin reductase (right) the reducing power of **NADPH** is passed to **FAD** and then to the small redox protein adrenodoxin. In this stereoscopic view the pyridine ring of the oxidized coenzyme NADP$^+$ and the tricyclic flavin ring of FAD are seen stacked against each other in the center; the adenylate ends stretching toward the top and bottom of the complex. From Ziegler and Schulz, (2000) *Biochemistry* **39**, 10986-10995.

Contents

767 .... **A. Pyridine Nucleotide Coenzymes and Dehydrogenases**
768 .... 1. Three-Dimensional Structures of Dehydrogenases
768 .... 2. Stereospecificity and Mechanism
770 .... *One step or two-step transfer?*
771 .... *Coenzyme and substrate analogs*
771 .... *Conformational changes during dehydrogenase action*
772 .... *Zinc-containing alcohol dehydrogenases*
774 .... *Other alcohol dehydrogenases and aldo-keto reductases*
775 .... 3. Dehydrogenation of Amino Acids and Amines
775 .... 4. Glyceraldehyde-3-Phosphate Dehydrogenase and the Generation of ATP in Fermentation Reactions
776 .... 5. Reduction of Carboxyl Groups
777 .... 6. Reduction of Carbon–Carbon Double Bonds
777 .... 7. Transient Carbonyl Groups in Catalysis
778 .... 8. ADP Ribosylation and Related Reactions of NAD$^+$ and NADP$^+$
779 .... 9. The Varied Chemistry of the Pyridine Nucleotides
779 .... *Addition to NAD$^+$ and NADP$^+$*
780 .... *Modification of NADH in acid*
780 .... *Other reactions of pyridine nucleotides*
780 .... **B. The Flavin Coenzymes**
783 .... 1. Flavoproteins and Their Reduction Potentials
783 .... 2. Typical Dehydrogenation Reactions Catalyzed by Flavoproteins
788 .... 3. More Flavoproteins
788 .... 4. Modified Flavin Coenzymes
789 .... 5. Mechanisms of Flavin Dehydrogenase Action
792 .... 6. Half-Reduced Flavins
794 .... 7. Metal Complexes of Flavins and Metalloflavoproteins
794 .... 8. Reactions of Reduced Flavins with Oxygen
795 .... **C. Lipoic Acid and the Oxidative Decarboxylation of α-Oxoacids**
795 .... 1. Chemical Reactions of Lipoic Acid
796 .... 2. Oxidative Decarboxylation of Alpha-Oxoacids
798 .... 3. Other Functions of Lipoic Acid
799 .... 4. Additional Mechanisms of Oxidative Decarboxylation
799 .... *Pyruvate oxidase*
799 .... *Pyruvate:ferredoxin oxidoreductase*
799 .... *Oxidative decarboxylation by hydrogen peroxide*
800 .... *Pyruvate formate-lyase reaction*
802 .... 5. Cleavage of α-Oxoacids and Substrate-Level Phosphorylation
802 .... **D. Tetrahydrofolic Acid and Other Pterin Coenzymes**
802 .... 1. Structure of Pterins
803 .... 2. Folate Coenzymes
804 .... 3. Dihydrofolate Reductase
805 .... *Inhibitors*
805 .... *Structure and mechanism*
808 .... 4. Single-Carbon Compounds and Groups in Metabolism
808 .... *Serine as a C1 donor*
808 .... *The glycine decarboxylase–synthetase system*
810 .... *Starting with formate or carbon dioxide*
810 .... *N^5-Formyl-THF (leucovorin)*
810 .... *Catabolism of histidine and purines*
810 .... *Thymidylate synthase*
813 .... *Synthesis of methyl groups*
813 .... **E. Specialized Coenzymes of Methanogenic Bacteria**
815 .... **F. Quinones, Hydroquinones, and Tocopherols**
815 .... 1. Pyrroloquinoline Quinone (PQQ)
816 .... 2. Copper Amine Oxidases
816 .... *Topaquinone (TPQ)*
817 .... *Lysine tyrosylquinone (LTQ)*
817 .... *Tryptophan tryptophanylquinone (TTQ)*
817 .... *Three-dimensional structures*
817 .... *Mechanisms*
818 .... 3. Ubiquinones, Plastoquinones, Tocopherols, and Vitamin K
819 .... 4. Quinones as Electron Carriers
820 .... 5. Vitamin K and γ-Carboxyglutamate Formation
822 .... 6. Tocopherols (Vitamin E) as Antioxidants
822 .... **G. Selenium-Containing Enzymes**

828 .... **References**
835 .... **Study Questions**

Boxes

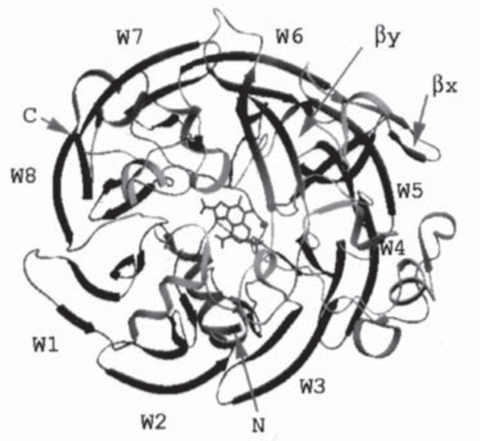
769 .... Box 15-A Nicotinic Acid and Nicotinamide
783 .... Box 15-B Riboflavin
786 .... Box 15-C Thioredoxin and Glutaredoxin
802 .... Box 15-D Folic Acid (Pteroylglutamic Acid)
812 .... Box 15-E Thymidylate Synthase, a Target Enzyme in Cancer Chemotherapy
821 .... Box 15-F The Vitamin K Family
823 .... Box 15-G The Vitamin E Family: The Tocopherols
826 .... Box 15-H Glutathione Peroxidase and Abnormalities of Red Blood Cells

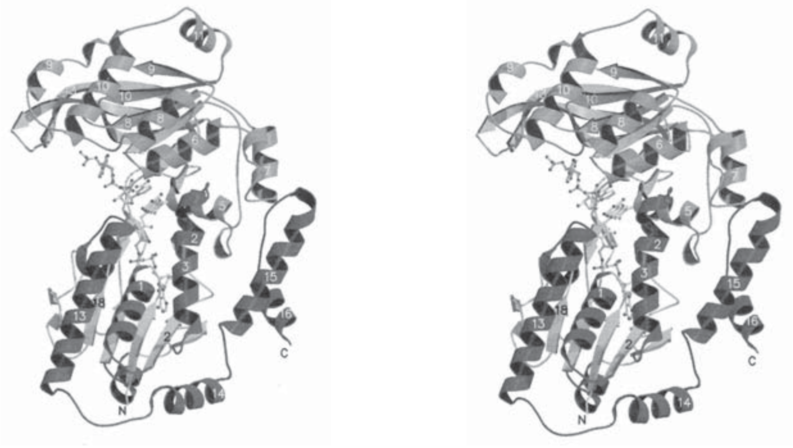
Tables

766 .... Table 15-1 Some Biochemical Hydrogen Transfer Reactions
784 .... Table 15-2 Some Dehydrogenation Reactions Catalyzed by Flavoproteins
808 .... Table 15-3 Single-Carbon Compounds in Order of Oxidation State of Carbon
825 .... Table 15-4 Selenium-Containing Proteins

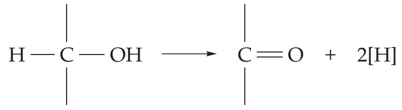
Coenzymes of Oxidation–Reduction Reactions

15





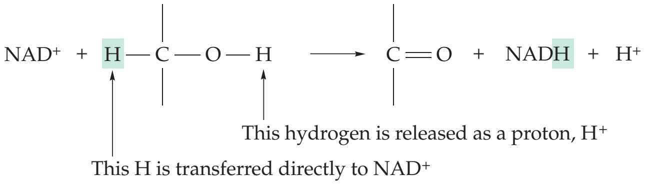
The **dehydrogenation** of an alcohol to a ketone or aldehyde (Eq. 15-1) is one of the most frequent biological oxidation reactions. Although the hydrogen atoms removed from the substrate are often indicated simply as 2[H], it was recognized early in the twentieth century that they are actually transferred to hydrogen-carrying coenzymes such as NAD^+, $NADP^+$, FAD, and **riboflavin 5'-phosphate** (FMN). This chapter deals with these coenzymes and also with a number of other organic oxidation–reduction coenzymes. They may be considered either as carriers of hydrogen or as carriers of electrons ($H = H^+ + e^-$) in metabolic reactions.



$$\text{H}-\overset{|}{\underset{|}{\text{C}}}-\text{OH} \longrightarrow \overset{|}{\underset{|}{\text{C}}}{=}\text{O} + 2[\text{H}]$$

(15-1)

When NAD^+ becomes reduced by dehydrogenation of an alcohol, one of the hydrogen atoms removed from the alcohol becomes firmly attached to the NAD^+, converting it to NADH. The other is released as a proton (Eq. 15-2). Study of ^2H-labeled alcohols and their oxidation by NAD^+ has shown that *dehydrogenases catalyze direct transfer to NAD^+ of the hydrogen that is attached to carbon in the alcohol*. There is never any exchange of this hydrogen atom with protons of the medium. At the same time, the hydrogen attached to the oxygen of the alcohol is released into the medium as H^+:



$$NAD^+ + \boxed{\text{H}}-\overset{|}{\underset{|}{\text{C}}}-\text{O}-\text{H} \longrightarrow \overset{|}{\underset{|}{\text{C}}}{=}\text{O} + \text{NAD}\boxed{\text{H}} + H^+$$

This hydrogen is released as a proton, H^+

This H is transferred directly to NAD^+

(15-2)

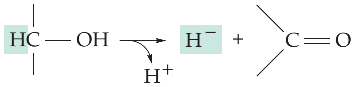
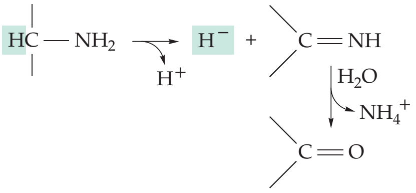
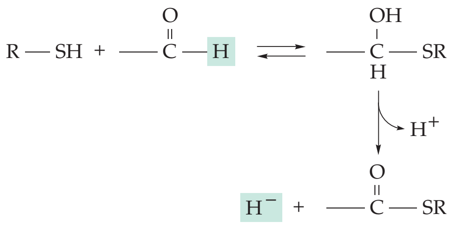
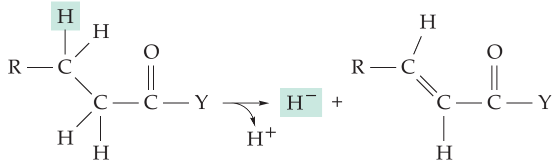
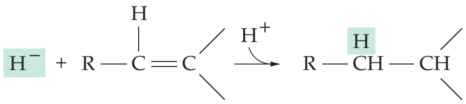
The foregoing observations suggested that these biological dehydrogenations may be viewed as removal of a hydride ion (H^-) together with a proton (H^+) rather than as removal of two hydrogen atoms. NAD^+ and $NADP^+$ are regarded as hydride ion-accepting coenzymes. However, it has been impossible to establish conclusively that the hydrogen atom and electron are transferred simultaneously as H^-. Transfer of the hydrogen atom to or from these coenzymes may conceivably be followed by or preceded by transfer of an electron. The situation is even less clear for the flavin coenzymes FAD and riboflavin phosphate for which intermediate **free radical** oxidation states are known to exist. However, regardless of the actual mechanism, it is convenient to classify most hydrogen transfer reactions of metabolism as if they occurred by transfer of a hydride ion. *The hydride ion can be regarded as a nucleophile which can add to double bonds or can be eliminated from substrates in reactions of types that we have already considered.* Some of these reactions are listed in Table 15-1.

Why are there *four* major hydrogen transfer coenzymes, NAD^+, $NADP^+$, FAD, and riboflavin phosphate (FMN), instead of just one? Part of the answer is that the **reduced pyridine nucleotides** NADPH and NADH are more powerful reducing agents than are reduced **flavins** (Table 6-7). Conversely, flavin coenzymes are more powerful oxidizing agents than are

NAD$^+$ and NADP$^+$. This difference reflects the chemical difference between the vitamins **riboflavin** and **nicotinamide** which form the oxidation–reduction centers of the coenzymes. Another difference is that NAD$^+$ and NADP$^+$ tend to be present in free forms within cells, diffusing from a site on one enzyme to a site on another. These coenzymes are *sometimes* tightly bound but flavin coenzymes are *usually* firmly bound to proteins, fixed, and unable to move. Thus, they tend to accept hydrogen atoms from one substrate and to pass them to a second substrate while attached to a single enzyme.

The oxidation–reduction potential of a pyridine nucleotide coenzyme system is determined by the standard redox potential for the free coenzyme (Table 6-8) together with the ratio of concentrations of oxidized to reduced coenzyme ([NAD$^+$] / [NADH], Eq. 6-64). If these concentrations are known, a redox

TABLE 15-1
Some Biochemical Hydrogen Transfer Reactions[a]

Reaction	Example (oxidant)
A Dehydrogenation of an alcohol 	Alcohol dehydrogenase (NAD$^+$)
B Dehydrogenation of an amine 	Amino acid dehydrogenases, amine oxidases (NAD$^+$ or flavin)
C Dehydrogenation of adduct of thiol and aldehyde 	Glyceraldehyde 3-phosphate dehydrogenase (NAD$^+$ or NADP$^+$)
D Dehydrogenation of acyl-CoA, acyl-ACP, or carboxylic acid  (Y = —S—CoA or —OH)	Acyl-CoA dehydrogenases (Flavin) Succinic dehydrogenase (Flavin) Opposite: enoyl reductase (NADPH)
E Reduction of desmosterol to cholesterol by NADPH 	Reduction of desmosterol to cholesterol by NADPH

[a] Reaction type 9 A,B of Table on the inside cover at the end of the book: A hypothetical hydride ion H$^-$ is transferred from the substrate to a coenzyme of suitable reduction potential such as NAD$^+$, NADP$^+$, FAD, or riboflavin 5′-phosphate. The reverse of hydrogenation is shown for E. Many of the reactions are reversible and often go spontaneously in the reverse direction from that shown here.

potential can be defined for the NAD$^+$ system within a cell. This potential may vary in different parts of the cell because of differences in the [NAD$^+$] / [NADH] ratio, but within a given region of the cell it is constant. On the other hand, the redox potentials of flavoproteins vary. Since the flavin coenzymes are not dissociable, two flavoproteins may operate at very different potentials even when they are physically close together.

Why are there two pyridine nucleotides, NAD$^+$ and NADP$^+$, differing only in the presence or absence of an extra phosphate group? One important answer is that they are members of two different oxidation–reduction systems, both based on nicotinamide but functionally independent. The experimentally measured ratio [NAD$^+$] / [NADH] is much higher than the ratio [NADP$^+$] / [NADPH]. Thus, these two coenzyme systems also can operate within a cell at different redox potentials. A related generalization that holds much of the time is that *NAD$^+$ is usually involved in pathways of catabolism, where it functions as an oxidant, while NADPH is more often used as a reducing agent in biosynthetic processes.* See Chapter 17, Section I for further discussion.

A. Pyridine Nucleotide Coenzymes and Dehydrogenases

In 1897, Buchner discovered that "yeast juice," prepared by grinding yeast with sand and filtering, catalyzed fermentation of sugar. This was a major discovery which excited the interest of many other biochemists.[1] Among them were Harden and Young, who, in 1904, showed that Buchner's cell-free yeast juice lost its ability to ferment glucose to alcohol and carbon dioxide when it was dialyzed. Apparently, fermentation depended upon a low-molecular-weight substance that passed out through the pores of the dialysis membrane. Fermentation could be restored by adding back to the yeast juice either the concentrated dialysate or boiled yeast juice (in which the enzyme proteins had been destroyed). The heat-stable material, which Harden and Young called **cozymase**, was eventually found to be a mixture of inorganic phosphate ions, thiamin diphosphate, and NAD$^+$. However, characterization of NAD$^+$ was not accomplished until 1935.

Pure NADP$^+$ was isolated from red blood cells in 1934 by Otto Warburg and W. Christian, who had been studying the oxidation of glucose 6-phosphate by erythrocytes.[1a] They demonstrated a requirement for a dialyzable coenzyme which they characterized and named **triphosphopyridine nucleotide** (TPN$^+$, but now officially NADP$^+$; Fig. 15-1). Thus, even before its recognition as an important vitamin in human nutrition, nicotinamide was identified as a component of NADP$^+$.

Warburg and Christian recognized the relationship of NADP$^+$ and NAD$^+$ (then called DPN$^+$) and proposed that both of these compounds act as hydrogen carriers through alternate reduction and oxidation of the pyridine ring. They showed that the coenzymes could be reduced either enzymatically or with sodium dithionite $Na_2S_2O_4$.

$$S_2O_4^{2-} + NAD^+ + H^+ \rightarrow NADH + 2\ SO_2 \tag{15-3}$$

The reduced coenzymes NADH and NADPH were characterized by a new light-absorption band at 340 nm. This is not present in the oxidized forms, which absorb maximally at 260 nm (Fig. 15-2). The reduced forms are stable in air, but their reoxidation was found

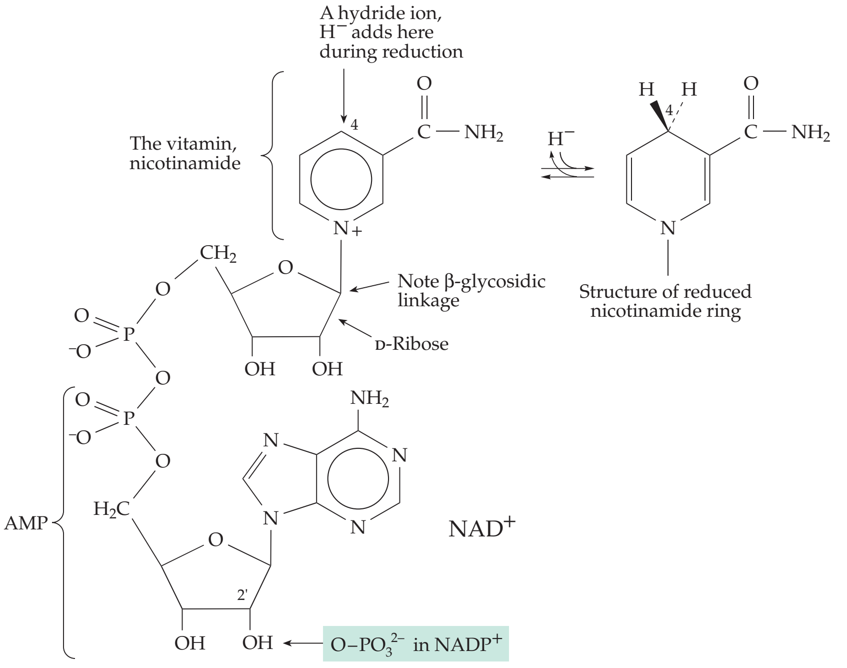

A hydride ion, H⁻ adds here during reduction
4
O
C — NH2
The vitamin, nicotinamide
N+
H⁻
H H
4
O
C — NH2
N
Structure of reduced nicotinamide ring
CH2
O
O
Note β-glycosidic linkage
O
P
⁻O
D-Ribose
OH OH
O
O
P
⁻O
O
NH2
N
N
N
N
H2C
AMP
NAD+
O
2'
OH OH
O–PO3 2− in NADP+

Figure 15-1 The hydrogen-carrying coenzymes NAD$^+$ (nicotinanide adenine dinucleotide) and NADP$^+$ (nicotinamide adenine dinucleotide phosphate). We use the abbreviations NAD$^+$ and NADP$^+$, even though the net charge on the entire molecule at pH 7 is negative because of the charges on oxygen atoms of the phospho groups.

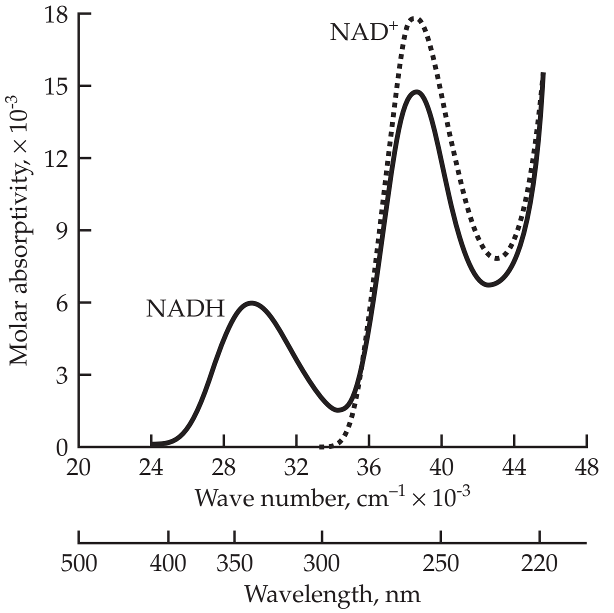


Figure 15-2 Absorption spectra of NAD$^+$ and NADH. Spectra of NADP$^+$ and NADPH are nearly the same as these. The difference in absorbance between oxidized and reduced forms at 340 nm is the basis for what is probably the single most often used spectral measurement in biochemistry. Reduction of NAD$^+$ or NADP$^+$ or oxidation of NADH or NADPH is measured by changes in absorbance at 340 nm in many methods of enzyme assay. If a pyridine nucleotide is not a reactant for the enzyme being studied, a **coupled assay** is often possible. For example, the rate of enzymatic formation of ATP in a process can be measured by adding to the reaction mixture the following enzymes and substrates: hexokinase + glucose + glucose-6-phosphate dehydrogenase + NADP$^+$. As ATP is formed, it phosphorylates glucose via the action of hexokinase. NADP$^+$ then oxidizes the glucose 6-phosphate that is formed with production of NADPH, whose rate of appearance is monitored at 340 nm.

to be catalyzed by certain yellow enzymes which were later identified as flavoproteins.

1. Three-Dimensional Structures of Dehydrogenases

Most NAD$^+$- or NADP$^+$- dependent dehydrogenases are dimers or trimers of 20- to 40-kDa subunits. Among them are some of the first enzymes for which complete structures were determined by X-ray diffraction methods. The structure of the 329-residue per subunit muscle (M_4) isoenzyme of **lactate dehydrogenases** (see Chapter 11) from the dogfish was determined to 0.25 nm resolution by Rossmann and associates in 1971.[2–4] More recently, structures have been determined for mammaliam muscle and heart type (H_4) isoenzymes,[5] for the testicular (C_4) isoenzyme from the mouse,[6] and for bacterial lactate dehydrogenases.[6–8] In all of these the polypeptide is folded nearly identically. The structures of the homologous cytosolic and mitochondrial isoenzymes of **malate dehydrogenase** are also similar,[9–11] as are those of the bacterial enzyme.[12,13] All of these proteins consist of two structural domains and the NAD$^+$ is bound to the nucleotide-binding domain in a similar manner as is shown in Fig. 2-13 for glyceraldehyde phosphate dehydrogenase. The coenzyme-binding domains of the dehydrogenases of known structures all have this nearly constant structural feature (often called the Rossmann fold) consisting of a six-stranded parallel sheet together with several α helical coils[14] (Figs. 2-13 and 2-27).

The coenzyme molecule curls around one end of the nucleotide-binding domains in a "C" conformation with the nicotinamide ring lying in a pocket (Figs. 2-13, 15-3). Even before the crystal structure of lactate dehydrogenase was known, the lack of pH dependence of coenzyme binding from pH 5 to 10 together with observed inactivation by butanedione suggested that the pyrophosphate group of NAD$^+$ binds to a guanidinium group of an arginine residue. This was identified by X-ray diffraction studies as Arg 101. This ion pairing interaction, as well as the hydrogen bond between Asp 53 and the 2' oxygen atom of a ribose ring (Fig. 15-3), is present in all of the lactate and malate dehydrogenases.

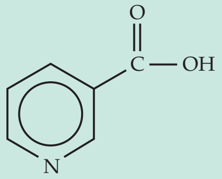
The adenine ring of the coenzyme is bound in a hydrophobic pocket with its amino group pointed out into the solvent. A second structural domain holds additional catalytic groups needed to form the active site.

2. Stereospecificity and Mechanism

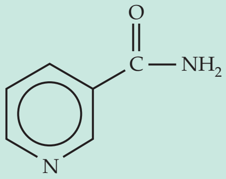
When NAD$^+$ is reduced in ^2H_2O by dithionite (Eq. 15-3) an atom of ^2H is introduced into the reduced pyridine. Chemical degradation showed the ^2H to be present at the 4 position of the ring para to the nitrogen atom[15] (see Fig. 15-1). As shown by Westheimer and associates, during enzymatic reduction of NAD$^+$ by deuterium-containing ethanol, $CH_3{-}C^2H_2OH$, one of the ^2H atoms is transferred into the NADH formed, thus establishing the *direct transfer of a hydrogen atom*.[16,17] When the NAD^{2}H formed in this way is reoxidized enzymatically with acetaldehyde, with regeneration of NAD$^+$ and ethanol, the ^2H is completely removed.

This was one of the first recognized examples of the ability of an enzyme to choose between two identical atoms at a *pro*-chiral center (Chapter 9). The two sides of the nicotinamide ring of NAD were designated A and B and the two hydrogen atoms at the 4 position of NADH as H_A (now known as *pro-R*) and H_B (*pro-S*). Alcohol dehydrogenase always removes the *pro-R* hydrogen. Malate, isocitrate, lactate, and D-glycerate dehydrogenases select the same hydrogen. However,

BOX 15-A NICOTINIC ACID AND NICOTINAMIDE



Nicotinic acid



Nicotinamide

Nicotinic acid was prepared in 1867 by oxidation of nicotine. Although it was later isolated by Funk and independently by Suzuki in 1911–1912 from yeast and rice polishings, it was not recognized as a vitamin. Its biological significance was established in 1935 when nicotinamide was identified as a component of NAD^+ by von Euler and associates and of $NADP^+$ by Warburg and Christian.[a] Both forms of the vitamin are stable, colorless compounds highly soluble in water.

In 1937, Elvehjem and coworkers demonstrated that nicotinic acid cured canine "blacktongue." In the same year it was found to cure human **pellagra**, a terrible disease characterized by weakness, indigestion, and loss of appetite followed by dermatitis, diarrhea, mental disorders, and eventual death. At that time pellagra was common in the United States, especially in the south. The U.S. Public Health Service estimated that during 1912–1916 there were 100,000 victims and 10,000 deaths a year.[b,c]

The daily requirement for an adult is about 7.5 mg. The amount is decreased by the presence in the diet of tryptophan, which can be converted partially to nicotinic acid (Chapter 25).[d,e] Tryptophan is about 1/60 as active as nicotinic acid itself. The one-time prevalence of pellagra in the southern United States was a direct consequence of a diet high in maize whose proteins have an unusually low tryptophan content.

---

[a] Schlenk, F. (1984) *Trends Biochem. Sci.* **9**, 286–288

[b] Wagner, A. F., and Folkers, K. (1964) *Vitamins and Coenzymes*, Wiley (Interscience), New York (p. 73)

[c] Rosenkrantz, B. G. (1974) *Science* **183**, 949–950

[d] Teply, L. J. (1993) *FASEB J.* **7**, 1300

[e] Sauberlich, H. E. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical and Medical Aspects*, Vol. B (Dolphin, D., Avramović, O., and Poulson, R., eds), pp. 599–626, Wiley (Interscience), New York

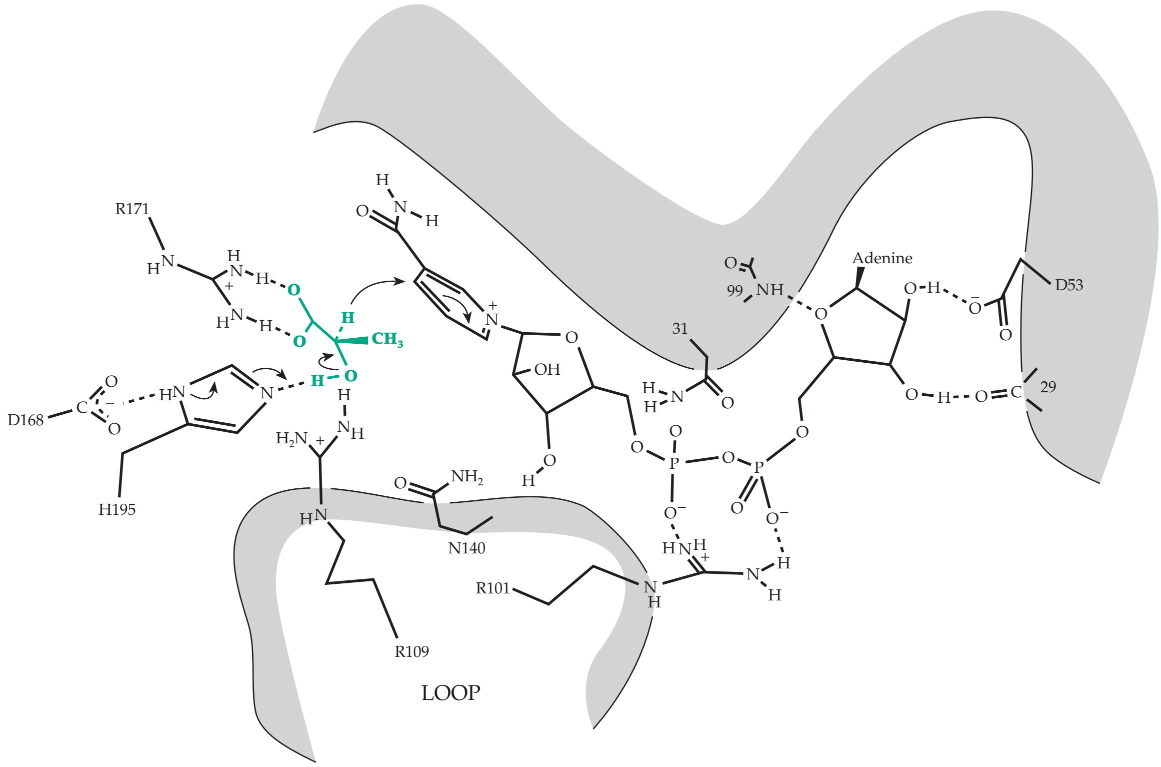
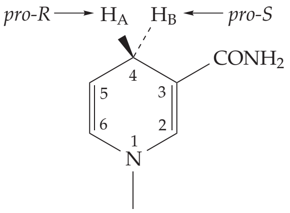


Figure 15-3 Diagrammatic structure of NAD^+ and L-lactate bound into the active site of lactate dehydrogenase. See Eventhoff *et al.*[5]

dehydrogenases acting on glucose 6-phosphate, glutamate, 6-phosphogluconate, and 3-phosphoglyceraldehyde remove the *pro-S* hydrogen. By 1979 this stereospecificity had been determined for 127 dehydrogenases, about half of which were found A-specific and half B-specific.[18] Isotopically labeled substrates have usually been used for this purpose but a simple NMR method has been devised.[19,20]



How complete is the stereospecificity? Does the enzyme sometimes make a mistake? *Lactate dehydrogenase displays nearly absolute specificity*, transferring a proton into the "incorrect" *pro-S* position no more than once in 5 x 10^7 catalytic cycles.[21] This suggests a difference in transition-state energies of about 40 kJ / mol for the two isomers. X-ray structural studies suggested that a major factor determining this stereoselectivity might be the location of hydrogen-bonding groups that hold the $-CONH_2$ group of the nicotinamide. If these are located in such a way that the nicotinamide ring has a *syn* orientation with respect to the ribose ring to which it is attached (Fig. 15-4), the A face of the coenzyme will be against the substrate undergoing oxidation or reduction. If the ring has an *anti* conformation, the B side will be against the substrate.[18,22] However, the explanation cannot be this simple. High specificity is still retained when NADH is modified, for example, by replacement of the $-CONH_2$ on the pyridine ring with $-COCH_3$ or $-CHO$ or by use of α-NADH, in which the ribose–nicotinamide linkage is α instead of β as in normal NADH.[23]

An intriguing idea is that the enzyme selectively stabilizes one of the boat conformations of NADH or NADPH. According to stereoelectronic principles the axially oriented hydrogen in such a boat structure (*pro-R* in Fig. 15-4) will be the most readily transferred. On the basis of this principle, together with the idea that enzymes adjust the Gibbs energies of intermediate states to achieve optimum catalytic rates, the following prediction was made by Benner and associates:[24] *The thermodynamically most easily reduced carbonyl compounds will react by enzymatic transfer of the hydride ion from the pro-R position of NADH.* Conversely, the most difficultly reduced carbonyl compounds will receive the hydride ion from the *pro-S* position of NADH. The proposal has been controversial and it has been argued that evolutionary relationships play a dominant role in determining stereoselectivity.[23] Theoretical computations suggest that the boatlike puckering of the ring is flexible, raising some doubts about the rigidity of the reactants in the active site.[12,25] Nevertheless, it is likely that the coenzyme as well as bound substrates must assume very specific conformations before the enzyme is able to move to the transition state for the reaction. A great deal of effort has been expended in trying to understand other factors that may explain the high stereospecificity of dehydrogenases.[12,22,26–28]

When a hydrogen atom is transferred by an enzyme from the 4 position of NADH or NADPH to an aldehyde or ketone to form an alcohol, the placement of the hydrogen atom on the alcohol is also stereospecific. Thus, alcohol dehydrogenase acting on NAD^{2}H converts acetaldehyde to (*R*)-mono-[^{2}H]ethanol (Eq. 9-73). Pyruvate is reduced by lactate dehydrogenase to L-lactate, and so on. Even mutations that disrupt the binding of the carboxylate of lactate, e.g., substitution of arginine 171 (see Fig. 15-3) by tryptophan or phenylalanine and introduction of a new arginine on the other side of the active site, do not alter the L-stereospecificity. However, the specificity for lactate is lost.[29]

One step or two-step transfer? Another major question about dehydrogenases is whether the hydrogen atom that is transferred moves as a hydride ion, as is generally accepted, or as a hydrogen atom with separate transfer of an electron and with an intermediate NAD or NADPH free radical. In one study para-substituted benzaldehydes were reduced with NADH and NAD^{2}H using yeast alcohol dehydrogenase as a catalyst.[30] This permitted the application of the Hammett equation (Box 6-C) to the rate data. For a series of benzaldehydes for which σ^+ varied widely, a value

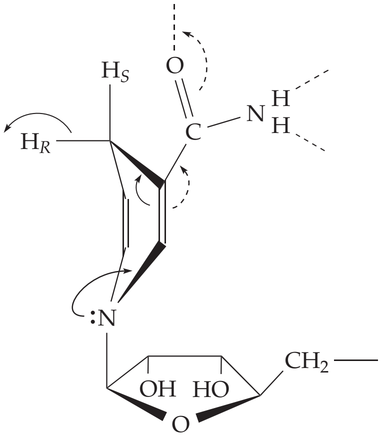


Figure 15-4 The nicotinamide ring of NADH in a *syn* boat conformation suitable for transfer of an axially oriented *pro-R* hydrogen atom from its A face as H^-. The flow of electrons is shown by the solid arrows. The dashed arrows indicate competing resonance which favors planarity of the ring and opposes the H^- transfer. Hydrogen bonds from the protein to the carboxamide group (dashed lines) affect both this tendency and the conformation of the nucleotide.

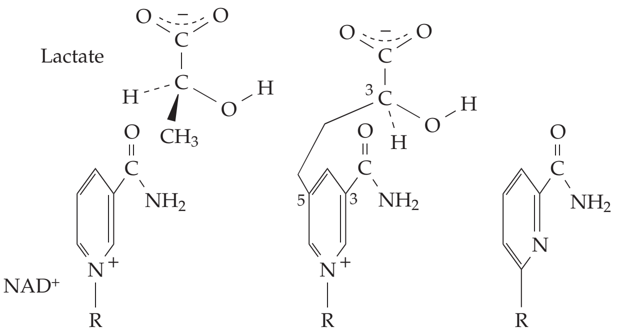
of ρ = +2.2 was observed for the rate constant with both NADH and NAD^{2}H. Thus, electron-accepting substituents in the para position hasten the reaction. While the significance of this observation is not immediately obvious, the relatively low value of ρ is probably incompatible with a mechanism that requires complete transfer of a single electron from NADH to acetaldehyde in the first step. A primary isotope effect on the rates was $k_H / k_D = 3.6$, indicating that the C–H bond in NADH is broken in the rate-limiting step. The fact that the isotope effect is the same for all of the substituted benzaldehydes also argued in favor of a hydride ion transfer.

Studies of the kinetics of nonenzymatic model reactions of NADH with quinones[31,32] have also been interpreted to favor a single-step hydride ion transfer. Application of Marcus theory (Chapter 9, Section D,4) to data from model systems also supports the hydride transfer mechanism.[33,34] Quantum mechanical tunneling may be involved in enzymatic transfer of protons, hydride ions, and electrons.[35–36] Tunneling is often recognized by unusually large primary or secondary kinetic isotope effects. According to semi-classical theory (Chapter 12), the maximum effects for deuterium and tritium are given by the following ratios:

$$(k_{2_\mathrm{H}} / k_{1_\mathrm{H}}) = (k_{3_\mathrm{H}} / k_{2_\mathrm{H}})^{3.3} \approx 7.$$

Higher ratios, which suggest tunneling, are frequently observed for dehydrogenases.[36a,b] Tunneling is apparently coupled to fluctuations in motion within the enzyme-substrate complex. Study of effects of pressure on reactions provides a new approach that can aid interpretation.[36c,36d]

Coenzyme and substrate analogs. The structures of enzyme•NAD$^+$•substrate complexes (Fig. 15-3) may be studied by X-ray crystallography under certain conditions or can be inferred from those of various stable enzyme-inhibitor complexes or from enzyme reconstituted with NAD$^+$ that has been covalently linked to the substrate. In the following diagram,[27] NAD$^+$ (I) is shown with L-lactate lying next to its A face, ready to transfer a hydride ion to the *pro-R* position in NADPH. Also shown (II) is 5-(*S*)-lac-NAD$^+$, with the covalently linked lactate portion in nearly the same position as in diagram I. This NAD derivative was used to obtain the first 0.27-nm structure with a bound substrate-like molecule in lactate dehydrogenase.[37] Since then the structures of many complexes with a variety of dehydrogenases have been studied. In coenzyme analog III the ring is bound to ribose with a C–C bond and it lacks the positive charge on the nicotinamide ring in NAD$^+$. It does not react with substrate. However, it binds to the coenzyme site of alcohol dehydrogenase and forms with ethanol a ternary complex whose structure has been solved.[38]

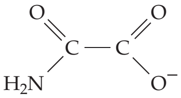


R = Ribose-*P*-*P*-adenosine

I II III

5-(*S*)-lac-NAD$^+$,
(3*S*)-5-(3-carboxy-3
hydroxypropyl)-NAD$^+$

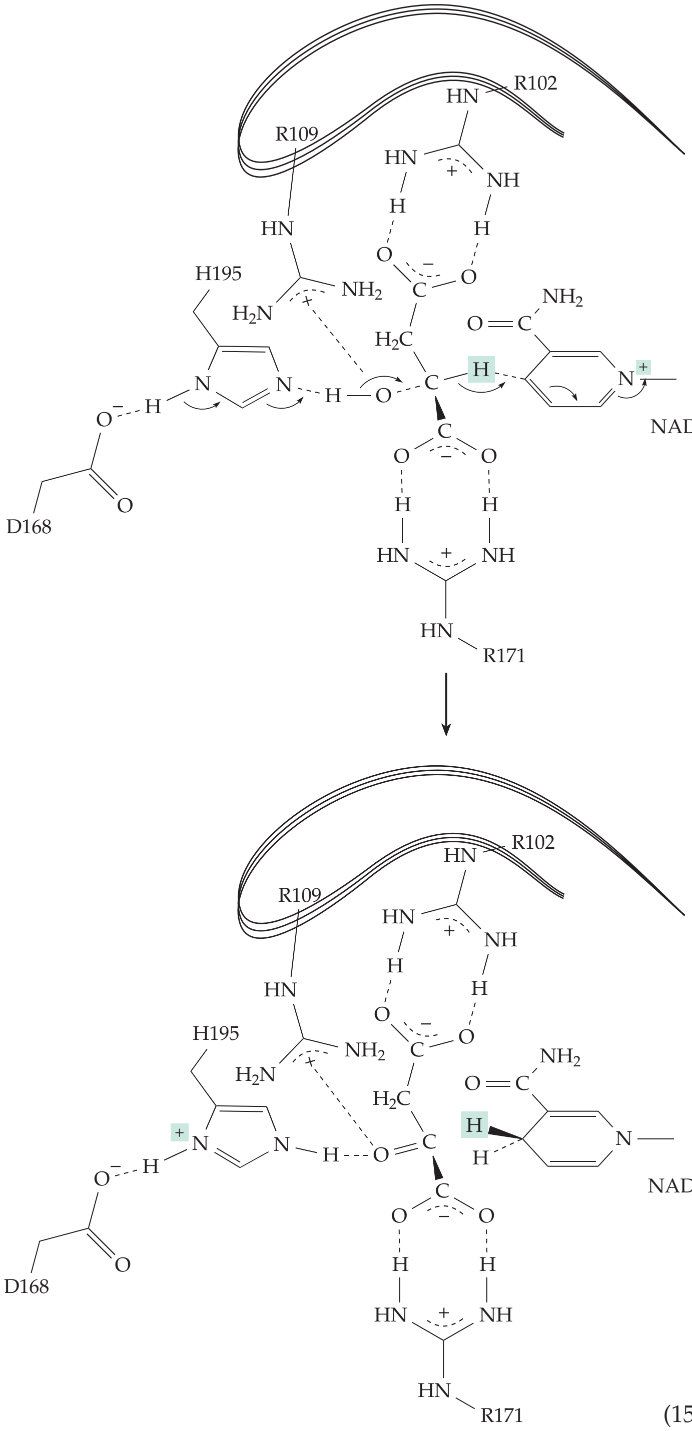
A related approach is to study complexes formed with normal NAD$^+$ but with an unreactive second substrate. An example is oxamate, which binds well to lactate dehydrogenase to form stable ternary complexes for which equilibrium isotope effects have been studied.[39]



Oxamate

In the structure of the lactate dehydrogenase active site shown in Fig. 15-3, the lactate carboxylate ion is held and neutralized by the guanidinium group of Arg 171, and the imidazole group of His 195 is in position to serve as a general base catalyst to abstract a proton from the hydroxyl group of the substrate (Eq. 15-4). This imidazole group is also hydrogen bonded to the carboxylate of Asp 168 as in the “charge-relay” system of the serine proteases (Fig. 12-10). The same features are present in the active site of malate dehydrogenases and have been shown essential by study of various mutant forms.[11,40–42] The His:Asp pair of the dehydrogenases is not part of the nucleotide-binding domain but is present in the second structural domain, the “catalytic domain.” This is another feature reminiscent of the serine proteases. A bacterial D-glycerate dehydrogenase also has a similar structure and the same catalytic groups. However, the placement of the catalytic histidine and the arginine that binds the α-carboxylate group of the substrate is reversed, allowing the enzyme to act on the D-isomer.[43]

Conformational changes during dehydrogenase action. Dehydrogenases bind coenzyme and substrate in an ordered sequence. The coenzyme binds first, then the oxidizable or reducible substrate. The binding of the coenzyme to lactate dehydrogenase is accompanied by a conformational change by which a loop, involving residues 98–120 and including one helix,



(15-4)

folds over the coenzyme like a lid.[41,44] This conformational change must occur during each catalytic cycle, just as in the previously discussed cases of citrate synthase (Chapter 13) and aspartate aminotransferase (Chapter 14). One effect of folding of the loop is to bring the side chain of Arg 109 into close proximity to His 195 and to the OH group of the bound lactate (Eq. 15-4). The closing of the loop also forces the positively charged nicotinamide ring more deeply into a relatively nonpolar pocket. This may induce a movement of the positive charge toward the 4' carbon of the ring, assisting in the transfer of the hydride ion (Eq. 15-4).[41] At the end of this transfer both Arg 109 and His 196 are positively charged and electrostatic repulsion between them may help to move the loop and release the products. If the reaction proceeds in the opposite direction the presence of two positive charges will assist in the hydride ion transfer. The importance of Arg 109 is demonstrated by the fact that a mutant with glutamine in place of Arg 109 has a value of k_{cat} only 1/400 that of native enzyme.[45]

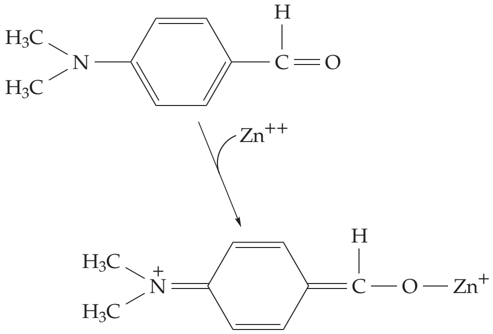
The loop closure also causes significant changes in the Raman spectra of the bound NAD$^+$, especially in vibrational modes that involve the carboxamide group of the nicotinamide ring. These have been interpreted as indicating an increased ridigity of binding of the coenzyme in the closed conformation.[46] As with many other enzymes acting on polar substrates, a characteristic of the pretransition state complex appears to be formation of a complex with a network of hydrogen bonds extending into the protein, exclusion of most water molecules, and tight packing of protein groups around the substrate. It is also significant that an overall net electrical charge on the ES complexes must be correct for tight binding of substrates to occur for lactate or malate dehydrogenase.[44] Positive and negative charges are balanced except for one excess positive charge which may be needed for catalysis.

Zinc-containing alcohol dehydrogenases. Liver alcohol dehydrogenase is a relatively nonspecific enzyme that oxidizes ethanol and many other alcohols. The much studied horse liver enzyme[47–60] is a dimer of 374-residue subunits, each of which contains a "catalytic" Zn^{2+} ion deeply buried in a crevice between the nucleotide-binding and catalytic domains. The enzyme also contains a "structural" Zn^{2+} ion[48] that is bound by four sulfur atoms from cysteine side chains but does not represent a conserved feature of all alcohol dehydrogenases. The catalytic Zn^{2+} is ligated by sulfur atoms from cysteines 46 and 174 and by a nitrogen atom of the imidazole group of His 67 (Fig. 15-5). In the free enzyme a water molecule is thought to occupy the 4th coordination position and its dissociation to form the Zn$^+$–HO$^-$ complex (Eq. 15-5) may account for a pK_a of 9.2 in the free enzyme and of 7.6 in the NAD$^+$ complex.[48] The apparent pK_a drops further to about 6.4 in the presence of the substrate

$$Zn^{2+}–OH_2 \rightarrow Zn^{+}–OH + H^{+} \qquad (15\text{-}5)$$

ethanol. The assignment of this pK_a value has been controversial.[48,49] Histidine 51 (Fig. 15-5), Glu 68 and Asp 49 have side chains close to the zinc and, as we have seen (Chapter 7), macroscopic pK_a values of proteins can be *shared* by two or more closely placed groups.

Substrate binding also induces a conformational change in this enzyme. When both coenzymes and substrate bind the "closed" conformation of the enzyme is formed by a rotation of the catalytic domains of the two subunits relative to the coenzyme-binding domains.[50–51a] Structures of ternary complexes with inhibitors and with substrates have also been established. For example, liver alcohol dehydrogenase was crystallized as the enzyme•NAD$^+$• *p*-bromobenzyl alcohol complex with saturating concentrations of substrates in an equilibrium mixture[51b] and studied at low resolution. Transient kinetic studies or direct spectroscopic determinations led to the conclusion that the internal equilibrium (E•NAD$^+$•alcohol = E•NADH•aldehyde) favors the NAD$^+$•alcohol complex.[52] Subsequently, the complex was studied at higher resolution, and the basic structural features were confirmed with a structure of the enzyme complexed with NAD$^+$ and 2,3,4,5,6-pentafluorobenzyl alcohol.[53] From the crystal structures of the NAD$^+$ • *p*-bromobenzyl alcohol complex and the previously mentioned complex with ethanol and analog III it appears that the oxygen of the alcohol substrate coordinates with the Zn^{2+}, displacing the bound water. Binding of the chromophoric aldehyde 4-*trans*-(*N*, *N*-dimethylamino) cinnamaldehyde shifts the wavelength of maximum absorbance by 66 nm, suggesting that Zn^{2+} binds directly to the oxygen of this ligand (Eq. 15-6).[54] Resonance-enhanced Raman spectroscopy (Chapter 23) of the complex of dimethylaminobenzaldehyde with alcohol dehydrogenase also supports an intermediate in which the Zn^{2+} becomes coordinated directly with the substrate oxygen (Eq. 15-6).[55] Rapid scanning spectrophotometry of complexes with another chromophoric substrate, 3-hydroxy-4-nitrobenzyl alcohol, also suggested that the alcohol first formed a complex with an undissociated alcoholic – OH group (Eq. 15-7, step *a*) and then lost a proton to form a zinc alcoholate complex (step *b*)[56] which could react by hydride ion transfer (step *c*).



(15-6)

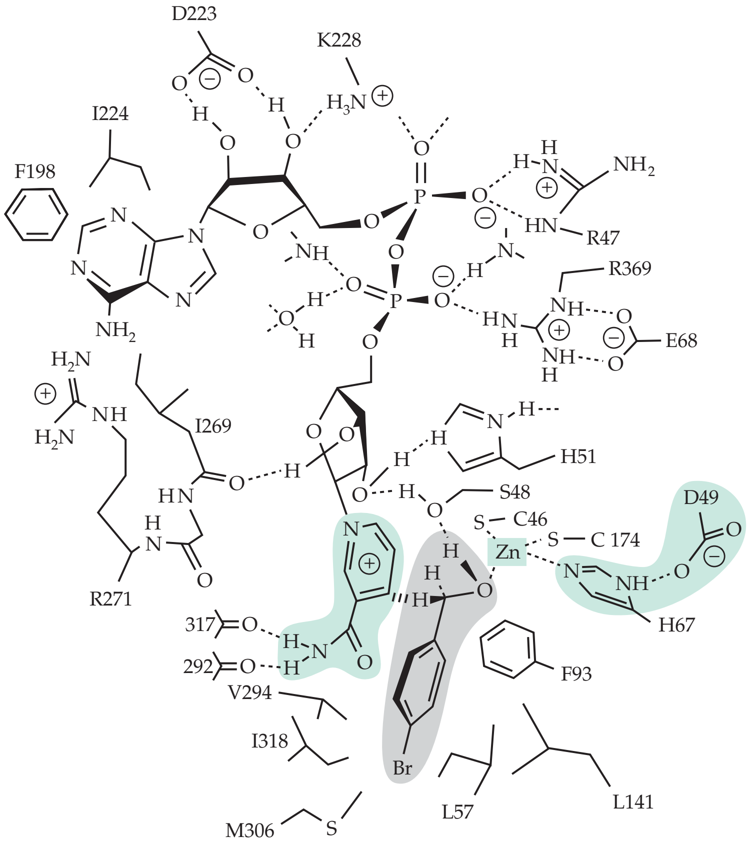
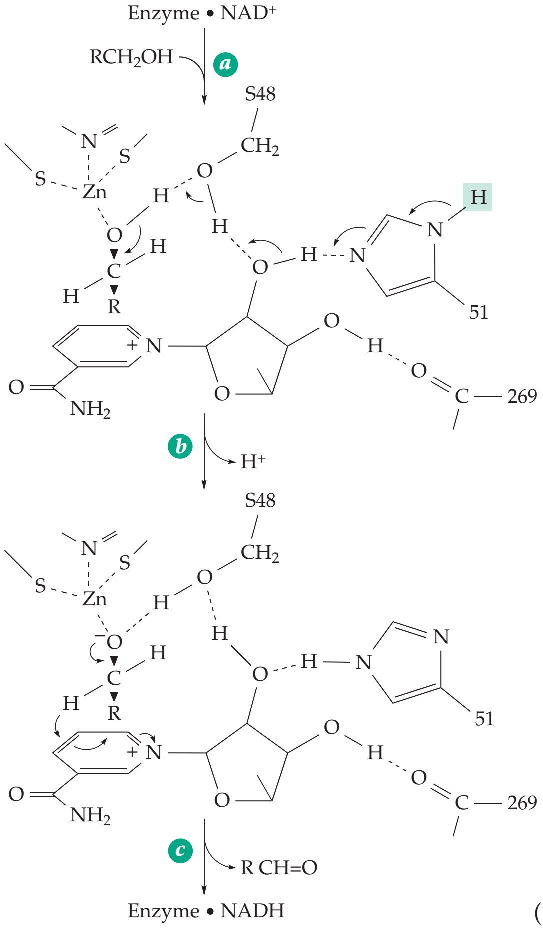


Figure 15-5 Structure of the complex of horse liver alcohol dehydrogenase with NAD$^+$ and the slow substrate *p*-bromobenzyl alcohol. The zinc atom and the nicotinamide ring of the bound NAD$^+$ are shaded. Adjacent to them is the bound substrate. Courtesy of Bryce Plapp. Based on Ramaswamy *et al.*[53]

Eklund *et al.* suggested that the side chains of Ser 48 and His 51 act as a proton relay system to remove the proton from the alcohol, in step *b* of Eq. 15-7, leaving the transient zinc-bound alcoholate ion, which can then transfer a hydride ion to NAD$^+$, in step *c*.[52] The shaded hydrogen atom leaves as H$^+$. The role of His 51 as a base is supported by studies of the inactivation of the horse liver enzyme by diethyl pyrocarbonate[57] and by directed mutation of yeast and liver enzymes. When His 51 was substituted by Gln the pK_a of 7 was abolished and the activity was decreased ten-fold.[58]

The functioning of zinc ions in enzymes has been controversial and other mechanisms have been proposed. Makinen *et al.* suggested a transient pentacoordinate Zn^{2+} complex on the basis of EPR measurements on en-



(15-7)

zyme containing Co^{2+}. Such a complex, in which the side chain of nearby Glu 68 would participate, would allow the coordinated water molecule to act as the base in deprotonation of the alcohol.[59] Molecular dynamics calculations indicate that Glu 68 can coordinate the zinc ion in this fashion, but Ryde suggests that its function may be to assist the exchange of ligands, i.e., the release of an alcohol or aldehyde product.[60] A variety of kinetic and spectroscopic studies have provided additional information that makes alcohol dehydrogenase one of the most investigated of all enzymes.

Liver alcohol dehydrogenase is important to the metabolism of ethanol by drinkers. Human beings exhibit small individual differences in their rates of alcohol metabolism which may reflect the fact that there are several isoenzymes and a number of genetic variants whose distribution differs from one person to the next as well as among tissues.[61–64] Inhibition of these isoenzymes by *uncompetitive* inhibitors, discussed in Chapter 9, is a goal in treatment of poisoning by methanol or ethylene glycol. Inhibition of the dehydrogenases slows the two-step oxidation of these substrates to toxic carboxylic acids.[65]

Yeast contains two cytosolic alcohol dehydrogenase isoenzymes.[66] Alcohol dehydrogenase I, present in large amounts in cells undergoing fermentation, functions to reduce acetaldehyde in the fermentation process. Alcohol dehydrogenase II is synthesized by cells growing on such carbon sources as ethanol itself and needing to oxidize it to obtain energy. A third isoenzyme is present in mitochondria. An alcohol dehydrogenase isoenzyme of green plants is induced by anaerobic stress such as flooding. It permits ethanolic fermentation to provide energy temporarily to submerged roots and other tissues.[67] Some bacteria contain an $NADP^+$-dependent, Zn^{2+}-containing alcohol dehydrogenase.[67a]

Other alcohol dehydrogenases and aldo-keto reductases. The oxidation of an alcohol to a carbonyl compound and the reverse reaction of reduction of a carbonyl group are found in so many metabolic pathways that numerous specialized dehydrogenases exist. A large group of "short-chain" dehydrogenases and reductases had at least 57 known members by 1995.[68–69a] Their structures and functions are variable. Most appear to be single-domain proteins with a large β sheet, a nucleotide-binding pocket, and a conserved pair of residues: Tyr 152 and Lys 156. These may function in a manner similar to that of the His-Asp pair in lactate dehydrogenase. A possible role of a cysteine side chain has been suggested for another member of this group, 3-hydroxyisobutyrate dehydrogenase, an enzyme of valine catabolism.[70]

Dehydrogenases often act primarily to reduce a carbonyl compound rather than to dehydrogenate an alcohol. These enzymes may still be called dehydrogenases. For example, in the lactic acid fermentation lactate is formed by reduction of pyruvate but we still call the enzyme lactate dehydrogenase. In our bodies this enzyme functions in both directions. However, some enzymes that act mainly in the direction of reduction are called reductases. An example is **aldose reductase**, a member of a family of **aldo-keto reductases**[71–73] which have $(\alpha / \beta)_8$-barrel structures.[74–76]

The normal physiological function of aldose reductase is uncertain but it can cause a problem in diabetic persons by reducing glucose to sorbitol (glucitol), ribose to ribitol, etc. The resulting sugar alcohols accumulate in the lens and are thought to promote cataract formation and may also be involved in the severe damage to retinas and kidneys that occurs in diabetes mellitus. Inhibitors of aldose reductase delay development of these complications in animals but the compounds tested are too toxic for human use.[76]

The active sites of aldo-keto reductases contain an essential tyrosine (Tyr 48) with a low pK_a value. The nearby His 110, Asp 43, and Lys 77 may also participate in catalysis.[76,77] The kinetics are unusual. Both NAD^+ and NADH are bound tightly and the overall rate of reduction of a substrate is limited by the rate of dissociation of NAD^+.[78] Citrate is a natural uncompetitive inhibitor of aldose reductase.[79]

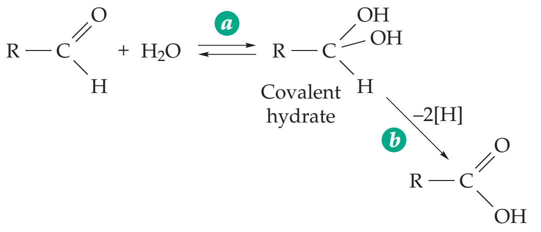
3. Dehydrogenation of Amino Acids and Amines

The dehydrogenation of an amine or the reverse reaction, the reduction of a Schiff base (Table 15-1, reaction type B), is another important pyridine nucleotide-dependent process. **Glutamate dehydrogenase,** a large oligomeric protein whose subunits contain 450 or more residues, is the best known enzyme catalyzing this reaction. An intermediate Schiff base of 2-oxoglutarate and NH_3 is a presumed intermediate.[80,81] Similar reactions are catalyzed by dehydrogenases for alanine,[82,83] leucine,[84] phenylalanine,[85] and other amino acids.[83,86,87]

4. Glyceraldehyde-3-Phosphate Dehydrogenase and the Generation of ATP in Fermentation Reactions

The oxidation of an aldehyde to a carboxylic acid, a highly exergonic process, often proceeds through a thioester intermediate whose cleavage can then be coupled to synthesis of ATP. This sequence is of central importance to the energy metabolism of cells (Chapters 10 and 17) and is shown in Fig. 15-6.

The best known enzyme catalyzing the first step of this reaction sequence is glyceraldehyde 3-phosphate dehydrogenase which functions in the glycolytic sequence (steps *a* and *b* of Fig. 10-3). It is present in both prokaryotes and eukaryotes as a tetramer of identical 36- to 37-kDa subunits. Three-dimensional structures have been determined for enzyme from several species, including lobster,[47,88] *E. coli,*[89] the thermophilic bacterium *Bacillus stearothermophilus* (Fig. 2-13),[90–90a] and trypanosomes.[91] Recall that aldehydes are in equilibrium with their covalent hydrates (Eq. 15-8, step *a*). Dehydrogenation of such a hydrate yields an acid (Eq. 15-8, step *b*) but such a mechanism offers no possibility of conserving the energy available from the reaction. However, during catalysis by glyceraldehyde-phosphate dehydrogenase the SH group of Cys 149, in the first step (step *a*, Fig. 15-6), adds to the substrate carbonyl group to form an adduct, a thiohemiacetal. This adduct is oxidized by NAD^+ to a thioester, an *S*-acyl enzyme (step *b*), which is then cleaved by the same enzyme through a displacement on carbon by an oxygen atom of P_i (phosphorolysis; step *c*). The sulfhydryl group of the enzyme is released simultaneously and the product, the acyl phosphate **1,3-bisphosphoglycerate**, is formed. The imidazole group of His 176 may catalyze both steps *a* and *b*.[89,92] A separate enzyme then transfers the phospho group from the 1 position of 1,3-bisphosphoglycerate to ADP to form ATP and 3-phosphoglycerate (step *d*). The overall sequence of Fig. 15-6 is the synthesis of one mole of ATP coupled to the oxidation of an aldehyde to a carboxylic acid and the conversion of NAD^+ to NADH.



(15-8)

In green plants and in some bacteria an $NADP^+$-dependent cytoplasmic glyceraldehyde 3-phosphate dehydrogenase *does not* use inorganic phosphate to form an acyl phosphate intermediate but gives 3-phosphoglycerate with a free carboxylate as in Eq. 15-8.[93,93a] Because it doesn't couple ATP cleavage to the dehydrogenation, it drives the $[NADPH] / [NADP^+]$ ratio to a high value favorable to biosynthetic processes (see Chapter 17).

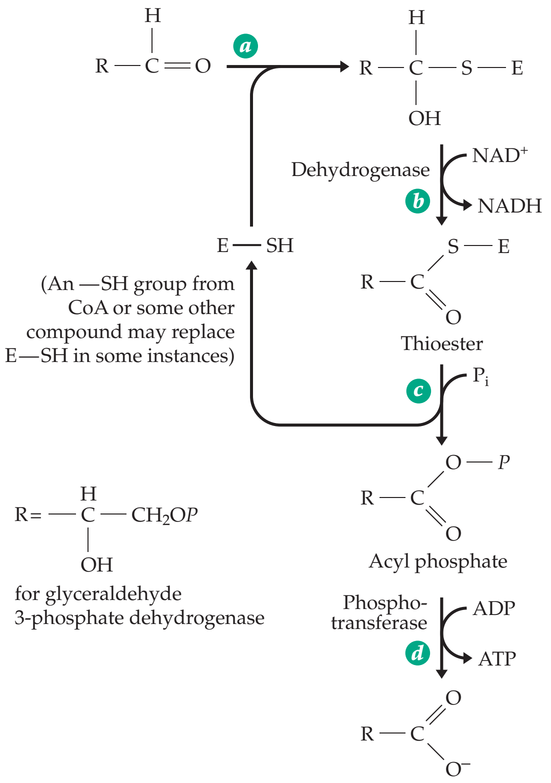

a
Dehydrogenase
NAD⁺
b
NADH
E — SH
(An —SH group from CoA or some other compound may replace E—SH in some instances)
Thioester
c
Pᵢ
Acyl phosphate
R = —CH(OH)—CH₂OP
for glyceraldehyde 3-phosphate dehydrogenase
Phosphotransferase
ADP
d
ATP

Figure 15-6 Generation of ATP coupled to oxidation of an aldehyde to a carboxylic acid. The most important known example of this sequence is the oxidation of glyceraldehyde 3-phosphate to 3-phosphoglycerate (Fig. 10-3, steps *a* and *b*). Other important sequences for "substrate-level" phosphorylation are shown in Eq. 14-23 and in Fig. 15-16.

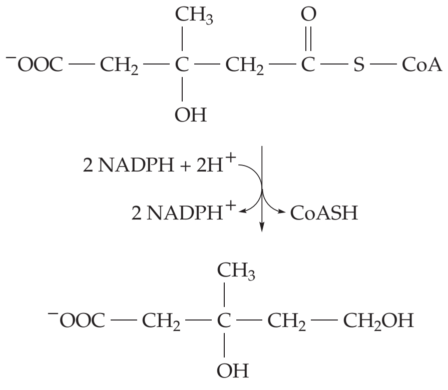
Animal tissues also contain aldehyde dehydrogenases of a nonspecific type which are thought to act to remove toxic aldehydes from tissues.[94,95] Like glyceraldehyde 3-phosphate, these enzymes form thioester intermediates as in Fig. 15-6 but which are hydrolyzed rather than being converted to acyl phosphates. A mutation (E487K) in the mitochondrial enzyme occurs in about 50% of the Asian population. Although the structural alteration is not at the active site, the enzyme activity is low. Individuals carrying the mutation are healthy but have an aversion to alcohol, whose consumption causes an elevated blood level of acetaldehyde, facial flushing, dizziness, and other symptoms. A similar effect is exerted by the drug disulfiram (Antabuse), which has been used to discourage drinking and whose metabolites are thought to inhibit aldehyde dehydrogenase.[95]

Alcohol dehydrogenases also oxidize aldehydes, probably most often as the geminal diol forms, according to Eq. 15-8. No ATP is formed. The same enzymes can catalyze the dismutation of aldehydes, with equal numbers of aldehyde molecules going to carboxylic acid and to the alcohol.[96–98]

5. Reduction of Carboxyl Groups

The last two reaction steps in Fig. 15-6, steps *c* and *d*, are in essence the reverse of the sequence used for synthesis of a thioester such as a fatty acyl coenzyme A. Thus, the chemistry by which ATP is generated during glycolysis and that by which it is utilized in biosynthesis is nearly the same. Furthermore, a standard biochemical method for reduction of carboxyl groups to aldehyde groups is conversion, in an ATP-requiring process, to a thioester followed by reduction of the thioester (Table 15-1, reaction type C). For example, the sequence of Fig. 15-6 is reversed during gluconeogenesis (see Fig. 17-17). The carboxyl group of the side chain of aspartate can be reduced in two steps to form the alcohol homoserine (Figs. 11-3; 24-13).

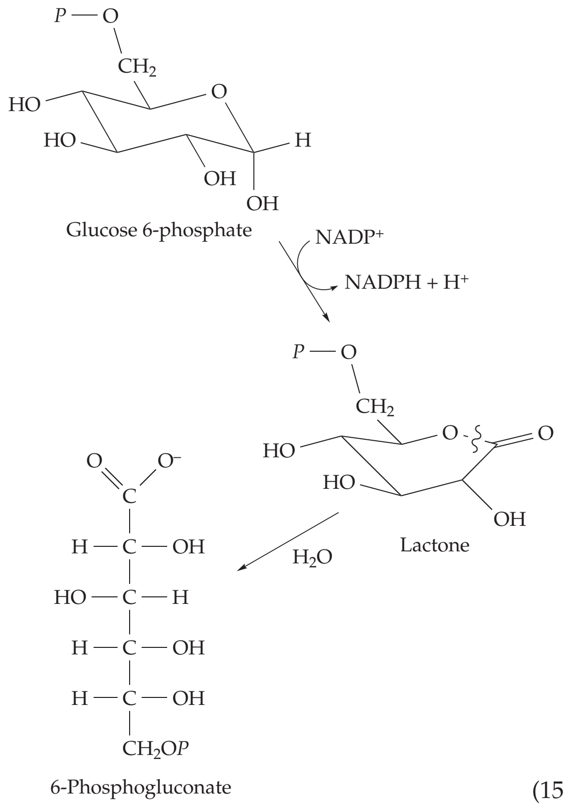
The aldehyde generated by reduction of a thioester is not always released from the enzyme but may be converted on to the alcohol in a second reduction step.[99] This is the case for **3-hydroxy-3-methylglutaryl-CoA reductase** (HMG-CoA reductase), a large 887-residue protein that synthesizes mevalonate (Eq. 15-9).[100–102] This highly regulated enzyme controls the rate of synthesis of cholesterol and is a major target of drugs designed to block cholesterol synthesis. The structure of a smaller 428-residue bacterial enzyme is known.[103,104] Aspartate 766 is a probable proton donor in both reduction steps and Glu 558 and His 865 may act as a catalytic pair that protonates the sulfur of coenzyme A.[100]



$$^{-}OOC-CH_2-\overset{\overset{CH_3}{|}}{\underset{\underset{OH}{|}}{C}}-CH_2-\overset{\overset{O}{\|}}{C}-S-CoA \xrightarrow[2\,NADPH^{+} \quad CoASH]{2\,NADPH + 2H^{+}} {}^{-}OOC-CH_2-\overset{\overset{CH_3}{|}}{\underset{\underset{OH}{|}}{C}}-CH_2-CH_2OH$$

(15-9)

A related oxidation reaction is catalyzed by **glucose-6-phosphate dehydrogenase**, the enzyme that originally attracted Warburg's attention and led to the discovery of $NADP^+$. The substrate, the hemiacetal ring form of glucose, is oxidized to a lactone which is then hydrolyzed to 6-phosphogluconate (Eq. 15-10).[104a] This oxidation of an aldehyde to a carboxylic acid is not linked directly to ATP synthesis as in Fig. 15-6. The ring-opening step ensures that the reaction goes to completion. This reaction is a major supplier of NADPH for reductive biosynthesis and the large Gibbs energy decrease for the overall reaction ensures that the ratio [NADPH] / [$NADP^+$] is kept high within cells. This is the only source of NADPH for mature erythrocytes and a deficiency of glucose 6-phosphate dehydrogenase is a common cause of drug- and food-

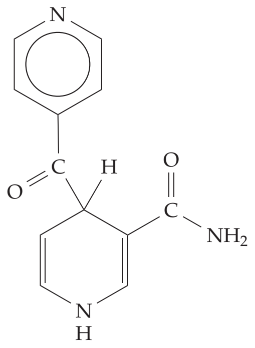


(15-10)

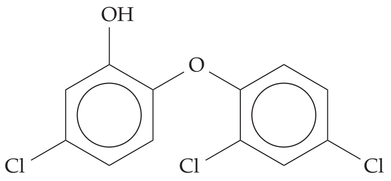
induced hemolytic anemia in human beings. About 400 variant forms of this enzyme are known.[105] Like the sickle cell trait (Box 7-B) some mutant forms of glucose-6-phosphate dehydrogenase appear to confer resistance to malaria.

6. Reduction of Carbon–Carbon Double Bonds

Neither NADP$^+$ nor NAD$^+$ is a strong enough oxidant to carry out the dehydrogenation of an acyl-CoA (reaction type D of Table 15-1). However, NADPH or NADH can participate in the opposite reaction. Thus, NADPH transfers a hydride ion to the β-carbon of an unsaturated acyl group during the biosynthesis of fatty acids (Chapter 17) and during elongation of shorter fatty acids (Chapter 21).[106–107b] A discovery of medical importance is that isonicotinyl hydrazide (INH), the most widely used antituberculosis drug, forms an adduct (of an INH anion or radical) with NAD$^+$ of long-chain **enoyl-acyl carrier protein reductase** (enoyl-ACP reductase).[133a,b] This enzyme utilizes NADH in reduction of a C=C double bond during synthesis of mycolic acids. The same enzyme is blocked by **triclosan**, an antibacterial compound used widely in household products such as antiseptic soaps, toothpastes, cosmetics, fabrics, and toys.[133c–e]

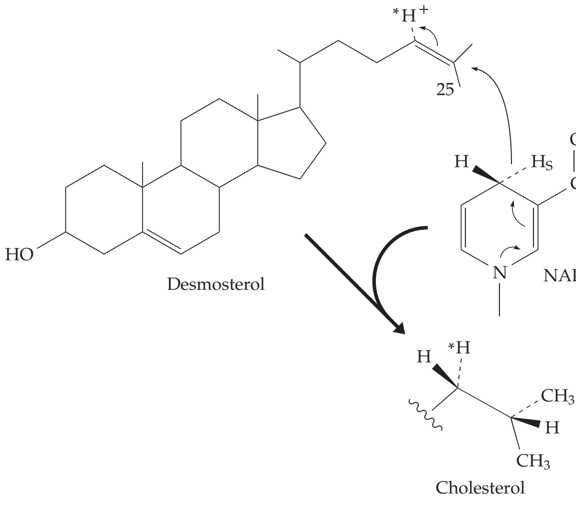


Adduct of isonicotinyl hydrazide anion with NAD$^+$



Triclosan

Less frequently NADPH is used to reduce an *isolated double bond*. An example is the hydrogenation of **desmosterol** by NADPH (Eq. 15-11), the final step in one of the pathways of biosynthesis of cholesterol (Fig. 22-7). In this and in other reactions of the same type, hydrogen transfer has been shown to be from the *pro-S* position in NADPH directly to C-25 of the sterol. The proton introduced from the medium (designated by the asterisks in Eq. 15-11) enters *trans* to the H$^-$ ion from NADPH. The proton always adds to the more electron-rich terminus of the double bond, i.e., it follows the Markovnikov rule. This result suggests that protonation of the double bond may precede H$^-$ transfer.[108]

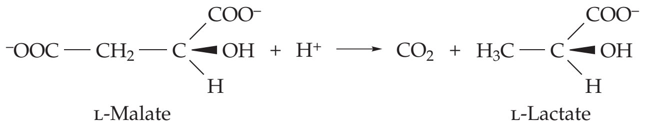


(15-11)

Additional pyridine nucleotide-dependent dehydrogenases include **glutathione reductase** (Figs. 15-10, 15-12), **dihydrofolate reductase** (Fig. 15-19), isocitrate dehydrogenase, ***sn*-glycerol-3-phosphate dehydrogenase** (Chapter 21), L-3-hydroxyacyl-CoA dehydrogenase (Chapter 21), **retinol dehydrogenase** (Chapter 23), and a bacterial quinone oxidoreductase.[109] Some of these also contain a flavin coenzyme.

7. Transient Carbonyl Groups in Catalysis

Some enzymes contain bound NAD$^+$ which oxidizes a substrate alcohol to facilitate a reaction step and is then regenerated. For example, the **malolactic enzyme** found in some lactic acid bacteria and also in *Ascaris* decarboxylates L-malate to lactate (Eq. 15-12). This reaction is similar to those of isocitrate dehydrogenase,[110–112] 6-phosphogluconate dehydrogenase,[113] and the malic enzyme (Eq. 13-45)[114] which utilize free NAD$^+$ to first dehydrogenate the substrate to a bound oxoacid whose β carbonyl group facilitates decarboxylation. Likewise, the bound NAD$^+$ of the malolactic

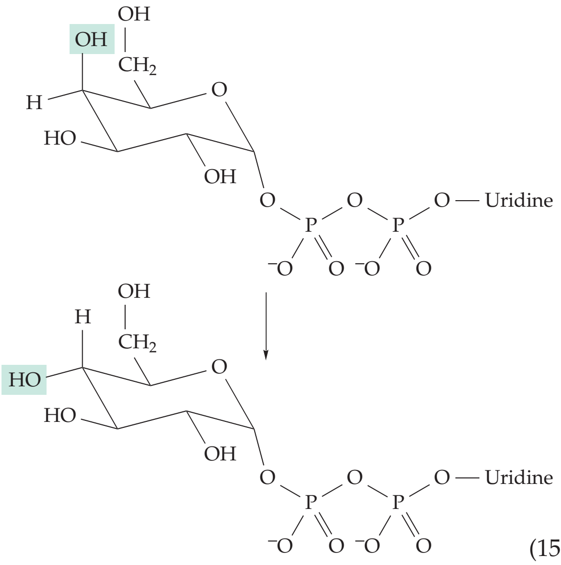


$$ {}^{-}OOC-CH_2-\overset{COO^-}{\underset{H}{C}}\blacktriangleleft OH + H^+ \longrightarrow CO_2 + H_3C-\overset{COO^-}{\underset{H}{C}}\blacktriangleleft OH $$

(15-12)

enzyme apparently dehydrogenates the malate to a bound oxaloacetate which is decarboxylated to pyruvate.[115] The latter remains in the active site and is reduced by the bound NADH to lactate which is released from the enzyme.

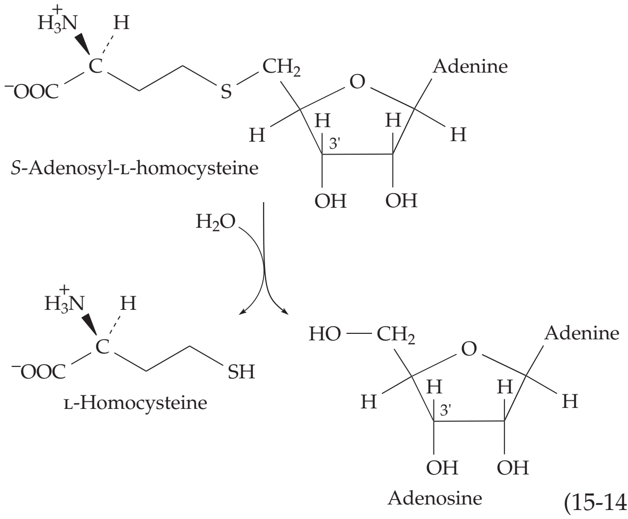
Another reason for introducing a carbonyl group is to form a symmetric intermediate in a reaction that inverts the configuration about a chiral center. An example is **UDP-galactose 4-epimerase**, an enzyme that converts UDP-galactose to UDP-glucose (Eq. 15-13; Chapter 20) and is essential in the metabolism of galactose in our bodies. The enzyme contains bound NAD^+ and forms a transient 4-oxo intermediate and bound NADH. Rotation of the intermediate allows nonstereospecific reduction by the NADH, leading to epimerization.[116–117a]



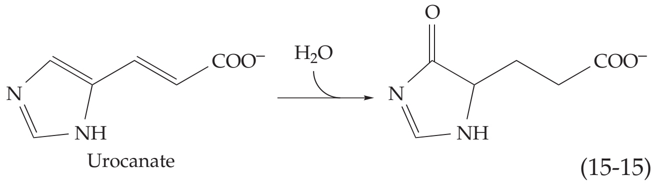
(15-13)

The enzyme is a member of the short-chain dehydrogenase group with a catalytic Tyr–Lys pair in the active site.[118] Another way that formation of an oxo group can assist in epimerization of a sugar is through enolization with nonstereospecific return of a proton to the intermediate enediol. A third possible mechanism of epimerization is through aldol cleavage followed by aldol condensation, with inversion of configuration. In each case the initial creation of an oxo group by dehydrogenation is essential.

Bound NAD^+ is also present in ***S*-adenosylhomocysteine hydrolase**,[119,120] which catalyzes the irreversible reaction of Eq. 15-14. Transient oxidation at the 3 position of the ribose ring facilitates the reaction. The reader can doubtless deduce the function that has been established for the bound NAD^+ in this enzyme. However, the role of NAD in the **urocanase** reaction (Eq. 15-15) is puzzling. This reaction, which is the second step in the catabolism of histidine, following Eq. 14-44, appears simple. However, there is no obvious mechanism and no obvious role for NAD^+. See Frey for a discussion.[117]



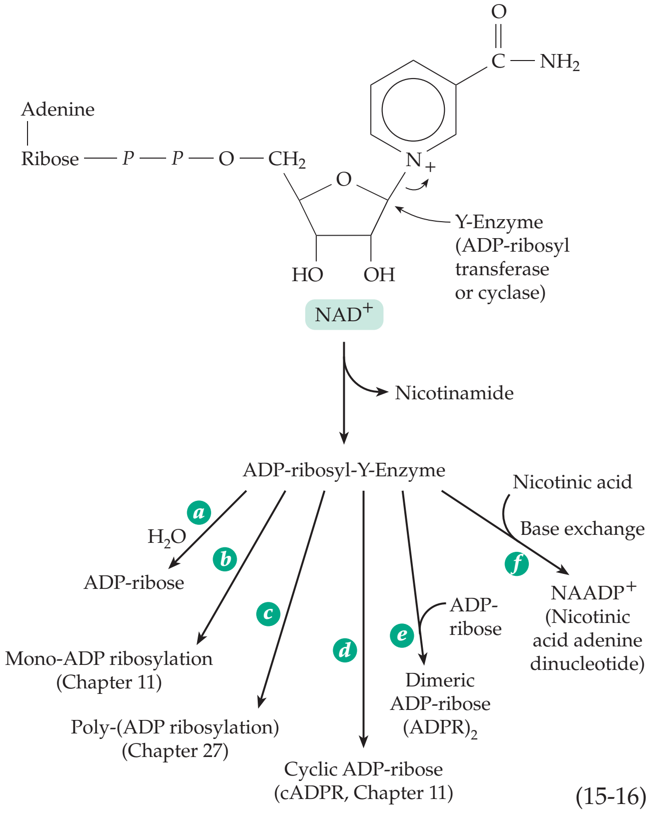
(15-14)



(15-15)

8. ADP Ribosylation and Related Reactions of NAD^+ and $NADP^+$

The linkage of nicotinamide to ribose in NAD^+ and $NADP^+$ is easily broken by nucleophilic attack on C-1 of ribose. In Chapter 11 enzyme-catalyzed ADP ribosylation, which can be shown as in Eq. 15-16, is discussed briefly. The nucleophilic group –Y from an ADP-ribosyltransferase carries the ADP-ribosyl group which can then be transferred by a second displacement onto a suitable nucleophilic acceptor group.[121,122] Hydrolysis (Eq. 15-16, step *a*) gives free ADP-ribose.[123] Other known products of enzymatic action are indicated in steps *c*–*f*. Poly-(ADP ribosylation) is discussed in Chapter 27. The structure of cyclic ADP-ribose (cADPR)[124] is shown in Chapter 11, Section E,2. The acceptor nucleophile is N-1 of the adenine ring which is made more nucleophilic by electron donation from the amino group. A similar reaction with ADP ribose (Eq. 15-16, step *e*) produces a dimeric ADP-ribose $(ADPR)_2$[125] while reaction with free nicotinic acid (step *f*) yields, in an overall base exchange, nicotinic acid adenine dinucleotide ($NAADP^+$).[126,127] Some of these compounds, e.g., cADPR, $NAADP^+$, and $(ADPR)_2$, are involved in signaling with calcium ions.[124]

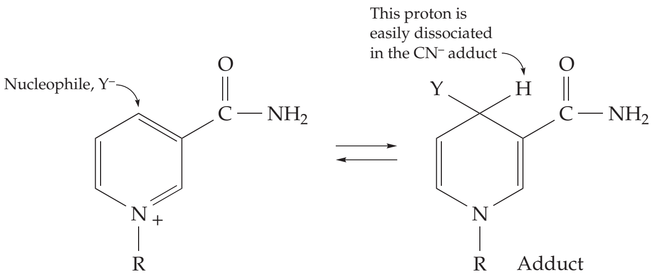

Adenine
Ribose — P — P — O — CH_2
O
C — NH_2
N+
HO
OH
Y-Enzyme
(ADP-ribosyl
transferase
or cyclase)
NAD+
Nicotinamide
ADP-ribosyl-Y-Enzyme
a
H_2O
ADP-ribose
b
Mono-ADP ribosylation
(Chapter 11)
c
Poly-(ADP ribosylation)
(Chapter 27)
d
Cyclic ADP-ribose
(cADPR, Chapter 11)
e
ADP-ribose
Dimeric
ADP-ribose
$(ADPR)_2$
Nicotinic acid
Base exchange
f
NAADP+
(Nicotinic
acid adenine
dinucleotide)

(15-16)

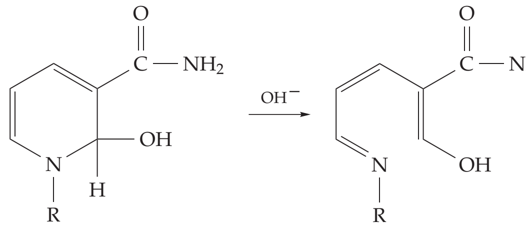
9. The Varied Chemistry of the Pyridine Nucleotides

Despite the apparent simplicity of their structures, the chemistry of the nicotinamide ring in NAD^+ and $NADP^+$ is surprisingly complex.[128,129] NAD^+ is extremely unstable in basic solutions, whereas NADH is just as unstable in slightly acidic media. These properties, together with the ability of NAD^+ to undergo condensation reactions with other compounds, have sometimes caused serious errors in interpretation of experiments and may be of significance to biological function.

Addition to NAD⁺ and NADP⁺. Many nucleophilic reagents add reversibly at the para (or 4) position (Eq. 15-17) to form adducts having structures resembling those of the reduced coenzymes. Formation of the cyanide adduct, whose absorption maximum is at 327 nm, has been used to introduce deuterium into the para position of the pyridine nucleotides. In the adduct, the hydrogen adjacent to the highly polarized C – N bond is easily dissociated as a proton. Thiolate anions and bisulfite also add. Dithionite ion, $S_2O_4^{2-}$, can lose SO_2 and acquire a proton to form the sulfoxylate ion HSO_2^- which also adds to the 4 position of the NAD^+ ring.[130] The resulting adduct is unstable and loses SO_2 to give NADH + H^+. Addition can also occur at the two ortho positions. The adducts of HO^- to the 4 position of NAD^+ are stable but those to the 2 position undergo ring opening (Eq. 15-18) in base-catalyzed reactions which are followed by further degradation.[128,131]

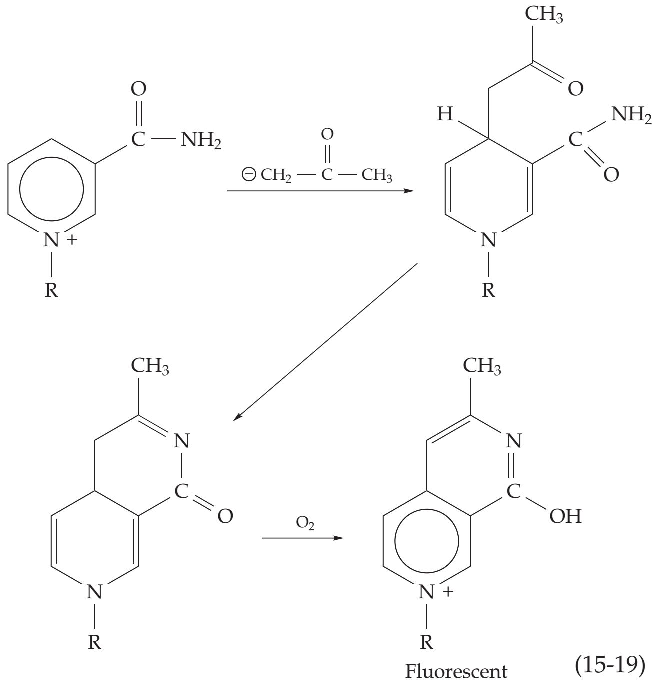

Nucleophile, Y^-
O
C — NH_2
N+
R
This proton is
easily dissociated
in the CN^- adduct
Y
H
O
C — NH_2
N
R
Adduct

(15-17)


O
C — NH_2
OH
N
H
R
OH^-
O
C — NH_2
N
OH
R

(15-18)

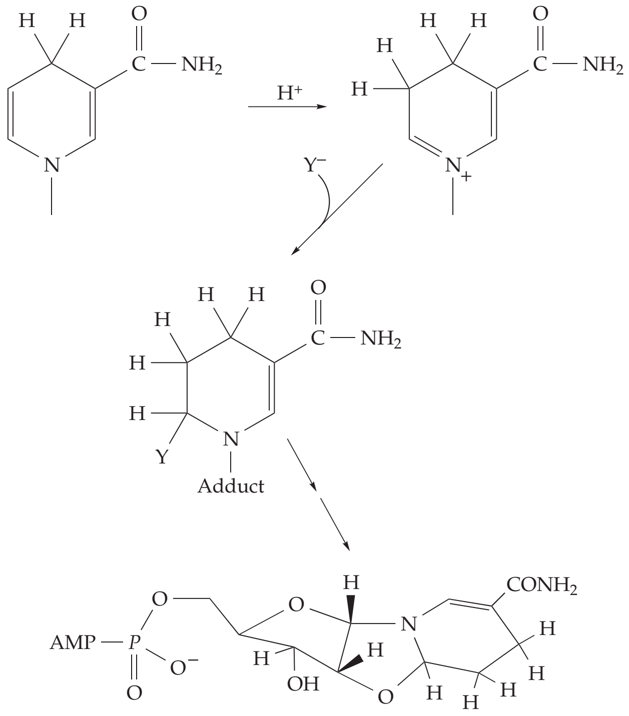
Another base-catalyzed reaction is the addition of enolate anions derived from ketones to the 4 position of the pyridine nucleotides (Eq. 15-19). The adducts undergo ring closure and in the presence of oxygen are converted slowly to fluorescent materials. While forming the basis for a useful analytical method for determination of NAD^+ (using 2-butanone), these reactions also have created a troublesome enzyme inhibitor from traces of acetone present in commercial NADH.[132]


O
C — NH_2
N+
R
O
⊖ CH_2 — C — CH_3
CH_3
O
H
NH_2
C
O
N
R
CH_3
N
C
O
N
R
O_2
CH_3
N
C
OH
N+
R
Fluorescent

(15-19)

The reactions of Eq. 15-19 occur nonenzymatically only under the influence of strong base but dehydrogenases often catalyze similar condensations relatively rapidly and reversibly. Pyruvate inhibits lactate dehydrogenase, 2-oxoglutarate inhibits glutamate dehydrogenase, and ketones inhibit a short-chain alcohol dehydrogenase in this manner.[133,69a]

Modification of NADH in acid. Reduced pyridine nucleotides are destroyed rapidly in dilute HCl and more slowly at pH 7 in reactions catalyzed by buffer acids.[128,131,134] Apparently the reduced nicotinamide ring is first protonated at C-5, after which a nucleophile Y^- adds at the 6 position (Eq. 15-20). The nucleophile may be OH^-, and the adduct may undergo further reactions. For example, water may add to the other double bond and the ring may open on either side of the nitrogen. The glycosidic linkage can be isomerized from β to α or can be hydrolytically cleaved. The early steps in the modification reaction are partially reversible, but the overall sequence is irreversible. One of the products, which has been characterized by crystal structure determination, is shown in Eq. 15-20.[128,135] It can arise if the group Y of Eq. 15-20 is the C-2' hydroxyl of the ribose ring and if the configuration of the glycosidic linkage is inverted (anomerized).



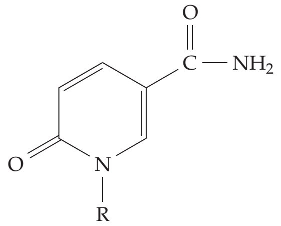
An acid modification product from NADH

(15-20)

The foregoing reactions have attracted interest because glyceraldehyde-3-phosphate dehydrogenase, in a side reaction, converts NADH to a substance referred to as NADH-X which has been shown to be the 6(*R*) adduct of Eq. 15-20, where Y is −OH. In an ATP-dependent reaction an enzyme from yeast reconverts NADH-X to NADH.[136]

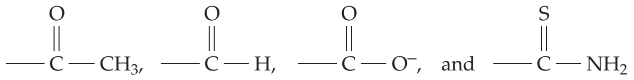
Mercury (II) ions can add in place of H^+ in the first step of Eq. 15-20 and subsequent reactions similar to those promoted by acid can occur.[137]

Other reactions of pyridine nucleotides. Alkaline hexacyanoferrate (III) oxidizes NAD^+ and $NADP^+$ to 2-,4-, and 6-pyridones. The 6-pyridone of *N*-methylnicotinamide is a well-known excretion product of nicotinic acid in mammals. Reoxidation of NADH and NADPH to NAD^+ and $NADP^+$ can be accomplished with hexacyanoferrate (III), quinones, and riboflavin but not by H_2O_2 or O_2. However, O_2 does react at neutral pH with uptake of a proton to form a peroxide derivative of NADH.[138] When heated in 0.1 N alkali at 100°C for 5 min, NAD^+ is hydrolyzed to nicotinamide and adenosine-diphosphate-ribose.



6-Pyridone

Treatment of NAD^+ with nitrous acid deaminates the adenine ring. The resulting deamino NAD^+ as well as synthetic analogs containing the following groups in place of the carboxyamide have been used widely in enzyme studies. In fact, almost every part of the coenzyme molecule has been varied systematically and the effects on the chemical and enzymatic properties have been investigated.[139–141] "Caged" NAD^+ and $NADP^+$ have also been made.[142] These compounds do not react as substrates until they are released ("uncaged") by photolytic action of a laser beam (see Chapter 23).



$$-\overset{\overset{\displaystyle O}{\|}}{C}-CH_3,\quad -\overset{\overset{\displaystyle O}{\|}}{C}-H,\quad -\overset{\overset{\displaystyle O}{\|}}{C}-O^-,\quad \text{and}\quad -\overset{\overset{\displaystyle S}{\|}}{C}-NH_2$$

B. The Flavin Coenzymes

Flavin adenine diphosphate (**FAD**, flavin adenine dinucleotide) and **riboflavin 5'-monophosphate** (**FMN**, flavin mononucleotide), whose structures are shown in Fig. 15-7, are perhaps the most versatile of all

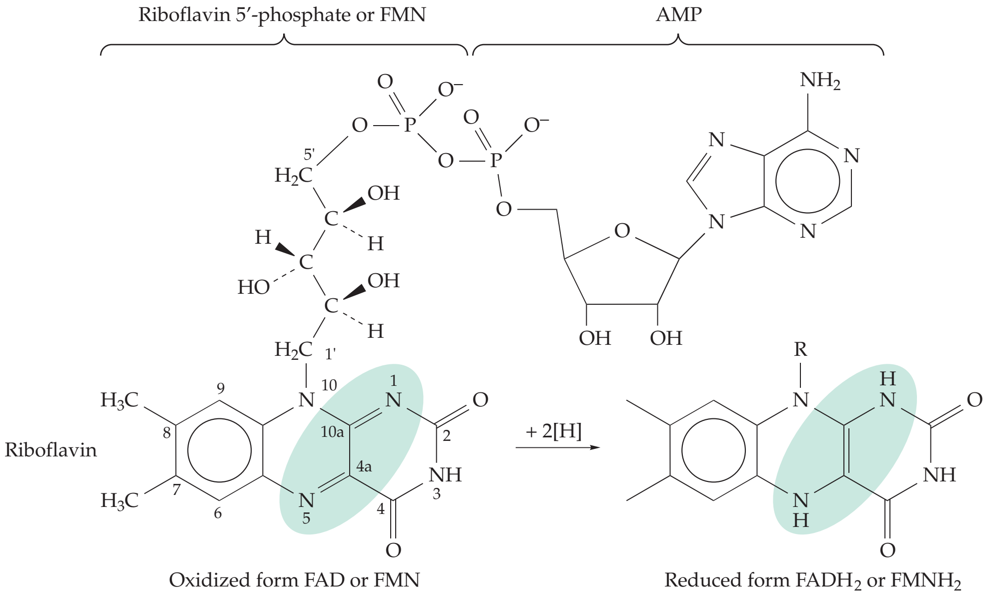

Riboflavin 5′-phosphate or FMN
AMP
Riboflavin
+ 2[H]
Oxidized form FAD or FMN
Reduced form FADH$_2$ or FMNH$_2$

Figure 15-7 The flavin coenzymes flavin adenine dinucleotide (FAD) and riboflavin 5'-phosphate (FMN). Dotted lines enclose the region that is altered upon reduction.

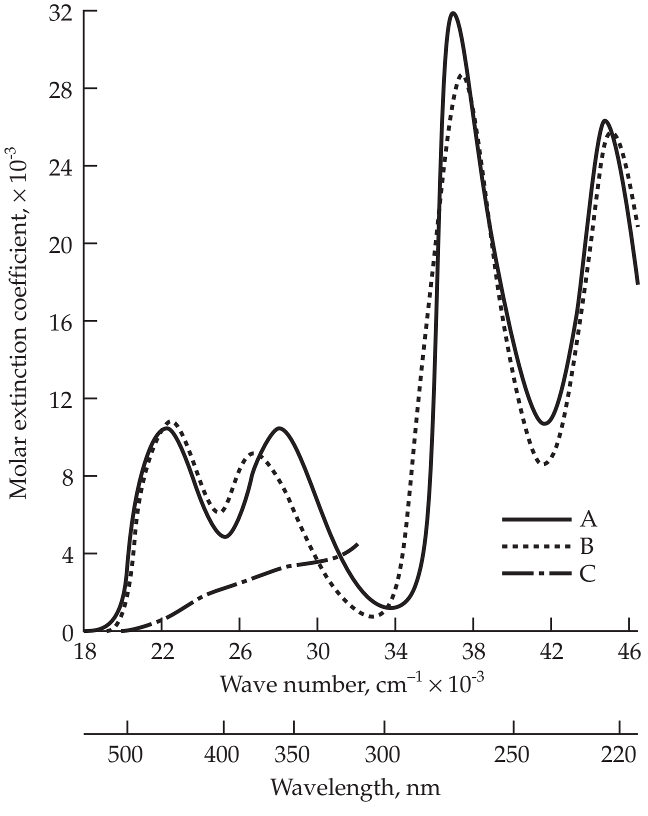


Figure 15-8 Absorption spectrum of neutral, uncharged riboflavin (A), the riboflavin anion (B), and reduced to the dihydro form (Fig. 15-7) by the action of light in the presence of EDTA (C). A solution of 1.1 x 10^{-4} M riboflavin containing 0.01 M EDTA was placed 11.5 cm from a 40-W incandescent lamp for 30 min.

the oxidation coenzymes. The name flavin adenine dinucleotide is not entirely appropriate because the D-ribityl group is not linked to the riboflavin in a glycosidic linkage. Hemmerich suggested that FAD be called flavin adenine diphosphate.[143]

The attention of biochemists was first attracted to flavins as a result of their color and fluorescence. The study of spectral properties of flavins (Fig. 15-8) has been of importance in understanding these coenzymes. The biochemical role of the flavin coenzymes was first recognized through studies of the "old yellow enzyme"[144,145] which was shown by Theorell to contain riboflavin 5'-phosphate. By 1938, FAD was recognized as the coenzyme of a different yellow protein, **D-amino acid oxidase** of kidney tissue. Like the pyridine nucleotides, the new flavin coenzymes were reduced by dithionite to nearly colorless dihydro forms (Figs. 15-7 and 15-8) revealing the chemical basis for their function as hydrogen carriers.

Flavins are also among the natural light receptors and display an interesting and much studied photochemistry (Fig. 3-5).[143] Flavins may function in some photoresponses of plants, and they serve as light emitters in bacterial bioluminescence (Chapter 23).[146,147]

Three facts account for the need of cells for both the flavin and pyridine nucleotide coenzymes: (1) Flavins are usually stronger oxidizing agents than is NAD$^+$. This property fits them for a role in the electron transport chains of mitochondria where a sequence of increasingly more powerful oxidants is needed and makes them ideal oxidants in a variety of other dehydrogenations. (2) Flavins can be reduced either by one- or two-electron processes. This enables them to participate in oxidation reactions involving free radicals and in reactions with metal ions. (3) Reduced flavins are "autooxidizable," i.e., they can be reoxidized directly and rapidly by O_2, a property shared with relatively few other organic substances. For example, NADH and NADPH are not spontaneously reoxidized by oxygen. Autooxidizability allows flavins of some enzymes to pass electrons directly to O_2 and also provides a basis for the functioning of flavins in hydroxylation reactions.

1. Flavoproteins and Their Reduction Potentials

Flavin coenzymes are usually bound tightly to proteins and cycle between reduced and oxidized states while attached to the same protein molecule. In a free unbound coenzyme the redox potential is determined by the structures of the oxidized and reduced forms of the couple. Both riboflavin and the pyridine nucleotides contain aromatic ring systems that are stabilized by resonance. Part of this resonance stabilization is lost upon reduction. The value of $E°'$ depends in part upon the varying amounts of resonance in the oxidized and reduced forms. The structures of the coenzymes have apparently evolved to provide values of $E°'$ appropriate for their biological functions.

The relative strengths of binding of oxidized and reduced flavin coenzymes to a protein also have strong effects upon the reduction potential of the coenzyme.[148] If the oxidized form is bound weakly, but the reduced form is bound tightly, a bound flavin will have a greater tendency to stay in the reduced form than it did when free. The reduction potential $E°'$ will be less negative than it is for the free flavin-dihydroflavin couple. On the other hand, if the oxidized form of the flavin is bound more tightly by the protein than is the reduced form, $E°'$ will be more negative and the flavoenzyme will be a less powerful oxidizing agent. In fact, the values of $E°'$ at pH 7 for flavoproteins span a remarkably wide range from -0.49 to $+0.19$ V. The state of protonation of the reduced flavin when bound to the enzyme will also have a major effect on the oxidation–reduction potential. For example, acyl-CoA dehydrogenases are thought to form an anionic species of reduced FAD ($FADH^-$) which is tightly bound to the protein.[149]

Every flavoprotein accepts electrons from the substrate that it oxidizes and passes these electrons on to another substrate, an oxidant. In the following sections we will consider for several enzymes how the electrons may get into the flavin from the oxidizable substrate and how they may flow out of the flavin into the final electron acceptor.

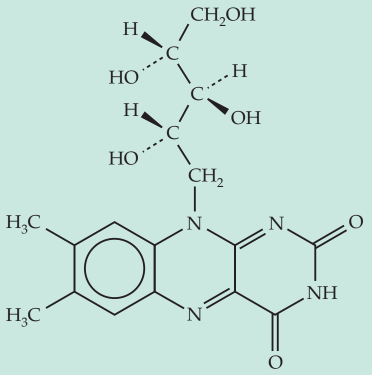
2. Typical Dehydrogenation Reactions Catalyzed by Flavoproteins

The functions of flavoprotein enzymes are numerous and diversified.[151–153a] A few of them are shown in Table 15-2 and are classified there as follows: (A) oxidation of hemiacetals to lactones, (B) oxidation of alcohols to aldehydes or ketones, (C) oxidation of amines to imines, (D) oxidation of carbonyl compounds or carboxylic acids to α,β-unsaturated compounds, (E) oxidation of NADH and NADPH in electron transport chains, and (F) oxidation of dithiols to disulfides or the reverse reaction. Three-dimensional structures are known for enzymes of each of these types.

Reactions of types A–C could equally well be catalyzed by pyridine nucleotide-requiring dehydrogenases. Recall that D-glucose-6-phosphate dehydrogenase uses $NADP^+$ as the oxidant (Eq. 15-10). The first product is the lactone which is hydrolyzed to 6-phosphogluconic acid. The similar reaction of free glucose (Table 15-2, reaction type A) is catalyzed by fungal **glucose oxidase**, a 580-residue FAD-containing enzyme.[154–155] A bacterial cholesterol oxidase has a similar structure.[156] The important plant enzyme **glycolate oxidase** is a dimer containing riboflavin 5'-phosphate.[157–159] It catalyzes a reaction of type B, which plays an important role in photorespiration (Chapter 23). **Amino acid oxidases** (reaction type C) are well-known. The peroxisomal D-amino acid oxidase from kidney was the source from which Warburg first isolated FAD and has been the subject of much investigation of mechanism and structure.[153,160–163] Many snake venoms contain an active 140-kDa L-amino-acid oxidase which contains FAD. Flavin-dependent **amine oxidases**, important in the human body, catalyze the related reaction with primary, secondary, or tertiary amines and in which a carboxyl group need not be present.[164–166a] Reduced flavin produced by all of these oxidases is reoxidized with molecular oxygen and hydrogen peroxide is the product. Nature has chosen to forego the use of an electron transport chain (Fig. 10-5), giving up the possible gain of ATP in favor of simplicity and a more direct reaction with oxygen. In some cases there is specific value to the organism in forming H_2O_2 (see Chapter 18).

In contrast to the flavin oxidases, flavin dehydrogenases pass electrons to carriers within electron transport chains and the flavin does *not* react with O_2. Examples include a bacterial **trimethylamine dehydrogenase** (Fig. 15-9) which contains an iron–sulfur cluster that serves as the immediate electron acceptor[167–169] and yeast **flavocytochrome b_2**, a lactate dehydrogenase that passes electrons to a built-in heme group which can then pass the electrons to an external acceptor, another heme in cytochrome *c*.[170–173] Like glycolate oxidase, these enzymes bind their flavin coenzyme at the ends of 8-stranded αβ barrels similar

BOX 15-B RIBOFLAVIN



The bright orange-yellow color and brilliant greenish fluorescence of riboflavin first attracted the attention of chemists. Blyth isolated the vitamin from whey in 1879 and others later obtained the same fluorescent, yellow compound from eggs, muscle, and urine. All of these substances, referred to as **flavins** because of their yellow color, were eventually recognized as identical. The structure of riboflavin was established in 1933 by R. Kuhn and associates, who had isolated 30 mg of the pure material from 30 kg of dried albumin from 10,000 eggs. The intense fluorescence assisted in the final stages of purification. The vitamin was synthesized in 1935 by P. Karrer.[a]

Riboflavin, a yellow solid, has a low solubility of ~100 mg / l at 25°C. Three crystalline forms are known. One of these, the "readily soluble form," is ten times more soluble than the others and can be used to prepare metastable solutions of higher concentration. One crystalline form is platelike and occurs naturally in the tapetum (Box 13-C) of the nocturnal lemur.

Discovery of the role of riboflavin in biological oxidation was an outgrowth of biochemists' interest in respiration. In the 1920s Warburg provided evidence that oxygen reacted with an iron-containing respiration catalyst and it was shown that the dye **methylene blue** could often substitute for oxygen as an oxidant (Box 18-A). Oxidation of glucose 6-phosphate by methylene blue within red blood cells required both a "ferment" (enzyme) and a "coferment," later identified as $NADP^+$. A yellow protein, isolated from yeast, was found to have the remarkable property of being decolorized by the reducing system of glucose 6-phosphate plus the protein and coferment from red blood cells.

Warburg and Christian showed that the color of this **old yellow enzyme** came from a flavin and proposed that its cyclic reduction and reoxidation played a role in cellular oxidation. When $NADP^+$ was isolated the proposal was extended to encompass a **respiratory chain**. The two hydrogen carriers $NADP^+$ and flavin would work in sequence to link dehydrogenation of glucose to the iron-containing catalyst that interacted with oxygen. While we still do not know the physiological function of the old yellow enzyme,[b] the concept of respiratory chain was correct.

Human beings require about 1.5 mg of riboflavin per day. Because of its wide distribution in food, a deficiency, which affects skin and eyes, is rarely seen. Riboflavin is produced commercially in large quantities by fungi such as *Eremothecium asbyii* which, apparently because of some metabolic anomaly, produce the vitamin in such copious amounts that it crystallizes in the culture medium.

When taken up by the body, riboflavin is converted into its coenzyme forms (Chapter 25) and any excess is quickly excreted in the urine. Urine also contains smaller amounts of metabolites. The ribityl group may be cut by the action of intestinal bacteria acting on riboflavin before it is absorbed. The resulting 10-hydroxyethyl flavin may sometimes be a major urinary product.[c,d] The related 10-formylmethyl flavin is also excreted,[c] as are small amounts of 7α- and 8α- hydroxyriboflavins, apparently formed in the body by hydroxylation. These may be degraded farther to the 7α- and 8α- carboxylic acids of lumichrome (riboflavin from which the ribityl side chain is totally missing).[e] A riboflavin glucoside has also been found in rat urine.[f]

The choroid layer of the eye (behind the retina) in many animals contains a high concentration of free riboflavin. Cats' eyes also contain a large amount of 7 α-hydroxyriboflavin (nekoflavin), as do their livers.[g] Nekoflavin is also present in human blood.[h] Hen egg white contains a 219-residue riboflavin-binding protein whose functions are thought to be storage of the vitamin and delivery to the developing embryo.[i–k] Most of the riboflavin in human blood is bound to proteins such as albumin and immunoglobulins. However, during pregnancy a riboflavin-binding protein similar to that of the chicken appears, apparently to carry riboflavin to the fetus.[j]

Riboflavin is stable to heat but is extremely sensitive to light, a fact of some nutritional significance. Do not leave bottled milk in the sunshine (see Fig. 15-8)! Many products of photolysis are formed (Fig. 3-5). Among them is lumichrome.

---

[a] Yagi, K. (1990) in *Flavins and Flavoproteins* (Curti, B., Ronchi, S., and Zanetti, G., eds), pp. 3–16, Walter de Gruyter, Berlin

[b] Kohli, R. M., and Massey, V. (1998) *J. Biol. Chem.* **273**, 32763–32770

[c] Owen, E. C., West, D. W., and Coates, M. E. (1970) *Br. J. Nutr.* **24**, 259–267

[d] Roughead, Z. K., and McCormick, D. B. (1991) *European Journal of Clinical Nutrition* **45**, 299–307

[e] Ohkawa, H., Ohishi, N., and Yagi, K. (1983) *J. Biol. Chem.* **258**, 5623–5628

[f] Ohkawa, H., Ohishi, N., and Yagi, K. (1983) *J. Nutr. Sci. Vitaminol.* **29**, 515–522

[g] Matsui, K., and Kasai, S. (1996) *J. Biochem.* **119**, 441–447

[h] Zempleni, J., Galloway, J. R., and McCormick, D. B. (1995) *Int. J. Vitamins Nutr. Res.* **66**, 151–157

[i] Matsui, K., Sugimoto, K., and Kasai, S. (1982) *J. Biochem.* **91**, 469–475

[j] Miura, R., Tojo, H., Fujii, S., Yamano, T., Miyake, Y., (1984) *J. Biochem.* **96**, 197–206

[k] Monaco, H. L. (1997) *EMBO J.* **16**, 1475–1483

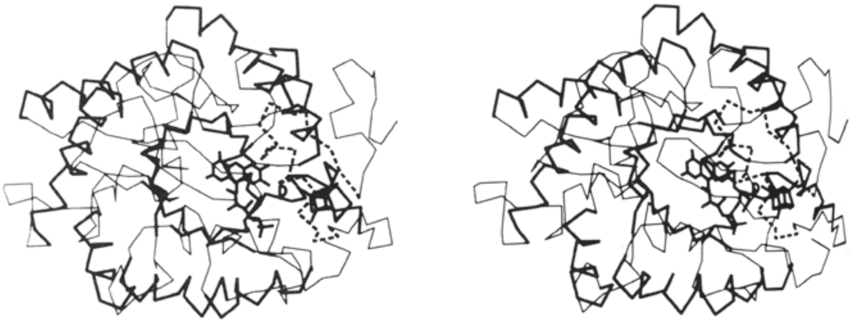


Figure 15-9 Stereoscopic view of the large domain (residues 1–383) of trimethylamine dehydrogenase from a methylotrophic bacterium. The helices and β strands of the $(\alpha\beta)_8$ barrel are drawn in heavy lines as are the FMN (center) and the Fe_4S_4 iron–sulfur cluster at the lower right edge. The α/β loop to which it is bound is drawn with dashed lines. The 733-residue protein also contains two other structural domains. From Lim *et al.*[150] Courtesy of F. S. Mathews.

to that of triose phosphate isomerase (Fig. 2-28). Flavocytochrome b_2 has an additional domain which carries the bound heme. Two additional domains of trimethylamine dehydrogenase have a topology resembling that of the FAD- and NADH-binding domains of glutathione reductase shown in Fig. 15-10. A bacterial **mandelate dehydrogenase** is structurally and mechanistically closely related to the glycolate oxidase family.[173a]

Reaction type D of Table 15-2, the dehydrogenation of an acyl-coenzyme A (CoA), could not be accomplished by a pyridine nucleotide system because the reduction potential ($E°'$, pH 7 = –0.32 V) is inappropriate. The more powerfully oxidizing flavin system is needed. (However, the reverse reaction, hydrogenation of a C=C bond, is often carried out biologically with a reduced pyridine nucleotide.) Dehydrogenation reactions of this type are important in the energy metabolism of aerobic cells. For example, the first oxidative step in the β oxidation of fatty acids (Fig. 17-1) is the α,β dehydrogenation of fatty acyl-CoA derivatives. The *pro-R* hydrogen atoms are removed from both the α- and β-carbon atoms to create the double bond (Table 15-2, type C).

TABLE 15-2
Some Dehydrogenation Reactions Catalyzed by Flavoproteins[a]

A

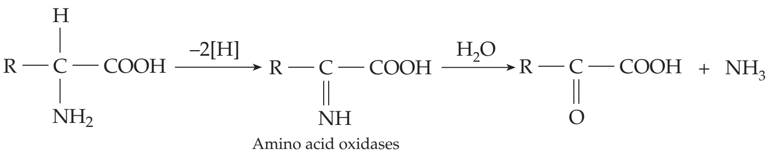
$$\text{D-Glucose} \xrightarrow[\text{Glucose oxidase}]{-2[\text{H}]} \text{Gluconolactone} \xrightarrow{\text{H}_2\text{O}} \text{D-Gluconic acid}$$

(See Eq. 17-12 for structures in a closely related reaction.)

B

$$\text{HOOC}-\text{CH}_2-\text{OH} \xrightarrow[\text{Glycolate oxidase}]{-2[\text{H}]} \text{HOOC}-\text{CH}=\text{O}$$

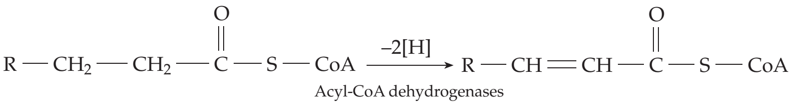
C



$$\text{R}-\underset{\text{NH}_2}{\overset{\text{H}}{\text{C}}}-\text{COOH} \xrightarrow{-2[\text{H}]} \text{R}-\underset{\text{NH}}{\underset{\|}{\text{C}}}-\text{COOH} \xrightarrow{\text{H}_2\text{O}} \text{R}-\underset{\text{O}}{\underset{\|}{\text{C}}}-\text{COOH} + \text{NH}_3$$

Amino acid oxidases

D

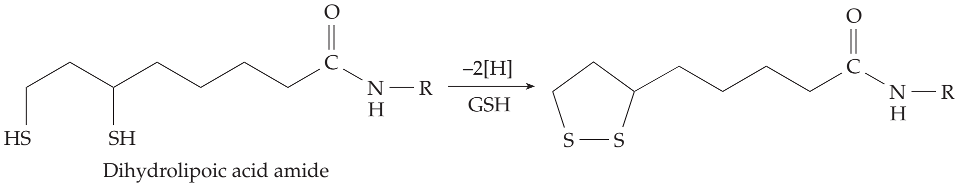


$$\text{R}-\text{CH}_2-\text{CH}_2-\overset{\text{O}}{\overset{\|}{\text{C}}}-\text{S}-\text{CoA} \xrightarrow[\text{Acyl-CoA dehydrogenases}]{-2[\text{H}]} \text{R}-\text{CH}=\text{CH}-\overset{\text{O}}{\overset{\|}{\text{C}}}-\text{S}-\text{CoA}$$

E

$$\text{NADH (or NADPH)} + \text{H}^+ \xrightarrow{-2[\text{H}]} \text{NAD}^+ \ (\text{NADP}^+)$$

F



---

[a] These are shown as removal of two H atoms [H] and may occur by transfer of H^-, $H^+ + e^-$, or $2H^+ + 2e^-$. They represent reaction type 9C of the table inside the back cover.

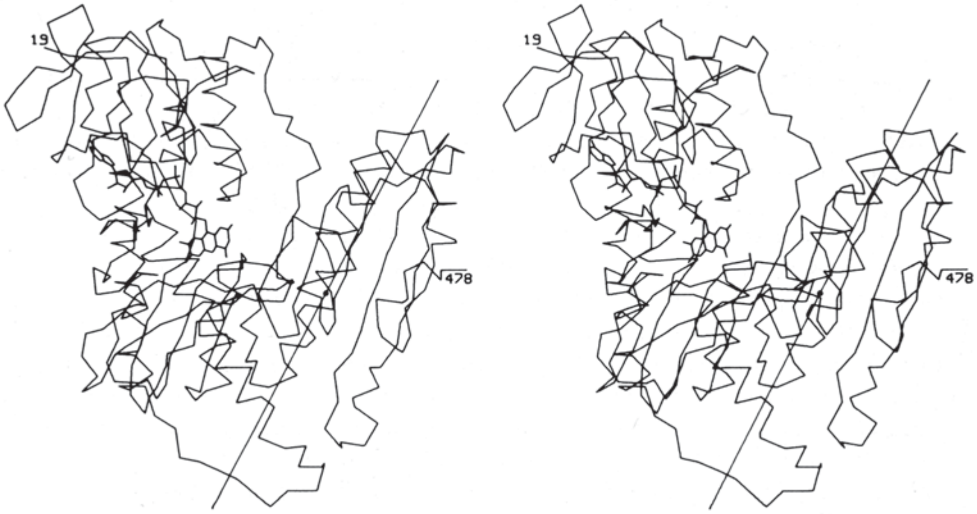
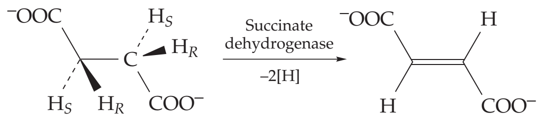


Figure 15-10 The three-dimensional structure of glutathione reductase. Bound FAD is shown. NAD^+ binds to a separate domain below the FAD. The two cysteine residues forming the reducible disulfide loop are indicated by dots. From Thieme *et al*.[182]

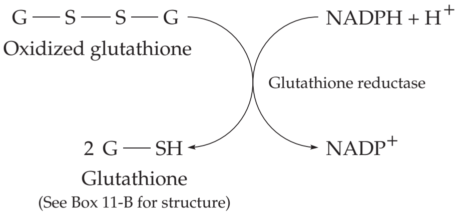
Animal mitochondria contain several different **acyl-CoA dehydrogenases** with differing preferences for chain length or branching pattern.[174–177] A related reaction that occurs in the citric acid cycle is dehydrogenation of succinate to fumarate by **succinate dehydrogenase** (Eq. 15-21).[177a] The dehydrogenation also involves *trans* removal of one of the two hydrogens, one *pro-S* hydrogen and one *pro-R*.[178] The enzyme has a large 621-residue flavoprotein subunit and a smaller 27-kDa iron–sulfur protein subunit.[179–181]



(15-21)

Neither the acyl-CoA dehydrogenases nor succinate dehydrogenase react with O_2. Acyl-CoA dehydrogenases pass the electrons removed from substrates to another flavoprotein, a soluble electron transferring flavoprotein (p. 794), which carries the electrons to an iron–sulfur protein embedded in the inner mitochondrial membrane where they enter the electron-transport chain. Succinate dehydrogenase as well as NADH dehydrogenase (Table 15-2, reaction E)[183] are embedded in the same membrane and also pass their electrons to iron–sulfur clusters and eventually to oxygen through the electron transport chain of the mitochondria (Chapter 18). **Fumarate reductase**[184] has properties similar to those of succinate dehydrogenase but catalyzes the opposite reaction in "anaerobic respiration" (Chapter 18),[184] as do similar reductases of bacteria[185] and of some eukaryotes.[186]

Dihydrolipoyl dehydrogenase (lipoamide dehydrogenase), **glutathione reductase**, and human **thioredoxin reductase**[187– 190] belong to a subclass of flavoproteins that act on dithiols or disulfides. The reaction catalyzed by the first of these is illustrated in Table 15-2 (reaction type F). The other two enzymes usually promote the reverse type of reaction, the reduction of a disulfide to two SH groups by NADPH (Eq. 15-22). Glutathione reductase splits its substrate into two halves while reduction of the small 12-kDa protein **thioredoxin** (Box 15-C) simply opens a loop in its peptide chain. The reduction of lipoic acid opens the small disulfide-containing 5-membered ring in that molecule. Each of these flavoproteins also contains within its structure a reducible disulfide group that participates in catalysis.

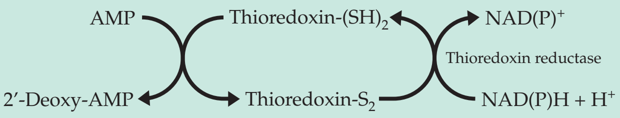


(15-22)

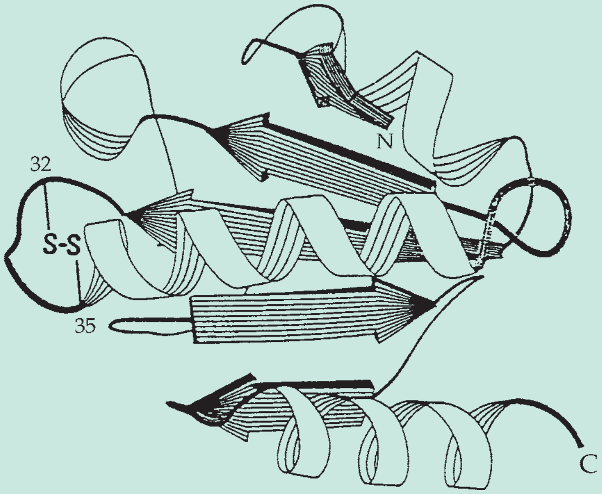
Each 50-kDa chain of the dimeric glutathione reductase is organized into three structural domains (Fig. 15-10).[191–193] Two of the domains each contain a nucleotide-binding motif resembling those of the NAD^+-dependent dehydrogenases. One of these domains binds NADPH and the other FAD. The latter domain also contains the reducible disulfide which is formed from Cys 58 and Cys 63. It serves as an intermediate hydrogen carrier which can in turn reduce oxidized glutathione. A **trypanothione reductase** from trypanosomes and related flagellated protozoa has a similar structure and acts on trypanothione, which replaces glutathione in these organisms.[188,194-196] Because it is unique to trypanosomes, this enzyme is a target for design of drugs against these organisms

BOX 15-C THIOREDOXIN AND GLUTAREDOXIN

The small proteins thioredoxin and glutaredoxin are present in relatively high concentrations in bacteria, plants, and animals. For example, thioredoxin has a concentration of 15 µM in *E. coli*. Both proteins were discovered by their role as reducing agents in conversion of the ribonucleotides AMP, GMP, CMP, and UMP to the corresponding 2-deoxyribonucleotides which are needed for synthesis of DNA:[a,b]



The redox group in thioredoxin is a **disulfide loop** located on a protrusion at one end of the molecule. In the 108-residue *E. coli* thioredoxin, as well as in the 105-residue human thioredoxin,[c] it is formed by cysteines 32 and 35 which are present in the conserved sequence CGPC. The –SH groups of these two interacting cysteines have pK_a values of ~6.9 and 7.5, the former belonging predominantly to the more exposed Cys 32.[d] The buried Asp 26 carboxyl group, which may be a proton donor to Cys 35 during reduction of the disulfide form,[e] forms a salt bridge with the Lys 57 –NH_3^+ and has a high pK_a of ~7.4.[e–g]

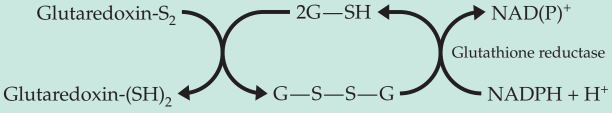


Oxidized *E. coli* thioredoxin. From Langsetmo *et al*.[f]

The pK_a assignments, which have been controversial, are discussed in Chapter 7. This disulfide loop is reduced by NADPH through the action of the flavoprotein enzyme **thioredoxin reductase**. The resulting thiol pair of the reduced thioredoxin is the reductant used for ribonucleotide reductase (Chapter 16). The standard redox potential $E°'$ of *E. coli* thioredoxin is –0.27 V, appropriately low for coupling to the NADPH / NADP$^+$ system.

Reduced thioredoxin has a variety of functions.[h] It is the reductant for conversion of methionine sulfoxide to methionine in bacteria, for reduction of sulfate in yeast, and for additional specific enzymatic reactions.[b,h] However, its major function may be to reduce disulfide linkages in various proteins.[i] Several photosynthetic enzymes are activated by reduction of disulfide linkages via photosynthetically generated reduced ferredoxin and thioredoxin[j,k] (Chapter 23). Reduced thioredoxin may also play a similar role in nonphotosynthetic cells. It may reduce mixed disulfides such as those formed between glutathione and proteins (Box 11-B).[l–n] Thioredoxin may participate in regulation of the level of nitric oxide (NO) in tissues[m] and it is needed in the assembly of filamentous bacteriophages.[b,o] For reasons that are not clear, thioredoxin is also an essential subunit for a virus-induced DNA polymerase formed in *E. coli* following infection by bacteriophage T7 (see Chapter 27).

It was a surprise to discover that a mutant of *E. coli* lacking thioredoxin can still reduce ribonucleotides. In the mutant cells thioredoxin is replaced by glutaredoxin, whose active site disulfide linkage is reduced by glutathione rather than directly by NADPH. Oxidized glutathione is, in turn, reduced by NADPH and glutathione reductase. Thus, the end result is the same with respect to ribonucleotide reduction.



However, the two proteins have significantly different specificities and functions. The disulfide loop in glutaredoxin, whose eukaryotic forms are often called **thioltransferases**,[p] has the sequence CPYC. Although glutaredoxins are weaker reductants of mixed disulfides of proteins with glutathione than are thioredoxins,[q–s] they are more specific.

Both thioredoxin and glutaredoxin are members of a larger group of thiol:disulfide oxidoreductases which are found in all known organisms. In *E. coli* there are one thioredoxin, three different glutaredoxins,[h,t] and the periplasmic protein disulfide

BOX 15-C (continued)

isomerases DsbA and DsbC, which were discussed in Chapter 10.[u–w] Similar enzymes are found in the endoplasmic reticulum of eukaryotic cells.[u,x] Like thioredoxins, they contain disulfide loops which contain the sequences CGHC and CGYC as well as CPHC. Thioredoxin reductase itself also can keep cytoplasmic bacterial proteins reduced.[y] Redox potentials vary within this family of proteins:[n,s,z,aa]

	$E^{\circ\prime}$, pH 7
Thioredoxins	–.27 V
Glutaredoxins	–.20 to –.23 V
Protein disulfide isomerase	–.18 V
DsbA	–.09 to –.12 V

These differences are correlated with differing functions. The intracellular reduced thioredoxins are thermodynamically the best reductants of disulfide linkages in proteins and they help keep intracellular proteins reduced. Glutaredoxin can be reduced efficiently by reduced glutathione or by NADPH and glutathione reductase and can, in turn, reduce cysteine and the oxidized form of vitamin C, **dehydroascorbic acid** (Box 18-D).[bb,cc] The periplasmic bacterial proteins DsbA and DsbC have the highest redox potentials and an unusually low first pK_a for the dithiol pair in their thioredoxinlike domains.[aa] The basis for these properties has been hard to understand but is consistent with their role in assisting formation of disulfide bridges in extracellular proteins and with the role of the related protein disulfide isomerase in the ER.[u,dd–ff] These disulfide exchanges are nucleophilic displacement reactions (Chapter 12).

---

[a] Holmgren, A. (1981) *Trends Biochem. Sci.* **6**, 26–29

[b] Holmgren, A. (1989) *J. Biol. Chem.* **264**, 13963–13966

[c] Forman-Kay, J. D., Clore, G. M., Wingfield, P. T., and Gronenborn, A. M. (1991) *Biochemistry* **30**, 2685–2698

[d] Dyson, H. J., Jeng, M.-F., Tennant, L. L., Slaby, I., Lindell, M., Cui, D.-S., Kuprin, S., and Holmgren, A. (1997) *Biochemistry* **36**, 2622–2636

[e] Jeng, M.-F., and Dyson, H. J. (1996) *Biochemistry* **35**, 1–6

[f] Langsetmo, K., Fuchs, J. A., and Woodward, C. (1991) *Biochemistry* **30**, 7603–7609

[g] Ladbury, J. E., Wynn, R., Hellinga, H. W., and Sturtevant, J. M. (1993) *Biochemistry* **32**, 7526–7530

[h] Gvakharia, B. O., Hanson, E., Koonin, E. K., and Mathews, C. K. (1996) *J. Biol. Chem.* **271**, 15307–15310

[i] Thomas, J. A., Poland, B., and Honzatko, R. (1995) *Arch. Biochem. Biophys.* **319**, 1–9

[j] Capitani, G., Markovic-Housley, Z., DelVal, G., Morris, M., Jansonius, J. N., and Schürmann, P. (2000) *J. Mol. Biol.* **302**, 135–154

[k] Buchanan, B. B., Schürmann, P., Decottignies, P., and Lozano, R. M. (1994) *Arch. Biochem. Biophys.* **314**, 257–260

[l] Wynn, R., Cocco, M. J., and Richrds, F. M. (1995) *Biochemistry* **34**, 11807–11813

[m] Nikitovic, D., and Holmgren, A. (1996) *J. Biol. Chem.* **271**, 19180–19185

[n] Prinz, W. A., Åslund, F., Holmgren, A., and Beckwith, J. (1997) *J. Biol. Chem.* **272**, 15661– 15667

[o] Russel, M., and Model, P. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 29–33

[p] Srinivasan, U., Mieyal, P. A., and Mieyal, J. J. (1997) *Biochemistry* **36**, 3199–3206

[q] Katti, S. K., Robbins, A. H., Yang, Y., and Wells, W. W. (1995) *Protein Sci.* **4**, 1998–2005

[r] Gravina, S. A., and Mieyal, J. J. (1993) *Biochemistry* **32**, 336–3376

[s] Nikkola, M., Gleason, F. K., and Eklund, H. (1993) *J. Biol. Chem.* **268**, 3845–3849

[t] Åslund, F., Nordstrand, K., Berndt, K. D., Nikkola, M., Bergman, T., Ponstingl, H., Jörnvall, H., Otting, G., and Holmgren, A. (1996) *J. Biol. Chem.* **271**, 6736–6745

[u] Freedman, R. B., Hirst, T. R., and Tuite, M. F. (1994) *Trends Biochem. Sci.* **19**, 331–335

[v] Rietsch, A., Belin, D., Martin, N., and Beckwith, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13048–13053

[w] Darby, N. J., and Creighton, T. E. (1995) *Biochemistry* **34**, 16770–16780

[x] Kanaya, E., Anaguchi, H., and Kikuchi, M. (1994) *J. Biol. Chem.* **269**, 4273–4278

[y] Derman, A. I., Prinz, W. A., Belin, D., and Beckwith, J. (1993) *Science* **262**, 1744–1747

[z] Chivers, P. T., Prehoda, K. E., and Raines, R. T. (1997) *Biochemistry* **36**, 4061–4066

[aa] Jacobi, A., Huber-Wunderlich, M., Hennecke, J., and Glockshuber, R. (1997) *J. Biol. Chem.* **272**, 21692–21699

[bb] Wells, W. W., Xu, D. P., Yang, Y., and Rocque, P. A. (1990) *J. Biol. Chem.* **265**, 15361–15364

[cc] Bischoff, R., Lepage, P., Jaquinod, M., Cauet, G., Acker-Klein, M., Clesse, D., Laporte, M., Bayol, A., Van Dorsselaer, A., and Roitsch, C. (1993) *Biochemistry* **32**, 725–734

[dd] Hwang, C., Sinskey, A. J., and Lodish, H. F. (1992) *Science* **257**, 1496–1502

[ee] Ruoppolo, M., Freedman, R. B., Pucci, P., and Marino, G. (1996) *Biochemistry* **35**, 13636–13646

[ff] Couprie, J., Vinci, F., Dugave, C., Quéméneur, E., and Mourtiez, M. (2000) *Biochemistry* **39**, 6732–6742

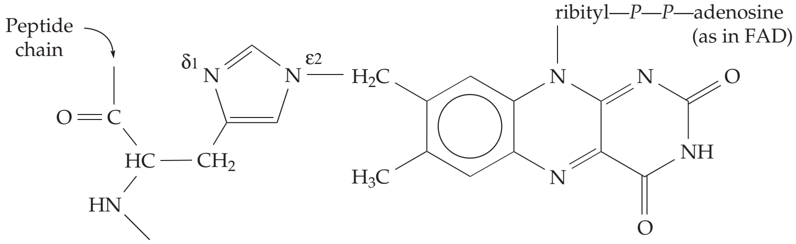
---

which cause such terrible diseases as African sleeping sickness and Chagas disease.[195,197]

Another flavoprotein constructed on the glutathione reductase pattern is the bacterial plasmid-encoded **mercuric reductase** which reduces the highly toxic Hg^{2+} to volatile elemental mercury, Hg^0. A reducible disulfide loop corresponding to that in glutathione reductase is present in this enzyme but there is also a second pair of cysteines nearby. All of these may participate in binding and reduction of Hg^{2+}.[198–199a]

3. More Flavoproteins

Flavoproteins function in virtually every area of metabolism and we have considered only a small fraction of the total number. Here are a few more. Flavin-dependent reductases use hydrogen atoms from NADH or NADPH to reduce many specific substances or classes of compounds. The FAD-containing ferredoxin: $NADP^+$ oxidoreductase catalyzes the reduction of free $NADP^+$ by reduced ferredoxin generated in the chloroplasts of green leaves.[200,201] Similar enzymes, some of which utilize reduced flavodoxins, are found in bacteria.[202,203] The FMN-containing subunit of NADH: ubiquinone oxidoreductase is an essential link in the mitochondrial electron transport chain for oxidation of NADH in plants and animals[183,204,205] and for related processes in bacteria. **Glutamate synthase**, a key enzyme in the nitrogen metabolism of plants and microorganisms, uses electrons from NADPH to reduce 2-oxoglutarate to glutamate in a complex glutamine-dependent process (see Fig. 24-5). The enzyme contains both FMN and FAD and three different iron–sulfur clusters.[205a] Flavin reductases use NADH or NADPH to reduce free riboflavin, FMN, or FAD needed for various purposes[206,206a] including emission of light by luminous bacteria.[207] They provide electrons to many enzymes that react with O_2 such as the cytochromes P450[208,209] and nitric oxide synthase (Chapter 18). An example is adrenodoxin reductase (see chapter banner, p. 764), which passes electrons from NADPH to cytochrome P450 via the small redox protein adrenodoxin. This system functions in steroid biosynthesis as is indicated in Fig. 22-7.[209a,b] Other flavin-dependent reductases have protective functions catalyzing the reduction of ascorbic acid radicals,[210,211] toxic quinones,[212–214] and peroxides.[215–218]



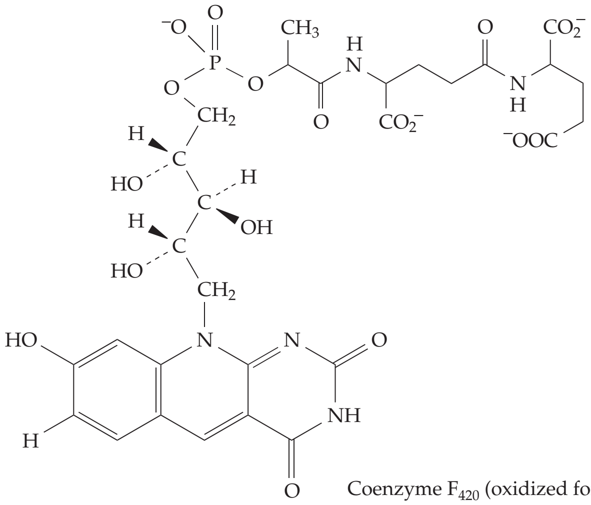
Covalently bound modified FAD of succinate dehydrogenase

4. Modified Flavin Coenzymes

Mitochondrial succinate dehydrogenase, which catalyzes the reaction of Eq. 15-21, contains a flavin prosthetic group that is covalently attached to a histidine side chain. This modified FAD was isolated and identified as 8α-($N^{\varepsilon 2}$-histidyl)-FAD.[219] The same prosthetic group has also been found in several other dehydrogenases.[220] It was the first identified member of a series of modified FAD or riboflavin 5'-phosphate derivatives that are attached by covalent bonds to the active sites of more than 20 different enzymes.[219]

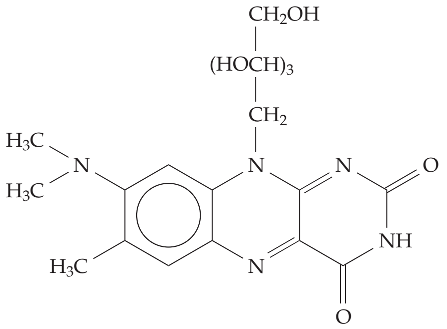
These include 8α-($N^{\varepsilon 2}$-histidyl)-FMN,[221] 8α-($N^{\delta 1}$-histidyl)-FAD,[222] 8α-(*O*-tyrosyl-FAD),[223] and 6-(*S*-cysteinyl)-riboflavin 5'-phosphate, which is found in trimethylamine dehydrogenase (Fig. 15-9).[224] An 8-hydroxy analog of FAD (–OH in place of the 8-CH_3) has been isolated from a bacterial electron-transferring flavoprotein.[225] Commercial FAD may contain some riboflavin 5'-pyrophosphate which activates some flavoproteins and inhibits others.[226]

Methanogenic bacteria contain a series of unique coenzymes (Section F) among which is **coenzyme F_{420}**, a 5-deazaflavin substituted by H at position 7 and –OH at position 8 (8-hydroxy-7,8-didemethyl-5-deazariboflavin).[227,228]



Coenzyme F_{420} (oxidized form)

This unique redox catalyst links the oxidation of H_2 or of formate to the reduction of $NADP^+$[229] and also serves as the reductant in the final step of methane biosynthesis (see Section E).[228] It resembles NAD^+ in having a redox potential of about –0.345 volts and the tendency to be only a two-electron donor. More recently free 8-hydroxy-7,8-didemethyl-5-deazariboflavin has been identified as an essential light-absorbing chromophore in DNA photolyase of *Methanobacterium*, other bacteria, and eukaryotic algae.[230] **Roseoflavin** is not a coenzyme but an antibiotic from *Streptomyces davawensis*.[231] Many synthetic flavins have been used in studies of mechanisms and for NMR[232] and other forms of spectroscopy.

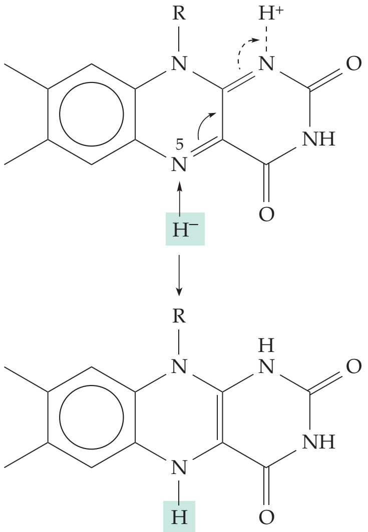


Roseoflavin

5. Mechanisms of Flavin Dehydrogenase Action

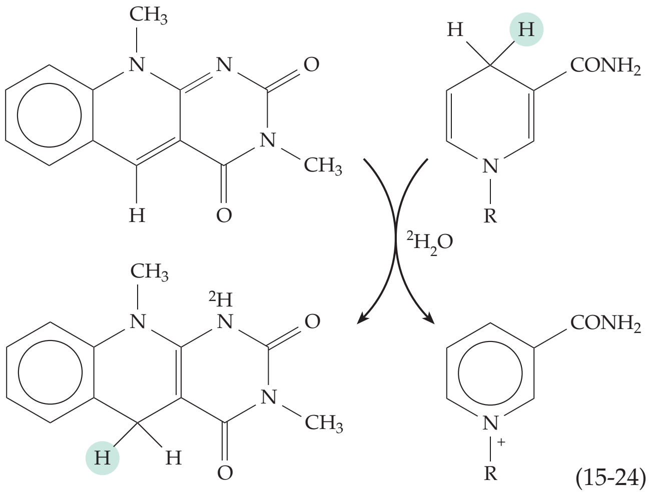
The chemistry of flavins is complex, a fact that is reflected in the uncertainity that has accompanied efforts to understand mechanisms. For flavoproteins at least four mechanistic possibilities must be considered.[153a,233] (a) A reasonable **hydride-transfer** mechanism can be written for flavoprotein dehydrogenases (Eq. 15-23). The hydride ion is donated at N-5 and a proton is accepted at N-1. The oxidation of alcohols, amines, ketones, and reduced pyridine nucleotides can all be visualized in this way. Support for such a mechanism came from study of the nonenzymatic oxidation of NADH by flavins, a reaction that occurs at moderate speed in water at room temperature. A variety of flavins and dihydropyridine derivatives have been studied, and the electronic effects observed for the reaction are compatible with the hydride ion mechanism.[234–236]

According to the mechanism of Eq. 15-23, a hydride ion is transferred directly from a carbon atom in a substrate to the flavin. However, a labeled hydrogen atom transferred to N-5 or N-1 would immediately exchange with the medium, rapid exchange being characteristic of hydrogens attached to nitrogen.



(15-23)

To avoid this problem, Brustlein and Bruice used a **5-deazaflavin** to oxidize NADH nonenzymatically.[237] When this reaction was carried out in ^2H_2O, no ^2H entered the product at C-5, indicating that a hydrogen atom (circled in Eq. 15-24) had been transferred directly from NADH to the C-5 position. Similar direct transfer of hydrogen to C-5 of 5-deazariboflavin 5'-phosphate is catalyzed by flavoproteins such as *N*-methylglutamate synthase[238] and acyl-CoA dehydrogenase.[237–239]



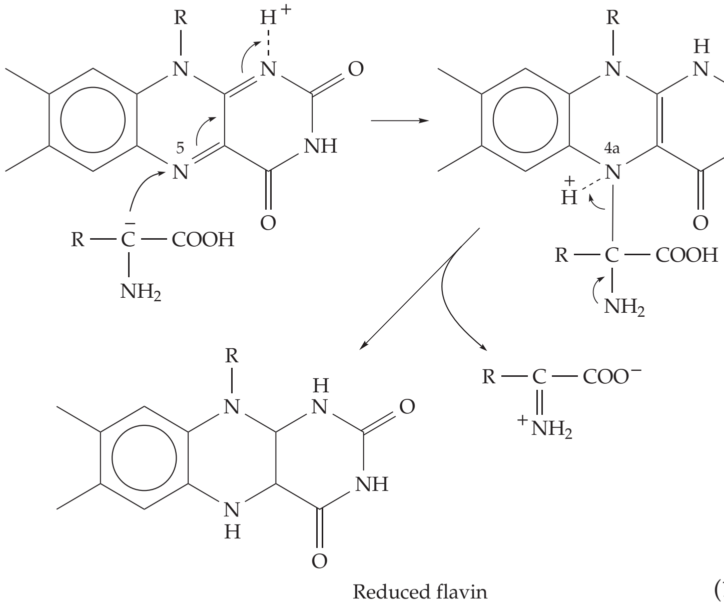
(15-24)

However, these experiments may not have established a mechanism for natural flavoprotein catalysis because the properties of 5-deazaflavins resemble those of NAD^+ more than of flavins.[239] Their oxidation–reduction potentials are low, they do not form stable free radicals, and their reduced forms don't react readily with O_2. Nevertheless, for an acyl-CoA dehydrogenase the rate of reaction of the deazaflavin is almost as fast as that of natural FAD.[238] For these enzymes a hydride ion transfer from the β CH (reaction type D of Table 15-1) is made easy by removal of the α-H of the acyl-CoA to form an enolate anion intermediate.

The three-dimensional structure of the medium chain acyl-CoA dehydrogenases with bound substrates and inhibitors is known.[174,175,240] A conserved glutamate side chain is positioned to pull the *pro-R* proton from the α carbon to create the initial enolate anion.[174,175,241] The *pro-R* β C–H lies by N-5 of the flavin ring seemingly ready to donate a hydride ion as in Eq. 15-23. NMR spectroscopy has been carried out with ^{13}C or ^{15}N in each of the atoms of the redox active part of the FAD. The results show directly the effects of strong hydrogen bonding to the protein at N-1, N-3, and N-5 and also suggest that the bound $FADH_2$ is really $FADH^-$ with the negative charge localized on N-1 by strong hydrogen bonding.[149] Many mutants have been made,[242] substrate analogs have been tested,[176,243] kinetic isotope effects have been measured,[242a] and potentiometric titrations have been done.[243] All of the results are compatible with the enolate anion hydride-transfer mechanism. Questions about the acidity of the α-H and the mechanism of its removal to form the enolate anion[242a] are similar to those discussed in Chapter 13, Section B.

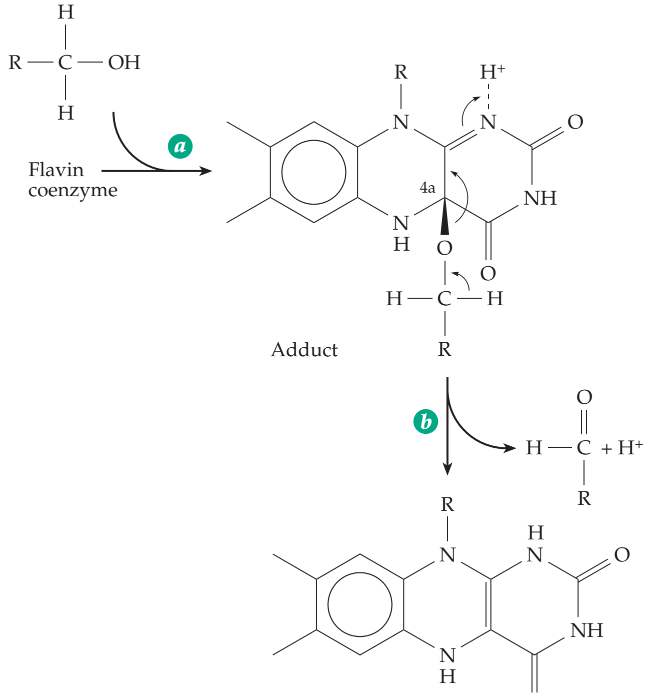
A peculiarity of several acyl-CoA dehydrogenases is a bright green color with an absorption maximum at 710 nm. This was found to result from tightly bound coenzyme A persulfide (CoA–S–S$^-$).[244,245]

(b) A second possible mechanism of flavin reduction is suggested by the occurrence of addition reactions involving the isoalloxazine ring of flavins. Sulfite adds to flavins by forming an N–S bond at the 5 position and nitroethane, which is readily dissociated to the carbanion $H_3C-CH^--NO_2$, acts as a substrate for D-amino acid oxidase.[246,247] This fact suggested a **carbanion mechanism** according to which normal D-amino acid substrates form carbanions by dissociation of the αH (Eq. 15-25). Ionization would be facilitated by binding of the substrate carboxylate to an adjacent arginine side chain and the carbanion could react at N-5 of the flavin as in Eq. 15-25. Similar mechanisms have been suggested for other flavin enzymes.[248,249]



(15-25)

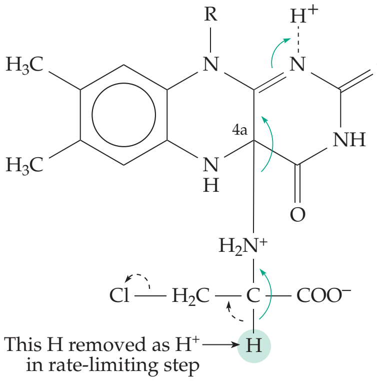
(c) The adducts with nitroethane and other compounds[250] pointed to reaction at N-5, but Hamilton[251] suggested that a better position for addition of nucleophiles is carbon 4a, which together with N-5 forms a cyclic Schiff base. He argued that other electrophilic centers in the flavin molecule, such as carbons 2, 4, and 10a, would be unreactive because of their involvement in amide or amidine-type resonance but an amine, alcohol, or other substrate could add to a flavin at position 4a (Eq. 15-26, step *a*). Cleavage of the newly formed C–O bond could then occur by movement of electrons from the alcohol part of the adduct into the flavin as indicated in step *b*. The products of this



(15-26)

4a adduct mechanism are the reduced flavin and an aldehyde, the same as would be obtained by the hydride ion mechanism. However, in Eq. 15-26, both hydrogens in the original substrate (that on oxygen and that on carbon) have dissociated as protons, the electrons having moved as a pair during the cleavage of the adduct. Hamilton argued that an isolated hydride ion has a large diameter while a proton is small and mobile; for this reason dehydrogenation may often take place by proton transfer mechanisms.

Experimental support for the mechanism of Eq. 15-26 has been obtained using D-chloroalanine as a substrate for D-amino acid oxidase.[252–254] Chloropyruvate is the expected product, but under anaerobic conditions pyruvate was formed. Kinetic data obtained with α-^{2}H and α-^{3}H substrates suggested a common intermediate for formation of both pyruvate and chloro-pyruvate. This intermediate could be an anion formed by loss of H$^+$ either from alanine or from a C-4a adduct. The anion could eliminate chloride ion as indicated by the dashed arrows in the following structure. This would lead to formation of pyruvate without reduction of the flavin. Alternatively, the electrons from the carbanion could flow into the flavin (green arrows), reducing it as in Eq. 15-26. A similar mechanism has been suggested for other flavoenzymes.[249,255] Objections to the carbanion mechanism are the expected



very high pK_a for loss of the α-H to form the carbanion[256] and the observed formation of only chloroalanine and no pyruvate in the reverse reaction of chloropyruvate, ammonia, and reduced flavoprotein.

A long-known characteristic of D-amino acid oxidase is its tendency to form charge-transfer complexes with amines, complexes in which a nonbonding electron has been transferred partially to the flavin. Complete electron transfer would yield a flavin radical and a substrate radical which could be intermediates in a **free radical mechanism**, as discussed in the next section.[256]

The three-dimensional structure of the complex of D-amino acid oxidase with the substrate analog benzoate has been determined. The carboxyl group of the inhibitor is bound by an arginine side chain (Fig. 15-11) that probably also holds the amino acid substrate. There is no basic group nearby in the enzyme that could serve to remove the α-H atom in Eq. 15-26 but the position is appropriate for a direct transfer of the hydrogen to the flavin as a hydride ion as in Eq. 15-23.[161,162,257] In spite of all arguments to the contrary the hydride ion mechanism could be correct! However, an adduct mechanism is still possible.

Experimental evidence supports a 4a adduct mechanism for glutathione reductase and related enzymes[191,258,258a] (Fig. 15-12). In this figure the reaction sequence is opposite to that in Eq. 15-26. The enzyme presumably functions as follows. NADPH binds next to the bound FAD and reduces it by transfer of the 4-*pro-S* hydrogen of the NADPH (Fig. 15-12, step *a*). The sulfur atom of Cys 63 is in van der Waals contact with the bound FAD at or near carbon atom 4a. In step *b* the nucleophilic center on atom C-4a of $FADH_2$ attacks a sulfur atom of the disulfide loop between cysteines 58 and 63 in the protein to create a C-4a adduct of a thiolate ion with oxidized FAD and to cleave the –S–S– linkage in the loop. In step *c* the thiol group of cysteine 63 is eliminated, after which the thiol of Cys 58 attacks the nearer sulfur atom of the oxidized glutathione in a nucleophilic displacement (step *d*) to give one reduced glutathione (GSH) and a mixed disulfide of glutathione and the enzyme (G-S-S-Cys 58). The thiolate anion of Cys 63, which is stabilized by interaction with the adjacent flavin ring, then attacks this disulfide (step *e*) to regenerate the internal disulfide and to free the second molecule of reduced glutathione. The imidazole group of the nearby His 467 of the second subunit presumably participates in catalysis as may some other side chains.[191] The disulfide exchange reactions are similar to those discussed in Chapter 12.

A variation is observed for *E. coli* thioredoxin reductase. The reducible disulfide and the NADPH binding site are both on the same side of the flavin rather than on opposite sides as in Fig. 15-12.[190,259] Mercuric reductase also uses NADPH as the reductant transferring the 4*S* hydrogen. The Hg^{2+} presumably binds to a sulfur atom of the reduced disulfide loop and there undergoes reduction. The observed geometry of the active site is correct for this mechanism.

6. Half-Reduced Flavins

A possible mechanism of flavin dehydrogenation consists of consecutive transfer of a hydrogen atom and of an electron with intermediate radicals being formed both on the flavin and on the substrate. Such

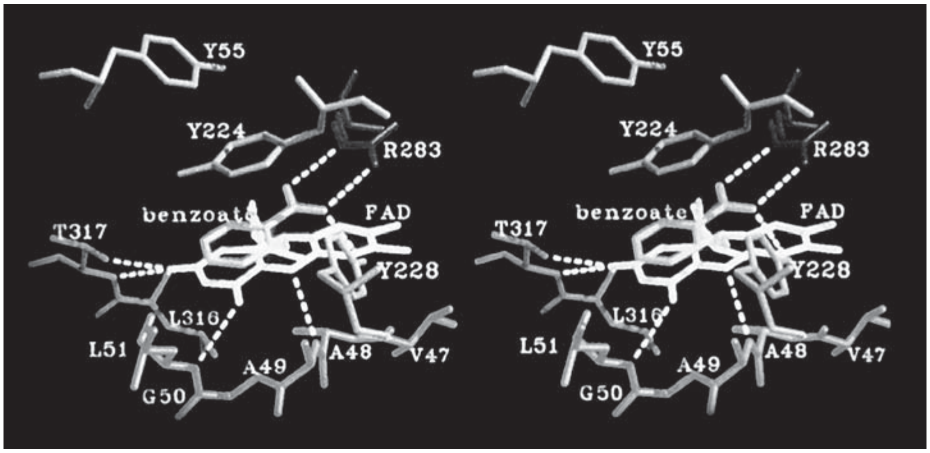


Figure 15-11 Stereoscopic view of the benzoate ion in its complex with D-amino acid oxidase. A pair of hydrogen bonds binds the carboxylate of the ligand to the guanidinium group of R283. Several hydrogen bonds to the flavin ring of the FAD are also indicated. Courtesy of Retsu Miura.[161]

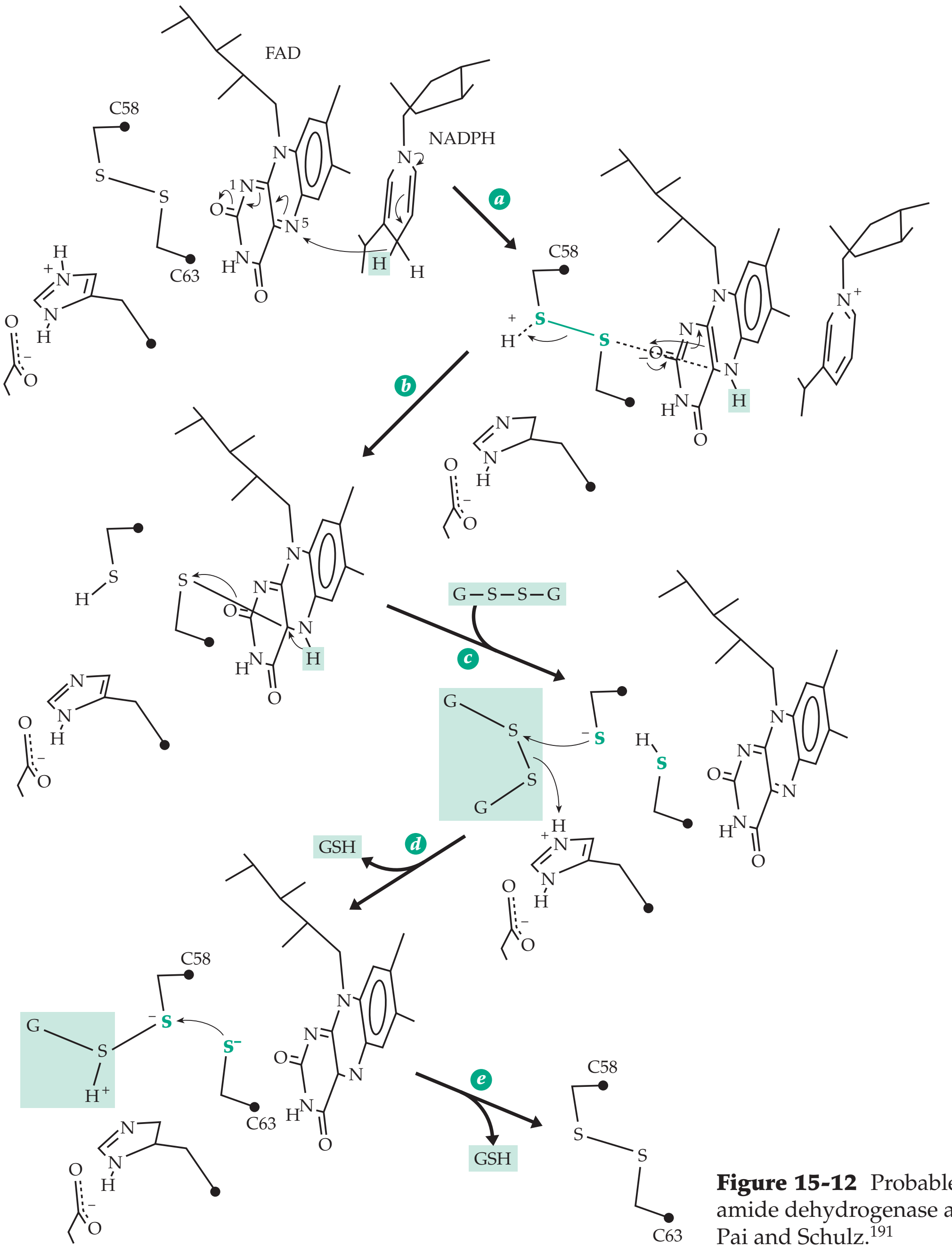


Figure 15-12 Probable reaction mechanism for lipoamide dehydrogenase and glutathione reductase. See Pai and Schulz.[191]

a mechanism takes full advantage of one of the most characteristic properties of flavins, their ability to accept single electrons to form **semiquinone** radicals. If the oxidized form Fl of a flavin is mixed with the reduced form FlH_2, a single hydrogen atom is transferred from FlH_2 to Fl to form two ${}^{\bullet}F1H$ radicals (Eq. 15-27).

$$Fl + FlH_2 \overset{K_f}{\rightleftharpoons} 2\ {}^{\bullet}FlH \tag{15-27}$$

The equilibrium represented by this equation is independent of pH, but because all three forms of the flavin have different pK_a values (Fig. 15-13) the apparent equilibrium constants relating total concentrations of oxidized, reduced, and radical forms vary with pH.[143,260–262] The fraction of radicals present is greater at low pH and at high pH than at neutrality. For a 3-alkylated flavin the formation constant K_f has been estimated as 2.3 x 10^{-2} and for riboflavin[260] as 1.5 x 10^{-2}. From these values and the pK_a values in Fig. 15-13, it is possible to estimate the amount of radical present at any pH.

Neutral flavin radicals have a blue color (the wavelength of the absorption maximum, λ_{max}, is ~560 nm) but either protonation at N-1 or dissociation of a proton from N-5 leads to red cation or anion radicals with λ_{max} at ~477 nm. Both blue and red radicals are

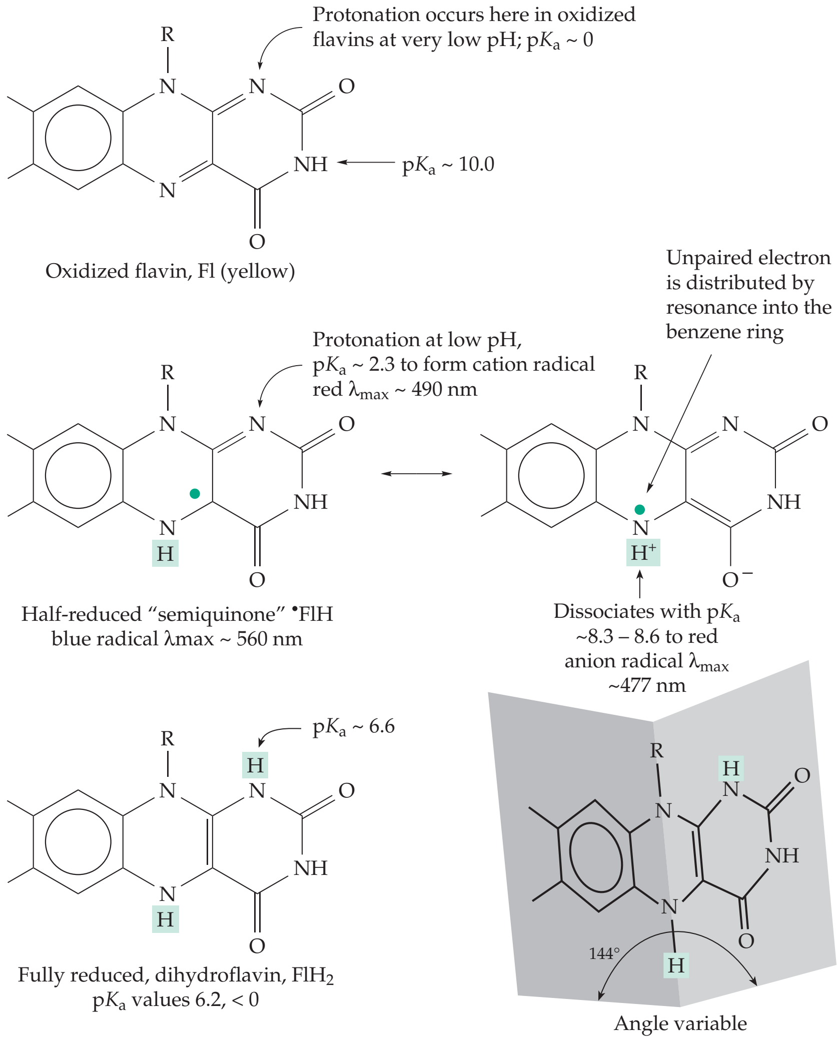
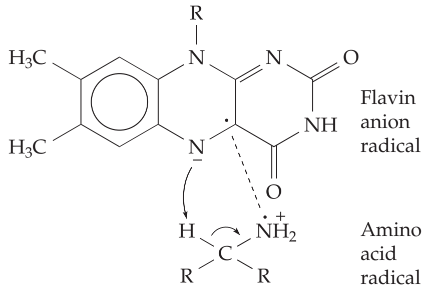

Protonation occurs here in oxidized flavins at very low pH; pK_a ~ 0
R
N
N
O
NH
pK_a ~ 10.0
O
Oxidized flavin, Fl (yellow)
Protonation at low pH, pK_a ~ 2.3 to form cation radical red λ_{max} ~ 490 nm
Unpaired electron is distributed by resonance into the benzene ring
H
H^+
O^-
Half-reduced "semiquinone" •FlH blue radical λmax ~ 560 nm
Dissociates with pK_a ~8.3 – 8.6 to red anion radical λ_{max} ~477 nm
pK_a ~ 6.6
Fully reduced, dihydroflavin, FlH_2 pK_a values 6.2, < 0
144°
Angle variable

Figure 15-13 Properties of oxidized, half-reduced, and fully reduced flavins. See Müller *et al.*[263,264]

observed in enzymes, with some enzymes favoring one and some the other. Hemmerich suggested that enzymes forming red radicals make a strong hydrogen bond to the proton in the 5 position of the flavin. This increases the basicity of N-1 leading to its protonation and formation of the red cation radicals.

It is possible that many flavoprotein oxygenases and dehydrogenases react via free radicals. For example, instead of the mechanism of Eq. 15-26, an electron could be transferred to the flavin, leaving a radical pair (at right). The crystallographic structure and modeling of the substrate complex support this possibility.[265,265a] In this pair the flavin radical would be more basic than in the fully oxidized form and the amino acid radical would be more acidic than in the uncharged form. A proton transfer as indicated together with coupling of the radical pair would yield the same product as the mechanism of Eq. 15-26.


R
H_3C
H_3C
N
N
O
NH
O
Flavin anion radical
H
$\dot{N}H_2^+$
C
R
R
Amino acid radical

Another alternative to a radical pair is a hydrogen atom transfer followed by a second electron transfer.[265]

If an enzyme binds a flavin radical much more tightly than the fully oxidized or reduced forms, reduction of the flavoprotein will take place in two one-electron steps. In such proteins the values of $E°'$ for the two steps may be widely separated. The best known examples are the small, low-potential electron-carrying proteins known as **flavodoxins**.[266–269a] These proteins, which carry electrons between pairs of other redox proteins, have a variety of functions in anaerobic and photosynthetic bacteria, cyanobacteria, and green algae. Their functions are similar to those of the **ferredoxins**, iron–sulfur proteins that are considered in Chapter 16. In some bacteria ferredoxin and flavodoxin are interchangeable and the synthesis of flavodoxin is induced if the bacteria become deficient in iron. Flavodoxins all contain riboflavin monophosphate, which functions by cycling between the fully reduced anionic form and a blue semiquinone radical.[270] The two reduction steps, from oxidized flavin to semiquinone and from semiquinone to dihydroflavin, are well separated. For example, the values of $E°'$ (pH 7) for the flavodoxin from *Megasphaera elsdenii* are −0.115 and −0.373 V, while those of the *Azotobacter vinlandii* flavodoxin (azotoflavin) are + 0.050 and – 0.495 V. The latter is the lowest known for any flavoprotein.

Flavodoxins are small proteins with an α/β structure resembling that of the nucleotide binding domain of dehydrogenases. According to ^{31}P NMR data, the phosphate group of the coenzyme bound to flavodoxin is completely ionized,[271] even though it is deeply buried in the protein and is not bound to any positively charged side chain but to the N terminus of an α helix and to four –OH groups of serine and threonine side chains. The flavin ring is partially buried near the surface of the 138-residue protein. An aromatic side chain, from tryptophan or tyrosine, lies against the flavin on the outside of the molecule. Flavodoxins can be crystallized in all three forms: oxidized, semiquinone, and fully reduced. In the crystals the flavin semiquinone, like the oxidized flavin, is nearly planar.

The **DNA photolyase** of *E. coli*, an enzyme that participates in the photochemical repair of damaged DNA (Chapter 23), contains a blue neutral FAD radical with a 580-nm absorption band and an appropriate ESR signal.[230,272] In contrast, the mitochondrial **electron-transferring flavoprotein** (ETF), a 57-kDa αβ dimer containing one molecule of FAD, functions as a single electron carrier cycling between oxidized FAD and the red anionic semiquinone.[273,274] The reduced forms of the acyl-CoA dehydrogenases transfer their electrons one at a time from their FAD to the FAD of two molecules of electron-transferring flavoprotein. Therefore, an intermediate enzyme-bound radical must be present in the FAD of acyl-CoA dehydrogenase at one stage of its catalytic cycle. A related ETF from a methylotrophic bacterium accepts single electrons from reduced trimethylamine dehydrogenase (Fig. 15-9).[275]

Another flavoprotein that makes use of both one- and two-electron transfer reactions is **ferredoxin-NADP⁺ oxidoreductase** (Eq. 15-28). Its bound FAD accepts electrons one at a time from each of the two

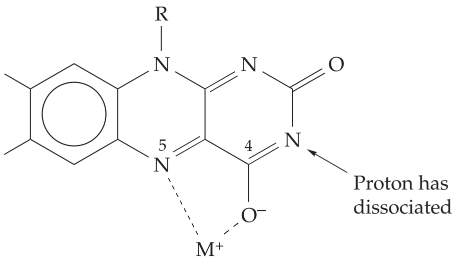
$$2Fd_{red} + NADP^+ + H^+ \rightarrow 2\,Fd_{ox}^+ + NADPH \qquad (15\text{-}28)$$

reduced ferredoxins (Fd_{red}) in chloroplasts and then presumably transfers a hydride ion to the $NADP^+$. The enzyme is organized into two structural domains,[201] one of which binds FAD and the other $NADP^+$. Similar single-electron transfers through flavoproteins also occur in many other enzymes. Chorismate mutase, an important enzyme in biosynthesis of aromatic rings (Chapter 25), contains bound FMN. Its function is unclear but involves formation of a neutral flavin radical.[276,277]

7. Metal Complexes of Flavins and Metalloflavoproteins

The presence of metal ions in many flavoproteins suggested a direct association of metal ions and flavins. Although oxidized flavins do not readily bind most metal ions, they form red complexes with Ag^+ and Cu^+ with a loss of a proton from N-3.[278] Flavin semiquinone radicals also form strong red complexes with many metals.[264] If the complexed metal ion can exist in more than one oxidation state, electron transfer between the flavin and a substrate could take place through the metal atom. However, *chelation by flavins in nature has not been observed*. Metalloflavoproteins probably function by having the metal centers close enough to the flavin for electron transfer to occur but not in direct contact. This is the case for a bacterial trimethylamine dehydrogenase in which the FeS cluster is bound about 0.4 nm from the alloxazine ring of riboflavin 5'-phosphate as shown in Fig. 15-9.

Some metalloflavoproteins contain heme groups. The previously mentioned **flavocytochrome b_2** of yeast is a 230-kDa tetramer, one domain of which carries riboflavin phosphate and another heme. A flavocytochrome from the photosynthetic sulfur bacterium *Chromatium* (cytochrome *c*-552)[279] is a complex of a 21-kDa cytochrome *c* and a 46-kDa flavoprotein containing 8α-(*S*-cysteinyl)-FAD. The 670-kDa **sulfite reductase** of *E. coli* has an $\alpha_8\beta_4$ subunit structure. The eight α chains bind four molecules of FAD and four of riboflavin phosphate, while the β chains bind three or four molecules of **siroheme** (Fig. 16-6) and also contain Fe_4S_4 clusters.[280,281] Many nitrate and some nitrite reductases are flavoproteins which also contain Mo or Fe prosthetic groups.[282,283] A group of **aldehyde oxidases** and **xanthine dehydrogenases** also contain molybdenum as well as iron (Chapter 16). In every case the metal ions are bound independently of the flavin.[283a]



Proton has dissociated

8. Reactions of Reduced Flavins with Oxygen

Free dihydroriboflavin reacts nonenzymatically in seconds and reduced flavin oxygenases react even faster with molecular oxygen to form hydrogen peroxide (Eq. 15-29).

$$FlH_2 + O_2 \rightarrow Fl + H_2O_2 \qquad (15\text{-}29)$$

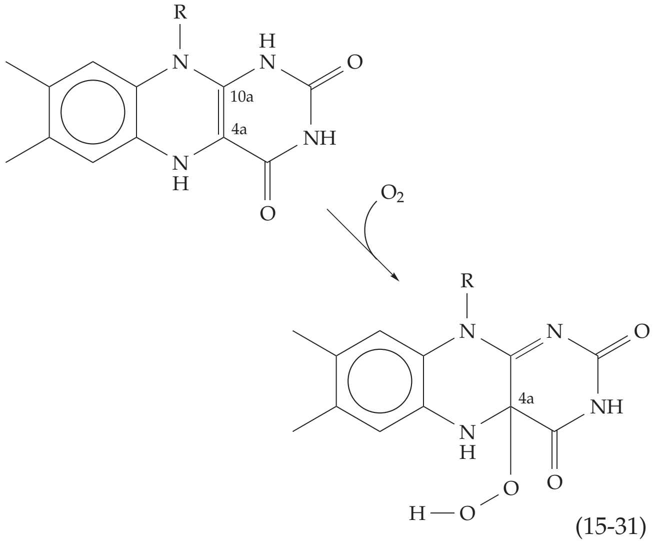
The reaction is more complex than it appears. As soon as a small amount of oxidized flavin is formed, it reacts with reduced flavin to generate flavin radicals $^\bullet FlH$ (Eq. 15-27). The latter react rapidly with O_2 each donating an electron to form superoxide anion radicals $^\bullet O_2^-$ (Eq. 15-30a) which can then combine with flavin radicals (Eq. 15-30b).[284]

$$^\bullet FlH + O_2 \rightarrow {}^\bullet O_2^- + H^+ + Fl \qquad (15\text{-}30a)$$

$$^\bullet O_2^- + {}^\bullet FlH + H^+ \rightarrow Fl + H_2O_2 \qquad (15\text{-}30b)$$

During the corresponding reactions of reduced flavoproteins with O_2, intermediates have been detected. For example, spectrophotometric studies of the FAD-containing bacterial *p*-hydroxybenzoate hydroxylase (Chapter 18) revealed the consecutive appearance of three intermediate forms.[285–287] The first, whose absorption maximum is at 380–390 nm, is thought to be an adduct at position 4a (Eq. 15-31). That such a 4a peroxide really forms with the riboflavin phosphate of the light-emitting bacterial **luciferase** (Chapter 23) was demonstrated using coenzyme enriched with ^{13}C at position 4a.[147] A large shift to lower frequency (from 104 to 83 ppm) accompanied formation of the transient adduct. Comparison with reference compounds showed that this change agreed with that predicted.

Other structures for O_2 adducts have also been considered, as has the possibility of rearrangements among these structures.[288] Nevertheless, the products observed from many different flavoprotein reactions can be explained on the basis of a 4a peroxide.[289]



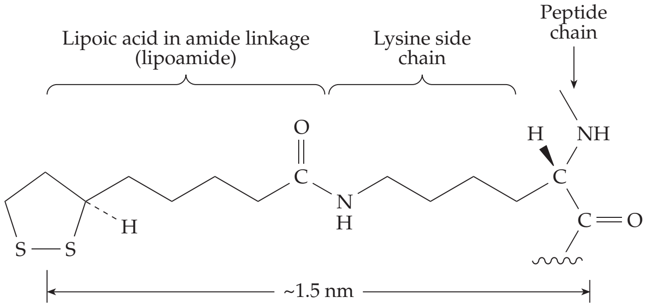
(15-31)

Formation of H_2O_2 by flavin oxidases can occur via elimination of a peroxide anion HOO^- from the adduct of Eq. 15-31 with regeneration of the oxidized flavin. In the active site of a hydroxylase, an OH group can be transferred from the peroxide to a suitable substrate (Eq. 18-42). Although radical mechanisms are likely to be involved, such hydroxylation reactions can also be viewed as transfer of OH^+ to the substrate together with protonation on the inner oxygen atom of the original peroxide to give a 4a –OH adduct. The latter is a covalent hydrate which can be converted to the oxidized flavin by elimination of H_2O. This hydrate is believed to be the third spectral intermediate identified during the action of *p*-hydroxybenzoate hydroxylase.[286,287,290]

C. Lipoic Acid and the Oxidative Decarboxylation of α-Oxoacids

The isolation of lipoic acid in 1951 followed an earlier discovery that the ciliate protozoan *Tetrahymena geleii* required an unknown factor for growth. In independent experiments acetic acid was observed to promote rapid growth of *Lactobacillus casei,* but it could be replaced by an unknown "acetate replacing factor." Another lactic acid bacterium *Streptococcus faecalis* was unable to oxidize pyruvate without addition of "pyruvate oxidation factor." By 1949, all three unknown substances were recognized as identical.[291,291a] After working up the equivalent of 10 tons of water-soluble residue from liver, Lester Reed and his collaborators isolated 30 mg of a fat-soluble acidic material which was named **lipoic acid** (or 6-thioctic acid).[292–294]

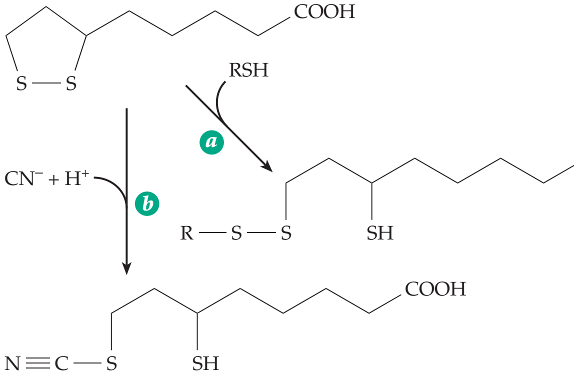
While *Tetrahymena* must have lipoic acid in its diet, we humans can make our own, and it is not considered a vitamin. Lipoic acid is present in tissues in extraordinarily small amounts. Its major function is to participate in the oxidative decarboxylation of α-oxoacids but it also plays an essential role in glycine catabolism in the human body as well as in plants.[295,296] The structure is simple, and the functional group is clearly the cyclic disulfide which swings on the end of a long arm. Like biotin, which is also present in tissues in very small amounts, lipoic acid is bound in covalent amide linkage to lysine side chains in active sites of enzymes:[296a]



1. Chemical Reactions of Lipoic Acid

The most striking chemical property of lipoic acid is the presence of ring strain of ~17–25 kJ mol^{-1} in the cyclic disulfide. Because of this, thiol groups and cyanide ions react readily with oxidized lipoic acid to give mixed disulfides (Eq. 15-32a) and isothiocyanates (Eq. 15-32b), respectively.

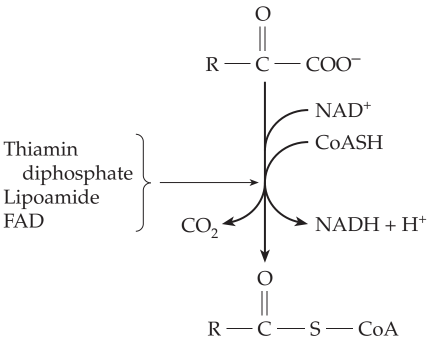
Another result of the ring strain is that the reduction potential $E°'$ (pH 7, 25°C), is –0.30 V, almost the same as that of reduced NAD (–0.32V). Thus, reoxidation of reduced lipoic acid amide by NAD^+ is thermodynamically feasible. Yet another property attributed to the ring strain in lipoic acid is the presence of an absorption maximum at 333 nm.



(15-32)

2. Oxidative Decarboxylation of Alpha-Oxoacids

The oxidative cleavage of an α-oxoacid is a major step in the metabolism of carbohydrates and of amino acids and is also a step in the citric acid cycle. In many bacteria and in eukaryotes the process depends upon both thiamin diphosphate and lipoic acid. The oxoacid anion is cleaved to form CO_2 and the remaining acyl group is combined with coenzyme A (Eq. 15-33). NAD^+ serves as the oxidant. The reaction is catalyzed by a complex of enzymes whose molecular mass varies from ~4 to 10 x 10^6, depending on the species and exact substrate.[297] Separate oxoacid dehydrogenase systems are known for pyruvate,[298–300] 2-oxoglutarate,[301] and the 2-oxoacids with branched side chains derived metabolically from leucine, isoleucine, and valine.[302,302a] In eukaryotes these enzymes are located in the mitochondria. The **pyruvate** and **2-oxoglutarate dehydrogenase** complexes of *E. coli* and *Azotobacter vinelandii* have been studied most. In both cases there are three major protein components. The first (E_1) is a **decarboxylase** (also referred to as a dehydrogenase) for which the thiamin diphosphate is the dissociable cofactor. The second (E_2) is a lipoic acid amide-containing "core" enzyme which is a **dihydrolipoyl transacylase.** The third (E_3) is the flavoprotein **dihydrolipoyl dehydrogenase**, a member of the glutathione reductase family with a three-dimensional structure and catalytic mechanism similar to those of glutathione reductase (Figs. 15-10, 15-12).[303–305]

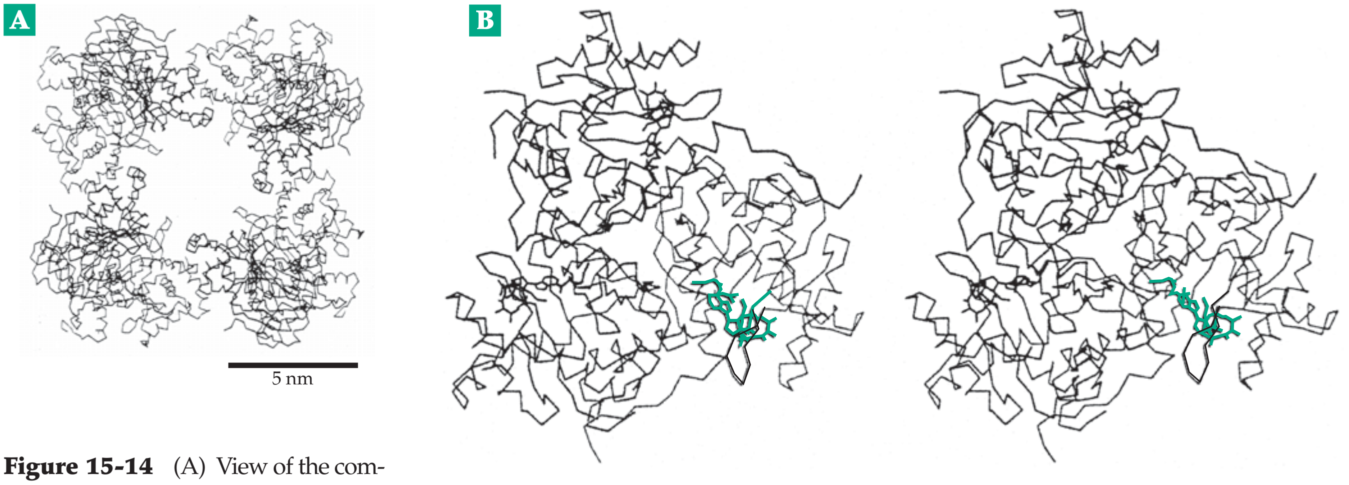


(15-33)

Electron microscopy of the core dihydrolipoyl transacylase from *E. coli* reveals a striking octahedral symmetry which has been confirmed by X-ray diffraction.[306–307a] The core from pyruvate dehydrogenase has a mass of ~2390 kDa and contains 24 identical 99.5-kDa E_2 subunits. The 2-oxoglutarate dehydrogenase from *E. coli* has a similar but slightly less symmetric structure. Each core subunit is composed of three domains. A lipoyl group is bound in amide linkage to lysine 42 and protrudes from one end of the domain. A second domain is necessary for binding to subunits E_1 and E_3, while the third major 250-residue domain contains the catalytic acyltransferase center.[308,309] This center closely resembles that of chloramphenicol acetyltransferase (Chapter 12).[310,311] The lipoyl[301,309,312] and catalytic[307] domains of the dihydrolipoyl succinyltransferase from 2-oxoglutarate dehydrogenase resemble those of pyruvate dehydrogenase and also of the branched chain oxoacid dehydrogenase. The three domains of the proteins are joined by long 25- to 30-residue segments rich in alanine, proline, and ionized hydrophilic side chains.[309] This presumably provides flexibility for the lipoyl groups which must move from site to site. The presence of unexpectedly sharp lines in the proton NMR spectrum of the core protein may be a result of this flexibility.[309]

To obtain the X-ray structure of the core protein it was necessary to delete the lipoyl- and $E_1(E_2)$- binding domains. The resulting 24-subunit structure is shown in Fig 15-14A,B.[306] It has been assumed for many years that 12 of the dimeric decarboxylase units (E_1) are bound to the 12 edges of the transacetylase cube, while six (50.6 x 2 kDa) flavoprotein (E_3) dimers bind on the six faces of the cube. The active centers of all three types of subunits are thought to come close together in the regions where the subunits touch, permitting the sequence of catalytic reactions indicated in Fig. 15-15 to take place. Eukaryotic as well as some bacterial pyruvate decarboxylases have a core of 60 subunits in an icosahedral array with 532 symmetry. This can be seen in the image reconstructions of the enzyme from *Saccharomyces cerevisiae* shown in Fig. 15-14C and D. A surprising discovery is that in this yeast enzyme the E_3 units are not on the outside of the 5-fold symmetric faces but protrude into the inner cavity.[299] Each of the 12 E_3 subunits is assisted in binding correctly to the E_2 core by a molecule of the 47-kDa **E_3-binding protein** (BP), also known as protein X.[313,314] Absence of this protein is associated with congenital lactic acidosis.

The unique function of lipoic acid is in the oxidation of the thiamin-bound active aldehyde (Fig. 15-15) in such a way that when the complex with thiamin breaks up, the acyl group formed by the oxidative decarboxylation of the oxoacid is attached to the



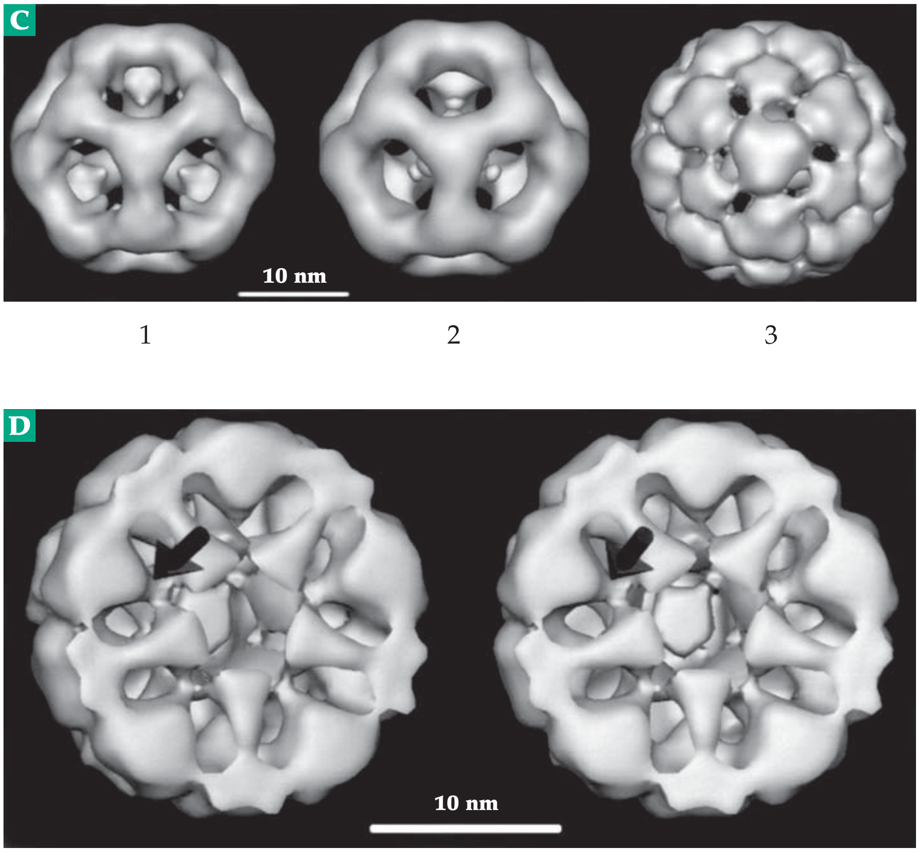


Figure 15-14 (A) View of the complex down a fourfold axis through the face of the cube. (B) Stereoscopic view of three subunits of the 24-subunit cubic core of the *E. coli* pyruvate dehydrogenase complex. The view is down the threefold axis through opposite vertices of the cube. Bound lipoate and CoA molecules are visible in each subunit. (A) and (B) from Knapp *et al*.[307a] (C) Image of the 60-subunit core of yeast pyruvate dehydrogenase reconstructed from electron micrographs of frozen-hydrated complexes. The five-, three-, and twofold symmetries can be seen. The view is along the threefold axis. 1. Truncated E_2; 2. Truncated E_2 plus binding proteins; 3. Truncated E_2 + binding protein + E_3. The reconstructions show that the E_3 and BP components are inside the truncated E_2 scaffold. The morphologies of E_3 and BP are not correctly represented since they are not bound with icosahedral symmetry. (D) Stereoscopic image of complex shown in B,3 with the front half cut away to show the interior. The conical objects protruding in are the E_3 molecules. The arrow points to a binding site for attachment of the binding protein to the core. Courtesy of J. K. Stoops *et al*.[299]

dihydrolipoyl group at the S-8 position.[315] The lipoic acid, which is attached to the flexible lipoyl domains of the core enzyme on a 1.5-nm-long arm, apparently first contacts the thiamin diphosphate site on one of the decarboxylase subunits. Bearing the acyl group, it now swings to the catalytic site on the core enzyme where CoA is bound. The acyl group is transferred to CoA producing a dihydrolipoyl group which then swings to the third subunit where the disulfide loop and a bound FAD of dihydrolipoamide dehydrogenase reoxidize the lipoyl group. The reduced flavin-disulfide enzyme is then oxidized by NAD^+ (Fig. 15-15) by the reverse of the mechanism depicted in Fig. 15-12.

Although the direct reaction of a lipoyl group with the thiamin-bound enamine (active aldehyde) is generally accepted, and is supported by recent studies,[315a] an alternative must be considered.[315] Hexacyanoferrate (III) can replace NAD^+ as an oxidant for pyruvate dehydrogenase and is also able to oxidize nonenzymatically thiamin-bound active acetaldehyde

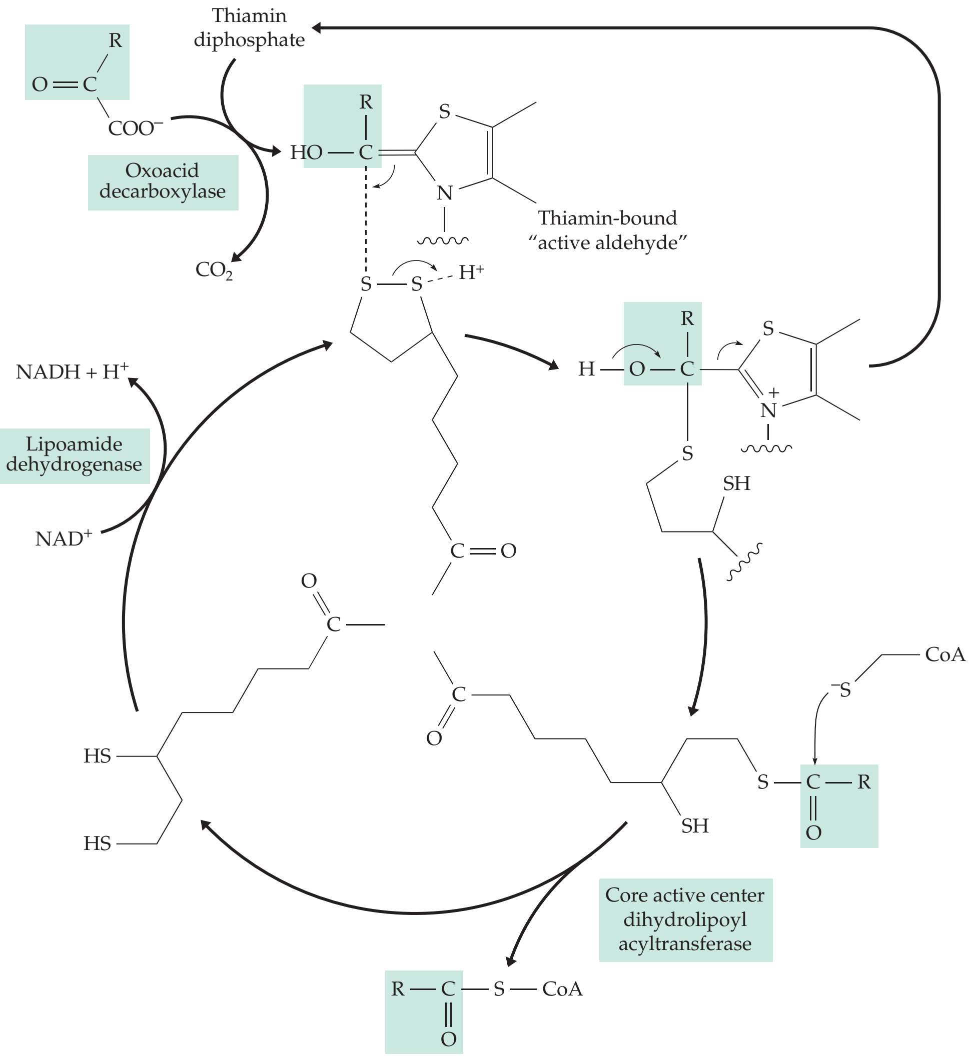


Figure 15-15 Sequence of reactions catalyzed by α-oxoacid dehydrogenases. The substrate and product are shown in boxes, and the path of the oxidized oxoacid is traced by the heavy arrows. The lipoic acid "head" is shown rotating about the point of attachment to a core subunit. However, a whole flexible domain of the core is also thought to move.

---

to 2-acetylthiamin, a compound in which the acetyl group has a high group transfer potential (Eq. 14-22). Thus, the lipoyl group could first oxidize the active aldehyde to a thiamin-bound acyl derivative, and then in the second step accept the acyl group by a nucleophilic displacement reaction. This mechanism fails to explain the unique role of lipoic acid in oxidative decarboxylation. However, as we will see in the next section, oxidative decarboxylation does not always require lipoic acid and acetylthiamin is probably an intermediate whenever lipoamide is not utilized.

Within many tissues the enzymatic activities of the pyruvate and branched chain oxoacid dehydrogenases complexes are controlled in part by a phosphorylation–dephosphorylation mechanism (see Eq. 17-9). Phosphorylation of the decarboxylase subunit by an ATP-dependent kinase produces an inactive phosphoenzyme. A phosphatase reactivates the dehydrogenase to complete the regulatory cycle (see Eq. 17-9 and associated discussion). The regulation is apparently accomplished, in part, by controlling the affinity of the protein for thiamin diphosphate.[315b] The lipoamide dehydrogenase component of all three dehydrogenase complexes appears to be the same.

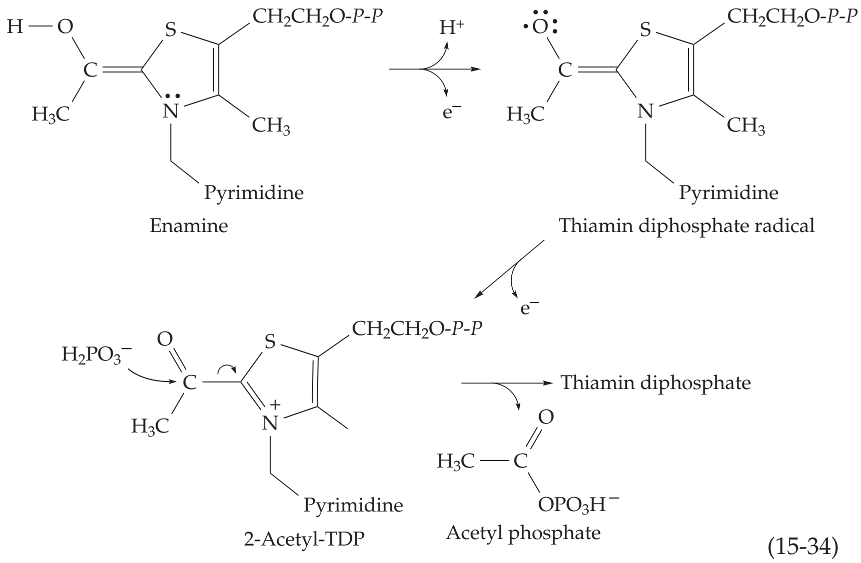
3. Other Functions of Lipoic Acid

In addition to its role in the oxidative decarboxylation of 2-oxoacids, lipoic acid functions in the human body as part of the essential mitochondrial glycine-cleavage system described in Section E. It may also participate in bacterial glycine reductase (Eq. 15-61) and other enzyme systems. Dihydrolipoamide dehydrogenase binds to G4-DNA structures of telomeres and may have a biological role in DNA-binding.[315c] Lipoic acid is being utilized as a nutritional supplement and appears to help in maintaining cellular levels of glutathione.[316]

4. Additional Mechanisms of Oxidative Decarboxylation

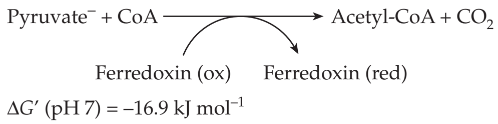
Pyruvate is a metabolite of central importance and a variety of mechanisms exist for its cleavage. The pyruvate dehydrogenase complex is adequate for strict aerobes and is very important in aerobic bacteria and in facultative anaerobes such as *E. coli*. Lactic acid bacteria, which lack cytochromes and other heme proteins, are able to carry out limited oxidation with flavoproteins such as **pyruvate oxidase**.[317] However, *strict* anaerobes have to avoid accumulation of reduced pyridine nucleotides and use a nonoxidative cleavage of pyruvate by **pyruvate formate-lyase**. The pyruvate dehydrogenase complex of *our* bodies generates NADH which can be oxidized in mitochondria to provide energy to our cells. However, NADH is a weak reductant. Some cells, such as those of nitrogen-fixing bacteria, require more powerful reductants such as reduced ferredoxin and utilize **pyruvate: ferredoxin oxidoreductase**. *Escherichia coli*, a facultative anaerobe, is adaptable and makes use of all of these types of pyruvate cleavage.[318]

Pyruvate oxidase. The soluble flavoprotein pyruvate oxidase, which was discussed briefly in Chapter 14 (Fig. 14-2, Eq. 14-22), acts together with a membrane-bound electron transport system to convert pyruvate to acetyl phosphate and CO_2.[319] Thiamin diphosphate is needed by this enzyme but lipoic acid is not. The flavin probably dehydrogenates the thiamin-bound intermediate to 2-acetylthiamin as shown in Eq. 15-34. The electron acceptor is the bound FAD and the reaction may occur in two steps as shown with a thiamin diphosphate radical intermediate.[319a] Reaction with inorganic phosphate generates the energy storage metabolite **acetyl phosphate**.



(15-34)

Pyruvate:ferredoxin oxidoreductase. Within clostridia and other strict anaerobes this enzyme catalyzes *reversible* decarboxylation of pyruvate (Eq. 15-35). The oxidant used by clostridia is the low-potential iron–sulfur ferredoxin.[320,320a] Clostridial ferredoxins contain two Fe–S clusters and are therefore two-electron oxidants. Ferredoxin substitutes for NAD^+ in Eq. 15-33 but the Gibbs energy decrease is much less (–16.9 vs – 34.9 kJ / mol. for oxidation by NAD^+).



$$\text{Pyruvate}^- + \text{CoA} \xrightarrow[\text{Ferredoxin (ox)} \quad \text{Ferredoxin (red)}]{} \text{Acetyl-CoA} + CO_2$$

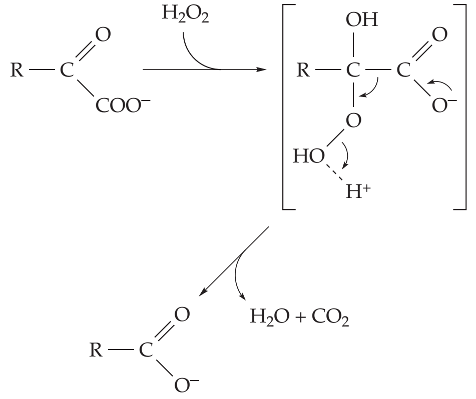
$$\Delta G' \text{ (pH 7)} = -16.9 \text{ kJ mol}^{-1}$$

(15-35)

The enzyme does not require lipoic acid. It seems likely that a thiamin-bound enamine is oxidized by an iron–sulfide center in the oxidoreductase to 2-acetylthiamin which then reacts with CoA. A free radical intermediate has been detected[318,321] and the proposed sequence for oxidation of the enamine intermediate is that in Eq. 15-34 but with the Fe–S center as the electron acceptor. Like pyruvate oxidase, this enzyme transfers the acetyl group from acetylthiamin to coenzyme A. Cleavage of the resulting acetyl-CoA is used to generate ATP. An indolepyruvate: ferredoxin oxidoreductase has similar properties.[322]

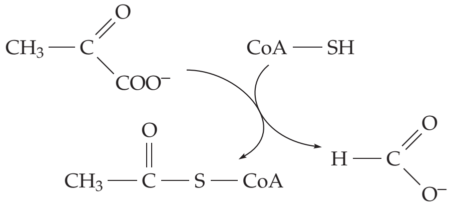
In methanogenic bacteria[320] the low-potential 5-deazaflavin coenzyme F_{420} serves as the reductant in a reversal of Eq. 15-35. A similar enzyme, **2-oxoglutarate synthase**, apparently functions in synthesis of 2-oxoglutarate from succinyl-CoA and CO_2 by photosynthetic bacteria.[323] Either reduced ferredoxin or, in *Azotobacter*, reduced flavodoxin is generated in nitrogen-fixing bacteria (Chapter 24) by cleavage of pyruvate and is used in the N_2 fixation process.

Oxidative decarboxylation by hydrogen peroxide. The nonenzymatic oxidative decarboxylation of α-oxoacids by H_2O_2 is well-known. The first step is the formation of an adduct, an organic peroxide, which breaks up as indicated in Eq. 15-36. An enzyme-catalyzed version of this reaction is promoted by **lactate monooxygenase**, a 360-kDa octameric flavoprotein obtained from *Mycobacterium smegmatis*[324] and a member of the glycolate oxidase family. Under anaerobic conditions, the enzyme produces pyruvate by a simple dehydrogenation. However, the pyruvate dissociates slowly and in the presence of oxygen it forms acetic acid, with one of the oxygen atoms of the carboxyl group coming from O_2.[324,325] Hydrogen peroxide is the usual product formed from oxygen by flavoprotein oxidases, and it seems likely that with lactate monooxygenase the hydrogen peroxide formed immediately oxidizes the pyruvate according to Eq. 15-36. The α-oxoacids formed by amino acid oxidases *in vitro* are also oxidized by accumulating hydrogen peroxide. However, if catalase (Chapter 16) is present it destroys the H_2O_2 and allows the oxoacid to accumulate.


H_2O_2
R — C
O
COO⁻
OH
O⁻
HO
H^+
$H_2O + CO_2$

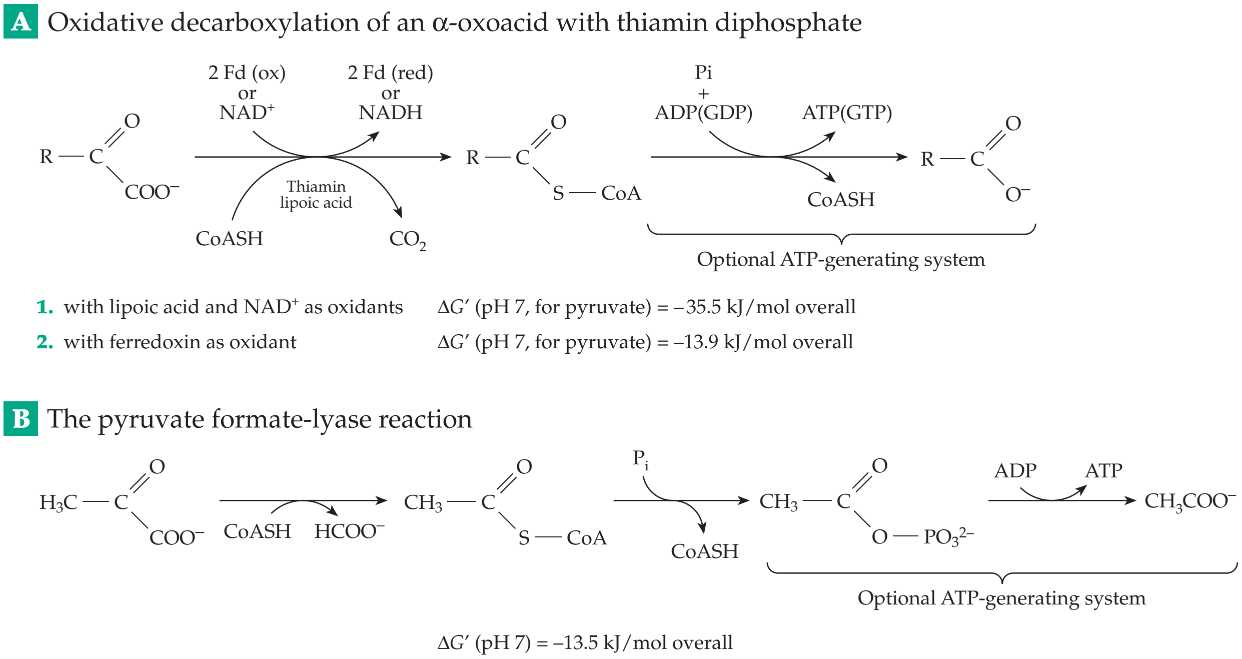
(15-36)

Pyruvate formate-lyase reaction. Anaerobic cleavage of pyruvate to acetyl-CoA and formate (Eq. 15-37) is essential to the energy economy of many cells, including those of *E. coli*. No external oxidant is needed, and the reaction does not require lipoic acid. The reaction has sometimes been called the "phosphoroclastic reaction" because the product acetyl-CoA usually reacts further with inorganic phosphate to form acetyl phosphate (see Fig. 15-16). The latter can then transfer its phospho group to ADP to form ATP.


CH_3 — C
O
COO⁻
CoA — SH
CH_3 — C — S — CoA
H — C
O⁻

ΔG' (pH 7) = –12.9 kJ mol⁻¹ (15-37)

The mechanism of the cleavage of the pyruvate in Eq. 15-37 is not obvious. Thiamin diphosphate is not involved, and free CO_2 is not formed. The first identified intermediate is an acetyl-enzyme containing a thioester linkage to a cysteine side chain. This is cleaved by reaction with CoA-SH to give the final product. A clue came when it was found by Knappe and coworkers that the active enzyme, which is rapidly inactivated by oxygen, contains a long-lived free radical.[326] Under anaerobic conditions cells convert the inactive form E_i to the active form E_a by an enzymatic reaction with *S*-adenosylmethionine and reduced flavodoxin Fd(red) as shown in Eq. 15-38.[327–329] A deactivase reverses the process.[330]


A Oxidative decarboxylation of an α-oxoacid with thiamin diphosphate

2 Fd (ox) or NAD^+
2 Fd (red) or NADH
R — C
O
COO⁻
Thiamin lipoic acid
CoASH
CO_2
S — CoA
Pi + ADP(GDP)
ATP(GTP)
CoASH
O⁻
Optional ATP-generating system

1. with lipoic acid and NAD^+ as oxidants ΔG′ (pH 7, for pyruvate) = –35.5 kJ/mol overall

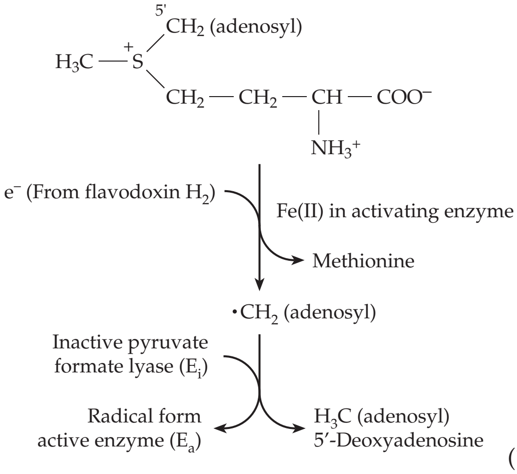
2. with ferredoxin as oxidant ΔG′ (pH 7, for pyruvate) = –13.9 kJ/mol overall

B The pyruvate formate-lyase reaction

H_3C — C
O
COO⁻
CoASH
HCOO⁻
CH_3 — C
S — CoA
P_i
CoASH
O — PO_3^{2-}
ADP
ATP
CH_3COO^-
Optional ATP-generating system

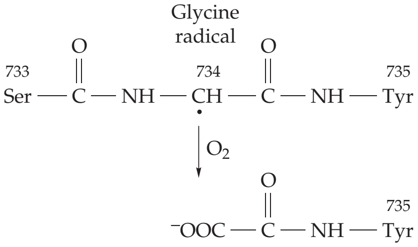
ΔG′ (pH 7) = –13.5 kJ/mol overall

Figure 15-16 Two systems for oxidative decarboxylation of α-oxoacids and for "substrate-level" phosphorylation. The value of ΔG′ = +34.5 kJ mol⁻¹ (Table 6-6) was used for the synthesis of ATP^{4-} from ADP^{3-} and HPO_4^{2-} in computing the values of ΔG′ given.


5'
CH2 (adenosyl)
H3C—S+
CH2—CH2—CH—COO−
NH3+
e− (From flavodoxin H2)
Fe(II) in activating enzyme
Methionine
•CH2 (adenosyl)
Inactive pyruvate formate lyase (Ei)
Radical form active enzyme (Ea)
H3C (adenosyl) 5′-Deoxyadenosine

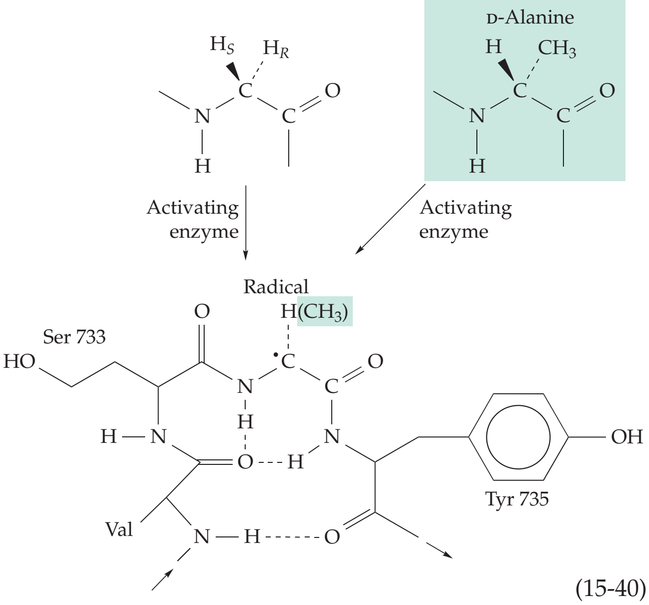
(15-38)

The activating enzyme, which is allosterically activated by pyruvate, is an iron–sulfur (Fe_4S_4) protein (Chapter 16).[331,331a] Formation of the observed radical may proceed via a 5'-deoxyadenosyl radical as has been proposed for lysine 2,3-aminomutase (Eq. 16-42).[332] The activation reaction also resembles the free radical-dependent reactions of vitamin B_{12} which are discussed in Chapter 16. When subjected to O_2 of air at 25°C pyruvate formate-lyase is destroyed with a half-life of ~10 s. The peptide chain is cleaved and sequence analysis and mass spectrometry of the resulting fragments show that the specific sequence Ser-Gly-Tyr at positions 733–755 is cut with formation of an oxalyl group on the N-terminal tyrosine of one fragment (Eq. 15-39). Various ^{13}C-containing amino acids were supplied to growing cells of *E. coli* and were incorporated into the proteins of the bacteria. Pyruvate formate-lyase containing ^{13}C in carbon-2 of glycine gave an EPR spectrum with hyperfine splitting arising from coupling of the unpaired electron of the radical with the adjacent ^{13}C nucleus.[333,333a] This experiment, together with the results described by Eq. 15-39, suggested that the radical is derived from Gly 734.


Glycine radical
733
Ser—C(=O)—NH—CH•—C(=O)—NH—Tyr
734
735
O2
−OOC—C(=O)—NH—Tyr
735

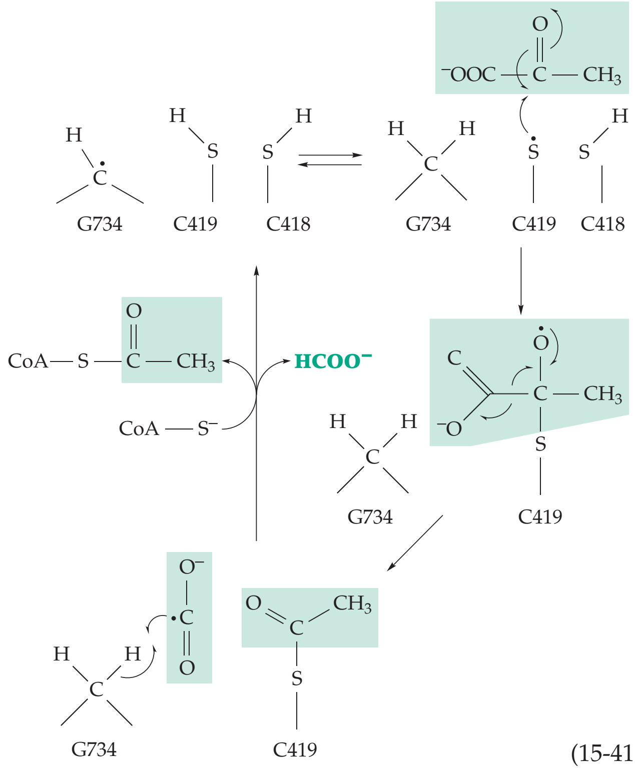
(15-39)

The activating enzyme will also generate radicals from short peptides such as Arg-Val-Ser-Gly-Tyr-Ala-Val, which corresponds to residues 731–737 of the pyruvate formate-lyase active site. If Gly 734 is replaced by L-alanine, no radical is formed, but radical *is* formed if D-alanine is in this position. This suggests that the *pro-S* proton of Gly 734 is removed by the activating enzyme[329] as illustrated in Eq. 15-40. It has been suggested that the peptide exists in the β-bend conformation shown.


HS
HR
N
C
C
O
H
D-Alanine
H
CH3
Activating enzyme
Activating enzyme
Radical
H(CH3)
O
Ser 733
HO
N
H
•C
O
C
N
H—N
O---H
Tyr 735
OH
Val
N—H------O

(15-40)

How does the enzyme work? The α-CH proton of the glycyl radical of the original *pro-R* proton undergoes an unexpected exchange with the deuterium of ^2H_2O. The exchange is catalyzed by the thiol group of Cys 419, suggesting that it is close to Gly 734 in the three-dimensional structure.[334] The mutants C419S and C418S are inactive but still allow formation of the Gly 734 radical.[334–335a] The mechanism has been proposed in Eq. 15-41.


H
•C
G734
H
S
C419
H
S
C418
H
H
C
G734
−OOC—C(=O)—CH3
•S
C419
H
S
C418
CoA—S—C(=O)—CH3
HCOO−
CoA—S−
H
H
C
G734
C
−O
•O
C—CH3
S
C419
O−
•C
O
H
H
C
G734
O
CH3
C
S
C419

(15-41)

5. Cleavage of α-Oxoacids and Substrate-Level Phosphorylation

The α-oxoacid dehydrogenases yield CoA derivatives which may enter biosynthetic reactions. Alternatively, the acyl-CoA compounds may be cleaved with generation of ATP. The pyruvate formate-lyase system also operates as part of an ATP-generating system for anaerobic organisms, for example, in the "mixed acid fermentation" of enterobacteria such as *E. coli* (Chapter 17). These two reactions, which are compared in Fig. 15-16, constitute an important pair of processes both of which accomplish substrate-level phosphorylation. They should be compared with the previously considered examples of substrate level phosphorylation depicted in Eq. 14-23 and Fig. 15-16.

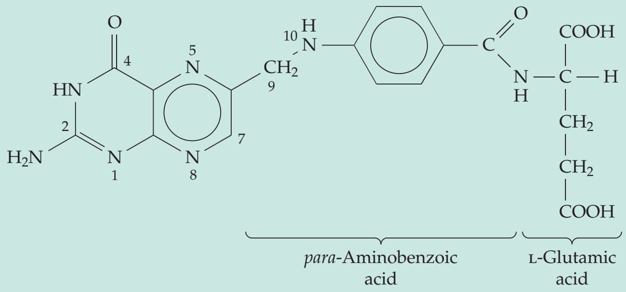
D. Tetrahydrofolic Acid and Other Pterin Coenzymes

In most organisms reduced forms of the vitamin **folic acid** serve as *carriers for one-carbon groups* at three different oxidation levels corresponding to **formic acid**, **formaldehyde**, and the **methyl group** and also facilitate their interconversion.[336,337] Moreover, folic acid is just one of the derivatives of the **pteridine** ring system that enjoy a widespread natural distribution.[338] One of these derivatives functions in the hydroxylation of aromatic amino acids and in nitric oxide synthase and yet another within several molybdenum-containing enzymes. Pteridines also provide coloring to insect wings and eyes and to the skins of amphibians and fish. They appear to act as protective filters in insect eyes and may function as light receptors. An example is a folic acid derivative found in some forms of the DNA photorepair enzyme **DNA photolyase** (Chapter 23).

1. Structure of Pterins

Because of the prevalence of its derivatives, 2-amino-4-hydroxypteridine has been given the trivial name **pterin**. Its structure resembles closely that of guanine, the 5-membered ring of the latter having been expanded to a 6-membered ring. In fact, pterins are derived

BOX 15-D FOLIC ACID (PTEROYLGLUTAMIC ACID)



In 1931, Lucy Wills, working in India, observed that patients with **tropical macrocytic anemia**, a disease in which the erythrocytes are enlarged but reduced in numbers, were cured by extracts of yeast or liver. The disease could be mimicked in monkeys fed the local diet, and a similar anemia could be induced in chicks. By 1938 it had also been shown that a factor present in yeast, alfalfa, and other materials was required for the growth of chicks. Isolation of the new vitamin came rapidly after it was recognized that it was also an essential nutrient for *Lactobacillus casei* and *Streptococcus faecalis* R.[a,b] Spinach was a rich source of the new compound, and it was named folic acid (from the same root as the word foliage).

The microbiological activity attributed to folic acid in extracts of natural materials was largely that of di- and triglutamyl derivatives, one of the facts that has led to the description of the history of folic acid as "the most complicated chapter in the story of the vitamin B complex."[a] Metabolic functions for folic acid were suggested by the observations that the requirement for *Streptococcus faecalis* could be replaced by thymine plus serine plus a purine base. Folic acid is required for the biosynthesis of all of these substances. A function in the interconversion of serine and glycine was suggested by the observation that certain mutants of *E. coli* required either serine or glycine for growth. Isotopic labeling experiments established that in the rat as well as in the yeast *Torulopsis* serine and glycine could be interconverted. It was also shown that the amount of interconversion decreased in the folate-deficient rat.

Deficiency of folic acid is a common nutritional problem of worldwide importance.[b] A recommended daily intake is 0.2 mg, but because of the association between low folic acid intake and neural tube defects in infants, women of child-bearing age should have 0.4 mg / day.[c–e] biosynthetically from guanine. The two-ring system of pterin is also related structurally and biosynthetically to that of riboflavin (Box 15-B).

---

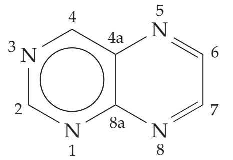
[a] Wagner, A. F., and Folkers, E. (1964) *Vitamins and Coenzymes*, Wiley (Interscience), New York

[b] Jukes, T. H. (1980) *Trends Biochem. Sci.* **5**, 112–113

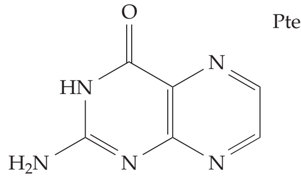
[c] Cziezel, A. E., and Dudás, I. (1992) *N. Engl. J. Med.* **327**, 1832–1835

[d] Jukes, T. H. (1997) *Protein Sci.* **6**, 254–256

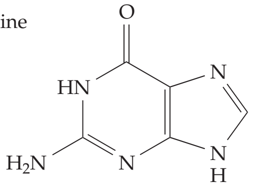
[e] Rosenquist, T. H., Ratashak, S. A., and Selhub, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15227–15232



Pteridine



Pterin: 2-amino-4-hydroxypteridine



Guanine

Interest in pteridines began with Frederick G. Hopkins, who in 1891 started his investigation of the yellow and white pigments of butterflies. Almost 50 years and a million butterflies later, the structures of the two pigments, **xanthopterin** and **leucopterin** (Fig. 15-17), were established.[339] These pigments are produced in such quantities as to suggest that their synthesis may be a means of deposition of nitrogenous wastes in dry form.

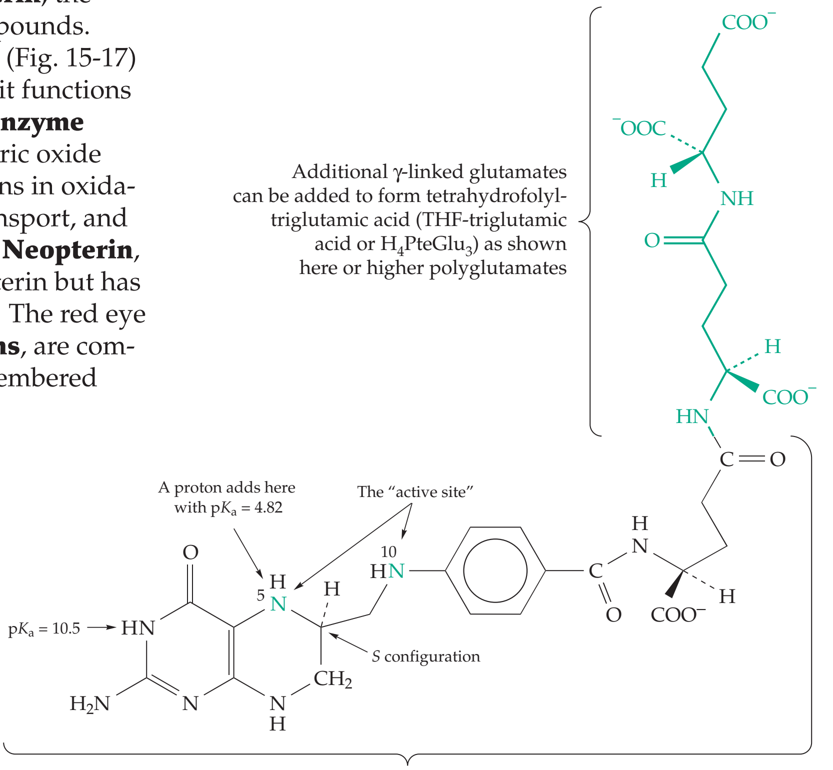
Among the simple pterins isolated from the eyes of *Drosophila*[340] is **sepiapterin** (Fig. 15-17), in which the pyrazine ring has been reduced in the 7,8 position and a short side chain is present at position 6. Reduction of the carbonyl group of sepiapterin with $NaBH_4$ followed by air oxidation produces **biopterin**, the most widely distributed of the pterin compounds. First isolated from human urine, biopterin (Fig. 15-17) is present in liver and other tissues where it functions in a reduced form as a **hydroxylation coenzyme** (see Chapter 18).[338] It is also present in nitric oxide synthase (Chapter 18).[341,342] Other functions in oxidative reactions, in regulation of electron transport, and in photosynthesis have been proposed.[343] **Neopterin**, found in honeybee larvae, resembles biopterin but has a D-*erythro* configuration in the side chain. The red eye pigments of *Drosophila*, called **drosopterins**, are complex dimeric pterins containing fused 7-membered rings (Fig. 15-17).[344,345]

In the pineal gland, as well as in the retina of the eye, light-sensitive pterins may be photochemically cleaved to generate such products as **6-formylpterin**, a compound that could serve as a metabolic regulator.[346] Another pterin acts as a chemical attractant for aggregation of the ameboid cells of *Dictyostelium lacteum*[347] (Box 11-C). **Molybdopterin** (Fig. 15-17) is a component of several Mo-containing enzymes discussed in Chapter 16,[348,349] and **methanopterin** is utilized by methanogenic bacteria.[350–353]

2. Folate Coenzymes

The coenzymes responsible for carrying single-carbon units in most organisms are derivatives of **5,6,7,8-tetrahydrofolic acid** (abbreviated H_4PteGlu, H_4folate or THF). However, in methanogenic bacteria the tetrahydro derivative of the structurally unique methanopterin (Fig. 15-17) is the corresponding single-carbon carrier.[353] Naturally occurring tetrahydrofolates contain a chiral center of the *S* configuration.[354,355] They exhibit negative optical rotation at 589 nm. The folate coenzymes are present in extremely low concentrations and the reduced ring is readily oxidized by air.

In addition to the single L-glutamate unit present in tetrahydrofolic acid, the coenzymes occur to the greatest extent as conjugates called **folyl polyglutamates** in which one to eight or more additional molecules of L-glutamic acid have been combined via amide linkages.[356–357a] The first two of the extra glutamates are always joined through the γ (side chain) carboxyl groups but in *E. coli* the rest are joined through their α carboxyls.[358] The distribution of the polyglutamates varies from one organism to the next. Some bacteria contain exclusively the triglutamate derivatives, while in others almost exclusively the tetraglutamate or octaglutamate derivatives predominate.[359] The serum



Additional γ-linked glutamates can be added to form tetrahydrofolyl-triglutamic acid (THF-triglutamic acid or H_4PteGlu$_3$) as shown here or higher polyglutamates

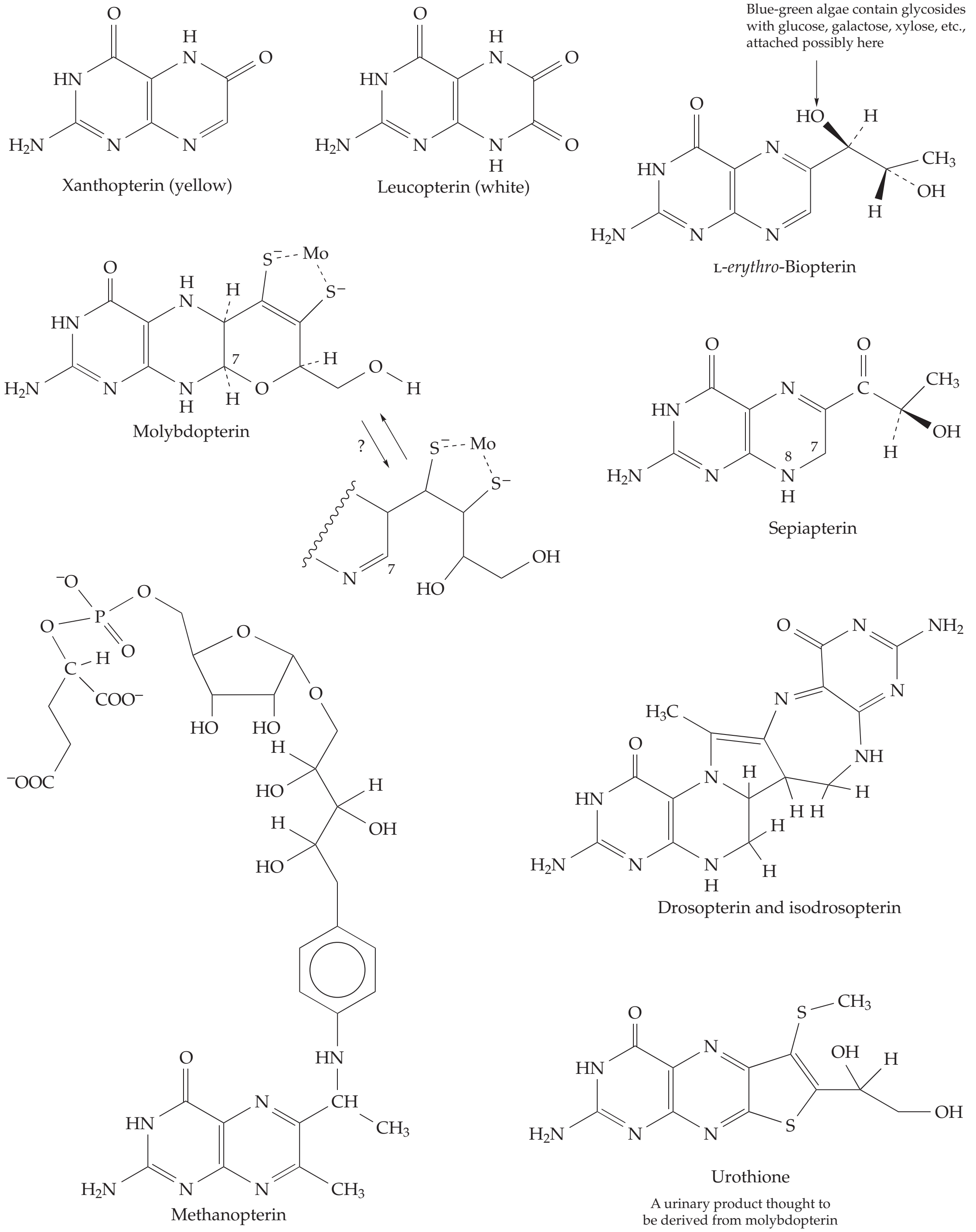
A proton adds here with pK_a = 4.82

The "active site"

pK_a = 10.5

S configuration

Tetrahydrofolic acid (THF) or tetrahydropteroylglutamic acid (H_4PteGlu)


Xanthopterin (yellow)
Leucopterin (white)
L-*erythro*-Biopterin
Molybdopterin
Sepiapterin
Drosopterin and isodrosopterin
Methanopterin
Urothione

Blue-green algae contain glycosides with glucose, galactose, xylose, etc., attached possibly here

A urinary product thought to be derived from molybdopterin

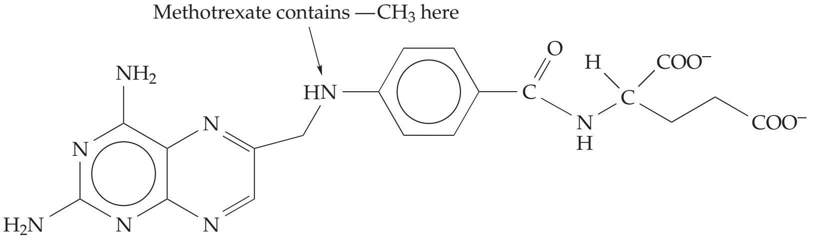
Figure 15-17 Structures of several biologically important pterins.

of many species contains only derivatives of folic acid itself but pteroylpentaglutamate is the major folate derivative present in rodent livers.[360,361] A large fraction of these pentaglutamates consists of the 5-methyl-THF derivative, while at the heptaglutamate level most consists of the free THF derivative. Folyl polyglutamate in its oxidized form is a component of some DNA photolyases (Chapter 23).[362]

3. Dihydrofolate Reductase

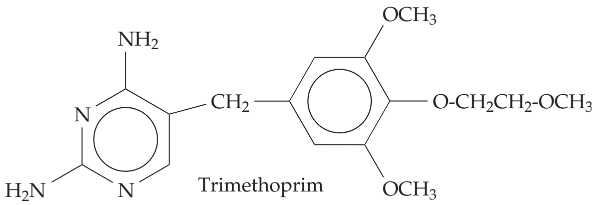
Folic acid and its polyglutamyl derivatives can be reduced to the THF coenzymes in two stages: the first step is a slow reduction with NADPH to 7,8-dihydrofolate (step *a*, Fig. 15-18). The same enzyme that catalyzes this reaction rapidly reduces the dihydrofolates to tetrahydrofolates (step *b*, Fig. 15-18). Again, NADPH is the reducing agent and the enzyme has been given the name **dihydrofolate reductase**. An unusual fact is that bacteriophage T4 not only carries a gene for its own dihydrofolate reductase but also the enzyme is a structural unit in the phage baseplate.[363]

Inhibitors. Aside from its role in providing reduced folate coenzymes for cells, this enzyme has attracted a great deal of attention because it appears to be a site of action of the important anticancer drugs **methotrexate** (amethopterin) and aminopterin.[293,364,365] These compounds inhibit dihydrofolate reductase in concentrations as low as 10^{-8} to 10^{-9} M. Methotrexate is also widely used as an **immunosuppresant** drug and in the treatment of parasitic infections.

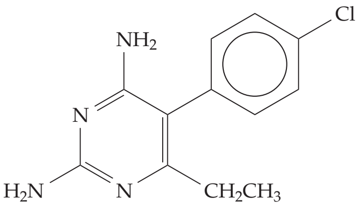

Methotrexate contains —CH_3 here

Aminopterin (4-amino-4-deoxyfolic acid)

Another dihydrofolate inhibitor **trimethoprim** is an important antibacterial drug, usually given together with a sulfonamide. Although it is not as close a structural analog of folic acid as is methotrexate, it is a potent inhibitor which binds far more tightly to the enzyme from bacteria than to that from humans.[366–368] The inhibitors **pyrimethamine** and **cycloguanil** are effective antimalarial drugs.[369,370]



Trimethoprim



Pyrimethamine

An enormous number of synthetic compounds have been prepared in the hope of finding still more effective inhibitors. By 1984, over 1700 different inhibitors of dihydrofolate reductase had been studied.[367] However, methotrexate and trimethoprim remain outstanding. Methotrexate has been in clinical use for nearly 50 years and is very effective against leukemia and some other cancers.[371] Before 1960 persons with acute leukemia lived no more than 3–6 months. However, with antifolate treatment some have lived for 5 years or more and complete cures of the relatively rare choriocarcinoma have been achieved. New methods of chemotherapy use the antifolates in combination with other drugs.

Folate coenzymes are required in the biosynthesis of both purines and thymine. Consequently, rapidly growing cancer cells have a high requirement for activity of this enzyme. However, since all cells require the enzyme the antifolates are toxic and cannot be used for prolonged therapy. An even more serious problem is the development of resistance to the drug by cancer cells, often through “amplification” of the dihydrofolate reductase gene[365,372–374] as discussed in Chapter 27. Cells may also become resistant to methotrexate and other antifolates as a result of mutations that prevent efficient uptake of the drug,[375] cause increased action of effux pumps in the cell,[376] reduce the affinity of dihydrofolate reductase for the drug,[377] or interfere with conversion of the drug to polyglutamate derivatives which are better inhibitors than free methotrexate.[378,379] Cell surface **folate receptors**, present in large numbers in some tumor cells, can be utilized to bring suitably designed antifolate drugs or even unrelated cytotoxic compounds into tumor cells.[379a] Some tumor cells are more active than normal cells in generating folyl polyglutamates, contributing to the effectiveness of methotrexate.[379] Resistance of *E. coli* cells to trimethoprin may result from acquisition by the bacteria of a new form of dihydrofolate reductase carried by a plasmid.[380]

Structure and mechanism. Dihydrofolate reductase from *E. coli* is a small 159-residue protein with a central parallel stranded sheet,[381–383] while that from higher animals is 20% larger. The three-dimensional structures of the enzymes from *Lactobacillus*,[384] chicken, mouse, and human are closely similar. The NADP binds at the C-terminal ends of β strands as in other dehydrogenases. In Fig. 15-19 the reduced nicotinamide end of $NADP^+$ is seen next to a molecule of bound dihydrofolate.[381] The side chain carboxylate of Asp 27 makes a pair of hydrogen bonds to the pterin ring as shown on the right-hand side of Eq. 15-42. As can be seen from Fig. 15-19, the nicotinamide ring of $NADP^+$ is correctly positioned to have donated a hydride ion to C-6 to form THF with the 6*S* configuration. Notice that in Eq. 15-42 the NH hydrogen atom at the 3' position

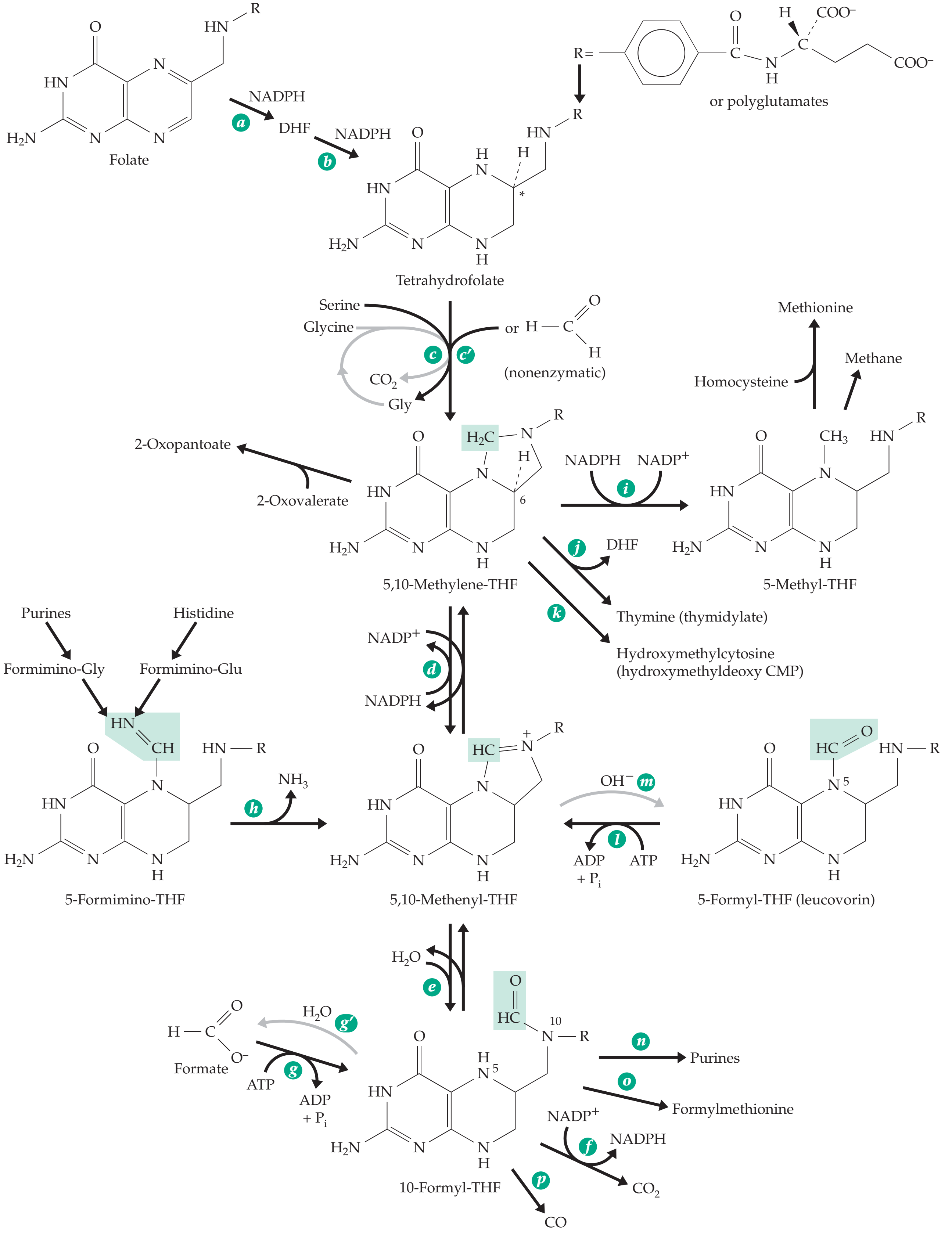
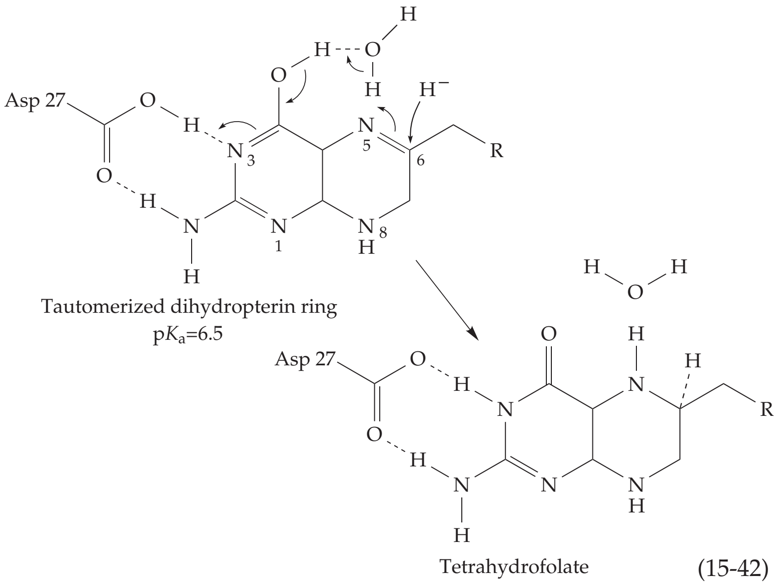

Folate
R=
or polyglutamates
a
NADPH
DHF
b
NADPH
Tetrahydrofolate
Serine
Glycine
c
c′
CO2
Gly
or H—C(=O)—H
(nonenzymatic)
Methionine
Methane
Homocysteine
2-Oxopantoate
2-Oxovalerate
5,10-Methylene-THF
NADPH
NADP+
i
5-Methyl-THF
j
DHF
k
Thymine (thymidylate)
Hydroxymethylcytosine
(hydroxymethyldeoxy CMP)
NADP+
d
NADPH
Purines
Histidine
Formimino-Gly
Formimino-Glu
NH3
h
5-Formimino-THF
5,10-Methenyl-THF
OH−
m
l
ADP
+ Pi
ATP
5-Formyl-THF (leucovorin)
H2O
e
Formate
H2O
g′
g
ATP
ADP
+ Pi
10-Formyl-THF
n
Purines
o
Formylmethionine
NADP+
f
NADPH
CO2
p
CO

Figure 15-18 Tetrahydrofolic acid and its one-carbon derivatives.

of the pterin makes a hydrogen bond to Asp 27. However, X-ray studies have shown that the pteridine ring of methotrexate is turned 180° so that its 4-NH_2 group forms hydrogen bonds to backbone carbonyls of Ala 97 and Leu 4 at the edge of the central pleated sheet while a protonated N-1 interacts with the side chain carboxylate of Asp 27.

Because of its significance in cancer therapy and its small size, dihydrofolate reductase is one of the most studied of all enzymes. Numerous NMR studies[385–387] and investigations of catalytic mechanism and of other properties[388] have been conducted. Many mutant forms have been created.[389–391] For example, substitution of Asp 27 (Asp 26 in *L. casei*) of the *E. coli*


Asp 27
O
H
N 3
O
H
N
H
N 1
N 8
H
N 5
6
R
H
O
H
H
H⁻
Tautomerized dihydropterin ring
pK_a=6.5
H
O
H
Asp 27
O
H
N
O
H
N
H
N
N
H
O
H
N
H
R
Tetrahydrofolate

(15-42)

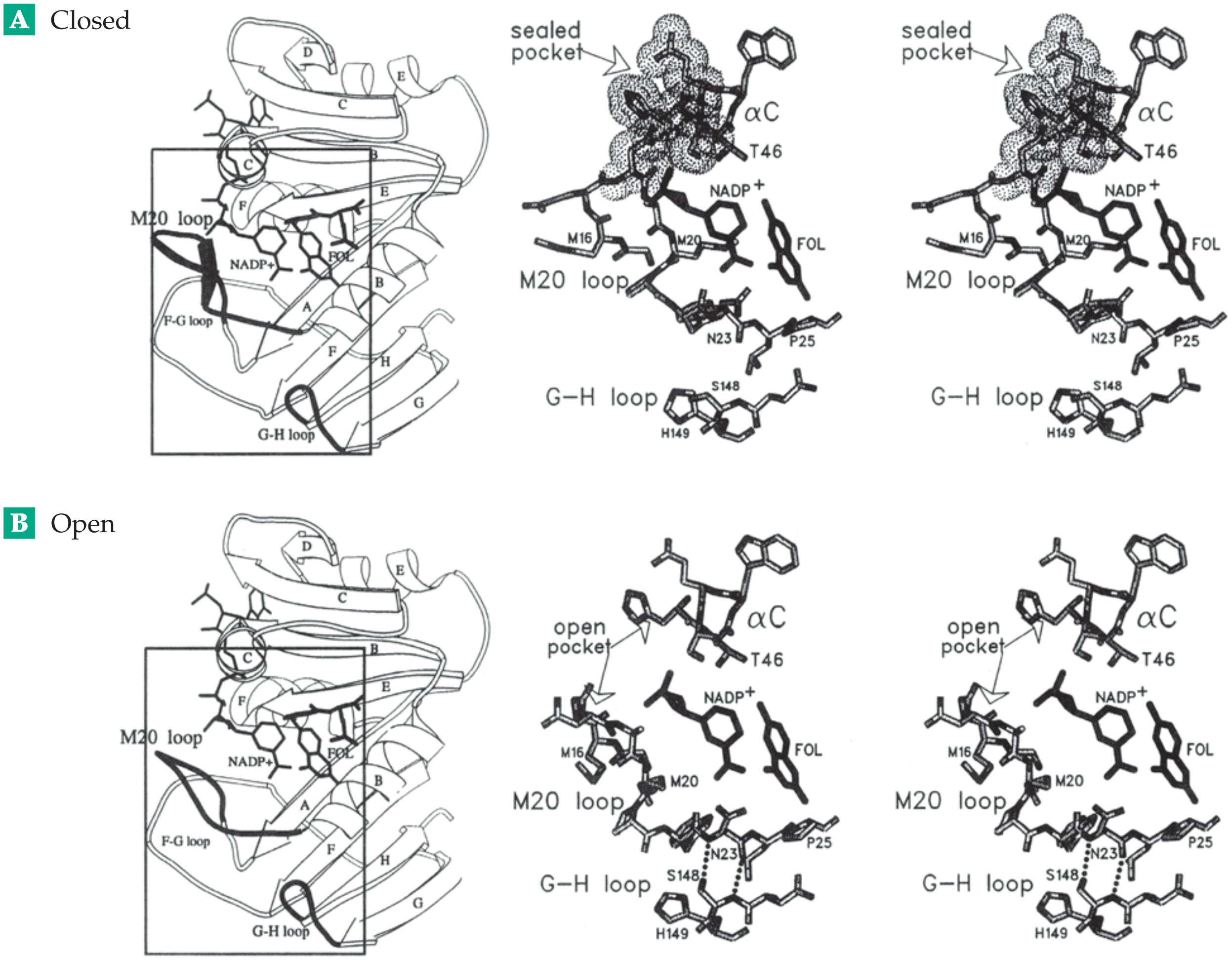
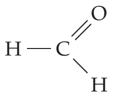
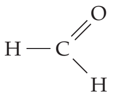
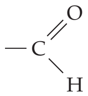
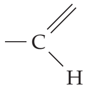
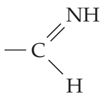

A Closed
D
E
C
C
B
E
M20 loop
F
NADP+
FOL
B
A
F-G loop
F
H
G
G-H loop
sealed pocket
αC
T46
NADP+
M16
M20
FOL
M20 loop
N23
P25
G–H loop
S148
H149
sealed pocket
αC
T46
NADP+
M16
M20
FOL
M20 loop
N23
P25
G–H loop
S148
H149
B Open
D
E
C
C
B
E
F
M20 loop
NADP+
FOL
B
A
F-G loop
F
H
G
G-H loop
open pocket
αC
T46
NADP+
M16
M20
FOL
M20 loop
N23
P25
G–H loop
S148
H149
open pocket
αC
T46
NADP+
M16
M20
FOL
M20 loop
N23
P25
G–H loop
S148
H149

Figure 15-19 Drawings of the active site of *E. coli* dihydrofolate reductase showing the bound ligands $NADP^+$ and tetrahydrofolate. Several key amino acid side chains are shown in the stereoscopic views on the right. The complete ribbon structures are on the left. (A) Closed form. (B) Open form into which substrates can enter and products can escape. From Sawaya and Kraut.[381] Courtesy of Joseph Kraut. Molscript drawings (Kraulis, 1991).

enzyme by Asn reduced the specific catalytic activity to 1/300 that of native enzyme indicating that this residue is important for catalysis.[390] The value of k_{cat} for reduction of dihydrofolate is highest at low pH and varies around a pK_a of ~6.5.[392–394] One interpretation is that this pK_a belongs to the Asp 27 carboxyl group and another is that it belongs to an N-5 protonated species of the coenzyme. As we have seen (Chapter 6) it is often impossible to assign a pK_a to a single group because there may be a mixture of interacting tautomeric species. Despite the enormous amount of study, we still don't know quite how the proton gets to N-5 in this reaction. Only one possibility is illustrated in Eq. 15-42. Asp 27 is protonated, the pterin ring is enolized, and a buried water molecule serves to relay a proton to N-5. The X-ray crystallographic studies on the *E. coli* enzyme show that after the binding of substrates a conformational change closes a lid over the active site (Fig. 15-19). This excludes water from N-5 but it may permit intermittent access, allowing transfer of a proton from a water molecule bound to O-4 as shown in Eq. 15-42.[381] During the reduction of folate to dihydrofolate a different mechanism of proton donation must be followed to allow protonation at C-7.

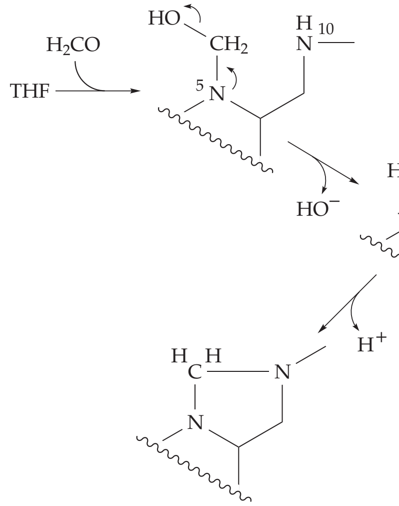
TABLE 15-3
Single-Carbon Compounds in Order of Oxidation State of Carbon

CH_4 Methane	$H_3C—Y$ Methyl groups bound to O, N, S	 Formaldehyde	 Formic acid	CO_2 H_2CO_3
		$—CH_2OH$ Hydroxymethyl group	 Formyl group	
		$—CH_2—$ Methylene group	 Methenyl group	
			 Formimino group	
			$C═O$ Carbon monoxide	

4. Single-Carbon Compounds and Groups in Metabolism

Some single-carbon compounds and groups important in metabolism are shown in Table 15-3 in order (from left to right) of increasing oxidation state of the carbon atom. Groups at three different oxidation levels, corresponding to **formic acid**, **formaldehyde**, and the **methyl group**, are carried by tetrahydrofolic acid coenzymes. While the most completely reduced compound, methane, cannot exist in a combined form, its biosynthesis depends upon reduced methanopterin, as does that of carbon monoxide. Figure 15-18 summarizes the metabolic interrelationships of the compounds and groups in this table.

Serine as a C1 donor. For many organisms, from *E. coli* to higher animals, **serine** is a major precursor of C-1 units.[395–397] The β-carbon of serine is removed as formaldehyde through direct transfer to tetrahydrofolate with formation of methylene-THF and glycine (Eq. 14-30, Fig. 15-18, step *c*). This is a stereospecific transfer in which the *pro-S* hydrogen on C-3 of serine enters the *pro-S* position also in methylene-THF.[398] The glycine formed in this reaction can, in turn, yield another single-carbon unit by loss of CO_2 under the influence of the THF and the PLP-requiring glycine cleavage system which is discussed in the next paragraph. Free formaldehyde in a low concentration can also combine with THF to form methylene-THF (Fig. 15-18, step *c*′ and Eq. 15-43).[399]



(15-43)

The glycine decarboxylase–synthetase system. Glycine is cleaved reversibly within mitochondria of plants and animals and also by bacteria to CO_2, NH_3, and a methylene group which is carried by tetrahydrofolic acid[296,400–403] (Fig. 15-20). Four proteins are required. The P-protein consists of two identical 100-kDa subunits, each containing a molecule of PLP. This protein is a **glycine decarboxylase** which, however, replaces the lost CO_2 by an electrophilic sulfur of lipoate rather than by a proton. Serine hydroxymethyltransferase can also catalyze this step of the sequence.[404] The lipoate is bound to a second protein, the H-protein. A third protein, the T-protein, carries bound tetrahydrofolate which displaces the aminomethyl group from the dihydrolipoate and converts it to N^5,N^{10}-methylene tetrahydrofolate with release of ammonia. The dihydrolipoate is then reoxidized by NAD^+ and dihydrolipamide dehydrogenase (Fig. 15-20). Glycine can be oxidized completely in liver mitochondria with the methylene group of methylene-THF being converted to CO_2 through reaction steps *d*, *e*, and *f* of Fig. 15-18. The glycine cleavage system is reversible and is used by some organisms to synthesize glycine.

Whether it arises from the hydroxymethyl group of serine or from glycine, the single-carbon unit of methylene-THF (which is at the formaldehyde level of oxidation) can either be oxidized further to 5,10-methenyl-THF and 10-formyl-THF (steps *d* and *e*, Fig. 15-18)

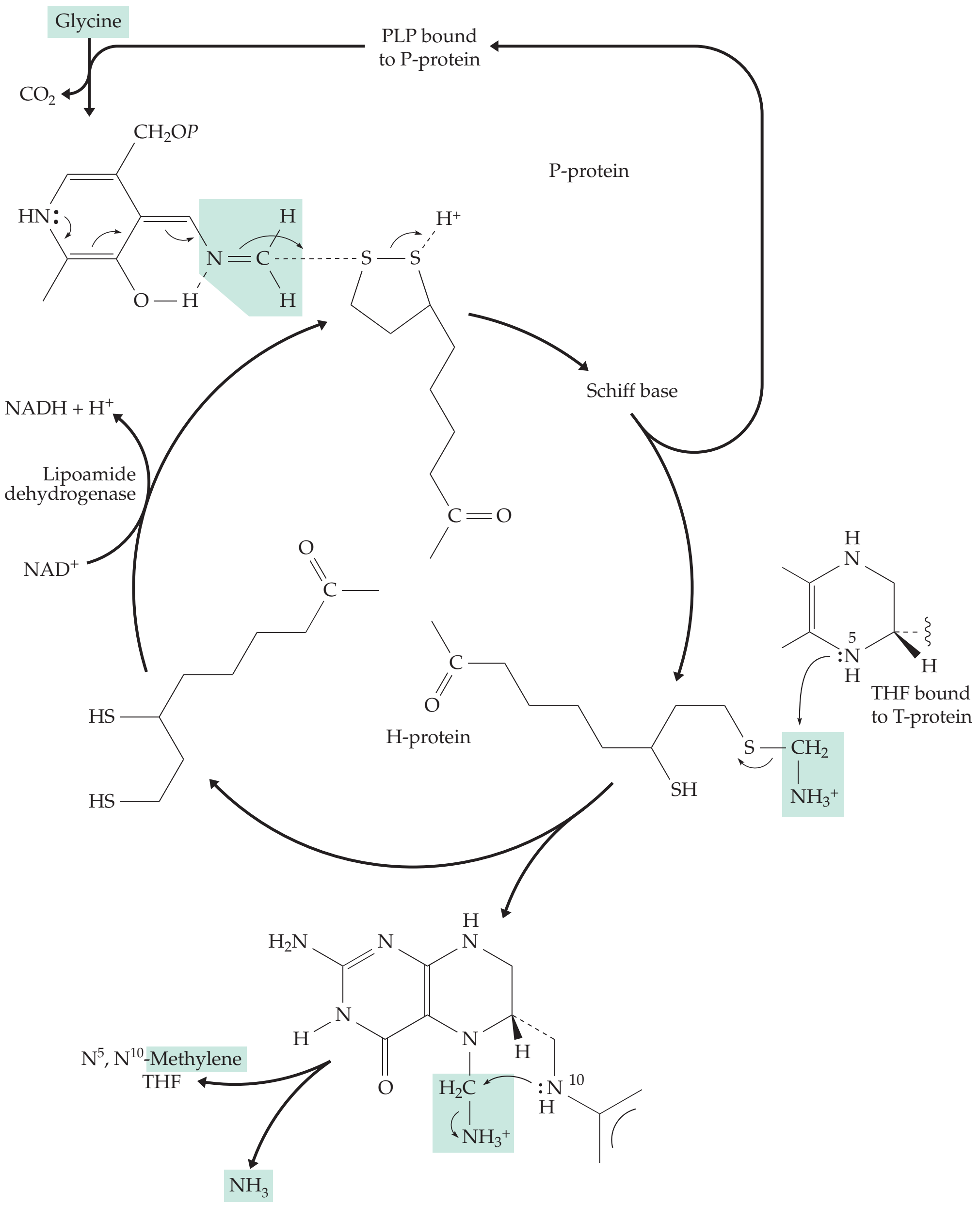

Glycine
PLP bound to P-protein
CO2
CH2OP
HN:
H
H+
N=C
S—S
O—H
H
P-protein
Schiff base
NADH + H+
Lipoamide dehydrogenase
NAD+
C=O
O
C
HS
HS
C
O
H-protein
H
N
5
:N
H
H
THF bound to T-protein
S—CH2
NH3+
SH
H
N
H2N
N
N
H
N
N
H
O
H2C
:N 10
H
NH3+
N5, N10-Methylene THF
NH3

Figure 15-20 The reversible glycine cleavage system of mitochondria. Compare with Fig. 15-15.

or it can be reduced to methyl-THF (step *i*). The reactions of folate metabolism occurs in both cytoplasm and mitochondria at rates that are affected by the length of the polyglutamate chain. The many complexities of these pathways are not fully understood.[395,397]

Starting with formate or carbon dioxide. Most organisms can, to some extent, also utilize formate as a source of single-carbon units. Human beings have a very limited ability to metabolize formate and the accumulation of formic acid in the body following ingestion of methanol is often fatal. However, many bacteria are able to subsist on formate as a sole carbon source. Archaea may generate the formate by reduction of CO_2. Utilization of formate begins with formation of 10-formyl-THF (Fig. 15-18, step *g*, lower left corner). 10-Formyl-THF can be reduced to methylene-THF and the single-carbon unit can be transferred to glycine to form serine. In some bacteria three separate enzymes are required to convert formate to methylene-THF: 10-formyl-THF synthetase (Fig. 15-18, step *g*, presumably via formyl phosphate),[405] methenyl-THF cyclohydrolase (step *e*, reverse), and methylene-THF dehydrogenase (step *d*, reverse). NADH is used by the acetogens but in *E. coli* and in most higher organisms NADPH is the reductant. In some bacteria and in some tumor cells, two of the three enzymes are present as a bifunctional enzyme.[406–407a] In most eukaryotes all three of the enzymatic activities needed for converting formate to methylene-THF are present in a single large (200-kDa) dimeric trifunctional protein called **formyl-THF synthetase**.[407,408]

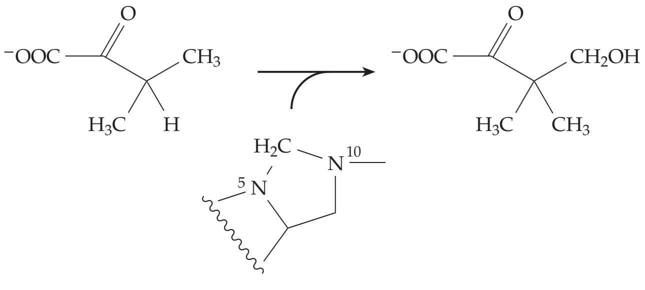
N^{10}-Formyl-THF serves as *a biological formylating agent* needed for two steps in the synthesis of purines[409–410b] (Chapter 25) and, in bacteria and in mitochondria and chloroplasts, for synthesis of formylmethionyl-tRNA[411] which initiates synthesis of all polypeptide chains in bacteria and in these organelles (Chapter 29).

N^5-Formyl-THF (leucovorin). The N^5-formyl derivative of THF (5-formyl-THF) is a growth factor for *Leuconostoc citrovorum.* Following this discovery in 1949 it was called the "citrovorum factor" or **leucovorin**. Its significance in metabolism is not clear. Perhaps it serves as a storage form of folate in cells that have a dormant stage, e.g., of seeds or spores.[412] It may also have a regulatory function.[413] In some ants and in certain beetles, it is stored and hydrolyzed to formic acid. The carabid beetle *Galerita lecontei* ejects a defensive spray that contains 80% formic acid.[414]

5-Formyl-THF can arise by transfer of a formyl group from formylglutamate and there is an enzyme that converts 5-formyl-THF to 5,10-methenyl-THF with concurrent cleavage of ATP.[412] 5-Formyl-THF is sometimes used in a remarkable way to treat certain highly malignant cancers. Following surgical removal of the tumor the patient is periodically given what would normally be a lethal dose of methotrexate, then, about 36 h later, the patient is rescued by injection of 5-formyl-THF. The mechanism of rescue is thought to depend upon the compound's rapid conversion to 10-formyl-THF within cells. Tumor cells are not rescued, perhaps because they have a higher capacity than normal cells for synthesis of polyglutamates of methotrexate.[379] This observation also suggests that the major anti-cancer effect of methotrexate may not be on dihydrofolate reductase but on enzymes of formyl group transfer that are inhibited by polyglutamate derivatives of methotrexate.

Catabolism of histidine and purines. Another source of single-carbon units in metabolism is the degradation of histidine which occurs both in bacteria and in animals via **formiminoglutamate**. The latter transfers the $-CH=NH$ group to THF forming 5-formimino-THF, which is in turn converted (step *h*, Fig. 15-18) to 5,10-methenyl-THF and ammonia. In bacteria that ferment purines, **formiminoglycine** is an intermediate. Again, the formimino group is transferred to THF and deaminated, the eventual product being 10-formyl-THF. In these organisms, the enzyme 10-formyl-THF synthase also has a very high activity.[415] It probably operates in reverse as a mechanism for synthesis of ATP in this type of fermentation. In other organisms excess 10-formyl-THF may be oxidized to CO_2 via an $NADP^+$-dependent dehydrogenase (step *f*, Fig. 15-16), providing a mechanism for detoxifying formic acid. In some organisms excess 10-formyl-THF may simply be hydrolyzed with release of formate (step *q*).[415a]

Thymidylate synthase. Methylene-THF serves as the direct precursor of the 5-methyl group of thymine as well as of the hydroxymethyl groups of **hydroxymethylcytosine**[416] and of **2-oxopantoate**, an intermediate in the formation of pantothenate and coenzyme A.[417] The latter is a simple hydroxymethyl transfer reaction (Eq. 15-44) that is related to an aldol condensation and which may proceed through an imine of the kind shown in Eq. 15-43.



$N^{5,10}$-Methylene-THF (15-44)

During thymine formation the coenzyme is oxidized to dihydrofolate, which must be reduced by dihydrofolate reductase to complete the catalytic cycle. A possible mechanistic sequence for **thymidylate synthase**, an enzyme of known three-dimensional structure,[354,418–421a] is given in Fig. 15-21. In the first step (*a*) a thiolate anion, from the side chain of Cys 198 of the 316-residue *Lactobacillus* enzyme, adds to the 5 position of the substrate 2'-deoxyuridine monophosphate (dUMP). As a consequence the 6 position becomes a nucleophilic center which can combine with methylene-THF as shown in step *b* of Fig. 15-21. After tautomerization (step *c*) this adduct eliminates THF (step *d*) to give a 5-methylene derivative of the dUMP. The latter immediately oxidizes the THF by a hydride ion transfer to form thymidylate and dihydrofolate.[422,422a]

In protozoa thymidylate synthase and dihydrofolate reductase exist as a single bifunctional protein.

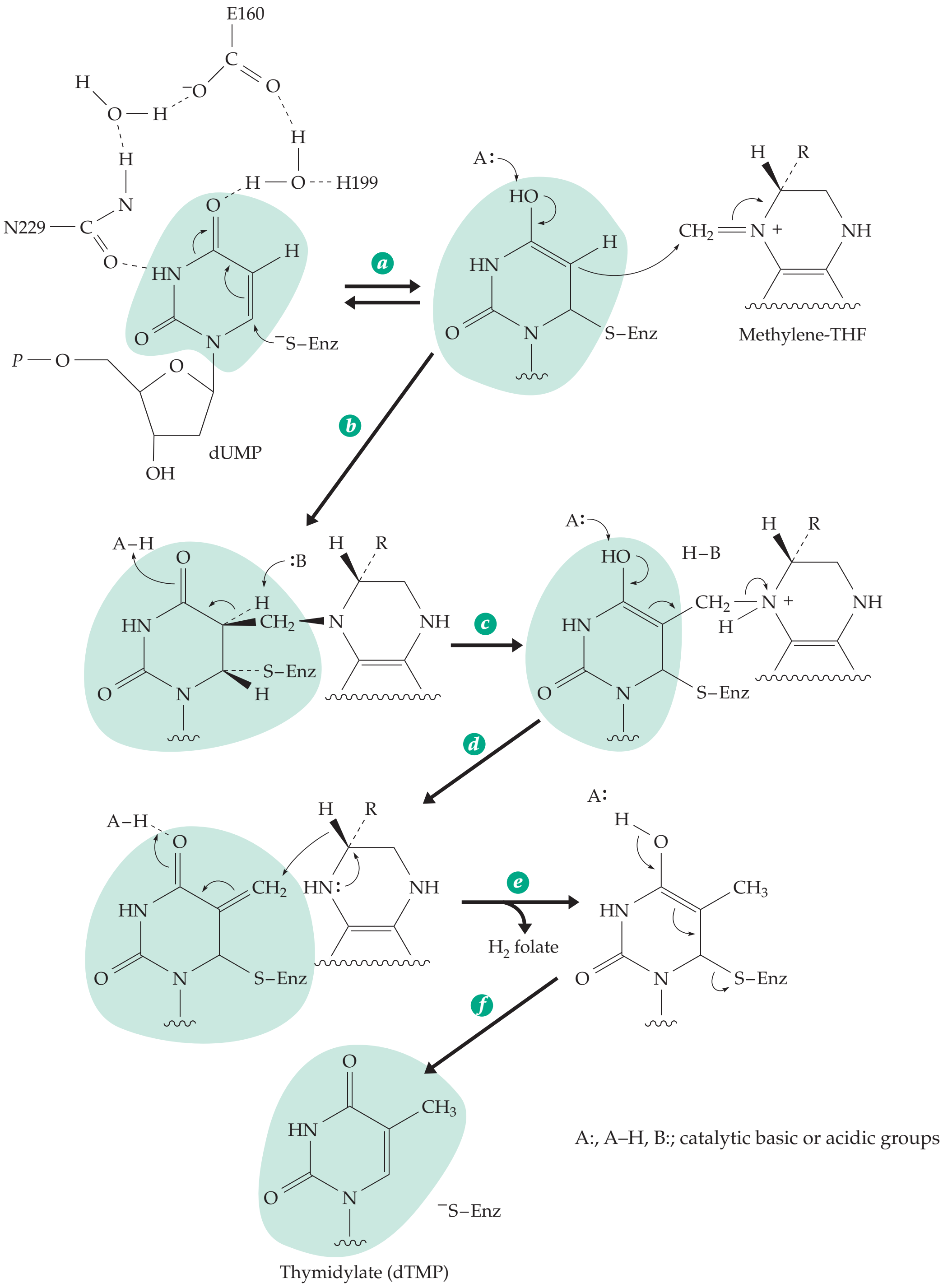
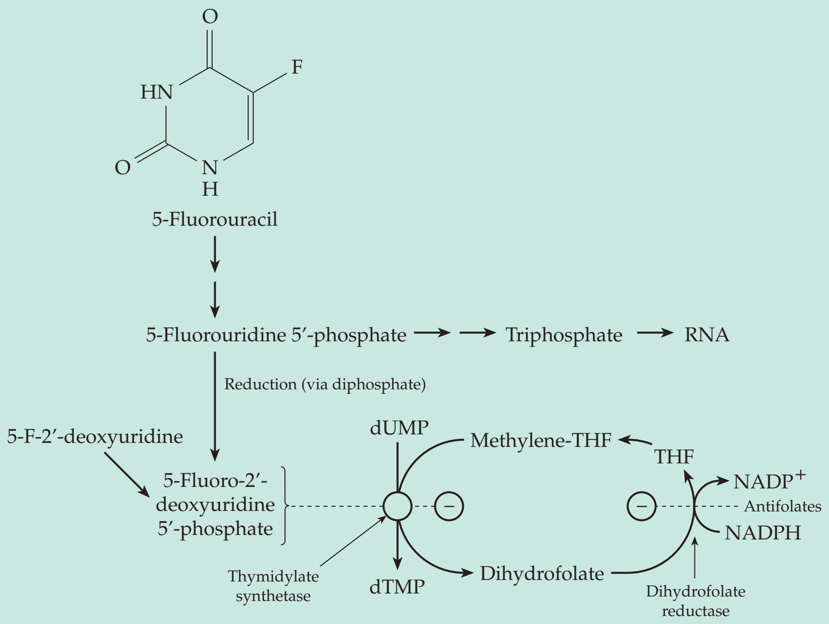


Figure 15-21 Probable mechanism of action of thymidylate synthase. After Huang and Santi[422] and Hyatt *et al.*[354]

Consequently, in methotrexate-resistant strains of *Leishmania* (and also in cancer cells) both enzyme activities are increased equally by gene amplification.[423] This presents a serious problem in the treatment of protozoal parasitic diseases for which few suitable drugs are known.

Cells of *E. coli*, when infected with T-even bacteriophage, convert dCMP to 5-hydroxymethyl-dCMP[423a]

BOX 15-E THYMIDYLATE SYNTHASE, A TARGET ENZYME IN CANCER CHEMOTHERAPY[a]



If an animal or bacterial cells are deprived of thymine they can no longer make DNA. However, synthesis of proteins and of RNA continues for some time. This can be demonstrated experimentally with thymine-requiring mutants. However, sooner or later such cells lose their vitality and die. The cause of this **thymineless death**[b] is not entirely clear. Perhaps thymine is needed to repair damage to DNA and if it is not available transcription eventually becomes faulty. Chromosome breakage is also observed.[c] Whatever the cause of death, the phenomenon provides the basis for some of the most effective chemotherapeutic attacks on cancer. Rapidly metabolizing cancer cells are especially vulnerable to thymineless death. Consequently, thymidylate synthase is an important target enzyme for inhibition. One powerful inhibitor is the monophosphate of **5-fluoro-2'-deoxyuridine**. The inhibition was originally discovered when 5-fluorouracil was recognized as a useful cancer chemotherapeutic agent.

Fluorouracil has many effects in cells, including incorporation into RNA,[b] but the inhibition of thymidylate synthase by the reduction product may be the most useful effect in chemotherapy. In fact, 5-fluoro-2'-deoxyuridine is a much less toxic and more potent drug than 5-fluorouracil. It binds into the active site of thymidylate synthase and reacts in the initial steps of catalysis. However, the 5-H of 2'-deoxyuridine, the normal substrate, must be removed as H^+ (step *c* of Fig. 15-21) in order for the reaction to continue. The 5-F atom cannot be removed in this way and a stable adduct is formed.[d,e] A crystal structure containing both the bound 5-fluorodeoxyuridine and methylene-THF supports this conclusion.[e]

Thymidylate synthase requires methylene tetrahydrofolate as a reductant and the reduction of dihydrofolate is also an important part of the process. In protozoa dihydrofolate reductase and thymidylate synthase occur as a single-chain bifunctional enzyme.[f] As has been pointed out in the main text, such folic acid analogs as methotrexate are among the most useful anticancer drugs. By inhibiting dihydrofolate reductase they deprive thymidylate synthase of an essential substrate.

Because 5-fluorouracil acts on normal cells as well as cancer cells, its usefulness is limited. Knowledge of the chemistry and three-dimensional structure of thymidylate synthase complexes is being used in an attempt to discover more specific and effective drugs that attack this enzyme.[g,h]

---

[a] Friedkin, M. (1973) *Adv. Enzymol.* **38**, 235–292

[b] Sahasrabudhe, P. V., and Gmeiner, W. H. (1997) *Biochemistry* **36**, 5981–5991

[c] Ayusawa, D., Shimizu, K., Koyama, H., Takeishi, K., and Seno, T. (1983) *J. Biol. Chem.* **258**, 12448–12454

[d] Huang, X. F., and Arvan, P. (1995) *J. Biol. Chem.* **270**, 20417–20423

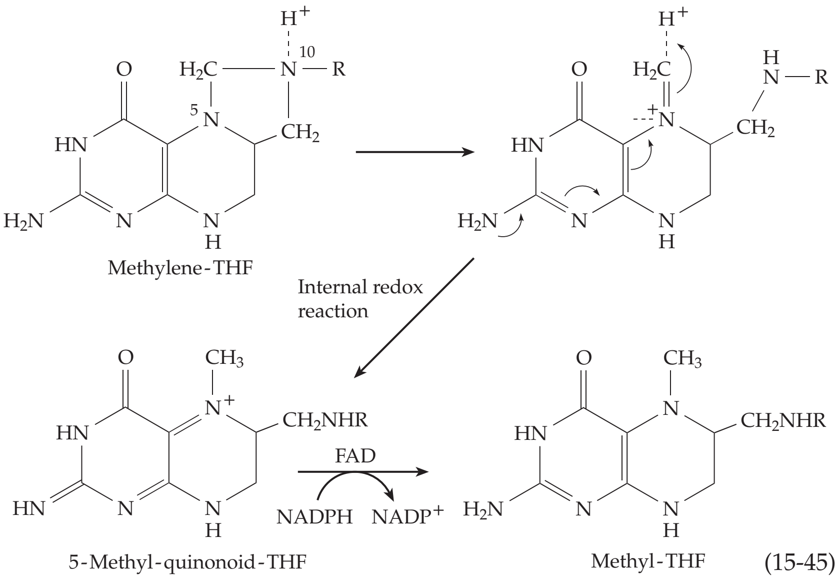
[e] Hyatt, D. C., Maley, F., and Montfort, W. R. (1997) *Biochemistry* **36**, 4585–4594

[f] Ivanetich, K. M., and Santi, D. V. (1990) *FASEB J.* **4**, 1591–1597

[g] Shoichet, B. K., Stroud, R. M., Santi, D. V., Kuntz, I. D., and Perry, K. M. (1993) *Science* **259**, 1445–1449

[h] Weichsel, A., Montfort, W. R., Ciesla, J., and Maley, F. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3493–3497

and suitably infected cells of *Bacillus subtilis* form 5-hydroxymethyl-dUMP.[424,425] These products can be formed by attack of OH^- on quinonoid intermediates of the type postulated for thymidylate synthetase (formed in step *d* of Fig. 15-21).[416] A related reaction is the posttranscriptional conversion of a single uracil residue in tRNA molecules of some bacteria to a thymine ring (a **ribothymidylic acid** residue). In this case, $FADH_2$ is used as a reducing agent to convert the initial adduct of Fig. 15-21 to the ribothymidylic acid and THF.[426]



(15-45)

Synthesis of methyl groups. The reduction of methylene-THF to 5-methyl-THF within all living organisms from bacteria to higher plants and animals provides the methyl groups needed in biosynthesis.[426a] These are required for formation of methionine and, from it, *S*-adenosylmethionine. The latter is used to modify proteins, nucleic acids, and other biochemicals through methylation of specific groups. In the methanogens, reduction of a corresponding methylene derivative of methanopterin gives rise to methane (Section F).

Mammalian methylene-THF reductase is a FAD-containing flavoprotein that utilizes NADPH for the reduction to 5-methyl-THF.[427,428] Matthews[429] suggested that the mechanism of this reaction involves an internal oxidation–reduction reaction that generates a 5-methyl-quinonoid dihydro-THF (Eq. 15-45). Methylene-THF reductase of acetogenic bacteria is also a flavoprotein but it contains Fe–S centers as well. The 237-kDa $\alpha_4\beta_4$ oligomer contains two molecules of FAD and four to six of both Fe and S^{2-} ions.[430,431]

The methyl group of methyl-THF is incorporated into methionine by the vitamin B_{12}-dependent methionine synthase which is discussed in Chapter 16. Matthews suggested that methionine synthase may also make use of the 5-methyl-quinonoid-THF of Eq. 15-45. An initial reduction step would precede transfer of the methyl group. Methyl-THF is a precursor to acetate in the acetogenic bacteria.[432] This corrinoid coenzyme-dependent process is also considered in Chapter 16.

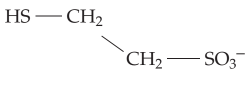
E. Specialized Coenzymes of Methanogenic Bacteria

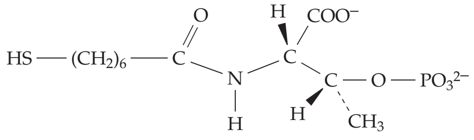
Methane-producing bacteria[433,434] obtain energy by reducing CO_2 with molecular hydrogen:

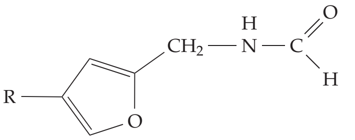
$$\begin{gathered} CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O \\ \Delta G^\circ = -131 \text{ kJ / mol} \end{gathered} \tag{15-46}$$

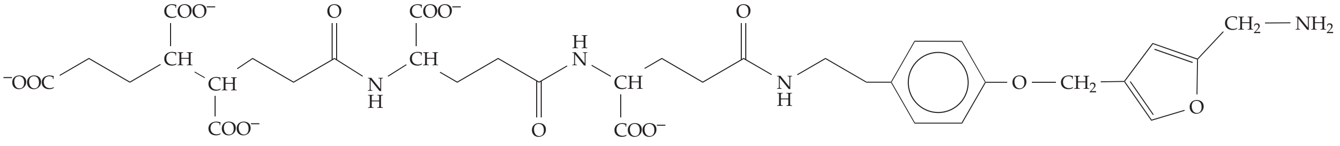
Some species are also able to utilize formate or formaldehyde as reducing agents.[435] These compounds are oxidized to CO_2, the reducing equivalents formed being used to reduce CO_2 to methane. Carbon monoxide can also be converted to CO_2.

$$\begin{gathered} CO + H_2O \rightarrow CO_2 + H_2 \\ \Delta G^\circ = -20 \text{ kJ / mol} \end{gathered} \tag{15-47}$$


Coenzyme M


7-Mercaptoheptanoylthreonine phosphate


Formylmethanofuran


Methanofuran

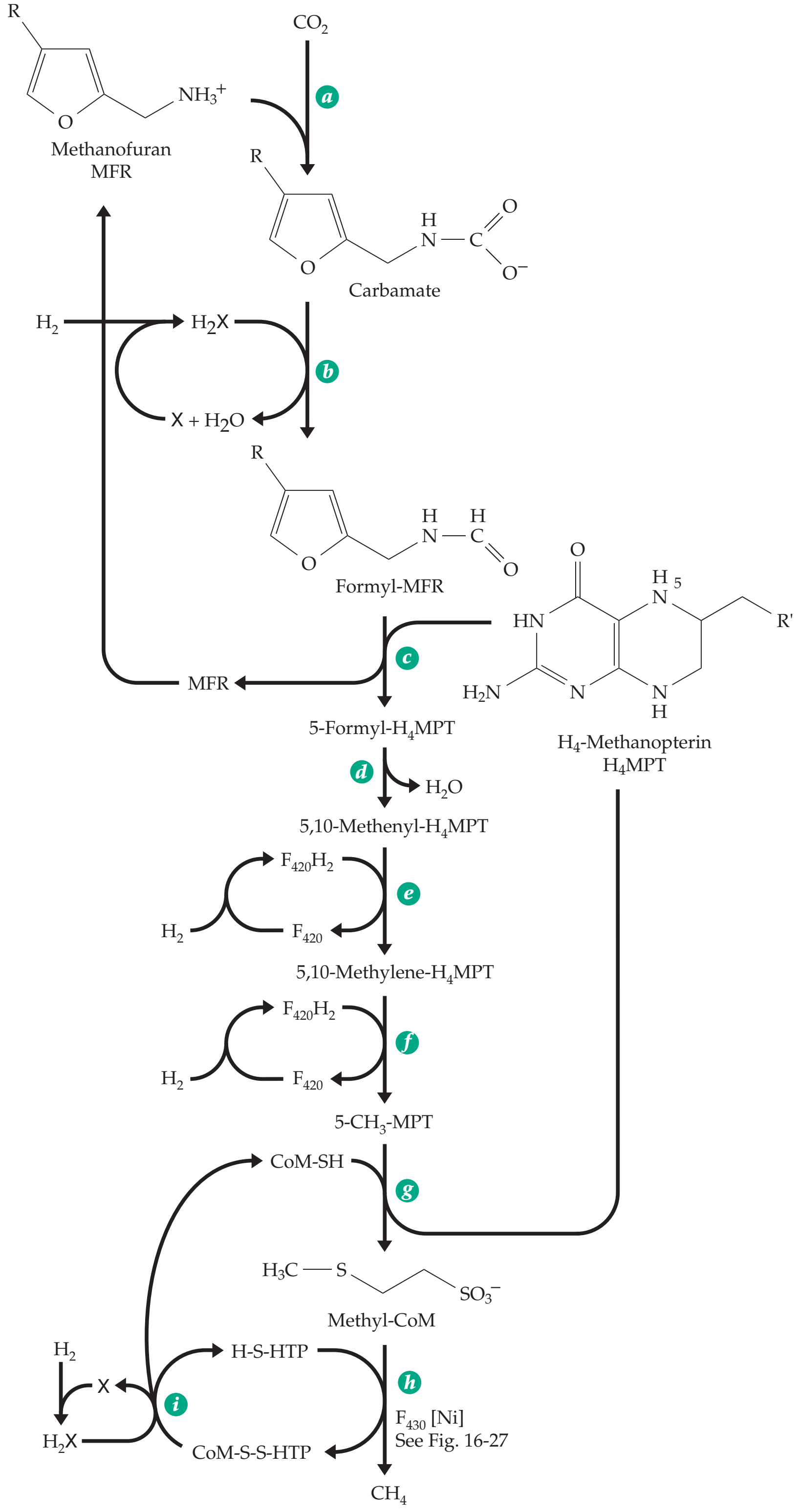

R
CO$_2$
NH$_3^+$
O
a
Methanofuran
MFR
R
H
N—C
O
O$^-$
O
Carbamate
H$_2$
H$_2$X
b
X + H$_2$O
R
H H
N—C
O
O
Formyl-MFR
O
H 5
N
R'
HN
c
MFR
H$_2$N
N
N
H
5-Formyl-H$_4$MPT
H$_4$-Methanopterin
H$_4$MPT
d
H$_2$O
5,10-Methenyl-H$_4$MPT
F$_{420}$H$_2$
e
H$_2$
F$_{420}$
5,10-Methylene-H$_4$MPT
F$_{420}$H$_2$
f
H$_2$
F$_{420}$
5-CH$_3$-MPT
CoM-SH
g
H$_3$C—S
SO$_3^-$
Methyl-CoM
H$_2$
H-S-HTP
h
X
i
F$_{430}$ [Ni]
See Fig. 16-27
H$_2$X
CoM-S-S-HTP
CH$_4$

Figure 15-22 Tentative scheme for reduction of carbon dioxide to methane by methanogens. After Rouvière *et al.*[352] and Thauer *et al.*[435]

By using the combined reactions of Eqs. 15-46 and 15-47 the bacteria can subsist on CO alone (Eq. 15-48):

$$4CO + 2\,H_2O \rightarrow 3CO_2 + CH_4$$
$$\Delta G^{\circ} = -211 \text{ kJ / mol of methane formed} \qquad (15\text{-}48)$$

Some species reduce methanol to methane via Eqs. 15-49 or 15-50.

$$CH_3OH + H_2 \rightarrow CH_4 + H_2O$$
$$\Delta G^{\circ} = -113 \text{ kJ / mol} \qquad (15\text{-}49)$$

$$4CH_3OH \rightarrow 3CH_4 + CO_2 + 2\,H_2O$$
$$\Delta G^{\circ} = -107 \text{ kJ / mol of methane formed} \qquad (15\text{-}50)$$

To accomplish these reactions a surprising variety of specialized cofactors are needed.[351,352,434] The first of these, **coenzyme M**, 2-mercaptoethane sulfonate, was discovered in 1974.[436] It is the simplest known coenzyme. Later, the previously described **5-deazaflavin F_{420}** (Section B), a **nickel tetrapyrrole F_{430}** (Chapter 16), **methanopterin** (Fig. 15-17),[437] the "carbon dioxide reduction factor" called **methanofuran**,[352,438] and **7-mercaptoheptanoylthreonine phosphate**[439,440] were also identified.

A sketch of the metabolic pathways followed in methane formation is given in Fig. 15-22.[352,435] In the first step (*a*) the amino group of methanofuran is thought to add to CO_2 to form a carbamate which is reduced to formylmethanofuran by H_2 and an intermediate carrier H_2X in step *b*. The formyl group is then transferred to tetrahydromethanopterin (H_4MPT) (step *c*)[440a] and is cyclized and reduced in two stages in steps *d*, *e*, and *f*. The reductant is the deazaflavin F_{420} and the reactions parallel those for conversion of formyl-THF to methyl-THF (Fig. 15-18).[431,440b,441] The methyl group of methyl-H_4MPT is then transferred to the sulfur atom of the thiolate anion of coenzyme M, from which it is reduced off as CH_4. This is a complex process requiring the nickel-containing F_{430}, FAD, and 7-mercaptoheptanoylthreonine phosphate (HS-HTP).

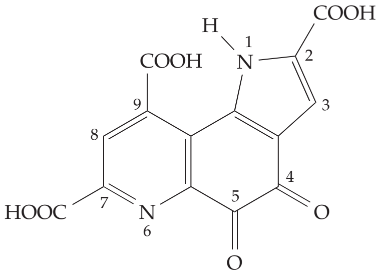
The HS-HTP may be the 2-electron donor for the reduction. The mixed-disulfide CoM-S-S-HTP appears to be an intermediate and also an allosteric effector for the first step, the reduction of CO_2 to formylmethanofuran. A complex of proteins that is unstable in oxygen is also needed. The principal component of the methyl reductase binds two molecules of the nickel-containing F_{430}. It also binds two moles of CoM-SH noncovalently. The binding process requires ATP as well as the presence of other proteins, notably those of the hydrogenase system.

Coenzyme M, previously found only in methanogenic archaeobacteria, has recently been discovered in both gram-negative and gram-positive alkene-oxidizing eubacteria. It seems to function in cleavage of epoxy rings and as a carrier of hydroxyalkyl groups.[441a] See also Chapter 17.

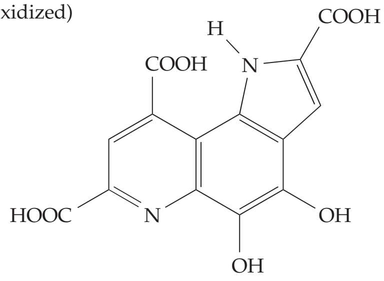
F. Quinones, Hydroquinones, and Tocopherols

1. Pyrroloquinoline Quinone (PQQ)

Bacteria that oxidize methane or methanol (**methylotrophs**) employ a periplasmic methanol dehydrogenase that contains as a bound coenzyme, the pyrroloquinoline quinone designated **PQQ** or methoxatin (Eq. 15-51).[442–444] This fluorescent *ortho*-quinone is released when the protein is denatured. Some other bacterial alcohol dehydrogenases[445,445a] and glucose dehydrogenases[446-446d] contain the same cofactor.



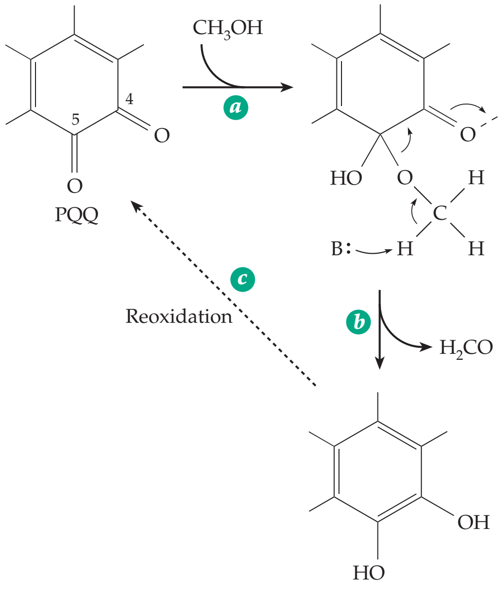
PQQ (oxidized)



$PQQH_2$ (reduced)

PQQ-containing methanol dehydrogenase from *Methylophilus* is a dimer of a 640-residue protein consisting of two disulfide-linked peptide chains with one noncovalently bound PQQ.[447] The large subunit contains a β propellor (Fig. 15-23), which is similar to that in the protein G_i shown in Fig. 11-7. The PQQ binding site lies above the propellor and is formed by a series of loops. The coenzyme interacts with several polar protein side chains and with a bound calcium ion.[448] Bacterial PQQ-dependent glucose dehydrogenases have also been studied intensively. The 450-residue soluble enzyme from *Acinetobacter calcoaceticus* is partially homologous to the methanol dehydrogenase.[446a,446b]

PQQ and the other quinone prosthetic groups described here all function in reactions that would be possible for pyridine nucleotide or flavin coenzymes. All of them, like the flavins, can exist in oxidized, half-reduced semiquinone and fully reduced dihydro forms. The questions to be asked are the same as we asked for flavins. How do the substrates react? How is the reduced cofactor reoxidized? In nonenzymatic reactions alcohols, amines, and enolate anions all add at C-5 of PQQ to give adducts such as that shown for methanol in Eq. 15-51, step *a*.[444,449,449a] Although many additional reactions are possible, this addition is a reasonable first step in the mechanism shown in Eq. 15-51. An enzymatic base could remove a proton as is indicated in step *b* to give $PQQH_2$. The pathway for reoxidation (step *c*) might involve a cytochrome *b*, cytochrome *c*, or bound ubiquinone.[445,446]



(15-51)

Although the soluble PQQ-dependent glucose dehydrogenase forms, with methylhydrazine, an adduct similar to that depicted in Eq. 15-51,[449b] the structure of a glucose complex of the $PPQH_2$- and Ca^{2+}-containing enzyme at a resolution of 0.2 nm suggests a direct hydride ion transfer. The only base close to glucose C1 is His 144. The C1-H proton lies directly above the $PQQH_2$ C5 atom. Oubrie *et al.*[446a] propose a deprotonation of the C1-OH by His 144 and H^- transfer to C5 of PQQ followed by tautomerization (Eq. 15-52). Theoretical calculations by Zheng and Bruice[449c] favor a simpler mechanism with hydride transfer to the oxygen atom of the C4 carbonyl (green arrows). In either case, the Ca^{2+} would assist by polarizing the C5 carbonyl group.

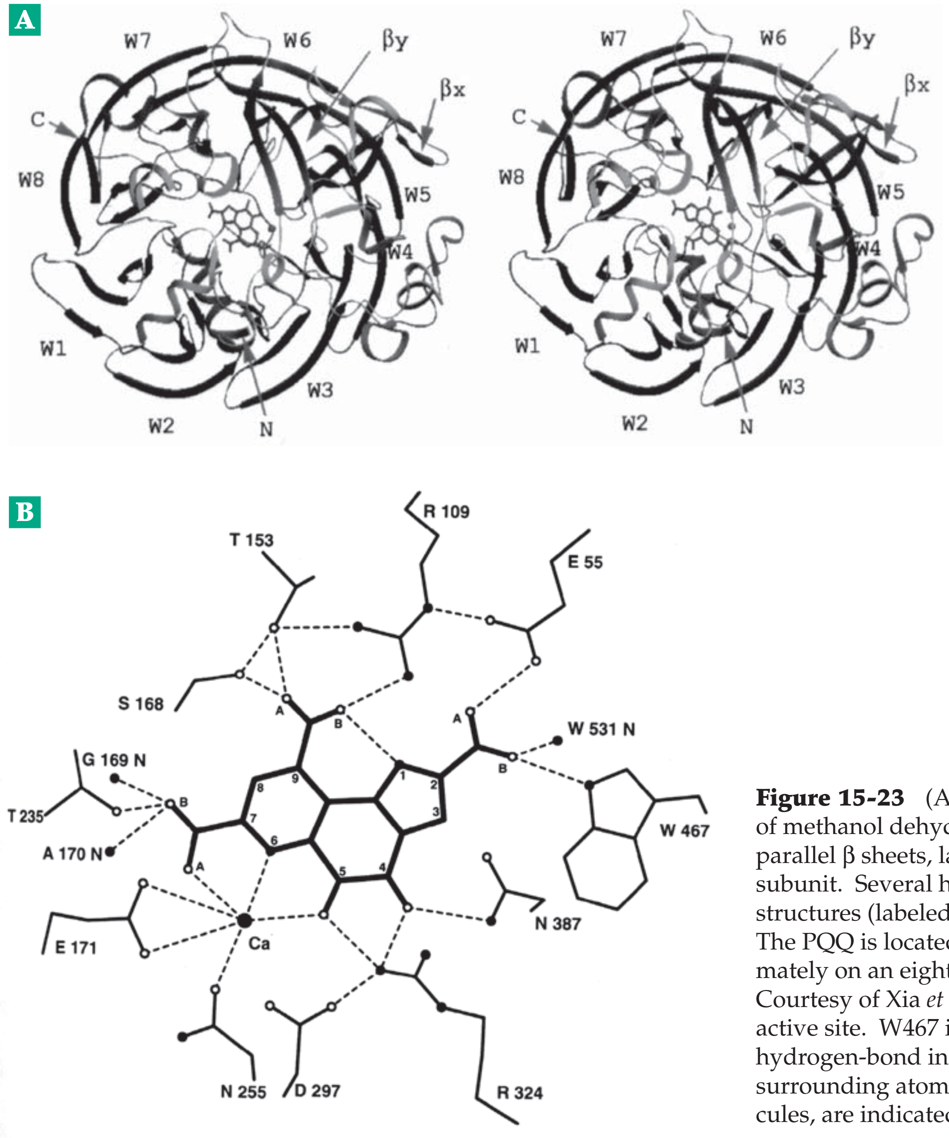
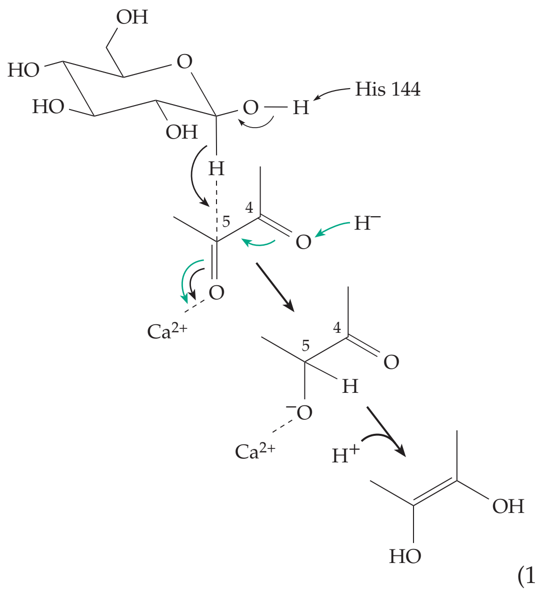


Figure 15-23 (A) Stereoscopic view of the H subunit of methanol dehydrogenase. Eight four-stranded antiparallel β sheets, labeled W1–W8, form the base of the subunit. Several helices and two additional β-sheet structures (labeled βx and βy) form a cap over the base. The PQQ is located in a funnel within the cap approximately on an eight-fold axis of pseudosymmetry. Courtesy of Xia *et al.*[447] (B) Schematic view of the active site. W467 is parallel to the plane of PQQ. All hydrogen-bond interactions between PQQ and its surrounding atoms, except for the three water molecules, are indicated. Courtesy of White *et al.*[448]

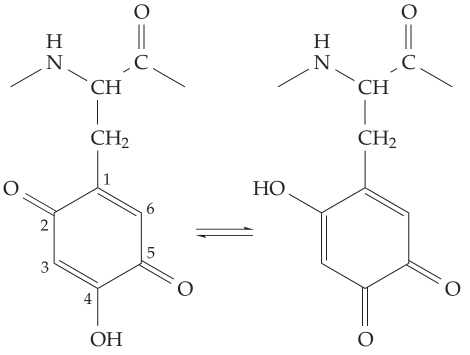


(15-52)

2. Copper Amine Oxidases

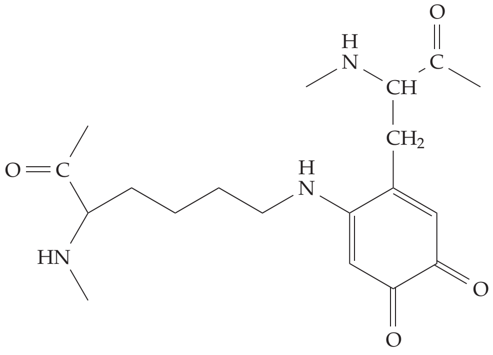
At first it appeared that PQQ had a broad distribution in enzymes, including eukaryotic amine oxidases. However, it was discovered, after considerable effort, that there are additional quinone cofactors that function in oxidation of amines. These are derivatives of tyrosyl groups of specific enzyme proteins. Together with enzymes containing bound PQQ they are often called **quinoproteins**.[450–454]

Topaquinone (TPQ). Both bacteria and eukaryotes contain amine oxidases that utilize bound copper ions and O_2 as electron acceptors and form an aldehyde, NH_3, and H_2O_2. The presence of an organic cofactor was suggested by the absorption spectra which was variously attributed to pyridoxal phosphate or PQQ. However, isolation from the active site of bovine serum amine oxidase established the structure of the reduced form of the cofactor as a trihydroxyphenylalanyl group covalently linked in the peptide chain.[455,456] Its oxidized form is called topaquinone:



Topaquinone (TPQ)

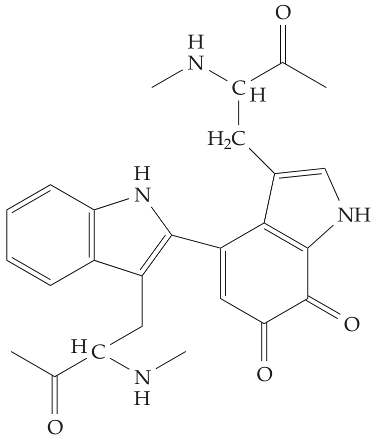
The same cofactor is present in an amine oxidase from *E. coli*[457–460] and in other bacterial[461–463] fungal,[463a] plant,[464] and mammalian[464a] amine oxidases.



Lysine tyrosylquinone (LTQ)

Lysine tyrosylquinone (LTQ). Another copper amine oxidase, **lysyl oxidase**, which oxidizes side chains of lysine in collagen and elastin (Eq. 8-8) contains a cofactor that has been identified as having a lysyl group of a different segment of the protein in place of the –OH in the 2 position of topaquinone.[465] Lysyl oxidase plays an essential role in the crosslinking of collagen and elastin.

Tryptophan tryptophanylquinone (TTQ). This recently discovered quinone cofactor is similar to the lysyl tyrosylquinone but is formed from two tryptophanyl side chains.[466] It has been found in **methylamine dehydrogenase** from methylotrophic gram-negative bacteria[467–469] and also in a bacterial aromatic amine dehydrogenase.[470]

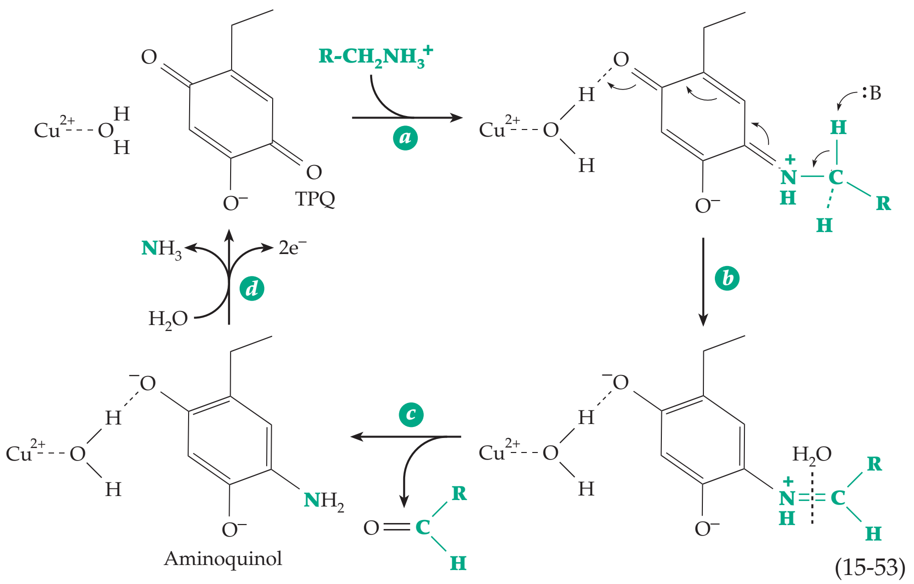


Tryptophan tryptophylquinone (TTQ)

Three-dimensional structures. The TPQ-containing amine oxidase from *E. coli* is a dimer of 727-residue subunits with one molecule of TPQ at position 402 in each subunit.[457,458] Methylamine dehydrogenase is also a large dimeric protein of two large 46.7-kDa subunits and two small 15.5-kDa subunits. Each large subunit contains a TTQ cofactor group.[468,471,471a] Reduced TTQ is reoxidized by the 12.5-kDa blue copper protein **amicyanin**. Crystal structures have been determined for complexes of methylamine dehydrogenase with amicyanin[471] and of these two proteins with a third protein, a small bacterial cytochrome *c*.[472,472a]

Mechanisms. Studies of model reactions[473–476] and of electronic, Raman,[466,477,478] ESR,[479,480] and NMR spectra and kinetics[481] have contributed to an understanding of these enzymes.[459,461,464,482,483] For these copper amine oxidases the experimental evidence suggests an aminotransferase mechanism.[450,453,474,474a-d] The structure of the *E.coli* oxidase shows that a single copper ion is bound by three histidine imidazoles and is located adjacent to the TPQ (Eq. 15-53). Asp 383 is a conserved residue that may be the catalytic base in Eq. 15-53.[474b] A similar mechanism can be invoked for LTQ and TTQ.

How is the reduced cofactor reoxidized? Presumably the copper ion adjacent to the TPQ functions in this process, passing electrons one at a time to the next carrier in a chain. There is no copper in the TTQ-containing subunits. Electrons apparently must jump about 1.6 nm to the copper ion of amicyanin, then another 2.5 nm to the iron ion of the cytochrome *c*.[472] Reoxidation of the aminoquinol formed in Eq. 15-53, step *d*, yields a Schiff base whose hydrolysis will release ammonia and regenerate the TTQ. Intermediate states with Cu$^+$ and a TTQ semiquinone radical have been observed.[483a]



(15-53)

3. Ubiquinones, Plastoquinones, Tocopherols, and Vitamin K

In 1955, R. A. Morton and associates in Liverpool announced the isolation of a quinone which they named ubiquinone for its ubiquitous occurrence.[484,485] It was characterized as a derivative of benzoquinone attached to an unsaturated polyprenyl (isoprenoid) side chain (Fig. 15-24). In fact, there is a family of ubiquinones: that from bacteria typically contains six prenyl units in its side chain, while most ubiquinones from mammalian mitochondria contain ten. Ubiquinone was also isolated by F. L. Crane and associates using isooctane extraction of mitochondria. These workers proposed that the new quinone, which they called **coenzyme Q**, might participate in electron transport. As is described in Chapter 18, this function has been fully established. Both the name ubiquinone and the abbreviation Q are in general use. A subscript indicates the number of prenyl units, e.g., Q_{10}. Ubiquinones can be reversibly reduced to the hydroquinone forms (Fig. 15-24), providing a basis for their function in electron transport within mitochondria and chloroplasts.[486–490]

While vitamin E (Box 15-G) and vitamin K (Box 15-F) are dietary essentials for humans, ubiquinone is apparently not. Animals are able to make ubiquinones in quantities adequate to meet the need for this essential component of mitochondria. However, an extra dietary supplement may sometimes be of value.[491] A reduced level of ubiquinone has been reported in gum tissues of patients with periodontal disease.[492] Ubiquinones are being tested as possible protectants against heart damage caused by lack of adequate oxygen or by drugs.[491]

A closely related series of **plastoquinones** occur in chloroplasts (Chapter 23). In these compounds the two methoxyl groups of ubiquinone are replaced by methyl groups (Fig. 15-24). The most abundant of these compounds, plastoquinone A, contains nine prenyl units.[493,494]

Addition of the hydroxyl group of the reduced ubiquinone or plastoquinone to the adjacent double bond leads to a chroman-6-ol structure. The compounds derived from ubiquinone in this way are called **ubichromanols** (Fig. 15-24). The corresponding ubichromenol (Fig. 15-24) has been isolated from human kidney. **Plastochromanols** are derived from plastoquinone. That from plastoquinone A, first isolated from tobacco, is also known as solanochromene.

A closely related and important family of chromanols are the **tocopherols** or vitamins E (Fig. 15-24, Box 15-G). Tocopherols are plant products found primarily in plant oils and are essential to proper nutrition of humans and other animals. α-Tocopherol is the most abundant form of the vitamin E family; smaller amounts of the β, δ, and γ forms occur, as do a series of **tocotrienols** which contain unsaturated isoprenoid units.[495] The configuration of α-tocopherol is 2*R*,4'*R*,8'*R* as indicated in Fig. 15-24. When α-tocopherol is oxidized, e.g., with ferric chloride, the ring can be opened by hydrolysis to give **tocopherolquinones** (Fig. 15-24), which can in turn be reduced to tocopherolhydroquinones. Large amounts of the tocopherolquinones have been found in chloroplasts.

Another important family of quinones, related in structure to those already discussed, are the **vitamins K** (Fig. 15-24, Box 15-F). These occur naturally as two families. The vitamins K_1 (**phylloquinones**) have only one double bond in the side chain and that is in the prenyl unit closest to the ring. This suggests again the possibility of chromanol formation. In the vitamin K_2 (**menaquinone**) series, a double bond is present in each of the prenyl units. A synthetic compound **menadione** completely lacks the polyprenyl side chain and bears a hydrogen in the corresponding position on the ring. Nevertheless, menadione serves as a synthetic vitamin K, apparently because it can be converted in the body to forms containing polyprenyl side chains.

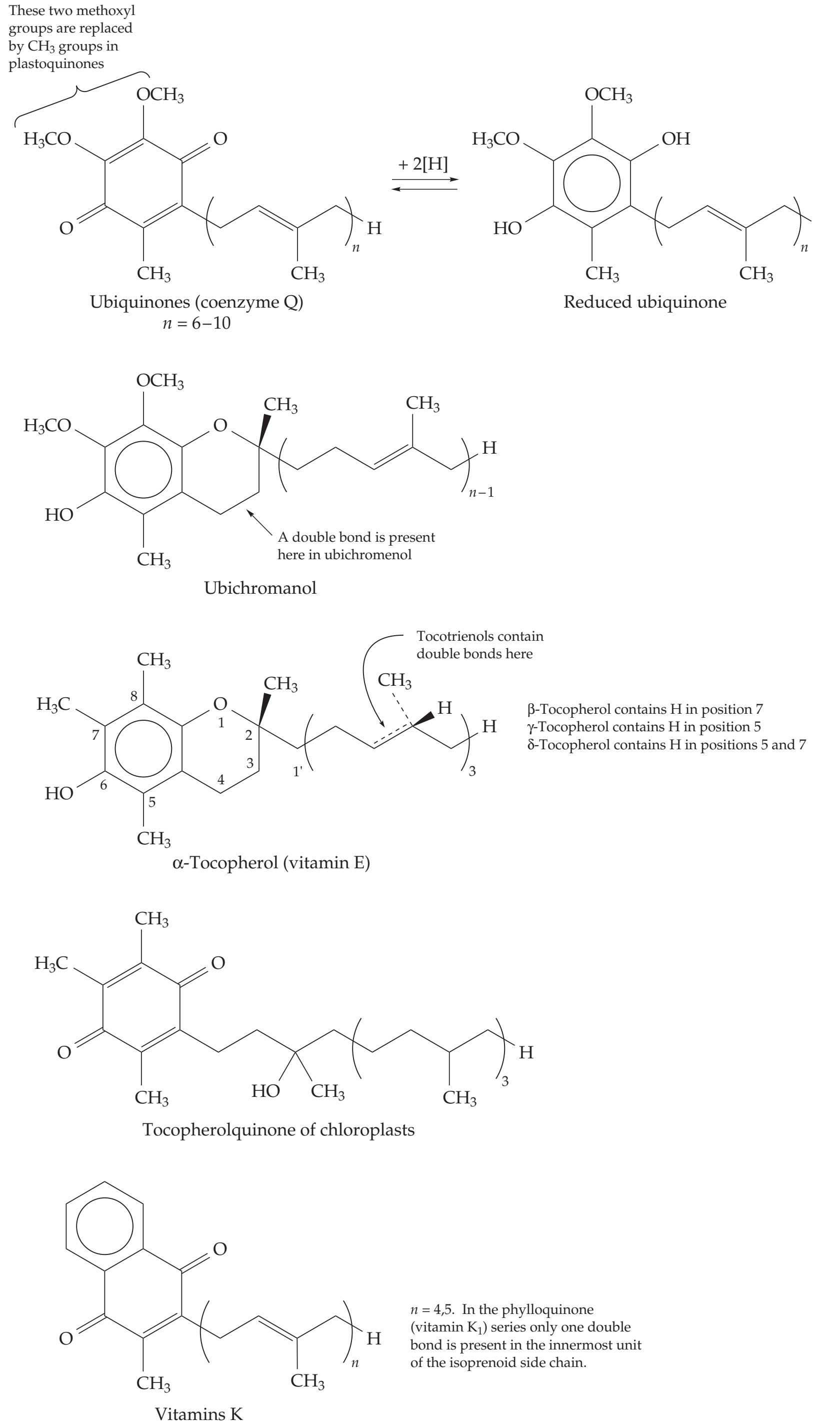

These two methoxyl groups are replaced by CH_3 groups in plastoquinones
OCH_3
H_3CO
O
O
CH_3
CH_3
H
n
Ubiquinones (coenzyme Q)
n = 6–10
+ 2[H]
OCH_3
H_3CO
OH
HO
CH_3
CH_3
H
n
Reduced ubiquinone
OCH_3
H_3CO
O
CH_3
CH_3
H
n–1
HO
CH_3
A double bond is present here in ubichromenol
Ubichromanol
Tocotrienols contain double bonds here
CH_3
H_3C
8
7
O
1
2
3
4
5
6
HO
CH_3
CH_3
CH_3
H
H
1'
3
β-Tocopherol contains H in position 7
γ-Tocopherol contains H in position 5
δ-Tocopherol contains H in positions 5 and 7
α-Tocopherol (vitamin E)
CH_3
H_3C
O
O
CH_3
HO
CH_3
CH_3
H
3
Tocopherolquinone of chloroplasts
O
O
CH_3
CH_3
H
n
n = 4,5. In the phylloquinone (vitamin K_1) series only one double bond is present in the innermost unit of the isoprenoid side chain.
Vitamins K

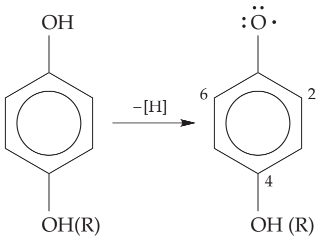
Figure 15-24 Structures of the isoprenoid quinones and vitamin E.

4. Quinones as Electron Carriers

Ubiquinones function as electron transport agents within the inner mitochondrial membranes[496] and also within the reaction centers of the photosynthetic membranes of bacteria (Eq. 23-32).[484,488,494] The plastoquinones also function in electron transport within these membranes. Within some mycobacteria vitamin K apparently participates in electron transport chains in the same way (see Chapter 18). Some bacteria contain both menaquinones and ubiquinones.

The vitamin E derivative **α-tocopherolquinone** (Fig. 15-24) can also serve as an electron carrier, being reversibly reduced to the hydroquinone form α-tocopherolquinol. Such a function has been proposed for the anaerobic rumen bacterium *Butyrovibrio fibrisolvens*.[497]

When a single hydrogen atom is removed from a hydroquinone or from a chromanol such as a tocopherol, a free radical is formed (Eq. 15-54). Phenols substituted in the 2, 4, and 6 positions give especially stable radicals.


OH
OH(R)
–[H]
:Ö·
6
2
4
OH (R)

(15-54)

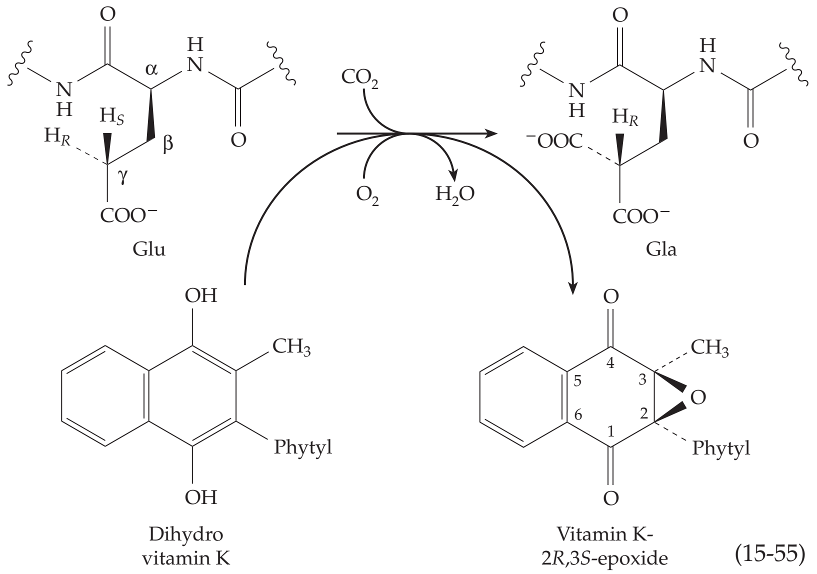
Both the presence of methyl substituents in the tocopherols and their chromanol structures increase the ability of these compounds to form relatively stable radicals.[498,499] This ability is doubtless probably important also in the function of ubiquinones and plastoquinones. Ubiquinone radicals (semiquinones) are probably intermediates in mitochondrial electron transport (Chapter 18) and radicals amounting to as much as 40% of the total ubiquinone in the NADH-ubiquinone reductase of heart mitochondria have been detected by EPR measurements.[500,501]

The equilibria governing semiquinone formation from quinones are similar to those for the flavin semiquinones which were discussed in Section B,6. Two consecutive one-electron redox steps can be defined. Their redox potentials will vary with pH because of a pK_a for the semiquinone in the pH 4.5 –6.5 region. For ubiquinone this pK_a is about 4.9 in water and 6.45 in methanol. A pK_a of over 13 in the hydroquinone form[502] will have little effect on redox potentials near pH 7. The potential for the one-electron reaction Q + e⁻ → Q⁻ is evaluated most readily. For this reaction $E°'$ (pH 7) is −0.074, −0.13, −0.17, and −0.23 for 2,3-dimethylbenzoquinone, plastoquinone, ubiquinone, and phylloquinone, respectively.

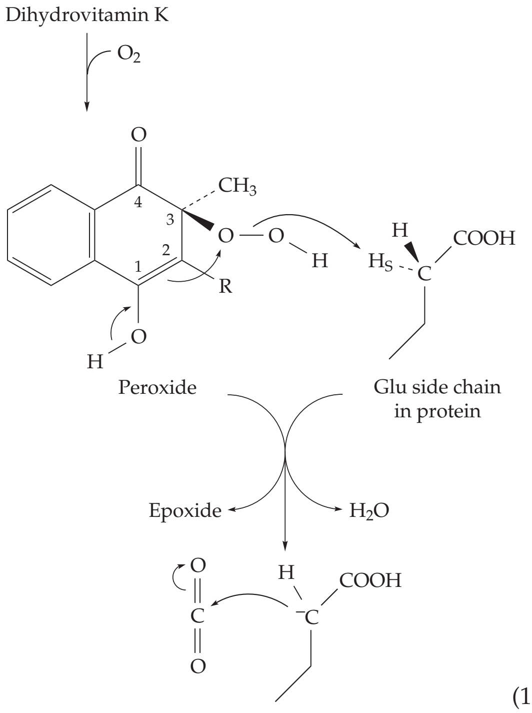
Why does this entire family of compounds have the long polyprenyl side chains? A simple answer is that they serve to anchor the compounds in the lipid portion of the cell membranes where they function. In the case of ubiquinones both the oxidized and the reduced forms may move freely through the lipid phase shuttling electrons between carriers.

5. Vitamin K and γ-Carboxyglutamate Formation

In higher animals the only known function of vitamin K is in the synthesis of γ-carboxyglutamate (Gla)-containing proteins, several of which are needed in blood clotting (Box 15-F, Chapter 12). Following the discovery of γ-carboxyglutamate, it was shown that liver microsomes were able to incorporate ^{14}C-containing bicarbonate or CO_2 into the Gla of prothrombin and could also generate Gla in certain simple peptides such as Phe-Leu-Glu-Val. Three enzymes are required. All are probably bound to the microsomal membranes.[503–507a] An NADPH-dependent reductase reduces vitamin K quinone to its hydroquinone form. Conversion of Glu residues to Gla residues requires this reduced vitamin K as well as O_2 and CO_2. During the carboxylation reaction the reduced vitamin K is converted into vitamin K 2,3-epoxide (Eq. 15-55).[508] The mechanism is uncertain but a peroxide intermediate such as that shown in Eq. 15-56 is probably involved. This could be used to generate a hydroxide ion adjacent to the *pro-S* -H of the glutamate side chain of the substrate. This hydrogen could be abstracted by the OH⁻ to form H_2O and a carbanion which would be stabilized by the adjacent carboxyl group[508–511] (Eq. 15-56).

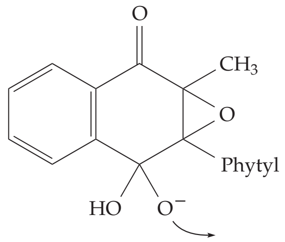


(15-55)



(15-56)

Dowd and coworkers raised doubts that a hydroxide ion released in the active site in this manner is a strong enough base to generate the anion shown in Eq. 15-52.[512,513] They hypothesized a *"base strength amplification" mechanism* that begins with a peroxide formed at C-4 followed by ring closure to form a dioxetane and rearrangement to the following hypothetical strong base.[507,512,514]



Hypothetical strong base

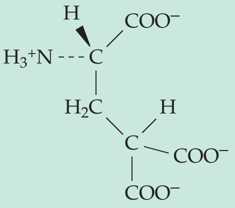
The proposal was supported by model experiments and also by observation of some incorporation of both atoms of $^{18}O_2$ into vitamin K epoxide. However, theoretical calculations support the simpler mechanism of Eq. 15-56.

The fact that L-*threo*-γ-fluoroglutamate residues, in which the fluorine atom is in the position corresponding to the *pro-S* hydrogen, are not carboxylated but that an *erythro*-γ-fluoroglutamate is carboxylated

BOX 15-F THE VITAMIN K FAMILY

The existence of an “antihemorrhagic factor” required in the diet of chicks to ensure rapid clotting of blood was reported in 1929 by Henrik Dam at the University of Copenhagen.[a,b] The fat-soluble material, later designated vitamin K, causes a prompt (2–6 h) decrease in the clotting time when administered to deficient animals and birds. The clotting time for a vitamin K-deficient chick may be greater than 240 s, but 6 h after injection of 2 µg of vitamin K_3 it falls to 50–100 s.[c] Pure vitamin K (Fig. 15-24), a 1,4-naphthoquinone, was isolated from alfalfa in 1939. Within a short time two series, the phylloquinones (vitamin K_1) and the menaquinones (vitamin K_2), were recognized. The most prominant phylloquinone contains the phytyl group, which is also present in the chlorophylls. For a human being a dietary intake of about 30 µg per day is recommended.[d] Additional vitamin K is normally supplied by intestinal bacteria.

The most obvious effect of a deficiency in vitamin K in animals is delayed blood clotting, which has been traced to a decrease in the activity of **prothrombin** and of clotting factors VII, IX, and X (Chapter 12, Fig. 12-17). Prothrombin formed by the liver in the absence of vitamin K lacks the ability to chelate calcium ions essential for the binding of prothrombin to phospholipids and to its activation to thrombin. The structural differences between this abnormal protein and the normal prothrombin have been pinpointed at the N terminus of the ~560 residue glycoprotein.[e,f] Tryptic peptides from the N termini differed in electrophoretic mobility. As detailed in Chapter 12, ten residues within the first 33, which were identified as glutamate residues by the sequence analysis on normal prothrombin, are actually **γ-carboxyglutamate** (Gla). The same amino acid is present near the N termini of clotting factors VII, IX, and X.

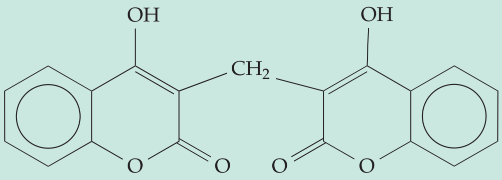


γ-Carboxyglutamate (Gla)

The fact that γ-carboxyglutamate had not been identified previously as a protein substituent is explained by its easy decarboxylation to glutamic acid during treatment with strong acid. The function of vitamin K is to assist in the incorporation of the additional carboxyl group into the glutamate residues of preformed prothrombin and other blood-clotting factors[g,h] with a resulting increase in calcium ion affinity.

Four other plasma proteins designated C, S, M, and Z contain γ-carboxyglutamate. The functions of proteins M and Z are unknown but protein C is a serine protease involved in regulation of blood coagulation and protein S is a cofactor that assists the action of protein C. Other proteins that require vitamin K for synthesis include the 49-residue **bone Gla protein** (or **osteocalcin**) and the 79-residue **matrix Gla protein** found in bone and cartilage.[i,j] These proteins contain three and five residues of Gla, respectively. Their possible functions in mineralization are considered in Chapter 8. At least two additional small human proline-rich Gla proteins of unknown function are synthesized in many tissues.[k] Gamma-carboxyglutamate also occurs in an invertebrate peptide from fish-hunting cone snails. This “sleeper peptide,” which induces sleep in mice after intracerebral injection, has the sequence GEE*E*LQE*NQE*LIRE*KSN. Here E* designates the 5 residues of Gla.[l]

An interesting facet of vitamin K nutrition and metabolism was revealed by the observation that cattle fed on spoiled sweet clover develop a fatal hemorrhagic disease. The causative agent is **dicoumarol**, a compound arising from coumarin, a natural constituent of clover. Dicourmarol and the closely related synthetic **warfarin** are both potent vitamin K antagonists. Warfarin is used both as a rat poison and in the treatment of thromboembolic disease. As rodenticides hydroxycourmarin derivatives are usually safe because a single accidental ingestion by a child or pet does little harm, whereas regular ingestion by rodents is fatal.



Dicoumarol

---

[a] Wasserman, R. H. (1972) *Ann. Rev. Biochem.* **41**, 179–202

[b] Tim Kim, X. (1979) *Trends Biochem. Sci.* **4**, 118–119

[c] Olson, R. E. (1964) *Science* **45**, 926–928

[d] Shils, M. E., Olson, J. A., and Shike, M., eds. (1994) *Modern Nutrition in Health and Disease*, 8th ed., Vol. 1, Lea & Febiger, Philadelphia, Pennsylvania (pp. 353–355)

[e] Friedman, P. A. (1984) *N. Engl. J. Med.* **310**, 1458–1460

[f] Stenflo, J. (1976) *J. Biol. Chem.* **251**, 355–363

[g] Wood, G. M., and Suttie, J. W. (1988) *J. Biol. Chem.* **263**, 3234–3239

[h] Wu, S.-M., Mutucumarana, V. P., Geromanos, S., and Stafford, D. W. (1997) *J. Biol. Chem.* **272**, 11718–11722

[i] Price, P. A., and Williamson, M. K. (1985) *J. Biol. Chem.* **260**, 14971–14975

[j] Price, P. A., Rice, J. S., and Williamson, M. K. (1994) *Protein Sci.* **3**, 822–830

[k] Kulman, J. D., Harris, J. E., Haldeman, B. A., Davie, E. W. (1997) *Proc. Natl. Acad. Sci., USA* **94**, 9058–9062

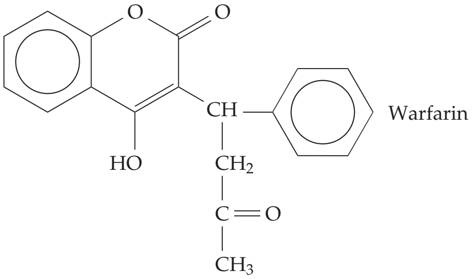
[l] Prorok, M., Warder, S. E., Blandl, T., and Castellino, F. J. (1996) *Biochemistry* **35**, 16528–16534

suggested the indicated stereospecificity.[515,516] This was confirmed by observation of a kinetic isotope effect when ^2H is present in the *pro-S* position.[517] Addition of the carbanion to CO_2 would generate the Gla residue. The two glutamates in the following sequence, in which X may be various amino acids, are carboxylated if the protein also carries a suitable N-terminal signal sequence: EXXXEXC. If a suitable glutamyl peptide is not available for carboxylation the postulated peroxide intermediate (Eq. 15-55) is still converted slowly to the 2,3-epoxide of vitamin K.

A third enzyme is required to reduce the epoxide to vitamin K (Eq. 15-57). The biological reductant is uncertain but dithiols such as dithiothreitol serve in the laboratory.[518] See also Eq. 18-47. Protonation of an intermediate enolate anion would give 3-hydroxy-2, 3-dihydrovitamin K, an observed side reaction product.

$$\text{Vitamin K epoxide} \quad + 2[\text{H}] \rightarrow H_2O + \text{Vitamin K} \qquad (15\text{-}57)$$

This reaction is of interest because of its specific inhibition by such coumarin derivatives as Warfarin:



This synthetic compound, as well as natural coumarin anticoagulants (Box 15-F), inhibits both the vitamin K reductase and the epoxide reductase.[518,519] The matter is of considerable practical importance because of the spread of warfarin-resistant rats in Europe and the United States. One resistance mutation has altered the vitamin K epoxide reductase so that it is much less susceptible to inhibition by warfarin.[519,520]

While glutamate residues in peptides of appropriate sequence are carboxylated by the vitamin K-dependent system, aspartate peptides scarcely react.[503] Beta-carboxyaspartate is present in protein C of the blood anticoagulant system (Fig. 12-17)[521] and in various other proteins containing EGF homology domains (Table 7-3),[522] but the mechanism of its formation is unknown.

6. Tocopherols (Vitamin E) as Antioxidants

The major function of the tocopherols is thought to be the protection of phospholipids of cell membranes against oxidative attack by free radicals and organic peroxides. Peroxidation of lipids, which is described in Chapter 21, can lead to rapid development of rancidity in fats and oils. However, the presence of a small amount of tocopherol inhibits this decomposition, presumably by trapping the intermediate radicals in the form of the more stable tocopherol radicals (Eq. 15-54) which may dimerize or react with other radicals to terminate the chain. Even though only one molecule of tocopherol is present for a thousand molecules of phospholipid, it is enough to protect membranes.[523] That vitamin E does function in this way is supported by the observation that much of the tocopherol requirement of some species can be replaced by *N,N'*-diphenyl-*p*-phenylenediamine, a synthetic antioxidant (see Table 18-5 for the structure of a related substance). Three generations of rats have been raised on a tocopherol-free diet containing this synthetic antioxidant. However, not all of the deficiency symptoms are prevented.

The antioxidant role of α-tocopherol in membranes is generally accepted.[524–526] It is thought to be critical to defense against oxidative injury and to help the body combat the development of tumors and to slow aging. Gamma-tocopherol may be more reactive than α-tocopherol in removing radicals created by NO and other nitrogen oxides.[526] Its actions are strongly linked to those of ascorbic acid (Box 18-D) and selenium. Ascorbate may reduce tocopherol semiquinone radicals, while selenium acts to enhance breakdown of peroxides as described in the next section.

G. Selenium-Containing Enzymes

In 1957, Schwartz and associates showed that the toxic element selenium was also a nutritional factor essential for prevention of the death of liver cells in rats.[527] Liver necrosis would be prevented by as little as 0.1 ppm of selenium in the diet. Similar amounts of selenium were shown to prevent a muscular dystrophy called “white muscle disease” in cattle and sheep grazing on selenium-deficient soil. Sodium selenite and other inorganic selenium compounds were more effective than organic compounds in which Se had replaced sulfur. **Keshan disease**, an often fatal heart condition that is prevalent among childen in Se-deficient regions of China, can be prevented by supplementation of the diet with $NaSeO_3$.[528] Even the little crustacean “water flea” *Daphnia* needs 0.1 part per billion of Se in its water.[529]

Selenium has long been known to enhance the antioxidant activity of vitamin E. Recent work suggests that vitamin E acts as a radical scavenger, preventing

BOX 15-G VITAMIN E: THE TOCOPHEROLS

Vitamin E was recognized in 1926 as a factor preventing sterility in rats that had been fed rancid lipids.[a–e] The curative factor, present in high concentration in wheat germ and lettuce seed oils, is a family of vitamin E compounds, the tocopherols (Fig. 15-24). The first of these was isolated by Evans and associates in 1936. Vitamin E deficiency in the rabbit or rat is accompanied by muscular degeneration (**nutritional muscular dystrophy**; see also Box 15-A) and a variety of other symptoms that vary from one species to another. Animals deficient in vitamin E display obvious physical deterioration followed by sudden death. Muscles of deficient rats show abnormally high rates of oxygen uptake, and abnormalities appear in the membranes of the endoplasmic reticulum as viewed with the electron microscope. It is thought that deterioration of lysosomal membranes may be the immediate cause of death.

The tocopherol requirement of humans is not known with certainty, but about 5 mg (7.5 IU) / day plus an additional 0.6 mg for each gram of polyunsaturated fatty acid consumed may be adequate. It is estimated that the average daily intake is about 14 mg, but the increasing use of highly refined foods may lead to dangerously low consumption. Recent interest[d–f] has been aroused by studies that show that much larger amounts of vitamin E (e.g., 100–400 mg/day) substantially reduce the risk of coronary disease and stroke in both women[g] and men[h] and also decrease oxidative modification of brain proteins.[i] The decrease in heart attacks and stroke may be in part an indirect effect of the anticlotting action of vitamin E quinone.[f] Plant oils are usually the richest sources of tocopherols, while animal products contain lower quantities.

To some extent the vitamin E requirement may be lessened by the presence in the diet of synthetic antioxidants and by selenium. Much evidence supports a relationship between the nutritional need for selenium and that for vitamin E. Lack of either causes muscular dystrophy in many animals as well as severe edema (exudative diathesis) in chicks. Since vitamin E-deficient rats have a low selenide (Se^{2-}) content, it has been suggested that vitamin E protects reduced selenium from oxidation.[j] Vitamin C (ascorbic acid), in turn, protects vitamin E.

---

[a] Sebrell, W. H., Jr., and Harris, R. S., eds. (1972) *The Vitamins*, Vol. 5, Academic Press, New York

[b] DeLuca, H. F., and Suttie, J. W., eds. (1970) *The Fat-Soluble Vitamins*, Univ. of Wisconsin Press, Madison, Wisconsin

[c] Machlin, L. J., ed. (1980) *Vitamin E*, Dekker, New York

[d] Mino, M., Nakamura, H., Diplock, A. T., and Kayden, H. J., eds. (1993) *Vitamin E: Its Usefulness in Health and in Curing Diseases*, Japan Scientific Socities Press, Tokyo

[e] Packer, L., and Fuchs, J., eds. (1993) *Vitamin E in Health and Disease*, Dekker, New York

[f] Dowd, P., and Zhend, Z. B. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8171–8175

[g] Stampfer, M. J., Hennekens, C. H., Manson, J. E., Colditz, G. A., Rosner, B., and Willett, W. C. (1993) *N. Engl. J. Med.* **328**, 1444–1449

[h] Rimm, E. B., Stampfer, M. J., Ascherio, A., Giovannucci, E., Colditz, G. A., and Willett, W. C. (1993) *N. Engl. J. Med.* **328**, 1450–1456

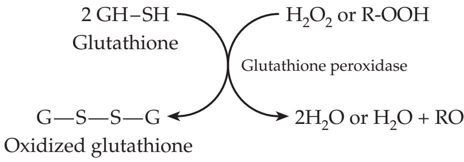
[i] Poulin, J. E., Cover, C., Gustafson, M. R., and Kay, M. M. B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5600–5603

[j] Diplock, A. T., and Lucy, J. A. (1973) *FEBS Lett.* **29**, 205–210

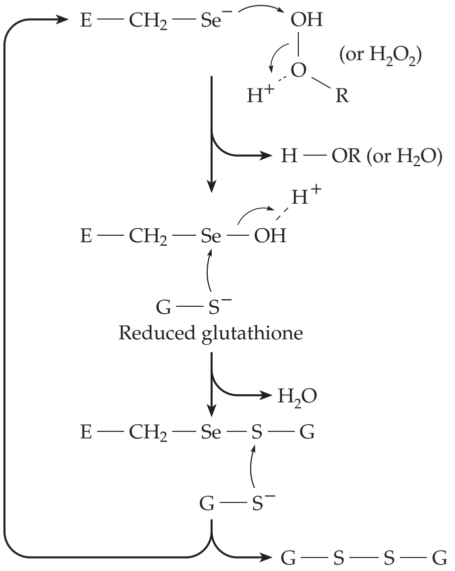
---

excessive peroxidation of membrane lipids, while selenium, in the enzyme **glutathione peroxidase**, acts to destroy the small amounts of peroxides that do form. This was the first established function of selenium in human beings, but there are others. If we include proteins from animals and bacteria, at least ten selenoproteins are known (Table 15-4).[530–534] Seven of these are enzymes and most catalyze redox processes. The active sites most often contain **selenocysteine**, whose selenol side chain is more acidic (pK_a ~5.2) than that of cysteine and exists as $-CH_2-Se^-$ at neutral pH.[530]

Glutathione peroxidases catalyze the reductive decomposition of H_2O_2 or of organic peroxides by glutathione (G–SH) according to Eq. 15-58. At least three isoenzyme forms have been identified in mammals: a cellular form,[531,535–537] a plasma form, and a form with a preference for organic peroxides derived from phospholipids.[327,538–540] A related selenoprotein has been found in a human poxvirus.[540a] Selenocysteine is present at a position (residues 41–47) near the N terminus of an α helix, in the ~ 180-residue polypeptides. A possible reaction mechanism involves attack by the selenol on the peroxide to give a selenic acid intermediate which is reduced by glutathione in two nucleophilic displacement steps (Eq. 15-59).

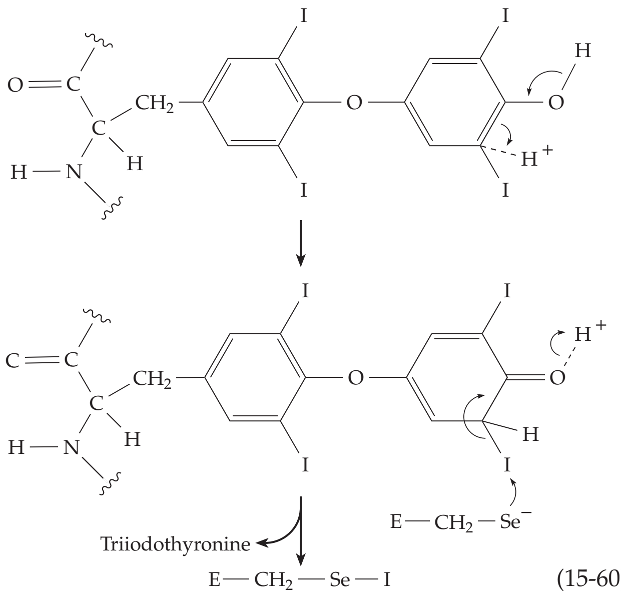


(15-58)


E — CH₂ — Se⁻
OH
O
H⁺
R
(or H₂O₂)
H — OR (or H₂O)
H⁺
E — CH₂ — Se — OH
G — S⁻
Reduced glutathione
H₂O
E — CH₂ — Se — S — G
G — S⁻
G — S — S — G

(15-59)

Three types of **iodothyronine deiodinase** remove iodine atoms from thyroxine to form the active thyroid hormone triiodothyronine and also to inactivate the hormone by removing additional iodine[531,541–546] (see also Chapter 25). In this case the $-CH_2-Se^-$ may attach the iodine atom, removing it as I^+ to form $-CH_2-Se-I$. The process could be assisted by the phenolic –OH group if it were first tautomerized (Eq. 15-60).

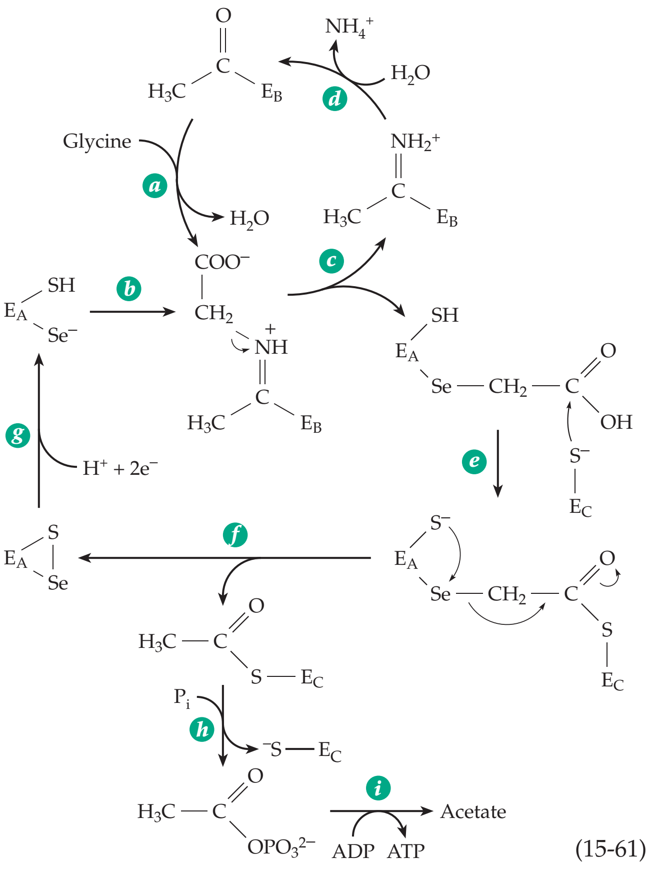

I
I
O
C
CH₂
C
H
H — N
O
O
H
H⁺
I
I
I
I
C
C
CH₂
C
H
H — N
O
O
H⁺
H
I
I
E — CH₂ — Se⁻
Triiodothyronine
E — CH₂ — Se — I

(15-60)

A recently discovered human selenoprotein is a **thioredoxin reductase** which is present in the T cells of the immune system as well as in placenta and other tissues.[189,547–549] The 55-kDa protein has one selenocysteine as the penultimate C-terminal residue. Another mammalian selenoprotein, of uncertain function, is the 57-kDa **selenoprotein P**. It contains over 60% of the selenium in rat plasma and is also present in the human body. Selenoprotein P contains ten selenocysteine residues.[550–552a] Some of these may be replaced by serine in a fraction of the molecules.[553] A smaller 9.6-kDa skeletal muscle protein, **selenoprotein W**, contains a single selenocysteine.[554–556] Another selenoprotein has been found in sperm cells, both in the tail and in a keratin-rich capsule that surrounds the mitochondria in the sperm midpiece.[557] Lack of this protein may be the cause of the abnormal immotile spermatozoa observed in Se-deficient rats and of reproductive difficulties among farm animals in Se-deficient regions.[558]

Several selenoproteins have been found in certain bacteria and archaea. A **hydrogenase** from *Methanococcus vannielii* contains selenocysteine.[559,560] This enzyme transfers electrons from H_2 to the C-5 *si* face of the 8-hydroxy-5-deazaflavin cofactor F_{420} (Section B,4). The same bacterium synthesizes two **formate dehydrogenases** (see Fig 15-23), one of which contains Se. Two Se-containing formate dehydrogenases are made by *E. coli*. One of them, which is coupled to a hydrogenase in the formate hydrogen-lyase system (see Eq. 15-37), is a 715-residue protein containing selenocysteine at position 140.[561–563] The second has selenocysteine at position 196 and functions with a nitrate reductase in anaerobic nitrate respiration.[561]

Glycine reductase is a complex enzyme[530,564–566] that catalyzes the reductive cleavage of glycine to acetyl phosphate and ammonia (Eq. 15-61) with the


O
H₃C
C
E_B
NH₄⁺
H₂O
d
Glycine
a
H₂O
NH₂⁺
H₃C
C
E_B
COO⁻
CH₂
NH
H₃C
C
E_B
b
c
SH
E_A
Se⁻
SH
E_A
Se — CH₂ — C
O
OH
S⁻
E_C
e
g
H⁺ + 2e⁻
f
E_A
S
Se
S⁻
E_A
Se — CH₂ — C
O
S
E_C
H₃C — C
O
S — E_C
P_i
h
⁻S — E_C
H₃C — C
O
OPO₃²⁻
i
Acetate
ADP ATP

(15-61)

subsequent synthesis of ATP (Eq. 14-43). Electrons for reduction of the disulfide that is formed are provided by NADH. A single selenocysteine residue is present in the small 12-kDa subunits. The enzyme contains a dehydroalanine residue (Chapter 14) in subunit B and a thiol group in subunit C.[566] An acetyl-enzyme derivative of subunit C, perhaps of its –SH group, has been identified.[546] The mechanism of action is uncertain but the steps in Eq. 15-61 have been suggested.[567] The subunits are designated E_A, E_B, and E_C. Step *e* is particularly hard to understand because formation of a thioester in this manner is not expected to occur spontaneously and must be linked in some way to other steps.

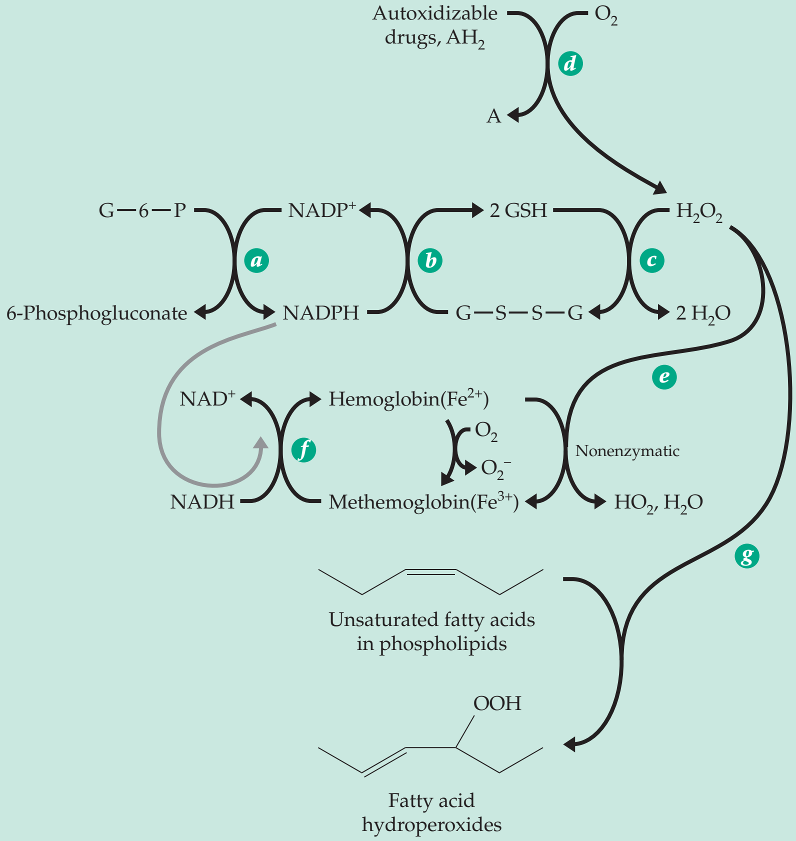
A selenium-containing **xanthine dehydrogenase** is present in purine-fermenting clostridia. Like other xanthine dehydrogenases (Chapter 16), it converts xanthine to uric acid and contains nonheme iron, molybdenum, FAD, and an Fe–S center. The selenium is probably present as Se^{2-} bound to Mo as is S^{2-} in xanthine oxidase (Fig. 16-32).[567a] A related reaction is catalyzed by the Fe–S protein **nicotinic acid hydroxylase** (Eq. 15-62) found in some clostridia.[568] Splitting of the Mo(V) EPR signal when ^{77}Se is present in the enzyme shows that the selenium is present as a ligand of molybdenum. Another member of the family is a purine hydroxylase that converts purine 2-hydroxypurine, or hypoxanthine to xanthine.[567a]

TABLE 15-4
Selenium-Containing Proteins

Enzyme	Source	Mass (kDa)	Subunit composition	Other cofactors
Glutathione peroxidases				None
Cellular	Mammals	21 x 4		
Plasma	Humans			
Phospholipid hydroperoxide	Pig, rat	18		
Iodothyronine deiodinases	Vertebrates			
Thioredoxin reductase	Humans	55 × 2		
Selenoprotein *P*	Mammals	57		None
Selenoprotein *W*	Rat	19.6		
Formate dehydrogenase	Bacteria - *E. coli* Archaea	600	$\alpha_2\beta_4\gamma_{2\text{-}4}$	Heme *b* Mo; molybdopterin
Hydrogenase	Bacteria Archaea		$\alpha_2\beta_4\gamma_2$	FAD; NiFeSe
Glycine reductase	Some clostridia		ABC	
Selenoprotein A		12		
Selenoprotein B		200		
Carbonyl protein C		250		Fe?
Nicotinic acid hydroxylase	Clostridia	300		FAD, FeS, Mo
Purine hydroxylase	Clostridia			FAD, FeS, Mo
Thiolase (contains selenomethionine)	*Clostridium kluyverii*	39 × 4		None
Carbon monoxide dehydrogenase	*Oligotropha carboxidovorans*	137 × 2	$\alpha_2\beta_2\gamma_2$	FAD; Mo; molybdopterin

BOX 15-H GLUTATHIONE PEROXIDASE AND ABNORMALITIES OF RED BLOOD CELLS

The processes by which hemoglobin is kept in the Fe(II) state and functioning normally within intact erythrocytes is vital to our health. Numerous hereditary defects leading to a tendency toward anemia have helped to unravel the biochemistry indicated in the accompanying scheme.[a]



About 90% of the glucose utilized by erythrocytes is converted by glycolysis to lactate, but about 10% is oxidized (via glucose 6-phosphate) to 6-phosphogluconate. The oxidation (reaction *a*) is catalyzed by glucose-6-phosphate dehydrogenase (Eq.15-10) using $NADP^+$. This is the principal reaction providing the red cell with NADPH for reduction of glutathione (Box 11-B) according to reaction *b*. Despite the important function of glucose-6-*P* dehydrogenase, ~400 million persons, principally in tropical and Mediterranean areas, have a hereditary deficiency of this enzyme. The genetic variations are numerous, with about 400 different ones having been identified. Although most individuals with this dificiency have no symptoms, the lack of the enzyme is truly detrimental and sometimes leads to excessive destruction of red cells and anemia during some sicknesses and in response to administration of various drugs.[b] The survival of the defective genes, like that for sickle cell hemoglobin (Box 7-B) is thought to result from increased resistance to malaria parasites.

Other erythrocyte defects that lead to drug sensitivity include a deficiency of glutathione (resulting from a decrease in its synthesis) and a deficiency of glutathione reductase (reaction *b*). The effects of drugs have been traced to the production of H_2O_2 (reaction *c*) in red blood cells; catalase, which converts H_2O_2 into H_2O and O_2, is thought to function in a similar way. Both enzymes are probably necessary for optimal health.

An excess of H_2O_2 can damage erythrocytes in two ways. It can cause excessive oxidation of functioning hemoglobin to the Fe(III)-containing methemoglobin. (Methemoglobin is also formed spontaneously during the course of the oxygen-carrying function of hemoglobin. It is estimated that normally as much as 3% of the hemoglobin may be oxidized to methemoglobin daily.) The methemoglobin formed is reduced back to hemoglobin through the action of **NADH-methemoglobin reductase** (reaction *f*). A smaller fraction of the methemoglobin is reduced by a similar enzyme requiring NADPH (as indicated by the colored arrow). A hereditary lack of the NADH-methemoglobin reductase is also known.

A second destructive function of H_2O_2 is attack on double bonds of unsaturated fatty acids of the phospholipids in cell membranes. The resulting fatty acid hydroperoxides can react further with C–C chain cleavage and disruption of the membrane. This is thought to be the principal cause of the hemolytic anemia induced by drugs in susceptible individuals. The selenium-containing glutathione peroxidase is thought to decompose these fatty acid hydroperoxides. Vitamin E (Box 15-G), acting as an antioxidant within membranes, is also needed for good health of erythrocytes.[c,d]

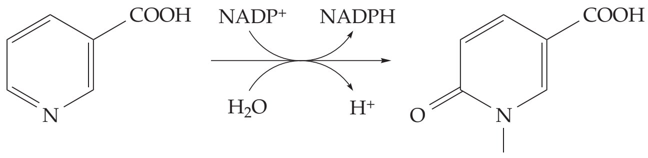
---

[a] Chanarin, I. (1970) in *Biochemical Disorders in Human Disease*, 3rd ed. (Thompson, R. H. S., and Wooton, I. D. P., eds), pp. 163–173, Academic Press, New York

[b] Luzzatto, L., and Mehta, A. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3367–3398, McGraw-Hill, New York

[c] Constantinescu, A., Han, D., and Packer, L. (1993) *J. Biol. Chem.* **268**, 10906–10913

[d] Liebler, D. C., and Burr, J. A. (1992) *Biochemistry* **31**, 8278–8284

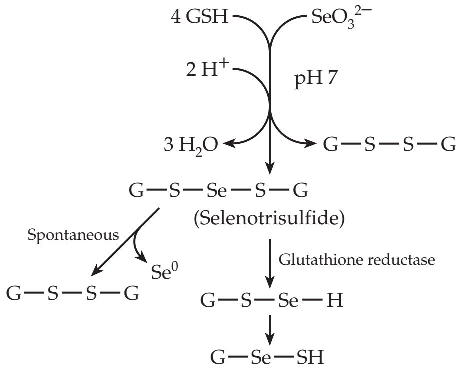


(15-62)

A **thiolase** (Eq. 13-35) from *Clostridium kluyveri* is one of only two known selenoproteins that contain selenomethionine.[569] However, the selenomethionine is incorporated randomly in place of methionine. This occurs in all proteins of all organisms to some extent and the toxicity of selenium may result in part from excessive incorporation of selenomethionine into various proteins.

Selenium is found to a minor extent wherever sulfur exists in nature. This includes the sulfur-containing modified bases of tRNA molecules. In addition to a small amount of nonspecific incorporation of Se into all S-containing bases there are, at least in bacteria, specific Se-containing tRNAs. In *E. coli* one of these is specific for lysine and one for glutamate. One of the modified bases has been identified as 5-methyl-aminomethyl-2-selenouridine.[570] It is present at the first position of the anticodon, the "wobble" position.[571]

Selenium has its own metabolism. Through the use of ^{75}Se as a tracer, normal rat liver has been shown to contain Se^{2-}, SeO_3^{2-}, and selenium in a higher oxidation state.[572] Glutathione may be involved in reduction of selenite to selenide.[573] The nonenzymatic reduction of selenite by glutathione yields a selenotrisulfide derivative (Eq. 15-63). The latter is spontaneously decomposed to oxidized glutathione and elemental selenium or by the action of glutathione reductase to glutathione and selenium. Selenocysteine can be converted to alanine + elemental selenium (Eq. 14-34). Some bacteria are able to oxidize elemental Se back to selenite.[574] Selenium undergoes biological methylation readily in bacteria, fungi, plants, and animals (Chapter 16).[575,576] This may in some way be related to the reported effect of selenium in protecting animals against the toxicity of mercury. Excess selenium may appear in the urine as trimethylselenonium ions.[577]



(15-63)

How is selenium incorporated into selenocysteine-containing proteins? This element does enter amino acids to a limited extent via the standard synthetic pathways for cysteine and methionine. However, the placement of selenocysteine into specific positions in selenoproteins occurs by the use of a minor serine-specific tRNA that acts as a suppressor of chain termination during protein synthesis.[532,533,578] (This topic is dealt with further in Chapter 29.) The genes for these and presumably for other selenocysteine-containing proteins have the "stop" codon TGA at the selenocysteine positions. However, when present in a suitable "context" the minor tRNA, carrying selenocysteine in place of serine, is utilized to place selenocysteine into the growing peptide chain. In bacteria, and presumably also in eukaryotes, the selenocysteinyl-tRNA is formed from the corresponding seryl-tRNA by a PLP-catalyzed β-replacement reaction. The selenium donor is not Se^{2-} but selenophosphate $Se-PO_3^{2-}$ in which the Se–P bond is quite weak.[579–581] After addition to the aminoacrylate intermediate in the PLP enzyme the Se–P bond may be hydrolytically cleaved to HPO_4^{2-} and selenocysteyl-tRNA.

References

1. Cornish-Bowden, A., ed. (1997) *New Beer in an Old Bottle (Eduard Buchner and the Growth of Biochemical Knowledge)*, Valencia, Universitat de València

1a. Kalckar, H. M. (1969) *Biological Phosphorylations*, Prentice-Hall, Englewood Cliffs, New Jersey (pp. 86–97)

2. Rossmann, M. G., Adams, M. J., Buehner, M., Ford, G. C., Hackert, M. L., Lentz, P. J., Jr., McPherson, A., Jr., Schevitz, R. W., and Smiley, I. E. (1971) *Cold Spring Harbor Symposia on Quant. Biol.* **36**, 179–191

3. Abad-Zapatero, C., Griffith, J. P., Sussman, J. L., and Rossmann, M. G. (1987) *J. Mol. Biol.* **198**, 445–467

4. Rossmann, M. (1974) *New Scientist* **61**, 266–268

5. Eventoff, W., Rossmann, M. G., Taylor, S. S., Torff, H.-J., Meyer, H., Keil, W., and Kiltz, H.-H. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2677–2681

6. Hogrefe, H. H., Griffith, J. P., Rossmann, M. G., and Goldberg, E. (1987) *J. Biol. Chem.* **262**, 13155–13162

7. Iwata, S., and Ohta, T. (1993) *J. Mol. Biol.* **230**, 21–27

8. Ostendorp, R., Auerbach, G., and Jaenicke, R. (1996) *Protein Sci.* **5**, 862–873

9. Birktoft, J. J., Rhodes, G., and Banaszak, L. J. (1989) *Biochemistry* **28**, 6065–6081

10. Gleason, W. B., Fu, Z., Birktoft, J., and Banaszak, L. (1994) *Biochemistry* **33**, 2078–2088

11. Goward, C. R., and Nicholls, D. J. (1994) *Protein Sci.* **3**, 1883–1888

12. Hall, M. D., Levitt, D. G., and Banaszak, L. J. (1992) *J. Mol. Biol.* **226**, 867–882

13. Kelly, C. A., Nishiyama, M., Ohnishi, Y., Beppu, T., and Birktoft, J. J. (1993) *Biochemistry* **32**, 3913–3922

14. Bellamacina, C. R. (1996) *FASEB J.* **10**, 1257–1269

15. Pullman, M. E., and Colowick, S. P. (1954) *J. Biol. Chem.* **206**, 129–141

16. Fisher, H. F., Conn, E. E., Vennesland, B., and Westheimer, F. H. (1953) *J. Biol. Chem.* **202**, 687–697

17. Westheimer, F. H. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical and Medical Aspects*, Vol. A (Dolphin, D., Avramovic, O., and Poulson, R., eds), pp. 253–322, Wiley (Interscience), New York

18. You, K.-s, Arnold, L. J., Jr., Allison, W. S., and Kaplan, N. O. (1978) *Trends Biochem. Sci.* **3**, 265–268

19. Nakajima, N., Nakamura, K., Esaki, N., Tanaka, H., and Soda, K. (1989) *J. Biochem.* **106**, 515–517

20. Esaki, N., Shimoi, H., Nakajima, N., Ohshima, T., Tanaka, H., and Soda, K. (1989) *J. Biol. Chem.* **264**, 9750–9752

21. Anderson, V. E., and LaReau, R. D. (1988) *J. Am. Chem. Soc.* **110**, 3695–3697

22. Almarsson, Ö., and Bruice, T. C. (1993) *J. Am. Chem. Soc.* **115**, 2125–2138

23. Oppenheimer, N. J. (1986) *J. Biol. Chem.* **261**, 12209–12212

24. Benner, S. A., Nambiar, K. P., and Chambers, G. K. (1985) *J. Am. Chem. Soc.* **107**, 5513–5517

25. Young, L., and Post, C. B. (1993) *J. Am. Chem. Soc.* **115**, 1964–1970

26. Almarsson, Ö, Karaman, R., and Bruice, T. C. (1992) *J. Am. Chem. Soc.* **114**, 8702–8704

27. LaReau, R. D., and Anderson, V. E. (1992) *Biochemistry* **31**, 4174–4180

28. Ohno, A., Tsutsumi, A., Kawai, Y., Yamazaki, N., Mikata, Y., and Okamura, M. (1994) *J. Am. Chem. Soc.* **116**, 8133–8137

29. Kallwass, H. K. W., Hogan, J. K., Macfarlane, E. L. A., Martichonok, V., Parris, W., Kay, C. M., Gold, M., and Jones, J. B. (1992) *J. Am. Chem. Soc.* **114**, 10704–10710

30. Klinman, J. P. (1972) *J. Biol. Chem.* **247**, 7977–7987

31. Ostovic, D., Roberts, R. M. G., and Kreevoy, M. M. (1983) *J. Am. Chem. Soc.* **105**, 7629–7631

32. Coleman, C. A., Rose, J. G., and Murray, C. J. (1992) *J. Am. Chem. Soc.* **114**, 9755–9762

33. Kong, Y. S., and Warshel, A. (1995) *J. Am. Chem. Soc.* **117**, 6234–6242

34. Lee, I.-S. H., Jeoung, E. H., and Kreevoy, M. M. (1997) *J. Am. Chem. Soc.* **119**, 2722–2728

35. Rucker, J., Cha, Y., Jonsson, T., Grant, K. L., and Klinman, J. P. (1992) *Biochemistry* **31**, 11489–11499

36. Huskey, W. P., and Schowen, R. L. (1983) *J. Am. Chem. Soc.* **105**, 5704–5706

36a. Kohen, A., Cannio, R., Bartolucci, S., and Klinman, J. P. (1999) *Nature (London)* **399**, 496–499

36b. Chin, J. K., and Klinman, J. P. (2000) *Biochemistry* **39**, 1278–1284

36c. Northrop, D. B. (1999) *J. Am. Chem. Soc.* **121**, 3521–3524

36d. Karsten, W. E., Hwang, C.-C., and Cook, P. F. (1999) *Biochemistry* **38**, 4398–4402

37. Grau, U. M., Trommer, W. E., and Rossmann, M. G. (1981) *J. Mol. Biol.* **151**, 289–307

38. Li, H., Hallows, W. H., Punzi, J. S., Pankiewicz, K. W., Watanabe, K. A., and Goldstein, B. M. (1994) *Biochemistry* **33**, 11734–11744

39. Gawlita, E., Paneth, P., and Anderson, V. E. (1995) *Biochemistry* **34**, 6050–6058

40. Clarke, A. R., Wilks, H. M., Barstow, D. A., Atkinson, T., Chia, W. N., and Holbrook, J. J. (1988) *Biochemistry* **27**, 1617–1622

41. Clarke, A. R., Atkinson, T., and Holbrook, J. J. (1989) *Trends Biochem. Sci.* **14**, 101–105

42. Clarke, A. R., Atkinson, T., and Holbrook, J. J. (1989) *Trends Biochem. Sci.* **14**, 145–148

43. Goldberg, J. D., Yoshida, T., and Brick, P. (1994) *J. Mol. Biol.* **236**, 1123–1140

44. Cortes, A., Emery, D. C., Halsall, D. J., Jackson, R. M., Clarke, A. R., and Holbrook, J. J. (1992) *Protein Sci.* **1**, 892–901

45. Clarke, A. R., Wigley, D. B., Chia, W. N., Barstow, D. A., Atkinson, T., and Holbrook, J. J. (1986) *Nature (London)* **324**, 699–702

46. Deng, H., Burgner, J., and Callender, R. (1992) *J. Am. Chem. Soc.* **114**, 7997–8003

47. Grau, W. M., Trommer, W. E., and Rossmann, M. G. (1981) *J. Mol. Biol.* **151**, 289–307

48. Pettersson, G. (1987) *CRC Critical Review of Biochemistry* **21**, 349–389

49. Maret, W., and Makinen, M. W. (1991) *J. Biol. Chem.* **266**, 20636–20644

50. Eklund, H., and Brändén, C.-I. (1987) in *Pyridine Nucleotide Coenzyme: Chemical, Biochemical and Medical Aspects*, Vol. 2, Part A (Dolphin, D., Poulson, R., and Avramovic, O., eds), pp. 51–98, Wiley, New York

51. Adolph, H. W., Kiefer, M., and Cedergren-Zeppezauer, E. (1997) *Biochemistry* **36**, 8743–8754

51a. Colby, T. D., Bahnson, B. J., Chin, J. K., Klinman, J. P., and Goldstein, B. M. (1998) *Biochemistry* **37**, 9295–9304

51b. Plapp, B. V., Eklund, H., and Brändén, C.-I. (1978) *J. Mol. Biol.* **122**, 23–32

52. Shearer, G. L., Kim, K., Lee, K. M., Wang, C. K., and Plapp, B. V. (1993) *Biochemistry* **32**, 11186–11194

53. Ramaswamy, S., Eklund, H., and Plapp, B. V. (1994) *Biochemistry* **33**, 5230–5237

54. Dunn, M. F., Dietrich, H., MacGibbon, A. K. H., Koerber, S. C., and Zeppezauer, M. (1982) *Biochemistry* **21**, 354–363

55. Jagodzinski, P. W., and Peticolas, W. L. (1981) *J. Am. Chem. Soc.* **103**, 234–236

56. MacGibbon, A. K. H., Koerber, S. C., Pease, K., and Dunn, M. F. (1987) *Biochemistry* **26**, 3058–3067

57. Hennecke, M., and Plapp, B. V. (1983) *Biochemistry* **22**, 3721–3728

58. Plapp, B. V. (1995) *Methods Enzymol.* **249**, 91–119

59. Makinen, M. W., Maret, W., and Yim, M. B. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2584–2588

60. Ryde, U. (1995) *Protein Sci.* **4**, 1124–1132

61. Ehrig, T., Muhoberac, B. B., Brems, D., and Bosron, W. F. (1993) *J. Biol. Chem.* **268**, 11721–11726

62. Hurley, T. D., Bosron, W. F., Stone, C. L., and Amzel, L. M. (1994) *J. Mol. Biol.* **239**, 415–429

63. Charlier, H. A., Jr., and Plapp, B. V. (2000) *J. Biol. Chem.* **275**, 11569–11575

64. Xie, P., Parsons, S. H., Speckhard, D. C., Bosron, W. F., and Hurley, T. D. (1997) *J. Biol. Chem.* **272**, 18558–18563

65. Cho, H., Ramaswamy, S., and Plapp, B. V. (1997) *Biochemistry* **36**, 382–389

66. Ganzhorn, A. J., Green, D. W., Hershey, A. D., Gould, R. M., and Plapp, B. V. (1987) *J. Biol. Chem.* **262**, 3754–3761

67. Hake, S., Kelley, P. M., Taylor, W. C., and Freeling, M. (1985) *J. Biol. Chem.* **260**, 5050–5054

67a. Korkhin, Y., Kalb-Gilboa, A. J., Peretz, M., Bogin, O., Burstein, Y., and Frolow, F. (1998) *J. Mol. Biol.* **278**, 967–981

68. Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàlez-Duarte, R., Jeffery, J., and Ghosh, D. (1995) *Biochemistry* **34**, 6003–6013

69. Ribas de Pouplana, L., and Fothergill-Gilmore, L. A. (1994) *Biochemistry* **33**, 7047–7055

69a. Benach, J., Atrian, S., Gonzàlez-Duarte, R., and Ladenstein, R. (1999) *J. Mol. Biol.* **289**, 335–355

70. Hawes, J. W., Crabb, D. W., Chan, R. M., Rougraff, P. M., and Harris, R. A. (1995) *Biochemistry* **34**, 4231–4237

71. Pawlowski, J. E., and Penning, T. M. (1994) *J. Biol. Chem.* **269**, 13502–13510

72. Wilson, D. K., Nakano, T., Petrash, J. M., and Quiocho, F. A. (1995) *Biochemistry* **34**, 14323–14330

73. Hoog, S. S., Pawlowski, J. E., Alzari, P. M., Penning, T. M., and Lewis, M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2517–2521

74. Wilson, D. K., Bohren, K. M., Gabbay, K. H., and Quiocho, F. A. (1992) *Science* **257**, 81–84

75. Rondeau, J.-M., Tête-Favier, F., Podjarny, A., Reymann, J.-M., Barth, P., Biellmann, J.-F., and Moras, D. (1992) *Nature (London)* **355**, 469–471

76. Tarle, I., Borhani, D. W., Wilson, D. K., Quiocho, F. A., and Petrash, J. M. (1993) *J. Biol. Chem.* **268**, 25687–25693

77. Grimshaw, C. E., Bohren, K. M., Lai, C.-J., and Gabbay, K. H. (1995) *Biochemistry* **34**, 14374–14384

78. Grimshaw, C. E. (1992) *Biochemistry* **31**, 10139–10145

79. Harrison, D. H., Bohren, K. M., Ringe, D., Petsko, G. A., and Gabbay, K. H. (1994) *Biochemistry* **33**, 2011–2020

80. Maniscalco, S. J., Saha, S. K., Vicedomine, P., and Fisher, H. F. (1996) *Biochemistry* **35**, 89–94

81. Saha, S. K., Maniscalco, S. J., and Fisher, H. F. (1996) *Biochemistry* **35**, 16483–16488

82. Smith, M. T., and Emerich, D. W. (1993) *J. Biol. Chem.* **268**, 10746–10753

83. Delforge, D., Devreese, B., Dieu, M., Delaive, E., Van Beeumen, J., and Remacle, J. (1997) *J. Biol. Chem.* **272**, 2276 – 2284
84. Sekimoto, T., Fukui, T., and Tanizawa, K. (1994) *J. Biol. Chem.* **269**, 7262 – 7266
85. Brunhuber, N. M. W., Banerjee, A., Jacobs, W. R., Jr., and Blanchard, J. S. (1994) *J. Biol. Chem.* **269**, 16203 – 16211
86. Britton, K. L., Baker, P. J., Engel, P. C., Rice, D. W., and Stillman, T. J. (1993) *J. Mol. Biol.* **234**, 938 – 945
87. Scapin, G., Blanchard, J. S., and Sacchettini, J. C. (1995) *Biochemistry* **34**, 3502 – 3512
88. Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C., and Rossmann, M. G. (1975) *J. Biol. Chem.* **250**, 9137 – 9162
89. Duée, E., Olivier-Deyris, L., Fanchon, E., Corbier, C., Branlant, G., and Dideberg, O. (1996) *J. Mol. Biol.* **257**, 814 – 838
90. Skarzynski, T., Moody, P. C. E., and Wonacott, A. J. (1987) *J. Mol. Biol.* **193**, 171 – 187

90a. Roitel, O., Sergienko, E., and Branlant, G. (1999) *Biochemistry* **38**, 16084 – 16091

91. Kim, H., and Hol, W. G. J. (1998) *J. Mol. Biol.* **278**, 5 – 11
92. Soukri, A., Mougin, A., Corbier, C., Wonacott, A., Branlant, C., and Branlant, G. (1989) *Biochemistry* **28**, 2568 – 2592
93. Habenicht, A., Hellman, U., and Cerff, R. (1994) *J. Mol. Biol.* **237**, 165 – 171

93a. Marchal, S., Rahuel-Clermont, S., and Branlant, G. (2000) *Biochemistry* **39**, 3327 – 3335

94. Sheikh, S., Ni, L., Hurley, T. D., and Weiner, H. (1997) *J. Biol. Chem.* **272**, 18817 – 18822
95. Steinmetz, C. G., Xie, P., Weiner, H., and Hurley, T. D. (1997) *Structure* **5**, 701 – 711
96. Henehan, G. T. M., and Oppenheimer, N. J. (1993) *Biochemistry* **32**, 735 – 738
97. Henehan, G. T. M., Chang, S. H., and Oppenheimer, N. J. (1995) *Biochemistry* **34**, 12294 – 12301
98. Shearer, G. L., Kim, K., Lee, K. M., Wang, C. K., and Plapp, B. V. (1993) *Biochemistry* **32**, 11186 – 11194
99. Feingold, D. S., and Franzen, J. S. (1981) *Trends Biochem. Sci.* **6**, 103 – 105
100. Frimpong, K., and Rodwell, V. W. (1994) *J. Biol. Chem.* **269**, 11478 – 11483
101. Omkumar, R. V., and Rodwell, V. W. (1994) *J. Biol. Chem.* **269**, 16862 – 16866
102. Friesen, J. A., and Rodwell, V. W. (1997) *Biochemistry* **36**, 1157 – 1162
103. Lawrence, C. M., Rodwell, V. W., and Stauffacher, C. V. (1995) *Science* **268**, 1758 – 1762
104. Friesen, J. A., Lawrence, C. M., Stauffacher, C. V., and Rodwell, V. W. (1996) *Biochemistry* **35**, 11945 – 11950

104a. Cosgrove, M. S., Naylor, C., Paludan, S., Adams, M. J., and Levy, H. R. (1998) *Biochemistry* **37**, 2759 – 2767

105. Luzzatto, L., and Mehta, A. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3367 – 3398, McGraw-Hill, New York
106. Wakil, S. J. (1989) *Biochemistry* **28**, 4523 – 4530
107. Smith, S. (1994) *FASEB J.* **8**, 1248 – 1259

107a. Barycki, J. J., O'Brien, L. K., Bratt, J. M., Zhang, R., Sanishvili, R., Strauss, A. W., and Banaszak, L. J. (1999) *Biochemistry* **38**, 5786 – 5798

107b. Fillgrove, K. L., and Anderson, V. E. (2000) *Biochemistry* **39**, 7001 – 7011

108. Watkinson, I. A., Wilton, D. L., Rahimtula, A. D., and Akhtar, M. M. (1971) *Eur. J. Biochem.* **23**, 1 – 6
109. Thorn, J. M., Barton, J. D., Dixon, N. E., Ollis, D. L., and Edwards, K. J. (1995) *J. Mol. Biol.* **249**, 785 – 799
110. Bolduc, J. M., Dyer, D. H., Scott, W. G., Singer, P., Sweet, R. M., Koshland, D. E., Jr., and Stoddard, B. L. (1995) *Science* **268**, 1312 – 1318
111. Lee, M. E., Dyer, D. H., Klein, O. D., Bolduc, J. M., Stoddard, B. L., and Koshland, D. E., Jr. (1995) *Biochemistry* **34**, 378 – 384
112. Schütz, M., and Radler, F. (1973) *Arch. Mikrobiol.* **91**, 183
113. Hwang, C.-C., and Cook, P. F. (1998) *Biochemistry* **37**, 15698 – 15702
114. Karsten, W. E., Gavva, S. R., Park, S.-H., and Cook, P. F. (1995) *Biochemistry* **34**, 3253 – 3260
115. Caspritz, G., and Radler, F. (1983) *J. Biol. Chem.* **258**, 4907 – 4910
116. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* **35**, 5137 – 5144
117. Frey, P. A. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical and Medical Aspects*, Vol. B (Dolphin, D., Avramovic, O., and Poulson, R., eds), pp. 461 – 512, Wiley (Interscience), New York

117a. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., and Holden, H. M. (2000) *Biochemistry* **39**, 5691 – 5701

118. Thoden, J. B., Hegeman, A. D., Wesenberg, G., Chapeau, M. C., Frey, P. A., and Holden, H. M. (1997) *Biochemistry* **36**, 6294 – 6304
119. Palmer, J. L., and Abeles, R. H. (1979) *J. Biol. Chem.* **254**, 1217 – 1226
120. Yuan, C.-S., Ault-Riché, D. B., and Borchardt, R. T. (1996) *J. Biol. Chem.* **271**, 28009 – 28016
121. Koch-Nolte, F., Petersen, D., Balasubramanian, S., Haag, F., Kahlke, D., Willer, T., Kastelein, R., Bazan, F., and Thiele, H.-G. (1996) *J. Biol. Chem.* **271**, 7686 – 7693
122. Takada, T., Iida, K., and Moss, J. (1995) *J. Biol. Chem.* **270**, 541 – 544
123. Rising, K. A., and Schramm, V. L. (1997) *J. Am. Chem. Soc.* **119**, 27 – 37
124. Muller-Steffner, H. M., Augustin, A., and Schuber, F. (1996) *J. Biol. Chem.* **271**, 23967 – 23972
125. De Flora, A., Guida, L., Franco, L., Zocchi, E., Bruzzone, S., Benatti, U., Damonte, G., and Lee, H. C. (1997) *J. Biol. Chem.* **272**, 12945 – 12951
126. Aarhus, R., Graeff, R. M., Dickey, D. M., Walseth, T. F., and Lee, H. C. (1995) *J. Biol. Chem.* **270**, 30327 – 30333
127. Chini, E. N., Beers, K. W., and Dousa, T. P. (1995) *J. Biol. Chem.* **270**, 3216 – 3223
128. Oppenheimer, N. J. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical and Medical Aspects*, Vol. A (Dolphin, D., Avramovic, O., and Poulson, R., eds), pp. 323 – 365, Wiley (Interscience), New York
129. Everse, J., Anderson, B., and You, K.-S., eds. (1982) *The Pyridine Nucleotide Coenzymes*, Academic Press, New York
130. Blankenhorn, G., and Moore, E. G. (1980) *J. Am. Chem. Soc.* **102**, 1092 – 1098
131. Bernofsky, C. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical and Medical Aspects*, Vol. B (Dolphin, D., Avramovic, O., and Poulson, R., eds), pp. 105 – 172, Wiley (Interscience), New York
132. Chaykin, S. (1967) *Ann. Rev. Biochem.* **36, (I)**, 149 – 170
133. Everse, J., and Kaplan, N. O. (1973) *Adv. Enzymol.* **37**, 61 – 133

133a. Rozwarski, D. A., Grant, G. A., Barton, D. H. R., Jacobs, W. R. J., and Sacchettini, J. C. (1998) *Science* **279**, 98 – 102

133b. Parikh, S., Moynihan, D. P., Xiao, G., and Tonge, P. J. (1999) *Biochemistry* **38**, 13623 – 13634

133c. McMurry, L. M., Oethinger, M., and Levy, S. B. (1998) *Nature (London)* **394**, 531 – 532

133d. Roujeinikova, A., and 14 other authors. (1999) *J. Mol. Biol.* **294**, 527 – 535

133e. Parikh, S. L., Xiao, G., and Tonge, P. J. (2000) *Biochemistry* **39**, 7645 – 7650

134. Johnson, R. W., Marschner, T. M., and Oppenheimer, N. J. (1988) *J. Am. Chem. Soc.* **110**, 2257 – 2263
135. Oppenheimer, N. J. (1973) *Biochem. Biophys. Res. Commun.* **50**, 683 – 690
136. Acheson, S. A., Kirkman, H. N., and Wolfenden, R. (1988) *Biochemistry* **27**, 7371 – 7375
137. Marshall, J. L., Booth, J. E., and Williams, J. W. (1984) *J. Biol. Chem.* **259**, 3033 – 3036
138. Bernofsky, C., and Wanda, S.-Y. C. (1982) *J. Biol. Chem.* **257**, 6809 – 6817
139. Woenckhaus, C. H. (1974) in *Topics in Current Chemistry (Fortschritte der chemischen Forschung*, Vol. 52 (Boschke, F. L., ed), pp. 209 – 233, Springer-Verlag, Berlin
140. Anderson, B. M. (1982) in *The Pyridine Nucleotide Coenzymes* (Everse, J., Anderson, B., and You, K.-S., eds), Academic Press, New York
141. Woenckhaus, C., and Jeck, R. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical and Medical Aspects*, Vol. A (Dolphin, D., Avramovic, O., and Poulson, R., eds), pp. 449 – 568, Wiley (Interscience), New York
142. Cohen, B. E., Stoddard, B. L., and Koshland, D. E., Jr. (1997) *Biochemistry* **36**, 9035 – 9044
143. Hemmerich, P. (1976) in *Progress in the Chemistry of Organic Natural Products*, Vol. 33 (Herz, W., Grisebach, H., and Kirby, G. W., eds), pp. 451 – 527, Springer-Verlag, New York
144. Fox, K. M., and Karplus, P. A. (1999) *J. Biol. Chem.* **274**, 9357–9362
145. Niino, Y. S., Chakraborty, S., Brown, B. J., and Massey, V. (1995) *J. Biol. Chem.* **270**, 1983 – 1991
146. Meighen, E. A. (1993) *FASEB J.* **7**, 1016 – 1022
147. Fisher, A. J., Raushel, F. M., Baldwin, T. O., and Rayment, I. (1995) *Biochemistry* **34**, 6581 – 6586
148. Van den Berghe-Snorek, S., and Stankovich, M. T. (1984) *J. Am. Chem. Soc.* **106**, 3685 – 3687
149. Miura, R., Nishina, Y., Sato, K., Fujii, S., Kuroda, K., and Shiga, K. (1993) *J. Biochem.* **113**, 106 – 113
150. Lim, L. W., Shamala, N., Mathews, F. S., Steenkamp, D. J., Hamlin, R., and Xuong, N. (1986) *J. Biol. Chem.* **261**, 15140 – 15146
151. Massey, V. (1995) *FASEB J.* **9**, 473 – 475
152. Müller, F. (1991) *Chemistry and Biochemistry of Flavoenzymes*, Vol. I, CRC Press, Boca Raton, Florida
153. Curti, B., Ronchi, S., and Zanetti, G., eds. (1990) *Flavins and Flavoproteins*, Walter de Gruyter, Berlin

153a. Fraaije, M. W., and Mattevi, A. (2000) *Trends Biochem. Sci.* **25**, 126 – 132

154. Kohen, A., Jonsson, T., and Klinman, J. P. (1997) *Biochemistry* **36**, 2603 – 2611

154a. Su, Q., and Klinman, J. P. (1999) *Biochemistry* **38**, 8572 – 8581

155. Bourdillon, C., Demaille, C., Moiroux, J., and Savéant, J.-M. (1993) *J. Am. Chem. Soc.* **115**, 2 – 10
156. Li, J., Vrielink, A., Brick, P., and Blow, D. M. (1993) *Biochemistry* **32**, 11507 – 11515
157. Lindqvist, Y., and Brändén, C.-I. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6855 – 6859
158. Lindqvist, Y., and Brändén, C.-I. (1989) *J. Biol. Chem.* **264**, 3624 – 3628
159. Stenberg, K., and Lindqvist, Y. (1997) *Protein Sci.* **6**, 1009 – 1015
160. Yagi, K., and Ozawa, T. (1989) *Biochim. Biophys. Acta.* **1000**, 203 – 206

160a. Curti, B., Ronchi, S., and Simonetta, M. P. (1990) in *Chemistry and Biochemistry of Flavoenzymes*, Vol. III (Müller, F., ed), pp. 69 – 94, CRC Press, Boca Raton, Florida

References

161. Mizutani, H., Miyahara, I., Hirotsu, K., Nishina, Y., Shiga, K., Setoyama, C., and Miura, R. (1996) *J. Biochem.* **120**, 14–17
162. Mattevi, A., Vanoni, M. A., Todone, F., Rizzi, M., Teplyakov, A., Coda, A., Bolognesi, M., and Curti, B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7496–7501
163. Vanoni, M. A., Cosma, A., Mazzeo, D., Mattevi, A., Todone, F., and Curti, B. (1997) *Biochemistry* **36**, 5624–5632
164. Tan, A. K., and Ramsay, R. R. (1993) *Biochemistry* **32**, 2137–2143
165. Woo, J. C. G., and Silverman, R. B. (1995) *J. Am. Chem. Soc.* **117**, 1663–1664
166. Walker, M. C., and Edmondson, D. E. (1994) *Biochemistry* **33**, 7088–7098
166a. Pawelek, P. D., Cheah, J., Coulombe, R., Macheroux, P., Ghisla, S., and Vrielink, A. (2000) *EMBO J.* **19**, 4204–4215
167. Barber, M. J., Neame, P. J., Lim, L. W., White, S., and Mathews, F. S. (1992) *J. Biol. Chem.* **267**, 6611–6619
168. Trickey, P., Basran, J., Lian, L.-Y., Chen, Z.-w, Barton, J. D., Sutcliffe, M. J., Scrutton, N. S., and Mathews, F. S. (2000) *Biochemistry* **39**, 7678–7688
168a. Jang, M.-H., Basran, J., Scrutton, N. S., and Hille, R. (1999) *J. Biol. Chem.* **274**, 13147–13154
169. Falzon, L., and Davidson, V. L. (1996) *Biochemistry* **35**, 12111–12118
170. Lindqvist, Y., Brändén, C.-I., Mathews, F. S., and Lederer, F. (1991) *J. Biol. Chem.* **266**, 3198–3207
171. Tegoni, M., and Cambillau, C. (1994) *Protein Sci.* **3**, 303–313
172. Gondry, M., and Lederer, F. (1996) *Biochemistry* **35**, 8587–8594
173. Tegoni, M., Gervais, M., and Desbois, A. (1997) *Biochemistry* **36**, 8932–8946
173a. Lehoux, I. E., and Mitra, B. (2000) *Biochemistry* **39**, 10055–10065
174. Schaller, R. A., Mohsen, A.-W. A., Vockley, J., and Thorpe, C. (1997) *Biochemistry* **36**, 7761–7768
175. Rudik, I., Ghisla, S., and Thorpe, C. (1998) *Biochemistry* **37**, 8437–8445
176. Mancini-Samuelson, G. J., Kieweg, V., Sabaj, K. M., Ghisla, S., and Stankovich, M. T. (1998) *Biochemistry* **37**, 14605–14612
176a. Srivastava, D. K., and Peterson, K. L. (1998) *Biochemistry* **37**, 8446–8456
177. Mohsen, A.-W. A., and Vockley, J. (1995) *Biochemistry* **34**, 10146–10152
177a. Ackrell, B., McIntire, B., and Vessey, D. (2000) *Trends Biochem. Sci.* **25**, 9–10
178. Rétey, J., Seibl, J., Arigoni, D., Cornforth, J. W., Ryback, G., Zeylemaker, W. P., and Veeger, C. (1970) *Eur. J. Biochem.* **14**, 232–242
179. Birch-Machin, M. A., Farnsworth, L., Ackrell, B. A. C., Cochran, B., Jackson, S., Bindoff, L. A., Aitken, A., Diamond, A. G., and Turnbull, D. M. (1992) *J. Biol. Chem.* **267**, 11553–11558
180. Schmidt, D. M., Saghbini, M., and Scheffler, I. E. (1992) *Biochemistry* **31**, 8442–8448
181. Sucheta, A., Ackrell, B. A. C., Cochran, B., and Armstrong, F. A. (1992) *Nature (London)* **356**, 361–362
182. Thieme, R., Pai, E. F., Schirmer, R. H., and Schulz, G. E. (1981) *J. Mol. Biol.* **152**, 763–782
183. Sled, V. D., Rudnitzky, N. I., Hatefi, Y., and Ohnishi, T. (1994) *Biochemistry* **33**, 10069–10075
184. Westenberg, D. J., Gunsalus, R. P., Ackrell, B. A. C., Sices, H., and Cecchini, G. (1993) *J. Biol. Chem.* **268**, 815–822
185. Hägerhäll, C., Fridén, H., Aasa, R., and Hederstedt, L. (1995) *Biochemistry* **34**, 11080–11089
186. Van Hellemond, J. J., Klockiewicz, M., Gaasenbeek, C. P. H., Roos, M. H., and Tielens, A. G. M. (1995) *J. Biol. Chem.* **270**, 31065–31070
187. Waksman, G., Krishna, T. S. R., Williams, J., CH, and Kuriyan, J. (1994) *J. Mol. Biol.* **236**, 800–816
188. Mulrooney, S. B., and Williams, C. H. (1994) *Biochemistry* **33**, 3148–3154
189. Arscott, L. D., Gromer, S., Schirmer, R. H., Becker, K., and Williams, C. H., Jr. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3621–3626
190. Lennon, B. W., and Williams, C. H., Jr. (1995) *Biochemistry* **34**, 3670–3677
191. Pai, E. F., and Schulz, G. E. (1983) *J. Biol. Chem.* **258**, 1752–1757
192. Karplus, P. A., and Schulz, G. E. (1987) *J. Mol. Biol.* **195**, 701–729
193. Mittl, P. R. E., and Schulz, G. E. (1994) *Protein Sci.* **3**, 799–809
194. Zhang, Y., Bond, C. S., Bailey, S., Cunningham, M. L., Fairlamb, A. H., and Hunter, W. N. (1996) *Protein Sci.* **5**, 52–61
195. Walsh, C., Bradley, M., and Nadeau, K. (1991) *Trends Biochem. Sci.* **16**, 305–309
196. Zheng, R., Cenas, N., and Blanchard, J. S. (1995) *Biochemistry* **34**, 12697–12703
197. Krauth-Siegel, R. L., and Schöneck, R. (1995) *FASEB J.* **9**, 1138–1146
198. Brown, N. L. (1985) *Trends Biochem. Sci.* **10**, 400–403
199. Schiering, N., Kabsch, W., Moore, M. J., Distefano, M. D., Walsh, C. T., and Pai, E. F. (1991) *Nature (London)* **352**, 168–172
199a. Engst, S., and Miller, S. M. (1999) *Biochemistry* **38**, 3519–3529
200. Serre, L., Vellieux, F. M. D., Gomez-Moreno, M. M. C., Fontecilla-Camps, J. C., and Frey, M. (1996) *J. Mol. Biol.* **263**, 20–39
201. Aliverti, A., Bruns, C. M., Pandini, V. E., Karplus, P. A., Vanoni, M. A., Curti, B., and Zanetti, G. (1995) *Biochemistry* **34**, 8371–8379
202. Ermler, U., Siddiqui, R. A., Cramm, R., and Friedrich, B. (1995) *EMBO J.* **14**, 6067–6077
203. Ingelman, M., Bianchi, V., and Eklund, H. (1997) *J. Mol. Biol.* **268**, 147–157
204. Menz, R. I., and Day, D. A. (1996) *J. Biol. Chem.* **271**, 23117–23120
205. Takano, S., Yano, T., and Yagi, T. (1996) *Biochemistry* **35**, 9120–9127
205a. Morandi, P., Valzasina, B., Colombo, C., Curti, B., and Vanoni, M. A. (2000) *Biochemistry* **39**, 727–735
206. Fieschi, F., Nivière, V., Frier, C., Décout, J.-L., and Fontecave, M. (1995) *J. Biol. Chem.* **270**, 30392–30400
206a. Ingelman, M., Ramaswamy, S., Nivière, V., Fontecave, M., and Eklund, H. (1999) *Biochemistry* **38**, 7040–7049
207. Tanner, J. J., Lei, B., Tu, S.-C., and Krause, K. L. (1996) *Biochemistry* **35**, 13531–13539
208. Porter, T. D. (1991) *Trends Biochem. Sci.* **16**, 154–158
209. Vang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kin, J.-J. P. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8411–8416
209a. Uhlmann, H., and Bernhardt, R. (1995) *J. Biol. Chem.* **270**, 29959–29966
209b. Ziegler, G. A., and Schulz, G. E. (2000) *Biochemistry* **39**, 10986–10995
210. Kobayashi, K., Tagawa, S., Sano, S., and Asada, K. (1995) *J. Biol. Chem.* **270**, 27551–27554
211. Sano, S., Miyake, C., Mikami, B., and Asada, K. (1995) *J. Biol. Chem.* **270**, 21354–21361
212. Tedeschi, G., Chen, S., and Massey, V. (1995) *J. Biol. Chem.* **270**, 1198–1204
213. Li, R., Bianchet, M. A., Talalay, P., and Amzel, L. M. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8846–8850
214. Chen, S., Clarke, P. E., Martino, P. A., Dang, P. S. K., Yeh, C.-H., Lee, T. D., Prochaska, H. J., and Talalay, P. (1994) *Protein Sci.* **3**, 1296–1304
215. Claiborne, A., Ross, R. P., and Parsonage, D. (1992) *Trends Biochem. Sci.* **17**, 183–186
216. Parsonage, D., and Claiborne, A. (1995) *Biochemistry* **34**, 435–441
217. Yeh, J. I., Claiborne, A., and Hol, W. G. J. (1996) *Biochemistry* **35**, 9951–9957
218. Poole, L. B. (1996) *Biochemistry* **35**, 65–75
219. Mewies, M., McIntire, W. S., and Scrutton, N. S. (1998) *Protein Sci.* **7**, 7–20
220. Brandsch, R., and Bichler, V. (1991) *J. Biol. Chem.* **266**, 19056–19062
221. Willie, A., Edmondson, D. E., and Jorns, M. S. (1996) *Biochemistry* **35**, 5292–5299
222. Kenney, W. C., Singer, T. P., Fukuyama, M., and Miyake, Y. (1979) *J. Biol. Chem.* **254**, 4689–4690
223. Kim, J., Fuller, J. H., Kuusk, V., Cunane, L., Chen, Z.-w, Mathews, F. S., and McIntire, W. S. (1995) *J. Biol. Chem.* **270**, 31202–31209
224. Mewies, M., Basran, J., Packman, L. C., Hille, R., and Scrutton, N. S. (1997) *Biochemistry* **36**, 7162–7168
225. Ghisla, S., and Mayhew, S. G. (1976) *Eur. J. Biochem.* **63**, 373–390
226. Hartman, H. A., Edmondson, D. E., and McCormick, D. B. (1992) *Anal. Biochem.* **202**, 348–355
227. Eirich, L. D., Vogels, G. D., and Wolfe, R. S. (1979) *J. Bacterial.* **140**, 20–27
228. Jacobson, F., and Walsh, C. (1984) *Biochemistry* **23**, 979–988
229. Jones, J. B., and Stadtman, T. C. (1980) *J. Biol. Chem.* **255**, 1049–1053
230. Sancar, A. (1994) *Biochemistry* **33**, 2–9
231. Otani, S., Takatsu, M., Nakano, M., Kasai, S., Miura, R., and Matsui, K. (1974) *The Journal of Antibiotics* **27**, 88–89
232. Murthy, Y. V. S. N., and Massey, V. (1995) *J. Biol. Chem.* **270**, 28586–28594
233. Ghisla, S., and Massey, V. (1989) *Eur. J. Biochem.* **2**, 243–289
234. Powell, M. F., and Bruice, T. C. (1983) *J. Am. Chem. Soc.* **105**, 1014–1021
235. Fox, J. L., and Tolin, G. (1966) *Biochemistry* **5**, 3865–3872
236. Lee, I.-S. H., Ostovíc, D., and Kreevoy, M. (1988) *J. Am. Chem. Soc.* **110**, 3989–3993
237. Brustlein, M., and Bruice, T. C. (1972) *J. Am. Chem. Soc.* **94**, 6548–6549
238. Jorns, M. S., and Hersch, L. B. (1975) *J. Biol. Chem.* **250**, 3620–3628
239. Ghisla, S., Thorpe, C., and Massey, V. (1984) *Biochemistry* **23**, 3154–3161
240. Kim, J.-J. P., Wang, M., and Paschke, R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7523–7527
241. Dakoji, S., Shin, I., Becker, D. F., Stankovich, M. T., and Liu, H.-w. (1996) *J. Am. Chem. Soc.* **118**, 10971–10979
242. Tompkins, L. S., and Falkow, S. (1995) *Science* **267**, 1621–1622
242a. Vock, P., Engst, S., Eder, M., and Ghisla, S. (1998) *Biochemistry* **37**, 1848–1860
243. Johnson, B. D., Mancini-Samuelson, G. J., and Stankovich, M. T. (1995) *Biochemistry* **34**, 7047–7055
244. Thorpe, C. (1989) *Trends Biochem. Sci.* **14**, 148–151
245. Djordjevic, S., Pace, C. P., Stankovich, M. T., and Kim, J.-J. P. (1995) *Biochemistry* **34**, 2163–2171
246. Porter, D. J. T., Voet, J. G., and Bright, H. J. (1973) *J. Biol. Chem.* **248**, 4400–4416
247. Alston, T. A., Porter, D. J. T., and Bright, H. J. (1983) *J. Biol. Chem.* **258**, 1136–1141
248. Gadda, G., Edmondson, R. D., Russell, D. H., and Fitzpatrick, P. F. (1997) *J. Biol. Chem.* **272**, 5563–5570
249. Lehoux, I., and Mitra, B. (1997) *FASEB J.* **11**, A889

250. Ghisla, S., and Massey, V. (1980) *J. Biol. Chem.* **255**, 5688–5696

251. Hamilton, G. H. (1971) *Prog. Bioorg. Chem.* **1**, 83–157

252. Walsh, C. T., Krodel, E., Massey, V., and Abeles, R. H. (1973) *J. Biol. Chem.* **248**, 1946–1951

253. Ghisla, S., and Massey, V. (1989) *Eur. J. Biochem.* **181**, 1–17

254. Curti, B., Ronchi, S., and Simonetta, M. P. (1992) in *Chemistry and Biochemistry of Flavoenzymes* (Müller, F., ed), pp. 69–94, CRC Press, Boca Raton, Florida

255. Lederer, F. (1992) *Protein Sci.* **1**, 540–548

256. Miura, R., and Miyake, Y. (1988) *Bioorg. Chem.* **16**, 97–110

257. Todone, F., Vanoni, M. A., Mozzarelli, A., Bolognesi, M., Coda, A., Curti, B., and Mattevi, A. (1997) *Biochemistry* **36**, 5853–5860

258. Rietveld, P., Arscott, L. D., Berry, A., Scrutton, N. S., Deonarain, M. P., Perham, R. N., and Williams, C. H. (1994) *Biochemistry* **33**, 13888–13895

258a. Krauth-Siegel, R. L., Arscott, L. D., Schönleben-Janas, A., Schirmer, R. H., and Williams, C. H., Jr. (1998) *Biochemistry* **37**, 13968–13977

259. Williams, C. H., Jr. (1995) *FASEB J.* **9**, 1267–1276

260. Barman, B. G., and Tollin, G. (1972) *Biochemistry* **11**, 4760–4765

261. Su, Y., and Tripathi, G. N. R. (1994) *J. Am. Chem. Soc.* **116**, 4405–4407

262. Zheng, Y.-J., and Ornstein, R. L. (1996) *J. Am. Chem. Soc.* **118**, 9402–9408

263. Müller, F., Hemmerich, P., and Ehrenberg, A. (1971) in *Flavins and Flavoproteins* (Kamin, H., ed), pp. 107–180, Univ. Park Press, Baltimore, Maryland

264. Müller, F., Eriksson, L. E. G., and Ehrenberg, A. (1970) *Eur. J. Biochem.* **12**, 93–103

265. Miura, R., Setoyama, C., Nishina, Y., Shiga, K., Mizutani, H., Miyahara, I., and Hirotsu, K. (1997) *FASEB J.* **11**, A1306

265a. Miura, R., Setoyama, C., Nishina, Y., Shiga, K., Mizutani, H., Miyahara, I., and Hirotsu, K. (1977) *J. Biochem.* **122**, 825–833

266. Smith, W. W., Burnett, R. M., Darling, G. D., and Ludwig, M. L. (1977) *J. Mol. Biol.* **117**, 195–225

267. Chang, F.-C., and Swenson, R. P. (1999) *Biochemistry* **38**, 7168–7176

268. Zhou, Z., and Swenson, R. P. (1996) *Biochemistry* **35**, 15980–15988

269. Ludwig, M. L., Pattridge, K. A., Metzger, A. L., and Dixon, M. M. (1997) *Biochemistry* **36**, 1259–1280

269a. Lostao, A., Gómez-Moreno, C., Mayhew, S. G., and Sancho, J. (1997) *Biochemistry* **36**, 14334–14344

270. Hoover, D. M., Drennan, C. L., Metzger, A. L., Osborne, C., Weber, C. H., Pattridge, K. A., and Ludwig, M. L. (1999) *J. Mol. Biol.* **294**, 725–743

271. Moonen, C. T. W., and Müller, F. (1982) *Biochemistry* **21**, 408–414

272. Jorns, M. S., Sancar, G. B., and Sancar, A. (1984) *Biochemistry* **23**, 2673–2679

273. McKean, M. C., Beckman, J. D., and Frerman, F. E. (1983) *J. Biol. Chem.* **258**, 1866–1870

274. Roberts, D. L., Salazar, D., Fulmer, J. P., Frerman, F. E., and Kim, J.-J. P. (1999) *Biochemistry* **38**, 1977–1989

275. Huang, L., Rohlfs, R. J., and Hille, R. (1995) *J. Biol. Chem.* **270**, 23958–23965

276. Balasubramanian, S., Coggins, J. R., and Abell, C. (1995) *Biochemistry* **34**, 341–348

277. Macheroux, P., Petersen, J., Bornemann, S., Lowe, D. J., and Thorneley, R. N. F. (1996) *Biochemistry* **35**, 1643–1652

278. Yu, M. W., and Fritchie, C. J. J. (1975) *J. Biol. Chem.* **250**, 946–951

279. Kenney, W. C., and Singer, T. P. (1977) *J. Biol. Chem.* **252**, 4767–4772

280. Ostrowski, J., Barber, M. J., Rueger, D. C., Miller, B. E., Siegel, L. M., and Kredich, N. M. (1989) *J. Biol. Chem.* **264**, 15796–15808

281. Covès, J., Zeghouf, M., Macherel, D., Guigliarelli, B., Asso, M., and Fontecave, M. (1997) *Biochemistry* **36**, 5921–5928

282. Campbell, W. H., and Kinghorn, J. R. (1990) *Trends Biochem. Sci.* **15**, 315–319

283. Hyde, G. E., Crawford, N. M., and Campbell, W. H. (1991) *J. Biol. Chem.* **266**, 23542–23547

283a. Terao, M., Kurosaki, M., Saltini, G., Demontis, S., Marini, M., Salmona, M., and Garattini, E. (2000) *J. Biol. Chem.* **275**, 30690–30700

284. Bruice, T. C. (1980) *Acc. Chem. Res.* **13**, 256–262

285. Entsch, B., Palfey, B. A., Ballou, D. P., and Massey, V. (1991) *J. Biol. Chem.* **266**, 17341–17349

286. Gatti, D. L., Entsch, B., Ballou, D. P., and Ludwig, M. L. (1996) *Biochemistry* **35**, 567–578

287. Schreuder, H. A., Hol, W. G. J., and Drenth, J. (1990) *Biochemistry* **29**, 3101–3108

288. Yamasaki, M., and Yamano, T. (1973) *Biochem. Biophys. Res. Commun.* **51**, 612–619

289. Merényi, G., and Lind, J. (1991) *J. Am. Chem. Soc.* **113**, 3146–3153

290. Entsch, B., and van Berkel, W. J. H. (1995) *FASEB J.* **9**, 476–483

291. Snell, E. E., and Broquist, H. P. (1949) *Arch. Biochem. Biophys.* **23**, 326–328

291a. Reed, L. J. (1998) *Protein Sci.* **7**, 220–224

292. Reed, L. J., DeBusk, B. G., Gunsalus, I. C., and Hornberger, C. S. (1951) *Science* **114**, 93–94

293. Jukes, T. H. (1997) *Protein Sci.* **6**, 254–256

294. Schmidt, U., Grafen, P., Altland, K., and Goedde, H. W. (1969) *Adv. Enzymol.* **32**, 432–469

295. Reed, K. E., Morris, T. W., and Cronan, J. E., Jr. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3720–3724

296. Pares, S., Cohen-Addad, C., Sieker, L., Neuburger, M., and Douce, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4850–4853

296a. Morris, T. W., Reed, K. E., and Cronan, J. E., Jr. (1994) *J. Biol. Chem.* **269**, 16091–16100

297. Roche, T. E., and Patel, M. E., eds. (1990) *Alpha-Keto Acid Dehydrogenase Complexes Organization, Regulation, and Biomedical Ramifications*, Vol. 573, New York Academy of Sciences, New York

298. Patel, M. S., and Roche, T. E. (1990) *FASEB J.* **4**, 3224–3233

299. Stoops, J. K., Cheng, R. H., Yazdi, M. A., Maeng, C.-Y., Schroeter, J. P., Klueppelberg, U., Kolodziej, S. J., Baker, T. S., and Reed, L. J. (1997) *J. Biol. Chem.* **272**, 5757–5764

300. Yi, J., Nemeria, N., McNally, A., Jordan, F., Machado, R. S., and Guest, J. R. (1996) *J. Biol. Chem.* **271**, 33192–33200

301. Berg, A., Vervoort, J., and de Kok, A. (1996) *J. Mol. Biol.* **261**, 432–442

302. Wynn, R. M., Davie, J. R., Zhi, W., Cox, R. P., and Chuang, D. T. (1994) *Biochemistry* **33**, 8962–8968

302a. Wynn, R. M., Davie, J. R., Chuang, J. L., Cote, C. D., and Chuang, D. T. (1998) *J. Biol. Chem.* **273**, 13110–13118

303. Mattevi, A., Obmolova, G., Kalk, K. H., van Berkel, W. J. H., and Hol, W. G. J. (1993) *J. Mol. Biol.* **230**, 1200–1215

304. Liu, T.-C., Korotchkina, L. G., Hyatt, S. L., Vettakkorumakankav, N. N., and Patel, M. S. (1995) *J. Biol. Chem.* **270**, 15545–15550

305. Guan, Y., Rawsthorne, S., Scofield, G., Shaw, P., and Doonan, J. (1995) *J. Biol. Chem.* **270**, 5412–5417

306. Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A., and Hol, W. G. J. (1992) *Science* **255**, 1544–1550

306a. Mattevi, A., Obmolova, G., Kalk, K. H., Westphal, A. H., de Kok, A., and Hol, W. G. J. (1993) *J. Mol. Biol.* **230**, 1183–1199

306b. Izard, T., AEvarsson, A., Allen, M. D., Westphal, A. H., Perhan, R. N., de Kok, A., and Hol, W. G. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1240–1245

307. Reed, L. J., and Hackert, M. L. (1990) *J. Biol. Chem.* **265**, 8971–8974

307a. Knapp, J. E., Mitchell, D. T., Yazdi, M. A., Ernst, S. R., Reed, L. J., and Hackert, M. L. (1998) *J. Mol. Biol.* **280**, 655–668

308. Dardel, F., Davis, A. L., Laue, E. D., and Perham, R. N. (1993) *J. Mol. Biol.* **229**, 1037–1048

309. Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T., and Guest, J. R. (1995) *J. Mol. Biol.* **248**, 328–343

310. Mattevi, A., Obmolova, G., Kalk, K. H., Teplyakov, A., and Hol, W. G. J. (1993) *Biochemistry* **32**, 3887–3901

311. Hendle, J., Mattevi, A., Westphal, A. H., Spee, J., de Kok, A., Teplyakov, A., and Hol, W. G. J. (1995) *Biochemistry* **34**, 4287–4298

312. Ricaud, P. M., Howard, M. J., Roberts, E. L., Broadhurst, R. W., and Perham, R. N. (1996) *J. Mol. Biol.* **264**, 179–190

313. Maeng, C.-Y., Yazdi, M. A., Niu, X.-D., Lee, H. Y., and Reed, L. J. (1994) *Biochemistry* **33**, 13801–13807

314. McCartney, R. G., Sanderson, S. J., and Lindsay, J. G. (1997) *Biochemistry* **36**, 6819–6826

315. Yang, Y.-S., and Frey, P. A. (1986) *Biochemistry* **25**, 8173–8178

315a. Pan, K., and Jordan, F. (1998) *Biochemistry* **37**, 1357–1364

315b. Hennig, J., Kern, G., Neef, H., Spinka, M., Bisswanger, H., and Hübner, G. (1997) *Biochemistry* **36**, 15772 – 15779

315c. Kee, K., Niu, L., and Henderson, E. (1998) *Biochemistry* **37**, 4224 – 4234

316. Han, D., Handelman, G., Marcocci, L., Sen, C. K., Roy, S., Kobuchi, H., Tritschler, H. J., Flohé, L., and Packer, L. (1997) *BioFactors* **6**, 321–338

317. Muller, Y. A., Schumacher, G., Rudolph, R., and Schulz, G. E. (1994) *J. Mol. Biol.* **237**, 315–335

318. Kerscher, L., and Oesterhelt, D. (1982) *Trends Biochem. Sci.* **7**, 371–374

319. Muller, Y. A., and Schulz, G. E. (1993) *Science* **259**, 965–967

319a. Tittmann, K., Golbik, R., Ghisla, S., and Hübner, G. (2000) *Biochemistry* **39**, 10747–10754

320. Menon, S., and Ragsdale, S. W. (1997) *Biochemistry* **36**, 8484–8494

320a. Bouchev, V. F., Furdui, C. M., Menon, S., Muthukumaran, R. B., Ragsdale, S. W., and McCracken, J. (1999) *J. Am. Chem. Soc.* **121**, 3724–3729

321. Smith, E. T., Blamey, J. M., and Adams, M. W. W. (1994) *Biochemistry* **33**, 1008–1016

322. Mai, X., and Adams, M. W. W. (1994) *J. Biol. Chem.* **269**, 16726–16732

323. Gehring, V., and Arnon, D. I. (1972) *J. Biol. Chem.* **247**, 6963–6969

324. Sun, W., Williams, C. H., Jr., and Massey, V. (1996) *J. Biol. Chem.* **271**, 17226–17233

325. Müh, U., Williams, C. H., and Massey, V. (1994) *J. Biol. Chem.* **269**, 7994–8000

326. Unkrig, V., Neugebauer, F. A., and Knappe, J. (1989) *Eur. J. Biochem.* **184**, 723–728

327. Maiorino, M., Chu, F. F., Ursini, F., Davies, K. J. A., Doroshow, J. H., and Esworthy, R. S. (1991) *J. Biol. Chem.* **266**, 7728–7732

328. Wong, K. K., Murray, B. W., Lewisch, S. A., Baxter, M. K., Ridky, T. W., Ulissi-DeMario, L., and Kozarich, J. W. (1993) *Biochemistry* **32**, 14102–14110

References

329. Frey, M., Rothe, M., Wagner, A. F. V., and Knappe, J. (1994) *J. Biol. Chem.* **269**, 12432–12437
330. Kraus, R. J., Foster, S. J., and Ganther, H. E. (1983) *Biochemistry* **22**, 5853–5858
331. Broderick, J. B., Duderstadt, R. E., Fernandez, D. C., Wojtuszewski, K., Henshaw, T. F., and Johnson, M. K. (1997) *J. Am. Chem. Soc.* **119**, 7390–7391
331a. Külzer, R., Pils, T., Kappl, R., Hüttermann, J., and Knappe, J. (1998) *J. Biol. Chem.* **273**, 4897–4903
332. Wu, W., Lieder, K. W., Reed, G. H., and Frey, P. A. (1995) *Biochemistry* **34**, 10532–10537
333. Wagner, A. F. V., Frey, M., Neugebauer, F. A., Schäfer, W., and Knappe, J. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 996–1000
333a. Gauld, J. W., and Eriksson, L. A. (2000) *J. Am. Chem. Soc.* **122**, 2035–2040
334. Parast, C. V., Wong, K. K., Lewisch, S. A., Kozarich, J. W., Peisach, J., and Magliozzo, R. S. (1995) *Biochemistry* **34**, 2393–2399
335. Parast, C. V., Wong, K. K., Kozarich, J. W., Peisach, J., and Magliozzo, R. S. (1995) *Biochemistry* **34**, 5712–5717
335a. Reddy, S. G., Wong, K. K., Parast, C. V., Peisach, J., Magliozzo, R. S., and Kozarich, J. W. (1998) *Biochemistry* **37**, 558–563
336. Mastropaolo, D., Camerman, A., and Camerman, N. (1980) *Science* **210**, 334–336
337. Bailey, L. B., ed. (1995) *Folate in Health and Disease*, Dekker, New York
338. Curtius, H.-C., Matasovic, A., Schoedon, G., Kuster, T., Guibaud, P., Giudici, T., and Blau, N. (1990) *J. Biol. Chem.* **265**, 3923–3930
339. Blakley, R. L., and Benkovic, S. J., eds. (1984–1986) *Folates and Pterins*, Wiley, New York (3 Vols.)
340. Hadorn, E. (1962) *Sci. Am.* **206**(April), 101–110
341. Ghosh, D. K., and Stuehr, D. J. (1995) *Biochemistry* **34**, 801–807
342. Witteveen, C. F. B., Giovanelli, J., and Kaufman, S. (1996) *J. Biol. Chem.* **271**, 4143–4147
343. Rembold, H., and Gyure, W. L. (1972) *Angew. Chem. Int. Ed. Engl.* **11**, 1061–1072
344. Jacobson, K. B., Dorsett, D., Pfleiderer, W., McCloskey, J. A., Sethi, S. K., Buchanan, M. V., and Rubin, I. B. (1982) *Biochemistry* **21**, 5700–5706
345. Wiederrecht, G. J., Paton, D. R., and Brown, G. M. (1984) *J. Biol. Chem.* **259**, 2195–2200
346. Cremer-Bartels, G., and Ebels, I. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2415–2418
347. Van Haastert, P. J. M., DeWitt, R. J. W., Grijpma, Y., and Konijn, T. M. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6270–6274
348. Irby, R. B., and Adair, J., WL. (1994) *J. Biol. Chem.* **269**, 23981–23987
349. Stiefel, E. I. (1996) *Science* **272**, 1599–1600
350. Escalante-Semerena, J. C., Rinehart, K. L., Jr., and Wolfe, R. S. (1984) *J. Biol. Chem.* **259**, 9447–9455
351. Wolfe, R. S. (1985) *Trends Biochem. Sci.* **10**, 396–399
352. Rouvière, P. E., and Wolfe, R. S. (1988) *J. Biol. Chem.* **263**, 7913–7916
353. White, R. H. (1993) *Biochemistry* **32**, 745–753
354. Hyatt, D. C., Maley, F., and Montfort, W. R. (1997) *Biochemistry* **36**, 4585–4594
355. Blakeley, R. L. (1984) in *Folates and Pterins*, Vol. I (Blakeley, P. L., and Bencovic, S. J., eds), pp. 191–253, Wiley, New York
356. Stover, P., and Schirch, V. (1993) *Trends Biochem. Sci.* **18**, 102–106
357. Lin, B.-F., and Shane, B. (1994) *J. Biol. Chem.* **269**, 9705–9713
357a. Cherest, H., Thomas, D., and Surdin-Kerjan, Y. (2000) *J. Biol. Chem.* **275**, 14056–14063
358. Keshavjee, K., Pyne, C., and Bognar, A. L. (1991) *J. Biol. Chem.* **266**, 19925–19929
359. Scott, J. M. (1976) *Biochem. Soc. Trans.* **4**, 845–850
360. Brody, T., Watson, J. E., and Stokstad, E. L. R. (1982) *Biochemistry* **21**, 276–282
361. Kim, D. W., Huang, T., Schirch, D., and Schirch, V. (1996) *Biochemistry* **35**, 15772–15783
362. Lipman, R. S. A., Bailey, S. W., Jarrett, J. T., Matthews, R. G., and Jorns, M. S. (1995) *Biochemistry* **34**, 11217–11220
363. Mathews, C. K. (1971) *Bacteriophage Biochemistry*, Van Nostrand-Reinhold, Princeton, New Jersey
364. Roth, B. (1986) *Fed. Proc.* **45**, 2765–2772
365. Schweitzer, B. I., Dicker, A. P., and Bertino, J. R. (1990) *FASEB J.* **4**, 2441–2452
366. Groom, C. R., Thillet, J., North, A. C. T., Pictet, R., and Geddes, A. J. (1991) *J. Biol. Chem.* **266**, 19890–19893
367. Blaney, J. M., Hansch, C., Silipo, C., and Vittonia, A. (1984) *Chem. Rev.* **84**, 333–407
368. Dale, G. E., Broger, C., D'Arcy, A., Hartman, P. G., DeHoogt, R., Jolidon, S., Kompis, I., Labhardt, A. M., Langen, H., Locher, H., Page, M. G. P., Stüber, D., Then, R. L., Wipf, B., and Oefner, C. (1997) *J. Mol. Biol.* **266**, 23–30
369. Sirawaraporn, W., Sathitkul, T., Sirawaraporn, R., Yuthavong, Y., and Santi, D. V. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1124–1129
370. Birdsall, B., Tendler, S. J. B., Arnold, J. R. P., Feeney, J., Griffin, R. J., Carr, M. D., Thomas, J. A., Roberts, G. C. K., and Stevens, M. F. G. (1990) *Biochemistry* **29**, 9660–9667
371. Jolivet, J., Cowan, K. H., Curt, G. A., Clendeninn, N. J., and Chabner, B. A. (1983) *N. Engl. J. Med.* **309**, 1094–1104
372. Federspiel, N. A., Beverley, S. M., Schilling, J. W., and Schimke, R. T. (1984) *J. Biol. Chem.* **259**, 9127–9140
373. Chu, E., Takimoto, C. H., Voeller, D., Grem, J. L., and Allegra, C. J. (1993) *Biochemistry* **32**, 4756–4760
374. Friesheim, J. H., and Matthews, D. A. (1984) in *Folate Antogonists as Therapeutic Agents*, Vol. 1 (Sirotnak, F. M., Burchall, J. J., Ensminger, W. D., and Montgomery, J. A., eds), pp. 69–131, Academic Press, Orlando, Florida
375. Fan, J., Vitols, K. S., and Huennekens, F. M. (1991) *J. Biol. Chem.* **266**, 14862–14865
376. Henderson, G. B., Hughes, T. R., and Saxena, M. (1994) *J. Biol. Chem.* **269**, 13382–13389
377. Thompson, P. D., and Freisheim, J. H. (1991) *Biochemistry* **30**, 8124–8130
378. Chen, L., Qi, H., Korenberg, J., Garrow, T. A., Choi, Y.-J., and Shane, B. (1996) *J. Biol. Chem.* **271**, 13077–13087
379. Roy, K., Mitsugi, K., and Sirotnak, F. M. (1997) *J. Biol. Chem.* **272**, 5587–5593
379a. Maziarz, K. M., Monaco, H. L., Shen, F., and Ratnam, M. (1999) *J. Biol. Chem.* **274**, 11086–11091
380. Park, H., Zhuang, P., Nichols, R., and Howell, E. E. (1997) *J. Biol. Chem.* **272**, 2252–2258
381. Sawaya, M. R., and Kraut, J. (1997) *Biochemistry* **36**, 586–603
382. Reyes, V. M., Sawaya, M. R., Brown, K. A., and Kraut, J. (1995) *Biochemistry* **34**, 2710–2723
383. Lee, H., Reyes, V. M., and Kraut, J. (1996) *Biochemistry* **35**, 7012–7020
384. Verma, C. S., Caves, L. S. D., Hubbard, R. E., and Roberts, G. C. K. (1997) *J. Mol. Biol.* **266**, 776–796
385. Cheung, H. T. A., Birdsall, B., Frenkiel, T. A., Chau, D. D., and Feeney, J. (1993) *Biochemistry* **32**, 6846–6854
386. Epstein, D. M., Benkovic, S. J., and Wright, P. E. (1995) *Biochemistry* **34**, 11037–11048
387. Meiering, E. M., and Wagner, G. (1995) *J. Mol. Biol.* **247**, 294–308
388. Cannon, W. R., Garrison, B. J., and Benkovic, S. J. (1997) *J. Am. Chem. Soc.* **119**, 2386–2395
389. Nakano, T., Spencer, H. T., Appleman, J. R., and Blakley, R. L. (1994) *Biochemistry* **33**, 9945–9952
390. Howell, E. E., Villafranca, J. E., Warren, M. S., Oatley, S. J., and Kraut, J. (1986) *Science* **231**, 1123–1128
391. Benkovic, S. J., Fierke, C. A., and Naylor, A. M. (1988) *Science* **239**, 1105–1110
392. Chen, Y.-Q., Kraut, J., Blakley, R. L., and Callender, R. (1994) *Biochemistry* **33**, 7023–7032
393. Basran, J., Casarotto, M. G., Barsukov, I. L., and Roberts, G. C. K. (1995) *Biochemistry* **34**, 2872–2882
394. Jeong, S.-S., and Gready, J. E. (1995) *Biochemistry* **34**, 3734–3741
395. Appling, D. R. (1991) *FASEB J.* **5**, 2645–2651
396. Stover, P. J., Chen, L. H., Suh, J. R., Stover, D. M., Keyomarsi, K., and Shane, B. (1997) *J. Biol. Chem.* **272**, 1842–1848
397. Rosenblatt, D. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. II (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3111–3128, McGraw-Hill, New York
398. Slieker, L. J., and Bencovic, S. J. (1984) *J. Am. Chem. Soc.* **106**, 1833–1838
399. Kallen, R. G., and Jencks, W. P. (1966) *J. Biol. Chem.* **241**, 5845–5850,5851–5863
400. Walker, J. L., and Oliver, D. J. (1986) *J. Biol. Chem.* **261**, 2214–2221
401. Kume, A., Koyata, H., Sakakibara, T., Ishiguro, Y., Kure, S., and Hiraga, K. (1991) *J. Biol. Chem.* **266**, 3323–3329
402. Pasternack, L. B., Laude, D. A., Jr., and Appling, D. R. (1992) *Biochemistry* **31**, 8713–8719
403. Kennard, O., and Salisbury, S. A. (1993) *J. Biol. Chem.* **268**, 10701–10704
404. Zieske, L. R., and Davis, L. (1983) *J. Biol. Chem.* **258**, 10355–10359
404a. Barber, R. D., and Donohue, T. J. (1998) *J. Mol. Biol.* **280**, 775–784
404b. Fernández, M. R., Biosca, J. A., Torres, D., Crosas, B., and Parés, X. (1999) *J. Biol. Chem.* **274**, 37869–37875
405. Mejillano, M. R., Jahansouz, H., Matsunaga, T. O., Kenyon, G. L., and Himes, R. H. (1989) *Biochemistry* **28**, 5136–5145
406. D'Ari, L., and Rabinowitz, J. C. (1991) *J. Biol. Chem.* **266**, 23953–23958
406a. Shen, B. W., Dyer, D. H., Huang, J.-Y., D'Ari, L., Rabinowitz, J., and Stoddard, B. L. (1999) *Protein Sci.* **8**, 1342–1349
407. Schmidt, A., Wu, H., MacKenzie, R. E., Chen, V. J., Bewly, J. R., Ray, J. E., Toth, J. E., and Cygler, M. (2000) *Biochemistry* **39**, 6325–6335
407a. Pawelek, P. D., and MacKenzie, R. E. (1998) *Biochemistry* **37**, 1109–1115
408. Wahls, W. P., Song, J. M., and Smith, G. R. (1993) *J. Biol. Chem.* **268**, 23792–23798
409. Klein, C., Chen, P., Arevalo, J. H., Stura, E. A., Marolewski, A., Warren, M. S., Benkovic, S. J., and Wilson, I. A. (1995) *J. Mol. Biol.* **249**, 153–175
409a. Greasley, S. E., Yamashita, M. M., Cai, H., Benkovic, S. J., Boger, D. L., and Wilson, I. A. (1999) *Biochemistry* **38**, 16783–16793
409b. Thoden, J. B., Firestine, S., Nixon, A., Benkovic, S. J., and Holden, H. M. (2000) *Biochemistry* **39**, 8791–8802
410. Caperelli, C. A., and Giroux, E. L. (1997) *Arch. Biochem. Biophys.* **341**, 98–103
411. Schmitt, E., Blanquet, S., and Mechulam, Y. (1996) *EMBO J.* **15**, 4749–4758

References

412. Kruschwitz, H. L., McDonald, D., Cossins, E. A., and Schirch, V. (1994) *J. Biol. Chem.* **269**, 28757 – 28763
413. Lutsenko, S., Daoud, S., and Kaplan, J. H. (1997) *J. Biol. Chem.* **272**, 5249 – 5255
414. Rossini, C., Attygalle, A. B., González, A., Smedley, S. R., Eisner, M., Meinwald, J., and Eisner, T. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6792 – 6797
415. Curthoys, N. P., and Rabinowitz, J. C. (1972) *J. Biol. Chem.* **247**, 1965 – 1971
415a. Krupenko, S. A., and Wagner, C. (1999) *J. Biol. Chem.* **274**, 35777 – 35784
416. Hardy, L. W., Graves, K. L., and Nalivaika, E. (1995) *Biochemistry* **34**, 8422 – 8432
417. Powers, S. G., and Snell, E. E. (1976) *J. Biol. Chem.* **251**, 3786 – 3793
418. Carreras, C. W., and Santi, D. V. (1995) *Ann. Rev. Biochem.* **64**, 721 – 762
419. Schiffer, C. A., Clifton, I. J., Davisson, V. J., Santi, D. V., and Stroud, R. M. (1995) *Biochemistry* **34**, 16279 – 16287
420. Birdsall, D. L., Finer-Moore, J., and Stroud, R. M. (1996) *J. Mol. Biol.* **255**, 522 – 535
421. Sage, C. R., Rutenber, E. E., Stout, T. J., and Stroud, R. M. (1996) *Biochemistry* **35**, 16270 – 16281
421a. Anderson, A. C., O'Neil, R. H., DeLano, W. L., and Stroud, R. M. (1999) *Biochemistry* **38**, 13829 – 13836
422. Huang, W., and Santi, D. V. (1997) *Biochemistry* **36**, 1869 – 1873
422a. Strop, P., Changchien, L., Maley, F., and Montfort, W. R. (1997) *Protein Sci.* **6**, 2504 – 2511
423. Knighton, E. R., Kan, C.-C., Howland, E., Janson, C. A., Hostomska, Z., Welsh, K. M., and Matthews, D. A. (1994) *Nature Struct. Biol.* **1**, 186 – 194
423a. Song, H. K., Sohn, S. H., and Suh, S. W. (1999) *EMBO J.* **18**, 1104 – 1113
424. Butler, M. M., Graves, K. L., and Hardy, L. W. (1994) *Biochemistry* **33**, 10521 – 10526
425. Graves, K. L., and Hardy, L. W. (1994) *Biochemistry* **33**, 13049 – 13056
426. Santi, D. V., and Hardy, L. W. (1987) *Biochemistry* **26**, 8599 – 8606
426a. Roje, S., Wang, H., McNeil, S. D., Raymond, R. K., Appling, D. R., Shachar-Hill, Y., Bohnert, H. J., and Hanson, A. D. (1999) *J. Biol. Chem.* **274**, 36089 – 36096
427. Green, J. M., Ballou, D. P., and Matthews, R. G. (1988) *FASEB J.* **2**, 42 – 47
428. Vanoni, M. A., Lee, S., Floss, H. G., and Matthews, R. G. (1990) *J. Am. Chem. Soc.* **112**, 3987 – 3992
429. Matthews, R. G. (1982) *Fed. Proc.* **41**, 2600 – 2604
430. te Brömmelstroet, B. W., Hensgens, C. M. H., Keltjens, J. T., van der Drift, C., and Vogels, G. D. (1990) *J. Biol. Chem.* **265**, 1852 – 1857
431. Clark, J. E., and Ljungdahl, L. G. (1984) *J. Biol. Chem.* **259**, 10845 – 10849
432. Zhao, S., Roberts, D. L., and Ragsdale, S. W. (1995) *Biochemistry* **34**, 15075 – 15083
433. Müller, V., Blaut, M., and Gottschalk, G. (1993) in *Methano-genesis: Ecology, Physiology, Biochemistry and Genetics* (Ferry, J. G., ed), pp. 360 – 406, Chapman and Hall, New York
434. Ferry, J. G., ed. (1993) *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics*, Chapman & Hall, New York
435. Thauer, R. K., Hedderich, R., and Fischer, R. (1993) in *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics* (Ferry, J. G., ed), pp. 209 – 252, Chapman and Hall, New York
436. Taylor, G. D., and Wolfe, R. S. (1974) *J. Biol. Chem.* **249**, 4879 –4885
437. Keltjens, J. T., Raemakers-Franken, P. C., and Vogels, G. D. (1993) in *Microbial Growth on C1 Compounds* (Murrell, J. C., and Kelly, D. P., eds), pp. 135 – 150, Intercept Ltd., Andover,UK
438. Leigh, J. A., Rinehart, K. L., Jr., and Wolfe, R. S. (1985) *Biochemistry* **24**, 995 – 999
439. Leigh, J. A., Rinehart, K. L., Jr., and Wolfe, R. S. (1984) *J. Am. Chem. Soc.* **106**, 3636 – 3640
440. White, R. H. (1989) *Biochemistry* **28**, 860 – 865
440a. Ermler, U., Merckel, M. C., Thauer, R. K., and Shima, S. (1997) *Structure* **5**, 635 – 646
440b. Shima, S., Warkentin, E., Grabarse, W., Sordel, M., Wicke, M., Thauer, R. K., and Ermler, U. (2000) *J. Mol. Biol.* **300**, 935 – 950
441. Mukhopadhyay, B., Purwantini, E., Pihl, T. D., Reeve, J. N., and Daniels, L. (1995) *J. Biol. Chem.* **270**, 2827 – 2832
441a. Allen, J. R., Clark, D. D., Krum, J. G., and Ensign, S. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8432 – 8437
442. Duine, J. A., and Frank, J. (1981) *Trends Biochem. Sci.* **6**, 278 – 280
443. Ohta, S., Fujita, T., and Tobari, J. (1981) *J. Biochem.* **90**, 205 – 213
444. Itoh, S., Ogino, M., Fukui, Y., Murao, H., Komatsu, M., Ohshiro, Y., Inoue, T., Kai, Y., and Kasai, N. (1993) *J. Am. Chem. Soc.* **115**, 9960 – 9967
445. de Jong, G. A. H., Caldeira, J., Sun, J., Jongejan, J. A., de Vries, S., Loehr, T. M., Moura, I., Moura, J. J. G., and Duine, J. A. (1995) *Biochemistry* **34**, 9451 – 9458
445a. Keitel, T., Diehl, A., Knaute, T., Stezowski, J. J., Höhne, W., and Görisch, H. (2000) *J. Mol. Biol.* **297**, 961 – 974
446. Matsushita, K., Shinagawa, E., Adachi, O., and Ameyama, M. (1989) *Biochemistry* **28**, 6276 – 6280
446a. Oubrie, A., Rozeboom, H. J., Kalk, K. H., Olsthoorn, A. J. J., Duine, J. A., and Dijkstra, B. W. (1999) *EMBO J.* **18**, 5187 – 5194
446b. Oubrie, A., Rozeboom, H. J., Kalk, K. H., Duine, J. A., and Dijkstra, B. W. (1999) *J. Mol. Biol.* **289**, 319 – 333
446c. Elias, M. D., Tanaka, M., Izu, H., Matsushita, K., Adachi, O., and Yamada, M. (2000) *J. Biol. Chem.* **275**, 7321 – 7326
446d. Dewanti, A. R., and Duine, J. A. (2000) *Biochemistry* **39**, 9384 – 9392
447. Xia, Z.-x, Dai, W.-w, Zhang, Y.-f, White, S. A., Boyd, G. D., and Mathews, F. S. (1996) *J. Mol. Biol.* **259**, 480 – 501
448. White, S., Boyd, G., Mathews, F. S., Xia, Z., Dai, W., Zhang, Y., and Davidson, V. L. (1993) *Biochemistry* **32**, 12955 – 12958
449. Ishida, T., Kawamoto, E., In, Y., Amano, T., Kanayama, J., Doi, M., Iwashita, T., and Nomoto, K. (1995) *J. Am. Chem. Soc.* **117**, 3278 – 3279
449a. Itoh, S., Kawakami, H., and Fukuzumi, S. (1998) *Biochemistry* **37**, 6562 – 6571
449b. Oubrie, A., Rozeboom, H. J., and Dijkstra, B. W. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11787 – 11791
449c. Zheng, Y.-J., and Bruice, T. C. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11881 – 11886
450. McIntire, W. S. (1994) *FASEB J.* **8**, 513 – 521
451. Anthony, C. (1996) *Biochem. J.* **320**, 697 – 711
452. Klinman, J. P., and Mu, D. (1994) *Ann. Rev. Biochem.* **63**, 299 – 344
453. Klinman, J. P. (1996) *J. Biol. Chem.* **271**, 27189 – 27192
454. Davidson, V. L. (1992) *Principles and Applications of Quinoproteins*, Dekker, New York
455. Janes, S. M., Mu, D., Wemmer, D., Smith, A. J., Kaur, S., Maltby, D., Burlingame, A. L., and Klinman, J. P. (1990) *Science* **248**, 981 – 987
456. Mu, D., Medzihradszky, K. F., Adams, G. W., Mayer, P., Hines, W. M., Burlingame, A. L., Smith, A. J., Cai, D., and Klinman, J. P. (1994) *J. Biol. Chem.* **269**, 9926 – 9932
457. Parsons, M. R., Convey, M. A., Wilmot, C. M., Yadav, K. D. S., Blakeley, V., Corner, A. S., Phillips, S. E. V., McPherson, M. J., and Knowles, P. F. (1995) *Structure* **3**, 1171 – 1184
458. Wilmot, C. M., Murray, J. M., Alton, G., Parsons, M. R., Convery, M. A., Blakeley, V., Corner, A. S., Palcic, M. M., Knowles, P. F., McPherson, M. J., and Phillips, S. E. V. (1997) *Biochemistry* **36**, 1608 – 1620
459. Moënne-Loccoz, P., Nakamura, N., Steinebach, V., Duine, J. A., Mure, M., Klinman, J. P., and Sanders-Loehr, J. (1995) *Biochemistry* **34**, 7020 – 7026
460. Parsons, M. R., Convery, M. A., Wilmot, C. M., Yadav, K. D., Blakley, V., Corner, A. S., Phillips, S. E. V., McPherson, J. J., and Knowles, P. F. (1995) *Structure* **3**, 1171 – 1184
461. Warncke, K., Babcock, G. T., Dooley, D. M., McGuirl, M. A., and McCracken, J. (1994) *J. Am. Chem. Soc.* **116**, 4028 – 4037
462. Choi, Y.-H., Matsuzaki, R., Fukui, T., Shimizu, E., Yorifuji, T., Sato, H., Ozaki, Y., and Tanizawa, K. (1995) *J. Biol. Chem.* **270**, 4712 – 4720
463. Choi, Y.-H., Matsuzaki, R., Suzuki, S., and Tanizawa, K. (1996) *J. Biol. Chem.* **271**, 22598 – 22603
463a. Schwartz, B., Green, E. L., Sanders-Loehr, J., and Klinman, J. P. (1998) *Biochemistry* **37**, 16591 – 16600
464. Medda, R., Padiglia, A., Pedersen, J. Z., Rotilio, G., Agrò, A. F., and Floris, G. (1995) *Biochemistry* **34**, 16375 – 16381
464a. Holt, A., Alton, G., Scaman, C. H., Loppnow, G. R., Szpacenko, A., Svendsen, I., and Palcic, M. M. (1998) *Biochemistry* **37**, 4946 – 4957
465. Wang, S. X., Mure, M., Medzihradszky, K. F., Burlingame, A. L., Brown, D. E., Dooley, D. M., Smith, A. J., Kagan, H. M., and Klinman, J. P. (1996) *Science* **273**, 1078 – 1084
465a. Wang, S. X., Nakamura, N., Mure, M., Klinman, J. P., and Sanders-Loehr, J. (1997) *J. Biol. Chem.* **272**, 28841 – 28844
466. Backes, G., Davidson, V. L., Huitema, F., Duine, J. A., and Sanders-Loehr, J. (1991) *Biochemistry* **30**, 9201 – 9210
466a. Zhu, Z., and Davidson, V. L. (1998) *J. Biol. Chem.* **273**, 14254 – 14260
467. McIntire, W. S., Wemmer, D. E., Chistoserdov, A., and Lidstrom, M. E. (1991) *Science* **252**, 817 – 824
468. Huizinga, E. G., van Zanten, B. A. M., Duine, J. A., Jongejan, J. A., Huitema, F., Wilson, K. S., and Hol, W. G. J. (1992) *Biochemistry* **31**, 9789 – 9795
469. Kuusk, V., and McIntire, W. S. (1994) *J. Biol. Chem.* **269**, 26136 – 26143
470. Hyun, Y.-L., and Davidson, V. L. (1995) *Biochemistry* **34**, 816 – 823
471. Chen, L., Durley, R., Poliks, B. J., Hamada, K., Chen, Z., Mathews, F. S., Davidson, V. L., Satow, Y., Huizinga, E., Vellieux, F. M. D., and Hol, W. G. J. (1992) *Biochemistry* **31**, 4959 – 4964
471a. Chen, L., Doi, M., Durley, R. C. E., Chistoserdov, A. Y., Lidstrom, M. E., Davidson, V. L., and Mathews, F. S. (1998) *J. Mol. Biol.* **276**, 131–149
472. Chen, L., Durley, R. C. E., Mathews, F. S., and Davidson, V. L. (1994) *Science* **264**, 86 – 90
472a. Zhu, Z., and Davidson, V. L. (1999) *Biochemistry* **38**, 4862 – 4867
473. Mure, M., and Klinman, J. P. (1995) *J. Am. Chem. Soc.* **117**, 8707 – 8718
474. Lee, Y., and Sayre, L. M. (1995) *J. Am. Chem. Soc.* **117**, 11823 – 11828

References

474a. Su, Q., and Klinman, J. P. (1998) *Biochemistry* **37**, 12513–12525

474b. Wilmot, C. M., Hajdu, J., McPherson, M. J., Knowles, P. F., and Phillips, S. E. V. (1999) *Science* **286**, 1724–1728

474c. Plastino, J., Green, E. L., Sanders-Loehr, J., and Klinman, J. P. (1999) *Biochemistry* **38**, 8204–8216

474d. Singh, V., Zhu, Z., Davidson, V. L., and McCracken, J. (2000) *J. Am. Chem. Soc.* **122**, 931–938

475. Itoh, S., Ogino, M., Haranou, S., Terasaka, T., Ando, T., Komatsu, M., Ohshiro, Y., Fukuzumi, S., Kano, K., Takagi, K., and Ikeda, T. (1995) *J. Am. Chem. Soc.* **117**, 1485–1493

476. Itoh, S., Kawakami, H., and Fukuzumi, S. (1997) *J. Am. Chem. Soc.* **119**, 439–440

477. Gorren, A. C. F., Moenne-Loccoz, P., Backes, G., de Vries, S., Sanders-Loehr, J., and Duine, J. A. (1995) *Biochemistry* **34**, 12926–12931

478. Moënne-Loccoz, P., Nakamura, N., Itoh, S., Fukuzumi, S., Gorren, A. C. F., Duine, J. A., and Sanders-Loehr, J. (1996) *Biochemistry* **35**, 4713–4720

479. Pedersen, J. Z., EL-Sherbini, S., Finazzi-Agrò, A., and Rotilio, G. (1992) *Biochemistry* **31**, 8–12

480. Warncke, K., Brooks, H. B., Lee, H.-i, McCracken, J., Davidson, V. L., and Babcock, G. T. (1995) *J. Am. Chem. Soc.* **117**, 10063–10075

481. Gorren, A. C. F., and Duine, J. A. (1994) *Biochemistry* **33**, 12202–12209

482. Hartmann, C., Brzovic, P., and Klinman, J. P. (1993) *Biochemistry* **32**, 2234–2241

483. Steinebach, V., de Vries, S., and Duine, J. A. (1996) *J. Biol. Chem.* **271**, 5580–5588

484. Morton, R. A. (1971) *Biol. Rev. Cambridge Philos. Soc.* **46**, 47–96

485. Morton, R. A. (1977) *Biol. Rev. Cambridge Philos. Soc.* **46**, 47–96

486. Morton, R. A. (1972) *Vitamins* **5**, 355–391

487. Suzuki, H., and King, T. E. (1983) *J. Biol. Chem.* **258**, 352–358

488. Matsuma, K., Bowyer, J. R., Ohnishi, T., and Dutton, L. P. (1983) *J. Biol. Chem.* **258**, 1571–1579

489. He, D.-Y., Yu, L., and Yu, C.-A. (1994) *J. Biol. Chem.* **269**, 27885–27888

490. He, D.-Y., Gu, L.-Q., Yu, L., and Yu, C.-A. (1994) *Biochemistry* **33**, 880–884

491. Folkers, K., Yamamuro, Y., and Ito, K., eds. (1980) *Biomedical and Clinical Aspects of Coenzyme Q*, Elsevier, Amsterdam

492. Morton, R. A. (1972) *Biochem. Soc. Symp.* **35**, 203–217

493. Gibbs, M. (1971) *Structure and Function of Chloroplasts*, Springer-Verlag, Berlin and New York

494. Trumpower, B. L., ed. (1982) *Function of Quinones in Energy Conserving Systems*, Academic Press, New York

495. Suzuki, Y. J., Tsuchiya, M., Wassall, S. R., Choo, Y. M., Govil, G., Kagan, V. E., and Packer, L. (1993) *Biochemistry* **32**, 10692–10699

496. Sun, I. L., Sun, E. E., Crane, F. L., Morré, D. J., Lindgren, A., and Löw, H. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11126–11130

497. Hughes, P. E., and Tove, S. B. (1980) *J. Biol. Chem.* **255**, 7095–7097

498. Valgimigli, L., Banks, J. T., Ingold, K. U., and Lusztyk, J. (1995) *J. Am. Chem. Soc.* **117**, 9966–9971

499. Nagaoka, S.-i, and Ishihara, K. (1996) *J. Am. Chem. Soc.* **118**, 7361–7366

500. van Belzen, R., Kotlyar, A. B., Moon, N., Dunham, R. D., and Albracht, S. P. J. (1997) *Biochemistry* **36**, 886–893

501. Salerno, J. C., Osgood, M., Liu, Y., Taylor, H., and Scholes, C. P. (1990) *Biochemistry* **29**, 6987–6993

502. Swallow, A. J. (1982) in *Function of Quinones in Energy Conserving Systems* (Trumpower, B. L., ed), pp. 59–71, Academic Press, New York

503. McTigue, J. J., Dhaon, M. K., Rich, D. H., and Suttie, J. W. (1984) *J. Biol. Chem.* **259**, 4272–4278

504. Suttie, J. W. (1993) *FASEB J.* **7**, 445–452

505. Morris, D. P., Soute, B. A. M., Vermeer, C., and Stafford, D. W. (1993) *J. Biol. Chem.* **268**, 8735–8742

506. Lingenfelter, S. E., and Berkner, K. L. (1996) *Biochemistry* **35**, 8234–8243

507. Kuliopulos, A., Hubbard, B. R., Lam, Z., Koski, I. J., Furie, B., Furie, B. C., and Walsh, C. T. (1992) *Biochemistry* **31**, 7722–7728

507a. Wu, S.-M., Mutucumarana, V. P., Geromanos, S., and Stafford, D. W. (1997) *J. Biol. Chem.* **272**, 11718–11722

508. Wood, G. M., and Suttie, J. W. (1988) *J. Biol. Chem.* **263**, 3234–3239

509. Metzler, D. E. (1977) *Biochemistry. The Chemical Reactions of Living Cells*, Academic Press, New York (p. 581)

510. Sadowski, J. A., Esmon, C. T., and Suttie, J. W. (1976) *J. Biol. Chem.* **251**, 2770–2776

511. Anton, D. L., and Friedman, P. A. (1983) *J. Biol. Chem.* **258**, 14084–14087

512. Naganathan, S., Hershline, R., Ham, S. W., and Dowd, P. (1994) *J. Am. Chem. Soc.* **116**, 9831–9839

513. Dowd, P., Hershline, R., Ham, S. W., and Naganathan, S. (1995) *Science* **269**, 1684–1691

514. Bouchard, B. A., Furie, B., and Furie, B. C. (1999) *Biochemistry* **38**, 9517–9523

514a. Zheng, Y.-J., and Bruice, T. C. (1998) *J. Am. Chem. Soc.* **120**, 1623–1624

515. Dubois, J., Gaudry, M., Bory, S., Azerad, R., and Marquet, A. (1983) *J. Biol. Chem.* **258**, 7897–7899

516. Dubois, J., Dugave, C., Fourès, C., Kaminsky, M., Tabet, J.-C., Bory, S., Gaudry, M., and Marquet, A. (1991) *Biochemistry* **30**, 10506–10512

517. Decottignies-Le Maréchal, P., Ducrocq, C., Marquet, A., and Azerad, R. (1984) *J. Biol. Chem.* **259**, 15010–15012

518. Lee, J. J., Principe, L. M., and Fasco, M. J. (1985) *Biochemistry* **24**, 7063–7070

519. Hildebrandt, E. F., and Suttie, J. W. (1982) *Biochemistry* **21**, 2406–2411

520. Fasco, M. J., Preusch, P. C., Hildebrandt, E., and Suttie, J. W. (1983) *J. Biol. Chem.* **258**, 4372–4380

521. Esmon, C. T. (1989) *J. Biol. Chem.* **264**, 4743–4746

522. Stenflo, J., Öhlin, A.-K., Owen, W. G., and Schneider, W. J. (1988) *J. Biol. Chem.* **263**, 21–24

523. Pryor, W. A. (1982) *Ann. N.Y. Acad. Sci.* **393**, 1–22

524. Ham, A.-J. L., and Liebler, D. C. (1995) *Biochemistry* **34**, 5754–5761

525. Valgimigli, L., Ingold, K. U., and Lusztyk, J. (1996) *J. Am. Chem. Soc.* **118**, 3545–3549

526. Christen, S., Woodall, A. A., Shigenaga, M. K., Southwell-Keely, P. T., Duncan, M. W., and Ames, B. N. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3217–3222

527. Schwarz, K., and Foltz, C. M. (1957) *J. Am. Chem. Soc.* **79**, 3292–3293

528. Shamberger, R. J. (1983) *Biochemistry of Selenium*, Plenum, New York

529. Keating, K. I., and Dagbusan, B. C. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3433–3437

530. Stadtman, T. C. (1991) *J. Biol. Chem.* **266**, 16257–16260

531. Stadtman, T. C. (1996) *Ann. Rev. Biochem.* **65**, 83–100

532. Burk, R. F. (1991) *FASEB J.* **5**, 2274–2279

533. Low, S. C., and Berry, M. J. (1996) *Trends Biochem. Sci.* **21**, 203–208

534. Ren, B., Huang, W., Åkesson, B., and Ladenstein, R. (1997) *J. Mol. Biol.* **268**, 869–885

535. Ladenstein, R., and Wendel, A. (1983) *Eur. J. Biochem.* **133**, 51–69

536. Shen, Q., Chu, F.-F., and Newburger, P. E. (1993) *J. Biol. Chem.* **268**, 11463–11469

537. Chu, F.-F., Doroshow, J. H., and Esworthy, R. S. (1993) *J. Biol. Chem.* **268**, 2571–2576

538. Duan, Y.-J., Komura, S., Fiszer-Szafarz, B., Szafarz, D., and Yagi, K. (1988) *J. Biol. Chem.* **263**, 19003–19008

539. Schnurr, K., Belkner, J., Ursini, F., Schewe, T., and Kühn, H. (1996) *J. Biol. Chem.* **271**, 4653–4658

540. Friedman, P. A. (1984) *N. Engl. J. Med.* **310**, 1458–1460

540a. Shisler, J. L., Senkevich, T. G., Berry, M. J., and Moss, B. (1998) *Science* **279**, 102–105

541. Visser, T. J. (1980) *Trends Biochem. Sci.* **5**, 222–224

542. Toyoda, N., Harney, J. W., Berry, M. J., and Larsen, P. R. (1994) *J. Biol. Chem.* **269**, 20329–20334

543. DePalo, D., Kinlaw, W. B., Zhao, C., Engelberg-Kulka, H., and St. Germain, D. L. (1994) *J. Biol. Chem.* **269**, 16223–16228

544. Davey, J. C., Becker, K. B., Schneider, M. J., St. Germain, D. L., and Galton, V. A. (1995) *J. Biol. Chem.* **270**, 26786–26789

545. Croteau, W., Whittemore, S. L., Schneider, M. J., and St. Germain, D. L. (1995) *J. Biol. Chem.* **270**, 16569–16575

546. St. Germain, D. L., Schwartzman, R. A., Groteau, W., Kanamori, A., Wang, Z., Brown, D. D., and Galton, V. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7767–7771

547. Tamura, T., and Stadtman, T. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1006–1011

548. Gladyshev, V. N., Jeang, K.-T., and Stadtman, T. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6146–6151

548a. Sun, Q.-A., Wu, Y., Zappacosta, F., Jeang, K.-T., Lee, B. J., Hatfield, D. L., and Gladyshev, V. N. (1999) *J. Biol. Chem.* **274**, 24522–24530

549. Marcocci, L., Flohé, L., and Packer, L. (1997) *BioFactors* **6**, 351–358

550. Hill, K. E., Lloyd, R. S., Yang, J.-G., Read, R., and Burk, R. F. (1991) *J. Biol. Chem.* **266**, 10050–10053

551. Himeno, S., Chittum, H. S., and Burk, R. F. (1996) *J. Biol. Chem.* **271**, 15769–15775

552. Steinert, P., Ahrens, M., Gross, G., and Flohé, L. (1997) *BioFactors* **6**, 311–319

552a. Saito, Y., Hayashi, T., Tanaka, A., Watanabe, Y., Suzuki, M., Saito, E., and Takahashi, K. (1999) *J. Biol. Chem.* **274**, 2866–2871

553. Read, R., Bellew, T., Yang, J.-G., Hill, K. E., Palmer, I. S., and Burk, R. F. (1990) *J. Biol. Chem.* **265**, 17899–17905

554. Vendeland, S. C., Beilstein, M. A., Chen, C. L., Jensen, O. N., Barofsky, E., and Whanger, P. D. (1993) *J. Biol. Chem.* **268**, 17103–17107

555. Vendeland, S. C., Beilstein, M. A., Yeh, J.-Y., Ream, W., and Whanger, P. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8749–8753

556. Yeh, J.-Y., Beilstein, M. A., Andrews, J. S., and Whanger, P. D. (1995) *FASEB J.* **9**, 392–396

557. Calvin, H. I., Cooper, G. W., and Wallace, E. (1981) *Gamete Research* **4**, 139–149

558. Shamberger, R. J. (1983) *Biochemistry of Selenium*, Plenum, New York (pp.47–49)

559. Yamazaki, S. (1982) *J. Biol. Chem.* **257**, 7926–7929

560. Yamazaki, S., Tsai, L., Stadtman, T. C., Teshima, T., and Nakaji, A. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1364–1366

561. Berg, B. L., Baron, C., and Stewart, V. (1991) *J. Biol. Chem.* **266**, 22386–22391

References

562. Gladyshev, V. N., Boyington, J. C., Khangulov, S. V., Grahame, D. A., Stadtman, T. C., and Sun, P. D. (1996) *J. Biol. Chem.* **271**, 8095–8100
563. Boyington, J. C., Gladyshev, V. N., Khangulov, S. V., Stadtman, T. C., and Sun, P. D. (1997) *Science* **275**, 1305–1306
564. Arkowitz, R. A., and Abeles, R. H. (1990) *J. Am. Chem. Soc.* **112**, 870–872
565. Arkowitz, R. A., and Abeles, R. H. (1991) *Biochemistry* **30**, 4090–4097
566. Stadtman, T. C., and Davis, J. N. (1991) *J. Biol. Chem.* **266**, 22147–22153
567. Arkowitz, R. A., and Abeles, R. H. (1989) *Biochemistry* **28**, 4639 –4644
567a. Self, W. T., and Stadtman, T. C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7208–7213
568. Gladyshev, V. N., Khangulov, S. V., and Stadtman, T. C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 232–236
569. Sliwkowski, M. X., and Stadtman, T. C. (1985) *J. Biol. Chem.* **260**, 3140–3144
570. Veres, Z., Tsai, L., Scholz, T. D., Politino, M., Balaban, R. S., and Stadtman, T. C. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2975–2979
571. Diamond, A. M., Choi, I. S., Crain, P. F., Hashizume, T., Pomerantz, S. C., Cruz, R., Steer, C. J., Hill, K. E., Burk, R. F., McCloskey, J. A., and Hatfield, D. L. (1993) *J. Biol. Chem.* **268**, 14215–14223
572. Diplock, A. T., and Lucy, J. A. (1973) *FEBS Lett.* **29**, 205–210
573. Sandholm, M., and Sipponen, P. (1973) *Arch. Biochem. Biophys.* **155**, 120–124
574. Sarathchandra, S. V., and Watkinson, J. H. (1981) *Science* **211**, 600–601
575. Reamer, D. C., and Zoller, W. H. (1980) *Science* **208**, 500–502
576. Mozier, N. M., McConnell, K. P., and Hoffman, J. L. (1988) *J. Biol. Chem.* **263**, 4527–4531
577. Harrison, D. J., Fluri, K., Seiler, K., Fan, Z., Effenhauser, C. S., and Manz, A. (1993) *Science* **261**, 895–897
578. Kromayer, M., Wilting, R., Tormay, P., and Böck, A. (1996) *J. Mol. Biol.* **262**, 413–420
579. Glass, R. S., Singh, W. P., Jung, W., Veres, Z., Scholz, T. D., and Stadtman, T. C. (1993) *Biochemistry* **32**, 12555–12559
580. Veres, Z., Kim, I. Y., Scholz, T. D., and Stadtman, T. C. (1994) *J. Biol. Chem.* **269**, 10597–10603
581. Mullins, L. S., Hong, S.-B., Gibson, G. E., Walker, H., Stadtman, T. C., and Raushel, F. M. (1997) *J. Am. Chem. Soc.* **119**, 6684–6685

Study Questions

1. *S*-adenosylmethionine is also a biological methyl group donor. The product of its methyl transferase reactions is *S*-adenosylhomocysteine. This product is further degraded by *S*-adenosylhomocysteine hydrolase, an enzyme that contains tightly bound NAD^+, to form homocysteine and adenosine.

   Write a step-by-step mechanism for the action of this hydrolase.

2. Compare the chemical mechanisms of enzyme-catalyzed decarboxylation of the following:
   a) a β-oxo-acid such as acetoacetate or oxaloacetate
   b) an α-oxo-acid such as pyruvate
   c) an amino acid such as L-glutamate

3. Describe the subunit structure of the enzyme pyruvate dehydrogenase. Discuss the functioning of each of the coenzymes that are associated with these subunits and write detailed mechanisms for each step in the pyruvate dehydrogenase reaction.

4. Free **formate** can be assimilated by cells via the intermediate **10-formyl-tetrahydrofolate** (10-formyl-THF).
   a) Describe the mechanism of synthesis of this compound from formate and tetrahydrofolate.
   b) Diagram a hypothetical transition state for the first step of this reaction sequence.
   c) Describe two or more uses that the human body makes of 10-formyl-THF.

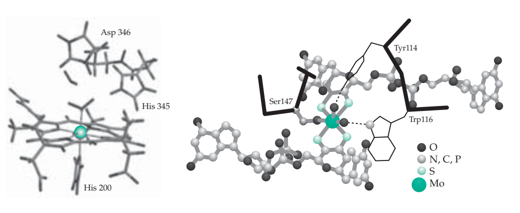
5. Using partial structural formulas, describe the reactions by which serine and methionine react to form *N*-formylmethionine needed for protein synthesis.

6. Write the equations for each of the reactions shown below. Using the $E^{0'}$ values below, calculate approximate Gibbs energies for each reaction, and show by the relative length of the arrows on which side of the reaction the equilibrium lies.
   a) The oxidation of malate by NAD^+
   b) The oxidation of succinate by NAD^+
   c) The oxidation of succinate by enzyme-bound FAD
   d) What can you say about the cofactor required for oxidation of succinate from your calculations?

   The values of $E^{0'}$ for several half reactions are given below. Everything has been rounded to one significant figure so that a calculator is unnecessary.

Reaction	$E^{0'}$ (volts)
$NAD^+ + H^+ + 2\,e^- \rightarrow NADH$	−0.3
enzyme bound FAD + $2\,H^+ + 2\,e^- \rightarrow$ enzyme bound $FADH_2$	0.0
fumarate + $2\,H^+ + 2\,e^- \rightarrow$ succinate	0.0
oxalacetate + $2\,H^+ + 2\,e^- \rightarrow$ malate	−0.2

7. Some acetogenic bacteria, which convert CO_2 to acetic acid, form pyruvate for synthesis of carbohydrates, etc., by formation of formaldehyde and conversion of the latter to glycine by reversal of the PLP and lipoic acid-dependent glycine decarboxylase, a 4-protein system. The glycine is then converted to serine, pyruvate, oxaloacetate, etc. Propose a detailed pathway for this sequence.



Transition metal ions function in many biological oxidation-reduction processes. (Left) A heme ring in **nitrite reductase** chelates an iron ion in its center, holding the Fe^{2+} or Fe^{3+} with bonds from four nitrogen atoms and a fifth bond from an imidazole ring below. Above the iron is a nitrite ion, NO_2^-, awaiting reduction to nitric oxide NO in a denitrifying bacterium. From Ranghino *et al.* (2000) *Biochemistry* **39**, 10958–10966. (Right) The active site of a bacterial **dimethylsulfoxide reductase** has an atom of molybdenum or tungsten at its center. The metal is held by four sulfur atoms from two molybdopterin molecules and an oxygen atom of a serine side chain. Two other oxygen atoms are bound as oxo groups and may participate in catalysis. From Stewart *et al.* (2000) *J. Mol. Biol.* **299**, 593–600.

Contents

837 .... **A. Iron**
837 .... 1. Uptake by Living Cells
838 .... *Siderophores*
839 .... *Uptake of iron by eukaryotic cells*
840 .... *Transferrins*
841 .... 2. Storage of Iron
843 .... 3. Heme Proteins
843 .... *Some names to remember*
844 .... *Hemes and heme proteins*
845 .... 4. The Cytochromes
846 .... *The c-type cytochromes*
847 .... *Cytochromes b, a, and o*
848 .... 5. Mechanisms of Biological Electron Transfer
848 .... *Electron-transfer pathways?*
849 .... *"Docking"*
849 .... *Coupling and gating of electron transfer*
850 .... *Effects of ionic equilibria on electron transfer*
850 .... 6. Reactions of Heme Proteins with Oxygen or Hydrogen Peroxide
850 .... *Oxygen-carrying proteins*
852 .... *Catalases and peroxidases*
853 .... *Mechanisms of catalase and peroxidase catalysis*
855 .... *Haloperoxidases*
857 .... 7. The Iron–Sulfur Proteins
857 .... *Ferredoxins, high-potential iron proteins, and rubredoxins*
858 .... *Rubredoxins*
859 .... *Chloroplast-type ferredoxins*
859 .... *The 3Fe – 4S clusters*
859 .... *Properties of iron–sulfur clusters*
860 .... *Functions of iron–sulfur enzymes*
862 .... 8. The (μ-oxo) Diiron Proteins
862 .... *Hemerythrin*
862 .... *Purple acid phosphatases*
863 .... *Diiron oxygenases and desaturases*
863 .... 9. Ribonucleotide Reductases
863 .... 10. Superoxide Dismutases
866 .... **B. Cobalt and Vitamin B_{12}**
867 .... 1. Coenzyme Forms
869 .... 2. Reduction of Cyanocobalamin and Synthesis of Alkyl Cobalamins
870 .... 3. Three Nonenzymatic Cleavage Reactions of Vitamin B_{12} Coenzymes
870 .... 4. Enzymatic Functions of B_{12} Coenzymes
871 .... *Cobalamin-dependent ribonucleotide reductase*
872 .... *The isomerization reactions*
874 .... *Stereochemistry of the isomerization reactions*
874 .... *Aminomutases*
875 .... *Transfer reactions of methyl groups*
876 .... *Corrinoid-dependent synthesis of acetyl-CoA*
877 .... **C. Nickel**
877 .... 1. Urease
878 .... 2. Hydrogenases
879 .... 3. Cofactor F_{430} and Methyl-Coenzyme M Reductase
881 .... 4. Tunichlorins
881 .... 5. Carbon Monoxide Dehydrogenases and Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase
882 .... **D. Copper**
883 .... 1. Electron-Transferring Copper Proteins
884 .... 2. Copper, Zinc-Superoxide Dismutase
884 .... 3. Nitrite and Nitrous Oxide Reductases
885 .... 4. Hemocyanins
885 .... 5. Copper Oxidases
887 .... 6. Cytochrome *c* Oxidase
887 .... **E. Manganese**
888 .... **F. Chromium**
889 .... **G. Vanadium**
890 .... **H. Molybdenum**
890 .... 1. Molybdenum Ions and Coenzyme Forms
892 .... 2. Enzymatic Mechanisms
893 .... 3. Nutritional Need for Mo
893 .... **I. Tungsten**

895 .... **References**
903 .... **Study Questions**

Boxes

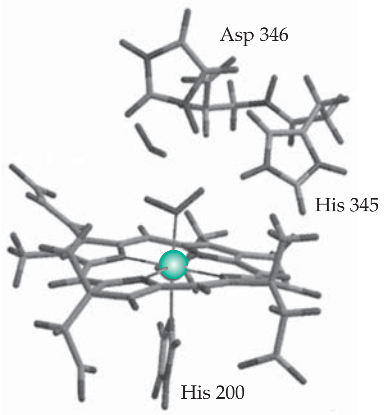
842 .... Box 16-A Magnetic Iron Oxide in Organisms
868 .... Box 16-B Cobalamin (Vitamin B_{12})

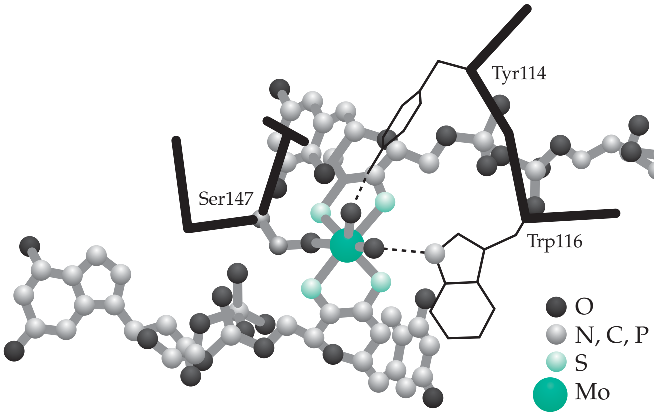
Tables

871 .... Table 16-1 Isomerization Reactions Involving Hydrogen Transfer and Dependent upon a Vitamin B_{12} Coenzyme

Transition Metals in Catalysis and Electron Transport

16





Although the amounts present within living cells are very small, the ions of the transition metals Fe, Co, Ni, Cu, and Mn are extremely active centers for catalysis, especially of reactions that take advantage of the ability of these metals to exist in more than one oxidation state.[1–4] Iron, copper, and nickel are also components of the electron carrier proteins that function as oxidants or reductants in many biochemical processes. These metals are all nutritionally essential, as are chromium and vanadium. Among the heavier transition elements molybdenum is a constituent of an important group of enzymes that includes the sulfite oxidase of human liver and nitrogenase of nitrogen-fixing bacteria. Tungsten occasionally substitutes for molybdenum.

A. Iron

Iron is one of the most abundant elements in the earth's crust, being present to the extent of ~4% in a typical soil. Its functions in living cells are numerous and diverse.[2,5–8] The average overall iron content of both bacteria and fungi is ~1 mmol / kg, but that of animal tissues is usually less. Seventy percent of the 3–5 g of iron present in the human body is located in the blood's erythrocytes, whose overall iron content is ~20 mM. In other tissues the total iron averages closer to 0.3 mM and consists principally of storage forms. The total concentration of iron in all of the iron-containing *enzymes* of tissues amounts to only about 0.01 mM. Although these concentrations are low, the iron is concentrated in oxidative enzymes of membranes and may attain much higher concentrations locally. Only a few parasitic or anaerobic bacteria, e.g. the lactic acid bacteria, possess no oxygen-requiring enzymes and are almost devoid of both iron and copper. All other organisms appear to require iron for life.

1. Uptake by Living Cells

A major problem for cells is posed by the relative insolubility of ferric hydroxide and other compounds from which iron must be extracted by the organism. A consequence is that iron is often taken up in a chelated form and is transferred from one organic ligand, often a protein, to another with little or no existence as free Fe^{3+} or Fe^{2+}. As can be calculated from the estimated solubility product of $Fe(OH)_3$ (Eq. 16-1),[7] the equilibrium concentration of Fe^{3+} at pH 7 is only 10^{-17} M.

$$K_{sp} = [Fe^{3+}][OH^-]^3 < 10^{-38}\ M^4 \qquad (16\text{-}1a)$$

$$\text{or } [Fe^{3+}] / [H^+]^3 < 10^4\ M^4 \text{ at } 25°\ C \qquad (16\text{-}1b)$$

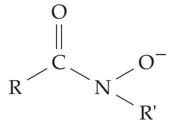
For a 2 μm^3 bacterial cell this amounts to just one free Fe^{3+} ion in almost 100 million cells at any single moment. The importance of chelated forms of iron becomes obvious. It is also evident from Eq. 16-1 that, in addition to chelation, a low external pH can also facilitate uptake of Fe^{3+} by organisms.

The values of the formation constants for chelates of Fe^{2+} typically lie between those of Mn^{2+} and Co^{2+} (Fig. 6-6, Table 6-9). For example, $K_1 = 10^{14.3}\ M^{-1}$ for formation of the Fe^{2+} chelate of EDTA. The smaller and more highly charged Fe^{3+} is bonded more strongly ($K_1 = 10^{25}\ M^{-1}$). These binding constants are independent of pH. However, the binding of any metal ion is affected by pH, as discussed in Chapter 6. A fact of considerable biochemical significance is the stronger binding of Fe^{3+} to oxygen-containing ligands than to nitrogen atoms, while Fe^{2+} tends to bind preferentially to nitrogen. It is also significant that Fe^{3+} bound to oxygen ligands tends to exchange readily with other ferric ions in the medium, whereas Fe^{3+} bound to nitrogen-containing ligands such as heme exchanges slowly. This fact is important for both iron-transport compounds and enzymes.

Siderophores. If a suitably high content of iron (e.g., 50 μM or more for *E. coli*) is maintained in the external medium, bacteria and other microorganisms have little problem with uptake of iron. However, when the external iron concentration is low, special compounds called siderophores are utilized to render the iron more soluble.[7–11c] For example, at iron concentrations below 2 μM, *E. coli* and other enterobacteria secrete large amounts of **enterobactin** (Fig. 16-1). The stable Fe^{3+}–enterobactin complex is taken up by a transport system that involves receptors on the outer bacterial membrane.[9,12,13] Siderophores from many bacteria have in common with enterobactin the presence of **catechol** (*ortho*-dihydroxybenzene) groups that chelate the iron.

The three catechol groups of enterobactin are carried on a cyclic serine triester structure. A variety of both cyclic and linear structures are found among other catechol siderophores.[14–19] For example, **parabactin** and **agrobactin** (Fig. 16-1) contain a backbone of spermidine[20] (Chapter 24). After the Fe^{3+}–enterobactin complex enters a bacterial cell the ester linkages of a siderophore are cleaved by an esterase. Because of the extremely high formation constant of $\sim 10^{52}\ \mathrm{M}^{-1}$ for the complex[11] the only way for a cell to release the Fe^{3+} is through this irreversible destruction of the iron carrier.[11d] Reduction to Fe^{2+} may be involved in release of iron from some siderophores.[11e]

The first known siderophore, isolated in 1952 by Neilands,[22] is **ferrichrome** (Fig. 16-1), a cyclic hexapeptide containing **hydroxamate** groups at the iron-binding centers. Oxygen atoms form the bonds to iron



Hydroxamate group

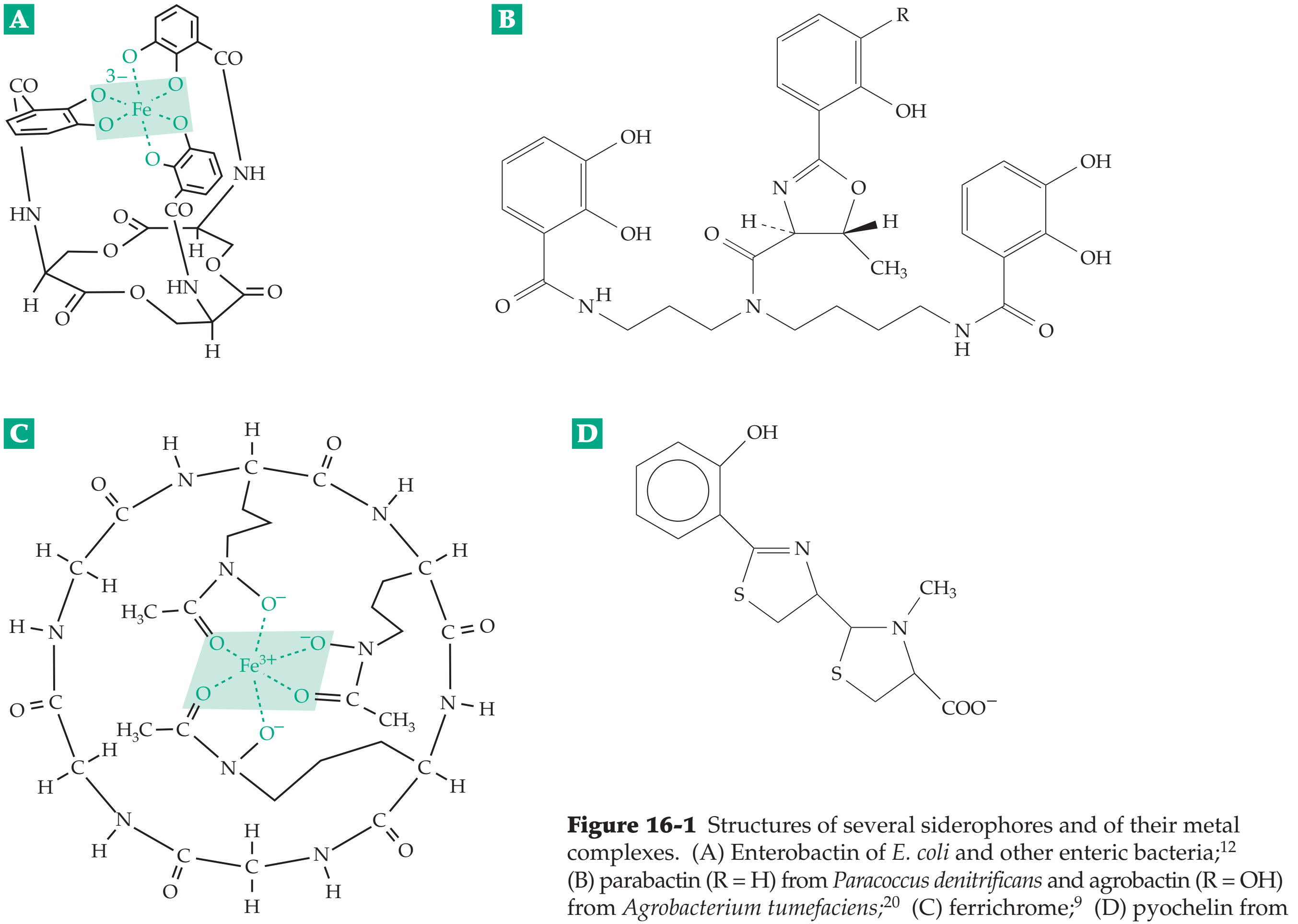


Figure 16-1 Structures of several siderophores and of their metal complexes. (A) Enterobactin of *E. coli* and other enteric bacteria;[12] (B) parabactin (R = H) from *Paracoccus denitrificans* and agrobactin (R = OH) from *Agrobacterium tumefaciens*;[20] (C) ferrichrome;[9] (D) pyochelin from *Pseudomonas aeruginosa*.[21]

in this compound also. Ferrichrome binds Fe^{3+} with a formation constant of ~10^{29} M^{-1}. The binding is not as tight as with enterobactin and the iron can be released by enzymatic reduction to Fe^{2+} which is much less tightly bound than is Fe^{3+}. The released ferrichrome can be secreted and used repeatedly to bring in more iron. Ferrichrome is produced by various fungi and bacilli and is only one of a series of known hydroxamate siderophores.[16] Since iron is essential to virtually all parasitic organisms the ability to obtain iron is often the limiting factor in establishing an infection.[23,24]

In *E. coli* there are seven different outer **membrane receptors** for siderophores.[25] One of these, the gated porin **FepA**, is specific for ferric enterochelin. With the assistance of another protein, **TonB**, it allows the ferric siderophore to penetrate the outer membrane.[26] A different receptor, **FhuA**, binds ferrichrome. Both FepA and FhaA are large 22-strand porins resembling the 16-strand porin shown in Fig. 8-20. However, they are nearly 7 nm long with internal diameters three times those of the 16-strand porins. In addition, loops of polypeptide chain on the outer edges can close while an N-terminal domain forms a "cork" that remains in place until the Fe^{3+}-siderophore complex enter the channel and binds. Like the hatches in an air lock on a spacecraft, the outer loops then close, after which the inner cork unwinds to allow the siderophore complex to enter the periplasmic space. The apparatus requires an energy supply, which apparently is provided by an additional complex consisting of proteins TonB, ExbA, and ExbD. They evidently couple the electrochemical gradient across the cell membrane (Chapter 8, B1 and C5 and Chapter 18) with the operation of the channel gates. Some bacteria use a different strategy for passage across the outer membrane.[11c] The channel of a receptor protein contains a molecule of an iron-free siderophore. When a molecule of Fe^{3+}-siderophore binds in the outside part of the channel the Fe^{3+} jumps to the inner siderophore, which then dissociates from the receptor, carrying the Fe^{3+}-siderophore complex into the periplasmic space. This mechanism also seems to be available in *E. coli*. The **ferric uptake regulation** (Fur) protein binds excess free Fe^{2+}, the resulting complex acting as a repressor of all of the iron uptake genes in *E. coli*.[11e]

Additional proteins are required for passage through the membrane.[11b,23,25,27] These are ABC transporters, which utilize hydrolysis of ATP as an energy source (Chapter 8, Section C,4). For uptake of the Fe^{3+}-ferrichrome complex protein **FhuD** is the periplasmic binding protein, **FhuC** is an integral membrane component, and the cytosolic **FhuC** contains the ATPase center.[11b] Another ABC transporter, found in many bacteria, carries unchelated Fe^{3+} across the inner membrane. The binding protein component for *Hemophilus influenzae*, designated Hit, resembles one lobe of mammalian transferrin (Fig. 16-2).[11b] The siderophore receptors of bacteria have been "parasitized" by various bacteriophages and toxic proteins. For example, FepA is also a receptor for the toxic colicins B and D (Box 8-D) and tonB is a receptor for bacteriophage T1.[9,13]

Some bacteria do not form siderophores but take up Fe^{2+}. Even *E. coli*, when grown anaerobically, synthesizes an uptake system for Fe^{2+}. It utilizes a 75-residue peptide encoded by gene *feoA* and a 773-residue protein encoded by *feoB*.[28,11b]

Uptake of iron by eukaryotic cells. The yeast *Saccharomyces cerevisiae* utilizes two systems for uptake of iron.[29,30] A low-affinity system transports Fe^{2+} with an apparent K_m of ~30 µM, while a high-affinity system has a K_m of ~0.15 µM. Study of these systems has been greatly assisted by the use of genetic methods developed for both bacteria and yeast (Chapter 26). The low-affinity iron uptake depends upon a protein transporter encoded by gene *FET*4 and on a reductase encoded by genes *FRE*1 and *FRE*2, proteins that are embedded in the cytoplasmic membrane.[29,31-33] It might seem reasonable that the *FET*3 copper oxidoreductase should keep Fe^{2+} *reduced* while it is transported. However, it appears to *oxidize* Fe^{2+} to Fe^{3+}. The high-affinity uptake system is more puzzling. It requires a permease encoded by *FTR I* and an additional protein encoded by *FET*3.[30,33–35] The Fet3 protein is a copper oxidoreductase related to **ceruloplasmin** (Section D). The protein **Fre1p** (encoded by *FRE*1) is a metalloreductase that reduces Cu^{2+} to Cu^{+}, as well as Fe^{3+} to Fe^{2+}. It is essential for copper uptake (Section D).[33] It has long been known that ceruloplasmin is required for mobilization of iron from mammalian tissues.[30] Hereditary ceruloplasmin deficiency causes accumulation of iron in tissues.[36] Yeast also contains both *mitochondrial* and *vacuolar* iron transporters.[37,37a,b]

The uptake of iron by animals is not as well understood[38–40] but it resembles that of yeast.[41] A general divalent cation transporter that is coupled to the membrane proton gradient is involved in intestinal iron uptake.[42,60] Ascorbic acid promotes the uptake of iron, presumably by reducing it to Fe(II), which is more readily absorbed than Fe(III), and also by promoting ferritin synthesis.[43] Uptake is also promoted by meat in the diet.[44] Within the body iron is probably transferred from one protein to another with only a transient existence as free Fe^{2+}. An average daily human diet contains ~15 mg of iron, of which ~1 mg is absorbed. This is usually enough to compensate for the small losses of the metal from the body, principally through the bile. Once it enters the body, iron is carefully conserved. The 9 billion red blood cells destroyed daily yield 20–25 mg of iron which is almost all reused or stored. The body apparently has no mechanism for excretion of large amounts of iron; a person's iron content is regulated almost entirely by the rate of uptake. This rate is increased during pregnancy and, in young women, to compensate for iron lost in menstrual bleeding. Nevertheless, control of iron uptake is imperfect and perhaps 500 million people around the world suffer from iron deficiency.[44,45] For others, an excessive intake of iron or a genetic defect lead to accumulation of iron to toxic levels, a condition called **hemochromatosis**.[44,46,46a,b] This condition may also arise in any disease that leads to excessive destruction of hemoglobin or accumulation of damaged erythrocytes. Examples are β-thalassemia (Chapter 28) and cerebral malaria.[47] Treatment with chelating agents designed to remove iron is often employed.[47,48]

Transferrins. Within the body iron is moved from one location to another while bound as Fe^{3+} to transferrins, a family of related 680- to 700-residue 80-kDa proteins.[38,49–53] Each transferrin molecule contains two Fe^{+3} binding sites, one located in each of two similar domains of the folded peptide chain. A dianion, usually CO_3^{2-}, is bound together with each Fe^{3+}. Milk transferrin (**lactoferrin**[51a,b,c] also found in leukocytes), hen egg transferrin (**ovotransferrin**),[52,52a] and rabbit and human serum transferrin[54,54a] all have similar structures. Each Fe^{3+} is bonded to oxygen anions from two tyrosine side chains, an aspartate carboxylate, an imidazole group, and the bound carbonate ion (Fig. 16-2B). Transferrin of blood plasma is encoded by a separate gene but has a similar structure. Transferrin of chickens appears to be identical to conalbumin of egg whites. The iron-binding proteins of body fluids are sometimes given the group name **siderophilins**. Transferrins may function not only in transport of iron throughout the body but also as iron buffers that provide a relatively constant iron concentration within tissues.

The entrance of iron into the body through the intestinal mucosal cells may involve the transferrin present in those cells[44] and the influx of iron may also be regulated by blood plasma transferrin. There is also a nontransferrin pathway.[42,55] Transferrins bind Fe^{2+} weakly and it is likely that a transferrin– Fe^{2+}–HCO_3^- complex formed initially undergoes oxidation to the Fe^{3+}–CO_3^{2-} complex within cells and within the bloodstream. A conformational change closes the protein around the iron ions.[56] In yeast the previously mentioned copper oxidoreductase encoded by the *FET*3 gene appears to not only oxidize Fe^{2+} but also transfer the resulting Fe^{3+} to transferrin. Ceruloplasmin may play a similar role in mammals.[33]

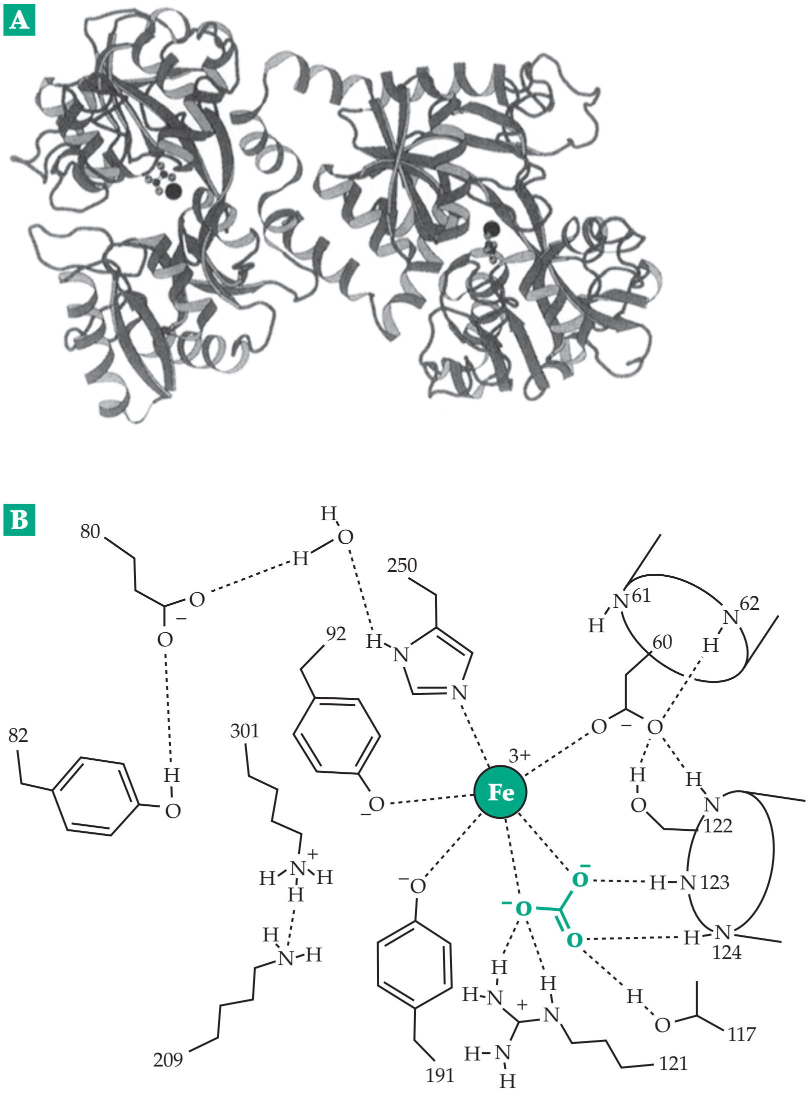


Figure 16-2 (A) Ribbon drawing of the polypeptide chain of a transferrin, human lactoferrin. The N lobe is to the left and the C lobe to the right. Each active site contains bound Fe^{3+} and a molecule of oxalate dianion which replaces the physiological CO_3^{2-}. From Baker *et al.*[51] Courtesy of Edward Baker. (B) Schematic diagram showing part of the hydrogen-bond network involved in binding the Fe^{3+} in the N lobe of hen ovotransferrin. Some side chain groups and water molecules have been omitted. The positions of hydrogen atoms and the charge state of acid–base groups are uncertain. Most of the hydrogen-bond distances (O - - O, O - - N, N - - N) indicated by dashed lines are between 0.27 and 0.3 nm. Release of the bound Fe^{3+} may be accomplished in part by protonation of the bound CO_3^{2-} to form HCO_3^-. See Kurokawa *et al.*[52]

Iron is transferred from the plasma transferrin into cells of the body following binding of the Fe^{3+}–transferrin complex to specific receptors. The surface of an immature red blood cell (reticulocyte) may contain 300,000 transferrin receptors, each capable of catalyzing the entry of ~36 iron ions per hour.[38] The receptor is a 180-kDa dimeric glycoprotein. When the Fe^{3+}–transferrin complex is bound, the receptors aggregate in coated pits and are internalized. The mechanism of release of the Fe^{3+} may occur by different mechanisms in the two lobes.[57] The pH of the endocytic vesicles containing the receptor complex is probably lowered to ~5.6 as in lysosomes. This may protonate the bound CO_3^{2-} in the complex[51c,54a,58] and assists in the release of the Fe^{3+}, possibly after reduction to Fe^{2+}. Both the apotransferrin and its receptor are returned to the cell surface for reuse, the apotransferrin being released into the blood. Chelating agents such as pyrophosphate, ATP, and citrate as well as simple anions[59] may also assist in removal of iron from transferrin. The same transmembrane transporter that is involved in intestinal iron uptake[42] is needed to remove iron from the endosome after release.[60]

2. Storage of Iron

Within tissues of animals, plants, and fungi much of the iron is packaged into the red-brown water-soluble protein **ferritin**, which stores Fe(III) in a soluble, nontoxic, and readily available form.[61–64] Although bacteria store very little iron,[65] some of them also contain a type of ferritin.[66–67a] On the other hand, the yeast *S. cerevisieae* stores iron in polyphosphate-rich granules, even though a ferritin is also present.[65] Ferritin contains 17–23% iron as a dense core of hydrated ferric oxide ~7 nm in diameter surrounded by a protein coat made up of twenty-four subunits of mo-

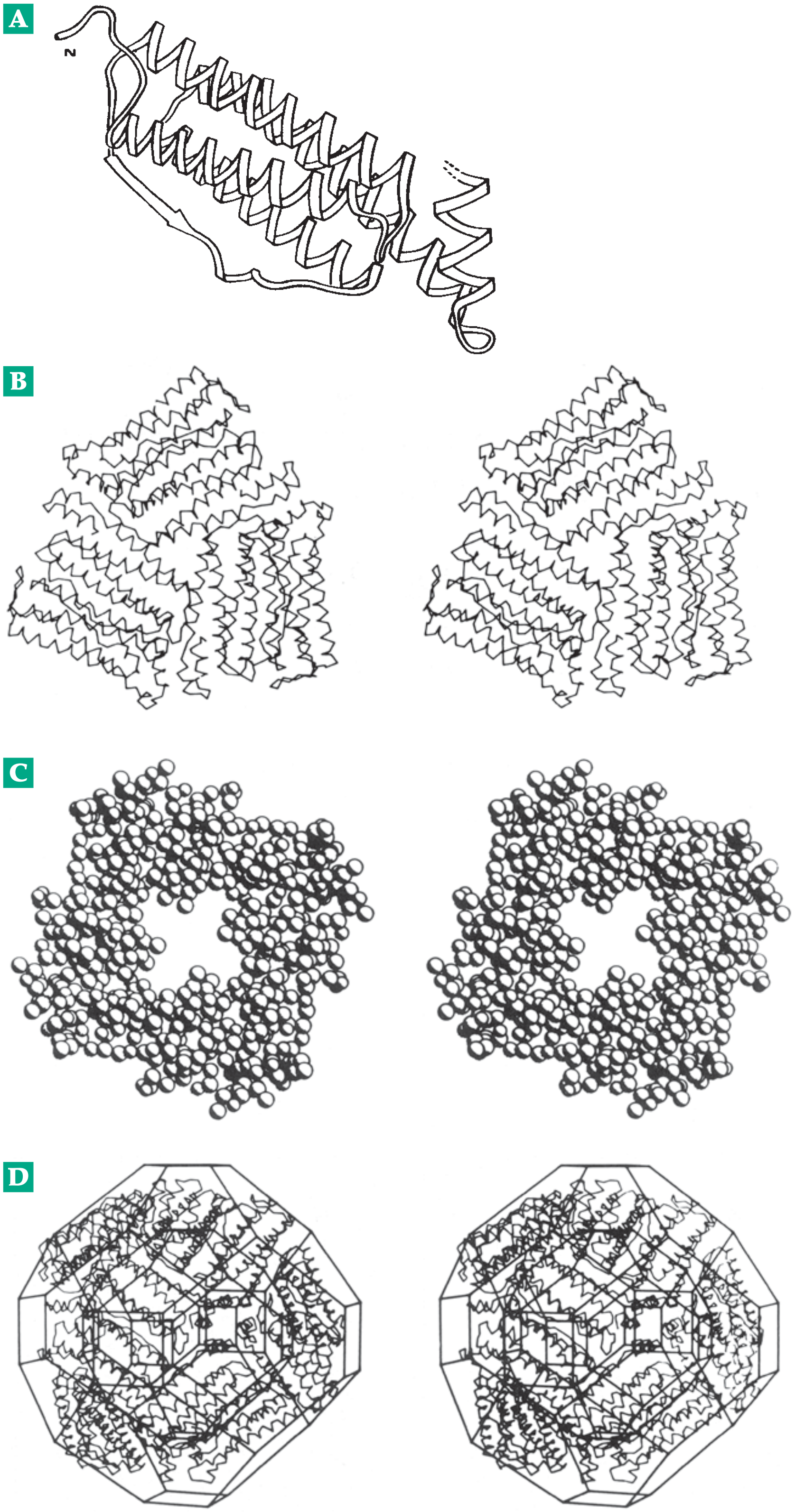


Figure 16-3 Structure of the protein shell of ferritin (apoferritin). (A) Ribbon drawing of the 163-residue monomer. From Crichton.[62] (B) Stereo drawing of a hexamer composed of three dimers. (C) A tetrad of four subunits drawn as a space-filling diagram and viewed down the four-fold axis from the exterior of the molecule. (D) A half molecule composed of 12 subunits inscribed within a truncated rhombic dodecahedron. B–D from Bourne *et al*.[74]

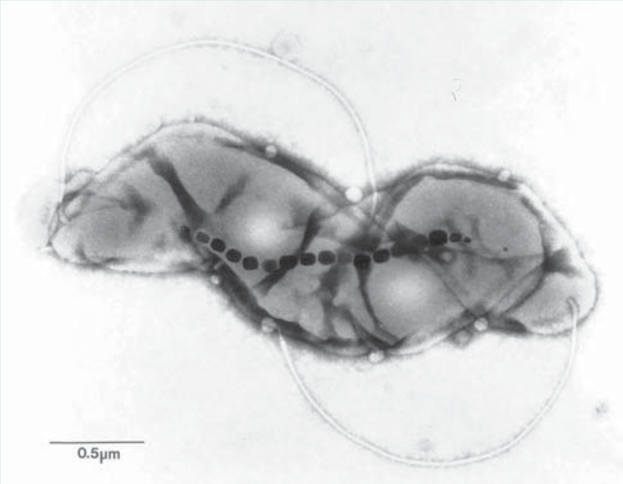
lecular mass 17- to 21-kDa. Each subunit is folded as a four-helix bundle (Fig. 16-3). Mammalian ferritins consist of combinations of subunits of two or more types. For example, human ferritins contain similarly folded 19-kDa L (light) and 21-kDa H (heavy) subunits.[68] The twenty-four subunits are arranged in a cubic array (Fig. 7-13, Fig. 16-3). The outer diameter of the 444-kDa apoferritin is ~12 nm. The completely filled ferritin molecule contains 23% Fe and over 2000 atoms of iron in a crystalline lattice. Larger ferritins may contain as many as 4500 atoms of iron with the approximate composition $[Fe(O)OH]_8FeOPO_3H_2$.[69] Phosphate ions are sometimes bound into surface layers of the ferritin cores.[69a] Ferritin cores are readily visible in the electron microscope, and ferritin is often used as a labeling reagent in microscopy. Another storage form of iron, **hemosiderin**, seems to consist of ferritin partially degraded by lysosomes and containing a higher iron content than does ferritin. Depositions of hemosiderin in the liver can rise to toxic levels if excessive amounts of iron are absorbed.

Iron can be deposited in ferritin by allowing apoferritin to stand with an Fe(II) salt and a suitable oxidant, which may be O_2. Physiological transfer of Fe(III) from transferrin to ferritin is thought to require prior reduction to Fe(II). The reoxidation by O_2 to Fe(III) for deposition in the ferritin core (Eq. 16-2) is catalyzed by **ferrooxidase sites** located in the centers of the helical bundles of the H-chains.[70–73]

$$2\,Fe^{2+} + O_2 + 4\,H_2O \;\rightarrow \\ 2\,Fe(O)OH_{core} + H_2O_2 + 4\,H^+ \qquad (16\text{-}2)$$

BOX 16-A MAGNETIC IRON OXIDE IN ORGANISMS

An unusual form of stored iron is the magnetic iron oxide **magnetite** (Fe_3O_4). Honeybees,[a,b] monarch butterflies,[b–c] homing pigeons,[c–e] migrating birds, and even magnetotactic bacteria[f] contain deposits of Fe_3O_4 that are suspected of being used in nagivation.[c,g] Some bacteria have magnetic iron sulfide particles.[h-j] Human beings have magnetic bones in their sinuses[k] and in their brains[l] and may be able to sense direction magnetically. A set of possible magnetoreceptor cells, as well as associated nerve pathways, have been identified in trout.[j] In the magnetotactic bacteria found in the Northern Hemisphere the magnetic domains are oriented parallel with the axis of motility of the bacteria which tend to swim toward the geomagnetic North and downward into sediments. Similar bacteria from the Southern Hemisphere prefer to swim south and downward. The magnetic polarity of the bacterial magnetite crystals can be reversed by strong magnetic pulses, after which the bacteria swim in the direction opposite to their natural one.[m] Magnetic ferritin can be produced artificially in the laboratory.[n] The resulting particles may have practical uses, for example, in medical magnetic imaging. Magnetic materials in the human body are of interest not only in terms of a possible sensory function but also because of possible effects of electromagnetic fields on human health and behavior.[o]

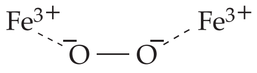


Magnetotactic soil bacterium containing 36 magnetite-containing magnetosomes. Courtesy of Dennis Bazylinski.

---

[a] Hsu, C.-Y., and Li, C.-W. (1994) *Science* **265**, 95–97
[b] Nichol, H., and Locke, M. (1995) *Science* **269**, 1888–1889
[bc] Etheredge, J. A., Perez, S. M., Taylor, O. R., and Jander, R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13845–13846
[c] Gould, J. L. (1982) *Nature (London)* **296**, 205–211
[d] Guilford, T. (1993) *Nature (London)* **363**, 112–113
[e] Moore, B. R. (1980) *Nature (London)* **285**, 69–70
[f] Blakemore, R. P., and Frankel, R. B. (1981) *Sci. Am.* **245** (Dec), 58–65
[g] Maugh, T. H., II,(1982) *Science* **215**, 1492 – 1493
[h] Dunin-Borkowski, R. E., McCartney, M. R., Frankel, R. B., Bazylinski, D. A., Pósfai, M., and Buseck, P. R. (1998) *Science* **282**, 1868–1870
[i] Pósfai, M., Buseck, P. R., Bazylinski, D. A., and Frankel, R. B. (1998) *Science* **280**, 880–883
[j] Walker, M. M., Diebel, C. E., Haugh, C. V., Pankhurst, P. M., Montgomery, J. C., and Green, C. R. (1997) *Nature (London)* **390**, 371–376
[k] Baker, R. R., Mather, J. G., and Kennaugh, J. H. (1983) *Nature (London)* **301**, 78–80
[l] Kirschvink, J. L., Kobayashi-Kirschvink, A., and Woodford, B. J. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7683–7687
[m] Blakemore, R. P., Frankel, R. B., and Kalmijn, A. J. (1980) *Nature (London)* **286**, 384–385
[n] Meldrum, F. C., Heywood, B. R., and Mann, S. (1992) *Science* **257**, 522–523
[o] Barinaga, M. (1992) *Science* **256**, 967

The ferrooxidase site is a dinuclear iron center (see Section 8) in which two iron ions (probably Fe^{2+}) are bound as in Fig. 16-4. They are then converted to Fe^{3+} ions by O_2, which may bind initially to the Fe^{2+}, forming a transient blue intermediate that is thought to have a peroxodiferric structure, perhaps of the following type.[71–73] Reaction of this intermediate with H_2O



$$Fe^{3+} \cdots {}^{-}O - O^{-} \cdots Fe^{3+}$$

may yield H_2O_2 plus a biomineral precursor $Fe^{2+} – O – Fe^{2+}$, which is incorporated into the core.[72a] Ferritin H subunits predominate in tissues with high oxygen levels, e.g., heart and blood cells, while the L subunits predominate in tissues with slower turnover of iron, e.g., liver.[72] The L subunits lack ferrooxidase activity but, in the centers of their helical bundles, contain polar side chains that may help to initiate growth of the mineral core.[64]

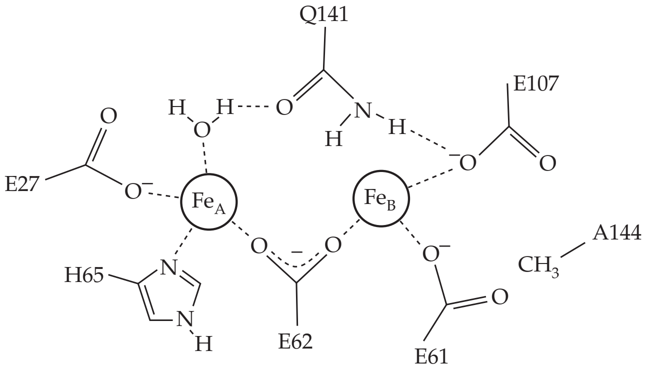


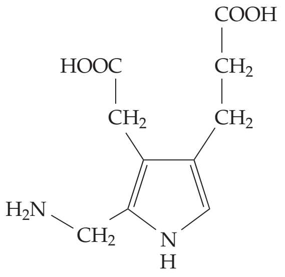
Figure 16-4 The dinuclear iron center or ferroxidase center of human ferritin based on the structure of a terbium(III) derivative.[73] Courtesy of Pauline Harrison.

Removal of Fe(III) from storage in ferritin cores may require reduction to Fe(II) again, possibly by ascorbic acid[75] or glutathione. Some bacterial ferritins contain a bound cytochrome *b* which may assist in reduction.[67,67a] Released iron in the Fe^{2+} state can be incorporated into iron-containing proteins or into heme. The enzyme **ferrochelatase**[76–76b] catalyzes the transfer of free Fe^{2+} into protoporphyrin IX (Section 3) to form protoheme (Fig. 16-5). Iron in the Fe(II) state may also be oxidized to Fe^{3+} through action of the copper-containing ceruloplasmin (Section D) and be incorporated into heme by direct transfer from ferritin.[53]

3. Heme Proteins

In 1879, German physiological chemist Hoppe-Seyler showed that two of the most striking pigments of nature are related. The red iron-containing **heme** from blood and the green magnesium complex **chlorophyll a** of leaves have similar ring structures. Later, H. Fischer proved their structures and provided them with the names and numbering systems that are used today. This information is summarized in the following section.

Some names to remember. **Porphins** are planar molecules which contain large rings made by joining four pyrrole rings with methine bridges. In the **chlorins**, found in the chlorophylls, one of the rings (ring D in chlorophyll, Fig. 23-20) is reduced. The specific class of porphins known as **porphyrins** have eight substituents around the periphery of the large ring. Like the chlorins and the **corrins** of vitamin B_{12} (Section B), the porphyrins are all formed biosynthetically from **porphobilinogen**. This compound is polymerized in two ways (see Fig. 24-21) to give porphyrins of types I and III (Fig. 16-5). In type I porphyrins, polymerization of porphobilinogen has taken place in a regular way so that the sequences of the carboxymethyl and carboxyethyl side chains (often referred to as acetic acid and propionic acid side chains, respectively) are the same all the way around the outside of the molecule. However, most biologically important porphyrins belong to type III, in which the first three rings A, B, and C have the same sequence of carboxymethyl and carboxyethyl side chains, but in which ring D has been incorporated in a reverse fashion. Thus, the carboxyethyl side chains of rings C and D are adjacent to each other (Fig. 16-5). Porphyrins containing all four carboxymethyl and four carboxyethyl side chains intact are known as **uroporphyrins**.



Porphobilinogen

Uroporphyrins I and III are both excreted in small amounts in the urine. Another excretion product is **coproporphyrin** III, in which all of the carboxymethyl side chains have been decarboxylated to methyl groups. The feathers of the tropical touraco are colored with copper(II) complex of coproporphyrin III and this porphyrin as well as others are commonly found in birds' eggs. The heme proteins are all derived from protoporphyrin IX, which is formed by decarboxylation and dehydrogenation of two of the carboxyethyl side chains of uroporphyrin III to vinyl groups (Fig. 16-5).

Hemes and heme proteins. Protoporphyrin IX contains a completely conjugated system of double bonds. In the center two hydrogen atoms are attached, one each to two of the nitrogens; they are free to move to other nitrogens in the center with rearrangement of the double bonds. Thus, there is tautomerism as well as resonance within the heme ring.[77–78a] The two central hydrogens can be replaced by many metal ions to form stable chelates. The complexes with Fe^{2+} are known as **hemes** and the Fe^{2+} complex with protoporphyrin IX as **protoheme**. Heme complexes of Fe^{2+} may be designated as **ferrohemes** and the Fe^{3+} compounds as **ferrihemes**. The Fe^{3+} protoporphyrin IX is also called **hemin**, and may be crystallized as a chloride salt.[78b] Iron tends to have a coordination number of six, and other ligands can attach to the iron from the two axial positions on opposite sides of the planar heme. If these are nitrogen ligands, such as pyridine or imidazole, the resulting compounds, called **hemochromes**, have characteristic absorption spectra. An example is cytochrome b_5, which contains two axial imidazole groups.

Several modifications of protoheme are indicated in Fig. 16-5. To determine which type of heme exists in a particular protein, it is customary to split off the heme by treatment with acetone and hydrochloric acid and to convert it by addition of pyridine to the pyridine hemochrome for spectral analysis. By this means, protoheme was shown to occur in hemoglobin, myoglobin, cytochromes of the *b* and P450 types, and catalases and many peroxidases. Cytochromes *a* and a_3 contain **heme *a***, while one of the terminal oxidase

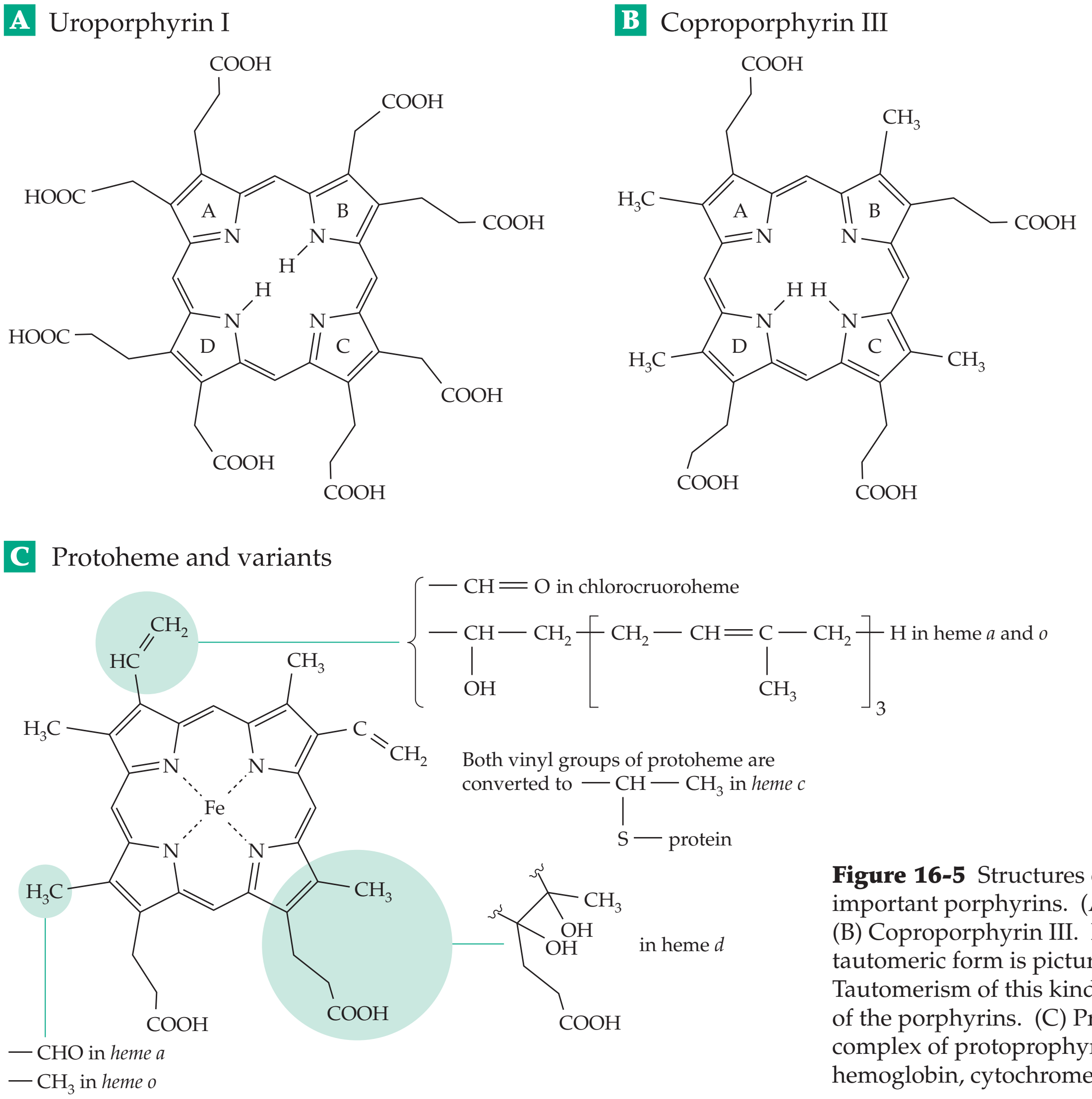

A Uroporphyrin I
COOH
HOOC
A
B
C
D
N
H
B Coproporphyrin III
CH3
H3C
H H
C Protoheme and variants
CH2
HC
CH3
H3C
C
CH2
N
Fe
—CH=O in chlorocruoroheme
—CH—CH2—[CH2—CH=C—CH2]3—H in heme a and o
OH
CH3
3
Both vinyl groups of protoheme are converted to —CH—CH3 in heme c
S—protein
OH
in heme d
—CHO in heme a
—CH3 in heme o

Figure 16-5 Structures of some biologically important porphyrins. (A) Uroporphyrin I. (B) Coproporphyrin III. Note that a different tautomeric form is pictured in B than in A. Tautomerism of this kind occurs within all of the porphyrins. (C) Protoheme, the Fe^{2+} complex of protoprophyrin IX, present in hemoglobin, cytochromes *b*, and other proteins.

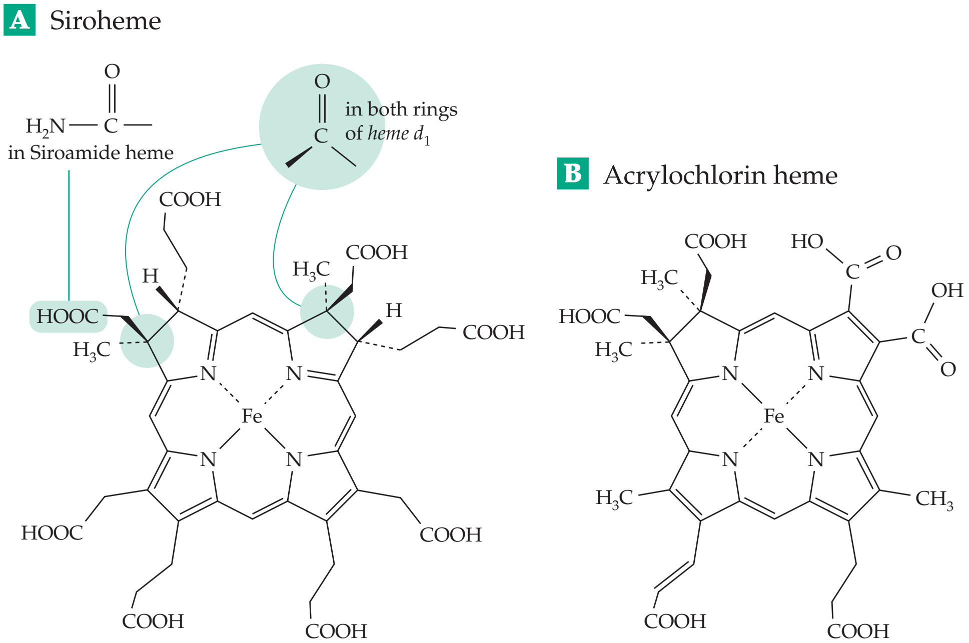


Figure 16-6 Structures of isobacteriochlorin prosthetic groups. (A) Siroheme from nitrite and sulfite reductases; (B) acrylochlorin heme from dissimilatory nitrite reductases of *Pseudomonas* and *Paracoccus*.

systems of enteric bacteria contains the closely related heme *o* (Fig. 16-5).[79,80] A second terminal oxidase of those same bacteria contains **heme *d*** (formerly a_2). **Heme *c*** (present in cytochromes *c* and *f*) is a variation in which two SH groups of the protein have added to the vinyl groups of protoheme to form two thioether linkages (Fig. 16-5). A few cytochromes *c* have only one such linkage. In myeloperoxidase (Section 6) three covalent linkages, different than those in cytochrome *c*, join the heme to the protein.[81,82] There is a possibility that heme *a* may sometimes form a Schiff base with a lysyl amino group through its formyl group.[83,84]

Heme *d* is a chlorin,[85] as is **acrylochlorin heme** from certain bacterial nitrite reductases (Fig. 16-6).[86,87] **Siroheme** (Fig. 16-6), which is found in both nitrite and sulfite reductases of bacteria (Chapter 24),[88,89] is an isobacteriochlorin in which both the A and B rings are reduced. It apparently occurs as an amide **siroamide** (Fig. 16-6) in *Desulfovibrio*.[90] Heme d_1 of nitrite reductases of denitrifying bacteria is a dioxobacteriochlorin derivative (Fig. 16-6).[91,92]

As in myoglobin, hemoglobin (Fig. 7-23), and cytochrome *c* (see Fig 16-8), one axial coordination position on the iron of most heme proteins (customarily called the *proximal* position) is occupied by an imidazole group of a histidine side chain. However, in cytochrome P450 and chloroperoxidase a thiolate ($-S^-$) group from a cysteinyl side chain, and in catalase a phenolate anion from a tyrosyl side chain, occupies the proximal position. The sixth or *distal* coordination position is occupied by the sulfur atom of methionine in cytochrome *c* and most other cytochromes with low-spin iron but cytochromes b_5 and c_3 have histidine. The high-spin heme proteins, such as cytochromes *c*′, globins, peroxidase, and catalase, usually have no ligand other than weakly bound H_2O in the distal position.[93]

Hemes are found in all organisms except the anaerobic clostridia and lactic acid bacteria. Heme proteins of blood carry oxygen reversibly, whereas those of **terminal oxidase systems**, **hydroxylases**, and **oxygenases** "activate" oxygen, catalyzing reactions with hydrogen ions and electrons or with carbon compounds. The heme-containing **peroxidases** and **catalases** catalyze reactions not with O_2 but with H_2O_2. Another group of heme proteins includes most of the cytochromes, which are purely electron-transferring compounds.

4. The Cytochromes

The iron in the small proteins known as cytochromes acts as an electron carrier, undergoing alternate reduction to the +2 state and oxidation to the +3 state. The cytochromes, discovered in 1884 by McMunn,[94] were first studied systematically in the 1920s by Keilin (Chapter 18) and have been isolated from many sources.[95–97] The classification into groups *a*, *b*, and *c* according to the position of the longest wavelength light absorption band (the α band; Fig. 16-7) follows a practice introduced by Keilin. However, it is now customary to designate a new cytochrome by giving the heme type (*a*, *b*, *c* or *d*) together with the wavelength of the α band, e.g., cytochrome c_{552} or cyt $b_{557.5}$.

Cytochromes of the *b* type including bacterial cytochrome *o* contain protoheme. Because the sixth position is ligated, most cytochromes *b* do not react with O_2. However, cytochromes *o* and *d* serve as terminal electron acceptors (cytochrome oxidases) and are oxidizable by O_2. Another protoheme-containing cytochrome, involved in hydroxylation (Chapter 18), is **cytochrome P450**. Here the 450 refers to the position of the intense "Soret band" (also called the γ band) of the spectrum (Fig. 16-7) in a difference spectrum run in the presence and absence of CO. Other properties are also used in arriving at designations for cytochromes. For example, cytochrome a_3 has a spectrum similar to that of cytochrome *a* but it reacts readily with both CO and O_2.

Another property that distinguishes various cytochromes is the redox potential $E°'$ (Table 6-8), which in this discussion is given for pH 7.0. Cytochromes carry electrons between other oxidoreductase proteins of widely varying values of $E^{0'}$. Because of the various heme environments cytochromes have greatly differing values of $E^{0'}$, allowing them to function in many different biochemical systems.[97a,97b] For mitochondrial cytochrome *c* the value of $E°'$ is ~ + 0.265 V but for the closely related cytochrome *f* of chloroplasts it is ~ +0.365 V and for cytochrome c_3 of *Desulfovibrio* about –0.330 V. There is more than an 0.6-volt difference between $E°'$ of cytochromes *f* and c_3. Cytochromes *b* tend to have lower $E°'$ values, close to zero, than most cytochromes *c*, while cytochrome a_3 has $E°'$ ~ +0.385 V.

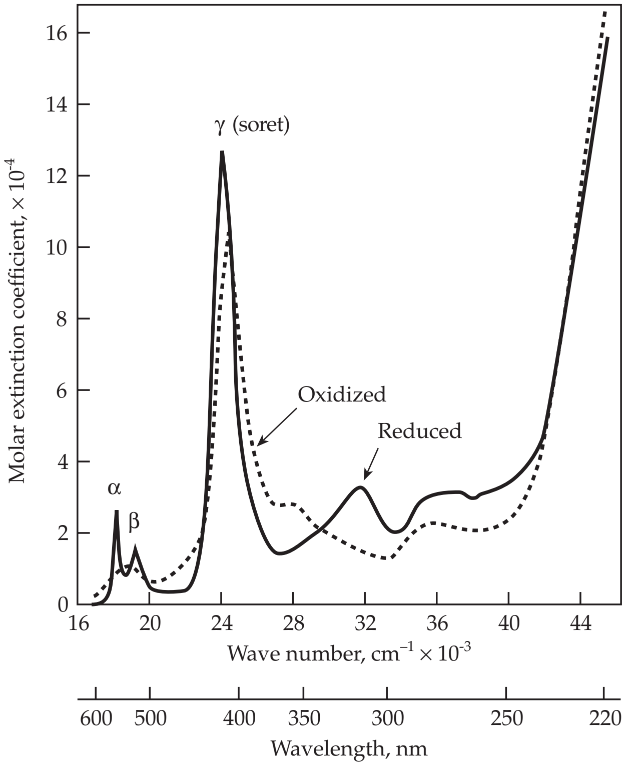


Figure 16-7 Absorption spectra of oxidized and reduced horse heart cytochrome *c* at pH 6.8. From data of Margoliash and Frohwirt.[98]

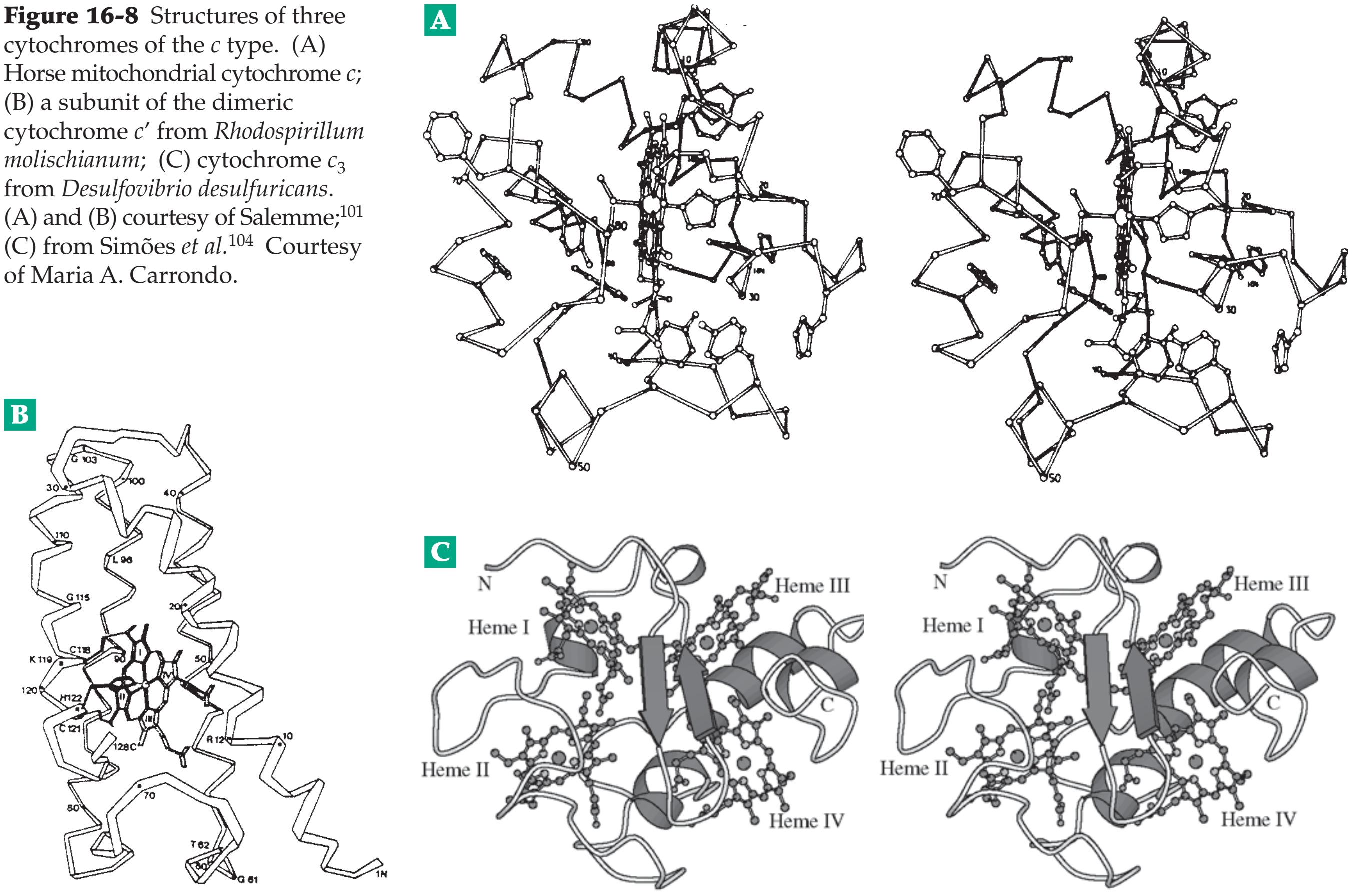
The c-type cytochromes. Mitochondrial cytochrome *c* is one of the few intracellular heme pigments that is soluble in water and that can be removed easily from membranes. A small 13-kDa protein typically containing about 104 amino acid residues, cytochrome *c* has been isolated from plants, animals, and eukaryotic microorganisms.[95–97,99,100] Complete amino acid sequences have been determined for over 100 species. Within the peptide chain 28 positions are invariant and a number of other positions contain only conservative substitutions. Cytochrome *c* was one of the first proteins to be used in attempting to trace evolutionary relationships between species by observing differences in sequence. Humans and chimpanzees have identical cytochrome *c*, but 12 differences in amino acid sequence occur between humans and the horses and 44 between human and *Neurospora*.[96] The related cytochrome c_2 of the photosynthetic bacterium *Rhodospillum rubrum* is thought to have diverged in evolution 2 x 10^9 years ago from the precursor of mammalian cytochrome *c*. Even so, 15 residues remain invariant.[101]

Structural studies[95–97,101–103] on cytochromes of the *c* and c_2 types show that the heme group provides a core around which the peptide chain is wound. The 104 residues of mitochondrial cytochrome *c* are enough to do little more than envelope the heme. In both the oxidized and reduced forms of the protein, methionine 80 (to the left in Fig. 16-8A) and histidine 18 (to the right) fill the axial coordination positions of the iron. The heme is nearly "buried" and inaccessible to the surrounding solvent.

The shorter chains of the 82- to 86-residue cytochromes c_{550} (from *Pseudomonas*[102]), c_{553},[102a] and c_{555} (from *Chlorobium*[105]) as well as the longer 112-residue polypeptide of cytochrome c_2 from *Rhodospirillum rubrum*[93] have nearly the same folding pattern as that in mitochondrial cytochrome *c*. However, the 128-residue chain of the dimeric cytochrome *c*′ from *Rhodospirillum molischianum* forms an antiparallel four-helix bundle (Fig. 16-8).[106–109] This is the same folding pattern present in the ferritin monomer (Fig. 16-3), hemerythrin (Fig. 2-22), and many other proteins including cytochrome b_{562} of *E. coli*.[110] Cytochrome *f*, which functions in photosynthetic electron transport, is also a *c*-type cytochrome but with a unique protein fold.[111,112]

Most cytochromes have only one heme group per polypeptide chain,[112] but the 115-residue cytochrome c_3 from the sulfate-reducing bacterium *Desulfovibrio* binds four hemes (Fig. 16-8C).[104,113–115] Each one seems to have a different redox potential in the –0.20 to –0.38 V range.[114] Another *c*-type cytochrome, also from *Desulfovibrio*, contains six hemes in a much larger 66-kDa protein and functions as a nitrite reductase.[116]

Figure 16-8 Structures of three cytochromes of the *c* type. (A) Horse mitochondrial cytochrome *c*; (B) a subunit of the dimeric cytochrome *c′* from *Rhodospirillum molischianum*; (C) cytochrome c_3 from *Desulfovibrio desulfuricans*. (A) and (B) courtesy of Salemme;[101] (C) from Simões *et al.*[104] Courtesy of Maria A. Carrondo.



Triheme and octaheme proteins are also known.[117]

Many cytochromes *c* are soluble but others are bound to membranes or to other proteins. A well-studied tetraheme protein binds to the reaction centers of many purple and green bacteria and transfers electrons to those photosynthetic centers.[118–120] Cytochrome c_2 plays a similar role in *Rhodobacter*, forming a complex of known three-dimensional structure.[121] Additional cytochromes participate in both cyclic and noncyclic electron transport in photosynthetic bacteria and algae (see Chapter 23).[120,122–124] Some bacterial membranes as well as those of mitochondria contain a **cytochrome *bc*$_1$ complex** whose structure is shown in Fig. 18-8.[125,126]

Cytochromes b, a, and o. Protoheme-containing cytochromes *b* are widely distributed.[127,128] There are at least five of them in *E. coli*. Whether in bacteria, mitochondria, or chloroplasts, the cytochromes *b* function within electron transport chains, often gathering electrons from dehydrogenases and passing them on to *c*-type cytochromes or to iron–sulfur proteins. Most cytochromes *b* are bound to or embedded within membranes of bacteria, mitochondria, chloroplasts, or endoplasmic reticulum (ER). For example, cytochrome b_5[129,129a] delivers electrons to a fatty acid desaturating system located in the ER of liver cells and to many other reductive biosynthetic enzymes.[130–132] The protein contains 132 amino acid residues plus another 85 largely hydrophobic N-terminal residues that provide a nonpolar tail which is thought to be buried in the ER membranes.[130] Solubilization of the protein causes loss of this N-terminal sequence. The heme in cytochrome b_5 is not covalently bonded to the protein but is held tightly between two histidine side chains. The polypeptide chain is folded differently than in either cytochrome *c* or myoglobin.

The folding pattern of cytochrome b_5 is also found in the complex heme protein **flavocytochrome b$_2$** from yeast (Chapter 15)[133] and probably also in liver **sulfite oxidase**[134,135] Both are 58-kDa peptides which can be cleaved by trypsin to 11-kDa fragments that have spectroscopic similarities and sequence homology with cytochrome b_5. Sulfite oxidase also has a molybdenum center (Section H). The 100-residue N-terminal portion of flavocytochrome b_2 has the cytochrome b_5 folding pattern but the next 386 residues form an eight-stranded $(\alpha / \beta)_8$ barrel that binds a molecule of FMN.[133,136] All of these proteins pass electrons to cytochrome *c*. In contrast, the folding of **cytochrome**

b_{562} of *E. coli* resembles that of cytochrome *c′* (Fig. 16-8).[110,137] However, it has methionine side chains as both the fifth and sixth iron ligands.

Cytochromes *b* of mitochondrial membranes are involved in passing electrons from succinate to ubiquinone in complex II[138] and also from reduced ubiquinone to cytochrome c_1 in the 248-kDa complex III (Fig. 18-8). A similar complex is present in photosynthetic purple bacteria.[123,139] Cytochrome b_{560} functions in the transport of electrons from succinate dehydrogenase to ubiquinone,[138] and cytochrome b_{561} of secretory vesicle membranes has a specific role in reducing ascorbic acid radicals.[140]

In bacteria some cytochromes *b* and d_1 serve as terminal electron carriers able to react with O_2, nitrite, or nitrate, while others act as carriers between redox systems.[141–143a] The aldehyde heme *a* is utilized by animals and by some bacteria in **cytochrome *c* oxidase**, a complex enzyme whose three-dimensional structure is known (see Fig. 18-10) and which is discussed further in Chapter 18.

5. Mechanisms of Biological Electron Transfer

The heme groups of the cytochromes as well as many other transition metal centers act as carriers of electrons. For example, cytochrome *c* may accept an electron from reduced cytochrome c_1 and pass it to cytochrome oxidase or cytochrome *c* peroxidase. The electron moves from one heme group to another over distances as great as 2 nm. Similar electron-transfer reactions between defined redox sites are met in photosynthetic reaction centers (Fig. 23-31), in metalloflavoproteins (Fig. 15-9), and in mitochondrial membranes.

What are the factors that determine the probability of an electron transfer reaction and the rate at which it may occur? They include: (1) The distance from the electron donor to the acceptor. (2) The thermodynamic driving force $\Delta G°$ for the reaction. This can be approximated using the difference in standard electrode potentials (as in Table 6-1) between donor and acceptor. $\Delta G° = -96.5\ \Delta E°$ kJ/mol at 25° C. (3) The chemical makeup of the material through which the electron transfer takes place. (4) Any changes in the geometry or charge state of the donor or acceptor that accompany the transfer. (5) The orientation of the acceptor and donor groups.[144–146] It is usually assumed that the Franck–Condon principle is obeyed, i.e., that the electron jump occurs so rapidly ($<10^{-12}$ s) that there is no change in the positions of atomic nuclei (see also Chapter 23). Subsequent rearrangement of nuclear positions may occur at rates that allow a rapid overall reaction.

The various factors that affect the rate of electron transfer were incorporated by Marcus into a quantitative theory. Electron transfer is often discussed in terms of this classic Marcus theory together with effects of quantum mechanical tunneling.[144,147–150] According to Marcus the electron transfer rate from a donor to an acceptor at a fixed separation depends upon $\Delta G°$, a nuclear reorganization parameter (λ), and the electronic coupling strength $|H_{AB}|$ between reactant and product in the transition state (Eq. 16-3):

$$k_{ET} = (4\pi^3 / h^2 \lambda k_B T)\,|H_{AB}|^2 \exp[-(\Delta G° + \lambda)^2 / 4\lambda k_B T] \tag{16-3}$$

Here $|H_{AB}|$ is a quantum mechanical matrix whose strength decreases exponentially with the distance of separation *R* as $e^{-\beta R}$ where β is a coefficient of the order of 9–14 nm^{-1}. At the closest contact (R = 0) the rate k_{ET}, by extrapolation from experimental data on small synthetic compounds, is close to the molecular vibration frequency of $10^{13}\ s^{-1}$.[151,152] At distances greater than 2 nm the rate would be negligible were it not for other factors.

Using mutant proteins as well as a variety of redox pairs and electron-transfer distances the validity of the Marcus equation with respect to the thermodynamic driving force and distance dependence has been verified.[153] This is even true for cytochrome *c* mutants functioning in living yeast cells.[146]

A huge amount of experimental work with proteins has been done to test and refine the theories of electron transport. For example, electron donor groups with various reduction potentials have been attached to various sites on the surface of a protein containing a heme or other electron accepting group. Ruthenium complexes such as Ru(III) $(NH_3)_5^{3+}$ form tight covalent linkages to imidazole nitrogens[154,155] such as that of His 33 of horse heart cytochrome *c*. This metal can be reduced rapidly to Ru(II) by an external reagent, after which the transfer of an electron from the Ru(II) across a distance of 1.2 nm to the heme Fe(III) can be followed spectroscopically. The reduction potentials *E°* for the Ru(III) / Ru(II) and Fe(III) / Fe(II) couples at pH 7 in these compounds are 0.16 and 0.27 V, respectively. Thus, an electron will jump spontaneously from the Ru(II) to the Fe(III) with $\Delta G° = -15.4$ kJ/mol. A rate constant of ~5 s^{-1}, which was nearly independent of temperature, was observed. Since the structures of Fe(II) and Fe(III) forms of cytochrome *c* differ only slightly,[156] the electron transfer apparently occurs with only a small amount of geometric rearrangement. The distribution of charges and dipoles within the protein may be such that the Fe^{2+} and Fe^{3+} complexes have almost equal thermodynamic stability.

Electron-transfer pathways? In spite of the success of the Marcus theory, rates of electron-transfer from the iron of cytochrome *c* have been found to vary for different pathways.[150,153,155] For example, transfer of an electron from Fe(II) in reduced cytochrome *c* to an Ru(III) complex on His 33 was fast (~440 s^{-1})[157] but

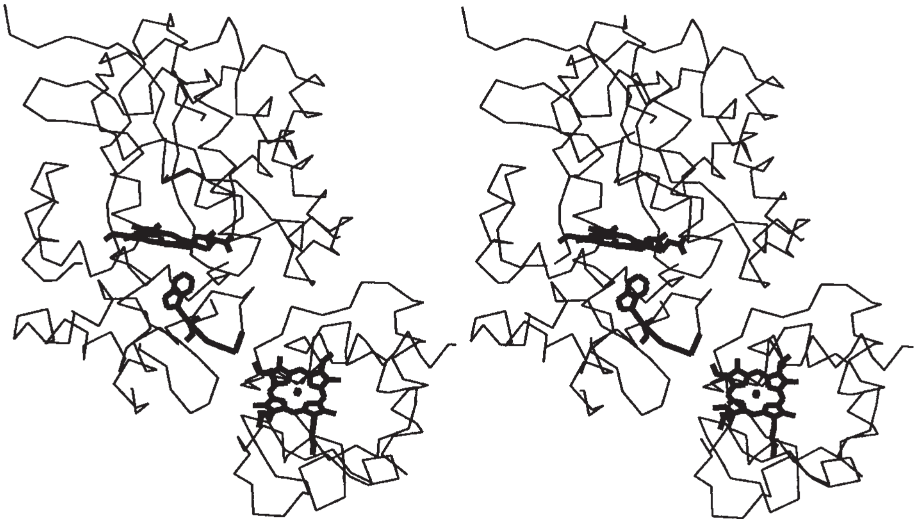


Figure 16-9 Stereoscopic α-carbon plot of yeast cytochrome *c* peroxidase (top) and yeast cytochrome *c* (below) as determined from a cocrystal by Pelletier and Kraut.[164] The heme rings of the two proteins appear in bold lines, as does the ring of tryptophan 191 and the backbone of residues 191–193 of the cytochrome *c* peroxidase. Drawing from Miller *et al.*[165]

the rate of transfer to an equidistant Fe(III) ion on Met 65 was at most 0.6 s^{-1}. These results suggested that distinct electron-transfer pathways exist. One suggestion was that the sulfur atom of Met 80 donates an electron to Fe^{3+} leaving an electron-deficient radical. The "hole" so created could be filled by an electron jumping in from the –OH group of the adjacent Tyr 67, which might then accept an electron from an external acceptor via Tyr 74 at the protein surface. Do electrons flow singly or as pairs from the surface through hydrogen-bonded paths? Use of both semisynthesis[155] and directed mutation[153] of cytochromes *c* is permitting a detailed study of these effects. A striking result is that substitution of the conserved residue phenylalanine 82 in a yeast cytochrome *c* with leucine or isoleucine retards electron transfer by a factor of ~10^4.

"Docking." It is now recognized that there are distinct "docking sites" on the surface of electron-transport proteins. For rapid electron transfer to occur the two electron carriers must be properly oriented and docked by formation of correct polar and nonpolar interactions. Early indications of the importance of docking came from study of modified cytochromes *c*. Each one of the 19 lysine side chains was individually altered by acylation or alkylation to remove the positive charge or to replace it with a negative charge. The rate of electron transfer into cytochrome *c* from hexacyanoferrate was decreased by a factor of 1.3–2.0 when any one lysine at positions 8, 13, 27, 72, or 79, which are clustered around the heme edge, was modified.[158] Modification of Lys 22, 55, 99, or 100, distant from this edge, had no effect. Electron transfer *into* cytochrome *c* from its natural electron donor ubiquinol:cytochrome *c* reductase was also strongly inhibited by modification of lysines that surround the heme edge.[159] Modification of these lysines also inhibited electron transfer *out of* cytochrome *c* into its natural acceptors, cytochrome *c* oxidase and cytochrome *c* peroxidase (Fig. 16-9).[160] A major factor in these effects is probably the large dipole moment in the cytochrome *c* that arises from the unequal distribution of surface charges. This charge distribution must assist in the proper docking of the cytochrome with its natural electron donors and acceptors. The positive surface charges presumably also facilitate the reaction with hexacyanoferrate (II) or ascorbate, both of which are negatively charged reductants that react rapidly with cytochrome *c*.

Measurements of many kinds have been made between natural donor–acceptor pairs such as cytochrome *c*–cytochrome b_5,[161,162] cytochrome *c*–cytochrome *c* peroxidase (Fig. 16-9),[153,163–166] trimethylamine dehydrogenase–FMN to Fe_4S_4 center (Fig. 15-9),[167] and methylamine dehydrogenase (TTQ radical)–amicyanin (Cu^{2+}).[168] Designed metalloproteins are being studied as well.[169] Femtosecond laser spectroscopy is providing a new approach.[169a]

Coupling and gating of electron transfer. Electrons are thought to be transferred into or out of cytochrome *c* through the exposed edge of the heme. The rate depends upon effective coupling, which in turn may depend upon orientation as well as the structure and dynamics of the protein.[170] Proteins with much β structure appear to provide stronger coupling than those that are largely composed of α helices.[144,171] The high nuclear reorganization energy λ of α helices may block electron transfer along some pathways.[153] A conformational change,[167,169] transfer of a proton, or binding of some other specific ion[172] before electron transfer occurs can be the "gating" process that determines the rate of electron transfer.[162,167,173] Electron transfer can also be "coupled" to an unfavorable, but fast, equilibrium.

Effects of ionic equilibria on electron transfer. The charge on an ion of Fe^{2+} in a heme is exactly balanced by two negative charges on the porphyrin ring. However, when the Fe^{2+} loses an electron to become Fe^{3+} an extra positive charge is suddenly present in the center of the protein. This change in charge will have a powerful electrostatic effect on charged groups in the immediate vicinity of the iron and even at the outer surface of the molecule. For example, an anion from the medium or from a neighboring protein molecule might become bound to the heme protein (Eq. 16-4).

$$X^- + [\mathrm{Fe(II)\ (heme)}] \rightarrow [\mathrm{Fe(III)\ (heme)}]^+ X^- + e^- \qquad (16\text{-}4)$$

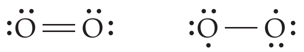
In this case the presence of a high concentration of X^- in the medium would favor the oxidation of Fe(II) to Fe(III). The reduced heme would be a better reducing agent and the oxidized form a weaker oxidant than in the absence of X^-. If –YH were a group in the protein the loss of an electron could cause –YH to dissociate so that Y^- and the Fe(III) would interact more tightly (Eq. 16-5).

$$\begin{aligned} &\mathrm{HY} - [\mathrm{Fe(II)\ (heme)}] \rightarrow \\ &Y^- - [\mathrm{Fe(III)\ (heme)}]^+ + H^+ + e^- \end{aligned} \qquad (16\text{-}5)$$

We see that electron transfer can be accompanied by loss of a proton and that $E°'$ may become pH dependent. (See also Eq. 16-18.) Even with cytochrome *c*, although there is little structural change upon electron transfer, there is an increased structural mobility in the oxidized form.[156] This may be important for coupling and could also facilitate associated proton-transfer reactions. For example, it is possible that in some cytochromes the imidazole ring in the fifth coordination position may become deprotonated upon oxidation. This possibility is of special interest because cytochromes are components of proton pumps in mitochondrial membranes (Chapter 18).

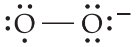
6. Reactions of Heme Proteins with Oxygen or Hydrogen Peroxide

As Ingraham remarked,[174] "Living in a bath of 20% oxygen, we tend to forget how reactive it is." From a thermodynamic viewpoint, all living matter is extremely unstable with respect to combustion by oxygen. Ordinarily, a high temperature is required and if we are careful with fire, we can expect to escape a catastrophe. However, one mole of properly chelated copper could catalyze consumption of all of the air in an average room within one second.[174] Biochemists are interested in both the fact that O_2 is kinetically stable and unreactive and also that oxidative enzymes such as cytochrome *c* oxidase are able to promote rapid reactions. Two oxygen atoms, each with six valence electrons, might reasonably be expected to form dioxygen, O_2, as a double-bonded structure with one σ and one π bond as follows (left):



Dioxygen

However, O_2 is paramagnetic and contains two unpaired electrons.[175] From this evidence O_2 might be assigned the structure on the right.[175] The oxygen molecule is very stable, and it is relatively difficult to add an electron to form the reactive **superoxide anion radical** O_2^-.

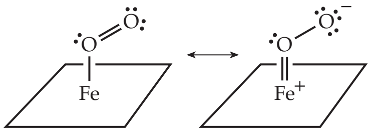


Superoxide anion radical

For this reason, oxidative attack by O_2 tends to be slow. However, once an electron has been acquired, it is easy for additional electrons to be added to the structure and further reduction occurs more easily. The biochemical question suggested is, "How can some heme proteins carry O_2 reversibly without any oxidation of the iron contained in them while others *activate* oxygen toward reaction with substrates?" Among this latter group, cytochrome *c* oxidase transfers electrons to both oxygen atoms so that only H_2O is a product, whereas the hydroxylases and oxygenases, which are discussed in Chapter 18, incorporate either one or two of the atoms of O_2, respectively, into an organic substrate. Before examining these reactions let us reconsider the heme oxygen carriers.

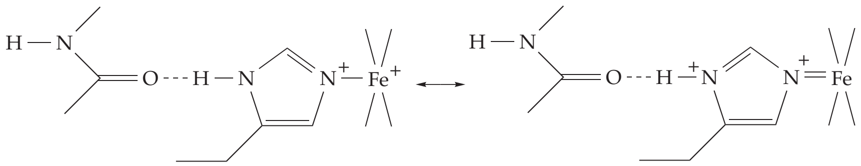
Oxygen-carrying proteins. In Chapter 7, we examined the behavior of hemoglobin in the cooperative binding of four molecules of O_2 and studied its structural relationship to the monomeric muscle protein myoglobin.[176] The iron in functional hemoglobin and myoglobin is always Fe(II) and is only very slowly converted by O_2 into the Fe(III) forms methemoglobin or metmyoglobin.[177,178] Erythrocytes contain an enzyme system for immediately reducing methemoglobin back to the Fe(II) state (see Box 15-H).

Binding of O_2 to the iron in the heme is usually considered not to cause a change in the oxidation state of the metal. However, oxygenated heme has some of the electronic characteristics of an Fe^{3+}–OO^- peroxide anion. Bonding of the heme iron to oxygen is thought to occur by donation of a pair of electrons by the oxygen to the metal. In deoxyhemoglobin the Fe(II) ion is in the "high-spin" state; four of the five 3*d* orbitals in the valence shell of the iron contain one unpaired electron and the fifth orbital contains two paired electrons. The binding of oxygen causes the iron to revert to the "low-spin" state in which all of the electrons are paired and the paramagnetism of hemoglobin is lost. The stability of heme–oxygen complexes is thought to be enhanced by "back-bonding," i.e., the donation of an electron pair from one of the filled *d* orbitals of the iron atom to form a bond with the adjacent oxygen.[179] This can be indicated symbolically as follows:



These structures, which have been formulated by assuming that one of the unshared electron pairs on O_2 forms the initial bond to the metal, are expected to lead to an angular geometry which has been observed in X-ray structures of model compounds,[180] in oxymyoglobin (Fig. 16-10),[181,182] and in oxyhemoglobin.[183] Neutron diffraction studies have shown that the outermost oxygen atom of the bound O_2 is hydrogen bonded to the H atom on the N^ε atom of the distal imidazole ring of His E7 (Fig. 16-10). Carbon monoxide binds with the C≡O axis perpendicular to the heme plane and unable to form a corresponding hydrogen bond.[184] This decreases the affinity for CO and helps to protect us from carbon monoxide poisoning.

All oxygen-carrying heme proteins have another imidazole group that binds to iron on the side opposite the oxygen site. Without this proximal imidazole group, heme does not combine with oxygen. Coordination with heterocyclic nitrogen compounds favors formation of low-spin iron complexes and simple synthetic compounds that closely mimic the behavior of myoglobin have been prepared by attaching an imidazole group by a chain of appropriate length to the edge of a heme ring.[179,185] Similar compounds bearing a pyridine ring in the fifth coordination position have a low affinity for oxygen. Thus, the polarizable imidazole ring itself seems to play a role in promoting oxygen binding. The π electrons of the imidazole ring may also participate in bonding to the iron as is indicated in the following structures.[179] The π bonding to the iron would allow the iron to back-bond more strongly to an O_2 atom entering the sixth coordination position. These diagrams illustrate another feature found frequently in heme proteins: The N–H group of the imidazole is hydrogen bonded to a peptide backbone carbonyl group.



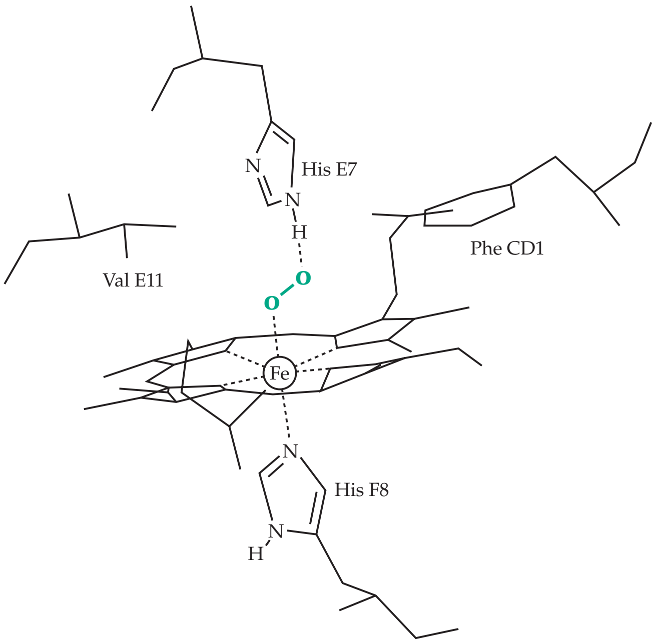


Figure 16-10 Geometry of bonding of O_2 to myoglobin and position of hydrogen bond to N^ε of the distal histidine E7 side chain. After Perutz.[182]

The coordination of the heme iron to histidine also appears to provide the basis for the cooperativity in binding of oxygen by hemoglobin.[186] The radius of high-spin iron, whether Fe(II) or Fe(III), is so large that the iron cannot fit into the center of the porphyrin ring but is displaced toward the coordinated imidazole group by a distance of ~0.04 nm for Fe(II).[187] Thus, in deoxyhemoglobin both iron and the imidazole group lie further from the center of the ring than they do in oxyhemoglobin. In the latter, the iron lies in the center of the porphyrin ring because the change to the low-spin state is accompanied by a decrease in ionic radius.[186,188] The change in protein conformation induced by this small shift in the position of the iron ion was described in Chapter 7. However, the exact nature of the linkage between the Fe position and the conformational changes is not clear.

The mechanical response to the movement of the iron and proximal histidine, described in Chapter 7, may explain this linkage. However, oxygenation may also induce a change in the charge distribution within the hydrogen-bond network of the protein. The carbonyl group shown in the foregoing structure is attached to the F helix (see Fig. 7-23) and is also hydrogen bonded to other amide groups. Electron withdrawal into the heme–oxygen complex would tend to strengthen the hydrogen bond as indicated by the resonance forms shown and also to weaken competing hydrogen bonds.[189,190] This could affect the charge distribution in the upper end of the F helix and could conceivably induce a momentary conformational change that could facilitate the rearrangement of structure that was discussed in Chapter 7 (Fig. 7-25). The $\alpha_1\beta_2$ contact in which a change of hydrogen bonding takes place is located nearby behind the F and G helices. In any event, it is remarkable that nature has so effectively made use of the subtle differences in the properties of iron induced by changes in the electron distribution within the *d* orbitals of this transition metal.

A few groups of invertebrates, e.g., the sipunculid worms, use a nonheme iron-containing protein, **hemerythrin**, as an oxygen carrier.[191,192] Its 113-residue subunits are often associated as octamers of C_4 symmetry, each peptide chain having a four-helix bundle structure (Fig. 2-22). Instead of a heme group, each monomer contains two atoms of high-spin Fe(II) held by a cluster of histidine and carboxylate side chains (see Fig. 16-20).[193,194] Hemerythrin is a member of a group of such diiron oxoproteins which are considered further in Section 8. The copper oxygen carrier **hemocyanin** is discussed in Section D.

Catalases and peroxidases. Many iron and copper proteins do not bind O_2 reversibly but "activate" it for further reaction. We will look at such metalloprotein oxidases in Chapter 18. Here we will consider heme enzymes that react not with O_2 but with peroxides. The peroxidases,[194a] which occur in plants, animals, and fungi, catalyze the following reactions (Eq. 16-6, 16-7):

$$H_2O_2 + AH_2 \rightarrow 2\ H_2O + A \qquad (16\text{-}6)$$

$$ROOH + AH_2 \rightarrow R\text{—}OH + H_2O + A \qquad (16\text{-}7)$$

Here AH_2 is an oxidizable organic compound such as an alcohol or a pair of one-electron donor molecules. Catalases, which are found in almost all aerobic cells,[194b] may sometimes account for as much as 1% of the dry weight of bacteria. The enzyme catalyzes the breakdown of H_2O_2 to water and oxygen by a mechanism similar to that employed by peroxidases. If Eq. 16-7 is rewritten with H_2O_2 for AH_2 and O_2 for A, we have the following equation:

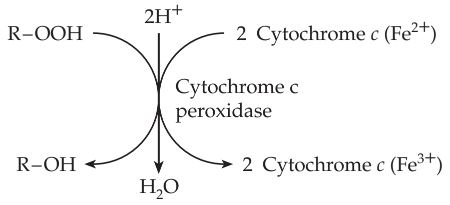
$$2\ H_2O_2 \xrightarrow{\text{Catalase}} 2\ H_2O + O_2 \qquad (16\text{-}8)$$

The action of catalase is very fast, almost 10^4 times faster than that of peroxidases. The molecular activity per catalytic center is about $2 \times 10^5\ s^{-1}$.

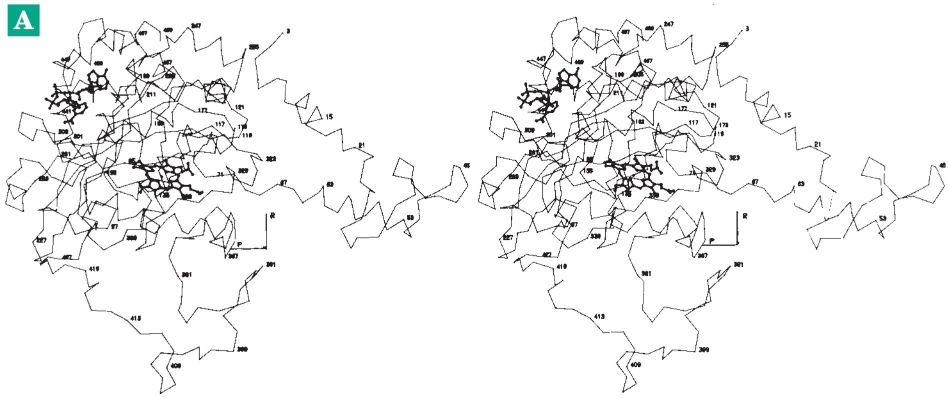
Catalase exerts a protective function by preventing the accumulation of H_2O_2 which might be harmful to cell constituents. The complete intolerance of obligate anaerobes to oxygen may result from their lack of this enzyme. Support for this protective function comes from the existence of the human hereditary condition **acatalasemia**.[195,196] Persons with extremely low catalase activity are found worldwide but are especially numerous in Korea. In Japan it is estimated that there are 1800 persons lacking catalase. Because about half of them have no symptoms, catalase might be judged unessential. However, many of the individuals affected develop ulcers around their teeth. Apparently, hydrogen peroxide produced by bacteria accumulates and oxidizes hemoglobin to methemoglobin (Box 15-H) depriving the tissues of oxygen.

Catalase from most eukaryotic species is tetrameric.[197] The protein from beef liver consists of 506-residue subunits.[198] Human catalase is similar.[198a] The proximal ligand to the heme Fe^{3+} is a tyrosinate anion (Tyr 358), while side chains of His 75 and Asn 148 lie close to the heme on the distal H_2O_2-binding side (Fig. 16-11). Larger ~650-residue fungal and bacterial catalases have a similar folding pattern but an extra C-terminal domain with a flavodoxin-like structure.[197,199] Catalase is gradually inactivated by its very reactive substrate. As isolated, beef liver catalase usually contains about two subunits in which the heme ring has been oxidatively cleaved to **biliverdin**[200] (Fig. 24-24) and various other alterations have been found.[197] Each subunit of mammalian catalases normally contains a bound molecule of NADPH which helps to protect against inactivation by H_2O_2.[201,202] Catalases from *Neurospora* and from *E. coli* contain heme *d* rather than protoporphyrin.[197,203] Some lactobacilli, lacking heme altogether, form a manganese-containing pseudocatalase.[204,205]

Of the plant peroxidases, which are found in abundance in the peroxisomes, the 40-kDa monomeric **horseradish peroxidase** has been studied the most.[206–208a] It occurs in over 30 isoforms and has an extracellular role in generating free radical intermediates for polymerization and crosslinking of plant cell wall components.[209] Secreted fungal peroxidases, e.g., such as those from *Coprinus*[210] and *Arthromyces*,[211] form a second class of peroxidases with related structures.[212] A third class is represented by **ascorbate peroxidase** from the cytosol of the pea[212–214] and by the small 34-kDa **cytochrome c peroxidase** from yeast mitochondria[215] (Fig. 16-11). The latter has a strong preference for reduced cytochrome *c* as a substrate (Eq. 16-9).[216,217]



(16-9)



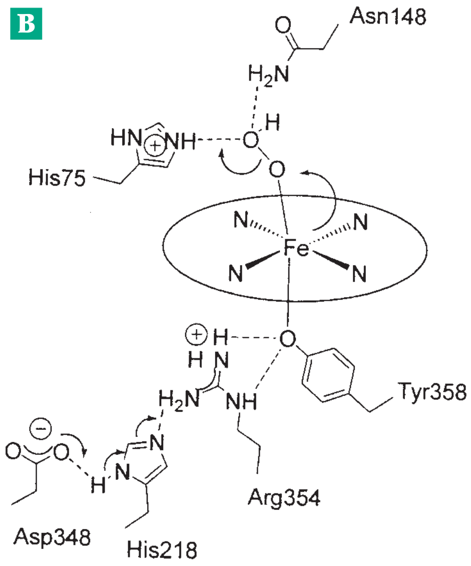


Figure 16-11 (A) Stereo drawing showing folding pattern for beef liver catalase and the positions of the NADPH (upper left) and heme (center). From Fita and Rossmann.[198] (B) Diagram of proposed structure of an Fe(III)–OOH ferric peroxide complex of human catalase (see also Fig. 16-14). A possible mechanism by which the peroxide is cleaved (step *b* in Fig. 16-14) is indicated by the arrows. His 75 and Asn 148 are directly involved, and a charge relay system below the ring may also participate. From Putnam *et al.*[198a]

Because the three-dimensional structures of the peroxidase, its reductant cytochrome *c*, and the complex of the two (Fig. 16-9) are known, cytochrome *c* peroxidase is the subject of much experimental study. Other fungal peroxidases, some of which contain manganese rather than iron, act to degrade lignin (Chapter 25).[218] A lignin peroxidase from the white wood-rot fungus *Phanerochaete chrysosporium* has a surface tryptophan with a specifically hydroxylated Cβ carbon atom which may have a functional role in catalysis.[218a,b]

The human body contains **lactoperoxidase**, a product of exocrine secretion into milk, saliva, tears, etc., and peroxidases with specialized functions in **saliva**, the **thyroid, eosinophils**,[219] and **neutrophils**.[220] The functions are largely protective but the enzymes also participate in biosynthesis. Mammalian peroxidases have heme covalently linked to the proteins, as indicated in Fig. 16-12.[220–222a]

The active site structure of peroxidases (Fig. 16-13) is quite highly conserved. As in myoglobin, an imidazole group is the proximal heme ligand, but it is usually hydrogen bonded to an aspartate carboxylate as a catalytic diad (Fig. 16-13).[223] In cytochrome *c* peroxidase there is also a buried tryptophan, which has already been highlighted in Fig. 16-9. A conserved and essential feature on the distal side is another histidine, which is hydrogen bonded to an asparagine[224] and which can also hydrogen bond to the substrate H_2O_2. Fungal peroxidases also have a conserved arginine on the distal side. However, even an octapeptide with a bound heme cut from cytochrome *c* acts as a "microperoxidase" with properties similar to those of natural peroxidases.[225]

Peroxidases and catalases contain high-spin Fe(III) and resemble metmyoglobin in properties. The enzymes are reducible to the Fe(II) state in which form they are able to combine (irreversibly) with O_2. We see that the same active center found in myoglobin and hemoglobin is present but its chemistry has been modified by the proteins. The affinity for O_2 has been altered drastically and a new group of catalytic activities for ferriheme-containing proteins has emerged.

Mechanisms of catalase and peroxidase catalysis. Attention has been focused on a series of strikingly colored intermediates formed in the presence of substrates. When a slight excess of H_2O_2 is added to a solution of horseradish peroxidase, the dark brown enzyme first turns olive green as **compound I** is formed, and then pale red as it turns into **compound II**. The latter reacts slowly with substrate AH_2 or with another H_2O_2 molecule to regenerate the original enzyme. This sequence of reactions is indicated by the colored arrows in Fig. 16-14, steps *a*–*d*.

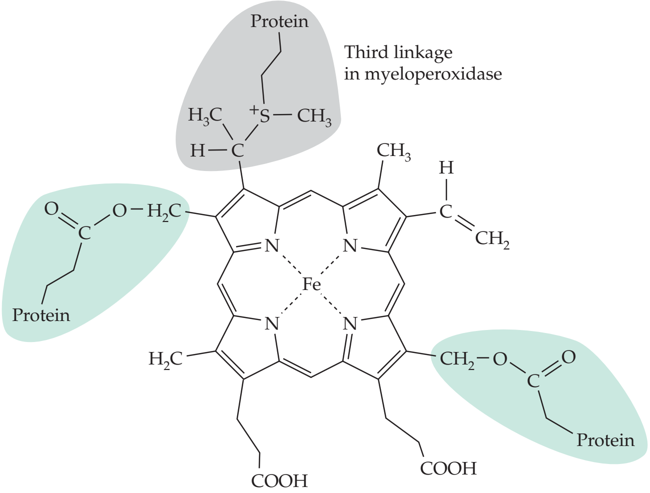


Figure 16-12 Linkage of heme to mammalian peroxidases. There are two ester linkages to carboxylate side chains from the protein.[220,221] Myeloperoxidase contains a third linkage.[222,222a]

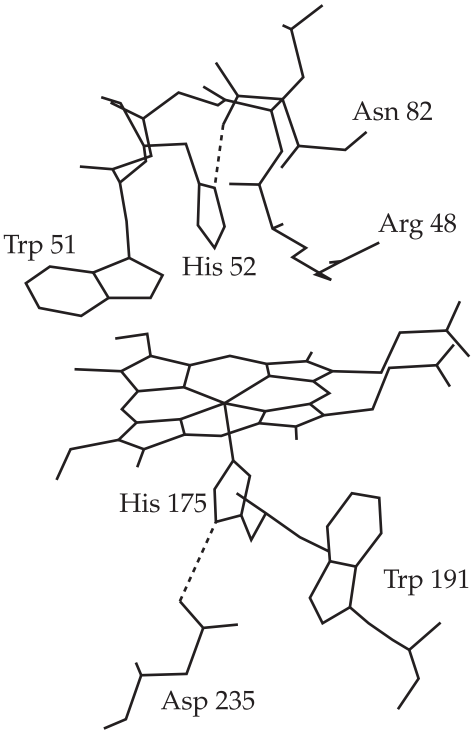


Figure 16-13 The active site of yeast cytochrome *c* peroxidase. Access for substrates is through a channel above the front edge of the heme ring as viewed by the reader. A pathway for entrance of electrons may be via Trp 191 and His 175. From Holzbaur *et al.*[206] Based on coordinates of Finzel *et al.*[215]

Titrations with such reducing agents as ferrocyanide or K_2IrBr_6 have established that compound I is converted into compound II by a one-electron reduction and compound II to free peroxidase by another one-electron reduction. Thus, the iron in compound I may formally be designated Fe(V) and that in II as Fe(IV). However, this does not tell us whether or not the oxygen atoms of H_2O_2 are present in compounds I and II. The enzyme in the Fe^{3+} form can be reduced to Fe^{2+} (Fig. 16-14, step *e*), as previously mentioned, and when the Fe^{2+} enzyme reacts with H_2O_2 it is apparently converted into compound II (Fig. 16-14, step *f*). This suggests that the latter is an Fe^{2+} complex of the peroxide anion. Here, P^{2-} represents the porphyrin ring:

$$P^{2-}—Fe^{+}(II)—OOH$$

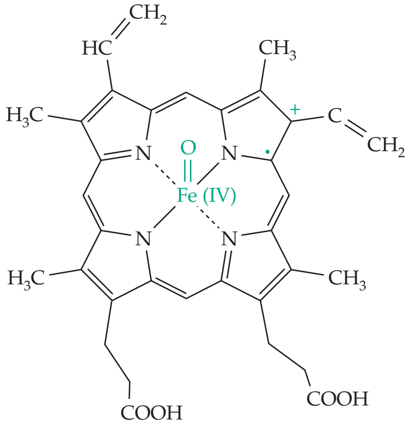
However, spectroscopic evidence suggests that compound II is a **ferryl iron** complex which could be derived from the preceding structure by addition of a proton and loss of water.[226,227]

$$P^{2-}—Fe^{2+}(IV)=O$$

Compound II

High concentrations of H_2O_2 convert II into compound III, which is thought to be the same as the **oxyperoxidase** that is formed upon addition of O_2 to the Fe(II) form of the free enzyme (Fig. 16-14, step *g*) and corresponds in structure to oxyhemoglobin.[228]

Compound I was at one time thought to be a complex of H_2O_2 or its anion with Fe(III), but its magnetic and spectral properties are inconsistent with this structure. Rather, it too appears to contain ferryl iron bound to an electron-deficient porphyrin π-cation radical.[229] The reaction with peroxide probably involves initial formation of a peroxide anion complex (Fig. 16-14, step *a*) which is cleaved with release of water (step *b*).[215,230] The resulting $Fe(V)^{+}=O$ compound is converted to compound I by transfer of a single elec-



Compound I. The unpaired electron and the positive charge are delocalized over the porphyrin ring and perhaps into the proximal histidine ring

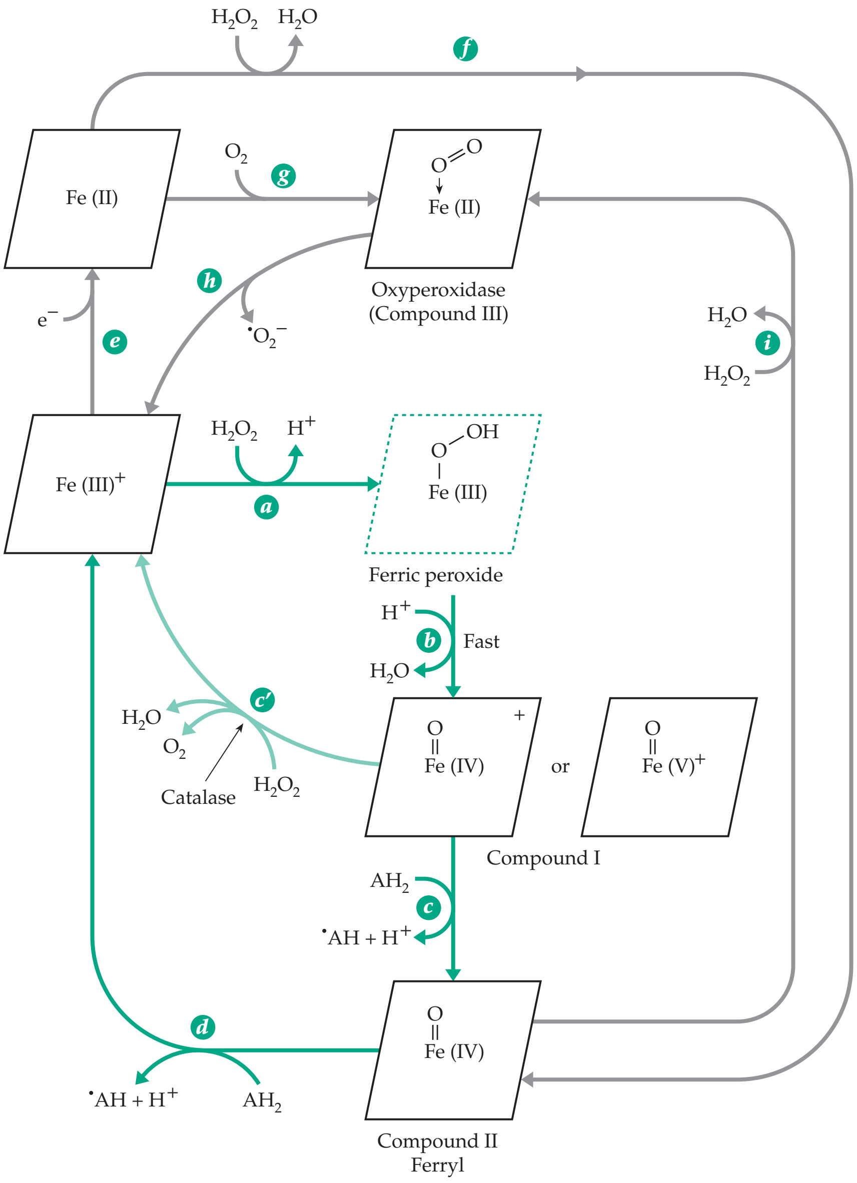


Figure 16-14 The catalytic cycles and other reactions of peroxidases and catalases. The principal cycle for peroxidases is given by the colored arrows. That of catalases is smaller, making use of step *a*, *b*, and *c*′, which is marked by a light green line.

tron from the porphyrin to the iron. In cytochrome *c* peroxidase compound I contains a free radical on the nearby Trp 191 ring instead of on the porphyrin radical.[216] Consistent with this is the fact that horseradish peroxidase contains phenylalanine in place of Trp 191.

If we consider the fate of substrate AH_2 during the action of a peroxidase, we see that donation of an electron to compound I to convert it into II (Fig. 16-14, step *c*) will generate a free radical •AH as well as a proton. The radical may then donate a second electron to II to form the free enzyme. Alternatively, a second molecule of AH_2 may react (Fig. 16-14, step *c*) to form a second radical •AH. The two •AH radicals may then disproportionate to form A and AH_2 or they may leave the enzyme and react with other molecules in their environment. Compound II of horse radish peroxidase is able to exchange the oxygen atom of its Fe(IV)=O center with water rapidly at pH 7, presumably by donation of a proton from the nearby histidine side chain (corresponding to His 52 of Fig. 16-13).[227,230a,b] This histidine presumably also functions in proton transfer during reactions with substrates (see Fig. 16-11B).[224]

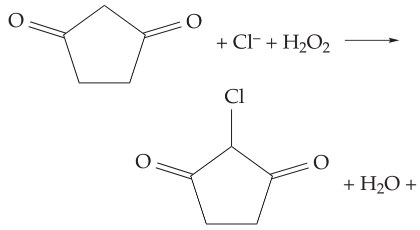
The catalase compound I appears to be converted in a two-electron reduction by H_2O_2 directly to free ferricatalase without intervention of compound II (Fig. 16-14, step *c*′). The catalytic histidine probably donates a proton to help form water from one of the oxygen

$$\underset{\text{Compound I}}{{}^{\bullet}\text{P}^{+}\text{—Fe(IV)=O}} + H_2O_2 \rightarrow H_2O + O_2 + \text{P—Fe(III)}^{+} \qquad (16\text{-}10)$$

atoms of the H_2O_2. Nevertheless, compound II does form slowly, especially if a slow substrate such as ethanol is present. The previously mentioned bound NADPH apparently reduces compound II formed in this way, converting the inactivated enzyme back to active catalase.[201,231] This may involve unusual one-electron oxidation steps for the NADPH.[231] Under some circumstances compound II of peroxidases reacts with substrates in a two-electron process[232–235] with transfer of an oxygen atom to the substrate, a characteristic also of reactions catalyzed by cytochromes P450 (see Eq. 18-57).

Haloperoxidases. Many specialized peroxidases are active in halogenation reactions. **Chloroperoxidases** from fungi[236,237] catalyze chlorination reactions like that of Eq. 16-11 using H_2O_2 and Cl^- as well as the usual peroxidase reaction.

Chloroperoxidase is isolated in a low-spin Fe(III) state. The reduced Fe(II) enzyme is a high-spin form

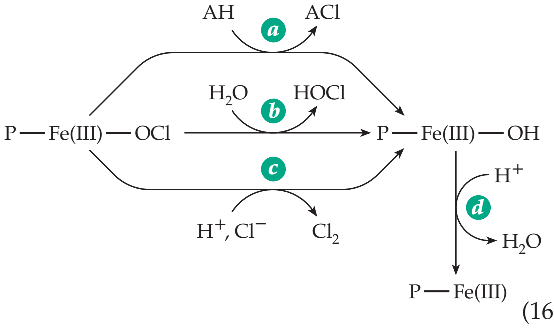


(16-11)

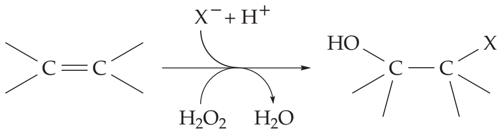
with spectroscopic properties similar to those of the oxygenases of the cytochrome P450 family, which are discussed in Chapter 18.[238,239] Like the cytochromes P450 chloroperoxidase contains a thiolate group of a cysteine side chain as the fifth iron ligand.[235,239,240] Chloroperoxidase forms compounds I and II, as do other peroxidases. A chloride ion may combine with compound I to form a complex of hypochlorite with the Fe(III) heme.

$$^{\bullet}P^{+}\text{—}Fe(IV)\text{=}O + Cl^{-} \rightarrow P\text{—}Fe(III)^{+}\text{—}OCl \qquad (16\text{-}12)$$

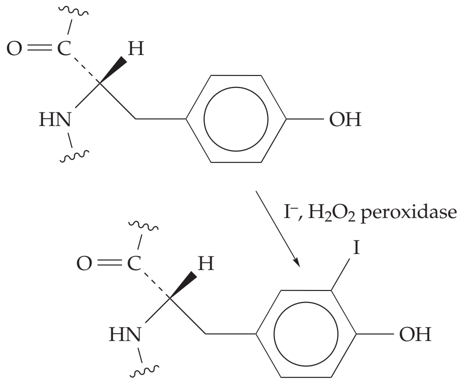
This intermediate could then halogenate substrate AH (Eq. 16-13, step *a*), lose HOCl (Eq. 16-13, step *b*), or generate Cl_2 by reaction with Cl^- (Eq. 16-13, step *c*). These are all well-established reactions for the enzyme. In each case the chlorine in the peroxidase complex can be viewed as an electrophile which is transferred to an attacking nucleophile. A fourth reaction that can go through the same intermediate is conversion of alkenes to α,β-halohydrins (Eq. 16-14).[241] **Lactoperoxidase** of milk, reacting with $I^- + H_2O_2$, promotes an analogous iodination of tyrosine and histidine residues of proteins. With radioactive $^{125}I^-$ or $^{131}I^-$ it provides a convenient and much used method for labeling of proteins in exposed surfaces of membranes (Eq. 16-15).[242] Iodinated tyrosine derivatives are formed in the thyroid gland by a similar reaction catalyzed by **thyroid peroxidase**.[232] Even horseradish peroxidase can oxidize iodide ions but neither it nor lactoperoxidase will carry out chlorination or bromination reactions.



(16-13)



(16-14)



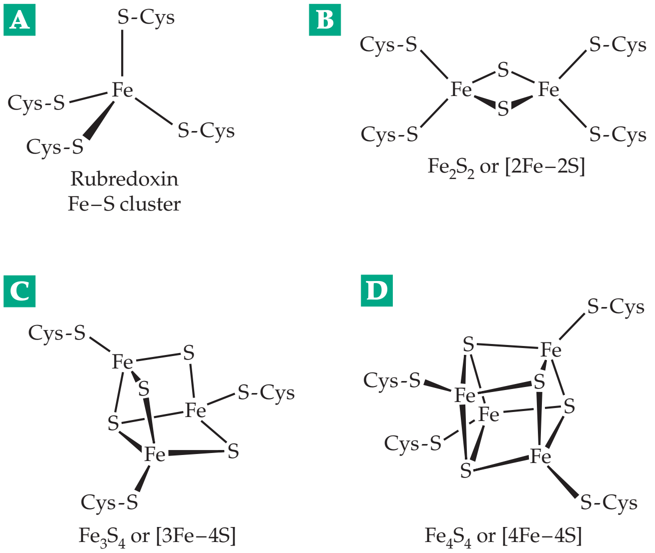
(16-15)

Myeloperoxidase, present in specialized lysosomes of polymorphonuclear leukocytes (neutrophils),[243] utilizes H_2O_2 and a halide ion to kill ingested bacteria.[244,245] Phagocytosis induces increased respiration by the leukocyte and generation of H_2O_2, partly by the membrane-bound NADPH oxidase described in Chapter 18. Some of the H_2O_2 is used by myeloperoxidase to attack the bacteria, apparently through generation of HOCl by peroxidation of Cl^-. Human myeloperoxidase is a tetramer of two 466-residue chains and two 108-residue chains, which carry the covalently linked heme.[222a,246] Another oxygen-dependent killing mechanism that may also be used by neutrophils is the generation of the reactive **singlet oxygen**.[247] This can occur by reaction of hypochlorite with H_2O_2 (Eq. 16-16) or from an enzyme-bound hypochlorite intermediate such as that shown in Eq. 16-13. Hereditary deficiency of myeloperoxidase is relatively common.[245]

$$OCl^- + H_2O_2 \rightarrow \underset{\text{singlet } O_2}{O_2(^1\Delta g)} + H_2O + Cl^- \qquad (16\text{-}16)$$

Lactoperoxidase[248] and chloroperoxidase[249] also generate singlet oxygen. The possible biological significance is discussed in Chapter 18. Eosinophil peroxidase appears to promote formation of **hydroxyl radicals**.[250] **Bromoperoxidases** are found in many red and green marine algae.[251] Many of them contain **vanadium** and function by a mechanism different than that used by heme peroxidases (see Section G).[252,253]

Another related nonheme enzyme is the selenoprotein glutathione peroxidase. It reacts by a mechanism very different (Eq. 15-59) from those discussed here, as does **NADPH peroxidase**, a flavoprotein with a cysteine sulfinate side chain in the active site.[254–255a] A lignin-degrading peroxidase from the white wood rot fungus *Phanerochaete chrysosporium* is a simple heme protein,[256] while other peroxidases secreted by this organism contain Mn.[257,258]



Rubredoxin Fe–S cluster

Fe_2S_2 or [2Fe–2S]

Fe_3S_4 or [3Fe–4S]

Fe_4S_4 or [4Fe–4S]

Figure 16-15 Four different iron–sulfur clusters of a type found in many proteins. From Beinert[259] with permission.

7. The Iron–Sulfur Proteins

Not all of the iron within cells is chelated by porphyrin groups. Hemerythrin (see Fig. 16-20) has been known for many years, but the general significance of nonheme iron proteins was not appreciated until large-scale preparation of mitochondria was developed by Crane in about 1945. The iron content of mitochondria was found to far exceed that of the heme proteins present. In 1960, Beinert, who was studying the mitochondrial dehydrogenase systems for succinate and for NADH, observed that when the electron transport chain was partially reduced by these substrates and the solutions were frozen at low temperature and examined, a strong EPR signal was observed at $g = 1.94$. The signal was obtained only upon reduction by substrate, and fractionation pointed to the nonheme iron proteins. Six or more proteins of this type are involved in the mitochondrial electron transport chain (Eq. 18-1), and numerous others have become recognized as members of the same large family of **iron–sulfur proteins** (Fe–S proteins).[259,260]

Ferredoxins, high-potential iron proteins, and rubredoxins. The presence of nonheme iron proteins is most evident in the anaerobic clostridia, which contain no heme. It was from these bacteria that the first Fe–S protein was isolated and named **ferredoxin**. This protein has a very low reduction potential of $E^{\circ\prime}$(pH 7) = –0.41 V. It participates in the pyruvate–ferredoxin oxidoreductase reaction (Eq. 15-35), in nitrogen fixation in some species, and in formation of H_2. A small green-brown protein, the ferredoxin of *Peptococcus aerogenes* contains only 54 amino acids but complexes eight atoms of iron. If the pH is lowered to ~1, eight molecules of H_2S are released. Thus, the protein contains eight "labile sulfur" atoms in an iron sulfide linkage. There are also eight iron atoms.

Another group of related electron carriers, the **high-potential iron proteins** (HIPIP) contain four labile sulfur and four iron atoms per peptide chain.[261–266] X-ray studies showed that the 86-residue polypeptide chain of the HIPIP of *Chromatium* is wrapped around a single **iron–sulfur cluster** which contains the side chains of four cysteine residues plus the four iron and four sulfur atoms (Fig. 16-15D).[261] This kind of cluster is referred to as [4Fe–4S], or as Fe_4S_4. Each cysteine sulfur is attached to one atom of Fe, with the four iron atoms forming an irregular tetrahedron with an Fe–Fe distance of ~0.28 nm. The four labile sulfur atoms (S^{2-}) form an interpenetrating tetrahedron 0.35 nm on a side with each of the sulfur atoms bonded to three iron atoms. The cluster is ordinarily able to accept only a single electron. The iron–sulfur cluster structure was a surprise, but after its discovery it was found that ions such as $[Fe_4S_4(S{-}CH_2CH_2COO^-)_4]^{6-}$ assemble spontaneously from their components and have a similar cluster structure.[259,266a] Thus, living things have simply improved upon a natural bonding arrangement.

The bacterial ferredoxins from *Peptococcus, Clostridium* (Fig. 16-16B),[267,268] *Desulfovibrio*, and other anaerobes each contain two Fe_4S_4 clusters with essentially the same structure as that of the *Chromatium* HIPIP.[267,269] Each cluster can accept one electron. Much of the amino acid sequence in the first half of the ferredoxin chain is repeated in the second half, suggesting that the chain may have originated as a result of gene duplication. Many invariant positions are present in the sequence, including those of the cysteine residues forming the Fe–S cluster. Ferredoxins with single Fe_4S_4 clusters are also known.[270]

The ferredoxins have reduction potentials $E^{\circ\prime}$ (pH 7) from about – 0.4 V to as low as – 0.6 V. However, the corresponding values for HIPIP proteins range from +0.05 to +0.50 V at pH 7.[271] This wide range of potentials initially seemed strange because the structures of the active centers of both the clostridial ferredoxins and the *Chromatium* HIPIP appear virtually identical.[272] Part of the explanation lies in the fact that Fe_4S_4 clusters can exist in three oxidation states (Eq. 16-17) that differ, one from another, by a single electron.[273]

$$\underset{\text{Reduced Fd}}{[Fe_4S_4]^+} \xrightarrow[e^-]{} \underset{\substack{\text{Oxidized Fd;}\\ \text{reduced high}\\ \text{potential iron}\\ \text{protein}}}{[Fe_4S_4]^{2+}} \xrightarrow[e^-]{} \underset{\substack{\text{Oxidized high}\\ \text{potential iron}\\ \text{protein; Super-}\\ \text{oxidized Fd}}}{[Fe_4S_4]^{3+}} \qquad (16\text{-}17)$$

Here the charges shown are those on the cluster. The cysteine ligands from the protein each add an additional negative charge. The *Chromatium* HIPIP and the clostridial ferredoxins have the middle oxidation state in common. The cluster is a little smaller in the more oxidized states; in the *Chromatium* HIPIP the Fe–Fe distance changes from 0.281 to 0.272 nm upon oxidation.

Rubredoxins. The simplest of the Fe–S proteins are the rubredoxins. These proteins contain iron but no labile sulfur. The rubredoxin of *Clostridium pasteurianum* is a 54-residue peptide containing four cysteines whose side chains form a distorted tetrahedron about a single iron atom (Fig. 16-16A).[274] Not shown for any

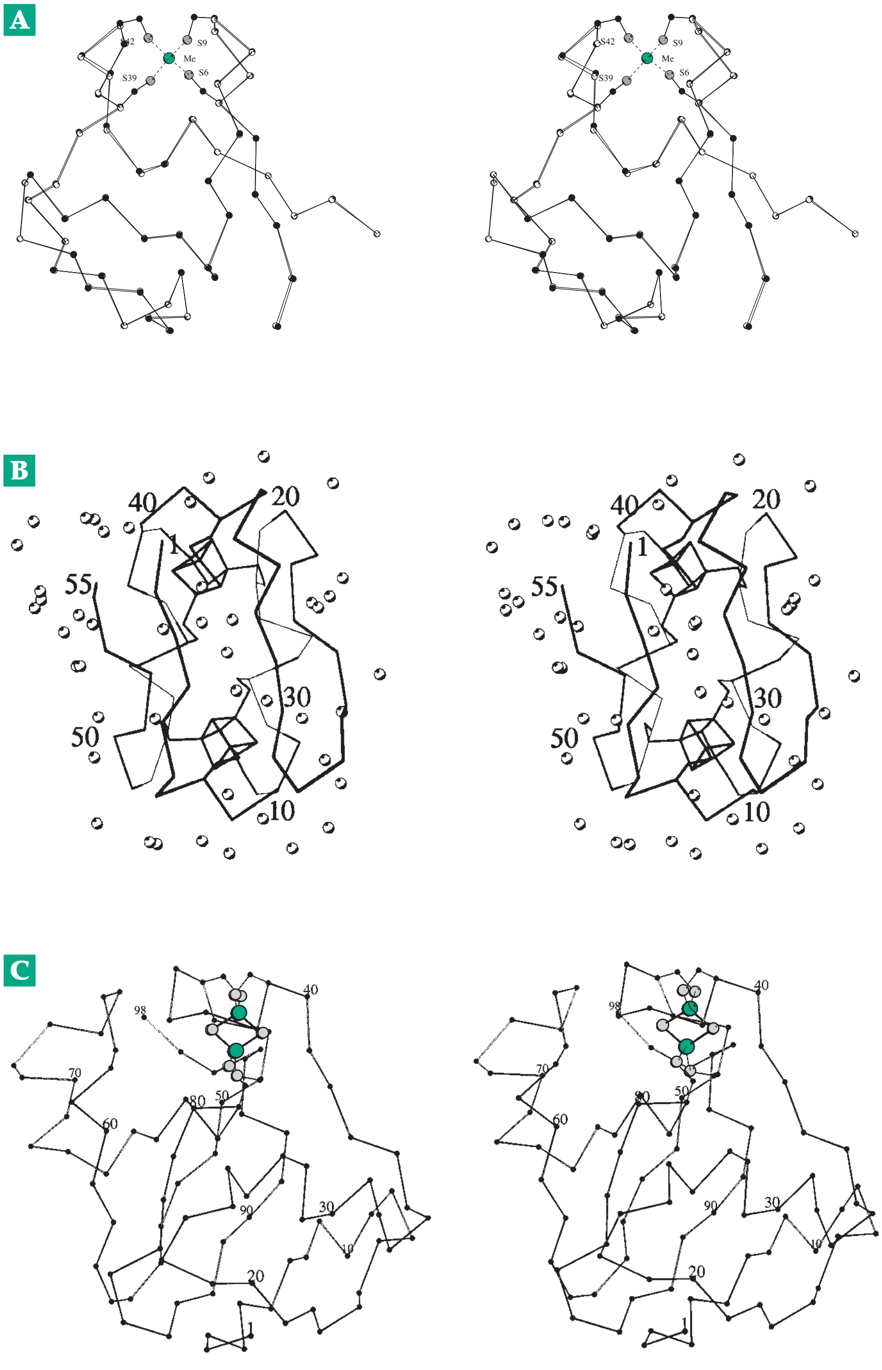
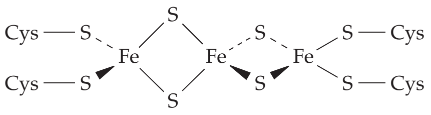


Figure 16-16 (A) Superimposed stereoscopic α-carbon traces of the peptide chain of rubredoxin from *Clostridium pasteurianum* with either Fe^{3+} (solid circles) or Zn^{2+} (open circles) bound by four cysteine side chains. From Dauter *et al.*[274] (B) Alpha-carbon trace for ferredoxin from *Clostridium acidurici*. The two Fe_4S_4 clusters attached to eight cysteine side chains are also shown. The open circles are water molecules. Based on a high-resolution X-ray structure by Duée *et al.*[267] Courtesy of E. D. Duée. (C) Polypeptide chain of a chloroplast-type ferredoxin from the cyanobacterium *Spirulina platensis*. The Fe_2S_2 cluster is visible at the top of the molecule. From Fukuyama *et al.*[276] Courtesy of K. Fukuyama.

of the structures in Fig. 16-16 are NH—S hydrogen bonds that connect backbone NH groups of the peptide chain to the sulfur atoms of the cysteine groups, forming the clusters.[275] These bonds may have important effects on properties of the cluster. Rubredoxins also participate in electron transport and can substitute for ferredoxins in some reactions. Larger 14-kDa and 18-kDa rubredoxins able to bind two iron ions participate in electron transport in a hydroxylase system of *Pseudomonas* (Chapter 18).[277] A smaller 7.9-kDa 2-Fe **desulforedoxin** functions in sulfate-reducing bacteria.[278]

Chloroplast-type ferredoxins. Members of a large class of [2Fe−2S] or Fe_2S_2 ferredoxins each contain two iron atoms and two labile sulfur atoms with the linear structure of Fig. 16-15B.[279,279a] Best known are the chloroplast ferredoxins, which transfer electrons from photosynthetic centers of chloroplasts to the flavoprotein reductase that reduces $NADP^+$ to NADPH.[280,280a] The structure of a cyanobacterial protein of this type is shown in Fig. 16-16C.[276] A second group of Fe_2S_2 ferredoxins are found in bacteria including *E. coli*[281] and in human mitochondria.[282] For example, in steroid hormone-forming tissues the ferredoxin **adrenodoxin**[282a] carries electrons to cytochromes P450. Its Fe_2S_2 center receives electrons from adrenodoxin reductase (Chapter 15 banner).[282b] The nitrogen-fixing *Clostridium pasteurianum* also contains a ferredoxin of this class.[283] In *Pseudomonas putida* (Chapter 18) the related 106-residue **putidaredoxin** transfers electrons to cytochromes P450.[284]

The 3Fe−4S clusters. A 106-residue ferredoxin from *Azotobacter vinelandii* contains seven iron atoms in two Fe−S clusters that operate at very different redox potentials of −0.42 and +0.32 V.[285] Other similar seven-iron proteins are known.[286] From EPR measurements it appeared that both clusters function between the 2^+ and 3^+ (oxidized and superoxidized) states of Eq. 16-17, despite the widely differing potentials. A super-reduced all-Fe(II) form with $E^{°′}$ (pH 7) = −0.70 V can also be formed.[286] X-ray crystallographic studies have revealed that the protein contains one Fe_4S_4 cluster and one Fe_3S_4 cluster (Fig. 16-15C).[286–288] The structures and environments of the Fe_3S_4 clusters are similar to those of Fe_4S_4 clusters but they lack one iron and one cysteine side chain. An Fe_4S_4 cluster may sometimes lose S^{2-} to form an Fe_3S_4 cluster such as the one in the *A. vinelandii* ferredoxin. A less likely possibility is isomerization to a linear Fe_3S_4 structure.[289]



Aconitase (Eq. 13-17) isolated under aerobic conditions contains an Fe_3S_4 cluster and is catalytically inactive. Incubation with Fe^{2+} activates the enzyme and reconverts the Fe_3S_4 to an Fe_4S_4 cluster (Fig. 13-4).[260,289,290]

Properties of iron−sulfur clusters. These clusters were viewed for many years as unstable and unable to exist outside of a protein. However, if protected from oxygen and manipulated in the presence of soluble organic thiols they are stable and "cofactor-like."[260] Intact clusters can be "extruded" from proteins by treatment with thiols in nonaqueous media. Both Fe_4S_4 and Fe_2S_2 clusters as well as more complex forms have been synthesized[291] and nonenzymatic cluster interconversions have been demonstrated. Binding to proteins stabilizes the clusters further, but some (Fe−S) proteins are labile and difficult to study. This is evidently because of partial exposure of the cluster to the surrounding solvent. Not only can O_2 cause oxidation of exposed clusters[291a] but also superoxide,[292] nitric oxide, and peroxynitrite can react with the iron. Aconitase has only three cysteine side chains available for coordination with Fe and the protein is unstable. Apparently, a superoxidized $[Fe_4S_4]^{3+}$ cluster is formed in the presence of O_2 but loses Fe^{2+} to give an $[Fe_3S_4]^+$ cluster.[293]

Another interesting cluster conversion is the joining of two Fe_2S_2 clusters in a protein to form a single Fe_4S_4 cluster at the interface between a dimeric protein. Such a cluster is present in the nitrogenase iron protein (Fig. 24-2) and probably also in biotin synthase.[294] The clusters in such proteins can also be split to release the monomers.

Synthetic iron−sulfur clusters have weakly basic properties[273] and accept protons with a pK_a of from 3.9 to 7.4. Similarly, one clostridial ferredoxin, in the oxidized form, has a pK_a of 7.4; it is shifted to 8.9 in the reduced form.[295] If we designate the low-pH oxidized form of such a protein as HOx^+ and the reduced form as HRed, we can depict the reduction of each Fe_4S_4 cluster as follows.

$$HOx^+ + e^- \rightarrow HRed \qquad (16\text{-}18)$$

Comparing with Eq. 6-64 and using the Michaelis pH functions (first two terms of Eq. 7-13) for HOx^+ and HRed, it is easy to show that the value of $E°′$ ($E_{1/2}$) at which equal amounts of oxidized protein (HOx^+ + Ox) and reduced protein (HRed + Red^-) are present is given by Eq. 16-19, in which K_{ox} and K_{red} are the K_a values for dissociation of the protonated oxidized and reduced forms, respectively.

$$E_{1/2} = E^0\ (\text{low pH}) + 0.0592 \log\left[\frac{(1 + K_{red}\ /[H^+])}{(1 + K_{ox}\ /[H^+])}\right] V \qquad (16\text{-}19)$$

At the high pH limit this becomes $E_{1/2} = E°$ (low pH) + 0.0592 (pK_{ox}–pK_{red}) and V = –0.371 + 0.0592 (7.4–8.9) V = –0.431 V. Thus, the value of $E_{1/2}$ changes from –0.371 to –0.460 V as the pH is increased. In the pH range between the pK_a values of 7.4 and 8.9 reduction of the protein will lead to binding of a proton from the medium and oxidation to loss of a proton. Human and other vertebrate ferredoxins also show pH-dependent redox potentials.[282] This suggests, as with the cytochromes, a possible role of Fe–S centers in the operation of proton pumps in membranes. Nevertheless, many ferredoxins, such as that of *C. pasteurianum*, show a constant value of $E°'$ from pH 6.3 to 10[296] and appear to be purely electron carriers.

Both the iron and labile sulfur can be removed from Fe–S proteins and the active proteins can often be reconstituted by adding sulfide and Fe^{2+} ions. Using this approach, the natural isotope ^{56}Fe (nuclear spin zero) has been exchanged with ^{57}Fe, which has a magnetic nucleus[297] and ^{32}S has been replaced by ^{77}Se. The resulting proteins appear to function naturally and give EPR spectra containing hyperfine lines that result from interaction of these nuclei with unpaired electrons in the Fe_4S_4 clusters (Fig. 16-17). These observations suggest that electrons accepted by Fe_4S_4 clusters are not localized on a single type of atom but interact with nuclei of both Fe and S. The native proteins as well as many mutant forms are being studied by NMR[299,300] and other spectroscopic techniques (Fig. 16-18),[260] by theoretical computations,[301–304] and by protein engineering and "rational design."[305,306]

Functions of iron–sulfur enzymes. Numerous iron–sulfur clusters are present within the membrane-bound electron transport chains discussed in Chapter 18. Of special interest is the Fe_2S_2 cluster present in a protein isolated from the cytochrome *bc* complex (complex III) of mitochondria. First purified by Rieske *et al.*,[307] this protein is often called the **Rieske iron–sulfur protein**.[308] Similar proteins are found in cytochrome *bc* complexes of chloroplasts.[125,300,309,310] In

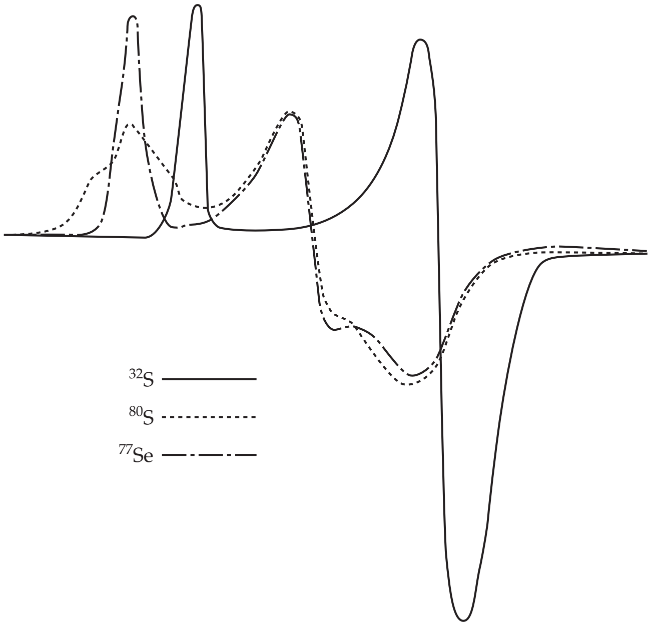


Figure 16-17 Electron paramagnetic resonance spectrum of the Fe–S protein putidaredoxin in the natural form (^{32}S) and with labile sulfur replaced by selenium isotopes. Well-developed shoulders are seen in the low-field end of the spectrum of the ^{77}Se (spin = 1/2)-containing protein. From Orme-Johnson *et al.*[298] Courtesy of W. H. Orme-Johnson.

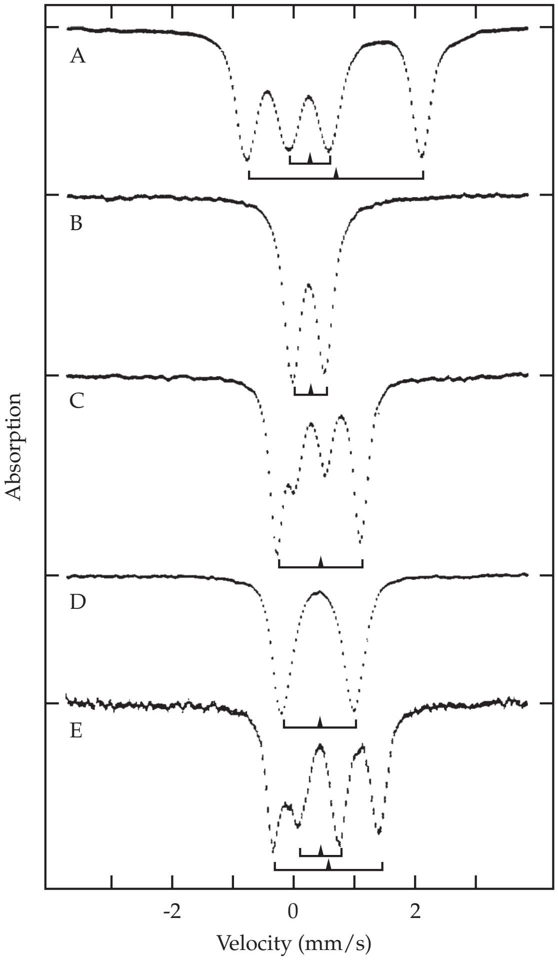
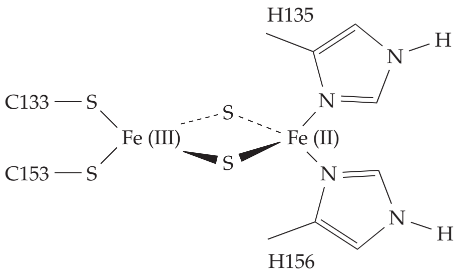


Figure 16-18 Mössbauer X-ray absorption spectra of iron–sulfur clusters. (See Chapter 23 for a brief description of the method.) Quadrupole doublets are indicated by brackets and isomer shifts are marked by triangles. (A) $[Fe_2S_2]^{1+}$ cluster of the Rieske protein from *Pseudomonas mendocina*, at temperature *T* = 200 K. (B) $[Fe_3S_4]^{1+}$ state of *D. gigas* ferredoxin II, *T* = 90 K. (C) $[Fe_3S_4]^0$ state of *D. gigas* ferredoxin II, *T* = 15 K. (D) $[Fe_4S_4]^{2+}$ cluster of *E. coli* FNR protein, *T* = 4.2 K. (E) $[Fe_4S_4]^{1+}$ cluster of *E. coli* sulfite reductase, *T* = 110 K. From Beinert *et al.*[260]

some bacteria Rieske-type proteins deliver electrons to oxygenases.[300] The 196-residue mitochondrial protein has an unusually high midpoint potential, E_m of ~ 0.30 V. The Fe_2S_2 cluster, which is visible in the atomic structure of complex III shown in Fig. 18-8, is coordinated by two cysteine thiolates and two histidine side chains. In a *Rhodobacter* protein they occur in the following conserved sequences:[311] C133-T-H-L-G-C138 and C153-P-C-H-G-S158. One iron is bound by C133 and C153 and the other by H135 and H156 as follows:

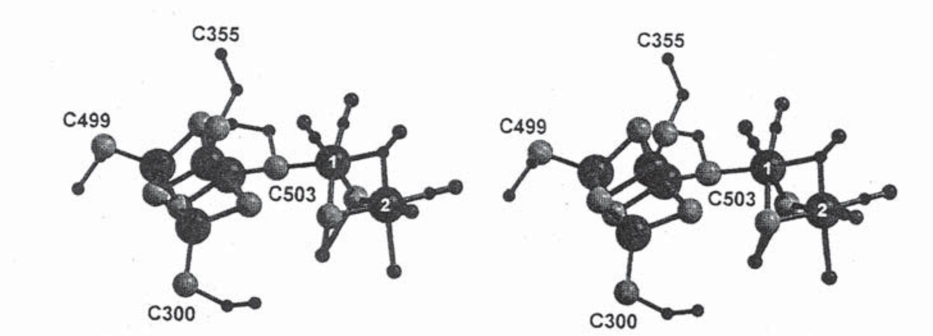


These proteins may also have an ionizable group with a pK_a of ~8.0, perhaps from one of the histidines that is linked to the oxidation–reduction reaction.

Iron–sulfur clusters are found in flavoproteins such as NADH dehydrogenase (Chapter 18) and trimethylamine dehydrogenase (Fig. 15-9) and in the siroheme-containing **sulfite reductases** and **nitrite reductases**.[312] These two reductases are found both in bacteria and in green plants. Spinach nitrite reductase,[313] which is considered further in Chapter 24, utilizes reduced ferredoxin to carry out a six-electron reduction of NO_2^- to NH_3 or of SO_3^{2-} to S^{2-}. The 61-kDa monomeric enzyme contains one siroheme and one Fe_4S_4 cluster. A sulfite reductase from *E. coli* utilizes NADPH as the reductant. It is a large $\beta_8\alpha_4$ oligomer.[312] The 66-kDa α chains contain bound flavin (4 FAD + 4 FMN),[312,314,315] while the 64-kDa β subunits contain both siroheme and a neighboring Fe_4S_4 cluster (Fig. 16-19).[89,304,312,316] The iron of the siroheme and the closest iron atom of the cluster are bridged by a single sulfur atom of a cysteine side chain.

A somewhat similar double cluster is present in **all-Fe hydrogenases** from *Clostridium pasteurianum*.[316a,b,c] These enzymes also contain two or three Fe_4S_4 clusters and, in one case, an Fe_2S_2 cluster.[316a] At the presumed active site a special **H cluster** consists of an Fe_3S_4 cluster with one cysteine sulfur atom shared by an adjoining Fe_2S_2 cluster: Because many hydrogenases contain nickel, their chemistry and functions are discussed in Section C,2.

The role of the iron–sulfur clusters in many of the proteins that we have just considered is primarily one of single-electron transfer. The Fe–S cluster is a place for an electron to rest while waiting for a chance to react. There may sometimes be an associated proton pumping action. In a second group of enzymes, exemplified by aconitase (Fig. 13-4), an iron atom of a cluster functions as a **Lewis acid** in facilitating removal of an –OH group in an α,β dehydration of a carboxylic acid (Chapter 13). A substantial number of other bacterial dehydratases as well as an important plant dihydroxyacid dehydratase also apparently use Fe–S clusters in a catalytic fashion.[317] Fumarases A and B from *E. coli*,[317] L-serine dehydratase of a *Peptostreptococcus* species,[317–319] and the dihydroxyacid



The hydrogenase Fe–S H cluster

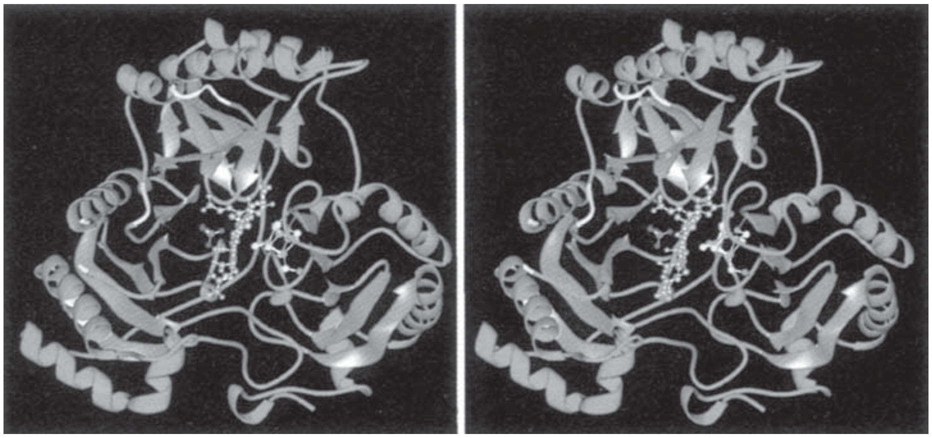


Figure 16-19 Stereoscopic view of *E. coli* assimilatory sulfite reductase. The siroheme (Fig. 16-6) is in the center with one edge toward the viewer and the Fe_4S_4 cluster is visible on its right side. A single S atom from a cysteine side chain bridges between the Fe of the siroheme and the Fe_4S_4 cluster. A phosphate ion is visible in the sulfite-binding pocket at the left center of the siroheme. From Crane *et al.*[312] Courtesy of E. D. Getzoff.

dehydratase[320,321] may all use their Fe_4S_4 clusters in a manner similar to that of aconitase (Eq. 13-17). However, the Fe–S enzymes that dehydrate *R*-lactyl-CoA to crotonyl-CoA,[322] 4-hydroxybutyryl-CoA to crotonyl-CoA,[323,324] and *R*-2-hydroxyglutaryl-CoA to *E*-glutaconyl-CoA[325] must act by quite different mechanisms, perhaps similar to those utilized in vitamin B_{12}-dependent reactions. In these enzymes, as in pyruvate formate lyase (Eq. 15-38), the Fe–S center may act as a radical generator.

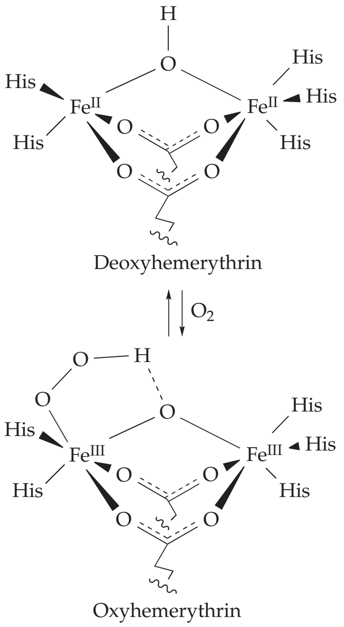
The molybdenum-containing enzymes considered in Section F also contain Fe–S clusters. Nitrogenases (Chapter 24) contain a more complex Fe–S–Mo cluster. Carbon monoxide dehydrogenase (Section C) contains 2 Ni, ~11 Fe, and 14 S^{2-} as well as Zn in a dimeric structure. In these enzymes the Fe–S clusters appear to participate in catalysis by undergoing alternate reduction and oxidation.

For a few enzymes such as aconitase and amido-transferases,[326] an Fe–S cluster plays a **regulatory role** in addition to or instead of a catalytic function.[327] Cytosolic aconitase is identical to **iron regulatory proteins 1** (IRP1), which binds to iron responsive elements in RNA to inhibit translation of genes asso-ciated with iron uptake. A high iron concentration promotes assembly of the Fe_4S_4 cluster (see Chapter 28, Section C,6). Another example is provided by the *E. coli* transcription factor SOXR. This protein, which controls a cellular defense system against oxygen-derived superoxide radicals, contains two Fe_2S_2 centers. Oxidation of these centers by superoxide radicals appears to induce the transcription of genes encoding superoxide dismutase and other proteins involved in protecting cells against oxidative injury (Chapter 18).[328–330]

8. The (μ-oxo) Diiron Proteins

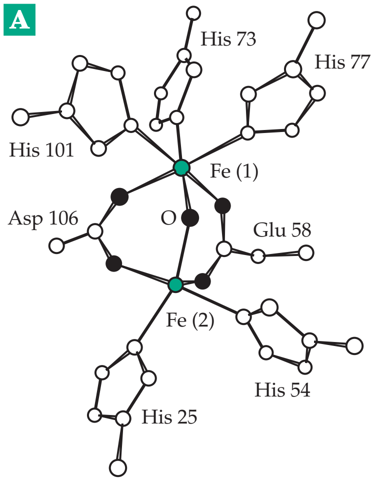
Both the ferroxidase center of ferritin (Fig. 16-4) and the oxygen-carrying hemerythrin are members of a family of diiron proteins with similar active site structures.[331,332] The pair of iron atoms, with the bridging (μ) ligands such as O_2, HO^-, HOO^-, O_2^{2-}, is often held between four α helices, as can be seen in Fig. 2-22 and in Fig. 16-20C. In most cases each iron is ligated by at least one histidine, one glutamate side chain, and frequently a tyrosinate side chain. As many as three side chains may bind to each iron. In addition, one or two carboxylate groups from glutamate or aspartate side chains bridge to both irons, as does the μ-oxo group, which is typically H_2O or ^-OH. Examples of this structure are illustrated in Figs. 16-4 and 16-20. Although the active sites of all of the diiron proteins appear similar, the chemistry of the catalyzed reactions is varied.

Hemerythrin. When both iron atoms are in the Fe(II) state, hemerythrin, like hemoglobin, functions as a carrier of O_2. In the oxidized Fe(III) **methemerythrin** form the iron atoms are only 0.32 nm apart. Three bridging (μ) groups lie between them: two carboxylate groups and a single oxygen atom which may be either O^{2-} or OH^-. One coordination position on one of the hexacoordinate iron atoms is open and appears to be the site of binding of oxygen. The O_2 is thought to accept two electrons, oxidizing the two iron atoms to Fe(III) and itself becoming a peroxide dianion O_2^{2-}. The process is completely reversible. The conversion of the oxygen to a bound peroxide ion is supported by studies of resonance Raman spectra (see Chapter 23) which also suggest that the peroxide group is protonated. In the diferrous protein the μ-oxo bridge is thought to be an ^-OH group. Upon oxygenation the proton could be shared with or donated to the peroxo group (Eq. 16-20).[334] A similar binding of O_2 as a peroxide dianion appears to occur in the copper-containing hemocyanins (Section C,3).



(16-20)

Purple acid phosphatases. Diiron-tyrosinate proteins with acid phosphatase activity occur in mammals, plants, and bacteria. Most are basic glycoproteins with an intense 510- to 550-nm light absorption band. Well-studied members come from beef spleen, from the uterine fluid of pregnant sows (**uteroferrin**),[335] and from human macrophages and osteoclasts.[336–336b] One of the two iron atoms is usually in the Fe(III) oxidation state, but the second can be reduced to Fe(II) by mild reductants such as ascorbate. This half-reduced form is enzymatically active and has a pink color and a characteristic EPR signal. Treatment with oxidants such as H_2O_2 or hexacyannoferrate (III)



Methemerythrin
(Fe^{3+}/Fe^{3+})

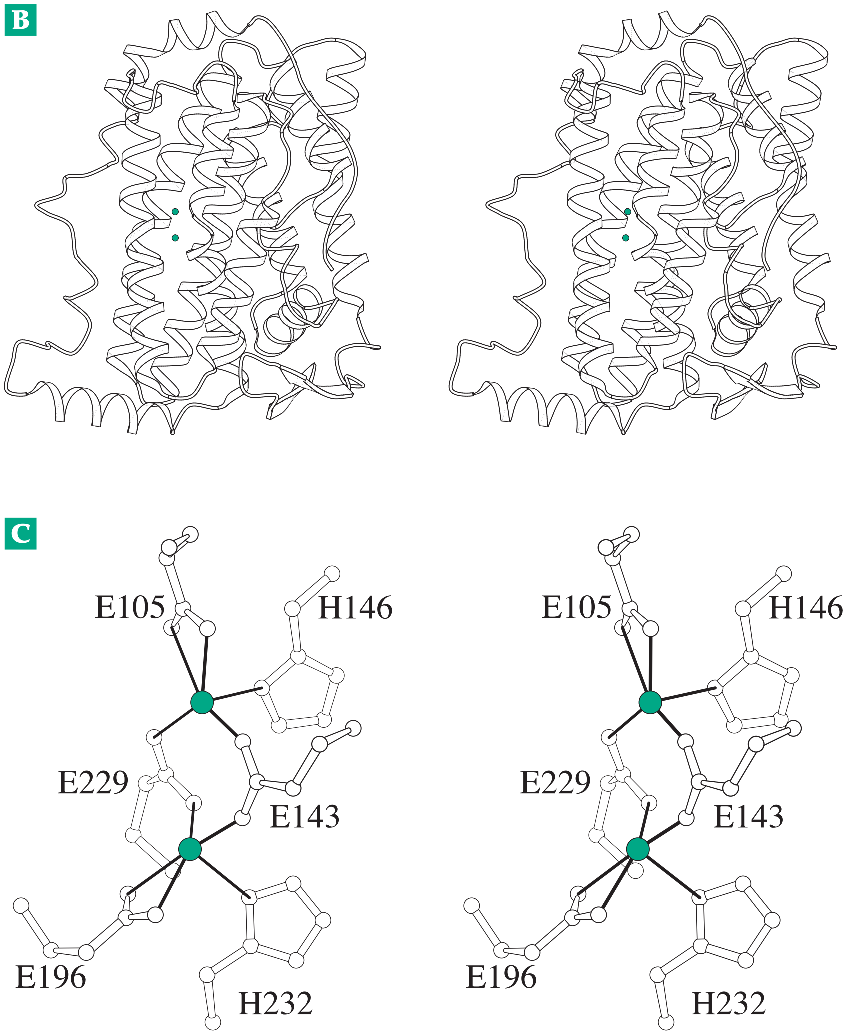


Figure 16-20 (A) The active site of hemerythrin showing the two iron atoms (green) and their ligands which include the μ oxo bridge and two bridging carboxylate groups. From Lukat *et al.*[193] The active site is between four parallel helices as shown in Fig. 2-22. (B) Stereoscopic view of the backbone structure of a Δ^9 stearoyl-acyl carrier protein desaturase which also contains a diiron center. Notice the nine-antiparallel-helix bundle. The diiron center is between four helices as in hemerythrin. (C) Stereoscopic view of the diiron center of the desaturase.[333] (B) and (C) courtesy of Ylva Lindqvist.

generates purple inactive forms which lack a detectable EPR signal.[337,338] The Fe^{3+} of the active enzyme can be replaced[339] with Ga^{3+} and the Fe^{2+} with Zn^{2+} with retention of activity; also, some plants contain phosphatases with Fe–Zn centers.[340,341] The catalytic mechanism resembles those of other metallophosphatases (Chapter 12) and the change of oxidation state of the Fe may play a regulatory role. On the other hand, a principal function of uteroferrin may be in transplacental transport of iron to the fetus.[342]

Diiron oxygenases and desaturases. In the (μ-oxo) diiron oxygenases O_2 is initially bound in a manner similar to that in hemerythrin but one atom of the bound O_2 is reduced to H_2O using electrons supplied by a cosubstrate such as NADPH. The other oxygen atom enters the substrate. This is illustrated by Eq. 16-21 for methane monooxygenase.[332,343–345] A toluene monooxygenase has similar properties.[346]

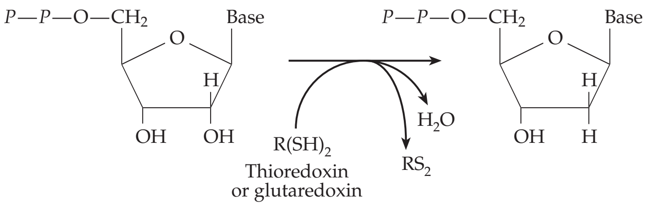
$$CH_4 + NADH + H^+ + O_2 \rightarrow CH_3OH + NAD^+ + H_2O \qquad (16\text{-}21)$$

In green plants a soluble Δ^9 stearoyl-acyl carrier protein desaturase uses O_2 and NADH or NADPH to introduce a double bond into fatty acids. The structure of this protein (Fig. 16-20B,C) is related to those of methane oxygenase and ribonucleotide reductase.[333,347] The desaturase mechanism is discussed in Chapter 21.

9. Ribonucleotide Reductases

Ribonucleotides are reduced to the 2'-deoxyribonucleotides (Eq. 16-21) that are needed for DNA synthesis by enzymes that act on either the di- or triphosphates of the purine and pyrimidine nucleosides[348–351] (Chapter 25). These ribonucleotide reductases utilize either thioredoxin or glutaredoxin (Box 15-C) as the immediate hydrogen donors (Eq. 16-22). The pair of closely spaced –SH groups in the reduced thioredoxin or glutaredoxin are converted into a disulfide bridge at the same time that the 2'-OH of the ribonucleotide di- (or tri-) phosphate is converted to H_2O. While some organisms employ a vitamin B_{12}-

dependent enzyme for this purpose, most utilize iron–tyrosinate enzymes (Class I ribonucleotide reductases). These are two-protein complexes of composition $\alpha_2\beta_2$. The enzyme from *E. coli* contains 761-residue α chains and 375-residue β chains. That from *Salmonella typhimurium*[351a] is similar, as are corresponding mammalian enzymes and a virus-encoded ribonucleotide reductase formed in *E. coli* following infection by T4 bacteriophage.[352] In every case the larger α_2 dimer, which is usually called the **R1 protein**, contains the substrate binding sites, allosteric effector sites, and redox-active SH groups. Each α chain is folded into an unusual $(\alpha / \beta)_{10}$ barrel.[353,354] As in the more familiar $(\alpha / \beta)_8$ barrels, the active site is at the N termini of the β strands.

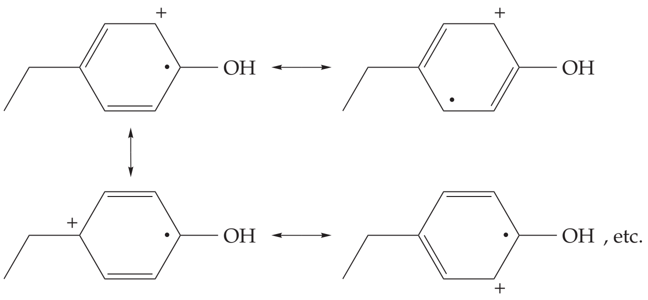


(16-22)

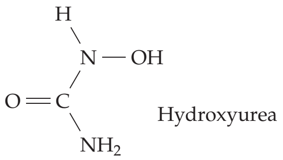
Each polypeptide chain of the β_2 dimer or **R2 protein** contains a diiron center which serves as a free radical generator.[354a,b,c] A few bacteria utilize a dimanganese center.[355] Oxygenation of this center is linked to the uptake of both a proton and an electron and to the removal of a hydrogen atom from the ring of tyrosine 166 to form H_2O and an organic radical (Eq. 16-23):[356–360]

$$\begin{aligned}&2\ \mathrm{Fe(II)}^{2+} + \mathrm{O_2} + \mathrm{H^+} + \mathrm{e^-} + \mathrm{Y122\text{-}OH} \ \rightarrow \\ &\quad \mathrm{Fe(III)}^{3+} - \mathrm{O}^{2-} - \mathrm{Fe(III)}^{3+} - \mathrm{Y122} - \mathrm{O}\bullet + \mathrm{H_2O}\end{aligned} \qquad (16\text{-}23)$$

The tyrosyl radical is used to initiate the ribonucleotide reduction at the active site in the R1 protein ~3.5 nm away. The tyrosyl radical is very stable and was discovered by a characteristic EPR spectrum of isolated enzyme. Alteration of this spectrum when bacteria were grown in deuterated tyrosine indicated that the radical is located on a tyrosyl side chain and that the spin density is delocalized over the tyrosyl ring.[361,362] Using protein engineering techniques the ring was located as Tyr 122 of the *E. coli* enzyme. A few of the resonance structures that can be used to depict the radical are the following:



A chain of hydrogen-bonded side chains apparently provides a pathway for transfer of an unpaired electron from the active site to the Tyr 122 radical and from there to the radical generating center.[363] The tyrosyl radical can be destroyed by removal of the iron by exposure to O_2 or by treatment of ribonucleotide reductases with hydroxyurea, which reduces the radical and also destroys catalytic activity:



A second group of ribonucleotide reductases (Class II), found in many bacteria, depend upon the cobalt-containing **vitamin B_{12} coenzyme** which is discussed in Section B. These enzymes are monomeric or homodimeric proteins of about the size of the larger α subunits of the Class I enzymes. The radical generating center is the 5'-deoxyadenosyl coenzyme.[350,364,365]

Class III or **anaerobic nucleotide reductases** are used by various anaerobic bacteria including *E. coli* when grown anaerobically[350,366–369a] and also by some bacteriophages.[370] Like the Class I reductases, they have an $\alpha_2\beta_2$ structure but each β subunit contains an Fe_4S_4 cluster which serves as the free radical generator,[369] that forms a stable glycyl radical at G580.[369a] In this respect the enzyme resembles pyruvate formate-lyase (Eq. 15-40). As with other enzymes using Fe_4S_4 clusters as radical generators, *S*-adenosylmethionine is also required. All ribonucleotide reductases may operate by similar radical mechanisms.[350,351]

When a 2'-Cl or -F analog of UDP was used in place of the substrate an irreversible side reaction occurred by which Cl^- or F^-, inorganic pyrophosphate, and uracil were released.[349] When one of these enzyme-activated inhibitors containing ^3H in the 3' position was tested, the tritium was shifted to the 2' position with loss of Cl^- and formation of a reactive 3'-carbonyl compound (Eq. 16-24) that can undergo β elimination at each end to give an unsaturated ketone which inactivates the enzyme. This suggested that the Fe-tyrosyl radical abstracts an electron (through a

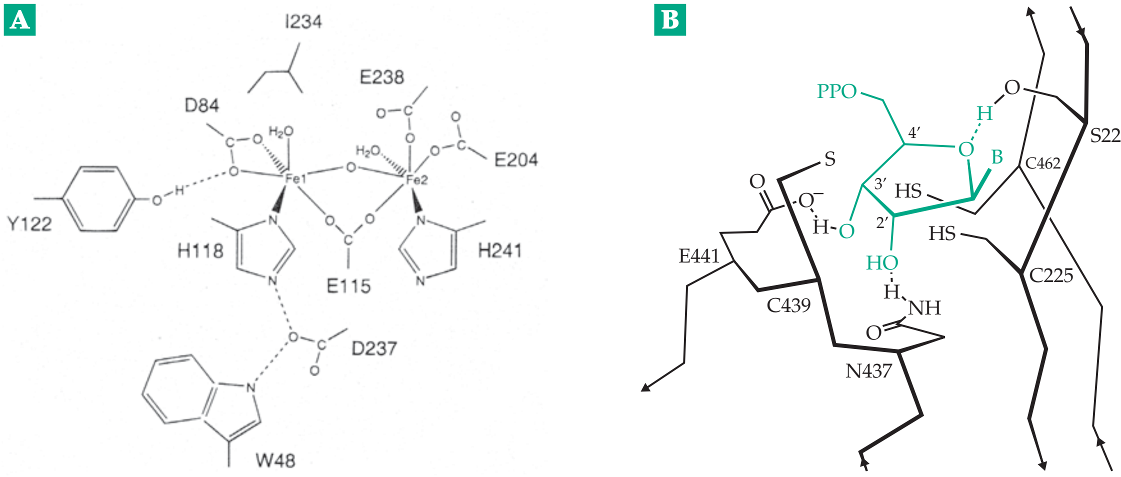
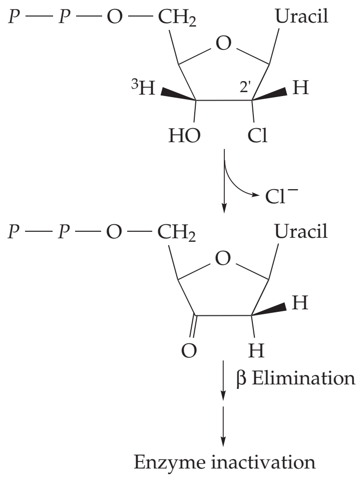
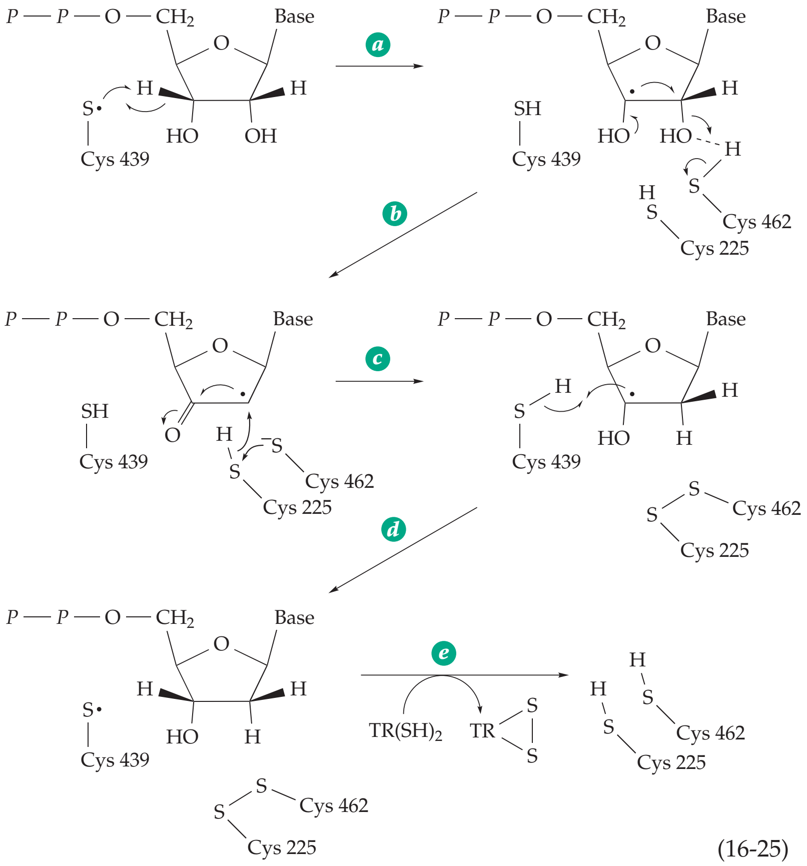

A
I234
E238
D84
H₂O
E204
Fe1
Fe2
Y122
H118
H241
E115
D237
W48
B
PPO
S
O
O⁻
H-O
4′
3′
2′
H
B
HS
C462
S224
HS
E441
C439
HO
H
NH
C225
O
N437

Figure 16-21 (A) Scheme showing the diiron center of the R2 subunit of *E. coli* ribonucleotide reductase. Included are the side chains of tyrosine 122, which loses an electron to form a radical, and of histidine 118, aspartate 237, and tryptophan 48. These side chains provide a pathway for radical transfer to the R1 subunit where the chain continues to tyrosines 738 and 737 and cysteine 429.[354a–c] From Andersson *et al.*[354c] (B) Schematic drawing of the active site region of the *E. coli* class III ribonucleotide reductase with a plausible position for a model-built substrate molecule. Redrawn from Lenz and Giese[373] with permission.

chain of intermediate groups) from an –SH group, now identified as C439 in the *E. coli* enzyme[371–372] (Fig. 16-21). The resulting thiyl radical is thought to abstract a hydrogen atom from C′-3 of a true substrate to form a substrate radical (Eq. 16-25, step *a*) which, with the help of the C462 –SH group, facilitates the loss of $^-$OH from C-2 in step *b*. The resulting C2 radical would be reduced by the nearby redox-active thiol pair C462 and C225 (Fig. 16-21 and Eq. 16-25, step *c*). In Eq. 16-25 the reaction is shown as a hydride transfer with an associated one-electron shift but the mechanism is uncertain. In step *d* of Eq. 16-25 the thiyl radical is regenerated and continues to function in subsequent rounds of catalysis. In the final step (step *e*) the redox active pair is reduced by reduced thioredoxin or glutaredoxin. The active site must open to release the product and to permit this reduction, which may involve participation of still other –SH groups in the protein.


P — P — O — CH₂
Uracil
O
³H
2′
H
HO
Cl
Cl⁻
P — P — O — CH₂
Uracil
O
H
O
H
β Elimination
Enzyme inactivation

(16-24)


P — P — O — CH₂
Base
O
S•
H
H
HO
OH
Cys 439
a
SH
Cys 439
HO
HO
H
S
Cys 462
H
S
Cys 225
b
c
O
H
⁻S
S
Cys 462
Cys 225
d
e
TR(SH)₂
TR
S
S
H
S
H
S
Cys 462
Cys 225

(16-25)

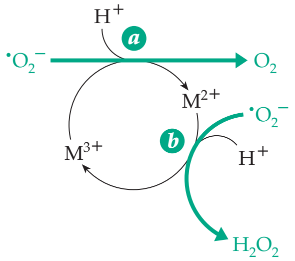
10. Superoxide Dismutases

Metalloenzymes of at least three different types catalyze the destruction of superoxide radicals that arise from reactions of oxygen with heme proteins, reduced flavoproteins, and other metalloenzymes. These superoxide dismutases (SODs) convert superoxide anion radicals $^\bullet O_2^-$ into H_2O_2 and O_2 (Eq. 16-26). The H_2O_2 can then be destroyed by catalase (Eq. 16-8).

$$2\,O_2^- + 2\,H^+ \longrightarrow H_2O_2 + O_2 \qquad (16\text{-}26)$$

The much studied Cu / Zn superoxide dismutase of eukaryotic cytoplasm is described in Section D. However, eukaryotic mitochondria contain manganese SOD and some eukaryotes also synthesize an iron-containing SOD. For example, the protozoan *Leishmania tropica*, which takes up residence in the phagolysosomes of a victim's macrophages, synthesizes an iron-containing SOD[374] to protect itself against superoxide generated by the macrophages. *Mycobacterium tuberculosis* secretes an iron SOD which assists its survival in living tissues and is also a target for the immune response of human hosts.[375]

Iron and manganese SODs have ~20-kDa subunits in each of which a single ion of Fe or Mn is bound by three imidazole groups and a carboxylate group.[376,377] The metal ion undergoes a cyclic change in oxidation state as illustrated by Eq. 16-27. Notice that two protons must be taken up for formation of H_2O_2. In Cu / Zn SOD the copper cycles between Cu^{2+} and Cu^+. The structure of the active site of an Fe SOD is shown in Fig. 16-22. That of Mn SOD is almost identical.[376] In addition to the histidine and carboxylate ligands, the metal binds a hydroxyl ion ^-OH or H_2O and has a site open for binding of $^\bullet O_2^-$. As indicated in Eq. 16-27, uptake of one proton is associated with each reaction step. As illustrated in Fig 16-22 for step *a* of Eq. 16-27, the first proton may be taken up to convert the bound ^-OH to H_2O. The enzyme in the Fe^{2+} form has a pK_a of 8.5 that has been associated with tyrosine 34. Perhaps this residue is involved in the proton uptake process.[378] A similarly located Tyr 41 from *Sulfolobus* is covalently modified, perhaps by phosphorylation.[378a]



(16-27)

B. Cobalt and Vitamin B_{12}

The human body contains only about 1.5 mg of cobalt, almost all of it is in the form of **cobalamin**, vitamin B_{12}. Ruminant animals, such as cattle and sheep, have a relatively high nutritional need for cobalt and in regions with a low soil cobalt content, such as Australia, cobalt deficiency in these animals is a serious problem. This need for cobalt largely reflects the high requirement of the microorganisms of the rumen (paunch) for vitamin B_{12}. All bacteria require vitamin B_{12} but not all are able to synthesize it. For example, *E. coli* lacks one enzyme in the biosynthetic

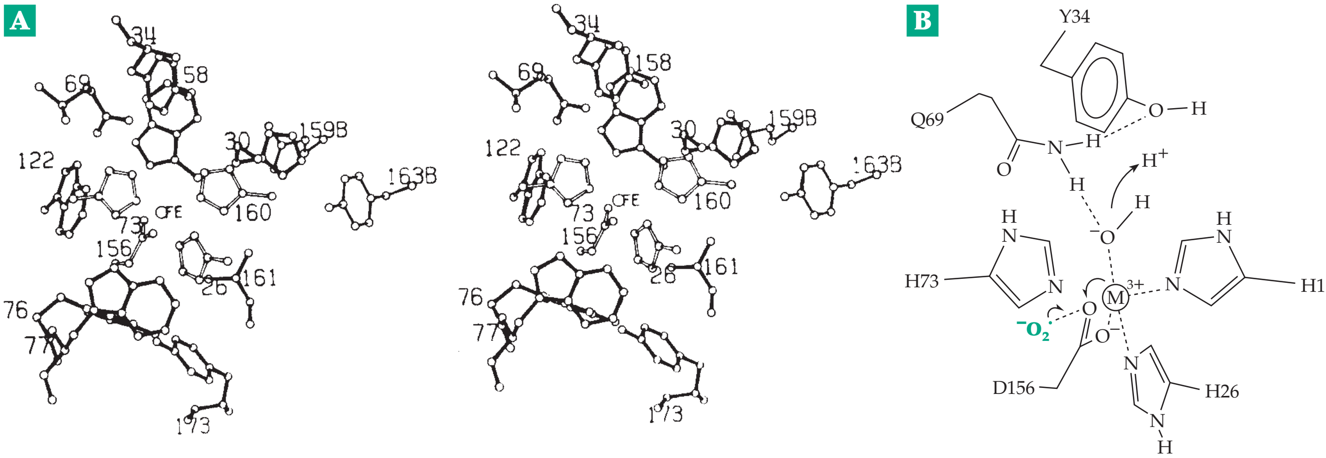
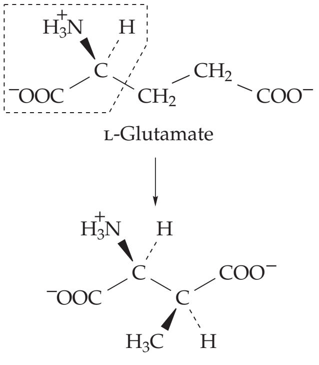


Figure 16-22 (A) Structure of the active site of iron superoxide dismutase from *E. coli*. From Carlioz *et al.*[379] Courtesy of M. Ludwig. (B) Interpretive drawing illustrating the single-electron transfer from a superoxide molecule to the Fe^{3+} of superoxide dismutase and associated proton uptake. Based on Lah *et al.*[376]

pathway and must depend upon other bacteria to complete the synthesis.

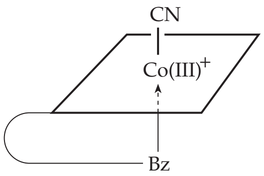
1. Coenzyme Forms

For several years after the discovery of cobalamin its biochemical function remained a mystery, a major reason being the extreme sensitivity of the coenzymes to decomposition by light. Progress came after Barker and associates discovered that the initial step in the anaerobic fermentation of glutamate by *Clostridium tetanomorphum* is rearrangement to β-methylaspartate[380,381] (Eq. 16-28).

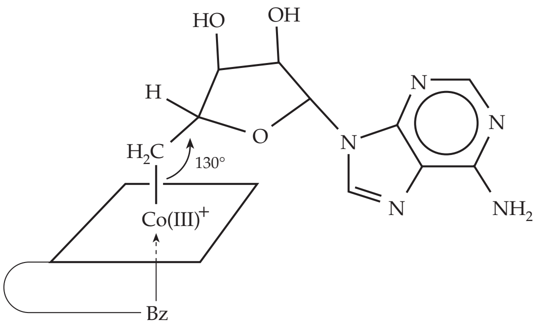


(16-28)

The latter compound can be catabolized by reactions that cannot be used on glutamate itself. Thus, the initial rearrangement is an indispensable step in the energy metabolism of the bacterium. A new coenzyme required for this reaction was isolated in 1958 after it was found that protection from light during the preparation was necessary. The coenzyme was characterized in 1961 by X-ray diffraction[382] as **5'-deoxyadenosylcobalamin**. It is related to cyanocobalamin (Box 16-B) by replacement of the CN group by a 5'-deoxyadenosyl group as indicated in the following abbreviated formulas.[383–385] Here the planes represent the corrin ring system and Bz the dimethylbenzimidazole that is coordinated with the cobalt from below the ring.



Cyanocobalamin

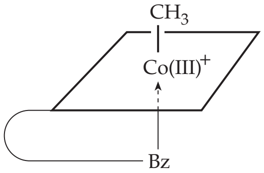


5′-Deoxyadenosylcobalamin

The most surprising structural feature is the Co–C single bond of length 0.205 nm. Thus, the coenzyme is an alkyl cobalt, the first such compound found in nature. In fact, alkyl cobalts were previously thought to be unstable. Vitamin B_{12} contains Co(III), and cyanocobalamin can be imagined as arising by replacement of the single hydrogen on the inside of the corrin ring by Co^{3+} plus CN^-. However, bear in mind that three other nitrogens of the corrin ring and a nitrogen of dimethylbenzimidazole also bind to the cobalt. Each nitrogen atom donates an electron pair to form coordinate covalent linkages. Because of resonance in the conjugated double-bond system of the corrin, all four of the Co–N bonds in the ring are nearly equivalent and the positive charge is distributed over the nitrogen atoms surrounding the cobalt.

The strength of the axial Co–C bond is directly influenced by the strength of bonding of the dimethylimidazole whose conjugate base has a microscopic pK_a of 5.5.[386] Protonation of this base breaks its bond to cobalt and may thereby strengthen the Co–C bond.[387] Steric factors are also important in determining the strength of this bond. NMR techniques are now playing an important role in investigation of these factors.[386–388]

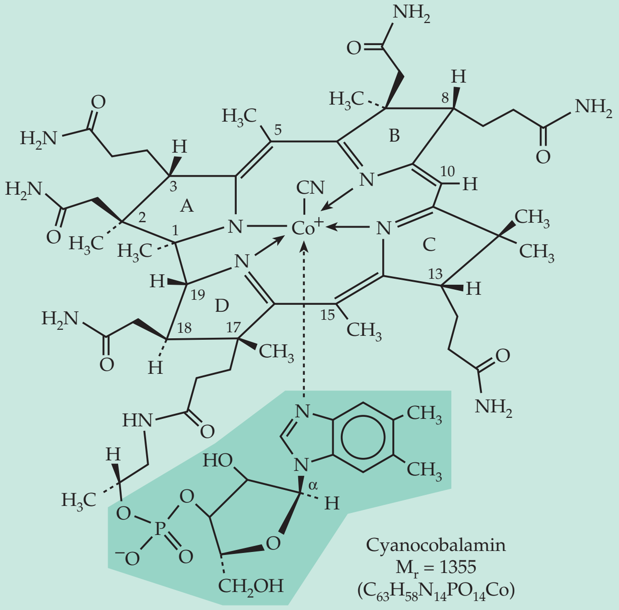
In both bacteria and liver, the 5'-deoxyadenosyl coenzyme is the most abundant form of vitamin B_{12}, while lesser amounts of **methylcobalamin** are present. Other naturally occurring analogs of the coenzymes include **pseudo vitamin B_{12}** which contains adenine in place of the dimethylbenzimidazole.



Methylcobalamin

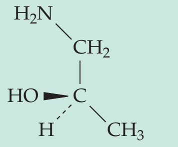
Like dimethylbenzimidazole, it is combined with ribose in the unusual α linkage. A compound called factor A is the vitamin B_{12} analog with 2-methyladenine. Related compounds have been isolated from such sources as sewage sludge which abounds in anaerobic bacteria. It has been suggested that plants may contain vitamin B_{12}-like materials which do not support growth of bacteria. Thus, we may not have discovered all of the alkyl cobalt coenzymes.

BOX 16-B COBALAMIN (VITAMIN B_{12})

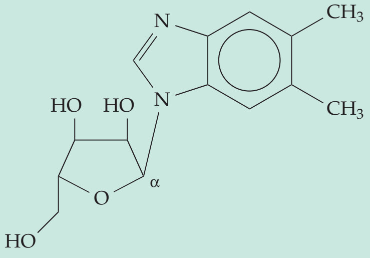


Cyanocobalamin
$M_r = 1355$
($C_{63}H_{58}N_{14}PO_{14}Co$)

The story of vitamin B_{12} began with pernicious anemia, a disease that usually affects only persons of age 60 or more but which occasionally strikes children.[a] Before 1926 the disease was incurable and usually fatal. Abnormally large, immature, and fragile red blood cells are produced but the total number of erythrocytes is much reduced from $4–6 \times 10^6$ mm^{-3} to $1–3 \times 10^6$ mm^{-3}. Within the bone marrow mitosis appears to be blocked and DNA synthesis is suppressed. The disease also affects other rapidly growing tissues such as the gastric mucous membranes (which stop secreting HCl) and nervous tissues. Demyelination of the central nervous system with loss of muscular coordination (ataxia) and psychotic symptoms is often observed.



R-1-Amino-2-propanol



α-Ribofuranoside of dimethylimidazole

In 1926, Minot and Murphy discovered that pernicious anemia could be controlled by eating one-half pound of raw or lightly cooked liver per day, a treatment which not all patients accepted with enthusiasm. Twenty-two years later vitamin B_{12} was isolated (as the crystalline derivative cyanocobalamin) and was shown to be the curative agent. It is present in liver to the extent of 1 mg kg^{-1} or ~10^{-6} M. Although much effort was expended in preparation of concentrated liver extracts for the treatment of pernicious anemia, the lack of an assay other than treatment of human patients made progress slow.

In the early 1940s nutritional studies of young animals raised on diets lacking animal proteins and maintained out of contact with their own excreta (which contained vitamin B_{12}) demonstrated the need for "animal protein factor" which was soon shown to be the same as vitamin B_{12}. The animal feeding experiments also demonstrated that waste liquors from streptomyces fermentations used in production of antibiotics were extremely rich in vitamin B_{12}. Later this vitamin was recognized as a growth factor for a strain of *Lactobacillus lactis* which responded with half-maximum growth to as little as 0.013 μg / l (10^{-11} M).

In 1948, red cobalt-containing crystals of vitamin B_{12} were obtained almost simultaneously by two pharmaceutical firms. Charcoal adsorption from liver extracts was followed by elution with alcohol and numerous other separation steps. Later fermentation broths provided a richer source. Chemical studies revealed that the new vitamin had an enormous molecular weight, that it contained one atom of phosphorus which could be released as P_i, a molecule of aminopropanol, and a ribofuranoside of dimethyl benzimidazole with the unusual α configuration.

Note the relationship of the dimethylbenzimidazole to the ring system of riboflavin (Box 15-B). Several molecules of ammonia could be released from amide linkages by hydrolysis, but all attempts to remove the cobalt reversibly from the ring system were unsuccessful. The structure was determined in 1956 by Dorothy C. Hodgkin and coworkers using X-ray diffraction.[b] At that time, it was the largest organic structure determined by X-ray diffraction. The complete laboratory synthesis was accomplished in 1972.[c]

The ring system of vitamin B_{12}, like that of porphyrins (Fig. 16-5), is made up of four pyrrole rings whose biosynthetic relationship to the corresponding rings in porphyrins is obvious from the structures. In addition, a number of "extra" methyl groups are present. A less extensive conjugated system of double bonds is present in the **corrin** ring of vitamin B_{12} than in porphyrins, and as a result, many chiral centers are found around the periphery

BOX 16-B (continued)

of the somewhat nonplanar rings.

Cyanocobalamin, the form of vitamin B_{12} isolated initially, contains cyanide attached to cobalt. It occurs only in minor amounts, if at all, in nature but is generated through the addition of cyanide during the isolation. **Hydroxocobalamin** (vitamin B_{12a}) containing OH^- in place of CN^- does occur in nature. However, the predominant forms of the vitamin are the coenzymes in which an alkyl group replaces the CN^- of cyanocobalamin.

Intramuscular injection of as little as 3–6 µg of crystalline vitamin B_{12} is sufficient to bring about a remission of pernicious anemia and 1 µg daily provides a suitable maintenance dose (often administered as hydroxocobalamin injected once every 2 weeks). For a normal person a dietary intake of 2–5 µg / day is adequate. There is rarely any difficulty in meeting this requirement from ordinary diets. Vitamin B_{12} has the distinction of being synthesized only by bacteria, and plants apparently contain none. Consequently, strict vegetarians sometimes have symptoms of vitamin B_{12} deficiency.

Pernicious anemia is usually caused by poor absorption of the vitamin. Absorption depends upon the **intrinsic factor**, a mucoprotein (or mucoproteins) synthesized by the stomach lining.[a,d–f] Pernicious anemia patients often have a genetic predilection toward decreased synthesis of the intrinsic factor. Gastrectomy, which decreases synthesis of the intrinsic factor, or infection with fish tapeworms, which compete for available vitamin B_{12} and interfere with absorption, can also induce the disease. Also essential are a plasma membrane receptor[g,h] and two blood transport proteins **transcobalamin**[d,i] and **cobalophilin**. The latter is a glycoprotein found in virtually every human biological fluid and which may protect the vitamin from photodegradation by light that penetrates tissues.[j] A variety of genetic defects involving uptake, transport, and conversion to vitamin B_{12} coenzyme forms are known.[f,k]

Normal blood levels of vitamin B_{12} are ~2 x 10^{-10} M or a little more, but in vegetarians the level may drop to less than one-half this value. A deficiency of folic acid can also cause megaloblastic anemia, and a large excess of folic acid can, to some extent, reverse the anemia of pernicious anemia and mask the disease.

---

[a] Karlson, P. (1979) *Trends Biochem. Sci.* **4**, 286

[b] Hodgkin, D. C. (1965) *Science* **150**, 979–988

[c] Maugh, T. H., II (1973) *Science* **179**, 266–267

[d] Gräsbeck, R., and Kouvonen, I. (1983) *Trends Biochem. Sci.* **8**, 203–205

[e] Allen, R. H., Stabler, S. P., Savage, D. G., and Lindenbaum, J. (1993) *FASEB J.* **7**, 1344–1353

[f] Fenton, W. A., and Rosenberg, L. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3129–3149, McGraw-Hill, New York

[g] Seetharam, S., Ramanujam, K. S., and Seetharam, B. (1992) *J. Biol. Chem.* **267**, 7421–7427

[h] Birn, H., Verroust, P. J., Nexo, E., Hager, H., Jacobsen, C., Christensen, E. I., and Moestrup, S. K. (1997) *J. Biol. Chem.* **272**, 26497–26504

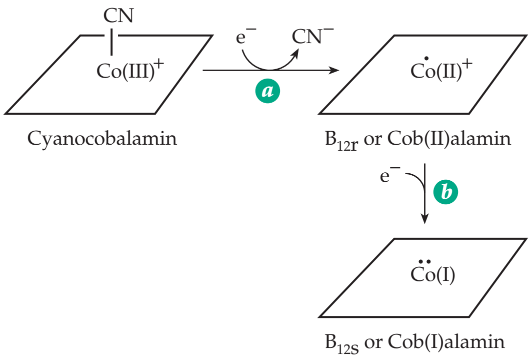
[i] Fedosov, S. N., Berglund, L., Nexo, E., and Petersen, T. E. (1999) *J. Biol. Chem.* **274**, 26015–26020

[j] Frisbie, S. M., and Chance, M. R. (1993) *Biochemistry* **32**, 13886–13892

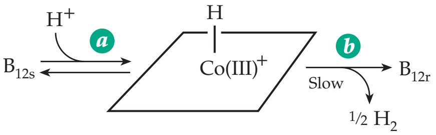
[k] Rosenblatt, D. S., Hosack, A., Matiaszuk, N. V., Cooper, B. A., and Laframboise, R. (1985) *Science* **228**, 1319–1321

2. Reduction of Cyanocobalamin and Synthesis of Alkyl Cobalamins

Cyanocobalamin can be reduced in two one-electron steps (Eq. 16-29).[385,388] The cyanide ion is lost in the first step (Eq. 16-29, step *a*), which may be accomplished with chromous acetate at pH 5 or by catalytic hydrogenation. The product is the brown paramagnetic compound B_{12r}, a tetragonal low-spin cobalt(II) complex. In the second step (Eq. 16-29, step *b*), an additional electron is added, e.g., from sodium borohydride or from chromous acetate at pH 9.5, to give the gray-green exceedingly reactive B_{12s}. The latter is thought to be in equilibrium with cobalt(III) hydride, as shown in Eq. 16-30, step *a*. The hydride is unstable and breaks down slowly to H_2 and B_{12r} (Eq. 16-30, step *b*).[389]

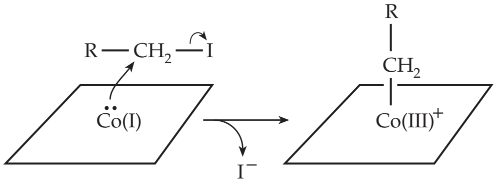


(16-29)



(16-30)

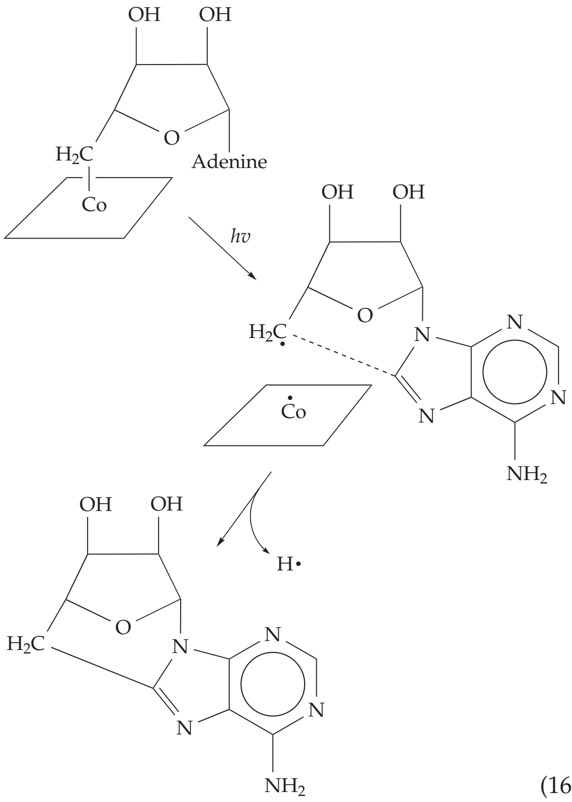
Vitamin B_{12s} reacts rapidly with alkyl iodides (e.g., methyl iodide or a 5'-chloro derivative of adenosine) via nucleophilic displacement to form the alkyl cobalt forms of vitamin B_{12} (Eq. 16-31). These reactions provide a convenient way of preparing isotopically labeled alkyl cobalamins, including those selectively enriched in ^{13}C for use in NMR studies.[390] The biosynthesis of 5'-deoxyadenosylcobalamin utilizes the same type of reaction with ATP as a substrate.[391] A **B_{12s} adenosyltransferase** catalyzes nucleophilic displacement on the 5' carbon of ATP with formation of the coenzyme and displacement of inorganic tripolyphosphate PPP_i.



(16-31)

3. Three Nonenzymatic Cleavage Reactions of Vitamin B_{12} Coenzymes

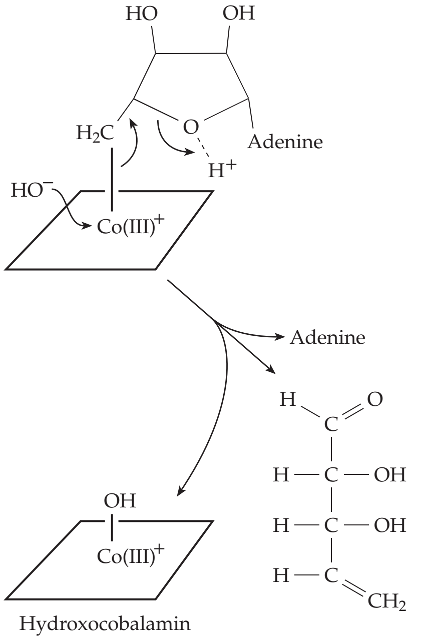
The 5'-deoxyadenosyl coenzyme is easily decomposed by a variety of agents. Anaerobic irradiation with visible light yields principally vitamin B_{12r} and a cyclic 5',8-deoxyadenosine which is probably formed through an intermediate radical[391a] (Eq. 16-32): Irradiation in the presence of air gives a variety of products.



(16-32)

Hydrolysis of deoxyadenosylcobalamin by acid (1 M HCl, 100°, 90 min) yields hydroxycobalamin, adenine, and an unsaturated sugar (Eq. 16-33). The initial reaction step is thought to be protonation of the oxygen of the ribose ring.

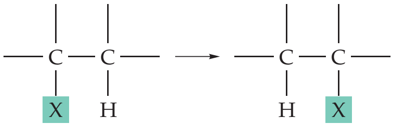
A related cleavage by alkaline cyanide can be viewed as a nucleophilic displacement of the deoxyadenosyl anion by cyanide. The end product is **dicyanocobalamin**, in which the loosely bound nucleotide containing dimethyl benzimidazole is replaced by a second cyanide ion. Methyl and other simple alkyl cobalamins are stable to alkaline cyanide. A number of other cleavage reactions of alkyl cobalamins are known.[392,393]



(16-33)

4. Enzymatic Functions of B_{12} Coenzymes

Three types of enzymatic reactions depend upon alkyl corrin coenzymes. The first is the reduction of ribonucleotide triphosphates by cobalamin-dependent ribonucleotide reductase, a process involving *intermolecular* hydrogen transfer (Eq. 16-21). The second type of reaction encompasses the series of isomerizations shown in Table 16-1. These can all be depicted as in Eq. 16-34. Some group X, which may be attached by a C–C, C–O, or C–N bond, is transferred to an adjacent carbon atom bearing a hydrogen. At the same time,



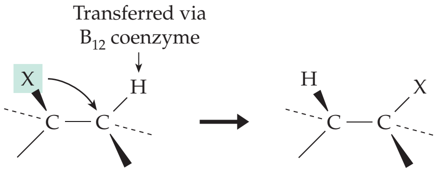
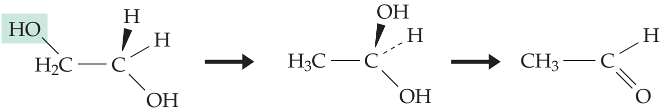
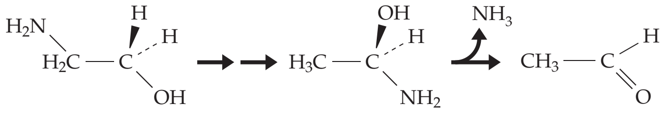
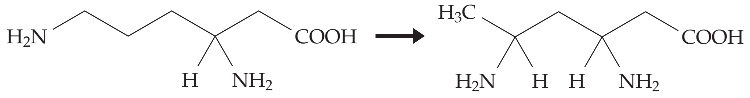
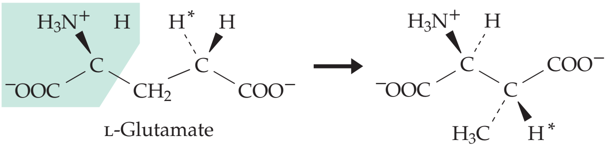
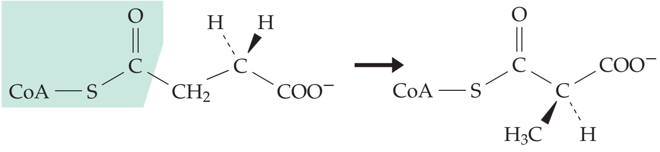
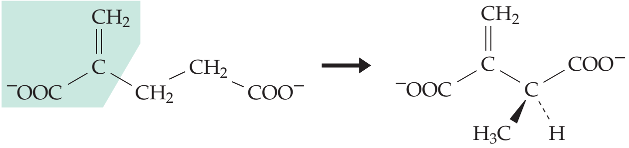
(16-34)

the hydrogen is transferred to the carbon to which X was originally attached. The third type of reaction is the transfer of methyl groups via methylcobalamin and some related bacterial metabolic reactions.

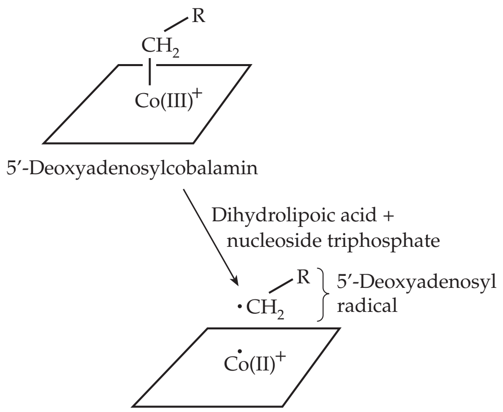
Cobalamin-dependent ribonucleotide reductase. Lactic acid bacteria such as *Lactobacillus leichmanni* and many other bacteria utilize a 5'-deoxyadenosylcobalamin-containing enzyme to reduce nucleoside triphosphates according to Eq. 16-21. Thioredoxin or dihydrolipoic acid can serve as the hydrogen donor. Early experiments showed that protons from water are reversibly incorporated at C-2' of the reduced nucleotide with retention of configuration. A more important finding was a large kinetic isotope effect of 1.8 when 3'-^{3}H-containing UTP was reduced by the enzyme.[394]

Reaction of the reductase with dihydrolipoic acid in the presence of deoxy-GTP, which apparently serves as an allosteric activator, leads to formation, within a few milliseconds, of a radical with a characteristic EPR spectrum that can be studied when the reaction mixture is rapidly cooled to 130°K. When GTP (a true substrate) is used instead of dGTP, the radical signal reaches a maximum in about 20 ms and then decays. Of the various oxidation states of cobalt (3+, 2+, and 1+)

TABLE 16-1
Isomerization Reactions Involving Hydrogen Transfer and Dependent upon a Vitamin B_{12} Coenzyme

	General reaction Migrating group is enclosed in a box
	Dioldehydratase Glycerol dehydratase catalyzes the same type of reaction
	Ethanolamine ammonia-lyase
	L-β-Lysine mutase D-α-Lysine mutase and ornithine mutase catalyze the same type of reaction
 L-Glutamate → *threo*-L-β-Methylaspartate	Glutamate mutase
 (R) Methylmalonyl-CoA	Methylmalonyl-CoA mutase Isobutyryl-CoA mutase catalyzes the same type of reaction
	α-Methyleneglutarate mutase

only the 2+ state of vitamin B$_{12r}$ is paramagnetic and gives rise to an EPR signal. The electronic absorption spectrum of the coenzyme of ribonucleotide reductase is also changed rapidly by substrate in a way that suggests formation of B$_{12r}$. Thus, it was proposed that a *homolytic* cleavage occurs to form B$_{12r}$ and a stabilized 5'-deoxyadenosyl radical (Eq. 16-35).[395–396a] However, H^3 is not transferred from the 3' position of the substrate into the deoxyadenosyl part of the coenzyme.[394] The enzyme has many properties in common with the previously discussed iron–tyrosinate ribonucleotide reductases (Fig. 16-21) including limited peptide sequence homology.[350] Stubbe and coworkers suggested that the deoxyadenosyl radical is formed as a radical chain initiator[394,396b] and that the mechanism of ribonucleotide reduction is as shown in Eq. 16-25. Studies of enzyme-activated inhibitors support this mechanism.[364]



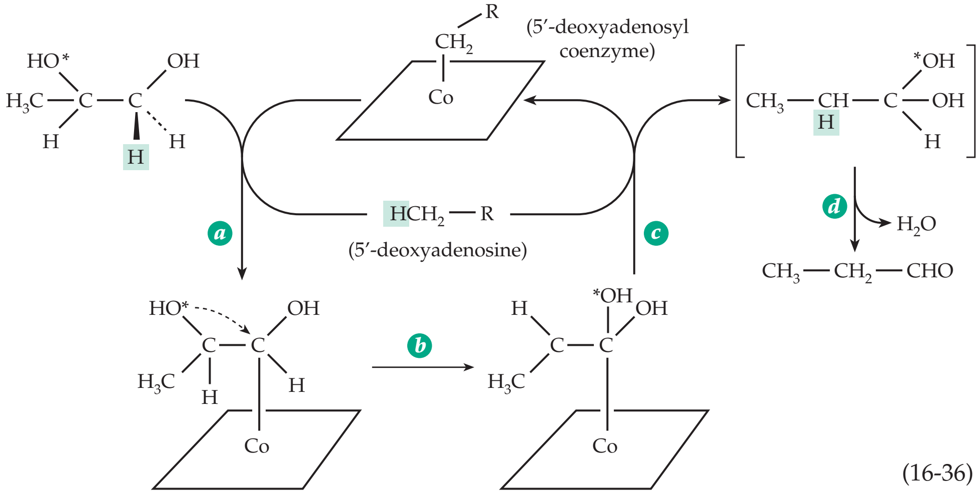
(16-35)

The isomerization reactions. At least 10 reactions of the type described by Eq. 16-34 are known[397] (Table 16-1). They can be subdivided into three groups. First, X = OH or NH$_2$ in Eq. 16-34; isomerization gives a *geminal*-diol or aminoalcohol that can eliminate H$_2$O or NH$_3$ to give an aldehyde. All of these enzymes, which are called **hydro-lyases** or **ammonia-lyases**, specifically require K$^+$ as well as the vitamin B$_{12}$ coenzyme. Second, X = NH$_2$ in Eq. 16-34; For this group of **aminomutases** PLP is required as a second coenzyme. Third, X is attached via a carbon atom; the enzymes are called **mutases**. **Methylmalonyl-Co mutase** is required for catabolism of propionate in the human body, and is one of only two known vitamin B$_{12}$-dependent enzymes. The related isobutyryl-CoA mutase participates in the microbial synthesis of such polyether antibiotics as monensin A.[397a,b] Other mutases are involved in anaerobic bacterial fermentations.

In these reactions the hydrogen is always transferred via the B$_{12}$ coenzyme. No exchange with the medium takes place. Since X may be an electronegative group such as OH, the reactions could all be treated formally as hydride ion transfers but it is more likely that they occur via homolytic cleavages. Such cleavage is indicated by the observation of EPR signals for several of the enzymes in the presence of their substrates.[398]

Abeles and associates showed that when **dioldehydratase** (Table 16-1) catalyzes the conversion of 1,2-[1-^{3}H]propanediol to propionaldehyde, tritium appears in the coenzyme as well as in the final product. When ^{3}H-containing coenzyme is incubated with unlabeled propanediol, the product also contains ^{3}H, which was shown by chemical degradation to be exclusively on C-5'. Synthetic 5'-deoxyadenosyl coenzyme containing ^{3}H in the 5' position transferred ^{3}H to product. Most important, using a mixture of propanediol and ethylene glycol, a small amount of *intermolecular* transfer was demonstrated; that is, ^{3}H was transferred into acetaldehyde, the product of dehydration of ethylene glycol. Similar results were also obtained with ethanolamine ammonia-lyase.[399]

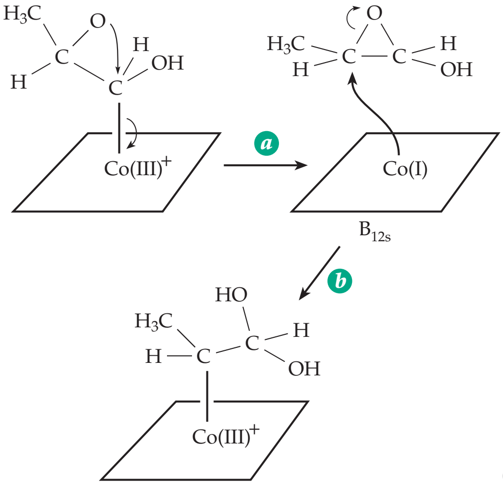
Another important experiment[398] showed that ^{18}O from [2-^{18}O]propanediol was transferred into the 1 position without exchange with solvent. Furthermore, ^{18}O from (*S*)-[1-^{18}O]propanediol was retained in the product while that from the (*R*) isomer was not. Thus, it appears that the enzyme stereospecifically dehydrates the final intermediate. From these and other experiments, it was concluded that initially a 5'-deoxyadenosyl radical is formed via Eq. 16-35. This radical then abstracts the hydrogen atom marked by a shaded box in Eq. 16-36 to form a substrate radical and 5'-deoxyadenosine. One proposal, illustrated in Eq. 16-36, is that the substrate radical immediately recombines with



(16-36)

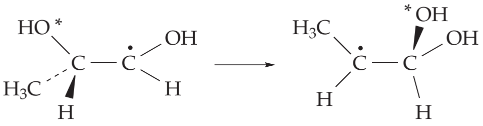
the Co(II) of the coenzyme to form an organo-cobalt substrate compound (step *a* of Eq. 16-36). The 5'-deoxyadenosine now contains hydrogen from the substrate; because of rotation of the methyl group this hydrogen becomes equivalent to the two already present in the coenzyme. The substrate–cobalamin compound formed in this step then undergoes isomerization, which, in the case of dioldehydratase, leads to intramolecular transfer of the OH group (step *b*). In step *c* the hydrogen atom is transferred back from the 5'-deoxyadenosine to its new location in the product and in step *d* the resulting *gem*-diol is dehydrated to form the aldehyde product.[401a]

Carbocation, carboanion, and free radical intermediates have all been proposed for the isomerization step in the reaction. A carbocation would presumably be cyclized to an epoxide which could react with the B_{12s} (Eq. 16-37, step *b*) to complete the isomerization.



(16-37)

At present most evidence favors, for the isomerization reactions, an enzyme-catalyzed rearrangement of the substrate radical produced initially during formation of the 5'-deoxyadenosine (Eq. 16-38).[400–401c]

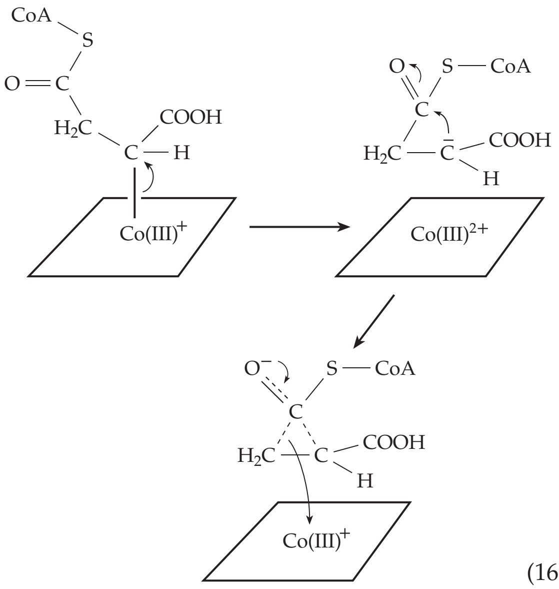


(16-38)

According to this mechanism the Co(II) of the B_{12r} formed in Eq. 16-35 has no active role in the isomerization and does not form an organocobalt intermediate as in Eq. 16-36. Its only role is to be available to recombine with the 5'-deoxyadenosyl radical at the end of the reaction sequence. Support for this interpretation has been obtained from study of model reactions and of organic radicals generated in other ways.[400]

Recently, EPR spectroscopy with ^{2}H- and ^{13}C-labeled glutamates as substrates for glutamate mutase permitted identification of a 4-glutamyl radical as a probable intermediate for that enzyme.[402,402a,b]

On the other hand, for methylmalonyl-CoA mutase Ingraham suggested cleavage of the Co–C bond of an organocobalt intermediate to form a carbanion, the substrate–cobalamin compound serving as a sort of "biological Grignard reagent." The carbanion would be stabilized by the carbonyl group of the thioester forming a "homoenolate anion" (Eq. 16-39). The latter could break up in either of two ways reforming a C–Co bond and causing the isomerization.[403] Some experimental results also favor an ionic or organocobalt pathway.[404]



(16-39)

The three-dimensional structure of methylmalonyl-CoA mutase from *Propionibacterium shermanii* shows that the vitamin B_{12} is bound in a base-off conformation with the dimethylbenzimidazole group bound to the protein far from the corrin ring (Fig. 16-23; see also Fig. 16-24).[405,405a] A histidine of the protein coordinates the cobalt, as also in methionine synthase. The entrance to the deeply buried active site is blocked by the coenzyme part of the substrate. The buried active site may be favorable for free radical rearrangement reactions. The structure of substrate complexes shows that the coenzyme is in the cob(II)alamin (B_{12r}) form with the 5'-deoxyadenosyl group detached from the cobalt, rotated, shifted, and weakly bound to the protein.[405a] Side chains from neighboring Y89, R207, and H244 all hydrogen bond to the substrate. The R207 guanidinium group makes an ion pair with the substrate carboxylate, and the phenolic and imidazole groups may have catalytic functions.[405b,c,d,406] Studies are also in progress on crystalline glutamate mutase.[406,406a]

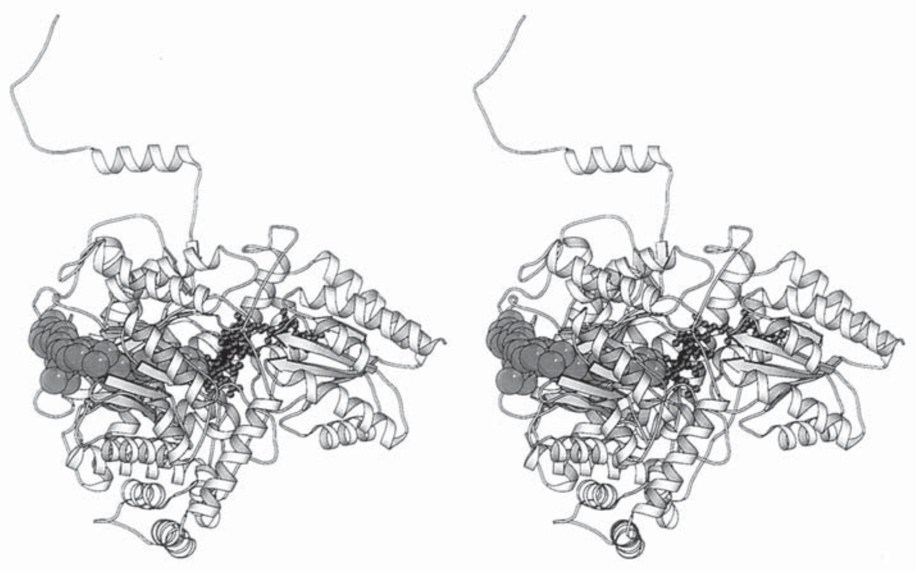
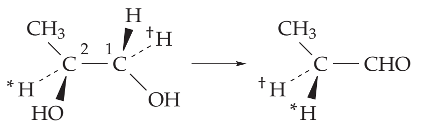


Figure 16-23 Three-dimensional structure of methylmalonyl-CoA mutase from *Propionobacterium shermanii*. The B_{12} coenzyme is deeply buried, as is the active site. A molecule of bound desulfo-coenzyme A, a substrate analog, blocks the active site entrance on the left side. From Mancia *et al*.[405] Courtesy of Philip R. Evans.

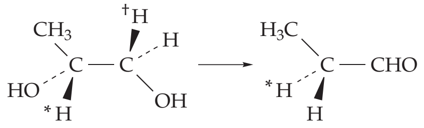
According to all of these mechanisms, 5'-deoxyadenosine is freed from its bond to cobalt during the action of the enzymes. Why then does the deoxyadenosine not escape from the coenzyme entirely, leading to its inactivation? Substrate-induced inactivation is not ordinarily observed with coenzyme B_{12}-dependent reactions, but some quasi-substrates do inactivate their enzymes. Thus, glycolaldehyde converts the coenzyme of dioldehydratase to 5'-deoxyadenosine and ethylene glycol does the same with ethanolamine deaminase. When 5'-deoxyinosine replaces the 5'-deoxyadenosine of the normal coenzyme in dioldehydratase, 5'-deoxyinosine is released quantitatively by the substrate. This suggests that the dehydratase may normally hold the adenine group of 5'-deoxyadenosine through hydrogen bonding to the amino group. Because the OH group of inosine tautomerizes to C=O, inosine may not be held as tightly. The deeply buried active site (Fig. 16-23) may also prevent escape of the deoxyadenosine.

Despite the evidence in its favor, there was initially some reluctance to accept 5'-deoxyadenosine as an intermediate in vitamin B_{12}-dependent isomerization reactions. It was hard to believe that a methyl group could exchange hydrogen atoms so rapidly. It was suggested that protonation of the oxygen of the ribose ring as in Eq. 16-33 might facilitate release of a hydrogen atom. However, substitution of the ring oxygen by CH_2 in a synthetic analog did not destroy the coenzymatic activity.[407] Another possibility is that the methyl group has an unusual reactivity if the cobalt is reduced to Co(II).

Stereochemistry of the isomerization reactions. Dioldehydratase acts on either the (*R*) or (*S*) isomers of 1,2-propanediol (Eqs. 16-40 and 16-41; asterisks and daggers mark positions of labeled atoms in specific experiments).



(*S*)-1,2-Propanediol (16-40)

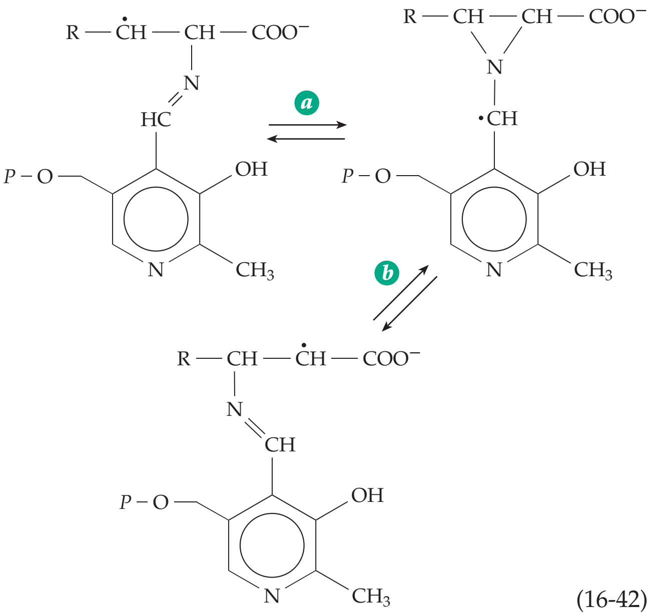


(*R*)-1,2-Propanediol (16-41)

In both cases the reaction proceeds with retention of configuration at C-2 and with stereochemical specificity[408] for one of the two hydrogens at C-1. The reaction catalyzed by methylmalonyl-CoA mutase likewise proceeds with retention of configuration at C-2 (Table 16-1)[409] but the glutamate mutase reaction is accompanied by inversion (Eq. 16-28).

Aminomutases. The enzymes **L-β-lysine mutase** (which is also **D-α-lysine mutase**) and **D-ornithine mutase** catalyze the transfer of an ω-amino group to an adjacent carbon atom[410] (Table 16-1). Two proteins are needed for the reaction; pyridoxal phosphate is required and is apparently directly involved in the amino group migration. In the β-lysine mutase the 6-amino group of L-β-lysine replaces the pro-*S* hydrogen at C-5 but with inversion at C-5 to yield (3*S*, 5*S*)-3,5-diaminohexanoic acid.[411] A bacterial D-lysine 5,6-aminomutase interconverts D-lysine with 2,5-diaminohexanoic acid.[411a] Another related enzyme is **L-leucine 2,3-aminomutase**, which catalyzes the reversible interconversion of L-leucine and β-leucine.[410,412] It was reported to be present in plants and also in the human body, but the latter could not be confirmed.[413]

The interconversion of L-α-lysine and L-β-lysine is catalyzed by a lysine 2,3-aminomutase found in certain clostridia.[414-415a] This enzyme also requires pyridoxal phosphate and catalyzes a reaction with the same stereochemistry as that of β-lysine mutase. However, it does not contain vitamin B_{12} but depends upon *S*-adenosylmethionine (AdoMet) and an iron–sulfur cluster. The adenosyl group of AdoMet may function in the same manner as does the deoxyadenosyl group of the adenosylcobalamin coenzyme. In these mutases a 5'-deoxyadenosyl radical may abstract a hydrogen atom from the β position of a Schiff base of PLP with the amino acid substrate. The radical isomerizes (Eq. 16-42) and then accepts a hydrogen atom back from the 5'-deoxyadenosine to complete the reaction. Its properties suggest that lysine 2,3-aminomutase is related to the pyruvate formate-lyase of *E. coli* (Eq. 15-37),[410] class I ribonucleotide reductases, and other enzymes that act by homolytic mechanisms.



(16-42)

Transfer reactions of methyl groups. The generation and utilization of methyl groups is a quantitatively important aspect of the metabolism of all cells. As we have seen (Fig. 15-18), methyl groups can be created by the reduction of one-carbon compounds attached to tetrahydrofolic acid. Methyl groups of methyltetrahydrofolic acid (N^5-CH_3THF) can then be transferred to the sulfur atom of homocysteine to form methionine (Eq. 16-43). The latter is converted to *S*-adenosylmethionine, the nearly universal methyl group donor for transmethylation reactions (Eq. 12-3). In some bacteria, fungi, and higher plants, the methyl-THF-homocysteine transmethylase does not depend upon vitamin B_{12} but is a metalloenzyme with zinc at the active center.[416] However, human beings share with certain strains of *E. coli* and other bacteria the need for methylcobalamin.

The structure of the *E. coli* enzyme (Fig. 16-24) shows methylcobalamin bound in a base-off conformation, with histidine 759 of the protein replacing dimethylbenzimidazole in the distal coordination position on the cobalt. This histidine is part of a sequence Asp-X-His-X-X-Gly that is found not only in methionine synthase but also in methylmalonyl-CoA mutase, glutamate mutase, and 2-methyleneglutarate mutase. However, diol dehydratase lacks this sequence and binds adenosylcobalamin with the dimethylbenzimidazole–cobalt bond intact.[417]

The coenzyme evidently functions in a cyclic process. The cobalt alternates between the +1 and +3 oxidation states as shown in Eq. 16-43. The first indication of such a cyclic process was the report by Weissbach that ^{14}C-labeled methylcobalamin could be isolated following treatment of the enzyme with such methyl donors as AdoMet and methyl iodide

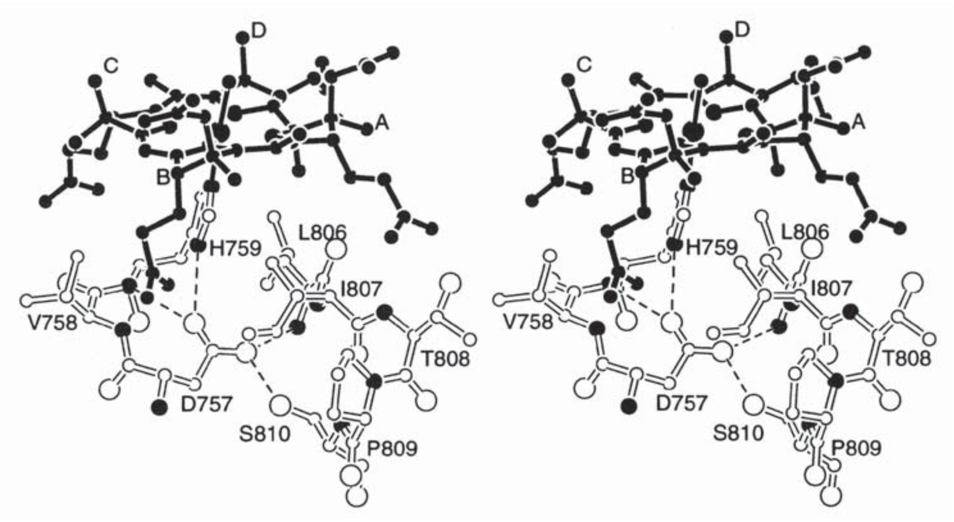
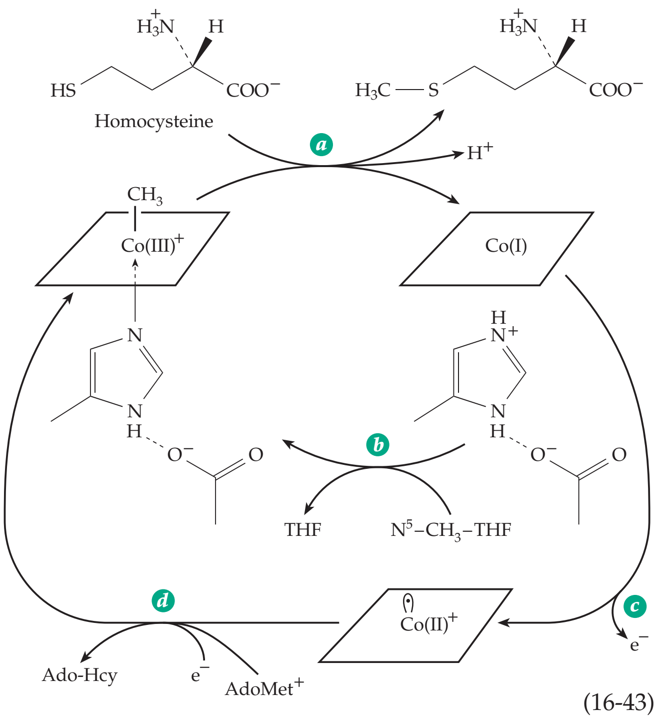


Figure 16-24 Stereoscopic views of the active site of methionine synthase from *E. coli*. Methylcobalamin (black) is in the active site with His 759 of the protein in the distal position of the coenzyme in a base-off conformation. The dimethylbenzimidole nucleotide has been omitted for clarity. Notice the hydrogen-bonded His 759 – D757 – S810 triad. From Jarret *et al*.[418a] Courtesy of R. G. Matthews.

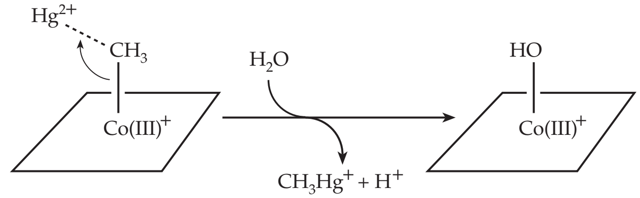


(16-43)

after reduction (e.g., with reduced riboflavin phosphate). The sequence parallels that of Eq. 16-31 for the laboratory synthesis of methylcobalamin. Nevertheless, the transmethylase demonstrates some complexities. Initially, it must be "activated" by AdoMet or methyl iodide, after which it cycles according to Eq. 16-43, steps *a* and *b*, but is gradually inactivated. This apparently happens by oxidation to a Co(II) form of the enzyme (Eq. 16-43, step *c*) that must be reductively methylated with AdoMet and reduced flavodoxin (step *d*) to regenerate the active form.[418] It is also possible that the methyl group is not transferred as a formal CH_3^+ as pictured in Eq. 16-43 but as a $^{\bullet}CH_3$ radical generated by homolytic cleavage of methylcobalamin to cob(II)alamin. Whatever the mechanism, a chiral methyl group is transferred from 5-methyl-THF to homocysteine with overall retention of configuration.[419]

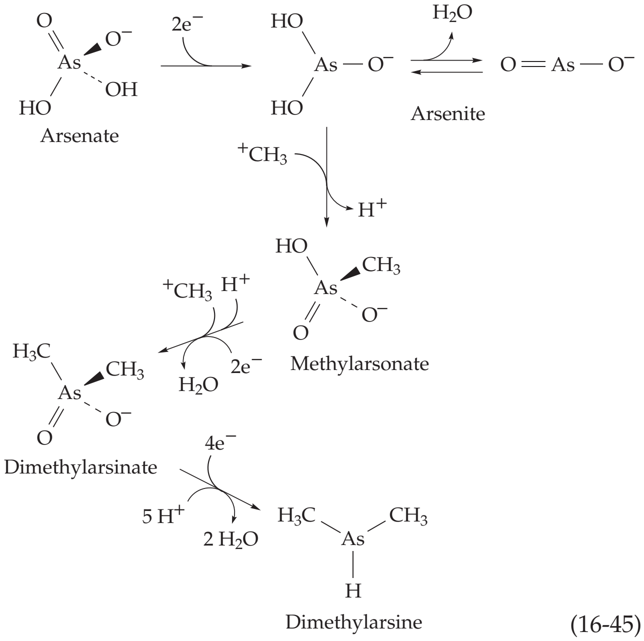
Another important group of methyl transfer reactions are those from methyl corrinoids to mercury, tin, arsenic, selenium, and tellurium. For example, Eq. 16-44 describes the methylation of Hg^{2+}. These reactions are of special interest because of the generation of toxic methyl and dimethyl mercury and dimethylarsine.

Notice that whereas in Eq. 16-43 the methyl group is transferred as CH_3^+ by nucleophilic displacement on a carbon atom, the transfer to Hg^{2+} in Eq. 16-44 is that of a carbanion, CH_3^-, with no valence change occurring in the cobalt. However, it is also possible that transfer occurs as a methyl radical.[420] Methyl corrinoids are able to undergo this type of reaction nonenzymatically, and the ability to transfer a methyl anion is a property of methyl corrinoids not shared by other transmethylating agents such as AdoMet. At the conclusion of the reaction in Eq. 16-44 the cobalt is in the +3 state. To be remethylated, it must presumably be reduced to Co(II). A second methyl group can be transferred by the same type of reaction to form $(CH_3)_2Hg$.



(16-44)

Methylation of arsenic is an important pollution problem because of the widespread use of arsenic compounds in insecticides and because of the presence of arsenate in the phosphate used in household detergents.[421,422] After reduction to arsenite, methylation occurs in two steps (Eq. 16-45). Additional reduction steps result in the formation of **dimethylarsine**, one of the principal products of action of methanogenic bacteria on arsenate. The methyl transfer is shown as occurring through CH_3^+, with an accompanying loss of a proton from the substrate. However, a CH_3 radical may be transferred with formation of a cobalt(II) corrinoid.[423]



(16-45)

Corrinoid-dependent synthesis of acetyl-CoA. The anaerobic bacterium *Clostridium thermoaceticum* obtains its energy for growth by reduction of CO_2 with hydrogen (Eq. 16-46). One of the CO_2 molecules is reduced to formate which is converted via 5-methyl-THF to the methyl corrinoid 5-methoxybenzimidazolylcobamide. The methyl group of the latter

$$4\ H_2 + 2\ CO_2 \longrightarrow CH_3COO^- + H^+ + 2\ H_2O \qquad \Delta G'\ (\text{pH } 7) = -94.9\ \text{kJ mol}^{-1} \tag{16-46}$$

combines with another CO_2 to form acetyl-CoA. The process, which requires the nickel-containing carbon monoxide dehydrogenase, is discussed in Section C.

C. Nickel

It was not until the 1970s that nickel was first recognized as a dietary essential for animals.[424–426] Nickel-deficient chicks grew poorly, had thickened legs, and developed dermatitis. Tissues of deficient animals contained swollen mitochondria and swollen perinuclear space suggesting a function for Ni in membranes. However, doubts have been raised about the conclusions based on these experiments.[427] Within tissues the nickel content ranges from 1 to 5 μg / l. Some of the metal in serum is present as complexes of low molecular mass and some is bound to serum albumin[428] and to a specific nickel-containing protein of the macroglobulin class, known as **nickeloplasmin**.[429] Nickel is also present in plants; in some, e.g., *Allysum*, it accumulates to high concentrations.[430] It is essential for legumes and possibly for all plants.[431] Nickel uptake proteins have been identified in bacteria and fungi.[432,432a] Because of its ubiquitous occurrence it is difficult to prepare a totally Ni-free diet. The acute toxicity of orally ingested Ni(II) is low, and homeostatic mechanisms exist in the animal body for regulating its concentration. However, the volatile nickel carbonyl $Ni(CO)_4$ is very toxic[426] and Ni from jewelry is a common cause of dermatitis.[428,433]

In its compounds nickel usually has the +2 oxidation state but the +3 and +4 states occur rarely in complexes. The Ni^{2+} ion contains eight 3d electrons, a configuration that favors square-planar coordination of four ligands. However, the ion is also able to form a complex with six ligands and an octahedral geometry. It has been suggested that this "ambivalence" may be of biochemical significance.

Nickel is found in at least four enzymes: **urease**, certain **hydrogenases, methyl-CoM reductase** (in its **cofactor F_{430}**) of methanogenic bacteria, and **carbon monoxide dehydrogenase** of acetogenic and methanogenic bacteria.[434]

1. Urease

Urease, which was first isolated from the jack bean has a special place in biochemical history as the first enzyme to be crystallized. This was accomplished by J. B. Sumner in 1926, and although Sumner eventually obtained the Nobel Prize, his first reports were greeted with skepticism and outright disbelief. The presence of two atoms of nickel in each molecule of urease[435] was not discovered until 1975. The metal ions had been overlooked previously, despite the fact that the absorption spectrum of the purified enzyme contains an absorption "tail" extending into the visible region with a shoulder at 425 nm and weak maxima at 725 and 1060 nm. Urease catalyzes the hydrolytic cleavage of urea to two molecules of ammonia and one of bicarbonate and is useful in the analytical determination of urea.

Jack bean urease is a trimer or hexamer of identical 91-kDa subunits while that of the bacterium *Klebsiella* has an $(\alpha\beta_2\gamma_2)_2$ stoichiometry. Nevertheless, the enzymes are homologous and both contain the same binickel catalytic center (Fig. 16-25).[435–437a] The three-dimensional structure of the *Klebsiella* enzyme revealed that the two nickel ions are bridged by a carbamyl group of a carbamylated lysine. Like ribulose bisphosphate carboxylase (Fig. 13-10), urease also requires CO_2 for formation of the active enzyme.[438] Formation of the metallocenter also requires four additional proteins, including a chaperonin and a nickel-binding protein.[438,439]

The mechanism of urease action is probably related to those of metalloproteases such as carboxypeptidase A (Fig. 12-16) and of the zinc-dependent carbonic

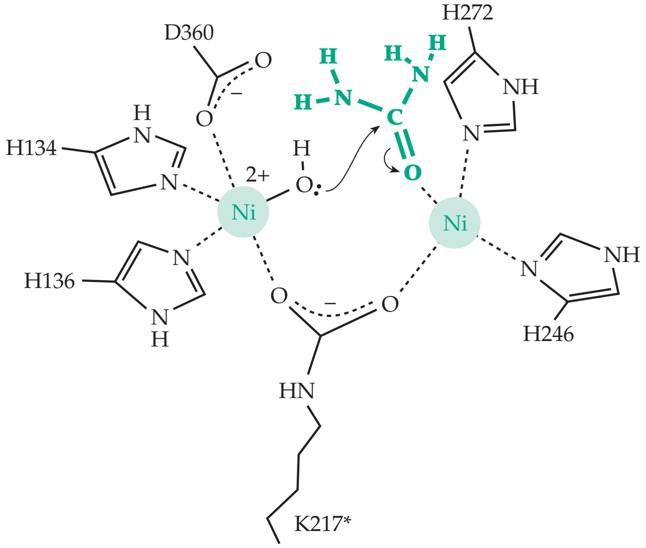
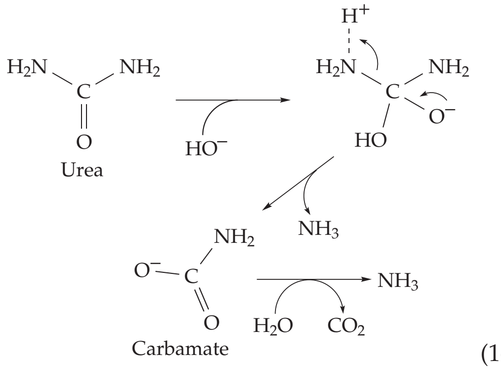


Figure 16-25 The active site of urease showing the two Ni^+ ions held by histidine side chains and bridged by a carbamylated lysine (K217*). A bound urea molecule is shown in green. It has been placed in an open coordination position on one nickel and is shown being attacked for hydrolytic cleavage by a hydroxyl group bound to the other nickel. Based on a structure by Jabri *et al*.[436] and drawing by Lippard.[437]



(16-47)

anhydrase (Fig. 13-1). In Fig. 16-25 one nickel ion is shown polarizing the carbonyl group, while the second provides a bound hydroxyl ion that serves as the attacking nucleophile. A probable intermediate product is a carbamate ion (Eq. 16-47).

Urease is an essential enzyme for bacteria and other organisms that use urea as a primary source of nitrogen. The peptic ulcer bacterium *Helicobacter pylori* uses urease to hydrolyze urea in order to defend against the high acidity of the stomach.[439a] The enzyme is also present in plant leaves and may play a necessary role in nitrogen metabolism.[431] In nitrogen-fixing legumes urea derivatives, the **ureides**, have an important function (Chapter 25) but urease may not be involved in the catabolism of these compounds.[440]

2. Hydrogenases

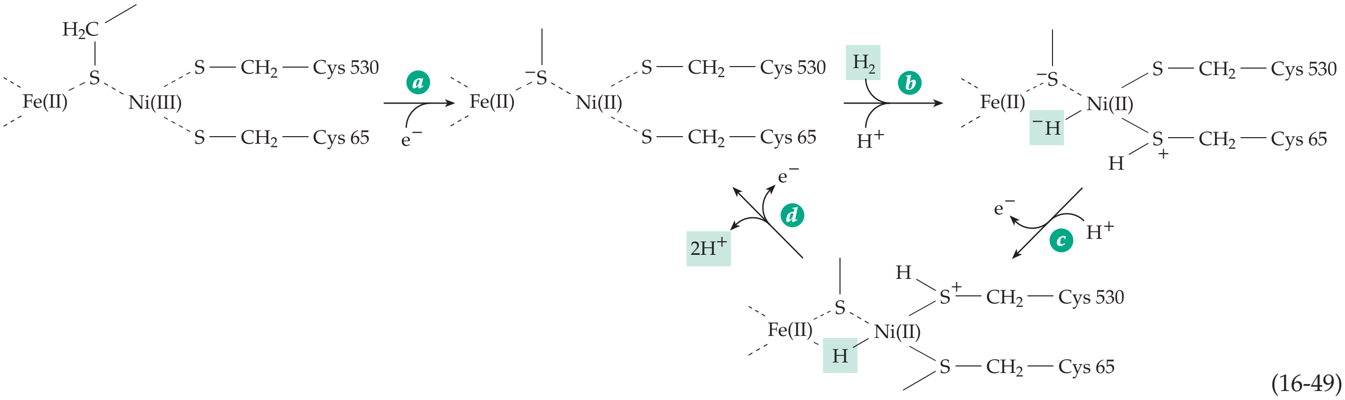
Many plants, animals, and microorganisms are able to evolve H_2 by reduction of hydrogen ions (Eq. 16-48) or to oxidize H_2 by the reverse of this reaction.[441,442]

$$2H^+ + 2e^- \rightarrow H_2 \qquad (16\text{-}48)$$

Hydrogenases have been classified into two main types: **Fe-hydrogenases**, which contain iron as the only metal,[443] and **Ni-hydrogenases**, which contain both iron and nickel.[444] In a few Ni-hydrogenases a selenocysteine residue replaces a conserved cysteine side chain.[445,446] Fe-hydrogenases are often extremely active and are utilized to rid organisms of an excess of electrons by evolution of H_2. Since they may also be used to acquire electrons by oxidation of H_2, they are often described as *bidirectional*. Ni-hydrogenases, as well as some Fe-hydrogenases, are involved primarily in *uptake* of H_2.[447] All hydrogenases contain one or more Fe–S centers in addition to the H_2-forming catalytic center.[441] Some hydrogenases are membrane bound and are often coupled through unidentified carriers to formate dehydrogenase (Chapter 17). In the strict anaerobes such as clostridia, hydrogenases are linked to ferredoxins. Hydrogenases are inactivated readily by O_2, which oxidizes the catalytic centers but can sometimes be reactivated by treatment with reducing agents.[448]

The 60-kDa all-iron monomeric hydrogenase I of *Clostridium pasteurianum* (mentioned on p. 861) contains ferredoxin-like Fe_4S_4 clusters plus additional Fe and sulfur atoms organized as a special H cluster. The EPR spectrum of the catalytic center, recognized because the spectrum is altered by the binding of carbon monoxide, is unusual. Its *g* values of 2.00, 2.04, and 2.10 are similar to those of oxidized high-potential iron proteins.[449] The *C. pasteurianum* hydrogenase II is a 53-kDa monomer containing eight Fe and eight S^{2-} ions. These are organized into one ferredoxin-like Fe_4S_4 cluster plus a three-Fe cluster and one iron ion in a unique environment.[449]

In contrast to these iron-only hydrogenases, the large periplasmic hydrogenase of *Desulfovibrio gigas* consists of one 28-kDa subunit and one 60-kDa subunit and contains two Fe_4S_4 clusters, one Fe_3S_4 cluster, and another dimetal center containing a single atom of Ni.[450,451] These are seen clearly in the three-dimensional structure depicted in Fig. 16-26. The three Fe–S clusters, at 5- to 6-nm intervals, form a chain from the external surface to the deeply buried nickel–iron center.[450–453a] A plausible pathway for transport of protons to the active center can also be seen.[450,451] The nickel center also contains an atom of Fe. Four cysteine side chains participate in forming the Ni–Fe cluster, two of them provide sulfur atoms that bridge between the metals while two others are ligands to Ni (Fig. 16-26).



(16-49)

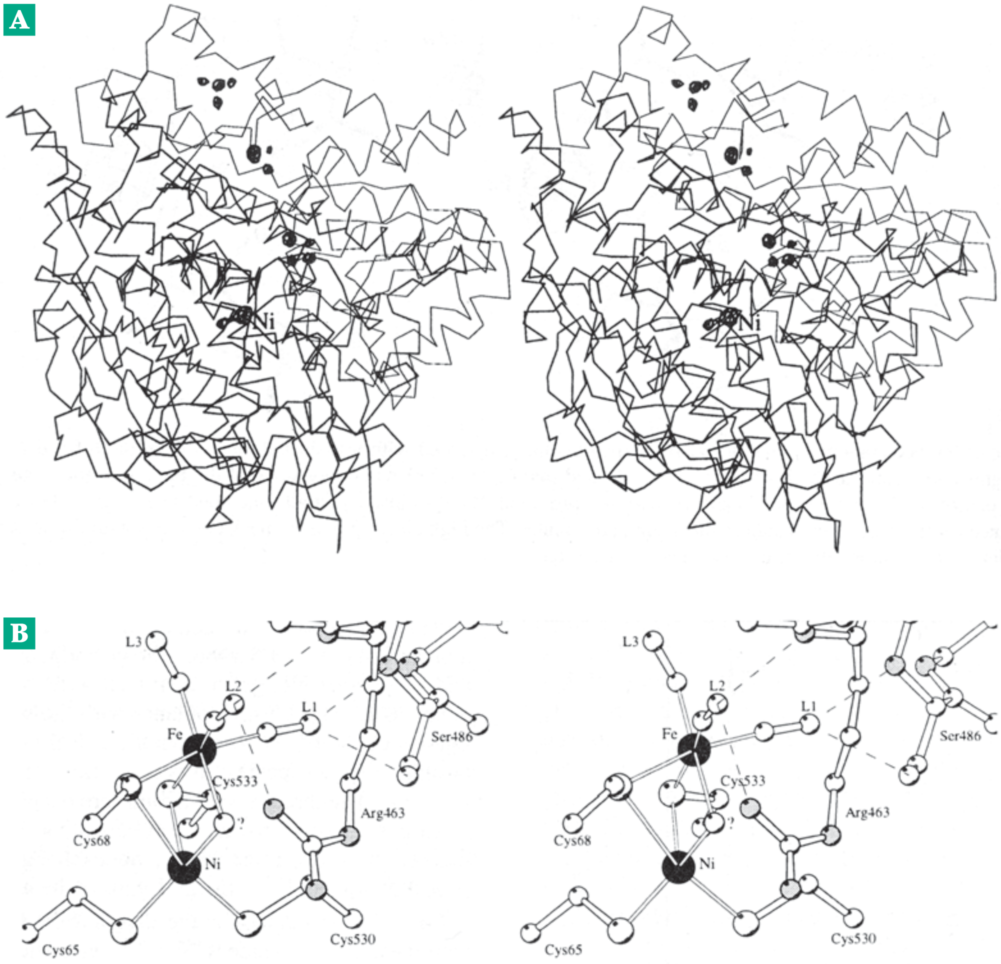


Figure 16-26 (A) Stereoscopic view of the structure of the *Desulfovibrio gigas* hydrogenase as an α-carbon plot. The electron density map at the high level of 8σ is superimposed and consists of dark spheres representing the Fe and Ni atoms. The iron atoms of the two Fe_4S_4 and one Fe_3S_4 clusters are seen clearly forming a chain from the surface of the protein to the Ni–Fe center. (B) The structure of the active site Ni–Fe pair. The two metals are bridged by two cysteine sulfur atoms and an unidentified atom, perhaps O, and the nickel is also coordinated by two additional cysteine sulfurs. Unidentified small molecules L1, L2, and L3 are also present. From Volbeda *et al.*[453] Courtesy of M. Frey.

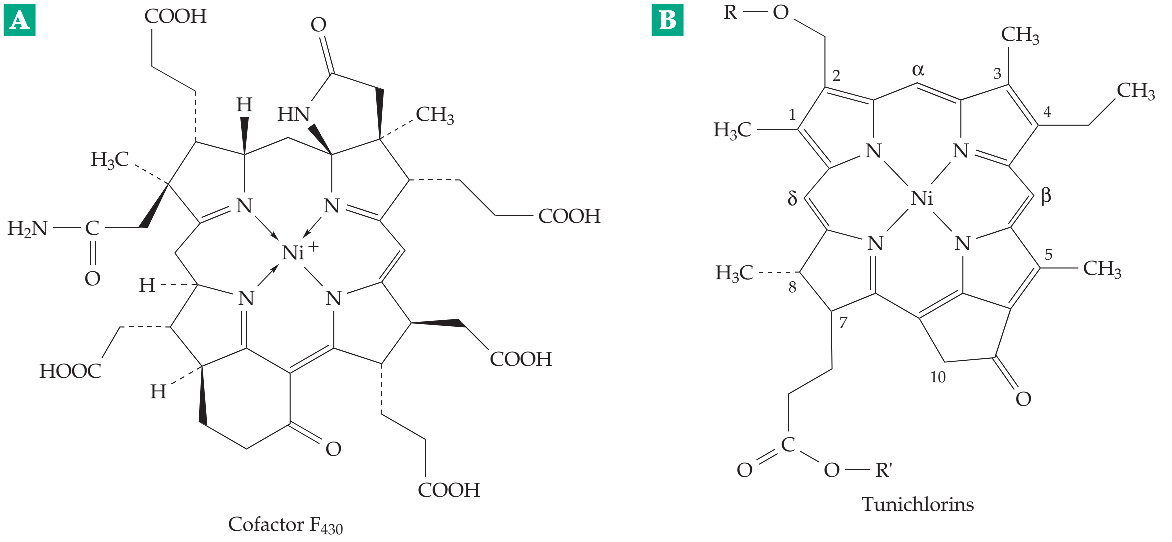
A hydrogenase from *chromatium vinosum* has a similar structure.[453b] There are still uncertainties about other nonprotein ligands such as H_2O.[452–453a]

All of the Ni-hydrogenases display an EPR signal that can be assigned to Ni(III).[452] However, the active enzyme from *D. gigas* contains Ni(II). A proposed mechanism[452] is indicated in Eq. 16-49. Step *a* of this equation is a reductive activation. In step *b* a molecule of H_2 is bound as a hydride ion on Ni and a proton on a nearby sulfur. Protonation of a second sulfur ligand to Ni is needed to promote the cleavage of H_2 prior to the two-step oxidation of the bound H^-. One of two Ni-containing hydrogenases of *Methanobacterium thermoautotrophicum* contains FAD as well as Fe–S clusters.[454] It specifically reduces the 5-deazaflavin cofactor F_{420} (Chapter 15). A major function of this deazaflavin is reduction of the nickel-containing cofactor F_{430}.

3. Cofactor F_{430} and Methyl-Coenzyme M Reductase

Cofactor F_{430} is a nickel tetrapyrrole with a structure (Fig. 16-27)[455,456] similar to those of vitamin B_{12} and of siroheme. The tetrapyrrole ring is the most highly reduced in cofactor F_{430}, which functions in reduction of methyl-CoM to methane in methanogens (Fig. 16-28). The methyl CoM reductase of *Methanobacterium thermoautotrophicum* is a large 300-kDa protein with subunit composition $\alpha_2\beta_2\gamma_2$ and containing two molecules of bound F_{430}. The nickel in F_{430} is first thought to be reduced, in an activation step, to Ni(I),[456a] which may attack the methyl group of methyl-CoM homolytically to yield a methyl nickel complex and a sulfur radical.[457,458] Alkyl nickel compounds react with protons, and in this case they would yield methane and would regenerate the Ni(II) form of the cofactor. The CoM radical could be reduced back to free CoM.

High-resolution crystal structures of the enzyme in two inactive Ni(II) forms[458] show the two F_{430} molecules. Each is bound in an identical channel about 3 nm in length and extending from the surface deep into the interior of the protein. The F_{430} lies at the bottom of this channel with its nickel atom coordinated with the oxygen atom of a glutamine side chain. In one form CoM lies directly above the nickel, with its


A
B
COOH
O
H
HN
CH3
H3C
N
N
Ni+
H2N
C
O
H
N
N
COOH
HOOC
H
COOH
O
COOH
R — O
CH3
2
α
3
CH3
4
H3C
1
N
N
Ni
δ
β
N
N
H3C
8
7
5
CH3
10
O
C
O
O — R'

Cofactor F_{430}

Tunichlorins

Figure 16-27 (A) Structure of the nickel-containing prosthetic group F_{430} as isolated in the esterified (methylated) form. From Pfaltz *et al.*[459] The "front" face, which reacts with methyl-coenzyme M, is toward the reader.[458] (B) Structure of a representative member of a family of tunichlorins isolated from marine tunicates.[460] For tunichlorin R = R′ = H. Related compounds have R′ = CH_3 and/or R = an alkyl group with 13–21 carbon atoms and up to six double bonds.

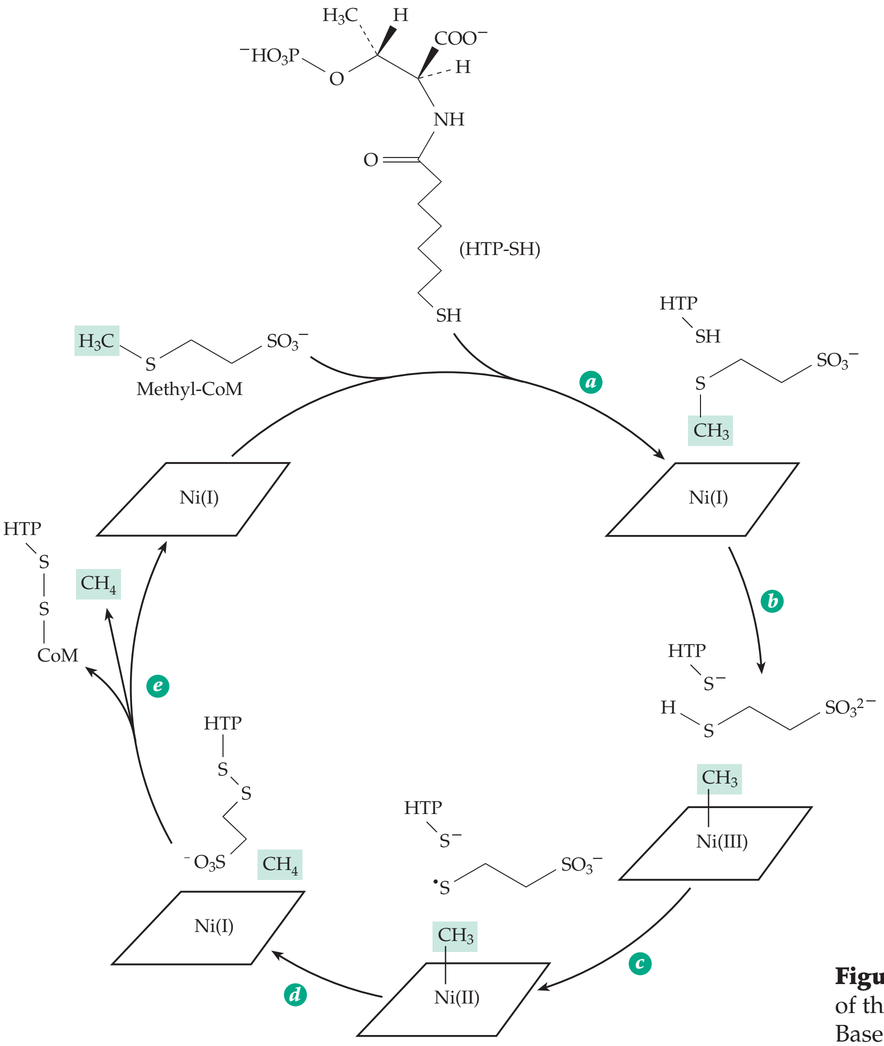

H3C
H
COO−
−HO3P
O
H
NH
O
(HTP-SH)
SH
H3C
S
SO3−
Methyl-CoM
a
HTP
SH
S
SO3−
CH3
Ni(I)
Ni(I)
b
HTP
S
S
CH4
CoM
e
HTP
S−
H
S
SO3 2−
HTP
S
S
−O3S
CH4
HTP
S−
•S
SO3−
CH3
Ni(III)
Ni(I)
CH3
Ni(II)
c
d

Figure 16-28 Proposed mechanism of action of the methane-forming coenzyme M reductase. Based on the crystal structure.

thiolate sulfur providing the sixth ligand for the nickel. The long –SH-containing side chain of **heptanoyl threonine phosphate** (HTP; also called coenzyme B; Chapter 15, Section E) also lies within the channel with its amino acid head group blocking the entrance. In a second crystal form the mixed disulfide HTP–S–S–CoM, an expected product of the reaction (Fig. 16-28), is present in the channel.[458] Because of the distance from the –SH group of HTP and the nickel atom it is clear that there must be some motion of the methyl-CoM and that the methane formed may stay trapped in the active site until the HTP–S–S–CoA product leaves.

A proposed mechanism for the catalytic cycle based on the X-ray results as well as previous chemical studies and EPR spectroscopy is shown in Fig. 16-28. The substrates enter in step *a*. The position of the HTP, with its extended side chain, is probably the same as that seen in the X-ray structures of the Ni(II) complexes but the conformation of the methyl-CoM is different. The methyl transfer in step *b* is reminiscent of that of methionine synthase (Eq. 16-43). Although the distance from the CoM sulfur and the HTP–SH is too great for direct proton transfer between the two, as indicated for step *b,* there are two tyrosine hydroxyls that could provide a pathway for proton transfer. The region around the surface of the nickel coenzyme is largely hydrophobic and could facilitate formation of the thiyl radical in step *c*. In the structure with the bound HTP–S–S–CoM heterodisulfide an oxygen atom of the sulfonate group of CoM is bonded tightly to the Ni(II). However, in the active Ni(I) form the nickel is nucleophilic and would probably repel the sulfonate, perhaps assisting the product release.[458]

The enzyme contains five posttranslationally modified amino acids near the active site: *N*-methylhistidine, 5-methylarginine, 2-methylglutamine, 2-methylcysteine, and thioglycine in a thiopeptide bond. The latter may be the site of radical formation.[458a,b]

4. Tunichlorins

Nickel is found in various marine invertebrates. In the tunicates (sea squirts and their relatives) it occurs in a fixed ratio with cobalt, suggesting a metabolic role.[460] A new class of nickel chelates called tunichlorins have been isolated. An example is shown in Fig. 16-27B. The function of tunichlorins is unknown but their existence suggests the possibility of unidentified biochemical roles for nickel.

5. Carbon Monoxide Dehydrogenases and Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase

There are several bacterial carbon monoxide dehydrogenases that catalyze the reversible oxidation of CO to CO_2:

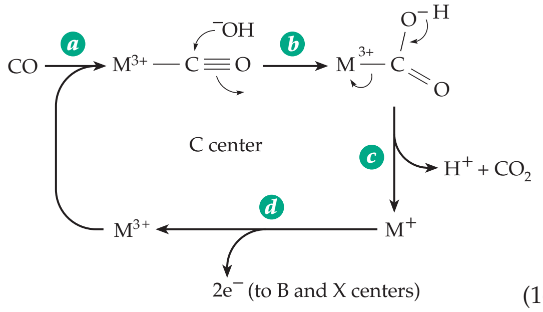
$$CO + H_2O \rightarrow CO_2 + H^+ + 2\,e^- \qquad (16\text{-}50a)$$

$$\begin{gathered} CO + H_2O \rightarrow CO_2 + H_2 \\ \Delta G^{\circ} = -38.1 \text{ kJ / mol} \end{gathered} \qquad (16\text{-}50b)$$

Some bacteria use CO as both a source of energy and for synthesis of carbon compounds. The purple photosynthetic bacterium *Rhodospirillum rubrum* employs a relatively simple monomeric Ni-containing CO dehydrogenase containing one atom of Ni and seven or eight iron atoms, apparently arranged in Fe_4S_4 clusters. These bacteria can grow anaerobically with CO as the sole source of both energy and carbon.[461] Some aerobic bacteria oxidize CO using a molybdenum-containing enzyme (Section H). However, the most studied CO dehydrogenase is a complex enzyme that also synthesizes, reversibly, acetyl-CoA from CO and a methyl corrin. Employed by methanogens, acetogens, and sulfate-reducing bacteria, it is at the heart of the **Wood–Ljundahl pathway** of autotrophic metabolism, which is discussed further in Chapter 17.

Both oxidation of CO to CO_2 and reduction of CO_2 to CO are important activities of CO dehydrogenase / acetyl-CoA synthase. During growth on CO, some CO must be oxidized to CO_2 and then reduced by the pathways of Fig. 15-22 to form a methyl-tetrahydropterin which can be used to form the methyl group of acetyl-CoA. During growth on any other carbon compound CO_2 must be reduced to CO to form the carbonyl group of acetyl-CoA which can serve as a precursor to all other carbon compounds. Native CO dehydrogenase/acetyl CoA synthase was isolated from cells of *Clostridium thermoaceticum* grown in the presence of radioactive ^{63}Ni. The protein is a 310-kDa $\alpha_2\beta_2$ oligomer. Each $\alpha\beta$ dimer contains 2 atoms of Ni, 1 of Zn, ~12 of Fe and ~12 sulfide ions,[462–464] which are organized into three metal clusters referred to as A, B, and C. Each cluster contains 4 Fe atoms and clusters A and C also contain 1 Ni each. Oxidation of CO occurs in the β subunits, each of which contains both cluster B, an Fe_4S_4 ferredoxin-type cluster, and cluster C, where the oxidation of CO is thought to occur. Cluster C contains 1 nickel ion as well as an Fe_4S_4 cluster that resembles that of aconitase (Fig. 13-4). Cluster A, which is in the α subunit, also contains 1 atom of Ni and 4 Fe ions and is probably the site of synthesis of acetyl-CoA.

Oxidation of CO may require cooperation of the nickel ion and the Fe_4S_4 group within the C cluster.



(16-51)

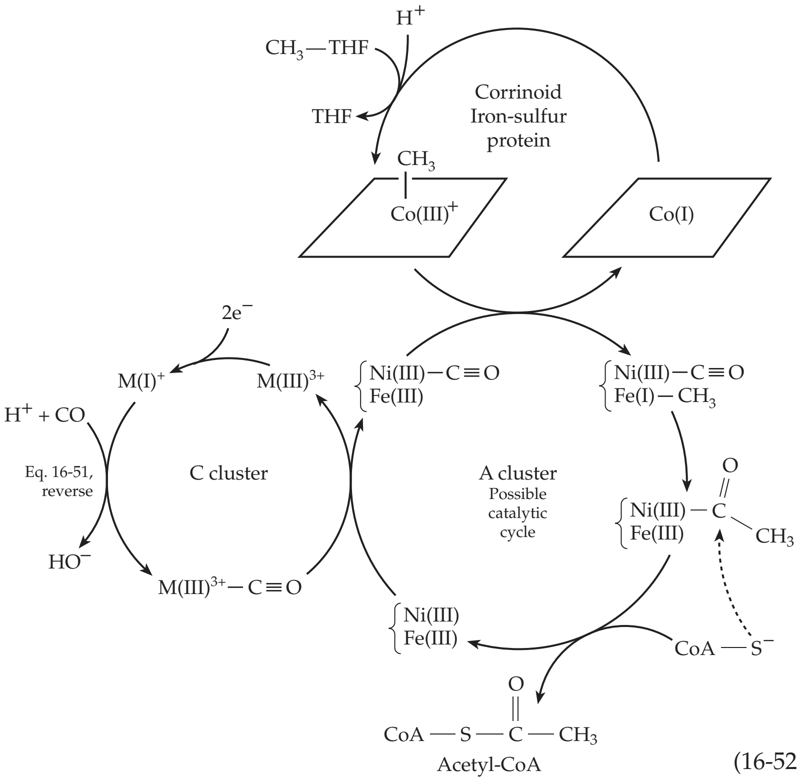
CO probably binds to one of these metals and is attacked by a hydroxyl ion (Eq. 16-51, step *b*) which may be donated by the other metal of the pair.[465,465a] CO_2 and a proton are released rapidly (step *c*), after which the reduced metal center is reoxidized (step *d*). One electron is thought to be transferred directly to the Fe_4S_4 cluster B and the second by an alternative route.[465] A multienzyme complex isolated from the methanogen *Methanosarcina* has an $(\alpha\beta\gamma\delta\varepsilon)_6$ structure, with the subunits having masses of 89, 60, 50, 48, and 20 kDa, respectively.[466] In this complex the CO dehydrogenase/acetyl-CoA synthase activity appears to reside in the $\alpha_2\varepsilon_2$ complex, the $\gamma\delta$ complex has a tetrahydropteridine: cob(I)amide–protein transferase, and the β subunit has an acetyltransferase that binds acetyl-CoA and transfers the acetyl group to a group on the β subunit.[466]

Although the *Clostridium* and *Methanosarcina* systems are not identical, similar mechanisms are presumably involved.[467,467a] To generate acetyl-CoA a methyl group is first transferred from a tetrahydropterin such as tetrahydrofolate or, in methanogens, tetrahydromethanopterin or tetrahydrosarcinapterin (Fig. 15-17) to form a methylcorrinoid. At the A center of the CO dehydrogenase a molecule of CO, which may be bound to the Ni, equilibrates with the methyl group, and with acetyl-CoA. As depicted in Eq. 16-52, acetyl-Ni may be an intermediate. Other details shown here are hypothetical. It is possible that the methyl group is transferred to the Ni atom in the M cluster before reaction with the CO, which might be bound to either Ni of cluster C or to Fe. This reaction of two transition-metal-bound ligands parallels a proposed industrial process for synthesis of acetic acid from methanol and CO and involving catalysis by rhodium metal and methyl iodide. It is thought that rhodium-bound CO is inserted into bound $Rh{-}CH_3$ to form an intermediate $Rh{-}\overset{O}{\overset{\|}{C}}{-}CH_3$. The acetyl group is released as acetyl iodide which is hydrolyzed to acetic acid. An acetyl-nickel intermediate may be involved in the corresponding biological reaction of Eq. 16-52. The stereochemistry of the sequence has been investigated using methyl-THF containing a chiral methyl group. Overall retention of the configuration of the methyl group in acetyl-CoA was observed.[468,469]

D. Copper

Copper was recognized as nutritionally essential by 1924 and has since been found to function in many cellular proteins.[470–474] Copper is so broadly distributed in foods that a deficiency has only rarely been observed in humans.[474a] However, animals may sometimes receive inadequate amounts because absorption of Cu^{2+} is antagonized by Zn^{2+} and because copper may be tied up by molybdate as an inert complex. There are copper-deficient desert areas of Australia where neither plants nor animals survive. Copper-deficient animals have bone defects, hair color is lacking, and hemoglobin synthesis is impaired. Cytochrome oxidase activity is low. The protein elastin of arterial walls is poorly crosslinked and the arteries are weak. Genetic defects in copper metabolism can have similar effects.

An adult human ingests ~2–5 mg of copper per day, about 30% of which is absorbed. The total body content of copper is ~100 mg (~2×10^{-4} mol/kg), and both uptake and excretion (via the bile) are regulated. Since an excess of copper is toxic, regulation is important.



(16-52)

Because Cu^{2+} is the most tightly bound metal ion in most chelating centers (Table 6-9), almost all of the copper present in living cells is complexed with proteins. Copper is transported in the blood by a 132-kDa, 1046-residue sky-blue glycoprotein called **ceruloplasmin**.[471,475–477] This one protein contains 3% of the total body copper.

Regulation of copper uptake has been studied in most detail in the yeast *Saccharomyces cerevisiae*. Uptake of Cu^{2+} is similar to that of Fe^{3+}. The same plasma membrane reductase system, consisting of proteins Fre1p and Fre2p (encoded by genes *FRE*1 and *FRE*2), acts to reduce both Fe^{3+} and Cu^{2+}.[478–481] These two genes are controlled in part by a transcriptional activator that responds to the internal copper concentration.[410,482,483] Similar regulation is thought to occur in both plants[484] and animals.

The human hereditary disorders **Wilson's disease** and **Menkes' disease** have provided further insight into copper metabolism.[485,486] In Wilson's disease the ceruloplasmin content is low and copper gradually accumulates to high levels in the liver and brain. In Menkes syndrome, there is also a low ceruloplasmin level and an accumulation of copper in the form of copper metallothionein.[487,488] Persons with this disease have abnormalities of hair, arteries, and bones and die in childhood of cerebral degeneration.[489,490] Similar symptoms are seen in some patients with Ehlers–Danlos syndrome (Box 8-E).[491] Genes for the proteins that are defective in both Wilson's and Menkes' diseases have been cloned and both proteins have been identified as P-type ATPase cation transporters (Chapter 8).[492–495c] The two proteins must be similar in structure as indicated by a 55% sequence identity.[496] Homologous genes involved in copper homeostasis have been located in both yeast[497] and the cyanobacterium *Synechococcus*.[498] The transporter encoded by this yeast gene, designated *Ccc*2, apparently functions to export copper from the cytosol into an extracytosolic compartment. In a similar way the Wilson and Menkes disease proteins, which reside in the *trans*-Golgi network, are thought to export copper or to provide copper for incorporation into essential proteins.[493,499] The Wilson disease protein is also found in a shortened form in mitochondrial membranes.[492] Other proteins associated with intracellular copper metabolism seem to be chaperones for Cu(I).[500–501a]

The ability of copper ions to undergo reversible changes in oxidation state permits them to function in a variety of oxidation–reduction processes. Like iron, copper also provides sites for reaction with O_2, with superoxide radicals, and with nitrite ions.

1. Electron-Transferring Copper Proteins

A large group of small, intensely blue copper proteins function as single-electron carriers within bacteria and plants. Best known is **plastocyanin**, which is ubiquitous in green plants and functions in the electron transport chain between the light-absorbing photosynthetic centers I and II of chloroplasts (Chapter 23). The bacterial **azurins**[502] are thought to carry electrons between cytochrome c_{441} and cytochrome oxidase. **Amicyanin** accepts electrons from the coenzyme TTQ of methylamine dehydrogenase of methylotrophic bacteria and passes them to a cytochrome *c* (Chapter 15).[168,503,504] A basic blue copper protein **phytocyanin** of uncertain function occurs in cucumber seeds.[505]

The 10.5-kDa peptide chain of plastocyanin is folded into an eight-stranded β barrel (Fig. 2-16), which contains a single copper atom. In poplar plastocyanin, the Cu is coordinated by the side chains of His 37, His 87, Met 92, and Cys 84 in a tetrahedral but distorted toward a trigonal bipyramidal geometry. Since copper-free apoplastocyanin has essentially the same structure, this geometry may be imposed by the protein onto the Cu^{2+}, which usually prefers square-planar or tetrahedral coordination (Chapter 7).[506] Calculations suggest that there is little or no strain and that the "Franck–Condon barrier" to electron transfer is low.[507,508] The three-dimensional structure and copper environment of azurin are similar to those of plastocyanin.[509–511] Messerschmidt *et al.*[512] suggested that the copper site in these proteins is perfectly adapted to its function because its geometry is a compromise between the optimal geometries of the Cu(I) and Cu(II) states between which it alternates. **Stellacyanin**, present in the Japanese lac tree and some other plants, is a mucoprotein; the 108-residue protein is over 40% carbohydrate.[513] While its spectrum resembles that of plastocyanins and azurins, stellacyanin contains no methionine and this amino acid cannot be a ligand to copper.[514] **Rusticyanin** functions in the periplasmic space of some chemolithotropic sulfur bacteria to transfer electrons from Fe^{2+} to cytochrome *c* as part of the energy-providing reaction for these organisms (Eq. 18-23).[515–517] **Halocyanin** functions in membranes of the archaeobacterium *Natronobacterium*,[518] and **aurocyanin** functions in green photosynthetic bacteria.[519]

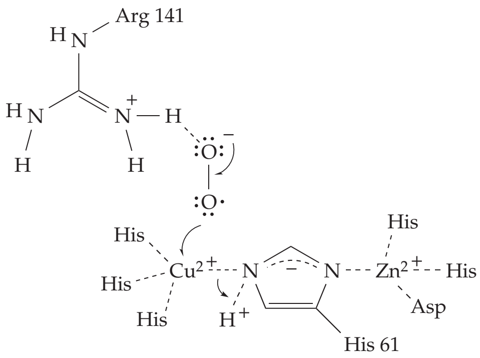
The blue color of these "type 1" copper proteins is much more intense than are the well known colors of the hydrated ion $Cu(H_2O)_4{}^{2+}$ or of the more strongly absorbing $Cu(NH_3)_4{}^{2+}$. The blue color of these simple complexes arises from a transition of an electron from one *d* orbital to another within the copper atom. The absorption is somewhat more intense in copper peptide chelates of the type shown in Eq. 6-85. However, the ~600 nm absorption bands of the blue proteins are an order of magnitude more intense, as is illustrated by the absorption spectrum of azurin (Fig. 23-8). The intense blue is thought to arise as a result of transfer of electronic charge from the cysteine thiolate to the Cu^{2+} ion.[520,521]

A third type of copper center, first recognized in cytochrome *c* oxidase (see Fig. 18-10) is called Cu_A or **purple CuA**. Each copper ion is bonded to an imidazole and two cysteines serve as bridging ligands. The two copper ions are about 0.24 nm apart, and the two Cu^{2+} ions together can accept a single electron from an external donor such as cytochrome *c* or azurin to give a half-reduced form.[521a,b]

2. Copper, Zinc-Superoxide Dismutase

Although of similar topology to the blue electron-transferring proteins, Cu, Zn-superoxide dismutase, has a different function. This dimeric 153-residue protein has been demonstrated in the cytoplasm of virtually all eukaryotic cells[522] and in the periplasmic space of some bacteria[523] where it converts superoxide ions $^{\bullet}O_2^-$ to O_2 and H_2O_2. The enzyme, which has a major protective role against oxidative damage to cells, presumably functions in a manner similar to that indicated in Eq. 16-27 for iron or manganese. However, copper cycles between Cu^{2+} and Cu^+, alternately accepting and donating electrons.

The active site of cytosolic superoxide dismutase (SOD) contains both Cu^{2+} and Zn^{2+}. The copper ion is of "type 2": nonblue and paramagnetic. It is surrounded by four imidazole groups with an irregular square planar geometry.[524–527] One of these imidazole groups (that of His 61) is shared with the Zn^{2+}, which is also bonded to two additional imidazole groups and a side chain carboxylate. The metal ions have evidently replaced the hydrogen atom that would otherwise be present on the imidazole of His 61 (see the following diagram). It has been suggested, as is also indicated in the diagram, that when the bound superoxide donates an electron to the Cu(II) to become O_2 (first step of Eq. 16-27), a proton becomes attached to the bridging imidazole with breakage of its linkage to the Cu(I). The structure of a new crystalline form of reduced yeast SOD shows that the Cu(I) has moved 0.1 nm away from the bridging imidazole in agreement with this possibility.[528] In the second half reaction the imidazole proton, together with a second proton from the medium and an electron from the Cu(I), would react to convert the second O_2^- into H_2O_2. The role of the Zn^{2+} may be in part structural but it may also serve to ensure that His 61 is protonated on the correct nitrogen atom. Arg 141 may assist in binding the O_2^- as is shown in the diagram. However, the fact that a mutant containing leucine in place of the active site arginine has over 10% of the activity of the native enzyme shows that the arginine is not absolutely essential.[529] Additional nearby positively charged arginine and lysine side chains may provide "electrostatic guidance" that increases the velocity of reaction of superoxide ions.[530,531] Cu, Zn-SOD is one of the fastest enzymes known.



In addition to the cytosolic SOD there is a longer ~222-residue extracellular form that binds to the proteoglycans found on cell surfaces.[522,527] Manganese SODs are found in mitochondria and in bacteria and iron SODs in plants and bacteria. They all appear to be important in protecting cells from superoxide radicals.[522,532,533] This importance was dramatically emphasized when it was found that a defective SOD is present in persons (about 1 in 100,000) with a hereditary form of **amylotrophic lateral sclerosis** (ALS), which is also called Lou Gehrig's disease after the baseball hero who was stricken with this terrible disease of motor neurons in 1939 at the age of 36.[534–536]

3. Nitrite and Nitrous Oxide Reductases

Copper enzymes participate in two important reactions catalyzed by denitrifying bacteria. Nitrite reductases from species of *Achromobacter*[537,538] and *Alcaligenes*[539–542] are trimeric proteins[543] made up of 37-kDa subunits, each of which contains one type 1 (blue) copper and one type 2 (nonblue) copper. The first copper serves as an electron acceptor from a small blue **pseudoazurin**.[544,544a] The second copper, which is in the active site, is thought to bind to nitrite throught its nitrogen atom and to reduce it to NO.

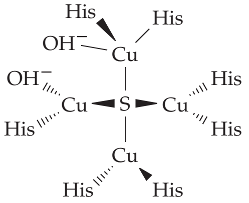
$$NO_2^- + 2H^+ + e^- \rightarrow NO + H_2O \qquad (16\text{-}53)$$

Crystallographic studies on the 343-residue *Alcaligenes* enzyme reveal two β barrel domains with the type 1 copper embedded in one of them and the type 2 copper in an interface between the domains. Studies of EPR[541] and ENDOR[545] spectra and of various mutant forms have shown that, as for other copper enzymes, the type 1 copper is an electron-transferring center, accepting electrons from the pseudoazurin and passing them to the type 2 copper which binds and reduces nitrite.[540,545a]

The immediate product of nitrite reductase is NO, which is reduced in two one-electron steps to N_2O, and then to N_2. The second of these steps is catalyzed by another copper enzyme, nitrous oxide reductase.[546]

$$N_2O + 2H^+ + 2e^- \rightarrow N_2 + H_2O \qquad (16\text{-}54)$$

The biological significance of these reactions is considered further in Chapters 18 and 24. The 132-kDa dimeric N_2O reductase from *Pseudomonas stutzeri* contains four copper atoms per subunit.[546] One of its copper centers resembles the Cu_A centers of cytochrome *c* oxidase. A second copper center consists of four copper ions, held by seven histidine side chains in a roughly tetrahedral array around one sulfide (S^{2-}) ion. Rasmussen *et al.* speculate that this copper-sulfide cluster may be an acceptor of the oxygen atoms of N_2O in the formation of N_2.[546a] There is also a cytochrome *cd_1* type of nitrite reductase.[143a]



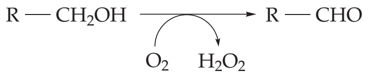
4. Hemocyanins

While many copper proteins are catalysts for oxidative reactions of O_2, hemocyanin reacts with O_2 reversibly. This water-soluble O_2 carrier is found in the blue blood of many molluscs and arthropods, including snails, crabs, spiders, and scorpions. Hemocyanins are large oligomers ranging in molecular mass from 450 to 13,000 kDa. Molluscan hemocyanins are cylindrical oligomers which have a striking appearance under the electron microscope. Simpler hemocyanins, found in arthropods, are hexamers of 660-residue 75-kDa subunits. Each subunit of the hemocyanin from the spiny lobster is folded into three distinct domains, one of which contains a pair of Cu(I) atoms which bind the O_2. Each copper ion is held by three imidazole groups without any bridging groups between them, the Cu–Cu distance being 0.36–0.46 nm.[547] *Octopus* hemocyanin has a different fold and forms oligomers of ten subunits. However, the active sites are very similar.[548] The O_2 is thought to bind between the two copper atoms. An allosteric mechanism may involve changes in the distance between the copper atoms.[549] The oxygenated compound is distinctly blue with a molar extinction coefficient 5–10 times greater than that of cupric complexes. This fact suggests that the Cu(I) has been oxidized to Cu(II) and that the O_2 has been reduced to the peroxide dianion O_2^{2-} in the complex.[550–552] Further support for this idea comes from the observation that treatment of oxygenated hemocyanin with glacial acetic acid leads to the formation of equal amounts of Cu^{2+} and Cu^+ and protonated superoxide.

$$(CuO_2Cu)^{2+} + H^+ \longrightarrow Cu^{2+} + Cu^+ + HO_2 \qquad (16\text{-}55)$$

5. Copper Oxidases

A large group of copper-containing proteins activate oxygen toward chemical reactions of dehydrogenation, hydroxylation, or oxygenation. **Galactose oxidase** (Fig. 16-29), from the mushroom *Polyporus,* is a dehydrogenase which converts the 6-hydroxymethyl group of galactose to an aldehyde while O_2 is reduced to H_2O_2.



$$R{-}CH_2OH \xrightarrow[O_2 \quad H_2O_2]{} R{-}CHO$$

(16-56)

Galactose oxidase has been used frequently to label glycoproteins of external cell membrane surfaces. Exposed terminal galactosyl or *N*-acetylgalactosaminyl residues are oxidized to the corresponding C-6 aldehydes and the latter are reduced under mild conditions with tritiated sodium borohydride.[553]

The single 639-residue polypeptide chain contains one type 2 copper ion.[554] Neither oxygen nor galactose affects the absorption spectrum of the light murky green enzyme, but the combination of the two does, suggesting that both substrates bind to the enzyme before a reaction takes place. A side reaction releases

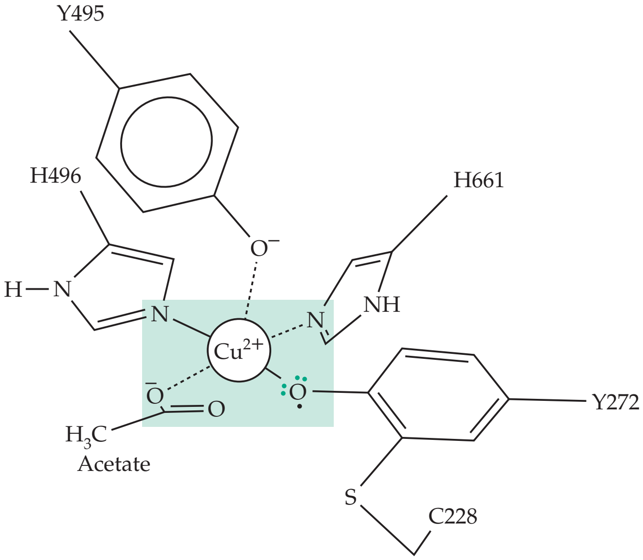


Figure 16-29 Drawing of the active site of galactose oxidase showing both the Cu(II) atom and the neighboring free radical on tyrosine 272, which has been modified by addition of the thiol of cysteine 228 and oxidation. See Halfen *et al.*[557] Based on a crystal structure of Ito *et al.*[558]

superoxide ion and leaves the enzyme in the inactive Cu(II) state. EPR spectroscopic observations on the enzyme were puzzling. The active enzyme shows no EPR signal but a one-electron reduction gives an inactive form with an EPR signal that arises from Cu(II). Experimental studies eventually pointed to the presence of a second reducible center which contains an organic free radical. In the active form this radical is **antiferromagnetically coupled** (spin-coupled) giving an "EPR-silent" enzyme able to accept two electrons.[555–557]

Another surprise was the discovery, from the X-ray structure,[558–559] that a tyrosine side chain at the active site has been modified by addition of a thiolate group from a cysteine residue. This structure, which is also the site of the organic free radical, is shown in Fig. 16-29. A possible mechanism of action is portrayed in Fig. 16-30. The substrate binds in step *a* and its –OH group is deprotonated, perhaps by transfer of H^+ to the phenolate oxygen of tyrosine 495 (Fig. 16-29) in step *b* of Fig. 16-30. In step *c* a free radical hydrogen transfer occurs to form a ketyl radical which is immediately oxidized by Cu(II) in step *d*. In steps *c* and *f* the oxidant O_2 is converted to H_2O and the aldehyde product is also released. A fungal glyoxal oxidase has similar characteristics.[560] The galactose oxidase proenzyme is self-processing. The Tyr-Cys cofactor arises as a result of copper-catalyzed oxidative modification via a tyrosine free radical.[560a]

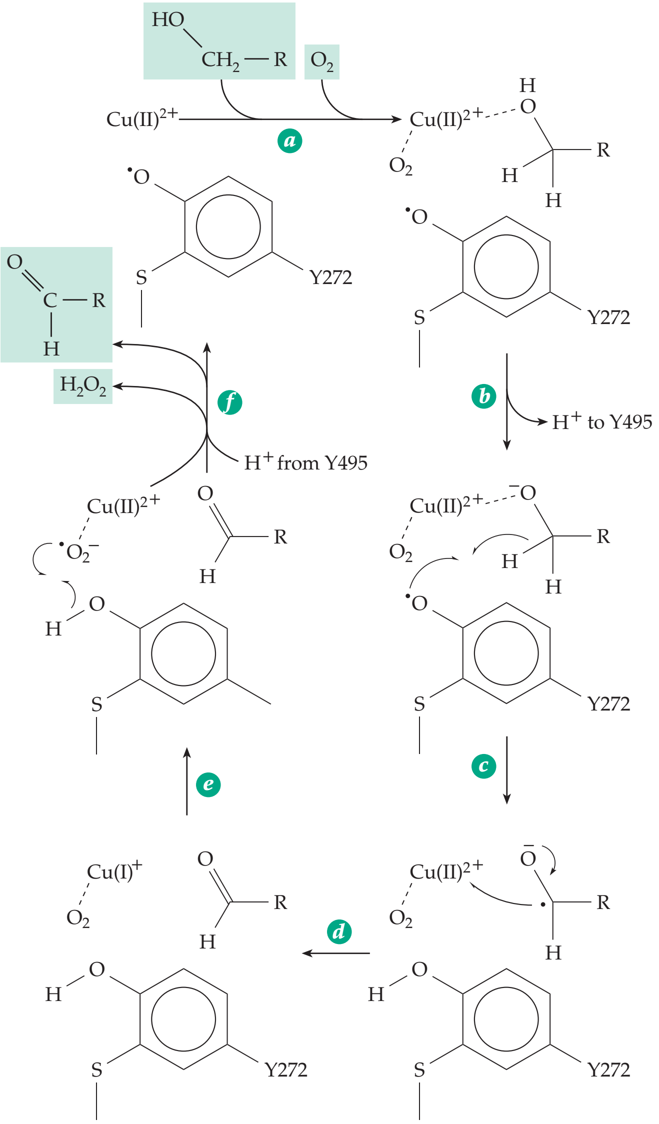
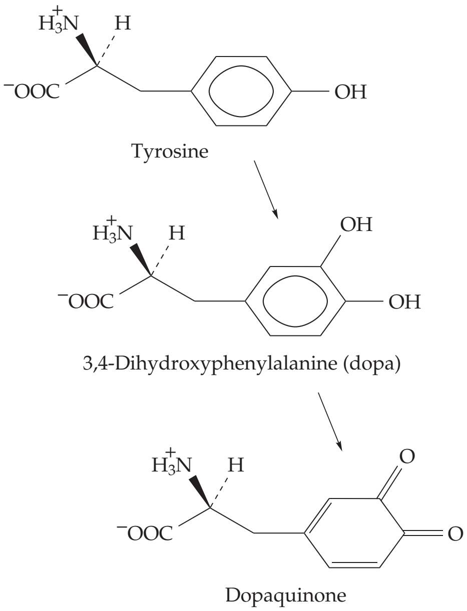


Figure 16-30 Possible reaction cycle for catalysis of the 6–OH of D-galactose or other suitable alcohol substrate by galactose oxidase. See Wachter *et al.*[566] and Whittaker *et al.*[567]

Several copper-containing amine oxidases[561,561a] convert amines to aldehydes and H_2O_2. They also contain one of the organic quinone cofactors discussed in Chapter 15. The dimeric **plasma amine oxidase** contains a molecule of the coenzyme TPQ and one Cu^{2+} per 90-kDa subunit.[562–563a] Whether the O_2 binds only to the single atom Cu(I) or also interacts directly with a cofactor radical in step *d* of Eq. 15-53 is uncertain.[562] **Lysyl oxidase**, which is responsible for conversion of ε-amino groups of side chains of lysine into aldehyde groups in collagen and elastin (Chapter 8) contains coenzyme LTQ as well as Cu.[564] The enzyme is specifically inhibited by β-aminopropionitrile (Box 8-G) and its activity is decreased in genetic diseases of copper metabolism. The **glycerol oxidase** of *Aspergillus* is a large 400 kDa protein containing one heme and two atoms of Cu.[565] It converts glycerol + O_2 into glyceraldehyde and H_2O_2. Also containing copper is **urate oxidase**, whose action is indicated in Fig. 25-18.

Tyrosinase catalyzes hydroxylation followed by dehydrogenation (Eq. 16-57). First identified in mushrooms, the enzyme has a widespread distribution in



(16-57)

nature. It is present in large amounts in plant tissues and is responsible for the darkening of cut fruits. In animals tyrosinase participates in the synthesis of **dihydroxyphenylalanine** (dopa) and in the formation of the black **melanin** pigment of skin and hair. Either a lack of or inhibition of this enzyme in the melanin-producing melanocytes causes **albinism** (Chapter 25).

The 46-kDa monomeric tyrosinase of *Neurospora* contains a pair of spin-coupled Cu(II) ions.[568,569] The structure of this copper pair (**type 3 copper**) has many properties in common with the copper pair in hemocyanin.[569a] For example, in the absence of other substrates, tyrosinase binds O_2 to form "oxytyrosinase," a compound with properties resembling those of oxyhemocyanin and containing a bound peroxide dianion.[569]

Tyrosinase is both an oxidase and a hydroxylase. Some other copper enzymes have only a hydroxylase function. One of the best understood of these is the **peptidylglycine α-hydroxylating monoxygenase**, which catalyzes the first step of the reaction of Eq. 10-11. The enzyme is a colorless two-copper protein but the copper atoms are 1.1 nm apart and do not form a binuclear center.[570] Ascorbate is an essential cosubstrate, with two molecules being oxidized to the semidehydroascorbate radical as both coppers are reduced to Cu(I). A ternary complex of reduced enzyme, peptide, and O_2 is formed and reacts to give the hydroxylated product.[570] A related two-copper enzyme is **dopamine β-monooxygenase**, which utilizes O_2 and ascorbate to hydroxylate dopamine to noradrenaline (Chapter 25).[571,572] These and other types of hydroxylases are compared in Chapter 18.

The **blue multicopper oxidases** couple the oxidation of substrates to the four-electron reduction of molecular oxygen to H_2O.[573] In this respect they resemble cytochrome *c* oxidase, which also contains copper. However, they do not contain iron. The best known member of this group is the plant enzyme **ascorbate oxidase**, which dehydrogenates ascorbic acid to dehydroascorbic acid (See Box 18-D). It is a dimeric blue copper with identical 70-kDa subunits. The three-dimensional structure revealed one type 1 copper ion held in a typical blue copper environment as in plastocyanin or azurin and also a **three-copper center**. In this center a pair of copper ions, each held by three imidazole groups, and bridged by a μ-oxo group as in hemerythrin (Fig. 16-20), lie 0.51 nm apart and 0.41–0.44 nm away from the third copper, which is held by two other imidazoles.[574] The type 1 copper shows typical intense 600-nm absorption and characteristic EPR signal, while the pair with the oxo bridge are antiferromagnetically coupled and EPR silent but with strong near ultraviolet light absorption (type 3 copper). The additional metal ion in the trinuclear center is a type 2 copper which lacks characteristic spectroscopic features.[575,576] Reduction of O_2 to 2 H_2O is thought to procede via superoxide radical intermediates. When substrate is added to the enzyme, the blue color fades and it can be shown that the copper is reduced to the +1 state. The reduced enzyme then reacts with O_2, converting it into two molecules of H_2O. Similar to ascorbate oxidase in structure and properties are **laccase**, found in the latex of the Japanese lac tree and in the mushroom *Polyporus*, and the previously discussed **ceruloplasmin.**[477] Laccase is a catalyst for oxidation of phenolic compounds by a free radical mechanism involving the trinuclear copper center.[577] Studied by Gabriel Bertrand in the 1890s, it was one of the first oxidative enzymes investigated.[578]

In addition to its previously mentioned role in copper transport, ceruloplasmin is an amine oxidase, a superoxide dismutase, and a ferrooxidase able to catalyze the oxidation of Fe^{2+} to Fe^{3+}. Ceruloplasmin contains three consecutive homologous 350-residue sequences which may have originated from an ancestral copper oxidase gene. Like ascorbate oxidase, this blue protein contains copper of the three different types. Blood clotting factors V and VIII (Fig. 12-17), and the iron uptake protein Fet3 (Section A,1) are also closely related.

6. Cytochrome *c* Oxidase

The most studied of all copper-containing oxidases is cytochrome *c* oxidase of mitochondria. This multisubunit membrane-embedded enzyme accepts four electrons from cytochrome *c* and uses them to reduce O_2 to 2 H_2O. It is also a proton pump. Its structure and functions are considered in Chapter 18. However, it is appropriate to mention here that the essential catalytic centers consist of two molecules of heme *a* (*a* and a_3) and three Cu^+ ions. In the fully oxidized enzyme two metal centers, one Cu^{2+} (of the two-copper center Cu_A) and one Fe^{3+} (heme *a*), can be detected by EPR spectroscopy. The other Cu^{2+} (Cu_B) and heme a_3 exist as an EPR-silent exchange-coupled pair just as do the two copper ions of hemocyanin and of other type 3 binuclear copper centers.

E. Manganese

Tissues usually contain less than one part per million of manganese on a dry weight basis, less than 0.01 mM in fresh tissues. This compares with a total content in animal tissues of the more abundant Mg^{2+} of 10 mM. A somewhat higher Mn content (3.5 ppm) is found in bone. Nevertheless, manganese is nutritionally essential[579,580] and its deficiency leads to well-defined symptoms. These include ovarian and testicular degeneration, shortening and bowing of legs, and other skeletal abnormalities such as the "slipped tendon disease" of chicks. In Mn deficiency the organic matrix of bones and cartilage develops poorly. The galactosamine, hexuronic acids, and chondroitin sulfates content of cartilage is decreased. Manganese is also essential for plant growth and plays a unique and essential role in the photosynthetic reaction centers of chloroplasts. Two magnetically coupled pairs of manganese ions bound in a protein act as the O_2 evolving center in photosynthetic system II. This function is considered in Chapter 23. An ABC transporter for managanese uptake has been identified in the cyanobacterium *Synechocystis*.[581]

Manganese lies in the center of the first transition series of elements. The stable Mn^{2+} (manganous ion) contains five 3*d* electrons in a high-spin configuration. The less stable Mn^{3+} (manganic ion) appears to be of importance in some enzymes and is essential to the photosynthetic evolution of oxygen. Many enzymes specifically require or prefer Mn^{2+}. These include galactosyl and *N*-acetylgalactosaminyltransferases[582] needed for synthesis of mucopolysaccharides (Chapter 20), lactose synthetase (Eq. 20-15), and a muconate-lactonizing enzyme (Eq. 13-23).[583]

Arginase, essential to the production of urea in the human body (Fig. 24-11), specifically requires Mn^{2+} which exists as a spin-coupled dimetal center with a bridging water or ^-OH ion. The Mn^{2+} may act much as does the Ni^{2+} ions of urease (Fig. 16-25).[584,585] Pyruvate carboxylase (Eq. 14-3) contains four atoms of tightly bound Mn^{2+}, one for each biotin molecule present. This manganese is essential for the transcarboxylation step in the action of this enzyme. Either Mn^{2+} or Mg^{2+} is also needed in the initial step of carboxylation of biotin (Eq. 14-5). Another Mn^{2+}-containing protein is the lectin concanavalin A (Chapter 4). The joining of *O*-linked oligosaccharides to secreted glycoproteins also seems to require manganese.[586]

Manganese is a component of a "pseudocatalase" of *Lactobacillus*,[204] of lignin-degrading peroxidases,[257,258] and of the wine-red superoxide dismutases found in bacteria and in the mitochondria of eukaryotes.[376,587] The dimeric dismutase from *E. coli* has a structure nearly identical to that of bacterial iron SOD (Fig. 16-22). The manganese ions are presumed to alternate between the +3 and +2 states during catalysis (Eq. 16-27). "Knockout" mice with inactivated Mn SOD genes live no more than three weeks, indicating that this enzyme is essential to life. However, mice lacking CuZn SOD appear normal in most circumstances.[588] Some dioxygenases contain manganese.[589] Many enzymes that require Mg^{2+} can utilize Mn^{2+} in its place, a fact that has been exploited in study of the active sites of enzymes.[590] The highly paramagnetic Mn^{2+} is the most useful ion for EPR studies (Box 8-C) and for investigations of paramagnetic relaxation of NMR signals. Manganese can also replace Zn^{2+} in some enzymes and may alter catalytic properties.

Manganese may function in the regulation of some enzymes. For example, glutamine synthetase (Fig. 24-7) in one form requires Mg^{2+} for activity but upon adenylylation binds Mn^{2+} tightly.[591] Nucleases and DNA polymerases often show altered specificity when Mn^{2+} substitutes for Mg^{2+}. However, the significance of these differences *in vivo* is uncertain. Manganese is mutagenic in living organisms, apparently because it diminishes the fidelity of DNA replication.[592]

A striking accumulation of Mn^{2+} often occurs within bacterial spores (Chapter 32). *Bacillus subtilus* absolutely requires Mn^{2+} for initiation of sporulation. During logarithmic growth the bacteria can concentrate Mn^{2+} from 1 µM in the external medium to 0.2 mM internally; during sporulation the concentrations become much higher.[593]

F. Chromium

Animals deficient in chromium grow poorly and have a reduced life span.[594–596] They also have decreased "glucose tolerance," i.e., glucose injected into the blood stream is removed only half as fast as it is normally.[597,598] This is similar to the effect of a deficiency of insulin. Fractionation of yeast led to the isolation of a chromium-containing **glucose tolerance factor** which appeared to be a complex of Cr^{3+}, nicotinic acid, and amino acids.[597] The chromium in this material is apparently well absorbed by the body but is probably not an essential cofactor.[596] Nevertheless, dietary supplementation with chromium appears to improve glucose utilization, apparently by enhancing the action of insulin.[596] Ingestion of glucose not only increases insulin levels in blood but also causes increased urinary loss of chromium,[599] perhaps as a result of insulin-induced mobilization of stored chromium.[596] It has been suggested that a specific **chromium-binding oligopeptide** isolated from mammalian liver[596,600] may be released in response to insulin and may activate a membrane phosphotyrosine phosphatase.[596]

Chromium concentrations in animal tissues are usually less than 2 µM but tend to be much higher in the caudate nucleus of the brain. High concentrations of Cr^{3+} have also been found in RNA–protein complexes.[601] While several oxidation states, including +2, +3, and +6, are known for chromium, only Cr(III) is found to a significant extent in tissues. The Cr(VI) complex ions, chromate and dichromate, are toxic and chronic exposure to chromate-containing dust can lead to lung cancer. Ascorbate is a principal biological reductant of chromate and can create mutagenic Cr(V) compounds that include a Cr(V)–ascorbate–peroxo complex.[602] However, Cr(III) compounds administered orally are not significantly toxic. Evidently, the Cr(VI) compounds can cross cell membranes and be reduced to Cr(III), which forms stable complexes with many constituents of cells including DNA.[603] The use of such "exchange-inert" Cr(III) complexes of ATP in enzymology was considered in Chapter 12.

Most forms of Cr(III) are not absorbed and utilized by the body. For this reason, and because of the increased use of sucrose and other refined foods, a marginal human chromium deficiency may be widespread.[604,605] This may result not only in poor utilization of glucose but also in other effects on lipid and protein metabolism.[597] However, questions have been raised about the use of chromium picolinate as a dietary supplement. High concentrations have been reported to cause chromosome damage[606] and there may be danger of excessive accumulation of chromium in the body.[607]

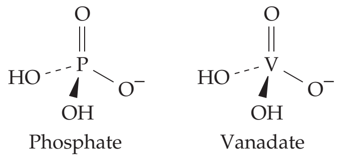
G. Vanadium

Vanadium is a dietary essential for goats and presumably also for human beings,[427] who typically consume ~2 mg / day. However, because vanadium compounds have powerful pharmacological effects it has been difficult to establish the nutritional requirement for animals.[427,604,607a] The adult body contains only about 0.1 mg of vanadium. Typical tissue concentrations are 0.1–0.7 µM[608,609] and serum concentration may be 10 nM or less. Vanadium can assume oxidation states ranging from +2 to +5, the vanadate ion VO_4^{3-} being the predominant form of V(V) in basic solution and in dilute solutions at pH 7. However, at millimolar concentrations, $V_3O_9^{3-}$, $V_4O_{12}^{4-}$, and other polynuclear forms predominate.[610–612] In plasma most exists as metavanadate, VO_2^-, but within cells it is reducted to the vanadyl cation VO^{2+} which is an especially stable double-bonded unit in compounds of V(IV).[609] Only at very low pH is V(III) stable.

The first suggestion of a possible biochemical function for vanadium came from the discovery that **vanadocytes**, the green blood cells of tunicates (sea squirts), contain ~1.0 M V(III) and 1.5–2 M H_2SO_4 .[613] It was proposed that a V-containing protein is an oxygen carrier. However, the V^{3+} appears not to be associated with proteins[612] and it does not carry O_2. It may be there to poison predators.[614] The vanadium-accumulating species also synthesize several complex, yellow catechol-type chelating agents (somewhat similar to enterobactin; Fig. 16-1) which presumably complex V(V) and perhaps also reduce it to V(III).[615] Vanadium is also accumulated by other marine organisms and by the mushroom *Amanita muscaria*.

Vanadoproteins are found in most marine algae and seaweed and in some lichens.[616] Among these are **haloperoxidases**,[252,253,617–618b] enzymes that are quite different from the corresponding heme peroxidases discussed in Section A,6. The vanadium is bound as hydrogen vanadate, HVO_4^{2-}, in trigonal bipyrimidal coordination with the three oxygens in equatorial positions and a histidine in one axial position. In the crystal structure an azide (N_3^-) ion occupies the other axial position, but it is presumably the site of interaction with peroxide.[619] The structure is similar to that of acid phosphatases inhibited by vanadate.[620,621] Many nitrogen-fixing bacteria contain genes for a vanadium-dependent nitrogenase that is formed only if molybdenum is not available.[622] The nitrogenases are discussed in Chapter 24.

Much of current interest in vanadium stems from the discovery that vanadate (HVO_4^{2-} at pH 7) is a powerful inhibitor of ATPases such as the sodium pump protein ($Na^+ + K^+$)ATPase (Chapter 8), of phosphatases,[623] and of kinases.[624] This can be readily understood from comparison of the structure of phosphate and vanadate ions.

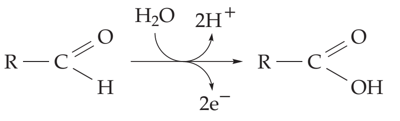


Other enzymes such as the cyclic AMP-dependent protein kinase are *stimulated* by vanadium.[624] Vanadate seems to inhibit most strongly those enzymes that form a phosphoenzyme intermediate. This inhibition may be diminished within cells because vanadate is readily reduced by glutathione and other intracellular reductants. The resulting vanadyl ion is a much weaker inhibitor and also stimulates several metabolic processes.[608]

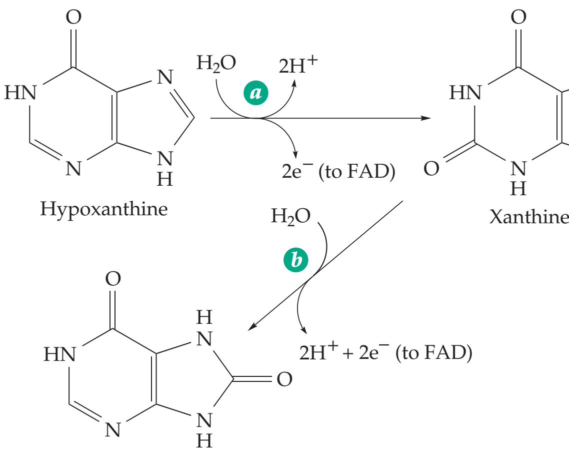
Also of great interest is an insulin-like action of vanadium[607a,625] and evidence that vanadium may be essential to proper cardiac function.[626] A role in lipid metabolism was suggested by the observation that in high doses vanadium inhibits cholesterol synthesis and lowers the phospholipid and cholesterol content of blood. Vanadium is reported to inhibit development of caries by stimulating mineralization of teeth. Unlike tungsten, vanadium does not compete with molybdenum in the animal body.[627] The sometimes dramatic effects of vanadate as an inhibitor, activator, and metabolic regulator are shared also by molybdate and tungstate.[628,629] Even greater effects are observed with vanadate, molybdate, or tungstate plus H_2O_2.[630] The resulting **pervanadate**, **permolybdate**, and **pertungstate** are often assumed to be monoperoxo compounds, e.g., vanodyl hydroperoxide. However, there is some uncertainty.[631]

H. Molybdenum

Long recognized as an essential element for the growth of plants, molybdenum has never been directly demonstrated as a necessary animal nutrient. Nevertheless, it is found in several enzymes of the human body, as well as in 30 or more additional enzymes of bacteria and plants.[632] **Aldehyde oxidases**,[633] **xanthine oxidase** of liver and the related **xanthine dehydrogenase**, catalyze the reactions of Eqs. 16-58 and 16-59 and contain molybdenum that is essential for catalytic activity. Xanthine oxidase also contains two Fe_2S_2 clusters and bound FAD. The enzymes can also oxidize xanthine further (Eq. 16-59, step *b*) by a repetition of the same type of oxidation process at positions 8 and 9 to form **uric acid**. The much studied xanthine dehydrogenase has been isolated from milk,[634,635] liver, fungi,[636] and some bacteria.[637] In the dehydrogenase NAD^+ is the electron acceptor that oxidizes the bound $FADH_2$ formed in Eq. 16-59. Xanthine dehydrogenase, in the absence of thiol compounds, is converted spontaneously into xanthine oxidase, probably as a result of a conformational change and formation of a disulfide bridge within the protein. Treatment with thiol compounds such as dithiothreitol reconverts the enzyme to the dehydrogenase. Evidently in the oxidase form the NAD^+ binding site has moved away from the FAD, permitting oxidation of $FADH_2$ by O_2 with formation of hydrogen peroxide.[635,638]



(16-58)



(16-59)

A purine hydroxylase from fungi,[639] bacterial quinoline and isoquinoline oxidoreductases,[640,641] and a selenium-containing nicotinic acid hydroxylase from *Clostridium barberei*[642] are members of the **xanthine oxidase family** (or molybdenum hydroxylase family).[632,641,643] Also included in the family are aldehyde oxidoreductases from the sulfate-reducing *Desulfovibrio gigas*[633] and from the tomato.[644]

Two other families of molybdoenzymes are the **sulfite oxidase family**[646a,b] and the **dimethylsulfoxide reductase family**.[632,641] **Nitrogenase** (Chapter 24) constitutes a fourth family. Sulfite oxidase (Eq. 16-60) is an essential human liver enzyme (see also Chapter 24).[645,646]

$$SO_3^{2-} + H_2O \rightarrow SO_4^{2-} + 2\ e^- + 2\ H^+ \qquad (16\text{-}60)$$

The **assimilatory nitrate reductase** (Eq. 16-61) of fungi and green plants (Chapter 24) also belongs to the sulfite oxidase family.

$$NO_3^{2}\ + 2\ e^- + 2\ H^+ \rightarrow NO_2^- + H_2O \qquad (16\text{-}61)$$

DMSO reductase reduces dimethylsulfoxide to dimethylsulfide (Eq. 16-62) as part of the biological sulfur cycle.[647–648d]

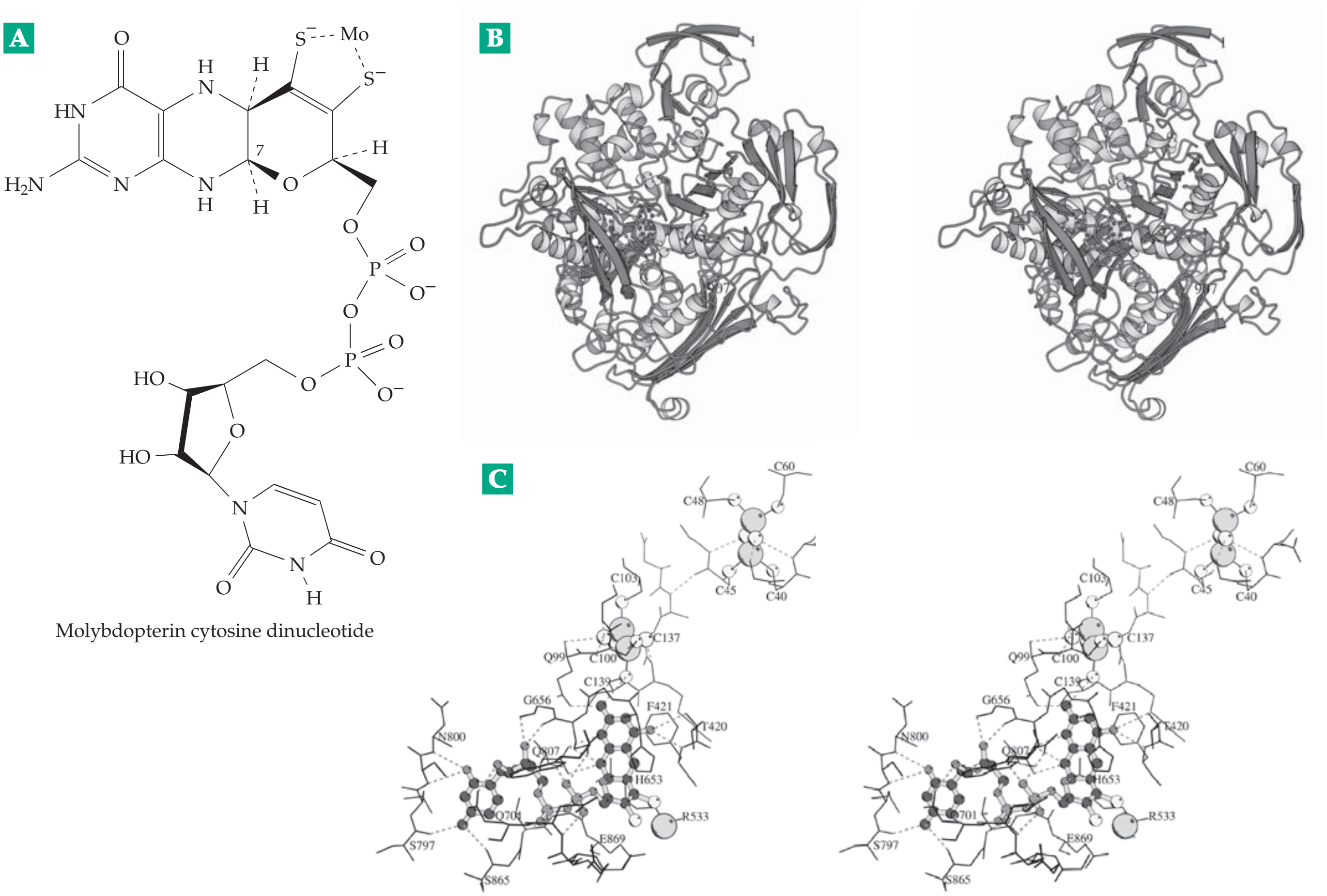
$$(H_3C)_2S{=}O + 2\ H^+ + 2\ e^- \rightarrow (CH_3)_2S \qquad (16\text{-}62)$$

A number of other reductases and dehydrogenases, including **dissimilatory nitrate reductases** of *E. coli* and of denitrifying bacteria (Chapter 18), belong to the DMSO reductase family. Other members are reductases for biotin *S*-oxide,[649] trimethylamine *N*-oxide, and polysulfides as well as **formate dehydrogenases** (Eq. 16-63), formylmethanofuran dehydrogenase (Fig. 15-22, step *b*), and arsenite oxidase.[632] Several other molybdoenzymes, such as pyridoxal oxidase, had not been classified by 1996.[632]

$$HCOO^- + 2\ e^- + 3\ H^+ \rightarrow CO_2 + 2\ H_2O \qquad (16\text{-}63)$$

1. Molybdenum Ions and Coenzyme Forms

Molybdenum is a metal of the second transition series, one of the few heavy elements known to be essential to life. Its most stable oxidation state, Mo(VI), has 4*d* orbitals available for coordination with anionic ligands. Coordination numbers of 4 and 6 are preferred, but molybdenum can accommodate up to eight ligands. Most of the complexes are formed from the oxycation $Mo(VI)O_2^{2+}$. If two molecules of water are coordinated with this ion, the protons are so acidic that they dissociate completely to give $Mo(VI)O_4^{2-}$, the molybdate ion. Other oxidation states vary from Mo(III) to Mo(V). In these lower oxidation states, the tendency for protons to dissociate from coordinated ligands is less, e.g., $Mo(III)(H_2O)_6^{3+}$ does not lose protons even in a very basic medium. Molybdenum tends to form dimeric



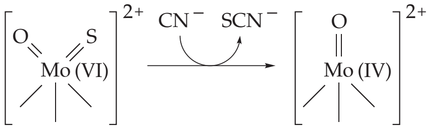
Molybdopterin cytosine dinucleotide

Figure 16-31 (A) Structure of molybdopterin cytosine dinucleotide complexed with an atom of molybdenum. (B) Stereoscopic ribbon drawing of the structure of one subunit of the xanthine oxidase-related aldehyde oxidoreductase from *Desulfovibrio gigas*. Each 907-residue subunit of the homodimeric protein contains two Fe_2S_2 clusters visible at the top and the molybdenum–molybdopterin coenzyme buried in the center. (C) Alpha-carbon plot of portions of the protein surrounding the molybdenum–molybdopterin cytosine dinucleotide and (at the top) the two plant-ferredoxin-like Fe_2S_2 clusters. Each of these is held by a separate structural domain of the protein. Two additional domains bind the molybdopterin coenzyme and there is also an intermediate connecting domain. In xanthine oxidase the latter presumably has the FAD binding site which is lacking in the *D. gigas* enzyme. From Romão *et al.*[633] Courtesy of R. Huber.

or polymeric oxygen-bridged ions. However, within the enzymes it exists as the unique **molybdenum coenzymes.** The Mo-containing enzymes usually also contain additional bound cofactors, including Fe–S clusters and flavin coenzymes or heme.

The recognition that the Mo in the molybdoproteins exists in organic cofactor forms came from studies of mutants of *Aspergillus* and *Neurospora*.[650] In 1964, Pateman and associates discovered mutants that lacked both nitrate reductase and xanthine dehydrogenase. Later, it was shown that acid-treated molybdoenzymes released a material that would restore activity to the inactived nitrate reductase from the mutant organisms. This new coenzyme, a phosphate ester of molybdopterin (Fig. 15-17), was characterized by Rajagopalan and coworkers.[650,651] A more complex form of the coenzyme, **molybdopterin cytosine dinucleotide** (Fig. 16-31), is found in the *D. gigas* aldehyde oxidoreductase. Related coenzyme forms include nucleotides of adenine, guanine (see chapter banner, p. 837), and hypoxanthine.[651a,651b] The structure of molybdopterin is related to that of **urothione** (Fig. 15-17), a normal urinary constituent. The relationship to urothione was strengthened by the fact that several children with severe neurological and other symptoms were found to lack both sulfite oxidase and xanthine dehydrogenase as well as the molybdenum cofactor and urinary urothione.[646,646a,646b]

Study by X-ray absorption spectroscopy of the extended **X-ray absorption fine structure** (EXAFS) has provided estimates of both the nature and the number of the nearest neighboring atoms around the Mo. The EXAFS spectra of xanthine dehydrogenase and of nitrate reductase from **Chlorella** confirmed the presence of both the $Mo(VI)O_2$ unit with Mo–O distances of 0.17 nm and two or three sulfur atoms at distances of 0.24 nm.[652,653] The two sulfur atoms were presumed to come from the molybdopterin. A peculiarity of the xanthine oxidase family is the presence on the molybdenum of a "cyanolyzable" sulfur.[654] This is a sulfide attached to the molybdenum, which is present as Mo(VI)OS rather than $Mo(VI)O_2$. Reaction with cyanide produces thiocyanate (Eq. 16-64).



(16-64)

The active site structures of the three classes of molybdenum-containing enzymes are compared in Fig. 16-32. In the DMSO reductase family there are two identical molybdopterin dinucleotide coenzymes complexed with one molybdenum. However, only one of these appears to be functionally linked to the Fe_2S_2 center.

Nitrogenase, which catalyzes the reduction of N_2 to two molecules of NH_3, has a different **molybdenum–iron cofactor (FeMo-co)**. It can be obtained by acid denaturation of the very oxygen-labile iron–molybdenum protein of nitrogenase followed by extraction with dimethylformamide.[655,656] The coenzyme is a complex Fe–S–Mo cluster also containing **homocitrate** with a composition $MoFe_7S_9$–homocitrate (see Fig. 24-3). Nitrogenase and this coenzyme are considered further in Chapter 24.

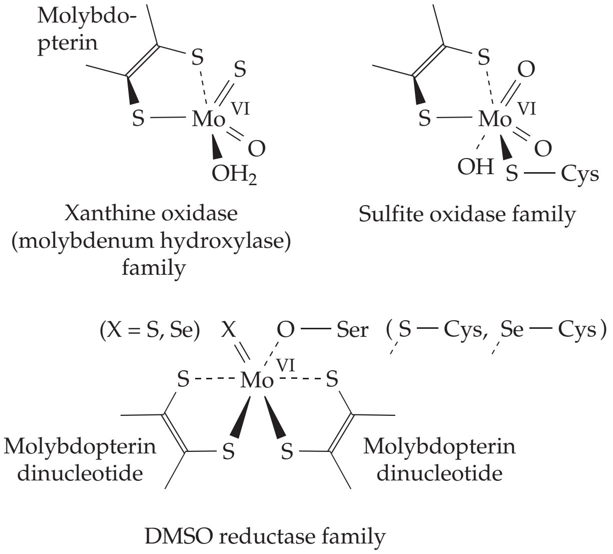


Figure 16-32 Structures surrounding molybdenum in three families of molybdoenzymes. See Hille.[632]

2. Enzymatic Mechanisms

Although several of the reactions catalyzed by molybdoenzymes are classified as dehydrogenases, all of them except nitrogenase involve H_2O as either a reactant or a product. The EXAFS spectra suggest that the $Mo(VI)O_2$ unit is converted to Mo(IV)O during reaction with a substrate Sub (Eq. 16-65, step *a*). Reaction of the Mo(IV)O with water (step *b*) completes the catalysis.

a

$$[\mathrm{Mo(VI)O_2}]^{2+} + \mathrm{Sub} \rightarrow [\mathrm{Mo(IV)O}]^{2+} + \mathrm{Sub–O}$$

b

$$[\mathrm{Mo(IV)O}]^{2+} + \mathrm{H_2O} \rightarrow 2\mathrm{H}^+ + 2e^- + [\mathrm{Mo(VI)}]^{2+}\mathrm{O_2}$$
$$\mathrm{SumSub} + \mathrm{H_2O} \rightarrow \mathrm{Sub–O} + 2\mathrm{H}^+ + 2e^- \quad (16\text{-}65)$$

Step *a* of all of these reactions can be regarded as an **oxo-transfer**.[653]

To complete the reaction, two electrons must be passed from Mo(IV) to a suitable acceptor, usually an Fe–S cluster or a bound heme group. FAD is also often present. Xanthine oxidase[634,635,643,657,657a] contains two Fe_2S_2 clusters and a FAD for each of the two atoms of Mo in the dimer. Since this enzyme acts like a typical flavin oxidase that generates H_2O_2 from O_2, it may be that electrons pass from Mo to the Fe–S center and then to the flavin. Since the EPR signal of the paramagnetic Mo(V), with its characteristic six-line hyperfine structure, is seen during the action of xanthine oxidase and other molybdenum-containing enzymes, single-electron transfers are probably involved.

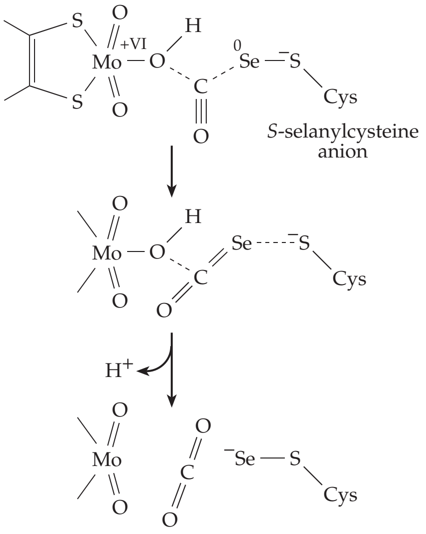
In bacteria such as *E. coli* a dissimilatory nitrate reductase allows nitrate to serve as an oxidant in place of O_2. An oxygen atom is removed from the nitrate to form nitrite as two electrons are accepted from a membrane-bound cytochrome *b*. The nitrate reductase consists of a 139-kDa Mo-containing catalytic subunit, a 58-kDa electron-transferring subunit that contains both Fe_3S_4 and Fe_4S_4 centers, and a 26-kDa heme-containing membrane anchor subunit.[658–660] The assimilatory nitrate reductase of fungi, green algae, and higher plants contains both a *b*-type cytochrome and FAD and a molybdenum coenzyme in a large oligomeric complex.[661–663a]

Formate dehydrogenases from many bacteria contain molybdopterin and also often selenium (Table 15-4).[664,665] A membrane-bound Mo-containing formate dehydrogenase is produced by *E. coli* grown anaerobically in the presence of nitrate. Under these circumstances it is coupled to nitrate reductase via an electron-transport chain in the membranes which permits oxidation of formate by nitrate (Eq. 18-26). This enzyme is also a multisubunit protein.[665,666] Two other Mo- and Se- containing formate dehydrogenases are produced by *E. coli*.[667,668] The three-dimensional structure is known for one of them, **formate dehydrogenase H**, a component of the anaerobic formate hydrogen lyase complex (Eq. 17-25).[669,670] The structure shows Mo held by the sulfur atoms of two molybdopterin molecules, as in DMSO reductase. The Se atom of SeCys 140 is also coordinated with the Mo atom, and the imidazole of His 141 is in close proximity. When ^{13}C-labeled formate was oxidized in ^{18}O-enriched water no ^{18}O was found in the released product, CO_2.[671] This suggested that formate may be bound to Mo and dehydrogenated, with Mo(VI) being reduced to Mo(IV). The formate hydrogen might be transferred as H^+ to the His 140 side chain. Mo(IV) could then be reoxidized by electron transfer in two one-electron steps.[670] However, recent X-ray absorption spectra suggest the presence in the enzyme of a selenosulfide ligand to Mo.[672] Mechanistic uncertainties remain!

A flavin-dependent formate dehydrogenase system found in *Methanobacterium* passes electrons from dehydrogenation of formate to FAD and then to the deazaflavin coenzyme F_{420}.[673] In contrast to these Mo-containing enzymes, the formate dehydrogenase from *Pseudomonas oxalaticus*, which oxidizes formate with NAD^+ (Eq. 16-66), contains neither Mo or Se.[674]

$$\mathrm{HCOO^- + NAD^+ \longrightarrow CO_2 + NADH} \qquad (16\text{-}66)$$

It is a large 315-kDa oligomer containing 2 FMN and ~20 Fe / S. Formate dehydrogenases of green plants and yeasts are smaller 70- to 80-kDa proteins lacking bound prosthetic groups.[674] A key enzyme in the metabolism of carbon monoxide-oxidizing bacteria is CO oxidase, another membrane-bound molybdoenzyme.[675–676c] It also contains selenium, which is attached to a cysteine side chain as **S-selanylcysteine.** A proposed reaction requence[676a] is shown in Eq. 16-67.



(16-67)

3. Nutritional Need for Mo

The first hint of an essential role of molybdenum in metabolism came from the discovery that animals raised on a diet deficient in molybdenum had decreased liver xanthine oxidase activity. There is no evidence that xanthine oxidase is essential for all life, but a human genetic deficiency of sulfite oxidase or of its molybdopterin coenzyme can be lethal.[646,646a,b] The conversion of molybdate into the molybdopterin cofactor in *E. coli* depends upon at least five genes.[677] In *Drosophila* the addition of the cyanolyzable sulfur (Eq. 16-64) is the final step in formation of xanthine dehydrogenase.[678] It is of interest that sulfur (S^0) can be transferred from rhodanese (see Eq. 24-45), or from a related mercaptopyruvate sulfurtransferase[679] into the desulfo form of xanthine oxidase to generate an active enzyme.[680]

Uptake of molybdate by cells of *E. coli* is accomplished by an ABC-type transport system.[681] In some bacteria, e.g., the nitrogen-fixing *Azotobacter*, molybdenum can be stored in protein-bound forms.[682]

I. Tungsten

For many years tungsten was considered only as a potential antagonist for molybdenum. However, in 1970 growth stimulation by tungsten compounds was observed for some acetogens, some methanogens, and a few hyperthermophilic bacteria. Since then over a dozen tungstoenzymes have been isolated.[683,683a] These can be classified into three categories: aldehyde oxidoreductases, formaldehyde oxidoreductases, and the single enzyme acetylene hydratase. In most cases the tungstoenzymes resemble the corresponding molybdoenzymes and in most instances organisms containing a tungsten-requiring enzyme also contain the corresponding molybdenum enzyme. However, a few hyperthermophilic archaea appear to require W and are unable to use Mo.

The aldehyde ferredoxin oxidoreductase from the hyperthermophile *Pyrococcus furiosus* was the first molybdopterin-dependent enzyme for which a three-dimensional structure became available.[683,684] The tungstoenzyme resembles that of the related molybdoenzyme (Fig. 16-31). A similar ferredoxin-dependent enzyme reduces glyceraldehyde-3-phosphate.[685] Another member of the tungstoenzyme aldehyde oxidoreductase family is **carboxylic acid reductase**, an enzyme found in certain acetogenic clostridia. It is able to use reduced ferredoxin to convert unactivated carboxylic acids into aldehydes, even though $E^{o'}$ for the acetaldehyde / acetate couple is –0.58 V.[686]

Tungsten- and sometimes Se-containing formate dehydrogenases together with *N*-formylmethanofuran dehydrogenases (Fig. 15-22, step *b*) form a second family.

Again, these appear to resemble the corresponding Mo-dependent enzymes. The unique **acetylene hydratase** from the acetylene-utilizing *Pelobacter acetylenicus* catalyzes the hydration of acetylene to acetaldehyde.[687]

$$H{-}C{\equiv}C{-}H + H_2O \rightarrow H_3C{-}CHO \tag{16-68}$$

In *Thermotoga maritima*, the most thermophilic organism known, tungsten promotes synthesis of an Fe-containing hydrogenase as well as some other enzymes but seems to have a regulatory rather than a structural role.[688]

References

1. Ochiai, E. I. (1977) *Bioinorganic Chemistry. An Introduction*, Allyn and Bacon, Boston
2. Fraústo da Silva, J. J. R., and Williams, R. J. P. (1991) *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*, Clarendon Press, Oxford
3. Lippard, S. J., and Berg, J. M. (1994) *Principles of Bioinorganic Chemistry*, Univ. Science Books, Mill Valley, California
4. Holm, R. H., Kennepohl, P., and Solomon, E. I. (1996) *Chem. Rev.* **96**, 2239 – 2314
5. Neilands, J. B., ed. (1974) *Microbial Iron Metabolism*, Academic Press, New York
6. Jacobs, A., and Woodwood, M., eds. (1974) *Iron in Biochemistry and Medicine*, Academic Press, New York
7. Neilands, J. B. (1973) in *Inorganic Biochemistry*, Vol. 1 (Eichhorn, G. L., ed), pp. 167 – 202, Elsevier, Amsterdam
8. Bergeron, R. J. (1986) *Trends Biochem. Sci.* **11**, 133 – 136
9. Neilands, J. B. (1995) *J. Biol. Chem.* **270**, 26723 – 26726
10. Sakaitani, M., Rusnak, F., Quinn, N. R., Tu, C., Frigo, T. B., Berchtold, G. A., and Walsh, C. T. (1990) *Biochemistry* **29**, 6789 – 6798
11. Harris, W. R., Carrano, C. J., Cooper, S. R., Sofen, S. R., Avdeef, A. E., McArdle, J. V., and Raymond, K. N. (1979) *J. Am. Chem. Soc.* **101**, 6097 – 6104

11a. Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K., and Welte, W. (1998) *Science* **282**, 2215 – 2220

11b. Braun, V., and Killmann, H. (1999) *Trends Biochem. Sci.* **24**, 104 – 109

11c. Stintzi, A., Barnes, C., Xu, J., and Raymond, K. N. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10691 – 10696

11d. Cohen, S. M., Meyer, M., and Raymond, K. N. (1998) *J. Am. Chem. Soc.* **120**, 6277 – 6286

11e. Adrait, A., Jacquamet, L., Le Pape, L., Gonzalez de Peredo, A., Aberdam, D., Hazemann, J.-L., Latour, J.-M., and Michaud-Soret, I. (1999) *Biochemistry* **38**, 6248 – 6260

12. Neilands, J. B., Erickson, T. J., and Rastetter, W. H. (1981) *J. Biol. Chem.* **256**, 3831 – 3832
13. Newton, S. M. C., Allen, J. S., Cao, Z., Qi, Z., Jiang, X., Sprencel, C., Igo, J. D., Foster, S. B., Payne, M. A., and Klebba, P. E. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4560 – 4565
14. Persmark, M., Expert, D., and Neilands, J. B. (1989) *J. Biol. Chem.* **264**, 3187 – 3193
15. Khalil-Rizvi, S., Toth, S. I., van der Helm, D., Vidavsky, I., and Gross, M. L. (1997) *Biochemistry* **36**, 4163 – 4171
16. Wong, G. B., Kappel, M. J., Raymond, K. N., Matzanke, B., and Winkelmann, G. (1983) *J. Am. Chem. Soc.* **105**, 810 – 815
17. Shanzer, A., Libman, J., Lifson, S., and Felder, C. E. (1986) *J. Am. Chem. Soc.* **108**, 7609 – 7619
18. Reid, R. T., Live, D. H., Faulkner, D. J., and Butler, A. (1993) *Nature (London)* **366**, 455 – 458
19. Llinás, M., Wilson, D. M., and Neilands, J. B. (1973) *Biochemistry* **12**, 3836 – 3843
20. Eng-Wilmot, D. L., and van der Helm, D. (1980) *J. Am. Chem. Soc.* **102**, 7719 – 7725
21. Cox, C. D., Rinehart, K. L., Jr., Moore, M. L., and Cook, J. C., Jr. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4256 – 4260
22. Neilands, J. B. (1952) *J. Am. Chem. Soc.* **74**, 4846 – 4847
23. Braun, V. (1985) *Trends Biochem. Sci.* **10**, 75 – 78
24. Letendre, E. D. (1985) *Trends Biochem. Sci.* **10**, 166 – 168
25. Köster, W., and Braun, V. (1990) *J. Biol. Chem.* **265**, 21407 – 21410
26. Jiang, X., Payne, M. A., Cao, Z., Foster, S. B., Feix, J. B., Newton, S. M. C., and Klebba, P. E. (1997) *Science* **276**, 1261 – 1264
27. Coy, M., and Neilands, J. B. (1991) *Biochemistry* **30**, 8201 – 8210
28. Kammler, M., Schoin, C., and Hantke, D. (1993) *J. Bacteriol.* **175**, 6212 – 6219
29. Dix, D. R., Bridgham, J. T., Broderius, M. A., Byersdorfer, C. A., and Eide, D. J. (1994) *J. Biol. Chem.* **269**, 26092 – 26099
30. Kaplan, J., and O’Halloran, T. V. (1996) *Science* **271**, 1510 – 1512
31. Dix, D., Bridgham, J., Broderius, M., and Eide, D. (1997) *J. Biol. Chem.* **272**, 11770 – 11777
32. de Silva, D., Davis-Kaplan, S., Fergestad, J., and Kaplan, J. (1997) *J. Biol. Chem.* **272**, 14208 – 14213
33. Radisky, D., and Kaplan, J. (1999) *J. Biol. Chem.* **274**, 4481 – 4484
34. Stearman, R., Yuan, D. S., Yamaguchi-Iwai, Y., Klausner, R. D., and Dancis, A. (1996) *Science* **271**, 1552 – 1557
35. Hassett, R. F., Romeo, A. M., and Kosman, D. J. (1998) *J. Biol. Chem.* **273**, 7628 – 7636
36. Mukhopadhyay, C., Attieh, Z. K., and Fox, P. L. (1998) *Science* **279**, 714 – 717
37. Li, L., and Kaplan, J. (1997) *J. Biol. Chem.* **272**, 28485 – 28493

37a. Lange, H., Kispal, G., and Lill, R. (1999) *J. Biol. Chem.* **274**, 18989 – 18996

37b. Urbanowski, J. L., and Piper, R. C. (1999) *J. Biol. Chem.* **274**, 38061 – 38070

38. Octave, J.-N., Schneider, Y.-J., Trouet, A., and Crichton, R. R. (1983) *Trends Biochem. Sci.* **8**, 217 – 222
39. Conrad, M. E., Umbreit, J. N., Moore, E. G., Peterson, R. D. A., and Jones, M. B. (1990) *J. Biol. Chem.* **265**, 5273 – 5279
40. Ponka, P., Schulman, H. M., Woodworth, R. C., and Richter, G. W., eds. (1990) *Iron Transport and Storage*, CRC Press, Boca Raton, Florida
41. Yu, J., and Wessling-Resnick, M. (1998) *J. Biol. Chem.* **273**, 6909 – 6915
42. Gunshin, H., Mackenzie, B., Berger, U. V., Gunshin, Y., Romero, M. F., Boron, W. F., Nussberger, S., Gollan, J. L., and Hediger, M. A. (1997) *Nature (London)* **388**, 482 – 488
43. Toth, I., Rogers, J. T., McPhee, J. A., Elliott, S. M., Abramson, S. L., and Bridges, K. R. (1995) *J. Biol. Chem.* **270**, 2846 – 2852
44. Finch, C. A., and Huebers, H. (1982) *N. Engl. J. Med.* **306**, 1520 – 1528
45. Scrimshaw, N. S. (1991) *Sci. Am.* **265**(Oct), 46 – 52
46. Kaplan, J., and Kushner, J. P. (2000) *Nature (London)* **403**, 711,713

46a. Kühn, L. C. (1999) *Trends Biochem. Sci.* **24**, 164 – 166

46b. Attieh, Z. K., Mukhopadhyay, C. K., Seshadri, V., Tripoulas, N. A., and Fox, P. L. (1999) *J. Biol. Chem.* **274**, 1116 – 1123

47. Gordeuk, V., Thuma, P., Brittenham, G., McLaren, C., Parry, D., Backenstose, A., Biemba, G., Msiska, R., Holmes, L., McKinley, E., Vargas, L., Gilkeson, R., and Poltera, A. A. (1992) *N. Engl. J. Med.* **327**, 1518 – 1521

48. Faller, B., and Nick, H. (1994) *J. Am. Chem. Soc.* **116**, 3860–3865

49. Baker, E. N., Rumball, S. V., and Anderson, B. F. (1987) *Trends Biochem. Sci.* **12**, 350–353

50. Welch, S. (1992) *Transferrin: The Iron Carrier*, CRC Press, Boca Raton, Florida

51. Baker, H. M., Anderson, B. F., Brodie, A. M., Shongwe, M. S., Smith, C. A., and Baker, E. N. (1996) *Biochemistry* **35**, 9007–9013

51a. Sharma, A. K., Paramasivam, M., Srinivasan, A., Yadav, M. P., and Singh, T. P. (1998) *J. Mol. Biol.* **289**, 303–317

51b. Peterson, N. A., Anderson, B. F., Jameson, G. B., Tweedie, J. W., and Baker, E. N. (2000) *Biochemistry* **39**, 6625–6633

51c. Bou Abdallah, F., and El Hage Chahine, J.-M. (2000) *J. Mol. Biol.* **303**, 255–266

52. Kurokawa, H., Mikami, B., and Hirose, M. (1995) *J. Mol. Biol.* **254**, 196–207

52a. Mizutani, K., Yamashita, H., Mikami, B., and Hirose, M. (2000) *Biochemistry* **39**, 3258–3265

53. Moore, S. A., Anderson, B. F., Groom, C. R., Haridas, M., and Baker, E. N. (1997) *J. Mol. Biol.* **274**, 222–236

54. Bailey, S., Evans, R. W., Garratt, R. C., Gorinsky, B., Hasnain, S., Horsburgh, C., Jhoti, H., Lindley, P. F., Mydin, A., Sarra, R., and Watson, J. L. (1988) *Biochemistry* **27**, 5804–5812

54a. He, Q.-Y., Mason, A. B., Tam, B. M., MacGillivray, R. T. A., and Woodworth, R. C. (1999) *Biochemistry* **38**, 9704–9711

55. Randell, E. W., Parkes, J. G., Olivieri, N. F., and Templeton, D. M. (1994) *J. Biol. Chem.* **269**, 16046–16053

56. Grossmann, J. G., Mason, A. B., Woodworth, R. C., Neu, M., Lindley, P. F., and Hasnain, S. S. (1993) *J. Mol. Biol.* **231**, 554–558

57. Zak, O., Tam, B., MacGillivray, R. T. A., and Aisen, P. (1997) *Biochemistry* **36**, 11036–11043

58. Zak, O., Aisen, P., Crawley, J. B., Joannou, C. L., Patel, K. J., Rafiq, M., and Evans, R. W. (1995) *Biochemistry* **34**, 14428–14434

59. Mecklenburg, S. L., Donohoe, R. J., and Olah, G. A. (1997) *J. Mol. Biol.* **270**, 739–750

60. Fleming, M. D., Romano, M. A., Su, M. A., Garrick, L. M., Garrick, M. D., and Andrews, N. C. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1148–1153

61. Theil, E. C. (1987) *Ann. Rev. Biochem.* **56**, 289–315

62. Crichton, R. R. (1984) *Trends Biochem. Sci.* **9**, 283–286

63. Trikha, J., Theil, E. C., and Allewell, N. M. (1995) *J. Mol. Biol.* **248**, 949–967

64. Hempstead, P. D., Yewdall, S. J., Fernie, A. R., Lawson, D. M., Artymiuk, P. J., Rice, D. W., Ford, G. C., and Harrison, P. M. (1997) *J. Mol. Biol.* **268**, 424–448

65. Crichton, R. R., Soruco, J.-A., Roland, F., Michaux, M.-A., Gallois, B., Précigoux, G., Mahy, J.-P., and Mansuy, D. (1997) *Biochemistry* **36**, 15049–15054

66. Watt, G. D., Frankel, R. B., Papaefthymiou, G. C., Spartalian, K., and Stiefel, E. I. (1986) *Biochemistry* **25**, 4330–4336

67. Garg, R. P., Vargo, C. J., Cui, X., and Kurtz, D. M., Jr. (1996) *Biochemistry* **35**, 6297–6301

67a. Romão, C. V., Regalla, M., Xavier, A. V., Teixeira, M., Liu, M.-Y., and Le Gall, J. (2000) *Biochemistry* **39**, 6841–6849

68. Dickey, L. F., Sreedharan, S., Theil, E. C., Didsbury, J. R., Wang, H.-H., and Kaufman, R. E. (1987) *J. Biol. Chem.* **262**, 7901–7907

69. Sayers, D. E., Theil, E. C., and Rennick, F. J. (1983) *J. Biol. Chem.* **258**, 14076–14079

69a. Johnson, J. L., Cannon, M., Watt, R. K., Frankel, R. B., and Watt, G. D. (1999) *Biochemistry* **38**, 6706–6713

70. Bauminger, E. R., Treffry, A., Quail, M. A., Zhao, Z., Nowik, I., and Harrison, P. M. (1999) *Biochemistry* **38**, 7791–7802

71. Yang, X., Chen-Barrett, Y., Arosio, P., and Chasteen, N. D. (1998) *Biochemistry* **37**, 9743–9750

72. Pereira, A. S., Small, W., Krebs, C., Tavares, P., Edmondson, D. E., Theil, E. C., and Huynh, B. H. (1998) *Biochemistry* **37**, 9871–9876

72a. Hwang, J., Krebs, C., Huynh, B. H., Edmondson, D. E., Theil, E. C., and Penner-Hahn, J. E. (2000) *Science* **287**, 122–125

73. Treffry, A., Zhao, Z., Quail, M. A., Guest, J. R., and Harrison, P. M. (1997) *Biochemistry* **36**, 432–441

74. Bourne, P. E., Harrison, P. M., Rice, D. W., Smith, J. M. A., and Stansfield, R. F. D. (1982) in *The Biochemistry and Physiology of Iron* (Saltman, P., and Hegenauer, J., eds), pp. 427–, Elsevier North Holland, Amsterdam

75. Bridges, K. R., and Hoffman, K. E. (1986) *J. Biol. Chem.* **261**, 14273–14277

76. Medlock, A. E., and Dailey, H. A. (2000) *Biochemistry* **39**, 7461–7467

76a. Franco, R., Ma, J.-G., Lu, Y., Ferreira, G. C., and Shelnutt, J. A. (2000) *Biochemistry* **39**, 2517–2529

76b. Lecerof, D., Fodje, M., Hansson, A., Hansson, M., and Al-Karadaghi, S. (2000) *J. Mol. Biol.* **297**, 221–232

77. Eaton, S. S., and Eaton, G. R. (1977) *J. Am. Chem. Soc.* **99**, 1601–1604

78. Crossley, M. J., Harding, M. M., and Sternhell, S. (1986) *J. Am. Chem. Soc.* **108**, 3608–3613

78a. Braun, J., Schwesinger, R., Williams, P. G., Morimoto, H., Wemmer, D. E., and Limbach, H.-H. (1996) *J. Am. Chem. Soc.* **118**, 11101–11110

78b. Treibs, A. (1979) *Trends Biochem. Sci.* **4**, 71–72

79. Wu, W., Chang, C. K., Varotsis, C., Babcock, G. T., Puustinen, A., and Wikström, M. (1992) *J. Am. Chem. Soc.* **114**, 1182–1187

80. Sone, N., Ogura, T., Noguchi, S., and Kitagawa, T. (1994) *Biochemistry* **33**, 849–855

81. Fenna, R., Zeng, J., and Davey, C. (1995) *Arch. Biochem. Biophys.* **316**, 653–656

82. Taylor, K. L., Strobel, F., Yue, K. T., Ram, P., Pohl, J., Woods, A. S., and Kinkade, J. M., Jr. (1995) *Arch. Biochem. Biophys.* **316**, 635–642

83. Callahan, P. M., and Babcock, G. T. (1983) *Biochemistry* **22**, 452–461

84. Petke, J. D., and Maggiora, G. M. (1984) *J. Am. Chem. Soc.* **106**, 3129–3133

85. Anraku, Y., and Gennis, R. B. (1987) *Trends Biochem. Sci.* **12**, 262–266

86. Margoliash, E., and Schejter, H. (1984) *Trends Biochem. Sci.* **9**, 364–367

87. Sotiriou, C., and Chang, C. K. (1988) *J. Am. Chem. Soc.* **110**, 2264–2270

88. Timkovich, R., Cork, M. S., and Taylor, P. V. (1984) *J. Biol. Chem.* **259**, 15089–15093

89. Kaufman, J., Spicer, L. D., and Siegel, L. M. (1993) *Biochemistry* **32**, 2853–2867

90. Matthews, J. C., Timkovich, R., Liu, M.-Y., and Le Gall, J. (1995) *Biochemistry* **34**, 5248–5251

91. Chang, C. K., Timkovich, R., and Wu, W. (1986) *Biochemistry* **25**, 8447–8453

92. Mylrajan, M., Andersson, L. A., Loehr, T. M., Wu, W., and Chang, C. K. (1991) *J. Am. Chem. Soc.* **113**, 5000–5005

93. Meyer, T. E., and Kamen, M. D. (1982) *Adv. Prot. Chem.* **35**, 105–212

94. Margoliash, E., and Schejter, A. (1984) *Trends Biochem. Sci.* **9**, 364–367

95. Dickerson, R. E., and Timkovich, R. (1975) in *The Enzymes*, 3rd ed., Vol. 11 (Boyer, P. D., ed), pp. 397–547, Academic Press, New York

96. Dickerson, R. E. (1972) *Sci. Am.* **226**(Apr), 58–72

97. Takano, T., Kallei, O. B., Swanson, R., and Dickerson, R. E. (1973) *J. Biol. Chem.* **248**, 5234–5255

97a. Sebban-Kreuzer, C., Blackledge, M., Dolla, A., Marion, D., and Guerlesquin, F. (1998) *Biochemistry* **37**, 8331–8340

97b. Dolla, A., Arnoux, P., Protasevich, I., Lobachov, V., Brugna, M., Giudici-Orticoni, M. T., Haser, R., Czjzek, M., Makarov, A., and Bruschi, M. (1999) *Biochemistry* **38**, 33–41

98. Margoliash, E., and Frohwirt, N. (1959) *Biochem. J.* **71**, 570–572

99. Margoliash, E., and Bosshard, H. A. (1983) *Trends Biochem. Sci.* **8**, 316–320

100. Poerio, E., Parr, G. R., and Taniuchi, H. (1986) *J. Biol. Chem.* **261**, 10976–10989

101. Salemme, F. R., Kraut, J., and Kamen, M. D. (1973) *J. Biol. Chem.* **248**, 7701–7716

102. Almassy, R. J., and Dickerson, R. E. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2674–2678

102a. Benini, S., González, A., Rypniewski, W. R., Wilson, K. S., Van Beeumen, J. J., and Ciurli, S. (2000) *Biochemistry* **39**, 13115–13126

103. Sogabe, S., and Miki, K. (1995) *J. Mol. Biol.* **252**, 235–247

104. Simões, P., Matias, P. M., Morais, J., Wilson, K., Dauter, Z., Carrondo, M. A., (1998) *Inorganic Chimica Acta* **273**, 213–224

105. Sczekan, S. R., and Joshi, J. G. (1987) *J. Biol. Chem.* **262**, 13780–13788

106. Weber, P. C. (1982) *Biochemistry* **21**, 5116–5119

107. La Mar, G. N., Jackson, J. T., Dugad, L. B., Cusanovich, M. A., and Bartsch, R. G. (1990) *J. Biol. Chem.* **265**, 16173–16180

108. Ren, Z., Meyer, T., and McRee, D. E. (1993) *J. Mol. Biol.* **234**, 433–445

109. Weber, P. C., Salemme, F. R., Mathews, F. S., and Bethge, P. H. (1981) *J. Biol. Chem.* **256**, 7702–7704

110. Hamada, K., Bethge, P. H., and Mathews, F. S. (1995) *J. Mol. Biol.* **247**, 947–962

111. Huang, D., Everly, R. M., Cheng, R. H., Heymann, J. B., Schägger, H., Sled, V., Ohnishi, T., Baker, T. S., and Cramer, W. A. (1994) *Biochemistry* **33**, 4401–4409

112. Prince, R. C., and George, G. N. (1995) *Trends Biochem. Sci.* **20**, 217–218

113. Higuchi, Y., Bando, S., Kusunoki, M., Matsurura, Y., Yasuska, N., Kakuda, M., Yamanaka, T., Yagi, T., and Inokuchi, H. (1981) *J. Biochem.* **89**, 1659–1662

114. Bruschi, M., Woudstra, M., Guigliarelli, B., Asso, M., Lojou, E., Retillot, Y., and Abergel, C. (1997) *Biochemistry* **36**, 10601–10608

115. Czjzek, M., Payan, F., Guerlesquin, F., Bruschi, M., and Haser, R. (1994) *J. Mol. Biol.* **243**, 653–667

116. Liu, M.-C., and Peck, H. D., Jr. (1981) *J. Biol. Chem.* **256**, 13159–13164

117. Banci, L., Bertini, I., Bruschi, M., Sompornpisut, P., and Turano, P. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14396–14400

118. Jenney, F. E., Jr., Prince, R. C., and Daldal, F. (1994) *Biochemistry* **33**, 2496–2502

119. Menin, L., Schoepp, B., Garcia, D., Parot, P., and Verméglio, A. (1997) *Biochemistry* **36**, 12175–12182

120. Menin, L., Schoepp, B., Parot, P., and Verméglio, A. (1997) *Biochemistry* **36**, 12183–12188

121. Adir, N., Axelrod, H. L., Beroza, P., Isaacson, R. A., Rongey, S. H., Okamura, M. Y., and Feher, G. (1996) *Biochemistry* **35**, 2535–2547

122. Kerfeld, C. A., Anwar, H. P., Interrante, R., Merchant, S., and Yeates, T. O. (1995) *J. Mol. Biol.* **250**, 627–647

123. Gao, F., Qin, H., Simpson, M. C., Shelnutt, J. A., Knaff, D. B., and Ondrias, M. R. (1996) *Biochemistry* **35**, 12812–12819

References

124. Okkels, J. S., Kjaer, B., Hansson, Ö., Svendsen, I., Møller, B. L., and Scheller, H. V. (1992) *J. Biol. Chem.* **267**, 21139–21145
125. Link, T. A., Hatzfeld, O. M., Unalkat, P., Shergill, J. K., Cammack, R., and Mason, J. R. (1996) *Biochemistry* **35**, 7546–7552
126. Xia, D., Yu, C.-A., Kim, H., Xia, J.-Z., Kachurin, A. M., Zhang, L., Yu, L., and Deisenhofer, J. (1997) *Science* **277**, 60–66
127. von Jagow, G., and Sebald, W. (1980) *Ann. Rev. Biochem.* **49**, 281–314
128. Kiel, J. L. (1995) *Type-B Cytochromes: Sensors and Switches*, CRC Press, Boca Raton, Florida
129. Mathews, F. S., Levine, M., and Argos, P. (1972) *J. Mol. Biol.* **64**, 449–464
129a. Kostanjevecki, V., Leys, D., Van Driessche, G., Meyer, T. E., Cusanovich, M. A., Fischer, U., Guisez, Y., and Van Beeumen, J. (1999) *J. Biol. Chem.* **274**, 35614–35620
130. Vergères, G., Ramsden, J., and Waskell, L. (1995) *J. Biol. Chem.* **270**, 3414–3422
131. Nagi, M., Cook, L., Prasad, M. R., and Cinti, D. L. (1983) *J. Biol. Chem.* **258**, 14823–14828
132. Nishida, H., Inaka, K., Yamanaka, M., Kaida, S., Kobayashi, K., and Miki, K. (1995) *Biochemistry* **34**, 2763–2767
133. Lindqvist, Y., Brändén, C.-I., Mathews, F. S., and Lederer, F. (1991) *J. Biol. Chem.* **266**, 3198–3207
134. Guiard, B., and Lederer, F. (1979) *J. Mol. Biol.* **135**, 639–650
135. Garrett, R. M., and Rajagopalan, K. V. (1996) *J. Biol. Chem.* **271**, 7387–7391
136. Tegoni, M., Begotti, S., and Cambillau, C. (1995) *Biochemistry* **34**, 9840–9850
137. Weber, P. C., Salemme, F. R., Mathews, F. C., and Bethge, P. H. (1981) *J. Biol. Chem.* **257**, 7702–7704
138. Yu, L., Wei, Y.-Y., Usui, S., and Yu, C.-A. (1992) *J. Biol. Chem.* **267**, 24508–24515
139. Knaff, D. B. (1990) *Trends Biochem. Sci.* **15**, 289–291
140. Jalukar, V., Kelley, P. M., and Njus, D. (1991) *J. Biol. Chem.* **266**, 6878–6882
141. Kita, K., Konishi, K., and Anraku, Y. (1984) *J. Biol. Chem.* **259**, 3368–3374
142. Nakamura, K., Yamaki, M., Sarada, M., Nakayama, S., Vibat, C. R. T., Gennis, R. B., Nakayashiki, T., Inokuchi, H., Kojima, S., and Kita, K. (1996) *J. Biol. Chem.* **271**, 521–527
143. Yang, X., Yu, L., and Yu, C.-A. (1997) *J. Biol. Chem.* **272**, 9683–9689
143a. Ranghino, G., Scorza, E., Sjögren, T., Williams, P. A., Ricci, M., and Hajdu, J. (2000) *Biochemistry* **39**, 10958–10966
144. Gray, H. B., and Winkler, J. R. (1996) *Ann. Rev. Biochem.* **65**, 537–561
145. Farver, O., and Pecht, I. (1991) *FASEB J.* **5**, 2554–2559
146. Komar-Panicucci, S., Sherman, F., and McLendon, G. (1996) *Biochemistry* **35**, 4878–4885
147. Page, C. C., Moser, C. C., Chen, X., and Dutton, P. L. (1999) *Nature (London)* **402**, 47–52
148. DeVault, D. (1984) *Quantum-Mechanical Tunnelling in Biological Systems*, Cambridge Univ. Press, Cambridge, UK
149. Closs, G. L., and Miller, J. R. (1988) *Science* **240**, 440–447
150. Evenson, J. W., and Karplus, M. (1993) *Science* **262**, 1247–1249
151. Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S., and Dutton, P. L. (1992) *Nature (London)* **355**, 796–802
152. Williams, R. J. P. (1992) *Nature (London)* **355**, 770–771
153. Pappa, H. S., and Poulos, T. L. (1995) *Biochemistry* **34**, 6573–6580
154. Mayo, S. L., Ellis, W. R., Jr., Crutchley, R. J., and Gray, H. B. (1986) *Science* **233**, 948–952
155. Wuttke, D. S., Bjerrum, M. J., Winkler, J. R., and Gray, H. B. (1992) *Science* **256**, 1007–1009
156. Banci, L., Bertini, I., De la Rosa, M. A., Koulougliotis, D., Navarro, J. A., and Walter, O. (1998) *Biochemistry* **37**, 4831–4843
157. Moreira, I., Sun, J., Cho, M. O.-K., Wishart, J. F., and Isied, S. S. (1994) *J. Am. Chem. Soc.* **116**, 8396–8397
158. Ahmed, A. J., and Millett, F. (1981) *J. Biol. Chem.* **256**, 1611–1615
159. Koppenol, W. H., and Margoliash, E. (1982) *J. Biol. Chem.* **257**, 4426–4437
160. Hahm, S., Miller, M. A., Geren, L., Kraut, J., Durham, B., and Millett, F. (1994) *Biochemistry* **33**, 1473–1480
161. Willie, A., Stayton, P. S., Sligar, S. G., Durham, B., and Millett, F. (1992) *Biochemistry* **31**, 7237–7242
162. Qin, L., and Kostic, N. M. (1994) *Biochemistry* **33**, 12592–12599
163. Liang, N., Mauk, A. G., Pielak, G. J., Johnson, J. A., Smith, M., and Hoffman, B. M. (1988) *Science* **240**, 311–314
164. Pelletier, H., and Kraut, J. (1992) *Science* **258**, 1748–1755
165. Miller, M. A., Vitello, L., and Erman, J. E. (1995) *Biochemistry* **34**, 12048–12058
166. Mei, H., Wang, K., Peffer, N., Weatherly, G., Cohen, D. S., Miller, M., Pielak, G., Durham, B., and Millett, F. (1999) *Biochemistry* **38**, 6846–6854
167. Falzon, L., and Davidson, V. L. (1996) *Biochemistry* **35**, 12111–12118
168. Merli, A., Brodersen, D. E., Morini, B., Chen, Z.-w, Durley, R. C. E., Mathews, F. S., Davidson, V. L., and Rossi, G. L. (1996) *J. Biol. Chem.* **271**, 9177–9180
169. Mutz, M. W., McLendon, G. L., Wishart, J. F., Gaillard, E. R., and Corin, A. F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9521–9526
169a. Castleman, A. W., Jr., Zhong, Q., and Hurley, S. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4219–4220
170. Daizadeh, I., Medvedev, E. S., and Stuchebrukhov, A. A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3703–3708
171. Langen, R., Chang, I.-J., Germanas, J. P., Richards, J. H., Winkler, J. R., and Gray, H. B. (1995) *Science* **268**, 1733–1735
172. Bishop, G. R., and Davidson, V. L. (1997) *Biochemistry* **36**, 13586–13592
173. Ivkovic-Jensen, M. M., and Kostic, N. M. (1997) *Biochemistry* **36**, 8135–8144
174. Ingraham, L. L. (1966) *Comprehensive Biochemistry* **14**, 424–446
175. Pauling, L. (1948) *The Nature of the Chemical Bond*, 2nd ed., Cornell Univ. Press, Ithaca, New York
176. Smulevich, G., Mantini, A. R., Paoli, M., Coletta, M., and Geraci, G. (1995) *Biochemistry* **34**, 7507–7516
177. Van Dyke, B. R., Saltman, P., and Armstrong, F. A. (1996) *J. Am. Chem. Soc.* **118**, 3490–3492
178. Sugawara, Y., Matsuoka, A., Kaino, A., and Shikama, K. (1995) *Biophys. J.* **69**, 583–592
179. Geibel, J., Chang, C. K., and Traylor, T. G. (1975) *J. Am. Chem. Soc.* **97**, 5924–5926
180. Gerothanassis, I. P., and Momenteau, M. (1987) *J. Am. Chem. Soc.* **109**, 6944–6947
181. Phillips, S. E. V., and Schoenborn, B. P. (1981) *Nature (London)* **292**, 81–84
182. Perutz, M. F. (1989) *Trends Biochem. Sci.* **14**, 42–44
183. Brzozowski, A., Derewenda, Z., Dodson, E., Dodson, G., Grabowski, M., Liddington, R., Skarzynski, T., and Vallely, D. (1984) *Nature (London)* **307**, 74–76
184. McMahon, M. T., deDios, A. C., Godbout, N., Salzmann, R., Laws, D. D., Le, H., Havlin, R. H., and Oldfield, E. (1998) *J. Am. Chem. Soc.* **120**, 4784–4797
185. Collman, J. P., Gagne, R. R., Reed, C. A., Halbert, T. R., Lang, G., and Robinson, W. T. (1975) *J. Am. Chem. Soc.* **97**, 1427–1439
186. Hoard, J. L. (1971) *Science* **174**, 1295–1302
187. Fermi, G., Perutz, M. F., and Shulman, R. G. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6167–6168
188. Klotz, I. M., Klippenstein, G. L., and Hendrickson, W. A. (1976) *Science* **192**, 335–344
189. Dou, T., Admiraal, S. J., Ikeda-Saito, M., Krzywda, S., Wilkinson, A. J., Li, T., Olson, J. S., Prince, R. C., Pickering, I. J., and George, G. N. (1995) *J. Biol. Chem.* **270**, 15993–16001
190. Shiro, Y., Iizuka, T., Marubayashi, K., Ogura, T., Kitagawa, T., Balasubramanian, S., and Boxer, S. G. (1994) *Biochemistry* **33**, 14986–14992
191. Kaminaka, S., Takizawa, H., Handa, T., Kihara, H., and Kitagawa, T. (1992) *Biochemistry* **31**, 6997–7002
192. Zhang, J.-H., and Kurtz, D. M., Jr. (1991) *Biochemistry* **30**, 9121–9125
193. Lukat, G. S., Kurtz, D. M. J., Shiemke, A. K., Loehr, T. M., and Sanders-Loehr, J. (1984) *Biochemistry* **23**, 6416–6422
194. Stenkamp, R. E., Sieker, L. C., Jensen, L. H., McCallum, J. D., and Sanders-Loehr, J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 713–716
194a. Dunford, H. B. (2000) *Heme Peroxidases*, Wiley-VCH, New York
194b. Maté, M. J., Zamocky, M., Nykyri, L. M., Herzog, C., Alzari, P. M., Betzel, C., Koller, F., and Fita, I. (1999) *J. Mol. Biol.* **268**, 135–149
195. Shaffer, J. B., Sutton, R. B., and Bewley, G. C. (1987) *J. Biol. Chem.* **262**, 12908–12911
196. Eaton, J. W., and Ma, M. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2371–2383, McGraw-Hill, New York
197. Bravo, J., Fita, I., Ferrer, J. C., Ens, W., Hillar, A., Switala, J., and Loewen, P. C. (1997) *Protein Sci.* **6**, 1016–1023
198. Fita, I., and Rossmann, M. G. (1985) *J. Mol. Biol.* **185**, 21–37
198a. Putnam, C. D., Arvai, A. S., Bourne, Y., and Tainer, J. A. (2000) *J. Mol. Biol.* **296**, 295–309
199. Vainshtein, B. K., Melik-Adamyan, W. R., Barynin, V. V., Vagin, A. A., and Grebenko, A. I. (1981) *Nature (London)* **293**, 411–412
200. Fita, I., and Rossmann, M. G. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1604–1608
201. Kirkman, H. N., Rolfo, M., Ferraris, A. M., and Gaetani, G. F. (1999) *J. Biol. Chem.* **274**, 13908–13914
202. Gouet, P., Jouve, H.-M., and Dideberg, O. (1995) *J. Mol. Biol.* **249**, 933–954
203. Jacob, G. S., and Orme-Johnson, W. H. (1979) *Biochemistry* **18**, 2975–2980
204. Kono, Y., and Fridovich, I. (1983) *J. Biol. Chem.* **258**, 6015–6019
205. Khangulov, S., Sivaraja, M., Barynin, V. V., and Dismukes, G. C. (1993) *Biochemistry* **32**, 4912–4924
206. Holzbaur, I. E., English, A. M., and Ismail, A. A. (1996) *J. Am. Chem. Soc.* **118**, 3354–3359
207. Ryan, O., Smyth, M. R., and Fágáin, C. O. (1994) *Essays in Biochemistry* **28**, 129–146
208. Newmyer, S. L., and Ortiz de Montellano, P. R. (1995) *J. Biol. Chem.* **270**, 19430–19438
208a. Rodríguez-López, J. N., Gilabert, M. A., Tudela, J., Thorneley, R. N. F., and García-Cánovas, F. (2000) *Biochemistry* **39**, 13201–13209

References

209. Rodriguez-Lopez, J. N., Hernández-Ruiz, J., Garcia-Cánovas, F., Thorneley, R. N. F., Acosta, M., and Arnano, M. B. (1997) *J. Biol. Chem.* **272**, 5469 – 5476
210. Abelskov, A. K., Smith, A. T., Rasmussen, C. B., Dunford, H. B., and Welinder, K. G. (1997) *Biochemistry* **36**, 9453 – 9463
211. Fukuyama, K., Sato, K., Itakura, H., Takahashi, S., and Hosoya, T. (1997) *J. Biol. Chem.* **272**, 5752 – 5756
212. Nissum, M., Neri, F., Mandelman, D., Poulos, T. L., and Smulevich, G. (1998) *Biochemistry* **37**, 8080 – 8087
213. Patterson, W. R., and Poulos, T. L. (1995) *Biochemistry* **34**, 4331 – 4341
214. Patterson, W. R., Poulos, T. L., and Goodin, D. B. (1995) *Biochemistry* **34**, 4342 – 4345
215. Finzel, B. C., Poulos, T. L., and Kraut, J. (1984) *J. Biol. Chem.* **259**, 13027 – 13036
216. Bonagura, C. A., Sundaramoorthy, M., Pappa, H. S., Patterson, W. R., and Poulos, T. L. (1996) *Biochemistry* **35**, 6107 – 6115
217. Wang, J., Larsen, R. W., Moench, S. J., Satterlee, J. D., Rousseau, D. L., and Ondrias, M. R. (1996) *Biochemistry* **35**, 453 – 463
218. Sinclair, R., Copeland, B., Yamazaki, I., and Powers, L. (1995) *Biochemistry* **34**, 13176 – 13182
218a. Doyle, W. A., Blodig, W., Veitch, N. C., Piontek, K., and Smith, A. T. (1998) *Biochemistry* **37**, 15097 – 15105
218b. Choinowski, T., Blodig, W., Winterhalter, K. H., and Piontek, K. (1999) *J. Mol. Biol.* **286**, 809 – 827
219. Yamaguchi, Y., Zhang, D.-E., Sun, Z., Albee, E. A., Nagata, S., Tenen, D. G., and Ackerman, S. J. (1994) *J. Biol. Chem.* **269**, 19410 – 19419
220. Andersson, L. A., Bylkas, S. A., and Wilson, A. E. (1996) *J. Biol. Chem.* **271**, 3406 – 3412
221. Rae, T. D., and Goff, H. M. (1996) *J. Am. Chem. Soc.* **118**, 2103 – 2104
222. Kooter, I. M., Moguilevsky, N., Bollen, A., van der Veen, L. A., Otto, C., Dekker, H. L., and Wever, R. (1999) *J. Biol. Chem.* **274**, 26794 – 26802
222a. Fiedler, T. J., Davey, C. A., and Fenna, R. E. (2000) *J. Biol. Chem.* **275**, 11964 – 11971
223. Nagano, S., Tanaka, M., Ishimori, K., Watanabe, Y., and Morishima, I. (1996) *Biochemistry* **35**, 14251 – 14258
224. Tanaka, M., Ishimori, K., and Morishima, I. (1998) *Biochemistry* **37**, 2629 – 2638
225. Low, D. W., Gray, H. B., and Duus, J. ø. (1997) *J. Am. Chem. Soc.* **119**, 1 – 5
226. Makino, R., Uno, T., Nishimura, Y., Iuzuka, T., Tsuboi, M., and Ishimura, Y. (1986) *J. Biol. Chem.* **261**, 8376 – 8382
227. Hashimoto, S., Tatsuno, Y., and Kitagawa, T. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2417 – 2421
228. Nakajima, R., and Yamazaki, I. (1987) *J. Biol. Chem.* **262**, 2576 – 2581
229. Morishima, I., Takamuki, Y., and Shiro, Y. (1984) *J. Am. Chem. Soc.* **106**, 7666 – 7672
230. Loew, G., and Dupuis, M. (1996) *J. Am. Chem. Soc.* **118**, 10584 – 10587
230a. Filizola, M., and Loew, G. H. (2000) *J. Am. Chem. Soc.* **122**, 18 – 25
230b. Roach, M. P., Ozaki, S.-i, and Watanabe, Y. (2000) *Biochemistry* **39**, 1446 – 1454
231. Olson, L. P., and Bruice, T. C. (1995) *Biochemistry* **34**, 7335 – 7347
232. Nakamura, M., Yamazaki, I., Kotani, T., and Ohtaki, S. (1985) *J. Biol. Chem.* **260**, 13546 – 13552
233. Kobayashi, S., Nakano, M., Kimura, T., and Schaap, A. P. (1987) *Biochemistry* **26**, 5019 – 5022
234. Ortiz de Montellano, P. R., Choe, Y. S., DePillis, G., and Catalano, C. E. (1987) *J. Biol. Chem.* **262**, 11641 – 11646
235. Dawson, J. H. (1988) *Science* **240**, 433 – 439
236. Libby, R. D., Thomas, J. A., Kaiser, L. W., and Hager, L. P. (1982) *J. Biol. Chem.* **257**, 5030 – 5037
237. Hosten, C. M., Sullivan, A. M., Palaniappan, V., Fitzgerald, M. M., and Terner, J. (1994) *J. Biol. Chem.* **269**, 13966 – 13978
238. Sono, M., Eble, K. S., Dawson, J. H., and Hager, L. P. (1985) *J. Biol. Chem.* **260**, 15530 – 15535
239. Sono, M., Dawson, J. H., and Hager, L. P. (1984) *J. Biol. Chem.* **259**, 13209 – 13216
240. Dawson, J. H., and Sono, M. (1987) *Chem. Rev.* **87**, 1255 – 1276
241. Geigert, J., Neidleman, S. L., and Dalietos, D. J. (1983) *J. Biol. Chem.* **258**, 2273 – 2277
242. Mueller, T. J., and Morrison, M. (1974) *J. Biol. Chem.* **259**, 7568 – 7573
243. Andersson, E., Hellman, L., Gullberg, U., and Olsson, I. (1998) *J. Biol. Chem.* **273**, 4747 – 4753
244. van Dalen, C. J., Winterbourn, C. C., Senthilmohan, R., and Kettle, A. J. (2000) *J. Biol. Chem.* **275**, 11638 – 11644
245. Nauseef, W. M., Brigham, S., and Cogley, M. (1994) *J. Biol. Chem.* **269**, 1212 – 1216
246. Davey, C. A., and Fenna, R. E. (1996) *Biochemistry* **35**, 10967 – 10973
247. Thomas, E. L., Bozeman, P. M., Jefferson, M. M., and King, C. C. (1995) *J. Biol. Chem.* **270**, 2906 – 2913
248. Kanofsky, J. R. (1983) *J. Biol. Chem.* **258**, 5991– 5993
249. Khan, A. U., Gebauer, P., and Hager, L. P. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5195 – 5197
250. McCormick, M. L., Roeder, T. L., Railsback, M. A., and Britigan, B. E. (1994) *J. Biol. Chem.* **269**, 27914 – 27919
251. Manthey, J. A., and Hager, L. P. (1981) *J. Biol. Chem.* **256**, 11232 – 11238
252. Colpas, G. J., Hamstra, B. J., Kampf, J. W., and Pecoraro, V. L. (1996) *J. Am. Chem. Soc.* **118**, 3469 – 3478
253. Messerschmidt, A., and Wever, R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 392 – 396
254. Mande, S. S., Parsonage, D., Claiborne, A., and Hol, W. G. J. (1995) *Biochemistry* **34**, 6985 – 6992
255. Crane, E. J., III, Parsonage, D., Poole, L. B., and Claiborne, A. (1995) *Biochemistry* **34**, 14114 – 14124
255a. Claiborne, A., Yeh, J. I., Mallett, T. C., Luba, J., Crane, E. J., III, Charrier, V., and Parsonage, D. (1999) *Biochemistry* **38**, 15407 – 15416
256. Miki, K., Renganathan, V., and Gold, M. H. (1986) *Biochemistry* **25**, 4790 – 4796
257. Sundaramoorthy, M., Kishi, K., Gold, M. H., and Poulos, T. L. (1994) *J. Biol. Chem.* **269**, 32759 – 32767
258. Mauk, M. R., Kishi, K., Gold, M. H., and Mauk, A. G. (1998) *Biochemistry* **37**, 6767 – 6771
259. Beinert, H. (1990) *FASEB J.* **4**, 2483 – 2491
260. Beinert, H., Holm, R. H., and Münck, E. (1997) *Science* **277**, 653 – 659
261. Carter, C. W., Jr., Kraut, J., Freer, S. T., Xuong, N., Alden, R. A., and Bartsch, R. G. (1974) *J. Biol. Chem.* **249**, 4212 – 4225
262. Breiter, D. R., Meyer, T. E., Rayment, I., and Holden, H. M. (1991) *J. Biol. Chem.* **266**, 18660 – 18667
263. Agarwal, A., Li, D., and Cowan, J. A. (1996) *J. Am. Chem. Soc.* **118**, 927 – 928
264. Banci, L., Bertini, I., Ciurli, S., Ferretti, S., Luchinat, C., and Piccioli, M. (1993) *Biochemistry* **32**, 9387 – 9397
265. Benning, M. M., Meyer, T. E., Rayment, I., and Holden, H. M. (1994) *Biochemistry* **33**, 2476 – 2483
266. Soriano, A., Li, D., Bian, S., Agarwal, A., and Cowan, J. A. (1996) *Biochemistry* **35**, 12479 – 12486
266a. Mulholland, S. E., Gibney, B. R., Rabanal, F., and Dutton, P. L. (1999) *Biochemistry* **38**, 10442 – 10448
267. Duée, E. D., Fanchon, E., Vicat, J., Sieker, L. C., Meyer, J., and Moulis, J.-M. (1994) *J. Mol. Biol.* **243**, 683 – 695
268. Dauter, Z., Wilson, K. S., Sieker, L. C., Meyer, J., and Moulis, J.-M. (1997) *Biochemistry* **36**, 16065 – 16073
269. Jensen, L. H. (1974) *Ann. Rev. Biochem.* **43**, 461 – 474
270. Séry, A., Housset, D., Serre, L., Bonicel, J., Hatchikian, C., Frey, M., and Roth, M. (1994) *Biochemistry* **33**, 15408 – 15417
270a. Kyritsis, P., Kümmerle, R., Huber, J. G., Gaillard, J., Guigliarelli, B., Popescu, C., Münck, E., and Moulis, J.-M. (1999) *Biochemistry* **38**, 6335 – 6345
271. Brereton, P. S., Verhagen, M. F. J. M., Zhou, Z. H., and Adams, M. W. W. (1998) *Biochemistry* **37**, 7351 – 7362
272. Bruschi, M. H., Guerlesquin, F. A., Bovier-Lapience, G. E., Bonicel, J. J., and Couchoud, P. M. (1985) *J. Biol. Chem.* **260**, 8292 – 8296
273. Bruice, T. C., Maskiewicz, R., and Job, R. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 231 – 234
274. Dauter, Z., Wilson, K. S., Sieker, L. C., Moulis, J.-M., and Meyer,J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8836 – 8840
275. Xiao, Z., Lavery, M. J., Ayhan, M., Scrofani, S. D. B., Wilce, M. C. J., Guss, J. M., Tregloan, P. A., George, G. N., and Wedd, A. G. (1998) *J. Am. Chem. Soc.* **120**, 4135 – 4150
276. Fukuyama, K., Hase, T., Matsumoto, S., Tsukihara, T., Katsube, Y., Tanaka, N., Kakudo, M., Wada, K., and Matsubara, H. (1980) *Nature (London)* **286**, 522 – 524
277. Kok, M., Oldenhuis, R., van der Linden, M. P. G., Meulenberg, C. H. C., Kingma, J., and Witholt, B. (1989) *J. Biol. Chem.* **264**, 5442 – 5451
278. Archer, M., Huber, R., Tavares, P., Moura, I., Moura, J. J. G., Carrondo, M. A., Sieker, L. C., LeGall, J., and Romao, M. J. (1995) *J. Mol. Biol.* **251**, 690 – 702
279. Hurley, J. K., Weber-Main, A. M., Stankovich, M. T., Benning, M. M., Thoden, J. B., Vanhooke, J. L., Holden, H. M., Chae, Y. K., Xia, B., Cheng, H., Markley, J. L., Martinez-Júlvez, M., Gómez-Moreno, C., Schmeits, J. L., and Tollin, G. (1997) *Biochemistry* **36**, 11100 – 11117
279a. Müller, J. J., Müller, A., Rottmann, M., Bernhardt, R., and Heinemann, U. (1999) *J. Mol. Biol.* **294**, 501 – 513
280. Hurley, J. K., Weber-Main, A. M., Hodges, A. E., Stankovich, M. T., Benning, M. M., Holden, H. M., Cheng, H., Xia, B., Markley, J. L., Genzor, C., Gomez-Moreno, C., Hafezi, R., and Tollin, G. (1997) *Biochemistry* **36**, 15109 – 15117
280a. Morales, R., Chron, M.-H., Hudry-Clergeon, G., Pétillot, Y., Norager, S., Medina, M., and Frey, M. (1999) *Biochemistry* **38**, 15764 – 15773
281. Ta, D. T., and Vickery, L. E. (1992) *J. Biol. Chem.* **267**, 11120 – 11125
282. Xia, B., Cheng, H., Skjeldal, L., Coghlan, V. M., Vickery, L. E., and Markley, J. L. (1995) *Biochemistry* **34**, 180 – 187
282a. Uhlmann, H., and Bernhardt, R. (1995) *J. Biol. Chem.* **270**, 29959 – 29966
282b. Ziegler, G. A., and Schulz, G. E. (2000) *Biochemistry* **39**, 10986 – 10995
283. Meyer, J., Fujinaga, J., Gaillard, J., and Lutz, M. (1994) *Biochemistry* **33**, 13642 – 13650
284. Pochapsky, T. C., and Ye, X. M. (1991) *Biochemistry* **30**, 3850 – 3856
285. Stout, C. D. (1988) *J. Biol. Chem.* **263**, 9256 – 9260
286. Duff, J. L. C., Breton, J. L. J., Butt, J. N., Armstrong, F. A., and Thomson, A. J. (1996) *J. Am. Chem. Soc.* **118**, 8593 – 8603

References

287. Schipke, C. G., Goodin, D. B., McRee, D. E., and Stout, C. D. (1999) *Biochemistry* **38**, 8228–8239
288. George, G. N., and George, S. J. (1988) *Trends Biochem. Sci.* **13**, 369–370
289. Kent, T. A., Emptage, M. H., Merkle, H., Kennedy, M. C., Beinert, H., and Münck, E. (1985) *J. Biol. Chem.* **260**, 6871–6881
290. Tong, J., and Feinberg, B. A. (1994) *J. Biol. Chem.* **269**, 24920–24927
291. Lane, R. W., Ibers, J. A., Frankel, R. B., and Holm, R. H. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2868–2872
291a. Camba, R., and Armstrong, F. A. (2000) *Biochemistry* **39**, 10587–10598
292. Gardner, P. R., and Fridovich, I. (1991) *J. Biol. Chem.* **266**, 19328–19333
293. Kennedy, M. C., Antholine, W. E., and Beinert, H. (1997) *J. Biol. Chem.* **272**, 20340–20347
294. Duin, E. C., Lafferty, M. E., Crouse, B. R., Allen, R. M., Sanyal, I., Flint, D. H., and Johnson, M. K. (1997) *Biochemistry* **36**, 11811–11820
295. Magliozzo, R. S., McIntosh, B. A., and Sweeney, W. V. (1982) *J. Biol. Chem.* **257**, 3506–3509
296. Prince, R. C., and Adams, M. W. W. (1987) *J. Biol. Chem.* **262**, 5125–5128
297. Orme-Johnson, W. H. (1972) *Ann. Rev. Biochem.* **42**, 159–204
298. Orme-Johnson, W. H., Hansen, R. E., Beinert, H., Tsibris, J. C. M., Bartholomaus, R. C., and Gunsalus, I. C. (1968) *Proc. Natl. Acad. Sci. U.S.A.* **60**, 368–372
299. Scrofani, S. D. B., Brereton, P. S., Hamer, A. M., Lavery, M. J., McDowall, S. G., Vincent, G. A., Brownlee, R. T. C., Hoogenraad, N. J., Sadek, M., and Wedd, A. G. (1994) *Biochemistry* **33**, 14486–14495
300. Holz, R. C., Small, F. J., and Ensign, S. A. (1997) *J. Biol. Chem.* **36**, 14690–14696
301. Mouesca, J.-M., Chen, J. L., Noodleman, L., Bashford, D., and Case, D. A. (1994) *J. Am. Chem. Soc.* **116**, 11898–11914
302. Swartz, P. D., Beck, B. W., and Ichiye, T. (1996) *Biophys. J.* **71**, 2958–2969
303. Kemper, M. A., Stout, C. D., Lloyd, S. E. J., Prasad, G. S., Fawcett, S., Armstrong, F. A., Shen, B., and Burgess, B. K. (1997) *J. Biol. Chem.* **272**, 15620–15627
304. Bominaar, E. L., Hu, Z., Münck, E., Girerd, J.-J., and Borshch, S. A. (1995) *J. Am. Chem. Soc.* **117**, 6976–6989
305. Scott, M. P., and Biggins, J. (1997) *Protein Sci.* **6**, 340–346
306. Coldren, C. D., Hellinga, H. W., and Caradonna, J. P. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6635–6640
307. Rieske, J. S., MacLennan, D. H., and Coleman, R. (1964) *Biochem. Biophys. Res. Commun.* **15**, 338–344
308. Beckmann, J. D., Ljungdahl, P. O., Lopez, J. L., and Trumpower, B. L. (1987) *J. Biol. Chem.* **262**, 8901–8909
309. Brasseur, G., Sled, V., Liebl, U., Ohnishi, T., and Daldal, F. (1997) *Biochemistry* **36**, 11685–11696
310. Mosser, G., Breyton, C., Olofsson, A., Popot, J.-L., and Rigaud, J.-L. (1997) *J. Biol. Chem.* **272**, 20263 – 20268
311. Liebl, U., Sled, V., Brasseur, G., Ohnishi, T., and Daldal, F. (1997) *Biochemistry* **36**, 11675–11684
312. Crane, B. R., Siegel, L. M., and Getzoff, E. D. (1995) *Science* **270**, 59–67
313. Wilkerson, J. O., Janick, P. A., and Siegel, L. M. (1983) *Biochemistry* **22**, 5048–5054
314. Crane, B. R., Siegel, L. M., and Getzoff, E. D. (1997) *Biochemistry* **36**, 12120–12137
315. Covès, J., Zeghouf, M., Macherel, D., Guigliarelli, B., Asso, M., and Fontecave, M. (1997) *Biochemistry* **36**, 5921–5928
316. Crane, B. R., Siegel, L. M., and Getzoff, E. D. (1997) *Biochemistry* **36**, 12101–12119
316a. Peters, J. W., Lanzilotta, W. N., Lemon, B. J., and Seefeldt, L. C. (1998) *Science* **282**, 1853–1858
316b. Popescu, C. V., and Münck, E. (1999) *J. Am. Chem. Soc.* **121**, 7877–7884
316c. Nicolet, Y., Lemon, B. J., Fontecilla-Camps, J. C., and Peters, J. W. (2000) *Trends Biochem. Sci.* **25**, 138–143
317. Grabowski, R., Hofmeister, A. E. M., and Buckel, W. (1993) *Trends Biochem. Sci.* **18**, 297–300
318. Hofmeister, A. E. M., Grabowski, R., Linder, D., and Buckel, W. (1993) *Eur. J. Biochem.* **215**, 341–349
319. Hofmeister, A. E. M., Berger, S., and Buckel, W. (1992) *Eur. J. Biochem.* **205**, 743–749
320. Flint, D. H., and Emptage, M. H. (1988) *J. Biol. Chem.* **263**, 3558–3564
321. Flint, D. H., Tuminello, J. F., and Miller, T. J. (1996) *J. Biol. Chem.* **271**, 16053–16067
322. Hofmeister, A. E. M., and Buckel, W. (1992) *Eur. J. Biochem.* **206**, 547–552
323. Müh, U., Cinkaya, I., Albracht, S. P. J., and Buckel, W. (1996) *Biochemistry* **35**, 11710–11718
324. Scherf, U., Söhling, B., Gottschalk, G., Linder, D., and Buckel, W. (1994) *Arch. Microbiol.* **161**, 239–245
325. Klees, A.-G., Linder, D., and Buckel, W. (1992) *Arch. Microbiol.* **158**, 294–301
326. Vollmer, S. J., Switzer, R. L., and Debrunner, P. G. (1983) *J. Biol. Chem.* **258**, 14284–14293
327. Ramsay, R. R., Dreyer, J.-L., Schloss, J. V., Jackson, R. H., Coles, C. J., Beinert, H., Cleland, W. W., and Singer, T. P. (1981) *Biochemistry* **20**, 7476–7482
328. Gaudu, P., and Weiss, B. (1996) *Proc. R. Soc. (London)* **93**, 10094–10098
329. Ding, H., Hidalgo, E., and Demple, B. (1996) *J. Biol. Chem.* **271**, 33173–33175
330. Hentze, M. W. (1996) *Trends Biochem. Sci.* **21**, 282–283
331. Feig, A. L., Masschelein, A., Bakac, A., and Lippard, S. J. (1997) *J. Am. Chem. Soc.* **119**, 334–342
332. Pulver, S. C., Froland, W. A., Lipscomb, J. D., and Solomon, E. I. (1997) *J. Am. Chem. Soc.* **119**, 387–395
333. Lindqvist, Y., Huang, W., Schneider, G., and Shanklin, J. (1996) *EMBO J.* **15**, 4081–4092
334. Que, L., Jr. (1991) *Science* **253**, 273–274
335. Doi, K., Gupta, R., and Aisen, P. (1987) *J. Biol. Chem.* **262**, 6982–6985
336. Hayman, A. R., and Cox, T. M. (1994) *J. Biol. Chem.* **269**, 1294–1300
336a. Lindqvist, Y., Johansson, E., Kaija, H., Vihko, P., and Schneider, G. (1999) *J. Mol. Biol.* **291**, 135–147
336b. Uppenberg, J., Lindqvist, F., Svensson, C., Ek-Rylander, B., and Andersson, G. (1999) *J. Mol. Biol.* **290**, 201–211
337. Wang, D. L., Holz, R. C., David, S. S., Que, L., Jr., and Stankovich, M. T. (1991) *Biochemistry* **30**, 8187–8194
338. Cohen, S. S. (1984) *Trends Biochem. Sci.* **9**, 334–336
339. Merkx, M., and Averill, B. A. (1998) *Biochemistry* **37**, 8490–8497
340. Klabunde, T., Sträter, N., Fröhlich, R., Witzel, H., and Krebs, B. (1996) *J. Mol. Biol.* **259**, 737–748
341. Sträter, N., Klabunde, T., Tucker, P., Witzel, H., and Krebs, B. (1995) *Science* **268**, 1489–1492
342. Baumbach, G. A., Ketcham, C. M., Richardson, D. E., Bazer, F. W., and Roberts, R. M. (1986) *J. Biol. Chem.* **261**, 12869–12878
343. Bender, C. J., Rosenzweig, A. C., Lippard, S. J., and Peisach, J. (1994) *J. Biol. Chem.* **269**, 15993–15998
344. Nesheim, J. C., and Lipscomb, J. D. (1996) *Biochemistry* **35**, 10240–10247
345. Shu, L., Nesheim, J. C., Kauffmann, K., Münck, E., Lipscomb, J. D., and Que, L., Jr. (1997) *Science* **275**, 515–518
346. Pikus, J. D., Studts, J. M., McClay, K., Steffan, R. J., and Fox, B. G. (1997) *Biochemistry* **36**, 9283–9289
347. Herold, S., and Lippard, S. J. (1997) *J. Am. Chem. Soc.* **119**, 145–156
348. Jordan, A., and Reichard, P. (1998) *Ann. Rev. Biochem.* **67**, 71–98
349. Stubbe, J. (1990) *J. Biol. Chem.* **265**, 5329–5332
350. Reichard, P. (1997) *Trends Biochem. Sci.* **22**, 81–85
351. Stubbe, J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2723–2724
351a. Eriksson, M., Jordan, A., and Eklund, H. (1998) *Biochemistry* **37**, 13359–13369
352. Sahlin, M., Petersson, L., Gräslund, A., Ehrenberg, A., Sjöberg, B.-M., and Thelander, L. (1987) *Biochemistry* **26**, 5541–5548
353. Uhlin, U., and Eklund, H. (1994) *Nature (London)* **370**, 533–539
354. Uhlin, U., and Eklund, H. (1996) *J. Mol. Biol.* **262**, 358–369
354a. Parkin, S. E., Chen, S., Ley, B. A., Mangravite, L., Edmondson, D. E., Huynh, B. H., and Bollinger, J. M., Jr. (1998) *Biochemistry* **37**, 1124–1130
354b. Rova, U., Adrait, A., Pötsch, S., Gräslund, A., and Thelander, L. (1999) *J. Biol. Chem.* **274**, 23746–23751
354c. Andersson, M. E., Högbom, M., Rinaldo-Matthis, A., Andersson, K. K., Sjöberg, B.-M., and Nordlund, P. (1999) *J. Am. Chem. Soc.* **121**, 2346–2352
355. Fieschi, F., Torrents, E., Toulokhonova, L., Jordan, A., Hellman, U., Barbe, J., Gibert, I., Karlsson, M., and Sjöberg, B.-M. (1998) *J. Biol. Chem.* **273**, 4329–4337
356. Ling, J., Sahlin, M., Sjöberg, B.-M., Loehr, T. M., and Sanders-Loehr, J. (1994) *J. Biol. Chem.* **269**, 5595–5601
357. Silva, K. E., Elgren, T. E., Que, L., Jr., and Stankovich, M. T. (1995) *Biochemistry* **34**, 14093–14103
358. Sturgeon, B. E., Burdi, D., Chen, S., Huynh, B.-H., Edmondson, D. E., Stubbe, J., and Hoffman, B. M. (1996) *J. Am. Chem. Soc.* **118**, 7551–7557
359. Kauppi, B., Nielsen, B. B., Ramaswamy, S., Larson, I. K., Thelander, M., Thelander, L., and Eklund, H. (1996) *J. Mol. Biol.* **262**, 706–720
360. Katterle, B., Sahlin, M., Schmidt, P. P., Pötsch, S., Logan, D. T., Gräslund, A., and Sjöberg, B.-M. (1997) *J. Biol. Chem.* **272**, 10414–10421
361. Sealy, R. C., Harman, L., West, P. R., and Mason, R. P. (1985) *J. Am. Chem. Soc.* **107**, 3401–3406
362. Sjöberg, B.-M., Karlsson, M., and Jörnvall, H. (1987) *J. Biol. Chem.* **262**, 9736–9743
363. Ekberg, M., Sahlin, M., Eriksson, M., and Sjöberg, B.-M. (1996) *J. Biol. Chem.* **271**, 20655–20659
364. Ong, S. P., McFarlan, S. C., and Hogenkamp, H. P. C. (1993) *Biochemistry* **32**, 11397–11404
365. Gerfen, G. J., Licht, S., Willems, J.-P., Hoffman, B. M., and Stubbe, J. (1996) *J. Am. Chem. Soc.* **118**, 8192–8197
366. Reichard, P. (1993) *J. Biol. Chem.* **268**, 8383–8386
367. Sun, X., Eliasson, R., Pontis, E., Andersson, J., Buist, G., Sjöberg, B.-M., and Reichard, P. (1995) *J. Biol. Chem.* **270**, 2443–2446

368. Sun, X., Ollagnier, S., Schmidt, P. P., Atta, M., Mulliez, E., Lepape, L., Eliasson, R., Gräslund, A., Fontecave, M., Reichard, P., and Sjöberg, B.-M. (1996) *J. Biol. Chem.* **271**, 6827–6831
369. Ollagnier, S., Mulliez, E., Schmidt, P. P., Eliasson, R., Gaillard, J., Deronzier, C., Bergman, T., Gräslund, A., Reichard, P., and Fontecave, M. (1997) *J. Biol. Chem.* **272**, 24216–24223
369a. Logan, D. T., Andersson, J., Sjöberg, B.-M., and Nordlund, P. (1999) *Science* **283**, 1499–1504
370. Young, P., Andersson, J., Sahlin, M., and Sjöberg, B.-M. (1996) *J. Biol. Chem.* **271**, 20770–20775
371. van der Donk, W. A., Stubbe, J., Gerfen, G. J., Bellew, B. F., and Griffin, R. G. (1995) *J. Am. Chem. Soc.* **117**, 8908–8916
371a. Lawrence, C. C., Bennati, M., Obias, H. V., Bar, G., Griffin, R. G., and Stubbe, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8979–8984
372. Covès, J., Le Hir de Fallois, L., Le Pape, L., Décout, J.-L., and Fontecave, M. (1996) *Biochemistry* **35**, 8595–8602
373. Lenz, R., and Giese, B. (1997) *J. Am. Chem. Soc.* **119**, 2784–2794
374. Trant, N. L., Meshnick, S. R., Kitchener, K., Eaton, J. W., and Cerami, A. (1983) *J. Biol. Chem.* **258**, 125–130
375. Cooper, J. B., McIntyre, K., Badasso, M. O., Wood, S. P., Zhang, Y., Garbe, T. R., and Young, D. (1995) *J. Mol. Biol.* **246**, 531–544
376. Lah, M. S., Dixon, M. M., Pattridge, K. A., Stallings, W. C., Fee, J. A., and Ludwig, M. L. (1995) *Biochemistry* **34**, 1646–1660
377. Stallings, W. C., Pattridge, K. A., Strong, R. K., and Ludwig, M. L. (1985) *J. Biol. Chem.* **260**, 16424–16432
378. Sorkin, D. L., Duong, D. K., and Miller, A.-F. (1997) *Biochemistry* **36**, 8202–8208
378a. Ursby, T., Adinolfi, B. S., Al-Karadaghi, S., De Vendittis, E., and Bocchini, V. (1999) *J. Mol. Biol.* **286**, 189–205
379. Carlioz, A., Ludwig, M. L., Stallings, W. C., Fee, J. A., Steinman, H. M., and Touati, D. (1988) *J. Biol. Chem.* **263**, 1555–1562
380. Barker, H. A. (1972) in *The Enzymes*, 3rd ed., Vol. 6 (Boyer, P. D., ed), pp. 509–537, Academic Press, New York
381. Barker, H. A. (1976) in *Relections on Biochemistry* (Kornberg, A., ed), pp. 95–, Pergamon Press, New York
382. Lenhert, P. G., and Hodgkin, D. C. (1961) *Nature (London)* **192**, 937–938
383. Lenhert, P. G., and Hodgkin, D. C. (1961) *Nature (London)* **192**, 937–938
384. Zagalak, B., and Friedrich, W., eds. (1979) *Vitamin B12*, de Gruyter, Berlin
385. Dolphin, D., ed. (1982) *B12*, Wiley, New York (2 vols.)
386. Brown, K. L. (1987) *J. Am. Chem. Soc.* **109**, 2277–2284
387. Anton, D. L., Hogenkamp, H. P. C., Walker, T. E., and Matwiyoff, N. A. (1982) *Biochemistry* **21**, 2372–2378
388. Wagner, F. (1966) *Ann. Rev. Biochem.* **35**, 405–428
389. Schrauzer, G. N., Deutsch, E., and Windgassen, R. J. (1968) *J. Am. Chem. Soc.* **90**, 2441–2442
390. Needham, T. E., Matwiyoff, N. A., Walker, T. E., and Hogenkamp, H. P. C. (1973) *J. Am. Chem. Soc.* **95**, 5019–5024
391. Parry, R. J., Ostrander, J. M., and Arzu, I. Y. (1985) *J. Am. Chem. Soc.* **107**, 2190–2191
391a. Walker, L. A., II, Shiang, J. J., Anderson, N. A., Pullen, S. H., and Sension, R. J. (1998) *J. Am. Chem. Soc.* **120**, 7286–7292
392. Hogenkamp, H. P. C. (1982) in B_{12}, Vol. I (Dolphin, D., ed), pp. 295–323, Wiley, New York
393. Hay, B. P., and Finke, R. G. (1986) *J. Am. Chem. Soc.* **108**, 4820–4829
394. Ashley, G. W., Harris, G., and Stubbe, J. (1986) *J. Biol. Chem.* **261**, 3958–3964
395. Sando, G. N., Blakely, R. L., Hogenkamp, H. P. C., and Hoffman, P. J. (1975) *J. Biol. Chem.* **250**, 8774–8779
396. Thelander, L., and Reichard, P. (1979) *Ann. Rev. Biochem.* **48**, 133–158
396a. Brown, K. L., and Li, J. (1998) *J. Am. Chem. Soc.* **120**, 9466–9474
396b. Licht, S. S., Lawrence, C. C., and Stubbe, J. (1999) *J. Am. Chem. Soc.* **121**, 7463–7468
397. Pratt, J. M. (1982) in B_{12}, Vol. I (Dolphin, D., ed), pp. 325–392, Wiley, New York
397a. Zerbe-Burkhardt, K., Ratnatilleke, A., Philippon, N., Birch, A., Leiser, A., Vrijbloed, J. W., Hess, D., Hunziker, P., and Robinson, J. A. (1998) *J. Biol. Chem.* **273**, 6508–6517
397b. Ratnatilleke, A., Vrijbloed, J. W., and Robinson, J. A. (1999) *J. Biol. Chem.* **274**, 31679–31685
398. Retéy, J., Umani-Ronchi, A., Seibl, J., and Arigoni, D. (1966) *Experientia* **22**, 502–503
399. Sato, K., Orr, J. C., Babior, B. M., and Abeles, R. H. (1976) *J. Biol. Chem.* **251**, 3734–3737
400. Wollowitz, S., and Halpern, J. (1984) *J. Am. Chem. Soc.* **106**, 8319–8321
401. Graves, S. W., Krouwer, J. S., and Babior, B. M. (1980) *J. Biol. Chem.* **255**, 7444–7448
401a. Smith, D. M., Golding, B. T., and Radom, L. (1999) *J. Am. Chem. Soc.* **121**, 5700–5704
401b. Ke, S.-C., and Warncke, K. (1999) *J. Am. Chem. Soc.* **121**, 9922–9927
401c. Warncke, K., Schmidt, J. C., and Ke, S.-C. (1999) *J. Am. Chem. Soc.* **121**, 10522–10528
402. Bothe, H., Darley, D. J., Albracht, S. P. J., Gerfen, G. J., Golding, B. T., and Buckel, W. (1998) *Biochemistry* **37**, 4105–4113
402a. Chih, H.-W., and Marsh, E. N. G. (1999) *Biochemistry* **38**, 13684–13691
402b. Roymoulik, I., Chen, H.-P., and Marsh, E. N. G. (1999) *J. Biol. Chem.* **274**, 11619–11622
403. Lowe, J. N., and Ingraham, L. L. (1971) *J. Am. Chem. Soc.* **93**, 3801–3802
404. He, M., and Dowd, P. (1998) *J. Am. Chem. Soc.* **120**, 1133–1137
405. Mancia, F., Keep, N. H., Nakagawa, A., Leadlay, P. F., McSweeney, S., Rasmussen, B., Bösecke, P., Diat, O., and Evans, P. R. (1996) *Structure* **4**, 339–350
405a. Mancia, F., Smith, G. A., and Evans, P. R. (1999) *Biochemistry* **38**, 7999–8005
405b. Thomä, N. H., Meier, T. W., Evans, P. R., and Leadlay, P. F. (1998) *Biochemistry* **37**, 14386–14393
405c. Smith, D. M., Golding, B. T., and Radom, L. (1999) *J. Am. Chem. Soc.* **121**, 9388–9399
405d. Maiti, N., Widjaja, L., and Banerjee, R. (1999) *J. Biol. Chem.* **274**, 32733–32737
406. Reitzer, R., Gruber, K., Jogl, G., Wagner, V. G., Bothe, H., Buckel, W., and Kratky, C. (1999) *Structure* **7**, 891–902
406a. Champloy, F., Jogl, G., Reitzer, R., Buckel, W., Bothe, H., Beatrix, B., Broeker, G., Michalowicz, A., Meyer-Klaucke, W., and Kratky, C. (1999) *J. Am. Chem. Soc.* **121**, 11780–11789
407. Stadtman, T. C. (1971) *Science* **171**, 859–867
408. Zagalak, B., Frey, P. A., Karabatsos, G. L., and Abeles, R. H. (1966) *J. Biol. Chem.* **241**, 3028–3035
409. Sprecher, M., Clark, M. J., and Sprinson, D. B. (1966) *J. Biol. Chem.* **241**, 872–877
410. Baker, J. R., and Stadtman, T. C. (1982) in B_{12}, Vol. II (Dolphin, D., ed), pp. 203–232, Wiley, New York
411. Kunz, F., Retéy, J., Arigoni, D., Tsai, L., and Stadtman, T. C. (1978) *Helv. Chim. Acta* **61**, 1139–1145
411a. Chang, C. H., and Frey, P. A. (2000) *J. Biol. Chem.* **275**, 106–114
412. Poston, J. M. (1977) *Science* **195**, 301–302
413. Stabler, S. P., Lindenbaum, J., and Allen, R. H. (1988) *J. Biol. Chem.* **263**, 5581–5588
414. Ballinger, M. D., Frey, P. A., Reed, G. H., and LoBrutto, R. (1995) *Biochemistry* **34**, 10086–10093
415. Wu, W., Lieder, K. W., Reed, G. H., and Frey, P. A. (1995) *Biochemistry* **34**, 10532–10537
415a. Wu, W., Booker, S., Lieder, K. W., Bandarian, V., Reed, G. H., and Frey, P. A. (2000) *Biochemistry* **39**, 9561–9570
416. Zhou, Z. S., Peariso, K., Penner-Hahn, J. E., and Matthews, R. G. (1999) *Biochemistry* **38**, 15915–15926
417. Yamanishi, M., Yamada, S., Muguruma, H., Murakami, Y., Tobimatsu, T., Ishida, A., Yamauchi, J., and Toraya, T. (1998) *Biochemistry* **37**, 4799–4803
418. Hall, D. A., Jordan-Starck, T. C., Loo, R. O., Ludwig, M. L., and Matthews, R. G. (2000) *Biochemistry* **39**, 10711–10719
418a. Jarrett, J. T., Amaratunga, M., Drennan, C. L., Scholten, J. D., Sands, R. H., Ludwig, M. L., and Matthews, R. G. (1996) *Biochemistry* **35**, 2464–2475
419. Zydowsky, T. M., Courtney, L. F., Frasca, V., Kobayashi, K., Shimizu, H., Yuen, L., Matthews, R. G., Benkovic, S. J., and Floss, H. G. (1986) *J. Am. Chem. Soc.* **108**, 3152–3153
420. Wood, J. M. (1982) in B_{12}, Vol. II (Dolphin, D., ed), pp. 151–164, Wiley, New York
421. Wood, J. M. (1974) *Science* **183**, 1049–1052
422. McBride, B. C., and Wolfe, R. S. (1971) *Biochemistry* **10**, 4312–4317
423. Parker, D. J., Wood, H. G., Ghambeer, R. K., and Ljungdahl, L. G. (1972) *Biochemistry* **11**, 3074–3080
424. Thauer, R. K., Diekert, G., and Schonheit, P. (1980) *Trends Biochem. Sci.* **5**, 304–306
425. Nielson, F. H. (1974) in *Trace Element Metabolism in Animals-2* (Hoekstra, W. G., Suttie, J. W., Ganther, H. E., and Mertz, W., eds), pp. 381–395, Univ. Park Press, Baltimore, Maryland
426. Schnegg, A., and Kirchgessner, M. (1976) *Int. J. Vitamins Nutr. Res.* **46**, 96–99
427. Nielsen, F. H. (1991) *FASEB J.* **5**, 2661–2667
428. Patel, S. U., Sadler, P. J., Tucker, A., and Viles, J. H. (1993) *J. Am. Chem. Soc.* **115**, 9285–9286
429. Nomoto, S., McNeely, M. D., and Sunderman, F. W., Jr. (1971) *Biochemistry* **10**, 1647–1651
430. Severne, B. C. (1974) *Nature (London)* **248**, 807–808
431. Eskew, D. L., Welch, R. M., and Cary, E. E. (1983) *Science* **222**, 621–623
432. Eitinger, T., and Friedrich, B. (1991) *J. Biol. Chem.* **266**, 3222–3227
432a. Eitinger, T., Degen, O., Böhnke, U., and Müller, M. (2000) *J. Biol. Chem.* **275**, 18029–18033
433. Frieden, E., ed. (1984) *Biochemistry of the Essential Ultratrace Elements*, Plenum, New York
434. Walsh, C. T., and Orme-Johnson, W. H. (1987) *Biochemistry* **26**, 4901–4906
435. Todd, M. J., and Hausinger, R. P. (1987) *J. Biol. Chem.* **262**, 5963–5967
436. Jabri, E., Carr, M. B., Hausinger, R. P., and Karplus, P. A. (1995) *Science* **268**, 998–1003
437. Lippard, S. J. (1995) *Science* **268**, 996–997
437a. Barrios, A. M., and Lippard, S. J. (1999) *J. Am. Chem. Soc.* **121**, 11751–11757
438. Park, I.-S., and Hausinger, R. P. (1995) *Science* **267**, 1156–1158

References

439. Lee, M. H., Pankratz, H. S., Wang, S., Scott, R. A., Finnegan, M. G., Johnson, M. K., Ippolito, J. A., Christianson, D. W., and Hausinger, R. P. (1993) *Protein Sci.* **2**, 1042–1052
439a. Doolittle, R. F. (1997) *Nature (London)* **388**, 515–516
440. Winkler, R. G., Blevins, D. G., Polacco, J. C., and Randall, D. D. (1988) *Trends Biochem. Sci.* **13**, 97–100
441. Mayhew, S. G., and O'Connor, M. E. (1982) *Trends Biochem. Sci.* **7**, 18–21
442. Elsden, S. R. (1981) *Trends Biochem. Sci.* **6**, 252–253
443. Fu, W., Drozdzewski, P. M., Morgan, T. V., Mortenson, L. E., Juszczak, A., Adams, M. W. W., He, S.-H., Peck, H. D., Jr., DerVartanian, D. V., LeGall, J., and Johnson, M. K. (1993) *Biochemistry* **32**, 4813–4819
444. Bagley, K. A., Van Garderen, C. J., Chen, M., Duin, E. C., Albracht, S. P. J., and Woodruff, W. H. (1994) *Biochemistry* **33**, 9229–9236
445. He, S. H., Teixeira, M., LeGall, J., Patil, D. S., Moura, I., Moura, J. J. G., DerVartanian, D. V., Huynh, B. H., and Peck, H. D., Jr. (1989) *J. Biol. Chem.* **264**, 2678–2682
446. Sorgenfrei, O., Duin, E. C., Klein, A., and Albracht, S. P. J. (1996) *J. Biol. Chem.* **271**, 23799–23806
447. Arp, D. J., and Burris, R. H. (1981) *Biochemistry* **20**, 2234–2240
448. Seefeldt, L. C., and Arp, D. J. (1989) *Biochemistry* **28**, 1588–1596
449. Telser, J., Benecky, M. J., Adams, M. W. W., Mortensen, L. E., and Hoffman, B. M. (1986) *J. Biol. Chem.* **261**, 13536–13541
450. Volbeda, A., Charon, M.-H., Piras, C., Hatchikian, E. C., Frey, M., and Fontecilla-Camps, J. C. (1995) *Nature (London)* **373**, 580–587
451. Cammack, R. (1995) *Nature (London)* **373**, 556–557
452. Dole, F., Fournel, A., Magro, V., Hatchikian, E. C., Bertrand, P., and Guigliarelli, B. (1997) *Biochemistry* **36**, 7847–7854
453. Volbeda, A., Garcin, E., Piras, C., de Lacey, A. L., Fernandez, V. M., Hatchikian, E. C., Frey, M., and Fontecilla-Camps, J. C. (1996) *J. Am. Chem. Soc.* **118**, 12989–12996
453b. Davidson, G., Choudhury, S. B., Gu, Z., Bose, K., Roseboom, W., Albracht, S. P. J., and Maroney, M. J. (2000) *Biochemistry* **39**, 7468–7479
453a. Amara, P., Volbeda, A., Fontecilla-Camps, J. C., and Field, M. J. (1999) *J. Am. Chem. Soc.* **121**, 4468–4477
454. Tau, S. L., Fox, J. A., Kojima, N., Walsh, C. T., and Orme-Johnson, W. H. (1984) *J. Am. Chem. Soc.* **106**, 3064–3066
455. Won, H., Olson, K. D., Wolfe, R. S., and Summers, M. F. (1990) *J. Am. Chem. Soc.* **112**, 2178–2184
456. Telser, J., Fann, Y.-C., Renner, M. W., Fajer, J., Wang, S., Zhang, H., Scott, R. A., and Hoffman, B. M. (1997) *J. Am. Chem. Soc.* **119**, 733–743
456a. Telser, J., Horng, Y.-C., Becker, D. F., Hoffman, B. M., and Ragsdale, S. W. (2000) *J. Am. Chem. Soc.* **122**, 182–183
457. Varadarajan, R., and Richards, F. M. (1992) *Biochemistry* **31**, 12315–12327
458. Ermler, U., Grabarse, W., Shima, S., Goubeaud, M., and Thauer, R. K. (1997) *Science* **278**, 1457–1462
458a. Selmer, T., Kahnt, J., Goubeaud, M., Shima, S., Grabarse, W., Ermler, U., and Thauer, R. K. (2000) *J. Biol. Chem.* **275**, 3755–3760
458b. Grabarse, W., Mahlert, F., Shima, S., Thauer, R. K., and Ermler, U. (2000) *J. Mol. Biol.* **303**, 329–344
459. Pfaltz, A., Jaun, B., Fässler, A., Eschenmoser, A., Jaenchen, R., Gilles, H. H., Diekert, G., and Thauer, R. K. (1982) *Helv. Chim. Acta* **65**, 828–865
460. Sings, H. L., Bible, K. C., and Rinehart, K. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10560–10565
461. Ensign, S. A. (1995) *Biochemistry* **34**, 5372–5381
462. Xia, J., Sinclair, J. F., Baldwin, T. O., and Lindahl, P. A. (1996) *Biochemistry* **35**, 1965–1971
463. Seravalli, J., Kumar, M., Lu, W.-P., and Ragsdale, S. W. (1997) *Biochemistry* **36**, 11241–11251
464. Xia, J., Hu, Z., Popescu, C. V., Lindahl, P. A., and Münck, E. (1997) *J. Am. Chem. Soc.* **119**, 8301–8312
465. Anderson, M. E., and Lindahl, P. A. (1996) *Biochemistry* **35**, 8371–8380
465a. Fraser, D. M., and Lindahl, P. A. (1999) *Biochemistry* **38**, 15706–15711
466. Grahame, D. A., and DeMoll, E. (1996) *J. Biol. Chem.* **271**, 8352–8358
467. Menon, S., and Ragsdale, S. W. (1998) *Biochemistry* **37**, 5689–5698
467a. Murakami, E., and Ragsdale, S. W. (2000) *J. Biol. Chem.* **275**, 4699–4707
468. Lebertz, H., Simon, H., Courtney, L. F., Benkovic, S. J., Zydowsky, L. D., Lee, K., and Floss, H. G. (1987) *J. Am. Chem. Soc.* **109**, 3173–3174
469. Raybuck, S. A., Bastian, N. R., Zydowsky, L. D., Kobayashi, K., Floss, H. G., Orme-Johnson, W. H., and Walsh, C. T. (1987) *J. Am. Chem. Soc.* **109**, 3171–3173
470. Howell, J. M., and Gawthorne, J. M., eds. (1987) *Copper in Animals and Man*, Vols. I and II, CRC Press, Boca, Raton, Florida
471. Frieden, E. (1968) *Sci. Am.* **218**(May), 103–114
472. Spiro, T. G., ed. (1981) *Copper Proteins*, Wiley, New York
473. Sigel, H., ed. (1981) *Metal Ions in Biological Systems*, Vol. 13, (Copper Proteins) Dekker, New York
474. Linder, M. C. (1991) *Biochemistry of Copper*, Plenum, New York
474a. Turnland, J. R. (1999) in *Modern Nutrition in Health and Disease*, 9th ed. (Shils, M. E., Olson, J. A., Shike, M., and Ross, A. C., eds), pp. 241–252, Williams & Wilkins, Baltimore, Maryland
475. Calabrese, L., Carbonaro, M., and Musci, G. (1988) *J. Biol. Chem.* **263**, 6480–6483
476. Yang, F., Friedrichs, W. E., Cupples, R. L., Bonifacio, M. J., Sanford, J. A., Horton, W. A., and Bowman, B. H. (1990) *J. Biol. Chem.* **265**, 10780–10785
477. Mukhopadhyay, C. K., Mazumder, B., Lindley, P. F., and Fox, P. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11546–11551
478. Hassett, R., and Kosman, D. J. (1995) *J. Biol. Chem.* **270**, 128–134
479. Georgatsou, E., Mavrogiannis, L. A., and Fragiadakis, G. S. (1997) *J. Biol. Chem.* **272**, 13786–13792
480. Glerum, D. M., Shtanko, A., and Tzagoloff, A. (1996) *J. Biol. Chem.* **271**, 20531–20535
481. Askwith, C., and Kaplan, J. (1998) *Trends Biochem. Sci.* **23**, 135–138
482. Yamaguchi-Iwai, Y., Serpe, M., Haile, D., Yang, W., Kosman, D. J., Klausner, R. D., and Dancis, A. (1997) *J. Biol. Chem.* **272**, 17711–17718
483. Zhu, Z., Labbé, S., Pena, M. M. O., and Thiele, D. J. (1998) *J. Biol. Chem.* **273**, 1277–1280
484. Kampfenkel, K., Kushnir, S., Babiychuk, E., Inzé, D., and Van Montagu, M. (1995) *J. Biol. Chem.* **270**, 28479–28486
485. Danks, D. M. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2211–2235, McGraw-Hill, New York
486. Lutsenko, S., and Cooper, M. J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6004–6009
487. Nielson, K. B., and Winge, D. R. (1984) *J. Biol. Chem.* **259**, 4941–4946
488. Hamer, D. H., Thiele, D. J., and Lemontt, J. E. (1985) *Science* **228**, 685–690
489. Holtzman, N. A. (1976) *Fed. Proc.* **35**, 2276–2280
490. Riordan, J. R., and Jolicoeur-Paquet, L. (1982) *J. Biol. Chem.* **257**, 4639–4645
491. Peltonen, L., Kuivaniemi, H., Palotie, A., Horn, N., Kaitila, I., and Kivirikko, K. I. (1983) *Biochemistry* **22**, 6156–6163
492. Hung, I. H., Suzuki, M., Yamaguchi, Y., Yuan, D. S., Klausner, R. D., and Gitlin, J. D. (1997) *J. Biol. Chem.* **272**, 21461–21466
493. Payne, A. S., and Gitlin, J. D. (1998) *J. Biol. Chem.* **273**, 3765–3770
494. Zhou, B., and Gitschier, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7481–7486
495. Lutsenko, S., Petrukhin, K., Cooper, M. J., Gilliam, C. T., and Kaplan, J. H. (1997) *J. Biol. Chem.* **272**, 18939–18944
495a. La Fontaine, S., Firth, S. D., Camakaris, J., Englezou, A., Theophilos, M. B., Petris, M. J., Howie, M., Lockhart, P. J., Greenough, M., Brooks, H., Reddel, R. R., and Mercer, J. F. B. (1998) *J. Biol. Chem.* **273**, 31375–31380
495b. Forbes, J. R., Hsi, G., and Cox, D. W. (1999) *J. Biol. Chem.* **274**, 12408–12413
495c. Cobine, P. A., George, G. N., Winzor, D. J., Harrison, M. D., Mogahaddas, S., and Dameron, C. T. (2000) *Biochemistry* **39**, 6857–6863
496. Yamaguchi, Y., Heiny, M. E., Suzuki, M., and Gitlin, J. D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14030–14035
497. Yuan, D. S., Stearman, R., Dancis, A., Dunn, T., Beeler, T., and Klausner, R. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2632–2636
498. Phung, L. T., Ajlani, G., and Haselkorn, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9651–9654
499. Klomp, L. W. J., Lin, S.-J., Yuan, D. S., Klausner, R. D., Culotta, V. C., and Gitlin, J. D. (1997) *J. Biol. Chem.* **272**, 9221–9226
500. Rae, T. D., Schmidt, P. J., Pufahl, R. A., Culotta, V. C., and O'Halloran, T. V. (1999) *Science* **284**, 805–808
500a. O'Halloran, T. V., and Culotta, V. C. (2000) *J. Biol. Chem.* **275**, 25057–25060
501. Valentine, J. S., and Gralla, E. B. (1997) *Science* **278**, 817–818
501a. Harrison, M. D., Jones, C. E., Solioz, M., and Dameron, C. T. (2000) *Trends Biochem. Sci.* **25**, 29–32
502. Hoitnik, C. W. G., Driscoll, P. C., Hill, H. A. O., and Canters, G. W. (1994) *Biochemistry* **33**, 3560–3571
503. Romero, A., Nar, H., Huber, R., Messerschmidt, A., Kalverda, A. P., Canters, G. W., Durley, R., and Mathews, F. S. (1994) *J. Mol. Biol.* **236**, 1196–1211
504. Kalverda, A. P., Wymenga, S. S., Lomman, A., van de Ven, F. J. M., Hilbers, C. W., and Canters, G. W. (1994) *J. Mol. Biol.* **240**, 358–371
505. Guss, J. M., Merritt, E. A., Phizackerley, R. P., and Freeman, H. C. (1996) *J. Mol. Biol.* **262**, 686–705
506. Garrett, T. P. J., Clingeleffer, D. J., Guss, J. M., Rogers, S. J., and Freeman, H. C. (1984) *J. Biol. Chem.* **259**, 2822–2825
507. Ryde, U., Olsson, M. H. M., Pierloot, K., and Roos, B. O. (1996) *J. Mol. Biol.* **261**, 586–596
508. Guckert, J. A., Lowery, M. D., and Solomon, E. I. (1995) *J. Am. Chem. Soc.* **117**, 2817–2844

509. Groeneveld, C. M., and Canters, G. W. (1988) *J. Biol. Chem.* **263**, 167 – 173
510. Karlsson, B. G., Tsai, L.-C., Nar, H., Sanders-Loehr, J., Bonander, N., Langer, V., and Sjölin, L. (1997) *Biochemistry* **36**, 4089 – 4095
511. van de Kamp, M., Canters, G. W., Wijmenga, S. S., Lommen, A., Hilbers, C. W., Nar, H., Messerschmidt, A., and Huber, R. (1992) *Biochemistry* **31**, 10194 – 10207
512. Messerschmidt, A., Prade, L., Kroes, S. J., Sanders-Loehr, J., Huber, R., and Canters, G. W. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3443 – 3448
513. Farver, O., Licht, A., and Pecht, I. (1987) *Biochemistry* **26**, 7317 – 7321
514. Vila, A. J., and Fernández, C. O. (1996) *J. Am. Chem. Soc.* **118**, 7291 – 7298
515. Casimiro, D. R., Toy-Palmer, A., Blake, R. C., II, and Dyson, H. J. (1995) *Biochemistry* **34**, 6640 – 6648
516. Grossmann, J. G., Engledew, W. J., Harvey, I., Strange, R. W., and Hasnain, S. S. (1995) *Biochemistry* **34**, 8406 – 8414
517. Botuyan, M. V., Toy-Palmer, A., Chung, J., Blake, R. C., II, Beroza, P., Case, D. A., and Dyson, H. J. (1996) *J. Mol. Biol.* **263**, 752 – 767
518. Hildebrandt, P., Matysik, J., Schrader, B., Scharf, B., and Engelhard, M. (1994) *Biochemistry* **33**, 11426 – 11431
519. McManus, J. D., Brune, D. C., Han, J., Sanders-Loehr, J., Meyer, T. E., Cusanovich, M. A., Tollin, G., and Blankenship, R. E. (1992) *J. Biol. Chem.* **267**, 6531 – 6540
520. Penfield, K. W., Gerwith, A. A., and Solomon, E. I. (1985) *J. Am. Chem. Soc.* **107**, 4519 – 4529
521. Blair, D. F., Campbell, G. W., Schoonover, J. R., Chan, S. I., Gray, H. B., Malmstrom, B. G., Pecht, I., Swanson, B. I., Woodruff, W. H., Cho, W. K., English, A. M., Fry, H. A., Lum, V., and Norton, K. A. (1985) *J. Am. Chem. Soc.* **107**, 5755 – 5766
521a. Robinson, H., Ang, M. C., Gao, Y.-G., Hay, M. T., Lu, Y., and Wang, A. H.-J. (1999) *Biochemistry* **38**, 5677 – 5683
521b. Holz, R. C., Bennett, B., Chen, G., and Ming, L.-J. (1998) *J. Am. Chem. Soc.* **120**, 6329 – 6335
522. Fridovich, I. (1995) *Ann. Rev. Biochem.* **64**, 97 – 112
523. Beck, B. L., Tabatabai, L. B., and Mayfield, J. E. (1990) *Biochemistry* **29**, 372 – 376
524. Tainer, J. A., Getzoff, E. D., Richardson, J. S., and Richardson, D. C. (1983) *Nature (London)* **306**, 284 – 287
525. Pesce, A., Capasso, C., Battistoni, A., Folcarelli, S., Rotilio, G., Desideri, A., and Bolognesi, M. (1997) *J. Mol. Biol.* **274**, 408 – 420
526. Rypniewski, W. R., Mangani, S., Bruni, B., Orioli, P. L., Casati, M., and Wilson, K. S. (1995) *J. Mol. Biol.* **251**, 282 – 296
527. Stenlund, P., Andersson, D., and Tibell, L. A. E. (1997) *Protein Sci.* **6**, 2350 – 2358
528. Ogihara, N. L., Parge, H. E., Hart, P. J., Weiss, M. S., Goto, J. J., Crane, B. R., Tsang, J., Slater, K., Roe, J. A., Valentine, J. S., Eisenberg, D., and Tainer, J. A. (1996) *Biochemistry* **35**, 2316 – 2321
529. Beyer, W. F., Jr., Fridovich, I., Mullenbach, G. T., and Hallewell, R. (1987) *J. Biol. Chem.* **262**, 11182 – 11187
530. Banci, L., Bertini, I., Bauer, D., Hallewell, R. A., and Viezzoli, M. S. (1993) *Biochemistry* **32**, 4384 – 4388
531. Getzoff, E. D., Cabelli, D. E., Fisher, C. L., Parge, H. E., Viezzoli, M. S., Banci, L., and Hallewell, R. A. (1992) *Nature (London)* **358**, 347 – 351
532. Afanas'ev, I. B. (1989,1991) *Superoxide Ion: Chemistry and Biological Implications*, Vol. 2, CRC Press, Boca Raton, Florida
533. Hiraishi, H., Terano, A., Razandi, M., Sugimoto, T., Harada, T., and Ivey, K. J. (1992) *J. Biol. Chem.* **267**, 14812 – 14817
534. Deng, H.-X., Hentati, A., Tainer, J. A., Iqbal, Z., Cayabyab, A., Hung, W.-Y., Getzoff, E. D., Hu, P., Herzfeldt, B., Roos, R. P., Warner, C., Deng, G., Soriano, E., Smyth, C., Parge, H. E., Ahmed, A., Roses, A. D., Hallewell, R. A., Pericak-Vance, M. A., and Siddique, T. (1993) *Science* **261**, 1047 – 1051
535. McNamara, J. O., and Fridovich, I. (1993) *Nature (London)* **362**, 20 – 21
536. Wiedau-Pazos, M., Goto, J. J., Rabizadeh, S., Gralla, E. B., Roe, J. A., Lee, M. K., Valentine, J. S., and Bredesen, D. E. (1996) *Science* **271**, 515 – 518
537. Adman, E. T., Godden, J. W., and Turley, S. (1995) *J. Biol. Chem.* **270**, 27458 – 27474
538. Suzuki, S., Kohzuma, T., Deligeer, Yamaguchi, K., Nakamura, N., Shidara, S., Kobayashi, K., and Tagawa, S. (1994) *J. Am. Chem. Soc.* **116**, 11145 – 11146
539. Murphy, M. E. P., Turley, S., and Adman, E. T. (1997) *J. Biol. Chem.* **272**, 28455 – 28460
540. Kukimoto, M., Nishiyama, M., Murphy, M. E. P., Turley, S., Adman, E. T., Horinouchi, S., and Beppu, T. (1994) *Biochemistry* **33**, 5246 – 5252
540a. Strange, R. W., Murphy, L. M., Dodd, F. E., Abraham, Z. H. L., Eady, R. R., Smith, B. E., and Hasnain, S. S. (1999) *J. Mol. Biol.* **287**, 1001 – 1009
541. Howes, B. D., Abraham, Z. H. L., Lowe, D. J., Brüser, T., Eady, R. R., and Smith, B. E. (1994) *Biochemistry* **33**, 3171 – 3177
542. Murphy, M. E. P., Turley, S., Kukimoto, M., Nishiyama, M., Horinouchi, S., Sasaki, H., Tanokura, M., and Adman, E. T. (1995) *Biochemistry* **34**, 12107 – 12117
543. Grossmann, J. G., Abraham, Z. H. L., Adman, E. T., Neu, M., Eady, R. R., Smith, B. E., and Hasnain, S. S. (1993) *Biochemistry* **32**, 7360 – 7366
544. Peters Libeu, C. A., Kukimoto, M., Nishiyama, M., Horinouchi, S., and Adman, E. T. (1997) *Biochemistry* **36**, 13160 – 13179
544a. Inoue, T., Nishio, N., Suzuki, S., Kataoka, K., Kohzuma, T., and Kai, Y. (1999) *J. Biol. Chem.* **274**, 17845 – 17852
545. Veselov, A., Olesen, K., Sienkiewicz, A., Shapleigh, J. P., and Scholes, C. P. (1998) *Biochemistry* **37**, 6095 – 6105
546. Neese, F., Zumft, W. G., Antholine, W. E., and Kroneck, P. M. H. (1996) *J. Am. Chem. Soc.* **118**, 8692 – 8699
546a. Rasmussen, T., Berks, B. C., Sanders-Loehr, J., Dooley, D. M., Zumft, W. G., and Thomson, A. J. (2000) *Biochemistry* **39**, 12753 – 12756
547. Gaykema, W. P. J., Hol, W. G. J., Vereijken, J. M., Soeter, N. M., Bak, H. J., and Beintema, J. J. (1984) *Nature (London)* **309**, 23 – 29
548. Cuff, M. E., Miller, K. I., van Holde, K. E., and Hendrickson, W. A. (1998) *J. Mol. Biol.* **278**, 855 – 870
549. Hazes, B., Magnus, K. A., Bonaventura, C., Bonaventura, J., Dauter, Z., Kalk, K. H., and Hol, W. G. J. (1993) *Protein Sci.* **2**, 597 – 619
550. Solomon, E. I., and Lowery, M. D. (1993) *Science* **259**, 1575 – 1581
551. Kitajima, N., Fujisawa, K., Fujimoto, C., Moro-oka, Y., Hashimoto, S., Kitagawa, T., Toriumi, K., Tatsumi, K., and Nakamura, A. (1992) *J. Am. Chem. Soc.* **114**, 1277 – 1291
552. Tian, G., and Klinman, J. P. (1993) *J. Am. Chem. Soc.* **115**, 8891 – 8897
553. Gahmberg, C. G. (1976) *J. Biol. Chem.* **251**, 510 – 515
554. Driscoll, J. J., and Kosman, D. J. (1987) *Biochemistry* **26**, 3429 – 3436
555. Clark, K., Penner-Hahn, J. E., Whittaker, M., and Whittaker, J. W. (1994) *Biochemistry* **33**, 12553 – 12557
556. Wachter, R. M., and Branchaud, B. P. (1996) *Biochemistry* **35**, 14425 – 14435
557. Halfen, J. A., Jazdzewski, B. A., Mahapatra, S., Berreau, L. M., Wilkinson, E. C., Que, L., Jr., and Tolman, W. B. (1997) *J. Am. Chem. Soc.* **119**, 8217 – 8227
558. Ito, N., Phillips, S. E. V., Yadav, K. D. S., and Knowles, P. F. (1994) *J. Mol. Biol.* **238**, 794 – 814
559. Baron, A. J., Stevens, C., Wilmot, C., Seneviratne, K. D., Blakeley, V., Dooley, D. M., Phillips, S. E. V., Knowles, P. F., and McPherson, M. J. (1994) *J. Biol. Chem.* **269**, 25095 – 25105
560. Whittaker, M. M., Kersten, P. J., Nakamura, N., Sanders-Loehr, J., Schweizer, E. S., and Whittaker, J. W. (1996) *J. Biol. Chem.* **271**, 681 – 687
560a. Rogers, M. S., Baron, A. J., McPherson, M. J., Knowles, P. F., and Dooley, D. M. (2000) *J. Am. Chem. Soc.* **122**, 990 – 991
561. Dooley, D. M., Scott, R. A., Knowles, P. F., Colangelo, C. M., McGuirl, M. A., and Brown, D. E. (1998) *J. Am. Chem. Soc.* **120**, 2599 – 2605
561a. Itoh, S., Taniguchi, M., Takada, N., Nagatomo, S., Kitagawa, T., and Fukuzumi, S. (2000) *J. Am. Chem. Soc.* **122**, 12087 – 12097
562. Scott, R. A., and Dooley, D. M. (1985) *J. Am. Chem. Soc.* **107**, 4348 – 4350
563. Suzuki, S., Sakurai, T., and Nakahara, A. (1986) *Biochemistry* **25**, 338 – 341
563a. Dove, J. E., Schwartz, B., Williams, N. K., and Klinman, J. P. (2000) *Biochemistry* **39**, 3690 – 3698
564. Williamson, P. R., and Kagan, H. M. (1987) *J. Biol. Chem.* **262**, 8196 – 8201
565. Uwajima, T., Shimizu, Y., and Terada, O. (1984) *J. Biol. Chem.* **259**, 2748 – 2753
566. Wachter, R. M., Montague-Smith, M. P., and Branchaud, B. P. (1997) *J. Am. Chem. Soc.* **119**, 7743 – 7749
567. Whittaker, M. M., Ballou, D. P., and Whittaker, J. W. (1998) *Biochemistry* **37**, 8426 – 8436
568. Toussaint, O., and Lerch, K. (1987) *Biochemistry* **26**, 8567 – 8571
569. Wilcox, D. E., Porras, A. G., Hwang, Y. T., Lerch, K., Winkler, M. E., and Solomon, E. I. (1985) *J. Am. Chem. Soc.* **107**, 4015 – 4027
569a. Decker, H., and Tuczek, F. (2000) *Trends Biochem. Sci.* **25**, 392 – 397
570. Prigge, S. T., Kolhekar, A. S., Eipper, B. A., Mains, R. E., and Amzel, L. M. (1997) *Science* **278**, 1300 – 1305
571. Menniti, F. S., Knoth, J., Peterson, D. S., and Diliberto, E. J., Jr. (1987) *J. Biol. Chem.* **262**, 7651 – 7657
572. Tian, G., Berry, J. A., and Klinman, J. P. (1994) *Biochemistry* **33**, 226 – 234
573. Meyer, T. E., Marchesini, A., Cusanovich, M. A., and Tollin, G. (1991) *Biochemistry* **30**, 4619 – 4623
574. Messerschmidt, A., Luecke, H., and Huber, R. (1993) *J. Mol. Biol.* **230**, 997 – 1014
575. Woolery, G. L., Powers, L., Peisach, J., and Spiro, T. G. (1984) *Biochemistry* **23**, 3428 – 3434
576. Farver, O., Wherland, S., and Pecht, I. (1994) *J. Biol. Chem.* **269**, 22933 – 22936
577. Palmer, A. E., Randall, D. W., Xu, F., and Solomon, E. I. (1999) *J. Am. Chem. Soc.* **121**, 7138 – 7149
578. Lehn, J.-M., Malmström, B. G., Selin, E., and Öblad, M. (1986) *Trends Biochem. Sci.* **11**, 228 – 230
579. Keen, C. L., Lonnerdal, B., and Hurley, L. S. (1984) in *Biochemistry of the Essential Ultratrace Elements* (Frieden, E., ed), pp. 89 – 132, Plenum, New York

References

580. Schramm, V. L., and Wedler, F. C., eds. (1986) *Manganese in Metabolism and Enzyme Function*, Academic Press, Orlando, Florida
581. Bartsevich, V. V., and Pakrasi, H. B. (1996) *J. Biol. Chem.* **271**, 26057– 26061
582. Baker, A. P., Griggs, L. J., Munro, J. R., and Finkelstein, J. A. (1973) *J. Biol. Chem.* **248**, 800 – 883
583. Goldman, A., Ollis, D. L., and Steitz, T. A. (1987) *J. Mol. Biol.* **194**, 147 – 153
584. Kanyo, Z. F., Scolnick, L. R., Ash, D. E., and Christianson, D. W. (1996) *Nature (London)* **383**, 554 – 557
585. Khangulov, S. V., Sossong, T. M., Jr., Ash, D. E., and Dismukes, G. C. (1998) *Biochemistry* **37**, 8539 – 8550
586. Zhou, G. W., Guo, J., Huang, W., Fletterick, R. J., and Scanlan, T. S. (1994) *Science* **265**, 1059 – 1064
587. Hirose, K., Longo, D. L., Oppenheim, J. J., and Matsushima, K. (1993) *FASEB J.* **7**, 361 – 368
588. MacMillan-Crow, L. A., Crow, J. P., and Thompson, J. A. (1998) *Biochemistry* **37**, 1613 – 1622
589. Boldt, Y. R., Whiting, A. K., Wagner, M. L., Sadowsky, M. J., Que, L., Jr., and Wackett, L. P. (1997) *Biochemistry* **36**, 2147 – 2153
590. Mildvan, A. S. (1974) *Ann. Rev. Biochem.* **43**, 357 – 399
591. Villafranca, J. J., and Wedler, F. C. (1974) *Biochemistry* **13**, 3286 – 3291
592. Goodman, M. F., Keener, S., Guidotti, S., and Branscomb, E. W. (1983) *J. Biol. Chem.* **258**, 3469 – 3475
593. Deuel, T. F., and Prusiner, S. (1974) *J. Biol. Chem.* **249**, 257 – 264
594. Schwartz, K., and Mertz, W. (1959) *Arch. Biochem. Biophys.* **85**, 292 – 295
595. Mertz, W. (1993) *J. Nutr. Sci. Vitaminol.* **123**, 626 – 633
596. Davis, C. M., Sumrall, K. H., and Vincent, J. B. (1996) *Biochemistry* **35**, 12963 – 12969
597. Saner, G. (1980) *Chromium in Nutrition and Disease*, Liss, New York
598. Frieden, E. (1972) *Sci. Am.* **227**(Jul), 52 – 60
599. Anderson, R. A., Bryden, N. A., and Canary, J. J. (1991) *Am. J. Clin. Nutr.* **51**, 864 – 868
600. Davis, C. M., and Vincent, J. B. (1997) *Arch. Biochem. Biophys.* **339**, 335 – 343
601. O'Dell, B. L., and Campbell, B. J. (1970) *Comprehensive Biochemistry* **21**, 179 – 226
602. Zhang, L., and Lay, P. A. (1996) *J. Am. Chem. Soc.* **118**, 12624 – 12637

602a. Thompson, K. H. (1999) *BioFactors* **10**, 43 – 51

603. Zhitkovich, A., Voitkun, V., and Costa, M. (1996) *Biochemistry* **35**, 7275 – 7282
604. Anderson, R. A. (1994) in *Risk Assessment of Essential Elements* (Merz, W., Abernathy, C. O., and Olin, S. S., eds), pp. 187 – 196, ISLI Press, Washington,D.C.
605. Conconi, A., Smerdon, M. J., Howe, G. A., and Ryan, C. A. (1996) *Nature (London)* **383**, 826 – 829
606. Stearns, D. M., Wise, J. P., Sr., Patierno, S. R., and Wetterhahn, K. E. (1995) *FASEB J.* **9**, 1643 – 1648
607. Stearns, D. M., BelBruno, J. J., and Wetterhahn, K. E. (1995) *FASEB J.* **9**, 1650 – 1657
608. Macara, I. G. (1980) *Trends Biochem. Sci.* **5**, 92 – 94
609. Nechay, B. R., Nanninga, L. B., Nechay, P. S. E., Post, R. L., Grantham, J. J., Macara, I. G., Kubena, L. F., Phillips, T. D., and Nielsen, F. H. (1986) *Fed. Proc.* **45**, 123 – 132
610. Chasteen, N. D. (1983) *Struct. Bonding* **53**, 104 – 138
611. Boyd, D. W., and Kustin, K. (1985) *Adv. Inorg. Biochem.* **6**, 311 – 365
612. Tullius, T. D., Gillum, W. O., Carlson, R. M. K., and Hodgson, K. O. (1980) *J. Am. Chem. Soc.* **102**, 5670 – 5676
613. Ryan, D. E., Grant, K. B., Nakanishi, K., Frank, P., and Hodgson, K. O. (1996) *Biochemistry* **35**, 8651 – 8661
614. Orgel, L. E. (1980) *Trends Biochem. Sci.* **5**, X (Aug.)
615. Ryan, D. E., Grant, K. B., and Nakanishi, K. (1996) *Biochemistry* **35**, 8640 – 8650
616. Chasteen, N. D., ed. (1990) *Vanadium in Biological Systems*, Kluwer Acad. Publ., Dordrecht, Germany
617. Winter, G. E. M., and Butler, A. (1996) *Biochemistry* **35**, 11805 – 11811
618. Wever, R., and Krenn, B. E. (1990) in *Vanadium in Biological Systems* (Chasteen, N. D., ed), pp. 81– 97, Kluwer Acad. Publ., Dordrecht, Germany

618a. Isupov, M. N., Dalby, A. R., Brindley, A. A., Izumi, Y., Tanabe, T., Murshudov, G. N., and Littlechild, J. A. (2000) *J. Mol. Biol.* **299**, 1035 – 1049

618b. Weyand, M., Hecht, H.-J., Kiess, M., Liaud, M.-F., Vilter, H., and Schomburg, D. (1999) *J. Mol. Biol.* **293**, 595 – 611

619. Haupts, U., Tittor, J., Bamberg, E., and Oesterhelt, D. (1997) *Biochemistry* **36**, 2 – 7
620. Lindquist, Y., Schneider, G., and Vihko, P. (1994) *Eur. J. Biochem.* **221**, 139 – 142
621. Hemrika, W., Renirie, R., Dekker, H. L., Barnett, P., and Wever, R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2145 – 2149
622. Chatterjee, R., Allen, R. M., Ludden, P. W., and Shah, V. K. (1997) *J. Biol. Chem.* **272**, 21604 – 21608
623. Zhang, M., Zhou, M., Van Etten, R. L., and Stauffacher, C. V. (1997) *Biochemistry* **36**, 15 – 23
624. Elberg, G., Li, J., and Shechter, Y. (1994) *J. Biol. Chem.* **269**, 9521 – 9527
625. Li, J., Elberg, G., Crans, D. C., and Shechter, Y. (1996) *Biochemistry* **35**, 8314 – 8318
626. Heyliger, C. E., Tahiliani, A. G., and McNeill, J. H. (1985) *Science* **227**, 1474 – 1476
627. Johnson, J. L., Rajagopalan, K. V., and Cohen, H. J. (1974) *J. Biol. Chem.* **249**, 859 – 866
628. Barberà, A., Rodríguez-Gil, J. E., and Guinovart, J. J. (1994) *J. Biol. Chem.* **269**, 20047 – 20053
629. Huyer, G., Liu, S., Kelly, J., Moffat, J., Payette, P., Kennedy, B., Tsaprailis, G., Gresser, M. J., and Ramachandran, C. (1997) *J. Biol. Chem.* **272**, 843 – 851
630. Li, J., Elberg, G., Gefel, D., and Shechter, Y. (1995) *Biochemistry* **34**, 6218 – 6225
631. Mikalsen, S.-O., and Kaalhus, O. (1998) *J. Biol. Chem.* **273**, 10036 – 10045
632. Hille, R. (1996) *Chem. Rev.* **96**, 2757 – 2816
633. Romão, M. J., Archer, M., Moura, I., Moura, J. J. G., LeGall, J., Engh, R., Schneider, M., Hof, P., and Huber, R. (1995) *Science* **270**, 1170 – 1176
634. Hille, R., and Nishino, T. (1995) *FASEB J.* **9**, 995 – 1003
635. Harris, C. M., and Massey, V. (1997) *J. Biol. Chem.* **272**, 8370 – 8379
636. Glatigny, A., Hof, P., Romao, M. J., Huber, R., and Scazzocchio, C. (1998) *J. Mol. Biol.* **278**, 431 – 438
637. Xiang, Q., and Edmondson, D. E. (1996) *Biochemistry* **35**, 5441 – 5450
638. Harris, C. M., and Massey, V. (1997) *J. Biol. Chem.* **272**, 28335 – 28341
639. Glatigny, A., and Scazzocchio, C. (1995) *J. Biol. Chem.* **270**, 3534 – 3550
640. Bläse, M., Bruntner, C., Tshisuaka, B., Fetzner, S., and Lingens, F. (1996) *J. Biol. Chem.* **271**, 23068 – 23079
641. Canne, C., Stephan, I., Finsterbusch, J., Lingens, F., Kappl, R., Fetzner, S., and Hüttermann, J. (1997) *Biochemistry* **36**, 9780 – 9790
642. Gladyshev, V. N., Khangulov, S. V., and Stadtman, T. C. (1996) *Biochemistry* **35**, 212 – 223
643. Huber, R., Hof, P., Duarte, R. O., Moura, J. J. G., Moura, I., Liu, M.-Y., LeGall, J., Hille, R., Archer, M., and Romao, M. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8846 – 8851
644. Ori, N., Eshed, Y., Pinto, P., Paran, I., Zamir, D., and Fluhr, R. (1997) *J. Biol. Chem.* **272**, 1019 – 1025
645. Garrett, R. M., Johnson, J. L., Graf, T. N., Feigenbaum, A., and Rajagopalan, K. V. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6394 – 6398
646. Johnson, J. L., and Wadman, S. K. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2271 – 2283, McGraw-Hill, New York

646a. Raitsimring, A. M., Pacheco, A., and Enemark, J. H. (1998) *J. Am. Chem. Soc.* **120**, 11263 – 11278

646b. Brody, M. S., and Hille, R. (1999) *Biochemistry* **38**, 6668 – 6677

647. Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K. V., and Rees, D. C. (1996) *Science* **272**, 1615 – 1621
648. Schneider, F., Löwe, J., Huber, R., Schindelin, H., Kisker, C., and Knäblein, J. (1996) *J. Mol. Biol.* **263**, 53 – 69

648a. Stewart, L. J., Bailey, S., Bennett, B., Charnock, J. M., Garner, C. D., and McAlpine, A. S. (2000) *J. Mol. Biol.* **299**, 593 – 600

648b. Rothery, R. A., Trieber, C. A., and Weiner, J. H. (1999) *J. Biol. Chem.* **274**, 13002 – 13009

648c. Temple, C. A., George, G. N., Hilton, J. C., George, M. J., Prince, R. C., Barber, M. J., and Rajagopalan, K. V. (2000) *Biochemistry* **39**, 4046 – 4052

648d. Li, H.-K., Temple, C., Rajagopalan, K. V., and Schindelin, H. (2000) *J. Am. Chem. Soc.* **122**, 7673 – 7680

649. Pollock, V. V., and Barber, M. J. (1997) *J. Biol. Chem.* **272**, 3355 – 3362
650. Rajagopalan, K. V. (1984) in *Biochemistry of the Essential Ultratrace Elements* (Freiden, E., ed), Plenum, New York
651. Kramer, S. P., Johnson, J. L., Ribeiro, A. A., Millington, D. S., and Rajagopalan, K. V. (1987) *J. Biol. Chem.* **262**, 16357 – 16363

651a. Rajagopalan, K. V., and Johnson, J. L. (1992) *J. Biol. Chem.* **267**, 10199 – 10202

651b. Luykx, D. M. A. M., Duine, J. A., and de Vries, S. (1998) *Biochemistry* **37**, 11366 – 11375

651c. Liu, M. T. W., Wuebbens, M. M., Rajagopalan, K. V., and Schindelin, H. (2000) *J. Biol. Chem.* **275**, 1814 – 1822

651d. Feng, G., Tintrup, H., Kirsch, J., Nichol, M. C., Kuhse, J., Betz, H., and Sanes, J. R. (1998) *Science* **282**, 1321 – 1324

652. Cramer, S. P., Solomonson, L. P., Adams, M. W. W., and Mortenson, L. E. (1984) *J. Am. Chem. Soc.* **106**, 1467 – 1471
653. Berg, J. M., and Holm, R. H. (1985) *J. Am. Chem. Soc.* **107**, 917 – 925
654. Cramer, S. P., Wahl, R., and Rajagopalan, K. V. (1981) *J. Am. Chem. Soc.* **103**, 7721 – 7727
655. Shah, V. K., Ugalde, R. A., Imperial, J., and Brill, W. J. (1984) *Ann. Rev. Biochem.* **53**, 231 – 257
656. Orme-Johnson, W. H. (1985) *Ann. Rev. Biophys. Biophys. Chem.* **14**, 419 – 459
657. Howes, B. D., Bray, R. C., Richards, R. L., Turner, N. A., Bennett, B., and Lowe, D. J. (1996) *Biochemistry* **35**, 1432 – 1443

657a. Caldeira, J., Belle, V., Asso, M., Guigliarelli, B., Moura, I., Moura, J. J. G., and Bertrand, P. (2000) *Biochemistry* **39**, 2700 – 2707

658. Rothery, R. A., Magalon, A., Giordano, G., Guigliarelli, B., Blasco, F., and Weiner, J. H. (1998) *J. Biol. Chem.* **273**, 7462 – 7469

References

659. Augier, V., Guigliarelli, B., Asso, M., Bertrand, P., Frixon, C., Giordano, G., Chippaux, M., and Blasco, F. (1993) *Biochemistry* **32**, 2013–2023

660. Magalon, A., Asso, M., Guigliarelli, B., Rothery, R. A., Bertrand, P., Giordano, G., and Blasco, F. (1998) *Biochemistry* **37**, 7363–7370

661. Campbell, W. H., and Kinghorn, J. R. (1990) *Trends Biochem. Sci.* **15**, 315–319

662. Garde, J., Kinghorn, J. R., and Tomsett, A. B. (1995) *J. Biol. Chem.* **270**, 6644–6650

663. Ratnam, K., Shiraishi, N., Campbell, W. H., and Hille, R. (1995) *J. Biol. Chem.* **270**, 24067–24072

663a. George, G. N., Mertens, J. A., and Campbell, W. H. (1999) *J. Am. Chem. Soc.* **121**, 9730 – 9731

664. Stadtman, T. C. (1991) *J. Biol. Chem.* **266**, 16257–16260

665. Berg, B. L., Baron, C., and Stewart, V. (1991) *J. Biol. Chem.* **266**, 22386–22391

666. Berg, B. L., Li, J., Heider, J., and Stewart, V. (1991) *J. Biol. Chem.* **266**, 22380–22385

667. Stewart, V. (1988) *Microbiol. Rev.* **52**, 190–232

668. Sawers, G., Heider, J., Zehelein, E., and Böck, A. (1991) *J. Bacteriol.* **173**, 4983–4993

669. Gladyshev, V. N., Boyington, J. C., Khangulov, S. V., Grahame, D. A., Stadtman, T. C., and Sun, P. D. (1996) *J. Biol. Chem.* **271**, 8095–8100

670. Boyington, J. C., Gladyshev, V. N., Khangulov, S. V., Stadtman, T. C., and Sun, P. D. (1997) *Science* **275**, 1305–1306

671. Khangulov, S. V., Gladyshev, V. N., Dismukes, G. C., and Stadtman, T. C. (1998) *Biochemistry* **37**, 3518–3528

672. George, G. N., Colangelo, C. M., Dong, J., Scott, R. A., Khangulov, S. V., Gladyshev, V. N., and Stadtman, T. C. (1998) *J. Am. Chem. Soc.* **120**, 1267–1273

673. Shuber, A. P., Orr, E. C., Recny, M. A., Schendel, P. F., May, H. D., Schauer, N. L., and Ferry, J. G. (1986) *J. Biol. Chem.* **261**, 12942–12947

674. Muller, U., Willnow, P., Ruschig, U., and Höpner, T. (1978) *Eur. J. Biochem.* **83**, 485–498

675. Rohde, M., Mayer, F., and Meyer, O. (1984) *J. Biol. Chem.* **259**, 14788–14792

676. Ribbe, M., Gadkari, D., and Meyer, O. (1997) *J. Biol. Chem.* **272**, 26627–26633

676a. Dobbek, H., Gremer, L., Meyer, O., and Huber, R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8884–8889

676b. Gremer, L., Kellner, S., Dobbek, H., Huber, R., and Meyer, O. (2000) *J. Biol. Chem.* **275**, 1864–1872

676c. Hänzelmann, P., Dobbek, H., Gremer, L., Huber, R., and Meyer, O. (2000) *J. Mol. Biol.* **301**, 1221–1235

677. Rajagopalan, K. V., and Johnson, J. L. (1992) *J. Biol. Chem.* **267**, 10199–10202

678. Wahl, R. C., Warner, C. K., Finnerty, V., and Rajagopalan, K. V. (1982) *J. Biol. Chem.* **257**, 3958–3962

679. Nagahara, N., Okazaki, T., and Nishino, T. (1995) *J. Biol. Chem.* **270**, 16230–16235

680. Nishino, T., Usami, C., and Tsushima, K. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1826–1829

681. Rech, S., Wolin, C., and Gunsalus, R. P. (1996) *J. Biol. Chem.* **271**, 2557–2562

682. Blanchard, C. Z., and Hales, B. J. (1996) *Biochemistry* **35**, 472–478

683. Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. W., and Rees, D. C. (1995) *Science* **267**, 1463–1469

683a. Hu, Y., Faham, S., Roy, R., Adams, M. W. W., and Rees, D. C. (1999) *J. Mol. Biol.* **286**, 899–914

684. Koehler, B. P., Mukund, S., Conover, R. C., Dhawan, I. K., Roy, R., Adams, M. W. W., and Johnson, M. K. (1996) *J. Am. Chem. Soc.* **118**, 12391–12405

685. Mukund, S., and Adams, M. W. W. (1995) *J. Biol. Chem.* **270**, 8389–8392

686. Johnson, M. K., Rees, D. C., and Adams, M. W. W. (1996) *Chem. Rev.* **96**, 2817–2839

687. Yadav, J., Das, S. K., and Sarkar, S. (1997) *J. Am. Chem. Soc.* **119**, 4315–4316

688. Juszczak, A., Aono, S., and Adams, M. W. W. (1991) *J. Biol. Chem.* **266**, 13834–13841

Study Questions

1. Describe one or more metabolic functions of ions or chelate complexes of ions derived from each of the following metallic elements: Ca, Mg, Fe, Cu, Ni, Co.

2. If the concentration of Cu, Zn-superoxide dismutase (SOD) in a yeast cell is 10 µM, the total copper (bound and free) is 70 µM, and the dissociation constant for loss of Cu^+ from SOD is 6 fM, what will be the concentration of free Cu^+ within the cell? If the cell volume is 10^{-14} liter, how many copper ions will be present in a single cell? See Roe *et al.* (1999) *Science* **284**, 805–808.

3. Outline the metabolic pathways that are utilized by acetic acid-producing bacteria (acetogens) in the stoichiometric conversion of one molecule of glucose into three molecules of acetic acid. Indicate briefly the nature of any unusual coenzymes or metalloproteins that are required.

4. Suppose that you could add a solution containing micromolar concentrations of Cu^{2+}, Mn^{2+}, Fe^{3+}, Co^{2+}, Zn^{2+}, and MoO_4^{2-} and millimolar amounts of Mg^{2+}, Ca^{2+}, and K^+ to a solution that contains a large excess of a mixture of many cellular proteins. What would be the characteristics of the sites that would become occupied by each of these metal ions? How tightly do you think they would be bound?

5. What factors affect the rate of electron transfer from an electron **donor** (atom or molecule) to an **acceptor**?

6. List some mechanism that cells can use to combat the toxicity of metal ions?

7. NADH peroxidase (p. 857, Eq. 15-59) does not contain a transition metal ion. Propose a reasonable detailed mechanism for its action and compare it with mechanisms of action of heme peroxidases.

Index to Volume 1

Page numbers set in **boldface** refer to major discussions.
The symbol *s* after a page number refers to a chemical structure.

A

ABC transporters 417, 419, 839
Abequose 180*s*
abl 572, 576
Abortive complexes 466. *See also* Nonproductive complexes
Absorption spectra
    of *N*-acetyltyrosine ethyl ester 123
    of adenosine 205
    of aromatic amino acid derivatives 123
    bands at 500 nm 750
    of cytidine 205
    of cytochrome *c* 847
    of proteins 122, 123
    of uridine 205
    of vitamin B_6 coenzyme 750
Acarbose 607, 607*s*
Acatalasemia 852
Acetabularia 21, 22
Acetaldehyde, hydration equilibrium 677
Acetan 179
Acetate
    binding constants to metal ions 311
    Gibbs energy of oxidation 297
Acetate kinase 661
Acetic acid
    pK_a values of 99, 293
Acetic anhydride, acylation by 124
Acetoacetate decarboxylase 705
Acetoacetyl-CoA 700*s*
Acetohydroxy acid isomeroreductase 712
Acetohydroxy acid synthase 734
Acetolactate 734*s*, 735
Acetone 705*s*
    tautomerism 45
Acetyl adenylate 661, 720
Acetylation
    coenzyme A for 720
    of histones 549
    of sulfanilamide 720
Acetylcholine
    as neurotransmitter 720
    receptor 422
Acetylcholinesterase 609, 634, 636
    pseudosubstrate of 636
Acetyl-CoA (acetyl-coenzyme A) 506*s*, 508, 881
    carboxylation of 724
    corrinoid-dependent synthesis 876, 877
    formation from fatty acids 511
    synthesis of 660
    reactions of, table 700
Acetyl-CoA carboxylase 725 – 729
    biotin subunit 726
Acetylene hydratase 894
Acetylenic fatty acids 381
Acetyl-enzyme
    in citrate lyase 704
    from serine proteases 617
N-Acetylglucosamine (GlcNAc) 164, 599
N-Acetyllactosamine 183
N-Acetylmuramic acid (MurNAc) 165*s*, 599
N-Acetylneuraminic acid (NeuAc) 165*s*
N-Acetylornithine 742
Acetyl phosphate 799
O-Acetylserine sulfhydrylase 742
S-Acetyltransferase 661
N-Acetyltyrosine amide 617*s*
Acid(s)
    enthalpy change of dissociation, table 293
    Gibbs energy of dissociation, table 293
    pK_a values of, table 293
    strengths of **95–96**
Acid – base catalysis 469, 486 – 491
    concerted 490
    of mutarotation 487
Acid – base chemistry
    of amino acids 51– 55, 304 – 307
    of enzymatic action 486 – 491, 691– 698
    of nucleic acids 204 – 211
Acid catalysis
    general 487– 491
    specific 489
Acid phosphatases 645
Acidity
    of benzoic acids 308
    of fatty acids 380
    of groups in enzymes 486
    of ketones 46
    of phenols 308
    of phenylacetic acids 308
Aconitase 676, 686 – 688, 861
    active site structure 689
    proton retention 686
cis-Aconitate 516*s*, 686*s*
    isomerization of 697
Acrasins 557
Acridine yellow 223*s*
Acrosomal process 370
Acrylochlorin heme 845, 845*s*
Acrylonitrile 115
ACTH 54*s*
    adenylate cyclase in action of 556
Actin 406
    F-actin structure 338
    thin filaments 369
Actinidin 618
α-Actinin 370, 405
Actinomycetes 6
Activation
    of enzyme(s) 471– 478
    of enzyme synthesis 536
    noncompetitive 473 – 475
Activation cascade, pancreatic proteases 609
Active transport 379, 410, 417– 420
    of Na^+ and K^+ 424
Activity of a solute 287
Activity coefficients 288
Acyl-ACP, dehydrogenation of 766
Acyl adenylate 507
Acyl carrier protein (ACP) 681, 704, 723
Acyl-CoA 526
Acyl-CoA dehydrogenase 766, 785, 789
Acyl-dihydrolipoyl derivative 736
Acyl-enzyme
    in glyceraldehyde phosphate dehydrogenase 775
    intermediates 610 – 613
Acyl groups table of names 380
Acyl phosphates 507, 509, **660 – 661**, 775*s*
Acylation 124
    of proteins 559
    in regulation 543
Acylglycerols 382
Acyltransferases 637
Adair equation 328, 354
Adaptor proteins
    Grb2 577
    Shc 577
Addition – elimination sequence 526
Addition reactions 526, 530, 677– 690
    of double bonds, adjacent to carboxylate 683 – 690
    of double bonds, carbon-carbon 681– 683
    of double bonds, isolated 688
    of double bonds, polarized **677– 690**
    of enolate to carbon dioxide 705 – 710
    facilitation by carbonyl group 681
    isomerization 690
    nucleophilic 123
    reversibility of 690

  stereochemistry 680, 681
Adducin 405
Adenine (Ade, A) 199*s*, 203
  binding constants to metal ions 311
Adenosine (Ado) 203, 778*s*
  absorption spectra of 205
  acid – base chemistry 204, 205
  hypermodified 235
  pK_a value of 293
Adenosine 5'-phosphate. *See* AMP
Adenosine diphosphate. *See* ADP
Adenosine monophosphate. *See* AMP
Adenosine triphosphate. *See* ATP
Adenosine-diphosphate-ribose. *See* ADP-ribose
S-Adenosylhomocysteine 592*s*, 778*s*
S-Adenosylhomocysteine hydrolase 778
S-Adenosylmethionine (AdoMet, SAM) 591, 592*s*, 813, 875
  decarboxylation 753, 754
  reaction with catechol *O*-methyltransferase 592
Adenosyltransferase, B_{12s} 870
Adenoviruses 247, 346
Adenylate cyclase **556–557**, 556*s*, 657
  characterization of 556
  toxic 548
Adenylate kinase 303, 654, 655*s*
  fold 658, 659
  functions of 655
Adenylate system **302 – 304**
  storage of energy 302
5'-Adenylic acid. *See* AMP
Adenylosuccinate lyase 685
Adenylylation 545
Adenylyl cyclases. *See* Adenylate cyclase
Adenylyl groups
  transfer from ATP 660
Adhesion discs (septate desmosomes) 29
Adhesion proteins 187
Adipocyte 379*s*
ADP (Adenosine diphosphate) 536
  in adenylate system 302 – 304
  complexes with metal ions 296
  dissociation as acid 288
  intracellular concentration 304
  P-31 NMR spectrum 642
  pK_a value of 293
  in regulation 535
ADP-ribose (ADPR) 315, 778, 780
ADP-ribosylation 545, **778**
ADP-ribosylation factors (ARFs) 559
Adrenaline (epinephrine) 534, 542, 553, 553*s*
  in adrenergic receptor 535
α-Adrenergic receptors 553, 558, 563
β-Adrenergic receptors 553, 554
  in asthma 553
  in heart failure 553
  receptor kinase 553
  structure (proposed) 534, 555
  topology 555
Adrenocorticotropin (ACTH) 54*s*
Adrenodoxin reductase 764, 788
Adsorption chromatography 103
Adsorption isotherm 326
Aequorin 314
Aerobes, obligate 8
Affinity chromatography 104 – 106
Affinity labeling 127, 477, 623
African sleeping sickness 787
Agammaglobulinemia 572
Agar-agar 170
Agarose 177*s*
  activated 105
  double helices in 177
  gel networks of 175
Agglutinins (lectins) 427
Aggrecan 181
Aggregation of cells 187
Aglycone 167
Agonists 479
Agrin 437
Agrobactin 838
Alamethicin 414
Alanine (Ala, A) 52*s*
  pK_a values of 293
L-Alanine (Ala, A) 52, 52*s*
  binding constants to metal ions 311
  biosynthesis 517
β-Alanine
  in coenzyme A 722
D-Alanine 429
L-Alanyl-L-alanine
  pK_a value of 293
D-Alanyl-D-alanyl carboxypeptidase 625
L-Alanyl-L-valyl-L-methionine 57*s*
Albinism 887
Alcohol(s)
  addition to carbonyl 677
  dehydrogenation of **765 – 767**
Alcohol dehydrogenase(s) 766
  isoenzymes 774
  stereospecificity of 479
  structure 773
  zinc in 680, 772 – 775
Aldaric acid 164
Aldehyde(s)
  covalent hydrates 775
  dehydrogenation of 766
  oxidation of 775
Aldehyde dehydrogenases 776
Aldehyde oxidases 794, 890
Aldehyde oxido-reductase 891*s*
Aldimine. 434s *See also* Schiff bases
  external, of pyridoxal phosphate 741
  internal, of pyridoxal phosphate 741
  from pyridoxal phosphate 742*s*
Aldo-keto reductases 774
Aldolase(s) 527, 679, **699–703**
  Schiff bases in 732
  type I 699
  type II 699, 700
Aldolase A
  enzyme-substrate complex 699
Aldol cleavage 698, 742
  in epimerization 778
Aldol condensation(s) 527, 529*s*, 699, 700
Aldonic acid 164
Aldose-ketose interconversions 692
Aldose reductase 774
Algae 20 – 22
Algin 21
Alginates 170, 178
Alkali metal ions **206**
Alkaline phosphatases 645
  active site of 645
  occurrence 645
Alkane(s) 382
Alkenes, hydration of 683
Alkyl shift, rearrangements with 527
Alkyl cobalt 867
Alkyl ethers 382
Alkyl cobalamins
  formation 870
  synthesis 869 – 870
Alkyl radicals, table of names 380
Alleles 184
Alliinase 742
Allolactose 168
Allose (All) 163*s*
Allosteric activation 539
Allosteric constant 476
Allosteric control
  of enzymes **539 – 541**
  Monod-Wyman-Changeux model 349
Allosteric effectors 357, 475 – 477, 536
Allosteric inhibition 539
Allosteric properties
  of enzymes **539 – 541**
  of repressor protein 539
Allosteric site 473
Allostery (allosterism) 475
Alloxan 121
Allylic rearrangement 527, 712
  with condensation 527
Allysine 429
Alpha-carbon plot 64
Alpha effect on reaction rates 590
Alpha helix *See* α Helix
Alport disease 438
Alternan 175
Alternative splicing of proteins 538
Altman, Sidney 84
Altrose (Alt) 163*s*
Aluminum
  hydroxide complex, soluble 310
  toxicity of **658**
Amadori rearrangement 169
Ameba 18
Amebic dystentery 187
Amebocytes 23, 25
Ameloblasts 442
Ameloblastin 442
Amelogenins 442
Amicyanin 817, 883
Amide(s)
  displacement reactions of 608
  linkages in proteins 55 – 57
  mechanism of formation 661, 662
  resonance energy 299
  weak basic properties 56
Amidination 124
Amiloride 421*s*
Amine(s)
  addition to carbonyl 677
  dehydrogenation of 766
Amine oxidases 766, 782
Amino acid(s) **51–55**
  abbreviations of, table 52, 53
  alpha (s) 41
  aromatic, light absorption 749
  C-terminal 57
  configurations **42**
  dehydrogenation of 775
  as dipolar ions 42, 51
  essential 52, 53
  N-terminal 57
  pK_a values of 55
    table 52, 53
  residues (definition) 57
  separation by ion exchange chromatography 104
  sequences of 57
  side chain groups 55
    table 52, 53

Amino acid decarboxylase 744
Amino acid dehydrogenase 766
Amino acid oxidases 782
D-Amino acid oxidase 478, 790, 791*s*
    of kidney 781
Amino acid sequence 57–59
    of aspartate aminotransferase 57, 202
    of insulin 118
Amino acid side chains
    torsion angles of 61, 62
Amino group(s)
    reactions of 123
Aminoacrylate 744*s*
Aminoacyl-tRNA synthetases 662
p-Aminobenzoate 473*s*
γ-Aminobutyric acid (Gaba) 744
Aminocyclopropane carboxylate (ACC) 744*s*
α-Aminoisobutyric acid 69
δ-Aminolevulinic acid 742, 745*s*
    biosynthesis 742, 745
    formation from glutamate-1-semialdehyde 742
Aminomutases (vitamin B_{12}) 872, 874
8-Amino-7-oxo-nonanoate 718, 742
Aminopeptidase(s) 117, 609, 627
β-Aminopropionitrile 438
Aminopterin, inhibition of dehydrofolate reductase 805
Aminotransferase(s) 737, 740–743, 747–753
    half reactions 737
Ammonia
    binding constants to metal ions 311
    pK_a value of 99, 293
Amoeba proteus 19
AMP (Adenosine 5'-phosphate) 203
    in adenylate system 302–304
    dissociation as acid 288
    pK_a value of 293
    as protein handle 719
Amphipathic character 380, 401
Amplification
    of DNA 260
    of signals 567
α-Amylase(s) 603, 605–608
    action patterns of 606
    calcium ion in 606
    chloride ion in 606
    pancreatic, 606s
    pH profile of 471
    subsites 606
β-Amylase(s) 603, 608
Amylopectin 165, 170, 172, 173*s*
Amylose 165, 170, 172
    double-helical form 173
Anaerobes
    facultative 8
    obligate 8
Anaerobic nucleotide reductases 864
Anaphylaxis 385
Anchors, membrane for proteins 402
    diphytanylglyceryl 402
    glycosylphosphatidylinositol (phosphatidylinositol-glycan) 403, 403*s*, 523
    myristoyl 402, 522
    palmitoyl 522
    polyprenyl 402
Angiosperms 29
Angiotensin II 563
Angiotensin-converting enzyme 625
Anhydrobiotic organisms 168
1,5-Anhydro-D-glucitol 165*s*
Animal cells, surroundings of 431–443
Animal tissue, composition of 31
Ankyrin 405
Annelida (segmented worms) 24
Annexins 316, 422, 571
Anomeric carbon atom 162
Anomeric effect 166
Antagonists 479
Antamanide 414
    Na$^+$ complex 415*s*
Anthranilate 747*s*
Anthrax 548
Anti conformation of nucleosides 211, 212
Antibiotics 414, 479
    anthracycline 223
Anticoagulation systems in blood 634
Anticodon 231
Anticodon loop in tRNA
    hypermodified base in 231
Anticooperativity 352
    in binding 329, 497
        to metal ions 307
        of protons 330
Antidotes for organic phosphates 636
Antienzymes 539
Antifreeze proteins 191
Antigen(s) 427
O-Antigen(s) 428
    of *Salmonella* 180
Antigenic determinants 427
Antihemophilic factor 632
Antimetabolites 473
Antioxidants 818–820, 822–824, 826
    vitamin E 721, 818, 819, 822, 823
Antiparallel β structures 62
α_2-Antiplasmin 634
Antiport process 410
Antithrombin 634
Antitrypsin 58
AP-1 (Activator Protein) 576, 577
Apoferritin 842
Apoptosis 574, 619
Apple motif of proteins 367
Apurinic acid 250
Apyrimidinic acid 255
Aquaporins 411, 412
Arabidopsis thaliana
    genome 12, 29
Arabinan 177
Arabinogalactan 431
L-Arabinose (Ara) 163
    binding protein 418
D-Arabinose (Ara) 163*s*
L-Arabinose isomerase 693
Arachidonic acid 381*s*, 565
Aragonite 440
Archaea 8
Archaeobacteria 7, 815
Area detector for X-rays 134
Arg-Gly-Asp (RGD) sequence 408
Arginase 888
    manganese in 888
Arginine (Arg, R) 53*s*, 745
Arginine kinase 656
Arginine phosphate 303
Argininosuccinate lyase 685
Aromatic clusters in proteins 75
Aromatic groups, formation by elimination 689
Arrhenius activation energy 483
Arrhenius equation 483
Arrhenius, Svante A. 9, 83
Arrows, drawing 528, 529
Arsenate
    phosphorylation, uncoupling of 596
    similarity to phosphate 596
Arsenic compounds 387, **596–597**
    toxicity of 596
Arsenicals in treatment of syphilis, 597
Arsenite, reaction with thiols 596
Arseno-ribofuranoside 596*s*
Arsenolysis 596
    of glucose 1-phosphate 595
Arsenophospholipid 387*s*
Arthropods 24
Ascaris 25
    chromosome number of 18
Ascarylose 180*s*
Ascomycetes 20
Ascorbate oxidase 887
Ascorbate peroxidase 852
Ascorbic acid 721
    nutritional requirement 756
Ascus 20
Ash, amount from tissue 31
Asparagine (Asn, N) 53s
Aspartase 526
Aspartate 737s
    binding constants to metal ions 311
    biosynthesis 711
    effect on biotin requirement 724
    metabolism of 539
    residues 312
Aspartate aminotransferase 57*s*, 135*s*, 753
    absorption spectra 749
    active site structure 744
    atomic structure 750
    catalytic intermediates, models 752
    NMR spectra 149
    quinonoid intermediate 750
    Ramachandran plot 61
    sequence 57
    transamination 742
Aspartate ammonia-lyase 685
Aspartate carbamoyltransferase 348*s*
    active sites 348
    regulation 540
Aspartate α-decarboxylase 753, 755
Aspartate β-decarboxylase 746
Aspartate racemase 741
Aspartic acid (Asp, D) 52, 53*s*
    biosynthesis 517
    pK_a value of 293, 487
Aspartic proteases **621–625**
Aspartyl aminopeptidase 621
β-Aspartyl phosphate 539, 540*s*
Assays of enzyme activity 456
Assembly core of virus shell 365
Assembly pathway
    Bacteriophage T4 367
    for virus φX174 366
Association constant. *See* Formation constant
Astacin 627
Asthma 26, 385
Asymmetric unit of crystal 134
Asymmetry of molecules 41–43
    handedness 41–43
    in oligomers 344
Ataxia telangiectasia 566, 574
Atomic force microscope 131
    diagram of 131
Atomic force microscopy 131, 179, 392
Atomic structures of macromolecules. *See* Three-dimensional structures

  - pyridoxal phosphate enzymes 750
- Atoms, table of sizes **41**
- ATP (Adenosine triphosphate) 536, **719–720**
  - in acetylation of sulfanilamide 720
  - in adenylate system 302–304
  - analogs of 655
  - binding constants to metal ions 311
  - as coenzymes 719
  - complexes with metal ions 296
  - dissociation as acid 288
  - exchange-inert complex 643
  - formation during oxidation of aldehyde, scheme 775
  - hydrolysis of 507
  - enthalpy of 296
  - Gibbs energy of 293, 297
  - pK_a values of 293
  - proteases, dependent 627, 628
  - in regulation 535
  - screw-sense isomer 643
  - synthesis of 303, **507–515**
- ATPase 638
  - Ca^{2+} 424, 425*s*
  - gastric H^+, K^+- 424
  - H^+, K^+- 424
  - Na^+, K^+- [(Na^+ + K^+)-activated] 423
  - P-type 423
  - traffic 417
  - vacuolar 425
- ATP, biosynthesis of
  - in fermentation reactions 661
  - uncoupling 636
- ATP-citrate lyase 703
- ATP synthase 425, 514, 659
- Atrial natriuretic factor 561
- Aurocyanin 883
- Autocatalytic reaction 470
- Autocrine effect 385, 574
- Autophosphorylation of receptor 562
- Auxilin 427
- Avidin 257, 728
- Avogadro's number 283
- Axis (axes) of symmetry 322, 323, 327–338
- Axons 11
- Azurins 883

B

- Bacillus, definition of 6
- *Bacillus subtilis*
  - genome 12
- Bacteria 2. *See also* Specific genus and species
  - acetic acid 8
  - aerobes 10
  - anaerobic 8
  - autotrophic 8
  - binding to cells 186
  - branched fatty acids of 381
  - chemoheterotrophic 7, 8
  - chemolithotrophic 7
  - classification of **6–8**
  - coats 431
  - composition of 31
  - electron micrograph of 4
  - flagella 6
  - gram-negative 6, 7, 180
    - transporters in 416–420
  - gram-positive 6, 7, 185, 190, 428
  - L forms 6
  - largest 8
  - methanogens 7
  - coenzymes of **813–815**
  - nutrition of 8
  - photoautotrophic 8
  - photosynthetic 7, 9
  - pili 335
- Bacterial dehydratases
  - Fe-S clusters in 861
- Bacterial diseases 7
- Bacteriophages
  - filamentous (fd, f1, M13) 129, 334, 335
  - helical 244
  - temperate 248
- Bacteriophage ϕX174 244
- Bacteriophage P22
  - tailspike protein 66, 66*s*
- Bacteriophages T-even 244, 363
- Bacteriophage T4(s) 363*s*
  - assembly pathway 367
  - lysozyme 599
- Bacteriorhodopsin 71
- Baculoviruses 247
- Baltimore, David 84
- Band 3 protein 404
  - of erythrocytes 420
- Barrels
  - α/β 77
  - α,α 71
  - β 65
- Basal body of flagella 15
- Basal metabolic rate 283
  - relationship to surface area 283
- Basal stem cells 439
- Base(s)
  - of nucleic acids 199
  - strengths of 95, 96
- Base catalysis, general 487–491
- Base composition of DNA and RNA 239
- Base pairs **207–211**
  - angle of twist 217
  - base inclination 217
  - of DNA 200
  - Hoogsteen 207, 231
  - propeller twist 217
  - strength of 209
  - tautomerism 211
  - thermodynamics of 209
  - Watson-Crick 207, 231
  - wobble 209, 231
- Base quartets 207–211
- Base triplets 207–211
- Basement membrane 15, 72, 431, 436
- Basidiomycetes 20
- Basophils 26, 188
- β Bend **72–74**
- *bcl* 572
- *Bdellovibrio* 3
- Benzaldehyde, resonance energy 299
- Benzene, resonance energy 299
- Benzoic acid
  - acidity of 308
  - pK_a value of 309
- Benzoylformate decarboxylase 734
- Berenil 225*s*
- Beriberi 721
- Beryllium, toxicity of 299
- BES buffer 99
  - pK_a value of 99
- Beta cleavage. *See* Cleavage, beta
- Beta condensation. *See* Condensation, beta
- Beta helix *See* β Helix
- Beta oxidation 511, 512
  - of fatty acids 511, 512
  - Knoop's theory **513**
  - pathway **506**
- Beta propeller 67, 560, 764
- Beta ribbons 241
- Beta sheet 64
- Beta spiral 74
- Bicarbonate ion
  - from carbon dioxide 677
  - in carboxylation 710
  - into carboxyl groups 711
  - rates of incorporation 710
- Bicarbonate sensor 557
- Bicinchoninic acid 102
- BICINE buffer 99
  - pK_a value of 99
- Biliverdin 852
- Binding. *See also* Association
  - by macromolecules 327–330
  - cooperative processes 330–332
  - equilibria **325–332**
  - of metal ions **307**
  - multiple sites of 327–330
  - ordered 464
  - of oxygen. *See* oxygen (O_2), binding of
  - random order 464, 467
- Binding constant(s) 325–327. *See also* Formation constant
  - anticooperativity 329, 330
  - intrinsic 328
  - for metal complexes, table 311
  - microscopic 328, 329
  - of protons to dianions, table **328**
- Binding curves 326
  - sigmoidal 331
- Binding energy
  - relation to maximum velocity 617
- Binding equilibria 325–332
  - dimerizing protein 350–352
- Binding protein(s) 418
  - for histidine 419*s*
  - for hormones 536
- Binding sites, multiple 327
- Bioactivation of insecticides 636
- Biochemical equilibria **304–316**
- Biochemical reactions
  - table of types **526, 527**
- Biocytin 724, 724*s*
- Biological recognition 29
  - glycoproteins, in 186
- Biomineral precursor 843
- Biopolymers, hydrolysis of 292
- Biopterin 803, 804*s*
  - as hydroxylation coenzyme 803
- Biosensors 122
- Biosynthesis 507, **515–517**
  - of isoprenoid compounds 689
  - scheme 506, 509
- Biosynthetic pathways
  - committed step 535
  - feedback inhibition 535
  - schemes 506, 509
- Biotin 515, 516, 711, 721, **723–730**, 723*s*
  - biosynthesis of 718, 745
  - in enzymes, table 724
  - mechanism of action 725–729
  - nutritional requirement 756
- Biotin-binding proteins **728**
- Biotin carboxylase 724
- Biotin carboxyl carrier protein 724
- Biotin holoenzyme synthetase 724

- Biotinidase 725
- BIS-TRIS buffer 99
  - pK_a value of 99
- 1,3-Bisphosphoglycerate 596*s*
  - in glyceraldehyde-3-phosphate dehydrogenase 775
- 2,3-Bisphosphoglycerate 357*s*, 361
  - effect on hemoglobin 357
- Bisphosphoglycerate mutase 654
- Bisulfite 779
  - reaction with uridine 251
    - with thiamin 731
- Biuret 102*s*
- Biuret method 102
- BK virus 247
- Black membrane 393
- Bleomycin 225*s*
- Bleomycin hydrolase 619
- Blood
  - carbonic anhydrase 678
  - cell(s) 26
  - complement system 58, 414
  - proteins of **58**
- Blood-brain barrier 29
- Blood clotting **631–634**
  - anticoagulation systems 634
  - cascade mechanism 632
  - cofactors 631
  - inhibition of 622
  - intrinsic mechanism 633
- Blood group determinants 182, 183
  - ABO system 184, 185
  - Cad antigen 185
  - Kell antigens 185
  - Lewis blood group 184, 185
  - MN antigens 185
  - P group 185
  - Rh antigen 185
- Blood plasma
  - electrolyte concentrations 206
  - proteins **58**
- Blood platelets 26, 384
- Blood type. *See* Blood group determinants
- Blue-green algae. *See* Cyanobacteria
- Blue multicopper oxidase 887
- Bohr effect 357
- Boltzmann constant 284
  - numerical value of 283
- Bond(s)
  - angles **39–40**
  - energies, table 299
  - lengths **40**
- Bonding
  - heterologous 333, 334
  - isologous 338
  - nonsymmetric 344
- Bone 26, 440, 441
- Bone morphogenic factor 443
- *Bordetella* 548
- Boric acid, pK_a value of 99
- Borohydride
  - in carbohydrate analysis 190
  - reduction by 124, 190
- Boronic acids 622
- Bradford method 102
- Brain
  - of molluscs, organization 24
  - transporters in 416, 417, 420
- Branch migration in DNA 229
- Breast cancers 572
- Brine shrimp 168
- Bromelain 618
- Bromoperoxidase(s) 856
- Brönsted equation in basicity 490
- Brönsted plot 309, 490
- Brown, Michael S. 84
- Brownian movement 461
- Brucellosis 7
- Bryophyta 29
- Buchner, Edward 83, 456
- Buffer(s) **97–98**
  - table 99
  - volatile 98
- β Bulge 66, 67*s*
- Burkitt's lymphoma 572, 576
- Butyrylcholinesterase 634

C

- Cacodylic acid, pK_a value of 99
- Cadherin(s) 187, 407
  - domains 408*s*
- Cadmium 317
  - toxicity of 680
- cADP ribose (cyclic ADP ribose) 564*s*
- *Caenorhabditis elegans*
  - apoptosis 619
  - cuticle 440
  - genome 12
  - number of cells 24
- Cage effect 462, 463
- Caged phosphate 137
- Calbindins 313
- Calcineurin 646
- Calciosomes 314, 563
- Calcite 440
- Calcitonin 314
- Calcium **314–315**
  - oscillations 315
  - release channel 314
  - in signaling **563–566**
- Calcium-binding protein(s) 312
- Calcium-binding sites 312
- Calcium carbonate
  - spicules 443
- Calcium channels 422
- Calcium ion(s) 31, 314, 315
  - activation of metabolic processes 314
  - concentration
    - in tissues 314
    - in serum 314
  - effects on equilibria 297
  - interaction with carboxylate groups 47
- Callose 170, 174
- Calmodulin 313, 315, 317*s*
  - in phosphorylase kinase 544
- Calnexin 188, 520
- Caloric values of food 282, 283
  - table 283
- Calorie, definition of 282
- Calorimeter
  - bomb 282
  - differential scanning 285
- Calorimetry
  - animal 283
  - differential scanning of membranes 394
- Calpactin *See* Annexin
- Calpains 619
- Calreticulins 313
- Calsequestrin 313, 314
- Calvin, Melvin 84
- Cambium layer 29
- Cambrian explosion 9
- cAMP 536, 544, 544*s*, 556, 557
  - second messenger 557
- Cancer **573–575**
  - chemotherapy 224, 812
  - p53 as suppressor 581
- Cancer cells
  - communicating junctions in 29
  - lack of contact inhibition 29
  - promoters 573
- Capsid 244
- Capsomers 244
- Capsule 6
- Captopril 624
- Carbamate group (*See also* carbamino group) 80
  - in ribulose bisphosphate carboxylase 707–709
- Carbamates
  - in proteins 80
- Carbamino groups 678 *See also* Carbamate group
  - in hemoglobin 358
  - in regulation 543
- Carbamoylation
  - by sodium cyanide
- Carbamylation *See also* Carbamoylation of lysine 707
  - of ribulose bisphosphate carboxylase 707
  - by sodium cyanate 124
- Carbanion mechanism of flavin dehydrogenase 790
- Carbinolamine(s) 677
  - configurations 680
  - formation of 526
- Carbocation(s)
  - from alkenes 683
  - oxocarbenium ion 598*s*
- Carbocation mechanism
  - of fumarate hydratase 684
- Carbocationic intermediate 595
- Carbodiimide 105*s*
- Carbohydrate(s)
  - analysis 188
  - caloric value 283
  - content, in tissues 31
  - hydrolysis of 189
  - methylation of 189
  - NMR spectra of 190–192
  - periodate oxidation of 190
- Carbohydrate-binding proteins. *See* lectins
- Carbon-14, properties of 110
- Carbon-carbon bonds
  - addition to 681–683
  - cleavage of 530
  - reduction of 777
- Carbon dioxide 358
  - carbon units, source of 810
  - hydration of 677, 710
  - reduction of 881
  - reduction, scheme 814
  - resonance energy 299
  - stability of 690
- Carbon monoxide 553
  - in hemoglobin 358
  - oxidation of 881
  - resonance energy 299
- Carbon monoxide dehydrogenase/acetyl-CoA synthase 881, 882

- Carbon monoxide oxidase 893
- Carbonic acid, pK_a value of 99
- Carbonic anhydrase 443, **676 – 678**, 710
  - active site structure 679
  - mechanism 678
  - turnover number of 458, 678
- Carbonium ion. *See* Carbocation
- 1,1'-Carbonyl-diimidazole 105*s*
- Carbonyl group
  - addition of enolate anion 699 – 704
  - addition to 677
  - in displacement reactions 608 – 637, 698
  - facilitation of addition or elimination 681
  - formation 723 – 730
  - in metabolic sequences 681
  - reactivity 608
  - transient formation of 777, 778
- *N*-1'-Carboxybiocytin
  - in action of enzyme 725
- *N*-1'-Carboxybiotin 725, 726, 726*s*
  - dimethyl ester of 725
  - intermediate 725
- γ-Carboxyglutamate 821, 821*s*
  - binding of calcium ions 80
  - formation of 820 – 822
- Carboxylation
  - beta 527, 727 – 729
  - biotin-assisted 724
  - driven by cleavage of ATP 705
- Carboxylesterases 635
- Carboxyl (carboxylate) groups
  - in proteins 312
  - reduction of 776
  - resonance energy 299
- Carboxylic acid
  - activation of 507
  - dehydrogenation of 766
- Carboxylic acid reductase 893
- Carboxyltransferase 724
  - of propionic acid bacteria 725
- Carboxymethylation
  - of proteins 594
- Carboxypeptidase(s) 117, 609, 625
  - zinc ion in 625
- Carboxypeptidase A 64*s*
  - active site 626*s*
  - pK_a values of 626
- Carboxypeptidase Y 610
- Carboxyphosphate 726, 726*s*
- Carcinogenic compounds
  - adducts with nucleic acids 235
- Carcinomas 573
- Cardiolipin 384
- Caricain 618
- Caries 442
- β-Carotene 10, 390*s*
- Carpal tunnel syndrome 756
- Carrageenan(s) 170, 177, 178*s*
- Carrageenin 22
- Cartilage 26, 436
  - articular 182
  - matrix protein 437
- Cascades
  - blood clotting 632
  - dephosphorylation 542
  - MAP kinase 576 – 578
  - pancreatic proteases 609
  - phosphorylation 542
  - regulatory 566, 567
- Casein kinase 544
- Caspases 619
- Catabolism 536
  - and ATP synthesis **507 – 515**
  - pathways of 506, **509**
- Catalase(s) 844, 852 – 855, 853*s*
  - catalytic cycles, scheme 855
  - compounds I, II 853 – 855
  - function of 852
  - hereditary deficiency 852
  - iron properties 853
  - mechanisms of catalysis 853 – 855
  - molecular activity of 852
  - turnover number of 458
- Catalysis (acid, base) 469, 486
- Catalysis (enzymatic)
  - acid and base 469, 486 – 491
  - covalent 494
  - mechanisms of **482 – 497**
  - microenvironment 495
  - microscopic reversibility 486
  - multiple attack 606
  - nucleophilic 494
  - proximity factor 495
  - stereoelectronic effect 495
  - strain and distortion in 496, 497
  - tautomeric 490, 491
  - transition state 482 – 486
- Catalyst, polyfunctional 490
- Catalytic cycles. *See also* Other specific enzymes and cell cycle
  - for catalase 855
  - for coenzyme M reductase 880
  - for glycine decarboxylase 809
  - for lysozyme 599
  - for α-oxoacid dehydrogenases 798
  - for serine protease 613
  - for vitamin B_{12}-dependent isomerases 872
- Catalytic dyad 620
- Catalytic subunits 348
- Catalytic triad 134, 611 – 614, 619, 635
  - of serine proteases 611 – 614
- Cataracts 169
- Catechol 838
- Catecholamine(s) 553
- Catechol *O*-methyltransferase (COMT) 591
- Cathepsin(s) 619, 621
  - G 610
  - K, in osteoclasts 619
- Cations. *See also* Individual ions
  - active transport 422
  - channels 421, 422
- Cat scratch disease 7
- Cauliflower mosaic virus 247
- CD (Cluster of Differentiation) 187, 407
- Cech, Thomas R. 84
- CED-3 619
- Cell(s)
  - aggregation of 187
  - columnar epithelium 26
  - composition of **30 – 33**
  - doubling time 470
  - fractionation of 98 – 100
  - junctions 28
  - metal binding sites in 311, 312
  - sizes of 11
  - squamus epithelium 26
- Cell adhesion molecules (CAMs) 406
- Cell constant organisms 24
- Cell contacts 28, 29
- Cell culture 26
- Cell cycle **578 – 581**
  - figure 580
- Cell differentiation antigens (cluster of differentiation, CD) 187, 407
- Cell membranes 3
  - anions 420, 421
- Cellobiohydrolases 602
- Cell recognition 29
- Cell respiration, inhibition 636
- Cell signaling, table of abbreviations 577
- Cell structure, compartments 552
- Cell types
  - animal 25
  - plant 29, 30
- Cell wall 5, 10, 15, **427 – 441**
  - of bacteria 427 – 431
  - of plants 5, 177, 179, 443
    - in micrograph 13
- Cellulases **602**, 603
  - structural features 602
- Cellulose 170, 171*s*, 172
  - content in spinach 31
- Cellulose I 171*s*, 172
- Cellulose II 172
- Cellulose-binding domains 602
- Cellusome 602
- Cementum 442
- Centipedes 24
- Centrifugation 100
  - in CsCl gradients 249
  - density gradients 101
  - zone 101
- Centrioles 10, 15
- Centromeres, DNA of 228
- Centrosome 627
- Cephalopods, mesozoa in kidneys of 23
- Ceramide 389*s*
- Ceramide aminethyl phosphonates 389*s*
- Cerebral ganglion in Platyhelminthes 23
- Cerebrosides 388, 389*s*
- Ceruloplasmin 58, 633, 839, 883, 887
- CFTR 417
- cGMP 561
- Chagas disease 787
- Chain elongation process 700
- Chain-terminating inhibitors for DNA sequencing 262
- Channeled substrates 552
- Chaperone (chaperonin) 339
  - calnexin 188
  - DnaK 518
  - GroEL 518
  - GroES 518
  - heat shock proteins 518
- Chargaff, Erwin 200
- Charge, effect on metal binding 311
- Charge-relay system 612
- Chelate complexes 310*s*
- Chelate effect 310
- Chelate stability 311
- Chemical force microscopy 131
- Chemical shift 138
  - hydrogen bonding, effect of 139
  - prediction 139
  - shielding of nuclei 138
  - table of NMR spectra 141
  - units of 138
- Chemoheterotrophic bacteria 8
- Chemolithotrophic bacteria 7
- Chemoprevention 575
- Chemoreceptors 562*s*
  - bacterial 561 – 563
- Chemostat 470
- Chemotaxis

  carboxylmethylation in 548, 563
  CheA 562, 563
  CheB 563
  CheR 563
  CheW 562, 563
  CheY 562, 563
  CheZ 563
Chiral acetates 704, 705
  nonenzymatic preparation 704
Chiral methyl group 591
Chirality 42
Chitin 19, 170, 175
  in exoskeleton of arthropods 24
Chitinases 599 – 602, 601, 603
Chitosan 175
Chlamydomonas 20, 22
Chloramphenicol 637*s*
Chloramphenicol acetyltransferase (CAT) 637
Chlorella 17, 22
Chloride channels 420, 421
  band 3 protein 420
Chloride ion, activation of α-amylases 606
Chlorin(s) 843
β-Chloroalanine 739
Chlorobenzoyl-CoA dehalogenase 682, 683
Chlorocruorins 362
Chlorocruoroheme 845*s*
Chloroperoxidase(s) 855
Chlorophyll(s) 13
  chelate effect 310
Chlorophyll a 10, 843
Chlorophyta 20
Chloroplast(s) 10, 14
  in micrograph 13
  ribulose bisphosphate carboxylase in 706
Cholera 7
Cholera toxin 546 – 548, 546*s*
  in ADP-ribosylation 546
  AMF images 131
Choleragenoid 333*s*
Cholesterol 390*s*, 392
  elevated 16
  as precursor of steroid hormones 696
  from reduction of desmosterol 777
Chondroitin sulfate(s) 170, 175, 176*s*
Chordata 25
Chorismate 485*s*, 690*s*
  biosynthesis of 689, 690
Chorismate mutase 485
Christian, W. 767
Christmas factor 632
Chromatin 11
Chromatography **102 – 105**. *See also* specific types
Chromium **888 – 889**
  complex with ATP 643
Chromosome(s) 11, 17
  bacterial 3
  homologous pair 18
  micrograph of 371
  numbers of 18
  sizes of 18
Chymopapain 618
Chymosin 621
Chymotrypsin 66, 609
  inhibition by TPCK 477
  specificity 117
  turnover number of 457
Chymotrypsinogen 609, 613, 615
  activation of 609, 615
  hydropathy index 78
Cibachron blue in affinity chromatography 106
Cilia 10, 15. *See also* flagella
  eukaryotic structure 15
Ciliophora 19
Cinnamic acid 756*s*
Cinnamoyl-CoA thiol esters 682
Circular dichroism (CD)
  of bound PLP and PMP 750
Cisplatin 218, 224*s*
Citrate 480*s*, 506*s*, 515, 516*s*, 686*s*
  binding constants to metal ions 311
  prochiral centers 480
Citrate cleaving enzymes **703 – 704**
Citrate synthase(s) 527, 676, 700–703, 702*s*
  active site 703
  mechanism 700
Citric acid
  pK_a values of 99
  synthesis of 700
Citric acid cycle 506, **508**, **515**
  discovery of 516
  priming of 711
  scheme 516
Citrulline 80
Citryl-CoA 703
Claisen condensation 699
Claisen rearrangement 485
Clathrin 426
Clathrin light chains 313
Cleavage
  aldol 698
  alpha 730
  of amino acid side chains 745
  beta **698 – 711**
Clone of cells 26
Cloning vehicle 263
Clostridium acetobutylicum 705
Clostripain, specificity 117
Clp (caseinolytic protease) 628
Cluster of differentiation, CD 187
CMP (Cytidine 5'-phosphate) 203
Cnidaria (Coelenterata) 23
Coated membrane 426*s*
Coated pits 426
Coated vesicles 426
Coatomer 426
Cobalamin 867. *See also* Cyanocobalamin, Vitamin B_{12}
  in ribonucleotide reductase 871
Cobalamin-binding proteins 58
Cobalophilin 869
Cobalt **866 – 877**
  complex with ATP 643
Cocarboxylase 721
Coccidia 19
Coccus, definition of 6
Cocoonase 610
CO dehydrogenase/acetyl-CoA synthase 881
Coding strand (non-transcribing) of DNA 204, 237
Codon(s) 4, 236
  initiation 236
  stop 236
  table 236
Coenzyme(s) 511, **719 – 757**, **765 – 829**
  ease of dissociation from proteins 719
  oxidative 719
Coenzyme A (CoA) 507, **720 – 723**, 722*s*
  functions of 722
Coenzyme A persulfide 790
Coenzyme A transferases **662**
Coenzyme F_{420} 788
Coenzyme M 813*s*, 814, 815
Coenzyme M reductase
  mechanism of 881
    scheme 880
Coenzyme Q 818. *See* Ubiquinones
Cofactor F_{430} 879 – 881, 880*s*
Coiled coil
  helical wheel 70
  heptad repeat 70
  superhelix 70
Colchicine 371, 371*s*
ColE1 248
Colicins 414, **418**
Colipase 635
Colitose 180*s*
Collagen 15, **72**, 181, 409, **432 – 441**
  alternative splicing 436
  amino acid composition 72
  biosynthesis of 433
  characteristics, table 434
  crosslinking 433, 434
  gene location, table 434
  genetic defects **438**
  hydrogen bonds in 72
  location, table 434
  sequence 72
  structure 72, 432, 433
  vertebrate types, table 434
Collagen fibril, electron micrograph 73
Collagen, type IV
  electron micrograph of 437
  structural model of 437
Collagenases 627
Collagenlike peptide 73*s*
Collenchyma 30
Color reagents 120
Combinatorial libraries 128, 129
Combustion, Gibbs energies of 297, 298
Communication between cells 26 – 29
Competitive inhibitors 471, 472
  slow, tight-binding 472
Complement system
  of blood 72, 414
  of thioester-containing proteins 630
Complementarity
  of amino and carbonyl 679
  of enzyme surfaces 478
  of substrate 478
Compound I, II
  of peroxidases, catalases 853, 855
COMT. *See* Catechol *O*-methyltransferase
Concanavalin A 186, 342
Concentration gradient(s) 289, 410
Concerted mechanism 490
Condensation
  aldol 699, 700
  alpha 730
  beta 527, **698 – 711**
Configuration
  of amino acids **41– 43**
  entgegen (*E*) 43
  rectus (*R*) 42
  RS notation **42**
  sinister (*S*) 42
  of sugars 163
  zusammen (*Z*) 43
Confocal micrograph
  of a cerebellum 130
  of metaphase chromosomes 361

Confocal scanning microscopy 129
Conformations of molecules **43–45**
  anti 43
  anti-periplanar 43
  boat (B) 166
  eclipsed 44
  envelope 166
  Fisher projection 42
  gauche 44
  Newman projection 44
  of nucleotides 211
  of polysaccharide chains 170
  quasi-equivalence 348
  random coil 69
  sickle 44
  skew (S) 166
  twist (T) 166
  5- and 6-membered rings 166
  C1 of sugars 166
  chair 166
Conformational changes
  cooperative **349–353**
  during dehydrogenase action 773
  in enzymes 481
  in proteins 81
  role of buried hydrogen bonds 81
  slow 477
Conformational maps 60
Conformational substates in proteins 496
Conidia 20
Conjugate bases 486
Connective tissue 72, 431
Connexons 29
Consensus sequences in DNA 265
Constitutive enzyme 538
Contact inhibition 573
  of cells 29
Control elements of metabolic reactions 536
Controllability coefficient 537
Coomassie brilliant blue 102, 122*s*
Cooperative binding 352
  of protons 331
  of substrate 476
Cooperative changes in conformation **349–353**
Cooperative processes 330–332, 334
Copper 65, 514, **882–887**
  as nutritional essential 882, 883
  uptake regulation 883
Copper amine oxidases **816–817**
  mechanisms 817
Copper oxidases 885–887
Copper proteins, electron-transferring 883, 884
Copper–sulfide cluster 884*s*
Coproporphyrin III 843, 845*s*
Cornified cell envelope 439
Corrin in transmethylation 592
Corrin ring 867, 868
Corrinoid-dependent synthesis
  of acetyl-CoA 876, 877
Cosmarium 22
COSY-NOESY diagram 143
Cotransport (symport) process 411, 416, 417
Coulomb 283
  definition of 47
Coulomb's law 47
Countercurrent distribution 102
Coupled enzyme assay 768
Coupling constants. *See also* NMR spectroscopy
Covalent catalysis 494, 591
Covalent modification
  in control of metabolism **541–545**
  by dephosphorylation 541–545
  by phosphorylation 541–545
  reactions, table 543
Covalent radii 40
Covalent structure of proteins 57
Cow, genome 12
Cozymase 767
Crambin, amide exchange rate 149
Creatine kinase 656
  transition state structure 640
Creatine phosphate 303
Crepenynic acid 381*s*
Cretinism 31
Creutzfeldt-Jakob disease 248
Crick, Francis H. C. 84, 200
Cristae of mitochondria 14
Crossing-over 18
Crosslinking 79
Crotonase. *See* Enoyl hydratase
Crowfoot Hodgkin, Dorothy M. 84
Cruciform structure in nucleic acids 229
Crustacea 24
Cruzain 619
Cryoenzymology 469
  elastase 616
Cryoprotectants 191
Crystallins 169
Crystallography 131–137
  electron 131
  X-ray 132–137
Crystals, liquid 392–394
Crystal systems 133
Cubic symmetry
  of oligomers 342
  of pyruvate dehydrogenase 796
Cupric ion, complexes with ammonia 307
Curdlan 170, 174
Cuticle
  of invertebrates 440
  of plants 15, 30
  in micrograph 13
Cutinase 134*s*, 635
Cyanase 717
Cyanate, carbamoylation by 360
Cyanide ion
  adduct with NAD$^+$ 779
  complex with hemoglobin 358
  reactivity of 590
β-Cyanoalanine 438
Cyanobacteria 7, 10
Cyanoborohydride 124
Cyanocobalamin 867*s*, 869. *See also* Vitamin B$_{12}$
  reduction 869, 870
Cyanogen bromide 117
  agarose, activation of 105
Cyclic adenosine 3',5'-monophosphate
  as attractant for *Dictyostelium* 20
Cyclic ADP ribose 778
Cyclic AMP. *See* cAMP
Cyclic GMP. *See* cGMP
Cyclic symmetry 333
Cyclins, mitotic 580
  cell cycle 524
  D1 572
  G1 or start 580
Cyclization
  PLP-dependent enzymes 741
Cyclodextrins 171
Cyclodextrin glucanotransferase 603, 607
Cycloglutamate inhibitors 739*s*
Cycloguanil 805
1,2-Cyclohexanedione, reaction with guanidinium groups, 126
Cyclophilin 488
  human 488*s*
D-Cycloserine 739*s*
Cyclosporin 488, 488*s*
β Cylinders 65, 66, 686
Cystathionine, 746*s*
  formation 746
Cystathionine β lyase 742
Cystathionine β-synthase 744
Cystathionine γ-synthase 743, 746
Cystatins 622, 629
Cysteamine disulfide (cystamine) 549
Cysteine (Cys, C) 52, 53*s*
  microscopic pK_a values 306
  zinc complex 680
Cysteine residues 53
  crosslinking of 80, 521, 522, 786, 787
Cysteine proteases **618–619**
Cystic fibrosis 16, 421
  transmembrane conductance regulatory 421
Cystine residues
  beta elimination 85
Cystinuria 16, 417
Cytidine (Cyd) 203
  absorption spectra of 205
  reaction with hydrazine 252
Cytidine 5'-phosphate. *See* CMP
Cytidine diphosphate choline 720
5'-Cytidylic acid. *See* CMP
Cytoadhesins 409
Cytochalasin B 416*s*
Cytochrome(s) **845–848**
Cytochromes *b* 847
Cytochrome b_{562} 847
Cytochrome *c* 846, 847, 847*s*
  absorption spectra 846
  structural studies 846
Cytochrome *c* oxidase 848, 887
Cytochrome *c* peroxidase 849*s*, 852
  active site of 853
Cytochrome c_3 847*s*
Cytochrome oxidase 513
Cytochrome P-450 846
Cytokines 188
Cytokinin 235
Cytomegalovirus
  micrograph 246
  protease of 614
Cytoplasm, definition of 3
Cytosine (Cyt, C) 199*s*, 203
  acid–base chemistry 204, 205
Cytoskeleton 15, **368–372**, 405
  intermediate filaments 369
  microfilaments 369, 370
  microtubules 370
Cytosol, definition of 14

D

Daltons 4
Dansyl chloride 118, 118*s*
Darwin, Charles 9
Daunomycin 222, 223*s*
  intercalation with DNA, figure 223

DDT dehydrochlorinase 551
Dead-end complexes 466
DEAE-cellulose 104
DEAE-Sephadex 106
Deamidation
  of asparagine 594
  of asparaginyl residues 84
5-Deazaflavin
  in hydrogenase 879
  nonenzymatic oxidation of NADH 789
Deazariboflavin 5'-phosphate 789
Debranching enzyme 606
Debye-Hückel equation 288
Decarboxylase(s) 753 – 755
  biotin-dependent 729
  dialkylglycine 745
  pyridoxal phosphate-dependent enzymes 744, 745
  pyruvoyl enzymes
  Schiff bases in 732
  thiamin-dependent 730 – 733
Decarboxylation 527, **705 – 710**
  catalysis by amines 705
  catalysis by metal ions 705
  of β-oxoacids 698, 705, 742
Decarboxylation-transamination reaction 748
Decarboxylative elimination 526
Degree of hydration, effect on metal binding 311
Dehydroalanine 755, 757*s*
Dehydroalanine residue 754, 756
Dehydroascorbate 787
  transport 416
Dehydroascorbic acid 416, 787
Dehydrogenase(s) 514, **765 – 802**
  conformational changes 771, 772
  3-dimensional structures 768
  mechanisms 768–775, **770 – 771**
  nucleotide-dependent 777
  stereospecificity 768 – 775
Dehydrogenation
  of alcohols 765 – 767
  of amines 775
  of amino acids 775
  reactions of flavoproteins, table 784
  types of, table 766
Dematin 405
Denaturation of proteins 82
  by guanidinium chloride 82
  by sodium dodecyl sulfate 82
  by urea 82
Dental plaque 442
Dentin 442
5'-Deoxyadenosyl cobalamin 867, 867*s*, 872
Deoxyhemerythrin, oxygenation of 862
Deoxyhemoglobin 850, 851
Deoxy-D-mannooctulosonic acid (KDO) 180*s*
Deoxyribonuclease **652 – 653**
Deoxyribonucleic acid. *See* DNA
2-Deoxyribose 200*s*
1-Deoxyxylulose 5-phosphate 735, 736
Dephosphorylation, cascade 542
Depsides 23
Depsipeptide 414
Derepression of enzyme synthesis 536
Dermatan sulfate 170, 175, 176*s*
Dermatosparaxis 438
Dermis 439
Desaturases 863
Desmid 22
Desmosine 436, 436*s*
Desmosome 10
Desmosterol 777*s*
  reduction to cholesterol 777
Desulforedoxin 859
Desulfuration of insectides 636
Dextrans 170, 174
Dextrin(s) 606
Diabetes mellitus 567
Diabetic nephropathy 438
Diacylglycerols 563, 564
  scheme 565
  in signaling 563 – 566
Dialkylglycine decarboxylase 743, 745
Dialkylnitrosamines 254*s*
Dialysis 100
1,4-Diaminopropane 745
Diastereoisomers (diastereomers) 43, 162
Diastereotopic atoms or groups 479
Diatomaceous earth 21
Diatoms 21
  motion of 21
Dicoumarol 821, 821*s*
Dictyosome in micrograph 13
Dictyostelium discoideum 20
  cAMP as messenger 557
  genome 12
Dicyanocobalamin 870
3,6-Dideoxyhexoses, formation of 747
Dielectric constant 47
Diesenhofer, Johan 84
Diethanolamine, pK_a value of 99
Diethylmalonic acid, pK_a value of 99
Diethylpyrocarbonate reaction with histidine 126
Difference electron density map 136
Differential scanning calorimetric curves 395
Differentiation
  among prokaryotes 9
  cluster of (CD) 187, 407
Diffractometer 134
Diffusion 109, 410
  exchange 410
  facilitated 410
  mobile carrier 414
  saturation effect 410
  through membranes 410, 414
Diffusion constant 461
Diffusion-controlled limit on reaction rates 462
Digestion 505, 507
Digoxigenin 257
Dihedral (D$_2$) symmetry 338
D-*gluco*-Dihydroacarbose 607*s*
Dihydrofolate reductase 804 – 808, 810, 811
  active site structure 807
  inhibitors 805
  mechanism 805 – 808
  structure 805 – 808
Dihydrolipoic acid amide 784*s*
Dihydrolipoyl dehydrogenase 785, 796
Dihydrolipoyl transacetylase 796*s*
Dihydropteroic acid 473
Dihydropyridine receptor 314
Dihydroriboflavin 794
Dihydro U loop of tRNA 231
Dihydrouridine 234
Dihydroxyacetone 164*s*
Dihydroxyacetone phosphate 694*s*
Dihydroxyacid 697
Dihydroxyacid dehydratase
  Fe-S clusters in 861
ortho-Dihydroxybenzene. *See also* Catechol
3,4-Dihydroxyphenylalanine (dopa) 80, 887, 887*s*
2,3-Dihydroxyproline 80
Diiron desaturases 863
Diiron oxygenases 863
Diiron proteins **862 – 864**
Diiron-tyrosinate proteins 862 – 863
Diisopropylfluorophosphate (DFP, diisopropylphosphofluoridate) 610*s*
Dimethylallyl diphosphate 712*s*
Dimethylallyl pyrophosphate, condensation of 527
Dimethylarsine 876
Dimethylglutaric acid, pK_a value of 99
Dimethylimidazole 868*s*
Dimethylsuberimidate 125*s*
  crosslinking of proteins 125
Dimethylsulfoxide reductase 836*s*
Dimethylsulfoxide reductase family 890
Dinoflagellate 19
Dioldehydratase 871
  stereochemistry 872, 874
Dioxetane 820
Dioxygen 850*s*
Dipeptidyl aminopeptidase 117
Dipeptidyl peptidase 619
Diphtheria toxin, ADP-ribosylation 548
Diphytanyl group 385
Diphytanyl tetraether lipids 387
Diploid cell 17
Diploid phase 18
Dipolar ion 41
Disaccharides 167–169
  hydrolysis of, Gibbs energies 169
Discoidin I 186
Diseases, bacterial 7
Displacement reactions, nucleophilic 123, 482, 530, **589 – 670**
  on carbonyl groups 608 – 637, 698
  multiple 660
  on phosphorus atom **637– 657** *See also* Kinases, phosphatases, transferases
    in-line 638
    permutational rearrangement 638
  proteases 608 – 637
  rates of **589 – 590**
  on sulfur atoms 659, 660
Dissipative structures 289
Dissociation constant(s) 304
  apparent 293
Disulfide bridges (crosslinkages) 65, 80
  cleavage of 115 – 116
  cleavage with performic acid 115
  locating 119
  in proteins 80
  reduction of 115, 785
Dithioerythritol 115
Dithionite ion 779
Dithiothreitol 98, 115, 822
Dityrosine linkages 81
DMSO reductase 890
DNA
  A form 213, 214*s*
  within ancient conserved regions 9
  B form 213, 214*s*
    hydration of 213
    right handed helical structure 213
    side-by-side structure 213
  base composition of 239
  base inclination 216
  base roll 216

bending of 218
cDNA 257
circular, in trypanosomes 219
cloning of 201
coding capacity 16
coding strand of 204, 237
conformational flexibility 216
consensus sequences 265
content in tissues 31
denaturation of 255
density gradient centrifugation 101
double helical structure 201
electron micrograph of 213
electron transfer in 218
electrophoresis 221, 249
extraction of 249
fingerprinting 259
footprinting 266
gyrase 220
H-DNA, proposed structure 226
helical twist 216
helix parameters 216
Holliday junction 228 – 230
hydration of 216
hydrolysis of 249, 251
hypersensitive sites 266
interaction with spermidine and spermine 218
intercalation **221**, **223**
intrachain NOEs 268
knot, electron micrograph 219
light absorption 209
linking number 220
major groove 201, 213
melting curve 256
minor groove 201, 213
mispaired bases in 218
mitochondrial 14
  map 262
nanochemical scale construction 230
nearest neighbor analysis 265
negative strand 244
NMR spectroscopy **266 – 270**
  ^{1}H spectrum 269
non-coding (transcribing) strand 237
open reading frame 237
operator 237
overstretched 217
palindromes in 238
plectonemically interwound 220
polymorphism 259
propeller twist 216
R-DNA 226
reassociation curves, figure 256
reassociation kinetics 255
relaxed 221
repetitive 11
replicative form 244
sequence determination 260 – 266
sequences, understanding 265
sequencing gel, figure 263
single-stranded binding (SSB) proteins 239
solenoidal 220
structures 214, 223, 225, 228, 229, 240 – 243
supercoiled **219 – 222**
superhelix density 221
tetraplex (quadruplex) structures **227**
topological isomers 221
toroidal (solenoidal) 220
triplex 227
typing 259
underwound 220, 221
Watson-Crick structure 200
Z form 214*s*, 215, 239
  enhancers in 216
DNA-binding motifs
  in proteins 241
DNA-binding proteins
  helix-turn-helix motif 239
DNA, circular 218, 219
  linking number 220
  writhe 220
DNA gyrase 220
DNA, mitochondrial 14
  physical map 262
DNA photolyase 794, 802
DNA polymerase(s) 657
  Thermus aquaticus (*Taq*) 260
  zinc in 680
DNA - protein interactions 266
DNA-RNA hybrids 257
DNA sequencing 260 – 266
  chain-terminating inhibitors 262
  method of Maxam and Gilbert 264
  Sanger dideoxy method. 262
DNA supercoiling estimating amounts 222
DNase. *See* Deoxyribonuclease
Docking protein for SRP 520
Domains in protein structure 74
Donnan equilibrium 420
Dopamine 744
Dopamine β-monooxygenase 887
Double displacement reactions 595
  ping-pong mechanism 595
  in ribonuclease 647
Double helices 201, 213 – 218
  conformational flexibility 216
  hybrid DNA-RNA 230
Double reciprocal plot 460
Drawing arrows 528, 529
Drosophilia melanogaster 24
  genome 12
Drosopterins 803, 804*s*
DSS or Tier's salt 138
Dyad axis 134, 213
Dye, reducible 300
Dynamic equilibrium 289
Dyneins 370
Dysentery, *Shigella dysenteriae* 7
Dystrophin 405

E

E (entgegen) configuration 43
E3 binding protein 796
Eadie-Hofstee plot 460
E-cadhedrin 574
Echinodermata 25
Ectoderm 23
Ectoenzymes 409
Ectopic proteins 573
Edelman, Gerald M. 84
Edman degradation 118
EDTA. *See* Ethylenediaminetetraacetic acid
Effector(s) of allosteric enzymes 473 – 475
EF-hand motif 313, 317
EGF (epithelial growth factor), definition of 577
EGF-like domains 367
Ehlers-Danlos syndrome 438
Eicosanoid 565
Eigen, Manfred 84
Elastase 66, 609, 610, 611*s*
  cryoenzymology 616
  β-cylinder in 78
Elastic fibers 436
Elastin 15, 72, 436
Electrical double layer 400
Electric field jump methods 468
Electrochemical gradient 410
Electrochemical transference 311
Electrode(s)
  calomel 300
  reference 300
Electrode potentials 300 – 302
  measurement of 302
  observed 301
  standard, table 300
Electrogenic 423
Electromotive force of electrochemical cell 300
Electron carriers 512 – 514, **767– 802**, 843 – 850, 883
Electron crystallography 131
Electron micrograph
  of bacteria 4
  of cell junctions 27
  of plant cell 13
  of starch granules 172
  of viruses 246
Electron microscopy 129 – 131
Electron microscope tomography 130
Electron-nuclear double resonance (ENDOR) 399
Electron paramagnetic resonance (EPR, ESR) spectra **398**, **399**
  first derivative plot 398
  of putidaredoxin 860
  of tetramethylpiperidine-1-oxyl 398
Electron transfer
  biological, mechanism of **848 – 850**
  coupling 849, 850
  gating 849, 850
  ionic equilibria, effects of 850
  pathways 848
Electron transferring flavoprotein (EFT) 514, 794
Electron transport chain 512 – 515
  of mitochondria 513
Electron-transport proteins 514, 767– 802, 819, 843 – 850
Electronic noses 122
Electropherogram, single erythrocyte 107
Electrophilic center
  definition of 589
  in enzymes 754 – 757
Electrophilic reagents
  reactions of nucleic acids 253 – 255
Electrophoresis 58, **106 – 108**
  agarose gel 106
  capillary 106, 107
  diagonal 120
  pulsed field 249, 250
  SDS-polyacrylamide gel 106, 113
  zone 106, 249
Electrophoretic molecular sieving 106
Electrostatic effects
  on pK_a values 330
  titration curves of proteins 330
  transmission through aromatic rings 330

Electrostatic theory 311
Element(s)
  essential 33
  in human body 31
  toxic 33
Elimination reactions 526, 530, **677–690**
  beta, of cystine residues 85
  conjugative **689**
  decarboxylative **689**
  facilitation by carbonyl group 681
  of γ substituent 746
  of PLP-dependent enzymes 742
  reversibility 690
Ellman's reagent 125, 125*s*
Elongation factor EF-Tu 558
Elongin complex 564
Embden-Myerhof-Parnas pathway 508, 510
Embryos, plant 30
Emphysema 629
  treatment of 631
Enamel 442
Enamelins 442
Enamelysin 627
Enamines, 528*s*
  formation of 526
Enantiomers 164
  definition of 41
Enantiotopic atoms or groups 479
Endergonic reactions 286
Endo ring conformation 212
Endocytosis 13, 425–427
Endoderm 23
Endoglucanases 602
Endoglycanases 593, 602
Endonexin *See* Annexin
Endoplasmic reticulum (ER) 10, 13, 14, 521
  cisternae of 14
  definition of 14
  in micrograph 13
  rough 14
  smooth 14
Endosomes 13, 426
Endosperm 30
Endothelial cell-derived relaxing factor. *See* Nitric oxide, NO
Endotoxins 431
  cis-Enediol 693*s*
Enediol; Enediolate 694*s*, 695*s*
  intermediate 693
Energy (*E*) 282
  solar 289
  sources of 507
  storage and utilization of 302, 303
Enhancer 576
Enkephalin 54*s*
Enkephalin convertase. *See* Carboxypeptidase N
Enolase 685
  active site 685*s*
  metal ion requirement 685
  relaxation of water protons 685
Enolase superfamily 686
Enolate anion 46, 46*s*, 681, 698–710
  formation of 526
  as nucleophile 527
  from oxosteroid 696
  phosphorylation of 706
  resonance-stabilized 494
Enolic forms 530
Enolic intermediates **691–699**
Enolpyruvoylshikimate-3-phosphate. *See* EPSP
Enoyl-acyl carrier protein reductase 777
Enoyl hydratase 681
Enoyl reductase 766
Entamoeba, PEP carboxytransphosphorylase of 706
Entamoeba histolytica 19
Enterobactin (enterochelin) 838*s*
Enterokinase. *See* Enteropeptidase
Enteron 24
Enteropeptidase 609
Enthalpy 282
  of combustion 282, 283
Enthalpy change(s) 282, 283
  estimation of 297
  of dissociation, table 293
Entner-Doudoroff pathway 697
Entropy (*S*) 284, 285
  of gases 284
  from heat capacities 285
  of racemization 284
  relationship to microscopic disorder 284
  table 285
  units, definition of 284
Environment (thermodynamics), definition of 282
Enzymatic reactions
  catalysis 482–497
  classification of 530
  enolic intermediates 691–699
  first order 457
  integrated rate equations 457–460
  pH, effect on 469–471
  progress curve 456
  reversible 458, 464–468
  second order 458
  stereoelectronic effect 495
  velocity of 455–458
Enzyme(s) 1, 2, **455–503**
  activation of **471–478**
  active site 491–494
  biotin-containing 724, 725
  table 724
  catalysis by 482–497. *See also* Catalysis (enzymatic)
  classification 497
  conformational changes in 481
  constitutive 538
  dead-end complexes 466
  debranching 606
  effectors of 474
  equilibrium 467
  feedback inhibition, figure 540
  history of 456
  hydrogen bonding in 491–494
  induced fit 481
  inducible 538
  inhibition of **471–478**
    competitive 471, 472
    irreversible 477, 478
    noncompetitive 473–475
  inverting 593
  isologous dimers 349
  kinetic parameters 459
  kinetics of action **455–471**
  mechanism
    iso 466
    ping-pong 465
    reciprocating (flip-flop) 497
  modifiers of 474
  molecular activity of 457
  molybdenum-containing 890–893
    Fe-S clusters in 862
  multisubstrate, kinetics of 464, 465
  oligomeric 475, 497
  pacemaker 535–537
  processive action 602
  prochiral centers 478–481
  proofreading 482
  pyridoxal phosphate, coenzyme table 743
  rate equation for 460
  rate of encounter with substrate 461–464
  reaction rate 455–458. *See* Kinetics
  regulation. *See* Enzyme regulation
  retaining 593
  saturation of 457
  selenium-containing 822–827
  specific activity of 457
  specificity of **478–482**
  starch-hydrolyzing 605–608
  stereospecificity 478–481
  substrate binding 463, 464
  tetrameric 342
  thiamin-dependent, structures 733
  thiamin diphosphate-dependent, table 735
Enzyme-activated inhibitors 682, 864
Enzyme assay, coupled 768
Enzyme-inhibitor complex 472
Enzyme regulation 535–537, 539–553, 571–581
  allosteric control 539–541
  by allosteric effectors 475–477
  by covalent modification 541–545
  by group transfer modification 545–549
  by hormones 553–563
  by ions 549
  by thiol-disulfide equilibria 549
Enzyme-substrate (ES) complexes
formation and reaction of 458–461
Enzyme synthesis, genetic control of. *See also* Protein synthesis
Eosinophils 26, 188
Epidermal growth factor (EGF) 544, 571
Epidermis 439
Epidermolysis bullosa 440
Epimerases 692
Epimerization of sugar, mechanisms 778
Epimers, definition of 163
Epinephrine (adrenaline) 542, 553, 553*s*
Episomes. *See* plasmid
Epithelial cells 29
Epitheliocytes 25
Epoxides, alkyation by 254
Epoxide hydrolases 591
EPR (electron paramagnetic resonance) spectroscopy 398, 399
  of glutamate mutase 873
  in study of phosphotransferases 639
EPSP (enolpyruvoylshikimate-3-phosphate) 687*s*
  synthase 687
Epstein-Barr virus 247
Equilibrium in carboxylation reactions 725
Equilibrium constant(s) 304
  for association **325–332**
  for binding of metal ions **307**
  for carboxylation of propionyl CoA 725
  changes with temperature 289
  from concentrations 288

    effect of pH 293 – 297
    for hydrolysis 292
    prediction by Hammett equation 309
    as ratio of rate constants 458
    relationship to Gibbs energy change 287
    stepwise, for complex formation 307
c-erb 571
v-erb 571, 572
Erhlich, Paul 473
ERK 577– 579
Erythrocruorins 68, 362
Erythrocytes 26, **404 – 406**
    cytoskeleton 405
    defects 826
    integrin 405, 406
    membrane skeleton 406
    NADP$^+$ from 767
    size of 11
Erythrocyte membranes
    band 3 protein 420
    proteins of 403
Erythrose 163*s*
Erythrulose 164*s*
Escherichia coli **3**
    composition of 31
    DNA of 4
    envelope structure 428
    genome 12
Essential trace elements 31, 33, 837, 877– 881, 887– 894
Ester(s)
    displacement reactions of 608
    mechanism of formation 661, 662
    resonance energy 299
Esterase(s) 526, **634 – 637**
Ethanol 508
Ethanolamine, pK_a value of 99
Ethanolamine ammonia-lyase 871
Ether lipids 382
Ether phosphatides 383*s*
Ethidium 223*s*
Ethidium bromide 221, 222
Ethylenediamine, binding constants to metal ions 311
Ethylenediaminetetraacetic acid (EDTA) 98, 311*s*
N-Ethylmorpholine, pK_a value of 99
ets 576
Eubacterium 3, 6
Euglena 19
Euglenophyta 20
Eukaryote(s), unicellular 18 – 22
Eukaryotic cells 11– 15
    definition of 2
    iron uptake 839
    structure 10
Evolution 9
    of eukaryotes 15 – 18
Exchange diffusion 410
Exchange rates, amide protons 148
Exchange reactions of enzymes 595
Exergonic reactions 285
Exo ring conformation 212
Exocytosis 14, 427
Exoglycanases 593
Exponential growth rate constant 470
Extensins
    glycosylation of 181
    in plant cell wall 72
Extracellular matrix (ECM) 406, **427– 441**

F

FAD (flavin adenine dinucleotide) 511, 511*s*, 765, 780, 781*s*
FADH$_2$ 529, **529*s***
Familial adenomatous polyposis 574
Faraday, numerical value of 283
Farnesyl group 402, 559
Fat(s). *See also* Triacylglycerol (triglyceride)
    composition of 380
    hydrolysis of 507
Fatty acid(s) **380 – 382**
    activation of 512
    acyl CoA, derivatives of 507
    biosynthesis of 722
    branched chain 381
    cyclopropane-containing 399
    essential 721
    in lipids 380
    names of, table **380**
    oxidation 511
    pK_a values of 380
    stability of 589
Fatty acyl-CoA derivatives 507
Fatty alcohols **380 – 382**, 382
Favin 64*s*
Feedback inhibition 536, 539
Feedback repression 536
Ferments 456
Fermentation(s) 8, 508 – 511, 535
    of glucose, Gibbs energy change 298
    glyceraldehyde-3-phosphate dehydrogenase in 775, 776
    reactions 298, 510
Ferns 18
    gametophyte and sporophyte forms 18
Ferredoxin(s) 793, 857, 858*s*
    from chloroplasts 859
    reduction potentials 857
Ferredoxin:NADP+ oxidoreductase 788, 794
Ferrichrome 838*s*, 839
Ferric uptake regulation protein 839
Ferriheme(s) 844
Ferritin 343*s*, 841*s*, 841, 842
Ferrochelatase 843
Ferroelectric crystals 486
Ferroheme(s) 844
Ferrous ion in aconitase 688
Ferroxidase center 843*s*
    in ferritin 842
Ferryl iron in peroxidase and catalase 854
Fetal hemoglobin 362
α-Fetoprotein 573
FhuC, D 839
Fiber cells 30
Fibrillin 440
Fibrin 631, 632
    stabilizing factor 632
Fibrinogen 58, 409, 578, 632
Fibroblasts 25, 433
    micrograph of 370
Fibrocytes 25
Fibronectin 407, 409
    domain structure 409
Fibrous dysplasia 576
Fibulin 437
Ficin 618
Field flow fractionation 103
Fifth disease 244
Filaggrin 370, 439
Filamin 370
Filaria worms 24
Fimbriae 6. *See also* Pili
Fingerprinting. *See also* Peptide mapping
    of DNA 259
    of proteins 118, 360
First Law of Thermodynamics 282
First order reactions 457
Fischer, Edmond H. 84
Fischer, Emil H. 42, 83
Fischer projection formula 42
    of monosaccharides 163
FK506 488
Flagella
    bacterial 3, 6, 562
    eukaryotic structure 15
Flatworms 23
Flavin(s)
    chromatography of 103
    electrophilic centers 790
    fluorescence 103
    half-reduced 791
    isoalloxazine ring 790
    metal complexes 794
    properties of 793
Flavin adenine dinucleotide. *See* FAD
Flavin adenine diphosphate. *See* FAD
Flavin coenzymes 766, **780 – 795**
    modified 788, 789
    reduced 794
Flavin radicals 792
    color of 794
    formation constant 794
Flavocytochrome *b*$_2$ 782, 794, 847
Flavodoxins 793, 799, 800
Flavoprotein(s) 513, 788
    dehydrogenation reactions 782 – 787
        table 784
    reduction potentials of 782
Flavoprotein dehydrogenases 789
    free radicals 793
    mechanisms 789 – 791
Flavoprotein oxygenases
    free radicals 793
Flow of materials within cells 289
Flow cytometry 107
Flukes 24
Fluorescamine 120*s*
Fluorescence
    in analysis 120
    of flavins 103
    recovery after photobleaching 400
    of vitamin B$_6$ coenzyme 750
Fluorescence probe
    polarity-dependent 396
Fluoride ion
    inhibition of acid phosphatases 645
5-Fluorouracil 812
Fluorographs 103, 252
5-Fluorouridine 574
Flux
    across membranes 410
    in metabolism (analysis) 536, 537
FMN (Riboflavin 5'-phosphate) 513, 765, 780, 781*s*
Focal adhesions 405, 406
Fodrin 405
Folate coenzymes, 803, 804. *See also* Folic acid
Folate receptors 805
3-Fold axis 333
Folding patterns
    of polypeptides **76**, **77**

  effects of sequence 78
Folic acid (pteroylglutamic acid) 473, 557, 721, 802
  deficiency of 802
  metabolic functions 802
  nutritional requirements 756
  in transmethylation 592
Folyl polyglutamates 803
Footprinting
  DNA 266
  hydroxyl radical 267
Foraminifera 19
Forces
  acting between molecules **46 – 51**
  electrostatic 47
  van der Waals 46
Formaldehyde 802
  hydration equilibrium 677
  reaction with nucleic acid bases 253
Formate. *See also* Formic acid
  carbon units, source of 810
Formate dehydrogenase 824, 890, 892, 893
  molybdopterin 892
Formate dehydrogenase H 893
Formation constants 304. *See also* Equilibrium constants
  macroscopic 328
  of metal complexes 312
  for Schiff bases 679
  stepwise 327, 328
Formic acid 802
  pK_a value of 99, 293
Formiminoglutamate 810
Formiminoglycine 810
Formyl-THF synthetase 810
Formylglycine in sulfatases 659
Formylmethanofuran 814*s*
N-Formylmethionine 231
6-Formylpterin 803
10-Formyltetrahydrofolic acid (l0-formyl THF) 810
Fos 70, 241, 572
  definition of 577
fos 572, 576, 579
Fossils of bacteria and blue-green algae 17
Fourier transform infrared spectroscopy
  of membranes 396
Fractionation
  of cells and tissues 98 – 100
Fragile X syndrome 228
Frame-shifting 538
Franck-Condon principle 848
Franklin, Rosalind N. 200
Free energy. *See* Gibbs energy
Free radicals. *See also* Superoxide anion radical, nitric oxide (NO)
  in action of peroxidase and catalase 855, 856
  in coenzyme M reductase 880
  flavin 791– 794
  in metalloenzymes 850 – 886
  in pyruvate formate-lyase 800 – 801
  in pyruvate oxidase 736, 799
  in reactions of vitamin B_{12} 869 – 876
  in ribonucleotide reductases 864, 865
Freeze-fractured membranes, figure 405
Freeze fracturing 130
FreIp metalloreductase 839, 883
Frictional coefficient 109
  of sphere 109
Fructofuranose (Fruf) 165*s*
Fructofuranosyl-enzyme 597*s*
Fructose (Fru) 162*s*, 175
Fructose 2,6-bisphosphate (Fru-2,6-P_2)
  in regulation of glycogen metabolism 537
  beta form 541*s*
Fructose bisphosphate aldolases 699
Fructose 6-phosphate 535, 693*s*
Fructose 6-phosphate kinase 656
Fructose-1,6-bisphosphatase 645
Fructose-2,6-bisphosphatase 646
Fruit fly. *See Drosophilia melanogaster*
Fucose (Fuc) 165*s*
L-Fucose isomerase 693
Fucosyltransferase 184
Fucoxanthin 22
Fumarase. *See* Fumarate hydratase
Fumarase-aspartase family 685
Fumarate 481*s*, 516*s*, 683*s*
Fumarate hydratase (fumarase) 526, 683, 688
  acid – base catalysis 471
  concerted reaction 685
  Fumarase A 688
  Fumarase B 688
  Fumarase C 683
  mechanism 683 – 685
  pH dependence 684
  rates of substrate exchange 684
  turnover number of 683
Fumarate reductase 785
Fumarylpyruvate 690*s*
Function of state *R* 476
Fungal infections 20
Fungi **20**
  cell walls 443
Furanose rings 161, 175. *See also* Sugars
Furfural 251, 167*s*

G

G-quartets 208, 227
G_0 phase of cell cycle 579
Gabaculline 738*s*
Galactan 177
Galactanase 602
Galactolipid 387*s*
Galactosamine 176
Galactose (Gal) 162, 163*s*
  metabolism of 778
Galactose chemoreceptor protein 186
  active site 187*s*
Galactose oxidase 885
  active site 886
  mechanism of 887
  tyrosine-cysteine cofactor 885, 886
Galactosemia 16, 169
β-Galactosidase 602, 603
Galectins 186
Gametes 17
Gametophyte form of fern 18
Gangliosides 388
  of cholera toxin 546
Gap junctions 29
Garrod, Archibald 16
Gas constant, numerical value of 283
Gastricsin 621
Gated pores 411
GDP dissociation inhibitors (GDIs) 559
Gel electrophoresis 112
  relative molecular mass by 112
Gel filtration 100, **112**
  oligosaccharides, separation of 100
  relative molecular mass by 112
Geminal diamine 751, 751*s*
Geminiviruses 244
Gene(s)
  amplification of 538
  average number of nucleotide pairs in 4
  human, number of 16
  imprinting of 234
  sequence for aspartate aminotransferase 202
  structural, definition 1
Gene therapy 16, 203
  sickle cell disease 361
Genetic code 1, **236 – 239**
  tables 236, 237
  variations 237
Genetic diseases 2, 16
Genetic recombination 6, 230
Genomes
  bacterial 3 – 5, 201
  changing 17
  definition of 4
  human 201
  sequences 12,
  table of 12
  of viruses 201
Geometrical isomers 43
Geometry of orbitals
  effect on metal binding 311
Geranylgeranyl group 402
Giardia lamblia 12, 19
Gibbs energy (Free energy) 285, 286
  of activation ($\Delta G^‡$) 483
  of association 325 – 332
  of dissociation 293
    of acetic acid 50
  of formation 286, **292 – 298**
    for biochemical substances, table 290
    of RNA helix, table 210
  of hydrolysis 283
    table 294
  in hydrophobic interactions 50
  linear relationships 308
  of oxidation
    for biochemical substances, table 290
    by NAD$^+$, table 290
  partial molar 287
  summing 286
Gibbs energy change(s)
  for dilution 287
  and equilibrium constant 287
  for fermentation 298
  for oxidation-reduction reactions 300
  for subunit association 333
Gingipains 619
Gla proteins 821
Glial cells 26
Glial filaments 369
Globin, evolution of gene 17
Globular proteins, properties 59
Globulin(s) 101, 58
Glucagon 542
  adenylate cyclase in action of 554
  receptor 554
β-Glucanases 602, 603
Glucans **172 – 175**
  β-1,3-linked 174

Glucarate dehydratase 686
D-Glucitol. *See* D-Sorbitol
Glucoamylase 603, 607
  α,α barrel structure, 77*s*
  active site structure 608
Glucocerebrosidase 603
Glucokinase 538, 568
Gluconeogenesis 515
  pathway 509
  regulation of 541
5-Gluconolactone 604
Gluconolactone 604*s*
Glucosamine (GlcN) 164, 176
Glucose (Glc)
  caloric value 283
Glucose (D-glucose, Glc) 161*s*, 162*s*
  concentration in blood 161
  diastereomers 162
  equilibrium mixture in water 162
  fermentation of 298
  ring form 162*s*
Glucose (L-glucose, L-Glc), 161*s*
Glucose 1-arsenate as intermediate in arsenolysis 595
Glucose 1,6-bisphosphate 653, 653*s*
Glucose dehydrogenase(s) 814
Glucose oxidase 782
Glucose-6-phosphatase 645
Glucose 6-phosphate 506*s*, 507, 535, 693, 693*s*
Gibbs energy of hydrolysis 293
  pK_a value of 293
Glucose-6-phosphate dehydrogenase 770, 776
Glucose-6-phosphate isomerase
  efficiency of 692
  turnover number of 692
Glucose tolerance factor
  chromium in 888
Glucose transporters 415, 416
GLUT1, topology diagram 416
GLUT4, response to insulin 416
Glucosidase
  alpha 603, 606
  alpha 1,6 606
  beta 602, 603
Glucosylasparaginase (aspartylglucosaminidase) 620
Glucuronic acid; glucuronate (GlcA) 164*s*, 165*s*, 176
  epimerization 177
GLUT1–GLUT7. *See* Glucose transporters
  L-Glutamate 737*s*
  biosynthesis 711
Glutamate dehydrogenase 770, 775
  inhibition by 2-oxoglutarate 780
Glutamate mutase 871
  stereochemistry 874
Glutamate-1-semialdehyde 742
Glutamate synthase 788
Glutamic acid (Glu, E) 52, 53*s*. *See also* Glutamate
  pK_a value of 487
Glutamine (Gln, Q) 53*s*
Glutamine PRPP amidotransferase 620
Glutamine synthetase 342, 662
Glutaredoxin 522, 549, 786, 787, 864
Glutathione 521, 549, **550–552**, 550*s*
  biosynthesis 551
  protective functions 551, 786, 826
Glutathione peroxidase 550, 826, 856
  selenium in 823
Glutathione reductase 785, 785*s*
  reaction mechanism 791
    scheme 792
Glutathione *S*-transferases 550, 682
Glycans **169–180**. *See also* Polysaccharides
  N-linked 183
Glycation 169
Glyceraldehyde 163*s*
Glyceraldehyde 3-phosphate 506*s*, 508, 694*s*
  oxidation of 508–510
Glyceraldehyde 3-phosphate dehydrogenase 63*s*, 766
  in fermentation reactions 775, 776
  formation of NADH-X 780
  in oxidation of aldehydes 775, 776
  thiol group of 677
D-Glycerate dehydrogenase, stereochemistry 768
Glycerol 380*s*
  accumulation in fish and insects 191
  in lipids 380
Glycerol 1-phosphate, pK_a value of 293
Glycerol dehydratase 871
Glycerol oxidase 886
sn-Glycerol 3-phosphate 383*s*
Glycerolteichoic acids 431*s*
L-*Glycero*-D-*manno*heptose 180*s*
Glycerophospholipids 383
Glycerylcysteine 428, 428*s*
Glycine (Gly, G) 52, 52*s*, 745
  binding constants to metal ions 311
  cleavage system, scheme 809
  decarboxylation of 745
  pK_a value of 99
  reductases 753, 755
Glycineamide, binding constant to metal ions 311
Glycine decarboxylase 743, 745, 809
Glycine decarboxylase-synthetase system 808
Glycine *N*-methyltransferase 591, 593*s*
Glycine reductase 824
Glycocalyx 6
Glycogen 165, 170, 173, 173*s*, 603
  breakdown of 537
  content in rat liver 31
  methylation analysis 196
  ^{13}C NMR 174
  particles in cytoplasm 10, 174
    STM image of 174
  storage of 537
  synthesis 509
Glycogen phosphorylase 541–545, 543*s*, **604**, **605**, 746
  catalytic site 605*s*
  conformational states 604
  control of **541–545**
  effect of insulin 571
  regulation of 539, 604
Glycogen synthase
  effect of insulin 571
Glycogenin 174
Glycogenolysis 545
Glycoglycerolipids 387
Glycolate 707
Glycolate oxidase 782
Glycolipids **387**, 392
  biological recognition 161
Glycolysis 506, 508–511, 699
  regulation of 541
Glycophorin A 404
Glycophorin C 405
Glycoproteins 15, **180–188**, 404, 405
  biological recognition 161
  blood group determinants 184
  high mannose type 183
  linkage of saccharides to 181
  O-linked 181, 182
  release of oligosaccharides 188
Glycosaminoglycan 181
  repeating disaccharide units of 176
Glycosidases
  beta 602
  endo 605
Glycosides
  displacement reactions of 593
N-Glycosides (glycosylamine) 169
Glycosidic acetal linkages 168
Glycosidic linkages 780
Glycosylamine 169
Glycosylation
  definition of 169
Glycosyl enzyme intermediates 595
Glycosyl esters 595
Glycosyl exchange reactions 593
Glycosylphosphatidylinositol (GPI) anchors 403
Glycosyltransferase(s) 184, 526, 604, 608
  catalytic groups, table 603
Glycyl endopeptidase 618
Glycylglycine
  binding constants to metal ions 311
  copper complexes 312*s*
  pK_a value of 99
Glycyl radical 864
Glyoxalase 527, 551, 698
Glyoxalase I, II 698
Glyoxylate 705*s*
Glyoxylate carboligase 736
Glyoxylic acid
  in polysaccharides 179
Glyoxysomes 14
Glyphosate 687*s*
GMP (Guanosine 5'-phosphate) 203
Gold 318
Goldstein, Joseph L. 84
Golgi apparatus 10, 521
  definition of 14
Golgi body in micrograph 13
Golgi, Camillo 83
Gonadatropin-releasing hormone
  NMR spectrum 145
Gonorrhea 7
Goodpasture syndrome 438
G protein(s) 534, 557–561
  bovine brain G_i 560*s*
  function, scheme 557
  monomeric 558
  three-dimensional structures 559–561
Gramicidin A 414
Gramicidin S 54*s*
Gram-negative bacteria 6, 7
  outer membrane 428–431
Gram-positive bacteria 6, 7, 185, 190, 428
Gram stain 6
Granulocytes 26
Granzymes 610
Grb2 577, 578, 580*s*
Green plant, composition of 31
GroEL 339, 518
GroES 339, 518
Ground substance between cells 26
Group transfer potential(s) 292, 293
  effect of metal ions 296, 297
  of pyrophosphates 636

Group translocation 411
Growth, control of 578
Growth factors **571–576**, 722
Growth rates of cells 470
Growth regulating hormones 567–571
GTP (guanosine triphosphate)
  as coenzyme 719
GTPase activating proteins (GAPs) 559
Guanidination 124
Guanidinium chloride 82*s*
Guanidinium groups 53, 55*s*
Guanidinoacetate methyltransferase 591
Guanine (Gua, G) 199*s*, 203
  hydration of 215
  quartet 208
Guanine nucleotide binding proteins. *See* G proteins
Guanine nucleotide dissociation stimulators (GDSs) 559
Guanosine (Guo) 203, 234*s*
  absorption spectra of 205
Guanosine 5'-(β,γ-imido) triphosphate (GppNp) 558*s*
Guanosine5’phosphate (GMP,Guanosine 5'-guanylic acid) 203
Guanylate cyclase 561, 657
Guanylin 548, 561
Guanylyl cyclase. *See* Guanylate cyclase
Guard cells 30
Gulose (Gul) 163*s*
Gustducin 558
Gutta-percha 389
Gymnosperms 29

H

Haemophilus influenzae
  genome sequence 12
Hageman factor 632
Hairpin ribozyme 649
Haldane, J. B. S. 9
Haldane relationship 464, 465
Half-life 457
Haloacid dehalogenase(s) 590
  mechanism of 590
Haloalkane dehalogenase(s) 591
  active site structure 591
Halocyanin 883
Haloperoxidases 855, 889
Hammerhead ribozyme 649, 651*s*
  mechanism of action 651
Hammett equation 308
  for dehydrogenation reactions 771
  for multisubstituted compounds 309
  table of constants 308
Handedness of molecules 41–43
Haploid
  definition of 5
  phase 18
Haptoglobin 58
Harden, Arthur 83
Haworth structural formulas 164
Hay fever 26
Heart disease 16
Heat 282
  of combustion 282
  flow of 284
Heat shock proteins (chaperonins) 518
Heisenberg uncertainty principle 138
HeLa strain of human cancer cells
  30 S ribosomal subunit 26
Helical structure
  polyethylene 44
  polyfluoroethylene 44
Helical wheel 69*s*, 70
Helix (helices). *See also* Superhelix
  3_{10} 69, 70
  amphipathic 79, 401
  hydrogen bonds in 68
  one-start 334
  Pi 69
  pitch of 68
  primary 334, 337
  properties of 69
  radial projection 334
  in serum albumin 71
  stacking of 70
  structures of 334
  torsion angles of 68–71
α Helix (helices) 68*s*, 69, 78
  as helical wheel 69*s*, 70
β Helix (helices)
  parallel 65
Helix bundles
  four helix 71
Helix-turn-helix motif 76
Hemagglutinin 186
Heme 68, 353, 514, 844, 850, 851
  covalent linkage to proteins 719
  imidazole group in 851
Heme *a* 844, 845*s*
Heme *c* 845, 845*s*
Heme *d* 845, 845*s*
Heme *o* 845*s*
Heme proteins **843–845**
  reactions with hydrogen peroxide **850–857**
  reactions with oxygen **850–857**
Heme-oxygen complexes 851
Hemerythrin 362, 852, 862
  active site, 863*s*
Hemiacetals 161, 162, 677. *See also* Glycosidic linkages
  formation of 526
Hemicelluloses 177, 602
Hemiketals 677
  formation of 526
Hemimercaptals 677
  formation of 526
Hemin 844
Hemochromatosis 840
Hemochrome(s) 844
Hemocyanins 885
Hemoflagellates 19
Hemoglobin(s) 17, **353–359**, 844
  abnormal 354, 359–362
  2,3-bisphosphoglycerate, effect 357, 358
  carbamino groups in 358
  comparative biochemistry 362
  competing ligands 358
  fetal 362
  fingerprints of 359
  folding pattern 354
  oxygen binding cooperativity 851
  in oxygen transport 354
  oxygenation curves 355
  pK_a values of 357
  salt bridges 357
  saturation curve 355
  structural changes during oxygenation 354–357
  subunits 353
  thermodynamic function for oxygenation, table 355
Hemoglobin mutants 354, 359–362
  F 362
  H 354
  M 359
  S 359, 360
Hemopexin 58, 627
Hemophilia A, B 633
Hemosiderin 842
Henderson-Hasselbalch equation 97
Heparan sulfate 175–177, 176*s*
Heparin 176*s*, 181
  anticoagulant properties of 177
  pentasaccharide segment 177*s*
Heparin cofactor II 177
Hepaticae 29
Hepatitis A virus 247
Hepatitis B virus 244
Hepatitis delta virus 247
HEPES buffer 99
  pK_a value of 99
Heptad in coiled coil 71
Heptanoyl threonine phosphate (HTP, coenzyme B) 881
Herbicides 678
Herbicides
  glyphosate 687
Herpesviruses 247
Heterocysts 10
Heteroduplex 256
Heterologous bonding interaction 333, 334
Heterologous interactions 342
Heterologous square 352*s*
Heteropolysaccharide(s)
  of animal body 175–177
  definition of 169
  of plants 177–179
Heterotropic interactions 476
Hexagonal phase
  of alkanes 394
  of lipid bilayer 394, 395
Hexokinase 467, 518, 654, 656
  in regulation 535
Hexoses 162
High performance liquid chromatography (HPLC) 103
High potential iron proteins (HIPIP) 857
Hill plot 331, 332
Hippuric acid 513*s*
Hirudin 634
Histamine 745
Histidine (His, H) 53*s*
  binding constants to metal ions 311
  catabolism of 810
  diethylpyrocarbonate, reaction with 127
  pK_a value of 487
  titration curve 97
Histidine ammonia-lyase 755
Histidine-binding protein 419*s*
Histidine decarboxylase
  active site 755
  mechanism 754
  proenzyme 754
Histidinohydroxymerodesmosine 434, 434*s*
Histone(s) 204
  complex with DNA 221
  H2A, ubiquitination of 525

HIV (human immunodeficiency virus) 248
HIV-1 protease 624*s*
    catalytic mechanism 625
Hofmeister series 311
Holley, Robert W. 84
Holliday junction 229*s*
    branch migration in 229
    in DNA **228–230**
    immobile 230
Homeostatic mechanisms for metal ions 307
Homoaconitase 688
Homocitrate 700*s*
L-Homocysteine 746, 746*s*, 778*s*
Homogenization 98
Homology, definition of 59
Homopolysaccharide(s) 175
    definition of 169
Homotropic interactions 476
Hoogsteen base pair 207
    reversed 208
Hookworms 24
Hopkins, Frederick G. 83
Hormone(s). *See also* Specific hormones
    receptors for 479, 536, **553–563**
    regulation of enzymes 553–563
Hormone-sensitive lipase
effect of insulin 571
Horseradish peroxidase 852, 856
Horseshoe crabs 24
Housing silk 67
HPLC. *See* High performance liquid chromatography
HSQC 144
Huber, Robert 84
Human body, elements in 31
Human genome 12, 201
Human glucose transporters 416
    figure 416
Human immunodeficiency virus. *See* HIV
Hyaluronan 170, 175–177, 176*s*
    hydrogen-bonding 176
    properties of 175
    repeating disaccharide units of 176
Hyaluronate lyase 686
Hyaluronic acid 15, 182. *See also* Hyaluronan
Hyaluronidase. *See* Hyaluronate lyase
Hybrid duplexes 256
Hycanthone 223*s*
Hydra 23, 24
Hydration
    of alkenes, mechanism 683
    effects on biochemical equilibria 50
    of ions 50, 311
    of nucleic acid bases 213
    of polar molecules 50
Hydration layer 51
Hydride ion
    shift of 527
    transfer of 765
Hydride-transfer mechanism 789
Hydrocarbon(s) **380–382**
α-Hydrogen
    loss from PLP Schiff bases 741
Hydrogen (H_2), *See also* Hydrogenase
    formation by bacteria 300
Hydrogenase 861, 878, 879
Hydrogen atoms, direct transfer of 768
Hydrogen bond(s) **47–49**, 69. *See also* Hydrogen bonding
    chains in proteins 56, 64
    in collagen 72
    competition for 82
    in hyaluronan 176
    in ion pairs 75
    lengths of 48, 493, 494
    low-barrier 494
    in polypeptides 56, 61
    in proteins 75
    strengths of 49, 494
    strong, in papain 619
Hydrogen-bonded network
    in carbonic anhydrase 678
    in proteins 75
Hydrogen bonding **491–494**. *See also* Hydrogen bond
    chemical shift, effect on 139
Hydrogen transfer reactions, table 766
Hydrogenase(s) 878, 879, 879*s*
    iron-only 878
    Ni containing 879
Hydrogenosomes 15
Hydrolases 589
    definition of 498
Hydrolysis 593, 608
    of biopolymers 116–118
    of carbohydrates 189
    of fats 507
    Gibbs energies of, table 294
    of proteins 116–118
Hydropathy index 78, 79
Hydrophobic effect 50
Hydrophobic interaction chromatography 103
Hydrophobic interactions **50**, **51**. *See also* Hydrophobic bonding
Hydrophobicity 51
Hydroquinone(s) 815–822
Hydroxamate 838
Hydroxide ion, binding constants to metal ions 311
Hydroxocobalamin 869
S-3-Hydroxy-3-methylglutaryl-CoA (HMG-CoA) 700*s*
4-Hydroxy-l-proline 432*s*
erythro-β-Hydroxyaspartate
    reaction with aminotransferase 750
Hydroxyl radicals 856
Hydroxylamine 251
    reactivity of 590
Hydroxylapatite 103, 441*s*
Hydroxylysine 432*s*
5-Hydroxymethylcytosine 234, 810
Hydroxymethylfurfural 167*s*
3-Hydroxy-3-methylglutaryl-CoA 699, 701*s*
    lyase 699
    reductase 776
    synthesis of 699
5-Hydroxymethyluracil 234
3-D-Hydroxymyristic acid 429
4-Hydroxynonenal 682
Hydroxyproline 52*s*, 72
2(α)-Hydroxypyridine 491*s*
    as catalyst 490, 491
    tautomerism 45
Hydroxypyridine crosslink in collagen 434
D-10-Hydroxystearate from hydration of oleic acid 688
Hyperfine structure of epr signal 398
Hyperpolarizability 70
Hyphae 20
Hypochromic effect 209, 255
Hypoxanthine
    tautomerization of 203

I

ICAM-1
    diagram 408
Ice
    hydrogen bonds in 49
    melting of 284
    reaction rates in 492
    structure of 49
Ice-nucleation proteins 191
Ichthyoses 439
Icosadeltahedra 346
Icosahedral viruses 345
Icosahedron 344
Idose (Ido) 163*s*
Iduronate, iduronic acid 176
    conformation 166
Image reconstruction 130
Imidazole
    binding constants to metal ions 311
    pK_a value of 99, 487
Imidazole group(s) 46*s*, 53
    hydrogen bonding 48
    numbering of ring atoms 46
    as a polarizable base 590
    in proteins 312
    protonation of 46
    tautomerism of 46
Imidazolium cation 46, 46*s*
Imine(s), 679. *See also* Schiff bases
    in enzymes 699, 740–746, 750
Immunoglobulin(s) 58, 74
    repeat 367
    superfamily 187
Immunosuppresant 488, 805
Indole, resonance energy 299
Indole-3-acetate 734
Induced fit 481
    model 351
    theory 349
Inducer(s) 539
    of gene expression 539
Inducible enzyme 538
Induction of enzyme synthesis 538, 539
Inflammation 384
Inhibition
    classical noncompetitive 474
    of enzyme synthesis 536
    of enzymes 471–478
    noncompetitive 473–475
    uncompetitive 475
Inhibition constant 472
Inhibitors
    competitive 471
    enzyme-activated 478, 864
    in study of enzyme mechanisms 475
    transition state 484
Initiation codon 236
Inorganic deposits by cells 15
Inosine 203
Inosinic acid 203
myo-Inositol 387*s*
Inositol 1,3,4,5-tetrakisphosphate ($InsP_4$) 564
Inositol 1,3,4,5,6-pentakisphosphate ($InsP_5$) 564
Inositol 1,4,5-trisphosphate ($InsP_3$) 563
    receptors 564
Inositol hexakisphosphate ($InsP_6$) 564
Inositol phosphates, scheme 565
Inositol polyphosphates

  in signaling 563 – 566
Inositol trisphosphate 542
Insecta 24
Insecticides **636**
Instantaneous velocity 456
Insulin 346*s*, 544, **567– 571**
  effects on acetyl-CoA carboxylase 568
  effects on ATP citrate-lyase 568
  effects on GLUT4 568
  effects on ribosomal subunit S6 568
  glucose uptake 570
  half-life 571
  hexameric, 347*s*
  metabolic effects of 568
  phosphorylation 571
  response element 578
  second messenger 569 – 571
Insulin receptor 568, 569
  autophosphorylation **569**
  down-regulation 571
  mutant forms 569
  protein kinase of 544
  structure (schematic) 569
  substrates 559
  substrate-1 (IRS-1) 570
  tyrosine kinase domain 569, 570*s*
Integral proteins of membrane 391, 401, 402
Integrated rate equations 457– 461
Integrin(s) 187, 405, 406
Interaction constant 352
Intercalation in DNA **222**, **223**
Interleukin-1β-converting enzyme 619
Internal energy 282
Internal equilibrium constants 640
International system of units 2
International unit of enzyme activity 457
Intervening sequences. *See* Intron
Intestinal mucosa
  alkaline phosphatase in 645
Intrinsic barrier 493, 691
Intrinsic factor in vitamin B_{12} uptake 869
  pernicious anemia 869
Introns 202
Inulin 170, 175
Inversion of configuration
  as criterion of mechanism 590, 591
  by glycosyltransferases 593
Involucrin 439
Iodacetamide 115
Iodide-starch complex 174*s*
Iodoacetate 115
Iodothyronine deiodinase 824
Ion(s)
  hydration of 50
  transport of **420 – 425**
    electrogenic 421
Ion channels 412 – 414, **420 – 425**
Ion exchange chromatography 103, 104
Ion exchange resins 104
Ion pairs 55
  hydrogen bonded 75
  in proteins 75
Ion pumps 422 – 425
  Ion
Ionic atmosphere of membranes 400
Ionic radius
  effect on metal binding 311
  for ions, table 310
Ionic strength 288, 297
Ionophores 414
Iron **837– 866**
  binding of O_2 850
  proximal histidine in heme 851
  in hemoglobin 850
  storage of 841– 843
  uptake by cells 837– 841
    non-transferrin pathway 840, 841
    transferrin pathway 840, 841
Iron regulatory factor 688
Iron regulatory protein 1 862
Iron-sulfur cluster(s) 514, 857, 857*s*
  3Fe-4S 859
  4Fe-4S 857
  reduction of 859
  in flavoproteins 861
  membrane proton pumps 860
  Mossbauer X-ray absorption spectra 860
  oxidation states 857
  properties of 859, 860
  radical generator 862
  reduction and oxidation 862
  single-electron transfer 861
Iron – sulfur enzymes
  functions of 860 – 862
Iron – sulfur proteins 686–688, **857– 862**
  electron paramagnetic resonance spectrum 860
Iron superoxide dismutase 866*s*
Iso-mechanisms 466
L-Isoaspartyl / D-aspartyl methyltransferase 594
Isoaspartyl group 84
Isobutyraldehyde, tautomerism 45
Isobutyryl-CoA mutase 872
Isocitrate 516*s*, 686*s*, 704*s*
  oxidation to 2-oxoglutarate 705
Isocitrate dehydrogenase
  phosphorylation of 545
  stereochemistry 768
Isocitrate lyase 704
Isocoumarins 623
Isodesmosine 436, 436*s*
Isoelectric focusing 108, 111
Isoelectric point 101, **106**
  of proteins 83
Isoenzymes (isozymes) 536, 538
Isoionic point **106**
Isolation of compounds 98 – 108
Isoleucine (Ile, I) 52*s*, 539
  biosynthesis 540
  branched fatty acids from 381
  configuration 43
Isologous interactions 337– 353
  in oligomers 349 – 353
  square 352*s*
Isomerases 498, 530, **692 – 697**
  aldose-ketose interconversions 692
  cis-trans, glutathione-dependent 690
  hydride shift mechanism 695
  3-oxosteroid 696, 697
  ring opening by 693
  triose phosphate 693 – 697
  xylose 695
Isomerization reactions 526, 527, 530, 690, 712
  of isopentenyl diphosphate 712
  stereochemistry 874
  vitamin B_{12}-dependent 873
    table 871
Isomers **41– 43**
Isomorphous replacement method 133
Isonicotinyl hydrazide 738
Isopentenyl diphosphate (prenyl pyrophosphate) 527, 689, 690*s*, 712*s*
Isopentenyladenosine 234*s*
Isopeptidases 525
Isopeptide linkage
  in proteins 80
  to ubiquitin 524
Isoprenoid alcohol phosphokinase 409
Isoprenoid compounds. *See also* Polyprenyl compounds
  biosynthesis 689, 700
Isoprenoid lipids **389**
Isoprenoid quinones, structures of 819
Isoprenylation. *See* Prenylation
Isopropylmalate 700*s*
Isopropylmalate isomerase 689
Isoproterenol 553, 553*s*
Isospartyl residues in proteins 594
Isotachophoresis 108
Isotope(s)
  properties of 110
  radioactive 110
  stable 110
Isotope effect(s)
  for alcohol dehydrogenases 771
  in displacement reactions
  kinetic 111, 486, 592, 593
Isotope exchange at equilibrium 467, 468
Isotopic labels (tracers) 110, 679
Isotopomer(s) 111
Isotopomer analysis 111

J

Joule 282
  definition of 2, 283
Jun 70, 241, 572
  definition of 577
jun 572, 576, 579

K

K$^+$ channel 413*s*
Karplus equation 139
Katal 457
KDEL receptor 554
KDO (ketodeoxyoctonate) 180*s*
Kelps 22
Kendrew, John C. 84
Keratan(s)
Keratan sulfate 175 – 177, 176*s*
  disulfide bridges 115
  N-linked 185
  terminal units 181
Keratin 70, 369
Keratinocytes 439
Keshan disease 822
Kethoxal 253
Ketimine 121, 744*s*. *See also* Schiff base
  from pyridoxal phosphate 742
  as electron acceptor 746, 747
α-Ketoacid. *See* 2-Oxoacid
Ketoamine 434*s*
Ketodeoxyoctonate. *See* KDO
Ketone(s), acidity of 46
Khorana, H. Gobind 84
Kidney cells, alkaline phosphatase in 645

- Kidney tubules 28
- Kilocalorie 282
- Kinases 638, **654 – 657**
  - cyclin-dependent 580
  - metal ions in 654
- Kinesin 370
- Kinetic(s)
  - high enzyme concentrations 461
  - of enzyme action **455 – 497**
  - of rapid reactions 468,469
- Kinetic constants 455
- Kinetic equations for complex mechanisms 466, 467
- Kinetic isotope effect (KIE) 76, 111, 486, **592**, **593**, 600, 789
- Kinetic parameters 459
- Kinetic theory of gases 462, 463
- Kinetisome 15
  - DNA of 15
- Klug, Aaron 84
- Knobs-into-holes bonding 71, 334
  - in coiled coil 71
- Knots
  - in peptide chains 74
  - in DNA 219
- Kojibiose 431
- Kornberg, Arthur 84
- Koshland
  - induced fit theory 349
  - model 352
- Krebs, Hans A. 83
- Kringles 367, 368*s*
- K systems 475
- Kuru 248
- Kynurenine 747*s*
  - conversion to alanine 746

L

- Lac repressor 240
- Laccase 887
- Lactacystin 620, 620*s*
- Lactate 506*s*
  - binding constants to metal ions 311
- Lactate dehydrogenase 342, 467, 768, 771
  - active site of 769
  - inhibition by pyruvate 780
  - isoenzymes of 538
- Lactate monooxygenase 799
- Lactic acid 508, 510
  - pK_a value of 293
- Lactobacillic acid 381*s*
- Lactoferrin 840, 840*s*
- Lactoperoxidase 853, 856
- Lactose (Lac) 167
  - α-Lactose 168*s*
  - sweetness of 168
  - β-Lactose 168
- Lactose (Lac) permease 417
- Lactosyl ceramide sulfate 389*s*
- *O*-Lactyl-GlcNAc. *See* Muramic acid
- Lactylthiamin 733*s*
- Lag phase of growth curve 470
- Lamellar (smectic) phase 392
- Laminin 71, 409, 437
- Lamins 369
- Landsteiner, blood groups 184
- Langmuir trough 394
- Langmuir-Adam film balance 394
- Langmuir-Blodgett layers 392, 394
- Large calorie, numerical value of 283
- Lateral phase separation 395
- Lathyrism 438
- Lavoisier 281
- Lead 31
- Leader peptidase 620
- Leadzyme 649
- Leaving group
  - definition of 589
  - effect on reaction rates 590
- Lecithin. *See* Phosphatidylcholine
- Lectin(s) 64*s*, 186, 186 – 188
  - domains in proteins 186
- Lederberg, Joshua 83
- Leghemoglobin 362
- Legionnaire's disease 7
- Leprosy 7
- Leucine (Leu, L) 52*s*
  - branched fatty acids from 381
  - biosynthesis of 527, 712
- L-Leucine 2,3-aminomutase 875
- Leucine aminopeptidase 627
- Leucine-zipper 69, 71, 241, 576
- Leucopterin 803, 804*s*
- Leucovorin 810
- Leukemias 573
- Leukocytes 26
  - adhesion 188
- Leupeptin 615, 622, 622*s*
- Lewis acid, Zn^{2+} 680
- Lewis blood group 184
- Lewis x antigen 188
- Lichen(s) 23
- Life, origin 9
- Ligase(s) 498
  - definition of 498
- Light absorption 122, 123. *See also* Absorption spectra
  - of DNA 209
- Lily, genome 12
- Lineweaver-Burk plot 460
- Link protein 181
- Linkage between genes 18
- Linoleic acid 381
- Lipase(s) **634 – 637**
  - hormone-sensitive 635
  - lipoprotein 635
- Lipid(s) **379 – 436**
  - bilayer 391*s*, 394
    - melting temperature 395
  - building blocks of **380 – 389**
  - caloric value 283
  - content in tissues 31
  - in membranes 392
    - physical properties of 394
  - neutral 382
- Lipid A 429, 430*s*
- Lipmann, Fritz A. 83, 720
- Lipoamide dehydrogenase
  - reaction mechanism, scheme 792
- Lipoarabinomannans 431
- Lipocortins *See* Annexin(s)
- Lipoic acid **795 – 802**
  - absorption maximum 795
  - arsenite, reaction with 596
  - chemical reactions 795
  - functions of 796 – 798, 798
  - proteins, attachment to 795*s*
  - reduction of 785
  - ring strain in 795
- Lipopolysaccharide 429
  - *O*-antigen of 180
  - structures of 430
- Lipoprotein(s) 58
  - bacterial 428
- Liposomes 392, 392 – 394
  - NMR spectra 396
- Liquid crystals 392 – 394
- Liquid crystalline phases 392
- Lithium salts, in treatment of manic-depressive illness 564
- Lithostatine 443
- Liverworts 29
- Lobry de Bruyn-Alberda van Ekenstein transformation 693
- Lock and key theory 478
- Log phase of growth 470
- Lon protease 628
- Loricin 439
- Lou Gehrig's disease 884
- Low barrier hydrogen bond 614
- Lowry method 102
- Luciferase 795
- Lungfish, genome 12
- Luteinizing hormone releasing factor (LRF) 54*s*
- Lutropin receptor 554
- Lyases 530
  - definition of 498
- Lymphocytes 25, 26
- Lymphomas 573
- Lynen, Feodor 84, 725
- Lyotropic series. *See* Hofmeister series
- Lysine (Lys, K) 700, 745
  - biosynthesis 53*s* , 540
  - pK_a value of 487
- D-Lysine 5,6-aminomutase 874
- D-α-Lysine mutase 871, 874
- L-β-Lysine mutase 871, 874
- Lysine tyrosylquinone (LTQ) 817, 817*s*
- Lysophosphatidylcholine 384
- Lysosome 10, 13, 98
- Lysozyme(s) **599 – 602**, 599*s*, 603
  - active site structure 600
  - catalytic side chain groups 599
  - kinetic isotope effect 600, 601
  - neighboring group assistance 601, 602
  - substrate distortion 601
  - transition state 601
  - HSQC NMR spectrum 146
  - titration curve 98
- Lysyl oxidase 434, 886
- α-Lytic protease 610
- Lyxose (Lyx) 163*s*

M

- M-phase of cell cycle 579
- Macrodipole of helix 70
- Macroglobulin(s) (IgM) 58, 629
  - α_2 630
- Macromolecular structures
  - self-assembly pathways **362 – 368**
- Macromolecules. *See also* Polymers
  - association of 325
  - binding by 327
  - packing of **332 – 348**
  - synthesis and turnover of **517– 527**
- Macrophages 26
- Macroscopic dissociation constants 305 – 307

Magnesium complex
    in kinase action 655
Magnesium ion(s) 31
    in adenylate system 302–304
    concentrations in tissues **299**
    effect on equilibria 296
    enzymes dependent upon 299
    intracellular concentration 299
Magnetic iron oxide 842
Magnetic resonance study
    of phosphorus atom displacement 639–642
Magnetite in organisms 842
Maize, genome 12
Major groove of DNA 201, 213
Malaria 19
L-Malate, 516*s*, 683*s*, 705*s* *See also* Malic acid
    cleavage to acetyl-CoA 703
Malate dehydrogenase 497, 768
Malathion 636
Maleylpyruvate 690*s*
Malic enzyme 705
Malolactic enzyme 777
Malonate decarboxylase 723
Maltoporin (LamB) 411
Maltose 81, 167*s*, 606
β-Maltose from β-amylase action 608
Mandelate dehydrogenase 784
Mandelate racemase 691
    active site 690
Manganese **887**, **888**
Manganese ion(s)
    effect on equilibria 297
    in enolase 685
Mannan(s) 170, 175
Mannitol 165, 191
Mannoprotein of yeast cell walls 185
Mannose (Man) 161*s*, 163*s*
Mannose-6-phosphate 185
Mannose-6-phosphate isomerase 693
MAP kinase cascade 576–578
    figure 579
    in insulin signaling 570
MAPK 578
    definition of 577
MAPKK, definition of 577
Marcus theory 771, 848
    of proton transfer 493
Marfan's syndrome 440
Markovnikov rule 777
Mass action ratio 535
Mass spectrometry
    electrospray 114
    MALDI 112–114
    in sequencing 119
Mass spectrum
    of Cu/Zn superoxide dismutase 114
    of myoglobin, apo 115
Mast cells 427
Mastigophora 19
Matrix of mitochondria 14
Matrix metalloproteases 627
McClintock, Barbara 84
Mechanocytes 25
Medusa form 23
Megakaryocytes 26
Meiosis 18
MEK 578, 579
Melanin 439, 887
Melanocyte(s) 439
    drawing 439
Melanophores 439
Melanosomes 439
Mellitin 414
Melting curve
    for DNA 256
    for a diacylglycerol 395
Membrane(s) 5, 390–410
    active transport 410
    anchors. *See* Anchors, membrane
    artificial 392–394, 393
    basement 10
    cation channels 421
    chemical compositions, table 393
    coated, 426*s*
    communication 427
    concentration gradient 206
    Davsen-Danielli model 393
    differential scanning calorimetry 394
    double, of gram-negative bacteria 5
    electrical properties 400
    fluid mosaic model 390
    fluidity 397
    freeze fractured 405
    inner side composition 401
    ionic atmosphere 400
    ionic gradients across 400
    lateral diffusion of lipids 399
    lateral phase separation 395
    lateral proton conduction 401
    lipid bilayer structure 390, 393
    lipid composition 392
    lipids, physical properties of 394
    NMR spectra 396
    outer side composition 401
    peripheral proteins 407
    permeability to neutral molecules 399
    pretransition temperature 394
    resting potential 400
    structure of **390–401**
    transport through **410–427**
    transporters 411–427
Membrane metabolism
    enzymes of 409
Membrane proteins **401–409**
    anchors 402
    glycoproteins 404
    integral 401
    positive-inside rule 401
Menadione 818
Menaquinones (vitamin K_2) 818, 821
Mengo virus 247
Menkes' disease 883
Mercaptans 624
Mercaptides 550
Mercaptoethanol 98
β-Mercaptoethylamine in coenzyme A 722*s*
7-Mercaptoheptanoylthreonine phosphate 814*s*
Mercapturic acid 550*s*
Mercuric reductase 787
Mercury 317
    methylation of 876
Mercury ions
    binding to proteins 125
    reduction of
Merrifield procedure, peptide synthesis 127, 128
Merrifield, R. Bruce 84
MES buffer 99
    pK_a value of 99
Mesosome 3
Mesozoa 23
Metabolic diseases. *See* Genetic diseases
Metabolic pathways **505**, **505–507**
    scheme 506
Metabolic regulation **535–581**
    control elements of 536
    sensitivity coefficient 537
Metabolism. *See also* Specific compounds
    activation 507, 508
    beta oxidation 511, 512
    control by covalent modification reactions
        table 543
    control by substrate cycles 567
    definition of 1, 2
    effects of insulin 568
    fermentation 508–511
    glycolysis 508–511
    introduction **505–527**
    oxidative phosphorylation 512–515
    regulation of. *See* Metabolic regulation
    study with isotopic tracers 110
Metabolite, activation 507, 508
Metal-binding proteins, chelate effect 310
Metal-binding sites 312
Metal complexes
    formation constants, table 311
    stability of, transition metal 311
Metal ion(s) 296
    binding of 307
    charge 311
    geometry of orbitals 311
    hydration of 311
    ionic radii 311
        table 310
    toxicity 33
Metalloenzymes 680
Metalloflavoproteins 794
Metalloproteases **625–627**
Metalloproteins 312
Metallothioneins **317**, 317*s*, 680
Metals, transition **837–895**
Metaphosphate ion 639
Metastasis 573
Metazoa 23–29
Methane formation 814
Methanofuran 814*s*
Methanogens 7, 813–815, 879–881
Methanol dehydrogenase 816*s*
Methanopterin 803, 804*s*
Methemerythrin 862
Methemoglobin 358
Methemoglobin reductase system 359
Methionine (Met, M) 53*s*, 539, 875
    biosynthesis of 540, 746
    sulfoxides 85
Methionine repressor 243
    operator complex 243*s*
Methionine sulfoxide 631*s*
Methionine synthase 813, 874*s*, 875
Method(s) **95–150**
    isolating compounds 98–108
    of peptide synthesis 127–129
    of carbohydrate analysis **188–192**
Methotrexate 574, 805
5-Methylarginine 881
3-Methylaspartate lyase 685
Methylation
    of arsenic 876
    of carbohydrates 189, 190
    of carboxyl groups 592
    of mercury 876
    of nucleic acids 591
    permethylation 189

  of protein side chains 79
  in regulation 543
  of uracil in tRNA 591
Methylcellulose 172
Methylcobalamin 867, 867*s*, 875
Methyl-coenzyme M reductase 879 – 881
Methyl corrinoids 876
β-Methylcrotonyl-CoA, carboxylation of 724
2-Methylcysteine 881
5-Methylcytosine 234
oxo-Methylene analog of transition state 626, 627*s*
α-Methyleneglutarate mutase 871
Methylene-THF reductase 813
Methylesterase 563
N-Methylglutamate synthase 789
2-Methylglutamine 881
Methylglyoxal 695, 695*s*
metabolism of 697
Methylglyoxal synthase 697
Methyl group(s) 802
  synthesis of 813
  transfer reactions of 875
N-Methylhistidine 881
4-Methylidine-imidazole-5-one (MIO) 757
S-Methylmalonyl-CoA
  formation from propionyl-CoA 725
    equilibrium constant 725
Methylmalonyl-CoA mutase 871
  mechanism 873
  stereochemistry 874
  three-dimensional structure 874
Methylotrophs 815
Methyltetrahydrofolic acid 875
Methyltransferases 526, 563, 591– 593
Mevalonic acid-5-diphosphate (pyrophosphate) 689, 690*s*
Micelles 390
Michael addition 681
Michaelis complex 751, 751*s*, 752
Michaelis constant 464
  definition of 459
Michaelis-Menten equation 459
Michaelis pH functions **96**, 296
Michel, Hartmut 84
Microbodies. *See* Peroxisomes
Microfilaments 369, 369,370
  behind desmosomes 29
Micropinocytosis 426
Microscopic binding constants 328, 329
Microscopic dissociation constants 305
  of pyridoxine 305
Microscopic reversibility, principle of 486
Microscopy
  atomic force 131
  chemical force 131
  confocal scanning optical microscopy 129
  electron 130
  near-field scanning optical 130
  negative contrast 130
  scanning tunneling 131
  STM image, glycogen 174
Microsome, definition of 14
Microtomes 130
Microtubules 15, 369, 370
  associated proteins 372
  growth of 373
  labile 334, 370
  organizing centers 372
  stable 370
Microvilli 10, 12
Miescher, Friedrich 199
Millipedes 24
Minor groove of DNA 201, 213
Mitchell, Peter 84
Mites 24
Mitochondria 14,15
  inner membrane 512
  matrix of 512
    definition of 14
  origin from bacteria 15
  of plant cell
    in micrograph 13
Mitochondrial DNA. *See* DNA, mitochondrial
Mitochondrial processing peptidase 625
Mitochondrion 10
Mitogen activated protein kinase (MAP kinase) cascade 576 – 578
  figure 579
Mitogenic activity 576
Mitomycin 225*s*
Mitosis 18, 371
Mobile defects in proteins 496
Modification of enzyme
  by group transfer 545 – 549
  by phosphorylation 541– 544
Modifiers of enzymes 473
Molal solution, definition of 287
Molecular activity (of enzyme) 457
Molecular biology, definition of 2
Molecular dynamics simulations of membranes 396
Molecular mass
  in daltons 4
  relative 4
Molecular size, as basis of separations 100
Molecular weight 4 *See also* Relative molecular mass
Molecules, properties of
  chiral 42
  conformations of **43 – 45**
  cyclic symmetry 332, 333
  packing of **41**
  structural principles **39 – 46**
Molluscs 24
Molybdenum **890 – 893**
  coenzyme forms 890 – 892
  nutritional need 893
  oxo-transfer reactions 892
Molybdenum-iron cofactor (FeMo-co) 892
Molybdoenzymes
  active site structures 892
  dimethylsulfoxide reductase family 890
  nitrogenase family 890
  sulfite oxidase family 890
  xanthine oxidase family 890
Molybdopterin 803, 804*s*, 825
  in formate dehydrogenases 892
Molybdopterin cytosine dinucleotide 891*s*
Monensin 414
  K^+ complex 415*s*
  synthesis 872
Monocytes 26, 188
Monod, Jacques 84
Monod-Wyman-Changeux (MWC) model 350, 476
  allosteric regulation 349
Monosaccharide(s) **161 – 167**
  characteristic reactions 167
  ^{13}C-NMR spectrum of 162
  reduction of 167
Morels 20
Morgan, Thomas H. 83
Mössbauer X-ray absorption spectra
  of iron-sulfur clusters 860
Mosses 29
Motion, in proteins 81
Mouse, genome 12
mRNA 5, 230, 257, 536
  splicing of 11
Mucins 181, 182
cis,cis-Muconate lactonizing enzyme 692
Mucopolysaccharidoses 659
Multicatalytic protease 339, 620. *See also* Proteasome
Multidrug resistance protein 417
Multilamellar vesicles (liposomes) 392
Multiple attack concept 606
Multisubstrate enzymes, kinetics of 464
Muramic acid (Mur) 165*s*
Murein 170, 428, 429*s*. *See also* Peptidoglycan
Musci 29
Muscle(s)
  cardiac 26
  smooth 26
  striated 26
  thin filaments 336
  electrolyte concentrations 206
Muscular dystrophy 822
  nutritional 822
Mutagens 222, 253, 254
Mutarotase 486, 526
Mutarotation 486
Mutases 637
  vitamin B_{12} dependent 653,654, 872
Mutation(s) 2
  addition 256
  autosomal recessive 16
  conservative 59
  deletion 256
  harmful dominant 16
  heteroduplex formation 256
  lethal 16
  rate of 16
myb 576
myc 572, 576, 577, 579
Mycelium 20
Mycobacteria 431
Mycobacterium tuberculosis 431
Mycolic acids 431
Mycoplasma(s) 2, 3
Mycoplasma genitalium 4
  genome 12
Mycorrhizae 20
Myelin 390, 399
Myeloperoxidase 856
Myoglobin 17, 68*s*, 70, **353 – 359**, 354*s*, 844
  bonding of dioxygen 851
  folding pattern 354
  mass spectrum 113
  mass spectrum of apo form 115
  NMR spectrum 148
  oxygenation curves 355
  packing of side chains 75
Myohemerythrin 71*s*
Myosin 369
  coiled coil structure of 71
  diffusion constant of 461
Myotonia congenita 420
Myriopoda 24
N-Myristoyl group 559
Myxococcus xanthis, genome 12
Myxomycetes 20

N

NAADP$^+$ 315, 779
Na$^+$, K$^+$-ATPase 423 – 425
    model 424
Nacrein 678
NAD$^+$ (NAD) 63*s*, 507*s*, 535, **765 – 771**, 767*s*, 778
    absorption spectrum of 768
    addition to 779
    containing ^{2}H 768
    determination of 779
NAD^{2}H 771
    oxidation by 298
NAD$^+$/ NADH ratio, redox potential 767
NADH (NADH + H$^+$; NADH$_2$) **765 – 771**
    absorption spectrum of 768
    modification in acid 780
NADH-X 780
NADH-methemoglobin reductase 826
NADH:ubiquinone oxidoreductase 788
    oxidation by ferricyanide 780
NADP$^+$ (NADP) 507, **765 – 771**, 767*s*, 779
    in catalase 853
    isolation of 767
NADPH (NADPH + H$^+$; NADPH$_2$) 507, **765 – 771**
NADPH/ NADP$^+$ ratio 776
Naphthalene, resonance energy 299
Nathans, Daniel 84
Natural killer cells 185
Near attack conformation 485
Near-field scanning microscope 130
Neighboring group assistance in displacement reactions 601, 602
Nekoflavin 783
Nematocyst 427
Nematodes 24, 25
Neocarzinostatin 218, 224*s*
Neopterin 803
Nephrogenic diabetes insipidus 412
Neprilysin 625
Nernst equation 400
Netropsin 225*s*
    binding to DNA 225
Neu (protein) 576
neu (oncogene) 572
Neuraminic acid anion (Neu) 165*s*
Neuraminidase 186
Neurofibromatosis 574
Neurofilaments 369
Neuroglia 25
Neuron(s) 25. *See also* Axon
    of squid 24
Neurospora crassa 20
    figure 21
Neurotransmitter(s) 29, 553
    receptors for 479
Neutron diffraction 137
Neutrophils 26, 188
Nick translation 257
Nickel **877– 882**
    dietary essential 877
    in plants 877
Nickeloplasmin 877
Nicotinamide 507, 721, 766, 767, 769*s*, 780. *See also* NAD$^+$, NADH; NADP$^+$, NADPH
    daily requirements 756, 769
Nicotinamide adenine dinucleotide. *See* NAD$^+$, NADH
Nicotinamide ring 770, 779
Nicotine 769
    oxidation to nicotinic acid 769
Nicotinic acid. 769*s* *See also* Nicotinamide
Nicotinic acid adenine dinucleotide (NAADP$^+$) 778
Nicotinic acid adenine dinucleotide phosphate (NAA 315
Nicotinic acid hydroxylase 825
Nidogen 437
Ninhydrin 120, 121*s*
Nitrate reductase(s)
    assimilatory 890
    dissimilatory 890
Nitric oxide, NO 553, 561
    binding to hemoglobin 358, 359
Nitrite reductase(s) 836*s*, 861
    copper in 884, 885
    cytochrome *cd*$_1$ type 885
Nitroform 126
Nitrogen mustards 574
Nitrogenase 890
Nitromethane, inhibition of ammonia-lyases 756
p-Nitrophenolate anion, as a leaving group 636
p-Nitrophenyl acetate 610, 610*s*
Nitrous acid, reaction with nucleic acids 253
Nitrous oxide reductase(s)
    copper in 884, 885
NMR spectra
    of ADP, ^{31}P 642
    of aspartate aminotransferase 149
    of carbohydrates 190 – 192
    of honey, ^{13}C 162
    of tRNA, ^{19}F 270
    of glycyl-L-histidylglycine 147
    of gonadatropin-releasing hormone 145
    of DNA, ^{1}H 269
    of lysozyme 146
    of myoglobin 147
    of muscle, ^{31}P 304
    of phospholipid phases, ^{31}P 397
    of phosphatidylcholine, ^{1}H 396
    of pyridoxal 5'-phosphate, ^{1}H 138
    pyruvate kinase equilibria 641
    of serine proteases, ^{1}H 613
    of tRNA 268
NMR spectroscopy **137– 149**, 494
    assignment of resonances 145
    band widths 138
    basic principles 137
    carbon-13 140
    chemical shift 138
        table 141
    correlation time, τ_c 142
    coupling constants 138, 139, 146
    deshielded 139
    deuterium 140
    distance constraints 145
    downfield 139
    DSS or Tier's salt 138
    exchangable protons 148
    fluorine-19 140
    Fourier transform 141
    free induction decay 141
    Heisenberg uncertainty principle 138
    HMQC (15N-multiple quantum coherence) 144
    HSQC (15N-multiple quantum coherence)
        lysozyme 146
    isotope-edited spectra 144, 145
    J coupling 139
    liposomes 396
    membranes 396
    multidimensional
    mixing time 143
        NOESY 142
    nitrogen-15 140
    NOE 140
    NOESY spectra 146
    nuclear Overhauser effect 140
    phosphorus-31 140, 295
        of adenylate system 303
        of muscle 304
        of phosphotransferases 640 – 643
    positional isotope exchange 641
    protein dynamics 145
    relaxation times 141
    residence time, τ_m 142
    ring current shifts 139
    saturation-transfer 641
    scalar coupling 139
    solid-state 149
    spin decoupling 139
    spin-lattice relaxation time, T$_1$ 142
    spin-spin interaction 139
    spin-spin relaxation time T$_2$ 142
    stable isotopes, use of 110, 111
    study of phosphotransferases 639
    tetramethylsilane (TMS) 138
    titrations 147
    transferred NOE 145
    two-dimensional 141
        correlation spectroscopy (COSY) 144
        TOCSY 144
    of various nuclei 141
    of vicinal protons 139
Nobel prizes, list 83
Nodes of Ranvier 29
Nojirimycin tetrazole 605, 605*s*
Non-coding (transcribing) strand in DNA 237
Noncompetitive inhibition 476, 477
Nonheme iron proteins. *See* Iron-sulfur and diiron proteins
Nonlinear equations 460
Nonmetallic ions, ionic radii, table 310
Nonproductive complexes 475
Norepinephrine (noradrenaline) 553, 553*s*
    in receptor 555*s*
Nuclear envelope 11
Nuclear magnetic resonance spectroscopy. *See* NMR spectroscopy
Nuclear Overhauser effect (NOE) 140. *See also* NMR spectroscopy
Nuclear pore 11
Nuclear relaxation
    by chromium ions 639
    by manganese ion 639
Nucleases 250
    micrococcal, active site 653
Nucleic acid(s) 199 – 213, 218
    abbreviated structures 204
    absorption of light 205
    acid – base chemistry 204, 205
    amount in cells 31
    bases, aromatic character of 204
        stacking of 209
    catenated 218
    cruciform structure in 229

  - dynamics 218
  - electrophilic reagents 253–255
  - homology of 256
  - hydrolysis of 249
  - interactions with proteins **239–244**
  - isolation of 249
  - knots in 218, 219
  - melting curves of 255
  - methods of study **249–270**
  - pK_a values of 204
  - separation of 249
  - tautomerism 204, 205
  - topology **218–230**
  - turnover of 527
- Nucleic acid bases
  - alkylation 253
  - aromatic character 207
  - chemical reactions of **207**
  - epoxides 254
  - halogenation 253
  - reactions 251–255
  - reactions with formaldehyde 253
  - reactions with kethoxal 253
  - reactions with nitrous acid 253
- Nucleoid 4
- Nucleolus 10, 11
- Nucleophilic catalysis 494. *See also* Covalent catalysis
- Nucleophilic character and bonding to metal ions 307
- Nucleophilic displacement reactions 526, **589–662**
  - carbocationic intermediates 598
  - cellulases 602
  - chitinase 599–602
  - glycogen phosphorylase 604, 605
  - β-glycosidases 602
  - glycosyltransferases 593–598
  - inversion 590, 591
  - kinetic isotope effects 592, 593
  - lysozymes 599–602
  - starch-hydrolyzing enzymes 605–608
  - transmethylation 591, 592
- Nucleophilic group, definition of 589
- Nucleophilic reactivity (nucleophilicity) 589
  - alpha effect 590
  - effect of polarizability on 590
  - of sulfur compounds 590
- Nucleoside(s)
  - conformations of 211, 212
  - names and abbreviations, table 203
  - torsion angles of 211
- Nucleoside diphosphate kinase 655
- Nucleoside Q (Queosine) 234*s*
- Nucleoside trialcohols, fluorographs 252
- Nucleoside triphosphates 302
- Nucleoside Y 234*s*
- Nucleosomes 239
  - DNA in 221
  - model 222
- Nucleotide(s) 199–213, 200*s*
  - names and abbreviations, table 203
  - chemistry of **203–213**
  - conformations of 211–213
  - modified 234
  - pairs 3
  - structures of 203–213
- Nucleotide binding domain 76
  - glyceraldehyde phosphate dehydrogenase 63*s*
- Nucleotide handles
  - of coenzymes 719
  - and specificity of metabolism 720
- Nucleotide kinases 655
- Nucleotide sequence
  - of aspartate aminotransferse 202
  - of DNA and RNA chains 200
  - of genes and genomes 12, 201
  - of mitochondrial DNA 201
  - of viruses 201
- Nucleotidyl transferases **657**
- Nucleotidylation 545
- Nucleus 10, 11
  - of plant cell, micrograph 13

O

- O-Antigen 180, 428–430
- *Obelia* 24
- Occlusion zones 29
- Octahedral geometry 311
- Octopus, intelligence of 24
- Oils, composition of 380
- Okadaic acid 545
- Old yellow enzyme 783
- Oleate hydratase 526, 688
- Oleic acid 381
- Olfactory responses 558
- Oligomer(s)
  - asymmetry 344–347
  - cubic symmetry 342–344
  - with dyad axes 337–348
  - heterologous tetramer 352
  - isologous tetramer 352
  - quasi-equivalence 344–347
- Oligomeric enzymes 475
- Oligonucleotides, 200–205
  - synthesis of 258
  - probe 119
- Oligosaccharide(s) **167–170**. *See also* Specific names
  - asparagine-linked (*N*-linked) 180–185
  - biological recognition of 161
  - N-linked 182–185, 183, 192*s*
  - release from glycoproteins 188
  - separation of 100
- Omega. *See* Oxidation
- OmpA protein 411, 428
- OmpC (osmoporin) 411
- OmpF 411, 412*s*
- Oncogenes 242, 544, **571–576**
  - in human tumors 571
  - table 572
- Oncogenic viruses 248
- One-start helix 334
- Oparin, I. V. 9
- Open systems 289
- Operator sequence of DNA
  - ^{1}H NMR spectrum 269
- Operons 240
- Opsin(s) 553
- Optical rotation 42
- Optimum rate for enzymes 469
- *d* Orbitals, participation in covalent bond formation 311
- Orcinol 251
- Ordered binding 464
- Ordered sequential mechanism 475
- Organelle(s) 11
  - separation 101
- Organic mechanisms, use of arrows 528
- Ornithine decarboxylase 342, 342*s*
- Ornithine mutases 871, 874
- *Oscillatoria* 22
- Osmate ester 393*s*
- Osmotic shock 417
- Osteoarthritis 438
- Osteoblasts 26, 441
  - alkaline phosphatase in 645
- Osteocalcin 821
- Osteoclasts 441
  - acid phosphatase in 645
- Osteocytes 441
  - alkaline phosphatase in 645
- Osteogenesis imperfecta 438
- Osteonectin 313
- Osteopetrosis 572, 678
- Osteopontin 409, 441, 572
- Osteoporosis 443, 619
- Ouabain 422
- Ovomucoids 183, 629
- Ovotransferrin 840
- Oxalate 699*s*
- Oxaloacetate 699*s*
- Oxaloacetate 506*s*, 515, 516*s*, 737*s*
  - acetylhydrolase 698
  - decarboxylases 527, 706
- 4-Oxalocrotonate decarboxylase 705
- 4-Oxalocrotonate tautomerase 697
- Oxalosuccinate 516*s*, 705
- Oxazoline 601*s*
- Oxidation
  - of biochemical substances, table 290
  - linked 705
  - by NAD$^+$, table 290
- β Oxidation pathway 506, 511
- Oxidation-reduction, by dehydration 697
- Oxidation states
  - of iron-sulfur clustsers 857
- Oxidation-reduction potentials. *See also* Reduction Potential
  - of pyridine nucleotide system 766
  - table 301
- Oxidation-reduction reactions
  - coenzymes of **765–827**
- Oxidative decarboxylation of α-oxoacids 511, 736, **796–802**
  - by hydrogen peroxide 799
  - mechanisms 799–801
  - with thiamin diphosphate 800
- Oxidative metabolism 535
- Oxidative phosphorylation 512
- Oxidizing power of redox couples 300
- Oxidoreductases 497
- α(2)-Oxoacid(s) (α-keto acids). *See also* specific acids
  - cleavage of 802
  - decarboxylation 730
  - oxidative decarboxylation **796–802**
- β-Oxoacid(s)
  - decarboxylation 705
- α(2)-Oxoacid dehydrogenases 348
  - sequence of reactions, diagram 798
- 3-Oxoacyl-CoA transferase 527
- α-Oxo alcohol, reaction with thiamin 732
- Oxocarbenium ion 598*s*, 604
- α(2)-Oxoglutarate (α-ketoglutarate) 516*s*, 737*s*
  - dehydrogenases 796
  - synthase 799
- 2-Oxoisovalerate 700*s*
- Oxonium ion. *See* Oxocarbenium ion
- 2-Oxopantoate 810

Oxophenarsine 597*s*
5-Oxoproline, 551*s*, 662*s*
Oxosteroid isomerase 526, 696
  molecular activity 696
Δ^5-3-Oxosteroid isomerase
  turnover number of 458
Oxyanion 613
Oxyanion hole 614, 615*s*
  formation of in trypsinogen 615
  of serine proteases 614
Oxygen (O_2)
  binding of, to heme proteins 353 – 358, 850 – 852
  to hemocyanin 885, 886
  paramagnetism of 850
  reactivity of 850
Oxygenases 863
  diiron 863
Oxyhemoglobin 851
Oxyperoxidase 854
Oxytocin 54*s*
Oysters, glycogen content 31

P

P Loop 425, 648, 658
Packing
  of macromolecules 332 – 348
  of molecules 41
Paget’s disease 443
Palindromes
  in DNA 238
  H- (mirror-repeats) 238
Pallidin 405
Palmitic acid 381
Palmitoyl group **559**
Pancreatic trypsin inhibitor 629
Pantetheine 722
Pantetheine 4'-phosphate 722
  in acyl carrier proteins 723
Pantoic acid 722, 723
  in coenzyme A 722
Pantothenic acid 721, 722, 722*s*
  nutritional requirement of 756
Papain 618
  active site 618
Paper chromatography 103
Papovaviruses 244
Parabactin 838, 838*s*
Parallel β structure 62
Paramagnetic ions, effects on NMR signal 640
Paramecium 17, 20
Paramylon 170, 174
Parathion 636*s*
Parathyroid hormone 314
Paratose 180*s*
Partition coefficient 410
Parvalbumin 312 – 316
Parvoviruses 244
Patterson map 135
Pauling, Linus C. 83, 84
Paxillin 406
p 53 cancer suppressor 574, 581
Pectate lyase 65, **686**
Pectate lyase C 686, 686*s*
Pectin(s) 15, 170, 177
Pellagra 721, 769
L-Penicillamine 739
Penicillin 20
Penicillin acylase 620
Penicillopepsin 621
Pentose phosphate pathway 700
  reductive 517
Pentoses 162
PEP. *See* Phosphoenolpyruvate
PEP carboxykinase. *See* Phosphoenolpyruvate carboxykinase
PEP carboxylase. *See* Phosphoenolpyruvate carboxylase
PEP carboxytransphosphorylase 706
PEP mutase 711
Pepsin 456, 621
  specificities 117
Pepsinogen 625
Pepstatin 623
Peptidase(s) 526, 609
  leader 620
  procollagen 433
  signal 620
  ripeptidyl 140
Peptide(s) **51– 59**
  β conformation of 60
  knotted chains 74
  torsion angles 60
Peptide deformylase 627
Peptide hormone(s). *See also* Chapter 30
  amidation of 522
Peptide linkage 51, 55*s*
  cis 56*s*
  dimensions 55
  resonance stabilization of 55
  trans 56
Peptide mapping 118
Peptide structure
  torsion angles of, table 61
Peptide synthesis **127– 129**
  solid-phase 127, 128
Peptide unit
  hydrogen bonds 56
  partial double bond character 55
  planarity of 55
  properties of **55**, **56**
  structure **55**, **56**
Peptidoglycan(s) 179, 428, 429*s*
Peptidyl transfer 234
Peptidyl transferases 637
Peptidyl-prolyl-(*cis-trans*) isomerases 488
Peptidylglycine α-hydroxylating monoxygenase 887
Performic acid 115
Perfusion chromatography 100, 101
Perinuclear space 11
Periodate oxidation
  of carbohydrates 190
Periodental disease 442
Periplasm 427, 428
Permeability coefficient 410
  for D-glucose 415
Permease 411
  for lactose 417
  periplasmic 417– 419
Pernicious anemia
  vitamin B_{12} treatment 868, 869
Peroxidase(s) 844, 845, 852 – 855
  ascorbate 852
  bromo 856
  catalytic cycles, scheme 855
  chloro 855
  compounds I, II 853 – 855
  cytochrome *c* 852
  glutathione 856
  halo 855, 889
  heme linkage 853
  horseradish 852, 856
  iron properties 853
  lacto 853, 856
  mammalian 853
  mechanisms of catalysis 853 – 855
  myelo 856
  structure 853
  thyroid 856
Peroxisomes (microbodies) 10, 14
Perutz, Max F. 84
PEST regions 527
pH 95 – 98
  of cytoplasm 549
  effects on enzymatic action 469
  effects on equilibria 293 – 297, 305
  effects on metal binding 310
  Michaelis pH functions 296
  within a cell 295
pH dependence
  of chymotrypsin 616
PH (plekstrin homology) domain 102, 367, 566, 571
Phaeophyta 22
Phage head 363
Phage lambda
  repressors 241
Phage tail 363
Phagocytosis 13
Phagosomes 426
Phantom atoms 42
Phase transition 284
Phase transitions
  of lipid bilayer 394
Phenol(s)
  acidity of 308
  pK_a value of 293, 309
  resonance energy 299
Phenolsulfonic acid
pK_a value of 99
Phenylacetic acid(s)
  acidity of 308
  pK_a value of 309
Phenylaceturic acid 513*s*
Phenylalanine (Phe, F) 52, 52*s*
  absorption spectrum of derivative 123
Phenylalanine ammonia-lyase **755**
Phenylephrine 563*s*
Phenylisothiocyanate 118
Phenylketonuria 16
Phenylpyruvate 689
Phenylpyruvate tautomerase 692
Phenylthiohydantoin 118
Pheromones 382
Phloem 30
Phorbol esters 564
Phosducin 558
Phosphagen 303
Phosphatase(s) 515, 526, 637, **645 – 647**
Phosphate anhydride 303
Phosphate, inorganic (P_i)
  in adenylate system 302, 304
Phosphatides 383, 384
  structures of 383
Phosphatidic acid 383*s*
Phosphatidyglycerol 383*s*, 384
Phosphatidylcholine 383, 383*s*, 391*s*
  in bilayer 391
  packing 391
Phosphatidylethanolamine 383, 383*s*

    bilayer, AFM image of 392
Phosphatidylinositol
    383*s*
Phosphatidylinositol (PtdIns) 563, 565, 566*s*
    in signalling **563 – 566**
Phosphatidylinositol 3-kinase 565
Phosphatidylinositol 4,5-bisphosphate 563
Phosphatidylserine 383*s*, **564**
    decarboxylation of 753
Phosphatidylserine decarboxylase 409, 755
3'-Phosphoadenosine 5'-phosphosulfate (PAPS) 659
Phosphoadenylation
    of protein side chains 79
Phosphodiesterase 556, 638
    for cGMP 558
Phosphoenolpyruvate (PEP) 510, 510*s*, 685*s*
    key metabolic intermediate 705, 706
    as stabilized enol 697
Phosphoenolpyruvate carboxykinase 549
Phosphoenolpyruvate carboxykinase (PEPCK) 706
Phosphoenolpyruvate carboxylase 527, 711
Phosphoenzyme 653, 703
Phosphofructokinase
    in regulation 535
Phosphofructokinase-2 541
Phosphoglucomutase 653
6-Phosphogluconate ion 164*s*
6-Phosphogluconate dehydratase 697
6-Phosphogluconate dehydrogenase 770
3-Phosphoglyceraldehyde dehydrogenase 770
2-Phosphoglycerate 510*s*, 685*s*
    dehydration 685
    dehydration to PEP 705
3-Phosphoglycerate 506*s*, 707*s*
    in chloroplasts 706
Phosphoglycerate kinase 654, 656
Phosphoglycerate mutase 653
Phospho (phosphoryl) group 292*s*
    chirality of 644
    as a handle 507
    transfer from ATP 660
    transfer of 292
    transfer of, by kinases 654
O-Phosphohomoserine 746
Phosphoketolase 736
Phosphokinases 526
Phospholipases 635
    A_1 566
    A_2 566, 635
    C **563**, 566
    D 566
Phospholipid(s) **379**, **380**, **382 – 387**, 392. *See also* Specific substances
    bilayer diagram 391
    head groups, functions of 396
    inverted hexagonal phase 397
    31P NMR 397
    non-bilayer structures 397
Phosphomannomutase 654
Phosphomutases 526
Phosphonamidate 626*s*
    as metalloprotease inhibitor 626
Phosphonate esters 626*s*
Phosphonates
    synthesis of 711
Phosphono group 711
Phosphonoethylamine 385*s*
Phosphonolipids 385
4'-Phosphopantetheine 723
4' - Phosphopantothenylcysteine
    decarboxylases 753, 755
Phosphophoryns 442
Phosphoporin (PhoE) 411, 412*s*
Phosphoric acid
    in lipids 380
    pK_a value of 99
Phosphorolysis, 593
Phosphorothioate
    phosphotransferases, study of 642
Phosphorus-32
    properties of 110
Phosphorylase a 541
Phosphorylase kinase 541, 544
Phosphorylation 541
    in bacteria 545
    cascade reactions 542
    oxidataive 303
    photosynthetic 303, *see also* Photosynthetic phosphorylation
    of protein side chains 79
    substrate level 775, 800
Phosphorylation, photosynthetic. *See* Photosynthetic phosphorylation
Phosphorylation reactions 303
Phosphorylation state ratio
    definition of 303
O-Phosphoserine 610*s*
Phosphoserine 545
Phosphothreonine 545
Phosphotransferase system
    bacterial 419, 420
Phosphotransferases 637
    in-line displacement 638
    P31 NMR of 642, 643
    pentacovalent intermediate 638
    stereochemistry 642, 643
Phosphotriesterase 646
    mechanism 646
Phosphotyrosine 545
Phosvitin 80
Photoautotrophic bacteria 8
Photofootprinting 266
Photorespiration 707
Photosynthesis 506, 517, 705
Photosynthetic bacteria 7
Photosynthetic phosphorylation 303, 517
Photosynthetic reaction centers 71
Phthaldialdehyde 120*s*
Phycoerythrin 22
Phycomycetes 20
Phylloquinones (vitamin K1) 818, 821
Phylogenetic approach
    to RNA structure 232
Phylogenetic tree 7, 8
Physarum polycephalum 20
Phytanyl group 385
Phytate. *See* Inositol hexakisphospate ($InsP_6$)
Phytochelins 318
Phytocyanin 883
Phytoglycolipids 388, 389*s*
Phytosphingosine 388, 388*s*
Picornaviruses 247
Pili 3, 6, 335, 336, 364. *See also* F pili
Pimeloyl-Coenzyme A
    conversion to biotin 719, 723
Pinacol 712*s*
Pinacolone 712*s*
Pinacol-pinacolone rearrangement 712
Ping-pong mechanism 465, 475
    rate equation for 466
    succinyl-CoA-acetoacetate CoA
        transferase 662
Pinocytosis 13
Piperazine, pK_a value of 99
Piperidine, pK_a value of 99
PIPES, pK_a value of 99
Pitch
    of a helix 68
    of B-DNA 213
Pituitary hormones 54
pK_a values **95 – 98**. *See also* Specific compounds
    of acids, table 293
    of amino acid side chains 55
    of buffers, table 99
    of buried groups 55
    of dicarboxylic acids, table 328
    effect on separation 104
    in hemoglobin 357
    of hydrogen-bonded protons 494
    in hemoglobin 357
    succinic acid 330
    of terminal – COOH 55
    of terminal – NH3$^+$ 55
    thiamin 330
Plague 7
Planarians 23
Plankton 22
Plant cell, electron micrograph 13
Plant cells and tissues **29**, **30**
    fiber 11
Plasma amine oxidase 886
Plasma antithrombin III 177
Plasma membrane 12, 379
Plasmalemma. *See* Plasma membrane
Plasmalogens 383*s*, 384
Plasmids 5, 248249
    ColE1 248
    drug resistance 248
Plasmin 634
Plasminogen 634
Plasmodesmata 10
    in micrograph 13
Plasmodium 19, 20
Plastids 14, 15
Plastochromanols 818
Plastocyanin 65*s*, 883
Plastoquinone(s) 818, 818*s*
Platelet(s) 26
    cyclic AMP in 557
Platelet activating factor **384**, 384*s*, **385**
Platelet-derived growth factor (PDGF) 544, 571
Platyhelminthes 23
Pleated sheet 62
    structure 62
P loop 425, 648
    consensus 559
    of Ras 559
PLP. *See* Pyridoxal phosphate (PLP)
PLP-dependent reactions. *See* Pyridoxal phosphate-dependent reactions
PNA, 227*s*
Poison(s). *See also* Insecticides, Antibiotics, Inhibitors
    cyanide 590
    hydroxylamine 590
Polar membrane 5
Polar molecules 50
    definition of 48
    hydration of 50
Polarizability 590

Polarization
  of water 47
Poliovirus(es) 247
Poly(ADP-ribose) synthase 545
Polyadenylation site
  in RNA 202
Polyamines 222
Polycarboxylate synthases, table 700
Polycarboxylic acid synthases **700 – 703**
Polyethylene
  helical structure of 45
Polyethyleneglycol (PEG) 102
Polyfluorethylene (Teflon)
  helical structure of 45
Polyglutamates 803
Polyglycine II 72, 73*s*
Polyhistidine "tag" 106
Poly-β-hydroxybutyric acid 3
Polymerase chain reaction (PCR) 260
  figure 261
Polymetaphosphate 3
Polymorphonuclear leukocytes 26
Polynucleotide(s) 199*s*, **226 – 228**. *See also* Nucleic acids, RNA, DNA
  hybridization 255
  hydrazine reactions 254
  interactions with ions 218
  melting of 255
  probes 255
  rotational movements of bases, figure 217
  triple-helical **226**
Polynucleotide kinase 258
Polynucleotide polymerases 638
Polyomavirus 247
Polypeptide(s) **56 – 59**. *See also* Proteins
  antibiotics 66
  chemical synthesis 85
  conformation of **59 – 61**, 78
  definition of 51
  torsion angles 59 – 62
Polypeptides. *See also* Peptides
Polyphosphates 302, 303
Polyprenyl compounds
  polymerization of 712
Polyproline 72, 73
Polyproteins 522, 628
Polysaccharide(s) **169 – 180**
  alternating, list of 170
  of bacterial surfaces 179, 180
  chains, conformation of 170
  list of, table 170
  main linkages, table 170
  structureal features of, table 170
Polysialic acid 187
Polyubiquitin 524
Porifera. *See* Sponges
Porins 65, 402, 411, 428
  FepA 839
  FhaA 839
Porphin(s) 843
Porphobilinogen 843*s*
Porphyrin(s) 843
  biosynthesis 742
  chelate effect 310
  tautomerism of 845
Posttranslational modification 402
  of proteins 521– 523
Potassium **206**
Potassium ion(s) 31
  activation of enzymes by 206
  binding to tetraplex DNA 206
  concentration in cells 206
  effect on equilibria 297
  replacement by proteins 206
  replacement by rubidium 206
  in tissues 299
Potentiometer 300
Poxvirus 247
PQQ *See* Pyrroloquinoline quinone
Preproteins 519 – 522. *See also* Proenzyme(s)
  peroxisomal 521
Precursor activation 540
Prenatal diagnosis 26
Prenyl diphosphate (pyrophosphate) 390*s*
Prenylation 402
  of proteins 559
Prenyltransferases 712
Prephenate 485, 690*s*
  conjugative elimination 689
Prepro-opiomelanocortin 522
Preproinsulin 519
Pressure jump method 468
Primary kinetic isotope effect
  for glucose-6-phosphate isomerase 592, 593, 693
Primary hydration sphere 311
Primary structure
  of nucleic acids 200 – 202
  of protein 56
Priming reaction in metabolism 507
Prion 248
Priority sequence 42
Procapsid 366
Processive mechanism
  of α-amylase 606
Prochiral centers 478, 768
  trigonal 480
Prochlorophytes 10
Procollagen 432, 432*s*
Procollagen peptidases 433
Product inhibition 464, 475
Proenzyme(s), 519, 609. *See also* Zymogen
  of histidine decarboxylase 754
Proflavine 223*s*
Progesterone 696
Programmed cell death. *See* Apoptosis
Progress curve for enzyme-catalyzed reaction 455
Proinsulin 519
Projection formula
  Fisher 42
  Newman 44
Prokaryotic cells, definition of 2
Proline (Pro, P) 52*s*
  in helices 69
  reductases 753, 755
Proline rings, torsion angles 62
Prontosil 473*s*
Proofreading
  by enzymes 482
1,2-Propanediol 874*s*
β Propeller 67, 560, 764
Propionic acid, pK_a value of 293
Propionic acid bacteria 725
PEP carboxytransphosphorylase 706
Propionyl-CoA, carboxylation of 724
Propranolol 553*s*
Prostacyclin 557
Prostaglandin(s) 565
Prostaglandin PGE$_2$ 381*s*
Prostanoids 381
Prostate gland, zinc content 680
Prosthetic groups 312, 719
Protamines 204, 239
Protease (proteinase) 609. *See also* Proteosome
  aspartic 621– 625
  ATP-dependent 627, 628
  in blood clotting 631– 634
  cysteine 618, 619
  in defense systems 628
  E. coli 628
  functions of 628
  metallo- 625 – 627
  multicatalytic 339, 525, 620
  pH dependence of 616
  serine. *See* Serine proteases
  specificities, table 117
  subsites of, nomenclature 617
  substrate specificities 616 – 618
  V-8, specificity 117
  stereoelectronic effect 616
Protease inhibitors **622**, 629 – 631
  alpha 1 629, 631
Proteasome 523
Protection mapping. *See* Footprinting
Protein(s)
  amino acid composition 59, 115 – 123
  anchors for 402
  aromatic clusters 75
  caloric value 283
  charged groups 101
  chemical modification **79 – 81**
  classification of residues 79
  compositions of specialized 59
  content
    from nitrogen analysis 31
    in tissues 31
  covalent structure 57
  crosslinking **79 – 81**
  crossover connections 74*s*
  degradation of 628
  disulfide crosslinkages 80
  domains 74
  engineering of 85
  families 76
  globular 59
  hydration of 50
  hydrogen bonding of side chains 75
  hydrogen bonds in 75
  hydrolysis 116 – 118
  idealized shapes 60
  induction of synthesis 538, 539
  intracellular degradation 523
  isoelectric point 83
  maturation 519
  methylation of 79
  molecular mass 59
  multienzyme complexes 348
  nucleotide-binding domain 76
  oxidation of 85
  oxygen carrying 850 – 853
  packing density of 75
  pH, effects of 83
  phosphoadenylation of 79
  phosphorylation of 79
  primary structure 56
  processing of 519
  repression of genes 538, 539
  salting in 101
  separation with ammonium sulfate 101
  selenium-containing, table 825
  side chain groups, reactions of **123 – 127**

  - solubility 101
  - solvent, effects of 83
  - sulfate esters of 79
  - thermostability 85
  - transport of 519–521
  - turnover 538, 539
  - water molecules in 76
- Protein(s), association
  - binding equilibria 350–352
  - dimerizing 350–352
  - interfaces 74
  - quaternary structure 332
  - regulatory subunits 348
  - subunits of 74, 332
  - tertiary structure 59
  - unsymmetrical dimerization 344
- Protein(s), conformational states
  - compact denatured state 82
  - conformational changes in 81
  - conformational substates 496
  - denaturation of 82
  - dynamic properties **81–85**
  - local unfolding 81
  - mobile defects in 496
  - packing defects 81
- Protein(s), folding patterns **76, 77**
  - α/α barrel 77
  - α/β barrel 77
  - architecture of folded **59–76**
  - folding intermediates 82
  - folding of 81–83, 518, 519
  - molten globule 82
  - motifs 367
  - β sandwich 64
  - secondary structure 63
  - seven-bladed β propeller 561
  - seven-helix 553–556
  - symmetry in 77
  - winged helix 243
- Protein C 634
- Protein coats of viruses 334
- Protein concentrations, estimation of 102
- Protein data bank 149
- Protein disulfide isomerases 83, 659, 787
- Protein kinase 541, 656, 659
  - cGMP-activated 544
- Protein kinase C 544, 564
- Protein kinase, cAMP-dependent active site 657
- Protein phosphatases 544, 646
- Protein S 634
- Protein sequenators 118
- Protein sequences from genes 119
- Protein synthesis 3, 538, 539
- Protein tyrosine kinases 544
- Protein-disulfide isomerase 83
- Protein-DNA interactions 266
- Proteinase. *See* Protease
- Proteoglycan(s) 181, 182. *See also* Glycosaminoglycans
  - aggregates of cartilage 182
  - molecular architecture 182
  - terminal unit of polysaccharides 181
- Proteolipid protein of myelin 401
- Proteome 116
- Proteosome 620, 627
  - 20S 339, 341*s*
  - 26S 340
  - electron micrograph 340
- Prothrombin 632, 821
- Prothrombinase 631
- Protista 18–22
- Protoheme(s) 844, 845*s*
- Proton(s)
  - conduction, lateral 401
  - dissociation, diffusion-controlled 493
  - intramolecular transfer 693
  - inventory technique 493, 618
- Proton transfer
  - coupled 493
  - Marcus theory 493
  - rates 492
  - ultra-fast 491
- Proto-oncogenes 571–576, 574
- Protoplast(s) 6
- Protozoa **18–20**
- Provirus 248
- Proximity factor in catalysis 495
- Pseudo first order reactions 457
- Pseudoazurin 884
- Pseudoknot in RNA 231
- Pseudomonad(s) 8, 688
- Pseudorotation 212, 638
- Pseudosubstrate of acetylcholinesterase 636
- Pseudouridine 234*s*
- Pseudo vitamin B$_{12}$ 867
- Psicose 164*s*
- Psoriasin 313
- Psoriasis 316, 440
- PTEN 566
- Pteridine 803
  - ring system 802–804
- Pteridophyta 29
- Pterin(s) 802
  - coenzymes **802–813**
  - structures 802–804
- Puffballs 20
- Pufferfish, genome 12
- Pullulan 170, 175
- Purine 199, 199*s*
  - catabolism of 810
- Purine bases
  - absorption spectra of 205
  - hydrogen-bonding possibilities 207
  - names of 203
  - pK_a values of 204, 205
- Purine hydroxylase 890
- Purple acid phosphatases 658, 862, 863
- Purple Cu$_A$ center 883, 884
- Putidaredoxin 859
  - EPR spectrum of 860
- Putrescine 363
- Pycnodysostosis 619
- Pyochelin 838*s*
- Pyranose rings 161, 162, 175. *See also* Sugars
- Pyrenoid 20
- Pyridine
  - binding constants to metal ions 311
  - pK_a value of 293
  - resonance energy 299
- Pyridine nucleotide(s) 765
  - chemistry 779, 780
  - coenzymes **767–780**
- 6-Pyridone(s) 780*s*
- Pyridoxal 738
- Pyridoxal phosphate (PLP) 121, 679, **737–753**, 740*s*. *See also* PLP-dependent reactions
  - absorption spectrum 749
  - as active site label 124
  - amino acid reactions 740
  - in aspartate aminotransferase 57
  - as coenzyme, table 743
  - dependent enzymes, cyclization 741
  - enzymes, atomic structures 750
  - in glycogen phosphorylase 604
  - mechanism of action 740, 741, 751–753
  - NMR spectrum of 138
  - optical properties 749, 750
  - Schiff bases, reactions of 699, 705, 742
  - transimination 741
- Pyridoxal phosphate-dependent enzymes, list 743
- Pyridoxal phosphate-dependent reactions 741–747
  - beta replacement 742–744
  - cyclization 741
  - decarboxylation 745
  - elimination 742
  - ketimine intermediate 745, 746
  - nonenzymatic models 737
  - racemization 741
  - side chain cleavage 745
  - stereochemistry of 749
  - transamination 741
- Pyridoxamine 738
  - orange product with ninhydrin 121
- Pyridoxamine phosphate (PMP) 741, 744*s*, 747
- Pyridoxamine:pyruvate aminotransferase, stereochemistry of 748
- Pyridoxine (vitamin B$_6$) 305*s*, 738
  - microscopic dissociation constants 305
  - pK_a values of 305
  - tautomerism 45
- ε-Pyridoxyllysine 750*s*
- Pyrimethamine 805*s*
- Pyrimidine 199*s*
- Pyrimidine bases
  - absorption spectra of 205
  - hydrogen-bonding possibilities 207
  - names of 203
  - pK_a values of 204, 205
- Pyrithiamin 731, 736
- Pyroglutamate aminopeptidase 619
- Pyroglutamyl group as N-terminus
  - in proteins 54, 522
- Pyrophosphatases 515, 646
- Pyrophosphate (PP$_i$) 508, 515
  - in adenylate system 302–304
- Pyrophospho group
  - transfer from ATP 660
- Pyrophosphoric acid, pK_a value of 99, 293
- Pyrrole, resonance energy 299
- Pyrroloquinoline quinone (PQQ) 764, 815
- Pyrrophyta 20
- Pyruvaldehyde. *See* Methylglyoxal
- Pyruvate 506*s*, 510
  - anaerobic cleavage 800
  - decarboxylation, reversible 799
- Pyruvate carboxylase 711, 724
  - control mechanisms 729
- Pyruvate decarboxylase 718, 734
- Pyruvate dehydrogenase 511, 796–798, 797*s*
  - effect of insulin 571
- Pyruvate:ferredoxin oxidoreductase 799
- Pyruvate formate-lyase 875
  - reaction 800, 801
- Pyruvate kinase 656, 706
  - equilibria by NMR 641
- Pyruvate oxidase 799
  - active site 734
- Pyruvic acid, pK_a value of 293
- Pyruvoyl groups in enzymes, 753–755
  - table 753

Q

Quanine quartet 208*s*
Quantum mechanical tunneling 494, 771, 848
Quasi-equivalence
  in oligomers 344
  in virus coats 346
Queosine (nucleoside Q) 234*s*
Quinone(s) 815–822
  as electron carriers 819
Quinonoid-carbanionic intermediate 741, 744*s*
  from pyridoxal phosphate 743, 744*s*
Quinoproteins 815–818
  copper 815

R

Rab protein 559
Rabies 247
Racemase(s) 284
  amino acid 692
  aspartate 741
Racemization 741
  of amino acids 284
  of lens proteins 85
  PLP-dependent enzymes 741
Rac protein 559
Radicals, free. *See* Free radicals
Radioautogram (radioautograph, autoradiograph) 111
Radioautography 110
Radioimmunoassays 110
Radioisotope 258
Radiolaria 18
Radius
  collision 462
  covalent **40, 41**
  van der Waals 40
Raf 578, 579
Raffinose 169
Ramachandran plot 60, 61
Random order of binding 464
Random coil conformation 69
Rapamycin 488
Rapid equilibrium assumption 467
Rapid photometric methods 468
Ras 577–579
  human c-H ras 560*s*
  P loop 559
Ras oncogenes (*ras*) 572, 576
Ras protein superfamily 558
Rat liver, composition of 31
Rates
  of approach to equilibrium 458
  of displacement reactions **589, 590**
  of enzymatic reactions 455–497. *See also* Kinetics
  of substrate binding 463
Rate constants 455
  apparent first order 458
  bimolecular 458, 462
  prediction by Hammett equation 309
Rate equation for enzymes 455
Reaction coordinate 482
Reaction rates, diffusion controlled limit 462
Reaction types in metabolism, 530
  table of **526, 527**
Reading frames in nucleic acids **236, 237**
Rearrangement reactions 527, 530, 712
  recA protein 219
Receptor(s) 1, 553–563
  acetylcholine 422
  alpha adrenergic. *See* α-Adrenergic receptors
  autophosphorylation of 562
  bacterial, for aspartate 561
  beta adrenergic. *See* β-Adrenergic receptors
  on cell surfaces 479
  down regulation of **571**
  for hormones **553–563**
  human adrenergic 555
  for surface proteins 407
  seven-helix structures 555
Receptor theory 479
Reciprocal plots
  for kinetics of enzymes 460, 465, 472
Recognition domains in proteins 367
Rectus configuration 42
Red blood cell. *See* Erythrocytes
Redox potential
  within cells 767
  of flavoproteins 767
Reduced flavins, reactions with oxygen 794
Reducing equivalents from citric acid cycle 515
Reducing powers of redox couples 300
Reduction potential(s)
  definition of 300
  table, 301
Reference electrode 300
Refinement of X-ray structure 136
Refolding of proteins 82
Regulation
  of activity of enzymes **539–553**
  of aspartate carbamoyltransferase 540
  control elements in 535–537
  of enzymatic activity **535–581**
  pacemaker enzymes 535–537
Regulators of G-protein signaling (RGS) 559
Regulatory cascade(s) 566, 567
  monocyclic (scheme) 567
Regulatory subunits 348, 540, 541
Relative molecular mass (M_r)
  determination of **108–115**
  estimation by gel filtration 109
  estimation by PAGE 109
Relaxation methods 468
Relaxation time
  approach to equilibrium 458, 468
  first order reactions 457
Renin 621
Reoviruses 248. *See also* front cover
Repair systems 16
Replication cycle. *See* Cell cycle
Repression of enzyme synthesis 536, 538, 539
Repressor(s) 76, 239
  of gene expression 539
Resolvase 229
  Tn3 219
Resonance **45, 46**, 46
  energy values, table 299
  in esters and amides 608
Respiration 300
Respiratory chain 783
Respiratory distress syndrome 386
Resting potential of cell membrane 400
Restriction endonuclease(s) 249, 653
Restriction maps 260
Resurrection plant 168
Retention signal 521
Retinoblastoma protein (Rb) 574
Retinol-binding protein 58
Retrieval signal KDEL 521
Retroviruses 248
Reverse transcriptase 248, 257, 657
Reversed phase columns 103
Reversible chemical reactions 284
  kinetics of 458
Rhabdoviruses 247
Rhamnogalacturonan 177
Rhamnose (Rha) 165*s*, 180
Rheumatoid arthritis 627
Rhinovirus 247
  icosahedral 344*s*
Rhizopoda 18
Rho protein 558
Rhodophyta 22
Ribbon drawing 64, 65, 71, 240, 243, 336, 338, 343, 347, 348, 372, 409, 412, 413, 419, 488, 560, 580, 599, 606, 614, 633, 708, 786, 807, 840, 841
Ribisco activase 707
Ribitol, conformation 44
Ribitolteichoic acids 431*s*
Riboflavin 721, 766, 783*s*
  absorption spectrum 781
  fluorescence 783
  light sensitivity 783
  nutritional requirement 756, 783
  properties 783
Riboflavin 5'-phosphate (flavin mononucleotide, FMN) 513, 765, 780, 781*s*
Ribonuclease 264, **647, 648**
  fungal 648
  pancreatic, diffusion constant of 461
  refolding of 82
Ribonuclease A
  mechanism of catalysis 647
  NMR spectra of 647
  pK_a values of 647
  reaction mechanism 648
  sequence of 647
  X-ray structural studies 647
  in-line mechanisms 647
Ribonuclease P 649–652
Ribonuclease S 647
Ribonucleoprotein domain 244
Ribonucleotide reductase 863–865
  active site 865
  cobalamin dependent 871
  enzyme-activated inhibitors 864
Ribose 1, 5-bisphosphate
  in brain 545
Ribose 5-phosphate isomerase 693
D-Ribose (Rib) 162, 163*s*, 200*s*
Ribosome(s) 3, 5, 11, 233–235
  structure of 233, 234
Ribosylthymidine (Thd) 203
Ribosylthymidine 5'-phosphate (Thd-5'P) 203
Ribothymidine 234
  of tRNA 231
Ribothymidylic acid residue 813
Ribozyme(s) 649–652, 239
  artificial 652
  hairpin 649, 651*s*
  hammerhead 649, 651
  leadzyme, 651*s*
Ribulose 164*s*
Ribulose bisphosphate 707*s*
Ribulose bisphosphate carboxylase (Rubisco)

  549, **706 – 710**, 708*s*
    carbamoylation of 707
    mechanism, scheme 709
    spinach, active site 708
Rice, genome 12
Ricin 186
Ricinoleic acid 381, 381*s*
Rickets 421, 721
    X-linked hypophosphatemic 625
Rickettsias 7
Rieske iron-sulfur protein 860
Ring conformations 332 – 337
    boat 598
    half-chair 598
Ring current 139
RNA 5, **230 – 236**
    antisense 237
    aptamers 235
    base composition of 239
    binding proteins 243
    content in tissues 31
    density gradient centrifugation 101
    electrophoresis 249
    extraction of 249
    forms of, table 232
    Gibbs energy of formation 210
    hydrolysis of 249, 251
    loops and turns 210*s*, 230
    pseudoknot 230
    figure 231
    self-splicing 649
    sequence determination 260 – 266
RNA polymerase 657
    zinc in 680
RNase T1 648
RNAworld 235
Rocky Mountain spotted fever 7
Rods (visual receptor cells) 390
Root hairs, dimensions of 30
Roseoflavin 788, 789*s*
Rossmann fold. *See* Nucleotide-binding domain
Rotamases 488
Rotary diffusion constant 463
Rotation of molecules 462, 463
Rotational barrier 44
Rotifers 24, 25
Roundworms (Nematoda) 24
R*p* (phosphorylation state ratio) 303
rRNA (ribosomal RNA) 5, 230, 233*s*
RS notation
    for configuration **42**
    for trigonal centers 480
Rubella 247
Rubidium 31
Rubisco *See* Ribulose bisphosphate carboxylase
Rubredoxin(s) 858[s], 859
    Fe-S cluster 857*s*
Ruhemann's purple 121*s*
Rusticyanin 883
Rusts 20
Ryanodine receptor 314, 422, 564

S

S-phase of cell cycle 579
Saccharomyces cerevisiae
    figure 20
    genome 12, 20
Salmonella 7
    PEP carboxylase in 711
    serotypes 180
Salt bridges 47
    of hemoglobin 357
Salting in/out of proteins 101
Salvarsan 597*s*
SAM. *See* *S*-adenosylmethionine (Adomet)
Samuelsson, Bergt I. 84
β Sandwich 64
Sanger, Frederick 83, 84, 118
Sarcodina 18
Sarcomas 573
Satellite tobacco necrosis virus 247
Saturation in binding 326
Saturation curve(s) 326 – 331, 458 – 461, 475 – 477
    hyperbolic 326
    sigmoidal 326
Saturation fraction
    definition of 326
    for cooperative processes 331
Scanning mutagenesis 402
Scanning tunneling microscope 131
Scatchard plot 327
Scenedesmus 22
Schiff base(s) 121, 752
    in aminotransferases 750
    of pyridoxal phosphate 742
    of pyridoxal phosphate, 740*s*
    reduction of
        by borohydride 124, 679
Schistosoma 24
Schistosomes
    protease inhibitors 622
Schistosomiasis 24, 222
Schizophyllan 170, 174
Schwann cells 390
Sclerenchyma 30
Scorpions 24
Scrapie 248
Screw-sense isomer of ATP complexes 643
Scurvy 721
SDS electrophoresis 111
SDS-PAGE 106
    of erythrocyte ghosts, figure 404
SDS-polyacrylamide gel electrophoresis. *See* SDS-PAGE
Sea cucumbers 25
Sea squirts. *See* tunicates
Sea urchin 25
    genome 12
Second Law of Thermodynamics 284285
Second messengers 553
    calcium ions 555
    cyclic ADP-ribose 555
    cyclic AMP 555
    cyclic GMP 555
    diadenosine 5' tetraphosphate 555
    guanosine 5' diphosphate, 3-diphosphate (ppGpp) 555
    inositol-1,4,5 trisphosphate 555
    for insulin 569 – 571
    Mn^{2+} 555
    table of 555
Second order reactions 458
Secondary kinetic isotope effect 592, 600
    on fumarate hydratase 684
Secondary plots for kinetics of multisubstrate enzymes 465
Secondary structure 63
    of a protein 63
Secretion gene (Se) 185
Secretion granules 10, 14
Sedimentation constant 100, 109
    of nucleic acids 112
    of proteins 112
Sedimentation equilibrium 108
Sedimentation velocity 108
Sedoheptulose 164*s*
Seeds 30
S-Selanylcysteine 893
Selectins 187, 188
Selenium
    in enzymes 822 – 827
    metabolism of 827
    in proteins, table 825
    proteins, incorporation into 827
Selenocysteine 237, 823, 824
    in hydrogenases 878
Selenocysteine lyase 746
Selenoprotein P 824
Selenoprotein W 824
Self-assembly of macromolecules 332
Seliwanoff's reagent 167
Semiquinone radicals of flavins 792, 794
Sensitivity coefficient 537
Sensor kinase/response regulator systems 545
Separations of compounds 98 – 108
    by chromatography 100 – 106
    by partition 102 – 103
Sephadex, gel networks of 175
Sepiapterin 803, 804*s*
Septate desmosomes 29
Septum 5
Sequence(s)
    of amino acids 57
    of nucleic acids
Sequencing primers 260
L-Serine (Ser, S) 52, 53*s*, 808
    as nucleophile 609, 610
    pK_a value of 487
Serine carboxypeptidases 610
Serine dehydratases 688, 742
Serine hydroxymethyltransferase 745
    quinonoid intermediate 750
Serine proteases **609 – 618**
    acyl-enzyme intermediates 610
    carboxypeptidase 609, 625
    catalytic cycle 613
    catalytic triad 614
    chymotrypsin 66, 457, 477, 609
    oxyanion hole 614, 615
    papain 618
    pepsin 456, 621
    specificities 117
    thrombin 632
    trypsin 66, 116, 457, 609
Serine *O*-sulfate 739
Serotonin(5-hydroxytryptamine) 744
Serpins 629
Serum albumin 58
    helices in 70
Serum response factor 578
Sex factors 249
SH2 and SH3 domains 367, 580
Sharp, Phillip A. 84
Shc 579
    as adaptor protein 570, 577
    definition of 577
β Sheet, properties of 64
Shells 440, 441

- *Shigella* 7
- Shikimate 3-phosphate 687*s*
- SH2 / phosphopeptide complex 368*s*
- Sialic acid 165*s*, 183, 546
- Sialidase 186
- Sickle cell disease 16, 359, 360
- Side chains packing 75
- Siderophilins 840
- Siderophore(s) 838, 838*s*
  - chelate effect 310
  - receptors 839
- Sieve plates 30
- Signal peptidase 522, 620
- Signal recognition particle (SRP) 520
- Signal sequences 519
- Signaling 1
- Silica gel, use in separation 103
- Silica skeleton
  - of radiolarian 178
- Silicon
  - in biological cross-linking 178
  - biological functions 178
- Silicon dioxide in "shells" of diatoms 21
- Silk 67
  - β structure in 63
  - fibroin, composition of 59
- Simian virus 40 (SV40) 244
- Single carbon compounds
  - in metabolism 808 – 813
  - table, according to oxidation state 808
- Singlet oxygen 856
- Sinister configuration 42
- Siroheme 794, 845, 845*s*, 861
- *sis* 576
- *v-sis* 571, 572
- Sizes of atoms, table **41**
- Skin **439**, **440**
- S-layers of bacteria 431
- Slime molds. *See Dictyostelium discoideum*
- Smith degradation of carbohydrates 190
- Smoluchowski equation 461, 462
- SNAPS 521
- SNARES 521
- Snell, Esmond 737
- Sodium **206**
- Sodium arsanilate 597*s*
- Sodium borohydride
  - action on aldoses and ketoses 167
  - radioactive 679
- Sodium borotritide 190
- Sodium channels 421
- Sodium dithionite 767
- Sodium dodecyl sulfate 82*s*
- Sodium ion 31
  - concentration in cells 206
- Sodium-potassium pump 423
- Solubility
  - of proteins 101
  - separation by 101
- Sonication (sonic oscillation) 98
- Sorbitol 164, 196, 774
- Sorbose 164*s*
- Sos 579
  - definition of 577
  - as exchange protein 578
- Southern blots 260
- Soybean lectin 186
- Space groups 133, 332
- Specific activity of an enzyme 457
- Specific hydroxyl ion catalysis 487
- Specificity of enzymatic action 478 – 482
- Spectrin 313, 405
- Sperm receptors 187
- Spermidine 204, 363
  - interaction with DNA 218
- Spermine
  - interaction with DNA 218
- Spherocytosis 405
- Spheroplast 6
- Sphinganine 388*s*
- Sphingoglycolipids 388
- Sphingolipids **388**, **389**
  - structures of 389
- Sphingomyelins 389*s*
- Sphingophospholipids 383, 388
- Sphingosine (sphingenine) 388*s*
  - biosynthesis of 745
- Sphingosine 1-phosphate 578
- Sphingosylphosphocholine 422
- Spicules 23
- Spiders 24, 67
- Spin decoupling. *See* NMR spectroscopy
- Spin labels 398, 399
- Spin-spin interaction. *See* NMR spectroscopy
- Spinach (*Spinacia oleracea*)
  - composition of 31
  - ribulose bisphosphate carboxylase in 706
- Spin-orbit coupling 398
- β Spiral 74
- Spirillum, definition of 6
- Spirochetes 7
- *Spirogyra* 21
- Sponges 23
  - species-specific reaggregation 187
- Spongin 23
- Spore(s), bacterial 8
- Sporophyte form of fern 18
- Sporozoa 19
- Squid 24
- *v-src*, *src* 572
- Src, definition of 577
- SRP receptor 520
- Stachyose 169
- Standard electrode potential 300
- Standard states 286
- Starch 172, 173*s*
- Starch granules 10
  - electron micrographs 172
  - layered structure 173
- Starch-hydrolyzing enzymes 605 – 608
- Starch-iodine complex 174*s*
  - helical structure 173
- Starfish 25
- State, thermodynamic 281
- Statherin 442
- Statine (Sta) 623*s*
- Stationary phase of growth curve 470
- Statistical effects
  - in binding 328, 329
- Steady state 1, 289
  - approximation 459
- Stearic acid 380*s*, 381
- Stearoyl ACP desaturase 863, 863*s*
- Stellacyanin 883
- Stem, section of, figure 31
- Stereochemical numbering system 480
- Stereochemical studies
  - using chiral acetates 704
- Stereochemistry
  - of chorismate formation 689
  - of enzymatic reactions 111
  - of phosphotransferases **642**, **643**
  - of pyridoxal phosphate-requiring enzymes 747
- Stereocilia 370
- Stereoelectronic control 598
- Stereoelectronic effect
  - in enzymatic catalysis 495
  - in proteases 616
- Stereospecificity
  - of dehydrogenases 768 – 775
  - prochiral centers 478
- Steroid hormones 536
- Sterol(s) **389**
- Stigma 20
- STM image of glycogen 174
- Stoichiometric dissociation constant 305
- Stokes-Einstein equation 461
- Stokes radius 109
- Stomach ulcer 187
- Stomata 30
- Stop codons 236
- Stopped flow technique 468
- *Stratum corneum* 439
- Strecker degradation 121
- Strength of binding. *See also* Formation constants, Dissociation constants, Gibbs
  - in metal complexes 307
- Streptavidin 728
- Stromelysins 627
- Strontium 19
- Structural domains, table 367
- β Structure 78
  - hydrogen bonds in 61
  - of proteins **61– 68**
  - in small peptides 66
  - twist in 63, 64
- Stuart factor 632
- Styrene, resonance energy 299
- Substituent constants
  - in Hammett equation 308
  - table 308
- Substitution reactions. *See* Displacement reactions
- Substrate
  - definition of 332
- Substrate binding
  - rate of 463
  - unequal 349, 350
- Substrate cycles 567
- Substrate saturation
  - of an enzyme 457
- Substrate specificity
  - of chymotrpsin 616
  - of serine proteases 616
  - of trypsin 617
- Substrate surrogate 463
- Subtilisin 610
  - active site 612
- Subunits in protein structures 74
- Succinate 478*s*
  - binding constants to metal ions 311
- Succinate dehydrogenase (succinic dehydrogenase) 478, 766, 785
- Succinic acid. *See also* Succinate
  - pK_a values of 99, 330
- Succinimide
  - racemization 85
- Succinimidyl group 84
- Succinyl-CoA– acetoacetate CoA transferase
  - ping-pong mechanism 662
- Succinyl-CoA synthetase 661
- Succinyl-CoA 516*s*

- *O*-Succinylhomoserine 746*s*
  - conversion to cystathionine 746
- Sucrose 167, 168*s*
- Sucrose-fructose exchange 595
- Sucrose phosphorylase 595
- Sugar(s) **161–167**. *See also* Monosaccharides, Oligosaccharides
  - configurations of 163
  - covalent hydrate **162**
  - in lipids 380
- Sugar isomerases 526
- Sugar phosphates
  - interconversions 508
- Sugar transporters
  - fructose 416
  - glucose 416
  - SGLT1-Na$^+$ dependent 417
- Sulfanilamide 473*s*, 720*s*
  - acetylation of 720
- Sulfatase(s) 526
- Sulfate esters
  - in proteins 79
- Sulfatides 388
- Sulfite oxidase 847
  - family 890
- Sulfite reductase(s) 794, 861, 861*s*
- Sulfoquinovose 165*s*
- Sulfogalactosylglycerolipid 387*s*
- Sulfometuron methyl 735*s*
- Sulfonamides 473, 473*s*
- Sulfonolipids 388
- Sulfonylureas 421
  - receptor 421
- Sulfotransferases 526, 659
- Sulfur displacements **659**
- Sulfur-35
  - properties of 110
- Summer, James B. 83
- Supercoiling
  - energy of 221
- Supercritical fluid chromatography 103
- Superhelix 70
  - density of DNA 221
- Superhelix density
  - effect of intercalation 223
- Superoxide anion radical 794, 826, 850, 851
- Superoxide dismutase(s) 866, 884
  - copper, zinc 884
  - β-cylinder in 78
  - mass spectrum 113, 114
  - zinc in 680
- Superoxide radical
  - reaction with FADH$_2$ 529
- Surface plasmon resonance 122
- Surfactant system of the lungs 386
- Surfactin 54*s*
- Sutherland, Earl W., Jr. 84
- Suzukacillin 414
- Svedberg equation 109
- Svedberg units 100
- Svedburg, Theodor 83
- Symbiotic relationship 17
- Symmetry 133, 134, 338 – 348
  - cubic 342 – 344
  - dihedral 338
  - space groups 133
- *Syn* conformation of nucleosides 211, 212
- Syn transfer of protons 696
- Synapse(s) 29
  - electrotonic 29
- Synchrotron radiation 137
  - as X-ray source 134
- Synemin 370
- Synthetic probe
  - for nucleic acids 257
- System, thermodynamic, definition of 281

T

- Taft equation 309
  - table of constants 308
- Tagatose, 164*s*
- Tailspike protein
  - of bacteriophage 66*s*
- Talin 406
- Talose (Tal) 163*s*
- Tamoxifen 575
- Tapetum 680
  - zinc in 680
- Tapeworms 24
- *Taq* polymerase 260
- Tartaric acid 43*s*
  - configuration of 43
- Tautomerases 692
- Tautomeric catalysis 490, 491
- Tautomeric ratios 45, 305
- Tautomerism **45**, **46**
  - in amides 45
  - in base pairing 211
  - of cysteine 306
  - equilibrium constant 45
  - in imidazole group 46
  - in 1-methyluracil 307
  - of nucleic acid bases 211
  - of nucleic acids 204, 205, 211
  - oxo-enol 45
  - of papain 306
  - of porphyrins 845
  - of pyridoxine 305
  - of pyrimidines 205
- Tautomerization
  - microscopic dissociation constants 305
- Taxol 371*s*
- Teeth 440, 441, **442**
  - diagram of 442
- Teichoic acids
  - in cell walls 431
- Teichuronic acids
  - in cell walls 431
- Telomeres 208
  - DNA in 227
- Temperature 283, 284
  - units of 283
- Temperature jump method 468
- Tenascin 409
- Tensin 407
- Tenuin 407
- Terminal bars 29
- *N*-Terminal nucleophile hydrolases
  - autoactivation of 621
- Termites, protozoa in 19
- Tertiary structure of a protein 59
- TES buffer 99
  - pK_a value of 99
- Testosterone 696
- 1,3,4,6-Tetrachloro-3α,6α-diphenylglycouril 126*s*
- Tetrahedral intermediate
  - in displacement reactions 609
  - stereochemistry of formation 616
- Tetrahedron 342
- Tetrahydrofolic acid (THF) 745, **802 – 813**, 803*s*,
  - derivatives, scheme 806
- *Tetrahymena* 19
- Tetrahymena ribozyme 650*s*
- Tetrameric enzyme 342
- 2,2,6,6-Tetramethylpiperidine-1-oxyl 398*s*
  - EPR spectrum 398
- Tetramethylsilane (TMS) 138
- Tetranitromethane 126
- Tetrodotoxin 736
- Theileriosis 544
- Thermochemistry 282, 283
  - standard states 286–289
- Thermodynamics **281 – 289**
  - base-pairing 209
  - biochemical standard, new 297
  - first law of **282**, **283**
  - irreversible 289
  - of life processes 289
  - nonequilibrium 289
  - second law of **284**, **285**
  - system 281
- Thermodynamic barrier 492
- Thermodynamic temperature scale 284, 285
- Thermolysin 625
  - specificity 117
- Thermostability
  - of enzymes 85
- Thiamin (vitamin B$_1$) *See also* Thiamin diphosphate
  - acid – base chemistry (scheme) 331
  - adducts, 733*s*
  - analogs 733
  - biosynthesis of 736
  - chemical properties of 730, 731
  - cleavage by bisulfite 731
  - cooperative binding of protons 331
  - degradation by base 731
  - nutritional requirement 756
  - pK_a values of 330
  - polyphosphates 736
  - tricyclic form 730, 737
  - yellow anion of 331
  - thiol form 730
- Thiamin coenzymes 730. *See also* Thiamin diphosphate
  - in nerve action 736
  - in transport of sodium 736
- Thiamin diphosphate (Pyrophosphate) 511, 718, **730 – 737**, 767
  - α cleavage 730, 730*s*
  - catalytic mechanisms 731– 733
  - enzymes, table 735
  - enzymatic reactions requiring 734– 736
  - half-reactions, scheme 735
  - minor tautomer 733
  - pyrimidine, role of 733
  - X-ray diffraction 733
- Thiazolium dipolar ion (ylid) 732, 733, 733*s*
- Thin layer chromatography 103
  - of maltooligosaccharides 189
  - *O*-methylated glucoses 190
- Thin-filaments. *See* Actin
- Thiochrome 730, 730*s*
- Thioesters 559, 775
  - displacement reactions of 608
  - formation from glyceraldehyde 3-phosphate 677
  - function of 700
  - mechanism of formation 661, 662
- Thioether 559

Thioglycine 881
Thiohemiacetal in glyceraldehyde phosphate dehydro 775
Thiol(s) 115
  addition to carbonyl group 677
  adduct, dehydrogenation of 766
  binding of mercury ions 125
Thiol – disulfide equilibria 549
Thiol – disulfide exchange 659
Thiol:disulfide oxidoreductases 786
Thiolases 699, 827
  biosynthetic 699
Thiolate anions 779
Thioltransferases 549, 659, 786
Thioredoxin 522, 549, 659, 786*s*, 787, 864
  disulfide loop 786
  functions of 786
Thioredoxin reductase 785, 786, 787, 824
  mechanism of 791
Thiosulfate: cyanide sulfurtransferase
  symmetry in 78
Thiouridine 234
Three-dimensional structures of
  aconitase 689
  adenylate kinase 655
  aldehyde oxido-reductase 891
  D-amino acid oxidase 791
  α-amylase, pancreatic 607
  aspartate aminotransferase 57, 135
    catalytic intermediates 752
  aspartate carbamyltransferase 348
  aspartate chemoreceptor 562
  bacteriophage P22 66
  cadherin 408
  calmodulin 317
  carbonic acid anhydrase I 679
  carboxypeptidase A 64
  catalase 853
  cholera toxin 333, 546
  chymotrypsin 611
  citrate synthase 702, 703
  cutinase 134
  cyclosporin 488
  cytochrome *c* 847
  cytochrome *c* peroxidase 849
  dihydrofolate reductase 807
  DNA 214, 223, 228, 229, 241
  DNA complex
    with methionine repressor 243
    with netropsin 225
    with Trp repressor protein 240
    with zinc finger protein 242
  elastase 611
  enolase 686
  F-actin 338
  favin 64
  ferredoxin 858
  ferritin 343, 841
  glucoamylase 77, 608
  glutathione reductase 785
  glyceraldehyde phosphate dehydrogenase 63
  glycine *N*-methyltransferase 593
  glycogen phosphorylase 605
  Grb2 580
  HIV-1 protease 624
  hydrogenase 879
  insulin 347
  insulin receptor 570
  iron superoxide dismutase 867
  K^+ channel 413
  kringle 2 motif 368
  lactoferrin 840
  lysozyme 599
  methanol dehydrogenase 816
  methionine synthase 874
  methylmalonyl-CoA mutase 874
  myohemerythrin 71
  P pilus 337
  pectate lyase C 686
  plastocyanin 65
  pyruvate dehydrogenase 797
  pyruvate oxidase 734
  rhinovirus 344
  ribulose bisphosphate carboxylase 708
  rRNA 233
  rubisco 708
  rubredoxin 858
  stearoyl-acyl carrier protein desaturase 863
  subtilisin 612
  sulfite reductase 861
  tissue factor 633
  trimethylamine dehydrogenase 784
  triose phosphate isomerase 77, 694
  tRNA 233
  trypsin 611
  trypsinogen 615
  tubulin 372
  virus fd protein sheath 335
  virus φX174 365
  L-*allo*-threonine 53*s*
Threonine (Thr, T) 53*s*, 539
  biosynthesis 540
  cleavage to acetaldehyde 745
  configuration 43
Threonine dehydratase 688, 742
Threonine synthase 746
Threose 163*s*
Thrombin 610, 632
Thrombocytes 26
Thrombomodulin 634
Thrombospondin 409
Thymidine (dThd) 203
Thymidylate synthase 810, 811
  mechanism of, scheme 811
Thymine (Thy) 199*s*, 203
  formation 810
Thymineless death 812
Thyroid peroxidase 856
Thyrotropin (thyrotropic hormone)-releasing factor 54*s*, 523
  receptor 554
Tight junctions 10, 28
Tissue(s)
  adipose 26
  animal 25
  collenchyma 29
  composition of 31
  connective 26
  epidermal of plants 30
  epithelial 26
  fractionation of 98 – 100
  meristematic 29
  muscle 26
  parenchyma 29
  sclerenchyma 29
  supporting 26
Tissue factor 631, 633*s*
  pathway 632
Tissue plasminogen activator 634
Titin 74
Titration curves **96**, **97**
  of proteins 330
Tobacco mosaic virus 247, 334*s*, 336*s*
Tocopherols 392, 815 – 822, 818
  as antioxidants 822
α-Tocopherol (vitamin E) 819*s*
Tocopherolquinone 818, 819, 819*s*
Tocotrienols 818
Togavirus 247
Toluenesulfonyl-amido-2-phenylethyl chloromethyl ketone (TPCK) 623*s*
Tomaymycin 224*s*
TonA 839
TonB 839
Tonegawa, Susumu 84
Tonoplast 11
Topaquinone (TPQ) 816, 817*s*
Topoisomerases 219, 575, 638, 657, 659
Topological theory of graphs 466
Topologies
  of nucleotide binding proteins 76
Topology of proteins 76
  diagram 76
Torsion angles 44
  definition of 59
  of nucleosides 212
  of peptides 63, 139
  of polynucleotides 211
  of side chain groups 61
  of sugar rings 212
Toxins
  channel-forming 414
Toxopyrimidine 739
TPQ *See* Topaquinone
Tracheids 30
Tracheophyta 29
Trachoma 7
Traffic ATPases 417
Transacylation 608
Transaldolase 700
Transaminases. *See* Aminotransferases
Transamination 737
  by PLP-dependent enzymes 741
Transcobalamin 869
Transcortin 58
Transcription
  definition of 5
  of genetic information 5
  regulation of 536
Transcription factor
  osteoblast-specific 441
Transcription factor(s) 76, 539
  AP-1 576
  binding to enhancer site 576
  GCN4 70, 241
  NF-κB, 243*s*
  Max 241, 70*s*
  proto-oncogenes as 576
  Sp1 202
Transcripts. *See* mRNA
Transducin 558
Transferases 498, 589 – 662
  aminotransferases 737, 740 – 743, 747– 753
  definition of 498
Transferrins 58, 840, 841
Transglutaminase 80, 633
Transimination
  in PLP-dependent enzymes 741, 751
Transition metals **837– 894**
Transition state 482 – 486
  diagram 482
  inhibitors 484

  theory 483
Transketolase 733, 736
Translation of genetic information, 5. *See also* Protein synthesis
  definition of 5
  nick 257
  regulation of 536
Translocation
  of proteins 519–521
Transmembrane proteins 391
Transmethylation **591**, **592**
  kinetic isotope effects on 592
Transmission coefficient κ 483
Transport, *See also* specific substrates
  of ions 420–425
  through membranes 400–415
  via pores and channels 401–403
Transporters
  12-helix 415–417
  membrane 411–427
  sugar 415, 416
Transposable genetic elements (transposons) 249
  Tn3 219
Transsialidase 187
Treadmilling
  of microtubules 372
Trehalose 167, 168*s*
  in fungi 168
  in insects 168
Triacylglycerol 381, 382*s*
Trialkyl lock 495*s*
Trichina 24
Tricholomic acid 739
Trichonympha 19
TRICINE buffer 99
  pK_a value of 99
Triclosan 777*s*, 777
Triethanolamine buffer 99
  pK_a value of 99
Trigonal bipyramid 638
  as transition state in pseudorotation 638
Trigonal carbon atoms 680, 681
Trigonal prochiral centers 480, 481
Triiodothyronine 572
Trimethoprim 805*s*
  inhibition of 805
Trimethylamine dehydrogenase 782, 784*s*
Trimethylarsonium lactic acid, 387*s*
Trimethyllysine
  in ubiquitin 525
Tripeptidyl peptidase 140, 610
Triose phosphate isomerase 693, 694*s*
  barrel structure 77*s*
  high catalytic activity 693
  target for antitrypanosomal drugs 693
Triphosphopyridine nucleotide (TPN$^+$, NADP$^+$) 767. *See also* NADP$^+$, NADPH
TRIS buffer
  pK_a value of 99
Tritium
  properties of 110
Triton X 403*s*
tRNA 230, 231, 231*s*, 233*s*
  ^{19}F NMR spectrum of 270
  ^{1}H NMR spectrum of 268
  wobble position 231
Tropical macrocytic anemia 802
Tropoelastin 436
Tropomodulin 406
Tropomyosin 370, 406
  coiled coil structure of 71
Troponin C 313, 314
Trp repressor protein 239, 240*s*
Truffles 20
Tryosine kinases 657
Trypanosome 19
  DNA circles in 219
  mitochondria of 14
Trypanothione 552
Trypanothione reductase 785
Tryparsamide 597*s*
Trypsin 66, 116, 609. *See also* Chymotrypsin
  hydrogen-bonding network, structure 612
  specificity 117
  turnover number of 457
Trypsinogen 609, 615*s*
  oxyanion hole, structure 615
Tryptases 610
Tryptic peptides 117
Tryptophan (Trp, W) 52*s*
  absorption spectrum of derivative 123
  C_β-hydroxylated 853
  nicotinamide activity 769
Tryptophan indole-lyase 742
  quinonoid intermediate 750
Tryptophan synthase 742
Tryptophan tryptophanylquinone (TTQ) 817, 817*s*
T system of tubules 12
TTQ *See* Tryptophan tryptophanylquinone
Tuberculosis 7
Tubular myelin 386
Tubulin 370, 372*s*
Tumor suppressor gene 407, 574
Tungsten 893
Tunicates
  vanadium in 25
Tunichlorin(s) 880*s*, 881
Tunneling
  quantum mechanical 494, 771, 848
  vibration-assisted 494
Turnover numbers of an enzyme 457
Turns
  beta 78
  in protein structures 72, 74
Twisted sheets 63
Two-fold (dyad) axes 134
  in oligomers 337–348
Typhoid fever
  S. typhi 7
Tyrosinase 886
Tyrosine (Tyr, Y) 52*s*
  absorption spectrum of derivative 123
  iodination 126
Tyrosine decarboxylase 737
Tyrosine phenol-lyase 742
Tyrosine protein kinases (MEK) 578
Tyrosine-*O*-sulfate 548
Tyrosyl radical 864
Tyvelose 180*s*

U

Ubichromanol 818, 819*s*
Ubiquinones (coenzyme Q) 392, 514, 818, 819*s*
Ubiquitin 81, 524*s*
  activating enzyme **524**, **525**
  carboxyl-terminal hydrolases 525
  conjugating enzymes 524
  genes 525
  protein ligase 524
UDP-galactose
  conversion to UDP glucose 778
UDP-galactose 4-epimerase 778
Ulothrix 21
Ultracentrifugation **108**, **109**
Ultracentrifuge 100
  analytical 108
  optical system, figure 109
Ultrafiltration 100
Ultrasensitive responses 567
Ultrasonic sound 468
UMP (Uridine 5'-phosphate) 200, 200*s*, 203
Unimolecular processes 457
Uniporters 414
Units, International System 2
Unsatisfied ends of hydrogen bonded chains 64
Uracil (Ura, U) 199*s*, 203
  tautomerism of 45
Urate oxidase 886
Urea 82*s*, 478*s*
Urea carboxylase 730
Urease 478, 877, 878
  active site of, 877*s*
  mechanism of 877
Ureido anion 726*s*
Uric acid 203*s*
  formation of 890
Uridine (Urd) 203, 234*s*
  absorption spectra of 205
  tautomer, minor 205*s*
Uridine 5'- phosphate. *See* UMP
Uridine diphosphate glucose (UDP-Glc) 515*s*, 720
Uridylate kinase 655
5'-Uridylic acid. *See* UMP
Urocanase reaction 778
Urocanic acid 755, 756*s*
Urokinase 634
Uronic acid 164
Uroporphyrin(s) 843
Uroporphyrin I, 845*s*
Urothione 804*s*, 891
Urticaria 385
Usher protein 364

V

Vacuole(s) 10, 11, 12
  definition of 3
Valine (Val, V) 52*s*
  biosynthesis of 527, 712
  branched fatty acids from 381
Valinomycin 414, 415*s*
  effect on potassium transport 414
van der Waals contact surfaces
  of purines and pyrimidines 207
van der Waals forces 46
van der Waals radii **40**, **41**
Vanadate
  as inhibitor of ATPases 889
  insulinlike action 889
Vanadium 856, **889**
Vanadocytes 889
Vanadoproteins 889

- van't Hoff equation 289
- Vapor phase chromatography 103
- Vascular plants 29
- Vasopressin 54*s*, 542, 563
  - receptor 554
- Vesicle(s), coated 426
- VHL (van Hippel-Lindau cancer syndrome) 564
- *Vibrio* 7
  - definition of 6
- Vimentin 369
- Vinblastine 371
- Vincristine 371
- Vinculin 406
- Vinylpyridine 116*s*
- Viral oncogenes 573
- Virion 244
- Viroids 247
- Virus(es) **244–249**. *See also* Bacteriophage
  - adeno 247
  - baculo 247
  - binding to cells 186
  - BK 247
  - cauliflower mosaic 247
  - characteristics, table 245
  - coat 345
  - DNA, table 245
  - dsRNA 244, 244–247
  - electron micrographs of 246
  - Epstein-Barr 247
  - hepatitis A 247
  - hepatitis delta 247
  - herpes 247
  - human immunodeficiency virus (HIV) 248
  - life cycles of 248
  - Mengo 247
  - oncogenic 248
  - papova 244
  - picorna 247
  - plant 346
  - polio 247
  - polyhedral structures 344
  - polyoma 247
  - pox 247
  - protein coat 334
  - rhabdo 247
  - rhino 247
  - RNA containing 247, 248
    - table 245
  - satellite tobacco mosaic 343*s*
  - satellite tobacco necrosis 247, 343*s*, 344
  - simian (SV40) 244
  - ssDNA 244
  - tobacco mosaic 247
  - toga 247
- Virus fd protein sheath 335*s*
- Virus ϕX174
  - assembly of virion structure 365
- Vitamin(s) 719. *See also* Individual vitamins
  - deficiency diseases 721
  - discovery of **721**
- Vitamin B complex 721
  - nutritional requirements 756
- Vitamin B_1 331. *See also* Thiamin
- Vitamin B_{12} 721, **866–877**. *See also* Cobalamin, Cyanocobalamin
  - adenosyltransferase 870
  - blood levels 869
  - cobalt in 866–877
  - coenzyme forms 867
  - nutritional requirements 756
- Vitamin B_{12} coenzyme 864. *See also* Cobalamin
  - dependent reactions, table 871
  - enzymatic functions 870–877
  - isomerization reactions 872
  - nonenzymatic cleavage 870
  - ribonucleotide reductase 871
- Vitamin B_6 family 721, **738**
  - nutritional requirements 756
- Vitamin C. *See* Ascorbic acid (ascorbate)
- Vitamin D 721
  - deposition of calcium in bones 314
- Vitamin E 721. *See also* Tocopherols
  - nutritional requirement 822
- Vitamin K 721, 818, 819*s*, **820–822**
  - in blood clotting 821
  - dihydro 820
  - epoxide 820
  - phylloquinone 821
- Vitronectin 409
- Voltage-gated K^+ channel 412
- Volvox 21
- von Szent-Györgyi, Albert 83
- von Willebrand factor 409, 633
- V system 475

W

- Wald, George 84
- Warburg, Otto H. 83, 767
- Warfarin 821, 822*s*
- Water
  - addition to carbonyl 677
  - clusters of molecules 49
  - content in cells 30
  - content in tissues 31
  - diffusion constant of 461
  - hydrogen bonding of 48
  - properties of **49–51**
  - structure of **49**, **50**
- Watson -Crick structure
  - of DNA 200
- Watson, James D. 84, 200
- Watson-Crick base pair 207, 208, 231
- Waxes 382, 382*s*
- WD repeat 67
- Wheat, genome 12
- Whooping cough 548
- Wilkins, Maurice F. 200
- Wilson's disease 883
- Winding number. *See* DNA, circular, linking number
- Wobble pairing 209
- Wood-Ljundahl pathway 881
- Work 282
  - chemical 282
  - to concentrate 1 mol of a substance 283
  - electrical 282, 302
  - electrochemical 300
  - mechanical 282
  - to raise 1 kg 1 m 283
  - to remove two charges 47
- World Wide Web
  - protein sequences/ structures 148
- Wortmannin 566*s*
- Wounds, healing of 29

X

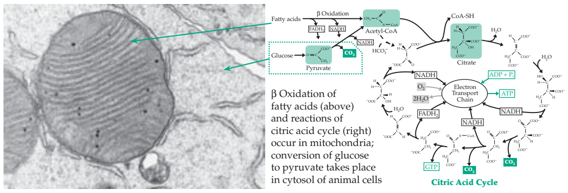
- X-ray diffraction **132–137**
  - difference electron density map 136
  - electron density map 135
  - isomorphous replacement method 133
  - MAD phasing 135
  - refinement 136
  - space groups 133
- X-rays
  - anomolous scattering 135
- Xanthan gum 179
- Xanthine 203*s*
- Xanthine dehydrogenases 794, 825, 890
  - molybdenum in 890
- Xanthine oxidase 890, 892
- Xanthopterin 803, 804*s*
- Xenobiotics **550**
- X-ray diffraction 611
- Xylanases 602, 603
- Xylans 165, 170, 175
- Xylem 30
- Xyloglucan 177
- α-D-Xylopyranose 695*s*
- Xylose (Xyl) 163*s*, 165*s*, 175
- Xylose isomerase 527, 693, 695
- D-Xylose (Xyl) 162
- Xylosidases 602
- Xylulose 164*s*

Y

- Yeasts 20
- Yellow fever 247
- Ylid 733*s*

Z

- Z (zusammen) configuration 43
- Zinc 317, **680**
  - in alcohol dehydrogenases 772–775
  - content of human 680
- Zinc finger protein 241, 242*s*, 243*s*, 680
- Zinc ion(s)
  - chelation by imidazole groups 625
  - in enzymes 773
  - replacements 680
- Zinc proteases 625–627
  - FtsH 628
- Zwitterion 41
- Zygote 17
- Zymogen 519, 609. *See also* Proenzyme; Proenzyme(s)
- Zymogen granules 609
- Zyxin 406



β Oxidation of fatty acids (above) and reactions of citric acid cycle (right) occur in mitochondria; conversion of glucose to pyruvate takes place in cytosol of animal cells

Metabolism, a complex network of chemical reactions, occurs in several different compartments in eukaryotic cells. Fatty acids, a major source of energy for many human cells, are oxidized in the mitochondria via β oxidation and the citric acid cycle. Glucose, a primary source of energy, is converted to pyruvate in the cytosol. Biosynthetic reactions occurring in both compartments form proteins, nucleic acids, storage polymers such as glycogen, and sparingly soluble lipid materials which aggregate to form membranes. Hydrophobic groups in proteins and other polymers also promote self-assembly of the cell. At the same time, oxidative processes, initiated by O_2, increase the water solubility of molecules, leading to metabolic turnover. Micrograph courtesy of Kenneth Moore.

Contents

939 **A. The Oxidation of Fatty Acids**
939 1. Beta Oxidation
941 *Peroxisomal beta oxidation*
941 *Unsaturated fatty acids*
942 *Branched-chain fatty acids*
942 *Oxidation of saturated hydrocarbons*
942 *Alpha oxidation and omega oxidation*
944 2. Carnitine and Mitochondrial Permeability
944 3. Human Disorders of Fatty Acid Oxidation
945 4. Ketone Bodies
947 **B. Catabolism of Propionyl Coenzyme A and Propionate**
947 1. The Malonic Semialdehyde Pathways
950 2. The Methylmalonyl-CoA Pathway of Propionate Utilization
950 **C. The Citric Acid Cycle**
950 1. A Clever Way to Cleave a Reluctant Bond
952 2. Synthesis of the Regenerating Substrate Oxaloacetate
952 3. Common Features of Catalytic Cycles
953 4. Control of the Cycle
957 5. Catabolism of Intermediates of the Citric Acid Cycle
958 **D. Oxidative Pathways Related to the Citric Acid Cycle**
958 1. The γ-Aminobutyrate Cycle
958 2. The Dicarboxylic Acid Cycle
960 **E. Catabolism of Sugars**
960 1. The Glycolysis Pathway
960 *Formation of pyruvate*
962 *The further metabolism of pyruvate*
962 2. Generation of ATP by Substrate Oxidation
963 3. The Pentose Phosphate Pathways
964 *An oxidative pentose phosphate cycle*
965 *Nonoxidative pentose phosphate pathways*
965 4. The Entner–Doudoroff Pathway
966 **F. Fermentation: “Life without Oxygen”**
966 1. Fermentations Based on the Embden–Meyerhof Pathway
966 *Homolactic and alcoholic fermentations*
966 *Energy relationships*
967 *Variations of the alcoholic and homolactic fermentations*
968 2. The Mixed Acid Fermentation
970 3. The Propionic Acid Fermentation
971 4. Butyric Acid and Butanol-Forming Fermentations
972 5. Fermentations Based on the Phosphogluconate and Pentose Phosphate Pathways
972 **G. Biosynthesis**
973 1. Metabolic Loops and Biosynthetic Families
973 2. Key Intermediates and Biosynthetic Families
973 **H. Harnessing the Energy of ATP for Biosynthesis**
974 1. Group Activation
976 2. Hydrolysis of Pyrophosphate
977 3. Coupling by Phosphorylation and Subsequent Cleavage by a Phosphatase
977 4. Carboxylation and Decarboxylation: Synthesis of Fatty Acids
978 **I. Reducing Agents for Biosynthesis**
980 1. Reversing an Oxidative Step with a Strong Reducing Agent
980 2. Regulation of the State of Reduction of the NAD and NADP Systems
981 3. Reduced Ferredoxin in Reductive Biosynthesis
982 **J. Constructing the Monomer Units**
982 1. Carbonyl Groups in Chain Formation and Cleavage
982 2. Starting with CO_2
985 3. Biosynthesis from Other Single-Carbon Compounds
987 4. The Glyoxylate Pathways
989 5. Biosynthesis of Glucose from Three-Carbon Compounds
990 6. Building Hydrocarbon Chains with Two-Carbon Units
990 7. The Oxoacid Chain Elongation Process
992 8. Decarboxylation as a Driving Force in Biosynthesis
992 9. Stabilization and Termination of Chain Growth by Ring Formation
992 10. Branched Carbon Chains
993 **K. Biosynthesis and Modification of Polymers**
993 1. Peptides and Proteins
994 2. Polysaccharides
995 3. Nucleic Acids
995 4. Phospholipids and Phosphate–Sugar Alcohol Polymers
995 5. Irreversible Modification and Catabolism of Polymers
996 **L. Regulation of Biosynthesis**
997 1. Glycogen and Blood Glucose
997 *Insulin*
999 *Glucagon*
999 2. Phosphofructo-1-Kinase in the Regulation of Glycolysis
1000 3. Gluconeogenesis
1000 4. Substrate Cycles
1000 5. Nuclear Magnetic Resonance, Isotopomer Analysis, and Modeling of Metabolism
1002 6. The Fasting State
1003 7. Lipogenesis

1006 **References**
1010 **Study Questions**

Boxes

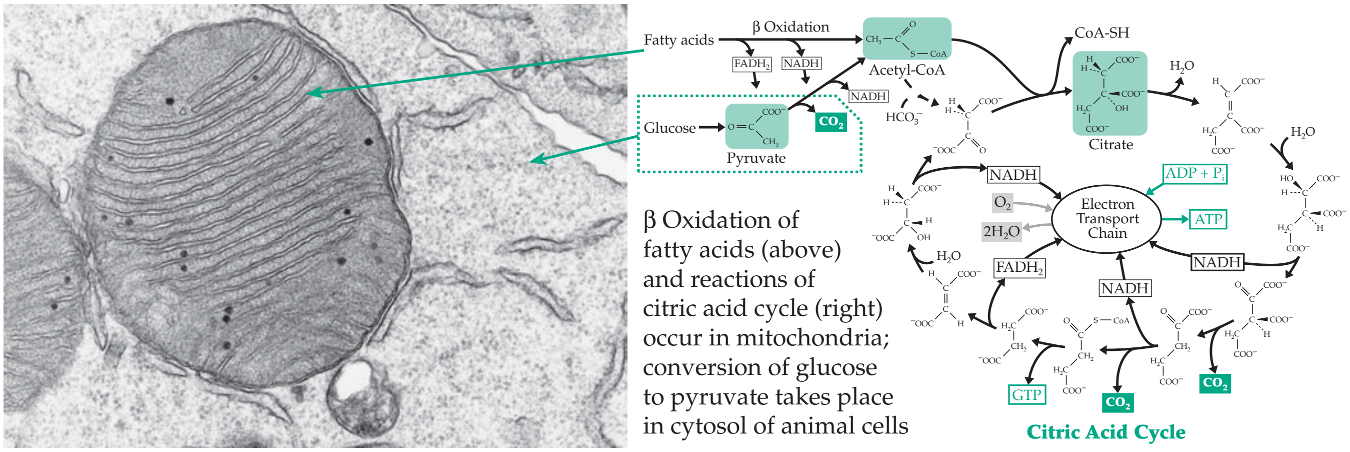
943 Box 17-A Refsum Disease
949 Box 17-B Methylmalonic Aciduria
954 Box 17-C Use of Isotopic Tracers in Study of the Tricarboxylic Acid Cycle
957 Box 17-D Fluoroacetate and “Lethal Synthesis”
985 Box 17-E ^{14}C and the Calvin–Benson Cycle
1002 Box 17-F Lactic Acidemia and Other Deficiencies in Carbohydrate Metabolism
1003 Box 17-G Diabetes Mellitus

Tables

968 Table 17-1 Products of the Mixed Acid Fermentation by *E. coli* at Low and High Values of pH
975 Table 17-2 “Activated” Groups Used in Biosynthesis
998 Table 17-3 Some Effects of Insulin on Enzymes

The Organization of Metabolism

17



β Oxidation of fatty acids (above) and reactions of citric acid cycle (right) occur in mitochondria; conversion of glucose to pyruvate takes place in cytosol of animal cells

Metabolism involves a bewildering array of chemical reactions, many of them organized as complex cycles which may appear difficult to understand. Yet, there is logic and orderliness. With few exceptions, metabolic pathways can be regarded as sequences of the reactions considered in Chapters 12–16 (and summarized in the table inside the back cover) which are organized to accomplish specific chemical goals. In this chapter we will examine the chemical logic of the major pathways of catabolism of foods and of cell constituents as well as some reactions of biosynthesis (anabolism). A few of the sequences have already been discussed briefly in Chapter 10.

A. The Oxidation of Fatty Acids

Hydrocarbons yield more energy upon combustion than do most other organic compounds, and it is, therefore, not surprising that one important type of food reserve, the fats, is essentially hydrocarbon in nature. In terms of energy content the component fatty acids are the most important. Most aerobic cells can oxidize fatty acids completely to CO$_2$ and water, a process that takes place within many bacteria, in the matrix space of animal mitochondria, in the peroxisomes of most eukaryotic cells, and to a lesser extent in the endoplasmic reticulum.

The carboxyl group of a fatty acid provides a point for chemical attack. The first step is a priming reaction in which the fatty acid is converted to a water-soluble acyl-CoA derivative in which the α hydrogens of the fatty acyl radicals are "activated" (step *a*, Fig. 17-1). This synthetic reaction is catalyzed by **acyl-CoA synthetases** (fatty acid:CoA ligases). It is driven by the hydrolysis of ATP to AMP and two inorganic phosphate ions using the sequence shown in Eq. 10-1 (p. 508). There are isoenzymes that act on short-, medium-, and long-chain fatty acids. Yeast contains at least five of these.[1] In every case the acyl group is activated through formation of an intermediate acyl adenylate; hydrolysis of the released pyrophosphate helps to carry the reaction to completion (see discussion in Section H).

1. Beta Oxidation

The reaction steps in the oxidation of long-chain acyl-CoA molecules to acetyl-CoA were outlined in Fig. 10-4. Because of the great importance of this β oxidation sequence in metabolism the steps are shown again in Fig. 17-1 (steps *b*–*e*). The chemical logic becomes clear if we examine the structure of the acyl-CoA molecule and consider the types of biochemical reactions available. If the direct use of O$_2$ is to be avoided, the only reasonable mode of attack on an acyl-CoA molecule is dehydrogenation. Removal of the α hydrogen as a proton is made possible by the activating effect of the carbonyl group of the thioester. The β hydrogen can be transferred from the intermediate enolate, as a hydride ion, to the bound FAD present in the **acyl-CoA dehydrogenases** that catalyze this reaction[2–5] (step *b*, Fig. 17-1; see also Eq. 15-23). These enzymes contain FAD, and the reduced coenzyme FADH$_2$ that is formed is reoxidized by an **electron transferring flavoprotein** (Chapter 15), which also contains FAD. This protein carries the electrons abstracted in the oxidation process to the inner membrane of the mitochondrion where they enter the mitochondrial electron transport system,[5a] as depicted in Fig. 10-5 and as discussed in detail in

Chapter 18.

The product of step *b* is always a **trans-Δ²-enoyl-CoA**. One of the few possible reactions of this unsaturated compound is nucleophilic addition at the β position. The reacting nucleophile is an HO^- ion from water. This reaction step (step *c*, Fig. 17-1) is completed by addition of H^+ at the α position. The resulting **β-hydroxyacyl-CoA** (3-hydroxyacyl-CoA) is dehydrogenated to a ketone by NAD^+ (step *d*).[5b] This series of three reactions is the β oxidation sequence.

At the end of this sequence, the β-oxoacyl-CoA derivative is cleaved (Fig. 17-1, step *e*) by a **thiolase** (see also Eq. 13-35). One of the products is acetyl-CoA, which can be catabolized to CO_2 through the citric acid cycle. The other product of the thiolytic cleavage is an acyl-CoA derivative that is *two carbon atoms shorter than the original acyl-CoA*. This molecule is recycled through the β oxidation process, a two-carbon acetyl unit being removed as acetyl-CoA during each turn of the cycle (Fig. 17-1). The process continues until the fatty acid chain is completely degraded. If the original fatty acid contained an *even* number of carbon atoms in a straight chain, acetyl-CoA is the only product. However, if the original fatty acid contained an *odd* number of carbon atoms, **propionyl-CoA** is formed at the end.

For every step of the β oxidation sequence there is a small family of enzymes with differing chain length preferences.[6,7] For example, in liver mitochondria one acyl-CoA dehydrogenase acts most rapidly on *n*-butyryl and other short-chain acyl-CoA; a second prefers a substrate of medium chain length such as *n*-octanoyl-CoA; a third prefers long-chain substrates such as palmitoyl-CoA; and a fourth, substrates with 2-methyl branches. A fifth enzyme acts specifically on isovaleryl-CoA. Similar preferences exist for the other enzymes of the β oxidation pathway. In *Escherichia coli*

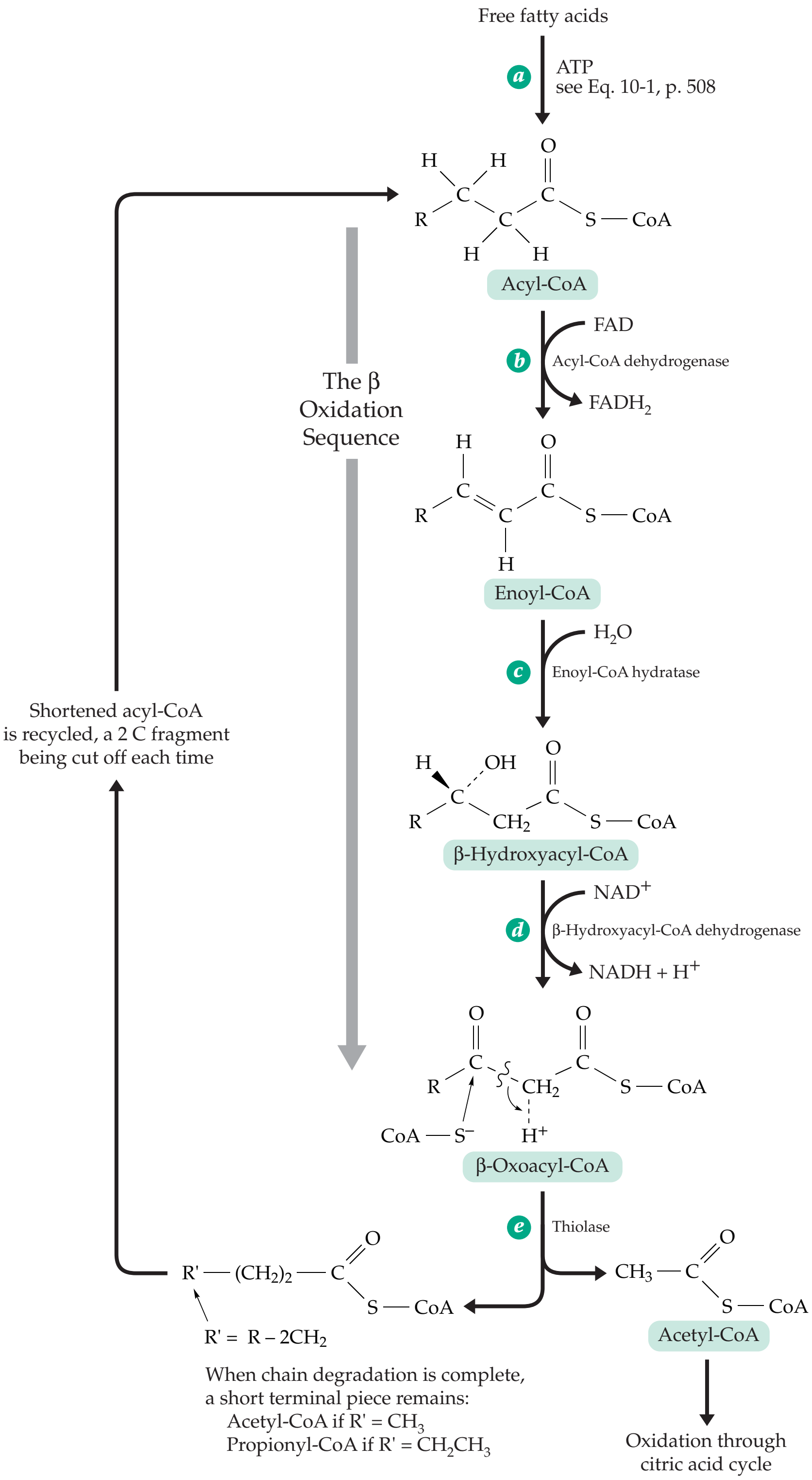

Free fatty acids
a
ATP
see Eq. 10-1, p. 508
Acyl-CoA
b
FAD
Acyl-CoA dehydrogenase
FADH2
Enoyl-CoA
c
H2O
Enoyl-CoA hydratase
β-Hydroxyacyl-CoA
d
NAD+
β-Hydroxyacyl-CoA dehydrogenase
NADH + H+
β-Oxoacyl-CoA
e
Thiolase
Acetyl-CoA
Oxidation through citric acid cycle
The β Oxidation Sequence
Shortened acyl-CoA is recycled, a 2 C fragment being cut off each time
R' = R – 2CH2
When chain degradation is complete, a short terminal piece remains:
Acetyl-CoA if R' = CH3
Propionyl-CoA if R' = CH2CH3

Figure 17-1 The β oxidation cycle for fatty acids. Fatty acids are converted to acyl-CoA derivatives from which 2-carbon atoms are cut off as acetyl-CoA to give a shortened chain which is repeatedly sent back through the cycle until only a 2- or 3-carbon acyl-CoA remains. The sequence of steps *b*, *c*, and *d* also occurs in many other places in metabolism.

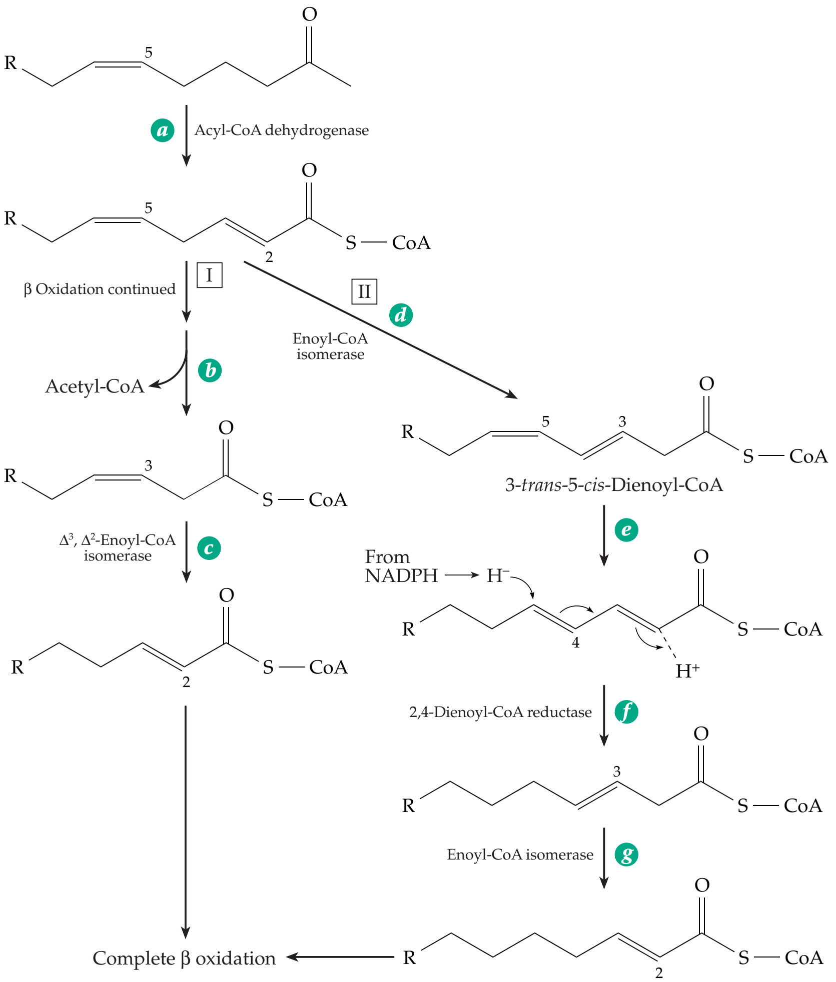
most of these enzymes are present as a complex of multifunctional proteins[8] while the mitochondrial enzymes may be organized as a multiprotein complex.[9,10]

Peroxisomal beta oxidation. In animal cells β oxidation is primarily a mitochondrial process,[5] but it also takes place to a limited extent within peroxisomes and within the endoplasmic reticulum.[11–14] This "division of labor" is still not understood well. Straight-chain fatty acids up to 18 carbons in length appear to be metabolized primarily in mitochondria, but in the liver fatty acids with very long chains are processed largely in peroxisomes.[13] There, a very long-chain acyl-CoA synthetase acts on fatty acids that contain 22 or more carbon atoms.[15] In yeast all β oxidation takes place in peroxisomes,[15,16] and in most organisms, including green plants,[17–18a] the peroxisomes are the most active sites of fatty acid oxidation. However, animal peroxisomes cannot oxidize short-chain acyl-CoA molecules; they must be returned to the mitochondria.[16] The activity of peroxisomes in β oxidation is greatly increased by the presence of a variety of compounds known as **peroxisome proliferators**. Among them are drugs such as aspirin and clofibrate and environmental xenobiotics such as the plasticizer bis-(2-ethyl-hexyl)-phthalate. They may induce as much as a tenfold increase in peroxisomal β oxidation.[11,12,19,19a]

Several other features also distinguish β oxidation in peroxisomes. The peroxisomal flavoproteins that catalyze the dehydrogenation of acyl-CoA molecules to unsaturated enoyl-CoAs (step *b* of Fig. 17-1) are **oxidases** in which the $FADH_2$ that is formed is reoxidized by O_2 to form H_2O_2.[13,20] In peroxisomes the enoyl-hydratase and the NAD^+-dependent dehydrogenase catalyzing steps *c* and *d* of Fig. 17-1 are present together with an enoyl-CoA isomerase (next section) as a trifunctional enzyme consisting of a single polypeptide chain.[21] As in mitochondrial β oxidation the 3-hydroxyacyl-CoA intermediates formed in both animal peroxisomes and plant peroxisomes (glyoxysomes) have the L configuration. However, in fungal peroxisomes as well as in *E. coli* they have the D configuration.[22,23] Further metabolism in these organisms requires an epimerase that converts the D-hydroxyacyl-CoA molecules to L.[24] In the past it has often been assumed that peroxisomal membranes are freely permeable to NAD^+, NADH, and acyl-CoA molecules. However, genetic experiments with yeast and other recent evidence indicate that they are impermeable *in vivo* and that carrier and shuttle mechanisms similar to those in mitochondria may be required.[14,25]

Unsaturated fatty acids. Mitochondrial β oxidation of such unsaturated acids as the Δ^9-oleic acid begins with removal of two molecules of acetyl-CoA to form a Δ^5-acyl-CoA. However, further metabolism is slow. Two pathways have been identified (Eq. 17-1).[26–29b] The first step for both is a normal dehydrogenation to a 2-*trans*-5-*cis*-dienoyl-CoA. In pathway I this intermediate reacts slowly by the normal β oxidation sequence to form a 3-*cis*-enoyl-CoA intermediate which must then be acted upon by an auxiliary enzyme, a *cis*-Δ^3-*trans*-Δ^2-enoyl-CoA isomerase (Eq. 17-1, step *c*), before β oxidation can continue.

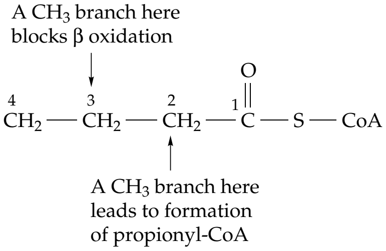
The alternative reductase pathway (II in Eq. 17-1) is often faster. It makes use of an additional isomerase which converts 3-*trans*, 5-*cis*-dienoyl-CoA into the 2-*trans*, 4-*trans* isomer in which the double bonds are conjugated with the carbonyl group.[29] This permits removal of one double bond by reduction with NADPH as shown (Eq. 17-1, step *f*).[29a,29b] The peroxisomal



(17-1)

pathway is similar.[21] However, the intermediate formed in step *e* of Eq. 17-1 may sometimes have the 2-*trans*, 4-*cis* configuration.[17] The NADH for the reductive step *f* may be supplied by an NADP-dependent isocitrate dehydrogenase.[29c] Repetition of steps *a*, *d*, *e*, and *f* of Eq. 17-1 will lead to β oxidation of the entire chain of polyunsaturated fatty acids such as linoleoyl-CoA or arachidonoyl-CoA. Important additional metabolic routes for polyunsaturated fatty acid derivatives are described in Chapter 21.

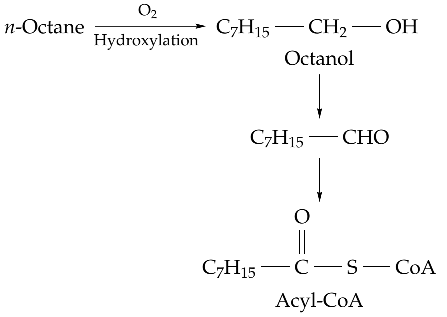
Branched-chain fatty acids. Most of the fatty acids in animal and plant fats have straight unbranched chains. However, branches, usually consisting of methyl groups, are present in lipids of some microorganisms, in waxes of plant surfaces, and also in polyprenyl chains. As long as there are not too many branches and if they occur only in the even-numbered positions (i.e., on carbons 2, 4, etc.) β oxidation proceeds normally. Propionyl-CoA is formed in addition to acetyl-CoA as a product of the chain degradation. On the other hand, if methyl groups occur in positions 3, 5, etc., β oxidation is blocked at step *d* of Fig. 17-1. A striking example of the effect of such blockage was provided by the synthetic detergents in common use until about 1966. These detergents contained a hydrocarbon chain with methyl groups distributed more or less randomly along the chain. Beta oxidation was blocked at many points and the result was a foamy pollution crisis in sewage plants in the United States and in some other countries. Since 1966, only biodegradable detergents having straight hydrocarbon chains have been sold.


A CH_3 branch here
blocks β oxidation
4
3
2
1
O
CH_2 — CH_2 — CH_2 — C — S — CoA
A CH_3 branch here
leads to formation
of propionyl-CoA

In fact, cells *are* able to deal with small amounts of these hard-to-oxidize substrates. The O_2-dependent reactions called α oxidation and ω oxidation are used. These are related also to the oxidation of hydrocarbons which we will consider next.

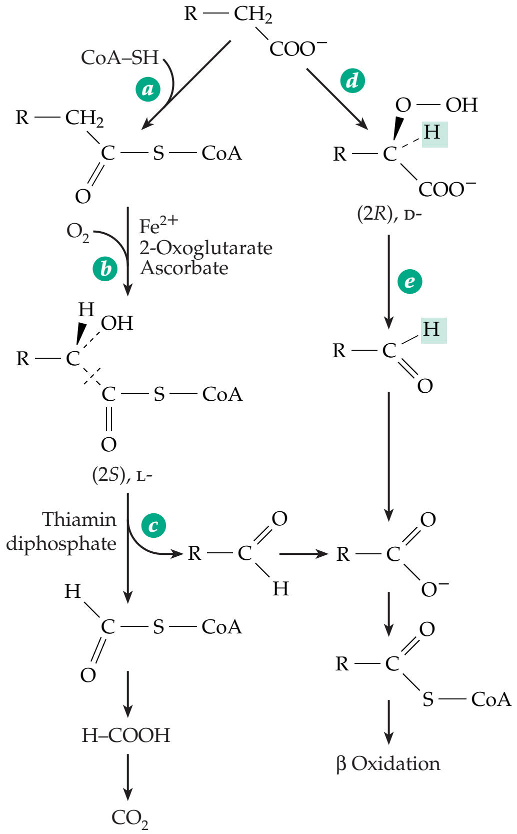
Oxidation of saturated hydrocarbons. Although the initial oxidation step is chemically difficult, the tissues of our bodies are able to metabolize saturated hydrocarbons such as *n*-heptane slowly, and some microorganisms oxidize straight-chain hydrocarbons rapidly.[30,31] Strains of *Pseudomonas* and of the yeast *Candida* have been used to convert petroleum into edible proteins.[9]

The first step in oxidation of alkanes is usually an O_2-requiring **hydroxylation** (Chapter 18) to a primary alcohol. Further oxidation of the alcohol to an acyl-CoA derivative, presumably via the aldehyde (Eq. 17-2), is a frequently encountered biochemical oxidation sequence.


n-Octane
O_2
Hydroxylation
C_7H_{15} — CH_2 — OH
Octanol
C_7H_{15} — CHO
O
C_7H_{15} — C — S — CoA
Acyl-CoA

(17-2)

Alpha oxidation and omega oxidation. Animal tissues degrade such straight-chain fatty acids as palmitic acid, stearic acid, and oleic acid almost entirely by β oxidation, but plant cells often oxidize fatty acids one carbon at a time. The initial attack may involve hydroxylation on the α-carbon atom (Eq. 17-3) to form either the D- or the L-2-hydroxy acid.[17,18,32,32a] The L-hydroxy acids are oxidized rapidly, perhaps by dehydrogenation to the oxo acids (Eq. 17-3, step *b*) and oxidative decarboxylation, possibly utilizing H_2O_2 (see Eq. 15-36). The D-hydroxy acids tend to accumulate


R — CH_2
COO^-
CoA–SH
a
d
R — CH_2
C — S — CoA
O
O — OH
H
R — C
COO^-
(2*R*), D-
O_2
Fe^{2+}
2-Oxoglutarate
Ascorbate
b
e
H
OH
R — C
C — S — CoA
O
(2*S*), L-
R — C
H
O
Thiamin
diphosphate
c
R — C
O
H
R — C
O
O^-
H
C — S — CoA
O
R — C
O
S — CoA
H–COOH
β Oxidation
CO_2

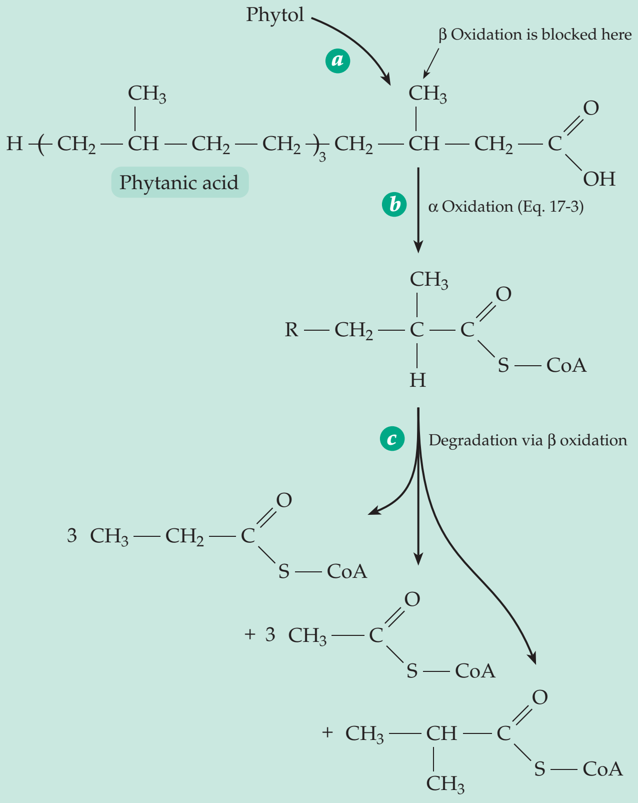
(17-3)

and are normally present in green leaves. However, they too are oxidized further, with retention of the α hydrogen as indicated by the shaded squares in Eq. 17-3, step *e*. This suggests a new type of dehydrogenation with concurrent decarboxylation. Alpha oxidation also occurs to some extent in animal tissues. For example, when β oxidation is blocked by the presence of a methyl side chain, the body may use α oxidation to get past the block (see **Refsum disease**, Box 17-A). As in plants, this occurs principally in the peroxisomes[33–35] and is important for degradation not only of polyprenyl chains but also bile acids. In the brain some of the fatty acyl groups of sphingolipids are hydroxylated to α-hydroxyacyl groups.[36] Alpha oxidation in animal cells occurs after conversion of free fatty acids to their acyl-CoA derivatives (Eq. 7-3, step *a*). This is followed by a 2-oxoglutarate-dependent hydroxylation (step *b*, see also Eq. 18-51) to form the 2-hydroxyacyl-CoA, which is cleaved in a standard thiamin diphosphate-requiring α cleavage (step *c*). The products are formyl-CoA, which is hydrolyzed and oxidized to CO_2, and a fatty aldehyde which is metabolized further by β oxidation.[34a]

In plants α-dioxygenases (Chapter 18) convert free fatty acids into 2(*R*)-hydroperoxy derivatives (Eq. 7-3, step *d*).[32a] These may be decarboxylated to fatty aldehydes (step *e*, see also Eq. 15-36) but may also give rise to a variety of other products. Compounds arising from linoleic and linolenic acids are numerous and include epoxides, epoxy alcohols, dihydroxy acids, short-chain aldehydes, divinyl ethers, and jasmonic acid (Eq. 21-18).[32a]

On other occasions, **omega (ω) oxidation** occurs at the opposite end of the chain to yield a dicarboxylic acid. Within the human body 3,6-dimethyloctanoic acid and other branched-chain acids are degraded largely via ω oxidation. The initial oxidative attack is by a hydroxylase of the cytochrome P450 group (Chapter 18). These enzymes act not only on fatty acids but also on prostaglandins, sterols, and many other lipids. In the animal body fatty acids are sometimes hydroxylated both at the terminal (ω) position and at the next (ω-2 or ω2) carbon. In plants hydroxylation may occur at the ω2, ω3, and ω4 positions as well.[17,37] Dicarboxylates resulting from ω oxidation of straight-chain fatty acids

BOX 17-A REFSUM DISEASE


Phytol
a
β Oxidation is blocked here
$H\text{-}(CH_2-CH(CH_3)-CH_2-CH_2)_3-CH_2-CH(CH_3)-CH_2-COOH$
Phytanic acid
b
α Oxidation (Eq. 17-3)
$R-CH_2-CH(CH_3)-C(=O)-S-CoA$
c
Degradation via β oxidation
$3\ CH_3-CH_2-C(=O)-S-CoA$
$+\ 3\ CH_3-C(=O)-S-CoA$
$+\ CH_3-CH(CH_3)-C(=O)-S-CoA$

In this autosomally inherited disorder of lipid metabolism the 20-carbon branched-chain fatty acid **phytanic acid** accumulates in tissues. Phytanic acid is normally formed in the body (step *a* in the accompanying scheme) from the polyprenyl plant alcohol **phytol**, which is found as an ester in the chlorophyll present in the diet (Fig. 23-20). Although only a small fraction of the ingested phytol is oxidized to phytanic acid, this acid accumulates to a certain extent in animal fats and is present in dairy products. Because β oxidation is blocked, the first step (step *b*) in degradation of phytanic acid is α oxidation in peroxisomes.[a] The remainder of the molecule undergoes β oxidation (step *c*) to three molecules of propionyl-CoA, three of acetyl-CoA, and one of isobutyryl-CoA. The disease, which was described by Refsum in 1946, causes severe damage to nerves and brain as well as lipid accumulation and early death.[b–d] This rare disorder apparently results from a defect in the initial hydroxylation. The causes of the neurological symptoms of Refsum disease are not clear, but it is possible that the isoprenoid phytanic acid interferes with prenylation of membrane proteins.[b]

---

[a] Singh, I., Pahan, K., Dhaunsi, G. S., Lazo, O., and Ozand, P. (1993) *J. Biol. Chem.* **268**, 9972–9979

[b] Steinberg, D. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2351–2369, McGraw-Hill, New York

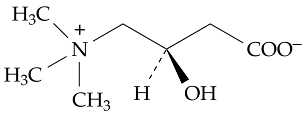
[c] Steinberg, D., Herndon, J. H., Jr., Uhlendorf, B. W., Mize, C. E., Avigan, J., and Milne, G. W. A. (1967) *Science* **156**, 1740–1742

[d] Muralidharan, V. B., and Kishimoto, Y. (1984) *J. Biol. Chem.* **259**, 13021–13026

can undergo β oxidation from both ends. The resulting short-chain dicarboxylates, which appear to be formed primarily in the peroxisomes,[38] may be converted by further β oxidation into succinyl-CoA and free succinate.[39] Incomplete β oxidation in mitochondria (Fig. 17-1) releases small amounts of 3(β)-hydroxy fatty acids, which also undergo ω oxidation and give rise to free 3-hydroxydicarboxylic acids which may be excreted in the urine.[40]

2. Carnitine and Mitochondrial Permeability

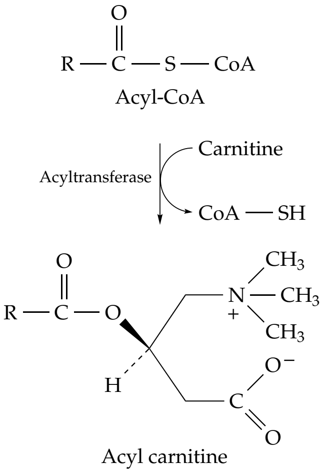
A major factor controlling the oxidation of fatty acids is the rate of entry into the mitochondria. While some long-chain fatty acids (perhaps 30% of the total) enter mitochondria as such and are converted to CoA derivatives in the matrix, the majority are “activated” to acyl-CoA derivatives on the inner surface of the outer membranes of the mitochondria. Penetration of these acyl-CoA derivatives through the mitochondrial inner membrane is facilitated by **L-carnitine**.[41–44]



L-Carnitine

Carnitine is present in nearly all organisms and in all animal tissues. The highest concentration is found in muscle where it accounts for almost 0.1% of the dry matter. Carnitine was first isolated from meat extracts in 1905 but the first clue to its biological action was obtained in 1948 when Fraenkel and associates described a new dietary factor required by the mealworm, *Tenebrio molitor*. At first designated **vitamin B_t**, it was identified in 1952 as carnitine. Most organisms synthesize their own carnitine from lysine side chains (Eq. 24-30). The inner membrane of mitochondria contains a long-chain acyltransferase (carnitine palmitoyltransferase I) that catalyzes transfer of the fatty acyl group from CoA to the hydroxyl group of carnitine (Eq. 17-4).[45–47a] Perhaps acyl carnitine derivatives pass through the membrane more easily than do acyl-CoA derivatives because the positive and negative charges can swing together and neutralize each other as shown in Eq. 17-4. Inside the mitochondrion the acyl group is transferred back from carnitine onto CoA (Eq. 17-4, reverse) by carnitine palmitoyltransferase II prior to initiation of β oxidation.

Tissues contain not only long-chain acylcarnitines but also **acetylcarnitine** and other short-chain acylcarnitines, some with branched chains.[41] By accepting acetyl groups from acetyl-CoA, carnitine causes the release of free coenzyme A which can then be reused. Thus, carnitine may have a regulatory function. In flight muscles of insects acetylcarnitine serves as a reservoir for acetyl groups. Carnitine acyltransferases that act on short-chain acyl-CoA molecules are also present in peroxisomes and microsomes, suggesting that carnitine may assist in transferring acetyl groups and other short acyl groups between cell compartments. For example, acetyl groups from peroxisomal β oxidation can be transferred into mitochondria where they can be oxidized in the citric acid cycle.[41]



(17-4)

3. Human Disorders of Fatty Acid Oxidation

Mitochondrial β oxidation of fatty acids is the principal source of energy for the heart. Consequently, inherited defects of fatty acid oxidation or of carnitine-assisted transport often appear as serious heart disease (inherited cardiomyopathy). These may involve heart failure, pulmonary edema, or sudden infant death. As many as 1 in 10,000 persons may inherit such problems.[48–50a] The proteins that may be defective include a plasma membrane carnitine transporter; carnitine palmitoyltransferases; carnitine/acylcarnitine translocase; long-chain, medium-chain, and short-chain acyl-CoA dehydrogenases; 2,4-dienoyl-CoA reductase (Eq. 17-1); and long-chain 3-hydroxyacyl-CoA dehydrogenase. Some of these are indicated in Fig. 17-2.

Several cases of genetically transmitted carnitine deficiency in children have been recorded. These children have weak muscles and their mitochondria oxidize long-chain fatty acids slowly. If the inner mitochondrial membrane carnitine palmitoyltransferase II is lacking, long-chain acylcarnitines accumulate in the mitochondria and appear to have damaging effects on membranes. In the unrelated condition of **acute myocardial ischemia** (lack of oxygen, e.g., during a heart attack) there is also a large accumulation of long-chain acylcarnitines.[51,52] These compounds may induce cardiac arrhythmia and may also account for sudden death from deficiency of carnitine palmitoyltransferase II. Treatment of disorders of carnitine metabolism with daily oral ingestion of several grams of carnitine is helpful, especially for deficiency of the plasma membrane transporter.[50a,53] Metabolic abnormalities may be corrected completely.[50a]

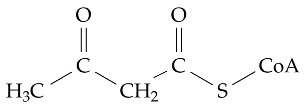
One of the most frequent defects of fatty acid oxidation is deficiency of a mitochondrial acyl-CoA dehydrogenase.[50] If the long-chain-specific enzyme is lacking, the rate of β oxidation of such substrates as octanoate is much less than normal and afflicted individuals excrete in their urine hexanedioic (adipic), octanedioic, and decanedioic acids, all products of ω oxidation.[54] Much more common is the lack of the mitochondrial *medium-chain* acyl-CoA dehydrogenase. Again, dicarboxylic acids, which are presumably generated by ω oxidation in the peroxisomes, are present in blood and urine. Patients must avoid fasting and may benefit from extra carnitine.

A deficiency of very long-chain fatty acid oxidation in peroxisomes is apparently caused by a defective transporter of the ABC type (Chapter 8).[55] The disease, **X-linked adrenoleukodystrophy (ALD)**, has received considerable publicity because of attempts to treat it with "Lorenzo's oil," a mixture of triglycerides of oleic and the C_{22} monoenoic **erucic acid**. The hope has been that these acids would flush out the very long-chain fatty acids that accumulate in the myelin sheath of neurons in the central nervous system and may be responsible for the worst consequences of the disease. However, there has been only limited success.[56,57]

Several genetic diseases involve the development of peroxisomes.[14,35,58,59] Most serious is the **Zellweger syndrome** in which there are no functional peroxisomes. Only "ghosts" of peroxisomes are present and they fail to take up proteins containing the C-terminal peroxisome-targeting sequence SKL.[60,60a] There are many symptoms and death occurs within the first year. Less serious disorders include the presence of catalaseless peroxisomes.[60a]

4. Ketone Bodies

When a fatty acid with an even number of carbon atoms is broken down through β oxidation the last intermediate before complete conversion to acetyl-CoA is the four-carbon **acetoacetyl-CoA**:


O
O
C
C
H3C
CH2
S
CoA

Acetoacetyl-CoA

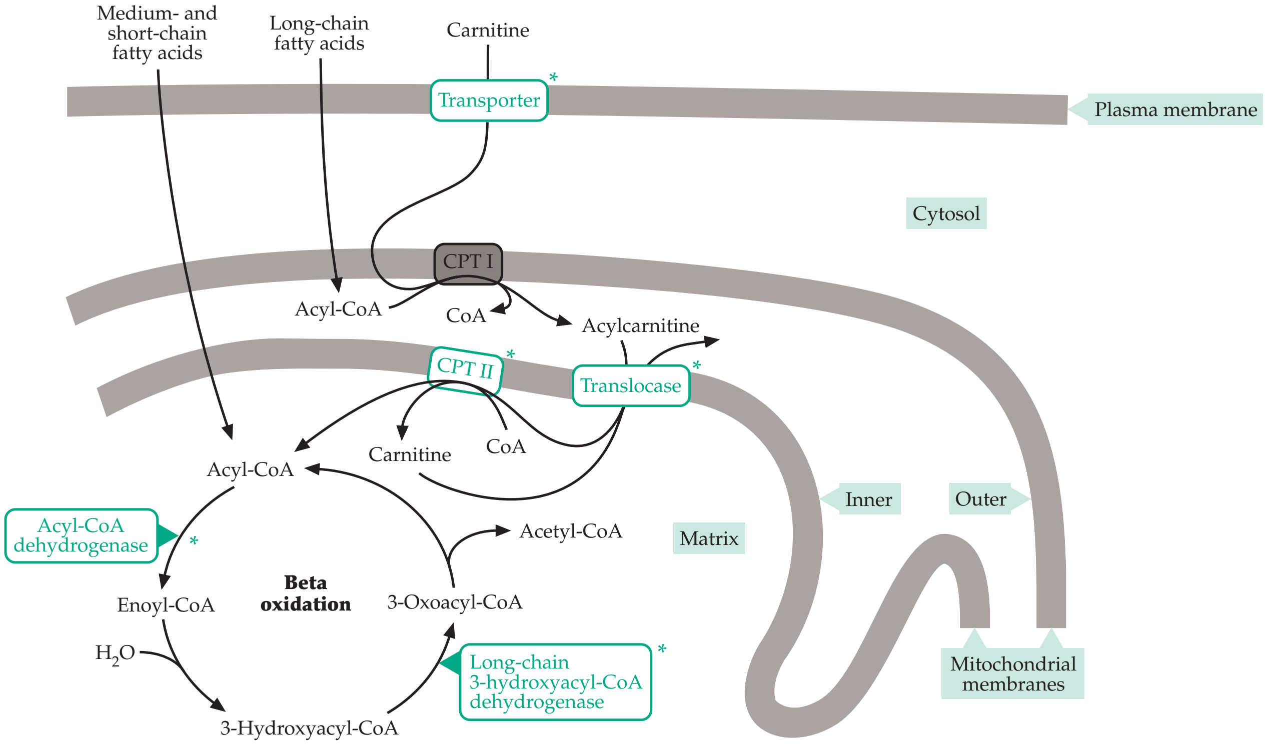
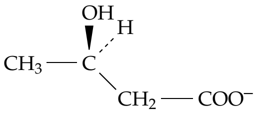

Medium- and short-chain fatty acids
Long-chain fatty acids
Carnitine
Transporter
*
Plasma membrane
Cytosol
CPT I
Acyl-CoA
CoA
Acylcarnitine
CPT II
*
Translocase
*
Carnitine
CoA
Acyl-CoA
Inner
Outer
Acyl-CoA dehydrogenase
*
Acetyl-CoA
Matrix
Beta oxidation
Enoyl-CoA
3-Oxoacyl-CoA
H2O
Long-chain 3-hydroxyacyl-CoA dehydrogenase
*
Mitochondrial membranes
3-Hydroxyacyl-CoA

Figure 17-2 Some specific defects in proteins of β oxidation and acyl-carnitine transport causing cardiomyopathy are indicated by the green asterisks. CPT I and CPT II are carnitine palmitoyltransferases I and II. After Kelly and Strauss.[48]

Acetoacetyl-CoA appears to be in equilibrium with acetyl-CoA within the body and is an important metabolic intermediate. It can be cleaved to two molecules of acetyl-CoA which can enter the citric acid cycle. It is also a precursor for synthesis of polyprenyl (isoprenoid) compounds, and it can give rise to free **acetoacetate**, an important constituent of blood. Acetoacetate is a β-oxoacid that can undergo decarboxylation to acetone or can be reduced by an NADH-dependent dehydrogenase to D-3-hydroxybutyrate. Notice that the configuration of this compound is opposite to that of L-3-hydroxybutyryl-CoA which is formed during β oxidation of fatty acids (Fig. 17-1). D-3-Hydroxybutyrate is sometimes stored as a polymer in bacteria (Box 21-D).

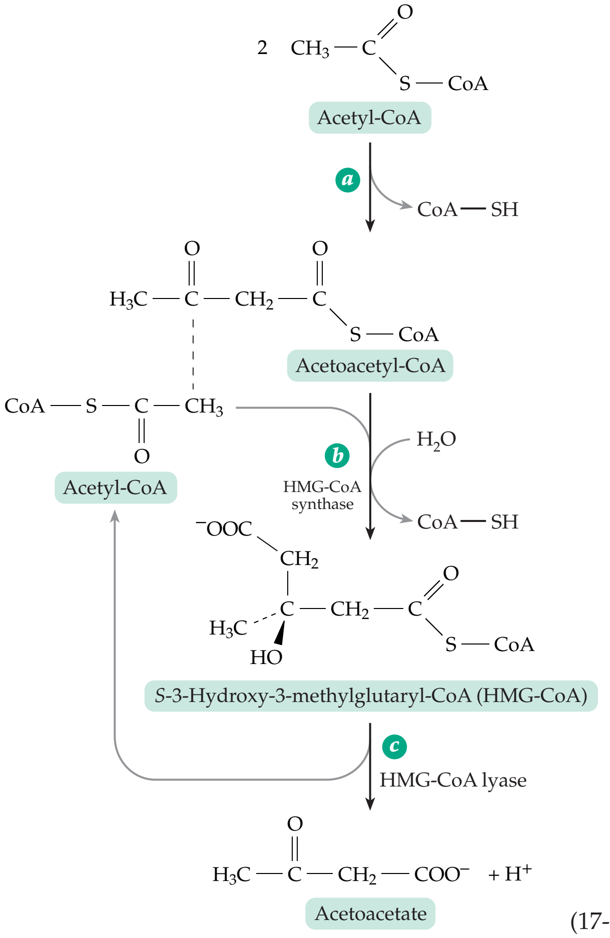


D-3-Hydroxybutyrate

The three compounds, acetoacetate, acetone, and 3-hydroxybutyrate, are known as **ketone bodies**.[60b] The inability of the animal body to form the glucose precursors, pyruvate or oxaloacetate, from acetyl units sometimes causes severe metabolic problems. The condition known as **ketosis**, in which excessive amounts of ketone bodies are present in the blood, develops when too much acetyl-CoA is produced and its combustion in the critic acid cycle is slow. Ketosis often develops in patients with Type I **diabetes mellitus** (Box 17-G), in anyone with high fevers, and during starvation. Ketosis is dangerous, if severe, because formation of ketone bodies produces hydrogen ions (Eq. 17-5) and acidifies the blood. Thousands of young persons with insulin-dependent diabetes die annually from ketoacidosis.

Rat blood normally contains about 0.07 mM acetoacetate, 0.18 mM hydroxybutyrate, and a variable amount of acetone. These amounts increase to 0.5 mM acetoacetate and 1.6 mM hydroxybutyrate after 48 h of starvation. On the other hand, the blood glucose concentration falls from 6 to 4 mM after 48 h starvation.[61] Under these conditions acetoacetate and hydroxybutyrate are an important alternative energy source for muscle and other tissues.[62,63] Acetoacetate can be thought of as a transport form of acetyl units, which can be reconverted to acetyl-CoA and oxidized in the citric acid cycle.

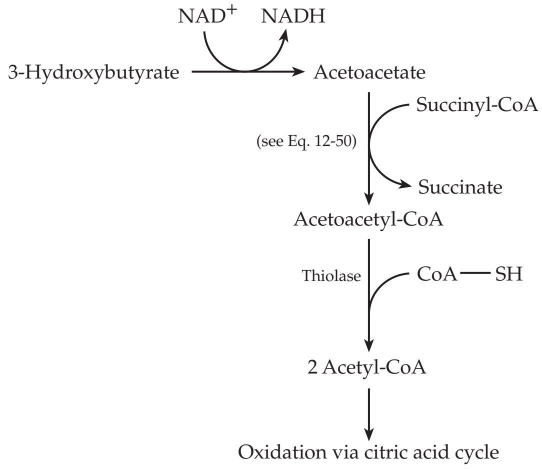
Some free acetoacetate is formed by direct hydrolysis of acetoacetyl-CoA. In rats, ~11% of the hydroxybutyrate that is excreted in the urine comes from acetoacetate generated in this way.[64] However, most acetoacetate arises in the liver indirectly in a two-step process (Eq. 17-5) that is closely related to the synthesis of cholesterol and other polyprenyl compounds. Step *a* of this sequence is a Claisen condensation, catalyzed by 3-hydroxy-3-methyl-glutaryl CoA synthase (HMG-CoA synthase)[64a–c] and followed by hydrolysis of one thioester linkage. It is therefore similar to the citrate synthase reaction (Eq. 13-38). Step *c* is a simple aldol cleavage. The overall reaction has the stoichiometry of a direct hydrolysis of acetoacetyl-CoA. Liver mitochondria contain most of the body's HMG-CoA synthase and are the major site of ketone body formation (**ketogenesis**). Cholesterol is synthesized from HMG-CoA that is formed in the cytoplasm (Chapter 22).



(17-5)

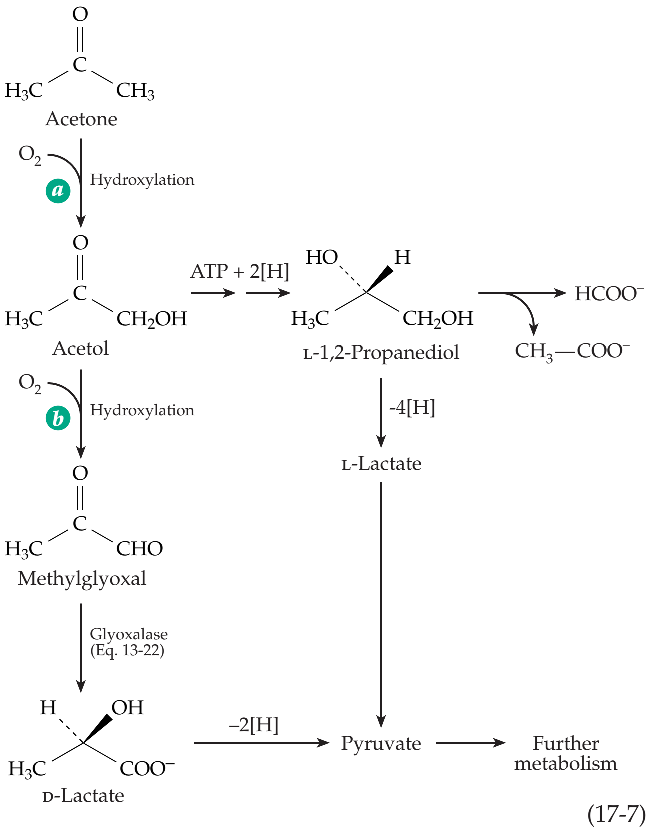
Utilization of 3-hydroxybutyrate or acetoacetate for energy requires their reconversion to acetyl-CoA as indicated in Eq. 17-6. All of the reactions of this sequence may be nearly at equilibrium in tissues that use ketone bodies for energy.[61]

Acetone, in the small amounts normally present in the body, is metabolized by hydroxylation to acetol (Eq. 17-7, step *a*), hydroxylation and dehydration to methylglyoxal (step *b*), and conversion to D-lactate and pyruvate. A second pathway via 1,2-propanediol and L-lactate is also shown in Eq. 17-7. During fasting the acetone content of human blood may rise to as much as 1.6 mM. As much as two-thirds of this may be converted to glucose.[65–69] Accumulation of acetone


NAD+
NADH
3-Hydroxybutyrate
Acetoacetate
Succinyl-CoA
(see Eq. 12-50)
Succinate
Acetoacetyl-CoA
Thiolase
CoA—SH
2 Acetyl-CoA
Oxidation via citric acid cycle

(17-6)

appears to induce the synthesis of the hydroxylases needed for methylglyoxal formation,[68] and the pyruvate formed by Eq. 17-7 may give rise to glucose by the gluconeogenic pathway. However, at high acetone concentrations most metabolism may take place through a poorly understood conversion of 1,2-propanediol to acetate and formate or CO_2.[69] No net conversion of acetate into glucose can occur in animals, but isotopic labels from acetate can enter glucose via acetyl-CoA and the citric acid cycle.


O
C
H3C
CH3
Acetone
O2
a
Hydroxylation
O
C
H3C
CH2OH
Acetol
ATP + 2[H]
HO
H
H3C
CH2OH
L-1,2-Propanediol
HCOO–
CH3—COO–
O2
b
Hydroxylation
-4[H]
L-Lactate
O
C
H3C
CHO
Methylglyoxal
Glyoxalase
(Eq. 13-22)
H
OH
H3C
COO–
D-Lactate
–2[H]
Pyruvate
Further metabolism

(17-7)

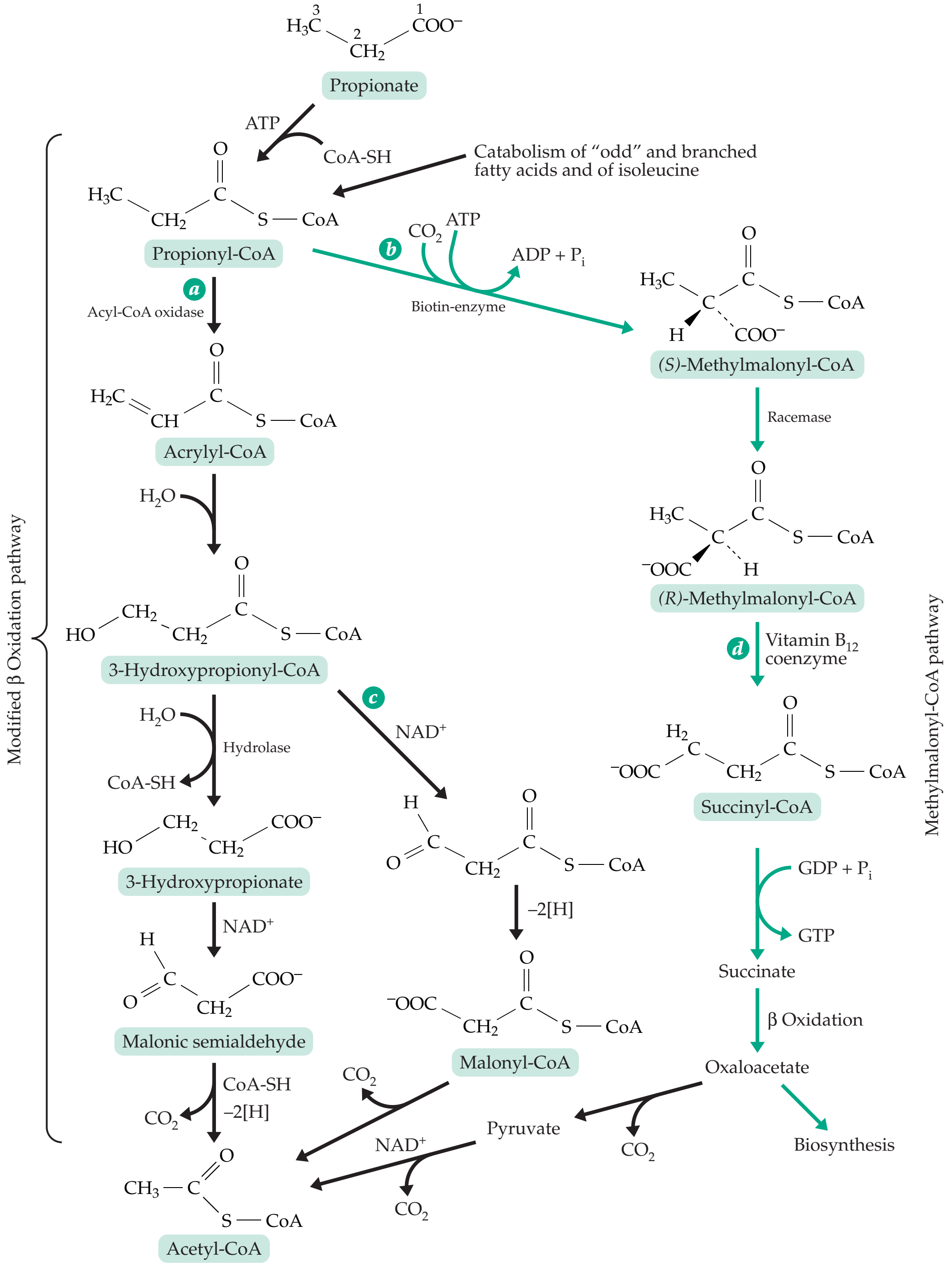
B. Catabolism of Propionyl Coenzyme A and Propionate

Beta oxidation of fatty acids with an odd number of carbon atoms leads to the formation of propionyl-CoA as well as acetyl-CoA. The three-carbon propionyl unit is also produced by degradation of cholesterol and other isoprenoid compounds and of isoleucine, valine, threonine, and methionine. Human beings ingest small amounts of free propionic acid, e.g., from Swiss cheese (which is cultivated with propionic acid-producing bacteria) and from propionate added to bread as a fungicide. In **ruminant** animals, such as cattle and sheep, the ingested food undergoes extensive fermentation in the **rumen**, a large digestive organ containing cellulose-digesting bacteria and protozoa. Major products of the rumen fermentations include acetate, propionate, and butyrate. Propionate is an important source of energy for these animals.

1. The Malonic Semialdehyde Pathways

The most obvious route of metabolism of propionyl-CoA is further β oxidation which leads to 3-hydroxypropionyl-CoA (Fig. 17-3, step *a*). This appears to be the major pathway in green plants.[17] Continuation of the β oxidation via steps *a–c* of Fig. 17-3 produces the CoA derivative of malonic semialdehyde. The latter can, in turn, be oxidized to malonyl-CoA, a β-oxoacid which can be decarboxylated to acetyl-CoA. The necessary enzymes have been found in *Clostridium kluyveri*,[70] but the pathway appears to be little used.

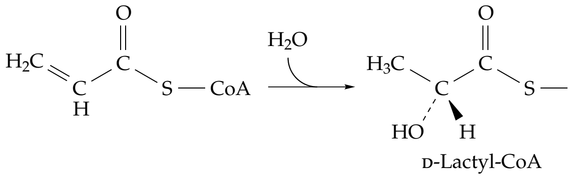
Nevertheless, malonyl-CoA is a major metabolite. It is an intermediate in fatty acid synthesis (see Fig. 17-12) and is formed in the peroxisomal β oxidation of odd chain-length dicarboxylic acids.[70a] Excess malonyl-CoA is decarboxylated in peroxisomes, and lack of the decarboxylase enzyme in mammals causes the lethal **malonic aciduria**.[70a] Some propionyl-CoA may also be metabolized by this pathway. The modified β oxidation sequence indicated on the left side of Fig. 17-3 is used in green plants and in many microorganisms. 3-Hydroxypropionyl-CoA is hydrolyzed to *free* β-hydroxypropionate, which is then oxidized to malonic semialdehyde and converted to acetyl-CoA by reactions that have not been completely described. Another possible pathway of propionate metabolism is the direct conversion to pyruvate via α oxidation into lactate, a mechanism that may be employed by some bacteria. Another route to lactate is through addition of water to acrylyl-CoA, the product of step *a* of Fig. 17-3. The water molecule adds in the "wrong way," the OH^- ion going to the α carbon instead of the β (Eq. 17-8). An enzyme with an active site similar to that of histidine ammonia-lyase (Eq. 14-48) could



3
1
H3C
2
CH2
COO−
Propionate
ATP
CoA-SH
Catabolism of "odd" and branched fatty acids and of isoleucine
O
H3C
CH2
C
S — CoA
Propionyl-CoA
a
Acyl-CoA oxidase
b
CO2
ATP
ADP + Pi
Biotin-enzyme
H3C
C
H
COO−
(S)-Methylmalonyl-CoA
Racemase
−OOC
H
(R)-Methylmalonyl-CoA
d
Vitamin B12 coenzyme
H2C
CH
Acrylyl-CoA
H2O
HO
3-Hydroxypropionyl-CoA
c
NAD+
Hydrolase
CoA-SH
3-Hydroxypropionate
−2[H]
Malonic semialdehyde
Malonyl-CoA
H2
−OOC
Succinyl-CoA
GDP + Pi
GTP
Succinate
β Oxidation
Oxaloacetate
Pyruvate
Biosynthesis
CO2
−2[H]
CH3
Acetyl-CoA
Modified β Oxidation pathway
Methylmalonyl-CoA pathway

Figure 17-3 Catabolism of propionate and propionyl-CoA. In the names for methylmalonyl-CoA the *R* and *S* refer to the methylmalonyl part of the structure. Coenzyme A is also chiral.

presumably catalyze such a reaction. Lactyl-CoA could be converted to pyruvate readily. *Clostridium propionicum* does interconvert propionate, lactate, and pyruvate via acrylyl-CoA and lactyl-CoA as part of a fermentation of alanine (Fig. 24-19).[71–74] The enzyme that catalyzes hydration of acrylyl-CoA in this case is a complex flavoprotein that may function via a free radical mechanism.[71,72,74]



H2C
C
H
O
C
S — CoA
H2O
H3C
HO
H
D-Lactyl-CoA

(17-8)

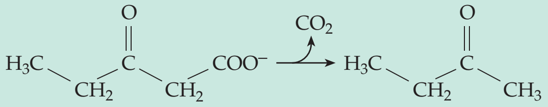
BOX 17-B METHYLMALONIC ACIDURIA

In this hereditary disease up to 1–2 g of methylmalonic acid per day (compared to a normal of <5 mg/day) is excreted in the urine, and a high level of the compound is present in blood. Two causes of the rare disease are known.[a–d] One is the lack of functional vitamin B_{12}-containing coenzyme. This can be a result of a mutation in any one of several different genes involved in the synthesis and transport of the cobalamin coenzyme.[e] Cultured fibroblasts from patients with this form of the disease contain a very low level of the vitamin B_{12} coenzyme (Chapter 16), and addition of excess vitamin B_{12} to the diet may restore coenzyme synthesis to normal. Among elderly patients a smaller increase in methylmalonic acid excretion is a good indicator of vitamin B_{12} deficiency. A second form of the disease, which does not respond to vitamin B_{12}, arises from a defect in the methylmalonyl mutase protein. Methylmalonic aciduria is often a very severe disease, frequently resulting in death in infancy. Surprisingly, some children with the condition are healthy and develop normally.[a,f]

A closely related disease is caused by a deficiency of propionyl-CoA carboxylase.[a] This may be a result of a defective structural gene for one of the two subunits of the enzyme, of a defect in the enzyme that attaches biotin to carboxylases, or of biotinitase, the enzyme that hydrolytically releases biotin from linkage with lysine (Chapter 14). The latter two defects lead to a multiple carboxylase deficiency and to methylmalonyl aciduria as well as ketoacidosis and propionic acidemia.[g]

Both methylmalonic aciduria and propionyl-CoA decarboxylase deficiency are usually accompanied by severe ketosis, hypoglycemia, and hyperglycinemia. The cause of these conditions is not entirely clear. However, methylmalonyl-CoA, which accumulates in methylmalonic aciduria, is a known inhibitor of pyruvate carboxylase. Therefore, ketosis may develop because of impaired conversion of pyruvate to oxaloacetate.

Patients with propionic or methylmalonic acidemia also secrete 2,3-butanediols (D-,L- or meso) and usually also 1,2-propandiol in their urine. Secretion of 1,2-propanediol is also observed during starvation and in diabetic ketoacidosis. Propanediol may be formed from acetone (Eq. 17-7), and butanediols may originate from acetoin, which is a side reaction product of pyruvate dehydrogenase. However, in the metabolic defects under consideration here, acetoin may be formed by hydroxylation of methylethyl ketone which can arise, as does acetone, by decarboxylation of an oxoacid precursor formed by β oxidation.[h]



Methylmalonic aciduria is rare and can be diagnosed incorrectly. In 1989 a woman in St. Louis, Missouri, was convicted and sentenced to life in prison for murdering her 5-month-old son by poisoning with ethylene glycol. While in prison she gave birth to another son who soon fell ill of methylmalonyl aciduria and was successfully treated. Reexamination of the evidence revealed that the first boy had died of the same disease and the mother was released.[i]

---

[a] Fenton, W. A., and Rosenberg, L. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1423–1449, McGraw-Hill, New York

[b] Matsui, S. M., Mahoney, M. J., and Rosenberg, L. E. (1983) *N. Engl. J. Med.* **308**, 857–861

[c] Hubbard, S. R., Wei, L., Ellis, L., and Hendrickson, W. A. (1994) *Nature (London)* **372**, 746–754

[d] Luschinsky Drennan, C., Matthews, R. G., Rosenblatt, D. S., Ledley, F. D., Fenton, W. A., and Ludwig, M. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5550–5555

[e] Fenton, W. A., and Rosenberg, L. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3129–3149, McGraw-Hill, New York

[f] Ledley, F. D., Levy, H. L., Shih, V. E., Benjamin, R., and Mahoney, M. J. (1984) *N. Engl. J. Med.* **311**, 1015–1018

[g] Wolf, B. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3151–3177, McGraw-Hill, New York

[h] Casazza, J. P., Song, B. J., and Veech, R. L. (1990) *Trends Biochem. Sci.* **15**, 26–30

[i] Zurer, P. (1991) *Chem. Eng. News* **69** Sep 30, 7–8

2. The Methylmalonyl-CoA Pathway of Propionate Utilization

Despite the simplicity and logic of the β oxidation pathway of propionate metabolism, higher animals use primarily the more complex methylmalonyl-CoA pathway (Fig. 17-3, step *b*). This is one of the two processes in higher animals presently known to depend upon vitamin B_{12}. This vitamin has never been found in higher plants, nor does the methylmalonyl pathway occur in plants. The pathway (Fig. 17-3) begins with the biotin- and ATP-dependent carboxylation of propionyl-CoA. The *S*-methylmalonyl-CoA so formed is isomerized to *R*-methylmalonyl-CoA, after which the methylmalonyl-CoA is converted to succinyl-CoA in a vitamin B_{12} coenzyme-requiring reaction step *d* (Table 16-1). The succinyl-CoA is converted to free succinate (with the formation of GTP compensating for the ATP used initially). The succinate, by β oxidation, is converted to oxaloacetate which is decarboxylated to pyruvate. This, in effect, removes the carboxyl group that was put on at the beginning of the sequence in the ATP-dependent step. Pyruvate is converted by oxidative decarboxylation to acetyl-CoA.

A natural question is "Why has this complex pathway evolved to do something that could have been done much more directly?" One possibility is that the presence of too much malonyl-CoA, the product of the β oxidation pathway of propionate metabolism (Fig. 17-3, pathways *a* and *c*), would interfere with lipid metabolism. Malonyl-CoA is formed in the cytosol during fatty acid biosynthesis and retards mitochondrial β oxidation by inhibiting carnitine palmitoyltransferase I.[46,70a,75] However, a relationship to mitochondrial propionate catabolism is not clear. On the other hand, the tacking on of an extra CO_2 and the use of ATP at the beginning suggests that the *methylmalonyl-CoA pathway* (Fig. 17-3) *is a biosynthetic rather than a catabolic route* (see Section H,4). The methylmalonyl pathway provides a means for converting propionate to oxaloacetate, a transformation that is chemically difficult.

In this context it is of interest that cows, whose metabolism is based much more on acetate than is ours, often develop a severe ketosis spontaneously. A standard treatment is the administration of a large dose of propionate which is presumably effective because of the ease of its conversion to oxaloacetate via the methylmalonyl-CoA pathway. It is possible that this pathway was developed by animals as a means of capturing propionyl units, scanty though they may be, for conversion to oxaloacetate and use in biosynthesis. In ruminant animals, the pathway is especially important. Whereas we have 5.5 mM glucose in our blood, the cow has only half as much, and a substantial fraction of this glucose is derived, in the liver, from the propionate provided by rumen microorganisms.[76] The need for vitamin B_{12} in the formation of propionate by these organisms also accounts for the high requirement for cobalt in the ruminant diet (Chapter 16).

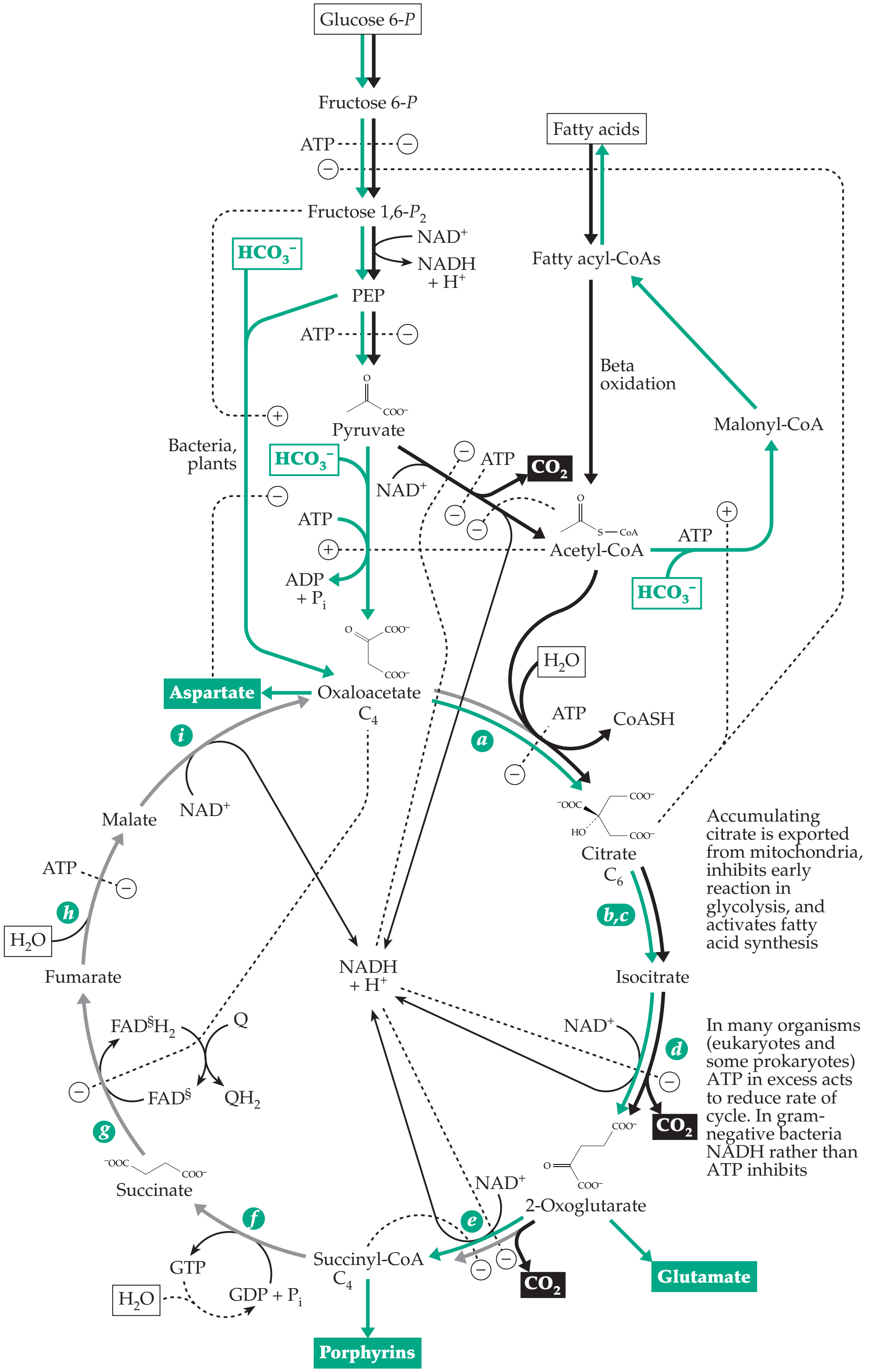
C. The Citric Acid Cycle

To complete the oxidation of fatty acids the acetyl units of acetyl-CoA generated in the β oxidation sequence must be oxidized to carbon dioxide and water.[77] The citric acid (or tricarboxylic acid) cycle by which this oxidation is accomplished is a vital part of the metabolism of almost all aerobic creatures. It occupies a central position in metabolism because of the fact that acetyl-CoA is also an intermediate in the catabolism of carbohydrates and of many amino acids and other compounds. The cycle is depicted in detail in Fig. 10-6 and in an abbreviated form, but with more context, in Fig. 17-4.

1. A Clever Way to Cleave a Reluctant Bond

Oxidation of the chemically resistant two-carbon acetyl group to CO_2 presents a chemical problem. As we have seen (Chapter 13), cleavage of a C–C bond occurs most frequently between atoms that are α and β to a carbonyl group. Such β cleavage is clearly impossible within the acetyl group. The only other common type of cleavage is that of a C–C bond adjacent to a carbonyl group (α cleavage), a thiamin-dependent process (Chapter 14). However, α cleavage would require the prior oxidation (hydroxylation) of the methyl group of acetate. Although many biological hydroxylation reactions occur, they are rarely used in the major pathways of rapid catabolism. Perhaps this is because the overall yield of energy obtainable via hydroxylation is less than that gained from dehydrogenation and use of an electron transport chain.[78]

The solution to the chemical problem of oxidizing acetyl groups efficiently is one very commonly found in nature; a catalytic cycle. Although direct cleavage is impossible, the two-carbon acetyl group of acetyl-CoA *can* undergo a Claisen condensation with a second compound that contains a carbonyl group. The condensation product has more than two carbon atoms, and a β cleavage to yield CO_2 is now possible. Since the cycle is designed to oxidize acetyl units we can regard acetyl-CoA as the **primary substrate** for the cycle. The carbonyl compound with which it condenses can be described as the **regenerating substrate**. To complete the catalytic cycle it is necessary that two carbon atoms be removed as CO_2 from the compound formed by condensation of the two substrates and that the remaining molecule be reconvertible to the original regenerating substrate. The reader may wish to play a


Glucose 6-P
Fructose 6-P
ATP
Fatty acids
Fructose 1,6-P_2
NAD^+
NADH
+ H^+
HCO_3^-
PEP
Fatty acyl-CoAs
ATP
Beta
oxidation
Pyruvate
Malonyl-CoA
Bacteria,
plants
HCO_3^-
ATP
CO_2
NAD^+
ATP
Acetyl-CoA
ATP
ADP
+ P_i
HCO_3^-
H_2O
Aspartate
Oxaloacetate
C_4
ATP
CoASH
i
a
Malate
NAD^+
Citrate
C_6
ATP
h
b,c
H_2O
Fumarate
NADH
+ H^+
Isocitrate
$FAD^§H_2$
Q
NAD^+
d
$FAD^§$
QH_2
CO_2
g
Succinate
NAD^+
2-Oxoglutarate
f
e
GTP
Succinyl-CoA
C_4
H_2O
GDP + P_i
CO_2
Glutamate
Porphyrins

Accumulating citrate is exported from mitochondria, inhibits early reaction in glycolysis, and activates fatty acid synthesis

In many organisms (eukaryotes and some prokaryotes) ATP in excess acts to reduce rate of cycle. In gram-negative bacteria NADH rather than ATP inhibits

Figure 17-4 The Krebs citric acid cycle. Some of its controlling interactions and its relationship to glycolysis. See also Figure 10-6. Positive and negative regulatory influences, whether arising by allosteric effects or via covalent modification, are indicated by ⊕ or ⊖. Some biosynthetic reaction pathways related to the cycle are shown in green. Steps are lettered to correspond to the numbering in Fig. 10-6, which shows more complete structural formulas. Three molecules of H_2O (boxed) enter the cycle at each turn, providing hydrogen atoms for generation of NADH + H^+ and reduced ubiquinone (QH_2). The covalently attached FAD is designated $FAD^§$.

game by devising suitable sequences of reactions for an acetyl-oxidizing cycle and finding the simplest possible regenerating substrate. Ask yourself whether nature could have used anything simpler than **oxaloacetate**, the molecule actually employed in the citric acid cycle.

The first step in the citric acid cycle (step *a*, Fig. 17-4) is the condensation of acetyl-CoA with oxaloacetate to form citrate. The synthase that catalyzes this condensation also removes the CoA by hydrolysis after it has served its function of activating a methyl hydrogen. This hydrolysis also helps to drive the cycle by virtue of the high group transfer potential of the thioester linkage that is cleaved. Before the citrate can be degraded through β cleavage, the hydroxyl group must be moved from its tertiary position to an adjacent carbon where, as a secondary alcohol, it can be oxidized to a carbonyl group. This is accomplished through steps *b* and *c*, both catalyzed by the enzyme aconitase (Eq. 13-17). Isocitrate can then be oxidized to the β-oxoacid **oxalosuccinate**, which does not leave the enzyme surface but undergoes decarboxylation while still bound (step *d*; also see Eq. 13-45).

The second carbon to be removed from citrate is released as CO_2 through catalysis by the thiamin diphosphate dependent **oxidative decarboxylation** of **2-oxoglutarate** (α-ketoglutarate; Chapter 15). To complete the cycle the four-carbon succinyl unit of succinyl-CoA must be converted back to oxaloacetate through a pathway requiring two more oxidation steps: Succinyl-CoA is converted to free succinate (step *f*) followed by a β oxidation sequence (steps *g*–*i*; Figs. 10-6 and 17-4). Steps *e* and *f* accomplish a substrate-level phosphorylation (Fig. 15-16). Succinyl-CoA is an unstable thioester with a high group transfer potential. Therefore, step *f* could be accomplished by simple hydrolysis. However, this would be energetically wasteful. The cleavage of succinyl-CoA is coupled to synthesis of ATP in *E. coli* and higher plants and to GTP in mammals. Some of the succinyl-CoA formed in mitochondria is used in other ways, e.g., as in Eq. 17-6 and for biosynthesis of porphyrins.

2. Synthesis of the Regenerating Substrate Oxaloacetate

The primary substrate of the citric acid cycle is acetyl-CoA. Despite many references in the biochemical literature to substrates "entering" the cycle as oxaloacetate (or as one of the immediate precursors succinate, fumarate, or malate), *these compounds are not consumed* by the cycle but are completely regenerated; hence the term *regenerating substrate*, which can be applied to any of these four substances. A prerequisite for the operation of a catalytic cycle is that a regenerating substrate be readily available and that its concentration be increased if necessary to accommodate a more rapid rate of reaction of the cycle. Oxaloacetate can normally be formed in any amount needed for operation of the citric acid cycle from **PEP** or from **pyruvate**, both of these compounds being available from metabolism of sugars.

In bacteria and green plants **PEP carboxylase** (Eq. 13-53), a highly regulated enzyme, is responsible for synthesizing oxaloacetate. In animal tissues **pyruvate carboxylase** (Eq. 14-3) plays the same role. The latter enzyme is almost inactive in the absence of the allosteric effector acetyl-CoA. For this reason, it went undetected for many years. In the presence of high concentrations of acetyl-CoA the enzyme is fully activated and provides for synthesis of a high enough concentration of oxaloacetate to permit the cycle to function. Even so, the oxaloacetate concentration in mitochondria is low, only 0.1 to 0.4 x 10^{-6} M (10–40 molecules per mitochondrion), and is relatively constant.[65,79]

3. Common Features of Catalytic Cycles

The citric acid cycle is not only one of the most important metabolic cycles in aerobic organisms, including bacteria, protozoa, fungi, higher plants, and animals, but also *a typical catalytic cycle*. Other cycles also have one or more primary substrates and at least one regenerating substrate. Associated with every catalytic cycle there must be a metabolic pathway that provides for synthesis of the regenerating substrate. Although it usually needs to operate only slowly to replenish regenerating substrate lost in side reactions, the pathway also provides *a mechanism for the net biosynthesis of any desired quantity of any intermediate in the cycle*. Cells draw off from the citric acid cycle considerable amounts of oxaloacetate, 2-oxoglutarate, and succinyl-CoA for synthesis of other compounds. For example, aspartate and glutamate are formed directly from oxaloacetate and 2-oxoglutarate, respectively, by transamination (Eq. 14-25).[79a,b] Citrate itself is exported from mitochondria and used for synthesis of fatty acids. It is often stated that the citric acid cycle functions in biosynthesis, but when intermediates in the cycle are drawn off for synthesis the complete cycle does not operate. Rather, *the pathway for synthesis of the regenerating substrate, together with some of the enzymes of the cycle, is used to construct a biosynthetic pathway.*

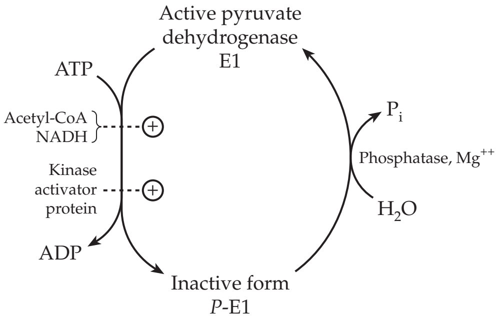
The word **amphibolic** is often applied to those metabolic sequences that are part of a catabolic cycle and at the same time are involved in a biosynthetic (anabolic) pathway. Another term, **anaplerotic**, is sometimes used to describe pathways for the synthesis of regenerating substrates. This word, which was suggested by H. L. Kornberg, comes from a Greek root meaning "filling up."[80]

4. Control of the Cycle

What factors determine the rate of oxidation by the citric acid cycle? As with most other important pathways of metabolism, several control mechanisms operate and different steps may become rate limiting under different conditions.[81] Factors influencing the flux through the cycle include (1) the rate of generation of acetyl groups, (2) the availability of oxaloacetate, and (3) the rate of reoxidation of NADH to NAD^+ in the electron transport chain. As indicated in Fig. 17-4, acetyl-CoA is a positive effector for conversion of pyruvate to oxaloacetate. Thus, acetyl-CoA "turns on" the formation of a substance required for its own further metabolism. However, when no pyruvate is available operation of the cycle may be impaired by lack of oxaloacetate. This may be the case when liver metabolizes high concentrations of ethanol. The latter is oxidized to acetate but it cannot provide oxaloacetate. Accumulating NADH reduces pyruvate to lactate, further interfering with formation of oxaloacetate.[82] In some individuals the accumulating acetyl units cannot all be oxidized in the cycle and instead are converted to the ketone bodies (Section A,4). A similar problem arises during metabolism of fatty acids by diabetic individuals with inadequate insulin. The accelerated breakdown of fatty acids in the liver overwhelms the system and results in ketosis, even though the oxaloacetate concentration remains normal.[83]

The rates of the oxidative steps in the citric acid cycle are limited by the rate of reoxidation of NADH and reduced ubiquinone in the electron transport chain which may sometimes be restricted by the availability of O_2. However, in aerobic organisms this rate is usually determined by the concentration of ADP and/or P_i available for conversion to ATP in the oxidative phosphorylation process (Chapter 18). If catabolism supplies an excess of ATP over that needed to meet the cell's energy needs, the concentration of ADP falls to a low level, cutting off phosphorylation. At the same time, ATP is present in high concentration and acts as a feedback inhibitor for the catabolism of carbohydrates and fats. This inhibition is exerted at many points, a few of which are indicated in Fig. 17-4. Important sites of inhibition are the **pyruvate dehydrogenase complex**,[84–85a] which converts pyruvate into acetyl-CoA; **isocitrate dehydrogenase**,[86,86a] which converts isocitrate into 2-oxoglutarate; and **2-oxoglutarate dehydrogenase**.[87] The enzyme **citrate synthase**, which catalyzes the first reaction of the cycle, is also inhibited by ATP.[88,89]

Mitochondrial pyruvate dehydrogenase, which contains a 60-subunit icosohedral core of dihydrolipoyltransacylase (Fig. 15-14), is associated with three molecules of a two-subunit kinase as well as six molecules of a structural **binding protein** which contains a lipoyl group that can be reduced and acetylated by other subunits of the core protein. The binding protein is apparently essential to the functioning of the dehydrogenase complex but not through its lipoyl group.[90,91] The specific pyruvate dehydrogenase kinase is thought to be one of the most important regulatory proteins involved in controlling energy metabolism in most organisms.[92–92b] Phosphorylation of up to three specific serine hydroxyl groups in the thiamin-containing decarboxylase subunit (designated E1) converts the enzyme into an inactive form (Eq. 17-9). A specific phosphatase reverses the inhibition. The kinase is most active on enzymes whose core lipoyl (E2) subunits are reduced and acetylated, a condition favored by high ratios of [acetyl-CoA] to free [CoASH] and of [NADH] to [NAD^+]. Since the kinase inactivates the enzyme the effect is to decrease the pyruvate dehydrogenase action when the system becomes saturated and NADH and acetyl-CoA accumulate. Conversely, a high [pyruvate] inhibits the kinase and increases the action of the dehydrogenase complex. This system also permits various external signals to be felt. For example, insulin has a pronounced stimulatory effect on mitochondrial energy.[65,92,93] One way in which this may be accomplished is through stimulation of the pyruvate dehydrogenase phosphatase, as indicated in Eq. 17-9. A **kinase activator protein** (Eq. 17-9) may also respond to various external stimuli and may be inhibited by insulin.[92]



(17-9)

The activities of 2-oxoglutarate dehydrogenase,[94] and to a lesser extent of pyruvate and isocitrate dehydrogenases, are increased by increases in the free Ca^{2+} concentration.[87] Calcium ions stimulate the phosphatase that dephosphorylates the deactivated phosphorylated pyruvate dehydrogenase and activate the other two dehydrogenases allosterically, increasing the affinities for the substrates.[87] Phosphorylation of the NAD^+-dependent isocitrate dehydrogenase also decreases its activity. In *E. coli* the isocitrate dehydrogenase kinase and a protein phosphatase exist as a bifunctional protein able to both deactivate the dehydrogenase and restore its activity.[86] For this organism, the decrease in activity forces substrate into the glyoxylate pathway (Section J,4) instead of the citric acid cycle.

BOX 17-C USE OF ISOTOPIC TRACERS IN STUDY OF THE TRICARBOXYLIC ACID CYCLE

The first use of isotopic labeling in the study of the citric acid cycle and one of the first in the history of biochemistry was carried out by Harland G. Wood and C. H. Werkman in 1941.[a,b] The aim was to study the fermentation of glycerol by propionic acid bacteria, a process that was not obviously related to the citric acid cycle. Some succinate was also formed in

$$\text{Glycerol} \rightarrow \text{propionate}^- + H^+ + H_2O$$
$$\Delta G' = -69 \text{ kJ/mol}$$

the fermentation, and on the basis of simple measurements of the fermentation balance reported in 1938 it was suggested that CO_2 was incorporated into oxaloacetate, which was then reduced to succinate. As we now know, this is indeed an essential step in the propionic acid fermentation (Section F,3). At the time ^{14}C was not available but the mass spectrometer, newly developed by A. O. Nier, permitted the use of the stable ^{13}C as a tracer. Wood and Werkman constructed a thermal diffusion column and used it to prepare bicarbonate enriched in ^{13}C and also built a mass spectrometer. By 1941 it was established unequivocally that carbon dioxide was incorporated into succinate by the bacteria.[c]

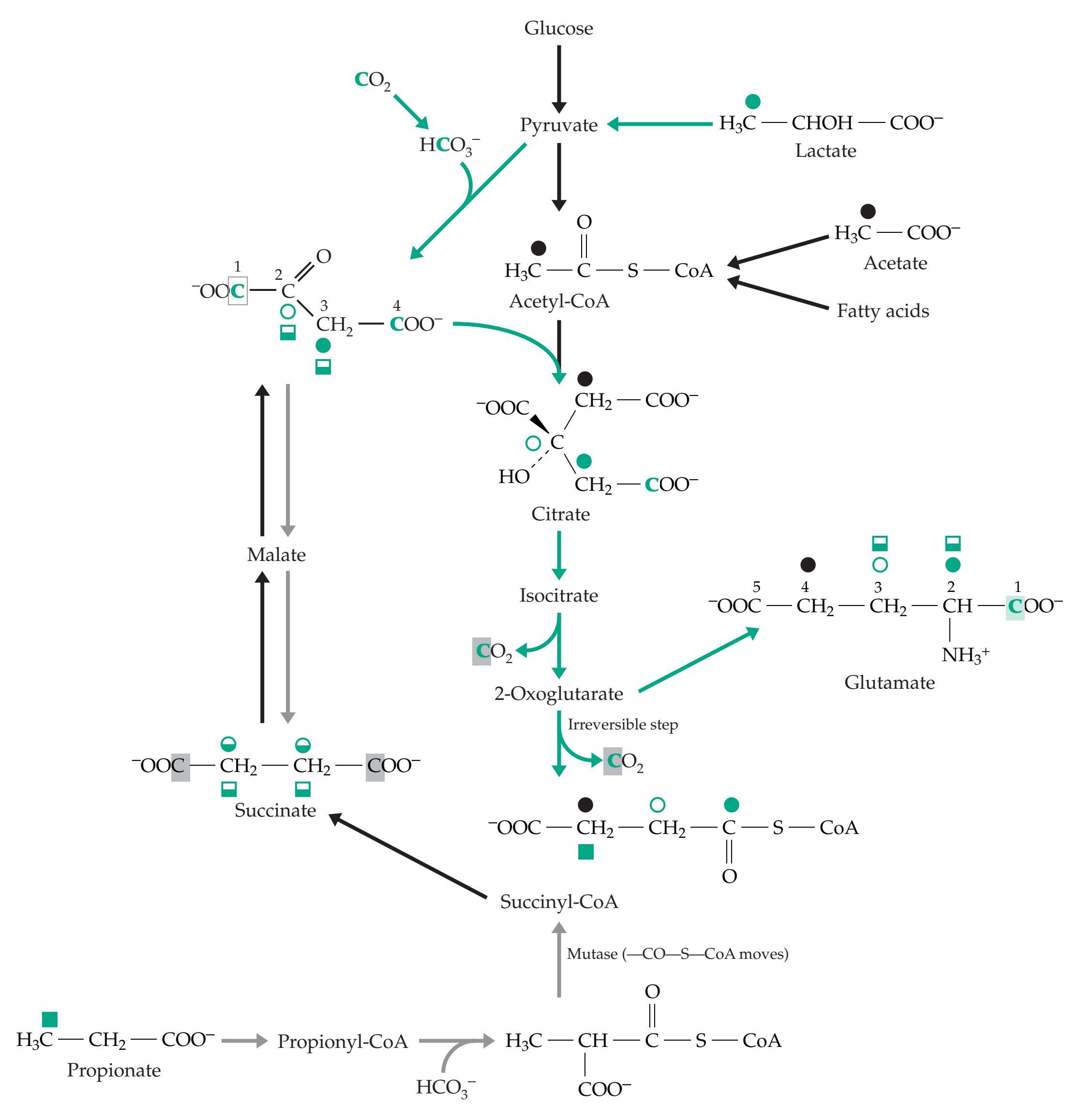
To test the idea that animal tissues could also incorporate CO_2 into succinate Wood examined the metabolism of a pigeon liver preparation to which malonate had been added to block succinate dehydrogenase (Box 10-B). Surprisingly, the accumulating succinate, which arose from oxalocetate via citrate, isocitrate, and 2-oxoglutarate (traced by green arrows in accompanying scheme), contained no ^{13}C. Soon, however, it was shown that CO_2 was incorporated into the carboxyl group of 2-oxoglutarate that is adjacent to the carbonyl group. That carboxyl is lost in conversion to succinate (Fig. 10-6) explains the lack of ^{13}C in succinate. It is of historical interest that these observations were incorrectly interpreted by many of the biochemists of the time. They agreed that *citrate could not be a member of the tricarboxylic acid cycle*. Since citrate is a symmetric compound it was thought that any ^{13}C incorporated into citrate would be present in equal amounts in both terminal carboxyl groups. This would necessarily result in incorporation of ^{13}C into succinate. It was not until 1948 that Ogston popularized the concept that by binding with substrates at three points, enzymes were capable of asymmetric attack upon symmetric substrates.[d] In other words, an enzyme could synthesize citrate with the carbon atoms from acetyl-CoA occupying one of the two $-CH_2COOH$ groups surrounding the prochiral center. Later, the complete stereochemistry of the citric acid cycle was elucidated through the use of a variety of isotopic labels (see p. 704). Some of the results are indicated by the asterisks and daggers in the structures in Fig. 10-6.

The operation of the citric acid cycle in living cells, organs, and whole animals has also been observed using NMR and mass spectroscopy with ^{13}C-containing compounds. For example, a heart can be perfused with a suitable oxygenated perfusion fluid[e] containing various ^{13}C-enriched substrates such as [U-^{13}C]fatty acids, [2-^{13}C]acetate, [3-^{13}C]L-lactate, or [2,3-^{13}C]propionate.[e – k] NMR spectroscopy allows direct and repeated observation of the ^{13}C nuclei from a given substrate and its entry into a variety of metabolic pathways. Because of the high dispersion of chemical shift values for ^{13}C the NMR resonance for the isotope can be seen at each position within a single compound.

A compound that is especially easy to observe is glutamate. This amino acid, most of which is found in the cytoplasm, is nevertheless in relatively rapid equilibrium with 2-oxoglutarate of the citric acid cycle in the mitochondria. The accompanying scheme shows where isotopic carbon from certain compounds will be located *when it first enters* the citric acid cycle and traces some of the labels into glutamate. For example, uniformly enriched fatty acids will introduce label into the two atoms of the *pro-S* arm of citrate and into 4- and 5-positions of glutamate whereas [2-^{13}C]acetate will introduce label only into the C4 position as marked by ● in the scheme. In the NMR spectrum a singlet resonance at 32.4 ppm will be observed. However, as successive turns of the citric acid cycle occur the isotope will appear in increasing amounts in the adjacent 3-position of glutamate. They will be recognized readily by the appearance of a multiplet. The initial singlet will be flanked by a pair of peaks that arise from spin–spin coupling with the adjacent 3-^{13}C of the [2,3-^{13}C]isotopomer (see accompanying figure). After longer periods of time the central resonance will weaken and the outer pair strengthen as the recycling occurs.

Metabolism with [U-^{13}C]fatty acids gives a labeling pattern similar to that with [2,3-^{13}C]acetate and it has been deduced that heart muscle normally metabolizes principally fatty acids for energy. What will happen to the glutamate C4 resonance if [3-^{13}C]lactate is added to the perfusion solution? It will enter both acetyl-CoA and oxaloacetate as indicated by ● in the following scheme. That will also intro-duce ^{13}C at C3 of glutamate. By looking at spectra at short times the relative amounts of lactate being oxidized via the cycle and that being converted

BOX 17-C (continued)

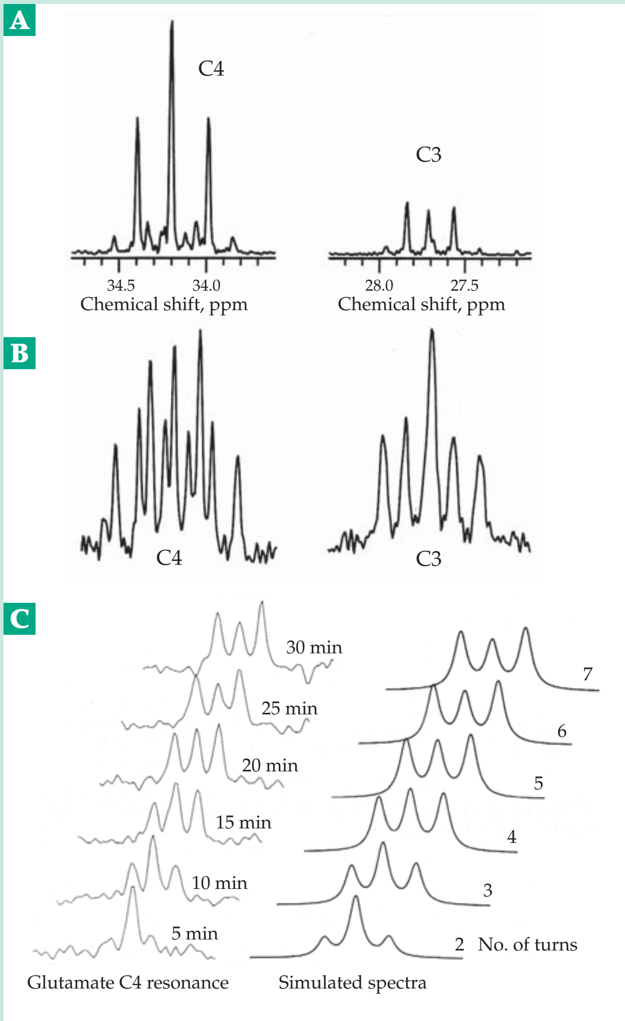

Glucose
CO₂
HCO₃⁻
Pyruvate
H₃C — CHOH — COO⁻
Lactate
O
H₃C — C — S — CoA
Acetyl-CoA
H₃C — COO⁻
Acetate
Fatty acids
1
2
O
⁻OOC — C
3
4
CH₂ — COO⁻
⁻OOC
CH₂ — COO⁻
C
HO
CH₂ — COO⁻
Citrate
Malate
Isocitrate
CO₂
2-Oxoglutarate
5
4
3
2
1
⁻OOC — CH₂ — CH₂ — CH — COO⁻
NH₃⁺
Glutamate
Irreversible step
CO₂
⁻OOC — CH₂ — CH₂ — COO⁻
Succinate
⁻OOC — CH₂ — CH₂ — C — S — CoA
O
Succinyl-CoA
Mutase (—CO—S—CoA moves)
H₃C — CH₂ — COO⁻
Propionate
Propionyl-CoA
HCO₃⁻
O
H₃C — CH — C — S — CoA
COO⁻

biosynthetically (anaplerotically) to glutamate can be estimated. There is a complication that has long been recognized. Oxaloacetate can be converted by exchange processes to succinate. Since succinate is symmetric the effect is to put 50% of the label into each of the central atoms of succinate (◒ in scheme). The exchange will then transfer label back into the C2 position of oxaloacetate (○) and through citric acid cycle reactions into C3 of glutamate. Now the C4 NMR resonance will contain an additional pair of peaks arising from spin–spin coupling with C2 but which will have a different coupling constant than that for coupling to C3.

If uniformly labeled [U-^{13}C]acetate is introduced the additional isotopomers, [3,4 - ^{13}C]glutamate and [3,4,5 - ^{13}C]glutamate, will be formed as will others with ^{13}C in the C1 and C5 positions but which will not affect the C4 resonance. A total of nine lines will be seen as illustrated in curve *a* of the accompanying figure. We see that the multiplet patterns arising from mass isotopomers are complex, but they can be predicted accurately with a computer program.[f] Isotopomers of succinate have also been analyzed.[g]

It is also of interest to introduce ^{13}C from propionate labeled in various positions. One of these is illustrated in the scheme. In this case the appearance of multiplets arising from [3,4-^{13}C] glutamate verifies the existence of end-to-end scrambling of the isotope in succinate. However, is the scrambling complete or are some molecules

BOX 17-C USE OF ISOTOPIC TRACERS IN STUDY OF THE TRICARBOXYLIC ACID CYCLE (cont.)



A. ^{13}C-NMR spectrum of extracts of Langendorff-perfused rat hearts perfused for 5 min with [1,2^{13}C] acetate, [3^{13}C]-lactate and glucose. Only the glutamate C4 (left) and C3 (right) resonances are shown. B. Spectrum after perfusion for 30 min. From Malloy *et al*.[r] C. The glutamate C4 resonance of an intact Langendorff-perfused rat heart supplied with 2 mM [2-^{13}C]acetate showing evolution of the multiplet as a function of time after introducing the label. The right panel shows glutamate C4 resonances generated by a computer simulation after turnover of citric acid cycle pools the indicated number of time. From Jeffrey *et al*.[f]

efficiently "channeled" through enzyme–enzyme complexes in such a way as to avoid scrambling? As shown in the scheme, full scrambling would give equal labeling of C2 and C3 of oxaloacetate and of glutamate. Experimentally greater labeling was seen at C3 than at C2 during the first few turns of the cycle suggesting that some channeling does occur.[e]

Isotopomer analysis can also be conducted by mass spectroscopy, which is more sensitive than NMR, using ^{13}C[h,k,l] or ^{2}H labeling.[j] Making use of a technique like that employed by Knoop (Box 10-A), a "chemical biopsy" can be performed on animals or on human beings, who may ingest gram quantities of sodium phenylacetate without harm. The phenylacetate is converted to an amide with glutamine (phenylacetylglutamine) which is excreted in the urine, from which it can easily be recovered for analysis.[l–n] This provides a non-invasive way of studying the operation of the citric acid cycle in the human body. Direct measurement on animal brains[o,p] and on human limbs or brain has also been accomplished by NMR spectroscopy[q] and may become more routine as instrumentation is improved.

---

[a] Wood, H. G. (1972) in *The Molecular Basis of Biological Transport* (Woessner, J. F., and Huijing, F., eds), pp. 1–54, Academic Press, New York

[b] Krampitz, L. O. (1988) *Trends Biochem. Sci.* **13**, 152–155

[c] Wood, H. G., Werkman, C. H., Hemingway, A., and Nier, A. O. (1941) *J. Biol. Chem.* **139**, 377–381

[d] Ogston, A. G. (1948) *Nature (London)* **162**, 936

[e] Sherry, A. D., Sumegi, B., Miller, B., Cottam, G. L., Gavva, S., Jones, J. G., and Malloy, C. R. (1994) *Biochemistry* **33**, 6268–6275

[f] Jeffrey, F. M. H., Rajagopal, A., Malloy, C. R., and Sherry, A. D. (1991) *Trends Biochem. Sci.* **16**, 5–10

[g] Jones, J. G., Sherry, A. D., Jeffrey, F. M. H., Storey, C. J., and Malloy, C. R. (1993) *Biochemistry* **32**, 12240–12244

[h] Des Rosiers, C., Di Donato, L., Comte, B., Laplante, A., Marcoux, C., David, F., Fernandez, C. A., and Brunengraber, H. (1995) *J. Biol. Chem.* **270**, 10027–10036

[i] Sherry, A. D., and Malloy, C. R. (1996) *Cell Biochem. Funct.* **14**, 259–268

[j] Yudkoff, M., Nelson, D., Daikhin, Y., and Erecinska, M. (1994) *J. Biol. Chem.* **269**, 27414–27420

[k] Beylot, M., Soloviev, M. V., David, F., Landau, B. R., and Brunengraber, H. (1995) *J. Biol. Chem.* **270**, 1509–1514

[l] Di Donato, L., Des Rosiers, C., Montgomery, J. A., David, F., Garneau, M., and Brunengraber, H. (1993) *J. Biol. Chem.* **268**, 4170–4180

[m] Magnusson, I., Schumann, W. C., Bartsch, G. E., Chandramouli, V., Kumaran, K., Wahren, J., and Landau, B. R. (1991) *J. Biol. Chem.* **266**, 6975–6984

[n] Chervitz, S. A., and Falke, J. J. (1995) *J. Biol. Chem.* **270**, 24043–24053

[o] Hyder, F., Chase, J. R., Behar, K. L., Mason, G. F., Siddeek, M., Rothman, D. L., and Shulman, R. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7612–7617

[p] Cerdan, S., Künnecke, B., and Seelig, J. (1990) *J. Biol. Chem.* **265**, 12916–12926

[q] Rothman, D. L., Novotny, E. J., Shulman, G. I., Howseman, A. M., Petroff, O. A. C., Mason, G., Nixon, T., Hanstock, C. C., Prichard, J. W., and Shulman, R. G. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9603–9606

[r] Malloy, C. R., Thompson, J. R., Jeffery, F. M. H., and Sherry, A. D. (1990) *Biochemistry* **29**, 6756–6761

Acting to counteract any drop in ATP level, accumulating ADP acts as a positive effector for isocitrate dehydrogenases.

Another way in which the phosphorylation state of the adenylate system can regulate the cycle depends upon the need for GDP in step *f* of the cycle (Fig. 17-4). Within mitochondria, GTP is used largely to reconvert AMP to ADP. Consequently, formation of GDP is promoted by AMP, a compound that arises in mitochondria from the utilization of ATP for activation of fatty acids (Eq. 13-44) and activation of amino acids for protein synthesis (Eq. 17-36).

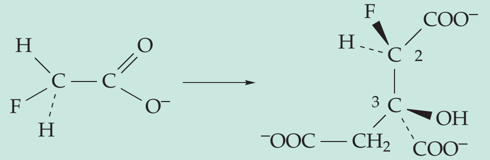
In *E. coli* and some other bacteria ATP does not inhibit citrate synthase but NADH does; the control is via the redox potential of the NAD$^+$ system rather than by the level of phosphorylation of the adenine nucleotide system.[95] Succinic dehydrogenase may be regulated by the redox state of ubiquinone (Chapter 15). Another mechanism of regulation may be the formation of specific protein–protein complexes between enzymes catalyzing reactions of the cycle.[96–97a] This may permit one enzyme to efficiently have a product of its action transferred to the enzyme catalyzing the next step in the cycle.

5. Catabolism of Intermediates of the Citric Acid Cycle

Acetyl-CoA is the only substrate that can be completely oxidized to CO$_2$ by the reaction of the citric acid cycle alone. Nevertheless, cells must sometimes

BOX 17-D FLUOROACETATE AND "LETHAL SYNTHESIS"

Among the most deadly of simple compounds is sodium fluoroacetate. The LD$_{50}$ (the dose lethal for 50% of animals receiving it) is only 0.2 mg/kg for rats, over tenfold less than that of the nerve poison diisopropylphosphofluoridate (Chapter 12).[a,b] Popular, but controversial, as the rodent poison "1080," fluoroacetate is also found in the leaves of several poisonous plants in Africa, Australia, and South America. Surprisingly, difluoroacetate HCF_2-COO^- is nontoxic and biochemical studies reveal that monofluoroacetate has no toxic effect on cells until it is converted metabolically in a "lethal synthesis" to 2*R*,3*R*-2-fluorocitrate, which is a competitive inhibitor of aconitase (aconitate hydratase, Eq. 13-17).[b–g] This fact was difficult to understand since citrate formed by the reaction of fluorooxaloacetate and acetyl-CoA has only weak inhibitory activity toward the same enzyme. Yet, it is the fluorocitrate formed from fluorooxaloacetate that contains a fluorine atom at a site that is attacked by aconitase in the citric acid cycle.



2*R*,3*R* 2-Fluoroacetate

The small van der Waals radius of fluorine (0.135 nm), comparable to that of hydrogen (0.12 nm), is often cited as the basis for the ability of fluoro compounds to "deceive" enzymes. However, the high electronegativity and ability to enter into hydrogen bonds may make F more comparable to –OH in metabolic effects. In the case of fluorocitrate it was proposed that the inhibitory isomer binds in the "wrong way" to aconitase in such a manner that the fluorine atom is coordinated with the ferric ion at the catalytic center.[c] However, 2*R*,3*R*-2-fluorocitrate is a simple competitive inhibitor of aconitase but an irreversible poison. It is especially toxic to nerves and also appears to affect mitochondrial membranes. Therefore, this poison may affect some other target, such as a citrate transporter.[d] Fluoroacetate is only one of many known naturally occurring fluorine compounds.[c]

Another example of lethal synthesis is seen in the use of 5-fluorouracil in cancer therapy (Box 15-E). In this compound and in many other fluorine-containing inhibitors the F atom replaces the H atom that is normally removed as H$^+$ in the enzymatic reaction. The corresponding F$^+$ cannot be formed.[h] Because of the high electronegativity of fluorine a C–F bond is polarized: $C^{\delta+}-F^{\delta-}$. This may have very large effects on reactivity at adjacent positions. For example, the reactivity of 2-fluoroglycosyl groups toward glycosyl transfer is decreased by several orders of magnitude (p. 597).

---

[a] Gibble, G. W. (1973) *J. Chem. Educ.* **50**, 460–462

[b] Elliott, K., and Birch, J., eds. (1972) *Carbon–Fluorine Compounds*, Elsevier, Amsterdam

[c] Glusker, J. P. (1971) in *The Enzymes*, 3rd ed., Vol. 5 (Boyer, P. D., ed), pp. 413–439, Academic Press, New York

[d] Kun, E. (1976) in *Biochemistry Involving Carbon–Fluorine Bonds* (Filler, R., ed), pp. 1–22, American Chemical Society, Washington, DC

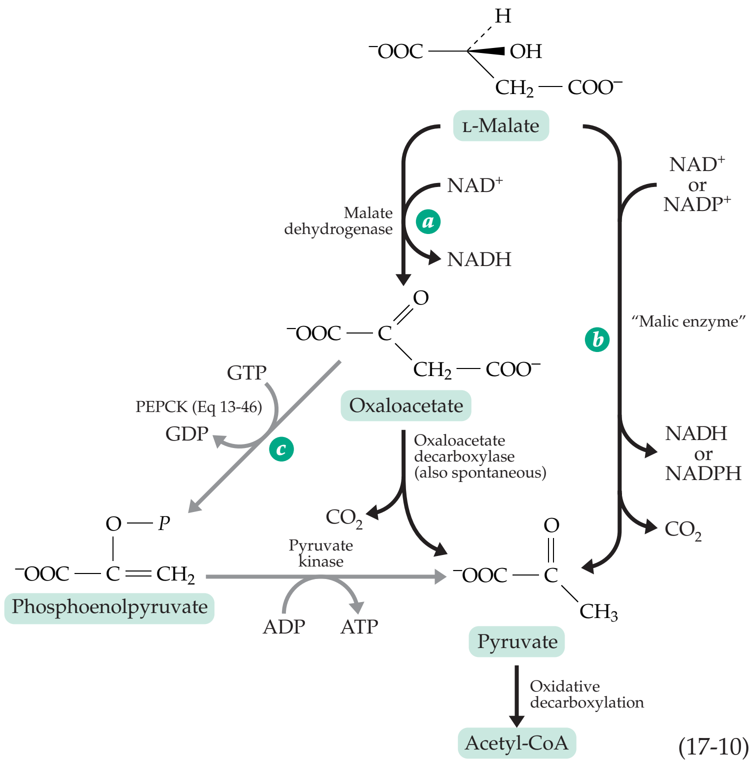
[e] Marletta, M. A., Srere, P. A., and Walsh, C. (1981) *Biochemistry* **20**, 3719–3723

[f] Rokita, S. E., and Walsh, C. T. (1983) *Biochemistry* **22**, 2821–2828

[g] Peters, R. A. (1957) *Adv. Enzymol.* **18**, 113–159

[h] Abeles, R. H., and Alston, T. A. (1990) *J. Biol. Chem.* **265**, 16705–16708

oxidize large amounts of one of the compounds found in the citric acid cycle to CO_2.[98,99] For example, bacteria subsisting on succinate as a carbon source must oxidize it for energy as well as convert some of it to carbohydrates, lipids, and other materials. Complete combustion of *any citric acid cycle intermediate* can be accomplished by conversion to malate followed by oxidation of malate to oxaloacetate (Eq. 17-10, step *a*) and decarboxylation (β cleavage) to pyruvate, or (Eq. 17-10, step *b*) oxidation and decarboxylation of malate by the **malic enzyme** (Eq. 13-45) without free oxaloacetate as an intermediate. Pathway *b* is probably the most important. It is catalyzed by two different malic enzymes present in animal mitochondria. One is specific for $NADP^+$ while the other reacts with NAD^+ as well.[100,101] They both have complex regulatory properties. For example, the less specific NAD^+-utilizing enzyme is allosterically inhibited by ATP but is activated by fumarate, succinate, or isocitrate.[100] Thus, accumulation of citric acid cycle intermediates "turns on" the malic enzyme, allowing the excess to leave the cycle and reenter as acetyl groups. Since the Michaelis constant for malate is high, this will not happen unless malate accumulates, signaling a need for acetyl-CoA. The $NADP^+$-dependent enzyme is activated by a high concentration of free CoA and is inhibited by NADH. Perhaps when glycolysis becomes slow the free CoA level rises and turns on malate oxidation.[101] On the other hand, rapid glycolysis increases the NADH concentration which inhibits the malic enzyme. The result is a buildup of the oxaloacetate concentration and an increase in activity of the citric acid cycle. The malic enzymes are also present in the cytoplasm, where one of them functions as part of an NADPH-generating cycle (Eq. 17-46).



(17-10)

D. Oxidative Pathways Related to the Citric Acid Cycle

In this section we will consider some other catalytic cycles as well as some noncyclic pathways of oxidation of one- and two-carbon substrates that are utilized by microorganisms.

1. The γ-Aminobutyrate Cycle

A modification of the citric acid cycle which involves glutamate and gamma (γ) aminobutyrate (GABA) has an important function in the brain (Fig. 17-5). Both glutamate and γ-aminobutyrate occur in high concentrations in brain (10 and 0.8 mM, respectively). Both are important neurotransmitters, γ-aminobutyrate being a principal neuronal inhibitory substance[102,103] (Chapter 30). In the γ-aminobutyrate cycle acetyl-CoA and oxaloacetate are converted into citrate (step *a*) in the usual way and the citrate is then converted into 2-oxoglutarate. The latter is transformed to L-glutamate either by direct amination (*b*) or by transamination (*c*), the amino donor being γ-aminobutyrate.

γ-Aminobutyrate is formed by decarboxylation of glutamate (Fig. 17-5, step *d*)[104] and is catabolized via transamination (step *e*)[105] to succinic semialdehyde, which is oxidized to succinate[106] and oxaloacetate. The two transamination steps in the pathways may be linked, as indicated in Fig. 17-5, to form a complete cycle that parallels the citric acid cycle but in which 2-oxoglutarate is oxidized to succinate via glutamate and γ-aminobutyrate. No thiamin diphosphate is required, but 2-oxoglutarate is reductively aminated to glutamate. The cycle is sometimes called the **γ-aminobutyrate shunt**, and it plays a significant role in the overall oxidative processes of brain tissue. This pathway is also prominent in green plants.[107–109] For example, under anaerobic conditions the radish *Raphanus sativus* accumulates large amounts of γ-aminobutyrate.[110] Most animal tissues contain very little γ-aminobutyrate, although it has been found in the oviducts of rats at concentrations that exceed those in the brain.[111]

2. The Dicarboxylic Acid Cycle

Some bacteria can subsist solely on glycolate, glycine, or oxalate, all of which

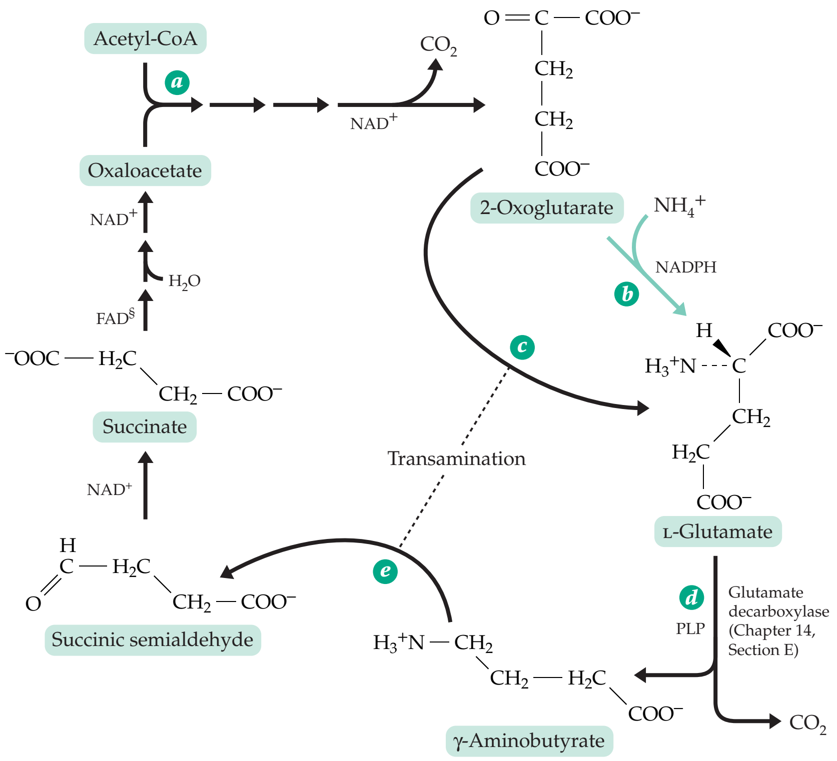

Acetyl-CoA
a
CO2
NAD+
O=C—COO−
CH2
CH2
COO−
2-Oxoglutarate
Oxaloacetate
NAD+
H2O
FAD§
−OOC—H2C
CH2—COO−
Succinate
NAD+
NH4+
NADPH
b
c
H
COO−
H3+N---C
CH2
H2C
COO−
L-Glutamate
Transamination
H
C—H2C
O
CH2—COO−
Succinic semialdehyde
e
H3+N—CH2
CH2—H2C
COO−
γ-Aminobutyrate
d
PLP
Glutamate decarboxylase (Chapter 14, Section E)
CO2

Figure 17-5 Reactions of the γ-aminobutyrate (GABA) cycle.

are converted to glyoxylate (Eq. 17-11). Glyoxylate is oxidized to CO_2 and water to provide energy to the bacteria and is also utilized for biosynthetic purposes. The energy-yielding process is found in the **dicarboxylic acid cycle** (Fig. 17-6), which catalyzes the complete oxidation of glyoxylate. Four hydrogen atoms are removed with generation of two molecules of NADH which can be oxidized by the respiratory chain to provide energy.[112,113] In the dicarboxylic acid cycle glyoxylate is the principal substrate and acetyl-CoA is the regenerating substrate rather than the principal substrate as it is for the citric acid cycle.

The logic of the dicarboxylic acid cycle is simple. Acetyl-CoA contains a potentially free carboxyl group. After the acetyl group of acetyl-CoA has been condensed with glyoxylate and the resulting hydroxyl group has been oxidized, the free carboxyl group appears in oxaloacetate in a position β to the carbonyl group. The carboxyl

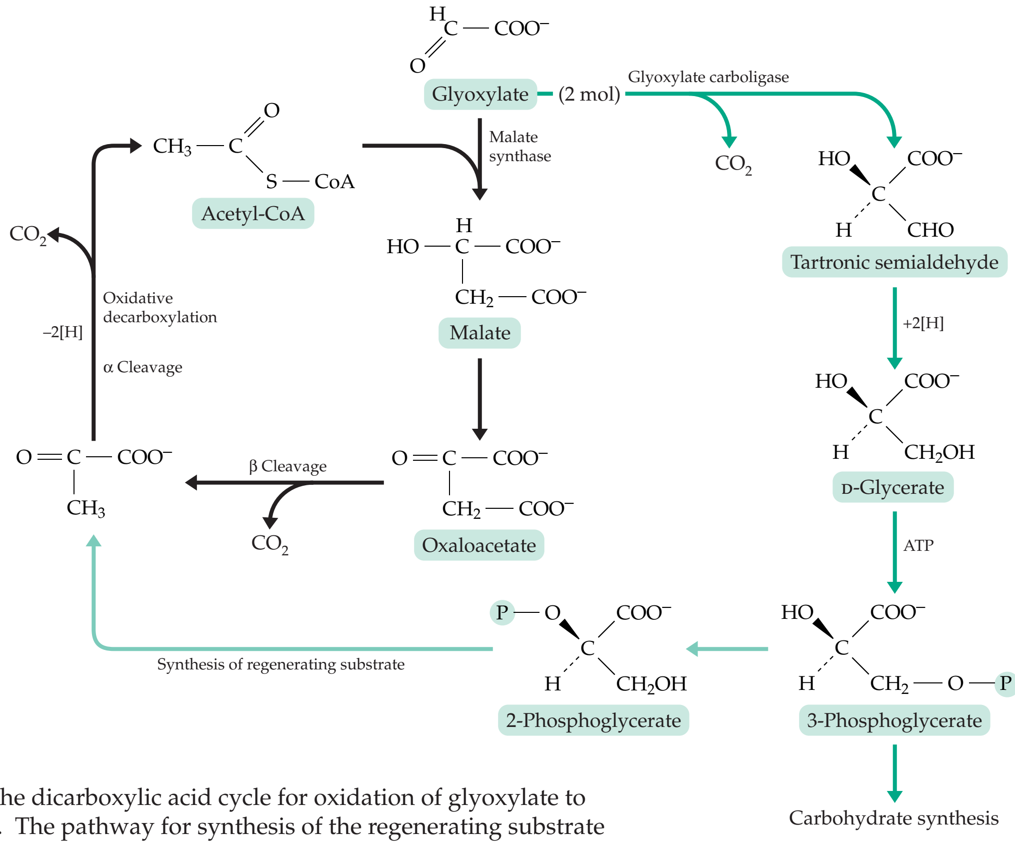
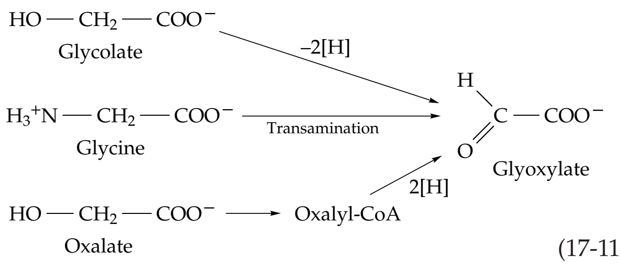

H
C—COO−
O
Glyoxylate
(2 mol)
Glyoxylate carboligase
CO2
HO
COO−
C
H
CHO
Tartronic semialdehyde
+2[H]
HO
COO−
C
H
CH2OH
D-Glycerate
ATP
HO
COO−
C
H
CH2—O—P
3-Phosphoglycerate
Carbohydrate synthesis
P—O
COO−
C
H
CH2OH
2-Phosphoglycerate
Synthesis of regenerating substrate
CH3—C
O
S—CoA
Acetyl-CoA
Malate synthase
H
HO—C—COO−
CH2—COO−
Malate
O=C—COO−
CH2—COO−
Oxaloacetate
β Cleavage
CO2
O=C—COO−
CH3
CO2
−2[H]
Oxidative decarboxylation
α Cleavage

Figure 17-6 The dicarboxylic acid cycle for oxidation of glyoxylate to carbon dioxide. The pathway for synthesis of the regenerating substrate is indicated by green lines. This pathway is also needed for synthesis of carbohydrates and all other cell constituents.



(17-11)

donated by the glyoxylate is still in the α position. A consecutive β cleavage and an oxidative α cleavage release the two carboxyl groups as carbon dioxide to reform the regenerating substrate. The cycle is simple and efficient. Like the citric acid cycle, it depends upon thiamin diphosphate, without which the α cleavage would be impossible. Comparing the citric acid cycle (Fig. 17-2) with the simpler dicarboxylic acid cycle, we see that in the former the initial condensation product citrate contains a hydroxyl group attached to a tertiary carbon atom. With no adjacent hydrogen it is impossible to oxidize it directly to the carbonyl group which is essential for subsequent chain cleavage; hence the dependence on aconitase to shift the OH to an adjacent carbon. Both cycles involve oxidation of a hydroxy acid to a ketone followed by β cleavage and oxidative α cleavage. In the citric acid cycle additional oxidation steps are needed to convert succinate back to oxaloacetate, corresponding to the fact that the citric acid cycle deals with a more reduced substrate than does the dicarboxylic acid cycle.

The synthetic pathway for the regenerating substrate of the dicarboxylic acid cycle is quite complex. Two molecules of glyoxylate undergo α condensation with decarboxylation by glyoxylate carboligase[114] (see also Chapter 14, Section D) to form **tartronic semialdehyde**. The latter is reduced to D-glycerate, which is phosphorylated to 3-phosphoglycerate and 2-phosphoglycerate. Since the phosphoglycerates are carbohydrate precursors, this **glycerate pathway** provides the organisms with a means for synthesis of carbohydrates and other complex materials from glyoxylate alone. At the same time, 2-phosphoglycerate can be converted to pyruvate and the pyruvate, by oxidative decarboxylation, to the regenerating substrate acetyl-CoA.

E. Catabolism of Sugars

In most sugars each carbon atom bears an oxygen atom which facilitates chemical attack by oxidation at any point in the carbon chain. Every sugar contains a potentially free aldehyde or ketone group, and the carbonyl function can be moved readily to adjacent positions by isomerases. Consequently, aldol cleavage is also possible at many points. For these reasons, the metabolism of carbohydrates is complex and varied.

A sugar chain can be cut in several places giving rise to a variety of metabolic pathways. However, in the energy economy of most organisms, including human beings, the **Embden–Meyerhof–Parnas** or **glycolysis pathway** by which hexoses are converted to pyruvate (Fig. 17-7) stands out above all others. We have already considered this pathway, which is also outlined in Figs. 10-2 and 10-3. Some history and additional important details follow.

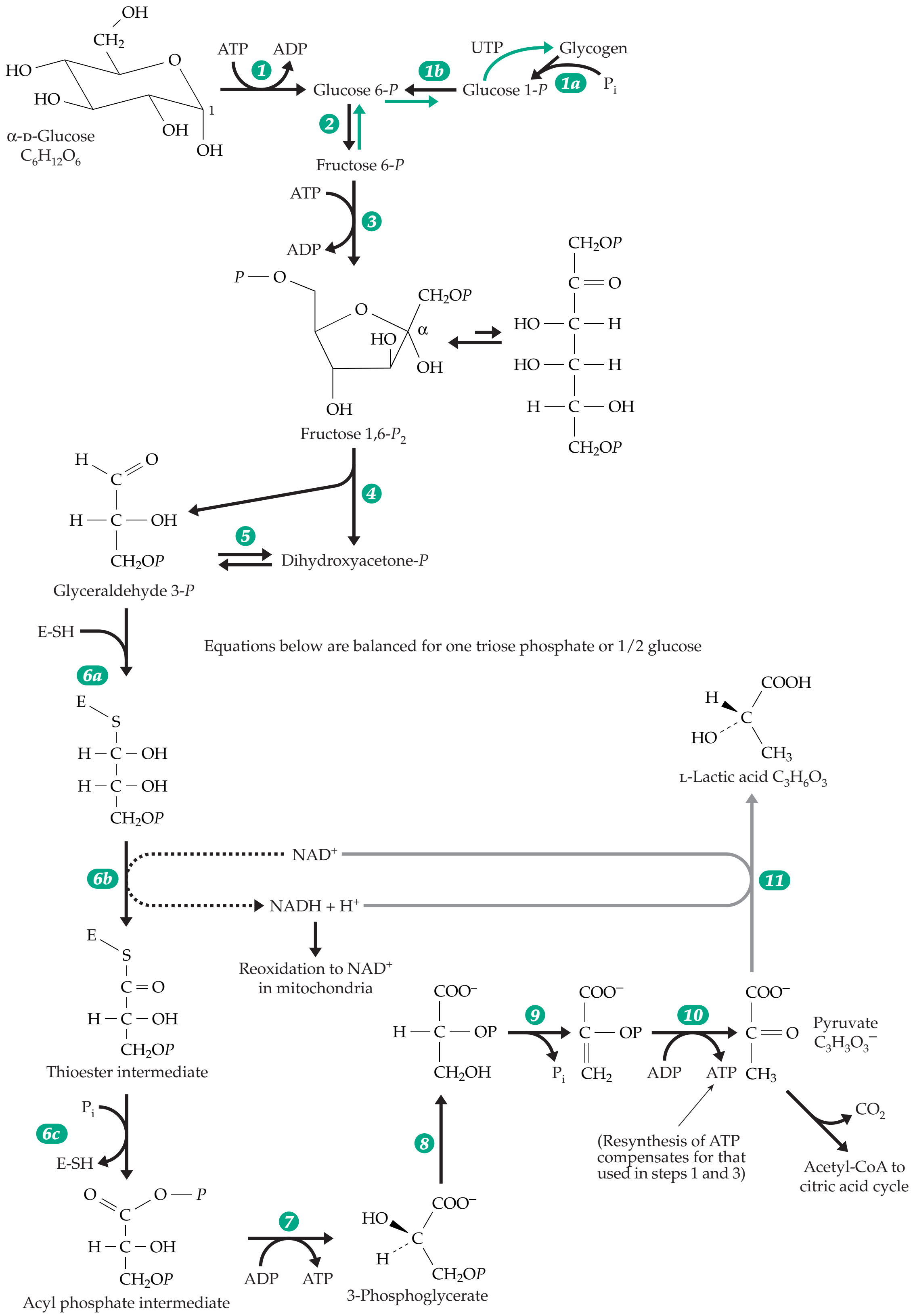
1. The Glycolysis Pathway

The discovery of glycolysis followed directly the early observations of Buchner and of Harden and Young on fermentation of sugar by yeast juice (p. 767). Another line of research, the study of muscle, soon converged with the investigations of alcoholic fermentation. Physiologists were interested in the process by which an isolated muscle could obtain energy for contraction in the absence of oxygen. It was shown by A. V. Hill that glycogen was converted to lactate to supply the energy, and Otto Meyerhof later demonstrated that the chemical reactions were related to those of alcoholic fermentation. The establishment of the structures and functions of the pyridine nucleotides in 1934 (Chapter 15) coincided with important studies by G. Embden in Frankfurt and of J. K. Parnas in Poland. The sequence of reactions in glycolysis soon became clear. All of the 15 enzymes catalyzing the individual steps in the sequence have been isolated and crystallized and are being studied in detail.[115]

Formation of pyruvate. The conversion of glucose to pyruvate requires ten enzymes (Fig. 17-7), and the sequence can be divided into four stages: preparation for chain cleavage (reactions *1–3*), cleavage and equilibration of triose phosphates (reactions *4* and *5*), oxidative generation of ATP (reactions *6* and *7*), and conversion of 3-phosphoglycerate to pyruvate (reactions *8–10*).

In preparation for chain cleavage, free glucose is phosphorylated to glucose 6-phosphate by ATP under the action of hexokinase (reaction *1*). Glucose 6-phosphate can also arise by cleavage of a glucosyl unit from glycogen by the consecutive action of glycogen phosphorylase (reaction 1*a*) and phosphoglucomutase, which transfers a phospho group from the oxygen at C-1 to that at C-6 (reaction 1*b*) (see also Eq. 12-39 and associated discussion of the mechanism of this enzyme). Why do cells attach phospho groups to sugars to initiate metabolism of the sugars? Four reasons can be given:

(a) The phospho group constitutes an electrically charged handle for binding the sugar phosphate to enzymes.
(b) There is a kinetic advantage in initiating a reaction sequence with a highly irreversible reaction


α-D-Glucose
C6H12O6
ATP
ADP
1
Glucose 6-P
1b
UTP
Glycogen
Glucose 1-P
1a
Pi
2
Fructose 6-P
ATP
3
ADP
Fructose 1,6-P2
α
4
Glyceraldehyde 3-P
5
Dihydroxyacetone-P
E-SH
6a
6b
NAD+
NADH + H+
11
Reoxidation to NAD+
in mitochondria
L-Lactic acid C3H6O3
Thioester intermediate
P i
6c
E-SH
Acyl phosphate intermediate
7
ADP
ATP
3-Phosphoglycerate
8
9
Pi
10
ADP
ATP
Pyruvate
C3H3O3−
CO2
Acetyl-CoA to
citric acid cycle

Equations below are balanced for one triose phosphate or 1/2 glucose

(Resynthesis of ATP compensates for that used in steps 1 and 3)

Figure 17-7 Outline of the glycolysis pathway by which hexoses are broken down to pyruvate. The ten enzymes needed to convert D-glucose to pyruvate are numbered. The pathway from glycogen using glycogen phosphorylase is also included, as is the reduction of pyruvate to lactate (step *11*). Steps *6a–7*, which are involved in ATP synthesis via thioester and acyl phosphate intermediates, are emphasized. See also Figures 10-2 and 10-3, which contain some additional information.

such as the phosphorylation of glucose.

(c) Phosphate esters are unable to diffuse out of cells easily and be lost.

(d) There is at least a possibility that the phospho groups may function in catalysis.

Reaction *2* of Fig. 17-7 is a simple isomerization that moves the carbonyl group to C-2 so that β cleavage to two three-carbon fragments can occur. Before cleavage a second phosphorylation (reaction *3*) takes place to form fructose 1,6-bisphosphate. This ensures that when fructose bisphosphate is cleaved by aldolase each of the two halves will have a phosphate handle. This second priming reaction (reaction *3*) is the first step in the series that is unique to glycolysis. The catalyst for the reaction, **phosphofructokinase**, is carefully controlled, as discussed in Chapter 11 (see Fig. 11-2).

Fructose bisphosphate is cleaved by action of an aldolase (reaction *4*) to give glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. These two triose phosphates are then equilibrated by triose phosphate isomerase (reaction *5*; see also Chapter 13). As a result, both halves of the hexose can be metabolized further via glyceraldehyde 3-*P* to pyruvate. The oxidation of glyceraldehyde 3-*P* to the corresponding carboxylic acid, 3-phosphoglyceric acid (Fig. 17-7, reactions *6* and *7*), is coupled to synthesis of a molecule of ATP from ADP and P_i. This means that two molecules of ATP are formed per hexose cleaved, and that two molecules of NAD^+ are converted to NADH in the process.

The conversion of 3-phosphoglycerate to pyruvate begins with transfer of a phospho group from the C-3 to the C-2 oxygen (reaction *8*) and is followed by dehydration through an α, β elimination catalyzed by **enolase** (reaction *9*). The product, phosphoenolpyruvate (PEP), has a high group transfer potential. Its phospho group can be transferred easily to ADP via the action of the enzyme **pyruvate kinase**, to leave the enol of pyruvic acid which is spontaneously converted to the much more stable pyruvate ion (see Eq. 7-59). Because two molecules of PEP are formed from each glucose molecule, the process provides for the recovery of the two molecules of ATP that were expended in the initial formation of fructose 1,6-bisphosphate from glucose. Several isoenzyme forms exist in mammals. Most of these are allosterically activated by fructose 1,6-bisphosphate.[115a,b] However, the enzyme from trypanosomes is activated by fructose 2,6-P_2.[115c]

The further metabolism of pyruvate. In the aerobic metabolism that is characteristic of most tissues of our bodies, pyruvate is oxidatively decarboxylated to acetyl-CoA, which can then be completely oxidized in the citric acid cycle (Fig. 17-4). The NADH produced in reaction *6* of Fig. 17-7, as well as in the oxidative decarboxylation of pyruvate and in subsequent reactions of the citric acid cycle, is reoxidized in the electron transport chain of the mitochondria as described in detail in Chapter 18 (see Fig. 18-5). An important alternative fate of pyruvate is to enter into fermentation reactions. For example, the enzyme lactate dehydrogenase (Fig. 17-7, reaction *11*) catalyzes reduction by NADH of pyruvate to L-lactate, or, for some bacteria, to D-lactate. This reaction can be coupled to the NADH-producing reaction *6* to give a balanced process by which glucose is fermented to lactic acid in the absence of oxygen (see also Eq. 10-3). In a similar process, yeast cells decarboxylate pyruvate (α cleavage) to acetaldehyde which is reduced to ethanol using the NADH produced in reaction *6*. These fermentation reactions are summarized in Fig. 10-3 and, along with many others, are discussed further in Section F of this chapter.

2. Generation of ATP by Substrate Oxidation

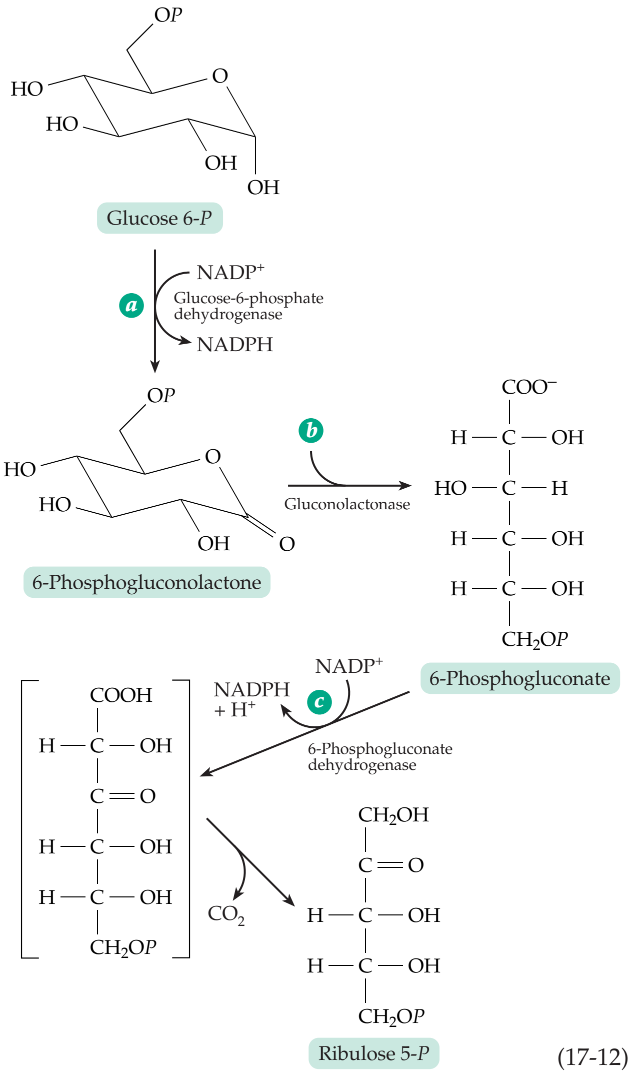
The formation of ATP from ADP and P_i is a vital process for all cells. It is usually referred to as "phosphorylation" and includes **oxidative phosphorylation** associated with the passage of electrons through an electron transport chain—usually in mitochondria; **photosynthetic phosphorylation**, a similar process occurring in chloroplasts under the influence of light; and **substrate-level phosphorylation**. Only the last is fully understood chemically. The dehydrogenation of glyceraldehyde 3-*P* and the accompanying ATP formation (reactions *6* and *7*, Fig. 17-7; Fig. 15-6) is the best known example of substrate-level phosphorylation and is tremendously important for yeasts and other microorganisms that live anaerobically. They depend upon this one reaction for their entire supply of energy. The conversion of glucose either to lactate or to ethanol and CO_2 is accompanied by a net synthesis of only two molecules of ATP and it is most logical to view these as arising from oxidation of glyceraldehyde 3-*P*. The formation of ATP from PEP and ADP in reaction *10* of Fig. 17-7 can be regarded as the recapturing of ATP "spent" in the priming reactions of steps *1* and *3*. With a gain of only two molecules of ATP for each molecule of hexose fermented, it is not surprising that yeast must ferment enormous amounts of sugar to meet its energy needs.

Each glucose unit of glycogen stored in our bodies can be converted to pyruvate with an apparent net gain of *three* molecules of ATP. However, two molecules of ATP were needed for the initial synthesis of each hexose unit of glycogen (Fig. 12-2). Therefore, the overall net yield for fermentation of stored polysaccharide is still only two ATP per hexose. The fermentation of glycogen accounts for the very rapid generation of lactic acid during intense muscular activity. However, in most circumstances within aerobic tissues reoxidation of NADH occurs via the electron transport chain of mitochondria with a much higher yield of ATP. Substrate-level phosphorylation can also follow oxidative decarboxylation of an α-oxoacid. For example, in the citric acid cycle GTP is formed following oxidative decarboxylation of 2-oxoglutarate (Fig. 17-4, steps *e* and *f*).

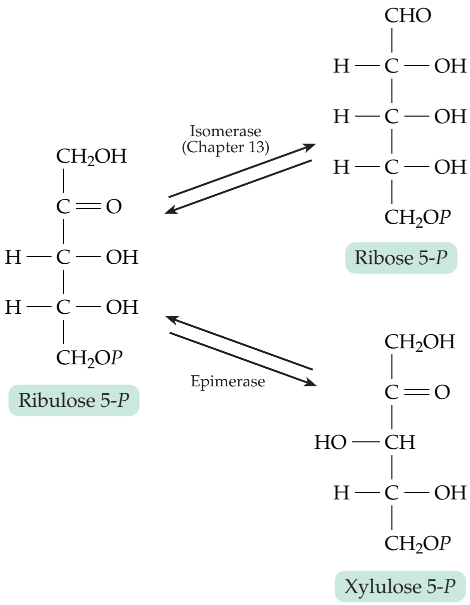
3. The Pentose Phosphate Pathways

A second way of cleaving glucose 6-phosphate utilizes sequences involving the five-carbon pentose sugars. They are referred to as **pentose phosphate pathways**, the phosphogluconate pathway, or the hexose monophosphate shunt. Historically, the evidence for such routes dates from the experiments of Warburg on the oxidation of glucose 6-*P* to 6-phosphogluconate (Chapter 15). For many years the oxidation remained an enzymatic reaction without a defined pathway. However, it was assumed to be part of an alternative method of degradation of glucose. Supporting evidence was found in the observation that tissues continue to respire in the presence of a high concentration of fluoride ion, a known inhibitor of the enolase reaction and capable of almost completely blocking glycolysis. Some tissues, e.g., liver, are especially active in respiration through this alternative pathway, whose details were elucidated by Horecker and associates.[116,117] We now know that the pentose phosphate pathways are multiple as well as multipurpose. They function in catabolism and also, when operating in the reverse direction, as a **reductive pentose phosphate pathway** that lies at the heart of the sugar-forming reactions of photosynthesis.

The oxidative pentose pathway provides *a means for cutting the chain of a sugar molecule one carbon at a time*, with the carbon removed appearing as CO_2. The enzymes required can be grouped into three distinct systems, all of which are found in the cytosol of animal cells: (i) a dehydrogenation–decarboxylation system, (ii) an isomerizing system, and (iii) a sugar rearrangement system. The dehydrogenation–decarboxylation system cleaves glucose 6-*P* to CO_2 and the pentose phosphate, ribulose 5-*P* (Eq. 17-12). Three enzymes are required, the first being glucose 6-*P* dehydrogenase[117a] (Eq. 17-12, step *a*; see also Eq. 15-10). The immediate product, a lactone, undergoes spontaneous hydrolysis. However, the action of **gluconolactonase** (Eq. 17-12, step *b*) causes a more rapid ring opening. A second dehydrogenation is catalyzed by **6-phosphogluconate dehydrogenase** (Eq. 17-12, step *c*),[117b] and this reaction is immediately followed by a β decarboxylation catalyzed by the same enzyme (as in Eq. 13-45). The value of $\Delta G°$ for oxidation of glucose 6-*P* to ribulose 5-*P* by $NADP^+$ according to Eq. 17-12 is –30.8 kJ mol^{-1}, a negative enough value to drive the [NADPH]/[$NADP^+$] ratio to an equilibrium value of over 2000 at a CO_2 tension of 0.05 atm.

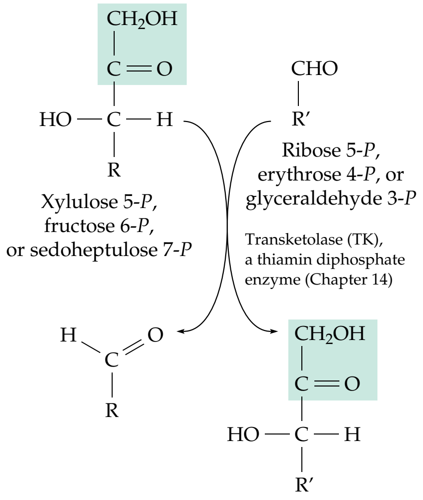

OP
HO
HO
O
OH
OH
Glucose 6-P
a
NADP+
Glucose-6-phosphate dehydrogenase
NADPH
OP
HO
HO
O
OH
O
6-Phosphogluconolactone
b
Gluconolactonase
COO–
H—C—OH
HO—C—H
H—C—OH
H—C—OH
CH2OP
6-Phosphogluconate
NADP+
NADPH + H+
c
6-Phosphogluconate dehydrogenase
COOH
H—C—OH
C=O
H—C—OH
H—C—OH
CH2OP
CO2
CH2OH
C=O
H—C—OH
H—C—OH
CH2OP
Ribulose 5-P

(17-12)

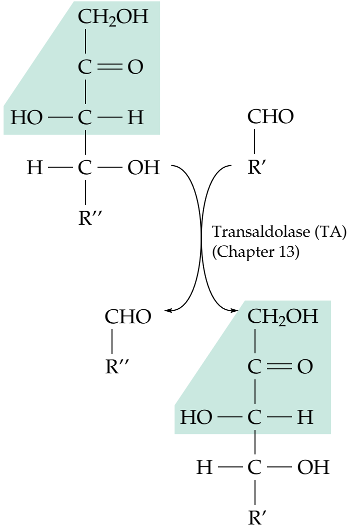

CH2OH
C=O
H—C—OH
H—C—OH
CH2OP
Ribulose 5-P
Isomerase (Chapter 13)
CHO
H—C—OH
H—C—OH
H—C—OH
CH2OP
Ribose 5-P
Epimerase
CH2OH
C=O
HO—CH
H—C—OH
CH2OP
Xylulose 5-P

(17-13)

The isomerizing system, consisting of two enzymes,


CH2OH
C=O
HO—C—H
R
Xylulose 5-P, fructose 6-P, or sedoheptulose 7-P
CHO
R′
Ribose 5-P, erythrose 4-P, or glyceraldehyde 3-P
Transketolase (TK), a thiamin diphosphate enzyme (Chapter 14)
H
C
O
R
CH2OH
C=O
HO—C—H
R′

(17-14)



(17-15)

interconverts three pentose phosphates (Eq. 17-13). As a consequence the three compounds exist as an equilibrium mixture. Both xylulose 5-*P* and ribose 5-*P* are needed for further reactions in the pathways.

The ingenious sugar rearrangement system uses two enzymes, **transketolase** and **transaldolase**. Both catalyze chain cleavage and transfer reactions (Eqs. 17-14 and 17-15) that involve the same group of substrates. These enzymes use the two basic types of C–C bond cleavage, adjacent to a carbonyl group (α) and one carbon removed from a carbonyl group (β). Both types are needed in the pentose phosphate pathways just as they are in the citric acid cycle. The enzymes of the pentose phosphate pathway are found in the cytoplasm of both animal and plant cells.[117c] Mammalian cells appear to have an additional set that is active in the endoplasmic reticulum and plants have another set in the chloroplasts.[117c]

An oxidative pentose phosphate cycle. Putting the three enzyme systems together, we can form a cycle that oxidizes hexose phosphates. Three carbon atoms are chopped off one at a time (Fig. 17-8A) leaving a three-carbon triose phosphate as the product. Since the dehydrogenation system works only on glucose 6-*P*, a part of the sugar rearrangement system must be utilized between each of the three oxidation steps. Notice that a C_5 unit (ribose 5-*P*) is used in the first reaction with transketolase but is regenerated at the end of the sequence. This C_5 unit is the regenerating substrate for the cycle. As indicated by the dashed arrows, it is formed readily in any quantity needed by oxidation of glucose 6-*P*. Before the C_5 unit that is formed in each oxidation step can be processed by the sugar rearrangement reactions, it must be isomerized[117c,118,118a,b] from ribulose 5-*P* to xylulose 5-*P*; before the C_5 unit, produced at the end of the sequence in Fig. 17-8, can be reutilized as a regenerating substrate, it must be isomerized to ribose 5-*P*. Thus, the cycle is quite complex. The same C_5 substrates appear at several points in Fig. 17-8A and substrates from different parts of the cycle become scrambled and the pathway does not degrade all the hexose molecules in a uniform manner. For this reason, Zubay described the pentose phosphate pathways as a "swamp."[119]

The oxidative pentose phosphate cycle is often presented as a means for complete oxidation of hexoses to CO_2. For this to happen the C_3 unit indicated as the product in Fig. 17-8A must be converted (through the action of aldolase, a phosphatase, and hexose phosphate isomerase) back to one-half of a molecule of glucose-6-*P* which can enter the cycle at the beginning. On the other hand, alternative ways of degrading the C_3 product glyceraldehyde-*P* are available. For example, using glycolytic enzymes, it can be oxidized to pyruvate and to CO_2 via the citric acid cycle.

As a general rule, NAD^+ is associated with catabolic reactions and it is somewhat unusual to find $NADP^+$ acting as an oxidant. However, in mammals the enzymes of the pentose phosphate pathway are specific for $NADP^+$. The reason is thought to lie in the need of NADPH for biosynthesis (Section I). On this basis, the occurrence of the pentose phosphate pathway in tissues having an unusually active biosynthetic function (liver and mammary gland) is understandable. In these tissues the cycle may operate as indicated in Fig. 17-8A with the C_3 product also being used in biosynthesis. Furthermore, any of the products from C_4 to C_7 may be withdrawn in any desired amounts without disrupting the smooth operation of the cycle. For example, the C_4 intermediate **erythrose 4-*P*** is required in synthesis of aromatic amino acids by bacteria and plants (but not in animals). **Ribose 5-*P*** is needed for formation of several amino acids and of nucleic acids by all organisms. In some circumstances the formation of ribose 5-*P* may be the only essential function for the pentose phosphate pathway.[120]

Several studies of the metabolism of isotopically labeled glucose[121–122a] have been in accord with

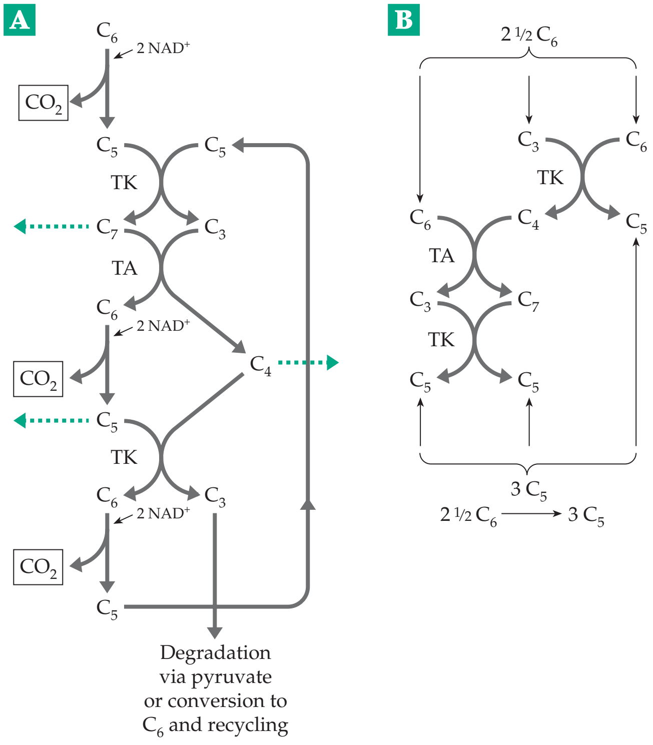


Figure 17-8 The pentose phosphate pathways. (A) Oxidation of a hexose (C_6) to three molecules of CO_2 and a three-carbon fragment with the option of removing C_3, C_4, C_5, and C_7 units for biosynthesis (dashed arrows). (B) Nonoxidative pentose pathways: 2 ½ $C_6 \rightarrow 3\ C_5$ or 2 $C_6 \rightarrow 3\ C_4$ or 3 ½ $C_6 \rightarrow 3\ C_7$.

---

operation of the pentose phosphate pathway as is depicted in Fig. 17-8. However, Williams and associates proposed a modification in the sugar rearrangement sequence in liver[123–126] to include the formation of arabinose 5-*P* (from ribose 5-*P*), an octulose bisphosphate, and an octulose 8-monophosphate. Many investigators argue that these additional reactions are of minor significance.[121,122,127] The measured concentrations of pentose phosphate pathway intermediates in rat livers are close to those predicated for a near-equilibrium state from equilibrium constants measured for the individual steps of Fig. 17-8.[128] Most of the concentrations are in the 4- to 10-µM range but the level of erythrose 4-*P*, which is predicted to be ~ 0.2 µM, is too low to measure.

In contrast to animals, the resurrection plant *Craterostigma plantaginenm* accumulates large amounts of a 2-oxo-octulose. This plant is one of a small group of angiosperms that can withstand severe dehydration and are able to rehydrate and resume normal metabolism within a few hours. During desiccation much of the octulose is converted into sucrose. The plant has extra transketolase genes which may be essential for this rapid interconversion of sugars.[129]

Nonoxidative pentose phosphate pathways. The sugar rearrangement system together with the glycolytic enzymes that convert glucose 6-*P* to glyceraldehyde 3-*P* can function to transform hexose phosphates into pentose phosphates (Fig. 17-8B; Eq. 17-16) which may be utilized for nucleic acid synthesis in erythrocytes and other cells.[130,131]

$$2\tfrac{1}{2}\ C_6 \rightarrow 3\ C_5 \tag{17-16}$$

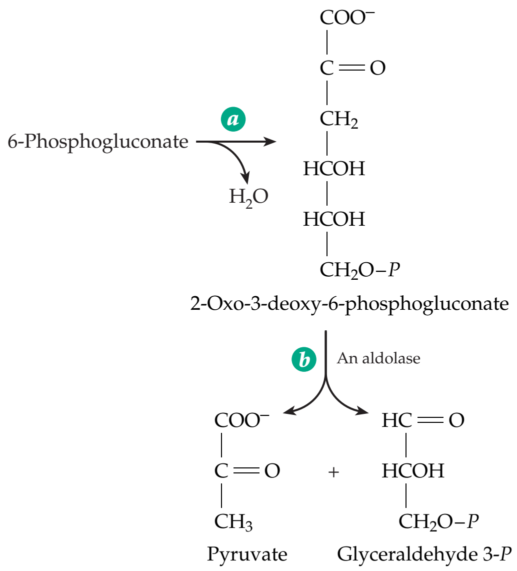
The reader can easily show that the same enzymes will catalyze the net conversion of hexose phosphate to erythrose 4-phosphate or to sedoheptulose 7-phosphate (Eq. 17-17):

$$2\ C_6 \rightarrow 3\ C_4; \quad 3\tfrac{1}{2}\ C_6 \rightarrow 3\ C_7 \tag{17-17}$$

An investigation of metabolism of the red lipid-forming yeast *Rhodotorula gracilis* (which lacks phosphofructokinase and is thus unable to break down sugars through the glycolytic pathway) indicated that 20% of the glucose is *oxidized* through the pentose phosphate pathways while 80% is *altered* by the nonoxidative pentose phosphate pathway.[100] However, it is not clear how the C_3 unit used in the nonoxidative pathway (Fig. 17-8B) is formed if glycolysis is blocked. A number of fermentations are also based on the pentose phosphate pathways (Section F,5).

4. The Entner–Doudoroff Pathway

An additional way of cleaving a six-carbon sugar chain provides the basis for the **Entner–Doudoroff pathway** which is used by *Zymomonas lindneri* and many other species of bacteria. Glucose 6-*P* is oxidized first to 6-phosphogluconate, which is converted by dehydration to a 2-oxo-3-deoxy derivative (Eq. 17-18,



(17-18)

step *a*). The resulting 2-oxo-3-deoxy sugar is cleaved by an aldolase (Eq. 17-18, step *b*) to pyruvate and glyceraldehyde 3-*P*, which are then metabolized in standard ways.

F. Fermentation: "Life without Oxygen"

Pasteur recognized in 1860 that fermentation was not a spontaneous process but a result of life in the near absence of air.[132] He realized that yeasts decompose much more sugar under anaerobic conditions than they do aerobically, and that the anaerobic fermentation was essential to the life of these organisms. In addition to the alcoholic fermentation of yeast, there are many other fermentations which have been attractive subjects for biochemical study. If life evolved at a time when no oxygen was available, the most primitive organisms must have used fermentations. They may be the oldest as well as the simplest ways in which cells obtain energy. The enzymes of the glycolysis pathway are found in the small genomes of *Mycoplasma, Haemophila*, and *Methanococcus*.[133,134]

Fermentation is also a vital process in the human body. Our muscles usually receive enough oxygen to oxidize pyruvate and to obtain ATP through aerobic metabolism, but there are circumstances in which the oxygen supply is inadequate. During extreme exertion, after most oxygen is consumed, muscle cells produce lactate by fermentation. White muscle of fish and fowl has little aerobic metabolism and normally yields L-lactic acid as a principal end product. Likewise, a variety of tissues within the human body, including the transparent lens and cornea of our eyes, are poorly supplied with blood and depend upon fermentation of glucose to lactic acid. Red blood cells and skin and sometimes adipose tissue are also major producers of lactic acid.[135] Of the ~115 g of lactic acid present in a 70-kg human body, about 29% comes from erythrocytes, 29% from skin, 17% from the brain, and 16% from skeletal muscle.[136] Because lactic acid lowers the pH of cells it must be removed efficiently.

Some of the lactic acid formed in muscle and most of the lactate formed in less aerobic tissues (e.g., adipose tissue)[136a] enters the bloodstream, which normally contains 1–2 mM lactate,[136] and is carried to the liver where it is reoxidized to pyruvate. Part of the pyruvate is then oxidized via the citric acid cycle while a larger part is reconverted to glucose (Section J,5). This glucose may be released into the bloodstream and returned to the muscles. The overall process is known as the **Cori cycle**. Lactic acid accumulates in muscle after vigorous exercise. It is exported to the liver slowly, but if mild exercise continues the lactate may be largely oxidized within muscle via the tricarboxylic acid cycle. Recent NMR studies have shown that lactic acid is formed rapidly during muscular contraction, even when exercise is mild.[136b] During the initial 15 ms of contraction the ATP utilized is regenerated from creatine phosphate (Eq. 6-67). During the remainder of the contraction (up to ~100 ms) glycogen is converted to lactic acid to provide ATP and to replenish the creatine phosphate. In the resting period following contraction most of the lactate is either dehydrogenated to pyruvate and oxidized in mitochondria or exported to other tissues. The glycogen stores in muscle are renewed by synthesis from blood glucose. Lactic acid is a convenient energy carrier and a precursor for gluconeogenesis which can be transferred between tissues easily.[136c] Cancer cells often take advantage of this opportunity to grow rapidly using fermentation of glucose to lactic acid as a source of energy.[136d]

Alcoholic fermentation allows roots of some plants to survive short periods of flooding. Ethanol does not acidify the tissues as does lactic acid, avoiding possible damage from low pH.[137,138] Goldfish can also use the ethanolic fermentation for short times, excreting the ethanol.[139]

1. Fermentations Based on the Embden–Meyerhof Pathway

Homolactic and alcoholic fermentations. The reactions by which glucose can be converted to lactate and, by yeast cells, to ethanol and CO_2 (Figs. 10-3 and 17-7) illustrate several features common to all fermentations. The NADH produced in the oxidation step is reoxidized in a reaction by which substrate is reduced to the final end product. The NAD alternates between oxidized and reduced forms. This coupling of oxidation steps with reduction steps in exact equivalence is characteristic of all true (anaerobic) fermentations. The formation of ATP from ADP and P_i by substrate-level phosphorylation is also common to all fermentations. The stoichiometry is often nearly exact and simple. For example, according to the reaction of Eq. 17-19, which is outlined step-by-step in Fig. 17-7, a net total of two moles of ATP is formed per mole of glucose fermented.

Energy relationships. If we disregard the synthesis of ATP, the equations for the lactic acid and ethanol fermentations are given by Eqs. 17-19 and 17-20.

$$\begin{gathered}\text{Glucose} \rightarrow 2\ \text{lactate}^- + 2\ \text{H}^+ \\ \Delta G'\ (\text{pH } 7) = -196\ \text{kJ mol}^{-1}\ (-46.8\ \text{kcal mol}^{-1})\end{gathered} \tag{17-19}$$

$$\begin{gathered}\text{Glucose} \rightarrow 2\ \text{CO}_2 + 2\ \text{ethanol} \\ \Delta G^\circ = -235\ \text{kJ mol}^{-1}\end{gathered} \tag{17-20}$$

The Gibbs energy changes are negative and of sufficient magnitude that the reactions will unquestionably go to completion. However, the synthesis of two molecules of ATP from inorganic phosphate and ADP, a reaction

(Eq. 17-21) for which $\Delta G'$ is substantially positive, is coupled to the fermentation.

$$\mathrm{ADP^{3-} + HPO_4^{\,2-} + H^+ \rightarrow ATP^{4-} + H_2O}$$
$$\Delta G' \text{ (pH 7)} = +\,34.5 \text{ kJ mol}^{-1} \text{ (Table 6-5)} \qquad (17\text{-}21)$$

To obtain the net Gibbs energy change for the complete reaction we must add 2 x 34.5 = +69.0 kJ to the values of $\Delta G'$ for Eqs. 17-19 and 17-20. When this is done we see that the net Gibbs energy changes are still highly negative, that the reactions will proceed to completion, and that these fermentations can serve as an usable energy source for organisms.

Biochemists sometimes divide ΔG for the ATP synthesis in a coupled reaction sequence (in this case +69 kJ) by the overall Gibbs energy decrease for the coupled process (196 or 235 kJ mol^{-1}) to obtain an "efficiency." In the present case the efficiency would be 35% and 29% for coupling of Eq. 17-21 (for 2 mol of ATP) to Eqs. 17-19 and 17-20, respectively. According to this calculation, nature is approximately one-third efficient in the utilization of available metabolic Gibbs energy for ATP synthesis. However, it is important to realize that this calculation of efficiency has no exact thermodynamic meaning. Furthermore, the utilization of ATP formed by a cell for various purposes is far from 100% efficient.

Why are the Gibbs energy decreases for Eqs. 17-19 and 17-20 so large? No overall oxidation takes place; there is only a rearrangement of the existing bonds between atoms of the substrate. Why does this rearrangement of bonds lead to a substantial negative ΔG? An answer is suggested by an examination of the numbers of each type of bond in the substrate and in the products. During the conversion from glucose to two molecules of lactate one C–C bond, one C–O bond, and one O–H bond are lost and one C–H bond and one C=O are gained. If we add up the bond energies for these bonds (Table 6-6) we find that the difference (ΔH) between substrate and products amounts to only about 20 kJ/mol. However, lactic acid contains a carboxyl group, and carboxyl groups have a special stability as a result of resonance. The extra resonance energy of a carboxyl group (Table 6-6) is ~117 kJ (28 kcal) per mole or 234 kJ/mol for two carboxyl groups. This is approximately the same as the Gibbs energy change (Eq. 17-19) for fermentation of glucose to lactate. Thus, the energy available results largely from the rearrangement of bonds by which the carboxyl groups of lactate are formed. Likewise, the resonance stabilization of CO_2 is given by Pauling as 151 kJ/mol, again of just the right magnitude to explain ΔG in alcoholic fermentations (Eq. 17-20).

On this basis we can state as a general rule that fermentations can occur when substrates consisting of largely singly bonded atoms and groups, such as the carbonyl groups that are not highly stabilized by resonance, are converted to products containing carboxyl groups or to CO_2. If we assume an efficiency of ~30%, the energy available will be about sufficient for synthesis of one ATP molecule for each carboxyl group or CO_2 created. Bear in mind that generation of ATP also depends upon availability of a mechanism. It is of interest that most synthesis of ATP is linked directly to the chemical processes by which carboxyl groups or CO_2 molecules are created in a fermentation process. The most important single reaction is the oxidation of the aldehyde group of glyceraldehyde 3-*P* to the carboxyl group of 3-phosphoglycerate (steps *6a*– *6c* and *7* in Fig. 17-7; see also Fig. 15-6).

Compare the fermentation of glucose with the complete oxidation of the sugar to carbon dioxide and water (Eq. 17-22), a process which yeast cells (as well as our own cells) carry out in the presence of air. The overall Gibbs energy change is over 10 times greater than that for fermentation, a fact that permits the cell

$$\mathrm{Glucose + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O}$$
$$\Delta G' = -2872 \text{ kJ } (-686.5 \text{ kcal}) \text{ mol}^{-1} \qquad (17\text{-}22)$$

to form an enormously greater quantity of ATP. The net gain in ATP synthesis, accompanying Eq. 17-22, is usually about 38 mol of ATP—19 times more than is available from fermentation of glucose. Thus, the explanation of Pasteur's observation that yeast decomposes much less sugar in the presence of air than in its absence is clear. Also, we can understand why a cell, living anaerobically, must metabolize a very large amount of substrate to grow. (Recall from Chapter 6 that ~1 mol of ATP energy is needed to produce 10 g of cells.)

Variations of the alcoholic and homolactic fermentations. The course of a fermentation is often affected drastically by changes in conditions. Many variations can be visualized by reference to Fig. 17-9, which shows a number of available metabolic sequences. We have already discussed the conversion of glucose to triose phosphate and via reaction pathway *a* to pyruvate, via reaction *c* to lactate, and via reaction *d* to ethanol.

If bisulfite is added to a fermenting culture of yeast, the acetaldehyde formed through reaction *d* is trapped as the bisulfite adduct blocking the reduction of acetaldehyde to ethanol, an essential part of the fermentation. Yeast cells accommodate this change by using the accumulating NADH to reduce half of the triose-*P* to glycerol through pathway *b*. Two enzymes are needed, a dehydrogenase and a phosphatase, to hydrolytically cleave off the phosphate. The balanced reaction is given by Eq. 17-23:

$$\mathrm{Glucose \rightarrow glycerol + acetaldehyde\ (trapped) + CO_2}$$
$$\Delta G' \text{ (pH 7)} = -105 \text{ kJ mol}^{-1} \qquad (17\text{-}23)$$

In this reaction only one molecule of CO_2 is produced but the overall Gibbs energy change is still adequate to make the reaction highly spontaneous. However (referring to Fig. 17-9), we see that the net synthesis of ATP is now apparently zero. The fermentation apparently does not permit cell growth. Nevertheless, it has been used industrially for production of glycerol.

Reduction of dihydroxyacetone phosphate to glycerol phosphate also occurs in insect flight muscle and apparently operates as an alternative to lactic acid formation in that tissue. There is no net gain of ATP in the conversion of free glucose to glycerol phosphate and pyruvate, but using stored glycogen in muscle as the starting material, the dismutation of triose-*P* to glycerol-*P* and pyruvate provides one ATP per glucose unit rapidly during the vigorous contraction of the powerful insect flight muscle. During the slower recovery phase, glycerol-*P* is thought to be reoxidized after entering the mitochondria of these highly aerobic cells. Thus, the transport of glycerol-*P* into mitochondria serves as a means for transporting reducing equivalents derived from reoxidation of NADH into the mitochondria. Indeed, the significance of glycerol-*P* to muscle metabolism may be more related to this function than to the rapid formation of ATP (see Chapter 18).

Why does the glycolysis sequence begin with phosphorylation of glucose by ATP? The phospho groups probably provide convenient handles and doubtless assist in substrate recognition. There may be a kinetic advantage but also a danger. Unless there is adequate regulation the "turbo design," in which ATP is used at the outset to drive glycolysis, may lead to accumulation of phosphorylated intermediates and to inadequate concentrations of ATP and inorganic phosphate.[139a,b] Yeast cells guard against this problem by synthesizing trehalose 6-phosphate, which acts as a feedback inhibitor of hexokinase.[139a] Trypanosomes utilize a different type of control. The enzymes that convert glucose into 3-phosphoglycerate are present in membrane-bounded organelles called **glycosomes**. Phosphoglycerate is exported from them into the cytosol where glycolysis is completed.[139b] Since inorganic phosphate is essential for ATP formation, if the P_i concentration falls too low the rate of fermentation by yeast juice is greatly decreased, an observation made by Harden and Young[139c] in 1906.

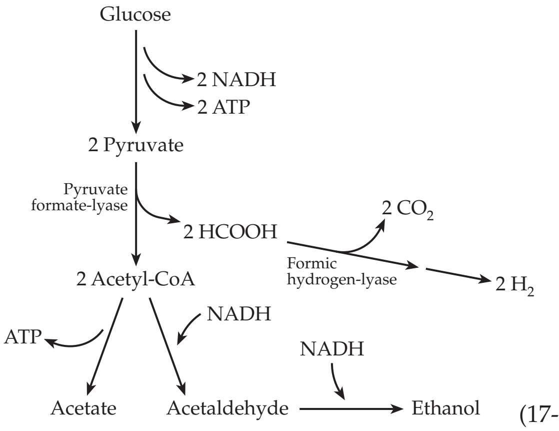
2. The Mixed Acid Fermentation

Enterobacteria, including *E. coli*, convert glucose to ethanol and acetic acid and either formic acid or CO_2 and H_2 derived from it. The stoichiometry is variable but the fermentation can be described in an idealized form as follows:

$$\begin{aligned}\text{Glucose} + H_2O \rightarrow \text{ethanol} + \text{acetate}^- \\ + H^+ + 2\,H_2 + 2\,CO_2\end{aligned}$$

$$\Delta G'\ (\text{pH } 7) = -225\ \text{kJ mol}^{-1} \qquad (17\text{-}24)$$

The details of the process and the oxidation–reduction balance can be pictured as in Eq. 17-25. Pyruvate is cleaved by the pyruvate formate–lyase reaction (Eq. 15-37) to acetyl-CoA and formic acid. Half of the acetyl-CoA is cleaved to acetate via acetyl-*P* with generation of ATP, while the other half is reduced in two steps to ethanol using the two molecules of NADH produced in the initial oxidation of triose phosphate (Eq. 17-25). The overall energy yield is three molecules of ATP per glucose. The "efficiency" is thus (3 x 34.5) ÷ 225 = 46%. Some of the glucose is converted to D-lactic and to succinic acids (pathway *f*, Fig. 17-9); hence the name **mixed acid fermentation**. Table 17-1 gives typical yields of the mixed acid fermentation of *E. coli*. Among the four major products are acetate, ethanol, H_2, and CO_2, as shown in Eq. 17-25. However, at high pH formate accumulated instead of CO_2.



(17-25)

TABLE 17-1
Products of the Mixed Acid Fermentation by *E. coli* at Low and High Values of pH[a]

Product (Millimole formed from 100 mmol of glucose)	pH 6.2	pH 7.8
Acetate	36	39
Ethanol	50	51
H_2	70	0.3
CO_2	88	1.7
Formate	2.4	86
Lactate	79	70
Succinate	11	15
Glycerol	1.4	0.3
Acetoin	0.1	0.2
Butanediol	0.3	0.2

[a] From Tempest, D. W. and Neijssel, O. M.[140] Based on data of Blackwood.[141]

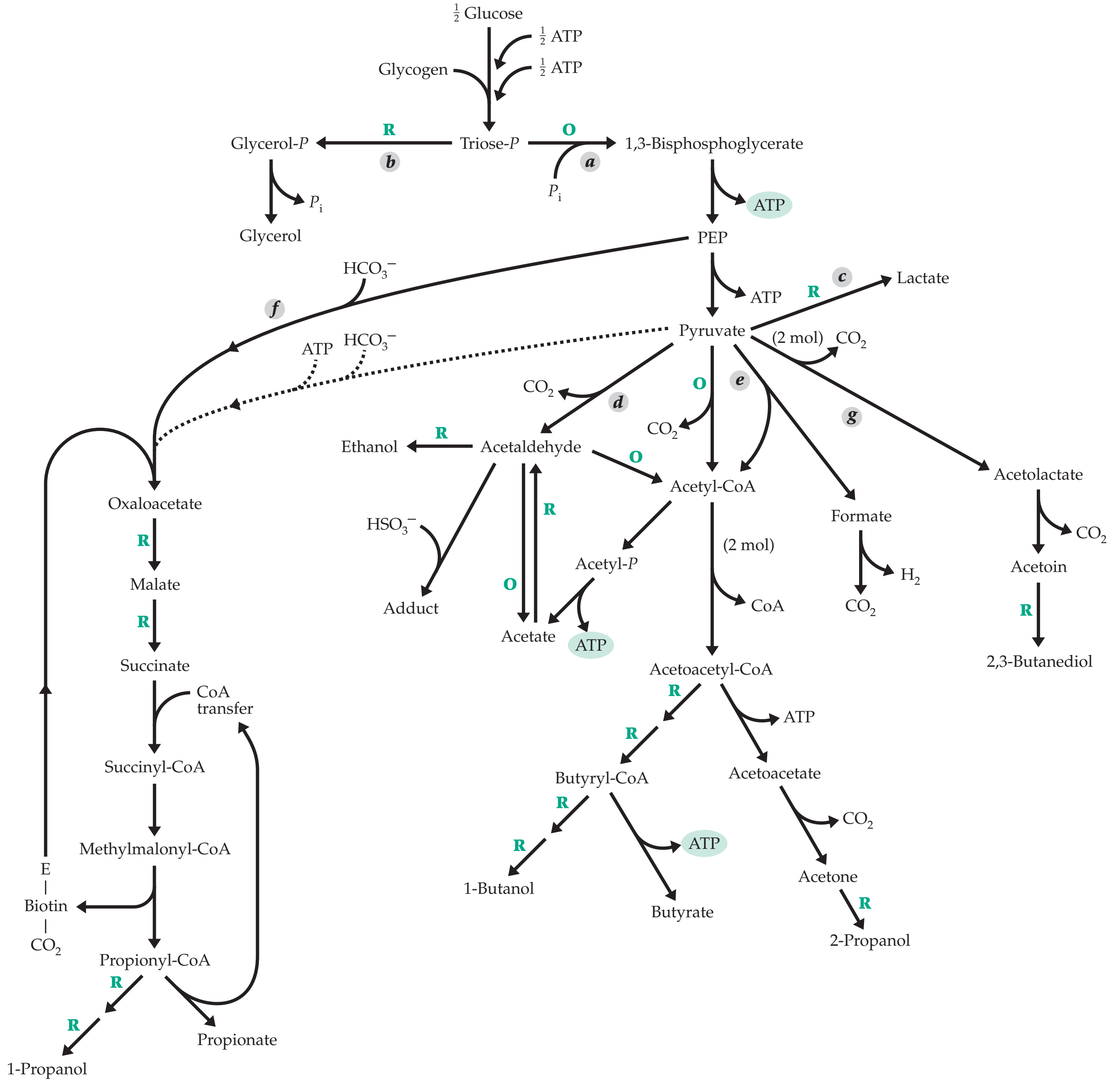
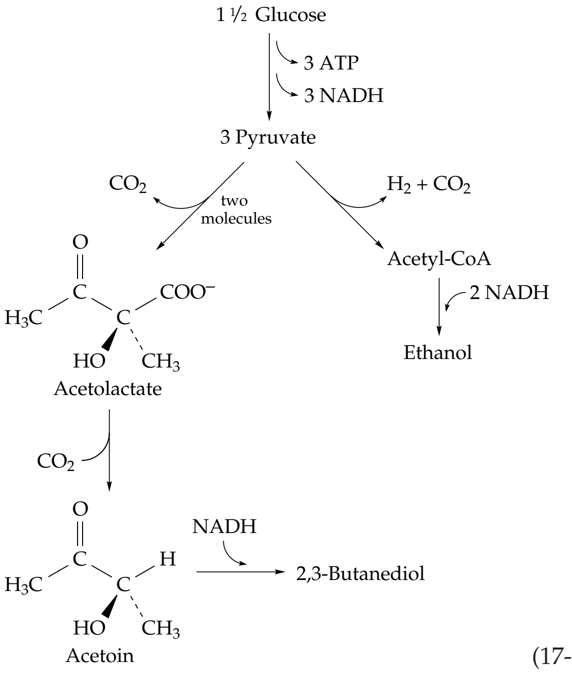

½ Glucose
½ ATP
½ ATP
Glycogen
Triose-P
R
b
Glycerol-P
Pi
Glycerol
O
a
Pi
1,3-Bisphosphoglycerate
ATP
PEP
HCO3−
f
ATP
c
R
Lactate
Pyruvate
(2 mol) CO2
ATP HCO3−
CO2
d
O
e
CO2
g
R
Ethanol
Acetaldehyde
O
Acetyl-CoA
Acetolactate
CO2
Acetoin
R
2,3-Butanediol
Formate
H2
CO2
Oxaloacetate
R
Malate
R
Succinate
CoA transfer
Succinyl-CoA
Methylmalonyl-CoA
E
Biotin
CO2
Propionyl-CoA
R
R
1-Propanol
Propionate
HSO3−
Adduct
R
O
Acetate
Acetyl-P
ATP
(2 mol)
CoA
Acetoacetyl-CoA
R
R
ATP
Butyryl-CoA
Acetoacetate
R
R
1-Butanol
ATP
Butyrate
CO2
Acetone
R
2-Propanol

Figure 17-9 Reaction sequences in fermentation based on the Embden–Meyerhof–Parnas pathway. Oxidation steps (producing NADH + H^+) are marked "O"; reduction steps (using NADH + H^+) are marked "R."

Over one-third of the glucose was fermented to lactate in both cases.

In some mixed acid fermentations (e.g., that of *Shigella*) formic acid accumulates, but in other cases (e.g., with *E. coli* at pH 6) it is converted to CO_2 and H_2 (Eq. 17-25). The equilibration of formic acid with CO_2 and hydrogen is catalyzed by the **formic hydrogen–lyase** system which consists of two iron–sulfur enzymes. The selenium-containing **formate dehydrogenase** (Eq. 16-63) catalyzes oxidation of formate to CO_2 by NAD^+, while a membrane-bound **hydrogenase** (Eq. 16-48) equilibrates NADH + H^+ with NAD^+ + H_2. Hydrogenase also serves to release H_2 from excess NADH. Krebs pointed out that an excess of NADH may arise because growth of cells requires biosynthesis of many components such as amino acids. When glucose is the sole source of carbon, biosynthetic reactions involve an excess of oxidation steps over reduction steps.[142] The excess of reducing equivalents may be released as H_2 or may be used to form highly reduced products such as succinate.

Among such genera as *Aerobacter* and *Serratia* part of the pyruvate formed is condensed with decarboxylation to form **S acetolactate**,[143] which is decarboxylated to acetoin (Eq. 17-26; pathway *g* of Fig. 17-9). The acetoin is reduced with NADH to **2,3-butanediol**, while a third molecule of pyruvate is converted to ethanol, hydrogen, and CO_2 (Eq. 17-26). The reaction provides the basis for industrial production of butanediol, which can be dehydrated nonenzymatically to butadiene.



(17-26)

Mixed acid fermentations are not limited to bacteria. For example, trichomonads, parasitic flagellated protozoa, have no mitochondria. They export pyruvate into the bloodstreams of their hosts and also contain particles called **hydrogenosomes** which can convert pyruvate to acetate, succinate, CO_2, and H_2.[144] Hydrogenosomes are bounded by double membranes and have a common evolutionary relationship with both mitochondria and bacteria. The enzyme that catalyzes pyruvate cleavage in hydrogenosomes apparently does not contain lipoate and may be related to the pyruvate–ferredoxin oxidoreductase of clostridia (Eq. 15-35). The hydrogenosomes also contain an active hydrogenase.

Many invertebrate animals are true facultative anaerobes, able to survive for long periods, sometimes indefinitely, without oxygen.[145–147] Among these are *Ascaris* (Fig. 1-14), oysters, and other molluscs. Succinate and alanine are among the main end products of anaerobic metabolism. The former may arise by a mixed acid fermentation that also produces pyruvate. The pyruvate is converted to acetate to balance the fermentation in *Ascaris lumbricoides*, which is in effect an obligate anaerobe. However, in molluscs the pyruvate may undergo transamination with glutamate to form alanine and 2-oxoglutarate; the oxoglutarate may be oxidatively decarboxylated to succinate. The reactions depend upon the availability of a store of glutamate or of other amino acids, such as arginine, that can give rise to glutamate.

3. The Propionic Acid Fermentation

Propionic (propanoic) acid-producing bacteria are numerous in the digestive tract of ruminants. Within the rumen some bacteria digest cellulose to form glucose, which is then converted to lactate and other products. The propionic acid bacteria can convert either glucose or lactate into propionic and acetic acids which are absorbed into the bloodstream of the host. Usually some succinic acid is also formed.

The basis of the propionic acid fermentation is conversion of pyruvate to oxaloacetate by carboxylation and the further conversion through succinate and succinyl-CoA to methylmalonyl-CoA and propionyl-CoA, reactions which are almost the exact reverse of those for the oxidation of propionate in the animal body (Fig. 17-3, pathway *d*). However, whereas the carboxylation of pyruvate to oxaloacetate in the animal body requires ATP, the propionic acid bacteria save one equivalent of ATP by using a carboxyltransferase (p. 725). This enzyme donates a carboxyl group from a preformed carboxybiotin compound generated in the decarboxylation of methylmalonyl-CoA in the next to final step of the reaction sequence (Fig. 17-10). A second molecule of ATP is saved by linking directly the conversion of succinate to succinyl-CoA to the cleavage of propionyl-CoA to propionate through the use of a CoA transferase (Eq. 12-50). To provide for oxidation–reduction balance, two-thirds of the glucose goes to propionate and one-third to acetate (Eq. 17-27):

$$\begin{aligned}&1\tfrac{1}{2}\,\text{Glucose} \rightarrow\\&\quad 2\ \text{propionate}^- + \text{acetate}^- + 3\,H^+ + CO_2 + H_2O\\&\Delta G'\ (\text{pH }7) = -465\ \text{kJ per }1\tfrac{1}{2}\ \text{mol of glucose}\end{aligned}\qquad(17\text{-}27)$$

More carboxyl groups and CO_2 molecules are formed in this fermentation ($2\tfrac{2}{3}$ per glucose molecule) than in the regular lactic acid fermentation. The yield of ATP (also $2\ ^2/_3$ mol/mol of glucose fermented) is correspondingly greater and $\Delta G'$ is more negative.

Using the same mechanism (Fig. 17-10), propionic acid bacteria are also able to ferment lactate, the product of fermentation by other bacteria, to propionate and acetate (Eq. 17-28). The net gain is one molecule of ATP. This reaction probably accounts for the niche

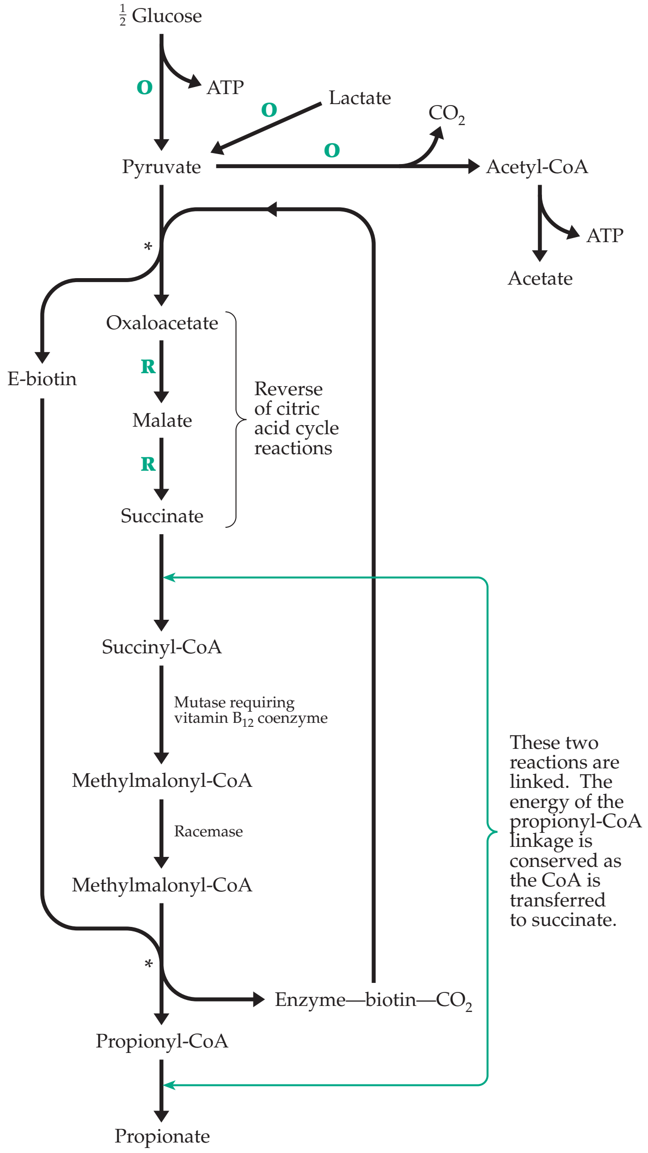

½ Glucose
O
ATP
O
Lactate
CO2
O
Pyruvate
Acetyl-CoA
ATP
Acetate
*
E-biotin
Oxaloacetate
R
Malate
R
Succinate
Reverse of citric acid cycle reactions
Succinyl-CoA
Mutase requiring vitamin B12 coenzyme
Methylmalonyl-CoA
Racemase
Methylmalonyl-CoA
These two reactions are linked. The energy of the propionyl-CoA linkage is conserved as the CoA is transferred to succinate.
*
Enzyme—biotin—CO2
Propionyl-CoA
Propionate

Figure 17-10 Propionic acid fermentation of *Propionobacteria* and *Veillonella*. Oxidation steps are designated by the symbol "O" and reduction steps by "R." The two coupled reactions marked by asterisks are catalyzed by carboxyltransferase.

$$\begin{aligned}&3\ \text{Lactate}^- \rightarrow 2\ \text{propionate}^- + \text{acetate}^- + H_2O + CO_2\\&\Delta G'\ (\text{pH 7}) = -171\ \text{kJ per 3 mol of lactate} \qquad (17\text{-}28)\end{aligned}$$

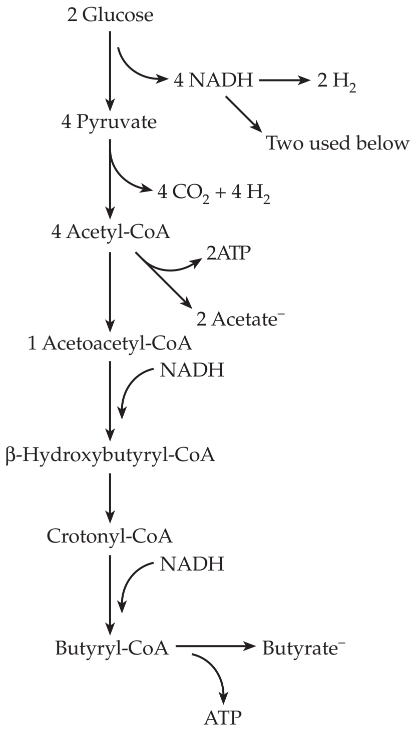
in the ecology of the animal rumen that is occupied by propionic acid bacteria.

4. Butyric Acid and Butanol-Forming Fermentations

A variety of fermentations are carried out by bacteria of the genus *Clostridium* and by the rumen organisms *Eubacterium (Butyribacterium)* and *Butyrivibrio*. For example, glucose may be converted to butyric and acetic acids together with CO_2 and H_2 (Eqs. 17-29 and 17-30).

$$\begin{aligned}&2\ \text{Glucose} + 2\ H_2O \rightarrow \text{butyrate}^- + 2\ \text{acetate}^-\\&\qquad\qquad + 4\ CO_2 + 6\ H_2 + 3\ H^+\\&\Delta G'\ (\text{pH 7}) = -479\ \text{kJ per 2 mol of glucose} \qquad (17\text{-}29)\end{aligned}$$

The yield of ATP (3½ mol/mol of glucose) is the highest we have discussed giving an efficiency of 50%. Another fermentation yields butanol, isopropanol, ethanol, and acetone.

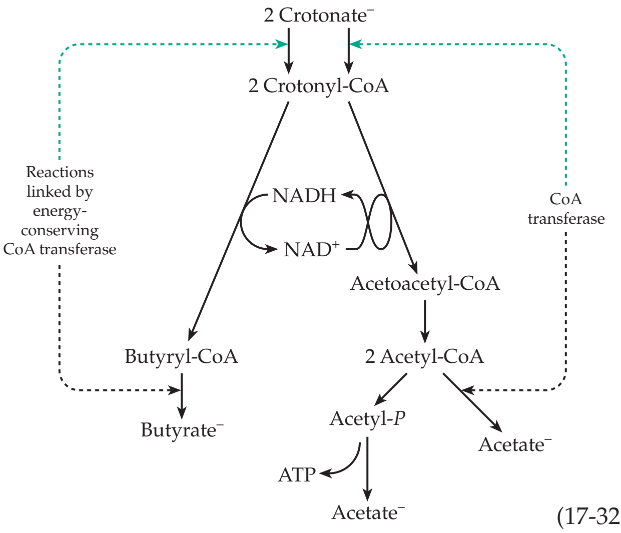

2 Glucose
4 NADH
2 H2
4 Pyruvate
Two used below
4 CO2 + 4 H2
4 Acetyl-CoA
2ATP
2 Acetate⁻
1 Acetoacetyl-CoA
NADH
β-Hydroxybutyryl-CoA
Crotonyl-CoA
NADH
Butyryl-CoA
Butyrate⁻
ATP

(17-30)

The fermentation of Eq. 17-31 is catalyzed by *Clostridium kluyveri*. The value of $-\Delta G'$ is one of the lowest that we have considered but is still enough to provide easily for the synthesis of one molecule of ATP.

$$\begin{aligned}&2\ \underset{\text{Crotonate}}{CH_3CH{=}CH{-}COO^-} + 2\ H_2O \rightarrow\\&\qquad\qquad \text{butyrate}^- + 2\ \text{acetate}^- + H^+\\&\qquad \Delta G'\ (\text{pH 7}) = -105\ \text{kJ mol}^{-1} \qquad (17\text{-}31)\end{aligned}$$

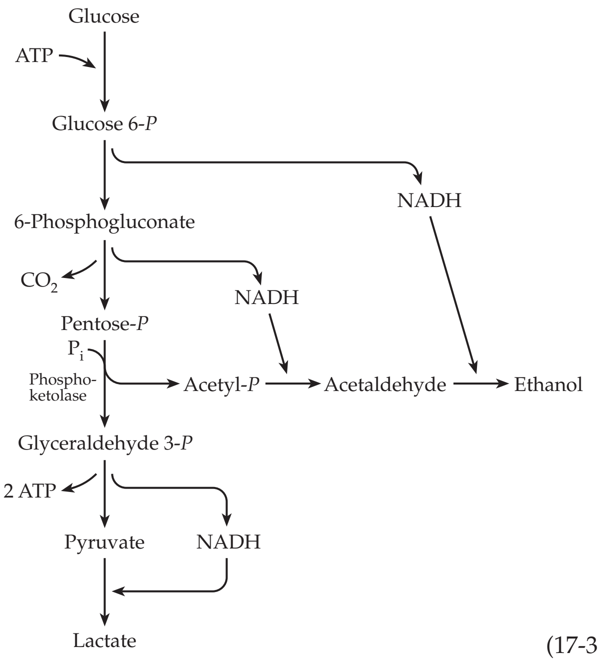
The energy of the butyryl-CoA linkage and of one of the acetyl-CoA linkages is conserved and utilized in the initial formation of crotonyl-CoA (Eq. 17-32). That leaves one acetyl-CoA which can be converted via acetyl-*P* to acetate with formation of ATP.


2 Crotonate$^-$
2 Crotonyl-CoA
Reactions linked by energy-conserving CoA transferase
NADH
NAD$^+$
CoA transferase
Acetoacetyl-CoA
Butyryl-CoA
2 Acetyl-CoA
Butyrate$^-$
Acetyl-*P*
Acetate$^-$
ATP
Acetate$^-$

(17-32)

5. Fermentations Based on the Phosphogluconate and Pentose Phosphate Pathways

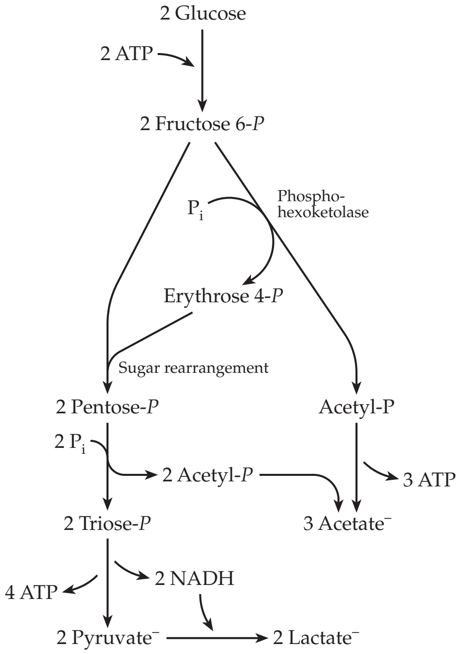
Some lactic acid bacteria of the genus *Lactobacillus*, as well as *Leuconostoc mesenteroides* and *Zymomonas mobilis*, carry out the **heterolactic** fermentation (Eq. 17-33) which is based on the reactions of the pentose phosphate pathway. These organisms lack aldolase, the key enzyme necessary for cleavage of fructose 1,6-bisphosphate to the triose phosphates. Glucose is converted to ribulose 5-*P* using the oxidative reactions of the pentose phosphate pathway. The ribulose-phosphate is cleaved by phosphoketolase (Eq. 14-23) to acetyl-phosphate and glyceraldehyde 3-phosphate, which are converted to ethanol and lactate, respectively. The overall yield is only one ATP per glucose fermented.


Glucose
ATP
Glucose 6-*P*
NADH
6-Phosphogluconate
CO$_2$
NADH
Pentose-*P*
P$_i$
Phospho-ketolase
Acetyl-*P*
Acetaldehyde
Ethanol
Glyceraldehyde 3-*P*
2 ATP
Pyruvate
NADH
Lactate

(17-33)

This is generated in the substrate level oxidative phosphorylation catalyzed by phosphoketolase. Metabolic engineering of *Zymomonas* was accomplished by transferring from other bacteria two operons that provide for assimilation of xylose and a complete set of enzymes for the pentose phosphate pathway. The engineered bacteria are able to convert pentose phosphates nonoxidatively (see Fig. 17-8) into glyceraldehyde 3-phosphate, which is converted to ethanol in high yield and with much greater synthesis of ATP than according to Eq. 17-33.[148]

A variation of the heterolactic fermentation is used by *Bifidobacterium* (Eq. 17-34).[149] Phosphoketolase and a **phosphohexoketolase**, which cleaves fructose 6-*P* to erythrose 4-*P* and acetyl-*P*, are required, as are the enzymes of the sugar rearrangement system (Section E,3). The net yield of ATP is 2 ½ molecules per molecule of glucose.


2 Glucose
2 ATP
2 Fructose 6-*P*
P$_i$
Phospho-hexoketolase
Erythrose 4-*P*
Sugar rearrangement
2 Pentose-*P*
Acetyl-P
2 P$_i$
2 Acetyl-*P*
3 ATP
2 Triose-*P*
3 Acetate$^-$
4 ATP
2 NADH
2 Pyruvate$^-$
2 Lactate$^-$

(17-34)

G. Biosynthesis

In this section and sections H – K the general principles and strategy of synthesis of the many carbon compounds found in living things will be considered. Since green plants and autotrophic bacteria are able to assemble all of their needed carbon compounds from CO$_2$, let us first examine the mechanisms by which this is accomplished. We will also need to ask how some organisms are able to subsist on such simple compounds as methane, formate, or acetate.

1. Metabolic Loops and Biosynthetic Families

As was pointed out in Chapter 10, routes of biosynthesis (anabolism) often closely parallel pathways of biodegradation (catabolism). Thus, catabolism begins with hydrolytic breakdown of polymeric molecules; the resulting monomers are then cleaved into small two- and three-carbon fragments. Biosynthesis begins with formation of monomeric units from small pieces followed by assembly of the monomers into polymers. The mechanisms of the individual reactions of biosynthesis and biodegradation are also often closely parallel. However, in most instances, there are clear-cut differences. A first principle of biosynthesis is that *biosynthetic pathways, although related to catabolic pathways, differ from them in distinct ways and are often catalyzed by completely different sets of enzymes*.

The sum of the pathways of biosynthesis and biodegradation form a continuous loop – a series of reactions that take place concurrently and often within the same part of a cell. Metabolic loops often begin in the central pathways of carbohydrate metabolism with three- or four-carbon compounds such as phosphoglycerate, pyruvate, and oxaloacetate. After loss of some atoms as CO_2 the remainder of the compound rejoins the "mainstream" of metabolism by entering a catabolic pathway leading to acetyl-CoA and oxidation in the citric acid cycle. Not all of the loops are closed within a given species. For example, *human beings are unable to synthesize the vitamins and the "essential amino acids."* We depend upon other organisms to make these compounds, but we do degrade them. Some metabolites, such as uric acid, are excreted by humans and are further catabolized by bacteria. From a chemical viewpoint the whole of nature can be regarded as an enormously complex set of branching and interconnecting metabolic cycles. Thus, the synthetic pathways used by autotrophs are all parts of metabolic loops terminating in oxidation back to CO_2.

It is often not possible to state at what point in a metabolic loop biosynthesis has been completed and biodegradation begins. An end product X that serves one need of a cell may be a precursor to another cell component Y which is then degraded to complete the loop. The reactions that convert X to Y can be regarded as either biosynthetic (for Y) or catabolic (for X).

2. Key Intermediates and Biosynthetic Families

In examining routes of biosynthesis it is helpful to identify some key intermediates. One of these is **3-phosphoglycerate**. This compound is a primary product of photosynthesis and may reasonably be regarded as the starting material from which all other carbon compounds in nature are formed. Phosphoglycerate, in most organisms, is readily interconvertible with both **glucose** and **phosphoenolpyruvate (PEP)**. Any of these three compounds can serve as the precursor for synthesis of other organic materials. A first stage in biosynthesis consists of those reactions by which 3-phosphoglycerate or PEP arise, whether it be from CO_2, formate, acetate, lipids, or polysaccharides.

The further biosynthetic pathways from 3-phosphoglycerate to the myriad amino acids, nucleotides, lipids, and miscellaneous compounds found in cells are complex and numerous. However, the basic features are relatively simple. Figure 17-11 indicates the origins of many substances including the 20 amino acids present in proteins, nucleotides, and lipids. Among the additional key biosynthetic precursors that can be identified from this chart are **glucose 6-phosphate**, **pyruvate**, **oxaloacetate**, **acetyl-CoA**, **2-oxoglutarate**, and **succinyl-CoA**.

The amino acid **serine** originates almost directly from 3-phosphoglycerate. **Aspartate** arises from oxaloacetate and **glutamate** from 2-oxoglutarate. These three amino acids each are converted to "families" of other compounds.[150] A little attention paid to establishing correct family relationships will make the study of biochemistry easier. Besides the serine, aspartate, and glutamate–oxoglutarate families, a fourth large family originates directly from pyruvate and a fifth (mostly lipids) from acetyl-CoA. The aromatic amino acids are formed from erythrose 4-*P* and PEP via the key intermediate **chorismic acid** (Box 9-E; Fig. 25-1). Other families of compounds arise from glucose 6-*P* and from the **pentose phosphates**. These groups have been set off roughly by the boxes outlined in green in Fig. 17-11.

H. Harnessing the Energy of ATP for Biosynthesis

In the past it seemed reasonable to think that some biosynthetic pathways involved exact reversal of catabolic pathways. For example, it was observed that glycogen phosphorylase catalyzed elongation of glycogen branches by transfer of glycosyl groups from glucose 1-phosphate. Likewise, the enzymes needed for the β oxidation of fatty acid derivatives, when isolated from mitochondria, catalyze formation of fatty acyl-CoA derivatives from acetyl-CoA and a reducing agent such as NADH. However, reactant concentrations within cells are rarely appropriate for reversal of a catabolic sequence. For catabolic sequences the Gibbs energy change is usually distinctly negative and reversal requires high concentrations of end products. However, the latter are often removed promptly from the cells. For example, NADH produced in degradation of fatty acids is oxidized to NAD^+ and is therefore never available in sufficient concentrations to reverse the β oxidation sequence.

Nature's answer to the problem of reversing a catabolic pathway lies in the coupling of cleavage of ATP to the biosynthetic reaction. The concept was introduced in Chapter 10, in which one sequence for linking hydrolysis of ATP to biosynthesis was discussed. However, living cells employ several different methods of harnessing the Gibbs energy of hydrolysis of ATP to drive biosynthetic processes. Many otherwise strange aspects of metabolism become clear if it is recognized that they provide a means for coupling ATP cleavage to biosynthesis. A few of the most important of these coupling mechanisms are summarized in this section.

1. Group Activation

Consider the formation of an ester (or of an amide) from a free carboxylic acid and an alcohol (or amine) by elimination of a molecule of water (Eq. 17-35). The reaction is thermodynamically unfavorable with values of $\Delta G'$ (pH 7) ranging from ~ +10 to 30 kJ mol^{-1} depending on conditions and structures of the specific compounds. Long ago, organic chemists learned that such reactions can be made to proceed by careful removal of the water that is generated (Eq. 17-35).

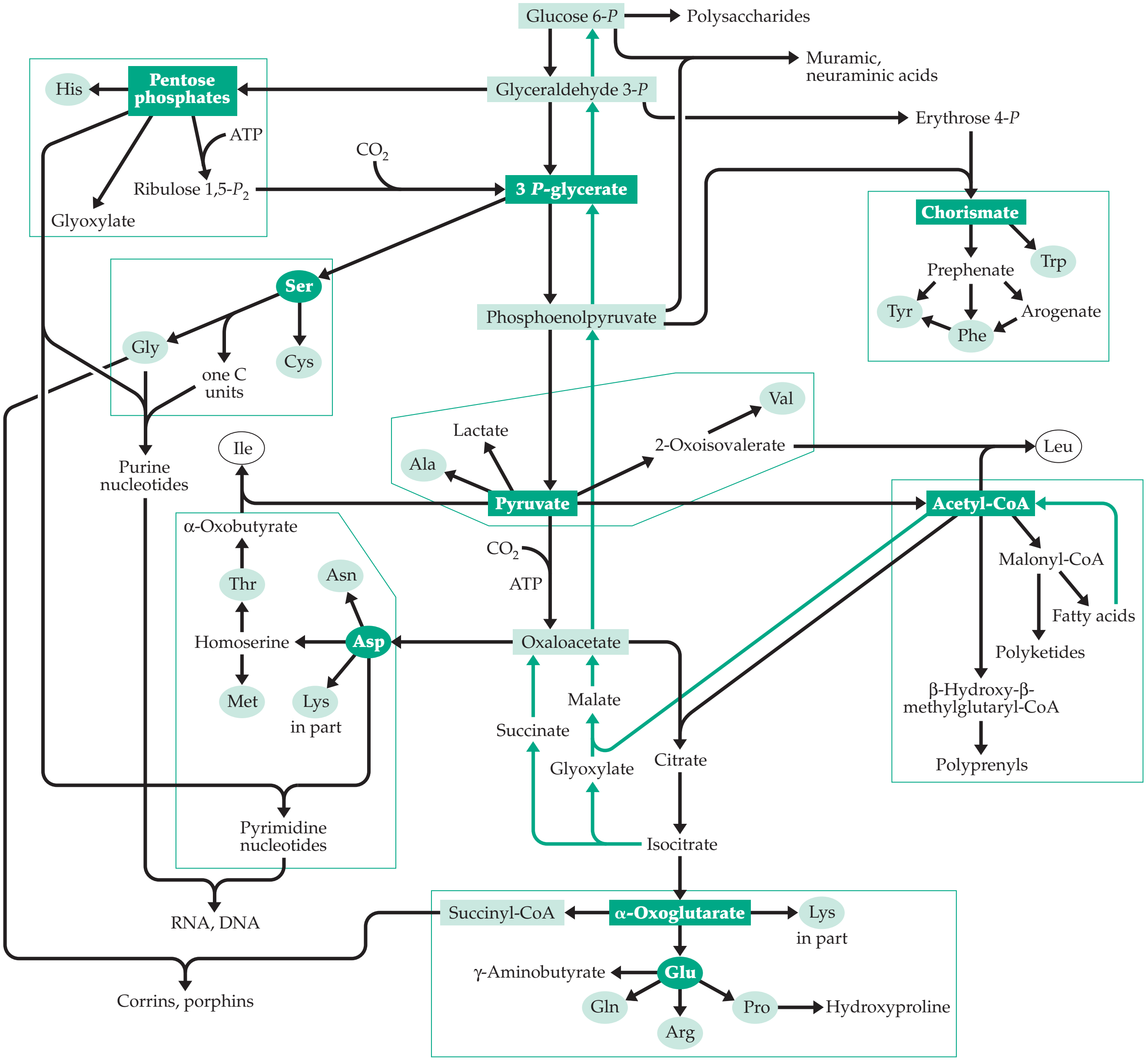
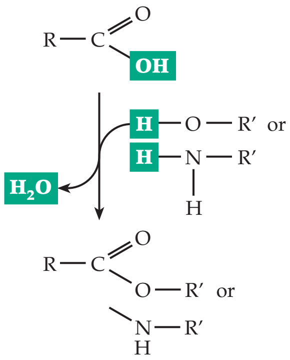
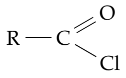

Glucose 6-P
Polysaccharides
Muramic, neuraminic acids
His
Pentose phosphates
Glyceraldehyde 3-P
Erythrose 4-P
ATP
CO2
Ribulose 1,5-P2
3 P-glycerate
Glyoxylate
Chorismate
Trp
Prephenate
Ser
Tyr
Arogenate
Phe
Phosphoenolpyruvate
Gly
Cys
one C units
Val
Lactate
Ile
Ala
2-Oxoisovalerate
Leu
Purine nucleotides
Pyruvate
Acetyl-CoA
α-Oxobutyrate
CO2
Malonyl-CoA
Asn
ATP
Thr
Fatty acids
Homoserine
Asp
Oxaloacetate
Polyketides
Met
Lys in part
Malate
β-Hydroxy-β-methylglutaryl-CoA
Succinate
Glyoxylate
Citrate
Polyprenyls
Pyrimidine nucleotides
Isocitrate
RNA, DNA
Succinyl-CoA
α-Oxoglutarate
Lys in part
γ-Aminobutyrate
Glu
Gln
Pro
Hydroxyproline
Arg
Corrins, porphins

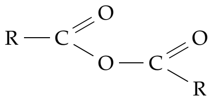
Figure 17-11 Some major biosynthetic pathways. Some key intermediates are enclosed in boxes and the 20 common amino acids of proteins are encircled. Key intermediates for each family are in shaded boxes or elipses. Green lines trace the reactions of the glyoxylate pathway and of glucogenesis.



(17-35)

However, it is often better to "activate" the carboxylic acid by conversion to an acyl chloride or an anhydride:


Acyl chloride

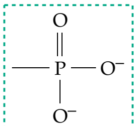
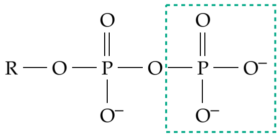
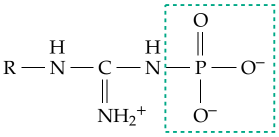
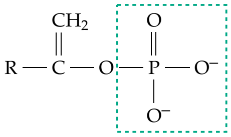
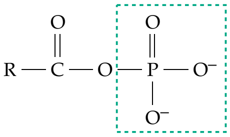
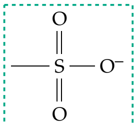
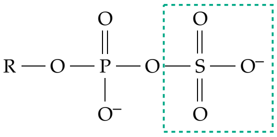
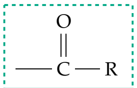
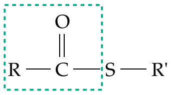
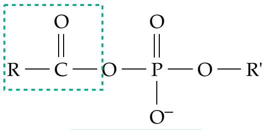
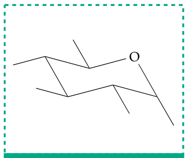
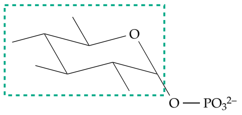
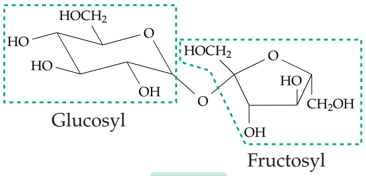
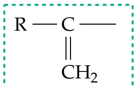
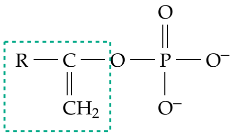
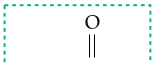
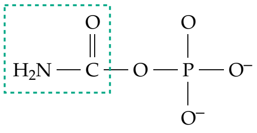
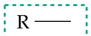
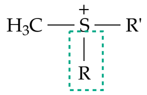

Acid anhydride

Nucleophilic attack on the carbonyl group of such a compound results in displacement of a good leaving group, Cl^- or $R-COO^-$. Nature has followed the same approach in forming from carboxylic acids **acyl phosphates** or **acyl-CoA** derivatives.

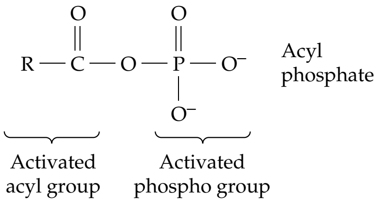
The virtue of these "activated" acyl compounds in biosynthetic reactions was considered in Chapter 12 and Table 15-1. Just as a carboxylic acid can be converted to an active acyl derivative, so other groups can be activated. ATP and other compounds with phospho groups of high group transfer potential are **active phospho** compounds. Sulfate is converted to a phosphosulfate anhydride, an **active sulfo** derivative. Sugars are converted to compounds such as glucose 1-*P* or sucrose, which contain **active glycosyl** groups. The group transfer potentials of the latter, though not as great as that of the phospho groups of ATP, are still high enough to make glucose 1-phosphate and sucrose effective glycosylating reagents. Table 17-2 lists several of the more important activated groups.

Group activation usually takes place at the expense of ATP cleavage.

TABLE 17-2
"Activated" Groups Used in Biosynthesis

Group	Typical activated forms
 **Phospho**	 Pyrophosphate  Guanidine phosphate  Enol phosphate  Acyl phosphate
 **Sulfo**	 in PAPS (3′-Phosphoadenosine-5′-phosphosulfate)
 **Acyl**	 Thioester  Acyl phosphate
 **Glycosyl**	 Glycosyl phosphate  Sucrose
 **Enoyl**	 Enol phosphate
 **Carbamoyl**	 Carbamoyl phosphate
 **Alkyl**	 *S*-Adenosylmethionine

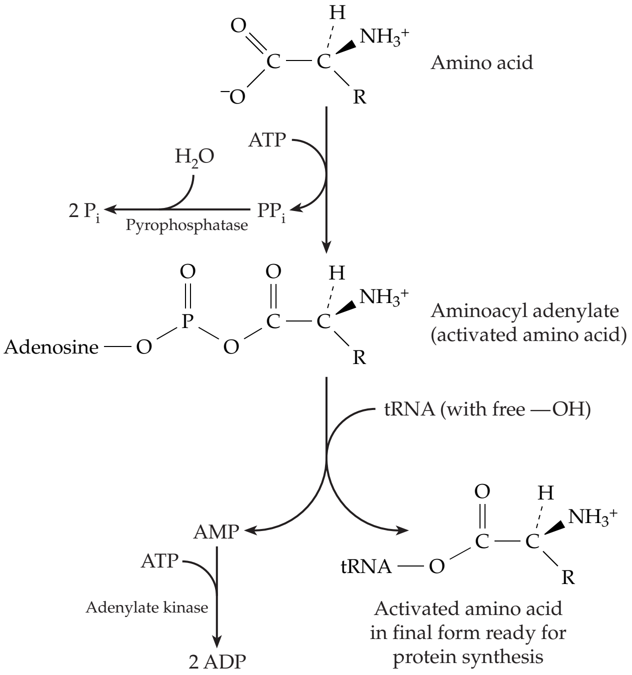
As pointed out in Chapter 12, acyl phosphates play a central role in metabolism by virtue of the fact that they contain both an activated acyl group and an activated phospho group. The high group transfer potential can be conserved in subsequent reactions *in either one group or the other* (but not in both). Thus, displacement on P by an oxygen of ADP will regenerate ATP and attack on C by an – SH will give a thioester.



Several of the other compounds in Table 17-2 can also be split in two ways to yield different activated groups, e.g., the phosphosulfate anhydride, enoyl phosphate, and carbamoyl phosphate. It is probably only through intermediates of this type that cleavage of ATP can be coupled to synthesis of activated groups. Such **common intermediates** are essential to the synthesis of ATP by substrate-level phosphorylation (Fig. 15-6).

2. Hydrolysis of Pyrophosphate

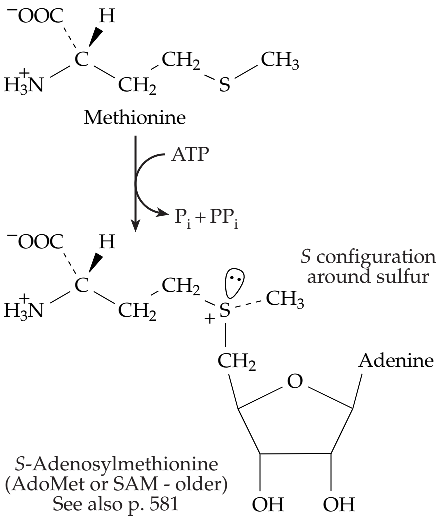
The splitting of inorganic pyrophosphate (PP_i) into two inorganic phosphate ions is catalyzed by **pyrophosphatases** (p. 636)[150a,b] that apparently occur universally. Their function appears to be simply to remove the product PP_i from reactions that produce it, shifting the equilibrium toward formation of a desired compound. An example is the formation of **aminoacyl-tRNA** molecules needed for protein synthesis. As shown in Eq. 17-36, the process requires the use of two ATP molecules to activate one amino acid. While the "spending" of two ATPs for the addition of one monomer unit to a polymer does not appear necessary from a thermodynamic viewpoint, it is frequently observed, and there is no doubt that hydrolysis of PP_i ensures that the reaction will go virtually to completion. Transfer RNAs tend to become saturated with amino acids according to Eq. 17-36 even if the concentration of free amino acid in the cytoplasm is low. On the other hand, kinetic considerations may be involved. Perhaps the biosynthetic sequence would move too slowly if it were not for the extra boost given by the removal of PP_i. Part of the explanation for the complexity may depend on control mechanisms which are only incompletely understood.



$$\text{Net: } 2\ \text{ATP} + \text{amino acid} + \text{tRNA} \longrightarrow \text{aminoacyl tRNA} + 2\ \text{ADP} + 2\ P_i \qquad (17\text{-}36)$$

In some metabolic reactions pyrophosphate esters are formed by consecutive transfer of the terminal phospho groups of two ATP molecules onto a hydroxyl group. Such esters often react with elimination of PP_i, e.g., in polymerization of prenyl units (reaction type 6B, Table 10-1; Fig. 22-1). Again, hydrolysis to P_i follows. Thus, *cleavage of pyrophosphate is a second very general method for coupling ATP cleavage to synthetic reactions*.

Although pyrophosphatases are ubiquitous, there are organisms in which PP_i is conserved by the cell and replaces ATP in several glycolytic reactions. These include *Propionibacterium*,[151,152] sulfate-reducing bacteria,[153] the photosynthetic *Rhodospirillum*, and the parasitic *Entamoeba histolytica*.[152,154] In the latter the internal concentration of PP_i is about 0.2 mM. Green plants also accumulate PP_i at concentrations of up to 0.2 mM.[155] Apparently, pyrophosphate is not always hydrolyzed immediately. Another mystery of metabolism is the accumulation of inorganic **polyphosphate** in chains of tens to many hundreds of phospho groups linked, as in pyrophosphate, by phosphoanhydride bonds. These polyphosphates are present in many bacteria, including *E. coli*, and also in fungi, plants, and animals.[156–156b] They constitute a store of energy as well as of phosphate. Various other functions have also been proposed. A polyphosphate kinase transfers a terminal phospho group from polyphosphate chains onto ADP to form ATP. This source of metabolic energy is evidently essential to the ability of *Pseudomonas aeruginosa* to form biofilms.[156a] Both endophosphatases and exophosphatases, of uncertain function, can degrade the chains hydrolytically. An exophosphatase from *E. coli* can completely hydrolyze polyphosphate chains of 1000 units processively without release of intermediates.[156b]

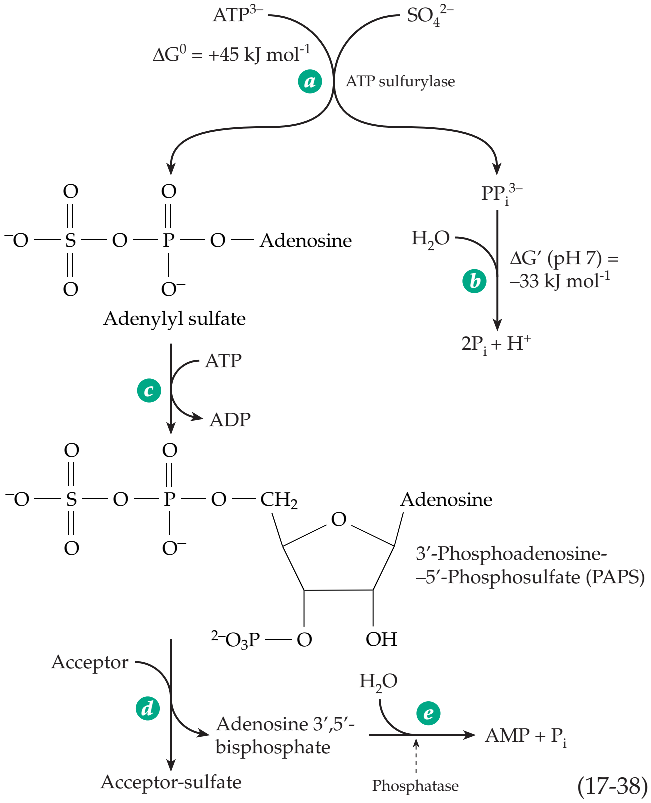

Methionine
ATP
$P_i + PP_i$
S configuration around sulfur
Adenine
S-Adenosylmethionine
(AdoMet or SAM - older)
See also p. 581

(17-37)

In a few instances group activation is coupled to cleavage of ATP at C-5' presumably with formation of bound tripolyphosphate (PPP_i). The latter is hydrolyzed to P_i and PP_i and ultimately to *three* molecules of P_i. An example is the formation of *S*-adenosylmethionine[157] shown in Eq. 17-37. The reaction is a displacement on the 5'-methylene group of ATP by the sulfur atom of methionine. While the initial product may be enzyme-bound PPP_i, it is P_i and PP_i that are released from the enzyme, the P_i arising from the terminal phosphorus (P_γ) of ATP.[157] The *S*-adenosylmethionine formed has the *S* configuration around the sulfur.[158]

3. Coupling by Phosphorylation and Subsequent Cleavage by a Phosphatase

A third general method for coupling the hydrolysis of ATP to drive a synthetic sequence is to transfer the terminal phospho group from ATP to a hydroxyl group *somewhere* on a substrate. Then, after the substrate has undergone a synthetic reaction, the phosphate is removed by action of a phosphatase. For example, in the activation of sulfate (Eq. 17-38),[159] the overall standard Gibbs energy change for steps *a* (catalyzed by **ATP sulfurylase**[160,161]) and *b* is distinctly positive (+12 kJ mol^{-1}). The equilibrium concentration of adenylyl sulfate formed in this group activation process is extremely low. Nature's solution to this problem is to spend another molecule of ATP to phosphorylate the 3' – OH of adenosine phosphosulfate. As the latter is formed, it is converted to 3'-phosphoadenosine-5'-phosphosulfate (Eq. 17-38, step *c*) by a kinase, which is often part of a bifunctional enzyme that also contains the active site of ATP sulfurylase.[162–163a] Since the equilibrium in this step lies far toward the right, the product accumulates in a substantial concentration (up to 1 mM in cell-free systems) and serves as the active sulfo group donor in formation of sulfate esters. The reaction cycle is completed by two more reactions. In Eq. 17-38, step *d*, the sulfo group is transferred to an acceptor, and in step *e* the extra phosphate group is removed from adenosine 3',5'-bisphosphate by a specific phosphatase. Since the reconversion of AMP to ADP requires expenditure of still a third high-energy linkage of ATP, the overall process makes use of three high-energy phosphate linkages for formation of one sulfate ester.

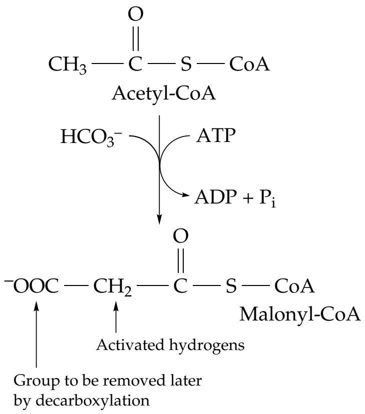

ATP^{3-}
SO_4^{2-}
ΔG^0 = +45 kJ mol^{-1}
a
ATP sulfurylase
PP_i^{3-}
H_2O
b
ΔG′ (pH 7) = –33 kJ mol^{-1}
$2P_i + H^+$
Adenosine
Adenylyl sulfate
ATP
c
ADP
Adenosine
3′-Phosphoadenosine-–5′-Phosphosulfate (PAPS)
Acceptor
d
Adenosine 3′,5′-bisphosphate
Acceptor-sulfate
H_2O
e
AMP + P_i
Phosphatase

(17-38)

An analogous use of ATP is found in photosynthetic reduction of carbon dioxide in which ATP phosphorylates ribulose 5-*P* to ribulose bisphosphate and the phosphate groups are removed later by phosphatase action on fructose bisphosphate and sedoheptulose bisphosphate (Section J,2). Phosphatases involved in synthetic pathways usually have a high substrate specificity and are to be distinguished from nonspecific phosphatases which are essentially digestive enzymes (Chapter 12).

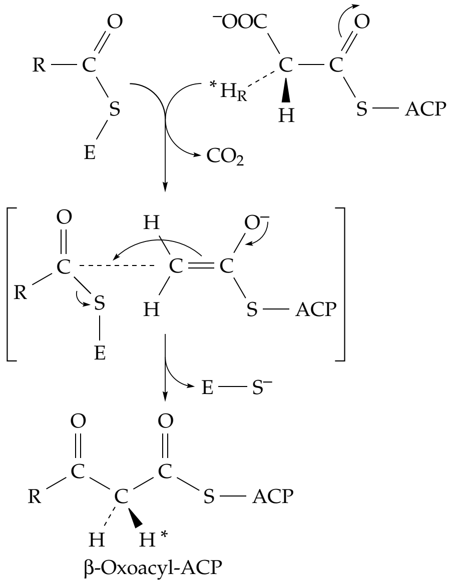
4. Carboxylation and Decarboxylation: Synthesis of Fatty Acids

A fourth way in which cleavage of ATP can be coupled to biosynthesis was recognized in about 1958 when Wakil and coworkers discovered that synthesis of fatty acids in animal cytoplasm is stimulated by carbon dioxide. However, when $^{14}CO_2$ was used in the experiment no radioactivity appeared in the fatty acids formed. Rather, it was found that acetyl-CoA was carboxylated to **malonyl-CoA** in an ATP- and biotin-requiring process (Eq. 17-39; see also Chapter 13). The carboxyl group formed in this reaction is later converted back to CO_2 in a decarboxylation (Fig. 17-12).


O
CH3 — C — S — CoA
Acetyl-CoA
HCO3−
ATP
ADP + Pi
O
−OOC — CH2 — C — S — CoA
Malonyl-CoA
Activated hydrogens
Group to be removed later
by decarboxylation

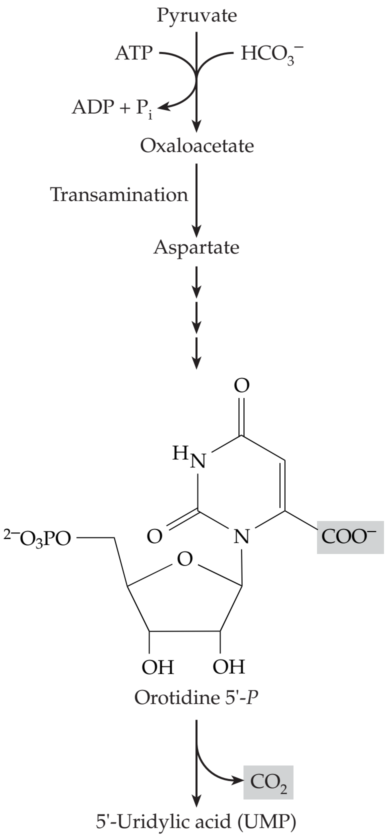
(17-39)

We know now that in most bacteria and green plants both an acetyl group of acetyl-CoA and a malonyl group of malonyl-CoA are transferred (steps *a* and *d* of Fig. 17-12) to the sulfur atoms of the phosphopantetheine groups of a low-molecular-weight **acyl carrier protein** (ACP; Chapter 14). The malonyl group of the malonyl-ACP is then condensed (step *f* of Fig. 17-12) with an acetyl group, which has been transferred from acetyl-ACP onto a thiol group of the enzyme (E in Eq. 17-40). The enolate anion indicated in this equation is generated by decarboxylation of the malonyl-ACP. It is this decarboxylation that drives the reaction to completion and, in effect, links C–C bond formation to the cleavage of the ATP required for the carboxylation step. A related sequence involving multifunctional proteins is used by animals and fungi[164] (Section J,6).


R — C
O
S
E
−OOC
O
*HR
C — C
H
S — ACP
CO2
O
H
O−
R
C
C=C
S
H
S — ACP
E
E — S−
O
O
R
C
C
C
S — ACP
H
H*
β-Oxoacyl-ACP

(17-40)

Carboxylation followed by a later decarboxylation is an important pattern in other biosynthetic pathways, too. Sometimes the decarboxylation follows the carboxylation by many steps. For example, pyruvate (or PEP) is converted to uridylic acid (Eq. 17-41; details are shown in Fig. 25-14):


Pyruvate
ATP
HCO3−
ADP + Pi
Oxaloacetate
Transamination
Aspartate
O
HN
2−O3PO
O
N
COO−
O
OH
OH
Orotidine 5'-P
CO2
5'-Uridylic acid (UMP)

(17-41)

I. Reducing Agents for Biosynthesis

Still another difference between biosynthesis of fatty acids and oxidation (in mammals) is that the former has an absolute requirement for NADPH (Fig. 17-12) while the latter requires NAD^+ and flavoproteins (Fig. 17-1). This fact, together with many other observations, has led to the generalization that *biosynthetic reduction reactions usually require NADPH rather than NADH*. Many measurements have shown that in the cytosol of eukaryotic cells the ratio $[NADPH]/[NADP^+]$ is high, whereas the ratio $[NADH]/[NAD^+]$ is low. Thus, the $NAD^+/NADH$ system is kept highly oxidized, in line with the role of NAD^+ as a principal biochemical oxidant, while the $NADP^+/NADPH$ system is kept reduced.

The use of NADPH in step *g* of Fig. 17-12 ensures that significant amounts of the β-oxoacyl-ACP derivative are reduced to the alcohol. Another difference between β oxidation and biosynthesis is that the alcohol formed in this reduction step in the biosynthetic process

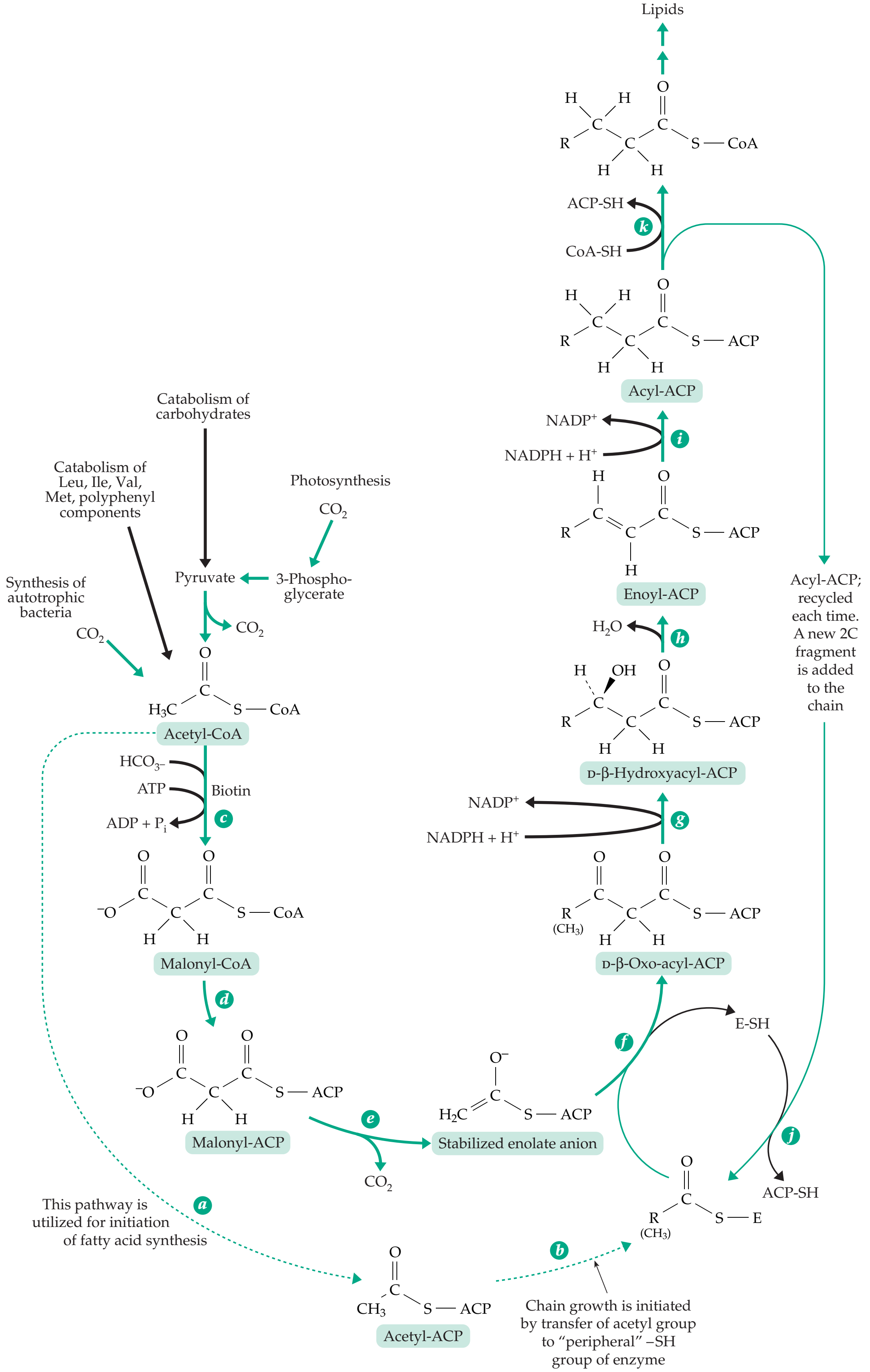

Lipids
CoA
ACP-SH
k
CoA-SH
ACP
Acyl-ACP
NADP+
i
NADPH + H+
ACP
Enoyl-ACP
H_2O
h
OH
ACP
D-β-Hydroxyacyl-ACP
NADP+
g
NADPH + H+
(CH3)
D-β-Oxo-acyl-ACP
Catabolism of carbohydrates
Catabolism of Leu, Ile, Val, Met, polyphenyl components
Photosynthesis
CO_2
Pyruvate
3-Phosphoglycerate
Synthesis of autotrophic bacteria
CO_2
CO_2
H_3C
S — CoA
Acetyl-CoA
HCO_3^-
ATP
Biotin
ADP + P_i
c
S — CoA
Malonyl-CoA
d
S — ACP
Malonyl-ACP
e
CO_2
H_2C
S — ACP
Stabilized enolate anion
f
E-SH
j
ACP-SH
S — E
(CH3)
Acyl-ACP; recycled each time. A new 2C fragment is added to the chain
This pathway is utilized for initiation of fatty acid synthesis
a
b
CH_3
S — ACP
Acetyl-ACP
Chain growth is initiated by transfer of acetyl group to "peripheral" –SH group of enzyme

Figure 17-12 The reactions of cytoplasmic biosynthesis of saturated fatty acids. Compare with pathway of β oxidation (Fig. 17-1).

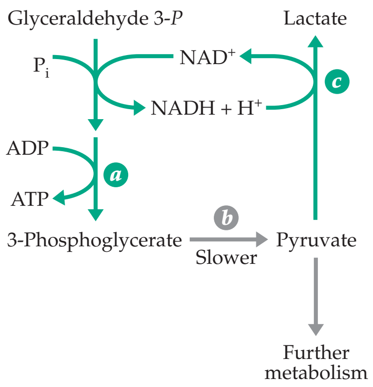
has the D configuration while the corresponding alcohol in β oxidation has the L configuration.

1. Reversing an Oxidative Step with a Strong Reducing Agent

The second reduction step in biosynthesis of fatty acids in the rat liver (step *i*) also required NADPH. The corresponding step in β oxidation utilizes FAD, but NADPH is a stronger reducing agent than $FADH_2$. Therefore, use of a reduced pyridine nucleotide again provides a thermodynamic advantage in pushing the reaction in the biosynthetic direction. Interesting variations have been observed among different species. For example, fatty acid synthesis in the rat requires only NADPH, but the multienzyme complexes from *Mycobacterium phlei*, *Euglena gracilis*, and the yeast *Saccharomyces cerevisiae* all give much better synthesis with a mixture of NADPH and NADH than with NADPH alone.[165] Apparently, NADPH is required in step *g* and NADH in step *i*. This seems reasonable because the equilibrium in step *i* lies far toward the product formation, and NADH at a very low concentration could carry out the reduction.

2. Regulation of the State of Reduction of the NAD and NADP Systems

The ratio $[NAD^+]/[NADH]$ appears to be maintained at a relatively constant value and in equilibrium with a series of different reduced and oxidized substrate pairs. Thus, it was observed that in the cytoplasm of rat liver cells, the dehydrogenations catalyzed by lactate dehydrogenase, *sn*-glycerol 3-phosphate dehydrogenase, and malate dehydrogenase are all at equilibrium with the same ratio of $[NAD^+]/[NADH]$.[166] In one experiment rat livers were removed and frozen in less than 8 s by "freeze-clamping" (Section L,2) and the concentrations of different components of the cytoplasm determined[167]; the ratio $[NAD^+]/[NADH]$ was found to be 634, while the ratio of [lactate]/[pyruvate] was 14.2. From these values an apparent equilibrium constant for reaction *c* of Eq. 17-42 was calculated as $K_c' = 9.0 \times 10^3$. The known equilibrium constant for the reaction (from *in vitro* experiments) is 8.8×10^3 (Eq. 17-43). In a similar way it was shown that several other dehydrogenation reactions are nearly at equilibrium. This conclusion has been confirmed more recently by NMR observations.[168]



(17-42)

$$K_c'\ (\text{pH 7, 38°C}) = \frac{[\text{lactate}]}{[\text{pyruvate}]} \times \frac{[NAD^+]}{[NADH]} = 8.8 \times 10^3 \qquad (17\text{-}43)$$

Now consider Eq. 17-42, step *a*, the ADP- and P_i-requiring oxidation of glyceraldehyde 3-phosphate (Fig. 15-6). Experimental measurements indicated that this reaction is also at equilibrium in the cytoplasm. In one series of experiments the measured phosphorylation state ratio $[ATP]/[ADP][P_i]$ was 709, while the ratio [3-phosphoglycerate]/[glyceraldehyde 3-phosphate] was 55.5. The overall equilibrium constant for Eq. 17-42*a* is given by Eq. 17-44. That calculated from known equilibrium constants is 60.

$$K_a'\ (\text{pH 7, 38°C}) = \frac{[ATP]}{[ADP][P_i]} \times \frac{[\text{3-phosphoglycerate}]}{[\text{glyceraldehyde phosphate}]} \times \frac{[NADH]}{[NAD^+]} = 709 \times 55.5 \times 1/634 = 62 \qquad (17\text{-}44)$$

From these data Krebs and Veech concluded that the oxidation state of the NAD system is determined largely by the phosphorylation state ratio of the adenylate system.[169] If the ATP level is high the equilibrium in Eq. 17-42*a* will be reached at a higher $[NAD^+]/[NADH]$ ratio and lactate may be oxidized to pyruvate to adjust the [lactate]/[pyruvate] ratio.

It is important not to confuse the reactions of Eq. 17-42 as they occur in an aerobic cell with the tightly coupled pair of redox reactions in the homolactate fermentation (Fig. 10-3; Eq. 17-19). The reactions of steps *a* and *c* of Eq. 17-42 are essentially at equilibrium, but the reaction of step *b* may be relatively slow. Furthermore, pyruvate is utilized in many other metabolic pathways and ATP is hydrolyzed and converted to ADP through innumerable processes taking place within the cell. Reduced NAD does not cycle between the two enzymes in a stoichiometric way and the "reducing equivalents" of NADH formed are, in large measure, transferred to the mitochondria. The proper view of the reactions of Eq. 17-42 is that the redox pairs represent a kind of **redox buffer system** that poises the $NAD^+/NADH$ couple at a ratio appropriate for its metabolic function.

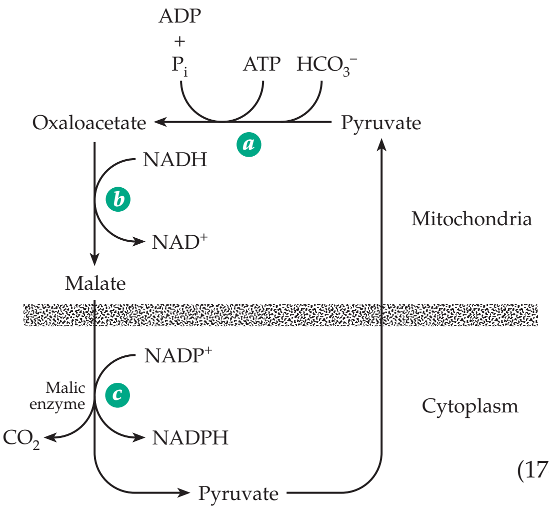
Somewhat surprisingly, within the mitochondria the ratio [NAD$^+$]/[NADH] is 100 times lower than in the cytoplasm. Even though mitochondria are the site of oxidation of NADH to NAD$^+$, the intense catabolic activity occurring in the β oxidation pathway and the citric acid cycle ensure extremely rapid production of NADH. Furthermore, the reduction state of NAD is apparently buffered by the low potential of the β-hydroxybutyrate–acetoacetate couple (Chapter 18, Section C,2). Mitochondrial pyridine nucleotides also appear to be at equilibrium with glutamate dehydrogenase.[169]

How is the cytoplasmic [NADPH]/[NADP$^+$] ratio maintained at a value higher than that of [NADH]/[NAD$^+$]? Part of the answer is from operation of the pentose phosphate pathway (Section E,3). The reactions of Eq. 17-12, if they attained equilibrium, would give a ratio of cytosolic [NADPH]/[NADP$^+$] > 2000 at 0.05 atm CO_2. Compare this with the ratio 1/634 for [NADH/[NAD$^+$] deduced from the observation on the reactions of Eq. 17-42.

Consider also the following **transhydrogenation** reaction (Eq. 17-45):

$$\text{NADH} + \text{NADP}^+ \rightarrow \text{NAD}^+ + \text{NADPH} \qquad (17\text{-}45)$$

There are soluble enzymes that catalyze this reaction for which *K* equals ~ 1. Within mitochondria an energy-linked system (Chapter 18) involving the membrane shifts the equilibrium to favor NADPH. However, within the cytoplasm, the reaction of Eq. 17-45 is driven by coupling ATP cleavage to the transhydrogenation via carboxylation followed by eventual decarboxylation. One cycle that accomplishes this is given in Eq. 17-46. The first step (step *a*) is ATP-dependent carboxylation of pyruvate to oxaloacetate, a reaction that occurs only within mitochondria (Eq. 14-3). Oxaloacetate can be reduced by malate dehydrogenase using NADH (Eq. 17-46, step *b*), and the resulting malate can be exported from the mitochondria. In the cytoplasm the malate is oxidized to pyruvate, with decarboxylation, by the malic enzyme (Eq. 13-45). The malic enzyme (Eq. 17-46, step *c*) is specific for NADP$^+$, is very active, and also appears to operate at or near equilibrium within the cytoplasm. On this basis, using known equilibrium constants, it is easy to show that the ratio [NADPH]/[NADP$^+$] will be ~ 10^5 times higher at equilibrium than the ratio [NADH]/[NAD$^+$].[169,170]



(17-46)

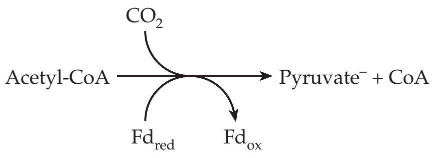
Since NADPH is continuously used in biosynthetic reactions, and is thereby reconverted to NADP$^+$, the cycle of Eq. 17-46 must operate continuously. As in Eq. 17-42, a true equilibrium does not exist but steps *b* and *c* are both essentially at equilibrium. These equilibria, together with those of Eq. 17-42 for the NAD system, ensure the correct redox potential of both pyridine nucleotide coenzymes in the cytoplasm.

Malate is not the only form in which C_4 compounds are exported from mitochondria. Much oxaloacetate is combined with acetyl-CoA to form citrate; the latter leaves the mitochondria and is cleaved by the ATP-dependent citrate-cleaving enzymes (Eq. 13-39). This, in effect, exports both acetyl-CoA (needed for lipid synthesis) and oxaloacetate which is reduced to malate within the cytoplasm. Alternatively, oxaloacetate may be transaminated to aspartate. The aspartate, after leaving the mitochondria, may be converted in another transamination reaction back to oxaloacetate. All of these are part of the nonequilibrium process by which C_4 compounds diffuse out of the mitochondria before completing the reaction sequence of Eq. 17-46 and entering into other metabolic processes. Note that the reaction of Eq. 17-46 leads to the *export* of reducing equivalents from mitochondria, the opposite of the process catalyzed by the malate–aspartate shuttle which is discussed in Chapter 18 (Fig. 18-18). The two processes are presumably active under different conditions.

While the difference in the redox potential of the two pyridine nucleotide systems is clear-cut in mammalian tissues, in *E. coli* the apparent potentials of the two systems are more nearly the same.[171]

3. Reduced Ferredoxin in Reductive Biosynthesis

Both the NAD$^+$ and NADP$^+$ systems have standard electrode potentials $E°'$ (pH 7) of –0.32 V. However, because of the differences in concentration ratios, the NAD$^+$ system operates at a less negative potential (–0.24 V) and the NADP$^+$ system at a more negative potential (–0.38 V) within the cytoplasm of eukaryotes. In green plants and in many bacteria a still more powerful reducing agent is available in the form of reduced ferredoxin. The value of $E°'$ (pH 7) for clostridial ferredoxin is –0.41 V, corresponding to a Gibbs energy change for the two-electron reduction of a substrate ~18 kJ mol^{-1} more negative than the corresponding value of $\Delta G'$ for reduction by NADPH. Using reduced ferredoxin (Fd) some photosynthetic bacteria and anaerobic bacteria are able to carry out reductions that are virtually impossible with the pyridine nucleotide system. For example, pyruvate and 2-oxoglutarate can be formed from acetyl-CoA (Eq. 15-35) and succinyl-CoA, respectively (Eq. 17-47).[172–173a] In our bodies the reaction of Eq. 17-47, with NAD^+ as the oxidant, goes only in the opposite direction and is essentially irreversible.



$\Delta G'$ (pH 7) = +17 kJ mol^{-1} (17-47)

J. Constructing the Monomer Units

Now let us consider the synthesis of the monomeric units from which biopolymers are made. How can simple one-carbon compounds such as CO_2 and formic acid be incorporated into complex carbon compounds? How can carbon chains grow in length or be shortened? How are branched chains and rings formed?

1. Carbonyl Groups in Chain Formation and Cleavage

Except for some vitamin B_{12}-dependent reactions, the cleavage or formation of carbon–carbon bonds usually depends upon the participation of carbonyl groups. For this reason, carbonyl groups have a central mechanistic role in biosynthesis. The activation of hydrogen atoms β to carbonyl groups permits β condensations to occur during biosynthesis. Aldol or Claisen condensations require the participation of two carbonyl compounds. Carbonyl compounds are also essential to thiamin diphosphate-dependent condensations and the aldehyde pyridoxal phosphate is needed for most C–C bond cleavage or formation within amino acids.

Because of the importance of carbonyl groups to the mechanism of condensation reactions, much of the assembly of either straight-chain or branched-carbon skeletons takes place between compounds in which the average oxidation state of the carbon atoms is similar to that in carbohydrates (or in formaldehyde, H_2CO). The diversity of chemical reactions possible with compounds at this state of oxidation is a maximum, a fact that may explain why carbohydrates and closely related substances are major biosynthetic precursors and why the average state of oxidation of the carbon in most living things is similar to that in carbohydrates.[174] This fact may also be related to the presumed occurrence of formaldehyde as a principal component of the earth's atmosphere in the past and to the ability of formaldehyde to condense to form carbohydrates.

In Fig. 17-13 several biochemicals have been arranged according to the oxidation state of carbon. Most of the important biosynthetic intermediates lie within ± 2 electrons per carbon atom of the oxidation state of carbohydrates. As the chain length grows, they tend to fall even closer. It is extremely difficult to move through enzymatic processes between 2C, 3C, and 4C compounds (i.e., vertically in Fig. 17-13) except at the oxidation level of carbohydrates or somewhat to its right, at a slightly higher oxidation level. On the other hand, it is often possible to move horizontally with ease using oxidation–reduction reactions. Thus, fatty acids are assembled from acetate units, which lie at the same oxidation state as carbohydrates and, after assembly, are reduced.

Among compounds of the same overall oxidation state, e.g., acetic acid and sugars, the oxidation states of individual carbon atoms can be quite different. Thus, in a sugar every carbon atom can be regarded as immediately derived from formaldehyde, but in acetic acid one end has been oxidized to a carboxyl group and the other has been reduced to a methyl group. Such internal oxidation–reduction reactions play an important role in the chemical manipulations necessary to assemble the carbon skeletons needed by a cell. Decarboxylation is a feature of many biosynthetic routes. Referring again to Fig. 17-13, notice that many of the biosynthetic intermediates such as pyruvate, oxoglutarate, and oxaloacetate are more oxidized than the carbohydrate level. However, their decarboxylation products, which become incorporated into the compounds being synthesized, are closer to the oxidation level of carbohydrates.

2. Starting with CO_2

There are three known pathways by which autotrophic organisms can use CO_2 to synthesize triose phosphates or 3-phosphoglycerate, three-carbon compounds from which all other biochemical substances can be formed.[175–177] The first of these is the **reductive tricarboxylic cycle.** This is a reversal of the oxidative citric acid cycle in which reduced ferredoxin is used as a reductant in the reaction of Eq. 17-47 to incorporate CO_2 into pyruvate. Succinyl-CoA can react with CO_2 in the same type of reaction to form 2-oxoglutarate, accomplishing the reversal of the only irreversible step in the citric acid cycle. Using these reactions photosynthetic bacteria and some anaerobes that can generate a high ratio of reduced to oxidized ferredoxin carry out the reductive tricarboxylic acid cycle. Together with

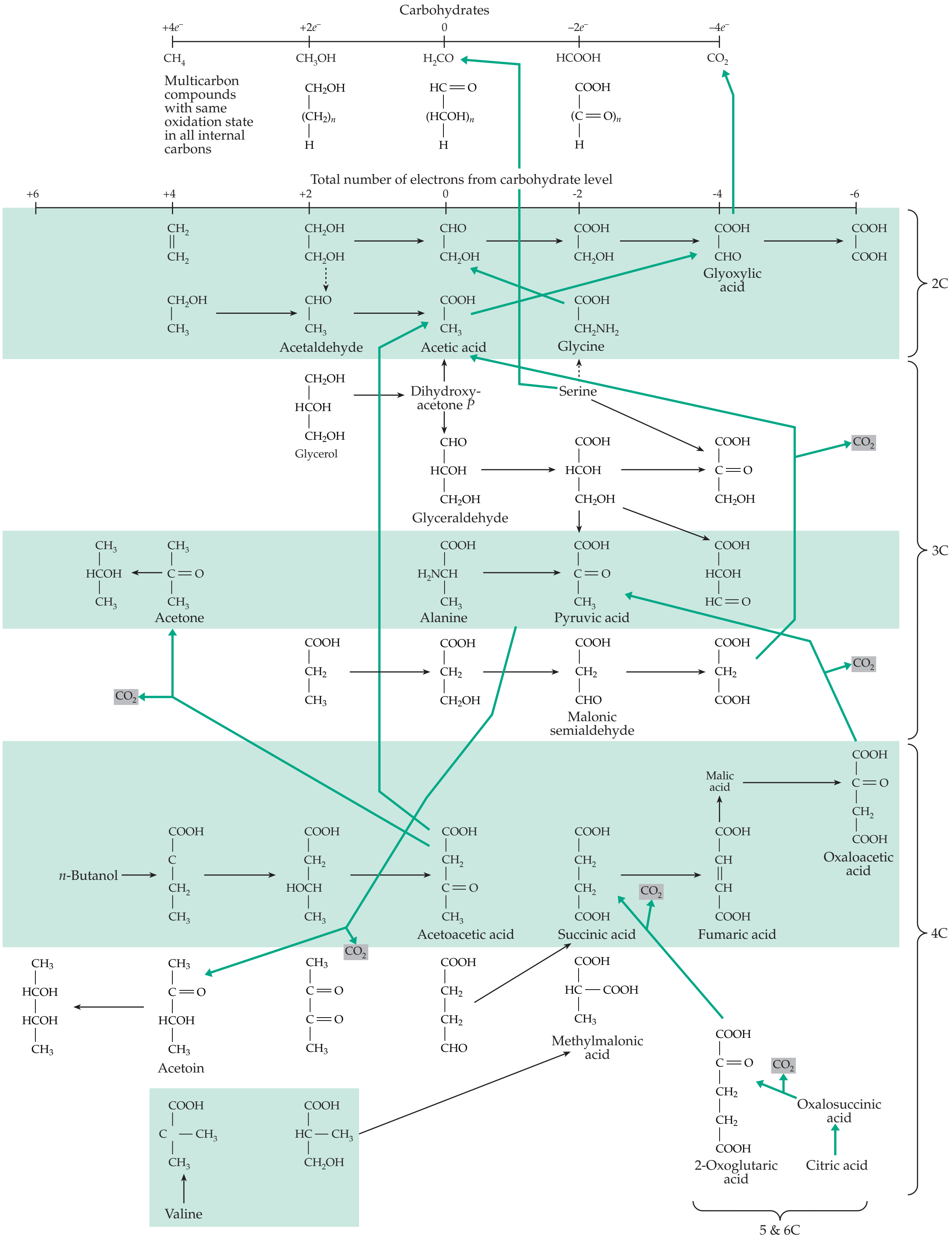

Carbohydrates
+4e⁻
+2e⁻
0
−2e⁻
−4e⁻
CH₄
CH₃OH
H₂CO
HCOOH
CO₂
Multicarbon compounds with same oxidation state in all internal carbons
CH₂OH (CH₂)n H
HC=O (HCOH)n H
COOH (C=O)n H
Total number of electrons from carbohydrate level
+6
+4
+2
0
-2
-4
-6
2C
Acetaldehyde
Acetic acid
Glycine
Glyoxylic acid
Glycerol
Dihydroxy-acetone P
Serine
Glyceraldehyde
Acetone
Alanine
Pyruvic acid
Malonic semialdehyde
3C
CO₂
n-Butanol
Malic acid
Oxaloacetic acid
Acetoacetic acid
Succinic acid
Fumaric acid
Acetoin
Methylmalonic acid
Valine
2-Oxoglutaric acid
Oxalosuccinic acid
Citric acid
4C
5 & 6C

Figure 17-13 Some biochemical compounds arranged in order of average oxidation state of the carbon atoms and by carbon-chain lengths. Black horizontal arrows mark some biological interconversions among compounds with the same chain length, while green lines show changes in chain length and are often accompanied by decarboxylation.

Eq. 17-47, the cycle provides for the complete synthesis of pyruvate from CO_2.[178,179]

A quantitatively much more important pathway of CO_2 fixation is the **reductive pentose phosphate pathway** (ribulose bisphosphate cycle or **Calvin–Benson cycle**; Fig. 17-14). This sequence of reactions, which takes place in the chloroplasts of green plants and also in many chemiautotrophic bacteria, is essentially a way of reversing the oxidative pentose phosphate pathway (Fig. 17-8). The latter accomplishes the complete oxidation of glucose or of glucose 1-phosphate by $NADP^+$ (Eq. 17-48):

$$\begin{aligned}&\text{Glucose 1-}P^{2-} + ATP^{4-} + 8\,H_2O + 12\,NADP^+ \rightarrow\\&\quad 6\,CO_2 + 12\,NADPH + ADP^{3-} + 2HPO_4^{\,2-} + 13\,H^+\\&\quad \Delta G'\ (\text{pH 7}) = -299\ \text{kJ mol}^{-1}\end{aligned} \qquad (17\text{-}48)$$

It would be almost impossible for a green plant to fix CO_2 using photochemically generated NADPH by an exact reversal of Eq. 17-48 because of the high positive Gibbs energy change. To solve this thermodynamic problem the reductive pentose phosphate pathway has been modified in a way that couples ATP cleavage to the synthesis.

The **reductive carboxylation** system is shown within the green shaded box of Fig. 17-14. Ribulose 5-phosphate is the starting compound and in the first step one molecule of ATP is expended to form **ribulose 1,5-bisphosphate**. The latter is carboxylated and cleaved to two molecules of 3-phosphoglycerate. This reaction was discussed in Chapter 13. The reductive step (step *c*) of the system employs NADPH together with ATP. Except for the use of the NADP system instead of the NAD system, it is exactly the reverse of the glyceraldehyde phosphate dehydrogenase reaction of glycolysis. Looking at the first three steps of Fig. 17-14 it is clear that in the reductive pentose phosphate pathway three molecules of ATP are utilized for each CO_2 incorporated. On the other hand, in the oxidative direction *no* ATP is generated by the operation of the pentose phosphate pathway.

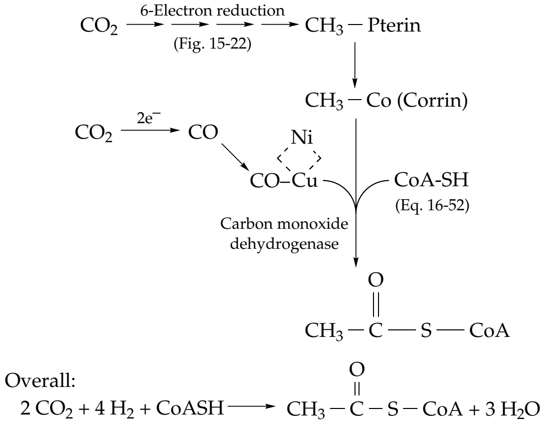
The reactions enclosed within the shaded box of Fig. 17-14 do not give the whole story about the coupling mechanism. A phospho group was transferred from ATP in step *a* and to complete the hydrolysis it must be removed in some future step. This is indicated in a general way in Fig. 17-14 by the reaction steps *d*, *e*, and *f*. Step *f* represents the action of specific phosphatases that remove phospho groups from the seven-carbon sedoheptulose bisphosphate and from fructose bisphosphate. In either case the resulting ketose monophosphate reacts with an aldose (via transketolase, step *g*) to regenerate ribulose 5-phosphate, the CO_2 acceptor. The overall reductive pentose phosphate cycle (Fig. 17-14B) is easy to understand as a reversal of the oxidative pentose phosphate pathway in which the oxidative decarboxylation system of Eq. 17-12 is replaced by the reductive carboxylation system of Fig. 17-14A. The scheme as written in Fig. 17-14B shows the incorporation of three molecules of CO_2. The reductive carboxylation system operates three times with a net production of one molecule of triose phosphate. As with other biosynthetic cycles, any amount of any of the intermediate metabolites may be withdrawn into various biosynthetic pathways without disruption of the flow through the cycle.

The overall reaction of carbon dioxide reduction in the Calvin–Benson cycle (Fig. 17-14) becomes

$$\begin{aligned}&6\,CO_2 + 12\,NADPH + 18\,ATP^{4-} + 11\,H_2O \rightarrow\\&\quad \text{glucose-1-}P^{2-} + 12\,NADP^+ + 18\,ADP^{3-}\\&\qquad\qquad 17\,HPO_4^{\,2-} + 6\,H^+\end{aligned} \qquad (17\text{-}49)$$

The Gibbs energy change $\Delta G'$ (pH 7) is now −357 kJ mol^{-1} instead of the +299 kJ mol^{-1} required to reverse the reaction of Eq. 17-48.

The third pathway for reduction of CO_2 to acetyl-CoA is utilized by acetogenic bacteria, by methanogens, and probably by sulfate-reducing bacteria.[179–181] This **acetyl-CoA pathway** (or **Wood–Ljungdahl pathway**) involves reduction by H_2 of one of the two molecules of CO_2 to the methyl group of methyltetrahydromethanopterin in methanogens and of methyltetra-hydrofolate in acetogens. The pathway utilized by methanogens is illustrated in Fig. 15-22.[182–184] A similar process utilizing H_2 as the reductant is employed by acetogens.[179,185–188a] In both cases a methyl corrinoid is formed and its methyl group is condensed with a molecule of carbon monoxide bound to a copper ion in a Ni–Cu cluster.[189a,b] The resulting acetyl group is transferred to a molecule of coenzyme A as illustrated in Eq. 16-52.[189] The bound CO is formed by reduction of CO_2, again using H_2 as the reductant.[190] The overall reaction for acetyl-CoA synthesis is given by Eq. 17-50. Conversion of acetyl-CoA to pyruvate via Eq. 17-47 leads into the glucogenic pathway.



$$2\,CO_2 + 4\,H_2 + CoASH \longrightarrow CH_3-\overset{\overset{\large O}{\|}}{C}-S-CoA + 3\,H_2O \qquad (17\text{-}50)$$

An alternative pathway by which some acetogenic bacteria form acetate is via reversal of the glycine decarboxylase reaction of Fig. 15-20. Methylene-THF is formed by reduction of CO_2, and together with NH_3 and CO_2 a lipoamide group of the enzyme and PLP forms glycine. The latter reacts with a second methylene-THF to form serine, which can be deaminated to pyruvate and assimilated. Methanogens may use similar pathways but ones that involve methanopterin (Fig. 15-17).[191]

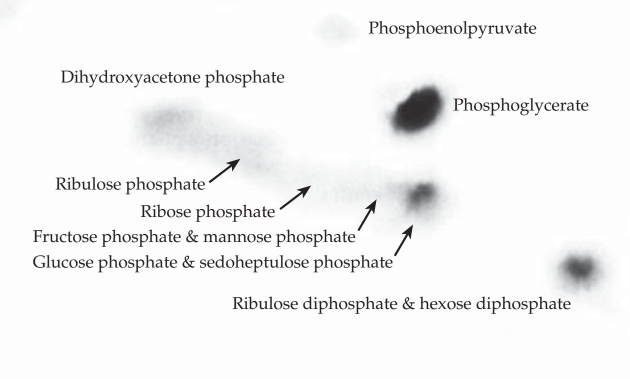
3. Biosynthesis from Other Single-Carbon Compounds

Various bacteria and fungi are able to subsist on such one-carbon compounds as methane, methanol, methylamine, formaldehyde, and formate.[192–197] Energy is obtained by oxidation to CO_2. **Methylotrophic bacteria** initiate oxidation of methane by hydroxylation (Chapter 18) and dehydrogenate the resulting methanol or exogenous methanol using the PPQ cofactor (Eq. 15-51).[198] Further dehydrogenation to formate and of formate to CO_2 via formate dehydrogenase (Eq. 16-63) completes the process. Some methylotrophic bacteria incorporate CO_2 for biosynthetic purposes via the ribulose bisphosphate (Calvin–Benson) cycle but many use pathways that begin with formaldehyde (or methylene-THF). Others employ variations of the reductive pentose phosphate pathway to convert formaldehyde to triose phosphate. In one of these, the **ribulose monophosphate cycle** or Quayle cycle,[192,193] ribulose 5-*P* undergoes an aldol condensation with formaldehyde to give a 3-oxo-hexulose 6-phosphate (Eq. 17-51, step *a*). The latter is isomerized to fructose 6-*P* (Eq. 17-51, step *b*). If this equation is applied to the

BOX 17-E ^{14}C AND THE CALVIN–BENSON CYCLE

The chemical nature of photosynthesis had intrigued chemists for decades but little was learned about the details until radioactive ^{14}C became available. Discovered in 1940 by Ruben and Kamen, the isotope was available in quantity by 1946 as a product of nuclear reactors. Initial studies of photosynthesis had been conducted by Ruben and Kamen using ^{11}C but ^{14}C made rapid progress possible. In 1946 Melvin Calvin and Andrew A. Benson began their studies that elucidated the mechanism of incorporation of CO_2 into organic materials.

A key development was two-dimensional paper chromatography with radioautography (Box 3-C). A suspension of the alga *Chlorella* (Fig. 1-11) was allowed to photosynthesize in air. At a certain time, a portion of $H^{14}CO_3$ was injected into the system, and after a few seconds of photosynthesis with ^{14}C present the suspension of algae was run into hot methanol to denature proteins and to stop the reaction. The soluble materials extracted from the algal cells were concentrated and chromatographed; radioautographs were then prepared. It was found that after 10 s of photosynthesis in the presence of $^{14}CO_2$, the algae contained a dozen or more ^{14}C labeled compounds. These included malic acid, aspartic acid, phosphoenolpyruvate, alanine, triose phosphates, and other sugar phosphates and diphosphates. However, *during the first five seconds a single compound, 3-phosphoglycerate, contained most of the radioactivity.*[a,b] This finding suggested that a two-carbon regenerating substrate might be carboxylated by $^{14}CO_2$ to phosphoglycerate. Search for this two-carbon compound was unsuccessful, but Benson, in Calvin's laboratory, soon identified ribulose bisphosphate,[c] which kinetic studies proved to be the true regenerating substrate.[c,d] Its carboxylation and cleavage[e] represent the first step in what has come to be known as the Calvin–Benson cycle (Fig. 17-14).[f]



Chromatogram of extract of the alga *Scenedesmus* after photosynthesis in the presence of $^{14}CO_2$ for 10 s. Courtesy of J. A. Bassham. The origin of the chromatogram is at the lower right corner.

---

[a] Benson, A. A., Bassham, J. A., Calvin, M., Goodale, T. C., Haas, V. A., and Stepka, W. (1950) *J. Am. Chem. Soc.* **72**, 1710–1718

[b] Benson, A. A. (1951) *J. Am. Chem. Soc.* **73**, 2971–2972

[c] Benson, A. A., Kawaguchi, S., Hayes, P., and Calvin, M. (1952) *J. Am. Chem. Soc.* **74**, 4477–4482

[d] Bassham, J. A., Benson, A. A., Kay, L. D., Harris, A. Z., Wilson, A. T., and Calvin, M. (1954) *J. Am. Chem. Soc.* **76**, 1760–1770

[e] Calvin, M., and Bassham, J. A. (1962) *The Photosynthesis of Carbon Compounds*, Benjamin, New York

[f] Fuller, R. C. (1999) *Photosynth Res.* **62**, 1–29

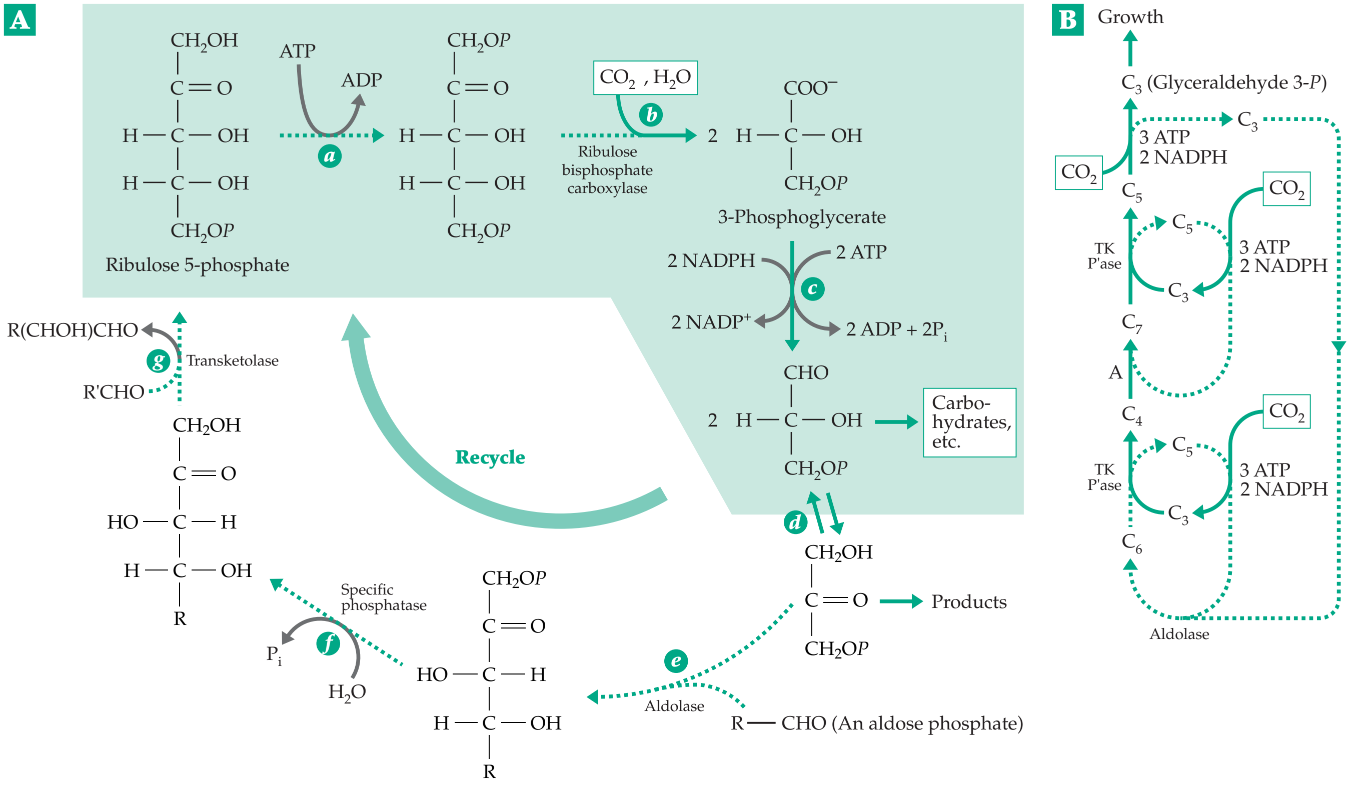

A
CH₂OH
C=O
H—C—OH
H—C—OH
CH₂OP
Ribulose 5-phosphate
ATP
ADP
a
CH₂OP
C=O
H—C—OH
H—C—OH
CH₂OP
CO₂ , H₂O
b
Ribulose bisphosphate carboxylase
COO⁻
2 H—C—OH
CH₂OP
3-Phosphoglycerate
2 NADPH
2 ATP
c
2 NADP⁺
2 ADP + 2Pᵢ
CHO
2 H—C—OH
CH₂OP
Carbo- hydrates, etc.
d
CH₂OH
C=O
CH₂OP
Products
e
Aldolase
R— CHO (An aldose phosphate)
CH₂OP
C=O
HO—C—H
H—C—OH
R
Specific phosphatase
f
Pᵢ
H₂O
CH₂OH
C=O
HO—C—H
H—C—OH
R
Recycle
R(CHOH)CHO
g
Transketolase
R'CHO
B
Growth
C₃ (Glyceraldehyde 3-P)
CO₂
3 ATP
2 NADPH
C₃
C₅
C₅
CO₂
3 ATP
2 NADPH
TK
P'ase
C₃
C₇
A
C₄
C₅
CO₂
3 ATP
2 NADPH
TK
P'ase
C₃
C₆
Aldolase

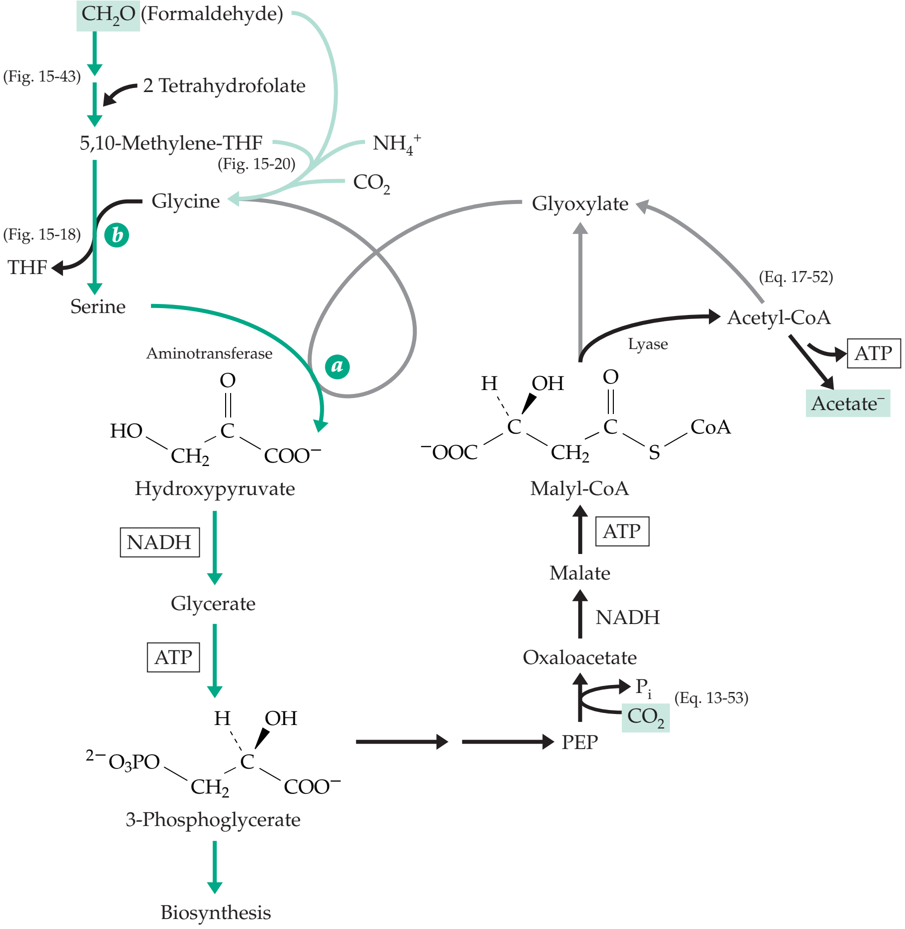
Figure 17-14 (A) The reductive carboxylation system used in reductive pentose phosphate pathway (Calvin–Benson cycle). The essential reactions of this system are enclosed within the dashed box. Typical subsequent reactions follow. The phosphatase action completes the phosphorylation–dephosphorylation cycle. (B) The reductive pentose phosphate cycle arranged to show the combining of three CO_2 molecules to form one molecule of triose phosphate. Abbreviations are RCS, reductive carboxylation system (from above); A, aldolase, Pase, specific phosphatase; and TK, transketolase.

three C_5 sugars three molecules of fructose 6-phosphate will be formed. One of these can be phosphorylated by ATP to fructose 1,6-bisphosphate, which can be cleaved by aldolase. One of the resulting triose phosphates can then be removed for biosynthesis and the second, together with the other two molecules of fructose 6-*P*, can be recycled through the sugar rearrangement sequence of Fig. 17-8B to regenerate the three ribulose 5-*P* molecules that serve as the regenerating substrate.

In bacteria, which lack formate dehydrogenase, formaldehyde can be oxidized to CO_2 to provide energy beginning with the reactions of Eq. 17-51. The resulting fructose 6-*P* is isomerized to glucose 6-*P*, which is then dehydrogenated via Eq. 17-12 to form CO_2 and the regenerating substrate ribulose 5-phosphate.

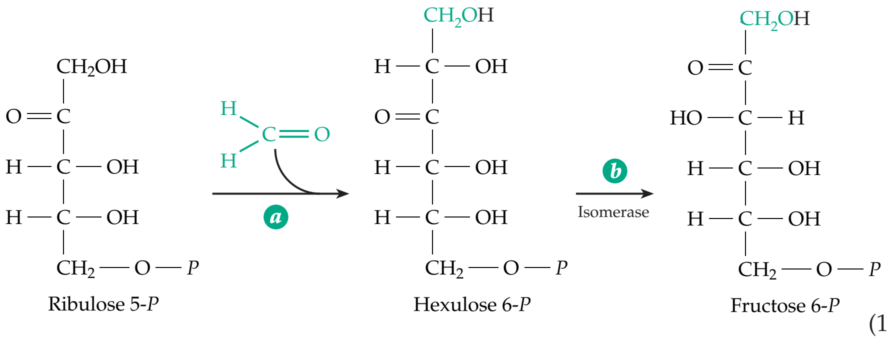
A number of pseudomonads and other bacteria convert C_1 compounds to acetate via tetrahydrofolic acid-bound intermediates and CO_2 using the **serine pathway**[179,192,193] shown in Fig. 17-15. This is a cyclic process for converting one molecule of formaldehyde (bound to tetrahydrofolate) plus one of CO_2 into acetate. The regenerating substrate is **glyoxylate**. Before condensation with the "active formaldehyde" of methylene THF, the glyoxylate undergoes transamination to glycine (Fig. 17-15, step *a*). The glycine plus formaldehyde forms serine (step *b*), which is then transaminated to hydroxypyruvate, again using step *a*. Glyoxylate plus formaldehyde could have been joined in a thiamin-dependent condensation. However, as in the γ-aminobutyrate shunt (Fig. 17-5), the coupled transamination step of Fig 17-15 permits use of PLP-dependent C–C bond formation.

Conversion of hydroxypyruvate to PEP (Fig. 17-15) involves reduction by NADH and phosphorylation by ATP to form 3-phosphoglycerate, which is converted to PEP as in glycolysis. The conversion of malate to acetate and glyoxylate via malyl-CoA and isocitrate lyase (Eq. 13-40) forms the product acetate and regenerates glyoxylate. As with other metabolic cycles, various intermediates, such as PEP, can be withdrawn for biosynthesis. However, there must be an independent route of synthesis of the regenerating substrate glyoxylate. One way in which this can be accomplished is to form glycine via the reversal of the glycine decarboxylase pathway as is indicated by the shaded green lines in Fig. 17-15.


CH₂O (Formaldehyde)
(Fig. 15-43)
2 Tetrahydrofolate
5,10-Methylene-THF
(Fig. 15-20)
NH₄⁺
CO₂
Glycine
(Fig. 15-18)
b
THF
Serine
Glyoxylate
(Eq. 17-52)
Acetyl-CoA
Lyase
ATP
Acetate⁻
Aminotransferase
a
Hydroxypyruvate
Malyl-CoA
NADH
ATP
Glycerate
Malate
ATP
NADH
Oxaloacetate
Pᵢ
(Eq. 13-53)
CO₂
PEP
3-Phosphoglycerate
Biosynthesis

Overall: $CH_2O + CO_2 + 2NADH \longrightarrow \text{acetate}^- + H^+ + H_2O$

Figure 17-15 One of the serine pathways for assimilation of one-carbon compounds.


Ribulose 5-P
a
b
Isomerase
Hexulose 6-P
Fructose 6-P

(17-51)

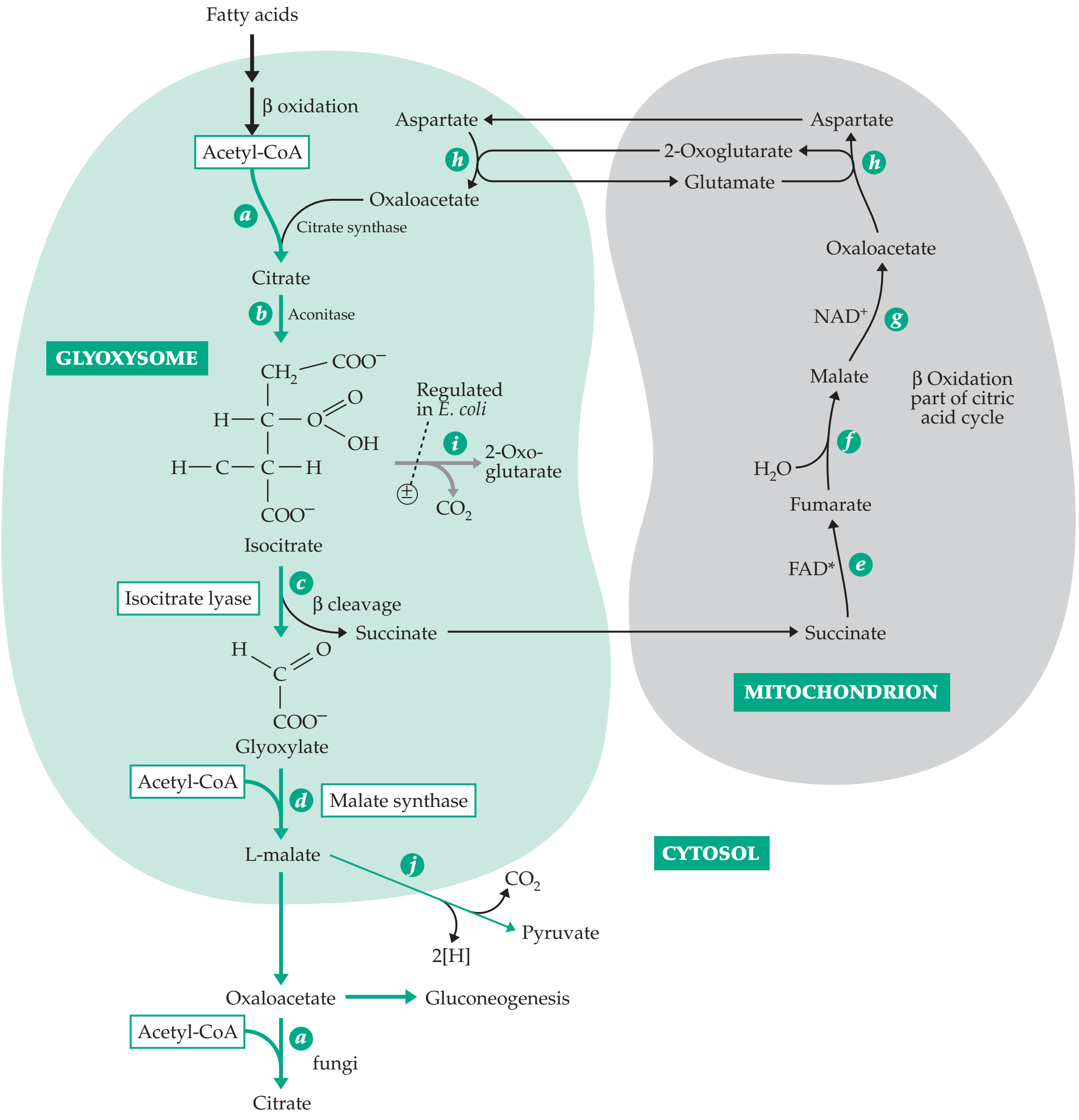
4. The Glyoxylate Pathways

The reductive carboxylation of acetyl-CoA to pyruvate (Eq. 17-47) occurs only in a few types of bacteria. For most species, from microorganisms to animals, the oxidative decarboxylation of pyruvate to acetyl-CoA is irreversible. This fact has many important consequences. For example, carbohydrate is readily converted to fat; because of the irreversibility of this process, excess calories lead to the deposition of fats. However, in animals fat cannot be used to generate most of the biosynthetic intermediates needed for formation of carbohydrates and proteins because those intermediates originate largely from C_3 units.

This limitation on the conversion of C_2 acetyl units to C_3 metabolites is overcome in many organisms by the **glyoxylate cycle** (Fig. 17-16), which converts *two* acetyl units into one C_4 unit. The cycle provides a way for organisms, such as *E. coli*,[111,199] *Saccharomyces*,[200] *Tetrahymena*, and the nematode *Caenorhabditis*,[201] to subsist on acetate as a sole or major carbon source. It is especially prominent in plants that store large amounts of fat in their seeds (**oil seeds**). In the germinating oil seed the glyoxylate cycle allows fat to be converted rapidly to sucrose, cellulose, and other carbohydrates needed for growth.

A key enzyme in the glyoxylate cycle is **isocitrate lyase**, which cleaves isocitrate (Eq. 13-40) to succinate and glyoxylate. The latter is condensed with a second acetyl group by the action of **malate synthase** (Eq. 13-38). The L-malate formed in this reaction is dehydrogenated to the regenerating substrate oxaloacetate. Some of the reaction steps are those of the citric acid cycle and it appears that in bacteria there is no spatial separation of the citric acid cycle and glyoxylate pathway. However, in plants the enzymes of the glyoxylate cycle are present in specialized peroxisomes known as **glyoxysomes**.[60] The glyoxysomes also contain the enzymes for β oxidation of fatty acids, allowing for efficient conversion of fatty acids to **succinate**. This compound is exported from the glyoxysomes and enters the mitochondria where it undergoes β oxidation to oxaloacetate. The latter can be converted by PEP carboxylase (Eq. 13-53) or by PEP carboxykinase (Eq. 13-46) to PEP.

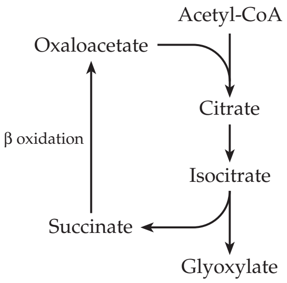
An **acetyl-CoA-glyoxylate** cycle, which catalyzes oxidation of acetyl groups to glyoxylate, can also be constructed from isocitrate lyase and citric acid cycle



Fatty acids
β oxidation
Acetyl-CoA
a
Citrate synthase
Oxaloacetate
Aspartate
h
2-Oxoglutarate
Glutamate
Aspartate
h
Oxaloacetate
Citrate
b
Aconitase
GLYOXYSOME
CH2
COO−
H
C
O
O
OH
H
C
C
H
COO−
Isocitrate
Regulated in E. coli
i
2-Oxo-glutarate
±
CO2
NAD+
g
Malate
β Oxidation part of citric acid cycle
H2O
f
Fumarate
FAD*
e
Isocitrate lyase
c
β cleavage
Succinate
Succinate
MITOCHONDRION
H
C
O
COO−
Glyoxylate
Acetyl-CoA
d
Malate synthase
L-malate
j
CO2
Pyruvate
2[H]
CYTOSOL
Oxaloacetate
Gluconeogenesis
Acetyl-CoA
a
fungi
Citrate

Figure 17-16 The glyoxylate pathway. The green line traces the pathway of labeled carbon from fatty acids or acetyl-CoA into malate and other products.

enzymes. Glyoxylate is taken out of the cycle as the product and succinate is recycled (Eq. 17-52). The independent pathway for synthesis of the regenerating substrate oxaloacetate is condensation of glyoxylate with acetyl-CoA (malate synthetase) to form malate and oxidation of the latter to oxaloacetate as in the main cycle of Fig. 17-16.


Acetyl-CoA
Oxaloacetate
Citrate
β oxidation
Isocitrate
Succinate
Glyoxylate

(17-52)

5. Biosynthesis of Glucose from Three-Carbon Compounds

Now let us consider the further conversion of PEP and of the triose phosphates to **glucose 1-phosphate**, the key intermediate in biosynthesis of other sugars and polysaccharides. The conversion of PEP to glucose 1-*P* represents a reversal of part of the glycolysis sequence. It is convenient to discuss this along with **gluconeogenesis**, the reversal of the complete glycolysis sequence from lactic acid. This is an essential part of the Cori cycle (Section F) in our own bodies, and the same process may be used to convert pyruvate derived from deamination of alanine or serine (Chapter 24) into carbohydrates.

Just as with the pentose phosphate cycle, an exact reversal of the glycolysis sequence (Eq. 17-53) is precluded on thermodynamic grounds. Even at very high values of the phosphorylation state ratio R_p, the reaction:

$$\begin{aligned}&2\ \text{Lactate}^- + 3\ \text{ATP}^{4-} + 2\ \text{H}_2\text{O} \rightarrow \text{C}_6\text{H}_{10}\text{O}_5\ (\text{glycogen})\\ &\quad + 3\ \text{ADP}^{3-} + 3\ \text{HPO}_4^{\ 2-} + \text{H}^+\\ &\Delta G'\ (\text{pH 7}) = +\ 107\ \text{kJ per glycosyl unit} \qquad (17\text{-}53)\end{aligned}$$

would be unlikely to go to completion. The actual pathways used for gluconeogenesis (Fig. 17-17, green lines) differ from those of glycolysis (black lines) in three significant ways. First, while glycogen breakdown is initiated by the reaction with inorganic phosphate catalyzed by phosphorylase (Fig. 17-17, step *a*), the biosynthetic sequence from glucose 1-*P*, via uridine diphosphate glucose (Fig. 17-17, step *b*; see also Eq. 17-56), is coupled to cleavage of ATP. Second, in the catabolic process (glycolysis) fructose 6-*P* is converted to fructose 1,6-P_2 through the action of a kinase (Fig. 17-17, step *c*), which is then cleaved by aldolase. The resulting triose phosphate is degraded to PEP. In glucogenesis a phosphatase is used to form fructose *P* from fructose P_2 (Fig. 17-17, step *d*). Third, during gly-colysis PEP is converted to pyruvate by a kinase with generation of ATP (Fig. 17-17, step *e*). During glucogenesis pyruvate is converted to PEP indirectly via oxaloacetate (Fig. 17-17, steps *f* and *g*) using pyruvate carboxylase (Eq. 14-3) and PEP carboxykinase (Eq. 13-46). This is another example of the coupling of ATP cleavage through a carboxylation–decarboxylation sequence. The net effect is to use two molecules of ATP (actually one ATP and one GTP) rather than *one* to convert pyruvate to PEP.

The overall reaction for reversal of glycolysis to form glycogen (Eq. 17-54) now has a comfortably negative standard Gibbs energy change as a result of coupling the cleavage of 7 ATP to the reaction.

$$\begin{aligned}&2\ \text{Lactate}^- + 7\ \text{ATP}^{4-} + 6\ \text{H}_2\text{O} \rightarrow\\ &\quad \text{glycogen} + 7\ \text{ADP}^{3-} + 7\ \text{HPO}_4^{\ 2-} + 5\ \text{H}^+\\[1ex] &\Delta G'\ (\text{pH 7}) = -\ 31\ \text{kJ mol}^{-1}\ \text{per glycosyl unit} \qquad (17\text{-}54)\end{aligned}$$

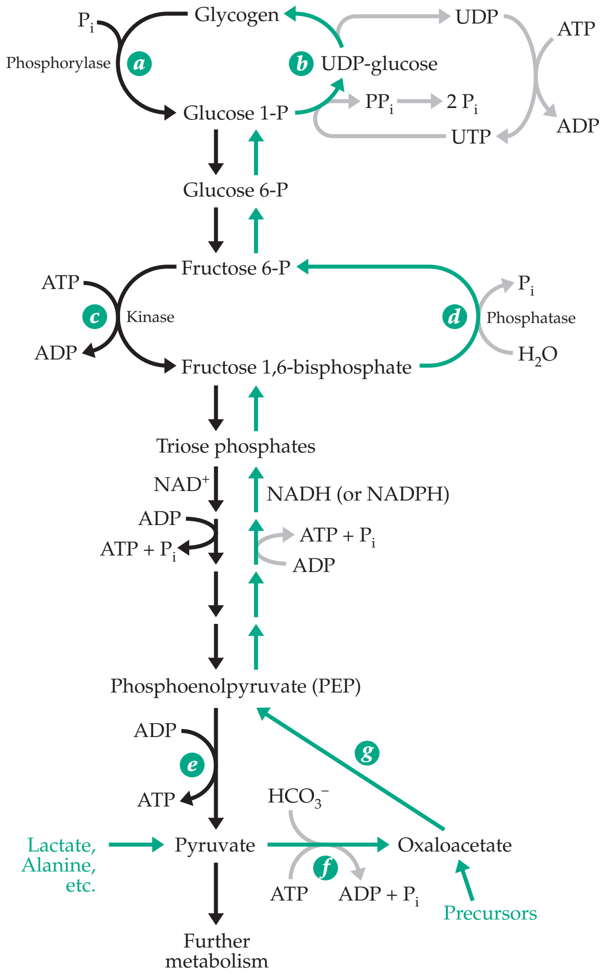
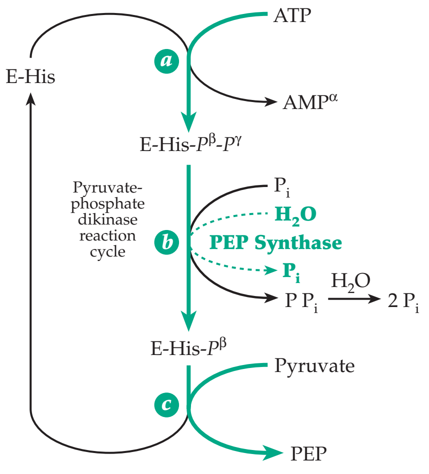

Pi
Glycogen
UDP
ATP
Phosphorylase
a
b
UDP-glucose
Glucose 1-P
PPi
2 Pi
UTP
ADP
Glucose 6-P
Fructose 6-P
ATP
c
Kinase
d
Pi
Phosphatase
ADP
H2O
Fructose 1,6-bisphosphate
Triose phosphates
NAD+
NADH (or NADPH)
ADP
ATP + Pi
ATP + Pi
ADP
Phosphoenolpyruvate (PEP)
ADP
e
g
ATP
HCO3−
Lactate, Alanine, etc.
Pyruvate
Oxaloacetate
f
ATP
ADP + Pi
Precursors
Further metabolism

Figure 17-17 Comparison of glycolytic pathway (left) with pathway of gluconeogenesis (right, green arrows).

Two enzymes that are able to convert pyruvate directly to PEP are found in some bacteria and plants. In each case, as in the animal enzyme system discussed in the preceding paragraph, the conversion involves expenditure of two high-energy linkages of ATP. The **PEP synthase** of *E. coli* first transfers a pyrophospho group from ATP onto an imidazole group of histidine in the enzyme (Eq. 17-55). A phospho group is hydrolyzed from this intermediate (dashed green line in Eq. 17-55, step *b*), ensuring that sufficient intermediate E–His–*P* is present. The latter reacts with pyruvate to form PEP.[202,203] **Pyruvate–phosphate dikinase** is a similar enzyme first identified in tropical grasses and known to play an important role in the CO_2 concentrating system of the so-called "C_4 plants" (Chapter 23).[204] The same enzyme participates in gluconeogenesis in *Acetobacter*. The reaction cycle for this enzyme is also portrayed in Eq. 17-55. In this case P_i, rather than water, is the attacking nucleophile in Eq. 17-55 and PP_i is a product. The latter is probably hydrolyzed by pyrophosphatase action, the end result being an overall reaction that is the same as with PEP synthase. Kinetic and positional isotope exchange studies suggest that the P_i must be bound to pyruvate–phosphate dikinase before the bound ATP can react with the imidazole group.[202] Likewise, AMP doesn't dissociate until P_i has reacted to form PP_i.



(17-55)

6. Building Hydrocarbon Chains with Two-Carbon Units

Fatty acid chains are taken apart two carbon atoms at a time by β oxidation. Biosynthesis of fatty acids reverses this process by using the two-carbon acetyl unit of acetyl-CoA as a starting material. The coupling of ATP cleavage to this process by a carboxylation–decarboxylation sequence, the role of acyl carrier protein (Section H,4), and the use of NADPH as a reductant (Section I) have been discussed and are summarized in Fig. 17-12, which gives the complete sequence of reactions for fatty acid biosynthesis. Why does β oxidation require CoA derivatives while biosynthesis requires the more complex acyl carrier protein (ACP)? The reason may involve control. ACP is a complex handle able to hold the growing fatty acid chain and to guide it from one enzyme to the next. In *E. coli* the various enzymes catalyzing the reactions of Fig. 17-12 are found in the cytosol and behave as independent proteins. The same is true for fatty acid synthases of higher plants which resemble those of bacteria.[205,205a]

It is thought that the ACP molecule lies at the center of the complex and that the growing fatty acid chain on the end of the phosphopantetheine prosthetic group moves from one subunit to the other.[164,206] The process is started by a **primer** which is usually acetyl-CoA in *E. coli*. Its acyl group is transferred first to the central molecule of ACP (step *a*, Fig. 17-12) and then to a "peripheral" thiol group, probably that of a cysteine side chain on a separate protein subunit (step *b*, Fig. 17-12). Next, a malonyl group is transferred (step *d*) from malonyl-CoA to the free thiol group on the ACP. The condensation (steps *e* and *f*) occurs with the freeing of the peripheral thiol group. The latter does not come into use again until the β-oxoacyl group formed has undergone the complete sequence of reduction reactions (steps *g*–*i*). Then the growing chain is again transferred to the peripheral –SH (step *j*) and a new malonyl unit is introduced on the central ACP.

After the chain reaches a length of 12 carbon atoms, the acyl group tends to be transferred off to a CoA molecule (step *k*) rather than to pass around the cycle again. Thus, chain growth is terminated. This tendency systematically increases as the chain grows longer.

In higher animals as well as in *Mycobacterium*,[207] yeast,[208] and *Euglena*, the **fatty acid synthase** consists of only one or two multifunctional proteins. The synthase from animal tissues has seven catalytic activities in a single 263-kDa 2500-residue protein.[209] The protein consists of a series of domains that contain the various catalytic activities needed for the entire synthetic sequence. One domain contains an ACP-like site with a bound 4'-phosphopantetheine as well as a cysteine side chain in the second acylation site. This synthase produces free fatty acids, principally the C_{16} palmitate. The final step is cleavage of the acyl-CoA by a thioesterase, one of the seven enzymatic activities of the synthase. See Chapter 21 for further discussion.

7. The Oxoacid Chain Elongation Process

As mentioned in Section 4, glyoxylate can be converted to oxaloacetate by condensation with acetyl-CoA (Fig. 17-16) and the oxaloacetate can be decarboxylated to pyruvate. This sequence of reactions resembles that of the conversion of oxaloacetate to 2-oxoglutarate in the citric acid cycle (Fig. 17-4). *Both* *are examples of a frequently used general chain elongation process for α-oxo acids*. This sequence, which is illustrated in Fig. 17-18, has four steps: (1) condensation of the α-oxo acid with an acetyl group, (2) isomerization by dehydration and rehydration (catalyzed by aconitase in the case of the citric acid cycle), (3) dehydrogenation, and (4) β decarboxylation. In many cases steps 3 and 4 are combined as a single enzymatic reaction. The isomerization of the intermediate hydroxy acid in step 2 is required because the hydroxyl group, which is attached to a tertiary carbon bearing no hydrogen, must be moved to the adjacent carbon atom before oxidation to a ketone can take place. However, in the case of glyoxylate, isomerization is not necessary because R = H.

It may be protested that the reaction of the citric acid cycle by which oxaloacetate is converted to oxoglutarate does not follow exactly the pattern of Fig. 17-18. The carbon dioxide removed in the decarboxylation step does not come from the part of the molecule donated by the acetyl group but from that formed from oxaloacetate. However, the end result is the same. Furthermore, there are two known citrate-forming enzymes with different stereospecificities (Chapter 13), one of which leads to a biosynthetic pathway strictly according to the sequence of Fig. 17-18.

At the bottom of Fig. 17-18 several stages of the α-oxo acid elongation process are arranged in tandem. We see that glyoxylate (a product of the acetyl-CoA–glyoxylate cycle) can be built up systematically to

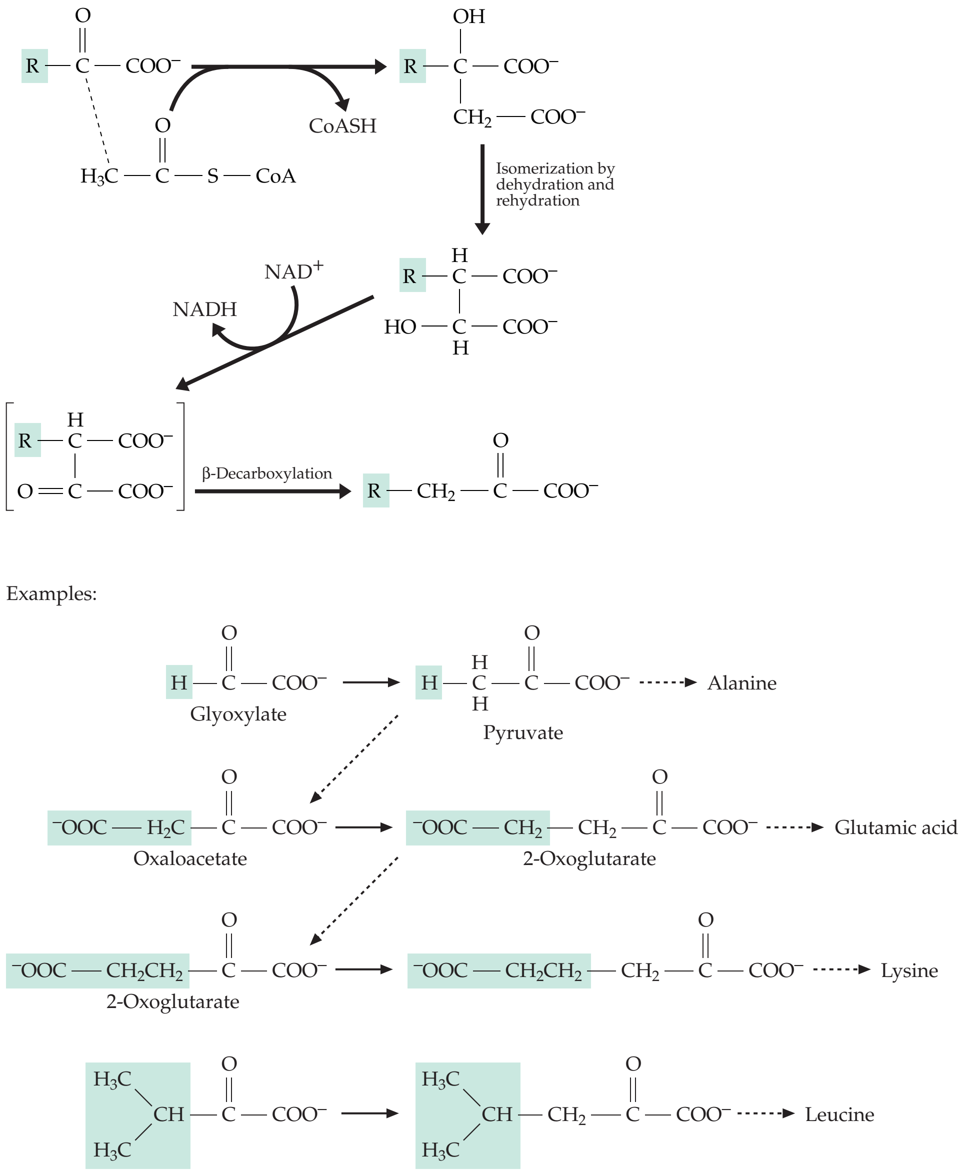
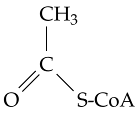

R
COO⁻
H₃C — C — S — CoA
CoASH
OH
CH₂ — COO⁻
Isomerization by dehydration and rehydration
HO
NAD⁺
NADH
β-Decarboxylation
R — CH₂ — C — COO⁻
Examples:
Glyoxylate
Pyruvate
Alanine
Oxaloacetate
2-Oxoglutarate
Glutamic acid
2-Oxoglutarate
Lysine
H₃C
CH
Leucine

Figure 17-18 The oxoacid chain elongation process.

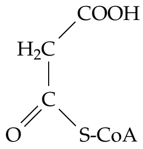
pyruvate, oxaloacetate, 2-oxoglutarate, and 2-oxoadipate (a precursor of lysine) using this one reaction sequence. Methanogens elongate 2-oxoadipate by one and two carbon atoms using the same sequence to give 7- and 8-carbon dicarboxylates.[210]

8. Decarboxylation as a Driving Force in Biosynthesis

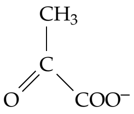
Consider the relationship of the following prominent biosynthetic intermediates one to another:



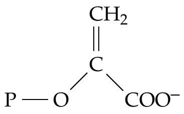
Acetyl-CoA



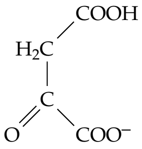
Malonyl-CoA



Pyruvate



PEP



Oxaloacetate (OAA)

Utilization of acetyl-CoA for the synthesis of long-chain fatty acids occurs via carboxylation to malonyl-CoA. *We can think of the malonyl group as a β-carboxylated acetyl group*. During synthesis of a fatty acid the carboxyl group is lost, and only the acetyl group is ultimately incorporated into the fatty acid. In a similar way *pyruvate can be thought of as an α-carboxylated acetaldehyde and oxaloacetate as an α- and β-dicarboxylated acetaldehyde*. During biosynthetic reactions these three- and four-carbon compounds also often undergo decarboxylation. Thus, they both can be regarded as "activated acetaldehyde units." Phosphoenolpyruvate is an α-carboxylated phosphoenol form of acetaldehyde and undergoes both decarboxylation and dephosphorylation before contributing a two-carbon unit to the final product.

It is of interest to compare two chain elongation processes by which two-carbon units are combined. In the synthesis of fatty acids the acetyl units are condensed and then are reduced to form straight hydrocarbon chains. In the oxo-acid chain elongation mechanism, the acetyl unit is introduced but is later decarboxylated. Thus, the chain is increased in length by one carbon atom at a time. These two mechanisms account for a great deal of the biosynthesis by chain extension. However, there are other variations. For example, glycine (a carboxylated methylamine), under the influence of pyridoxal phosphate and with accompanying decarboxylation, condenses with succinyl-CoA (Eq. 14-32) to extend the carbon chain and at the same time to introduce an amino group. Likewise, serine (a carboxylated ethanolamine) condenses with palmitoyl-CoA in biosynthesis of sphingosine (as in Eq. 14-32). Phosphatidylserine is decarboxylated to phosphatidylethanolamine in the final synthetic step for that phospholipid (Fig. 21-5).

9. Stabilization and Termination of Chain Growth by Ring Formation

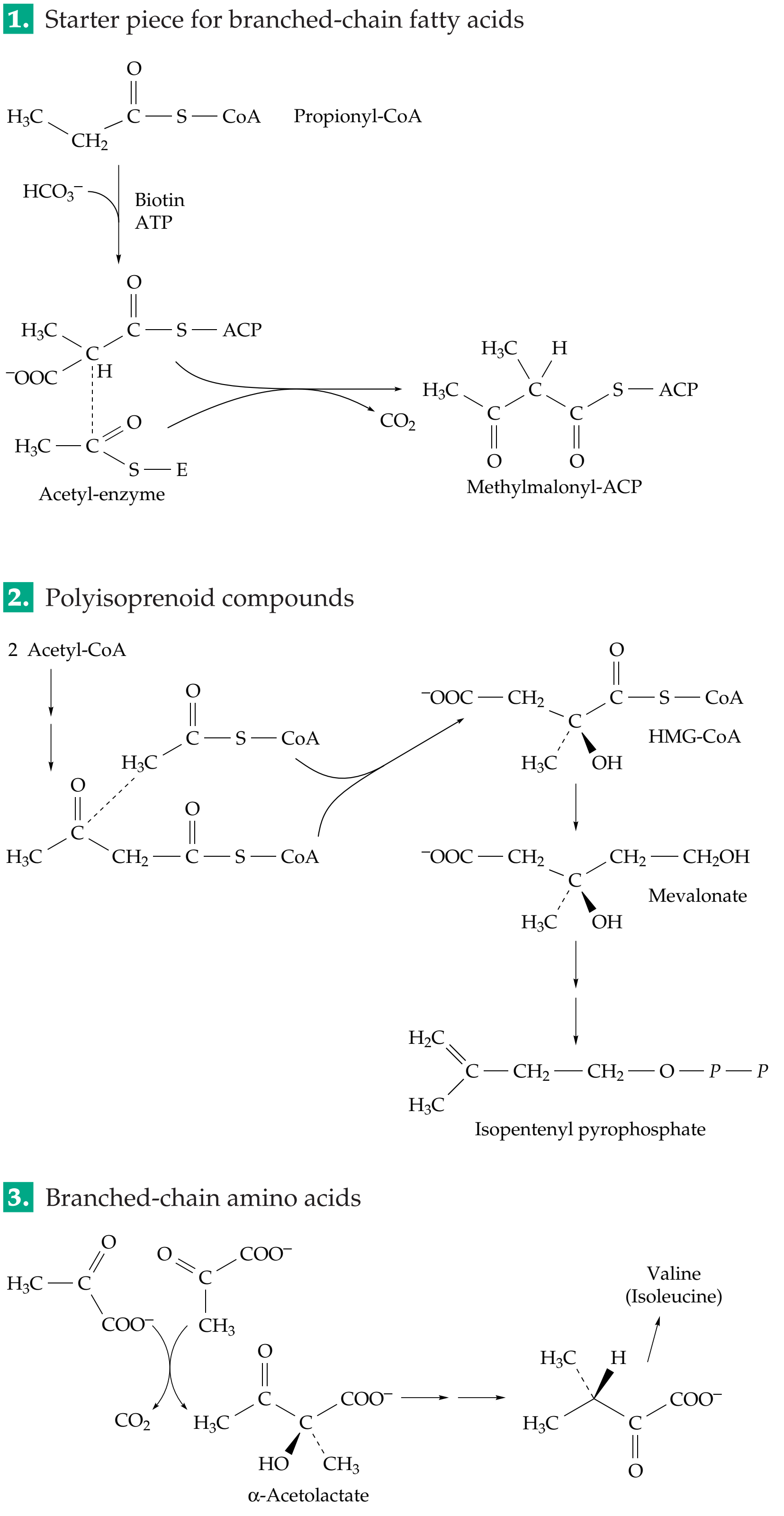
Biochemical substances frequently undergo cyclization to form stable five- and six-atom ring structures. The three-carbon glyceraldehyde phosphate exists in solution primarily as the free aldehyde (and its covalent hydrate) but glucose 6-phosphate exists largely as the cyclic hemiacetal. In this ring form no carbonyl group is present and further chain elongation is inhibited. When the hemiacetal of glucose 6-*P* is enzymatically isomerized to glucose 1-*P* the ring is firmly locked. Glucose 1-*P*, in turn, serves as the biosynthetic precursor of polysaccharides and related compounds, in all of which the sugar rings are stable. Ring formation can occur in lipid biosynthesis, too. Among the **polyketides** (Chapter 21), polyprenyl compounds (Chapter 22), and aromatic amino acids (Chapter 25) are many substances in which ring formation has occurred by ester or aldol condensations followed by reduction and elimination processes. This is a typical sequence for biosynthesis of highly stable aromatic rings.

10. Branched Carbon Chains

Branched carbon skeletons are formed by standard reaction types but sometimes with addition of rearrangement steps. Compare the biosynthetic routes to three different branched five-carbon units (Fig. 17-19) The first is the use of a **propionyl group** to initiate formation of a branched-chain fatty acid. Propionyl-CoA is carboxylated to methylmalonyl-CoA, whose acyl group is transferred to the acyl carrier protein before condensation. Decarboxylation and reduction yields an acyl-CoA derivative with a methyl group in the 3-position.

The second five-carbon branched unit, in which the branch is one carbon further down the chain, is an intermediate in the biosynthesis of **polyprenyl** (isoprenoid) compounds and steroids. Three two-carbon units are used as the starting material with decarboxylation of one unit. Two acetyl units are first condensed to form acetoacetyl-CoA. Then a third acetyl unit, which has been transferred from acetyl-CoA onto an SH group of the enzyme, is combined with the acetoacetyl-CoA through an ester condensation. The thioester linkage to the enzyme is hydrolyzed to free the product **3-hydroxy-3-methylglutaryl-CoA** (HMG-CoA). This sequence is illustrated in Eq. 17-5. The thioester group of HMG-CoA is reduced to the alcohol **mevalonic acid**, a direct precursor to isopentenyl pyrophosphate, from which the polyprenyl compounds are formed (Fig. 22-1).

The third type of carbon-branched unit is 2-oxoisovalerate, from which valine is formed by transamination. The starting units are two molecules of pyruvate which combine in a thiamin diphosphate-dependent α condensation with decarboxylation. The resulting α-acetolactate contains a branched chain but is quite unsuitable for formation of an α amino acid. A rearrangement moves the methyl group to the β position (Fig. 24-17), and elimination of water from the diol forms the enol of the desired α-oxo acid (Fig. 17-19). The precursor of isoleucine is formed in an analogous way by condensation, with decarboxylation of one molecule of pyruvate with one of 2-oxobutyrate.


1. Starter piece for branched-chain fatty acids

2. Polyisoprenoid compounds

3. Branched-chain amino acids

Figure 17-19 Biosynthetic origins of three five-carbon branched structural units. Notice that decarboxylation is involved in driving each sequence.

K. Biosynthesis and Modification of Polymers

There are three chemical problems associated with the assembly of a protein, nucleic acid, or other biopolymer. The first is *to overcome thermodynamic barriers*. The second is *to control the rate of synthesis*, and the third is *to establish the pattern or sequence in which the monomer units are linked together*. Let us look briefly at how these three problems are dealt with by living cells.

1. Peptides and Proteins

Activation of amino acids for incorporation into oligopeptides and proteins can occur via two routes of acyl activation. In the first of these an **acyl phosphate** (or acyl adenylate) is formed and reacts with an amino group to form a peptide linkage (Eq. 13-4). The tripeptide **glutathione** is formed in two steps of this type (Box 11-B). In the second method of activation **aminoacyl**

adenylates are formed. They transfer their activated aminoacyl groups onto specific tRNA molecules during synthesis of proteins (Eq. 17-36). In other cases activated aminoacyl groups are transferred onto –SH groups to form intermediate **thioesters**. An example is the synthesis of the antibiotic **gramicidin S** formed by *Bacillus brevis*. The antibiotic is a cyclic decapeptide with the following five-amino-acid sequence repeated twice in the ringlike molecule[211]:

(–D-Phe–L-Pro–L-Val–L-Orn–L-Leu–)$_2$

The soluble enzyme system responsible for its synthesis contains a large 280-kDa protein that not only activates the amino acids as aminoacyl adenylates and transfers them to thiol groups of 4'-phosphopantetheine groups covalently attached to the enzyme but also serves as a template for joining the amino acids in proper sequence.[211–214] Four amino acids—proline, valine, ornithine (Orn), and leucine—are all bound. A second enzyme (of mass 100 kDa) is needed for activation of phenylalanine. It is apparently the activated phenylalanine (which at some point in the process is isomerized from L- to D-phenylalanine) that initiates polymer formation in a manner analogous to that of fatty acid elongation (Fig. 17-12). Initiation occurs when the amino group of the activated phenylalanine (on the second enzyme) attacks the acyl group of the aminoacyl thioester by which the activated proline is held. Next, the freed imino group of proline attacks the activated valine, etc., to form the pentapeptide. Then two pentapeptides are joined and cyclized to give the antibiotic. The sequence is absolutely specific, and it is remarkable that this relatively small enzyme system is able to carry out each step in the proper sequence. Many other peptide antibiotics, such as the bacitracins, tyrocidines,[215] and enniatins, are synthesized in a similar way,[213,216,217] as are depsipeptides and the immunosuppresant cyclosporin. A virtually identical pattern is observed for formation of **polyketides**,[218,219] whose chemistry is considered in Chapter 21.

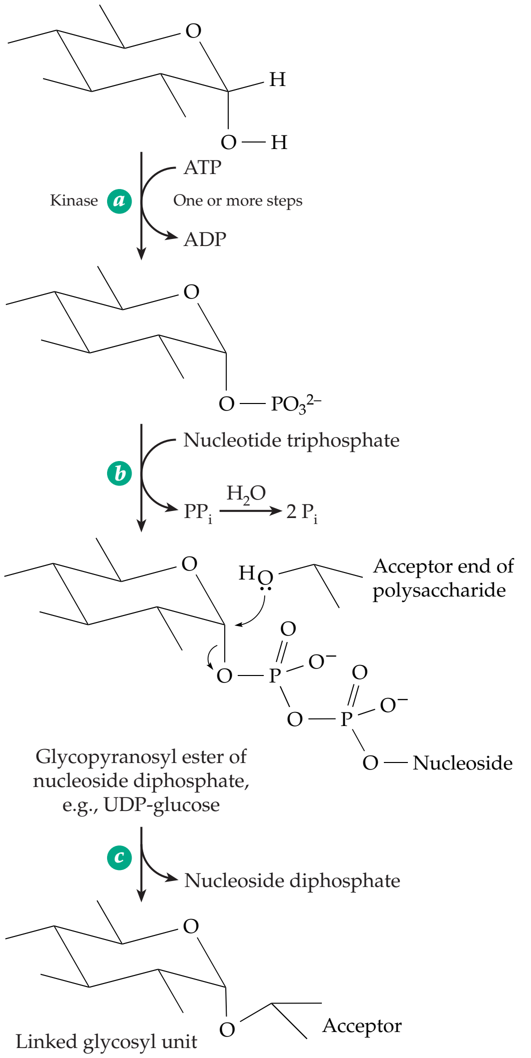
While peptide antibiotics are synthesized according to enzyme-controlled polymerization patterns, both proteins and nucleic acids are made by **template mechanisms**. The sequence of their monomer units is determined by genetically encoded information. A key reaction in the formation of proteins is the transfer of activated aminoacyl groups to molecules of tRNA (Eq. 17-36). The tRNAs act as carriers or adapters as explained in detail in Chapter 29. Each **aminoacyl-tRNA synthetase** must recognize the correct tRNA and attach the correct amino acid to it. The tRNA then carries the activated amino acid to a ribosome, where it is placed, at the correct moment, in the active site. **Peptidyltransferase**, using a transacylation reaction, in an *insertion mechanism* transfers the C terminus of the growing peptide chain onto the amino group of the new amino acid to give a tRNA-bound peptide one unit longer than before.

2. Polysaccharides

Incorporation of a sugar monomer into a polysaccharide also involves cleavage of two high-energy phosphate linkages of ATP. However, the activation process has its own distinctive pattern (Eq. 17-56). Usually a sugar is first phosphorylated by a kinase or a kinase plus a phosphomutase (Eq. 17-56, step *a*). Then a nucleoside triphosphate (NuTP) reacts under the influence of a second enzyme with elimination of pyrophosphate and formation of a **glycopyranosyl ester** of the nucleoside diphosphate, more often known as a **sugar nucleotide** (Eq. 17-56, step *b*). The inorganic pyrophosphate is hydrolyzed by pyrophosphatase while the sugar nucleotide donates the activated glycosyl group for polymerization (Eq. 17-56, step *c*). In this step the glycosyl group is transferred with displacement of the nucleoside diphosphate. Thus, the overall process involves first the cleavage of ATP to ADP and P_i, and then the cleavage of a nucleoside triphosphate to a nucleoside diphosphate plus P_i. The nucleoside triphosphate in Eq. 17-56, step *b* is sometimes ATP, in which case the overall result is the splitting of two molecules of ATP to ADP. However, as detailed in Chapter 20, the whole series of nucleotide "handles" serve to carry various activated glycosyl units.

What determines the pattern of incorporation of sugar units into polysaccharides? Homopolysaccharides, like cellulose and the linear amylose form of starch, contain only one monosaccharide component in only one type of linkage. A single synthetase enzyme can add unit after unit of an activated sugar (UDP glucose or other sugar nucleotide) to the growing end. However, at least two enzymes are needed to assemble a branched molecule such as that of the glycogen molecule. One is the synthetase; the second is a **branching enzyme**, a transglycosylase. After the chain ends attain a length of about ten monosaccharide units the branching enzyme attacks a glycosidic linkage somewhere in the chain. Acting much like a hydrolase, it forms a glycosyl enzyme (or a stabilized carbocation) intermediate. The enzyme does not release the severed chain fragment but transfers it to another nearby site on the branched polymer. In the synthesis of glycogen, the chain fragment is joined to a free 6-hydroxyl group of the glycogen, creating a new branch attached by an α-1,6-linkage.

Other carbohydrate polymers consist of **repeating oligosaccharide units**. Thus, in hyaluronan units of glucuronic acid and N-acetyl-D-glucosamine alternate (Fig. 4-11). The "O antigens" of bacterial cell coats (p. 180) contain repeating subunits made up of a "block" of four or five different sugars. In these and

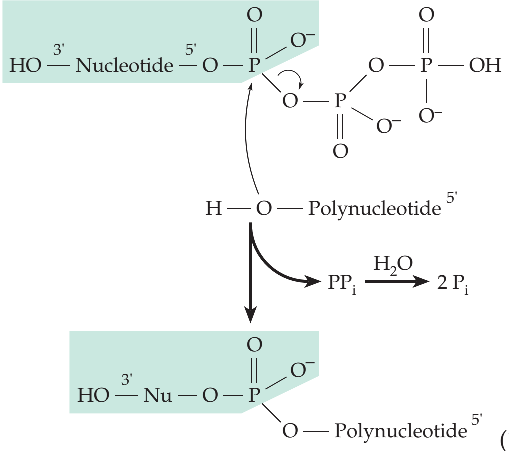

O
H
O — H
ATP
Kinase
a
One or more steps
ADP
O
O — PO_3^{2-}
Nucleotide triphosphate
b
$PP_i \xrightarrow{H_2O} 2\ P_i$
O
HO
Acceptor end of polysaccharide
O
O — P — O⁻
O
O — P — O⁻
O — Nucleoside
Glycopyranosyl ester of nucleoside diphosphate, e.g., UDP-glucose
c
Nucleoside diphosphate
O
O
Acceptor
Linked glycosyl unit

(17-56)

many other cases the pattern of polymerization is established by the specificities of individual enzymes. An enzyme capable of joining an activated glucosyl unit to a growing polysaccharide will do so only if the proper structure has been built up to that point. In cases where a block of sugar units is transferred it is usually *inserted* at the nonreducing end of the polymer, which may be covalently attached to a protein. Notice that the insertion mode of chain growth exists for lipids, polysaccharides, and proteins.

3. Nucleic Acids

The activated nucleotides are the nucleoside 5'-triphosphates. The ribonucleotides ATP, GTP, UTP, and CTP are needed for RNA synthesis and the 2'-deoxyribonucleotide triphosphates, dATP, dTTP, dGTP, and dCTP for DNA synthesis. In every case, the addition of activated monomer units to a growing polynucleotide chain is catalyzed by an enzyme that binds to the template nucleic acid. The choice of the proper nucleotide unit to place next in the growing strand is determined by the nucleotide already in place in the complementary strand, a matter that is dealt with in Chapters 27 and 28. The chemistry is a simple displacement of pyrophosphate (Eq. 17-57). The 3'-hydroxyl of the polynucleotide attacks the phosphorus atom of the activated nucleoside triphosphate. Thus, *nucleotide chains always grow from the 5′ end, with new units being added at the 3′ end*.


HO — Nucleotide — O — P
3'
5'
O
O⁻
O — P
O
O⁻
O — P — OH
O
O⁻
H — O — Polynucleotide
5'
$PP_i \xrightarrow{H_2O} 2\ P_i$
HO — Nu — O — P
3'
O
O⁻
O — Polynucleotide
5'

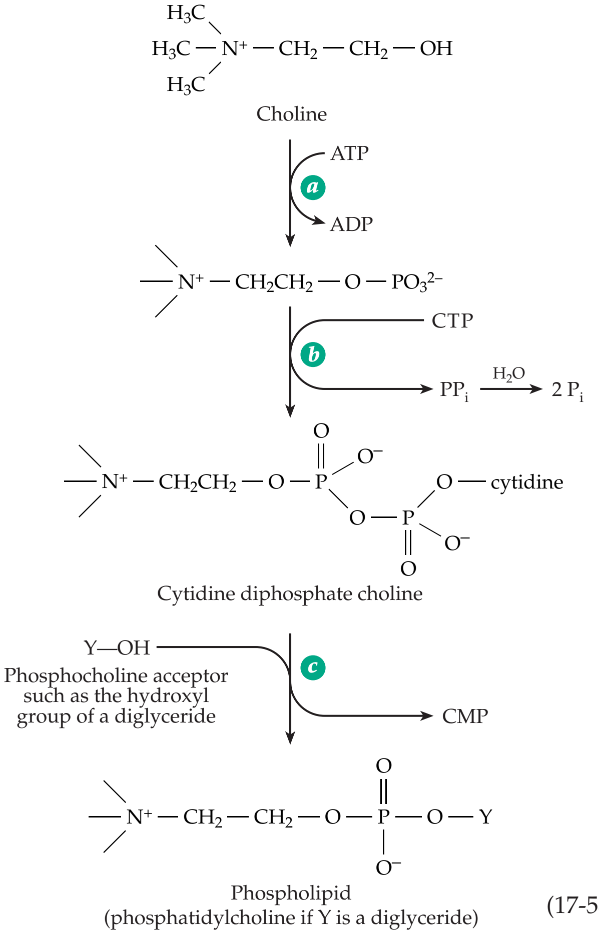
(17-57)

4. Phospholipids and Phosphate–Sugar Alcohol Polymers

Choline and ethanolamine are activated in much the same way as are sugars. For example, choline can be phosphorylated using ATP (Eq. 17-58, step *a*) and the phosphocholine formed can be further converted (Eq. 17-58, step *b*) to **cytidine diphosphate choline**. Phosphocholine is transferred from the latter onto a suitable acceptor to form the final product (Eq. 17-58, step *c*). The polymerization pattern differs from that for polysaccharide synthesis. When the sugar nucleotides react, the entire nucleoside diphosphate is eliminated (Eq. 17-56), but CDP-choline and CDP-ethanolamine react with elimination of CMP (Eq. 17-58, step *c*), leaving one phospho group in the final product. The same thing is true in the synthesis of the bacterial teichoic acids (Chapter 8). Either CDP-glycerol or CDP-ribitol is formed first and polymerization takes place with elimination of CMP to form the alternating phosphate–sugar alcohol polymer.[220]

5. Irreversible Modification and Catabolism of Polymers

While polymers are being synthesized continuously by cells, they are also being modified and torn down. Nothing within a cell is static. As discussed in Chapters


Choline
ATP
a
ADP
CTP
b
PP$_i$ $\xrightarrow{H_2O}$ 2 P$_i$
cytidine
Cytidine diphosphate choline
Y—OH
Phosphocholine acceptor such as the hydroxyl group of a diglyceride
c
CMP
Phospholipid
(phosphatidylcholine if Y is a diglyceride)

(17-58)

10 and 29, everything turns over at a slower or faster rate. Hydrolases attack all of the polymers of which cells are composed, and active catabolic reactions degrade the monomers formed. Membrane surfaces are also altered, for example, by hydroxylation and glycosylation of both glycoproteins and lipid head groups. It is impossible to list all of the known modification reactions of biopolymers. They include hydrolysis, methylation, acylation, isopentenylation, phosphorylation, sulfation, and hydroxylation. Precursor molecules are cut and trimmed and often modified further to form functional proteins or nucleic acids. Phosphotransferase reactions splice RNA transcripts to form mRNA and a host of alterations convert precursors into mature tRNA molecules (Chapter 28). Even DNA, which remains relatively unaltered, undergoes a barrage of chemical attacks. Only because of the presence of an array of repair enzymes (Chapter 27) does our DNA remain nearly unchanged so that faithful copies can be provided to each cell in our bodies and can be passed on to new generations.

L. Regulation of Biosynthesis

A simplified view of metabolism is to consider a cell as a "bag of enzymes." Indeed, much of metabolism can be explained by the action of several thousand enzymes promoting specific reactions of their substrates. These reactions are based upon the natural chemical reactivities of the substrates. However, the enzymes, through the specificity of their actions and through association with each other,[96,221–223] channel the reactions into a selected series of metabolic pathways. The reactions are often organized as cycles which are inherently stable. We have seen that biosynthesis often involves ATP-dependent reductive reactions. *It is these reductive processes that produce the less reactive nonpolar lipid groupings and amino acid side chains so essential to the assembly of insoluble intracellular structures.* Oligomeric proteins, membranes, microtubules, and filaments are all the natural result of aggregation caused largely by hydrophobic interactions with electrostatic forces and hydrogen bonding providing specificity. A major part of metabolism is the creation of complex molecules that aggregate spontaneously to generate structure. This structure includes the lipid-rich cytoplasmic membranes which, together with embedded carrier proteins, control the entry of substances into cells. Clearly, the cell is now much more than a bag of enzymes, containing several compartments, each of which contains its own array of enzymes and other components. Metabolite concentrations may vary greatly from one compartment to another.

The reactions that modify lipids and glycoproteins provide a driving force that assists in moving membrane materials generated internally into the outer surface of cells. Other processes, including the breakdown by lysosomal enzymes, help to recycle membrane materials. Oxidative attack on hydrophobic materials such as the sterols and the fatty acids of membrane lipids results in their conversion into more soluble substances which can be degraded and completely oxidized. The flow of matter within cells tends to occur in metabolic loops and some of these loops lead to formation of membranes and organelles and to their turnover. This flow of matter, which is responsible for growth and development of cells, is driven both by hydrolysis of ATP coupled to biosynthesis and by irreversible degradative alterations of polymers and lipid materials. It also provides for transient formation and breakup of complexes of macromolecules, which may be very large, in response to varying metabolic needs.

Anything that affects the rate of a reaction involved in either biosynthesis or degradation of any component of the cell will affect the overall picture in some way. Thus, every chemical reaction that contributes to a quantitatively significant extent to metabolism has some controlling influence. Since molecules interact with each other in so many ways, reactions of metabolic control are innumerable. Small molecules act on macromolecules as effectors that influence conformation and reactivity. Enzymes act on each other to break covalent bonds, to oxidize, and to crosslink. Transferases add phospho, glycosyl, methyl, and other groups to various sites. The resulting alterations often affect catalytic activities. The number of such interactions significant to metabolic control within an organism may be in the millions. Small wonder that biochemical journals are filled with a confusing number of postulated control mechanisms.

Despite this complexity, some regulatory mechanisms stand out clearly. The control of enzyme synthesis through feedback repression and the rapid control of activity by feedback inhibition (Chapter 11) have been considered previously. Under some circumstances, in which there is a constant growth rate, these controls may be sufficient to ensure the harmonious and proportional increase of all constitutents of a cell. Such may be the case for bacteria during logarithmic growth (Box 9-B) or for a mammalian embryo growing rapidly and drawing all its nutrients from the relatively constant supply in the maternal blood.

Contrast the situation in an adult. Little growth takes place, but the metabolism must vary with time and physiological state. The body must make drastic readjustments from normal feeding to a starvation situation and from resting to heavy exercise. The metabolism needed for rapid exertion is different from that needed for sustained work. A fatty diet requires different metabolism than a high-carbohydrate diet. The necessary control mechanisms must be rapid and sensitive.

1. Glycogen and Blood Glucose

Two special features of glucose metabolism in animals are dominant.[224] The first is the storage of glycogen for use in providing muscular energy rapidly. This is a relatively short-term matter but the rate of glycolysis can be intense: The entire glycogen content of muscle could be exhausted in only 20 s of anaerobic fermentation or in 3.5 min of oxidative metabolism.[225] There must be a way to turn on glycolysis quickly and to turn it off when it is no longer needed. At the same time, it must be possible to reconvert lactate to glucose or glycogen (gluconeogenesis). The glycogen stores of the muscle must be repleted from glucose of the blood. If insufficient glucose is available from the diet or from the glycogen stores of the liver, it must be synthesized from amino acids.

The second special feature of glucose metabolism is that certain tissues, including brain, blood cells, kidney medulla, and testis, ordinarily obtain most of their energy through oxidation of glucose.[226,227] For this reason, the glucose level of blood cannot be allowed to drop much below the normal 5 mM. The mechanism of regulation of the blood glucose level is complex and incompletely understood. A series of hormones are involved.

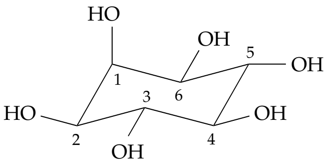
Insulin. This 51-residue cross-linked polypeptide (Fig. 7-17) is synthesized in the pancreatic islets of Langerhans, a tissue specialized for synthesis and secretion into the bloodstream of a series of small peptide hormones. One type of islet cells, the β cells, forms primarily insulin which is secreted in response to high (> 5mm) blood [glucose].[228] Insulin has a wide range of effects on metabolism,[228a] which are discussed in Chapter 11, Section G. Most of these effects are thought to arise from binding to insulin receptors (Figs. 11-11 and 11-12) and are mediated by cascades such as that pictured in Fig. 11-13.[229–232c] The end result is to increase or decrease activities of a large number of enzymes as is indicated in Table 17-3. Some of those are also shown in Fig. 17-20, which indicates interactions with the tricarboxylic acid cycle and lipid metabolism. Binding of insulin to the extracellular domain of its dimeric receptor induces a conformational change that activates the intracellular tyrosine kinase domains of the two subunits. Recent studies suggest that in the activated receptor the two transmembrane helices and the internal tyrosine kinase domains move closer together, inducing the essential autophosphosphorylation.[232b] The kinase domain of the phosphorylated receptor, in turn, phosphorylates several additional proteins, the most important of which seem to be the insulin receptor substrates IRS-1 and IRS-2. Both appear to be essential in different tissues.[232d,e] Phosphorylated forms of these proteins initiate a confusing variety of signaling cascades.[232f–i]

One of the most immediate effects of insulin is to stimulate an increased rate of uptake of glucose by muscle and adipocytes (fat cells) and other insulin-sensitive tissues. This uptake is accomplished largely by movement of the glucose transporter GLUT4 (Chapter 8) from internal “sequestered” storage vesicles located near the cell membrane into functioning positions in the membranes.[232f,j–m] Activation of this translocation process apparently involves IRS-1 and phosphatidylinositol (PI) 3-kinase, which generates PI-3,4,5-P_3 (Fig. 11-9).[232c,n,o] The latter induces the translocation. However, the mechanism remains obscure. The process may also require a second signaling pathway which involves action of the insulin receptor kinase on an adapter protein known as **CAP**, a transmembrane caveolar protein **flotillin**, and a third protein **Cbl**, a known cellular protooncogene. Phosphorylated Cbl forms a complex with CAP and flotillin in a “lipid raft” which induces the exocytosis of the sequestered GLUT4 molecules.[232o,p]

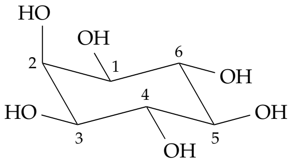
TABLE 17-3
Some Effects of Insulin on Enzymes

Name of Enzyme	Type of Regulation
A. Activity increased	
Enzymes of glycolysis	
Glucokinase	Transcription induced
Phosphofructokinase	via 2,6-fructose P_2
Pyruvate kinase	Dephosphorylation
6-Phosphofructo-2-kinase	Dephosphorylation
Enzymes of glycogen synthesis	
Glucokinase	Transcription
Glycogen synthase (muscle)	Dephosphorylation
Enzymes of lipid synthesis	
Pyruvate dehydrogenase (adipose)	Dephosphorylation (Eq. 17-9)
Acetyl-CoA carboxylase	Dephosphorylation
ATP-citrate lyase	Phosphorylation
Fatty acid synthase	
Lipoprotein lipase	
Hydroxymethylglutaryl-CoA reductase	
B. Activity decreased	
Enzymes of gluconeogenesis	
Pyruvate carboxylase	
PEP carboxykinase	Transcription inhibited
Fructose 1,6-bisphosphate	
Glucose 6-phosphatase	
Enzymes of lipolysis	
Triglyceride lipase	
(hormone-sensitive lipase)	Dephosphorylation
Enzymes of glycogenolysis	
Glycogen phosphorylase	
C. Other proteins affected by insulin	
Glucose transporter GLUT4	Redistribution
Ribosomal protein S6	Phosphorylation by $p90^{rsk}$
IGF-II receptor	Redistribution
Transferrin receptor	Redistribution
Calmodulin	Phosphorylation

A clue to another possible unrecognized mechanism of action for insulin comes from the observation that urine of patients with non-insulin-dependent diabetes contains an unusual isomer of inositol, D-*chiro*-inositol.[233,234]



D-*chiro*-Inositol



myo-Inositol
(numbered as D)

Plasma of such individuals contains an antagonist of insulin action, an inositol phosphoglycan containing *myo*-inositol as a cyclic 1,2-phosphate ester and galactosamine and mannose in a 1:1:3 ratio.[235] This appears to be related to the glycosyl phosphatidylinositol (GPI) membrane anchors (Fig. 8-13). It has been suggested that such a glycan, perhaps containing *chiro*-inositol, is released in response to insulin and serves as a second messenger for insulin.[235–236a] This hypothesis remains unproved.[237] However, insulin does greatly stimulate a GPI-specific phospholipase C, at least in yeast.[237a] Another uncertainty surrounds the possible cooperation of chromium (Chapter 16) in the action of insulin.

How do the insulin-secreting pancreatic β cells sense a high blood glucose concentration? Two specialized proteins appear to be involved. The sugar transporter **GLUT2** allows the glucose in blood to equilibrate with the free glucose in the β cells,[237b] while **glucokinase** (hexokinase IV or D) apparently serves as the glucose sensor.[228,238] Despite the fact that glucokinase is a monomer, it displays a cooperative behavior toward glucose binding, having a low affinity at low [glucose] and a high affinity at high [glucose]. Mutant mice lacking the glucokinase gene develop early onset diabetes which is mild in heterozygotes but severe and fatal within a week of birth for homozygotes.[239,240] These facts alone do not explain how the sensor works and there are doubtless other components to the signaling system.

A current theory is that the increased rate of glucose catabolism in the β cells when blood [glucose] is high leads to a high ratio of [ATP]/[ADP] which induces closure of ATP-sensitive K^+ channels and opening of voltage-gated Ca^{2+} channels.[241] This could explain the increase in [Ca^{2+}] within β cells which has been associated with secretion of insulin[242,243] and which is thought to induce the exocytosis in insulin storage granules. The internal [Ca^{2+}] in pancreatic islet cells is observed to oscillate in a characteristic way that is synchronized with insulin secretion.[243]

Glucagon. This 29-residue peptide is the principal hormone that counteracts the action of insulin. Glucagon acts primarily on liver cells (hepatocytes) and adipose tissue and is secreted by the α cells of the islets of Langerhans in the pancreas, the same tissue whose β cells produce insulin, if the blood glucose concentration falls much below 2 mM.[244–250] Like the insulin-secreting β cells, the pancreatic α cells contain glucokinase, which may be involved in sensing the drop in glucose concentration. However, the carrier GLUT2 is not present and there is scant information on the sensing mechanism.[248]

Glucagon promotes an increase in the blood glucose level by stimulating breakdown of liver glycogen, by inhibiting its synthesis, and by stimulating gluconeogenesis. All of these effects are mediated by cyclic AMP through cAMP-activated protein kinase (Fig. 11-4) and through fructose 2,6-P_2 (Fig. 11-4 and next section). Glucagon also has a strong effect in promoting the release of glucose into the bloodstream. **Adrenaline** has similar effects, again mediated by cAMP. However, glucagon affects the liver, while adrenaline affects many tissues. **Glucocorticoids** such as cortisol (Chapter 22) also promote gluconeogenesis and the accumulation of glycogen in the liver through their action on gene transcription.

The release of glucose from the glycogen stores in the liver is mediated by **glucose 6-phosphatase**, which is apparently embedded within the membranes of the endoplasmic reticulum. A labile enzyme, it consists of a 357-residue catalytic subunit,[251,252] which may be associated with other subunits that participate in transport.[252,253] A deficiency of this enzyme causes the very severe type 1a **glycogen storage disease** (see Box 20-D).[251,253] Only hepatocytes have significant glucose 6-phosphatase activity.

2. Phosphofructo-1-Kinase in the Regulation of Glycolysis

The metabolic interconversions of glucose 1-*P*, glucose 6-*P*, and fructose 6-*P* are thought to be at or near equilibrium within most cells. However, the phosphorylation by ATP of fructose 6-*P* to fructose 1,6-P_2 catalyzed by phosphofructose-1-kinase (Fig. 11-2, step *b*; Fig. 17-17, top center) is usually far from equilibrium. This fact was established by comparing the mass action ratio [fructose 1,6-P_2] [ADP]/[fructose 6-P] [ATP] measured within tissues with the known equilibrium constant for the reaction. At equilibrium this mass action ratio should be equal to the equilibrium constant (Section I,2). The experimental techniques for determining the four metabolite concentrations that are needed for evaluation of the mass action ratio in tissues are of interest. The tissues must be frozen very rapidly. This can be done by compressing them between large liquid nitrogen-cooled aluminum clamps. For details see Newsholme and Start,[225] pp. 30 – 32. Tissues can be cooled to – 80°C in less than 0.1 s in this manner. The frozen tissue is then powdered, treated with a frozen protein denaturant such as perchloric acid, and analyzed. From data obtained in this way, a mass action ratio of 0.03 was found for the phosphofructo-1-kinase reaction in heart muscle.[225] This is much lower than the equilibrium constant of over 3000 calculated from the value of $\Delta G'$ (pH 7) = -20.1 kJ mol^{-1}. Thus, like other biochemical reactions that are nearly irreversible thermodynamically, this reaction is far from equilibrium in tissues.

The effects of ATP, AMP, and fructose 2,6-bisphosphate on phosphofructokinase have been discussed in Chapter 11, Section C. Fructose 2,6-P_2 is a potent allosteric activator of phosphofructokinase and a strong competitive inhibitor of fructose 1,6-bisphosphatase (Fig. 11-2). It is formed from fructose 6-*P* and ATP by the 90-kDa bifunctional phosphofructo-2-kinase/fructose 2,6-bisphosphatase. Thus, the same protein forms and destroys this allosteric effector. Since the bifunctional enzyme is present in very small amounts, the rate of ATP destruction from the substrate cycling is small.

Glucagon causes the concentration of *liver* fructose 2,6-P_2 to drop precipitously from its normal value. This, in turn, causes a rapid drop in glycolysis rate and shifts metabolism toward gluconeogenesis. At the same time, liver glycogen breakdown is inceased and glucose is released into the bloodstream more rapidly. The effect on fructose 2,6-P_2 is mediated by a cAMP-dependent protein kinase which phosphorylates the bifunctional kinase/phosphatase in the liver.[254] This modification greatly reduces the kinase activity and strongly activates the phosphatase, thereby destroying the fructose 2,6-P_2. The changes in activity appear to be largely a result of changes in the appropriate K_m values which are increased for fructose 6-*P* and decreased for fructose 2,6-P_2.[255]

3. Gluconeogenesis

If a large amount of lactate enters the liver, it is oxidized to pyruvate which enters the mitochondria. There, part of it is oxidized through the tricarboxylic acid cycle. However, if [ATP] is high, pyruvate dehydrogenase is inactivated by phosphorylation (Eq. 17-9) and the amount of pyruvate converted to oxaloacetate and malate (Eq. 17-46) may increase. Malate may leave the mitochondrion to be reoxidized to oxaloacetate, which is then converted to PEP and on to glycogen (heavy green arrows in Fig. 17-20). When [ATP] is high, phosphofructokinase is also blocked, but the fructose 1,6-bisphosphatase, which hydrolyzes one phosphate group from fructose 1,6-P_2 (Fig. 11-2, step *d*), is active. If the glucose content of blood is low, the glucose 6-*P* in the liver is hydrolyzed and free glucose is secreted. Otherwise, most of the glucose 6-*P* is converted to glycogen. Muscle is almost devoid of glucose 6-phosphatase, the export of glucose not being a normal activity of that tissue.

Gluconeogenesis in liver is strongly promoted by glucagon and adrenaline. The effects, mediated by cAMP, include stimulation of fructose 1,6-bisphosphatase and inhibition of phosphofructo-1-kinase, both caused by the drop in the level of fructose 2,6-P_2.[254,256] The conversion of pyruvate to PEP via oxaloacetate is also promoted by glucagon. This occurs primarily by stimulation of pyruvate carboxylase (Eq. 14-3).[257,258] However, it has been suggested that the most important mechanism by which glucagon enhances gluconeogenesis is through stimulation of mitochondrial respiration, which in turn may promote gluconeogenesis.[257]

The conversion of oxaloacetate to PEP by PEP-carboxykinase (PEPCK, Eq. 14-43; Fig. 17-20) is another control point in gluconeogenesis. Insulin inhibits gluconeogenesis by decreasing transcription of the mRNA for this enzyme.[259–261a] Glucagon and cAMP stimulate its transcription. The activity of PEP carboxykinase[262] is also enhanced by Mn^{2+} and by very low concentrations of Fe^{2+}. However, the enzyme is readily inactivated by Fe^{2+} and oxygen.[263] Any regulatory significance is uncertain.

Although the regulation of gluconeogenesis in the liver may appear to be well understood, some data indicate that the process can occur efficiently in the presence of high average concentrations of fructose 2,6-P_2. A possible explanation is that liver consists of several types of cells, which may contain differing concentrations of this inhibitor of gluconeogenesis.[264] However, mass spectroscopic studies suggest that glucose metabolism is similar throughout the liver.[265]

4. Substrate Cycles

The joint actions of phosphofructokinase and fructose 1,6-bisphosphatase (Fig. 11-2, steps *b* and *c*; see also Fig. 17-20) create a substrate cycle of the type discussed in Chapter 11, Section F. Such cycles apparently accomplish nothing but the cleavage of ATP to ADP and P_i (ATPase activity). There are many cycles of this type in metabolism and the fact that they do not ordinarily cause a disastrously rapid loss of ATP is a consequence of the tight control of the metabolic pathways involved. In general, only one of the two enzymes of Fig. 11-2, steps *b* and *c*, is fully activated at any time. Depending upon the metabolic state of the cell, degradation may occur with little biosynthesis or biosynthesis with little degradation. Other obvious substrate cycles involve the conversion of glucose to glucose 6-*P* and hydrolysis of the latter back to glucose (Fig. 17-20, upper left-hand corner), the synthesis and breakdown of glycogen (upper right), and the conversion of PEP to pyruvate and the reconversion of the latter to PEP via oxaloacetate and malate (partially within the mitochondria).

While one might suppose that cells always keep substrate cycling to a bare minimum, experimental measurements on tissues *in vivo* have indicated surprisingly high rates for the fructose 1,6-bisphosphatase–phosphofructokinase cycle in mammalian tissues when glycolytic flux rates are low and also for the pyruvate → oxaloacetate → PEP → pyruvate cycle.[266] As pointed out in Chapter 11, by maintaining a low rate of substrate cycling under conditions in which the carbon flux is low (in either the glycolytic or glucogenic direction) the system is more sensitive to allosteric effectors than it would be otherwise. However, when the flux through the glycolysis pathway is high the relative amount of cycling is much less and the amount of ATP formed approaches the theoretical 2.0 per glucose.[267]

Substrate cycles generate heat, a property that is apparently put to good use by cold bumblebees whose thoracic temperature must reach at least 30°C before they can fly. The insects apparently use the fructose bisphosphatase–phosphofructokinase substrate cycle (Fig. 11-2, steps *b* and *c*) to warm their flight muscles.[268] It probably helps to keep us warm, too.

5. Nuclear Magnetic Resonance, Isotopomer Analysis, and Modeling of Metabolism

As as been pointed out in Boxes 3-C and 17-C, the use of ^{13}C and other isotopic tracers together with NMR and mass spectroscopy have provided powerful tools for understanding the complex interrelationships among the various interlocking pathways of metabolism. In Box 17-C the application of ^{13}C NMR to the citric acid cycle was described. Similar approaches have been used to provide direct measurement of the glucose concentration in human brain (1.0 ± 0.1 mM; 4.7 ± 0.3 mM in plasma)[226] and to study gluconeogenesis[269–271] as well as fermentation.[271a] Similar investigations have been made using mass spectroscopy.[272] The metabolism of acetate through the glycoylate pathway in yeast has been observed by ^{13}C NMR.[200] Data obtained from such experiments are being used in attempts to model metabolism and to understand how flux rates through the various pathways are altered in response to varying conditions.[65,273–276]

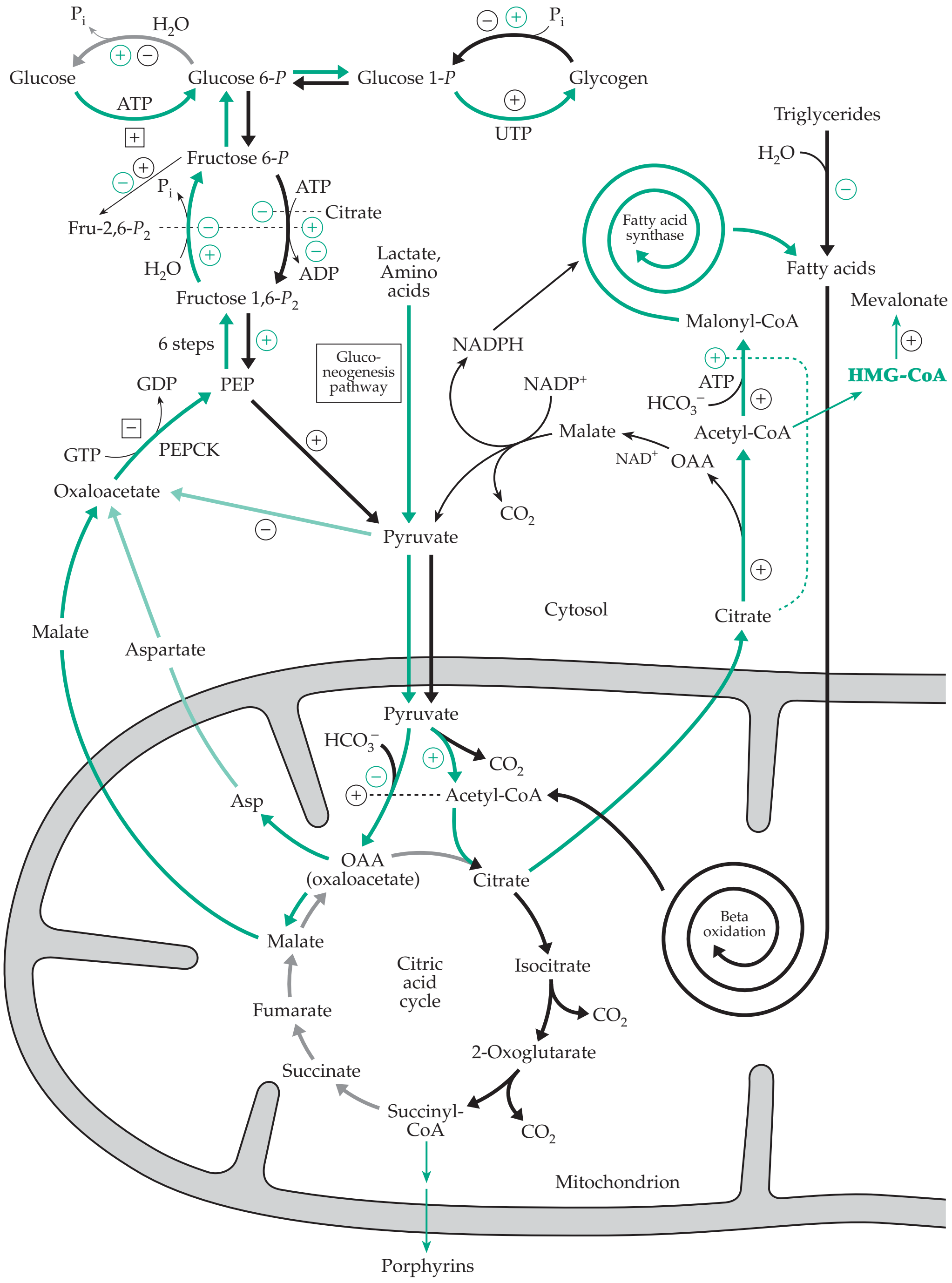

P_i
H_2O
⊕ ⊖
Glucose
Glucose 6-P
ATP
⊞
Glucose 1-P
⊖ ⊕
P_i
Glycogen
⊕
UTP
Triglycerides
H_2O
⊖
Fatty acids
Fatty acid synthase
Fructose 6-P
⊖ ⊕
P_i
ATP
⊖
Citrate
⊕
Fru-2,6-P_2
⊖
⊖
⊕
H_2O
ADP
Fructose 1,6-P_2
Lactate, Amino acids
6 steps
⊕
Gluco-neogenesis pathway
NADPH
NADP⁺
Malonyl-CoA
Mevalonate
⊕
HMG-CoA
GDP
PEP
⊟
GTP
PEPCK
⊕
ATP
⊕
HCO_3^-
⊕
Malate
Acetyl-CoA
NAD^+
OAA
Oxaloacetate
CO_2
⊖
Pyruvate
⊕
Malate
Aspartate
Cytosol
Citrate
Pyruvate
HCO_3^-
⊕
CO_2
⊖
Asp
⊕
Acetyl-CoA
OAA (oxaloacetate)
Citrate
Beta oxidation
Malate
Citric acid cycle
Isocitrate
CO_2
Fumarate
2-Oxoglutarate
Succinate
Succinyl-CoA
CO_2
Mitochondrion
Porphyrins

Figure 17-20 The interlocking pathways of glycolysis, gluconeogenesis, and fatty acid oxidation and synthesis with indications of some aspects of control in hepatic tissues. (⟶) Reactions of glycolysis, fatty acid degradation, and oxidation by the citric acid cycle. (⟶) Biosynthetic pathways. Some effects of insulin via indirect action on enzymes ⊕, ⊖, or on transcription ⊞, ⊟. Effects of glucagon ⊕, ⊖.

6. The Fasting State

During prolonged fasting, glycogen supplies are depleted throughout the body and fats become the principal fuels. Both glucose and pyruvate are in short supply. While the hydrolysis of lipids provides some glycerol (which is phosphorylated and oxidized to dihydroxyacetone-*P*), the quantity of glucose precursors formed in this way is limited. Since the animal body cannot reconvert acetyl-CoA to pyruvate, there is a continuing need for both glucose and pyruvate. The former is needed for biosynthetic processes, and the latter is a precursor of oxaloacetate, the regenerating substrate of the citric acid cycle. For this reason, during fasting the body readjusts its metabolism. As much as 75% of the glucose need of the brain can gradually be replaced by ketone bodies derived from the breakdown of fats (Section A,4).[277] Glucocorticoids (e.g., cortisol; Chapter 22) are released from the adrenal glands. By inducing enzyme synthesis, these hormones increase the amounts of a variety of enzymes within the cells of target organs such as the liver. Glucocorticoids also appear to increase the sensitivity of cell responses to cAMP and hence to hormones such as glucagon.[268]

The overall effects of glucocorticoids include an increased release of glucose from the liver (increased activity of glucose 6-phosphatase), an elevated blood glucose and liver glycogen, and a decreased synthesis of mucopolysaccharides. The reincorporation of amino acids released by protein degradation is inhibited and synthesis of enzymes degrading amino acids is enhanced. Among these enzymes are tyrosine and alanine aminotransferases, enzymes that initiate amino acid degradation which gives rise to the glucogenic precursors fumarate and pyruvate.

The inability of the animal body to form the glucose

BOX 17-F LACTIC ACIDEMIA AND OTHER DEFICIENCIES IN CARBOHYDRATE METABOLISM

The lactate concentration in blood can rise from its normal value of 1–2 mM to as much as 22 mM after very severe exercise such as sprinting, but it gradually returns to normal, requiring up to 6–8 h, less if mild exercise is continued. However, continuously high lactic acid levels are observed when enzymes of the gluconeogenic pathway are deficient or when oxidation of pyruvate is partially blocked.[a,b] Severe and often lethal deficiencies of the four key gluconeogenic enzymes pyruvate carboxylase, PEP carboxykinase, fructose 1,6-bisphosphatase, and glucose 6-phosphatase are known.[b] Pyruvate carboxylase deficiency may be caused by a defective carboxylase protein, by an absence of the enzyme that attaches biotin covalently to the three mitochondrial biotin-containing carboxylases (Chapter 14, Section C), or by defective transport of biotin from the gut into the blood. The latter types of deficiency can be treated successfully with 10 mg biotin per day.

Deficiency of pyruvate dehydrogenase is the most frequent cause of lactic acidemia.[a,c] Since this enzyme has several components (Fig. 15-15), a number of forms of the disease have been observed. Patients are benefitted somewhat by a high-fat, low-carbohydrate diet. Transient lactic acidemia may result from infections or from heart failure. One treatment is to administer dichloroacetate, which stimulates increased activity of pyruvate dehydrogenase, while action is also taken to correct the underlying illness.[d] Another problem arises if a lactate transporter is defective so that lactic acid accumulates in muscles.[e]

A different problem results from deficiency of enzymes of glycolysis such as phosphofructokinase (see Box 20-D), phosphoglycerate mutase, and pyruvate kinase. Lack of one isoenzyme of phosphoglycerate mutase in muscle leads to intolerance to strenuous exercise.[f] A deficiency in pyruvate kinase is one of the most common defects of glycolysis in erythrocytes and leads to a shortened erythrocyte lifetime and hereditary hemolytic anemia.[g]

Deficiency of the first enzyme of the pentose phosphate pathway, glucose 6-phosphate dehydrogenase, is widespread.[h] Its geographical distribution suggests that, like the sickle-cell trait, it confers some resistance to malaria. A partial deficiency of 6-phosphogluconolactonase (Eq. 17-12, step *b*) has also been detected within a family and may have contributed to the observed hemolytic anemia.[i]

---

[a] Robinson, B. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1479–1499, McGraw-Hill, New York

[b] Robinson, B. H. (1982) *Trends Biochem. Sci.* **7**, 151–153

[c] McCartney, R. G., Sanderson, S. J., and Lindsay, J. G. (1997) *Biochemistry* **36**, 6819–6826

[d] Stacpoole, P. W., 17 other authors, and Dichloroacetate-Lactic Acidosis Study Group (1992) *N. Engl. J. Med.* **327**, 1564–1569

[e] Fishbein, W. N. (1986) *Science* **234**, 1254–1256

[f] DiMauro, S., Mirando, A. F., Khan, S., Gitlin, K., and Friedman, R. (1981) *Science* **212**, 1277–1279

[g] Tanaka, K. R., and Paglia, D. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3485–3511, McGraw-Hill, New York

[h] Pandolfi, P. P., Sonati, F., Rivi, R., Mason, P., Grosveld, F., and Luzzatto, L. (1995) *EMBO J.* **14**, 5209–5215

[i] Beutler, E., Kuhl, W., and Gelbart, J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3876–3878

precursors pyruvate or oxaloacetate from acetyl units is sometimes a cause of severe metabolic problems. Ketosis, which was discussed in Section A,4, develops when too much acetyl-CoA is produced and not efficiently oxidized in the citric acid cycle. Ketosis occurs during starvation, with fevers, and in insulin-dependent diabetes (see also Box 17-G). In cattle, whose metabolism is based much more on acetate than is ours, spontaneously developing ketosis is a frequent problem.

7. Lipogenesis

A high-carbohydrate meal leads to an elevated blood glucose concentration. The glycogen reserves within cells are filled. The ATP level rises, blocking the citric acid cycle, and citrate is exported from mitochondria (Fig. 17-20). Outside the mitochondria citrate is cleaved by the ATP-requiring citrate lyase (Eq. 14-37) to acetyl-CoA and oxaloacetate. The oxaloacetate can be reduced to malate and the latter oxidized with $NADP^+$ to pyruvate (Eq. 17-46), which can again enter the mitochondrion. In this manner acetyl groups are exported from the mitochondrion as acetyl-CoA which can be carboxylated, under the activating influence of citrate, to form malonyl-CoA, the precursor of fatty acids. The NADPH formed from oxidation of the malate provides part of the reducing equivalents needed for fatty acid synthesis. Additional NADPH is available from the pentose phosphate pathway. Thus, excess carbohydrate is readily converted into fat by our bodies. These reactions doubtless occur to some extent in most cells, but they are quantitatively

BOX 17-G DIABETES MELLITUS

The most prevalent metabolic problem affecting human beings is diabetes mellitus.[a–c] Out of a million people about 400 develop **type I** (or juvenile-onset) **insulin-dependent diabetes mellitus** (IDDM) between the ages of 8 and 12. Another 33,000 (over 3%) develop diabetes by age 40–50, and by the late 70s over 7% are affected. A propensity toward diabetes is partially hereditary, and recessive susceptibility genes are present in a high proportion of the population. The severity of the disease varies greatly. About half of the type I patients can be treated by diet alone, while the other half must receive regular insulin injections because of the atrophy of the insulin-producing cells of the pancreas. Type I diabetes sometimes develops very rapidly with only a few days of ravenous hunger and unquenchable thirst before the onset of ketoacidosis. Without proper care death can follow quickly. This suggested that a virus infection might cause the observed death of the insulin-secreting β cells of the pancreatic islets. However, the disease appears to be a direct result of an autoimmune response (Chapter 31). Antibodies directed against such proteins as insulin, glutamate decarboxylase,[d,e] and a tyrosine phosphatase[f] of the patient's own body are present in the blood. There may also be a direct attack on the β cells by T cells of the immune system (see Chapter 31).[g,h] The events that trigger such autoimmune attacks are not clear, but there is a strong correlation with susceptibility genes, in both human beings[i,j] and mice.[k,l]

Adults seldom develop type I diabetes but often suffer from **type II** or **non-insulin-dependent diabetes mellitus** (NIDDM). This is not a single disease but a syndrome with many causes. There is usually a marked decrease in sensitivity to insulin (referred to as **insulin resistance**) and poor uptake and utilization of glucose by muscles.[m] In rare cases this is a result of a mutation in the gene for the insulin molecule precursors (Eq. 10-8) or in the gene regulatory regions of the DNA.[n,o] Splicing of the mRNA[p] may be faulty or there may be defects in the structure or in the mechanisms of activation of the insulin receptors (Figs. 11-11 and 11-12).[q] The number of receptors may be too low or they may be degraded too fast to be effective. About 15% of persons with NIDDM have mutations in the insulin substrate protein IRS-1 (Chapter 11, Section G) but the significance is not clear.[m,r] Likewise, the causes of the loss of sensitivity of insulin receptors as well as other aspects of insulin resistance are still poorly understood.[s] In addition, prolonged high glucose concentrations result in decreased insulin synthesis or secretion, both of which are also complex processes. After synthesis the insulin hexamer is stored as granules of the hexamer $(\text{insulin})_6Zn_2$ (Fig. 7-18) in vesicles at low pH. For secretion to occur the vesicles must first dock at membrane sites and undergo exocytosis. The insulin dissolves, releasing the Zn^{2+}, and acts in the monomeric form.[t] Because the mechanisms of action of insulin are still not fully understood, it is difficult to interpret the results of the many studies of diabetes mellitus.

A striking symptom of diabetes is the high blood glucose level which may range from 8 to 60 mM. Lower values are more typical for mild diabetes because when the glucose concentration exceeds the renal threshold of ~8 mM the excess is secreted into the urine. Defective utilization of glucose seems to be tied to a failure of glucose to exert proper

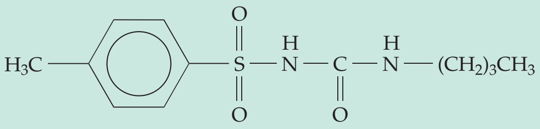
BOX 17-G DIABETES MELLITUS (continued)

feedback control. The result is that gluconeogenesis is increased with corresponding breakdown of proteins and amino acids. The liver glycogen is depleted and excess nitrogen from protein degradation appears in the urine. In IDDM diabetes the products of fatty acid degradation accumulate, leading to ketosis. The volume of urine is excessive and tissues are dehydrated.

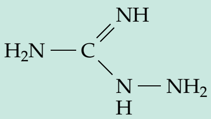
Although the acute problems of diabetes, such as coma induced by ketoacidosis, can usually be avoided, it has not been possible to prevent long-term complications that include cataract formation and damage to the retina and kidneys. Most diabetics eventually become blind and half die within 15–20 years. Many individuals with NIDDM develop insulin-dependent diabetes in later life as a result of damage to the pancreatic β cells. The high glucose level in blood appears to be a major cause of these problems. The aldehyde form of glucose reacts with amino groups of proteins to form Schiff bases which undergo the Amadori rearrangement to form ketoamines (Eq. 4-8). The resulting modified proteins tend to form abnormal disulfide crosslinkages. Crosslinked aggregates of lens proteins may be a cause of cataract. The accumulating glucose-modified proteins may also induce autoimmune responses that lead to the long-term damage to kidneys and other organs. Another problem results from reduction of glucose to sorbitol (Box 20-A). Accumulation of sorbitol in the lens may cause osmotic swelling, another factor in the development of cataracts.[w,x] Excessive secretion of the 37-residue polypeptide **amylin**, which is synthesized in the β cells along with insulin, is another frequent complication of diabetes.[u,v] Amylin precipitates readily within islet cells to form **amyloid deposits** which are characteristic of NIDDM.

For many persons with diabetes regular injections of insulin are essential. Insulin was discovered in 1921 in Toronto by Banting and Best, with a controversial role being played by Professor J. J. R. Macleod, who shared the Nobel prize with Banting in 1923.[y,z] In 1922 the first young patients received pancreatic extracts and a new, prolonged life.[z–bb] Persons with IDDM are still dependent upon daily injections of insulin, but attempts are being made to treat the condition with transplanted cells from human cadavers.[cc] Animal insulins are suitable for most patients, but allergic reactions sometimes make **human insulin** essential. The human hormone, which differs from bovine insulin in three positions (Thr in human vs Ala in bovine at positions 8 of the A chain and 30 of the B chain and Ile vs Val at position 10 of the A chain), is now produced in bacteria using recombinant DNA. Nonenzymatic laboratory synthesis of insulin has also been achieved, but it is difficult to place the disulfide crosslinks in the proper positions. New approaches mimic the natural synthesis, in which the crosslinking takes place in proinsulin (Fig. 10-7).

NIDDM is strongly associated with obesity,[dd] and dieting and exercise often provide adequate control of blood glucose. Sulfonylurea drugs such as the following induce an increase in the number of insulin receptors formed and are also widely used in treatment of the condition.[ee,ff] These drugs bind to and inhibit ATP-sensitive K^+ channels in the β cell membranes. A defect in this sulfonylurea receptor has been associated with excessive insulin secretion in infants.[17] New types of drugs are being tested.[gg–kk] These include inhibitors of aldose reductase,[ii] which forms sorbitol; compounds such as aminoguanidine, which inhibit formation of advanced products of glycation and newly discovered fungal metabolites that activate insulin receptors.[jj]



1-Butyl-*p*-tolylsulfonylurea



Aminoguanidine

---

[a] Atkinson, M. A., and Maclaren, N. K. (1990) *Sci. Am.* **263**(Jul), 62–71

[b] Taylor, S. I. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 843–896, McGraw-Hill, New York

[c] Draznin, B., and LeRoith, D., eds. (1994) *Molecular Biology of Diabetes, Parts I and II*, Humana Press, Totowa, New Jersey

[d] Baekkeskov, S., Aanstoot, H.-J., Christgau, S., Reetz, A., Solimena, M., Cascalho, M., Folli, F., Richter-Olesen, H., and Camilli, P.-D. (1990) *Nature (London)* **347**, 151–156

[e] Nathan, B., Bao, J., Hsu, C.-C., Aguilar, P., Wu, R., Yarom, M., Kuo, C.-Y., and Wu, J.-Y. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 242–246

[f] Lu, J., Li, Q., Xie, H., Chen, Z.-J., Borovitskaya, A. E., Maclaren, N. K., Notkins, A. L., and Lan, M. S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2307–2311

[g] Solimena, M., Dirkx, R., Jr., Hermel, J.-M., Pleasic-Williams, S., Shapiro, J. A., Caron, L., and Rabin, D. U. (1996) *EMBO J.* **15**, 2102–2114

[h] MacDonald, H. R., and Acha-Orbea, H. (1994) *Nature (London)* **371**, 283–284

[i] Todd, J. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8560–8565

BOX 17-G (continued)

[j] Davies, J. L., Kawaguchi, Y., Bennett, S. T., Copeman, J. B., Cordell, H. J., Pritchard, L. E., Reed, P. W., Gough, S. C. L., Jenkins, S. C., Palmer, S. M., Balfour, K. M., Rowe, B. R., Farrall, M., Barnett, A. H., Bain, S. C., and Todd, J. A. (1994) *Nature (London)* **371**, 130–136

[k] Todd, J. A., Aitman, T. J., Cornall, R. J., Ghosh, S., Hall, J. R. S., Hearne, C. M., Knight, A. M., Love, J. M., McAleer, M. A., Prins, J.-B., Rodrigues, N., Lathrop, M., Pressey, A., DeLarato, N. H., Peterson, L. B., and Wicker, L. S. (1991) *Nature (London)* **351**, 542–547

[l] Leiter, E. H. (1989) *FASEB J.* **3**, 2231–2241

[m] Kim, J. K., Gavrilova, O., Chen, Y., Reitman, M. L., and Shulman, G. I. (2000) *J. Biol. Chem.* **275**, 8456–8460

[n] Zhao, L., Cissell, M. A., Henderson, E., Colbran, R., and Stein, R. (2000) *J. Biol. Chem.* **275**, 10532–10537

[o] Catasti, P., Chen, X., Moyzis, R. K., Bradbury, E. M., and Gupta, G. (1996) *J. Mol. Biol.* **264**, 534–545

[p] Wang, J., Shen, L., Najafi, H., Kolberg, J., Matschinsky, F. M., Urdea, M., and German, M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4360–4365

[q] Schmid, E., Hotz-Wagenblatt, A., Hack, V., and Dröge, W. (1999) *FASEB J.* **13**, 1491–1500

[r] Thomas, P. M., Cote, G. J., Wohllk, N., Haddad, B., Mathew, P. M., Rabl, W., Aguilar-Bryan, L., Gagel, R. F., and Bryan, J. (1995) *Science* **268**, 426–429

[s] Nakajima, K., Yamauchi, K., Shigematsu, S., Ikeo, S., Komatsu, M., Aizawa, T., and Hashizume, K. (2000) *J. Biol. Chem.* **275**, 20880–20886

[t] Aspinwall, C. A., Brooks, S. A., Kennedy, R. T., and Lakey, J. R. T. (1997) *J. Biol. Chem.* **272**, 31308–31314

[u] Leighton, B., and Cooper, G. J. S. (1990) *Trends Biochem. Sci.* **15**, 295–299

[v] Lorenzi, A., Razzaboni, B., Weir, G. C., and Yankner, B. A. (1994) *Nature (London)* **368**, 756–760

[w] De Winter, H. L., and von Itzstein, M. (1995) *Biochemistry* **34**, 8299–8308

[x] Wilson, D. K., Bohren, K. M., Gabbay, K. H., and Quiocho, F. A. (1992) *Science* **257**, 81–84

[y] Stevenson, L. G. (1979) *Trends Biochem. Sci.* **4**, N158–N160

[z] Broad, W. J. (1982) *Science* **217**, 1120–1122

[aa] Orci, L., Vassalli, J.-D., and Perrelet, A. (1988) *Sci. Am.* **259**(Sep), 85–94

[bb] Marliss, E. B. (1982) *N. Engl. J. Med.* **306**, 362–364

[cc] Lacy, P. E. (1995) *Sci. Am.* **273**(Jul), 50–58

[dd] Simoneau, J.-A., Colberg, S. R., Thaete, F. L., and Kelley, D. E. (1995) *FASEB J.* **9**, 273–278

[ee] Aguilar-Bryan, L., Nichols, C. G., Wechsler, S. W., Clement, J. P., IV, Boyd, A. E., III, González, G., Herrera-Sosa, H., Nguy, K., Bryan, J., and Nelson, D. A. (1995) *Science* **268**, 423–426

[ff] Eliasson, L., Renström, E., Ämmälä, C., Berggren, P.-O., Bertorello, A. M., Bokvist, K., Chibalin, A., Deeney, J. T., Flatt, P. R., Gäbel, J., Bromada, J., Larsson, O., Lindström, P., Rhodes, C. J., and Rorsman, P. (1996) *Science* **271**, 813–815

[gg] Keen, H. (1994) *N. Engl. J. Med.* **331**, 1226–1227

[hh] Clark, C. M., Jr., and Lee, D. A. (1995) *N. Engl. J. Med.* **332**, 1210–1216

[ii] Bohren, K. M., Grimshaw, C. E., Lai, C.-J., Harrison, D. H., Ringe, D., Petsko, G. A., and Gabbay, K. H. (1994) *Biochemistry* **33**, 2021–2032

[jj] Qureshi, S. A., Ding, V., Li, Z., Szalkowski, D., Biazzo-Ashnault, D. E., Xie, D., Saperstein, R., Brady, E., Huskey, S., Shen, X., Liu, K., Xu, L., Salituro, G. M., Heck, J. V., Moller, D. E., Jones, A. B., and Zhang, B. B. (2000) *J. Biol. Chem.* **275**, 36590–36595

[kk] Moler, D. E. (2001) *Nature (London)* **414**, 821–827

most important in the liver, in fat cells of adipose tissue, and in mammary glands. The process is also facilitated by insulin, which promotes the activation of pyruvate dehydrogenase (Eq. 17-9) and of fatty acid synthase of adipocytes.[277a] Activity of fatty acid synthase seems to be regulated by the rate of transcription of its gene, which is controlled by a transcription factor designated either as **adipocyte determination and differentiation factor-1** (**ADD-1**) or **sterol regulatory element-binding protein-1c** (**SREBP-1c**). This protein (ADD-1/SREBP-1c) may be a general mediator of insulin action.[277b] The nuclear DNA-binding protein known as **peroxisome proliferator-activated receptor gamma** (**PPAR**$_\gamma$) is also involved in the control of insulin action, a conclusion based directly on discovery of mutations in persons with type II diabetes.[277c] A newly discovered hormone **resistin**, secreted by adipocytes, may also play a role.[277d] Another adipocyte hormone, **leptin**, impairs insulin action.[277e] Recent evidence suggests that both insulin and leptin may have direct effects on the brain which also influence blood glucose levels.[277f] Malonyl-CoA, which may also play a role in insulin secretion,[278,279] inhibits carnitine palmitoyltransferase I (CPT I; Fig. 17-2), slowing fatty acid catabolism.[280]

References

1. Knoll, L. J., Schall, O. F., Suzuki, I., Gokel, G. W., and Gordon, J. I. (1995) *J. Biol. Chem.* **270**, 20090 – 20097
2. Thorpe, C., and Kim, J.-J. P. (1995) *FASEB J.* **9**, 718 – 725
3. Vock, P., Engst, S., Eder, M., and Ghisla, S. (1998) *Biochemistry* **37**, 1848 – 1860
4. Aoyama, T., Ueno, I., Kamijo, T., and Hashimoto, T. (1994) *J. Biol. Chem.* **269**, 19088 – 19094
5. Eaton, S., Bartlett, K., and Pourfarzan, M. (1996) *Biochem. J.* **320**, 345 – 357

5a. Dwyer, T. M., Mortl, S., Kemter, K., Bacher, A., Fauq, A., and Frerman, F. E. (1999) *Biochemistry* **38**, 9735 – 9745

5b. Barycki, J. J., O'Brien, L. I., Strauss, A. W., and Banaszak, L. J. (2001) *J. Biol. Chem.* **276**, 36718 – 36726

6. Ikeda, Y., Okamure-Ikeda, K., and Tanaka, K. (1985) *J. Biol. Chem.* **260**, 1311 – 1325
7. Yang, S.-Y., Bittman, R., and Schultz, H. (1985) *J. Biol. Chem.* **260**, 2862 – 2888
8. Yang, S.-Y., Yang, X.-Y. H., Healy-Louie, G., Schulz, H., and Elzinga, M. (1990) *J. Biol. Chem.* **265**, 10424 – 10429
9. Wyatt, J. M. (1984) *Trends Biochem. Sci.* **9**, 20 – 23
10. Nada, M. A., Rhead, W. J., Sprecher, H., Schulz, H., and Roe, C. R. (1995) *J. Biol. Chem.* **270**, 530 – 535
11. Osumi, T., and Hashimoto, T. (1984) *Trends Biochem. Sci.* **9**, 317 – 319
12. Kaikaus, R. M., Sui, Z., Lysenko, N., Wu, N. Y., Ortiz de Montellano, P. R., Ockner, R. K., and Bass, N. M. (1993) *J. Biol. Chem.* **268**, 26866 – 26871
13. Chu, R., Varanasi, U., Chu, S., Lin, Y., Usuda, N., Rao, M. S., and Reddy, J. K. (1995) *J. Biol. Chem.* **270**, 4908 – 4915
14. Masters, C., and Crane, D. (1995) *The Peroxisome: A Vital Organelle*, Cambridge Univ. Press, London
15. Luo, Y., Karpichev, I. V., Kohanski, R. A., and Small, G. M. (1996) *J. Biol. Chem.* **271**, 12068 – 12075
16. Elgersma, Y., van Roermund, C. W. T., Wanders, R. J. A., and Tabak, H. F. (1995) *EMBO J.* **14**, 3472 – 3479
17. Gerhardt, B. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 528 – 565, CRC Press, Boca Raton, Florida
18. Moore, T. S., Jr., ed. (1993) *Lipid Metabolism in Plants*, CRC Press, Boca Raton, Florida

18a. Geisbrecht, B. V., Zhang, D., Schulz, H., and Gould, S. J. (1999) *J. Biol. Chem.* **274**, 21797 – 21803

19. Novikov, D. K., Vanhove, G. F., Carchon, H., Asselberghs, S., Eyssen, H. J., Van Veldhoven, P. P., and Mannaerts, G. P. (1994) *J. Biol. Chem.* **269**, 27125 – 27135

19a. Lopez-Huertas, E., Charlton, W. L., Johnson, B., Graham, I. A., and Baker, A. (2000) *EMBO J.* **19**, 6770 – 6777

20. Orci, L., Vassalli, J.-D., and Perrelet, A. (1988) *Sci. Am.* **259**(Sep), 85 – 94
21. Luthria, D. L., Baykousheva, S. P., and Sprecher, H. (1995) *J. Biol. Chem.* **270**, 13771 – 13776
22. Filppula, S. A., Sormunen, R. T., Hartig, A., Kunau, W.-H., and Hiltunen, J. K. (1995) *J. Biol. Chem.* **270**, 27453 – 27457
23. Smeland, T. E., Cuebas, D., and Schulz, H. (1991) *J. Biol. Chem.* **266**, 23904 – 23908
24. Preisig-Müller, R., Gühnemann-Schäfer, K., and Kindl, H. (1994) *J. Biol. Chem.* **269**, 20475 – 20481
25. van Roermund, C. W. T., Elgersma, Y., Singh, N., Wanders, R. J. A., and Tabak, H. F. (1995) *EMBO J.* **14**, 3480 – 3486
26. Malila, L. H., Siivari, K. M., Mäkelä, M. J., Jalonen, J. E., Latipää, P. M., Kunau, W.-H., and Hiltunen, J. K. (1993) *J. Biol. Chem.* **268**, 21578 – 21585
27. Chen, L.-S., Jin, S.-J., and Tserng, K.-Y. (1994) *Biochemistry* **33**, 10527 – 10534
28. Chen, L.-S., Jin, S.-J., Dejak, I., and Tserng, K.-Y. (1995) *Biochemistry* **34**, 442 – 450
29. Zhang, D., Liang, X., He, X.-Y., Alipui, O. D., Yang, S.-Y., and Schulz, H. (2001) *J. Biol. Chem.* **276**, 13622 – 13627

29a. Mursula, A. M., van Aalten, D. M. F., Hiltunen, J. K., and Wierenga, R. K. (2001) *J. Mol. Biol.* **309**, 845 – 853

29b. Fillgrove, K. L., and Anderson, V. E. (2001) *Biochemistry* **40**, 12412 – 12421

29c. Henke, B., Girzalsky, W., Berteaux-Lecellier, V., and Erdmann, R. (1998) *J. Biol. Chem.* **273**, 3702 – 3711

30. Johnson, M. J. (1967) *Science* **155**, 1515 – 1519
31. Eggink, G., Engel, H., Meijer, W. G., Otten, J., Kingma, J., and Witholt, B. (1988) *J. Biol. Chem.* **263**, 13400 – 13405
32. Stumpf, P. K., ed. (1987) *The Biochemistry of Plants; A Comprehensive Treatise*, Vol. 9, Academic Press, Orlando, Florida

32a. Hamberg, M., Sanz, A., and Castresana, C. (1999) *J. Biol. Chem.* **274**, 24503 – 24513

33. Singh, H., Beckman, K., and Poulos, A. (1994) *J. Biol. Chem.* **269**, 9514 – 9520
34. Vanhove, G. F., Van Veldhoven, P. P., Fransen, M., Denis, S., Eyssen, H. J., Wanders, R. J. A., and Mannaerts, G. P. (1993) *J. Biol. Chem.* **268**, 10335 – 10344

34a. Foulon, V., Antonenkov, V. D., Croes, K., Waelkens, E., Mannaerts, G. P., Van Veldhoven, P. P., and Casteels, M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10039 – 10044

35. Lazarow, P. B., and Moser, H. W. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2287 – 2324, McGraw-Hill, New York
36. Kaya, K., Ramesha, C. S., and Thompson, G. A., Jr. (1984) *J. Biol. Chem.* **259**, 3548 – 3553
37. Hardwick, J. P., Song, B.-J., Huberman, E., and Gonzalez, F. J. (1987) *J. Biol. Chem.* **262**, 801 – 810
38. Cerdan, S., Künnecke, B., Dölle, A., and Seelig, J. (1988) *J. Biol. Chem.* **263**, 11664 – 11674
39. Tserng, K.-Y., and Jin, S.-J. (1991) *J. Biol. Chem.* **266**, 2924 – 2929
40. Jin, S.-J., Hoppel, C. L., and Tserng, K.-Y. (1992) *J. Biol. Chem.* **267**, 119 – 125
41. Bieber, L. L., Emaus, R., Valkner, K., and Farrell, S. (1982) *Fed. Proc.* **41**, 2858 – 2862
42. Bieber, L. L. (1988) *Ann. Rev. Biochem.* **57**, 261 – 283
43. Rebouche, C. J. (1992) *FASEB J.* **6**, 3379 – 3386
44. Ramsay, R. R. (1994) in *Essays in Biochemistry*, Vol. 28 (Tipton, K. F., ed), pp. 47 – 62, Portland Press, London and Chapel Hill, North Carolina
45. Farrell, S. O., Fiol, C. J., Reddy, J. K., and Bieber, L. L. (1984) *J. Biol. Chem.* **259**, 13089 – 13095
46. Brady, P. S., Ramsay, R. R., and Brady, L. J. (1993) *FASEB J.* **7**, 1039 – 1044

46a. Schmalix, W., and Bandlow, W. (1993) *J. Biol. Chem.* **268**, 27428 – 27439

47. Brown, N. F., Weis, B. C., Husti, J. E., Foster, D. W., and McGarry, J. D. (1995) *J. Biol. Chem.* **270**, 8952 – 8957

47a. Yamazaki, N., Shinohara, Y., Kajimoto, K., Shindo, M., and Terada, H. (2000) *J. Biol. Chem.* **275**, 31739 – 31746

48. Kelly, D. P., and Strauss, A. W. (1994) *N. Engl. J. Med.* **330**, 913 – 919
49. Marinetti, G. V. (1990) *Disorders of Lipid Metabolism*, Plenum, New York

49a. Tamai, I., Ohashi, R., Nezu, J.-i, Sai, Y., Kobayashi, D., Oku, A., Shimane, M., and Tsuji, A. (2000) *J. Biol. Chem.* **275**, 40064 – 40072

50. Roe, C. R., and Coates, P. M. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1501 – 1533, McGraw-Hill, New York

50a. Wang, Y., Ye, J., Ganapathy, V., and Longo, N. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2356 – 2360

51. Ford, D. A., Han, X., Horner, C. C., and Gross, R. W. (1996) *Biochemistry* **35**, 7903 – 7909
52. Requero, M. A., Goñi, F. M., and Alonso, A. (1995) *Biochemistry* **34**, 10400 – 10405
53. Treem, W. R., Stanley, C. A., Finegold, D. N., Hale, D. E., and Coates, P. M. (1988) *N. Engl. J. Med.* **319**, 1331 – 1336
54. Rhead, W. J., Amendt, B. A., Fritchman, K. S., and Felts, S. J. (1983) *Science* **221**, 73 – 75
55. Mosser, J., Douar, A.-M., Sarde, C.-O., Kioschis, P., Feil, R., Moser, H., Poustka, A.-M., Mandel, J.-L., and Aubourg, P. (1993) *Nature (London)* **361**, 726 – 730
56. Aubourg, P., Adamsbaum, C., Lavallard-Rousseau, M.-C., Rocchiccioli, F., Cartier, N., Jambaqué, I., Jakobezak, C., Lemaitre, A., Boureau, F., Wolf, C., and Bougnères, P.-F. (1993) *N. Engl. J. Med.* **329**, 745 – 752
57. Moser, H. W., Smith, K. D., and Moser, A. B. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2325 – 2349, McGraw-Hill, New York
58. Yahraus, T., Braverman, N., Dodt, G., Kalish, J. E., Morrell, J. C., Moser, H. W., Valle, D., and Gould, S. J. (1996) *EMBO J.* **15**, 2914 – 2923
59. Street, J. M., Evans, J. E., and Natowicz, M. R. (1996) *J. Biol. Chem.* **271**, 3507 – 3516
60. Wolins, N. E., and Donaldson, R. P. (1994) *J. Biol. Chem.* **269**, 1149 – 1153

60a. Fujiwara, C., Imamura, A., Hashiguchi, N., Shimozawa, N., Suzuki, Y., Kondo, N., Imanaka, T., Tsukamoto, T., and Osumi, T. (2000) *J. Biol. Chem.* **275**, 37271 – 37277

60b. Quant, P. A. (1994) in *Essays in Biochemistry*, Vol. 28 (Tipton, K. F., ed), Portland Press, Chapel Hill, North Carolina

61. Krebs, H. A., Williamson, D. H., Bates, M. W., Page, M. A., and Hawkins, R. A. (1971) *Adv. Enzyme Regul.* **9**, 387 – 409
62. McGarry, J. D., and Foster, D. W. (1980) *Ann. Rev. Biochem.* **49**, 395 – 420
63. Endemann, G., Goetz, P. G., Edmond, J., and Brunengraber, H. (1982) *J. Biol. Chem.* **257**, 3434 – 3440
64. Ohgaku, S., Brady, P. S., Schumann, W. C., Bartsch, G. E., Margolis, J. M., Kumaran, K., Landau, S. B., and Landau, B. R. (1982) *J. Biol. Chem.* **257**, 9283 – 9289

64a. Misra, I., and Miziorko, H. M. (1996) *Biochemistry* **35**, 9610 – 9616

64b. Chun, K. Y., Vinarov, D. A., and Miziorko, H. M. (2000) *Biochemistry* **39**, 14670 – 14681

64c. Chun, K. Y., Vinarov, D. A., Zajicek, J., and Miziorko, H. M. (2000) *J. Biol. Chem.* **275**, 17946 – 17953

65. Sato, K., Kashiwaya, Y., Keon, C. A., Tsuchiya, N., King, M. T., Radda, G. K., Chance, B., Clarke, K., and Veech, R. L. (1995) *FASEB J.* **9**, 651 – 658
66. Casazza, J. P., Felver, M. E., and Veech, R. L. (1984) *J. Biol. Chem.* **259**, 231 – 236
67. Argilés, J. M. (1986) *Trends Biochem. Sci.* **11**, 61 – 63
68. Landau, B. R., and Brunengraber, H. (1987) *Trends Biochem. Sci.* **12**, 113 – 114

References

69. Gavino, V. C., Somma, J., Philbert, L., David, F., Garneau, M., Bélair, J., and Brunengraber, H. (1987) *J. Biol. Chem.* **262**, 6735 – 6740

70. Vagelos, P. R. (1960) *J. Biol. Chem.* **235**, 346 – 350

70a. Sacksteder, K. A., Morrell, J. C., Wanders, R. J. A., Matalon, R., and Gould, S. J. (1999) *J. Biol. Chem.* **274**, 24461 – 24468

71. Hofmeister, A. E. M., and Buckel, W. (1992) *Eur. J. Biochem.* **206**, 547 – 552

72. Buckel, W. (1992) *FEMS Microbiol. Rev.* **88**, 211 – 232

73. Vagelos, P. R., Earl, J. M., and Stadtman, E. R. (1959) *J. Biol. Chem.* **234**, 490 – 497, 765–769

74. Kuchta, R. D., and Abeles, R. H. (1985) *J. Biol. Chem.* **260**, 13181 – 13189

75. Taylor, S. I. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 843 – 896, McGraw–Hill, New York

76. Weigand, E., Young, J. W., and McGilliard, A. D. (1972) *Biochem. J.* **126**, 201 – 209

77. Baldwin, J. E., and Krebs, H. (1981) *Nature (London)* **291**, 381 – 385

78. Nishimura, J. S., and Grinnell, F. (1972) *Adv. Enzymol.* **36**, 183 – 202

79. Rolleson, F. S. (1972) *Curr. Top. Cell. Regul.* **5**, 47 – 75

79a. Jucker, B. M., Lee, J. Y., and Shulman, R. G. (1998) *J. Biol. Chem.* **273**, 12187 – 12194

79b. Petersen, S., de Graaf, A. A., Eggeling, L., Möllney, M., Wiechert, W., and Sahm, H. (2000) *J. Biol. Chem.* **275**, 35932 – 35941

80. Kornberg, H. L. (1966) *Essays Biochem.* **2**, 1 – 31

81. Newsholme, E. A., and Start, C. (1973) *Regulation in Metabolism*, Wiley, New York (pp. 124 – 145)

82. Lieber, C. S. (1976) *Sci. Am.* **234**(Mar), 25 – 33

83. Kalnitsky, G., and Tapley, D. F. (1958) *Biochem. J.* **70**, 28 – 34

84. Rahmatullah, M., and Roche, T. E. (1987) *J. Biol. Chem.* **262**, 10265 – 10271

85. Yudkoff, M., Nelson, D., Daikhin, Y., and Erecinska, M. (1994) *J. Biol. Chem.* **269**, 27414 – 27420

85a. Reed, L. J. (2001) *J. Biol. Chem.* **276**, 38329 – 38336

86. Stueland, C. S., Eck, K. R., Stieglbauer, K. T., and LaPorte, D. C. (1987) *J. Biol. Chem.* **262**, 16095 – 16099

86a. Panisko, E. A., and McAlister-Henn, L. (2001) *J. Biol. Chem.* **276**, 1204 – 1210

87. Rashed, H. M., Waller, F. M., and Patel, T. B. (1988) *J. Biol. Chem.* **263**, 5700 – 5706

88. Chiang, P. K., and Sacktor, B. (1975) *J. Biol. Chem.* **250**, 3399 – 3408

89. Krebs, H. A. (1970) *Adv. Enzyme Regul.* **8**, 335 – 353

90. Maeng, C.-Y., Yazdi, M. A., and Reed, L. J. (1996) *Biochemistry* **35**, 5879 – 5882

91. Harris, R. A., Bowker-Kinley, M. M., Wu, P., Jeng, J., and Popov, K. M. (1997) *J. Biol. Chem.* **272**, 19746 – 19751

92. Patel, M. S., and Roche, T. E. (1990) *FASEB J.* **4**, 3224 – 3233

92a. Yang, D., Gong, X., Yakhnin, A., and Roche, T. E. (1998) *J. Biol. Chem.* **273**, 14130 – 14137

92b. McCartney, R. G., Sanderson, S. J., and Lindsay, J. G. (1997) *Biochemistry* **36**, 6819 – 6826

93. Bessman, S. P., Mohan, C., and Zaidise, I. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5067 – 5070

94. Wan, B., LaNoue, K. F., Cheung, J. Y., and Scaduto, R. C., Jr. (1989) *J. Biol. Chem.* **264**, 13430 – 13439

95. Srere, P. A. (1971) *Adv. Enzyme Regul.* **9**, 221 – 233

96. Ovádi, J., and Srere, P. A. (1996) *Cell Biochem. Funct.* **14**, 249 – 258

97. Sherry, A. D., and Malloy, C. R. (1996) *Cell Biochem. Funct.* **14**, 259 – 268

97a. Vélot, C., Mixon, M. B., Teige, M., and Srere, P. A. (1997) *Biochemistry* **36**, 14271 – 14276

98. Palmer, T. N., and Sugden, M. C. (1983) *Trends Biochem. Sci.* **8**, 161 – 162

99. Loeber, G., Infante, A. A., Maurer-Fogy, I., Krystek, E., and Dworkin, M. B. (1991) *J. Biol. Chem.* **266**, 3016 – 3021

100. Mandella, R. D., and Sauer, L. A. (1975) *J. Biol. Chem.* **250**, 5877 – 5884

101. Macrae, A. R. (1971) *Biochem. J.* **122**, 495 – 501

102. Huh, T.-L., Casazza, J. P., Huh, J.-W., Chi, Y.-T., and Song, B. J. (1990) *J. Biol. Chem.* **265**, 13320 – 13326

103. Nathan, B., Bao, J., Hsu, C.-C., Aguilar, P., Wu, R., Yarom, M., Kuo, C.-Y., and Wu, J.-Y. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 242 – 246

104. Nathan, B., Hsu, C.-C., Bao, J., Wu, R., and Wu, J.-Y. (1994) *J. Biol. Chem.* **269**, 7249 – 7254

105. Toney, M. D., Pascarella, S., and De Biase, D. (1995) *Protein Sci.* **4**, 2366 – 2374

106. Hearl, W. G., and Churchich, J. E. (1985) *J. Biol. Chem.* **260**, 16361 – 16366

107. Baum, G., Lev-Yadum, S., Fridmann, Y., Arazi, T., Katsnelson, H., Zik, M., and Fromm, H. (1996) *EMBO J.* **15**, 2988 – 2996

108. Baum, G., Chen, Y., Arazi, T., Takatsuji, H., and Fromm, H. (1993) *J. Biol. Chem.* **268**, 19610 – 19617

109. Satyanarayan, V., and Nair, P. M. (1990) *Phytochemistry* **29**, 367 – 375

110. Streeter, J. G., and Thompson, J. F. (1972) *Plant Physiol.* **49**, 579 – 584

111. del Rio, R. M. (1981) *J. Biol. Chem.* **256**, 9816 – 9819

112. Bartley, W., Kornberg, H. L., and Quayle, J. R. (1970) *Essays in Cell Metabolism*, Wiley-Interscience, New York (p. 125)

113. Quayle, J. R. (1963) *Biochem. J.* **89**, 492 – 503

114. Chang, Y.-Y., Wang, A.-Y., and Cronan, J. E., Jr. (1993) *J. Biol. Chem.* **268**, 3911 – 3919

115. Fothergill-Gilmore, L. A. (1986) *Trends Biochem. Sci.* **11**, 47 – 51

115a. Bond, C. J., Jurica, M. S., Mesecar, A., and Stoddard, B. L. (2000) *Biochemistry* **39**, 15333 – 15343

115b. Valentini, G., Chiarelli, L., Fortin, R., Speranza, M. L., Galizzi, A., and Mattevi, A. (2000) *J. Biol. Chem.* **275**, 18145 – 18152

115c. Rigden, D. J., Phillips, S. E. V., Michels, P. A. M., and Fothergill-Gilmore, L. A. (1999) *J. Mol. Biol.* **291**, 615 – 635

116. Horecker, B. L., Gibbs, M., Klenow, H., and Smyrniotis, P. Z. (1954) *J. Biol. Chem.* **207**, 393 – 403

117. Wood, T. (1985) *The Pentose Phosphate Pathway*, Academic Press, Orlando, Florida

117a. Vought, V., Ciccone, T., Davino, M. H., Fairbairn, L., Lin, Y., Cosgrove, M. S., Adams, M. J., and Levy, H. R. (2000) *Biochemistry* **39**, 15012 – 15021

117b. Karsten, W. E., Chooback, L., and Cook, P. F. (1998) *Biochemistry* **37**, 15691 – 15697

117c. Kopriva, S., Koprivova, A., and Süss, K.-H. (2000) *J. Biol. Chem.* **275**, 1294 – 1299

118. Bublitz, C., and Steavenson, S. (1988) *J. Biol. Chem.* **263**, 12849 – 12853

118a. Johnson, A. E., and Tanner, M. E. (1998) *Biochemistry* **37**, 5746 – 5754

118b. Chen, Y.-R., Larimer, F. W., Serpersu, E. H., and Hartman, F. C. (1999) *J. Biol. Chem.* **274**, 2132 – 2136

119. Zubay, G. L., Parsons, W. W., and Vance, D. E. (1995) *Principles of Biochemistry*, W.C.Brown, Dubuque, Iowa (p. 276)

120. Reitzer, L. J., Wice, B. M., and Kennell, D. (1980) *J. Biol. Chem.* **255**, 5616 – 5626

121. Scofield, R. E., Kosugi, K., Chandramouli, V., Kumaran, K., Schumann, W. C., and Landau, B. R. (1985) *J. Biol. Chem.* **260**, 15439 – 15444

122. Magnusson, I., Chandramouli, V., Schumann, W. C., Kumaran, K., Wahren, J., and Landau, B. R. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4682 – 4685

122a. Kurland, I. J., Alcivar, A., Bassilian, S., and Lee, W.-N. P. (2000) *J. Biol. Chem.* **275**, 36787 – 36793

123. Longenecker, J. P., and Williams, J. F. (1980) *Biochem. J.* **188**, 847 – 857

124. Williams, J. F. (1980) *Trends Biochem. Sci.* **5**, 315 – 320

125. Williams, J. F. (1983) *Trends Biochem. Sci.* **8**, 275 – 277

126. Flanigan, I., Collins, J. G., Arora, K. K., Macleod, J. K., and Williams, J. F. (1993) *Eur. J. Biochem.* **213**, 477 – 485

127. Landau, B. R., and Wood, H. G. (1983) *Trends Biochem. Sci.* **8**, 292 – 296; 312 – 313

128. Casazza, J. P., and Veech, R. L. (1986) *J. Biol. Chem.* **261**, 690 – 698

129. Bernacchia, G., Schwall, G., Lottspeich, F., Salamini, F., and Bartels, D. (1995) *EMBO J.* **14**, 610 – 618

130. Berthon, H. A., Kuchel, P. W., and Nixon, P. F. (1992) *Biochemistry* **31**, 12792 – 12798

131. Horecker, B. L. (1965) *J. Chem. Educ.* **42**, 244 – 253

132. Decker, K., Jungermann, K., and Thauer, R. K. (1970) *Angew. Chem. Int. Ed. Engl.* **9**, 138 – 158

133. Mushegian, A. R., and Koonin, E. V. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10268 – 10273

134. Bult, C. J., and 39 other authors. (1996) *Science* **273**, 1058 – 1073

135. DiGirolamo, M., Newby, F. D., and Lovejoy, J. (1992) *FASEB J.* **6**, 2405 – 2412

136. Robinson, B. H. (1982) *Trends Biochem. Sci.* **7**, 151 – 153

136a. DiGirolamo, M., Newby, F. D., and Lovejoy, J. (1992) *FASEB J.* **6**, 2405 – 2412

136b. Shulman, R. G., and Rothman, D. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 457 – 461

136c. Gladden, L. B. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 395 – 397

136d. Bouzier, A.-K., Goodwin, R., Macouillard-Poulletier de Gannes, F., Valeins, H., Voisin, P., Canioni, P., and Merle, M. (1998) *J. Biol. Chem.* **273**, 27162 – 27169

137. Roberts, J. K. M., Callis, J., Jardetzky, O., Walbot, V., and Freeling, M. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6029 – 6033

138. Hake, S., Kelley, P. M., Taylor, W. C., and Freeling, M. (1985) *J. Biol. Chem.* **260**, 5050 – 5054

139. Shoubridge, E. A., and Hochachka, P. W. (1980) *Science* **209**, 308 – 309

139a. Teusink, B., Walsh, M. C., van Dam, K., and Westerhoff, H. V. (1998) *Trends Biochem. Sci.* **23**, 162 – 169

139b. Bakker, B. M., Mensonides, F. I. C., Teusink, B., van Hoek, P., Michels, P. A. M., and Westerhoff, H. V. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2087 – 2092

139c. Manchester, K. L. (2000) *Trends Biochem. Sci.* **25**, 89 – 92

140. Tempest, D. W., and Neijssel, O. M. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 797 – 806, Am. Soc. for Microbiology, Washington, D.C.

141. Blackwood, A. C., Neish, A. C., and Ledingham, G. A. (1956) *J. Bacteriol.* **72**, 497 – 499

142. Krebs, H. A. (1972) *Essays Biochem.* **8**, 1 – 34

143. Tse, J. M.-T., and Schloss, J. V. (1993) *Biochemistry* **32**, 10398 – 10403

References

144. Bui, E. T. N., Bradley, P. J., and Johnson, P. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9651–9656
145. Hochachka, P. W., and Mustafa, T. (1972) *Science* **178**, 1056–1060
146. Hochachka, P. W., and Somero, G. N. (1973) *Strategies of Biochemical Adaptation*, Saunders, Philadelphia, Pennsylvania (pp. 46–61)
147. Hochachka, P. W. (1980) *Living Without Oxygen*, Harvard Univ. Press, Cambridge, Massachusetts
148. Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., and Picataggio, S. (1995) *Science* **267**, 240–243
149. Stanier, R. Y., Doudoroff, M., and Adelberg, E. A. (1970) *Microbial World*, 3rd ed., Prentice-Hall, Englewood Cliffs, New Jersey (p. 191)
150. Switzer, R. L. (1974) in *The Enzymes*, 3rd ed., Vol. 10 (Boyer, P. D., ed), pp. 607–629, Academic Press, New York

150a. Baykov, A. A., Fabrichniy, I. P., Pohjanjoki, P., Zyryanov, A. B., and Lahti, R. (2000) *Biochemistry* **39**, 11939–11947

150b. Ahn, S., Milner, A. J., Fütterer, K., Konopka, M., Ilias, M., Young, T. W., and White, S. A. (2001) *J. Mol. Biol.* **313**, 797–811

151. Wood, H. G., and Goss, N. H. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 312–315
152. Wood, H. G. (1977) *Fed. Proc.* **36**, 2197–2205
153. Liu, C.-L., Hart, N., and Peck, H. D., Jr. (1982) *Science* **217**, 363–364
154. Weinbach, E. C. (1981) *Trends Biochem. Sci.* **6**, 254–257
155. Takeshige, K., and Tazawa, M. (1989) *J. Biol. Chem.* **264**, 3262–3266
156. Kornberg, A., Rao, N. N., and Ault-Riché, D. (1999) *Ann. Rev. Biochem.* **68**, 89–125

156a. Rashid, M. H., Rumbaugh, K., Passador, L., Davies, D. G., Hamood, A. N., Iglewski, B. H., and Kornberg, A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9636–9641

156b. Bolesch, D. G., and Keasling, J. D. (2000) *J. Biol. Chem.* **275**, 33814–33819

157. Mudd, S. H. (1973) *The Enzymes*, 3rd ed., Vol. 8, Academic Press, New York (pp. 121–154)
158. Cornforth, J. W., Reichard, S. A., Talalay, P., Carrell, H. L., and Glusker, J. P. (1977) *J. Am. Chem. Soc.* **99**, 7292–7301
159. Peck, H. D., Jr. (1974) in *The Enzymes*, 3rd ed., Vol. 10 (Boyer, P. D., ed), pp. 651–669, Academic Press, New York
160. Geller, D. H., Henry, J. G., Belch, J., and Schwartz, N. B. (1987) *J. Biol. Chem.* **262**, 7374–7382
161. Liu, C., Martin, E., and Leyh, T. S. (1994) *Biochemistry* **33**, 2042–2047
162. Li, H., Deyrup, A., Mensch, J. R., Jr., Domowicz, M., Konstantinidis, A. K., and Schwartz, N. B. (1995) *J. Biol. Chem.* **270**, 29453–29459
163. Moréra, S., Chiadmi, M., LeBras, G., Lascu, I., and Janin, J. (1995) *Biochemistry* **34**, 11062–11070

163a. MacRae, I. J., Segel, I. H., and Fisher, A. J. (2000) *Biochemistry* **39**, 1613–1621

164. Wakil, S. J., Stoops, J. K., and Joshi, V. C. (1983) *Ann. Rev. Biochem.* **52**, 537–579
165. White, H. B., III, Mitsuhashi, O., and Bloch, K. (1971) *J. Biol. Chem.* **246**, 4751–4754
166. Williamson, D. H., Lund, P., and Krebs, H. A. (1967) *Biophys. J.* **103**, 514–527
167. Stubbs, M., Veech, R. L., and Krebs, H. A. (1972) *Biophys. J.* **126**, 59–65
168. Chung, Y., and Jue, T. (1992) *Biochemistry* **31**, 11159–11165
169. Krebs, H. A., and Veech, R. L. (1969) *Mitochondria: Structure & Function*, Academic Press, New York (pp. 101–109)
170. Veech, R. L., Guynn, R., and Veloso, D. (1972) *Biochem. J.* **127**, 387–397
171. Lundquist, R., and Olivera, B. M. (1971) *J. Biol. Chem.* **246**, 1107–1116
172. Buchanan, B. B. (1969) *J. Biol. Chem.* **244**, 4218–4223
173. Gehring, U., and Arnon, D. I. (1972) *J. Biol. Chem.* **247**, 6963–6969

173a. Furdui, C., and Ragsdale, S. W. (2002) *Biochemistry* **41**, 9921–9937

174. Stanier, R. Y., Doudoroff, M., and Adelberg, E. A. (1970) *The Microbial World*, 3rd ed., Prentice-Hall, Englewood Cliffs, New Jersey
175. Stanier, R. Y., Ingraham, J. L., Wheelis, M. L., and Painter, P. R. (1986) *The Microbial World*, 5th ed., Prentice-Hall, Englewood Cliffs, New Jersey
176. Brock, T. D., and Madigan, M. T. (1988) *Microbiology*, Prentice-Hall, Englewood Cliffs, New Jersey
177. Gottschalk, G. (1979) *Bacterial Metabolism*, Springer-Verlag, New York
178. Buchanan, B. B., Schürmann, P., and Shanmugam, K. T. (1972) *Biochim. Biophys. Acta.* **283**, 136–145
179. Wood, H. G., Ragsdale, S. W., and Pezacka, E. (1986) *Trends Biochem. Sci.* **11**, 14–17
180. Ferry, J. G., ed. (1993) *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics*, Chapman & Hall, New York
181. Thauer, R. K., Hedderich, R., and Fischer, R. (1993) in *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics* (Ferry, J. G., ed), pp. 209–252, Chapman & Hall, New York
182. Thauer, R. K., Schwörer, B., and Zirngibl, C. (1993) in *Microbial Growth on C1 Compounds* (Murrell, J. C., and Kelly, D. P., eds), pp. 151–162, Intercept Ltd., Andover, UK
183. Müller, V., Blaut, M., and Gottschalk, G. (1993) in *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics* (Ferry, J. G., ed), pp. 360–406, Chapman & Hall, New York
184. Simpson, P. G., and Whitman, W. B. (1993) in *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics* (Ferry, J. G., ed), pp. 445–472, Chapman & Hall, New York
185. Ljungdahl, L., Irion, E., and Wood, H. G. (1965) *Biochemistry* **4**, 2771–2780
186. Wood, H. G. (1991) *FASEB J.* **5**, 156–163
187. Drake, H. L. (1993) in *Microbial Growth on C1 Compounds* (Murrell, J. C., and Kelly, D. P., eds), pp. 493–507, Intercept Ltd., Andover, UK
188. Wood, H. G., and Ljungdahl, L. G. (1991) in *Variations in Autotrophic Life* (Shively, J. M., and Barton, L. L., eds), pp. 201–250, Academic Press, San Diego, California

188a. Tan, X. S., Sewell, C., and Lindahl, P. A. (2002) *J. Am. Chem. Soc.* **124**, 6277–6284

189. Qiu, D., Kumar, M., Ragsdale, S. W., and Spiro, T. G. (1996) *J. Am. Chem. Soc.* **118**, 10429–10435

189a. Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W., and Drennan, C. L. (2002) *Science* **298**, 567–572

189b. Peters, J. W. (2002) *Science* **298**, 552–553

190. Menon, S., and Ragsdale, S. W. (1996) *Biochemistry* **35**, 12119–12125
191. Hemming, A., and Blotevogel, K. H. (1985) *Trends Biochem. Sci.* **10**, 198–200
192. Salem, A. R., Hacking, A. J., and Quayle, J. R. (1973) *Biochem. J.* **136**, 89–96
193. Quayle, J. R. (1980) *Biochem. Soc. Trans.* **8**, 1–10
194. Anthony, C. (1982) *The Biochemistry of Methylotrophs*, Academic Press, New York
195. Colby, J., Dalton, H., and Whittenbury, R. (1979) *Ann. Rev. Microbiol.* **33**, 481–517
196. Haber, C. L., Allen, L. N., Zhao, S., and Hanson, R. S. (1983) *Science* **221**, 1147–1153
197. Higgins, I. J., Best, D. J., and Hammond, R. C. (1980) *Nature (London)* **286**, 561–567
198. Anthony, C. (1996) *Biochem. J.* **320**, 697–711
199. Stueland, C. S., Gorden, K., and LaPorte, D. C. (1988) *J. Biol. Chem.* **263**, 19475–19479
200. Dickinson, J. R., Dawes, I. W., Boyd, A. S. F., and Baxter, R. L. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5847–5851
201. Liu, F., Thatcher, J. D., and Epstein, H. F. (1997) *Biochemistry* **36**, 255–260
202. Cook, A. G., and Knowles, J. R. (1985) *Biochemistry* **24**, 51–58
203. Pocalyko, D. J., Carroll, L. J., Martin, B. M., Babbitt, P. C., and Dunaway-Mariano, D. (1990) *Biochemistry* **29**, 10757–10765
204. Xu, Y., Yankie, L., Shen, L., Jung, Y.-S., Mariano, P. S., and Dunaway-Mariano, D. (1995) *Biochemistry* **34**, 2181–2187
205. Ohlrogge, J. B. (1982) *Trends Biochem. Sci.* **7**, 386–387

205a. Heath, R. J., Su, N., Murphy, C. K., Rock, C. O. (2000) *J. Biol. Chem.* **275**, 40128–40133

206. Holak, T. A., Kearsley, S. K., Kim, Y., and Prestegard, J. H. (1988) *Biochemistry* **27**, 6135–6142
207. Wood, W. I., Peterson, D. O., and Bloch, K. (1978) *J. Biol. Chem.* **253**, 2650–2656
208. Kolodziej, S. J., Penczek, P. A., Schroeter, J. P., and Stoops, J. K. (1996) *J. Biol. Chem.* **271**, 28422–28429
209. Smith, S. (1994) *FASEB J.* **8**, 1248–1259
210. White, R. H. (1989) *Arch. Biochem. Biophys.* **270**, 691–697
211. Lipmann, F. (1971) *Science* **173**, 875–884
212. Kurahashi, K. (1974) *Ann. Rev. Biochem.* **43**, 445–459
213. Kleinkauf, H., and von Döhren, H. (1983) *Trends Biochem. Sci.* **8**, 281–283
214. Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R., and Morris, H. R. (1996) *J. Biol. Chem.* **271**, 15428–15435
215. Pfeifer, E., Pavela-Vrancic, M., von Döhren, H., and Kleinkauf, H. (1995) *Biochemistry* **34**, 7450–7459
216. Pavela-Vrancic, M., Pfeifer, E., Schröder, W., von Döhren, H., and Kleinkauf, H. (1994) *J. Biol. Chem.* **269**, 14962–14966
217. Haese, A., Pieper, R., von Ostrowski, T., and Zocher, R. (1994) *J. Mol. Biol.* **243**, 116–122
218. Kao, C. M., Pieper, R., Cane, D. E., and Khosla, C. (1996) *Biochemistry* **35**, 12363–12368
219. Cortes, J., Wiesmann, K. E. H., Roberts, G. A., Brown, M. J. B., Staunton, J., and Leadlay, P. F. (1995) *Science* **268**, 1487–1489
220. Pooley, H. M., and Karamata, D. (1994) in *Bacterial Cell Wall (New Comprehensive Biochemistry)*, Vol. 27 (Ghuysen, J.-M., and Hakenbeck, R., eds), pp. 187–198, Elsevier, Amsterdam
221. Knull, H., and Minton, A. P. (1996) *Cell Biochem. Funct.* **14**, 237–248
222. Low, P. S., Rathinavelu, P., and Harrison, M. L. (1993) *J. Biol. Chem.* **268**, 14627–14631
223. Hardin, C. D., and Roberts, T. M. (1995) *Biochemistry* **34**, 1323–1331
224. Milner, Y., and Wood, H. G. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 2463–2468
225. Newsholme, E. A., and Start, C. (1973) *Regulation in Metabolism*, Wiley, New York
226. Gruetter, R., Novotny, E. J., Boulware, S. D., Rothman, D. L., Mason, G. F., Shulman, G. I., Shulman, R. G., and Tamborlane, W. V. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1109–1112
227. Lapidot, A., and Gopher, A. (1994) *J. Biol. Chem.* **269**, 27198–27208
228. Efrat, S., Tal, M., and Lodish, H. F. (1994) *Trends Biochem. Sci.* **19**, 535–538

228a. Saltiel, A. R., and Kahn, C. R. (2001) *Nature (London)* **414**, 799 – 806

229. Paz, K., Voliovitch, H., Hadari, Y. R., Roberts, C. T., Jr., LeRoith, D., and Zick, Y. (1996) *J. Biol. Chem.* **271**, 6998 – 7003

230. Moxham, C. M., Tabrizchi, A., Davis, R. J., and Malbon, C. C. (1996) *J. Biol. Chem.* **271**, 30765 – 30773

231. Guilherme, A., Klarlund, J. K., Krystal, G., and Czech, M. P. (1996) *J. Biol. Chem.* **271**, 29533 – 29536

232. Tsakiridis, T., Taha, C., Grinstein, S., and Klip, A. (1996) *J. Biol. Chem.* **271**, 19664 – 19667

232a. Krook, A., Whitehead, J. P., Dobson, S. P., Griffiths, M. R., Ouwens, M., Baker, C., Hayward, A. C., Sen, S. K., Maassen, J. A., Siddle, K., Tavaré, J. M., and O'Rahilly, S. (1997) *J. Biol. Chem.* **272**, 30208 – 30214

232b. Ottensmeyer, F. P., Beniac, D. R., Luo, R. Z.-T., and Yip, C. C. (2000) *Biochemistry* **39**, 12103 – 12112

232c. Aguirre, V., Uchida, T., Yenush, L., Davis, R., and White, M. F. (2000) *J. Biol. Chem.* **275**, 9047 – 9054

232d. Alper, J. (2000) *Science* **289**, 37,39

232e. Withers, D. J., Gutierrez, J. S., Towery, H., Burks, D. J., Ren, J.-M., Previs, S., Zhang, Y., Bernal, D., Pons, S., Shulman, G. I., Bonner-Weir, S., and White, M. F. (1998) *Nature (London)* **391**, 900 – 904

232f. Czech, M. P., and Corvera, S. (1999) *J. Biol. Chem.* **274**, 1865 – 1868

232g. Lavan, B. E., Fantin, V. R., Chang, E. T., Lane, W. S., Keller, S. R., and Lienhard, G. E. (1997) *J. Biol. Chem.* **272**, 21403 – 21407

232h. Inoue, G., Cheatham, B., Emkey, R., and Kahn, C. R. (1998) *J. Biol. Chem.* **273**, 11548 – 11555

232i. Qiao, L.-y, Goldberg, J. L., Russell, J. C., and Sun, X. J. (1999) *J. Biol. Chem.* **274**, 10625 – 10632

232j. Pessin, J. E., Thurmond, D. C., Elmendorf, J. S., Coker, K. J., and Okada, S. (1999) *J. Biol. Chem.* **274**, 2593 – 2596

232k. Shepherd, P. R., and Kahn, B. B. (1999) *N. Engl. J. Med.* **341**, 248 – 257

232l. Rea, S., Martin, L. B., McIntosh, S., Macaulay, S. L., Ramsdale, T., Baldini, G., and James, D. E. (1998) *J. Biol. Chem.* **273**, 18784 – 18792

232m.Lee, W., Ryu, J., Souto, R. P., Pilch, P. F., and Jung, C. Y. (1999) *J. Biol. Chem.* **274**, 37755 – 37762

232n. Kanzaki, M., Watson, R. T., Artemyev, N. O., and Pessin, J. E. (2000) *J. Biol. Chem.* **275**, 7167 – 7175

232o. Czech, M. P. (2000) *Nature (London)* **407**, 147 – 148

232p. Baumann, C. A., Ribon, V., Kanzaki, M., Thurmond, D. C., Mora, S., Shigematsu, S., Bickel, P. E., Pessin, J. E., and Saltiel, A. R. (2000) *Nature (London)* **407**, 202 – 207

233. Ostlund, R. E., Jr., Seemayer, R., Gupta, S., Kimmel, R., Ostlund, E. L., and Sherman, W. R. (1996) *J. Biol. Chem.* **271**, 10073 – 10078

234. Ostlund, R. E., Jr., McGill, J. B., Herskowitz, I., Kipnis, D. M., Santiago, J. V., and Sherman, W. R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9988 – 9992

235. Galasko, G. T. F., Abe, S., Lilley, K., Zhang, C., and Larner, J. (1996) *J. Clin. Endocrinol. Metab.* **81**, 1051 – 1057

236. Romero, G., Luttrell, L., Rogol, A., Zeller, K., Hewlett, E., and Larner, J. (1988) *Science* **240**, 509 – 511

236a. Frick, W., Bauer, A., Bauer, J., Wied, S., and Müller, G. (1998) *Biochemistry* **37**, 13421 – 13436

237. Saltiel, A. R. (1994) *FASEB J.* **8**, 1034 – 1040

237a. Müller, G., Grey, S., Jung, C., and Bandlow, W. (2000) *Biochemistry* **39**, 1475 – 1488

237b. Thorens, B., Guillam, M.-T., Beermann, F., Burcelin, R., and Jaquet, M. (2000) *J. Biol. Chem.* **275**, 23751 – 23758

238. Van Schaftingen, E., Detheux, M., and Veiga da Cunha, M. (1994) *FASEB J.* **8**, 414 – 419

239. Terauchi, Y., Sakura, H., Yasuda, K., Iwamoto, K., Takahashi, N., Ito, K., Kasai, H., Suzuki, H., Ueda, O., Kamada, N., Jishage, K., Komeda, K., Noda, M., Kanazawa, Y., Taniguchi, S., Miwa, I., Akanuma, Y., Kodama, T., Yazaki, Y., and Kadowaki, T. (1995) *J. Biol. Chem.* **270**, 30253 – 30256

240. Agius, L., Peak, M., Newgard, C. B., Gomez-Foix, A. M., and Guinovart, J. J. (1996) *J. Biol. Chem.* **271**, 30479 – 30486

241. Webb, D.-L., Islam, M. S., Efanov, A. M., Brown, G., Köhler, M., Larsson, O., and Berggren, P.-O. (1996) *J. Biol. Chem.* **271**, 19074 – 19079

242. Deeney, J. T., Cunningham, B. A., Chheda, S., Bokvist, K., Juntti-Berggren, L., Lam, K., Korchak, H. M., Corkey, B. E., and Berggren, P.-O. (1996) *J. Biol. Chem.* **271**, 18154 – 18160

243. Bergsten, P., Grapengiesser, E., Gylfe, E., Tangholm, A., and Hellman, B. (1994) *J. Biol. Chem.* **269**, 8749 – 8753

244. Buggy, J. J., Livingston, J. N., Rabin, D. U., and Yoo-Warren, H. (1995) *J. Biol. Chem.* **270**, 7474 – 7478

245. Walajtys-Rhode, E., Zapatero, J., Moehren, G., and Hoek, J. B. (1992) *J. Biol. Chem.* **267**, 370 – 379

246. de Duve, C. (1994) *FASEB J.* **8**, 979 –981

247. Rubanyi, G. M., and Botelho, L. H. P. (1991) *FASEB J.* **5**, 2713 – 2720

248. Heimberg, H., De Vos, A., Moens, K., Quartier, E., Bouwens, L., Pipeleers, D., Van Schaftingen, E., Madsen, O., and Schuit, F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7036 – 7041

249. Granner, D., and Pilkis, S. (1990) *J. Biol. Chem.* **265**, 10173 – 10176

250. Unson, C. G., Wu, C.-R., and Merrifield, R. B. (1994) *Biochemistry* **33**, 6884 – 6887

251. Lei, K.-J., Pan, C.-J., Liu, J.-L., Shelly, L. L., and Chou, J. Y. (1995) *J. Biol. Chem.* **270**, 11882 – 11886

252. Berteloo, A., St-Denis, J.-F., and van de Werve, G. (1995) *J. Biol. Chem.* **270**, 21098 – 21102

253. Chen, Y.-T., and Burchell, A. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 935 – 965, McGraw-Hill, New York

254. Pilkis, S. J., El-Maghrabi, M. R., and Claus, T. C. (1988) *Ann. Rev. Biochem.* **57**, 755 – 783

255. Tornheim, K. (1988) *J. Biol. Chem.* **263**, 2619 – 2624

256. Scheffler, J. E., and Fromm, H. J. (1986) *Biochemistry* **25**, 6659 – 6665

257. Kraus-Friedmann, N. (1986) *Trends Biochem. Sci.* **11**, 276 – 279

258. Blackard, W. G., and Clore, J. N. (1988) *J. Biol. Chem.* **263**, 16725 – 16730

259. Liu, J., Park, E. A., Gurney, A. L., Roesler, W. J., and Hanson, R. W. (1991) *J. Biol. Chem.* **266**, 19095 – 19102

260. Beale, E. G., Chrapkiewicz, N. B., Scoble, H. A., Metz, R. J., Quick, D. P., Noble, R. L., Donelson, J. E., Biemann, K., and Granner, D. K. (1985) *J. Biol. Chem.* **260**, 10748 – 10760

261. Scott, D. K., Mitchell, J. A., and Granner, D. K. (1996) *J. Biol. Chem.* **271**, 31909 – 31914

261a. Yeagley, D., Moll, J., Vinson, C. A., and Quinn, P. G. (2000) *J. Biol. Chem.* **275**, 17814 – 17820

262. Höppner, W., Beckert, L., Buck, F., and Seitz, H.-J. (1991) *J. Biol. Chem.* **266**, 17257 – 17260

263. Punekar, N. S., and Lardy, H. A. (1987) *J. Biol. Chem.* **262**, 6714 – 6719

264. Kuwajima, M., Golden, S., Katz, J., Unger, R. H., Foster, D. W., and McGarry, J. D. (1986) *J. Biol. Chem.* **261**, 2632 – 2637

265. Cline, G. W., and Shulman, G. I. (1995) *J. Biol. Chem.* **270**, 28062 – 28067

266. Rognstad, R., and Katz, J. (1972) *J. Biol. Chem.* **247**, 6047 – 6054

267. Clark, M. G., Bloxham, D. P., Holland, P. C., and Lardy, H. A. (1973) *Biochem. J.* **134**, 589 – 597

268. Exton, J. H., Friedmann, N., Wong, E. H.-A., Brineaux, J. P., Corbin, J. D., and Park, C. R. (1972) *J. Biol. Chem.* **247**, 3579 – 3588

269. Katz, J., Wals, P., and Lee, W.-N. P. (1993) *J. Biol. Chem.* **268**, 25509 – 25521

270. Des Rosiers, C., Di Donato, L., Comte, B., Laplante, A., Marcoux, C., David, F., Fernandez, C. A., and Brunengraber, H. (1995) *J. Biol. Chem.* **270**, 10027 – 10036

270a. Previs, S. F., Hallowell, P. T., Neimanis, K. D., David, F., and Brunengraber, H. (1998) *J. Biol. Chem.* **273**, 16853 – 16859

271. Previs, S. F., Fernandez, C. A., Yang, D., Soloviev, M. V., David, F., and Brunengraber, H. (1995) *J. Biol. Chem.* **270**, 19806 – 19815

271a. Zhang, B. L., Yunianta, and Martin, M. L. (1995) *J. Biol. Chem.* **270**, 16023 – 16029

272. Neese, R. A., Schwarz, J.-M., Faix, D., Turner, S., Letscher, A., Vu, D., and Hellerstein, M. K. (1995) *J. Biol. Chem.* **270**, 14452 – 14463

273. Martin, G., Chauvin, M.-F., Dugelay, S., and Baverel, G. (1994) *J. Biol. Chem.* **269**, 26034 – 26039

274. Fernandez, C. A., and Des Rosiers, C. (1995) *J. Biol. Chem.* **270**, 10037 – 10042

275. Ni, T.-C., and Savageau, M. A. (1996) *J. Biol. Chem.* **271**, 7927 – 7941

276. Shiraishi, F., and Savageau, M. A. (1993) *J. Biol. Chem.* **268**, 16917 – 16928

277. Wang, H.-C., Ciskanik, L., Dunaway-Mariano, D., van der Saal, W., and Villafranca, J. J. (1988) *Biochemistry* **27**, 625 – 633

277a. Moon, Y. S., Latasa, M.-J., Kim, K.-H., Wang, D., and Sul, H. S. (2000) *J. Biol. Chem.* **275**, 10121 – 10127

277b. Flier, J. S., and Hollenberg, A. N. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14191 – 14192

277c. Barroso, I., Gurnell, M., Crowley, V. E. F., Agostini, M., Schwabe, J. W., Soos, M. A., Li Maslen, G., Williams, T. D. M., Lewis, H., Schafer, A. J., Chatterjee, V. K. K., and O'Rahilly, S. (1999) *Nature (London)* **402**, 880 – 883

277d. Steppan, C. M., Bailey, S. T., Bhat, S., Brown, E. J., Banerjee, R. R., Wright, C. M., Patel, H. R., Ahima, R. S., and Lazar, M. A. (2001) *Nature (London)* **409**, 307 – 312

277e. Müller, G., Ertl, J., Gerl, M., and Preibisch, G. (1997) *J. Biol. Chem.* **272**, 10585 – 10593

277f. Schwartz, M. W. (2000) *Science* **289**, 2066 – 2067

278. Newgard, C. B., and McGarry, J. D. (1995) *Ann. Rev. Biochem.* **64**, 689 – 719

279. Corkey, B. E., Glennon, M. C., Chen, K. S., Deeney, J. T., Matschinsky, F. M., and Prentki, M. (1989) *J. Biol. Chem.* **264**, 21608 – 21612

280. Kudo, N., Barr, A. J., Barr, R. L., Desai, S., and Lopaschuk, G. D. (1995) *J. Biol. Chem.* **270**, 17513 – 17520

281. Munir, E., Yoon, J. J., Tokimatsu, T., Hattori, T., and Shimada, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11126 – 11130

Study Questions

1. Write out a complete step-by-step mechanism for the reactions by which citrate can be synthesized from pyruvate and then exported from mitochondria for use in the biosynthesis of fatty acids. Include a chemically reasonable mechanism for the action of ATP–citrate lyase, which catalyzes the following reaction:

   ATP + citrate + CoA-SH → Acetyl-CoA + oxaloacetate + ADP + P_i

   Show how this reaction can be incorporated into an ATP-driven cyclic pathway for generating NADPH from NADH.

2. Show which parts (if any) of the citric acid cycle are utilized in each of the following reactions and what, if any, additional enzymes are needed in each case.
   a. Oxidation of acetyl-CoA to CO_2
   b. Catabolism of glutamate to CO_2
   c. Biosynthesis of glutamate from pyruvate
   d. Formation of propionate from pyruvate

3. Here is a possible metabolic reaction for a fungus.

   L-Leucine + 2-oxoglutarate^{2-} + 2 $^1/_2$ O_2 → L-glutamate$^-$ + citrate^{3-} + H_2O + $2H^+$
   $\Delta G°'$ (pH 7) = –1026 kJ/mol

   Suggest a metabolic pathway for this reaction. Is it thermodynamically feasible?

4. It has been suggested that in *Escherichia coli* pyruvate may act as the regenerating substrate for a catalytic cycle by which glyoxylate, OHC–COO$^-$, is oxidized to CO_2. Key enzymes in this cycle are thought to be a 2-oxo-4-hydroxyglutarate aldolase and 2-oxoglutarate dehydrogenase. Propose a detailed pathway for this cycle.

5. Some bacteria use a "dicarboxylic acid cycle" to oxidize glyoxylate OHC–COO$^-$ to CO_2. The regenerating substrate for this cycle is acetyl-CoA. It is synthesized from glyoxylate by a complex pathway that begins with conversion of two molecules of glyoxylate to tartronic semialdehyde: $^-$OOC–CHOH–CHO. The latter is then dehydrogenated to D-glycerate.

   Write out a detailed scheme for the dicarboxylate cycle. Also indicate how glucose and other cell constituents can be formed from intermediates created in this biosynthetic pathway.

6. Some bacteria that lack the usual aldolase produce ethanol and lactic acid in a 1:1 molar ratio via the "heterolactic fermentation." Glucose is converted to ribulose 5-phosphate via the pentose phosphate pathway enzymes. A thiamin diphosphate-dependent "phosphoketolase" cleaves xylulose 5-phosphate in the presence of inorganic phosphate to acetyl phosphate and glyceraldehyde 3-phosphate.

   Propose a mechanism for the phosphoketolase reaction and write a balanced set of equations for the fermentation.

7. Bacteria of the genera *Aerobacter* and *Serratia* ferment glucose according to the following equation:

   1 $^1/_2$ Glucose → 2,3-butanediol + 3 CO_2 + H_2 + ethanol

   Write out a detailed pathway for the reactions. Use the pyruvate formate–lyase reaction. What yield of ATP do you expect per molecule of glucose fermented?

8. Some Clostridia ferment glucose as follows:

   2 Glucose + 2 H_2O → butyrate$^-$ + 2 acetate$^-$ + 4 CO_2 + 6 H_2 + 3 H^+
   $\Delta G°'$ (pH 7) = –479 kJ/mol of butyrate formed

   Write out detailed pathways. How much ATP do you think can be formed per glucose molecule fermented?

9. Propionic acid bacteria use the following fermentation:

   1 $^1/_2$ Glucose → 2 propionate + acetate$^-$ + 3 H^+ + CO_2 + H_2O
   $\Delta G°'$ = –464 kJ/mol of acetate formed

   Write out a detailed pathway for the reactions. How much ATP can be formed per molecule of glucose?

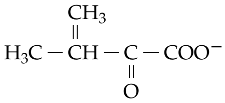
10. Consider the following reaction which can occur in the animal body:

    2 Lactate$^-$ ($C_3H_5O_3^-$) + CoASH + 2 NAD^+ → butyryl-CoA (C_4H_7O–S–CoA) + 2 CO_2 + 2 NADH + H_2O
    $\Delta G°'$ (pH 7) ~ 70 kJ/mol of butyryl-CoA formed

    Outline the sequence of reactions involved in this transformation. Do you think that any ATP will either be used or generated in the reaction? Explain.

11. Leucine, an essential dietary constituent for human beings, is synthesized in many bacteria and plants using pyruvate as a starting material. Outline this pathway of metabolism and illustrate the chemical reaction mechanisms involved in each step.

12. Write a step-by-step sequence for all of the chemical reactions involved in the biosynthesis of L-leucine from 2-oxoisovalerate:



$$H_3C-\overset{\overset{\displaystyle CH_3}{\|}}{CH}-\underset{\underset{\displaystyle O}{\|}}{C}-COO^-$$

Notice that this compound contains one carbon atom less than leucine. Start by condensing 2-oxoisovalerate with acetyl-CoA in a reaction similar to that of citrate synthase. Use structural formulas. Show all intermediate structures and indicate what coenzymes are needed. Use curved arrows to indicate the flow of electrons in each step.

13. Some fungi synthesize lysine from 2-oxoglutarate by elongating the chain using a carbon atom derived from acetyl-CoA to form the 6-carbon 2-oxoadipate. The latter is converted by an ATP-dependent reduction to the ε-aldehyde. Write out reasonable mechanisms for the conversion of 2-oxoglutarate to the aldehyde. The latter is converted on to L-lysine by a non–PLP-dependent transamination via saccharopine (Chapter 24).

14. Outline the pathway for biosynthesis of L-leucine from glucose and NH_4^+ in autotrophic organisms. In addition, outline the pathways for degradation of leucine to CO_2, water, and NH_4^+ in the human body. For this overall pathway or "metabolic loop," mark the locations (one or more) at which each of the following processes occurs.
    a. Synthesis of a thioester by dehydrogenation
    b. Substrate-level phosphorylation
    c. Thiamin-dependent α condensation
    d. Oxoacid chain-elongation process
    e. Transamination
    f. Oxidative decarboxylation of an α-oxoacid
    g. Partial β oxidation
    h. Thiolytic cleavage
    i. Claisen condensation
    j. Biotin-dependent carboxylation

15. A photosynthesizing plant is exposed to $^{14}CO_2$. On which carbon atoms will the label first appear in glucose?

16. The Calvin–Benson cycle and the pentose phosphate pathway (Eq. 17-12) have many features in common but run in opposite directions. Since the synthesis of glucose from CO_2 requires energy, the energy expenditure for the two processes will obviously differ. Describe the points in each pathway where a Gibbs energy difference is used to drive the reaction in the desired direction.

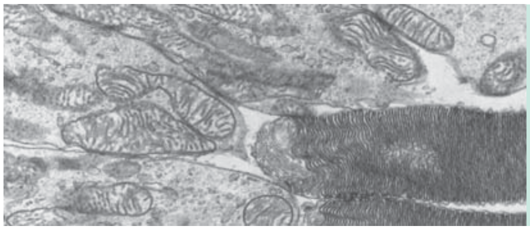
17. Draw the structure of ribulose 1(^{14}C), 5-bisphosphate. Enter an asterisk (*) next to carbon 1 to show that this position is ^{14}C-labeled.

    Draw the structures of the products of the ribulose bisphosphate carboxylase reaction, indicating the radioactive carbon position with an asterisk.

18. A wood-rotting fungus is able to convert glucose to oxalate approximately according to the following equation:

    $$\text{Glucose} + 5\ O_2 \rightarrow 2\ \text{Oxalic acid}\ (C_2O_4H_2) + 2\ CO_2 + 4\ H_2O$$

    Propose a mechanism. See Munir *et al.*[281] for details.



Electrons flowing through the electron transport chains in the membranes of the mitochondria, (at the left) in this thin section through the retina of a kangaroo rat (*Dipodomys ordi*) generates ATP. The ATP provides power for the synthesis and functioning of the stacked photoreceptor disks seen at the right. The outer segment of each rod cell (See Fig. 23-40), which may be 15–20 mm in length, consists of these disks, whose membranes contain the photosensitive protein pigment rhodopsin. Absorption of light initiates an electrical excitation which is sent to the brain. Micrograph from Porter and Bonneville (1973) Fine Structure of Cells and Tissues, Lea and Febiger, Philadelphia, Pennsylvania. Courtesy of Mary Bonneville.

Contents

1013 .... **A. The Architecture of the Mitochondrion**
1014 .... 1. The Mitochondrial Membranes and Compartments
1015 .... 2. The Chemical Activities of Mitochondria
1016 .... 3. The Mitochondrial Genome
1018 .... 4. Growth and Development
1019 .... **B. Electron Transport Chains**
1019 .... 1. The Composition of the Mitochondrial Electron Transport System
1019 .... 2. The Sequence of Electron Carriers
1020 .... *Submitochondrial particles and complexes*
1023 .... *Mitochondrial electron transport in plants and fungi*
1025 .... *Electron transport chains of bacteria*
1026 .... 3. Structures and Functions of the Individual Complexes I - IV and Related Bacterial Assemblies
1026 .... *Complex I, NADH-ubiquinone oxidoreductase*
1027 .... *Complex II, succinate-ubiquinone oxidoreductase*
1027 .... *Complex III (ubiquinol-cytochrome c oxidoreductase or cytochrome bc_1 complex)*
1028 .... *Complex IV. Cytochrome c oxidase (Ubiquinol-cytochrome c oxidoreductase)*
1032 .... **C. Oxidative Phosphorylation**
1033 .... 1. The Stoichiometry (P/O Ratio) and Sites of Oxidative Phosphorylation
1033 .... *Respiratory control and uncoupling*
1033 .... *"States" of mitochondria and spectrophotometric observation*
1034 .... 2. Thermodynamics and Reverse Electron Flow
1036 .... 3. The Mechanism of Oxidative Phosphorylation
1036 .... *Search for chemical intermediates*
1037 .... *Peter Mitchell's chemiosmotic theory*
1039 .... *Observed values of E_m and pH*
1039 .... *Proton pumps driven by electron transport*
1041 .... 4. ATP Synthase
1041 .... *The synthase structure*
1043 .... *How is ATP made?*
1044 .... *Paul Boyer's binding change mechanism*
1044 .... *Rotational catalysis*
1045 .... 5. ATP-driven Proton Pumps
1046 .... 6. Uncouplers and Energy-linked Processes in Mitochondria
1047 .... **D. Transport and Exchange across Mitochondrial Membranes**
1050 .... **E. Energy from Inorganic Reactions**
1051 .... 1. Reduced Inorganic Compounds as Substrates for Respiration
1051 .... *The hydrogen-oxidizing bacteria*
1051 .... *Nitrifying bacteria*
1052 .... *The sulfur-oxidizing bacteria*
1054 .... 2. Anaerobic Respiration
1054 .... *Nitrate as an electron acceptor*
1054 .... *Reduction of nitrite: denitrification*
1056 .... *Sulfate-reducing and sulfur-reducing bacteria*
1057 .... *Methane bacteria*
1057 .... **F. Oxygenases and Hydroxylases**
1057 .... 1. Dioxygenases
1059 .... 2. Monooxygenases
1059 .... *Flavin-containing monooxygenases*
1061 .... *Reduced pteridines as cosubstrates*
1062 .... *Hydroxylation-induced migration*
1063 .... *2-oxoglutarate as a decarboxylating cosubstrate*
1064 .... *Copper-containing hydroxylases*
1065 .... *Hydroxylation with cytochrome P450*
1068 .... *Other iron-containing oxygenases*
1069 .... *Nitric oxide and NO synthases*
1072 .... **G. Biological Effects of Reduced Oxygen Compounds**
1072 .... 1. The Respiratory Burst of Neutrophils
1074 .... 2. Oxidative Damage to Tissues
1074 .... *Antioxidant systems*
1075 .... *Transcriptional regulation of antioxidant proteins*

1076 .... **References**
1085 .... **Study Questions**

Boxes

1022 .... Box 18-A Historical Notes on Respiration
1024 .... Box 18-B Defects of Mitochondrial DNA
1048 .... Box 18-C Using Metabolism to Generate Heat: Thermogenic Tissues
1066 .... Box 18-D Vitamin C: Ascorbic Acid
1070 .... Box 18-E The Toxicity of Acetaminophen

Tables

1016 .... Table 18-1 Catalog of Mitochondrial Genes
1018 .... Table 18-2 Alterations in the Genetic Code in the DNA of Animal Mitochondria
1019 .... Table 18-3 Ratios of Components in the Electron Transport Chain of Mitochondria
1021 .... Table 18-4 Some Well-Known Respiratory Inhibitors
1025 .... Table 18-5 Some Artificial Electron Acceptors
1034 .... Table 18-6 Wavelengths of Light Used to Measure States of Oxidation of Carriers in the Electron Transport Chain of Mitochondria
1035 .... Table 18-7 Electrode Potentials of Mitochondrial Electron Carriers and Gibbs Energy Changes Associated with Passage of Electrons
1050 .... Table 18-8 Some Mitochondrial Membrane Transporters

Electron Transport, Oxidative Phosphorylation, and Hydroxylation

18

In this chapter we will look at the processes by which reduced carriers such as NADH and $FADH_2$ are oxidized within cells. Most familiar to us, because it is used in the human body, is **aerobic respiration**. Hydrogen atoms of NADH, $FADH_2$, and other reduced carriers appear to be transferred through a chain of additional carriers of increasingly positive reduction potential and are finally combined with O_2 to form H_2O. In fact, the hydrogen nuclei move freely as protons (or sometimes as H^- ions); it is the *electrons* that are deliberately transferred. For this reason, the chain of carriers is often called the **electron transport chain**. It is also referred to as the **respiratory chain**.

Because far more energy is available to cells from oxidation of NADH and $FADH_2$ than can be obtained by fermentation, the chemistry of the electron transport chain and of the associated reactions of ATP synthesis assumes great importance. A central question becomes "How is ATP generated by flow of electrons through this series of carriers"? Not only is most of the ATP formed in aerobic and in some anaerobic organisms made by this process of **oxidative phosphorylation**, but the solar energy captured during photosynthesis is used to form ATP in a similar manner. The mechanism of ATP generation may also be intimately tied to the function of membranes in the transport of ions. In a converse manner, the mechanism of oxidative phosphorylation may be related to the utilization of ATP in providing energy for the contraction of muscles.

In some organisms, especially bacteria, energy may be obtained through oxidation of H_2, H_2S, CO, or Fe^{2+} rather than of the hydrogen atoms removed from organic substrates. Furthermore, some bacteria use **anaerobic respiration** in which NO_3^-, SO_4^{2-}, or CO_2 act as oxidants either of reduced carriers or of reduced inorganic substances. In the present chapter, we will consider these energy-yielding processes as well as the chemistry of reactions of oxygen that lead to incorporation of atoms from O_2 into organic compounds.

The oxidative processes of cells have been hard to study, largely because the enzymes responsible are located in or on cell membranes. In bacteria the sites of electron transport and oxidative phosphorylation are on the inside of the plasma membrane or on membranes of mesosomes. In eukaryotes they are found in the inner membranes of the mitochondria and, to a lesser extent, in the endoplasmic reticulum. For this reason we should probably start with a closer look at mitochondria, the "power plants of the cell."

A. The Architecture of the Mitochondrion

Mitochondria are present in all eukaryotic cells that use oxygen in respiration, but the number per cell and the form and size vary.[1–4] Certain tiny trypanosomes have just *one* mitochondrion but some oocytes have as many as 3 x 10^5. Mammalian cells typically contain several hundred mitochondria and liver cells[5] more than 1000. Mammalian sperm cells may contain 50–75 mitochondria,[6] but in some organisms only one very large helical mitochondrion, formed by the fusion of many individual mitochondria, wraps around the base of the tail. Typical mitochondria appear to be about the size of cells of *E. coli*. However, study of ultrathin serial sections of a single yeast cell by electron microscopy has shown that, under some growth conditions, all of the mitochondria are interconnected.[7]

In every case a mitochondrion is enclosed by two concentric membranes, an *outer* and an *inner* membrane, each ~5–7 nm thick (Figs. 18-1, 18-2). The inner membrane is folded to form the **cristae**. The number of cristae, the form of the cristae, and the relative amount of the internal **matrix** space are variable. In liver there is little inner membrane and a large matrix space, while in heart mitochondria there are more folds and a higher rate of oxidative phosphorylation. The enzymes catalyzing the tricarboxylic acid cycle are also unusually active in heart mitochondria. A typical heart mitochondrion has a volume of 0.55 μm^3; for every cubic micrometer of mitochondrial volume there are 89 μm^2 of inner mitochondrial membranes.[9]

Mitchondria can swell and contract, and forms other than that usually seen in osmium-fixed electron micrographs have been described. In some mitochondria the cristae are swollen, the matrix volume is much reduced, and the **intermembrane space** between the membranes is increased. Rapidly respiring mitochondria fixed for electron microscopy exhibit forms that have been referred to as "energized" and "energized-twisted."[10] The micrograph (Fig. 18-1) and drawing (Fig. 18-2) both show a significant amount of intermembrane space. However, electron micrographs of mitochondria from rapidly frozen aerobic tissues show almost none.[11] Recent studies by electron microscopic tomography show cristae with complex tubular structures. The accepted simple picture of mitochondrial structure (Fig. 18-1) is undergoing revision.[12–12b] The isolated mitochondria that biochemists have studied may be fragments of an interlinked **mitochondrial reticulum** that weaves its way through the cell.[12b] However, this reticulum may not be static but may break and reform. The accepted view that the mitochondrial matrix space is continuous with the internal space in the cristae is also the subject of doubts. Perhaps they are two different compartments.[12a]

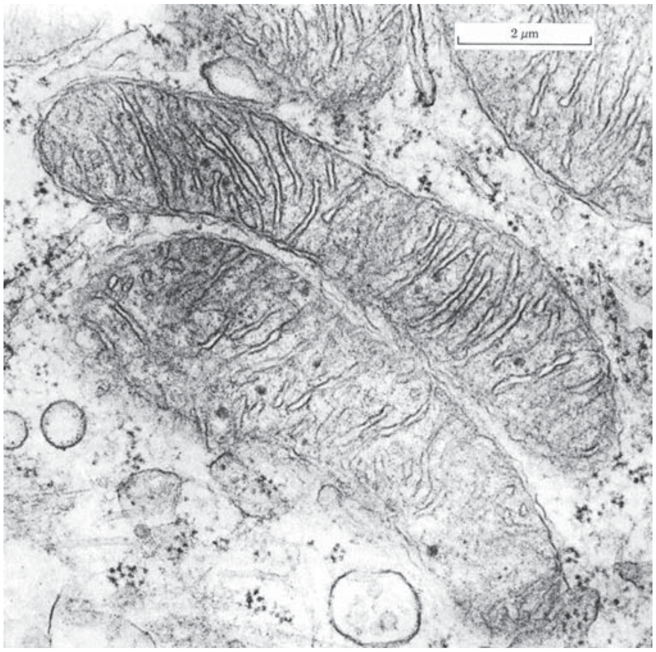


Figure 18-1 Thin section of mitochondria of a cultured kidney cell from a chicken embryo. The small, dark, dense granules within the mitochondria are probably calcium phosphate. Courtesy of Judie Walton.

1. The Mitochondrial Membranes and Compartments

The outer membranes of mitochondria can be removed from the inner membranes by osmotic rupture.[13] Analyses on separated membrane fractions show that *the outer membrane* is less dense (density ~1.1 g / cm^3) than the inner (density ~1.2 g / cm^3). It is highly permeable to most substances of molecular mass 10 kDa or less because of the presence of pores of ~2 nm diameter. These are formed by **mitochondrial porins**,[14–17] which are similar to the outer membrane porins of gram-negative bacteria (Fig. 8-20). The ratio of phospholipid to protein (~0.82 on a weight basis) is much higher than in the inner membrane. Extraction of the phospholipids by acetone destroys the membrane. Of the lipids present, there is a low content of cardiolipin, a high content of phosphatidylinositol and cholesterol, and no ubiquinone.

The inner membrane is impermeable to many substances. Neutral molecules of <150 Da can penetrate the membranes, but the permeability for all other materials including small ions such as H^+, K^+, Na^+, and Cl^- is tightly controlled. The ratio of phospholipid to protein in the inner membrane is ~0.27, and cardiolipin makes up ~20% of the phospholipid present. Cholesterol is absent. Ubiquinone and other components of the respiratory chain are all found in the inner membrane. Proteins account for 75% of the mass of the membrane.

Another characteristic of the inner mitochondrial membrane is the presence of projections on the inside surface, which faces the mitochondrial matrix. See Fig. 18-14. These spherical 85-kDa particles, discovered by Fernandez Moran in 1962 and attached to the membrane through a "stalk", display ATP-hydrolyzing (ATPase) activity. The latter was a clue that the knobs might participate in the *synthesis* of ATP during oxidative phosphorylation. In fact, they are now recognized as a complex of proteins called **coupling factor 1** (F_1) or **ATP synthase**.

In addition to bacterialike mitochondrial ribosomes and small circular molecules of DNA, mitochondria may contain variable numbers of dense granules of calcium phosphate, either $Ca_3(PO_4)_2$ or hydroxylapatite (Fig. 8-34),[4,18] as well as of phospholipoprotein.[4]

Quantitatively the major constituents of the matrix space are a large number of proteins. These account for about 56% by weight of the matrix material and exist in a state closer to that in a protein crystal than in a true solution.[19–20a] Mitochondrial membranes also contain proteins, both tightly bound relatively nonpolar intrinsic proteins and extrinsic proteins bound to the membrane surfaces. The other mitochondrial compartment, the intermembrane space, may normally be very small but it is still "home" for a few enzymes.

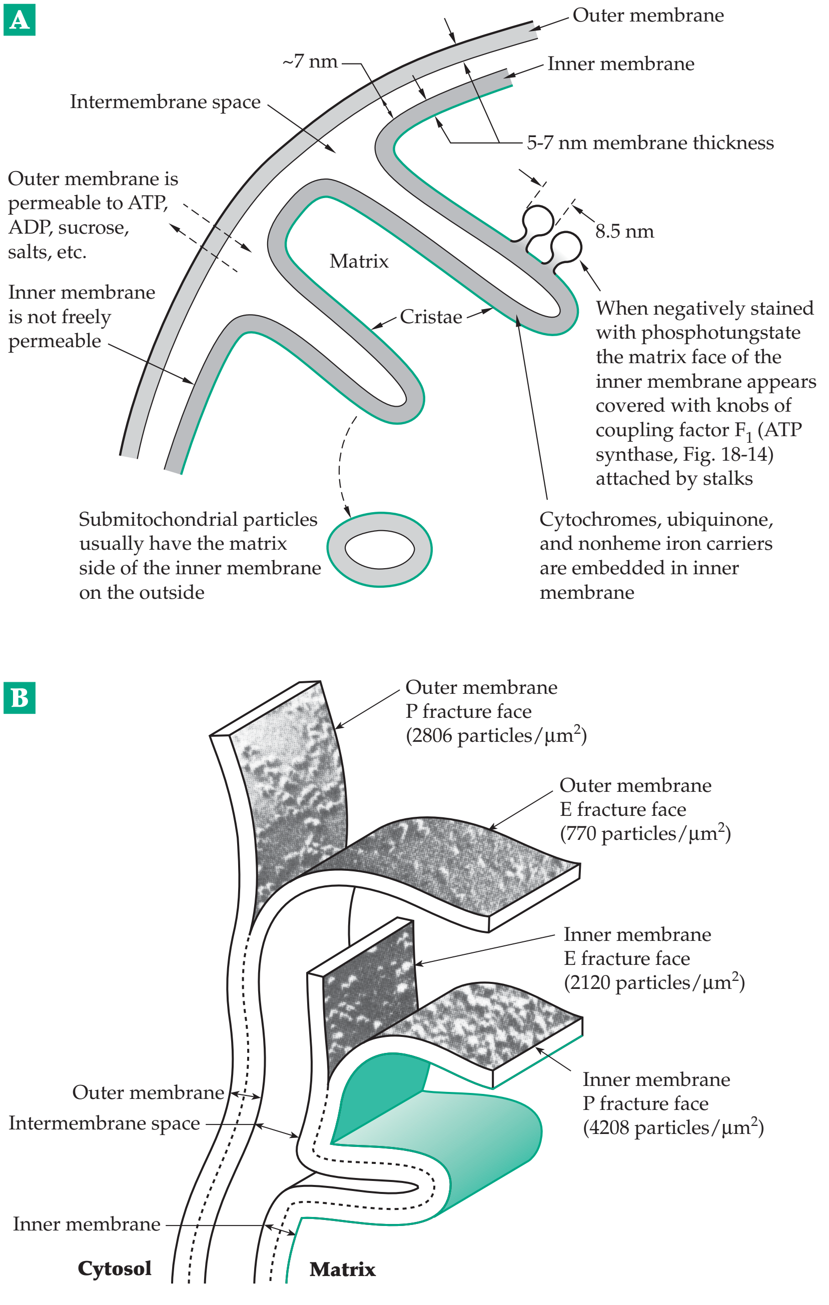

A
Outer membrane
~7 nm
Inner membrane
Intermembrane space
5-7 nm membrane thickness
Outer membrane is permeable to ATP, ADP, sucrose, salts, etc.
8.5 nm
Matrix
Inner membrane is not freely permeable
Cristae
When negatively stained with phosphotungstate the matrix face of the inner membrane appears covered with knobs of coupling factor F_1 (ATP synthase, Fig. 18-14) attached by stalks
Submitochondrial particles usually have the matrix side of the inner membrane on the outside
Cytochromes, ubiquinone, and nonheme iron carriers are embedded in inner membrane
B
Outer membrane P fracture face (2806 particles/µm^2)
Outer membrane E fracture face (770 particles/µm^2)
Inner membrane E fracture face (2120 particles/µm^2)
Inner membrane P fracture face (4208 particles/µm^2)
Outer membrane
Intermembrane space
Inner membrane
Cytosol
Matrix

Figure 18-2 (A) Schematic diagram of mitochondrial structure. (B) Model showing organization of particles in mitochondrial membranes revealed by freeze-fracture electron microscopy. The characteristic structural features seen in the four half-membrane faces (EF and PF) that arise as a result of fracturing of the outer and inner membranes are shown. The four smooth membrane surfaces (ES and PS) are revealed by etching. From Packer.[8]

2. The Chemical Activities of Mitochondria

Mention of mitochondria usually brings to the mind of the biochemist the **citric acid cycle,** the **β oxidation pathway** of fatty acid metabolism, and **oxidative phosphorylation**. In addition to these major processes, many other chemical events also occur. Mitochondria concentrate Ca^{2+} ions and control the entrance and exit of Na^+, K^+, dicarboxylates, amino acids, ADP, P_i and ATP, and many other substances.[16] Thus, they exert regulatory functions both on catabolic and biosynthetic sequences. The glycine decarboxylase system (Fig. 15-20) is found in the mitochondrial matrix and is especially active in plant mitochondria (Fig. 23-37). Several cytochrome P450-dependent hydroxylation reactions, important to the biosynthesis and catabolism of steroid hormones and to the metabolism of vitamin D, take place within mitochondria. Mitochondria make only a few of their own proteins but take in several hundred other proteins from the cytoplasm as they grow and multiply.

Where within the mitochondria are specific enzymes localized? One approach to this question is to see how easily the enzymes can be dissociated from mitochondria. Some enzymes come out readily under hypotonic conditions. Some are released only upon sonic oscillation, suggesting that they are inside the matrix space. Others, including the cytochromes and the flavoproteins that act upon succinate and NADH, are so firmly embedded in the inner mitochondrial membranes that they can be dissociated only through the use of non-denaturing detergents.

Because the enzymes of the citric acid cycle[20a] (with the exception of succinate dehydrogenase) and β oxidation are present in the matrix, the reduced electron carriers must approach the inner membrane from the matrix side (the M side). Thus, the embedded enzymes designed to oxidize NADH, succinate, and other reduced substrates must be accessible from the matrix side. However, *sn*-glycerol 3-phosphate dehydrogenase, a flavoprotein, is accessible from the “outside” of the cytoplasmic (C side) of the inner membrane.[21] Fluorescent antibodies to cytochrome *c* bind only to the cytoplasmic (intermembrane) side of the inner membrane, but antibodies to cytochrome oxidase label both sides, which suggested that this protein complex spans the membrane.[22] However, oxidation of cytochrome *c* by cytochrome oxidase occurs only on the cytoplasmic surface.[22] Antibodies to the ATP synthase that makes up the knobs bind strictly to the matrix side.

The outer mitochondrial membrane contains monoamine oxidase, cytochrome b_5, fatty acyl-CoA synthase, and enzymes of cardiolipin synthesis[22a] as well as other proteins. Cardiolipin (diphosphatidylglycerol; Fig. 21-4) is found only in the inner mitochondrial membrane and in bacteria. It is functionally important for several mitochondrial enzymes including cytochrome oxidase and cytochome bc_1.[22a–c] It is also essential to photosynthetic membranes for which an exact role in an interaction between the lipid membrane and the associated protein has been revealed by crystallography.[22d] In other respects the composition of the inner mitochondrial membrane resembles that of the membranes of the endoplasmic reticulum. Isoenzyme III of adenylate kinase, a key enzyme involved in equilibrating ATP and AMP with ADP (Eq. 6-65), is one of the enzymes present in the intermembrane space. A number of other kinases, as well as sulfite oxidase, are also present between the membranes.[4]

As mentioned in Box 6-D, mitochondria sometimes take up calcium ions. The normal total concentration of Ca^{2+} is ~1 mM and that of free Ca^{2+} may be only ~0.1 μM.[22e,f] However, under some circumstances mitochondria accumulate large amounts of calcium, perhaps acting as a Ca^{2+} buffer.[22g,f] The so called ryanodine receptors (Fig. 19-21), prominent in the endoplasmic reticulum, have also been found in heart mitochondria, suggesting a function in control of calcium oscillations.[22i,j] On the other hand, accumulation of calcium by mitochondria may be pathological and the activation of Ca^{2+}-dependent proteases may be an initial step in apoptosis.[22h,22k]

TABLE 18-1
Catalog of Mitochondrial Genes[a]

Name and symbol	*Homo sapiens*	*Reclinomonas americana*	*Saccharomyces cerevisiae*	*Arabidopsis thaliana*
Ribosomal RNA				
s rRNA (small, 12s)	1	1	1	1
l rRNA (large, 16s)	1	1	1	1
5 S RNA		1		1
Transfer RNAs	22[b]	26	24	22
NADH dehydrogenase				
Subunits ND1–6, ND4L	7	12	0	9
Cytochrome *b*	1	1	1	1
Cytochrome oxidase				
Subunits I, II, III	3	3	3	3
ATP synthase				
Subunits 6, 8, others	2	5	3	4
Total protein coding genes	13	62	8	27
Total genes	37	92	35	53
Size of DNA (kbp)	16.596	69	75	367

[a] Data from Palmer, J. D. (1997) *Nature (London)* **387**, 454–455.[1]
[b] One for each amino acid of the genetic code but two each for serine and leucine.

3. The Mitochondrial Genome

Each mitochondrion contains several molecules of DNA (**mtDNA**), usually in a closed, circular form, as well as the ribosomes, tRNA molecules, and enzymes needed for protein synthesis.[1,23–26] With rare exceptions almost all of the mitochondrial DNA in a human cell is inherited from the mother.[6,26a] The size of the DNA circles varies from 16–19 kb in animals[27] to over 200 kb in many higher plants. Complete sequences of many mitochondrial DNAs are known.[28,28a] Among these are the 16,569 bp human mtDNA,[29] the 16,338 bp bovine mtDNA, the 16,896 bp mtDNA of the wallaroo *Macropus robustus*,[30] and the 17,533 bp mtDNA of the amphibian *Xenopus laevis*.[31,32] The sea urchin *Paracentotus lividus* has a smaller 15,697 bp genome. However, the order of the genes in this and other invertebrate mtDNA is different from that in mammalian mitochondria.[33] Protozoal mtDNAs vary in size from ~5900 bp for the parasitic malaria organism *Plasmodium falciparum*[34,35] to 41,591 bp for *Acanthamoeba castellani*[36] and 69,034 bp for the fresh water flagellate *Reclinomonas americana*.[26,37]

All of the mammalian mtDNAs are organized as shown in Fig. 18-3. The two strands of the DNA can be separated by virtue of their differing densities. The heavy (H) strand has a 5'→3' polarity in a counterclockwise direction in the map of Fig. 18-3, while the light (L) strand has a clockwise polarity. From the sequences 13 genes for specific proteins, 2 genes for ribosomal RNA molecules, and 22 genes for transfer RNAs have been identified. The genes are listed in Table 18-1 and have also been marked on the map in Fig. 18-3. The map also shows the tRNA genes, labeled with standard one-letter amino acid abbreviations, and the directions of transcription. Most of the protein genes are on the H-strand. One small region, the D-loop, contains an origin of replication and control signals for transcription (see Chapters 27 and 28).

The genes in mammalian mtDNA are closely packed with almost no nucleotides between them. However, the 19.5-kb mtDNA of *Drosophila* contains an (A+T)-rich region, which varies among species.[38] In the much larger 78-kb genome of yeast *Saccharomyces cerevisiae* many (A+T)-rich spacer regions are present.[39] The yeast mitochondrial genome also contains genes for several additional proteins. Mitochondria of *Reclinomonas americana* contain 97 genes, including those for 5S RNA, the RNA of ribonuclease P as well as a variety of protein coding genes. Perhaps this organism is primitive, resembling the original progenitor of eukaryotic life.[26] The mtDNA of trypanosomes is present in the large mitochondrion or kinetoplast as 40–50 "maxicircles" ~20–35 kb in size, together with 5000–10,000 "minicircles", each of 645–2500 bp (see Fig. 5-16). The latter encode **guide RNA** for use in RNA editing (Chapter 28). The large mitochondrial DNAs of higher plants, e.g., *Arabidopsis* (Table 18-1), also contain additional protein genes as well as large segments of DNA between the genes. The genome of the turnip (*Brassica campestris*) exists both as a 218-kb **master chromosome** and smaller 83- and 135-kb incomplete chromosomes, a pattern existing for most land plants.[40–42] The muskmelon contains 2500 kb of mitochondrial DNA (mtDNA). On the other hand, most mtDNA of the liverwort *Marchantia polymorpha* consists of 186-kb linear duplexes.[42a]

The compact size of the mammalian genome is dependent, in part, on alterations in the genetic code, as shown in Table 18-2, and a modification of tRNA structures that permits mitochondria to function with a maximum of 22 tRNAs (Chapter 28).[43–45] However, the more primitive *Reclinomonas* utilizes the standard genetic code in its mitochondria.[26] The mammalian mitochondrial genes contain no introns, but some yeast mitochondrial genes do. Furthermore, some of the introns contain long open reading frames. At least two of these genes-within-genes encode enzymes that excise the introns.

Why does mtDNA contain *any* protein genes, or why does mtDNA even exist? It seems remarkable that the cells of our bodies make the 100 or so extra proteins (encoded in the nucleus) needed for replication, transcription, amino acid activation, and mitochondrial ribosome formation and bring these into the mitochondria for the sole purpose of permitting the synthesis there of 13 proteins. The explanation is not evident. What are the 13 proteins?

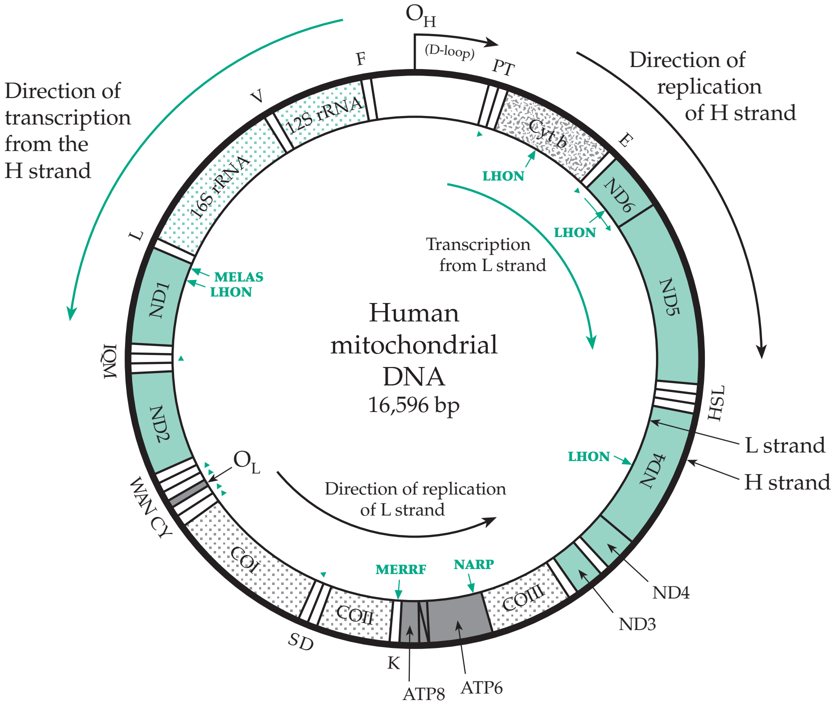


Figure 18-3 Genomic map of mammalian mitochondrial DNA. The stippled areas represent tRNA genes which are designated by the single-letter amino acid code; polarity is counterclockwise except for those marked by green arrow heads. All protein-coding genes are encoded on the H strand (with counterclockwise polarity), with the exception of ND6, which is encoded on the L strand. COI, COII, and COIII: cytochrome oxidase subunits I, II, and III; Cyt *b*: cytochrome *b*; ND: NADH dehydrogenase; ATP:ATP synthase. O_H and O_L: the origins of H and L strand replication, respectively. After Wallace[46] and Shoffner and Wallace.[45] Positions of a few of many known mutations that cause serious diseases are marked using abbrevations defined in Box 18-B.

TABLE 18-2
Alterations in the Genetic Code in the DNA of Animal Mitochondria

Codons	Nuclear DNA[a]	Mitochondrial DNA
AGA, ACG	Arg	Termination
AUA	Ile	Met
UGA	Termination	Trp

[a] See Table 5-5 for the other "standard" codons.

Three are the large functional subunits of cytochrome oxidase, one is cytochrome *b*, and seven are subunits of the NADH dehydrogenase system (Complex I). Two are subunits of ATP synthase. These are all vitally involved in the processes of electron transport and oxidative phosphorylation, but so are other proteins that are imported from the cytoplasm.

One gene in yeast mtDNA is especially puzzling. The *var 1* gene encodes a mitochondrial ribosomal protein, whose sequence varies with the strain of yeast. The gene is also involved in unusual recombinational events.[47] Another unusual aspect of yeast mitochondrial genetics is the frequent appearance of "petite" mutants, which grow on an agar surface as very small colonies. These have lost a large fraction of their mitochondrial DNA and, therefore, the ability to make ATP by oxidative phosphorylation. The remaining petite mtDNA may sometimes become integrated into nuclear DNA.[48] A few eukaryotes that have no aerobic metabolism also have no mitochondria.[49]

4. Growth and Development

Mitochondria arise by division and growth of preexisting mitochondria. Because they synthesize only a few proteins and RNA molecules, they must import many proteins and other materials from the cytoplasm. A mitochondrion contains at least 100 proteins that are encoded by nuclear genes.[50,50a] The mechanisms by which proteins are taken up by mitochondria are complex and varied. Many of the newly synthesized proteins carry, at the N terminus, presequences that contain **mitochondrial targeting signals**[51–53] (Chapter 10). These amino acid sequences often lead the protein to associate with receptor proteins on the outer mitochondrial membrane and subsequently to be taken up by the mitochondria. While the targeting sequences are usually at the N terminus of a polypeptide, they are quite often internal. The N-terminal sequences are usually removed by action of the **mitochondrial processing peptidase** (MPP) in the matrix, but internal targeting sequences are not removed.[52] Targeting of proteins to mitochondria may be assisted by **mRNA binding proteins** that guide appropriate mRNAs into the vicinity of mitochondria or other organelles.[53]

In addition to targeting signals, polypeptides destined for the inner mitochondrial membrane contain additional **topogenic signals** that direct the polypeptide to its destination. These topogenic signals are distinct from the targeting signals, which they sometimes follow. Topogenic signals are usually hydrophobic sequences, which may become transmembrane segments of the protein in its final location.[52,54] The uptake of many proteins by mitochondria requires the electrical potential that is usually present across the inner membrane (Section E). The fact that mitochondrial proteins usually have higher isoelectric points and carry more positive charges at neutral pH favors uptake.[55] In addition, chaperonins assist in

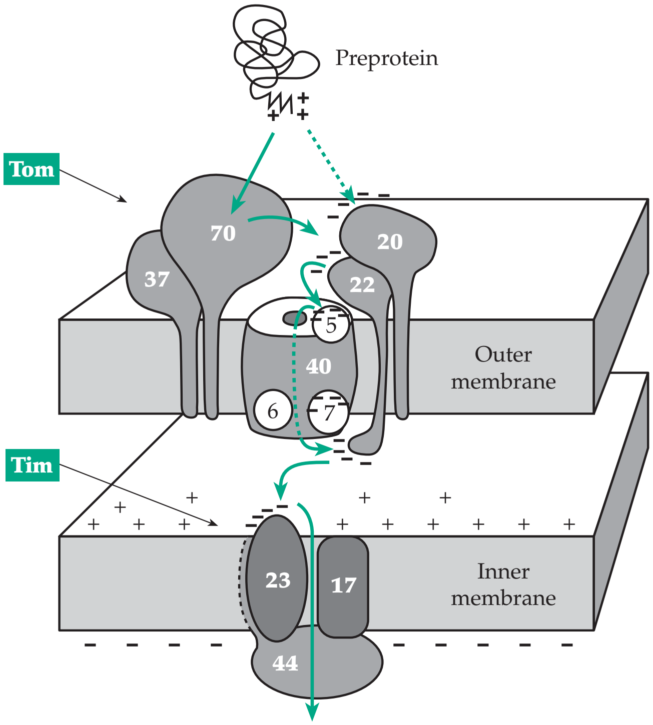


Figure 18-4 Schematic diagram of the protein transport machinery of mitochondrial membranes labeled according to the uniform nomenclature.[57] Subunits of outer membrane receptors and translocase are labeled Tom (translocase of outer membrane) and those of the inner membrane Tim (translocase of inner membrane). They are designated Tom70, etc., according to their sizes in kilodaltons (kDa). Preproteins are recognized by receptor Tom70•Tom 37 and / or by Tom22•Tom20. Clusters of negative charges on many components help guide the preprotein through the uptake pores in one or both membranes.[50,58] See Pfanner *et al.*,[57] Schatz,[50,50a] and Gabriel *et al.*[50b]

unfolding the protein to be taken up, assist in transport of some proteins,[50a] and may help the imported proteins to assemble into oligomeric structures.[51–53,56]

Protein uptake also requires a set of special proteins described as the **translocase of the outer mitochondrial membrane** (**Tom**) and **translocase of the mitochondrial inner membrane** (**Tim**). Subunits that form the receptor targets and transport pores are designated, according to their approximate molecular masses in kD as Tom70, Tim23, etc. (Fig. 18-4).[57] Preproteins are recognized by the receptor complexes Tom70 • Tom37 and / or Tom22 • Tom20 on the mitochondrial surface. They then enter the **general import pore** formed by Tom40, Tom6, and Tom7 with the assistance of a small integral membrane protein Tom5, which has a positively charged C-terminal membrane anchor segment and a negatively charged N-terminal portion that may bind to the positively charged mitochondrial targeting sequences.[50,59] A number of other translocase components, including Tom20 and Tom22 of the outer membrane and Tim23 of the inner membrane, also have acidic extramembranous domains.[58] This suggested an "acid chain" hypothesis according to which the targeting signal interacts consecutively with a series of acidic protein domains that help to guide it across the two membranes.[50a,58,59] A series of small proteins, Tim8, 9, 10, 12, 13, function in yeast mitochondria to mediate the uptake of metabolite transporters. A defect in the human nuclear gene for a protein that resembles Tim8 causes **deafness dystonia**, a recessive X-linked neurodegenerative disorder.[59a,b]

B. Electron Transport Chains

During the 1940s, when it had become clear that formation of ATP in mitochondria was coupled to electron transport, the first attempts to pick the system apart and understand the molecular mechanism began. This effort led to the identification and at least partial characterization of several flavoproteins, iron-sulfur centers, ubiquinones, and cytochromes, most of which have been described in Chapters 15 and 16. It also led to the picture of mitochondrial electron transport shown in Fig. 10-5 and which has been drawn in a modern form in Fig. 18-5.

1. The Composition of the Mitochondrial Electron Transport System

Because of the difficulty of isolating the electron transport chain from the rest of the mitochondrion, it is easiest to measure ratios of components (Table 18-3). Cytochromes a, a_3, b, c_1, and c vary from a 1:1 to a 3:1 ratio while flavins, ubiquinone, and nonheme iron occur in relatively larger amounts. The much larger amount of pyridine nucleotides is involved in carrying electrons from the various soluble dehydrogenases of the matrix to the immobile carriers in the inner membrane, while ubiquinone has a similar function within the lipid bilayer of mitochondrial membranes.

TABLE 18-3
Ratios of Components in the Electron Transport Chain of Mitochondria[a,b]

Electron carrier	Rat liver mitochondria	Beef heart mitochondria
Cytochrome a_3	1.0	1.1
Cytochrome a	1.0	1.1
Cytochrome b	1.0	1.0
Cytochrome c_1	0.63	0.33–0.51
Cytochrome c	0.78	0.66–0.85
Pyridine nucleotides	24	
Flavins	3	1
Ubiquinone	3–6	7
Copper		2.2
Nonheme iron		5.5

[a] From Wainio, W. W. (1970) *The Mammalian Mitochondrial Respiratory Chain*, Academic Press, New York, and references cited therein.

[b] Molecular ratios are given. Those for the cytochromes refer to the relative numbers of heme groups.

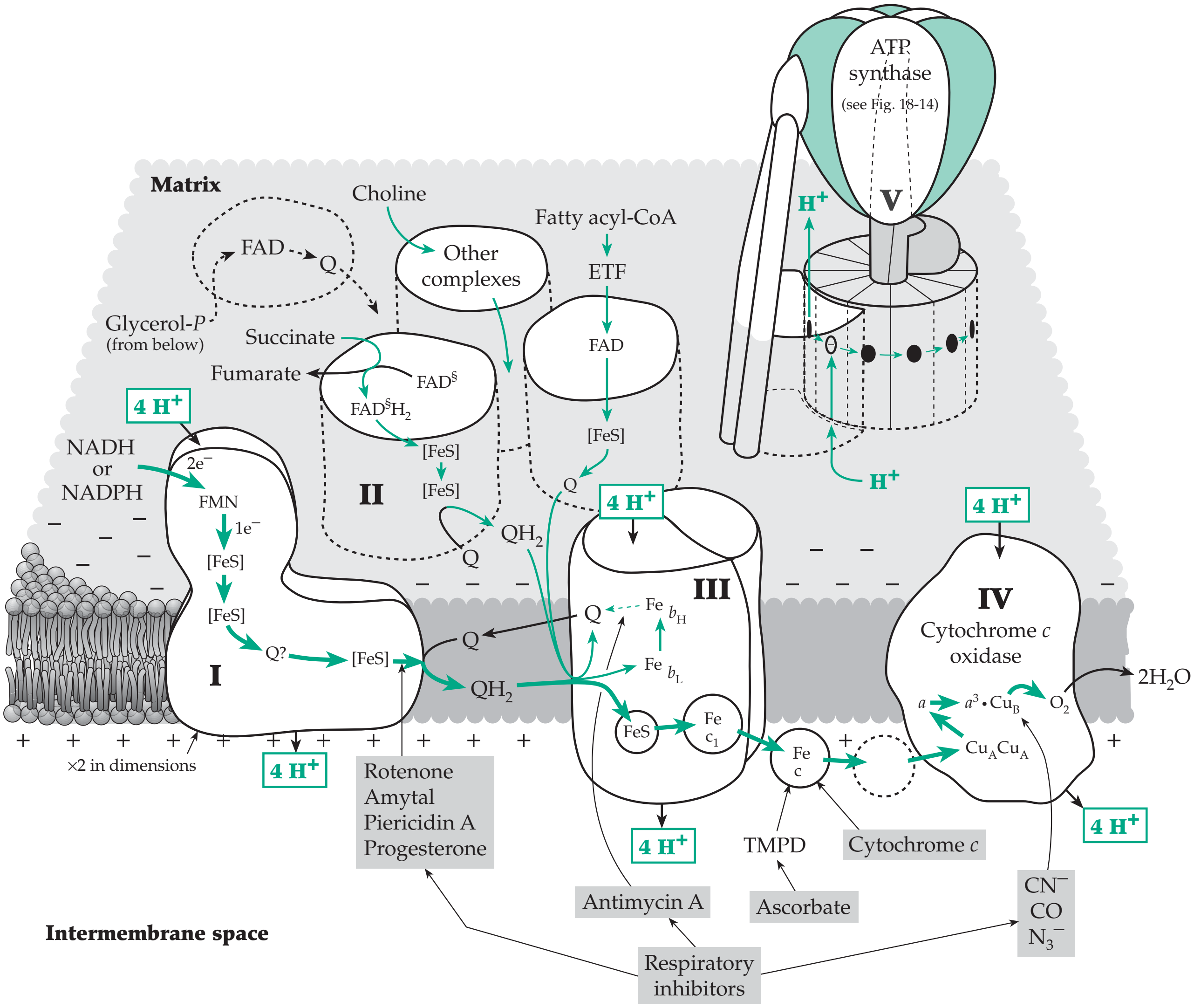
What are the molar concentrations of the electron carriers in mitochondrial membranes? In one experiment, cytochrome b was found in rat liver mitochondria to the extent of 0.28 μmol/g of protein. If we take a total mitochondrion as about 22% protein, the average concentration of the cytochrome would be ~0.06 mM. Since all the cytochromes are concentrated in the inner membranes, which may account for 10% or less of the volume of the mitochondrion, the concentration of cytochromes may approach 1 mM in these membranes. This is sufficient to ensure rapid reactions with substrates.

2. The Sequence of Electron Carriers

Many approaches have been used to deduce the sequence of carriers through which electron flow takes place (Fig. 18-5). In the first place, it seemed reasonable to suppose that the carriers should lie in order of increasing oxidation–reduction potential going from left to right of the figure. However, since the redox potentials existing in the mitochondria may be somewhat different from those in isolated enzyme preparations, this need not be strictly true.

The development by Chance of a dual wavelength spectrophotometer permitted easy observation of the state of oxidation or reduction of a given carrier within mitochondria.[60] This technique, together with the study of specific inhibitors (some of which are indicated in Fig. 18-5 and Table 18-4), allowed some electron transport sequences to be assigned. For example, blockage with **rotenone** and **amytal** prevented reduction of the cytochrome system by NADH but allowed reduction by succinate and by other substrates having their own flavoprotein components in the chain. Artificial electron acceptors, some of which are shown in Table 18-5, were used to bypass parts of the chain as indicated in Fig. 18-5.

Submitochondrial particles and complexes. Many methods have been employed to break mitochondrial membranes into submitochondrial particles that retain an ability to catalyze some of the reactions of the chain.[61] For example, the Keilin–Hartree preparation of heart muscle is obtained by homogenizing mitochondria and precipitation at low pH.[62] The resulting particles have a low cytochrome *c* content and do not carry out oxidative phosphorylation.



ATP synthase
(see Fig. 18-14)
V
H+
H+
Matrix
Choline
Other complexes
Fatty acyl-CoA
ETF
FAD
FAD
Q
Glycerol-P
(from below)
Succinate
Fumarate
FAD§
FAD§H2
[FeS]
[FeS]
[FeS]
II
Q
QH2
Q
4 H+
NADH or NADPH
2e−
FMN
1e−
[FeS]
[FeS]
Q?
[FeS]
I
Q
QH2
4 H+
III
Q
Fe bH
Fe bL
FeS
Fe c1
Fe c
4 H+
IV
Cytochrome c oxidase
a
a3 • CuB
O2
CuACuA
2H2O
4 H+
×2 in dimensions
4 H+
Rotenone
Amytal
Piericidin A
Progesterone
4 H+
TMPD
Cytochrome c
Antimycin A
Ascorbate
CN−
CO
N3−
Respiratory inhibitors
Intermembrane space

Figure 18-5 A current concept of the electron transport chain of mitochondria. Complexes I, III, and IV pass electrons from NADH or NADPH to O_2, one NADH or two electrons reducing one O to H_2O. This electron transport is coupled to the transfer of about 12 H^+ from the mitochondrial matrix to the intermembrane space. These protons flow back into the matrix through ATP synthase (V), four H^+ driving the synthesis of one ATP. Succinate, fatty acyl-CoA molecules, and other substrates are oxidized via complex II and similar complexes that reduce ubiquinone Q, the reduced form QH_2 carrying electrons to complex III. In some tissues of some organisms, glycerol phosphate is dehydrogenated by a complex that is accessible from the intermembrane space.

However, they do transport electrons and react with O_2. Other electron transport particles have been prepared by sonic oscillation. Under the electron microscope such particles appear to be small membranous vesicles resembling mitochondrial cristae.

Many detergents are strong denaturants of proteins, but some of them disrupt mitochondrial membranes without destroying enzymatic activity. A favorite is **digitonin** (Fig. 22-12), which causes disintegration of the outer membrane. The remaining fragments of inner membrane retain activity for oxidative phosphorylation. Such submitochondrial particles can be fractionated further by chemical treatments. Separate complexes can be obtained by treating the inner membranes with the nondenaturing detergent cholate (Fig. 22-10) and isolating the complexes by differential salt fractionation using ammonium sulfate. The isolated complexes I – IV catalyze reactions of four different portions of the electron transport process[63–65] as indicated in Eq. 18-1:

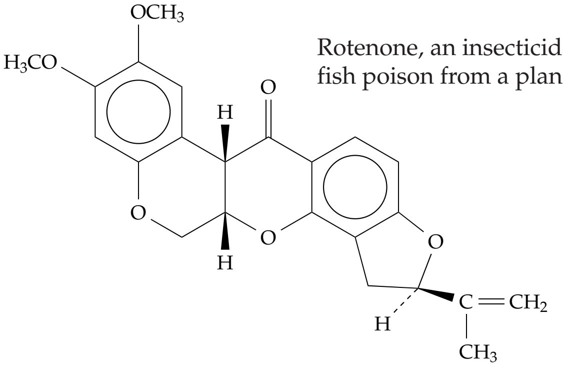
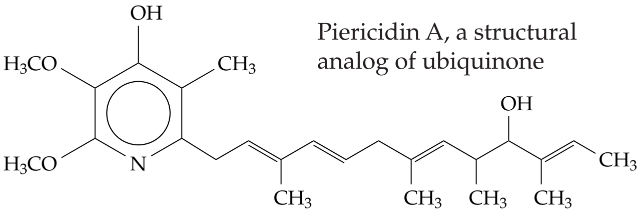
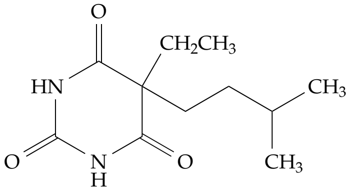
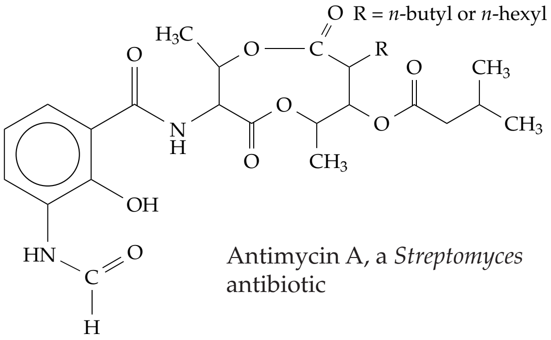
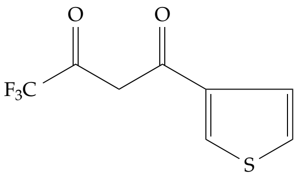
$$\begin{tikzcd} \text{Succinate} \arrow[dr, "\text{II}"] & & & \\ \text{NADH} \arrow[r, "\text{I}"] & \text{Q} \arrow[r, "\text{III}"] & \text{cytochrome } c \arrow[r, "\text{IV}"] & O_2 \end{tikzcd} \tag{18-1}$$

These complexes are usually named as follows: I, **NADH-ubiquinone oxidoreductase**; II, **succinate-ubiquinone oxidoreductase**; III, **ubiquinol-cytochrome *c* oxidoreductase**; IV, **cytochrome *c* oxidase**. The designation **complex V** is sometimes applied to ATP synthase (Fig. 18-14). Chemical analysis of the electron transport complexes verified the probable location of some components in the intact chain. For example, a high iron content was found in both complexes I and II and copper in complex IV.

We now recognize not only that these complexes are discrete structural units but also that they are functional units. Complete X-ray crystallographic structures are available for complexes III and IV and for much of the ATP synthase complex. As is indicated in Fig. 18-5, complexes I – IV are linked by two soluble electron carriers, ubiquinone and cytochrome *c*.

The lipid-soluble ubiquinone (Q) is present in both bacterial and mitochondrial membranes in relatively large amounts compared to other electron carriers (Table 18-2). It seems to be located at a point of convergence of the NADH, succinate, glycerol phosphate, and choline branches of the electron transport chain. Ubiquinone plays a role somewhat like that of NADH, which carries electrons between dehydrogenases in the cytoplasm and from soluble dehydrogenases in the aqueous mitochondrial matrix to flavoproteins embedded in the membrane. Ubiquinone transfers electrons plus protons between proteins within the

TABLE 18-4
Some Well-Known Respiratory Inhibitors[a]

 OCH$_3$ H$_3$CO O H O O H O H C=CH$_2$ CH$_3$	Rotenone, an insecticide and fish poison from a plant root
 OH H$_3$CO CH$_3$ H$_3$CO N OH CH$_3$ CH$_3$ CH$_3$ CH$_3$ CH$_3$	Piericidin A, a structural analog of ubiquinone
 O CH$_2$CH$_3$ HN CH$_3$ O N H O CH$_3$	Amytal (amobarbital)
Progesterone	(See Fig. 22-11)
 O R = *n*-butyl or *n*-hexyl H$_3$C O C R O O CH$_3$ N H O O O CH$_3$ CH$_3$ OH HN C O H	Antimycin A, a *Streptomyces* antibiotic
	Thenoyltrifluoroacetone (TTB, 4,4,4-trifluoro-1-(2-thienyl)-1,3-butadione
Cyanide	$^-C\equiv N$
Azide	$N\equiv N\equiv N^-$
Carbon monoxide	CO

[a] See Fig. 18-5 for sites of inhibition.

BOX 18-A HISTORICAL NOTES ON RESPIRATION

Animal respiration has been of serious interest to scientists since 1777, when Lavoisier concluded that foods undergo slow combustion within the body, supposedly in the blood. In 1803–1807, Spallanzani established for the first time that the tissues were the actual site of respiration, but his conclusions were largely ignored. In 1884, MacMunn discovered that cells contain the heme pigments, which are now known as **cytochromes**. However, the leading biochemists of the day dismissed the observations as experimental error, and it was not until the present century that serious study of the chemistry of biological oxidations began.[a,b]

Recognition that substrates are oxidized by **dehydrogenation** is usually attributed to H. Wieland. During the years 1912–1922 he showed that synthetic dyes, such as methylene blue, could be substituted for oxygen and would allow respiration of cells in the absence of O_2. Subsequent experiments (see Chapter 15) led to isolation of the soluble pyridine nucleotides and flavoproteins and to development of the concept of an electron transport chain.

Looking at the other end of the respiratory chain, Otto Warburg[c,d] noted in 1908 that all aerobic cells contain iron. Moreover, iron-containing charcoal prepared from blood catalyzed nonenzymatic oxidation of many substances, but iron-free charcoal prepared from cane sugar did not. Cyanide was found to inhibit tissue respiration at low concentrations similar to those needed to inhibit nonenzymatic catalysis by iron salts. On the basis of these investigations, Warburg proposed in 1925 that aerobic cells contain an iron-based *Atmungsferment* (respiration enzyme), which was later called **cytochrome oxidase**. It was inhibited by carbon monoxide.

Knowing that carbon monoxide complexes of hemes are dissociated by light, Warburg and Negelein, in 1928, determined the photochemical **action spectrum** (see Chapter 23) for reversal of the carbon monoxide inhibition of respiration of the yeast *Torula utilis*. The spectrum closely resembled the absorption spectrum of known heme derivatives (Fig. 16-7). Thus, it was proposed that O_2, as well as CO, combines with the iron of the heme group in the Atmungsferment.

Meanwhile, during 1919–1925, David Keilin, while peering through a microscope equipped with a spectroscopic ocular at thoracic muscles of flies and other insects, observed a pigment with four distinct absorption bands. At first he thought it was derived by some modification of hemoglobin, but when he found the same pigment in fresh baker's yeast, he recognized it as an important new substance. In his words:[e]

> *One day while I was examing a suspension of yeast freshly prepared from a few bits of compressed yeast shaken vigorously with a little water in a test-tube, I failed to find the characteristic four-banded absorption spectrum, but before I had time to remove the suspension from the field of vision of the microspectroscope the four absorption bands suddenly reappeared. This experiment was repeated time after time and always with the same result: the absorption bands disappeared on shaking the suspension with air and reappeared within a few seconds on standing.*
>
> *I must admit that this first visual perception of an intracellular respiratory process was one of the most impressive spectacles I have witnessed in the course of my work. Now I have no doubt that cytochrome is not only widely distributed in nature and completely independent of haemoglobin, but that it is an intracellular respiratory pigment which is much more important than haemoglobin.*

Keilin soon realized that three of the absorption bands, those at 604, 564, and 550 nm (*a*, *b*, and *c*), represented different pigments, while the one at 521 nm was common to all three. Keilin proposed the names cytochromes *a*, *b*, and *c*. The idea of an electron transport or respiratory chain followed[e] quickly as the flavin and pyridine nucleotide coenzymes were recognized to play their role at the dehydrogenase level. Hydrogen removed from substrates by these carriers could be used to oxidize reduced cytochromes. The latter would be oxidized by oxygen under the influence of cytochrome oxidase.

In 1929, Fiske and Subbarow,[d,f–h] curious about the occurrence of purine compounds in muscle extracts, discovered and characterized ATP. It was soon shown (largely through the work of Lundsgaard and Lohman)[f] that hydrolysis of ATP provided energy for muscular contraction. At about the same time, it was learned that synthesis of ATP accompanied glycolysis. That ATP could also be formed as a result of electron transport became clear following an observation of Engelhardt[h,i] in 1930, that methylene blue stimulated ATP synthesis by tissues.

The study of electron transport chains and of oxidative phosphorylation began in earnest after Kennedy and Lehninger,[j] in 1949, showed that mitochondria were the site not only of ATP synthesis but also of the operation of the citric acid cycle and fatty acid oxidation pathways. By 1959, Chance had introduced elegant new techniques of spectrophotometry that led to formulation of the electron

BOX 18-A (continued)

transport chain as follows:

$$\text{Substrate} \rightarrow \text{pyridine nucleotides} \rightarrow \text{flavoprotein} \rightarrow \text{cyt } b \rightarrow \text{cyt } c \rightarrow \text{cyt } a \rightarrow \text{cyt } a_3 \rightarrow O_2$$

Since that time, some new components have been added, notably the ubiquinones and iron-sulfur proteins, but the basic form proposed for the chain was correct.

[a] Kalckar, H. M. (1969) *Biological Phosphorylations*, Prentice-Hall, Englewood Cliffs, New Jersey

[b] Kalckar, H. M. (1991) *Ann. Rev. Biochem.* **60**, 1–37

[c] Edsall, J. T. (1979) *Science* **205**, 384–385

[d] Fiske, C. H., and Subbarow, Y. (1929) *Science* **70**, 381–382

[e] Keilin, D. (1966) *The History of Cell Respiration and Cytochrome*, Cambridge Univ. Press, London and New York

[f] Kalckar, H. (1980) *Trends Biochem. Sci.* **5**, 56–57

[g] Schlenk, F. (1987) *Trends Biochem. Sci.* **12**, 367–368

[h] Saraste, M. (1998) *Science* **283**, 1488–1493

[i] Slater, E. C. (1981) *Trends Biochem. Sci.* **6**, 226–227

[j] Talalay, P., and Lane, M. D. (1986) *Trends Biochem. Sci.* **11**, 356–358

membrane bilayer. Membranes also contain **ubiquinone-binding proteins**,[66,67] which probably hold the ubiquinone that is actively involved in electron transport. Perhaps some ubiquinone molecules function as fixed carriers. There is also uncertainty about the number of sites at which ubiquinone functions in the chain.

Mitochondrial electron transport in plants and fungi. Plant mitochondria resemble those of mammals in many ways, but they contain additional dehydrogenases and sometimes utilize alternative pathways of electron transport,[68–73] as do fungi.[74] Mitochondria are impermeable to NADH and NAD^+. Animal mitochondria have shuttle systems (see Fig. 18-16) for bringing the reducing equivalents of NADH into mitochondria and to the NADH dehydrogenase that faces the matrix side of the inner membrane. However, plant mitochondria also have an NADH dehydrogenase on the *outer* surface of the inner membrane (Fig. 18-6). This enzyme transfers electrons to ubiquinone, is not inhibited by rotenone (see Fig. 18-5), and also acts on NADPH. Inside the mitochondria a high-affinity NADH dehydrogenase resembles complex I of animal mitochondria and is inhibited by rotenone.[73] There is also a low-affinity NADH dehydrogenase, which is insensitive to rotenone. Some plant mitochondria respire slowly in the presence of cyanide. They utilize an **alternative oxidase** that replaces complex III and cytochome *c* oxidase and which is not inhibited by antimycin or by cyanide (Fig. 18-6).[68,71,75] It is especially active in thermogenic plant tissues (Box 18-C). A

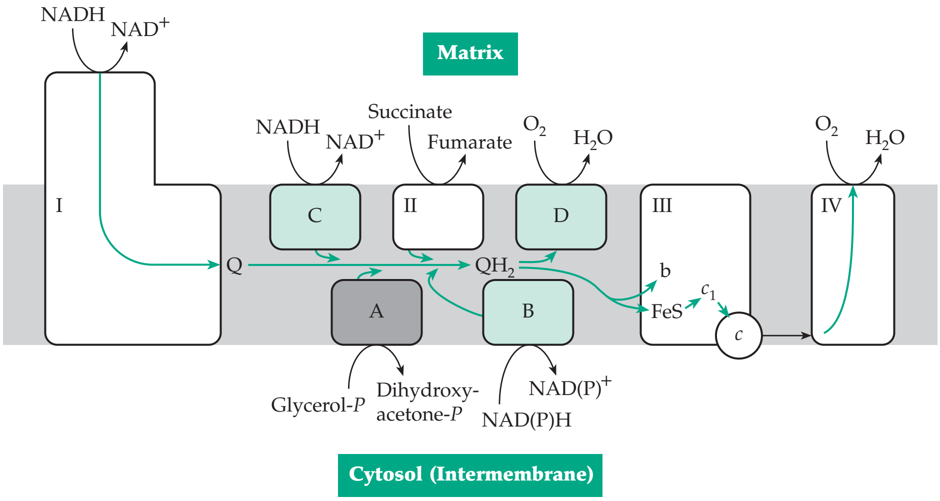


Figure 18-6 Schematic diagram of some mitochondrial dehydrogenase and oxidase complexes of plants and also the glycerol phosphate dehydrogenase of animals, which is embedded in the inner membrane. Complexes I–IV are also shown. (A) The glycerol phosphate dehydrogenase of some animal tissues. It is accessible from the intermembrane space on the cytosolic side. (B) The rotenone-insensitive NAD(P)H dehydrogenase of the external membrane surface of plants. (C) The rotenone-insensitive NADH dehydrogenase facing the matrix side in some plants. (D) The plant alternative oxidase. Ubiquinone, Q. The three green stippled dehydrogenases are not coupled to proton pumps or ATP synthesis. After Hoefnagel *et al.*[73]

BOX 18-B DEFECTS OF MITOCHONDRIAL DNA

A mutation in any of the 13 protein subunits, the 22 tRNAs, or the two rRNAs whose genes are carried in mitochondrial DNA may possibly cause disease. The 13 protein subunits are all involved in electron transport or oxidative phosphorylation. The syndromes resulting from mutations in mtDNA frequently affect oxidative phosphorylation (OXPHOS) causing what are often called “OXPHOS diseases.”[a–g] Mitochondrial oxidative phosphorylation also depends upon ~100 proteins encoded in the nucleus. Therefore, OXPHOS diseases may result from defects in either mitochondrial or nuclear genes. The former are distinguished by the fact that they are inherited almost exclusively maternally. Most mitochondrial diseases are rare. However, mtDNA is subject to rapid mutation, and it is possible that accumulating mutants in mtDNA may be an important component of aging.[h–k]

The first recognition of mitochondrial disease came in 1959. A 30-year old Swedish woman was found to have an extremely high basal metabolic rate (180% of normal), a high caloric intake (>3000 kcal / day), and an enormous perspiration rate. She had developed these symptoms at age seven. Examination of her mitochondria revealed that electron transport and oxidative phosphorylation were very loosely coupled. This explains the symptoms. However, the disease (Luft disease) is extremely rare and the underlying cause isn’t known.[i] Its recognition did focus attention on mitochondria, and by 1988, 120 different mtDNA defects had been described.[e,i]

Some OXPHOS disorders, including Luft disease, result from mutations in nuclear DNA. A second group arise from point mutations in mtDNA and a third group involve deletions, often very large, in mtDNA. Persons with these deletions survive because they have both mutated and normal mtDNA, a condition of **heteroplasmy** of mtDNA. As these persons age their disease may become more severe because they lose many normal mitochondria.[d,e]

The names of mitochondrial diseases are often complex and usually are described by abbreviations. Here are a few of them: **LHON**, Lebers hereditary optical neuropathy; **MERRF**, myoclonic epilepsy and ragged-red-fiber disease; **MELAS**, mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes; **NARP**, neurological muscle weakness, ataxia, and retinitis pigmentosa; **Leigh disease — SNE**, subacute necrotizing encephalomyelopathy; **KSS**, Kearns–Sayre syndrome; **CPEO**, chronic progressive external ophthalmoplegia. LHON is a hereditary disease that often leads to sudden blindness from death of the optic nerve especially among males. Any one of several point mutations in subunits ND1, 2, 4, 5, and 6 of NADH dehydrogenase (complex I; Figs. 18-3 and 18-5), cytochrome *b* of complex II, or subunit I of cytochrome oxidase may cause this syndrome. Most frequent is an R340H mutation of the ND4 gene at position 11,778 of mtDNA (Fig. 18-3).[e,l,m] It may interfere with reduction of ubiquinone.[n] Mutations in the ND1 gene at position 3460 and in the ND6 gene at position 14484 or in the cytochrome *b* gene at position 15257 cause the same disease.[l] The most frequent (80–90%) cause of MERRF, which is characterized by epilepsy and by the appearance of ragged red fibers in stained sections of muscle, is an A → G substitution at position 8344 of mtDNA in the TψC loop (Fig. 5-30) of mitochondrial tRNA^{Lys}. A similar disease, MELAS, is accompanied by strokes (not seen in MERRF) and is caused in 80% of cases by an A → G substitution in the dihydrouridine loop (Fig. 5-30) of mitochondrial tRNA^{Leu}.[o] CPEO, Leigh disease, and KSS often result from large deletions of mtDNA.[p] NARP and related conditions have been associated with an L156R substitution in the ATPase 6 gene of ATP synthase.[q]

Can mitochondrial diseases be treated? Attempts are being made to improve the function of impaired mitochondria by adding large amounts of ubiquinone, vitamin K, thiamin, riboflavin, and succinate to the diet.[e] One report suggests that mitochondrial decay during aging can be reversed by administration of *N*-acetylcarnitine.[k]

---

[a] Palca, J. (1990) *Science* **249**, 1104–1105

[b] Capaldi, R. A. (1988) *Trends Biochem. Sci.* **13**, 144–148

[c] Darley-Usmar, V., ed. (1994) *Mitochondria: DNA, Proteins and Disease*, Portland Press, London

[d] Wallace, D. C. (1999) *Science* **283**, 1482–1488

[e] Shoffner, J. M., and Wallace, D. C. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1535–1609, McGraw-Hill, New York

[f] Schon, E. A. (2000) *Trends Biochem. Sci.* **25**, 555–560

[g] Tyler, D. D. (1992) *The Mitochondrion in Health and Disease*, VCH Publ., New York

[h] Wallace, D. C. (1992) *Science* **256**, 628–632

[i] Luft, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8731–8738

[j] Tanhauser, S. M., and Laipis, P. J. (1995) *J. Biol. Chem.* **270**, 24769–24775

[k] Shigenaga, M. K., Hagen, T. M., and Ames, B. N. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10771–10778

[l] Hofhaus, G., Johns, D. R., Hurko, O., Attardi, G., and Chomyn, A. (1996) *J. Biol. Chem.* **271**, 13155–13161

[m] Brown, M. D., Trounce, I. A., Jun, A. S., Allen, J. C., and Wallace, D. C. (2000) *J. Biol. Chem.* **275**, 39831–39836

[n] Zickermann, V., Barquera, B., Wikström, M., and Finel, M. (1998) *Biochemistry* **37**, 11792–11796

[o] Hayashi, J.-I., Ohta, S., Kagawa, Y., Takai, D., Miyabayashi, S., Tada, K., Fukushima, H., Inui, K., Okada, S., Goto, Y., and Nonaka, I. (1994) *J. Biol. Chem.* **269**, 19060–19066

[p] Moraes, C. T., and 19 other authors. (1989) *N. Engl. J. Med.* **320**, 1293–1299

[q] Hartzog, P. E., and Cain, B. D. (1993) *J. Biol. Chem.* **268**, 12250–12252

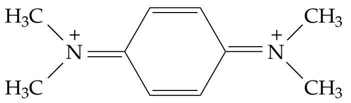
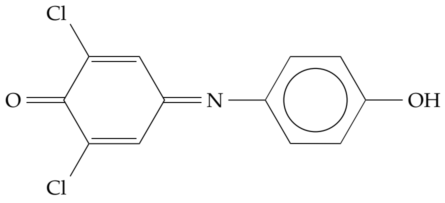
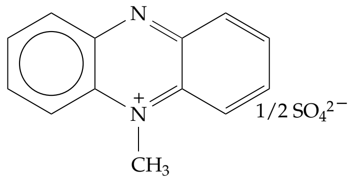
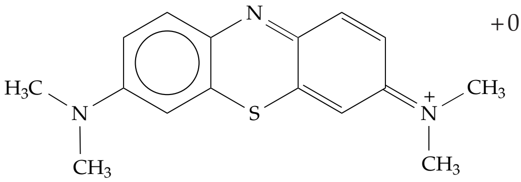
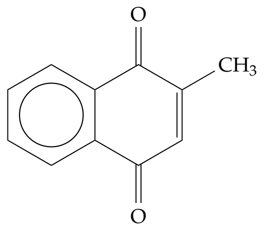
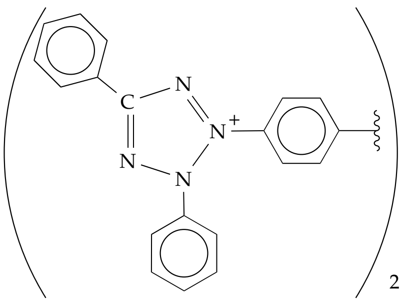
similar oxidase is present in trypanosomes.[72] Neither the rotenone-insensitive dehydrogenases nor the alternative oxidases are coupled to synthesis of ATP.

Electron transport chains of bacteria. The bacterial electron transport systems are similar to that of mitochondria but simpler. Bacteria also have a variety of alternative pathways that allow them to adapt to various food sources and environmental conditions.[76,77] The gram-negative soil bacterium *Paracoccus denitrificans*, which has been called "a free-living mitochondrion," has a mammalian-type respiratory system. Its complexes I–IV resemble those of animals and of fungi,[78–79] but *Paracoccus* has fewer subunits in each complex. Complex I of *E. coli* is also similar to that of our own bodies.[79–80] However, other major flavoprotein dehydrogenases in *E. coli* act on D-lactate and *sn*-3-glycerol phosphate.[81] Pyruvate is oxidized by a membrane-bound flavoprotein (Fig. 14-2). All of these enzymes pass electrons to ubiquinone-8 (Q_8).[82] Succinate dehydrogenase of *E. coli* resembles that of mitochondria,[83] and the ubiquinol oxidase of *Paracoccus* resembles complexes III + IV of mitochondria. It can be resolved into a three-subunit bc_1 complex, a three-subunit c_1aa_3 complex, and another 57-kDa peptide.[84] The last contains a 22-kDa cytochrome c_{552}, which is considerably larger than mitochondrial cytochrome *c*.

The cytochrome aa_3 terminal oxidase is produced constitutively, i.e., under all conditions. However, when cells are grown on succinate or H_2 another set of enzymes is produced with the *b*-type cytochrome o_3 as the terminal **quinol oxidase** (Eq. 18-2).[85]

$$\begin{tikzcd}
 & H_2 \arrow[d] & \\
\text{Succinate} \arrow[r] & Q \arrow[r] & \text{cytochrome } b \arrow[d] \\
 & & \text{cytochrome } o_3 \arrow[d] \\
 & & O_2
\end{tikzcd} \quad (18\text{-}2)$$

TABLE 18-5
Some Artificial Electron Acceptors[a,b]

Compound	Structure	E°′(pH 7) 30°C
Ferricyanide	$Fe(CN)_6^{3-}$	+0.36 V (25°C)
Oxidized form of tetramethyl-*p*-phenylenediamine		+0.260 V
2,6-Dichlorophenol-indophenol (DCIP)		+0.217 V
Phenazine methosulfate (PMS)		+0.080 V
Ascorbate	(See Box 18-D)	+0.058 V
Methylene blue		+0.011 V
Menadione		+0.008 V (°25C)
Tetrazolium salts, e.g., "neotetrazolium chloride"		−0.125 V

---

[a] From Wainio, W. W. (1970) *The Mammalian Mitochondrial Respiratory Chain*, Academic Press, New York, pp. 106–111.

[b] See Fig. 18-5 for sites of action.

Two terminal quinol oxidase systems, both related to cytochrome *c* oxidase, are utilized by *E. coli* to oxidize ubiquinol-8. When cultured at high oxygen tensions, cytochrome bo_3 (also called cytochrome *bo*) is the major oxidase. It utilizes heme *o* (Fig. 16-5) instead of heme *a*. However, at low oxygen tension, e.g., in the late logarithmic stage of growth, the second oxidase, cytochrome *bd*, is formed.[76,86–88a] It contains two molecules of the chlorin heme *d* (Fig. 16-5), which appear to be involved directly in binding O_2. This terminal oxidase system is present in many bacteria and can utilize either O_2 or nitrite as the oxidant. A simpler electron transport chain appears to be involved in the oxidation of pyruvate by *E. coli*. The flavoprotein pyruvate oxidase passes electrons to Q_8, whose reduced form can pass electrons directly to cytochrome *d*. Incorporation of these two pure protein complexes and ubiquinone-8 into phospholipid vesicles has given an active reconstituted chain.[82] Other bacteria utilize a variety of quinol oxidase systems, which contain various combinations of cytochromes: aa_3, caa_3, *cao*, bo_3, and ba_3.[88b,c]

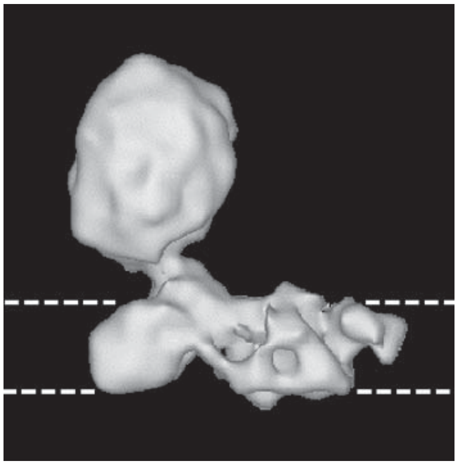


Figure 18-7 Three-dimensional image of bovine NADH-Ubiquinone oxidoreductase (complex I) reconstructed from individual images obtained by electron cyro-microscopy. The resolution is 2.2 nm. The upper portion projects into the mitochondrial matrix while the horizontal part lies within the membrane as indicated. Courtesy of N. Grigorieff.[90]

3. Structures and Functions of the Individual Complexes I – IV and Related Bacterial Assemblies

What are the structures of the individual electron transport complexes? What are the subunit compositions? What cofactors are present? How are electrons transferred? How are protons pumped? We will consider these questions for each of complexes I–IV, as found in both prokaryotes and eukaryotes.[88d,e]

Complex I, NADH-ubiquinone oxidoreductase. Complex I oxidizes NADH, which is generated within the mitochondrial matrix by many dehydrogenases. Among these are the pyruvate, 2-oxoglutarate, malate, and isocitrate dehydrogenases, which function in the tricarboxylic acid cycle; the β-oxoacyl-CoA dehydrogenase of the β oxidation system for fatty acids; and 2-hydroxybutyrate, glutamate, and proline dehydrogenases. All produce NADH, which reacts with the flavoprotein component of complex I. Whether from bacteria,[79] fungal mitochondria,[89] or mammalian mitochondria[89a,90] complex I exists as an L-shaped object, of which each of the two arms is ~23 nm long. One arm projects into the matrix while the other lies largely within the inner mitochondrial membrane (Fig. 18-7). The mitochondrial complex, which has a mass of ~1 MDa, has the same basic structure as the 530-kDa bacterial complex. However, the arms are thicker in the mitochondrial complex. Analysis of the denatured proteins by gel electrophoresis revealed at least 43 peptides.[78,90] Bound to some of these are the electron carriers FMN, Fe_2S_2, and Fe_4S_4 clusters, ubiquinone or other quinones, and perhaps additional unidentified cofactors.[79] Complex I from *E. coli* is smaller, containing only 14 subunits. These are encoded by a cluster of 14 genes, which can be directly related by their sequences to subunits of mitochondrial complex I and also to the corresponding genes of *Paracoccus denitrificans*.[80,91] Complex I of *Neurospora* contains at least 35 subunits.[89] The 14 subunits that are present both in bacteria and in mitochondria probably form the structural core of the complex. The other subunits thicken, strengthen, and rigidify the arms. Some of the "extra" subunits have enzymatic activities that are not directly related to electron transport. Among these are a 10-kDa prokaryotic type acyl carrier protein (ACP), which may be a relic of a bacterial fatty acid synthase, reflecting the endosymbiotic origin of mitochondria.[92] Also present is a 40-kDa NAD(P)H dependent reductase / isomerase, which may be involved in a biosynthetic process, e.g., synthesis of a yet unknown redox group.[79,92]

In all cases, FMN is apparently the immediate acceptor of electrons from NADH. From the results of extrusion of the Fe–S cores (Chapter 16) and EPR measurements it was concluded that there are three tetranuclear (Fe_4S_4) iron–sulfur centers and at least two binuclear (Fe_2S_2) centers[93,94] as well as bound ubiquinone.[95] Chemical analysis of iron and sulfide suggested up to eight Fe–S clusters per FMN, while gene sequences reveal potential sites for formation of six Fe_4S_4 clusters and two Fe_2S_2 clusters.[78] Treatment of complex I with such "chaotropic agents" as 2.5 M urea or 4 M sodium trichloroacetate followed by fractionation with ammonium sulfate[95] gave three fractions:

(1) A soluble NADH dehydrogenase consisting of a 51-kDa peptide that binds both the FMN and also one tetranuclear Fe – S cluster (designated N3) and a 24-kDa peptide that carries a binuclear Fe–S center designated N1b. (2) A 75-kDa peptide bearing two binuclear Fe–S centers, one of which is called N1a and also 47-, 30-, and 13-kDa peptides. One of these carries tetranuclear center N4. (3) A group of insoluble, relatively nonpolar proteins, one of which carries tetranuclear cluster N2. It may be the immediate donor of electrons to a ubiquinone held by a ubiquinone-binding protein designated QP-N. In bacteria seven of these are homologs of the seven NADH dehydrogenase subunits encoded by mtDNA (Fig. 18-3). A 49-kDa subunit of complex I in the yeast *Yarrowia lipolytica* is strikingly similar to the hydrogen reactive subunit of NiFe hydrogenases (Fig. 16-26).[95a] These proteins are thought to lie within the membrane arm and to form ~55 transmembrane α helices. Ubiquinone may also function as a carrier within complex I,[96,97] and there may be a new redox cofactor as well.[79] The following tentative sequence (Eq. 18-3) for electron transfer within complex I (with apparent $E^{\circ\prime}$ values of carriers) has been suggested. By equilibration with external redox systems, the redox potentials of these centers within the mitochondria have been estimated and are given (in V) in Eq. 18-3. The presence of a large fraction of the bound ubiquinone as a free radical suggests that the quinone functions as a one-electron acceptor rather than a two-electron acceptor. A characteristic of complex I is inhibition by rotenone or piericidin, both of which block electron transport at the site indicated in Fig. 18-5.

$$\begin{array}{ccccccccccccccc}
\mathrm{NADH} & \rightarrow & \mathrm{FMN} & \rightarrow & \mathrm{N1a} & \rightarrow & (\mathrm{N3}, & \mathrm{N1b}, & \mathrm{N4}) & \rightarrow & \mathrm{N2} & \rightarrow & \mathrm{QP\text{-}N} & \xrightarrow[\ominus]{\text{rotenone}} & \mathrm{Q} \\
 & & & & Fe_2S_2 & & Fe_4S_4 & Fe_2S_2 & Fe_4S_4 & & Fe_4S_4 & & & & \\
-0.32 & & & & -0.38 & & & -0.24 & & & -0.21\ \mathrm{V} & & & &
\end{array} \qquad (18\text{-}3)$$

Complex II, succinate-ubiquinone oxidoreductase.

Complex II, which carries electrons from succinate to ubiquinone, contains covalently linked 8^{α}-(*N*-histidyl)-FAD (Chapter 15) as well as Fe–S centers and one or more ubiquinone-binding sites. There are four subunits whose structures and properties have been highly conserved among mitochondria and bacteria and also in **fumarate reductases**. The latter function in the opposite direction during anaerobic respiration with fumarate as the terminal oxidant, both in bacteria[98–99a] and in parasitic helminths and other eukaryotes that can survive prolonged anaerobic conditions (Chapter 17, Section F,2).[100] Complex II from *E. coli* consists of 64-, 27-, 14-, and 13-kDa subunits, which are encoded by genes *sdhCDAB* of a single operon.[101–103] The two larger hydrophilic subunits associate to form the readily soluble succinate dehydrogenase. The 64-kDa subunit carries the covalently bound FAD while the 27-kDa subunit carries three Fe–S centers. The two small 13- and 14-kDa subunits form a hydrophobic anchor and contain a ubiquinol-binding site (QD-S)[103] as well as a heme that may bridge the two subunits[102] to form cytochrome b_{556}. The functions of the heme is uncertain. The soluble mammalian succinate dehydrogenase resembles closely that of *E. coli* and contains three Fe – S centers: binuclear S1 of $E^{\circ\prime}$ 0 V, and tetranuclear S2 and S3 of −0.25 to −0.40 and + 0.065 V, respectively. Center S3 appears to operate between the −2 and −1 states of Eq. 16-17 just as does the cluster in the *Chromatium* high potential iron protein. The function of the very low potential S2 is not certain, but the following sequence of electron transport involving S1 and S3 and the bound ubiquinone QD–S[66] has been proposed (Eq. 18-4).

$$\text{Succinate} \rightarrow (\text{FAD}^*, \text{S1}) \rightarrow \text{S3} \rightarrow \text{QP–S} \rightarrow \text{Q} \qquad (18\text{-}4)$$

In addition to complexes I and II several other membrane-associated FAD-containing dehydrogenase systems also send electrons to soluble ubiquinone. These include dehydrogenases for choline, *sn*-glycerol 3-phosphate, and the electron-transfer protein (ETF) of the fatty acyl-CoA β oxidation system (Fig. 18-5). The last also accepts electrons from dehydrogenases for sarcosine (*N*-methylglycine), dimethylglycine, and other substrates. The *sn*-glycerol 3-phosphate dehydrogenase is distinguished by its accessibility from the intermembrane (cytosolic) face of the inner mitochondrial membrane (Fig. 18-6).

Complex III (ubiquinol-cytochrome c oxidoreductase or cytochrome bc_1 complex).

Mitochondrial complex III is a dimeric complex, each subunit of which contains 11 different subunits with a total molecular mass of ~240 kDa per monomer.[104–107] However, in many bacteria the complex consists of only three subunits, cytochrome *b*, cytochrome c_1, and the high potential (~0.3 V) Rieske iron-sulfur protein, which is discussed in Chapter 16, Section A,7. These three proteins are present in all bc_1 complexes. In eukaryotes the 379-residue cytochrome *b* is mitochondrially encoded. Although there is only one cytochrome *b* gene in the mtDNA, two forms of cytochrome *b* can be seen in absorption spectra: b_H (also called b_{562} or b_K) and lower potential b_L (also called b_{566} or b_T).[107a,b]

X-ray diffraction studies have revealed the complete 11-subunit structure of bovine bc_1 complex[104,106–107] as well as a nearly complete structure of the chicken bc_1 complex (Fig. 18-8).[105] The bovine complex contains 2166 amino acid residues per 248-kDa monomer and exists in crystals as a 496-kDa dimer and probably functions as a dimer.[106-107] The two hemes of cytochrome *b* are near the two sides of the membrane, and the Fe–S and cytochrome c_1 subunits are on the surface next to the intermembrane space (Fig. 18-8). On the matrix side (bottom in Fig. 18-8A) are two large ~440 residue "core" subunits that resemble subunits of the mitochondrial processing protease. They may be evolutionary relics of that enzyme.[106,108,108a] Mitochondrial cytochrome b_H has an $E°'$ value of +0.050 V, while that of b_L is −0.090 V at pH 7.[109] That of the Rieske Fe–S protein is + 0.28 V.[110]

The sequence of electron transport within complex III has been hard to determine in detail. For reasons discussed in Section C, the "Q-cycle" shown in Fig. 18-9 has been proposed.[111–114a] As is indicated in Fig. 18-9, complex II accepts electrons from QH_2 and passes them consecutively to the Fe–S protein, cytochrome c_1, and the external cytochrome *c*. However, half of the electrons are recycled through the two heme groups of cytochrome *b*, as is indicated in the figure and explained in the legend. The X-ray structure (Fig. 18-8) is consistent with this interpretation. Especially intriguing is the fact that the Fe_2S_2 cluster of the Rieske protein subunit has been observed in two or three different conformations.[105–107,114a-c] In Fig. 18-8C the structures of two conformations are superimposed. The position of the long helix at the right side is unchanged but the globular domain at the top can be tilted up to bring the Fe_2S_2 cluster close to the heme of cytochrome c_1, or down to bring the cluster close to heme b_L. Movement between these two positions is probably part of the catalytic cycle.[115]

The simpler cytochrome bc_1 complexes of bacteria such as *E. coli*,[102] *Paracoccus dentrificans*,[116] and the photosynthetic *Rhodobacter capsulatus*[117] all appear to function in a manner similar to that of the large mitochondrial complex. The bc_1 complex of *Bacillus subtilis* oxidizes reduced **menaquinone** (Fig. 15-24) rather than ubiquinol.[118] In chloroplasts of green plants photochemically reduced **plastoquinone** is oxidized by a similar complex of cytochrome *b*, *c*-type cytochrome *f*, and a Rieske Fe–S protein.[119–120a] This cytochrome b_6f complex delivers electrons to the copper protein plastocyanin (Fig. 23-18).

The electron acceptor for complex III is cytochrome *c*, which, unlike the other cytochromes, is water soluble and easily released from mitochondrial membranes. Nevertheless, it is usually present in a roughly 1:1 ratio with the fixed cytochromes, and it seems unlikely that it is as free to diffuse as are ubiquinone and NAD^+.[121,122] However, a small fraction of the cytochrome *c* may diffuse through the intermembrane space and accept electrons from cytochrome b_5, which is located in the outer membrane.[123] Cytochrome *c* forms a complex with cardiolipin (diphosphatidylglycerol), a characteristic component of the inner mitochondrial membrane.[124]

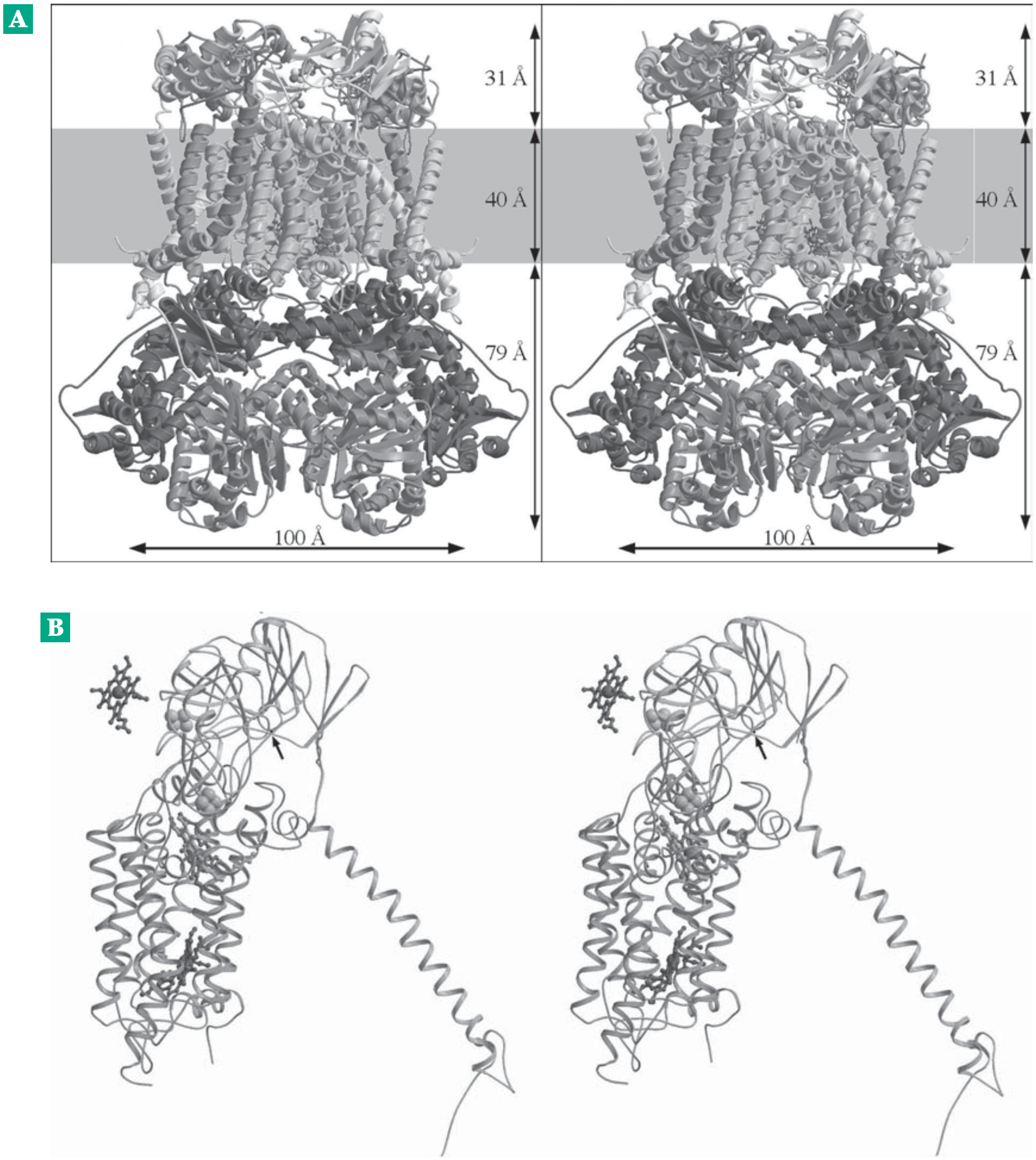
Complex IV. Cytochrome c oxidase (ubiquinol-cytochrome c oxidoreductase). Complex IV from mammalian mitochondria contains 13 subunits. All of them have been sequenced, and the three-dimensional structure of the complete complex is known (Fig. 18-10).[125–127] The simpler cytochrome *c* oxidase from *Paracoccus denitrificans* is similar but consists of only three subunits. These are homologous in sequence to those of the large subunits I, II, and III of the mitochondrial complex. The three-dimensional structure of the *Paracoccus* complex is also known. Its basic structure is nearly identical to that of the catalytic core of subunits I, II, and III of the mitochondrial complex (Fig. 18-10,A).[128] All three subunits have transmembrane helices. Subunit III seems to be structural in function, while subunits I and II contain the oxidoreductase centers: two hemes *a* (*a* and a_3) and two different copper centers, Cu_A (which contains two Cu^{2+}) and a third Cu^{2+} (Cu_B) which exists in an EPR-silent exchange coupled pair with a_3. Bound Mg^{2+} and Zn^{2+} are also present in the locations indicated in Fig. 18-10.

The Cu_A center has an unusual structure.[130–132] It was thought to be a single atom of copper until the three-dimensional structure revealed a dimetal center, whose structure follows. The Cu_B-cytochrome a_3 center is also unusual. A histidine ring is covalently attached to tyrosine.[133–135a] Like the tyrosine in the active site of galactose oxidase (Figs. 16-29, 16-30), which carries a covalently joined cysteine, that of cytochrome oxidase may be a site of tyrosyl radical formation.[135]

Cytochrome *c* oxidase accepts four electrons, one at a time from cytochrome *c*, and uses them to reduce O_2 to two H_2O. Electrons enter the oxidase via the Cu_A center and from there pass to the cytochrome *a* and on to the cytochrome a_3 – Cu_B center where the reduction of O_2 takes place. A possible sequence of steps in the catalytic cycle is given in Fig. 18-11. Reduction of O_2 to two H_2O requires four electrons and also four protons. An additional four protons are evidently pumped across the membrane for each catalytic cycle.[136–138] The overall reaction is:

$$4\,e^- + 8\,H^+_{IN} + O_2 \longrightarrow 2\,H_2O + 4\,H^+_{OUT} \tag{18-5}$$

The reaction of O_2 with cytochrome *c* oxidase to form the oxygenated species A (Fig. 18-11) is very rapid, occurring with apparent lifetime τ (Eq. 9-5) of ~8–10 μs.[139] Study of such rapid reactions has depended upon a flow-flash technique developed by Greenwood and Gibson.[136,140,141] Fully reduced cytochrome oxidase is allowed to react with carbon monoxide, which binds to the iron in cytochrome a_3 just as does O_2. In fact, it was the spectroscopic observation that only half of the


A
31 Å
40 Å
79 Å
100 Å
31 Å
40 Å
79 Å
100 Å
B

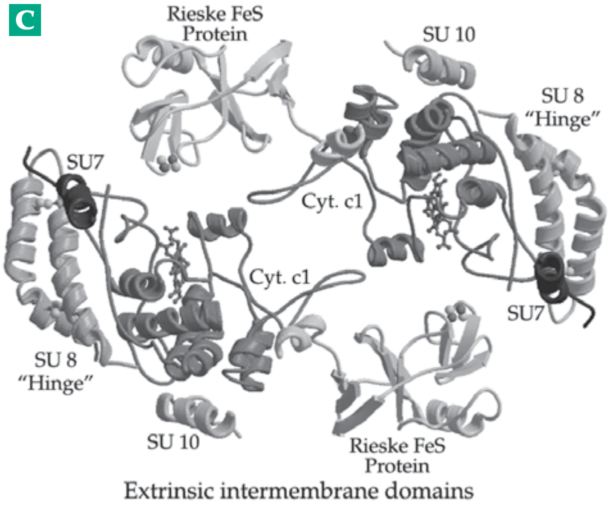

C
Rieske FeS Protein
SU 10
SU 8 "Hinge"
SU7
Cyt. c1
Cyt. c1
SU7
SU 8 "Hinge"
SU 10
Rieske FeS Protein
Extrinsic intermembrane domains

Figure 18-8 Stereoscopic ribbon diagrams of the chicken bc_1 complex (A) The native dimer. The molecular twofold axis runs vertically between the two monomers. Quinones, phospholipids, and detergent molecules are not shown for clarity. The presumed membrane bilayer is represented by a gray band. (B) Isolated close-up view of the two conformations of the Rieske protein (top and long helix at right) in contact with cytochrome *b* (below), with associated heme groups and bound inhibitors, stigmatellin, and antimycin. The isolated heme of cytochrome c_1 (left, above) is also shown. (C) Structure of the intermembrane (external surface) domains of the chicken bc_1 complex. This is viewed from within the membrane, with the transmembrane helices truncated at roughly the membrane surface. Ball-and-stick models represent the heme group of cytochrome c_1, the Rieske iron–sulfur cluster, and the disulfide cysteines of subunit 8. SU, subunit; cyt, cytochrome. From Zhang *et al.*[105]

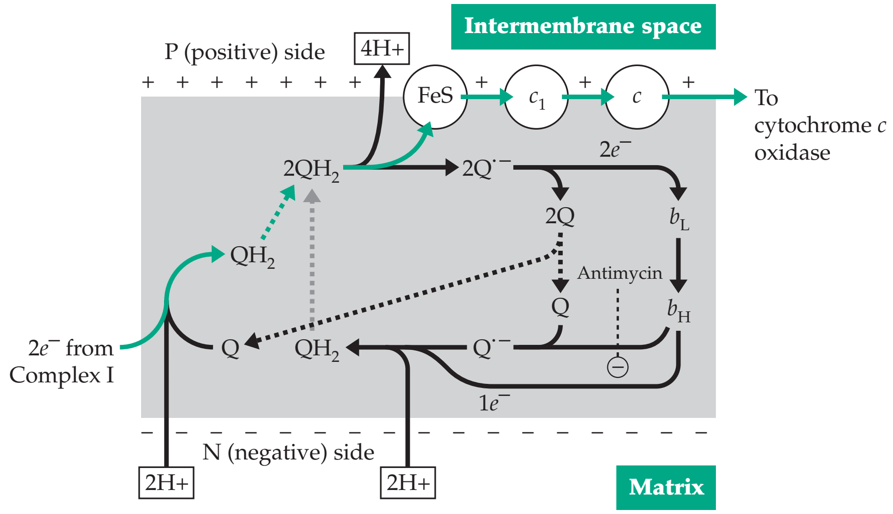
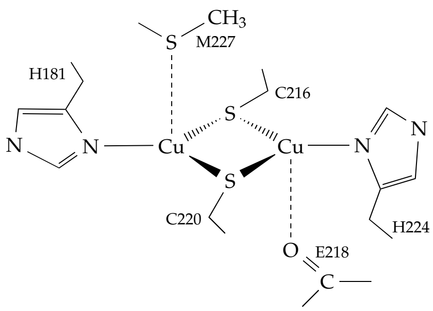
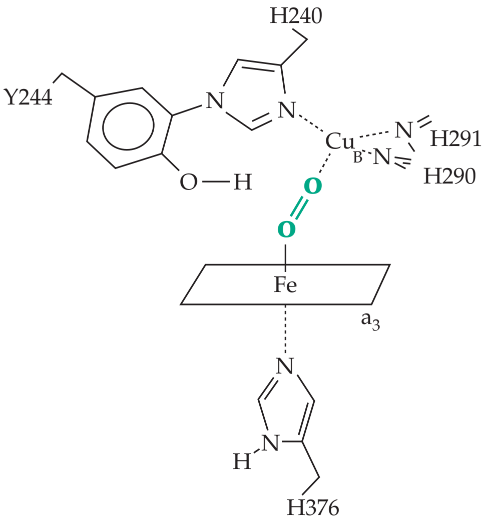


Figure 18-9 Proposed routes of electron transfer in mitochondrial complex III according to Peter Mitchell's Q cycle. Ubiquinone (Q) is reduced to QH_2 by complex I (left side of diagram) using two H^+ taken up from the matrix (leaving negative charges on the inner membrane surface). After diffusing across the bilayer (dashed line) the QH_2 is oxidized in the two steps with release of the two protons per QH_2 on the positive (P) side of the membrane. In the two-step oxidation via anionic radical Q^- one electron flows via the Rieske Fe–S protein and the cytochrome c_1 heme to external cytochrome *c*. The other electron is transferred to heme b_L of cytochrome *b*, then across the membrane to heme b_H which now reduces Q to Q^-. A second QH_2 is dehydrogenated in the same fashion and the electron passed through the cytochrome *b* centers is used to reduce Q^- to QH_2 with uptake from the matrix of 2 H^+. The resulting QH_2 diffuses back across the membrane to function again while the other Q diffuses back to complex I. The net result is pumping of 4 H^+ per 2 e^- passed through the complex. Notice that in the orientation used in this figure the matrix is at the bottom, not the top as in Figs. 18-4 and 18-5.



The Cu_A center of cytochrome oxidase



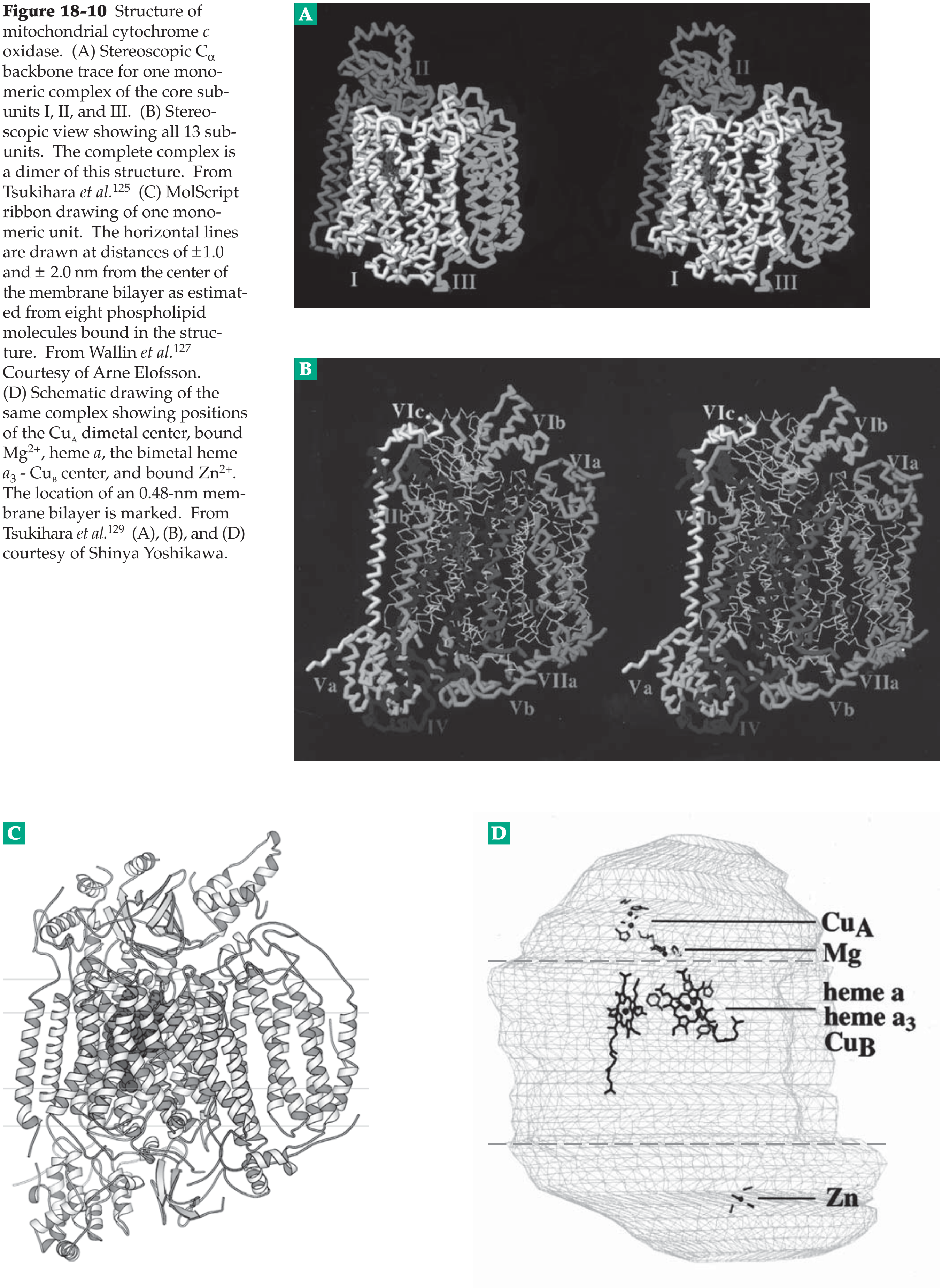
The Cu_B•A_3 center of cytochrome oxidase

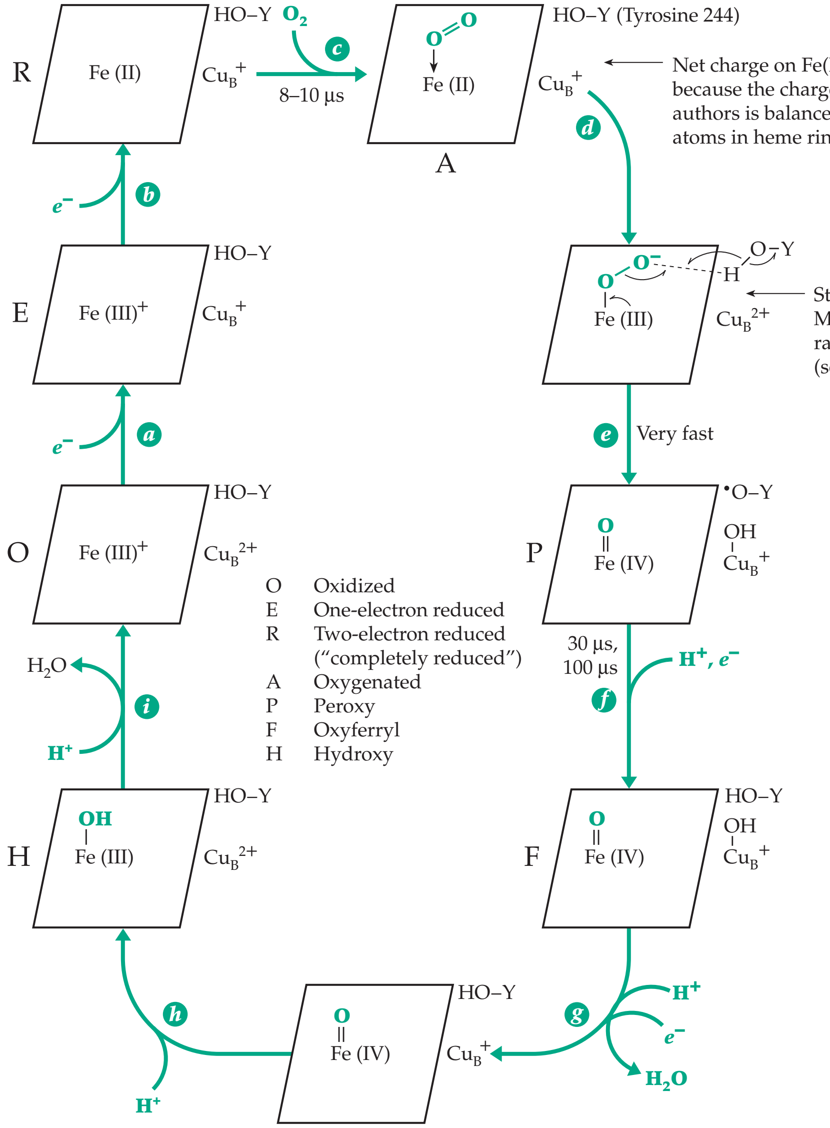
cytochrome *a* combined with CO that led Keilin to designate the reactive component a_3. This CO complex is mixed with O_2-containing buffer and irradiated with a laser pulse to release the CO and allow O_2 to react. The first rapid reaction observed is the binding of O_2 (step *c* in Fig. 18-11). Formation of a peroxy intermediate from the initial oxygenated form (A in Fig. 18-11) is very fast. The O–O bond of O_2 has already been cleaved in form P (Fig. 18-11), which has until recently been thought to be the peroxy intermediate. In fact, spectroscopic measurements indicate that form P contains an oxo-ferryl ion with the second oxygen of the original O_2 converted to an OH ion and probably coordinated with Cu_B.[136a,136c,142,142a–c] P may also contain an organic radical, perhaps formed from tyrosine 244 as indicated in Fig. 18-11.

A second relaxation time of τ = 32–45 μs has been assigned[139] to the conversion of the peroxide intermediate P to P′. A third relaxation time (τ = 100–140 μs) is associated with the oxidation of Cu_A by *a* (not shown in Fig. 18-11).[143] This electron transfer step limits the rate of step *f* of Fig. 18-11. Another reduction step with τ ~ 1.2 ms is apparently associated with electron transfer in step *h*. This slowest step still allows a first-order reaction rate of ~ 800 s^{-1}.

When O_2 reacts with cytochrome *c* oxidase, it may be bound initially to either the a_3 iron or to Cu_B, but in the peroxy intermediate P it may bind to both atoms. Oxyferryl compound F (Fig. 8-11) as well as radical species, can also be formed by treatment of the oxidized

Figure 18-10 Structure of mitochondrial cytochrome *c* oxidase. (A) Stereoscopic C_α backbone trace for one monomeric complex of the core subunits I, II, and III. (B) Stereoscopic view showing all 13 subunits. The complete complex is a dimer of this structure. From Tsukihara *et al.*[125] (C) MolScript ribbon drawing of one monomeric unit. The horizontal lines are drawn at distances of ±1.0 and ± 2.0 nm from the center of the membrane bilayer as estimated from eight phospholipid molecules bound in the structure. From Wallin *et al.*[127] Courtesy of Arne Elofsson. (D) Schematic drawing of the same complex showing positions of the Cu_A dimetal center, bound Mg^{2+}, heme *a*, the bimetal heme a_3 - Cu_B center, and bound Zn^{2+}. The location of an 0.48-nm membrane bilayer is marked. From Tsukihara *et al.*[129] (A), (B), and (D) courtesy of Shinya Yoshikawa.


A
II
I
III
II
I
III
B
VIc
VIb
VIa
VIIb
Va
VIIa
Vb
IV
VIc
VIb
VIa
VIIb
Va
VIIa
Vb
IV
C
D
CuA
Mg
heme a
heme a3
CuB
Zn


R
Fe (II)
HO–Y
Cu_B^+
O_2
c
8–10 µs
O=O
Fe (II)
HO–Y (Tyrosine 244)
Cu_B^+
A
d
e⁻
b
E
Fe (III)⁺
HO–Y
Cu_B^+
O
O⁻
Fe (III)
H
O–Y
Cu_B^{2+}
e⁻
a
e
Very fast
O
Fe (III)⁺
HO–Y
Cu_B^{2+}
P
O
Fe (IV)
•O–Y
OH
Cu_B^+
O Oxidized
E One-electron reduced
R Two-electron reduced (“completely reduced”)
A Oxygenated
P Peroxy
F Oxyferryl
H Hydroxy
H_2O
i
H⁺
30 µs, 100 µs
f
H⁺, e⁻
H
OH
Fe (III)
HO–Y
Cu_B^{2+}
F
O
Fe (IV)
HO–Y
OH
Cu_B^+
h
H⁺
O
Fe (IV)
HO–Y
Cu_B^+
g
H⁺
e⁻
H_2O

Net charge on Fe(II)a_3 is shown as zero because the charge 2⁺ shown by some authors is balanced by 2⁻ from nitrogen atoms in heme ring.

Structure originally proposed for form P. Movement of one electron to form tyrosinate radical is indicated by single-headed arrows (see p. 1030)

Figure 18-11 Possible catalytic cycle of cytochrome *c* oxidase at the cytochrome $a_3 – Cu_B$ site. The fully oxidized enzyme (O; left center) receives four electrons consecutively from the cyt *c* → Cu_A → cyt *a* chain. In steps *a* and *b* both heme *a* and Cu_B, as well as the Cu_A center and cyt a_3, are reduced to give the fully reduced enzyme (R). In the very fast step *c* the cyt a_3 heme becomes oxygenated and in step *d* is converted to a peroxide with oxidation of both the Fe and Cu. Intermediate P was formerly thought to be a peroxide but is now thought to contain ferryl iron and an organic radical. This radical is reduced by the third electron in step *f* to give the ferryl form F, with Cu^{2+} participating in the oxidation. The fourth electron reduces Cu_B again (step *g*) allowing reduction to the hydroxy form H in step *h*. Protonation to form H_2O (step *i*) completes the cycle which utilizes 4 e^- + 4 H^+ + O_2 to form 2 H_2O. Not shown is the additional pumping of 4 H^+ across the membrane from the matrix to the intermembrane space.

enzyme O with hydrogen peroxide.[143a-144] Use of various inhibitors has also been important in studying this enzyme. Cyanide, azide, and sulfide ions, as well as carbon monoxide, are powerful inhibitors. Cyanide specifically binds to the Fe^{3+} form of cytochrome a_3 preventing its reduction,[145] while CO competes with O_2 for its binding site. A much-used reagent that modifies carboxyl groups in proteins, and which inhibits many proton translocating proteins, is dicyclohexyl carbodiimide (Eq. 3-10).[146] The step-by-step flow of electrons through cytochrome *c* oxidase seems quite well defined. However, one of the most important aspects is unclear. How is the pumping of protons across the membrane coupled to electron transport?[137,138,142,147,147a] Many recent studies have employed directed mutation of residues in all four subunits to locate possible proton pathways or channels.[148–152] Most ideas involve movement through hydrogen bonded chains (Eq. 9-94), which may include the carboxylate groups of the bound hemes.[153] Conformational changes may be essential to the gating of proton flow by electron transfers.[143]

The surface of the matrix side of cytochrome oxidase contains histidine and aspartate side chains close together. It has been suggested that they form a proton collecting antenna that contains groups basic enough to extract protons from the buffered matrix and guide them to a proton conduction pathway.[154] Calcium ions also affect proton flow.[153a,b] We will return to this topic in Section C,3 (p. 1040).

C. Oxidative Phosphorylation

During the 1940s when it had become clear that formation of ATP from ADP and inorganic phosphate was coupled to electron transport in mitochondria, intensive efforts were made to discover the molecular mechanisms. However, nature sometimes strongly resists attempts to pry out her secrets, and the situation which prevailed was aptly summarized by Ephraim Racker: "Anyone who is not confused about oxidative phosphorylation just doesn't understand the situation."[155] The confusion is only now being resolved.

1. The Stoichiometry (P/O Ratio) and Sites of Oxidative Phosphorylation

Synthesis of ATP *in vitro* by tissue homogenates was demonstrated in 1937 by Kalckar, who has written a historical account.[156] In 1941, Ochoa[157] obtained the first reliable measurement of the P/O ratio, the *number of moles of ATP generated per atom of oxygen utilized* in respiration. The P/O ratio is also equal to the number of moles of ATP formed for each pair of electrons passing through an electron transport chain. Ochoa established that for the oxidation of pyruvate to acetyl-CoA and CO_2, with two electrons passed down the mitochondrial electron transport chain, the P/O ratio was ~3. This value has since been confirmed many times.[158–160] However, experimental difficulties in measuring the P/O ratio are numerous.[161] Many errors have been made, even in recent years, and some investigators[162] have contended that this ratio is closer to 2.5 than to 3. One method for measuring the P/O ratio is based on the method of determining the amount of ATP used that is described in the legend to Fig. 15-2.

The experimental observation of a P/O ratio of ~3 for oxidation of pyruvate and other substrates that donate NADH to the electron transport chain led to the concept that there are *three sites for generation of ATP*. It was soon shown that the P/O ratio was only 2 for oxidation of succinate. This suggested that one of the sites (site I) is located between NADH and ubiquinone and precedes the diffusion of QH_2 formed in the succinate pathway to complex III.

In 1949, Lehninger used ascorbate plus tetramethylphenylene-diamine (TMPD, Table 18-4) to introduce electrons into the chain at cytochrome *c*. The sequence ascorbate → TMPD → cytochrome *c* was shown to occur nonenzymatically. Later, it became possible to use cytochrome *c* as an electron donor directly. In either case only one ATP was generated, as would be anticipated if only site III were found to the right of cytochrome *c*. Site I was further localized by Lardy, who used hexacyanoferrate (III) (ferricyanide) as an artificial oxidant to oxidize NADH in the presence of antimycin *a*. Again a P/O ratio of one was observed. Finally, in 1955, Slater showed that passage of electrons from succinate to cytochrome c also gave only one ATP, the one generated at site II. The concept of three sites of ATP formation became generally accepted. However, as we shall see, these sites are actually proton-pumping sites, and there may be more than three of them.

Respiratory control and uncoupling. With proper care relatively undamaged mitochondria can be isolated. Such mitochondria are said to be **tightly coupled**. By this we mean that electrons cannot pass through the electron transport chain without generation of ATP. If the concentration of ADP or of P_i becomes too low, both phosphorylation and respiration cease. This **respiratory control** by ADP and P_i is a property of undamaged mitochondria. It may seem surprising that damaged mitochondria or submitochondrial particles are often able to transfer electrons at a faster rate than do undamaged mitochondria. However, electron transfer in damaged mitochondria occurs *without synthesis of ATP* and with no slowdown as the ADP concentration drops. A related kind of **uncoupling** of electron transport from ATP synthesis is brought about by various lipophilic anions called **uncouplers**, the best known of which is **2,4-dinitrophenol**. Even before the phenomenon of uncoupling was discovered, it had been known that dinitrophenol substantially increased the respiration rates of animals. The compound had even been used (with some fatal results) in weight control pills. The chemical basis of uncoupling will be considered in Section D.

"States" of mitochondria and spectrophotometric observation. Chance and Williams defined five **states** of tightly coupled mitochondria[60,163]; of these, states 3 and 4 are most often mentioned. If no oxidizable substrate or ADP is added the mitochondria have a very low rate of oxygen uptake and are in state 1. If oxidizable substrate and ADP are added rapid O_2 uptake is observed, the rate depending upon the rate of flow of electrons through the electron transport chain. This is state 3. As respiration occurs the coupled phosphorylation converts ADP into ATP, exhausting the ADP. Respiration slows to a very low value and the mitochondria are in state 4. If the substrate is present in excess, addition of more ADP will return the mitochondria to state 3.

Chance and associates employed spectrophotometry on intact mitochondria or submitochondrial particles to investigate both the sequence of carriers and the sites of phosphorylation. Using the dual wavelength spectrophotometer, the light absorption at the absorption maximum (λ_{max}) of a particular component was followed relative to the absorption at some other reference wavelength (λ_{ref}). The principal wavelengths used are given in Table 18-6. From these measurements the state of oxidation or reduction of each one of the carriers could be observed in the various states and in the presence of inhibitors. The

TABLE 18-6
Wavelengths of Light Used to Measure States of Oxidation of Carriers in the Electron Transport Chain of Mitochondria[a]

Carrier	λ_{max} (nm)[b]	λ_{ref} (nm)
NADH	340	374
Flavins	465	510
Cytochromes		
b^{2+}	564(α)	575
	530(β)	
	430(γ)	
c_1^{2+}	534(α)	
	523(β)	
	418(γ)	
c^{2+}	550(α)	540
	521(β)	
	416(γ)	
a^{2+}	605(α)	630(590)
	450(γ)[a]	
a_3^{2+}	600(α)[a]	
	445(γ)	455

[a] After Chance, B. and Williams, G. R. (1955) *J. Biol. Chem,* **217**, 409–427; (1956) Adv. Enzymol. **17**, 409–427.

[b] The wavelengths used for each carrier in dual wavelength spectroscopy appear opposite each other in the two columns. Some positions of other absorption bands of cytochromes are also given.

experiments served to establish that electrons passing down the chain do indeed reside for a certain length of time on particular carriers. That is, in a given state each carrier exists in a defined ratio of oxidized to reduced forms ([ox] / [red]). Such a result would not be seen if the entire chain functioned in a cooperative manner with electrons passing from the beginning to the end in a single reaction. By observing changes in the ratio [ox] / [red] under different conditions, some localization of the three phosphorylation sites could be made. In one experiment antimycin *a* was added to block the chain ahead of cytochrome c_1. Then tightly coupled mitochondria were allowed to go into state 4 by depletion of ADP. Since the concentration of oxygen was high and cytochrome a_3 has a low K_m for O_2 (~3 μM) cytochrome a_3 was in a highly oxidized state. Cytochrome *a* was also observed to be oxidized, while cytochrome c_1 and *c* remained reduced. The presence of this **crossover point** suggested at the time that cytochrome *c* might be at or near one of the "energy conservation sites." Accounts of more recent experiments using the same approach are given by Wilson *et al*.[164]

2. Thermodynamics and Reverse Electron Flow

From Table 6-8 the value of $\Delta G'$ for oxidation of one mole of NADH by oxygen (1 atm) is –219 kJ. At a pressure of ~10^{-2} atm O_2 in tissues the value is –213 kJ. However, when the reaction is coupled to the synthesis of three molecules of ATP ($\Delta G' = +34.5$ kJ mol^{-1}) the net Gibbs energy change for the overall reaction becomes $\Delta G' = -110$ kJ mol^{-1}. This is still very negative. However, we must remember that the concentrations of ATP, ADP, and P_i can depart greatly from the 1:1:1 ratio implied by the $\Delta G'$ value.

An interesting experiment is to allow oxidative phosphorylation to proceed until the mitochondria reach state 4 and to measure the **phosphorylation state ratio R_p**, which equals the value of [ATP] / [ADP][P_i] that is attained. This mass action ratio, which has also been called the "phosphorylation ratio" or "phosphorylation potential" (see Chapter 6 and Eq. 6-29), often reaches values greater than 10^4–10^5 M^{-1} in the cytosol.[164] An extrapolated value for a zero rate of ATP hydrolysis of log $R_p = 6.9$ was estimated. This corresponds (Eq. 6-29) to an increase in group transfer potential (ΔG of hydrolysis of ATP) of 39 kJ/mol. It follows that the overall value of ΔG for oxidation of NADH in the coupled electron transport chain is less negative than is $\Delta G'$. If synthesis of three molecules of ATP is coupled to electron transport, the system should reach an equilibrium when $R_p = 10^{6.4}$ at 25°C, the difference in ΔG and $\Delta G'$ being 3RT ln $Rp = 3 \times 5.708 \times 6.4 = 110$ kJ mol^{-1}. This value of *Rp* is, within experimental error, the same as the maximum value observed.[165] There apparently is an almost true equilibrium among NADH, O_2, and the adenylate system if the P/O ratio is 3.

Within more restricted parts of the chain it is possible to have *reversed electron flow*. Consider the passage of electrons from NADH, partway through the chain, and back out to fumarate, the oxidized form of the succinate–fumarate couple. The Gibbs energy change $\Delta G'$ (pH 7) for oxidation of NADH by fumarate is –67.7 kJ mol^{-1}. In uncoupled mitochondria electron flow would always be from NADH to fumarate. However, in tightly coupled mitochondria, in which ATP is being generated at site I, the overall value of $\Delta G'$ becomes much less negative. If $R_p = 10^4$ M^{-1}, $\Delta G'$ for the coupled process becomes approximately zero (–67.7 + 68 kJ mol^{-1}). Electron flow can easily be reversed so that succinate reduces NAD^+. Such ATP-driven reverse flow occurs under some physiological conditions within mitochondria of living cells, and some anaerobic bacteria generate all of their NADH by reversed electron flow (see Section E).

Another experiment involving equilibration with the electron transport chain is to measure the "observed potential" of a carrier in the chain as a function of the concentrations of ATP, ADP, and P_i. The observed potential E is obtained by measuring log([ox] / [red]) and applying Eq. 18-6 in which $E°'$ is the known midpoint potential of the couple (Table 6-8) and n is the number of electrons required to reduce one molecule of the carrier. If the system is equilibrated with a

$$E = \frac{-\Delta G}{nF} = E°' + \frac{0.0592}{n} \log \frac{[\text{ox}]}{[\text{red}]}$$
$$= \text{observed potential of carrier} \quad (18\text{-}6)$$

"redox buffer" (Chapter 6), E can be fixed at a preselected value. For example, a 1:1 mixture of succinate and fumarate would fix E at +0.03 V while the couple 3-hydroxybutyrate-acetoacetate in a 1:1 ratio would fix it at $E°' = -0.266$ V. Consider the potential of cytochrome b_{562} (b_H), which has an $E°'$ value of 0.030 V. Substituting this in Eq. 18-7 and using $E = -0.266$ V (as obtained by equilibration with 3-hydroxybutyrate–acetoacetate), it is easy to calculate that at equilibrium the ratio [ox] / [red] for cytochrome b_{562} is about 10^{-5}. In other words, in the absence of O_2 this cytochrome will be kept almost completely in the reduced form in an uncoupled mitochondrion.

However, if the electron transport between 3-hydroxybutyrate and cytochrome b_{562} is tightly coupled to the synthesis of one molecule of ATP, the observed potential of the carrier will be determined not only by the imposed potential E_i of the equilibrating system but also by the phosphorylation state ratio of the adenylate system (Eq. 18-7). Here $\Delta G'_{ATP}$ is the group transfer potential ($-\Delta G'$ of hydrolysis) of ATP at pH 7 (Table 6-6), and n' is the number of electrons passing through the chain required to synthesize one ATP. In the upper part of the equation n is the number of electrons required to reduce the carrier, namely one in the case of cytochrome b_{562}.

From Eq. 18-7 it is clear that in the presence of a high phosphorylation state ratio a significant fraction of cytochrome b_{562} may remain in the reduced form at equilibrium. Thus, if $R_p = 10^4$, if $E°'$ for cytochrome b_{562} is 0.030 V, if $n' = 2$, and the potential E is fixed at −0.25 V using the hydroxybutyrate–acetoacetate couple, we calculate, from Eq. 18-7, that the ratio [ox] / [red] for cytochrome b_{562} will be 1.75. Now, if R_p is varied the observed potential of the carrier should change as predicted by Eq. 18-7. This variation has been observed.[164] For a tenfold change in R_p the observed potential of cytochrome b_{562} changed by 0.030 V, just that predicted if $n' = 2$. On the other hand, the observed potential of cytochrome c varied by 0.059 V for every tenfold change in the ratio. This is just as expected if $n' = 2$, and if synthesis of two molecules of

TABLE 18-7
Electrode Potentials of Mitochondrial Electron Carriers and Gibbs Energy Changes Associated with Passage of Electrons[a]

	Electron carrier	$E°'$ (pH 7) isolated	$E°'$ (pH 7.2) in mitochondria	ΔG (kJ mol^{-1}) for 2 e^- flow to O_2 at 10^{-2} atm, carriers at pH 7
	NADH / NAD+	–0.320		–213
Group I	Flavoprotein		~ –0.30	
~ –0.30 V	Fe–S protein		~ –0.305	
	β-Hydroxybutyrate–acetoacetate	–0.266		–203
	Lactate–pyruvate	–0.185		–187
	Succinate-fumarate	0.031		–146
Group II	Flavoprotein		~ –0.045	
~ 0 V	Cytochrome b_T		–0.030	
	Cu		0.001	
	Fe–S protein		0.030	
	Cytochrome b_K		0.030	
	Ubiquinone	0.10	0.045	–132
	Cytochrome a_3 + ATP		0.155	
Group III	Cytochrome c_1		0.215	
	Cytochrome c	0.254	0.235	–102
	Cytochrome b_T + ATP		0.245	
	Cytochrome a	0.29	0.210	
	Cu		0.245	
	Fe–S protein		0.28	
Group IV	Cytochrome a_3		0.385	–77
	O_2 (10^{-2} atm)	0.785		0.00
	1 atm	0.815		

[a] Data from Wilson, D. F., Dutton, P. L., Erecinska, M., Lindsay, J. G., and Soto, N. (1972) *Acc. Chem. Res.* **5**, 234–241 and Wilson, D. F., Erecinska, M., and Sutoon, P. L. (1974) *Ann. Rev. Biophys. Bioeng.* **3**, 203–230.

$$\begin{aligned} E(\text{observed}) &= E°' + \frac{0.0592}{n}\log_{10}\frac{[\text{ox}]}{[\text{red}]} \\ &= E_i + \frac{\Delta G'_{\text{ATP}}}{96.5n'} + \frac{RT}{n'F}\ln\frac{[\text{ATP}]}{[\text{ADP}][\text{P}_\text{i}]} \\ &= E_i + \frac{0.358}{n'} + \frac{0.0592}{n'}\log_{10} R_\text{P} \end{aligned} \tag{18-7}$$

ATP is coupled to the electron transport to cytochrome *c*. Thus, we have experimental evidence that when one-electron carriers such as the cytochromes are involved, the passage of *two electrons* is required to synthesize one molecule of ATP. Furthermore, from experiments of this type it was concluded that the sites of phosphorylation were localized in or related to complexes I, III, and IV.

Another kind of experiment is to equilibrate the electron transport chain with an external redox pair of known potential using *uncoupled* mitochondria. The value of $E°'$ of a particular carrier can then be measured by observation of the ratio [ox] / [red] and applying Eq. 18-7. While changes in the equilibrating potential E will be reflected by changes in [ox] / [red] the value of $E°'$ will remain constant. The $E°'$ values of Fe–S proteins and copper atoms in the electron transport chain have been obtained by equilibrating mitochondria, then rapidly freezing them in liquid nitrogen, and observing the ratios [ox] / [red] by EPR at 77K (Table 18-7).

The values of $E°'$ of the mitochondrial carriers fall into four **isopotential groups** at ~–0.30, ~0, ~+0.22, and ~+0.39 V (Table 18-7). When tightly coupled mitochondria are allowed to go into state 4 (low ADP, high ATP, O_2 present but low respiration rate), the observed potentials change. That of the lowest isopotential group (which includes NAD^+ / NADH) falls to ~–0.38 V, corresponding to a high state of reduction of the carriers to the left of the first phosphorylation site in Fig. 18-4. Groups 2 and 3 remain close to their midpoint potentials at ~–0.05 and +0.26 V. In this condition the potential difference between each successive group of carriers amounts to ~0.32 V, just enough to balance the formation of one molecule of ATP for each two electrons passed at a ratio $R_\text{p} \approx 10^4$ M^{-1} (Eq. 18-7).

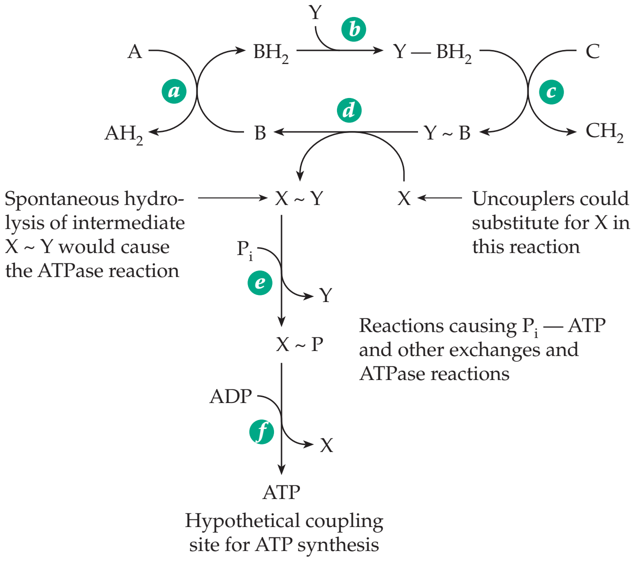
Two cytochromes show exceptional behavior and appear twice in Table 18-7. The midpoint potential $E°'$ of cytochrome b_{566} (b_L) changes from –0.030 V in the absence of ATP to +0.245 V in the presence of a high concentration of ATP. On the other hand, $E°'$ for cytochrome a_3 drops from +0.385 to 0.155 V in the presence of ATP. These shifts in potential must be related to the coupling of electron transport to phosphorylation.

3. The Mechanism of Oxidative Phosphorylation

It was natural to compare mitochondrial ATP synthesis with substrate-level phosphorylations, in which high-energy intermediates are generated *by the passage of electrons through the substrates*. The best known example is oxidation of the aldehyde group of glyceraldehyde 3-phosphate to an acyl phosphate, which, after transfer of the phospho group to ADP, becomes a carboxylate group (Fig. 15-6). The Gibbs energy of oxidation of the aldehyde to the carboxylate group provides the energy for the synthesis of ATP. However, this reaction differs from mitochondrial electron transport in that *the product, 3-phosphoglycerate, is not reconverted to glyceraldehyde 3-phosphate*. Electron carriers of the respiratory chain must be regenerated in some cyclic process. Because of this, it was difficult to imagine practical mechanisms for oxidative phosphorylation that could be related to those of substrate level phosphorylation. Nevertheless, many efforts were made over a period of several decades to find such high-energy intermediates.

Search for chemical intermediates. An early hypothetical model, proposed by Lipmann,[166] is shown in Fig. 18-12. Here A, B, and C are three electron carriers in the electron transport chain. Carrier C is a better oxidizing agent than B or A. Carrier B has some special chemistry that permits it, in the reduced state, to react with group Y of a protein (step *b*) to form $Y–BH_2$. The latter, an unidentified adduct, is converted by oxidation with carrier C (step *c*) to a "high energy" oxidized form indicated as Y ~ B. Once the possibility of generating such an intermediate is conceded, it is easy to imagine plausible ways in which the energy of this intermediate could be transferred into forms with which we are already familiar. For example, another protein X could react (step *d*) to form X ~ Y in which the X ~ Y linkage could be a thioester, an acyl phosphate, or other high-energy form. Furthermore, it might not be necessary to have two proteins; X and Y could be different functional groups of the same protein. They might be nonprotein components, e.g., Y might be a phospholipid.

Generation of ATP by the remaining reactions (steps *e* and *f* of Fig. 18-12) is straightforward. For example, if X ~ Y were a thioester the reactions would be the reverse of Eq. 12-48. These reaction steps would also be responsible for observed exchange reactions, for example, the mitochondrially catalyzed exchange of inorganic phosphate ($H^{32}PO_4^{2-}$) into the terminal position of ATP. Mitochondria and submitochondrial particles also contain ATP-hydrolyzing **(ATPase)** activity, which is thought to depend upon the same machinery that synthesizes ATP in tightly coupled mitochondria. In the scheme of Fig. 18-12, ATPase



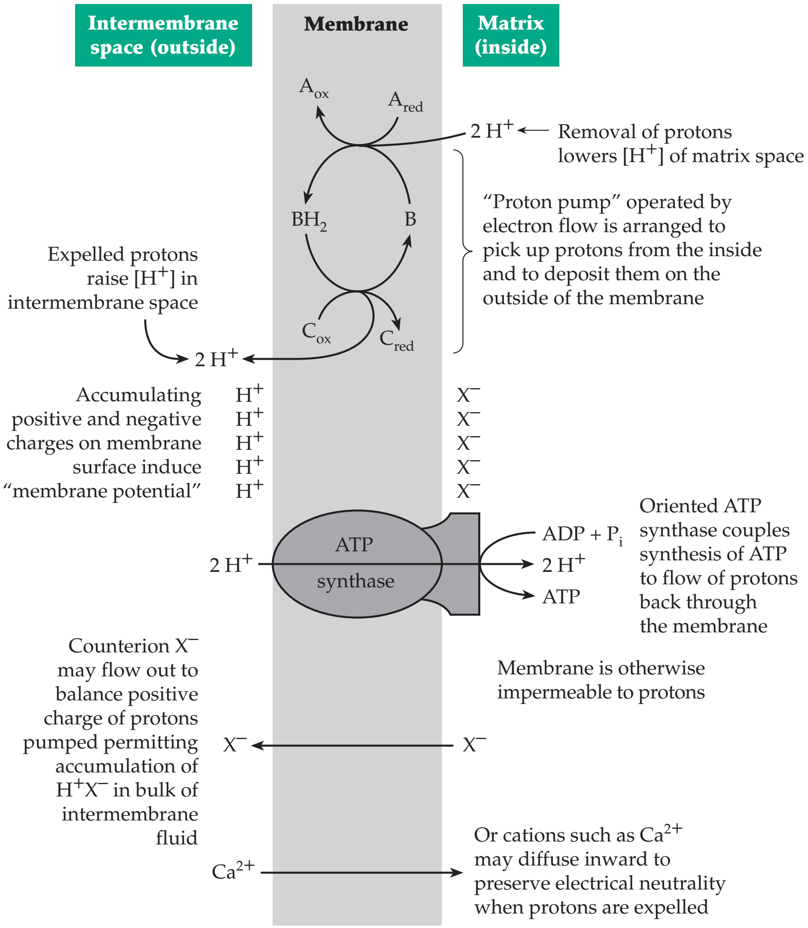
A
BH$_2$
Y
b
Y — BH$_2$
C
a
c
d
AH$_2$
B
Y ~ B
CH$_2$
Spontaneous hydrolysis of intermediate X ~ Y would cause the ATPase reaction
X ~ Y
X
Uncouplers could substitute for X in this reaction
P$_i$
e
Y
X ~ P
Reactions causing P$_i$ — ATP and other exchanges and ATPase reactions
ADP
f
X
ATP
Hypothetical coupling site for ATP synthesis

Figure 18-12 An early proposal for formation of ATP via "high-energy" chemical intermediates.

activity would be observed if hydrolysis of X ~ Y were to occur. Partial disruption of the system would lead to increased ATPase reactivity, as is observed. Uncouplers such as the dinitrophenolate ion or arsenate ion, acting as nucleophilic displacing groups, could substitute for a group such as X. Spontaneous breakdown of labile intermediates would permit oxidation to proceed unimpaired. Since there are three different sites of phosphorylation, we might expect to have three different enzymes of the type Y in the scheme of Fig. 18-12, but it would be necessary to have only one X.

In Lipmann's original scheme group Y was visualized as adding to a carbon–carbon double bond to initiate the sequence. Isotopic exchange reactions ruled out the possibility that either ADP or P$_i$ might serve as Y, but it was attractive to think that a bound phosphate ion, e.g., in a phospholipid or coenzyme, could be involved. Y ~ B of Fig. 18-12 would be similar in reactivity to an acyl phosphate or thioester. However, whatever the nature of Y ~ B, part of group Y would be left attached to B after the transfer of Y to X. For example, if Y were Y'OH compound X ~ OY' would be formed, and the carrier would be left in step *d* in the form of B–OH. Elimination of a hydroxyl group would be required to regenerate B. Perhaps nature has shunned this mechanism because there is no easy way to accomplish such an elimination. Many variations on the scheme of Fig. 18-12 were proposed,[166] and some were discussed in the first edition of this textbook.[167] However, as attractive as these ideas may have seemed, *all attempts to identify discrete intermediates that might represent X ~ Y failed. Furthermore, most claims to have seen Y ~ B by any means have been disproved*.

Peter Mitchell's chemiosmotic theory. To account for the inability to identify high energy intermediates as well as the apparent necessity for an intact membrane, Peter Mitchell, in 1961, offered his **chemiosmotic theory** of oxidative phosphorylation.[168–175a] aThis theory also accounts for the existence of **energy-linked processes** such as the accumulation of cations by mitochondria. The principal features of the Mitchell theory are illustrated in Fig. 18-13. Mitchell proposed



Intermembrane space (outside)
Membrane
Matrix (inside)
A$_{ox}$
A$_{red}$
2 H$^+$
Removal of protons lowers [H$^+$] of matrix space
BH$_2$
B
"Proton pump" operated by electron flow is arranged to pick up protons from the inside and to deposit them on the outside of the membrane
Expelled protons raise [H$^+$] in intermembrane space
C$_{ox}$
C$_{red}$
2 H$^+$
Accumulating positive and negative charges on membrane surface induce "membrane potential"
H$^+$
H$^+$
H$^+$
H$^+$
H$^+$
X$^-$
X$^-$
X$^-$
X$^-$
X$^-$
ATP synthase
2 H$^+$
ADP + P$_i$
2 H$^+$
ATP
Oriented ATP synthase couples synthesis of ATP to flow of protons back through the membrane
Membrane is otherwise impermeable to protons
Counterion X$^-$ may flow out to balance positive charge of protons pumped permitting accumulation of H$^+$X$^-$ in bulk of intermembrane fluid
X$^-$
X$^-$
Or cations such as Ca^{2+} may diffuse inward to preserve electrical neutrality when protons are expelled
Ca^{2+}

Figure 18-13 Principal features of Mitchell's chemiosmotic theory of oxidative phosphorylation.

that the inner membrane of the mitochondrion is a closed, proton-impermeable **coupling membrane**, which contains **proton pumps** operated by electron flow and which cause protons to be expelled through the membrane from the matrix space. As indicated in Fig. 18-13, an oxidized carrier B, upon reduction to BH_2, acquires two protons. These protons do not necessarily come from reduced carrier AH_2, and Mitchell proposed that they are picked up from the solvent on the matrix side of the membrane. Then, when BH_2 is reoxidized by carrier C, protons are released on the outside of the membrane. On the basis of existing data, *Mitchell assumed a stoichiometry of two protons expelled for each ATP synthesized*. It followed that there should be three different proton pumps in the electron transport chain corresponding to the three phosphorylation sites.

The postulated proton pumps would lead either to bulk accumulation of protons in the intermembrane space and cytoplasm, with a corresponding drop in pH, *or to an accumulation of protons along the membrane itself*. The latter would be expected if counterions X^- do not pass through the membrane with the protons. The result in such a case would be the development of a **membrane potential**, a phenomenon already well documented for nerve membranes (Chapter 8).

A fundamental postulate of the chemiosmotic theory is the presence of an oriented ATP synthase that utilizes the Gibbs energy difference of the proton gradient to drive the synthesis of ATP (Fig. 18-9). Since $\Delta G'$ (pH 7) for ATP synthesis is +34.5 kJ mol^{-1} and, if as was assumed by Mitchell, the passage of two protons through the ATP synthase is required to form one ATP, the necessary pH gradient (given by Eq. 6-25 or Eq. 18-9 with $E_m = 0$) would be $34.5/(2 \times 5.708) = 3.0$ pH units at 25°C. On the other hand, if the phosphorylation state ratio is $\sim 10^4\ M^{-1}$, the pH difference would have to be 5 units. Most investigators now think that 4 H^+ per ATP are needed by the synthase. If so, a pH difference of 2.5 units would be adequate. Various experiments have shown that passage of electrons does induce a pH difference, and that an artifically induced pH difference across mitochondrial membranes leads to ATP synthesis. However, pH gradients of the required size have not been observed. Nevertheless, if the membrane were charged as indicated in Fig. 18-13, without accumulation of protons in the bulk medium, a membrane potential would be developed, and this could drive the ATP synthase, just as would a proton gradient.

The mitochondrial membrane potential E_m (or $\Delta\psi$) is the potential difference measured across a membrane relative to a reference electrode present in the surrounding solution.[176] For both mitochondria and bacteria E_m normally has a negative value. The Gibbs energy change $\Delta\psi_H+$ for transfer of one mole of H^+ from the inside of the mitochondrion to the outside, against the concentration and potential gradients, is given by Eq. 18-8. This equation follows directly from Eqs. 6-25

$$\begin{aligned}\Delta G_H+ &= 2.303\ RT\ \Delta pH - E_m F\\ &= 5.708\ \Delta pH - 96.5\ E_m\ kJ/mol \text{ at } 25°C\\ &\text{where } \Delta pH = pH\ (inside) - pH\ (outside)\end{aligned} \tag{18-8}$$

and 6-63 with n = 1. The same information is conveyed in Eq. 18-9, which was proposed by Mitchell for what he calls the **total protonic potential difference Δ*p***.

$$\begin{aligned}\Delta p\ (volts) &= E_m(volts) - 2.303\ \frac{RT}{F}\ \Delta\ pH\\ \Delta p\ (mV) &= E_m(mV) - 59.2\ \Delta\ pH \text{ at } 25°C\\ E_m &= \Delta\psi\end{aligned} \tag{18-9}$$

Mitchell was struck by the parallel between the force and flow of electrons, which we call electricity, and the force and flow of protons, which he named **proticity**.[174] This led one headline writer in *Nature*[177] to describe Mitchell as "a man driven by proticity," but if Mitchell is right, as seems to be the case, we are all driven by proticity! Mitchell also talked about **protonmotive** processes and referred to Δp as the **protonmotive force**. Although it is a potential rather than a force, this latter name is a popular designation for Δp.

The reader should be aware that considerable confusion exists with respect to names and definitions.[176] For example, the ΔG_H+ of Eq. 18-8 can also be called the **proton electrochemical potential** $\Delta\mu_H+$, which is analogous to the chemical potential μ of an ion (Eq. 6-24) and has units of kJ/mol (Eq. 18-10).

$$\begin{aligned}-\Delta G_H+ = \Delta\mu_H+ &= F\ \Delta p\\ &= 96.5\ \Delta p\ kJ/mol \text{ at } 25°C\end{aligned} \tag{18-10}$$

However, many authors use $\Delta\mu_H+$ as identical to the protonmotive force Δp.

From Eq. 18-9 or Eq. 18-10 it can be seen that a membrane potential E_m of −296 mV at 25°C would be equivalent to a 5.0 unit change in pH and would be sufficient, if coupled to ATP synthesis via 2 H^+, to raise R_p to $10^4\ M^{-1}$. Any combination of ΔpH and E_m providing Δp of −296 mV would also suffice. If the ratio H^+/ATP = 4, Δp of −148 mV would suffice.

The chemiosmotic hypothesis had the great virtue of predicting the following consequences which could be tested: (1) electron-transport driven proton pumps with defined stoichiometries and (2) a separate ATP synthase, which could be driven by a pH gradient or membrane potential. Mitchell's hypothesis was initially greeted with skepticism but it encouraged many people, including Mitchell and his associate Jennifer Moyle, to test these predictions, which were soon found to be correct.[178]

Observed values of E_m and pH. One of the problems[179] in testing Mitchell's ideas has been the difficulty of reliably measuring Δp. To evaluate the pH term in Eq. 18-10 measurements have been made with microelectrodes and indicator dyes. However, the most reliable approach has been to observe the distribution of weak acids and bases across the mitochondrial membrane.[180] This is usually done with a suspension of freshly isolated active mitochondria. The method has been applied widely using, for example, methylamine. A newer method employs an isotope exchange procedure to measure the pH-sensitive carbonic anhydrase activity naturally present in mitochondria.[181]

The measurement of E_m ($\Delta\psi$) is also difficult.[179] Three methods have been used: (1) measurement with microelectrodes; (2) observation of fluorescent probes; (3) distribution of permeant ions. Microelectrodes inserted into mitochondria[182] have failed to detect a significant value for E_m. Fluorescent probes are not very reliable,[179,183] leaving the distribution of permeant ions the method of choice. In this method a mitochondrial suspension is exposed to an ion that can cross the membrane but which is not pumped or subject to other influences that would affect its distribution. Under such conditions the ion will be distributed according to Eq. 18-11. The most commonly used ions are K^+, the same ion that is thought to reflect the membrane potential of nerve axons (Chapter 30), or Rb^+. To make the inner mitochondrial membrane permeable to K^+, valinomycin (Fig. 8-22) is added. The membrane potential, with n = 1 in Eq. 9-1, becomes:

$$E_m = -59.2\ ([K^+]_{inside} / [K^+]_{outside})\ \text{volts} \tag{18-11}$$

In these experiments respiring mitochondria are observed to take up the K^+ or Rb^+ to give a high ratio of K^+ inside to that outside and consequently a negative E_m. There are problems inherent in the method. The introduction of a high concentration of ion perturbs the membrane potential, and there are uncertainties concerning the contribution of the Donnan equilibrium (Eq. 8-5) to the observed ion distribution.[184]

In most instances, either for mitochondrial suspensions or whole bacteria, ΔpH is less negative than −0.5 unit making a contribution of, at most, −30 mV to Δp. The exception is found in the thylakoid membranes of chloroplasts (Chapter 23) in which protons are pumped into the thylakoid vesicles and in which the internal pH falls dramatically upon illumination of the chloroplasts.[185] The ΔpH reaches a value of −3.0 or more units and Δp is ~180 mV, while E_m remains ~0. Reported values of E_m for mitochondria and bacteria range from −100 to −168 mV and Δp from −140 to −230 mV.[172,179] Wilson concluded that E_m for actively respiring mitochondria, using malate or glutamate as substrates, attains maximum (negative) values of E_m= −130 mV and Δp = −160 mV.[179] However, Tedeschi and associates[183,184] argued that E_m is nearly zero for liver mitochondria and seldom becomes more negative than −60 mV for any mitochondria.

A crucially important finding is that submitochondrial particles or vesicles from broken chloroplasts will synthesize ATP from ADP and P_i, when an artificial pH gradient is imposed.[172,186] Isolated purified F_1F_0 ATPase from a thermophilic *Bacillus* has been coreconstituted into liposomes with the light-driven proton pump **bacteriorhodopsin** (Chapter 23). Illumination induced ATP synthesis.[187] These observations support Mitchell's proposal that the ATP synthase is both spatially separate from the electron carriers in the membrane and utilizes the protonmotive force to make ATP. Thus, the passage of protons from the outside of the mitochondria back in through the ATP synthase induces the formation of ATP. What is the stoichiometry of this process?

It is very difficult to measure the flux of protons across the membrane either out of the mitochondria into the cytoplasm or from the cytoplasm through the ATP synthase into the mitochondria. Therefore, estimates of the stoichiometry have often been indirect. One argument is based on thermodynamics. If Δp attains values no more negative than −160 mV and R_p within mitochondria reaches at least $10^4\ M^{-1}$, we must couple ΔG_H of −15.4 kJ/ mol to ΔG of formation of ATP of +57.3 kJ/ mol. To do this four H^+ must be translocated per ATP formed. Recent experimental measurements with chloroplast ATP synthase[188] also favor four H^+. It is often proposed that one of these protons is used to pump ADP into the mitochondria via the ATP-ADP exchange carrier (Section D). Furthermore, if R_p reaches $10^6\ M^{-1}$ in the cytoplasm, it must exceed $10^4\ M^{-1}$ in the mitochondrial matrix.

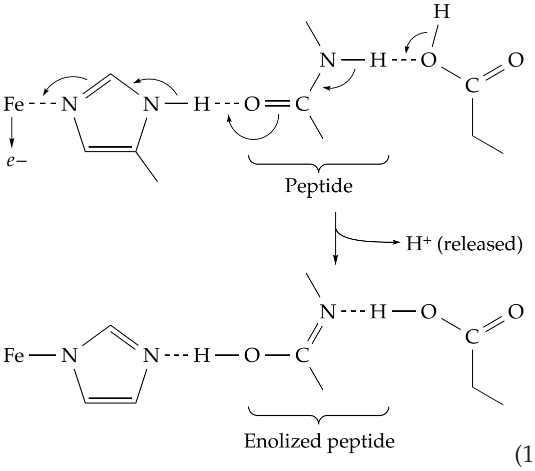
Proton pumps driven by electron transport. What is the nature of the proton-translocating pumps that link Δp with electron transport? In his earliest proposals Mitchell suggested that electron carriers, such as flavins and ubiquinones, each of which accepts two protons as well as two electrons upon reduction, could serve as the proton carriers. Each pump would consist of a pair of oxidoreductases. One, on the inside (matrix side) of the coupling membrane, would deliver two electrons (but no protons) to the carrier (B in Fig. 18-13). The two protons needed for the reduction would be taken from the solvent in the matrix. The second oxidoreductase would be located on the outside of the membrane and would accept two electrons from the reduced carrier (BH_2 in Fig. 18-13) leaving the two released protons on the outside of the membrane. To complete a "loop" that would allow the next carrier to be reduced, electrons would have to be transferred through fixed electron carriers embedded in the membrane from the reduced electron acceptor (C_{red} in Fig. 18-13) to the oxidized form of the oxidoreductase to be used as reductant for the next loop. These loops, located in complexes I and III of Fig. 18-5, would pump three protons per electron or six H^+/O. With a P/O ratio of three this would provide two H^+ per ATP formed. Mitchell regarded this stoichiometry as appropriate.

The flavin of NAD dehydrogenase was an obvious candidate for a carrier, as was ubiquinone. However, the third loop presented a problem. Mitchell's solution was the previously discussed **Q cycle**, which is shown in Fig. 18-9. This accomplishes the pumping in complex III of 2 H^+/e^-, the equivalent of two loops.[111] However, as we have seen, the magnitude of Δp suggests that 4 H^+, rather than 2 H^+, may be coupled to synthesis of one ATP. If this is true, mitochondria must pump 12 H^+/O rather than six when dehydrogenating NADH, or eight H^+/O when dehydrogenating succinate.

The stoichiometry of proton pumping was measured by Lehninger and associates using a fast-responding O_2 electrode and a glass pH electrode.[189,190] They observed an export of eight H^+/O for oxidation of succinate rat liver mitochondria in the presence of a permeant cation that would prevent the buildup of E_m, and four H^+/O (2 H^+/e^-) for the cytochrome oxidase system. These are equivalent to two H^+/e^- at each of sites II and III as is indicated in Fig. 18-4. Some others have found lower H^+/e^- ratios.

If two H^+/e^- are pumped out of mitochondria, where do we find the pumping sites? One possibility is that protons are pumped through the membrane by a **membrane Bohr effect**, so named for its similarity to the Bohr effect observed upon oxygenation of hemoglobin. In the latter case (Chapter 7), the pK_a values of certain imidazole and terminal amino groups are decreased when O_2 binds. This may result, in part, from an electrostatic effect of O_2 in inducing a partial positive charge in the heme. This partial charge may then cause a decrease in the pK_a values of nearby groups. Similarly, complete loss of an electronic charge from a heme group or an iron–sulfur protein in the electron transport chain would leave a positive charge, an electron "hole," which could induce a large change in the pK_a of a neighboring group. One manifestation of this phenomenon may be a strong pH dependence of the reduction potential (Eq. 16-19).

Protons that could logically be involved in a membrane Bohr effect are those present on imidazole rings coordinated to Fe or Cu in redox proteins. Removal of an electron from the metal ion could be accompanied by displacement of electrons within the imidazole, within a peptide group that is hydrogen-bonded to an imidazole, or within some other acidic group. A hypothetical example is illustrated in Eq. 18-12 in which a carboxyl group loses a proton when "handed" a second. If the transiently enolized peptide linkage formed in this process is tautomerized back to its original state before the iron is reduced again, the proton originally present on the carboxyl group will be released. It is easy to imagine that a proton could then be "ferried" in (as in Eq. 9-96) from the opposite side of the membrane to reprotonate the imidazole group and complete the pumping process.



(18-12)

In view of the large number of metal-containing electron carriers in the mitochondrial chain, there are many possible locations for proton pumps. However, the presence of the three isopotential groups of Table 18-7 suggests that the pumps are clustered in complexes I-III as pictured in Fig. 18-5. One site of pumping is known to be in the cytochrome *c* oxidase complex. When reconstituted into phospholipid, the purified complex does pump protons in response to electron transport, H^+/e^- ratios of ~1 being observed.[136,137,147,191] As mentioned in Section B,3 a large amount of experimental effort has been devoted to identifying proton transport pathways in cytochrome *c* oxidase and also in the cytochrome bc_1 (complex II).[192] Proton pumping appears to be coupled to chemical changes occurring between intermediates P and F of Fig. 18-11, between F and O,[136,193] and possibly between O and R.[137,138] Mechanisms involving direct coupling of chemical changes at the A_3Cu_B center and at the Cu_A dimetal center have been proposed.[147,194]

How do protons move from the pumping sites to ATP synthase molecules? Since protons, as H_3O^+, are sufficiently mobile, ordinary diffusion may suffice. Because of the membrane potential they will tend to stay close to the membrane surface, perhaps being transported on phosphatidylethanolamine head groups (see Chapter 8). According to the view of R. J. P. Williams protons are not translocated across the entire membrane by the proton pumps, but flow through the proteins of the membrane to the ATP synthase.[195] There the protons induce the necessary conformational changes to cause ATP synthesis. A related idea is that transient high-energy intermediates generated by electron transport within membranes are proton-carrying conformational isomers. When an electron is removed from an electron-transporting metalloprotein, the resulting positively charged "hole" could be stable for some short time, while the protein diffused within the membrane until it encountered an F_0 protein of an ATP synthase. Then it might undergo an induced conformational change at the same time that it "handed" the Bohr effect proton of Eq. 18-14 to the F_0 protein and simultaneously induced a conformational change in that protein. The coupling of proton transport to conformational changes seems plausible, when we recall that the induction of conformational changes within proteins almost certainly involves rearrangement of hydrogen bonds.

A consequence of the chemiosmotic theory is that there is no need for an integral stoichiometry between protons pumped and ATP formed or for an integral P/O ratio. There are bound to be inefficiencies in coupling, and Δp is also used in ways other than synthesis of ATP.

4. ATP Synthase

In 1960, Racker and associates[196,197] discovered that the "knobs" or "little mushrooms" visible in negatively stained mitochondrial fragments or fragments of bacterial membranes possess ATP-hydrolyzing (**ATPase)** activity. Earlier the knob protein had been recognized as one of several **coupling factors** required for reconstitution of oxidative phosphorylation by submitochondrial particles.[197] Electron micrographs showed that the submitochondrial particles consist of closed vesicles derived from the mitochondrial cristae, and that the knobs (Fig. 18-14A) are on the *outside* of the vesicles. They can be shaken loose by ultrasonic oscillation with loss of phosphorylation and can be added back with restoration of phosphorylation. The knob protein became known as **coupling factor F_1**. Similar knobs present on the outside of the thylakoids became **CF_1**and those inside thermophilic bacteria **TF_1**. The ATPase activity of F_1 was a clue that *the knobs were really ATP synthase*. It also became clear that a portion of the ATP synthase is firmly embedded in the membranes. This part became known as **F_0**. Both the names F_1F_0 ATP synthase and F_1F_0 ATPase are applied to the complex, the two names describing different catalytic activities. The ATPase activity is usually not coupled to proton pumping but is a readily measurable property of the F_1 portion. In a well-coupled submitochondrial particle the ATPase activity will be coupled to proton transport and will represent a reversal of the ATP synthase activity.

The synthase structure. The F_1 complex has been isolated from *E. coli*,[204,205] other bacteria,[206,207] yeast,[208a,b] animal tissues,[199,209–211] and chloroplasts.[212–214] In every case it consists of five kinds of subunits with the stoichiometry $\alpha_3\beta_3\gamma\delta\varepsilon$.[214a,b] The F_0 complex of *E. coli* contains three subunits designated a,b, and c. All of these proteins are encoded in one gene cluster, the *unc* operon (named for uncoupled mutants), with the following order:

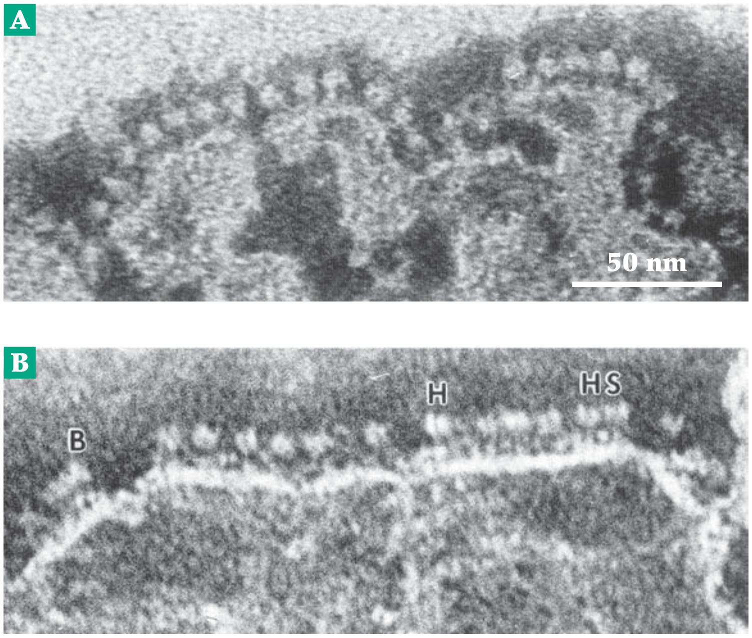
I B E F H A G D C – – – Gene symbol
i a c b δ α γ β ε – – – Subunit symbol

Here I is the regulatory gene (as in Fig. 28-1). The *E. coli* F_0 appears to have approximately the unusual stoichiometry $ab_2c_{9\text{-}11}$. This suggested the possibility that 12 c subunits form a ring with D_6 or D_{12} symmetry, the latter being illustrated in the structural proposal shown in Fig. 18-14E. However, crystallographic evidence suggests that there may be 10, not 12 subunits.[214c]

Mitochondrial ATP synthase of yeast contains at least 13 different kinds of subunits[208] and that of animals[215] 16, twice as many as in *E. coli*. Subunits α, β, γ, a, b, and c of the mitochondrial synthase correspond to those of *E. coli*. However, the mitochondrial homolog of *E. coli* δ is called the **oligomycin-sensitivity-conferring protein** (OSCP).[216–218] It makes the ATPase activity sensitive to oligomycin. The mitochondrial δ subunit corresponds to ε of *E. coli* or of chloroplasts.[217,219] Mitochondrial ε has no counterpart in bacteria.[209,220] In addition,[209,215] mitochondria contain subunits called d, e, f, g, A6L, F6, and IF_1, the last being an 84-residue inhibitor, a regulatory subunit.[221] The subunits of yeast ATP synthase correspond to those of the animal mitochondrial synthase but include one additional protein (**h**).[208a]

	F_1							F_0								
E. coli	α	β	γ	δ	ε			a	b	c						
Mitochondria	α	β	γ	OSCP	δ	ε	IF_1	a	b	c	d	e	f	g	A6L	F6

Six of the relatively large (50–57 kDa) α and β subunits associate to an $\alpha_3\beta_3$ complex that constitutes the knobs.[202,210] Chemical crosslinking, directed mutation, electron cryomicroscopy,[222,222a] and high-resolution X-ray diffraction measurements[199,207,211,223,224] have established that the α and β subunits alternate in a quasisymmetric cyclic head that contains active sites for ATP formation in the three β subunits (Fig. 18-14C–E). The α subunits also contain ATP-binding sites, but they are catalytically inactive, and their bound MgATP does not exchange readily with external ATP and can be replaced by the nonhydrolyzable AMP-PNP (Fig. 12-31) with retention of activity. The $\alpha_3\beta_3$ complex is associated with the F_0 part by a slender **central stalk**



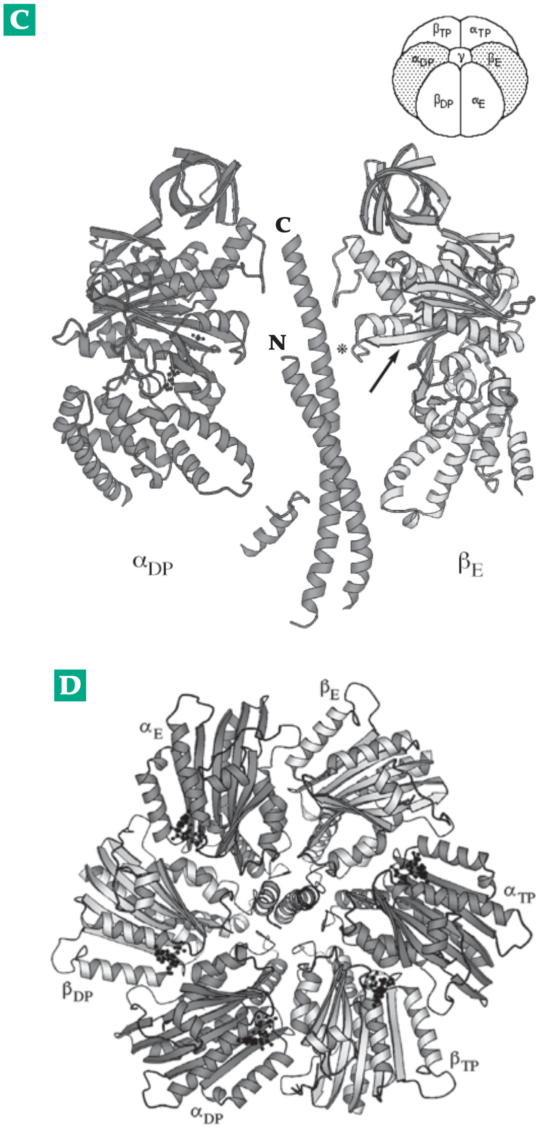
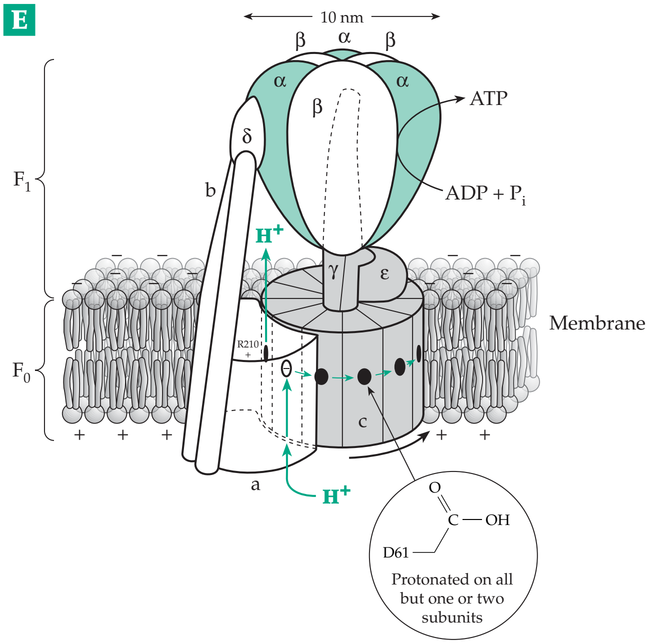


Figure 18-14 ATP synthase and vacuolar ATPase. (A) "Knobs" of ATP synthase on mitochondrial membranes negatively stained with phosphotungstate. (B) Vacuolar proton-pumping ATPase from an intact vacuolar membrane stained in the same way. Some images have been marked to indicate well-resolved head groups (H), stalks (S), and basal components (B). (A) and (B) are from Dschida and Bowman.[198] Courtesy of Barry J. Bowman. (C) Ribbon model of the atomic structure of the F_1 part of bovine heart mitochondrial ATP synthase. This section through the knob, which is drawn schematically at the upper right, shows one α subunit, containing bound ADP, α_{DP}, to the left and one empty β subunit, β_E, to the right. In the center are the N and C termini of subunit γ. The arrow points to a disruption in the β sheet of the subunit structure in the β_E conformation. The asterisk marks a loop that would collide with the C-terminal part of subunit γ if the latter were rotated. (D) View of the F_1 ATP synthase from the membrane side. The section shown contains the nucleotide binding sites. Subunits with empty sites are labeled α_E and β_E. Those with bound ADP are labeled α_{DP} and β_{DP}, and those with bound ATP α_{TP} and β_{TP}. Deviation from perfect threefold symmetry can be seen in this view. (C) and (D) are from Abrahams *et al.*[199] Courtesy of John E. Walker. (E) A recent model of the *E. coli* ATP synthase. The $\alpha_3\beta_3$ head contains the ATP-synthesizing sites in the β subunits. The $\alpha_3\beta_3$ complex and also subunits a, b, and δ remain stationary and form the stator for a protic engine, whose rotor consists of 12 c subunits and attached γ and ε subunits. Rotation is induced by the membrane potential or difference in H^+ activity on the two sides of the membrane. The carboxylate group of Asp 61 in each c subunit must be protonated to –COOH in order for it to move away from the entrance channel in subunit a. The presence of the positively charged Arg 210 near the exit channel in subunit a induces release of the proton when the c subunit has rotated almost 360°. According to this picture, 12 protons are required for one rotation with synthesis of three molecules of ATP. After Elston *et al.*[200] and Zhou *et al.*[201] See also Junge *et al.*[202] and Engelbrecht and Junge.[203]



or **shaft**. Much effort has gone into establishing the subunit composition of the shaft and the F_0 parts of the structure. As is indicated in Fig. 18-14E, subunits γ and ε of the *E. coli* enzyme are both part of the central shaft.[219,225,226] The same is true for the mitochondrial complex, in which the δ subunit corresponds to bacterial ε.[227] The role of this subunit is uncertain. It is part of the shaft but is able to undergo conformational alterations that can permit its C-terminal portion to interact either with F_0 or with the $\alpha_3\beta_3$ head.[227,227a] The unique ε subunit of mitochondrial ATPase appears also to be part of the shaft.[220]

The most prominent component of the central shaft is the 270-residue subunit γ, which associates loosely with the $\alpha_3\beta_3$ head complex but more tightly with F_0. About 40 residues at the N terminus and 60 at the C terminus form an α-helical coiled coil, which is visible in Fig. 18-14E[199,211] and which protrudes into the central cavity of the $\alpha_3\beta_3$ complex. Because it is asymmetric, the γ subunit apparently acts as a rotating camshaft to physically alter the α and β subunits in a cyclic manner. Asymmetries are visible in Fig. 18-14D.[211] The central part of subunit γ forms a more globular structure, which bonds with the c subunits of F_0.[205] Exact structures are not yet clear.

The δ subunit of *E. coli* ATP synthase (OSCP of mitochondria) was long regarded as part of the central stalk. However, more recent results indicate that it is found in a **second stalk**, which joins the $\alpha_3\beta_3$ complex to F_0. The central stalk rotates, relative to the second stalk. The second stalk may be regarded as stationary and part of a **stator** for a protic engine.[228,229] This stalk has been identified[230] in electron micrographs of chloroplast F_1F_0 and by crosslinking studies. As is depicted in Fig. 18-14E, a major portion of the second stalk is formed by two molecules of subunit b. Recent results indicate that bacterial subunit δ (mitochondrial OSCP) extends further up than is shown in Fig. 18-14, and together with subunit F6 may form a cap at the top of the $\alpha_3\beta_3$ head.[230a, 230b]

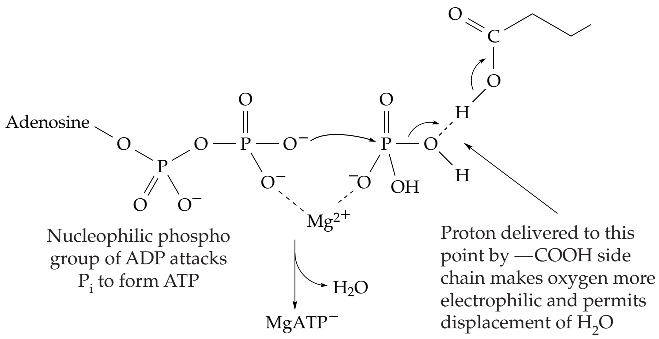
The F_0 portion of bacterial ATP synthase, which is embedded in the membrane, consists of one 271-residue subunit a, an integral membrane protein probably with five transmembrane helices,[231,232] two 156-residue b subunits, and ~ twelve 79-residue c subunits. The c subunit is a proteolipid, insoluble in water but soluble in some organic solvents. The structure of monomeric c in chloroform:methanol:H_2O (4:4:1) solution has been determined by NMR spectroscopy. It is a hairpin consisting of two antiparallel α helices.[233] Twelve of the c subunits are thought to assemble into a ring with both the N and C termini of the subunit chains in the periplasmic (or intermembrane) face of the membrane.[234,235] The ratio of c to a subunits has been difficult to measure but has been estimated as 9–12. The fact that both genetically fused c_2 dimers and c_3 trimers form function F_0 suggested that they assemble to a c_{12} ring as shown in Fig. 18-14E.[236] However, the recent crystallographic results that revealed a C_{10} ring[214c] raise questions about stoichiometry.

Since ATP synthesis takes place in F_1, it has long been thought that the F_0 part of the ATP synthase contains a "proton channel," which leads from the inside of the mitochondria to the F_1 assemblage.[146] Such a channel would probably not be an open pore but a chain of hydrogen-bonded groups, perhaps leading through the interior of the protein and able to transfer protons via icelike conduction. One residue in the c subunit, Asp 61, which lies in the center of the second of the predicted transmembrane helices, is critical for proton transport.[236a] Natural or artificial mutants at this position (e.g., D61G or D61N) do not transport protons. This carboxyl group also has an unusually high reactivity and specificity toward the protein-modifying reagent dicyclohexylcarbodiimide (DCCD; see Eq. 3-10).[146,237] Modification of a single c subunit with DCCD blocks the proton conductance.

An interesting mutation is replacement of alanine 62 of the c subunit with serine. This mutant will support ATP synthase using Li^+ instead of H^+.[237] Certain alkylophilic bacteria, such as *Propionigenium modestum*, have an ATP synthase that utilizes the membrane potential and a flow of Na^+ ions rather than protons through the c subunits.[238–240c] The sodium transport requires glutamate 65, which fulfills the same role as D61 in *E. coli*, and also Q32 and S66. Study of mutants revealed that the polar side chains of all three of these residues bind Na^+, that E65 and S66 are needed to bind Li^+, and that only E65 is needed for function with H^+.

The a subunit is also essential for proton translocation.[231,232,241] Structural work on this extremely hydrophobic protein has been difficult, but many mutant forms have been studied. Arginine 210 is essential as are E219 and H245. However, if Q252 is mutated to glutamate, E219 is no longer essential.[241] One of the OXPHOS diseases (NARP; Box 18-B) is a result of a leucine-to-arginine mutation in human subunit a.[241a] The b subunit is an elongated dimer, largely of α-helical structure.[242,243] Its hydrophobic N terminus is embedded in the membrane,[229,244] while the hydrophilic C-terminal region interacts with subunit σ of F_1, in the stator structure (Fig. 18-14E). Some of the F_0 subunits (d, e, f, g, A6L) may form a collar around the lower end of the central stalk.[210a,b]

How is ATP made? No covalent intermediates have been identified, and isotopic exchange studies indicate a direct dehydration of ADP and P_i to form bound ATP.[245] For the nucleophilic terminal phospho group of ADP to generate a high-energy linkage directly by attack on the phosphorus atom of P_i an OH^- ion must be eliminated (Eq. 18-13). This is not a probable reaction at pH 7, but it would be reasonable at low pH. Thus, one function of the oriented ATP synthase might be to deliver one or more protons flowing in from F_0 specifically to the oxygen that is to be eliminated (Eq. 18-13). As we have seen (Section 2), on thermodynamic grounds 3–4 protons would probably be needed. Perhaps they could be positioned nearby to exert a large electrostatic effect, or they could assist in releasing the ATP formed from the synthase by inducing a conformational change. However, it isn't clear how protons could be directed to the proper spots.



Nucleophilic phospho group of ADP attacks P_i to form ATP

Proton delivered to this point by —COOH side chain makes oxygen more electrophilic and permits displacement of H_2O

(18-13)

Paul Boyer's binding change mechanism. Boyer and associates suggested that ATP synthesis occurs rapidly and reversibly in a closed active site of the ATP synthase in an environment that is essentially anhydrous. ATP would then be released by an energy-dependent conformational change in the protein.[245–249] Oxygen isotope exchange studies verified that a rapid interconversion of bound ADP, P_i, and ATP does occur. Studies of soluble ATP synthase, which is necessarily uncoupled from electron transport or proton flow, shows that ATP is exceedingly tightly bound to F_1 as expected by Boyer's mechanism.[248] According to his **conformational coupling** idea, protons flowing across the membrane into the ATP synthase would in some way induce the conformational change necessary for release of ATP.

The idea of conformational coupling of ATP synthesis and electron transport is especially attractive when we recall that ATP is used in muscle to carry out mechanical work. Here we have the hydrolysis of ATP coupled to motion in the protein components of the muscle. It seems reasonable that ATP should be formed as a result of motion induced in the protein components of the ATPase. Support for this analogy has come from close structural similarities of the F_1 ATPase β subunits and of the active site of ATP cleavage in the muscle protein myosin (Chapter 19).

A simple version of Boyer's binding change mechanism is shown in Figure 18-15. The three F_1 β subunits are depicted in three different conformations. In O the active site is open, in T it is closed, and if ATP is present in the active site it is tightly bound. In the low affinity L conformation ligands are bound weakly. In step *a* MgADP and P_i enter the L site while MgATP is still present in the T site. In step *b* a protonic-energy-dependent step causes synchronous conformational changes in all of the subunits. The tight site opens and MgATP is free to leave. At the same time MgADP and P_i in the T site are converted spontaneously to tightly bound ATP. The MgATP is in reversible equilibrium with MgADP + P_i, which must be bound less tightly than is MgATP. That is, the high positive value of $\Delta G'$ for formation of ATP must be balanced by a corresponding negative $\Delta G'$ for a conformational or electronic reorganization of the protein in the T conformation. Opening of the active site in step *b* of Fig. 18-15 will have a high positive $\Delta G'$ unless it is coupled to proton flow throught F_0. Of three sites in the subunits, one binds MgATP very tightly ($K_d \sim 0.1\ \mu M$) while the other sites bind less tightly ($K_d \sim 20\ \mu M$).[250,251] However, it has been very difficult to establish binding constants or K_m values for the ATPase reaction.[248] Each of the three β sites probably, in turn, becomes the high-affinity site, consistent with an ATP synthase mechanism involving protein conformational changes.

Rotational catalysis. Boyer suggested that there is a cyclic rotation in the conformations of the three β subunits of the ATP synthase, and that this might involve rotation about the stalk. By 1984, it had been shown that bacterial flagella are rotated by a protonic motor (Chapter 19), and a protic rotor for ATP synthase had been proposed by Cox *et al.*[252] and others.[245] However, the b subunits were thought to be in the central stalk.[222] More recently chemical crosslinking experiments,[201,253] as well as electron microscopy, confirmed the conclusion that an intact stator structure must also be present as in Fig. 18-14E.[202] The necessary second stalk is visible in CF_1F_0 ATPase of chloroplasts[230] and also in the related vacuolar ATPase, a proton or Na^+ pump from a clostridium.[254] See also Section 5. Another technique, **polarized absorption recovery after photobleaching**, was applied after labeling of Cys 322, the penultimate residue at the C terminus of the γ subunit with the dye eosin. After photobleaching with a laser beam the polarization of the light absorption by the dye molecule relaxed because of rotation. Relaxation was observed when ATP was added but not with ADPPNP.[202,255,256]

The most compelling experiments were performed by Noji *et al.*[202,257–260] They prepared the $\alpha_3\beta_3\gamma$ subcomplex of ATPase from a thermophilic bacterium. The complex was produced in *E. coli* cells from the cloned genes allowing for some "engineering" of the proteins. A ten-histidine "tag" was added at the N termini of the β subunits so that the complex could be "glued" to a microscope coverslip coated with a nickel complex with a high affinity for the His tags. The γ subunit shafts protrude upward as shown in Fig. 18-16. The γ subunit was mutated to replace its

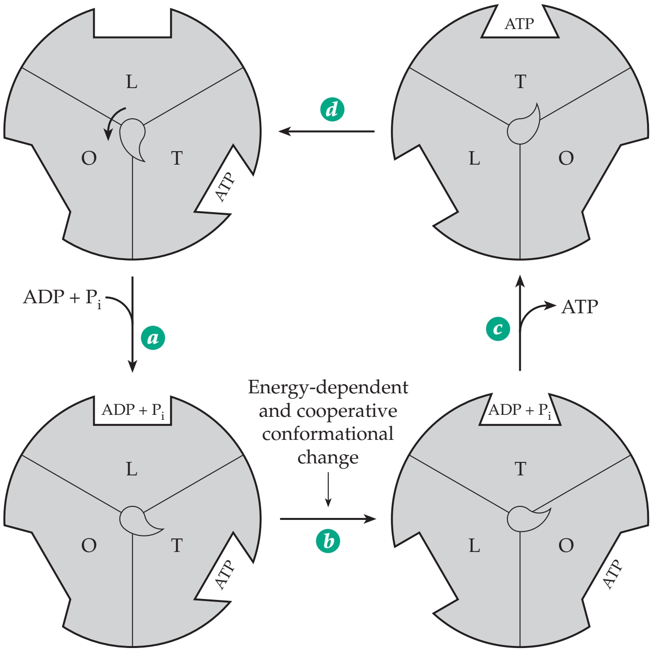


Figure 18-15 Boyer's binding change mechanism for ATP synthase in a simple form. After Boyer[245] but modified to include a central camshaft which may drive a cyclic alteration in conformations of the subunits. The small "pointer" on this shaft is not to be imagined as real but is only an indicator of rotation with induced conformational changes. The rotation could occur in 120° steps rather than the smaller steps suggested here.

only cysteine by serine and to introduce a cysteine in place of Ser 107 of the stalk region of γ. The new cysteine was biotinylated and attached to streptavidin (see Box 14-B) which was also attached to a fluorescently labeled actin filament (Fig. 7-10) ~1–3 µm in length as shown in Fig. 18-16. The actin fiber rotated in a counterclockwise direction when ATP was added but did not rotate with AMPPNP. At low ATP concentrations the rotation could be seen to occur in discrete 120° steps.[258,261,262] Each 120° step seems to consist of ~90° and ~30° substeps, each requiring a fraction of a millisecond.[262a] The ATPase appears to be acting as a **stepper motor**, hydrolysis of a single ATP turning the shaft 120°. Rotation at a rate of ~14 revolutions per second would require the hydrolysis of ~42 ATP per second. If the motor were attached to the F_0 part it would presumably pump four (or perhaps three) H^+ across a membrane for each ATP hydrolyzed. Acting in reverse, it would make ATP. A modification of the experiment of Fig. 18-16 was used to demonstrate that the c subunits also rotate with respect to the $\alpha_3\beta_3$ head.[262b] Other experiments support rotation of the c ring relative to subunit a.[262c,d]

Still to be answered are important questions. How does ATP hydrolysis turn the shaft? Are four H^+ pumped for each step, or are there smaller single proton substeps? Is the simple picture in Fig. 18-15 correct or, as proposed by some investigators,[263–265] must all three β subunits be occupied for maximum catalytic activity?[266] How is the coupling of H^+ transport to mechanical motion accomplished?[267,267a–d]

5. ATP-driven Proton Pumps

Not all proton pumps are driven by electron transport. ATP synthase is reversible, and if Δp is low, hydrolysis of ATP can pump protons out of mitochondria or across bacterial plasma membranes.[268] Cells of *Streptococcus faecalis*, which have no respiratory chain

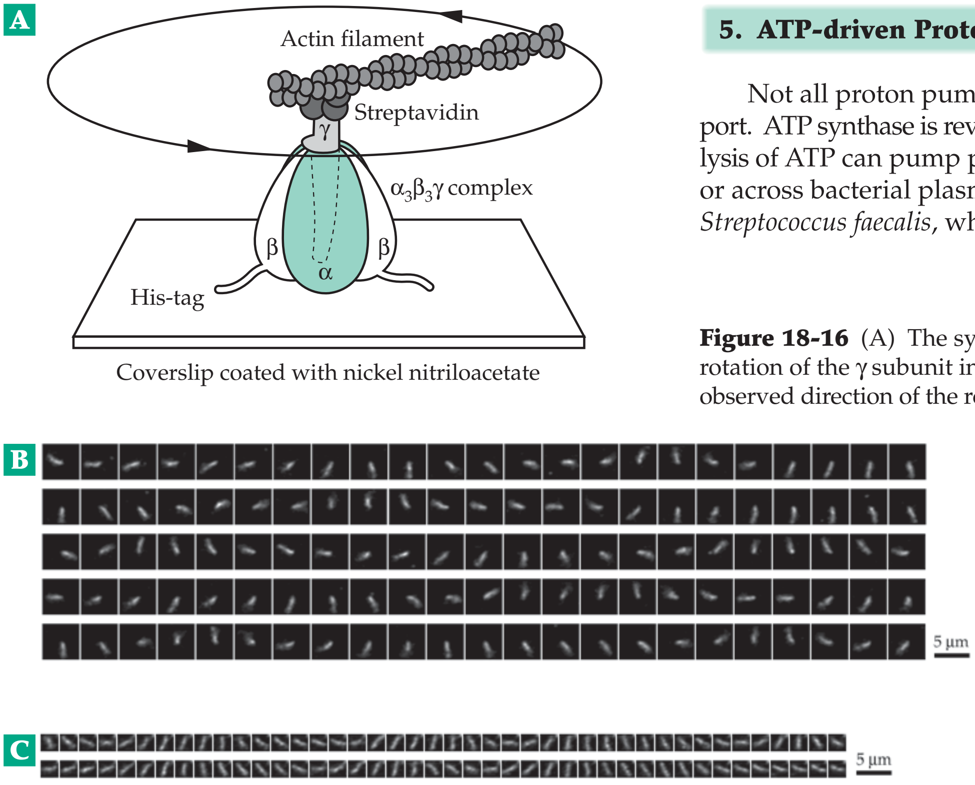

A
Actin filament
Streptavidin
γ
$\alpha_3\beta_3\gamma$ complex
β
β
α
His-tag
Coverslip coated with nickel nitriloacetate
B
5 µm
C
5 µm

Figure 18-16 (A) The system used for observation of the rotation of the γ subunit in the $\alpha_3\beta_3\gamma$ subcomplex. The observed direction of the rotation of the γ subunit is indicated by the arrows. (B) Sequential images of a rotating actin filament attached as in (A). (C) Similar images obtained with the axis of rotation near the middle of the filament. The images correspond to the view from the top in (A). Total length of the filament, 2.4 µm; rotary rate, 1.3 revolutions per second; time interval between images, 33 ms. From Noji *et al.*[258]

and form ATP by glycolysis, use an F_1F_0 ATP synthase complex to pump protons out to help regulate cytoplasmic pH.[268] Similar **vacuolar (V-type) H^+-ATPases** or V_1V_0 ATPases pump protons into vacuoles, Golgi and secretory vesicles, coated vesicles, and lysosomes[198,269-270] in every known type of eukaryotic cell.[271,272] These proton pumps are similar in appearance (Fig. 18-14,B) and in structure to F_1F_0 ATPases.[272a–c] The 65- to 77-kDa A subunits and 55- to 60-kDa B subunits are larger than the corresponding F_1F_0 α and β subunits. Accessory 40-, 39-, and 33-kDa subunits are also present in V_1. The V_0 portion appears to contain a hexamer of a 16-kDa proteolipid together with 110- and 21-kDa subunits.[271] V-type ATPases are also found in archaebacteria[271,273] and also in some clostridia[254] and other eubacteria.[273a] A type of proton pump, the **V-PPase**, uses hydrolysis of inorganic pyrophosphate as a source of energy.[274] It has been found in plants, in some phototrophic bacteria, and in acidic calcium storage vesicles (acidocalcisomes) of trypanosomes.[274a]

Other ATP-dependent proton pumps are present in the plasma membranes of yeast and other fungi[274b] and also in the acid-secreting parietal cells of the stomach (Fig. 18-17). The H^+-ATPase of *Neurospora* pumps H^+ from the cytoplasm without a counterion. It is electrogenic.[275,275a] However, the gastric H^+,K^+-ATPase exchanges H_3O^+ for K^+ and cleaves ATP with formation of a phosphoenzyme.[276] It belongs to the family of P-type ion pumps that includes the mammalian Na^+,K^+-ATPase (Fig. 8-25) and Ca^{2+}-ATPase (Fig. 8-26). These are discussed in Chapter 8. The H^+,K^+-ATPases, which are widely distributed within eukaryotes, are also similar, both in sequence and in the fact that a phospho group is transferred from ATP onto a carbox-ylate group of an aspartic acid residue in the protein. All of them, including a Mg-ATPase of *Salmonella*, are two-subunit proteins. A large catalytic α subunit contains the site of phosphorylation as well as the ATP- and ion-binding sites. It associates noncovalently with the smaller heavily glycosylated β subunit.[276–278] For example, the rabbit H^+,K^+-ATPase consists of a 1035-residue α chain which has ten transmembrane segments and a 290-residue β chain with a single transmembrane helix and seven N-linked glycosylation sites.[278]

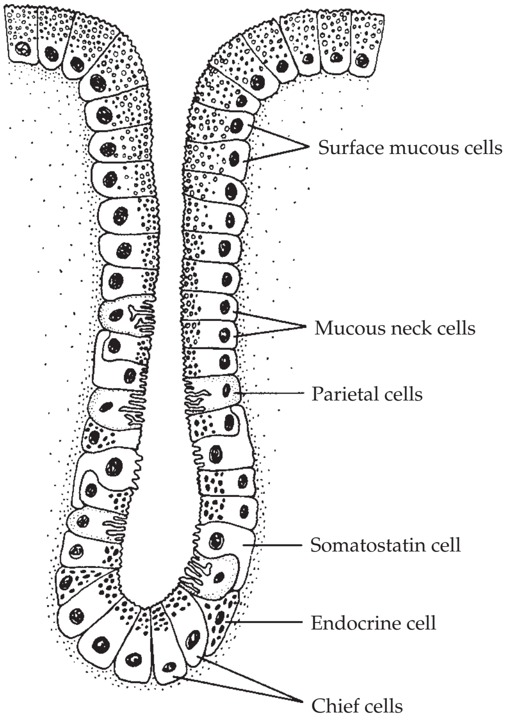
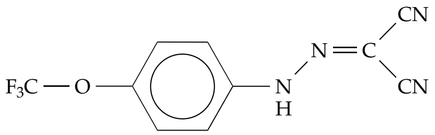


Figure 18-17 Schematic diagram of an acid-producing oxyntic gland of the stomach. The normal human stomach contains about 10^9 parietal (oxyntic) cells located in the walls of these glands. From Wolfe and Soll.[279] Modified from Ito. These glands also produce mucus, whose role in protecting the stomach lining from the high acidity is uncertain.[280]

6. Uncouplers and Energy-linked Processes in Mitochondria

Many compounds that uncouple electron transport from phosphorylation, like 2,4-dinitrophenol, are weak acids. Their anions are nucleophiles. According to the scheme of Fig. 18-12, they could degrade a high energy intermediate, such as Y ~ B, by a nucleophilic attack on Y to give an inactive but rapidly hydrolyzed derivative of Y. On the other hand, according to Mitchell's hypothesis uncouplers facilitate the transport of protons back into the mitochondria thereby destroying Δp. The fact that the anions of the uncouplers are large, often aromatic, and therefore soluble in the lipid bilayer supports this interpretation; the protonated uncouplers can diffuse into the mitochondria and the anion can diffuse back out. Mitochondria can also be uncoupled by a combination of ionophores, e.g., a mixture of valinomycin (Fig. 8-22), which carries K^+ into the mitochondria, plus nigericin, which catalyzes an exchange of K^+ (out) for H^+ (in).[172]

The uncoupler carbonyl cyanide *p*-(trifluoromethoxy)phenylhydrazone (FCCP) and related compounds are widely used in biochemical studies. Their action can be explained only partially by increased proton conduction.

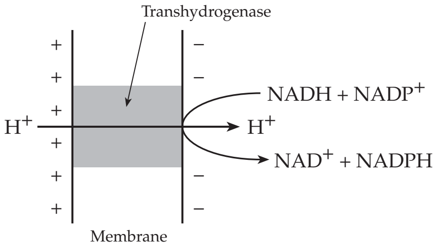


Carbonyl cyanide *p*-(trifluoromethoxy)phenylhydrazone (FCCP)

Uncoupling is sometimes important to an organism. The generation of heat by uncoupling is discussed in Box 18-C. The fungus *Bipolaris maydis* caused a crisis in maize production when it induced pore formation in mitochondrial membranes of a special strain used in production of hybrid seeds.[281,282]

Synthesis of ATP by mitochondria is inhibited by oligomycin, which binds to the OSCP subunit of ATP synthase. On the other hand, there are processes that require energy from electron transport and that are not inhibited by oligomycin. These **energy-linked processes** include the transport of many ions across the mitochondrial membrane (Section E) and reverse electron flow from succinate to NAD^+ (Section C,2). Dinitrophenol and many other uncouplers block the reactions, but oligomycin has no effect. This fact can be rationalized by the Mitchell hypothesis if we assume that Δp can drive these processes.

Another energy-linked process is the **transhydrogenase** reaction by which NADH reduces $NADP^+$ to form NADPH. In the cytoplasm various other reactions are used to generate NADPH (Chapter 17, Section I), but within mitochondria a membrane-bound transhydrogenase has this function.[283–286a] It couples the transhydrogenation reaction to the transport of one (or possibly more than one) proton back into the mitochondria (Eq. 18-14). A value of Δp of −180 mV could increase the ratio of [NADPH] / [$NADP^+$] within mitochondria by a factor of as much as 1000.



(18-14)

Transhydrogenases function in a similar way within bacteria. Whether from *E. coli*, photosynthetic bacteria, or bovine mitochondria, transhydrogenases have similar structures.[285] Two 510-residue α subunits associate with two 462-residue β subunits to form an $\alpha_2\beta_2$ tetramer with 10–14 predicted transmembrane helices. The α subunits contain separate NAD(H) and NADP(H) binding sites. A conformational change appears to be associated with the binding or release of the $NADP^+$ or NADPH.[287]

D. Transport and Exchange across Mitochondrial Membranes

Like the external plasma membrane of cells, the inner mitochondrial membrane is selective. Some nonionized materials pass through readily but the transport of ionic substances, including the anions of the dicarboxylic and tricarboxylic acids, is restricted. In some cases energy-dependent active transport is involved but in others one anion passes inward in exchange for another anion passing outward. In either case specific translocating carrier proteins are needed.

Solutes enter mitochondria through pores in thousands of molecules of the **voltage-gated anion-selective channel VDAC**, also known as mitochondrial porin.[15,16,288,289] In the absence of a membrane potential these pores allow free diffusion to molecules up to ~1.2 kDa in mass and may selectively permit passage of anions of 3- to 5-kDa mass. However, a membrane potential greater than ~20 mV causes the pores to close. NADH also decreases permeability. In the closed state the outer membrane becomes almost impermeant to ATP.[289,290]

An example of energy-dependent transport is the uptake of Ca^{2+} by mitochondria. As indicated in the lower part of Fig. 18-13, there are two possibilities for preservation of electrical neutrality according to the chemiosmotic theory. Counterions X^- may flow out to balance the protons discharged on the outside. On the other hand, if a cation such as Ca^{2+} flows inward to balance the two protons flowing outward, neutrality will be preserved and the mitochondrion will accumulate calcium ions. Experimentally such accumulations via a **calcium uniporter**[4] are observed to accompany electron transport. In the presence of a suitable ionophore energy-dependent accumulation of potassium ions also takes place.[291] In contrast, an **electroneutral exchange** of one Ca^{2+} for two Na^+ is mediated by a Na^+–Ca^{2+} exchanger.[292,293] It permits Ca^{2+} to leave mitochondria. A controversial role of mitochondria in accumulating Ca^{2+} postulates a special **rapid uptake mode** of exchange (see p. 1049).[294]

It is thought that glutamate enters mitochondria as the monoanion Glu^- in exchange for the dianion of aspartate Asp^{2-}. Like the uptake of Ca^{2+} this exchange is driven by Δp. Since a membrane potential can be created by this exchange in the absence of Δp, the process is electrogenic.[4] In contrast, an **electroneutral** exchange of malate^{2-} and 2-oxoglutarate^{2-} occurs by means of carriers that are not energy-linked.[295,296] This dicarboxylate transporter is only one of 35 structurally related mitochondrial carriers identified in the complete genome of yeast.[296,297] Another is the **tricarboxylic transporter** (citrate transport protein) which exchanges the dianionic form of citrate for malate, succinate, isocitrate, phosphoenolpyruvate, etc.[298,298a,b]

The important **adenine nucleotide carrier** takes ADP into the mitochondrial matrix for phosphorylation in a 1:1 ratio with ATP that is exported into the cytoplasm.[299-300b] This is one of the major rate-determining processes in respiration. It has been widely accepted that the carrier is electrogenic,[300]

BOX 18-C USING METABOLISM TO GENERATE HEAT: THERMOGENIC TISSUES

A secondary but important role of metabolism in warm-blooded animals is to generate heat. The heat evolved from ordinary metabolism is often sufficient, and an animal can control its temperature by regulating the heat exchange with the environment. Shivering also generates heat and is used from birth by pigs.[a] However, this is insufficient for many newborn animals, for most small mammals of all ages, and for animals warming up after hibernation. The need for additional heat appears to be met by **brown fat**, a tissue which contrasts strikingly with the more abundant white adipose tissue. Brown fat contains an unusually high concentration of blood vessels, many mitochondria with densely packed cristae, and a high ratio of cytochrome *c* oxidase to ATP synthase. Also present are a large number of sympathetic nerve connections, which are also related to efficient generation of heat. Newborn humans have a small amount of brown fat, and in newborn rabbits it accounts for 5–6% of the body weight.[b–d] It is especially abundant in species born without fur and in hibernating animals. Swordfish also have a large mass of brown adipose tissue that protects their brains from rapid cooling when traveling into cold water.[a]

The properties of brown fat pose an interesting biochemical question. Is the energy available from electron transport in the mitochondria dissipated as heat because ATP synthesis is uncoupled from electron transport? Or does ATP synthesis take place but the resulting ATP is hydrolyzed wastefully through the action of ATPases? Part of the answer came from the discovery that mitochondria of brown fat cells synthesize a 32-kDa **uncoupling protein** (UCP1 or thermogenin). It is incorporated into the inner mitochondrial membranes where it may account for 10–15% of total protein.[d–f] This protein, which is a member of a family of mitochondrial membrane metabolic carriers (Table 18-8), provides a “short-circuit” that allows the protonmotive force to be dissipated rapidly, perhaps by a flow of protons out through the uncoupling protein.[g–i] Synthesis of the uncoupling protein is induced by exposure to cold, but when an animal is warm the uncoupling action is inhibited.

The uncoupling protein resembles the ATP/ADP and phosphate *anion* carriers (Table 18-8),[g,i] which all have similar sizes and function as homodimers. Each monomeric subunit has a triply repeated ~100-residue sequence, each repeat forming two transmembrane helices. Most mitochondrial transporters carry anions, and UCP1 will transport Cl^-.[h,i] However, the relationship of chloride transport to its real function is unclear. Does the protein transport H^+ into, or does it carry HO^- out from, the mitochondrial matrix?[g,h] Another possibility is that a fatty acid anion binds H^+ on the intermembrane surface and carries it across into the matrix as an unionized fatty acid. The fatty acid anion could then pass back out using the anion transporter function and assisted by the membrane potential.[h,i]

The uncoupling protein is affected by several control mechanisms. It is inhibited by nucleotides such as GDP, GTP, ADP, and ATP which may bind at a site corresponding to that occupied by ATP or ADP in the ADP/ATP carrier.[i] Uncoupling is stimulated by noradrenaline,[f] which causes a rapid increase in heat production by brown fat tissues, apparently via activation of adenylate cyclase. Uncoupling is also stimulated by fatty acids.[j] Recently UCP1 and related uncoupling proteins have been found to require both fatty acids and **ubiquinone** for activity.[jj,jk]

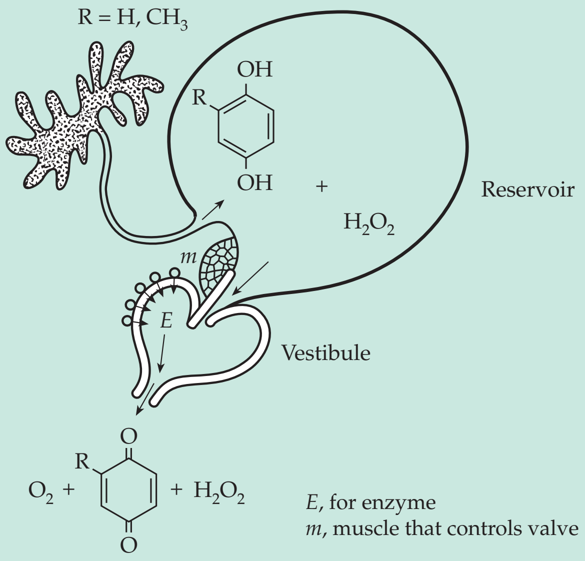
It has been suggested that brown adipose tissue may also function to convert excess dietary fat into heat and thereby to resist obesity.[k–m] Mice lacking the gene for the mitochondrial uncoupling protein are cold-sensitive but not obese. However, other proteins, homologous to UCP1, have been discovered. They may partially compensate for the loss.[m,h]

The bombardier beetle generates a hot, quinone-containing defensive discharge, which is sprayed in a pulsed jet from a special reaction chamber at a temperature of 100°C.[n–p] The reaction mixture of 25% hydrogen peroxide and 10% hydroquinone plus methylhydroquinone is stored in a reservoir as shown in the accompanying figure and reacts with explosive force when it comes into contact with catalase and peroxidases in the reaction chamber. The synthesis and storage of 25% H_2O_2 poses interesting biochemical questions!

Some plant tissues are thermogenic. For example, the spadix (or inflorescence, a sheathed floral spike) of the skunk cabbage *Symplocarpus foetidus* can maintain a 15–35°C higher temperature than that of the surrounding air.[q] The voodoo lily in a single day heats the upper end of its long spadex to a temperature 22°C above ambient, volatilizing a foul smelling mixture of indoles and amines.[r,s] This is accomplished using the alternative oxidase system[s] (Box D in Fig. 18-6). The lotus flower maintains a temperature of 30–35°C, while the ambient temperature may vary from 10–30°C.[t] While the volatilization of insect attractants may be the primary role for thermogenesis in plants, the warm flowers may also offer an important reward to insect pollinators. Beetles and bees require thoracic temperatures above 30°C to initiate flight and, therefore,

BOX 18-C (continued)

benefit from the warm flowers.[t] While in flight bees vary their metabolic heat production by altering their rate of flight, hovering, and other changes in physical activity.[u]



Reservoir and reaction vessel of the bombardier beetle. From D. J. Aneshansley, *et al.*[n]

---

[a] Tyler, D. D. (1992) *The Mitochondrion in Health and Disease*, VCH Publ., New York

[b] Dawkins, M. J. R., and Hull, D. (1965) *Sci. Am.* **213**(Aug), 62–67

[c] Lindberg, O., ed. (1970) *Brown Adipose Tissue*, Am. Elsevier, New York

[d] Nicholls, D. G., and Rial, E. (1984) *Trends Biochem. Sci.* **9**, 489–491

[e] Cooney, G. J., and Newsholme, E. A. (1984) *Trends Biochem. Sci.* **9**, 303–305

[f] Ricquier, D., Casteilla, L., and Bouillaud, F. (1991) *FASEB J.* **5**, 2237–2242

[g] Klingenberg, M. (1990) *Trends Biochem. Sci.* **15**, 108–112

[h] Jaburek, M., Varecha, M., Gimeno, R. E., Dembski, M., Jezek, P., Zhang, M., Burn, P., Tartaglia, L. A., and Garlid, K. D. (1999) *J. Biol. Chem.* **274**, 26003–26007

[i] González-Barroso, M. M., Fleury, C., Levi-Meyrueis, C., Zaragoza, P., Bouillaud, F., and Rial, E. (1997) *Biochemistry* **36**, 10930–10935

[j] Hermesh, O., Kalderon, B., and Bar-Tana, J. (1998) *J. Biol. Chem.* **273**, 3937–3942

[jj] Echtay, K. S., Winkler, E., and Klingenberg, M. (2000) *Nature (London)* **408**, 609–613

[jk] Echtay, K. S., Winkler, E., Frischmuth, K., and Klingenberg, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1416–1421

[k] Rothwell, N. J., and Stock, M. J. (1979) *Nature (London)* **281**, 31–35

[l] Tai, T.-A. C., Jennermann, C., Brown, K. K., Oliver, B. B., MacGinnitie, M. A., Wilkison, W. O., Brown, H. R., Lehmann, J. M., Kliewer, S. A., Morris, D. C., and Graves, R. A. (1996) *J. Biol. Chem.* **271**, 29909–29914

[m] Enerbäck, S., Jacobsson, A., Simpson, E. M., Guerra, C., Yamashita, H., Harper, M.-E., and Kozak, L. P. (1997) *Nature (London)* **387**, 90–94

[n] Aneshansley, D. J., Eisner, T., Widom, J. M., and Widom, B. (1969) *Science* **165**, 61–63

[o] Eisner, T., and Aneshansley, D. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9705–9709

[p] Dean, J., Aneshansley, D. J., Edgerton, H. E., and Eisner, T. (1990) *Science* **248**, 1219–1221

[q] Knutson, R. M. (1974) *Science* **186**, 746–747

[r] Diamond, J. M. (1989) *Nature (London)* **339**, 258–259

[s] Rhoads, D. M., and McIntosh, L. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2122–2126

[t] Seymour, R. S., and Schultze-Motel, P. (1996) *Nature (London)* **383**, 305

[u] Harrison, J. F., Fewell, J. H., Roberts, S. P., and Hall, H. G. (1996) *Science* **274**, 88 – 90

---

bringing in ADP^{3-} and exporting ATP^{4-} in an exchange driven by Δp. However, an electroneutral exchange, e.g., of ADP^{3-} for ATP^{3-}, may also be possible. The carrier is an ~300-residue 32-kDa protein, which is specifically inhibited by the plant glycoside **atractyloside** or the fungal antibiotic **bongkrekate**. The carrier is associated with bound cardiolipin.[301] This one transporter accounts for ~10% of all of the mitochondrial protein.[302,303]

A separate dimeric carrier allows P_i to enter, probably as $H_2PO_4^-$.[304-305a] This ion enters mitochondria in an electroneutral fashion, either in exchange for OH^- or by cotransport with H^+. A less important carrier[306] exchanges HPO_4^{2-} for malate^{2-}. Several other transporters help to exchange organic and inorganic ions. One of them allows pyruvate to enter mitochondria in exchange for OH^- or by cotransport with H^+. Some of the identified carriers are listed in Table 18-8. As discussed in Chapter 8, exchange carriers are also important in plasma membranes of organisms from bacteria to human beings. For example, many metabolites enter cells by cotransport with Na^+ using the energy of the Na^+ gradient set up across the membrane by the Na^+,K^+ pump.

Under some circumstances the inner membrane develops one or more types of large-permeability pore. An increase in Ca^{2+} may induce opening of an unselective pore which allows rapid uptake of Ca^{2+}.[294,307,307a] A general anion-specific channel may be involved in volume homeostasis of mitochondria.[308]

Mitochondria are not permeable to NADH. However, reactions of glycolysis and other dehydrogenations in the cytoplasm quickly reduce available NAD^+ to NADH. For aerobic metabolism to occur, the "reducing equivalents" from the NADH must be transferred into the mitochondria. Fungi and green plants have solved the problem by providing *two* NADH dehydrogenases embedded in the inner mitochondrial membranes (Fig. 18-6). One faces the matrix space and oxidizes the NADH produced in the matrix while the second faces outward to the intermembrane space and is able to oxidize the NADH formed in the cytoplasm. In animals the reducing equivalents from NADH enter the mitochondria indirectly. There are several mechanisms, and more than one may function simultaneously in a tissue.

In insect flight muscle, as well as in many mammalian tissues, NADH reduces dihydroxyacetone phosphate. The resulting *sn*-3-glycerol *P* passes through the permeable outer membrane of the mitochondria, where it is reoxidized to dihydroxyacetone phosphate by the FAD-containing glycerol-phosphate dehydrogenase embedded in the outer surface of the inner membrane (Figs. 18-5, 18-6). The dihydroxyacetone can then be returned to the cytoplasm. The overall effect of this **glycerol-phosphate shuttle** (Fig. 18-18A) is to provide for mitochondrial oxidation of NADH produced in the cytoplasm. In heart and liver the same function is served by a more complex **malate–aspartate shuttle** (Fig. 18-18B).[309] Reduction of oxaloacetate to malate by NADH, transfer of malate into mitochondria, and reoxidation with NAD^+ accomplishes the transfer of reduction equivalents into the mitochondria. Mitochondrial membranes are not very permeable to oxaloacetate. It returns to the cytoplasm mainly via transamination to aspartate, which leaves the mitochondria together with 2-oxoglutarate. At the same time glutamate enters the mitochondria in exchange for aspartate. The 2-oxoglutarate presumably exchanges with the entering malate as is indicated in Fig. 18-18B. The export of aspartate may be energy-linked as a result of the use of an electrogenic carrier that exchanges $\text{glutamate}^- + H^+$ entering mitochondria for aspartate^- leaving the mitochondria. Thus, Δp may help to expel aspartate from mitochondria and to drive the shuttle.

The glycerol-phosphate shuttle, because it depends upon a mitochondrial flavoprotein, provides ~ 2 ATP per electron pair ($P/O = 2$), whereas the malate–aspartate shuttle may provide a higher yield of ATP. The glycerol-phosphate shuttle is essentially irreversible, but the reactions of the malate–aspartate shuttle can be reversed and utilized in gluconeogenesis (Chapter 17).

TABLE 18-8
Some Mitochondrial Membrane Transporters[a]

Ion Diffusing In	Ion Diffusing Out	Comment[b]
ADP^{3-}	ATP^{3-}	
or ADP^{3-}	ATP^{4-}	Electrogenic symport
$H_2PO_4^- + H^+$		Electroneutral symport
or $H_2PO_4^-$	OH^-	
HPO_4^{2-}	Malate^{2-}	
Malate^{2-}	$\text{2-Oxoglutarate}^{2-}$	
$\text{Glutamate}^{2-} + H^+$	Aspartate^{2-}	
Glutamate^-	OH^-	
Pyruvate^-	OH^-	
or $\text{Pyruvate}^- + H^+$		Electroneutral symport
$\text{Citrate}^{3-} + H^+$	Malate^{2-}	
Ornithine^+	H^+	
Acylcarnitine	Carnitine	
2 Na^+	Ca^{2+}	
H^+	K^+	
H^+	Na^+	
General transporters		
VDAC (porin)		Outer membrane
Large anion pores		Inner membrane

[a] From Nicholls and Ferguson[172] and Tyler[4].

[b] Unless indicated otherwise the transporters are *antiporters* that catalyze an electroneutral ion exchange.

E. Energy from Inorganic Reactions

Some bacteria obtain all of their energy from inorganic reactions. These **chemolithotrophs** usually have a metabolism that is similar to that of heterotrophic organisms, but they also have the capacity to obtain all of their energy from an inorganic reaction. In order to synthesize carbon compounds they must be able to fix CO_2 either via the reductive pentose phosphate cycle or in some other way. The chloroplasts of green plants, using energy from sunlight, supply the organism with both ATP and the reducing agent NADPH (Chapter 17). In a similar way the lithotrophic bacteria obtain both energy and reducing materials from inorganic reactions.

Chemolithotrophic organisms often grow slowly, making study of their metabolism difficult.[310] Nevertheless, these bacteria usually use electron transport chains similar to those of mitochondria. ATP is formed by oxidative phosphorylation, the amount formed per electron pair depending upon the number of proton-pumping sites in the chain. This, in turn, depends upon the electrode potentials of the reactions involved. For example, H_2, when oxidized by O_2, leads to passage of electrons through the entire electron transport chain with synthesis of ~3 molecules of ATP per electron pair. On the other hand, oxidation by O_2 of nitrite, for which $E°'$ (pH 7) = +0.42 V, can make use only of the site III part of the chain. Not only is the yield of ATP less than in the oxidation of H_2 but also there is another problem. Whereas reduced pyridine

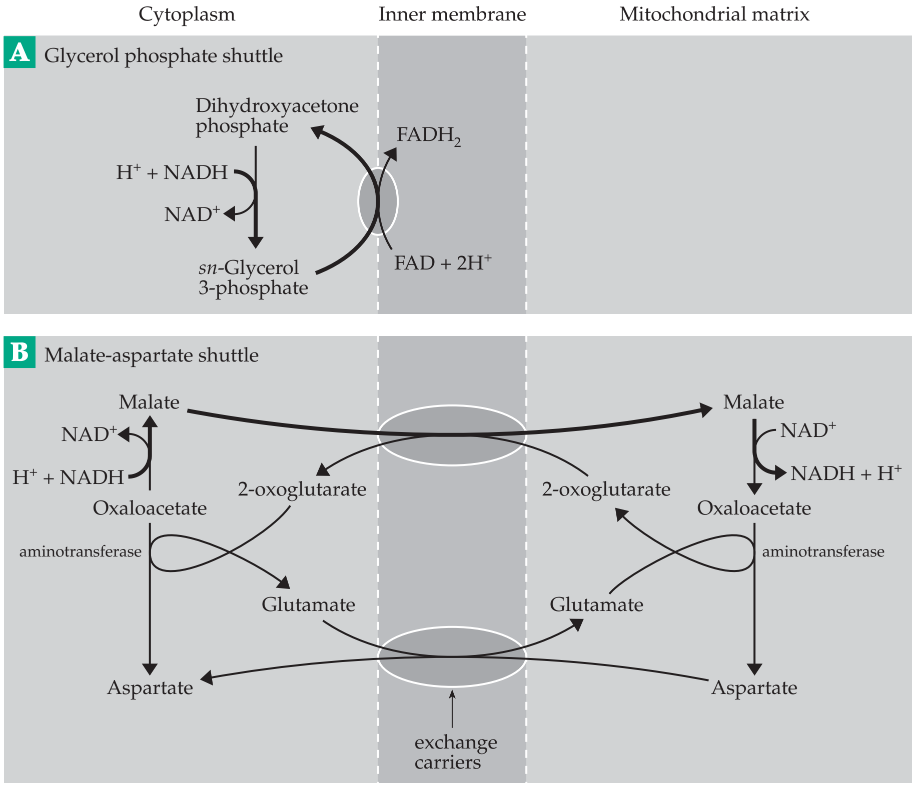

Cytoplasm
Inner membrane
Mitochondrial matrix
A Glycerol phosphate shuttle
Dihydroxyacetone phosphate
H+ + NADH
NAD+
sn-Glycerol 3-phosphate
FADH2
FAD + 2H+
B Malate-aspartate shuttle
Malate
NAD+
H+ + NADH
Oxaloacetate
aminotransferase
2-oxoglutarate
Glutamate
Aspartate
Malate
NAD+
NADH + H+
Oxaloacetate
aminotransferase
2-oxoglutarate
Glutamate
Aspartate
exchange carriers

Figure 18-18 (A) The glycerol-phosphate shuttle and (B) the malate–aspartate shuttle for transport from cytoplasmic NADH into mitochondria. The heavy arrows trace the pathway of the electrons (as 2H) transported.

nucleotides needed for biosynthesis can be generated readily from H_2, nitrite is not a strong enough reducing agent to reduce NAD^+ to NADH. The only way that reducing agents can be formed in cells utilizing oxidation of nitrite as an energy source is via *reverse electron flow* driven by hydrolysis of ATP or by Δp. Such reverse electron flow is a common process for many chemolithotrophic organisms.

Let us consider the inorganic reactions in two groups: (1) oxidation of reduced inorganic compounds by O_2 and (2) oxidation reactions in which an inorganic oxidant, such as nitrate or sulfate, substitutes for O_2. The latter reactions are often referred to as **anaerobic respiration**.

1. Reduced Inorganic Compounds as Substrates for Respiration

The hydrogen-oxidizing bacteria. Species from several genera including *Hydrogenomonas*, *Pseudomonas*, and *Alcaligenes* oxidize H_2 with oxygen:

$$H_2 + 1/2\ O_2 \rightarrow H_2O \qquad \Delta G^\circ = -237.2 \text{ kJ mol}^{-1} \qquad (18\text{-}15)$$

Some can also oxidize carbon monoxide:

$$CO + 1/2\ O_2 \rightarrow CO_2 \qquad \Delta G^\circ = -257.1 \text{ kJ mol}^{-1} \qquad (18\text{-}16)$$

The hydrogen bacteria can also oxidize organic compounds using straightforward metabolic pathways. The key enzyme is a membrane-bound nickel-containing hydrogenase (Fig. 16-26), which delivers electrons from H_2 into the electron transport chain.[310a] A second soluble hydrogenase (sometimes called **hydrogen dehydrogenase**) transfers electrons to $NADP^+$ to form NADPH for use in the reductive pentose phosphate cycle and for other biosynthetic purposes.

Nitrifying bacteria. Two genera of soil bacteria oxidize ammonium ion to nitrite and nitrate (Eqs. 18-17 and 18-18).[311]

$$NH_4^+ + 3/2\ O_2 \rightarrow NO_2^- + 2\ H^+ + H_2O\ (\textit{Nitrosomonas}) \qquad \Delta G'\ (\text{pH } 7) = -272 \text{ kJ mol}^{-1} \qquad (18\text{-}17)$$

$$NO_2^- + 1/2\ O_2 \rightarrow NO_3^-\ (\textit{Nitrobacter}) \qquad \Delta G'\ (\text{pH } 7) = -76 \text{ kJ mol}^{-1} \qquad (18\text{-}18)$$

The importance of these reactions to the energy metabolism of the bacteria was recognized in 1895 by Winogradsky, who first proposed the concept of chemiautotrophy. Because the nitrifying bacteria grow slowly (generation time ~10–12 h) it has been hard to get enough cells for biochemical studies and progress has been slow. The reaction catalyzed by *Nitrosomonas* (Eq. 18-17) is the more complex; it occurs in two or more stages and is catalyzed by two enzymes as illustrated in Fig. 18-19. The presence of hydrazine blocks oxidation of hydroxylamine (NH_2OH) in step *b* and permits that intermediate to accumulate. The oxidation of ammonium ion by O_2 to hydroxylamine (step *a*) is endergonic with $\Delta G'$ (pH 7) = 16 kJ mol^{-1} and is incapable of providing energy to the cell. It occurs by a hydroxylation mechanism (see Section G). On the other hand, the oxidation of hydroxylamine to nitrite by O_2 in step *b* is highly exergonic with $\Delta G'$ (pH 7) = –228 kJ mol^{-1}. The hydroxylamine oxidoreductase that catalyzes this reaction is a trimer of 63-kDa subunits, each containing seven *c*-type hemes and an unusual heme P450, which is critical to the enzyme's function[312–314a] and which is covalently linked to a tyrosine as well as to two cysteines.

The electrode potentials for the two- and four-electron oxidation steps are indicated in Fig. 18-19. It is apparent that step *b* can feed four electrons into the electron transport system at about the potential of ubiquinone. Two electrons are needed to provide a cosubstrate (Section G) for the ammonia monooxygenase and two could be passed on to the terminal cytochrome aa_3 oxidase. The stoichiometry of proton pumping in complexes III and IV is uncertain, but if it is assumed to be as shown in Fig. 18-19 and similar to that in Fig. 18-5, there will be ~13 protons available to be passed through ATP synthase to generate ~3 ATP per NH_3 oxidized. However, to generate NADH for reductive biosynthesis *Nitrosomonas* must send some electrons to NADH dehydrogenase (complex I) using reverse electron transport, a process that depends upon Δp to drive the reaction via a flow of protons through the NADH dehydrogenase from the periplasm back into the bacterial cytoplasm (Fig. 18-19).

Nitrobacter depends upon a simpler energy-yielding reaction (Eq. 18-18) with a relatively small Gibbs energy decrease. The two-electron oxidation delivers electrons to the electron transport chain at $E°' = +0.42$ V. The third oxygen in NO_3^- originates from H_2O, rather than from O_2 as might be suggested by Eq. 18-18.[316,317] It is reasonable to anticipate that a single molecule of ATP should be formed for each pair of electrons reacting with O_2. However, *Nitrobacter* contains a confusing array of different cytochromes in its membranes.[311] Some of the ATP generated by passage of electrons from nitrite to oxygen must be used to drive a reverse flow of electrons through both a bc_1-type complex and NADH dehydrogenases. This generates reduced pyridine nucleotides required for biosynthetic reactions (Fig. 18-20).

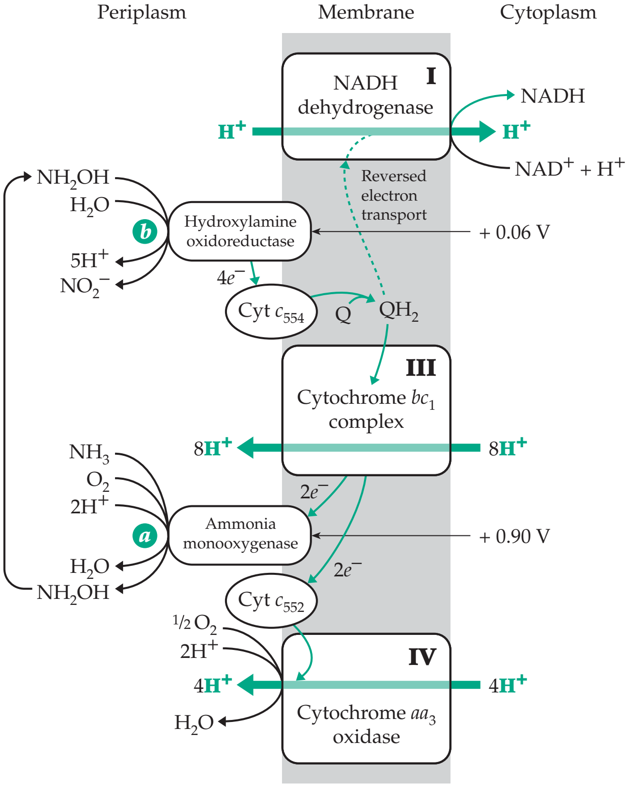


Figure 18-19 The ammonia oxidation system of the bacterium *Nitrosomonas*. Oxidation of ammonium ion (as free NH_3) according to Eq. 18-17 is catalyzed by two enzymes. The location of ammonia monooxygenase (step *a*) is uncertain but hydroxylamine oxidoreductase (step *b*) is periplasmic. The membrane components resemble complexes I, III, and IV of the mitochondrial respiratory chain (Fig. 18-5) and are assumed to have similar proton pumps. Solid green lines trace the flow of electrons in the energy-producing reactions. This includes flow of electrons to the ammonia monoxygenase. Complexes III and IV pump protons out but complex I catalyzes reverse electron transport for a fraction of the electrons from hydroxylamine oxidoreductase to NAD^+. Modified from Blaut and Gottschalk.[315]

An interesting feature of the structure of *Nitrobacter* is the presence of several double-layered membranes which completely envelop the interior of the cell. Nitrite entering the cell is oxidized on these membranes and cannot penetrate to the interior, where it might have toxic effects.

The sulfur-oxidizing bacteria. Anaerobic conditions prevail in marine sediments, in poorly stirred swamps, and around hydrothermal vents at the bottom of the sea. Sulfate-reducing bacteria form high concentrations (up to mM) of H_2S (in equilibrium with HS^- and S^{2-})[318–320] This provides the substrate for bacteria of the genus *Thiobacillus*, which are able to oxidize sulfide, elemental sulfur, thiosulfate, and sulfite to sulfate and live where the aerobic and anaerobic regions meet.[311,321–323] Most of these small gram-negative

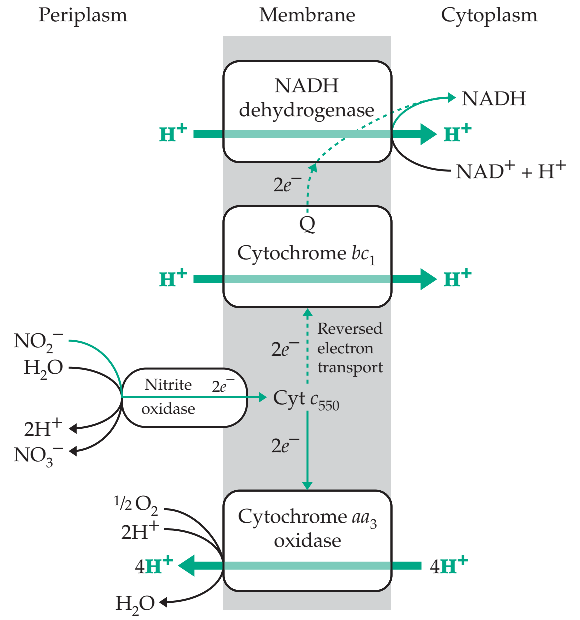
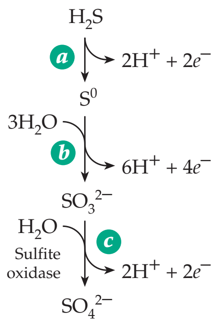

Periplasm
Membrane
Cytoplasm
NADH dehydrogenase
H+
H+
NADH
NAD+ + H+
2e−
Q
Cytochrome bc1
H+
H+
Reversed electron transport
2e−
NO2−
H2O
Nitrite oxidase
2e−
Cyt c550
2H+
NO3−
2e−
1/2 O2
2H+
Cytochrome aa3 oxidase
4H+
4H+
H2O

Figure 18-20 Electron transport system for oxidation of the nitrite ion to the nitrate ion by *Nitrobacter*. Only one site of proton pumping for oxidative phosphorylation is available. Generation of NADH for biosynthesis requires two stages of reverse electron transport.

organisms, found in water and soil, are able to grow in a simple salt medium containing an oxidizable sulfur compound and CO_2. One complexity in understanding their energy-yielding reactions is the tendency of sulfur to form chain molecules. Thus, when sulfide is oxidized, it is not clear that it is necessarily converted to monoatomic elemental sulfur as indicated in Eq. 18-19. Elemental sulfur (S_8^0) often precipitates. In *Beggiotoa*, another sulfide-oxidizing bacterium, sulfur is often seen as small globules within the cells. Fibrous sulfur precipitates are often abundant in the sulfide-rich layers of ponds, lakes, and oceans.[318]


H2S
a
2H+ + 2e−
S0
3H2O
b
6H+ + 4e−
SO3 2−
H2O
c
Sulfite oxidase
2H+ + 2e−
SO4 2−

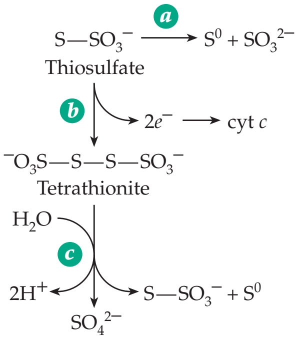
(18-19)

The reactions of Eq. 18-19 occur in the periplasmic space of some species.[315,323a,324] Steps *a* and *b* of Eq. 18-19 are catalyzed by a 67-kDa sulfide dehydrogenase in the periplasm of a purple photosynthetic bacterium.[324] The enzyme consists of a 21-kDa subunit containing two cytochrome *c*-like hemes, presumably the site of binding of S^{2-}, and a larger 46-kDa FAD-containing flavoprotein resembling glutathione reductase.[324] The molybdenum-containing sulfite oxidase (Fig. 16-32), which is found in the intermembrane space of mitochondria, may be present in the periplasmic space of these bacteria. However, there is also an intracellular pathway for sulfite oxidation (see Eq. 18-22).

The sulfide-rich layers inhabited by the sulfur oxidizers also contain thiosulfate, $S_2O_3^{2-}$. It may arise, in part, from reaction of glutathione with elemental sulfur:

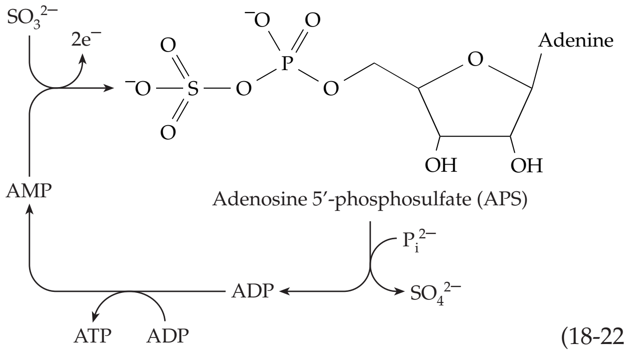
$$\mathrm{GSH} + S_8^0 \rightarrow \mathrm{G{-}S{-}S_8{-}H} \tag{18-20}$$

The linear polysulfide obtained by this reaction may be oxidized, the sulfur atoms being removed from the chain either one at a time to form sulfite or two at a time to form thiosulfate.[322,322a] Thiosulfate is oxidized by all species, the major pathway beginning with cleavage to S^0 and SO_3^{2-} (Eq. 18-21, step *a*). At high thiosulfate concentrations some may be oxidized to tetrathionate (step *b*), which is hydrolyzed to sulfate (step *c*).


S—SO3−
Thiosulfate
a
S0 + SO3 2−
b
2e−
cyt c
−O3S—S—S—SO3−
Tetrathionite
H2O
c
2H+
S—SO3− + S0
SO4 2−

(18-21)

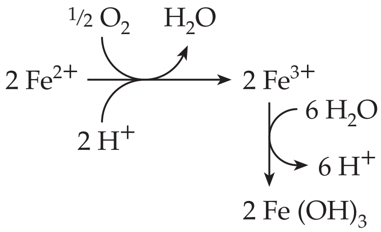
Oxidation of sulfite to sulfate within cells occurs by a pathway through **adenosine 5'-phosphosulfate** (**APS**, adenylyl sulfate). Oxidation via APS (Eq. 18-22) provides a means of substrate-level phosphorylation,


SO3 2−
2e−
Adenine
AMP
Adenosine 5′-phosphosulfate (APS)
OH
OH
Pi 2−
ADP
SO4 2−
ATP
ADP

(18-22)

the only one known among chemolithotrophic bacteria. No matter which of the two pathways of sulfite oxidation is used, thiobacilli also obtain energy via electron transport. With a value of $E°'$ (pH 7) of −0.454 V [$E°'$ (pH 2) = −0.158 V] for the sulfate–sulfite couple an abundance of energy may be obtained. Oxidation of sulfite to sulfate produces hydrogen ions. Indeed, pH 2 is optimal for the growth of *Thiobacillus thiooxidans*, and the bacterium withstands 5% sulfuric acid.[322]

The "iron bacterium" *Thiobacillus ferrooxidans* obtains energy from the oxidation of Fe^{2+} to Fe^{3+} with subsequent precipitation of ferric hydroxide (Eq. 18-23). However, it has been recognized recently that a previously unknown species of Archaea is much more important than *T. ferrooxydans* in catalysis of this reaction.[324a]



(18-23)

Since the reduction potential for the Fe(II) / Fe(III) couple is + 0.77 V at pH 7, the energy obtainable in this reaction is small. These bacteria always oxidize reduced sulfur compounds, too. Especially interesting is their oxidation of **pyrite**, ferrous sulfide (Eq. 18-24). The Gibbs energy change was calculated from published data[325] using a value of $G_f°$ for $Fe(OH)_3$ of

$$2\ FeS + 4\tfrac{1}{2}\ O_2 + 5\ H_2O \rightarrow 2\ Fe(OH)_3 + 2\ SO_4^{2-} + 4\ H^+$$
$$\Delta G'\ (pH\ 2) = -1340\ kJ \qquad (18\text{-}24)$$

−688 kJ mol^{-1} estimated from its solubility product. Because sulfuric acid is generated in this reaction, a serious water pollution problem is created by the bacteria living in abandoned mines. Water running out of the mines often has a pH of 2.3 or less.[326]

Various invertebrates live in S^{2-}-containing waters. Among these is a clam that has symbiotic sulfur-oxidizing bacteria living in its gills. The clam tissues apparently carry out the first step in oxidation of the sulfide.[327] Among the animals living near sulfide-rich thermal vents in the ocean floor are giant 1-meter-long tube worms. Both a protective outer tube and symbiotic sulfide-oxidizing bacteria protect them from toxic sulfides.[319,320]

2. Anaerobic Respiration

Nitrate as an electron acceptor. The use of nitrate as an alternative oxidant to O_2 is widespread among bacteria. For example, *E. coli* can subsist anaerobically by reducing nitrate to nitrite (Eq. 18-25).[311,328] The respiratory (dissimilatory) nitrate reductase that

$$NO_3^- + 2\ H^+ + 2e^- \rightarrow NO_2^- + H_2O$$
$$E°'\ (pH7) = +0.421\ V \qquad (18\text{-}25)$$

catalyzes the reaction is a large three-subunit molybdenum-containing protein. The enzyme is present in the plasma membrane, and electrons flow from ubiquinone through as many as six heme and Fe–S centers to the molybdenum atom.[328-329] A second molybdoenzyme, formate dehydrogenase (discussed in Chapter 16), appears to be closely associated with nitrate reductase. Formate is about as strong a reducing agent as NADH (Table 6-8) and is a preferred electron donor for the reduction of NO_3^- (Eq. 18-26).[329a,b] Since cytochrome *c* oxidase of the electron transport chain is bypassed, one less ATP is formed than when O_2 is the oxidant. Nitrate is the oxidant preferred by bacteria grown under anaerobic conditions. The

$$HCOO^- + NO_3^- + H^+ \rightarrow NO_2^- + CO_2 + H_2O$$
$$\Delta G'\ (pH\ 7) = 165\ kJ\ mol^{-1} \qquad (18\text{-}26)$$

presence of NO_3^- induces the synthesis of nitrate reductase and represses the synthesis of alternative enzymes such as **fumarate reductase**,[330,331] which reduces fumarate to succinate (see also p. 1027). On the other hand, if NO_3^- is absent and fumarate, which can be formed from pyruvate, is present, synthesis of fumarate reductase is induced. Although it is a much weaker oxidant than is nitrate ($E^{o'}$ = 0.031 V), fumarate is still able to oxidize H_2 or NADH with oxidative phosphorylation. Like the related succinate dehydrogenase, fumarate reductase of *E. coli* is a flavoprotein with associated Fe / S centers. It contains covalently linked FAD and Fe_2S_2, Fe_4S_4, and three-Fe iron–sulfur centers.[332] In some bacteria a soluble periplasmic cytochrome c_3 carried out the fumarate reduction step.[332a] Trimethylamine *N*-oxide[330,333] or dimethylsulfoxide (DMSO; Eq. 16-62)[334,335] can also serve as alternative oxidants for anaerobic respiration using appropriate molybdemum-containing reductases (Chapter 16).

Reduction of nitrite: denitrification. The nitrite formed in Eq. 18-25 is usually reduced further to ammonium ions (Eq. 18-27). The reaction may not be important to the energy metabolism of the bacteria, but it provides NH_4^+ for biosynthesis. This six-electron reduction is catalyzed by a hexaheme protein containing six *c*-type hemes bound to a single 63-kDa polypeptide chain.[336,337]

$$NO_2^- + 8\ H^+ + 6\ e^- \rightarrow 2\ H_2O + NH_4^+ \qquad (18\text{-}27)$$

Several types of **denitrifying bacteria**[315,338–340] use either nitrate or nitrite ions as oxidants and reduce nitrite to N_2. A typical reaction for *Micrococcus denitrificans* is oxidation of H_2 by nitrate (Eq. 18-28). *Thiobacillus denitrificans*, like other thiobacilli, can oxidize

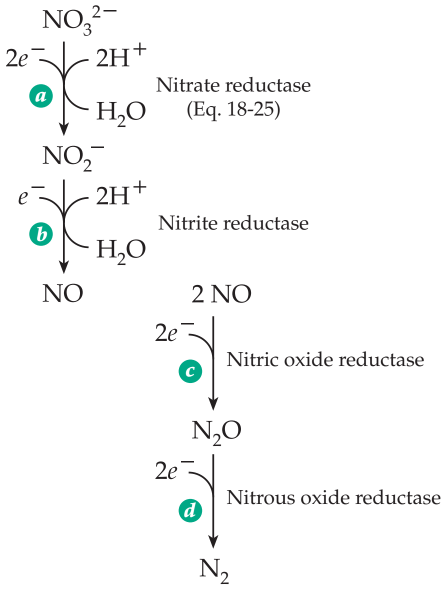
$$\begin{aligned} &5\ H_2 + 2\ NO_3^- + 2\ H^+ \rightarrow N_2 + 6\ H_2O \\ &\Delta G'\ (\text{pH } 7) = -561\ \text{kJ mol}^{-1} \text{ of nitrate reduced or} \\ &\qquad -224\ \text{kJ mol}^{-1} \text{ of } H_2 \text{ oxidized} \end{aligned} \tag{18-28}$$

sulfur as well as H_2S or thiosulfate using nitrate as the oxidant (Eq. 18-29):

$$\begin{aligned} &5\ S + 6\ NO_3^- + 2\ H_2O \rightarrow 5\ SO_4^{2-} + 3\ N_2 + 4\ H^+ \\ &\Delta G'\ (\text{pH } 7) = -455\ \text{kJ mol}^{-1} \text{ of nitrate reduced or} \\ &\qquad -546\ \text{kJ mol}^{-1} \text{ of S oxidized} \end{aligned} \tag{18-29}$$

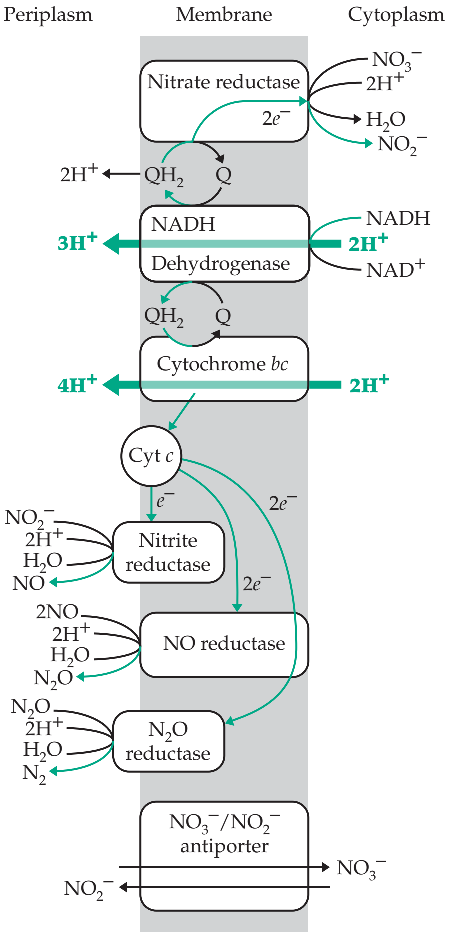
The reactions begin with reduction of nitrate to nitrite (Eq. 18-25) and continue with further reduction of nitrite to nitric oxide, **NO**; nitrous oxide, **N_2O**; and dinitrogen, **N_2**. A probable arrangement of the four enzymes needed for the reactions of Eq. 18-30 in *Paracoccus denitrificans* is shown in Fig. 18-21. See also pp. 884, 885.

Two types of dissimilatory nitrite reductases catalyze step *b* of Eq. 18-30. Some bacteria use a copper-containing enzyme, which contains a type 1 (blue) copper bound to a β barrel domain of one subunit and a type 2 copper at the catalytic center. The type 1 copper is thought to receive electrons from the small copper-containing carrier pseudoazurin (Chapter 16).[341–342b]



NO_3^{2-}
$2e^-$
$2H^+$
a
H_2O
Nitrate reductase (Eq. 18-25)
NO_2^-
e^-
$2H^+$
b
H_2O
Nitrite reductase
NO
2 NO
$2e^-$
c
Nitric oxide reductase
N_2O
$2e^-$
d
Nitrous oxide reductase
N_2

(18-30)



Periplasm
Membrane
Cytoplasm
Nitrate reductase
NO_3^-
$2H^+$
$2e^-$
H_2O
NO_2^-
$2H^+$
QH_2
Q
NADH
Dehydrogenase
$3H^+$
NADH
$2H^+$
NAD^+
QH_2
Q
Cytochrome *bc*
$4H^+$
$2H^+$
Cyt *c*
e^-
$2e^-$
NO_2^-
$2H^+$
H_2O
NO
Nitrite reductase
$2e^-$
2NO
$2H^+$
H_2O
N_2O
NO reductase
N_2O
$2H^+$
H_2O
N_2
N_2O reductase
NO_3^-/NO_2^- antiporter
NO_3^-
NO_2^-

Figure 18-21 Organization of the nitrate reduction system in the outer membrane of the bacterium *Paracoccus dinitrificans* as outlined by Blaut and Gottschalk.[315] The equations are not balanced as shown but will be balanced if two NO_3^- ions are reduced to N_2 by five molecules of NADH (see also Eq. 18-28). Although this will also require seven protons, about 20 additional H^+ will be pumped to provide for ATP synthesis.

More prevalent is **cytochrome *cd*$_1$** nitrite reductase.[340,343–346] The water-soluble periplasmic enzyme is a homodimer of ~60-kDa subunits, each containing a *c*-type heme in a small N-terminal domain and **heme *d*$_1$**, a ferric dioxoisobacteriochlorin (Fig. 16-6). The latter is present in the central channel of an eight-bladed β-sheet propeller[345-346g] similar to that in Fig. 15-23A. The heme d_1 is unusual in having its Fe atom ligated by a tyrosine hydroxyl oxygen, which may be displaced to allow binding of NO_3^-. The electron required for the reduction is presumably transferred from the electron transfer chain in the membrane to the heme *d*, via the heme *c* group.[347] Cytochrome cd_1 nitrite reductases have an unexpected second enzymatic activity. They catalyze the four-electron reduction of O_2 to H_2O, as does cytochrome *c* oxidase. However, the rate is much slower than that of nitrite reduction.[340,348]

The enzyme catalyzing the third step of Eq. 18-30 (step *c*), **nitric oxide reductase**, is an unstable membrane-bound protein cytochrome *bc* complex.[349,350]

It has been isolated as a two-subunit protein, but genetic evidence suggests the presence of additional subunits.[350] The small subunit is a cytochrome *c*, while the larger subunit is predicted to bind two protohemes as well as a nonheme iron center. This protein also shows sequence homology with cytochrome *c* oxidase. It contains no copper, but it has been suggested that a heme *b*–nonheme Fe center similar to the heme *a*–Cu_B center of cytochrome *c* oxidase may be present. It may be the site at which the nitrogen atoms of two molecules of NO are joined to form N_2O.[350,351] A different kind of NO reductase is utilized by the denitrifying fungus *Fusarium oxysporum*. It is a cytochrome P450 but with an unusually low redox potential (–0.307 V). This **cytochrome P450$_{nor}$** does not react with O_2 (as in Eq. 18-57) but binds NO to its heme Fe^{3+}, reduces the complex with two electrons from NADH, then reacts with a second molecule of NO to give N_2O and H_2O.[352]

Reduction of N_2O to N_2 by bacteria (Eq. 18-30, step *d*) is catalyzed by the copper-containing nitrous oxide reductase. The purple enzyme is a dimer of 66-kDa subunits, each containing four atoms of Cu.[353] It has spectroscopic properties similar to those of cytochome *c* oxidase and a dinuclear copper–thiolate center similar to that of Cu_A in cytochrome *c* oxidase (p. 1030). The nature of the active site is uncertain.[354]

Sulfate-reducing and sulfur-reducing bacteria.

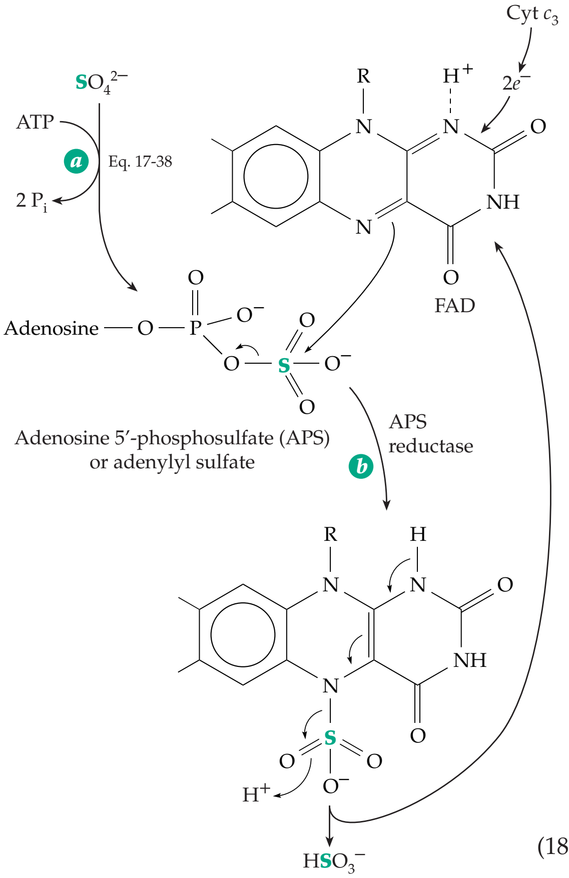
A few obligate anaerobes obtain energy by using sulfate ion as an oxidant.[355–356a] For example, *Desulfovibrio desulfuricans* catalyzes a rapid oxidation of H_2 with reduction of sulfate to H_2S (Eq. 18-31).

$$4\ H_2 + SO_4^{2-} + 2\ H^+ \rightarrow H_2S + 4\ H_2O$$
$$\Delta G'\ (\text{pH } 7) = -154\ \text{kJ mol}^{-1} \text{ of sulfate reduced} \tag{18-31}$$

While this may seem like an esoteric biological process, the reaction is quantitatively significant. For example, it has been estimated that within the Great Salt Lake basin bacteria release sulfur as H_2S in an amount of 10^4 metric tons (10^7 kg) per year.[357]

The reduction potential for sulfate is extremely low ($E°'$, pH 7 = –0.454 V), and organisms are not known to reduce it directly to sulfite. Rather, a molecule of ATP is utilized to form adenosine 5'-phosphosulfate (APS) through the action of **ATP sulfurylase** (ATP:sulfate adenylyltransferase, Eq. 17-38).[358,359] APS is then reduced by cytochrome c_3 (Eq. 18-32, step *b*). The 13-kDa low-potential ($E°'$, pH 7 = 0.21 V) cytochrome c_3 contains four heme groups (Figure 16-8C) and is found in high concentration in sulfate-reducing bacteria.[360,361] Some of these bacteria have larger polyheme cytochromes *c*.[361a] For example, *Desulfovibrio vulgaris* forms a 514-residue protein carrying 16 hemes organized as four cytochrome c_3-like domains.[362] Each heme in cytochrome c_3 has a distinct redox potential within the range −0.20 to −0.40 V.[361–363]

APS is reduced (Eq. 18-32, step *b*) by **APS reductase**, a 220-kDa iron–sulfur protein containing FAD and several Fe–S clusters. An intermediate in the reaction may be the adduct of sulfite with FAD, which may be formed as in Eq. 18-32. The initial step in this hypothetical mechanism is displacement on sulfur by a strong nucleophile generated by transfer of electrons from reduced ferredoxin to cytochrome c_3 to the flavin.[364]



(18-32)

Bisulfite produced according to Eq. 18-32 is reduced further by a **sulfite reductase,** which is thought to receive electrons from flavodoxin, cyt c_3, and a hydrogenase. ATP synthesis is coupled to the reduction. Sulfite reductases generally contain both siroheme and Fe_4S_4 clusters (Fig. 16-19). They appear to be able to carry out the 6-electron reduction to sulfide without accumulation of intermediates.[365,366] However, in contrast to the assimilatory sulfite reductases present in many organisms, the dissimilatory nitrite reductases of sulfur-reducing bacteria may also release some thiosulfate $S_2O_3^{2-}$.[367] A possible role of menaquinone (vitamin K_2), present in large amounts in *Desulfovibrio*, has been suggested.[311] Although *Desulfovibrio* can obtain their energy from Eq. 18-31, they are not true autotrophs and must utilize compounds such as acetate together with CO_2 as a carbon source.

Some thermophilic archaeobacteria are able to live with CO_2 as their sole source of carbon and reduction of elemental sulfur with H_2 (Eq. 18-33) as their sole source of energy.[368,369]

$$\begin{gathered} H_2 + S^0 \rightarrow H_2S \\ \Delta G^\circ = -27.4 \text{ kJ mol}^{-1} \end{gathered} \tag{18-33}$$

This is remarkable in view of the small standard Gibbs energy decrease. Some species of the archaeobacterium *Sulfolobus* are able either to live aerobically oxidizing sulfide to sulfate with O_2 (Eq. 18-22) or to live anaerobically using reduction of sulfur by Eq. 18-33 as their source of energy.[369]

The sulfate-reducing bacterium *Desulfovibrio sulfodismutans* carries out what can be described as "inorganic fermentations" which combine the oxidation of compounds such as sulfite or thiosulfate (as observed for sulfur-oxidizing bacteria; Eq. 18-22), with reduction of the same compounds (Eq. 18-34).[370,370a] Dismutation of $S_2O_3^{2-}$ plus H_2O to form SO_4^{2-} and H_2S also occurs but with a less negative Gibbs energy change.

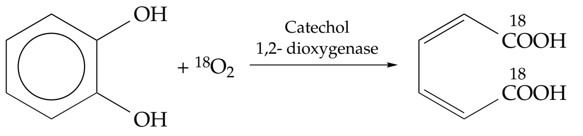
$$\begin{gathered} 4\ SO_3^{2-} + H^+ \rightarrow 3\ SO_4^{2-} + HS^- \\ \Delta G' \text{ (pH 7)} = -185 \text{ kJ/mol} \end{gathered} \tag{18-34}$$

A strain of *Pseudomonas* obtains all of its energy by reducing sulfate using phosphite, which is oxidized to phosphate.[370b]

Methane bacteria. The methane-producing bacteria (Chapter 15) are also classified as chemiautotrophic organisms. While they can utilize substances such as methanol and acetic acid, they can also reduce CO_2 to methane and water using H_2 (Fig. 15-22). The electron transport is from hydrogenase, perhaps through ferredoxin to formate dehydrogenase and via the deazaflavin F_{420} and $NADP^+$ to the methanopterin-dependent dehydrogenases that carry out the stepwise reduction of formate to methyl groups (Fig. 16-28). Generation of ATP probably involves proton pumps, perhaps in internal coupling membranes .[315,371]

F. Oxygenases and Hydroxylases

For many years the idea of dehydrogenation dominated thinking about biological oxidation. Many scientists assumed that the oxygen found in organic substances always came from water, e.g., by addition of water to a double bond followed by dehydrogenation of the resulting alcohol. Nevertheless, it was observed that small amounts of O_2 were essential, even to anaerobically growing cells.[372] In 1955, Hayaishi and Mason independently demonstrated that ^{18}O was sometimes incorporated into organic compounds directly from $^{18}O_2$ as in Eq. 18-35. Today a bewildering variety of **oxygenases** are known to function in forming such essential metabolites as sterols, prostaglandins, and active derivatives of vitamin D. Oxygenases are also needed in the catabolism of many substances, often acting on nonpolar groups that cannot be attacked readily by other types of enzyme.[372]

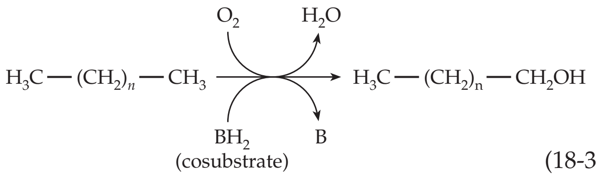


Catechol

Cis,cis-muconic acid

(18-35)

Oxygenases are classified either as **dioxygenases** or as **monooxygenases**. The monooxygenases are also called mixed function oxidases or **hydroxylases**. Dioxygenases catalyze incorporation of two atoms of oxygen as in Eq. 18-35, but monooxygenases incorporate only one atom. The other oxygen atom from the O_2 is converted to water. A typical monooxygenase-catalyzed reaction is the hydroxylation of an alkane to an alcohol (Eq. 18-36).



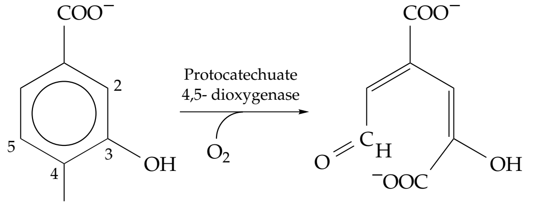
$$H_3C-(CH_2)_n-CH_3 \xrightarrow[BH_2 \text{ (cosubstrate)} \ \rightarrow \ B]{O_2 \ \rightarrow \ H_2O} H_3C-(CH_2)_n-CH_2OH \tag{18-36}$$

A characteristic of the monooxygenases is that an additional reduced substrate, a **cosubstrate** (BH_2 in Eq. 18-36), is usually required to reduce the second atom of the O_2 molecule to H_2O.

Since O_2 exists in a "triplet" state with two unpaired electrons, it reacts rapidly only with transition metal ions or with organic radicals (Chapter 16). For this reason, most oxygenases contain a transition metal ion, usually of iron or copper, or contain a cofactor, such as FAD, that can easily form a radical or act on a cosubstrate or substrate to form a free radical.

1. Dioxygenases

Among the best known of the oxygenases that incorporate both atoms of O_2 into the product are those that participate in the biological degradation of aromatic compounds by cleaving double bonds at positions between two OH groups as in Eq. 18-35 or adjacent to one OH group of an *ortho* or *para* hydroxyl pair.[373] A much studied example is **protocatechuate 3,4-dioxygenase**,[373–375] which cleaves its substrate between the two OH groups (*intradiol cleavage*) as in Eq. 18-35. A different enzyme, **protocatechuate 4,5-dioxygenase**,[357] cleaves the same substrate next to just one of the two OH groups (*extradiol cleavage*; Eq. 18-37) to form the aldehyde α-hydroxy-δ-carboxymuconic semialdehyde. Another extradiol cleaving enzyme, **protocatechuate 2,3-dioxygenase**, acts on the same substrate. Many other dioxygenases attack related substrates.[376–380] Intradiol-cleaving enzymes are usually iron-tyrosinate proteins (Chapter 16) in which the iron is present in the Fe(III) oxidation state and remains in this state throughout the catalytic cycle.[375] The enzymes usually have two subunits and no organic prosthetic groups. For example, a protocatechuate 3,4-dioxygenase from *Pseudomonas* has the composition $(\alpha\beta Fe)_{12}$ with subunit masses of 23 (α) and 26.5 (β) kDa. The iron is held in the active site cleft between the α and β subunits by Tyr 408, Tyr 447, His 460, and His 462 of the β subunit and a water molecule.[375] These enzymes and many other oxygenases probably assist the substrate in forming radicals that can react with O_2 to form organic peroxides. Some plausible intermediate species are pictured in Fig. 18-22. The reactions are depicted as occurring in two-electron steps. However, O_2 is a diradical, and it is likely that the Fe^{3+}, which is initially coordinated to both phenolate groups of the ionized substrate, assists in forming an organic free radical that reacts with O_2.



(18-37)

Extradiol dioxygenases have single Fe^{2+} ions in their active sites. The O_2 probably binds to the Fe^{2+} and may be converted transiently to an Fe^{3+}-superoxide complex which adds to the substrate. Some extradiol dioxygenases require an Fe_2S_2 ferredoxin to reduce any Fe^{3+}-enzyme that is formed as a side reaction back to the Fe^{2+} state.[381] Possible intermediates are given in Fig. 18-22 (left side) with two-electron steps used to save space and to avoid giving uncertain details about free radical intermediates. Formation of the organic radical is facilitated by the iron atom, which may be coordinated initially to both phenolate groups of the ionized substrate. The peroxide intermediates, for both types of dioxygenases, may react and be converted to various final products by several mechanisms.[382]

Tryptophan dioxygenase (indoleamine 2,3-dioxygenase)[383] is a heme protein which catalyzes the reaction of Eq. 18-38. The oxygen atoms designated by

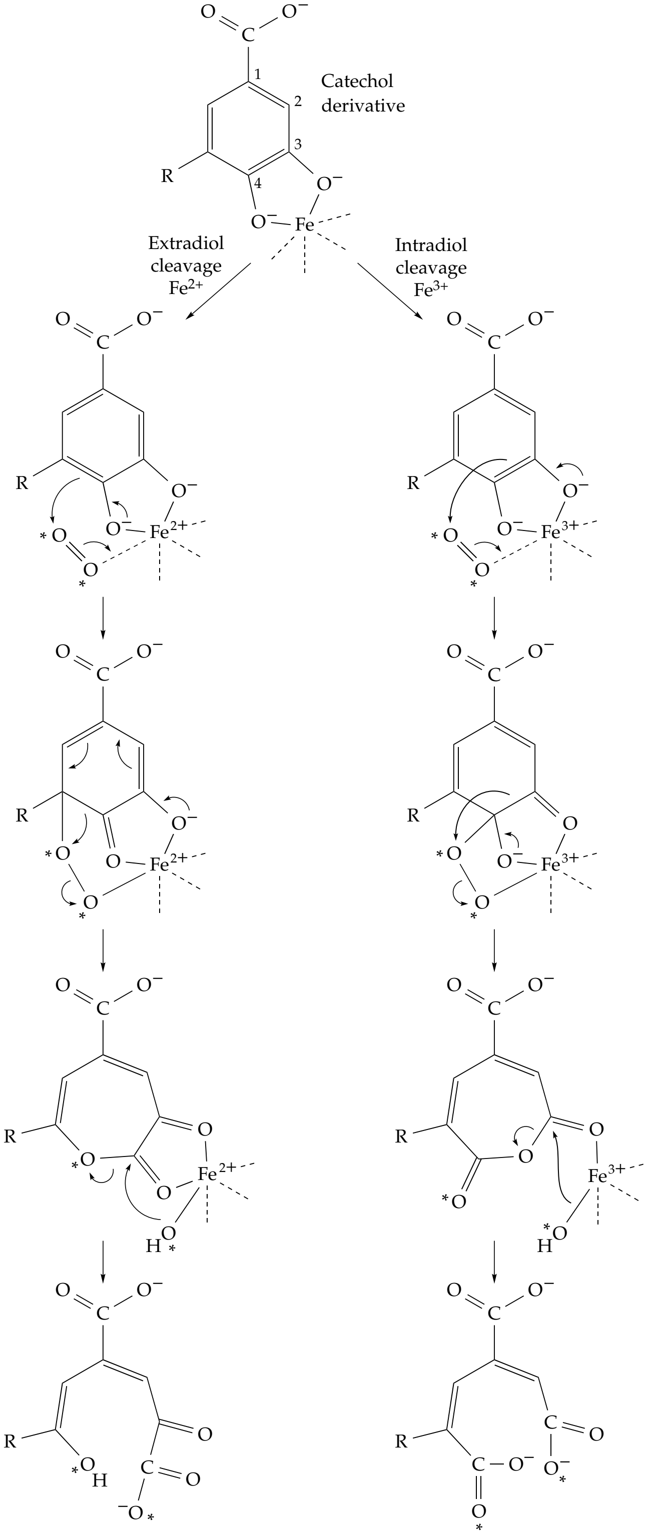
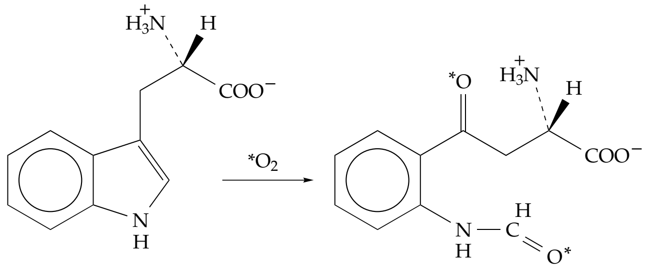


Figure 18-22 Some possible intermediates in the action of extradiol (left) and intradiol (right) aromatic dioxygenases. Although the steps depict the flow of pairs of electrons during the formation and reaction of peroxide intermediates, the mechanisms probably involve free radicals whose formation is initiated by O_2. The asterisks show how two atoms of labeled oxygen can be incorporated into final products. After Ohlendorf *et al.*[374]

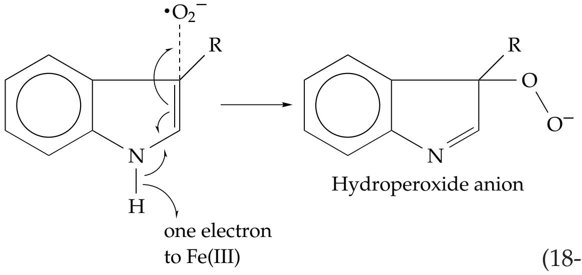
the asterisks are derived from O_2. Again, the first step is probably the formation of a complex between Fe(II) and O_2, but tryptophan must also be present before this can occur. At 5°C the enzyme, tryptophan, and O_2 combine to give an altered spectrum reminiscent of that of compound III of peroxidase (Fig. 16-14). This oxygenated complex may, perhaps, then be converted to a complex of Fe(II) and superoxide ion.



Tryptophan Formylkynurenine

(18-38)

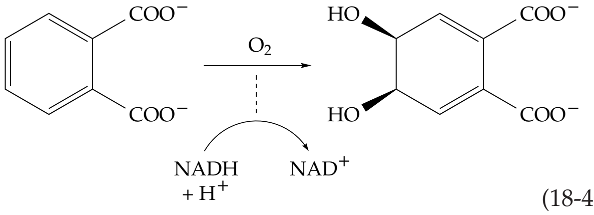
There is much evidence, including inhibition by superoxide dismutase and stimulation by added potassium superoxide,[384] that the superoxide anion radical is the species that attacks the substrate (Eq. 18-39). In this reaction one electron is returned to the Fe(III) form of the enzyme to regenerate the original Fe(II) form. Subsequent reaction of the hydroperoxide anion would give the observed products.



Hydroperoxide anion

(18-39)

Some dioxygenases require a cosubstrate. For example, **phthalate dioxygenase**[385] converts phthalate to a *cis*-dihydroxy derivative with NADH as the cosubstrate (Eq. 18-40). Similar double hydroxylation reactions catalyzed by soil bacteria are known for benzene, benzoate,[386] toluene, naphthalene, and several other aromatic compounds.[386a] The formation of the *cis*-glycols is usually followed by dehydrogenation or oxidative decarboxylation by NAD^+ to give a catechol, whose ring is then opened by another dioxygenase reaction (Chapter 25). An elimination of Cl^- follows dioxygenase action on *p*-chlorophenylacetate and produces 3,4-dihydroxyphenylacetate as a product. Pthalate dioxygenases consist of two subunits. The 50-kDa dioxygenase subunits receive electrons from reductase subunits that contain a Rieske-type Fe–S centers and bound FMN.[387] The dioxygenase also contains an Fe_2S_2 center, and electrons flow from NADH to FMN and through the two Fe–S centers to the Fe^{2+} of the active site.[387–388]



(18-40)

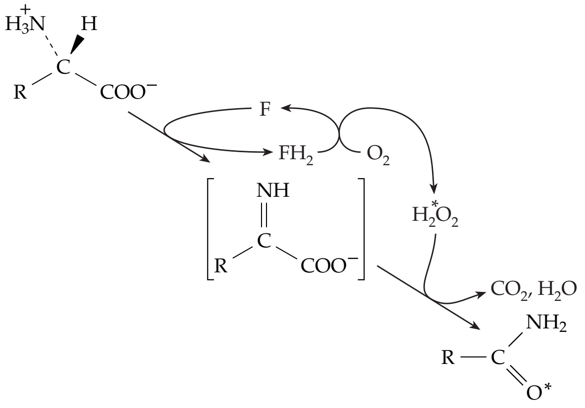
Lipoxygenases catalyze oxidation of polyunsaturated fatty acids in plant lipids. Within animal tissues the lipoxygenase-catalyzed reaction of arachidonic acid with O_2 is the first step in formation of **leukotrienes** and other mediators of inflammation. These reactions are discussed in Chapter 21.

2. Monooxygenases

Two classes of monooxygenases are known. Those requiring a cosubstrate (BH_2 of Eq. 18-36) in addition to the substrate to be hydroxylated are known as **external monooxygenases**. In the other group, the **internal monooxygenases**, some portion of the substrate being hydroxylated also serves as the cosubstrate. Many internal monooxygenases contain flavin cofactors and are devoid of metal ions.

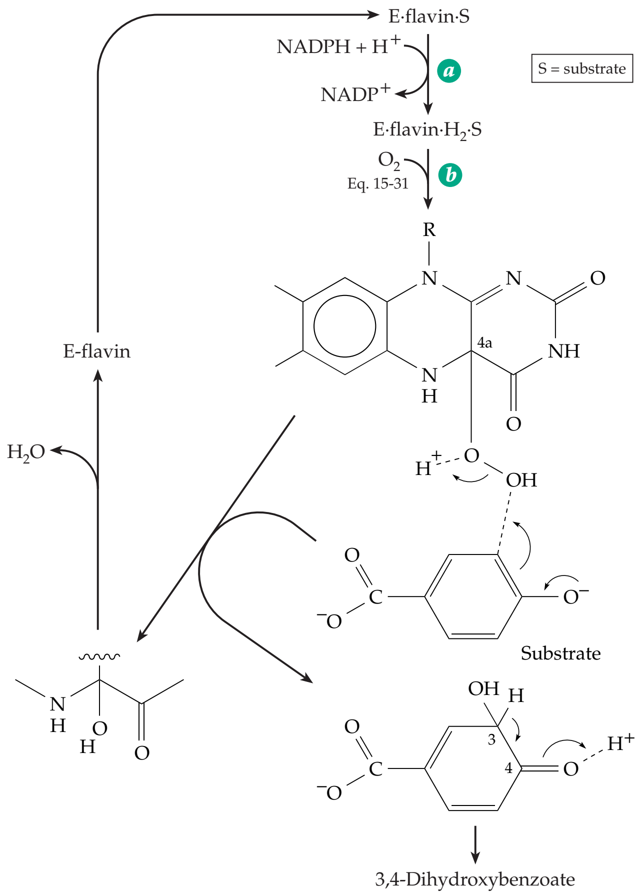
Flavin-containing monooxygenases. One group of flavin-dependent monooxygenases form H_2O_2 by reaction of O_2 with the reduced flavin and use the H_2O_2 to hydroxylate 2-oxoacids. An example is **lactate monooxygenase,** which apparently dehydrogenates lactate to pyruvate and then oxidatively decarboxylates the pyruvate to acetate with H_2O_2 (Eq. 15-36). One atom of oxygen from O_2 is incorporated into the acetate formed.[389,390] In a similar manner, the FAD-containing bacterial **lysine monooxygenase** probably catalyzes the sequence of reactions shown in Eq. 18-41.[391] When native lysine monooxygenase was treated with sulfhydryl-blocking reagents the resulting modified enzyme produced a 2-oxoacid, ammonia, and H_2O_2, just the products predicted from the hydrolytic decomposition of the bracketed intermediate of Eq. 18-41. Similar bacterial enzymes act on tryptophan and phenylalanine.[392]

NADPH can serve as a cosubstrate of flavoprotein monooxygenase by first reducing the flavin, after which the reduced flavin can react with O_2 to generate the hydroxylating reagent.[393] An example is the bacterial **4-hydroxybenzoate hydroxylase** which forms 3,4-dihydroxybenzoate.[394] The 43-kDa protein consists of three domains, the large FAD-binding domain being folded in nearly the same way as that of glutathione reductase (Fig. 15-10). The 4-hydroxybenzoate binds


$H_3\overset{+}{N}$
H
C
R
COO^-
F
FH_2
O_2
NH
C
R
COO^-
$H_2\overset{*}{O}_2$
CO_2, H_2O
NH_2
R — C
O*

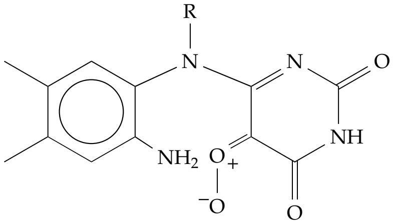
(18-41)

first into a deep cleft below the N-5 edge of the isoalloxazine ring of the FAD; then the NADH binds. Spectroscopic studies have shown the existence of at least three intermediates. The first of these has been identified as the 4a-peroxide whose formation (Eq. 15-31) is discussed in Chapter 15. The third intermediate is the corresponding 4a-hydroxyl compound. The substrate hydroxylation must occur in a reaction with the flavin peroxide, presumably with the phenolate anion form of the substrate (Eq. 18-42).[395] The initial hydroxylation product is tautomerized to form the product 3,4-dihydroxybenzoate.

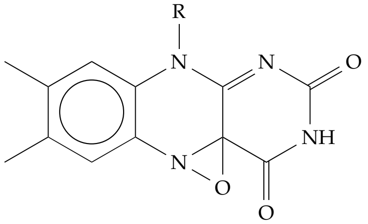

E·flavin·S
NADPH + H^+
$NADP^+$
a
S = substrate
E·flavin·H_2·S
O_2
Eq. 15-31
b
R
N
N
O
4a
NH
N
H
O
H^+
O
OH
E-flavin
H_2O
O
C
^-O
O^-
Substrate
N
H
O
H
O
OH
H
3
4
O
H^+
3,4-Dihydroxybenzoate

(18-42)

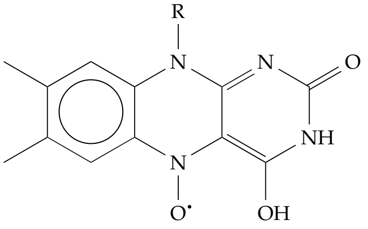
According to this mechanism, one of the two oxygen atoms in the hydroperoxide reacts with the aromatic substrate, perhaps as OH^+ or as a superoxide radical. A variety of mechanisms for activating the flavin peroxide to give a more potent hydroxylating reagent have been proposed. These include opening of the central ring of the flavin to give a carbonyl oxide intermediate which could transfer an oxygen atom to the substrate,[396] elimination of H_2O to form an **oxaziridine,**[397] or rearrangment to a **nitroxyl radical**.[398] Any of these might be an active electrophilic hydroxylating reagent. However, X-ray structural studies suggest that conformational changes isolate the substrate–FAD-enzyme complex from the medium stabilizing the 4a peroxide via hydrogen bonding[399–400] in close proximity to the substrate. Reaction could occur by the simple mechanism of Eq. 18-42, a mechanism also supported by ^{19}F NMR studies with fluorinated substrate analogs[401] and other investigations.[401a,b]


R
N
N
O
NH
NH_2
O^+
^-O
O

Carbonyl oxide

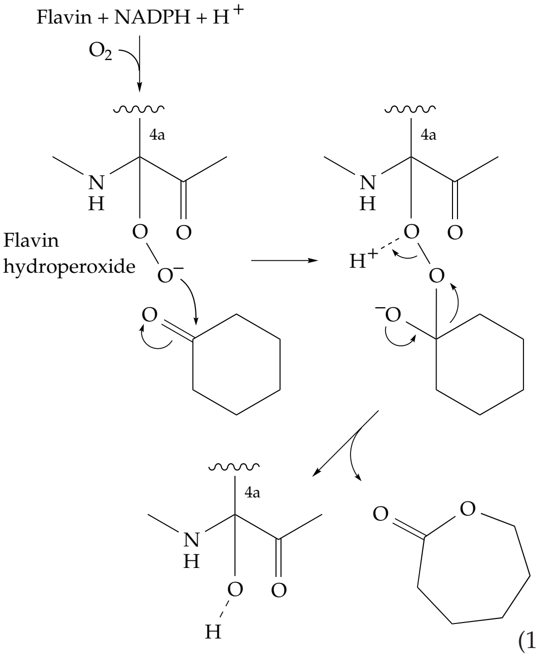


An oxaziridine


R
N
N
O
NH
N
O•
OH

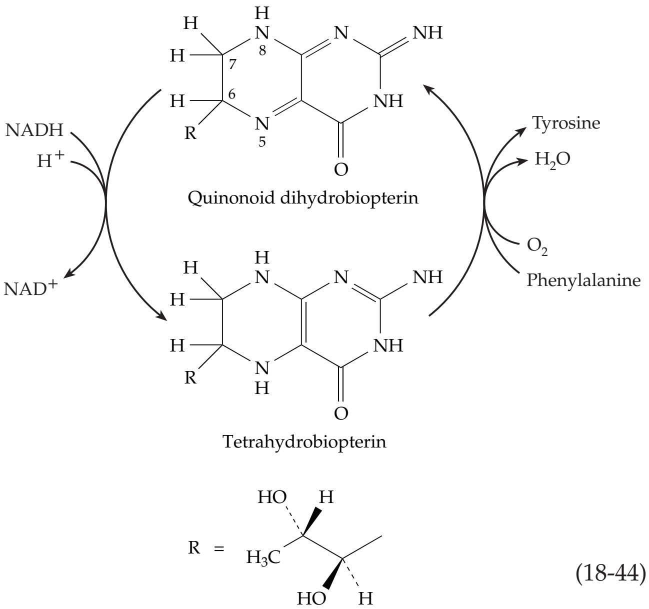
A nitroxyl radical

Related flavin hydroxylases act at nucleophilic positions on a variety of molecules[393,402] including phenol,[403] salicylate,[404] anthranilate,[405] *p*-cresol,[406] 4-hydroxyphenylacetate,[407,408] and 4-aminobenzoate.[409] Various microsomal flavin hydroxylases are also known.[410] Flavin peroxide intermediates are also able to hydroxylate some electrophiles.[411] For example, the bacterial **cyclohexanone oxygenase** catalyzes the ketone to lactone conversion of Eq. 18-43.[411a] The mechanism presumably involves the nucleophilic attack of the flavin hydroperoxide on the carbonyl group of the substrate followed by rearrangement. This parallels the Baeyer–Villiger rearrangement that results from treatment of ketones with peracids.[393] Cyclohexanone oxygenase also catalyzes a variety of other reactions,[412] including conversion of sulfides to sulfoxides.



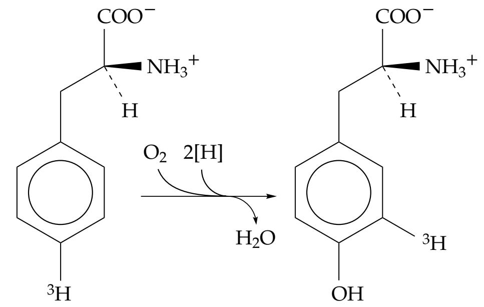
(18-43)

Reduced pteridines as cosubstrates. A dihydro form of biopterin (Fig. 15-17) serves as a cosubstrate, that is reduced by NADPH (Eq. 18-44) in hydroxylases that act on phenylalanine, tyrosine, and tryptophan.



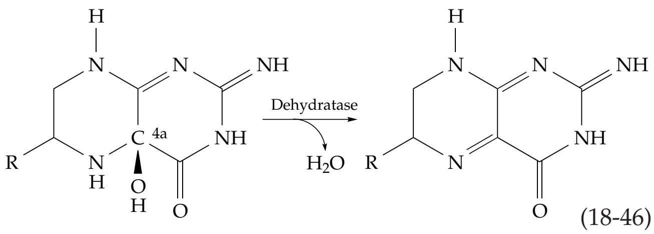
(18-44)

The tetrahydrobiopterin formed in this reaction is similar in structure to a reduced flavin. The mechanism of its interaction with O_2 could reasonably be the same as that of 4-hydroxybenzoate hydroxylase. However, **phenylalanine hydroxylase**, which catalyzes the formation of tyrosine (Eq. 18-45), a dimer of 451-residue subunits, contains one Fe per subunit,[413–415a] whereas flavin monooxygenases are devoid of iron. **Tyrosine hydroxylase**[416–419a] and **tryptophan hydroxylase**[420] have very similar properties. All three enzymes contain regulatory, catalytic, and tetramerization domains as well as a common Fe-binding motif in their active sites.[413,421,421a]



(18-45)

The role of the iron atom in these enzymes must be to accept an oxygen atom from the flavin peroxide, perhaps forming a reactive ferryl ion and transferring the oxygen atom to the substrate, e.g., as do cytochromes P450 (see Eq. 18-57). The 4*a*-hydroxytetrahydrobiopterin, expected as an intermediate if the mechanism parallels that of Eq. 18-42, has been identified by its ultraviolet absorption spectrum.[422] A ring-opened intermediate has also been ruled out for phenylalanine hydroxylase.[423] However, the 4*a* –OH adduct has been observed by ^{13}C-NMR spectroscopy. Its absolute configuration is 4*a*(S) and the observation of an ^{18}O-induced shift in the ^{13}C resonance of the 4*a*-carbon atom[424] confirms the origin of this oxygen from $^{18}O_2$ (see Eq. 18-42). A "stimulator protein" needed for rapid reaction of phenylalanine hydroxylase has been identified as a **4*a*-carbinolamine dehydratase** (Eq. 18-46).[425–426] This protein also has an unexpected function as part of a complex with transcription factor HNF1 which is found in nuclei of liver cells.[425a,426]



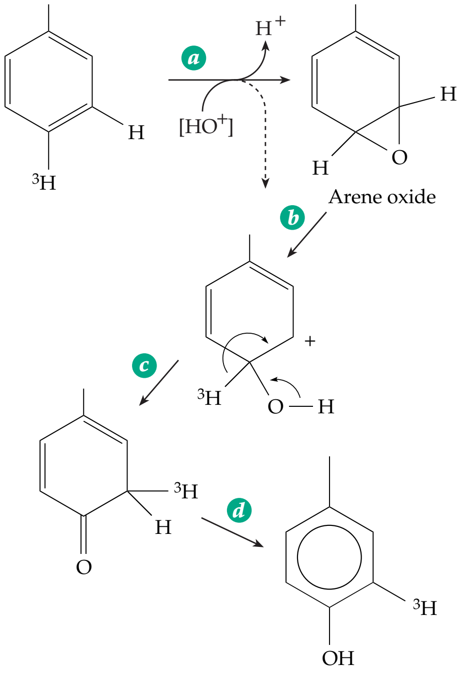
(18-46)

Dihydrobiopterin can exist as a number of isomers. The quinonoid form shown in Eqs. 18-44 and 18-46 is a tautomer of 7,8-dihydrobiopterin, the form generated by dihydrofolate reductase (Chapter 15). A pyridine nucleotide-dependent **dihydropteridine reductase**[427–429] catalyzes the left-hand reaction of Eq. 18-44.

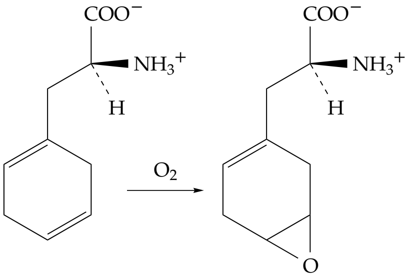
The hereditary absence of phenylalanine hydroxylase, which is found principally in the liver, is the cause of the biochemical defect **phenylketonuria** (Chapter 25, Section B).[430,430a] Especially important in the metabolism of the brain are tyrosine hydroxylase, which converts tyrosine to 3,4-dihydroxyphenylalanine, the rate-limiting step in biosynthesis of the catecholamines (Chapter 25), and tryptophan hydroxylase, which catalyzes formation of 5-hydroxytryptophan, the first step in synthesis of the neurotransmitter 5-hydroxytryptamine (Chapter 25). All three of the pterin-dependent hydroxylases are under complex regulatory control.[431,432] For example, tyrosine hydroxylase is acted on by at least four kinases with phosphorylation occurring at several sites.[431,433,433a] The kinases are responsive to nerve growth factor and epidermal growth factor,[434] cAMP,[435] Ca^{2+} + calmodulin, and Ca^{2+} + phospholipid (protein kinase C).[436] The hydroxylase is inhibited by its endproducts, the catecholamines,[435] and its activity is also affected by the availability of tetrahydrobiopterin.[436]

Hydroxylation-induced migration. A general result of enzymatic hydroxylation of aromatic compounds is the intramolecular migration of a hydrogen atom or of a substituent atom or group as is shown for the ^3H atom in Eq. 18-45.[437] Dubbed the NIH shift (because the workers discovering it were in a National Institutes of Health laboratory), the migration tells us something about possible mechanisms of hydroxylation. In Eq. 18-45 a tritium atom has shifted in response to the entering of the hydroxyl group. The migration can be visualized as resulting from electrophilic attack on the aromatic system, e.g., by an oxygen atom from Fe(N)=O or by OH^+ (Eq. 18-47).

Such an attack could lead in step *a* either to an **epoxide** (**arene oxide**) or directly to a carbocation as shown in Eq. 18-47. Arene oxides can be converted, via the carbocation step *b*, to end products in which the NIH shift has occurred.[438] The fact that phenylalanine hydroxylase also catalyzes the conversion of the special substrate shown in Eq. 18-48 to a stable epoxide, which cannot readily undergo ring opening, also supports this mechanism.

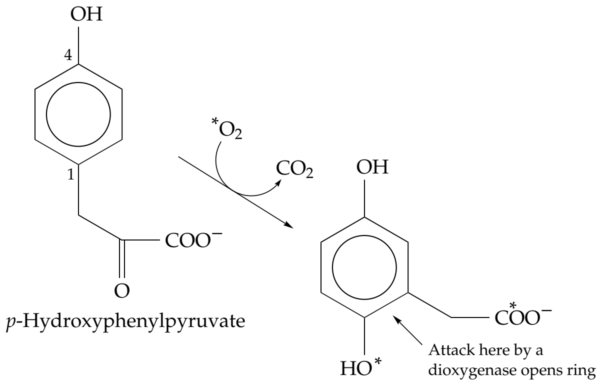


(18-47)



(18-48)

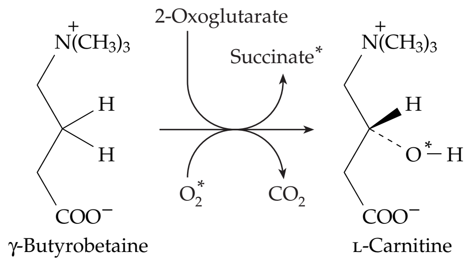
Operation of the NIH shift can cause migration of a large substituent as is illustrated by the hydroxylation of 4-hydroxyphenylpyruvate (Eq. 18-49), a key step in the catabolism of tyrosine (Chapter 25). Human 4-hydroxyphenylpyruvate dioxygenase is a dimer of 43-kDa subunits.[439] A similar enzyme from *Pseudomonas* is a 150-kDa tetrameric iron-tyrosinate protein, which must be maintained in the reduced Fe(II) state for catalytic activity.[440] Although this enzyme is a dioxygenase, it is probably related in its mechanism of action to the 2-oxoglutarate-dependent monooxygenases discussed in the next section (Eqs. 18-51, 18-52). It probably uses the oxoacid side chain of the substrate to generate a reactive oxygen intermediate such as Fe(IV)=O by the decarboxylative mechanism of Eqs. 18-50 and 18-51. The iron-bound oxygen attacks C1 of



(18-49)

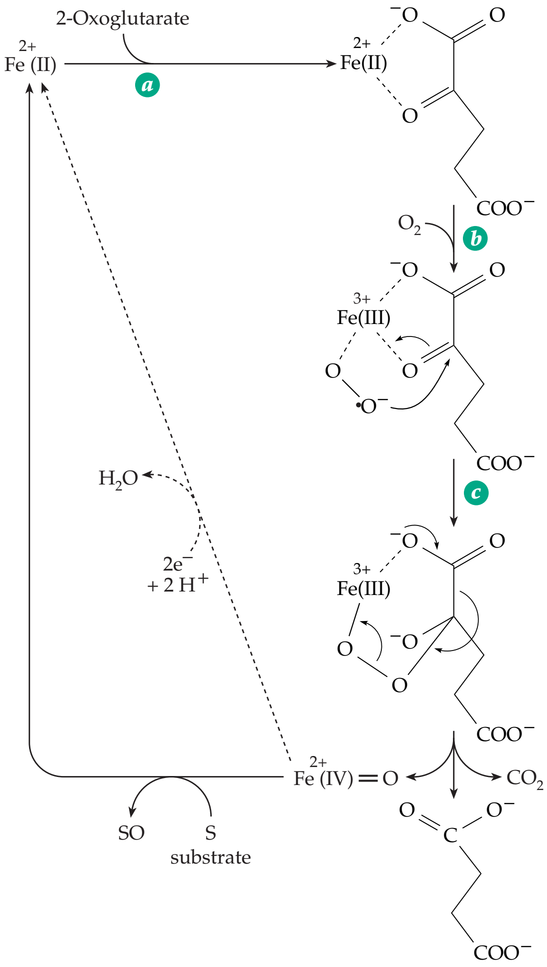
the aromatic ring, the electron-donating *p*-hydroxyl group assisting. This generates a hydroxylated carbocation of the type shown in Eq. 18-47 in which the whole two-carbon side chain undergoes the NIH shift.

2-Oxoglutarate as a decarboxylating cosubstrate. Several oxygenases accept hydrogen atoms from 2-oxoglutarate, which is decarboxylated in the process to form succinate. Among these are enzymes catalyzing hydroxylation of residues of proline in both the 3- and 4-positions (Eq. 8-6)[441–444] and of lysine in the 5-position (Eq. 8-7)[445,446] in the collagen precursor **procollagen**. The hydroxylation of prolyl residues also takes place within the cell walls of plants.[447] Similar enzymes hydroxylate the β-carbon of aspartyl or asparaginyl side chains in EGF domains (Table 7-3) of proteins.[441] Thymine[448] and taurine[449,449a] are acted on by related dioxygenases. A bacterial oxygenase initiates the degradation of the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) using another 2-oxoglutarate-dependent hydroxylase.[450,451] In the human body a similar enzyme hydroxylates γ-butyrobetaine to form carnitine (Eq. 18-50).[452] All of these enzymes contain iron and require ascorbate, whose function is apparently to prevent the oxidation of the iron to the Fe(III) state.


2-Oxoglutarate
Succinate*
O_2^*
CO_2
$\overset{+}{N}(CH_3)_3$
H
H
COO⁻
γ-Butyrobetaine
$\overset{+}{N}(CH_3)_3$
H
O*–H
COO⁻
L-Carnitine

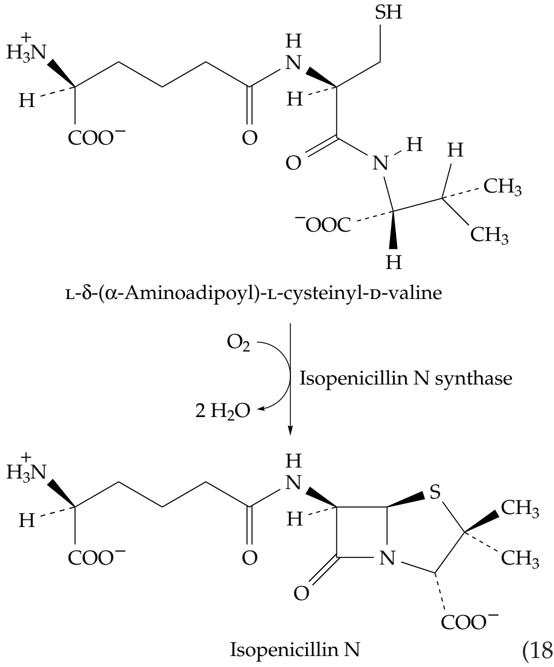
(18-50)

When $^{18}O_2$ is used for the hydroxylation of γ-butyrobetaine (Eq. 18-51), one atom of ^{18}O is found in the carnitine and one in succinate. The reaction is stereospecific and occurs with retention of configuration at C-3, the *pro*-R hydrogen being replaced by OH while the *pro*-S hydrogen stays.[453] Under some conditions these enzymes decarboxylate 2-oxoglutarate in the absence of a hydroxylatable substrate, the iron being oxidized to Fe^{3+} and ascorbate being consumed stoichiometrically.[454] A plausible mechanism (Eq. 18-51) involves formation of an Fe(II)–O_2 complex, conversion to $Fe(III)^+\ {}^\bullet O_2^-$, and addition of the superoxide ion to 2-oxoglutarate to form an adduct.[451,455] Decarboxylation of this adduct could generate the oxidizing reagent, perhaps Fe(IV)=O. In the absence of substrate S the ferryl iron could be reconverted to Fe(II) by a suitable reductant such as ascorbate. In the absence of ascorbate the Fe(IV) might be reduced to a catalytically inert Fe(III) form.


2-Oxoglutarate
Fe^{2+} (II)
a
Fe^{2+}(II)
⁻O
O
O
COO⁻
O_2
b
Fe^{3+}(III)
•O⁻
COO⁻
c
H_2O
2e⁻
+ 2 H⁺
Fe^{3+}(III)
COO⁻
Fe^{2+} (IV)=O
CO_2
SO
S
substrate
C
O⁻
COO⁻

(18-51)

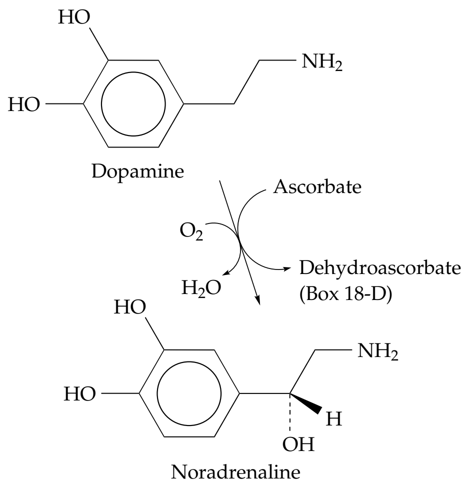
An unusual oxygenase with a single Fe^{2+} ion in its active site closes the four-membered ring in the biosynthesis of penicillins (Eq. 18-52). It transfers four


SH
$H_3\overset{+}{N}$
H
N
H
COO⁻
O
O
N
H
H
CH_3
CH_3
⁻OOC
H
L-δ-(α-Aminoadipoyl)-L-cysteinyl-D-valine
O_2
Isopenicillin N synthase
2 H_2O
$H_3\overset{+}{N}$
H
N
H
S
CH_3
CH_3
COO⁻
O
O
N
COO⁻
Isopenicillin N

(18-52)

hydrogen atoms from its dipeptide substrate to form two molecules of water and the product isopenicillin N.[456,457] Sequence comparison revealed several regions including the Fe-binding sites that are homologous with the oxoacid-dependent oxygenases. A postulated mechanism for **isopenicillin N synthase** involves formation of an Fe^{3+} superoxide anion complex as in Eq. 18-51. However, instead of attack on an oxoacid as in Eq. 18-51, it removes a hydrogen from the substrate to initiate the reaction sequence.[457] Other related oxygenases include **aminocyclopropane-1-carboxylate oxidase** (Eq. 24-35); **deacetoxycephalosporin C synthase**,[457a] an enzyme that converts penicillins to cephalosporins (Box 20-G); and **clavaminate synthase**,[458,459] an enzyme needed for synthesis of the β-lactamase inhibitor clavulanic acid, and **clavaminate synthase**.[458,459] This 2-oxoglutarate-dependent oxygenase catalyzes three separate reactions in the synthesis of the clinically important β-lactamase inhibitor clavulanic acid. The first step is similar to that in Eq. 18-50. The second is an oxidative cyclization and the third a desaturation reaction.

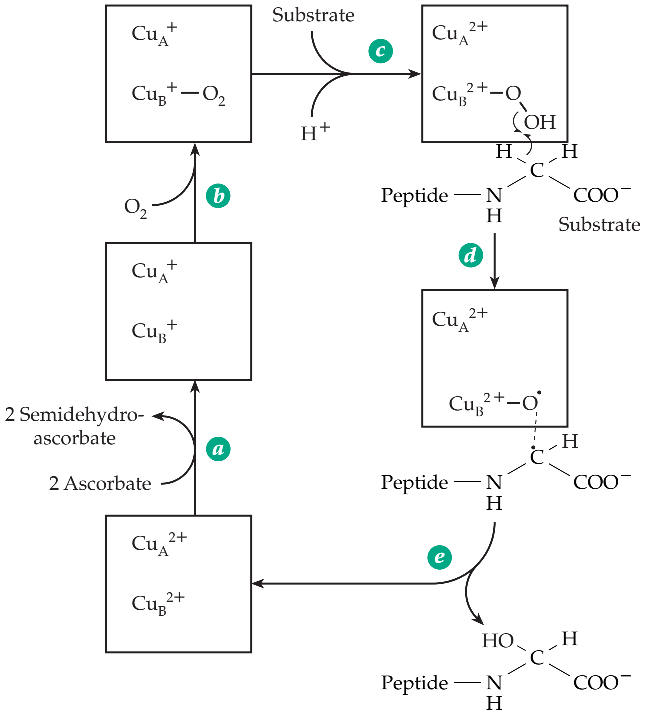
Copper-containing hydroxylases. Many Fe(II)-containing hydroxylases require a reducing agent to maintain the iron in the reduced state, and ascorbate is often especially effective. In addition, ascorbate is apparently a true cosubstrate for the copper-containing **dopamine β-hydroxylase**, an enzyme required in the synthesis of noradrenaline according to Eq. 18-53. This reaction takes place in neurons of the brain and in the adrenal gland, a tissue long known as especially rich in ascorbic acid. The reaction requires two molecules of ascorbate, which are converted in two one-electron steps to **semidehydroascorbate**.[460] Both the structure of this free radical and that of the fully oxidized form of vitamin C, **dehydroascorbic acid**, are shown in Box 18-D. Dopamine β-hydroxylase is a 290-kDa tetramer, consisting of a pair of identical disulfide-crosslinked homodimers, which contains two Cu ions per subunit.[461]



(18-53)

A similar copper-dependent hydroxylase constitutes the N-terminal domain of the **peptidylglycine α-amidating enzyme** (Eq. 10-11). This bifunctional enzyme hydroxylates C-terminal glycines in a group of neuropeptide hormones and other secreted peptides. The second functional domain of the enzyme cleaves the hydroxylated glycine to form a C-terminal amide group and glyoxylate.[462–464b] The three-dimensional structure of a 314-residue catalytic core of the hydroxylase domain is known.[463] Because of similar sequences and other properties, the structures of this enzyme and of dopamine β-hydroxylase are thought to be similar. The hydroxylase domain of the α-amidating enzyme is folded into two eight-stranded antiparallel jelly-roll motifs, each of which binds one of the two copper ions. Both coppers can exist in a Cu(II) state and be reduced by ascorbate to Cu(I). One Cu (Cu_A) is held by three imidazole groups and is thought to be the site of interaction with ascorbate. The other copper, Cu_B, which is 1.1 nm away from Cu_A, is held by two imidazoles. The substrate binds adjacent to Cu_B.[463]

The reaction cycle of these enzymes begins with reduction of both coppers from Cu(II) to Cu(I) (Eq. 18-54, step *a*). Both O_2 and substrate bind (steps *b* and *c*, but not necessarily in this order). The O_2 bound to Cu_B is reduced to a peroxide anion that remains bound to Cu_B. Both Cu_A and Cu_B donate one electron, both being oxidized to Cu(II). These changes are also included in step *c* of Eq. 18-54. One proposal is that the resulting peroxide is cleaved homolytically while removing the *pro-S* hydrogen of the glycyl residue.



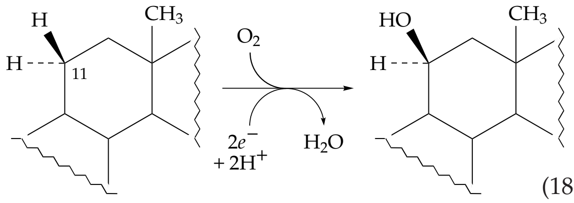
(18-54)

The resulting glycyl radical couples with the oxygen radical that is bound to Cu_B (step *e*).

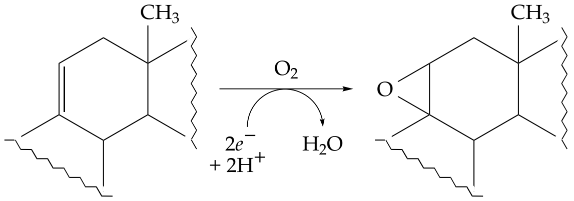
A variety of other copper hydroxylases are known. For example, **tyrosinase**, which contains a binuclear copper center, catalyzes both hydroxylation of phenols and aromatic amines and dehydrogenation of the resulting catechols or *o*-aminophenols (Eq. 16-57). As in hemocyanin, the O_2 is thought to be reduced to a peroxide which bridges between the two copper atoms. Methane-oxidizing bacteria, such as *Methylococcus capsulatus*, oxidize methane to methanol to initiate its metabolism. They do this with a copper-containing membrane-embedded monooxygenase whose active site is thought to contain a trinuclear copper center. Again a bridging peroxide may be formed and may insert an oxygen atom into the substrate.[465,466] The same bacteria produce a soluble methane monooxygenase containing a binuclear iron center.

Hydroxylation with cytochrome P450. An important family of heme-containing hydroxylases, found in most organisms from bacteria to human beings, are the cytochromes P450. The name comes from the fact that in their reduced forms these enzymes form a complex with CO that absorbs at 450 nm. In soil bacteria cytochromes P450 attack compounds of almost any structure. In the adrenal gland they participate in steroid metabolism,[467,468] and in the liver microsomal cytochromes P450 attack drugs, carcinogens, and other xenobiotics (foreign compounds).[469–471] They convert cholesterol to bile acids[472] and convert vitamin D,[473] prostaglandins, and many other metabolites to more soluble and often biologically more active forms. In plants cytochromes P450 participate in hydroxylation of fatty acids at many positions.[474] They play a major role in the biosynthetic phenylpropanoid pathway (Fig. 25-8) and in lignin synthesis.[475] More than 700 distinct isoenzyme forms have been described.[476,476a]

Cytochromes P450 are monooxygenases whose cosubstrates, often NADH or NADPH, deliver electrons to the active center heme via a separate flavoprotein and often via an iron-sulfur protein as well.[476a,b] A typical reaction (Eq. 18-55) is the 11 β-hydroxylation of a steroid, an essential step in the biosynthesis of steroid hormones (Fig. 22-11). The hydroxyl group is introduced without inversion of configuration. The same enzyme converts unsaturated derivatives to epoxides (Eq. 18-56), while other cytochromes P450 epoxidize olefins.[477] Epoxide hydrolases, which act by a mechanism related to haloalkane dehalogenase (Fig. 12-1), convert the epoxides to diols.[478] Cytochromes P450 are able to catalyze a bewildering array of other reactions[479–481] as well. Most of these, such as conversion of amines and thioesters to *N*- or *S*-oxides, also involve transfer of an oxygen atom to the substrate. Others, such as the reduction of epoxides, *N*-oxides, or nitro compounds, are electron-transfer reactions.



(18-55)

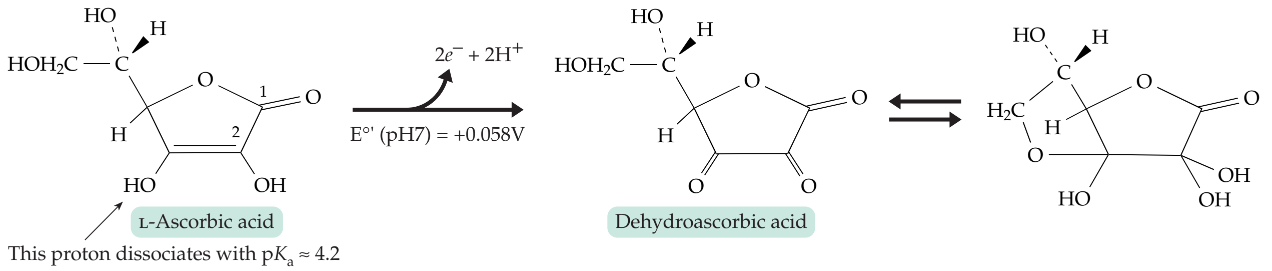


(18-56)

Several different cytochromes P450 are present in mammalian livers.[470] All are bound to membranes of the endoplasmic reticulum and are difficult to solubilize. Biosynthesis of additional forms is induced by such agents as phenobarbital,[470] 3-methylcholanthrene,[469] dioxin,[482] and ethanol.[483] These substances may cause as much as a 20-fold increase in P450 activity. Another family of cytochrome P450 enzymes is present in mitochondria.[483a] A large number of cytochrome P450 genes have been cloned and sequenced. Although they are closely related, each cytochrome P450 has its own gene. There are at least ten families of known P450 genes and the total number of these enzymes in mammals may be as high as 200. Microorganisms, from bacteria to yeast, produce many other cytochromes P450.

Microsomal cytochromes P450 receive electrons from an **NADPH-cytochrome P450** reductase, a large 77-kDa protein that contains one molecule each of FAD and FMN.[484–485a] It is probably the FAD which accepts electrons from NADPH and the FMN which passes them on to the heme of cytochrome P450. Cytochrome b_5 is also reduced by this enzyme,[486] and some cytochromes P450 may accept one electron directly from the flavin of the reductase and the second electron via cytochrome b_5. However, most bacterial and mitochondrial cytochromes P450 accept electrons only from small iron–sulfur proteins. Those of the adrenal gland receive electrons from the 12-kDa **adrenodoxin**.[487,488] This small protein of the ferredoxin class contains one Fe_2S_2 cluster and is, therefore, able to transfer electrons one at a time from the FAD-containing NADPH-adrenodoxin reductase[489] to the cytochrome P450. The camphor 5-monoxygenase from *Pseudomonas putida* consists of three components: an FAD-containing reductase, the Fe_2S_2 cluster-containin **putidaredoxin**,[489a] and cytochrome $P450_{cam}$.[490] Some other bacterial

BOX 18-D VITAMIN C: ASCORBIC ACID



L-Ascorbic acid

Dehydroascorbic acid

This proton dissociates with $pK_a \approx 4.2$

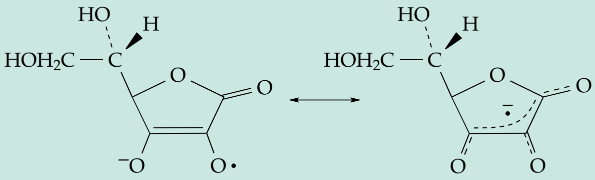
Hemorrhages of skin, gums, and joints were warnings that death was near for ancient sea voyagers stricken with **scurvy**. It was recognized by the year 1700 that the disease could be prevented by eating citrus fruit, but it was 200 years before efforts to isolate vitamin C were made. Ascorbic acid was obtained in crystalline form in 1927,[a–e] and by 1933 the structure had been established. Only a few vertebrates, among them human beings, monkeys, guinea pigs, and some fishes, require ascorbic acid in the diet; most species are able to make it themselves. Compared to that of other vitamins, the nutritional requirement is large.[e] Ten milligrams per day prevents scurvy, but subclinical deficiency, as judged by fragility of small capillaries in the skin, is present at that level of intake. "Official" recommendations for vitamin C intake have ranged from 30 to 70 mg / day. A more recent study[f] suggests 200 mg / day, a recommendation that is controversial.[g]

The biological functions of vitamin C appear to be related principally to its well-established reducing properties and easy one-electron oxidation to a free radical or two-electron reduction to **dehydroascorbic acid**. The latter is in equilibrium with the hydrated hemiacetal shown at the beginning of this box as well as with other chemical species.[h–j] Vitamin C is a weak acid which also has metal complexing properties.

Ascorbate, the anion of ascorbic acid, tends to be concentrated in certain types of animal tissues and may reach 3 mM or more in leukocytes, in tissues of eyes and lungs, in pituitary, adrenal, and parotid glands,[k,l] and in gametes.[m] Uptake into vesicles of the endoplasmic reticulum may occur via glucose transporters.[n] Ascorbate concentrations are even higher in plants and may exceed 10 mM in chloroplasts.[o] In animals the blood plasma ascorbate level of 20–100 µM is tightly controlled.[p,q] Cells take up ascorbate but any excess is excreted rapidly in the urine.[q] Both in plasma and within cells most vitamin C exists as the reduced form, ascorbate. When it is formed, the oxidized dehydroascorbate is reduced back to ascorbate or is degraded. The lactone ring is readily hydrolyzed to 2,3-dioxogulonic acid, which can undergo decarboxylation and oxidative degradation, one product being oxalate (see Fig. 20-2).[r] Tissues may also contain smaller amounts of L-ascorbic acid 2-sulfate, a compound originally discovered in brine shrimp. It is more stable than free ascorbate and may be hydrolyzed to ascorbate in tissues.[s]

In the chromaffin cells of the adrenal glands and in the neurons that synthesize catecholamines as neurotransmitters, ascorbate functions as a cosubstrate for dopamine β-hydroxylase (Eq. 18-53).[t,u] In fibroblasts it is required by the prolyl and lysyl hydroxylases and in hepatocytes by homogentisate dioxygenase (Eq. 18-49). Any effect of ascorbic acid in preventing colds may be a result of increased hydroxylation of procollagen and an associated stimulation of procollagen secretion.[v] High levels of ascorbate in guinea pigs lead to more rapid healing of wounds.[w] An important function of ascorbate in the pituitary and probably in other endocrine glands is in the α-amidation of peptides (Eq. 10-11).[x,y] Together with Fe(II) and O_2 ascorbate is a powerful nonenzymatic hydroxylating reagent for aromatic compounds. Like hydroxylases, the reagent attacks nucleophilic sites, e.g., converting phenylalanine to tyrosine. Oxygen atoms from $^{18}O_2$ are incorporated into the hydroxylated products. While H_2O_2 is formed in the reaction mixture, it cannot replace ascorbate. The relationship of this system to biochemical functions of ascorbate is not clear. An unusual function for vitamin C has been proposed for certain sponges that are able to etch crystalline quartz (SiO_2) particles from sand or rocks.[z]

Ascorbate is a major antioxidant, protecting cells and tissues from damage by free radicals, peroxides, and other metabolites of O_2.[p,r,aa,bb] It is chemically suited to react with many biologically important radicals and is present in high enough concentrations to be effective. It probably functions in cooperation with glutathione (Box 11-B),[cc] α-tocopherol (Fig. 15-24),[dd] and lipoic acid.[ee] Ascorbate can react with radicals in one-election transfer reactions to give the monodehydroascorbate radical[aa]:



BOX 18-D (continued)

Two ascorbate radicals can react with each other in a disproportionation reaction to give ascorbate plus dehydroascorbate. However, most cells can reduce the radicals more directly. In many plants this is accomplished by NADH + H^+ using a flavoprotein **monodehydroascorbate reductase**.[o] Animal cells may also utilize NADH or may reduce dehydroascorbate with reduced glutathione.[cc,ff] Plant cells also contain a very active blue copper ascorbate oxidase (Chapter 16, Section D,5), which catalyzes the opposite reaction, formation of dehydroascorbate.[gg] A heme ascorbate oxidase has been purified from a fungus.[hh] Action of these enzymes initiates an oxidative degradation of ascorbate, perhaps through the pathway of Fig. 20-2.

Ascorbate can also serve as a signal. In cultured cells, which are usually deficient in vitamin C, addition of ascorbate causes an enhanced response to added iron, inducing synthesis of the iron storage protein ferritin.[ii] Ascorbate indirectly stimulates transcription of procollagen genes[v] and decreases secretion of insulin by the pancreas.[jj] However, since its concentration in blood is quite constant this effect is not likely to cause a problem for a person taking an excess of vitamin C.

Should we take extra vitamin C to protect us from oxygen radicals and slow down aging? Linus Pauling, who recommended an intake of 0.25–10 g / day, maintained that ascorbic acid also has a specific beneficial effect in preventing or ameliorating symptoms of the common cold.[j,kk] However, critics point out that unrecognized hazards may exist in high doses of this seemingly innocuous compound. Ascorbic acid has antioxidant properties, but it also promotes the generation of free radicals in the presence of Fe(III) ions, and it is conceivable that too much may be a bad thing.[ll] Catabolism to oxalate may promote formation of calcium oxalate kidney stones. Under some conditions products of dehydroascorbic acid breakdown may accumulate in the lens and contribute to cataract formation.[i,mm,nn] However, dehydroascorbate, or its decomposition products, apparently *protects* low-density lipoproteins against oxidative damage.[bb] Pauling pointed out that nonhuman primates synthesize within their bodies many grams of ascorbic acid daily, and that there is little evidence for toxicity. Pauling's claim that advanced cancer patients are benefited by very high (10 g daily) doses of vitamin C has been controversial, and some studies have failed to substantiate the claim.[oo,pp]

---

[a] Hughes, R. E. (1983) *Trends Biochem. Sci.* **8**, 146–147

[b] Staudinger, H. J. (1978) *Trends Biochem. Sci.* **3**, 211–212

[c] Szent-Györgyi, A. (1963) *Ann. Rev. Biochem.* **32**, 1–14

[d] Bradford, H. F. (1987) *Trends Biochem. Sci.* **12**, 344–347

[e] Packer, L., and Fuchs, J., eds. (1997) *Vitamin C in Health and Disease*, Dekker, New York

[f] Levine, M., Conry-Cantilena, C., Wang, Y., Welch, R. W., Washko.PW, Dhariwal, K. R., Park, J. B., Lazarev, A., Graumlich, J. F., King, J., and Cantilena, L. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3704–3709

[g] Young, V. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14344–14348

[h] Hroslif, J., and Pederson, B. (1979) *Acta Chem. Scand.* **B33**, 503–511

[i] Counsell, J. N., and Horing, D. H., eds. (1981) *Vitamin C*, Applied Sciences Publ., London

[j] Burns, J. J., Rivers, J. M., and Machlin, L. J., eds. (1987) *Third Conference on Vitamin C*, Vol. 498, New York Academy of Sciences, New York

[k] Washko, P. W., Wang, Y., and Levine, M. (1993) *J. Biol. Chem.* **268**, 15531–15535

[l] Rosen, G. M., Pou, S., Ramos, C. L., Cohen, M. S., and Britigan, B. E. (1995) *FASEB J.* **9**, 200–209

[m] Moreau, R., and Dabrowski, K. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10279–10282

[n] Bánhegyi, G., Marcolongo, P., Puskás, F., Fulceri, R., Mandl, J., and Benedetti, A. (1998) *J. Biol. Chem.* **273**, 2758–2762

[o] Sano, S., Miyake, C., Mikami, B., and Asada, K. (1995) *J. Biol. Chem.* **270**, 21354–21361

[p] May, J. M., Qu, Z.-c, and Whitesell, R. R. (1995) *Biochemistry* **34**, 12721–12728

[q] Vera, J. C., Rivas, C. I., Velásquez, F. V., Zhang, R. H., Concha, I. I., and Golde, D. W. (1995) *J. Biol. Chem.* **270**, 23706–23712

[r] Rose, R. C., and Bode, A. M. (1993) *FASEB J.* **7**, 1135 – 1142

[s] Benitez, L. V., and Halver, J. E. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 5445–5449

[t] Dhariwal, K. R., Shirvan, M., and Levine, M. (1991) *J. Biol. Chem.* **266**, 5384–5387

[u] Tian, G., Berry, J. A., and Klinman, J. P. (1994) *Biochemistry* **33**, 226–234

[v] Chojkier, M., Houglum, K., Solis-Herruzo, J., and Brenner, D. A. (1989) *J. Biol. Chem.* **264**, 16957–16962

[w] Harwood, R., Grant, M. E., and Jackson, D. S. (1974) *Biochem. J.* **142**, 641–651

[x] Bradbury, A. F., and Smyth, D. G. (1991) *Trends Biochem. Sci.* **16**, 112–115

[y] Eipper, B. A., Milgram, S. L., Husten, E. J., Yun, H.-Y., and Mains, R. E. (1993) *Protein Sci.* **2**, 489–497

[z] Bavestrello, G., Arillo, A., Benatti, U., Cerrano, C., Cattaneo-Vietti, R., Cortesogno, L., Gaggero, L., Giovine, M., Tonetti, M., and Sarà, M. (1995) *Nature (London)* **378**, 374–376

[aa] Kobayashi, K., Harada, Y., and Hayashi, K. (1991) *Biochemistry* **30**, 8310–8315

[bb] Retsky, K. L., Freeman, M. W., and Frei, B. (1993) *J. Biol. Chem.* **268**, 1304–1309

[cc] Ishikawa, T., Casini, A. F., and Nishikimi, M. (1998) *J. Biol. Chem.* **273**, 28708–28712

[dd] May, J. M., Qu, Z.-c, and Morrow, J. D. (1996) *J. Biol. Chem.* **271**, 10577–10582

[ee] Lykkesfeldt, J., Hagen, T. M., Vinarsky, V., and Ames, B. N. (1998) *FASEB J.* **12**, 1183–1189

[ff] May, J. M., Cobb, C. E., Mendiratta, S., Hill, K. E., and Burk, R. F. (1998) *J. Biol. Chem.* **273**, 23039–23045

[gg] Gaspard, S., Monzani, E., Casella, L., Gullotti, M., Maritano, S., and Marchesini, A. (1997) *Biochemistry* **36**, 4852–4859

[hh] Kim, Y.-R., Yu, S.-W., Lee, S.-R., Hwang, Y.-Y., and Kang, S.-O. (1996) *J. Biol. Chem.* **271**, 3105–3111

[ii] Toth, I., and Bridges, K. R. (1995) *J. Biol. Chem.* **270**, 19540–19544

[jj] Bergsten, P., Sanchez Moura, A., Atwater, I., and Levine, M. (1994) *J. Biol. Chem.* **269**, 1041–1045

[kk] Pauling, L. (1970) *Vitamin C and the Common Cold*, Freeman, San Francisco, California

[ll] Halliwell, B. (1999) *Trends Biochem. Sci.* **24**, 255–259

[mm] Russell, P., Garland, D., Zigler, J. S., Jr., Meakin, S. O., Tsui, L.-C., and Breitman, M. L. (1987) *FASEB J.* **1**, 32–35

[nn] Nagaraj, R. H., Sell, D. R., Prabhakaram, M., Ortwerth, B. J., and Monnier, V. M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10257–10261

[oo] Moertel, C. G., Fleming, T. R., Creagan, E. T., Rubin, J., O'Connell, M. J., and Ames, M. M. (1985) *N. Engl. J. Med.* **312**, 142–146

[pp] Lee, S. H., Oe, T., and Blair, I. A. (2001) *Science* **292**, 2083–2086

cytochrome P450s, such as a soluble fatty acid hydroxylase from *Bacillus megaterium*, have reductase domains with tightly bound FMN and FAD bound to the same polypeptide chain as is the heme.[491]

All cytochromes P450 appear to have at their active sites a molecule of heme with a thiolate anion as an axial ligand in the fifth position (Fig. 18-23). These relatively large heme proteins of ~ 45- to 55-kDa mass may consist of as many as 490 residues. Only a few three-dimensional structures are known,[490,492–494] and among these there are significant differences. However, on the basis of a large amount of experimental effort[487,495,496] it appears that all cytochromes P450 act by basically similar mechanisms.[474,496a,b,497] As indicated in Eq. 18-57, the substrate AH binds to the protein near the heme, which must be in the Fe(III) form. An electron delivered from the reductase then reduces the iron to the Fe(II) state (Eq. 18-57, step *b*). Then O_2 combines with the iron, the initial oxygenated complex formed in step *c* being converted to an Fe(III)-superoxide complex (Eq. 18-57, step *d*). Subsequent events are less certain.[497a] Most often a second electron is transferred in from the reductase (Eq. 18-57, step *e*) to give a peroxide complex of Fe(III), which is then converted in step *f* to a ferryl iron form, as in the action of peroxidases (Fig. 16-14). This requires transfer of two H^+ into the active site. The ferryl Fe(IV)=O donates its oxygen atom to the substrate regenerating the Fe(III) form of the heme (step *g*) and releasing the product (step *h*).

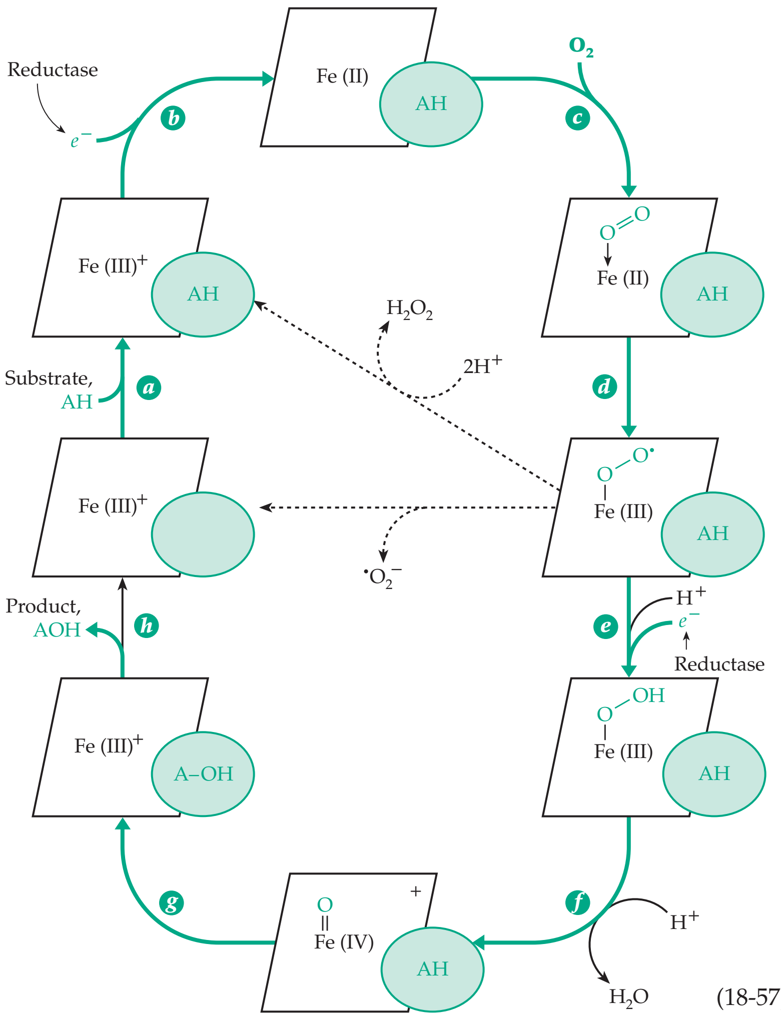
Microsomal cytochromes P450 often form hydrogen peroxide as a side product. This may arise directly from the $Fe–O–O^-$ intermediate shown in Eq. 18-57. Some cytochromes P450 use this reaction in reverse to carry out hydroxylation utilizing peroxides instead of O_2 (Eq. 18-58).

$$AH + ROOH \rightarrow AOH + ROH \qquad (18\text{-}58)$$

Cytochromes P450 often convert drugs or other foreign compounds to forms that are more readily excreted.[499] However, the result is not always beneficial. For example, 3-methylcholanthrene, a strong inducer of cytochrome P450, is converted to a powerful carcinogen by the hydroxylation reaction.[500] See also Box 18-E.

Other iron-containing oxygenases. Hydroxylases with properties similar to those of cytochrome P450 but containing nonheme iron catalyze ω-oxidation of alkanes and fatty acids in certain bacteria, e.g., *Pseudomonas oleovarans*. A flavoprotein rubredoxin reductase, is also required.[501] The methylotrophs *Methylococcus* and *Methylosinus* hydroxylate methane using as cosubstrate NADH or NADPH (Eq. 18-59). A soluble complex consists of 38-kDa reductase containing FAD and an Fe_2S_2 center, a small 15-kDa component, and a 245-kDa hydroxylase with an $(\alpha\beta\gamma)_2$ composition and a three-dimensional structure[502–503] similar to that of ribonucleotide reductase (Chapter 16, Section A,9). Each large α subunit contains a diiron center similar to that shown in Fig. 16-20C. It is likely that O_2 binds between the two iron atoms in the Fe(II) oxidation state and, oxidizing both irons to Fe(III), is converted to a bridging peroxide group as shown in Eq. 18-60. In this intermediate, in which the two metals are held rigidly by the surrounding ligands including a bridging carboxylate side chain, the O–O bond may be broken as in Eq. 18-60, steps *a* and *b*, to generate an Fe(IV)–O$^•$ radical that may

$$CH_4 + O_2 + NADH + H^+ \rightarrow CH_3OH + H_2O + NAD^+ \qquad (18\text{-}59)$$



(18-57)

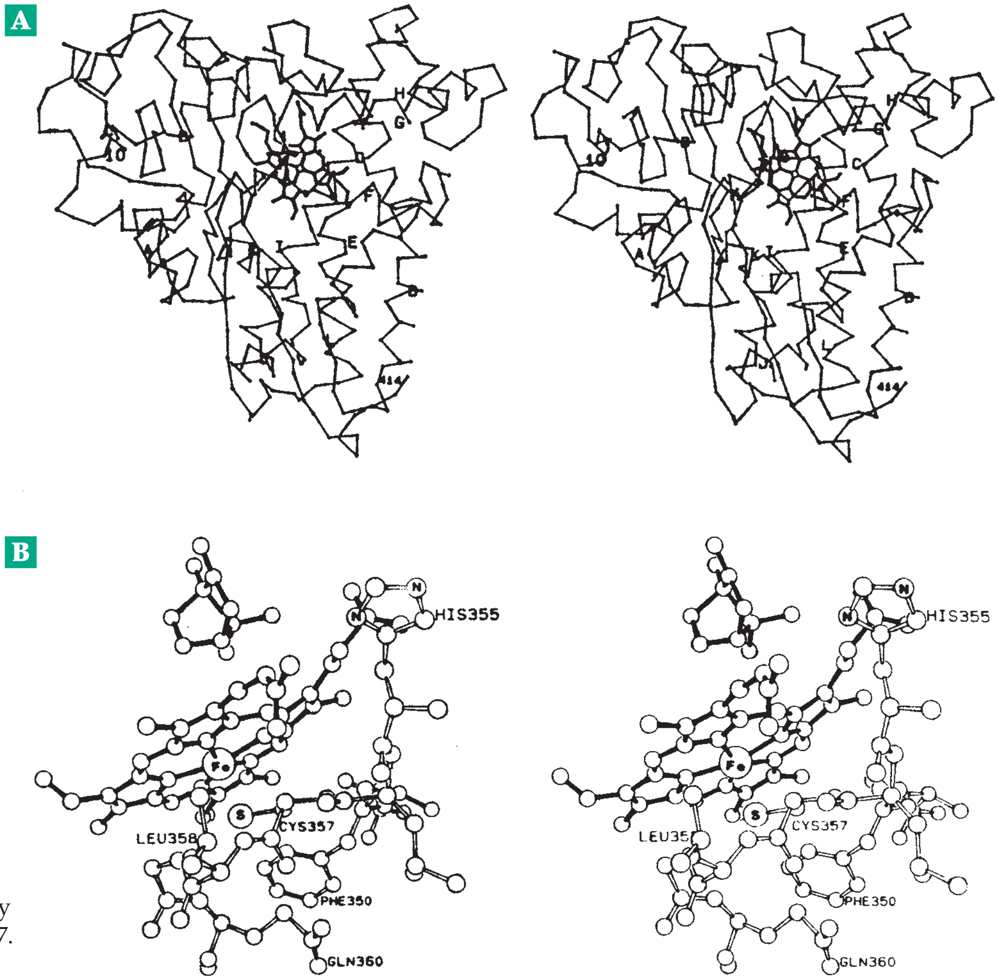
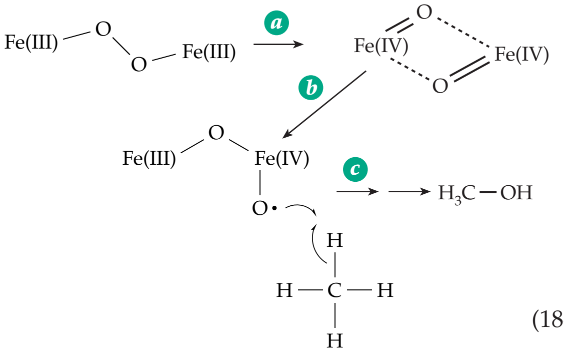


Figure 18-23 (A) Stereoscopic α-carbon backbone model of cytochrome $P450_{cam}$ showing the locations of the heme and of the bound camphor molecule. (B) View in the immediate vicinity of the thiolate ligand from Cys 357. From Poulos *et al.*[498]

remove a hydrogen atom from the substrate (step *c*) and undergo subsequent reaction steps analogous to those in the cytochrome P450 reaction cycle.[504–506]



(18-60)

A group of related bacterial enzymes hydroxylate alkanes,[507] toluene,[508] phenol,[509] and other substrates.[509a] Eukaryotic fatty acid desaturases (Fig. 16-20B) belong to the same family.[508] Some bacteria use cytochrome P450 or other oxygenase to add an oxygen atom to an alkene to form an epoxide. For example, propylene may be converted to either *R* or *S* epoxypropane which may be hydrolyzed, rearranged by a coenzyme M-dependent reaction, and converted to acetoacetate, which can be used as an energy source.[509a,b]

Nitric oxide and NO synthases. Nitric oxide (NO) is a reactive free radical whose formula is often written as $^{\bullet}$NO to recognize this characteristic. However, NO is not only a toxic and sometimes dangerous metabolite but also an important hormone with functions in the circulatory system, the immune system, and the brain.[510–512] The hormonal effects of NO are discussed in Chapter 30, but it is appropriate here to mention a few reactions. Nitric oxide reacts rapidly with O_2 to form nitrite (Eq. 18-61).

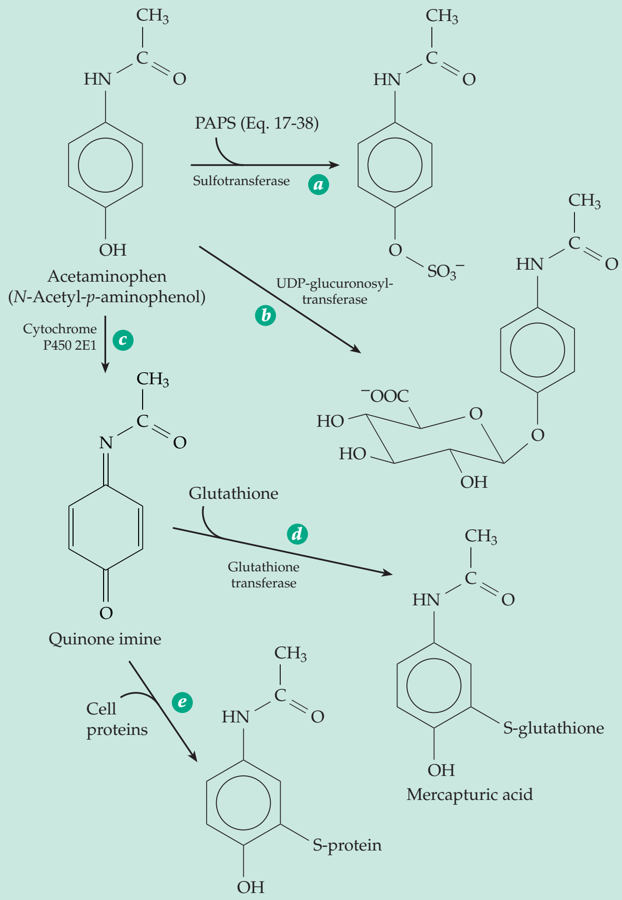
$$4\ NO + O_2 + 2\ H_2O \rightarrow 4\ NO_2^- + 4\ H^+ \qquad (18\text{-}61)$$

It also combines very rapidly with superoxide anion radical to form **peroxynitrite** (Eq. 18-62).[513] This is another reactive oxidant which, because of its relatively high pK_a of 6.8, is partially protonated and able to diffuse through phospholipids within cells.[514,515]

BOX 18-E THE TOXICITY OF ACETAMINOPHEN

Most drugs, as well as toxins and other xenobiotic compounds, enter the body through membranes of the gastrointestinal tract, lungs, or skin. Drugs are frequently toxic if they accumulate in the body. They are often rather hydrophobic and are normally converted to more polar, water-soluble substances before elimination from the body. Two major types of reaction take place, usually in the liver. These are illustrated in the accompanying scheme for acetaminophen (*N*-acetyl-*p*-aminophenol), a widely used analgesic and antipyretic (fever relieving) nonprescription drug sold under a variety of trade names: (1) A large water-soluble group such as sulfate[a] or glucuronate is transferred onto the drug by a nucleophilic displacement reaction (steps *a* and *b* of scheme). (2) Oxidation, demethylation, and other alterations are catalyzed by one or more of the nearly 300 cytochrome P450 monooxygenases present in the liver (step *c*). Oxidation products may be detoxified by glutathione S-transferases, step *d* (see also Box 11-B).[b,c,cc]

These reactions protect the body from the accumulation of many compounds but in some cases can cause serious problems. The best known of these involves acetaminophen. Its oxidation by cytochrome P450 2E1 or by prostaglandin H synthase[d] yields a highly reactive quinone imine which reacts with cell proteins.[e] Since the cytochrome P450 oxidation can occur in two steps, a reactive intermediate radical is also created.[c,f] At least 20 drug-labeled proteins arising in this way have been identified.[c] Both addition of thiol groups of proteins to the quinone imine (step *e* of scheme) and oxidation of protein thiols occur.[g] Mitochondria suffer severe damage,[h] some of which is related to induction of Ca^{2+} release.[i]



Acetaminophen is ordinarily safe at the recommended dosages, but large amounts exhaust the reserve of glutathione and may cause fatal liver damage. By 1989, more than 1000 cases of accidental or intentional (suicide) overdoses had been reported with many deaths. Prompt oral or intravenous administration of *N*-acetylcysteine over a 72-hour period promotes synthesis of glutathione and is an effective antidote.[j]

Similar problems exist for many other drugs. Both acetaminophen and phenacetin, its ethyl ether derivative, may cause kidney damage after many years of use.[k,l] Metabolism of phenacetin and several other drugs varies among individuals. Effective detoxification may not occur in individuals lacking certain isoenzyme forms of cytochrome P450.[m] Use of the anticancer drugs daunomycin (daunorubicin; Figs. 5-22 and 5-23) and adriamycin is limited by severe cardiac toxicity arising from free radicals generated during oxidation of the drugs.[n] These are only a few examples of the problems with drugs, pesticides, plasticizers, etc.

---

[a] Klaassen, C. D., and Boles, J. W. (1997) *FASEB J.* **11**, 404–418

[b] Lee, W. M. (1995) *N. Engl. J. Med.* **333**, 1118–1127

[c] Qiu, Y., Benet, L. Z., and Burlingame, A. L. (1998) *J. Biol. Chem.* **273**, 17940–17953

[cc] Chen, W., Shockcor, J. P., Tonge, R., Hunter, A., Gartner, C., and Nelson, S. D. (1999) *Biochemistry* **38**, 8159 – 8166

[d] Potter, D. W., and Hinson, J. A. (1987) *J. Biol. Chem.* **262**, 974–980

[e] Lee, S. S. T., Buters, J. T. M., Pineau, T., Fernandez-Salguero, P., and Gonzalez, F. J. (1996) *J. Biol. Chem.* **271**, 12063–12067

[f] Rao, D. N. R., Fischer, V., and Mason, R. P. (1990) *J. Biol. Chem.* **265**, 844–847

[g] Tirmenstein, M. A., and Nelson, S. D. (1990) *J. Biol. Chem.* **265**, 3059–3065

[h] Burcham, P. C., and Harman, A. W. (1991) *J. Biol. Chem.* **266**, 5049–5054

[i] Weis, M., Kass, G. E. N., Orrenius, S., and Moldéus, P. (1992) *J. Biol. Chem.* **267**, 804–809

[j] Smilkstein, M. J., Knapp, G. L., Kulig, K. W., and Rumack, B. H. (1988) *N. Engl. J. Med.* **319**, 1557–1562

[k] Stolley, P. D. (1991) *N. Engl. J. Med.* **324**, 191–193

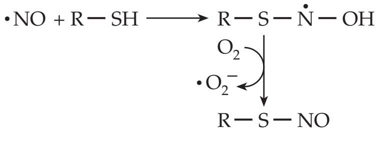
[l] Rocha, G. M., Michea, L. F., Peters, E. M., Kirby, M., Xu, Y., Ferguson, D. R., and Burg, M. B. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5317–5322

[m] Distlerath, L. M., Reilly, P. E. B., Martin, M. V., Davis, G. G., Wilkinson, G. R., and Guengerich, F. P. (1985) *J. Biol. Chem.* **260**, 9057–9067

[n] Davies, K. J. A., and Doroshow, J. H. (1986) *J. Biol. Chem.* **261**, 3060–3067

$$^{\bullet}NO + {}^{\bullet}O_2^- \rightarrow O{=}N{-}O{-}O^- \qquad (18\text{-}62)$$

NO binds to the iron atoms in accessible heme groups such as those of hemoglobin[516] and of guanylate cylases,[517,518] and in some Fe – S proteins such as aconitase. Some blood-sucking insects utilize salivary heme proteins called **nitrophorins** to carry NO into host tissues where it activates guanylate cyclase causing vasodilation. Nitrophorins also bind histamine and inhibit blood coagulation, which assists feeding.[518a,b] In the presence of a suitable oxidant such as O_2, nitric oxide reacts with thiol groups of proteins and small molecules to give ***S*-nitrosothiols** (Eq. 18-63).[511,519-520]



S-Nitrosothiol (18-63)

However, the physiological mechanisms of formation of these *S*-nitroso compounds is not clear.[520–521] One mechanism may involve conversion by O_2 to nitrous anhydride.[511,522]

$$2\ NO + O_2 \rightarrow ON{-}O{-}NO \qquad (18\text{-}64)$$

Nucleophilic attack on this compound by $-SH$, $-NH_2$, and other nucleophiles would yield *S*-nitroso and *N*-nitroso compounds with release of nitrous oxide N_2O.

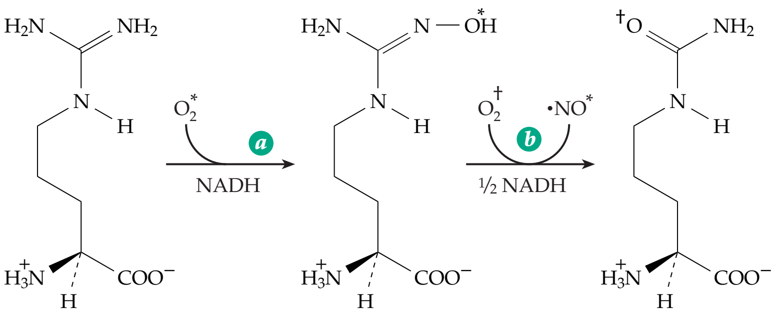
The relatively stable *S*-nitrosothiols derived from glutathione, cysteine, and proteins such as hemoglobin[516] may be important storage and transport forms of NO. If so, mechanisms of release of NO are important. A simple homolytic cleavage of R–S–NO to NO plus a thiyl radical $R{-}S^{\bullet}$ has often been assumed. However, rapid cleavage requires catalysis by a transition metal ion or reaction with reducing agents such as ascorbate or other thiols.[523,524] *S*-Nitrosothiols may also give rise to nitrosonium (NO^+) or nitroxy (NO^-) ions.[524]

NO synthases are oxygenases that carry out a two-step oxidation of L-arginine to L-citrulline with production of NO. In the first step, a normal monooxygenase reaction, L-N^{γ}-hydroxyarginine is formed (Eq. 18-65, step *a*). In the second step (Eq. 18-65, step *b*) NO is formed in a three-electron oxidation. In this equation the symbols * and † indicate positions of incorporation of labeled O_2 atoms in the intermediate and final products.

The human body contains three types of nitric oxide synthase known as **neuronal** (nNOS or NOS1), **inducible** (iNOS or NOS2), and **endothelial** (eNOS or NOS3).[511a–f] These enzymes have a broad distribution within various tissues, but NOS1 is especially active in neurons and NOS3 in endothelial cells. The inducible NOS2 originally discovered in macrophages is transcriptionally regulated. When these phagocytic cells are at rest the activity of NOS2 is very low, but it becomes highly active after induction by cytokines or by the lipopolysaccharides of bacterial cell membranes.[525,526] Both NOS1 and NOS2 are constitutively expressed but require calcium ions, which bind to a calmodulin domain of the protein. The inducible NOS2 doesn't require added Ca^{2+} but does contain the calmodulin domain. NOS3 carries an N-terminal myristoyl group as well as cysteines that may be palmitoylated. It is located in caveolae of plasma membranes and in Golgi complexes.[527,528]

Nucleotide sequences revealed a close similarity of NO synthases to cytochrome P450 reductase.[525,529,530] Study of spectra suggested that NO synthases, in which the heme is held by a thiolate sulfur of a cysteine side chain, might be specialized cytochromes P450.[531] However, although the heme in NO synthases is bound by a thiolate group the protein fold is unlike that of cytochromes P450.[532,533] The NO synthases all share a three-component structure.[534] In the NOS2 from macrophages residues 1–489 form the catalytic oxygenase domain, residues 499–530 bind calmodulin, and residues 531–1144 form the reductase domain.[535] The last contains a binding site for NADPH as well as bound FMN and FAD. The reductase domain structure can be modeled after that of cytochrome P450 reductase.[533] Electrons from NADPH are apparently transferred to FAD, then to FMN, and then to the catalytic site.[525]

The least understood aspect of NO synthases is the requirement for tetrahydrobiopterin, BH_4, the same coenzyme required by the other pterin-dependent monooxygenases (Eq. 18-44). The presence of this coenzyme in the reduced BH_4 form is essential for step *a* of Eq. 18-65 but not for step *b*. This suggests that in step *a* an organic peroxide might be generated by BH_4 and used to form an oxo-iron hydroxylating reagent.

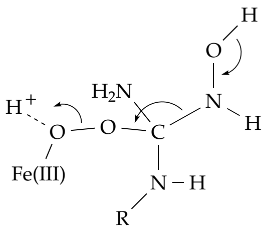


L-Arginine L-N^{γ}-hydroxyarginine L-Citrulline

(18-65)

However, there is no evidence for the expected quinonoid dihydropterin, and the three-dimensional structure suggested that BH_4 plays a structural role in mediating essential conformational changes.[532,535a,b] Nevertheless, newer data indicate a role in electron transfer.[535c]

Step *b* of Eq. 18-65 is an unusual three-electron oxidation, which requires only one electron to be delivered from NADPH by the reductase domain. Hydrogen peroxide can replace O_2 in this step.[536] A good possibility is that a peroxo or superoxide complex of the heme in the Fe(III) state adds to the hydroxyguanidine group. For example, the following structure could arise from addition of Fe(III)–O–O⁻:

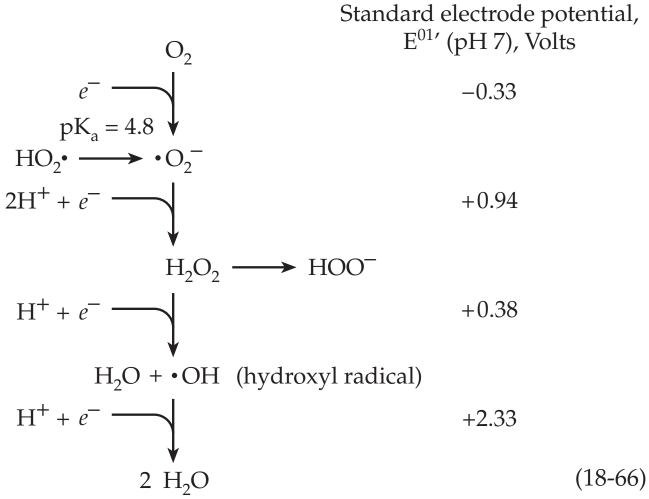


Breakup as indicated by the arrows on this structure would give Fe(III)–OH, citrulline, and O=N–H, **nitroxyl**. This is one electron (e^- + H^+) more reduced than •NO. Perhaps the adduct forms from Fe(III)–O–O•. On the other hand, there is evidence that NO synthases may produce nitroxyl or nitroxyl ion NO^- as the initial product.[537–538] NO and other products such as N_2O and NO_2^- may arise rapidly in subsequent reactions. Nitrite is a major oxidation product of NO in tissues.[538a] The chemistry of NO in biological systems is complex and not yet fully understood. See also pp. 1754, 1755.

G. Biological Effects of Reduced Oxygen Compounds

Although molecular oxygen is essential to the aerobic mode of life, it is toxic at high pressures. Oxidative damage from O_2 appears to be an important cause of aging and also contributes to the development of cancer. Reduced forms of oxygen such as superoxide, hydrogen peroxide, and hydroxyl radicals are apparently involved in this toxicity.[539,540] The same agents are deliberately used by phagocytic cells such as the neutrophils (polymorphonuclear leukocytes) to kill invading bacteria or fungi and to destroy malignant cells.[541]

The reactions shown with vertical arrows in Eq. 18-66 can give rise to the reduced oxygen compounds. The corresponding standard redox potential at pH 7 for each is also given.[539,542–544] As indicated by the low value of the redox potential for the $O_2 / O_2^{\bullet -}$ couple, the formation of superoxide by reduction of O_2 is spontaneous only for strongly reducing one-electron donors. Superoxide ion is a strong reductant, but at the same time a powerful one-electron oxidant, as is indicated by the high electrode potential of the $O_2^{\bullet -}$/ H_2O_2 couple.



(18-66)

1. The Respiratory Burst of Neutrophils

Some 25 x 10^9 neutrophils circulate in an individual's blood, and an equal number move along the surfaces of red blood cells. Invading microorganisms are engulfed after they are identified by the immune system as foreign. Phagocytosis is accompanied by a rapid many-fold rise in the rate of oxygen uptake as well as an increased glucose metabolism. One purpose of this **respiratory burst**[545–548] is the production of reduced oxygen compounds that kill the ingested microorganisms. In the very serious **chronic granulomatous disease** the normal respiratory burst does not occur, and bacteria are not killed.[549] The respiratory burst seems to be triggered not by phagocytosis itself, but by stimulation of the neutrophil by chemotactic formylated peptides such as formyl-Met-Leu-Phe[550] and less rapidly by other agonists such as phorbol esters.

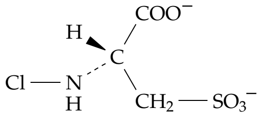
The initial product of the respiratory burst appears to be superoxide ion O_2^-. It is formed by an **NADPH oxidase,** which transports electrons from NADPH to O_2 probably via a flavoprotein and cytochrome b_{558}. Either the flavin or the cytochrome b_{558} must donate one electron to O_2 to form the superoxide anion.

$$\text{NADPH} \rightarrow \text{FAD} \rightarrow \text{cyt } b_{558} \rightarrow O_2 \qquad (18\text{-}67)$$

Flavocytochrome b_{558} (also called b_{-245}) has the unusually low redox potential of −0.245 V. It exists in phagocytic cells as a heterodimer of membrane-associated subunits p22-*phox* and gp91-*phox* where *phox* indicates phagocytic oxidase. The larger 91-kDa subunit contains two heme groups as well as one FAD and the presumed NADPH binding site.[548,551–552a] The mechanism of interaction with O_2 is unclear. Unlike hemoglobin but like other cytochromes *b*, cytochrome b_{558} does not form a complex with CO.[553] NADPH oxidase also requires two cytosolic components p47-*phox* and p67-*phox*. In resting phagocytes they reside in the cytosol as a 240-kDa complex with a third component, p40-*phox*, which may serve as an inhibitor.[548,554] Upon activation of the phagocyte in response to chemotactic signals the cytosolic components undergo phosphorylation at several sites,[554] and protein p47-*phox* and p67-*phox* move to the membrane and bind to and with the assistance of the small G protein Rac[552a,554a] activate flavocytochrome b_{558}. Phosphorylation of p47-*phox* may be especially important.[555]

In the X-chromosome-linked type of chronic granulomatous disease flavocytochrome b_{558} is absent or deficient, usually because of mutation in gp91-*phox*.[556,557] In an autosomal recessive form the superoxide-forming oxidase system is not activated properly. In some patients protein kinase C fails to phosphorylate p47-*phox*.[556,558] Less severe symptoms arise from deficiences in myeloperoxidase, chloroperoxidase (Chapter 16), glucose 6-phosphate dehydrogenase, glutathione synthetase, and glutathione reductase. The importance of these enzymes can be appreciated by examination of Fig. 18-24, which illustrates the relationship of several enzymatic reactions to the formation of superoxide anion and related compounds. Not only neutrophils but monocytes, macrophages, **natural killer cells** (NK cells), and other phagocytes apparently use similar chemistry in attacking ingested cells (Chapter 31).[559] Superoxide-producing NADH oxidases have also been found in nonphagocytic cells in various tissues.[559a]

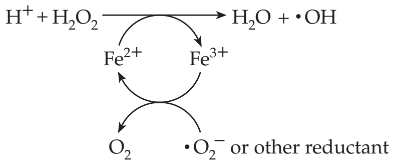
What kills the ingested bacteria and other microorganisms? Although superoxide anion is relatively unreactive, its protonated form $HO_2\bullet$, is very reactive. Since its pK_a is 4.8, there will be small amounts present even at neutral pH. Some of the $^\bullet O_2^-$ may react with NO to form peroxynitrite (Eq. 18-62).[559b] Peroxynitrite, in turn, can react with the ubitquitous CO_2 to give $^\bullet CO_3^-$ and $^\bullet NO_2$ radicals.[559c] Peroxynitrite anion also reacts with metalloenzyme centers[559d] and causes nitration and oxidation of aromatic residues in proteins.[559d,e] However, neutrophils contain active superoxide dismutases, and most of the superoxide that is formed is converted quickly to O_2 and H_2O_2. The latter may diffuse into the phagosomes as well as into the extracellular space. The H_2O_2 itself is toxic, but longer lived, more toxic oxidants are also formed. Reaction of H_2O_2 with **myeloperoxidase** (Chapter 16) produces hypochlorous acid, **(HOCl;** Eqs. 16-12, Fig. 18-24) and **chloramines** such as NH_2Cl, RNHCl, and $RNCl_2$. An important intracellular chloramine may be that of taurine.



Chloramine formed from taurine

Human neutrophils use HOCl formed by myeloperoxidase to oxidize α-amino acids such as tyrosine to reactive aldehydes that form adducts with –SH, $–NH_2$, imidazole, and other nucleophilic groups.[560] They also contain NO synthases, which form NO, peroxynitrite (Fig. 18-24), and nitrite.[561,562]

Hydroxyl radicals •OH, which attack proteins, nucleic acids, and a large variety of other cellular constituents, may also be formed. Although too reactive to diffuse far, they can be generated from H_2O_2 by Eq. 18-68. This reaction involves catalysis by Fe ions as shown.[562a]



(18-68)

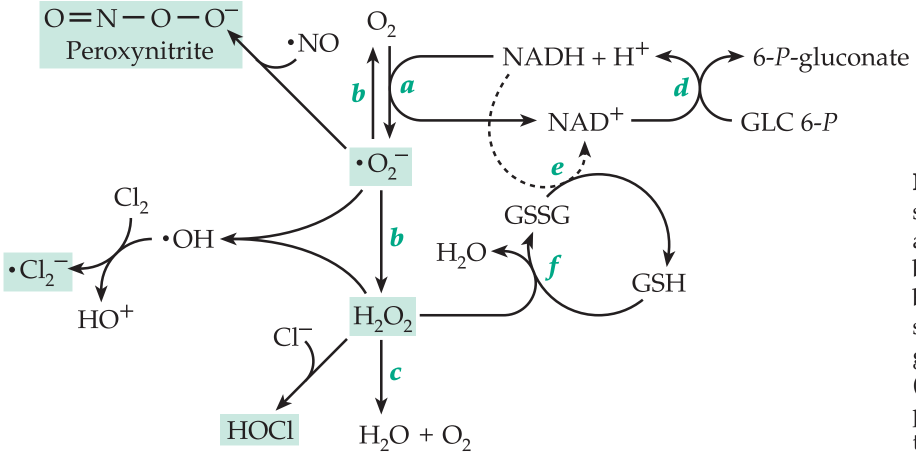


Figure 18-24 Some reactions by which superoxide anions, hydrogen peroxide and related compounds are generated by neutrophils and to a lesser extent by other cells: (*a*) NADPH oxidase, (*b*) superoxide dismutase, (*c*) catalase, (*d*) glucose-6-phosphate dehydrogenase, (*e*) glutathione reductase, (*f*) glutathione peroxidase. Abbrevations: GSH, glutathione; GSSB, oxidized glutathione.

Because Fe^{3+} is present in such low concentrations, there is uncertainty as to the biological significance of this reaction.[563] However, other iron compounds may function in place of Fe^{3+} and Fe^{2+} in Eq. 18-68.[564] A mixture of ferrous salts and H_2O_2 (Fenton reagent) has long been recognized as a powerful oxidizing mixture, which generates •OH or compounds of similar reactivity.[564–567] Ascorbate, and various other compounds, can also serve as the reductant in Eq. 18-68.[568]

Eosinophils, whose presence is stimulated by parasitic infections, have a peroxidase which acts preferentially on Br^- to form HOBr.[569] This compound can react with H_2O_2 more efficiently than does HOCl (Eq. 16-16) to form the very reactive **singlet oxygen.**[570] Singlet oxygen can also be generated from H_2O_2 and $\bullet O_2$ by Eq. 18-69[571] and also photochemically.[572]

$$\bullet O_2^- + H_2O_2 \rightarrow {}^1O_2({}^1\Delta_g) + OH + OH^- \qquad (18\text{-}69)$$

Additional killing mechanisms used by phagocytes include acidification of the phagocytic lysosomes with the aid of a proton pump[573] and formation of toxic peptides. For example, bovine neutrophils produce the bacteriocidal peptide RLCRIVVIRVCR which has a disulfide crosslinkage between the two cysteine residues.[574] Microorganisms have their own defenses against the oxidative attack by phagocytes. Some bacteria have very active superoxide dismutases. The protozoan *Leishmania* produces an acid phosphatase that shuts down the production of superoxide of the host cells in response to activating peptides.[575]

A respiratory burst accompanies fertilization of sea urchin eggs.[576,577] In this case, the burst appears to produce H_2O_2 as the major or sole product and is accompanied by release of **ovoperoxidase** from cortical granules. This enzyme uses H_2O_2 to generate **dityrosine crosslinkages** between tyrosine side chains during formation of the fertilization membrane. Defensive respiratory bursts are also employed by plant cells.[578,579] See also Box 18-B.

2. Oxidative Damage to Tissues

Superoxide anion radicals are formed not only in phagocytes but also as an accidental by-product of the action of many flavoproteins,[580,581] heme enzymes, and other transition metal-containing proteins. An example is xanthine oxidase. It is synthesized as xanthine dehydrogenase which is able to use NAD^+ as an oxidant, but upon aging, some is converted into the $^\bullet O_2$-utilizing xanthine oxidase (Chapter 16). This occurs extensively during ischemia. When oxygen is readmitted to a tissue in which this conversion of xanthine dehydrogenase to xanthine oxidase has occurred, severe oxidative injury may occur.[582] In animals the intravenous administration of superoxide dismutase or pretreatment with the xanthine oxidase inhibitor **allopurinol** (Chapter 25) prevents much of the damage, suggesting that superoxide is the culprit.

Hydrogen peroxide is also generated within cells[583] by flavoproteins and metalloenzymes and by the action of superoxide dismutase on $^\bullet O_2^-$. Since H_2O_2 is a small uncharged molecule, it can diffuse out of cells and into other cells readily. If it reacts with Fe(II), it can be converted within cells to $^\bullet OH$ radicals according to Eq. 18-68. Such radicals and others have been detected upon readmission of oxygen to ischemic animal hearts.[584,585] It has also been suggested that NADH may react with Fe(III) compounds in the same way as does O_2^- in Eq. 18-68 to provide a mechanism for producing hydroxyl radicals from H_2O_2.[539] Nitric oxide, formed by the various NO synthases in the cytosol and in mitochondria[586] and by some cytochromes P450,[587] is almost ubiquitous and can also lead to formation of peroxynitrate (Eq. 18-62). Thus, the whole range of reduced oxygen compounds depicted in Eq. 18-24 are present in small amounts throughout cells.

There is little doubt that these compounds cause extensive damage to DNA, proteins, lipids, and other cell constituents.[539,540,563,588] For example, one base in 150,000 in nuclear DNA is apparently converted from guanine to 8-hydroxyguanine presumably as a result of attack by oxygen radicals.[589] In mitochondrial DNA one base in 8000 undergoes this alteration. This may be a result of the high rate of oxygen metabolism in mitochondria and may also reflect the lack of histones and the relatively inefficient repair of DNA within mitochondria. Proteins undergo chain cleavage, crosslinking, and numerous side chain modification reactions.[588] Dissolved O_2 can react directly with exposed glycyl residues in protein backbones to create glycyl radicals which may lead to chain cleavage as in Eq. 15-39.[588a] Iron–sulfur clusters, such as the Fe_4S_4 center of aconitase (Fig. 13-4), are especially sensitive to attack by superoxide anions.[588a–c] "Free iron" released from the Fe–S cluster may catalyze formation of additional damaging radicals.

Antioxidant systems. Cells have numerous defenses against oxidative damage.[563,590,591] Both within cells and in extracellular fluids superoxide dismutase (Eq. 16-27) decomposes superoxide to O_2 and H_2O_2. The H_2O_2 is then broken down by catalase (Eq. 16-8) to O_2 and H_2O. In higher animals the selenoenzyme glutathione peroxidase (Chapter 16) provides another route for decomposition of H_2O_2 and lipid peroxides of membranes. The oxidized glutathione formed is reduced by NADPH. The system has a critical role within erythrocytes (Box 15-H). In chloroplasts an analogous system utilizes ascorbate peroxidase, ascorbate, and glutathione to break down peroxides.[592]

$$\text{NADPH} + \text{H}^+ \rightarrow \text{glutathione} \rightarrow \text{ascorbate} \rightarrow \text{H}_2\text{O}_2 \quad (18\text{-}70)$$

Ascorbate,[593–594] glutathione, NADPH,[594a] and tocopherols (Box 15-G)[595] all act as scavengers of free radicals such as O_2^-, $^\bullet OH$ and $ROO^\bullet$, $^\bullet CO_3^-$, and of singlet oxygen. Antioxidant protection is needed in extracellular fluids as well as within cells. In addition to glutathione and ascorbate, bilirubin,[596] uric acid,[597] melatonin,[598,598a] circulating superoxide dismutase, and the copper protein ceruloplasmin (Chapter 16) all act as antioxidants. Methionine residues of proteins may have a similar function.[599] Various proteins and small chelating compounds such as citrate tie up Fe^{3+} preventing it from promoting radical formation. Tocopherols, ubiquinols, and lipoic acid[600,600a–c] protect membranes. Beta carotene (Fig. 22-5), another lipid-soluble antioxidant, is the most effective quencher of singlet O_2 that is known. Even nitric oxide, usually regarded as toxic, sometimes acts as an antioxidant.[601] Trehalose protects plants against oxidative damage.[601a]

An increasing number of proteins are being recognized as protectants against oxidative damage. The exposed –SH and $-SCH_3$ groups of cysteine and methionine residues in proteins may function as appropriately located scavengers which may donate electrons to destroy free radicals or react with superoxide ions to become sulfonated. The thioredoxin (Box 15-C) and glutathione (Eq. 18-70) systems, in turn, reduce the protein radicals formed in this way.[601b–d] Methionine sulfoxide, both free and in polypeptides, is reduced by **methionine sulfoxide reductase** in organisms from bacteria to humans.[601e–g] Biotin, together with biotin sulfoxide reductase,[601h] may provide another antioxidant system. Some bacteria utilize glutathione-independent **alkylperoxide reductases** to scavenge organic peroxides.[601i] while mammals accomplish the same result with **peroxiredoxins** and with thioredoxin.[601j] Many other proteins will doubtless be found to participate in defense against oxidative damage. Oxygen is always present and its reactions in our bodies are essential. Generation of damaging reduced oxygen compounds and radicals is inevitable. Evolution will select in favor of many proteins that have been modified to minimize the damage.

Antioxidant enzymes do not always protect us. There was great excitement when it was found that victims of a hereditary form of the terrible neurological disease **amyotrophic lateral sclerosis** (**ALS**; see also Chapter 30) carry a defective gene for Cu / Zn-superoxide dismutase (SOD; Eq. 16-27).[602–603b] This discovery seemed to support the idea that superoxide anions in the brain were killing neurons. However, it now appears that in some cases of ALS the defective SOD is *too active*, producing an excess of H_2O_2, which damages neurons.

Transcriptional regulation of antioxidant proteins. Certain proteins with easily accessible Fe–S clusters, e.g., aconitase, are readily inactivated by oxidants such as peroxynitrite.[540,604] At least two proteins of this type function as transcription factors in *E. coli*. These are known as **SoxR** and **OxyR**. The SoxR protein is sensitive to superoxide anion, which carries out a one-electron oxidation on its Fe_2S_2 centers.[540,605–607] In its oxidized form SoxR is a transcriptional regulator that controls 30–40 genes, among them several that are directly related to "**oxidative stress.**"[608] These include genes for manganese SOD, glucose-6-phosphate dehydrogenase, a DNA repair nuclease, and aconitase (to replace the inactivated enzyme). The OxyR protein, which responds to elevated $[H_2O_2]$, is activated upon oxidation of a pair of nearby –SH groups to form a disulfide bridge.[607,609] It controls genes for catalase, glutathione reductase, an alkyl hydroperoxide reductase,[610,610a] and many others. Similar transcriptional controls in yeast result in responses to low doses of H_2O_2 by at least 167 different proteins.[608] Animal mitochondria also participate in sensing oxidant levels.[611,612] (See also Chapter 28, Section C,6.)

References

1. Tzagoloff, A. (1982) *Mitochondria*, Plenum, New York
1a. Scheffler, I. E. (1999) *Mitochondria*, Wiley-Liss, New York
2. Chappell, J. B. (1979) *The Energetics of Mitochondria*, 2nd ed., Oxford Univ. Press, London (Carolina Biology Reader No. 19)
3. Lee, C. P., Schatz, G., and Dallner, G., eds. (1981) *Mitochondria and Microsomes*, Addison-Wesley, Reading, Massachusetts
4. Tyler, D. D. (1992) *The Mitochondrion in Health and Disease*, VCH Publ., New York
5. Noble, R. W., and Gibson, Q. H. (1970) *J. Biol. Chem.* **245**, 2409 – 2413
6. Ankel-Simons, F., and Cummins, J. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13859 – 13863
7. Champion, P. M., Münck, E., Debrunner, P. G., Hollenberg, P. F., and Hager, L. P. (1973) *Biochemistry* **12**, 426 – 435
8. Packer, L. (1973) in *Mechanisms in Bioenergetics* (Azzone, G. F., Ernster, L., Papa, S., Quagliariello, E., and Siliprandi, N., eds), pp. 33 – 52, Academic Press, New York
9. Smith, L. D., and Gholson, R. K. (1969) *J. Biol. Chem.* **244**, 68 – 71
10. Harris, R. A., Williams, C. H., Caldwell, M., Green, D. E., and Valdivia, E. (1969) *Science* **165**, 700 – 703
11. Malhotra, S. K., and Sikewar, S. S. (1983) *Trends Biochem. Sci.* **8**, 358 – 359
12. Frey, T. G., and Mannella, C. A. (2000) *Trends Biochem. Sci.* **25**, 319 – 324
12a. Varmus, H. E. (1985) *Nature (London)* **314**, 583 – 584
12b. Rutter, G. A., and Rizzuto, R. (2000) *Trends Biochem. Sci.* **25**, 215 – 221
13. Ernster, L., and Drahota, Z., eds. (1969) *Mitochondria: Structure and Function*, Vol. 17, FEBS Symposium, (pp. 5 – 31)
14. Pfaller, R., Freitag, H., Harmey, M. A., Benz, R., and Neupert, W. (1985) *J. Biol. Chem.* **260**, 8188 – 8193
15. Heins, L., Mentzel, H., Schmid, A., Benz, R., and Schmitz, U. K. (1994) *J. Biol. Chem.* **269**, 26402 – 26410
15a. Bölter, B., and Soll, J. (2001) *EMBO J.* **20**, 935 – 940
16. Mannella, C. A. (1992) *Trends Biochem. Sci.* **17**, 315 – 320
17. Ha, H., Hajek, P., Bedwell, D. M., and Burrows, P. D. (1993) *J. Biol. Chem.* **268**, 12143 – 12149
18. Sutfin, L. V., Holtrop, M. E., and Ogilvie, R. E. (1971) *Science* **174**, 947– 949
19. Srere, P. A. (1982) *Trends Biochem. Sci.* **7**, 375 – 377
20. Srere, P. A. (1981) *Trends Biochem. Sci.* **6**, 4 – 7
20a. Haggie, P. M., and Brindle, K. M. (1999) *J. Biol. Chem.* **274**, 3941 – 3945
21. McCabe, E. R. B. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1631– 1652, McGraw-Hill, New York
22. Hackenbrock, C. R., and Hammon, K. M. (1975) *J. Biol. Chem.* **250**, 9185 – 9197
22a. Jiang, F., Ryan, M. T., Schlame, M., Zhao, M., Gu, Z., Klingenberg, M., Pfanner, N., and Greenberg, M. L. (2000) *J. Biol. Chem.* **275**, 22387 – 22394
22b. Sedlák, E., and Robinson, N. C. (1999) *Biochemistry* **38**, 14966 – 14972
22c. Gomez, B., Jr., and Robinson, N. C. (1999) *Biochemistry* **38**, 9031 – 9038
22d. McAuley, K. E., Fyfe, P. K., Ridge, J. P., Isaacs, N. W., Cogdell, R. J., and Jones, M. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14706 – 14711
22e. Bazhenova, E. N., Deryabina, Y. I., Eriksson, O., Zvyagilskaya, R. A., and Saris, N.-E. L. (1998) *J. Biol. Chem.* **273**, 4372 – 4377
22f. Horikawa, Y., Goel, A., Somlyo, A. P., and Somlyo, A. V. (1998) *Biophys. J.* **74**, 1579 – 1590
22g. Territo, P. R., French, S. A., Dunleavy, M. C., Evans, F. J., and Balaban, R. S. (2001) *J. Biol. Chem.* **276**, 2586 – 2599
22h. Smaili, S. S., Stellato, K. A., Burnett, P., Thomas, A. P., and Gaspers, L. D. (2001) *J. Biol. Chem.* **276**, 23329 – 23340
22i. Beutner, G., Sharma, V. K., Giovannucci, D. R., Yule, D. I., and Sheu, S.-S. (2001) *J. Biol. Chem.* **276**, 21482 – 21488
22j. Arnaudeau, S., Kelley, W. L., Walsh, J. V., Jr., and Demaurex, N. (2001) *J. Biol. Chem.* **276**, 29430 – 29439
22k. Berridge, M. J., Bootman, M. D., and Lipp, P. (1998) *Nature (London)* **395**, 645 – 648
23. Clayton, D. A. (1984) *Ann. Rev. Biochem.* **53**, 573 – 594
24. Slonirnski, P., Borst, P., and Attardi, G., eds. (1982) *Mitochondrial Genes*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
25. Attardi, G. (1981) *Trends Biochem. Sci.* **6**, 86 – 89; 100 – 103
26. Palmer, J. D. (1997) *Nature (London)* **387**, 454 – 455
26a. Schwartz, M., and Vissing, J. (2002) *N. Engl. J. Med.* **347**, 576 – 580
27. Wolstenholme, D. R. (1992) in *Mitochondrial Genomes—International Review of Cytology*, Vol. 141 (Wolstenholme, D. R., and Jeon, K. W., eds), pp. 173 – 216, Academic Press, San Diego, California
28. Wolstenholme, D. R., and Jeon, K. W., eds. (1992) *Mitochondrial Genomes*, Vol. 141, Academic Press, San Diego, California
28a. Wolfsberg, T. G., Schafer, S., Tatusov, R. L., and Tatusova, T. A. (2001) *Trends Biochem. Sci.* **26**, 199 – 203
29. Borst, P., and Grivell, L. A. (1981) *Nature (London)* **290**, 443 – 444
30. Janke, A., Xu, X., and Arnason, U. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1276 – 1281
31. Anderson, S., Bankier, A. T., Barrell, B. G., deBruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981) *Nature (London)* **290**, 457 – 470
32. Anderson, S., de Bruijn, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F., and Young, I. G. (1984) *J. Mol. Biol.* **156**, 683 – 717
33. Cantatore, P., Roberti, M., Rainaldi, G., Gadaleta, M. N., and Saccone, C. (1989) *J. Biol. Chem.* **264**, 10965 – 10975
34. Gardner, M. J., and 26 other authors (1998) *Science* **282**, 1126 – 1132
35. Köhler, S., Delwiche, C. F., Denny, P. W., Tilney, L. G., Webster, P., Wilson, R. J. M., Palmer, J. D., and Roos, D. S. (1997) *Science* **275**, 1485 – 1489
36. Burger, G., Plante, I., Lonergan, K. M., and Gray, M. W. (1995) *J. Mol. Biol.* **245**, 522 – 537
37. Lang, B. F., Burger, G., O'Kelly, C. J., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M., and Gray, M. W. (1997) *Nature (London)* **387**, 493 – 497
38. de Bruijn, M. H. L. (1983) *Nature (London)* **304**, 234 – 241
39. Bernardi, G. (1982) *Trends Biochem. Sci.* **7**, 404 – 408
40. Palmer, J. D., and Shields, C. R. (1984) *Nature (London)* **307**, 437 – 440
41. Leblanc, C., Boyen, C., Richard, O., Bonnard, G., Grienenberger, J.-M., and Kloareg, B. (1995) *J. Mol. Biol.* **250**, 484 – 495
42. Oldenburg, D. J., and Bendich, A. J. (1998) *J. Mol. Biol.* **276**, 745 – 758
42a. Oldenburg, D. J., and Bendich, A. J. (2001) *J. Mol. Biol.* **310**, 549 – 562
43. Barrell, B. G., Bankier, A. T., and Drouin, J. (1979) *Nature (London)* **282**, 189 – 194
44. Wallace, D. C. (1982) *Microbiol. Rev.* **46**, 208 – 240
45. Shoffner, J. M., and Wallace, D. C. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1535 – 1609, McGraw-Hill, New York
46. Wallace, D. C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8739 – 8746
47. Butow, R. A., Perlman, P. S., and Grossman, L. I. (1985) *Science* **228**, 1496 – 1501
48. Douglas, M., and Takeda, M. (1985) *Trends Biochem. Sci.* **10**, 192 – 194
49. Cavalier-Smith, T. (1987) *Nature (London)* **326**, 332 – 333
50. Schatz, G. (1997) *Nature (London)* **388**, 121 – 122
50a. Tokatlidis, K., and Schatz, G. (1999) *J. Biol. Chem.* **274**, 35285 – 35288
50b. Gabriel, K., Buchanan, S. K., and Lithgow, T. (2001) *Trends Biochem. Sci.* **26**, 36 – 40
51. Schatz, G., and Dobberstein, B. (1996) *Science* **271**, 1519 – 1526
52. Stuart, R. A., and Neupert, W. (1996) *Trends Biochem. Sci.* **21**, 261 – 267
53. Lithgow, T., Cuezva, J. M., and Silver, P. A. (1997) *Trends Biochem. Sci.* **22**, 110 – 113
54. Rojo, E. E., Guiard, B., Neupert, W., and Stuart, R. A. (1998) *J. Biol. Chem.* **273**, 8040 – 8047
55. Hartmann, C., Christen, P., and Jaussi, R. (1991) *Nature (London)* **352**, 762 – 763
56. Gärtner, F., Voos, W., Querol, A., Miller, B. R., Craig, E. A., Cumsky, M. G., and Pfanner, N. (1995) *J. Biol. Chem.* **270**, 3788 – 3795
57. Pfanner, N., Douglas, M. G., Endo, T., Hoogenraad, N. J., Jensen, R. E., Meijer, M., Neupert, W., Schatz, G., Schmitz, U. K., and Shore, G. C. (1996) *Trends Biochem. Sci.* **21**, 51 – 52
58. Komiya, T., Rospert, S., Koehler, C., Looser, R., Schatz, G., and Mihara, K. (1998) *EMBO J.* **17**, 3886 – 3898
59. Dietmeier, K., Hönlinger, A., Bömer, U., Dekker, P. J. T., Eckerskorn, C., Lottspeich, F., Kübrich, M., and Pfanner, N. (1997) *Nature (London)* **388**, 195 – 200
59a. Koehler, C. M., Leuenberger, D., Merchant, S., Renold, A., Junne, T., and Schatz, G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2141 – 2146
59b. Wallace, D. C., and Murdock, D. G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1817 – 1819
60. Chance, B., and Williams, G. R. (1955) *J. Biol. Chem.* **217**, 409 – 427
61. Prebble, J. N. (1981) *Mitochondria Chloroplasts and Bacterial Membranes*, Longman, London and New York
62. King, T. E. (1967) *Methods Enzymol.* **10**, 202 – 208
63. Ragen, C. I., and Racker, E. (1973) *J. Biol. Chem.* **248**, 2563 – 2569
64. Hafeti, Y. (1985) *Ann. Rev. Biochem.* **54**, 1015 – 1019
65. Yu, L., and Yu, C.-A. (1982) *J. Biol. Chem.* **257**, 2016 – 2021
66. Yu, C.-A., and Yu, L. (1982) *J. Biol. Chem.* **257**, 6127 – 6131
67. Wakabayashi, S., Takao, T., Shimonishi, Y., Kuramitsu, S., Matsubara, H., Wang, T., Zhang, Z., and King, T. E. (1985) *J. Biol. Chem.* **260**, 337 – 343
68. Moore, A. L., and Rich, P. R. (1980) *Trends Biochem. Sci.* **5**, 284 – 288
69. Palmer, J. M., and Moller, I. M. (1982) *Trends Biochem. Sci.* **7**, 258 – 261

References

70. Douce, R., and Day, D. A., eds. (1985) *Higher Plant Cell Respiration*, Vol. 18, Springer-Verlag, New York
71. Rhoads, D. M., and McIntosh, L. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2122 – 2126
72. Clarkson, A. B., JR, Bienen, E. J., Pollakis, G., and Grady, R. W. (1989) *J. Biol. Chem.* **264**, 17770 – 17776
73. Hoefnagel, M. H. N., Atkin, O. K., and Wiskich, J. T. (1998) *Biochim. Biophys. Acta.* **1366**, 235 – 255
74. Albury, M. S., Affourtit, C., and Moore, A. L. (1998) *J. Biol. Chem.* **273**, 30301– 30305
75. Rhoads, D. M., Umbach, A. L., Sweet, C. R., Lennon, A. M., Rauch, G. S., and Siedow, J. N. (1998) *J. Biol. Chem.* **273**, 30750 – 30756
76. Anraku, Y. (1988) *Ann. Rev. Biochem.* **57**, 101 – 132
77. Trumpower, B. L., and Gennis, R. B. (1994) *Ann. Rev. Biochem.* **63**, 675 – 716
78. Yamaguchi, M., Belogrudov, G. I., and Hatefi, Y. (1998) *J. Biol. Chem.* **273**, 8094 – 8098

78a. Grivennikova, V. G., Kapustin, A. N., and Vinogradov, A. D. (2001) *J. Biol. Chem.* **276**, 9038 – 9044

79. Guénebaut, V., Schlitt, A., Weiss, H., Leonard, K., and Friedrich, T. (1998) *J. Mol. Biol.* **276**, 105 – 112

79a. Hellwig, P., Scheide, D., Bungert, S., Mäntele, W., and Friedrich, T. (2000) *Biochemistry* **39**, 10884 – 10891

80. Weidner, U., Geier, S., Ptock, A., Friedrich, T., Leif, H., and Weiss, H. (1993) *J. Mol. Biol.* **233**, 109 – 122
81. Schryvers, A., Lohmeier, E., and Weiner, J. H. (1978) *J. Biol. Chem.* **253**, 783 – 788
82. Koland, J. G., Miller, M. J., and Gennis, R. B. (1984) *Biochemistry* **23**, 445 – 453
83. Condon, C., Cammack, R., Patil, D. S., and Owen, P. (1985) *J. Biol. Chem.* **260**, 9427 – 9434
84. Berry, E. A., and Trumpower, B. L. (1985) *J. Biol. Chem.* **260**, 2458 – 2467
85. John, P. (1981) *Trends Biochem. Sci.* **6**, 8 – 10
86. Kaysser, T. M., Ghaim, J. B., Georgiou, C., and Gennis, R. B. (1995) *Biochemistry* **34**, 13491 – 13501
87. Anraku, Y., and Gennis, R. B. (1987) *Trends Biochem. Sci.* **12**, 262 – 266
88. Sun, J., Kahlow, M. A., Kaysser, T. M., Osborne, J. P., Hill, J. J., Rohlfs, R. J., Hille, R., Gennis, R. B., and Loehr, T. M. (1996) *Biochemistry* **35**, 2403 – 2412

88a. Zhang, J., Hellwig, P., Osborne, J. P., Huang, H.-w, Moënne-Loccoz, P., Konstantinov, A. A., and Gennis, R. B. (2001) *Biochemistry* **40**, 8548 – 8556

88b. Gerscher, S., Döpner, S., Hildebrandt, P., Gleissner, M., and Schäfer, G. (1996) *Biochemistry* **35**, 12796 – 12803

88c. Das, T. K., Gomes, C. M., Teixeira, M., and Rousseau, D. L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9591 – 9596

88d. Saraste, M. (1999) *Science* **283**, 1488 – 1493

88e. Villani, G., Capitanio, N., Bizzoca, A., Palese, L. L., Carlino, V., Tattoli, M., Glaser, P., Danchin, A., and Papa, S. (1999) *Biochemistry* **38**, 2287 – 2294

89. Guénebaut, V., Vincentelli, R., Mills, D., Weiss, H., and Leonard, K. R. (1997) *J. Mol. Biol.* **265**, 409 – 418

89a. Schuler, F., Yano, T., Di Bernardo, S., Yagi, T., Yankovskaya, V., Singer, T. P., and Casida, J. E. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4149 – 4153

90. Grigorieff, N. (1998) *J. Mol. Biol.* **277**, 1033 – 1046
91. Di Bernardo, S., Yano, T., and Yagi, T. (2000) *Biochemistry* **39**, 9411 – 9418
92. Schneider, R., Brors, B., Massow, M., and Weiss, H. (1997) *FEBS Lett.* **407**, 249 – 252
93. Ohnishi, T., Ragan, C. I., and Hatefi, Y. (1985) *J. Biol. Chem.* **260**, 2782 – 2788
94. Yano, T., Sled, V. D., Ohnishi, T., and Yagi, Y. (1996) *J. Biol. Chem.* **271**, 5907 – 5913
95. Ragan, C. I., Galante, Y. M., Hatefi, Y., and Ohnishi, T. (1984) *Biochemistry* **21**, 590 – 594

95a. Kashani-Poor, N., Zwicker, K., Kerscher, S., and Brandt, U. (2001) *J. Biol. Chem.* **276**, 24082 – 24087

96. Heinrich, H., Azevedo, J. E., and Werner, S. (1992) *Biochemistry* **31**, 11420 – 11424
97. Ohshima, M., Miyoshi, H., Sakamoto, K., Takegami, K., Iwata, J., Kuwabara, K., Iwamura, H., and Yagi, T. (1998) *Biochemistry* **37**, 6436 – 6445
98. Kowal, A. T., Werth, M. T., Manodori, A., Cecchini, G., Schröder, I., Gunsalus, R. P., and Johnson, M. K. (1995) *Biochemistry* **34**, 12284 – 12293

98a. Iverson, T. M., Luna-Chavez, C., Cecchini, G., and Rees, D. C. (1999) *Science* **284**, 1961 – 1966

98b. Doherty, M. K., Pealing, S. L., Miles, C. S., Moysey, R., Taylor, P., Walkinshaw, M. D., Reid, G. A., and Chapman, S. K. (2000) *Biochemistry* **39**, 10695 – 10701

98c. Lancaster, C. R. D., Kröger, A., Auer, M., and Michel, H. (1999) *Nature (London)* **402**, 377 – 385

98d. Lancaster, C. R. D., Gross, R., Haas, A., Ritter, M., Mäntele, W., Simon, J., and Kröger, A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13051 – 13056

98e. Matsson, M., Tolstoy, D., Aasa, R., and Hederstedt, L. (2000) *Biochemistry* **39**, 8617 – 8624

99. Westenberg, D. J., Gunsalus, R. P., Ackrell, B. A. C., Sices, H., and Cecchini, G. (1993) *J. Biol. Chem.* **268**, 815 – 822

99a. Mowat, C. G., Pankhurst, K. L., Miles, C. S., Leys, D., Walkinshaw, M. D., Reid, G. A., and Chapman, S. K. (2002) *Biochemistry* **41**, 11990 – 11996

100. Van Hellemond, J. J., Klockiewicz, M., Gaasenbeek, C. P. H., Roos, M. H., and Tielens, A. G. M. (1995) *J. Biol. Chem.* **270**, 31065 – 31070
101. Nakamura, K., Yamaki, M., Sarada, M., Nakayama, S., Vibat, C. R. T., Gennis, R. B., Nakayashiki, T., Inokuchi, H., Kojima, S., and Kita, K. (1996) *J. Biol. Chem.* **271**, 521 – 527
102. Vibat, C. R. T., Cecchini, G., Nakamura, K., Kita, K., and Gennis, R. B. (1998) *Biochemistry* **37**, 4148 – 4159
103. Yang, X., Yu, L., and Yu, C.-A. (1997) *J. Biol. Chem.* **272**, 9683 – 9689
104. Xia, D., Yu, C.-A., Kim, H., Xia, J.-Z., Kachurin, A. M., Zhang, L., Yu, L., and Deisenhofer, J. (1997) *Science* **277**, 60 – 66
105. Zhang, Z., Huang, L., Shulmeister, V. M., Chi, Y.-I., Kim, K. K., Hung, L.-W., Crofts, A. R., Berry, E. A., and Kim, S.-H. (1998) *Nature (London)* **392**, 677 – 684

105a. Crofts, A. R., Hong, S., Zhang, Z., and Berry, E. A. (1999) *Biochemistry* **38**, 15827 – 15839

106. Iwata, S., Lee, J. W., Okada, K., Lee, J. K., Iwata, M., Rasmussen, B., Link, T. A., Ramaswamy, S., and Jap, B. K. (1998) *Science* **281**, 64 – 71
107. Smith, J. L. (1998) *Science* **281**, 58 – 59

107a. Baymann, F., Robertson, D. E., Dutton, P. L., and Mäntele, W. (1999) *Biochemistry* **38**, 13188 – 13199

107b. Lange, C., and Hunte, C. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2800 – 2805

108. Braun, H.-P., and Schmitz, U. K. (1995) *Trends Biochem. Sci.* **20**, 171 – 175

108a. Deng, K., Shenoy, S. K., Tso, S.-C., Yu, L., and Yu, C.-A. (2001) *J. Biol. Chem.* **276**, 6499 – 6505

109. Saribas, A. S., Ding, H., Dutton, P. L., and Daldal, F. (1995) *Biochemistry* **34**, 16004 – 16012
110. Denke, E., Merbitz-Zahradnik, T., Hatzfeld, O. M., Snyder, C. H., Link, T. A., and Trumpower, B. L. (1998) *J. Biol. Chem.* **273**, 9085 – 9093
111. Mitchell, P. (1976) *J. Theor. Biol.* **62**, 327 – 367
112. Trumpower, B. L. (1990) *J. Biol. Chem.* **265**, 11409 – 11412
113. Orii, Y., and Miki, T. (1997) *J. Biol. Chem.* **272**, 17594 – 17604

113a. Bartoschek, S., Johansson, M., Geierstanger, B. H., Okun, J. G., Lancaster, C. R. D., Humpfer, E., Yu, L., Yu, C.-A., Griesinger, C., and Brandt, U. (2001) *J. Biol. Chem.* **276**, 35231 – 35234

114. Jünemann, S., Heathcote, P., and Rich, P. R. (1998) *J. Biol. Chem.* **273**, 21603 – 21607

114a. Darrouzet, E., Moser, C. C., Dutton, P. L., and Daldal, F. (2001) *Trends Biochem. Sci.* **26**, 445 – 451

114b. Darrouzet, E., Valkova-Valchanova, M., Moser, C. C., Dutton, P. L., and Daldal, F. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4567 – 4572

114c. Crofts, A. R., Barquera, B., Gennis, R. B., Kuras, R., Guergova-Kuras, M., and Berry, E. A. (1999) *Biochemistry* **38**, 15807 – 15826

115. Tian, H., Yu, L., Mather, M. W., and Yu, C.-A. (1998) *J. Biol. Chem.* **273**, 27953 – 27959
116. Meinharat, S. M., Yang, X., Trumpower, B. L., and Ohnishi, T. (1987) *J. Biol. Chem.* **262**, 8702 – 8706
117. Saribas, A. M., Valkova-Valchanova, M., Tokito, M. K., Zhang, Z., Berry, E. A., and Daldal, F. (1998) *Biochemistry* **37**, 8105 – 8114
118. Yu, J., and Le Brun, N. E. (1998) *J. Biol. Chem.* **273**, 8860 – 8866
119. Barber, J. (1984) *Trends Biochem. Sci.* **9**, 209 – 211
120. Joliot, P., and Joliot, A. (1998) *Biochem. Soc. Trans.* **37**, 10404 – 10410

120a. Finazzi, G. (2002) *Biochemistry* **41**, 7475 – 7482

121. Hochman, J., Ferguson-Miller, S., and Scindler, M. (1985) *Biochemistry* **24**, 2509 – 2516
122. Hackenbrock, C. R. (1981) *Trends Biochem. Sci.* **6**, 151 – 154
123. Bernardi, P., and Azzone, G. F. (1981) *J. Biol. Chem.* **256**, 7187 – 7192
124. Vincent, J. S., and Levin, I. W. (1986) *J. Am. Chem. Soc.* **108**, 3551 – 3554
125. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., and Yoshikawa, S. (1996) *Science* **272**, 1136 – 1144
126. Gennis, R., and Ferguson-Miller, S. (1995) *Science* **269**, 1063 – 1064
127. Wallin, E., Tsukihara, T., Yoshikawa, S., Von Heijne, G., and Elofsson, A. (1997) *Protein Sci.* **6**, 808 – 815
128. Iwata, S., Ostermeier, C., Ludwig, B., and Michel, H. (1995) *Nature (London)* **376**, 660 – 669
129. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., and Yoshikawa, S. (1995) *Science* **269**, 1069 – 1074
130. Luchinat, C., Soriano, A., Djinovic-Carugo, K., Saraste, M., Malmström, B. G., and Bertini, I. (1997) *J. Am. Chem. Soc.* **119**, 11023 – 11027
131. Salgado, J., Warmerdam, G., Bubacco, L., and Canters, G. W. (1998) *Biochemistry* **37**, 7378 – 7389
132. Blackburn, N. J., de Vries, S., Barr, M. E., Houser, R. P., Tolman, W. B., Sanders, D., and Fee, J. A. (1997) *J. Am. Chem. Soc.* **119**, 6135 – 6143
133. Yoshikawa, S., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., Yamashita, E., Inoue, N., Yao, M., Fei, M. J., Libeu, C. P., Mizushima, T., Yamaguchi, H., Tomizaki, T., and Tsukihara, T. (1998) *Science* **280**, 1724 – 1729

References

134. Ostermeier, C., Harrenga, A., Ermler, U., and Michel, H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10547 – 10553
135. Stubbe, J., and Riggs-Gelasco, P. (1998) *Trends Biochem. Sci.* **23**, 438 – 443
135a. Buse, G., Soulimane, T., Dewor, M., Meyer, H. E., and Blüggel, M. (1999) *Protein Sci.* **8**, 985 – 990
136. Babcock, G. T., and Wikström, M. (1992) *Nature (London)* **356**, 301 – 309
136a. Babcock, G. T. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12971 – 12973
136b. Wikström, M. (2000) *Biochemistry* **39**, 3515 – 3519
136c. Morgan, J. E., Verkhovsky, M. I., Palmer, G., and Wilström, M. (2001) *Biochemistry* **40**, 6882 – 6892
137. Gennis, R. B. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12747 – 12749
138. Michel, H. (1999) *Biochemistry* **38**, 15129 – 15140
139. Sucheta, A., Georgiadis, K. E., and Einarsdóttir, O. (1997) *Biochemistry* **36**, 554 – 565
140. Gibson, Q., and Greenwood, C. (1963) *Biochem. J.* **86**, 541 – 554
141. Greenwood, C., and Gibson, Q. H. (1967) *J. Biol. Chem.* **242**, 1782 – 1787
142. Fabian, M., and Palmer, G. (2001) *Biochemistry* **40**, 1867 – 1874
142a. Han, S., Takahashi, S., and Rousseau, D. L. (2000) *J. Biol. Chem.* **275**, 1910 – 1919
142b. Karpefors, M., Ädelroth, P., Namslauer, A., Zhen, Y., and Brzezinski, P. (2000) *Biochemistry* **39**, 14664 – 14669
142c. Capitanio, N., Capitanio, G., Minuto, M., De Nitto, E., Palese, L. L., Nicholls, P., and Papa, S. (2000) *Biochemistry* **39**, 6373 – 6379
143. Karpefors, M., Ädelroth, P., Zhen, Y., Ferguson-Miller, S., and Brzezinski, P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13606 – 13611
143a. Rigby, S. E. J., Jünemann, S., Rich, P. R., and Heathcote, P. (2000) *Biochemistry* **39**, 5921 – 5928
143b. Pecoraro, C., Gennis, R. B., Vygodina, T. V., and Konstantinov, A. A. (2001) *Biochemistry* **40**, 9695 – 9708
144. Zaslavsky, D., Sadoski, R. C., Wang, K., Durham, B., Gennis, R. B., and Millett, F. (1998) *Biochemistry* **37**, 14910 – 14916
145. Nicholls, P. (1983) *Trends Biochem. Sci.* **8**, 353
146. Solioz, M. (1984) *Trends Biochem. Sci.* **9**, 309 – 312
147. Musser, S. M., and Chan, S. I. (1995) *Biophys. J.* **68**, 2543 – 2555
147a. Siletsky, S., Kaulen, A. D., and Konstantinov, A. A. (1999) *Biochemistry* **38**, 4853 – 4861
148. Aagaard, A., Gilderson, G., Mills, D. A., Ferguson-Miller, S., and Brzezinski, P. (2000) *Biochemistry* **39**, 15847 – 15850
148a. Brändén, M., Sigurdson, H., Namslauer, A., Gennis, R. B., Ädelroth, P., and Brzezinski, P. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5013 – 5018
149. Mitchell, D. M., Fetter, J. R., Mills, D. A., Adelroth, P., Pressler, M. A., Kim, Y., Aasa, R., Brzezinski, P., Malmström, B. G., Alben, J. O., Babcock, G. T., Ferguson-Miller, S., and Gennis, R. B. (1996) *Biochemistry* **35**, 13089 – 13093
150. Qian, J., Shi, W., Pressler, M., Hoganson, C., Mills, D., Babcock, G. T., and Ferguson-Miller, S. (1997) *Biochemistry* **36**, 2539 – 2543
151. Ädelroth, P., Gennis, R. B., and Brzezinski, P. (1998) *Biochemistry* **37**, 2470 – 2476
152. Konstantinov, A. A., Siletsky, S., Mitchell, D., Kaulen, A., and Gennis, R. B. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9085 – 9090
153. Behr, J., Hellwig, P., Mäntele, W., and Michel, H. (1998) *Biochemistry* **37**, 7400 – 7406
153a. Riistama, S., Laakkonen, L., Wikström, M., Verkhovsky, M. I., and Puustinen, A. (1999) *Biochemistry* **38**, 10670 – 10677
153b. Jünemann, S., Meunier, B., Fisher, N., and Rich, P. R. (1999) *Biochemistry* **38**, 5248 – 5255
154. Marantz, Y., Nachliel, E., Aagaard, A., Brzezinski, P., and Gutman, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8590 – 8595
155. Lehninger, L. A. (1972) in *International Symposium of Biochemistry and Biophysics of Mitochondrial Membranes* (Azzone, G. F., ed), p. 1, Academic Press, New York
156. Kalckar, H. M. (1969) *Biological Phosphorylations*, Prentice-Hall, Englewood Cliffs, New Jersey
157. Ernster, L. (1993) *FASEB J.* **7**, 1520 – 1524
158. Lemasters, J. J. (1984) *J. Biol. Chem.* **259**, 13123 – 13130
159. Flatt, J. P., Pahud, P., Ravussin, E., and Jéquier, E. (1984) *Trends Biochem. Sci.* **9**, 466 – 468
160. Hinkle, P. C., Kumar, M. A., Resetar, A., and Harris, D. L. (1991) *Biochemistry* **30**, 3576 – 3582
161. Ferguson, S. J. (1986) *Trends Biochem. Sci.* **11**, 351 – 353
162. Berry, E. A., and Hinkle, P. C. (1983) *J. Biol. Chem.* **258**, 1474 – 1486
163. Chance, B., and Williams, G. R. (1956) *Adv. Enzymol.* **17**, 65 – 134
164. Wilson, D. F., Erecinska, M., and Dutton, P. L. (1974) *Annu Rev Biophys Bioeng.* **3**, 203 – 230
165. Veech, R. L., Lawson, J. W. R., Cornell, N. W., and Krebs, H. A. (1979) *J. Biol. Chem.* **254**, 6538 – 6547
166. Lardy, H. A., and Ferguson, S. M. (1969) *Ann. Rev. Biochem.* **38**, 991– 1034
167. Metzler, D. E. (1977) *Biochemistry; The Chemical Reactions of Living Cells*, Academic Press, New York (pp. 598-600)
168. Mitchell, P. (1961) *Nature (London)* **191**, 144 – 148
169. Mitchell, P. (1968) *Chemiosmotic Coupling and Energy Transduction*, Glynn Res., Bodmin., Cornwall, England
170. Hinkle, P. C., and McCarty, R. E. (1978) *Sci. Am.* **238**(Mar), 104 – 123
171. Skulachev, V. P., and Hinckle, P. C., eds. (1981) *Chemiosmotic Protein Circuits in Biological Membranes*, Addison-Wesley, Reading, Massachusetts
172. Nicholls, D. G., and Ferguson, S. J. (1992) *Bioenergetics 2*, Academic Press, London
173. Mitchell, P. (1966) *Biol. Rev. Cambridge Phil. Soc.* **41**, 445 – 502
174. Mitchell, P. (1979) *Science* **206**, 1148 – 1159
175. Mitchell, P. (1978) *Trends Biochem. Sci.* **3**, N58 – N61
175a. Williams, R. J. P., and Prebble, J. (2002) *Trends Biochem. Sci.* **27**, 393 – 395
176. Lowe, A. G., and Jones, M. N. (1984) *Trends Biochem. Sci.* **9**, 11 – 13
177. Garland, P. (1978) *Nature (London)* **276**, 8 – 9
178. Mitchell, P., and Moyle, J. (1969) *Eur. J. Biochem.* **7**, 471 – 484
179. Wilson, D. F. (1980) in *Membrane Structure and Function*, Vol. 1 (Bittar, E. E., ed), Wiley, New York (pp. 153 – 195)
180. Waddell, W. J., and Butler, T. C. (1959) *J. Clin. Invest.* **38**, 770 –
181. Dodgson, S. J., Forster, R. E., II, and Storey, B. T. (1982) *J. Biol. Chem.* **257**, 1705 – 1711
182. Bowman, C., and Tedeschi, H. (1979) *Nature (London)* **280**, 597 – 598
183. Kinnally, K. W., Tedeschi, H., and Maloff, B. L. (1978) *Biochemistry* **17**, 3419 – 3427
184. Tedeschi, H. (1980) *Trends Biochem. Sci.* **5**, VIII – IX
185. Jagendorf, A. T. (1975) in *Bioenergetics of Photosynthesis* (Govindjee, ed), Academic Press, New York (pp. 413 – 492)
186. Jagendorf, A. T., and Uribe, E. (1966) *Proc. Natl. Acad. Sci. U.S.A.* **55**, 170 – 177
187. Richard, P., Pitard, B., and Rigaud, J.-L. (1995) *J. Biol. Chem.* **270**, 21571 – 21578
188. van Walraven, H. S., Strotmann, H., Schwarz, O., and Rumberg, B. (1996) *FEBS Lett.* **379**, 309 – 313
189. Costa, L. E., Reynafarje, B., and Lehninger, A. L. (1984) *J. Biol. Chem.* **259**, 4802 – 4811
190. Reynafarje, B., Costa, L. E., and Lehninger, A. L. (1986) *J. Biol. Chem.* **261**, 8254 – 8262
191. Capitanio, N., Capitanio, G., Demarinis, D. A., De Nitto, E., Massari, S., and Papa, S. (1996) *Biochemistry* **35**, 10800 – 10806
192. Wang, Y., Howton, M. M., and Beattie, D. S. (1995) *Biochemistry* **34**, 7476 – 7482
193. Wikström, M. (1989) *Nature (London)* **338**, 776 – 778
194. Wikström, M., Bogacher, A., Finel, M., Morgan, J. E., Puustinen, M., Raitio, M. L., Verkhovskaya, M. L., and Verkhovsky, M. I. (1994) *Biochim. Biophys. Acta.* **1187**, 106 – 111
195. Williams, R. J. P. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 71 – 97
196. Penefsky, H. S., Pullman, M. E., Datta, A., and Racker, E. (1960) *J. Biol. Chem.* **235**, 3330 – 3336
197. Racker, E. (1976) *A New Look at Mechanisms in Bioenergetics*, Academic Press, New York
198. Dschida, W. J., and Bowman, B. J. (1992) *J. Biol. Chem.* **267**, 18783 – 18789
199. Abrahams, J. P., Leslie, A. G. W., Lutter, R., and Walker, J. E. (1994) *Nature (London)* **370**, 621 – 628
200. Elston, T., Wang, H., and Oster, G. (1998) *Nature (London)* **391**, 510 – 513
201. Zhou, Y., Duncan, T. M., and Cross, R. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10583 – 10587
202. Junge, W., Lill, H., and Engelbrecht, S. (1997) *Trends Biochem. Sci.* **22**, 420 – 423
203. Engelbrecht, S., and Junge, W. (1997) *FEBS Letters* **414**, 485 – 491
204. Amzel, L. M., and Pedersen, P. L. (1983) *Ann. Rev. Biochem.* **52**, 801 – 824
205. Hausrath, A. C., Grüber, G., Matthews, B. W., and Capaldi, R. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13697 – 13702
206. Matsui, T., Muneyuki, E., Honda, M., Allison, W. S., Dou, C., and Yoshida, M. (1997) *J. Biol. Chem.* **272**, 8215 – 8221
207. Shirakihara, Y., Leslie, A. G. W., Abrahams, J. P., Walker, J. E., Ueda, T., Sekimoto, Y., Kambara, M., Saika, K., Kagawa, Y., and Yoshida, M. (1997) *Structure* **5**, 825 – 836
208. Spannagel, C., Vaillier, J., Chaignepain, S., and Velours, J. (1998) *Biochemistry* **37**, 615 – 621
208a. Arnold, I., Pfeiffer, K., Neupert, W., Stuart, R. A., and Schägger, H. (1999) *J. Biol. Chem.* **274**, 36 – 40
208b. Soubannier, V., Rusconi, F., Vaillier, J., Arselin, G., Chaignepain, S., Graves, P.-V., Schmitter, J.-M., Zhang, J. L., Mueller, D., and Velours, J. (1999) *Biochemistry* **38**, 15017 – 15024
209. Collinson, I. R., Runswick, M. J., Buchanan, S. K., Fearnley, I. M., Skehel, J. M., van Raaij, M. J., Griffiths, D. E., and Walker, J. E. (1994) *Biochemistry* **33**, 7971 – 7978
210. Pedersen, P. L., and Amzel, L. M. (1993) *J. Biol. Chem.* **268**, 9937 – 9940
210a. Hee Ko, Y., Hullihen, J., Hong, S., and Pedersen, P. L. (2000) *J. Biol. Chem.* **275**, 32931 – 32939
210b. Karrasch, S., and Walker, J. E. (1999) *J. Mol. Biol.* **290**, 379 – 384
211. Bianchet, M. A., Hullihen, J., Pedersen, P. L., and Amzel, L. M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11065 – 11070

212. Possmayer, F. E., and Gräber, P. (1994) *J. Biol. Chem.* **269**, 1896–1904
213. Sokolov, M., and Gromet-Elhanan, Z. (1996) *Biochemistry* **35**, 1242–1248
213a. Groth, G., and Pohl, E. (2001) *J. Biol. Chem.* **276**, 1345–1352
214. Hammes, G. G. (1983) *Trends Biochem. Sci.* **8**, 131–134
214a. Noji, H., and Yoshida, M. (2001) *J. Biol. Chem.* **276**, 1665–1668
214b. Schnitzer, M. J. (2001) *Nature (London)* **410**, 878–881
214c. Stock, D., Leslie, A. G. W., and Walker, J. E. (1999) *Science* **286**, 1700–1705
215. Belogrudov, G. I., Tomich, J. M., and Hatefi, Y. (1996) *J. Biol. Chem.* **271**, 20340–20345
216. Collinson, I. R., van Raaij, M. J., Runswick, M. J., Fearnley, I. M., Skehel, J. M., Orriss, G. L., Miroux, B., and Walker, J. E. (1994) *J. Mol. Biol.* **242**, 408–421
217. Golden, T. R., and Pedersen, P. L. (1998) *Biochemistry* **37**, 13871–13881
218. Joshi, S., Cao, G. J., Nath, C., and Shah, J. (1997) *Biochemistry* **36**, 10936–10943
219. Uhlin, U., Cox, G. B., and Guss, J. M. (1997) *Structure* **5**, 1219–1230
220. Gabellieri, E., Strambini, G. B., Baracca, A., and Solaini, G. (1997) *Biophys. J.* **72**, 1818–1827
221. Gordon-Smith, D. J., Carbajo, R. J., Yang, J.-C., Videler, H., Runswick, M. J., Walker, J. E., and Neuhaus, D. (2001) *J. Mol. Biol.* **308**, 325–339
222. Capaldi, R. A., Aggeler, R., Turina, P., and Wilkens, S. (1994) *Trends Biochem. Sci.* **19**, 284–289
222a. Capaldi, R. A., and Aggeler, R. (2002) *Trends Biochem. Sci.* **27**, 154–160
223. van Raaij, M. J., Abrahams, J. P., Leslie, A. G. W., and Walker, J. E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6913–6917
224. Abrahams, J. P., Buchanan, S. K., van Raaij, M. J., Fearnley, I. M., Leslie, A. G. W., and Walker, J. E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9420–9424
225. Wilkens, S., and Capaldi, R. A. (1998) *J. Biol. Chem.* **273**, 26645–26651
226. Bulygin, V. V., Duncan, T. M., and Cross, R. L. (1998) *J. Biol. Chem.* **273**, 31765–31769
227. Pan, W., Ko, Y. H., and Pedersen, P. L. (1998) *Biochemistry* **37**, 6911–6923
227a. Tsunoda, S. P., Rodgers, A. J. W., Aggeler, R., Wilce, M. C. J., Yoshida, M., and Capaldi, R. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6560–6564
228. Rodgers, A. J. W., Wilkens, S., Aggeler, R., Morris, M. B., Howitt, S. M., and Capaldi, R. A. (1997) *J. Biol. Chem.* **272**, 31058–31064
229. Sawada, K., Kuroda, N., Watanabe, H., Moritani-Otsuka, C., and Kanazawa, H. (1997) *J. Biol. Chem.* **272**, 30047–30053
230. Böttcher, B., Schwarz, L., and Gräber, P. (1998) *J. Mol. Biol.* **281**, 757–762
230a. McLachlin, D. T., Coveny, A. M., Clark, S. M., and Dunn, S. D. (2000) *J. Biol. Chem.* **275**, 17571–17577
230b. Wilkens, S., Zhou, J., Nakayama, R., Dunn, S. D., and Capaldi, R. A. (2000) *J. Mol. Biol.* **295**, 387–391
231. Long, J. C., Wang, S., and Vik, S. B. (1998) *J. Biol. Chem.* **273**, 16235–16240
232. Jiang, W., and Fillingame, R. H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6607–6612
233. Girvin, M. E., Rastogi, V. K., Abildgaard, F., Markley, J. L., and Fillingame, R. H. (1998) *Biochemistry* **37**, 8817–8824
234. Jones, P. C., Jiang, W., and Fillingame, R. H. (1998) *J. Biol. Chem.* **273**, 17178–17185
235. Groth, G., Tilg, Y., and Schirwitz, K. (1998) *J. Mol. Biol.* **281**, 49–59
236. Jones, P. C., and Fillingame, R. H. (1998) *J. Biol. Chem.* **273**, 29701–29705
236a. Dmitriev, O. Y., Abildgaard, F., Markley, J. L., and Fillingame, R. H. (2002) *Biochemistry* **41**, 5537–5547
237. Zhang, Y., and Fillingame, R. H. (1995) *J. Biol. Chem.* **270**, 87–93
238. Kaim, G., Matthey, U., and Dimroth, P. (1998) *EMBO J.* **17**, 688–695
239. Kaim, G., Wehrle, F., Gerike, U., and Dimroth, P. (1997) *Biochemistry* **36**, 9185–9194
240. Kaim, G., and Dimroth, P. (1998) *EMBO J.* **17**, 5887–5895
240a. Kaim, G., and Dimroth, P. (1999) *EMBO J.* **18**, 4118–4127
240b. Dimroth, P., Wang, H., Grabe, M., and Oster, G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4924–4929
240c. Aufurth, S., Schägger, H., and Müller, V. (2000) *J. Biol. Chem.* **275**, 33297–33301
241. Vik, S. B., and Antonio, B. J. (1994) *J. Biol. Chem.* **269**, 30364–30369
241a. Nijtmans, L. G. J., Henderson, N. S., Attardi, G., and Holt, I. J. (2001) *J. Biol. Chem.* **276**, 6755–6762
242. Sorgen, P. L., Bubb, M. R., McCormick, K. A., Edison, A. S., and Cain, B. D. (1998) *Biochemistry* **37**, 923–932
243. Dunn, S. D., and Chandler, J. (1998) *J. Biol. Chem.* **273**, 8646–8651
244. McLachlin, D. T., Bestard, J. A., and Dunn, S. D. (1998) *J. Biol. Chem.* **273**, 15162–15168
245. Boyer, P. D. (1993) *Biochim. Biophys. Acta.* **1140**, 215–250
246. Zhou, J.-M., and Boyer, P. D. (1993) *J. Biol. Chem.* **268**, 1531–1538
247. Boyer, P. D. (1995) *FASEB J.* **9**, 559–561
248. Boyer, P. D. (1997) *Ann. Rev. Biochem.* **66**, 717–749
249. Smith, L. T., Rosen, G., and Boyer, P. D. (1983) *J. Biol. Chem.* **258**, 10887–10894
250. Souid, A.-K., and Penefsky, H. S. (1995) *J. Biol. Chem.* **270**, 9074–9082
251. Penefsky, H. S., and Cross, R. L. (1991) *Adv. Enzymol.* **64**, 173–214
252. Cox, G. B., Jans, D. A., Fimmel, A. L., Gibson, F., and Hatch, L. (1984) *Biochim. Biophys. Acta.* **768**, 201–208
253. Aggeler, R., Ogilvie, I., and Capaldi, R. A. (1997) *J. Biol. Chem.* **272**, 19621–19624
254. Boekema, E. J., Ubbink-Kok, T., Lolkema, J. S., Brisson, A., and Konings, W. N. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14291–14293
255. Sabbert, D., Engelbrecht, S., and Junge, W. (1996) *Nature (London)* **381**, 623–625
256. Sabbert, D., Engelbrecht, S., and Junge, W. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4401–4405
257. Noji, H. (1998) *Science* **282**, 1844–1845
258. Noji, H., Yasuda, R., Yoshida, M., and Kinosita, K., Jr. (1997) *Nature (London)* **386**, 299–302
259. Kato-Yamada, Y., Noji, H., Yasuda, R., Kinosita, K., Jr., and Yoshida, M. (1998) *J. Biol. Chem.* **273**, 19375–19377
260. Berg, H. C. (1998) *Nature (London)* **394**, 324–325
261. Yasuda, R., Noji, H., Kinosita, K. J., Jr., and Yoshida, M. (1998) *Cell* **93**, 1117–1124
262. Kinosita, K. J., Jr., Yasuda, R., Noji, H., Ishiwata, S., and Yoshida, S. (1998) *Cell* **93**, 21–24
262a. Yasuda, R., Noji, H., Yoshida, M., Kinosita, K., Jr., and Itoh, H. (2001) *Nature (London)* **410**, 898–904
262b. Sambongi, Y., Iko, Y., Tanabe, M., Omote, H., Iwamoto-Kihara, A., Ueda, I., Yanagida, T., Wada, Y., and Futai, M. (1999) *Science* **286**, 1722–1724
262c. Hutcheon, M. L., Duncan, T. M., Ngai, H., and Cross, R. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8519–8524
262d. Nishio, K., Iwamoto-Kihara, A., Yamamoto, A., Wada, Y., and Futai, M. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13448–13452
263. Ko, Y. H., Bianchet, M., Amzel, L. M., and Pedersen, P. L. (1997) *J. Biol. Chem.* **272**, 18875–18881
264. Dou, C., Grodsky, N. B., Matsui, T., Yoshida, M., and Allison, W. S. (1997) *Biochemistry* **36**, 3719–3727
265. Grodsky, N. B., Dou, C., and Allison, W. S. (1998) *Biochemistry* **37**, 1007–1014
266. Dou, C., Fortes, P. A. G., and Allison, W. S. (1998) *Biochemistry* **37**, 16757–16764
267. Wang, H., and Oster, G. (1998) *Nature (London)* **396**, 279–282
267a. Rastogi, V. K., and Girvin, M. E. (1999) *Nature (London)* **402**, 263–268
267b. Jones, P. C., Hermolin, J., Jiang, W., and Fillingame, R. H. (2000) *J. Biol. Chem.* **275**, 31340–31346
267c. Le, N. P., Omote, H., Wada, Y., Al-Shawi, M. K., Nakamoto, R. K., and Futai, M. (2000) *Biochemistry* **39**, 2778–2783
267d. Ko, Y. H., Hong, S., and Pedersen, P. L. (1999) *J. Biol. Chem.* **274**, 28853–28856
268. Kobayashi, H., Suzuki, T., and Unemoto, T. (1986) *J. Biol. Chem.* **261**, 627–630
269. Forgac, M. (1999) *J. Biol. Chem.* **274**, 12951–12954
269a. Oka, T., Toyomura, T., Honjo, K., Wada, Y., and Futai, M. (2001) *J. Biol. Chem.* **276**, 33079–33085
269b. Müller, M. L., Jensen, M., and Taiz, L. (1999) *J. Biol. Chem.* **274**, 10706–10716
269c. Kawamura, Y., Arakawa, K., Maeshima, M., and Yoshida, S. (2000) *J. Biol. Chem.* **275**, 6515–6522
270. Oka, T., Yamamoto, R., and Futai, M. (1998) *J. Biol. Chem.* **273**, 22570–22576
271. van Hille, B., Richener, H., Evans, D. B., Green, J. R., and Bilbe, G. (1993) *J. Biol. Chem.* **268**, 7075–7080
272. Nelson, H., Mandiyan, S., and Nelson, N. (1994) *J. Biol. Chem.* **269**, 24150–24155
272a. Wilkens, S., Vasilyeva, E., and Forgac, M. (1999) *J. Biol. Chem.* **274**, 31804–31810
272b. Boekema, E. J., van Breemen, J. F. L., Brisson, A., Ubbink-Kok, T., Konings, W. N., and Lolkema, J. S. (1999) *Nature (London)* **401**, 37–38
272c. Sagermann, M., Stevens, T. H., and Matthews, B. W. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7134–7139
273. Steinert, K., Wagner, V., Kroth-Pancic, P. G., and Bickel-Sandkötter, S. (1997) *J. Biol. Chem.* **272**, 6261–6269
273a. Yokoyama, K., Muneyuki, E., Amano, T., Mizutani, S., Yoshida, M., Ishida, M., and Ohkuma, S. (1998) *J. Biol. Chem.* **273**, 20504–20510
274. Zhen, R.-G., Kim, E. J., and Rea, P. A. (1997) *J. Biol. Chem.* **272**, 22340–22348
274a. Scott, D. A., and Docampo, R. (2000) *J. Biol. Chem.* **275**, 24215–24221
274b. Miranda, M., Allen, K. E., Pardo, J. P., and Slayman, C. W. (2001) *J. Biol. Chem.* **276**, 22485–22490
275. Pederson, P. L., and Carafoli, E. (1987) *Trends Biochem. Sci.* **12**, 146–150
275a. Kühlbrandt, W., Zeelen, J., and Dietrich, J. (2002) *Science* **297**, 1692–1696
276. Lambrecht, N., Corbett, Z., Bayle, D., Karlish, S. J. D., and Sachs, G. (1998) *J. Biol. Chem.* **273**, 13719–13728
277. Asano, S., Tega, Y., Konishi, K., Fujioka, M., and Takeguchi, N. (1996) *J. Biol. Chem.* **271**, 2740–2745
278. Melle-Milovanovic, D., Milovanovic, M., Nagpal, S., Sachs, G., and Shin, J. M. (1998) *J. Biol. Chem.* **273**, 11075–11081

References

279. Wolfe, M. M., and Soll, A. H. (1988) *N. Engl. J. Med.* **319**, 1707 – 1715

280. Priver, N. A., Rabon, E. C., and Zeidel, M. L. (1993) *Biochemistry* **32**, 2459 – 2468

281. Nicholls, D. G., and Ferguson, S. J. (1992) *Bioenergetics 2*, Academic Press, London (p. 139)

282. Levings, C. S., III. (1990) *Science* **250**, 942 – 947

283. Hatefi, Y., and Yamaguchi, M. (1996) *FASEB J.* **10**, 444 – 452

284. Glavas, N. A., Hou, C., and Bragg, P. D. (1995) *Biochemistry* **34**, 7694 – 7702

285. Meuller, J., and Rydström, J. (1999) *J. Biol. Chem.* **274**, 19072 – 19080

286. Grimley, R. L., Quirk, P. G., Bizouarn, T., Thomas, C. M., and Jackson, J. B. (1997) *Biochemistry* **36**, 14762 – 14770

286a. Venning, J. D., Rodrigues, D. J., Weston, C. J., Cotton, N. P. J., Quirk, P. G., Errington, N., Finet, S., White, S. A., and Jackson, J. B. (2001) *J. Biol. Chem.* **276**, 30678 – 30685

287. Fjellström, O., Axelsson, M., Bizouarn, T., Hu, X., Johansson, C., Meuller, J., and Rydström, J. (1999) *J. Biol. Chem.* **274**, 6350 – 6359

288. Lee, A.-C., Zizi, M., and Colombini, M. (1994) *J. Biol. Chem.* **269**, 30974 – 30980

289. Rostovtseva, T., and Colombini, M. (1996) *J. Biol. Chem.* **271**, 28006 – 28008

290. Rostovtseva, T., and Colombini, M. (1997) *Biophys. J.* **72**, 1954 – 1962

291. Beavis, A. D., Lu, Y., and Garlid, K. D. (1993) *J. Biol. Chem.* **268**, 997 – 1004

292. Li, W., Shariat-Madar, Z., Powers, M., Sun, X., Lane, R. D., and Garlid, K. D. (1992) *J. Biol. Chem.* **267**, 17983 – 17989

293. Cox, D. A., and Matlib, M. A. (1993) *J. Biol. Chem.* **268**, 938 – 947

294. Sparagna, G. C., Gunter, K. K., Sheu, S.-S., and Gunter, T. E. (1995) *J. Biol. Chem.* **270**, 27510 – 27515

295. Bisaccia, F., Capobianco, L., Brandolin, G., and Palmieri, F. (1994) *Biochemistry* **33**, 3705 – 3713

296. Fiermonte, G., Palmieri, L., Dolce, V., Lasorsa, F. M., Palmieri, F., Runswick, M. J., and Walker, J. E. (1998) *J. Biol. Chem.* **273**, 24754 – 24759

297. Nelson, D. R., Felix, C. M., and Swanson, J. M. (1998) *J. Mol. Biol.* **277**, 285 – 308

298. Kaplan, R. S., Mayor, J. A., and Wood, D. O. (1993) *J. Biol. Chem.* **268**, 13682 – 13690

298a. Xu, Y., Kakhniashvili, D. A., Gremse, D. A., Wood, D. O., Mayor, J. A., Walters, D. E., and Kaplan, R. S. (2000) *J. Biol. Chem.* **275**, 7117 – 7124

298b. Bandell, M., and Lolkema, J. S. (2000) *J. Biol. Chem.* **275**, 39130 – 39136

299. Müller, V., Basset, G., Nelson, D. R., and Klingenberg, M. (1996) *Biochemistry* **35**, 16132 – 16143

300. Broustovetsky, N., Bamberg, E., Gropp, T., and Klingenberg, M. (1997) *Biochemistry* **36**, 13865 – 13872

300a. Wiedemann, N., Pfanner, N., and Ryan, M. T. (2001) *EMBO J.* **20**, 951 – 960

300b. Heimpel, S., Basset, G., Odoy, S., and Klingenberg, M. (2001) *J. Biol. Chem.* **276**, 11499 – 11506

301. Beyer, K., and Klingenberg, M. (1985) *Biochemistry* **24**, 3821 – 3826

302. Muller, M., Krebs, J. J. R., Cherry, R. J., and Kawato, S. (1984) *J. Biol. Chem.* **259**, 3037 – 3043

303. Ziegler, M., and Penefsky, H. S. (1993) *J. Biol. Chem.* **268**, 25320 – 25328

304. Schroers, A., Burkovski, A., Wohlrab, H., and Krämer, R. (1998) *J. Biol. Chem.* **273**, 14269 – 14276

305. Briggs, C., Mincone, L., and Wohlrab, H. (1999) *Biochemistry* **38**, 5096 – 5102

305a. Majima, E., Ishida, M., Miki, S., Shinohara, Y., and Terada, H. (2001) *J. Biol. Chem.* **276**, 9792 – 9799

306. Kaplan, R. S., Pratt, R. D., and Pederson, P. L. (1986) *J. Biol. Chem.* **261**, 12767 – 12773

307. Nicolli, A., Petronilli, V., and Bernardi, P. (1993) *Biochemistry* **32**, 4461 – 4465

307a. Scorrano, L., Petronilli, V., Colonna, R., Di Lisa, F., and Bernardi, P. (1999) *J. Biol. Chem.* **274**, 24657 – 24663

308. Liu, G., Hinch, B., Davatol-Hag, H., Lu, Y., Powers, M., and Beavis, A. D. (1996) *J. Biol. Chem.* **271**, 19717 – 19723

309. Sugano, T., Handler, J. A., Yoshihara, H., Kizaki, Z., and Thurman, R. G. (1990) *J. Biol. Chem.* **265**, 21549 – 21553

310. Schlegel, H. G., and Eberhardt, U. (1972) *Adv. Microbiol. Physiol.* **7**, 205 – 242

310a. Pierik, A. J., Roseboom, W., Happe, R. P., Bagley, K. A., and Albracht, S. P. J. (1999) *J. Biol. Chem.* **274**, 3331 – 3337

311. Bothe, H., and Trebst, A., eds. (1981) *Biology of Inorganic Nitrogen and Sulfur*, Springer, Berlin

312. Arciero, D. M., Golombek, A., Hendrich, M. P., and Hooper, A. B. (1998) *Biochemistry* **37**, 523 – 529

313. Igarashi, N., Moriyama, H., Fujiwara, T., Fukumori, Y., and Tanaka, N. (1997) *Nature Struct. Biol.* **4**, 276 – 284

313a. Schalk, J., de Vries, S., Kuenen, J. G., and Jetten, M. S. M. (2000) *Biochemistry* **39**, 5405 – 5412

314. Hendrich, M. P., Petasis, D., Arciero, D. M., and Hooper, A. B. (2001) *J. Am. Chem. Soc.* **123**, 2997 – 3005

314a. Hendrich, M. P., Upadhyay, A. K., Riga, J., Arciero, D. M., and Hooper, A. B. (2002) *Biochemistry* **41**, 4603 – 4611

315. Blaut, M., and Gottschalk, G. (1997) in *Bioenergetics* (Gräber, P., and Milazzo, G., eds), pp. 139 – 211, Birkhäuser Verlag, Basel

316. DiSpirito, A. A., and Hooper, A. B. (1986) *J. Biol. Chem.* **261**, 10534 – 10537

317. Logan, M. S. P., and Hooper, A. B. (1995) *Biochemistry* **34**, 9257 – 9264

318. Taylor, C. D., and Wirsen, C. O. (1997) *Science* **277**, 1483 – 1485

319. Felbeck, H., and Somero, G. N. (1982) *Trends Biochem. Sci.* **7**, 201 – 204

320. Gaill, F. (1993) *FASEB J.* **7**, 558 – 565

321. Trudinger, P. A. (1969) *Ad. Microbiol. Physiol.* **3**, 111 – 158

322. Roy, A. B., and Trudinger, P. A. (1970) *The Biochemistry of Inorganic Compounds of Sulfur*, Cambridge Univ. Press, London and New York

322a. Cheesman, M. R., Little, P. J., and Berks, B. C. (2001) *Biochemistry* **40**, 10562 – 10569

323. Postgate, J. R., and Kelly, D. P., eds. (1982) *Sulphur Bacteria*, The Royal Society,

323a. Griesbeck, C., Schütz, M., Schödl, T., Bathe, S., Nausch, L., Mederer, N., Vielreicher, M., and Hauska, G. (2002) *Biochemistry* **41**, 11552 – 11565

324. Chen, Z.-w, Koh, M., Van Driessche, G., Van Beeumen, J. J., Bartsch, R. G., Meyer, T. E., Cusanovich, M. A., and Mathews, F. S. (1994) *Science* **266**, 430 – 432

324a. Edwards, K. J., Bond, P. L., Gihring, T. M., and Banfield, J. F. (2000) *Science* **287**, 1796 – 1799

325. Hodgman, C. D., ed. (1967 – 1968) *CRC Handbook of Chemistry and Physics*, 48th ed., Chem. Rubber Publ. Co., Cleveland, Ohio

326. Dugan, P. R. (1972) *Biochemical Ecology of Water Pollution*, Plenum, New York

327. Powell, M. A., and Somero, G. N. (1986) *Science* **233**, 563 – 566

328. Rothery, R. A., Blasco, F., and Weiner, J. H. (2001) *Biochemistry* **40**, 5260 – 5268

328a. Anderson, L. J., Richardson, D. J., and Butt, J. N. (2001) *Biochemistry* **40**, 11294 – 11307

329. Magalon, A., Asso, M., Guigliarelli, B., Rothery, R. A., Bertrand, P., Giordano, G., and Blasco, F. (1998) *Biochemistry* **37**, 7363 – 7370

329a. Richardson, D., and Sawers, G. (2002) *Science* **295**, 1842 – 1843

329b. Jormakka, M., Törnroth, S., Byrne, B., and Iwata, S. (2002) *Science* **295**, 1863 – 1868

330. Iuchi, S., and Lin, E. C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3901 – 3905

331. Spiro, S., and Guest, J. R. (1991) *Trends Biochem. Sci.* **16**, 310 – 314

332. Cecchini, G., Thompson, C. R., Ackrell, B. A. C., Westenberg, D. J., Dean, N., and Gunsalus, R. P. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8898 – 8902

332a. Turner, K. L., Doherty, M. K., Heering, H. A., Armstrong, F. A., Reid, G. A., and Chapman, S. K. (1999) *Biochemistry* **38**, 3302 – 3309

333. Czjzek, M., Dos Santos, J.-P., Pommier, J., Giordano, G., Méjean, V., and Haser, R. (1998) *J. Mol. Biol.* **284**, 435 – 447

334. Trieber, C. A., Rothery, R. A., and Weiner, J. H. (1994) *J. Biol. Chem.* **269**, 7103 – 7109

335. Rothery, R. A., and Weiner, J. H. (1996) *Biochemistry* **35**, 3247 – 3257

336. Costa, C., Moura, J. J. G., Moura, I., Liu, M. Y., Peck, H. D., Jr., LeGall, J., Wang, Y., and Huynh, B. H. (1990) *J. Biol. Chem.* **265**, 14382 – 14387

337. Blackmore, R. S., Gibson, Q. H., and Greenwood, C. (1992) *J. Biol. Chem.* **267**, 10950 – 10955

338. Ferguson, S. J. (1987) *Trends Biochem. Sci.* **12**, 353 – 357

339. Wang, Y., and Averill, B. A. (1996) *J. Am. Chem. Soc.* **118**, 3972 – 3973

339a. Suharti, Strampraad, M. J. F., Schröder, I., and de Vries, S. (2001) *Biochemistry* **40**, 2632 – 2639

340. Nurizzo, D., Cutruzzolà, F., Arese, M., Bourgeois, D., Brunori, M., Cambillau, C., and Tegoni, M. (1998) *Biochemistry* **37**, 13987 – 13996

341. Kukimoto, M., Nishiyama, M., Tanokura, M., Adman, E. T., and Horinouchi, S. (1996) *J. Biol. Chem.* **271**, 13680 – 13683

342. Murphy, M. E. P., Turley, S., and Adman, E. T. (1997) *J. Biol. Chem.* **272**, 28455 – 28460

342a. Peters-Libeu, C. A., Kukimoto, M., Nishiyama, M., Horinouchi, S., Adman, E. T. (1997) *Biochemistry* **36**, 13160 – 13179

342b. Inoue, T., Nishio, N., Suzuki, S., Kataoka, K., Kohzuma, T., and Kai, Y. (1999) *J. Biol. Chem.* **274**, 17845 – 17852

343. Weeg-Aerssens, E., Wu, W., Ye, R. W., Tiedje, J. M., and Chang, C. K. (1991) *J. Biol. Chem.* **266**, 7496 – 7502

344. Williams, P. A., Fülöp, V., Garman, E. F., Saunders, N. F. W., Ferguson, S. J., and Hajdu, J. (1997) *Nature (London)* **389**, 406 – 412

345. Baker, S. C., Saunders, N. F. W., Willis, A. C., Ferguson, S. J., Hajdu, J., and Fülöp, V. (1997) *J. Mol. Biol.* **269**, 440 – 455

346. Nurizzo, D., Silvestrini, M.-C., Mathieu, M., Cutruzzolà, F., Bourgeois, D., Fülöp, V., Hajdu, J., Brunori, M., Tegoni, M., and Cambillau, C. (1997) *Structure* **5**, 1157 – 1171

346a. Cutruzzolà, F., Brown, K., Wilson, E. K., Bellelli, A., Arese, M., Tegoni, M., Cambillau, C., and Brunori, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2232 – 2237

346b. Sjögren, T., Svensson-EK, M., Hajdu, J., and Brzezinki, P. (2000) *Biochemistry* **39**, 10967 – 10974

346c. Einsle, O., Messerschmidt, A., Stach, P., Bourenkov, G. P., Bartunik, H. D., Huber, R., and Kroneck, P. M. H. (1999) *Nature (London)* **400**, 476 – 480

346d. Einsle, O., Stach, P., Messerschmidt, A., Simon, J., Kröger, A., Huber, R., and Kroneck, P. M. H. (2000) *J. Biol. Chem.* **275**, 39608 – 39616

346e. Sjögren, T., and Hajdu, J. (2001) *J. Biol. Chem.* **276**, 29450 – 29455

346f. Ranghino, G., Scorza, E., Sjögren, T., Williams, P. A., Ricci, M., and Hajdu, J. (2000) *Biochemistry* **39**, 10958 – 10966

References

346g. Kobayashi, K., Koppenhöfer, A., Ferguson, S. J., Watmough, N. J., and Tagawa, S. (2001) *Biochemistry* **40**, 8542–8547

347. Kobayashi, K., Koppenhöfer, A., Ferguson, S. J., and Tagawa, S. (1997) *Biochemistry* **36**, 13611–13616

348. Cheesman, M. R., Ferguson, S. J., Moir, J. W. B., Richardson, D. J., Zumft, W. G., and Thomson, A. J. (1997) *Biochemistry* **36**, 16267–16276

349. Sakurai, N., and Sakurai, T. (1997) *Biochemistry* **36**, 13809–13815

350. Hendriks, J., Warne, A., Gohlke, U., Haltia, T., Ludovici, C., Lübben, M., and Saraste, M. (1998) *Biochemistry* **37**, 13102–13109

351. Moënne-Loccoz, P., and de Vries, S. (1998) *J. Am. Chem. Soc.* **120**, 5147–5152

352. Shiro, Y., Fujii, M., Isogai, Y., Adachi, S.-i, Iizuka, T., Obayashi, E., Makino, R., Nakahara, K., and Shoun, H. (1995) *Biochemistry* **34**, 9052–9058

353. Neese, F., Zumft, W. G., Antholine, W. E., and Kroneck, P. M. H. (1996) *J. Am. Chem. Soc.* **118**, 8692–8699

354. Farrar, J. A., Zumft, W. G., and Thomson, A. J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9891–9896

355. Postgate, J. R., FRS. (1984) in *The Sulfate-Reducing Bacteria*, 2nd ed., pp. 56–63, Cambridge University Press, Cambridge

356. Postgate, J. R. (1984) *The Sulfate-reducing Bacteria*, 2nd ed., Cambridge Univ. Press, London

356a. Michaelis, W., and 16 other authors. (2002) *Science* **297**, 1013–1015

357. Grey, D. C., and Jensen, M. L. (1972) *Science* **177**, 1099–1100

358. Gavel, O. Y., Bursakov, S. A., Calvete, J. J., George, G. N., Moura, J. J. G., and Moura, I. (1998) *Biochemistry* **37**, 16225–16232

359. Klenk, H.-P., and 50 other authors. (1997) *Nature (London)* **390**, 364–370

360. Morais, J., Palma, P. N., Frazao, C., Caldeira, J., LeGall, J., Moura, I., Moura, J. J. G., and Carrondo, M. A. (1995) *Biochemistry* **34**, 12830–12841

361. Louro, R. O., Catarino, T., Turner, D. L., Picarra-Pereira, M. A., Pacheco, I., LeGall, J., and Xavier, A. V. (1998) *Biochemistry* **37**, 15808–15815

361a. Fritz, G., Griesshaber, D., Seth, O., and Kroneck, P. M. H. (2001) *Biochemistry* **40**, 1317–1324

362. Florens, L., Ivanova, M., Dolla, A., Czjzek, M., Haser, R., Verger, R., and Bruschi, M. (1995) *Biochemistry* **34**, 11327–11334

363. Banci, L., Bertini, I., Bruschi, M., Sompornpisut, P., and Turano, P. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14396–14400

364. Feng, Y., and Swenson, R. P. (1997) *Biochemistry* **36**, 13617–13628

365. Lui, S. M., and Cowan, J. A. (1994) *J. Am. Chem. Soc.* **116**, 11538–11549

366. Crane, B. R., Siegel, L. M., and Getzoff, E. D. (1997) *Biochemistry* **36**, 12101–12119

367. Moura, I., LeGall, J., Lino, A. R., Peck, H. D., Jr., Fauque, G., Xavier, A. V., Der Vertanian, D. V., Moura, J. J. G., and Huynh, B. H. (1988) *J. Am. Chem. Soc.* **110**, 1075–1082

368. Fischer, F., Zillig, W., Stetter, K. O., and Schreiber, G. (1983) *Nature (London)* **301**, 511–515

369. Kelly, D. P. (1985) *Nature (London)* **313**, 734

370. Bak, F., and Cypionka, H. (1987) *Nature (London)* **326**, 891–892

370a. Morelli, X., Czjzek, M., Hatchikian, C. E., Bornet, O., Fontecilla-Camps, J. C., Palma, N. P., Moura, J. J. G., and Guerlesquin, F. (2000) *J. Biol. Chem.* **275**, 23204–23210

370b. Costas, A. M. C., White, A. K., and Metcalf, W. W. (2001) *J. Biol. Chem.* **276**, 17429–17436

371. Müller, V., Blaut, M., and Gottschalk, G. (1993) in *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics* (Ferry, J. G., ed), pp. 360–406, Chapman and Hall, New York

372. Hayaishi, O., and Nozaki, M. (1969) *Science* **164**, 389–405

373. Orville, A. M., and Lipscomb, J. D. (1993) *J. Biol. Chem.* **268**, 8596–8607

374. Ohlendorf, D. H., Orville, A. M., and Lipscomb, J. D. (1994) *J. Mol. Biol.* **244**, 586–608

375. Orville, A. M., Lipscomb, J. D., and Ohlendorf, D. H. (1997) *Biochemistry* **36**, 10052–10066

376. Sanvoisin, J., Langley, G. J., and Bugg, T. D. H. (1995) *J. Am. Chem. Soc.* **117**, 7836–7837

377. Miller, M. A., and Lipscomb, J. D. (1996) *J. Biol. Chem.* **271**, 5524–5535

378. Han, S., Eltis, L. D., Timmis, K. N., Muchmore, S. W., and Bolin, J. T. (1995) *Science* **270**, 976–980

379. Shu, L., Chiou, Y.-M., Orville, A. M., Miller, M. A., Lipscomb, J. D., and Que, L., Jr. (1995) *Biochemistry* **34**, 6649–6659

380. Senda, T., Sugiyama, K., Narita, H., Yamamoto, T., Kimbara, K., Fukuda, M., Sato, M., Yano, K., and Mitsui, Y. (1996) *J. Mol. Biol.* **255**, 735–752

381. Hugo, N., Armengaud, J., Gaillard, J., Timmis, K. N., and Jouanneau, Y. (1998) *J. Biol. Chem.* **273**, 9622–9629

382. Fraser, M. S., and Hamilton, G. A. (1982) *J. Am. Chem. Soc.* **104**, 4203–4211

383. Leeds, J. M., Brown, P. J., McGeehan, G. M., Brown, F. K., and Wiseman, J. S. (1993) *J. Biol. Chem.* **268**, 17781–17786

384. Ohnishi, T., Hirata, F., and Hayaishi, O. (1977) *J. Biol. Chem.* **252**, 4643–4647

385. Gassner, G. T., Ballou, D. P., Landrum, G. A., and Whittaker, J. W. (1993) *Biochemistry* **32**, 4820–4825

386. Yamaguchi, M., and Fujisawa, H. (1982) *J. Biol. Chem.* **257**, 12497–12502

386a. Senda, T., Yamada, T., Sakurai, N., Kubota, M., Nishizaki, T., Masai, E., Fukuda, M., and Mitsui, Y. (2000) *J. Mol. Biol.* **304**, 397–410

387. Gassner, G. T., Ludwig, M. L., Gatti, D. L., Correll, C. C., and Ballou, D. P. (1995) *FASEB J.* **9**, 1411–1418

387a. Coulter, E. D., Moon, N., Batie, C. J., Dunham, W. R., and Ballou, D. P. (1999) *Biochemistry* **38**, 11062–11072

388. Bertrand, P., More, C., and Camensuli, P. (1995) *J. Am. Chem. Soc.* **117**, 1807–1809

389. Lockridge, O., Massey, V., and Sullivan, P. A. (1972) *J. Biol. Chem.* **247**, 8097–8106

390. Sanders, S. A., Williams, C. H., Jr., and Massey, V. (1999) *J. Biol. Chem.* **274**, 22289–22295

391. Flashner, M. I. S., and Massey, V. (1974) *J. Biol. Chem.* **249**, 2579–2586

392. Emanuele, J. J., and Fitzpatrick, P. F. (1995) *Biochemistry* **34**, 3710–3715

393. Massey, V. (1994) *J. Biol. Chem.* **269**, 22459–22462

394. Entsch, B., and van Berkel, W. J. H. (1995) *FASEB J.* **9**, 476–483

395. Gatti, D. L., Entsch, B., Ballou, D. P., and Ludwig, M. L. (1996) *Biochemistry* **35**, 567–578

396. Hamilton, G. A. (1971) *Prog. Bioorg. Chem.* **1**, 83

397. Orf, H. W., and Dolphin, D. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 2646–2650

398. Wagner, W. R., Spero, D. M., and Rastetter, W. H. (1984) *J. Am. Chem. Soc.* **106**, 1476–1480

399. Moran, G. R., Entsch, B., Palfey, B. A., and Ballou, D. P. (1997) *Biochemistry* **36**, 7548–7556

399a. Frederick, K. K., Ballou, D. P., and Palfey, B. A. (2001) *Biochemistry* **40**, 3891–3899

400. Eppink, M. H. M., Schreuder, H. A., and van Berkel, W. J. H. (1998) *J. Biol. Chem.* **273**, 21031–21039

401. van der Bolt, F. J. T., van den Heuvel, R. H. H., Vervoort, J., and van Berkel, W. J. H. (1997) *Biochemistry* **36**, 14192–14201

401a. Eppink, M. H. M., Overkamp, K. M., Schreuder, H. A., and Van Berkel, W. J. H. (1999) *J. Mol. Biol.* **292**, 87–96

401b. Ortiz-Maldonado, M., Ballou, D. P., and Massey, V. (2001) *Biochemistry* **40**, 1091–1101

402. Eppink, M. H. M., Schreuder, H. A., and van Berkel, W. J. H. (1997) *Protein Sci.* **6**, 2454–2458

403. Maeda-Yorita, K., and Massey, V. (1993) *J. Biol. Chem.* **268**, 4134–4144

404. Einarsdottir, G. H., Stankovich, M. T., and Tu, S.-C. (1988) *Biochemistry* **27**, 3277–3285

405. Powlowski, J., Ballou, D. P., and Massey, V. (1990) *J. Biol. Chem.* **265**, 4969–4975

406. Mathews, F. S., Chen, Z.-w, and Bellamy, H. D. (1991) *Biochemistry* **30**, 238–247

407. Arunachalam, U., Massey, V., and Miller, S. M. (1994) *J. Biol. Chem.* **269**, 150–155

408. Prieto, M. A., and Garcia, J. L. (1994) *J. Biol. Chem.* **269**, 22823–22829

409. Ysuji, H., Ogawa, T., Bando, N., Kimoto, M., and Sasaoka, K. (1990) *J. Biol. Chem.* **265**, 16064–16067

410. Itagaki, K., Carver, G. T., and Philpot, R. M. (1996) *J. Biol. Chem.* **271**, 20102–20107

411. Jones, K. C., and Ballou, D. P. (1986) *J. Biol. Chem.* **261**, 2553–2559

411a. Sheng, D., Ballou, D. P., and Massey, V. (2001) *Biochemistry* **40**, 11156–11167

412. Branchaud, B. P., and Walsh, C. T. (1985) *J. Am. Chem. Soc.* **107**, 2153–2161

413. Erlandsen, H., Flatmark, T., Stevens, R. C., and Hough, E. (1998) *Biochemistry* **37**, 15638–15646

414. Loeb, K. E., Westre, T. E., Kappock, T. J., Mitic, N., Glasfeld, E., Caradonna, J. P., Hedman, B., Hodgson, K. O., and Solomon, E. I. (1997) *J. Am. Chem. Soc.* **119**, 1901–1915

415. Chen, D., and Frey, P. A. (1998) *J. Biol. Chem.* **273**, 25594–25601

415a. Teigen, K., Froystein, N. Å., and Martínez, A. (1999) *J. Mol. Biol.* **294**, 807–823

416. Goodwill, K. E., Sabatier, C., and Stevens, R. C. (1998) *Biochemistry* **37**, 13437–13445

417. Francisco, W. A., Tian, G., Fitzpatrick, P. F., and Klinman, J. P. (1998) *J. Am. Chem. Soc.* **120**, 4057–4062

418. Michaud-Soret, I., Andersson, K. K., and Que, L., Jr. (1995) *Biochemistry* **34**, 5504–5510

419. Hillas, P. J., and Fitzpatrick, P. F. (1996) *Biochemistry* **35**, 6969–6975

419a. Almås, B., Toska, K., Teigen, K., Groehn, V., Pfleiderer, W., Martínez, A., Flatmark, T., and Haavik, J. (2000) *Biochemistry* **39**, 13676–13686

420. Moran, G. R., Daubner, S. C., and Fitzpatrick, P. F. (1998) *J. Biol. Chem.* **273**, 12259–12266

421. Ramsey, A. J., Hillas, P. J., and Fitzpatrick, P. F. (1996) *J. Biol. Chem.* **271**, 24395–24400

421a. Fitzpatrick, P. F. (1999) *Ann. Rev. Biochem.* **68**, 355–381

422. Lazarus, R. A., DeBrosse, C. W., and Benkovic, S. J. (1982) *J. Am. Chem. Soc.* **104**, 6869–6871

423. Pike, D. C., Hora, M. T., Bailey, S. W., and Ayling, J. E. (1986) *Biochemistry* **25**, 4762–4771

424. Dix, T. A., Bollag, G. E., Domanico, P. L., and Benkovic, S. J. (1985) *Biochemistry* **24**, 2955–2958

425. Rebrin, I., Thöny, B., Bailey, S. W., and Ayling, J. E. (1998) *Biochemistry* **37**, 11246–11254

425a. Endrizzi, J. A., Cronk, J. D., Wang, W., Crabtree, G. R., and Alber, T. (1995) *Science* **268**, 556–559

426. Ficner, R., Sauer, U. H., Stier, G., and Suck, D. (1995) *EMBO J.* **14**, 2034–2042

References

427. Su, Y., Varughese, K. I., Xuong, N. H., Bray, T. L., Roche, D. J., and Whiteley, J. M. (1993) *J. Biol. Chem.* **268**, 26836 – 26841
428. Varughese, K. I., Xuong, N. H., Kiefer, P. M., Matthews, D. A., and Whiteley, J. M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5582 – 5586
429. Kiefer, P. M., Grimshaw, C. E., and Whiteley, J. M. (1997) *Biochemistry* **36**, 9438 – 9445
430. Scriver, C. R., Kaufman, S., Eisensmith, R. C., and Woo, S. L. C. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1015 – 1075, McGraw-Hill, New York

430a. Kobe, B., Jennings, I. G., House, C. M., Michell, B. J., Goodwill, K. E., Santarsiero, B. D., Stevens, R. C., Cotton, R. G. H., and Kemp, B. E. (1999) *Nature Struct. Biol.* **6**, 442 – 448

431. Ginns, E. L., Rehavi, M., Martin, B. M., Weller, M., O'Malley, K. L., LeMarca, M. E., McAllister, C. G., and Paul, S. M. (1988) *J. Biol. Chem.* **263**, 7406 – 7410
432. Grenett, H. E., Ledley, F. D., Reed, L. L., and Woo, S. L. C. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5530 – 5534
433. Pigeon, D., Ferrara, P., Gros, F., and Thibault, J. (1987) *J. Biol. Chem.* **262**, 6155 – 6158

433a. Itagaki, C., Isobe, T., Taoka, M., Natsume, T., Nomura, N., Horigome, T., Omata, S., Ichinose, H., Nagatsu, T., Greene, L. A., and Ichimura, T. (1999) *Biochemistry* **38**, 15673 – 15680

434. McTigue, M., Cremins, J., and Halegoua, S. (1985) *J. Biol. Chem.* **260**, 9047 – 9056
435. Okuno, S., and Fujisawa, H. (1985) *J. Biol. Chem.* **260**, 2633 – 2635
436. Albert, K. A., Helmer-Matyjek, E., Nairn, A. C., Muller, T. H., Waycock, J. W., Greene, L. A., Goldstein, M., and Greengard, P. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7713 – 7717
437. Guroff, G., Daly, J. W., Jerina, D. M., Renson, J., Witkop, B., and Udenfriend, S. (1967) *Science* **157**, 1524 – 1530
438. Kasperek, G. J., Bruice, T. C., Yagi, H., Kaubisch, N., and Jerina, D. M. (1972) *J. Am. Chem. Soc.* **94**, 7876 – 7882
439. Endo, F., Awata, H., Tanoue, A., Ishiguro, M., Eda, Y., Titani, K., and Matsuda, I. (1992) *J. Biol. Chem.* **267**, 24235 – 24240
440. Bradley, F. C., Lindstedt, S., Lipscomb, J. D., Que, L., Jr., Roe, A. L., and Rundgren, M. (1986) *J. Biol. Chem.* **261**, 11693 – 11696
441. McGinnis, K., Ku, G. M., VanDusen, W. J., Fu, J., Garsky, V., Stern, A. M., and Friedman, P. A. (1996) *Biochemistry* **35**, 3957 – 3962
442. Lamberg, A., Pihlajaniemi, T., and Kivirikko, K. I. (1995) *J. Biol. Chem.* **270**, 9926 – 9931
443. Myllyharju, J., and Kivirikko, K. I. (1997) *EMBO J.* **16**, 1173 – 1180
444. Annunen, P., Autio-Harmainen, H., and Kivirikko, K. I. (1998) *J. Biol. Chem.* **273**, 5989 – 5992
445. Myllylä, R., Pihlajaniemi, T., Pajunen, L., Turpeenniemi-Hujanen, T., and Kivirikko, K. I. (1991) *J. Biol. Chem.* **266**, 2805 – 2810
446. Valtavaara, M., Szpirer, C., Szpirer, J., and Myllylä, R. (1998) *J. Biol. Chem.* **273**, 12881 – 12886
447. Bolwell, G. P., Robbins, M. P., and Dixon, R. A. (1985) *Biochem. J.* **229**, 693 – 699
448. Thornburg, L. D., and Stubbe, J. (1993) *Biochemistry* **32**, 14034 – 14042
449. Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997) *J. Biol. Chem.* **272**, 23031 – 23036

449a. Elkins, J. M., Ryle, M. J., Clifton, I. J., Hotopp, J. C. D., Lloyd, J. S., Burzlaff, N. I., Baldwin, J. E., Hausinger, R. P., and Roach, P. L. (2002) *Biochemistry* **41**, 5185 – 5192

450. Whiting, A. K., Que, L., Jr., Saari, R. E., Hausinger, R. P., Frederick, M. A., and McCracken, J. (1997) *J. Am. Chem. Soc.* **119**, 3413 – 3414
451. Hotopp, J. C. D., and Hausinger, R. P. (2002) *Biochemistry* **41**, 9787 – 9794
452. Henderson, L. M., Nelson, P. J., and Henderson, L. (1982) *Fed. Proc.* **41**, 2843 – 2847
453. Englard, S., Blanchard, J. S., and Midelfort, C. F. (1985) *Biochemistry* **24**, 1110 – 1116
454. Myllyla, R., Majamaa, K., Gunzler, V., Hanauske-Abel, H. M., and Kivirikko, K. I. (1984) *J. Biol. Chem.* **259**, 5403 – 5405
455. Ho, R. Y. N., Mehn, M. P., Hegg, E. L., Liu, A., Ryle, M. J., Hausinger, R. P., and Que, L., Jr. (2001) *J. Am. Chem. Soc.* **123**, 5022 – 5029
456. Roach, P. L., Clifton, I. J., Fülöp, V., Harlos, K., Barton, G. J., Hajdu, J., Andersson, I., Schofield, C. J., and Baldwin, J. E. (1995) *Nature (London)* **375**, 700 – 704
457. Roach, P. L., Clifton, I. J., Hensgens, C. M. H., Shibata, N., Schofield, C. J., Hajdu, J., and Baldwin, J. E. (1997) *Nature (London)* **387**, 827 – 830

457a. Lloyd, M. D., Lee, H.-J., Harlos, K., Zhang, Z.-H., Baldwin, J. E., Schofield, C. J., Charnock, J. M., Garner, C. D., Hara, T., Terwisscha van Scheltinga, A. C., Valegård, K., Viklund, J. A. C., Hajdu, J., Andersson, I., Danielsson, Å., and Bhikhabhai, R. (1999) *J. Mol. Biol.* **287**, 943 – 960

458. Busby, R. W., Chang, M. D.-T., Busby, R. C., Wimp, J., and Townsend, C. A. (1995) *J. Biol. Chem.* **270**, 4262 – 4269
459. Zhou, J., Kelly, W. L., Bachmann, B. O., Gunsior, M., Townsend, C. A., and Solomon, E. I. (2001) *J. Am. Chem. Soc.* **123**, 7388 – 7398
460. Tian, G., Berry, J. A., and Klinman, J. P. (1994) *Biochemistry* **33**, 226 – 234
461. Klinman, J. P., Krueger, M., Brenner, M., and Edmondson, D. E. (1984) *J. Biol. Chem.* **259**, 3399 – 3402
462. Merkler, D. J., Kulathila, R., Consalvo, A. P., Young, S. D., and Ash, D. E. (1992) *Biochemistry* **31**, 7282 – 7288
463. Prigge, S. T., Kolhekar, A. S., Eipper, B. A., Mains, R. E., and Amzel, L. M. (1997) *Science* **278**, 1300 – 1305

463a. Jaron, S., and Blackburn, N. J. (2001) *Biochemistry* **40**, 6867 – 6875

464. Francisco, W. A., Merkler, D. J., Blackburn, N. J., and Klinman, J. P. (1998) *Biochemistry* **37**, 8244 – 8252

464a. Jaron, S., and Blackburn, N. J. (1999) *Biochemistry* **38**, 15086 – 15096

464b. Kolhekar, A. S., Bell, J., Shiozaki, E. N., Jin, L., Keutmann, H. T., Hand, T. A., Mains, R. E., and Eipper, B. A. (2002) *Biochemistry* **41**, 12384 – 12394

465. Wilkinson, B., Zhu, M., Priestley, N. D., Nguyen, H.-H. T., Morimoto, H., Williams, P. G., Chan, S. I., and Floss, H. G. (1996) *J. Am. Chem. Soc.* **118**, 921 – 922
466. Elliott, S. J., Zhu, M., Tso, L., Nguyen, H.-H. T., Yip, J. H.-K., and Chan, S. I. (1997) *J. Am. Chem. Soc.* **119**, 9949 – 9955
467. Swinney, D. C., and Mak, A. Y. (1994) *Biochemistry* **33**, 2185 – 2190
468. Imai, T., Yamazaki, T., and Kominami, S. (1998) *Biochemistry* **37**, 8097 – 8104
469. Denison, M. S., and Whitlock, J. P., Jr. (1995) *J. Biol. Chem.* **270**, 18175 – 18178
470. Coon, M. J., Ding, X., Pernecky, S. J., and Vaz, A. D. N. (1992) *FASEB J.* **6**, 669 – 673
471. Gonzalez, F. J., and Lee, Y.-H. (1996) *FASEB J.* **10**, 1112 – 1117
472. Nishimoto, M., Gotoh, O., Okuda, K., and Noshiro, M. (1991) *J. Biol. Chem.* **266**, 6467 – 6471
473. Dilworth, F. J., Scott, I., Green, A., Strugnell, S., Guo, Y.-D., Roberts, E. A., Kremer, R., Calverley, M. J., Makin, H. L. J., and Jones, G. (1995) *J. Biol. Chem.* **270**, 16766 – 16774
474. Cabello-Hurtado, F., Batard, Y., Salaün, J.-P., Durst, F., Pinot, F., and Werck-Reichhart, D. (1998) *J. Biol. Chem.* **273**, 7260 – 7267
475. Meyer, K., Shirley, A. M., Cusumano, J. C., Bell-Lelong, D. A., and Chapple, C. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6619 – 6623
476. Harris, D. L., and Loew, G. H. (1998) *J. Am. Chem. Soc.* **120**, 8941 – 8948

476a. French, K. J., Strickler, M. D., Rock, D. A., Rock, D. A., Bennett, G. A., Wahlstrom, J. L., Goldstein, B. M., and Jones, J. P. (2001) *Biochemistry* **40**, 9532 – 9538

476b. Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, R. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000) *Science* **287**, 1615 – 1622

477. Vaz, A. D. N., McGinnity, D. F., and Coon, M. J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3555 – 3560
478. Tzeng, H.-F., Laughlin, L. T., and Armstrong, R. N. (1998) *Biochemistry* **37**, 2905 – 2911
479. De Montellano, P. R. O., ed. (1995) *Cytochrome P450. Structure, Mechanism, and Biochemistry*, 2nd ed., Plenum, New York
480. Coon, M. J., Vaz, A. D. N., and Bestervelt, L. L. (1996) *FASEB J.* **10**, 428 – 434
481. Negishi, M., Uno, T., Darden, T. A., Sueyoshi, T., and Pedersen, L. G. (1996) *FASEB J.* **10**, 683 – 689
482. Gilday, D., Gannon, M., Yutzey, K., Bader, D., and Rifkind, A. B. (1996) *J. Biol. Chem.* **271**, 33054 – 33059
483. Khani, S. C., Zophiropoulos, P. G., Fujita, V. S., Porter, T. D., Koop, D. R., and Coon, M. J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 638 – 642

483a. Pikuleva, I. A., Puchkaev, A., and Björkhem, I. (2001) *Biochemistry* **40**, 7621 – 7629

484. Porter, T. D. (1991) *Trends Biochem. Sci.* **16**, 154 – 158
485. Hubbard, P. A., Shen, A. L., Paschke, R., Kasper, C. B., and Kim, J.-J. P. (2001) *J. Biol. Chem.* **276**, 29163 – 29170

485a. Gutierrez, A., Lian, L.-Y., Wolf, C. R., Scrutton, N. S., and Roberts, G. C. K. (2001) *Biochemistry* **40**, 1964 – 1975

486. Perret, A., and Pompon, D. (1998) *Biochemistry* **37**, 11412 – 11424
487. Vidakovic, M., Sligar, S. G., Li, H., and Poulos, T. L. (1998) *Biochemistry* **37**, 9211 – 9219
488. Takemori, S., and Kominami, S. (1984) *Trends Biochem. Sci.* **9**, 393 – 396
489. Kobayashi, K., Miura, S., Miki, M., Ichikawa, Y., and Tagawa, S. (1995) *Biochemistry* **34**, 12932 – 12936

489a. Pochapsky, T. C., Jain, N. U., Kuti, M., Lyons, T. A., and Heymont, J. (1999) *Biochemistry* **38**, 4681 – 4690

490. Poulos, T. L., and Raag, R. (1992) *FASEB J.* **6**, 674 – 679
491. Murataliev, M. B., Klein, M., Fulco, A., and Feyereisen, R. (1997) *Biochemistry* **36**, 8401 – 8412
492. Ravichandran, K. G., Boddupalli, S. S., Hasemann, C. A., Peterson, J. A., and Deisenhofer, J. (1993) *Science* **261**, 731 – 736
493. Hasemann, C. A., Ravichandran, K. G., Peterson, J. A., and Deisenhofer, J. (1994) *J. Mol. Biol.* **236**, 1169 – 1185
494. Harris, D. L., and Loew, G. H. (1996) *J. Am. Chem. Soc.* **118**, 6377 – 6387

References

495. Higgins, L., Bennett, G. A., Shimoji, M., and Jones, J. P. (1998) *Biochemistry* **37**, 7039 – 7046
496. Toy, P. H., Newcomb, M., and Hollenberg, P. F. (1998) *J. Am. Chem. Soc.* **120**, 7719 – 7729
496a. French, K. J., Strickler, M. D., Rock, D. A., Rock, D. A., Bennett, G. A., Wahlstrom, J. L., Goldstein, B. M., and Jones, J. P. (2001) *Biochemistry* **40**, 9532 – 9538
496b. Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, R. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000) *Science* **287**, 1615 – 1622
497. Harris, D., Loew, G., and Waskell, L. (1998) *J. Am. Chem. Soc.* **120**, 4308 – 4318
497a. Kupfer, R., Liu, S. Y., Allentoff, A. J., and Thompson, J. A. (2001) *Biochemistry* **40**, 11490 – 11501
498. Poulos, T. L., Finzel, B. C., Gunsalus, I. C., Wagner, G. C., and Kraut, J. (1985) *J. Biol. Chem.* **260**, 16122 – 16130
499. Wang, H., Dick, R., Yin, H., Licad-Coles, E., Kroetz, D. L., Szklarz, G., Harlow, G., Halpert, J. R., and Correia, M. A. (1998) *Biochemistry* **37**, 12536 – 12545
500. Custer, L., Zajc, B., Sayer, J. M., Cullinane, C., Phillips, D. R., Cheh, A. M., Jerina, D. M., Bohr, V. A., and Mazur, S. J. (1999) *Biochemistry* **38**, 569 – 581
501. Shanklin, J., Achim, C., Schmidt, H., Fox, B. G., and Münck, E. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2981 – 2986
502. Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993) *Nature (London)* **366**, 537 – 543
502a. Whittington, D. A., Sazinsky, M. H., and Lippard, S. J. (2001) *J. Am. Chem. Soc.* **123**, 1794 – 1795
503. Elango, N., Radhakrishnan, R., Froland, W. A., Wallar, B. J., Earhart, C. A., Lipscomb, J. D., and Ohlendorf, D. H. (1997) *Protein Sci.* **6**, 556 – 568
504. Shu, L., Nesheim, J. C., Kauffmann, K., Münck, E., Lipscomb, J. D., and Que, L., Jr. (1997) *Science* **275**, 515 – 518
505. Siegbahn, P. E. M., and Crabtree, R. H. (1997) *J. Am. Chem. Soc.* **119**, 3103 – 3113
505a. Valentine, A. M., LeTadic-Biadatti, M.-H., Toy, P. H., Newcomb, M., and Lippard, S. J. (1999) *J. Biol. Chem.* **274**, 10771 – 10776
505b. Chang, S.-L., Wallar, B. J., Lipscomb, J. D., and Mayo, K. H. (2001) *Biochemistry* **40**, 9539 – 9551
506. Liu, Y., Nesheim, J. C., Paulsen, K. E., Stankovich, M. T., and Lipscomb, J. D. (1997) *Biochemistry* **36**, 5223 – 5233
507. Katopodis, A. G., Wimalasena, K., Lee, J., and May, S. W. (1984) *J. Am. Chem. Soc.* **106**, 7928 – 7935
508. Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Münck, E., Steffan, R. J., McClay, K., and Fox, B. G. (1996) *Biochemistry* **35**, 9106 – 9119
509. Qian, H., Edlund, U., Powlowski, J., Shingler, V., and Sethson, I. (1997) *Biochemistry* **36**, 495 – 504
509a. Eichhorn, E., van der Ploeg, J. R., and Leisinger, T. (1999) *J. Biol. Chem.* **274**, 26639 – 26646
509b. Allen, J. R., Clark, D. D., Krum, J. G., and Ensign, S. A. (1999) *Biochemistry* **96**, 8432 – 8437
510. Mayer, B., and Hemmens, B. (1997) *Trends Biochem. Sci.* **22**, 477 – 481
511. Stamler, J. S., Singel, D. J., and Loscalzo, J. (1992) *Science* **258**, 1898 – 1902
511a. Wolthers, K. R., and Schimerlik, M. I. (2001) *Biochemistry* **40**, 4722 – 4737
511b. Ledbetter, A. P., McMillan, K., Roman, L. J., Masters, BsS., Dawson, J. H., and Sono, M. (1999) *Biochemistry* **38**, 8014 – 8021
511c. Li, H., Raman, C. S., Martásek, P., Masters, B. S. S., and Poulos, T. L. (2001) *Biochemistry* **40**, 5399 – 5406
511d. Marletta, M. A. (2001) *Trends Biochem. Sci.* **26**, 519 – 521
511e. Rusche, K. M., and Marletta, M. A. (2001) *J. Biol. Chem.* **276**, 421 – 427
511f. Crane, B. R., Arvai, A. S., Ghosh, S., Getzoff, E. D., Stuehr, D. J., and Tainer, J. A. (2000) *Biochemistry* **39**, 4608 – 4621
512. Culotta, E., and Koshland, D. E., Jr. (1992) *Science* **258**, 1862 – 1863
513. Richeson, C. E., Mulder, P., Bowry, V. W., and Ingold, K. U. (1998) *J. Am. Chem. Soc.* **120**, 7211 – 7219
514. Pfeiffer, S., Gorren, A. C. F., Schmidt, K., Werner, E. R., Hansert, B., Bohle, D. S., and Mayer, B. (1997) *J. Biol. Chem.* **272**, 3465 – 3470
515. Marla, S. S., Lee, J., and Groves, J. T. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14243 – 14248
516. Chan, N.-L., Rogers, P. H., and Arnone, A. (1998) *Biochemistry* **37**, 16459 – 16464
517. Friebe, A., Schultz, G., and Koesling, D. (1996) *EMBO J.* **15**, 6863 – 6868
518. Schelvis, J. P. M., Zhao, Y., Marletta, M. A., and Babcock, G. T. (1998) *Biochemistry* **37**, 16289 – 16297
518a. Andersen, J. F., and Montfort, W. R. (2000) *J. Biol. Chem.* **275**, 30496 – 30503
518b. Roberts, S. A., Weichsel, A., Qiu, Y., Shelnutt, J. A., Walker, F. A., and Montfort, W. R. (2001) *Biochemistry* **40**, 11327 – 11337
519. DeMaster, E. G., Quast, B. J., Redfern, B., and Nagasawa, H. T. (1995) *Biochemistry* **34**, 11494 – 11499
519a. Lai, T. S., Hausladen, A., Slaughter, T. F., Eu, J. P., Stamler, J. S., and Greenberg, C. S. (2001) *Biochemistry* **40**, 4904 – 4910
520. Gow, A. J., Buerk, D. G., and Ischiropoulos, H. (1997) *J. Biol. Chem.* **272**, 2841 – 2845
520a. Nedospasov, A., Rafikov, R., Beda, N., and Nudler, E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13543 – 13548
521. Goldstein, S., and Czapski, G. (1996) *J. Am. Chem. Soc.* **118**, 3419 – 3425
522. Caulfield, J. L., Wishnok, J. S., and Tannenbaum, S. R. (1998) *J. Biol. Chem.* **273**, 12689 – 12695
523. Singh, R. J., Hogg, N., Joseph, J., and Kalyanaraman, B. (1996) *J. Biol. Chem.* **271**, 18596 – 18603
524. Wong, P. S.-Y., Hyun, J., Fukuto, J. M., Shirota, F. N., DeMaster, E. G., Shoeman, D. W., and Nagasawa, H. T. (1998) *Biochemistry* **37**, 5362 – 5371
525. Ghosh, D. K., and Stuehr, D. J. (1995) *Biochemistry* **34**, 801 – 807
526. Nathan, C., and Xie, Q.-w. (1994) *J. Biol. Chem.* **269**, 13725 – 13728
527. Rodríguez-Crespo, I., Gerber, N. C., and Ortiz de Montellano, P. R. (1996) *J. Biol. Chem.* **271**, 11462 – 11467
528. Liu, J., García-Cardena, G., and Sessa, W. C. (1996) *Biochemistry* **35**, 13277 – 13281
529. Bredt, D. S., Hwang, P. M., Glatt, C. E., Lowenstein, C., Reed, R. R., and Snyder, S. H. (1991) *Nature (London)* **351**, 714 – 718
530. Masters, B. S. S., McMillan, K., Sheta, E. A., Nishimura, J. S., Roman, L. J., and Martasek, P. (1996) *FASEB J.* **10**, 552 – 558
531. Hurshman, A. R., and Marletta, M. A. (1995) *Biochemistry* **34**, 5627 – 5634
532. Crane, B. R., Arvai, A. S., Ghosh, D. K., Wu, C., Getzoff, E. D., Stuehr, D. J., and Tainer, J. A. (1998) *Science* **279**, 2121 – 2126
533. Poulos, T. L., Raman, C. S., and Li, H. (1998) *Structure* **6**, 255 – 258
534. Villeret, V., Huang, S., Zhang, Y., and Lipscomb, W. N. (1995) *Biochemistry* **34**, 4307 – 4315
535. González, D. H., and Andreo, C. S. (1989) *Trends Biochem. Sci.* **14**, 24 – 27
535a. Reif, A., Fröhlich, L. G., Kotsonis, P., Frey, A., Bömmel, H. M., Wink, D. A., Pfleiderer, W., and Schmidt, H. H. H. W. (1999) *J. Biol. Chem.* **274**, 24921 – 24929
535b. Heller, R., Unbehaun, A., Schellenberg, B., Mayer, B., Werner-Felmayer, G., and Werner, E. R. (2001) *J. Biol. Chem.* **276**, 40 – 47
535c. Hurshman, A. R., and Marletta, M. A. (2002) *Biochemistry* **41**, 3439 – 3456
536. Pufahl, R. A., Wishnok, J. S., and Marletta, M. A. (1995) *Biochemistry* **34**, 1930 – 1941
537. Rusche, K. M., Spiering, M. M., and Marletta, M. A. (1998) *Biochemistry* **37**, 15503 – 15512
537a. Miranda, K. M., Espey, M. G., Yamada, K., Krishna, M., Ludwick, N., Kim, S. M., Jourd'heuil, D., Grisham, M. B., Feelisch, M., Fukuto, J. M., and Wink, D. A. (2001) *J. Biol. Chem.* **276**, 1720 – 1727
538. Schmidt, H. H. H. W., Hofmann, H., Schindler, U., Shutenko, Z. S., Cunningham, D. D., and Feelisch, M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14492 – 14497
538a. Burner, U., Furtmüller, P. G., Kettle, A. J., Koppenol, W. H., and Obinger, C. (2000) *J. Biol. Chem.* **275**, 20597 – 20601
539. Imlay, J. A., and Linn, S. (1988) *Science* **240**, 1302 – 1309
540. Fridovich, I. (1997) *J. Biol. Chem.* **272**, 18515 – 18517
541. Rosen, G. M., Pou, S., Ramos, C. L., Cohen, M. S., and Britigan, B. E. (1995) *FASEB J.* **9**, 200 – 209
542. Malmstrom, B. G. (1982) *Ann. Rev. Biochem.* **51**, 21 – 59
543. Naqui, A., and Chance, B. (1986) *Ann. Rev. Biochem.* **55**, 137 – 166
544. Wood, P. M. (1987) *Trends Biochem. Sci.* **12**, 250 – 251
545. Segal, A. W., and Abo, A. (1993) *Trends Biochem. Sci.* **18**, 43 – 47
546. Chanock, S. J., Benna, J. E., Smith, R. M., and Babior, B. M. (1994) *J. Biol. Chem.* **269**, 24519 – 24522
547. Baggiolini, M., Boulay, F., Badwey, J. A., and Curnutte, J. T. (1993) *FASEB J.* **7**, 1004 – 1010
548. Han, C.-H., Freeman, J. L. R., Lee, T., Motalebi, S. A., and Lambeth, J. D. (1998) *J. Biol. Chem.* **273**, 16663 – 16668
549. Forehand, J. R., Nauseef, W. M., Curnutte, J. T., and Johnston, R. B., Jr. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3995 – 4028, McGraw-Hill, New York
550. El Benna, J., Dang, P. M.-C., Gaudry, M., Fay, M., Morel, F., Hakim, J., and Gougerot-Pocidalo, M.-A. (1997) *J. Biol. Chem.* **272**, 17204 – 17208
551. Huang, J., Hitt, N. D., and Kleinberg, M. E. (1995) *Biochemistry* **34**, 16753 – 16757
552. Cross, A. R., Rae, J., and Curnutte, J. T. (1995) *J. Biol. Chem.* **270**, 17075 – 17077
552a. Dang, P. M.-C., Cross, A. R., and Babior, B. M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3001 – 3005
553. Isogai, Y., Iizuka, T., and Shiro, Y. (1995) *J. Biol. Chem.* **270**, 7853 – 7857
554. Bouin, A.-P., Grandvaux, N., Vignais, P. V., and Fuchs, A. (1998) *J. Biol. Chem.* **273**, 30097 – 30103
554a. Di-Poi, N., Fauré, J., Grizot, S., Molnár, G., Pick, E., and Dagher, M.-C. (2001) *Biochemistry* **40**, 10014 – 10022
555. Park, J.-W., Hoyal, C. R., El Benna, J., and Babior, B. M. (1997) *J. Biol. Chem.* **272**, 11035 – 11043

References

556. Kramer, I. M., Verhoeven, A. J., van der Bend, R. L., Weening, R. S., and Roos, D. (1988) *J. Biol. Chem.* **263**, 2352 – 2357
557. Yoshida, L. S., Saruta, F., Yoshikawa, K., Tatsuzawa, O., and Tsunawaki, S. (1998) *J. Biol. Chem.* **273**, 27879 – 27886
558. Segal, A. W., Heyworth, P. G., Cockcroft, S., and Barrowman, M. M. (1985) *Nature (London)* **316**, 547 – 549
559. Cox, F. E. G. (1983) *Nature (London)* **302**, 19
559a. Shiose, A., Kuroda, J., Tsuruya, K., Hirai, M., Hirakata, H., Naito, S., Hattori, M., Sakaki, Y., and Sumimoto, H. (2001) *J. Biol. Chem.* **276**, 1417 – 1423
559b. Lee, C.-i, Miura, K., Liu, X., and Zweier, J. L. (2000) *J. Biol. Chem.* **275**, 38965 – 38972
559c. Bonini, M. G., and Augusto, O. (2001) *J. Biol. Chem.* **276**, 9749 – 9754
559d. Alvarez, B., Ferrer-Sueta, G., Freeman, B. A., and Radi, R. (1999) *J. Biol. Chem.* **274**, 842 – 848
559e. Zhang, H., Joseph, J., Feix, J., Hogg, N., and Kalyanaraman, B. (2001) *Biochemistry* **40**, 7675 – 7686
560. Hazen, S. L., Hsu, F. F., d’Avignon, A., and Heinecke, J. W. (1998) *Biochemistry* **37**, 6864 – 6873
561. Evans, T. J., Buttery, L. D. K., Carpenter, A., Springall, D. R., Polak, J. M., and Cohen, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9553 – 9558
562. Eiserich, J. P., Hristova, M., Cross, C. E., Jones, A. D., Freeman, B. A., Halliwell, B., and van der Vliet, A. (1998) *Nature (London)* **391**, 393 – 397
562a. Henle, E. S., Han, Z., Tang, N., Rai, P., Luo, Y., and Linn, S. (1999) *J. Biol. Chem.* **274**, 962 – 971
563. Halliwell, B., and Gutteridge, J. M. C. (1985) *Free Radicals in Biology and Medicine*, Clarendon Press, Oxford
564. Yamazaki, I., and Piette, L. H. (1990) *J. Biol. Chem.* **265**, 13589 – 13594
565. Wink, D. A., Nims, R. W., Saavedra, J. E., Utermahlen, W. E. J., and Ford, P. C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6604 – 6608
566. Pogozelski, W. K., McNeese, T. J., and Tullius, T. D. (1995) *J. Am. Chem. Soc.* **117**, 6428 – 6433
567. Luo, Y., Henle, E. S., and Linn, S. (1996) *J. Biol. Chem.* **271**, 21167 – 21176
568. Hlavaty, J. J., and Nowak, T. (1997) *Biochemistry* **36**, 15514 – 15525
569. Weiss, S. J., Test, S. T., Eckmann, C. M., Roos, D., and Regiani, S. (1986) *Science* **234**, 200 – 203
570. Kanofsky, J. R., Hoogland, H., Wever, R., and Weiss, S. J. (1988) *J. Biol. Chem.* **263**, 9692 – 9696
571. Khan, A. U., and Kasha, M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12365 – 12367
572. Stratton, S. P., and Liebler, D. C. (1997) *Biochemistry* **36**, 12911 – 12920
573. Mollinedo, F., Manara, F. S., and Scheider, D. L. (1986) *J. Biol. Chem.* **261**, 1077 – 1082
574. Romeo, D., Skerlavaj, B., Bolognesi, M., and Gennaro, R. (1988) *J. Biol. Chem.* **263**, 9573 – 9575
575. Remaley, A. T., Kuhns, D. B., Basford, R. E., Glew, R. H., and Kaplan, S. S. (1984) *J. Biol. Chem.* **259**, 11173 – 11175
576. Heinecke, J. W., Meier, K. E., Lorenzen, J. A., and Shapiro, B. M. (1990) *J. Biol. Chem.* **265**, 7717 – 7720
577. Shapiro, B. M. (1991) *Science* **252**, 533 – 536
578. Jabs, T., Tschöpe, M., Colling, C., Hahlbrock, K., and Scheel, D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4800 – 4805
579. Chandra, S., and Low, P. S. (1997) *J. Biol. Chem.* **272**, 28274 – 28280
580. Imlay, J. A. (1995) *J. Biol. Chem.* **270**, 19767 – 19777
581. Zhang, L., Yu, L., and Yu, C.-A. (1998) *J. Biol. Chem.* **273**, 33972 – 33976
582. McCord, J. M. (1985) *N. Engl. J. Med.* **312**, 159 – 163
583. González-Flecha, B., and Demple, B. (1995) *J. Biol. Chem.* **270**, 13681 – 13687
584. Zweier, J. L. (1988) *J. Biol. Chem.* **263**, 1353 – 1357
585. Karoui, H., Hogg, N., Fréjaville, C., Tordo, P., and Kalyanaraman, B. (1996) *J. Biol. Chem.* **271**, 6000 – 6009
586. Giulivi, C., Poderoso, J. J., and Boveris, A. (1998) *J. Biol. Chem.* **273**, 11038 – 11043
587. Jousserandot, A., Boucher, J.-L., Henry, Y., Niklaus, B., Clement, B., and Mansuy, D. (1998) *Biochemistry* **37**, 17179 – 17191
588. Berlett, B. S., and Stadtman, E. R. (1997) *J. Biol. Chem.* **272**, 20313 – 20316
588a. Rauk, A., and Armstrong, D. A. (2000) *J. Am. Chem. Soc.* **122**, 4185 – 4192
588b. Messner, K. R., and Imlay, J. A. (1999) *J. Biol. Chem.* **274**, 10119 – 10128
588c. Srinivasan, C., Liba, A., Imlay, J. A., Valentine, J. S., and Gralla, E. B. (2000) *J. Biol. Chem.* **275**, 29187 – 29192
589. Beckman, K. B., and Ames, B. N. (1997) *J. Biol. Chem.* **272**, 19633 – 19636
590. Chow, C. K., ed. (1988) *Cellular Antioxidant Defense Mechanisms*, CRC Press, Boca Raton, Florida (3 volumes)
591. Halliwell, B., and Gutteridge, J. M. C. (1986) *Trends Biochem. Sci.* **11**, 372 – 375
592. Dalton, D. A., Russell, S. A., Hanus, F. J., Pascae, G. A., and Evans, H. J. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3811 – 3815
593. Berger, T. M., Polidori, M. C., Dabbagh, A., Evans, P. J., Halliwell, B., Morrow, J. D., Roberts, L. J., II, and Frei, B. (1997) *J. Biol. Chem.* **272**, 15656 – 15660
593a. Kirsch, M., and de Groot, H. (2000) *J. Biol. Chem.* **275**, 16702 – 16708
594. Conklin, P. L., Williams, E. H., and Last, R. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9970 – 9974
594a. Kirsch, M., and de Groot, H. (2001) *FASEB J.* **15**, 1569 – 1574
595. Christen, S., Woodall, A. A., Shigenaga, M. K., Southwell-Keely, P. T., Duncan, M. W., and Ames, B. N. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3217 – 3222
596. Stocker, R., Yamamoto, Y., McDonagh, A. F., Glazer, A. N., and Ames, B. N. (1987) *Science* **235**, 1043 – 1046
597. Peden, D. B., Hohman, R., Brown, M. E., Mason, R. T., Berkebile, C., Fales, H. M., and Kaliner, M. A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 7638 – 7642
598. Reiter, R. J. (1995) *FASEB J.* **9**, 526 – 533
598a. Martín, M., Macías, M., Escames, G., León, J., and Acuña-Castroviejo, D. (2000) *FASEB J.* **14**, 1677 – 1679
599. Levine, R. L., Mosoni, L., Berlett, B. S., and Stadtman, E. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15036 – 15040
600. Schultz, J. R., Ellerby, L. M., Gralla, E. B., Valentine, J. S., and Clarke, C. F. (1996) *Biochemistry* **35**, 6595 – 6603
600a. Beyer, R. E., Segura-Aguilar, J., Di Bernardo, S., Cavazzini, M., Fato, R., Fiorentini, D., Galli, M. C., Setti, M., Landi, L., and Lenaz, G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2528 – 2532
600b. Lass, A., and Sohal, R. S. (2000) *FASEB J.* **14**, 87 – 94
600c. Suh, J. H., Shigeno, E. T., Morrow, J. D., Cox, B., Rocha, A. E., Frei, B., and Hagen, T. M. (2001) *FASEB J.* **15**, 700 – 706
601. Wink, D. A., Hanbauer, I., Krishna, M. C., DeGraff, W., Gamson, J., and Mitchell, J. B. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9813 – 9817
601a. Benaroudj, N., Lee, D. H., and Goldberg, A. L. (2001) *J. Biol. Chem.* **276**, 24261 – 24267
601b. Sun, Q.-A., Kirnarsky, L., Sherman, S., and Gladyshev, V. N. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3673 – 3678
601c. Kanzok, S. M., Fechner, A., Bauer, H., Ulschmid, J. K., Müller, H.-M., Botella-Munoz, J., Schneuwly, S., Schirmer, R. H., and Becker, K. (2001) *Science* **291**, 643 – 646
601d. Lee, S.-R., Bar-Noy, S., Kwon, J., Levine, R. L., Stadtman, T. C., and Rhee, S. G. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2521 – 2526
601e. Lowther, W. T., Brot, N., Weissbach, H., and Matthews, B. W. (2000) *Biochemistry* **39**, 13307 – 13312
601f. St. John, G., Brot, N., Ruan, J., Erdjument-Bromage, H., Tempst, P., Weissbach, H., and Nathan, C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9901 – 9906
601g. Boschi-Muller, S., Azza, S., Sanglier-Cianferani, S., Talfournier, F., Van Dorsselear, A., and Branlant, G. (2000) *J. Biol. Chem.* **275**, 35908 – 35913
601h. Pollock, V. V., and Barber, M. J. (2001) *Biochemistry* **40**, 1430 – 1440
601i. Bieger, B., and Essen, L.-O. (2001) *J. Mol. Biol.* **307**, 1 – 8
601j. Seo, M. S., Kang, S. W., Kim, K., Baines, I. C., Lee, T. H., and Rhee, S. G. (2000) *J. Biol. Chem.* **275**, 20346 – 20354
602. Deng, H.-X., Hentati, A., Tainer, J. A., Iqbal, Z., Cayabyab, A., Hung, W.-Y., Getzoff, E. D., Hu, P., Herzfeldt, B., Roos, R. P., Warner, C., Deng, G., Soriano, E., Smyth, C., Parge, H. E., Ahmed, A., Roses, A. D., Hallewell, R. A., Pericak-Vance, M. A., and Siddique, T. (1993) *Science* **261**, 1047 – 1051
602a. Yim, M. B., Kang, J.-H., Yim, H.-S., Kwak, H.-S., Chock, P. B., and Stadtman, E. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5709 – 5714
603. Goto, J. J., Gralla, E. B., Valentine, J. S., and Cabelli, D. E. (1998) *J. Biol. Chem.* **273**, 30104 – 30109
603a. Estévez, A. G., Crow, J. P., Sampson, J. B., Reiter, C., Zhuang, Y., Richardson, G. J., Tarpey, M. M., Barbeito, L., and Beckman, J. S. (1999) *Science* **286**, 2498 – 2500
603b. Goto, J. J., Zhu, H., Sanchez, R. J., Nersissian, A., Gralla, E. B., Valentine, J. S., and Cabelli, D. E. (2000) *J. Biol. Chem.* **275**, 1007 – 1014
604. Gardner, P. R., Raineri, I., Epstein, L. B., and White, C. W. (1995) *J. Biol. Chem.* **270**, 13399 – 13405
605. Gaudu, P., and Weiss, B. (1996) *Proc. Roy. Soc. (London)* **93**, 10094 – 10098
606. Hidalgo, E., Bollinger, J. M., JR, Bradley, T. M., Walsh, C. T., and Demple, B. (1995) *J. Biol. Chem.* **270**, 20908 – 20914
607. Demple, B. (1998) *Science* **279**, 1655 – 1656
608. Godon, C., Lagniel, G., Lee, J., Buhler, J.-M., Kieffer, S., Perrot, M., Boucherie, H., Toledano, M. B., and Labarre, J. (1998) *J. Biol. Chem.* **273**, 22480 – 22489
609. Zheng, M., Åslund, F., and Storz, G. (1998) *Science* **279**, 1718 – 1721
610. Ellis, H. R., and Poole, L. B. (1997) *Biochemistry* **36**, 13349 – 13356
610a. Fuangthong, M., and Helmann, J. D. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6690 – 6695
611. Duranteau, J., Chandel, N. S., Kulisz, A., Shao, Z., and Schumacker, P. T. (1998) *J. Biol. Chem.* **273**, 11619 – 11624
612. Wenger, R. H. (2002) *FASEB J.* **16**, 1151 – 1162
613. DeLong, E. F. (2002) *Nature (London)* **419**, 676 – 677
614. Sinninghe Damsté, J. S., Strous, M., Rijpstra, W. I. C., Hopmans, E. C., Geenevasen, J. A. J., van Duin, A. C. T., van Niftrik, L. A., and Jetten, M. S. M. (2002) *Nature (London)* **419**, 708 – 712

Study Questions

1. Reticulocytes (immature red blood cells) contain mitochondria that are capable of both aerobic and anaerobic oxidation of glucose. In an experiment using these cells, incubated in oxygenated Krebs–Ringer solution with 10 mM glucose, the addition of antimycin A produced the following changes in metabolite concentration after 15 min (From Ghosh, A. K. and Sloviter, H. A. (1973) *J. Biol. Chem.* **248**, 3035–3040). Interpret the observed changes in ATP, ADP, and AMP concentrations (see tabulation). Express the concentration of each component after addition of antimycin as a percentage of that before addition. Then plot the resulting figures for each compound in the sequence found in glycolysis, i.e., label the X axis as follows:

G6P
F6P
FDP
TP
etc.

Metabolite	Abbreviation	Concentrations (mmol/1 of cells) Before addition of antimycin	After addition of antimycin
Glucose 6-phosphate	G6P	460	124
Fructose 6-phosphate	F6P	150	30
Fructose 1,6-bisphosphate	FBP	8	33
Triose phosphates	TP	18	59
3-Phosphoglycerate	3PGA	45	106
2-Phosphoglycerate	2PGA	26	19
Phosphoenolpyruvate	PEP	46	34
Pyruvate	Pyr	126	315
Lactate	Lac	1125	8750
ATP		2500	1720
ADP		280	855
AMP		36	206

2. The following problem can be solved using standard reduction potentials (Table 6-8). Use $E°'$ (pH 7) values for NAD^+, enzyme-bound FAD, and fumarate of −0.32, 0.0, and −0.03 volts, respectively. Values of numerical constants are given in Table 6-1.

   a) Derive an equation relating the equilibrium constant for a reaction, K_{eq}, to differences in E_0'.

   b) Calculate the numerical values of K_{eq} for the reactions

   $$\text{Succinate} + NAD^+ \rightarrow \text{Fumarate} + NADH + H^+$$

   $$\text{Succinate} + FAD \rightarrow \text{Fumarate} + FADH_2$$

   at pH 7 and 25°C. The values should be calculated for succinate and the oxidant in the numerator.

3. Compare the catalytic cycles of the following enzymes:

   Peroxidase
   Cytochrome *c* oxidase
   Cytochrome P450

4. What chemical properties are especially important for the following compounds in the electron transport complexes of mitochondria?

   FAD or FMN
   Ubiquinone (coenzyme Q)
   Cytochrome *c*

5. Describe the operation of the F_1F_0ATP synthase of mitochondrial membranes.

6. In studies of mitochondrial function the following stoichiometric ratios have been measured.

   a) The P/O ratio: number of molecules of ATP formed for each atom of oxygen (as O_2) taken up by isolated mitochondria under specified conditions.
   b) The ratio of H^+ ions translocated across a mitochondrial inner membrane to the molecules of ATP formed.
   c) The ratio of H^+ ions pumped out of a mitochondrion to the number of molecules of ATP formed.

   Discuss the experimental difficulties in such measurements. How do uncertainties affect conclusions about the mechanism of ATP synthase? Are the ratios in (b) and (c) above necessarily equal? Explain.

7. Compare P/O ratios observed for mitochondrial respiration with the following substrates and conditions:

a) Oxidation of NADH by O_2.
b) Oxidation of succinate by O_2.
c) Dehydrogenation of ascorbate by O_2.

How would the ratio of ATP formed to the number of electrons passing from NADH through the respiratory chain differ for these three oxidants: O_2, fumarate, nitrite?

8. What is the mitochondrial glycerol phosphate shuttle? Is it utilized by plant cells? Explain.

9. What chemical reactions are included in these two important components of the nitrogen cycle (see also Fig. 24-1)?

   Nitrification
   Denitrification

10. What is the difference between a dioxygenase and a monooxygenase? What is meant by a cosubstrate for a monooxygenase?

11. The enzyme *p*-hydroxybenzoate hydroxylase utilizes a cosubstrate together with O_2 to form 3,4-dihydroxybenzoate. Indicate the mechanisms by which the bound FAD cofactor participates in the reaction.

12. What pterin-dependent hydroxylation reactions are important to the human body? Point out similarities and differences between flavin and pterin hydroxylase mechanisms.

13. Describe the basic properties of nitric oxide synthases (NOSs) and their varied functions in the body. What are the three different types of NOS? In what ways do they differ?

14. List several compounds that cause oxidative stress in cells and describe some chemical and physiological characteristics of each.

15. Propylene glycol is metabolized by several aerobic bacteria to acetoacetate, which can be catabolized as an energy source (see references 509a and 509b). The first step is conversion to an epoxide which reacts further in coenzyme M-dependent and CO_2-dependent reactions to form acetoacetate. Can you propose chemical mechanisms?

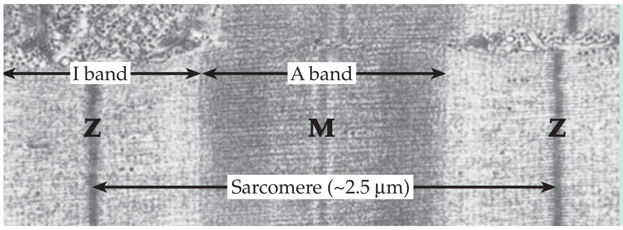
16. A group of slow-growing denitrifying bacteria obtain energy by oxidizing ammonium ions anaerobically with nitrite ions.[613,614]

$$NO_2^- + NH_4^+ \rightarrow N_2 + 2\,H_2O$$

Intermediate metabolites are hydroxylamine (H_2NOH) and hydrazine (N_2H_4). The reaction takes place within internal vesicles known as **anammoxosomes**. Unusual cyclobutane- and cyclohexane-based lipids in their membranes are thought to partially prevent the escape of the toxic intermediates from the anammoxosomes.[614]

Four protons may move from the cytoplasm into the vesicles for each ammonium ion oxidized. Can you write a reaction sequence? What is the Gibbs energy change for the reaction? How is ATP generated? See p. 1052.

Notes



Electron micrograph of a thin longitudinal section of a myofibril from pig muscle. The basic contractile unit is the **sarcomere**, which extends from one Z line to the next. Thin **actin** filaments are anchored at the M lines and the thick **myosin** filaments at the Z lines. The (anisotropic) A bands are regions of overlap of interdigitated thick and thin filaments, while the I ("isotropic") bands are devoid of thick filaments. The ATP-driven contraction of muscle involves sliding of the interdigitated filaments and shortening of the sarcomere to ~1.8 µm. Micrograph courtesy of Marvin Stromer

Contents

1089 .... **A. Motility of Bacteria**
1089 .... 1. The Structure and Properties of Bacterial Flagella
1090 .... *Quasiequivalence*
1091 .... *Growth of flagella*
1091 .... 2. Rotation of Flagella
1093 .... 3. Chemotaxis
1096 .... **B. Muscle**
1096 .... 1. The Structural Organization of Striated Muscle
1096 .... 2. Muscle Proteins and Their Structures
1096 .... *Actin and the thin filaments*
1099 .... *Titin and nebulin*
1099 .... *Proteins of the M-line and Z-disc*
1099 .... *The regulatory proteins troponin and tropomyosin*
1099 .... *Myosins*
1101 .... *The thick filaments*
1104 .... 3. Actomyosin and Muscle Contraction
1104 .... *The rowing hypothesis*
1105 .... *Why two heads?*
1105 .... *Structure of the myosin heads*
1106 .... *How does actin bind?*
1107 .... *Kinesins and other molecular motors*
1108 .... *The ATPase cycle of actomyosin and of the kinesins*
1108 .... *Observing single molecules*
1110 .... *Thinking about chemical mechanisms*
1114 .... 4. Control of Muscle Contraction
1114 .... *Calcium ions in muscle*
1117 .... *Smooth muscle*
1117 .... 5. The Phosphagens
1118 .... **C. Motion in Nonmuscle Cells**
1118 .... 1. Actin-Based Motility
1119 .... 2. Transport along Microtubules by Kinesin and Dynein
1121 .... 3. Eukaryotic Cilia and Flagella
1122 .... 4. Chemotaxis
1122 .... 5. Other Forms of Movement
1122 .... **References**
1127 .... **Study Questions**

Boxes

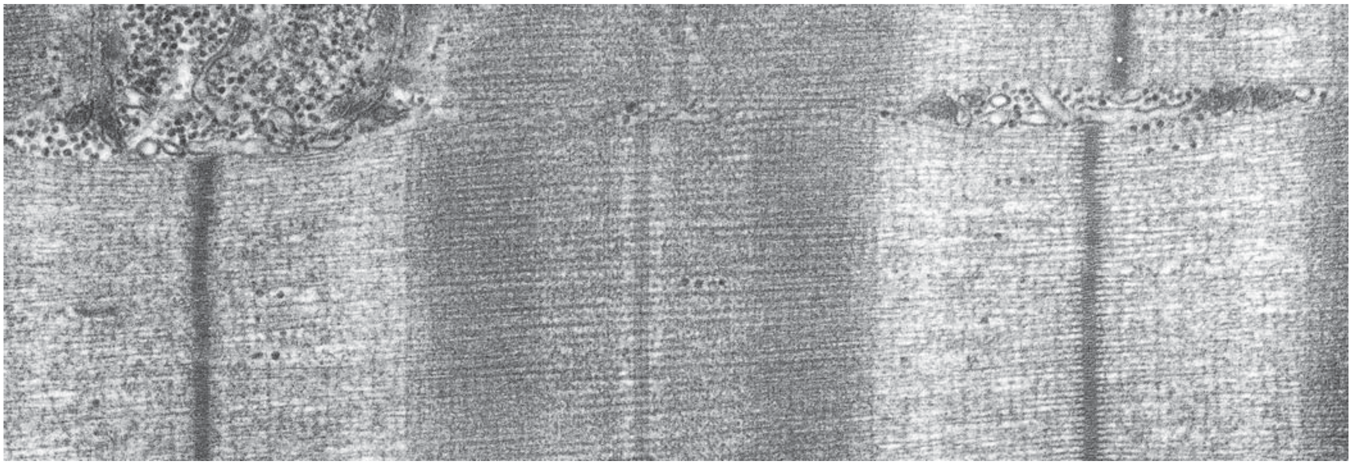
1112 .... Box 19-A Hereditary Diseases of Muscle
1115 .... Box 19-B Malignant Hyperthermia and Stress-Prone Pigs
1120 .... Box 19-C Actin-Based Motility and Bacterial Invasion

Tables

1118 .... Table 19-1 Some Actin-Binding Proteins

The Chemistry of Movement

19



The swimming of bacteria, the flowing motion of the ameba, the rapid contraction of voluntary muscles, and the slower movements of organelles and cytoplasm within cells all depend upon transduction of chemical energy into mechanical work.

A. Motility of Bacteria

The smallest organs of propulsion are the bacterial flagella (Figs. 1-1, 1-3), and we have been able to unravel some of the mystery of movement by looking at them. When a cell of *E. coli* or *Salmonella* swims smoothly, each flagellum forms a left-handed superhelix with an ~2.3 µm pitch. Rotation of these "propellors" at rates of 100–200 revolutions / s (100–200 Hz) or more[1,2] in a counterclockwise direction, as viewed from the distal end of the flagellum, drives the bacterium forward in a straight line.[3–8] Several flagella rotate side-by-side as a bundle.[4] The observed velocities of 20–60 µm / s are remarkably high in comparison with the dimensions of the bacteria. Also remarkable is the fact that a cell may travel straight for a few seconds, but then tumble aimlessly for about 0.1 s before again moving in a straight line in a different direction. The tumbling occurs when the flagellum reverses its direction of rotation and also changes from a left-handed to a right-handed superhelix, which has just half the previous pitch.

Such behavior raised many questions. What causes reversal of direction of the propellor? Why do the bacteria tumble? How does a bacterium "decide" when to tumble? How is the flagellum changed from a left-handed to a right-handed superhelix? How does this behavior help the bacterium to find food? Most intriguing of all, what kind of motor powers the flagella? The answers are complex, more than 50 genes being needed to specify the proteins required for assembly and operation of the motility system of *E. coli* or *Salmonella typhimurium*.[9]

1. The Structure and Properties of Bacterial Flagella

Twenty or more structural proteins are present from the base to the tip of a complete bacterial flagellum. However, over most of their length the long thin shafts (Figs. 1-1, 19-1) are composed of subunits of single proteins called **flagellins**. Flagellin molecules have a high content of hydrophobic amino acids and, in *Salmonella*, contain one residue of the unusual N^{ε}-methyllysine. The subunits are arranged in a helix of outside diameter ~20 nm in which they also form 11 nearly longitudinal rows or **protofilaments**.[10–12a] Each subunit gives rise to one of the projections seen in the stereoscopic view in Fig. 19-1B. The flagella usually appear under the electron microscope to be supercoiled (Fig 19-1C–E) with a long "wavelength" (pitch) of ~2.5 µm. The supercoiled structure is essential for function, and mutant bacteria with straight flagella are nonmotile. Under some conditions and with some mutant flagellins, straight flagella, of the type shown in Fig. 19-1B, are formed. There is a central hole which is surrounded by what appears to be inner and outer tubes with interconnecting "spokes." However, all of the 494-residue flagellin subunits presumably have identical conformations, and each subunit contributes to both inner and outer tubes as well as to the outer projections. **Basal bodies** (Fig. 19-2) anchor the flagella to the cell wall and plasma membrane and contain the protic motors (Fig. 19-3) that drive the flagella.[14–16]

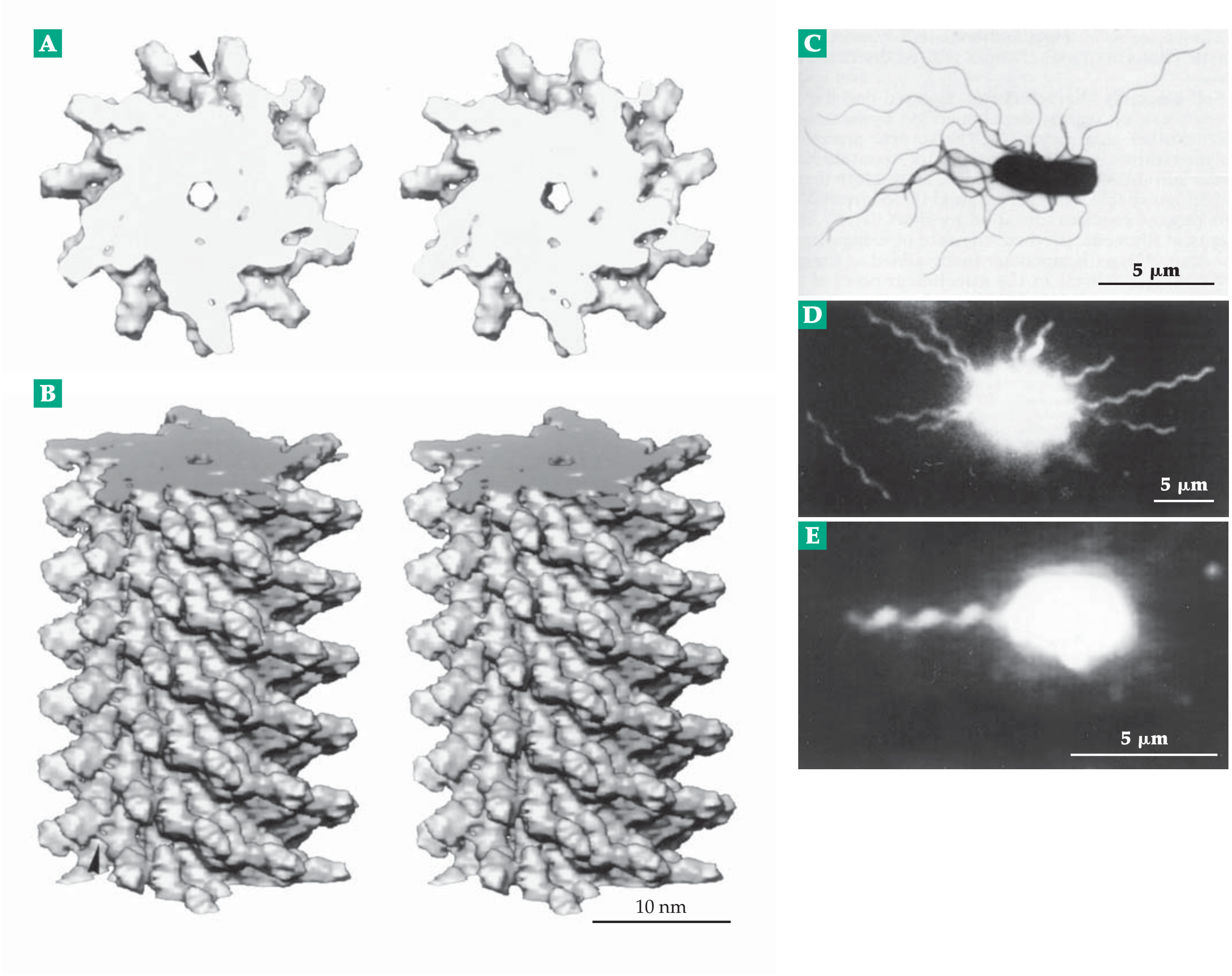


Figure 19-1 (A) Axial view of a 5-nm thick cross-section of the flagellar filament shown in (B). The 11 subunits form two turns of the one-start helix. (B) Stereoscopic oblique view of a 30-nm long section of a flagellum of *Salmonella typhimurium*. This is a straight flagellum from a nonmotile strain of bacteria. The structure was determined to a resolution of 0.9 nm by electron cryomicroscopy. From Mimori *et al.*[11] Courtesy of Keiichi Namba. (C) Electron micrograph of a cell of *S. typhimurium* showing peritrichous (all-around) distribution of flagella. Courtesy of S. Aizawa.[3] (D) Dark-field light micrograph of a flagellated cell of *S. typhimurium* with flagella dispersed during tumbling (see text). Courtesy of R. M. Macnab.[3] (E) Image of a cell of *Vibrio alginolyticus* obtained with dark-field illumination showing the single polar flagellum.[13] Because the illumination was strong, the size of the cell body and the thickness of the flagellum in the image appear large. Courtesy of Michio Homma.

Quasiequivalence. There are two distinct types of straight flagella: one (R) in which the protofilaments have a right-handed twist (as in Fig. 19-1) and the other (L) in which the protofilaments have a left-handed twist. These arise from two different conformations of the subunit proteins. Native supercoiled flagella contain a mixture of flagellins in the R- and L-states with all subunits in a given protofilament being in the same state. The supercoiling of the filament cannot be explained by stacking of identical subunits but is thought to arise because of an asymmetric distribution of protofilaments in a given state around the filament.[17–19a] Here, as with the icosahedral viruses (Chapter 7), quasiequivalence permits formation of a structure that would be impossible with full equivalence of subunits. The corkscrew shape of the flagellum is essential to the conversion of the motor's torque into a forward thrust.[18] Certain mutants of *Salmonella* have "curly" flagella with a superhelix of one-half the normal pitch. The presence of *p*-fluorophenylalanine in the growth medium also produces curly flagella, and normal flagella can be transformed to curly ones by a suitable change of pH. More important for biological function, the transformation from normal to curly also appears to take place during the tumbling of bacteria associated with chemotaxis.[17]

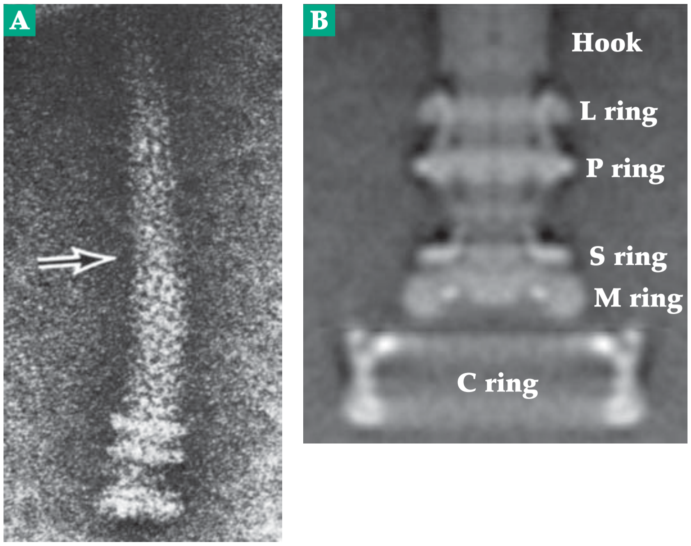


Figure 19-2 (A) Electron micrograph of a flagellum from *E. coli* stained with uranyl acetate. The M-and S-rings are seen at the end. Above them are the P-ring, thought to connect to the peptidoglycan layer, and the L-ring, thought to connect to the outer membrane or lipopolysaccharide layer (see Fig. 8-28). An arrow marks the junction between hook and thinner filament. From DePamphilis and Adler.[14] The hook is often bent to form an elbow. (B) Average of ~100 electron micrographs of frozen-hydrated preparations of basal bodies showing the cytoplasmic C-ring (see Fig. 19-3) extending from the thickened M-ring. From DeRosier.[16]

Growth of flagella. Iino added *p*-fluorophenylalanine to a suspension of bacteria, whose flagella had been broken off at various distances from the body.[20] Curly ends appeared as the flagella grew out. Unlike the growth of hairs on our bodies, the flagella grew from the outer ends. Because no free flagellin was found in the surrounding medium, it was concluded that the flagellin monomers are synthesized within the bacterium, then pass out, perhaps in a partially unfolded form, through the 2- to 3-nm diameter hole[10,12] in the flagella, and bind at the ends.[21] Flagella of *Salmonella* grow at the rate of 1 µm in 2–3 min initially, then more slowly until they attain a length of ~15 µm. More recent studies have provided details. The hook region (Fig. 19-3) grows first to a length of ~55 nm by addition to the basal-body rod of ~140 subunits of protein **FlgE**. During growth a **hook cap** formed from subunits **FlgD** prevents the FlgE subunits from passing out into the medium.[22,23] Hook subunits are added beneath the cap, moving the cap outward. Hook growth is terminated by protein **FlgK** (also called hook-associated protein Hap1). This protein displaces the hook cap and initiates growth of the main filament.[24] The first 10–20 subunits added are those of the FlgK (Fig. 19-3). These are followed by 10–20 subunits of **FlgL** (Hap3), a modified flagellin whose mechanical properties can accomodate the stress induced in the flagella by their rotation.[25] FlgJ is also needed for rod formation.[25a]

Growth of the flagellum to a length of up to 20 µm continues with subunits of **FliC** that are added at the tip, which is now covered by a dodecamer of the **cap protein FliD** (HAP2).[24,26,26a,b] Its 5-fold rotational symmetry means that this "star-cap" does not form a perfect plug for the 11-fold screw-symmetry of the flagellum, a fact that may be important in allowing new flagellin subunits to add at the growing tip. If the cap protein is missing, as in some *FliD* mutants, a large amount of flagellin leaks into the medium.[24]

Still unclear is how the protein synthesis that is taking place on the ribosomes in the bacterial cytoplasm is controlled and linked to "export machinery" at the base of the flagellin. As indicated in Fig. 19-3, the genetically identified proteins FlhA, FliH, and FliI are involved in the process that sends the correct flagellin subunits through the growing flagellum at the appropriate time. FliI contains an ATPase domain.[26c] FliS protein may be an export chaperone.[26d]

2. Rotation of Flagella

A variety of experiments showed that the flagellum is a rigid propellor that is rotated by a "motor" at the base. For example, a bacterium, artificially linked by means of antibodies to a short stub of a flagellum of another bacterium, can be rotated by the second bacterium. Rotation of cells tethered to a cover slip has also been observed. Although it is impossible to see individual flagella on live bacteria directly, bundles of flagella and even single filaments (Fig. 19-1C) can be viewed by dark-field light microscopy.[8,29] Normal flagella appear to have a left-handed helical form, but curly *Salmonella* flagella, which have a superhelix of one-half the normal pitch, form a right-handed helix.[5] Normal bacteria swim in straight lines but periodically "tumble" before swimming in a new random direction. This behavior is part of the system of **chemotaxis** by which the organism moves toward a food supply.[30] Curly mutants tumble continuously. When bacteria tumble the flagella change from normal to curly. The pitch is reversed and shortened. A proposed mechanism for the change of pitch involves propagation of cooperative conformational changes down additional

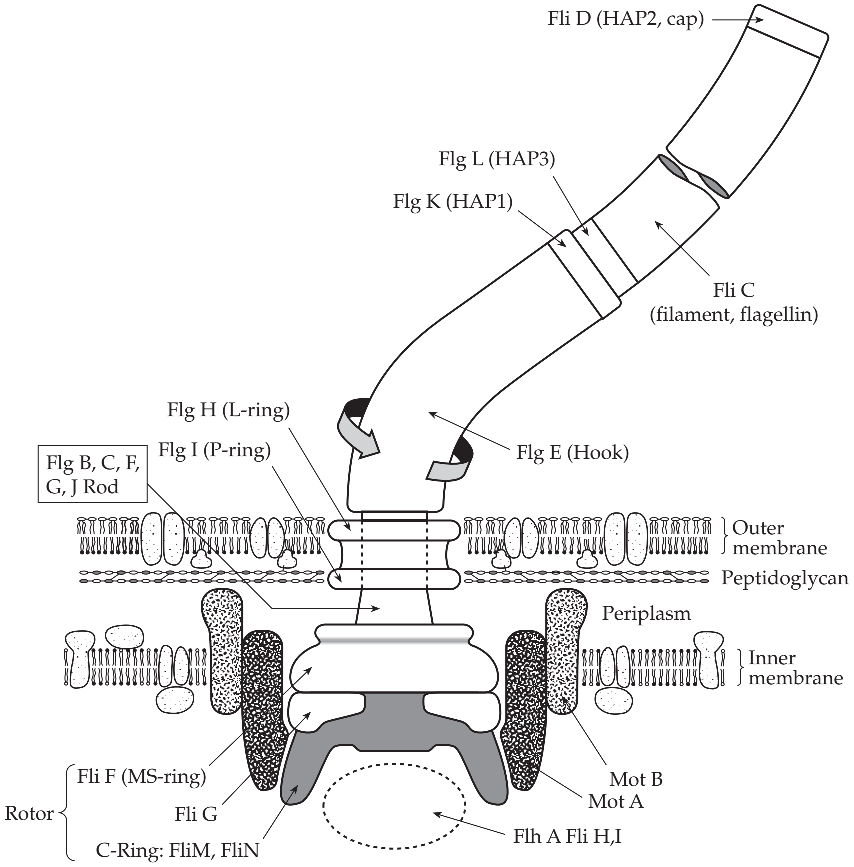


Figure 19-3 Schematic drawing of bacterial flagellar motor. Based on drawings of Berg,[27] Zhou and Blair,[28] and Elston and Oster.[1]

rows of flagellin subunits.[31]

There are no muscle-type proteins in the flagella. By incubating flagellated bacteria with penicillin and then lysing them osmotically, Eisenbach and Adler obtained cell envelopes whose flagella would rotate in a counterclockwise fashion if a suitable artificial electron donor was added.[32] This and other evidence showed that ATP is not needed. Rather, the torque developed is proportional to the **protonmotive force** and, under some circumstances, to ΔpH alone. It is the flow of protons from the external medium into the cytoplasm that drives the flagella.[8] Movement of *E. coli* cells in a capillary tube can also be powered by an external voltage.[33] In alkalophilic strains of *Bacillus* and some *Vibrio* species a sodium ion gradient will substitute.[13] Several hundred protons or Na^+ ions must pass through the motor per revolution.[8] Some estimates, based on energy balance,[29] are over 1000. However, Na^+-dependent rotation at velocities of up to 1700 Hz has been reported for the polar flagellum of *Vibrio alginolyticus*. It is difficult to understand how the bacterium could support the flow of 1000 Na^+ per revolution to drive the flagellum.[2]

What kind of protic motor can be imagined for bacterial flagella? Electron microscopy reveals that the flagellar hook is attached to a rod that passes through the cell wall and is, in turn, attached to a thin disc, the **M-ring** (or MS-ring), which is embedded in the cytoplasmic membrane both for gram-positive and gram-negative bacteria (Fig. 19-3). Two additional rings are present above the M-ring of flagella from gram-negative bacteria. The P-ring interacts with the peptidoglycan layer, and the L-ring contacts the outer membrane (lipopolysaccharide; Fig. 19-3). A logical possibility is that the M-ring, which lies within the plasma membrane, is the rotor, and a ring of surrounding protein subunits is the stator for the motor (Fig. 19-4).[34,35] Glagolev and Skulachev suggested in 1978 that attraction between $-COO^-$ and $-NH_3^+$ groups provides the force for movement.[34] Protons passing down an H^+-conducting pathway from the outer surface could convert $-NH_2$ groups to $-NH_3^+$, which would then be attracted to the $-COO^-$ groups on the stator subunits. When these two oppositely charged groups meet, a proton could be transferred from $-NH_3^+$ to $-COO^-$ destroying the electrostatic attraction. At the same time, movement of the M-ring would bring the next $-NH_2$ group to the H^+-conducting pathway from the outside. The $-COOH$ of the stator would now lose its proton through a conducting pathway to the inside of the bacterium, the proximity of the new $-NH_3^+$ assisting in this proton release. Since that time, other models based on electrostatic interactions have been advanced.[1,29,36]

Approximately 40 genes are required for assembly of the flagella, but mutations in only five motility genes have produced bacteria with intact flagella that do not rotate. Among these genes are *motA*, *mot B*, *FliG*, *FliM*, and *FliN*.[16,29,37,37a] Infection with a lambda transducing bacteriophage carrying functional *motB* genes restores motility to *motB* mutants by inducing synthesis of the motB protein. Block and Berg observed rotation of single bacteria tethered to a coverslip by their flagella. As the synthesis of the motB protein increased,the flagellar rotation rate increased in as many as 16 steps. This suggested that as many as 16 subunits of the *motB* gene product may contribute to the operation of the motor.[38] Later studies suggest eight subunits[39] rather than 16.

Both the M-ring and the thin S-ring, which lies directly above it and is now usually referred to as the MS-ring, are formed from ~20–25 subunits of the 61-kDa **FliF** protein.[39] Both the MotA and MotB proteins are embedded in the inner bacterial membrane and appear to form a circular array of "studs" around the M-ring.[16] MotA has a large cytosolic domain as well as four predicted trans-membrane helices[40] while MotB has a large periplasmic domain and probably binds to the peptidoglycan.[37,41,41a] The MotA and MotB proteins, which bind to each other, are thought to form the ~8 functional units in the stator of the motor.[37] Proteins FliG, FliM, and FliN are evidently parts of the rotor assembly. FliM and FliN form an additional ring, the cytoplasmic or **C-ring**, which had been difficult to see in early electron microscopy. As many as 40 of each of these subunits may be present in the ring.[42,43] A ring of FliG subunits joins the C-ring to the MS-ring (Fig. 19-3). FliE is also a part of the basal body.[25a]

From study of mutants it has been concluded that three charged residues of FliG, R279, D286, and D287 are directly involved in generation of torque by the motor.[44] Side chains of these residues may interact with the cytoplasmic domains of MotA and MotB. Residues R90 and E98 of MotA may be involved in controlling proton flow through the motor units.[28,44] The two prolines P173 and P122 are also essential for torque generation.[28]

There are obvious similarities between the flagellar motors and the protic turbines of ATP synthases (Fig. 18-14), but there are also substantial differences. It apparently takes about 12 protons for one revolution of the ATP synthase but about 1000, or ~125 per motor unit, for rotation of a bacterial flagellum. Elston and Oster propose an ion turbine more complex than that of ATP synthase. They suggest that the rotor might contain about 60 slanted rows of positively charged groups spaced as shown in Fig. 19-4. The motor is reversible, i.e., it can rotate in either direction. One possibility is that the subunits alter their conformations cooperatively in such a way that the slant of the rows of charged groups is reversed. Other possibilities for altering the constellation of charges via conformational changes can be imagined.[1] See also Thomas *et al.*[44a]

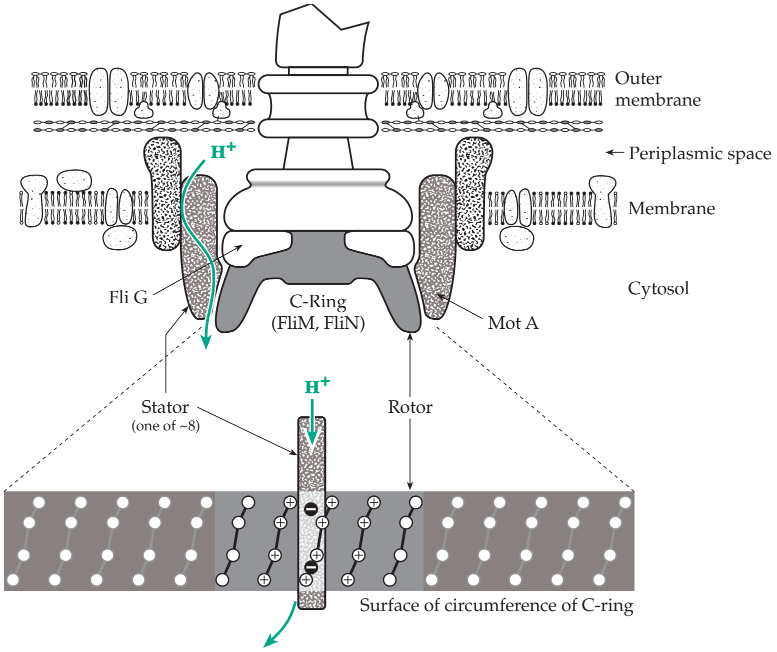


Figure 19-4 Schematic drawing of a hypothetical configuration of rotor and one stator unit in a flagellar motor as proposed by Elston and Oster.[1] The rotor can hold up to 60 positive charges provided by protons flowing from the periplasm through the stator motor units that surround the C-ring and hopping from one site to the next along the slanted lines. The rotor is composed of 15 repeating units, each able to accommodate four protons. Negative charges on the stator units are 0.5 nm from the rotor charges at their closest approach. For details see the original paper.

3. Chemotaxis

The flagellar motor is reversible, and in response to some signal from the bacterium it will turn in the opposite direction. At the same time, the flagellin subunits and those of the hook undergo conformational changes that change the superhelical twist. Perhaps synchronous conformational changes in the M-ring also are associated with the change in direction of rotation and are induced by interaction with a **switch complex** that lies below the M-ring. This consists of proteins FliG, FliM, and FliN.[44b] Mutations in any one of these proteins lead to the following four phenotypes: absence of flagella, paralyzed flagella, or flagella with the switch biased toward clockwise or toward counterclockwise rotation.[45]

What signals a change in direction of rotation? The answer lies in the attraction of bacteria to compounds that they can metabolize. Bacteria will swim toward such compounds but away from repellent substances, a response known as **chemotaxis**. Cells of *E. coli* swim toward higher concentrations of L-serine (but not of D-serine), of L-aspartate, or of D-ribose.

Phenol and Ni^{2+} ions are repellent.[46–48] By what mechanism can a minuscule prokaryotic cell sense a concentration gradient? It is known that the plasma membrane contains receptor proteins, whose response is linked to control of the flagella. Since the dimensions of a bacterium are so small, it would probably be impossible for them to sense the difference in concentration between one end and the other end of the cell. The chemotatic response apparently results from the fact that a bacterium swims for a relatively long time without tumbling when it senses that the concentration of the attractant is increasing with *time*. When it swims in the opposite direction and the concentration of attractant decreases, it tumbles sooner.[49]

Koshland[47] proposed that as the membrane receptors become increasingly occupied with the attractant molecule, the rate of formation v_f of some compound X, within the membrane or within the bacterium, is increased (Eq. 19-1). When [X] rises higher than a threshold level, tumbling is induced. At the same time, X is destroyed at a velocity of v_d.

$$\xrightarrow{v_f} [X] \xrightarrow{v_d} \qquad (19\text{-}1)$$

Subsequently, a readjustment of v_f and v_d occurs such that the concentration of X falls to its normal steady state level. X would act directly on the flagellar motor.

The receptors for L-serine[50–51a] and L-aspartate[52,53] are 60-kDa proteins encoded by genes *tsr* and *tar* in *Salmonella* or *E. coli*.[46,54] These proteins span the inner plasma membrane of the bacteria as shown in Figs. 11-8 and 19-5. The functioning of the receptor has been discussed in some detail in Chapter 11. However, there is still much that is not understood. The symmetric head, whose structure is known (Fig. 11-8), has two binding sites, but the aspartate receptor binds only one aspartate tightly. There is substantial evidence that suggests a piston-type sliding of one helix toward the cytoplasm as part of the signaling mechanism.[54a] While the flagella are distributed around the cell, the receptors appear to be clustered at the cell poles.[55]

Proteins encoded by genes *cheA*, *cheW*, *cheY* and *cheZ*, *cheB*, and *ChR* are all involved in controlling chemotaxis.[48,56] Their functions are indicated in the scheme of Fig. 19-5. All of the corresponding protein products have been isolated and purified, and the whole chemotaxis system has been reconstituted in phospholipid vesicles.[57] Gene *CheA* encodes a 73-kDa protein kinase, which binds as a dimer to the cytoplasmic domains of the related aspartate, serine, and ribose/galactose receptors with the aid of a coupling protein, cheW (Fig. 19-5). A great deal of effort has been expended in trying to understand how binding of an attractant molecule to the periplasmic domain of the receptor can affect the activity of the CheA kinase, but the explanation is unclear. There is a consensus that a small but distinct conformational alteration is transmitted through the receptor.[58–61a] An apparently α-helical region containing methylation sites (Fig. 19-5) appears to be critically involved in the signaling, responding not only to occupancy of the receptor site but also to intracellular pH and temperature and to methylation. Mutation of the buried Gly 278 found in this region to branched hydrophobic amino acids, such as Val or leucine, locks the receptor in state with a superactivated CheA kinase, while *substitution* of Gly 278 with aspartate leaves the kinase *inactive*.[61] Occupancy of the normal receptor site with ligand (aspartate, serine, etc.,) dramatically decreases the kinase activity.

The CheA protein is an autokinase which, upon activation by the receptor, becomes phosphorylated on N^{ε} of the imidazole ring of His 48. It then transfers this phospho group from His 48 to the carboxylate of Asp 57 of the 654-residue protein **CheY**, which is known as the **response regulator**.[62–65d] The unregulated flagellum rotates counterclockwise (CCW). *Phospho-CheY* (CheY-*P*, which qualifies as X in Eq. 19-1) *carries the message to the flagellar motor to turn clockwise (CW).* This is apparently accomplished through the binding of CheY-*P* to the N-terminal portion of protein FliM. This presumably induces a conformational change, which is propagated to FliG and to all of the proteins of the rotor and flagellar rod, hook, and filament.[45,65,66,66a] The flagella fly apart, and the bacterium tumbles and heads randomly in a new direction.

Tumbling occurs most often when receptors are unoccupied, and the bacteria change directions often, as if lost. However, if a receptor is occupied by an attractant, the activity of CheY is decreased and less CheY-*P* will be made. The carboxyl phosphate linkage in this compound is labile and readily hydrolyzed, a process hastened by the phosphatase **CheZ**.[67–69] Consequently, in the presence of a high enough attractant concentration the tumbling frequency is decreased, CCW flagellar rotation occurs, and the bacterium swims smoothly for a relatively long time.

There are still other important factors. Occupancy of the receptor by a ligand makes the receptor protein itself a substrate for the chemotaxis-specific methyltransferase encoded by the *cheR* gene.[62,70,71] This enzyme transfers methyl groups from *S*-adenosylmethionine to specific glutamate side chains of the receptor to form methyl esters. In the aspartate receptor there are four such glutamate residues in a large cytoplasmic domain that includes the C terminus. Two of these glutamates are initially glutamines and can undergo methylation only if they are deaminated first.[72] An esterase encoded by the *cheB* gene[72] removes the methyl ester groupings as methanol.

The action of the CheR methyltransferase is apparently unregulated, but the esterase activity of CheB is controlled by the phosphorylation state of the

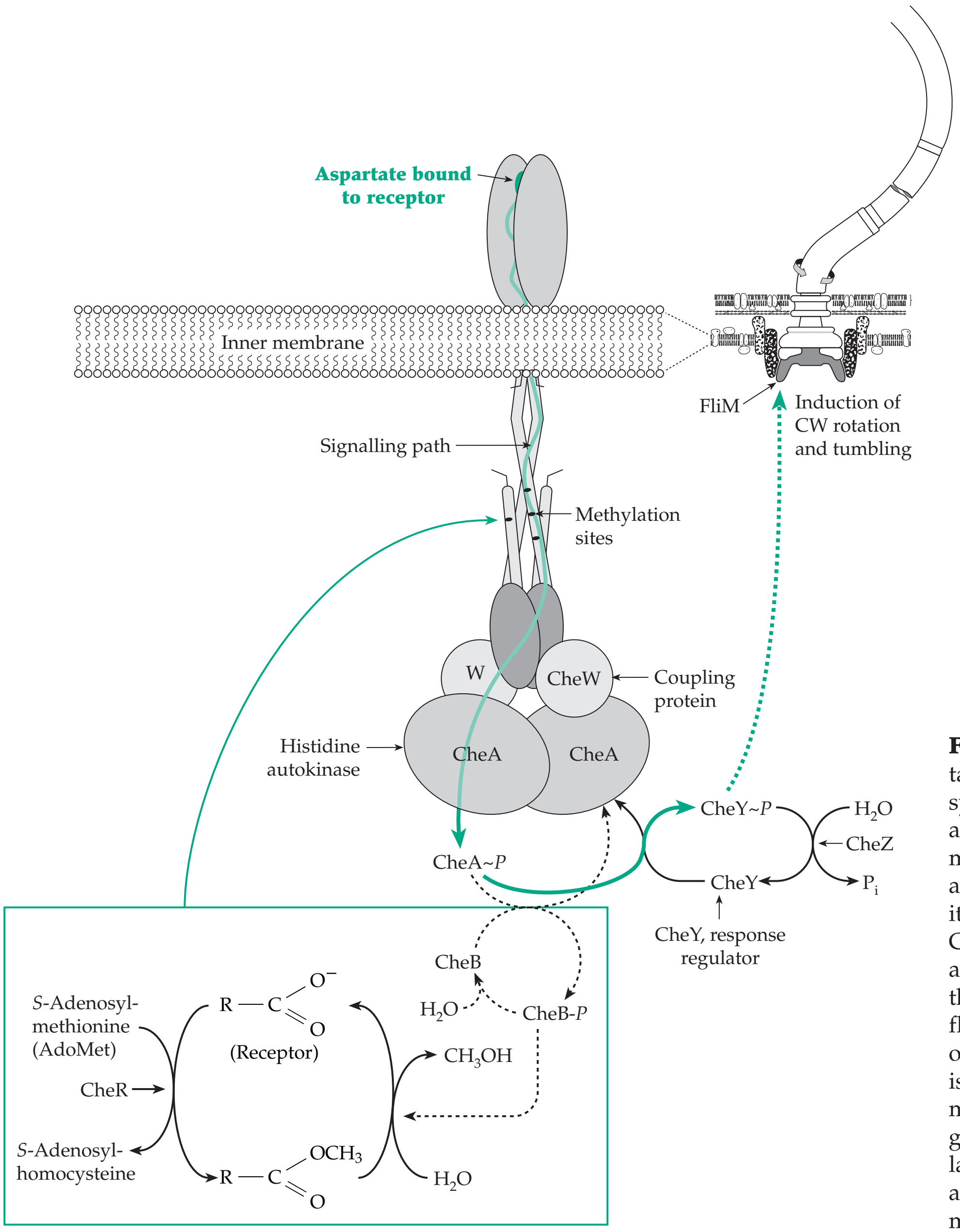

Aspartate bound to receptor
Inner membrane
FliM
Induction of CW rotation and tumbling
Signalling path
Methylation sites
W
CheW
Coupling protein
Histidine autokinase
CheA
CheA
CheY~P
H₂O
CheZ
CheY
Pᵢ
CheY, response regulator
CheA~P
CheB
H₂O
CheB-P
S-Adenosyl-methionine (AdoMet)
R — C
O⁻
O
(Receptor)
CheR
CH₃OH
S-Adenosyl-homocysteine
OCH₃
R — C
O
H₂O

Figure 19-5 Schematic representation of an important chemotactic system of *E. coli*, *S. typhimurium,* and other bacteria. The transmembrane receptor activates the autokinase CheA, which transfers its phospho group to proteins CheY and CheB to form CheY-*P* and CheB-*P*. CheY-*P* regulates the direction of rotation of the flagella, which are distributed over the bacterial surface. CheR is a methyltransferase which methylates glutamate carboxyl groups in the receptor and modulates the CheA activity. CheZ is a phosphatase and CheB-*P* a methylesterase.

autokinase CheA. CheB competes with CheY (Fig. 19-5), and CheB-*P* is the active form of the esterase. After a chemotactic stimulus the level of CheA-*P* falls and so does the activity of the methylesterase. The number of methyl groups per receptor rises making the CheA kinase more active and opposing the decrease in kinase activity caused by receptor occupancy. The system is now less sensitive to the attractant; the bacterium has *adapted* to a higher attractant concentration.[62,73,73a] It tumbles more often unless the attractant concentration rises; if it is headed toward food tumbling is still inhibited. If it is headed away from the attractant the levels of both CheY-*P* and ChB-*P* rise. A high level of fumarate within the cell also acts on the switch–motor complex and favors CW rotation.[74]

For some bacterial attractants such as D-galactose, D-ribose, maltose, and dipeptides[75] the corresponding binding proteins,[38,76] which are required for the sugar uptake (e.g., Fig. 4-18A), are also necessary for chemotaxis. The occupied binding proteins apparently react with membrane-bound receptors to trigger the chemotactic response. The aspartate receptor (*tar* gene product) appears also to be the receptor for the maltose-binding protein complex,[47] and both the aspartate and the serine receptor (*tsr* gene product) also mediate thermotaxis and pH taxis.[77,77a] Clusters of identical receptors may function cooperatively to provide high sensitivity and dynamic range.[77b]

B. Muscle

There is probably no biological phenomenon that has excited more interest among biochemists than the movement caused by the contractile fibers of muscles. Unlike the motion of bacterial flagella, the movement of muscle is directly dependent on the hydrolysis of ATP as its source of energy. Several types of muscle exist within our bodies. **Striated** (striped) **skeletal muscles** act under voluntary control. Closely related are the **involuntary striated heart muscles,** while **smooth involuntary muscles** constitute a third type. Further distinctions are made between fast-twitch and slow-twitch **fibers**. **Fast-twitch fibers** have short isometric contraction times, high maximal velocities for shortening, and high rates of ATP hydrolysis. They occur predominately in white muscle. Because of the absence of the strong oxidative metabolism found in red muscles, fast-twitch fibers fatigue rapidly. Although red muscle sometimes contains fast-twitch fibers, it more often consists of **slow-twitch fibers,** which have a longer contraction time, low shortening velocity, and low ATPase. They are more resistant to fatigue[56] than fast-twitch fibers.[78] Embryonic muscle contains fast-twitch fibers as well as embryonic forms which contract slowly.[79] Some organisms contain specialized types of muscle. For example, the asynchronous flight muscles of certain insects cause the wings to beat at rates of 100–1000 Hz. In these muscles nerve impulses are used only to start and to stop the action; otherwise the cycle of contraction and relaxation continues automatically.[80] The adductor muscles, which close the shells of oysters and clams, can sustain large tensions for long periods of time with little expenditure of energy. This is accomplished by a **catch mechanism**.[81]

1. The Structural Organization of Striated Muscle

Skeletal muscles consist of bundles of long **muscle fibers,** which are *single cells* of diameter 10–100 µm formed by the fusion of many embryonic cells. The lengths are typically 2–3 cm in mammals but may sometimes be as great as 50 cm. Each fiber contains up to 100–200 nuclei. Typical cell organelles are present but are often given special names. Thus, the plasma membrane (plasmalemma) of muscle fibers is called the **sarcolemma**. The cytoplasm is **sarcoplasm,** and mitochondria may be called **sarcosomes**. The major characteristic of muscle is the presence of the contractile **myofibrils**, organized bundles of proteins 1–2 µm in diameter and not separated by membranes from the cytoplasm. Since they occupy most of the cytoplasm, a substantial number of myofibrils are present in each muscle fiber.

In the light microscope cross striations with a repeating distance of ~2.5 µm can be seen in the myofibrils (Figs. 19-6 and 19-7). The space between two of the dense **Z-discs** (Z lines) defines the **sarcomere**, the basic contractile unit. In the center of the sarcomere is a dense **A-band** (anisotropic band). The name refers to the intense birefringence of the band when viewed with plane polarized light. Straddling the Z-discs are less dense **I-bands** (the abbreviation stands for isotropic, a misnomer, for although the bands lack birefringence, they are not isotropic). Weakly staining **M-lines** (usually visible only with an electron microscope) mark the centers of the A-bands and of the sarcomeres.

The fine structure of the sarcomere was a mystery until 1953, when H. E. Huxley, examining thin sections of skeletal muscle with the electron microscope, discovered a remarkably regular array of interdigitated protein filaments.[82,83] **Thick filaments**, 12–16 nm in diameter and ~1.6 µm long, are packed in a hexagonal array on 40- to 50-nm centers throughout the A-bands (Fig. 19-6B). Between these thick filaments are **thin filaments** only 8 nm in diameter and extending from the Z-line for a length of ~1.0 µm. When contracted muscle was examined, it was found that the I-bands had become so thin that they had nearly disappeared and that the amount of overlap between the thick and the thin filaments had increased. This indicated that contraction had consisted of the sliding movement of the thick and thin filaments with respect to each other.[84] In skeletal muscle the sarcomere shortens to a length of ~2 µm, but in insect flight muscle a much smaller shortening occurs repetitively at a very high rate.

2. Muscle Proteins and Their Structures

The myofibrillar proteins make up 50–60% of the total protein of muscle cells. Insoluble at low ionic strengths, these proteins dissolve when the ionic strength exceeds ~0.3 and can be extracted with salt solutions. Analysis of isolated mammalian myofibrils[86] shows that nine proteins account for 96% or more of the protein; **myosin**, which constitutes the bulk of the thick filaments, accounts for 43% and **actin**, the principal component of the thin filaments, 22%.

Actin and the thin filaments. There are at least six forms of actin in adult mammalian tissues: α-cardiac, α-skeletal muscle, α- and γ-smooth muscle, β- and γ-cytoplasmic.[87–89] All of them are closely homologous, e.g., the 42-kDa α-skeletal muscle actin differs in only 4 of 375 residues from the α-cardiac form and only in 6 residues from the γ-smooth form. In almost all organisms actins contain one residue of N^{δ}-methylhistidine at position 73.[87,88,90] Actin is an unusual protein in that

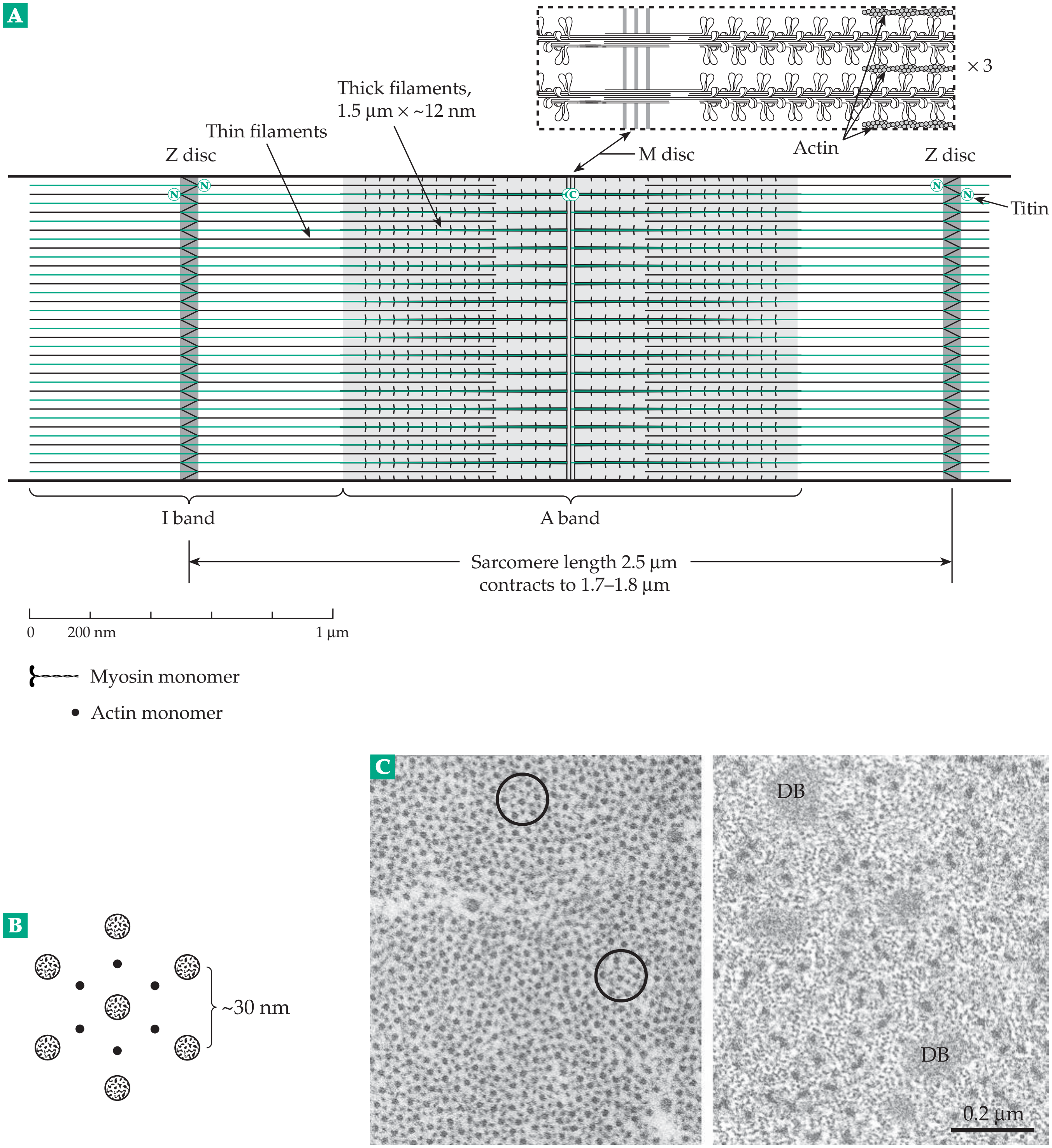

A
Thick filaments,
1.5 µm × ~12 nm
Thin filaments
Z disc
× 3
M disc
Actin
Z disc
N
N
C
N
N
Titin
I band
A band
Sarcomere length 2.5 µm
contracts to 1.7–1.8 µm
0
200 nm
1 µm
Myosin monomer
Actin monomer
B
~30 nm
C
DB
DB
0.2 µm

Figure 19-6 (A) The structure of a typical sarcomere of skeletal muscle. The longitudinal section depicted corresponds to that of the electron micrograph, Fig. 19-7A. The titin molecules in their probable positions are colored green. The heads of only a fraction of the myosin molecules are shown protruding toward the thin actin filaments with which they interact. A magnified section at the top is after Spudich.[85] It shows the interactions of the myosin heads with the thin filaments at the right-hand edge. (B) A sketch showing the arrangement of thick and thin filaments as seen in a transverse section of a striated muscle fiber. (C) Left: electron micrograph of a transverse section of a glycerated rabbit psoas muscle. The hexagonal arrangement of six thin filaments around one thick filament can be seen in the center of the circle. Six other thick filaments form a larger concentric circle as in (B). Right: transverse section of a smooth muscle fiber. Notice the irregular arrangement of thick and thin filaments. Filaments of intermediate diameter are also present, as are dense bodies (DB). The latter are characteristically present in smooth muscle.

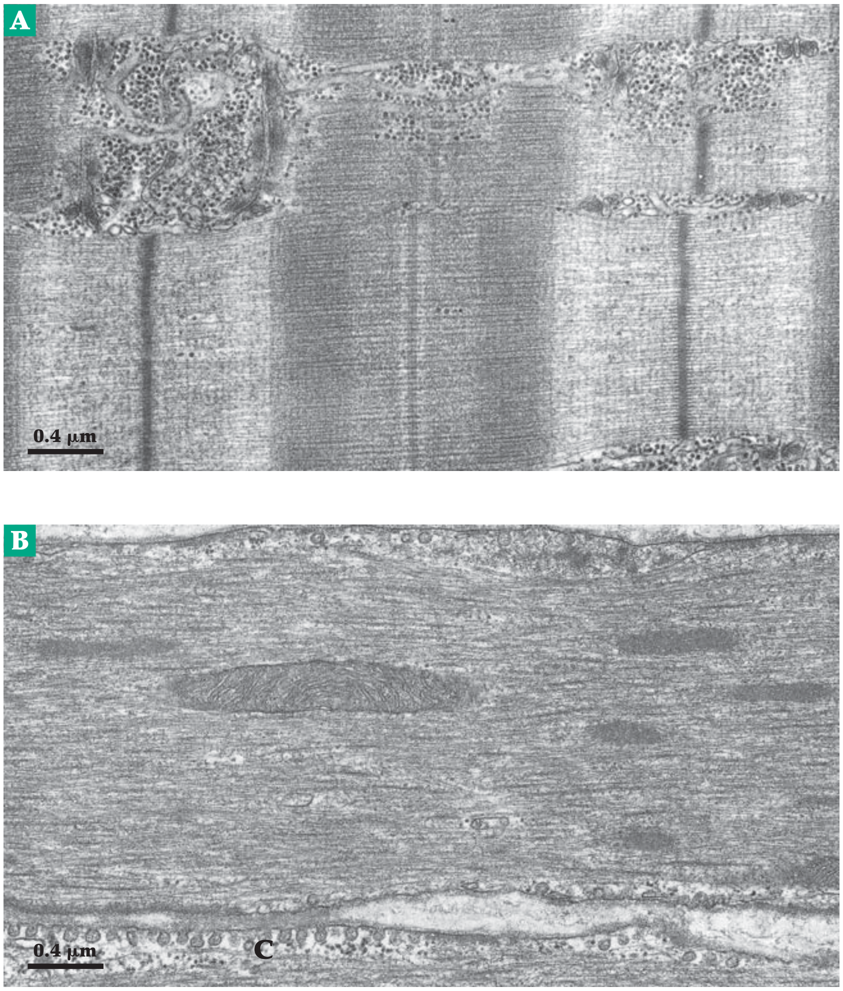


Figure 19-7 (A) Electron micrograph of a longitudinal section of a mammalian skeletal muscle (pig biceps muscle). The tissue was doubly fixed, first with formaldehyde and glutaraldehyde, then with osmium tetroxide. It was then stained with uranyl acetate and lead citrate. The section shows a white muscle fiber containing few mitochondria and narrow Z-lines. The Z-discs (marked Z), M-line, A- and I-bands, and thick and thin filaments can all be seen clearly. The periodicity of ~40 nm along the thin filaments corresponds to the length of the tropomyosin molecules, and the cross striation is thought to represent bound tropomyosin and troponin. The numerous dense particles in the upper part of the micrograph are glycogen granules, while the horizontal membranous structures are longitudinal tubules of the sarcoplasmic reticulum (endoplasmic reticulum). These come into close apposition to the T tubules leading from the surface of the muscle fiber. The T tubules (T) are visible in longitudinal section at the upper left of the micrograph on both sides of the Z-line and in cross-section in the upper right-hand corner. There a T tubule is seen lying between two lateral cisternae of the sarcoplasmic reticulum. (B) Longitudinal section of smooth muscle (chicken gizzard) fixed as in A. Thick filaments (Th), which are considerably thicker than those in striated muscle and less regular, can be seen throughout the section. They are surrounded by many thin filaments, which are often joined to dense bodies (DB). A mitochondrion (Mi) is seen in the center of the micrograph, and at the lower edge is a boundary between two adjacent cells. Notice the caveolae (C), which are present in large numbers in the plasma membrane and which are extremely active in smooth muscle. Micrographs courtesy of Marvin Stromer.

it can exist in both a filamentous and a soluble state. The interconversion between them is of great physiological importance. Actin filaments dissolve in a low ionic strength medium containing ATP to give the soluble, monomeric **G-actin**. Each G-actin monomer usually contains one molecule of bound ATP and a calcium ion.

Because of its tendency to polymerize, G-actin has been difficult to crystallize. However, it forms crystalline complexes with several other proteins, e.g., deoxyribonuclease I,[91] a fragment of gelsolin, and profilin,[92] which block polymerization and it has recently been crystallized as the free ADP complex.[92a] The three-dimensional structure of the actin is nearly the same in all cases. The molecule folds into four domains, the ATP binding site being buried in a deep cleft. The atomic structure (Fig. 7-10) resembles that of hexokinase, of glycerol kinase, and of an ATP-binding domain of a chaperonin of the Hsp 70 family.[90] As with the kinases, actin can exist in a closed and more open conformations, one of which is seen in the profilin complex. Addition of 1 mM Mg^{2+} or 0.1 M KCl to a solution of G-actin leads to spontaneous transformation into filaments of **F-actin** (Figs. 7-10 and 19-9) each containing 340–380 actin monomers and resembling the thin filaments of muscle.[93-94a] The ATP bound in the F-actin filament is hydrolyzed within ~100 s to ADP and P$_i$. However, the hydrolysis is not as rapid as polymerization so that a "cap" of ATP-containing monomers may be found at each end of the filament.[94,95,95a] There is a striking similarity to the binding of nucleotides to microtubule subunits (Fig. 7-33) and in the contractile tail sheath of bacteriophage (Box 7-C).

The two ends of the F-actin filaments have different surfaces of the monomer exposed and grow at different rates. This has been demonstrated by allowing the myosin fragment called heavy meromyosin (HMM; see Fig. 19-10) to bind to (or "decorate") an actin filament. The myosin heads bind at an angle, all pointed in one direction. This gives a "pointed" appearance to one end and a "barbed" appearance at the other. When monomeric actin is added to such an HMM-decorated F-actin filament the barbed ends grow much faster than the pointed ends.[94,96] In the intact sarcomere the ends that become pointed when decorated are free, while the opposite barbed ends of the filaments are attached at the Z-line (Fig. 19-6A). The existence in the cytoplasm of proteins that "cap" the fast-growing end of actin filaments thereby preventing further growth[96,97] suggests that cap proteins may be present at the ends of the thin filaments of the myofibrils.

Titin and nebulin. The third most abundant protein (10%), titin (also called **connectin**),[98–100a] is one of the largest of known proteins. Titin cDNA from human cardiac muscle encodes a 26,926-residue chain. Several tissue-specific isoforms of the protein are created by alternative mRNA splicing.[101] A single titin molecule stretches ~1200 nm from the Z-disc, where the N terminus is bound, to the M-line, where the C-terminal domain is attached (Fig. 19-8A). Thoughout much of the A-band titin binds to the thick filament and appears to be part of a scaffold for maintenance of the sarcomere structure. The I-band portion of titin has elastic properties that allow it to lengthen greatly or to shorten as the sarcomere changes length.[98,100,102]

Under the electron microscope titin appears as a flexible beaded string ~4 nm in diameter. Most of the molecule is made up of repetitive domains of two types. In human cardiac titin there are 132 folded domains that resemble type III fibronectin repeats and 112 immunoglobulin-like domains.[98] In a "PEVK region," between residues 163 and 2174, 70% of the residues are Pro, Glu, Val, or Lys. The titin molecule may be organized as polyproline helices in this elastic region.[102a] At the C terminus of titin 800 residues, including a Ser / Thr protein kinase domain, are found within the M-line.

Another very large protein, **nebulin** (3% of the total protein),[103] appears to be stretched alongside the thin filaments. In the electron microscope it appears as a flexible, beaded string ~4 nm in diameter. Ninety-seven percent of the 6669-residue human nebulin is organized as 185 copies of an ~ 35-residue module.[104,105] Nebulin has a proline residue at about every 35th position, possibly corresponding in length to the pitch of the actin helix (Fig. 7-10). At the C terminus is an SH3 domain (see Fig. 11-14), which is preceded by a 120-residue segment rich in potential phosphorylation sites.[106] This part of the peptide chain is anchored in the Z-discs (Fig. 19-8B, C). The three extreme N-terminal modules of nebulin bind to tropomodulin, which caps the pointed ends of thin filments.[106a] Avian cardiac muscle contains a much shorter 100-kDa protein called **nebulette**, which resembles the C-terminal parts of nebulin. Nebulin has been described as encoding a blueprint for thin filament architecture.[99,103]

Proteins of the M-line and Z-disc. The M-line region contains the structural protein **myomesin**, which binds to both titin and myosin and holds the two together.[107] Fast skeletal and cardiac fibers also contain another **M-protein**, which may bridge between myosin filaments. Both the C-terminal region of nebulin and the N termini of pairs of titin molecules meet in the Z-disc, where they are bound into a lattice containing **α-actinin**[98,108-109b] and other proteins (Fig. 19-8B). The dimeric α-actinin, a member of the spectrin family, has a subunit mass of ~97 kDa.[109a] Found primarily in the Z-discs, it is also present in nonmuscle cells in stress fibers and at other locations in the cytoskeleton (Chapter 7). It may anchor actin filaments to various structures outside of the sarcomere.[110] In the dense Z-disc of insect flight muscle a regular hexagonal lattice of α-actinin[111] and a large (500–700 kDa) modular protein called **kettin**[112,112a,b] bind the thin filaments of opposite polarity together.

The **C-protein** (thick filaments), myomesin (M-line protein), and α-actinin (Z-line protein)[110,113,114] each provide 2% of the protein in the myofibril. Less than 1% each of 11 or more other proteins may also be present within the sarcomere.[86,115] Several of these, including the cytoskeletal proteins **desmin** and **vimentin**, and **synemin** surround the Z-discs.[116,116a]

The regulatory proteins troponin and tropomyosin. These two proteins are also associated with the filaments, each one contributing ~5% to the total protein of myofibrils. Tropomyosin is an elongated α-helical coiled-coil molecule, each molecule of which associates with seven actin subunits of an actin filament. Troponin consists of three subunits known as troponins C, T, and I. The elongated troponin T binds to tropomyosin. Troponin I is an inhibitor of the interaction of myosin and actin necessary for muscle contraction. Troponin C, a member of the calmodulin family (Fig. 6-8), binds Ca^{2+} and induces conformational changes that relieve the inhibition and allow contraction to occur. Nebulin is also thought to bind to tropomyosin. A possible arrangement of one of the tropomyosin–troponin–nebulin complexes that lie along the length of the thin filaments is shown schematically and as a three-dimensional model in Figs. 19-8C and D. These proteins are discussed further in Section 4. Figure 19-9 shows a model of the thin filaments with tropomyosin coiled-coil molecules on each side. The troponin subunits are not shown.

Myosins. There are 15 distinct families of proteins within the myosin superfamily.[117–120] They vary greatly in size, but all of them bind and hydrolyze ATP, and all bind to actin. Most have C-terminal tails. At their N

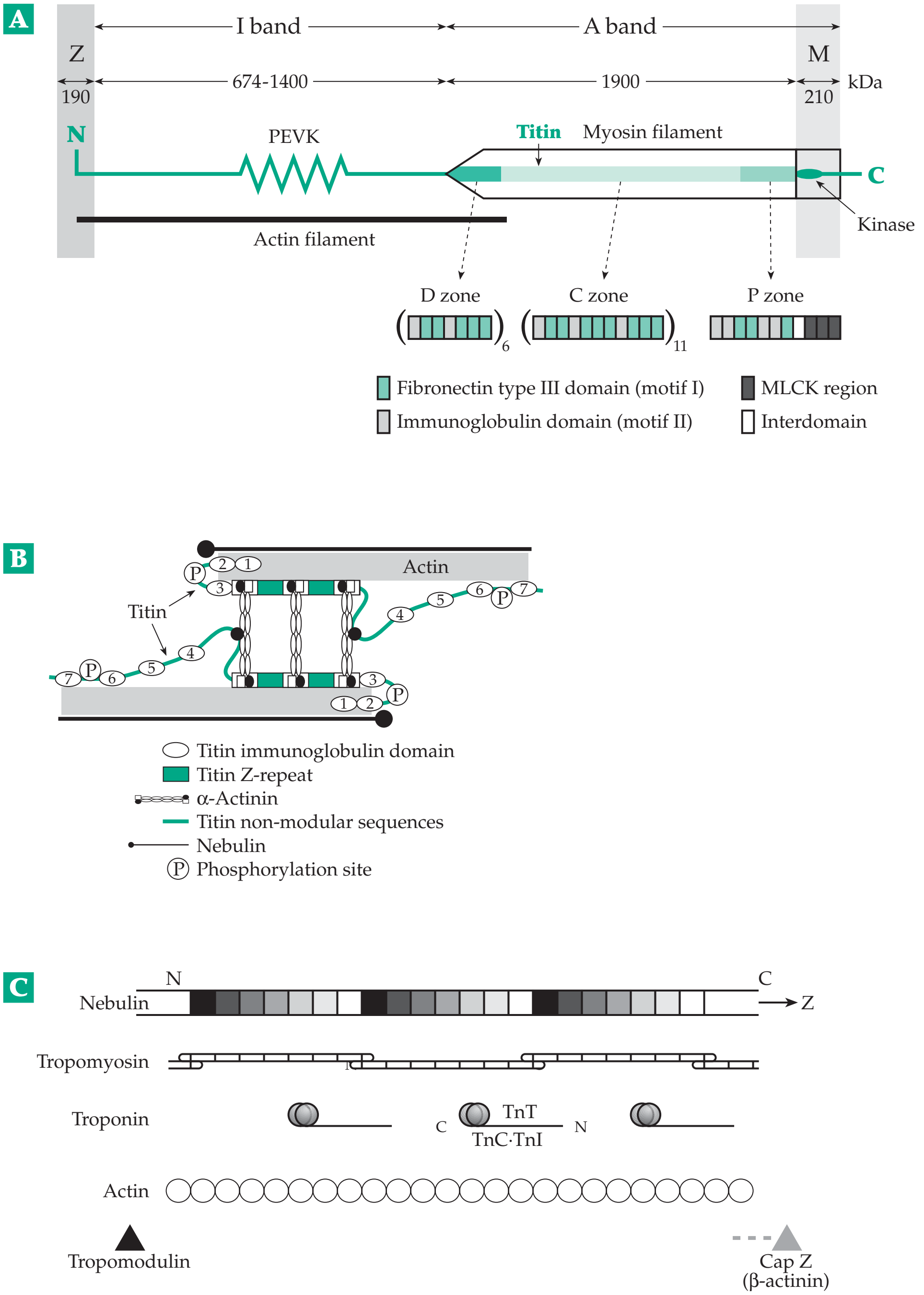

A
I band
A band
Z
M
190
674-1400
1900
210
kDa
N
PEVK
Titin
Myosin filament
C
Kinase
Actin filament
D zone
C zone
P zone
6
11
Fibronectin type III domain (motif I)
Immunoglobulin domain (motif II)
MLCK region
Interdomain
B
Actin
Titin
P
1
2
3
4
5
6
7
Titin immunoglobulin domain
Titin Z-repeat
α-Actinin
Titin non-modular sequences
Nebulin
Phosphorylation site
C
Nebulin
N
C
Z
Tropomyosin
Troponin
TnT
TnC·TnI
C
N
Actin
Tropomodulin
Cap Z
(β-actinin)

Figure 19-8 (A) Schematic drawing showing one molecule of titin (connectin) in a half sarcomere and its relationship to thick myosin filaments and thin actin filaments. The complex repeat patterns of fibronectin type III, immunoglobulin, in the three zones D, C, and P are also indicated.[98] See Maruyama.[98,98a] (B) Schematic drawing of the molecular structure of the sarcomere Z-disc. Titin, which is thought to parallel the thin filaments through the I-band, consists of various modules that are numbered from the N termini. In the Z-disc titin binds to α-actin, shown here as three vertical rods, and also to actin or actin-binding proteins. The SH3 domain (shown as a sphere) of nebulin and the N terminus of titin may interact. Regulatory phosphorylation sites are marked P. From Young *et al.*[108] Courtesy of Mathias Gautel. (C) Hypothetical model of a composite regulatory complex containing nebulin, tropomyosin, and troponin on the thin filaments of a skeletal muscle sarcomere. Each seven-module nebulin super-repeat (squares with graded shading) binds one tropomyosin, possible through the seven charge clusters along the length of each tropomyosin, and one troponin complex (shaded spheres with a tail). This complex consists of TnT, TnI, and TnC in orientations indicated by the N and C termini. Each nebulin super-repeat binds to seven actin monomers (open circles) along the thin filament. **Tropomodulin** caps the pointed ends of actin filaments and Cap Z, the “barbed ends.” From Wang *et al.*[103]

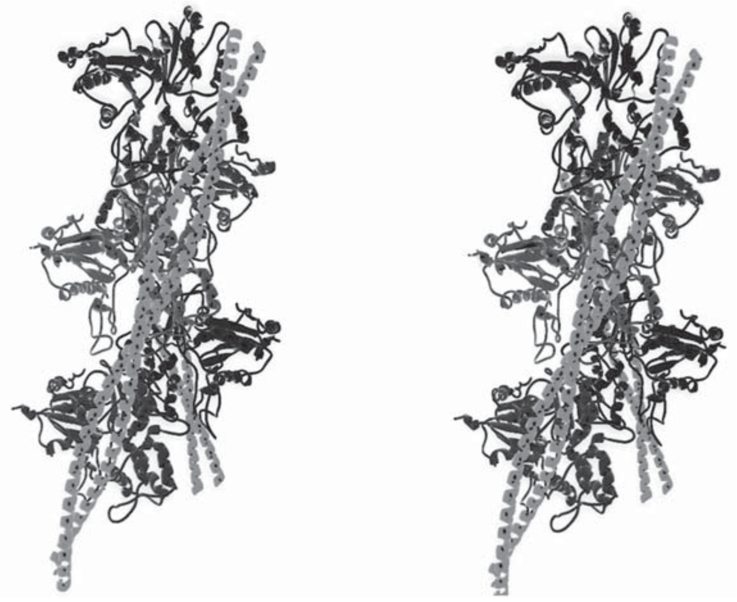


Figure 19-9 Stereoscopic ribbon drawing of the proposed structure of a thin actin filament with tropomyosin coiled-coils bound on opposing sides.[124] Five actin nomomers are assembled in the structure as is also illustrated in Fig. 7-10. From Lorenz *et al*.[125] Courtesy of Michael Lorenz.

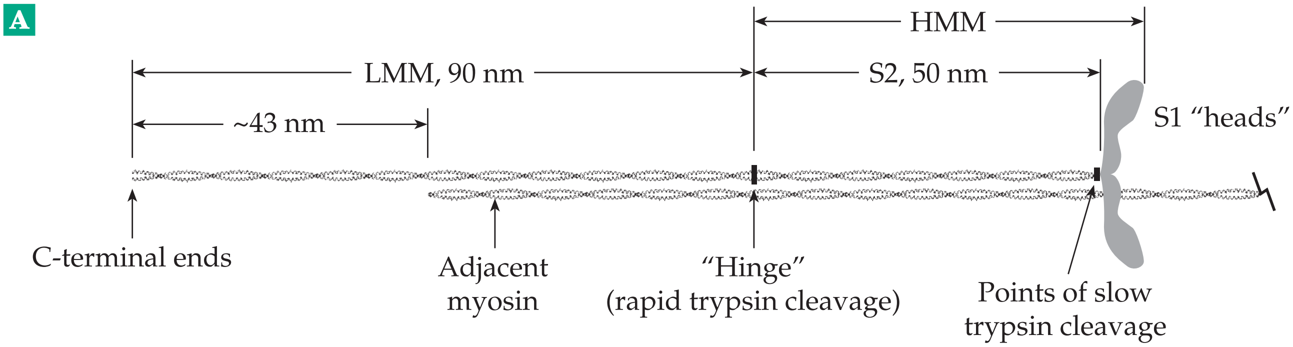
termini are one or two globular heads, which contain the catalytic centers in which ATP hydrolysis occurs. Sizes vary from 93 kDa for a myosin with a very short tail from *Toxoplasma*[118] to over 300 kDa. Myosins I, found in ameboid organisms and also in our own bodies (for example in the microvilli of the brush border of intestinal epithelial cells), are small single-headed molecules.[117,121] Myosins II are the "conventional" myosins of myofibrils and are often referred to simply as myosin. However, each of the three muscle types (skeletal, cardiac, and smooth) has its own kind of myosin II.[121a–c] Likewise, at least six different genes have been identified for the light chains of the myosin heads.[122] Fast and slow muscle as well as embryonic muscle have their own light chains. Each myosin II molecule consists of two identical ~230-kDa **heavy chains,** which are largely α-helical, together with two pairs of smaller 16- to 21-kDa **light chains**. Human skeletal muscle heavy chains contain 1938 residues of which the first ~850 are folded into pear-shaped heads, which contain the catalytic sites involved in harnessing ATP cleavage to movement. Following proline 850 nearly all of the remaining 1088 residues form an α-helical coiled-coil rod of dimensions ~160 x 2 nm (Fig. 19-10) in which the two chains coil around each other. The two heavy chains are parallel, each having its N terminus in one of the two heads and its C terminus bound in the shaft of the thick filament.

Myosins II from other sources have similar structures. For example, analysis of the DNA sequence for a heavy chain gene from the nematode *Caenorhabditis* showed that the protein contains 1966 residues, 1095 of which contain an amino acid sequence appropriate for a 160-nm long coiled coil.[123] There are no prolines within this sequence, which lies between Pro 850 and Pro 1944. Although there are many bands containing positively and negatively charged side chains along the myosin rod, the interactions between the two coiled helices are largely nonpolar. In *Drosophila* 15 different heavy-chain isoforms are created by splicing of a single mRNA.[123a]

While the C-terminal portions of the two parallel myosin heavy chains form a rod, the N-terminal portions fold into two separate heads. Each head also contains two smaller 16- to 21-kDa peptide chains which belong to the calmodulin family. One of these, the **essential light chain**, is tightly bound to the heavy chain. The second, the **regulatory light chain**, is able to bind Ca^{2+} and is less tightly bound to the rest of the head. A short treatment with trypsin or papain cuts the myosin molecule into two pieces. The tail end gives rise to **light meromyosin (LMM)**, a molecule ~90 nm in length. The remainder of the molecule including the heads is designated **heavy meromysin (HMM)**. A longer trypsin treatment leads to cleavage of HMM into one ~62-kDa **S2** fragment 40 nm long, and two ~130-kDa **S1** fragments, each of the latter representing one of the two heads (Fig. 19-10).

The junction of the head and tail portions of myosin appears rigid in Fig. 19-10A. However, there must be considerable conformational flexibility and perhaps some uncoiling of the helices to allow the two heads to interact with a single thin filament as is observed by electron microscopy.[126,128] There also appears to be a hinge between the S2 and LMM segments (Figs. 19-10A and 19-14).

The thick filaments. Dissociated myosin molecules can be induced to aggregate into rods similar to the thick filaments of muscle.[129] Since the filaments


A
HMM
LMM, 90 nm
S2, 50 nm
~43 nm
S1 "heads"
C-terminal ends
Adjacent myosin
"Hinge" (rapid trypsin cleavage)
Points of slow trypsin cleavage

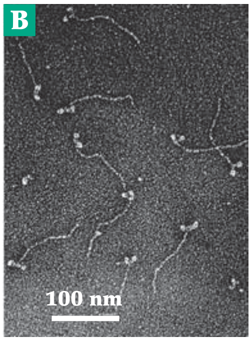


Figure 19-10 (A) An approximate scale drawing of the myosin molecule. The "hinge" is a region that is rapidly attacked by trypsin to yield the light and heavy meromyosins (LMM and HMM). Total length ~160 nm, molecular mass, 470 kDa; two ~200-kDa heavy chains; two pairs of 16- to 21-kDa light chains; heads: ~15 x 4 x 3 nm. (B) Electron micrograph of rabbit myosin monomers that became dissociated from thick filaments in the presence of ATP, fixed and shadowed with platinum.[127] Courtesy of Tsuyoshi Katoh.

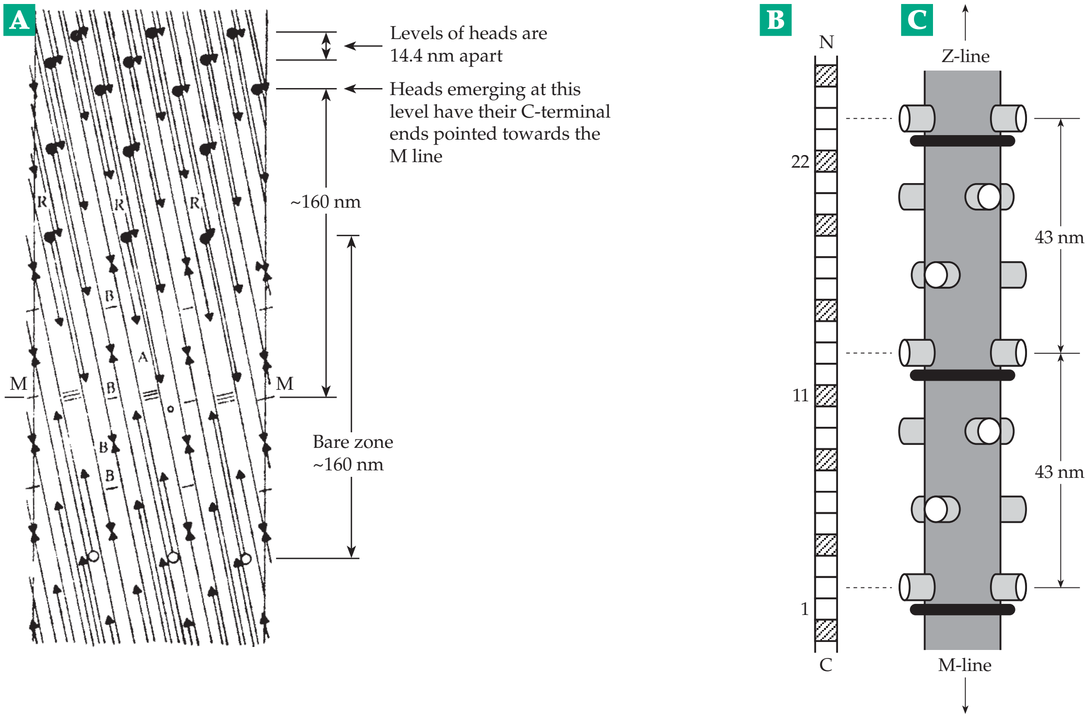

A
Levels of heads are 14.4 nm apart
Heads emerging at this level have their C-terminal ends pointed towards the M line
~160 nm
Bare zone ~160 nm
R
R
R
B
A
B
B
B
M
M
B
N
22
11
1
C
C
Z-line
43 nm
43 nm
M-line

Figure 19-11 (A) Radial projection illustrating packing of myosin rods as suggested by Squire[130] for thick filaments of vertebrate skeletal muscle. The region of the bare zone at the M-line is shown. The filled circles represent the head ends of the myosin molecules and the arrowheads represent the other end of the rod, i.e., the end of the LMM portion. Antiparallel molecules interacting with overlaps of 43 and 130 nm are shown joined by single and triple cross-lines, respectively. Positions where two arrowheads meet are positions of end-to-end butting. O is an "up" molecule (thin lines) and A a "down" molecule (thick lines). The molecules move from the core at the C-terminal end to the filament surface at the head end. The levels marked B may be the levels of attachment of M-bridge material to the myosin filament. The level M-M is the center of the M-line and of the whole filament. The lateral scale is exaggerated more than threefold. (B) A segment of titin showing the 43-nm 11-domain super-repeat. (C) Model of a segment of a thick filament showing the 43-nm repeat, the C-protein, also bound at 43 nm intervals.[99] (B) and (C) Courtesy of John Trinick.

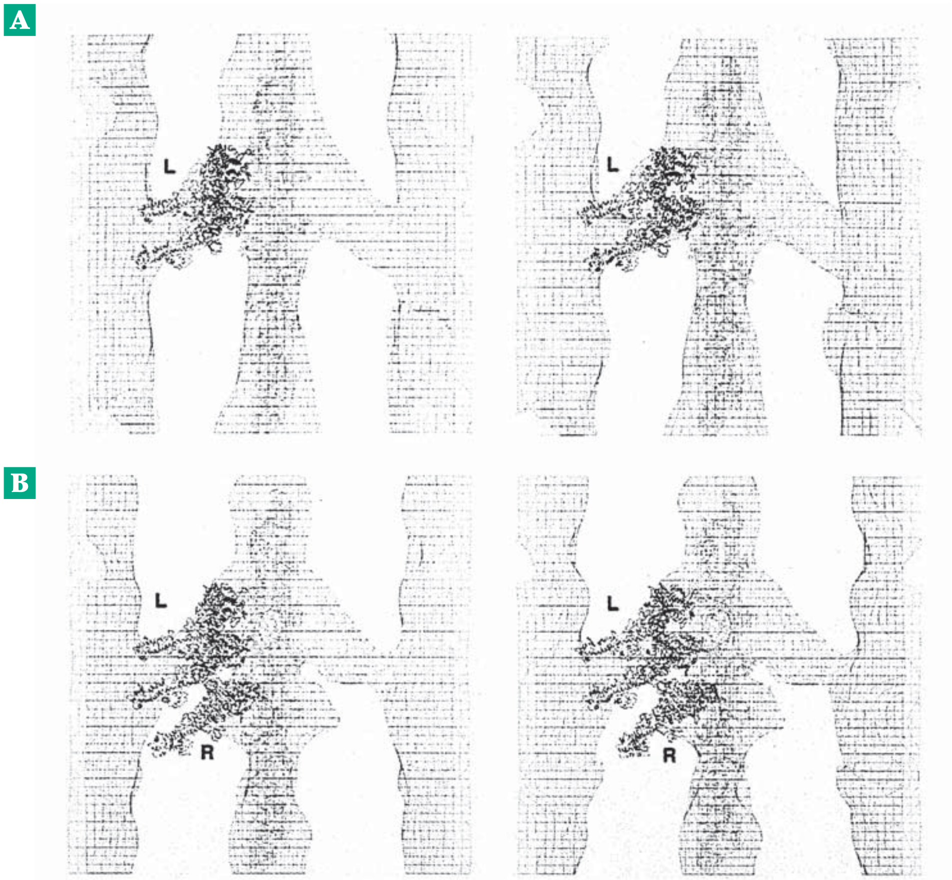


Figure 19-12 (A) Stereoscopic views of computer-assisted reconstructions of images of myosin heads attached to an F-actin filament centered between two thick filaments. Atomic structures of actin (Fig. 7-10) and of myosin heads (Fig. 19-15) have been built into the reconstructed images obtained by electron microscopy. (A) With the nonhydrolyzable ATP analog ATPPNP bound in the active sites. (B) Rigor. Two myosin heads are apparently bound to a single actin filament in (A). If they belong to the same myosin molecule the two C-terminal ends must be pulled together from the location shown here. In (B) a third head is attached, presumably from another myosin rod. This configuration is often seen in rigor. From Winkler *et al.*[134] Courtesy of K. A. Taylor.

have a diameter of ~14 to 20 nm, a large number of the thin 2-nm myosin molecules must be packed together. Electron microscopy reveals the presence of the heads projecting from the thick filaments at intervals of ~43.5 nm. However, there is a bare zone centered on the M-line, a fact that suggests tail-to-tail aggregation of the myosin monomers at the M-line in the centers of the thick filaments (see magnified section of Fig. 19-6, A). A helical packing arrangement involving about 300 myosin molecules (up to 30 rods in a single cross section) in close packing with a small central open core has been proposed for skeletal muscle myosin[130,131] and is illustrated in Fig. 19-11A,B. There are approximately three heads per turn of the helix, each group of three heads spaced 14.3 nm from the preceding one along the thick filament. It is apparently the zones of positive and negative charge, which are especially prevalent in the LMM segment toward the C termini, that lock the successive myosin molecules into this 14.3-nm spacing.[116,132] Titin also binds to the LMM segment of the myosin rod,[99,133] and its 11-domain super repeat of IgG-like and fibronectin-like modules are also 43 nm in length.[98,101] There are typically 47–49 of these super repeats in titin, and if each fits to a turn of the helix, as shown in Fig. 19-11B, there would be 147 myosin molecules in one-half of the thick filament.

Not all muscles have the thick filament structure of Fig. 19-11. In the tarantula muscle, which has a particularly well-ordered structure, there are four myosins per turn.[135,135a] Figure 19-13 shows a reconstruction of scallop myosin which has a 7-fold rotational symmetry. The thick filaments often contain

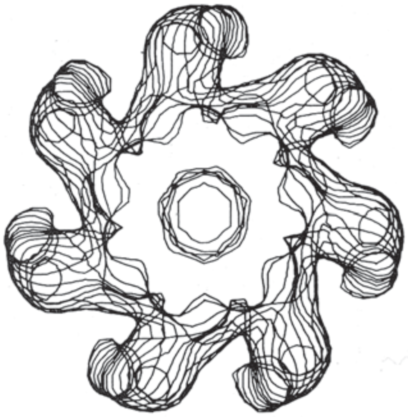


Figure 19-13 Superimposed sections for the 14 nm thickness of a computer-assisted reconstruction of the myosin filaments of the scallop adductor muscle. From Vibert and Craig.[137]

other proteins in addition to myosin. Thus, skeletal muscle contains the C-protein in a series of helical bands along the thick rod.[135b,c] In nematodes, molluscs, and insects the thick filament has a cylindrical core of **paramyosin**, another protein with a structure resembling that of the myosin rod. A minor component of *Drosophila* myosin, the **myosin rod protein**, lacks heads but is transcribed from the myosin heavy chain gene.[136]

3. Actomyosin and Muscle Contraction

That actin and myosin are jointly responsible for contraction was demonstrated long before the fine structure of the myofibril became known. In about 1929, ATP was recognized as the energy source for muscle contraction, but it was not until 10 years later that Engelhardt and Ljubimowa showed that isolated myosin preparations catalyzed the hydrolysis of ATP.[138] Szent-Györgi[139,140] showed that a combination of the two proteins actin (discovered by F. Straub[141]) and myosin was required for Mg^{2+}-stimulated ATP hydrolysis (ATPase activity). He called this combination **actomyosin**.

Under the electron microscope the myosin heads can sometimes be seen to be attached to the nearby thin actin filaments as **crossbridges**. When skeletal muscle is relaxed (not activated by a nerve impulse), the crossbridges are not attached, and the muscle can be stretched readily. The thin filaments are free to move past the thick filaments, and the muscle has some of the properties of a weak rubber band. However, when the muscle is activated and under tension, the crossbridges form more frequently. When ATP is exhausted (e.g., after death) muscle enters the state of **rigor** in which the crossbridges can be seen by electron microscopy to be almost all attached to thin filaments, accounting for the complete immobility of muscle in rigor (Figs. 19-12, 14).[134]

In rigor the crossbridges are almost all firmly attached to the thin actin filaments, making an approximately 45° angle to the actin filaments.[142–144] However, the addition of ATP causes their instantaneous release and the relaxation of the muscle fiber. In contrast, activation by a nerve impulse, with associated release of calcium ions (Section B,4), causes the thin filaments to slide between the thick filaments with shortening of the muscle. An activated muscle shortens if a low tension is applied to the muscle, but at a higher tension it maintains a constant length. Because the maximum tension developed is proportional to the length of overlap between the thick and thin filaments, it was natural to identify the individual crossbridges as the active centers for generation of the force needed for contraction.

The rowing hypothesis. H. E. Huxley[145,146] and A. F. Huxley and R. M. Simmons[147] independently proposed that during contraction the myosin heads attach themslves to the thin actin filaments. The hydrolysis of ATP is then coupled to the generation of a tension that causes the thick and thin filaments to be pulled past each other. The heads then release themselves and become attached at new locations on the actin filament. Repetition of this process leads to the sliding motion of the filaments (Fig. 19-14). The evidence in favor of this "rowing" or "swinging bridge" hypothesis was initially based largely on electron microscopy. For example, contracting muscle was frozen rapidly and fixed for microscopy in the frozen state.[148] Relaxed muscle shows no attached crossbridges, but contracting muscle has many. However,

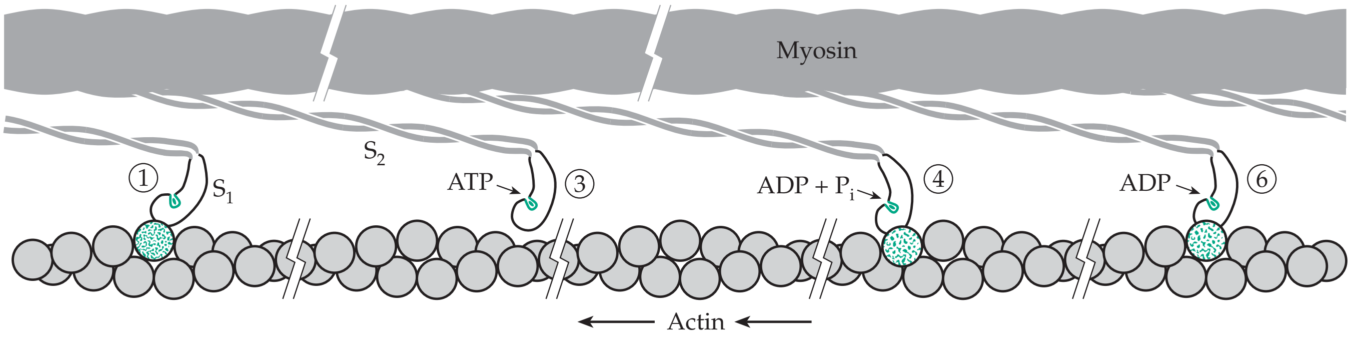


Figure 19-14 A model for the coupling of ATP hydrolysis to force production in muscle based on proposals of H. E. Huxley, and A. F. Huxley and Simmons. The power stroke is depicted here as a rotation of the crossbridge from a 90° to a 45° configuration. Four representative stages are shown: (1) the rigor complex, (3) the dissociated myosin ATP complex, (4) the actomyosin ADP pre-power stroke state in which the actin–myosin band has reformed but with a different actin subunit, which may be distant from that in (1), and (6) the actomyosin ADP post-power stroke state. Force production and contraction result from crossbridges passing cyclically through the steps depicted from left to right. Numbering of the stages corresponds approximately to that in Fig. 19-18. After H. Huxley.[146]

their appearance was distinct from that seen in rigor. The model was also supported by indirect physical methods.

An impressive demonstration that myosin heads do move along the actin filaments was provided by Sheetz and Spudich, who found that myosin-coated fluorescent beads ~0.7 µm in diameter will move along actin filaments from cells of the alga *Nitella* in an ATP-dependent fashion at velocities similar to those required in muscle.[149] The myosin heads literally glide along the thick cables of parallel actin filaments present in these algae.

Why two heads? The actin filament is a two-start helix, and it is natural to ask whether the two myosin heads bind to just one or simultaneously to both of the actin strands. Most evidence supports a 1:1 interaction of a single head with just one strand of actin. However, the other actin strand may associate with heads from a different thick filament. Another question concerns the role of the pairs of myosin heads. Could the two heads bind sequentially to the actin and exert their pull in a fixed sequence? In the reconstruction of the actomyosin complex in rigor (Fig. 19-12B) two different images are seen for the crossbridges. This suggests the existence of two different conformations for the attached myosin heads. Similar images for smooth muscle heavy meromyosin in its inactive (resting) dephosphorylated state (see p. 1116) show the two heads in very different orientations with one binding to the other of the pair and blocking its movement.[121b] Perhaps one head is tightly bound at the end of the power stroke while the other is at a different stage of the catalytic cycle. Nevertheless, single-headed myosin from *Acanthamoeba* will propell organelles along actin filaments,[150] and actin filaments will slide across a surface coated with single-headed myosin formed by controlled proteolysis.[151] The additional interactions seen in rigor may be peculiar to that state.

Structure of the myosin heads. Myosin and myosin fragments can be isolated in large quantities, but they have been difficult to crystallize. However, Rayment and coworkers purified S1 heads cleaved from chicken myosin by papain and subjected them to reductive methylation (using a dimethylamine–borane complex; see also Eq. 3-34). With most of the lysine side chain amino groups converted to dimethylamine groups, high-quality crystals were obtained, and a structure was determined by X-ray diffraction.[152] Since that time various forms of both modified and unmodified myosin heads from several species have been studied by X-ray crystallography.[153–160] Especially clear results were obtained with unmodified myosin from the ameba *Dictyostelium discoideum*. The head structure, shown in Fig. 19-11, includes a 95-kDa piece of the heavy chain and both light chains. A clearer picture of the neck region containing the light chains was provided by the structure of the "regulatory domain" of scallop myosin.[161] Unlike mammalian or avian myosins, molluscan myosins are regulated by binding of Ca^{2+} to a site in the essential light chain, but the structures are similar to those in Figs. 19-10 and 19-15.

Cleavage of the ~850-residue S1 heads with trypsin yields mainly three large fragments that correspond to structural domains of the intact protein as shown in Fig. 19-15. They are known as the 25-kDa (N-terminal), 50-kDa, and 20-kDa fragments, and for myosin from *D. discoideum* correspond to residues 1 to 204, 216 to 626, and 647 to 843, respectively. The ATP-binding site is in a deep cleft between the 20-kDa and 50-kDa

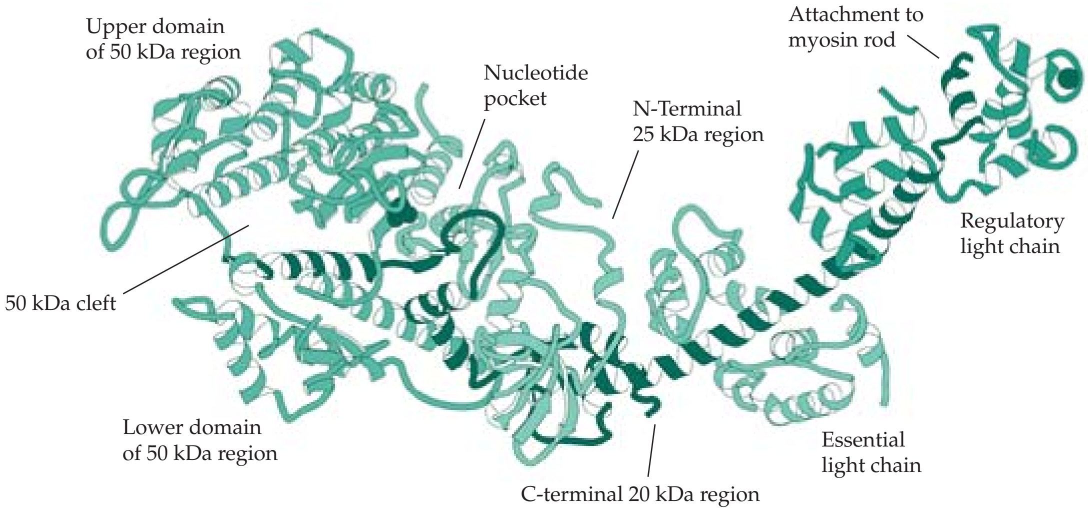
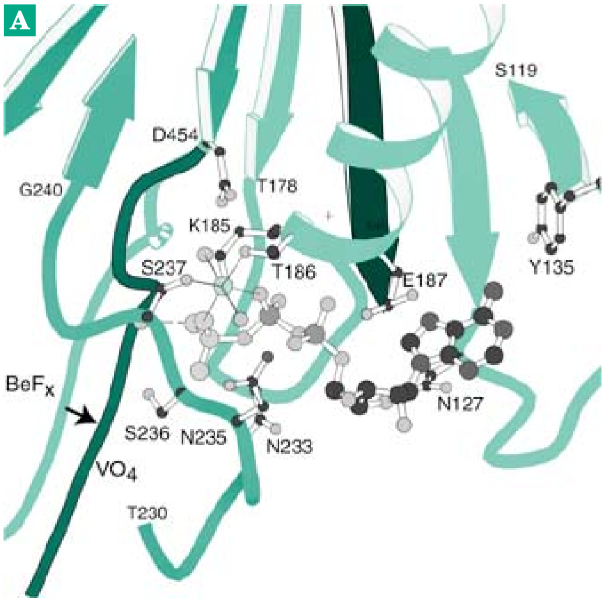


Figure 19-15 Ribbon representation of chicken skeletal myosin subfragment-1 showing the major domains and tryptic fragments. Prepared with the program MolScript. From Rayment.[157]



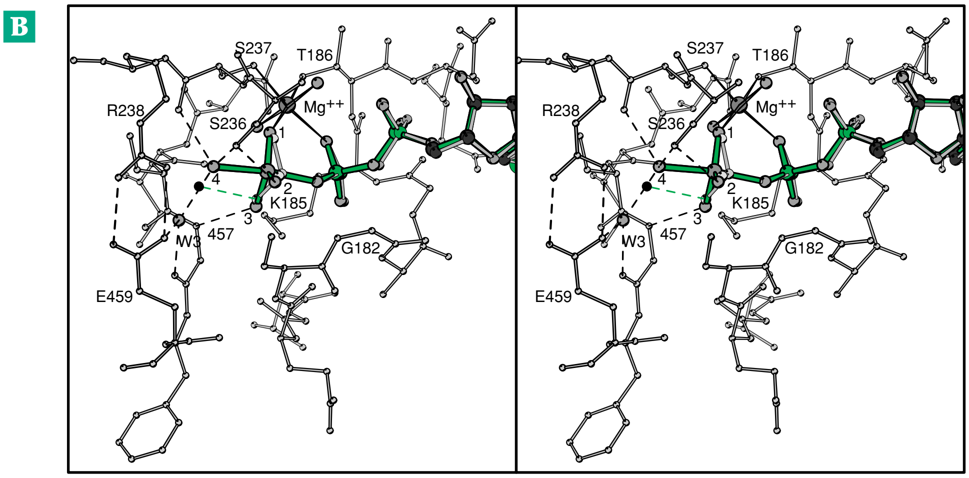


Figure 19-16 (A) The nucleotide binding site of myosin with MgADP•BeF_x bound in a conformation thought to mimic that of ATP prior to hydrolysis. The β-sheet strands are contributed by both the 25-kDa and 50-kDa domains. The P-loop lies between T178 and E187. The conserved N233 to G240 loop, which also contributes important ATP-binding residues, comes from the 50-kDa region. (B) Stereoscopic view of the γ-phospho group binding pocket with the bound MgADP•VO_4 (vanadate) complex. The coordinated Mg^{2+} and associated water molecules are seen clearly. Courtesy of Ivan Rayment.[157]

regions. Figure 19-16 illustrates the binding of an ATP analog, the beryllium fluoride complex of MgADP, in the active site. As can be seen, the ATP binds to loops at the C termini of the β strands of the 8-stranded β sheet from the 25-kDa domain. The conserved P-loop (Chapter 12, E), which lies between T178 and E187, curls around the α and β phospho groups, and has the sequence G(179)ESGAGKT. A second conserved loop N(233)SNSSR-G(240) from the 50-kDa domain contributes to the binding of ATP.

The actin-binding region of the myosin head is formed largely by the 50-kDa segment, which is split by a deep cleft into two separate domains (Fig. 19-15), both of which are thought to participate in binding to actin. A surface loop (loop 1) near the ATP-binding site at the junction of the 25- and 50-kDa regions affects the kinetic properties of myosin, probably by influencing product release. A second loop (loop 2, residues 626–647) at the junction of the 50- and 20-kDa regions interact with actin. Loop 2 contains a GKK sequence whose positive charges may interact with negative charges in the N-terminal part of actin.[162–164]

The C-terminal fragment of myosin contains a globular domain that interacts with both the 20-kDa and 50-kDa regions and contains an α-helical neck that connects to the helix of the coiled-coil myosin rod. This helical region is surrounded by the two myosin light chains (Fig. 19-15).[157] A pair of reactive thiol groups (from C697 and C707) in the globular domain are near the active site. Crosslinking of these cysteines by an –S–S– bridge has been utilized to trap nucleotide analogs in the active site.[165]

How does actin bind? The actin monomer consists of four subdomains, 1, 2, 3, and 4 numbered from the N terminus (Fig. 7-10). The negatively charged N-terminal region of actin contains the sequence

$$\mathrm{D-E-D-\overset{4}{E}\text{---}\overset{24}{D}-\overset{25}{D}.}$$

It may interact with loop 1 of myosin, which contains five lysines. However, to form a strong interaction with the myosin head a conformational change must occur in the myosin. A change may also occur in actin. Modeling suggests that a large nonpolar contact region involves actin residues A144, I341, I345, L349, and F352 and myosin residues P529, M530, I535, M541, F542, and P543. A conformational change in actin, which might involve largely the highly conserved actin subdomain 2, may also be required for tight interaction.[142,166–168]

Kinesins and other molecular motors. Before considering further how the myosin motor may work, we should look briefly at the **kinesins**, a different group of motor molecules,[168a] which transport various cellular materials along microtubule "rails." They also participate in organization of the mitotic spindle and other microtubule-dependent activities.[168a,b,c] See Section C,2 for further discussion. More than 90 members of the family have been identified. Kinesin heads have much shorter necks than do the myosin heads. A myosin head is made up of ~850 residues, but the motor domain of a kinesin contains only ~345. Like myosin, the 950- to 980-residue kinesins have a long coiled-coil C- terminal region that forms a "neck" of ~50 residues, a "stalk" of ~190 and ~330 residue segments with a Pro / Gly-rich hinge between them, and an ~45 residue "tail."[169–171]

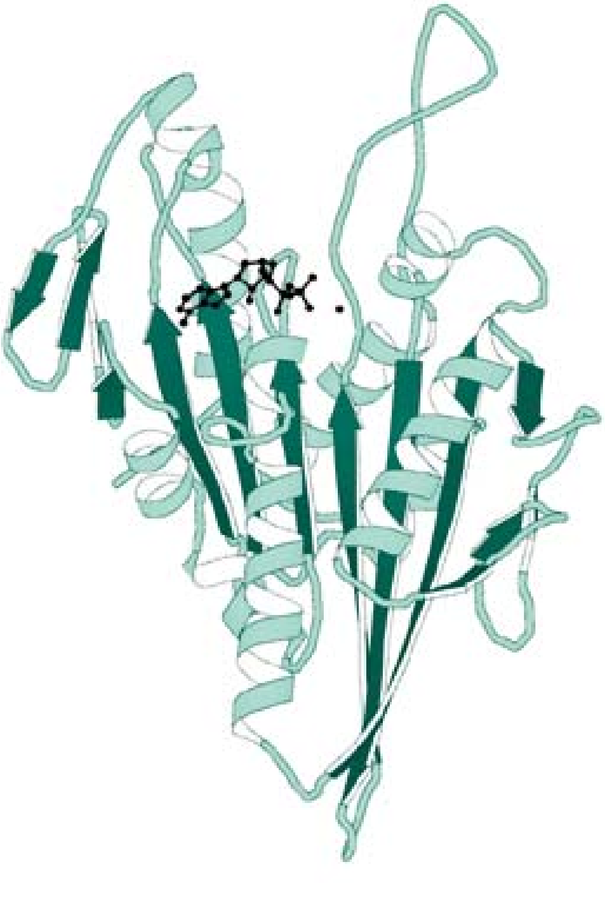


Figure 19-17 Ribbon drawing of human kinesin with bound Mg•ADP. From Gulick *et al.*[174] Courtesy of Ivan Rayment and Andy Gulick.

Crystal structures are known for motor domains of human kinesin[172] and of a kinesin from rat brain.[169,173] The structures of one of six yeast kinesins,[174] a protein called **Kar3**, and also of a *Drosophila* motor molecule designated **Ncd** have also been determined.[175] The last was identified through study of a *Drosophila* mutant called non-claret disjunctional (Ncd). The motor domains of various members of the kinesin family show ~40% sequence identity and very close structural identity (Fig. 19-17).[174] Although the sequences are different from those of the myosin heads or of G proteins, the folding pattern in the core structures is similar in all cases. An 8-stranded β sheet is flanked by three α helices on each side and a P-loop crosses over the ATP-binding site as in Fig. 19-16. Further similarity is found in the active site structures, which, for a monomeric kinesin KIF1A,[174a] have been determined both with bound ADP and with a nonhydrolyzable analog of ATP.[174b,174c] Although there is little similarity in amino acid sequences the structures in the catalytic core are clearly related to each other, to those of dimeric kinesins,[174d] to those of myosins, and to those of the GTP-hydrolyzing G proteins.

A puzzling discovery was that the motor domain of kinesin, which binds primarily to the β subunits of tubulin (Fig. 7-34) and moves toward the fast growing *plus* end of the microtubule,[176] is located at the N terminus of the kinesin molecule, just as is myosin. However, the Ncd and Kar3 motor domains are at the C-terminal ends of their peptide chains and move their "cargos" toward the *minus* ends of microtubules.[174] Nevertheless, the structures of all the kinesin heads are conserved as are the basic chemical mechanisms. The differences in directional preference are determined by a short length of peptide chain between the motor domain and the neck, which allows quite different geometric arrangements when bound to microtubules.[173,177] Like Ncd, myosin VI motor domains also move "backwards" toward the pointed (minus) ends of actin filaments.[178–179a]

Other major differences between kinesins and myosin II heads involve kinetics[180,181] and processivity.[173] Dimeric kinesin is a **processive** molecule. It moves rapidly along microtubules in 8-nm steps but remains attached.[182,182a] Myosins V and VI are also processive[183–183e] but myosin II is not. It binds, pulls on actin, and then releases it. The many myosin heads interacting with each actin filament accomplish muscle contraction with a high velocity in spite of the short time of attachment. Ncd and Kar3 are also nonprocessive and slower than the *plus* end-oriented kinesins.[184]

The ATPase cycles of actomyosin and of the kinesins. The properties of the protein assemblies found in muscle have been described in elegant details, but the most important question has not been fully answered. How can the muscle machinery use the Gibbs energy of hydrolysis of ATP to do mechanical work? Some insight has been obtained by studying the ATPase activity of isolated myosin heads (S1) alone or together with actin. Results of numerous studies of ATP binding, hydrolysis, and release of products using fast reaction techniques[185–191] and cryoenzymology[191a] are summarized in Fig. 19-18. In resting muscle the myosin heads swing freely in the ~20-nm space between the thick and thin filaments. However, in activated muscle some heads are bound tightly to actin as if in rigor (complex A•M in Fig. 19-18). When ATP is added MgATP binds into the active site of the myosin (Fig. 19-18, step *a*) inducing a conformational change to form A•M*•ATP in which the bond between actin and myosin is weakened greatly, while that between myosin and ATP is strengthened. The complex dissociates (step *b*) to give free actin and (M*•ATP), which accumulates at –15°C. However, at higher temperatures the bound ATP is hydrolyzed rapidly (step *c*) to a form M**•ADP•P_i in which the ATP has been cleaved to ADP + P_i but in which the split products remain bound at the active site.[116,192,192a,b] All of these reactions are reversible. That is, the split products can recombine to form ATP. This fact suggests that most of the Gibbs energy of hydrolysis of the ATP must be stored, possibly through a conformational change in the myosin head or through tighter bonding to ATP. As long as calcium ions are absent, there is only a slow release of the bound ADP and P_i and replacement with fresh ATP takes place. Thus, myosin alone shows a very weak ATPase activity.

On the other hand, in activated muscle the head with the split ATP products will bind to actin (step *d*), probably at a new position. The crossbridges that form appear to be attached almost at right angles to the thin filaments. In step *e*, P_i is released following a conformational alteration that is thought to open a "back door" to allow escape of the phosphate ion.[193] In the final two steps (*f* and *g*) the stored energy in the myosin head (or in the actin) is used to bring about another conformational change that alters the angle of attachment of myosin head to the thin filament from ~90° to ~45°. At least some indication of such a change can be observed directly by electron microscopy.[144] Such a change in angle is sufficient to cause the actin filament to move ~10 nm with respect to the thick filaments to complete the movement cycle (Fig. 19-18), if the head is hinged at the correct place. However, the existence of at least four different conformational states suggests a more complex sequence.[193a,193b] Examination of the three-dimensional structures available also suggest a complex sequence of alterations in structure and geometry. X-ray crystallographic structures of myosin heads, in states thought to correspond to states 1 and 3 of Figs. 19-14 and 19-18, are also in agreement on an ~10 (5–12) nm movement of the lever arm.[194,195] Six states of the actomyosin complex are depicted in Fig. 19-18, but a complete kinetic analysis requires at least eight and possibly 12 states.[196,197]

Observing single molecules. A major advance in the study of molecular motors has been the development of ways to observe and study single macromolecules. The methods make use of **optical traps** (optical "tweezers") that can hold a very small (~1 µm diameter) polystyrene or silica bead near the waist of a laser beam focused through a microscope objective.[198–202] In one experimental design an F-actin filament is stretched between two beads, held in a pair of optical traps. The filament is pulled taut and lowered onto a stationary silica bead to which a few myosin HMM fragments have been attached (Fig. 19-19). If ATP is present, short transient movements along the filament are detected by observation of displacements of one of the beads when the actin filament contacts HMM heads. An average lateral displacement of 11 nm was observed. Each HMM head exerted a force of 3–4 pN, a value consistent with expectations for the swinging bridge model.[200] From the duration of a single displacement (≤7 ms) and an estimated k_{cat} for ATP hydrolysis of 10 s^{-1}, the fraction of time that the head was attached during one catalytic cycle of the head was therefore only 0.07. This ratio, which is called the **duty ratio**, is low for actomyosin. However, many myosin heads bind to each actin filament in a muscle. Each head exerts its pull for a short time. but the actin is never totally unattached.[203] Similar measurements with smooth muscle revealed similar displacements but with a 10-fold slower sliding velocity and a 4-fold increase in the duty ratio. This may perhaps account for an observed 3-fold increase in force as compared with skeletal muscle.[204,204a]

Other single molecule techniques involve direct observation of motor molecules or of S1 myosin fragments tagged with highly fluorescent labels.[205,206] All measurements of single molecule movement are subject to many errors. Brownian motion of the beads makes measurements difficult.[207] Not all results are in agreement, and some are difficult to understand.[207a] Most investigators agree that there is a step size of ~4–10 nm. Kitamura *et al.* found 5.3 nm as the average.[206] However, they also reported the puzzling observation that some single S1 molecules moved 11–30 nm in two to five rapid successive steps during the time of hydrolysis of a single molecule at ATP. They suggested that some of the energy of ATP hydrolysis may be stored in S1 or in the actin filament and be released in multiple steps. Veigel *et al.*[208] observed that a brush border myosin I from chicks produced ~ 6 nm movements, each of which was followed by an additional ~ 5.5 nm step within ~5 ms. They attribute these steps to two stages in the power stroke, e.g., to steps *f* and *g* of Fig. 19-18. A value of ~10nm was reported recently by Piazzesi *et al*.[208a] Myosin V moves along actin filaments with very large 36-nm steps.[208b]

Motion of kinesin heads has been observed by movement of microtubules over biotinylated kinesin fixed to a steptavidin-coated surface,[209] by direct observation of fluorescent kinesin moving along microtubules,[171] and by optical trap interferometry.[210] Kinesin heads move by 8-nm steps, evidently the exact length

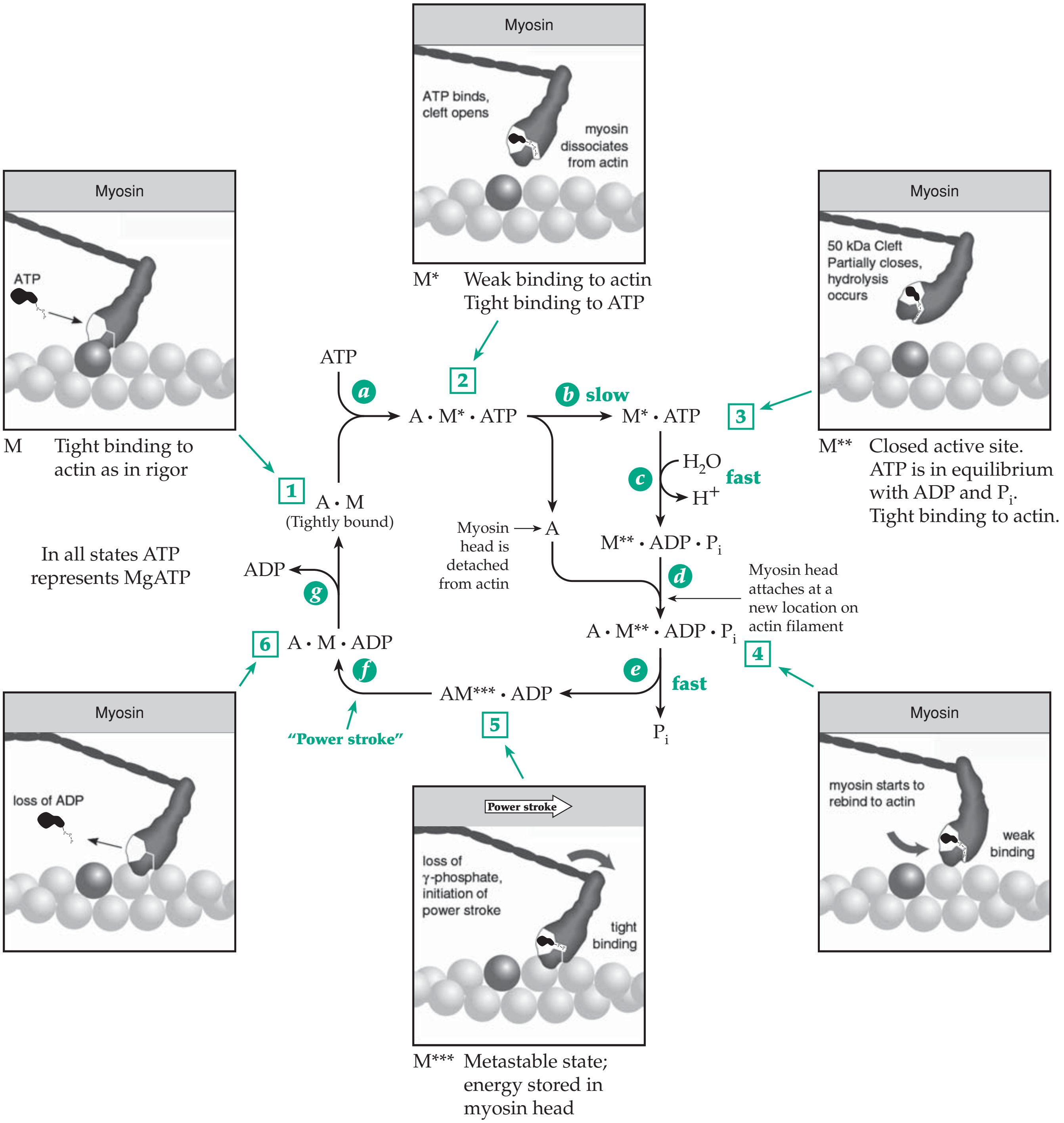

Myosin
ATP binds, cleft opens
myosin dissociates from actin
M* Weak binding to actin Tight binding to ATP
Myosin
ATP
M Tight binding to actin as in rigor
Myosin
50 kDa Cleft Partially closes, hydrolysis occurs
M** Closed active site. ATP is in equilibrium with ADP and P_i. Tight binding to actin.
ATP
a
2
A • M* • ATP
b slow
M* • ATP
3
1
A • M (Tightly bound)
c
H_2O
fast
H^+
Myosin head is detached from actin
A
M** • ADP • P_i
In all states ATP represents MgATP
ADP
g
d
Myosin head attaches at a new location on actin filament
A • M** • ADP • P_i
4
6
A • M • ADP
f
e
fast
AM*** • ADP
5
P_i
"Power stroke"
Myosin
loss of ADP
Myosin
Power stroke
loss of γ-phosphate, initiation of power stroke
tight binding
M*** Metastable state; energy stored in myosin head
Myosin
myosin starts to rebind to actin
weak binding

Figure 19-18 Simplified view of the ATP hydrolysis cycle for actomyosin. A similar cycle can be written for kinesins and dyneins. Here A stands for fibrous actin and M, M*, M**, and M*** for four different conformations of the myosin heads. As indicated by the numbers in squares, four of the six states of actomyosin shown here can also be correlated with those in Fig. 19-14.

of an αβ tubulin dimer in the microtubule structure (Fig. 7-33). One ATP is apparently hydrolyzed for each 8-nm step. However, shorter steps of ~5 nm have sometimes been seen.[211,212]

The movement is processive, kinesin motors typically taking 100 steps before dissociating from the microtubule.[201,212a] Kinesin is bound to the microtubule continuously. Its duty ratio is nearly 1.0 (the same is true for the bacterial flagellar motor; Fig. 19-4).[212b] However, single kinesin heads, which lack the coiled-coil neck region, have a duty ratio of <0.45. The movement is nonprocessive.[213] The Ncd motor is also nonprocessive.[214–216] As mentioned previously, the Ncd and kinesin motor domains are at opposite ends of the peptide chain, and the motors move in opposite directions along microtubules.[217,218] The critical difference between the two motor molecules was found in the neck domains, which gave rise to differing symmetrics in the two heads.[219] The latter are shown in Fig. 19-20, in which they have been docked onto the tubulin protofilament structure. One head, both of ncd and of kinesin, occupies a similar position on the microtubule, but the other head points toward the microtubular plus end for kinesin but toward the minus end for Ncd. Cryoelectron microscopy also supports this interpretation.[220]

Still not fully understood is the processive action of kinesin.[221–224] It is often assumed that this protein moves in a hand-over-hand fashion with the two heads binding alternatively to the microtubule. Some substantial reorganization of the peptide chain in the hinge region at the end of the neck is presumably involved.[173] An alternative "inchworm" mechanism has been suggested.[220a]

Thinking about chemical mechanisms. We have now examined the active sites of kinases that cleave ATP (Chapter 12), ATPases that pump ions by cleaving ATP, ATP synthases that form ATP from ADP and P_i (Chapter 18), and GTP hydrolyzing enzymes that cause movement and shape changes that control metabolic processes (Chapter 11). It is striking that the active site regions where the ATP or GTP bind have such a highly conserved structure.[225] This suggests that the secret of movement can be found in the very strong interactions of the nucleotides and their split products with the proteins. In every case there is at least one tight binding or closed conformation in which a large number of hydrogen bonds and ionic interactions bind the nucleotide. This is shown for a kinase in Fig. 12-32 and for myosin in Fig. 19-16. During the actomyosin reaction several conformational changes must occur. Not only does the affinity for the bound nucleotide vary, but also the binding of actin to myosin can be strong, as in the nucleotide-free state or

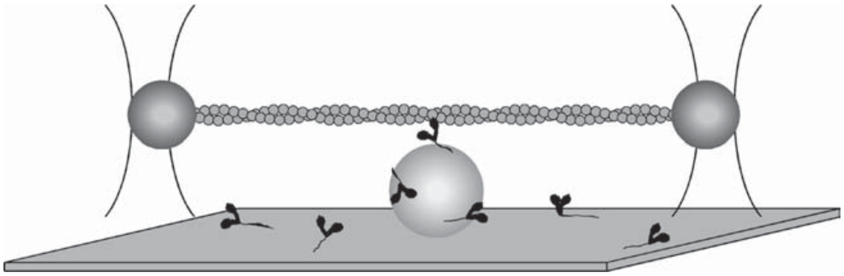


Figure 19-19 Schematic drawing (not to scale) illustrating the use of two optical traps that are focused on beads attached to a single actin filament. The filament is lowered onto a stationary silica bead sparsely coated with HMM fragments of myosin. In the presence of ATP the myosin heads bind transiently for a few milliseconds to the actin, moving it in one direction and displacing the beads from their positions in the optical traps. An image of one of the beads is projected onto photodiode detectors capable of measuring small displacements. The displacing force can also be recorded. For details of the experiments and of the optical traps and measuring devices see Finer *et al.*[200] Courtesy of J. A. Spudich.

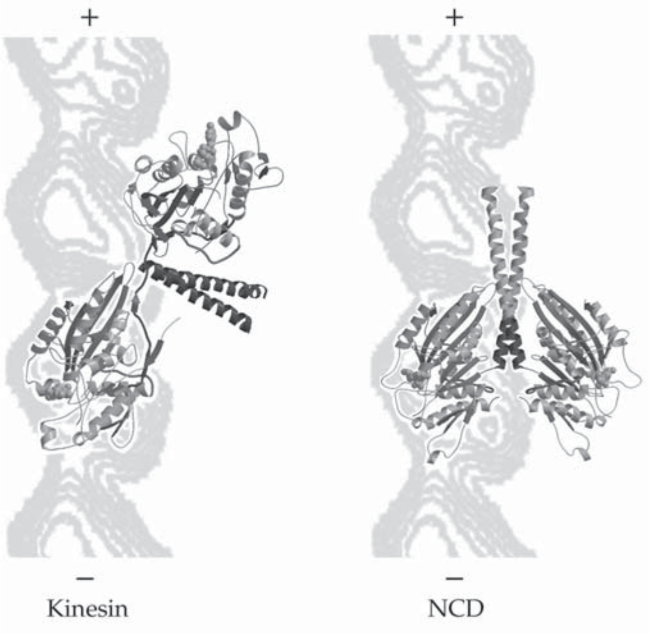
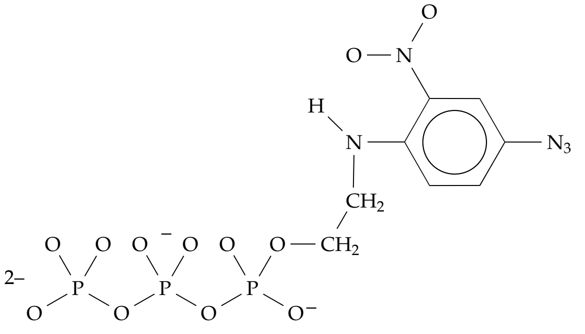


Figure 19-20 Model showing the ncd and kinesin dimer structures docked onto a tubulin protofilament. The bound ncd and kinesin heads are positioned similarly. Because of the distinct architectures of the kinesin and ncd necks, the unbound kinesin head points toward the plus end, whereas the unbound ncd head is tilted toward the minus end of the protofilament. From Sablin *et al.*[219] Courtesy of Ronald Vale.

in the presence of bound ADP. Binding is weak when ATP or the split product ADP + P$_i$ are in the active site.

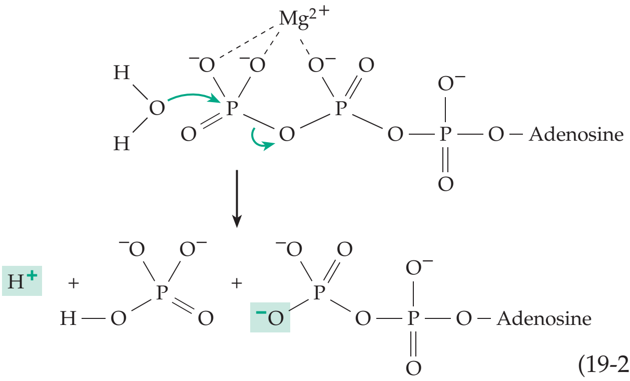
To understand these differences we should look at the structure of ATP itself. The triphosphate group has many negative charges repelling each other. What must happen to allow the binding of ATP to myosin to break the actin-myosin bond? The electrostatic attraction of these phospho groups for active site groups is doubtless one cause of the observed conformational changes. Could it be that electrostatic repulsion, via a proton shuttle mechanism, is also induced at the right point in the actin–myosin interface? Many studies with analogs of ATP have contributed to our understanding. Neither the purine nor the ribose ring of ATP is absolutely essential. The compound 2-[(4-azido-2-nitrophenyl) amino] diphosphate (NANDP) and related nonnucleotide analogs[165,196,226] support muscle contraction and relaxation in the same way as does ATP. An analog with a rigid five-membered ring, 2',3'-dideoxydidehydro-ATP, is also active.[227] A comparison of kinetic data and X-ray structural data supports the proposal that the ATP must be bound in the conformation shown for MgADP•BeF$_x$ in Fig. 19-16A.[196] When the two SH groups of C697 and C707 of the myosin head are crosslinked by various reagents,[227a] this NANDP analog can be trapped at the active site. Because of the presence of its azide ($-N_3$) group the trapped compound can serve as a photoaffinity label, attaching itself to a tryptophan side chain upon activation with visible light (Eq. 23-27).



NANTP

How can cleavage of ATP to bound ADP + P$_i$ create a metastable high-energy state of the myosin head ready to hold onto and pull the actin chain? This may be compared with the inverse problem of generating ATP in oxidative phosphorylation, in which ADP and P$_i$ coexist in equilibrium with ATP in a closed active site (Fig. 18-14). Comparison should also be made with the GTP-hydrolyzing G proteins (Fig. 11-7).[227b,c] During hydrolysis of GTP by the Ras protein binding to the protein induces a shift of negative charge from the γ- to β-oxygens of GTP facilitating bond cleavage as in Eq. 19-2. G proteins also couple substrate hydrolysis to mechanical motion. We should also think about the fact that when ATP is cleaved within myosin there will necessarily be a flow of electrical charge from the water to the ADP (Eq. 19-2). This will be followed by some accomodation of the protein to the new charge constellation. As we have seen previously, movement of protons is often the key to conformational changes. In this case, the initial change must be to create a high-energy state of myosin which, following loss of the orthophosphate ion, can cause the major conformational change that swings the lever arm of the myosin. The conformational changes may occur in several steps in which the packing of groups within the myosin head is always tight in some places and rather loose in others. Movement within the head is being observed not only by X-ray crystallography but by **fluorescence resonance energy transfer** (FRET; Chapter 23)[227d] and by the newer **luminescence resonance energy transfer** (LRET). For example, a terbium chelate of azide-ATP was photochemically bound in the active site, and a fluorescent dye was attached to Cys 108 in the regulatory light chain. The terbium ion was irradiated, and fluorescence of the dye was observed. Distance changes, measured in the absence and presence of ATP, were consistent with the swinging arm model.[228] Dyes have been attached to –SH groups engineered into various locations in the myosin molecule to permit other distance measurements.[229,230] In another elegant application of the FRET technique the **green fluorescent protein** of *Aequoria* (Box 23-A) was fused to the C terminus of the motor domain of myosin giving a fluorescent lever arm. Energy transfer to blue fluorescent protein fused to the N terminus of the S1 head was measured. The distance between these was estimated by the FRET technique and was also consistent with expectations for the "rowing model."[231]

The "rowing model" is generally accepted, but other quite different processes have been proposed to account for the elementary cycle of muscle contraction. Muscle contracts nearly *isovolumetrically*; thus, anything that expands the sarcomere will cause a contraction. It has been suggested that the hydrolysis of ATP deposits negatively charged phospho groups on the actin filaments, and that the electrostatic repulsion is responsible for



(19-2)

BOX 19-A HEREDITARY DISEASES OF MUSCLE

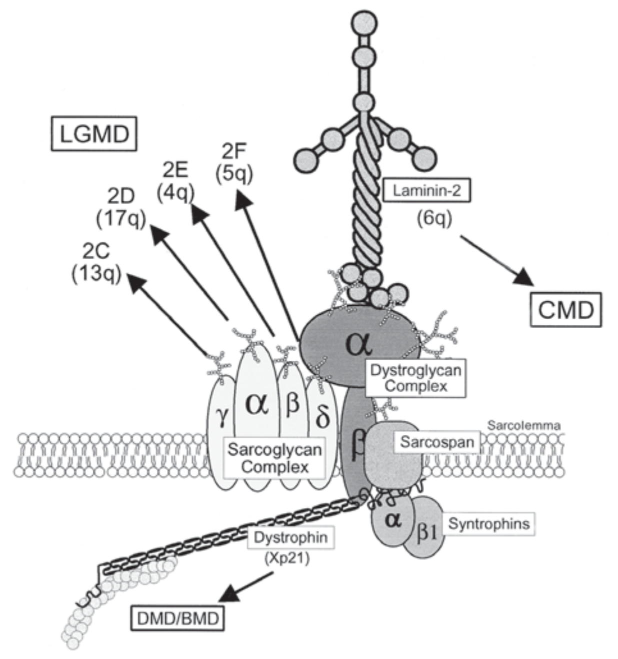
Considering the numerous specialized proteins in muscle it is not surprising that many rare hereditary muscle diseases are known. The most frequent and most studied of these is **Duchenne muscular dystrophy**. An X-linked disease of boys, it may not be recognized until two to three years of age, but victims are usually in a wheelchair by age 12 and die around age 20. Individual muscle fibers disintegrate, die, and are replaced by fibrous or fatty tissue.[a–d] The disease strikes about 1 out of 3500 boys born. The less serious **Becker muscular dystrophy** arises from defects in the same gene but affects only 1 in 30,000 males, some of whom have a normal life span. Because of its frequency and the knowledge that the gene must lie in the X-chromosome, an intensive search for the gene was made. It was found in 1986 after a five-year search.[a,e] This was the first attempt to find a faulty gene whose protein product was totally unknown. The project, which relied upon finding restriction fragment polymorphisms (Chapter 26) that could serve as markers in the genome, made use of the DNA from patients with a range of related diseases. The very rare female patients in whom the faulty gene had been translocated from the X-chromosome to an autosome also provided markers. DNA probes obtained from a young man with a large X-chromosome deletion that included genes related to retinitis pigmentosa and several other diseases provided additional markers. The result was a triumph of "reverse genetics" which has since been applied to the location of many other disease genes.[e]

The muscular dystrophy gene may be the largest human gene. It consists of 2.3 million base pairs, which include 79 exons which encode a huge 427 kDa protein named **dystrophin**. The protein consists of four main domains.[a,f,g] The N-terminal domain binds to actin and is homologous to β-actinin. The central domain is an elongated rod resembling spectrin. It contains repetitive coiled-coil segments and four hinge regions. The third domain is rich in cysteine and binds Ca^{2+}, while the fourth domain has a structure that is shared by several other proteins of the dystrophin family. Dystrophin is quantitatively a minor protein of muscle. It forms part of the cytoskeleton, lying adjacent to the sarcolemma (cell membrane) along with β-spectrin and vinculin (see Fig. 8-16).

While one end of the dystrophin molecule binds to actin filaments, the C-terminal domain associates with several additional proteins to form a **dystrophin–glycoprotein complex** (see figure).[f,h–k] Dystrophin is linked directly to the membrane-spanning protein **β-dystroglycan,** which in the outer membrane surfaces associates with a glycoprotein **α-dystroglycan**. The latter binds to laminin-2 (Fig. 8-33), a protein that binds the cell to the basal lamina. Four other membrane-spanning proteins, α-, β-, γ-, and δ-**sarcoglycans**, are among additional members of the complex.[k–m]

Patients with Duchenne muscular dystrophy are deficient not only in dystrophin but also in the dystroglycan and sarcoglycan proteins.[f,n] Evidently, dystrophin is needed for formation of the complex which plays an essential role in muscle. In both types of X-linked muscular dystrophy there are individuals with a wide range of point mutations, frame-shift mutations, and deletions in the dystrophin gene.[d] The essential function of dystrophin and associated proteins is uncertain but may be related to the linkage from actin filaments through the membrane to laminin. Individuals with Becker muscular dystrophy also have defects in dystrophin, but the protein is partially functional. Some other muscular dystrophies are caused by defects in the autosomal genes of any of the four sarcoglycan subunits.[j,k,o] or in that of laminin α1 chain.[p,q] The arrows in the accompanying drawing indicate chromosome locations of the sarcoglycan subunits, which are sites for mutations causing **limb girdle muscular dystrophy**.[k]

Dystrophin, shorter isoforms, and related proteins are found in many tissues including the brain.[s] One related protein, **utrophin** (dystrophin-related



Schematic model of the dystrophin–glycoprotein complex. Courtesy of Kevin P. Campbell. See Lim and Campbell[r] Abbreviations: LGMD, Limb Girdle muscular dystrophy; CMD, congenital muscular dystrophy; DMD / BMD, Duchenne / Becker muscular dystrophies.

BOX 19-A (continued)

protein 1), is present in the neuromuscular junction of adult skeletal muscle. One approach to therapy of Duchenne muscular dystrophy is to stimulate a higher level of expression of the utrophin gene.[t] Because the dystrophin gene is so large treatment by gene transfer is not practical, but transfer of parts of the gene may be. Myoblast transfer has not been successful, but new approaches will be devised.[d]

Myotonic dystrophy is a generalized adult-onset disorder with muscular spasms, weakness, and many other symptoms.[u–y] It is one of the **triple-repeat diseases** (Table 26-4). The affected gene encodes a protein kinase of unknown function. The corresponding mRNA transcript has ~2400 nucleotides. The gene has a CTG repeat $(CTG)_n$ near the 3'-end with $n < 30$ normally. For persons with the mildest cases of myotonic dystrophy n may be over 50 while in severe cases it may be as high as 2000. As in other triple-repeat diseases the repeat number tends to increase in successive generations of people as does the severity of the disease.[x]

For some individuals, muscular dystrophy causes no obvious damage to skeletal muscle but affects the heart producing a severe **cardiomyopathy,** and persons with Duchenne muscular dystrophy often die from heart failure. Heart failure from other causes, some hereditary, is a major medical problem, especially among older persons. Hereditary forms are often autosomal dominant traits that may cause sudden death in young persons. At least seven genes for cardiac sarcomeric proteins including actin,[z] myosin, both heavy and light chains,[aa–dd] three subunits of troponin,[ee] tropomyosin, and protein C (p. 1104) may all carry mutations that cause cardiomyopathy.

A hereditary disease common in Japan results from a defect in migration of neurons and is associated with brain malformation as well as muscular dystrophy.[ff] In **nemaline myopathy** a defect in nebulin leads to progressive weakness and often to death in infancy. A characteristic is the appearance of "nemaline bodies" or thickened Z-discs containing Z-disc proteins.[gg] Some hereditary diseases involve nonmuscle myosins. Among these is **Usher syndrome**, the commonest cause of deaf-blindness. The disease, which results from a defect in the myosin VIA gene, typically causes impairment of hearing and retinitis pigmentosa (Chapter 23).[hh]

---

[a] Anderson, M. S., and Kunkel, L. M. (1992) *Trends Biochem. Sci.* **17**, 289–292

[b] Emery, A. E. H., and Emery, M. L. H. (1995) *The History of a Genetic Disease: Duchenne Muscular Dystrophy or Meryon's Disease*, Royal Society of Medicine, London

[c] Brown, S. C., and Lucy, J. A., eds. (1997) *Dystrophin: Gene, Protein and Cell Biology*, Cambridge Univ. Press, London

[d] Worton, R. G., and Brooke, M. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4195–4226, McGraw-Hill, New York

[e] Rowland, L. P. (1988) *N. Engl. J. Med.* **318**, 1392–1394

[f] Tinsley, J. M., Blake, D. J., Zuellig, R. A., and Davies, K. E. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8307–8313

[g] Fabbrizio, E., Bonet-Kerrache, A., Leger, J. J., and Mornet, D. (1993) *Biochemistry* **32**, 10457–10463

[h] Ervasti, J. M., and Campbell, K. P. (1991) *Cell* **66**, 1121–1131

[i] Madhavan, R., and Jarrett, H. W. (1995) *Biochemistry* **34**, 12204–12209

[j] Sweeney, H. L., and Barton, E. R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13464–13466

[k] Jung, D., and 13 others. (1996) *J. Biol. Chem.* **271**, 32321–32329

[l] Winder, S. J. (2001) *Trends Biochem. Sci.* **26**, 118–124

[m] McDearmon, E. L., Combs, A. C., and Ervasti, J. M. (2001) *J. Biol. Chem.* **276**, 35078–35086

[n] Durbeej, M., and Campbell, K. P. (1999) *J. Biol. Chem.* **274**, 26609–26616

[o] Yang, B., Ibraghimov-Beskrovnaya, O., Moomaw, C. R., Slaughter, C. A., and Campbell, K. P. (1994) *J. Biol. Chem.* **269**, 6040–6044

[p] Noguchi, S., and 17 others. (1995) *Science* **270**, 819–822

[q] Tiger, C.-F., Champliaud, M.-F., Pedrosa-Domellof, F., Thornell, L.-E., Ekblom, P., and Gullberg, D. (1997) *J. Biol. Chem.* **272**, 28590–28595

[r] Lim, L. E., and Campbell, K. P. (1998) *Curr. Opin. Neurol.* **11**, 443–452

[s] Dixon, A. K., Tait, T.-M., Campbell, E. A., Bobrow, M., Roberts, R. G., and Freeman, T. C. (1997) *J. Mol. Biol.* **270**, 551–558

[t] Campbell, K. P., and Crosbie, R. H. (1996) *Nature (London)* **384**, 308–309

[u] Harper, P. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4227–4252, McGraw-Hill, New York

[v] Ptacek, L. J., Johnson, K. J., and Griggs, R. C. (1993) *N. Engl. J. Med.* **328**, 482–489

[w] Pearson, C. E., and Sinden, R. R. (1996) *Biochemistry* **35**, 5041–5053

[x] Tapscott, S. J., and Thornton, C. A. (2001) *Science* **293**, 816–817

[y] Miller, J. W., Urbinati, C. R., Teng-umnuay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A., and Swanson, M. S. (2000) *EMBO J.* **19**, 4439–4448

[z] Olson, T. M., Michels, V. V., Thibodeau, S. N., Tai, Y.-S., and Keating, M. T. (1998) *Science* **280**, 750–755

[aa] Rayment, I., Holden, H. M., Sellers, J. R., Fananapazir, L., and Epstein, N. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3864–3868

[bb] Roopnarine, O., and Leinwand, L. A. (1998) *Biophys. J.* **75**, 3023–3030

[cc] Yanaga, F., Morimoto, S., and Ohtsuki, I. (1999) *J. Biol. Chem.* **274**, 8806–8812

[dd] Martinsson, T., Oldfors, A., Darin, N., Berg, K., Tajsharghi, H., Kyllerman, M., and Wahlström, J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14614–14619

[ee] Miller, T., Szczesna, D., Housmans, P. R., Zhao, J., de Freitas, F., Gomes, A. V., Culbreath, L., McCue, J., Wang, Y., Xu, Y., Kerrick, W. G. L., and Potter, J. D. (2001) *J. Biol. Chem.* **276**, 3743–3755

[ff] Kobayashi, K., and 17 others. (1998) *Nature (London)* **394**, 388–392

[gg] Pelin, K., and 19 others. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2305–2310

[hh] Weil, D., and 18 others. (1995) *Nature (London)* **374**, 60–61

lateral expansion of the sarcomere.[232] Still other ideas have been advanced.[233,234]

4. Control of Muscle Contraction

Skeletal muscle must be able to rest without excessive cleavage of ATP but able to act rapidly with a high expenditure of energy upon nervous excitation. Even a simple physical activity requires that a person's muscles individually contract and relax in rapid response to nerve impulses from the brain. To allow for this control the endoplasmic reticulum (**sarcoplasmic reticulum, SR**) of striated muscle fibers is organized in a striking regular manner.[235–237] Interconnecting tubules run longitudinally through the fibers among the bundles of contractile elements. At regular intervals they come in close contact with infoldings of the outer cell membrane (the **T system** of membranes, Fig. 19-21; see also Fig. 19-7A). A nerve impulse enters the muscle fiber through the neuromuscular junctions and travels along the sarcolemma and into the T tubules. At the points of close contact the signal is transmitted to the longitudinal tubules of the sarcoplasmic reticulum, which contain a high concentration of calcium ions.

Calcium ions in muscle. A nerve signal arriving at a muscle causes a sudden release of the calcium ions into the cytoplasm from cisternae of the sarcoplasmic reticulum (SR) that are located adjacent to the T-tubules. Diffusion of the Ca^{2+} into the myofibrils initiates contraction. In smooth muscle the signals do not come directly from the nervous system but involve hormonal regulation.[238] Again, calcium ions play a major role, which is also discussed in Chapter 6, Section E, and in Box 6-D. Muscle contains a large store of readily available Ca^{2+} in lateral cisternae of the SR. The free intracellular Ca^{2+} concentration is kept low by a very active ATP-dependent calcium ion pump (Fig. 8-26), which is embedded in the membranes of the SR.[238a] Within the vesicles Ca^{2+} is held loosely by the ~63-kDa protein **calsequestrin,** which binds as many as 50 calcium ions per molecule. When the cytoplasmic concentration of free Ca^{2+} falls below ~10^{-6} M, contraction ceases. In fast-contracting skeletal muscles the Ca^{2+}-binding protein **parvalbumin** (Fig. 6-7) may

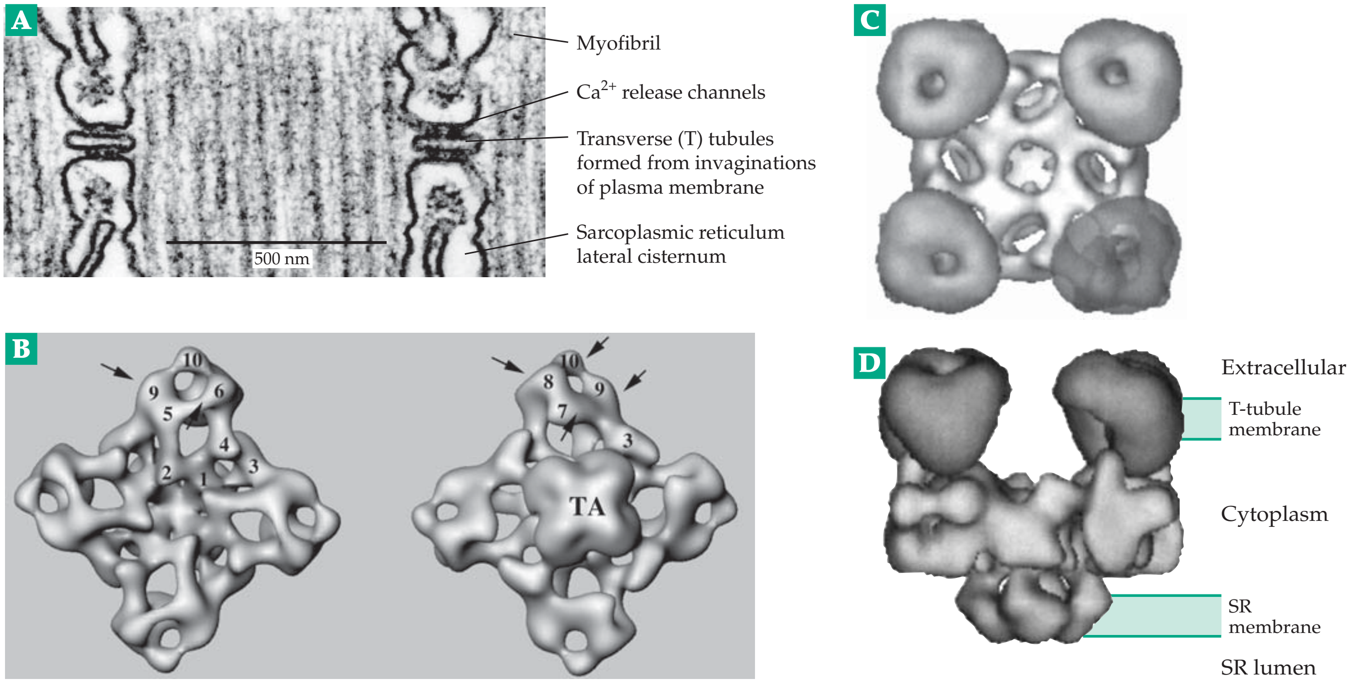
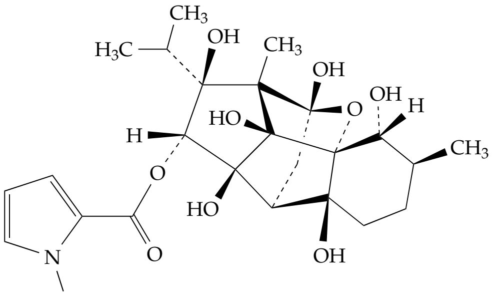


Figure 19-21 (A) Electron micrograph showing two transverse tubules (T-tubules) that are formed by infolding of the plasma membrane. They wrap around a skeletal muscle fiber and carry nerve impulses to all parts of the fiber. From Alberts *et al.*[237] Courtesy of Clara Franzini-Armstrong. (B) Three-dimensional surface representation of the calcium release channels known as ryanodine receptors, type RyR1 based on cryoelectron microscopy and image reconstruction at a resolution of 4.1 nm. The image to the left shows the surface that would face the cytoplasm while that to the right shows the surface that would interact with the sarcoplasmic reticulum, TA representing the transmembrane portion. Notice the fourfold symmetry of the particle, which is composed of four 565-kDa subunits. From Sharma *et al.*[239] Courtesy of Manjuli Rani Sharma. (C), (D) Model showing proposed arrangement of ryanodine receptors and dihydropyridine receptors (round) in the T tubule and SR membranes. From Serysheva *et al.*[245a]

assist in rapid removal of free Ca^{2+} from the cytoplasm. Contraction is activated when Ca^{2+} is released from the SR through the **calcium release channels**,[240–244] which are often called **ryanodine receptors**. The name arises from their sensitivity to the insecticidal plant alkaloid ryanodine, which at low concentration (≤60 µM) causes the channels to open, but a high concentration closes the channels.[243] These calcium release channels consist of tetramers of ~5000-residue proteins. The bulk of the 565-kDa polypeptides are on the cytosolic surface of the SR membranes, where they form a complex "foot" structure (Fig. 19-21B) that spans the ~12 nm gap between the SR vesicles and the T-tubule membrane.[239,240] Ryanodine receptor function is modulated by NO, which apparently binds to –SH groups within the Ca^{2+} channel.[243a,243b] Some ryanodine receptors are activated by cyclic ADP-ribose (cADPR, p. 564).[243c] Some have an oxidoreductase-like structural domain.[243d]

The release channels open in response to an incompletely characterized linkage to the **voltage sensor** that is present in the T-tubule membrane and is known as the **dihydropyridine receptor**.[240,245] This too, is a Ca^{2+} channel, which opens in response to arrival of an **action potential** (nerve impulse; see Chapter 30) that move along the T-tubule membrane. Because the action potential arrives almost simultaneously throughout the T-tubules of the muscles, the dihydropyridine receptors all open together. It isn't clear whether the linkage to the calcium channels is via stimulation from released Ca^{2+} passing from the dihydropyridine receptor to the surfaces of the feet of the release channels, or is a result of depolarization of the T-tubule membrane, or involves direct mechanically linked conformational changes.[240,245] The close cooperation of the Ca^{2+} release channel and the voltage sensor is reflected in their close proximity. In the sarcoplasmic reticulum every second release channel is adjacent to a voltage sensor in the opposing T-tubule membrane.[240,245a] The essential nature of the voltage



Ryanodine

BOX 19-B MALIGNANT HYPERTHERMIA AND STRESS-PRONE PIGS

Very rarely during surgery the temperature of a patient suddenly starts to rise uncontrollably. Even when heroic measures are taken, sudden death may follow within minutes. This **malignant hyperthermia syndrome** is often associated with administration of halogenated anesthetics such as a widely used mixture of halothane (2-bromo-2-chloro-1,1,1-trifluoroethane) and succinylcholine.[a–d] There is often no warning that the patient is abnormally sensitive to anesthetic. However, development of an antidote together with increased alertness to the problem has greatly decreased the death rate. Nevertheless, severe damage to nerves and kidneys may still occur.[c] Biochemical investigation of the hyperthermia syndrome has been facilitated by the discovery of a similar condition that is prevalent among certain breeds of pigs. Such "stress-prone" pigs are likely to die suddenly of hyperthermia induced by some stress such as shipment to market. The sharp rise in temperature with muscles going into a state of rigor is accompanied by a dramatic lowering of the ATP content of the muscles.

The problem, both in pigs and in humans susceptible to malignant hyperthermia, was found in the Ca^{2+} release channels (ryanodine receptors). Study of inheritance in human families together with genetic studies in pigs led to the finding that the stress-prone pigs have cysteine replacing arginine 615 in the Ca^{2+} channel protein. This modification appears to facilitate opening of the channels but to inhibit their closing.[e] A similar mutation has been found in some human families in which the condition has been recognized. However, there is probably more than one site of mutation in humans.[c,f] Similar mutations in the nematode *C. elegans* are being investigated with the hope of shedding light both on the problem of hyperthermia and on the functioning of the Ca^{2+} release channels.[g]

---

[a] Gordon, R. A., Britt, B. A., and Kalow, W., eds. (1973) *International Symposium on Malignant Hyperthermia*, Thomas, Springfield, Illinois

[b] Clark, M. G., Williams, C. H., Pfeifer, W. F., Bloxham, D. P., Holland, P. C., Taylor, C. A., and Lardy, H. A. (1973) *Nature (London)* **245**, 99–101

[c] MacLennan, D. H., and Phillips, M. S. (1992) *Science* **256**, 789–794

[d] Simon, H. B. (1993) *N. Engl. J. Med.* **329**, 483–487

[e] Fujii, J., Otsu, K., Zorzato, F., de Leon, S., Khanna, V. K., Weiler, J. E., O'Brien, P. J., and MacLennan, D. H. (1991) *Science* **253**, 448–451

[f] MacLennan, D. H., Duff, C., Zorzato, F., Fujii, J., Phillips, M., Korneluk, R. G., Frodis, W., Britt, B. A., and Worton, R. G. (1990) *Nature (London)* **343**, 559–561

[g] Sakube, Y., Ando, H., and Kagawa, H. (1997) *J. Mol. Biol.* **267**, 849–864

sensor is revealed by a lethal mutation (**muscular disgenesis**) in mice. Animals with this autosomal recessive trait generate normal action potentials in the sarcolemma but Ca^{2+} is not released and no muscular contraction occurs. They lack a 170-kDa dihydropyridine-binding subunit of the sensor.[246]

Some aspects of regulation by calcium ions are poorly understood. The frequent observations of oscillations in [Ca^{2+}] in cells is described in Box 6-D. Another phenomenom is the observation of Ca^{2+} "sparks," detected with fluorescent dyes and observation by confocal microscopy.[247] These small puffs of Ca^{2+} have been seen in cardiac muscle[247] and in a somewhat different form in smooth muscle.[248] They may represent the release of Ca^{2+} from a single release channel or a small cluster of channels. When the calcium release channels open, Ca^{2+} ions flow from the cisternae of the SR into the cytoplasm, where they activate both the troponin–tropomyosin system and also the Ca^{2+}-calmodulin-dependent **light chain protein kinase,** which acts on the light chains of the myosin head. These light chains resemble calmodulin in their Ca^{2+}-binding properties. The function of light chain phosphorylation of skeletal muscle myosin is uncertain but it is very important in smooth muscle.[248a]

The regulatory complex of tropomyosin and troponin is attached to the actin filaments as indicated in Fig. 19-8D and also in Fig. 19-9. The latter shows a model at near atomic resolution but without side chains on the tropomyosin and without the troponin components. When the regulatory proteins are completely removed from the fibrils, contraction occurs until the ATP is exhausted. However, in the presence of the regulatory proteins and in the absence of calcium, both contraction and hydrolysis of ATP are blocked. Tropomyosin (Tm) is a helical coiled-coil dimer, a 40-nm rigid rod, in which the two 284-residue 33-kDa monomers have a parallel orientation (Fig. 19-9)[249] as in the myosin tail. However, an 8– to 9– residue overlap at the ends may permit end-to-end association of the Tm molecules bound to the actin filament. As with other muscle proteins there are several isoforms,[250,251] whose distribution differs in skeletal and smooth muscle and in platelets. The elongated Tm rods appear to fit into the grooves between the two strands of actin monomers in the actin filament.[252–254] In resting muscle the Tm is thought to bind to actin near the site at which the S1 portion of the myosin binds. As a consequence, the Tm rod may block the attachment of the myosin heads to actin and prevent actin-stimulated hydrolysis of ATP. The 40-nm Tm rod can contact about seven actin subunits at once (Fig. 19-9). Thus, one Tm–troponin complex controls seven actin subunits synchronously.

Troponin (Tn) consists of three polypeptides (TnC, TnI, TnT) that range in mass from 18 to 37 kDa. The complex binds both to Tm and to actin.[255,256] Peptide TnT binds tightly to Tm and is thought to link the TnI•TnC complex to Tm.[256,257] TnI interacts with actin and inhibits ATPase activity in the absence of Ca^{2+}.[258–261] It may work with the other two peptides to keep the Tm in the proper position to inhibit ATP hydrolysis. TnC binds calcium ions. This ~160-residue protein has a folding pattern almost identical to that of calmodulin (Fig. 6-8) with four Ca^{2+}-binding domains arranged in two pairs at the ends of a long 9-turn helix. When Ca^{2+} binds to TnC, a conformational change occurs[258,259,262-265] (p. 313). This induces changes in the troponin•tropomyosin•actin complex, releases the inhibition of actomyosin ATPase, and allows contraction to occur.[265a] In the heart additional effects are exerted by β-adrenergic stimulation, which induces phosphorylation of two sites on TnI by the action of the cAMP-dependent protein kinase PKA. Dephosphorylation by protein phosphatase 2A completes a regulatory cycle in which the doubly phosphorylated TnI has a decreased sensitivity to calcium ions.[266] Cardiac muscle also contains a specialized protein called **phospholamban**. An oligomer of 52-residue subunits, it controls the calcium ion pump in response to β-adrenergic stimulation. Unphosphorylated phospholamban inhibits the Ca ATPase, keeping [Ca^{2+}] high in the cytoplasm. Phosphorylation of phospholamban by cAMP and/or calmodulin-dependent protein kinase activates the Ca^{2+} pump,[267–268a] removing Ca^{2+} and ending contraction.

X-ray diffraction and electron microscopy in the 1970s suggested that when calcium binds to troponin the tropomyosin moves through an angle of ~ 20° away from S1, uncovering the active site for the myosin–ATP–actin interactions.[252,253] Tropomyosin could be envisioned as rolling along the surface of the actin, uncovering sites on seven actin molecules at once. Side-chain knobs protruding from the tropomyosin like teeth on a submicroscopic gear might engage complementary holes in the actin. At the same time a set of magnesium ion bridges between zones of negative charge on tropomyosin and actin could hold the two proteins together. This proposal has been difficult to test. Although the older image reconstruction is regarded as unreliable, recent work still supports this **steric blocking** model.[255,269-270c] Image reconstruction from electron micrographs of thin filaments shows that, in the presence of Ca^{2+}, the tropomyosin does move 25° away from the position in low [Ca^{2+}]. However, instead of two states of the thin filament, "on" and "off," there may be at least three, which have been called "blocked," "closed," and "open."[255,269,271,271a] The closed state may be attained in rigor.[269] In addition, the possibility that changes in the conformation of actin as well as of myosin occur during the contraction cycle must be considered.[255]

Smooth muscle. The primary regulation of smooth muscle contraction occurs via phosphorylation of the Ser 19 –OH group in the 20-kDa regulatory light chains of each myosin head.[121b,160,272–274] The phosphorylated form is active, participating in the contraction process. Removal of Ca^{2+} by the calcium pump and dephosphorylation of the light chains by a protein phosphatase[275] restores the muscle to a resting state. The N-terminal part of the myosin light chain kinase binds to actin, while the catalytic domain is in the center of the protein. The C-terminal part binds to myosin, and this binding also has an activating effect.[276]

Another protein, **caldesmon**, binds to smooth muscle actin and blocks actomyosin ATPase.[271,277–278a] It is present in smooth muscle in a ratio of actin:tropomyosin:caldesmon of ~14:2:1. Inhibition can be reversed by Ca^{2+}, but there is no agreement on the function of caldesmon.[277] It is an elongated ~756-residue protein with N-terminal domain, which binds to myosin, and a C-terminal domain, which binds to actin, separated by a long helix.[278] Caldesmon may be a substitute for troponin in a tropomyosin-type regulatory system, or it may promote actomyosin assembly. Another possibility is that it functions in a **latch state**, an energy-economic state of smooth muscle at low levels of ATP hydrolysis.[278,279] In molluscan muscles Ca^{2+} binds to a myosin light chain and activates contraction directly. Some molluscan smooth muscles (**catch muscles**) also have a latch state, which enables these animals to maintain muscular tension for long periods of time, e.g., holding their shells tightly closed, with little expenditure of energy.[279a] Catch muscles contain myosin plus a second protein, **catchin**, which is formed as a result of alternative mRNA splicing. Catchin contains an N-terminal sequence that may undergo phosphorylation as part of a regulatory mechanism.[280] However, recent experiments indicate that twitchin (see next paragraph), rather than catchin, is essential to the catch state and is regulated by phosphorylation.[280a] Regulation of the large groups of unconventional myosins is poorly understood. Phosphorylation of groups on the myosin heavy chains is involved in ameba myosins and others.[281]

An unexpected aspect of regulation was discovered from study of the 40 or more genes of *Caenorhabditis elegans* needed for assembly and function of muscle. The mutants designated *unc-22* showed a constant twitch arising from the muscles in the nematode's body. The gene was cloned using transposon tagging (Chapter 27) and was found to encode a mammoth 753-kDa 6839-residue protein which has been named **twitchin**.[282–285] Twitchin resembles titin (Fig. 19-8) and like titin has a protein kinase domain, which is normally inhibited by the end of its peptide chain, which folds over the active site of the kinase. Perhaps the protein kinase activities of twitchin, titin, and related proteins[285] are required in assembly of the sarcomere.

5. The Phosphagens

ATP provides the immediate source of energy for muscles but its concentration is only ~5 mM. As discussed in Chapter 6, Section D, **phosphagens**, such as **creatine phosphate**, are also present and may

$$\text{Creatine-}P + \text{ADP} \longrightarrow \text{Creatine} + \text{ATP}$$

attain a concentration of 20 mM in mammalian muscle. This provides a reserve of high-energy phospho groups and keeps the adenylate system buffered at a high phosphorylation state ratio[286] (see Eq. 6-67).

The concentration of both ATP and creatine phosphate as well as their rates of interconversion can be monitored by ^{31}P NMR within living muscles (Figs. 6-4 and 19-22). Phospho groups were observed to be

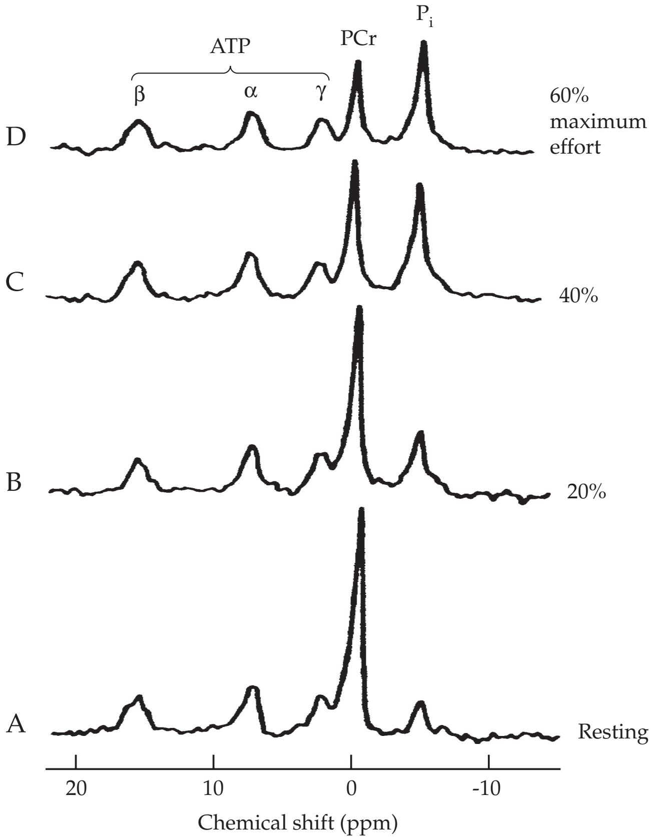


Figure 19-22 Phosphorus-31 magnetic resonance spectra of wrist flexor muscles of the forearm of a trained long-distance runner at rest and during contraction at three different levels of exercise. Ergometer measurements indicating the percent of initial maximum strength (% max) were recorded over each 6-min period. Spectra were obtained during the last 3 min of each period. Times of spectral data collection: A, resting; B, 4–6 min; C, 10–12 min; and D, 16–18 min. The pH ranged from 6.9 to 7.0. From Park *et al*.[288]

transferred from creatine phosphate to ADP to form ATP with a flux of 13 mmol $kg^{-1}s^{-1}$ in rat legs.[287] The reverse reaction must occur at about the same rate because little cleavage of ATP to P_i was observed in the anaesthetized rats. Use of surface coils has permitted direct observation of the operation of this shuttle system in human muscle (Fig. 19-22)[288] as well as in animal hearts (see Chapter 6). Only a fraction of the total creatine present within cells participates in the shuttle, however.[289]

C. Motion in Nonmuscle Cells

At one time actin and myosin were thought to be present only in muscles, but we know now that both actin proteins of the myosin family are present in all eukaryotic cells. Ameboid movement, the motion of cilia and flagella, and movement of materials along microtubules within cells also depend upon proteins of this group.

1. Actin-Based Motility

Ameboid movements of protozoa and of cells from higher organisms, the ruffling movements of cell membranes, phagocystosis, and the cytoplasmic streaming characteristic of many plant cells[289a] have all been traced to actin filaments or actin cables rather

TABLE 19-1
Some Actin-Binding Proteins

Function	Name
Bind and stabilize monomeric actin	Profilin[a,b,c,d,e,f,g,h]
	ADF/Cofilin[i,j,h,k]
	Thymosin[f,l]
Cap actin filament ends	CapZ[m,n]
barbed end	Fragmin[o,p]
pointed end	β-Actin[q]
	Tropomodulin[r]
	Arp2/3, a complex of seven polypeptides
Sever or dissociate actin filament	Gelsolin[s,t,u,v,w]
	Depactin
	Profilin[d,e]
	ADF/cofilin[h,k,x,y]
Crosslink actin filaments or monomers	
Tight bundles	Villin[c,z,aa]
Loose bundles	α-Actinin[bb]
Spectrin[bb]	
	Fascin[cc]
	MARCKS[a]
Network	Filamin[bb,c]
	Gelactins
Bind actin filaments to membrane	Talin[dd]
	"ERM" proteins[ee,ff]

[a] Aderem, A. (1992) *Trends Biochem. Sci.* **17**, 438–443
[b] Mannherz, H. G. (1992) *J. Biol. Chem.* **267**, 11661–11664
[c] Way, M., and Weeds, A. (1990) *Nature (London)* **344**, 292–293
[d] Eads, J. C., Mahoney, N. M., Vorobiev, S., Bresnick, A. R., Wen, K.-K., Rubenstein, P. A., Haarer, B. K., and Almo, S. C. (1998) *Biochemistry* **37**, 11171–11181
[e] Vinson, V. K., De La Cruz, E. M., Higgs, H. N., and Pollard, T. D. (1998) *Biochemistry* **37**, 10871–10880
[f] Kang, F., Purich, D. L., and Southwick, F. S. (1999) *J. Biol. Chem.* **274**, 36963–36972
[g] Gutsche-Perelroizen, I., Lepault, J., Ott, A., and Carlier, M.-F. (1999) *J. Biol. Chem.* **274**, 6234–6243
[h] Nodelman, I. M., Bowman, G. D., Lindberg, U., and Schutt, C. E. (1999) *J. Mol. Biol.* **294**, 1271–1285
[i] Lappalainen, P., Fedorov, E. V., Fedorov, A. A., Almo, S. C., and Drubin, D. G. (1997) *EMBO J.* **16**, 5520–5530
[j] Rosenblatt, J., and Mitchison, T. J. (1998) *Nature (London)* **393**, 739–740
[k] Chen, H., Bernstein, B. W., and Bamburg, J. R. (2000) *Trends Biochem. Sci.* **25**, 19–23
[l] Carlier, M.-F., Didry, D., Erk, I., Lepault, J., Van Troys, M. L., Vandekerckhove, J., Perelroizen, I., Yin, H., Doi, Y., and Pantaloni, D. (1996) *J. Biol. Chem.* **271**, 9231–9239
[m] Barron-Casella, E. A., Torres, M. A., Scherer, S. W., Heng, H. H. Q., Tsui, L.-C., and Casella, J. F. (1995) *J. Biol. Chem.* **270**, 21472–21479
[n] Kuhlman, P. A., and Fowler, V. M. (1997) *Biochemistry* **36**, 13461–13472
[o] Steinbacher, S., Hof, P., Eichinger, L., Schleicher, M., Gettemans, J., Vandekerckhove, J., Huber, R., and Benz, J. (1999) *EMBO J.* **18**, 2923–2929
[p] Khaitlina, S., and Hinssen, H. (1997) *Biophys. J.* **73**, 929–937
[q] See main text
[r] Gregorio, C. C., Weber, A., Bondad, M., Pennise, C. R., and Fowler, V. M. (1995) *Nature (London)* **377**, 83–86
[s] Azuma, T., Witke, W., Stossel, T. P., Hartwig, J. H., and Kwiatkowski, D. J. (1998) *EMBO J.* **17**, 1362–1370
[t] De Corte, V., Demol, H., Goethals, M., Van Damme, J., Gettemans, J., and Vandekerckhove, J. (1999) *Protein Sci.* **8**, 234–241
[u] McGough, A., Chiu, W., and Way, M. (1998) *Biophys. J.* **74**, 764–772
[v] Sun, H. Q., Yamamoto, M., Mejillano, M., and Yin, H. L. (1999) *J. Biol. Chem.* **274**, 33179–33182
[w] Robinson, R. C., Mejillano, M., Le, V. P., Burtnick, L. D., Yin, H. L., and Choe, S. (1999) *Science* **286**, 1939–1942
[x] Carlier, M.-F., Ressad, F., and Pantaloni, D. (1999) *J. Biol. Chem.* **274**, 33827–33830
[y] McGough, A., and Chiu, W. (1999) *J. Mol. Biol.* **291**, 513–519
[z] Markus, M. A., Matsudaira, P., and Wagner, G. (1997) *Protein Sci.* **6**, 1197–1209
[aa] Vardar, D., Buckley, D. A., Frank, B. S., and McKnight, C. J. (1999) *J. Mol. Biol.* **294**, 1299–1310
[bb] Matsudaira, P. (1991) *Trends Biochem. Sci.* **16**, 87–92
[cc] Ono, S., Yamakita, Y., Yamashiro, S., Matsudaira, P. T., Gnarra, J. R., Obinata, T., and Matsumura, F. (1997) *J. Biol. Chem.* **272**, 2527–2533
[dd] McLachlan, A. D., Stewart, M., Hynes, R. O., and Rees, D. J. G. (1994) *J. Mol. Biol.* **235**, 1278–1290
[ee] Tsukita, S., Yonemura, S., and Tsukita, S. (1997) *Trends Biochem. Sci.* **22**, 53–58
[ff] Tsukita, S., and Yonemura, S. (1999) *J. Biol. Chem.* **274**, 34507–34510

than to microtubules.[289b] Actin is one of the most abundant proteins in all eukaryotes. Its network of filaments is especially dense in the lamellipodia of cell edges, in microvilli, and in the specialized stereocilia and acrosomal processes (see also pp. 369-370).[289c] Actin filaments and cables are often formed rapidly and dissolve quickly. When actin filaments grow, the monomeric subunits with bound ATP are added most rapidly at the "barbed end" and dissociate from the filament at the "pointed end" (see Section B,2).[94,290] The rate of growth may be ~20–200 nm / s, which requires the addition of 10–100 subunits / s.[291] Various actin-binding proteins control the growth and stability of the filaments. The actin-related proteins **Arp2** and **Arp3**, as a complex Arp2/3, together with recently recognized **formins**.[291a], provide nuclei for rapid growth of new actin filaments as branches near the barbed ends.[290,292–293c]

Growth of the barbed ends of actin filaments is stimulated by phosphoinositides and by members of the Rho family of G proteins (p. 559)[293d] through interaction with proteins of the **WASp** group.[293b,d,e] The name WAS comes from the immune deficiency disorder Wiskott–Aldrich Syndrome. Yet another family, the **Ena/VASP** proteins, are also implicated in actin dynamics. They tend to localize at focal adhesions and edges of lamellipodia.[293e,f] Profilin (Table 19-1) stabilizes a pool of monomeric actin when the barbed ends of actin filaments are capped. However, it catalyzes both the addition of actin monomers to uncapped barbed ends and rapid dissociation of subunits from pointed ends, leading to increased **treadmilling**.[294,295] Actin-severing proteins such as the **actin depolymerizing factor** (**ADF** or **cofilin**, Table 19-1) promote breakdown of the filaments.[296–297a] Treadmilling in the actin filaments of the lamellipods of crawling cells or pseudopods of amebas provides a motive force for many cells[291,298–299] ranging from those of *Dictyostelium* to human leukocytes. A series of proteins known as the **ezrin**, **radixin**, **moesin** (**ERM**) group attach actin to integral membrane proteins (Fig. 8-17)[292,300,301] and may interact directly with membrane lipids.[301a,b] Bound ATP in the actin subunits is essential for polymerization, and excess ATP together with crosslinking proteins stabilize the filaments. However, when the bound ATP near the pointed ends is hydrolyzed to ADP the filaments become unstable and treadmilling is enhanced. Thus, as in skeletal muscle, ATP provides the energy for movement.

Bacteria also contain filamentous proteins that resemble F-actin and which may be utilized for cell-shape determination.[301c] Actin-based motility is used by some bacteria and other pathogens during invasion of host cells (Box 19-C). It is employed by sperm cells of *Ascaris* and of *C. elegans*, which crawl by an ameboid movement that utilizes treadmilling of filaments formed from a motile sperm protein, which does not resemble actin.[302,302a] Cells are propelled on a glass surface at rates up to ~1 µm / s.

Various nonmuscle forms of myosin also interact with actin without formation of the myofibrils of muscle.[299] In most higher organisms nonmuscle myosins often consist of two ~200-kDa subunits plus two pairs of light chains of ~17 and 24 kDa each. These may form bipolar aggregates, which may bind to pairs of actin filaments to cause relative movement of two parts of a cell.[303] Movement depending upon the cytoskeleton is complicated by the presence of a bewildering array of actin-binding proteins, some of which are listed in Table 19-1.

2. Transport along Microtubules by Kinesin and Dynein

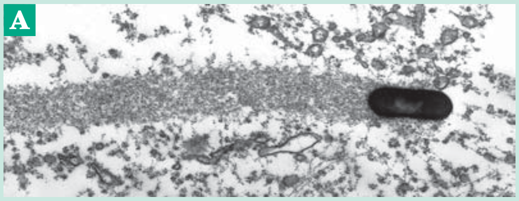
Many materials are carried out from the cell bodies of neurons along microtubules in the axons, which in the human body may be as long as 1 m. The rates of this **fast axonal transport** in neurons may be as high as 5 µm / s or 0.43 m / day. The system depends upon ATP and kinesin (Fig. 19-17) and permits small vesicles to be moved along single microtubules.[304–305b] Movement is from the minus end toward the plus end of the microtubule as defined in Figs. 7-33, 7-34. Slower **retrograde axonal transport** carries vesicles from the synapses at the ends of the axons (Fig. 30-8) back toward the cell body. This retrograde transport depends upon the complex motor molecule **cytoplasmic dynein** which moves materials from the plus end of the microtubule toward the minus end.[305,305c] In addition to these movements, as mentioned in Chapter 7, microtubules often grow in length rapidly or dissociate into their tubulin subunits. Growth occurs at one end by addition of tubulin subunits with their bound GTP. The fast growing *plus*-ends of the microtubules are usually oriented toward the cell periphery, while the *minus*-ends are embedded in the **centrosome** or **microtubule-organizing center** (p. 372).[306] Just as with actin, in which bound ATP is hydrolyzed to ADP, the bound GTP in the β-tubulin subunits of microtubules is hydrolyzed to GDP[307–310] decreasing the stability of the microtubules, a phenomenon described as **dynamic instability**. Various **microtubule-associated proteins** (MAPs) have strong effects on this phenomenon.[311] The MAPs are often regulated by phosphorylation–dephosphorylation cycles involving serine / threonine kinases. Microtubules also undergo posttranslational alterations not seen in other proteins. These include addition or removal of tyrosine at the C terminus.[312] Polyglycyl groups containing 3–34 glycine residues may be bound covalently to γ-carboxyl groups of glutamate side chains in both α- and β-tubulins.[312,313] This stabilizes the microtubules and is important to the long-lived microtubules of the axonomes of flagella and

BOX 19-C ACTIN-BASED MOTILITY AND BACTERIAL INVASION

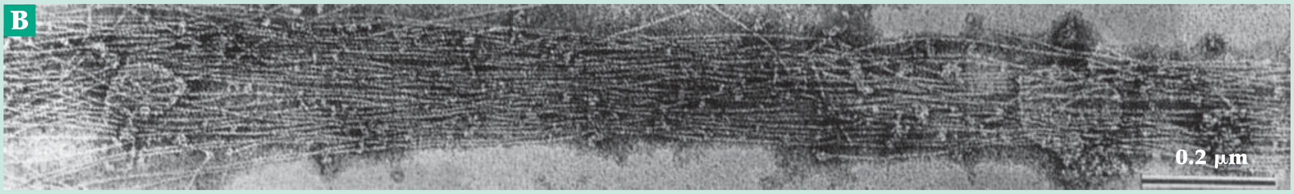
Listeria monocytogenes is a dangerous food-borne bacterium that has become a major problem in the United States. This is one of the best understood *intracellular* pathogens. It is able to enter cells, escape from phagocytic vesicles, spread from cell to cell, and cross intestinal, blood–brain, and placental barriers.[a–c] Within cells these bacteria move using actin-based motility. Actin subunits polymerize at one end of a bacterium leaving a “comet tail” of crosslinked fibrous actin behind (see micrographs). Actin polymerization occurs directly behind the bacteria with subunits of monomeric actin adding to the fast growing “barbed end” (see Section B,2) of the actin strands. Growth has been described as a “Brownian ratchet.”[c,d] Continual Brownian movement opens up spaces behind the bacteria, spaces that are immediately filled by new actin subunits. This provides a propulsive force adequate to move the bacteria ahead at velocities of about 0.3 μm / s.

Polymerization of actin is induced by interaction of a dimer of a 610-residue bacterial protein **ActA** with proteins of the host cell.[a,e–h] ActA is a composite protein with an N-terminal region that protrudes from the bacterial cell, a central region of proline-rich repeats that appear to be essential for recognition by host cells, and a C-terminal hydrophobic membrane anchor. There are also regions of close sequence similarity to the human actin-binding proteins vinculin and zyxin. The number of host proteins needed in addition to monomeric actin are:[i,j] the two actin-related proteins, **Arp2** and **Arp3**, which stimulate actin polymerization and branching;[h] **ADF/ cofilin**, which increases the rates, both of growth at the barbed ends and dissociation from the pointed ends of the filaments; and **Cap Z**, which caps barbed ends (Table 19-1). The need for ADF/ cofilin and Cap Z seems paradoxical. Cap Z may cap mostly older and slower growing filaments, restricting rapid filament assembly to the region closest to the bacterium. The need for ADF/ cofilin is unclear.[i] Growth rates are also enhanced by the human protein called **VASP** (vasodilator-stimulated phosphoprotein). The proline-rich region of the bacterial ActA may bind to VASP to initiate polymerization.[g] Both profilin (Table 19-1) and the crosslinking protein α-actinin also stimulate comet tail growth. Myosin does not participate in actin-based motility, but the hydrolysis of ATP drives the process through its link to actin polymerization.[i]

Although *Listeria* has been studied most, actin-based motility is employed by other pathogens as well, e.g., *Shigella flexneri* (the dysentery bacterium),[k] *Rickettsia*,[l] and vaccinia virus.[l] Although enteropathogenic *E. coli* do not use this method of movement, they induce accumulation of actin beneath the bacteria. They also promote formation of actin-rich adherent pseudopods and highly organized cytoskeletal structures that presumably assist the bacteria in entering a cell.[m]



(A) *Listeria* cell with “comet tail” of cross-linked actin filaments. From Kocks *et al*. (1992) *Cell* **68**, 521–531. Courtesy of Pascale Cossart.



(B) Enlarged section of a thin comet tail of high resolution showing the actin filaments. From Sechi *et al*.[b] Courtesy of Antonio Sechi.

---

[a] Cossart, P., and Lecuit, M. (1998) *EMBO J.* **17**, 3797–3806

[b] Sechi, A. S., Wehland, J., and Small, J. V. (1997) *J. Cell Biol.* **137**, 155–167

[c] Pantaloni, D., Le Clainche, C., and Carlier, M.-F. (2001) *Science* **292**, 1502–1506

[d] Mogilner, A., and Oster, G. (1996) *Biophys. J.* **71**, 3030–3045

[e] Mourrain, P., Lasa, I., Gautreau, A., Gouin, E., Pugsley, A., and Cossart, P. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10034–10039

[f] Southwick, F. S., and Purich, D. L. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5168–5172

[g] Niebuhr, K., Ebel, F., Frank, R., Reinhard, M., Domann, E., Carl, U. D., Walter, U., Gertler, F. B., Wehland, J., and Chakraborty, T. (1997) *EMBO J.* **16**, 5433–5444

[h] Welch, M. D., Rosenblatt, J., Skoble, J., Portnoy, D. A., and Mitchison, T. J. (1998) *Science* **281**, 105–108

[i] Loisel, T. P., Boujemaa, R., Pantaloni, D., and Carlier, M.-F. (1999) *Nature (London)* **401**, 613–616

[j] Machesky, L. M., and Cooper, J. A. (1999) *Nature (London)* **401**, 542–543

[k] Bourdet-Sicard, R., Rüdiger, M., Jockusch, B. M., Gounon, P., Sansonetti, P. J., and Tran Van Nhieu, G. (1999) *EMBO J.* **18**, 5853–5862

[l] Cudmore, S., Cossart, P., Griffiths, G., and Way, M. (1995) *Nature (London)* **378**, 636–638

[m] Rosenshine, I., Ruschkowski, S., Stein, M., Reinscheid, D. J., Mills, S. D., and Finlay, B. B. (1996) *EMBO J.* **15**, 2613–2624

cilia. Polyglutamyl groups of 6–7 glutamates are also often added.[314]

Both dynein and several kinesins act as motors for formation of the spindle and for movement of chromosomes toward the minus ends of spindle microtubules during mitosis and meiosis (Fig. 26-11).[314a] In the genome of *Saccharomyces cerevisiae* there is only one dynein gene, but genes for six different kinesin-type motor molecules are present.[315] In higher organisms there may be even more genes for kinesins but there is apparently only one dynein in most species.[316] **Axonemal dyneins** drive the motion of eukaryotic flagella and cilia. As with the cytoplasmic dyneins a complete molecule consists of two or three heavy chains with molecular mass ~520 kDa, some localized in the dynein tail, and several lighter chains.[305a,317–321] Like myosin dynein is an ATPase.

3. Eukaryotic Cilia and Flagella

The motion of eukaryotic flagella (Fig. 1-8) involves a sliding of the microtubular filaments somewhat analogous to the sliding of muscle filaments.[305,322–325] Sliding between the outer doublet microtubules (Fig. 19-23) via their inner and outer arms (dynein compounds) is thought to provide the characteristic bending waves.[325a,b] The movement is powered by dynein and ATP hydrolysis. Force and displacement measurements made by optical trapping nanometry suggest that the characteristic rhythmic beating of flagella results from an oscillatory property of the dynein.[326] The extremely complex structure of flagella is illustrated in Fig. 19-23. About 250 individual axonemal proteins have been detected in flagella of the alga *Chlamydomonas* (Fig. 1-11),[327] and a large number of mutants with various defects in their flagella have been isolated. The radial spokes (Fig. 19-23) alone contain 17 different proteins. These spokes protrude at ~29-nm intervals while the dynein molecules lie between pairs of the outer microtubule doublets at ~24-nm intervals. The dynein "arms" protrude from the "A" microtubule of each outer doublet and make contact with the incomplete "B" microtubule of the next doublet (Fig. 19-12). Although the shapes of the molecules are quite different, the basic chemistry of the ATPase activity of the dynein–microtubule system resembles that of actomyosin. However, the complexity of the dynein arms,[328]

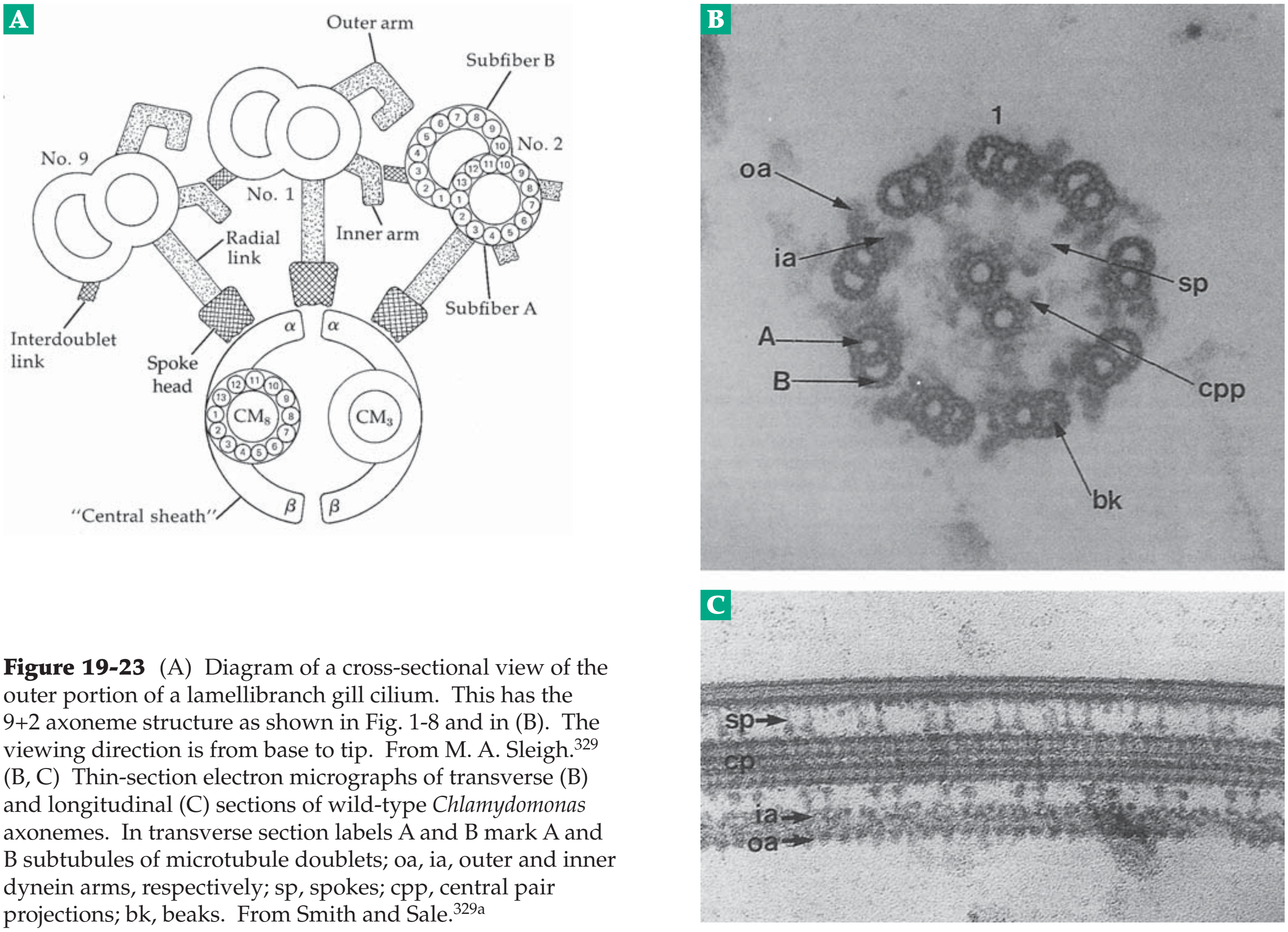


Figure 19-23 (A) Diagram of a cross-sectional view of the outer portion of a lamellibranch gill cilium. This has the 9+2 axoneme structure as shown in Fig. 1-8 and in (B). The viewing direction is from base to tip. From M. A. Sleigh.[329] (B, C) Thin-section electron micrographs of transverse (B) and longitudinal (C) sections of wild-type *Chlamydomonas* axonemes. In transverse section labels A and B mark A and B subtubules of microtubule doublets; oa, ia, outer and inner dynein arms, respectively; sp, spokes; cpp, central pair projections; bk, beaks. From Smith and Sale.[329a]

which exist in two types, inner and outer, suggests a complex contraction cycle.

4. Chemotaxis

As described in Box 11-C, the ameboid cells of the slime mold *Dictyostelium discoideum* are attracted to nutrients such as folic acid during their growth stage. Later, as the cells undergo developmental changes they become attracted by pulses of cyclic AMP.[330] Occupancy of 7-helix receptors for cAMP on the outer plasma membrane appears to induce methylation of both proteins and phospholipids and a rise in cytosolic Ca^{2+} and changes in the cytoskeleton that result in preferential extension of the actin-rich pseudopods toward the chemoattractant.[331]

In a similar manner, human ameboid leukocytes are attracted to sites of inflammation by various **chemotactic factors**.[332] These include the 74-residue cleavage product C5a formed from the fifth component of complement (Chapter 31),[333] various **lymphokines** (Chapter 31) secreted by lymphocytes, and peptides such as VGSE and AGSE, as well as larger peptides released by mast cells, basophils, or stimulated monocytes[334] and oxoicosenoids.[335] Polymorphonuclear leukocytes, upon engulfing sodium urate crystals in gouty joints, release an 8.4-kDa chemotactic protein which may cause a damaging response in this arthritic condition. Leukotriene B is a potent chemotactic agent as are a series of specific bacterial products, formylated peptides such as *N*-formyl-MLF.[332,336]

Neutrophils, monocytes, macrophages, eosinophils, basophils, and polymorphonuclear leukocytes are all affected by several or all of these factors. Binding to specific receptors results in a variety of changes in the cells. These include alterations in membrane potential, cyclic nucleotide levels, and ion fluxes (Na^+, K^+, Ca^{2+}) as well as increased methylation of specific proteins. A reorganization of microtubules and actin fibrils occurs, probably in response to an altered gradient of Ca^{2+}. The morphology of the cells changes, and they begin immediately to crawl toward the chemoattractants. It appears that these ameboid cells detect a gradient of attractant concentration between one end of the cell and the other, even though the anticipated difference may amount to only 0.1% of the total.[337,338]

5. Other Forms of Movement

Movement is characteristic of life and is caused not only by motor proteins but by various springs and ratchets which may be energized in a number of ways.[339] A striking example, which any one with a microscope and some fresh pond water can observe, is contraction of the stalk of protozoa of the genus *Vorticella*. Apparently first reported in 1676 by Leeuwenhoek the organism's 2–3 mm-long stalk contracts into a coiled spring (see p. 1 and also p. 281) when the animal is disturbed. Application of calcium ions causes contraction within a few milliseconds to ~40% of the original length. The process reverses slowly after a few seconds. Contraction is caused by a spring-like organ the **spasmoneme**, which is a bundle of short 2 nm-diameter fibrils inside the stalk. The fibrils are thought to be weakly crosslinked and held in the extended state by electrostatic repulsion between the negatively charged rods. Addition of Ca^{2+} neutralizes the charges permitting an entropy-driven collapse of the fibers.[339]

Another remarkable example is extension of the acrosomal process from a sperm cell of the horseshoe crab *Limulus polyphemus* at fertilization. A bundle of actin filaments in a crystalline state lies coiled around the base of the nucleus. At fertilization the bundle uncoils and slides through a tunnel in the nucleus to form a 60 µm-long acrosomal process within a few seconds. The uncoiled bundle is also crystalline. The coiled bundle is apparently overtwisted and an actin crosslinking protein **scruin** mediates the conformational alteration that takes place.[339] A somewhat related process may be involved in contraction of bacteriophage tails (pp. 363, 364)

Some bacteria glide with a twitching movement induced by rapid retraction of pili.[340] Another type of movement involves the pinching off of vesicles, e.g., of clathrin-coated pits (Fig. 8-27). This is a GTP-driven process that requires a mechanoenzyme called **dynamin**.[341,342]

References

1. Elston, T. C., and Oster, G. (1997) *Biophys. J.* **73**, 703–721
2. Magariyama, Y., Sugiyama, S., Muramoto, K., Maekawa, Y., Kawagishi, I., Imae, Y., and Kudo, S. (1994) *Nature (London)* **371**, 752
3. Macnab, R. M. (1987) in *Escherichia coli and Salmonella typhimurium* (Niedhardt, F. C., ed), pp. 70–83, Am. Soc. for Microbiology, Washington, D.C.
4. Berg, H. C. (1975) *Sci. Am.* **233**(Aug), 36–44
5. Shimada, K., Kamiya, R., and Asakura, S. (1975) *Nature (London)* **254**, 332–334
6. Berg, H. C. (1975) *Nature (London)* **254**, 389–392
7. Macnab, R. M. (1985) *Trends Biochem. Sci.* **10**, 185–188
8. Macnab, R. M., and Aizawa, S.-J. (1984) *Annu Rev Biophys Bioeng.* **13**, 51–83
9. Sharp, L. L., Zhou, J., and Blair, D. F. (1995) *Biochemistry* **34**, 9166–9171
10. Morgan, D. G., Owen, C., Melanson, L. A., and DeRosier, D. J. (1995) *J. Mol. Biol.* **249**, 88–110
11. Mimori, Y., Yamashita, I., Murata, K., Fujiyoshi, Y., Yonekura, K., Toyoshima, C., and Namba, K. (1995) *J. Mol. Biol.* **249**, 69–87
12. Yamashita, I., Vonderviszt, F., Mimori, Y., Suzuki, H., Oosawa, K., and Namba, K. (1995) *J. Mol. Biol.* **253**, 547–558

12a. Samatey, F. A., Imada, K., Nagashima, S., Vonderviszt, F., Kumasaka, T., Yamamoto, M., and Namba, K. (2001) *Nature (London)* **410**, 331–337

13. Muramoto, K., Kawagishi, I., Kudo, S., Magariyama, Y., Imae, Y., and Homma, M. (1995) *J. Mol. Biol.* **251**, 50–58
14. DePamphilis, M. L., and Adler, J. (1971) *J. Bacteriol.* **105**, 396–407

References

15. Stallmeyer, M. J. B., Aizawa, S.-I., Macnab, R. M., and DeRosier, D. J. (1989) *J. Mol. Biol.* **205**, 519 – 528
16. DeRosier, D. J. (1998) *Cell* **93**, 17– 20
17. Trachtenberg, S., and DeRosier, D. J. (1992) *J. Mol. Biol.* **226**, 447 – 454
18. Trachtenberg, S., DeRosier, D. J., Zemlin, F., and Beckmann, E. (1998) *J. Mol. Biol.* **276**, 759 – 773
19. Hasegawa, K., Yamashita, I., and Namba, K. (1998) *Biophys. J.* **74**, 569 – 575

19a. Macnab, R. M. (2001) *Nature (London)* **410**, 321 – 322

20. Iino, T. (1969) *Bacteriol. Rev.* **33**, 454 – 475
21. Kubori, T., Shimamoto, N., Yamaguchi, S., Namba, K., and Aizawa, S.-I. (1992) *J. Mol. Biol.* **226**, 433 – 446
22. Vonderviszt, F., Závodszky, P., Ishimura, M., Uedaira, H., and Namba, K. (1995) *J. Mol. Biol.* **251**, 520 – 532
23. Muramoto, K., Makishima, S., Aizawa, S.-I., and Macnab, R. M. (1998) *J. Mol. Biol.* **277**, 871 – 882
24. Ikeda, T., Oosawa, K., and Hotani, H. (1996) *J. Mol. Biol.* **259**, 679 – 686
25. Fahrner, K. A., Block, S. M., Krishnaswamy, S., Parkinson, J. S., and Berg, H. C. (1994) *J. Mol. Biol.* **238**, 173 – 186

25a. Hirano, T., Minamino, T., and Macnab, R. M. (2001) *J. Mol. Biol.* **312**, 359 – 369

26. Maki, S., Vonderviszt, F., Furukawa, Y., Imada, K., and Namba, K. (1998) *J. Mol. Biol.* **277**, 771 – 777

26a. Yonekura, K., Maki, S., Morgan, D. G., DeRosier, D. J., Vonderviszt, F., Imada, K., and Namba, K. (2000) *Science* **290**, 2148 – 2152

26b. Macnab, R. M. (2000) *Science* **290**, 2086 – 2087

26c. Minamino, T., Tame, J. R. H., Namba, K., and Macnab, R. M. (2001) *J. Mol. Biol.* **312**, 1027 – 1036

26d. Auvray, F., Thomas, J., Fraser, G. M., and Hughes, C. (2001) *J. Mol. Biol.* **308**, 221 – 229

27. Berg, H. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14225 – 14228
28. Lloyd, S. A., Whitby, F. G., Blair, D. F., and Hill, C. P. (1999) *Nature (London)* **400**, 472 – 475
29. Schuster, S. C., and Khan, S. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 509 – 539
30. Adler, J. (1976) *Sci. Am.* **234** (Apr), 40 – 47
31. Calladine, C. R. (1975) *Nature (London)* **255**, 121 – 124
32. Eisenbach, M., and Adler, J. (1981) *J. Biol. Chem.* **256**, 8807 – 8814
33. Fung, D. C., and Berg, H. C. (1995) *Nature (London)* **375**, 809 – 812
34. Glagolev, A. N., and Skulachev, V. P. (1978) *Nature (London)* **273**, 280 – 282
35. Chun, S. Y., and Parkinson, J. S. (1988) *Science* **239**, 276 – 278
36. Laüger, P. (1990) *Commun. Theor. Biol.* **2**, 99 – 123
37. Tang, H., Braun, T. F., and Blair, D. F. (1996) *J. Mol. Biol.* **261**, 209 – 221

37a. Ko, M., and Park, C. (2000) *J. Mol. Biol.* **303**, 371 – 382

38. Block, S. M., and Berg, H. C. (1984) *Nature (London)* **309**, 470 – 472
39. Ueno, T., Oosawa, K., and Aizawa, S.-I. (1994) *J. Mol. Biol.* **236**, 546 – 555
40. Zhou, J., Fazzio, R. T., and Blair, D. F. (1995) *J. Mol. Biol.* **251**, 237 – 242
41. Garza, A. G., Biran, R., Wohlschlegel, J. A., and Manson, M. D. (1996) *J. Mol. Biol.* **258**, 270 – 285

41a. Van Way, S. M., Hosking, E. R., Braun, T. F., and Manson, M. D. (2000) *J. Mol. Biol.* **297**, 7 – 24

42. Zhao, R., Pathak, N., Jaffe, H., Reese, T. S., and Khan, S. (1996) *J. Mol. Biol.* **261**, 195 – 208
43. Marykwas, D. L., Schmidt, S. A., and Berg, H. C. (1996) *J. Mol. Biol.* **256**, 564 – 576
44. Zhou, J., Lloyd, S. A., and Blair, D. F. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6436 – 6441

44a. Thomas, D. R., Morgan, D. G., and DeRosier, D. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10134 – 10139

44b. Lux, R., Kar, N., and Khan, S. (2000) *J. Mol. Biol.* **298**, 577 – 583

45. Toker, A. S., and Macnab, R. M. (1997) *J. Mol. Biol.* **273**, 623 – 634
46. Kehry, M. R., Doak, T. G., and Dahlquist, F. W. (1984) *J. Biol. Chem.* **259**, 11828 – 11835
47. Koshland, D. E., Jr. (1988) *Biochemistry* **27**, 5829 – 5834
48. Stock, J., and Stock, A. (1987) *Trends Biochem. Sci.* **12**, 371 – 375
49. Segall, J. E., Block, S. M., and Berg, H. C. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8987 – 8991
50. Jeffery, C. J., and Koshland, D. E., Jr. (1993) *Protein Sci.* **2**, 559 – 566
51. Li, J., Li, G., and Weis, R. M. (1997) *Biochemistry* **36**, 11851 – 11857

51a. Isaac, B., Gallagher, G. J., Balazs, Y. S., and Thompson, L. K. (2002) *Biochemistry* **41**, 3025 – 3036

52. Scott, W. G., Milligan, D. L., Milburn, M. V., Privé, G. G., Yeh, J., Koshland, D. E., Jr., and Kim, S.-H. (1993) *J. Mol. Biol.* **232**, 555 – 573
53. Yeh, J. I., Biemann, H.-P., Privé, G. G., Pandit, J., Koshland, D. E., Jr., and Kim, S.-H. (1996) *J. Mol. Biol.* **262**, 186 – 201
54. Foster, D. L., Mowbray, S. L., Jap, B. K., and Koshland, D. E., Jr. (1985) *J. Biol. Chem.* **260**, 11706 – 11710

54a. Falke, J. J., and Hazelbauer, G. L. (2001) *Trends Biochem. Sci.* **26**, 257 – 265

55. Maddock, J. R., and Shapiro, L. (1993) *Science* **259**, 1717 – 1723
56. Wolfe, A. J., Conley, M. P., and Berg, H. C. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6711 – 6715
57. Ninfa, E. G., Stock, A., Mowbray, S., and Stock, J. (1991) *J. Biol. Chem.* **266**, 9764 – 9770
58. Kim, S.-H. (1994) *Protein Sci.* **3**, 159 – 165
59. Chervitz, S. A., and Falke, J. J. (1995) *J. Biol. Chem.* **270**, 24043 – 24053
60. Ottemann, K. M., Thorgeirsson, T. E., Kolodziej, A. F., Shin, Y.-K., and Koshland, D. E., Jr. (1998) *Biochemistry* **37**, 7062 – 7069
61. Trammell, M. A., and Falke, J. J. (1999) *Biochemistry* **38**, 329 – 336

61a. Hirschman, A., Boukhvalova, M., VanBruggen, R., Wolfe, A. J., and Stewart, R. C. (2001) *Biochemistry* **40**, 13876 – 13887

62. Spiro, P. A., Parkinson, J. S., and Othmer, H. G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7263 – 7268
63. Volz, K. (1993) *Biochemistry* **32**, 11741 – 11753
64. Bellsolell, L., Cronet, P., Majolero, M., Serrano, L., and Coll, M. (1996) *J. Mol. Biol.* **257**, 116 – 128
65. Jiang, M., Bourret, R. B., Simon, M. I., and Volz, K. (1997) *J. Biol. Chem.* **272**, 11850 – 11855

65a. Halkides, C. J., McEvoy, M. M., Casper, E., Matsumura, P., Volz, K., and Dahlquist, F. W. (2000) *Biochemistry* **39**, 5280 – 5286

65b. Lee, S.-Y., Cho, H. S., Pelton, J. G., Yan, D., Berry, E. A., and Wemmer, D. E. (2001) *J. Biol. Chem.* **276**, 16425 – 16431

65c. Cho, H. S., Lee, S.-Y., Yan, D., Pan, X., Parkinson, J. S., Kustu, S., Wemmer, D. E., and Pelton, J. G. (2000) *J. Mol. Biol.* **297**, 543 – 551

65d. Kim, C., Jackson, M., Lux, R., and Khan, S. (2001) *J. Mol. Biol.* **307**, 119 – 135

66. Bren, A., and Eisenbach, M. (1998) *J. Mol. Biol.* **278**, 507 – 514

66a. Bren, A., and Eisenbach, M. (2001) *J. Mol. Biol.* **312**, 699 – 709

67. Sanna, M. G., and Simon, M. I. (1996) *J. Biol. Chem.* **271**, 7357 – 7361
68. Blat, Y., and Eisenbach, M. (1996) *J. Biol. Chem.* **271**, 1226 – 1231
69. Silversmith, R. E., Appleby, J. L., and Bourret, R. B. (1997) *Biochemistry* **36**, 14965 – 14974
70. Weis, R. M., Chasalow, S., and Koshland, D. E., Jr. (1990) *J. Biol. Chem.* **265**, 6817 – 6826
71. Wu, J., Li, J., Li, G., Long, D. G., and Weis, R. M. (1996) *Biochemistry* **35**, 4984 – 4993
72. West, A. H., Martinez-Hackert, E., and Stock, A. M. (1995) *J. Mol. Biol.* **250**, 276 – 290
73. Alon, U., Surette, M. G., Barkai, N., and Leibler, S. (1999) *Nature (London)* **397**, 168 – 171

73a. Bray, D. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7 – 9

74. Prasad, K., Caplan, S. R., and Eisenbach, M. (1998) *J. Mol. Biol.* **280**, 821 – 828
75. Nickitenko, A. V., Trakhanov, S., and Quiocho, F. A. (1995) *Biochemistry* **34**, 16585 – 16595
76. Lux, R., Jahreis, K., Bettenbrock, K., Parkinson, J. S., and Lengeler, J. W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11583 – 11587
77. Gebert, J. F., Overhoff, B., Manson, M. D., and Bods, W. (1988) *J. Biol. Chem.* **263**, 16652 – 16660

77a. Nishiyama, S.-i, Maruyama, I. N., Homma, M., and Kawagishi, I. (1999) *J. Mol. Biol.* **286**, 1275 – 1284

77b. Kim, S.-H., Wang, W., and Kim, K. K. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11611 – 11615

78. Unsworth, B. R., Witzmann, F. A., and Fitts, R. H. (1982) *J. Biol. Chem.* **257**, 15129 – 15136
79. Kelly, A. M., and Rubinstein, N. A. (1980) *Nature (London)* **288**, 267 – 269
80. Bullard, B. (1983) *Trends Biochem. Sci.* **8**, 68 – 70
81. Cohen, C. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3176 – 3178
82. Huxley, H., and Hanson, J. (1954) *Nature (London)* **173**, 973 – 976
83. Huxley, H. E. (1958) *Sci. Am.* **199**, 67 – 82
84. Huxley, H. E. (1990) *J. Biol. Chem.* **265**, 8347 – 8350
85. Glickson, J. D., Phillips, W. D., and Rupley, J. A. (1971) *J. Am. Chem. Soc.* **93**, 4031 – 4038
86. Yates, L. D., and Greaser, M. L. (1983) *J. Mol. Biol.* **168**, 123 – 141
87. Yao, X., Grade, S., Wriggers, W., and Rubenstein, P. A. (1999) *J. Biol. Chem.* **274**, 37443 – 37449
88. Rubenstein, P. A. (1990) *BioEssays* **12**, 309 – 315
89. Allen, P. G., Shuster, C. B., Käs, J., Chaponnier, C., Janmey, P. A., and Herman, I. M. (1996) *Biochemistry* **35**, 14062 – 14069
90. Kabsch, W., and Holmes, K. C. (1995) *FASEB J.* **9**, 167 – 174
91. Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F., and Holmes, K. C. (1990) *Nature (London)* **347**, 37 – 44
92. Chik, J. K., Lindberg, U., and Schutt, C. E. (1996) *J. Mol. Biol.* **263**, 607 – 623

92a. Otterbein, L. R., Graceffa, P., and Dominguez, R. (2001) *Science* **293**, 708 – 711

93. Pollard, T. D., and Craig, S. W. (1982) *Trends Biochem. Sci.* **7**, 55 – 58
94. Teubner, A., and Wegner, A. (1998) *Biochemistry* **37**, 7532 – 7538

94a. Orlova, A., Galkin, V. E., VanLoock, M. S., Kim, E., Shvetsov, A., Reisler, E., and Egelman, E. H. (2001) *J. Mol. Biol.* **312**, 95 – 106

95. Carlier, M.-F., and Pantaloni, D. (1988) *J. Biol. Chem.* **263**, 817 – 825

95a. Sablin, E. P., Dawson, J. F., VanLoock, M. S., Spudich, J. A., Egelman, E. H., and Fletterick, R. J. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10945 – 10947

References

96. Maruta, H., Knoerzer, W., Hinssen, H., and Isenberg, G. (1984) *Nature (London)* **312**, 424–427
97. Pollard, T. D., and Cooper, J. A. (1986) *Ann. Rev. Biochem.* **55**, 987–1035
98. Maruyama, K. (1997) *FASEB J.* **11**, 341–345
98a. Maruyama, K. (2002) *Trends Biochem. Sci.* **27**, 264–266
99. Trinick, J. (1994) *Trends Biochem. Sci.* **19**, 405–409
100. Labeit, S., and Kolmerer, B. (1995) *Science* **270**, 293–296
100a. Amodeo, P., Fraternali, F., Lesk, A. M., and Pastore, A. (2001) *J. Mol. Biol.* **311**, 283–296
101. Kolmerer, B., Olivieri, N., Witt, C. C., Herrmann, B. G., and Labeit, S. (1996) *J. Mol. Biol.* **256**, 556–563
102. Tskhovrebova, L., and Trinick, J. (2001) *J. Mol. Biol.* **310**, 755–771
102a. Ma, K., Kan, L.-s, and Wang, K. (2001) *Biochemistry* **40**, 3427–3438
103. Wang, K., Knipfer, M., Huang, Q.-Q., van Heerden, A., Hsu, L. C.-L., Gutierrez, G., Quian, X.-L., and Stedman, H. (1996) *J. Biol. Chem.* **271**, 4304–4314
104. Kalverda, A. P., Wymenga, S. S., Lomman, A., van de Ven, F. J. M., Hilbers, C. W., and Canters, G. W. (1994) *J. Mol. Biol.* **240**, 358–371
105. Millevoi, S., Trombitas, K., Kolmerer, B., Kostin, S., Schaper, J., Pelin, K., Granzier, H., and Labeit, S. (1998) *J. Mol. Biol.* **282**, 111–123
106. Politou, A. S., Millevoi, S., Gautel, M., Kolmerer, B., and Pastore, A. (1998) *J. Mol. Biol.* **276**, 189–202
106a. McElhinny, A. S., Kolmerer, B., Fowler, V. M., Labeit, S., and Gregorio, C. C. (2001) *J. Biol. Chem.* **276**, 583–592
107. Obermann, W. M. J., Gautel, M., Weber, K., and Fürst, D. O. (1997) *EMBO J.* **16**, 211–220
108. Young, P., Ferguson, C., Banuelos, S., and Gautel, M. (1998) *EMBO J.* **17**, 1614–1624
109. Sorimachi, H., Freiburg, A., Kolmerer, B., Ishiura, S., Stier, G., Gregorio, C. C., Labeit, D., Linke, W. A., Suzuki, K., and Labeit, S. (1997) *J. Mol. Biol.* **270**, 688–695
109a. Tang, J., Taylor, D. W., and Taylor, K. A. (2001) *J. Mol. Biol.* **310**, 845–858
109b. Joseph, C., Stier, G., O'Brien, R., Politou, A. S., Atkinson, R. A., Bianco, A., Ladbury, J. E., Martin, S. R., and Pastore, A. (2001) *Biochemistry* **40**, 4957–4965
110. Isobe, Y., Warner, F. D., and Lemanski, L. F. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6758–6762
111. Deatherage, J. F., Cheng, N., and Bullard, B. (1989) *J. Cell Biol.* **108**, 1775–1782
112. van Straaten, M., Goulding, D., Kolmerer, B., Labeit, S., Clayton, J., Leonard, K., and Bullard, B. (1999) *J. Mol. Biol.* **285**, 1549–1562
112a. Kolmerer, B., Clayton, J., Benes, V., Allen, T., Ferguson, C., Leonard, K., Weber, U., Knekt, M., Ansorge, W., Labeit, S., and Bullard, B. (2000) *J. Mol. Biol.* **296**, 435–448
112b. Fukuzawa, A., Shimamura, J., Takemori, S., Kanzawa, N., Yamaguchi, M., Sun, P., Maruyama, K., and Kimura, S. (2001) *EMBO J.* **20**, 4826–4835
113. Yamaguchi, M., Izumimoto, M., Robson, R. M., and Stromer, M. H. (1985) *J. Mol. Biol.* **184**, 621–644
114. Baron, M. D., Davison, M. D., Jones, P., Patel, B., and Critchley, D. R. (1987) *J. Biol. Chem.* **262**, 2558–2561
115. Pan, K.-M., Roelke, D. L., and Greaser, M. L. (1986) *J. Biol. Chem.* **261**, 9922–9928
116. McLachlan, A. D. (1984) *Ann. Rev. Biophys. Bioeng.* **13**, 167–189
116a. Bellin, R. M., Huiatt, T. W., Critchley, D. R., and Robson, R. M. (2001) *J. Biol. Chem.* **276**, 32330–32337
117. Carragher, B. O., Cheng, N., Wang, Z.-Y., Korn, E. D., Reilein, A., Belnap, D. M., Hammer, J. A., III, and Steven, A. C. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15206–15211
118. Heintzelman, M. B., and Schwartzman, J. D. (1997) *J. Mol. Biol.* **271**, 139–146
119. Baker, J. P., and Titus, M. A. (1997) *J. Mol. Biol.* **272**, 523–535
120. Hasson, T., and Mooseker, M. S. (1996) *J. Biol. Chem.* **271**, 16431–16434
121. Jontes, J. D., and Milligan, R. A. (1997) *J. Mol. Biol.* **266**, 331–342
121a. Ajtai, K., Garamszegi, S. P., Park, S., Dones, A. L. V., and Burghardt, T. P. (2001) *Biochemistry* **40**, 12078–12093
121b. Wendt, T., Taylor, D., Trybus, K. M., and Taylor, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4361–4366
121c. Andersen, J. L., Schjerling, P., and Saltin, B. (2000) *Sci. Am.* **283**(Sep), 48–55
122. Weiss, A., Schiaffino, S., and Leinwand, L. A. (1999) *J. Mol. Biol.* **290**, 61–75
123. Robillard, G., and Schulman, R. G. (1972) *J. Mol. Biol.* **71**, 507–511
123a. Swank, D. M., Bartoo, M. L., Knowles, A. F., Iliffe, C., Bernstein, S. I., Molloy, J. E., and Sparrow, J. C. (2001) *J. Biol. Chem.* **276**, 15117–15124
124. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics* **6**, 13–27
125. Lorenz, M., Poole, K. J. V., Popp, D., Rosenbaum, G., and Holmes, K. C. (1995) *J. Mol. Biol.* **246**, 108–119
126. Knight, P. J. (1996) *J. Mol. Biol.* **255**, 269–274
127. Katoh, T., Konishi, K., and Yazawa, M. (1998) *J. Biol. Chem.* **273**, 11436–11439
128. Offer, G., and Knight, P. (1996) *J. Mol. Biol.* **256**, 407–416
129. Davis, J. S. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 217–239
130. Squire, J. M. (1973) *J. Mol. Biol.* **77**, 291–323
131. Padrón, R., Alamo, L., Murgich, J., and Craig, R. (1998) *J. Mol. Biol.* **275**, 35–41
132. Sohn, R. L., Vikstrom, K. L., Strauss, M., Cohen, C., Szent-Gyorgyi, A. G., and Leinwand, L. A. (1997) *J. Mol. Biol.* **266**, 317–330
133. Bennett, P. M., and Gautel, M. (1996) *J. Mol. Biol.* **259**, 896–903
134. Winkler, H., Reedy, M. C., Reedy, M. K., Tregear, R., and Taylor, K. A. (1996) *J. Mol. Biol.* **264**, 302–322
135. Crowther, R. A., Padron, R., and Craig, R. (1985) *J. Mol. Biol.* **184**, 429–439
135a. Offer, G., Knight, P. J., Burgess, S. A., Alamo, L., and Padrón, R. (2000) *J. Mol. Biol.* **298**, 239–260
135b. Flavigny, J., Souchet, M., Sébillon, P., Berrebi-Bertrand, I., Hainque, B., Mallet, A., Bril, A., Schwartz, K., and Carrier, L. (1999) *J. Mol. Biol.* **294**, 443–456
135c. Witt, C. C., Gerull, B., Davies, M. J., Centner, T., Linke, W. A., and Thierfelder, L. (2001) *J. Biol. Chem.* **276**, 5353–5359
136. Standiford, D. M., Davis, M. B., Miedema, K., Franzini-Armstrong, C., and Emerson, C. P. J. (1997) *J. Mol. Biol.* **265**, 40–55
137. Vibert, P. and Craig, R. (1983) *J. Mol. Biol.* **165**, 303–320
138. Engelhardt, W. A., and Ljubimowa, M. N. (1939) *Nature (London)* **144**, 668–669
139. Szent-Gyorgyi, A. (1941) *Studies Inst. Med. Chem., Univ. Szeged* **1**, 17–26 (reprinted in H. M. Kalckar ed., *Biological Phosphorylations*, pp. 465–472, Prentice-Hall, Englewood Cliffs, New Jersey 1969)
140. Szent-Gyorgyi, A. (1947) *Chemistry of Muscular Contraction*, Academic Press, New York
141. Straub, F. B. (1969) in *Biological Phosphorylation* (Kalckar, H. M., ed), pp. 474–483, Prentice-Hall, Englewood Cliffs, New Jersey (reprinted from Studies Inst. Med. Chem. Univ. Szeged, 2, 3–15)
142. Milligan, R. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 21–26
143. Schmitz, H., Reedy, M. C., Reedy, M. K., Tregear, R. T., Winkler, H., and Taylor, K. A. (1996) *J. Mol. Biol.* **264**, 279–301
144. Katayama, E. (1998) *J. Mol. Biol.* **278**, 349–367
145. Huxley, H. (1969) *Science* **114**, 1356–1366
146. Huxley, H. E. (1998) *Trends Biochem. Sci.* **23**, 84–87
147. Huxley, A. F., and Simmons, R. M. (1969) *Nature (London)* **233**, 533–538
148. Tsukita, S., and Yano, M. (1985) *Nature (London)* **317**, 182–184
149. Sheetz, M. P., and Spudich, J. A. (1983) *Nature (London)* **303**, 31–35
150. Adams, R. J., and Pollard, T. D. (1986) *Nature (London)* **322**, 754–756
151. Harada, Y., Noguchi, A., Kishino, A., and Yanagida, T. (1987) *Nature (London)* **326**, 805–808
152. Rayment, I., Rypniewski, W. R., Schmidt-Bäse, K., Smith, R., Tomchick, D. R., Benning, M. M., Winkelmann, D. A., Wesenberg, G., and Holden, H. M. (1993) *Science* **261**, 50–58
153. Fisher, A. J., Smith, C. A., Thoden, J. B., Smith, R., Sutoh, K., Holden, H. M., and Rayment, I. (1995) *Biochemistry* **34**, 8960–8972
154. Smith, C. A., and Rayment, I. (1995) *Biochemistry* **34**, 8973–8981
155. Fisher, A. J., Smith, C. A., Thoden, J., Smith, R., Sutoh, K., Holden, H. M., and Rayment, I. (1995) *Biophys. J.* **68**, 19s–29s
156. Rayment, I., Holden, H. M., Whittaker, M., Yohn, C. B., Lorenz, M., Holmes, K. C., and Milligan, R. A. (1993) *Science* **261**, 58–65
157. Rayment, I. (1996) *J. Biol. Chem.* **271**, 15850–15853
158. Smith, C. A., and Rayment, I. (1996) *Biochemistry* **35**, 5404–5417
159. Gulick, A. M., Bauer, C. B., Thoden, J. B., and Rayment, I. (1997) *Biochemistry* **36**, 11619–11628
160. Murphy, R. A. (1994) *FASEB J.* **8**, 311–318
161. Xie, X., Harrison, D. H., Schlichting, I., Sweet, R. M., Kalabokis, V. N., Szent-Györgyi, A. G., and Cohen, C. (1994) *Nature (London)* **368**, 306–312
162. Furch, M., Geeves, M. A., and Manstein, D. J. (1998) *Biochemistry* **37**, 6317–6326
163. Van Dijk, J., Fernandez, C., and Chaussepied, P. (1998) *Biochemistry* **37**, 8385–8394
164. Yengo, C. M., Chrin, L., Rovner, A. S., and Berger, C. L. (1999) *Biochemistry* **38**, 14515–14523
165. Grammer, J. C., Kuwayama, H., and Yount, R. G. (1993) *Biochemistry* **32**, 5725–5732
166. Miller, C. J., Wong, W. W., Bobkova, E., Rubenstein, P. A., and Reisler, E. (1996) *Biochemistry* **35**, 16557–16565
166a. Hansen, J. E., Marner, J., Pavlov, D., Rubenstein, P. A., and Reisler, E. (2000) *Biochemistry* **39**, 1792–1799
166b. Bertrand, R., Derancourt, J., and Kassab, R. (2000) *Biochemistry* **39**, 14626–14637
166c. Sasaki, N., Ohkura, R., and Sutoh, K. (2000) *J. Biol. Chem.* **275**, 38705–38709
167. Orlova, A., Chen, X., Rubenstein, P. A., and Egelman, E. H. (1997) *J. Mol. Biol.* **271**, 235–243
167a. Prochniewicz, E., and Thomas, D. D. (2001) *Biochemistry* **40**, 13933–13940

168. Belmont, L. D., Orlova, A., Drubin, D. G., and Egelman, E. H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 29–34
168a. Higgins, S. J., and Banting, G., eds. (2000) *Molecular Motors*, Vol. Essays in Biochemistry No. 35, Portland Press, London
168b. Vale, R. D., and Milligan, R. A. (2000) *Science* **288**, 88–95
168c. Verhey, K. J., and Rapoport, T. A. (2001) *Trends Biochem. Sci.* **26**, 545–550
169. Sack, S., Müller, J., Marx, A., Thormählen, M., Mandelkow, E.-M., Brady, S. T., and Mandelkow, E. (1997) *Biochemistry* **36**, 16155–16165
169a. Song, Y.-H., Marx, A., Müller, J., Woehlke, G., Schliwa, M., Krebs, A., Hoenger, A., and Mandelkow, E. (2001) *EMBO J.* **20**, 6213–6225
169b. Gilbert, S. P. (2001) *Nature (London)* **414**, 597–598
170. Stock, M. F., Guerrero, J., Cobb, B., Eggers, C. T., Huang, T.-G., Li, X., and Hackney, D. D. (1999) *J. Biol. Chem.* **274**, 14617–14623
171. Vale, R. D., Funatsu, T., Pierce, D. W., Romberg, L., Harada, Y., and Yanagida, T. (1996) *Nature (London)* **380**, 451–453
172. Kull, F. J., Sablin, E. P., Lau, R., Fletterick, R. J., and Vale, R. D. (1996) *Nature (London)* **380**, 550–555
173. Mandelkow, E., and Johnson, K. A. (1998) *Trends Biochem. Sci.* **23**, 429–433
174. Gulick, A. M., Song, H., Endow, S. A., and Rayment, I. (1998) *Biochemistry* **37**, 1769–1776
174a. Miki, H., Setou, M., Kaneshiro, K., and Hirokawa, N. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7004–7011
174b. Kikkawa, M., Sablin, E. P., Okada, Y., Yajima, H., Fletterick, R. J., and Hirokawa, N. (2001) *Nature (London)* **411**, 439–445
174c. Schliwa, M., and Woehlke, G. (2001) *Nature (London)* **411**, 424–425
174d. Yun, M., Zhang, X., Park, C.-G., Park, H.-W., and Endow, S. A. (2001) *EMBO J.* **20**, 2611–2618
175. Sablin, E. P., Kull, F. J., Cooke, R., Vale, R. D., and Fletterick, R. J. (1996) *Nature (London)* **380**, 555–559
176. Thormählen, M., Marx, A., Müller, S. A., Song, Y.-H., Mandelkow, E.-M., Aebi, U., and Mandelkow, E. (1998) *J. Mol. Biol.* **275**, 795–809
177. Block, S. M. (1998) *Cell* **93**, 5–8
178. Wells, A. L., Lin, A. W., Chen, L.-Q., Safer, D., Cain, S. M., Hasson, T., Carragher, B. O., Milligan, R. A., and Sweeney, H. L. (1999) *Nature (London)* **401**, 505–508
179. Schliwa, M. (1999) *Nature (London)* **401**, 431–432
179a. Homma, K., Yoshimura, M., Saito, J., Ikebe, R., and Ikebe, M. (2001) *Nature (London)* **412**, 831–834
180. Ma, Y.-Z., and Taylor, E. W. (1995) *Biochemistry* **34**, 13242–13251
181. Ma, Y.-Z., and Taylor, E. W. (1997) *J. Biol. Chem.* **272**, 717–723
182. Visscher, K., Schnitzer, M. J., and Block, S. M. (1999) *Nature (London)* **400**, 184–189
182a. Tomishige, M., Klopfenstein, D. R., and Vale, R. D. (2002) *Science* **297**, 2263–2267
183. Mehta, A. D., Rock, R. S., Rief, M., Spudich, J. A., Mooseker, M. S., and Cheney, R. E. (1999) *Nature (London)* **400**, 590–593
183a. Rief, M., Rock, R. S., Mehta, A. D., Mooseker, M. S., Cheney, R. E., and Spudich, J. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9482–9486
183b. Trybus, K. M., Krementsova, E., and Freyzon, Y. (1999) *J. Biol. Chem.* **274**, 27448–27456
183c. De La Cruz, E. M., Wells, A. L., Sweeney, H. L., and Ostap, E. M. (2000) *Biochemistry* **39**, 14196–14202
183d. Rock, R. S., Rice, S. E., Wells, A. L., Purcell, T. J., Spudich, J. A., and Sweeney, H. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13655–13659
183e. Karcher, R. L., Roland, J. T., Zappacosta, F., Huddleston, M. J., Annan, R. S., Carr, S. A., and Gelfand, V. I. (2001) *Science* **293**, 1317–1320
184. Goldman, Y. E. (1998) *Cell* **93**, 1–4
185. Taylor, E. W. (1991) *J. Biol. Chem.* **266**, 294–302
186. Houadjeto, M., Travers, F., and Barman, T. (1992) *Biochemistry* **31**, 1564–1569
187. White, H. D., Belknap, B., and Webb, M. R. (1997) *Biochemistry* **36**, 11828–11836
188. Spudich, J. A. (1994) *Nature (London)* **372**, 515–518
189. Cooke, R. (1993) *FASEB J.* **9**, 636–642
190. Ostap, E. M., Barnett, V. A., and Thomas, D. D. (1995) *Biophys. J.* **69**, 177–188
191. Kambara, T., Rhodes, T. E., Ikebe, R., Yamada, M., White, H. D., and Ikebe, M. (1999) *J. Biol. Chem.* **274**, 16400–16406
191a. Lionne, C., Stehle, R., Travers, F., and Barman, T. (1999) *Biochemistry* **38**, 8512–8520
192. Harrington, W. F., and Rodgers, M. E. (1984) *Ann. Rev. Biochem.* **53**, 35–73
192a. Uyeda, T. Q. P., Tokuraku, K., Kaseda, K., Webb, M. R., and Patterson, B. (2002) *Biochemistry* **41**, 9525–9534
192b. Himmel, D. M., Gourinath, S., Reshetnikova, L., Shen, Y., Szent-Györgyi, A. G., and Cohen, C. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12645–12650
193. Yount, R. G., Lawson, D., and Rayment, I. (1995) *Biophys. J.* **68**, 445–495
193a. Burghardt, T. P., Park, S., and Ajtai, K. (2001) *Biochemistry* **40**, 4834–4843
193b. Borejdo, J., Ushakov, D. S., Moreland, R., Akopova, I., Reshetnyak, Y., Saraswat, L. D., Kamm, K., and Lowey, S. (2001) *Biochemistry* **40**, 3796–3803
194. Dominguez, R., Freyzon, Y., Trybus, K. M., and Cohen, C. (1998) *Cell* **94**, 559–571
195. Highsmith, S. (1999) *Biochemistry* **38**, 9791–9797
196. Gulick, A. M., Bauer, C. B., Thoden, J. B., Pate, E., Yount, R. G., and Rayment, I. (2000) *J. Biol. Chem.* **275**, 398–408
197. Geeves, M. A., and Holmes, K. C. (1999) *Ann. Rev. Biochem.* **68**, 687–728
198. Block, S. M. (1992) *Nature (London)* **360**, 493–495
199. Mehta, A. D., Rief, M., Spudich, J. A., Smith, D. A., and Simmons, R. M. (1999) *Science* **283**, 1689–1695
200. Finer, J. T., Simmons, R. M., and Spudich, J. A. (1994) *Nature (London)* **368**, 113–119
201. Molloy, J. E., Burns, J. E., Kendrick-Jones, J., Tregear, R. T., and White, D. C. S. (1995) *Nature (London)* **378**, 209–212
202. Mehta, A. D., Rief, M., and Spudich, J. A. (1999) *J. Biol. Chem.* **274**, 14517–14520
203. Howard, J. (1997) *Nature (London)* **389**, 561–567
204. Guilford, W. H., Dupuis, D. E., Kennedy, G., Wu, J., Patlak, J. B., and Warshaw, D. M. (1997) *Biophys. J.* **72**, 1006–1021
204a. Rosenfeld, S. S., Xing, J., Whitaker, M., Cheung, H. C., Brown, F., Wells, A., Milligan, R. A., and Sweeney, H. L. (2000) *J. Biol. Chem.* **275**, 25418–25426
205. Sase, I., Miyata, H., Ishiwata, S., and Kinosita, K., Jr. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5646–5650
206. Kitamura, K., Tokunaga, M., Iwane, A. H., and Yanagida, T. (1999) *Nature (London)* **397**, 129–134
207. Smith, D. A. (1998) *Biophys. J.* **75**, 2996–3007
207a. Yanagida, T., and Iwane, A. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9357–9359
208. Veigel, C., Coluccio, L. M., Jontes, J. D., Sparrow, J. C., Milligan, R. A., and Molloy, J. E. (1999) *Nature (London)* **398**, 530–533
208a. Piazzesi, G., Reconditi, M., Linari, M., Lucii, L., Sun, Y.-B., Narayanan, T., Boesecke, P., Lombardi, V., and Irving, M. (2002) *Nature (London)* **415**, 659–662
208b. Tanaka, H., Homma, K., Iwane, A. H., Katayama, E., Ikebe, R., Saito, J., Yanagida, T., and Ikebe, M. (2002) *Nature (London)* **415**, 192–195
209. Berliner, E., Mahtani, H. K., Karki, S., Chu, L. F., Cronan, J. E., Jr., and Gelles, J. (1994) *J. Biol. Chem.* **269**, 8610–8615
210. Schnitzer, M. J., and Block, S. M. (1997) *Nature (London)* **388**, 386–390
211. Coppin, C. M., Finer, J. T., Spudich, J. A., and Vale, R. D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1913–1917
212. Irving, M., and Goldman, Y. E. (1999) *Nature (London)* **398**, 463–465
212a. Shimizu, T., Thorn, K. S., Ruby, A., and Vale, R. D. (2000) *Biochemistry* **39**, 5265–5273
212b. Ryu, W. S., Berry, R. M., and Berg, H. C. (2000) *Nature (London)* **403**, 444–447
213. Young, E. C., Mahtani, H. K., and Gelles, J. (1998) *Biochemistry* **37**, 3467–3479
214. Foster, K. A., Correia, J. J., and Gilbert, S. P. (1998) *J. Biol. Chem.* **273**, 35307–35318
215. deCastro, M. J., Ho, C.-H., and Stewart, R. J. (1999) *Biochemistry* **38**, 5076–5081
216. Platts, J. A., Howard, S. T., and Bracke, B. R. F. (1996) *J. Am. Chem. Soc.* **118**, 2726–2733
217. Sosa, H., and Milligan, R. A. (1996) *J. Mol. Biol.* **260**, 743–755
218. Hirose, K., Cross, R. A., and Amos, L. A. (1998) *J. Mol. Biol.* **278**, 389–400
219. Sablin, E. P., Case, R. B., Dai, S. C., Hart, C. L., Ruby, A., Vale, R. D., and Fletterick, R. J. (1998) *Nature (London)* **395**, 813–816
220. Hirose, K., Lockhart, A., Cross, R. A., and Amos, L. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9539–9544
220a. Hua, W., Chung, J., and Gelles, J. (2002) *Science* **295**, 844–848
221. Rice, S., Lin, A. W., Safer, D., Hart, C. L., Naber, N., Carragher, B. O., Cain, S. M., Pechatnikova, E., Wilson-Kubalek, E. M., Whittaker, M., Pate, E., Cooke, R., Taylor, E. W., Milligan, R. A., and Vale, R. D. (1999) *Nature (London)* **402**, 778–784
222. Hancock, W. O., and Howard, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13147–13152
223. Crevel, I., Carter, N., Schliwa, M., and Cross, R. (1999) *EMBO J.* **18**, 5863–5872
224. Brendza, K. M., Rose, D. J., Gilbert, S. P., and Saxton, W. M. (1999) *J. Biol. Chem.* **274**, 31506–31514
225. Pate, E., Naber, N., Matuska, M., Franks-Skiba, K., and Cooke, R. (1997) *Biochemistry* **36**, 12155–12166
226. Moschcovich, L., Peyser, Y. M., Salomon, C., Burghardt, T. P., and Muhlrad, A. (1998) *Biochemistry* **37**, 15137–15143
227. Gopal, D., Pavlov, D. I., Levitsky, D. I., Ikebe, M., and Burke, M. (1996) *Biochemistry* **35**, 10149–10157
227a. Kliche, W., Pfannstiel, J., Tiepold, M., Stoeva, S., and Faulstich, H. (1999) *Biochemistry* **38**, 10307–10317
227b. Furch, M., Fujita-Becker, S., Geeves, M. A., Holmes, K. C., and Manstein, D. J. (1999) *J. Mol. Biol.* **290**, 797–809
227c. Allin, C., and Gerwert, K. (2001) *Biochemistry* **40**, 3037–3046
227d. Nyitrai, M., Hild, G., Lukács, A., Bódis, E., and Somogyi, B. (2000) *J. Biol. Chem.* **275**, 2404–2409

References

228. Xiao, M., Li, H., Snyder, G. E., Cooke, R., Yount, R. G., and Selvin, P. R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15309–15314
229. Palm, T., Sale, K., Brown, L., Li, H., Hambly, B., and Fajer, P. G. (1999) *Biochemistry* **38**, 13026–13034
230. Corrie, J. E. T., Brandmeier, B. D., Ferguson, R. E., Trentham, D. R., Kendrick-Jones, J., Hopkins, S. C., van der Heide, U. A., Goldman, Y. E., Sabido-David, C., Dale, R. E., Criddle, S., and Irving, M. (1999) *Nature (London)* **400**, 425–430
231. Suzuki, Y., Yasunaga, T., Ohkura, R., Wakabayashi, T., and Sutoh, K. (1998) *Nature (London)* **396**, 380–383
232. Ashley, R. (1972) *J. Theor. Biol.* **36**, 339–354
233. McClare, C. W. F. (1972) *J. Theor. Biol.* **35**, 569–595
234. Davydov, A. S. (1973) *J. Theor. Biol.* **38**, 559–569
235. Porter, K. R., and Franzini-Armstrong, C. (1965) *Sci. Am.* **212**(Mar), 73–80
236. Hoyle, G. (1970) *Sci. Am.* **222**(Apr), 85–93
237. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) *Molecular Biology of the Cell*, 3rd ed., Garland, New York
238. Macrez-Leprêtre, N., Kalkbrenner, F., Schultz, G., and Mironneau, J. (1997) *J. Biol. Chem.* **272**, 5261–5268
238a. Toyoshima, C., and Nomura, H. (2002) *Nature (London)* **418**, 605–611
239. Sharma, M. R., Penczek, P., Grassucci, R., Xin, H.-B., Fleischer, S., and Wagenknecht, T. (1998) *J. Biol. Chem.* **273**, 18429–18434
240. Bers, D. M., and Fill, M. (1998) *Science* **281**, 790–791
241. Sonnleitner, A., Conti, A., Bertocchini, F., Schindler, H., and Sorrentino, V. (1998) *EMBO J.* **17**, 2790–2798
242. Du, G. G., and MacLennan, D. H. (1998) *J. Biol. Chem.* **273**, 31867–31872
243. Bidasee, K. R., and Besch, H. R., Jr. (1998) *J. Biol. Chem.* **273**, 12176–12186
243a. Sun, J., Xin, C., Eu, J. P., Stamler, J. S., and Meissner, G. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11158–11162
243b. Feng, W., Liu, G., Allen, P. D., and Pessah, I. N. (2000) *J. Biol. Chem.* **275**, 35902–35907
243c. Patel, S., Churchill, G. C., and Galione, A. (2001) *Trends Biochem. Sci.* **26**, 482–489
243d. Baker, M. L., Serysheva, I. I., Sencer, S., Wu, Y., Ludtke, S. J., Jiang, W., Hamilton, S. L., and Chiu, W. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12155–12160
244. Takeshima, H., Komazaki, S., Hirose, K., Nishi, M., Noda, T., and Lino, M. (1998) *EMBO J.* **17**, 3309–3316
245. Zhou, J., Cribbs, L., Yi, J., Shirokov, R., Perez-Reyes, E., and Ríos, E. (1998) *J. Biol. Chem.* **273**, 25503–25509
245a. Serysheva, I. I., Ludtke, S. J., Baker, M. R., Chiu, W., and Hamilton, S. L. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10370 –10375
246. Knudson, C. M., Chaudhari, N., Sharp, A. H., Powell, J. A., Beam, K. G., and Campbell, K. P. (1989) *J. Biol. Chem.* **264**, 1345–1348
247. Smith, G. D., Keizer, J. E., Stern, M. D., Lederer, W. J., and Cheng, H. (1998) *Biophys. J.* **75**, 15–32
248. Fay, F. S. (1995) *Science* **270**, 588–589
248a. Kamm, K. E., and Stull, J. T. (2001) *J. Biol. Chem.* **276**, 4527–4530
249. Xie, X., Rao, S., Walian, P., Hatch, V., Phillips, G. N., Jr., and Cohen, C. (1994) *J. Mol. Biol.* **236**, 1212–1226
250. Dufour, C., Weinberger, R. P., Schevzov, G., Jeffrey, P. L., and Gunning, P. (1998) *J. Biol. Chem.* **273**, 18547–18555
251. Kagawa, H., Sugimoto, K., Matsumoto, H., Inoue, T., Imadzu, H., Takuwa, K., and Sakube, Y. (1995) *J. Mol. Biol.* **251**, 603–613
252. Amos, L. A. (1985) *Ann. Rev. Biophys. Biophys. Chem.* **14**, 291–313
253. Zot, A. S., and Potter, J. D. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 535–559
254. Saeki, K., Sutoh, K., and Wakabayashi, T. (1996) *Biochemistry* **35**, 14465–14472
255. Squire, J. M., and Morris, E. P. (1998) *FASEB J.* **12**, 761–771
256. Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) *J. Mol. Biol.* **271**, 728–750
257. Malnic, B., Farah, C. S., and Reinach, F. C. (1998) *J. Biol. Chem.* **273**, 10594–10601
258. Hernández, G., Blumenthal, D. K., Kennedy, M. A., Unkefer, C. J., and Trewhella, J. (1999) *Biochemistry* **38**, 6911–6917
259. Vassylyev, D. G., Takeda, S., Wakatsuki, S., Maeda, K., and Meéda, Y. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4847–4852
260. McKay, R. T., Tripet, B. P., Pearlstone, J. R., Smillie, L. B., and Sykes, B. D. (1999) *Biochemistry* **38**, 5478–5489
260a. Mercier, P., Spyracopoulos, L., and Sykes, B. D. (2001) *Biochemistry* **40**, 10063–10077
261. Kobayashi, T., Kobayashi, M., Gryczynski, Z., Lakowicz, J. R., and Collins, J. H. (2000) *Biochemistry* **39**, 86–91
262. Strynadka, N. C. J., Cherney, M., Sielecki, A. R., Li, M. X., Smillie, L. B., and James, M. N. G. (1997) *J. Mol. Biol.* **273**, 238–255
263. da Silva, A. C. R., and Reinach, F. C. (1991) *Trends Biochem. Sci.* **16**, 53–57
264. Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry* **36**, 4386–4392
265. Li, H.-C., and Fajer, P. G. (1998) *Biochemistry* **37**, 6628–6635
265a. Lehman, W., Rosol, M., Tobacman, L. S., and Craig, R. (2001) *J. Mol. Biol.* **307**, 739–744
266. Reiffert, S. U., Jaquet, K., Heilmeyer, L. M. G., Jr., and Herberg, F. W. (1998) *Biochemistry* **37**, 13516–13525
267. Simmerman, H. K. B., Kobayashi, Y. M., Autry, J. M., and Jones, L. R. (1996) *J. Biol. Chem.* **271**, 5941–5946
268. Reddy, L. G., Jones, L. R., and Thomas, D. D. (1999) *Biochemistry* **38**, 3954–3962
268a. Asahi, M., Green, N. M., Kurzydlowski, K., Tada, M., and MacLennan, D. H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10061–10066
269. Holmes, K. C. (1995) *Biophys. J.* **68**, 2s–7s
270. Vibert, P., Craig, R., and Lehman, W. (1997) *J. Mol. Biol.* **266**, 8–14
270a. Moraczewska, J., and Hitchcock-DeGregori, S. E. (2000) *Biochemistry* **39**, 6891–6897
270b. Narita, A., Yasunaga, T., Ishikawa, T., Mayanagi, K., and Wakabayashi, T. (2001) *J. Mol. Biol.* **308**, 241–261
270c. Craig, R., and Lehman, W. (2001) *J. Mol. Biol.* **311**, 1027–1036
271. Hnath, E. J., Wang, C.-L. A., Huber, P. A. J., Marston, S. B., and Phillips, G. N., Jr. (1996) *Biophys. J.* **71**, 1920–1933
271a. Gerson, J. H., Kim, E., Muhlrad, A., and Reisler, E. (2001) *J. Biol. Chem.* **276**, 18442–18449
272. Allen, B. G., and Walsh, M. P. (1994) *Trends Biochem. Sci.* **19**, 362–368
273. Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) *Science* **256**, 632–638
274. Wu, X., Clack, B. A., Zhi, G., Stull, J. T., and Cremo, C. R. (1999) *J. Biol. Chem.* **274**, 20328–20335
275. Lee, M. R., Li, L., and Kitazawa, T. (1997) *J. Biol. Chem.* **272**, 5063–5068
276. Ye, L.-H., Kishi, H., Nakamura, A., Okagaki, T., Tanaka, T., Oiwa, K., and Kohama, K. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6666–6671
277. Lehman, W., Vibert, P., and Craig, R. (1997) *J. Mol. Biol.* **274**, 310–317
278. Graceffa, P. (1997) *Biochemistry* **36**, 3792–3801
278a. Wang, Z., and Yang, Z.-Q. (2000) *Biochemistry* **39**, 11114–11120
279. Gollub, J., Cremo, C. R., and Cooke, R. (1999) *Biochemistry* **38**, 10107–10118
279a. Stafford, W. F., Jacobsen, M. P., Woodhead, J., Craig, R., O'Neall-Hennessey, E. O., and Szent-Györgyi, A. G. (2001) *J. Mol. Biol.* **307**, 137–147
280. Yamada, A., Yoshio, M., Oiwa, K., and Nyitray, L. (2000) *J. Mol. Biol.* **295**, 169–178
280a. Yamada, A., Yoshio, M., Kojima, H., and Oiwa, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6635–6640
281. Brzeska, H., and Korn, E. D. (1996) *J. Biol. Chem.* **271**, 16983–16986
282. Benian, G. M., Kiff, J. E., Neckelmann, N., Moerman, D. G., and Waterston, R. H. (1989) *Nature (London)* **342**, 45–50
283. Heierhorst, J., Kobe, B., Feil, S. C., Parker, M. W., Benian, G. M., Weiss, K. R., and Kemp, B. E. (1996) *Nature (London)* **380**, 636–639
284. Lei, J., Tang, X., Chambers, T. C., Pohl, J., and Benian, G. M. (1994) *J. Biol. Chem.* **269**, 21078–21085
285. Johnson, K. A., and Quiocho, F. A. (1996) *Nature (London)* **380**, 585–587
286. Bessman, S. P., and Carpenter, C. L. (1985) *Ann. Rev. Biochem.* **54**, 831–862
287. Balaban, R. S., Kantor, H. L., and Ferretti, J. A. (1983) *J. Biol. Chem.* **258**, 12787–12789
288. Park, J. H., Brown, R. L., Park, C. R., McCully, K., Cohn, M., Haselgrove, J., and Chance, B. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8976–8980
289. Savabi, F. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7476–7480
289a. Kashiyama, T., Ito, K., and Yamamoto, K. (2001) *J. Mol. Biol.* **311**, 461–466
289b. Pantaloni, D., Le Clainche, C., and Carlier, M.-F. (2001) *Science* **292**, 1502–1506
289c. Pollard, T. D. (2000) *Trends Biochem. Sci.* **25**, 607–611
290. Machesky, L. M., and Way, M. (1998) *Nature (London)* **394**, 125–126
291. Carlier, M.-F., and Pantaloni, D. (1997) *J. Mol. Biol.* **269**, 459–467
291a. Pruyne, D., Evangelista, M., Yang, C., Bi, E., Zigmond, S., Bretscher, A., and Boone, C. (2002) *Science* **297**, 612–615
292. Yin, H. L., and Stull, J. T. (1999) *J. Biol. Chem.* **274**, 32529–32530
292a. Robinson, R. C., Turbedsky, K., Kaiser, D. A., Marchand, J.-B., Higgs, H. N., Choe, S., and Pollard, T. D. (2001) *Science* **294**, 1679–1684
292b. Dayel, M. J., Holleran, E. A., and Mullins, R. D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14871–14876
293. Svitkina, T. M., and Borisy, G. G. (1999) *Trends Biochem. Sci.* **24**, 432–436
293a. Boldogh, I. R., Yang, H.-C., Nowakowski, W. D., Karmon, S. L., Hays, L. G., Yates, J. R., III, and Pon, L. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3162–3167
293b. Kim, A. S., Kakalis, L. T., Abdul-Manan, N., Liu, G. A., and Rosen, M. K. (2000) *Nature (London)* **404**, 151–158
293c. Higgs, H. (2001) *Trends Biochem. Sci.* **26**, 219
293d. Kaibuchi, K., Kuroda, S., and Amano, M. (1999) *Ann. Rev. Biochem.* **68**, 459–486
293e. Castellano, F., Le Clainche, C., Patin, D., Carlier, M.-F., and Chavrier, P. (2001) *EMBO J.* **20**, 5603–5614

References

293f. Reinhard, M., Jarchau, T., and Walter, U. (2001) *Trends Biochem. Sci.* **26**, 243–249

294. Gutsche-Perelroizen, I., Lepault, J., Ott, A., and Carlier, M.-F. (1999) *J. Biol. Chem.* **274**, 6234–6243

295. Kang, F., Purich, D. L., and Southwick, F. S. (1999) *J. Biol. Chem.* **274**, 36963–36972

296. Chen, H., Bernstein, B. W., and Bamburg, J. R. (2000) *Trends Biochem. Sci.* **25**, 19–23

297. Carlier, M.-F., Ressad, F., and Pantaloni, D. (1999) *J. Biol. Chem.* **274**, 33827–33830

297a. Ono, S., McGough, A., Pope, B. J., Tolbert, V. T., Bui, A., Pohl, J., Benian, G. M., Gernert, K. M., and Weeds, A. G. (2001) *J. Biol. Chem.* **276**, 5952–5958

298. Theriot, J. A., and Mitchison, T. J. (1991) *Nature (London)* **352**, 126–131

299. Stossel, T. P. (1993) *Science* **260**, 1086–1094

300. Tsukita, S., and Yonemura, S. (1999) *J. Biol. Chem.* **274**, 34507–34510

301. Hanakam, F., Gerisch, G., Lotz, S., Alt, T., and Seelig, A. (1996) *Biochemistry* **35**, 11036–11044

301a. Niggli, V. (2001) *Trends Biochem. Sci.* **26**, 604–611

301b. Caroni, P. (2001) *EMBO J.* **20**, 4332–4336

301c. van den Ent, F., Amos, L. A., and Löwe, J. (2001) *Nature (London)* **413**, 39–44

302. Bullock, T. L., Roberts, R. M., and Stewart, M. (1996) *J. Mol. Biol.* **263**, 284–296

302a. Villeneuve, A. M. (2001) *Science* **291**, 2099–2101

303. Spudich, J. D., and Lord, K. (1974) *J. Biol. Chem.* **249**, 6013–6020

304. Allen, R. D. (1987) *Sci. Am.* **256**(Feb), 42–47

305. Gibbons, I. R. (1988) *J. Biol. Chem.* **263**, 15837–15840

305a. Setou, M., Nakagawa, T., Seog, D.-H., and Hirokawa, N. (2000) *Science* **288**, 1796–1802

305b. Goldstein, L. S. B. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6999–7003

305c. Fan, J.-S., Zhang, Q., Tochio, H., Li, M., and Zhang, M. (2001) *J. Mol. Biol.* **306**, 97–108

306. Rodionov, V., Nadezhdina, E., and Borisy, G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 115–120

307. Burgess, S. A. (1995) *J. Mol. Biol.* **250**, 52–63

308. Caplow, M., and Shanks, J. (1998) *Biochemistry* **37**, 12994–13002

309. Dougherty, C. A., Himes, R. H., Wilson, L., and Farrell, K. W. (1998) *Biochemistry* **37**, 10861–10865

310. Hyams, J. S., and Lloyd, C. W., eds. (1994) *Microtubules*, Wiley-Liss, New York

311. Drewes, G., Ebneth, A., and Mandelkow, E.-M. (1998) *Trends Biochem. Sci.* **23**, 307–311

312. Vinh, J., Langridge, J. I., Bré, M.-H., Levilliers, N., Redeker, V., Loyaux, D., and Rossier, J. (1999) *Biochemistry* **38**, 3133–3139

313. Redeker, V., Levilliers, N., Schmitter, J.-M., Le Caer, J.-P., Rossier, J., Adoutte, A., and Bré, M.-H. (1994) *Science* **266**, 1688–1691

314. Regnard, C., Audebert, S., Desbruyères, E., Denoulet, P., and Eddé, B. (1998) *Biochemistry* **37**, 8395–8404

314a. Sharp, D. J., Rogers, G. C., and Scholey, J. M. (2000) *Nature (London)* **407**, 41–47

315. Hoyt, M. A., Hyman, A. A., and Bähler, M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12747–12748

316. Barton, N. R., and Goldstein, L. S. B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1735–1742

317. Habura, A., Tikhonenko, I., Chisholm, R. L., and Koonce, M. P. (1999) *J. Biol. Chem.* **274**, 15447–15453

318. Samsó, M., Radermacher, M., Frank, J., and Koonce, M. P. (1998) *J. Mol. Biol.* **276**, 927–937

319. Shimizu, T., Toyoshima, Y. Y., Edamatsu, M., and Vale, R. D. (1995) *Biochemistry* **34**, 1575–1582

320. King, S. M., Barbarese, E., Dillman, J. F., III, Patel-King, R. S., Carson, J. H., and Pfister, K. K. (1996) *J. Biol. Chem.* **271**, 19358–19366

321. Koonce, M. P., Köhler, J., Neujahr, R., Schwartz, J.-M., Tikhonenko, I., and Gerisch, G. (1999) *EMBO J.* **18**, 6786–6792

322. Sleigh, M. A., ed. (1974) *Cilia and Flagella*, Academic Press, New York

323. Smith, E. F., and Sale, W. S. (1994) in *Microtubules* (Hyams, J. S., and Lloyd, C. W., eds), pp. 381–392, Wiley-Liss, New York

324. Tanaka-Takiguchi, Y., Itoh, T. J., and Hotani, H. (1998) *J. Mol. Biol.* **280**, 365–373

325. Satir, P. (1999) *FASEB J.* **13**, S235–S237

325a. Fang, Y.-I., Yokota, E., Mabuchi, I., Nakamura, H., and Obizumi, Y. (1997) *Biochemistry* **36**, 15561–15567

325b. Sakakibara, H., Kojima, H., Sakai, Y., Katayama, E., and Oiwa, K. (1999) *Nature (London)* **400**, 586–590

326. Shingyoji, C., Higuchi, H., Yoshimura, M., Katayama, E., and Yanagida, T. (1998) *Nature (London)* **393**, 711–714

327. Hunt, A. J. (1998) *Nature (London)* **393**, 624–625

328. Benashski, S. E., Patel-King, R. S., and King, S. M. (1999) *Biochemistry* **38**, 7253–7264

329. Sleigh, M. A., ed. (1974) *Cilia and Flagella*, p. 14 Academic Press, New York

329a. Smith, E. F., and Sale, W. S. (1994) in *Microtubules* (Hyams, J. S., and Lloyd, C. W., eds), pp. 381–392, Wiley-Liss, New York

330. Klein, P. S., Sun, T. J., Saxe, C. L., III, Kimmel, A. R., Johnson, R. L., and Devreotes, P. N. (1988) *Science* **241**, 1467–1472

331. Caterina, M. J., and Devreotes, P. N. (1991) *FASEB J.* **5**, 3078–3085

332. Snyderman, R., and Goetzl, E. J. (1981) *Science* **213**, 830–837

333. Gerard, N. P., Bao, L., Xiao-Ping, H., Eddy, R. L., Jr., Shows, T. B., and Gerard, C. (1993) *Biochemistry* **32**, 1243–1250

334. Thelen, M., Peveri, P., Kernen, P., von Tscharner, V., Walz, A., and Baggiolini, M. (1988) *FASEB J.* **2**, 2702–2706

335. Schwenk, U., and Schröder, J.-M. (1995) *J. Biol. Chem.* **270**, 15029–15036

336. Hsu, M. H., Chiang, S. C., Ye, R. D., and Prossnitz, E. R. (1997) *J. Biol. Chem.* **272**, 29426–29429

337. Jin, T., Zhang, N., Long, Y., Parent, C. A., and Devreotes, P. N. (2000) *Science* **287**, 1034–1036

338. Chung, C. Y., Funamoto, S., and Firtel, R. A. (2001) *Trends Biochem. Sci.* **26**, 557–566

339. Mahadevan, L., and Matsudaira, P. (2000) *Science* **288**, 95–99

340. Merz, A. J., So, M., and Sheetz, M. P. (2000) *Nature (London)* **407**, 98–102

341. Kirchhausen, T. (1999) *Nature (London)* **398**, 470–471

342. McNiven, M. A., Cao, H., Pitts, K. R., and Yoon, Y. (2000) *Trends Biochem. Sci.* **25**, 115–120

Study Questions

1. Describe briefly major aspects of the structure, properties, locations, and functions of each of the following proteins of skeletal muscle.

    Actin
    Myosin
    Titin
    Nebulin
    α-Actinin
    C-protein
    Tropomyosin
    Troponin
    Myomesin
    Desmin
    Vimentin

2. Describe the generally accepted sliding filament model of muscle contraction. List some uncertainties in this description.

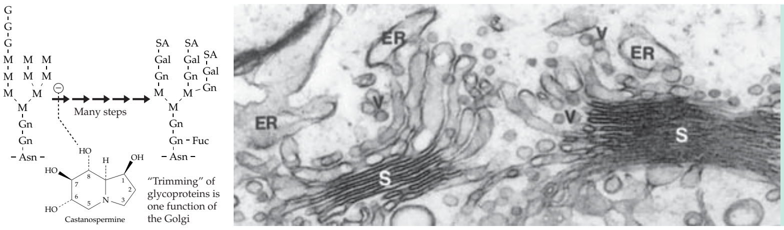
3. Compare mechanisms that regulate contraction in skeletal muscle and in smooth muscle.

4. Compare myosin with kinesins and dyneins. What features do they have in common? What differences can you describe?

5. Compare the properties of actin in skeletal muscle and in nonmuscle cells. What is meant by “treadmilling?” What is “actin-based motility?”

6. The human genome contains more than 100 genes for proteins of the kinesin superfamily. Why?

7. Describe some of the major diseases that involve muscle proteins.



"Trimming" of glycoproteins is one function of the Golgi

The branched oligosaccharides of glycoprotein surfaces are formed on asparagine side chains of selected cell surface proteins. The oligosaccharide at the left is formed in the ER and is transferred intact (Fig. 20-6) to an acceptor asparagine. It is then trimmed by removal of glucose and mannose units and residues of glucosamine, galactose, and fucose are added as in Fig. 20-7. These reactions begin in the ER and continue in the Golgi apparatus (right). See also Fig. 20-8.

Contents

1129 .... **A. Interconversions of Monosaccharides**
1130 .... 1. The Metabolism of Galactose
1131 .... 2. Inositol
1132 .... 3. D-Glucuronic Acid, Ascorbic Acid, and Xylitol
1135 .... 4. Transformations of Fructose 6-Phosphate
1136 .... 5. Extending a Sugar Chain with Phosphoenolpyruvate (PEP)
1137 .... 6. Synthesis of Deoxy Sugars
1140 .... **B. Synthesis and Utilization of Oligosaccharides**
1143 .... **C. Synthesis and Degradation of Polysaccharides**
1143 .... 1. Glycogen and Starch
1146 .... 2. Cellulose, Chitin, and Related Glycans
1147 .... *An insertion mechanism for synthesis of cellulose*
1148 .... *Callose and other β-1,3-linked glycans*
1148 .... *Chitin*
1149 .... *Cell walls of plants*
1150 .... 3. Patterns in Polysaccharide Structures
1152 .... *Dextrans*
1152 .... *Lipid-dependent synthesis of polysaccharides*
1153 .... **D. Proteoglycans and Glycoproteins**
1153 .... 1. Glycosaminoglycans
1155 .... 2. *O*-Linked Oligosaccharides
1155 .... 3. Assembly of *N*-Linked Oligosaccharides on Glycosyl Carrier Lipids
1156 .... *Trimming of glycoprotein oligosaccharides*
1157 .... *Extensions and terminal elaborations*
1157 .... *The perplexing Golgi apparatus*
1159 .... *Lysosomal enzymes*
1160 .... *The hepatic asialoglycoprotein (Gal) receptor*
1160 .... **E. Biosynthesis of Bacterial Cell Walls**
1161 .... *The O-antigens and lipid A*
1161 .... *Gram-positive bacteria*
1168 .... **F. Biosynthesis of Eukaryotic Glycolipids**
1168 .... 1. Glycophosphatidylinositol (GPI) Anchors
1168 .... 2. Cerebrosides and Gangliosides
1169 .... **G. The Intracellular Breakdown of Polysaccharides and Glycolipids**
1169 .... 1. Mucopolysaccharidoses
1170 .... 2. Sphingolipidoses
1170 .... 3. Causes of Lysosomal Diseases
1172 .... 4. Can Lysosomal Diseases Be Treated?

1173 .... **References**
1179 .... **Study Questions**

Boxes

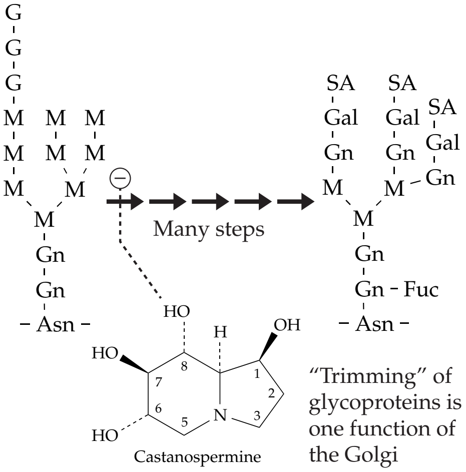
1131 .... Box 20-A Fructose for Sperm Cells via the Polyol Pathway
1139 .... Box 20-B The Biosynthesis of Streptomycin
1142 .... Box 20-C Osmotic Adaptation
1145 .... Box 20-D Genetic Diseases of Glycogen Metabolism
1149 .... Box 20-E Oligosaccharides in Defensive and Other Responses of Plants
1151 .... Box 20-F What Does Boron Do?
1164 .... Box 20-G Penicillins and Related Antibiotics
1166 .... Box 20-H Antibiotic Resistance and Vancomycin

Tables

1171 .... Table 20-1 Lysosomal Storage Diseases: Sphingolipidoses and Mucopolysaccharidoses

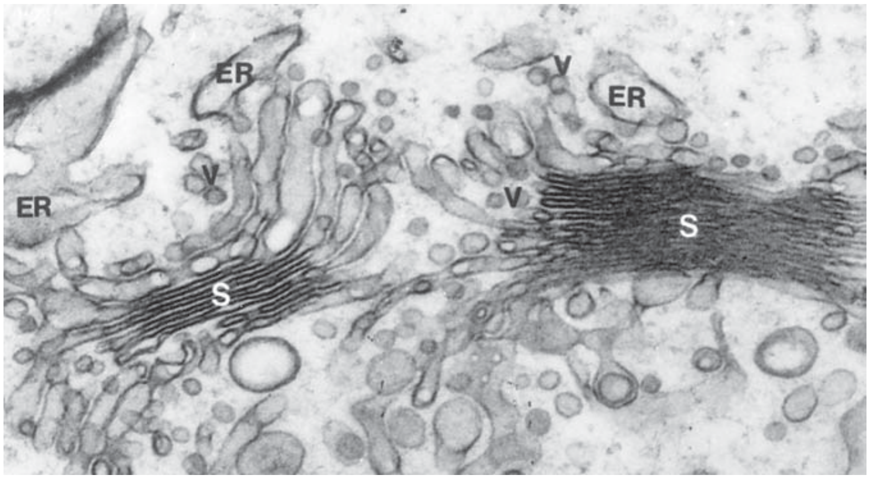
Some Pathways of Carbohydrate Metabolism

20



Castanospermine

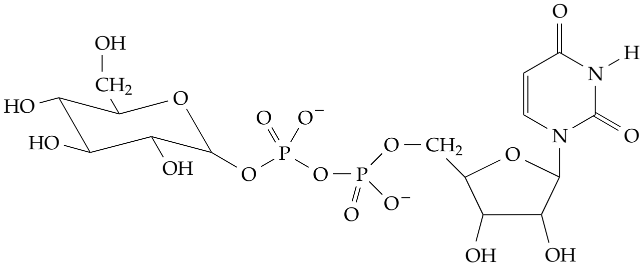
“Trimming” of glycoproteins is one function of the Golgi



The general principles of biosynthesis, as well as the pathways of formation of major carbohydrate and lipid precursors, are considered in Chapter 17. Also described are the processes of gluconeogenesis, the synthesis of glucose 6-phosphate and fructose 6-phosphate from free glucose, and typical polymerization pathways for formation of polysaccharides. In this chapter, additional aspects of the metabolism of monosaccharides, oligosaccharides, polysaccharides, glycoproteins, and glycolipids are considered. These are metabolic transformations that affect the physical properties of cell surfaces and body fluids. They are essential to signaling between cells, to establishment of the immunological identity of individuals, and to the development of strong cell wall materials. Some of the differences in carbohydrates found in bacteria, fungi, green plants, and mammals are considered.

A. Interconversions of Monosaccharides

Chemical interconversions between compounds are easiest at the level of oxidation of carbohydrates. Consequently, many reactions by which one sugar can be changed into another are known. Most of the transformations take place in the “sugar nucleotide derivatives” (see also Eq. 17-56). The first of this group of compounds to be recognized was **uridine diphosphate glucose** (UDPG), which was discovered around 1950 by L. F. Leloir[1,2] during his investigation of the metabolism of galactose 1-*P*. The fact that interconversions of hexoses take place largely at the sugar nucleotide level was unknown at the time. Leloir’s studies led to the characterization of both UDP-glucose and UDP-galactose.



Uridine diphosphate D-glucose

Figure 20-1 summarizes pathways by which glucose 6-phosphate or fructose 6-phosphate can be converted into many of the other sugars found in living things. Galactose and mannose can also be interconverted with the other sugars. A kinase forms **mannose 6-phosphate** which equilibrates with fructose 6-phosphate. Galactokinase converts free galactose to **galactose 1-phosphate**, which can be isomerized to glucose 1-phosphate by the reactions of Eq. 20-1. Fructose, an important human dietary constituent[3] derived largely by hydrolysis of sucrose, can also be formed in tissues via the **sorbitol pathway**[4] (Box 20-A). Fructose can be phosphorylated to fructose 1-phosphate by liver **fructokinase**. We have no mutase able to convert fructose 1-*P* to fructose 6-*P*, but a special aldolase cleaves fructose 1-*P* to dihydroxyacetone phosphate and free glyceraldehyde. Lack of this aldolase leads to occasionally observed cases of fructose intolerance.[5,6] The glyceraldehyde formed from fructose can be metabolized by reduction to glycerol followed by phosphorylation (glycerol kinase) and reoxidation to dihydroxyacetone phosphate. Some phosphorylation of fructose 1-*P* to fructose 1,6-P_2 apparently also occurs.[7] Interconversion of ribose 5-*P* and other sugar phosphates is a central part of the pentose phosphate pathway (Fig. 17-8). Free ribose can be phosphorylated by a **ribokinase**.[8]

Oxidation of UDP-glucose in two steps[9,9a] by NAD$^+$ yields **UDP-glucuronic** acid, which can be epimerized to **UDP-galacturonic** acid. Likewise (see bottom of Fig. 20-1), **guanosine disphosphate-mannose** (GDP-mannose) is oxidized to **GDP-mannuronic acid**, which undergoes 4-epimerization to **GDP-guluronic acid**. Looking again at the top of the scheme, notice that UDP-D-glucuronic acid may be epimerized at the 5 position to **UDP-L-iduronic acid**. However, the iduronic acid residues in dermatan sulfate arise by inversion at C-5 of D-glucuronic acid residues in the polymer.[10] The mechanism of these reactions, like that of the decarboxylation of UDP-glucuronic acid to UDP-xylose (near the top of Fig. 20-1), apparently have not been well investigated.

Notice that glucuronic acid is abbreviated GlcA, in accord with IUB recommendations. However, many authors use GlcUA, ManUA, etc., for the uronic acids.

1. The Metabolism of Galactose

The reactions of galactose have attracted biochemists' interest because of the occurrence of the rare (30 cases / million births) hereditary disorder **galactosemia**. When this defect is present, the body cannot transform galactose into glucose metabolites but reduces it to the sugar alcohol **galactitol** or oxidizes it to **galactonate**, both products being excreted in the urine. Unfortunately, severe gastrointestinal troubles often appear within a few days or weeks of birth. Growth is slow and cataracts develop in the eyes, probably as a result of the accumulating galactitol. Death may come quickly from liver damage. Fortunately, galactose-free diets can be prepared for young infants, and if the disease is diagnosed promptly the most serious damage can be avoided. However, it has not been possible to prevent long-term effects that include speech difficulties, learning disabilities, and ovarian dysfunction.[1,11]

In some less seriously affected galactosemic patients **galactokinase** (Eq. 20-1, step *a*) is absent, but it is more often **galactose-1-phosphate uridyltransferase** (Eq. 20-1, step *b*) that is missing or inactive.[12–15a] This enzyme transforms galactose 1-*P* to UDP-galactose by displacing glucose 1-*P* from UDP-glucose. The UDP-galactose is then isomerized by the NAD$^+$-dependent

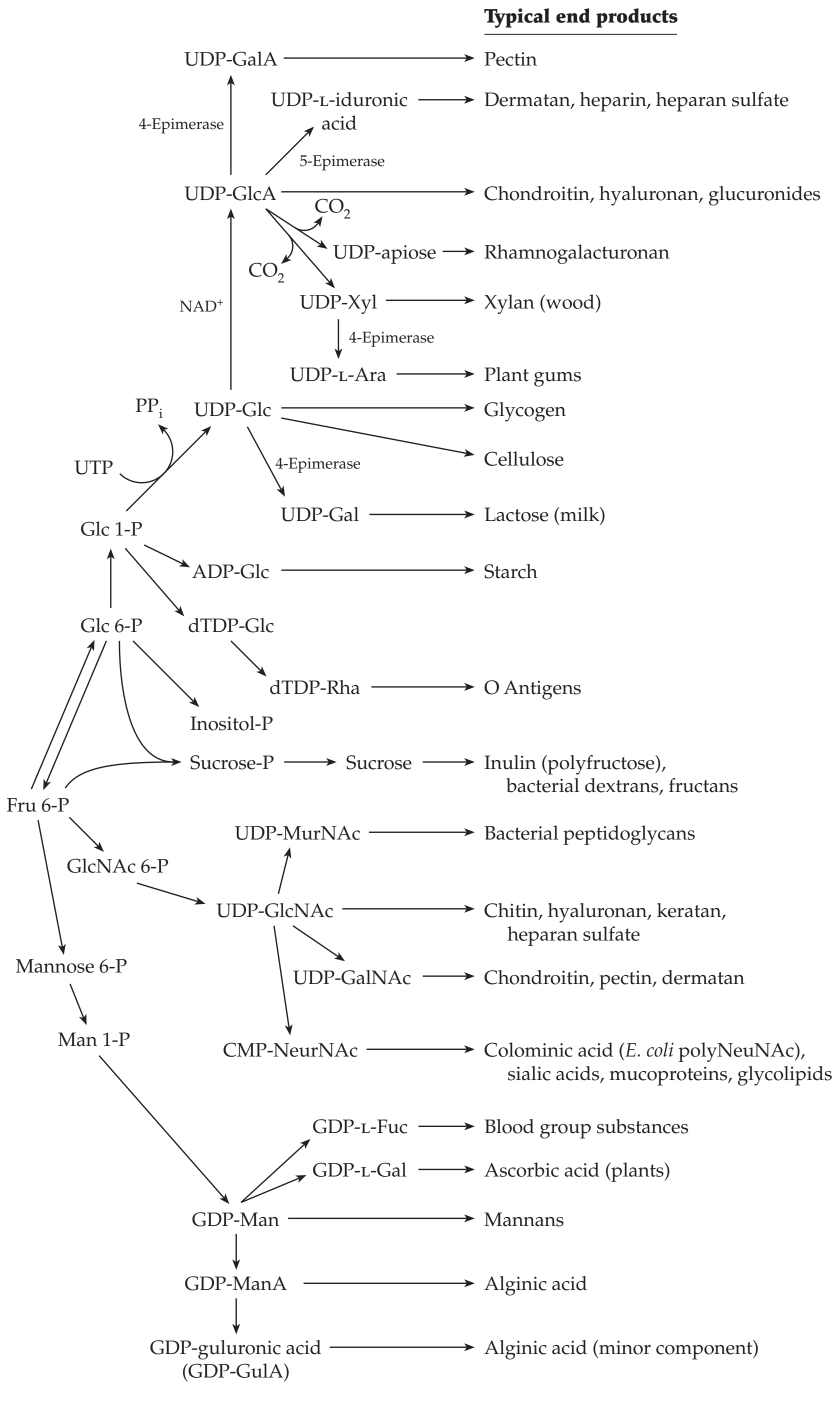
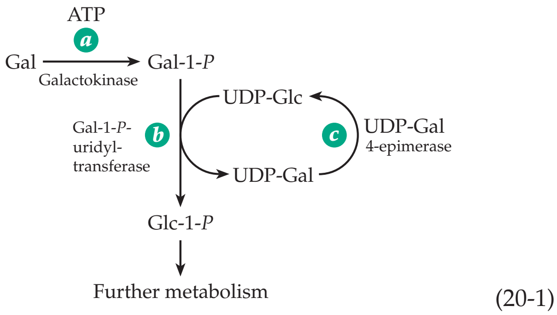

Typical end products
UDP-GalA
Pectin
4-Epimerase
UDP-L-iduronic acid
Dermatan, heparin, heparan sulfate
5-Epimerase
UDP-GlcA
Chondroitin, hyaluronan, glucuronides
CO₂
UDP-apiose
Rhamnogalacturonan
CO₂
NAD⁺
UDP-Xyl
Xylan (wood)
4-Epimerase
UDP-L-Ara
Plant gums
PPi
UDP-Glc
Glycogen
Cellulose
UTP
4-Epimerase
UDP-Gal
Lactose (milk)
Glc 1-P
ADP-Glc
Starch
Glc 6-P
dTDP-Glc
dTDP-Rha
O Antigens
Inositol-P
Sucrose-P
Sucrose
Inulin (polyfructose), bacterial dextrans, fructans
Fru 6-P
UDP-MurNAc
Bacterial peptidoglycans
GlcNAc 6-P
UDP-GlcNAc
Chitin, hyaluronan, keratan, heparan sulfate
Mannose 6-P
UDP-GalNAc
Chondroitin, pectin, dermatan
Man 1-P
CMP-NeurNAc
Colominic acid (E. coli polyNeuNAc), sialic acids, mucoproteins, glycolipids
GDP-L-Fuc
Blood group substances
GDP-L-Gal
Ascorbic acid (plants)
GDP-Man
Mannans
GDP-ManA
Alginic acid
GDP-guluronic acid (GDP-GulA)
Alginic acid (minor component)

Figure 20-1 Some routes of interconversion of monosaccharides and of polymerization of the activated glycosyl units.



(20-1)

UDP-Gal 4-**epimerase**[16–16b] (Eq. 20-1, step *c*; see also Eq. 15-13 and accompanying discussion). Absence of this enzyme also causes galactosemia.[11] The overall effect of the reactions of Eq. 20-1 is to transform galactose into glucose 1-*P*. At the same time, the 4-epimerase can operate in the reverse direction to convert UDP-glucose to UDP-galactose, when the latter is needed for biosynthesis (Fig. 20-1).

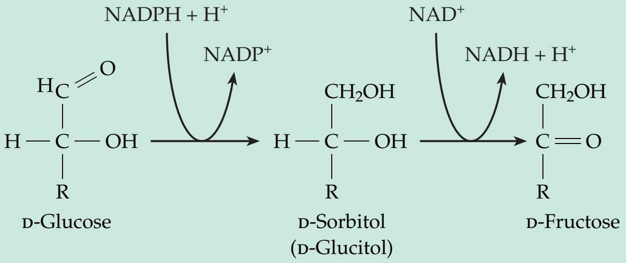
Another enzyme important to galactose metabolism, at least in *E. coli*, is **galactose mutarotase**.[17] Cleavage of lactose by β-galactosidase produces β-D-galactose which must be converted to the α-anomer by the mutarotase before it can be acted upon by galactokinase. Galactose is present in most glycoproteins and glycolipids in the pyranose ring form. However, in bacterial O-antigens, in cell walls of mycobacteria and fungi, and in some protozoa glactose occurs in the furanose form. The precursor is UDP-Gal*f*, which is formed from UDP-Gal*p* by **UDP-Gal*p* mutase**.[17a]

2. Inositol

Related to the monosaccharides is the hexahydroxycyclohexane ***myo*-inositol** (Eq. 20-2). This **cyclitol**, which is apparently present universally within cells (Fig. 11-9), can be formed from glucose-6-*P* according to Eq. 20-2 using a synthase that contains bound

BOX 20-A FRUCTOSE FOR SPERM CELLS VIA THE POLYOL PATHWAY

An interesting example of the way in which the high $[NADPH]/[NADP^+]$ and $[NAD^+]/[NADH]$ ratios in cells can be used to advantage is found in the metabolism of sperm cells. Whereas D-glucose is the commonest sugar used as an energy source by mammalian cells, spermatozoa use principally D-fructose, a sugar that is not readily metabolized by cells of surrounding tissues.[a–c] Fructose, which is present in human semen at a concentration of 12 mM, is made from glucose by cells of the seminal vesicle by reduction with NADPH to the sugar alcohol D-sorbitol, which in turn is oxidized in the 2 position by NAD^+. The combination of high $[NADPH]/[NADP^+]$ and high $[NAD^+]/[NADH]$ ratio is sufficient to shift the equilibrium far toward fructose formation.[d]



The polyol pathway is an active bypass of the dominant glycolysis pathway in many organisms.[e] Sorbitol and other polyols such as glycerol, erythritol, threitol, and ribitol serve as cryoprotectants in plants, insects, and other organisms.[f] Sorbitol is also an important osmolyte in some organisms (see Box 20-C). On the other hand, accumulation of sorbitol in lenses of diabetic individuals has often been blamed for development of cataract. However, doubts have been raised about this conclusion. The polyol pathway is more active than normal in diabetes, and there is evidence that the increased flow in this pathway may lead to an increase in oxidative damage to the lens. This may result, in part, from the depletion of NADPH needed for reduction of oxidized glutathione in the antioxidant system.[g] Aldose reductase inhibitors, which reduce the rate of sorbitol formation, decrease cataract formation. However, the reason for this is not yet clear.[h]

---

[a] McGilvery, R. W. (1970) *Biochemistry, A Functional Approach*, Saunders, Philadelphia, Pennsylvania (pp. 631–632)

[b] Hers, H. G. (1960) *Biochim. Biophys. Acta.* **37**, 127–

[c] Gitzelmann, R., Steinmann, B., and Van den Berghe, G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 905–935, McGraw-Hill, New York

[d] Prendergast, F. G., Veneziale, C. M., and Deering, N. G. (1975) *J. Biol. Chem.* **250**, 1282–1289

[e] Luque, T., Hjelmqvist, L., Marfany, G., Danielsson, O., El-Ahmad, M., Persson, B., Jörnvall, H., and Gonzàlez-Duarte, R. (1998) *J. Biol. Chem.* **273**, 34293–34301

[f] Podlasek, C. A., and Serianni, A. S. (1994) *J. Biol. Chem.* **269**, 2521–2528

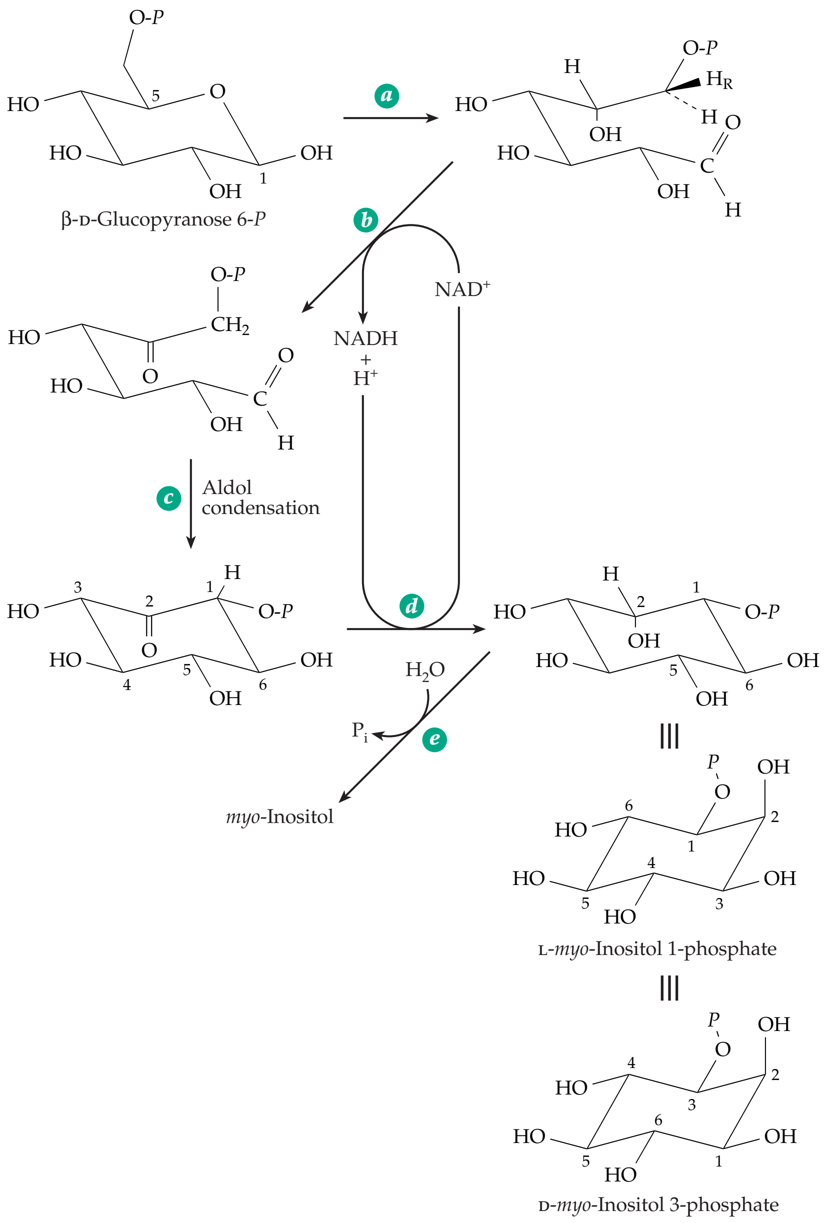
[g] Lee, A. Y. W., and Chung, S. S. M. (1999) *FASEB J.* **13**, 23–30

[h] Srivastava, S., Watowich, S. J., Petrash, J. M., Srivastava, S. K., and Bhatnagar, A. (1999) *Biochemistry* **38**, 42–54

NAD^+. In addition to the two redox steps (Eq. 20-2, *b* and *d*), this enzyme catalyzes both the conversion of the β anomer of glucose 6-*P* to the open-chain aldehyde form and the internal aldol condensation of Eq. 20-2, step *c*.[18-19b] The pro-*R* hydrogen at C-6 of glucose 1-*P* is lost in step *b* while the pro-*S* hydrogen is retained.[20] The ring numbering system is different for glucose and for the inositols, C-5 of glucose 1-*P* becoming C-2 of L-*myo*-inositol. Since *myo*-inositol contains a plane of symmetry D- and L- forms are identical. However, they are numbered differently (Eq. 20-2). The phosphoinositides and inositol polyphosphates are customarily numbered as derivatives of D-*myo*-inositol.

Synthesis of inositol by animals is limited and *myo*-inositol is sometimes classified as a vitamin. Mice grow poorly and lose some of their hair if deprived of dietary inositol. Various phosphate esters of inositol occur in nature. For example, large amounts of the hexaphosphate (**phytic acid**) are present in grains, usually as the calcium or mixed Ca^{2+}–Mg^{2+} salts known as **phytin**. The two apical cells of the 28-cell larvae of mesozoa (Fig. 1-12A) contain enough magnesium phytate in granular form to account for up to half of the weight of the larvae.[21] Inositol pentaphosphate is an allosteric activator for hemoglobin in birds and turtles (p. 358). Di-*myo*-inositol-1,1'-phosphate is an osmolyte in some hypothermophilic archaea.[19a] Inositol is a component of **galactinol**, the β glycoside of D-galactose with inositol (Eq. 20-15). Galactinol, as well as free inositol, circulates in human blood and in plants and may be a precursor of cell wall polysaccharides. However, in our own bodies the greatest importance for inositol doubtless lies in the inositol-containing phospholipids known as **phosphoinositides** (Figs. 8-2, 11-9, 21-5). Their function in generation of "second messengers" for various hormones is dealt with in Chapters 11 and 21.

A person typically ingests daily about one gram of inositol, some in the free form, some as phosphoinositides, and some as phytin. As much as four grams of inositol per day may be synthesized in the kidneys.[22] Breast milk is rich in inositol and dietary supplementation with inositol has increased survival of premature infants with respiratory distress syndrome.[22] The action of insulin is reported to be improved by administration of D-*chiro*-inositol (p. 998) to women with polycystic ovary syndrome.[22a]



(20-2)

3. D-Glucuronic Acid, Ascorbic Acid, and Xylitol

In bacteria, as well as in animal kidneys,[23] inositol may be converted to D-glucuronic acid (Fig. 20-1) with the aid of an oxygenase. Free glucuronic acid may also be formed by animals from glucose or from UDP-glucose (Fig. 20-2). Within the animal body glucuronic acid can be reduced with NADH (Fig. 20-2, step *a*) to yield **L-gulonic acid**, an aldonic acid that could also be formed by oxidation at the aldehyde end of the sugar **gulose**. Because C-6 of the glucuronic acid has become C-1 of gulonic acid, the latter belongs to the L family of

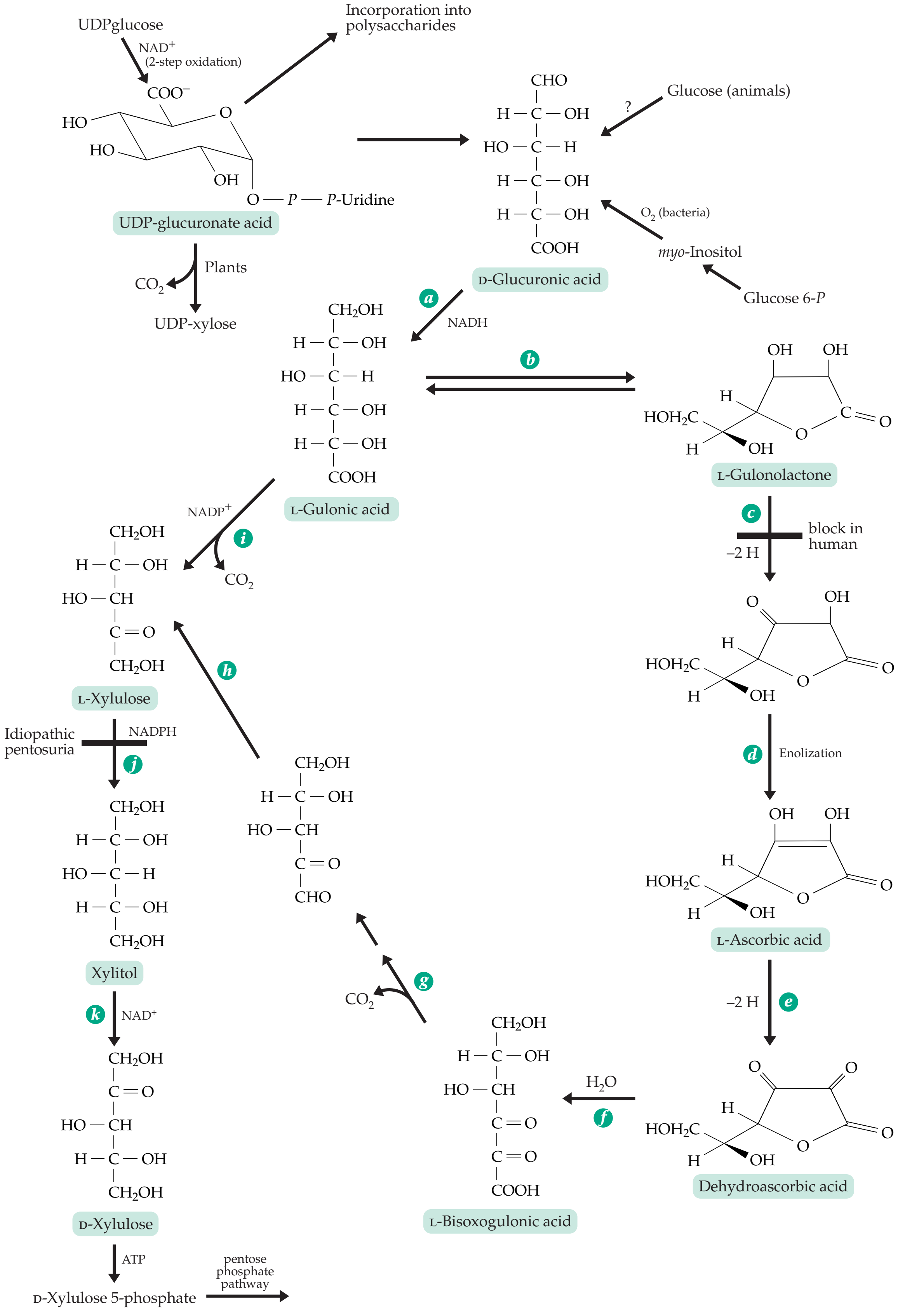
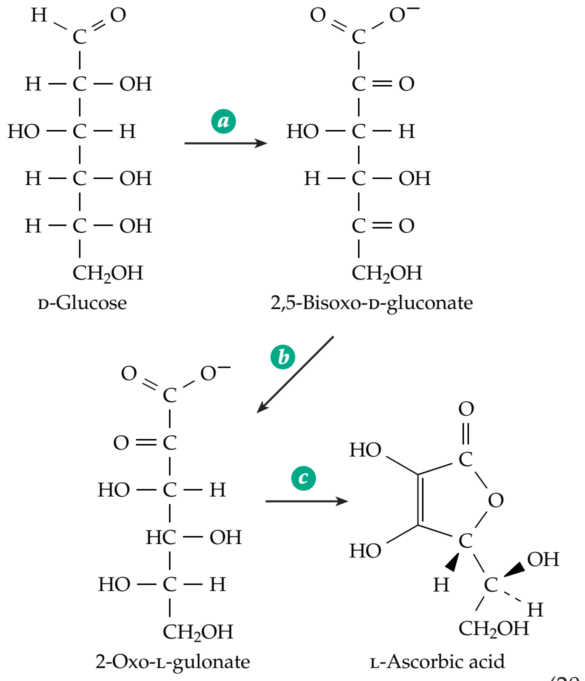

UDPglucose
NAD+
(2-step oxidation)
Incorporation into polysaccharides
UDP-glucuronate acid
O — P — P-Uridine
Plants
CO2
UDP-xylose
D-Glucuronic acid
Glucose (animals)
?
O2 (bacteria)
myo-Inositol
Glucose 6-P
a
NADH
L-Gulonic acid
b
L-Gulonolactone
c
block in human
–2 H
d
Enolization
L-Ascorbic acid
e
–2 H
Dehydroascorbic acid
f
H2O
L-Bisoxogulonic acid
g
CO2
h
i
NADP+
CO2
L-Xylulose
Idiopathic pentosuria
NADPH
j
Xylitol
k
NAD+
D-Xylulose
ATP
D-Xylulose 5-phosphate
pentose phosphate pathway

Figure 20-2 Some pathways of metabolism of D-glucuronic acid and of ascorbic acid, vitamin C.

sugars. Gulonic acid can be converted to a cyclic lactone (step *b*) which, in a two-step process involving dehydrogenation and enolization (steps *c* and *d*), is converted to **L-ascorbic acid**. This occurs in most higher animals.[24] However, the dehydrogenation step is lacking in human beings and other primates, in the guinea pig, and in a few other species. One might say that we and the guinea pig have a genetic defect at this point which obliges us to eat relatively large quantities of plant materials to satisfy our bodily needs for ascorbic acid (see Box 18-D). Gulonolactone oxidase is one of the enzymes containing covalently bound 8α-(N^1-histidyl)riboflavin.[25] The defective human gene for this enzyme has been identified, isolated, and sequenced. It is found to have accumulated a large number of mutations, which have rendered it inactive and now only a pseudogene.[26] Mice with an inactivated gulonolactone oxidase have a dietary requirement for vitamin C similar to that of humans. They suffer severe vascular damage on diets marginal in ascorbic acid.[26a] Even in rodents Na^+-dependent ascorbic acid transporters are present in metabolically active tissues to bring the vitamin from the blood into cells.[26b]

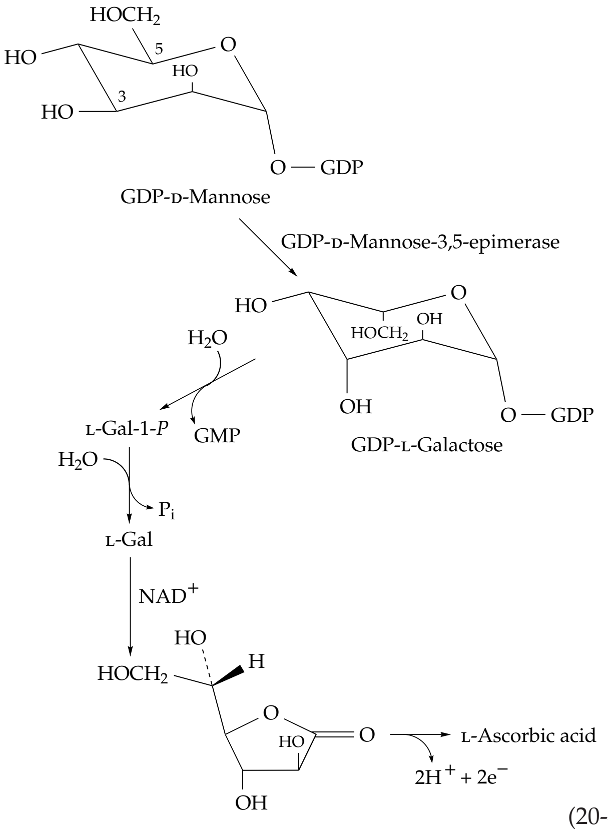
A clever bit of genetic engineering has permitted the conversion of D-glucose to 2-oxo-L-gulonate in the enzymatic sequence of Eq. 20-3, *a*, *b*.



(20-3)

The bacterium *Erwinia herbicola* naturally has the ability to oxidize glucose to 2,5-bisoxo-D-gluconate (Eq. 20-3, step *a*) but cannot carry out the next step, the stereospecific reduction to 2-oxo-L-gulonate. However, a gene encoding a suitable reductase was isolated from a genomic library from *Corynebacterium*. The cloned gene was fused to an *E. coli trp* promoter (see Chapter 28) and was introduced in a multicopy plasmid into *E. herbicola*. The resultant organism can carry out both steps *a* and *b* of Eq. 20-3 leaving only step *c*, a nonenzymatic acid-catalyzed reaction, to complete an efficient synthesis of vitamin C from glucose.[27]

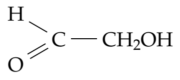
Higher plants make large amounts of L-ascorbate, which in leaves may account for 10% of the soluble carbohydrate content.[28] However, the pathway of synthesis differs from that in Fig. 20-2. Both D-mannose and L-galactose are efficient precursors. The pathway in Eq. 20-4, which starts with GDP-D-mannose and utilizes known enzymatic processes, has been suggested.[28,29] The GDP-D-mannose-3,5-epimerase is a well documented but poorly understood enzyme. Multistep mechanisms related to that of UDP-glucose 4-epimerase (Eqs. 20-1, 15-14) can be envisioned.



(20-4)

Ascorbic acid is readily oxidized to dehydroascorbic acid (Box 18-D; Fig. 20-2, step *e*), which may be hydrolyzed to L-bisoxogulonic acid (step *f*). The latter, after decarboxylation and reduction, is converted to L-xylulose (steps *g* and *h*), a compound that can also be formed by a standard oxidation and decarboxylation sequence on L-gulonic acid (step *i*). Reduction of xylulose to xylitol and oxidation of the latter with NAD^+ (steps *j* and *k*) produces D-xylulose, which can be phosphorylated with ATP and enter the pentose phosphate pathways. A metabolic variation produces a condition called **idiopathic pentosuria**. Affected individuals cannot reduce xylulose to xylitol and, hence, excrete large amounts of the pentose into the urine, especially if the diet is rich in glucuronic acid. The "defect" seems to be harmless, but the sugar in the urine can cause the condition to be mistaken for diabetes mellitus.[30]

Xylitol is as sweet as sucrose and has been used as a food additive. Because it does not induce formation of dental plaque, it is used as a replacement for sucrose in chewing gum. It appeared to be an ideal sugar substitute for diabetics. However, despite the fact that it is already naturally present in the body, ingestion of large amounts of xylitol causes bladder tumors as well as oxalate stones in rats and mice. Its use has, therefore, been largely discontinued. A possible source of the problem may lie in the conversion by fructokinase of some of the xylitol to D-xylulose 1-*P*, which can be cleaved by the xylulose 1-*P* aldolase to dihydroxyacetone *P* and glycolaldehyde.

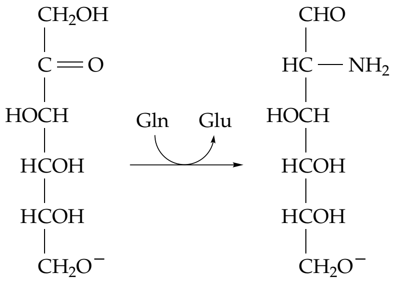


Glycolaldehyde

The latter can be oxidized to oxalate and may also be carcinogenic. As indicated in the upper left corner of Fig. 20-2, UDP-glucuronate can be decarboxylated to UDP-xylose.

4. Transformations of Fructose 6-Phosphate

Biosynthesis of **D-glucosamine 6-phosphate** is accomplished by reaction of fructose 6-*P* with glutamine (Eq. 20-5):

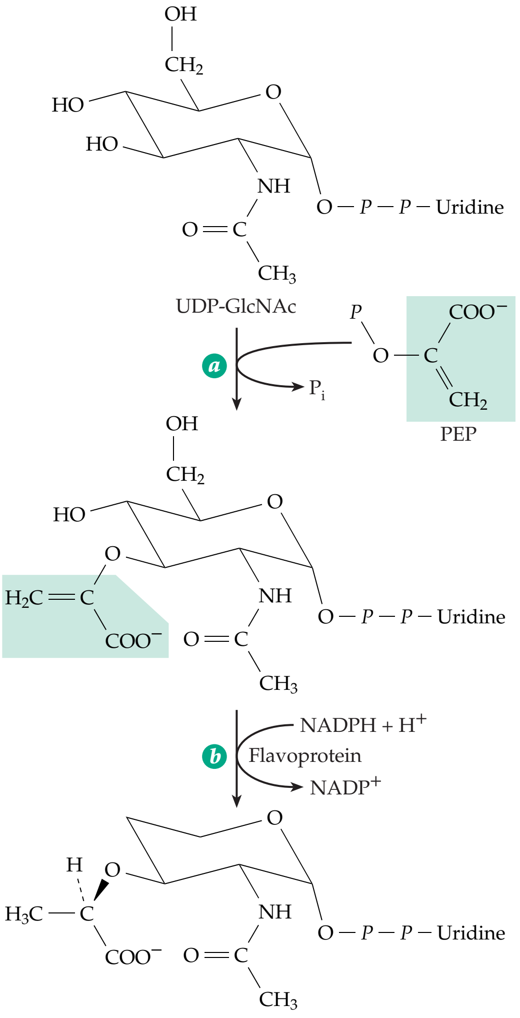


Fructose 6-*P*　D-Glucosamine 6-*P*

(20-5)

Glutamine is one of the principal combined forms of ammonia that is transported throughout the body (Chapter 24). Glucosamine 6-phosphate synthase, which catalyzes the reaction of Eq. 20-5, is an amidotransferase of the N-terminal nucleophile hydrolase superfamily (Chapter 12).[31] It hydrolyzes the amide linkage of glutamine. The released ammonia presumably reacts with the carbonyl group of fructose 6-*P* to form an imine,[32–34a] which then undergoes a reaction analogous to that catalyzed by sugar isomerases.[35] The resulting D-glucosamine 6-*P* is acetylated on its amino group by transfer of an acetyl group,[36] and a mutase moves the phospho group to form *N*-acetylglucosamine 1-*P*. In *E. coli* acetylation occurs on GlcN 1-*P* and is catalyzed by a bifunctional enzyme that also has mutase activity.[37-37b] The resulting *N*-acetylglucosamine 1-*P* is converted to UDP-*N*-acetylglucosamine (UDP-GlcNAc) with cleavage of UTP to inorganic pyrophosphate as in the synthesis of UDP-glucose (Eq. 17-56). Cells of *E. coli* are also able to catabolize glucosamine 6-phosphate. A **deaminase,** with many properties similar to those of GlcN 6-*P* synthase, catalyzes a reaction resembling the reverse of Eq. 20-5 but releasing NH_3.[38,39]

One of the compounds formed from UDP-GlcNAc is **UDP-*N*-acetylmuramic acid**. The initial step in its synthesis is an unusual type of displacement reaction on the α-carbon of PEP by the 3-hydroxyl group of the sugar (Eq. 20-6, step *a*).[40-41c] Inorganic phosphate is



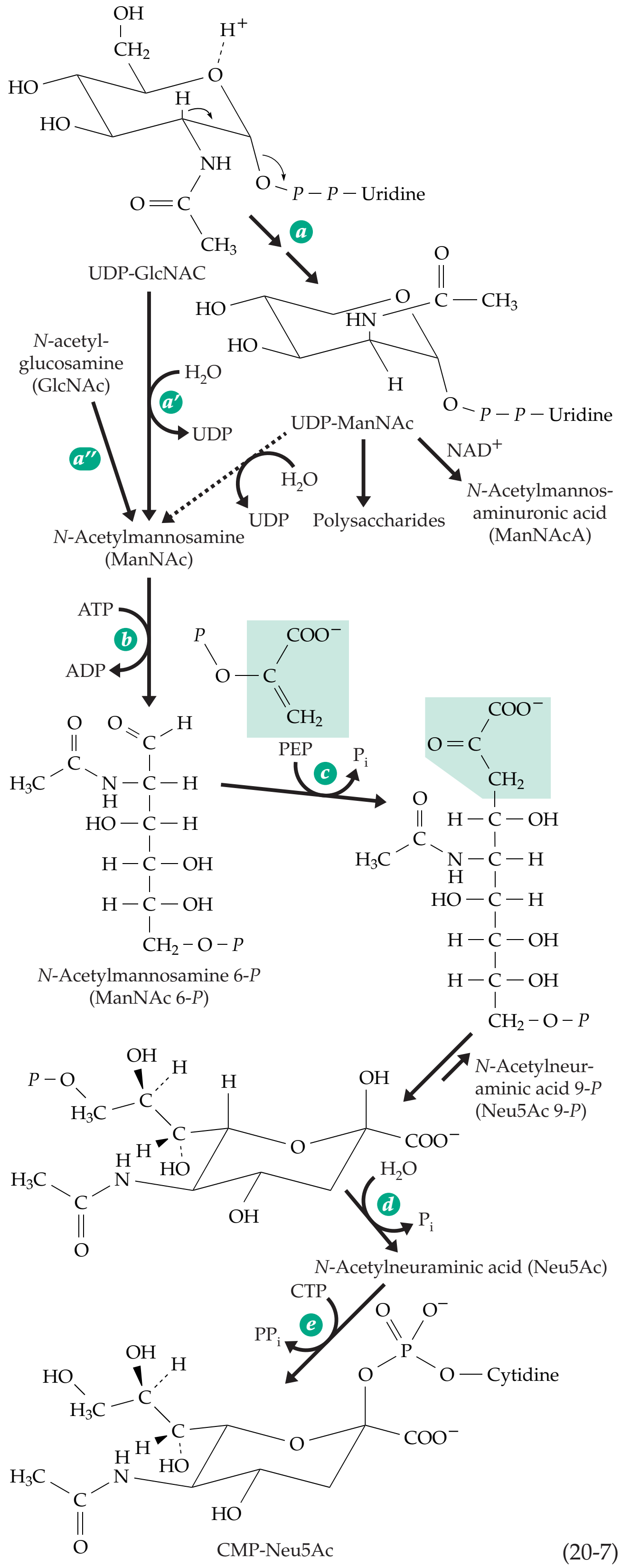
UDP-*N*-acetylmuramic acid
(UDP-MurNAc)

(20-6)

displaced with formation of an enolpyruvyl derivative of UDP-GlcNAc. This derivative is then reduced by NADPH (Eq. 20-6, step *b*).[42-43a] A second sugar nucleotide formed from UDP-GlcNAc is **UDP-*N*-acetylgalactosamine** (UDP-GalNAc), which may be created by the same 4-epimerase that generates UDP-Gal (Eq. 20-1).[44] Some animal tissues such as kidney and liver also have a **GalNAc kinase** that may salvage, for reuse, GalNAc that arises from the degradation of complex polysaccharides.[44] Bacteria may dehydrogenate UDP-GalNAc to UDP-*N*-acetylgalactosaminuric acid (UDP-GalNAcA).[44a]

UDP-GlcNAc can be converted to UDP-*N*-acetylmannosamine (UDP-ManNAc) with concurrent elimination of UDP (Eq. 20-7).[45–47b] This unusual epimerization occurs without creation of an adjacent carbonyl group that would activate the 2-H for removal as a proton. As indicated by the small arrows in Eq. 20-7, step *a′*, the UDP is evidently eliminated. In a bacterial enzyme it remains in the E–S complex and is returned after a conformational change involving the acetamido group. This allows the transient C1–C2 double bond to be protonated from the opposite side (Eq. 20-7, step *a*).[47] In bacteria the UDP-ManNAc may be dehydrogenated to UDP-*N*-acetylmannos-aminuronic acid (ManNAcA). Both ManNAc and ManNAcA are components of bacterial capsules.[47]

In mammals the epimerase (Eq. 20-7, step *a′*) probably utilizes a similar chemical mechanism but eliminates UDP and replaces it with HO^- to give free *N*-acetylmannosamine, which is then phosphorylated on the 6-hydroxyl (Eq. 20-7, step *b*). ManNAc may also be formed from free GlcNAc by another 2-epimerase (step a′′).[47c,d]

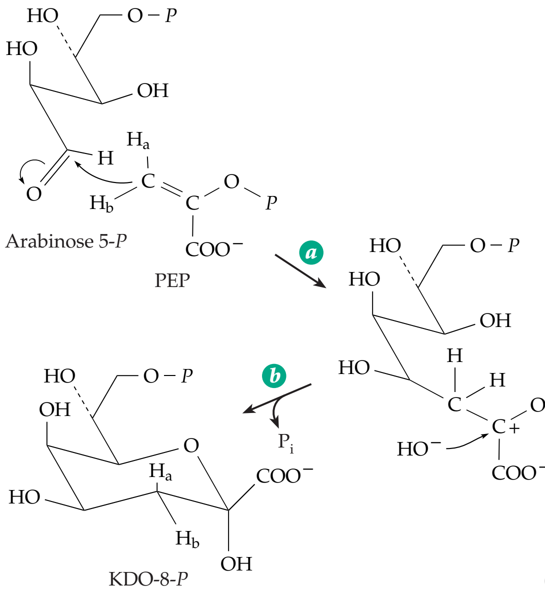

UDP-GlcNAC
a
UDP-ManNAc
N-acetyl-glucosamine (GlcNAc)
a′
a′′
H2O
UDP
NAD+
N-Acetylmannos-aminuronic acid (ManNAcA)
Polysaccharides
N-Acetylmannosamine (ManNAc)
ATP
b
ADP
PEP
c
Pi
N-Acetylmannosamine 6-P (ManNAc 6-P)
N-Acetylneuraminic acid 9-P (Neu5Ac 9-P)
d
N-Acetylneuraminic acid (Neu5Ac)
CTP
e
PPi
Cytidine
CMP-Neu5Ac

(20-7)

5. Extending a Sugar Chain with Phosphoenolpyruvate (PEP)

The six-carbon chain of ManNAc 6-*P* can be extended by three carbon atoms using an aldol-type condensation with a three-carbon fragment from PEP (Eq. 20-7, step *c*) to give ***N*-acetylneuraminic acid** (sialic acid).[48] The nine-carbon chain of this molecule can cyclize to form a pair of anomers with 6-membered rings as shown in Eq. 20-7. In a similar manner, arabinose 5-*P* is converted to the 8-carbon **3-deoxy-D-*manno*-octulosonic acid** (**KDO**) (Fig. 4-15), a component of the lipopolysaccharide of gram-negative bacteria (Fig. 8-30), and D-Erythrose 4-*P* is converted to 3-deoxy-D-*arabino*-heptulosonate 7-*P*, the first metabolite in the shikimate pathway of aromatic synthesis (Fig. 25-1).[48a] The arabinose-*P* used for KDO synthesis is formed by isomerization of D-ribulose 5-*P* from the pentose phosphate pathway, and erythrose 4-*P* arises from the same pathway.

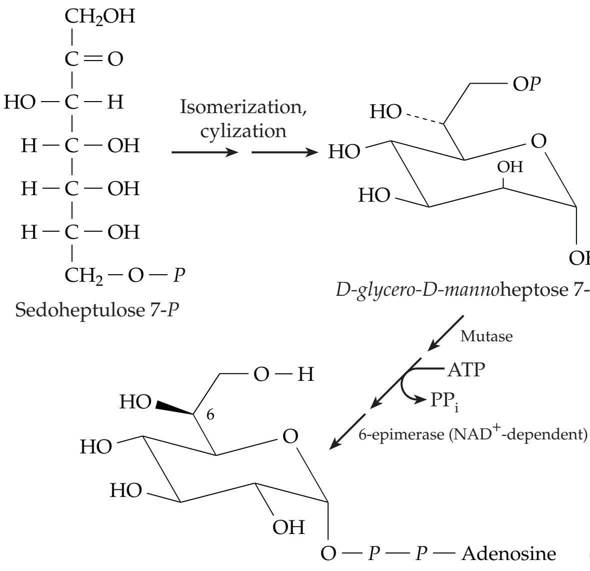
The mechanism of the aldol condenstion that forms these sugars is somewhat unexpected. A reactive enolate anion can be formed from PEP by hydrolytic attack on the phospho group with cleavage of the O–P bond. However, in reactions such as step *a* of



(20-8)

Eq. 20-6, step *c* of Eq. 20-7,and also in EPSP synthase (Eq. 25-4) the initial condensation does not involve O–P cleavage. NMR studies of the action of KDO synthase reveal that the C–O bond of PEP is cleaved as is indicated in Eq. 20-8.[49-52b] The *si* face of PEP faces the *re* face of the carbonyl group of the sugar phosphate. A carbanionic center is generated at C-3 of PEP with possible participation of the phosphate oxygen as well as electrostatic stabilization of the carbocation formed in step *a*. Ring closure (step *b*) occurs with loss of P_i. The immediate product of the aldol condensation, in Eq. 20-7, is *N*-acetylneuraminic acid 9-phosphate, which is cleaved through phosphatase action (step *d*) and is activated to the CMP derivative by reaction with CTP (Eq. 20-7,step *e*).[52c] Further alterations may occur. For example, CMP-Neu5Ac is hydroxylated to form CMP-*N*-glycolylneuraminic acid.[53] Furthermore, an additional type of sialic acid, 2-oxo-3-deoxy-D-*glycero*-D-*galacto*-nononic acid (**KDN**), has been found in human developmentally regulated glycoproteins and also in many other organisms.[54-55a] It has an –OH group in the 5-position rather than the acetamido group of the other sialic acids. Like NeuNAc it is activated by reaction with CTP forming CMP-KDN. These activated monosaccharides differ from most others in being derivatives of a CMP rather than of CDP. More than 40 different naturally occuring variations of sialic acid have been identified.[55b]

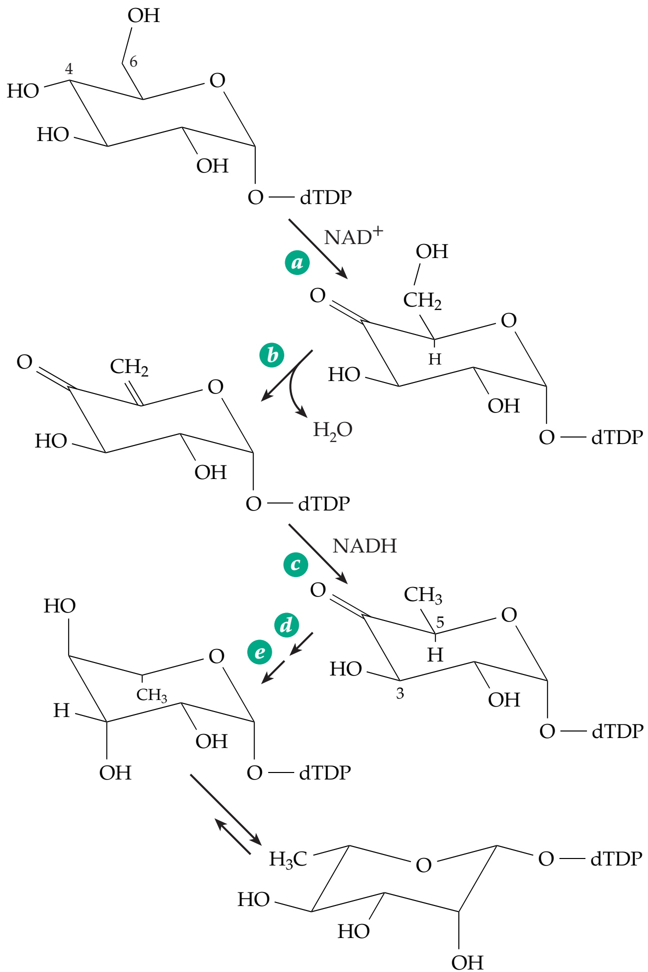
In a similar fashion, KDO is converted to the β-linked **CMP-KDO**,[56-56b] which is incorporated into lipid A as shown in Fig. 20-10. The ADP derivative of the **L-*glycero*-D-*manno*-heptose** (Fig. 4-15), which is also present in the lipopolysaccharide of gram-negative bacteria, is formed from sedoheptulose 7-*P* in a five-step process (Eq. 20-9).[57-58b]


Sedoheptulose 7-P
Isomerization, cylization
D-glycero-D-mannoheptose 7-P
Mutase
ATP
PP_i
6-epimerase (NAD$^+$-dependent)
O — P — P — Adenosine

(20-9)

6. Synthesis of Deoxy Sugars

Metabolism of sugars often involves dehydration to α,β-unsaturated carbonyl compounds. An example is the formation of 2-oxo-3-deoxy derivatives of sugar acids (Eq. 14-36). Sometimes a carbonyl group is

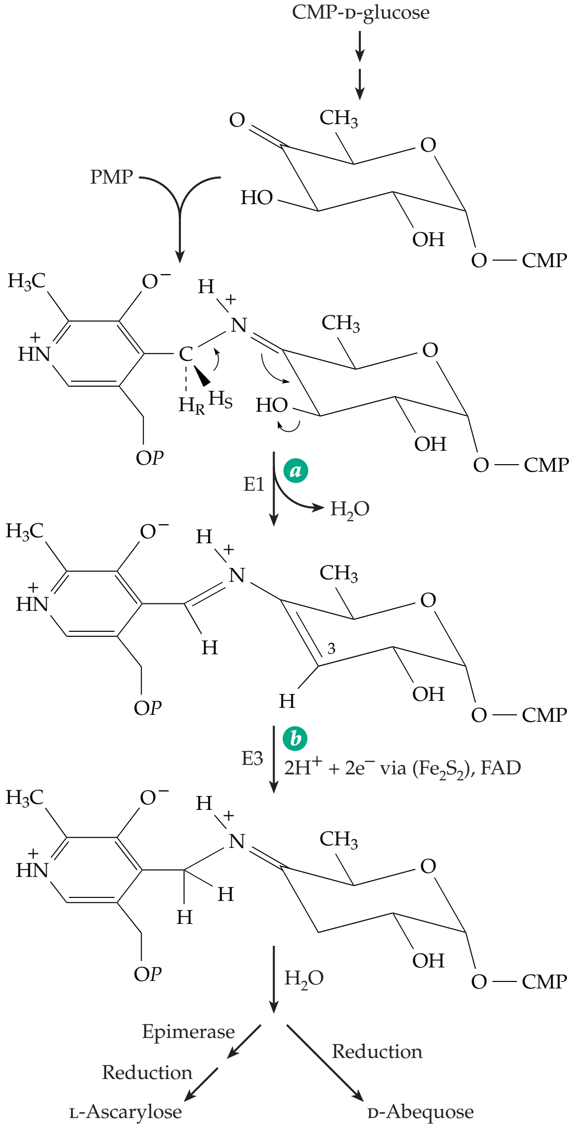


(20-10)

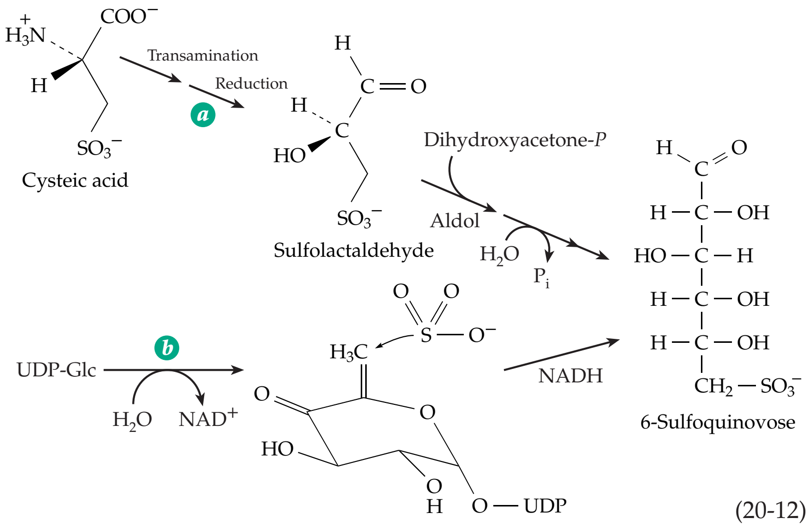
created by oxidation of an –OH group, apparently for the sole purpose of promoting dehydration. For example, the biosynthesis of L-rhamnose from D-glucose is a multistep process (Eq. 20-10) that takes place while the sugars are attached to deoxythymidine diphosphate.[59,59a,b] Introduction of the carbonyl group by dehydrogenation with tightly bound NAD^+ (Eq. 20-10, step *a*) is followed by dehydration (step *b*).[59c,d] To complete the sequence, the double bond formed by dehydration is reduced (step *c*) by the NADH produced in step *a*. A separate enzyme, a 3,5-epimerase catalyzes inversion at both C-3 and C-5 (step *d*).[59e] Finally, a third enzyme is needed for a second reduction (step *e*) using NADPH.[59f] The biosynthesis of **GDP-L-fucose** from GDP-D-mannose occurs by a parallel sequence.[60-61b]

The metabolism of free L-fucose (6-deoxy-L-galactose), which is present in the diet and is also generated by degradation of glycoproteins, resembles the Entner–Doudoroff pathway of glucose metabolism (Eq. 17-18). Similar degradative pathways act on D-arabinose and L-galactose.[60]

Bacterial surface polysaccharides contain a variety of dideoxy sugars. The four 3,6-dideoxy sugars **D-paratose** (3,6-dideoxy-D-glucose), **D-abequose** (3,6-dideoxy-D-galactose), **D-tyvelose** (3,6-dideoxy-D-mannose), and **L-ascarylose** (3,6-dideoxy-L-mannose), whose structures are shown in Fig. 4-15, arise from CDP-glucose.[60a] This substrate is first converted, in reactions parallel to the first three steps of Eq. 20-10, to 4-oxo-6-deoxy-CDP-glucose which reacts in two steps with pyridoxamine 5'-phosphate (PMP) and NADH (Eq. 20-11). This unusual reaction[62–65] is catalyzed by a two-enzyme complex. The first component, E1, catalyzes the formation of a Schiff base of the substrate with PMP and a transamination, which also accomplishes dehydration, to give an unsaturated sugar ring (Eq. 20-11, step *a*). The protein also contains an Fe_2S_2 center suggesting a possible one-electron transfer. The second component, E_3, contains both an Fe_2S_2 plant type ferredoxin center and bound FAD.[65] Observation by EPR spectroscopy revealed accumulation of an organic free radical[64] that may be an intermediate in step *b* of Eq. 20-11. Hydrolysis, epimerization at C-5, and reduction yields L-ascarylose. A similar reaction sequence without the last epimerization would yield D-abequose. CDP-D-tyvelose arises by C-2 epimerization of CDP-D-paratose.[65a] Other unusual sugars[66–68] are formed from intermediates in Eq. 20-11. One is a **3-amino-3,4,6-trideoxyhexose** in which the amino group has been provided by transamination[67] (see also Box 20-B).


CMP-D-glucose
O
CH3
O
HO
OH
O — CMP
PMP
H3C
O−
H
+
N
CH3
O
+
HN
C
HR HS
HO
OH
O — CMP
OP
E1
a
H2O
H3C
O−
H
+
N
CH3
O
+
HN
H
3
H
OH
O — CMP
OP
E3
b
2H+ + 2e− via (Fe2S2), FAD
H3C
O−
H
+
N
CH3
O
+
HN
H
H
OH
O — CMP
OP
H2O
Epimerase
Reduction
Reduction
L-Ascarylose
D-Abequose

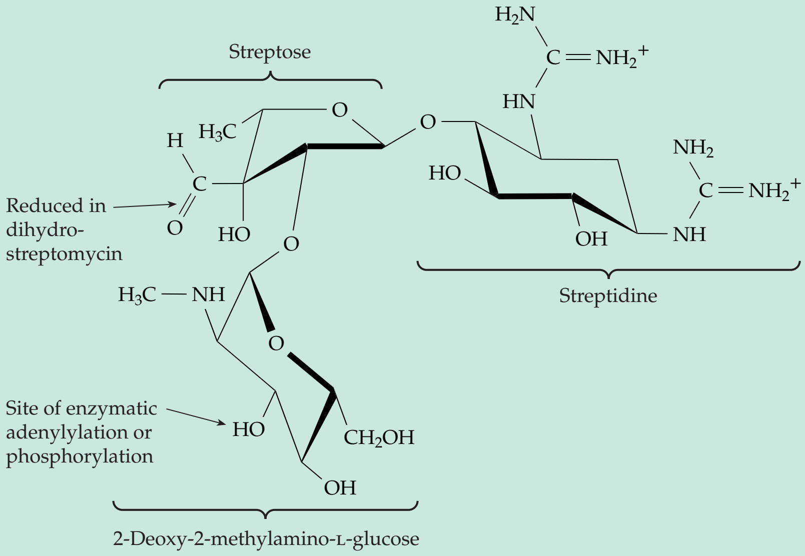
(20-11)


+
H3N
COO−
H
SO3−
Cysteic acid
Transamination
Reduction
a
H
C=O
H
C
HO
SO3−
Sulfolactaldehyde
Dihydroxyacetone-P
Aldol
H2O
Pi
H
C
O
H — C — OH
HO — C — H
H — C — OH
H — C — OH
CH2 — SO3−
6-Sulfoquinovose
UDP-Glc
b
H2O
NAD+
O
O
S
O−
H3C
O
O
HO
O
H
O — UDP
NADH

(20-12)

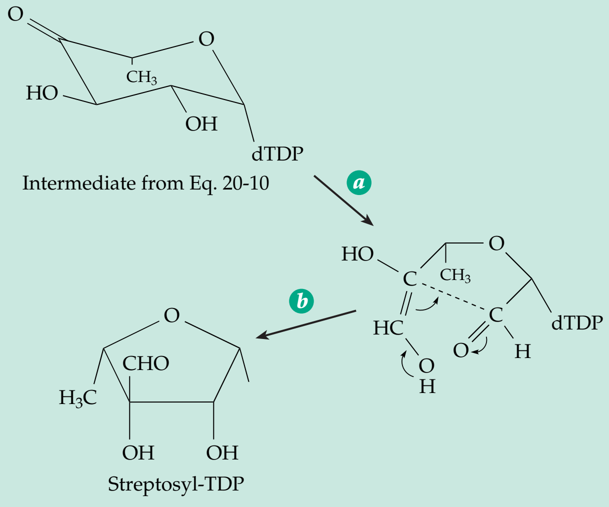
The unusual sulfur-containing sugar **6-sulfoquinovose** is present in

BOX 20-B THE BIOSYNTHESIS OF STREPTOMYCIN


Streptose
Streptidine
Reduced in dihydro-streptomycin
Site of enzymatic adenylylation or phosphorylation
2-Deoxy-2-methylamino-L-glucose

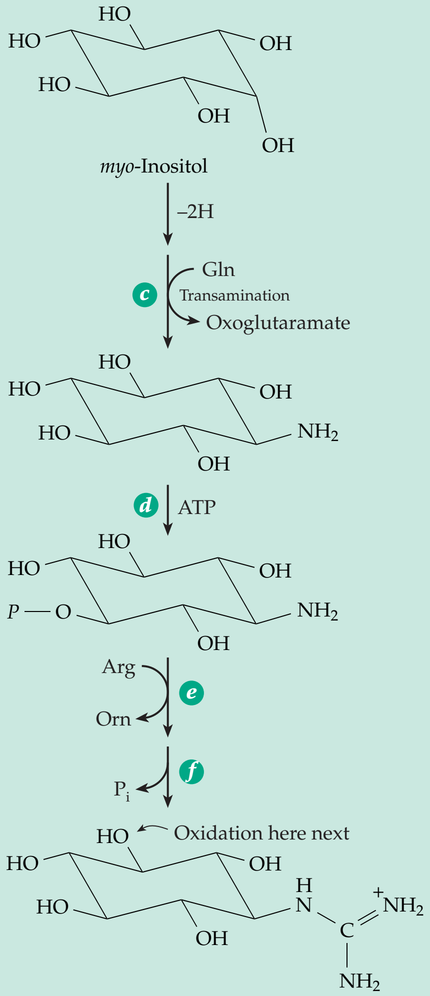
Streptomycin

Streptomycin, the kanamycins, neomycins, and gentamycins form a family of medically important **aminoglycoside antibiotics**.[a] They are all water-soluble basic carbohydrates containing three or four unusual sugar rings. D-Glucose is a precursor of streptomycin, all three rings being derived from it. While the route of biosynthesis of 2-deoxy-2-methylamino-L-glucose is not entirely clear, the pathways to L-streptose and streptidine, the other two rings, have been characterized.[b–d] The starting material for streptidine synthesis is a nucleoside diphosphate sugar, which is an intermediate in the synthesis of L-rhamnose (Eq. 20-10). The carbon–carbon chain undergoes an aldol cleavage as shown in step *a* of the following equation:


Intermediate from Eq. 20-10
a
b
dTDP
Streptosyl-TDP

The ring-open product is written here as an enediol, which is able to recyclize in an aldol condensation (step *b*) to form a five-membered ring with a branch at C-3. The L-streptosyl nucleoside diphosphate formed in this way serves as the donor of streptose to streptomycin.

The basic cyclitol streptidine is derived from *myo*-inositol, which has been formed from glucose 6-*P* (Eq. 20-2). The guanidino groups are introduced by oxidation of the appropriate hydroxyl group to a carbonyl group followed by transamination from a specific amino donor. In the first step, illustrated by the following equation, glutamine is the amino donor for the transamination, the oxoacid product being α-oxoglutaramic acid.


myo-Inositol
–2H
c
Gln
Transamination
Oxoglutaramate
d
ATP
e
Arg
Orn
f
Pi
Oxidation here next

BOX 20-B THE BIOSYNTHESIS OF STREPTOMYCIN (continued)

The amino group on the ring now receives an amidine group, which is transferred from arginine by nucleophilic displacement[e] in a reaction resembling that in the synthesis of urea (see Fig. 24-10, step *h*). However, there is first a phosphorylation at the 2 position. After the amidine transfer has occurred to form the guanidino group, the phospho group is hydrolyzed off by a phosphatase. This is another phosphorylation–dephosphorylation sequence (p. 977) designed to drive the reaction to completion in the desired direction. The second guanidino group is introduced in an analogous way by oxidation at the 3 position followed by transamination, this time with the amino group being donated by alanine. Again, a phosphorylation is followed by transfer of an amidine group from arginine. The final hydrolytic removal of the phospho group (which this time is added at C-6) does not occur until the two other sugar rings have been transferred on from nucleoside diphosphate precursors to form streptomycin phosphate.

As with other antibiotics,[f–i] streptomycin is subject to inactivation by enzymes encoded by genetic resistance factors (Chapter 26). Among these are enzymes that transfer phospho groups or adenylyl groups onto streptomycin at the site indicated by the arrow in the structure.[j,k] Thus, dephosphorylation at one site generates the active antibiotic as the final step in the biosynthesis, while phosphorylation at another site inactives the antibiotic.

---

[a] Benveniste, R., and Davies, J. (1973) *Ann. Rev. Biochem.* **42**, 471–506

[b] Luckner, M. (1972) *Secondary Metabolism in Plants and Animals*, Academic Press, New York (pp.78–80)

[c] Walker, J. B., and Skorvaga, M. (1973) *J. Biol. Chem.* **248**, 2441–2446

[d] Marquet, A., Frappier, F., Guillerm, G., Azoulay, M., Florentin, D., and Tabet, J.-C. (1993) *J. Am. Chem. Soc.* **115**, 2139–2145

[e] Fritsche, E., Bergner, A., Humm, A., Piepersberg, W., and Huber, R. (1998) *Biochemistry* **37**, 17664–17672

[f] Cox, J. R., and Serpersu, E. H. (1997) *Biochemistry* **36**, 2353–2359

[g] McKay, G. A., and Wright, G. D. (1996) *Biochemistry* **35**, 8680–8685

[h] Thompson, P. R., Hughes, D. W., Cianciotto, N. P., and Wright, G. D. (1998) *J. Biol. Chem.* **273**, 14788–14795

[i] Gerratana, G., Cleleand, W. W., and Reinhardt, L. A. (2001) *Biochemistry* **40**, 2964–2971

[j] Roestamadji, J., Grapsas, I., and Mobashery, S. (1995) *J. Am. Chem. Soc.* **117**, 80–84

[k] Thompson, P. R., Hughes, D. W., and Wright, G. D. (1996) *Biochemistry* **35**, 8686–8695

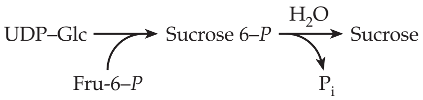
---

the sulfolipid of chloroplasts (p. 387).[69] A possible biosynthetic sequence begins with transamination of cysteic acid to 3-sulfopyruvate, reduction of the latter to sulfolactaldehyde, and aldol condensation with dihydroxyacetone-*P* as indicated in Eq. 20-12a.[70] See also Eq. 24-47 and Fig. 4-4. However, biosynthesis in chloroplasts appears to start with action of a 4,6-dehydratase on UDP-glucose followed by addition of sulfite and reduction (Eq. 20-12*b*).[70a,b] The sulfite is formed by reduction of sulfate via adenylyl sulfate (Fig. 24-25). However, biosynthesis in chloroplasts appears to start with action of a 4,6-dehydratase on UDP-glucose followed by addition of sulfite and reduction (Eq. 20-12*b*).[70a,b] The sulfite is formed by reduction of sulfate via adenylyl sulfate (Fig. 24-25).

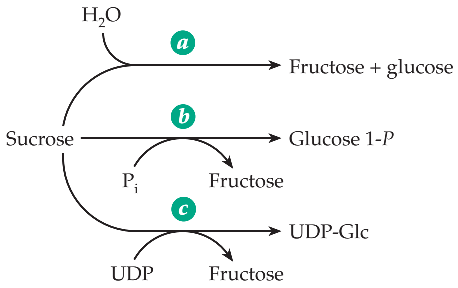
B. Synthesis and Utilization of Oligosaccharides

Our most common food sugar **sucrose** is formed in all green plants and nowhere else. It is made both in the chloroplasts and in the vicinity of other starch deposits. It serves both as a transport sugar and, dissolved within vacuoles, as an energy store. Sucrose is very soluble in water and is chemically inert because the hemiacetal groups of both sugar rings are blocked. However, sucrose is thermodynamically reactive, the glucosyl group having a group transfer potential of 29.3 kJ mol^{-1}. It is extremely sensitive toward hydrolysis catalyzed by acid. Transport of sugar in the form of a disaccharide provides an advantage to plants in that the disaccharide has a lower osmotic pressure than would the same amount of sugar in monosaccharide form.

Biosynthesis of sucrose[71,71a] utilizes both UDP-glucose and fructose 6-*P* (Eq. 20-13). Reaction of UDP-glucose with fructose can also occur to give sucrose directly.[72] Because this reaction is reversible, sucrose serves as a source of UDP-glucose for synthesis of cellulose and other polysaccharides in plants. Metabolism of sucrose in the animal body begins with the action of **sucrase** (invertase), which hydrolyzes the disaccharide to fructose and glucose (Eq. 20-14, step *a*). The same enzyme is also found in higher plants and fungi. Mammalian sucrase is one of several carbohydrases that are anchored to the external surfaces of the microvilli of the small intestines. Sucrose is bound



(20-13)



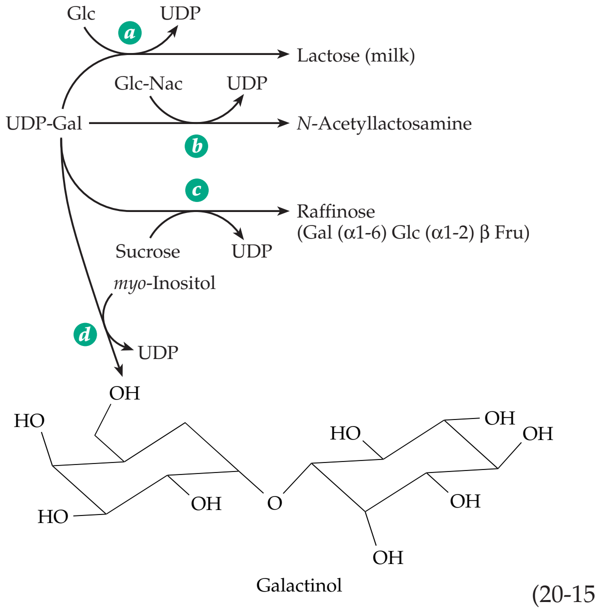
(20-14)

tightly but noncovalently to **isomaltase,** which hydrolyzes the α-1,6-linked isomaltose and related oligosaccharides. A nonpolar N-terminal segment of the isomaltase anchors the pair of enzymes to the microvillus membrane. The two-protein complex arises naturally because the two enzymes are synthesized as a single polypeptide, which is cleaved by intestinal proteases.[73,74]

Because of the relatively high group transfer potential of either the glucosyl or fructosyl parts, sucrose is a substrate for glucosyltransferases such as sucrose phosphorylase (Eq. 20-14, step *b*; see also Eq. 12-7 and associated discussion). In certain bacteria this reaction makes available the activated glucose 1-*P* which may enter catabolic pathways directly. Cleavage of sucrose for biosynthetic purposes can occur by reaction 20-14, step *c*, which yields UDP-glucose in a single step.

A disaccharide with many of the same properties as sucrose is **trehalose,** which consists of two α-glucopyranose units in 1,1 linkage (p. 168). The biosynthetic pathway from UDP-glucose and glucose 6-*P* parallels that for synthesis of sucrose (Eq. 20-13).[75,76] In *E. coli* the genes for the needed glucosyltransferase and phosphatase are part of a single operon. Its transcription is controlled in part by glucose-mediated catabolite repression (Chapter 28) and also by a repressor of the Lac family.[76,76a,77] The repressor is allosterically activated by trehalose 6-*P*, the intermediate in the synthesis. Trehalose formation in bacteria, fungi, plants, and microscopic animals is strongly induced during conditions of high osmolality (see Box 20-C).[77] Both trehalose and maltose can also be taken up via an ABC type transporter (p. 417).[77a,b]

Lactose, the characteristic sugar of milk, is formed by transfer of a galactosyl unit from UDP-galactose directly to glucose (Eq. 20-15, reaction *a*). The similar transfer of a galactosyl unit to *N*-acetylglucosamine to form *N*-acetyllactosamine (Eq. 20-15, reaction *b*) occurs in many animal tissues. An interesting regulatory mechanism is involved. The transferase catalyzing Eq. 20-15, reaction *b*, forms a complex with **α-lactalbumin** to become **lactose synthase**,[78–80b] the enzyme that catalyzes reaction *a*. Lactalbumin was identified as a milk constituent long before its role as a regulatory protein was recognized.

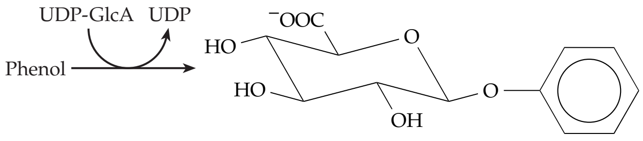


(20-15)

A very common biochemical problem is intolerance to lactose.[81] This results from the inability of the intestinal mucosa to make enough **lactase** to hydrolyze the sugar to its monosaccharide components galactose and glucose. Among most of the peoples of the earth only infants have a high lactase level, and the use of milk as a food for adults often leads to a severe diarrhea. The same is true for most animals. In fact, baby seals and walruses, which drink lactose-free milk, become very ill if fed cow's milk.

The plant trisaccharide **raffinose** arises from UDP-galactose by transfer of a galactosyl unit onto the 6-hydroxyl of the glucose ring of sucrose (Eq. 20-15, reaction *c*). Transfer of a galactosyl unit onto *myo*-inositol (Eq. 20-15, reaction *d*) produces **galactinol**, whose occurrence is widespread within the plant kingdom. Galactinol, in turn, can serve as a donor of activated galactosyl groups. Thus, many plants contain **stachyose** and higher homologs, all of which are formed by transfer of additional α-D-galactosyl units onto the 6-hydroxyl of the galactose unit of raffinose. These sugars appear to serve as antifreeze agents in the plants. The concentration of stachyose in soy beans can be as high as that of sucrose. Some seeds, e.g., those of maize, are coated with a glassy sugar mixture of sucrose and raffinose in a ratio of ~3:1.[81a]

Besides the oligosaccharides, living organisms form a great variety of glycosides that contain nonsugar components. Among these are the **glucuronides** (glucosiduronides), excretion products found in urine and derived by displacement of UDP from UDP-glucuronic acid by such compounds as phenol, benzoic acid, and sterols.[81b,c] Phenol is converted to phenyl glucuronide (Eq. 20-16), while benzoic acid (also excreted in part as hippuric acid, Box 10-A) yields an ester by the same type of displacement reaction. Many other aromatic or aliphatic compounds containing –OH, –SH, $-NH_2$, or $-COO^-$ groups also form glucuronides.[82]

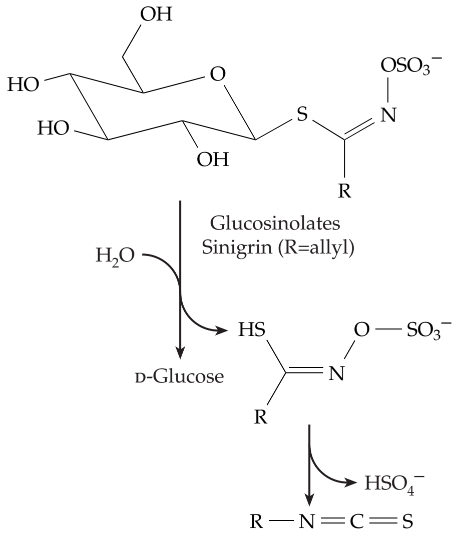


Phenyl glucuronide

(20-16)

Among these is bilirubin (Fig. 24-24). UDP-glucuronosyltransferases responsible for their synthesis are present in liver microsomes.

Among the many glycosides and glycosylamines made by plants are the anthocyanin and flavonoid pigments of flowers (Box 21-E), cyanogenic glycosides such as amygdalin (Box 25-B), and antibiotics (e.g., see Box 20-B).[83,84] Some are characteristic of certain families of plants. For example, more than 100 β-thioglucosides known as **glucosinolides** are found in the Cruciferae (cabbages, mustard, rapeseed). The compounds impart the distinctive flavors and aromas of the plants. However, some are toxic and may cause goiter or liver damage. The enzyme **myrosinase** hydrolyzes these compounds releasing isothiocyanates, thiocyanates, and nitriles (Eq. 20-17).[85-86a] L-Ascorbate acts as a cofactor for this enzyme, evidently providing a catalytic base.[86a]



(20-17)

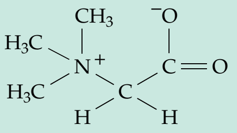
BOX 20-C OSMOTIC ADAPTATION

Bacteria, plants of many kinds, and a variety of other organisms are forced to adapt to conditions of variable osmotic pressure.[a,b] For example, plants must resist drought, and some must adapt to increased salinity. Some organisms live in saturated brine ~6 M in NaCl.[c] The **osmotic pressure** Π in dilute aqueous solutions is proportional to the total molar concentration of solute particles, c_s, as follows.

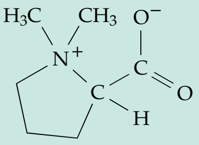
$\Pi = RTc_s$ where R is the gas constant and T the Kelvin temperature

At higher solute concentrations c_s must be replaced by the “effective molar concentration,” which is called **osmolarity** (OsM) and has units of molarity (see Record *et al.* for discussion).[b] An osmolarity difference across a membrane of 0.04 OsM results in a turgor pressure of ~ 1 atmosphere. To adapt to changes in the environmental osmolarity organisms must alter their internal solute concentrations.

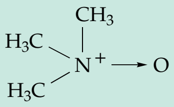
Cells of *E. coli* can adapt to at least 100-fold changes in osmolarity. Because of the porosity of the bacterial outer membrane the osmolarity of the periplasmic space is normally the same as that of the external medium. However, the inner membrane is freely permeable only to water and a few solutes such as glycerol.[b] The bacterial cells avoid loss of water when the external osmolarity is high by accumulating K^+ together with anions such as glutamate⁻ and nonionic **osmoprotectants** such as trehalose, sucrose, and oligosaccharides. *E. coli* cells will also take up other osmoprotectants such as **glycine betaine**, dimethylglycine, choline, proline, and proline betaine. Some methanogens accumulate N^ε-acetyl-β-lysine as well as glycine betaine.[d]



Betaine (glycine betaine)



Proline betaine



Trimethylamine *N*-oxide

Functioning in a somewhat different way in *E. coli* are 6- to 12-residue **periplasmic membrane-derived oligosaccharides**. These are β-1,2- and β-1,6-linked glucans covalently linked to *sn*-1-phosphoglycerol, phosphoethanolamine, or succinate (see Fig. 8-28).[b,e,f] They accumulate in the periplasm when cells are placed in a medium of low osmolarity. The resulting increased turgor in the periplasm is thought to buffer the cytoplasm against the loss of external osmolarity and to protect the periplasmic space from being eliminated by expansion of the plasma membrane. Related cyclic glucans, which are attached to *sn*-1-phosphoglycerol or *O*-succinyl ester residues, are accumulated by rhizobia.[g]

C. Synthesis and Degradation of Polysaccharides

Polysaccharides are all formed by transfer of glycosyl groups onto initiating molecules or onto growing polymer chains. The initiating molecule is usually a glycoprotein. However, let us direct our attention first to the growth of polysaccharide chains. The glycosyl are transferred by the action of glycosyltransferases from substrates such as UDP-glucose, other sugar nucleotides, and sometimes sucrose. The glycosyltransferases act by mechanisms discussed in Chapter 12 and are usually specific with respect both to substrate structure and to the type of linkage formed.

1. Glycogen and Starch

The bushlike glycogen molecules grow at their numerous nonreducing ends by the transfer of glucosyl units from UDP-glucose (Eq. 17-56)[87,87a] or in bacteria from ADP-glucose.[88–90] Utilization of glycogen by the cell involves removal of glucose units as glucose 1-*P* by the action of glycogen phosphorylase. The combination of growth and degradation from the same chain ends provides a means of rapidly storing and utilizing glucose units. The synthesis and breakdown of glycogen in mammalian muscle (Fig. 11-4) involves one of the first studied[91] and best known metabolic control systems. Various aspects have been discussed in Chapters 11, 12, and 17. The mechanism[92] and regulatory features[93–96b] have been described. An important recent development is the observation of glycogen concentrations in human muscles *in vivo* with ^{13}C NMR. This can be coupled with observation of glucose 6-*P* by ^{31}P NMR. The concentration of the latter is ~ 1 mM but increases after intense exercise.[94]

Glycogen phosphorylase and glycogen synthase alone are insufficient to synthesize and degrade glycogen. Synthesis also requires the action of the **branching enzyme** amylo-(1,4 → 1,6-transglycosylase,[97]

BOX 20-C (continued)

Fungi, green algae, and higher plants more often accumulate glycerol,[h,i] sorbitol, sucrose,[j] trehalose,[k] or proline.[a,l,m] These compounds are all "compatible solutes" which tend not to disrupt cellular structure.[n] Betaines and proline are especially widely used by a variety or organisms. How is it then that some desert rodents, some fishes, and other creatures accumulate **urea**, a well-known protein denaturant? The answer is that they also accumulate methylamine or trimethylamine *N*-oxide in an approximately 2:1 ratio of urea to amine. The mixture of compounds is compatible, the stabilizing effects of the amines offsetting the destabilizing effect of urea.[c,o]

Adaptation to changes in osmotic pressure involves sensing and signaling pathways that have been partially elucidated for *E. coli*[p] and yeasts.[i,q] Major changes in structure and metabolism may result. For example, in *E. coli* the outer membrane porin OmpF (Fig. 8-20) is replaced by OmpC (osmoporin), which has a smaller pore.[r]

A "resurrection plant" that normally contains an unusual 2-octulose converts this sugar almost entirely into sucrose when desiccated. This is one of a small group of plants that are able to withstand severe desiccation but can, within a few hours, reverse the changes when rehydrated.[j]

---

[a] Le Rudulier, D., Strom, A. R., Dandekar, A. M., Smith, L. T., and Valentine, R. C. (1984) *Science* **224**, 1064–1068

[b] Record, M. T., Jr., Courtenay, E. S., Cayley, D. S., and Guttman, H. J. (1998) *Trends Biochem. Sci.* **23**, 143–148

[c] Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D., and Somero, G. N. (1982) *Science* **217**, 1214–1222

[d] Sowers, K. R., Robertson, D. E., Noll, D., Gunsalus, R. P., and Roberts, M. F. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9083–9087

[e] Kennedy, E. P. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 672–679, Am. Soc. for Microbiology, Washington,DC

[f] Fiedler, W., and Rotering, H. (1988) *J. Biol. Chem.* **263**, 14684–14689

[g] Weissborn, A. C., Rumley, M. K., and Kennedy, E. P. (1991) *J. Biol. Chem.* **266**, 8062–8067

[h] Ben-Amotz, A., and Avron, M. (1981) *Trends Biochem. Sci.* **6**, 297–299

[i] Davenport, K. R., Sohaskey, M., Kamada, Y., Levin, D. E., and Gustin, M. C. (1995) *J. Biol. Chem.* **270**, 30157–30161

[j] Bernacchia, G., Schwall, G., Lottspeich, F., Salamini, F., and Bartels, D. (1995) *EMBO J.* **14**, 610–618

[k] Dijkema, C., Kester, H. C. M., and Visser, J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 14–18

[l] García-Ríos, M., Fujita, T., LaRosa, P. C., Locy, R. D., Clithero, J. M., Bressan, R. A., and Csonka, L. N. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8249–8254

[m] Verbruggen, N., Hua, X.-J., May, M., and Van Montagu, M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8787–8791

[n] Higgins, C. F., Cairney, J., Stirling, D. A., Sutherland, L., and Booth, I. R. (1987) *Trends Biochem. Sci.* **12**, 339–344

[o] Lin, T.-Y., and Timasheff, S. N. (1994) *Biochemistry* **33**, 12695–12701

[p] Racher, K. I., Voegele, R. T., Marchall, E. V., Culham, D. E., Wood, J. M., Jung, H., Bacon, M., Cairns, M. T., Ferguson, S. M., Liang, W.-J., Henderson, P. J. F., White, G., and Hallett, F. R. (1999) *Biochemistry* **38**, 1676–1684

[q] Shiozaki, K., and Russell, P. (1995) *EMBO J.* **14**, 492–502

[r] Kenney, L. J., Bauer, M. D., and Silhavy, T. J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8866–8870

an enzyme with dual specificity. After the chain ends attain a length of about ten glucose units, the branching enzyme attacks a 1,4-glycosidic linkage somewhere in the chain. Acting much as does a hydrolase, it forms a glycosyl enzyme or a stabilized carbocation intermediate. The enzyme does not release the severed chain fragment but transfers it to another nearby site on the glycogen molecule. There the enzyme rejoins the bound oligosaccharide chain that it carries to a free 6-hydroxyl group of the glycogen creating a new branch attached in α-1,6-linkage. Degradation of glycogen requires **debranching** after the long nonreducing ends of the polysaccharide have been shortened until only four glycosyl residues remain at each branch point. This is accomplished by **amylo-1,6-glucosidase / 4- α-glucanotransferase**. This 165-kDa bifunctional enzyme transfers a trisaccharide unit from each branch end to the main chain and also removes hydrolytically the last glucosyl residue at each branch point.[98-99a]

How are new glycogen molecules made? There is some evidence that a 37-kDa protein primer **glycogenin** is needed to initiate their formation.[100-101a] Thus, glycogen synthesis may be analogous to that of the glycosaminoglycans considered in Section D,1. Muscle glycogenin is a self-glycosylating protein, which catalyzes attachment of ~7 to 11 glucose units in α–1,4 linkage to the hydroxyl group of Tyr 194. The glucose units are added one at a time and when the chain is long enough it becomes a substrate for glycogen synthase.[100,102] The role of glycogenin in liver has been harder to demonstrate,[103] but a second glycogenin gene, which is expressed in liver, has been identified.[104] Genes for several glycogenins or glycogenin-like proteins have been identified in yeast, *Caenorhabditis elegans*, and *Arabidopsis*.[101a,105]

In contrast to animals, bacteria such as *E. coli* synthesize glycogen via ADP-glucose rather than UDP-glucose.[88] ADP-glucose is also the glucosyl donor for synthesis of starch in plants. The first step in the biosynthesis (Eq. 20-18) is catalyzed by the enzyme ADP-glucose pyrophosphorylase (named for the reverse reaction).

$$\text{Glucose 1-}P + \text{ATP} \rightarrow \text{ADP-glucose} + \text{PP}_i \qquad (20\text{-}18)$$

In bacteria this enzyme is usually inhibited by AMP and ADP and activated by glycolytic intermediates such as fructose 1,6-P_2, fructose 6-P, or pyruvate. In higher plants, green algae, and cyanobacteria the enzyme is usually activated by 3-phosphoglycerate, a product of photosynthetic CO_2 fixation, and is inhibited by inorganic phosphate (P_i).[106–108].

In eukaryotic plants starch is deposited within chloroplasts or in the cytoplasm as granules (Fig. 4-6) in a specifically differentiated and physically fragile plastid, the **amyloplast**.[108–110] Within the granules the starch is deposited in layers ~ 9 nm in thickness. About two-thirds of the thickness consists of nearly crystalline arrays, probably of double helical amylopectin side chains (Figs. 4-7, 4-8, 20-3) with "amorphous" segments between the layers.[111–114] In maize there are at least five starch synthases, one of which forms the straight chain amylose.[115–117] There are also at least three branching enzymes[118] and two or three debranching enzymes.[119,120] As in the synthesis of glycogen the molecules of amylopectin may grow at the many nonreducing ends. A current model, which is related to the broom-like cluster model of French (Fig. 4-7), is shown in Fig. 20-3. The branches are thought to arise, in part, by transglycosylation within the double helical strands. After branching the two chains remain in a double helix but the cut chain can now grow. Only double helical parts of strands pack well in the crystalline layer. A recent suggestion is that debranching enzymes then trim the molecule, removing single-stranded regions.[112]

The location (within the granule) of amylose, which makes up 15–30% by weight of many starches,[121] is uncertain. It may fill in the amorphous layers. It may be cut and provide primer pieces for new amylopectin molecules.[122,122a] Another possibility is that it grows by an insertion mechanism such as that portrayed for cellulose in Fig. 20-5 and is extruded inward from the membrane of the amyloplast. This mechanism might explain a puzzling question about starch. The branched amylopectin presumably grows in much the same way as does glycogen. A branching enzyme transfers part of the growing glycan chain to the $-CH_2-OH$ group of

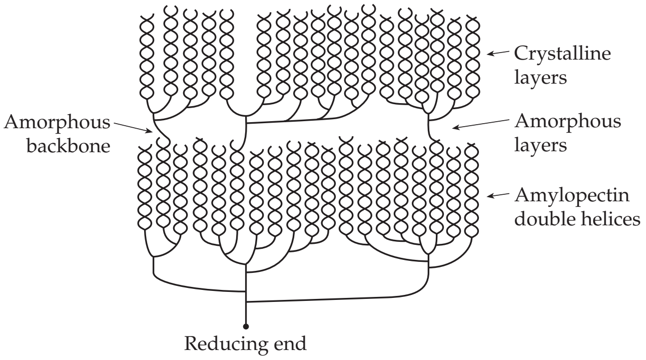


Figure 20-3 Proposed structure of a molecule of amylopectin in a starch granule. The highly branched molecule lies within 9 nm thick layers, about 2 / 3 of which contains parallel double helices of the kind shown in Fig. 4-8 in a semicrystalline array. The branches are concentrated in the amorphous region.[113,114,121] Some starch granules contain no amylose, but it may constitute up to 30% by weight of the starch. It may be found in part in the amorphous bands and in part intertwined with the amylopectin.[122]

BOX 20-D GENETIC DISEASES OF GLYCOGEN METABOLISM

In 1951, B. McArdle described a patient who developed pain and stiffness in muscles after moderate exericse.[a] Surprisingly, this person completely lacked muscle glycogen phosphorylase. Since that time several hundred others have been found with the same defect. Glycogen accumulates in muscle tissue in this disease, one of the several types of **glycogen storage disease**.[b] Severe exercise is damaging, but steady moderate exercise can be tolerated. Until the time of McArdle's discovery, it was assumed that glycogen was synthesized by reversal of the phosphorylase reaction. No hint of the UDP-glucose pathway had appeared, and it was, therefore, not obvious how glycogen could accumulate in the muscles of these patients. Leloir's discovery of UDP-glucose at about the same time provided the answer. Persons with McArdle syndrome are greatly benefitted by a high-protein diet, presumably because amino acids such as alanine and glutamine are converted efficiently to glucose and because branched-chain amino acids may serve as a direct source of muscle energy.[c,d]

Several other rare heritable diseases also lead to accumulation of glycogen because of some block in its breakdown through the glycolysis pathway. The enzyme deficiencies include those of muscle phosphofructokinase,[e] liver phosphorylase kinase, liver phosphorylase, and liver glucose-6-phosphatase. In the last case, glycogen accumulates because the liver stores cannot be released to the blood as free glucose.[b,f,g] This is a dangerous disease because blood glucose concentrations may fall too low at night. The prognosis improved greatly when methods were devised for providing the body with a continuous supply of glucose. The simplest treatment is ingestion of uncooked cornstarch which is digested slowly.[b,h] In one of the storage diseases the branching enzyme of glycogen synthesis is lacking, and glycogen is formed with unusually long outer branches. In another the debranching enzyme is lacking, and only the outer branches of glycogen can be removed readily.[i]

The most serious of the storage diseases involve none of the enzymes mentioned above. Pompe disease is a fatal generalized glycogen storage disease in which a lysosomal α-1,4-glucosidase is lacking. This observation suggested the existence of a new and essential pathway of degradation of glycogen to free glucose in the lysosomes. A few cases of glycogen synthase deficiency have been reported. Little or no glycogen is stored in muscle or liver, and patients must eat at regular intervals to prevent hypoglycemia. Severe diseases in which glycogen synthesis is impaired include deficiencies of the gluconeogenic enzymes pyruvate carboxylase and PEP carboxykinase.

The following tabulation includes deficiencies of glycogen metabolism, glycolysis, and gluconeogenesis.[a] Glycogen storage diseases are often designated as Types I–V and these terms are included.

Deficiency	Organ	Severity
Glycogen phosphorylase (Type V), McArdle disease	Muscle	Moderate, late onset
Glycogen phosphorylase	Liver	Very mild
Phosphorylase kinase	Liver	Very mild
Debranching enzyme (Type III)	Liver	Mild
Lysosomal α-glucosidase (Type II)		Lethal, infant and adult form
Phosphofructokinase	Muscle	Moderate, late onset
Phosphoglycerate mutase[j]	Muscle	Moderate
Pyruvate carboxylase		Lethal
PEP carboxykinase		Lethal
Fructose-1,6-bisphosphatase	Muscle	Severe
Glycogen synthase	Liver	Mild
Branching enzyme (Type IV)		Lethal, liver transplantation
Glucose-6-phosphatase (Type I)		Severe if untreated

[a] Huijing, F. (1979) *Trends Biochem. Sci.* **4**, 192

[b] Chen, Y.-T., and Burchell, A. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 935–965, McGraw-Hill, New York

[c] Slonim, A. E., and Goans, P. J. (1985) *N. Engl. J. Med.* **312**, 355–359

[d] Goldberg, A. L., and Chang, T. W. (1987) *Fed. Proc.* **37**, 2301–2307

[e] Raben, N., Sherman, J., Miller, F., Mena, H., and Plotz, P. (1993) *J. Biol. Chem.* **268**, 4963–4967

[f] Nordlie, R. C., and Sukalski, K. A. (1986) *Trends Biochem. Sci.* **11**, 85–88

[g] Lei, K.-J., Shelly, L. L., Pan, C.-J., Sidbury, J. B., and Chou, J. Y. (1993) *Science* **262**, 580–583

[h] Chen, Y.-T., Cornblath, M., and Sidbury, J. B. (1984) *N. Engl. J. Med.* **310**, 171–175

[i] Thon, V. J., Khalil, M., and Cannon, J. F. (1993) *J. Biol. Chem.* **268**, 7509–7513

[j] Shanske, S., Sakoda, S., Hermodson, M. A., DiMauro, S., and Schon, E. A. (1987) *J. Biol. Chem.* **262**, 14612–14617

a glucose unit in an adjacent polysaccharide chain that lies parallel to the first, possibly in a double helix. Since amylose and amylopectin are intimately intermixed in the starch granules, it seems strange that the branching enzyme never transfers a branch to molecules of the straight-chain amylose. However, if the linear amylose chains are oriented in the opposite direction from the amylopectin chains, the nonreducing ends of the amylose molecules would be located toward the center of the starch granule. Growth could occur by an insertion mechanism at the reducing ends and the ends could move out continually with the amyloplast membrane as the granule grows.[123] Recent evidence from ^{14}C labeling indicates that both amylose and amylopectin too may grow by insertion at the reducing end of glucose units from ADP-glucose.[123a,b] Branching could occur to give the structure of Fig. 20-3. Starch synthesis in leaves occurs by day but at night the starch is degraded by amylases, α-glucosidases, and starch phosphorylase. Both the starch synthases and catabolic enzymes are present within the amyloplasts where they may be associated with regulatory proteins of the 14-3-3 class.[122a]

Digestion of dietary glycogen and starch in the human body begins with the salivary and pancreatic amylases, which cleave α-1,4 linkages at random. It continues with a **glucoamylase** found in the brush border membranes of the small intestine where it occurs as a complex with **maltase**.[74] Carbohydrases are discussed in Chapter 12, Section B.

2. Cellulose, Chitin, and Related Glycans

Cellulose synthases transfer glucosyl units from UDP-glucose, while chitin synthases utilize UDP-*N*-acetylglucosamine. Not only green plants but some fungi and a few bacteria form cellulose. The ameba *Dictyostelium discoideum* also coats its spores with cellulose.[124] Electron microscopic investigations suggest that both in bacteria[125] and in plants[126] multienzyme aggregates located at the plasma membrane synthesize many polymer chains side by side to generate hydrogen-bonded microfibrils which are extruded through the membrane. Both green plants and fungi also form important β-1,3-linked glycans.

The bacterial cellulose synthase from *Acetobacter xylinum* can be solubilized with detergents, and the resulting enzyme generates characteristic 1.7 nm cellulose fibrils (Fig. 20-4) from UDP-glucose.[125,127–129] These are similar, but not identical, to the fibrils of cellulose I produced by intact bacteria.[125,130] Each native fibril appears as a left-handed helix which may contain about nine parallel chains in a crystalline array. Three of these helices appear to coil together (Fig. 20-4) to form a larger 3.7-nm left-handed helical fibril. Similar fibrils are formed by plants. In both

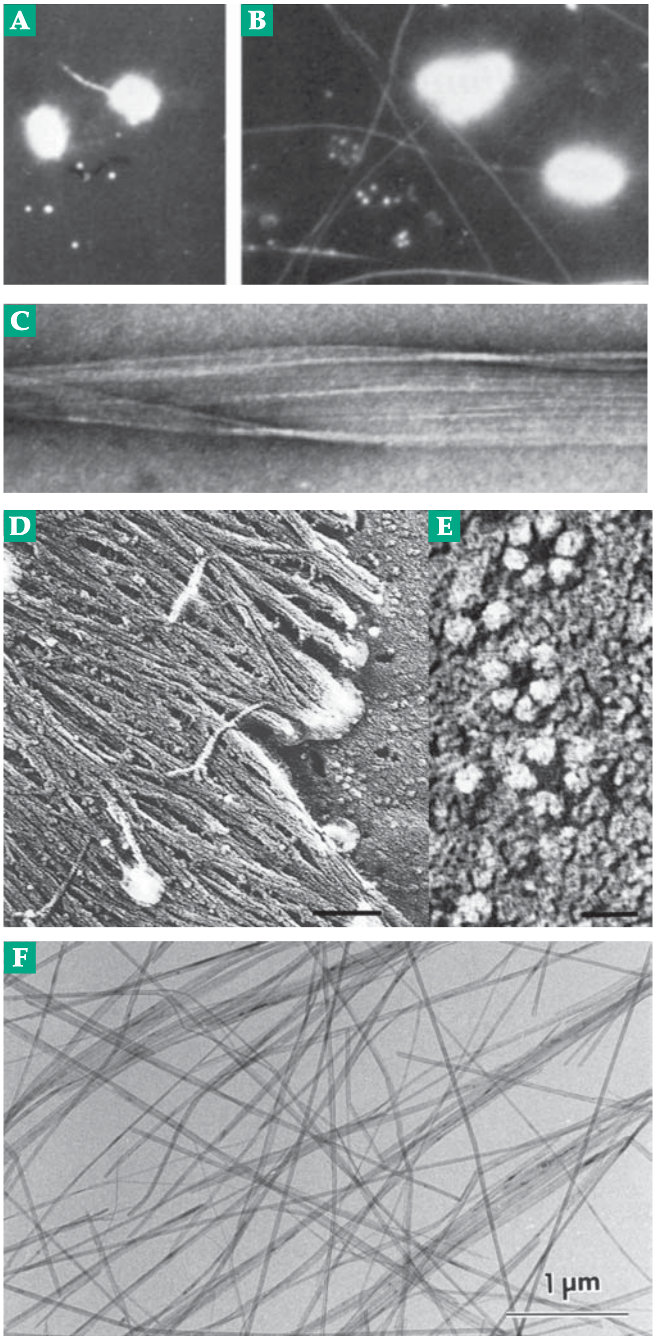


Figure 20-4 Cellulose microfibrils being formed by *Acetobacter xylinum*.[127] (A) Dark-field light micrograph after five minutes of cellulose production (x 1250). (B) After 15 minutes a pellicle of cellulose fibers is forming (x 2000) (C) Negatively stained cellulose ribbon. At the right the subdivision into microfibrils is visible. Courtesy of R. Malcolm Brown, Jr. (D) Cellulose microfibrils overlaying the plasma membrane in the secondary cell wall of a tracheary element of *Zinnia elegans*. Bar = 100 nm. (E) Rosettes in the plasma membrane underlying the cellulose-rich secondary cell wall thickening in *A. elegans*. Bar = 30 nm. (D) and (E) from Haigler and Blanton.[132] Courtesy of Candace H. Haigler. (F) Chitin microfibrils purified from protective tubes of the tube-worm *Lamellibrachia satsuma*.[137] Courtesy of Junji Sugiyama.

bacteria and plants the cellulose I fibrils that are formed are highly crystalline, contain parallel polysaccharide chains (Fig. 4-5), and have the tensile strength of steel. Electron micrographs show that the cell envelope of *A. xylinum* contains 5–80 pores, through which the cellulose is extruded, lying along the long axis of the cell.[129] The biosynthetic enzymes are probably bound to the plasma membrane. Similar, but more labile, cellulose synthases are present in green plants.[131] In *Arabidopsis* there are ten genes. The encoded cellulose synthases appear to be organized as rosettes on some cell surfaces (Fig. 20-4E).[131a-133a] The rosettes may be assembled to provide parallel synthesis of ~36 individual cellulose chains needed to form a fibril.[131a]

Because of the insolubility of cellulose fibrils it has been difficult to determine whether they grow from the reducing ends or the nonreducing ends of the chains. From silver staining of reducing ends and micro electron diffraction of cellulose fibrils attached to bacteria,[134] Koyama *et al.* concluded that the reducing ends are extruded from cells. New glucosyl rings would be added at the *nonreducing ends*, which remain attached noncovalently to the cells.[134] From amino acid sequence similarities it was also concluded that the same is true for *Arabidopsis*.[131,133] A single cellulose chain has a twofold screw axis, each residue being rotated 180° from the preceding residue (Fig. 4-5). It was postulated that two synthases act cooperatively to add cellobiose units. Another suggestion is that sitosterol β-glucoside acts in some fashion as a primer for cellulose synthesis in plants.[133b]

An insertion mechanism for synthesis of cellulose. Using ^{14}C "pulse and chase" labeling Han and Robyt found that new glucosyl units are added at the *reducing ends* of cellulose chains formed by cell membrane preparations from *A. xylinum*.[135] This conclusion is in accord with the generalization that extracellular polysaccharides made by bacteria usually grow from the reducing end by an insertion mechanism that depends upon a polyprenyl alcohol present in the cell membrane.[136] This lipid alcohol, often the C$_{55}$

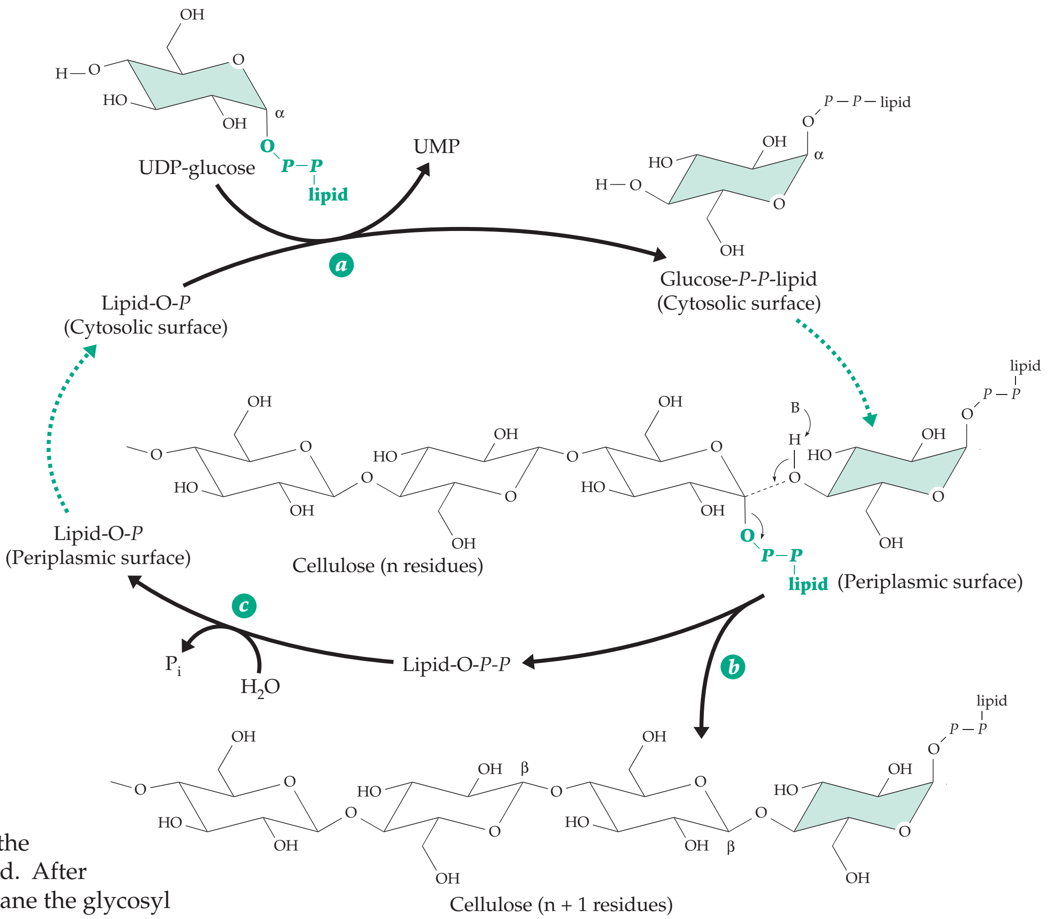
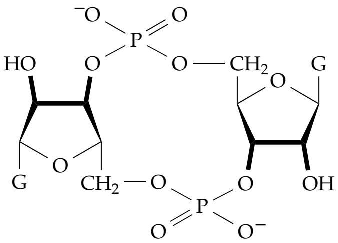


Figure 20-5 Proposed insertion mechanism for biosynthesis of cellulose. Three enzymatic steps are involved: a nucleophilic displacement reaction of a lipid phosphate on UDP-glucose yields a glucosyl diphosphate lipid in which the α-glycosyl linkage is retained. After passage through the membrane the glycosyl group is inserted into the reducing end of a cellulose chain, which is covalently attached by a pyrophosphate linkage to another lipid. The first lipid diphosphate is released and is hydrolyzed (step *c*) to the monophosphate, which crosses the membrane to complete the cycle. After Han and Robyt.[135] As throughout this book *P* represents the phospho group –PO$_3$H. The H may be replaced by groups which may contain oxygen atoms. This explains why an O is included in Lipid-O-*P* but no O is shown between the *P*'s in -O-*P*-*P*.

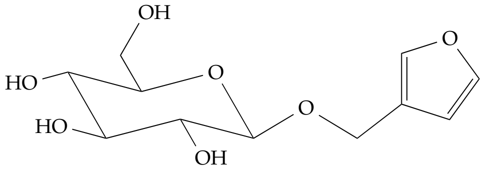
bactoprenol, reacts with UDP-glucose (or other glycosyl donor) to give a lipopyrophospho-glucose (step *a*, Fig. 20-5). The α linkage of the UDP-glucose is retained in this compound. The growing cellulose chain is attached at the reducing end by a similar linkage to a second lipid molecule. Then, in a displacement on the anomeric carbon of the first glucosyl residue of the cellulose chain, the new glucosyl unit is inserted with inversion of the α linkage to β. In step *c* the pyrophosphate linkage of the lipid diphosphate is hydrolyzed to regenerate the lipid monophosphate and to drive the reaction toward completion. Two of the steps in the cycle involve transport across the bacterial membrane. The first involves the lipid –O-*P*-*P*-glucose and the second the lipid monophosphate. This type of insertion mechanism is a common feature of polyprenol phosphate-dependent synthetic cycles for extracellular polysaccharides (Figs. 20-6, 20-9 and Eq. 20-20). However, further verification is needed for cellulose synthesis.

Regulation of cellulose synthesis in bacteria depends on allosteric activator of cellulose synthase, **cyclic diguanylate** (c-di-GMP), and a Ca^{2+}-activated phosphodiesterase that degrades the activator.[129,138–139a] Sucrose is the major transport form of glucose in plants. Synthesis of both cellulose and starch is reduced in mutant forms of maize deficient in sucrose synthase (Eq. 20-13). This synthase, acting in the reverse direction, forms UDP-Glc from sucrose.[140,141] The enzymatic degradation of cellulose is an important biological reaction, which is limited to certain bacteria, to fungi, and to organisms such as termites that obtain cellulases from symbiotic bacteria or by ingesting fungi.[142] These enzymes are discussed in Chapter 12, Section B,6. Genetic engineering methods now offer the prospect of designing efficient cellulose-digesting yeasts[143] that may be used to produce useful fermentation products from cellulose wastes.



Cyclic diguanylate

Callose and other β-1,3-linked glycans. Attempts to produce cellulose from UDP-Glc using enzymes of isolated plasma membranes from higher plants have usually yielded the β-1,3-linked glucan (callose) instead. This is a characteristic polysaccharide of plant wounds which, as healing occurs, is degraded and replaced by cellulose.[140,144] Callose formation is induced by a specific activator β-furfuryl-β-glucoside, and callose synthase is virtually inactive unless both the activator and Ca^{2+} are present.[144]



β-Furfuryl-β-glucoside

Beta-1,3-linked glycans are major components of the complex layered cell wall of yeasts and other fungi. In the fission yeast *Saccharomyces pombe* ~55% of the cell wall carbohydrate consists of β-1,3-linked glucan with some β-1,4-linked branches, ~ 28% is α-1,3-linked glucan, ~6% is α-1,6-linked glucan, and ~0.5% is chitin. There are two carbohydrate layers, the outer one appearing amorphous. The inner layer contains interwoven fibrils of both α-1,3-linked and α-1,4-linked glucans and holds the shape of the cell. The β-1,3 glucan synthase is localized on the inner side of the cell membrane and is activated by GTP and a small subunit of the Rho family of G proteins.[145]

Plants synthesize 1,3-β-glucanases that hydrolyze the glycans of fungal cell walls. Synthesis is induced by wounding as a defense reaction (see Box 20-E). These glycanases also function in the removal of callose.[146]

Chitin. Like cellulose synthase, fungal chitin synthases are present in the plasma membrane and extrude microfibrils of chitin to the outside.[147–150] In the fungus *Mucor* the majority of the chitin synthesized later has its *N*-acetyl groups removed hydrolytically to form the deacetylated polymer **chitosan**.[151,152] Chitin is also a major component of insect exoskeletons. For this reason, chitin synthase is an appropriate target enzyme for design of synthetic insecticides.[153]

Chitin hydrolyzing enzymes are formed by fungi and in marine bacteria.[154] Chitinases are also present in plant vacuoles, where they participate in defense against fungi and other pathogens[155] (Box 20-E). More recently a chitinase has been identified in human activated macrophages.[156] Another unanticipated discovery was that a developmental gene designated *DG*42, from *Xenopus*, has a sequence similar to that of the *NodC* gene. The latter encodes a synthase for chitin oligosaccharides (Nod factors) that serve as nodulation factors in *Rhizobia* (Chapter 24). The enzyme is synthesized for only a short time during early embryonic development.[157] The significance of this discovery is not yet clear. Synthesis of both the bacterial Nod factors and chitin oligosaccharides in zebrafish embryos occurs by transfer of GlcNAc residues from UDP-GlcNAc at the *nonreducing ends* of the chains.[158] Whether the same is true of chitin in fungi or arthropods remains uncertain.

Cell walls of plants. The thick walls of higher plant cells (Figs. 1-7, 4-14, and 20-4D) provide strength and rigidity to plants and, at the same time, allow rapid elongation during periods of growth.[159–163a] Northcote[164] likened the wall structure to glass fiber-reinforced plastic (fiber glass). Thus, the cell wall contains microfibrils of cellulose and other polysaccharides embedded in a matrix, also largely polysaccharide. The **primary cell wall** laid down in green plants during early stages of growth contains loosely interwoven cellulose fibrils ~10 nm in diameter and with an ~4 nm crystalline center. The cellulose in these fibrils has a degree of polymerization of 8000–12,000 glucose units. As the plant cell matures, a secondary cell wall is laid down on the inside of the primary wall. This contains many layers of closely packed microfibrils, alternate layers often being laid down at different angles to one another (Fig. 20-4D). The microfibrils in green plants are most often cellulose but may contain other polysaccharides as well. Some algae are rich in fibrils of xylan and mannan.

The materials present in the matrix phase vary with the growth period of the plant. During initial phases **pectin** (polygalacturonic acid derivatives) predominate but later xylans and a variety of other polysaccharides known as **crosslinking glycans** (or hemicelluloses) appear. Primary cell wall constituents of dicotyledons include **xyloglucans** (linear glucan chains with xylose, galactose, and fucose units in

BOX 20-E OLIGOSACCHARIDES IN DEFENSIVE AND OTHER RESPONSES OF PLANTS

Plants that are attacked by bacteria, fungi, or arthropods respond by synthesizing broad-spectrum antibiotics called **phytoalexins**,[a,b] by strengthening their cell walls with lignin and hydroxyproline-rich proteins called **extensins**,[c] and by making **protease inhibitors** and other proteins that help to block the chemical attack.[d] These plant responses seem to be initiated by the release from an invading organism of **elicitors**, which are often small oligosaccharide fragments, sometimes called **oligosaccharins**.[e] These include β-1,6-linked glucans that carry β-1,3-linked branches as well as chitin and chitosan oligomers, derived from fungal cell walls.[f] Other elicitors include galacturonic acid oligomers released from damaged plant cell walls,[g] metabolites such as arachidonic acid and glutathione,[h] and bacterial toxins.[i] Any of these may serve as signals to plants to take defensive measures.

Phytoalexins are often isoflavonoid derivatives (Box 21-E). Their synthesis, like that of lignin, occur via 4-coumarate (4-hydroxycinnamate, Fig. 25-8). The ligase which forms the thioester of 4-coumarate with coenzyme A is one of the **pathogenesis-related proteins** whose synthesis is induced.[j] A second induced enzyme is chalcone synthase, which condenses three acetyl units onto 4-coumaroyl-CoA as shown in Box 21-E. Its induction by elicitors acting on bean cells requires only five minutes.[h] Another rapidly induced gene is that of cinnamoyl alcohol dehydrogenase,[k] essential to lignin synthesis. Other proteins formed in response to infections include **chitinases** that are able to attack invading fungi[l,m] as well as the protease inhibitors. Their synthesis is induced via derivatives of **jasmonate**, a product of the octadecenoic acid pathway (Eq. 21-18).[a] As yet, little is known about the mechanism by which elicitors induce the defensive responses, but the presence of receptors, of phosphorylation, and of release of second messengers have been suggested.[d]

Lipooligosaccharides known as Nod factors (p. 1365) are another group of signaling molecules. These chitin-related *N*-acylated oligomers of *N*-acetylglucosamine (GlcNAc) do not defend against infection but invite infection of roots of legumes by appropriate species of *Rhizobia*[n–p] leading to formation of nitrogen-fixing root nodules.

---

[a] Blechert, S., Brodschelm, W., Hölder, S., Kammerer, L., Kutchan, T. M., Mueller, M. J., Xia, Z.-Q., and Zenk, M. H. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4099–4105

[b] Ebel, J., and Grisebach, H. (1988) *Trends Biochem. Sci.* **13**, 23–27

[c] Kieliszewski, M. J., O'Neill, M., Leykam, J., and Orlando, R. (1995) *J. Biol. Chem.* **270**, 2541–2549

[d] Ryan, C. A. (1988) *Biochemistry* **27**, 8879–8883

[e] Ryan, C. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1–2

[f] Baureithel, K., Felix, G., and Boller, T. (1994) *J. Biol. Chem.* **269**, 17931–17938

[g] Reymond, P., Grünberger, S., Paul, K., Müller, M., and Farmer, E. E. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4145–4149

[h] Dron, M., Clouse, S. D., Dixon, R. A., Lawton, M. A., and Lamb, C. J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6738–6742

[i] Bidwai, A. P., and Takemoto, J. Y. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6755–6759

[j] Douglas, C., Hoffmann, H., Schulz, W., and Hahlbrock, K. (1987) *EMBO J.* **6**, 1189–1195

[k] Walter, M. H., Grima-Pettenati, J., Grand, C., Boudet, A. M., and Lamb, C. J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5546–5550

[l] Legrand, M., Kauffmann, S., Geoffroy, P., and Fritig, B. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6750–6754

[m] Payne, G., Ahl, P., Moyer, M., Harper, A., Beck, J., Meins, F., Jr., and Ryals, J. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 98–102

[n] Cedergren, R. A., Lee, J., Ross, K. L., and Hollingsworth, R. I. (1995) *Biochemistry* **34**, 4467–4477

[o] Jabbouri, S., Relic, B., Hanin, M., Kamalaprija, P., Burger, U., Promé, D., Promé, J. C., and Broughton, W. J. (1998) *J. Biol. Chem.* **273**, 12047–12055

[p] Dénarié, J., Debellé, F., and Promé, J.-C. (1996) *Ann. Rev. Biochem.* **65,** 503–535

branches),[164a] other crosslinking glycans, and galacturonic acid-rich pectic materials.[163a] The xyloglycans, which comprise 20% of the cell wall in some plants, have a backbone of α-1,4-linked glucose units with numerous α-1,6-linked xylose rings, some of which carry attached L-arabinose, galactose, or fucose. The structures, which vary from species to species, are organized as repeating blocks with a continuous glucan backbone. Another crosslinking glycan is **glucuronoarabinoxylan**. The backbone is β-1,4-linked xylose. Less abundant glucomannans, galactomannans, and galactoglucomannans, with β-1,4-linked mannan backbone structures, are also present in most angiosperms.[163a]

Pectins form a porous gel on the inside surface of plant cell walls.[163a,164a] A major component is a **homogalacturonan**, which consists of α-1,4-linked galacturonic acid (GalA). A second is rhamnogalacturonan I, an alternating polymer of (2-L-Rhaα1 → 4GalAα →) units. The most interesting pectin component is **rhamnogalacturonan II**, one of the less abundant constuents of pectin. It is obtained by hydrolytic cleavage of pectin by a polygalacturonidase. Before such release it forms parts (hairy regions) of pectin molecules that are largely homogalacturonans (in smooth regions). A rhamnogalacturonan II segment consists of 11 different monomer units.[164b–f] Attached to the polygalacturonic acid backbone are four oligosaccharides, consisting of rhamnose, galactose and fucose as well as some unusual sugars (see structure in Box 20-E). This polysaccharide is apparently present in all higher plants and is unusually stable, accumulating, for example, in red wine.[164e] It contains two residues of the branched chain sugar **apiose**, one of which is a site of crosslinking by boron (Box 20-F). A borate diol ester linkage binds two molecules of the pectin together as a dimer, perhaps controlling the porosity of the pectin gel. All of the complex cell wall polysaccharides bind, probably through multiple hydrogen bonds, to the cellulose microfibrils (Fig. 4-14). The resulting structures are illustrated in drawings of Carpita and McCann,[163a] which are more current than is Fig. 4-14. The cellulose plus crosslinking glycans form one network in the cell wall. The pectic substances form a second independent network. Some covalent crosslinking occurs, but most interactions are noncovalent.[163a] The site of biosynthesis of pectins and hemicelluloses is probably Golgi vesicles which pass to the outside via exocytosis. However, the cellulose fibrils as well as the chitin in fungi are apparently extruded from the plasma membrane.

Although the principal cell wall components of plants are carbohydrates, proteins account for 5–10% of the mass.[165] Predominant among these are glycoprotein **extensins**. Like collagen, they are rich in 4-hydroxyproline which is glycosylated with arabinose oligosaccharides and galactose (p. 181). Other hydroxyproline-containing proteins with the characteristic sequence (hydroxyproline)$_4$-Ser are also found, e.g., in soybean cell walls.[166] Some plant cell walls contain glycine-rich structural proteins. One in the petunia consists of 67% glycine residues.[167] During advanced stages of formation, as the walls harden into wood, large amounts of **lignins** are laid down in some plant cells. These chemically resistant phenylpropanoid polymers contain many crosslinked aromatic rings (Fig. 25-9).[163a]

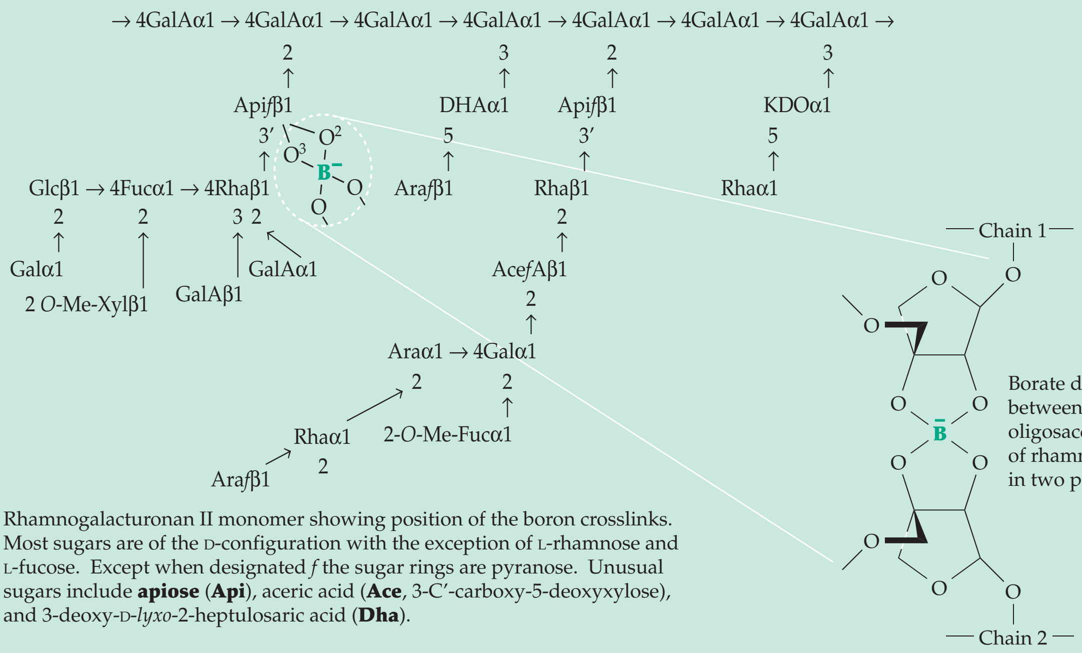
A remarkable aspect of primary plant cell walls is their ability to be elongated extremely rapidly during growth. While the driving force for cell expansion is thought to be the development of pressure within the cell, the manner in which the wall expands is closely regulated. After a certain point in development, elongation occurs in one direction only and under the influence of plant hormones. Most striking is the effect of the **gibberellins** (Eq. 22-5), which cause very rapid elongation. Elongation of plant cell walls may depend to some extent upon chemical cleavage and reforming of crosslinking polysaccharides. However, the cellulose fibrils probably remain intact and slide past each other.[161] A curious effect, which is mediated by proteins called **expansins**,[133a,168] is the ability of plant tissues to extend rapidly when incubated in a mildly acidic buffer of pH <5.5. Expansins are also involved in ripening of fruit. They may disrupt noncovalent bonding between cellulose fibrils and the hemicelluloses.[169,170] The β-expansins of grasses are allergens found in grass pollens.[133a,168] The borate diol ester linkages in the pectin may also facilitate expansion.

3. Patterns in Polysaccharide Structures

How can the many complex polysaccharides found in nature be synthesized? Are there genetically determined patterns? How are these controlled? The answer can be found in the *specificities* of the hundreds of known *glycosyltransferases*[171,172] and in the *patterns of expression of the genes* for transferases and other proteins. As a consequence, a great variety of structurally varied polysaccharide structures arise, especially on cell surfaces. The structures are not random but depend upon the assortment of glycosyltransferases available at the particular stage of development in a tissue. The numerous possibilities can account for much of the variation observed between species, between tissues, and also among individuals.

The simplest pattern is the growth of straight-chain homopolysaccharides such as amylose, cellulose, and chitin. The glycosyltransferases must recognize both the glycosyl donor, e.g., ADP-glucose, UDP-glucose, and also the correct end of the growing polymer, always adding the same monomer unit.

BOX 20-F WHAT DOES BORON DO?



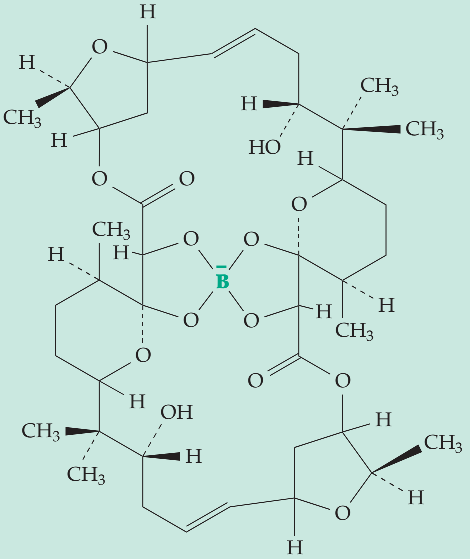
Rhamnogalacturonan II monomer showing position of the boron crosslinks. Most sugars are of the D-configuration with the exception of L-rhamnose and L-fucose. Except when designated *f* the sugar rings are pyranose. Unusual sugars include **apiose** (**Api**), aceric acid (**Ace**, 3-C'-carboxy-5-deoxyxylose), and 3-deoxy-D-*lyxo*-2-heptulosaric acid (**Dha**).

Borate diol ester linkage between apiose rings in oligosaccharide clusters of rhamnogalacturonan II in two pectin chains.

For 75 years or more it has been known that boron is essential for growth of green plants.[a,b] In its absence root tips fail to elongate normally, and synthesis of DNA and RNA is inhibited. Boron in the form of boric acid, $B(OH)_3$, is absorbed from soil. Although deficiency is rare it causes disintegration of tissues in such diseases as "heart rot" of beets and "drought spot" of apples. The biochemical role has been obscure, but is usually thought to involve formation of borate esters with sugar rings or other molecules with adjacent pairs of –OH groups (as in the accompanying structures). A regulatory role involving the plant hormones auxin, gibberelic acid, and cytokinin has also been suggested.

Diatoms also require boron, which is incorporated into the silicon-rich cell walls.[a] Some strains of *Streptomyces griseus* produce boron-containing macrolide antibiotics such as **aplasmomycin** (right).[c] Recently a function in plant cell walls has been identified (see also main text) as crosslinking of rhamnogalacturonan portions of pectin chains by borate diol ester linkages as illustrated.

It was long thought that boron was not required by human beings, but more recent studies suggest that we may need ~30 µg / day.[d] The possible functions are uncertain. Animals deprived of boron show effects on bone, kidney, and brain as well as a relationship to the metabolism of calcium, copper, and nitrogen. Nielson proposed a signaling function, perhaps via phosphoinositides, in animals.[b]



Aplasmomycin, a boron-containing antibiotic

---

[a] Salisbury, F. B., and Ross, C. W. (1992) *Plant Physiology*, 4th ed., Wadsworth, Belmont, California

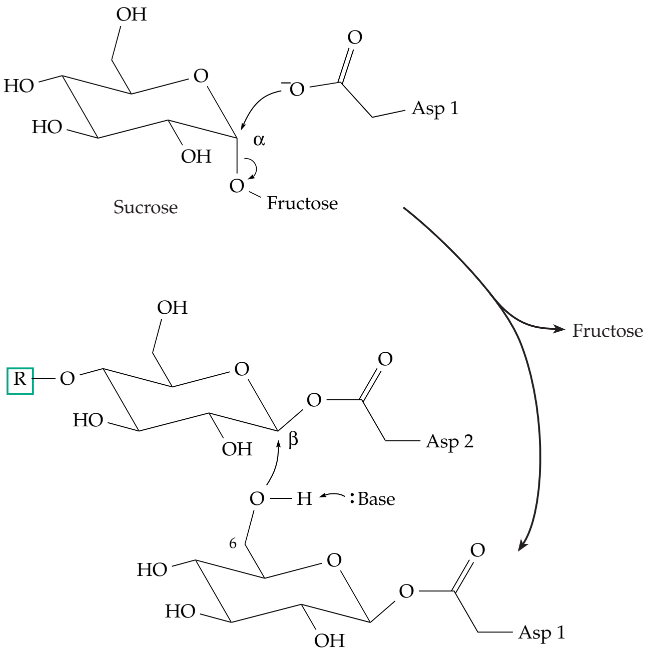
[b] Nielsen, F. H. (1991) *FASEB J.* **5**, 2661–2667

[c] Lee, J. J., Dewick, P. M., Gorst-Allman, C. P., Spreafico, F., Kowal, C., Chang, C.-J., McInnes, A. G., Walter, J. A., Keller, P. J., and Floss, H. G. (1987) *J. Am. Chem. Soc.* **109**, 5426–5432

[d] Nielsen, F. H. (1999) in *Modern Nutrition in Health and Disease*, 9th ed. (Shils, M. E., Olson, J. A., Shike, M., and Ross, A. C., eds), pp. 283–303, Williams & Wilkins, Baltimore, Maryland

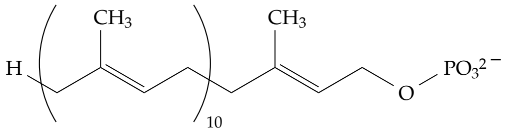
In contrast, hyaluronan and the polysaccharide chains of **glycosaminoglycans** (Fig. 4-11) have an alternating pattern. For a hyaluronan chain growing at the reducing end, one active site of hyaluronan synthase must be specific for UDP-GlcNAc and transfer the sugar unit only to the end of a glucuronic acid ring. A second active site must be specific for UDP-glucuronic acid but attach it only to the end of an acetylglucosamine unit.[172a,172b] There is still uncertainty about the direction of growth of hyaluronan.[173–175] Some hyaluronan synthases are lipid-dependent and their mechanism may resemble that proposed for cellulose synthesis (Fig. 20-5).

Dextrans. Some polysaccharides, such as the bacterial dextrans, are synthesized outside of cells by the action of secreted enzymes. An enzyme of this type, **dextran sucrase** of *Leuconostoc* and *Streptococcus*, adds glucosyl units at the *reducing* ends of the dextran chains (p. 174). Sucrose is the direct donor of the glucosyl groups, which are added by an insertion mechanism.[121,176–178] However, it is not dependent upon a membrane lipid as is that of Fig. 20-5. The glucosyl groups are transferred from sucrose to one of a pair of carboxylate groups of aspartate side chains in the active site.[179,180] If both carboxylates are glucosylated, a dextran chain can be initiated by insertion of one glucosyl group into the second (Eq. 20-19). The dextran grows alternating binding sites between the two carboxylates. Chain growth can be terminated by reaction with a sugar or oligosaccharide that fits into the active site and acts in place of the glucosyl group attached to Asp 1 as pictured in Eq. 20-19. The α-1,3-linked branches can be formed when a 3-OH group of a second dextran chain enters the catalytic site, serving as the glycosyl acceptor. See Robyt for a detailed discussion of synthesis of dextrans and related polysaccharides such as **alternan** and the α-1,3-linked **mutan** (p. 175).[121,176] Some bacteria form β-2,6-linked **fructans** by a similar mechanism, with glucose being released by displacement on C2 of sucrose.[181] Fructans are also formed in green plants, apparently from reaction of two molecules of sucrose with release of glucose to form the trisaccharide Fru*f*β2 → 1Fru*f*β2–1α-Glc*p*, which then transfers a fructosyl group to the growing chain.



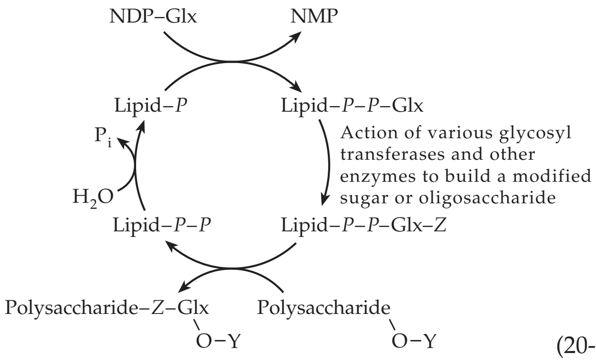
R = H or growing α–1,6-linked chain (20-19)

Lipid-dependent synthesis of polysaccharides. Insertion of monomer units at the base of a chain is a major mechanism of polymerization that is utilized for synthesis not only of polysaccharides but also of proteins (Chapter 29). For most carbohydrates the synthesis is dependent upon a polyprenyl lipid alcohol. In bacteria this is often the 55-carbon **undecaprenol** or **bactoprenol**,[136] which functions as a phosphate ester:



Undecaprenol phosphate

It serves as a membrane anchor for the growing polysaccharide. We have already discussed one example in the hypothetical cellulose synthase mechanism of Fig. 20-5. For some polysaccharides the mechanism is better established. The synthetic cycles all resemble that of Fig. 20-5 and can be generalized as in Eq. 20-20. Here NDP-Glx is a suitable nucleotide disphosphate derivative of sugar Glx, and Z-Glx is the repeating unit of the polysaccharide formed by the action of glycosyltransferases and other enzymes.

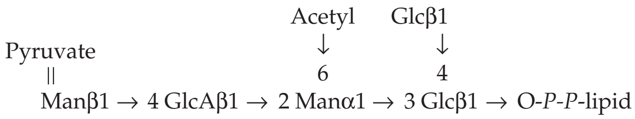


(20-20)

For example, the biosynthesis of alginate involves GDP-mannuronic acid (GDP-ManA) as NDP-Glx, bactoprenol as the lipid, and a glycosyltransferase that inserts a second mannuronate residue (as Z).

An additional transferase that uses acetyl-CoA as a substrate sometimes acetylates one mannuronate unit. The disaccharide units are then inserted into the growing chain. An additional modification, which occurs after polymerization, is random C5 epimerization of unacetylated D-mannuronate residues to L-guluronate.[136,182] Formation of alginate is of medical interest because infections by alginate-forming bacteria are a major cause of respiratory problems in cystic fibrosis.[182]

Sometimes an oligosaccharide assembled on the polyprenol phosphate represents a substantial block in assembly of a repeating polymer. For example, the xanthan gum (p. 179) produced by the bacterium *Xanthomonas campestris* is formed by several successive glycosyl transfers to bactoprenol-*P*-*P*-Glc. A second glucose is transferred onto the first from UDP-Glc, forming a pair of glucosyl groups in β-1,4 linkage. Mannose is then transferred from GDP-Man and joined in an α-1,3 linkage to the first GDP-Man to form a branch point. A glucuronate residue is then transferred from UDP-GlcA and another mannose from GDP-Man. The last mannose is modified by reaction with PEP to form a ketal (Eq. 4-9). The product of this assembly is the following lipid-bound oligosaccharide block.



$$\overset{\text{Pyruvate}}{\overset{\|}{\text{Man}}}\beta 1 \rightarrow 4\ \text{GlcA}\beta 1 \rightarrow 2\ \overset{\text{Acetyl}}{\overset{\downarrow}{\overset{6}{\text{Man}}}}\alpha 1 \rightarrow 3\ \overset{\text{Glc}\beta 1}{\overset{\downarrow}{\overset{4}{\text{Glc}}}}\beta 1 \rightarrow \text{O-}P\text{-}P\text{-lipid}$$

This is inserted into the growing polysaccharide using the free 4-OH on the second glucose to link the units in a cellulose type chain. The twelve separate genes needed for synthesis of xanthan gum are contained in a 16-kb segment of the *X. campestris* genome.[136] Lipid-bound intermediates are also involved in synthesis of peptidoglycans (Fig. 20-9) and in the assembly of bacterial O-antigens (Fig. 8-30). Both of these also yield "block polymers."

D. Proteoglycans and Glycoproteins

The glycoproteins contain oligosaccharides attached to the protein either through *O*-glycosidic linkages with hydroxyl groups of side chains of serine, threonine, hydroxyproline, or hydroxylysine (*O*-linked) or via glycosylaminyl linkages to asparagine side chains (*N*-linked). The "core proteins" of the proteoglycans carry long polysaccharide chains, which are usually *O*-linked and are usually described as glycosaminoglycans.

1. Glycosaminoglycans

Synthesis of the alternating polysaccharide hyaluronan has been discussed in Section C,3 and may occur by an insertion mechanism. However, other glycosaminoglycans (sulfate esters of **chondroitin**, **dermatan**, **keratan**, **heparan**, and **heparin**) grow at their nonreducing ends.[183,184] Their synthesis is usually initiated by the hydroxyl group of serine or threonine side chains at special locations within several secreted proteins.[185] These proteins are synthesized in the rough ER and then move to the Golgi. Addition of the first sugar ring begins in the ER with transfer of single xylosyl residues to the initiating –OH groups.[186–190b] This reaction is catalyzed by the first of a group of special glycosyltransferases of high specificity that form the special terminal units (Chapter 4, Section D,1), that anchor the alternating polysaccharide represented here as $(X–Y)_n$:

$$(\text{X} \rightarrow \text{Y})_n \rightarrow 4\ \text{GlcA}\beta 1 \rightarrow 3\ \text{Gal}\beta 1 \rightarrow$$
$$3\ \text{Gal}\beta 1 \rightarrow 4\ \text{Xyl}\beta 1 \rightarrow \text{O-Ser / Thr}$$

After transfer of the xylosyl residue from UDP-xylose to the –OH group in the protein,[190a] a second enzyme with proper specificity transfers a galactosyl group from UDP-galactose, joining it in β-1,4 linkage. A third enzyme transfers another galactosyl group onto the first one in β-1,3 linkage. A fourth enzyme, with a specificity different from that used in creating the main chain, then transfers a glucuronosyl group from UDP-glucuronic acid onto the chain terminus to complete the terminal unit.[190c] Then two more enzymes transfer the alternating units in sequence to form the repeating polymer with lengths of up to 100 or more monosaccharide residues. The sequence $(X–Y)_n$ in the preceding formula is:

$$(\rightarrow 4\ \text{GlcA}\beta 1 \rightarrow 3\ \text{GalNAc}\beta 1 \rightarrow)_n$$
for chondroitin[172a,191,192]

$$(\rightarrow 4\ \text{GlcA}\beta 1 \rightarrow 4\ \text{GlcNAc}\alpha 1 \rightarrow)_n$$
for heparan sulfate and heparin

$$(\rightarrow 3\ \text{Gal}\beta 1 \rightarrow 4\ \text{GlcNAc}\beta 1 \rightarrow)_n$$
for dermatan sulfate

Subsequent modifications of the polymers involve extensive formation of *O*-sulfate esters,[190a,193–197] *N*-deacetylation and *N*-sulfation,[198,199] and epimerization at C5.[10] In some tissues almost all GluA is epimerized.[200] The modifications are especially extensive in dermatan, heparan sulfates, and heparin (see also p. 177).[196,201–203b] The modifications are not random and follow a defined order. *N*-Deacetylation must precede *N*-sulfation, and *O*-sulfation is initiated only after *N*-sulfation of the entire chain is complete. The modifications occur within the Golgi (see Fig. 20-7) but not all of the glycosyltransferases, PAPS (3'-phosphoadenosine 5'-phosphosulfate)-dependent sulfotransferases, and epimerases are present within a single compartment. Nevertheless, an entire glycosaminoglycan chain can be synthesized within 1–3 min.[189]

The completed polymers are modified uniformly. There are clusters of sulfo groups with unusual structures in chondroitin from squid and shark cartilages[204,205] and fucosylated chondroitin from echinoderms.[206] Similar modifications are present less extensively in vertebrates. One of the best known modifications forms the unique pentasaccharide sequence shown in Fig. 4-13, which is essential to the anticoagulant activity of heparin. This sequence has been synthesized in the laboratory as have related longer heparin chains. A sequence about 17 residues in length containing an improved synthetic version of the unique pentasaccharide binds tightly to both thrombin and antithrombin (Chapter 12, Section C,9).[207,208] This opens the door to the development of improved substitutes for the medically important heparin. Heparan sulfate chains are found on proteoglycans throughout the body, but the highly modified heparin does not circulate in the blood. It is largely sequestered in cytoplasmic granules within mast cells and is released as needed.[208a] Heparin binds to many different proteins (p. 177). Among them is the glycoprotein selenoprotein P (p. 824), which may impart antioxidant properties to the extracellular matrix.[208b]

Although glycosaminoglycans are most often attached to *O*-linked terminal units, chondroitin sulfate chains can also be synthesized with *N*-linked oligosaccharides attached to various glycoproteins serving as initiators.[209] At least one form of keratan sulfate, found in the cornea, is linked to its initiator protein via GlcNAc-Man to *N*-linked oligosaccharides of the type present in many glycoproteins (Section D).

At least 25 different proteins that are secreted into the extracellular spaces of the mammalian body carry glycosaminoglycan chains.[183,210,211] Most of these proteins can be described as (1) **small leucine-rich proteoglycans** with 36- to 42-kDa protein cores and (2) **large modular proteoglycans** whose protein cores have molecular masses of 40 to 500 kDa.[210] The most studied of the second group is **aggrecan**, a major component of cartilage. This 220-kDa protein carries ~100 chondroitin chains, each averaging about 100 monosaccharide residues and ~100 negative charges from the carboxylate and sulfate groups. Aggrecan has three highly conserved globular domains near the N and C termini.[212–213a] The G1 domain near the N terminus is a **lectin** (p. 186), which, together with a small link protein that is structurally similar to the G1 domain, binds to a decasaccharide unit of hyaluronan. One hyaluronan molecule of 500- to1000-kDa mass (~2500–5000 residues) may bind 100 aggrecan and link molecules to form an ~200,000-kDa particle such as that shown in Fig. 4-16. These enormous highly negatively charged molecules, together with associated counterions, draw in water and preserve osmotic balance. It is these molecules that keep our joints mobile and which deteriorate by proteolytic degradation in the common **osteoarthritis**.[214,215] The keratan sulfate content of cartilage varies with age, and the level in serum and in synovial fluid is increased in osteoarthritis.[215] Keratan sulfate is also found in the cornea and the brain. Its content is dramatically decreased in the cerebral cortex of patients with Alzheimer disease.[216]

Other modular core proteins[210] include **versican** of blood vessels and skin,[210,213a,217,217a] **neurocam** and **brevican** of brain, **perlecan** of basement membranes,[218] **agrin** of neuromuscular junctions, and **testican** of seminal fluid. However, several of these have a broader distribution than is indicated in the foregoing description. The sizes vary from 44 kDa for testican to greater than 400 kDa for perlecan. The numbers of glycosaminoglycan chains are smaller than for aggrecan, varying from 1 to 30. Another of the chondroitin sulfate-bearing core protein is **appican**, a protein found in brain and one of the splicing variants of the amyloid precursor protein that gives rise to amyloid deposits in Alzheimer disease (Chapter 30).[217a,b]

The core proteins of the leucine-rich proteoglycans have characteristic horseshoe shapes and are constructed from ~28-residue repeats, each containing a β turn and an α helix. The three-dimensional structures are doubtless similar to that of a ribonuclease inhibitor of known structure which contains 15 tandem repeats.[219,220] A major function of these proteoglycans seems to be to interact with collagen fibrils, which have distinct proteoglyan-binding sites,[221] and also with fibronectin.[222] The small leucine-rich proteoglycans have names such as **biglycan**, **decorin**,[222a] **fibromodulin**, **lumican**, **keratoglycan**, **chondroadherin**, **osteoglycin**, and **osteoadherin**.[219,223–223c] The distribution varies with the tissue and the stage of development. For example, biglycan may function in early bone formation; decorin, which has a high affinity for type I collagen, disappears from bone tissue as mineralization takes place. Osteoadherin is found in mature osteoblasts.[223] **Phosphocan**, another brain proteoglycan, has an unusually high content (about one residue per mole) of L-isoaspartyl residues (see Box 12-A).[224]

Proteoglycans bind to a variety of different proteins and polysaccharides. For example, the large extracellular matrix protein **tenascin**, which is important to adhesion, cell migration, and proliferation, binds to chondroitin sulfate proteoglycans such as neurocan.[225] **Syndecan**, a transmembrane proteoglycan, carries both chondroitin and heparan sulfate chains, enabling it to interact with a variety of proteins that mediate cell–matrix adhesion.[185]

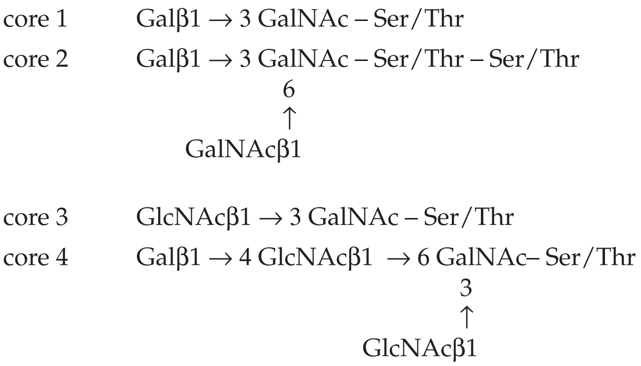
The ability of dissociated cells of sponges to aggregate with cells only of a like type (p. 29) depends upon large extracellular proteoglycans. That of *Microciona prolifera* appears to be an aggregate of about three hundred 35-kDa core protein molecules with equal masses of attached carbohydrate. This **aggregation factor** has a total mass of ~2×10^4 kDa.[226,227] It apparently interacts specifically, in the presence of Ca^{2+} ions, with a 210-kDa cell matrix protein to hold cells of the same species together.[227]

2. *O*-Linked Oligosaccharides

A variety of different oligosaccharides are attached to hydroxyl groups on appropriate residues of serine, threonine, hydroxylysine, and hydroxyproline in many different proteins (Chapter 4). Such oligosaccharides are present on external cell surfaces, on secreted proteins, and on some proteins of the cytosol and the nucleus.[228–231b] The rules that determine which –OH groups are to become glycosylated are not yet clear.[232] Glycosylation occurs in the ER, and, just as during synthesis of the long carbohydrate chains of proteoglycans, the sugar rings are added directly to an –OH group, either of the protein or of the growing oligosaccharide. The first glycosyl group transferred is most often **GalNAc** for external and secreted proteins[233] but more often **GlcNAc** for cytosolic and nuclear proteins.[228,231,233a–c] Glycosylation of protein –OH groups can occur on either the luminal or cytosolic faces of the ER membranes.[234] The external *O*-linked glycoproteins often have large clusters of oligosaccharides attached to –OH groups of serine or threonine, but cytosolic proteins may carry only a small number of small oligosaccharides.

Of great importance are the **blood group determinants** which are discussed in Box 4-C. The ABO determinants are found at the nonreducing ends of *O*-linked oligosaccharides. Conserved Ser/Thr sites in the epidermal growth factor domains (Table 7-3) of various proteins carry ***O*-linked fucose**.[235]

The secreted **mucins** are unique in having clusters of large numbers of oligosaccharides linked by ***N*-acetylgalactosamine** to serine or threonine of the polypeptide.[236] The following core structures predominate.[237] These may be lengthened or further branched by the particular variety of glycosyltransferases present in a tissue and by their specificities.[233] The human genome contains at least nine mucin genes.[238] The large apomucins contain central domains with tandem repeats rich in Ser, Thr, Gly, and Ala and flanked at the ends by cysteine-rich domains.[239] For example, porcine submaxillary mucins are encoded by a gene with at least three alleles that encode 90, 125, or 133 repeats. The polypeptide may contain as many as 13,288 residues. N-terminal cysteine-rich regions are involved in dimer formation.[240]



core 1 $\text{Gal}\beta1 \rightarrow 3\ \text{GalNAc} - \text{Ser/Thr}$

core 2 $\text{Gal}\beta1 \rightarrow 3\ \underset{\substack{6 \\ \uparrow \\ \text{GalNAc}\beta1}}{\text{GalNAc}} - \text{Ser/Thr} - \text{Ser/Thr}$

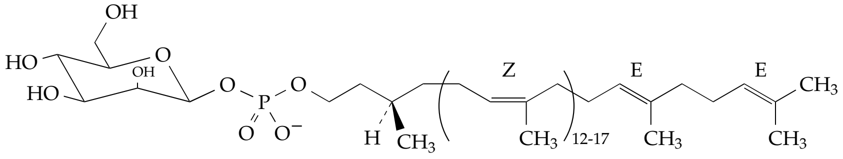
core 3 $\text{GlcNAc}\beta1 \rightarrow 3\ \text{GalNAc} - \text{Ser/Thr}$

core 4 $\text{Gal}\beta1 \rightarrow 4\ \text{GlcNAc}\beta1 \rightarrow 6\ \underset{\substack{3 \\ \uparrow \\ \text{GlcNAc}\beta1}}{\text{GalNAc}} - \text{Ser/Thr}$

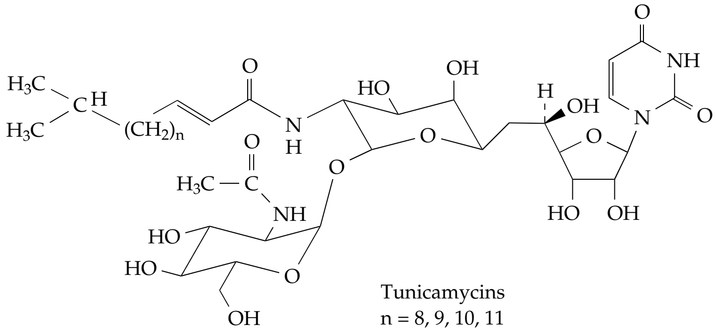
3. Assembly of *N*-Linked Oligosaccharides on Glycosyl Carrier Lipids

In eukaryotes the biosynthesis of the *N*-linked oligosaccharides of glycoproteins depends upon the polyprenyl alcohols known as **dolichols**, which are present in membranes of the endoplasmic reticulum. They contain 16–20 prenyl units, of which the one bearing the OH group is completely saturated as a result of the action of an NADPH-dependent reductase on the unsaturated precursor.[241] The predominant dolichol in mammalian cells contains 19 prenyl units. The structure of its mannosyl phosphate ester, one of the intermediates in the oligosaccharide synthesis, is illustrated below. The fully extended 95-carbon dolichol has a length of almost 10 nm, four times greater than that of oleic acid and twice the thickness of the nonpolar membrane bilayer core. The need for this great length is not clear nor is it clear why the first prenyl unit must be saturated for good acceptor activity.

The assembly of the oligosaccharides that will become linked to Asn residues in proteins occurs on the phosphate head of dolichol-*P*. The process begins on the cytoplasmic face of the membrane and within the lumen of the rough or smooth ER and continues within cisternae of the Golgi apparatus.[234,242–245] The initial transfer of GlcNAc-*P* to dolichol-*P* (Fig. 20-6, step *a*) appears to occur on the cytoplasmic face of the ER and is specifically inhibited by **tunicamycin**.[246,247] As the first "committed reaction" of *N*-glycosylation, it is regulated by a variety of hormonal and other factors.[248,249] The reaction takes place cotranslationally as the still unfolded peptide chain leaves the ribosome.[242] The oligosaccharide, still attached to the dolichol, continues to grow on the cytosolic surface of the ER membrane by transfer of GlcNAc and five residues of mannose from their sugar nucleotide forms (Fig. 20-6, steps *b* and *c*).[249a] The intermediate Dol-*P*-*P*-$GlcNAc_2Man_5$ crosses the membrane bilayer (Fig. 20-6, step *d*), after which mannosyl and glucosyl units are added (steps *e* and *f*). These sugars are carried across the membrane while attached to dolichol.



Man-*P*-dolichol



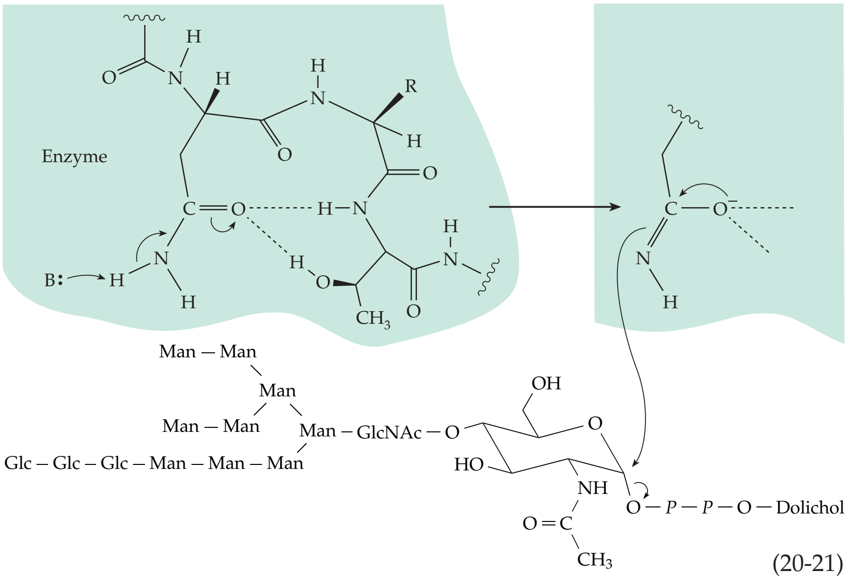
Tunicamycins
n = 8, 9, 10, 11

The completed 14-residue branched oligosaccharide Glc$_3$Man$_9$GlcNAc$_2$, with the structure indicated in Fig. 20-6, is then transferred to a suitable asparagine side chain (step *g*). This may be on a newly synthesized protein or on a still-growing polypeptide chain that is being extruded through the membrane into the lumenal space of the rough ER (Eq. 20-21; Fig. 20-6). The glycosylation site is often at the sequence Asn-X-Ser(Thr), which is likely to be present at a beta bend in the folded protein. Bends of the type illustrated in Eq. 20-21 and stabilized by the asparagine side chain are apparently favored.[250] In such a bend the –OH of the serine or threonine helps to polarize the amide group of the Asn side chain, perhaps enolizing it and generating a nucleophilic center that can participate in a displacement reaction[250,251] as indicated in Eq. 20-21. The **oligosaccharyltransferase** that catalyzes the reaction is a multisubunit protein. As many as eight different subunits have been reported for the enzyme in yeast. Genes for at least five of these are essential.[250–252b] One subunit serves to recognize suitable glycosylation sites.

Trimming of glycoprotein oligosaccharides. After transfer to glycoproteins the newly synthesized oligosaccharides undergo trimming, the hydrolytic removal of some of the sugar units, followed by addition of new sugar units to create the finished glycoproteins. The initial glycosylation process ensures that the glycoproteins remain in the lumen of the ER or within vesicles or cisternae separated from the cytoplasm. The subsequent processing appears to allow the cell to **sort** the proteins. Some remain attached to membranes and take up residence within ER, Golgi, or plasma membrane. Others are passed outward into transfer vesicles, Golgi, and secretion vesicles. A third group enter **lysosomes**. A series of specific inhibitors of trimming reactions, some of whose structures are shown in Fig. 20-7, has provided important insights.[253–255] Use of these inhibitors, together with immunochemical methods and study of yeast mutants,[250,252,256] is enabling us to learn many details of glycoprotein biosynthesis.

Whereas the formation of dolichol-linked oligosaccharides occurs in an identical manner in virtually all eukaryotic cells, trimming is highly variable as is the addition of new monosaccharide units.[257–257b] The major pathway for mammalian glycoproteins is shown in Fig. 20-7. Specific hydrolases in the ER remove all of the glucosyl units and one to three mannosyl units.[258] Removal of additional mannosyl residues occurs in the cis Golgi, to give the pentasaccharide core Man$_3$GlcNAc$_2$ which is common to all of the complex N-linked oligosaccharides. However, partial trimming without additional glycosylation produces some "high mannose" oligosaccharides.[258a] Removal of glucose may be necessary to permit some glycoproteins to leave the ER.

Sulfate groups and in some cases fatty acyl groups[259] may also be added. The exact composition of the oligosaccharides may depend upon the condition of the cell and may be altered in response to external influences.[257] Oligosaccharides attached to proteins that remain in the ER membranes may undergo



(20-21)

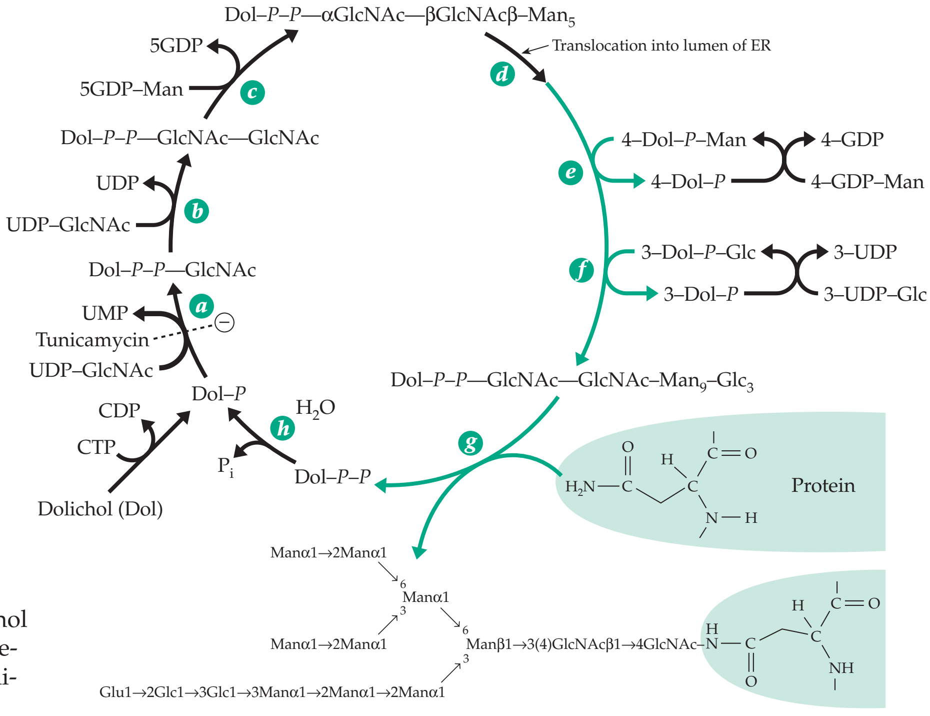

Dol–P–P—αGlcNAc—βGlcNAcβ–Man5
Translocation into lumen of ER
5GDP
5GDP–Man
c
d
Dol–P–P—GlcNAc—GlcNAc
UDP
b
UDP–GlcNAc
4–Dol–P–Man
4–GDP
e
4–Dol–P
4–GDP–Man
3–Dol–P–Glc
3–UDP
f
3–Dol–P
3–UDP–Glc
Dol–P–P—GlcNAc
UMP
a
Tunicamycin
UDP–GlcNAc
Dol–P
Dol–P–P—GlcNAc—GlcNAc–Man9–Glc3
CDP
H2O
CTP
h
g
Pi
Dol–P–P
Dolichol (Dol)
O
H
C=O
H2N—C
C
Protein
N—H
Manα1→2Manα1
6
Manα1
3
6
Manα1→2Manα1
Manβ1→3(4)GlcNAcβ1→4GlcNAc–N—C
H
C=O
C
O
NH
3
Glu1→2Glc1→3Glc1→3Manα1→2Manα1→2Manα1

Figure 20-6 Biosynthesis of the dolichol diphosphate-linked oligosaccharide precursor to glycoproteins. The site of inhibition by tunicamycin is indicated.

very little trimming. However, the three glucosyl residues are usually removed by the glucosidases present in the rough ER. Some plant glycoproteins of the high-mannose type undergo no further processing.

Extensions and terminal elaborations. Even before trimming is completed, addition of new residues begins within the medial cisternae.[260] In mammalian Golgi, *N*-acetylglucosamine is added first. Galactose, sialic acid, and often fucose are then transferred from their activated forms to create such elaborate oligosaccharides as that shown in Fig. 4-17. As many as 500 glycosyltransferases, having different specificities for glycosyl donor and glucosyl acceptor, may be involved.[245] Extensions of the basic oligosaccharide structure often contain polylactosamine chains, branches with fucosyl residues, and sulfate groups.[245,261] More than 14 sialyltransferases place sialic acid residues, often in terminal positions, on these cell surface oligosaccharides.[262,263]

The cell wall of the yeast *Saccharomyces* is rich in **mannoproteins** that contain 50–90% mannose.[264] The ~250-residue mannan chains consist of an α-1,6-linked backbone with mono-, di-, and tri-mannosyl branches. These are attached to the same core structure as that of mammalian oligosaccharides. All of the core structures are formed in a similar way.[258,265] The mannoproteins may serve as a "filler" to occupy spaces in a cell wall constructed from β-1,3- and β-1,6-linked glycans and chitin. All of the four components, including the mannoproteins, are covalently linked together.[266] As was emphasized in Chapter 4 (pp. 186–188) glycoproteins serve many needs in biological recognition. The N-linked oligosaccharides play a major role in both animals and plants.[266a–c] Use of mass spectrometry, new automated methods of oligosaccharide synthesis,[266d] and development of new synthetic inhibitors[266c] are all contributing to current studies of what is commonly called "glycobiology."

The perplexing Golgi apparatus. First observed by Camillo Golgi[267,268] in 1898, the stacked membranes, now referred to simply as Golgi, remain somewhat mysterious.[268–271] There are at least three functionally distinct sets of Golgi cisternae, the **cis** (nearest the nucleus), **medial**, and **trans**. An additional series of tubules referred to as the **trans Golgi network** lies between the Golgi and the cell surface and may be the site at which lysosomal enzymes are sorted from proteins to be secreted.[260,272] Immunochemical staining directed toward specific glycosidases and glycosyl transferases suggested that the trimming reactions of glycoproteins start in the ER and continue as the proteins pass outward successively from one compartment of the Golgi to the next (Fig. 20-8). This has been the conventional view since the 1970s. The movement of the glycoproteins between compartments is thought to take place in small vesicles using a rather elaborate system of specialized proteins. Some of these coat the vesicles[273,274] while others target the vesicles to specific locations, e.g., the lysosomes[275] or the plasma membranes where they may be secreted.[273,277] A host of regulatory G-proteins assist these complex processes and drive them via hydrolysis of GTP.[271,278]

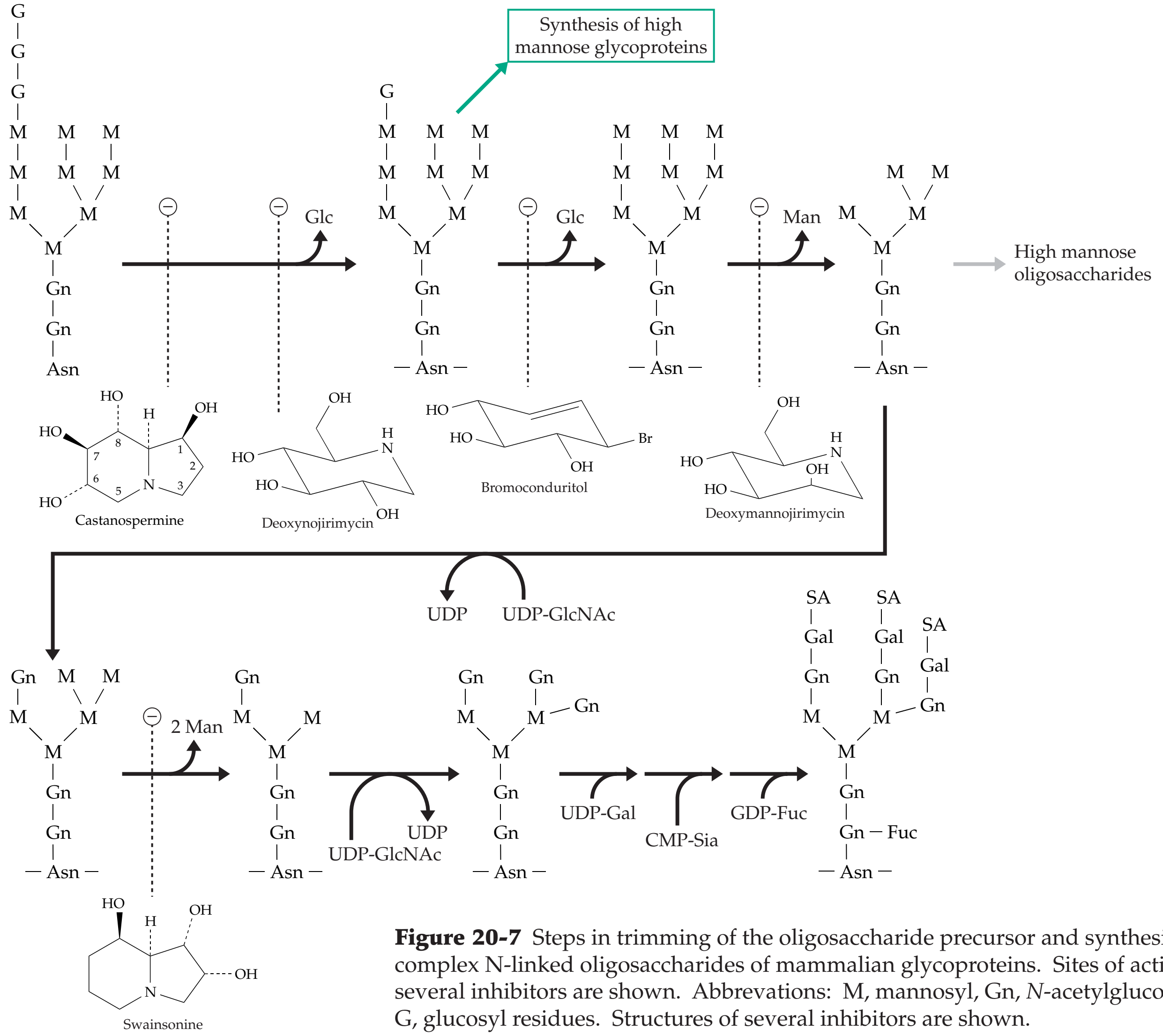

Synthesis of high mannose glycoproteins
G
M
Gn
Asn
Glc
Man
High mannose oligosaccharides
Castanospermine
Deoxynojirimycin
Bromoconduritol
Deoxymannojirimycin
UDP
UDP-GlcNAc
2 Man
UDP-Gal
CMP-Sia
GDP-Fuc
SA
Gal
Fuc
Swainsonine

Figure 20-7 Steps in trimming of the oligosaccharide precursor and synthesis of complex N-linked oligosaccharides of mammalian glycoproteins. Sites of action of several inhibitors are shown. Abbrevations: M, mannosyl, Gn, *N*-acetylglucosaminyl, G, glucosyl residues. Structures of several inhibitors are shown.

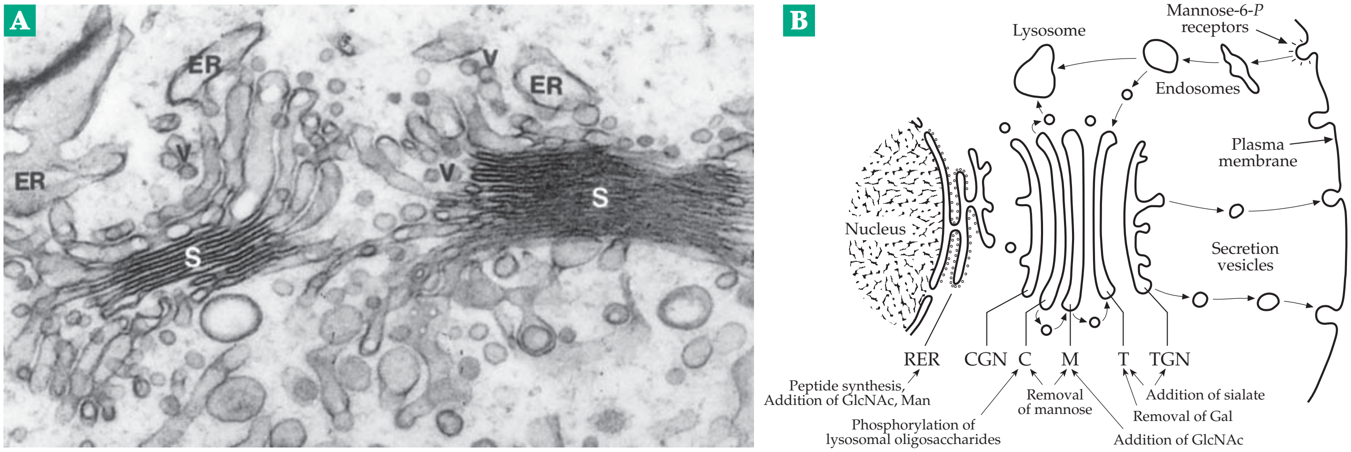
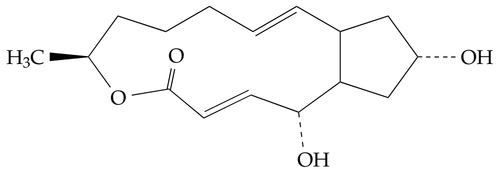

A
ER
V
S
B
Lysosome
Mannose-6-*P* receptors
Endosomes
Plasma membrane
Nucleus
Secretion vesicles
RER
CGN
C
M
T
TGN
Peptide synthesis, Addition of GlcNAc, Man
Phosphorylation of lysosomal oligosaccharides
Removal of mannose
Addition of sialate
Removal of Gal
Addition of GlcNAc

Figure 20-8 (A) Electron micrograph showing a transverse section through part of the Golgi apparatus of an early spermatid. Cisternae of the ER, Golgi stacks (S), and vesicles (V) can be seen. Curved arrows point to associated tubules. Magnification X45,000.[276] Courtesy of Y. Clermont. (B) Scheme showing functions of endoplasmic reticulum, transfer vesicles, Golgi apparatus, and secretion vesicles in the metabolism of glycoproteins.

While most proteins synthesized in the ER follow the exocytotic pathway through the Golgi, some are retained in the ER and some that pass on through the Golgi are returned to the ER.[279] In fact, such **retrograde transport** can carry some proteins taken up by endocytosis through the plasma membrane and through the Golgi to the ER where they undergo *N*-glycosylation. Retrograde transport is essential for recycling of plasma membrane proteins and lipids. The forward flow of glycoproteins and membrane components from the ER to the Golgi can be blocked by the fungal macrocyclic lactone **brefeldin A**. In cells treated with this drug, which inactivates a small CTP-binding protein,[280] the Golgi apparatus is almost completely resorbed into the ER by retrograde transport. Proteins remaining in the ER undergo increased *O*-glycosylation as well as unusual types of *N*-glycosylation.

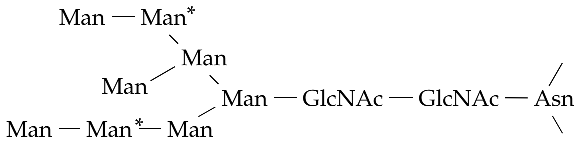


Brefeldin A

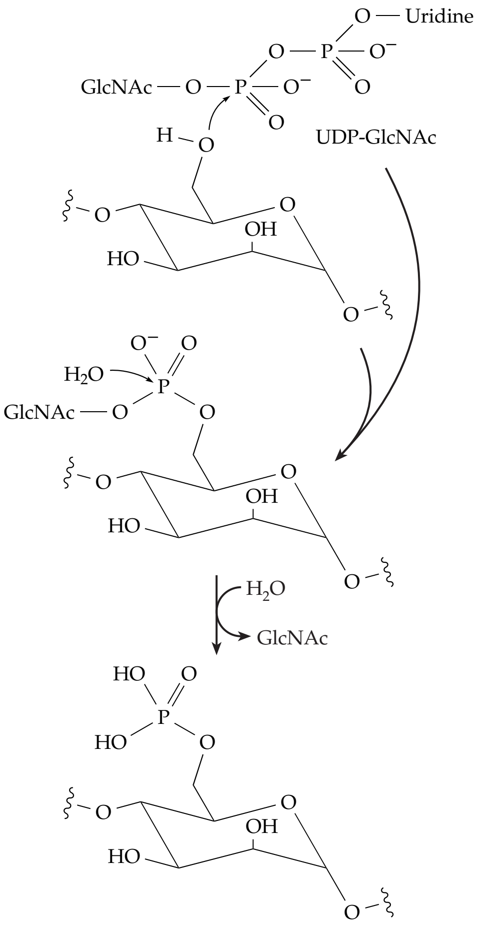
Although the conventional view of flow through the Golgi is generally accepted, it is difficult to distinguish it from an alternative explanation: *The Golgi compartments may move outward continuously while retrograde transport occurs via the observed vesicles.*[268,272] Some evidence for this **cisternal maturation model** has been known for many years but was widely regarded as reflecting unusual exceptions to the conventional model. In fact, both views could be partially correct; vesicular transport may function in both directions.[280a] High-resolution tomographic images are also altering our view of the Golgi.[280b]

The proteins of Golgi membranes are largely integral membrane proteins and peripheral proteins associated with the cytosolic face. Some of the integral membrane proteins are the oligosaccharide-modifying enzymes, which protrude into the Golgi lumen.[280c,280d] Many other proteins participate in transport,[280d–f] docking, membrane fusion,[280g,h] and acidification of Golgi compartments.[280i] Many of the first studies of vesicular transport were conducted with synaptic vesicles and are considered in Chapter 30 (see Fig. 30-20). Other aspects of membrane fusion and transport are discussed in Chapter 8. A group of specialized Golgi proteins, the **golgins**, are also present. They are designated golgin-84, -95, -160, -245, and -376 (giantin or macrogolgin) and were identified initially as human **autoantigens** (Chapter 31), appearing in the blood of persons with autoimmune disorders such as Sjögren's syndrome.[281,282] Another protein of molecular mass ~130 kDa, and which appears to be specific to the trans-Golgi network, has been found in human serum of patients with renal vasculitis.[283]

Lysosomal enzymes. Various **sorting signals** are encoded within proteins. These include the previously mentioned C-terminal KDEL (mammals) or HDEL (yeast) amino acid sequence, which serves as a retrieval signal for return of proteins from the Golgi to the ER (p. 521).[279,284] Other sorting signals are provided by the varied structures of the oligosaccharides attached to glycoproteins. These sugar clusters convey important biological information, which is "decoded" in the animal body by interaction with various **lectins** that serve as receptors.[285] This often leads to endocytosis of the glycoprotein. An example is provided by the more than 50 proteins that are destined to become lysosomal enzymes and which undergo phosphorylation of 6-OH groups of the mannosyl residue marked by asterisks on the following structure. This is an *N*-linked oligosaccharide that has been partially trimmed. The phosphorylation is accomplished in two steps by enzymes present in the cis Golgi compartment (Eq. 20-22) An ***N*-acetylglucosaminylphosphotransferase** transfers phospho-GlcNAc units from UDP-GlcNAc onto the 6-OH groups of mannosyl residues. These must be recognized by the phosphotransferase as appropriate.[286,287] Then a hydrolase cleaves off GlcNAc.



The proteins carrying the mannose 6-phosphate groups bind to one of two different **mannose 6-*P* receptors** present in the Golgi membranes and are subsequently transported in clathrin-coated vesicles to endosomes where the low pH causes the proteins to dissociate from the receptors, which may be recycled.[288–290a] The hydrolytic enzymes are repackaged in lysosomes. The same mannose 6-*P* receptors also appear on the external surface of the plasma membrane allowing many types of cells to take up lysosomal enzymes that have "escaped" from the cell. These proteins, too, are transported to the lysosomes. The mannose 6-*P* receptors have a dual function, for they also remove insulinlike growth factor from the circulation, carrying it to the lysosomes for degradation.[287,290] Most Man 6-*P* groups are removed from proteins once they reach the lysosomes but this may not always be true.[291] Not all lysosomal proteins are recruited by the mannose 6-*P* receptors. Some lysosomal membrane proteins are sorted by other mechanisms.[292]

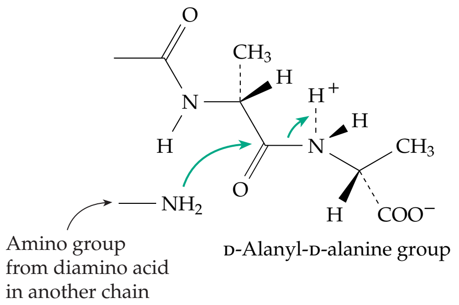


(20-22)

The hepatic asialoglycoprotein (Gal) receptor. A variety of proteins are taken out of circulation in the blood by the hepatocytes of the liver. Serum glycoproteins bearing sialic acid at the ends of their oligosaccharides (see Fig. 20-7) have relatively long lives, but if the sialic acid is removed by hydrolysis, the exposed galactosyl residues are recognized by the multisubunit **asialoglycoprotein receptor**.[293–295] The bound proteins are then internalized rapidly via the coated pit pathway and are degraded in the lysosomes. Other receptors, including those that recognize transferrin, low-density lipoprotein, α_2 macroglobulin, and T lymphocyte antigens, also depend upon interaction with oligosaccharides.[296]

E. Biosynthesis of Bacterial Cell Walls

The outer surfaces of bacteria are rich in specialized polysaccharides. These are often synthesized while attached to lipid membrane anchors as indicated in a general way in Eq. 20-20.[136,296a] One of the specific biosynthetic cycles (Fig. 20-9) that depends upon undecaprenol phosphate is the formation of the **peptidoglycan** (murein) layer (Fig. 8-29) of both gram-negative and gram-positive bacterial cell walls. Synthesis begins with attachment of L-alanine to the OH of the lactyl group of UDP-*N*-acetylmuramic acid in a typical ATP-requiring process (Fig. 20-9, step *a*).[297] Next D-glutamic acid, *meso*-diaminopimelate (Fig. 8-29), or L-lysine, and D-alanyl-D-alanine are joined in sequence, each in another ATP-requiring step.[298–301d] The entire unit assembled in this way is transferred to undecaprenol phosphate with creation of a pyrophosphate linkage (step *e*). An *N*-acetylglucosamine unit is added by action of another transferase (step *f*), and in an ATP-requiring process ammonia is sometimes added to cap the free α-carboxyl group of the D-glutamyl residue (step *g*). In *Staphylococcus aureus* and related gram-positive bacteria five glycyl units are also added, each from a molecule of glycyl-tRNA (green arrows in Fig. 20-9). The completely assembled repeating unit, together with the connecting peptide chain needed in the crosslinking reaction, is transferred onto the growing chain (step *h*). As in formation of dextrans, growth is by insertion of the repeating unit at the reducing end of the chain. The polyprenyl diphosphate is released, and the cycle is completed by the action of a pyrophosphatase (step *i*). This step is blocked by **bacitracin**, an antibiotic which forms an unreactive complex with the polyprenyl diphosphate carrier. Completion of the peptidoglycan requires crosslinking. This is accomplished by displacement of the terminal D-alanine of the pentapeptide by attack by the $-NH_2$ group of the diaminopimelate or lysine or other diamino acid (see also Fig. 8-29).[301e]



Because the peptidoglycan layer must resist swelling of the bacteria in media of low osmolarity, it must be strong and must enclose the entire bacterium. At the same time the bacterium must be able to grow in size and also to divide. For these reasons bacteria must continuously not only synthesize peptidoglycan but also degrade it.[302,303] The latter is accomplished by hydrolytic cleavage using cell wall enzymes to the ***N*-acetylglucosamine-anhydro-*N*-acetylmuramate-tripeptide** (GlcNAc-1,6-anhydro-MurNAc-L-Ala-D-Glu-A_2pm) fragment.[304,305] A hydrolase cuts the peptide bridge.[305a] This process is probably essential to formation of new growing points for expansion of the murein layer. Most of the peptide fragments that are released in the periplasm are transported back into the cytosol.

The anhydroMurNAc is removed, and new UDP-MurNAc and D-Ala-D-Ala units are added salvaging the tripeptide unit. The repaired UDP-MurNA-pentapeptide can then reenter the biosynthetic pathway (Fig. 20-9).

The O-antigens and lipid A. A cluster of sugar units of specific structure makes up the repeating unit of the "O-antigen" of *Salmonella*. The many structural variations in this surface polysaccharide account for the over two thousand serotypes of *Salmonella* (p. 180).[121,306] As is illustrated in Fig. 8-30, the O-antigen is a repeating block polymer that is attached to a complex lipopolysaccharide "core" and a hydrophobic membrane anchor known as lipid A (Figs. 8-28 and 8-30).[307–308a] Lipid A and the attached core and O-antigen are synthesized inside the bacterial cell by enzymes found in the cytoplasmic membrane.[309] The complete lipopolysaccharide units are then translocated from the inner membrane to the outer membrane of the bacteria. The synthesis of the O-antigen is understood best. Consider the following group E3 antigen, where Abe is abequose (Fig. 4-15) and Rha is rhamnose:

$$\begin{array}{l} \quad\text{Abe}\alpha 1 \\ \qquad\quad\downarrow \\ \qquad\quad 4 \\ (\rightarrow 6\,\text{Gal}\beta 1\rightarrow 6\,\text{Man}1\rightarrow 4\,\text{Rha}1\rightarrow)_n \end{array}$$

Assembly of this repeating unit begins with the transfer of a *phosphogalactosyl* unit from UDP-Gal to the phospho group of the lipid carrier undecaprenol phosphate. The basic reaction cycle is much like that in Fig. 20-9 for assembly of a peptidoglycan.

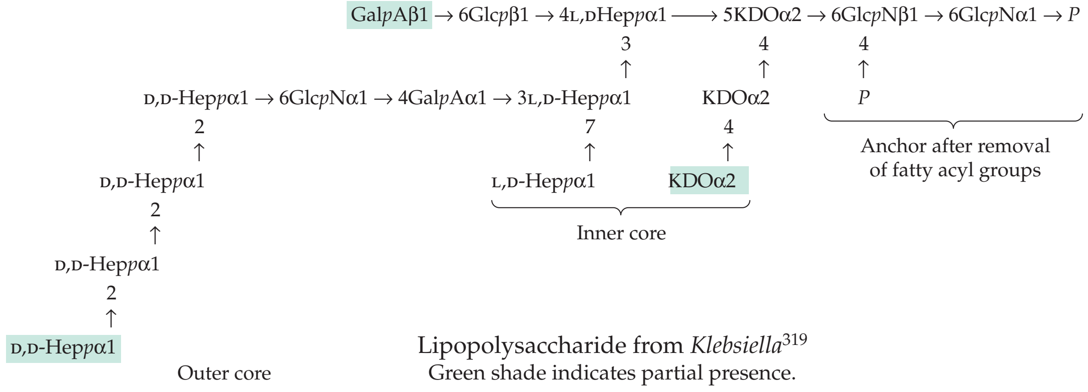
The oligosaccharide repeating unit of the O-antigen is constructed by the consecutive transferring action of three more transferases. For the antigen shown above, one enzyme transfers a rhamnosyl unit, another a mannosyl unit, and another an abequosyl unit from the appropriate sugar nucleotides. Then the entire growing O-antigen chain, which is attached to a second molecule of undecaprenol diphosphate, is transferred onto the end of the newly assembled oligosaccharide unit. In effect, the newly formed oligosaccharide is inserted at the reducing end of the growing chain just as in Fig. 20-9. Elongation continues by the transfer of the entire chain onto yet another tetrasaccharide unit. As each oligosaccharide unit is added, an undecaprenol diphosphate unit is released and a phosphatase cleaves off the terminal phospho group to regenerate the original undecaprenol phosphate carrier. When the O-antigen is long enough, it is attached to the rest of the lipopolysaccharide.

The lipid A anchor is also based on a carbohydrate skeleton. Its assembly in *E. coli*,[307] which requires nine enzymes, is depicted in Fig. 20-10. *N*-Acetylglucosamine 6-*P* is acylated at the 3-position[310] and after deacetylation[311] at the 2-position. As shown in this figure, acylation is accomplished by transfer of hydroxymyristoyl groups from acyl carrier protein (ACP). Two molecules of the resulting UDP-2,3-diacyl-GlcN are then joined via the reactions shown to give the acylated disaccharide precursor to lipid A. Stepwise transfer of KDO, L-*glycero*-D-*manno*-heptose, and other monosaccharide units from the appropriate sugar nucleotides and further acylation follows (Fig. 20-10). The assembled O-antigen chain is transferred from undecaprenol diphosphate onto the lipopolysaccharide core. This apparently occurs on the periplasmic surface of the plasma membrane. If so, the core lipid domain must be flipped across the plasma membrane before the O-antigen chain is attached.[312] Less is known about the transport of the completed lipopolysaccharide across the periplasmic space and into the outer membrane.

The core structures of the lipopolysaccharides vary from one species to another or even from one strain of bacteria to another. All three domains (lipid A, core, and O-antigen) contribute to the antigenic properties of the bacterial surface[313] and to the virulence of the organism.[313,314] Nitrogen-fixing strains of *Rhizobium* require their own peculiar lipopolysaccharides for successful symbiosis with a host plant.[315] However, there are some features common to most lipopolysaccharides. Two to three residues of KDO are usually attached to the acylated diglucosamine anchor, and these are often followed by 3 – 4 heptose rings.[316–318]

The structure of the inner core regions of a typical lipopolysaccharide from *E. coli* is indicated in Fig. 8-30. The complete structure of the lipopolysaccharide from a strain of *Klebsiella* is shown at the top of the next page.[319] Here L, D-Hep*p* is D-*manno*-heptopyranose and D, D-Hep*p* is D-*glycero*-D-*manno*-heptose. As in this case, the outer core often contains several different hexoses. The lipopolysaccharide of *Neisseria meningititis* has sialic acid at the outer end.[320] However, the major virulence factor for this organism, which is a leading cause of bacterial meningitis in young children, is a capsule of poly(ribosyl)ribitol phosphate that surrounds the cell.[321] *Haemophilus influenzae*, a common cause of ear infections and meningitis in children, has no *O*-antigen but a more highly branched core oligosaccharide than is present in *E. coli*.[321a] *Legionella* has its own variations.[321b]

Gram-positive bacteria. Although their outer coatings are extremely varied, all gram-positive bacteria have a peptidoglycan similar to that of gram-negative bacteria but often containing the intercalated pentaglycine bridge indicated in Fig. 8-29. However, the peptidoglycan of gram-positive bacteria is 20–50 nm thick, as much as ten times thicker than that of *E. coli*. Furthermore, the peptidoglycan is intertwined with the anionic polymers known as teichoic acids and


GalpAβ1 → 6Glcpβ1 → 4L,DHeppα1 → 5KDOα2 → 6GlcpNβ1 → 6GlcpNα1 → P
3
4
4
D,D-Heppα1 → 6GlcpNα1 → 4GalpAα1 → 3L,D-Heppα1
KDOα2
P
2
7
4
D,D-Heppα1
L,D-Heppα1
KDOα2
2
D,D-Heppα1
2
D,D-Heppα1
Anchor after removal of fatty acyl groups
Inner core
Outer core

Lipopolysaccharide from *Klebsiella*[319]
Green shade indicates partial presence.

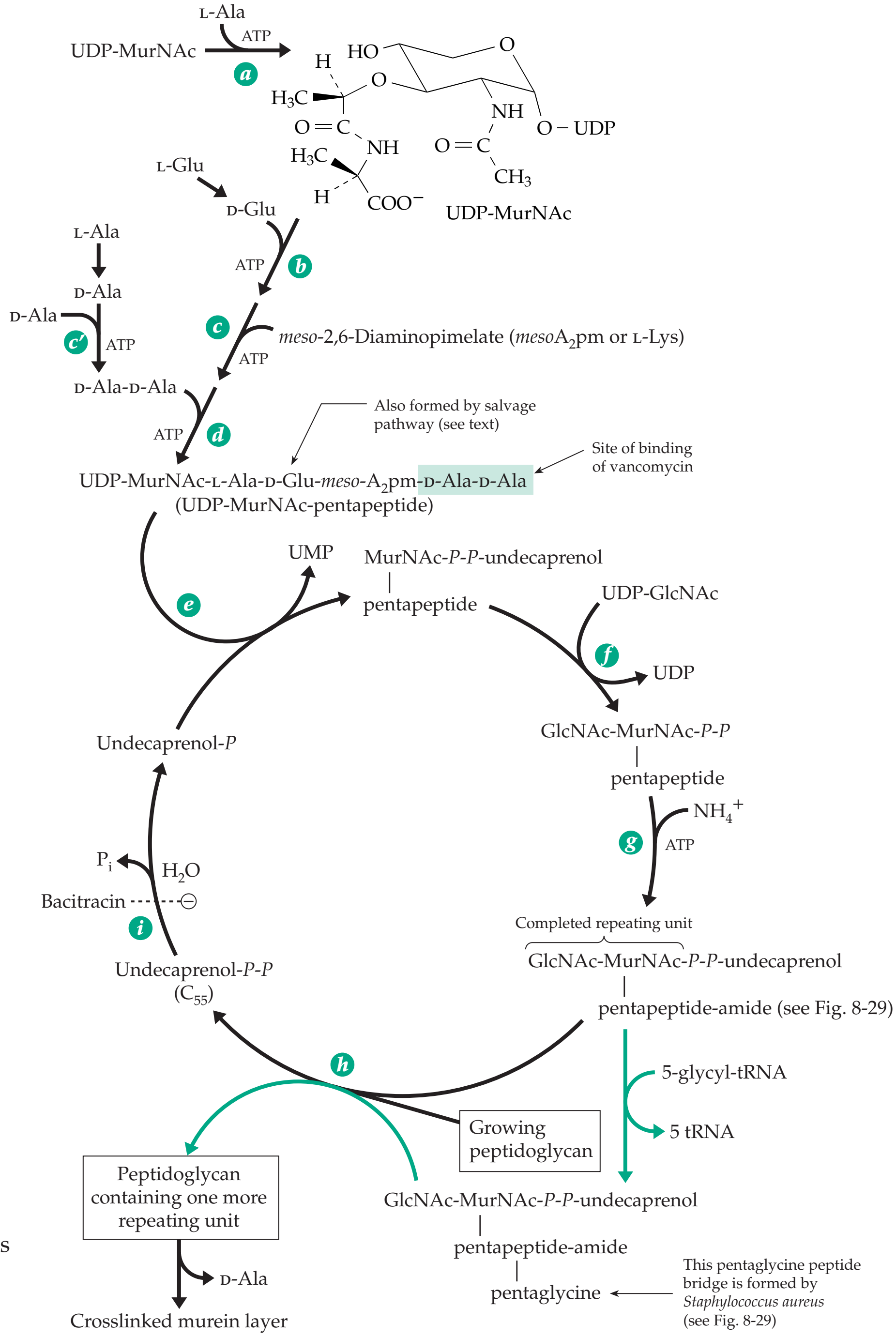

L-Ala
ATP
UDP-MurNAc
a
HO
O
H
O
H_3C
NH
O – UDP
O=C
NH
O=C
H_3C
CH_3
H
COO⁻
UDP-MurNAc
L-Glu
D-Glu
L-Ala
ATP
b
D-Ala
D-Ala
c
meso-2,6-Diaminopimelate (mesoA$_2$pm or L-Lys)
c′
ATP
ATP
D-Ala-D-Ala
ATP
d
Also formed by salvage pathway (see text)
Site of binding of vancomycin
UDP-MurNAc-L-Ala-D-Glu-meso-A$_2$pm-D-Ala-D-Ala
(UDP-MurNAc-pentapeptide)
UMP
MurNAc-P-P-undecaprenol
pentapeptide
e
UDP-GlcNAc
f
UDP
Undecaprenol-P
GlcNAc-MurNAc-P-P
pentapeptide
NH_4^+
g
ATP
P_i
H_2O
Bacitracin
i
Completed repeating unit
GlcNAc-MurNAc-P-P-undecaprenol
pentapeptide-amide (see Fig. 8-29)
Undecaprenol-P-P
(C_{55})
h
5-glycyl-tRNA
5 tRNA
Growing peptidoglycan
Peptidoglycan containing one more repeating unit
GlcNAc-MurNAc-P-P-undecaprenol
pentapeptide-amide
pentaglycine
This pentaglycine peptide bridge is formed by Staphylococcus aureus (see Fig. 8-29)
D-Ala
Crosslinked murein layer

Figure 20-9 Biosynthesis of bacterial peptidoglycans. See Fig. 8-29 for details of the peptidoglycan structures. Green arrows show alternative route used by gram-positive bacteria.

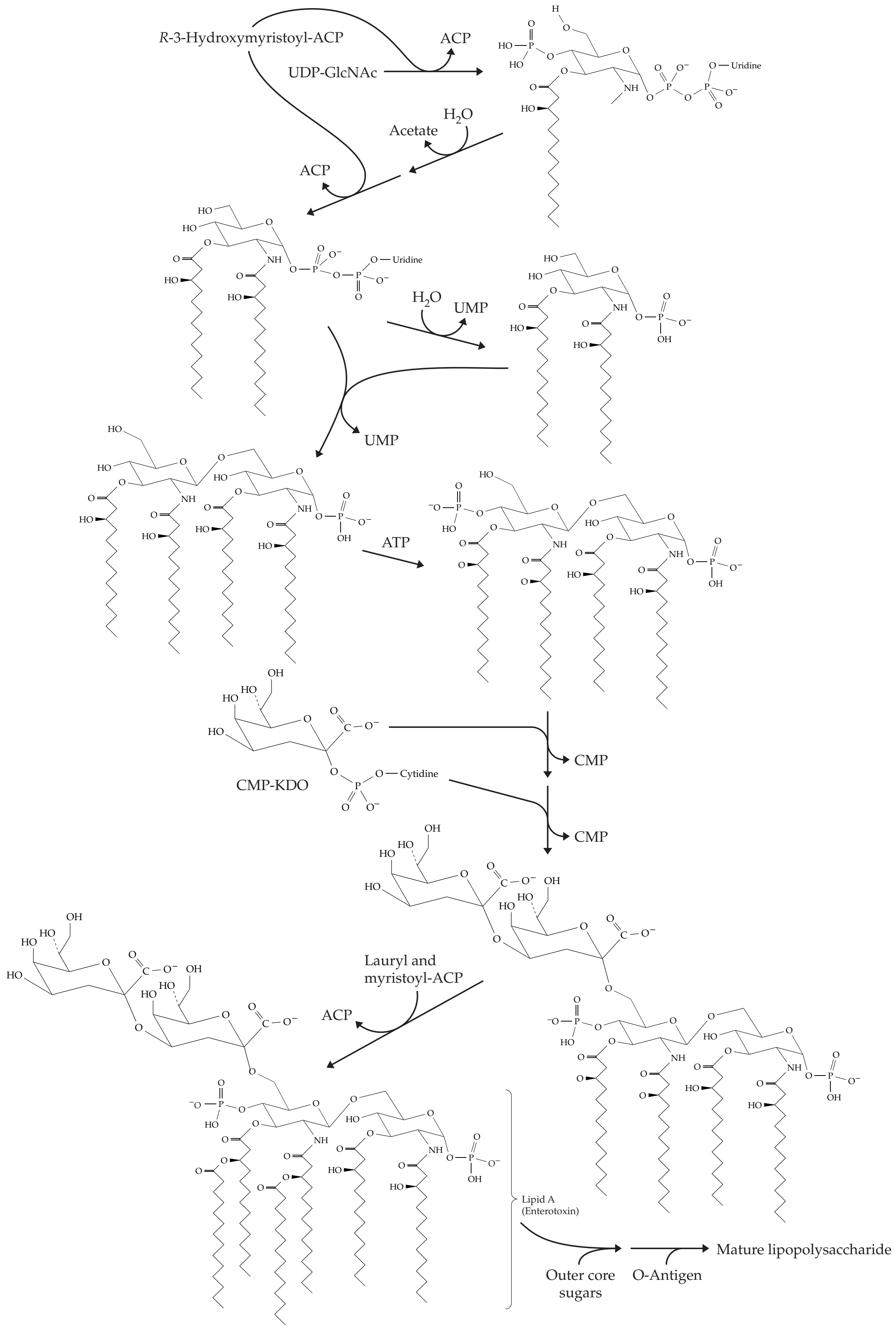
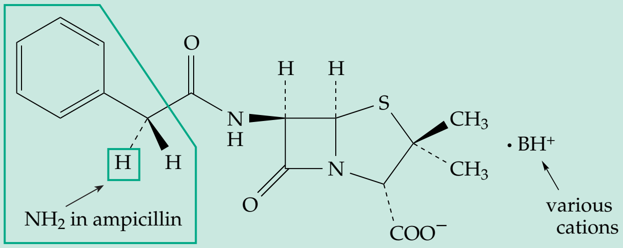

R-3-Hydroxymyristoyl-ACP
ACP
UDP-GlcNAc
Uridine
H2O
Acetate
ACP
Uridine
H2O
UMP
UMP
ATP
CMP-KDO
Cytidine
CMP
CMP
Lauryl and
myristoyl-ACP
ACP
Lipid A
(Enterotoxin)
Outer core
sugars
O-Antigen
Mature lipopolysaccharide

Figure 20-10 Proposed biosynthetic route for synthesis of lipid A and the mature lipopolysaccharide of the *E. coli* cell wall. After C. R. H. Raetz *et al.*[307]

BOX 20-G PENICILLINS AND RELATED ANTIBIOTICS



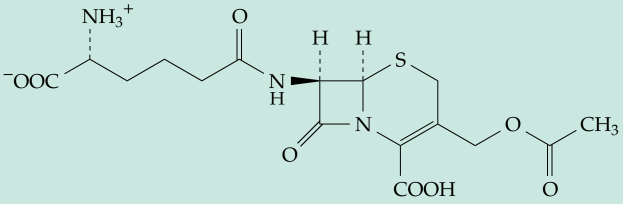
Penicillin G
(Benzylpenicillin)

Many organisms produce chemical substances that are toxic to other organisms. Some plants secrete from their roots or leaves compounds that block the growth of other plants. More familiar to us are the medicinal antibiotics produced by fungi and bacteria. The growth inhibition of one kind of organism upon another was well known in the last century, e.g., as reported by Tyndall[a,b] in 1876. The beginning of modern interest in the phenomenon is usually attributed to Alexander Fleming, who, in 1928, noticed the inhibition of growth of staphylococci by *Penicillium notatum*. His observation led directly to the isolation of penicillin, which was first used on a human patient in 1930. The early history as well as the subsequent purification, characterization, synthesis, and development as the first major antibiotic has been recorded in numerous books and articles.[c–i] During the same time period, Rene Dubos isolated the peptide antibiotics **gramicidin** and **tyrocidine**.[j] A few years later **actinomycin** (Box 28-A) and **streptomycin** (Box 20-B) were isolated from soil actinomycetes (streptomyces) by Waksman, who coined the name **antibiotic** for these compounds. Streptomycin was effective against tuberculosis, a finding that helped to stimulate an intensive search for additional antibacterial substances. Since that time, new antibiotics have been discovered at the rate of more than 50 a year. More than 100 are in commercial production.

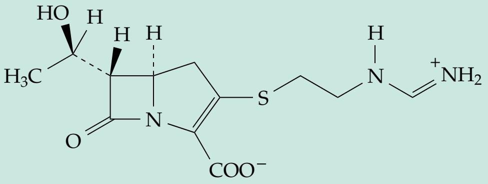
Major classes of antibiotics include more than 200 peptides such as the gramicidins, bacitracin, tyrocidines and valinomycin (Fig. 8-22)[k]; more than 150 **penicillins, cephalosporins,** and related compounds; **tetracyclines** (Fig. 21-10); the **macrolides**, large ring lactones such as the **erythromycins** (Fig. 21-11); and the **polyene** antibiotics (Fig. 21-10).

Penicillin was the first antibiotic to find practical use in medicine. Commercial production began in the early 1940s and benzylpenicillin (penicillin G), one of several natural penicillins that differ in the R group boxed in the structure above, became one of the most important of all drugs. Most effective against gram-positive bacteria, at higher concentrations it also attacks gram-negative bacteria including *E. coli*. The widely used semisynthetic penicillin **ampicillin** (R = D-α-aminobenzyl) attacks both gram-negative and gram-positive organisms. It shares with penicillin extremely low toxicity but some danger of allergic reactions. Other semisynthetic penicillins are resistant to β-lactamases, enzymes produced by penicillin-resistant bacteria which cleave the four-membered β-lactam ring of natural penicillins and inactivate them.

Closely related to penicillin is the antibiotic **cephalosporin C**. It contains a D-α-aminoadipoyl side chain, which can be replaced to form various semisynthetic cephalosporins. **Carbapenems** have similar structures but with CH_2 replacing S and often a different chirality in the lactam ring.



Cephalosporin C



Imipenem, a carbapenem antibiotic of last resort

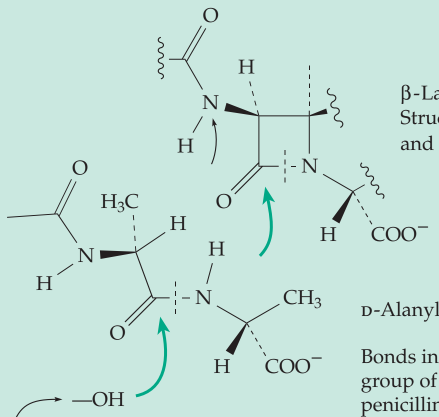
These and other related β-lactams are medically important antibacterial drugs whose numbers are increasing as a result of new isolations, synthetic modifications, and utilization of purified biosynthetic enzymes.[c,l,m]

How do antibiotics act? Some, like penicillin, block specific enzymes. Peptide antibiotics often form complexes with metal ions (Fig. 8-22) and disrupt the control of ion permeability in bacterial membranes. Polyene antibiotics interfere with proton and ion transport in fungal membranes. Tetracyclines and many other antibiotics interfere directly with protein synthesis (Box 29-B). Others intercalate into DNA molecules (Fig. 5-23; Box 28-A). There is no single mode of action. The search for suitable antibiotics for human use consists in finding compounds highly toxic to infective organisms but with low toxicity to human cells.

Penicillin kills only growing bacteria by preventing proper crosslinking of the peptidoglycan

BOX 20-G (continued)

layer of their cell walls. An amino group from a diamino acid in one peptide chain of the peptidoglycan displaces a D-alanine group in a transpeptidation (acyltransferase) reaction. The transpeptidase is also a hydrolase, a DD-carboxypeptidase. Penicillins are structural analogs of D-alanyl-D-alanine and bind to the active site of the transpeptidase.[l,n–p] The β-lactam ring of penicillins is unstable, making penicillins powerful acylating agents. The transpeptidase apparently acts by a double displacement mechanism, and the initial attack of a nucleophilic serine hydroxyl group of the enzyme on penicillin bound at the active site leads to formation of an inactivated, penicillinoylated enzyme.[q,r] More than one protein in a bacterium is derivatized by penicillin.[s] Therefore, more than one site of action may be involved in the killing of bacterial cells.



β-Lactam
Structure of penicillin
and related antibiotics

D-Alanyl-D-alanine group

Bonds in D-alanyl-D-alanine group of peptidoglycan or in penicillins are cleaved as indicated by dashed lines

Serine —OH group of transpeptidase.

Several classes of **β-lactamases**, often encoded in transmissible plasmids, have spread worldwide rapidly among bacteria, seriously decreasing the effectivenss of penicillins and other β-lactam antibiotics.[t–y] Most β-lactamases (classes A and C) contain an active site serine and are thought to have evolved from the DD transpeptidases, but the B type[y] has a catalytic Zn^{2+}. The latter, as well as a recently discovered type A enzyme,[z] hydrolyze imipenem, currently one of the antibiotics of last resort used to treat infections by penicillin-resistant bacteria. Some β-lactam antibiotics are also powerful inhibitors of β-lactamases.[u,aa,bb] These antibiotics may also have uses in inhibition of serine proteases[cc,dd] such as elastase. Some antibiotic-resistant staphylococci produce an extra penicillin-binding protein that protects them from beta lactams.[ee] Because of antibiotic resistance the isolation of antibiotics from mixed populations of microbes from soil, swamps, and lakes continues. Renewed efforts are being made to find new targets for antibacterial drugs and to synthesize new compounds in what will evidently be a never-ending battle. We also need better antibiotics against fungi and protozoa.

---

[a] Tyndall, J. (1876) *Phil. Trans. Roy. Soc. London B* **166**, 27

[b] Reese, K. M. (1980) *Chem. Eng. News, Sept. 29*, p64

[c] Abraham, E. P. (1981) *Sci. Am.* **244** (Jun), 76–86

[d] Sheehan, J. C. (1982) *The Enchanted Ring: The Untold Story of Penicillin*, MIT Press, Cambridge, Massachusetts

[e] Hobby, G. L. (1985) *Penicillin: Meeting the Challenge*, Yale Univ. Press,

[f] Chain, B. (1991) *Nature (London)* **353**, 492–494

[g] Williams, T. I. (1984) *Howard Florey: Penicillin and After*, Oxford Univ. Press, London

[h] Moberg, C. L. (1991) *Science* **253**, 734–735

[i] Nayler, J. H. C. (1991) *Trends Biochem. Sci.* **16**, 195–197

[j] Crease, R. P. (1989) *Science* **246**, 883–884

[k] Gabay, J. E. (1994) *Science* **264**, 373–374

[l] Nayler, J. H. C. (1991) *Trends Biochem. Sci.* **16**, 234–237

[m] Mourey, L., Miyashita, K., Swarén, P., Bulychev, A., Samama, J.-P., and Mobashery, S. (1998) *J. Am. Chem. Soc.* **120**, 9382–9383

[n] Yocum, R. R., Waxman, D. J., and Strominger, J. L. (1980) *Trends Biochem. Sci.* **5**, 97–101

[o] Kelly, J. A., Moews, P. C., Knox, J. R., Frere, J.-M., and Ghuysen, J.-M. (1982) *Science* **218**, 479–480

[p] Kelly, J. A., and Kuzin, A. P. (1995) *J. Mol. Biol.* **254**, 223–236

[q] Englebert, S., Charlier, P., Fonzè, E., To'th, Y., Vermeire, M., Van Beeumen, J., Grandchamps, J., Hoffmann, K., Leyh-Bouille, M., Nguyen-Distèche, M., and Ghuysen, J.-M. (1994) *J. Mol. Biol.* **241**, 295–297

[r] Kuzin, A. P., Liu, H., Kelly, J. A., and Knox, J. R. (1995) *Biochemistry* **34**, 9532–9540

[s] Thunnissen, M. M. G. M., Fusetti, F., de Boer, B., and Dijkstra, B. W. (1995) *J. Mol. Biol.* **247**, 149–153

[t] Siemers, N. O., Yelton, D. E., Bajorath, J., and Senter, P. D. (1996) *Biochemistry* **35**, 2104–2111

[u] Swarén, P., Massova, I., Bellettini, J. R., Bulychev, A., Maveyraud, L., Kotra, L. P., Miller, M. J., Mobashery, S., and Samama, J.-P. (1999) *J. Am. Chem. Soc.* **121**, 5353–53595

[v] Guillaume, G., Vanhove, M., Lamotte-Brasseur, J., Ledent, P., Jamin, M., Joris, B., and Frère, J.-M. (1997) *J. Biol. Chem.* **272**, 5438–5444

[w] Adediran, S. A., Deraniyagala, S. A., Xu, Y., and Pratt, R. F. (1996) *Biochemistry* **35**, 3604–3613

[x] Brown, R. P. A., Aplin, R. T., and Schofield, C. J. (1996) *Biochemistry* **35**, 12421–12432

[y] Carfi, A., Pares, S., Duée, E., Galleni, M., Duez, C., Frère, J. M., and Dideberg, O. (1995) *EMBO J.* **14**, 4914–4921

[z] Swarén, P., Maveyraud, L., Raquet, X., Cabantous, S., Duez, C., Pédelacq, J.-D., Mariotte-Boyer, S., Mourey, L., Labia, R., Nicolas-Chanoine, M.-H., Nordmann, P., Frère, J.-M., and Samama, J.-P. (1998) *J. Biol. Chem.* **273**, 26714–26721

[aa] Trehan, I., Beadle, B. M., and Shoichet, B. K. (2001) *Biochemistry* **40**, 7992–7999

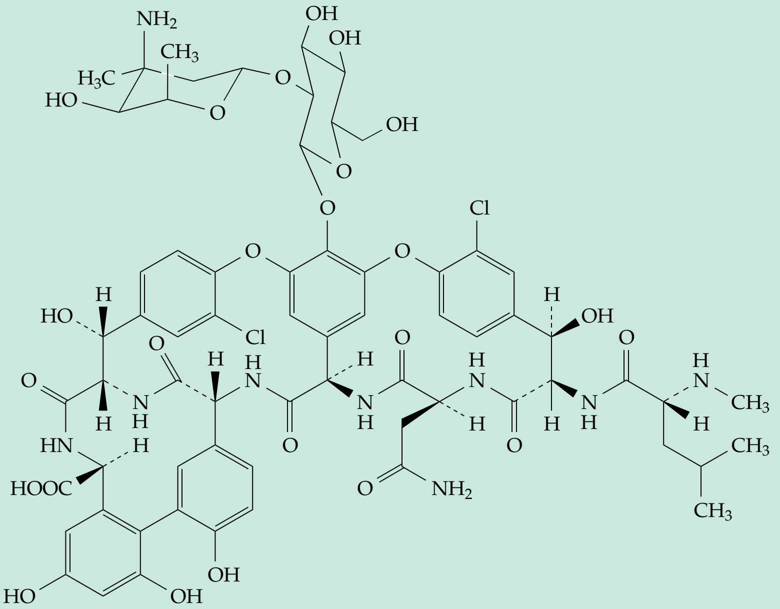
[bb] Swarén, P., Golemi, D., Cabantous, S., Bulychev, A., Maveyraud, L., Mobashery, S., and Samama, J.-P. (1999) *Biochemistry* **38**, 9570–9576

[cc] Wilmouth, R. C., Kassamally, S., Westwood, N. J., Sheppard, R. J., Claridge, T. D. W., Aplin, R. T., Wright, P. A., Pritchard, G. J., and Schofield, C. J. (1999) *Biochemistry* **38**, 7989–7998

[dd] Taylor, P., Anderson, V., Dowden, J., Flitsch, S. L., Turner, N. J., Loughran, K., and Walkinshaw, M. D. (1999) *J. Biol. Chem.* **274**, 24901–24905

[ee] Zhang, H. Z., Hackbarth, C. J., Chansky, K. M., and Chambers, H. F. (2001) *Science* **291**, 1962–1965

BOX 20-H ANTIBIOTIC RESISTANCE AND VANCOMYCIN



Vancomycin

As antibiotics came into widespread use, an unanticipated problem arose in the rapid development of resistance by bacteria. The problem was made acute by the fact that resistance genes are easily transferred from one bacterium to another by the infectious R-factor plasmids.[a–d] Since resistance genes for many different antibiotics may be carried on the same plasmid, "super bacteria," resistant to a large variety of antibiotics, have developed, often in hospitals.

The problem has reached the crisis stage, perhaps most acutely for tuberculosis. Drug-resistant *Mycobacteria tuberculosis* have emerged, especially, in patients being treated for HIV infection (see Box 21-C). Mechanisms of resistance often involve inactivation of the antibiotics. Aminoglycosides such as streptomycin, spectinomycin, and kanamycin (Box 20-B) are inactivated by en-zymes catalyzing phosphorylation or adenylylation of hydroxyl groups on the sugar rings.[e–g] Penicillin and related antibiotics are inactivated by β-lactamases (Box 20-G). Chloramphenicol (Fig. 25-10) is inactivated by acylation on one or both of the hydroxyl groups.

What is the origin of the drug resistance factors? Why do genes for inactivation of such unusual molecules as the antibiotics exist widely in nature? Apparently the precursor to drug resistance genes fulfill normal biosynthetic roles in nature. An antibiotic-containing environment, such as is found naturally in soil, leads to selection of mutants of such genes with drug-inactivating properties. Never-theless, it is not entirely clear why drug resistance factors have appeared so promptly in the population. Overuse of antibiotics in treating minor infections is one apparent cause.[h,i] Another is probably the widespread use on farms.[j] A nationwide effort to decrease the use of erythromycin in Finland had a very favorable effect in decreasing the incidence of erythromycin-resistant group A streptocci.[h]

Because of the rapid development of resistance, continuous efforts are made to alter antibiotics by semisynthesis (see Box 20-G) and to identify new targets for antibiotics or for synthetic antibacterial compounds.[d] An example is provided by the discovery of vancomycin. Like the penicillins, this antibiotic interferes with bacterial cell wall

teichuronic acids (p. 431). Both proteins and neutral polysaccharides, sometimes covalently bound, may also be present.[303] Like peptidoglycans, teichoic and teichuronic acids are assembled on undecaprenyl phosphates[303] or on molecules of diacylglycerol.[322] Either may serve as an anchor. A "linkage unit" may be formed by transfer of several glycosyl rings onto an anchor unit. For example, in synthesis of ribitol teichoic acid sugar rings are transferred from UDP-GlcNAc, UDP-ManNAC, and CDP-Gal to form the following linkage unit:

$$\text{Lipid}-P-P-\text{GlcNAc}-\text{ManNAc}-(P\text{-Gal})_{2,3}$$

Then many ribitol phosphate units are added by transfer from CDP-ribitol. Finally, the chain is capped by transfer of a glucose from UDP-Glc. Lipoteichoic acids often carry covalently linked D-alanine in ester linkage, altering the net electrical charge on the cell surface.[322a] The completed teichoic acid may then be transferred to a peptidoglycan, releasing the lipid phosphate for reuse.[303] Glycerol teichoic acid may be formed in a similar fashion.[322] Teichuronic acids arise by alternate transfers of *P*-GalNAc from UDP-GalNAc and of GlcA from UDP-GlcA.[303]

Gram-positive bacteria often carry surface proteins that interact with host tissues in establishing human infections. Protein A of *Staphylococcus* is a well-known

BOX 20-H (continued)

synthesis but does so by binding tightly to the D-alanyl-D-alanine termini of peptidoglycans that are involved in crosslinking (Fig. 8-29, Fig. 20-9).[k,l] Like penicillin, vancomycin prevents crosslinking but is unaffected by β-lactamases. Initially bacteria seemed unable to develop resistance to vancomycin, and this antibiotic was for 25 years the drug of choice for β-lactam resistant streptococci or staphylococci. However, during this period bacteria carrying a plasmid with nine genes on the transposon Tn1546 (see Fig. 27-30) developed resistance to vancomycin and were spread worldwide.[l] Vancomycin-resistant bacteria are able to sense the presence of the antibiotic and to synthesize an altered **D-alanine:D-alanine ligase**, the enzyme that joins two D-alanine molecules in an ATP-dependent reaction to form the D-alanyl-D-alanine needed to permit peptidoglycan crosslinking (Fig. 20-9, step *c′*). The altered enzyme adds D-lactate rather than D-alanine providing an –OH terminus in place of $-NH_3^+$. This prevents the binding of vancomycin[l,m] and allows crosslinking of the peptidoglycan via depsipeptide bonds. Another gene in the transposon encodes an oxoacid reductase which supplies the D-lactate.[l] A different resistant strain synthesizes a D-Ala-D-Ser ligase.[n] High-level vancomycin resistance is not attained unless the bacteria also synthesize a D-alanyl-D-alanine dipeptidase.[o]

The D-Ala-D-Ala ligase does provide yet another attractive target for drug design.[p,q] Still another is the D-Ala-D-Ala adding enzyme (Fig. 20-8, step *d*; encoded by the *E. coli MurF* gene).[r,s] Strategies for combatting vancomycin resistance include synthesis of new analogs of the antibiotic[t,u] and simultaneous administration of small molecules that catalyze cleavage of the D-Ala-D-lactate bond formed in cell wall precursors of resistant bacteria.[v]

Another possibility is to use **bacteriophages** directly as antibacterial medicines. This approach was introduced as early as 1919 and has enjoyed considerable success. It is now regarded as a promising alternative to the use of antibiotics in many instances.[w]

---

[a] Davies, J. (1994) *Science* **264**, 375–382
[b] Benveniste, R., and Davies, J. (1973) *Ann. Rev. Biochem.* **42**, 471–506
[c] Clowes, R. C. (1973) *Sci. Am.* **228**(Apr), 19–27
[d] Neu, H. C. (1992) *Science* **257**, 1064–1073
[e] McKay, G. A., and Wright, G. D. (1996) *Biochemistry* **35**, 8680–8685
[f] Thompson, P. R., Hughes, D. W., Cianciotto, N. P., and Wright, G. D. (1998) *J. Biol. Chem.* **273**, 14788–14795
[g] Cox, J. R., and Serpersu, E. H. (1997) *Biochemistry* **36**, 2353–2359
[h] Seppälä, H., Klaukka, T., Vuopio-Varkila, J., Muotiala, A., Helenius, H., Lager, K., and Huovinen, P. (1997) *N. Engl. J. Med.* **337**, 441–446
[i] Gorbach, S. L. (2001) *N. Engl. J. Med.* **345**, 1202–1203
[j] Witte, W. (1998) *Science* **279**, 996–997
[k] Sheldrick, G. M., Jones, P. G., Kennard, O., Williams, D. H., and Smith, G. A. (1978) *Nature (London)* **271**, 223–225
[l] Walsh, C. T. (1993) *Science* **261**, 308–309
[m] Sharman, G. J., Try, A. C., Dancer, R. J., Cho, Y. R., Staroske, T., Bardsley, B., Maguire, A. J., Cooper, M. A., O'Brien, D. P., and Williams, D. H. (1997) *J. Am. Chem. Soc.* **119**, 12041–12047
[n] Park, I.-S., Lin, C.-H., and Walsh, C. T. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10040–10044
[o] Aráoz, R., Anhalt, E., René, L., Badet-Denisot, M.-A., Courvalin, P., and Badet, B. (2000) *Biochemistry* **39**, 15971–15979
[p] Fan, C., Moews, P. C., Walsh, C. T., and Knox, J. R. (1994) *Science* **266**, 439–443
[q] Fan, C., Park, I.-S., Walsh, C. T., and Knox, J. R. (1997) *Biochemistry* **36**, 2531–2538
[r] Duncan, K., van Heijenoort, J., and Walsh, C. T. (1990) *Biochemistry* **29**, 2379–2386
[s] Anderson, M. S., Eveland, S. S., Onishi, H. R., and Pompliano, D. L. (1996) *Biochemistry* **35**, 16264–16269
[t] Walsh, C. (1999) *Science* **284**, 442–443
[u] Ge, M., Chen, Z., Onishi, H. R., Kohler, J., Silver, L. L., Kerns, R., Fukuzawa, S., Thompson, C., and Kahne, D. (1999) *Science* **284**, 507–511
[v] Chiosis, G., and Boneca, I. G. (2001) *Science* **293**, 1484–1487
[w] Stone, R. (2002) *Science* **298**, 728–731

---

example. After synthesis in the cytoplasm, it enters the secretory pathway. An N-terminal hydrophobic leader sequence and a 35-residue C-terminal sorting signal guide it to the correct destination. There a free amino group of an unlinked pentaglycyl group of the peptidoglycan carries out a transamidation reaction with an LPXTG sequence in the proteins, cutting the chain between the threonine and glycine residues, and anchoring the protein A to the peptidoglycan.[323]

Group A streptococci, which are serious human pathogens, form α-helical coiled-coil threads whose C termini are anchored in the cell membrane. They protrude through the peptidoglycan layers and provide a hairlike layer around the bacteria. A variable region at the N termini provides many antigens, some of which escape the host's immune system allowing infection to develop.[324] Group B streptococci form carbohydrate antigens linked to teichoic acid.[325] Streptococci, which are normally present in the mouth, utilize their carbohydrate surfaces as receptors for adhesion, allowing them to participate in formation of dental plaque.[326]

Cell walls of mycobacteria are composed of a peptidoglycan with covalently attached galactan chains. Branched chains of **arabinan**, a polymer of the furanose ring form of arabinose with covalently attached **mycolic acids**, are glycosidically linked to the galactan.[327] Shorter **glycopeptidolipids**, containing modified glucose and rhamnose rings as well as fatty acids, contribute to the complexity of mycobacterial surfaces.[328]

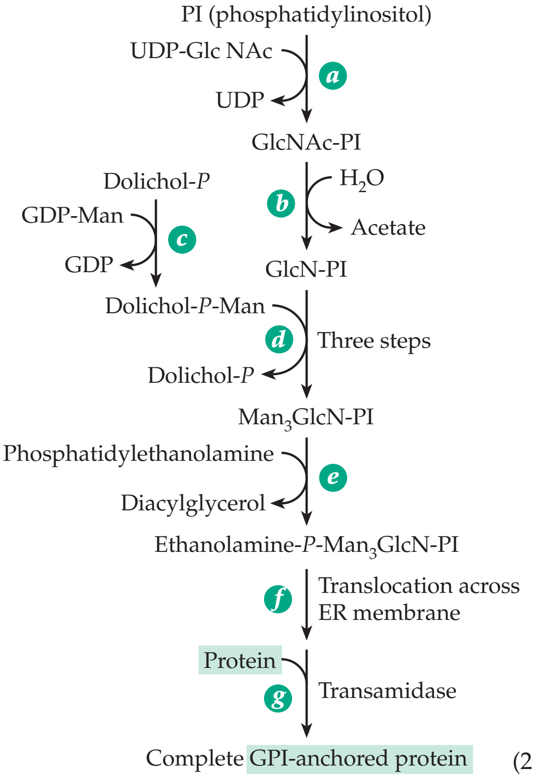
These examples describe only a small sample of the great diversity of cell coats found in the prokaryotic world. Some bacteria also provide themselves with additional protection in the form of external sheaths of crystalline arrays of proteins known as S-layers.[329]

F. Biosynthesis of Eukaryotic Glycolipids

Glycolipids may be thought of as membrane lipids bearing external oligosaccharides. In this sense, they are similar to glycoproteins both in location and in biological significance. Like the glycoproteins, glycolipids are synthesized in the ER, then transported into the Golgi apparatus and eventually outward to join the outer surface of the plasma membrane. Some glycolipids are attached to proteins by covalent linkage.

1. Glycophosphatidylinositol (GPI) Anchors

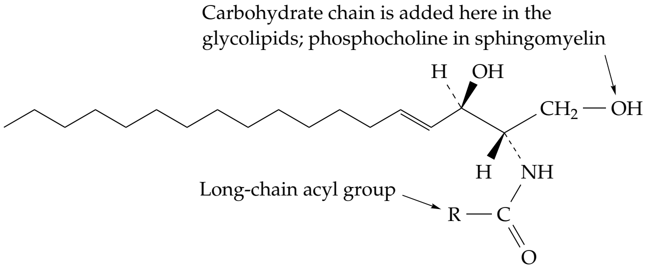
More than 100 different human proteins are attached to phosphatidylinositol anchors of the type shown in Fig. 8-13.[329a] Similar anchors are prevalent in yeast and in protozoa including *Leishmania* and *Trypanosoma*,[330–334b] and *Plasmodium* [334c] where they often bind major surface proteins to the plasma membrane.[335] They are also found in mycobacteria.[336] The structures of the hydrophobic anchor ends are all similar.[337] Two or three fatty acyl groups hold the molecule to the bilayer. Variations are found in the attached glycan portion, both in the number of sugar rings and in the structures of the covalently attached phosphoethanol-amine groups.[337–339] A typical assembly pathway is shown in Eq. 20-23. The first step (step *a* in Eq. 20-23), the transfer of an *N*-acetylglucosamine residue to phosphatidylinositol, is surprisingly complex, requiring at least three proteins.[340] The hydrolytic deacetylation (step *b*) helps to drive the synthetic process. Step *c* provides dolichol-*P*-mannose for the GPI anchors as well as for glycoproteins. The phosphoethanolamine part of the structure is added from phosphatidylethanolamine, apparently via direct nucleophilic displacement.[333] In this way the C terminus of the protein forms an amide linkage with the $-NH_2$ group of ethanolamine in the GPI anchor. Another unexpected finding was that this completed anchor unit undergoes "remodeling" during which the fatty acyl chains of the original phosphatidylinositol are replaced by other fatty acids.[339,339a]



(20-23)

2. Cerebrosides and Gangliosides

These two groups of glycolipids are derived from the *N*-acylated sphingolipids known as **ceramides**. Some biosynthetic pathways from sphingosine to these substances are indicated in Fig. 20-11. Acyl, glycosyl, and sulfo groups are transferred from appropriate derivatives of CoA, CDP, UDP, CMP, and from PAPS to form more than 40 different gangliosides.[341,342] The biosynthesis of a sphingomyelin is also shown in this scheme but is discussed in Chapter 21.

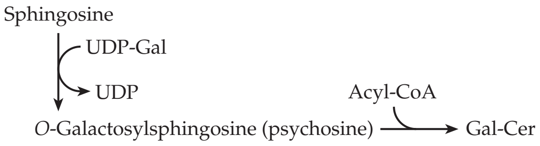


Carbohydrate chain is added here in the glycolipids; phosphocholine in sphingomyelin

Long-chain acyl group

A ceramide

Each biosynthetic step in Fig. 20-11 is catalyzed by a specific transferase. Most of these enzymes are present in membranes of the ER and the Golgi.[343–346] Furthermore, the sequence by which the transferases act may not always be fixed, and a complete biosynthetic scheme would be far more complex than is shown in the figure. For example, one alternative sequence is the synthesis of galactosyl ceramide by transfer of galactose to sphingosine followed by acylation (Eq. 20-24). However, the pathway shown in Fig. 20-11 is probably more important.



$$\text{Sphingosine} \xrightarrow{\text{UDP-Gal} \rightarrow \text{UDP}} O\text{-Galactosylsphingosine (psychosine)} \xrightarrow{\text{Acyl-CoA}} \text{Gal-Cer}$$

(20-24)

G. The Intracellular Breakdown of Polysaccharides and Glycolipids

The attention of biochemists has been drawn to the importance of pathways of degradation of complex polysaccharides through the existence of at least 35 inherited **lysosomal storage diseases**.[347–351] In many of these diseases one of the 40-odd lysosomal hydrolases is defective or absent.

1. Mucopolysaccharidoses

There are at least seven mucopolysaccharidoses (Table 20-1) in which glycosaminoglycans such as hyaluronic acid accumulate to abnormal levels in tissues and may be excreted in the urine. The diseases cause severe skeletal defects; varying degrees of mental retardation; and early death from liver, kidney, or cardiovascular problems. As in other lysosomal diseases, undegraded material is stored in intracellular inclusions lined by a single membrane. Various tissues are affected to different degrees, and the diseases tend to progress with time.

First described in 1919 by Hurler, **mucopolysaccharidosis I** (MPS I, the **Hurler syndrome**) leads to accumulation of partially degraded dermatan and heparan sulfates (Fig. 4-11).[347,352,353] A standard procedure in the study of diseases of this type is to culture fibroblasts from a skin biopsy. Such cells cultured from patients with the Hurler syndrome accumulate the polysaccharide, but when fibroblasts from a normal person are cultured in the same vessel the defect is "corrected." It was shown that a protein secreted by the normal fibroblasts is taken up by the defective fibroblasts, permitting them to complete the degradation of the stored polysaccharide. This "Hurler corrective factor" was identified as an **α-L-iduronidase**. In the **Hunter syndrome** (MPS II) dermatan sulfate and heparan sulfate accumulate. The missing enzyme is a **sulfatase** for 2-sulfated iduronate residues.[354,355] The diagram at the bottom of the page illustrates the need for both of these enzymes as well as three others in the degradation of dermatan.[352,356,357]

The **Sanfilippo disease** type A (MPS III) corrective factor is a heparan *N*-sulfatase. However, as is true for many other metabolic diseases, the same symptoms

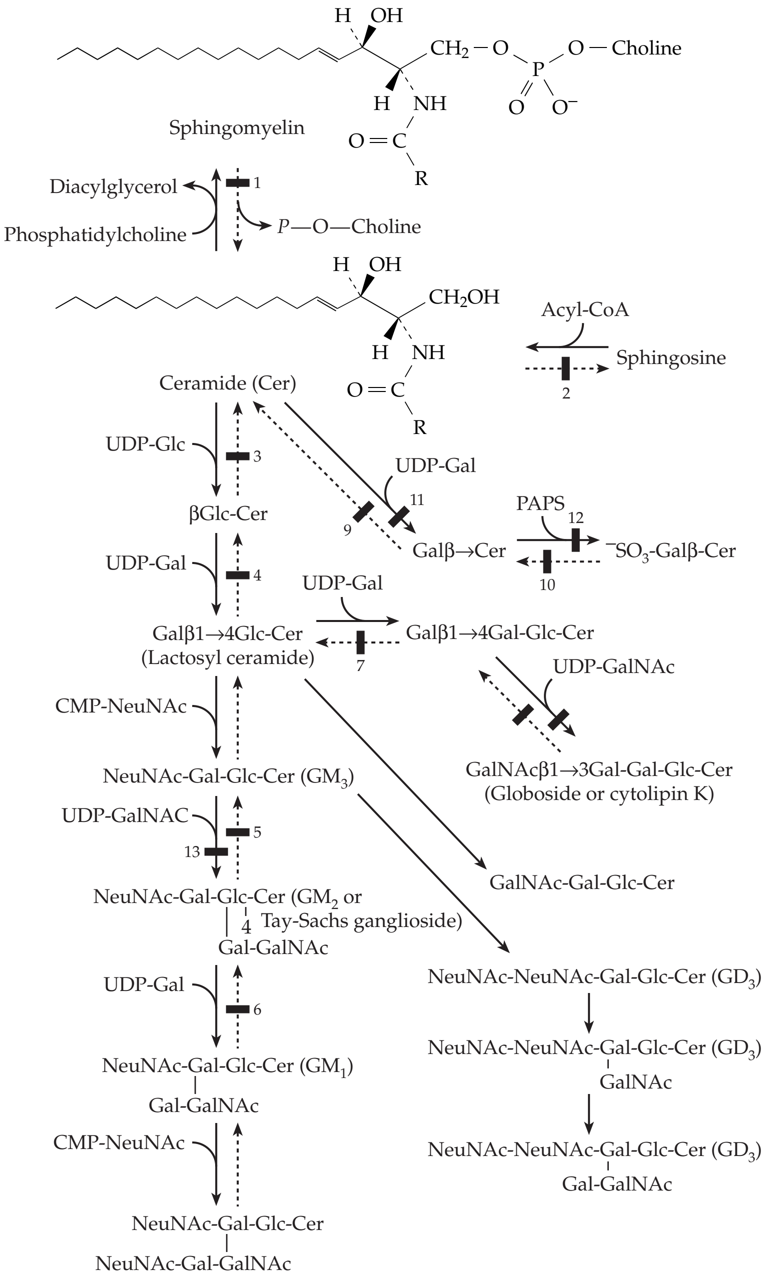

H OH
CH2 – O
P
O – Choline
O O–
H NH
Sphingomyelin
O=C
R
Diacylglycerol
Phosphatidylcholine
1
P—O—Choline
H OH
CH2OH
Acyl-CoA
Sphingosine
2
H NH
O=C
R
Ceramide (Cer)
UDP-Glc
3
βGlc-Cer
UDP-Gal
11
9
PAPS
12
Galβ→Cer
–SO3-Galβ-Cer
10
UDP-Gal
4
UDP-Gal
Galβ1→4Glc-Cer
(Lactosyl ceramide)
7
Galβ1→4Gal-Glc-Cer
UDP-GalNAc
CMP-NeuNAc
GalNAcβ1→3Gal-Gal-Glc-Cer
(Globoside or cytolipin K)
NeuNAc-Gal-Glc-Cer (GM3)
UDP-GalNAC
5
13
GalNAc-Gal-Glc-Cer
NeuNAc-Gal-Glc-Cer (GM2 or
4 Tay-Sachs ganglioside)
Gal-GalNAc
NeuNAc-NeuNAc-Gal-Glc-Cer (GD3)
UDP-Gal
6
NeuNAc-NeuNAc-Gal-Glc-Cer (GD3)
GalNAc
NeuNAc-Gal-Glc-Cer (GM1)
Gal-GalNAc
CMP-NeuNAc
NeuNAc-NeuNAc-Gal-Glc-Cer (GD3)
Gal-GalNAc
NeuNAc-Gal-Glc-Cer
NeuNAc-Gal-GalNAc

Figure 20-11 Biosynthesis and catabolism of glycosphingolipids. The heavy bars indicate metabolic blocks in known diseases.

may arise from several causes. Thus, Sanfilippo diseases B and D arise from lack of an *N*-acetylglucosaminidase and of a sulfatase for GlcNac-6-sulfate, respectively.[354] In Sanfilippo disease C the missing or defective enzyme is an acetyl transferase that transfers an acetyl group from acetyl-CoA onto the amino groups of glucosamine residues in heparan sulfate fragments. All four of these enzymes are needed to degrade the glucosamine-uronic acid pairs of heparan. The *N*-sulfate groups must be removed by the *N*-sulfatase. The free amino groups formed must then be acetylated before the *N*-acetylglucosaminidase can cut off the GlcNac groups. Remov-al of the 6-sulfate groups requires the fourth enzyme. Completion of the degradation also requires both β-glucuronidase and α-L-iduronidase. Another lysosomal enzyme deficiency, which is most prevalent in Finland, is the absence of **aspartylglucosaminidase**, an N-terminal nucleophile hydrolase (Chapter 12, Section C,3) that cleaves glucosamine from aspartate side chains to which oligosaccharides were attached in glycoproteins.[358–360]

Some hereditary diseases are characterized by lack of two or more lysosomal enzymes. In **I-cell disease** (mucolipidosis II), which resembles the Hurler syndrome, at least ten enzymes are absent or are present at much reduced levels.[350,361] The biochemical defect is the absence from the Golgi cisternae of the *N*-acetylglucosaminyl phosphotransferase that transfers *P*-GlcNAc units from UDP-GlcNAc onto mannose residues (Eq. 20-22) of glycoproteins marked for use in lysosomes.

2. Sphingolipidoses

There are at least ten lysosomal storage diseases, known as sphingolipidoses, that involve the metabolism of the glycolipids. Their biochemical bases are indicated in Fig. 20-11 and in Table 20-1. **Gaucher disease**[362–365] is a result of an autosomal recessive trait that permits glucosyl ceramide to accumulate in macrophages. The liver and spleen are seriously damaged, the latter becoming enlarged to four or five times normal size in the adult form of the disease. In the more severe juvenile form mental retardation occurs. By 1965, it was established that cerebroside is synthesized at a normal rate in the individuals affected, but that a lysosomal hydrolase was missing. This blocked the catabolic pathway indicated by dashed arrows in Fig. 20-11 (block No. 3 in the figure). In many patients a single base change causing a Leu → Pro substitution accounts for the defect. In **Fabry disease** an X-linked gene that provides for removal of galactosyl residues from cerebrosides is defective.[350] This leads to accumulation of the triglycosylceramide whose degradation is blocked at point 7 in Fig. 20-11.

The best known and the commonest sphingolipidosis is **Tay–Sachs disease**.[366–368] Several hundred cases have been reported since it was first described in 1881. A terrible disease, it is accompanied by mental deterioration, blindness, paralysis, dementia, and death by the age of three. About 15 children a year are born in North America with this condition, and the world figure must be 5–7 times this. The defect is in the α subunit of the β-hexosaminidase A (point 7 in Fig. 20-10)[366,366a] with accumulation of ganglioside G_{M2}. Somewhat less severe forms of the disease are caused by different mutations in the same gene[369] or in a protein activator. **Sandhoff disease**, which resembles Tay–Sachs disease, is caused by a defect in the β subunit, which is present in both β-hexosaminidases A and B.[368] Mutant "knockout" mice that produce only ganglioside GM_3 as the major ganglioside in their central nervous system die suddenly from seizures if they hear a loud sound. This provides further evidence of the essential nature of these components of nerve membranes.[369a]

3. Causes of Lysosomal Diseases

The descriptions given here have been simplified. For many lysosomal diseases there are mild and severe forms and infantile or juvenile forms to be contrasted with adult forms. Some of the enzymes exist as multiple isozymes. An enzyme may be completely lacking or

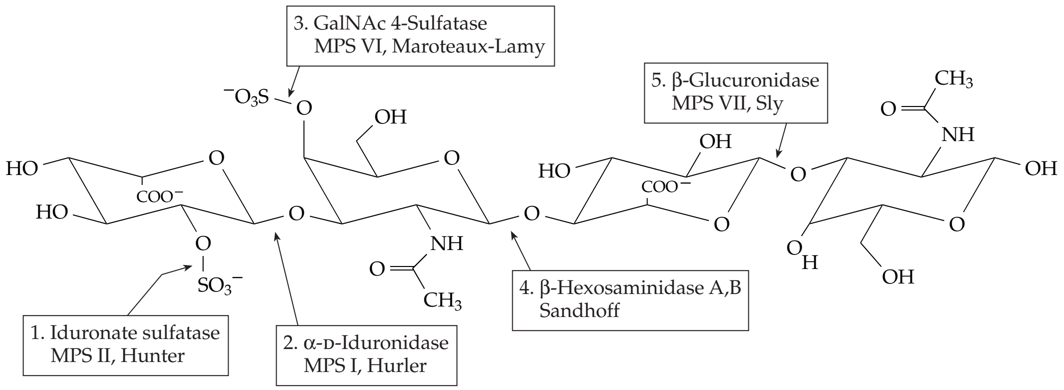


TABLE 20-1
Lysosomal Storage Diseases: Sphingolipidoses and Mucopolysaccharidoses[a]

No. in Fig. 20-11	Name	Defective enzyme
1.	Niemann–Pick disease[b]	Sphingomyelinase
2.	Farber disease (lipogranulomatosis)	Ceramidase
3.	Gaucher disease[c]	β-Glucocerebrosidase
4.	Lactosyl ceramidosis	β-Galactosyl hydrolase
5.	Tay–Sachs disease[c]	β-Hexosaminidase A
6.	G_{M1} gangliosidosis[d]	β-Galactosidase
7.	Fabry disease[c]	α-Galactosidase
8.	Sandhoff disease[e]	β-Hexosaminidases A and B
9.	Globoid cell leukodystrophy	Galactocerebrosidase
10.	Metachromatic leukodystrophy	Arylsulfatase A
13.	Hematoside (G_{M3}) accumulation	G_{M3}-*N*-acetylgalactosaminyltransferase
	Pompe disease[f]	α-Glucosidase
	Hurler syndrome (MPS I)[c]	α-L-Iduronidase
	Hunter syndrome (MPS II)[c]	Iduronate 2-sulfate sulfatase
	Sanfilippo disease[c,g]	
	Type A (MPS III)	Heparan *N*-sulfatase
	Type B	*N*-Acetylglucosaminidase
	Type C	Acetyl-CoA: α-glucosaminide *N*-acetyltransferase
	Type D	GlcNAc-6-sulfate sulfatase
	Maroteaux–Lamy syndrome (MPS VI)[g]	Arylsulfatase B
	Sly syndrome (MPS VII)[g,h]	β-Glucuronidase
	Aspartylglycosaminuria[i]	Aspartylglucosaminidase
	Mannosidosis	β-Mannosidase
	Fucosidosis	α-L-Fucosidase
	Mucolipidosis	α-*N*-Acetylneuraminidase
	Sialidosis[j]	

[a] A general reference is Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds. (1995) *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1, McGraw-Hill, New York (pp. 2427–2879)

[b] Wenger, D. A., Sattler, M., Kudoh, T., Snyder, S. P., and Kingston, R. S. (1980) *Science* **208**, 1471–1473

[c] See main text

[d] Hoogeveen, A. T., Reuser, A. J. J., Kroos, M., and Galjaard, H. (1986) *J. Biol. Chem.* **261**, 5702–5704

[e] Gravel, R. A., Clarke, J. T. R., Kaback, M. M., Mahuran, D., Sandhoff, K., and Suzuki, K. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2839–2879, McGraw-Hill, New York

[f] See Box 20-D

[g] Neufeld, E. F., and Muenzer, J. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2465 – 2494, McGraw-Hill, New York

[h] Wu, B. M., Tomatsu, S., Fukuda, S., Sukegawa, K., Orii, T., and Sly, W. S. (1994) *J. Biol. Chem.* **269**, 23681–23688

[i] Mononen, I., Fisher, K. J., Kaartinen, V., and Aronson, N. N., Jr. (1993) *FASEB J.* **7**, 1247–1256

[j] Seppala, R., Tietze, F., Krasnewich, D., Weiss, P., Ashwell, G., Barsh, G., Thomas, G. H., Packman, S., and Gahl, W. A. (1991) *J. Biol. Chem.* **266**, 7456–7461

may be low in concentration. The causes of the deficiencies may include total absence of the gene, absence of the appropriate mRNA, impaired conversion of a proenzyme to active enzyme, rapid degradation of a precursor or of the enzyme itself, incorrect transport of the enzyme precursor to its proper destination, presence of mutations that inactivate the enzyme, or absence of protective proteins. Several lysosomal hydrolases require auxiliary **activator proteins** that allow them to react with membrane-bound substrates.[350,351,365,370]

4. Can Lysosomal Diseases Be Treated?

There has been some success in using enzyme replacement therapy for lysosomal deficiency diseases.[347,371-371b] One approach makes use of the fact that the mannose-6-*P* receptors of the plasma membrane take up suitably marked proteins and transfer them into lysosomes (Section C,2). The missing enzyme might simply be injected into the patient's bloodstream from which it could be taken up into the lysosomes.[372] This carries a risk of allergic reaction, and it may be safer to attempt microencapsulation of the enzyme, perhaps in ghosts from the patient's own erythrocytes.[347] A second approach, which has had limited success, is transplantation of an organ[371] or of bone marrow[373] from a donor with a normal gene for the missing enzyme. This is dangerous and is little used at present. However, new hope is offered by the possibility of transferring a gene for the missing enzyme into some of the patient's cells. For example, the cloned gene for the transferase missing in Gaucher disease has been transferred into cultured cells from Gaucher disease patients with apparent correction of the defect.[374] Long-term correction of the Hurler syndrome in bone marrow cells also provides hope for an effective therapy involving gene transfer into a patient's own bone marrow cells[375] or transplantation of selected hematopoietic cells.[375a]

In the cases of Gaucher disease and Fabry disease, it is hoped that treatment of infants and young children may prevent brain damage. However, in Tay–Sachs disease the primary sites of accumulation of the ganglioside GM_2 are the ganglion and glial cells of the brain. Because of the "blood–brain barrier" and the severity of the damage it seems less likely that the disease can be treated successfully.

The approach presently used most often consists of identifying carriers of highly undesirable genetic traits and offering genetic counseling. For example, if both parents are carriers the risk of bearing a child with Tay–Sachs disease is one in four. Women who have borne a previous child with the disease usually have the genetic status of the fetus checked by **amniocentesis**. A sample of the amniotic fluid surrounding the fetus is withdrawn during the 16th to 18th week of pregnancy. The fluid contains fibroblasts that have become detached from the surface of the fetus. These cells are cultured for 2–3 weeks to provide enough cells for a reliable assay of the appropriate enzymes. Such tests for a variety of defects are becoming faster and more sensitive as new techniques are applied.[376] In the case of Tay–Sachs disease, most women who have one child with the disease choose abortion if a subsequent fetus has the disease.

The diseases considered here affect only a small fraction of the problems in the catabolism of body constituents. On the other hand, fewer cases are on record of deficiencies in biosynthetic pathways. These are more often absolutely lethal and lead to early spontaneous abortion. However, blockages in the biosynthesis of cerebrosides are known in the special strains of mice known as Jimpy, Quaking, and msd (myelin synthesis deficient).[377,378] The transferases (points 11 and 12 of Fig. 20-11) are not absent but are of low activity. The mice have distinct neurological defects and poor myelination of nerves in the brain. A human ailment involving impaired conversion of GM_3 to GM_2 (with accumulation of the former; point 13 of Fig. 20-11) has been reported. Excessive synthesis of sialic acid causes the rare human **sialuria.**[379] This is apparently a result of a failure in proper feedback inhibition.

Animals suffer many of the same metabolic diseases as humans. Among these are a large number of lysosomal deficiency diseases.[380] Their availability means that new methods of treating the diseases may, in many cases, be tried first on animals. For example, enzyme replacement therapy for the Hurler syndrome is being tested in dogs.[381] Bone marrow transplantation for human **α-mannosidosis** is being tested in cats with a similar disease.[382] Mice with a hereditary deficiency of β-glucuronidase are being treated by gene transfer from normal humans.[357]

References

1. Cardini, C. E., Paladini, A. C., Caputto, R., and Leloir, L. F. (1950) *Nature (London)* **165**, 191–192
2. Grisolía, S. (1988) *Nature (London)* **331**, 212
3. Hallfrisch, J. (1990) *FASEB J.* **4**, 2652–2660
4. Jefferey, J., and Jornvall, H. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 901–905
5. Cox, T. M. (1994) *FASEB J.* **8**, 62–71
6. Gitzelmann, R., Steinmann, B., and Van den Berghe, G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 905–935, McGraw-Hill, New York
7. Gopher, A., Vaisman, N., Mandel, H., and Lapidot, A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5449–5453
8. Hope, J. N., Bell, A. W., Hermodson, M. A., and Groarke, J. M. (1986) *J. Biol. Chem.* **261**, 7663–7668
9. Ridley, W. P., Houchins, J. P., and Kirkwood, S. (1975) *J. Biol. Chem.* **250**, 8761–8767

9a. Campbell, R. E., Mosimann, S. C., van de Rijn, I., Tanner, M. E., and Strynadka, N. C. J. (2000) *Biochemistry* **39**, 7012–7023

10. Li, J.-p, Hagner-McWhirter, Å., Kjellén, L., Palgi, J., Jalkanen, M., and Lindahl, U. (1997) *J. Biol. Chem.* **272**, 28158–28163
11. Segal, S., and Berry, G. T. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 967–1000, McGraw-Hill, New York
12. Ruzicka, F. J., Wedekind, J. E., Kim, J., Rayment, I., and Frey, P. A. (1995) *Biochemistry* **34**, 5610–5617
13. Wedekind, J. E., Frey, P. A., and Rayment, I. (1996) *Biochemistry* **35**, 11560–11569
14. Thoden, J. B., Ruzicka, F. J., Frey, P. A., Rayment, I., and Holden, H. M. (1997) *Biochemistry* **36**, 1212–1222
15. Frey, P. A. (1996) *FASEB J.* **10**, 461–470

15a. Lai, K., Willis, A. C., and Elsas, L. J. (1999) *J. Biol. Chem.* **274**, 6559–6566

16. Thoden, J. B., and Holden, H. M. (1998) *Biochemistry* **37**, 11469–11477

16a. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., and Holden, H. M. (2001) *J. Biol. Chem.* **276**, 15131–15136

16b. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., and Holden, H. M. (2001) *J. Biol. Chem.* **276**, 20617–20623

17. Beebe, J. A., and Frey, P. A. (1998) *Biochemistry* **37**, 14989–14997

17a. Zhang, Q., and Liu, H.-w. (2001) *J. Am. Chem. Soc.* **123**, 6756–6766

18. Wong, Y.-H. H., and Sherman, W. R. (1985) *J. Biol. Chem.* **260**, 11083–11090
19. Migaud, M. E., and Frost, J. W. (1996) *J. Am. Chem. Soc.* **118**, 495–501

19a. Tian, F., Migaud, M. E., and Frost, J. W. (1999) *J. Am. Chem. Soc.* **121**, 5795–5796

19b. Chen, L., Zhou, C., Yang, H., and Roberts, M. F. (2000) *Biochemistry* **39**, 12415–12423

20. Loewus, M. W., Loewus, F. A., Brillinger, G. U., Otsuka, H., and Floss, H. G. (1980) *J. Biol. Chem.* **255**, 11710–11712
21. Lapan, E. A. (1975) *Exp. Cell. Res.* **94**, 277–282
22. Holub, B. J. (1992) *N. Engl. J. Med.* **326**, 1285–1287

22a. Nestler, J. E., Jakubowicz, D. J., Reamer, P., Gunn, R. D., and Allan, G. (1999) *N. Engl. J. Med.* **340**, 1314–1320

23. Reddy, C. C., Swan, J. S., and Hamilton, G. A. (1981) *J. Biol. Chem.* **256**, 8510–8518
24. Nishikimi, M., and Yagi, K. (1996) in *Subcellular Biochemistry*, Vol. 25 (Harris, J. R., ed), pp. 17–39, Plenum, New York
25. Koshizaka, T., Nishikimi, M., Ozawa, T., and Yagi, K. (1988) *J. Biol. Chem.* **263**, 1619–1621
26. Nishikimi, M., Fukuyama, R., Minoshima, S., Shimizu, N., and Yagi, K. (1994) *J. Biol. Chem.* **269**, 13685–13688

26a. Maeda, N., Hagihara, H., Nakata, Y., Hiller, S., Wilder, J., and Reddick, R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 841–846

26b. Tsukaguchi, H., Tokui, T., Mackenzie, B., Berger, U. V., Chen, X.-Z., Wang, Y., Brubaker, R. F., and Hediger, M. A. (1999) *Nature (London)* **399**, 70–75

27. Miller, J. V., Estell, D. A., and Lazarus, R. A. (1987) *J. Biol. Chem.* **262**, 9016–9020
28. Wheeler, G. L., Jones, M. A., and Smirnoff, N. (1998) *Nature (London)* **393**, 365–369
29. Conklin, P. L., Norris, S. R., Wheeler, G. L., Williams, E. H., Smirnoff, N., and Last, R. L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4198–4203
30. Hiatt, H. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1001–1010, McGraw-Hill, New York
31. Massière, F., Badet-Denisot, M.-A., René, L., and Badet, B. (1997) *J. Am. Chem. Soc.* **119**, 5748–5749
32. Bearne, S. L., and Wolfenden, R. (1995) *Biochemistry* **34**, 11515–11520
33. Badet, B., Vermoote, P., Haumont, P. Y., Lederer, F., and Le Goffic, F. (1987) *Biochemistry* **26**, 1940–1948
34. Golinelli-Pimpaneau, B., Le Goffic, F., and Badet, B. (1989) *J. Am. Chem. Soc.* **111**, 3029–3034

34a. Bearne, S. L., and Blouin, C. (2000) *J. Biol. Chem.* **275**, 135–140

35. Leriche, C., Badet-Denisot, M.-A., and Badet, B. (1996) *J. Am. Chem. Soc.* **118**, 1797–1798
36. Mio, T., Yamada-Okabe, T., Arisawa, M., and Yamada-Okabe, H. (1999) *J. Biol. Chem.* **274**, 424–429
37. Mengin-Lecreulx, D., and van Heijenoort, J. (1996) *J. Biol. Chem.* **271**, 32–39

37a. Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F., and Bourne, Y. (2001) *J. Biol. Chem.* **276**, 11844–11851

37b. Olsen, L. R., and Roderick, S. L. (2001) *Biochemistry* **40**, 1913–1921

38. Montero-Morán, G. M., Horjales, E., Calcagno, M. L., and Altamirano, M. M. (1998) *Biochemistry* **37**, 7844–7849
39. Wolosker, H., Kline, D., Bian, Y., Blackshaw, S., Cameron, A. M., Fralich, T. J., Schnaar, R. L., and Snyder, S. H. (1998) *FASEB J.* **12**, 91–99
40. Ramilo, C., Appleyard, R. J., Wanke, C., Krekel, F., Amrhein, N., and Evans, J. N. S. (1994) *Biochemistry* **33**, 15071–15079
41. Kim, D. H., Lees, W. J., Kempsell, K. E., Lane, W. S., Duncan, K., and Walsh, C. T. (1996) *Biochemistry* **35**, 4923–4928

41a. Krekel, F., Oecking, C., Amrhein, N., and Macheroux, P. (1999) *Biochemistry* **38**, 8864–8878

41b. Krekel, F., Samland, A. K., Macheroux, P., Amrhein, N., and Evans, J. N. S. (2000) *Biochemistry* **39**, 12671–12677

41c. Samland, A. K., Etezady-Esfarjani, T., Amrhein, N., and Macheroux, P. (2001) *Biochemistry* **40**, 1550–1559

42. Benson, T. E., Walsh, C. T., and Hogle, J. M. (1997) *Biochemistry* **36**, 806–811
43. Benson, T. E., Walsh, C. T., and Massey, V. (1997) *Biochemistry* **36**, 796–805

43a. Benson, T. E., Harris, M. S., Choi, G. H., Cialdella, J. I., Herberg, J. T., Martin, J. P., Jr., and Baldwin, E. T. (2001) *Biochemistry* **40**, 2340–2350

44. Pastuszak, I., O'Donnell, J., and Elbein, A. D. (1996) *J. Biol. Chem.* **271**, 23653–23656

44a. Zhao, X., Creuzenet, C., Bélanger, M., Egbosimba, E., Li, J., and Lam, J. S. (2000) *J. Biol. Chem.* **275**, 33252–33259

45. Schauer, R. (1985) *Trends Biochem. Sci.* **10**, 357–360
46. Simon, E. S., Bednarski, M. D., and Whitesides, G. M. (1988) *J. Am. Chem. Soc.* **110**, 7159–7163
47. Morgan, P. M., Sala, R. F., and Tanner, M. E. (1997) *J. Am. Chem. Soc.* **119**, 10269–10277

47a. Campbell, R. E., Mosimann, S. C., Tanner, M. E., and Strynadka, N. C. J. (2000) *Biochemistry* **39**, 14993–15001

47b. Keppler, O. T., Hinderlich, S., Langner, J., Schwartz-Albiez, R., Reutter, W., and Pawlita, M. (1999) *Science* **284**, 1372–1376

47c. Itoh, T., Mikami, B., Maru, I., Ohta, Y., Hashimoto, W., and Murata, K. (2000) *J. Mol. Biol.* **303**, 733–744

47d. Jacobs, C. L., Goon, S., Yarema, K. J., Hinderlich, S., Hang, H. C., Chai, D. H., and Bertozzi, C. R. (2001) *Biochemistry* **40**, 12864–12874

48. Stäsche, R., Hinderlich, S., Weise, C., Effertz, K., Lucka, L., Moormann, P., and Reutter, W. (1997) *J. Biol. Chem.* **272**, 24319–24324

48a. Jordan, P. A., Bohle, D. S., Ramilo, C. A., and Evans, J. N. S. (2001) *Biochemistry* **40**, 8387–8396

49. Dotson, G. D., Dua, R. K., Clemens, J. C., Wooten, E. W., and Woodard, R. W. (1995) *J. Biol. Chem.* **270**, 13698–13705
50. Sheflyan, G. Y., Howe, D. L., Wilson, T. L., and Woodard, R. W. (1998) *J. Am. Chem. Soc.* **120**, 11028–11032
51. Kaustov, L., Kababya, S., Du, S., Baasov, T., Gropper, S., Shoham, Y., and Schmidt, A. (2000) *Biochemistry* **39**, 14865–14876
52. Radaev, S., Dastidar, P., Patel, M., Woodard, R. W., and Gatti, D. L. (2000) *J. Biol. Chem.* **275**, 9476–9484

52a. Duewel, H. S., Radaev, S., Wang, J., Woodard, R. W., and Gatti, D. L. (2001) *J. Biol. Chem.* **276**, 8393–8402

52b. Wagner, T., Kretsinger, R. H., Bauerle, R., and Tolbert, W. D. (2000) *J. Mol. Biol.* **301**, 233–238

52c. Mosimann, S. C., Gilbert, M., Dombroswki, D., To, R., Wakarchuk, W., and Strynadka, N. C. J. (2001) *J. Biol. Chem.* **276**, 8190–8196

53. Kawano, T., Koyama, S., Takematsu, H., Kozutsumi, Y., Kawasaki, H., Kawashima, S., Kawasaki, T., and Suzuki, A. (1995) *J. Biol. Chem.* **270**, 16458–16463
54. Nishino, S., Kuroyanagi, H., Terada, T., Inoue, S., Inoue, Y., Troy, F. A., and Kitajima, K. (1996) *J. Biol. Chem.* **271**, 2909–2913
55. Inoue, S., Kitajima, K., and Inoue, Y. (1996) *J. Biol. Chem.* **271**, 24341–24344

55a. Angata, T., Nakata, D., Matsuda, T., Kitajima, K., and Troy, F. A., II. (1999) *J. Biol. Chem.* **274**, 22949–22956

55b. Lawrence, S. M., Huddleston, K. A., Pitts, L. R., Nguyen, N., Lee, Y. C., Vann, W. F., Coleman, T. A., and Betenbaugh, M. J. (2000) *J. Biol. Chem.* **275**, 17869–17877

56. Kohlbrenner, W. E., Nuss, M. M., and Fesik, S. W. (1987) *J. Biol. Chem.* **262**, 4534–4537

56a. Royo, J., Gómez, E., and Hueros, G. (2000) *J. Biol. Chem.* **275**, 24993–24999

56b. Jelakovic, S., and Schulz, G. E. (2001) *J. Mol. Biol.* **312**, 143–155

57. Kontrohr, T., and Kocsis, B. (1981) *J. Biol. Chem.* **256**, 7715–7718

References

58. Ding, L., Seto, B. L., Ahmed, S. A., and Coleman, W. G., Jr. (1994) *J. Biol. Chem.* **269**, 24384 – 24390
58a. Ni, Y., McPhie, P., Deacon, A., Ealick, S., and Coleman, W. G., Jr. (2001) *J. Biol. Chem.* **276**, 27329 – 27334
58b. Kneidinger, B., Graninger, M., Puchberger, M., Kosma, P., and Messner, P. (2001) *J. Biol. Chem.* **276**, 20935 – 20944
59. Sharon, N. (1975) *Complex Carbohydrates*, Addison-Wesley, Reading, Massachusetts (pp. 131 – 138)
59a. Blankenfeldt, W., Asuncion, M., Lam, J. S., and Naismith, J. H. (2000) *EMBO J.* **19**, 6652 – 6663
59b. Hegeman, A. D., Gross, J. W., and Frey, P. A. (2001) *Biochemistry* **40**, 6598 – 6610
59c. Gross, J. W., Hegeman, A. D., Gerratana, B., and Frey, P. A. (2001) *Biochemistry* **40**, 12497 – 12504
59d. Allard, S. T. M., Giraud, M.-F., Whitfield, C., Graninger, M., Messner, P., and Naismith, J. H. (2001) *J. Mol. Biol.* **307**, 283 – 295
59e. Christendat, D., Saridakis, V., Dharamsi, A., Bochkarev, A., Pai, E. F., Arrowsmith, C. H., and Edwards, A. M. (2000) *J. Biol. Chem.* **275**, 24608 – 24612
59f. Kneidinger, B., Graninger, M., Adam, G., Puchberger, M., Kosma, P., Zayni, S., and Messner, P. (2001) *J. Biol. Chem.* **276**, 5577 – 5583
60. Chan, J. Y., Nwokoro, N. A., and Schachter, H. (1979) *J. Biol. Chem.* **251**, 7060 – 7068
60a. He, X. M., and Liu, H.-w. (2002) *Ann. Rev. Biochem.* **71**, 701 – 754
61. Bonin, C. P., Potter, I., Vanzin, G. F., and Reiter, W.-D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2085 – 2090
61a. Menon, S., Stahl, M., Kumar, R., Xu, G.-Y., and Sullivan, F. (1999) *J. Biol. Chem.* **274**, 26743 – 26750
61b. Rosano, C., Bisso, A., Izzo, G., Tonetti, M., Sturla, L., De Flora, A., and Bolognesi, M. (2000) *J. Mol. Biol.* **303**, 77 – 91
62. Rubenstein, P. A., and Strominger, J. L. (1974) *J. Biol. Chem.* **249**, 3776 – 3781
63. Pieper, P. A., Guo, Z., and Liu, H.-w. (1995) *J. Am. Chem. Soc.* **117**, 5158 – 5159
64. Johnson, D. A., Gassner, G. T., Bandarian, V., Ruzicka, F. J., Ballou, D. P., Reed, G. H., and Liu, H.-w. (1996) *Biochemistry* **35**, 15846 – 15856
65. Chen, X. M. H., Ploux, O., and Liu, H.-w. (1996) *Biochemistry* **35**, 16412 – 16420
65a. Hallis, T. M., Zhao, Z., and Liu, H.-w. (2000) *J. Am. Chem. Soc.* **122**, 10493 – 10503
66. Chang, C.-W. T., Chen, X. H., and Liu, H.-w. (1998) *J. Am. Chem. Soc.* **120**, 9698 – 9699
67. Zhao, L., Que, N. L. S., Xue, Y., Sherman, D. H., and Liu, H.-w. (1998) *J. Am. Chem. Soc.* **120**, 12159 – 12160
68. Toshima, K., Nozaki, Y., Mukaiyama, S., Tamai, T., Nakata, M., Tatsuta, K., and Kinoshita, M. (1995) *J. Am. Chem. Soc.* **117**, 3717 – 3727
69. Benning, C., Beatty, J. T., Prince, R. C., and Somerville, C. R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1561 – 1565
70. Benson, A. A. (1963) *Adv. Lipid Res.* **1**, 387 – 394
70a. Mulichak, A. M., Theisen, M. J., Essigmann, B., Benning, C., and Garavito, R. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13097 – 13102
70b. Sanda, S., Leustek, T., Theisen, M. J., Garavito, R. M., and Benning, C. (2001) *J. Biol. Chem.* **276**, 3941 – 3946
71. Preiss, J. (1984) *Trends Biochem. Sci.* **9**, 24 – 27
71a. Lunn, J. E., Ashton, A. R., Hatch, M. D., and Heldt, H. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12914 – 12919
72. Singh, A. N., Hester, L. S., and Raushel, F. M. (1987) *J. Biol. Chem.* **262**, 2554 – 2557
73. Hunziker, W., Spiess, M., Semenza, G., and Lodish, H. F. (1986) *Cell* **46**, 227 – 234
74. Naim, H. Y., Sterchi, E. E., and Lentze, M. J. (1988) *J. Biol. Chem.* **263**, 19709 – 19717
75. Wolschek, M. F., and Kubicek, C. P. (1997) *J. Biol. Chem.* **272**, 2729 – 2735
76. Horlacher, R., and Boos, W. (1997) *J. Biol. Chem.* **272**, 13026 – 23032
76a. Hars, U., Horlacher, R., Boos, W., Welte, W., and Diederichs, K. (1998) *Protein Sci.* **7**, 2511 – 2521
77. Kandror, O., DeLeon, A., and Goldberg, A. L. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9727 – 9732
77a. Diederichs, K., Diez, J., Greller, G., Müller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W., and Welte, W. (2000) *EMBO J.* **19**, 5951 – 961
77b. Diez, J., Diederichs, K., Greller, G., Horlacher, R., Boos, W., and Welte, W. (2001) *J. Mol. Biol.* **305**, 905 – 915
78. Shaper, N. L., Hollis, G. F., Douglas, J. G., Kirsch, I. R., and Shaper, J. H. (1988) *J. Biol. Chem.* **263**, 10420 – 10428
79. Rajput, B., Shaper, N. L., and Shaper, J. H. (1996) *J. Biol. Chem.* **271**, 5131 – 5142
80. Yadav, S. P., and Brew, K. (1991) *J. Biol. Chem.* **266**, 698 – 703
80a. Gastinel, L. N., Cambillau, C., and Bourne, Y. (1999) *EMBO J.* **18**, 3546 – 3557
80b. Ramakrishnan, B., Shah, P. S., and Qasba, P. K. (2001) *J. Biol. Chem.* **276**, 37665 – 37671
81. Montgomery, R. K., Büller, H. A., Rings, E. H. H. M., and Grand, R. J. (1991) *FASEB J.* **5**, 2824 – 2832
81a. Potera, C. (1998) *Science* **281**, 1793
81b. Barbier, O., Girard, C., Breton, R., Bélanger, A., and Hum, D. W. (2000) *Biochemistry* **39**, 11540 – 11552
81c. Lévesque, E., Turgeon, D., Carrier, J.-S., Montminy, V., Beaulieu, M., and Bélanger, A. (2001) *Biochemistry* **40**, 3869 – 3881
82. Lewis, D. A., and Armstrong, R. N. (1983) *Biochemistry* **22**, 6297 – 6303
83. Hanessian, S., and Haskell, T. H. (1970) in *The Carbohydrates*, 2nd ed., Vol. 2A (Pigman, W., and Horton, D., eds), pp. 139 – 211, Academic Press, New York
84. Snell, J. F. (1966) *Biosynthesis of Antibiotics*, Vol. 1, Academic Press, New York
85. Botti, M. G., Taylor, M. G., and Botting, N. P. (1995) *J. Biol. Chem.* **270**, 20530 – 20535
86. Cottaz, S., Henrissat, B., and Driguez, H. (1996) *Biochemistry* **35**, 15256 – 15259
86a. Burmeister, W. P., Cottaz, S., Rollin, P., Vasella, A., and Henrissat, B. (2000) *J. Biol. Chem.* **275**, 39385 – 39393
87. Lomako, J., Lomako, W. M., and Whelan, W. J. (1988) *FASEB J.* **2**, 3097 – 3103
87a. Cid, E., Gomis, R. R., Geremia, R. A., Guinovart, J. J., and Ferrer, J. C. (2000) *J. Biol. Chem.* **275**, 33614 – 33621
88. Baecker, P. A., Greenberg, E., and Preiss, J. (1986) *J. Biol. Chem.* **261**, 8738 – 8743
89. Hill, M. A., Kaufmann, K., Otero, J., and Preiss, J. (1991) *J. Biol. Chem.* **266**, 12455 – 12460
90. Furukawa, K., Tagaya, M., Tanizawa, K., and Fukui, T. (1994) *J. Biol. Chem.* **269**, 868 – 871
91. Ochoa, S. (1985) *Trends Biochem. Sci.* **10**, 147 – 150
92. Kim, S. C., Singh, A. N., and Raushel, F. M. (1988) *J. Biol. Chem.* **263**, 10151 – 10154
93. Skurat, A. V., Wang, Y., and Roach, P. J. (1994) *J. Biol. Chem.* **269**, 25534 – 25542
94. Shulman, R. G., Bloch, G., and Rothman, D. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8535 – 8542
95. Villar-Palasi, C., and Guinovart, J. J. (1997) *FASEB J.* **11**, 544 – 558
96. Printen, J. A., Brady, M. J., and Saltiel, A. R. (1997) *Science* **275**, 1475 – 1478
96a. Halse, R., Rochford, J. J., McCormack, J. G., Vandenheede, J. R., Hemmings, B. A., and Yeaman, S. J. (1999) *J. Biol. Chem.* **274**, 776 – 780
96b. Oikonomakos, N. G., Schnier, J. B., Zographos, S. E., Skamnaki, V. T., Tsitsanou, K. E., and Johnson, L. N. (2000) *J. Biol. Chem.* **275**, 34566 – 34573
97. Thon, V. J., Khalil, M., and Cannon, J. F. (1993) *J. Biol. Chem.* **268**, 7509 – 7513
98. Takrama, J., and Madsen, N. B. (1988) *Biochemistry* **27**, 3308 – 3314
99. Yang, B.-Z., Ding, J.-H., Enghild, J. J., Bao, Y., and Chen, Y.-T. (1992) *J. Biol. Chem.* **267**, 9294 – 9299
99a. Nakayama, A., Yamamoto, K., and Tabata, S. (2001) *J. Biol. Chem.* **276**, 28824 – 28828
100. Alonso, M. D., Lomako, J., Lomako, W. M., and Whelan, W. J. (1995) *FASEB J.* **9**, 1126 – 1137
101. Blumenfeld, M. L., and Krisman, C. R. (1985) *J. Biol. Chem.* **260**, 11560 – 11566
101a. Pederson, B. A., Cheng, C., Wilson, W. A., and Roach, P. J. (2000) *J. Biol. Chem.* **275**, 27753 – 27761
102. Alonso, M. D., Lomako, J., Lomako, W. M., and Whelan, W. J. (1995) *J. Biol. Chem.* **270**, 15315 – 15319
103. Ercan, N., Gannon, M. C., and Nuttall, F. Q. (1994) *J. Biol. Chem.* **269**, 22328 – 22333
104. Mu, J., and Roach, P. J. (1998) *J. Biol. Chem.* **273**, 34850 – 34856
105. Mu, J., Skurat, A. V., and Roach, P. J. (1997) *J. Biol. Chem.* **272**, 27589 – 27597
106. Charng, Y.-y, Iglesias, A. A., and Preiss, J. (1994) *J. Biol. Chem.* **269**, 24107 – 24113
107. Nakata, P. A., Anderson, J. M., and Okita, T. W. (1994) *J. Biol. Chem.* **269**, 30798 – 30807
108. Van den Koornhuyse, N., Libessart, N., Delrue, B., Zabawinski, C., Decq, A., Iglesias, A., Carton, A., Preiss, J., and Ball, S. (1996) *J. Biol. Chem.* **271**, 16281 – 16287
109. Pozueta-Romero, J., Frehner, M., Viale, A. M., and Akazawa, T. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5769 – 5773
110. Yu, Y., Mu, H. H., Mu-Forster, C., and Wasserman, B. P. (1998) *Plant Physiol.* **116**, 1451 – 1460
111. Calvert, P. (1997) *Nature (London)* **389**, 338 – 339
112. Ball, S., Guan, H.-P., James, M., Myers, A., Keeling, P., Mouille, G., Buléon, A., Colonna, P., and Preiss, J. (1996) *Cell* **86**, 349 – 352
113. Waigh, T. A., Hopkinson, I., Donald, A. M., Butler, M. F., Heidelbach, F., and Riekel, C. (1997) *Macromolecules* **30**, 3813 – 3820
114. Gallant, D. J., Bouchet, B., and Baldwin, P. M. (1997) *Carbo. Polymers* **32**, 177 – 191
115. Fontaine, T., D'Hulst, C., Maddelein, M.-L., Routier, F., Pépin, T. M., Decq, A., Wieruszeski, J.-M., Delrue, B., Van den Koornhuyse, N., Bossu, J.-P., Fournet, B., and Ball, S. (1993) *J. Biol. Chem.* **268**, 16223 – 16230
116. Gao, M., Wanat, J., Stinard, P. S., James, M. G., and Myers, A. M. (1998) *Plant Cell* **10**, 399 – 412
117. Cao, H., Imparl-Radosevich, J., Guan, H. P., Keeling, P. L., James, M. G., and Myers, A. M. (1999) *Plant Physiol.* **120**, 1 – 11
118. Rahman, A., Wong, K.-s, Jane, J.-l, Myers, A. M., and James, M. G. (1998) *Plant Physiol.* **117**, 425 – 435
119. Guan, H., Kuriki, T., Sivak, M., and Preiss, J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 964 – 967
120. Beatty, M. K., Rahman, A., Cao, H., Woodman, W., Lee, M., Myers, A. M., and James, M. G. (1999) *Plant Physiol.* **119**, 255 – 266
121. Robyt, J. F. (1998) *Essentials of Carbohydrate Chemistry*, Springer, New York

122. Ball, S. G., van de Wal, M. H. B. J., and Visser, R. G. F. (1998) *Trends Plant Sci.* **3**, 462 – 467

122a. Sehnke, P. C., Chung, H.-J., Wu, K., and Ferl, R. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 765 – 770

123. French, D. and Robyt, J. F. (1973) *Abstr., 166th Natl. Meet., Am. Chem. Soc.*, Abstract 65 BIOL

123a. Robyt, J. F. (2000) *Abstr., 220th Natl. Meet., Am. Chem. Soc.*, Abstract 84 CARB

123b. Mukerjea, R., Yu, L., and Robyt, J. F. (2002) *Carbohydr. Res.* **337**, 1015 – 1022

124. McGuire, V., and Alexander, S. (1996) *J. Biol. Chem.* **271**, 14596 – 14603

125. Bureau, T. E., and Brown, R. M., Jr. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6985 – 6989

126. Lloyd, C. (1980) *Nature (London)* **284**, 596 – 597

127. Brown, R. M., Jr., Willison, J. H. M., and Richardson, C. L. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 4565 – 4569

128. Lin, F. C., Brown, R. M., Jr., Cooper, J. B., and Delmer, D. P. (1985) *Science* **230**, 822 – 825

129. Ross, P., Mayer, R., and Benziman, M. (1991) *Microbiol. Rev.* **55**, 35 – 58

130. Bokelman, G. H., Ruben, G. C., and Krakow, W. (1988) *J. Cell Biol.* **107**, 147a

131. Carpita, N., and Vergara, C. (1998) *Science* **279**, 672 – 673

131a. Scheible, W.-R., Eshed, R., Richmond, T., Delmer, D., and Somerville, C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10079 – 10084

131b. Kurek, I., Kawagoe, Y., Jacob-Wilk, D., Doblin, M., and Delmer, D. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11109 – 11114

132. Haigler, C. H., and Blanton, R. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12082 – 12085

133. Arioli, T., Peng, L., Betzner, A. S., Burn, J., Wittke, W., Herth, W., Camilleri, C., Höfte, H., Plazinski, J., Birch, R., Cork, A., Glover, J., Redmond, J., and Williamson, R. E. (1998) *Science* **279**, 717 – 720

133a. Cosgrove, D. J. (2000) *Nature (London)* **407**, 321 – 326

133b. Peng, L., Kawagoe, Y., Hogan, P., and Delmer, D. (2002) *Science* **295**, 147 – 150

134. Koyama, M., Helbert, W., Imai, T., Sugiyama, J., and Henrissat, B. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9091 – 9095

135. Han, N. S., and Robyt, J. F. (1998) *Carbohydr. Res.* **313**, 125 – 133

136. Sutherland, I. W. (1993) in *Industrial Gums*, 3rd ed. (Whistler, R. L., and BeMiller, J. N., eds), pp. 69 – 85, Academic Press, San Diego, California

137. Sugiyama, J., Boisset, C., Hashimoto, M., and Watanabe, T. (1999) *J. Mol. Biol.* **286**, 247 – 255

138. Mayer, R., Ross, P., Weinhouse, H., Amikam, D., Volman, G., Ohana, P., Calhoon, R. D., Wong, H. C., Emerick, A. W., and Benziman, M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5472 – 5476

139. Egli, M., Gessner, R. V., Williams, L. D., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Rich, A., and Frederick, C. A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3235 – 3239

139a. Chang, A. L., Tuckerman, J. R., Gonzalez, G., Mayer, R., Weinhouse, H., Volman, G., Amikam, D., Benziman, M., and Gilles-Gonzalez, M.-A. (2001) *Biochemistry* **40**, 3420 – 3426

140. Amor, Y., Haigler, C. H., Johnson, S., Wainscott, M., and Delmer, D. P. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9353 – 9357

141. Nakai, T., Tonouchi, N., Konishi, T., Kojima, Y., Tsuchida, T., Yoshinaga, F., Sakai, F., and Hayashi, T. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14 – 18

142. Martin, M. M., and Martin, J. S. (1978) *Science* **199**, 1453 – 1455

143. Skipper, N., Sutherland, M., Davies, R. W., Kilburn, D., Miller, R. C., Jr., Warren, A., and Wong, R. (1985) *Science* **230**, 958 – 960

144. Ohana, P., Delmer, D. P., Steffens, J. C., Matthews, D. E., Mayer, R., and Benziman, M. (1991) *J. Biol. Chem.* **266**, 13742 – 13745

145. Arellano, M., Durán, A., and Pérez, P. (1996) *EMBO J.* **15**, 4584 – 4591

146. Hrmova, M., Garrett, T. P. J., and Fincher, G. B. (1995) *J. Biol. Chem.* **270**, 14556 – 14563

147. Cabib, E. (1987) *Adv. Enzymol.* **59**, 59 – 101

148. Orlean, P. (1987) *J. Biol. Chem.* **262**, 5732 – 5739

149. Silverman, S. J., Sburlati, A., Slater, M. L., and Cabib, E. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4735 – 4739

150. Machida, S., and Saito, M. (1993) *J. Biol. Chem.* **268**, 1702 – 1707

151. Davis, L. L., and Bartnicki-Garcia, S. (1984) *Biochemistry* **23**, 1065 – 1073

152. Kafetzopoulos, D., Martinou, A., and Bouriotis, V. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2564 – 2568

153. Marx, J. L. (1977) *Science* **197**, 1170 – 1172

154. Bassler, B. L., Gibbons, P. J., Yu, C., and Roseman, S. (1991) *J. Biol. Chem.* **266**, 24268 – 24275

155. Sticher, L., Hofsteenge, J., Milani, A., Neuhaus, J.-M., and Meins, F., Jr. (1992) *Science* **257**, 655 – 657

156. Boot, R. G., Renkema, G. H., Verhoek, M., Strijland, A., Bliek, J., de Meulemeester, T. M. A. M. O., Mannens, M. M. A. M., and Aerts, J. M. F. G. (1998) *J. Biol. Chem.* **273**, 25680 – 25685

157. Semino, C. E., Specht, C. A., Raimondi, A., and Robbins, P. W. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4548 – 4553

158. Kamst, E., Bakkers, J., Quaedvlieg, N. E. M., Pilling, J., Kijne, J. W., Lugtenberg, B. J. J., and Spaink, H. P. (1999) *Biochemistry* **38**, 4045 – 4052

159. Gibeaut, D., and Carpita, N. C. (1994) *FASEB J.* **8**, 904 – 915

160. Darvill, S., McNeil, M., Albersheim, P., and Delmer, D. P. (1980) in *The Biochemistry of Plants*, Vol. I (Tolbert, N. E., ed), pp. 91 – 162, Academic Press, New York

161. Preston, R. D. (1979) *Ann. Rev. Plant Physiol.* **30**, 55

162. Brett, C. T., and Hillman, J. R., eds. (1985) *Biochemistry of Plant Cell Walls*, Cambridge Univ. Press, Cambridge

163. MacKay, A. L., Wallace, J. C., Sasaki, K., and Taylor, I. E. P. (1988) *Biochemistry* **27**, 1467 – 1473

163a. Carpita, N., and McCann, M. (2000) in *Biochemistry and Molecular Biology of Plants* (Buchanan, B. B., Gruissem, W., and Jones, R. L., eds), pp. 52 – 108, American Society of Plant Physiologists, Rockville, Maryland

164. Northcote, D. H. (1972) *Ann. Rev. Plant Physiol.* **23**, 113

164a. Perrin, R. M., DeRocher, A. E., Bar-Peled, M., Zeng, W., Norambuena, L., Orellana, A., Raikhel, N. V., and Keegstra, K. (1999) *Science* **284**, 1976 – 1979

164b. Kofod, L. V., Kauppinen, S., Christgau, S., Andersen, L. N., Heldt-Hansen, H. P., Dörreich, K., and Dalbøge, H. (1994) *J. Biol. Chem.* **269**, 29182 – 29189

164c. O'Neill, M. A., Warrenfeltz, D., Kates, K., Pellerin, P., Doco, T., Darvill, A. G., and Albersheim, P. (1996) *J. Biol. Chem.* **271**, 22923 – 22930

164d. Ishii, T., Matsunaga, T., Pellerin, P., O'Neill, M. A., Darvill, A., and Albersheim, P. (1999) *J. Biol. Chem.* **274**, 13098 – 13104

164e. Höfte, H. (2001) *Science* **294**, 795 – 797

164f. O'Neill, M. A., Eberhard, S., Albersheim, P., and Darvill, A. G. (2001) *Science* **294**, 846 – 849

165. Robertson, D., Mitchell, G. P., Gilroy, J. S., Gerrish, C., Bolwell, G. P., and Slabas, A. R. (1997) *J. Biol. Chem.* **272**, 15841 – 15848

166. Averyhart-Fullard, V., Datta, K., and Marcus, A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1082 – 1085

167. Condit, C. M., and Meagher, R. B. (1986) *Nature (London)* **323**, 178 – 181

168. Cosgrove, D. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5504 – 5505

169. Rose, J. K. C., Lee, H. H., and Bennett, A. B. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5955 – 5960

170. Fleming, A. J., McQueen-Mason, S., Mandel, T., and Kuhlemeier, C. (1997) *Science* **276**, 1415 – 1418

171. Stanley, P., and Ioffe, E. (1995) *FASEB J.* **9**, 1436 – 1444

172. Baenziger, J. U. (1994) *FASEB J.* **8**, 1019 – 1025

172a. DeAngelis, P. L., and Padgett-McCue, A. J. (2000) *J. Biol. Chem.* **275**, 24124 – 24129

172b. Pummill, P. E., Kempner, E. S., and DeAngelis, P. L. (2001) *J. Biol. Chem.* **276**, 39832 – 39835

173. Laurent, T. C., and Fraser, J. R. E. (1992) *FASEB J.* **6**, 2397 – 2404

174. Tlapak-Simmons, V. L., Baggenstoss, B. A., Kumari, K., Heldermon, C., and Weigel, P. H. (1999) *J. Biol. Chem.* **274**, 4246 – 4253

175. Spicer, A. P., and McDonald, J. A. (1998) *J. Biol. Chem.* **273**, 1923 – 1932

176. Robyt, J. F. (1979) *Trends Biochem. Sci.* **4**, 47 – 49

177. Robyt, J. F. (1995) *Adv. Carbohydr. Chem. Biochem.* **51**, 133 – 168

178. Robyt, J. F., and Martin, P. J. (1983) *Carbohydr. Res.* **113**, 301 – 315

179. Mooser, G., Hefta, S. A., Paxton, R. J., Shively, J. E., and Lee, T. D. (1991) *J. Biol. Chem.* **266**, 8916 – 8922

180. Funane, K., Shiraiwa, M., Hashimoto, K., Ichishima, E., and Kobayashi, M. (1993) *Biochemistry* **32**, 13696 – 13702

181. Sprenger, N., Bortlik, K., Brandt, A., Boller, T., and Wiemken, A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11652 – 11656

182. Beale, J. M., Jr., and Foster, J. L. (1996) *Biochemistry* **35**, 4492 – 4501

183. Hardingham, T. E., and Fosang, A. J. (1992) *FASEB J.* **6**, 861 – 870

184. Huang, S., Wang, Y.-X., and Draper, D. E. (1996) *J. Mol. Biol.* **258**, 308 – 321

185. Kokenyesi, R., and Bernfield, M. (1994) *J. Biol. Chem.* **269**, 12304 – 12309

186. Kearns, A. E., Campbell, S. C., Westley, J., and Schwartz, N. B. (1991) *Biochemistry* **30**, 7477 – 7483

187. Shworak, N. W., Shirakawa, M., Mulligan, R. C., and Rosenberg, R. D. (1994) *J. Biol. Chem.* **269**, 21204 – 21214

188. Vertel, B. M., Walters, L. M., Flay, N., Kearns, A. E., and Schwartz, N. B. (1993) *J. Biol. Chem.* **268**, 11105 – 11112

189. Fernández, C. J., and Warren, G. (1998) *J. Biol. Chem.* **273**, 19030 – 19039

190. Lugemwa, F. N., Sarkar, A. K., and Esko, J. D. (1996) *J. Biol. Chem.* **271**, 19159 – 19165

190a. Bai, X., Wei, G., Sinha, A., and Esko, J. D. (1999) *J. Biol. Chem.* **274**, 13017 – 13024

190b. Götting, C., Kuhn, J., Zahn, R., Brinkmann, T., and Kleesiek, K. (2000) *J. Mol. Biol.* **304**, 517 – 528

190c. Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., and Negishi, M. (2000) *J. Biol. Chem.* **275**, 34580 – 34585

191. Sugahara, K., Ohkita, Y., Shibata, Y., Yoshida, K., and Ikegami, A. (1995) *J. Biol. Chem.* **270**, 7204 – 7212

References

192. Pavão, M. S. G., Aiello, K. R. M., Werneck, C. C., Silva, L. C. F., Valente, A.-P., Mulloy, B., Colwell, N. S., Tollefsen, D. M., and Mourão, P. A. S. (1998) *J. Biol. Chem.* **273**, 27848–27857
193. Sugumaran, G., Katsman, M., and Drake, R. R. (1995) *J. Biol. Chem.* **270**, 22483–22487
194. Yamauchi, S., Hirahara, Y., Usui, H., Takeda, Y., Hoshino, M., Fukuta, M., Kimura, J. H., and Habuchi, O. (1999) *J. Biol. Chem.* **274**, 2456–2463
195. Kobayashi, M., Habuchi, H., Habuchi, O., Saito, M., and Kimata, K. (1996) *J. Biol. Chem.* **271**, 7645–7653
196. Razi, N., and Lindahl, U. (1995) *J. Biol. Chem.* **270**, 11267–11275
197. Shworak, N. W., Liu, J., Petros, L. M., Zhang, L., Kobayashi, M., Copeland, N. G., Jenkins, N. A., and Rosenberg, R. D. (1999) *J. Biol. Chem.* **274**, 5170–5184
198. Toma, L., Berninsone, P., and Hirschberg, C. B. (1998) *J. Biol. Chem.* **273**, 22458–22465
199. Aikawa, J.-i, and Esko, J. D. (1999) *J. Biol. Chem.* **274**, 2690–2695
200. Scott, J. E., Heatley, F., and Wood, B. (1995) *Biochemistry* **34**, 15467–15474
201. Yanagashita, M., and Hascall, V. C. (1992) *J. Biol. Chem.* **267**, 9451–9454
202. Lane, D. A., and Björk, I., eds. (1992) *Heparin and Related Polysaccharides*, Plenum, New York
203. Lindahl, U., Kusche-Gullberg, M., and Kjellén, L. (1998) *J. Biol. Chem.* **273**, 24979–24982
203a. Safaiyan, F., Lindahl, U., and Salmivirta, M. (2000) *Biochemistry* **39**, 10823–10830
203b. Zhang, L., Lawrence, R., Schwartz, J. J., Bai, X., Wei, G., Esko, J. D., and Rosenberg, R. D. (2001) *J. Biol. Chem.* **276**, 28806–28813
204. Kinoshita, A., Yamada, S., Haslam, S. M., Morris, H. R., Dell, A., and Sugahara, K. (1997) *J. Biol. Chem.* **272**, 19656–19665
205. Nadanaka, S., Clement, A., Masayama, K., Faissner, A., and Sugahara, K. (1998) *J. Biol. Chem.* **273**, 3296–3307
206. Mourão, P. A. S., Pereira, M. S., Pavão, M. S. G., Mulloy, B., Tollefsen, D. M., Mowinckel, M.-C., and Abildgaard, U. (1996) *J. Biol. Chem.* **271**, 23973–23984
207. Petitou, M., Hérault, J.-P., Bernat, A., Driguez, P.-A., Duchaussoy, P., Lormeau, J.-C., and Herbert, J.-M. (1999) *Nature (London)* **398**, 417–422
208. Sinaÿ, P. (1999) *Nature (London)* **398**, 377–378
208a. Zehnder, J. L., and Galli, S. J. (1999) *Nature (London)* **400**, 714–715
208b. Hondal, R. J., Ma, S., Caprioli, R. M., Hill, K. E., and Burk, R. F. (2001) *J. Biol. Chem.* **276**, 15823–15831
209. Spiro, R. C., Casteel, H. E., Laufer, D. M., Reisfeld, R. A., and Harper, J. R. (1989) *J. Biol. Chem.* **264**, 1779–1786
210. Iozzo, R. V., and Murdoch, A. D. (1996) *FASEB J.* **10**, 598–614
211. Cheng, F., Heinegård, D., Fransson, L.-Å., Bayliss, M., Bielicki, J., Hopwood, J., and Yoshida, K. (1996) *J. Biol. Chem.* **271**, 28572–28580
212. Hauser, N., Paulsson, M., Heinegård, D., and Mörgelin, M. (1996) *J. Biol. Chem.* **271**, 32247–32252
213. Zheng, J., Luo, W., and Tanzer, M. L. (1998) *J. Biol. Chem.* **273**, 12999–13006
213a. Olin, A. I., Mörgelin, M., Sasaki, T., Timpl, R., Heinegård, D., and Aspberg, A. (2001) *J. Biol. Chem.* **276**, 1253–1261
214. Arner, E. C., Pratta, M. A., Trzaskos, J. M., Decicco, C. P., and Tortorella, M. D. (1999) *J. Biol. Chem.* **274**, 6594–6601
215. Brown, G. M., Huckerby, T. N., Bayliss, M. T., and Nieduszynski, I. A. (1998) *J. Biol. Chem.* **273**, 26408–26414
216. Lindahl, B., Eriksson, L., Spillmann, D., Caterson, B., and Lindahl, U. (1996) *J. Biol. Chem.* **271**, 16991–16994
217. Dours-Zimmermann, M. T., and Zimmermann, D. R. (1994) *J. Biol. Chem.* **269**, 32992–32998
217a. Yang, B. L., Cao, L., Kiani, C., Lee, V., Zhang, Y., Adams, M. E., and Yang, B. B. (2000) *J. Biol. Chem.* **275**, 21255–21261
217b. Pangalos, M. N., Efthimiopoulos, S., Shioi, J., and Robakis, N. K. (1995) *J. Biol. Chem.* **270**, 10388–10391
218. Halfter, W., Dong, S., Schurer, B., and Cole, G. J. (1998) *J. Biol. Chem.* **273**, 25404–25412
219. Scott, J. E. (1996) *Biochemistry* **35**, 8795–8799
220. Weber, I. T., Harrison, R. W., and Iozzo, R. V. (1996) *J. Biol. Chem.* **271**, 31767–31770
221. Tai, G.-H., Huckerby, T. N., and Nieduszynski, I. A. (1996) *J. Biol. Chem.* **271**, 23535–23546
222. Ungefroren, H., and Krull, N. B. (1996) *J. Biol. Chem.* **271**, 15787–15795
222a. Yang, V. W.-C., LaBrenz, S. R., Rosenberg, L. C., McQuillan, D., and Höök, M. (1999) *J. Biol. Chem.* **274**, 12454–12460
223. Sommarin, Y., Wendel, M., Shen, Z., Hellman, U., and Heinegård, D. (1998) *J. Biol. Chem.* **273**, 16723–16729
223a. Iozzo, R. V. (1999) *J. Biol. Chem.* **274**, 18843–18846
223b. Saika, S., Shiraishi, A., Saika, S., Liu, C.-Y., Funderburgh, J. L., Kao, C. W.-C., Converse, R. L., and Kao, W. W.-Y. (2000) *J. Biol. Chem.* **275**, 2607–2612
223c. Bengtsson, E., Aspberg, A., Heinegård, D., Sommarin, Y., and Spillmann, D. (2000) *J. Biol. Chem.* **275**, 40695–40702
224. David, C. L., Orpiszewski, J., Zhu, X.-C., Reissner, K. J., and Aswad, D. W. (1998) *J. Biol. Chem.* **273**, 32063–32070
225. Grumet, M., Milev, P., Sakurai, T., Karthikeyan, L., Bourdon, M., Margolis, R. K., and Margolis, R. U. (1994) *J. Biol. Chem.* **269**, 12142–12146
226. Fernàndez-Busquets, X., Kammerer, R. A., and Burger, M. M. (1996) *J. Biol. Chem.* **271**, 23558–23565
227. Varner, J. A. (1996) *J. Biol. Chem.* **271**, 16119–16125
228. Hart, G. W., Holt, G. D., and Haltiwanger, R. S. (1988) *Trends Biochem. Sci.* **13**, 380–384
229. Sadler, J. E. (1984) in *"Biology of Carbohydrates"*, Vol. 2 (Ginsburg, V., and Robbins, P. W., eds), pp. 199–288, Wiley, New York
230. Jentoft, N. (1990) *Trends Biochem. Sci.* **15**, 291–294
231. Dong, D. L.-Y., Xu, Z.-S., Chevrier, M. R., Cotter, R. J., Cleveland, D. W., and Hart, G. W. (1993) *J. Biol. Chem.* **268**, 16679–16687
231a. Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., and Marth, J., eds. (1999) *Essentials of Glycobiology*, Cold Spring Harbor Lab. Press, Plainview, New York
231b. Roseman, S. (2001) *J. Biol. Chem.* **276**, 41527–41542
232. Hanisch, F.-G., Müller, S., Hassan, H., Clausen, H., Zachara, N., Gooley, A. A., Paulsen, H., Alving, K., and Peter-Katalinic, J. (1999) *J. Biol. Chem.* **274**, 9946–9954
233. Schwientek, T., Nomoto, M., Levery, S. B., Merkx, G., van Kessel, A. G., Bennett, E. P., Hollingsworth, M. A., and Clausen, H. (1999) *J. Biol. Chem.* **274**, 4504–4512
233a. Comer, F. I., and Hart, G. W. (2000) *J. Biol. Chem.* **275**, 29179–29182
233b. Wells, L., Vosseller, K., and Hart, G. W. (2001) *Science* **291**, 2376–2378
233c. Hanover, J. A. (2001) *FASEB J.* **15**, 1865–1876
234. Abeijon, C., and Hirschberg, C. B. (1992) *Trends Biochem. Sci.* **17**, 32–36
235. Moloney, D. J., Lin, A. I., and Haltiwanger, R. S. (1997) *J. Biol. Chem.* **272**, 19046–19050
236. Allen, A. (1983) *Trends Biochem. Sci.* **8**, 169–173
237. Yeh, J.-C., Ong, E., and Fukuda, M. (1999) *J. Biol. Chem.* **274**, 3215–3221
238. Gum, J. R., Jr., Ho, J. J. L., Pratt, W. S., Hicks, J. W., Hill, A. S., Vinall, L. E., Roberton, A. M., Swallow, D. M., and Kim, Y. S. (1997) *J. Biol. Chem.* **272**, 26678–6686
239. Desseyn, J.-L., Buisine, M.-P., Porchet, N., and Aubert, J.-P. (1998) *J. Biol. Chem.* **273**, 30157–30164
240. Perez-Vilar, J., and Hill, R. L. (1998) *J. Biol. Chem.* **273**, 34527–34534
241. Sagami, H., Kurisaki, A., and Ogura, K. (1993) *J. Biol. Chem.* **268**, 10109–10113
242. Opdenakker, G., Rudd, P. M., Ponting, C. P., and Dwek, R. A. (1993) *FASEB J.* **7**, 1330–1337
243. Gahmberg, C. G., and Tolvanen, M. (1996) *Trends Biochem. Sci.* **21**, 308–311
244. Manzella, S. M., Hooper, L. V., and Baenziger, J. U. (1996) *J. Biol. Chem.* **271**, 12117–12120
245. Drickamer, K., and Taylor, M. E. (1998) *Trends Biochem. Sci.* **23**, 321–324
246. Elbein, A. D. (1981) *Trends Biochem. Sci.* **6**, 219–221
247. Zhu, X., Zeng, Y., and Lehrman, M. A. (1992) *J. Biol. Chem.* **267**, 8895–8902
248. Rajput, B., Muniappa, N., and Vijay, I. K. (1994) *J. Biol. Chem.* **269**, 16054–16061
249. Zara, J., and Lehrman, M. A. (1994) *J. Biol. Chem.* **269**, 19108–19115
249a. Ünligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J. M. (2000) *EMBO J.* **19**, 5269–5280
250. Silberstein, S., and Gilmore, R. (1996) *FASEB J.* **10**, 849–858
251. Imperiali, B., and Shannon, K. L. (1991) *Biochemistry* **30**, 4374–4380
252. Yan, Q., Prestwich, G. D., and Lennarz, W. J. (1999) *J. Biol. Chem.* **274**, 5021–5025
252a. Chen, X., VanValkenburgh, C., Liang, H., Fang, H., and Green, N. (2001) *J. Biol. Chem.* **276**, 2411–2416
252b. Karaoglu, D., Kelleher, D. J., and Gilmore, R. (2001) *Biochemistry* **40**, 12193–12206
253. Schwarz, R. T., and Datema, R. (1984) *Trends Biochem. Sci.* **9**, 32–34
254. Elbein, A. D. (1987) *Ann. Rev. Biochem.* **56**, 497–534
255. Kaushal, G. P., Pan, Y. T., Tropea, J. E., Mitchell, M., Liu, P., and Elbein, A. D. (1988) *J. Biol. Chem.* **263**, 17278–17283
256. Kukuruzinska, M. A., Bergh, M. L. E., and Jackson, B. J. (1987) *Ann. Rev. Biochem.* **56**, 915–944
257. Hatton, M. W. C., Marz, L., and Regoeczi, E. (1983) *Trends Biochem. Sci.* **8**, 287–291
257a. Vallée, F., Karaveg, K., Herscovics, A., Moremen, K. W., and Howell, P. L. (2000) *J. Biol. Chem.* **275**, 41287–41298
257b. Van Petegem, F., Contreras, H., Contreras, R., and Van Beeumen, J. (2001) *J. Mol. Biol.* **312**, 157–165
258. Herscovics, A., and Orlean, P. (1993) *FASEB J.* **7**, 540–550
258a. Dell, A., and Morris, H. R. (2001) *Science* **291**, 2351–2356
259. Dolci, E. D., and Palade, G. E. (1985) *J. Biol. Chem.* **260**, 10728–10735
260. Gonatas, J. O., Mezitis, S. G. E., Stieber, A., Fleischer, B., and Gonatas, N. K. (1989) *J. Biol. Chem.* **264**, 646–653
261. Hooper, L. V., Hindsgaul, O., and Baenziger, J. U. (1995) *J. Biol. Chem.* **270**, 16327–16332
262. Okajima, T., Fukumoto, S., Miyazaki, H., Ishida, H., Kiso, M., Furukawa, K., Urano, T., and Furukawa, K. (1999) *J. Biol. Chem.* **274**, 11479–11486

References

263. Lee, Y.-C., Kaufmann, M., Kitazume-Kawaguchi, S., Kono, M., Takashima, S., Kurosawa, N., Liu, H., Pircher, H., and Tsuji, S. (1999) *J. Biol. Chem.* **274**, 11958 – 11967
264. Trimble, R. B., and Atkinson, P. H. (1986) *J. Biol. Chem.* **261**, 9815 – 9824
265. Chiba, Y., Suzuki, M., Yoshida, S., Yoshida, A., Ikenaga, H., Takeuchi, M., Jigami, Y., and Ichishima, E. (1998) *J. Biol. Chem.* **273**, 26298 – 26304
266. Kollár, R., Reinhold, B. B., Petrákova, E., Yeh, H. J. C., Ashwell, G., Drgonová, J., Kapteyn, J. C., Klis, F. M., and Cabib, E. (1997) *J. Biol. Chem.* **272**, 17762 – 17775
266a. Helenius, A., and Aebi, M. (2001) *Science* **291**, 2364 – 2369
266b. Lehrman, M. A. (2001) *J. Biol. Chem.* **276**, 8623 – 8626
266c. Bertozzi, C. R., and Kiessling, L. L. (2001) *Science* **291**, 2357 – 2364
266d. Plante, O. J., Palmacci, E. R., and Seeberger, P. H. (2001) *Science* **291**, 1523 – 1527
267. Mazzarello, P., and Bentivoglio, M. (1998) *Nature (London)* **392**, 543 – 544
268. Featherstone, C. (1998) *Science* **282**, 2172 – 2174
269. Berger, E. G., and Roth, J., eds. (1997) *The Golgi Apparatus*, Birkhäuser Verlag, Basel
270. Driouich, A., Faye, L., and Staehelin, L. A. (1993) *Trends Biochem. Sci.* **18**, 210 – 214
271. Lazar, T., Götte, M., and Gallwitz, D. (1997) *Trends Biochem. Sci.* **22**, 468 – 472
272. Farquhar, M. G., and Hauri, H.-P. (1997) in *The Golgi Apparatus* (Berger, E. G., and Roth, J., eds), pp. 63 – 129, Birkhäuser Verlag, Basel, Switzerland
273. Springer, S., and Schekman, R. (1998) *Science* **281**, 698 – 700
274. Walter, D. M., Paul, K. S., and Waters, M. G. (1998) *J. Biol. Chem.* **273**, 29565 – 29576
275. Dice, J. F. (1990) *Trends Biochem. Sci.* **15**, 305 – 309
276. Rambourg, A., and Clermont, Y. (1997) in *The Golgi Apparatus* (Berger, E. G., and Roth, J., eds), pp. 37 – 61, Birkhäuser Verlag, Basel, Switzerland
277. Ungermann, C., and Wickner, W. (1998) *EMBO J.* **17**, 3269 – 3276
278. Wu, S.-K., Zeng, K., Wilson, I. A., and Balch, W. E. (1996) *Trends Biochem. Sci.* **21**, 472 – 476
279. Webb, R. J., East, J. M., Sharma, R. P., and Lee, A. G. (1998) *Biochemistry* **37**, 673 – 679
280. Ivessa, N. E., De Lemos-Chiarandini, C., Gravotta, D., Sabatini, D. D., and Kreibich, G. (1995) *J. Biol. Chem.* **270**, 25960 – 25967
280a. Todorow, Z., Spang, A., Carmack, E., Yates, J., and Schekman, R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13643 – 13648
280b. Marsh, B. J., Mastronarde, D. N., Buttle, K. F., Howell, K. E., and McIntosh, J. R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2399 – 2406
280c. Jakymiw, A., Raharjo, E., Rattner, J. B., Eystathioy, T., Chan, E. K. L., and Fujita, D. J. (2000) *J. Biol. Chem.* **275**, 4137 – 4144
280d. Hirschberg, C. B., Robbins, P. W., and Abeijon, C. (1998) *Ann. Rev. Biochem.* **67**, 49 – 69
280e. Gao, X.-D., and Dean, N. (2000) *J. Biol. Chem.* **275**, 17718 – 17727
280f. Bell, A. W., and 16 other authors. (2001) *J. Biol. Chem.* **276**, 5152 – 5165
280g. Alvarez, C., Garcia-Mata, R., Hauri, H.-P., and Sztul, E. (2001) *J. Biol. Chem.* **276**, 2693 – 2700
280h. Charest, A., Lane, K., McMahon, K., and Housman, D. E. (2001) *J. Biol. Chem.* **276**, 29456 – 29465
280i. Demaurex, N., Furuya, W., D'Souza, S., Bonifacino, J. S., and Grinstein, S. (1998) *J. Biol. Chem.* **273**, 2044 – 2051
281. Fritzler, M. J., Lung, C.-C., Hamel, J. C., Griffith, K. J., and Chan, E. K. L. (1995) *J. Biol. Chem.* **270**, 31262 – 31268
282. Bascon, R. A., Srinivasan, S., and Nussbaum, R. L. (1999) *J. Biol. Chem.* **274**, 2953 – 2962
283. Kain, R., Angata, K., Kerjaschki, D., and Fukuda, M. (1998) *J. Biol. Chem.* **273**, 981 – 988
284. Johannes, L., Tenza, D., Antony, C., and Goud, B. (1997) *J. Biol. Chem.* **272**, 19554 – 19561
285. Drickamer, K. (1988) *J. Biol. Chem.* **263**, 9557 – 9560
286. Cantor, A. B., and Kornfeld, S. (1992) *J. Biol. Chem.* **267**, 23357 – 23363
287. Kornfeld, S. (1992) *Ann. Rev. Biochem.* **61**, 307 – 330
288. Pohlmann, R., Boeker, M. W. C., and von Figura, K. (1995) *J. Biol. Chem.* **270**, 27311 – 27318
289. Sleat, D. E., and Lobel, P. (1997) *J. Biol. Chem.* **272**, 731 – 738
290. York, S. J., Arneson, L. S., Gregory, W. T., Dahms, N. M., and Kornfeld, S. (1999) *J. Biol. Chem.* **274**, 1164 – 1171
290a. Zhu, Y., Doray, B., Poussu, A., Lehto, V.-P., and Kornfeld, S. (2001) *Science* **292**, 1716 – 1718
291. Sleat, D. E., Sohar, I., Lackland, H., Majercak, J., and Lobel, P. (1996) *J. Biol. Chem.* **271**, 19191 – 19198
292. Klionsky, D. J., and Emr, S. D. (1990) *J. Biol. Chem.* **265**, 5349 – 5352
293. Lodish, H. F. (1991) *Trends Biochem. Sci.* **16**, 374 – 377
294. Chao, W., Liu, H., Hanahan, D. J., and Olson, M. S. (1992) *J. Biol. Chem.* **267**, 6725 – 6735
295. Chiu, M. H., Thomas, V. H., Stubbs, H. J., and Rice, K. G. (1995) *J. Biol. Chem.* **270**, 24024 – 24031
296. Reichner, J. S., Whiteheart, S. W., and Hart, G. W. (1988) *J. Biol. Chem.* **263**, 16316 – 16326
296a. Filipe, S. R., Severina, E., and Tomasz, A. (2001) *J. Biol. Chem.* **276**, 39618 – 39628
297. Emanuele, J. J., Jr., Jin, H., Yanchunas, J., Jr., and Villafranca, J. J. (1997) *Biochemistry* **36**, 7264 – 7271
298. Anderson, M. S., Eveland, S. S., Onishi, H. R., and Pompliano, D. L. (1996) *Biochemistry* **35**, 16264 – 16269
299. Duncan, K., van Heijenoort, J., and Walsh, C. T. (1990) *Biochemistry* **29**, 2379 – 2386
300. Van Heijenoort, J. (1994) in *Bacterial Cell Wall (New Comprehensive Biochemistry)*, Vol. 27 (Ghuysen, J.-M., and Hakenbeck, R., eds), pp. 39 – 54, Elsevier, Amsterdam
301. Matsuhashi, M. (1994) in *Bacterial Cell Wall (New Comprehensive Biochemistry)*, Vol. 27 (Ghuysen, J.-M., and Hakenbeck, R., eds), pp. 55 – 71, Elsevier, Amsterdam
301a. Ha, S., Chang, E., Lo, M.-C., Men, H., Park, P., Ge, M., and Walker, S. (1999) *J. Am. Chem. Soc.* **121**, 8417 – 8426
301b. Marmor, S., Petersen, C. P., Reck, F., Yang, W., Gao, N., and Fisher, S. L. (2001) *Biochemistry* **40**, 12207 – 12214
301c. Gordon, E., Flouret, B., Chantalat, L., van Heijenoort, J., Mengin-Lecreulx, D., and Dideberg, O. (2001) *J. Biol. Chem.* **276**, 10999 – 11006
301d. Yan, Y., Munshi, S., Leiting, B., Anderson, M. S., Chrzas, J., and Chen, Z. (2000) *J. Mol. Biol.* **304**, 435 – 445
301e. Lee, W., McDonough, M. A., Kotra, L. P., Li, Z.-H., Silvaggi, N. R., Takeda, Y., Kelly, J. A., and Mobashery, S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1427 – 1431
302. Koch, A. L. (1985) *Trends Biochem. Sci.* **10**, 11 – 14
303. Archibald, A. R., Hancock, I. C., and Harwood, C. R. (1993) in *Bacillus subtilis and Other Gram-Positive Bacteria* (Sonenshein, A. L., Hoch, J. A., and Losick, R., eds), pp. 381 – 410, American Society for Microbiology, Washington, D. C.
304. Jacobs, C. (1997) *Science* **278**, 1731 – 1732
305. Jacobs, C., Huang, L.-j, Bartowsky, E., Normark, S., and Park, J. T. (1994) *EMBO J.* **13**, 4684 – 4694
305a. Templin, M. F., Ursinus, A., and Höltje, J.-V. (1999) *EMBO J.* **18**, 4108 – 4117
306. Keenleyside, W. J., and Whitfield, C. (1996) *J. Biol. Chem.* **271**, 28581 – 28592
307. Raetz, C. R. H., Ulevitch, R. J., Wright, S. D., Sibley, C. H., Ding, A., and Nathan, C. F. (1991) *FASEB J.* **5**, 2652 – 2660
308. Odegaard, T. J., Kaltashov, I. A., Cotter, R. J., Steeghs, L., van der Ley, P., Khan, S., Maskell, D. J., and Raetz, C. R. H. (1997) *J. Biol. Chem.* **272**, 19688 – 19696
308a. Raetz, C. R. H., and Whitfield, C. (2002) *Ann. Rev. Biochem.* **71**, 635 – 700
309. Yethon, J. A., Heinrichs, D. E., Monteiro, M. A., Perry, M. B., and Whitfield, C. (1998) *J. Biol. Chem.* **273**, 26310 – 26316
310. Raetz, C. R. H., and Roderick, S. L. (1995) *Science* **270**, 997 – 1000
311. Jackman, J. E., Raetz, C. R. H., and Fierke, C. A. (1999) *Biochemistry* **38**, 1902 – 1911
312. Zhou, Z., White, K. A., Polissi, A., Georgopoulos, C., and Raetz, C. R. H. (1998) *J. Biol. Chem.* **273**, 12466 – 12475
313. Luk, J. M. C., Lind, S. M., Tsang, R. S. W., and Lindberg, A. A. (1991) *J. Biol. Chem.* **266**, 23215 – 23225
314. Kerwood, D. E., Schneider, H., and Yamasaki, R. (1992) *Biochemistry* **31**, 12760 – 12768
315. Basu, S. S., York, J. D., and Raetz, C. R. H. (1999) *J. Biol. Chem.* **274**, 11139 – 11149
316. Rietschel, E. T., Kirikae, T., Schade, F. U., Mamat, U., Schmidt, G., Loppnow, H., Ulmer, A. J., Zähringer, U., Seydel, U., Di Padova, F., Schreier, M., and Brade, H. (1994) *FASEB J.* **8**, 217 – 225
317. White, K. A., Kaltashov, I. A., Cotter, R. J., and Raetz, C. R. H. (1997) *J. Biol. Chem.* **272**, 16555 – 16563
318. Brooke, J. S., and Valvano, M. A. (1996) *J. Biol. Chem.* **271**, 3608 – 3614
319. Süsskind, M., Brade, L., Brade, H., and Holst, O. (1998) *J. Biol. Chem.* **273**, 7006 – 7017
320. Pavliak, V., Brisson, J.-R., Michon, F., Uhrín, D., and Jennings, H. J. (1993) *J. Biol. Chem.* **268**, 14146 – 14152
321. Phillips, N. J., Apicella, M. A., Griffiss, J. M., and Gibson, B. W. (1993) *Biochemistry* **32**, 2003 – 2012
321a. White, K. A., Lin, S., Cotter, R. J., and Raetz, C. R. H. (1999) *J. Biol. Chem.* **274**, 31391 – 31400
321b. Kooistra, O., Lüneberg, E., Lindner, B., Knirel, Y. A., Frosch, M., and Zähringer, U. (2001) *Biochemistry* **40**, 7630 – 7640
322. Ganfield, M.-C. W., and Pieringer, R. A. (1980) *J. Biol. Chem.* **255**, 5164 – 5169
322a. Volkman, B. F., Zhang, Q., Debabov, D. V., Rivera, E., Kresheck, G. C., and Neuhaus, F. C. (2001) *Biochemistry* **40**, 7964 – 7972
323. Ton-That, H., Labischinski, H., Berger-Bächi, B., and Schneewind, O. (1998) *J. Biol. Chem.* **273**, 29143 – 29149
324. Fischetti, V. A. (1991) *Sci. Am.* **264**(Jun), 58 – 65
325. Michon, F., Brisson, J.-R., Dell, A., Kasper, D. L., and Jennings, H. J. (1988) *Biochemistry* **27**, 5341 – 5351
326. Cassels, F. J., Fales, H. M., London, J., Carlson, R. W., and van Halbeek, H. (1990) *J. Biol. Chem.* **265**, 14127 – 14135
327. Wolucka, B. A., McNeil, M. R., de Hoffmann, E., Chonjnacki, T., and Brennan, P. J. (1994) *J. Biol. Chem.* **269**, 23328 – 23335
328. Besra, G. S., McNeil, M. R., Rivoire, B., Khoo, K.-H., Morris, H. R., Dell, A., and Brennan, P. J. (1993) *Biochemistry* **32**, 347 – 355

References

329. Beveridge, T. J., and Koval, S. F., eds. (1993) *Advances in Bacterial Paracrystalline Surface Layers*, Plenum, New York

329a. Maeda, Y., Watanabe, R., Harris, C. L., Hong, Y., Ohishi, K., Kinoshita, K., and Kinoshita, T. (2001) *EMBO J.* **20**, 250– 261

330. Takeda, J., and Kinoshita, T. (1995) *Trends Biochem. Sci.* **20**, 367–371

331. Udenfriend, S., and Kodukula, K. (1995) *Ann. Rev. Biochem.* **64**, 563–591

332. Tarutani, M., Itami, S., Okabe, M., Ikawa, M., Tezuka, T., Yoshikawa, K., Kinoshita, T., and Takeda, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7400–7405

333. Menon, A. K., and Stevens, V. L. (1992) *J. Biol. Chem.* **267**, 15277–15280

334. Medof, M. E., Nagarajan, S., and Tykocinski, M. L. (1996) *FASEB J.* **10**, 574–586

334a. Ferguson, M. A. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10673–10675

334b. Mahoney, A. B., Sacks, D. L., Saraiva, E., Modi, G., and Turco, S. J. (1999) *Biochemistry* **38**, 9813–9823

334c. Smith, T. K., Gerold, P., Crossman, A., Paterson, M. J., Borissow, C. N., Brimacombe, J. S., Ferguson, M. A. J., and Schwarz, R. T. (2002) *Biochemistry* **41**, 12395–12406

335. Ralton, J. E., and McConville, M. J. (1998) *J. Biol. Chem.* **273**, 4245–4257

336. Gilleron, M., Nigou, J., Cahuzac, B., and Puzo, G. (1999) *J. Mol. Biol.* **285**, 2147–2160

337. Benghezal, M., Benachour, A., Rusconi, S., Aebi, M., and Conzelmann, A. (1996) *EMBO J.* **15**, 6575–6583

338. Zawadzki, J., Scholz, C., Currie, G., Coombs, G. H., and McConville, M. J. (1998) *J. Mol. Biol.* **282**, 287–299

339. Smith, T. K., Sharma, D. K., Crossman, A., Dix, A., Brimacombe, J. S., and Ferguson, M. A. J. (1997) *EMBO J.* **16**, 6667–6675

339a. Morita, Y. S., Acosta-Serrano, A., Buxbaum, L. U., and Englund, P. T. (2000) *J. Biol. Chem.* **275**, 14147–14154

340. Leidich, S. D., and Orlean, P. (1996) *J. Biol. Chem.* **271**, 27829–27837

341. Chien, J.-L., and Hogan, E. L. (1983) *J. Biol. Chem.* **258**, 10727–10730

342. van Echten, G., and Sandhoff, K. (1993) *J. Biol. Chem.* **268**, 5341–5344

343. Sprong, H., Kruithof, B., Leijendekker, R., Slot, J. W., van Meer, G., and van der Sluijs, P. (1998) *J. Biol. Chem.* **273**, 25880–25888

344. Schwientek, T., Almeida, R., Levery, S. B., Holmes, E. H., Bennett, E., and Clausen, H. (1998) *J. Biol. Chem.* **273**, 29331–29340

345. Lannert, H., Gorgas, K., Meissner, I., Wieland, F. T., and Jeckel, D. (1998) *J. Biol. Chem.* **273**, 2939–2946

346. Jaskiewicz, E., Zhu, G., Bassi, R., Darling, D. S., and Young, W. W., Jr. (1996) *J. Biol. Chem.* **271**, 26395–26403

347. Neufeld, E. F., Lim, T. W., and Shapiro, L. J. (1975) *Ann. Rev. Biochem.* **44**, 357–376

348. Neufeld, E. F. (1991) *Ann. Rev. Biochem.* **60**, 257–280

349. Tager, J. M. (1985) *Trends Biochem. Sci.* **10**, 324–326

350. Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds (1995) *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1, McGraw-Hill, New York (pp. 2427–2879)

351. von Figura, K., and Hasilik, A. (1984) *Trends Biochem. Sci.* **9**, 29–31

352. Neufeld, E. F., and Muenzer, J. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2465–2494, McGraw-Hill, New York

353. Stoltzfus, L. J., Sosa-Pineda, B., Moskowitz, S. M., Menon, K. P., Dlott, B., Hooper, L., Teplow, D. B., Shull, R. M., and Neufeld, E. F. (1992) *J. Biol. Chem.* **267**, 6570–6575

354. Kresse, H., Paschke, E., von Figura, K., Gilberg, W., and Fuchs, W. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6822–6826

355. Wilson, P. J., Morris, C. P., Anson, D. S., Occhiodoro, T., Bielicki, J., Clements, P. R., and Hopwood, J. J. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8531–8535

356. Wicker, G., Prill, V., Brooks, D., Gibson, G., Hopwood, J., von Figura, K., and Peters, C. (1991) *J. Biol. Chem.* **266**, 21386–21391

357. Wu, B. M., Tomatsu, S., Fukuda, S., Sukegawa, K., Orii, T., and Sly, W. S. (1994) *J. Biol. Chem.* **269**, 23681–23688

358. McGovern, M. M., Aula, P., and Desnick, R. J. (1983) *J. Biol. Chem.* **258**, 10743–10747

359. Mononen, I., Fisher, K. J., Kaartinen, V., and Aronson, N. N., Jr. (1993) *FASEB J.* **7**, 1247–1256

360. Saarela, J., Laine, M., Tikkanen, R., Oinonen, C., Jalanko, A., Rouvinen, J., and Peltonen, L. (1998) *J. Biol. Chem.* **273**, 25320–25328

361. Rothman, J. E., and Lenard, J. (1984) *Trends Biochem. Sci.* **9**, 176–178

362. Beutler, E. (1992) *Science* **256**, 794–799

363. Beutler, E. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5384–5390

364. Beutler, E., and Grabowski, G. A. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2641–2670, McGraw-Hill, New York

365. Sano, A., Radin, N. S., Johnson, L. L., and Tarr, G. E. (1988) *J. Biol. Chem.* **263**, 19597–19601

366. Myerowitz, R., and Costigan, F. C. (1988) *J. Biol. Chem.* **263**, 18587–18589

366a. Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G., and James, M. N. G. (2001) *J. Biol. Chem.* **276**, 10330 –10337

367. Neufeld, E. F. (1989) *J. Biol. Chem.* **264**, 10927–10930

368. Gravel, R. A., Clarke, J. T. R., Kaback, M. M., Mahuran, D., Sandhoff, K., and Suzuki, K. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2839–2879, McGraw-Hill, New York

369. Paw, B. H., Moskowitz, S. M., Uhrhammer, N., Wright, N., Kaback, M. M., and Neufeld, E. F. (1990) *J. Biol. Chem.* **265**, 9452–9457

369a. Kawai, H., Allende, M. L., Wada, R., Kono, M., Sango, K., Deng, C., Miyakawa, T., Crawley, J. N., Werth, N., Bierfreund, U., Sandhoff, K., and Proia, R. L. (2001) *J. Biol. Chem.* **276**, 6885–6888

370. Hama, Y., Li, Y.-T., and Li, S.-C. (1997) *J. Biol. Chem.* **272**, 2828–2833

371. Beutler, E. (1981) *Trends Biochem. Sci.* **6**, 95–97

371a. Schiffmann, R., and 21 other authors (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 365 –370

371b. Sly, W. S., Vogler, C., Grubb, J. H., Zhou, M., Jiang, J., Zhou, X. Y., Tomatsu, S., Bi, Y., and Snella, E. M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2205 – 2210

372. Furbish, F. S., Steer, C. J., Barranger, J. A., Jones, E. A., and Brady, R. O. (1978) *Biochem. Biophys. Res. Commun.* **81**, 1047–1053

373. Krivit, W., Pierpont, M. E., Ayaz, K., Tsai, M., Ramsay, N. K. C., Kersey, J. H., Weisdorf, S., Sibley, R., Snover, D., McGovern, M. M., Schwartz, M. F., and Desnick, R. J. (1984) *N. Engl. J. Med.* **311**, 1606–1611

374. Sorge, J., Kuhl, W., West, C., and Beutler, E. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 906–909

375. Fairbairn, L. J., Lashford, L. S., Spooncer, E., McDermott, R. H., Lebens, G., Arrand, J. E., Arrand, J. R., Bellantuono, I., Holt, R., Hatton, C. E., Cooper, A., Besley, G. T. N., Wraith, J. E., Anson, D. S., Hopwood, J. J., and Dexter, T. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2025–2030

375a. Qin, G., Takenaka, T., Telsch, K., Kelley, L., Howard, T., Levade, T., Deans, R., Howard, B. H., Malech, H. L., Brady, R. O., and Medin, J. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3428– 3433

376. Beaudet, A. L., Scriver, C. R., Sly, W. S., and Valle, D. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 53–118, McGraw-Hill, New York

377. Brenkert, A., Arora, R. C., Radin, N. S., Meier, H., and MacPike, A. D. (1972) *Brain Res.* **36**, 195–202

378. Morell, P., and Constantino-Ceccarini, E. (1972) *Lipids* **7**, 266–268

379. Seppala, R., Tietze, F., Krasnewich, D., Weiss, P., Ashwell, G., Barsh, G., Thomas, G. H., Packman, S., and Gahl, W. A. (1991) *J. Biol. Chem.* **266**, 7456–7461

380. Spellacy, E., Shull, R. M., Constantopoulos, G., and Neufeld, E. F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6091–6095

381. Shull, R. M., Kakkis, E. D., McEntee, M. F., Kania, S. A., Jonas, A. J., and Neufeld, E. F. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12937–12941

382. Walkley, S. U., Thrall, M. A., Dobrenis, K., Huang, M., March, P. A., Siegel, D. A., and Wurzelmann, S. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2970–2974

Study Questions

1. Constitution of cell surface oligosaccharides or polysaccharides includes the following:
    - D-Glucose
    - D-Mannose
    - D-Galactose
    - L-Arabinose
    - L-Fucose
    - D-Glucuronic acid
    - D-Neuraminic acid

    Outline pathways for biosynthesis of these compounds from glucose.

2. Decarboxylation steps are required for synthesis of UDP-xylulose and UDP-apiose (Fig. 21-1). Propose chemical mechanisms for these reactions.

3. How do animals and plants differ with respect to transport and storage of glucose?

4. Comment on unresolved questions about the biosynthesis of cellulose, amylose, and amylopectin. What glycosyl carrier groups are required?

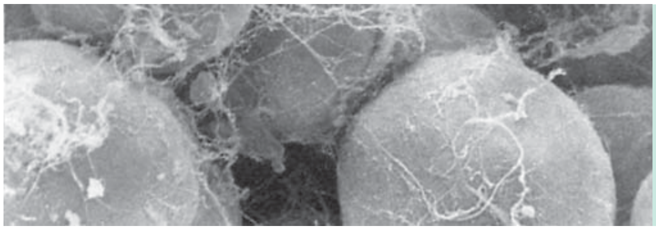
5. Most 5- and 6-membered sugars are found in nature as pyranose ring forms. Why is ribose in RNA in the furanose ring form?

6. If the ratio $[NAD^+]/[NADH]$ in a cell were 500 and the ratio $[NADP^+]/[NADPH]$ were 0.002, what concentrations of fructose and sorbitol would be in equilibrium with 0.1 mM glucose? See Box 20-A and Table 6-4.

7. Write a balanced equation for reaction of boric acid (H_3BO_3) with two sugar rings to give a borate diol ester linkage (Box 20-E).

8. Describe in general terms the process by which *N*-linked oligosaccharides are synthesized and attached to proteins. What are the functions of the ER and the Golgi?

9. What, if any, restrictions do you think should be applied to the use of antibiotics on farms?



Two large fat-filled adipose cells are seen in the foreground of this scanning electron micrograph. They are part of a larger cluster of cells from rat tissue. Delicate strands of connective tissue fibers intertwine the cells and hold them together. While most of the connective tissue substance has been washed away during preparation of the specimen, the remnants give a realistic impression of the soft, loose nature of the intercellular material. From Porter and Bonneville (1973) *Fine Structure of Cells and Tissues*, Lea and Febiger, Philadelphia, Pennsylvania. Courtesy of Mary Bonneville.

Contents

1181 .... **A. Digestion, Synthesis, and Distribution of Triacylglycerols in the Human Body**
1181 .... 1. Plasma Lipoproteins
1184 .... 2. Movement of Lipid Materials Between Cells
1185 .... **B. The Biosynthesis of Fatty Acids and Their Esters**
1185 .... 1. Fatty Acid Synthases
1188 .... 2. Control of Chain Length
1189 .... 3. Starter Pieces and Branches
1189 .... 4. Synthesis by the Oxoacid Chain Elongation Process
1191 .... 5. Unsaturated Fatty Acids
1193 .... 6. Cyclopropane Fatty Acids and Mycolic Acids
1195 .... 7. The Lipids of Skin and Other Surfaces
1196 .... *Surface lipids of plants*
1196 .... *Formation of hydrocarbons*
1196 .... *Insect waxes, hydrocarbons, and pheromones*
1196 .... **C. Synthesis of Triacylglycerols, Glycolipids, and Phospholipids**
1197 .... 1. Phospholipids
1201 .... 2. The Ether-Linked Lipids
1202 .... 3. Sphingolipids
1202 .... 4. Complex Lipids in Signaling
1204 .... 5. Peroxidation of Lipids and Rancidity
1205 .... 6. Some Nutritional Questions
1206 .... **D. Prostaglandins and Related Prostanoid Compounds**
1206 .... 1. Metabolism of the Prostaglandins
1208 .... 2. Thromboxanes and Prostacyclins
1208 .... 3. Lipoxygenases
1210 .... 4. Leukotrienes, Lipoxins, and Related Compounds
1210 .... 5. Physiological Effects of the Prostanoids
1211 .... 6. Inflammation
1212 .... 7. Plant Lipoxygenases and Jasmonic Acid
1212 .... **E. The Polyketides**

1218 .... **References**
1223 .... **Study Questions**

Boxes

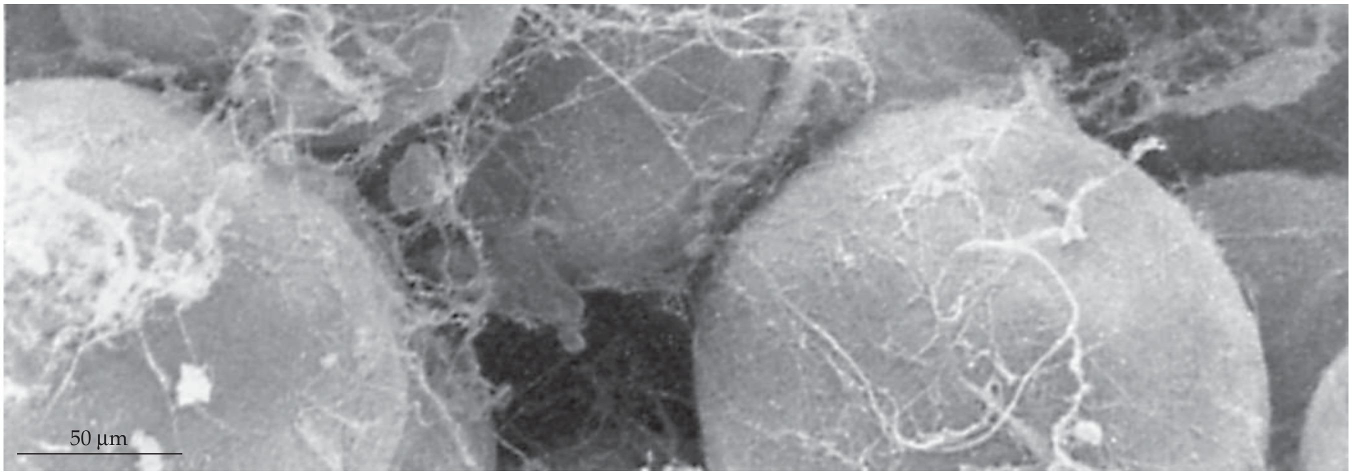
1186 .... Box 21-A Lipocalins, Fatty Acid-Binding Proteins, and Lipophorins
1190 .... Box 21-B The Essential Fatty Acids
1194 .... Box 21-C Tuberculosis
1200 .... Box 21-D Poly-β-Hydroxybutyrate and Biodegradable Plastics
1214 .... Box 21-E How the Flowers Make Their Colors

Tables

1183 .... Table 21-1 Classes of Lipoprotein Particles
1183 .... Table 21-2 Properties of Major Plasma Apolipoproteins
1189 .... Table 21-3 Starter Pieces for Biosynthesis of Fatty Acids

Specific Aspects of Lipid Metabolism

21



The basic pathways for both synthesis of fatty acids and for their β-oxidation (Fig. 17-1) have been described in Chapter 17. However, there are many variations to these pathways, and additional sets of enzymes are needed to synthesize the complex array of lipids present in most organisms. We will consider these details in this chapter. Like most other organisms, human beings are able to synthesize triacylglycerols (triglycerides), phospholipids, and glycolipids needed for cell membranes. Glucose can serve as the starting material. However, dietary lipids are also a major source. For this reason, we will start with a discussion of the digestion and uptake of lipids and of the distribution by way of the bloodstream of ingested lipids and of lipids synthesized in the liver or in other tissues.

A. Digestion, Synthesis, and Distribution of Triacylglycerols in the Human Body

Digestion of triglycerides begins in the stomach with emulsification and partial digestion by gastric lipase. Within the small intestine the ~100-residue protein called **colipase**[1–2a] binds to the surface of the fat droplets and provides an attachment site for the 449-residue **pancreatic lipase**. This Ca^{2+}-dependent serine esterase cleaves each triglyceride to two molecules of fatty acid and one of a 2-monoacylglycerol.[3–5] These products are emulsified by bile salts (Fig. 22-10) and are then taken up by the cells of the intestinal lining. The fatty acids are converted to acyl-CoA esters which transfer their acyl groups to the monoacylglycerols to resynthesize the triacylglycerols.[6] The latter are incorporated into the very large lipoprotein particles called **chylomicrons** (Table 21-1) and enter the bloodstream via the lymphatic system (Fig. 21-1).[7] Free fatty acids are also transported as complexes with serum albumin.

Synthesis of lipids from carbohydrates is an efficient process, which occurs largely in the liver and also in intestinal epithelial cells.[6] The newly synthesized triacylglycerols, together with smaller amounts of phospholipids and cholesterol, combine with specific **apolipoproteins**, which are also synthesized in the liver, to form **very low density lipoprotein (VLDL)** particles which are secreted into the blood stream. They transport the newly formed triacylglycerols from the liver to other body cells including the adipocytes, which store excess fat (Fig. 21-1).

1. Plasma Lipoproteins

The small particles of plasma lipoprotein, which carry triacylglycerols, can be separated according to their buoyant densities by centrifugation. They have been classified into five groups of increasing density but smaller size as **chylomicrons,** very low density lipoproteins **(VLDL)**, intermediate density lipoproteins (**IDL**), low density **(LDL),** and high density lipoproteins **(HDL)** (Table 21-1 and Fig. 21-2). Each lipoprotein particle contains one or more apolipoproteins (Table 21-2), whose sizes vary from the enormous 4536-residue apoB-100 to apoC-II and apoC-III, each of which contains just 79 residues[7a] and the 57-residue apoC-I.[7b]

The larger lipoproteins are spherical micelles containing a core of triacylglycerols and esters of cholesterol surrounded by a 2- to 3-nm-thick layer consisting of phospholipid, free cholesterol, and the apolipoprotein components.[8] The size of the lipoprotein particles also varies from a 200- to 500-nm diameter for chylomicrons to as little as 5 nm for the smallest HDL particles. The difference in volume is more impressive. If, as has been estimated,[9] a 22-nm diameter LDL particle contains about 2000 cholesterol and cholesteryl ester molecules and 800 phospholipids, a small HDL particle of 7-nm diameter will have room for only about 60 molecules of cholesterol and 90 of phospholipid, while a chylomicron may carry 10 million molecules of triacylglycerol. HDL particles are quite heterogeneous. As is indicated in Table 21-1, they are sometimes dividing into HDL2 and HDL3 density groups. In addition, there is a pre-HDL with lower phospholipid content and discoid forms low in cholesterol. Models of a reconstituted lipoprotein disc contain two molecules of apoA-I and ~160 phosphatidylcholines that form a bilayer core.[10–10b]

Each apolipoprotein has one or more distinct functions. The apoB proteins probably stabilize the lipoprotein micelles. In addition, apoB-100 is essential to recognition of LDL by its receptors. The 79-residue apoC-II has a specific function of activating the lipoprotein lipase that hydrolyses the triacylglycerols of chylomicrons and VLDL. Lack of either C-II or the lipase results in a very high level of triacylglycerols in the blood.[11]

The large apolipoprotein B-100 is synthesized in the liver and is a principal component of VLDL, IDL, and LDL. It is the sole protein in LDL, accounting for nearly 20% of the mass of LDL particles. Partly because of its insolubility in water, its detailed structure is uncertain. If it were all coiled into an α helix, it would be 680 nm long and could encircle the LDL particle nearly 10 times! While the true structure of apoB-100 is unknown, it is thought to be extended and to span at least a hemisphere of the LDL surface.[12] It consists of at least five domains. Sixteen cysteines are present in the first 25 residues at the N terminus, forming a crosslinked high-cysteine region. There are also 16 *N*-glycosylated sites. Domain IV (residues 3071–4011) is thought to contain the site that binds to its specific receptor, the LDL receptor.[12] Heterogeneity in the amide I band of the infrared absorption spectrum (Fig. 23-3) suggests that about 24% is α helix, 23% β sheet, and that a large fraction consists of turns, and unordered and extended peptide structures.[13]

In intestinal epithelial cells the same apoB gene that is used to synthesize apoB-100 in the liver is used to make the shorter **apoB-48** (48%) protein. This is accomplished in an unusual way that involves "editing" of the mRNA that is formed. Codon 2153 in the mRNA for the protein is CAA, encoding glutamine. However, the cytosine of the triplet is acted on by a deaminase, an editing enzyme, to form UAA, a chain termination codon.[14,15] A third form of apoB is found in **lipoprotein(a)** (Lpa). This LDL-like particle contains apoB-100 to which is covalently attached by a single disulfide linkage (probably to Cys 3734 of apoB-100) a second protein, **apo(a)**. The latter consists largely of a chain of from 11 to over 50 kringle domains resembling the 78-residue kringle-4 of plasminogen (see Fig. 7-30C)[16–19a] as well as a protease domain.[20] This additional chain may cause tighter binding to LDL receptors and may cause lipoprotein(a) to displace plasminogen from cell surface receptors.[21] The amount of Lp(a) varies over 1000-fold among individuals and is genetically determined. The number of kringle domains also varies.[17] Although the presence of high Lp(a) is associated with a high risk of atherosclerosis and stroke,[21a] many healthy 100-year olds also have high serum Lp(a).[22]

Apolipoprotein A-I is the primary protein component of HDL.[23–25b] Most of the 243 residues consist of a nearly continuous amphipathic α helix with kinks at regularly spaced proline residues.[26–28] Two disulfide-linked ApoA-I molecules may form a belt that encircles the discoid lipoprotein.[25b] ApoA-II is the second major HDL protein, but no clearly specialized function has been identified.[29,30] ApoA-I, II, and IV, apoC-I, II, and III, and apoE all have multiple repeats of 22 amino acids with sequences that suggest amphipathic helices. The 391-residue ApoA-IV has 13 tandem 22-residue repeats. Proline and glycine are present in intervening hinge regions.[23] This may enable these proteins to spread over and penetrate the surfaces of the lipoprotein micelles. Most of these proteins are encoded by a related multigene family.[7,30a]

The 299-residue apolipoprotein E plays a key role in metabolism of both triacylglycerols and cholesterol. Like apoB-100 it binds to cell surface receptors.[31–33a] Absence of functional apoE leads to elevated plasma triacylglycerol and cholesterol, a problem that is considered in Chapter 22, Section D. The N-terminal domain, from residues 23 to 164, forms a 6.5 nm-long four-helix bundle, which binds to the LDL receptors.[32] There are three common isoforms of alipoprotein E (apoE2, apoE3, and apoE4). ApoE3 is most common.[33b] The presence of apoE4 is associated with an increased risk of Alzheimer disease (Chapter 30).

The major lipoproteins of insect hemolymph, the **lipophorins**, transport diacylglycerols. The apolipophorins have molecular masses of ~250, 80, and sometimes 18 kDa.[34–37a] The three-dimensional structure of a small 166-residue lipophorin (apolipophorin-III) is that of a four-helix bundle. It has been suggested that it may partially unfold into an extended form, whose amphipathic helices may bind to a phospholipid surface of the lipid micelle of the lipophorin.[35] A similar behavior may be involved in binding of mammalian apolipoproteins. Four-helix lipid-binding proteins have also been isolated from plants.[38] See also Box 21-A. Specialized lipoproteins known as **lipovitellins**

TABLE 21-1
Classes of Lipoprotein Particles

Class	Diameter (nm)	Density (g/ml)	Composition (weight percent)[a]				
			Surface components			Core lipids	
			Protein	Phospholipid	Cholesterol	Cholesteryl esters	Triacylglycerol
Chylomicrons	75–1200	0.93	2	7	2	3	86
VLDL	30–80	0.93–1.006	8	18	7	12	55
IDL	25–35	1.006–1.019	19	19	9	29	23
LDL	18–25	1.019–1.063	22	22	8	42	6
HDL2	9–12	1.063–1.125	40	33	5	17	5
HDL3	5–9	1.125–1.21	45	35	4	13	3
Lp (a); slow pre-β	25–30	1.04–1.09					

[a] Data from Havel, R. J., and Kane, J. P. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. II (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1841–1852, McGraw-Hill, New York.

These are averages and there is considerable variation.

TABLE 21-2
Properties of Major Plasma Apolipoproteins

Designation	No. residues	Mass (kDa)	Source	Function
A-I	243	29		Major HDL protein
A-II	—	17.4	Liver and intestine	
A-IV	376	44.5		
B-100	4536	513	Liver	VLDL formation; ligand for LDL receptor
B-48	2152	241	Intestine	Chylomicron formation ligand for liver chylomicron receptor
C-I	57	6.6		
C-II	79	8.9	Liver	Cofactor for lipoprotein lipase
C-III	79	8.8		
D	—	31	Many tissues	A lipocalin
E	299	34	Liver, VLDL	Ligand for LDL receptor
(a)	Variable			Ligand for liver chylomicron receptor

store phospholipid in eggs whether from nematodes, frogs, or chickens.[39] There is some sequence similarity to that of human apolipoprotein B-100.

2. Movement of Lipid Materials Between Cells

After the synthesis and release of chylomicrons into the lymphatic circulation, various exchange processes occur by which apolipoproteins, as well as enzymes and other proteins, may be added or removed. These very complex and incompletely understood phenomena are presented in simplified form in Fig. 21-1. Chylomicrons donate apolipoproteins of the A and C families to HDL particles which, in turn, donate apoE and may also return some apoC protein to the chylomicrons.

Both chylomicrons and VLDL particles undergo similar processes in the capillary blood vessels, where their triacylglycerols are hydrolyzed to glycerol and free fatty acids by **lipoprotein lipase**.[40–42a] This enzyme requires for its activity the apolipoprotein C-II which is present in the chylomicrons and VLDL particles. Lipoprotein lipase is also known as the "clearing

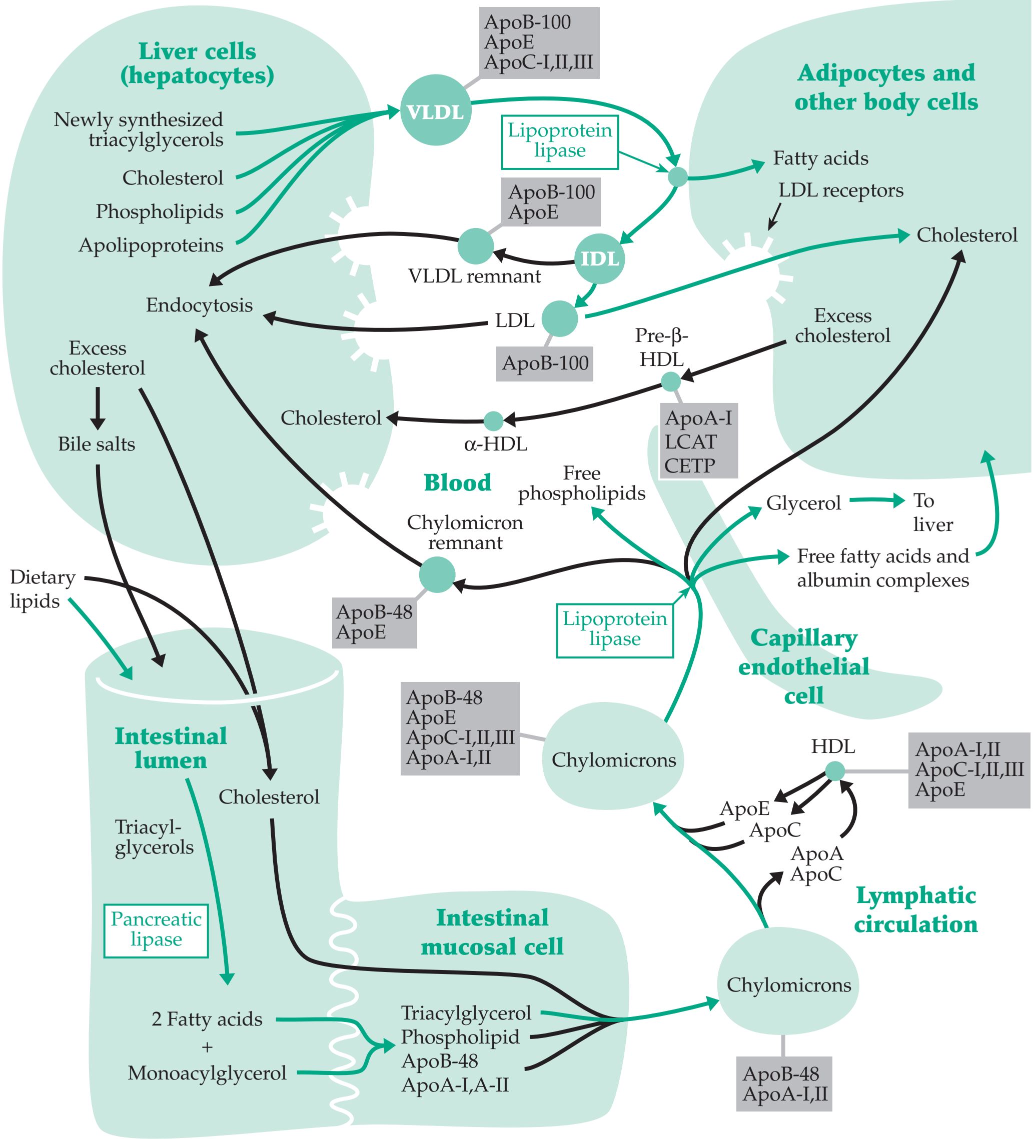

ApoB-100
ApoE
ApoC-I,II,III
Liver cells (hepatocytes)
VLDL
Adipocytes and other body cells
Newly synthesized triacylglycerols
Cholesterol
Phospholipids
Apolipoproteins
Lipoprotein lipase
Fatty acids
LDL receptors
ApoB-100
ApoE
IDL
Cholesterol
VLDL remnant
Endocytosis
LDL
ApoB-100
Excess cholesterol
Pre-β-HDL
Excess cholesterol
Cholesterol
α-HDL
ApoA-I
LCAT
CETP
Bile salts
Blood
Free phospholipids
Glycerol
To liver
Chylomicron remnant
Free fatty acids and albumin complexes
Dietary lipids
ApoB-48
ApoE
Lipoprotein lipase
Capillary endothelial cell
ApoB-48
ApoE
ApoC-I,II,III
ApoA-I,II
Chylomicrons
Intestinal lumen
HDL
ApoA-I,II
ApoC-I,II,III
ApoE
Cholesterol
ApoE
ApoC
Triacyl-glycerols
ApoA
ApoC
Lymphatic circulation
Pancreatic lipase
Intestinal mucosal cell
Chylomicrons
2 Fatty acids
+
Monoacylglycerol
Triacylglycerol
Phospholipid
ApoB-48
ApoA-I,A-II
ApoB-48
ApoA-I,II

Figure 21-1 Movement of triacylglycerols from liver and intestine to body cells and lipid carriers of blood. VLDL; very low density lipoprotein which contains triacylglycerols, phospholipids, cholesterol, and apolipoproteins B, and C. IDL; intermediate density lipoproteins found in human plasma. LDL; low density lipoproteins which have lost most of their triacylglycerols. ApoB-100, etc., are apolipoproteins listed in Table 21-2. LCAT, lecithin: cholesterol acyltransferase; CETP, cholesteryl ester transfer protein (see Chapter 22).

factor" because it clears the milky chylomicron-containing lymph. It is secreted by adipocytes and other cells and becomes attached to heparan sulfate proteglycans on surfaces of capillary endothelial cells, a major site of its action.[43] Hereditary absence of functional lipoprotein lipase causes **chylomicronemia**, a massive buildup of chylomicrons in plasma.[41,44] The condition does not cause atherosclerosis but may lead to pancreatitis if not treated. Restriction of dietary fat to 20 g / day or less usually prevents problems. Naturally occurring mutations of lipoprotein lipase involving both the aspartate of the catalytic triad (p. 635)[45] and the flexible loop that covers the active site[46] have been discovered.

Both lipoprotein lipase and the less well understood **hepatic lipase** are related structurally to pancreatic lipase.[42,42b] In addition to hydrolysis of the triacylglycerols, the uptake of materials from lipoproteins probably involves shedding of intact phospholipids, perhaps as liposome-like particles.[40]

The free fatty acids and glycerol are taken up by mammalian tissue cells leaving the cholesterol and some of the phospholipids of the VLDL particles as LDL. In humans **intermediate density lipoproteins** (IDL) are formed initially, but some are converted to LDL later. Both LDL particles and the shrunken **chylomicron remnants** and **VLDL remnants** are taken up by endocytosis in coated pits and are degraded by body cells, principally of the liver.[47,48] The best known of these receptors is the 839-residue **LDL receptor,** which has a specific affinity for ApoB-100. The related **VLDL receptor** (apoE receptor) has a higher affinity for apoE[48–50] and may function in uptake of both VLDL and chylomicron remnants. The **LDL receptor-related protein** functions as a third lipoprotein receptor.[51] In addition, a series of **scavenger receptors**, found in abundance in macrophages, take up oxidized lipoproteins and other materials.[51,52] Scavenger receptor B1 (SR-B1), which is also found in liver cells, is involved in uptake of cholesterol from HDL particles by hepatocytes [53] (see also Chapter 22). Liver cells, and other cells as well, contain **lipocalins** and **fatty acid binding proteins** (Box 21-A) that help to carry these relatively insoluble acids to their destinations within the cells. Serum albumin (Box 2-A) is also a major carrier of free fatty acids.[53a] Within the adipocytes the fatty acids are reconverted to triacylglycerols. The low density (LDL) and high density (HDL) lipoproteins are involved primarily in transporting cholesterol to and from cells, a topic that is discussed in Chapter 22, Section D,2.

Fatty acids are carried to tissues for use in synthesis of triacylglycerols, phospholipids, and other membrane lipids. The mobilization of fatty acids from triacylglycerol stores and from cholesterol esters depends upon **hormone-sensitive lipase** (p. 635).[53b, 53c] This enzyme is activated by cAMP-dependent phosphorylation and moves from the cytoplasm to the surfaces of lipid droplets in response to catecholamines and other lipolytic hormones. Fatty acids are a major fuel for aerobic cells. Their conversion to acyl-CoA derivative and oxidation to CO_2 by beta oxidation (Fig. 17-1) and other pathways are discussed in Chapter 17 (pp. 939 – 950).

B. The Biosynthesis of Fatty Acids and Their Esters

The synthesis of fatty acids two carbon atoms at a time from acetyl-CoA has been considered in Chapter 17 and is outlined in Fig. 17-12. In this pathway, which resembles the β oxidation sequence in reverse, the products are saturated fatty acids with an even number of carbon atoms as shown in Fig. 21-2. In this section, we will consider some of the factors that lead to variations in the chain lengths and types of fatty acids.

1. Fatty Acid Synthases

Both bacteria and plants have separate enzymes that catalyze the individual steps in the biosynthetic sequence (Fig. 17-12). The fatty acyl group grows while attached to the small acyl carrier protein (ACP).[54–58] Control of the process is provided, in part, by the existence of isoenzyme forms. For example, in *E. coli* there are three different β-oxoacyl-ACP synthases. They carry out the transfer of any acyl primer from ACP to the enzyme, decarboxylate malonyl-ACP, and carry out the Claisen condensation (steps *b*, *e*, and *f* in Eq. 17-12)[58a – e] One of the isoenzymes is specialized for the initial elongation of acetyl-ACP and also provides feedback regulation.[58c] The other two function specifically in synthesis of unsaturated fatty acids.

In a few bacteria and protozoa and in higher animals the fatty acid synthase consists of only one or two multifunctional proteins. That from animal tissues contains six enzymes and an acyl carrier protein (ACP) domain as well. The human enzyme contains 2504 amino acid residues organized as a series of functional domains.[59–59b] Pairs of the 272-kDa chains associate to form 544-kDa dimers. The complex protein may have arisen via an evolutionary process involving fusion of formerly separate genes.[60] The enzyme contains an ACP-like site with a bound 4'-phosphopantetheine near the C terminus as well as a cysteine side chain near the N terminus in the second acylation site. Since the two –SH groups can be crosslinked by dibromopropanone,[61,62] an antiparallel linear arrangement of the two chains was proposed.[63–65] Locations of the six enzymatic activities in each chain are indicated on page 1187. According to this picture,

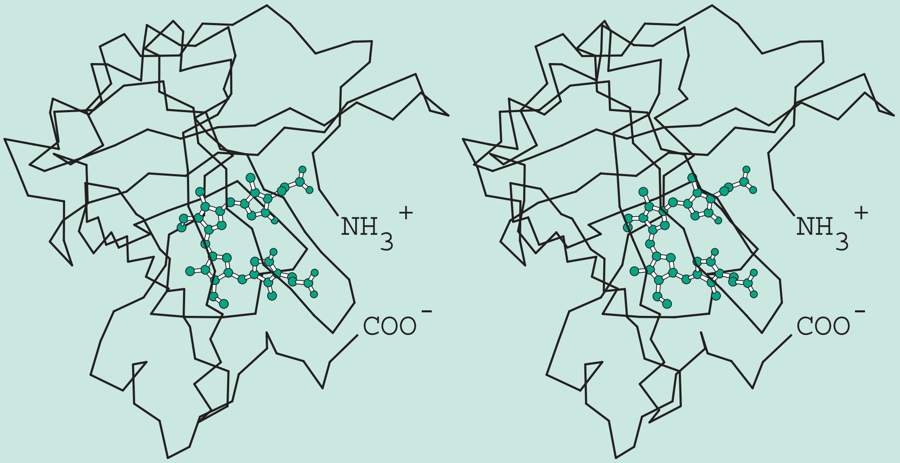
BOX 21-A LIPOCALINS, FATTY ACID-BINDING PROTEINS, AND LIPOPHORINS

Small hydrophobic molecules, which might easily bind in biologically undesirable ways, are chaperoned in animals, plants, and bacteria by binding proteins that provide hydrophobic cavities or crevices appropriate for holding these molecules in readily releasable forms. The **lipocalins**, most of which are extracellular proteins, have a conserved structural motif consisting of an 8-stranded β barrel arranged as two stacked orthogonal sheets with a C-terminal α helix that blocks one end. The other end is able to open and allow a small hydrophobic molecule to bind in the internal cavity[a–c] (see figure). Only three short amino acid sequences are conserved within a large family of lipocalins[b,d] which includes **plasma retinol-binding protein**,[e] mammalian **odorant-binding proteins**,[f] **α-lactalbumin**, **apolipoprotein D**,[a] and the blue biliverdin-binding protein **insecticyanin** of insect hemolymph.[g,h] Most lipocalins are soluble, but some such as the plasma **α1-microglobulin**,[i] which plays a role in the immune system (Chapter 31), have additional functions that require them to bind to other proteins or to cell surfaces.[j] The **gelatinase-associated lipocalin** of human neutrophils binds bacterially derived *N*-formylpeptides that act as chemotactic agents (Chapter 19) and induces release of materials from intracellular granules.[c] A few lipocalins have enzymatic activity. For example, **prostaglandin D synthase** is both an enzyme and a carrier of bile pigments and thyroid hormones.[k] Most lipocalins have been found in higher animals, but at least a few bacterial proteins belong to the family.[d] One is the 77-residue *E. coli* outer membrane lipoprotein.[a]

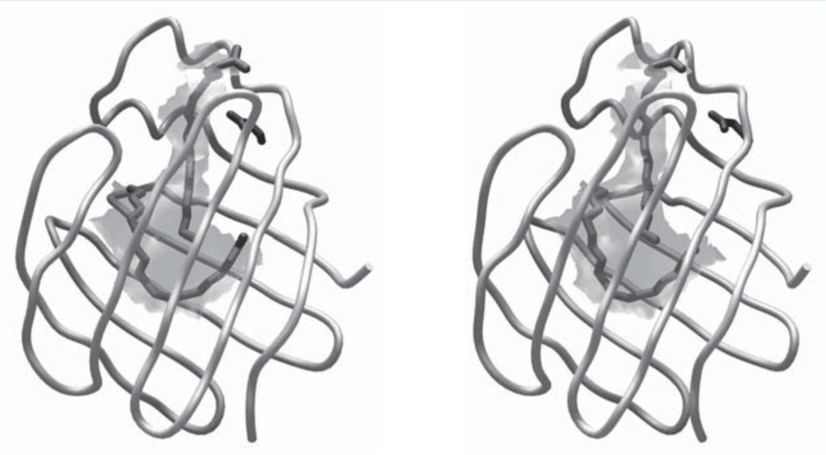
A related family of proteins are represented by **fatty acid-binding proteins**[l–o] and by the intracellular **retinol-** and **retinoic acid-binding proteins** (see also Box 22-A).[p] These are 10-stranded antiparallel β-barrels with two helices blocking an end (see Figure).

A third group of lipid-binding proteins have a four-helix bundle structure. They include the insect **lipophorins**, which transport diacylglycerols in the hemolymph (see main text), and nonspecific lipid carriers of green plants.[q] An 87-residue four-helix protein with a more open structure binds acyl-coenzyme A molecules in liver.[r]

A small 98-residue sterol-binding protein from



Stereoscopic view of an α-carbon model of an insecticyanin subunit with the bound biliverdin. The N and C termini are labeled NH_3^+ and COO^-, respectively. The positions of several amino acid residues are indicated. From Holden *et al.*[g] Courtesy of Hazel Holden.



Structure of a crystalline fatty acid-binding protein from liver with two molecules of bound oleate (dark rods). The lower molecule is more deeply embedded in the protein and more tightly bound than the second molecule, which is closer to the outer surface of the protein. Semitransparent grey marks the solvent-accessible surface of the binding cavity. An unknown molecule, perhaps butanoic acid (as modeled), binds also at top of the protein. See Thompson *et al.*[n] Courtesy of Leonard Banaszak.

BOX 21-A (continued)

the fungus *Phytophthora*, an agriculturally important plant pathogen, has a very different folding pattern. The sterol binds into a cavity formed by six helices and two loops. The protein is not only a sterol carrier but an **elicitin**, which induces a defensive response in the invaded plant. The function of the protein for the invader may be to acquire sterols for the fungus, which is unable to synthesize them.[s]

Many larger lipid carrier proteins are known. The 476-residue plasma cholesteryl ester transfer protein is discussed briefly in Chapter 22. Plasma phospholipid transfer proteins are of similar size.[t,u] A 456-residue human phospholipid-binding protein interacts with the lipopolysaccharide of the surfaces of gram-negative bacteria (Fig. 8-30) and participates in the immune response to the bacteria. It has an elongated boomerang shape with two cavities, both of which bind a molecule of phosphatidylcholine. Other plasma lipid transfer proteins may have similar structures.[v]

---

[a] Bishop, R. E., Penfold, S. S., Frost, L. S., Höltje, J.-V., and Weiner, J. H. (1995) *J. Biol. Chem.* **270**, 23097–23103

[b] Flower, D. R., North, A. C. T., and Attwood, T. K. (1993) *Protein Sci.* **2**, 753–761

[c] Coles, M., Diercks, T., Muehlenweg, B., Bartsch, S., Zölzer, V., Tschesche, H., and Kessler, H. (1999) *J. Mol. Biol.* **289**, 139–157

[d] Flower, D. R., Sansom, C. E., Beck, M. E., and Attwood, T. K. (1995) *Trends Biochem. Sci.* **20**, 498–499

[e] Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1990) *Proteins* **8**, 44–61

[f] Spinelli, S., Ramoni, R., Grolli, S., Bonicel, J., Cambillau, C., and Tagoni, M. (1998) *Biochemistry* **37**, 7913–7918

[g] Holden, H. M., Rypniewski, N. R., Law, J. H., and Rayment, I. (1987) *EMBO J.* **6**, 1565–1570

[h] Huber, R., Schneider, M., Mayr, I., Müller, R., Deutzmann, R., Suter, F., Zuber, H., Falk, H., and Kayser, H. (1987) *J. Mol. Biol.* **198**, 499–513

[i] Åkerström, B., and Lögdberg, L. (1990) *Trends Biochem. Sci.* **15**, 240–243

[j] Bishop, R. E., and Weiner, J. H. (1996) *Trends Biochem. Sci.* **21**, 127

[k] Beuckmann, C. T., Aoyagi, M., Okazaki, I., Hiroike, T., Toh, H., Hayaishi, O., and Urade, Y. (1999) *Biochemistry* **38**, 8006 – 8013

[l] Sacchettini, J. C., and Gordon, J. I. (1993) *J. Biol. Chem.* **268**, 18399–18402

[m] Wiesner, S., Kurian, E., Prendergast, F. G., and Halle, B. (1999) *J. Mol. Biol.* **286**, 233–246

[n] Thompson, J., Winter, N., Terwey, D., Bratt, J., and Banaszak, L. (1997) *J. Biol. Chem.* **272**, 7140–7150

[o] Hohoff, C., Börchers, T., Rüstow, B., Spener, F., and van Tilbeurgh, H. (1999) *Biochemistry* **38**, 12229–12239

[p] Thompson, J. R., Bratt, J. M., and Banaszak, L. J. (1995) *J. Mol. Biol.* **252**, 433–446

[q] Heinemann, B., Andersen, K. V., Nielsen, P. R., Bech, L. M., and Poulsen, F. M. (1996) *Protein Sci.* **5**, 13–23

[r] Andersen, K. V., and Poulsen, F. M. (1992) *J. Mol. Biol.* **226**, 1131–1141

[s] Boissy, G., O'Donohue, M., Gaudemer, O., Perez, V., Pernollet, J.-C., and Brunie, S. (1999) *Protein Sci.* **8**, 1191–1199

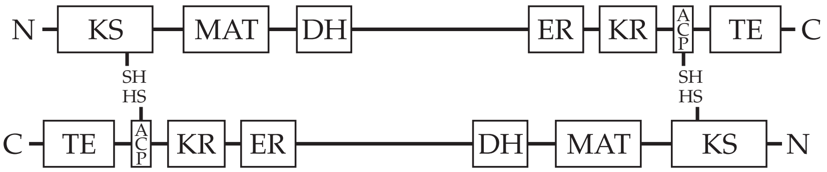
[t] Wirtz, K. W. A. (1991) *Ann. Rev. Biochem.* **60**, 73–99

[u] Tall, A. (1995) *Ann. Rev. Biochem.* **64**, 235–257

[v] Beamer, L. J., Carroll, S. F., and Eisenberg, D. (1997) *Science* **276**, 1861–1864

the ACP domain of one chain would cooperate with the β-oxacyl synthase (KS) domain of the second chain. However, more recent studies indicate greater flexibility with the ACP, MAS, and KS domains of a single chain also being able to function together.[58,62] Animal fatty acid synthases produce free fatty acids, principally the C_{16} palmitate. The final step is cleavage of the acyl-CoA by a thioesterase, one of the six enzymatic activities of the synthase.

Yeast fatty acid synthase[66,67] has an $\alpha_6\beta_6$ structure where the 208-kDa α subunit contains the ACP-like site, the active site –SH, and three catalytic activities. The 220-kDa β subunit has five catalytic activities. The yeast enzyme contains the FMN thought to act as $FMNH_2$ in the second reduction step. As in bacteria, the products of the complex are molecules of acyl-CoA of chain lengths C_{14}, C_{16}, and C_{18}.[68]



Abbreviation	Enzymatic activity	Residue numbers
KS	β-Oxoacyl (ketoacyl) synthase	1 – 406
MAT	Malonyl and acetyl transferase	428 – 815
DH	Dehydratase	829 – 969
Central region	Structural core (?)	970 – 1629
ER	Enoyl reductase	1630 – 1850
KR	β-Oxoacyl (ketoacyl) reductase	1870 – 2100
ACP	Acyl carrier protein	2114 – 2190
TE	Thioesterase	2200 – 2505

Organization of eukaryotic fatty acid synthase. From Joshi *et al.*[61]

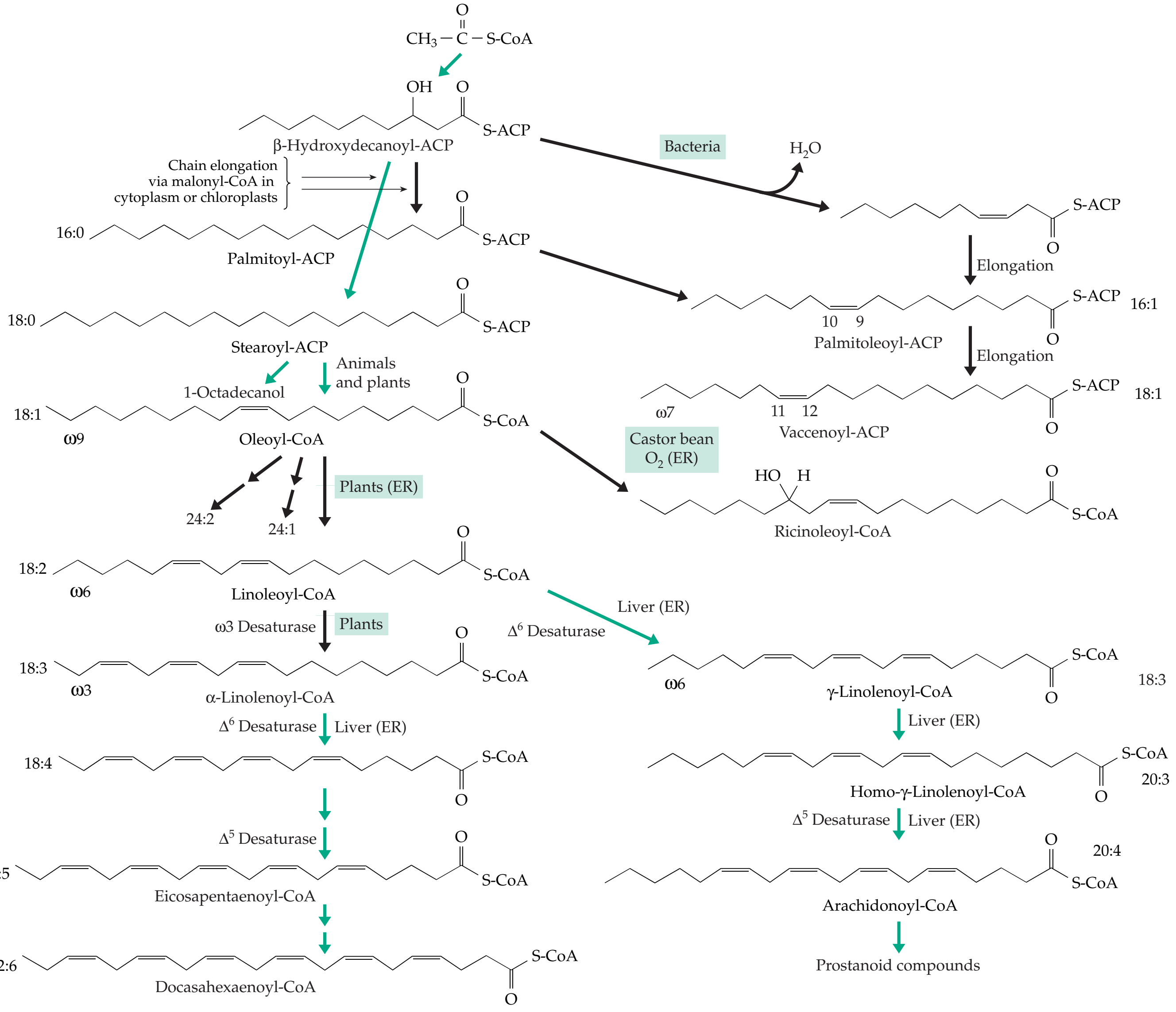

O
CH3 – C – S-CoA
OH O
S-ACP
β-Hydroxydecanoyl-ACP
Chain elongation via malonyl-CoA in cytoplasm or chloroplasts
Bacteria
H2O
S-ACP
O
16:0
S-ACP
Palmitoyl-ACP
Elongation
S-ACP
16:1
10 9
Palmitoleoyl-ACP
18:0
S-ACP
Stearoyl-ACP
Elongation
1-Octadecanol
Animals and plants
S-ACP
18:1
ω7
11 12
Vaccenoyl-ACP
18:1
S-CoA
ω9
Oleoyl-CoA
Castor bean
O2 (ER)
HO H
S-CoA
Ricinoleoyl-CoA
Plants (ER)
24:2
24:1
18:2
S-CoA
ω6
Linoleoyl-CoA
Liver (ER)
ω3 Desaturase
Plants
Δ6 Desaturase
18:3
S-CoA
ω3
α-Linolenoyl-CoA
S-CoA
ω6
γ-Linolenoyl-CoA
18:3
Δ6 Desaturase
Liver (ER)
18:4
S-CoA
Liver (ER)
S-CoA
20:3
Homo-γ-Linolenoyl-CoA
Δ5 Desaturase
Liver (ER)
Δ5 Desaturase
20:5
S-CoA
Eicosapentaenoyl-CoA
20:4
S-CoA
Arachidonoyl-CoA
22:6
S-CoA
Docasahexaenoyl-CoA
Prostanoid compounds

Figure 21-2 Some biosynthetic reactions of fatty acids. Green arrows indicate transformations carried out by the human body.

2. Control of Chain Length

The length of fatty acid chains is controlled largely by the enzymatic activity that releases the fatty acyl-CoA molecules or the free fatty acids from the synthase complex. In the animal enzymes the thioesterase, which is built into the synthase molecule, favors the release primarily of the 16-carbon saturated (16:0) palmitic acid. However, in mammary glands and in the uropygial glands (preen glands) of waterfowl shorter chain fatty acids predominate. These are released from the synthase by reaction with a second thioesterase, a 29-kDa protein[69,70] that catalyzes the otherwise premature release of shorter fatty acids. Cow's milk contains significant amounts of C_4-C_{14} acids as well as those with longer chains, whereas rabbit's milk contains largely C_8 to C_{10} fatty acids.[71]

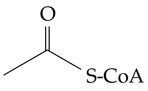
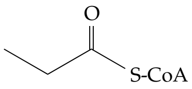
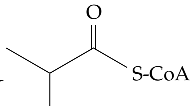
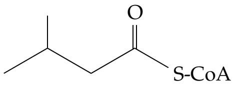
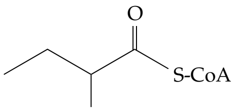
In plants most biosynthesis occurs in the chloroplasts or in the protoplastids of seeds.[72–75] There are two different synthase systems in chloroplasts, one that forms primarily the 16:0 palmitoyl-ACP and the other the 18:0 stearoyl-ACP. Hydrolysis of the palmitoyl-ACP releases palmitate, one major product of chloroplasts. However, the stearoyl-ACP is desaturated to oleoyl-ACP[75a] before hydrolysis to free oleate or conversion to oleoyl-CoA. In many species oleic acid is almost the sole fatty acid exported by the chloroplasts. However, it undergoes a variety of modification reactions in the plant cytosol.

Plants, animals, and fungi all have fatty acid elongation systems in the endoplasmic reticulum. Using malonyl-CoA and NADPH,[76,77] chain lengths of fatty acids may be increased to C_{20} to C_{26}. Elongation of fatty acids can also occur in mitochondria by reactions that are essentially the reverse of β oxidation. The only deviation from an exact reversal of oxidation is the use of NADPH as the reductant for enoyl-CoA reductase. Elongation of fatty acids in the *outer* membrane of mitochondria, followed by transport of the elongated chains into the mitochondria, may even constitute another shuttle for transport of reducing equivalents from NADH into mitochondria (Chapter 18, Section D).[78] Elongation reactions may also occur in peroxisomes.[78a]

3. Starter Pieces and Branches

Acetyl-CoA is most often the primer or starter piece for fatty acid synthesis, but butyryl-CoA is a better primer for rabbits. Butyryl-CoA arises from acetyl-CoA by a reversal of β oxidation, the necessary enzymes occurring in significant amounts in the cytosol.[79] If either acetyl-CoA or butyryl-CoA is the starter piece, chain elongation via malonyl-CoA (Fig. 18-12) leads to fatty acids with an even number of carbon atoms. However, degradation of the branched chain amino acids valine, isoleucine, and leucine creates a series of branched starter pieces (Table 21-3), whose utilization leads to formation of branched fatty acids of the iso and anteiso series. These are found in bacteria, in the lipids of tobacco and wool, in the "sound lens" of echo-locating porpoises,[80] and in many other materials.[81] Propionyl-CoA serves as an intermediate for introduction (via methylmalonyl-CoA) of branches at various other points in a fatty acid chain.[82] For example, 2*R*- and 4*R*-methylhexanoic acids, 2,4,6,8-tetramethyldecanoic acid, and a variety of other branched chain acids are esterified with long- chain alcohols (mainly 1-octadecanol) to form the waxes of the preen glands of ducks and geese.[83] The C_{32} **mycocerosic acid** of *Mycobacterium tuberculosis* is also formed using both malonyl-CoA and methylmalonyl-CoA for chain elongation.[84] This acid is present in mycobacterial cell walls esterified with long-chain diols (Box 21-C).[85]

TABLE 21-3
Starter Pieces for Biosynthesis of Fatty Acids

Starter piece	Fatty acid products
 Acetyl-CoA	Acid with even number of carbon atoms
 Propionyl-CoA	Acid with odd number of carbon atoms
Valine →  Isobutyryl-CoA	Iso series (even)
Leucine →  Isovaleryl-CoA	Iso series (odd)
Isoleucine → 	Anteiso series (odd)

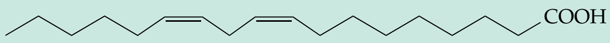
4. Synthesis by the Oxoacid Chain Elongation Process

The carbon skeleton of leucine is derived from that of valine by elongation of the corresponding oxoacid by a single carbon atom that is derived from acetyl-CoA (Fig. 17-18). At least in plants some branched-chain fatty acids of medium length are formed via the same oxoacid elongation process, which extends the chain one carbon atom at a time using 2-oxobutyrate as a starting compound. The same process can be used to form medium length straight-chain fatty acids (up to ~C_{12}) with either an odd or an even number of carbon atoms.[86]

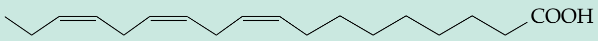
2-Oxoglutarate can also serve as a starter piece for elongation by the oxoacid pathway. Extension by three carbon atoms yields 2-**oxosuberate** (Eq. 21-1). This dicarboxylate is converted by reactions shown in Eq. 24-39 into biotin and in archaebacteria into the coenzyme 7-mercaptoheptanoylthreonine phosphate (HTP), Eq. 21-1.[87] Lipoic acid is also synthesized from a fatty acid, the eight-carbon octanoate.[88,89] A fatty acid synthase system that utilizes a mitochondrial ACP may have as its primary function the synthesis of octanoate for lipoic acid formation.[90] The mechanism of insertion of the two sulfur atoms to form lipoate (Chapter 15) is uncertain. It requires an iron–sulfur protein[91,91a,b] and is probably similar to the corresponding process in the synthesis of biotin (Eq. 24-39)[92-93a] and in formation of HTP (Eq. 21-1). One component of the archaebacterial cofactor **methanofuran** (Chapter 15) is a tetracarboxylic acid that is formed from 2-oxoglutarate by successive condensations with two malonic acid units as in fatty acid synthesis.[94]

BOX 21-B THE ESSENTIAL FATTY ACIDS

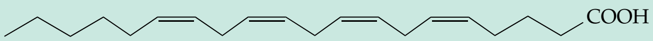
In 1930, George and Mildred Burr reported that the C18:2 $(\Delta^{9,12})$ **linoleic acid**, a fatty acid of exclusively plant origin, cured a disease condition observed in rats raised on a highly purified fat-free diet.[a,b] These animals grew poorly, developed a scaly dermatitis, and suffered kidney damage and impaired fertility. The symptoms could be prevented if 1% of the dietary energy was provided by linoleic acid. This C18:2 fatty acid can be converted in animals into a series of other fatty acids by chain elongation and desaturation. All of this series have a double bond six carbon atoms from the $-CH_3$ terminus and form an ω6 (or n-6) family.[c]



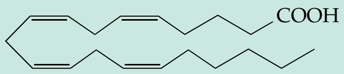
C18:2, ω6 Linoleic acid



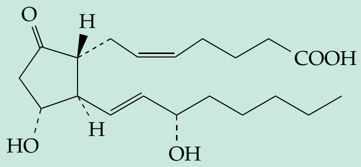
C18:3, ω3 α-Linolenic acid



C20:4, ω6 Arachidonic acid



Arachidonic acid – folded



Prostaglandin E_2

$$\begin{array}{lccccccccccc} \omega6: & C18{:}2 & \rightarrow & C18{:}3 & \rightarrow & C20{:}3 & \rightarrow & C20{:}4 & \rightarrow & C22{:}4 & \rightarrow & C22{:}5\\ & & & & & & & \text{Arachidonate} & & & & \\ & & & & & \downarrow & & \downarrow & & & & \\ & & & & & PGE_1, & & PGE_2, & & & & \\ & & & & & PGF_{1\alpha} & & PGF_{2\alpha} & & & & \end{array}$$

The major known essential function of linoleic acid is conversion to the C20:4 $(\Delta^{5,8,11,14})$ **arachidonic acid**, the major precursor to prostaglandins and other prostanoid compounds (Section D). This conversion occurs in infants as well as adults,[d] but the rate may not always be adequate, and arachidonic acid is usually classified as an essential fatty acid. It is not clear whether linoleic acid has any essential role of its own. However, while arachidonic acid can be converted into the prostaglandins designated PGE_2 and $PGF_{2\alpha}$, linoleic acid can also give rise, via the C20:30 **dihomolinolenic acid**, to PGE_1 and $PGF_{1\alpha}$ (see Eq. 21-16). The C18:3 $(\Delta^{9,12,15})$ **α-linolenic acid**, another plant acid, can partially replace linoleic acid and can be converted into PGE_3 and $PGF_{3\alpha}$.

$$\begin{array}{lccccccccccc} \omega3: & C18{:}3 & \rightarrow & C18{:}4 & \rightarrow & C20{:}4 & \rightarrow & C20{:}5 & \rightarrow & C22{:}5 & \rightarrow & C22{:}6\\ & & & & & & & \downarrow & & \downarrow & & \downarrow\\ & & & & & & & PGE_3, & & PG_4 & & PG_5\\ & & & & & & & PGE_{3\alpha} & & & & \end{array}$$

Thus, it is not surprising that the three acids are not completely equivalent.[e–g]

Recent interest has focused on the C20:5 **eicosapentaenoic acid** (EPA) and the C22:6 **docosahexaenoic acids** (DHA). These ω3 (or n-3) polyunsaturated acids are formed from linolenic acid by marine algae and are found in fish oils.[h] The C22:5 and C22:6 acids can be converted to prostaglandins of the PG_4 and PG_5 series. DHA together with the ω6 C22:4 acid constitutes over 30% of the fatty acids in brain phospholipids. In the retina DHA accounts for over 60% of the total in the rod outer segments (Fig. 23-40). DHA may be formed in the human body from α-linolenic acid obtained from plant sources (Fig. 21-2). However, the rate of synthesis may be inadequate especially in infants and in persons of old age.[i–m] Deficiency of either EPA or DHA may lead to poor brain development during prenatal and infant life. Formation of new synapses between neurons as well as growth of new neurons in some areas of the brain is associated with thinking and with memory formation (Chapter 30). Therefore, a lack of the essential ω3 and ω6 fatty acids may contribute to mental deterioration in older adults. Eskimo populations, which consume large amounts of fish, have a very low incidence of coronary heart disease. An inverse relation between fish consumption and heart disease has also been demonstrated in other populations.[n] Inclusion of fish oils to 20–30% of total caloric intake in the diet causes a marked decrease in plasma triacylglycerols and very low density lipoproteins (VLDL).[o] This effect has been attributed to the altered composition of the prostanoid compounds known as thromboxanes and prostacyclins (PGI). For similar reasons the ω3 acids may have an antiinflammatory effect.[p,q] Administration of fish oil to patients with kidney disease has proved beneficial[r] and may decrease risk of some cancers.[s] However, a diet high in ω3 fatty acids has also been reported to increase cancer risk.[f] Long-chain ω3 fatty acids may protect against sudden death from heart disease.[t] They may promote lateral phase separation within membranes to form regions low in cholesterol (see references 95a and 119d).

BOX 21-B (continued)

[a] Burr, G. O., and Burr, M. M. (1930) *J. Biol. Chem.* **86**, 587–621

[b] Burr, G. (1980) *Trends Biochem. Sci.* **5**, 28

[c] Horrobin, D. F., ed. (1990) *Omega-6 Essential Fatty Acids*, Wiley-Liss, New York

[d] Salem, N., JR, Wegher, B., Mena, P., and Uauy, R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 49–54

[e] Leat, W. M. F. (1981) *Trends Biochem. Sci.* **6**, IX–X

[f] Cave, W. T., Jr. (1991) *FASEB J.* **5**, 2160–2166

[g] Lands, W. E. M. (1992) *FASEB J.* **6**, 2530–2536

[h] Lees, R. S., and Karel, M., eds. (1990) *Omega-3 Fatty Acids in Health and Disease*, Dekker, New York

[i] Cho, H. P., Nakamura, M., and Clarke, S. D. (1999) *J. Biol. Chem.* **274**, 37335–37339

[j] Farkas, T., Kitajka, K., Fodor, E., Csengeri, I., Lahdes, E., Yeo, Y. K., Krasznai, Z., and Halver, J. E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6362–6366

[k] Qiu, X., Hong, H., and MacKenzie, S. L. (2001) *J. Biol. Chem.* **276**, 31561–31566

[l] Kang, Z. B., Ge, Y., Chen, Z., Cluette-Brown, J., Laposata, M., Leaf, A., and Kang, J. X. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4050–4054

[m] Carper, J. (2000) *Your Miracle Brain*, Harper Collins Publ., New York

[n] Kromhout, D., Bosschieter, E. B., and Coulander, C. L. (1985) *N. Engl. J. Med.* **312**, 1205–1209

[o] Phillipson, B. E., Rothrock, D. W., Connor, W. E., Harris, W. S., and Illingworth, D. R. (1985) *N. Engl. J. Med.* **312**, 1210–1216

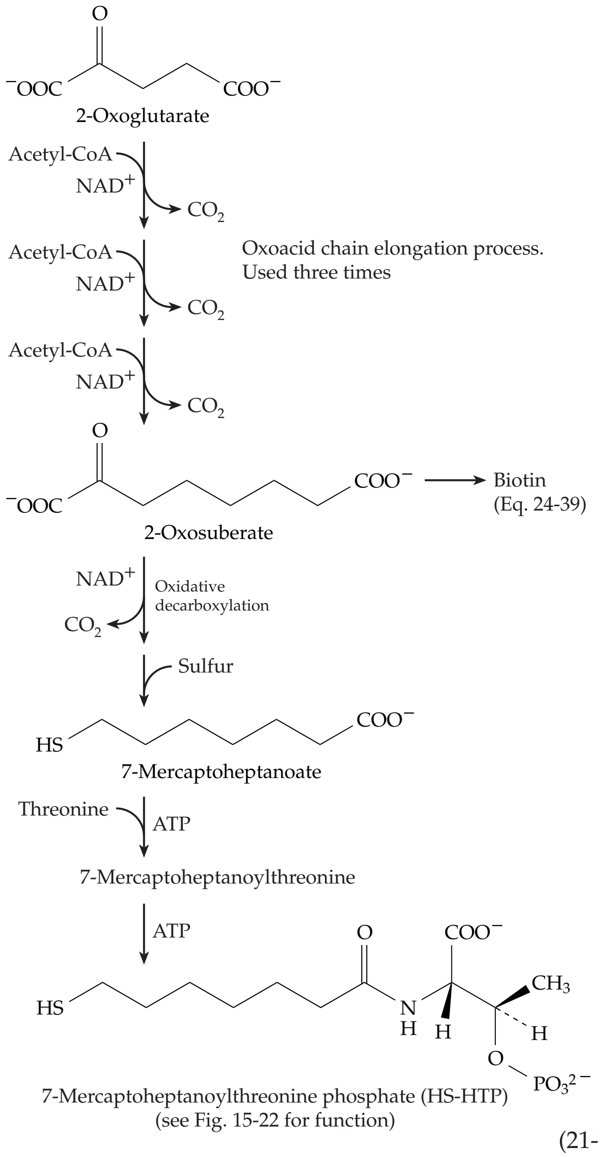
[p] Lee, T. H., Hoover, R. L., Williams, J. D., Sperling, R. I., Ravalese, J., III, Spur, B. W., Robinson, D. W., Corey, E. J., Lewis, R. A., and Austen, K. F. (1985) *N. Engl. J. Med.* **312**, 1217–1224

[q] Hwang, D. (1989) *FASEB J.* **3**, 2052–2061

[r] Donadio, J. V., Jr., Bergstralh, E. J., Offord, K. P., Spencer, D. C., and Holley, K. E. (1994) *N. Engl. J. Med.* **331**, 1194–1199

[s] Hilakivi-Clarke, L., Clarke, R., Onojafe, I., Raygada, M., Cho, E., and Lippman, M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9372–9377

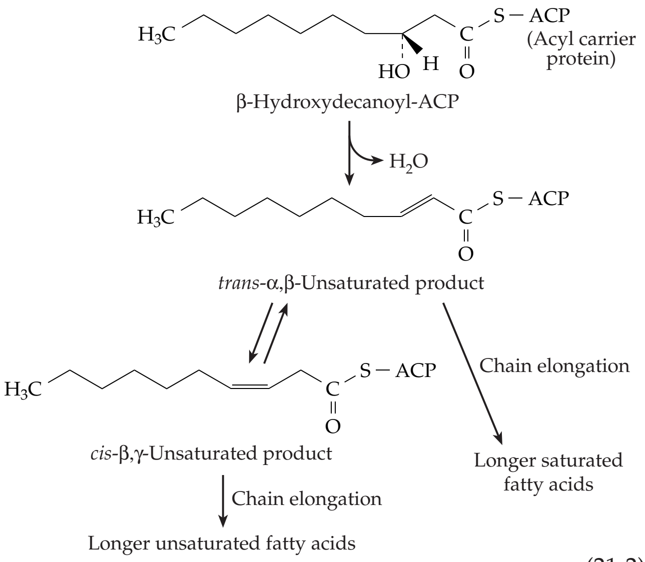
[t] Albert, C. M., Campos, H., Stampfer, M. J., Ridker, P. M., Manson, J. E., Willett, W. C., Ma, J. (2002) *N. Engl. J. Med.* **346**, 1113 – 1118 See also **347**, 531 – 533


2-Oxoglutarate
Acetyl-CoA
NAD+
CO2
Oxoacid chain elongation process. Used three times
2-Oxosuberate
Biotin (Eq. 24-39)
Oxidative decarboxylation
Sulfur
7-Mercaptoheptanoate
Threonine
ATP
7-Mercaptoheptanoylthreonine
7-Mercaptoheptanoylthreonine phosphate (HS-HTP) (see Fig. 15-22 for function)

(21-1)

5. Unsaturated Fatty Acids

Fatty acids containing one or more double bonds provide necessary fluidity to cell membranes[95,95a] and serve as precursors to other components of cells. Significant differences in the methods of introduction of double bonds into fatty acids are observed among various organisms. Bacteria such as *E. coli* that can live anaerobically often form **vaccenic acid** as the principal unsaturated fatty acid. It is formed by chain elongation after introduction of a cis double bond at the C_{10} stage of synthesis. The bacteria possess a **β-hydroxydecanoyl thioester dehydratase,** which catalyzes elimination of a β-hydroxyl group to yield primarily the *cis*-β,γ rather than the *trans*-α,β-unsaturated product (Eq. 21-2).[96] The mechanism may resem-


β-Hydroxydecanoyl-ACP
(Acyl carrier protein)
H2O
trans-α,β-Unsaturated product
cis-β,γ-Unsaturated product
Chain elongation
Longer saturated fatty acids
Chain elongation
Longer unsaturated fatty acids

(21-2)

ble that of enoyl hydratase (Eq. 13-7), the indicated *trans*-α,β-unsaturated intermediate (enzyme bound) being isomerized to the *cis*-β,γ-unsaturated product through an allylic rearrangement. The product can then be elongated to the C_{16} **palmitoleoyl-ACP** and C_{18} **vaccenoyl-ACP** derivatives (Fig. 21-2, right side). However, dehydration of β-hydroxydecanoyl-ACP lies at a branch point in the biosynthetic sequences. The *trans*-α,β-unsaturated fatty acyl compound lies on the usual route of chain elongation to palmitoyl-CoA (left side, Fig. 21-2).

In higher plants, animals, protozoa, and fungi, saturated fatty acids are acted upon by **desaturases** to introduce double bonds, usually of the cis (Z) configuration. The substrates may be fatty acyl-ACP, fatty acyl-CoA molecules, membrane phospholipids,[97] or glycolipids.[98] The Δ^9 desaturase, isolated from liver or from yeast, converts stearoyl-CoA to oleoyl-CoA (Eq. 21-3).[99–102] This membrane-associated enzyme system utilizes NADH as a reductant, passing electrons via cytochrome b_5 reductase and cytochrome b_5 itself to the desaturase.[103,104] The pro-*R* hydrogens are removed at both C-9 and C-10.

$$O_2 + 2\ H^+ + 2\ e^- + \text{stearoyl-CoA} \rightarrow \text{oleoyl-CoA} + 2\ H_2O \qquad (21\text{-}3)$$

In plants a similar enzyme catalyzes formation of the first double bond in a fatty acyl group converting stearoyl-ACP into oleoyl-ACP in the chloroplasts.[72,75a,105–108] The soluble Δ^9 stearoyl-ACP desaturase has a diiron-oxo active site (Fig. 16-20, B, C).[109,110] Electrons are donated from light-generated reduced ferredoxin (see Chapter 23). In addition to the Δ^9 desaturase both plants and cyanobacteria usually desaturate C_{18} acids also at the Δ^{12} and Δ^{15} positions and C_{16} acids at the Δ^7, Δ^{10}, and Δ^{13} (ω3) positions.[111,112] Desaturation of oleate occurs primari-

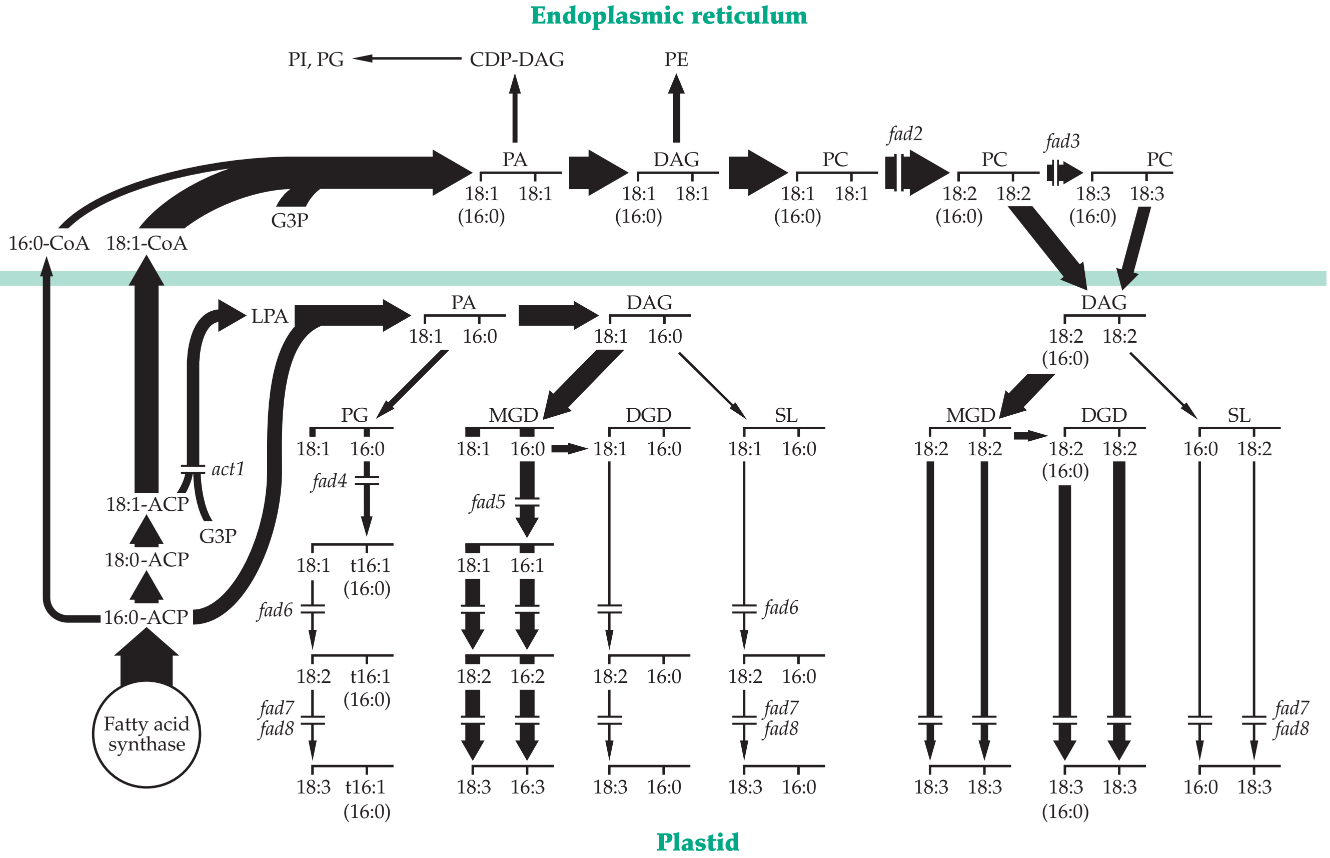

Endoplasmic reticulum
PI, PG
CDP-DAG
PE
PA
18:1 18:1
(16:0)
DAG
18:1 18:1
(16:0)
PC
18:1 18:1
(16:0)
fad2
PC
18:2 18:2
(16:0)
fad3
PC
18:3 18:3
(16:0)
G3P
16:0-CoA
18:1-CoA
LPA
PA
18:1 16:0
DAG
18:1 16:0
DAG
18:2 18:2
(16:0)
PG
18:1 16:0
fad4
18:1 t16:1
(16:0)
fad6
18:2 t16:1
(16:0)
fad7
fad8
18:3 t16:1
(16:0)
MGD
18:1 16:0
fad5
18:1 16:1
18:2 16:2
18:3 16:3
DGD
18:1 16:0
18:2 16:0
18:3 16:0
SL
18:1 16:0
fad6
18:2 16:0
fad7
fad8
18:3 16:0
MGD
18:2 18:2
18:3 18:3
DGD
18:2 18:2
(16:0)
18:3 18:3
(16:0)
SL
16:0 18:2
fad7
fad8
16:0 18:3
act1
18:1-ACP
G3P
18:0-ACP
16:0-ACP
Fatty acid synthase
Plastid

Figure 21-3 Major pathways of synthesis of fatty acids and glycerolipids in the green plant *Arabidopsis*. The major site of fatty acid synthesis is chloroplasts. Most is exported to the cytosol as oleic acid (18:1). After conversion to its coenzyme A derivative it is converted to phosphatidic acid (PA), diacylglycerol (DAG), and the phospholipids: phosphatidylcholine (PC), phosphatidylinositol (PI), phosphatidylglycerol (PG), and phosphatidylethanolamine (PE). Desaturation also occurs, and some linoleic and linolenic acids are returned to the chloroplasts. See text also. From Sommerville and Browse.[106] See also Figs. 21-4 and 21-5. Other abbreviations: monogalactosyldiacylglycerol (MGD), digalactosyldiacylglycerol (DGD), sulfolipid (SL), glycerol 3-phosphate (G3P), lysophosphatidic acid (LPA), acyl carrier protein (ACP), cytidine diphosphate-DAG (CDP-DAG).

ly in the ER after conversion of the free acid to its coenzyme A derivative or to phosphatidylcholine. Consecutive introduction of two double bonds forms linoleoyl-CoA (18:2, Δ^9,^{12}) and linolenoyl-CoA (18:3, Δ^9,^{12},^{15}; see Fig. 21-3).[108] All double bonds are cis. The membrane lipids of chloroplasts contain both linoleic and linolenic acids, which have apparently been returned to the chloroplasts from the cytosol[72] as indicated in Fig. 21-3. Plants grown at colder temperatures have a higher content of these trienoic acids than those grown at higher temperature.[72a]

The origin of ricinoleic acid, an abundant constitutuent of castor beans, is also shown in Fig. 21-2. It is formed by an **oleate hydroxylase** that has an amino acid sequence similar to those of oleate desaturases.[113] Both hydroxylation and desaturation are reactions catalyzed by diiron centers.[114] Other fatty acid hydroxylases act on the alpha[115] and the omega positions. The latter are members of the cytochrome P450 family.[116,117]

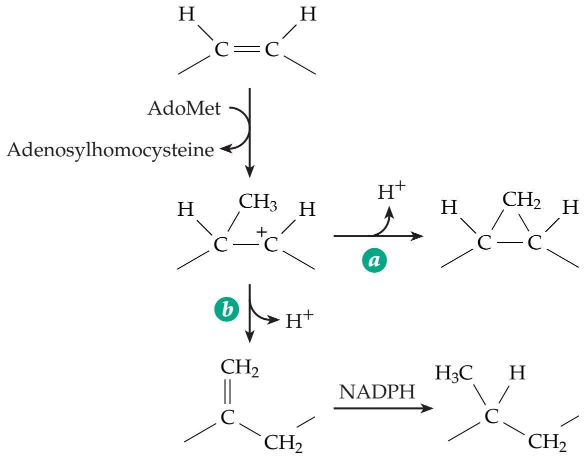
The conversion of oleoyl-CoA to linoleoyl-CoA is accomplished by some insects[118] but does not take place in most animals. As a result of this biosynthetic deficiency, polyunsaturated fatty acids such as linoleic, linolenic, and the C_{20} arachidonic acid are necessary in the diet (Box 21-B). One essential function of linoleic acid is to serve as a precursor of **prostaglandins** and related **prostanoids** (Section D). Dietary linoleate is converted to its CoA derivative and then by sequential Δ^6 desaturation,[119] elongation, and then Δ^5 desaturation, to the 20:4 ($\Delta^{5,8,11,14}$) arachidonoyl-CoA (Fig. 21-2, lower right). These acids are referred to as ω6 because of the position of the last double bond. Linolenic acid can be converted in an analogous fashion to the CoA derivative of the 20:5 ($\Delta^{5,8,11,14,17}$ ω6) eicosapentaenoic acid (**EPA**). The 22:6 docasahexaenoic acid (**DHA**; Fig. 21-2) is apparently formed by elongation of the 22:5 acyl-CoA to 24:5, desaturation, transfer to a peroxisome or mitochondrion, and β oxidation to shorten the chain.[95a]

These acids are very important in human nutrition.[119a–d] (See also Box 21-B.) In the absence of adequate essential fatty acids oleate is desaturated and elongated in a similar sequence to the unusual 20:3 ($\Delta^{5,8,11}$ ω9) acid. Vertebrate tissues also carry out desaturation at the Δ^4 and Δ^3 positions.[111] Lepidoptera, which synthesize a great diversity of pheromones, are rich in unusual desaturases such as the gland-specific acyl-CoA Δ^{11} desaturase of cabbage looper moths.[120]

Polyunsaturated fatty acids, containing 4, 5, or 6 double bonds and chain lengths of up to C_{36}, are found in phosphatidylcholine of vertebrate retinas.[121] Although the double bonds are rarely in conjugated positions in food fats and in animal bodies, some plants convert oleic or linoleic acid into fatty acids with as many as three or four conjugated double bonds.[121a] **Conjugated linoleic acid** (9*c*, 11*t* 18:2) can be formed from 11-*trans*-octadecenoate in the human body.[121b] This compound is also found in meat and dairy products. It has been reported to have anticancer properties[121c] and may be another beneficial dietary constituent. An isomerase isolated from red algae converts polyunsaturated acids into forms with conjugated double bonds. For example, arachidonic acid (5*Z*, 8*Z*, 11*A*, 14*Z*) is converted to (5*Z*, 7*E*, 9*E*, 14*Z*)-eicosatetraenoic acid.[122,123]

6. Cyclopropane Fatty Acids and Mycolic Acids

Fatty acids containing one or more cyclopropane rings are present in many bacteria (p. 381).[124,125] The extra carbon of the cyclopropane ring is added from *S*-adenosylmethionine (AdoMet) at the site of a cis double bond in a fatty acyl group of a phosphatidylethanolamine molecule in a membrane (Eq. 21-4).[126,126a] The same type of intermediate carbocation can yield either a cyclopropane fatty acid (Eq. 21-4, step *a*) or a methenyl fatty acid (Eq. 21-4, step *b*). The latter can be reduced to a branched fatty acid. This is an alternative way of introducing methyl branches that is used by some bacteria.[127]



(21-4)

Mycobacteria are rich in cyclopropane-containing fatty acids. These **mycolic acids** are major components of the cell walls and may account for 30% of the dry weight of the cells.[128] The most abundant mycolic acid of *M. tuberculosis* consists of C_{52} fatty acid containing two cyclopropane rings joined via a Claisen-type condensation with a C_{26} carboxylate fatty acid (Eq. 21-5). A similar mycolic acid formed by *M. smegmatis* has double bonds instead of cyclopropane rings as indicated below Eq. 21-5.[128,129] There are other variations. In place of a double bond or cyclopropane group there may be –OH, $-OCH_3$, C=O, epoxide, or CH_3[127]

Cyclopropane fatty acids are catabolized via β oxidation,[130] which is modified as in Eq. 21-6 when the chain degradation reaches the cyclopropane ring. The

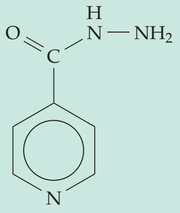
BOX 21-C TUBERCULOSIS

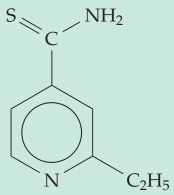
As many as one-third of the inhabitants of the earth are infected by *Mycobacterium tuberculosis*. For most the infection is dormant, but in some the slow-growing bacteria cause a progressive and deadly destruction of the lungs. There are still about three million deaths annually, and *M. tuberculosis* now, as in the past, kills more people then any other pathogen.[a,b] The development of drug-resistant strains of the bacterium and the threat of a world-wide resurgence of tuberculosis[c,d] has spurred new efforts to understand the unusual metabolism of mycobacteria and to develop new drugs. The complete 4.4 million base-pair sequence of the circular genome is known.[c] An unusually large fraction of the genes encode enzymes involved in synthesis and breakdown of lipids, including the synthesis of **mycolic acids** (Eq. 21-5) which are characteristic of mycobacteria.

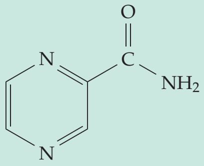
The mycobacterial cell wall, discussed in Chapter 8, contains mycolic acids bound covalently at the nonreducing ends of arabinogalactans that are attached to the inner peptidoglycan layer,[e] as well as phosphatidylinositol-anchored lipoarabinomannans. Other unusual lipids that are also present and account for some of the difficulty of treatment with antibiotics include esters of **mycocerosic acid** with long-chain diols known as phenolphthiocerols and phthiocerols.[f,fg]

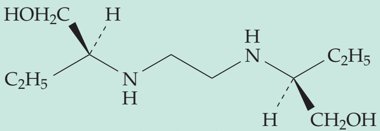
Streptomycin (Boxes 20-B, 20-H) was introduced into clinical use against tuberculosis in about 1943. However, resistant mutants always survived until newer drugs were developed. Isonicotinylhydrazide (**isoniazid**) is especially effective in combinations with suitable antibiotics and other drugs.[g] The four-drug combination isoniazid, rifampicin (Box 28-A), pyrazinamide, and ethambutol is often used. Nevertheless, bacteria resistant to all of these have developed.

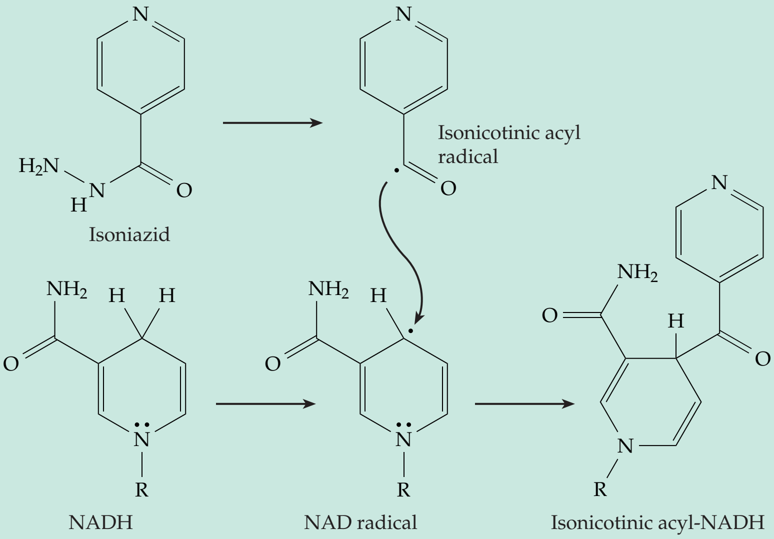
Although isoniazid has been in use for about 45 years, the enzyme that it inhibits has been recognized only recently. It is a specific NADH-dependent **enoyl reductase** involved in synthesis of mycolic acids.[h,i] The isoniazid must be activated by action of a bacterial catalase-peroxidase.[j,k] This enzyme may convert the drug to a reactive radical that combines with a NADH-derived radical to form an adduct in the active site of the enzymes. One possible reaction sequence follows.[h] However, the mechanisms are not clear.

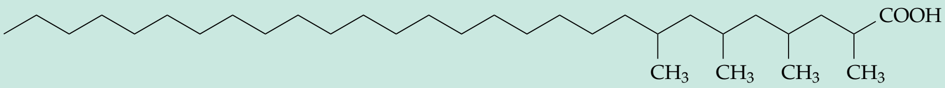

Isonicotinylhydrazide (Isoniazid)


Ethionamide


Pyrazinamide


Ethambutol




2*R*,4*R*,6*R*,8*R*-Tetramethyloctacosanoic acid (mycocerosic acid)

BOX 21-C (continued)

[a] Young, D. B. (1998) *Nature (London)* **393**, 515–516

[b] Venisse, A., Rivière, M., Vercauteren, J., and Puzo, G. (1995) *J. Biol. Chem.* **270**, 15012–15021

[c] Cole, S. T., and and 41 other authors. (1998) *Nature (London)* **393**, 537–544

[d] Iseman, M. D. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2428–2429

[e] Scherman, M. S., Kalbe-Bournonville, L., Bush, D., Xin, Y., Deng, L., and McNeil, M. (1996) *J. Biol. Chem.* **271**, 29652–29658

[f] Fitzmaurice, A. M., and Kolattukudy, P. E. (1998) *J. Biol. Chem.* **273**, 8033–8039

[fg] Patterson, J. H., McConville, M. J., Haites, R. E., Coppel, R. L., and Billman-Jacobe, H. (2000) *J. Biol. Chem.* **275**, 24900–24906

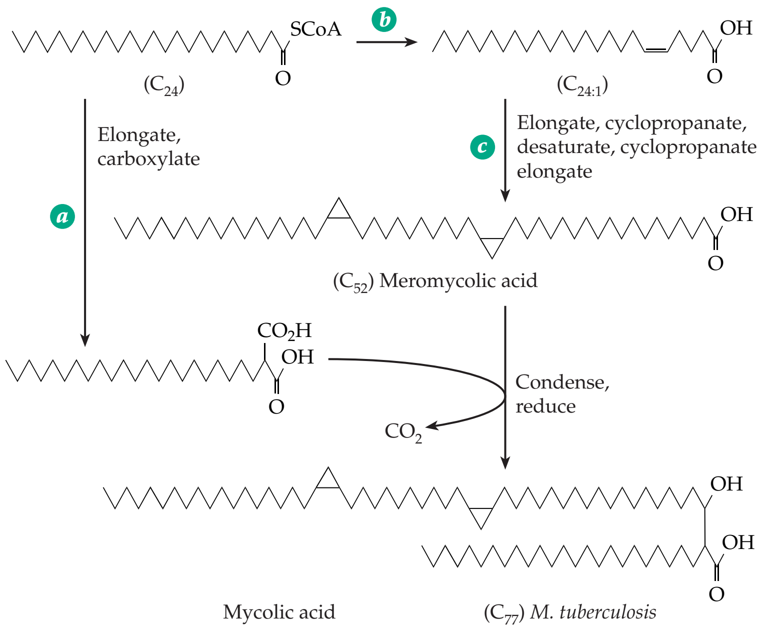
[g] Blanchard, J. S. (1996) *Ann. Rev. Biochem.* **65**, 215–239

[h] Rozwarski, D. A., Grant, G. A., Barton, D. H. R., Jacobs, W. R. J., and Sacchettini, J. C. (1998) *Science* **279**, 98–102

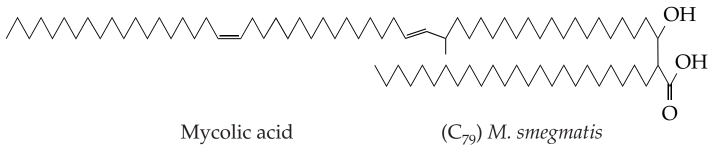
[i] Baldock, C., Rafferty, J. B., Stuitje, A. R., Slabas, A. R., and Rice, D. W. (1998) *J. Mol. Biol.* **284**, 1529–1546

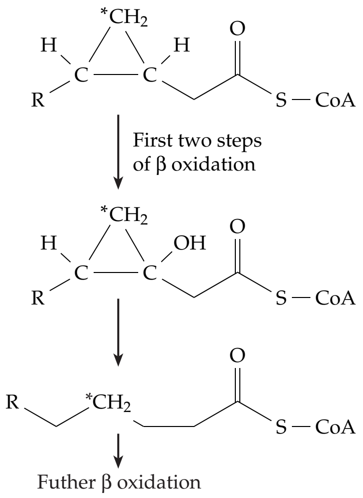
[j] Sherman, D. R., Mdluli, K., Hickey, M. J., Arain, T. M., Morris, S. L., Barry, C. E., III, and Stover, C. K. (1996) *Science* **272**, 1641–1643

[k] Wengenack, N. L., Lopes, H., Kennedy, M. J., Tavares, P., Pereira, A. S., Moura, I., Moura, J. J. G., and Rusnak, F. (2000) *Biochemistry* **39**, 11508–11513


SCoA
b
OH
(C24)
O
(C24:1)
O
Elongate, carboxylate
a
c
Elongate, cyclopropanate, desaturate, cyclopropanate elongate
OH
O
(C52) Meromycolic acid
CO2H
OH
O
Condense, reduce
CO2
OH
OH
O
Mycolic acid
(C77) M. tuberculosis

(21-5)

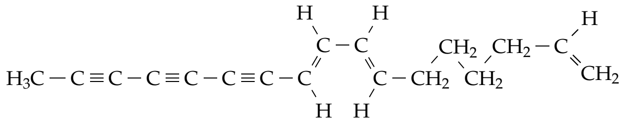

OH
OH
O
Mycolic acid
(C79) M. smegmatis


*CH2
H
H
O
R
C
C
S − CoA
First two steps of β oxidation
*CH2
H
OH
O
R
C
C
S − CoA
O
R
*CH2
S − CoA
Futher β oxidation

(21-6)

ring opening of the cyclopropanol derivatives occurs readily, even with mild nonenzymatic acid–base catalysis.

Another alteration of unsaturated fatty acids is the formation of acetylenic groups (–C≡C–). This apparently occurs by dehydrogenation of –CH=CH–. Examples of naturally occurring acetylenes are **crepenynic acid** (p. 381), **alloxanthin** (p. 1240), and the following remarkable hydrocarbon from the common cornflower *Centaurea cyanus*[131]:


H3C − C≡C − C≡C − C≡C − C
H
H
C − C
H
H
C − CH2
CH2
CH2
CH2 − C
H
CH2

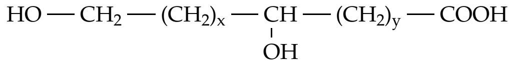
7. The Lipids of Skin and Other Surfaces

Special fatty materials are often secreted to form external surfaces of organisms.[81,132] An example already mentioned is the secretion of the uropygial glands (preen glands) of water fowl. In the goose 90% of this material is a wax consisting of monoesters of various acids with predominantly **1-octadecanol** as the long-chain fatty alcohol.[81] The latter is formed by reduction of stearoyl-CoA as indicated in Fig. 21-2. Waxes are also important constituents of marine environments, where they are not limited to surfaces. For example, copepods, which constitute a major component of marine zooplankton, may contain up to 70% of their dry weight as wax esters. Some marine animals, such as sperm whales, accumulate the same esters in major amounts as energy stores.[133]

Among the compounds present in the lipids of human skin are a variety of branched fatty acids, both free and combined. They may play a role in maintaining the ecological balance among microorganisms of the skin, and they also impart to each individual a distinct odor or "chemical fingerprint."[132] Some of the skin lipids are incorporated into the cornified outer skin surface (Box 8-F).[134] See also Section C, 3.

Surface lipids of plants. The thick cuticle (Fig. 1-6) that covers the outer surfaces of green plants consists largely of waxes and other lipids but also contains a complex polymeric matrix of **cutin** (stems and leaves) or **suberin** (roots and wound surfaces).[135,135a] Plant waxes commonly have $C_{10} - C_{30}$ chains in both acid and alcohol components. Methyl branches are frequently present. A major function of the waxes is to inhibit evaporation of water and to protect the outer cell layer. In addition, the methyl branched components may inhibit enzymatic breakdown by microbes. Free fatty acids, free alcohols, aldehydes, ketones, β-diketones, and alkanes are also present in plant surface waxes. Chain lengths are usually $C_{20} - C_{35}$.[136] Hydrocarbon formation can occur in other parts of a plant as well as in the cuticle. Thus, normal **heptane** constitutes up to 98% of the volatile portion of the turpentine of *Pinus jeffreyi*.[81]

Cutin is largely a polyester with a high content of ω-hydroxypalmitic acid and related fatty acids, which are also hydroxylated at a second position:



$$HO-CH_2-(CH_2)_x-\underset{\displaystyle OH}{\underset{|}{CH}}-(CH_2)_y-COOH$$

Cutin monomers. C_{16} acids in which y = 8, 7, 6, or 5 and x + y = 13

This allows branching of the polymer. Monomers of other chain lengths as well as aromatic components related to lignin are also present and polymerized into a high molecular mass branched structure. Suberin is a more complex ligninlike polymer with a high content of phenolic constituents[135] such as vanillin (Fig. 25-8).

Formation of hydrocarbons. Alkanes and alkenes occur in plants, in preen gland secretions, and in insects. The alkanes of plant cuticle are thought to be formed by elongation of a C_{16} acid followed by loss of the carboxyl group. The mechanisms are not obvious. However, these hydrocarbons are often two carbon atoms shorter than the starting fatty acid. The pathway between them might begin by α-oxidation to form an α-peroxy acid which would decarboxylate to form an aldehyde, a reaction similar to that of Eq. 15-36. Alternatively, a long-chain acyl-CoA may be reduced directly to an aldehyde. In fact, when suitable inhibitors are present aldehydes do accumulate in tissues that are forming hydrocarbons.[137] Conversion of an aldehyde intermediate to an alkane may occur by **decarbonylation** (loss of CO). This has been demonstrated in pea (*Pisum sativum*) leaves,[138] in uropygial glands,[139] in flies, and in a colonial green alga, *Botyrococcus braunii*.[138] In the last case 32% of the dry weight of the cells is C_{27}, C_{29}, and C_{31} hydrocarbons. They appear to be formed by action of a decarbonylase that apparently contains a cobalt porphyrin.[137] Plants require cobalt for growth, but an enzymatic function has not previously been established.

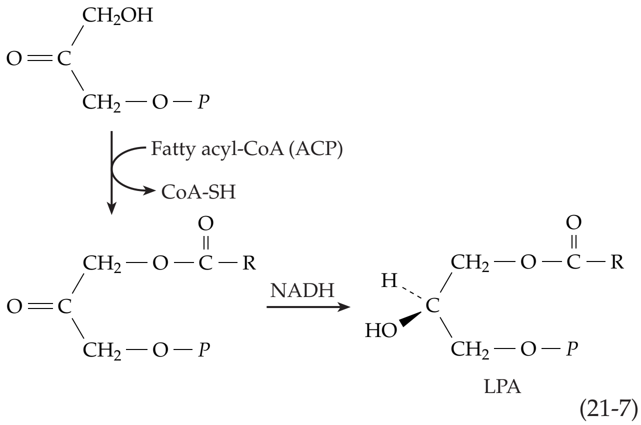
In contrast, the sex pheromone of the female housefly is (Z)-9-tricosene, a hydrocarbon apparently formed by an oxidative decarboxylative process from a precursor aldehyde by an enzyme that requires NADPH and O_2 and is apparently a cytochrome P450.[140] Oxidative deformylation by a cytochrome P450 converts aldehydes to alkenes, presumably via a peroxo intermediate.[117] Formation of an alkene by decarboxylation has also been proposed,[141] but a mechanism is not obvious.

Insect waxes, hydrocarbons, and pheromones. Surface lipids of many insects contain esters of long-chain (as long as 66 carbon atoms) alcohols and long-chain acids.[142] On the other hand, waxes of the tobacco hornworm consist largely of 11- and 12-oxo derivatives of a C_{28} alcohol, which may be esterified to short-chain acids.[142] A major hydrocarbon of cockroaches is 6,9-heptacosadiene.[81]

Insects communicate through the use of a great variety of volatile pheromones. As mentioned in Chapter 8, Section A,1, some moths utilize acetate esters of various isomers of Δ^7 and Δ^{11} unsaturated C_{14} fatty acids as sex pheromones. Some other moths convert the *trans*-11-tetradecenyl acetate into the corresponding C_{14} aldehyde or alcohol, while others use similar compounds of shorter ($C_{11} - C_{12}$) chain length.[143] Some ants use ketones, such as 4-methyl-3-heptanone, as well as various isoprenoid compounds and pyrazines as volatile signaling compounds.[144] Other insects also utilize isoprenoids,[145] alkaloids,[146] and aromatic substances as pheromones.

C. Synthesis of Triacylglycerols, Glycolipids, and Phospholipids

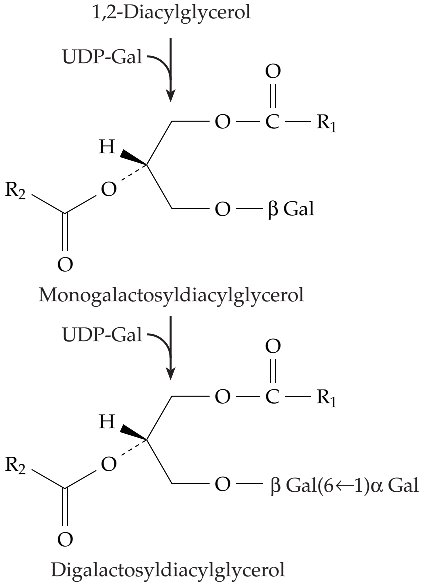
Reduction of dihydroxyacetone phosphate yields *sn*-glycerol 3-phosphate, the starting compound for formation of the glycerol-containing lipids (Fig. 21-4 step *a*).[146a,b] Transfer of two acyl groups from ACP or CoA to the hydroxyl groups of this compound (steps *b* and *c*) yields 1,2-diacylglycerol 3-phosphate (phosphatidic acid). Two different acyltransferases are required.[147] Unsaturated fatty acids are incorporated preferentially into the 2-position. The intermediate 1-acyl-*sn*-glycero-3-phosphate, often called **lysophosphatidic acid** (LPA), is formed in excess in activated platelets and has a variety of signaling activities.[148,149] LPA for signaling is derived by turnover of existing phospholipids. An alternative route of LPA formation in liver is the transfer of one acyl group onto dihydroxyacetone phosphate and reduction prior to addition of the second acyl group (Eq. 21-7).



(21-7)

Phosphatidic acid lies at a metabolic branch point. On the one hand, the phospho group can be removed by a specific phosphatase (step *d*)[150] and another acyl group (most often an unsaturated acyl group) may be transferred onto the resulting diacylglycerol (DAG, diglyceride, step *e*)[150a,b] to form a **triacylglycerol** (triglyceride). Alternatively, the phosphatidic acid may be converted to a **CDP-diacylglycerol** (step *g*), a key intermediate in phospholipid synthesis both in eukaryotes and in bacteria.[151] Not only can phosphatidic acid be hydrolyzed to 1,2-diacylglycerols, but the reverse process can occur by action of a kinase. This presumably permits recycling of the diacylglycerol formed by turnover of membrane phospholipids.[152]

Diacylglycerols can also be converted to a variety of glycolipids such as the **galactolipids** of chloroplasts (Eq. 21-8). See also Chapter 8. These are the major lipids of photosynthetic membranes.[98,153–155] Some bacteria, e.g., the mycoplasma *Acholeplasma laidlawii*, contain both monoglucosyl- and diglucosyl-DAG. Changes in the ratio of these two membrane components may regulate the phase equilibrium between bilayer and nonbilayer forms.[156] 1,2-Diacylglycerol can also react with UDP-sulfoquinovose (Eq. 20-12) to form the characteristic sulfolipid of chloroplasts.[154]



(21-8)

In animals a principal regulatory point for lipid synthesis is in the activation of acetyl-CoA carboxylase by citrate (Fig. 17-20).[156a,b] Beyond that, a complex hormonal control is exerted on both biosynthesis and the catabolism of triglycerides stored in liver and adipose tissues.[157] For example, adrenaline and glucagon, by stimulating production of cAMP, stimulate acetyl-CoA carboxylase,[158] activate lipases that cleave triacylglycerols, and mobilize depot fats.[159] Insulin, on the other hand, promotes lipid storage. It increases the activity of the enzymes of lipogenesis from the ATP-dependent citrate cleavage enzyme (Eq. 13-39) and inhibits cAMP production, thus blocking lipolysis within cells. At the transcriptional level sterols bind to activator proteins (**sterol regulatory element binding proteins,** SREBPs) and activate genes for acetyl-CoA carboxylase[158] and for stearoyl-CoA desaturases.[160] Fatty acid synthases, which play a central role in lipid formation, are controlled by both hormonal and nutritional factors at the transcriptional[161] and translational[162] levels. Environmental factors also have indirect effects. For example, the Δ^9 fatty acid desaturase activity of poikilothermic (cold-blooded) animals is increased at low temperatures. The resulting increased synthesis of unsaturated fats leads to increased fluidity of the membrane bilayer.[163] As mentioned on p. 1193; the same is true for green plants.

1. Phospholipids

Bacterial and also some eukaryotic phospholipids are formed following conversion (Fig. 21-4, step *g*) of phosphatidic acids to CDP-diacylglycerols, which are able to react with a variety of nucleophiles with displacement of CMP.[164–166] Reaction with L-serine (step *h*)[167] leads to **phosphatidylserine,** and reaction with glycerol 3-phosphate (step *i*),[168] which enters cells via a special transporter,[168a] produces **phosphatidyglycerol 3-*P*.** The enzyme catalyzing the formation of phosphatidylserine appears to occur naturally as an integral membrane protein of the ER. Some is also bound to ribosomes and to mitochondria.[169,169a] In contrast, most of the other enzymes of phospholipid formation are closely associated with or embedded in the cytoplasmic membrane. One of these, a pyruvoyl group dependent enzyme (Chapter 14, Section F), catalyzes decarboxylation of phosphatidylserine to **phosphatidylethanolamine** (PE, step *j*, Fig. 21-4).[170] This reaction had been thought unimportant in animals, but

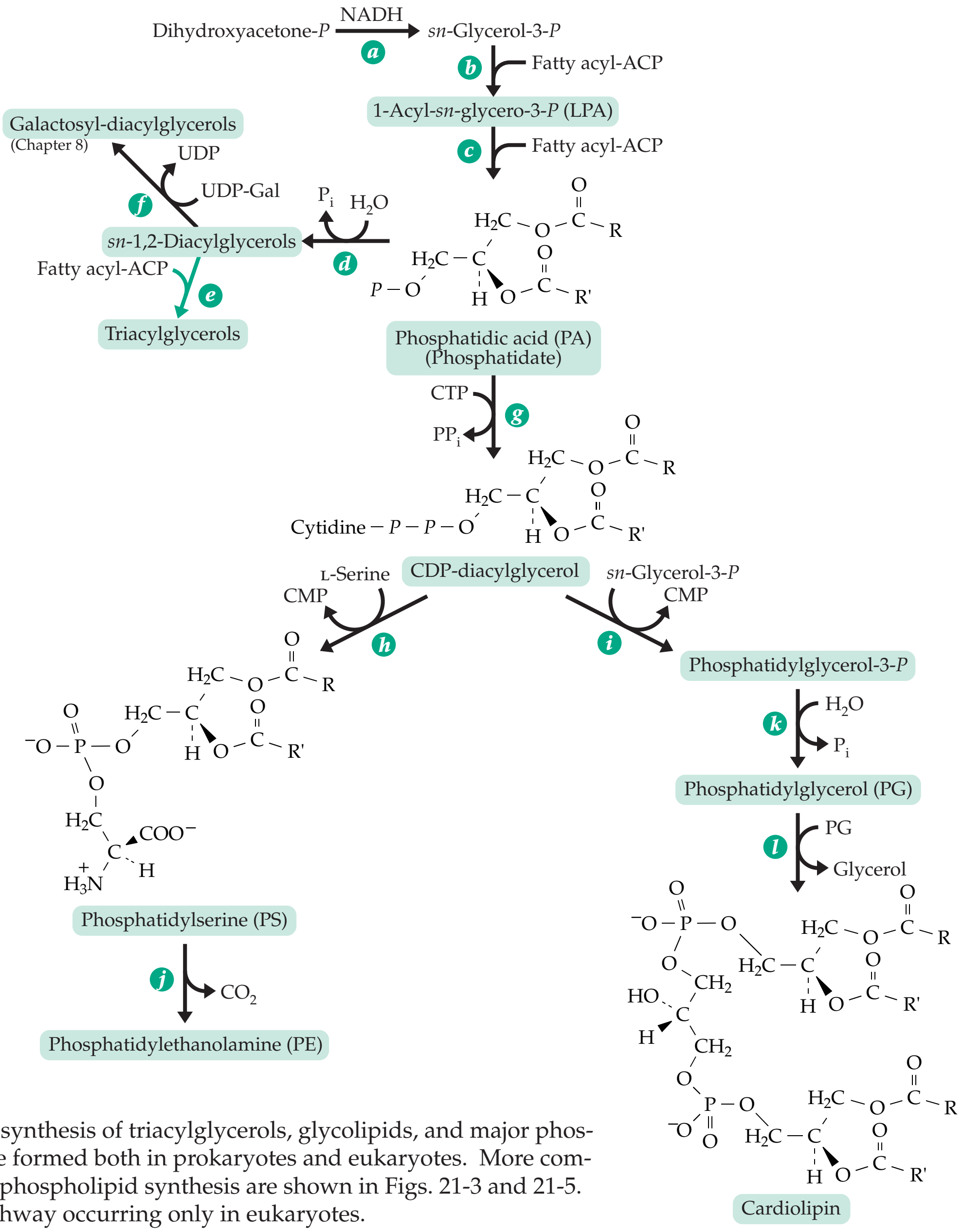

Dihydroxyacetone-*P*
NADH
a
sn-Glycerol-3-*P*
b
Fatty acyl-ACP
1-Acyl-*sn*-glycero-3-*P* (LPA)
c
Fatty acyl-ACP
Galactosyl-diacylglycerols
(Chapter 8)
UDP
UDP-Gal
f
P_i H_2O
sn-1,2-Diacylglycerols
d
Fatty acyl-ACP
e
Triacylglycerols
Phosphatidic acid (PA)
(Phosphatidate)
CTP
PP_i
g
Cytidine – *P* – *P* – O
L-Serine
CMP
CDP-diacylglycerol
sn-Glycerol-3-*P*
CMP
h
i
Phosphatidylglycerol-3-*P*
H_2O
k
P_i
Phosphatidylglycerol (PG)
PG
l
Glycerol
Phosphatidylserine (PS)
j
CO_2
Phosphatidylethanolamine (PE)
Cardiolipin

Figure 21-4 Biosynthesis of triacylglycerols, glycolipids, and major phospholipids that are formed both in prokaryotes and eukaryotes. More complete schemes of phospholipid synthesis are shown in Figs. 21-3 and 21-5. Green arrow: pathway occurring only in eukaryotes.

results with cultured cells show that decarboxylation of phosphatidylserine is often the major route of formation of phosphatidylethanolamine in mammalian cells.[171,172] This phospholipid also accounts for 75% of total phospholipid of the *E. coli* cell envelope. It is synthesized on the cytosolic side of the inner membrane, but it is also translocated to the outer membrane, where it is a major constituent of the inner bilayer leaflet.[173] PE is essential for viability of *E. coli* cells.[174] It provides dipolar ionic head groups and apparently serves as a chaperone for folding of some membrane proteins.[175]

After removal of a phosphate from phosphatidylglycerol 3-*P*, the resulting phosphatidylglycerol can be converted to **diphosphatidylglycerol** (known as **cardiolipin**). One manner in which this is accomplished in bacteria is indicated by step *l* of Fig. 21-4. One molecule of glycerol is displaced as two molecules of phosphatidylglycerol are coupled. The alternative pathway of Eq. 21-9 is followed in eukaryotic mitochondria and perhaps in some bacteria. The entire phosphatidic acid group is transferred from CDP-diacylglycerol to phosphatidylglycerol with displacement of CMP.[176–178] Gram-negative bacteria also synthesize a second set of membrane phospholipids, compounds such as **lipid A** (Figs. 8-30, 20-10) that are based on acylated glucosamine.[165]

Phosphatidylcholine, which is rarely present in bacteria, is formed in eukaryotes from phosphatidylethanolamine by three consecutive steps of methylation

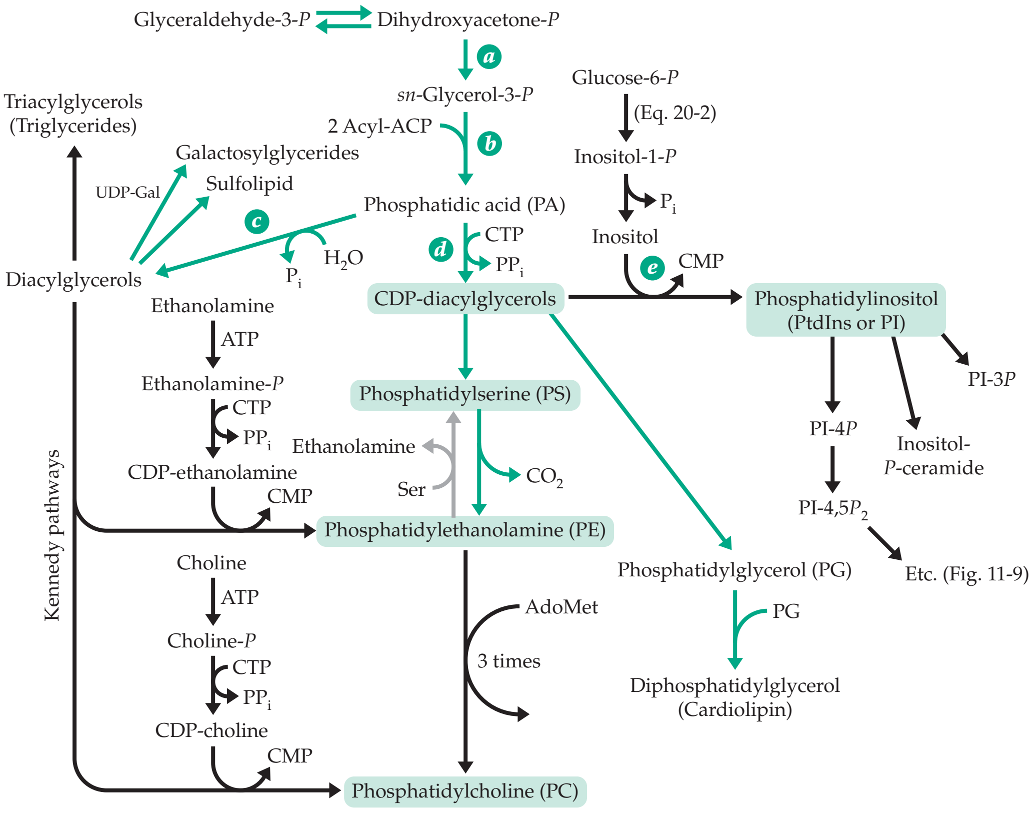
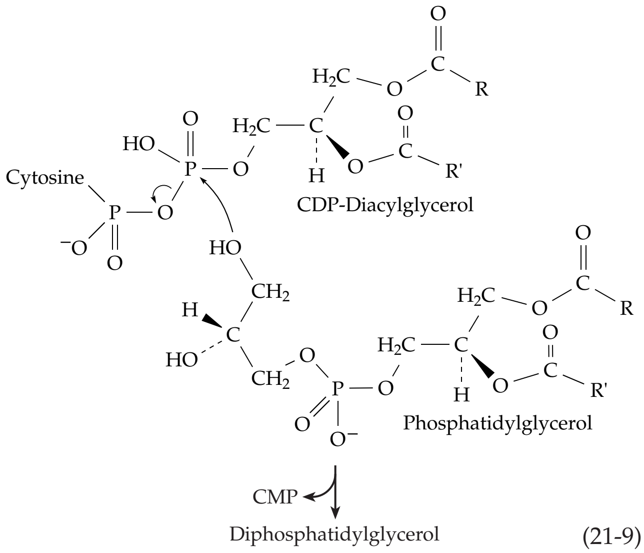

Glyceraldehyde-3-P
Dihydroxyacetone-P
a
sn-Glycerol-3-P
2 Acyl-ACP
b
Glucose-6-P
(Eq. 20-2)
Inositol-1-P
P_i
Inositol
Triacylglycerols (Triglycerides)
Galactosylglycerides
Sulfolipid
UDP-Gal
Phosphatidic acid (PA)
c
H_2O
P_i
d
CTP
PP_i
e
CMP
Diacylglycerols
Ethanolamine
ATP
Ethanolamine-P
CTP
PP_i
CDP-ethanolamine
CMP
CDP-diacylglycerols
Phosphatidylinositol (PtdIns or PI)
PI-3P
PI-4P
Inositol-P-ceramide
PI-4,5P_2
Etc. (Fig. 11-9)
Phosphatidylserine (PS)
Ethanolamine
Ser
CO_2
Phosphatidylethanolamine (PE)
Phosphatidylglycerol (PG)
PG
Diphosphatidylglycerol (Cardiolipin)
Kennedy pathways
Choline
ATP
Choline-P
CTP
PP_i
CDP-choline
CMP
AdoMet
3 times
Phosphatidylcholine (PC)

Figure 21-5 A more complete outline of the biosynthesis of triacylglycerols, glycolipids, and phospholipids including characteristic eukaryotic pathways. Green lines indicate pathways utilized by both bacteria and eukaryotes. Structures of some of the compounds are shown in Fig. 21-4. The gray arrows show the formation of phosphatidylserine by exchange with ethanolamine (Eq. 21-10).

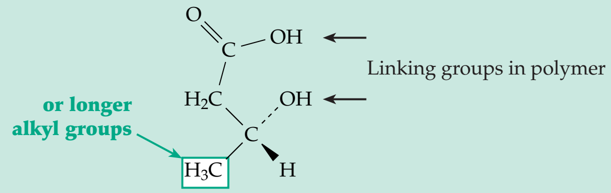

Cytosine
CDP-Diacylglycerol
Phosphatidylglycerol
CMP
Diphosphatidylglycerol

(21-9)

by S-adenosylmethionine (Fig. 21-5). This pathway is of major importance in eukaryotic cells.[179,180] However, alternative pathways (the Kennedy pathways),[166] which are represented by black lines on the left side of Fig. 21-5, are also used for formation of both phosphatidylcholine and phosphatidylethanolamine. In both cases, the free base, choline, or ethanolamine[180a,b] is phosphorylated with ATP. Choline phosphate formed in this manner is then converted by reaction with CTP to CDP-choline (Eq. 17-58).[181] Phosphatidylcholine is formed from this intermediate[181a,b] while CDP-ethanolamine is used to form phosphatidylethanolamine (Fig. 21-5). These synthetic reactions occur within cell nuclei as well as on surfaces of cytoplasmic membranes.[181c]

The formation of phosphatidylserine and possibly other phospholipids in animal tissues may also be accomplished by exchange reactions (Eq. 21-10, step *a*).[182,183] At the same time, decarboxylation of phosphatidylserine back to phosphatidylethanolamine (Eq. 21-10, step *b*) also takes place, the net effect being a catalytic cycle for decarboxylation of serine to ethanolamine. The latter can react with CTP to initiate synthesis of new phospholipid molecules or can be converted to phosphatidylcholine (step *c*). However, unless there is an excess of methionine and folate in the diet, choline is an essential human nutrient.[184]

BOX 21-D POLY-β-HYDROXYBUTYRATE AND BIODEGRADABLE PLASTICS



D-β-Hydroxybutyric acid

The important bacterial storage material polyhydroxybutyric acid is related metabolically and structurally to the lipids. This highly reduced polymer is made up of D-β-hydroxybutyric acid units in ester linkage, about 1500 residues being present per chain. The structure is that of a compact right-handed coil with a twofold screw axis and a pitch of 0.60 nm.[a] Within bacteria it often occurs in thin lamellae ~5.0 nm thick. Since a chain of 1500 residues stretches to 440 nm, there must be ~88 folds in a single chain. Present in both cytoplasmic granules and in membranes,[b] polyhydroxybutyrate can account for as much as 50% of the total carbon of some bacteria.[c] In *E. coli* and many other bacteria polyhydroxybutyrate is present in a lower molecular mass form bound to calcium polyphosphates, proteins, or other macromolecules.[d,e] It has also been extracted from bovine serum albumin and may be ubiquitous in both eukaryotes and prokaryotes.[d,e] The polymer may function in formation of Ca^{2+} channels in membranes.[b,d]

Biosynthesis occurs from 3-hydroxybutyryl-CoA. Some bacteria incorporate other β-hydroxyacids into the polymer.[f] Apparently various hydroxyacyl-CoAs can be diverted from the β oxidation pathway to polymer synthesis,[g] and synthases that will accept a variety of β-hydroxyacyl-CoA substrates have been isolated.[h,i] More than 80 different hydroxyacyl groups can be incorporated into the polymer.[i] A bacterially produced copolymer of β-hydroxybutyrate and β-hydroxyvalerate resembles polypropylene but is biodegradeable. It not only can be used for sutures and other medical implants[j] but also could compete with petroleum-derived plastics[h] and be derived from renewable sources. To this end the synthase genes have been cloned, engineered, and transferred into other microorganisms and plants.[k–n] Transgenic cotton plants incorporate polyhydroxybutyrate granules into the cotton fibers altering the properties of the fibers.[m] The polyhydroxybutyrate synthases apper to be related mechanistically to bacterial lipases.[o]

---

[a] Okamura, K., and Marchessault, R. H. (1967) in *Conformation of Biopolymers*, Vol. 2 (Ramachandran, G. N., ed), pp. 709–720, Academic Press, New York

[b] Reusch, R. N., and Sadoff, H. L. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4176–4180

[c] Jacob, G. S., Garbow, J. R., and Schaefer, J. (1986) *J. Biol. Chem.* **261**, 16785–16787

[d] Reusch, R. N., Huang, R., and Bramble, L. L. (1995) *Biophys. J.* **69**, 754–766

[e] Huang, R., and Reusch, R. N. (1996) *J. Biol. Chem.* **271**, 22196–22202

[f] Peoples, O. P., and Sinskey, A. J. (1989) *J. Biol. Chem.* **264**, 15298–15303

[g] de Waard, P., van der Wal, H., Huijberts, G. N. M., and Eggink, G. (1993) *J. Biol. Chem.* **268**, 315–319

[h] Müh, U., Sinskey, A. J., Kirby, D. P., Lane, W. S., and Stubbe, J. (1999) *Biochemistry* **38**, 826–837

[i] Gerngross, T. U., and Martin, D. P. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6279–6283

[j] Pool, R. (1989) *Science* **245**, 1187–1189

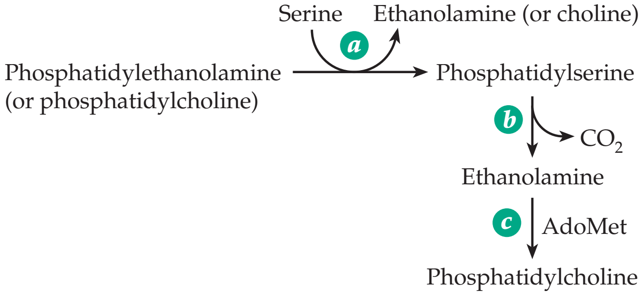
[k] Poirier, Y., Dennis, D. E., Klomparens, K., and Somerville, C. (1992) *Science* **256**, 520–523

[l] Mittendorf, V., Robertson, E. J., Leech, R. M., Krüger, N., Steinbüchel, A., and Poirier, Y. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13397–13402

[m] John, M. E., and Keller, G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12768–12773

[n] García, B., Olivera, E. R., Minambres, B., Fernández-Valverde, M., Canedo, L. M., Prieto, M. A., García, J. L., Martínez, M., and Luengo, J. M. (1999) *J. Biol. Chem.* **274**, 29228–29241.

[o] Crandall, W. V., and Lowe, M. E. (2001) *J. Biol. Chem.* **276**, 12505–12512


Serine
Ethanolamine (or choline)
a
Phosphatidylethanolamine (or phosphatidylcholine)
Phosphatidylserine
b
CO_2
Ethanolamine
c
AdoMet
Phosphatidylcholine

(21-10)

Apparently the synthesis via serine and phosphatidylserine cannot provide an adequate amount of choline, which is present in the body not only in phosphatidylcholine but in plasmalogens, sphingomyelins, and the neurotransmitter **acetylcholine**.[185]

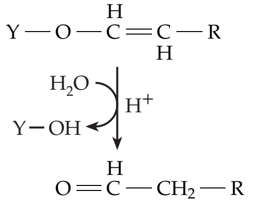
Phosphatidylinositol (PI), a major component of membrane lipids, is formed by displacement of CMP from CMD-dialylglycerol by *myo*-inositol.[186] It is also converted into a variety of less abundant phosphorylated derivatives that engage in signaling activities (see Fig. 11-9). In addition, PI is a component of the glycosylphosphatidylinositol (GPI) membrane anchors for suface proteins (Fig. 8-13). Free GPI anchors, lacking bound proteins, are also present in membranes.

They are especially abundant in many parasitic protozoa and may carry additional glycosyl groups.[186a]

Regulation of phospholipid synthesis, which has been studied in detail in yeast,[187–191] is complex but highly coordinated. The committed step in the synthesis of PE and PC is the hydrolysis of phosphatidate (PA) by a phosphatase to generate diacylglycerols (Fig. 21-5, step *c*). Reaction of PA with CTP (step *d*) also affects synthesis of the other major phospholipids. Much of the coordinate regulation arises at the transcriptional level. For example, genes for the synthesis of PC or PI are repressed by inositol alone and in combination with choline.[187,188] Regulation of CTP synthetase controls the formation of CDP-diacylglycerols.[190] In mammalian cells PC synthesis appears to occur only during the S-phase of the cell cycle (Fig. 11-15).[191] The CTP: phosphocholine cytidyltransferase that catalyzes CDP-choline formation is controlled by storage in a resevoir in the nucleus from which it is transferred to ER membranes.[181b]

2. The Ether-Linked Lipids

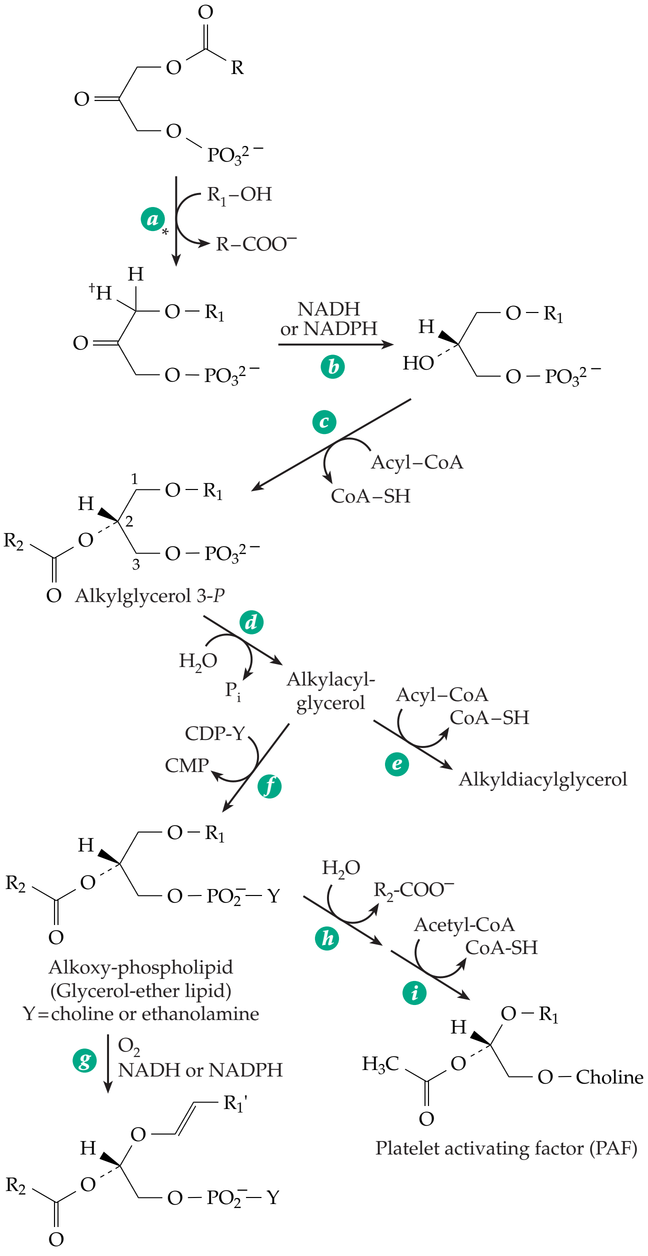
Closely related to both the triacylglycerols and phospholipids, the ether-linked lipids contain in place of one ester group an *alkoxy* (–OR) or *alkenyl* (–O–CH=CH–R) group.[192] Phospholipids containing the alk-1-enyl group, the **plasmalogens**, were first recognized in 1924 by Feulgen and Voit, who were developing histological staining procedures. They observed that treatment of tissue slices with acid resulted in the liberation of aldehydes, which were later shown to be formed by breakdown of the alkenyl lipids (Eq. 21-11). Over 10% of the lipid in the human central nervous system is plasmalogen and about 1% is alkoxy lipid. Among the latter is the **platelet activating factor** (Box 8-A).[193-194a] In many mammalian cells the ethanolamine plasmalogen **plasmenylethanolamine** represents the major storage depot of arachidonic acid.[195]



$$\mathrm{Y{-}O{-}\overset{H}{C}{=}\underset{H}{C}{-}R} \xrightarrow[\mathrm{H_2O}\ \to\ \mathrm{Y{-}OH}]{\mathrm{H^+}} \mathrm{O{=}\overset{H}{C}{-}CH_2{-}R} \qquad (21\text{-}11)$$

Ether-linked lipids constitute up to 35% of the total phospholipid in molluscs. Although they are usually regarded as animal constituents, small amounts of ether-linked lipids have been identified in plants. The major phospholipids of archaebacterial membranes are ether-linked derivatives of the polyprenyl phytanyl group and of the dimeric biphytanyl group (Chapter 8).[196–198]

Biosynthesis of ether lipids begins with formation of fatty acyl derivatives of dihydroxyacetone phosphate. The acyl group is then displaced, along with the oxygen atom to which it is attached, by an alkoxy group of a long-chain fatty alcohol (Eq. 21-12, step *a*), which is formed by reduction of the corresponding acyl-CoA.[199] The oxygen of the alcohol (designated by an asterisk) is retained in the product.[200,201] The reaction differs significantly from displacements discussed in Chapter 12. The pro-*R* hydrogen atom (marked by the dagger, †) at C-1 exchanges with the medium during the reaction suggesting that enolization of the dihydroxyacetone phosphate takes place. A possible mechanism would be to add the incoming R–O⁻, generated as in serine proteases, to the double bond of



Alkylglycerol 3-*P*

Alkoxy-phospholipid (Glycerol-ether lipid) Y = choline or ethanolamine

Platelet activating factor (PAF)

Plasmalogen: Y = ethanolamine

(21-12)

the enol at C-1. This would generate a transient carbanion on C-2. It could then eliminate the carboxylate containing the original C-1 acyl group, and the enol could then ketonize.

Once an alkoxy derivative of dihydroxyacetone is formed, reduction to the 2-OH form, further acylation, and conversion to various alkyl phospholipids and neutral lipids can occur. The pathways (Eq. 21-12, steps *b*–*f*) are closely akin to those of Fig. 21-4. The conversion of alkoxy lipids to plasmalogens occurs by oxidative desaturation (Eq. 21-12, step *f*).[202] The initial steps in the synthesis of ether-linked lipids take place principally in the peroxisomes. Enzymes catalyzing both the acylation of dihydroxyacetone phosphate and the synthesis of alkyl-dihydroxyacetone-*P* (step *a*, Eq. 21-12) are found in high amounts in animal peroxisomes. In the rare autosomal recessive disorder known as the **Zellweger syndrome** peroxisomes are completely lacking.[203] Both the synthesis of ether-linked lipids[204] and the β oxidation of very-long-chain fatty acids are depressed. These acids, principally C26:0 and C26:1, accumulate in tissues[205,206] of patients with this severe disease, which is usually fatal during the first four months of life.

The platelet activating factor (PAF, Box 8-A) is formed in neutrophils and macrophages from alkylacyl-*sn*-glycero-3-phosphocholine by the action of phospholipase A_2. This enzyme removes the C2 acyl group, which is then replaced by an *acetyl group* transferred from acetyl-CoA to form PAF. Alternatively, a phosphocholine group may be transferred onto 1-alkyl-2-acyl-*sn*-glycerol from CDP-choline as in the formation of phosphatidylcholine (Fig. 21-5). PAF can undergo hydrolytic removal of its acetyl group in tissues but can also transfer it to such acceptors as lysoplasmalogens or sphingosine.[194] Hydrolytic loss of the acetyl group from PAF destroys biological activities including induction of allergic and inflammatory responses.[194a] The various signaling activities of PAF arise from binding to G-protein linked receptors in many cells and tissues.[207]

3. Sphingolipids

Sphinoglipids are phospholipids and glycolipids derived from **sphingosine** and other "long-chain bases."[208] At least 60 bases of this type have been identified.[209] They vary in chain length from C_{14} to C_{26} and include members of the iso and anteiso series. Up to two double bonds may be present. The C_{18} compound, usually called sphingosine, is derived from condensation of palmitoyl-CoA with serine.[209a] Carbon dioxide is lost from the serine during the condensation reaction (Fig. 21-6, step *a*; Chapter 14), and the resulting ketone is reduced with NADPH (step *b*) to form **sphinganine**, a common component of animal sphingolipids. It may be hydroxylated to phytosphingosine in plants and fungi (step *c*).[210,211] Sphinganine is converted to long-chain amides by acyl transfer from acyl-CoA (step *d*) and then undergoes desaturation (step *e*)[212–213a] to form **ceramides**, the precursors to more than 100 gangliosides (Fig. 20-11),[214,215] to the phospholipid **sphingomyelins** (Fig. 21-6, step *g*), and also to free sphingosine (step *h*). This last reaction is degradative and on the pathway of breakdown of gangliosides (Fig. 20-11). Further catabolism of sphingosine is thought to take place by a PLP-mediated chain cleavage to palmitaldehyde.[216]

The essential functions of sphingolipids, including the complex gangliosides, are only now being clarified.[215,217,218] The latter are abundant in brain and are thought to function in cell–cell recognition. On blood cell surfaces they carry blood group antigens (Box 4-C). They play an essential role in spermatogenesis[218] and may function in various signaling processes.[218a] In the outer cornified layers of skin, ceramides with very long chain (C_{28}–C_{36}) fatty acyl groups undergo ω hydroxylation (Fig. 21-6) and become esterified to glutamate side chains of specific skin proteins called **involucrins**. The long hydrocarbon chains are thought to pass entirely through the lipid bilayer to form rigid lamellae of a water-impermeable outer skin barrier.[134] An important hypothesis is that sphingolipids associate with cholesterol to form "lipid rafts," which float in a sea of glycerolipids and serve as bases for various signaling processes. The long hydrophobic acyl chains of the sphingolipids pack well with cholesterol to form a rigid lipid structure of high melting temperature.[218b,c]

4. Complex Lipids in Signaling

While pathways of synthesis of complex lipids have been described, we are far from understanding the dynamics of the synthesis and turnover of the membranous structures built from them. The fact that the lipid bilayer of a cell membrane is so thin means that any sudden changes in composition at a particular location will cause changes in physical properties and a wave of diffusion that will travel along the membrane. The membrane seems to be ideally structured to receive and propagate messages from outer surface or internal receptors, or messages sent along the bilayer.

One of the most studied examples of signaling with membrane lipids is provided by the **phosphoinositide cascade**, which is pictured in Fig. 11-9. Six or more phosphate esters of phosphatidylinositol (PI) are generated by the action of kinases.[219,220] More than 100 extracellular signaling molecules activate specific isozyme forms of **phospholipase C**,[221–224] releasing 20 or more different inositol phosphates from these

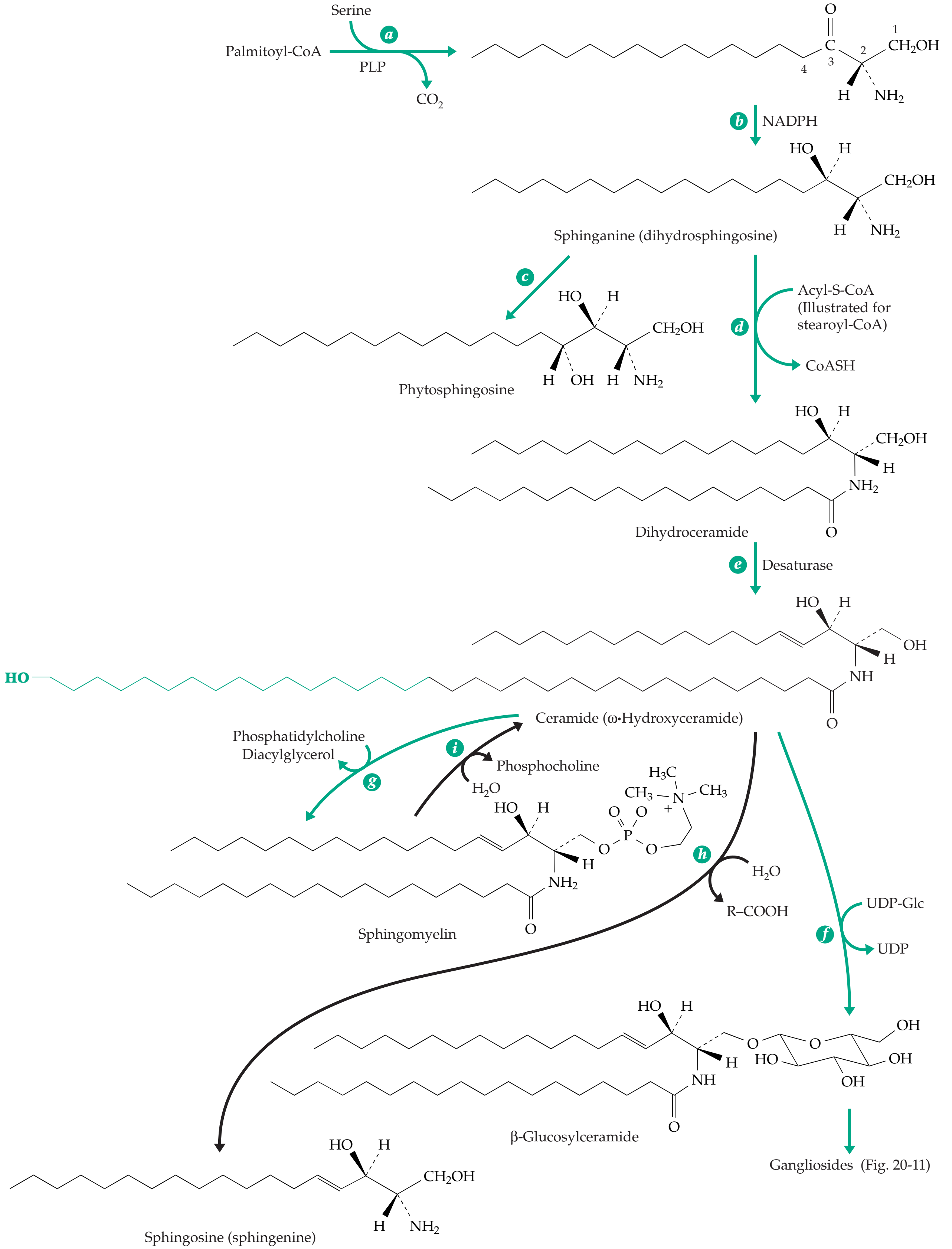

Serine
a
Palmitoyl-CoA
PLP
CO_2
O
C
1
CH_2OH
2
3
4
H
NH_2
b
NADPH
HO
H
CH_2OH
H
NH_2
Sphinganine (dihydrosphingosine)
c
HO
H
CH_2OH
H
OH
H
NH_2
Phytosphingosine
d
Acyl-S-CoA
(Illustrated for
stearoyl-CoA)
CoASH
HO
H
CH_2OH
H
NH_2
O
Dihydroceramide
e
Desaturase
HO
H
OH
H
HO
NH
O
Ceramide (ω-Hydroxyceramide)
Phosphatidylcholine
Diacylglycerol
g
i
Phosphocholine
H_2O
H_3C
CH_3
N
CH_3
+
HO
H
O
O
P
O
O
H
NH_2
O
Sphingomyelin
h
H_2O
R–COOH
UDP-Glc
f
UDP
HO
H
O
O
OH
H
HO
OH
NH
OH
O
β-Glucosylceramide
Gangliosides (Fig. 20-11)
HO
H
CH_2OH
H
NH_2
Sphingosine (sphingenine)

Figure 21-6 Pathways of synthesis and metabolism of sphingolipids. Gray arrows indicate catabolic pathways. See also Fig. 20-11. The green extension on the ceramide structure is that of a long-chain ω-hydroxyceramide that is covalently bound to protein in human skin.

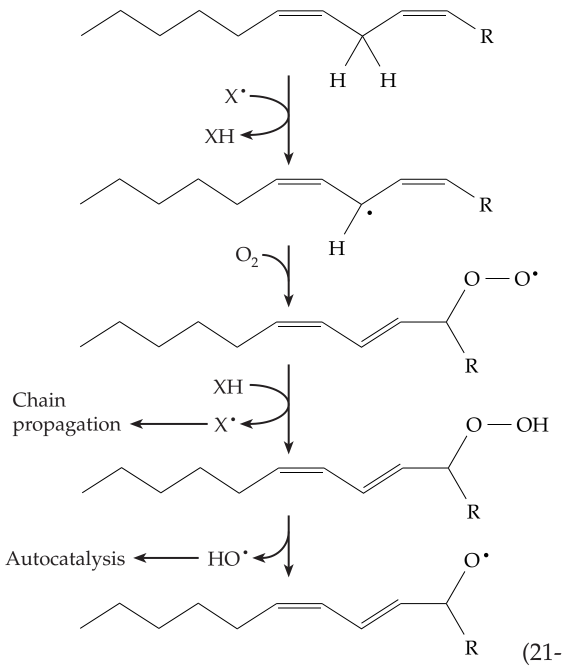
phosphoinositide esters. The released inositol phosphates, which act as water-soluble messengers, are further modified by the action of several phosphatases (Fig. 11-9).[225] At the same time, **diacylglycerols** are left in the membrane. With loss of the negative charges of the PI phosphates there will be immediate electrostatic effects in the membrane, which may alter the ionic environment, open ion channels, etc. The diacylglycerols, which diffuse within the membrane, may lose arachidonic acid from the *sn*-2 position to supply substrate for the arachidonate cascades described in Section D (Eq. 21-16). Diacylglycerols also activate the 11 isozyme forms of **protein kinase C**.[226–228] Some of these enzymes not only are activated by diacylglycerols, but also require **phosphatidylserine** for activity.

Other lipid-based signaling cascades arise from reactions that modify phosphatidylcholine molecules. Unsaturated fatty acids in the *sn*-2 position are readily oxidized by free radicals with cleavage of the hydrocarbon chains to form alcohols, aldehydes, and carboxylic acids. These mimic PAF in their biological activities.[194a,228a] Phosphatidate molecules with saturated or monounsaturated fatty acids in the *sn*-2 position arise from breakdown of phosphatidylcholine catalyzed by **phospholipase D**.[229] Phosphatidate may also be formed by a family of lipid **diacylglycerol kinases**.[230,230a] Phosphatidates containing saturated or monounsaturated fatty acids also have a variety of signaling activities.[230a] An arachidonoyl-diacylglycerol kinase is thought to function in many processes. An example is a PI-mediated cycle in invertebrate vision.[231] Sphingomyelin breakdown (black arrows in Fig. 21-6) releases diffusible ceramides that have been implicated as signaling molecules in cell proliferation, differentiation, growth arrest, and apoptosis.[232–237a] Sphingosine and sphingosine 1-*P* also have signaling functions.[211,237b-e]

Phospholipids have been shown to exchange between different membranes, e.g., of mitochondria and the ER. Exchanges of phosphatidylcholine, phosphatidylinositol, and sphingomyelin are catalyzed by specific **exchange proteins** (Box 21-A).[238,238a] These proteins may also participate in signaling, but their major function may be to transport the phospholipids from their sites of synthesis to the various membranes of the cell.

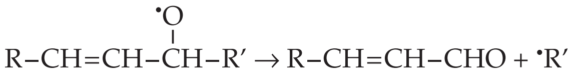
5. Peroxidation of Lipids and Rancidity

Storage of fats and oils leads to **rancidity**, a largely oxidative deterioration that causes development of unpleasant tastes, odors, and toxic compounds.[239] Similar chemical changes account for the "drying" of oil-based paints and varnishes. These reactions occur most readily with polyunsaturated fatty acids, whether free or in ester linkage within triacylglycerols. The reactions are initiated by free radicals, which may be generated by oxidative enzymes within or outside of cells, or by nonenzymatic reactions catalyzed by traces of transition metals or by environmental pollutants. Characteristic of rancidity is an autocatalytic chain reaction (Eq. 21-13).[239–241]



(21-13)

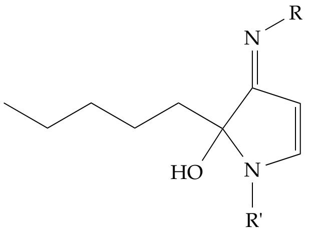
Radical $X^{\bullet}$, which initiates the reaction, is regenerated in a chain propagation sequence that, at the same time, produces an organic peroxide. The latter can be cleaved to form two additional radicals, which can also react with the unsaturated fatty acids to set up the autocatalytic process. Isomerization, chain cleavages, and radical coupling reactions also occur, especially with polyunsaturated fatty acids. For example, reactive unsaturated aldehydes can be formed (Eq. 21-14).



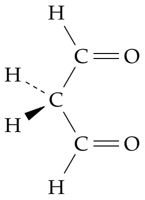
$$R{-}CH{=}CH{-}\overset{{}^{\bullet}O}{\overset{|}{C}}H{-}R' \rightarrow R{-}CH{=}CH{-}CHO + {}^{\bullet}R'$$

(21-14)

An intermediate in Eq. 21-13 may be converted to **4-hydroxy-2-nonenal**, a prominent product of the peroxidation of arachidonic or linoleic acids (Eq. 21-15).[242–243a] However, other biosynthetic pathways to this compound are possible.[244,244a] 4-Hydroxy-2-nonenal can react with side chains of lysine, cysteine, and histidine[245] to form fluorescent products such as the following cyclic compound generated by an oxidative reaction.[246]



Polyenoic acids also give rise to malondialdehyde, a reactive mutagenic compound, which can be reduced



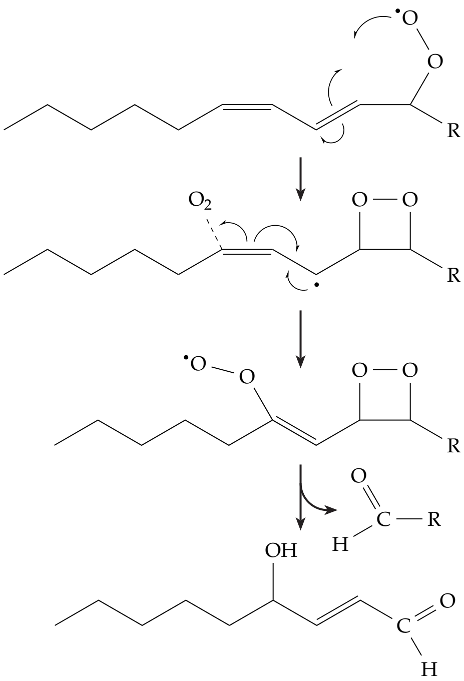
Malondialdehyde

and dehydrated to acrolein, a toxic compound which also reacts with both lysine and serine to

CH_2=CH–CHO

Acrolein

produce products excreted in the urine.[247,248] More dangerous are similar reactions of these aldehydes with proteins of the body in conditions such as diabetes or renal insufficiency.[242] The bifunctional malondialdehyde forms Schiff bases with protein amino groups and acts as a crosslinking agent.[249] **Age pigments** (also called **lipofuscin**), which tend to accumulate within neurons and other cells, are thought to represent precipitated lipid-protein complexes resulting from such reactions.[250] The reactions are similar to those of proteins with the products of sugar breakdown (glycation; p. 69).[250a] Organisms have developed multiple enzymatic mechanisms for detoxification of products of both glycation and oxidative degradation.[243a]



(21-15)

The oxidative degradation represented by the foregoing reactions is referred to as peroxidation. Peroxidation can lead to rapid development of rancidity in fats and oils. However, the presence of a small amount of tocopherol inhibits this decomposition, presumably by trapping the intermediate radicals in the form of the more stable tocopherol radicals (Eq. 15-54), which may dimerize or react with other radicals to terminate the chain.

Catalytic hydrogenation of vegetable oils is widely used to form harder fats and to decrease the content of polyunsaturated fatty acyl groups. The products have a greatly increased resistance to rancidity. However, they also contain fats with trans double bonds as well as isomers with double bonds in unusual positions.[251–253] Such compounds may interfere with normal fatty acid metabolism and also appear to affect serum lipoprotein levels adversely. Trans fatty acids are present in some foods. One hundred grams of butter contain 4–8 g, but hydrogenated fats often contain much more. It has been estimated that in the United States trans fatty acids account for 6–8% of total dietary fat.[253]

6. Some Nutritional Questions

While many of the poorer people on earth starve to death the problems of atherosclerosis and obesity affect many in wealthier societies.[253a–c] The fat content of foods is often blamed, and, as discussed in Boxes 21-B and 22-B, the quality of fatty acids in the diet is very important. However, like fatty acids, carbohydrates are also metabolized via acetyl-CoA and can readily be converted to both fatty acids and cholesterol.[253d] Obesity is largely a problem of excessive total caloric intake.

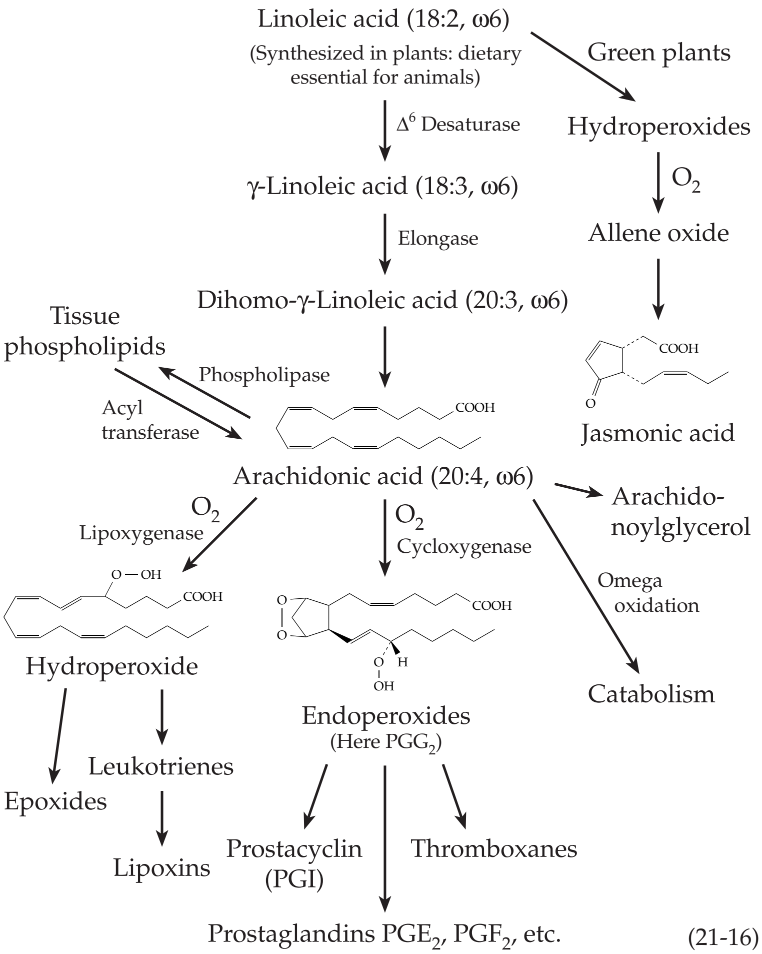
Why do some people stay slim while others become obese? What are the regulatory mechanisms that affect appetite and body composition? The human body weight tends to be stable or to increase slowly during adult life.[253e] Is there a natural set point for each individual? No, an apparent set point is just a result of action of a multitude of factors including genetic variations[253f] and psychological factors that affect exercise levels, eating habits, etc.[253g] It is worthwhile to recognize that the basal metabolic rate, which is also affected by many factors, accounts for a very high fraction of a person's energy expenditure (p. 283).

The following are among specific biochemical factors that act on the energy balance: the activity of acetyl-CoA carboxylase and the associated level of malonyl-CoA[253h,i]; the activity of mitochondrial uncoupling proteins (Box 18-C)[253j,k]; actions of the hormone leptin[253l] (which have been hard to interpret)[253m,n]; and other hormones including cholecystokinins and neuropeptide Y (Chapter 30).

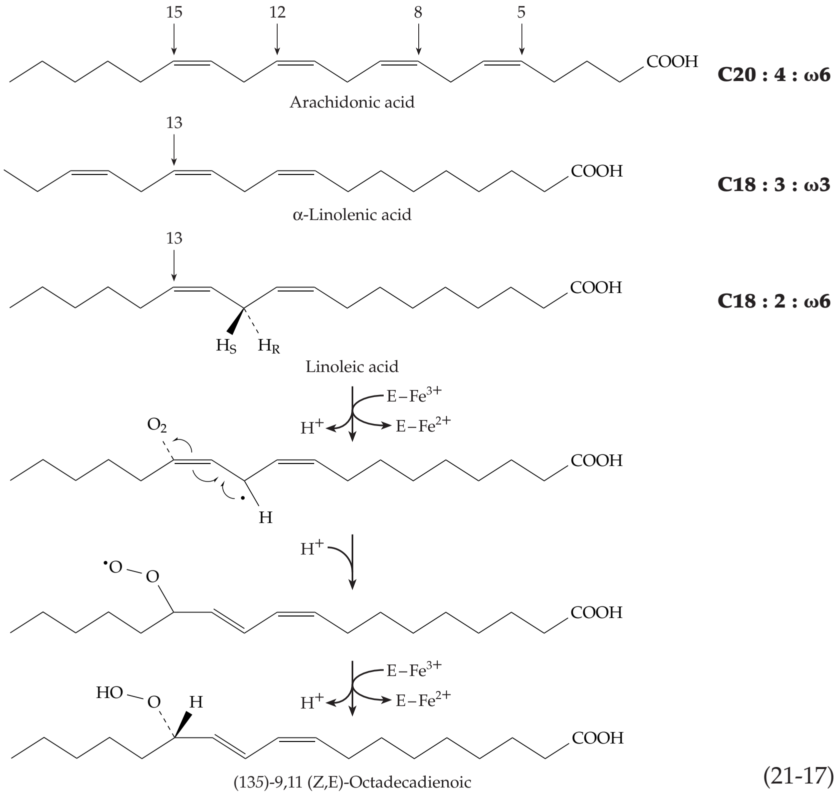
D. Prostaglandins and Related Prostanoid Compounds

Lipid peroxidation has often been regarded simply as an undesirable side reaction, but it is also a normal part of metabolism. Initiated by enzymatically generated radicals, peroxidation occurs as specific metabolic pathways, such as the **arachidonate cascade**, which leads to a variety of local hormones and other substances (Eq. 21-16).[254–256a]

As early as 1930, it was recognized that seminal fluid contains materials that promote contraction of uterine muscles. The active compounds, the **prostaglandins**, were isolated and crystallized in 1960 and were identified shortly thereafter.[257,258] As many as 14 closely related compounds are found in human seminal fluid, one of the richest known sources. Prostaglandins are present in seminal fluid at a total concentration of ~1 mM, but their action on smooth muscles has been observed at a concentration as low as 10^{-9} M. The structures and biosynthetic pathways of several of the prostaglandins are indicated in Fig. 21-7. Prostaglandins are usually abbreviated PG with an additional letter and numerical subscript added to indicate the type. The E type are β-hydroxyketones, the F type 1,3-diols, and the A type α, β-unsaturated ketones. Series 2 prostaglandins arise from arachidonic acid, while series 1 and 3 arise from fatty acids containing one fewer or one more double bond, respectively (Fig. 21-7). Additional forms are known.[257,259]


Linoleic acid (18:2, ω6)
(Synthesized in plants: dietary essential for animals)
Green plants
Δ6 Desaturase
Hydroperoxides
γ-Linoleic acid (18:3, ω6)
O2
Elongase
Allene oxide
Dihomo-γ-Linoleic acid (20:3, ω6)
Tissue phospholipids
Phospholipase
Acyl transferase
COOH
COOH
O
Jasmonic acid
Arachidonic acid (20:4, ω6)
Arachido-noylglycerol
O2
Lipoxygenase
O2
Cycloxygenase
Omega oxidation
O—OH
COOH
O
O
COOH
O H
OH
Hydroperoxide
Endoperoxides
(Here PGG2)
Catabolism
Leukotrienes
Epoxides
Lipoxins
Prostacyclin (PGI)
Thromboxanes
Prostaglandins PGE2, PGF2, etc.

(21-16)


15
12
8
5
COOH
C20 : 4 : ω6
Arachidonic acid
13
COOH
C18 : 3 : ω3
α-Linolenic acid
13
COOH
C18 : 2 : ω6
HS HR
Linoleic acid
E–Fe3+
H+
E–Fe2+
O2
COOH
H
H+
•O–O
COOH
E–Fe3+
HO–O H
H+
E–Fe2+
COOH
(135)-9,11 (Z,E)-Octadecadienoic

(21-17)

1. Metabolism of the Prostaglandins

Prostaglandins are not stored by cells but are synthesized in response to external stimuli. Arachidonic acid and other polyenoic acids are present in relatively small amounts (e.g., ~1% of total plasma fatty acids), but they are concentrated in the 2-position of phospholipids. This is in part a result of phospholipid "remodeling." Acyl groups are hydrolyzed from the *sn*-2 position by action of phospholipase A_2. An acyltransferase with a preference for arachidonoyl groups then transfers esterified arachidonic acid from

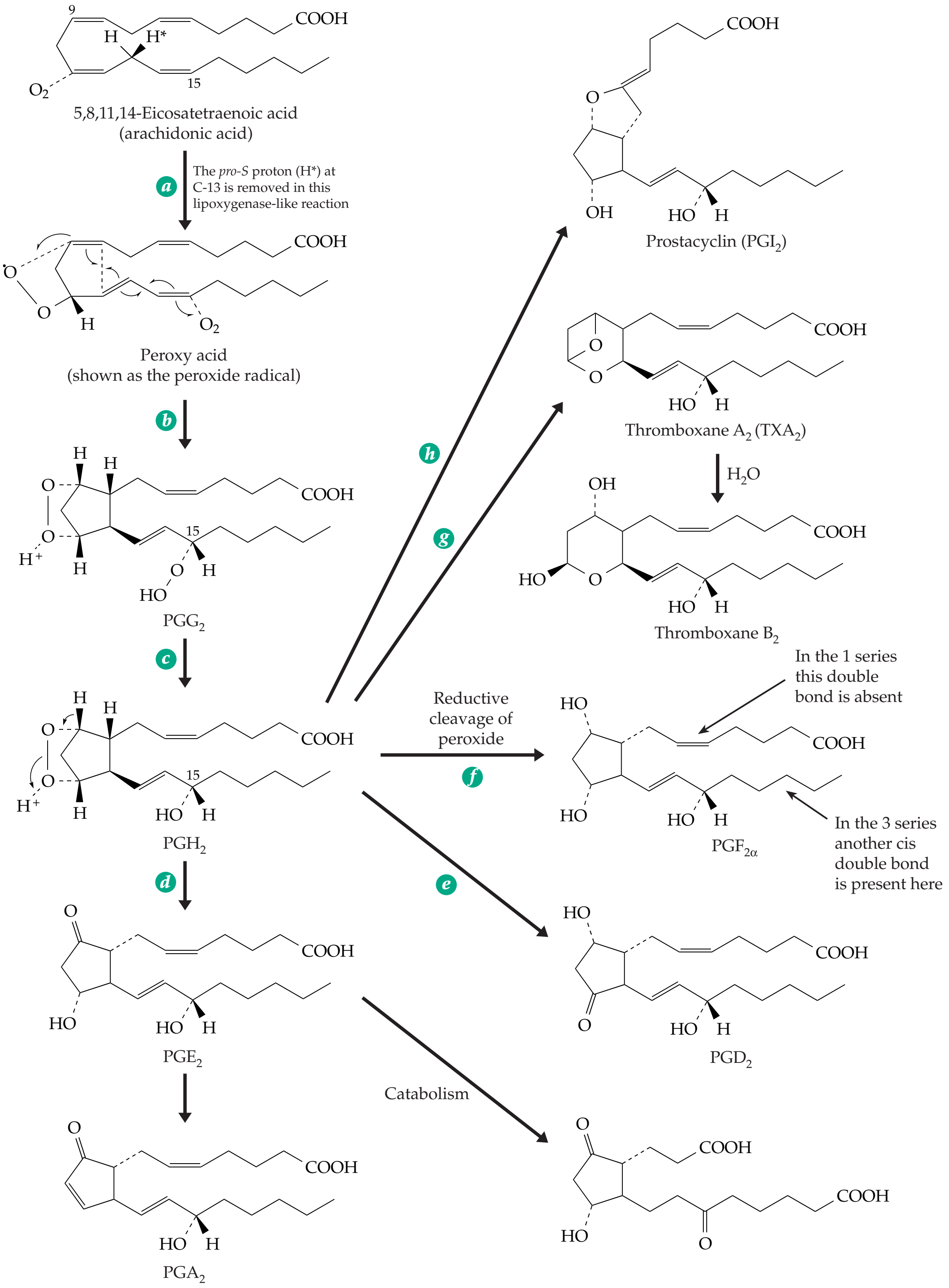

9
COOH
H
H*
O2
15
5,8,11,14-Eicosatetraenoic acid
(arachidonic acid)
a
The pro-S proton (H*) at C-13 is removed in this lipoxygenase-like reaction
O
O
H
O2
COOH
Peroxy acid
(shown as the peroxide radical)
b
H
H
O
O
H+
H
15
O
H
HO
COOH
PGG2
c
H
H
O
O
H+
H
15
HO
H
COOH
PGH2
d
O
HO
HO
H
COOH
PGE2
O
HO
H
COOH
PGA2
h
g
Reductive cleavage of peroxide
f
e
Catabolism
COOH
O
OH
HO
H
Prostacyclin (PGI2)
O
O
COOH
HO
H
Thromboxane A2 (TXA2)
H2O
OH
HO
O
COOH
HO
H
Thromboxane B2
In the 1 series this double bond is absent
HO
HO
COOH
HO
H
PGF2α
In the 3 series another cis double bond is present here
HO
O
COOH
HO
H
PGD2
O
COOH
HO
COOH
O

Figure 21-7 Pathway of synthesis and catabolism of the prostaglandins.

phosphatidylcholine and other phospholipids to the lyso forms of the phospholipids, which lack a 2-acyl group. The enzyme has a strong preference for the lyso-ethanolamine plasmalogens. As a consequence, in the plasmenylethanolamine of platelets arachidonoyl groups account for 66% of the acyl residues at the 2-position.[260] An arachidonate-specific acyl-CoA synthetase rapidly reconverts any free arachidonate that is not used for prostanoid synthesis back into phospholipids.[261]

The synthesis of prostaglandins, which was elucidated by Samuelsson,[258–260] begins with the release of arachidonate and other polyenoic acid precursors from phospholipids through the action of phospholipase A_2. The released arachidonate is then acted upon by **prostaglandin H synthases,** which catalyze two consecutive reactions at adjacent but distinct sites in a single protein.[262–262d] The first, **cyclooxygenase** or prostaglandin endoperoxide synthase reaction, forms PGG_2 from arachidonate and the second, a **peroxidase** reaction, generates PGH_2. There are two major mammalian isozymes of prostaglandin synthase, which are often called cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2). From studies of stereospecifically synthesized ^3H-containing fatty acid precursors, it was established that the first step in cyclooxygenase action involves removal of the *pro*-S proton at C-13 of the fatty acid (step *a*, Fig. 21-7). The O_2-requiring cyclooxygenase resembles lipoxygenase (Eq. 21-17).[181] The product is a peroxy acid, possibly in the form of the peroxide radical shown in Fig. 21-7. This radical (or peroxide anion) undergoes cyclization with synchronous attack by a separate O_2 molecule at C-15 (Fig. 21-7, step *b*) to give the endoperoxide PGG. Reduction of the latter to an OH group by the NADPH-dependent peroxidase (step *c*) yields PGH. The entire sequence is catalyzed by the single 70-kDa PGH synthase, which contains a single heme prosthetic group.[263] During the cyclooxygenase reaction the enzyme appears to accept electrons from cytochrome b_5. During the peroxidase step the heme group undergoes formation of the characteristic peroxidase intermediate compounds I and II[263] (Fig. 16-14). It has been suggested that a tyrosyl radical is generated in the peroxidase active site (on Y385 of COX-1) and is used to form an arachidonate radical that reacts with O_2 in the cyclooxygenase reaction.[264–266b] Alternatively, a carbocation mechanism is also possible.[267]

PGH can break down in three ways to give the E and F series of prostaglandins.[268] In one the proton at C-9 is eliminated (step *d*) as indicated by the small arrows by the PGH_2 structure of Fig. 21-7. An alternative isomerization (step *e*) gives PGD_2. The F prostaglandins are formed by reductive cleavage of the endoperoxide (step *f*). The A series and other prostaglandins arise by secondary reactions, one of which is shown in Fig. 21-7.

A biochemical characteristic of the prostaglandins is rapid catabolism. The product shown in Fig. 21-7 (lower right) arises by oxidation of the 15-OH to a carbonyl group, permitting reduction of the adjacent trans double bond. Two steps of β oxidation as well as ω oxidation are also required[269] to produce the dicarboxylic acid product shown. However, a series of products appears, and the distribution varies among species. Catabolism of prostaglandins is especially active in the lungs, and any prostaglandins entering the bloodstream are removed by a single pass through the lungs. This observation has led to the conclusion that prostaglandins are not hormones in the classical sense but act on a more local basis.

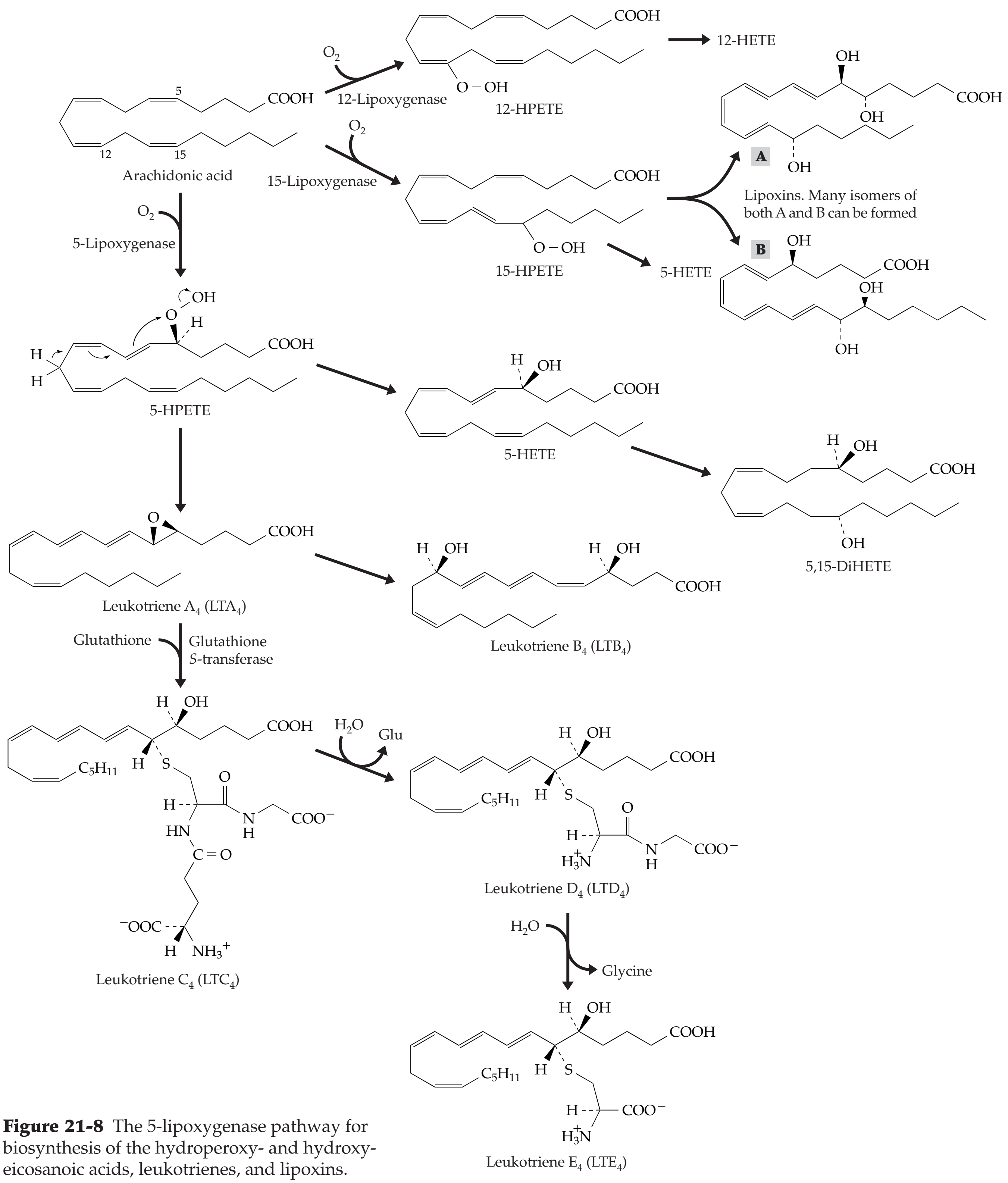
2. Thromboxanes and Prostacyclins

In blood platelets and in some other tissues PGG is also transformed to another series of compounds, the **thromboxanes**,[270] which were identified in 1975. Labile hemiacetals, the thromboxanes A (TXA, Fig. 21-7), are derived by rearrangement of PGH (step *g*). Thromboxane synthase,[271–273] which catalyzes the reaction, has characteristics of a cytochrome P450. Cytochromes P450 are known to react with peroxides as well as with O_2, and the endoperoxide of PGH may be opened by the synthase prior to rearrangement to TXA.[273] Thromboxane A_2 is so unstable that its half-life at 37°C in water is ~36 s. It is spontaneously converted to TXB_2 (Fig. 21-7), which contains an –OH group at C-15. The thromboxanes B are much more stable than TXA but are not very active physiologically.

By 1976, Vane and associates had identified another prostanoid compound, **prostacyclin** (or PGI_2).[274–275a] This compound also arises from PGH_2 by action of a cytochrome P450-like prostacyclin synthase (Fig. 21-7).[273,275,276] It is thought to be an important vasoprotective molecule. As with the thromboxanes, prostacyclin undergoes rapid inactivation[277] by hydrolysis to the physiologically inactive 6-oxo-$PGF_1α$.

3. Lipoxygenases

Lipoxygenases, of which the enzyme from soy beans has been studied the most, also catalyze oxidation of polyunsaturated fatty acids in lipids as indicated in Eq. 21-17. Formation of the hydroperoxide product is accompanied by a shift of the double bond and conversion from cis to trans configuration. Soybean lipoxygenase is a member of a family of related lipoxygenases that are found in all eukaryotes. All appear to have similar iron- or manganese-containing active sites and to act by similar mechanisms.[278–280e] The major substrate in animals is arachidonic acid (probably as the arachidonate ion). As marked on the structures above Eq. 21-17, there are 5-, 8-, 12-, and 15-lipoxygenases, which catalyze reaction with dioxygen at the indicated places.[281,282] Linoleic and linolenic acids are the primary substrates in plants. Soybean lipoxygenase acts on the 13-position of linoleic acid as shown in Eq. 21-17. However, this enzyme is often referred to as a *15-lipoxygenase* because it acts on arachidonate at C15. The 100-kDa enzyme from



5
COOH
12
15
Arachidonic acid
O2
12-Lipoxygenase
O – OH
12-HPETE
12-HETE
O2
15-Lipoxygenase
COOH
O – OH
15-HPETE
5-HETE
OH
COOH
OH
A
OH
Lipoxins. Many isomers of both A and B can be formed
B
OH
COOH
OH
OH
O2
5-Lipoxygenase
OH
O
H
H
H
COOH
5-HPETE
H
OH
COOH
5-HETE
H
OH
COOH
OH
5,15-DiHETE
O
COOH
Leukotriene A4 (LTA4)
H OH
H OH
COOH
Leukotriene B4 (LTB4)
Glutathione
Glutathione S-transferase
H OH
COOH
C5H11
H
S
O
H
N
H
COO–
HN
C = O
–OOC
H NH3+
Leukotriene C4 (LTC4)
H2O
Glu
H OH
COOH
C5H11
H
S
O
H
H3N+
N
H
COO–
Leukotriene D4 (LTD4)
H2O
Glycine
H OH
COOH
C5H11
H
S
H
COO–
H3N+
Leukotriene E4 (LTE4)

Figure 21-8 The 5-lipoxygenase pathway for biosynthesis of the hydroperoxy- and hydroxy-eicosanoic acids, leukotrienes, and lipoxins.

soybeans contains one atom of Fe(II), which is bound by a cluster of His and Tyr side chains.[283] It must be oxidized to the Fe(III) state before becoming active.[279,284] The initial reaction with O_2 may occur via an intermediate radical.

4. Leukotrienes, Lipoxins, and Related Compounds

Yet another series of products results from the action on arachidonate of tissue lipoxygenases, which compete with the prostaglandin-forming cyclooxygenases. The 5-lipoxygenase (Fig. 21-8) produces the unstable peroxide **5-hydroperoxy 6,8,11,14-eicosatetraenoic acid** (usually abbreviated **5-HPETE**). This enzyme requires ATP and Ca^{2+} and appears to be regulated by a series of other metabolites.[285] The 12- and 15-lipoxygenases, whose distribution varies among different mammalian organs and tissues, form the corresponding 12- and 15-HPETES as well as 5, 15-, 8,15-, 14,15-diHPETEs.[255,286] Some of these peroxides have physiological effects of their own, but they are largely transformed by peroxidases to more stable compounds such as the corresponding alcohols (hydroxy-icosatetraenoic acids or HETEs; Fig. 21-8).

The **leukotrienes** are formed from 5-HPETE (Fig. 21-8).[287,288] Dehydration of HPETE produces the unstable epoxide **leukotriene A_4** (LTA_4), which can be hydrolyzed enzymatically by leukocytes to the diol **leukotriene B_4** (LTB_4).[289,290] Alternatively leukotriene synthase, present in many cells, catalyzes the addition of glutathione (Box 11-B) to the LTA. This is a ring-opening reaction of the epoxide that can be visualized as a nucleophilic displacement by the thiolate anion of glutathione at C-6 (Fig. 21-8). The product is **leukotriene C_4** (LTC_4),[291] which can undergo consecutive removal of glutamate and glycine to form **leukotrienes D** and **E** (LTD_4, LTE_4), respectively. Removal of the glutamate occurs by the action of γ-glutamyl transpeptidase (Box 11-B), whereas removal of the glycine is hydrolytic. LTC_4 and the more potent LTD_4 have been identified as the **slow-reacting substance of anaphylaxis** (SRS-A), a long-sought mediator of bronchial asthma.[292,293] Leukotrienes can be formed from polyunsaturated acids other than arachidonic acid. Thus, eicosapentaenoic acid yields LTC_5 and LTD_5.[294] A lipoxygenase-derived product from the C18:2 linoleic acid is 13-hydroxylinoleic acid, which is made principally by endothelial cells that line blood vessels. It may contribute to resistance to blood clotting.[294]

Products of the 15-lipoxygenase pathway include a group of trihydroxytetraenes formed by leukocytes.[295] Several routes of biosynthesis, which may involve epoxide intermediates, are known.[296,297] The structures of two of these compounds, **lipoxin A** and **lipoxin B,** are shown in Fig. 21-8. Several stereoisomers and cis–trans isomers can be formed. These compounds can all arise from 15-HPETE, either by enzyme action or nonenzymatically. In fact, the entire series of prostanoid compounds arise by reactions related to but more specific than those that occur during nonenzymatic autoxidation of arachidonate.[255,298] Cytochrome P450-catalyzed reaction with O_2 can convert arachidonic acid into four different **epoxytrienoic acids** (**EETs**), which may also exist as stereoisomers. They are vasodilators which affect a variety of signaling pathways.[298a,b]

5. Physiological Effects of the Prostanoids

The release of arachidonate and initiation of the arachidonate cascade is induced by hormones, various inflammatory and immunological stimuli, and even mechanical agitation. Tissues do not all behave the same in response to the arachidonate cascade.[299] Blood platelets form largely thromboxane A_2, whereas tissues of the aorta form prostacyclin. Prostaglandin D_2 is a major prostanoid in the central nervous system.[300] Biological functions of prostanoids are also varied.[256a,301] The **primary prostaglandins** PGE and PGF were first recognized as mediators of inflammation. However, PGE_2 and PGF_2 sometimes have opposite effects. The unstable precursors PGG_2 and PGH_2, which have half-lives of only a few minutes, are much more powerful than the more stable PGEs and PGFs. Prostaglandin D_2, which is released in lungs during attacks of asthma, is thought to be a major bronchoconstrictor, but it may also serve as a neurotransmitter.[300] Thromboxanes released from platelets cause smooth muscle contraction and aggregation of the platelets, the first step in blood clot formation. Thromboxanes have half life-times of only seconds but are extremely potent not only in inducing platelet aggregation but also in causing contraction of blood vessels. Prostacyclin has the opposite effect, being a potent vasodilator that causes relaxation of smooth muscle. Upon release from blood vessel walls it acts to prevent clot formation.

The lipoxygenase pathway (Fig. 21-8) leading to the leukotrienes, lipoxins, and other products is especially active in leukocytes and in mast cells.[302] The leukotrienes promote inflammation, but lipoxins A_4 and B_4 are antiinflammatory.[303] The release of leukotrienes LTC_4, LTD_4, and LTE_4 in lung tissue is correlated with the long-lasting contractions of smooth muscle of the bronchi that are characteristic of asthma.[304] Leukotriene LTC_4 is ~1000 times more powerful than histamine in inducing such contraction. Leukotrienes have also been found in the central nervous system.[305]

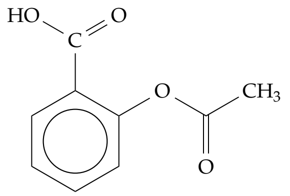
Some effects of prostaglandins are mediated through cell surface G-protein coupled receptors (see Chapter 11).[306] Some other prostanoids bind to and activate nuclear peroxisome proliferator-activated receptors.[306] PGJ_2 may inhibit fatty acid synthesis and fat deposition in adipose tissue through these receptors. Some of the prostanoid derivatives enter membranes and may become incorporated into phospholipids and exert their effects there.

A number of medical uses of prostaglandins have been discovered and more will probably be developed. While prostaglandins may be required for conception, small amounts of PGE_2 or $PGF_{2\alpha}$ induce abortion. $PGF_{2\alpha}$ is also used to induce labor. Prostaglandins are widely employed to control breeding of farm animals, to synchronize their estrus cycles, and to improve the efficiency of artificial insemination.[307]

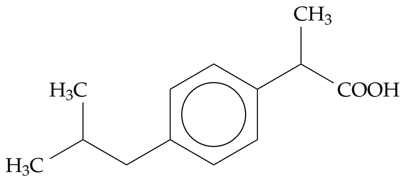
6. Inflammation

Special interest in the prostaglandins has focused on pain of inflammation and allergic responses. The medical significance is easy to see. Five million Americans have **rheumatoid arthritis**, an inflammatory disease. Bronchial asthma and other allergic diseases are equally important. Our most common medicine is **aspirin**, an anti-inflammatory drug. Both the inflammatory response and the immune response are normal parts of the defense mechanisms of the body, but both are potentially harmful, and it is their regulation that is probably faulty in rheumatoid arthritis and asthma. Overproduction of prostaglandins may be a cause of menstrual cramps.[308]

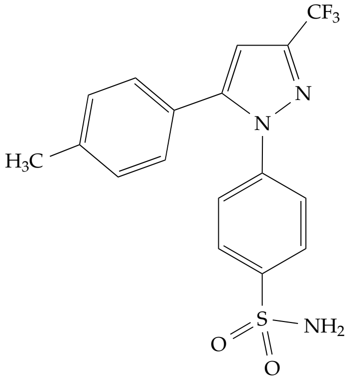
Prostaglandins have been implicated both in the induction of inflammation and in its relief. In inflammation small blood vessels become dilated, and fluid and proteins leak into the interstitial spaces to produce the characteristic swelling (edema). Many polymorphonuclear leukocytes attracted by chemotactic factors that include LTB_4[309] (Chapter 19) migrate into the inflamed area, engulfing dead tissue and bacteria. In this process lysosomes of the leukocytes release phospholipase A, which hydrolyzes phospholipids and initiates the arachidonate cascade. The leukotrienes that are formed promote the inflammatory response. However, cAMP can suppress inflammation, and PGE_2 has a similar effect. Indeed, E prostaglandins, when inhaled in small amounts, relieve asthma.



Aspirin (acetylsalicylic acid)



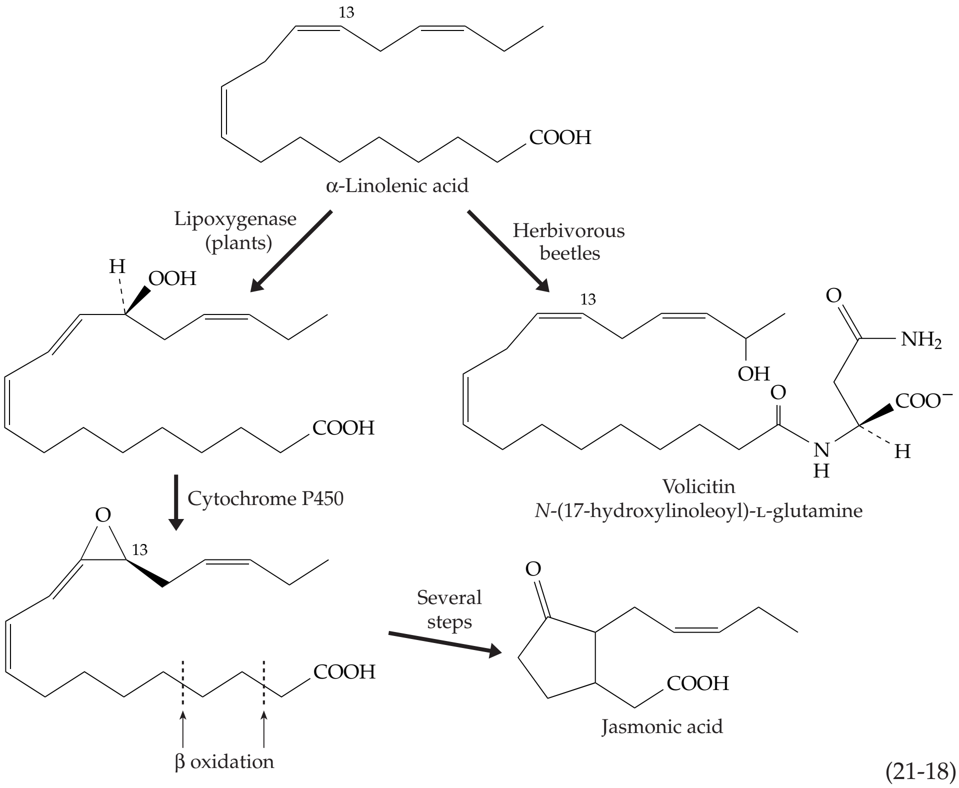
Ibuprofen



Celecoxib, a new COX-2 inhibitor

The synthesis of prostaglandins is inhibited by aspirin[310] and many other analgesic drugs. Aspirin is an acetylating reagent, and the inhibition has been traced to acetylation of the side chain –OH group of a single serine residue, Ser 530 of COX-1 or Ser 516 of COX-2 in the arachidonate binding channel.[311–313] Other nonsteroidal antiinflammatory drugs (NSAIDs), e.g., ibuprofen, are competitive inhibitors of the cyclooxygenases.[314] COX-3, a variant form of COX-1, may be the target for acetaminohen (Box 18-E).[314a] However, the same drugs also inhibit activation of neutrophils and may thus exercise their anti-inflammatory action in more than one way.[315] Since PGE_1 is a potent **pyrogen** (fever-inducing agent), a relationship to the ability of aspirin to reduce fever is also suggested. Unfortunately, all of these drugs inhibit COX-1 of platelets. Small regular doses of aspirin may be useful in preventing blood clots in persons with arterial disease, but they can be disastrous. Thousands of people die annually of hemorrhage caused by aspirin.[316,317]

Recently it was recognized that COX-1 provides eicosinoids for homeostatic purposes, while it is COX-2 that is inducible and generates prostoglandins for production of leukotrienes and induction of the inflammatory response. Now there is a major effort, with the first drugs already in use, to develop specific inhibitors for COX-2, which do not inhibit COX-1. It is hoped that these will be safer than aspirin.[256a,266a,311,317,317a-c] However, these drugs can also cause dangerous side effects.[317d] COX-2 of macrophages is also inhibited by γ-tocopherol, a major form of vitamin E.[317e]



(21-18)

7. Plant Lipoxygenases and Jasmonic Acid

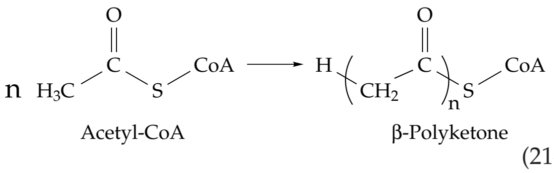
An octadecenoid signaling pathway (Eq. 21-18), which resembles the arachidonate cascade in some respects, plays an important role in green plants.[318–320] Alpha-linolenic acid is acted upon by a lipoxgenase in plastids to form a 13-hydroperoxy derivative. This is dehydrated and cyclized by allene oxide synthase. Although this doesn't appear to be an oxidation–reduction process, the enzyme seems to be a cytochrome P450 and to initiate the cyclization to the unstable epoxide **allene oxide** by homolytic cleavage of the peroxy group.[321,321a] Allene oxide synthase and the cyclase that acts in the next step may be cytosolic, while the β oxidation that shortens the chain occurs in plants exclusively in peroxisomes or glyoxysomes.[319]

Jasmonic acid is a plant growth regulator that affects many aspects of plant development as well as responses to environmental signals. A very important function is mobilization of plant defenses in response to damage by herbivores, by bacterial or fungal pathogens, or by ultraviolet light.[322,323] The synthesis of protease inhibitors as well as phytoalexins is induced. A curious variant of the jasmonate pathway is the acquisition of α-linoleic acid from plants by chewing caterpillars. The linoleic acid is hydroxylated by the insect, and conjugated with glutamine to form **volicitin** (*N*-(17-hydroxylinoleoyl)-L-glutamine; Eq. 21-18). Some of this compound reenters the plant from material regurgitated into wounds by the caterpillars. Volicitin induces the plant to release volatile terpenes and other compounds. The value to the caterpillars is not clear, but not only does volicitin induce defensive reactions in plants but also the released volatile compounds attract wasps that parasitize the caterpillars.[324,325] Plants also form a group of **isoprostanes** E$_1$ from α-linolenic acid.[326]

Allene oxides are unusual biological products. However, they are formed from arachidonic acid by some corals and are evidently precursors to prostaglandin esters, which may be present in high concentrations.[327] Allene oxides are also present in starfish oocytes.[321,328]

E. The Polyketides

In 1907, Collie proposed that polymers of ketene ($CH_2{=}C{=}O$) might be precursors of such compounds as **orsellinic** acid, a common constituent of lichens. The hypothesis was modernized in 1953 by Birch and Donovan, who proposed that several molecules of acetyl-CoA are condensed (Eq. 21-19) but *without the two reduction steps required in biosynthesis of fatty acids* (Fig. 17-12).[329] As we now know they were correct in assuming that the condensation occurs via malonyl-CoA and an acyl carrier group of an enzyme. The resulting **β-polyketone** can react in various ways to give the large group of compounds known as polyketides.

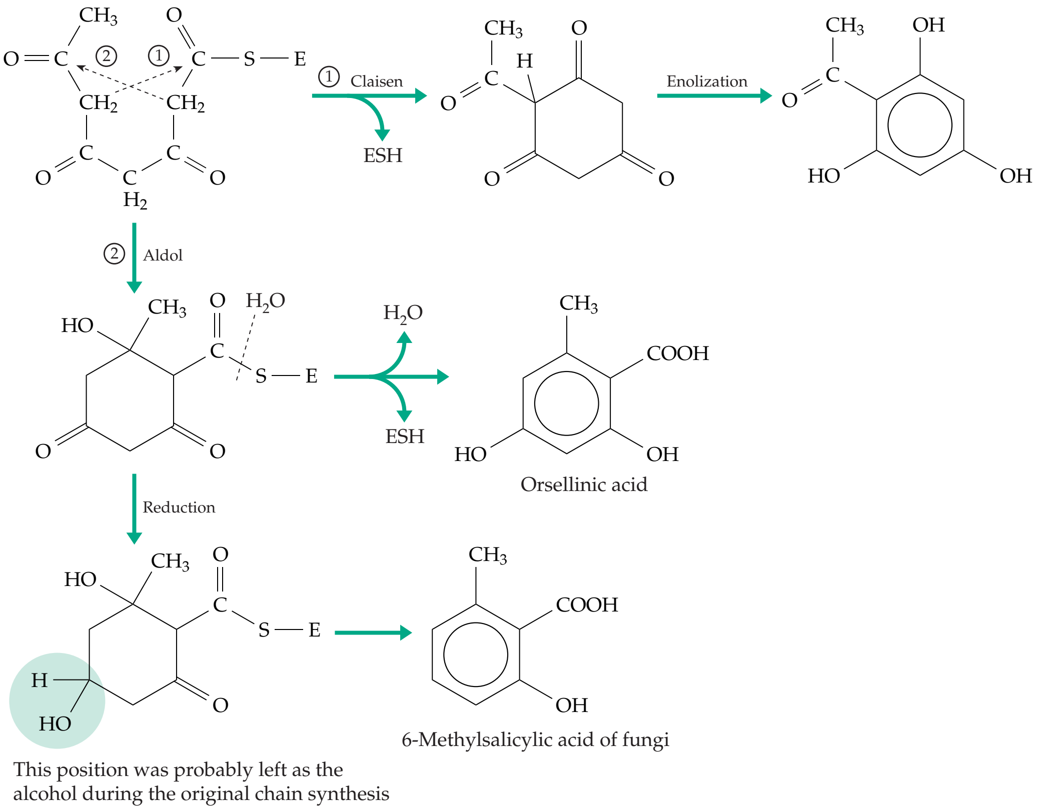


(21-19)

β-Polyketones can be stabilized by ring formation through ester or aldol condensations. Remaining carbonyl groups can be reduced (prior to or after cyclization) to hydroxyl groups, and the latter can be eliminated as water to form benzene or other aromatic rings. Figure 21-9 illustrates two ways in which cyclization can occur. One involves a Claisen ester condensation during which the enzyme and its SH group are eliminated. Enolization of the product gives a trihydroxy-acetophenone. The second cyclization reaction is the aldol condensation. Following the condensation water is eliminated, and the product is hydrolyzed and enolized to form orsellinic acid. Another product of fungal metabolism is **6-methylsalicylic acid,** which lacks one OH group of orsellinic acid. This synthesis can be explained by assuming that the carbonyl group at C-5 of the original β-polyketone was reduced to an OH group at some point during the biosynthesis. Elimination of two molecules of water together with enolization of the remaining ring carbonyl gives the product (Fig. 21-9).[330]

By allowing a few variations in the basic polyketone structure, the biosynthesis of a large number of unusual compounds can be explained. Extra oxygen atoms can be added by hydroxylation, and methyl groups may be transferred from *S*-adenosylmethionine to form methoxyl groups.[331] Occasionally a methyl group may be transferred directly to the carbon chain. Glycosyl groups may also be attached.[332,333] Many starter pieces other than acetyl-CoA may initiate polyketide synthesis. These include the branched-chain acids of Table 21-3, nicotinic and benzoic acids, 4-coumaroyl-CoA, and a 14:1 Δ^9-ACP. The last of these starter pieces is formed by desaturation of the corresponding 14:0-ACP and is converted via polyketide synthesis to one of a family of **anacardic acids,** which provide pest resistance to a variety of dicotyledenous plants (Fig. 21-10, bottom).[334] The CoA derivative of malonic acid amide is the starter piece for synthesis of the antibiotic **tetracycline** as indicated in Fig. 21-10).[335] Polyketide origins of some other antibiotics are also indicated in this figure.

The cloning and sequencing of genes for enzymes involved in synthesis of polyketides of fungi and actinomycetes has shown that these enzymes are closely related to the fatty acid synthases and, like the latter, have a multidomain structure (Fig. 21-11). The possibilities of engineering these genes, together with the urgent need for new antibiotics, has led to an



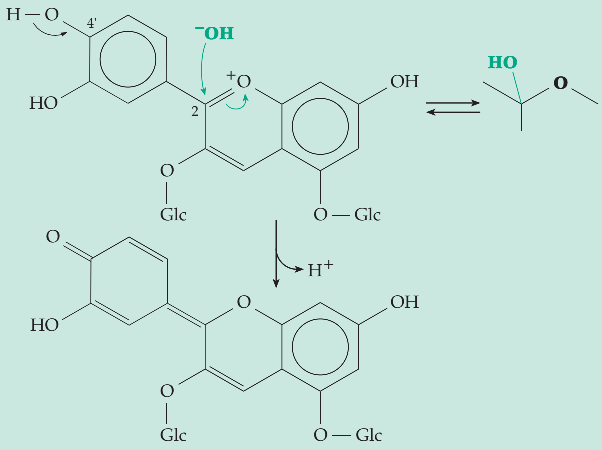
This position was probably left as the alcohol during the original chain synthesis

Figure 21-9 Postulated origin of orsellinic acid and other polyketides.

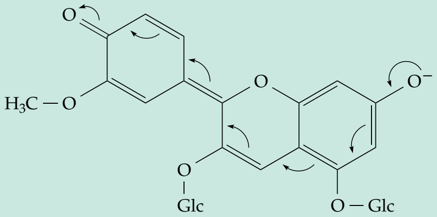
BOX 21-E HOW THE FLOWERS MAKE THEIR COLORS

Most of the pigments of flowers arise from a single polyketide precursor. Phenylalanine is converted to ***trans*-cinnamic acid** (Eq. 14-45) and then to cinnamoyl-CoA. The latter acts as the starter piece for chain elongation via malonyl-CoA (step *a* in the accompanying scheme). The resulting β-polyketone derivative can cyclize in two ways. The aldol condensation (step *b*) leads to **stilbenecarboxylic acid** and to such compounds as **pinosylvin** of pine trees. The Claisen condensation (step *c*) produces **chalcones**, **flavonones**, and **flavones**. These, in turn, can be converted to the yellow **flavonol pigments** and to the red, purple, and blue **anthocyanidins**.[a–c]

At the bottom of the synthetic scheme on the next page the structures and names of three common anthocyanidins are shown. The names are derived from those of flowers from which they have been isolated. The colors depend upon the number of hydroxyl groups and on the presence or absence of methylation and glycosylation. In addition to the three pigments indicated in the diagram, three other common anthocyanidins are formed by methylation. **Peonidin** is 3'-methylcyanidine. Methylation of delphinidin at position 3' yields **petunidin**, while methylation at both the 3' and 5' positions gives **malvidin**. There are many other anthocyanidins of more limited distribution. Anthocyanidins are nearly insoluble, but they exist in plants principally as glycosides known as **anthocyanins**. The number of different glycosides among the many species of flowering plants is large. Both the 3 and 5 –OH groups may be glycosylated with Glc, Gal, Rha, Ara, and by a large variety of oligosaccharides. The colors of the anthocyanins vary from red to violet and blue and are pH dependent. For example, **cyanin** (diglucosyl cyanidin) is red in acid solution and becomes violet upon dissociation of the 4'-hydroxyl group:

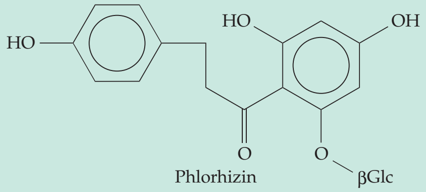


Dissociation of the 7–OH generates an anion with an extended conjugated π electron system, which will favor absorption of long-wavelength light (Chapter 23) and a blue color. Notice that a large number of resonance structures can be drawn for both the anthocyanin and the dissociated forms. Formation of complexes of Mg^{2+} or other metal ions with the 4' $-O^-$and adjacent OH groups may also stabilize blue colored forms.[d]



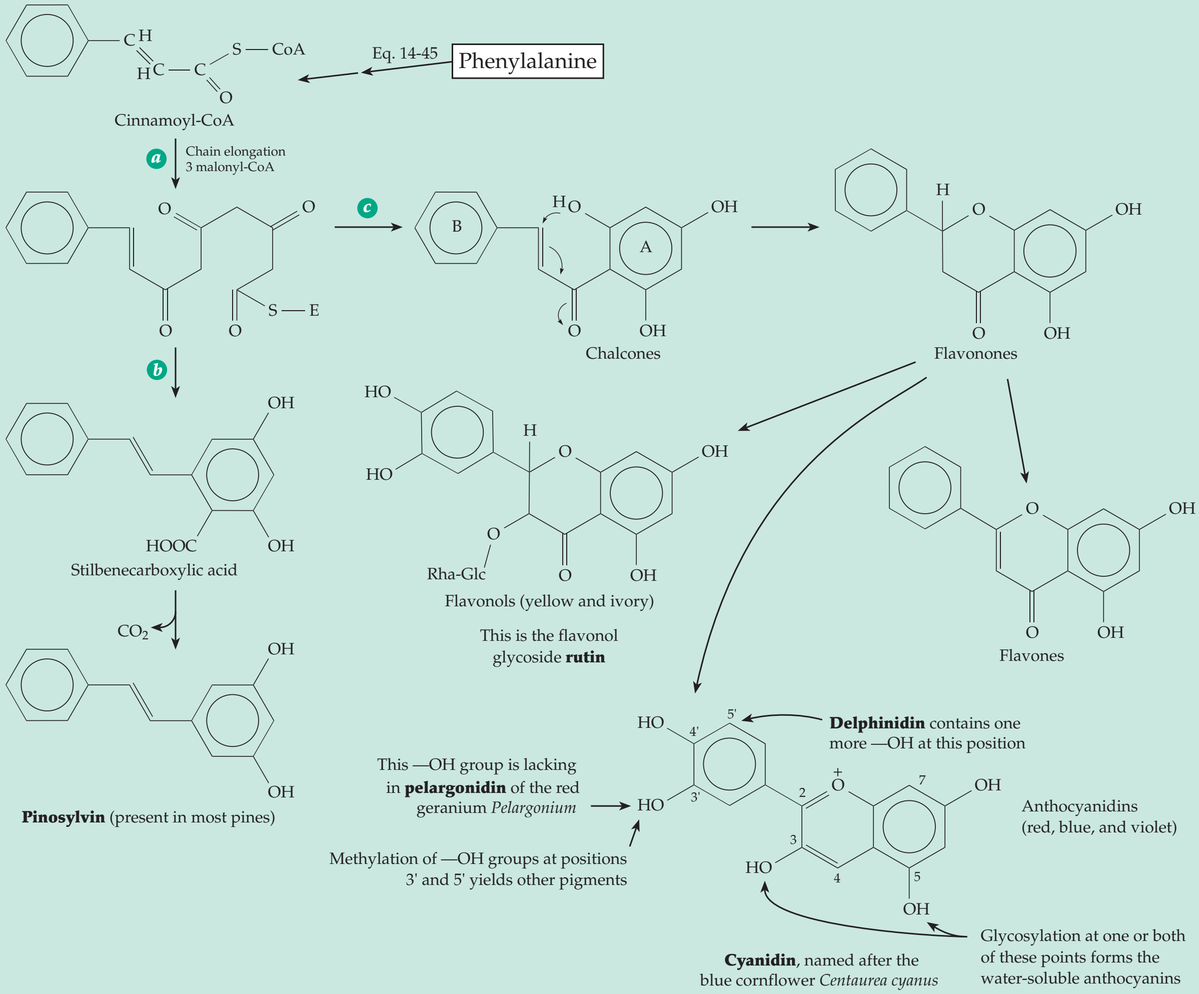
Most blue flower pigments are based on delphinidin,[b] but the "heavenly blue" of the morning glory is a peonidin with a complex caffeoylglucose-containing glycosyl group on the 3-position. Its blue color has been attributed to the relatively high pH of ~7.7 in vacuoles.[e] The aromatic rings within the glycosyl group of this and other complex anthocyanins may fold over the primary chromophore and stabilize the colored forms. A competing reaction, which is indicated in green on the first structure in this box, is the addition of a hydroxyl ion at C-2 to give a nearly colorless adduct.[f]

The yellow pigments of flowers are usually flavonols. The most common of all is **rutin**, the 3α-rhamnosyl-D-glucosyl derivative of **quercitin** (see diagram). An extraordinary number of other flavonols, flavones, and related compounds are found throughout the plant kingdom.[g] One of these is **phlorhizin**, a dihydrochalcone found in the root bark of pears, apples, and other plants of the rose family. Phlorhizin specifically blocks resorption of glucose by kidney tubules. As a result, the drug induces a strong glucosuria. The biochemical basis is uncertain, but the action on kidney tubules may be related to inhibition of mutarotase.[h]



Phlorhizin
(Phloridzin)

BOX 21-E (continued)


Eq. 14-45
Phenylalanine
Cinnamoyl-CoA
a
Chain elongation
3 malonyl-CoA
c
B
A
Chalcones
Flavonones
b
Stilbenecarboxylic acid
CO_2
Pinosylvin (present in most pines)
Flavonols (yellow and ivory)
Flavones
Anthocyanidins
(red, blue, and violet)
Rha-Glc
S — CoA
S — E

This is the flavonol glycoside **rutin**

Delphinidin contains one more —OH at this position

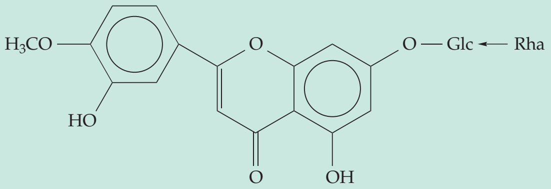
This —OH group is lacking in **pelargonidin** of the red geranium *Pelargonium*

Methylation of —OH groups at positions 3' and 5' yields other pigments

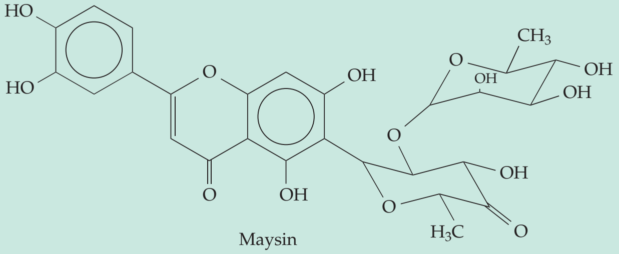
Cyanidin, named after the blue cornflower *Centaurea cyanus*

Glycosylation at one or both of these points forms the water-soluble anthocyanins

The flavone glycoside **hesperidin** makes up 80% of the dry weight of orange peels. It has been claimed (but not proved) that this compound, also known as **vitamin P** and **citrus bioflavonoid**, is essential to good health. Another flavone, **maysin**, is a resistance factor for the corn earworm and is present in silks of resistant strains of *Zea mays*.[i]


H_3CO
HO
O — Glc ← Rha
OH

Hesperidin


HO
HO
OH
CH_3
H_3C

Maysin

---

[a] Clevenger, S. (1964) *Sci. Am.* **210**(Jun), 85–92

[b] Harborne, J. B. (1988) in *Plant Pigments* (Goodwin, T. W., ed), pp. 299–343, Academic Press, London

[c] Lloyd, A. M., Walbot, V., and Davis, R. W. (1992) *Science* **258**, 1773–1775

[d] Kondo, T., Yoshida, K., Nakagawa, A., Kawai, T., Tamura, H., and Goto, T. (1992) *Nature (London)* **358**, 515–518

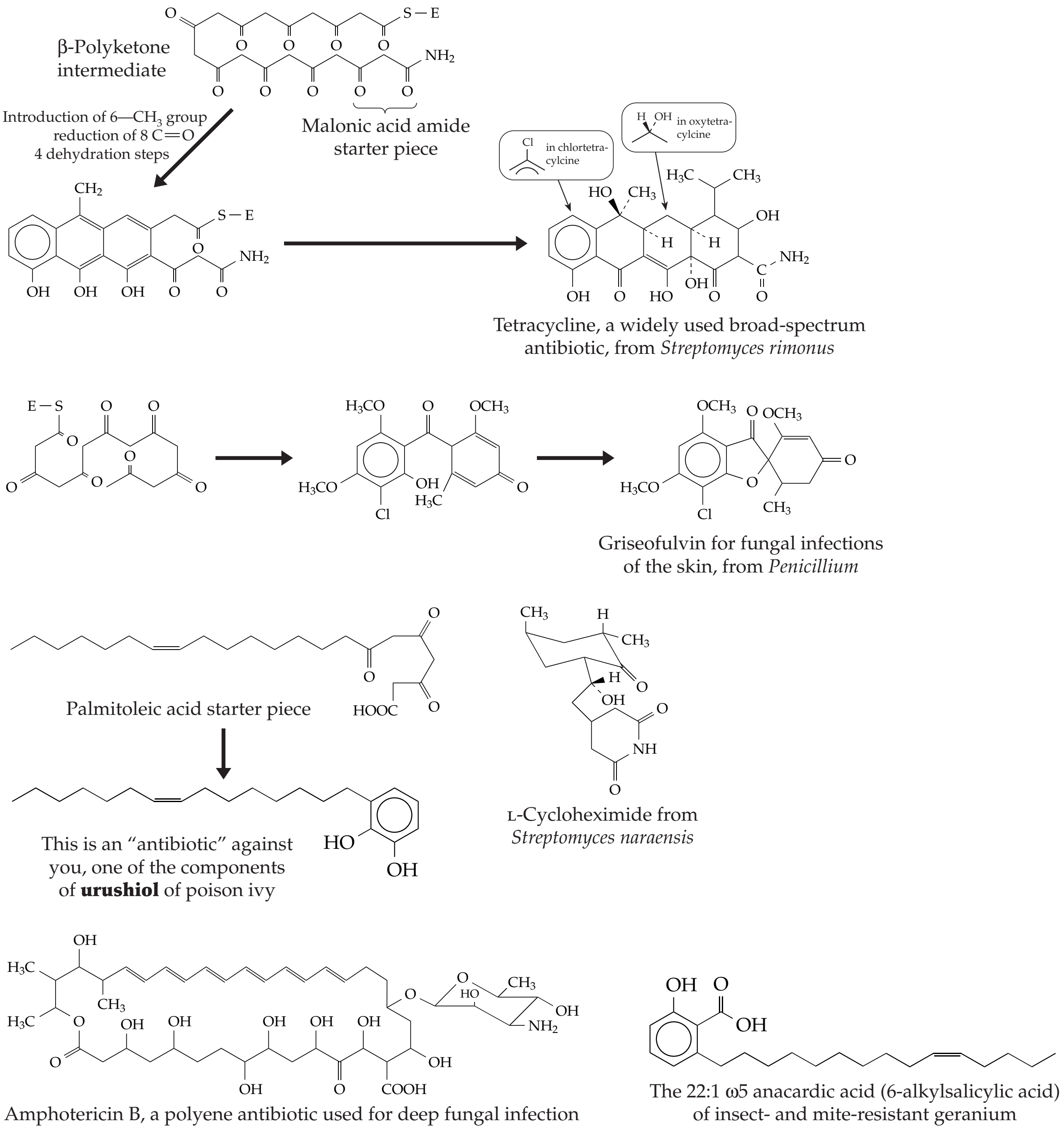
[e] Yoshida, K., Kondo, T., Okazaki, Y., and Katou, K. (1995) *Nature (London)* **373**, 291

[f] Figueiredo, P., Elhabiri, M., Saito, N., and Brouillard, R. (1996) *J. Am. Chem. Soc.* **118**, 4788–4793

[g] Nakayama, T., Yonekura-Sakakibara, K., Sato, T., Kikuchi, S., Fukui, Y., Fukuchi-Mizutani, M., Ueda, T., Nakao, M., Tanaka, Y., Kusumi, T., and Nishino, T. (2000) *Science* **290**, 1163–1166.

[h] White, A., Handler, P., and Smith, E. L. (1973) *Principles of Biochemistry*, 5th ed., McGraw-Hill, New York (pp. 415–416)

[i] Byrne, P. F., McMullen, M. D., Snook, M. E., Musket, T. A., Theuri, J. M., Widstrom, N. W., Wiseman, B. R., and Coe, E. H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8820–8825


β-Polyketone intermediate
Introduction of 6—CH$_3$ group
reduction of 8 C=O
4 dehydration steps
Malonic acid amide starter piece
Cl in chlortetracylcine
H OH in oxytetracylcine
Palmitoleic acid starter piece
This is an "antibiotic" against you, one of the components of **urushiol** of poison ivy

Tetracycline, a widely used broad-spectrum antibiotic, from *Streptomyces rimonus*

Griseofulvin for fungal infections of the skin, from *Penicillium*

L-Cycloheximide from *Streptomyces naraensis*

Amphotericin B, a polyene antibiotic used for deep fungal infection

The 22:1 ω5 anacardic acid (6-alkylsalicylic acid) of insect- and mite-resistant geranium

Figure 21-10 Some important polyketide antibiotics and plant defensive compounds.

explosion of information about polyketide synthases.[336,336a–c]

A 26-kb gene cluster encoding enzymes for synthesis of the blue antibiotic **actinorhodin** by *Streptomyces coelicolor* has been cloned and sequenced.[332,337] The three large ~10-kb genes required for formation of the broad-spectrum antibiotic **erythromycin** by *Saccharopolyspora erythraea* have also been cloned and sequenced.[337–339] In both cases, the genes encode large proteins with structures resembling those of the eukaryotic fatty acid synthases (Section B,1). However, a new feature is evident. As shown in Fig. 21-11, each of the three polypeptides of the deoxyerythronolide synthase, which synthesizes the aglycone of erythromycin, consists of two multidomain modules, each able to catalyze one round of reaction with a new molecule of malonyl-CoA. When reduction of an oxo group or dehydration and

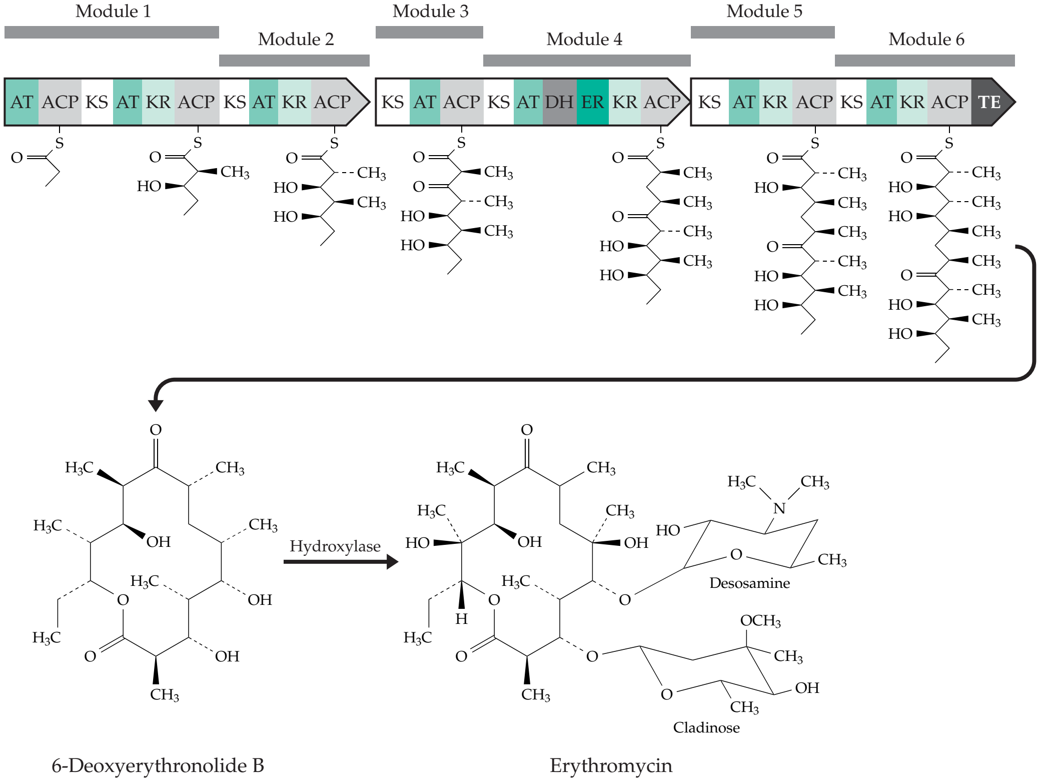


Figure 21-11 Catalytic domains within three polypeptide chains of the modular polyketide synthase that forms 6-deoxyerythronolide B, the aglycone of the widely used antibiotic erythromycin. The domains are labeled as for fatty acid synthases; AT, acyltransferase; ACP, acyl carrier protein; KS, β-ketoacyl-ACP synthase; KR, ketoreductase; DH, dehydrase; ER, enoylreductase; TE, thioesterase. After Pieper *et al.*[338] Courtesy of Chaitan Khosla.

reduction of an enoyl-CoA are not needed in a round, the KR, DH, and ER domains are absent (as in module 3 of Fig. 21-11). A final domain contains a thioesterase that releases and cyclizes the product. The “assembly line” sequence of synthetic steps, beginning with a propionyl group from propionyl-CoA, is pictured in Fig. 21-11. Two hydroxylation steps[340] and transfer of two unusual glycosyl groups complete the synthesis of the antibiotic.

Other medically important polyketides include the antibiotics **doxorubicin** (14-hydroxydaunomycin; Fig. 5-23),[341] rifamycin (Box 28-A),[342] and the antifungal **pimaricin**,[343] **griseofulvin**, and **amphotericin** (Fig. 21-10), the HMG-CoA reductase inhibitor **lovastatin**,[344] the 2-butanyl-4-methylthreonine of cyclosporin A (Box 9-F),[345] and other immunosuppressants such as **rapamycin**.[346] Many characteristic plant products, including **stilbenes**[347] and **chalcones**[348,348a] (Box 21-E), are polyketides. A variety of different polyketides serve as phytoalexins.[349] Some such as **aflatoxin**[350] are dangerous toxins. Ants and ladybird beetles make toxic polyamine alkaloids using a polyketide pathway.[350a]

Avermectin (Fig. 30–25), a widely used antibiotic against canine heartworms, is formed by a polyketide synthase with an unusually broad specificity for starter units. More than 40 alternative carboxylic acids are accepted. By grafting the first multidomain module of the erythromycin-forming synthase (of Fig. 21-11) onto the wide-specificity loading module of the avermectin-forming synthase, a whole new series of antibiotics have been created.[351] This is only one of many steps being taken to create new aliphatic and aromatic, linear and macrocyclic polyketides by genetic engineering.[336,352,353] Combinatorial biosynthesis (see Chapter 3) is also being developed[336,354] and has even been discovered in nature.[355]

References

1. Ayvazian, L., Crenon, I., Hermoso, J., Pignol, D., Chapus, C., and Kerfelec, B. (1998) *J. Biol. Chem.* **273**, 33604 – 33609
2. Bezzine, S., Ferrato, F., Ivanova, M. G., Lopez, V., Verger, R., and Carrière, F. (1999) *Biochemistry* **38**, 5499 – 5510
2a. Crandall, W. V., and Lowe, M. E. (2001) *J. Biol. Chem.* **276**, 12505 – 12512
3. Hermoso, J., Pignol, D., Kerfelec, B., Crenon, I., Chapus, C., and Fontecilla-Camps, J. C. (1996) *J. Biol. Chem.* **271**, 18007 – 18016
4. Kobayashi, J., Applebaum-Bowden, D., Dugi, K. A., Brown, D. R., Kashyap, V. S., Parrott, C., Duarte, C., Maeda, N., and Santamarina-Fojo, S. (1996) *J. Biol. Chem.* **271**, 26296 – 26301
5. Winkler, F. K., D'Arcy, A., and Hunziker, W. (1990) *Nature (London)* **343**, 771 – 774
6. Levy, E., Mehran, M., and Seidman, E. (1995) *FASEB J.* **9**, 626 – 635
7. Havel, R. J., and Kane, J. P. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1841 – 1852, McGraw-Hill, New York
7a. Liu, H., Talmud, P. J., Lins, L., Brasseur, R., Olivecrona, G., Peelman, F., Vandekerckhove, J., Rosseneu, M., and Labeur, C. (2000) *Biochemistry* **39**, 9201 – 9212
7b. Rozek, A., Sparrow, J. T., Weisgraber, K. H., and Cushley, R. J. (1999) *Biochemistry* **38**, 14475 – 14484
8. Orlova, E. V., Sherman, M. B., Chiu, W., Mowri, H., Smith, L. C., and Gotto, A. M., Jr. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8420 – 8425
9. Brown, M. S., and Goldstein, J. L. (1974) *Sci. Am.* **251**(Nov), 58 – 66
10. Phillips, J. C., Wriggers, W., Li, Z., Jonas, A., and Schulten, K. (1997) *Biophys. J.* **73**, 2337 – 2346
10a. Segrest, J. P., Jones, M. K., Klon, A. E., Sheldahl, C. J., Hellinger, M., De Loof, H., and Harvey, S. C. (1999) *J. Biol. Chem.* **274**, 31755 – 31758
10b. Tricerri, M. A., Agree, A. K. B., Sanchez, S. A., Bronski, J., and Jonas, A. (2001) *Biochemistry* **40**, 5065 – 5074
11. Rosseneu, M., and Labeur, C. (1995) *FASEB J.* **9**, 768 – 776
12. Chan, L. (1992) *J. Biol. Chem.* **267**, 25621 – 25624
13. Ferguson, S. J. (1987) *Trends Biochem. Sci.* **12**, 353 – 357
14. MacGinnitie, A. J., Anant, S., and Davidson, N. O. (1995) *J. Biol. Chem.* **270**, 14768 – 14775
15. Richardson, N., Navaratnam, N., and Scott, J. (1998) *J. Biol. Chem.* **273**, 31707 – 31717
16. Lawn, R. M. (1992) *Sci. Am.* **266**(Jun), 54 – 60
17. Utermann, G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1887 – 1912, McGraw-Hill, New York
18. Trieu, V. N., and McConathy, W. J. (1995) *J. Biol. Chem.* **270**, 15471 – 15474
19. Mochalkin, I., Cheng, B., Klezovitch, O., Scanu, A. M., and Tulinsky, A. (1999) *Biochemistry* **38**, 1990 – 1998
19a. Wang, J., and White, A. L. (1999) *J. Biol. Chem.* **274**, 12883 – 12889
20. Fless, G. M., Santiago, J. Y., Furbee, J., Jr., and Meredith, S. C. (1997) *Biochemistry* **36**, 11304 – 11313
21. Williams, K. J., Fless, G. M., Petrie, K. A., Snyder, M. L., Brocia, R. W., and Swenson, T. L. (1992) *J. Biol. Chem.* **267**, 13284 – 13292
21a. Goldstein, J. L., and Brown, M. S. (2001) *Science* **292**, 1310 – 1312
22. Baggio, G., Donazzan, S., Monti, D., Mari, D., Martini, S., Gabelli, C., Dalla Vestra, M., Previato, L., Guido, M., Pigozzo, S., Cortella, I., Crepaldi, G., and Franceschi, C. (1998) *FASEB J.* **12**, 433 – 437
23. Bergeron, J., Frank, P. G., Scales, D., Meng, Q.-H., Castro, G., and Marcel, Y. L. (1995) *J. Biol. Chem.* **270**, 27429 – 27438
24. Deeb, S. S., Cheung, M. C., Peng, R., Wolf, A. C., Stern, R., Albers, J. J., and Knopp, R. H. (1991) *J. Biol. Chem.* **266**, 13654 – 13660
25. Laccotripe, M., Makrides, S. C., Jonas, A., and Zannis, V. I. (1997) *J. Biol. Chem.* **272**, 17511 – 17522
25a. Gorshkova, I. N., Liu, T., Zannis, V. I., and Atkinson, D. (2002) *Biochemistry* **41**, 10529–10539
25b. Klon, A. E., Segrest, J. P., and Harvey, S. C. (2002) *Biochemistry* **41**, 10895–10905
26. Borhani, D. W., Rogers, D. P., Engler, J. A., and Brouillette, C. G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12291 – 12296
27. Wang, G., Sparrow, J. T., and Cushley, R. J. (1997) *Biochemistry* **36**, 13657 – 13666
28. Rogers, D. P., Roberts, L. M., Lebowitz, J., Engler, J. A., and Brouillette, C. G. (1998) *Biochemistry* **37**, 945 – 955
29. Lopez, J., Roghani, A., Bertrand, J., Zanni, E., Kalopissis, A., Zannis, V. I., and Chambaz, J. (1994) *Biochemistry* **33**, 4056 – 4064
30. Boisfer, E., Lambert, G., Atger, V., Tran, N. Q., Pastier, D., Benetollo, C., Trottier, J.-F., Beaucamps, I., Antonucci, M., Laplaud, M., Griglio, S., Chambaz, J., and Kalopissis, A.-D. (1999) *J. Biol. Chem.* **274**, 11564 – 11572
30a. Pennacchio, L. A., Olivier, M., Hubacek, J. A., Cohen, J. C., Cox, D. R., Fruchart, J.-C., Krauss, R. M., and Rubin, E. M. (2001) *Science* **294**, 169–173
31. Dong, L.-M., Wilson, C., Wardell, M. R., Simmons, T., Mahley, R. W., Weisgraber, K. H., and Agard, D. A. (1994) *J. Biol. Chem.* **269**, 22358 – 22365
32. Wilson, C., Wardell, M. R., Weisgraber, K. H., Mahley, R. W., and Agard, D. A. (1991) *Science* **252**, 1817 – 1822
33. Mahley, R. W., and Rall, S. C. J. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1953 – 1980, McGraw-Hill, New York
33a. Kypreos, K. E., Morani, P., van Dijk, K. W., Havekes, L. M., and Zannis, V. I. (2001) *Biochemistry* **40**, 6027 – 6035
33b. Morrow, J. A., Arnold, K. S., Dong, J., Balestra, M. E., Innerarity, T. L., and Weisgraber, K. H. (2000) *J. Biol. Chem.* **275**, 2576 – 2580
34. Law, J. H., and Wells, M. A. (1989) *J. Biol. Chem.* **264**, 16335 – 16338
35. Breiter, D. R., Kanost, M. R., Benning, M. M., Wesenberg, G., Law, J. H., Wells, M. A., Rayment, I., and Holden, H. M. (1991) *Biochemistry* **30**, 603 – 608
36. Zhang, Y., Lewis, R. N. A. H., McElhaney, R. N., and Ryan, R. O. (1993) *Biochemistry* **32**, 3942 – 3952
37. Soulages, J. L., Rivera, M., Walker, F. A., and Wells, M. A. (1994) *Biochemistry* **33**, 3245 – 3251
37a. Dettloff, M., Weers, P. M. M., Niere, M., Kay, C. M., Ryan, R. O., and Wiesner, A. (2001) *Biochemistry* **40**, 3150 – 3157
38. Heinemann, B., Andersen, K. V., Nielsen, P. R., Bech, L. M., and Poulsen, F. M. (1996) *Protein Sci.* **5**, 13 – 23
39. Banaszak, L., Sharrock, W., and Timmins, P. (1991) *Ann. Rev. Biophys. Biophys. Chem.* **20**, 221 – 246
40. Deckelbaum, R. J., Ramakrishnan, R., Eisenberg, S., Olivecrona, T., and Bengtsson-Olivecrona, G. (1992) *Biochemistry* **31**, 8544 – 8551
41. Brunzell, J. D. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1913 – 1932, McGraw-Hill, New York
42. van Tilbeurgh, H., Egloff, M.-P., Martinez, C., Rugani, N., Verger, R., and Cambillau, C. (1993) *Nature (London)* **362**, 814 – 820
42a. Borén, J., Lookene, A., Makoveichuk, E., Xiang, S., Gustafsson, M., Liu, H., Talmud, P., and Olivecrona, G. (2001) *J. Biol. Chem.* **276**, 26916 – 26922
42b. Hime, N. J., Barter, P. J., and Rye, K.-A. (1998) *J. Biol. Chem.* **273**, 27191 – 27198
43. Lookene, A., Savonen, R., and Olivecrona, G. (1997) *Biochemistry* **36**, 5267 – 5275
44. Dichek, H. L., Fojo, S. S., Beg, O. U., Skarlatos, S. I., Brunzell, J. D., Cutler, G. B., Jr., and Brewer, H. B., Jr. (1991) *J. Biol. Chem.* **266**, 473 – 477
45. Ma, Y., Bruin, T., Tuzgol, S., Wilson, B. I., Roederer, G., Liu, M.-S., Davignon, J., Kastelein, J. J. P., Brunzell, J. D., and Hayden, M. R. (1992) *J. Biol. Chem.* **267**, 1918 – 1923
46. Dugi, K. A., Dichek, H. L., Talley, G. D., Brewer, J., HB, and Santamarina-Fojo, S. (1992) *J. Biol. Chem.* **267**, 25086 – 25091
47. Mamo, J. C. L., Elsegood, C. L., Gennat, H. C., and Yu, K. (1996) *Biochemistry* **35**, 10210 – 10214
48. Strickland, D. K., Kounnas, M. Z., and Argraves, W. S. (1995) *FASEB J.* **9**, 890 – 898
49. Sakai, J., Hoshino, A., Takahashi, S., Miura, Y., Ishii, H., Suzuki, H., Kawarabayasi, Y., and Yamamoto, T. (1994) *J. Biol. Chem.* **269**, 2173 – 2182
50. Willnow, T. E., Sheng, Z., Ishibashi, S., and Herz, J. (1994) *Science* **264**, 1471 – 1474
51. Hajjar, D. P., and Haberland, M. E. (1997) *J. Biol. Chem.* **272**, 22975 – 22978
52. Suzuki, K., Yamada, T., and Tanaka, T. (1999) *Biochemistry* **38**, 1751 – 1756
53. Krieger, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4077 – 4080
53a. Bhattacharya, A. A., Grüne, T., and Curry, S. (2000) *J. Mol. Biol.* **303**, 721 – 732
53b. Saltiel, A. R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 535 – 537
53c. De Simone, G., Galdiero, S., Manco, G., Lang, D., Rossi, M., and Pedone, C. (2000) *J. Mol. Biol.* **303**, 761 – 771
54. Holak, T. A., Kearsley, S. K., Kim, Y., and Prestegard, J. H. (1988) *Biochemistry* **27**, 6135 – 6142
55. Heath, R. J., and Rock, C. O. (1995) *J. Biol. Chem.* **270**, 15531 – 15538
56. Huang, W., Jia, J., Edwards, P., Dehesh, K., Schneider, G., and Lindqvist, Y. 1998) *EMBO J.* **17**, 1183 – 1191
57. Heath, R. J., and Rock, C. O. (1995) *J. Biol. Chem.* **270**, 26538 – 26542
58. Joshi, A. K., Witkowski, A., and Smith, S. (1998) *Biochemistry* **37**, 2515 – 2523
58a. Zhang, Y.-M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O., and White, S. W. (2001) *J. Biol. Chem.* **276**, 8231 – 8238
58b. Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A., and Wright, H. T. (2001) *J. Biol. Chem.* **276**, 20516 – 20522
58c. Qiu, X., Janson, C. A., Smith, W. W., Head, M., Lonsdale, J., and Konstantinidis, A. K. (2001) *J. Mol. Biol.* **307**, 341 – 356

References

58d. McGuire, K. A., Siggaard-Andersen, M., Bangera, M. G., Olsen, J. G., and von Wettstein-Knowles, P. (2001) *Biochemistry* **40**, 9836 – 9845

58e. Moche, M., Dehesh, K., Edwards, P., and Lindqvist, Y. (2001) *J. Mol. Biol.* **305**, 491 – 503

59. Jayakumar, A., Tai, M.-H., Huang, W.-Y., Al-Feel, W., Hsu, M., Abu-Elheiga, L., Chirala, S. S., and Wakil, S. J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8695 – 8699

59a. Chirala, S. S., Jayakumar, A., Gu, Z.-W., and Wakil, S. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3104 – 3108

59b. Witkowski, A., Joshi, A. K., and Smith, S. (2002) *Biochemistry* **41**, 10877–10887

60. McCarthy, A. D., and Hardie, D. G. (1984) *Trends Biochem. Sci.* **9**, 60 – 63

61. Joshi, A. K., Witkowski, A., and Smith, S. (1997) *Biochemistry* **36**, 2316 – 2322

62. Witkowski, A., Joshi, A. K., Rangan, V. S., Falick, A. M., Witkowska, H. E., and Smith, S. (1999) *J. Biol. Chem.* **274**, 11557 – 11563

63. Wakil, S. J. (1989) *Biochemistry* **28**, 4523 – 4530

64. Jayakumar, A., Chirala, S. S., and Wakil, S. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12326 – 12330

65. Smith, S. (1994) *FASEB J.* **8**, 1248 – 1259

66. Wakil, S. J., Stoops, J. K., and Joshi, V. C. (1983) *Ann. Rev. Biochem.* **52**, 537 – 579

67. Kolodziej, S. J., Penczek, P. A., Schroeter, J. P., and Stoops, J. K. (1996) *J. Biol. Chem.* **271**, 28422 – 28429

68. Orme, T. W., McIntyre, J., Lynen, F., Kühn, L., and Schweizer, E. (1972) *Eur. J. Biochem.* **24**, 407 – 415

69. Poulose, A. J., Rogers, L., Cheesbrough, T. M., and Kolattukudy, P. E. (1985) *J. Biol. Chem.* **260**, 15953 – 15958

70. Witkowska, H. E., Green, B. N., and Smith, S. (1990) *J. Biol. Chem.* **265**, 5662 – 5665

71. Smith, S., and Abraham, S. (1975) *Adv. Lipid Res.* **13**, 195 – 239

72. Stumpf, P. K. (1981) *Trends Biochem. Sci.* **6**, 173 – 176

72a. Murakami, Y., Tsuyama, M., Kobayashi, Y., Kodama, H., and Iba, K. (2000) *Science* **287**, 476 – 479

73. Ohlrogge, J. B., Jaworski, J. G., and Post-Beittenmiller, D. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 3 – 32, CRC Press, Boca Raton, Florida

74. Heinz, E. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 33 – 90, CRC Press, Boca Raton, Florida

75. Stumpf, P. K. (1984) in *Fatty Acid Metabolism and its Regulation: New Comprehensive Biochemistry*, Vol. 7 (Numa, S., ed.), pp. 155 – 179, Elsevier, Amsterdam

75a. Lyle, K. S., Möenne-Loccoz, P., Ai, J., Sanders-Loehr, J., Loehr, T. M., and Fox, B. G. (2000) *Biochemistry* **39**, 10507 – 10513

76. Choi, J.-Y., and Martin, C. E. (1999) *J. Biol. Chem.* **274**, 4671 – 4683

77. Watkins, P. A., Lu, J.-F., Steinberg, S. J., Gould, S. J., Smith, K. D., and Braiterman, L. T. (1998) *J. Biol. Chem.* **273**, 18210 – 18219

78. Whereat, A. F., Orishimo, M. W., Nelson, J., and Phillips, S. J. (1969) *J. Biol. Chem.* **244**, 6498 – 6506

78a. Das, A. K., Uhler, M. D., and Hajra, A. K. (2000) *J. Biol. Chem.* **275**, 24333 – 24340

79. Lin, C. Y., and Kumar, S. (1971) *J. Biol. Chem.* **246**, 3284 – 3290

80. Varanasi, U., Feldman, H. R., and Malins, D. C. (1975) *Nature (London)* **255**, 340 – 343

81. Kolattukudy, P. E. (1968) *Science* **159**, 498 – 505

82. Dillwith, J. W., Nelson, J. H., Pomonis, J. G., Nelson, D. R., and Blomquist, G. J. (1982) *J. Biol. Chem.* **257**, 11305 – 11314

83. Odham, G., and Stenhagen, E. (1971) *Acc. Chem. Res.* **4**, 121 – 128

84. Cole, S. T., and and 41 other authors. (1998) *Nature (London)* **393**, 537 – 544

85. Azad, A. K., Sirakova, T. D., Fernandes, N. D., and Kolattukudy, P. E. (1997) *J. Biol. Chem.* **272**, 16741 – 16745

86. Kroumova, A. B., Xie, Z., and Wagner, G. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11437 – 11441

87. Howell, D. M., Harich, K., Xu, H., and White, R. H. (1998) *Biochemistry* **37**, 10108 – 10117

88. Hayden, M. A., Huang, I. Y., Iliopoulos, G., Orozco, M., and Ashley, G. W. (1993) *Biochemistry* **32**, 3778 – 3782

89. Sulo, P., and Martin, N. C. (1993) *J. Biol. Chem.* **268**, 17634 – 17639

90. Wada, H., Shintani, D., and Ohlrogge, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1591 – 1596

91. Busby, R. W., Schelvis, J. P. M., Yu, D. S., Babcock, G. T., and Marletta, M. A. (1999) *J. Am. Chem. Soc.* **121**, 4706 – 4707

91a. Miller, J. R., Busby, R. W., Jordan, S. W., Cheek, J., Henshaw, T. F., Ashley, G. W., Broderick, J. B., Cronan, J. E., Jr., and Marletta, M. A. (2000) *Biochemistry* **39**, 15166 – 15178

91b. Gueguent, V., Macherel, D., Jaquinod, M., Douce, R., and Bourguignon, J. (2000) *J. Biol. Chem.* **275**, 5016 – 5025

92. Duin, E. C., Lafferty, M. E., Crouse, B. R., Allen, R. M., Sanyal, I., Flint, D. H., and Johnson, M. K. (1997) *Biochemistry* **36**, 11811 – 11820

93. Hayden, M. A., Huang, I., Bussiere, D. E., and Ashley, G. W. (1992) *J. Biol. Chem.* **267**, 9512 – 9515

93a. Ugulava, N. B., Sacanell, C. J., and Jarrett, J. T. (2001) *Biochemistry* **40**, 8352 – 8358

94. White, R. H. (1987) *Biochemistry* **26**, 3163 – 3167

95. deMendoza, D., and Cronan, J. E., Jr. (1983) *Trends Biochem. Sci.* **8**, 49 – 52

95a. Wallis, J. G., Watts, J. L., and Browse, J. (2002) *Trends Biochem. Sci.* **27**, 467–473

96. Henderson, B. S., Larsen, B. S., and Schwab, J. M. (1994) *J. Am. Chem. Soc.* **116**, 5025 – 5034

97. Roughan, G., and Slack, R. (1984) *Trends Biochem. Sci.* **9**, 383 – 386

98. Sperling, P., Linscheid, M., Stöcker, S., Mühlback, H.-P., and Heinz, E. (1993) *J. Biol. Chem.* **268**, 26935 – 26940

99. Jeffcoat, R., and James, A. T. (1984) in *Fatty Acid Metabolism and its Regulation: New Comprehensive Biochemistry*, Vol. 7 (Numa, S., ed), pp. 85 – 112, Elsevier, Amsterdam

100. Thiede, M. A., and Strittmatter, P. (1985) *J. Biol. Chem.* **260**, 14459 – 14463

101. Shanklin, J., Whittle, E., and Fox, B. G. (1994) *Biochemistry* **33**, 12787 – 12794

102. Buist, P. H., and Behrouzian, B. (1998) *J. Am. Chem. Soc.* **120**, 871 – 876

103. Vergères, G., Ramsden, J., and Waskell, L. (1995) *J. Biol. Chem.* **270**, 3414 – 3422

104. Mitchell, A. G., and Martin, C. E. (1995) *J. Biol. Chem.* **270**, 29766 – 29772

105. McKeon, T. A., and Stumpf, P. K. (1982) *J. Biol. Chem.* **257**, 12141 – 12147

106. Somerville, C., and Browse, J. (1991) *Science* **252**, 80 – 87

107. Töpfer, R., Martini, N., and Schell, J. (1995) *Science* **268**, 681 – 686

108. Browse, J., McConn, M., James, D., Jr., and Miquel, M. (1993) *J. Biol. Chem.* **268**, 16345 – 16351

109. Lindqvist, Y., Huang, W., Schneider, G., and Shanklin, J. (1996) *EMBO J.* **15**, 4081 – 4092

110. Fox, B. G., Shanklin, J., Ai, J., Loehr, T. M., and Sanders-Loehr, J. (1994) *Biochemistry* **33**, 12776 – 12786

111. Spychalla, J. P., Kinney, A. J., and Browse, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1142 – 1147

112. Mustardy, L., Los, D. A., Gombos, Z., and Murata, N. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10524 – 10527

113. Broun, P., Shanklin, J., Whittle, E., and Somerville, C. (1998) *Science* **282**, 1315 – 1317

114. van de Loo, F. J., Broun, P., Turner, S., and Somerville, C. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6743 – 6747

115. Kaya, K., Ramesha, C. S., and Thompson, G. A., Jr. (1984) *J. Biol. Chem.* **259**, 3548 – 3553

116. Ortiz de Montellano, P. R., Chan, W. K., Tuck, S. F., Kaikaus, R. M., Bass, N. M., and Peterson, J. A. (1992) *FASEB J.* **6**, 695 – 699

117. Vaz, A. D. N., McGinnity, D. F., and Coon, M. J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3555 – 3560

118. de Renobales, M., Cripps, C., Stanley-Samuelson, D. W., Jurenka, R. A., and Blomquist, G. J. (1987) *Trends Biochem. Sci.* **12**, 364 – 366

119. Cho, H. P., Nakamura, M. T., and Clarke, S. D. (1999) *J. Biol. Chem.* **274**, 471 – 477

119a. Bernoud-Hubac, N., Davies, S. S., Boutaud, O., Montine, T. J., and Roberts, L. J., II. (2001) *J. Biol. Chem.* **276**, 30964 – 30970

119b. Mirnikjoo, B., Brown, S. E., Kim, H. F. S., Marangell, L. B., Sweatt, J. D., and Weeber, E. J. (2001) *J. Biol. Chem.* **276**, 10888 – 10896

119c. Brzustowicz, M. R., Cherezov, V., Zerouga, M., Caffrey, M., Stillwell, W., and Wassall, S. R. (2002) *Biochemistry* **41**, 12509–12519

119d. Albert, C. M., Campos, H., Stampfer, M. J., Ridker, P. M., Manson, J. E., Willett, W. C., and Ma, J. (2002) *M. Engl. J. Med.* **346**, 1113–1118 (and discussion in Vol. 347, pp. 531–533)

120. Knipple, D. C., Rosenfield, C.-L., Miller, S. J., Liu, W., Tang, J., Ma, P. W. K., and Roelofs, W. L. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15287 – 15292

121. Aveldaño, M. I., and Sprecher, H. (1997) *J. Biol. Chem.* **262**, 1180 – 1186

121a. Cahoon, E. B., Ripp, K. G., Hall, S. E., and Kinney, A. J. (2001) *J. Biol. Chem.* **276**, 2637 – 2643

121b. Adlof, R. O., Duval, S., and Emken, E. A. (2000) *Lipids* **35**, 131 – 135

121c. Majumder, B., Wahle, K. W. J., Moir, S., Schofield, A., Choe, S.-N., Farquharson, A., Grant, I., and Heys, S. D. (2002) *FASEB J.* **16**, 1447-1449

122. Wise, M. L., Hamberg, M., and Gerwick, W. H. (1994) *Biochemistry* **33**, 15223 – 15232

123. Wise, M. L., Rossi, J., and Gerwick, W. H. (1997) *Biochemistry* **36**, 2985 – 2992

124. Cox, G. S., Thomas, E., Kaback, H. R., and Weissbach, H. (1973) *Arch. Biochem. Biophys.* **158**, 667 – 676

125. Wang, A.-Y., Grogan, D. W., and Cronan, J. E., Jr. (1992) *Biochemistry* **31**, 11020 – 11028

126. Packter, N. M. (1973) *Biosynthesis of Acetate-Derived Compounds*, Wiley, New York

126a. Glickman, M. S., Cahill, S. M., and Jacobs, W. R., Jr. (2001) *J. Biol. Chem.* **276**, 2228 – 2233

127. Yuan, Y., Crane, D. C., Musser, J. M., Sreevatsan, S., and Barry, C. E., III. (1997) *J. Biol. Chem.* **272**, 10041 – 10049

128. Yuan, Y., Lee, R. E., Besra, G. S., Belisle, J. T., and Barry, C. E., III. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6630 – 6634

129. Liu, J., and Nikaido, H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4011 – 4016

130. Tipton, C. L., and Al-Shathir, N. M. (1974) *J. Biol. Chem.* **249**, 886 – 889

131. Bohlmann, F., Burkhardt, T., and Zdero, C. (1973) *Naturally Occuring Acetylenes*, Academic Press, New York

References

132. Nicolaides, N. (1974) *Science* **186**, 19 – 26
133. Bauermeister, A., and Sargent, J. R. (1979) *Trends Biochem. Sci.* **4**, 209 – 211
134. Nemes, Z., Marekov, L. N., Fésüs, L., and Steinert, P. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8402 – 8407
135. Kolattukudy, P. E. (1980) *Science* **208**, 990 – 1000
135a. Wellesen, K., Durst, F., Pinot, F., Benveniste, I., Nettesheim, K., Wisman, E., Steiner-Lange, S., Saedler, H., and Yephremov, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9694 – 9699
136. von Wettstein-Knowles, P. M. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 127 – 166, CRC Press, Boca Raton, Florida
137. Dennis, M., and Kolattukudy, P. E. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5306 – 5310
138. Cheesbrough, T. M., and Kolattukudy, P. E. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6613 – 6617
139. Cheesbrough, T. M., and Kolattukudy, P. E. (1988) *J. Biol. Chem.* **263**, 2738 – 2743
140. Reed, J. R., Vanderwel, D., Choi, S., Pomonis, J. G., Reitz, R. C., and Blomquist, G. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10000 – 10004
141. Gorgen, G., and Boland, W. (1989) *Eur. J. Biochem.* **185**, 237 – 242
142. Buckner, J. S., Nelson, D. R., Fatland, C. L., Hakk, H., and Pomonis, J. G. (1984) *J. Biol. Chem.* **259**, 8461 – 8470
143. Morse, D., and Meighen, E. (1984) *Science* **226**, 1434 – 1436
144. Hölldobler, B. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 19 – 22
145. Prestwich, G. D. (1987) *Science* **237**, 999 – 1006
146. Dussourd, D. E., Harvis, C. A., Meinwald, J., and Eisner, T. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9224 – 9227
146a. Athenstaedt, K., and Daum, G. (2000) *J. Biol. Chem.* **275**, 235 – 240
146b. Kalhan, S. C., Mahajan, S., Burkett, E., Reshef, L., and Hanson, R. W. (2001) *J. Biol. Chem.* **276**, 12928 – 12931
147. Rock, C. O., Goelz, S. E., and Cronan, J. E., Jr. (1981) *J. Biol. Chem.* **256**, 736 – 742
148. Eberhardt, C., Gray, P. W., and Tjoelker, L. W. (1997) *J. Biol. Chem.* **272**, 20299 – 20305
149. Guo, Z., Liliom, K., Fischer, D. J., Bathurst, I. C., Tomei, L. D., Kiefer, M. C., and Tigyi, A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14367 – 14372
150. Kai, M., Wada, I., Imai, S.-i, Sakane, F., and Kanoh, H. (1997) *J. Biol. Chem.* **272**, 24572 – 24578
150a. Buhman, K. K., Chen, H. C., and Farese, R. V., Jr. (2001) *J. Biol. Chem.* **276**, 40369 – 40372
150b. Abo-Hashema, K. A. H., Cake, M. H., Power, G. W., and Clarke, D. (1999) *J. Biol. Chem.* **274**, 35577 – 35582
151. Shen, H., Heacock, P. N., Clancey, C. J., and Dowhan, W. (1996) *J. Biol. Chem.* **271**, 789 – 795
152. Badola, P., and Sanders, C. R., II. (1997) *J. Biol. Chem.* **272**, 24176 – 24182
153. Maréchal, E., Block, M. A., Joyard, J., and Douce, R. (1994) *J. Biol. Chem.* **269**, 5788 – 5798
154. Joyard, J., Block, M. A., Malherbe, A., Maréchal, E., and Douce, R. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 231 – 258, CRC Press, Boca Raton, Florida
155. Dörmann, P., Balbo, I., and Benning, C. (1999) *Science* **284**, 2181 – 2184
156. Karlsson, O. P., Rytömaa, M., Dahlqvist, A., Kinnunen, P. K. J., and Wieslander, A. (1996) *Biochemistry* **35**, 10094 – 10102
156a. Abu-Elheiga, L., Matzuk, M. M., Abo-Hashema, K. A. H., and Wakil, S. J. (2001) *Science* **291**, 2613 – 2616
156b. Jelenska, J., Crawford, M. J., Harb, O. S., Zuther, E., Haselkorn, R., Roos, D. S., and Gornicki, P. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2723 – 2728
157. Thompson, G. A., Jr. (1992) *The Regulation of Membrane Lipid Metabolism*, 2nd ed., CRC Press, Boca Raton, Florida
158. Lopez, J. M., Bennett, M. K., Sanchez, H. B., Rosenfeld, J. M., and Osborne, T. F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1049 – 1053
159. Honnor, R. C., Dhillon, G. S., and Londos, C. (1985) *J. Biol. Chem.* **260**, 15130 – 15138
160. Tabor, D. E., Kim, J. B., Spiegelman, B. M., and Edwards, P. A. (1999) *J. Biol. Chem.* **274**, 20603 – 20610
161. Soncini, M., Yet, S.-F., Moon, Y., Chun, J.-Y., and Sul, H. S. (1995) *J. Biol. Chem.* **270**, 30339 – 30343
162. Dudek, S. M., and Semenkovich, C. F. (1995) *J. Biol. Chem.* **270**, 29323 – 29329
163. Tiku, P. E., Gracey, A. Y., Macartney, A. I., Beynon, R. J., and Cossins, A. R. (1996) *Science* **271**, 815 – 818
164. Jackson, B. J., Gennity, J. M., and Kennedy, E. P. (1986) *J. Biol. Chem.* **261**, 13464 – 13468
165. Raetz, C. R. H., and Dowhan, W. (1990) *J. Biol. Chem.* **265**, 1235 – 1238
166. Kennedy, E. P. (1992) *Ann. Rev. Biochem.* **61**, 1 – 28
167. Bae-Lee, M. S., and Carman, G. M. (1984) *J. Biol. Chem.* **259**, 10857 – 10862
168. Kawasaki, K., Kuge, O., Chang, S.-C., Heacock, P. N., Rho, M., Suzuki, K., Nishijima, M., and Dowhan, W. (1999) *J. Biol. Chem.* **274**, 1828 – 1834
168a. Auer, M., Kim, M. J., Lemieux, M. J., Villa, A., Song, J., Li, X.-D., and Wang, D.-N. (2001) *Biochemistry* **40**, 6628 – 6635
169. Louie, K., and Dowhan, W. (1980) *J. Biol. Chem.* **255**, 1124 – 1127
169a. Stone, S. J., and Vance, J. E. (2000) *J. Biol. Chem.* **275**, 34534 – 34540
170. Trotter, P. J., and Voelker, D. R. (1995) *J. Biol. Chem.* **270**, 6062 – 6070
171. Shiao, Y.-J., Lupo, G., and Vance, J. E. (1995) *J. Biol. Chem.* **270**, 11190 – 11198
172. Heikinheimo, L., and Somerharju, P. (1998) *J. Biol. Chem.* **273**, 3327 – 3335
173. Huijbregts, R. P. H., de Kroon, A. I. P. M., and de Kruijff, B. (1998) *J. Biol. Chem.* **273**, 18936 – 18942
174. Rietveld, A. G., Chupin, V. V., Koorengevel, M. C., Wienk, H. L. J., Dowhan, W., and de Kruijff, B. (1994) *J. Biol. Chem.* **269**, 28670 – 28675
175. Bogdanov, M., Umeda, M., and Dowhan, W. (1999) *J. Biol. Chem.* **274**, 12339 – 12345
176. Ohtsuka, T., Nishijima, M., and Akamatsu, Y. (1993) *J. Biol. Chem.* **268**, 22908 – 22911
177. Zhao, M., Schlame, M., Rua, D., and Greenberg, M. L. (1998) *J. Biol. Chem.* **273**, 2402 – 2408
178. Chang, S.-C., Heacock, P. N., Mileykovskaya, E., Voelker, D. R., and Dowhan, W. (1998) *J. Biol. Chem.* **273**, 14933 – 14941
179. McMaster, C. R., and Bell, R. M. (1994) *J. Biol. Chem.* **269**, 14776 – 14783
180. Kent, C., and Carman, G. M. (1999) *Trends Biochem. Sci.* **24**, 146 – 150
180a. Dowd, S. R., Bier, M. E., and Patton-Vogt, J. L. (2001) *J. Biol. Chem.* **276**, 3756 – 3763
180b. Lykidis, A., Wang, J., Karim, M. A., and Jackowski, S. (2001) *J. Biol. Chem.* **276**, 2174 – 2179
181. Vance, D. E., and Pelech, S. L. (1984) *Trends Biochem. Sci.* **9**, 17 – 20
181a. Henneberry, A. L., Wistow, G., and McMaster, C. R. (2000) *J. Biol. Chem.* **275**, 29808 – 29815
181b. Cornell, R. B., and Northwood, I. C. (2000) *Trends Biochem. Sci.* **25**, 441 – 447
181c. Hunt, A. N., Clark, G. T., Attard, G. S., and Postle, A. D. (2001) *J. Biol. Chem.* **276**, 8492 – 8499
182. Kuge, O., Saito, K., and Nishijima, M. (1997) *J. Biol. Chem.* **272**, 19133 – 19139
183. Vance, J. E. (1998) *Trends Biochem. Sci.* **23**, 423 – 428
184. Zeisel, S. H., da Costa, K.-A., Franklin, P. D., Alexander, E. A., LaMont, J. T., Sheard, N. F., and Beiser, A. (1991) *FASEB J.* **5**, 2093 – 2098
185. Lee, H.-C., Fellenz-Maloney, M.-P., Liscovitch, M., and Krzysztof Blusztajn, J. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10086 – 10090
186. Nikawa, J.-i, Kodaki, T., and Yamashita, S. (1987) *J. Biol. Chem.* **2 262**, 4876 – 4881
186a. Ilgoutz, S. C., Zawadzki, J. L., Ralton, J. E., and McConville, M. J. (1999) *EMBO J.* **18**, 2746 – 2755
187. Wu, W.-I., and Carman, G. M. (1996) *Biochemistry* **35**, 3790 – 3796
188. Griac, P., Swede, M. J., and Henry, S. A. (1996) *J. Biol. Chem.* **271**, 25692 – 25698
189. Kim, K.-H., Voelker, D. R., Flocco, M. T., and Carman, G. M. (1998) *J. Biol. Chem.* **273**, 6844 – 6852
190. Ostrander, D. B., O'Brien, D. J., Gorman, J. A., and Carman, G. M. (1998) *J. Biol. Chem.* **273**, 18992 – 19001
191. Wieprecht, M., Wieder, T., Paul, C., Geilen, C. C., and Orfanos, C. E. (1996) *J. Biol. Chem.* **271**, 9955 – 9961
192. Snyder, F., ed. (1972) *Ether Lipids: Chemistry and Biology*, Academic Press, New York
193. Lee, T.-c, Malone, B., and Snyder, F. (1988) *J. Biol. Chem.* **263**, 1755 – 1760
194. Karasawa, K., Qiu, X., and Lee, T.-c. (1999) *J. Biol. Chem.* **274**, 8655 – 8661
194a. Min, J.-H., Wilder, C., Aoki, J., Arai, H., Inoue, K., Paul, L., and Gelb, M. H. (2001) *Biochemistry* **40**, 4539 – 4549
195. Ford, D. A., Rosenbloom, K. B., and Gross, R. W. (1992) *JBC* **267**, 11222 – 11228
196. Langworthy, T. A. (1985) in *The Bacteria* (Woese, C. R., and Wolfe, R. S., eds), pp. 459 – 497, Academic Press, New York
197. Jones, W. J., Nagle, D. P., Jr., and Whitman, W. B. (1989) *Microbiol. Rev.* **51**, 135 – 177
198. DeRosa, M. A., Gambacorta, A., and Gliozzi, A. (1986) *Microbiol. Rev.* **50**, 70 – 80
199. Burdett, K., Larkins, L. K., Das, A. K., and Hajra, A. K. (1991) *J. Biol. Chem.* **266**, 12201 – 12206
200. Friedberg, S. J., Weintraub, S. T., Singer, M. R., and Greene, R. C. (1982) *J. Biol. Chem.* **258**, 136 – 142
201. Brown, A. J., and Snyder, F. (1983) *J. Biol. Chem.* **258**, 4184 – 4189
202. Paltauf, F., and Holasek, A. (1973) *J. Biol. Chem.* **248**, 1609 – 1615
203. Borst, P. (1983) *Trends Biochem. Sci.* **8**, 269 – 272
204. de Vet, E. C. J. M., Ijlst, L., Oostheim, W., Wanders, R. J. A., and van den Bosch, H. (1998) *J. Biol. Chem.* **273**, 10296 – 10301
205. Moser, A. E., Singh, I., Brown, F. R., Solish, G. I., Kelley, R. I., Benke, P. J., and Moser, H. W. (1984) *N. Engl. J. Med.* **310**, 1141 – 1146
206. Lazarow, P. B., and Moser, H. W. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2287 – 2324, McGraw-Hill, New York
207. Chao, W., Liu, H., Hanahan, D. J., and Olson, M. S. (1992) *J. Biol. Chem.* **267**, 6725 – 6735
208. Lynch, D. V. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 285 – 308, CRC Press, Boca Raton, Florida

209. Barenholz, Y., and Gatt, S. (1982) in *New Comprehensive Biochemistry*, Vol. 4 (Hawthorne, J. N., and Ansell, G. B., eds), pp. 129–177, Elsevier, Amsterdam

209a. Ikushiro, H., Hayashi, H., and Kagamiyama, H. (2001) *J. Biol. Chem.* **276**, 18249–18256

210. Stoffel, W., and Melzner, I. (1980) *Z. Physiol. Chem.* **361**, 755–771

211. Grilley, M. M., Stock, S. D., Dickson, R. C., Lester, R. L., and Takemoto, J. Y. (1998) *J. Biol. Chem.* **273**, 11062–11068

212. Sperling, P., Zähringer, U., and Heinz, E. (1998) *J. Biol. Chem.* **273**, 28590–28596

213. Toledo, M. S., Levery, S. B., Straus, A. H., Suzuki, E., Momany, M., Glushka, J., Moulton, J. M., and Takahashi, H. K. (1999) *Biochemistry* **38**, 7294–7306

213a. Savile, C. K., Fabriàs, G., and Buist, P. H. (2001) *J. Am. Chem. Soc.* **123**, 4382–4385

214. Yamakawa, T. (1988) *Trends Biochem. Sci.* **13**, 452–454

215. van Echten, G., and Sandhoff, K. (1993) *J. Biol. Chem.* **268**, 5341–5344

216. Wiegandt, H. (1971) *Adv. Lipid Res.* **9**, 249–289

217. Hakomori, S.-i. (1990) *J. Biol. Chem.* **265**, 18713–18716

218. Takamiya, K., Yamamoto, A., Furukawa, K., Zhao, J., Fukumoto, S., Yamashiro, S., Okada, M., Haraguchi, M., Shin, M., Kishikawa, M., Shiku, H., Aizawa, S., and Furukawa, K. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12147–12152

218a. Mylvaganam, M., and Lingwood, C. A. (1999) *J. Biol. Chem.* **274**, 20725–20732

218b. Uragami, M., Tokutake, N., Yan, X., and Regen, S. L. (2001) *J. Am. Chem. Soc.* **123**, 5124–5125

218c. Brown, D. A., and London, E. (2000) *J. Biol. Chem.* **275**, 17221–17224

219. Rameh, L. E., and Cantley, L. C. (1999) *J. Biol. Chem.* **274**, 8347–8350

220. Anderson, R. A., Boronenkov, I. V., Doughman, S. D., Kunz, J., and Loijens, J. C. (1999) *J. Biol. Chem.* **274**, 9907–9910

221. Rhee, S. G., and Bae, Y. S. (1997) *J. Biol. Chem.* **272**, 15045–15048

222. Heinz, D. W., Ryan, M., Bullock, T. L., and Griffith, O. H. (1995) *EMBO J.* **14**, 3855–3863

223. Essen, L.-O., Perisic, O., Katan, M., Wu, Y., Roberts, M. F., and Williams, R. L. (1997) *Biochemistry* **36**, 1704–1718

224. Hondal, R. J., Zhao, Z., Kravchuk, A. V., Liao, H., Riddle, S. R., Yue, X., Bruzik, K. S., and Tsai, M.-D. (1998) *Biochemistry* **37**, 4568–4580

225. Majerus, P. W., Kisseleva, M. V., and Norris, F. A. (1999) *J. Biol. Chem.* **274**, 10669–10672

226. Newton, A. C. (1995) *J. Biol. Chem.* **270**, 28495–28498

227. Newton, A. C., and Johnson, J. E. (1998) *Biochim. Biophys. Acta.* **1376**, 155–172

228. Epand, R. M., Stevenson, C., Bruins, R., Schram, V., and Glaser, M. (1998) *Biochemistry* **37**, 12068–12073

228a. McIntyre, T. M., Zimmerman, G. A., and Prescott, S. M. (1999) *J. Biol. Chem.* **274**, 25189–25192

229. Hodgkin, M. N., Pettitt, T. R., Martin, A., Michell, R. H., Pemberton, A. J., and Wakelam, M. J. O. (1998) *Trends Biochem. Sci.* **23**, 200–204

230. Topham, M. K., and Prescott, S. M. (1999) *J. Biol. Chem.* **274**, 11447–11450

230a. Jones, D. R., Pettitt, T. R., Sanjuán, M. A., Mérida, I., and Wakelam, M. J. O. (1999) *J. Biol. Chem.* **274**, 16846–16852

231. Walsh, J. P., Suen, R., and Glomset, J. A. (1995) *J. Biol. Chem.* **270**, 28647–28653

232. Schissel, S. L., Keesler, G. A., Schuchman, E. H., Williams, K. J., and Tabas, I. (1998) *J. Biol. Chem.* **273**, 18250–18259

233. Testi, R. (1996) *Trends Biochem. Sci.* **21**, 468–471

234. Spiegel, S., and Merrill, A. H., Jr. (1996) *FASEB J.* **10**, 1388–1397

235. Hannun, Y. A., and Obeid, L. M. (1995) *Trends Biochem. Sci.* **20**, 73–77

236. Usta, J., El Bawab, S., Roddy, P., Szulc, Z. M., Hannun, Y. A., and Bielawska, A. (2001) *Biochemistry* **40**, 9657–9668

236a. Hannun, Y. A., Luberto, C., and Argraves, K. M. (2001) *Biochemistry* **40**, 4893–4903

237. Mandala, S. M., Thornton, R., Tu, Z., Kurtz, M. B., Nickels, J., Broach, J., Menzeleev, R., and Spiegel, S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 150–155

237a. Liu, Y.-Y., Han, T.-Y., Giuliano, A. E., and Cabot, M. C. (2001) *FASEB J.* **15**, 719–730

237b. Windh, R. T., Lee, M.-J., Hla, T., An, S., Barr, A. J., and Manning, D. R. (1999) *J. Biol. Chem.* **274**, 27351–27358

237c. Ammit, A. J., Hastie, A. T., Edsall, L. C., Hoffman, R. K., Amrani, Y., Krymskaya, V. P., Kane, S. A., Peters, S. P., Penn, R. B., Spiegel, S., and Panettieri, R. A., Jr. (2001) *FASEB J.* **15**, 1212–1214

237d. Dickson, R. C. (1998) *Ann. Rev. Biochem.* **67**, 27–48

237e. Hla, T., Lee, M.-J., Ancellin, N., Paik, J. H., and Kluk, M. J. (2001) *Science* **294**, 1875–1878

238. Westerman, J., de Vries, K.-J., Somerharju, P., Timmermans-Hereijgers, J. L. P. M., Snoek, G. T., and Wirtz, K. W. A. (1995) *J. Biol. Chem.* **270**, 14263–14266

238a. Yoder, M. D., Thomas, L. M., Tremblay, J. M., Oliver, R. L., Yarbrough, L. R., and Helmkamp, G. M., Jr. (2001) *J. Biol. Chem.* **276**, 9246–9252

239. Gutteridge, J. M. C., and Halliwell, B. (1990) *Trends Biochem. Sci.* **15**, 129–135

240. Porter, N. A., and Wujek, D. G. (1984) *J. Am. Chem. Soc.* **106**, 2626–2629

241. Wagner, B. A., Buettner, G. R., and Burns, C. P. (1994) *Biochemistry* **33**, 4449–4453

242. Al-Abed, Y., Liebich, H., Voelter, W., and Bucala, R. (1996) *J. Biol. Chem.* **271**, 2892–2896

243. Uchida, K., Shiraishi, M., Naito, Y., Torii, Y., Nakamura, Y., and Osawa, T. (1999) *J. Biol. Chem.* **274**, 2234–2242

243a. Burczynski, M. E., Sridhar, G. R., Palackal, N. T., and Penning, T. M. (2001) *J. Biol. Chem.* **276**, 2890–2897

244. Gardner, H. W., and Hamberg, M. (1993) *J. Biol. Chem.* **268**, 6971–6977

244a. Schneider, C., Tallman, K. A., Porter, N. A., and Brash, A. R. (2001) *J. Biol. Chem.* **276**, 20831–20838

245. Kato, Y., Mori, Y., Makino, Y., Morimitsu, Y., Hiroi, S., Ishikawa, T., and Osawa, T. (1999) *J. Biol. Chem.* **274**, 20406–20414

246. Tsai, L., Szweda, P. A., Vinogradova, O., and Szweda, L. I. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7975–7980

247. McGirr, L. G., Hadley, M., and Draper, H. H. (1985) *J. Biol. Chem.* **260**, 15427–15431

248. Hadley, M., and Draper, H. H. (1988) *FASEB J.* **2**, 138–140

249. Slatter, D. A., Paul, R. G., Murray, M., and Bailey, A. J. (1999) *J. Biol. Chem.* **274**, 19661–19669

250. Armstrong, D., Dimmitt, S., Boehme, D. H., Leonberg, S. C., Jr., and Vogel, W. (1974) *Science* **186**, 155–156

250a. Onorato, J. M., Jenkins, A. J., Thorpe, S. R., and Baynes, J. W. (2000) *J. Biol. Chem.* **275**, 21177–21184

251. Hill, E. G., Johnson, S. B., Lawson, L. D., Mahfouz, M. M., and Holman, R. T. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 953–957

252. Holman, R. T., Pusch, F., Svingen, B., and Dutton, H. J. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4830–4834

253. Mensink, R. P., and Katan, M. B. (1990) *N. Engl. J. Med.* **323**, 439–445

253a. Kopelman, P. G. (2000) *Nature (London)* **404**, 635–643

253b. Unger, R. H., and Orci, L. (2001) *FASEB J.* **15**, 312–321

253c. Taubes, G. (2001) *Science* **291**, 2536–2545

253d. Koo, S.-H., Dutcher, A. K., and Towle, H. C. (2001) *J. Biol. Chem.* **276**, 9437–9445

253e. Weigle, D. S. (1994) *FASEB J.* **8**, 302–310

253f. Barsh, G. S., Farooqi, I. S., and O'Rahilly, S. (2000) *Nature (London)* **404**, 644–651

253g. Harris, R. B. S. (1990) *FASEB J.* **4**, 3310–3318

253h. Ruderman, N., and Flier, J. S. (2001) *Science* **291**, 2558–2559

253i. Lee, J.-J., Moon, Y.-A., Ha, J.-H., Yoon, D.-J., Ahn, Y.-H., and Kim, K.-S. (2001) *J. Biol. Chem.* **276**, 2576–2585

253j. Gura, T. (1998) *Science* **280**, 1369–1370

253k. Jucker, B. M., Ren, J., Dufour, S., Cao, X., Previs, S. F., Cadman, K. S., and Shulman, G. I. (2000) *J. Biol. Chem.* **275**, 39279–39286

253l. Clément, K., Vaisse, C., Lahlou, N., Cabrol, S., Pelloux, V., Cassuto, D., Gourmelen, M., Dina, C., Chambaz, J., Lacorte, J.-M., Basdevant, A., Bougnères, P., Lebouc, Y., Froguel, P., and Guy-Grand, B. (1998) *Nature (London)* **392**, 398–401

253m. Gura, T. (2000) *Science* **287**, 1738–1741

253n. Lee, Y., Wang, M.-Y., Kakuma, T., Wang, Z.-W., Babcock, E., McCorkle, K., Higa, M., Zhou, Y.-T., and Unger, R. H. (2001) *J. Biol. Chem.* **276**, 5629–5635

254. Makita, K., Falck, J. R., and Capdevila, J. H. (1996) *FASEB J.* **10**, 1456–1463

255. Schewe, T., and Kühn, H. (1991) *Trends Biochem. Sci.* **16**, 369–73

256. Capdevila, J. H., Falck, J. R., and Estabrook, R. W. (1992) *FASEB J.* **6**, 731–736

256a. Funk, C. D. (2001) *Science* **294**, 1871–1875

257. Bergström, S. (1967) *Science* **157**, 382–391

258. Samuelsson, B., Granström, E., Green, K., Hamberg, M., and Hammarström, S. (1975) *Ann. Rev. Biochem.* **44**, 669–695

259. Newton, R. F., and Roberts, S. M., eds. (1982) *Prostaglandins and Thromboxanes*, Butterworth Scientific, London

260. Neufeld, E. J., Bross, T. E., and Majerus, P. W. (1984) *J. Biol. Chem.* **259**, 1986–1992

261. Kramer, R. M., and Deykin, D. (1983) *J. Biol. Chem.* **258**, 13806–13811

262. Picot, D., Loll, P. J., and Garavito, R. M. (1994) *Nature (London)* **367**, 243–249

262a. Seibold, S. A., Cerda, J. F., Mulichak, A. M., Song, I., Garavito, R. M., Arakawa, T., Smith, W. L., and Babcock, G. T. (2000) *Biochemistry* **39**, 6616–6624

262b. Klefer, J. R., Pawlitz, J. L., Moreland, K. T., Stegeman, R. A., Hood, W. F., Glerse, J. K., Stevens, A. M., Goodwin, D. C., Rowlinson, S. W., Marnett, L. J., Stallings, W. C., and Kurumball, R. G. (2000) *Nature (London)* **405**, 97–101

262c. Thuresson, E. D., Malkowski, M. G., Lakkides, K. M., Rieke, C. J., Mulichak, A. M., Ginell, S. L., Garavito, R. M., and Smith, W. L. (2001) *J. Biol. Chem.* **276**, 10358–10365

262d. Peng. S., Okeley, N. M., Tsai, A.-L., Wu, G., Kulmacz, R. J., and van der Donk. W. A. (2002) *J. Am. Chem. Soc.* **124**, 10785–10796

263. Lambeir, A.-M., Markey, C. M., Dunford, H. B., and Marnett, L. J. (1985) *J. Biol. Chem.* **260**, 14894–14896

264. Tsai, A.-l, Kulmacz, R. J., and Palmer, G. (1995) *J. Biol. Chem.* **270**, 10503–10508

References

265. Smith, W. L., Garavito, R. M., and DeWitt, D. L. (1996) *J. Biol. Chem.* **271**, 33157 – 33160
266. Tsai, A.-l, Palmer, G., Xiao, G., Swinney, D. C., and Kulmacz, R. J. (1998) *J. Biol. Chem.* **273**, 3888 – 3894
266a. Marnett, L. J., Rowlinson, S. W., Goodwin, D. C., Kalgutkar, A. S., and Lanzo, C. A. (1999) *J. Biol. Chem.* **274**, 22903 – 22906
266b. Malkowski, M. G., Ginell, S. L., Smith, W. L., and Garavito, R. M. (2000) *Science* **15**, 1933 – 1937
267. Dean, A. M., and Dean, F. M. (1999) *Protein Sci.* **8**, 1087 – 1098
268. Jakobsson, P.-J., Thorén, S., Morgenstern, R., and Samuelsson, B. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7220 – 7225
269. Williams, D. W., Hale, S. E., Okita, R. T., and Masters, B. S. S. (1984) *J. Biol. Chem.* **259**, 14600 – 14608
270. Hamberg, M., Svensson, J., and Samuelsson, B. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2994 – 2998
271. Hecker, M., and Ullrich, V. (1989) *J. Biol. Chem.* **264**, 141 – 150
272. Ruan, K.-H., Li, D., Ji, J., Lin, Y.-Z., and Gao, X. (1998) *Biochemistry* **37**, 822 – 830
273. Haurand, M., and Ullrich, V. (1985) *J. Biol. Chem.* **260**, 15059 – 15067
274. Moncada, S., Gryglewski, R., Bunting, S., and Vane, J. R. (1976) *Nature (London)* **263**, 663 – 665
275. Vane, J. R., and Bergström, S., eds. (1979) *Prostacyclin*, Raven Press, New York
275a. Cheng, Y., Austin, S. C., Rocca, B., Koller, B. H., Coffman, T. M., Grosser, T., Lawson, J. A., and FitzGerald, G. A. (2002) *Science* **296**, 539– 541
276. Shyue, S.-K., Ruan, K.-H., Wang, L.-H., and Wu, K. K. (1997) *J. Biol. Chem.* **272**, 3657 – 3662
277. Wong, P. Y.-K., Malik, K. U., Taylor, B. M., Schneider, W. P., McGiff, J. C., and Sun, F. F. (1985) *J. Biol. Chem.* **260**, 9150 – 9153
278. Kramer, J. A., Johnson, K. R., Dunham, W. R., Sands, R. H., and Funk, M. O., Jr. (1994) *Biochemistry* **33**, 15017 – 15022
279. Glickman, M. H., and Klinman, J. P. (1996) *Biochemistry* **35**, 12882 – 12892
280. Nelson, M. J., Chase, D. B., and Seitz, S. P. (1995) *Biochemistry* **34**, 6159 – 6163
280a. Brash, A. R. (1999) *J. Biol. Chem.* **274**, 23679 – 23682
280b. Skrzypczak-Jankun, E., Bross, R. A., Carroll, R. T., Dunham, W. R., and Funk, M. O., Jr. (2001) *J. Am. Chem. Soc.* **123**, 10814 – 10820
280c. Knapp, M. J., Seebeck, F. P., and Klinman, J. P. (2001) *J. Am. Chem. Soc.* **123**, 2931 – 2932
280d. Mogul, R., and Holman, T. R. (2001) *Biochemistry* **40**, 4391 – 4397
280e. Tomchick, D. R., Phan, P., Cymborowski, M., Minor, W., and Holman, T. R. (2001) *Biochemistry* **40**, 7509–7517
281. Funk, C. D., Keeney, D. S., Oliw, E. H., Boeglin, W. E., and Brash, A. R. (1996) *J. Biol. Chem.* **271**, 23338 – 23344
282. Hamberg, M., Su, C., and Oliw, E. (1998) *J. Biol. Chem.* **273**, 13080 – 13088
283. Boyington, J. C., Gaffney, B. J., and Amzel, L. M. (1993) *Science* **260**, 1482 – 1486
284. Moiseyev, N., Rucker, J., and Glickman, M. H. (1997) *J. Am. Chem. Soc.* **119**, 3853 – 3860
285. Ford-Hutchinson, A. W., Gresser, M., and Young, R. N. (1994) *Ann. Rev. Biochem.* **63**, 383 – 417
286. Glasgow, W. C., Harris, T. M., and Brash, A. R. (1986) *J. Biol. Chem.* **261**, 200 – 204
287. Hammarström, S. (1983) *Ann. Rev. Biochem.* **52**, 355 – 377
288. Samuelsson, B., Dahlén, S.-E., Lindgren, J. Å., Rouzer, C. A., and Serhan, C. N. (1987) *Science* **237**, 1171 – 1176
289. Samuelsson, B., and Funk, C. D. (1989) *J. Biol. Chem.* **264**, 19469 – 19472
290. Blomster Andberg, M., Hamberg, M., and Haeggström, J. Z. (1997) *J. Biol. Chem.* **272**, 23057 – 23063
291. Carter, B. Z., Wiseman, A. L., Orkiszewski, R., Ballard, K. D., Ou, C.-N., and Lieberman, M. W. (1997) *J. Biol. Chem.* **272**, 12305 – 12310
292. Lewis, R. A., Austen, K. F., and Soberman, R. J. (1990) *N. Engl. J. Med.* **323**, 645 – 655
293. Lynch, K. R., O'Neill, G. P., Liu, Q., Im, D.-S., Sawyer, N., Metters, K. M., Coulombe, N., Abramovitz, M., Figueroa, D. J., Zeng, Z., Connolly, B. M., Bai, C., Austin, C. P., Chateauneuf, A., Stocco, R., Greig, G. M., Kargman, S., Hooks, S. B., Hosfield, E., Williams, D. L., Jr., Ford-Hutchinson, A. W., Caskey, C. T., and Evans, J. F. (1999) *Nature (London)* **399**, 789 – 793
294. Hammarström, S. (1983) *J. Biol. Chem.* **258**, 1427 – 1430
295. Serhan, C. N., Hamberg, M., and Samuelsson, B. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5335 – 5339
296. Rowley, A. F., Lloyd-Evans, P., Barrow, S. E., and Serhan, C. N. (1994) *Biochemistry* **33**, 856 – 863
297. Chang, M. S., Boeglin, W. E., Guengerich, F. P., and Brash, A. R. (1996) *Biochemistry* **35**, 464 – 471
298. Brash, A. R., Porter, A. T., and Maas, R. L. (1985) *J. Biol. Chem.* **260**, 4210 – 4216
298a. Zeldin, D. C. (2001) *J. Biol. Chem.* **276**, 36059 – 36062
298b. Node, K., Huo, Y., Ruan, X., Yang, B., Spiecker, M., Ley, K., Zeldin, D. C., and Liao, J. K. (1999) *Science* **285**, 1276 – 1279
299. Goetzl, E. J., An, S., and Smith, W. L. (1995) *FASEB J.* **9**, 1051 – 1058
300. Urade, Y., Fujimoto, N., and Hayaishi, O. (1985) *J. Biol. Chem.* **260**, 12410 – 12415
301. Brock, T. G., McNish, R. W., and Peters-Golden, M. (1999) *J. Biol. Chem.* **274**, 11660 – 11666
302. MacMillan, D. K., Hill, E., Sala.A, Sigal, E., Shuman, T., Henson, P. M., and Murphy, R. C. (1994) *J. Biol. Chem.* **269**, 26663 – 26668
303. Maddox, J. F., Colgan, S. P., Clish, C. B., Petasis, N. A., Fokin, V. V., and Serhan, C. N. (1998) *FASEB J.* **12**, 487 – 494
304. Dahlén, S.-E., Hansson, G., Hedqvist, P., Björck, T., Granström, E., and Dahlén, B. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1712 – 1716
305. Lindgren, J. Å., Hökfelt, T., Dahlén, S.-E., Patrono, C., and Samuelsson, B. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6212 – 6216
306. Reginato, M. J., Krakow, S. L., Bailey, S. T., and Lazar, M. A. (1998) *J. Biol. Chem.* **273**, 1855 – 1858
307. Nelson, N. A., Kelly, R. C., and Johnson, R. A. (1982) *Chem. Eng. News* **Aug 16**, 30 – 44
308. Marx, J. L. (1979) *Science* **205**, 175 – 176
309. Wheelan, P., and Murphy, R. C. (1995) *J. Biol. Chem.* **270**, 19845 – 19852
310. Weissmann, G. (1991) *Sci. Am.* **264**(Jan), 84 – 90
311. Kalgutkar, A. S., Crews, B. C., Rowlinson, S. W., Garner, C., Seibert, K., and Marnett, L. J. (1998) *Science* **280**, 1268 – 1270
312. Vane, J. (1994) *Nature (London)* **367**, 215 – 217
313. Lecomte, M., Laneuville, O., Ji, C., DeWitt, D. L., and Smith, W. L. (1994) *J. Biol. Chem.* **269**, 13207 – 13215
314. Loll, P. J., Picot, D., Ekabo, O., and Garavito, R. M. (1996) *Biochemistry* **35**, 7330 – 7340
314a. Warner, T. D., and Mitchell, J. A. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13371-13373
315. Abramson, S., Korchak, H., Ludewig, R., Edelson, H., Haines, K., Levin, R. I., Herman, R., Rider, L., Kimmel, S., and Weissman, G. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7227 – 7231
316. Patrono, C. (1994) *N. Engl. J. Med.* **330**, 1287 – 1294
317. DuBois, R. N., Abramson, S. B., Crofford, L., Gupta, R. A., Simon, L. S., van de Putte, L. B. A., and Lipsky, P. E. (1998) *FASEB J.* **12**, 1063 – 1073
317a. Selinsky, B. S., Gupta, K., Sharkey, C. T., and Loll, P. J. (2001) *Biochemistry* **40**, 5172 – 5180
317b. FitzGerald, G. A., and Patrono, C. (2001) *N. Engl. J. Med.* **345**, 433 – 442
317c. Killen, J. P., Nzerue, C. M., Rich, S. A., FitzGerald, G. A., and Patrono, C. (2001) *N. Engl. J. Med.* **345**, 1708 – 1709
317d. Vane, J. R. (2002) *Science* **296**, 474–475
317e. Jiang, Q., Elson-Schwab, I., Courtemanche, C., and Ames, B. N. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11494 – 11499
318. Vick, B. A. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 167 – 194, CRC Press, Boca Raton, Florida
319. Schaller, F., and Weiler, E. W. (1997) *J. Biol. Chem.* **272**, 28066 – 28072
320. Vijayan, P., Shockey, J., Lévesque, C. A., Cook, R. J., and Browse, J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7209 – 7214
321. Song, W.-C., Baertschi, S. W., Boeglin, W. E., Harris, T. M., and Brash, A. R. (1993) *J. Biol. Chem.* **268**, 6293 – 6298
321a. Ziegler, J., Stenzel, I., Hause, B., Maucher, H., Hamberg, M., Grimm, R., Ganal, M., and Wasternack, C. (2000) *J. Biol. Chem.* **275**, 19132 – 19138
322. Conconi, A., Smerdon, M. J., Howe, G. A., and Ryan, C. A. (1996) *Nature (London)* **383**, 826 – 829
323. Xie, D.-X., Feys, B. F., James, S., Nieto-Rostro, M., and Turner, J. G. (1998) *Science* **280**, 1091 – 1094
324. Alborn, H. T., Turlings, T. C. J., Jones, T. H., Stenhagen, G., Loughrin, J. H., and Tumlinson, J. H. (1997) *Science* **276**, 945 – 949
325. Paré, P. W., Alborn, H. T., and Tumlinson, J. H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13971 – 13975
326. Parchmann, S., and Mueller, M. J. (1998) *J. Biol. Chem.* **273**, 32650 – 32655
327. Koljak, R., Boutaud, O., Shieh, B.-H., Samel, N., and Brash, A. R. (1997) *Science* **277**, 1994 – 1996
328. Brash, A. R., Hughes, M. A., Hawkins, D. J., Boeglin, W. E., Song, W.-C., and Meijer, L. (1991) *J. Biol. Chem.* **266**, 22926 – 22931
329. Birch, A. J. (1967) *Science* **156**, 202 – 206
330. Spencer, J. B., and Jordan, P. M. (1992) *Biochemistry* **31**, 9107 – 9116
331. Fu, H., Alvarez, M. A., Khosla, C., and Bailey, J. E. (1996) *Biochemistry* **35**, 6527 – 6532
332. Kendrew, S. G., Harding, S. E., Hopwood, D. A., and Marsh, E. N. G. (1995) *J. Biol. Chem.* **270**, 17339 – 17343
333. Quirós, L. M., and Salas, J. A. (1995) *J. Biol. Chem.* **270**, 18234 – 18239
334. Schultz, D. J., Cahoon, E. B., Shanklin, J., Craig, R., Cox-Foster, D. L., Mumma, R. O., and Medford, J. I. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8771 – 8775
335. Michal, G., ed. (1999) *Biochemical Pathways*, Wiley - Spektrum Academischer Verlag, New York - Heidelberg, Germany
336. Hutchinson, C. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3336 – 3338
336a. Bao, W., Sheldon, P. J., and Hutchinson, C. R. (1999) *Biochemistry* **38**, 9752 – 9757

References

336b. Jez, J. M., Ferrer, J.-L., Bowman, M. E., Dixon, R. A., and Noel, J. P. (2000) *Biochemistry* **39**, 890–902
336c. Khosla, C., Gokhale, R. S., Jacobsen, J. R., and Cane, D. E. (1999) *Ann. Rev. Biochem.* **68**, 219–253
337. Cane, D. E. (1994) *Science* **263**, 338–340
338. Pieper, R., Gokhale, R. S., Luo, G., Cane, D. E., and Khosla, C. (1997) *Biochemistry* **36**, 1846–1851
339. Kao, C. M., Pieper, R., Cane, D. E., and Khosla, C. (1996) *Biochemistry* **35**, 12363–12368
340. Lambalot, R. H., Cane, D. E., Aparicio, J. J., and Katz, L. (1995) *Biochemistry* **34**, 1858–1866
341. Kendrew, S. G., Katayama, K., Deutsch, E., Madduri, K., and Hutchinson, C. R. (1999) *Biochemistry* **38**, 4794–4799
342. Kim, C.-G., Yu, T.-W., Fryhle, C. B., Handa, S., and Floss, H. G. (1998) *J. Biol. Chem.* **273**, 6030–6040
343. Aparicio, J. F., Colina, A. J., Ceballos, E., and Martín, J. F. (1999) *J. Biol. Chem.* **274**, 10133–10139
344. Kennedy, J., Auclair, K., Kendrew, S. G., Park, C., Vederas, J. C., and Hutchinson, C. R. (1999) *Science* **284**, 1368–1372
345. Offenzeller, M., Santer, G., Totschnig, K., Su, Z., Moser, H., Traber, R., and Schneider-Scherzer, E. (1996) *Biochemistry* **35**, 8401–8412
346. Schwecke, T., Aparicio, J. F., Molnár, I., König, A., Ee Khaw, L., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortés, J., Lester, J. B., Böhm, G. A., Staunton, J., and Leadlay, P. F. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7839–7843
347. Preisig-Müller, R., Gehlert, R., Melchior, F., Stietz, U., and Kindl, H. (1997) *Biochemistry* **36**, 8349–8358
348. Schröder, J., Raiber, S., Berger, T., Schmidt, A., Schmidt, J., Soares-Sello, A. M., Bardshiri, E., Strack, D., Simpson, T. J., Veit, M., and Schröder, G. (1998) *Biochemistry* **37**, 8417–8425
348a. Jez, J. M., Bowman, M. E., and Noel, J. P. (2001) *Biochemistry* **40**, 14829–14838
349. Guo, L., Dixon, R. A., and Paiva, N. L. (1994) *J. Biol. Chem.* **269**, 22372–22378
350. Brown, D. W., Adams, T. H., and Keller, N. P. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14873–14877
350a. Schroeder, F. C., Farmer, J. J., Smedley, S. R., Attygalle, A. B., Eisner, T., and Meinwald, J. (2000) *J. Am. Chem. Soc.* **122**, 3628–3634
351. Marsden, A. F. A., Wilkinson, B., Cortés, J., Dunster, N. J., Staunton, J., and Leadlay, P. F. (1998) *Science* **279**, 199–202
352. Shen, Y., Yoon, P., Yu, T.-W., Floss, H. G., Hopwood, D., and Moore, B. S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3622–3627
353. Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999) *Science* **284**, 482–485
354. Reynolds, K. A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12744–12746
355. Schröder, F. C., Farmer, J. J., Attygalle, A. B., Smedley, S. R., Eisner, T., and Meinwald, J. (1998) *Science* **281**, 428–431
356. Metz, J. G., Roessler, P., Facciotti, D., Levering, C., Dittrich, F., Lassner, M., Valentine, R., Lardizabal, K., Domergue, F., Yamada, A., Yazawa, K., Knauf, V., and Browse, J. (2001) *Science* **293**, 290–293
357. Nugent, C., Prins, J. B., Whitehead, J. P., Wentworth, J. M., Chatterjee, V. K. K., and O'Rahilly, S. (2001) *J. Biol. Chem.* **276**, 9149–9157
358. Gibbons, H. S., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2000) *J. Biol. Chem.* **275**, 32940–32949
359. Sohlenkamp, C., de Rudder, K. E. E., Röhrs, V., López-Lara, I. M., and Geiger, O. (2000) *J. Biol. Chem.* **275**, 18919–18925
360. Moody, J. S., Kozak, K. R., Ji, C., and Marnett, L. J. (2001) *Biochemistry* **40**, 861–866
361. Nakajima, J.-i, Tanaka, Y., Yamazaki, M., and Saito, K. (2001) *J. Biol. Chem.* **276**, 25797–25803

Study Questions

1. Outline possible pathways of metabolism of dietary fats. Consider digestion, transport of fatty acids, storage, conversion to prostaglandins, steroids, etc. Will any of the fat be converted into glucose?

2. What are the functions in the human body of the following?

   Pancreatic lipase
   Lipocalins
   Lipoprotein lipase
   Very low density lipoprotein (VLDL)
   Hormone-sensitive lipase
   Chylomicrons
   Apolipoproteins

3. Describe two different types of fatty acid synthase. Compare the basic chemical reactions that are involved. Also, compare these with the reactions of fatty acid oxidation.

4. Discuss the different types of fatty acids found in the human body and the synthetic pathways by which they are formed.

5. What mechanisms are utilized for incorporation of double bonds into fatty acids? Propose a mechanism that makes use of polyketide synthase domains (Fig. 21-11) in the synthesis of polyunsaturated fatty acids. See Metz *et al.*[356]

6. In what locations would you expect to find the following?

   Tripalmitin
   Mycolic acids
   Arachidonic acid
   Propionic acid
   Docosahexaenoic acid

7. Substitution of a small percentage of ω6 fatty acids in the diet of insulin-resistant rodents with ω3 unsaturated fatty acids normalized insulin action.[357] Can you suggest possible mechanisms? Is this result significant to human nutrition?

8. Formation of the 3-hydroxymyristoyl groups of lipid A (Fig. 8-30) requires O_2. Comparisons of amino acid sequences suggest that an Fe^{2+}/2-oxoglutarate-dependent oxygenase is involved.[358] Write a balanced equation for this reaction.

9. Phosphatidylcholine can be formed by two pathways as described on pp. 1198–1199. A third pathway, used by some bacteria, involves a direct one-step reaction of choline with CDP-diacylglycerol.[359] Write a reasonable chemical mechanism.

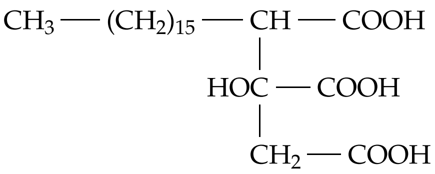
Study Questions

10. ^{14}C-Carboxyl labeled palmitic acid is fed to a fasted rat. There is no increase in liver glycogen, but the glucose units of the glycogen contain ^{14}C.
    a) Outline, using appropriate equations, the reaction sequence by which the carbon atoms of glucose become labeled.
    b) Explain why there is no net synthesis of glycogen from the fatty acid.

11. a) Write the reactions that most *dietary* tripalmitin will undergo in the body of an adult human in order to be deposited in adipose tissue as tripalmitin.
    b) What is the minimum amount of ATP (high energy bonds) normally required to deposit the one mole of dietary tripalmitin in adipose tissue? Count only ATP involved in tripalmitin metabolism and consider the source of glycerol in the adipose tissue.

12. Describe the biochemical effects on lipid metabolism of injecting into a normal animal
    a) insulin
    b) glucagon
    c) epinephrine

13. Suggest a biosynthetic pathway for formation of the fungal metabolite **agaricic acid**:



$$\begin{array}{l} CH_3-(CH_2)_{15}-CH-COOH \\ \qquad\qquad\qquad\quad | \\ \qquad\qquad\quad HOC-COOH \\ \qquad\qquad\qquad\quad | \\ \qquad\qquad\qquad\quad CH_2-COOH \end{array}$$

14. The ketone **palmiton** $CH_3(CH_2)14$ – *CO – $(CH_2)_{14}CH_3$ is formed by mycobacteria. The carbon marked by an asterisk was found to be labeled after feeding of [1-^{14}C]palmitic acid. Suggest a biosynthetic pathway.

15. The following reaction occurs in the biosynthesis of fatty acids.

    Crotonyl-ACP + NADPH + H^+ → butyryl-ACP + $NADP^+$

    The reduction half-reactions for crotonyl-ACP and NADPH are

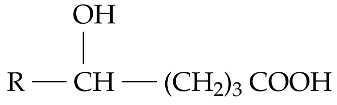
    Crotonyl-ACP + 2 H^+ + 2 e^- → butyryl-ACP
    $E^{o\prime} = -0.015$ V
    NADP + 2 H^+ + 2 e^- → NADPH + H
    $E^{o\prime} = -0.320$ V

    What is $\Delta G^{o\prime}$ for this reaction? What is the equilibrium constant for the reaction?

16. How does the inhibition of citrate synthase affect fatty acid synthesis?

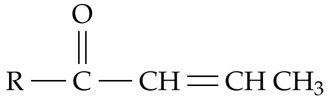
17. Malonyl CoA is an allosteric effector of carnitine acyl transferase. What kind of effector is it, i.e., activator or inhibitor, and what is the logic behind the interaction?

18. Compound II is formed in a series of enzymatic reactions from compound I. Propose a mechanistically realistic sequence, showing by name any cofactors required.



$$R-\overset{\displaystyle OH}{\overset{|}{CH}}-(CH_2)_3\,COOH$$

I



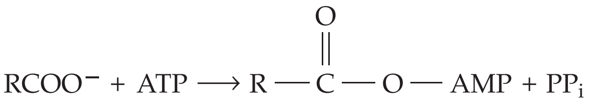
$$R-\overset{\displaystyle O}{\overset{\|}{C}}-CH=CH\,CH_3$$

II

19. Fatty acid biosynthesis requires NADPH. Where does the NADPH come from?

20. An individual has been found who is missing malic enzyme in his cytoplasm. He has instead an enzyme that converts the oxalacetate made from the citrate lyase reaction directly to pyruvate and CO_2. Discuss this patient in terms of the likely effect of these changes on his ability to synthesize fatty acids.

21. The $\Delta G^{o\prime}$ values for the hydrolysis of any P – O – P bond of ATP, inorganic pyrophosphate, or any acyl CoA thiolester are all about –34 kJ / mole, while the corresponding figure for the hydrolysis of a mixed carboxylic phosphate anhydride is about –55 kJ / mole. Calculate the value of $\Delta G^{o\prime}$ for the following reaction describing the activation of fatty acids to the fatty acyl adenylate.



$$RCOO^- + ATP \longrightarrow R-\overset{\displaystyle O}{\overset{\|}{C}}-O-AMP + PP_i$$

Study Questions

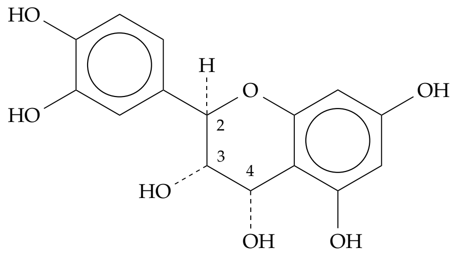
22. Fatty acid biosynthesis is made irreversible by the specific input of energy. Name the reactions or give equations for those steps in the pathway that require ATP. It is important that you consider both the mitochondrial and cytosolic components of the pathway.

23. The fatty acid biosynthesis pathway communicates with at least three other metabolic pathways either by sharing common intermediates or by regulatory mechanisms. Fill in the table below. List four molecules that have this function. You should name the additional pathway where each of these is found and briefly describe what it does in this second pathway. Do **not** use a redox cofactor as one of your choices.

Molecule	**Other Pathway**	**Role in Second Pathway**

24. The endogenous cannabinoid 2-arachidonoylglycerol is thought to play important roles both in the brain (Chapter 30) and in the immune system (Chapter 31). Leukocyte 12-oxygenase acts on this compound.[360] What products would be expected?

25. In the synthesis of anthocyanidins (Box 21-E) another Fe^{2+}/2-oxoglutarate-dependent oxygenase acts on the colorless leukoanthrocyanidin, which is then converted to the colored anthocyanidin:



Leucoanthocyanidin (colorless)

$\downarrow O_2$

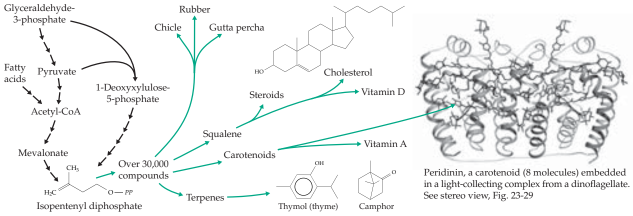
Anthocyanidin (see Box 21-E)

Propose a reasonable sequence for this reaction. See Nakajima *et al*.[361]

26. Why does a dietary deficiency of choline in humans induce a fatty liver, i.e., a liver in which the hepatocytes contain excess triglycerides?

27. Name two important functions of citrate in the conversion of acetate to palmitate by adipose tissues.

28. Outline the steps by which linoleic acid can be converted to prostaglandins in the human body.


Glyceraldehyde-3-phosphate
Fatty acids
Pyruvate
1-Deoxyxylulose-5-phosphate
Acetyl-CoA
Mevalonate
Isopentenyl diphosphate
Over 30,000 compounds
Rubber
Chicle
Gutta percha
Cholesterol
Steroids
Vitamin D
Squalene
Carotenoids
Vitamin A
Terpenes
Thymol (thyme)
Camphor

Peridinin, a carotenoid (8 molecules) embedded in a light-collecting complex from a dinoflagellate. See stereo view, Fig. 23-29

Starting with the simple compounds acetyl-CoA, glyceraldehyde-3-phosphate, and pyruvate, which arise via the central pathways of metabolism, the key intermediate **isopentenyl diphosphate** is formed by two independent routes. It is then converted by bacteria, fungi, plants, and animals into thousands of different naturally occurring products. These include high polymers, such as rubber, as well as vitamins, sterols, carotenoids, and over 30,000 different terpenes and related compounds. Many of the latter are found only in specific plants where they may function as defensive compounds or pheromones.

Contents

1227 .... **A. Isopentenyl Diphosphate and Polyprenyl Synthases**
1229 .... 1. An Alternative Pathway for Isoprenoid Synthesis
1230 .... 2. Isomerization and Isoprene Formation
1230 .... 3. Polyprenyl Compounds
1231 .... 4. Prenylation of Proteins and Other Compounds
1232 .... **B. Terpenes**
1232 .... 1. Biosynthesis of Monoterpenes
1234 .... 2. Sesquiterpenes and Diterpenes
1236 .... 3. Formation of the Symmetric Terpenes, Squalene and Phytoene
1237 .... **C. Carotenes and Their Derivatives**
1240 .... 1. Xanthophylls and Other Oxidized Carotenes
1243 .... 2. Properties and Functions of Carotenes
1243 .... **D. Steroid Compounds**
1244 .... 1. Biosynthesis of Sterols
1245 .... *Formation of cholesterol*
1247 .... 2. Metabolism of Cholesterol in the Human Body
1248 .... *Serum cholesterol*
1250 .... *The LDL and related receptors*
1251 .... *Abnormalities of cholesterol metabolism*
1251 .... *Bile acids*
1252 .... **E. The Steroid Hormones**
1253 .... 1. Progestins
1253 .... 2. Adrenal Cortical Steroids
1254 .... 3. Androgens
1260 .... 4. Estrogens
1263 .... 5. The Steroid Receptor Family
1265 .... **F. Other Steroids**

1266 .... **References**
1271 .... **Study Questions**

Boxes

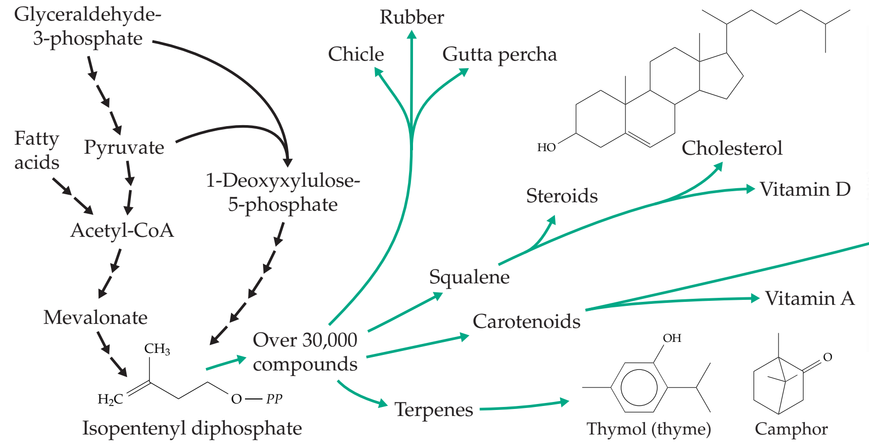
1241 .... Box 22-A Vitamin A
1249 .... Box 22-B Atherosclerosis
1257 .... Box 22-C Vitamin D
1261 .... Box 22-D The Renin–Angiotensin–Aldosterone System and the Regulation of Blood Pressure

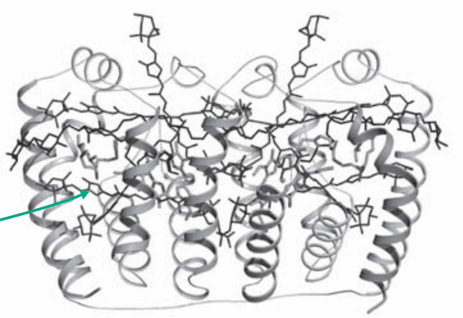
Tables

1264 .... Table 22-1 Known Members of the Steroid Receptor Family

22

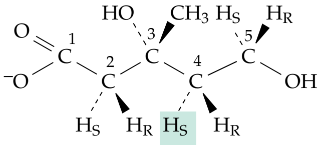
Polyprenyl (Isoprenoid) Compounds


Glyceraldehyde-3-phosphate
Fatty acids
Pyruvate
Acetyl-CoA
1-Deoxyxylulose-5-phosphate
Mevalonate
CH_3
H_2C
O — *PP*
Isopentenyl diphosphate
Over 30,000 compounds
Rubber
Chicle
Gutta percha
Squalene
Steroids
HO
Cholesterol
Vitamin D
Carotenoids
Vitamin A
Terpenes
OH
Thymol (thyme)
O
Camphor

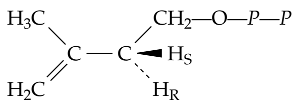


Peridinin, a carotenoid (8 molecules) embedded in a light-collecting complex from a dinoflagellate. See stereo view, Fig. 23-29

The **terpenes**, **carotenoids**, **steroids**, and many other compounds arise in a direct way from the prenyl group of **isopentenyl diphosphate** (Fig. 22-1).[1–6a] Biosynthesis of this five-carbon branched unit from **mevalonate** has been discussed previously (Chapter 17, Fig. 17-19) and is briefly recapitulated in Fig. 22-1. Distinct isoenzymes of 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase) in the liver produce HMG-CoA destined for formation of ketone bodies (Eq. 17-5) or mevalonate.[7,8] A similar cytosolic enzyme is active in plants which, collectively, make more than 30,000 different isoprenoid compounds.[9,10] However, many of these are formed by an alternative pathway that does not utilize mevalonate but starts with a thiamin diphosphate-dependent condensation of glyceraldehyde 3-phosphate with pyruvate (Figs. 22-1, 22-2).


O
$^-$O
1
2
3
4
5
HO
CH_3
H_S
H_R
OH
H_S
H_R
H_S
H_R

3(*R*) Mevalonate


H_3C
H_2C
CH_2—O—*P*—*P*
H_S
H_R

Isopentenyl diphosphate

The two-step reduction of HMG-CoA to mevalonate (Fig. 22-1, step *a*)[11–15] is highly controlled, a major factor in regulating cholesterol synthesis in the human liver.[12,16,17] The N-terminal portion of the 97-kDa 888-residue mammalian HMG-CoA reductase is thought to be embedded in membranes of the ER, while the C-terminal portion is exposed in the cytoplasm.[16] The enzyme is sensitive to feedback inhibition by cholesterol (see Section D, 2). The regulatory mechanisms include a phosphorylation–dephosphorylation cycle and control of both the rates of synthesis and of proteolytic degradation of this key enzyme.[14,15,18–20]

A. Isopentenyl Diphosphate and Polyprenyl Synthases

In animals all isoprenoid compounds are apparently synthesized from mevalonate, which is converted by the consecutive action of two kinases[21–23] into mevalonate 5-diphosphate (Fig. 22-1, step *b*). Mevalonate kinase is found predominantly in peroxisomes, which are also active in other aspects of steroid synthesis in humans.[21,24] A deficiency of this enzyme is associated with mevalonic aciduria, a serious hereditary disease in which both blood and urine contain very high concentrations of mevalonate.[23] Mevalonate diphosphate kinase, which is also a decarboxylase, catalyzes phosphorylation of the 3-OH group of mevalonate (step *c*, Fig. 22-1) and decarboxylative elimination of phosphate (step *d*)[25] to form isopentenyl diphosphate.

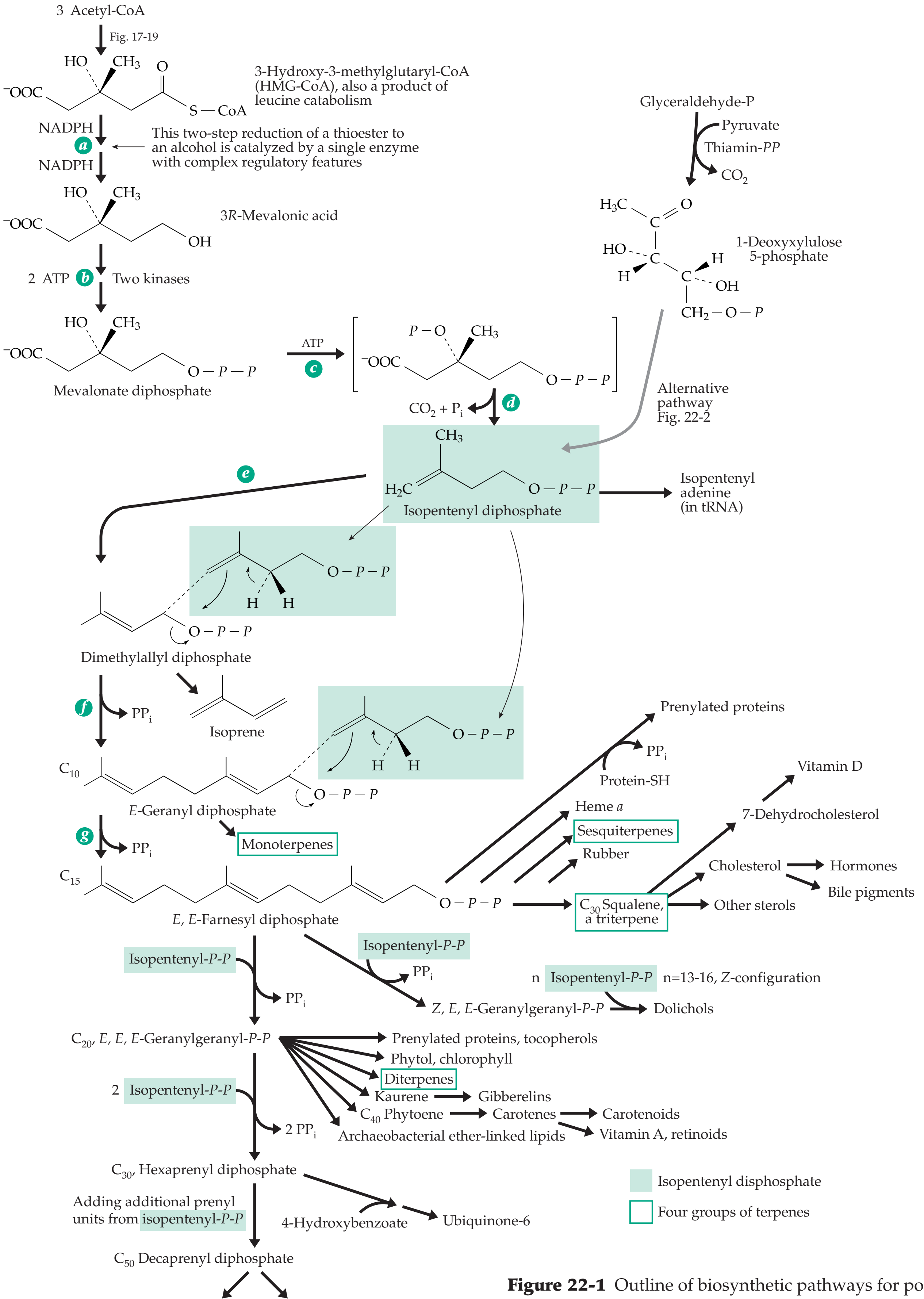

3 Acetyl-CoA
Fig. 17-19
HO
CH3
O
⁻OOC
S – CoA
3-Hydroxy-3-methylglutaryl-CoA (HMG-CoA), also a product of leucine catabolism
NADPH
a
NADPH
This two-step reduction of a thioester to an alcohol is catalyzed by a single enzyme with complex regulatory features
3R-Mevalonic acid
OH
2 ATP
b
Two kinases
O – P – P
Mevalonate diphosphate
ATP
c
P – O
d
CO2 + Pi
Glyceraldehyde-P
Pyruvate
Thiamin-PP
CO2
H3C
C
HO
H
C
OH
CH2 – O – P
1-Deoxyxylulose 5-phosphate
Alternative pathway Fig. 22-2
H2C
Isopentenyl diphosphate
Isopentenyl adenine (in tRNA)
e
H
H
Dimethylallyl diphosphate
f
PPi
Isoprene
C10
E-Geranyl diphosphate
Monoterpenes
g
C15
E, E-Farnesyl diphosphate
Prenylated proteins
Protein-SH
Heme a
Sesquiterpenes
Rubber
Vitamin D
7-Dehydrocholesterol
Cholesterol
Hormones
Bile pigments
C30 Squalene, a triterpene
Other sterols
Isopentenyl-P-P
Z, E, E-Geranylgeranyl-P-P
n Isopentenyl-P-P n=13-16, Z-configuration
Dolichols
C20, E, E, E-Geranylgeranyl-P-P
Prenylated proteins, tocopherols
Phytol, chlorophyll
Diterpenes
Kaurene
Gibberelins
C40 Phytoene
Carotenes
Carotenoids
Vitamin A, retinoids
Archaeobacterial ether-linked lipids
2 Isopentenyl-P-P
2 PPi
C30, Hexaprenyl diphosphate
4-Hydroxybenzoate
Ubiquinone-6
Adding additional prenyl units from isopentenyl-P-P
C50 Decaprenyl diphosphate
Chicle
Gutta-percha
Isopentenyl disphosphate
Four groups of terpenes

Figure 22-1 Outline of biosynthetic pathways for polyprenyl compounds. Steps *e*, *f*, and *g* and many subsequent steps are catalyzed by isoprenyl diphosphate synthases.

1. An Alternative Pathway for Isoprenoid Synthesis

It is generally agreed that mevalonate is the precursor to sterols in higher plants as well as in animals and is also the precursor to plant carotenoids. However, it is poorly incorporated into monoterpenes and into some diterpenes such as those of the taxane group.[26,27] The alternative **glyceraldehyde 3-phosphate:pyruvate pathway** explains this result. The pathway also operates in some bacteria and apparently is the sole source of isoprenoid compounds for the unicellular alga *Scenedesmus*.[28] The pathway is outlined in Fig. 22-2. Pyruvate is decarboxylated by a thiamin diphosphate-dependent enzyme,[29] and the resulting enamine is condensed with D-glyceraldehyde 3-phosphate to form 1-deoxyxylulose 5-phosphate.[28,30–31a] The latter undergoes an isomeroreductase rearrange-

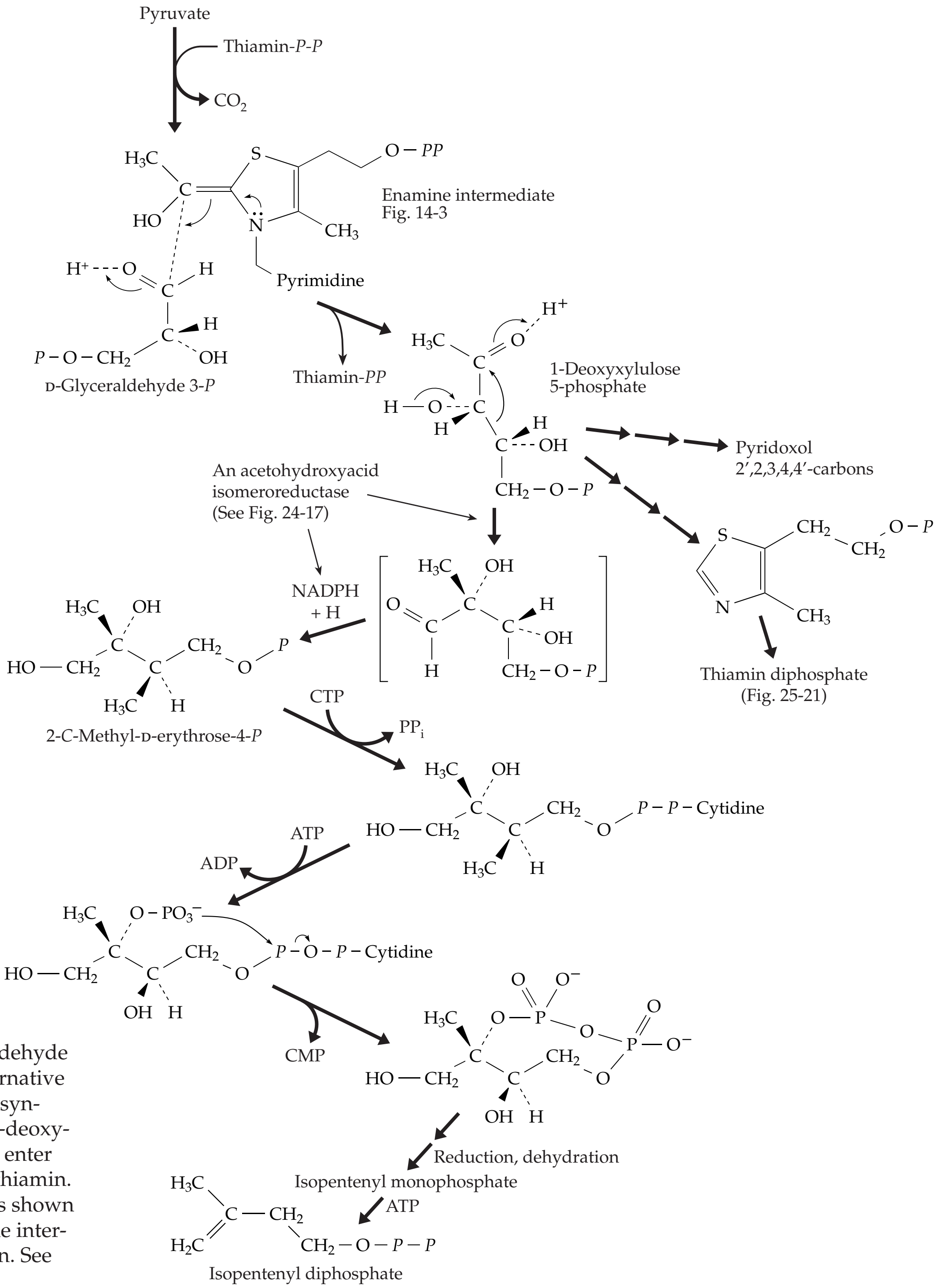
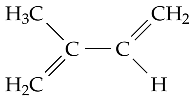

Pyruvate
Thiamin-*P*-*P*
CO_2
H_3C
S
O – *PP*
C
HO
N
CH_3
Enamine intermediate
Fig. 14-3
Pyrimidine
H^+
O
H
C
C
H
P – O – CH_2
OH
D-Glyceraldehyde 3-*P*
Thiamin-*PP*
H^+
H_3C
O
1-Deoxyxylulose
5-phosphate
H — O
H
C
H
C --- OH
CH_2 – O – *P*
Pyridoxol
2′,2,3,4,4′-carbons
An acetohydroxyacid
isomeroreductase
(See Fig. 24-17)
NADPH
+ H
H_3C
OH
O
C
H
OH
CH_2 – O – *P*
S
CH_2
O – *P*
CH_2
N
CH_3
Thiamin diphosphate
(Fig. 25-21)
HO — CH_2
CH_2
O
P
H_3C
H
2-*C*-Methyl-D-erythrose-4-*P*
CTP
PP_i
P – *P* – Cytidine
ATP
ADP
O – PO_3^-
P – O – *P* – Cytidine
OH H
CMP
O
O^-
O — P
P — O^-
Reduction, dehydration
Isopentenyl monophosphate
ATP
H_2C
C — CH_2
CH_2 – O – *P* – *P*
Isopentenyl diphosphate

Figure 22-2 The glyceraldehyde 3-phosphate:pyruvate alternative pathway of isoprenoid biosynthesis. The intermediate 1-deoxyxylulose 5-phosphate may enter terpenes, vitamin B_6, and thiamin. Isopentenyl diphosphate is shown as the final product, but the intermediate steps are uncertain. See Lange *et al.*[32g]

ment of the type that occurs in the biosynthesis of valine and isoleucine (Fig. 24-17),[32] but the additional steps on to isopentenyl diphosphate are not obvious. However, some intermediate compounds have been identified as is indicated in Fig. 22-2.[32a–g] 1-Deoxyxylulose 5-phosphate has also been identified as an intermediate in the biosynthesis of vitamin B$_6$ in *E. coli*.[32a,h,33] It gives rise to the 2', 2, 3, 4, and 4' carbon atoms of pyridoxine and also to the pyrimidine ring of thiamin. See Chapter 25.

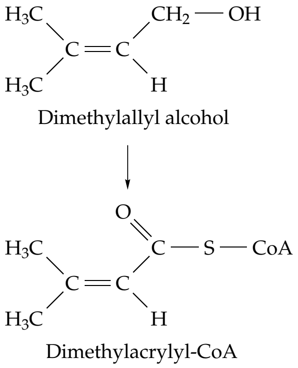
2. Isomerization and Isoprene Formation

Before polymer formation begins, one molecule of isopentenyl diphosphate must be isomerized to **dimethylallyl diphosphate** (Fig. 22-1, step *e*, Eq. 13-56).[10,33a,b] In this process the hydrogen that was in the 4-*pro-S* position of mevalonic acid (the *pro-R* position of isopentenyl pyrophosphate) is lost. Dimethylallyl diphosphate is unstable and can undergo acid-catalyzed elimination of PP$_i$ to form isoprene, apparently by a carbocation mechanism.



Isoprene

This evidently accounts for the presence of isoprene in the breath.[34] Isoprene is also formed by many plants and is released into the atmosphere in large amounts, which contribute to photochemical formation of haze. A Mg^{2+}-dependent enzyme catalyzes the elimination of pyrophosphate.[35] Isoprene emissions rise with increasing temperature, and it has been suggested that the isoprene may dissolve in chloroplast membranes and in some way confer increased heat resistance.[36,37] Hydrolytic dephosphorylation can lead to dimethylallyl alcohol, which is oxidized in the liver to dimethylacrylyl-CoA (Eq. 22-1).

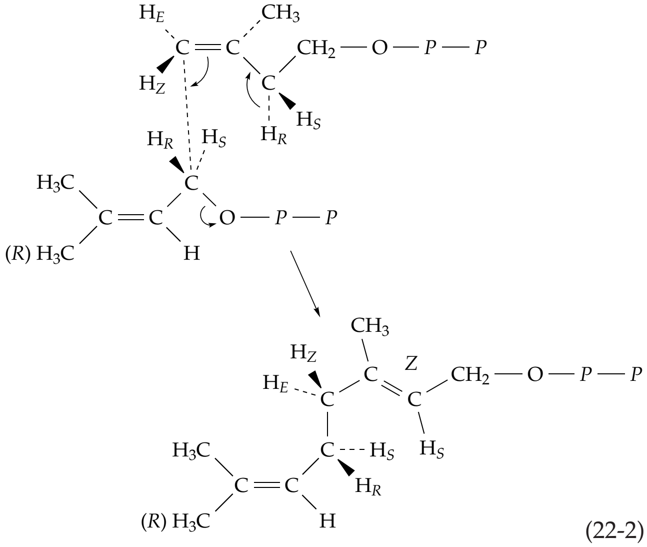


(22-1)

The latter is also a catabolite of leucine and can be converted back to HMG-CoA via a biotin-dependent carboxylation (see Fig. 24-18). This provides a means of recycling the dimethylallyl alcohol back to the polyprenyl pathway.[38]

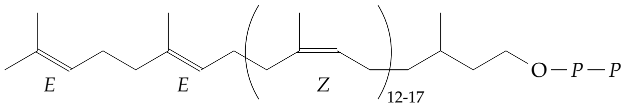
3. Polyprenyl Compounds

Dimethylallyl diphosphate serves as the starter piece for most polyprenyl compounds. Additional prenyl units are added, with elimination of pyrophosphate, by the action of **polyprenyl diphosphate synthases** as indicated in Fig. 22-1. Many of the products have all-trans (*E*) double bonds. A substantial number of these synthases are known and are distinguished by their chain length specificity and stereochemical properties.[39–46b] The most studied is farnesyl diphosphate synthase. The three-dimensional structure of an avian form is known.[47,47a] It catalyzes steps *f* and *g* of Fig. 22-1, joining three prenyl groups with the *E*,*E* (trans, trans) configuration. This protein, which consists almost entirely of packed α helices, has a large central cavity with conserved lysines and two aspartate-rich sequences (DDXXD) along its walls. These polar groups, together with magnesium ions, probably bind the pyrophosphate groups of the substrates.[48] Aspartates 224 and 225 of a bacterial form of the enzyme appear to be essential for catalytic activity.[49] The reaction is thought to be initiated by elimination of PP$_i$ to form a carbocation to which the second prenyl unit adds as in Eq. 22-2.[50,51] For each prenyl unit a hydrogen atom that was originally the 4-*pro-S* hydrogen of mevalonic acid is lost as a proton.[52,53] Addition of another prenyl unit gives *E*,*E*,*E* geranylgeranyl diphosphate.[46a]



(22-2)

The chain length of the polyprenyl compounds appears to be determined by the protein structure of the synthase.[45,54,54a,b] Polymerization of prenyl units can continue with the formation of high molecular weight polyprenyl alcohols such as the **dolichols** and bacterial decaprenols (Chapter 20) or of the high polymers **rubber** (all *Z* configuration), chicle, and **gutta-percha**.[6a] Dolichols, which function in the biosynthesis of glycoproteins, consist of 16–21 prenyl units and are synthesized in the endoplasmic reticulum as the diphosphates.[55–57] Farnesyl diphosphate is elongated to *Z,E,E* geranylgeranyl diphosphate, and polymerization continues with addition of 13–18 more units, all with the *Z* (cis) configuration.[46b,57a,b] However, after dephosphorylation[58] the double bond of the last unit added becomes saturated.[59] Partial absence of the required reductase causes a serious human deficiency disease involving faulty glycoprotein synthesis.[60]



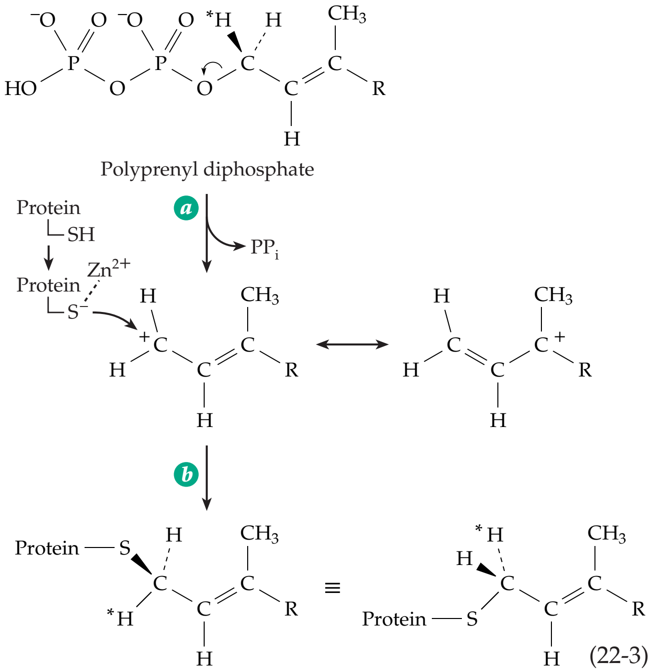
A fully extended 19-unit dolichol (dolichol-19) would have a length of about 10 nm, twice that of the bilayer in which it is dissolved. It has been suggested that the central part of the molecule has a helical structure, while the ends are more flexible. Dolichols also appear to increase the fluidity of membrane bilayers.[61] Bacterial undecaprenyl diphosphate, which has a similar function, contains only one *E* and ten *Z* double bonds[62–63a] (see p. 1152).

Rubber also contains almost entirely *Z* double bonds. Consistent with this fact is the finding that the prenyltransferases catalyzing formation of rubber promote loss of the *pro-R* proton rather than the *pro-S* proton of mevalonic acid (see Eq. 22-2). There appear to be two types of prenyltransferase in animal mitochondria giving rise to *E* and *Z* double bonds, respectively.[64] In contrast, the rubber tree contains a 137-residue protein, the **rubber elongation factor**. This small protein binds to *E* prenyltransferases causing them to form *Z* double bonds.[65] The bacterium *Micrococcus luteus* synthesizes all *E* polyprenyl alcohol diphosphates up to the C_{45} nonaprenyl compound **solanesyl diphosphate**.[66]

Chain elongation during polymerization of prenyl units can be terminated in one of a number of ways. The pyrophosphate group may be hydrolyzed to a monophosphate or to a free alcohol. Alternatively, two polyprenyl compounds may join "head to head" to form a symmetric dimer. The C_{30} terpene **squalene**, the precursor to cholesterol, arises in this way from two molecules of farnesyl diphosphate as does **phytoene**, precursor of the C_{40} carotenoids, from *E,E,E* geranylgeranyl diphosphate. The phytanyl groups of archaebacterial lipids (p. 385) arise rather directly from geranylgeranyl diphosphate by transfer of the polyprenyl group to the $–CH_2OH$ group of *sn*-glycerol 3-phosphate.[67,67a] This is followed by hydrogenation of the double bonds. Formation of diphytanyl group (p. 388) must involve additional crosslinking reactions.

4. Prenylation of Proteins and Other Compounds

Polyprenyl groups are often transferred onto thiolate ions of cysteine side chains of certain proteins that bind to membranes (p. 559).[68,69] We have previously considered the Ras family (Chapter 11). Recoverin, an important protein in the visual cycle (Fig. 23-43), is another example of a prenylated protein. Both **farnesyltransferases**[70–76f] and **geranylgeranyltransferases**[72,77–78b] have been characterized, and the three- dimensional structure of the former has been established.[73,75–76] The two-domain protein contains a seven-helix crescent-shaped hairpin domain and an α,α-barrel similar to that in Fig. 2-29. A bound zinc ion in the active site may bind the $–S^-$ group of the substrate protein after the farnesyl diphosphate has been bound into the active site.[76,79] These enzymes are thought to function by a carbocation mechanism as shown in Eq. 22-3 and with the indicated inversion of configuration.[71]



(22-3)

Inhibition of these prenyltransferases blocks growth of tumor cells. Many prenyltransferase inhibitors are apparently nontoxic to normal cells and are undergoing human clinical trials as anticancer drugs.[76a, 79a,b] Among other important polyprenyl compounds are the side chains of vitamin K, the ubiquinones, plastoquinones, tocopherols, and the phytyl group of chlorophyll. In all cases, a diphosphate of a polyprenyl alcohol serves as an alkyl group donor. Introduction of the polyprenyl chain into aromatic groups, such as those of the quinones (Fig. 15-24), occurs at a position ortho to a hydroxyl group in the reduced quinone (hydroquinone). The reader should be able to propose a reasonable prenyltransferase mechanism involving participation of the hydroxy group. The monoprenyl compound dimethylallyl diphosphate prenylates the N^6 position of adenine in a specific site in many tRNA molecules (Fig. 5-33)[80] as well as the C-4 position of L-tryptophan in the synthesis of ergot alkaloids.[81]

B. Terpenes

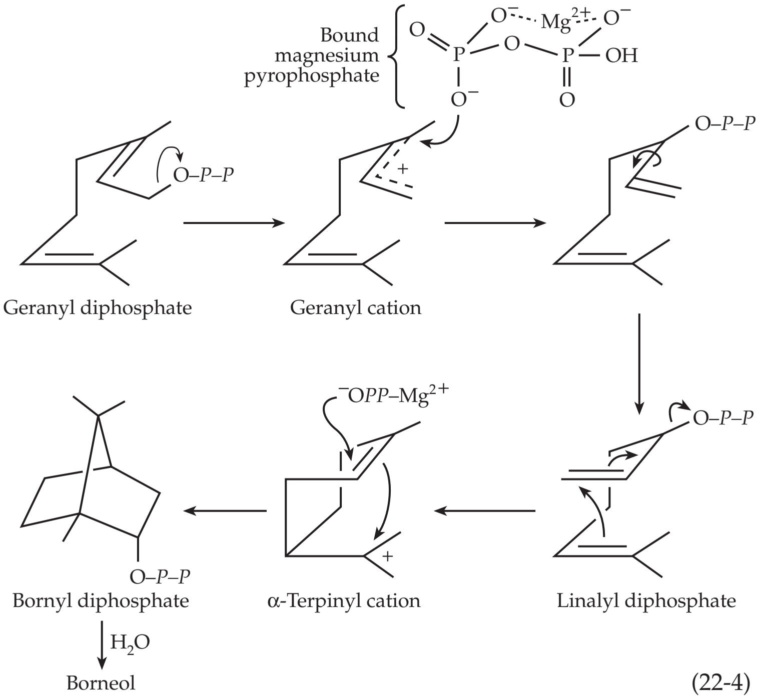
The number of small compounds that arise from isopentenyl diphosphate and are found in plants, animals, and bacteria is staggering. Just a few of these “terpenes” are shown in Figs. 22-3 and 22-4. The biosynthetic pathways have been worked out by “feeding” radioactively labeled acetate to plants and studying the characteristic labeling patterns in the terpenes. Many of the enzymes involved have been identified and studied. A given plant usually contains a large number of different terpenes, which are often concentrated in specialized “oil glands” or resinous duct tissues. Lesser amounts, often as glycosides of terpene alcohols, may be present within cells. Some terpenes occur in truly enormous amounts. For example, turpentine may contain 64% of **α-pinene** and juniper oil 65% **α-terpineol**.[82] The large quantities of α-pinene released into the air from pine trees are a major cause of photochemical smog.[83]

Terpenes have a variety of functions. Plant terpenes may deter herbivores and attract pollinators. They may participate in competition among plants and may act as antibiotics, called **phytoalexins**, to protect plants from bacteria and fungi.[84] In invertebrate animals terpenes serve as hormones, pheromones, and defensive repellants (Figs. 22-3, 22-4). The terpene squalene is the precursor to sterols. Some terpenes are toxic. For example, thujone (Fig. 22-3), which is present in the liqueur absinthe, causes serious chronic poisoning.[84]

1. Biosynthesis of Monoterpenes

The compounds of Figure 22-3 each contain ten carbon atoms and are called monoterpenes. They occur largely in plants, but some function in arthropods as pheromones. As with chain elongation, the cyclization of geranyl diphosphate to the various monoterpenes appears to occur through loss of pyrophosphate (as PP_i) with formation of an intermediate carbocation such as that depicted in Equation 22-3.[85–88a] Similar mechanisms initiate cyclization of sesquiterpenes and diterpenes. Numerous terpene cyclases have been isolated, and several have been studied carefully. A stereochemical view of the formation of borneol is illustrated in Eq. 22-4. Both linalyl-*PP* and bornyl-*PP* are intermediates. Croteau and associates suggested that a tight ion pair between carbocation and magnesium pyrophosphate is maintained at each stage.[86,89]

As is indicated in Fig. 22-3, the same intermediate cation can yield a variety of end products. For example, pure geranyl diphosphate: pinene cyclase catalyzes formation of several other terpenes in addition to α-pinene.[89] Another aspect of terpene synthesis is that insects may convert a plant terpene into new compounds for their own use. For example, **myrcene**, which is present in pine trees, is converted by bark beetles to **ipsenol** (Fig. 22-3), a compound that acts as an aggregation pheromone.[90]



(22-4)

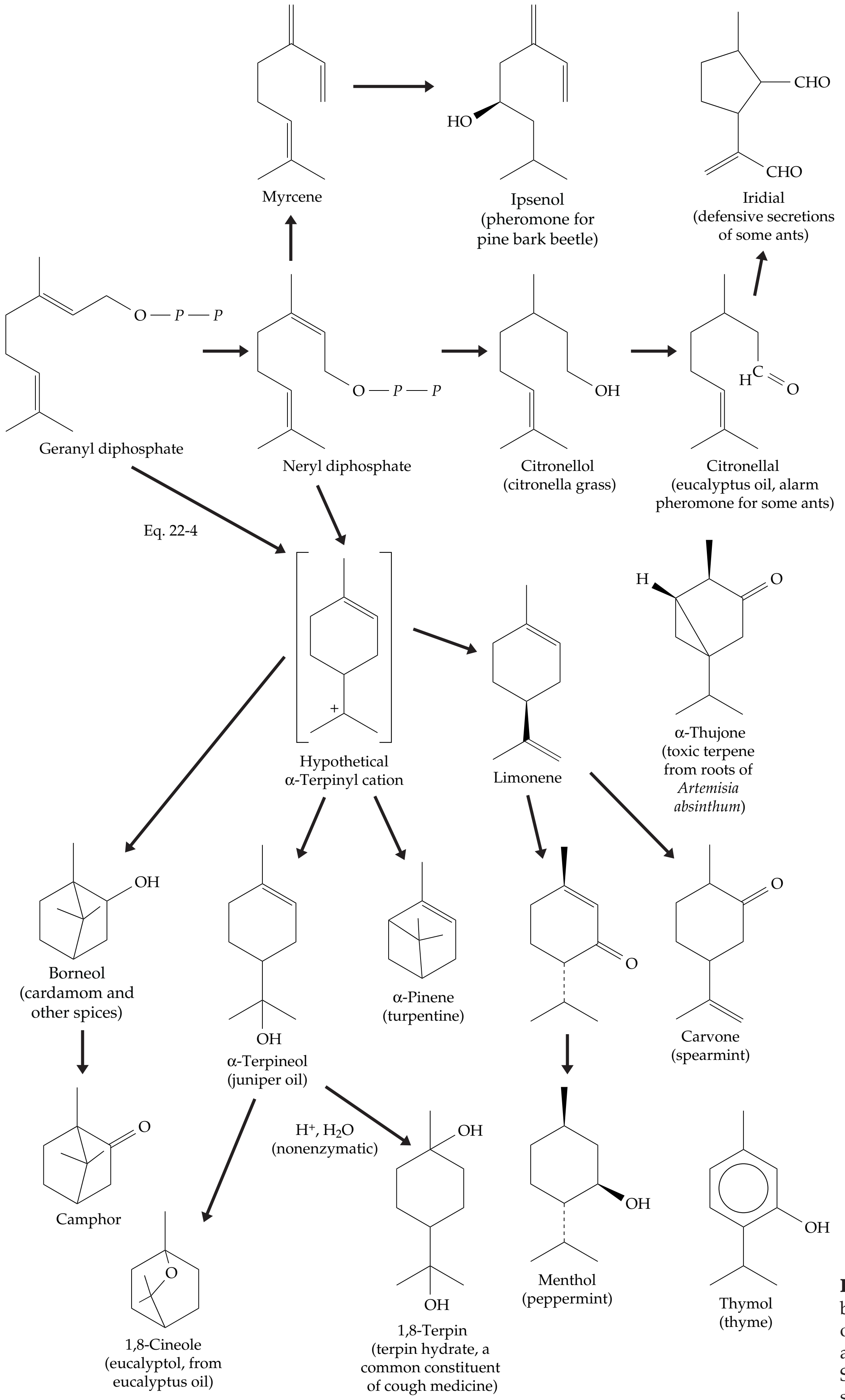
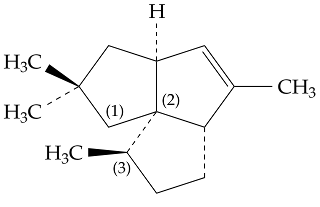

CHO
HO
CHO
Myrcene
Ipsenol
(pheromone for
pine bark beetle)
Iridial
(defensive secretions
of some ants)
O — P — P
O — P — P
OH
H C O
Geranyl diphosphate
Neryl diphosphate
Citronellol
(citronella grass)
Citronellal
(eucalyptus oil, alarm
pheromone for some ants)
Eq. 22-4
H
O
+
Hypothetical
α-Terpinyl cation
Limonene
α-Thujone
(toxic terpene
from roots of
*Artemisia
absinthum*)
OH
O
O
Borneol
(cardamom and
other spices)
OH
α-Terpineol
(juniper oil)
α-Pinene
(turpentine)
Carvone
(spearmint)
O
Camphor
H^+, H_2O
(nonenzymatic)
OH
OH
OH
O
1,8-Cineole
(eucalyptol, from
eucalyptus oil)
OH
1,8-Terpin
(terpin hydrate, a
common constituent
of cough medicine)
Menthol
(peppermint)
Thymol
(thyme)

Figure 22-3 Probable biosynthetic pathways of some monoterpenes and related substances. Some of the natural sources are indicated.

2. Sesquiterpenes and Diterpenes

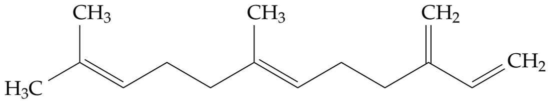
Most of the compounds shown in Figure 22-4 are derived from the C_{15} farnesyl diphosphate. There are more than 300 known cyclic structures among these **sesquiterpenes,** and many sesquiterpene synthases have been characterized.[91,91a] **Aristolochene** is formed by the action of a 38-kDa cyclase that has been isolated from species of *Penicillium* and *Aspergillus.*[92–94] Notice that the synthesis must involve two cyclization steps and migration of a methyl group. Three-dimensional structures are known for at least two terpene synthases,[95,96] and comparison of gene sequences suggests that many others have similar structures. The **5-epi-aristolochene synthase** of tobacco makes the 5-epimer of aristolochene (Fig. 22-4). It binds the diphosphate group within a central cavity using two Mg^{2+} ions held by carboxylates, some of which are in the DDXXD sequence found also in polyprenyl diphosphate synthases. The enzyme active sites of both the epi-aristolochene synthase and a **pentalenene synthase** from *Streptomyces*[96] are rich in polar groups that form hydrogen-bonded networks and which participate in proton abstraction and donation during the rearrangement reactions that must occur.



Pentalenene

Aromatic groups that are also present may assist in stabilizing intermediate carbocationic species. Deprotonation by an aspartate side chain in the epi-aristolochene synthase has been proposed to assist the cyclization; subsequent reprotonation by an Asp•Tyr•Asp triad would generate a new carbocation and promote the necessary methyl group migration. A detailed step-by-step mechanism has been proposed.[95] The fungal pentalenene synthase has an active site histidine, which is proposed to serve as proton acceptor and donor for the several steps of the reaction of farnesyl diphosphate.[96] The carbon atoms originating from C1 to C3 of the precursor have been marked on the pentalenene structure as well as on the aristolochene structure in Fig. 22-4. We now see how synthases can guide the terpene cyclase reactions to give specific products.

Another sesquiterpene synthase forms **trichodiene**, the parent compound for a family of mycotoxins and antibiotics.[97,97a,b] A different sesquiterpene synthase, present in peppermint and also in a wide range of other plants and animals, forms the acyclic **(E)-β-farnesene**.[98] It serves as alarm pheromone for aphids, has a variety of signaling functions in other insects, and is a urinary pheromone in mice.[98]



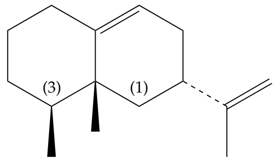
(*E*)-β-Farnesene

Abscisic acid, one of five known types of plant hormone of general distribution throughout higher plants, is not regarded as a true terpene because it arises by degradation of a carotenoid.[99–99b] However, its structure (Fig. 22-4) is that of a sesquiterpene.

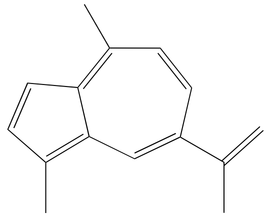
The C_{20} **diterpenes** are derived from geranylgeranyl-*PP*. Among the best known members are another group of plant hormones, the **gibberellins**.[100–102a] The first gibberellin was isolated as a product of plants infected with a *Fusarium* fungus. The rice plants grew in an abnormally tall, weak form. Subsequently, this multimembered class of over 50 highly modified diterpenes have been shown to have a variety of regulatory functions in all higher plants. For example, gibberellins are essential for stem elongation.

Equation 22-5 gives an abbreviated biosynthetic sequence for gibberellin A_1. The ring closure of step *a*, Eq. 22-5, may be initiated by protonation of the double bond at the left of the first structural formula. The resultant carbocation could initiate the consecutive closure of the two rings and the loss of a proton from a methyl group (step *b*) to yield copalyl-*PP*. Steps *c* and *d*, Eq. 22-5, each represent several reactions. In step *c*, pyrophosphate is eliminated, and the methyl group that becomes a methylene in *ent*-kaurene (enantiotopic kaurene) undergoes migration.[102b] Step *d* involves several hydroxylation and oxidation steps as well as a ring contraction through which one of the original methylene groups ends up as a carboxyl group in the final product.[102,103,103a] Deactivation of the hormone is initiated by a 2-oxoglutarate-dependent dioxygenase.[103b]

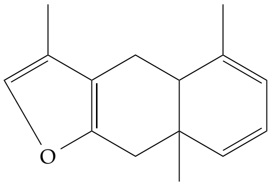
The **juvenile hormone** of insects (Fig. 22-4) is also of polyprenyl origin.[104,105] However, two of the methyl groups have been converted to (or replaced by) ethyl groups. The isolation and identification of the structure of the juvenile hormone was a difficult task. After its completion it was a surprise to researchers to discover that a large variety of synthetic compounds, sometimes with only a small amount of apparent structural similarity, also serve as juvenile hormones, keeping insects in the larval stage or preventing insect eggs from hatching. Furthermore, a number of plant products such as **juvebione** (Fig. 22-4),[106] which was originally isolated from paper, have the same effect. Thus, in nature products of plant metabolism have a profound effect upon the development of insects that



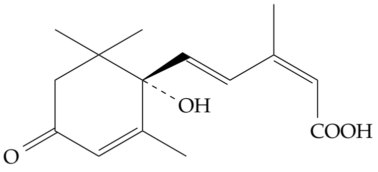
Aristolochene, parent hydrocarbon of many fungal toxins



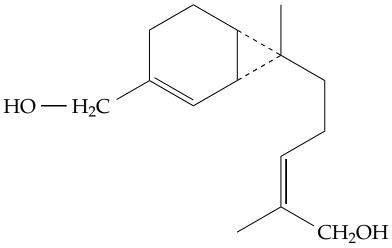
Lactarazulene, dye from the mushroom *Lactarius deliciosus*



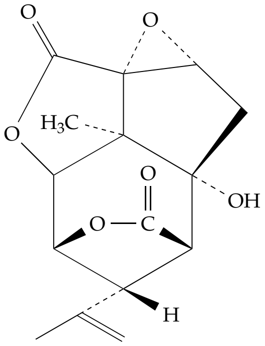
Furanoeudesma-1,3-diene, a sesquiterpene from myrrh with analgesic properties



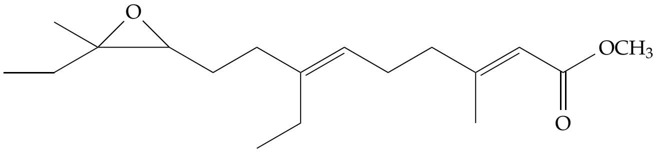
Abscisic acid, an important hormone of green plants



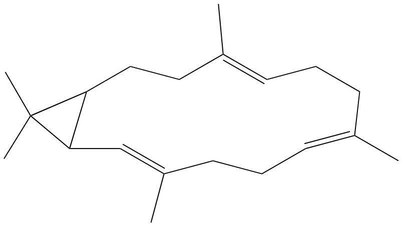
Sirenin, a sperm attractant produced by female gametes of the phycomycete *Allomyces*



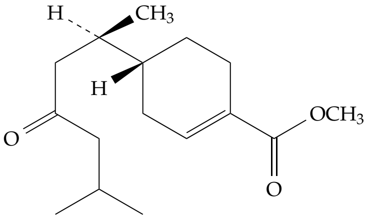
Picrotoxin, a plant toxin sometimes used to kill fish



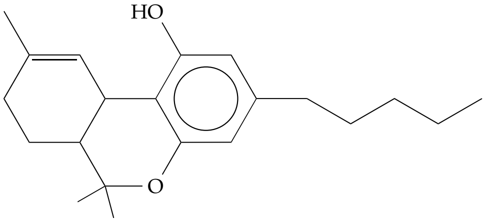
Juvenile hormone, a substance acting to maintain the larval stage of insects



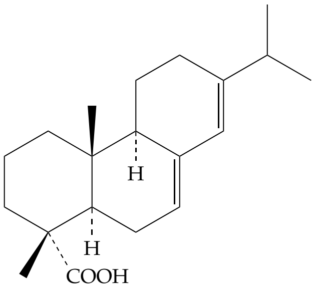
Casbene, a macrocyclic phytoalexin



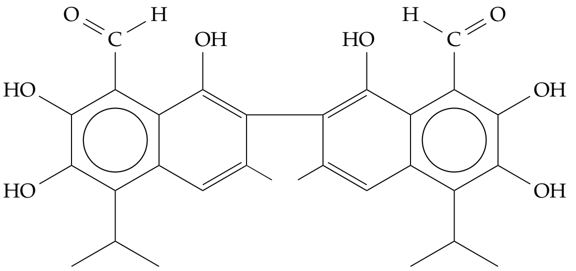
Juvebione, a product of the balsam fir with juvenile hormone activity



Δ'-3,4-*trans*-Tetrahydrocannabinol (Δ'-THC) a psychotropic component of marijuana

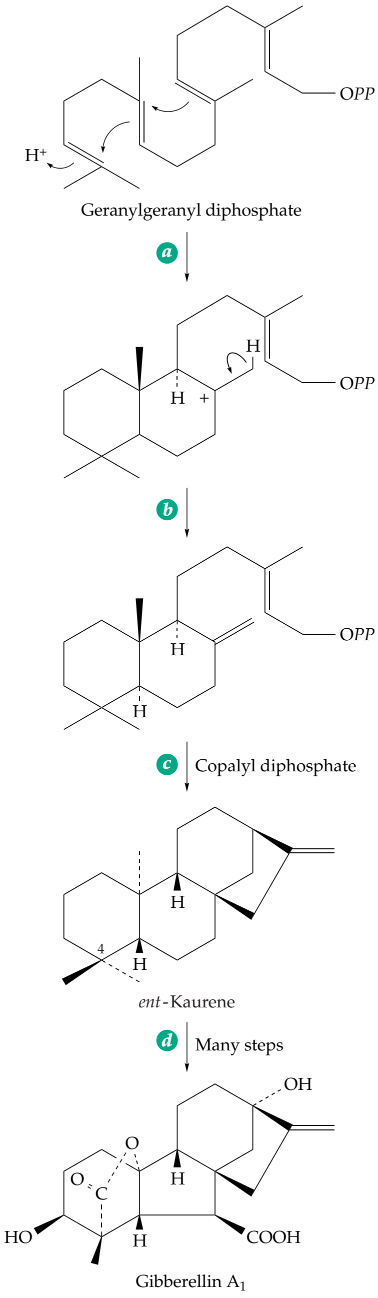


Abietic acid, a component of rosin, from conifers



Gossypol, toxic, dimeric sesquiterpene from cottonseed

Figure 22-4 More terpenes and related substances. The numbers in parentheses on the aristolochene structure are those of atoms in the precursor farnesyl diphosphate.



(22-5)

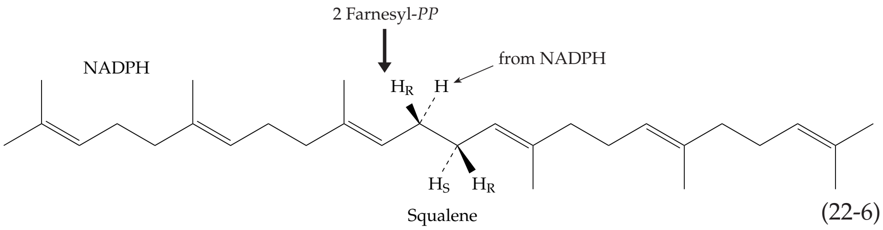
eat the plants. There is interest in the possible use of juvenile hormone, or of synthetic compounds mimicking its action, as insecticides.

Many conifers secrete **oleoresin** (pitch) in response to attack by bark beetles. The oleoresin contains approximately equal amounts of turpentine (a mixture of monoterpenes and sesquiterpenes) and diterpenoid rosin including **abietic acid** (Fig. 22-4).[91,107,107a] The oleoresin is toxic to beetles and, after evaporation of the turpentine, forms a hard rosin seal over the wounds. **Casbene** (Fig. 22-4) is a diterpene produced by castor beans as a phytoalexin (Chapter 31), an antifungal antibiotic.[108] The synthesis of the anticancer compound taxol (Box 7-D) from geranylgeranyl diphosphate involves extensive oxidative and other modification.[109]

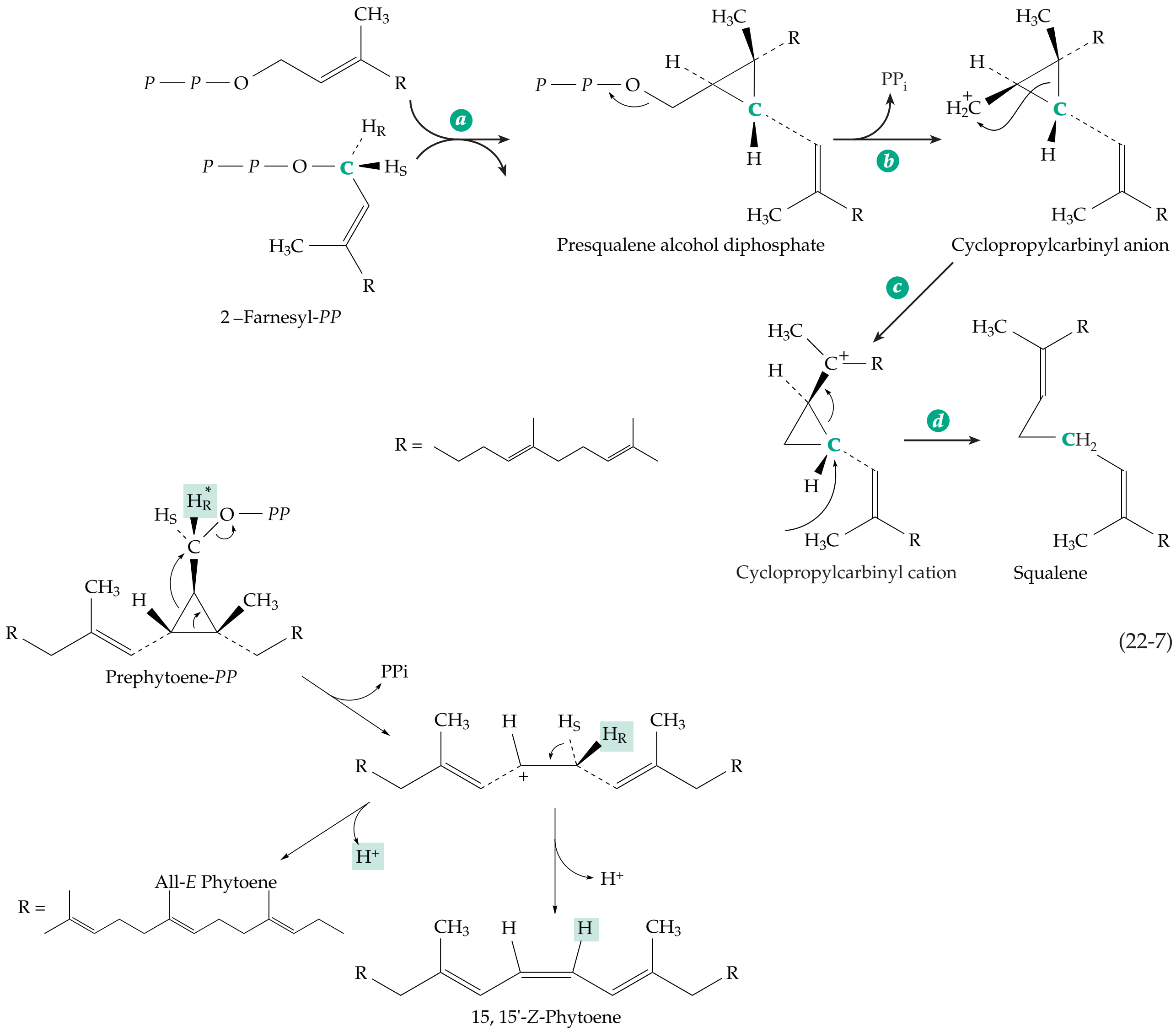
3. Formation of the Symmetric Terpenes, Squalene and Phytoene

Two molecules of the C_{15} farnesyl diphosphate can be joined "head to head" to form the C_{30} squalene (Eq. 22-6). Similarly, two C_{20} **geranylgeranyl-*PP*** molecules can be joined to form the C_{40} phytoene (Fig. 22-5), a precursor of carotenoid pigments of plants.[110] In the synthesis of squalene both pyrophosphate groups are eliminated from the precursor molecules, and one proton from C-1 of one of the molecules of farnesyl-*PP* is lost. The other three C-1 hydrogens are retained. At the same time, one proton is introduced from the *pro-S* position of NADPH. Squalene synthase has been difficult to obtain,[111,112] and its mechanism has been uncertain. However, there is strong evidence in favor of carbocationic intermediates.[113–114b] The first step (Eq. 22-7, step *a*) involves reaction of the initial carbocation with the double bond of the second farnesyl-*PP* to form the cyclopropane derivative **presqualene alcohol-*PP***, which was first isolated from yeast as the free alcohol by Rilling and associates.[115] The loss of the second pyrophosphate (Eq. 22-7, step *b*) generates a cyclopropylcarbinyl cation, which can rearrange (step *c*) to a more stable tertiary cation.[112,113] The latter is reduced by NADPH with opening of the cyclopropane ring (Eq. 22-7, step *d*). Once formed squalene diffuses within and between membranes with the help of cytosolic protein carriers.[116]

Phytoene (Fig. 22-5) is apparently formed from geranylgeranyl-*PP* via **prephytoene-*PP*,** whose structure is entirely analogous to that of presqualene-*PP*.[44,117] However, no reduction by NADH is required (Eq. 22-8). It is known that the 5-*pro-R* hydrogen atoms of mevalonate are retained in the phytoene as indicated by a shaded box in Eq. 22-8. Elimination of the other (*pro-S*) hydrogen yields 15,15'–*Z* phytoene (*cis*-phytoene), while elimination of the *pro-R* hydrogen yields all-*E* (trans) phytoene. Higher plants and fungi form mostly *cis*-phytoene, but some bacteria produce the all-*E* isomer.[118]



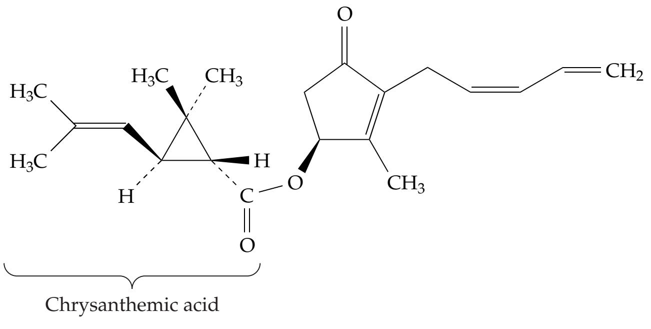
(22-6)


2 –Farnesyl-*PP*
a
Presqualene alcohol diphosphate
b
Cyclopropylcarbinyl anion
c
Cyclopropylcarbinyl cation
d
Squalene
R =
Prephytoene-*PP*
PPi
All-*E* Phytoene
R =
H+
15, 15'-Z-Phytoene

(22-7)

(22-8)

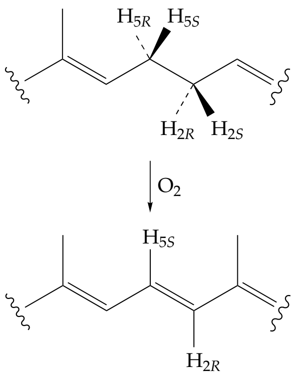
Another polyprenyl compound formed by a head to head condensation is **chrysanthemic acid**. This monoterpene component of the pyrethrum insecticides is formed by chrysanthemums from two molecules of dimethyallyl-*PP* via an intermediate analogous to presqualene alcohol-*PP*.[118a]


Chrysanthemic acid
Pyrethrin I

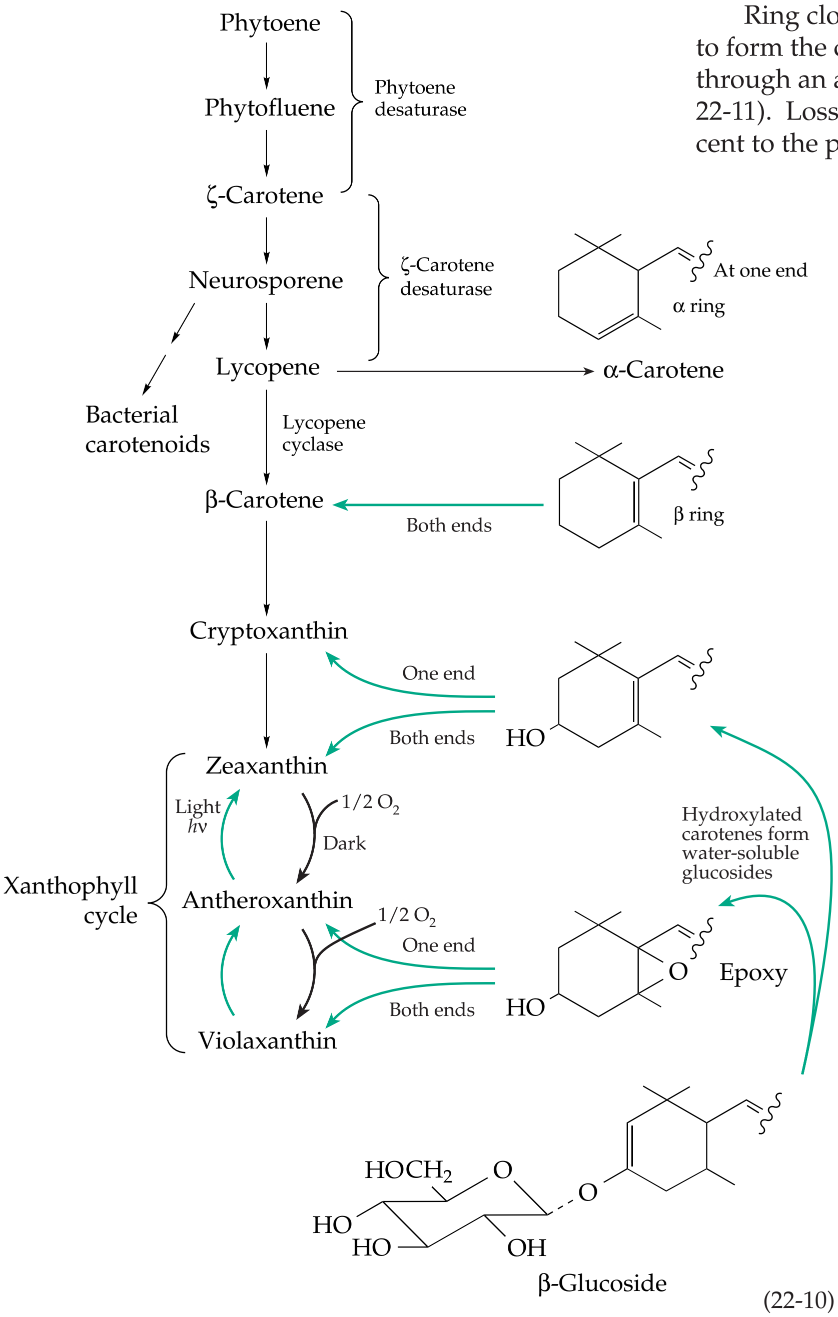
A quite different "tail to tail" condensation, whose chemistry is still obscure,[118b] must occur in archaebacteria whose lipids contain the C_{40} diphytanyl alcohol. An example is the diglyceryltetraether,[119,120] whose structure is shown on p. 388.

C. Carotenes and Their Derivatives

Phytoene can be converted to the carotenes by pathways indicated in part in Fig. 22-5 and Eq. 22-10. One of the first products is **lycopene**, the red pigment of tomatoes and watermelons, which is an all-trans compound. If 15-*Z* phytoene is formed, it must, at some point, be isomerized to an all-*E* isomer, and four additional double bonds must be introduced. The isomerization *may* be nonenzymatic. The double bonds are created by an oxygen-dependent **desaturation,** which occurs through the trans loss of hydrogen atoms.


H5R H5S
H2R H2S
O2
H5S
H2R

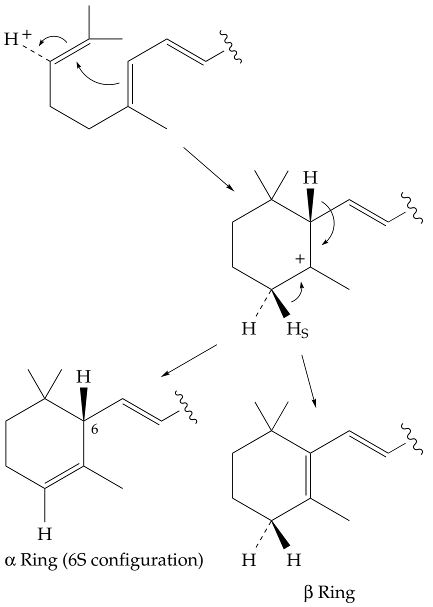
(22-9)


Phytoene
Phytofluene
Phytoene desaturase
ζ-Carotene
Neurosporene
ζ-Carotene desaturase
Lycopene
Bacterial carotenoids
Lycopene cyclase
α-Carotene
At one end
α ring
β-Carotene
Both ends
β ring
Cryptoxanthin
One end
Both ends
HO
Zeaxanthin
Light
hv
1/2 O2
Dark
Xanthophyll cycle
Antheroxanthin
1/2 O2
One end
Both ends
HO
O
Epoxy
Violaxanthin
Hydroxylated carotenes form water-soluble glucosides
HOCH2
O
O
HO
HO
OH
β-Glucoside

(22-10)

Desaturation takes place in a stepwise fashion, and many intermediate compounds with fewer double bonds are known (Eq. 22-10).[118,121–123] The enzymes required have not been characterized well until recently. Plant enzymes are present in small amounts, and isolation has been difficult. However, the genes for carotenoid biosynthesis in such bacteria as the purple photosynthetic *Rhodobacter*,[118,124] *Rhodospirillum*,[125] and *Rubrivarax*,[126] the cyanobacterium *Synechococcus*,[127] and the nonphotosynthetic *Erwinia*[44,118] have been cloned, sequenced, and used to produce enzymes in quantities that can be studied. Matching genes from higher plants have also been cloned and expressed in bacteria.[123]

Ring closure at the ends of the lycopene molecule to form the carotenes can be formulated most readily through an acid-catalyzed carbocation mechanism (Eq. 22-11). Loss of one or the other of two protons adjacent to the positive charge leads to the β ring of β-carotene or to the α ring of α-carotene.[110,123a] Compounds with only one ring may also be formed.[123b] In many bacteria these rings are not formed at all, but the open-chain (acyclic) cartenoids may be modified in ways similar to those of higher plants.[118,124,125]


H+
H
+
H
HS
H
6
H
α Ring (6S configuration)
H
H
β Ring

(22-11)

A genetic engineering success is the transfer of genes for synthesis of β-carotene into rice. The resulting "golden rice" contains enough carotene in its endosperm to make a significant contribution to the vitamin A needs of people for whom

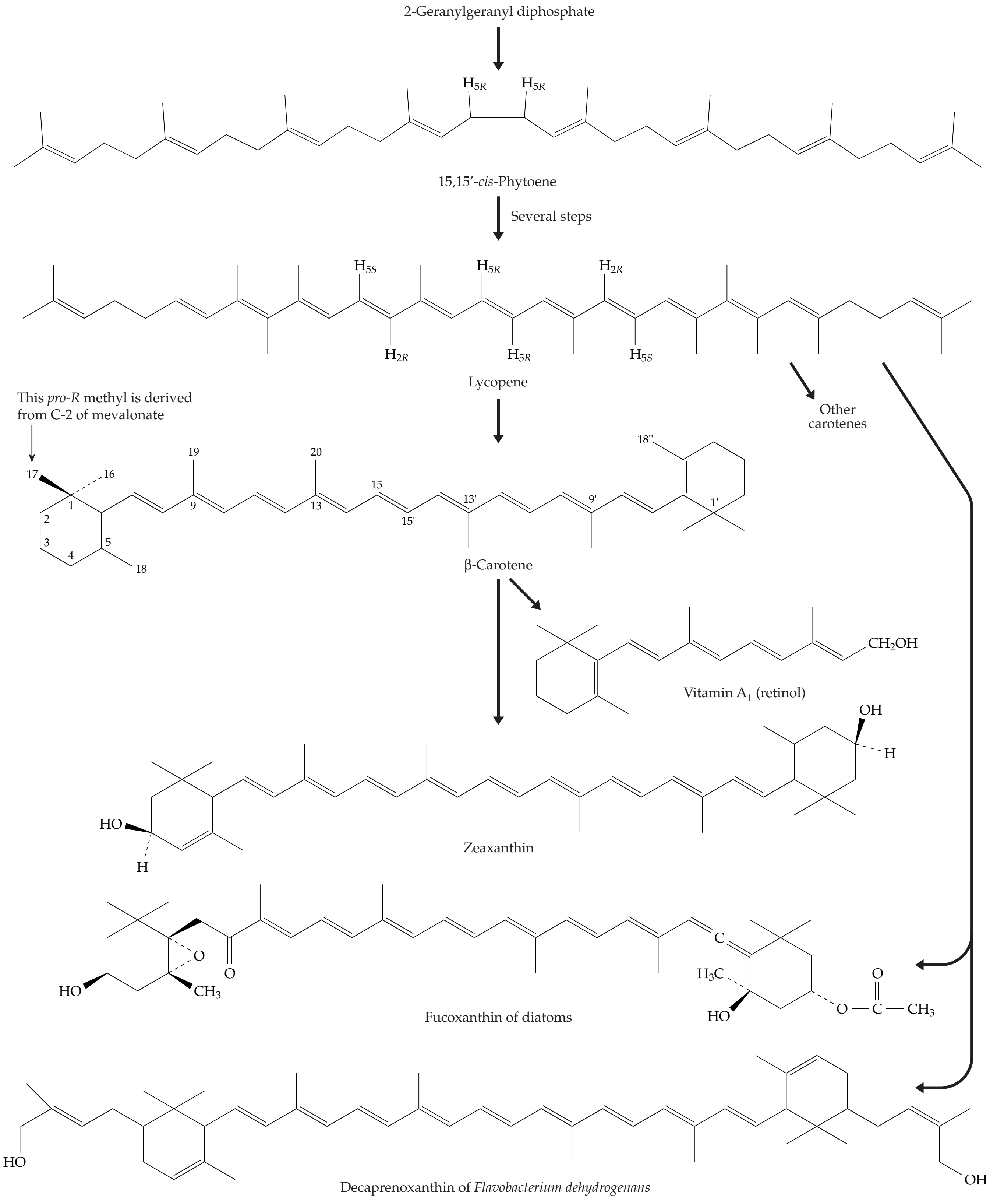
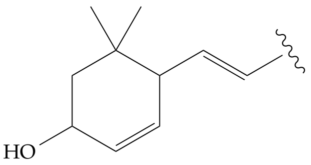

2-Geranylgeranyl diphosphate
H_{5R}
H_{5R}
15,15′-*cis*-Phytoene
Several steps
H_{5S}
H_{5R}
H_{2R}
H_{2R}
H_{5R}
H_{5S}
Lycopene
Other carotenes
This *pro-R* methyl is derived from C-2 of mevalonate
17
16
1
2
3
4
5
18
19
9
20
13
15
15'
13'
9'
18''
1'
β-Carotene
CH_2OH
Vitamin A_1 (retinol)
OH
H
HO
H
Zeaxanthin
O
O
HO
CH_3
C
H_3C
HO
O
O — C — CH_3
Fucoxanthin of diatoms
HO
OH
Decaprenoxanthin of *Flavobacterium dehydrogenans*

Figure 22-5 Structures and partial biosynthetic pathways for a few of the more than 600 known carotenoid compounds. The origin of some hydrogen atoms from mevalonate is shown, using the numbering for mevalonate. The numbering system for C_{40} carotenoids is also indicated.

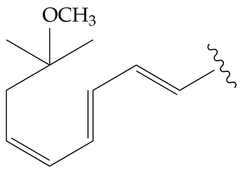
rice is a major food. There are estimated to be over 100 million vitamin A-deficient children in the world. As many as one-quarter of a million of these go blind each year.[127a] It is hoped that the golden rice will help to alleviate the problem. At the same time, an ongoing program is supplying vitamin A, which is stored in the liver, at regular intervals of time to many children.[127b]

1. Xanthophylls and Other Oxidized Carotenes

Carotenes can be hydroxylated and otherwise modified in a number of ways.[110,128–131] The structure of zeaxanthin, one of the resulting **xanthophylls**, is indicated in Fig. 22-5. Some other xanthophylls are shown in Eq. 22-10. Lutein resembles zeaxanthin, but the ring at one end of the chain has been isomerized by a shift in double bond position to the accompanying structure. The photosynthetic bacterium *Rhodospirillum rubrum* has its own special carotenoid spirilloxanthin, which has the accompanying structure at both ends of the chain.



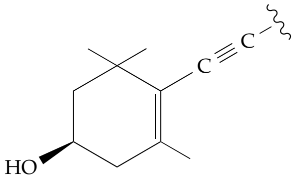
Lutein (structure at one end of molecule)



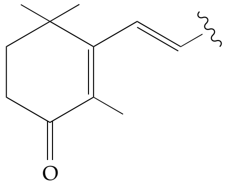
Spirilloxanthin

Fucoxanthin (Fig. 22-5) is the characteristic brown pigment of diatoms. One end of the molecule has an epoxide, also formed by the action of O_2, while the other end contains an **allene** structure rare in nature. Even so, fucoxanthin may be the most abundant carotenoid of all. The structure of the allene-containing end of the fucoxanthin molecule (turned over from that shown in Fig. 22-5) is also given in Eq. 22-12. Figure 22-5 does not indicate the stereochemistry of the allene group correctly; the carotenoid chain protrudes behind the ring as drawn in the equation. **Violaxanthin** contains epoxide groups in the rings at both ends of the molecule (Eq. 22-12). An isomerase in algae converts violaxanthin into **neoxanthin** (Eq. 22-12), which contains the allene structure at one end. Subsequent acetylation yields fucoxanthin.

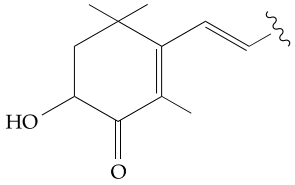
Other algal carotenoids contain acetylenic triple bonds. For example, **alloxanthin** has the following structure at both ends of the symmetric molecule. The symmetric carotenoids **canthaxanthin** and **astaxanthin** have oxo groups at both ends:



Alloxanthin



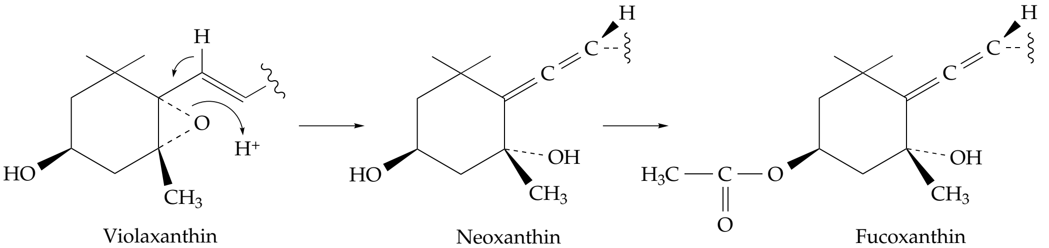
Canthaxanthin



Astaxanthin

These carotenoids have a limited distribution and occur as complexes, perhaps in Schiff base linkage, with proteins. Astaxanthin–protein complexes with absorption maxima ranging from 410 nm to 625 nm or more provide the color to the lobster's exoskeleton.[130,132] Whereas most naturally occurring carotenoids have all-*E* double bonds, mono-*Z* isomers of canthaxanthin are found in the colored carotenoproteins of the brine shrimp *Artemesia*.[133]

Some bacteria synthesize C_{50} carotenoids such as decaprenoxanthin (Fig. 22-5), the extra carbon atoms at each end being donated from additional prenyl groups, apparently at the stage of cyclization of lycopene.[134] Thus, a carbocation derived by elimination of pyrophosphate from dimethylallyl-*PP* could replace the H^+ shown in the first step of Eq. 22-11. The foregoing descriptions deal with only a few of the many known structural modifications of carotenoids.[2,135,136]



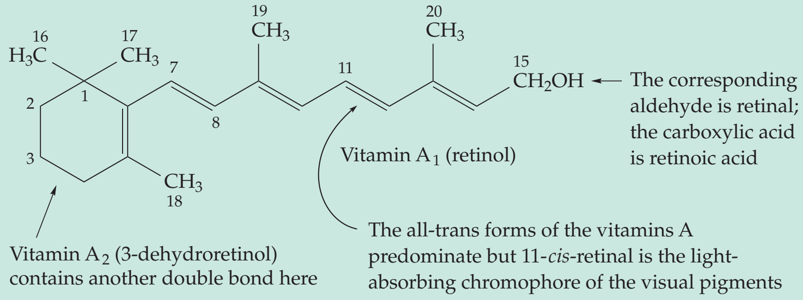
Violaxanthin (ring structure present at both ends)

Neoxanthin (this structure occurs at one end only)

Fucoxanthin

(22-12)

BOX 22-A VITAMIN A



Vitamin A$_1$ (retinol)

The corresponding aldehyde is retinal; the carboxylic acid is retinoic acid

Vitamin A$_2$ (3-dehydroretinol) contains another double bond here

The all-trans forms of the vitamins A predominate but 11-*cis*-retinal is the light-absorbing chromophore of the visual pigments

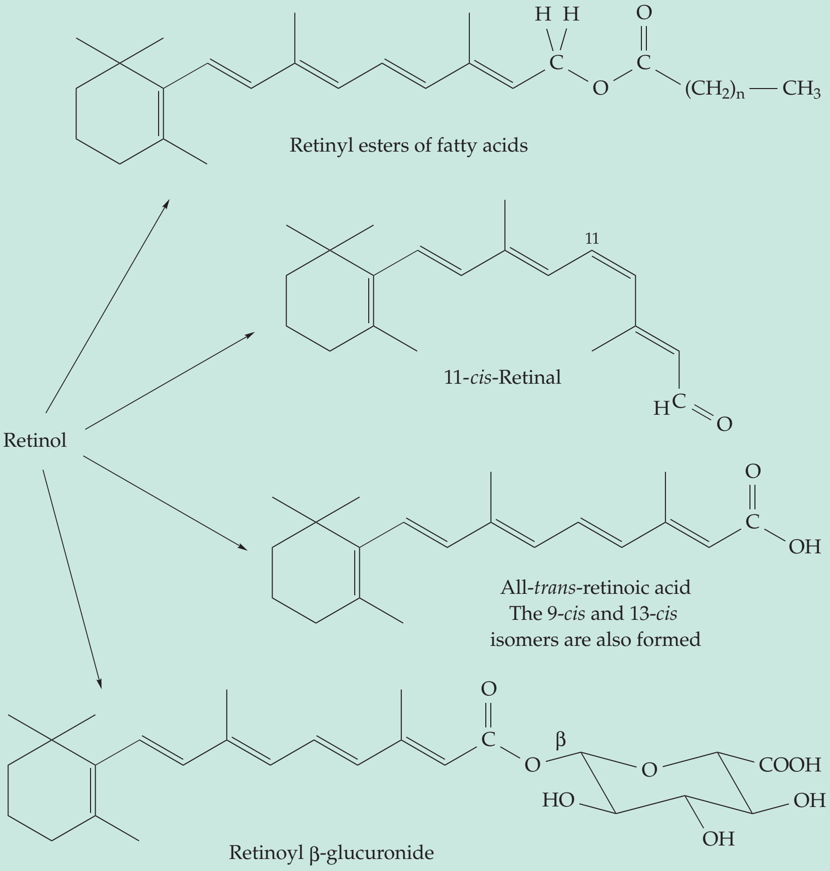
The recognition in 1913 of vitamin A (Box 14-A) was soon followed by its isolation from fish liver oils.[a–c] Both vitamin A$_1$ (**retinol**) and vitamin A$_2$ are 20-carbon polyprenyl alcohols. They are formed by cleavage of the 40-carbon β-carotene (Fig. 22-5) or other carotenoids containing a β-ionone ring. While the carotenes are plant products, vitamin A is produced only in animals, primarily within cells of the intestinal mucosa.[d–f] The carotene chains are cleaved in the center, and to some extent in other positions,[g] by oxygenases; β-carotene yields as many as two molecules of the vitamin A aldehyde **retinal**.[h–i] The retinal is reduced by NADH to retinol which is immediately esterified, usually with saturated fatty acids, by transfer of an acyl group from a fatty acyl-CoA or from phosphatidylcholine. The resulting retinyl esters are transported in chylomicrons. They remain in the chylomicron remnants (see Fig. 21-1), which are taken up by the liver where both hydrolysis and reesterification occur.[j] Vitamin A is one of the few vitamins that can be stored in animals in relatively large quantities. It accumulates in the liver, mainly as retinyl palmitate, in special storage cells termed stellate cells.[f] The human body usually contains enough vitamin A to last for several months.

Free retinol is released from the liver as a 1:1 complex of retinol with the 21-kDa **retinol-binding protein**.[k,l] This protein is normally almost saturated with retinol and is bound to another serum protein, the 127-residue **transthyretin** (prealbumin).[m,n] Some of the retinol is oxidized to **retinoic acid**. Both all-*trans* and 13-*cis*-retinoic acids as well as 5,6-epoxyretinoic acid are found in tissue.[f,o,p,pq]

Another metabolite, which may be very important, is **retinoyl β-glucuronide**.[q,r,rs]

Cell surfaces of body tissues appear to contain receptors for the retinol-binding protein. Many cells also contain cytoplasmic retinol-binding proteins[s–u] as well as proteins that bind retinoic acid.[u–y] These proteins are members of the large superfamily of hydrophobic transporter molecules described in Box 21-A. This includes the milk protein **β-lactoglobulin,** which also forms a complex with retinol.[z,aa]

A strikingly early symptom of vitamin A deficiency is **night blindness**. A variety of other symptoms include dry skin and hair, conjunctivitis of the eyes, retardation of growth, and low resistance to infection. The skin symptoms are particularly noticeable in the internal respiratory passages and alimentary canal lining. About 0.7 mg/day of vitamin A is required by an adult. The content of vitamin A in foods is often expressed in terms of international units: 1.0 mg of retinol equals 3333 I.U.



Retinyl esters of fatty acids

11-*cis*-Retinal

All-*trans*-retinoic acid
The 9-*cis* and 13-*cis* isomers are also formed

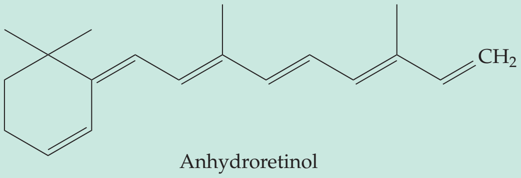
Retinoyl β-glucuronide

BOX 22-A VITAMIN A (continued)

Vitamin A, as retinal, has a clearly established role in vision (Chapter 23) and apparently has a specialized function in reproduction. In vitamin A deficiency no sperm cells are formed in males, and fetal resorption occurs in females. Rats deprived of vitamin A but fed retinoic acid become blind and sterile but otherwise appear healthy.[e,bb] Evidently either the alcohol or the aldehyde has an essential function in reproduction, whereas bone growth and maintenance of mucous secretions requires only retinoic acid. Indeed, retinoic acid is 100 to 1000 times more active than other forms of vitamin A in these differentiation functions.[r]

In vitamin A deficiency the internal epithelial surfaces of lungs and other tissues, which are usually rich in mucous secreting cells and in ciliated cells, develop thick layers of keratinizing squamous cells similar to those on the external surface of the body. The synthesis of some mannose- and glucosamine-containing glycoproteins consequently decreases.[cc] The major effects of retinoic acid is evidently through regulation of transcription (Chapter 28). In developing lungs retinoic acid promotes the transformation of undifferentiated epithelial cells into mucus-secreting cells.[dd]

Do we know all of the special chemistry of vitamin A that is involved in its functions? Retinal could form Schiff bases with protein groups as it does in the visual pigments. Redox reactions could occur. Conjugative elimination of water from retinol to form **anhydroretinol** is catalyzed nonenzymatically by HCl. Anhydroretinol occurs in nature and may serve as an inhibitor of the action of 14-hydroxy-*retro*-retinol in lymphocyte differentiation.[ee,ff]



Anhydroretinol

Much recent interest has been aroused by the fact that retinoid compounds, including both retinol and retinoic acid, reduce the incidence of experimentally induced cancer. In addition, 13-*cis*-retinoic acid taken orally is remarkably effective in treatment of severe cystic acne.[gg] However, both vitamin A and retinoic acid in large doses are **teratogenic**, i.e., they cause fetal abnormalities. The use of 13-*cis*-retinoic acid during early phases of pregnancy led to a high incidence of major malformations in infants born.[hh]

---

[a] Moore, T. (1981) *Trends Biochem. Sci.* **6**, 115–116

[b] Moore, T. (1957) *Vitamin A*, Elsevier, Amsterdam

[c] Olson, J. A. (1968) *Vitam. Horm.(N. Y.)* **26**, 1–63

[d] Blomhoff, R., ed. (1994) *Vitamin A in Health and Disease*, Dekker, New York

[e] Goodman, D. S. (1984) *N. Engl. J. Med.* **310**, 1023–1031

[f] Blomhoff, R., Green, M. H., Berg, T., and Norum, K. R. (1990) *Science* **250**, 399–404

[g] Tang, G., Wang, X.-D., Russell, R. M., and Krinsky, N. I. (1991) *Biochemistry* **30**, 9829–9834

[h] Redmond, T. M., Gentleman, S., Duncan, T., Yu, S., Wiggert, B., Gantt, E., and Cunningham, F. X., Jr. (2001) *J. Biol. Chem.* **276**, 6560–6565

[hi] Kiefer, C., Hessel, S., Lampert, J. M., Vogt, K., Lederer, M. O., Breithaupt, D. E., and von Lintig, J. (2001) *J. Biol. Chem.* **276**, 14110–14116

[i] Symons, M. (1996) *Trends Biochem. Sci.* **21**, 178–181

[j] Blomhoff, R., Rasmussen, M., Nilsson, A., Norum, K. R., Berg, T., Blaner, W. S., Kato, M., Mertz, J. R., Goodman, D. S., Eriksson, U., and Peterson, P. A. (1985) *J. Biol. Chem.* **260**, 13560–13565

[k] Zanotti, G., Ottonello, S., Berni, R., and Monaco, H. L. (1993) *J. Mol. Biol.* **230**, 613–624

[l] Zanotti, G., Marcello, M., Malpeli, G., Folli, C., Sartori, G., and Berni, R. (1994) *J. Biol. Chem.* **26**, 29613–29620

[m] Monaco, H. L., Rizzi, M., and Coda, A. (1995) *Science* **268**, 1039–1041

[n] Blake, C. C. F., Geisow, M. J., Swan, I. D. A., Rerat, C., and Rerat, B. (1974) *J. Mol. Biol.* **88**, 1–12

[o] Napoli, J. L., Khalil, H., and McCormick, A. M. (1982) *Biochemistry* **21**, 1942–1949

[p] Napoli, J. (1996) *FASEB J.* **10**, 993–1001

[pq] Tryggvason, K., Romert, A., and Eriksson, U. (2001) *J. Biol. Chem.* **276**, 19253–19258

[q] Barua, A. B., and Olson, J. A. (1986) *Am. J. Clin. Nutr.* **43**, 481–485

[r] Formelli, F., Barua, A. B., and Olson, J. A. (1996) *FASEB J.* **10**, 1014–1024

[rs] Cullum, M. E., and Zile, M. H. (1985) *J. Biol. Chem.* **260**, 10590–10596

[s] Rong, D., Lovey, A. J., Rosenberger, M., d'Avignon, A., Ponder, J., and Li, E. (1993) *J. Biol. Chem.* **268**, 7929–7934

[t] Quadro, L., Blaner, W. S., Salchow, D. J., Vogel, S., Piantedosi, R., Gouras, P., Freeman, S., Cosma, M. P., Colantuoni, V., and Gottesman, M. E. (1999) *EMBO J.* **18**, 4633–4644

[u] Newcomer, M. E. (1995) *FASEB J.* **9**, 229–239

[v] Ross, A. C. (1993) *FASEB J.* **7**, 317–327

[w] Thompson, J. R., Bratt, J. M., and Banaszak, L. J. (1995) *J. Mol. Biol.* **252**, 433–446

[x] Chen, X., Tordova, M., Gilliland, G. L., Wang, L., Li, Y., Yan, H., and Ji, X. (1998) *J. Mol. Biol.* **278**, 641–653

[y] Mansfield, S. G., Cammer, S., Alexander, S. C., Muehleisen, D. P., Gray, R. S., Tropsha, A., and Bollenbacher, W. E. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6825–6830

[z] Godovac-Zimmermann, J. (1988) *Trends Biochem. Sci.* **13**, 64–66

[aa] Cho, Y., Batt, C. A., and Sawyer, L. (1994) *J. Biol. Chem.* **269**, 11102–11107

[bb] Clamon, G. H., Sporn, M. B., Smith, J. M., and Saffiotti, V. (1974) *Nature (London)* **250**, 64–66

[cc] Rossa, G. C., Bendrick, C. J., and Wolf, G. (1981) *J. Biol. Chem.* **256**, 8341–8347

[dd] Chytil, F. (1996) *FASEB J.* **10**, 986–992

[ee] Buck, J., Derguini, F., Levi, E., Nakanishi, K., and Hammerling, U. (1991) *Science* **254**, 1654–1656

[ff] Grün, F., Noy, N., Hämmerling, U., and Buck, J. (1996) *J. Biol. Chem.* **271**, 16135–16138

[gg] Greenwald, D., DeWys, W., Black, G., Malone, W., Sporn, M., and Underwood, B. A. (1984) *Science* **224**, 338

[hh] Lammer, E. J., Chen, D. T., Hoar, R. M., Agnish, N. D., Benke, P. J., Braun, J. T., Curry, C. J., Fernhoff, P. M., Grix, A. W., Jr., Lott, I. T., Richard, J. M., and Sun, S. C. (1985) *N. Engl. J. Med.* **313**, 837–841

2. Properties and Functions of Carotenes

The most characteristic property of carotenoids is the striking color, most often yellow to red, which is used by birds as a sexual attractant and by plants to attract pollinators.[137,138] The associated light absorption fits these compounds for a role in photosynthetic light-harvesting,[139] in photoprotection, and in photoreception,[140] matters that are dealt with in Chapter 23. One aspect of photoprotection, which involves the **xanthophyll cycle,** is also indicated in Eq. 22-10. The cycle allows green plants to adjust to varying light intensity by altering the amount of zeaxanthin available for quenching excessive amounts of photoexcited chlorophyll (Chapter 23). Zeaxanthin undergoes epoxidation by O_2 to form antheroxanthin and violaxanthin as shown in Eq. 22-10. The process requires NADPH and reduced ferredoxin.[131] When light intensity is high the process is reversed by an ascorbate-dependent violaxanthin **de-epoxidase**.[99,128,141]

Violaxanthin also functions as a precursor to the plant hormone abscisic acid. Compare the structure of the latter (Fig. 22-4) with those of carotenoids. Oxidative cleavage of violaxanthin or related epoxy-carotenoids initiates the pathway of synthesis of this hormone.[142,143]

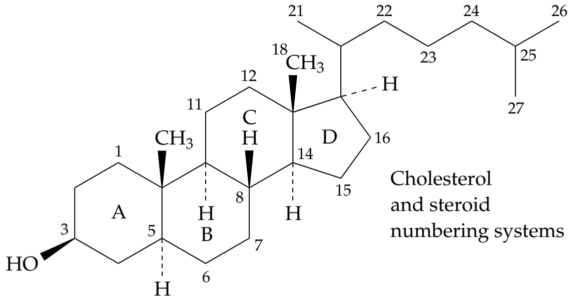
The system of conjugated double bonds responsible for carotenoid colors also helps to impart specific shapes to these largely hydrophobic molecules and ensures that they occupy the appropriate niches in the macromolecular complexes with which they associate. Information on stereochemistry is provided in a short review by Britton.[138]

β-Carotene, which can serve as an antioxidant at low oxygen pressures and can quench singlet oxygen,[144,145] has been associated with a reduced incidence of lung cancer.[137,146] While most animals do not synthesize carotenoids, they use them to make vitamin A and related retinoids and also as colorants. Yellow and red pigments of bird feathers[147] and the colors of tissues of salmon and of lobsters and other invertebrates are derived from dietary carotenoids, which are often modified further by the new host. The lobster accumulates astaxanthin, as a blue protein complex,[148] and the flamingo uses the astaxanthin of shrimp to color its feathers.[130]

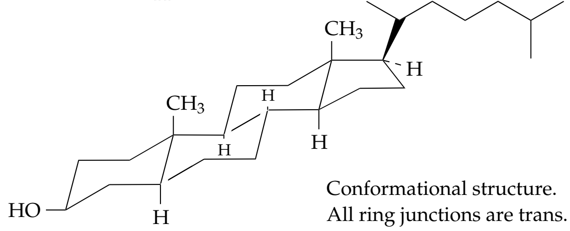
Dietary carotenes and carotenoids are absorbed and transported in the plasma of humans and animals by lipoproteins.[149] The conversion of carotenes to vitamin A (Box 22-A) provides the aldehyde **retinal** for synthesis of visual pigments (Chapter 23) and **retinoic acid**, an important regulator of gene transcription and development (Chapter 32).[150–152c] See also Section E,5.

D. Steroid Compounds

The large class of **steroids** contains a characteristic four-ring nucleus consisting of three fused six-membered rings and one five-membered ring.[153] **Cholestanol** (dihydrocholesterol) may be taken as a representative steroid alcohol or **sterol**. Most sterols, including cholestanol, contain an 8- to 10-carbon side chain at position 17. The polyprenyl origin of the side chain is suggested by the structure. Steroid compounds usually contain an oxygen atom at C-3. This atom is present in an –OH group in the sterols and frequently in a carbonyl group in other steroids. Most steroids contain two axially oriented methyl groups, the "angular methyl groups,"which are attached to the ring system and numbered C-18 and C-19. In the customary projection formulas they are to be thought of as extending forward toward the viewer. In the same manner, the equatorially oriented 3-OH group of cholestanol and the side chain at C-17 also project forward toward the viewer in the projection formula.



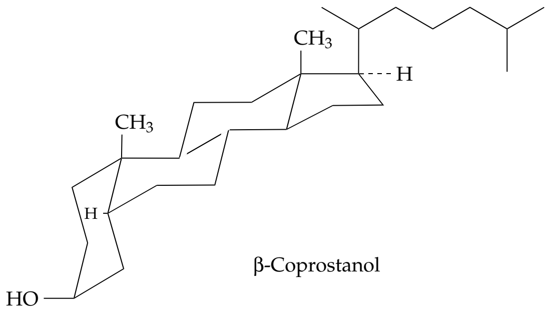
Cholesterol and steroid numbering systems



Conformational structure. All ring junctions are trans.

The angular methyl groups, the 3-OH groups, and the side chain of cholestanol are all on the same side of the steroid ring in the projection formula and are all said to have a **β orientation**. Substituents projecting from the opposite side of the ring system are **α oriented**. While the methyl groups (C-18 and C-19) almost always have the β orientation, the 3-OH group has the α orientation in some sterols. Dashed lines are customarily used to connect α-oriented substituents, and solid lines are used for β-oriented substituents in structural formulas. Cholesterol is chiral and its enantiomer does not support life for *C. elegans* and presumably for other organisms.[153a]

In cholestanol the ring fusions between rings A and B, B and C, and C and D are all trans; that is, the hydrogen atoms or methyl groups attached to the bridgehead carbon atoms project on opposite sides of the ring system. This permits all three of the six-membered rings to assume relatively unstrained chair conformations. However, the introduction of a double bond alters the shape of the molecule significantly. Thus, in cholesterol the double bond between C-5 and C-6 (Δ^5) distorts both the A and B rings from the chair conformation. In some steroid compounds the junction between rings A and B is cis. This greatly alters the overall shape of the steroid from the relatively flat one of cholestanol to one that is distinctly bent. An example is **β-coprostanol**, a product of bacterial action on cholesterol and a compound occurring in large amounts in the feces. In some sterols, notably the estrogenic hormones, ring A is completely aromatic and the methyl group at C-19 is absent.



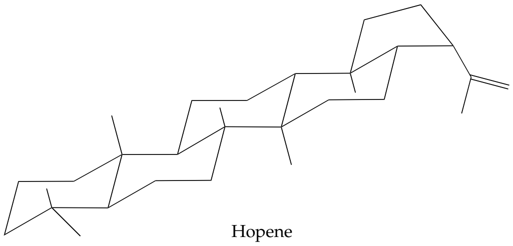
β-Coprostanol

1. Biosynthesis of Sterols

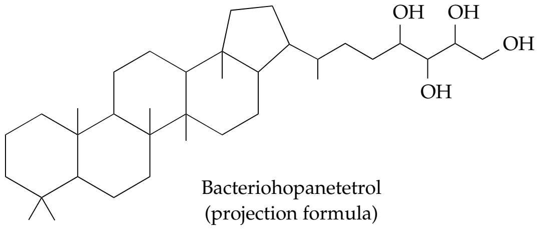
Most animal steroids arise from cholesterol, which in turn is derived from squalene. This C_{30} triterpene, whose biosynthesis is described in Section B, is named after the dogfish *Squalus* in whose liver it accumulates as a result of blockage in oxidation to cholesterol. Squalene is also a prominent constituent of human skin lipids. Its conversion to **cholesterol**, which takes place in most animal tissues,[117,154–156] is initiated by a microsomal enzyme system that utilized O_2 and NADPH to form **squalene 2,3-oxide** (Fig. 22-6, step *a*). The subsequent cyclization reaction, which probably takes place through a carbocation created by attack of a proton on the oxygen atom of the epoxide ring (Fig. 22-6, step *b*), is catalyzed by the large 70- to 80-kDa **oxidosqualene cyclase**.[157–159] The enzyme from rat liver consists of 733 residues and contains a highly conserved sequence with two consecutive aspartates that are thought to be at the active site. The sequence is somewhat similar to that of prenyltransferases and sesquiterpene cyclases.[157] The cyclization step appears to require that the enzyme hold the substrate in a rigid conformation as indicated in Fig. 22-6. The flow of electrons effects the closure of all four rings. The carbocation created at C-2 of squalene (C-4 of the sterol ring that is formed) by opening of the epoxide ring reacts with electrons from the 6,7 double bond to close ring A leaving a carbocation at C-6. This in turn reacts with the 10,11 double bond leaving a carbocation at C-10, etc. At the end of this cascade a carbocation is left on C-19 of squalene, which is numbered C-20 in the incipient sterol. The closures of rings A and B both follow the Markovnikov rule by generating relatively stable tertiary carbocations. Thus, the natural chemical reactivities of the substrates are followed in these enzymatic reaction steps. However, this is not the case in the closure of ring C to form a 6-membered ring instead of a 5-membered ring. This presumably happens because the enzyme imposes the correct geometry for a 6-membered ring on the squalene and the correct stereochemistry on the ring closure.[160]

The rearrangement of this initially created C-20 carbocation to **lanosterol** (Fig. 22-6, step *c*) is also a remarkable reaction that requires the shift of a hydride ion and of two methyl groups, as indicated by the arrows in the figure. In addition, a hydrogen at C-9 (sterol numbering) is lost as a proton. Lanosterol is named for its occurrence in lanolin, the waxy fat in wool. Although the principal component of lanolin is cholesterol, lanosterol is its precursor both in sheep and in all other animals. Cholesterol is in turn the precursor to other animal sterols. The cholesterol biosynthetic pathway also provides cells with a variety of important signaling molecules.[160a]

In green plants, which contain little or no cholesterol, **cycloartenol** is the key intermediate in sterol biosynthesis.[161–162a] As indicated in Fig. 22-6, step *c*′, cycloartenol can be formed if the proton at C-9 is shifted (as a hydride ion) to displace the methyl group from C-8. A proton is lost from the adjacent methyl group to close the cyclopropane ring. There are still other ways in which squalene is cyclized,[162,163,163a] including some that incorporate nitrogen atoms and form alkaloids.[163b] One pathway leads to the **hopanoids**. These triterpene derivatives function in bacterial membranes, probably much as cholesterol does in our membranes. The three-dimensional structure of a bacterial hopene synthase is known.[164,164a] Like glucoamylase (Fig. 2-29) and farnesyl transferase, the enzyme has an $(\alpha,\alpha)_6$-barrel structure in one domain and a somewhat similar barrel in a second domain. The active site lies in a large interior cavity. The properties of the hopene synthase are similar to those of oxidosqualene synthase, and it appears to function by a similar mechanism, which resembles that of Fig. 22-6 but does not depend upon O_2. Hopene lacks polar groups, but these are provided in the hopanoids by a polyol side chain. One of these compounds, **bacteriohopanetetrol**, may be one of the most abundant compounds on earth.[160,165,166] Hopanoids appear to originate from mevalonate synthesized via the 1-deoxyxylulose pathway (Fig. 22-2). The polyol side chain is probably formed from ribose.[166]



Hopene



Bacteriohopanetetrol
(projection formula)

Formation of cholesterol. The conversion of lanosterol to cholesterol requires at least 19 steps,[167,168] which are catalyzed by enzymes bound to membranes of the ER. The removal of the three methyl groups of lanosterol, the migration of the double bond within the B ring, and the saturation of the double bond in the side chain may occur in more than one sequence, two of which are indicated in Fig. 22-7. The predominant pathway in many organisms including humans is the oxidative demethylation at the C / D ring junction (C-14) by a cytochrome P450 called **lanosterol 14α-demethylase**. This single enzyme catalyzes three consecutive O_2 and NADPH-dependent reactions that convert the methyl group to hydroxymethyl, formyl, and then free formate (right side of Fig. 22-8).[169–172b] Steps *a* and *b* are typical cytochrome P450 oxygenation reactions. In step *b* a geminal diol is formed and is dehydrated to the formyl derivative. The third step is atypical. Shyadehi *et al.* proposed the sequence depicted in steps *c-f* in which an Fe(III) peroxo intermediate reacts as shown.[172] This mechanism is supported by the fact that both ^{18}O present in the formyl group and ^{18}O from $^{18}O_2$ appear in the liberated formate.

The corresponding reactions of the methyl groups at C-4 on the A ring[167,168,173] are depicted on the left side of Fig. 22-8. The 4α methyl group is first hydroxylated by a microsomal (ER) system similar to cytochrome P450 but able to accept electrons from NADH and cytochrome b_5 rather than NADPH.[173] The two-step oxidation of the resulting alcohol to a carboxylic acid is catalyzed by the same enzyme. A second enzyme catalyzes the dehydrogenation of the 3-OH group to a ketone allowing for efficient β-decarboxylation (Fig. 22-8, steps *j* and *k*).[173a] Inversion of configuration at C-4, assisted by the 3-carbonyl group (step *l*), places the second 4-methyl group in the α orientation. After reduction of the 3-carbonyl by a third enzyme the sequence is repeated on this second methyl group.

In addition to the enzymes that are embedded in the membranes of the ER, conversion of lanosterol to cholesterol depends upon soluble cytoplasmic carrier proteins.[174] See also Box 21-A. Other sterols formed in

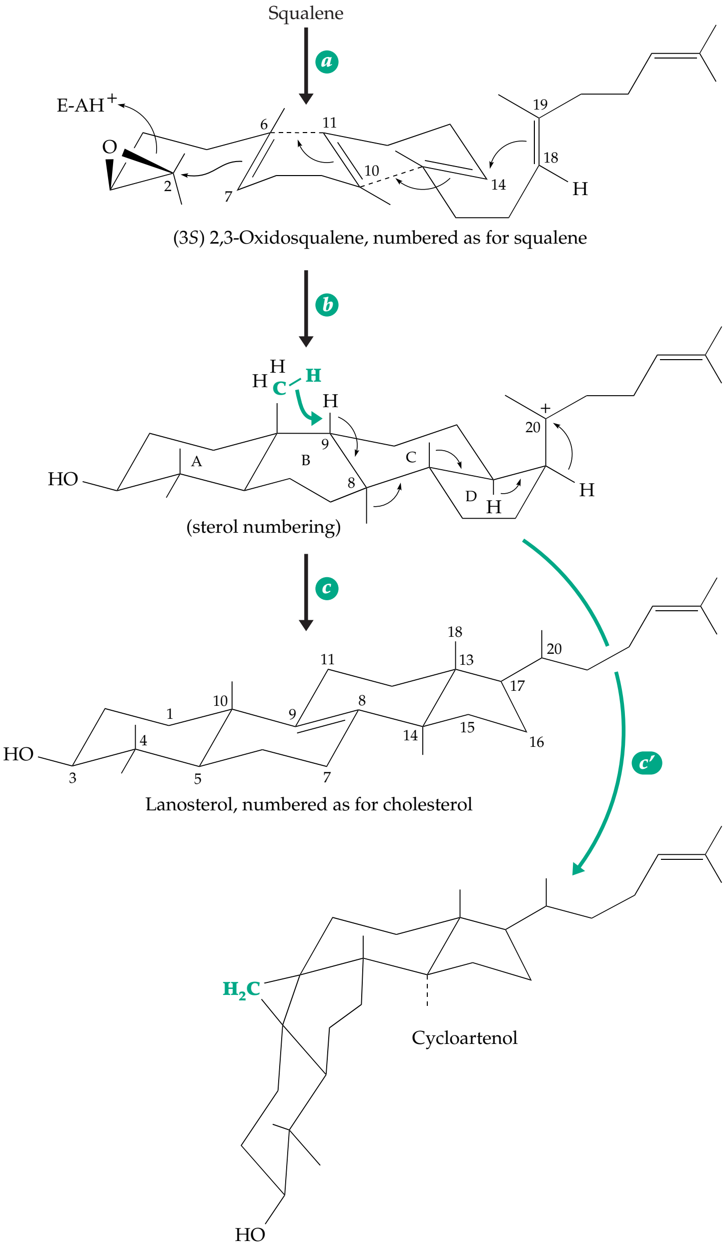
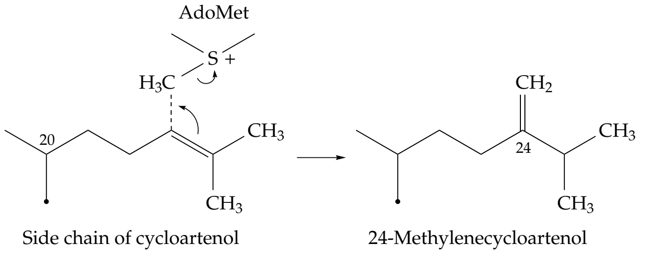


Figure 22-6 The cyclization of all-*trans* squalene to lanosterol and cycloartenol.

the animal body are **7-dehydrocholesterol**, prominent in skin and a precursor of vitamin D. Both β-cholestanol and its isomer β-coprostanol are formed by bacteria in the intestinal tract, and small amounts of cholesterol are converted to cholestanol within tissues. **Ergosterol**, the most common sterol in fungi, contains the $\Delta^{5,7}$ ring system of 7-dehydrocholesterol as well as an extra double bond in the side chain. It arises from zymosterol (Fig. 22-7).[173,174a]

As indicated previously, plant sterols are thought to be formed in most cases through cycloartenol which is often converted to **24-methylenecycloartenol**, a substance present in grapefruit peel and in many other plants. The methylene carbon is donated by *S*-adenosylmethionine (AdoMet) as shown in Eq. 22-13, which implies a transient intermediate carbocation. Saturation of the side chain and oxidative demethylation similar to that shown in Fig. 22-8[175] and introduction of a double bond[176] leads to **campesterol** (Fig. 22-9). It has the Δ^5-unsaturated ring of cholesterol but, like many other plant sterols, the side chain has one additional methyl group, which is also donated from *S*-adenosylmethionine.[161,177,178] Several more steps are required to convert campesterol into the plant steroid hormone **brassinolide**.[179–180] Among higher plants, **sitosterol** and **stigmasterol** are the most common sterols. Each contains an extra ethyl group in the side chain. Sitosterol is formed by the methylation (by AdoMet) of ergosterol. For the guinea pig stigmasterol is a vitamin, the"antistiffness factor" necessary to prevent stiffening of the joints. Some other plant steroids arise without addition of the extra carbons at C-23 or C-24 but usually via a different cyclization of squalene. Of these, the cucurbitacins (Fig. 22-9) are among the bitterest substances known.[181]


AdoMet
S +
H_3C
20
CH_3
CH_3
Side chain of cycloartenol
CH_2
24
CH_3
CH_3
24-Methylenecycloartenol

(22-13)

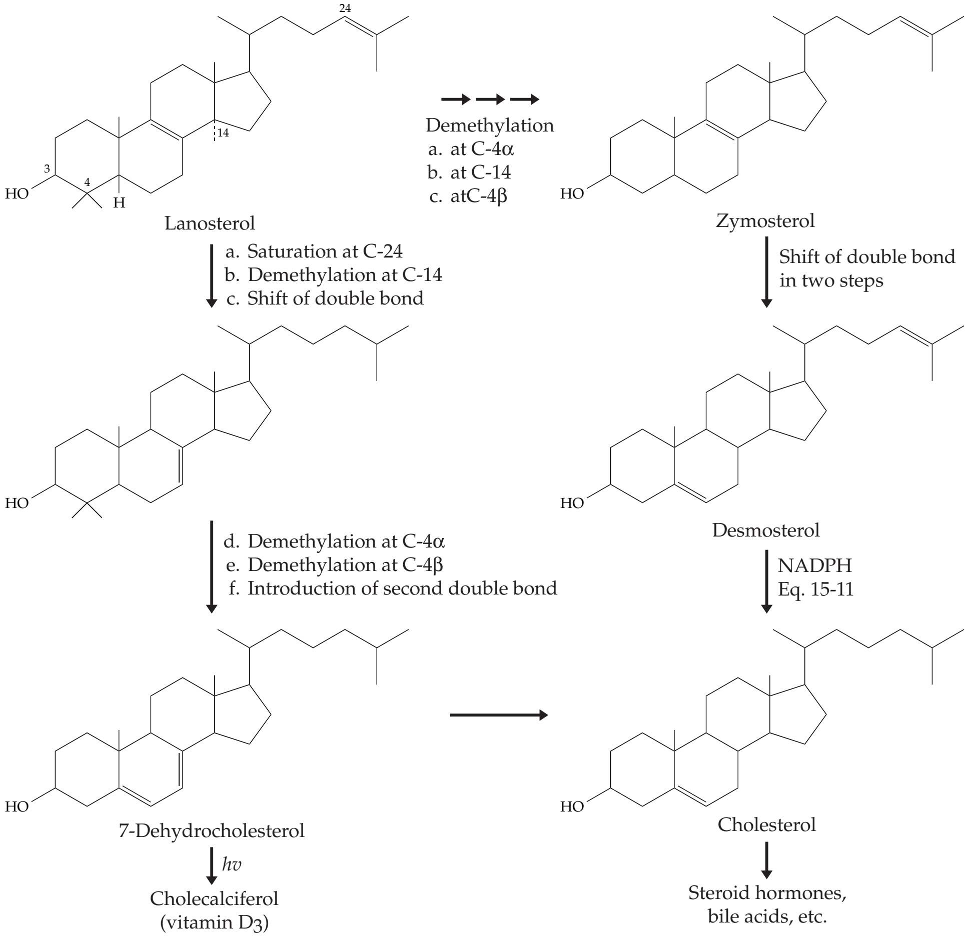

24
14
3
4
HO
H
Lanosterol
Demethylation
a. at C-4α
b. at C-14
c. atC-4β
HO
Zymosterol
a. Saturation at C-24
b. Demethylation at C-14
c. Shift of double bond
Shift of double bond
in two steps
HO
HO
Desmosterol
d. Demethylation at C-4α
e. Demethylation at C-4β
f. Introduction of second double bond
NADPH
Eq. 15-11
HO
HO
7-Dehydrocholesterol
Cholesterol
hv
Cholecalciferol
(vitamin D3)
Steroid hormones,
bile acids, etc.

Figure 22-7 Conversion of lanosterol to cholesterol. Two of many possible sequences are shown.

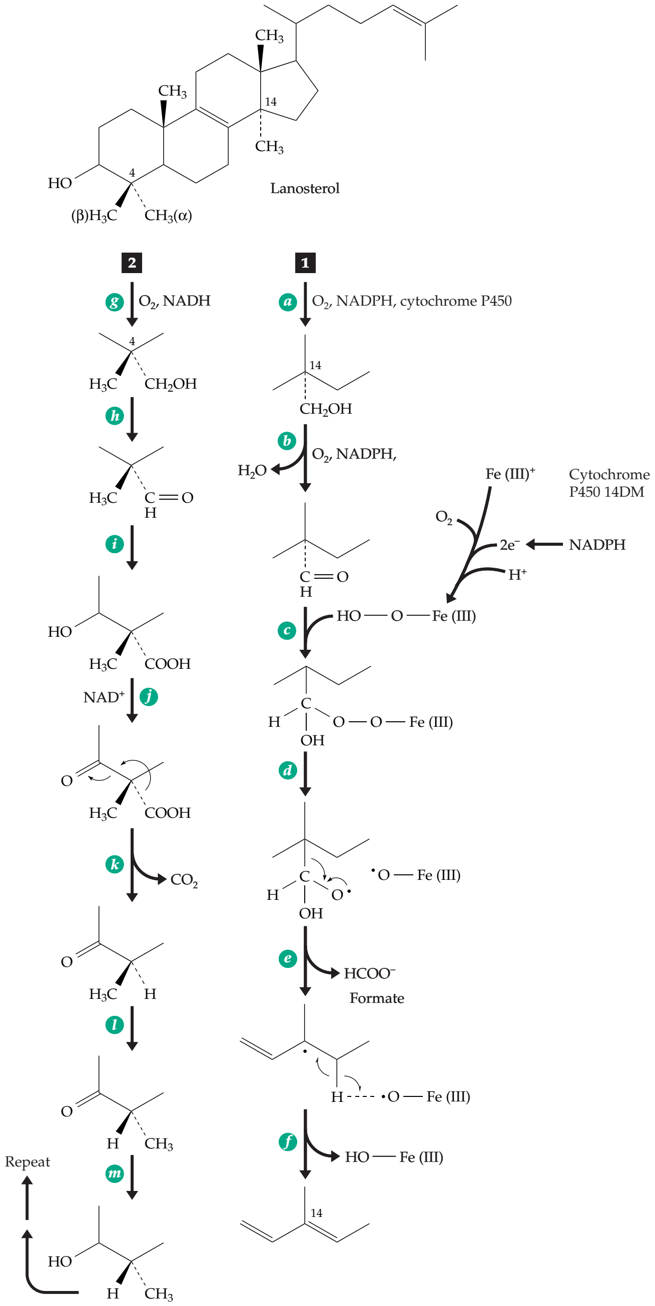


Figure 22-8 Steps in the demethylation of lanosterol. The most frequent sequence, labeled [1], begins with demethylation at C-14 by the action of a cytochrome P450 and is followed [2] by the successive demethylation of the α-CH_3 and β-CH_3 at C-4 by an NADH-dependent oxygenase.

2. Metabolism of Cholesterol in the Human Body

Cholesterol is both absorbed from the intestinal tract and synthesized from acetate via squalene, principally in the liver. The quantities produced are substantial. Daily biosynthesis is ~600 mg, and dietary uptake may supply another 300 mg.[182] Not only is there a large amount of cholesterol in the brain and other nervous tissues but also about 1.7 g of cholesterol per liter is present in blood plasma, about two-thirds of it being esterified principally to unsaturated fatty acids. The cholesterol content of plasma varies greatly with diet, age, and sex. By age 55 it averages 2.5 g / liter and may be considerably higher. Women up to the age of menopause have distinctly lower blood cholesterol than do men. Cholesterol regulates its own abundance by a variety of feedback mechanisms.[183,184] These include inhibition of the synthesis by means of reduced activities (step *a* of Fig. 22-1) of HMG-CoA reductase, farnesyl diphosphate synthase (step *g* of Fig. 22-1), and squalene synthase. All of these reactions are essential steps in cholesterol synthesis.[16,185] On the other hand, cholesterol induces an increase in acyl-CoA:cholesterol acyltransferase.

Dietary cholesterol, together with triacylglycerols, is absorbed from the intestinal tract and enters the large lipoprotein chylomicrons (see Fig. 21-1). Absorption of cholesterol is incomplete, usually amounting to less than 40% of that in the diet. Absorption requires bile salts and is influenced by other factors.[186] As it is needed cholesterol is taken from the plasma lipoproteins into cells by endocytosis. Much of the newly absorbed cholesterol is taken up by the liver. The liver also secretes cholesterol, in the form of esters with fatty acids, into the bloodstream.

Cholesterol is synthesized in the ER and other internal membranes by most cells of the body.[187,188] Newly formed cholesterol is sorted from the ER into the various membranes of the cell, the greatest abundance being in plasma membranes where cholesterol plays an essential role in decreasing fluidity. Cholesterol also aggregates with sphingolipids to form rigid lipid "rafts" floating in the plasma membrane. These rafts are thought to have important functions in signaling, in distribution of lipid materials,[188a] and in influencing protein translocation.[188b] Caveolae in cell surfaces may also arise from cholesterol-rich rafts.[188c] However, cholesterol must also be able to move out of the internal membranes back into the interior ER of the cell to provide for homeostasis and to allow formation of cholesteryl esters for transport, bile acids for excretion by liver, and the steroid hormones.[183,184,189] Movement between organelles occurs with the aid of sterol carrier proteins.[190–190c]

Liver and some intestinal cells export cholesterol into the bloodstream, together with triacylglycerols and phospholipids in the form of VLDL particles, for uptake by other tissues (see Fig. 21-1). Cholesteryl esters are formed in the ER by **lecithin:cholesterol acyltransferase** (LCAT), an enzyme that transfers the central acyl group from phosphatidylcholine to the hydroxyl group of cholesterol.[191,191a] This enzyme is also secreted by the liver and acts on free cholesterol in lipoproteins.[192] Tissue acyltransferases also form cholesteryl esters from fatty acyl-CoAs.[192a]

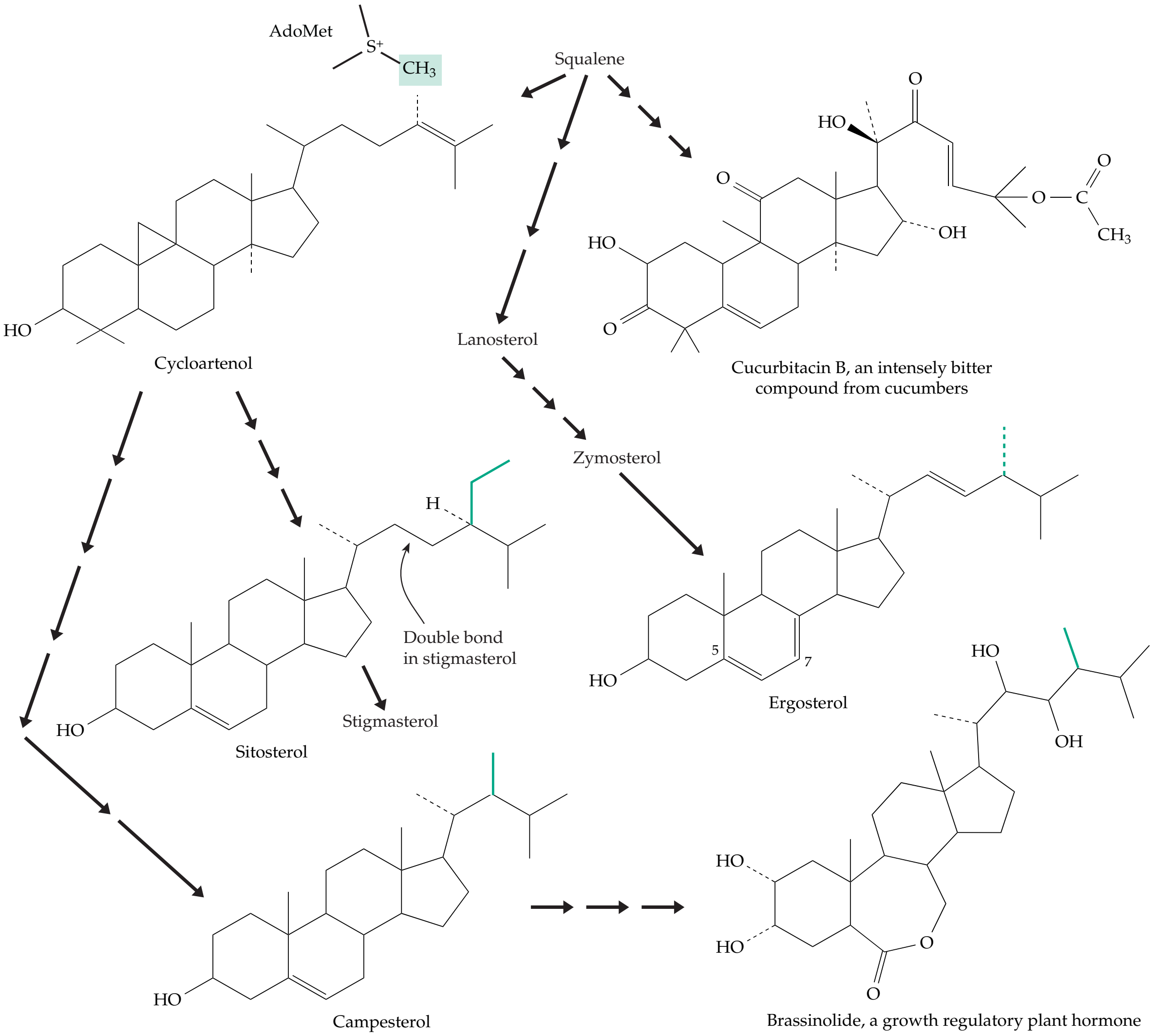

AdoMet
S+
CH3
Squalene
O
HO
O
OH
O
O — C
CH3
HO
O
Cucurbitacin B, an intensely bitter compound from cucumbers
HO
Cycloartenol
Lanosterol
Zymosterol
H
Double bond in stigmasterol
Stigmasterol
HO
Sitosterol
5
7
HO
Ergosterol
HO
OH
HO
Campesterol
HO
HO
O
O
Brassinolide, a growth regulatory plant hormone

Figure 22-9 Structures and routes of biosynthesis for a few plant steroids.

In the brain a complex of cholesterol with apolipoprotein E (Table 21-2) promotes the formation of new synapses (Chapter 30). Synthesis of cholesterol for this purpose appears to occur within glial cells.[192b,c]

Serum cholesterol. Most cholesterol is carried in the blood by low density lipoprotein (LDL, Tables 21-1, 21-2), which delivers the cholesteryl esters directly to cells that need cholesterol. Both a 74-kDa **cholesteryl ester transfer protein**[193–195a] and a **phospholipid transfer protein**[196,196a] are also involved in this process. **Cholesterol esterases**, which release free cholesterol, may act both on lipoproteins and on pancreatic secretions.[197–199]

The LDL–cholesterol complex binds to LDL receptors on the cell surfaces.[167,168,200–202] These receptors are specific for apolipoprotein B-100 present in the LDL. The occupied LDL–receptor complexes are taken up by endocytosis through coated pits; the apolipoproteins are degraded in lysosomes, while the cholesteryl esters are released and cleaved by a specific **lysosomal acid lipase**[203,204] to form free cholesterol.

While the primary role of LDL appears to be the transport of esterified cholesterol to tissues, the high density lipoproteins (HDL) carry excess cholesterol away from most tissues to the liver.[205–207] The apoA-I present in the HDL particle not only binds lipid but activates LCAT, which catalyzes formation of cholesteryl esters which migrate into the interior of the HDL and are carried to the liver.

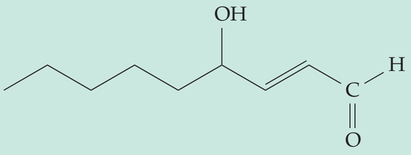
Unlike other lipoproteins, HDL particles are assembled outside of cells from lipids and proteins, some of which may be donated from chylomicrons (see Fig. 21-1) or other lipoprotein particles. HDL has a higher protein content than other lipoproteins and is more heterogeneous. The major HDL protein is apolipoprotein A-I, but many HDL particles also contain A-II,[205,208–210] and apolipoproteins A-IV, D, and E may also be present. A low plasma level of HDL cholesterol is associated with a high risk of atherosclerosis.[205,207]

BOX 22-B ATHEROSCLEROSIS

Our most common lethal disease is atherosclerosis, which causes constriction and blockage of arteries of the heart, brain, and other organs. In the United States, Europe, and Japan half of all deaths can be attributed to this ailment.[a,b] There seems to be a variety of causes. However, there is agreement that the disease begins with injury to the endothelial cells that form the inner lining of the arteries.[a,c,d] This is followed by the aggregation of blood platelets at the sites of injury and infiltration of smooth muscle cells, which may be attracted by 12-hydroxyeicosotetraenoic acid and other chemoattractants formed by activated platelets.[c] **"Foam cells"** laden with cholesterol and other lipids appear, and the lesions enlarge to become the characteristic plaques (**atheromas**).

The best understood cause of atherosclerosis is the genetic defect **familial hypercholesterolemia**, an autosomal dominant trait carried by one person in 500 all over the world.[e] Males with the defective gene tend to develop atherosclerosis when they are 35–50 years of age. The approximately one in a million persons *homozygous* for the trait develop coronary heart disease in their teens or earlier. Cultured fibroblasts from these patients have 40- to 60-fold higher levels of HMG-CoA reductase (Eq. 15-9) than are present normally, and the rate of cholesterol synthesis is increased greatly. The LDL level is very high and, as shown by Brown and Goldstein,[f–i] the LDL receptor gene is defective. Genetic defects associated with a low HDL level are also associated with atherosclerosis[b,j–l] as is a genetic variant of the metalloproteinase **stromelysin**.[m]

Other factors favoring development of atherosclerosis include hypertension and smoking. Chickens infected with a herpes virus (Marek disease virus) develop the disease after infection, and it is possible that artery damage in humans can also be caused by viruses[n] or bacterial infections.[o] In recent years it has been established that oxidative modification of the phospholipids in LDL induces the uptake of LDL by scavenger receptors of macrophages. This appears to trigger the development of foam cells and atherosclerotic plaques.[c,p,q] The initial damage is thought to be caused by lipid peroxides in the diet or generated by lipoxygenases in platelets and other cells.[c,p,r] Unsaturated fatty acids in lipoproteins can undergo oxidation (Chapter 21), especially in the presence of Cu^{2+} ions,[s,t] to yield malondialdehyde, 4-hydroxynonenal (Eq. 21-15), and other reactive compounds, which may damage the lipoproteins and cause them to have too high an affinity for their receptors in the smooth muscle cells of artery walls.[p,u] The 17β-hydroperoxy derivative of cholesterol has also been found in atherosclerotic lesions and may account for some of the toxicity of oxidized LDL.[r] Ascorbic acid may help to prevent formation of these oxidation products.[p,q,v] Chlorinated sterols may also be produced by the myeloperoxidase of the phagocytic macrophages that are abundant in atherosclerotic plaque.[w] Trans fatty acids, which are abundant in some margarines, and other hydrogenated fats raise both cholesterol and LDL levels.[x] Another cause of artery disease may be the presence of excessive homocysteine,[y,z] which can accumulate as a result of marginal deficiencies of folate, vitamin B_6, or vitamin B_{12}.



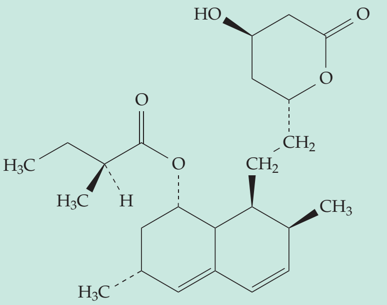
4-Hydroxy-2-nonenal

What can be done to prevent atherosclerosis? For persons with a high LDL level there is little doubt that a decreased dietary intake of cholesterol and a decrease in caloric intake are helpful. While such dietary restriction may be beneficial to the entire population, controlled studies of the effect of dietary modification on atherosclerosis have been disappointing and confusing.[aa] A diet that is unhealthy for some may be healthy for others. For example, an 88-year old man who ate 25 eggs a day for many years had a normal plasma cholesterol level of 150–200 mg / deciliter (3.9–5.2 mM)![bb] Comparisons of diets rich in unsaturated fatty acids, palmitic acid, or stearic acid have also been confusing.[cc,cd,dd] Can it be true that palmitic acid from tropical oils and other plant sources promotes atherogenesis, but that both unsaturated fatty acids and stearic acid from animal fats are less dangerous?

One of the best therapeutic approaches may be to prevent absorption of cholesterol from the intestines by inclusion of a higher fiber content in the diet.[ee] Supplementation with a cholesterol-binding resin may provide additional protection. Plant sterols also interfere with cholesterol absorption. Incorporation of esters of **sitostanol** into margarine provides an easy method of administration.[ff] Supplemental vitamin E may also be of value.[q] Another effective approach is to decrease the rate of cholesterol synthesis by administration of drugs that inhibit the synthesis of cholesterol. Inhibitors of HMG-CoA reductase,[gg,hh] (e.g., vaLostatin) isopentenyl-*PP* isomerase, squalene synthase (e.g.,

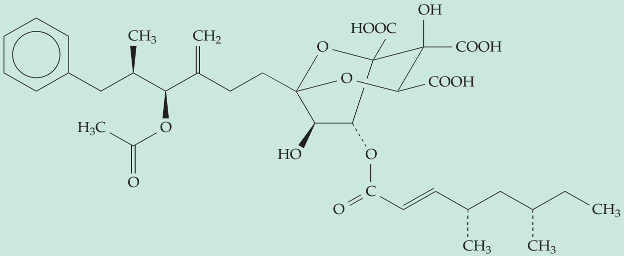
BOX 22-B ATHEROSCLEROSIS (continued)

squalestatin S1),[ii] and other enzymes in the biosynthetic pathway are targets for drug treatment.



Lovastatin
(Mevinolin)

Questions of possible long-term toxicity remain. Since 1976 there has been a greater than 25% decrease in the incidence of ischemic heart disease in the United States.[jj] Increased exercise, a decreased severity of influenza epidemics, and fluoridation of water[kk] have been suggested as explanations.



Squalestatin S1

---

[a] Ross, R. (1993) *Nature (London)* **362**, 801–809
[b] Krieger, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4077–4080
[c] Yagi, K. (1986) *Trends Biochem. Sci.* **11**, 18–19
[d] Breslow, J. L. (1996) *Science* **272**, 685–688
[e] Goldstein, J. L., Hobbs, H. H., and Brown, M. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1981–2030, McGraw-Hill, New York
[f] Brown, M. S., and Goldstein, J. L. (1974) *Sci. Am.* **251**(Nov), 58–66
[g] Brown, M. S., Kovanen, P. T., and Goldstein, J. L. (1981) *Science* **212**, 628–635
[h] Brown, M. S., and Goldstein, J. L. (1986) *Science* **232**, 34–47
[i] Motulsky, A. G. (1986) *Science* **231**, 126–128
[j] Breslow, J. L. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8314–8318
[k] Bergeron, J., Frank, P. G., Scales, D., Meng, Q.-H., Castro, G., and Marcel, Y. L. (1995) *J. Biol. Chem.* **270**, 27429–27438
[l] Breslow, J. L. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2031–2052, McGraw-Hill, New York
[m] Ye, S., Eriksson, P., Hamsten, A., Kurkinen, M., Humphries, S. E., and Henney, A. M. (1996) *J. Biol. Chem.* **271**, 13055–13060
[n] Benditt, E. P., Barrett, T., and McDougall, J. K. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6386–6389
[o] Gura, T. (1998) *Science* **281**, 35–37
[p] Holvoet, P., and Collen, D. (1994) *FASEB J.* **8**, 1279–1284
[q] Steinberg, D. (1997) *J. Biol. Chem.* **272**, 20963–20966
[r] Chisolm, G. M., Ma, G., Irwin, K. C., Martin, L. L., Gunderson, K. G., Linberg, L. F., Morel, D. W., and DiCorleto, P. E. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11452–11456
[s] Lynch, S. M., and Frei, B. (1995) *J. Biol. Chem.* **270**, 5158–5163
[t] Leeuwenburgh, C., Rasmussen, J. E., Hsu, F. F., Mueller, D. M., Pennathur, S., and Heinecke, J. W. (1997) *J. Biol. Chem.* **272**, 3520–3526
[u] Itabe, H., Yamamoto, H., Suzuki, M., Kawai, Y., Nakagawa, Y., Suzuki, A., Imanaka, T., and Takano, T. (1996) *J. Biol. Chem.* **271**, 33208–33217
[v] Retsky, K. L., Freeman, M. W., and Frei, B. (1993) *J. Biol. Chem.* **268**, 1304–1309
[w] Hazen, S. L., Hsu, F. F., Duffin, K., and Heinecke, J. W. (1996) *J. Biol. Chem.* **271**, 23080–23088
[x] Mensink, R. P., and Katan, M. B. (1990) *N. Engl. J. Med.* **323**, 439–445
[y] Chen, Z., Crippen, K., Gulati, S., and Banerjee, R. (1994) *J. Biol. Chem.* **269**, 27193–27197
[z] Kokame, K., Kato, H., and Miyata, T. (1996) *J. Biol. Chem.* **271**, 29659–29665
[aa] Kolata, G. (1985) *Science* **227**, 40–41
[bb] Kern, F., Jr. (1991) *N. Engl. J. Med.* **324**, 896–899
[cc] Bonanome, A., and Grundy, S. M. (1988) *N. Engl. J. Med.* **318**, 1244–1248
[cd] Merkel, M., Velez-Carrasco, W., Hudgins, L. C., and Breslow, J. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13294 – 13299
[dd] Hayes, K. C., and Khosla, P. (1992) *FASEB J.* **6**, 2600–2607
[ee] Jenkins, D. J. A., Wolever, T. M. S., Rao, A. V., Hegele, R. A., Mitchell, S. J., Ransom, T. P. P., Boctor, D. L., Spadafora, P. J., Jenkins, A. L., Mehling, C., Relle, L. K., Connelly, P. W., Story, J. A., Furumoto, E. J., Corey, P., and Würsch, P. (1993) *N. Engl. J. Med.* **329**, 21–26
[ff] Pedersen, T. R. (1995) *N. Engl. J. Med.* **333**, 1350–1351
[gg] Grundy, S. M. (1988) *N. Engl. J. Med.* **319**, 24–33
[hh] Brown, M. S., and Goldstein, J. L. (1996) *Science* **272**, 629
[ii] Kelly, M. J., and Roberts, S. M. (1995) *Nature (London)* **373**, 192–193
[jj] Stallones, R. A. (1980) *Sci. Am.* **243**(Nov), 53–59
[kk] Taves, D. R. (1978) *Nature (London)* **272**, 361–362

The LDL and related receptors. The LDL receptor gene extends over 50 kb of DNA and appears to be a mosaic of exons shared by several other genes that seem to have nothing to do with cholesterol metabolism.[201,211] The 839-residue receptor protein consists of five structural domains. The N-terminal domain that binds the LDL consists of seven repeated ~40-residue cysteine-rich modules.[212,213] This is followed by a large domain that is homologous to a precursor of the epidermal growth factor, a 48-residue domain containing many *O*-glycosylated serine and threonine residues, a 27-residue hydrophobic region that spans the membrane, and a 50-residue C-terminal cytoplasmic domain.[200,214–216] Synthesis of LDL receptors is regulated by a feedback mechanism, the cholesterol released within cells inhibiting the synthesis of new receptors.

As mentioned in Chapter 21, there are several related receptors with similar structures. Two of them have a specificity for apolipoprotein E and can accept remnants of VLDL particles and chylomicrons.[216–220] The LDL receptor-related protein is a longer-chain receptor.[216,221] LDL particles, especially when present in excess or when they contain oxidized lipoproteins, may be taken up by endocytosis into macrophages with the aid of the quite different **scavenger receptors**.[221–225] The uptake of oxidized lipoproteins by these receptors may be a major factor in promoting development of atherosclerosis (Box 22-B). On the other hand, **scavenger receptor SR-B1**, which is also present in liver cells, was recently identified as the receptor for HDL and essential to the "reverse cholesterol transport" that removes excess cholesterol for excretion in the bile.[213,213a]

Abnormalities of cholesterol metabolism. A variety of genetic problems have been identified, many of them being associated with atherosclerosis (Box 22-B).[218,226–230] In the commonest form of **familial hypercholesterolemia** a mutation in the LDL receptor protein prevents normal synthesis, binding, clustering into coated pits, or uptake of LDL and its cholesteryl esters. Over 600 mutations have been identified.[229,229a] Some of these are present in a Ca^{2+}-binding region of the 5th cysteine-rich module.[230] Other disorders that raise the plasma LDL level include a defective apoB-100 protein (see p. 1182)[230a] and deficiency of a protein that seems to be involved in incorportation of LDL receptors into clathrin-coated pits during endocytosis or in receptor recycling.[229a] In a **cholesteryl ester storage disease** the lysosomal lipase is lacking. Absence of lecithin:cholesterol acyltransferase from plasma causes corneal opacity and often kidney failure.[231]

In the very rare and fatal **Niemann–Pick C1** disease lysosomes in cells of the central nervous system and the viscera accumulate LDL-derived cholesterol. Study of the DNA of patients led to discovery of a 1278-residue integral membrane protein, which may be required for the Golgi-mediated transport of unesterified cholesterol from lysosomes to the ER.[189,232–234c]

Some people with elevated lipoprotein levels have VLDL that migrates on electrophoresis in the β band rather than the pre-β band (see Box 2-A). The presence of the β-VLDL is associated with a high incidence of artery disease,[218] which is most likely to develop in persons homozygous for a genetic variant of apolipoprotein E. The problem may arise because apo-E is required for receptor-mediated uptake of VLDL, which interacts both with tissue LDL receptors and with hepatic apo-E receptors. Genes for many of the apolipoproteins are polymorphic, and numerous alleles are present in a normal population.[218,235] In the rare **Tangier disease** apolipoprotein A-I is catabolized too rapidly, and the HDL level is depressed, resulting in accumulation of cholesterol esters in macrophages, Schwann cells, and smooth muscles. Orange-yellow enlarged tonsils are characteristic of the disease.[236] An ABC type transporter that allows cholesterol to leave cells is defective.[236a,b,c] The faulty component is known as the **cholesterol-efflux regulatory protein**. Another ABC transporter, apparently encoded by a pair of genes, which are expressed predominantly in liver and intestinal cells, prevents excessive accumulation of plant sterols such as sitosterol (Fig. 22-9).[236d,e]

Bile acids. Among the metabolites of cholesterol the bile acids (Fig. 22-10)[182,237,238] are quantitatively the most important (100–400 mg / day). These powerful emulsifying agents are formed in the liver and flow into the bile duct and the small intestine. A large fraction is later reabsorbed in the duodenum and is returned to the liver for reuse.[238a] Formation of the bile acids involves the removal of the double bond of cholesterol, inversion at C-3 to give a 3α-hydroxyl group, followed by hydroxylation and oxidation of the side chain.[238b–f] The principal human bile acids are **cholic acid** and **chenodeoxycholic acid** (Fig. 22-10). The free bile acids are then converted to CoA derivatives and conjugated with glycine and taurine to form **bile salts**, such as **glycocholic** and **taurocholic acids**.[238d] Several rare lipid-storage diseases are associated with defective bile acid formation.[239–241] In one of these, **cerebrotendinous xanthomatosis**, cholestanol is deposited both in tendons throughout the body and in the brain. Oxidation of the cholesterol side chain is incomplete with excretion, as glucuronides, of large amounts of bile alcohols (precursors to the bile acids). The synthesis of bile acids is regulated by feedback inhibition by the bile acids, but in this disease the inhibition is absent and the rates of both cholesterol biosynthesis and oxidation are increased. The problem is not one of storage of cholesterol but of the cholestanol that arises as a minor product of the pathway. A proper ratio of bile salt, phosphatidylcholine, and cholesterol in the bile is important to prevention of **cholesterol gallstones**.[207]

A variety of other oxidative modifications of cholesterol take place in tissues to give small amounts of diols.[242] Hydroperoxides of cholesterol may also be formed.[243] Some of the products are probably toxic, but others may be essential. One of these is 26-hydroxycholesterol, a minor component of plasma but a major neonatal excretion product.[244]

The body contains sulfate esters of cholesterol and other sterols,[245] sometimes in quite high concentrations relative to those of unesterified steroids. These esters are presumably soluble transport forms. They

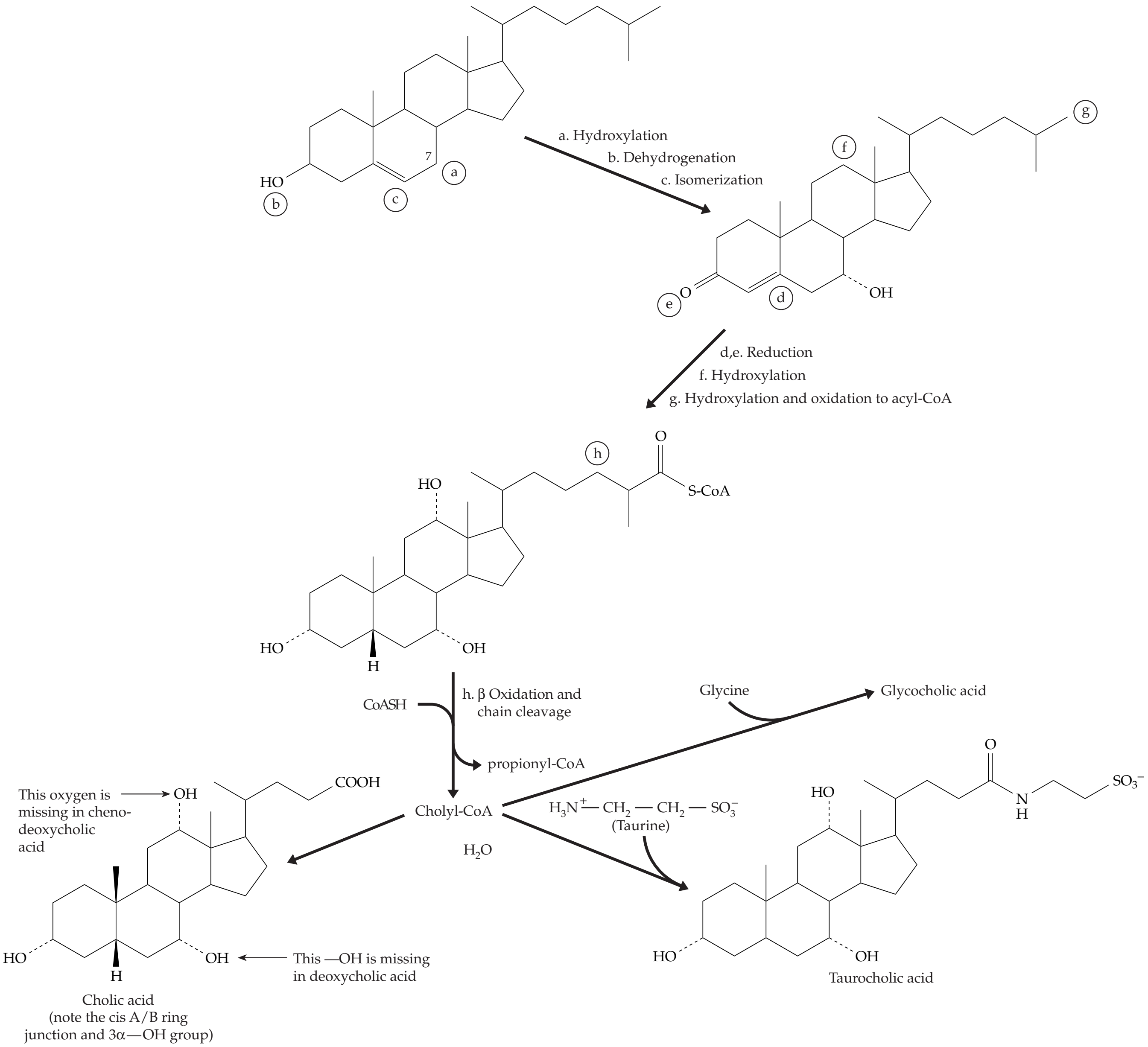

HO
b
c
7
a
a. Hydroxylation
b. Dehydrogenation
c. Isomerization
f
g
e
O
d
OH
d,e. Reduction
f. Hydroxylation
g. Hydroxylation and oxidation to acyl-CoA
h
O
S-CoA
HO
HO
H
OH
CoASH
h. β Oxidation and chain cleavage
propionyl-CoA
Glycine
Glycocholic acid
Cholyl-CoA
H₂O
H₃N⁺— CH₂ — CH₂ — SO₃⁻
(Taurine)
This oxygen is missing in cheno-deoxycholic acid
OH
COOH
HO
OH
H
This —OH is missing in deoxycholic acid
Cholic acid
(note the cis A/B ring junction and 3α—OH group)
O
N
H
SO₃⁻
HO
HO
OH
Taurocholic acid

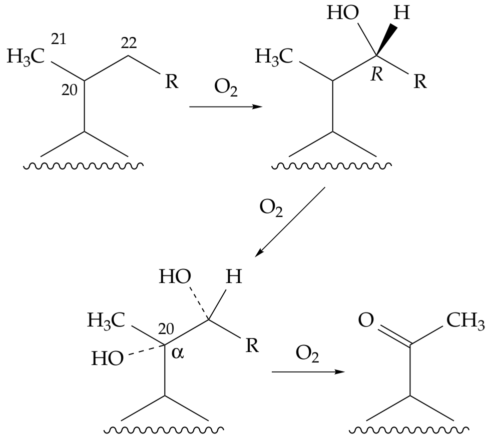
Figure 22-10 Formation of the bile acids.

are hydrolyzed by a steroid sulfatase present within cells and whose absence causes **X-linked ichthyosis**, one of the commoner forms of scaly skin. The defect may also lead to corneal opacity but appears not to cause any other problems.[246] Glucuronides of sterols are another group of water-soluble metabolites.[247] We tend to think of cholesterol as an inert structural component of cell membranes. However, it has been found in ester linkage to a signaling protein of the "hedgehog" family, important in embryo development.[248]

E. The Steroid Hormones

In the animal body three important groups of hormones are formed by the metabolism of cholesterol: the **progestins**, the **sex hormones**, and the **adrenal cortical hormones**.[249] Their synthesis occurs principally in mitochondria of the adrenal cortex and the gonads. Steroid hormone synthesis is regulated by hormones, such as **corticotropin** (ACTH), from the anterior pituitary[250] (see Chapter 30) and is also dependent upon the recently discovered **steroidogenic acute regulatory protein,** which in some way promotes the movement of stored cholesterol into mitochondrial membranes.[251,252] Some major pathways of biosynthesis are outlined in Fig. 22-11. The side chain is shortened to two carbon atoms through hydroxylation and oxidative cleavage to give the key intermediate **pregnenolone**. The reaction is initiated by the mitochondrial **cytochrome P450$_{ssc}$**, which receives electrons from NADPH and adrenodoxin (Chapter 16).[253] Hydroxylation occurs sequentially on C-22 and C-20 (Eq. 22-14). The chain cleavage is catalyzed by the same enzyme, an overall 6-electron oxidation occurring in three O_2-dependent steps. Dehydrogenation of the 3-OH group of pregnenolone to C=O is followed by a shift in the double bond, the oxosteroid isomerase reaction (Eq. 13-30, step *b*). In bacteria these two steps are catalyzed by different proteins, but a single human 3β-hydroxysteroid / Δ^5-Δ^4 isomerase catalyzes both reactions.[254,255] The product is the α,β-unsaturated ketone **progesterone**.

Most steroid hormones exist in part as sulfate esters and may also become esterified with fatty acids.[256] The fatty acid esters may have relatively long lives within tissues.[256] A special sex hormone-binding globulin transports sex hormones in the blood and regulates their access to target cells.[256a,b]



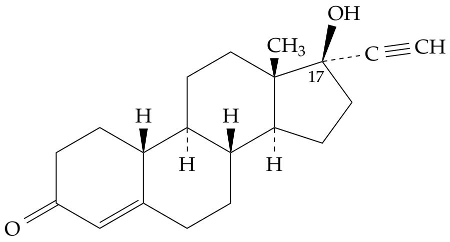
Pregnenolone

(22-14)

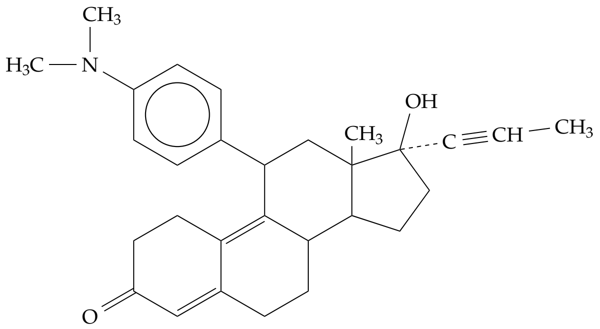
1. Progestins

Progesterone is the principal hormone of the **corpus luteum**, the endocrine gland that develops in the ovarian follicle after release of an ovum. Progesterone is also formed in the adrenals, testes, and placenta. It is metabolized rapidly, largely by reduction to alcohols, which may then be conjugated and excreted as glucuronides (see Eq. 20-16).[247] Reduction of the double bond within the A ring of progesterone leads to complete loss of activity, an indication that the α,β-unsaturated ketone group may play an essential role in the action of the hormone. Progesterone has a special role in the maintenance of pregnancy, and together with the estrogenic hormones it regulates the menstrual cycle. It is also essential for reproduction in lower animals such as birds and amphibians.[257,258]

The synthetic progesterone agonist **norethindrone** is widely used as one component of contraceptive pills. Having an opposite effect is **mifepristone** (also known as RU 486), a powerful antagonist of both progesterone and glucocorticoids. It is capable of inducing abortion and has other medical uses.[258–260] It is an effective emergency contraceptive agent that prevents implantation of an embryo.[261,262]



Norethindrone
(19-Nor-17α-ethynyltestosterone)



Mifepristone (RU 486)

2. Adrenal Cortical Steroids

Within the adrenal cortex (the outer portion of the adrenal glands) progesterone is converted into two groups of hormones of which **cortisol** and **aldosterone** are representative.[263] Two different cytochrome P450 hydroxylases, found in the ER and specific for C-21 and C-17α, respectively, together with a mitochondrial cytochrome P450 specific for C-11β (Eq. 18-55) participate in formation of cortisol.[264] Two of the same enzymes together with additional hydroxylases are required to form aldosterone.

Absence of the C-21 hydroxylase is one of the commonest of hereditary metabolic defects and is one of several enzymatic deficiencies that lead to **congenital adrenal hyperplasia**.[265–269] Cortisol, the synthesis of which is controlled by ACTH, is secreted by the adrenals in amounts of 15–30 mg daily in an adult. The hormone, which is essential to life, circulates in the blood, largely bound to the plasma protein **transcortin**. Cortisol, in turn, exerts feedback inhibition on ACTH production, and it is this feedback loop that fails when the C-21 hydroxylase is missing. Normally the circulating cortisol binds to receptors in both the pituitary and the hypothalamus of the brain to inhibit release of both ACTH and its hypothalamic releasing hormone (corticotropin-releasing hormone, CRF; see also Chapter 30). Girls are especially seriously affected by adrenal hyperplasia because, as ACTH production increases, the adrenal glands swell and produce an excessive amount of androgens. This occurs during the prenatal period of androgen release that initiates sexual differentiation. Girls with this deficiency are born with a masculine appearance of their external genitalia and continue to develop a masculine appearance. For reasons that are not clear the gene for the 21-hydroxylase is located within the HLA region (Chapter 31) of human chromosome 6.

Cortisol is a glucocorticoid which promotes gluconeogenesis and the accumulation of glycogen in the liver (Chapter 17). While it induces increased protein synthesis in the liver, it inhibits protein synthesis in muscle and many other tissues and leads to breakdown of fats to free fatty acids in adipose tissue.

Cortisol and its close relative **cortisone**, which was discovered by Kendall and Reichstein in the late 1940s, are probably best known for their anti-inflammatory effect in the body.[268,270] The effect depends upon several factors including inhibition of protein synthesis by fibroblasts, neutrophils, and antibody-forming cells. Migration of neutrophils into the inflamed area is also suppressed. Because of this action cortisone and synthetic analogs such as prednisolone and dexamethasone are among the modern "wonder drugs." They are used in controlling acute attacks of arthritis and of serious inflammations of the eyes and other organs. However, prolonged therapy can have serious side effects including decreased resistance to infections, wasting of muscle, and resorption of bone. The last results from a specific inhibition of calcium absorption from the gastrointestinal tract, glucocorticoids being antagonistic to the action of vitamin D (Box 22-C).

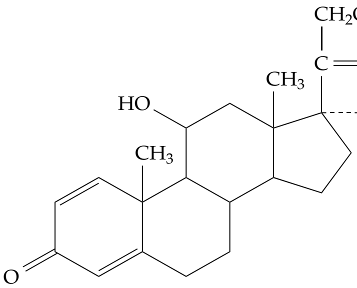
Aldosterone, which is classified as a **mineralocorticoid**, is produced under the control of the **renin–angiotensin** hormone system (Box 22-D), which is stimulated when sodium ion receptors in the kidneys detect an imbalance. It is synthesized in vascular cells of the body as well as in the adrenal cortex.[271] Aldosterone promotes the resorption of sodium ions in the kidney tubules and thus regulates water and electrolyte metabolism.[267,268] Glucocorticoids also have weak mineralocorticoid activity, and most patients with adrenocortical insufficiency (**Addison's disease**) can be maintained with glucocorticoids alone if their salt intake is adequate. Addison's disease develops when the adrenals are destroyed, most often by autoimmune disease or by tuberculosis.

3. Androgens

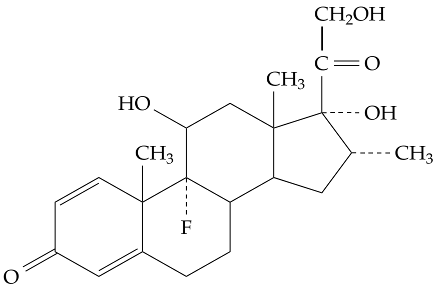
The principal **androgenic** or male sex hormone is **testosterone** formed from pregnenolone through removal of the side chain at C-17. The first step in the conversion is 17α-hydroxylation by a cytochrome P450 which may also mediate the further oxidative cleavage to **Δ^4-androstenedione** (Fig. 22-11). Reduction of the 17-carbonyl group forms testosterone. About 6–10 mg are produced daily in men, and smaller amounts (~0.4 mg) are synthesized in women. Testosterone is carried in the blood as a complex with a β-globulin and affects a variety of target tissues including the reproductive organs. Another striking effect is stimulation of the growth of the beard. Testosterone also causes premature death of follicles of head hair in genetically susceptible individuals. However, a bald man can usually grow a full beard, and follicles of the beard type, when transplanted to the head, remain immune to the action of androgen. No one knows what regulatory differences explain this fact. Baldness might be cured by use of suitable antagonists of the androgenic hormones, but the beard might fall out and sexual interest could be lost.

Androgen synthesis in the human male fetus begins at the age of about 70 days when the testes enlarge and go through an important period of activity that begins the conversion of the infant body to a male type. Other bursts of testosterone synthesis occur during infancy, but there is little further synthesis until the onset of puberty.[268,272]

Within many target tissues testosterone is converted by an NADPH-dependent 5α-reductase into **5α-dihydrotestosterone**. That this transformation is important is shown by the fact that absence of one of the two 5α-reductase isomers causes a form of **pseudohermaphroditism** in which male children are often mistaken as female and raised as girls. However, at puberty they become unmistakably male.[268,273–275] Many other metabolites of testosterone are known. These include the

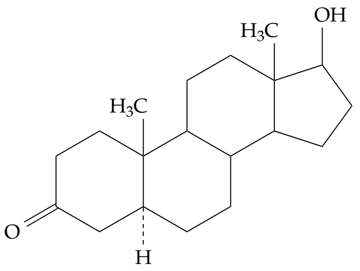


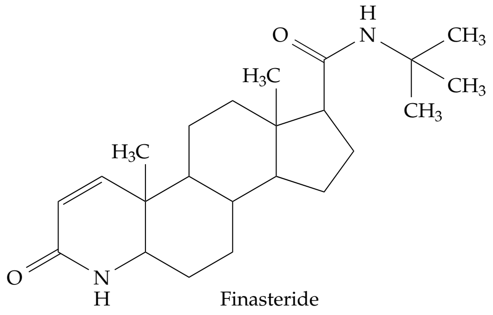
Prednisolone



Dexamethasone

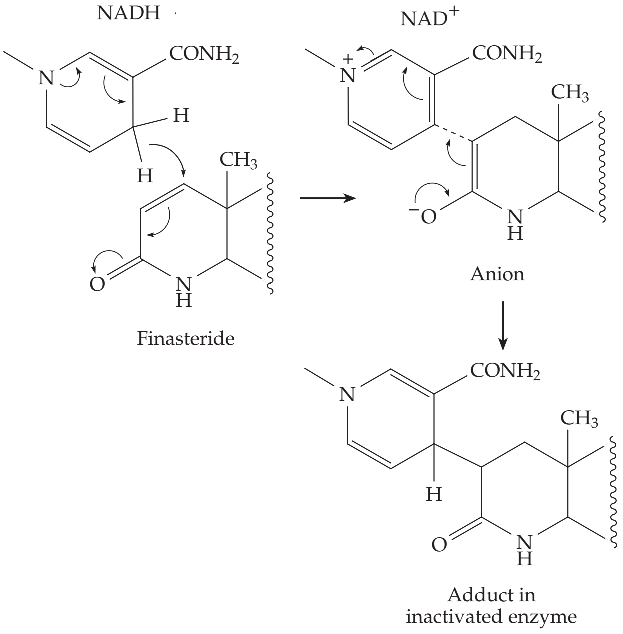
isomeric 5β-dihydrotestosterone and 5α-androstane-3β, 17β-diol which arises by reduction of the carbonyl group of 5α-dihydrotestosterone. Testosterone and dihydroxytestosterone have distinct roles in the body.


5α-Dihydrotestosterone


Finasteride

For example, testostosterone is required for sperm cell formation, voice deepening, and growth of pubic hair while dihydroxytestosterone stimulates development of the prostate gland and male pattern hair growth.[276]

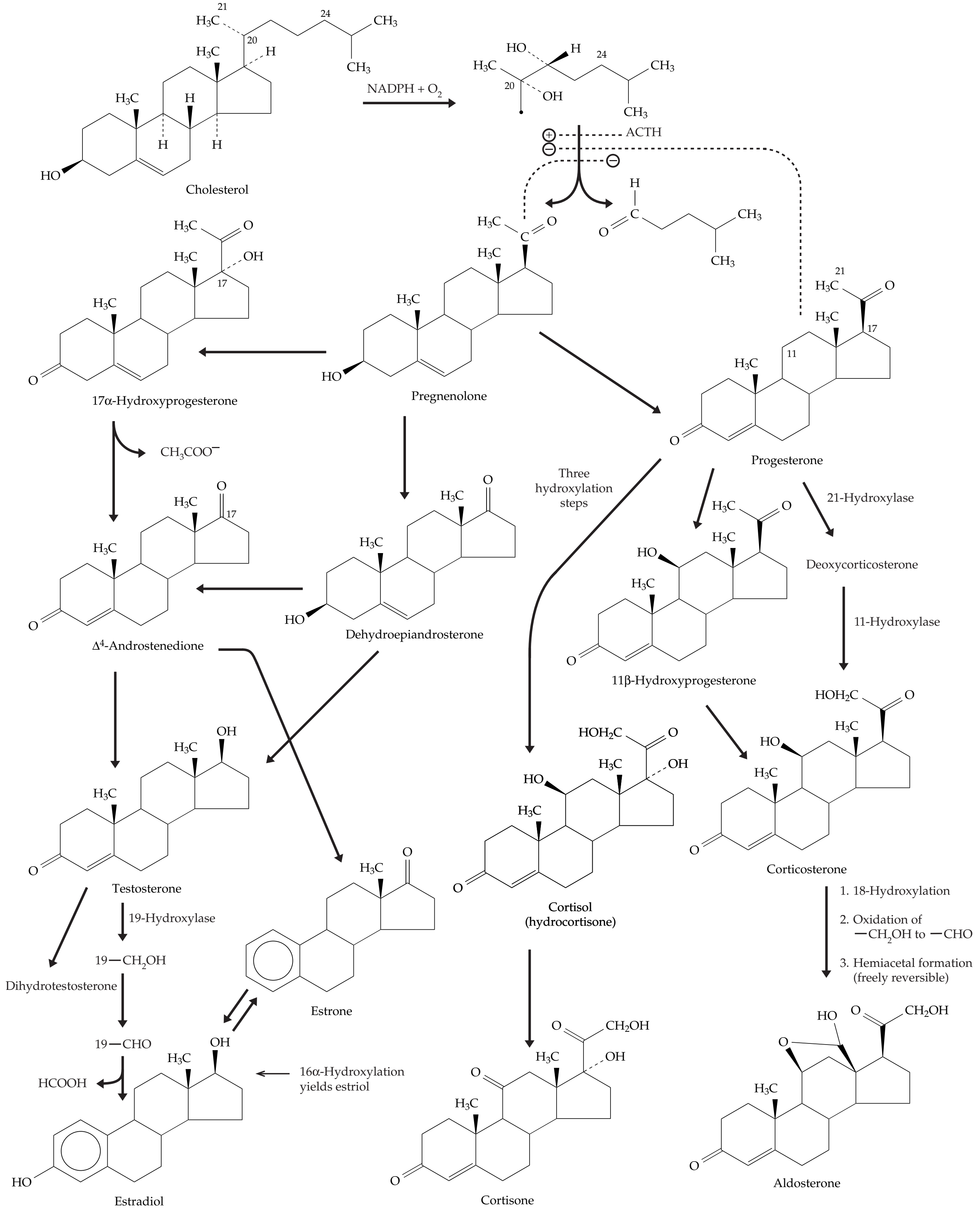
Males with deficiency of the 5α-reductase isoenzyme do not develop acne, male pattern baldness, or enlarged prostates.[274] The last fact was some of the impetus for development of the steroid 5α-reductase inhibitor **finasteride,** which is widely used to treat benign prostate enlargement.[274,277,278] It is an enzyme-activated inhibitor in which the NADH reduces the C=C bond in the A ring, which is not in the same position as in the substrate. The resulting anion cannot become protonated but instead adds to the NAD^+ as shown in Eq. 22-15.



(22-15)

A number of other androgens are present in the body. The adrenal glands make **dehydroepiandrosterone** (DHEA; Fig. 22–11), which circulates in human blood as its sulfate ester in higher concentration than that of any other androgen.[279–283] However, this steroid is absent from most species. DHEA can be taken up by tissues and converted to testosterone, estrogens, or other steroids (Fig. 22–11).[280] Recent attention has been focused on this hormone because it reaches a peak plasma concentration at age 20–25 years and by age 70 has fallen to 1/5 this value or less.[283] Should older men supplement their circulating DHEA by oral ingestion of 25–50 mg per day of DHEA sulfate? The hormone depresses blood cholesterol and lowers blood glucose in diabetic individuals.[280,282] It seems to promote increased energy metabolism.[281] It may fight obesity and atherosclerosis,[282] increase levels of estrogen and other steroid hormones in the brain,[284] and enhance memory and immune function.[285–285c] However, the hormone may be metabolized differently in different tissues, and its pathway of biosynthesis in the brain is uncertain.[279,286] Will men synthesize more testosterone from DHEA or make more estrogens in their adipose tissues? Young women should not take DHEA. It may increase the testosterone and dihydrotestosterone levels in the blood manyfold, and the women may become hirsute and masculinized.[280] However, most DHEA is converted to estrogen which may be of value to older women.[285a]

In addition to their role in sexual development androgens have a generalized "anabolic" effect causing increased protein synthesis, especially in muscles.[287–289] They promote bone growth, and the adolescent growth spurt in both males and females is believed to result from androgens. The greater height attained by men results in part from the higher concentration of androgen than is present in women. Many synthetic steroids have been made in an attempt to find "anabolic hormones" with little or no androgenic activity. The effort has been at least partially successful, and the use of anabolic hormones by athletes has become both widespread and controversial.

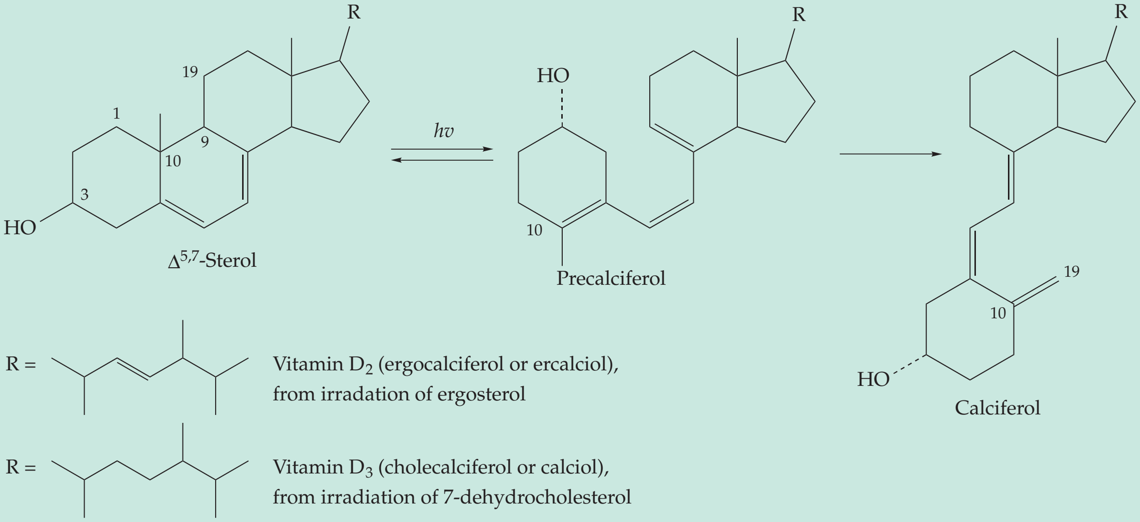


Cholesterol
$NADPH + O_2$
ACTH
Pregnenolone
17α-Hydroxyprogesterone
Progesterone
CH_3COO^-
Δ4-Androstenedione
Dehydroepiandrosterone
Three hydroxylation steps
11β-Hydroxyprogesterone
21-Hydroxylase
Deoxycorticosterone
11-Hydroxylase
Testosterone
Estrone
Cortisol (hydrocortisone)
Corticosterone
19-Hydroxylase
$19—CH_2OH$
Dihydrotestosterone
$19—CHO$
HCOOH
16α-Hydroxylation yields estriol
1. 18-Hydroxylation
2. Oxidation of $—CH_2OH$ to $—CHO$
3. Hemiacetal formation (freely reversible)
Estradiol
Cortisone
Aldosterone

Figure 22-11 Biosynthesis of some steroid hormones.

BOX 22-C VITAMIN D

A lack of vitamin D causes **rickets**, a disease of humans and other animals in which the bones are soft, deformed, and poorly calcified. Rickets was recognized by some persons to result from a dietary deficiency well over a hundred years ago, and the use of cod liver oil to prevent the disease was introduced in about 1870. By 1890 an association of rickets with a lack of sunlight had been made.



$\Delta^{5,7}$-Sterol

Precalciferol

Calciferol

Vitamin D_2 (ergocalciferol or ercalciol), from irradiation of ergosterol

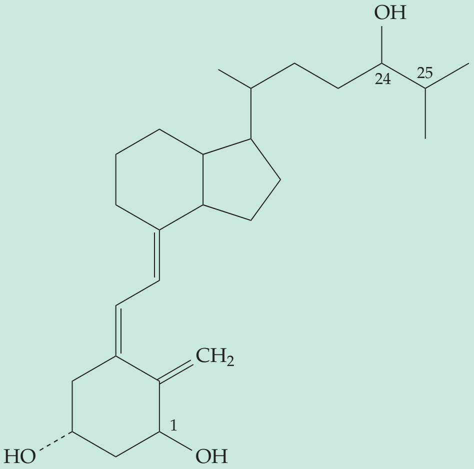
Vitamin D_3 (cholecalciferol or calciol), from irradiation of 7-dehydrocholesterol

However, it was not until 1924, when Steenbock and Hess showed that irradiation of certain foods generated protective activity against the disease, that vitamin D (**calciferol**) was recognized as a second lipid-soluble vitamin. Vitamin D is a family of compounds formed by the irradiation of $\Delta^{5,7}$-unsaturated sterols such as ergosterol and 7-dehydrocholesterol.[a–e] The former yields **ergocalciferol** (vitamin D_2) and the latter **cholecalciferol** (vitamin D_3).

At low temperature the intermediate **precalciferol** can be isolated. Irradiation sets up a photochemical steady-state equilibrium between the $\Delta^{5,7}$-sterol and the precalciferol. At higher temperatures the latter is converted to calciferol.[f] Other products, including toxic ones, are produced in slower photochemical side reactions. Therefore, the irradiation of ergosterol for food fortification must be done with care.

With normal exposure to sunlight enough 7-dehydrocholesterol is converted to cholecalciferol in the skin that it was concluded that no dietary vitamin D is required by most adults except during pregnancy. However, recently it has been recognized that old and sick adults probably need 400–600 I.U. per day to maintain calcium absorption and to prevent osteoporesis and fractures.[g,gh] It is usually recommended that children receive ~20 µg (400 I.U.) of ergocalciferol per day in their diet. Larger amounts are undesirable, and at a tenfold higher level vitamin D is seriously toxic.[h]

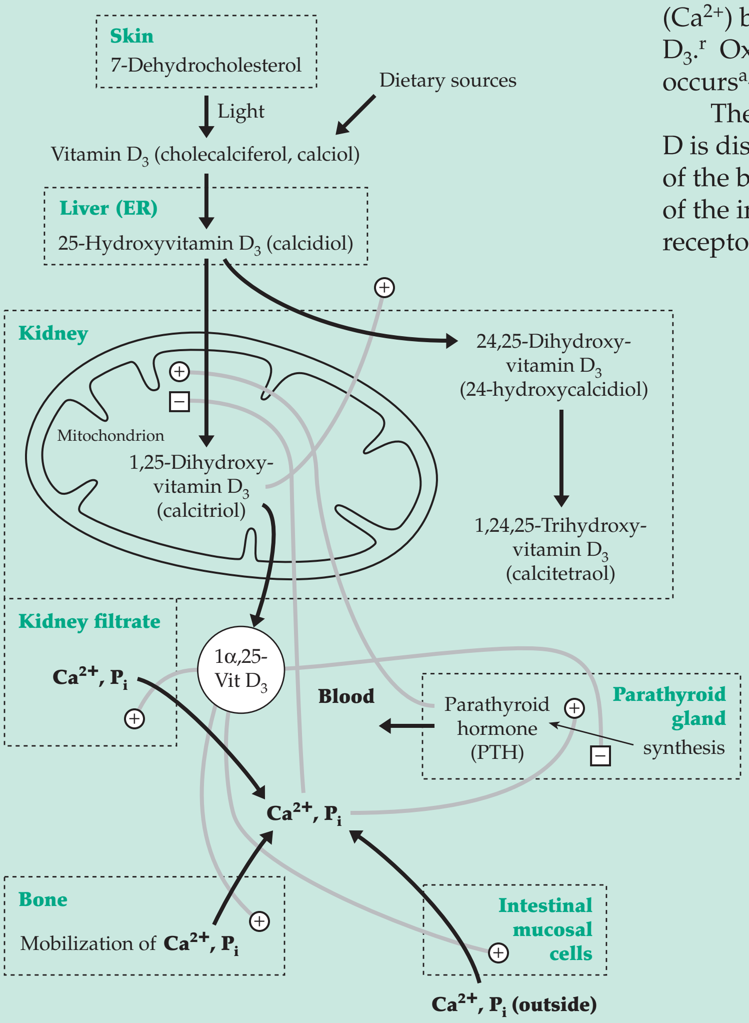
The principal function of vitamin D is in the control of calcium metabolism. This is accomplished through the mediation of polar, hydroxylated metabolites, the most important of which is **1α,25-dihydroxyvitamin D3** (1,25-dihydroxycholecalciferol or calcitriol).[a,i,j] This compound may be properly described as a steroid hormone and vitamin D itself as a hormone precursor formed in the skin. The major hydroxylation reactions of vitamin D are summarized in the accompanying scheme. The first hydroxylation to 25-hydroxyvitamin D_3 occurs largely in the liver,[jk] but the subse-



1 α,25-Dihydroxyvitamin D_3

BOX 22-C VITAMIN D (continued)

quent cytochrome P450 catalyzed 1α-hydroxylation takes place almost entirely in the kidneys.[k,l] Since it is 1,25-dihydroxy derivative that is essential for control of calcium ion metabolism, human patients with damaged kidneys often suffer severe demineralization of their bones (renal osteodystrophy). Administration of synthetic 1α,25-dihydroxyvitamin D_3 provides an effective treatment for these persons and also for children with an inherited defect in production of this hormone.[a,m] However, large doses, tested as an antileukemia drug, caused a severe hypercalcinemia limiting its use.[n]


Skin
7-Dehydrocholesterol
Light
Dietary sources
Vitamin D_3 (cholecalciferol, calciol)
Liver (ER)
25-Hydroxyvitamin D_3 (calcidiol)
⊕
Kidney
⊕
⊟
Mitochondrion
1,25-Dihydroxy-vitamin D_3 (calcitriol)
24,25-Dihydroxy-vitamin D_3 (24-hydroxycalcidiol)
1,24,25-Trihydroxy-vitamin D_3 (calcitetraol)
Kidney filtrate
Ca^{2+}, P_i
⊕
1α,25-Vit D_3
Blood
Parathyroid hormone (PTH)
⊕
⊟
Parathyroid gland
synthesis
Ca^{2+}, P_i
Bone
Mobilization of Ca^{2+}, P_i
⊕
Intestinal mucosal cells
⊕
Ca^{2+}, P_i (outside)

A second major vitamin D metabolite is 24*R*,25-dihydroxyvitamin D_3, a compound that circulates in the blood at a concentration 10 times higher than that of the 1,25-isomer.[a,b] However, no biological function has been discovered, and like a series of other polar metabolites (>30) it is probably on a pathway of inactivation and degradation of vitamin D. 1α,25-Dihydroxyvitamin D is also hydroxylated at C-24.[o] Additional hydroxylations occur at the 23- and 26-positions.[p,q] The 24-OH is often converted to an oxo group. Oxidation at C-26 together with cyclization yields 26,23 lactol and lactone species. The 25-hydroxy-26,23-lactone suppresses serum (Ca^{2+}) by competing with 1,25-dihydroxyvitamin D_3.[r] Oxidative cleavage of the side chains also occurs[a,s] as in the metabolism of cholesterol.

The hormonally active 1α,25-dihydroxyvitamin D is distributed through the bloodstream to all parts of the body. It is taken up rapidly by nuclei of cells of the intestinal lining where it binds to a 55-kDa receptor protein. In response, the cells synthesize **calbindins** (Chapter 6), Ca^{2+}-binding proteins which facilitate the uptake of calcium ions by the body.[t–v] (see also Fig. 6-7).

Other target organs for the action of 1,25-dihydroxyvitamin D include the kidneys, bone, muscle,[vw]and skin. The hormone promotes reabsorption of both Ca^{2+} and inorganic phosphate by kidney tubules. In bone it binds to a specific receptor where it promotes the mobilization of calcium ions. This effect may result in part from stimulation of calcium-activated ATPase of the outer membrane of bone cells. Dissolution of bone also requires the presence of **parathyroid hormone** (**PTH**), the 83-residue hormone secreted by the parathyroid gland. In women past the age of menopause and in elderly men the production of 1,25-dihydroxyvitamin D decreases.[w] This may be a cause of the serious bone loss (**osteoporosis**) frequently observed. Treatment with 1,25-dihydroxyvitamin D_3 or a synthetic analog seems to be helpful to such individuals.[x,xy] See also Chapter 30, Section A,5.

There is another important member of the Ca^{2+} homeostatic system. While vitamin D and PTH act together to increase the calcium level of the blood, **calcitonin**, a hormone of the thyroid gland, lowers the level of Ca^{2+} by promoting deposition of calcium in bone by the osteoblasts. The overall effect is to hold the concentration of Ca^{2+} in the blood at 2.2–2.6 mM in most animals with bones serving as a mobile reserve. Another role for vitamin D is suggested by the observation that 1,25-dihydroxyvitamin D_3 inhibits the growth promoting effect of interleukin-2 on mitogen-activated lymphocytes (Chapter 31).

BOX 22-C (continued)

Like other steroid hormones 1,25-dihydroxyvitamin D_3 acts to regulate gene transcription. It binds to a specific receptor, a member of the v-*erb*-A superfamily of ligand-activated transcription factors and a relative of the steroid receptor family.[a,j,y] Like other hormone receptors this vitamin D_3 receptor (VDR) has some rapid "nongenomic" actions such as causing an increase in intracellular [Ca^{2+}] as well as slower effects on transcription.[y–bb] Like the steroid receptors it is found in both cytoplasm and nucleus.[y] When occupied by 1,25-dihydroxyvitamin D, the receptor binds tightly to specific dihydroxyvitamin D_3 response elements in the DNA. These are found in promoter sequences for genes such as that of bone protein **osteocalcin**.[cc] In some promoters the response element binds heterodimers of VDR with another receptor, e.g., the retinoid X receptor (Table 22-1).[cc,dd] One effect of this response element is to activate the gene for the 24-hydroxylase involved in degradation of 1,25-dihydroxyvitamin D.[dd] Heterodimers of VDR with the thyroid hormone receptor activate the transcription of genes for calbindins D_{28K} and D_{9K}.[u] VDR also binds to repressor sequences, e.g., in the parathyroid gland.[ee] See also Chapter 28 and Table 22-1.

---

[a] DeLuca, H. F. (1988) *FASEB J.* **2**, 224–236

[b] Reichel, H., Koeffler, H. P., and Norman, A. W. (1989) *N. Engl. J. Med.* **320**, 980–991

[c] Lawson, D. E. M. (1978) *Vitamin D*, Academic Press, New York

[d] Norman, A. W. (1979) *Vitamin D, the Calcium Homeostatic Steroid Hormone*, Academic Press, New York

[e] DeLuca, H. F. (1979) *Vitamin D – Metabolism and Function*, Springer, New York

[f] Holick, M. F., Tian, X. Q., and Allen, M. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3124–3126

[g] Utiger, R. D. (1998) *N. Engl. J. Med.* **338**, 828–829

[gh] Thomas, M. K., Lloyd-Jones, D. M., Thadhani, R. I., Shaw, A. C., Deraska, D. J., Kitch, B. T., Vamvakas, E. C., Dick, I. M., Prince, R. L., and Finkelstein, J. S. (1998) *N. Engl. J. Med.* **338**, 777–783

[h] Jacobus, C. H., Holick, M. F., Shao, Q., Chen, T. C., Holm, I. A., Kolodny, J. M., Fuleihan, G. E.-H., and Seely, E. W. (1992) *N. Engl. J. Med.* **326**, 1173–1177

[i] Holick, M. F., Schnoes, T. H. K., DeLuca, H. F., Suda, T., and Cousins, R. J. (1971) *Biochemistry* **10**, 2799–2804

[j] Minghetti, P. P., and Norman, A. W. (1988) *FASEB J.* **2**, 3043–3053

[jk] Hosseinpour, F., and Wikvall, K. (2000) *J. Biol. Chem.* **275**, 34650–34655

[k] Takeyama, K.-i, Kitanaka, S., Sato, T., Kobori, M., Yanagisawa, J., and Kato, S. (1997) *Science* **277**, 1827–1830

[l] Menaa, C., Vrtovsnik, F., Friedlander, G., Corvol, M., and Garabédian, M. (1995) *J. Biol. Chem.* **270**, 25461–25467

[m] Hughes, M. R., Malloy, P. J., Kieback, D. G., Kesterson, R. A., Pike, J. W., Feldman, D., and O'Malley, B. W. (1988) *Science* **242**, 1702–1705

[n] Lee, N. E., Reddy, G. S., Brown, A. J., and Williard, P. G. (1997) *Biochemistry* **36**, 9429–9437

[o] Miyamoto, Y., Shinki, T., Yamamoto, K., Ohyama, Y., Iwasaki, H., Hosotani, R., Kasama, T., Takayama, H., Yamada, S., and Suda, T. (1997) *J. Biol. Chem.* **272**, 14115–14119

[p] Mayer, E., Bishop, J. E., Chandraratna, R. A. S., Okamura, W. H., Kruse, J. R., Popjak, G., Ohnuma, N., and Norman, A. W. (1983) *J. Biol. Chem.* **258**, 13458–13465

[q] Koszewski, N. J., Reinhardt, T. A., Napoli, J. L., Beitz, D. C., and Horst, R. L. (1988) *Biochemistry* **27**, 5785–5790

[r] Yamada, S., Nakayama, K., Takayama, H., Shinki, T., Takasaki, Y., and Suda, T. (1984) *J. Biol. Chem.* **259**, 884–889

[s] Jones, G., Kano, K., Yamada, S., Furusawa, T., Takayama, H., and Suda, T. (1984) *Biochemistry* **23**, 3749–3754

[t] Szebenyi, D. M. E., and Moffat, K. (1986) *J. Biol. Chem.* **261**, 8761–8777

[u] Schräder, M., Müller, K. M., Nayeri, S., Kahlen, J.-P., and Carlberg, C. (1994) *Nature (London)* **370**, 382–386

[v] Heizmann, C. W., and Hunziker, W. (1991) *Trends Biochem. Sci.* **16**, 98–103

[vw] Morelli, S., Buitrago, C., Vazquez, G., De Boland, A. R., and Boland, R. (2000) *J. Biol. Chem.* **275**, 36021–36028

[w] Slovik, D. M., Adams, J. S., Neer, R. M., Holick, M. F., and Potts, J. T., Jr. (1981) *N. Engl. J. Med.* **305**, 372–374

[x] Tilyard, M. W., Spears, G. F. S., Thomson, J., and Dovey, S. (1992) *N. Engl. J. Med.* **326**, 357–362

[xy] Shevde, N. K., Plum, L. A., Clagett-Dame, M., Yamamoto, H., Pike, J. W., and Deluca, H. F. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13487–13491

[y] Barsony, J., Renyi, I., and McKoy, W. (1997) *J. Biol. Chem.* **272**, 5774–5782

[z] Yukihiro, S., Posner, G. H., and Guggino, S. E. (1994) *J. Biol. Chem.* **269**, 23889–23893

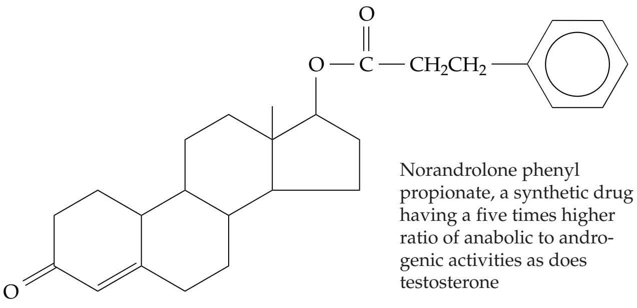
[aa] Lissoos, T. W., Beno, D. W. A., and Davis, B. H. (1993) *J. Biol. Chem.* **268**, 25132–25138

[bb] de Boland, A. R., Morelli, S., and Boland, R. (1994) *J. Biol. Chem.* **269**, 8675–8679

[cc] Schräder, M., Müller, K. M., and Carlberg, C. (1994) *J. Biol. Chem.* **269**, 5501–5504

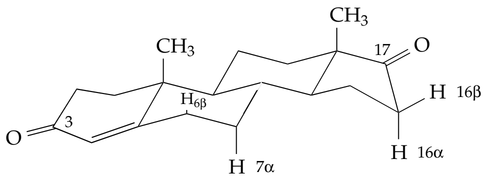
[dd] Allegretto, E. A., Shevde, N., Zou, A., Howell, S. R., Boehm, M. F., Hollis, B. W., and Pike, J. W. (1995) *J. Biol. Chem.* **270**, 23906–23909

[ee] Kremer, R., Sebag, M., Champigny, C., Meerovitch, K., Hendy, G. N., White, J., and Goltzman, D. (1996) *J. Biol. Chem.* **271**, 16310–16316



Norandrolone phenyl propionate, a synthetic drug having a five times higher ratio of anabolic to androgenic activities as does testosterone

Several sex-dependent differences have been observed in the action of cytochrome P450 isoenzymes on steroid hormones.[290,290a] Thus, androstenedione is hydroxylated by rat liver enzymes specific for the 6β, 7α, 16α, and 16β positions.[291] Of these the 16 hydroxylase is synthesized only in males, and synthesis of the 6 hydroxylase is also largely suppressed in females.



Δ4-Androstenedione

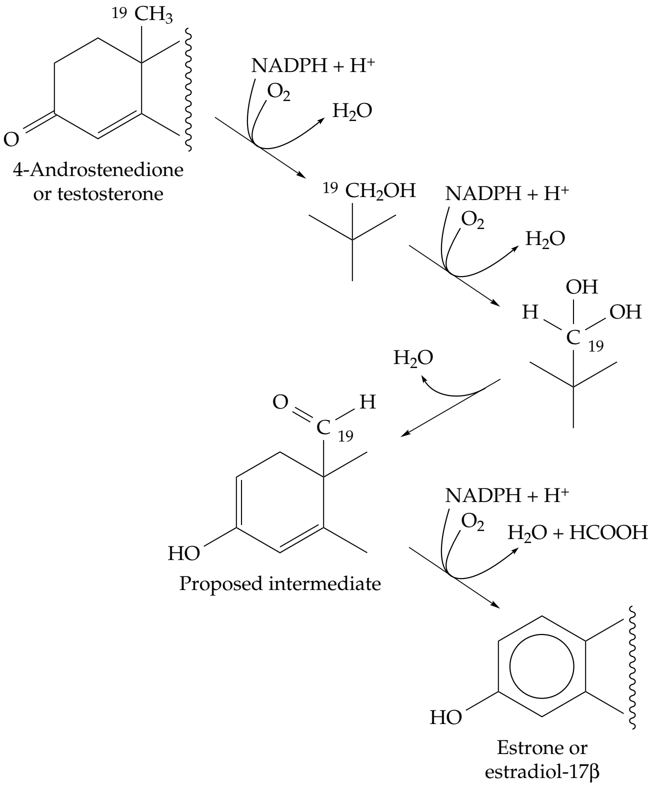
A female-specific 15β hydroxylase acts on steroid sulfates such as corticosterone sulfate and forms the major urinary excretion product of that hormone in female rats.[292] These sex-specific differences in enzymes are thought to be related to secretions of growth hormone that are in turn controlled by the "programming" of the hypothalamus by androgen during the neonatal period in rats[272] or during human fetal development.

4. Estrogens

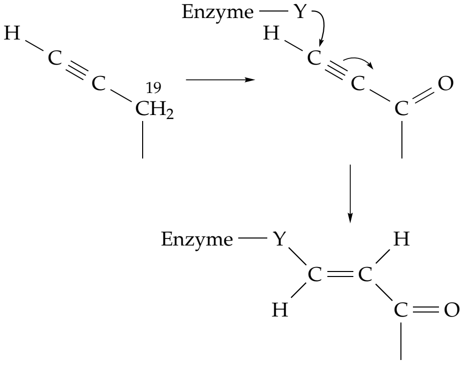
The principal **estrogenic** or female hormone is estradiol-17β. It is formed by oxidative removal of C-19 of testosterone followed by aromatization of the A ring.[268,293–295] All of the estrogenic hormones have this aromatic ring. Its formation involves three steps of hydroxylation followed by elimination of formate (Eq. 22-16). This **aromatase** appears to be a unique cytochrome P450, which catalyzes all of the steps of Eq. 22-16. It accepts electrons from NADPH via the flavoprotein NADH-cytochrome P450 reductase, which serves as the intermediate electron carrier.[296] It probably acts by a mechanism related to that illustrated for lanosterol 14α-demethylase on the right-hand side of Fig. 22-8.[294] This enzyme is the target for synthetic enzyme-activated inhibitors.[293] One of these is an androstenedione derivative with an acetylenic group attached to C-19. Passage through the first two steps of Eq. 22-16 generates a conjugated ketone to which a nucleophilic group of the enzyme can add irreversibly to inactivate the enzyme (Eq. 22-17). The C-17 acetylenic progesterone antagonist (norethyndrone) is also an enzyme activated inhibitor of the aromatase.

Estrogens are formed largely in the ovary and during pregnancy in the placenta. Estrogens are also synthesized in the testes, and the estrogen content of the horse testis is the highest of any endocrine organ. Target tissues for estrogens include the mammary glands, the uterus, and many other tissues throughout the body. Estrogens act on the growing ends of the long bones to stop growth and are therefore responsible, in part, for the shorter stature of females as compared to males. They are responsible for the overall higher fat content of the female body and for the smoother skin of the female. Recent attention has been focused on the effects of estrogens on brain neurons. Estrogens stimulate sprouting of axons and dendrites in neurons in cell cultures, and there is preliminary evidence that the hormone improves memory in healthy women and in those with Alzheimer disease.[285]

The cooperative action of progesterone and estradiol regulate the menstrual cycle. At the beginning of the cycle the levels of both estrogen and progesterone are low. Estrogen synthesis increases as a result of release of **follicle-stimulating hormone** (FSH) from the anterior pituitary. This hormone stimulates growth of the graafian follicles of the ovary which in turn produce estrogen. At about the midpoint of the cycle, as a result of the action of the pituitary **luteinizing hormone** (LH), an ovum is released and progesterone secretion begins. The latter is essential to maintenance of pregnancy. If a blastocyst is not implanted, hormone production decreases and menstruation occurs.



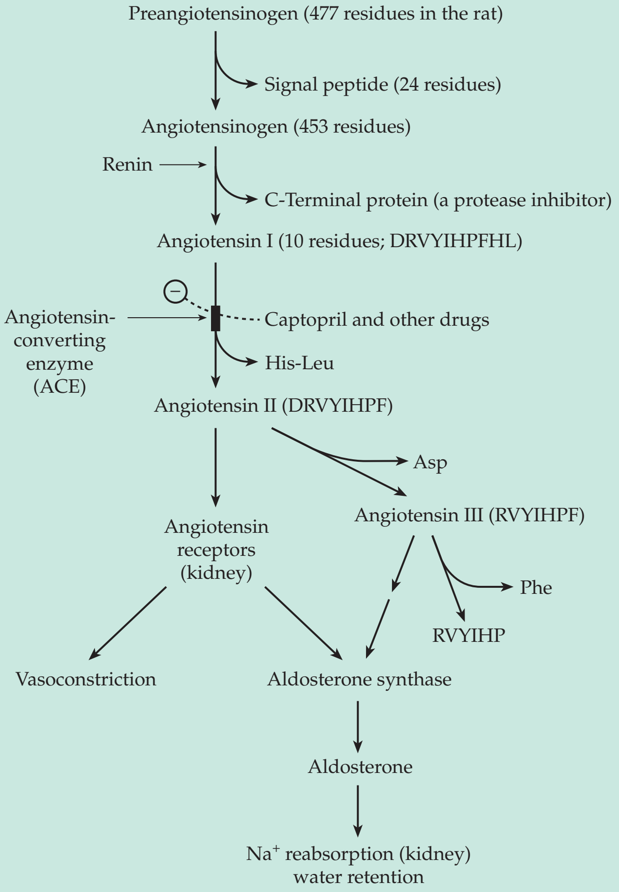
(22-16)



(22-17)

BOX 22-D THE RENIN-ANGIOTENSIN-ALDOSTERONE SYSTEM AND THE REGULATION OF BLOOD PRESSURE

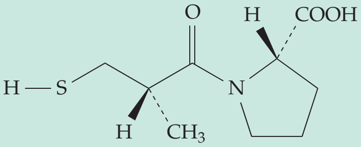
The mineralocorticoid **aldosterone** was isolated and identified in 1953. Although the function of adrenal cortical hormones in regulation of electrolytes had been known for many years, the special role of aldosterone had been overlooked.[a] Aldosterone works in concert with the aspartate protease **renin** and the octapeptide **angiotensin II** to regulate blood pressure. Angiotensin II, which is the most potent pressor substance known, is formed in the liver from the 477-residue (in the rat) **preangiotensinogen** as shown in the following cascade:


Preangiotensinogen (477 residues in the rat)
Signal peptide (24 residues)
Angiotensinogen (453 residues)
Renin
C-Terminal protein (a protease inhibitor)
Angiotensin I (10 residues; DRVYIHPFHL)
−
Angiotensin-converting enzyme (ACE)
Captopril and other drugs
His-Leu
Angiotensin II (DRVYIHPF)
Asp
Angiotensin III (RVYIHPF)
Angiotensin receptors (kidney)
Phe
RVYIHP
Vasoconstriction
Aldosterone synthase
Aldosterone
Na+ reabsorption (kidney) water retention

This **renin–angiotensin system** is a peptide hormone-generating system that operates in blood plasma rather than within tissues.[b,c]

Angiotensinogen, which is secreted by the liver and circulates in the blood, is converted to the physiologically inactive decapeptide angiotensin I by cleavage of a Leu-Leu peptide bond by the 328-residue renin.[d–f] Its precursor **preprorenin** is produced in the kidneys by the juxtaglomerular cells as well as in some other tissues and undergoes several steps of processing before the active enzyme is formed.[g–j] Active renin is released from the kidney cells into the bloodstream in response to various stimuli which include low arterial pressure resulting from constriction of the renal arteries or loss of blood.[k] Parathyroid hormone, glucagon, other adrenergic agonists, cAMP, some prostaglandins, low levels of aldosterone or Na$^+$, or high K$^+$ all induce secretion of renin. High blood pressure, α-adrenergic agonists, some prostaglandins, angiotensin, vasopressin, high Na$^+$, or low K$^+$ concentration decrease secretion of renin. It has been suggested that these diverse effects may be modified by a rise or fall in the Ca^{2+} concentration, high Ca^{2+} inhibiting secretion of renin, and low Ca^{2+} favoring secretion.

The only known physiological substrate for renin is angiotensinogen,[e] but it may also act on related proteins in the brain and other organs.[g] The inactive angiotensin I is converted to angiotensin II by the metal-containing carboxydipeptidase known as **angiotensin-converting enzyme**.[l,m] This enzyme is a target for drugs such as captopril, which is used to control high blood pressure (hypertension).[i] A zinc-dependent aminopeptidase may cut off the N-terminal aspartate to form angiotension III,[n] and degradation of angiotensins II and III can be initiated by removal of the C-terminal phenylalanine by a prolylcarboxypeptidase.[o]


O
H
COOH
H
S
N
H
CH3

Captopril

Angiotensin II has a variety of effects. By constricting blood vessels it raises blood pressure, and by stimulating thirst centers in the brain it increases blood volume. Both angiotensins II and III also act on the adrenal gland to promote the synthesis and release of aldosterone. Most of the effects of angiotension II are mediated by 359-residue seven-helix G-protein linked receptors which activate phospholipase C.[p,q,r] Like other steroid hormones aldosterone acts,via mineralocorticoid receptors, to control transcription of a certain set of proteins. The end effect is to increase the transport of Na$^+$ across the renal tubules and back into the blood. Thus, aldosterone acts to decrease the loss of Na$^+$ from the body. It promotes retention of water and raises

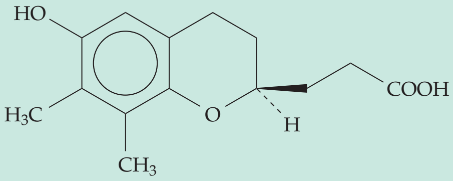
BOX 22-D THE RENIN–ANGIOTENSIN–ALDOSTERONE SYSTEM AND THE REGULATION OF BLOOD PRESSURE (continued)

blood pressure.[c,r,s] Its primary function is to provide adequate Na^+ to cells.[s] Dietary sodium appears to have little or no effect on blood pressure.[t]

The control of blood pressure is considerably more complex than it is described here. Another hormone system involving the peptide **bradykinin** and prostaglandins acts to lower blood pressure. Bradykin is also cleaved by the angiotensin-converting enzyme but is *inactivated* by the cleavage.[u,v] At least ten human genes have been shown to affect blood pressure.[c] One of these is the structural gene for angiotensin-converting enzyme, which has been linked to hypertension in both rats and humans.[w,x]

While several antagonists of angiotensin-converting enzyme are widely used to treat hypertension, they are not free of harmful side effects.[u,y,z] One alternative approach is to inhibit renin.[v,y]

While the angiotensins promote release of aldosterone, the **atrial natriuretic hormone**[r,aa–cc] inhibits release. This group of 21- to 33-residue polypeptides, secreted by cells of the atria (auricles) of the heart, also inhibits release of renin and promotes secretion of both Na^+ and water. Thus, they antagonize the action of aldosterone, which promotes Na^+ retention. However, there is uncertainty as to the significance of these peptides. The following metabolite of γ-tocopherol (Fig. 15-24) has been isolated from urine and is proposed as a new endogenous natriuretic factor.[dd]



2, 7, 8-trimethyl-2-(β-carboxyethyl)-6-hydroxychroman

---

[a] Tait, J. F., and Tait, S. A. S. (1978) *Trends Biochem. Sci.* **3**, N273–N275

[b] Inagami, T. (1994) in *Essays in Biochemistry*, Vol. 28 (Tipton, K. F., ed), pp. 147– , Portland Press, Chapel Hill, North Carolina

[c] Lifton, R. P. (1996) *Science* **272**, 676–680

[d] Kem, D. C., and Brown, R. D. (1990) *N. Engl. J. Med.* **323**, 1136–1137

[e] Wang, W., and Liang, T. C. (1994) *Biochemistry* **33**, 14636–14641

[f] Tong, L., Pav, S., Lamarre, D., Pilote, L., LaPlante, S., Anderson, P. C., and Jung, G. (1995) *J. Mol. Biol.* **250**, 211– 222

[g] Yanai, K., Saito, T., Kakinuma, Y., Kon, Y., Hirota, K., Taniguchi-Yanai, K., Nishijo, N., Shigematsu, Y., Horiguchi, H., Kasuya, Y., Sugiyama, F., Yagami, K.-i, Murakami, K., and Fukamizu, A. (2000) *J. Biol. Chem.* **275**, 5–8

[h] Smith, E. L., Hill, R. L., Lehman, I. R., Lefkowitz, R. J., Handler, P., and White, A. (1983) *Principles of Biochemistry, Mammalian Biochemistry*, 7th ed., McGraw-Hill, New York (pp. 157–163)

[i] Bull, H. G., Thornberry, N. A., Cordes, M. H. J., Patchett, A. A., and Cordes, E. H. (1985) *J. Biol. Chem.* **260**, 2952–2962

[j] Derkx, F. H. M., Schalekamp, M. P. A., and Schalekamp, M. A. D. H. (1987) *J. Biol. Chem.* **262**, 2472–2477

[k] Fray, J. C. S., Lush, D. J., and Valentine, A. N. D. (1983) *Fed. Proc.* **42**, 3150–3154

[l] Tipnis, S. R., Hooper, N. M., Hyde, R., Karran, E., Christie, G., and Turner, A. J. (2000) *J. Biol. Chem.* **275**, 33238–33243

[m] Ehlers, M. R. W., Schwager, S. L. U., Scholle, R. R., Manji, G. A., Brandt, W. F., and Riordan, J. F. (1996) *Biochemistry* **35**, 9549–9559

[n] Vazeux, G., Wang, J., Corvol, P., and Llorens-Cortès, C. (1996) *J. Biol. Chem.* **271**, 9069–9074

[o] Tan, F., Morris, P. W., Skidgel, R. A., and Erdös, E. G. (1993) *J. Biol. Chem.* **268**, 16631–16638

[p] Noda, K., Feng, Y.-H., Liu, X.-p, Saad, Y., Husain, A., and Karnik, S. S. (1996) *Biochemistry* **35**, 16435–16442

[q] Boucard, A. A., Wilkes, B. C., Laporte, S. A., Escher, E., Guillemette, G., and Leduc, R. (2000) *Biochemistry* **39**, 9662–9670

[qr] Heerding, J. N., Hines, J., Fluharty, S. J., and Yee, D. K. (2001) *Biochemistry* **40**, 8369–8377

[r] Flier, J. S., and Underhill, L. H. (1985) *N. Engl. J. Med.* **313**, 1330–1340

[s] Berger, S., Bleich, M., Schmid, W., Cole, T. J., Peters, J., Watanabe, H., Kriz, W., Warth, R., Greger, R., and Schütz, G. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9424–9429

[t] Taubes, G. (1998) *Science* **281**, 898–907

[u] Williams, G. H. (1988) *N. Engl. J. Med.* **319**, 1517– 1525

[v] Dealwis, C. G., Frazao, C., Badasso, M., Cooper, J. B., Tickle, I. J., Driessen, H., Blundell, T. L., Murakami, K., Miyazaki, H., Sueiras-Diaz, J., Jones, D. M., and Szelke, M. (1994) *J. Mol. Biol.* **236**, 342–360

[w] Hilbert, P., Lindpaintner, K., Beckmann, J. S., Serikawa, T., Soubrier, F., Dubay, C., Cartwright, P., De Gouyon, B., Julier, C., Takahasi, S., Vincent, M., Ganten, D., Georges, M., and Lathrop, G. M. (1991) *Nature (London)* **353**, 521–529

[x] Kreutz, R., Hübner, N., James, M. R., Bihoreau, M.-T., Gauguier, D., Lathrop, G. M., Ganten, D., and Lindpaintner, K. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8778–8782

[y] Tong, L., Pav, S., Lamarre, D., Simoneau, B., Lavallée, P., and Jung, G. (1995) *J. Biol. Chem.* **270**, 29520–29524

[z] Warren, J. B., and Loi, R. K. (1995) *FASEB J.* **9**, 411–418

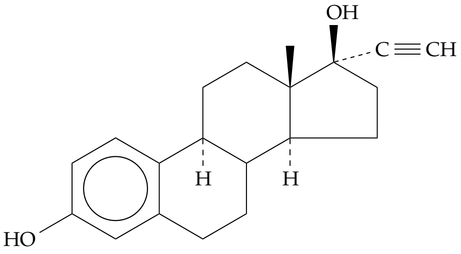
[aa] Sagnella, G. A., and MacGregor, G. A. (1986) *Trend Biochem. Sci.* II, 299–302

[bb] Cantin, M., and Genest, J. (1986) *Sci. Am.* **254**(February), 76–81

[cc] Lopez, M. J., Garbers, D. L., and Kuhn, M. (1997) *J. Biol. Chem.* **272**, 23064– 23068

[dd] Wechter, W. J., Kantoci, D., Murray, E. D., Jr., D'Amico, D. C., Jung, M. E., and Wang, W.-H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6002–6007

Administration of estrogens and progestins inhibits FSH and LH secretion from the pituitary (feedback inhibition) and hence ovulation. This effect is the action of contraceptive pills. A small amount of the synthetic estrogen 17-ethynylestradiol may be taken daily for 10–15 days followed by a combination of estrogen plus a progestin such as norethindrone for 0–15 days. Alternatively, a progestin alone may be

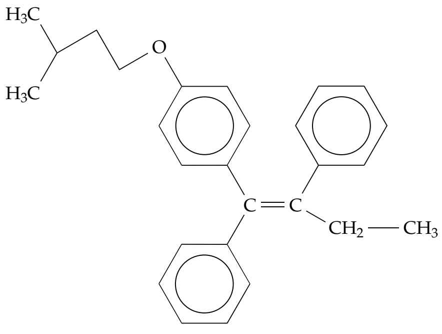


17-Ethynylestradiol

ingested over the entire period. Another synthetic compound with estrogenic activity is diethylstilbestrol. Its once widespread use in promoting growth of cattle and other animals has been discontinued because of carcinogenic action in rats fed large amounts of the compound.

Human cancers of the breast and endometrium are stimulated by estrogen. However, the mechanism is unknown.[297,298] It has usually been assumed that the proliferation of cells induced by estrogens leads to mutations and cancer. However, estrogens can form adducts with DNA after oxidation to 2- and 4-hydroxy derivatives and further oxidation to quinones.[297,298] In a similar manner prostate cancer is promoted by androgens. It has also been observed that in the United States the incidence of cancers of the ovary and endometrium has declined substantially during the past 35 years, perhaps as a result of the anti-estrogenic activity of the progestin in the widely used oral contraceptives.[299] This observation led to the idea of **hormonal chemoprevention**, the deliberate use of hormone antagonists to slow cancer formation.[299a] This may be especially attractive to persons carrying known cancer-susceptibility genes. The synthetic antiestrogen **tamoxifen** is being evaluated as a chemopreventive agent for breast cancer.[299,299b] However, a planned large-scale trial was postponed because of uncertainities about safety.[300]

In addition to steroids there are plant flavonoids (Box 21-E) that have estrogenic activity. These labile compounds are among the **"environmental estrogens."** In addition, there are many much more stable compounds, including the insecticides atrazine and DDT, PCBs, and phenolic softeners for plastics, that have weak estrogenic activity. Alarm has been sounded by some who maintain that these **xenoestrogens** are contributing to breast cancer, to reproductive difficulties in animals, and to low sperm counts in men.[301] It seems surprising that such small amounts of weakly estrogenic compounds could have such large effects. Some experiments suggest that two weak xenoestrogens may cooperate to give larger effects,[302,303] but this concept needs further evaluation.



trans-Tamoxifen

Although estrogens are usually regarded as female hormones, they are present in small amounts in male blood and in high concentration in semen. Male reproductive tissues contain estrogen receptors, and mice deficient in these receptors are sterile and their testes degenerate.[304] Furthermore, as in females estrogen stops growth of long bones in late puberty. A few men lacking estrogen or estrogen receptors have grown very tall (>2.1 m and still growing).[305]

5. The Steroid Receptor Family

The principal mode of action of steroid hormones is to stimulate transcription of specific sets of genes. The plasma concentrations of these hormones are low, typically ~10^{-9} M, but they have a high affinity for their protein receptors, some of which are located initially in the cytosol but are found largely in the nucleus. The earliest identification of steroid receptors was accomplished with radioactive ^{3}H-labeled progesterone, estrogens, and glucocorticoids.[306,307] Autoradiographs of thin sections of tissue made after the uptake of the hormones revealed that the radioactivity was concentrated in the nuclei. The protein receptors were then isolated and were found to bind both to the hormone and to specific sequences in DNA, the hormone **response elements**.[308] The progesterone, estrogen, and glucocorticoid receptors are multidomain proteins with two of the domains having highly conserved sequences and structures. One of these domains binds to DNA and the other to the steroid hormone. Their amino acid sequences are also related to those of the v-*erb-A* oncogene (Table 11-3).[309] With this knowledge available and through use of methods of "reverse genetics," receptors for the other steroid hormones and also for vitamin D, retinoic acid, and thyroid hormone were identified as members of the family as were several "orphan" receptors of as yet unknown function (Table 22-1). Invertebrates have similar receptors. At least eight are present in *Drosophila*,[307] and the family is present in the nematode *Caenorhabditis*.[310]

Because of their hydrophobic character the steroid hormones or other nonpolar ligands diffuse through membranes into cells. There they may encounter a variety of binding proteins that affect their access to a receptor.[307,311] Some receptors, including glucocorticoid receptors, are found in the cytosol. After a hor-

TABLE 22-1
Known Members of the Steroid Receptor Family[a]

Glucocorticoid[b–d]
Mineralocorticoid[e,f]
Progesterone[g,h]
Androgen[i,j,k]
Estrogen[l–o]
Vitamin D_3[p,q]
Thyroid α, β_1, β_2[r]
Retinoic acid α, β, γ[s]
Retinoid-X α,β[t–w]
Peroxisome proliferators[x,y]
Farnesoid X: bile acids[z,aa]
Orphan receptors, 8 or more

[a] See Fuller (1991) *FASEB J.* **5**, 3092–3099

[b] Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., and Sigler, P. B. (1991) *Nature (London)* **352**, 497–505

[c] Eriksson, M. A. L., Härd, T., and Nilsson, L. (1995) *Biophys. J.* **68**, 402–426

[d] La Baer, J., and Yamamoto, K. R. (1994) *J. Mol. Biol.* **239**, 664–688

[e] Funder, J. W. (1993) *Science* **259**, 1132–1133

[f] Geller, D. S., Farhi, A., Pinkerton, N., Fradley, M., Moritz, M., Spitzer, A., Meinke, G., Tsai, F. T. F., Sigler, P. B., and Lifton, R. P. (2000) *Science* **289**, 119–123

[g] Williams, S. P., and Sigler, P. B. (1998) *Nature (London)* **393**, 392–396

[h] Knotts, T. A., Orkiszewski, R. S., Cook, R. G., Edwards, D. P., and Weigel, N. L. (2001) *J. Biol. Chem.* **276**, 8475–8483

[i] Grossmann, M. E., Lindzey, J., Blok, L., Perry, J. E., Kumar, M. V., and Tindall, D. J. (1994) *Biochemistry* **33**, 14594–14600

[j] Matias, P. M., and 13 other authors. (2000) *J. Biol. Chem.* **275**, 26164–26171

[k] Sack, J. S., Kish, K. F., Wang, C., Attar, R. M., Kiefer, S. E., An, Y., Wu, G. Y., Scheffler, J. E., Salvati, M. E., Krystek, S. R., Jr., Weinmann, R., and Einspahr, H. M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4904–4909

[l] Ekena, K., Weis, K. E., Katzenellenbogen, J. A., and Katzenellenbogen, B. S. (1996) *J. Biol. Chem.* **271**, 20053–20059

[m] Tanenbaum, D. M., Wang, Y., Williams, S. P., and Sigler, P. B. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5998–6003

[n] Brzozowski, A. M., Pike, A. C. W., Dauter, Z., Hubbard, R. E., Bonn, T., Engström, O., Öhman, L., Greene, G. L., Gustafsson, J.-Å., and Carlquist, M. (1997) *Nature (London)* **389**, 753–758

[o] Kahlert, S., Nuedling, S., van Eickels, M., Vetter, H., Meyer, R., and Grohé, C. (2000) *J. Biol. Chem.* **275**, 18447–18453

[p] Craig, T. A., Veenstra, T. D., Naylor, S., Tomlinson, A. J., Johnson, K. L., Macura, S., Juranic, N., and Kumar, R. (1997) *Biochemistry* **36**, 10482–10491

[q] Hendick, M., and Carlberg, C. (2000) *J. Mol. Biol.* **304**, 793–801

[r] Kostrouch, Z., Kostrouchova, M., and Rall, J. E. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 156–159

[s] Klaholz, B. P., Mitschler, A., and Moras, D. (2000) *J. Mol. Biol.* **302**, 155–170

[t] Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995) *Nature (London)* **375**, 377–382

[u] Holmbeck, S. M. A., Foster, M. P., Casimiro, D. R., Sem, D. S., Dyson, J., and Wright, P. E. (1998) *J. Mol. Biol.* **281**, 271–284

[v] Kersten, S., Dong, D., Lee, W.-y, Reczek, P. R., and Noy, N. (1998) *J. Mol. Biol.* **284**, 21–32

[w] Chawla, A., Repa, J. J., Evans, R. M., and Mangelsdorf, D. J. (2001) *Science* **294**, 1866–1870

[x] Nolte, R. T., Wisely, G. B., Westin, S., Cobb, J. E., Lambert, M. H., Kurokawa, R., Rosenfeld, M. G., Willson, T. M., Glass, C. K., and Milburn, M. V. (1998) *Nature (London)* **395**, 137–143

[y] Rosen, E. D., and Spiegelman, B. M. (2001) *J. Biol. Chem.* **276**, 37731–37734

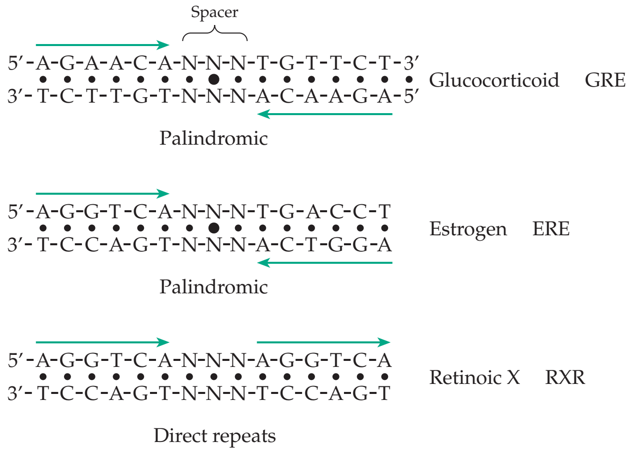
[z] Makishima, M., Okamoto, A. Y., Repa, J. J., Tu, H., Learned, R. M., Luk, A., Hull, M. V., Lustig, K. D., Mangelsdorf, D. J., and Shan, B. (1999) *Science* **284**, 1362–1365

[aa] Parks, D. J., Blanchard, S. G., Bledsoe, R. K., Chandra, G., Consler, T. G., Kliewer, S. A., Stimmel, J. B., Willson, T. M., Zavacki, A. M., Moore, D. D., and Lehmann, J. M. (1999) *Science* **284**, 1365–1368

mone binds to a cytosolic receptor, the complex apparently undergoes time and temperature dependent alterations that activate the receptor before it diffuses into the nucleus and binds to its proper response element in the DNA.[312] This process has been observed directly for glucocorticoid receptors labeled by fusion with green fluorescent protein (Box 23-A).[313]

The conserved 68-residue DNA binding domain of the glucocorticoid receptor contains two Zn^{2+} ions, each coordinated by four cysteine $-CH_2-S^-$ groups with tetrahedral geometry. These two consecutive motifs form structures somewhat similar to those of the "zinc fingers" shown in Fig. 5-38.[314–316] However, the overall folding pattern is different from those considered in Chapter 5. The two zinc-binding sites lie at the N termini of a pair of helices that cross at right angles near their centers. One of these is a DNA-recognition helix that fits into the major groove of DNA thereby allowing interaction of its amino acid side chains with the bases of the DNA response elements.

The response elements for glucocorticoids and estrogen receptors contain short palindromic sequences with various three-nucleotide "spacer" sequences in the center as follows.[308,314,316–318] Two receptor proteins bind to the palindromic DNA forming a homodimeric receptor pair. For the 9-*cis* retinoic acid receptor RXR-α the response element contains a pair of direct repeats of a 6-base consensus sequence with a two-base pair spacer:

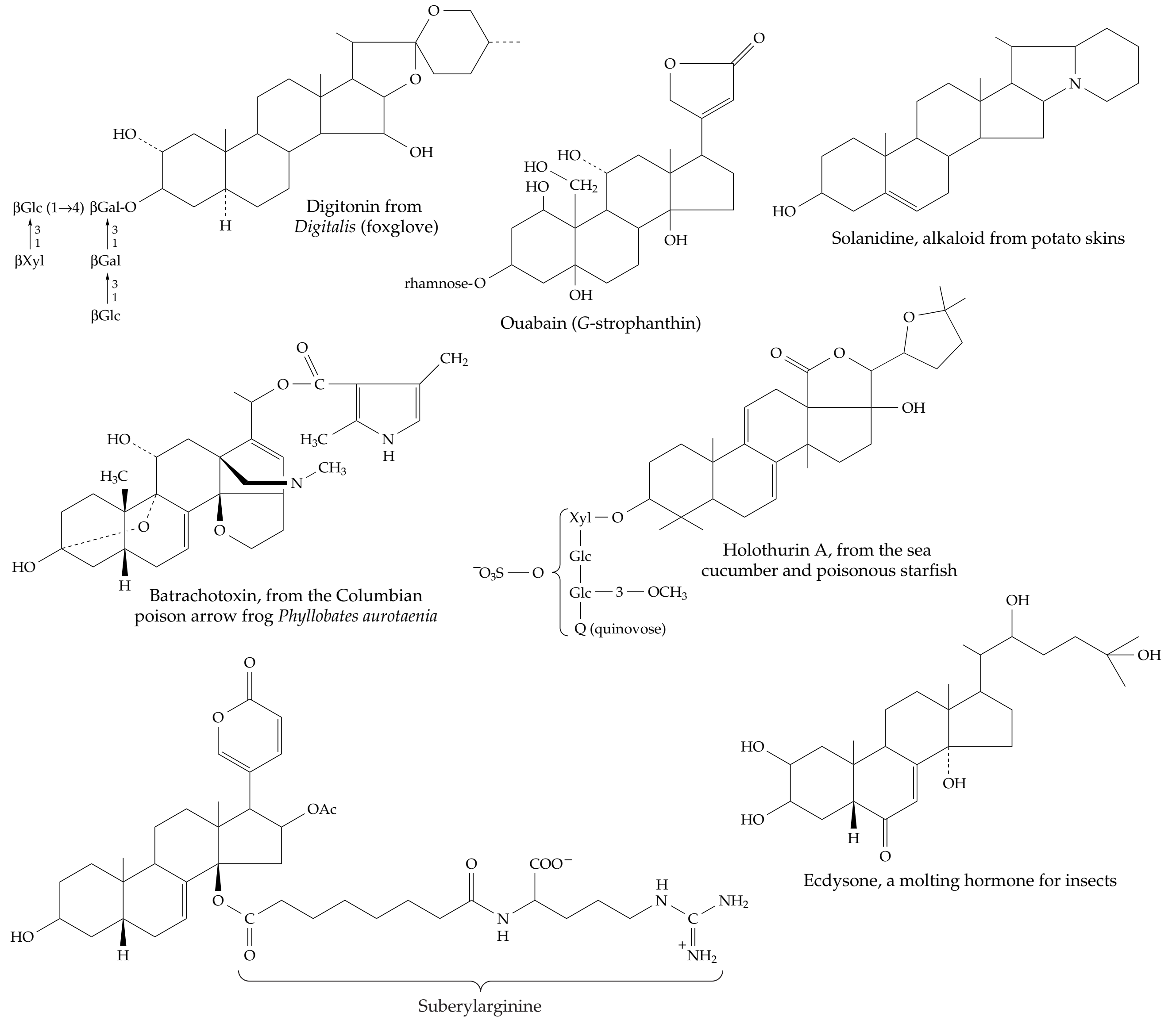

Spacer
5′-A-G-A-A-C-A-N-N-N-T-G-T-T-C-T-3′
3′-T-C-T-T-G-T-N-N-N-A-C-A-A-G-A-5′
Glucocorticoid GRE
Palindromic
5′-A-G-G-T-C-A-N-N-N-T-G-A-C-C-T
3′-T-C-C-A-G-T-N-N-N-A-C-T-G-G-A
Estrogen ERE
Palindromic
5′-A-G-G-T-C-A-N-N-N-A-G-G-T-C-A
3′-T-C-C-A-G-T-N-N-N-T-C-C-A-G-T
Retinoic X RXR
Direct repeats

The RXR-α receptor binds differently and tends to form *heterodimeric* pairs with other receptors.[319,320] All of these receptors undergo conformational changes when agonists or antagonists bind.[320,321] Estrogens can also bind to androgen receptors, perhaps in playing their essential role in male reproductive physiology.[322] There are more than one type of receptor for each group of steroid hormones, and these may interact differently with the various response elements in DNA[323] making the effects of hormones complex and hard to analyze. Interactions with additional proteins also affect the response of a cell to hormones.[324] Furthermore, steroid hormones have **transcription-independent effects**. For example, progesterone binds to oxytocin receptors[325] as well as to other steroid receptors, which affect a broad range of biochemical processes.[326]

F. Other Steroids

The **saponins** are a series of steroid glycosides with detergent properties that are widespread among higher plants.[327] Some are toxic, and among these toxic materials are compounds of extraordinary medical importance. Best known are the steroid glycosides of *Digitalis*, among them **digitonin** (Fig. 22-12). The particular arrangement of sugar units in this molecule imparts a specificity toward heart muscle. The compound is extremely toxic; in small amounts it acts to increase the tone of heart muscle and is widely used in treatment of congestive heart failure.[328] The maintenance dose is only 0.1 mg / day. Another toxic glycoside and heart stimulant is **ouabain** (Fig. 22-12). Ouabain is a specific inhibitor of the membrane-bound ($Na^+ + K^+$)-ATPase believed to be the ion pump that keeps intracellular K^+ concentrations high and Na^+ concentrations low (Chapter 8). Similar glycosides account for the extreme toxicity of the leaves of the oleander and the roots of the lily of the valley. A steroid glycoside from red squill is used as a rat poison. A number of alkaloids (nitrogenous bases) are derived from steroids. An example is **solanidine** (Fig. 22-12), which is present in the skins and sprouts of potatoes, making both quite toxic.

Some animals also contain toxic steroids. **Batrachotoxin** of the Columbian poison arrow frog (Fig. 22-12) is present in amounts of only 50 µg per frog.[329,330]


Digitonin from *Digitalis* (foxglove)

Ouabain (*G*-strophanthin)

Solanidine, alkaloid from potato skins

Batrachotoxin, from the Columbian poison arrow frog *Phyllobates aurotaenia*

Holothurin A, from the sea cucumber and poisonous starfish

Ecdysone, a molting hormone for insects

Suberylarginine

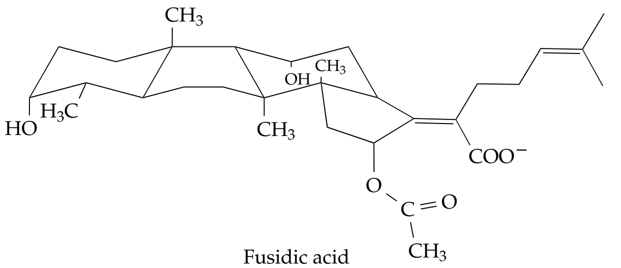
Bufotoxin, from the skin glands of toads

Figure 22-12 More steroids, mostly toxic.

The toxin acts on nerves to block transmissions to the muscle by increasing the permeability of membranes to sodium ions. It is specifically antagonized by **tetrodotoxin** (Fig. 30-16). Batrachotoxin alkaloids are present also in certain birds.[330a] Some echinoderms make powerful steroid toxins such as **holothurin A** (Fig. 22-12), a surface active agent that causes irreversible destruction of the excitability of neuromuscular tissues. The common toad produces in its skin steroid toxins such as **bufotoxin** (Fig. 22-12), which are sufficiently powerful to teach a dog to leave toads alone.

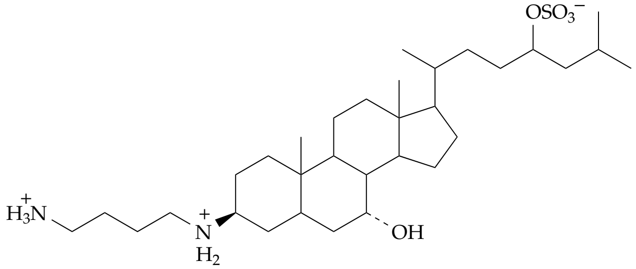
Ecdysone, a highly hydroxylated steroid (Fig. 22-12), is a molting hormone for insects.[331,332] Several molecules with ecdysone activity are known, and some of these are produced by certain plants. Although ecdysones are needed by insects for larval molting, they are toxic in excess. Perhaps plants protect themselves from insects by synthesizing these substances.

Among the antibiotics **fusidic acid** is a steroid. An inhibitor of chloramphenicol acetyltransferase[333] it is highly inhibitory to staphyococci but almost noninhibitory to *E. coli*. Note the boat conformation of the B ring.



Fusidic acid

Another steroid antibiotic, **squalamine**,[334] was isolated from the stomachs of sharks. It is effective against gram-positive and gram-negative bacteria, and some fungi as well.[335]



Squalamine

References

1. Luckner, M. (1984) *Secondary Metabolism in Plants and Animals*, 2nd ed., Academic Press, New York
2. Britton, G. (1976) in *Chemistry and Biochemistry of Plant Pigments*, 2nd ed., Vol. 1 (Goodwin, T. W., ed), pp. 262–327, Academic Press, New York
3. Haslam, E. (1985) *Metabolites and Metabolism*, Oxford Univ. Press, London
4. Herbert, R. B. (1989) *The Biosynthesis of Secondary Metabolites*, 2nd ed., Chapman and Hall, New York
5. Towers, G. H. N., and Stafford, H. A., eds. (1990) *Biochemistry of the Mevalonic Acid Pathway to Terpenoids*, Vol. 24, Plenum Press, New York and London
6. Nes, W. D., ed. (1994) *Isopentenoids and Other Natural Products*, American Chemical Society, Washington D.C.

6a. Wang, K., and Ohnuma, S.-i. (1999) *Trends Biochem. Sci.* **24**, 445–451

7. Misra, I., Narasimhan, C., and Miziorko, H. M. (1993) *J. Biol. Chem.* **268**, 12129–12135
8. Quant, P. A. (1994) in *Essays in Biochemistry*, Vol. 28 (Tipton, K. F., ed), Portland Press, Chapel Hill, North Carolina
9. Denbow, C. J., Lång, S., and Cramer, C. L. (1996) *J. Biol. Chem.* **271**, 9710–9715
10. Street, I. P., Coffman, H. R., Baker, J. A., and Poulter, C. D. (1994) *Biochemistry* **33**, 4212–4217
11. Frimpong, K., and Rodwell, V. W. (1994) *J. Biol. Chem.* **269**, 11478–11483
12. Bischoff, K. M., and Rodwell, V. W. (1997) *Protein Sci.* **6**, 156–161
13. Misra, I., and Miziorko, H. M. (1996) *Biochemistry* **35**, 9610–9616
14. Omkumar, R. V., and Rodwell, V. W. (1994) *J. Biol. Chem.* **269**, 16862–16866
15. Lawrence, C. M., Rodwell, V. W., and Stauffacher, C. V. (1995) *Science* **268**, 1758–1762
16. Luskey, K. L., and Stevens, B. (1985) *J. Biol. Chem.* **260**, 10271–10277
17. Goldstein, J. L., and Brown, M. S. (1990) *Nature (London)* **343**, 425–430
18. Meigs, T. E., Roseman, D. S., and Simoni, R. D. (1996) *J. Biol. Chem.* **271**, 7916–7922
19. Berkhout, T. A., Simon, H. M., Patel, D. D., Bentzen, C., Niesor, E., Jackson, B., and Suckling, K. E. (1996) *J. Biol. Chem.* **271**, 14376–14382
20. Correll, C. C., Ng, L., and Edwards, P. A. (1994) *J. Biol. Chem.* **269**, 17390–17393
21. Biardi, L., Sreedhar, A., Zokaei, A., Vartak, N. B., Bozeat, R. L., Shackelford, J. E., Keller, G.-A., and Krisans, S. K. (1994) *J. Biol. Chem.* **269**, 1197–1205
22. Potter, D., Wojnar, J. M., Narasimhan, C., and Miziorko, H. M. (1997) *J. Biol. Chem.* **272**, 5741–5746
23. Schafer, B. L., Bishop, R. W., Kratunis, V. J., Kalinowski, S. S., Mosley, S. T., Gibson, K. M., and Tanaka, R. D. (1992) *J. Biol. Chem.* **267**, 13229–13238
24. Paton, V. G., Shackelford, J. E., and Krisans, S. K. (1997) *J. Biol. Chem.* **272**, 18945–18950
25. Dhe-Paganon, S., Magrath, J., and Abeles, R. H. (1994) *Biochemistry* **33**, 13355 – 13362
26. Eisenreich, W., Menhard, B., Hylands, P. J., Zenk, M. H., and Bacher, A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6431–6436
27. Kuo, S.-M., and Aronson, P. S. (1996) *J. Biol. Chem.* **271**, 15491–15497
28. Rohmer, M., Seemann, M., Horbach, S., Bringer-Meyer, S., and Sahm, H. (1996) *J. Am. Chem. Soc.* **118**, 2564–2566
29. Sprenger, G. A., Schörken, U., Wiegert, T., Grolle, S., de Graaf, A. A., Taylor, S. V., Begley, T. P., Bringer-Meyer, S., and Sahm, H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12857–12862
30. Himmeldirk, K., Sayer, B. G., and Spenser, I. D. (1998) *J. Am. Chem. Soc.* **120**, 3581–3589
31. Taylor, S. V., Kelleher, N. L., Kinsland, C., Chiu, H.-J., Costello, C. A., Backstrom, A. D., McLafferty, F. W., and Begley, T. P. (1998) *J. Biol. Chem.* **273**, 16555–16560

31a. Rohdich, F., Hecht, S., Gärtner, K., Adam, P., Krieger, C., Amslinger, S., Arigoni, D., Bacher, A., and Eisenreich, W. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1158–1163

32. Arigoni, D., Sagner, S., Latzel, C., Eisenreich, W., Bacher, A., and Zenk, M. H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10600–10605

32a. Cane, D. E. (2000) *Science* **287**, 818–819

32b. Arigoni, D., Eisenreich, W., Latzel, C., Sagner, S., Radykewicz, T., Zenk, M. H., and Bacher, A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1309–1314

32c. Rohdich, F., Wungsintaweekul, J., Fellermeier, M., Sagner, S., Herz, S., Kis, K., Eisenreich, W., Bacher, A., and Zenk, M. H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11758–11763

32d. Herz, S., and 10 other authors. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2486–2490

32e. Rohdich, F., Wungsintaweekul, J., Eisenreich, W., Richter, G., Schuhr, C. A., Hecht, S., Zenk, M. H., and Bacher, A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6451–6456

32f. Lange, B. M., and Croteau, R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13714–13719

32g. Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13172–13177

32h. Zeidler, J., Ullah, N., Gupta, R. N., Pauloski, R. M., Sayer, B. G., and Spenser, I. D. (2002) *J. Am. Chem. Soc.* **124**, 4542–4543

33. Hill, R. E., Himmeldirk, K., Kennedy, I. A., Pauloski, R. M., Sayer, B. G., Wolf, E., and Spenser, I. D. (1996) *J. Biol. Chem.* **271**, 30426–30435

33a. Paton, V. G., Shackelford, J. E., and Krisans, S. K. (1997) *J. Biol. Chem.* **272**, 18945–18950

References

33b. Kaneda, K., Kuzuyama, T., Takagi, M., Hayakawa, Y., and Seto, H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 932–937

34. Deneris, E. S., Stein, R. A., and Mead, J. F. (1985) *J. Biol. Chem.* **260**, 1382–1385

35. Silver, G. M., and Fall, R. (1995) *J. Biol. Chem.* **270**, 13010–13016

36. Sharkey, T. D., and Singsaas, E. L. (1995) *Nature (London)* **374**, 769

37. Mlot, C. (1995) *Science* **268**, 641–642

38. Brady, P. S., Scofield, R. F., Schumann, W. C., Ohgaku, S., Kumaran, K., Margolis, J. M., and Landau, B. R. (1982) *J. Biol. Chem.* **257**, 10742–10746

39. Chen, A., Kroon, P. A., and Poulter, C. D. (1994) *Protein Sci.* **3**, 600–607

40. Cunillera, N., Arró, M., Delourme, D., Karst, F., Boronat, A., and Ferrer, A. (1996) *J. Biol. Chem.* **271**, 7774–7780

41. Tarshis, L. C., Proteau, P. J., Kellogg, B. A., Sacchettini, J. C., and Poulter, C. D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15018–15023

42. Ohnuma, S.-i, Narita, K., Nakazawa, T., Ishida, C., Takeuchi, Y., Ohto, C., and Nishino, T. (1996) *J. Biol. Chem.* **271**, 30748–30754

43. Ohnuma, S.-i, Nakazawa, T., Hemmi, H., Hallberg, A.-M., Koyama, T., Ogura, K., and Nishino, T. (1996) *J. Biol. Chem.* **271**, 10087–10095

44. Math, S. K., Hearst, J. E., and Poulter, C. D. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6761–6764

45. Ohnuma, S.-i, Hirooka, K., Hemmi, H., Ishida, C., Ohto, C., and Nishino, T. (1996) *J. Biol. Chem.* **271**, 18831–18837

46. Carattoli, A., Romano, N., Ballario, P., Morelli, G., and Macino, G. (1991) *J. Biol. Chem.* **266**, 5854–5859

46a. Kuzuguchi, T., Morita, Y., Sagami, I., Sagami, H., and Ogura, K. (1999) *J. Biol. Chem.* **274**, 5888–5894

46b. Kharel, Y., Zhang, Y.-W., Fujihashi, M., Miki, K., and Koyama, T. (2001) *J. Biol. Chem.* **276**, 28459–28464

47. Tarshis, L. C., Yan, M., Poulter, C. D., and Sacchettini, J. C. (1994) *Biochemistry* **33**, 10871–10877

47a. Fernandez, S. M. S., Kellogg, B. A., and Poulter, C. D. (2000) *Biochemistry* **39**, 15316–15321

48. Sacchettini, J. C., and Poulter, C. D. (1997) *Science* **277**, 1788–1789

49. Koyama, T., Tajima, M., Sano, H., Doi, T., Koike-Takeshita, A., Obata, S., Nishino, T., and Ogura, K. (1996) *Biochemistry* **35**, 9533–9538

50. Poulter, C. D., and Rilling, H. C. (1978) *Acc. Chem. Res.* **11**, 307–313

51. Poulter, C. D., Wiggins, D. L., and Le, A. T. (1981) *J. Am. Chem. Soc.* **103**, 3926–3927

52. Suga, T., Hirata, T., Aoki, T., and Shishibori, T. (1983) *J. Am. Chem. Soc.* **105**, 6178–6179

53. Davisson, V. J., Neal, T. R., and Poulter, C. D. (1993) *J. Am. Chem. Soc.* **115**, 1235–1245

54. Berdis, A. J., and Benkovic, S. J. (1996) *Biochemistry* **35**, 9253–9265

54a. Zhang, Y.-W., Li, X.-Y., and Koyama, T. (2000) *Biochemistry* **39**, 12717–12722

54b. Oh, S. K., Han, K. H., Ryu, S. B., and Kang, H. (2000) *J. Biol. Chem.* **275**, 18482–18488

55. Matsuoka, S., Sagami, H., Kurisaki, A., and Ogura, K. (1991) *J. Biol. Chem.* **266**, 3464–3468

56. Keller, R. K. (1987) *Trends Biochem. Sci.* **12**, 443–445

57. Ericsson, J., Scallen, T. J., Chojnacki, T., and Dallner, G. (1991) *J. Biol. Chem.* **266**, 10602–10607

57a. Pan, J.-J., Chiou, S.-T., and Liang, P.-H. (2000) *Biochemistry* **39**, 10936–10942

57b. Fujihashi, M., Zhang, Y.-W., Higuchi, Y., Li, X.-Y., Koyama, T., and Miki, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4337–4342

58. Frank, D. W., and Waechter, C. J. (1998) *J. Biol. Chem.* **273**, 11791–11798

59. Sagami, H., Igarashi, Y., Tateyama, S., Ogura, K., Roos, J., and Lennarz, W. J. (1996) *J. Biol. Chem.* **271**, 9560–9566

60. Ohkura, T., Fukushima, K., Kurisaki, A., Sagami, H., Ogura, K., Ohno, K., Hara-Kuge, S., and Yamashita, K. (1997) *J. Biol. Chem.* **272**, 6868–6875

61. Murgolo, N. J., Patel, A., Stivala, S. S., and Wong, T. K. (1989) *Biochemistry* **28**, 253–260

62. Takahashi, I., Ogura, K., and Seto, S. (1980) *J. Biol. Chem.* **255**, 4539–4543

63. Wolucka, B. A., and de Hoffmann, E. (1995) *J. Biol. Chem.* **270**, 20151–20155

63a. Schulbach, M. C., Mahapatra, S., Macchia, M., Barontini, S., Papi, C., Minutolo, F., Bertini, S., Brennan, P. J., and Crick, D. C. (2001) *J. Biol. Chem.* **276**, 11624–11630

64. Runquist, M., Ericsson, J., Thelin, A., Chojnacki, T., and Dallner, G. (1994) *J. Biol. Chem.* **269**, 5804–5809

65. Dennis, M. S., Henzel, W. J., Bell, J., Kohr, W., and Light, D. R. (1989) *J. Biol. Chem.* **264**, 18618–18626

66. Ohnuma, S.-i, Koyama, T., and Ogura, K. (1991) *J. Biol. Chem.* **266**, 23706–23713

67. Chen, A., Zhang, D., and Poulter, C. D. (1993) *J. Biol. Chem.* **268**, 21701–21705

67a. Morii, H., Nishihara, M., and Koga, Y. (2000) *J. Biol. Chem.* **275**, 36568–36574

68. Zhang, F. L., and Casey, P. J. (1996) *Ann. Rev. Biochem.* **65**, 241–269

69. Gelb, M. H. (1997) *Science* **275**, 1750–1751

70. Vogt, A., Sun, J., Qian, Y., Tan-Chiu, E., Hamilton, A. D., and Sebti, S. M. (1995) *Biochemistry* **34**, 12398–12403

71. Mu, Y. Q., Omer, C. A., and Gibbs, R. A. (1996) *J. Am. Chem. Soc.* **118**, 1819–1823

72. Del Villar, K., Mitsuzawa, H., Yang, W., Sattler, I., and Tamanoi, F. (1997) *J. Biol. Chem.* **272**, 680–687

73. Park, H.-W., Boduluri, S. R., Moomaw, J. F., Casey, P. J., and Beese, L. S. (1997) *Science* **275**, 1800–1804

74. Mathis, J. R., and Poulter, C. D. (1997) *Biochemistry* **36**, 6367–6376

75. Dunten, P., Kammlott, U., Crowther, R., Weber, D., Palermo, R., and Birktoft, J. (1998) *Biochemistry* **37**, 7907–7912

76. Long, S. B., Casey, P. J., and Beese, L. S. (1998) *Biochemistry* **37**, 9612–9618

76a. Crespo, N. C., Ohkanda, J., Yen, T. J., Hamilton, A. D., and Sebti, S. M. (2001) *J. Biol. Chem.* **276**, 16161–16167

76b. Long, S. B., Hancock, P. J., Kral, A. M., Hellinga, H. W., and Beese, L. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12948–12953

76c. Micali, E., Chehade, K. A. H., Isaacs, R. J., Andres, D. A., and Spielmann, H. P. (2001) *Biochemistry* **40**, 12254–12265

76d. Huang, C.-c, Hightower, K. E., and Fierke, C. A. (2000) *Biochemistry* **39**, 2593–2602

76e. Pang, Y.-P., Xu, K., El Yazal, J., and Prendergast, F. G. (2000) *Protein Sci.* **9**, 1857–1865

76f. Long, S. B., Casey, P. J., and Beese, L. S. (2002) *Nature (London)* **419**, 645–650

77. Yokoyama, K., McGeady, P., and Gelb, M. H. (1995) *Biochemistry* **34**, 1344–1354

78. Witter, D. J., and Poulter, C. D. (1996) *Biochemistry* **35**, 10454–10463

78a. Clausen, V. A., Edelstein, R. L., and Distefano, M. D. (2001) *Biochemistry* **40**, 3920–3930

78b. Dursina, B., Thomä, N. H., Sidorovitch, V., Niculae, A., Iakovenko, A., Rak, A., Albert, S., Ceacareanu, A.-C., Kölling, R., Herrmann, C., Goody, R. S., and Alexandrov, K. (2002) *Biochemistry* **41**, 6805–6816

79. Fu, H.-W., Beese, L. S., and Casey, P. J. (1998) *Biochemistry* **37**, 4465–4472

79a. Ashar, H. R., James, L., Gray, K., Carr, D., Black, S., Armstrong, L., Bishop, W. R., and Kirschmeier, P. (2000) *J. Biol. Chem.* **275**, 30451–30457

79b. Huber, H. E., Robinson, R. G., Watkins, A., Nahas, D. D., Abrams, M. T., Buser, C. A., Lobell, R. B., Patrick, D., Anthony, N. J., Dinsmore, C. J., Graham, S. L., Hartman, G. D., Lumma, W. C., Williams, T. M., and Heimbrook, D. C. (2001) *J. Biol. Chem.* **276**, 24457–24465

80. Leung, H.-C. E., Chen, Y., and Winkler, M. E. (1997) *J. Biol. Chem.* **272**, 13073–13083

81. Gebler, J. C., Woodside, A. B., and Poulter, C. D. (1992) *J. Am. Chem. Soc.* **114**, 7354–7360

82. Clause, E. P., Tyler, V. E., and Brady, L. R. (1970) *Pharmacognosy*, 6th ed., Lea & Febiger, Philadelphia, Pennsylvania

83. Loreto, F., Ciccioli, P., Brancaleoni, E., Cecinato, A., Frattoni, M., and Sharkey, T. D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9966–9969

84. Harborne, J. B., and Tomas-Barberan, F. A., eds. (1991) *Ecological Chemistry and Biochemistry of Plant Terpenoids*, Clarendon Press, Oxford

85. Gershenzon, J., and Croteau, R. B. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 339–388, CRC Press, Boca Raton, Florida

86. Bohlmann, J., Meyer-Gauen, G., and Croteau, R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4126–4133

87. Bohlmann, J., Steele, C. L., and Croteau, R. (1997) *J. Biol. Chem.* **272**, 21784–21792

88. Wise, M. L., Savage, T. J., Katahira, E., and Croteau, R. (1998) *J. Biol. Chem.* **273**, 14891–14899

88a. Williams, D. C., McGarvey, D. J., Katahira, E. J., and Croteau, R. (1998) *Biochemistry* **37**, 12213–12220

89. Croteau, R., and Satterwhite, D. M. (1989) *J. Biol. Chem.* **264**, 15309–15315

90. Seybold, S. J., Quilici, D. R., Tillman, J. A., Vanderwel, D., Wood, D. L., and Blomquist, G. J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8393–8397

91. Steele, C. L., Crock, J., Bohlmann, J., and Croteau, R. (1998) *J. Biol. Chem.* **273**, 2078–2089

91a. Rising, K. A., Starks, C. M., Noel, J. P., and Chappell, J. (2000) *J. Am. Chem. Soc.* **122**, 1861–1866

92. Procter, R. H., and Hohn, T. M. (1993) *J. Biol. Chem.* **268**, 4543–4548

93. Cane, D. E., and Tsantrizos, Y. S. (1996) *J. Am. Chem. Soc.* **118**, 10037–10040

93a. Caruthers, J. M., Kang, I., Rynkiewicz, M. J., Cane, D. E., and Christianson, D. W. (2000) *J. Biol. Chem.* **275**, 25533–25539

94. Back, K., and Chappell, J. (1995) *J. Biol. Chem.* **270**, 7375–7381

95. Starks, C. M., Back, K., Chappell, J., and Noel, J. P. (1997) *Science* **277**, 1815–1820

96. Lesburg, C. A., Zhai, G., Cane, D. E., and Christianson, D. W. (1997) *Science* **277**, 1820–1824

97. Cane, D. E., and Xue, Q. (1996) *J. Am. Chem. Soc.* **118**, 1563–1564

97a. Rynkiewicz, M. J., Cane, D. E., and Christianson, D. W. (2002) *Biochemistry* **41**, 1732–1741

References

97b. Peters, R. J., and Croteau, R. B. (2002) *Biochemistry* **41**, 1836–1842
98. Crock, J., Wildung, M., and Croteau, R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12833 – 12838
99. Marin, E., Nussaume, L., Quesada, A., Gonneau, M., Sotta, B., Hugueney, P., Frey, A., and Marion-Poll, A. (1996) *EMBO J.* **15**, 2331 – 2342
99a. Qin, X., and Zeevaart, J. A. D. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 15354 – 15361
99b. Seo, M., Peeters, A. J. M., Koiwai, H., Oritani, T., Marion-Poll, A., Zeevaart, J. A. D., Koornneef, M., Kamiya, Y., and Koshiba, T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12908 – 12913
100. He, X., Saint-Jeannet, J.-P., Woodgett, J. R., Varmus, H. E., and Dawid, I. B. (1995) *Nature (London)* **374**, 617 – 622
101. Spray, C. R., Kobayashi, M., Suzuki, Y., Phinney, B. O., Gaskin, P., and MacMillan, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10515 – 10518
102. Xu, Y.-L., Li, L., Wu, K., Peeters, A. J. M., Gage, D. A., and Zeevaart, J. A. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6640 – 6644
102a. Rojas, M. C., Hedden, P., Gaskin, P., and Tudzynski, B. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5838 – 5843
102b. Kawaide, H., Sassa, T., and Kamiya, Y. (2000) *J. Biol. Chem.* **275**, 2276 – 2280
103. Lange, T., Hedden, P., and Graebe, J. E. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8552 – 8556
103a. Helliwell, C. A., Chandler, P. M., Poole, A., Dennis, E. S., and Peacock, W. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2065 – 2070
103b. Thomas, S. G., Phillips, A. L., and Hedden, P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4698 – 4703
104. Feyereisen, R., and Farnsworth, D. E. (1987) *J. Biol. Chem.* **262**, 2676 – 2681
105. Toong, Y. C., Schooley, D. A., and Baker, F. C. (1988) *Nature (London)* **333**, 170 – 171
106. Bohlmann, J., Crock, J., Jetter, R., and Croteau, R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6756 – 6761
107. Vogel, B. S., Wildung, M. R., Vogel, G., and Croteau, R. (1996) *J. Biol. Chem.* **271**, 23262 – 23268
107a. Peters, R. J., and Croteau, R. B. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 580 – 584
108. Mau, C. J. D., and West, C. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8497 – 8501
109. Lin, X., Hezari, M., Koepp, A. E., Floss, H. G., and Croteau, R. (1996) *Biochemistry* **35**, 2968 – 2977
110. Goodwin, T. W., ed. (1988) *Plant Pigments*, Academic Press, London
111. Mookhtiar, K. A., Kalinowski, S. S., Zhang, D., and Poulter, C. D. (1994) *J. Biol. Chem.* **269**, 11201 – 11207
112. Gu, P., Ishii, Y., Spencer, T. A., and Shechter, I. (1998) *J. Biol. Chem.* **273**, 12515 – 12525
113. Jarstfer, M. B., Blagg, B. S. J., Rogers, D. H., and Poulter, C. D. (1996) *J. Am. Chem. Soc.* **118**, 13089 – 13090
114. Zurer, P. (1997) *Chem. Eng. News* January 6, 6 – 7
114a. Radisky, E. S., and Poulter, C. D. (2000) *Biochemistry* **39**, 1748 – 1760
114b. Jarstfer, M. B., Zhang, D.-L., and Poulter, C. D. (2002) *J. Am. Chem. Soc.* **124**, 8834–8845
115. Musico, F., Carlson, J. P., Kuehl, L., and Rilling, H. C. (1974) *J. Biol. Chem.* **249**, 3746 – 3749
116. Kojima, Y., Friedlander, E. J., and Bloch, K. (1981) *J. Biol. Chem.* **256**, 7235 – 7239
117. Torssell, K. B. G. (1983) *Natural Product Chemistry*, Wiley, New York (pp. 181 – 215)
118. Armstrong, G. A., and Hearst, J. E. (1996) *FASEB J.* **10**, 228 – 237
118a. Rivera, S. B., Swedlund, B. D., King, G. J., Bell, R. N., Hussey, C. E., Jr., Shattuck-Eidens, D. M., Wrobel, W. M., Peiser, G. D., and Poulter, C. D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4373 – 4378
118b. Soderberg, T., Chen, A., and Poulter, C. D. (2001) *Biochemistry* **40**, 14847 – 14854
119. Heathcock, D. H., Finkelstein, B. L., Aoki, T., and Poulter, C. D. (1985) *Science* **229**, 862 – 863
120. Poulter, C. D., Aoki, T., and Daniels, L. (1988) *J. Am. Chem. Soc.* **110**, 2620 – 2624
121. Britton, G. (1988) in *Plant Pigments* (Goodwin, T. W., ed), pp. 133 – 182, Academic Press, London
122. Chamovitz, D., Sandmann, G., and Hirschberg, J. (1993) *J. Biol. Chem.* **268**, 17348 – 17353
123. Bartley, G. E., Viitanen, P. V., Pecker, I., Chamovitz, D., Hirschberg, J., and Scolnik, P. A. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 6532 – 6536
123a. Arrach, N., Fernández-Martín, R., Cerdá-Olmedo, E., and Avalos, J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1687 – 1692
123b. Cunningham, F. X., Jr., and Gantt, E. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2905 – 2910
124. Armstrong, G. A., Schmidt, A., Sandmann, G., and Hearst, J. E. (1990) *J. Biol. Chem.* **265**, 8329 – 8338
125. Komori, M., Ghosh, R., Takaichi, S., Hu, Y., Mizoguchi, T., Koyama, Y., and Kuki, M. (1998) *Biochemistry* **37**, 8987 – 8994
126. Ouchane, S., Picaud, M., Vernotte, C., Reiss-Husson, F., and Astier, C. (1997) *J. Biol. Chem.* **272**, 1670 – 1676
127. Langer, M., Pauling, A., and Rétey, J. (1995) *Angew. Chem. Int. Ed. Engl.* **34**, 1464 – 1465
127a. Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., Beyer, P., and Potrykus, I. (2000) *Science* **287**, 303 – 305
127b. A contribution of $40 to UNICEF (www.unicefusa.com) will provide vitamin A for a year to 1000 young children, protecting them from blindness.
128. Demmig-Adams, B., Gilmore, A. M., and Adams, W. W., III. (1996) *FASEB J.* **10**, 403 – 412
129. Sun, Z., Gantt, E., and Cunningham, F. X., Jr. (1996) *J. Biol. Chem.* **271**, 24349 – 24352
130. Fraser, P. D., Miura, Y., and Misawa, N. (1997) *J. Biol. Chem.* **272**, 6128 – 6135
131. Bouvier, F., d'Harlingue, A., Hugueney, P., Marin, E., Marion-Poll, A., and Camara, B. (1996) *J. Biol. Chem.* **271**, 28861 – 28867
132. Schroeder, W. A., and Johnson, E. A. (1995) *J. Biol. Chem.* **270**, 18374 – 18379
133. Nelis, H. J. C. F., Lavens, P., Moens, L., Sorgeloos, P., Jonckheere, J. A., Criel, G. R., and DeLeenheer, A. P. (1984) *J. Biol. Chem.* **259**, 6063 – 6066
134. Swift, I. E., and Milborrow, B. V. (1981) *J. Biol. Chem.* **256**, 11607 – 11611
135. Searcy, D. G. (1982) *Trends Biochem. Sci.* **7**, 183 – 185
136. Woese, C. R., Kandler, O., and Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576 – 4579
137. Olson, J. A., and Krinsky, N. I. (1995) *FASEB J.* **9**, 1547 – 1550
138. Britton, G. (1995) *FASEB J.* **9**, 1551 – 1558
139. Zuber, H. (1986) *Trends Biochem. Sci.* **11**, 414 – 419
140. Quiñones, M. A., Lu, Z., and Zeiger, E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2224 – 2228
141. Bugos, R. C., and Yamamoto, H. Y. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6320 – 6325
142. Rock, C. D., and Zeevaart, J. A. D. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7496 – 7499
143. Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A. D., and McCarty, D. R. (1997) *Science* **276**, 1872 – 1874
144. Kennedy, T. A., and Liebler, D. C. (1992) *J. Biol. Chem.* **267**, 4658 – 4663
145. Everett, S. A., Dennis, M. F., Patel, K. B., Maddix, S., Kundu, S. C., and Willson, R. L. (1996) *J. Biol. Chem.* **271**, 3988 – 3994
146. Burton, G. W., and Ingold, K. U. (1984) *Science* **224**, 569 – 573
147. Brush, A. H. (1990) *FASEB J.* **4**, 2969 – 2977
148. Weesie, R. J., Jansen, F. J. H. M., Merlin, J. C., Lugtenburg, J., Britton, G., and de Groot, H. J. M. (1997) *Biochemistry* **36**, 7288 – 7296
149. Parker, R. S. (1996) *FASEB J.* **10**, 542 – 551
150. Hoffman, M. (1990) *Science* **250**, 372 – 373
151. Gudas, L. J. (1994) *J. Biol. Chem.* **269**, 15399 – 15402
152. Means, A. L., and Gudas, L. J. (1995) *Ann. Rev. Biochem.* **64**, 201 – 233
152a. Mertz, J. R., Shang, E., Piantedosi, R., Wei, S., Wolgemuth, D. J., and Blaner, W. S. (1997) *J. Biol. Chem.* **272**, 11744 – 11749
152b. White, J. A., Guo, Y.-D., Baetz, K., Beckett-Jones, B., Bonasoro, J., Hsu, K. E., Dilworth, F. J., Jones, G., and Petkovich, M. (1996) *J. Biol. Chem.* **271**, 29922 – 29927
152c. Lamb, A. L., and Newcomer, M. E. (1999) *Biochemistry* **38**, 6003 – 6011
153. Parish, E. J., and Nes, W. D., eds. (1997) *Biochemistry and Function of Sterols*, CRC Press, Boca Raton, Florida
153a. Crowder, C. M., Westover, E. J., Kumar, A. S., Ostlund, R. E., Jr., and Covey, D. F. (2001) *J. Biol. Chem.* **276**, 44369 – 44372
154. van Tamelen, E. E. (1968) *Acc. Chem. Res.* **1**, 111 – 120
155. Heftmann, E. (1969) *Steroid Biochemistry*, Academic Press, New York
156. Templeton, W. (1969) *An Introduction to the Chemistry of Terpenoids and Steroids*, Butterworths, London
157. Abe, I., and Prestwich, G. D. (1994) *J. Biol. Chem.* **269**, 802 – 804
158. Abe, I., and Prestwich, G. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9274 – 9278
159. Shi, Z., Buntel, C. J., and Griffin, J. H. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7370 – 7374
160. Nes, W. D., and Venkatramesh, M. (1994) in *Isopentenoids and Other Natural Products* (Nes, W. D., ed), pp. 55 – 89, American Chemical Society, Washington D.C.
160a. Edwards, P. A., and Ericsson, J. (1999) *Ann. Rev. Biochem.* **68**, 157 – 185
161. Goad, L. J. (1991) in *Ecological Chemistry and Biochemistry of Plant Terpenoids* (Harborne, J. B., and Tomas-Barberan, F. A., eds), pp. 209 – 229, Clarendon Press, Oxford
162. Nes, W. D., Parker, S. R., Crumley, F. G., and Ross, S. A. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 389 – 426, CRC Press, Boca Raton, Florida
162a. Herrera, J. B. R., Wilson, W. K., and Matsuda, S. P. T. (2000) *J. Am. Chem. Soc.* **122**, 6765 – 6766
163. Nes, W. D. (1990) in *Biochemistry of the Mevalonic Acid Pathway to Terpenoids* (Towers, G. H. N., and Stafford, H. A., eds), pp. 283 – 327, Plenum, New York
163a. Kushiro, T., Shibuya, M., Masuda, K., and Ebizuka, Y. (2000) *J. Am. Chem. Soc.* **122**, 6816 – 6824
163b. Heathcock, C. H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14323 – 14327
164. Wendt, K. U., Poralla, K., and Schulz, G. E. (1997) *Science* **277**, 1811 – 1815
164a. Wendt, K. U., Lenhart, A., and Schulz, G. E. (1999) *J. Mol. Biol.* **286**, 175 – 187
165. Prince, R. C. (1987) *Trends Biochem. Sci.* **12**, 455 – 456

166. Rohmer, M., and Bisseret, P. (1994) in *Isopentenoids and Other Natural Products* (Nes, W. D., ed), pp. 31–43, American Chemical Society, Washington D.C.

167. Fukushima, H., Grinstead, G. F., and Gaylor, J. L. (1981) *J. Biol. Chem.* **256**, 4822–4826

168. Krieger, M., Kingsley, D., Sege, R., Hobbie, L., and Kozarsky, K. (1985) *Trends Biochem. Sci.* **10**, 447–452

169. Williams, J. B., and Napoli, J. L. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4658–4662

170. Lamb, D. C., Kelly, D. E., Schunck, W.-H., Shyadehi, A. Z., Akhtar, M., Lowe, D. J., Baldwin, B. C., and Kelly, S. L. (1997) *J. Biol. Chem.* **272**, 5682–5688

171. Trzaskos, J. M., Ko, S. S., Magolda, R. L., Favata, M. F., Fischer, R. T., Stam, S. H., Johnson, P. R., and Gaylor, J. L. (1995) *Biochemistry* **34**, 9670–9676

172. Shyadehi, A. Z., Lamb, D. C., Kelly, S. L., Kelly, D. E., Schunck, W.-H., Wright, J. N., Corina, D., and Akhtar, M. (1996) *J. Biol. Chem.* **271**, 12445–12450

172a. Podust, L. M., Poulos, T. L., and Waterman, M. R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3068–3073

172b. Bellamine, A., Mangla, A. T., Nes, W. D., and Waterman, M. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8937–8942

173. Bard, M., Bruner, D. A., Pierson, C. A., Lees, N. D., Biermann, B., Frye, L., and Koegel, C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 186–190

173a. Taton, M., Husselstein, T., Benveniste, P., and Rahier, A. (2000) *Biochemistry* **39**, 701–711

174. Seedorf, U., Brysch, P., Engel, T., Schrage, K., and Assmann, G. (1994) *J. Biol. Chem.* **269**, 21277–21283

174a. Gachotte, D., Barbuch, R., Gaylor, J., Nickel, E., and Bard, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13794–13799

175. Pascal, S., Taton, M., and Rahier, A. (1993) *J. Biol. Chem.* **268**, 11639–11654

176. Rahier, A., and Taton, M. (1996) *Biochemistry* **35**, 7069–7076

177. Nes, W. D., Janssen, G. G., and Bergenstrahle, A. (1991) *J. Biol. Chem.* **266**, 15202–15212

178. Shi, J., Gonzalas, R. A., and Bhattacharyya, M. K. (1996) *J. Biol. Chem.* **271**, 9384–9389

179. Russell, D. W. (1996) *Science* **272**, 370–371

179a. Rouleau, M., Marsolais, F., Richard, M., Nicolle, L., Voigt, B., Adam, G., and Varin, L. (1999) *J. Biol. Chem.* **274**, 20925–20930

180. Li, J., Biswas, M. G., Chao, A., Russell, D. W., and Chory, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3554–3559

181. Metcalf, R. L., Metcalf, R. A., and Rhodes, A. M. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3769–3772

182. Russell, D. W., and Setchell, K. D. R. (1992) *Biochemistry* **31**, 4737–4749

183. Lange, Y., and Steck, T. L. (1994) *J. Biol. Chem.* **269**, 29371–29374

184. Liscum, L., and Underwood, K. W. (1995) *J. Biol. Chem.* **270**, 15443–15446

185. Guan, G., Dai, P.-H., Osborne, T. F., Kim, J. B., and Shechter, I. (1997) *J. Biol. Chem.* **272**, 10295–10302

186. Mackay, K., Starr, J. R., Lawn, R. M., and Ellsworth, J. L. (1997) *J. Biol. Chem.* **272**, 13380–13389

187. Lange, Y., and Muraski, M. F. (1988) *J. Biol. Chem.* **263**, 9366–9373

188. Bretscher, M. S., and Munro, S. (1993) *Science* **261**, 1280–1281

188a. Simons, K., and Ikonen, E. (2000) *Science* **290**, 1721–1726

188b. Nilsson, I., Ohvo-Rekilä, H., Slotte, J. P., Johnson, A. E., and von Heijne, G. (2001) *J. Biol. Chem.* **276**, 41748–41754

188c. Parpal, S., Karlsson, M., Thorn, H., and Strålfors, P. (2001) *J. Biol. Chem.* **276**, 9670–9678

189. Lange, Y., Ye, J., and Steck, T. L. (1998) *J. Biol. Chem.* **273**, 18915–18922

190. Puglielli, L., Rigotti, A., Greco, A. V., Santos, M. J., and Nervi, F. (1995) *J. Biol. Chem.* **270**, 18723–18726

190a. López García, F., Szyperski, T., Dyer, J. H., Choinowski, T., Seedorf, U., Hauser, H., and Wüthrich, K. (2000) *J. Mol. Biol.* **295**, 595–603

190b. Schroeder, F., Frolov, A., Starodub, O., Atshaves, B. B., Russell, W., Petrescu, A., Huang, H., Gallegos, A. M., McIntosh, A., Tahotna, D., Russell, D. H., Billheimer, J. T., Baum, C. L., and Kier, A. B. (2000) *J. Biol. Chem.* **275**, 25547–25555

190c. Choinowski, T., Hauser, H., and Piontek, K. (2000) *Biochemistry* **39**, 1897–1902

191. Yang, H., Bard, M., Bruner, D. A., Gleeson, A., Deckelbaum, R. J., Aljinovic, G., Pohl, T. M., Rothstein, R., and Sturley, S. L. (1996) *Science* **272**, 1353–1356

191a. Spady, D. K., Willard, M. N., and Meidell, R. S. (2000) *J. Biol. Chem.* **275**, 27005–27012

192. Jauhiainen, M., and Dolphin, P. J. (1986) *J. Biol. Chem.* **261**, 7032–7043

192a. Seo, T., Oelkers, P. M., Giattina, M. R., Worgall, T. S., Sturley, S. L., and Deckelbaum, R. J. (2001) *Biochemistry* **40**, 4756–4762

192b. Mauch, D. H., Nägler, K., Schumacher, S., Göritz, C., Müller, E.-C., Otto, A., and Pfrieger, F. W. (2001) *Science* **294**, 1354–1357

192c. Barres, B. A., and Smith, S. J. (2001) *Science* **294**, 1296–1297

193. Jiang, X.-c, and Bruce, C. (1995) *J. Biol. Chem.* **270**, 17133–17138

194. Agellon, L. B., Quinet, E. M., Gillette, T. G., Drayna, D. T., Brown, M. L., and Tall, A. R. (1990) *Biochemistry* **29**, 1372–1376

195. Epps, D. E., Greenlee, K. A., Harris, J. S., Thomas, E. W., Castle, C. K., Fisher, J. F., Hozak, R. R., Marschke, C. K., Melchior, G. W., and Kézdy, F. J. (1995) *Biochemistry* **34**, 12560–12569

195a. Föger, B., Chase, M., Amar, M. J., Vaisman, B. L., Shamburek, R. D., Paigen, B., Fruchart-Najib, J., Paiz, J. A., Koch, C. A., Hoyt, R. F., Brewer, H. B., Jr., and Santamarina-Fojo, S. (1999) *J. Biol. Chem.* **274**, 36912–36920

196. Tall, A. (1995) *Ann. Rev. Biochem.* **64**, 235–257

196a. Desrumaux, C., Labeur, C., Verhee, A., Tavernier, J., Vandekerckhove, J., Rosseneu, M., and Peelman, F. (2001) *J. Biol. Chem.* **276**, 5908–5915

197. Feaster, S. R., Quinn, D. M., and Barnett, B. L. (1997) *Protein Sci.* **6**, 73–79

198. Lopez-Candales, A., Bosner, M. S., Spilburg, C. A., and Lange, L. G. (1993) *Biochemistry* **32**, 12085–12089

199. Sutton, L. D., Froelich, S., Hendrickson, H. S., and Quinn, D. M. (1991) *Biochemistry* **30**, 5888–5893

200. Brown, M. S., and Goldstein, J. L. (1986) *Science* **232**, 34–47

201. Motulsky, A. G. (1986) *Science* **231**, 126–128

202. Brown, M. S., and Goldstein, J. L. (1983) *Ann. Rev. Biochem.* **52**, 223–261

203. Griffiths, G., and Simons, K. (1986) *Science* **234**, 438–443

204. Golgi, C. (1898) *Arch. Ital. Biol.* **30**, 60 and 278

205. Breslow, J. L. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2031–2052, McGraw-Hill, New York

206. Dietschy, J. M. (1997) *Am. J. Clin. Nutr.* **65**, 1581S–1589S

207. Krieger, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4077–4080

208. Lopez, J., Roghani, A., Bertrand, J., Zanni, E., Kalopissis, A., Zannis, V. I., and Chambaz, J. (1994) *Biochemistry* **33**, 4056–4064

209. Bergeron, J., Frank, P. G., Scales, D., Meng, Q.-H., Castro, G., and Marcel, Y. L. (1995) *J. Biol. Chem.* **270**, 27429–27438

210. Gong, E. L., Stoltzfus, L. J., Brion, C. M., Murugesh, D., and Rubin, E. M. (1996) *J. Biol. Chem.* **271**, 5984–5987

211. Südhof, T. C., Goldstein, J. L., Brown, M. S., and Russell, D. W. (1985) *Science* **228**, 815–822

212. Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P. A., and Smith, R. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6334–6338

213. Kozarsky, K. F., Donahee, M. H., Rigotti, A., Iqbal, S. N., Edelman, E. R., and Krieger, M. (1997) *Nature (London)* **387**, 414–417

213a. Krieger, M. (1999) *Ann. Rev. Biochem.* **68**, 523–558

214. Brown, M. S., and Goldstein, J. L. (1974) *Sci. Am.* **251**(Nov), 58–66

215. Lehrman, M. A., Russell, D. W., Goldstein, J. L., and Brown, M. S. (1987) *J. Biol. Chem.* **262**, 3354–3361

216. Strickland, D. K., Kounnas, M. Z., and Argraves, W. S. (1995) *FASEB J.* **9**, 890–898

217. Chappell, D. A., Inoue, I., Fry, G. L., Pladet, M. W., Bowen, S. L., Iverius, P.-H., Lalouel, J.-M., and Strickland, D. K. (1994) *J. Biol. Chem.* **269**, 18001–18006

218. Rosseneu, M., and Labeur, C. (1995) *FASEB J.* **9**, 768–776

219. Sakai, J., Hoshino, A., Takahashi, S., Miura, Y., Ishii, H., Suzuki, H., Kawarabayasi, Y., and Yamamoto, T. (1994) *J. Biol. Chem.* **269**, 2173–2182

220. Takahashi, S., Suzuki, J., Kohno, M., Oida, K., Tamai, T., Miyabo, S., Yamamoto, T., and Nakai, T. (1995) *J. Biol. Chem.* **270**, 15747–15754

221. Krieger, M., and Herz, J. (1994) *Ann. Rev. Biochem.* **63**, 601–637

222. Doi, T., Kurasawa, M., Higashino, K.-i, Imanishi, T., Mori, T., Naito, M., Takahashi, K., Kawabe, Y., Wada, Y., Matsumoto, A., and Kodama, T. (1994) *J. Biol. Chem.* **269**, 25598–25604

223. Krieger, M., Acton, S., Ashkenas, J., Pearson, A., Penman, M., and Resnick, D. (1993) *J. Biol. Chem.* **268**, 4569–4572

224. Krieger, M. (1992) *Trends Biochem. Sci.* **17**, 141–146

225. Hajjar, D. P., and Haberland, M. E. (1997) *J. Biol. Chem.* **272**, 22975–22978

226. Schaefer, E. J., and Levy, R. I. (1985) *N. Engl. J. Med.* **312**, 1300–1310

227. Breslow, J. L. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8314–8318

228. Havel, R. J., and Kane, J. P. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1841–1852, McGraw-Hill, New York

229. Goldstein, J. L., Hobbs, H. H., and Brown, M. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1981–2030, McGraw-Hill, New York

229a. Goldstein, J. L., and Brown, M. S. (2001) *Science* **292**, 1310–1312

229b. Garcia, C. K., and 12 other authors. (2001) *Science* **292**, 1394–1398

230. Fass, D., Blacklow, S., Kim, P. S., and Berger, J. M. (1997) *Nature (London)* **388**, 691–693

230a. Borén, J., Ekström, U., Ågren, B., Nilsson-Ehle, P., and Innerarity, T. L. (2001) *J. Biol. Chem.* **276**, 9214–9218

References

231. Glomset, J. A., Assmann, G., Gjone, E., and Norum, K. R. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1933–1952, McGraw-Hill, New York

232. Pennisi, E. (1997) *Science* **277**, 180–181

233. Carstea, E. D., and 37 other authors. (1997) *Science* **277**, 228–231

234. Loftus, S. K., Morris, J. A., Carstea, E. D., Gu, J. Z., Cummings, C., Brown, A., Ellison, J., Ohno, K., Rosenfeld, M. A., Tagle, D. A., Pentchev, P. G., and Pavan, W. J. (1997) *Science* **277**, 232–235

234a. Davies, J. P., Ioannou,Y. A. (2000) *J. Biol. Chem.* **275**, 24367–24374

234b. Davies, J. P., Chen, F. W., and Ioannou, Y. A. (2000) *Science* **290**, 2295–2298

234c. Naureckiene, S., Sleat, D. E., Lackland, H., Fensom, A., Vanier, M. T., Wattiaux, R., Jadot, M., and Lobel, P. (2000) *Science* **290**, 2298–2301

235. Assmann, G., von Eckardstein, A., and Brewer, H. B., Jr. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2053–2072, McGraw-Hill, New York

236. Lohse, P., Kindt, M. R., Rader, D. J., and Brewer, H. B., Jr. (1991) *J. Biol. Chem.* **266**, 13513–13518

236a. Gura, T. (1999) *Science* **285**, 814–815

236b. Scott, J. (1999) *Nature (London)* **400**, 816–819

236c. Liscovitch, M., and Lavie, Y. (2000) *Trends Biochem. Sci.* **25**, 530–534

236d. Berge, K. E., Tian, H., Graf, G. A., Yu, L., Grishin, N. V., Schultz, J., Kwiterovich, P., Shan, B., Barnes, R., and Hobbs, H. H. (2000) *Science* **290**, 1771–1775

236e. Allayee, H., Laffitte, B. A., and Lusis, A. J. (2000) *Science* **290**, 1709–1711

237. Bloch, K. (1982) *Trends Biochem. Sci.* **7**, 334–336

238. Tanaka, N., Nonaka, T., Tanabe, T., Yoshimoto, T., Tsuru, D., and Mitsui, Y. (1996) *Biochemistry* **35**, 7715–7730

238a. Kramer, W., Sauber, K., Baringhaus, K.-H., Kurz, M., Stengelin, S., Lange, G., Corsiero, D., Girbig, F., König, W., and Weyland, C. (2001) *J. Biol. Chem.* **276**, 7291–7301

238b. Lundell, K., Hansson, R., and Wikvall, K. (2001) *J. Biol. Chem.* **276**, 9606–9612

238c. del Castillo-Olivares, A., and Gil, G. (2000) *J. Biol. Chem.* **275**, 17793–17799

238d. Steinberg, S. J., Mihalik, S. J., Kim, D. G., Cuebas, D. A., and Watkins, P. A. (2000) *J. Biol. Chem.* **275**, 15605–15608

238e. Kotti, T. J., Savolainen, K., Helander, H. M., Yagi, A., Novikov, D. K., Kalkkinen, N., Conzelmann, E., Hiltunen, J. K., and Schmitz, W. (2000) *J. Biol. Chem.* **275**, 20887–20895

238f. Pullinger, C. R., Eng, C., Salen, G., Shefer, S., Batta, A. K., Erickson, S. K., Verhagen, A., Rivera, C. R., Mulvihill, S. J., Malloy, M. J., and Kane, J. P. (2002) *J. Clin. Invest.* **110**, 109–117

239. Berginer, V. M., Salem, G., and Shefer, S. (1984) *N. Engl. J. Med.* **311**, 1649–1652

240. Björkhem, I., and Boberg, K. M. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2073–2099, McGraw-Hill, New York

241. de Vree, J. M. L., Jacquemin, E., Sturm, E., Cresteil, D., Bosma, P. J., Aten, J., Deleuze, J.-F., Desrochers, M., Burdelski, M., Bernard, O., Oude Elferink, R. P. J., and Hadchouel, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 282–287

242. Breuer, O., and Björkhem, I. (1995) *J. Biol. Chem.* **270**, 20278–20284

243. Korytowski, W., Geiger, P. G., and Girotti, A. W. (1996) *Biochemistry* **35**, 8670–8679

244. Javitt, N. B., Kok, E., Burstein, S., Cohen, B., and Kutscher, J. (1981) *J. Biol. Chem.* **256**, 12644–12646

245. Anderson, C. J., Lucas, L. J. H., and Widlanski, T. S. (1995) *J. Am. Chem. Soc.* **117**, 3889–3890

246. Ballabio, A., Parenti, G., Carrozzo, R., Sebastio, G., Andria, G., Buckle, V., Fraser, N., Craig, I., Rocchi, M., Romeo, G., Jobsis, A. C., and Persico, M. G. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4519–4523

247. Beaulieu, M., Lévesque, E., Hum, D. W., and Bélanger, A. (1996) *J. Biol. Chem.* **271**, 22855–22862

248. Porter, J. A., Young, K. E., and Beachy, P. A. (1996) *Science* **274**, 255–258

249. Briggs, M. H., and Brotherton, J. (1970) *Steriod Biochemistry and Pharmacology*, Academic Press, New York

250. Waterman, M. R., and Bischof, L. J. (1997) *FASEB J.* **11**, 419–427

251. Waterman, M. R. (1995) *Science* **267**, 1780–1781

252. Lin, D., Sugawara, T., Strauss, J. F., III, Clark, B. J., Stocco, D. M., Saenger, P., Rogol, A., and Miller, W. L. (1995) *Science* **267**, 1828–1831

253. Heyl, B. L., Tyrrell, D. J., and Lambeth, J. D. (1986) *J. Biol. Chem.* **261**, 2743–2749

254. Rhéaume, E., Sanchez, R., Mébarki, F., Gagnon, E., Carel, J.-C., Chaussain, J.-L., Morel, Y., Labrie, F., and Simard, J. (1995) *Biochemistry* **34**, 2893–2900

255. Bain, P. A., Yoo, M., Clarke, T., Hammond, S. H., and Payne, A. H. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8870–8874

256. Borg, W., Shackleton, C. H. L., Pahuja, S. L., and Hochberg, R. B. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1545–1549

256a. Avvakumov, G. V., Muller, Y. A., and Hammond, G. L. (2000) *J. Biol. Chem.* **275**, 25920–25925

256b. Grishkovskaya, I., Avvakumov, G. V., Sklenar, G., Dales, D., Hammond, G. L., and Muller, Y. A. (2000) *EMBO J.* **19**, 504–512

257. Mester, J., and Baulieu, E.-E. (1984) *Trends Biochem. Sci.* **9**, 56–59

258. Baulieu, E.-E. (1989) *Science* **245**, 1351–1357

259. Ulmann, A., Teutsch, G., and Philibert, D. (1990) *Sci. Am.* **262**(Jun), 42–48

260. Beck, C. A., Weigel, N. L., Moyer, M. L., Nordeen, S. K., and Edwards, D. P. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4441–4445

261. Glasier, A., Thong, K. J., Dewar, M., Mackie, M., and Baird, D. T. (1992) *N. Engl. J. Med.* **327**, 1041–1044

262. Grimes, D. A., and Cook, R. J. (1992) *N. Engl. J. Med.* **327**, 1088–1089

263. Brann, D. W., and Mahesh, V. B. (1991) *FASEB J.* **5**, 2691–2698

264. Narasimhulu, S., Eddy, C. R., Dibartolomeis, M., Kowluru, R., and Jefcoate, C. R. (1985) *Biochemistry* **24**, 4287–4294

265. White, P. C. (1994) *N. Engl. J. Med.* **331**, 250–258

266. Phillips, I. R., and Shepard, E. A. (1985) *Nature (London)* **314**, 130–131

267. Goldsworthy, G. J., Robinson, J., and Mordue, W. (1981) *Endocrimology*, Blackie, Glascow

268. Makin, H. L. J., ed. (1984) *Biochemistry of the Steroid Hormones*, 2nd ed., Blackwell, Oxford

269. New, M. I., ed. (1985) *Congenital Adrenal Hyperplasia (Ann. N. York Acad. Sci.)*, Vol. 458, New York Acad. Sci., New York

270. Chrousos, G. P. (1995) *N. Engl. J. Med.* **332**, 1351–1362

271. Hatakeyama, H., Miyamori, I., Fujita, T., Takeda, Y., Takeda, R., and Yamamoto, H. (1994) *J. Biol. Chem.* **269**, 24316–24320

272. McEwen, B. (1981) *Nature (London)* **291**, 610

273. Moore, R. J., Griffin, J. E., and Wilson, J. D. (1975) *J. Biol. Chem.* **250**, 7168–7172

274. Bull, H. G., Garcia-Calvo, M., Andersson, S., Baginsky, W. F., Chan, H. K., Ellsworth, D. E., Miller, R. R., Stearns, R. A., Bakshi, R. K., Rasmusson, G. H., Tolman, R. L., Myers, R. W., Kozarich, J. W., and Harris, G. S. (1996) *J. Am. Chem. Soc.* **118**, 2359–2365

275. Russell, D. W., and Wilson, J. D. (1994) *Ann. Rev. Biochem.* **63**, 25–61

276. Lin, T.-M., and Chang, C. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4988–4993

277. Rittmaster, R. S. (1994) *N. Engl. J. Med.* **330**, 120–125

278. Gormley, G. J., Stoner, E., Bruskewitz, R. C., Imperato-McGinley, J., Walsh, P. C., McConnell, J. D., Andriole, G. L., Geller, J., Bracken, B. R., Tenover, J. S., Vaughan, E. D., Pappas, F., Taylor, A., Binkowitz, B., and Ng, J. (1992) *N. Engl. J. Med.* **327**, 1185–1191

279. Baulieu, E.-E., and Robel, P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4089–4091

280. Lardy, H., Partridge, B., Kneer, N., and Wei, Y. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6617–6619

281. Berdanier, C. D., Parente, J. A., Jr., and Mcintosh, M. K. (1993) *FASEB J.* **7**, 414–419

282. Nestler, J. E., Clore, J. N., and Blackard, W. G. (1992) *FASEB J.* **6**, 3073–3075

283. Berr, C., Lafont, S., Debuire, B., Dartigues, J.-F., and Baulieu, E.-E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13410–13415

284. Rose, K. A., Stapleton, G., Dott, K., Kieny, M. P., Best, R., Schwarz, M., Russell, D. W., Björkhem, I., Seckl, J., and Lathe, R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4925–4930

285. Wickelgren, I. (1997) *Science* **276**, 675–678

285a. Yen, S. S. C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8167–8169

285b. Baulieu, E.-E., Thomas, G., Legrain, S., Lahlou, N., Roger, M., Debuire, B., Faucounau, V., Girard, L., Hervy, M.-P., Latour, F., Leaud, M.-C., Mokrane, A., Pitti-Ferrandi, H., and 10 other authors. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4279–4284

285c. Rose, K., Allan, A., Gauldie, S., Stapleton, G., Dobbie, L., Dott, K., Martin, C., Wang, L., Hedlund, E., Seckl, J. R., Gustafsson, J.-Å., and Lathe, R. (2001) *J. Biol. Chem.* **276**, 23937–23944

286. Cascio, C., Prasad, V. V. K., Lin, Y. Y., Lieberman, S., and Papadopoulos, V. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2862–2867

287. Kochakian, C. D. (1989) *Foundations of Endocrinology*, Univ. of Alabama School of Medicine, Tuscaloosa

288. Kochakian, C. D. (1987) *Trends Biochem. Sci.* **12**, 446–450

289. Bardin, C. W. (1996) *N. Engl. J. Med.* **335**, 52

290. Matsunaga, T., Nomoto, M., Kozak, C. A., and Gonzalez, F. J. (1990) *Biochemistry* **29**, 1329–1341

290a. Yamada, A., Yamada, M., Fujita, Y., Nishigami, T., Nakasho, K., and Uematsu, K. (2001) *J. Biol. Chem.* **276**, 4604–4610

291. Waxman, D. J., Dannan, G. A., and Guengerich, F. P. (1985) *Biochemistry* **24**, 4409–4417

292. MacGeoch, C., Morgan, E. T., Halper, J., and Gustafsson, J.-A. (1984) *J. Biol. Chem.* **259**, 15433–15439

293. Covey, D. F., Hood, W. F., Beusen, D. D., and Carrell, H. L. (1984) *Biochemistry* **23**, 5398–5406

294. Vaz, A. D. N., Kessell, K. J., and Coon, M. J. (1994) *Biochemistry* **33**, 13651–13661

References

295. Harada, N., Ogawa, H., Shozu, M., Yamada, K., Suhara, K., Nishida, E., and Takagi, Y. (1992) *J. Biol. Chem.* **267**, 4781 – 4785
296. Graham-Lorence, S., Khalil, M. W., Lorence, M. C., Mendelson, C. R., and Simpson, E. R. (1991) *J. Biol. Chem.* **266**, 11939 – 11946
297. Terashima, I., Suzuki, N., Itoh, S., Yoshizawa, I., and Shibutani, S. (1998) *Biochemistry* **37**, 8803 – 8807
298. Service, R. F. (1998) *Science* **279**, 1631 – 1633
299. Henderson, B. E., Ross, R. K., and Pike, M. C. (1993) *Science* **259**, 633 – 638

299a. Hong, W. K., and Sporn, M. B. (1997) *Science* **278**, 1073 – 1077

299b. Shimotakahara, S., Gorin, A., Kolbanovskiy, A., Kettani, A., Hingerty, B. E., Amin, S., Broyde, S., Geacintov, N., and Patel, D. J. (2000) *J. Mol. Biol.* **302**, 377 – 393

300. Marshall, E. (1994) *Science* **264**, 1524 – 1527
301. Davis, D. L., and Bradlow, H. L. (1995) *Sci. Am.* **273**(Oct), 166 – 172
302. Kaiser, J. (1996) *Science* **272**, 1418
303. Arnold, S. F., Klotz, D. M., Collins, B. M., Vonier, P. M., Guillette, J., LJ, and McLachlan, J. A. (1996) *Science* **272**, 1489 – 1492
304. Hess, R. A., Bunick, D., Lee, K.-H., Bahr, J., Taylor, J. A., Korach, K. S., and Lubahn, D. B. (1997) *Nature (London)* **390**, 509 – 512
305. Sharpe, R. M. (1997) *Nature (London)* **390**, 447 – 448
306. O'Malley, B. W., and Schrader, W. T. (1976) *Sci. Am.* **234**(Feb), 32 – 43
307. Fuller, P. J. (1991) *FASEB J.* **5**, 3092 – 3099
308. La Baer, J., and Yamamoto, K. R. (1994) *J. Mol. Biol.* **239**, 664 – 688
309. Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.-M., Argos, P., and Chambon, P. (1986) *Nature (London)* **320**, 134 – 139
310. Kostrouch, Z., Kostrouchova, M., and Rall, J. E. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 156 – 159
311. Scherrer, L. C., Dalman, F. C., Massa, E., Meshinchi, S., and Pratt, W. B. (1990) *J. Biol. Chem.* **265**, 21397 – 21400
312. Tsai, M.-J., and O'Malley, B. W. (1994) *Ann. Rev. Biochem.* **63**, 451 – 486
313. Htun, H., Barsony, J., Renyi, I., Gould, D. L., and Hager, G. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4845 – 4850
314. Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., and Sigler, P. B. (1991) *Nature (London)* **352**, 497 – 505
315. Eriksson, M. A. L., Härd, T., and Nilsson, L. (1995) *Biophys. J.* **68**, 402 – 426
316. Schwabe, J. W. R., and Rhodes, D. (1991) *Trends Biochem. Sci.* **16**, 291 – 296
317. Hyder, S. M., Nawaz, Z., Chiappetta, C., Yokoyama, K., and Stancel, G. M. (1995) *J. Biol. Chem.* **270**, 8506 – 8513
318. Rastinejad, F., Perlmann, T., Evans, R. M., and Sigler, P. B. (1995) *Nature (London)* **375**, 203 – 211
319. Lee, M.-O., Dawson, M. I., Picard, N., Hobbs, P. D., and Pfahl, M. (1996) *J. Biol. Chem.* **271**, 11897 – 11903
320. Vivat, V., Zechel, C., Wurtz, J.-M., Bourguet, W., Kagechika, H., Umemiya, H., Shudo, K., Moras, D., Gronemeyer, H., and Chambon, P. (1997) *EMBO J.* **16**, 5697 – 5709
321. Brzozowski, A. M., Pike, A. C. W., Dauter, Z., Hubbard, R. E., Bonn, T., Engström, O., Öhman, L., Greene, G. L., Gustafsson, J.-Å., and Carlquist, M. (1997) *Nature (London)* **389**, 753 – 758
322. Yeh, H., Miyamoto, H., Shima, H., and Chang, C. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5527 – 5532
323. Paech, K., Webb, P., Kuiper, G. G. J. M., Nilsson, S., Gustafsson, J.-Å., Kushner, P. J., and Scanlan, T. S. (1997) *Science* **277**, 1508 – 1510
324. Oñate, S. A., Tsai, S. Y., Tsai, M.-J., and O'Malley, B. W. O. (1995) *Science* **70**, 1354 – 1357
325. Grazzini, E., Guillon, G., Mouillac, B., and Zingg, H. H. (1998) *Nature (London)* **392**, 509 – 512
326. Kester, H. A., van der Leede, B.-jM., van der Saag, P. T., and van der Burg, B. (1997) *J. Biol. Chem.* **272**, 16637 – 16643
327. Bowyer, P., Clarke, B. R., Lunness, P., Daniels, M. J., and Osbourn, A. E. (1995) *Science* **267**, 371 – 374
328. Smith, T. W. (1988) *N. Engl. J. Med.* **318**, 358 – 365
329. Albuquerque, E. X., Daly, J. W., and Witkop, B. (1971) *Science* **172**, 995 – 1002
330. Daly, J. W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9 – 13

330a. Dumbacher, J. P., Spande, T. F., and Daly, J. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12970–12975

331. Coudron, T. A., Law, J. H., and Koeppe, J. K. (1981) *Trends Biochem. Sci.* **6**, 248 – 252
332. Pongs, O. (1985) in *Interaction of Steroid Hormone Receptors with DNA* (Sluyser, M., ed), pp. 226 – 240, Ellis Horwood, Chichester
333. Murray, I. A., Cann, P. A., Day, P. J., Derrick, J. P., Sutcliffe, M. J., Shaw, W. V., and Leslie, A. G. W. (1995) *J. Mol. Biol.* **254**, 993 – 1005
334. Moore, K. S., Wehrli, S., Roder, H., Rogers, M., Forrest, J. N., Jr., McCrimmon, D., and Zasloff, M. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1354 – 1358
335. Stone, R. (1993) *Science* **259**, 1125

Study Questions

1. Outline the sequence and chemical mechanisms of the reactions involved in the conversion of acetyl-CoA into mevalonate.

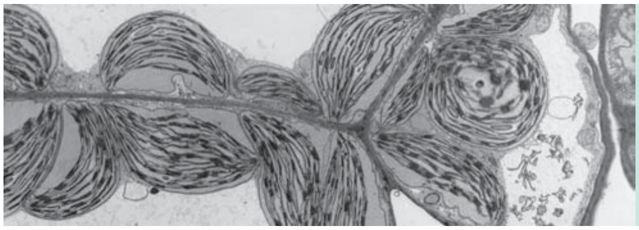
2. a) Show the structures of the reactants for the hydroxymethylglutaryl CoA synthase reaction.
   b) Free coenzyme A is liberated in the above reaction. From which molecule did it come? Explain the metabolic purpose behind the liberation of free CoA.

3. List as many substances as you can that are of polyprenyl origin and are present
   a) in foods
   b) in various commercial products

4. What distinctly different functions do 3-hydroxy-3-methylglutaryl-CoA synthases serve in the cytosol and in mitochondria of the liver?

5. Outline the functions of mitochondrial enzymes in the conversion of fructose into cholesterol in the liver.

6. How do you think that hydroxycitrate, an inhibitor of ATP citrate lyase, would affect the ability of liver to convert dietary fructose into bile acids?



Chloroplasts fill most of the cytoplasm around the junction of three cells of *Arabidopsis thaliana* seen in this micrograph. Both grana stacks and stroma lamellae (pictured in more detail in Fig. 23-19) can be seen. Also present are several small mitochondria. Portions of the large vacuoles, characteristic of plant cells, are seen at top, right, and bottom. Micrograph courtesy of Kenneth Moore.

Contents

1273 .... **A. Properties of Light**
1275 .... **B. Absorption of Light**
1275 .... 1. Quantitative Measurement of Light Absorption, Spectroscopy
1276 .... 2. The Energy Levels of Molecules
1276 .... 3. Infrared (IR) and Raman Spectra
1277 .... *Vibrational frequencies of amide groups*
1278 .... *Raman spectra*
1280 .... 4. Electronic Spectra
1280 .... *Shapes of absorption bands*
1282 .... *Classification of electronic transitions*
1283 .... *Use of plane polarized light*
1283 .... *Relationship of absorption positions and intensity to structures*
1284 .... *Spectra of proteins and nucleic acids*
1285 .... *Difference spectra and derivative spectra*
1286 .... 5. Circular Dichroism and Optical Rotatory Dispersion
1287 .... 6. Photoacoustic Spectroscopy
1288 .... 7. X-Ray Absorption and Mössbauer Spectroscopies
1288 .... **C. Fluorescence and Phosphorescence**
1288 .... 1. Excitation and Emission Spectra
1290 .... *Fluorescence lifetimes*
1290 .... *Time-resolved fluorescence spectroscopy*
1290 .... *Triplet states, phosphorescence, and quenching*
1291 .... *Anisotropy*
1291 .... 2. Fluorescence Resonance Energy Transfer (FRET)
1293 .... 3. Energy-Selective Spectroscopic Methods
1293 .... 4. Analytical Applications of Fluorescence
1294 .... **D. Photochemistry**
1295 .... 1. Chemical Equilibria in the Excited State
1296 .... 2. Photoreactions of Nucleic Acid Bases
1297 .... 3. Sunburn, Cancer, and Phototherapy
1297 .... 4. Photoaffinity Labels
1297 .... 5. Microphotolysis and Ultrafast Light-Induced Reactions
1298 .... 6. Optical Tweezers, Light-Directed Synthesis, and Imaging
1298 .... **E. Photosynthesis**
1299 .... 1. Two Photosystems, the Z Scheme, and Reaction Centers
1299 .... *The Z scheme*
1300 .... *Electron transport and photophosphorylation*
1301 .... *Bacterial photosynthesis*
1301 .... 2. Chloroplast Structure
1301 .... *Chloroplast membranes*
1302 .... *Photosynthetic pigments and their environments*
1305 .... 3. The Light-Receiving Complexes
1305 .... *Phycobilinosomes*
1306 .... *Purple photosynthetic bacteria*
1306 .... *Green sulfur and nonsulfur bacteria*
1308 .... *Eukaryotic plants and cyanobacteria*
1310 .... 4. The Reaction Centers and Their Photochemistry
1310 .... *Reaction centers of purple bacteria*
1312 .... *Reaction center kinetics*
1314 .... *Cyclic photophosphorylation in purple bacteria*
1314 .... *Comparison with other reaction centers*
1314 .... *PSI of cyanobacteria and green plants*
1317 .... *PSII and formation of oxygen*
1318 .... *ATP synthesis in chloroplasts*
1319 .... *Protection of chloroplasts against radiation and oxygen*
1319 .... 5. Control of Photosynthesis
1320 .... *Light-induced transcription*
1320 .... *Light-induced control via the ferredoxin/thioredoxin system*
1321 .... 6. Photorespiration; C_3 and C_4 Plants
1321 .... *Metabolism of glycolic acid*
1322 .... *The C_4 cycle for concentration of carbon dioxide*
1323 .... *Metabolism in the family Crassulaceae*
1324 .... 7. Photosynthetic Formation of Hydrogen
1324 .... **F. Vision**
1324 .... 1. Visual Pigments
1325 .... *Transmembrane structure*
1326 .... *The visual chromophores*
1329 .... 2. The Light-Induced Transformation
1330 .... 3. The Nerve Impulse
1330 .... *Transducin, cyclic GMP, and phosphodiesterase*
1331 .... *Rhodopsin kinase, recoverin, and arrestin*
1332 .... *Some details about cone cells and invertebrate vision*
1332 .... 4. Regeneration of Visual Pigments; the Retinal Cycle
1332 .... 5. Diseases of the Retina
1333 .... 6. Proteins of the Lens
1333 .... **G. Bacteriorhodopsin and Related Ion Pumps and Sensors**
1335 .... *Halorhodopsin*
1335 .... *Sensory rhodopsins*
1336 .... *The photoactive yellow proteins (xanthopsins)*
1336 .... *Stentorin*
1337 .... **H. Phytochrome**
1338 .... **I. Some Blue Light Responses**
1339 .... 1. Cryptochromes
1339 .... 2. Photolyases
1340 .... **J. Bioluminescence**

1347 .... **References**
1357 .... **Study Questions**

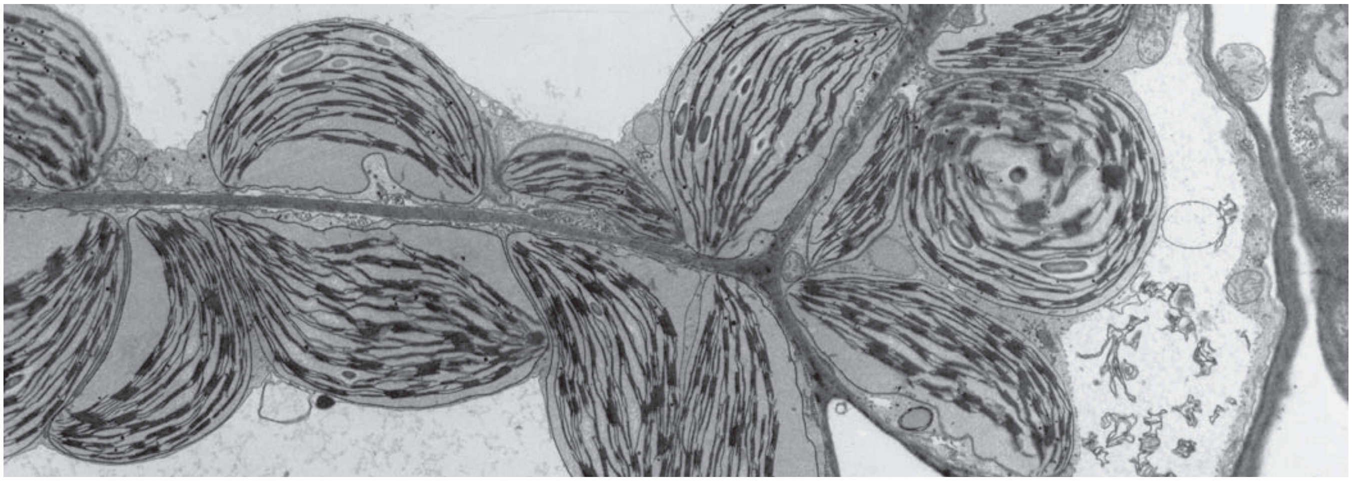
Boxes

1344 .... Box 23-A The Green Fluorescent Protein and Other Light-Emitting Antennas

Tables

1273 .... Table 23-1 Some Properties of Light
1303 .... Table 23-2 Approximate Composition of Photosynthetic Units in a Spinach Chloroplast
1305 .... Table 23-3 The Common Bilin Pigments Present in Phycobiliniproteins
1313 .... Table 23-4 Properties of Various Reaction Centers

23
Light and Life



Light plays a pervasive role in human life. The earth is bathed in light from the sun, and from this light comes not only warmth but also energy for all living organisms. Of the 3 x 10^4 kJ m^{-2} of light energy falling on the earth each day,[1,2] ~30 kJ m^{-2} are captured by photosynthesis.[3] Light penetrating the atmosphere allows us to see and provides color to our environment. It controls the flowering of plants, the germination of seeds and spores, the greening of seedlings, and the daily cycles of many organisms. High in the stratosphere ultraviolet light reacts with oxygen to create a protective blanket of ozone. The ultraviolet light that is not screened out by the ozone layer kills bacteria, tans our skin, and often mutates our DNA, inducing many cancers.[4,5] Organisms, from bacteria to higher plants, display **phototaxis**, the ability to move toward a source of light or to orient themselves with respect to a source of light. In plants the chloroplasts assume an orientation that maximizes efficiency of light absorption. Plants grow toward light (**phototropism**), and some organisms avoid light. Many organisms emit light.

Many of our most important experimental techniques involve the use of light or of other forms of electromagnetic radiation of a wide range of energies. X-rays, ultraviolet light, infrared light, and microwaves all serve in the study of biomolecules.

A. Properties of Light

Light is a form of electromagnetic radiation and possesses characteristics of both waves and particles (**photons**). The energy of a photon is usually measured by the frequency (or by the wavelength in a vacuum to which it is inversely related, Table 23-1). A portion of the electromagnetic spectrum is shown on a logarithmic scale in Fig. 23-1.[1] At the high-energy end (off the scale of the figure to the right) are cosmic rays and gamma rays, while at the low-energy end radio waves extend to wavelengths of many kilometers. The narrow range of wavelengths from about 100 nm to a few micrometers, which is the subject of this chapter, includes the ultraviolet, visible, and near infrared

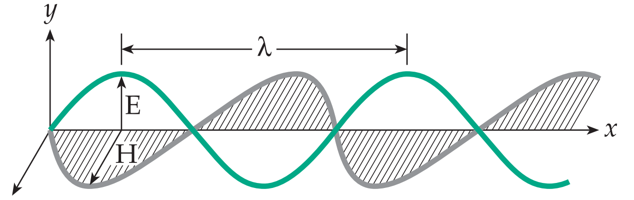
TABLE 23-1
Some Properties of Light

Velocity of light in a vacuum	$c = 2.998 \times 10^8$ m s^{-1}
Velocity of light in a medium	$c' = c/n$ where n = refractive index
Wave number (in nm)	$\bar{\nu} = 1/\lambda$; $\bar{\nu}$ (in cm^{-1}) = $10^7/\lambda$
Frequency	$\nu = c/\lambda = c\bar{\nu}$ ν (in hertz) = $2.998 \times 10^{10}\ \bar{\nu}$ (cm^{-1}) in a vacuum
Energy of quantum	$E = h\nu = hc\bar{\nu}$ E (joules) = $1.986 \times 10^{-23}\ \bar{\nu}$ (cm^{-1}) E (eV) = $1.240 \times 10^{-4}\ \bar{\nu}$ (cm^{-1})
Energy of einstein	$E = Nh\nu = Nhc\bar{\nu}$ $= 6.023 \times 10^{23}\ hc\bar{\nu}$ E (joules) = $11.961\ \bar{\nu}$ (cm^{-1}) E (kcal) = $2.859 \times 10^{-3}\ \bar{\nu}$ (cm^{-1})

ranges. The second line of Fig. 23-1 shows this region expanded. Note that the range of light reaching the earth's surface is narrow, largely being confined to wavelengths of 320–1100 nm. The human eye responds to an even more limited range of 380–760 nm, in which all of the colors of the rainbow can be found. The aromatic rings of proteins and nucleic acids absorb maximally at 280 and 260 nm, respectively. Even though these wavelengths are largely screened out by the ozone layer of the stratosphere, enough light penetrates to cause mutations and to damage the skin of the unwary sunbather.

The energy of a quantum of light is proportional to the **wave number** or **frequency**. The wave number $\bar{\nu}$ is the reciprocal of wavelength and is customarily given in units of cm^{-1} (**reciprocal centimeters**). Most of the absorption spectra in this book are plotted against wave number in cm^{-1}. The frequency ν in **hertz** is equal to $c'\nu$, where c' is the velocity of light in a medium. (The velocity of light in a vacuum is designated c and is equal to 3.00×10^8 m s^{-1}.) The energy of a quantum of light E is equal to $h\nu$, where h is Planck's constant, 6.626×10^{-34} J s^{-1}. From a chemical viewpoint, we are more interested in the energy of one **einstein**, i.e., one "mole" of light (6.023×10^{23} quanta). The energy in kJ per einstein is 11,960 $\bar{\nu}$ (in cm^{-1}, vacuum). Energy relationships are summarized in Table 23-1. The lower three scales of Fig. 23-1 also show the relationships of $\bar{\nu}$, ν, and E to wavelength.

The light wave is characterized by oscillating electrical and magnetic fields.[2,3,6] For propagation of light in the x direction the electric field vector **E**, which is customarily plotted in the y direction, is a function of the wavelength λ and the time (Eq. 23-1).



$$E_y = A \sin 2\pi(x/\lambda - \nu t + \phi) \tag{23-1}$$

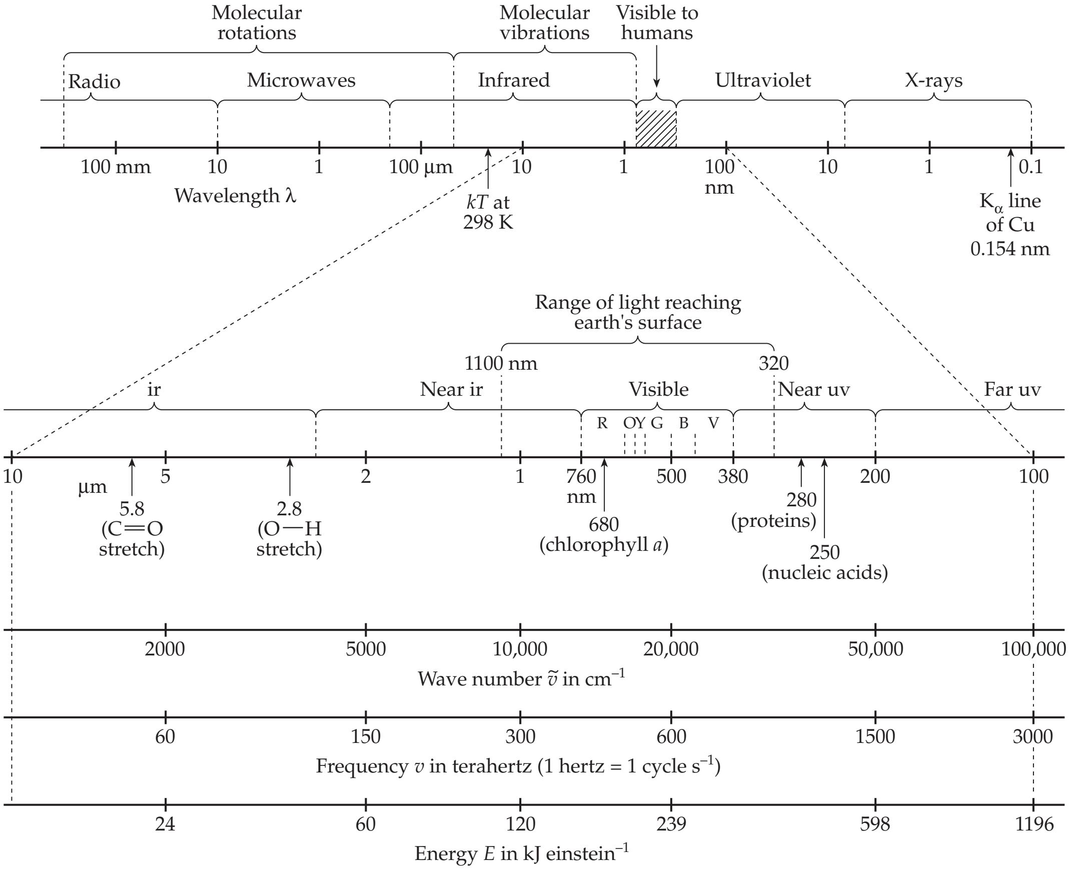

Molecular rotations
Molecular vibrations
Visible to humans
Radio
Microwaves
Infrared
Ultraviolet
X-rays
100 mm
10
1
100 µm
10
1
100 nm
10
1
0.1
Wavelength λ
kT at 298 K
Kα line of Cu 0.154 nm
Range of light reaching earth's surface
1100 nm
320
ir
Near ir
Visible
Near uv
Far uv
R
O
Y
G
B
V
10
µm
5
2
1
760 nm
500
380
200
100
5.8 (C=O stretch)
2.8 (O—H stretch)
680 (chlorophyll a)
280 (proteins)
250 (nucleic acids)
2000
5000
10,000
20,000
50,000
100,000
Wave number ṽ in cm⁻¹
60
150
300
600
1500
3000
Frequency v in terahertz (1 hertz = 1 cycle s⁻¹)
24
60
120
239
598
1196
Energy E in kJ einstein⁻¹

Figure 23-1 A part of the electromagnetic spectrum. The letters V, B, G, Y, O, R over the visible part of the spectrum refer to the colors of the light. The position marked "K_a line of Cu" is the wavelength of X-rays and most widely employed in X-ray diffraction studies of proteins and other organic materials.

The magnetic vector H is at right angles to the electric vector and is given by Eq. 23-2.

$$H_z = (\varepsilon / \mu)\text{-}\, A \sin 2\pi(x / \lambda - \upsilon t + \phi) \qquad (23\text{-}2)$$

The velocity of propagation of light in a medium, *c′* (Eq. 23-3), depends upon both ε, the **dielectric constant** of the medium, and μ, the **magnetic permeability**.

$$\text{Velocity in a medium: } c' = c / (\varepsilon/\mu)\text{-} = c / n \qquad (23\text{-}3)$$

The **refractive index** of a medium relative to a vacuum is given the symbol *n*. It is the factor by which the velocity of light in a vacuum is diminished in a medium. It is a function of wavelength. For the 589 nm sodium line *n* is 1.00029 for air and 1.33 for water at 25°C.

The term ϕ in Eqs. 23-1 and 23-2 is a **phase factor**. Most light is called **incoherent** because ϕ varies for the many photons making up the beam. **Coherent light** produced by **lasers** contains photons all with the same phase relationship. If the electric vectors of all the photons in a beam of light are in the same plane (as will be the case for light emerging after passage through certain kinds of crystals), the light is called **plane polarized**. The direction of polarization is that of the electric vector E. **Circularly polarized light**, in which the electric vector rotates and traces out either a left-handed or right-handed helix, can also be generated. A beam of left-handed circularly polarized light, together with a comparable beam polarized in the right-handed direction, is equivalent to a beam of plane polarized light. Conversely, plane polarized light can be resolved mathematically into right- and left-handed circularly polarized components.

B. Absorption of Light

Absorption of light is fundamental to all aspects of photochemistry and provides the basis for absorption spectroscopy.[3,5–11] Light absorption is always **quantized**. It can take place only when the energy *h*ν of a quantum is equal to the difference in energy between two energy levels of the absorbing molecule (Eq. 23-4).

$$E_2 - E_1 = h\nu \qquad (23\text{-}4)$$

Not only must the difference $E_2 - E_1$ be correct for absorption but also there must always be a change in the dipole moment of the molecule in going from one energy level to another. Only when this is true can the electric field of the light wave interact with the molecule. A further limitation comes from the symmetry properties of the wave functions associated with each energy level. Quantum mechanical considerations show that transitions between certain energy levels are allowed, while others are forbidden. Consideration of such matters is beyond the scope of this book, but the student should be aware that the quantum mechanical selection rules that express this fact are an important determinant of light absorption.

Many types of light source are used in chemical measurements. Of great importance is the recent development of lasers that deliver very short pulses of light. Pulses as short as five femtoseconds (5 fs)[11a,11b] and even less[11c] are being utilized for very rapid spectroscopy and excitation of fluorescence. Structures are being determined by ultrafast electron diffraction[11d] or X-ray diffraction.[11e,11f] It takes 200 fs or more for a chemical bond to stretch and break during a reaction. The cleavage and formation of bonds during this time can be observed using 5-fs pulses. Lasers with pulses in the attosecond range may soon be used to observe movements of electrons.[11g]

1. Quantitative Measurement of Light Absorption, Spectroscopy

An absorption spectrum is a plot of some measure of the intensity of absorption as a function of wavelength or wave number. The **transmittance** of a sample held in a **cell** (or **cuvette**) is the fraction of incident light that is transmitted, i.e., transmittance = I / I_o where I_o is the intensity of light entering the sample and *I* is that of the emerging light. The transmittance is usually defined for a single wavelength, i.e., for **monochromatic** light. The absorbance (or optical density) is defined by Eq. 23-5, which also states the **Beer–Lambert law**. The length (in centimeters) of the

$$\text{Absorbance} = A = \log_{10} (I_o / I) = \varepsilon c l \qquad (23\text{-}5)$$

light path through the sample is *l*, *c* is the concentration in moles per liter, and ε is the **molar extinction coefficient** (molar absorptivity or molar absorption coefficient), whose units are liter mol^{-1} cm^{-1} (or $\mathbf{M^{-1}\ cm^{-1}}$). The reader can derive Eq. 23-5 by assuming that in a thin layer of thickness *dx* the number of light quanta absorbed is proportional to the number of absorbing molecules in the layer. Integration from *x* = 0 to *l* gives the Beer–Lambert law. Equation 23-5 generally holds very well for solutions containing single ionic or molecular forms. However, it is usually valid only for monochromatic light. Furthermore, making precise measurements of absorbance is not easy. At *A* = 1 only 10% of the incident light is transmitted, and the utmost care is required to obtain a value of *A* good to within ±0.05. At *A* = 2 only 1% of incident light is transmitted, and the value of *A* will be much less reliable. Very low absorbances are also difficult to measure. In view of the importance that spectrometry has played in biochemistry, it may seem surprising that the first reliable commercial laboratory ultraviolet–visible spectrophotometers became available in 1940 and the first commercial infrared spectrometer in 1942.[12]

2. The Energy Levels of Molecules

The energy of molecules consists of **kinetic** (translational), **rotational**, **vibrational**, and **electronic** components. The corresponding rotational, vibrational, and electronic energy levels are always quantized. Light quanta of wavelengths 0.2–20 mm (50–0.5 cm^{-1}; frequencies of 1.5 x 10^{12} to 1.5 x 10^{10} s^{-1}) with energies of 0.6–0.006 kJ/einstein are sufficient to excite molecules from a given rotational energy level to a higher one. Spectra in this "far infrared" or "microwave" region often consist of a closely spaced series of lines. For example, the rotational spectrum of gaseous HCl is a series of lines at 20.7 cm^{-1} intervals beginning at that wave number and reaching a maximum at about 186 cm^{-1} (54 µm). The energies involved in absorption of such light are far lower than energies of activation for common chemical reactions and lower than the average translational energy of molecules in solution at ordinary temperatures (3/2 k_BT or 3.7 kJ/mol at 25°C). However, they are still much higher than energies involved in the nuclear transitions of NMR spectra (Chapter 3). Compare 500 Hz for a proton resonating at 1 ppm in a 500 MHz NMR spectrometer with the 10^{10}–10^{12} s^{-1} frequencies of microwave spectra.

Vibrational energies range from about 6 to as much as 100 kJ mol^{-1} with corresponding wave numbers of ~500–8000 cm^{-1}. The resulting absorption spectra are in the infrared region. Excited electronic energy levels are ~120 – 200 kJ mol^{-1}, and the spectral transitions are at 10,000–100,000 cm^{-1} (1000–100 nm wavelengths) in the visible and ultraviolet region.

3. Infrared (IR) and Raman Spectra

Absorption in the near infrared region is dominated by changes in vibrational energy levels. A typical wave number is that of the "amide A" band at 3300 cm^{-1} (3.0 µm wavelength), approximately 10^{14} s^{-1}. First let us consider the stretching vibrations of a diatomic molecule. The two nuclei of the molecule can be thought of as connected with a spring. The energy of oscillation is approximately that of a harmonic oscillator. Application of quantum theory shows that the discrete energy levels that can be assumed by the oscillator are equally spaced. The difference between each pair of successive energy levels is $h\nu$, where ν is the frequency of light that must be absorbed to raise the energy from one level to the next. In the ground state (unexcited state) the molecule still possesses a **zero-point energy,** $E = 1/2\, h\nu_o$, equal to half the energy needed to induce a transition.

While the harmonic oscillator is a good approximation to the behavior of a molecule in the lower vibrational energy states, marked deviations occur at higher energies. At the lower energy levels the change in the distance between the atomic centers during the course of the vibration amounts to only ±10% or less, but as the energy becomes greater the bond stretches more and the motion becomes **anharmonic**. The energy states of molecules are often in the form of **Morse curves** in which energy is plotted against internuclear distance (Fig. 23-2). As the internuclear distance becomes very short, the energy rises steeply. As the bond is stretched, there comes a point at which addition of more energy ruptures the bond. A diatomic molecule will dissociate into atoms and more complex molecules into fragments. Vibrational energy levels can be portrayed as horizontal lines at appropriate heights on the Morse curve (Fig. 23-2).

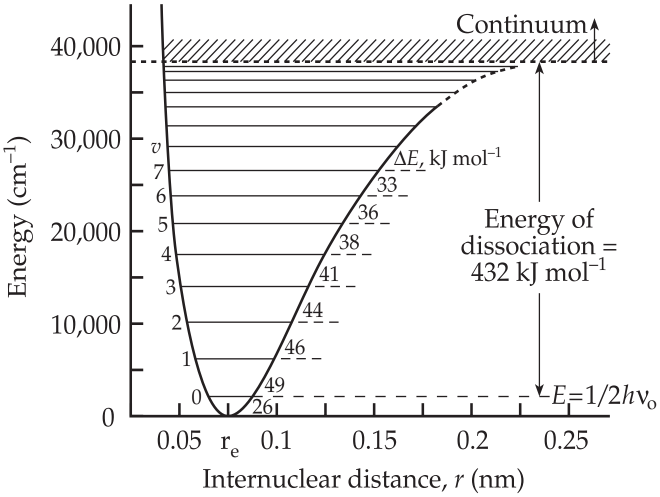


Figure 23-2 The potential energy of the hydrogen molecule as a function of internuclear distance, and the position of its vibrational energy levels. ΔE values are energy differences between successive levels; v designates vibrational quantum numbers. Adapted from Calvert and Pitts,[2] p. 135.

Because there are many rotational energy levels corresponding to each vibrational level, IR spectra contain absorption bands resulting from simultaneous changes in both the vibrational and rotational energy levels of molecules. Instead of single peaks corresponding to single transitions in vibrational energy, progressions of sharp bands at closely spaced intervals are observed. An example is provided by the band corresponding to the stretching frequency of the H–Cl bond in gaseous HCl at 2886 cm^{-1} (3.46 μm). There is actually no band at this wave number but a series of almost equally spaced bands on either side of the fundamental frequency from ~2600 to ~3100 cm^{-1} at intervals of ~21 cm^{-1}, i.e., the wave number of the rotational frequency seen in the microwave spectrum (Herzberg,[8] p. 55). The effect is to broaden the band as seen in a low-resolution spectrum. This is only one cause of the broadening of IR bands in solution. Another cause is interaction with solvent to provide a heterogeneity in the environments of the absorbing molecules.

The IR spectra of diatomic molecules are relatively easy to interpret, but for more complex substances the infrared absorption bands often cannot be associated with individual chemical bonds. Instead, they correspond to the **fundamental vibrations** (normal vibrations) of the *molecule*. Fundamental vibrations are those in which the center of gravity does not change. For a molecule containing n atoms, there are $3n - 6$ such vibrations. They are sometimes dominated by a vibration of a single bond, but often involve synchronous motion of many atoms. The fundamental vibrations of a molecule are described by such words as *stretching*, *bending* (in-plane and out-of-plane), *twisting*, and *deformation*. Rarely are all $3n - 6$ bands seen in an infrared spectrum. Some of the vibrations, e.g., the symmetric stretching of the linear CO_2 molecule, are not accompanied by any change in dipole moment, while other bands may simply be too weak to be observed clearly.

Vibrations involving many atoms in a molecule, i.e., **skeletal vibrations**, are often found in the region of 700–1400 cm^{-1} (14–7 μm). Vibrational frequencies that are dominantly those of individual functional groups can often be identified in the range 1000–5000 cm^{-1} (10–2 μm). Examples of the latter are the stretching frequencies of C–H, N–H, and O–H bonds, which have wave numbers of ~2900, 3300, and 3600 cm^{-1}, respectively. The energy (and frequency) of the vibrations increases as the difference in electronegativity between the two atoms increases. When a bond connects two heavier atoms, the frequency is lower, e.g., the wave number for C–O in a primary alcohol is ~1053 cm^{-1}. For a double bond it increases; for C=O it is ~1700 cm^{-1}. This C=O stretching frequency usually gives rise to one of the strongest bands observed in IR spectra. Hydrogen bonding has a strong and characteristic effect. Thus, the O–H frequency at ~3600 cm^{-1} is decreased to ~3500 cm^{-1} by hydrogen bonding.

Theory predicts that for a harmonic oscillator only a change from one vibrational energy level to the next higher is allowed, but for anharmonic oscillators weaker transitions to higher vibrational energy levels can occur. The resulting "overtones" are found at approximate multiples of the frequency of the fundamental. Combination frequencies representing sums and differences of frequencies of individual IR bands may also be seen. The intensities of these bands are low, but their presence at relatively high energies in the near IR region (4000–12,500 cm^{-1}) means that they may be easier to observe than the fundamental frequencies in the more crowded IR region.[12a] Development of the very sensitive **Fourier-transform infrared spectroscopy** (FTIR) has made it possible to record the complex IR spectra of macromolecules in dilute aqueous solutions rapidly. The water spectrum is substracted digitally.[13–16] FTIR has been utilized to study amide groups in peptides, carboxyl groups in proteins,[17] conformations of sugar rings in DNA,[18] and the ionization state of phosphate groups.[19] New computational methods involving use of two-dimensional representations provide simplifications in interpretation of IR spectra.[19a] Another variant is **total reflection FTIR**, a technique that records spectra of thin films and has permitted the recording of transient changes in protein spectra with microsecond time resolution.[20,21]

Vibrational frequencies of amide groups. The IR absorption bands of amide groups, which are present in both proteins and in the purine and pyrimidine bases, have attracted a great deal of attention.[13,22–23a] The **amide I** band at ~1680 cm^{-1} is associated with an in-plane normal mode of vibration that involves primarily C=O stretching. (The band is designated I′ if the N–H has been exchanged to form N–^{2}H.)[16] The **amide II** band at ~1500 $^{-1}$ and the **amide III** band at ~1250 cm^{-1} both arise from in-plane modes that involve N–H bending, while the higher frequency **amide A** band at ~3450 cm^{-1} involves N–H stretching. It is shifted to ~3300 cm^{-1} when the N–H is hydrogen bonded. Examples of IR spectra of proteins, including the amide bands A, I, and II, are shown in Fig. 23-3. The band shapes are complex. Those of the amide bands I and III depend upon the conformation of the peptide chain. For example, (Fig 23-3B) amide groups in α helices have an amide I band about 20 cm^{-1} higher than do those in β structures.

If peptide chains can be oriented in a regular fashion, it may be useful to measure **infrared linear dichroism**.[24,25] Absorption spectra are recorded by passing plane polarized light through the protein in two mutually perpendicular directions, with the electric vector either parallel to the peptide chains or perpendicular to the chains. Such a pair of spectra is shown in Fig. 23-3A for oriented fibrils of insulin. In this instance, the insulin molecules are thought to assume a β conformation and to be stacked in such a way that they extend transverse to the fibril axis (a cross-β structure). When the electric vector is parallel to the fibril axis, it is perpendicular to the peptide chains. Since the amide I band is dominated by a carbonyl stretching motion that is perpendicular to the peptide chains in the β structure, this band is enhanced when the electric vector is also perpendicular to the peptide chains and is diminished when the electric vector is parallel to the peptide chains (perpendicular to the fibril axis, Fig. 23-3A). The same is true of the amide A band which is dominated by an N–H stretch. On the other hand, the dichroism of the amide II band is the opposite because it tends to be dominated by an N–H bending, which is in the plane of the peptide group but is longitudinal in direction. In **isotope-edited FTIR**, heavy atoms such as ^{13}C are introduced to shift IR bands and assist in their identification. The method can be combined with measurement of linear dichroism of oriented peptides.[25]

The loss of the amide II band in D_2O is one of the major tools for studying protein dynamics.[13,26] (see Chapter 3, Section I,5). In some cases the four main secondary structures, α helix, β sheet, β turn, and random coil, can be distinguished.[27] The amide bands of pyrimidines can also be observed in IR spectra of pyrimidines. Figure 23-4A shows the spectrum of 1-methyluracil in H_2O and also in D_2O. Notice that the amide II band is totally lacking in D_2O. The same figure also shows the IR spectrum of 1-methyluracil containing ^{18}O in the 4 position. The shift of 9 cm^{-1} in the amide II position is part of the evidence that the NH bending vibration is extensively coupled to C=O and C=C stretching modes.

Raman spectra. In a collision between a photon and a molecule, the photon may undergo **elastic collision** in which the photon loses no energy but changes its direction of travel. Such scattering is known as **Rayleigh scattering** and forms the basis for a method of molecular mass determination. Sometimes **inelastic** collisions occur in which both the

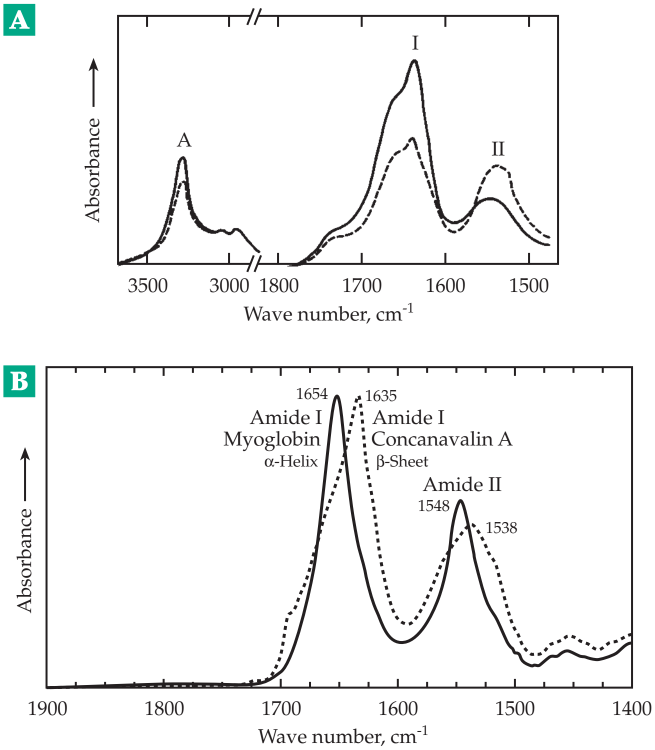


Figure 23-3 Infrared absorbance spectra of the amide regions of proteins. (A) Spectra of insulin fibrils illustrating dichroism. Solid line, electric vector parallel to fibril axis; broken line, electric vector perpendicular to fibril axis. From Burke and Rougvie.[24] Courtesy of Malcolm Rougvie. See also Box 29-E. (B) Fourier transform infrared (FTIR) spectra of two soluble proteins in aqueous solution obtained after subtraction of the background H_2O absorption. The spectrum of myoglobin, a predominantly α-helical protein, is shown as a continuous line. That of concanavalin A, a predominantly β-sheet containing protein, is shown as a broken line. From Haris and Chapman.[14] Courtesy of Dennis Chapman.

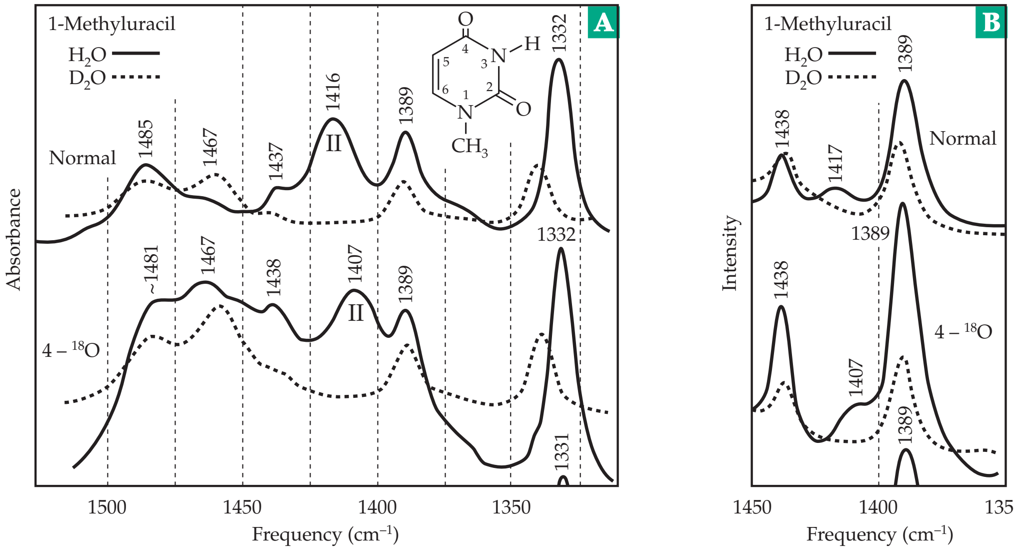


Figure 23-4 (A) Infrared and (B) Raman spectra of 1-methyluracil in H_2O and D_2O. Spectra for normal 1-methyluracil and for the specific isotopic derivatives with ^{18}O in the 4 position are shown. From Miles *et al.*[28]

molecule and the photon undergo changes in energy. Since such changes must be quantized and involve vibrational and rotational levels of the molecule, the spectrum of the scattered light (Raman spectrum) contains much of the same information as an ordinary IR spectrum. However, the selection rules are not the same. Some transitions are "infrared active" and others are "Raman active." IR-active transitions can occur only when the dipole moment varies with time as a molecule vibrates, whereas Raman-active transitions require that the polarizability change with time. For this reason, it, s useful to measure both IR and Raman spectra on the same sample. Until recently, Raman spectroscopy was not used much in biochemistry because of the low intensity of the scattered light. However, with laser excitation the technique is practical[7,13,23,28,29] and can be applied to aqueous or nonaqueous solutions, and to solid or dispersed forms of macromolecules.

Both amide I and amide III bands are seen in Raman spectra of proteins.[30] Lippert *et al.* devised the following method for estimating the fractions of α-helix, β sheet, and random coil conformations in proteins.[31] The amide I Raman bands are recorded at 1632 and 1660 cm^{-1} in D_2O (amide I′). The amide III band, which is weak in D_2O, is measured at 1240 cm^{-1} in H_2O. The intensities of the three bands relative to the intensity of an internal standard (the 1448 cm^{-1} CH_2 deformation) are related to those of standard poly-L-lysine in known conformations. See also Craig and Gaber.[32] The Raman spectrum of 2-methyluracil is shown in Fig. 23-4B. Note the low intensity of the amide II band relative to that of the amide I band, a characteristic of Raman spectra. Linear dichroism observed by polarized Raman microspectrophotometry has provided information about orientation of indole rings of tryptophan in filamentous virus particles.[33,33a]

In **resonance Raman spectroscopy**[34–37] a laser beam of a wavelength that is absorbed in an electronic transition is used. The scattered light is often strongly enhanced at frequencies differing from that of the laser by Raman frequencies of groups within the chromophore or of groups in another molecule adjacent to the chromophore. The resonance effect not only increases the sensitivity of Raman spectroscopy but also allows a person to study specifically the vibrational spectrum of a selected aromatic group or other structure within a macromolecule. Problems associated with the technique are fluorescence, which may be 10^6 times as strong as the Raman emission, and photochemical damage from the intense laser beam. Fluorescence is often quenched with KI (see Section C,1).

If the exciting laser has a frequency ν_0 and the frequency of a vibrationally excitation in a molecule is ν_1 the Raman spectrum will contain a pair of bands,

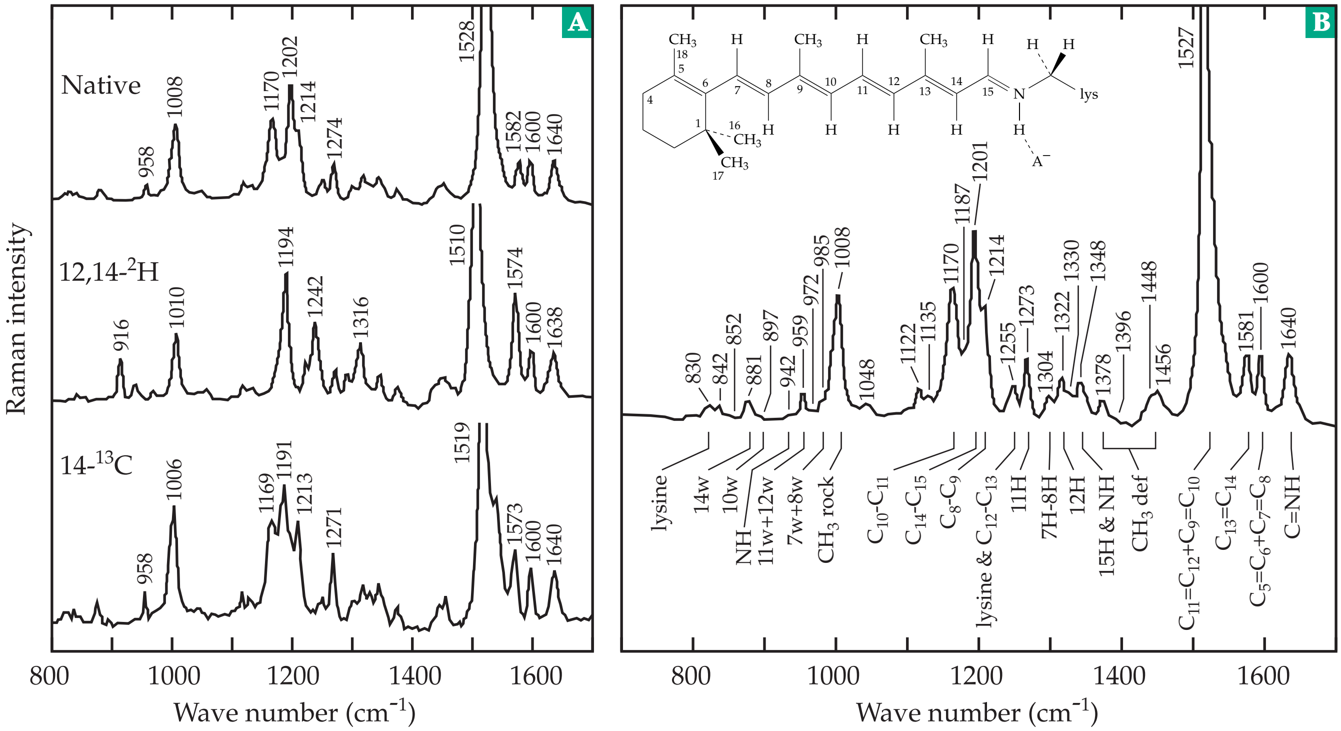


Figure 23-5 Resonance Raman spectra. (A) Of the retinaldehyde-containing bacteriorhodopsin bR_{568} (see Fig. 23-45) and its 12,14-^{2}H and 14-^{13}C isotopic derivatives. (B) Of bR_{568} labeled with the dominant internal coordinates that contribute to the normal modes. From Lugtenburg *et al.*[37]

the stronger one or “stokes” band of energy $h\ (\nu_0 - \nu_1)$ and a weaker one or “anti-stokes” band of $h\ (\nu_0 + \nu_1)$. Special techniques such as **coherent anti-stokes** Raman scattering provide a means of getting around the fluorescence problem.[13,36,38] Raman spectroscopy is also possible with excitation of ultraviolet absorption bands. It can be applied to peptides,[39–41] tryptophan or tyrosine residues of proteins,[42] nucleic acid bases,[43–45] heme proteins,[46–48] other metalloproteins,[49] flavin coenzymes,[50–52] pyridoxal phosphate,[53,54] flavoproteins,[54a] carotenoid-containing proteins,[55] and to substrates undergoing cleavage in the active site of papain.[56] Resonance Raman spectroscopy is very useful in the study of adducts of CO, NO, and O_2 with heme proteins because coupled vibrational modes of both the porphyrin rings and axial ligands can be observed.[56a,b] Resonance Raman spectra are strongly amplified for samples adsorbed to specially prepared colloidal silver particles.[57] Examples of resonance Raman spectra are shown in Fig. 23-5.

4. Electronic Spectra

Biochemists make extensive use of spectroscopy in the ultraviolet (UV) and visible range. Visible light begins at the red end at ~12,000 cm^{-1} (800 nm) and extends to 25,000 cm^{-1} (400 nm). The ultraviolet range begins at this point and extends upward, the upper limit accessible to laboratory spectrophotometers being ~55,000 cm^{-1} (180 nm). The energies covered in the visible–UV range are from ~140 to ~660 kJ/mol. The latter is greater than the bond energy of all but the strongest double and triple bonds (Table 6-7). It is understandable that UV light is effective in inducing photochemical reactions. Even the lower energy red light, which is used by plants in photosynthesis, contains enough energy per einstein to make it feasible to generate ATP, to reduce $NADP^+$, and to carry out other photochemical processes. Although the energies of light absorbed in electronic transitions are large, the geometry of molecules in the excited states is often only slightly altered from that in the ground state. The amount of vibration is increased, and the molecule usually expands moderately in one or more dimensions.

The significance of light absorption in biochemical studies lies in the great sensitivity of electronic energy levels of molecules to their immediate environment and to the fact that spectrophotometers are precise and sensitive. The related measurements of circular dichroism and fluorescence also have widespread utility for study of proteins, nucleic acids, coenzymes, and many other biochemical substances that contain intensely absorbing groups or **chromophores**.[58]

Shapes of absorption bands. Electronic absorption bands are usually quite broad, the width of the band at half-height often being 3000–4000 cm^{-1}. The breadth arises largely from the coupling of electronic excitation to changes in the vibrational and rotational energy levels. Inhomogeneity of environments in the solvent also contributes. Shapes of absorption bands are to a large extent determined by the **Franck–Condon principle**, which states that no significant change in the positions of the atomic nuclei of the molecule occurs during the time of the electronic transition. Since the frequency of light absorbed during these transitions is ~10^{15} to $10^{16}\ s^{-1}$, the absorption of light energy occurs within 10^{-15} to 10^{-16} s, the time equivalent to the passage of one wavelength of light. During this period the vibrational motions of the nuclei are almost insignificant because of the much lower frequencies of vibration. Two types of potential energy curves for excited states of molecules are shown in Fig. 23-6.[2] In the first the geometry of the molecule is little changed between ground state and excited state. At room temperature most molecules are in the lowest energy states of at least the most energetic of the various vibrational modes of the molecule ($3/2\ k_B T$ ~300 cm^{-1}). Therefore, the most probable transitions occur from the lowest vibrational states of the ground electronic states. The most probable internuclear distance for a molecule in the ground state is the equilibrium distance r_e (Fig. 23-2). Since that distance is the same in all of the vibrational levels of the electronically excited state, transitions to any of these states may occur. The transition to the first vibrational level of the excited state is most likely. The result is an absorption spectrum in which the sharp band representing the “0–0 transition” is most intense and in which there are progressively weaker bands corresponding to the 0–1, 0–2, 0–3, etc., transitions (Fig. 23-6A). Many organic dyes with long series of conjugated double bonds have spectra of this type.[59,60]

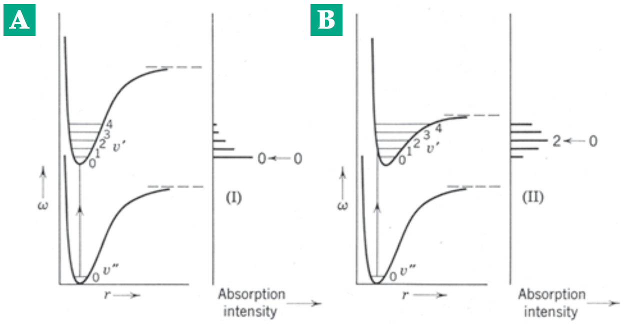


Figure 23-6 Typical potential energy curves for two types of band spectra: (A) For a transition in which the equilibrium internuclear distances r_e are about equal in the ground and excited states. (B) For a transition in which r_e' (excited state) > r_e (ground state). From Calvert and Pitts,[2] p. 179.

A second type of spectrum is illustrated in Fig. 23-6B. In this instance, the molecule has expanded in the excited state, and r_e is greater than in the ground state. The Franck–Condon principle suggests that a transition is likely only to those vibrational levels of the excited state in which the internuclear distance is compressed for a significant fraction of the time, approximately to that of r_e in the ground state. Examination of Fig. 23-6B explains why the resulting absorption spectra tend to have weak 0–0 bands and stronger bands corresponding to transitions to higher levels.

For real spectra of polyatomic molecules the situation is more complex. Some molecules in the ground state do occupy higher vibrational levels of the less energetic modes. Therefore, there will be weaker lines, some of which lie on the low-energy side of the 0–0 transition. Since in polyatomic molecules there are several normal modes of vibration, there will be other progressions of absorption bands paralleling those shown in Fig. 23-6 and filling in the valleys between them. All of the bands are broadened by rotational coupling and by interactions with solvent.

An example of a molecule giving a spectrum of the type shown in Fig. 23-6B is toluene. The vapor phase spectrum contains a large number of sharp lines, some of which can be seen in the low-resolution spectrum of Fig. 23-7. Several progressions can be identified.[61] One begins with the intense 0–0 line at 37.48 x 10^3 cm^{-1} and in which spacing of ~930 cm^{-1} between lines corresponds to a vibration causing symmetric expansion of the ring (ring breathing frequency), a frequency that can also be observed in the infrared spectrum. Other progressions beginning at the 0–0 line involve additional modes of vibration with frequencies (in the excited electronic state) of 460, 520, and 1190 cm^{-1}. Additional weaker bands are "buried" in the valleys in Fig. 23-6. When the spectrum of toluene is measured in solution, the sharp lines are broadened, but there are still indications of vibrational structure.

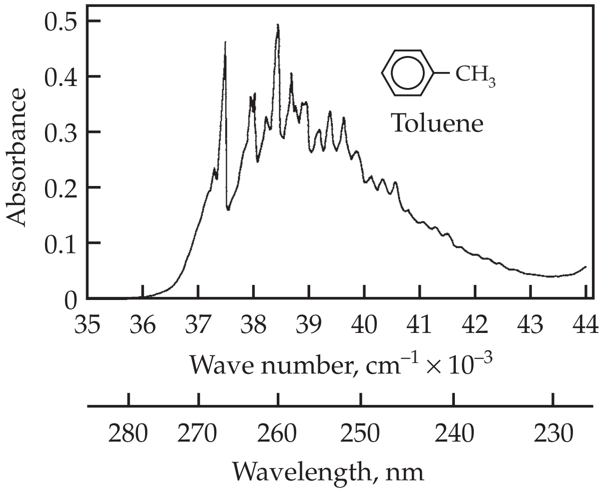


Figure 23-7 The spectrum of the first electronic transition of toluene vapor at low resolution. Cary 1501 spectrophotometer.

Gaussian curves (normal distribution functions) can sometimes be used to describe the shape of the overall envelope of the many vibrationally induced subbands that make up one electronic absorption band, e.g., for the absorption spectrum of the copper-containing blue protein of *Pseudomonas* (Fig. 23-8) Gaussian bands are appropriate. They permit resolution of the spectrum into components representing individual electronic transitions. Each transition is described by a **peak position**, **height** (molar extinction coefficient), and **width** (as measured at the half-height, in cm^{-1}). However, most absorption bands of organic compounds are not symmetric but are skewed

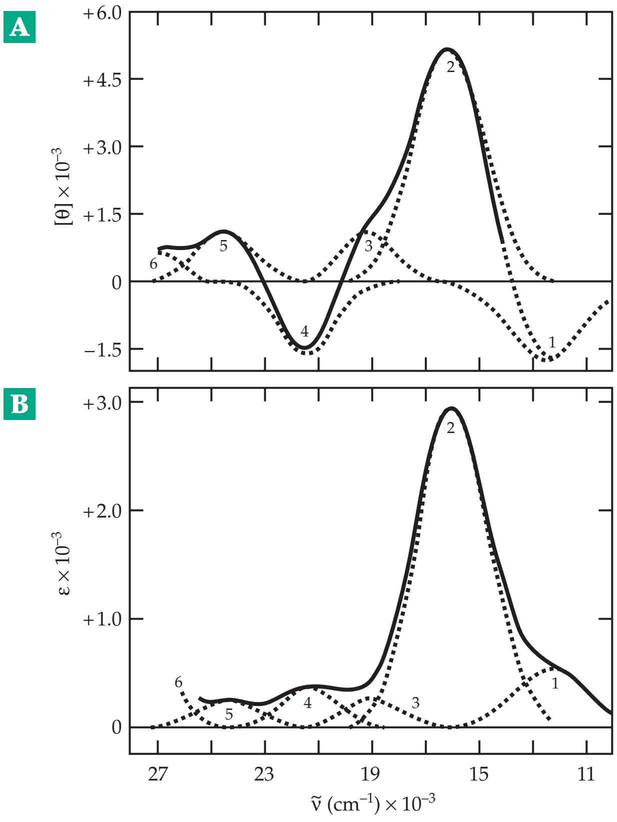


Figure 23-8 Resolution of the visible circular dichroism (ellipticity) spectrum (A) and absorption spectrum (B) of the *Pseudomonas* blue protein into series of overlapping Gaussian bands (—). The numbers 1 to 6 refer to bands of identical position and width in both spectra. Absorption envelopes resulting from the sum of the set of overlapping Gaussian bands (—) correspond within the error of the measurement to the experimental spectra. The *dashed part* of the CD envelope above 700 nm was completed by a curve fitter with the use of a band in the position of *band 1* of the absorption spectrum. From Tang *et al.*[68]

toward the high-energy side. It is best to fit such bands with a skewed function such as the **log normal** distribution curve.[62–64] In addition to position, height, and width, a fourth parameter provides a measure of **skewness**. Computer-assisted fitting with log normal curves gives precise values for the positions, widths, and intensities. In general, the peak position is somewhat to the high-energy side of the 0–0 transition.

Absorption spectra plotted as a linear function of wavelength are sometimes fitted with Gaussian curves. However, Gaussian curves only occasionally give a good fit for such spectra, and it is undesirable to measure bandwidths in nanometers. It is wave number that is proportional to energy. Spectral bands tend to have similar widths across the visible–ultraviolet range when plotted against wave number but not when ploted against wavelength. Another approach to the quantitative description of spectra is to fit the major progressions of vibrational subbands with series of narrow Gaussian curves.[65–67]

Classification of electronic transitions. The intense 600-nm absorption band of the copper blue protein in Fig. 23-8 is attributed to a *d–d* transition of an electron in the metal ion from one *d* orbital to another of higher energy.[68] The intensity is thought to arise from transfer of an electron from a cysteine thiolate to the copper (p. 883). The electronic transitions in most organic molecules are of a different type. Transitions lying at energies <55,000 cm^{-1} are classified as either $\pi-\pi^*$ or $n-\pi^*$. In the $\pi-\pi^*$ transitions an electron is moved from a bonding π molecular orbital to an antibonding (π^*) orbital. Such a transition is present in ethylene at 61,540 cm^{-1} (162.5 nm) with a maximum molar extinction coefficient ε_{max} of ~15,000 M^{-1} cm^{-1}. An $n-\pi^*$ transition results from the raising of an electron in an unshared pair of an oxygen or nitrogen atom into a π^* antibonding orbital. These transitions are invariably weak. For example, acetone in H_2O shows an $n-\pi^*$ transition at 37,740 cm^{-1} (265 nm). The value of ε_{max} is ~240 and the width is about 6400 cm^{-1}. A characteristic of $n-\pi^*$ transitions is a strong shift to lower energies as the compound is moved from water into less polar solvents. For example, the peak of the acetone band lies at 36,920 cm^{-1} in methanol and at 35,970 cm^{-1} (278 nm) in hexane. Such a solvent shift is often taken as diagnostic of an $n-\pi^*$ transition, and it is often stated that the $\pi-\pi^*$ bands shift in the opposite direction upon change of solvent character. However, the latter is not true for many of the polar chromophores found in biochemical substances. Thus, the $\pi-\pi^*$ bands of tyrosine also shift to lower energies when the molecule is moved from water into hexane. However, the magnitude of the shift is much less than for the $n-\pi^*$ band of acetone.

A molecule can have several different excited states of increasing energies. In benzene and its derivatives there are three easily detectable $\pi-\pi^*$ transitions (see Fig. 3-13). The first is a weak band centered at ~ 260 nm in toluene (Fig. 23-7) and ~275 nm in tyrosine (Fig. 3-13) with $\varepsilon = 10^2$ to 10^3. The second is a band at a higher frequency (at 1.35 0.10 times the frequency of the first band) with ε_{max} often as high as 10^4. The third band is found at still higher energies with ε_{max} reaching 5 x 10^4. The excited-state energy levels represented by these transitions were labeled ^1L_b, ^1L_a, and ^1B_a by Platt. Other authors described the levels in terms of the symmetries of the molecular orbitals, the ground state being $^1A_{1g}$, and the three excited states $^1B_{2u}$, $^1B_{1u}$, and $^1E_{1u}$. In these symbols the superscript 1 indicates that the excited states are **singlet** in nature; that is, the electrons remain paired in the excited states. Absorption of visible and ultraviolet light almost always leads to singlet excited states initially. For more complex ring systems the number of possible electronic transitions increases, but attempts are often made to relate these transitions back to those of benzene.

The intensities of electronic transitions vary greatly. The area ($\mathcal{A}$) under the absorption band, when ε is plotted against wave number $\bar{\nu}$, is directly proportional (Eq. 23-6) to a dimensionless quantity called the **oscillator strength** *f*.

$$f = \frac{2.303\, m_e c^2}{\pi N e^2} F\mathcal{A} = 4.32 \times 10^{-9}\, F\mathcal{A} \tag{23-6}$$

In this equation m_e and e are the mass and charge of the electron, c is the velocity of light, N is Avogadro's number, and $\mathcal{A}$ is the area in a plot of ε vs ν in cm^{-1}; F is a dimensionless correction factor that is related to the refractive index of the medium and is near unity for aqueous solutions. If the area is approximated as that of a triangle of height ε_{max} and width (at half-height) W, we find that for a typical absorption band of $\varepsilon_{max} = 10^4$ and $W = 3000\ cm^{-1}$, $f = 0.13$.

The oscillator strength is related to the probability of a transition and can become approximately 1 only for the strongest electronic transitions. However, it is rarely this high. For example, the oscillator strength is ~10^{-4} for Cu^{2+} and ~2 x 10^{-3} for the toluene absorption band shown in Fig. 23-7. The low intensity of absorption bands of benzene derivatives is related to the fact that these transitions are quantum mechanically forbidden for a completely symmetric molecule. It is only because of coupling with asymmetric vibrations of the ring that the ^1L_b transition of benzene becomes weakly allowed. In the benzene spectrum the 0–0 transition is completely absent, and only those progressions involving uptake of an additional 520 cm^{-1} of a nonsymmetric vibrational energy are observed. In the case of toluene and phenylalanine, the asymmetry of the ring introduced by the substituents permits the 0–0 transition to occur and leads to a higher oscillator strength than that observed with benzene. The ^1L_a transition of benzenoid derivatives is also partially forbidden by selection rules, and only the third band begins to approach an oscillator strength of one.

Use of plane polarized light. The intensity of a spectral transition is directly related to the **transition dipole moment** (or simply the transition moment), a vector quantity that depends upon the dipole moments of the ground and excited states. For aromatic ring systems, the transition dipole moments of the $\pi-\pi^*$ transitions lie in the plane of the ring. However, both the directions and intensities for different $\pi-\pi^*$ transitions within a molecule vary.

The transition moment has a dimension of length (usually given in angstroms) and can be thought of as a measure of the extent of the charge migration during the transition. Light is absorbed best when the directions of polarization (i.e., of the electric vector of the light) and of the transition moment coincide. This fact can easily be verified by light absorption measurements on crystals. As with infrared spectra of oriented peptide chains (Fig. 23-3), the electronic spectra of crystals display a distinct dichroism. Crystals of coenzyme-containing proteins (Fig. 23-9) are very appropriate for spectroscopy with polarized light because the chromophores are spaced far enough apart to avoid electronic interaction and have absorbances low enough to record with crystals of the order of 0.1 mm thickness.[69,70]

In contrast to $\pi-\pi^*$ transitions, the $n-\pi^*$ transitions of heterocyclic compounds and carbonyl-containing rings are often polarized in a direction perpendicular to the plane of the ring. Linear dichroism of cytosine, adenine, and other nucleic acid bases has been measured on single crystals and in partially oriented polymer films.[71] Magnetically induced linear dichroism provides a new tool for study of metalloproteins.[72]

Relationship of absorption positions and intensity to structures. While quantum mechanical calculations permit prediction of the correct number and approximate positions of absorption bands, they are imprecise. For this reason, electronic spectroscopy also relies upon a combination of empirical rules and atlases of spectra that can be used for comparison purposes.[74–76] The following may help to orient the student. The position of an absorption band shifts **bathochromically** (to longer wavelength, lower energy) when the number of conjugated double bonds increases. Thus, **butadiene** absorbs at 46,100 cm^{-1} (217 nm) vs the 61,500 cm^{-1} of ethylene. As the number of double bonds increases further, the bathochromic shifts become progressively smaller (but remain more nearly constant in terms of wavelength than wave number). For **lycopene** (Fig. 23-10) with 11

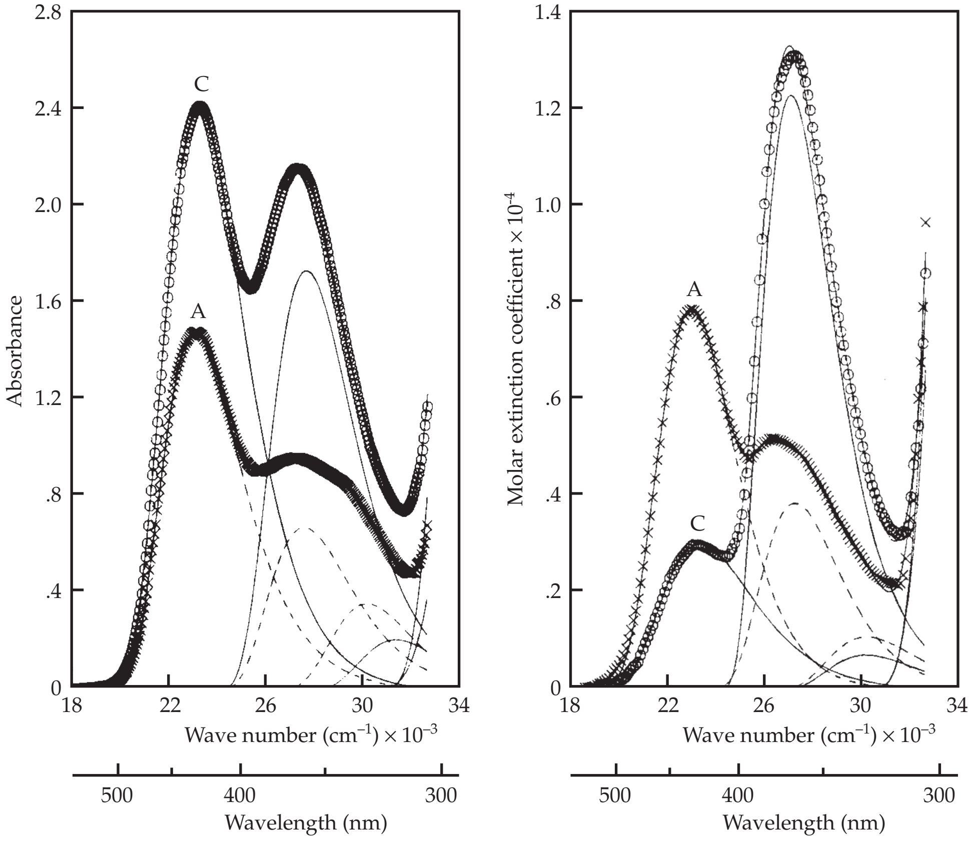


Figure 23-9 Polarized absorption spectra of orthorhombic crystals of cytosolic aspartate aminotransferase. The light beam passed through the crystals along the *b* axis with the plane of polarization parallel to the *a* axis (A) or the *c* axis (C). Left, native enzyme at pH 5.4; right, enzyme soaked with 300 mM 2-methylaspartate at pH 5.9. The band at ~430 nm represents the low pH protonated Schiff base form of the enzyme. Upon soaking with 2-methylaspartate the coenzyme rotates ~30° to form a Schiff base with this quasisubstrate. The result is a large change in the *c*/*a* polarization ratio. The ~364 nm band in the complex represents the free enzyme active site in the second subunit of the dimeric enzyme.[70,73] Courtesy of C. M. Metzler.

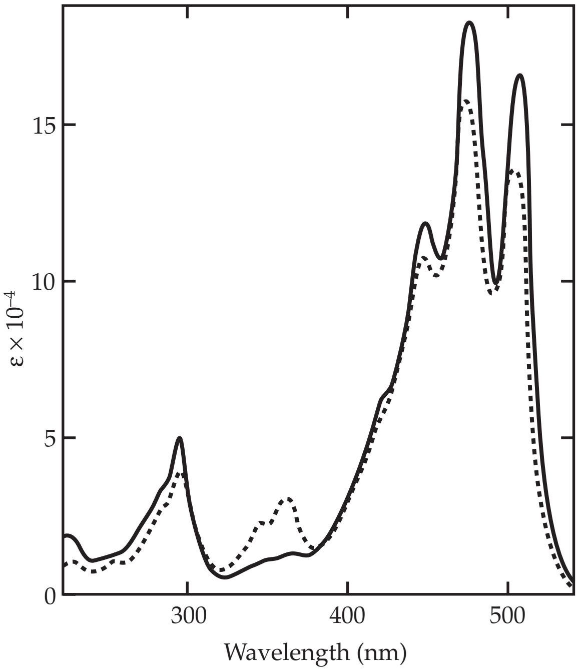


Figure 23-10 The absorption spectrum of lycopene (plotted vs wavelength). Note the vibrational structure, which has a spacing of ~1200–1500 cm^{-1}. The solid line is for all-translycopene while the dashed line is that of the sample after refluxing 45 min in the dark. The new peak at ~360 nm arises from isomers containing some cis double bonds.

conjugated double bonds the absorption band is located at 21,300 cm^{-1} and displays distinct vibrational structure (Fig. 23-10). Certain ring molecules such as the porphyrins and chlorophylls have spectra that can be related back to those of the linear polyenes. Note (Fig. 16-7) that the porphyrin α and β bands represent vibrational structure of a single electronic transition, whereas the intense Soret band results from a different transition.

Substituted benzenes almost invariably absorb at lower energies than the parent hydrocarbon. The stronger the electron withdrawing or donating ability of the substituent, the larger the bathochromic shift. The magnitude of the shift has been correlated with the Hammett σ constants. Thus, the first absorption band of tyrosine in water is shifted 2600 cm^{-1} toward the red from that of benzene, while that of the dissociated tyrosine anion is shifted 4700 cm^{-1}, very roughly in proportion to the σ_p values of Box 6-C. Especially large shifts are observed when functional groups of opposite types (that is, an electron donating group vs an electron accepting group) are both present in the same ring. The effects of ortho and meta substituent pairs are closely similar (in contrast to the differing electronic effects of ortho and meta pairs in chemical reactivity). Substituent pairs in para positions yield somewhat different spectral shifts. When there are more than two substituents, the two strongest groups often dominate in determining the character of the spectrum. Useful empirical rules have been developed.[77,78]

Spectra of proteins and nucleic acids. Most proteins have a strong light absorption band at 280 nm (35,700 cm^{-1}) which arises from the aromatic amino acids tryptophan, tyrosine, and phenylalanine (Fig. 3-14). The spectrum of phenylalanine resembles that of toluene (Fig. 23-7),[61] whose 0–0 band comes at 37.32 x 10^3 cm^{-1}. The vibrational structure of phenylalanine can be seen readily in the spectra of many proteins (e.g., see Fig. 23-11A). The spectrum of tyrosine is also similar (Fig. 3-13), but the 0–0 peak is shifted to a lower energy of ~35,500 cm^{-1} (in water). Progressions with spacings of 1200 and 800 cm^{-1} are prominent.[79] The low-energy band of tryptophan consists of two overlapping transitions ^1L_a and ^1L_b.[65] The ^1L_b transition has well-resolved vibrational subbands, whereas those of the ^1L_a transition are more diffuse. Tryptophan derivatives in hydrocarbon solvents show 0–0 bands for both of these transitions at approximately 289.5 nm (34,540 cm^{-1}). However, within proteins the ^1L_a band may be shifted 3–10 nm (up to 1100 cm^{-1}) toward lower energies, probably as a result of hydrogen bonding to other groups in the protein. The largest shifts can occur when the NH group of the indole ring is hydrogen bonded to COO^-, a ring nitrogen of histidine, or a carbonyl group of amides.[80] In an aqueous medium the ^1L_b band of tryptophan is shifted to higher energies and the ^1L_a band to lower energies than in a hydrocarbon solvent.

In addition to the three aromatic amino acids, disulfide bonds absorb in the near ultraviolet region as indicated in Fig. 3-14. Since the absorption characteristics depend upon the dihedral angles in the disulfide bridges, it is difficult to accurately evaluate the contribution of this chromophore to the 280-nm band.

Tyrosine, tryptophan, and phenylalanine all have additional transitions in the high-energy UV region of the spectrum (Fig. 3-13). Even more intense are the absorption bands of the amide groups, which become significant above 45,000 cm^{-1}.[81] These include a weak $n-\pi^*$ transition at ~45,500 cm^{-1} (210 nm) overlapped by a strong $\pi-\pi^*$ transition at ~52,000 cm^{-1} (192 nm).[82] Histidine also has absorption bands in this region.

As with polypeptides, the light absorption properties of polynucleotides reflect those of the individual components. The spectra of the purine and pyrimidine bases as ribonucleosides are shown in Fig. 5-5. The number of individual electronic transitions and their origins are not immediately obvious, but many measurements in solutions and in crystals, as well as theoretical computations,[7,83,84] have been made. Cytosine has $\pi-\pi^*$ transitions at ~275, 230, 200, and

185 nm,[83] the two highest energy bands being overlapped. Adenine derivatives have seven $\pi-\pi^*$ transitions.[71] Spectra of flavins contain at least four intense transitions (Fig. 15-8).[85]

Whereas proteins have their low energy absorption band at ~280 nm, polynucleotides typically have maxima at ~260 nm (38,500 cm^{-1}). A phenomenon of particular importance in the study of nucleic acids is the **hypochromic effect**. In a denatured polynucleotide the absorption is approximately the sum of that of the individual components. However, when a double helical structure is formed and the bases are stacked together, there is as much as a 34% depression in the absorbance at 260 nm. This provides the basis for optical measurement of DNA melting curves (Fig. 5-45).[45,86] The physical basis for the hypochromic effect is found in dipole–dipole interactions between the closely stacked base pairs.[7,86,87]

Difference spectra and derivative spectra.

Changes in light-absorbing properties of proteins and nucleic acids are often measured as a function of some quantity such as pH, temperature, ionic environment, or the presence or absence of another interacting molecule. The induced changes in the spectrum are small but can be seen if the *difference* between the two spectra, one "unperturbed" and the other in the presence of some "perturbant," is recorded. The perturbant might be an additional reagent, an altered solvent (e.g., with added glycerol, D_2O), a change in pH, or temperature. The difference spectrum shown in Fig. 23-11B arises from the binding of an inhibitor succinate together with a substrate carbamoyl phosphate to the catalytic subunit of aspartate carbamoyltransferase (Fig. 7-20).[88] The difference spectrum appears as a pair of peaks and a valley in the aromatic amino acid region. With proper interpretation (caution!) difference spectra can be used to infer something about the change in environment of aromatic amino acids in a protein.[89]

Difference spectra are usually recorded by placing the unperturbed spectrum in the *reference* light beam of a spectrophotometer and the perturbed solution in the *sample* beam in carefully matched cuvettes. However, the spectrum shown in Fig. 23-11B was obtained by recording the two spectra independently and subtracting them with the aid of a computer. The same data have been treated in another way by fitting two log normal curves (p. 1283) to the absorption bands and plotting the differences between the mathematically

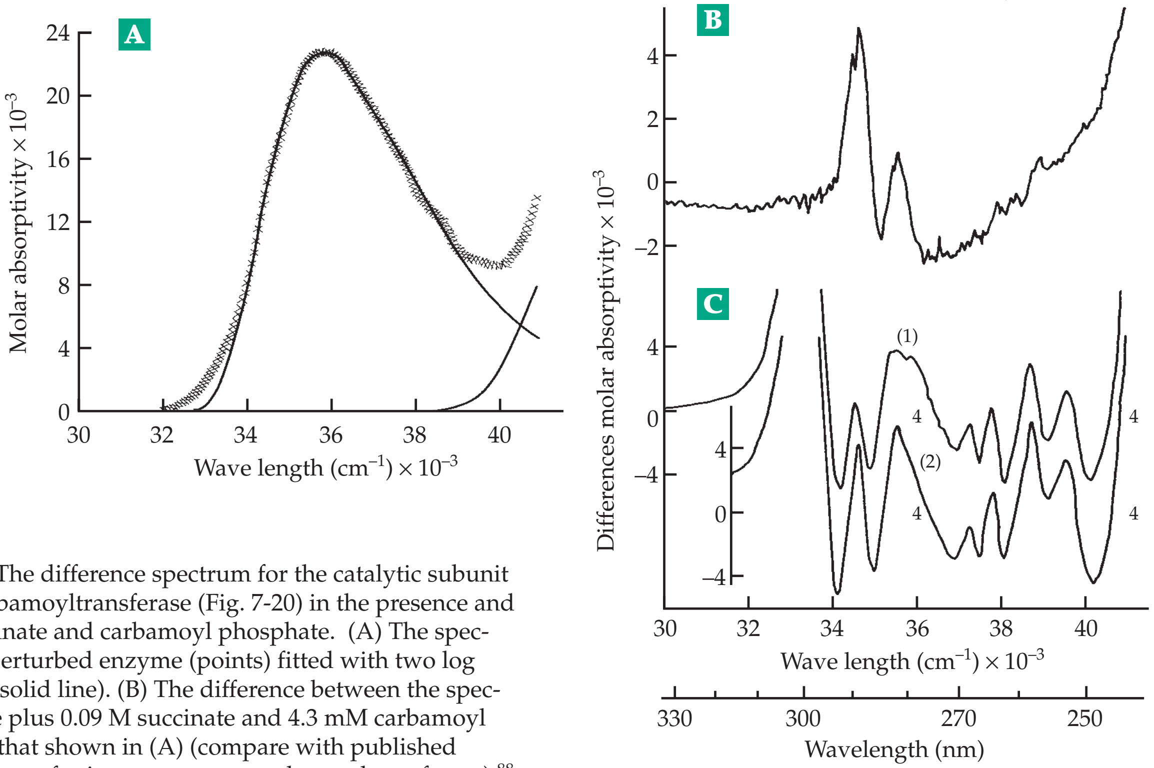


Figure 23-11 The difference spectrum for the catalytic subunit of aspartate carbamoyltransferase (Fig. 7-20) in the presence and absence of succinate and carbamoyl phosphate. (A) The spectrum of the unperturbed enzyme (points) fitted with two log normal curves (solid line). (B) The difference between the spectrum of enzyme plus 0.09 M succinate and 4.3 mM carbamoyl phosphate and that shown in (A) (compare with published difference spectrum for intact aspartate carbamoyltransferase).[88] (C) Curve 1, "fine structure plot" obtained by subtracting the spectrum in A from the smooth curve obtained by summing the two log normal curves. Curve 2, a similar plot for enzyme plus succinate and carbamoyl phosphate. The enzyme was supplied by G. Nagel and H. K. Schachman and the spectra were recorded by I.-Y. Yang.

smooth fitted curve and the experimental points taken at close intervals[90,91] as shown in Fig. 23-11C. The two "fine structure plots" obtained in this way are an alternative way of representing the same data that gave rise to the difference spectrum. The method has the advantage that information about the overall band shape is obtained from the computer-assisted curve fitting process. Thus, the binding of succinate and carbamoyl phosphate caused an almost insignificant shift (of 20 cm^{-1}) in the overall band position and a very slight broadening. The principal effect is an enhancement in the vibrational structure at 34,600 cm^{-1} in the 0–0 band of the two tryptophan residues present in the subunit. The cause of this change is not entirely obvious, a weakness of difference spectroscopy. Another way of examining a spectrum such as that in Fig. 23-11A is to plot a mathematical derivative of the curve. Both second[92] and fourth derivatives[91,93,94] yield curves similar to the difference plots of Fig. 23-11C.

5. Circular Dichroism and Optical Rotatory Dispersion

The circular dichroism of a sample is the difference between the molar extinction coefficients for left-handed and right-handed polarized light (Eq. 23-7) and is observed only for chiral molecules.[7,95–97]

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R \text{ (units are } M^{-1}\ cm^{-1}) \quad (23\text{-}7)$$

The **dichrograph** gives a direct measure of $\Delta\varepsilon$. A circular dichroism (CD) spectrum often resembles an absorption spectrum, the peaks coming at the same positions as the peaks in the absorption spectrum of the same sample. However, the CD can be either positive or negative and may be positive for one transition and negative for another (Fig. 23-8). It is most convenient to plot $\Delta\varepsilon$ directly as a function of wavelength or wave number. However, much of the literature makes use of the **molar ellipticity** (Eq. 23-8):

$$\text{Molar ellipticity} = [\theta] = 3299\ \Delta\varepsilon$$
$$\text{(units are degrees cm}^2\text{ decimole}^{-1}) \quad (23\text{-}8)$$

The **rotational strength** may also be evaluated (Eq. 23-9):

$$\text{Rotational strength} = \int [(\Delta\in)/\lambda] d\lambda \quad (23\text{-}9)$$

The integration is carried out over the entire absorption band for a given transition.

Circular dichroism is closely related to **optical rotatory dispersion**, the variation of optical rotation with wavelength. Optical rotation depends upon the difference in refractive index ($\eta_L - \eta_R$) between left-handed and right-handed polarized light. Rotation α is measured as an angle in degrees or radians. Data are customarily reported in terms of **specific rotation**, that of a hypothetical solution containing 1 g / ml in a 1 dm (decimeter) tube. Specific rotation is calculated (Eq. 23-10) from the observed rotation, the concentration c' in g ml^{-1}, and the length of the tube l' in decimeters.

$$\text{Specific rotation} = [\alpha] = \alpha_{obs} / c'l' \quad (23\text{-}10)$$

The **molecular rotation** is defined by Eq. 23-11 in which M_r is the molecular mass and c and l are in moles per liter and cm, respectively.

$$\text{Molecular rotation} = [\phi] = 100\ \alpha_{obs} / cl$$
$$= [\alpha]\ M_r / 100 \quad (23\text{-}11)$$

It is often multiplied by a factor of $3 / (\eta^2 + 2)$ to correct for a minor effect of the polarizability of the field acting on the molecules. The rotation in the radians per centimeter of light path can be related (Eq. 23-12) directly to the wavelength of the light and the refractive indices η_L and η_R.

$$\alpha \text{ (radians / cm)} = [\alpha]\ c' / 1800 = \pi / \lambda [\eta_L - \eta_R] \quad (23\text{-}12)$$

In contrast to circular dichroism, optical rotary dispersion (ORD) extends far from absorption bands into spectral regions in which the compound is transparent. As an absorption band is approached, the optical rotation increases in either the positive or negative direction. Then, within the absorption band it drops abruptly through zero and assumes the opposite sign on the other side of the band (the Cotton effect). Although the occurrence of optical rotation in nonabsorbing regions of the spectrum provides an advantage to ORD measurements, the interpretation of ORD spectra is more complex than that of CD spectra. In principle, the two can be related mathematically and both are able to give the same kind of chemical information.[7]

The CD in the *d*–*d* bands of the blue copper protein (Fig. 23-8) arises in part from the fact that within the protein the copper ion is in an asymmetric environment. For a similar reason, the aromatic amino acids of proteins often give rise to circular dichroism. In the case of tyrosine, the sign of the CD bands can be either positive or negative but is the same throughout a given transition. The CD bands are similar in shape to the absorption bands.[36,98] The behavior of phenylalanine is more complex. The progression of vibrational subbands at 930 cm^{-1} intervals above the 0–0 band all have the same sign, and the intensities relative to that of the 0–0 band are similar to those in absorption. However, the vibrations of wave numbers equal to that of the 0–0 transition plus 180 and 520 cm^{-1} sometimes give rise to CD bands of the opposite sign, and the relative intensity relationships are variable.[61,98]

The binding of a symmetric chromophore to a protein or nucleic acid often induces CD in that chromophore. For example, the bands of enzyme-bound pyridoxal and pyridoxamine phosphates shown in Fig. 14-9 are positively dichroic in CD, but the band of the quinonoid intermediate at 20,400 cm^{-1} (490 nm) displays negative CD. When "transimination" occurs to form a substrate Schiff base (Eq. 14-26), the CD is greatly diminished. While the coenzyme ring is known to change its orientation (Eq. 14-39; Fig. 14-10), it is not obvious how the change in environment is related to the change in CD.

A series of octant rules make it possible to predict the sign and magnitude of CD to be expected for $n-\pi^*$ transitions of simple carbonyl compounds.[99] Theoretical approaches to the CD and ultraviolet absorption of proteins in the high-energy ultraviolet region have also been developed. In a regular β structure, in an α helix or in a crystalline array, the transitions of adjacent amide groups may be **coupled**, the excitation energy being delocalized. This **exciton** delocalization leads to a splitting (Davydov splitting) into two transitions of somewhat different energies and polarized in different directions.[7,9] The amide absorption band at 52,600 cm^{-1} is split in an α helix into components at ~48,500 and 52,600 cm^{-1}. Furthermore, low-energy $\pi-\pi^*$ and $n-\pi^*$ states are close together in energy, a fact that allows mixing of the two states and appearance of rotational strength in the $\pi-\pi^*$ band with a sign opposite to that in the $n-\pi^*$ band.[100]

Both the sign and intensity of the CD bands of peptides also depend upon conformation. Well-defined differences are observed among α helices, β structure, and random-coil conformations. Measurements may be extended into the "vacuum ultraviolet" region—up to 60,000 cm^{-1} in aqueous solutions.[101] A useful empirical approach is to deduce spectra of helices, β structures, and unordered peptide chains from measured spectra together with an examination of actual structures obtained by X-ray crystallography[7,95,97,102,103] (Fig. 23-12). Note that the CD curve for the α helix has a deep minimum at 222 nm, whereas the β form has a shallower minimum. The random structure has almost no CD at the same wavelength. The approximate helix content of a protein is often estimated from the depth of the trough at 222 nm in the CD spectrum. Better predictions can be made by using a computer-assisted comparison of an experimental CD spectrum with those of a series of proteins of known 3-D structure.[7]

The circular dichroism of polynucleotides at 275 nm is a linear function of both the helix winding angle and the base pair twist.[103] Measurement of CD spectra on large polynucleotides or large molecular aggregates is complicated by differential light scattering of right and left circularly polarized light. However, the phenomenon can also provide new structural information.[104] Progress has been made in attempts to predict the optical rotation of molecules from quantitative values for the polarizabilities of individual atoms.[105–109]

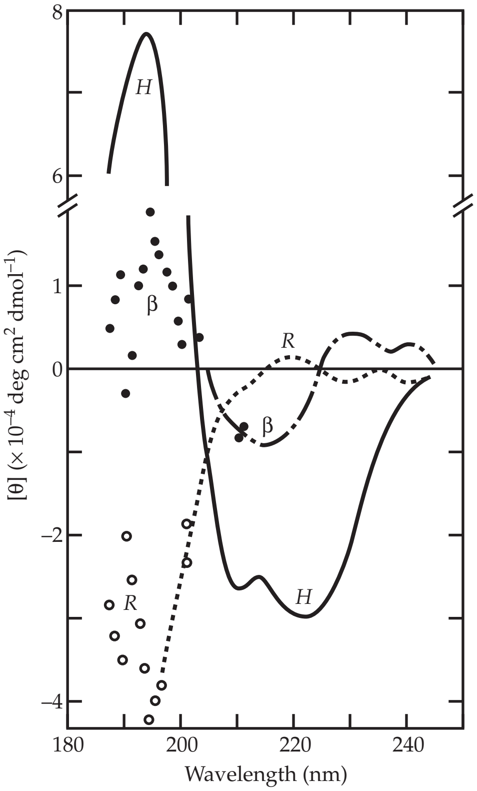


Figure 23-12 Circular dichroism of the helix (H), β, and unordered (R) form computed from the CD of five proteins. Points are plotted when a smooth curve could not be drawn. From Y.-H. Chen *et al*.[102]

Vibrational circular dichroism involves IR absorption bands. The technique has been applied to sugars,[110] oligosaccharides,[111] proteins,[112] and nucleic acids.[113] The related **vibrational Raman optical activity** has also been applied to polyribonucleotides.[114]

6. Photoacoustic Spectroscopy

Photoacoustic spectra are recordings of the energy emitted as heat after absorption of monochromatic light. The sample is placed in a closed photoacoustic cell. The light beam, which is chopped at an audiofrequency, induces a periodic heating and cooling of the gas in contact with the sample in the cell. This is sensed as sound by a sensitive microphone. The resulting electrical signals are sent to a computer for analysis. The output is an absorption spectrum resembling that measured optically. The samples do not have to be transparent.[5] A related technique is photoacoustic calorimetry.[115,116]

7. X-Ray Absorption and Mössbauer Spectroscopies

The importance of X-ray diffraction (discussed in Chapter 3) to biochemistry is obvious, but techniques related to absorption of X-rays and γ-rays have also come into widespread use.[117] Abbreviations such as **XANES** and **EXAFS** are common in the metalloprotein literature. The names arise from the sharp increase in the absorption coefficient for X-rays as their energy is increased to what is called the **K absorption edge**. At slightly lower energies absorption of an X-ray by an atom leads to expulsion of an electron or the raising of an electron to an excited state. Absorption of X-rays will expel all except the inner 1*s* electrons. As the energy is increased further, the stepwise increase in absorption that constitutes the edge is observed. At higher energies the absorption decreases. However, with a high-resolution instrument distinct oscillations are observed on the high-energy side of the edge, extending for ~20 eV. This is **X-ray absorption near-edge structure** (XANES). When an X-ray absorbing atom in a molecule is surrounded by other atoms, a fine structure that depends upon the nature of these atoms and their distances from the absorbing atom is observed over a range of several hundred electron volts above the edge. This is **extended X-ray absorption fine structure** (EXAFS).[117]

The EXAFS technique has been especially useful for metalloproteins. It has often provided the first clues as to the identity of atoms (O, N, S) surrounding a metal atom and either covalently bonded to it or coordinated with it (Chapter 16). Interpretations are often difficult, and a common approach is to try to simulate the observed spectrum by calculation from a proposed structure.[118] Tautomerism in crystalline Schiff bases (see Eq. 23-24) has been studied by near-edge X-ray absorption fine structure (NEXAFS) employing soft X-rays.[119]

Mössbauer spectroscopy, also called recoil-free nuclear resonance absorption, depends upon resonant absorption of γ-rays emitted by a radioactive source by atomic nuclei.[120] The phenomenon was initially difficult to observe, but the German physicist Mössbauer devised a way in which to record the absorption of a quantum of energy equal to the difference in two energy states of the atomic nucleus. The method depends upon a Doppler effect observed when the sample or source moves. Consequently, Mössbauer spectra, such as that in Fig. 16-18, are plots of absorption versus velocity. Mössbauer spectroscopy has been applied to numerous metalloproteins, especially those containing iron centers. It is a major tool in investigation of Fe–S proteins.[121,122] Since ^{56}Fe is "silent" in Mössbauer spectroscopy, proteins are often enriched with ^{57}Fe for observation.

C. Fluorescence and Phosphorescence

An electronically excited molecule is able to lose its excitation energy and return to the ground state in several ways. One of these is to reemit a quantum of light as fluorescence.[7,123–127] The intensity and spectral properties of fluorescent emission can be measured by illuminating a sample in a cuvette with four clear faces with the measuring photomultiplier set at right angles to the exciting light beam. In absorption spectrophotometry we measure a difference between the light intensity of the beam entering the sample and that emerging from the sample. In fluorescence spectroscopy we measure the absolute intensity of the light emitted. Although this intensity is small, the measurement can be made extremely sensitive, far more so than can light absorption. For this reason, fluorescence is widely used for detection and analysis, e.g., in DNA sequencers. Enzyme kinetics can be studied with fluorescent substrates at very low concentrations.[127a] Fluorescent antibodies, DNA chips, and numerous bioassay and imaging methods are dependent upon measurement of fluorescence. Fluorescence can also yield a wealth of information about the chemical and physical properties of electronically excited states of molecules.

1. Excitation and Emission Spectra

Measurements of the intensity of fluorescence at any wavelength vs the wavelength of monochromatic light used to excite the fluorescence give a fluorescence **excitation spectrum**. The excitation spectrum is an example of an **action spectrum**, which is a measure of any response to absorbed light. At very low concentrations of pure substances, action spectra tend to be identical to absorption spectra. However, since the observed response (fluorescence in this case) is proportional to light absorbed, action spectra should be compared to plots of $1-T$ (where T = transmittance, Section B,1) vs wavelength rather than to plots of ε vs λ. The two plots are proportional at low concentrations. For a discussion of action spectra see Clayton.[123]

A fluorescence **emission spectrum** is a record of fluorescence intensity vs wavelength for a constant intensity of exciting light. Excitation and emission spectra for a flavin and for the indole ring of tryptophan are both given in Fig. 23-13. The heights of the

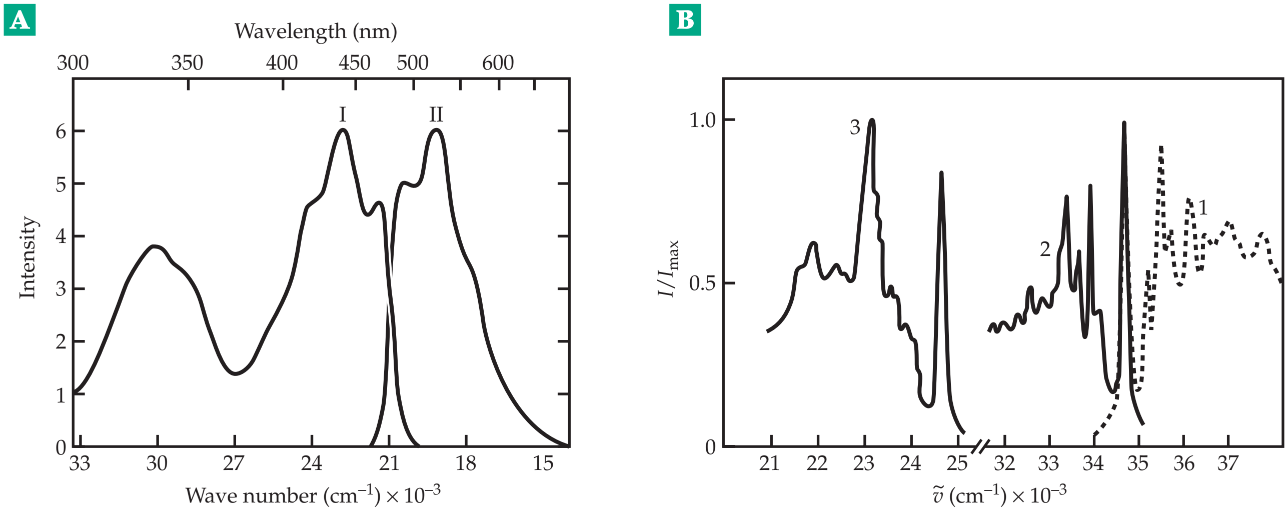


Figure 23-13 (A) Corrected emission and excitation spectra of riboflavin tetrabutyrate in *n*-heptane. Concentration, about 0.4 mg l^{-1}. Curve 1: excitation spectrum; emission at 525 nm. Curve 2: emission spectrum; excitation at 345 nm. From Kotaki and Yagi.[128] (B) Indole in cyclohexane, T = 196 K. 1, Fluorescence excitation spectrum; 2, fluorescence spectrum; and 3, phosphorescence spectrum. From Konev.[125]

emission spectra have been adjusted to the same scale as that of absorption. The fluorescent emission is always at a lower energy than that of the absorbed light. The excitation and emission spectra overlap only slightly, and the emission spectrum is an approximate mirror image of excitation spectrum. To understand this, refer to the diagram in Fig. 23-14. Absorption usually leads to a higher vibrational energy state after light absorption than before. However, most of the excess vibrational energy is dissipated before much fluorescent emission occurs. The excited molecule finds itself in the lowest vibrational state of the upper electronic state, and it is from this state that the bulk of the fluorescent emission takes place. Furthermore, whereas absorption usually occurs from the lowest vibrational state of the ground electronic level, fluorescence can populate many excited vibrational states of the ground electronic state (Fig. 23-14). Consequently, as indicated in the figure, the fluorescent emission spectrum consists of a series of subbands at lower energies than those observed in the absorption spectrum. The two spectra have only the 0–0 transition in common. As can be seen from Fig. 23-14, even the two 0–0 transitions do not coincide exactly. The peak of emission is shifted toward slightly lower energies than that of absorption because during or immediately following absorption of a photon there is some rearrangement of solvent molecules around the absorbing molecule to an energetically more stable arrangement. Just as excess vibrational energy is dissipated in the excited state, so relaxation of these solvent molecules around the excited chromophore leads to a small shift in energy. A similar relaxation

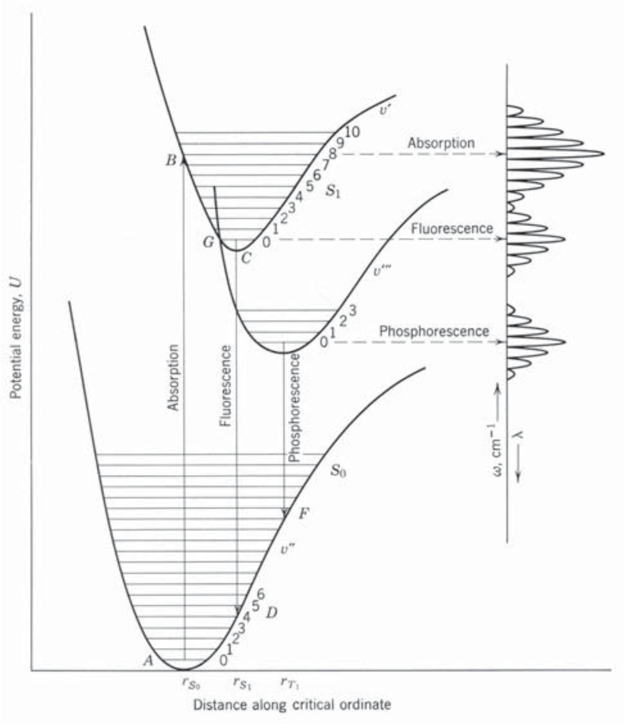


Figure 23-14 Potential energy diagram for the ground state S_0 and the first excited singlet S_1 and triplet T_1 states of a representative organic molecule in solution. G is a point of intersystem crossing $S_1 \rightarrow T_1$. For convenience in representation, the distances *r* were chosen $r_{S_0} < r_{S_1} < r_{T_1}$; thus, the spectra are spread out. Actually, in complex, fairly symmetric molecules, $r_{S_0} \sim r_{S_1} < r_{T_1}$ and the 0–0 absorption and fluorescence bands almost coincide, but phosphorescence bands are significantly displaced to the lower wavelengths. From Calvert and Pitts,[2] p. 274.

occurs in the ground state of a molecule that has just emitted a photon as fluorescence. This also contributes to the shift in position of the 0–0 band in fluorescence (see Parker,[124] p. 13).

Several molecular properties can be measured using emission and excitation spectra. These include fluorescence lifetime, efficiency, anisotropy of the emitted light, mobility of chromophores, rates of quenching, and energy transfer to other chromophores.

Fluorescence lifetimes. Why are some molecules fluorescent, while others are not? The possibility for fluorescent emission is limited by the radiative lifetime τ_r, which is related by Eq. 23-13 to the first-order rate constant k_f for exponential decay of the excited state by fluorescence.

$$\tau_r = 1 / k_f \qquad (23\text{-}13)$$

The radiative lifetime is a function of the wavelength of the light and of the oscillator strength of the transition. For molecules absorbing in the near UV, the approximation of Eq. 23-14 is often made.

$$1 / \tau_r \sim 10^4 \varepsilon_{max}) \qquad (23\text{-}14)$$

Thus, if $\varepsilon = 10{,}000$, the radiative lifetime (the time in which the fluorescence decays to $1/e$, its initial value) is ~10^{-8} s (10 ns). If the absorption is more intense, the lifetime is shorter, and if it is less intense, it is longer. Other modes of deexcitation compete with fluorescence; therefore, the shorter the radiative lifetime the more likely that fluorescence will be observed.

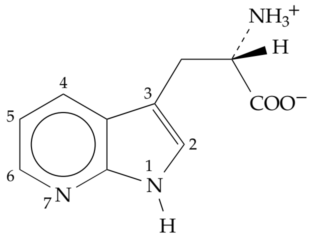
The actual lifetime τ of an excited molecule is usually less than τ_r because of the competing nonradiative processes. The sum of their rate constants can be designated k_{nr}. The **fluorescence efficiency** (or **quantum yield**) ϕ_F is given by Eq. 23-15.

$$\phi_F = k_f / k_f + k_{nr}) = k_r \tau$$
$$\text{where } \tau^{-1} = k_f + k_{nr} \qquad (23\text{-}15)$$

For a highly fluorescent molecule such as riboflavin, ϕ_f may be 0.25 or more.[129] For tryptophan in water it is about 0.14, and in proteins it varies from near zero to 0.35.[130]

Time-resolved fluorescence spectroscopy. The fluorescence lifetime τ can be measured with either of two different types of fluorometer.[7,127,131–133] **Pulse fluorometers** use pulsed lasers that can deliver pulses of light lasting as little as one picosecond or less. This permits rapid excitation and permits the direct observation of emitted light, using photon countings, over the entire range of time from a few picoseconds to milliseconds required for decay of the fluorescence.[131] The observed value of τ for riboflavin 5'-phosphate (ε_{max} = 12,200 at 450 nm) at 25°C is ~5 ns.[134] That for tryptophan is 3 ns.

Phase fluorometers utilize continuous irradiation by a beam of light that is sinusoidally modulated. If the frequency of the modulation is set correctly, there will be a phase difference in the modulation of the fluorescent emission that will depend upon τ. Phase fluorometry can yield the same information as does pulse fluorometry.[127,132,133] By using two or more modulation frequencies the decay rates and fluorescence lifetimes for various chromophores in a protein can be observed. For example, the protein **colicin A** (Box 8-D) contains three tryptophans W86, W130, and W140. Their fluorescence decays with lifetimes τ_1, τ_2, τ_3 of ~0.6–0.9 ns, 2.0–2.2 ns, and 4.2–4.9 ns at pH 7. While τ_3 originates mainly from W140, both of the other tryptophans contribute to τ_1 and τ_2. Changes in fluorescence intensity with pH reflect a pK_a value of 5.2.[135] Tryptophan, which often occurs at only one or a few places in a protein, is a useful fluorescent probe for study of protein dynamics. The optical properties of 7-azatryptophan, 2-azotryptophan, and 5-hydroxytryptophan are even better because their absorption maxima occur at longer wavelengths. These amino acids can be biosynthetically introduced in place of tryptophan in proteins.[136–138] The maximum fluorescence of tryptophan in one protein is at 350 nm, but for 7-azatryptophan in the same protein it was shifted to 380 nm.[136]



7-Azatryptophan

Triplet states, phosphoresence, and quenching. In addition to emitting fluorescent radiation, molecules can often pass from the excited singlet state to a lower energy **triplet state**, in which two electrons are now unpaired and the molecule assumes something of the character of a diradical (see Fig. 23-14). This process, known as **intersystem crossing**, competes directly with fluorescence and shortens the fluorescence lifetime. The triplet state is long-lived (e.g., for tryptophan in water at 20°C it is 1.2 ms[139]) and is responsible for much of the photochemical behavior of molecules. It also gives rise to the delayed light emission known as **phosphorescence**, as is illustrated in Figure 23-14. Other processes that compete with fluorescence are **photochemical reactions** of the singlet excited state and **internal conversion**. The latter is the process by which a molecule moves from the lowest vibrational state of the upper electronic level to some high vibrational state of the unexcited electronic level. This is the principal means of depopulating the electronic state and competes directly with fluorescence.

The rate of relaxation by nonradiative pathways can be increased by addition of **quenchers**. Quenching of fluorescence occurs by several mechanisms, many of which involve collision of the excited chromophore with the quenching molecule. Some substances such as iodide ion are especially effective quenchers. The fluorescence efficiency of a substance in the absence of a quencher can be expressed (Eq. 23-16) in terms of the rate constants for fluorescence (k_f), for nonradiative decay (k_{nr}), and for phosphorescence (k_p):

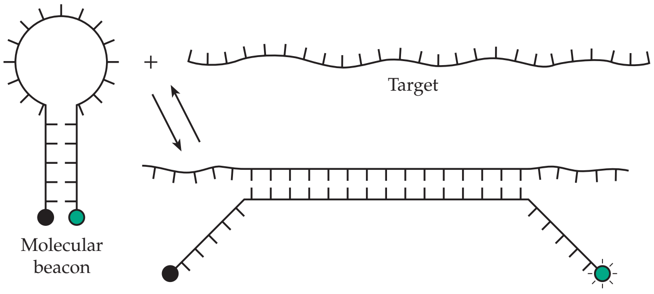
$$\phi_f = k_f / (k_f + k_{nr} + k_p) \qquad (23\text{-}16)$$

In the presence of a quencher, Q, there is an additional rate process for relaxation. The ratio of the fluorescence efficiency in the absence of (ϕ_f^0) and the presence of a quencher is given by the **Stern–Volmer equation**.[7,140]

$$\phi_f^0 / \phi_f = 1 + K[\mathrm{Q}] = 1 + k_Q \tau_0 [\mathrm{Q}] \qquad (23\text{-}17)$$

The constant K is known as the Stern–Volmer quenching constant; k_Q is the rate constant for the quenching reaction, and τ_0 the lifetime in the absence of quencher. Fluorescence quenching of tryptophan in proteins by acrylamide or O_2 has been used to determine whether tryptophan side chains are accessible to solvent or are "buried" in the protein.[141,142] The long-lived phosphorescence of tryptophan can be studied in a similar way.[143]

A recent application of fluorescence quenching is the development of "**molecular beacons**" for detection of viruses such as the AIDS viruses HIV-1 and HIV-2.[144] A single-stranded oligonucleotide is synthesized with a 25- or 33-nucleotide sequence complementary to a sequence in the target viral RNA. At the ends of this sequence are two 6-nucleotide arms with complementary sequences that will form a stable double-helical stem at the annealing temperature used for PCR amplification of the viral nucleic acid. The end of one arm carries a covalently bonded fluorescent dye, e.g., a fluorescein or rhodamine derivative. The other arm carries a potent covalently linked fluorescence quencher such as 4-(4'-dimethylaminophenylazo)benzoic acid. When the arms form a duplex, the quencher will be next to the fluorophore and no fluorescence will be seen upon irradiation with light of a suitable exciting wavelength. However, if viral DNA is present it will hydridize with the central polynucleotide, keeping the fluorophore and quenchers far apart and allowing the beacon to signal the presence of a virus (Eq. 23-18). As few as ten retroviral genomes could be detected. By using a series of molecular beacons with different colored fluorescence and specific for different viruses, it is possible to test for more than one virus simultaneously.



(23-18)

Anisotropy. Light emitted from excited molecules immediately after absorption is always partially polarized, whether or not the exciting beam consists of plane polarized light. When light polarized in a vertical plane is used for excitation, part of the emitted light (of intensity I_v) will have its electric vector parallel to that of the exciting light. The remainder of intensity I_h, will be polarized in a horizontal plane. The **polarization *P*** of the emitted radiation is defined by Eq. 23-19 and the **anisotropy *R*** by Eq. 23-20. After excitation by a laser pulse both the fluorescence and its anisotropy decay with time and can be measured. The decay of R (but not of P) can usually be described as the sum

$$P = (I_v - I_h) / (I_v + I_h) \qquad (23\text{-}19)$$

$$R = (I_v - I_h) / (I_v + 2\,I_h) \qquad (23\text{-}20)$$

of simple exponential curves, which are readily obtained by phase fluorometry. These can in turn be related to specific types of motion, such as rotation of the emitting molecule or group.[7,145,146] Rotation of tryptophan rings, both free and restricted, has been studied in a variety of proteins.[145,147] However, interpretation is difficult.[130,148] The rotational rates obtained from anisotropy measurements are strongly affected by the viscosity of the medium (see Eq. 9-35).

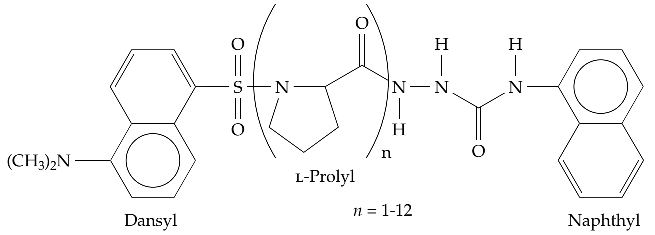
2. Fluorescence Resonance Energy Transfer (FRET)

Electronic excitation of one chromophore sometimes elicits fluorescence from a different chromophore that is located nearby. For example, excitation of a monomolecular layer of dye can induce fluorescence in a layer of another dye spaced 5 nm away. Excitation of tyrosine residues in proteins can lead to fluorescence from tryptophan, and excitation of tryptophan can cause fluorescence in dyes attached to the surface of a protein or in an embedded coenzyme.[134] Such fluorescence resonance energy transfer (**FRET**) is expected for molecules, when the fluorescence spectrum of one overlaps the absorption spectrum of the other. The mechanism is not one of fluorescence emission and absorption but of nonradiative resonant transfer of energy. Resonant transfer of energy is of major biological significance in photosynthesis (Section E). Most of the chlorophyll molecules, which absorb light in a chloroplast, transfer the absorbed energy in a stepwise fashion to a **reaction center**.

Förster[149] calculated that the rate of energy transfer k_t should be proportional to the rate of fluorescence k_f, to an orientation factor K^2, to the spectral overlap interval J, to the inverse fourth power of the refractive index n, and to the inverse sixth power of the distance r separating the two chromophores.

$$k_t \propto k_f K^2 J n^{-4} r^{-6} \tag{23-21}$$

Besides predicting the inverse sixth power dependence of energy transfer, Förster provided a formula for calculating R_0, the distance between chromophores at which 50% efficient singlet-singlet energy transfer takes place. R_0 is commonly of the order of 2.0 nm. Making use of these relationships, Stryer proposed a method of measuring distances between chromophores. He calibrated the method by constructing a series of molecules containing various lengths of the rigid threefold polyproline helix to which dansyl groups were attached at one end and naphthyl groups at the other.[150] By exciting the naphthyl group, which has the higher energy absorption band and is strongly fluorescent, the characteristic lower energy emission of the dansyl group could be observed if energy transfer took place. Since the fluorescent emission band of the naphthyl group overlaps the absorption band of the dansyl group, efficient transfer was expected. The results of a plot of transfer efficiency against distance is shown in Fig. 23-15. The inverse sixth power dependence was followed quite accurately with a value of R_0 ~3.4 nm. Having calibrated his "spectroscopic ruler," Stryer turned his attention to biochemical macromolecules. Attaching the same kinds of fluorescent probe to the visual light receptor rhodopsin, Wu and Stryer were able to estimate distances between specific parts of the molecule and to draw some conclusions about the overall shape.[151]



More recently the FRET technique has been widely applied to a broad range of biochemical problems. Sensitivity has been improved to the extent that fluorescence of single molecules can be detected.[152–154] Use of **terbium** (Tb^{3+}) or **europium** (Eu^{3+}) ions, which can provide luminous labels for metal-binding sites, has provided another advance. These ions absorb light poorly and are therefore only weakly fluorescent. However, they can be excited by resonance energy transfer and become brilliantly luminous. This **luminescence resonance energy transfer** (**LRET**) is a variant of FRET, which allows distances up to ~10 nm to be measured.[155–157] Another advance is the ability to graft into specific proteins fluorescent tags such as the intact **green fluorescent protein** (Section J)[158,159] or an amino acid sequence such as CCXXCC in which the four –SH groups of the cysteines serve to trap an arsenic derivative of fluorescein (see Box 12-B).[159] **Confocal laser scanning microscopy** (Chapter 3) is basic to many applications.[160]

Specific applications of FRET and LRET include observation of myosin movement (Fig. 19-14),[157] measurement of distances between binding sites on tubulin,[161] determining stoichiometry of subunit assembly in a γ-aminobutyrate receptor of brain,[162] association of proteins in peroxisomes,[160] study of hydridization of deoxyribonucleotides,[163] verifying the handedness of various forms of DNA,[164] and other studies of DNA and RNA.[164a,b]

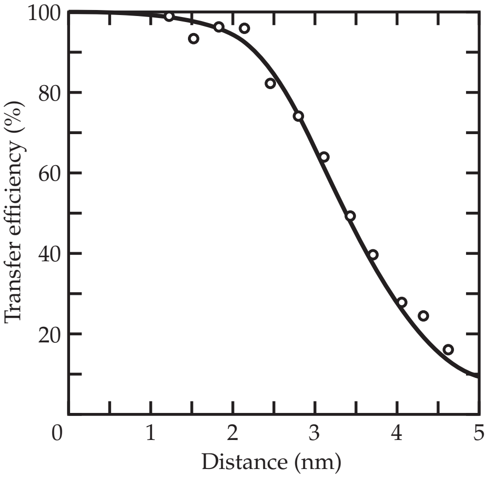


Figure 23-15 Efficiency of energy transfer as a function of distance between α-naphthyl and dansyl groups at the ends of a polyprolyl "rod" (L-prolyl)$_n$. The observed efficiencies of transfer for n = 1 to 12 are shown as points. The solid line corresponds to an r^{-6} distance dependence. From L. Stryer.[150]

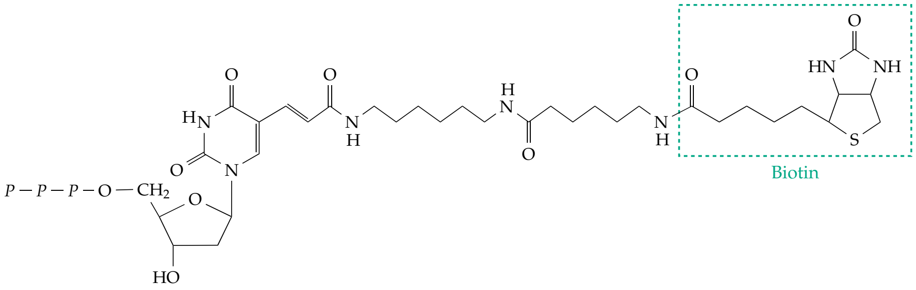
Using the Förster equation the distance between the two calcium-binding sites in parvalbumin (Fig. 6-7) has been estimated by energy transfer from Eu(III) in one site to Tb(III) in the other[165] to within 10–15% of the distance of 1.18 nm based on X-ray crystallography.

3. Energy-Selective Spectroscopic Methods

At low enough temperatures vibrational fine structure of aromatic chromophores may be well resolved, especially if they are embedded in a suitable matrix such as argon or N_2, which is deposited on a transparent surface at 15 K. This **matrix isolation spectroscopy**[77,166] may reveal differences in spectra of conformers or, as in Fig. 23-16, of tautomers. In the latter example the IR spectra of the well-known amino-oxo and amino-hydroxy tautomers of cytosine can both be seen in the matrix isolation IR spectrum. Figure 23-16 is an IR spectrum, but at low temperatures electronic absorption spectra may display sharp vibrational structure. For example, aromatic hydrocarbons dissolved in *n*-heptane or *n*-octane and frozen often have absorption spectra, and therefore fluorescence excitation spectra, which often consist of very narrow lines. A laser can be tuned to excite only one line in the absorption spectrum. For example, in the spectrum of the carcinogen 11-methylbenz(*a*)anthrene in frozen octane three major transitions arise because there are three different environments for the molecule. Excitation of these lines separately yields distinctly different emission spectra.[77] Likewise, in complex mixtures of different hydrocarbons emission can be excited from each one at will and can be used for estimation of amounts. Other related methods of energy-selective laser spectroscopy include **fluorescence line narrowing**[167] and **spectral hole burning**.[167,168]

4. Analytical Applications of Fluorescence

Because of the high sensitivity with which fluorescence can be detected, its measurement is important as an analytical tool. As a result of improved techniques **fluorescence microscopy** has become one of the most important of all tools in biological studies.[168a] New types of microscopes (see also pp. 129–131) have increased resolution beyond what was thought possible.[168b–d] Studies such as those of lipid metabolism in the transparent zebrafish are possible using substrates that carry fluorescent labels.[168e] As mentioned in the preceding paragraph, many aromatic compounds can be detected by their fluorescence. The relatively weak fluorescence of proteins and nucleic acids can be greatly enhanced by the binding of a highly fluorescent dye to the macromolecules. Fluorescent antibodies are widely employed for this purpose. Fluorescent labels are rapidly replacing radioisotopes in analysis of nucleic acids. For example, biotin may be attached to a pyrimidine base of a nucleoside triphosphate by a long spacer arm. The modified base can then be incorporated enzymatically into polynucleotides, e.g., in the synthesis of probes used for hybridization. The attached biotin can be detected by binding to avidin or streptavidin (Box 14-B) and use of fluorescent antibodies to this



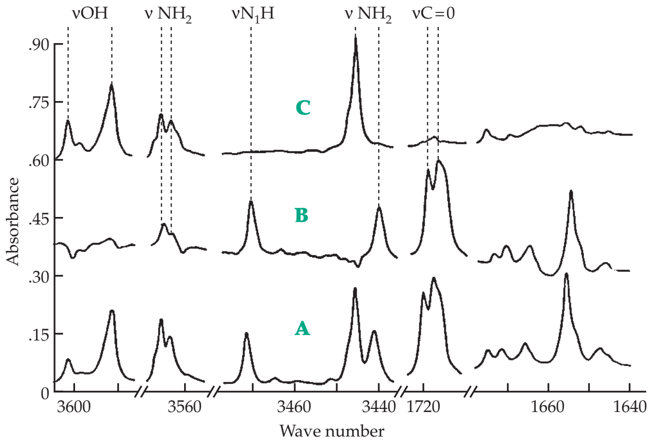
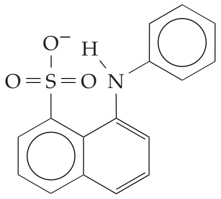


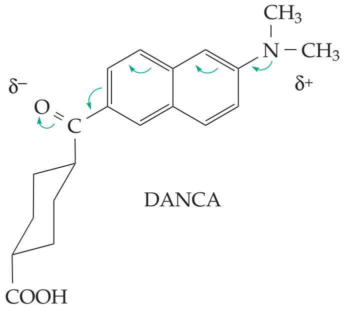
Fig. 23-16 Infrared spectrum of cytosine in the –OH, NH, and C=O stretching region in an argon matrix in 15 K. (A) Initial spectrum of a sample containing both tautomers. (B) Spectrum of the amino-oxo tautomer calculated by substraction of the spectrum of the amino-hydroxy tautomer. (C) Spectrum of the aminohydroxy tautomer obtained by irradiating the initial sample in order to photochemically convert it into the aminohydroxy tautomer. Spectra (B) and (C) were also corrected for residual amounts of the second tautomer. From Szczesniak *et al.*[166] Courtesy of W. B. Person.

protein.[169,170] Fluorescent dyes can also be covalently attached to nucleotides. Fluorescent dideoxynucleoside triphosphates are used as chain terminators in DNA sequencing (Chapter 5). Using a different dye that fluoresces at a different wavelength for each of the four dideoxynucleosides, polynucleotides can be sequenced automatically using a single column rather than four parallel lanes as in Fig. 5-49.

Fluorescent "probes" such as **1-anilinonapthalene-8-sulfonate** or 1,6-diphenyl-1,3,5-hexatriene embedded in membranes, contractile fibers, etc., can reveal changes in mobility that accompany alterations in physiological conditions. For example, molecular changes occurring in membranes during nerve conduction and in mitochondria during electron transport can be observed.[146,171] Another type of probe is exemplified by 2'-(*N*,*N*-dimethylamino)-6-4-*trans*-cyclohexamoic (DANCA). Its emission maximum shifts from 390 nm in cyclohexane to 520 nm in water, presumably as a consequence of increased polarization of the molecule in the excited state, as indicated by the green arrows in the accompanying structural formula.[172] DANCA can be used to obtain some idea of the polarity of sites within macromolecules to which it binds.

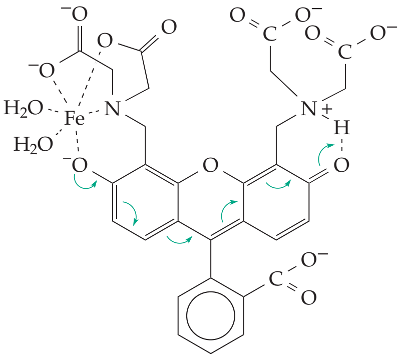


1-Anilinonaphthalene-8-sulfonate



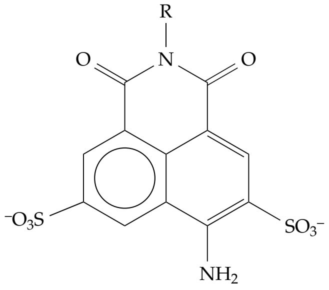
DANCA

Study of calcium ions in living cells has been immensely aided by calcium fluorophores[173] (Box 6-D), which are often derivatives of EDTA (Table 6-9). One of these, **calcein**,[174] is also very specific toward Fe^{3+}. The natural calcium-dependent luminous protein, aequorin (Section J), is also widely used.



Fe^{III} - calcein • H^+

Nontoxic but highly fluorescent dyes are used to study diffusion within and between cells. The so called Lucifer dyes have been employed to trace shapes and branching patterns in neurons.[175]



Structural formula for Lucifer dyes
For Lucifer yellow
R = —NH—CO—NH—NH_2

The recently developed **fluorescence correlation spectroscopy** permits studies of molecular association in one femtoliter of solution using a confocal or two-photon microscope. Two lasers are used to excite two fluorophores of different colors, each one on a different type of molecule. Fluorescence of single molecules can be detected, and molecular associations can be detected by changes in the distribution of the fluctuations in fluorescence intensity caused by Brownian motion.[176–178] A different type of advance is development of computer programs that analyze chromosomes stained with a mixture of dyes with overlapping spectra and display the result as if each chromosome were painted with a specific color.[179–180a] Yet another advance is development of **semiconductor nanocrystals** (or "quantum dots") with narrow absorption bands and intense fluorescence. The wavelength of absorption and fluorescence depends upon the size of the crystals. For example, CdSe crystals of diameter 2–5 nm coated with silica or with a surfactant fluoresce across the visible range. They have a variety of uses in biological staining.[181,182]

D. Photochemistry

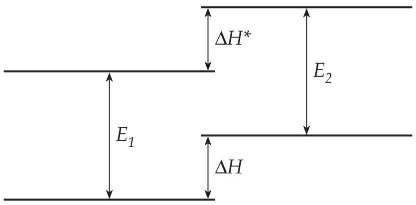
Because of their high energy, molecules in either the singlet or triplet photoexcited state undergo a greater variety of chemical reactions than do molecules in the ground state.[5,183,184] Many of these photochemical reactions arise from the triplet state that is formed from the singlet by intersystem crossing. Selection rules forbid transitions between excited triplet state and ground state; therefore, the radiative lifetime of the triplet state is long. The diradical character of the triplet state also makes it unusually reactive. Despite its forbidden character, nonradiative deexcitation of the triplet state is possible, and phosphorescence is observed for most molecules at low temperatures if the solvent is immobilized as a glass. The intense light from lasers can also induce a variety of photochemical processes that arise from absorption of two or more photons.

One of the simplest of photochemical processes is the dissociation or uptake of a proton by an excited molecule as a result of a change in the pK_a value of a functional group. Various other bond cleavages may lead to dissociation into ions or radicals. Photoelimination and photoaddition reactions both occur. Molecules may be isomerized, a process of importance in visual receptors. Excited molecules may become strong oxidizing agents able to accept hydrogen atoms or electrons from other molecules. An example is the **photooxidation** of EDTA by riboflavin (which undergoes photoreduction as shown in Fig. 15-8). A biologically more important example is in photosynthesis, during which excited chlorophyll molecules carry out **photoreduction** of another molecule and are themselves transiently oxidized. A frustrating aspect of investigation of photochemical reactions is that the variety of reactions possible often leads to a superabundance of photochemical products, e.g., see the thin layer chromatogram of cleavage products of riboflavin in Fig. 3-5. However, biological photoprocesses are usually much more specific.

1. Chemical Equilibria in the Excited State

When pyridoxamine with a dipolar ionic ring structure (Fig. 14-9) and an absorption peak at 30,700 cm^{-1} (326 nm) is irradiated, fluorescence emission is observed at 25,000 cm^{-1} (400 nm). When basic pyridoxamine with an anionic ring structure and an absorption peak at 32,500 cm^{-1} (308 nm) is irradiated, fluorescence is observed at 27,000 cm^{-1} (370 nm), again shifted ~5500 cm^{-1} from the absorption peak. However, when the same molecule is irradiated in acidic solution, where the absorption peak is at 34,000 cm^{-1} (294 nm), the luminescent emission at 25,000 cm^{-1} is the same as from the neutral dipolar ionic form and abnormally far shifted (9000 cm^{-1}) from the 34,000 cm^{-1} absorption peak.[185,186] The phenomenon, which is observed for most phenols, results from rapid dissociation of a proton from the phenolic group in the photoexcited state. A phenolic group is more acidic in the excited state than in the ground state, and the excited pyridoxamine cation in acid solution is rapidly converted to a dipolar ion.

The variation of fluorescence intensity with pH can provide direct information about the pK_a in the excited state. Förster suggested the following indirect procedure for estimating excited-state pK_a values for phenols. Let E_1 represent the energy of the 0–0 transition (preferably measured as the mean of the observed 0–0 transition energies in absorption and fluorescent emission spectra); let E_2 represent the energy of the 0–0 transition in the dissociated (anionic in the case of a phenol) form, while ΔH and ΔH^* represent the enthalpies of dissociation in the ground and excited states, respectively. It is evident from the diagram that Eq. 23-22 holds.



$$E_1 - E_2 = \Delta H - \Delta H^* \tag{23-22}$$

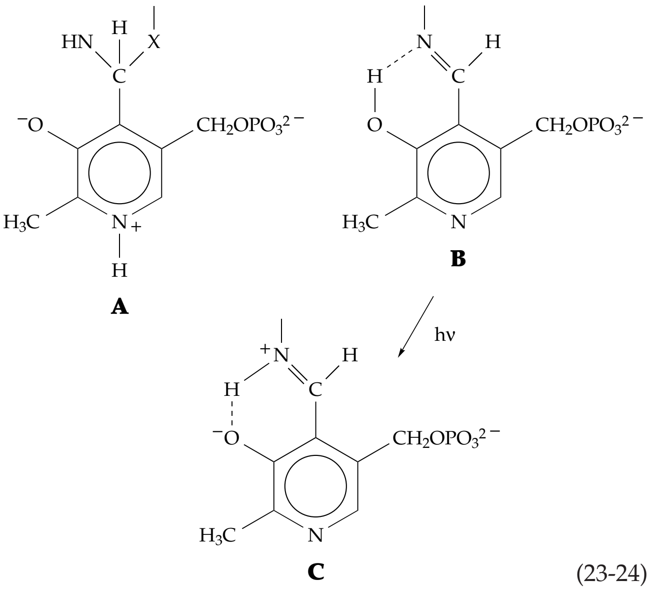
If we assume that the changes in entropy for the reaction are the same in the ground and excited state, Eq. 23-23 follows.

$$\log_{10}(K^*/K) = Nh\,(\Delta\,\bar{\nu})/(2.3RT)$$
or
$$\text{p}K^* = \text{p}K - (2.1 \times 10^3)\Delta\nu\ (\text{cm}^{-1}) \text{ at } 25°\text{C} \tag{23-23}$$

For example, a shift in the spectrum of the basic form by 1000 cm^{-1} to a lower wave number compared with the acid form corresponds to a decrease of 2.1 units in pK_a for dissociation of the acid form. Whereas it is desirable to use both absorption and fluorescent measurements to locate the approximate positions of 0–0 bands, absorption measurements alone are often used, and the positions of the band maxima are taken. Thus, for pyridoxamine the shift in absorption maximum from 34,000 cm^{-1} in the protonated form to 30,700 cm^{-1} in the dissociated form suggests that the pK_a of pyridoxamine of 3.4 in the ground state is shifted by 6.9 units to −3.5 in the excited state. Bridges *et al.*[186] evaluated this same pK_a from the pH dependence of fluorescence as pK^* ~ −4.1.

While phenols and amines are usually more acidic in the singlet excited state than in the ground state, some substances, e.g., aromatic ketones, may become more *basic* in the photoexcited state.

Observation of an abnormally large shift in the position of fluorescent emission of pyridoxal phosphate (PLP) in glycogen phosphorylase answered an interesting chemical question.[187,188] A 330 nm (30,300 cm^{-1}) absorption band could be interpreted either as arising from an adduct of some enzyme functional group with the Schiff base of PLP and a lysine side chain (structure A) or as a nonionic tautomer of a Schiff base in a hydrophobic environment (structure B, Eq. 23-24). For structure A, the fluorescent emission would be expected at a position similar to that of pyridoxamine. On the other hand, Schiff bases of the



(23-24)

type indicated by structure B would be expected to undergo a photoinduced proton shift (phototautomerization) to form structure C[124,187] with an absorption band at 430 nm (23,300 cm^{-1}) and fluorescent emission at a still lower energy. Since the observed fluorescence was at 530 nm, it was judged that the chromophore does have structure B.

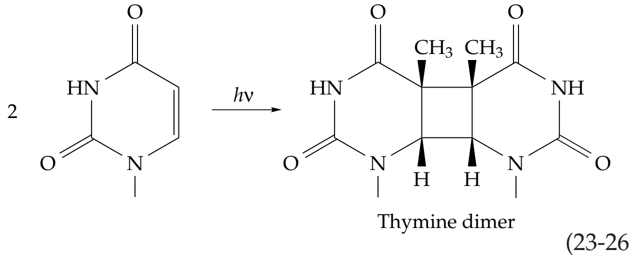
The rate of proton dissociations from the excited states of molecules can be measured directly by nanosecond fluorimetry.[189]

2. Photoreactions of Nucleic Acid Bases

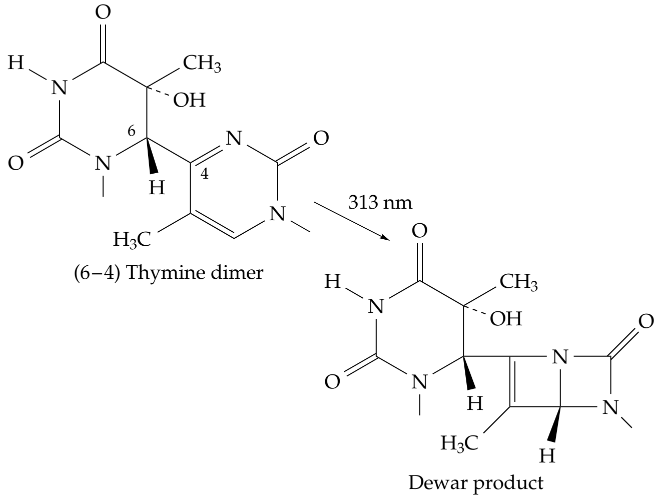
Photochemical reactions of the purines and pyrimidines assume special significance because of the high molar extinction coefficients of the nucleic acids present in cells. Light is likely to be absorbed by nucleic acids and to induce photoreactions that lead to mutations.[190] Both pyrimidines and purines undergo photochemical alterations, but purines are only about one-tenth as sensitive as pyrimidines. **Photohydration** of cytidine (Eq. 23-25) is observed readily. The reaction is the photochemical analog of the hydration of α,β-unsaturated carboxylic acids. Uracil derivatives also undergo photohydration.

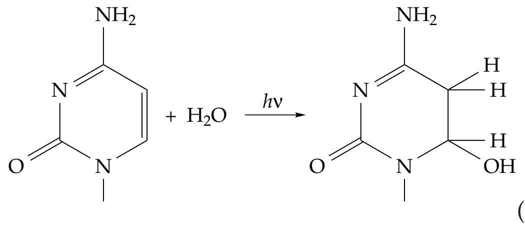
A more important reaction is the photodimerization of thymine (Eq. 23-26), a reaction also observed with uracil. A variety of stereoisomers of the resulting cyclobutane-linked structure are formed. The one shown in the equation predominates after irradiation of frozen thymine. Another important isomer is the 6–4 thymine dimer.[191,191a] Both of these types of **cyclobutane dimers** block DNA replication. This accounts for much of the lethal and mutagenic effect of ultraviolet radiation on organisms. The matter is sufficiently important that a special "excision repair" process is used by cells to cut out the thymine dimers (Chapter 27, Section E). In addition, light-dependent **photolyases**, discussed in Section I, act to reverse the dimerization reactions.

Light can also cause addition and other crosslinking reactions between DNA and proteins or other cell constituents.[192,193] One use of such reactions in the laboratory is DNA "photo footprinting" (Fig. 5-50), a technique which reveals contact regions between DNA and associated proteins.[190] Another type of cellular damage is caused by photosensitization of DNA by a light-absorbing intercalating agent such as **8-methoxypsoralen**.[194,195] DNA as well as adjacent proteins can be damaged.

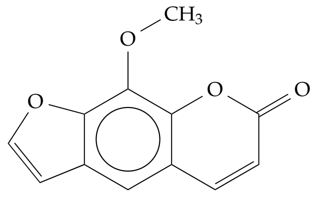


(23-26)





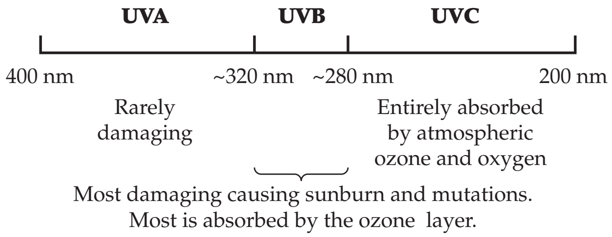
(23-25)



8-Methoxypsoralen

3. Sunburn, Cancer, and Phototherapy

Ultraviolet light is sometimes classified according to its energy and capacity for damaging cells as follows:[195a]


UVA
UVB
UVC
400 nm
~320 nm
~280 nm
200 nm
Rarely damaging
Entirely absorbed by atmospheric ozone and oxygen
Most damaging causing sunburn and mutations. Most is absorbed by the ozone layer.

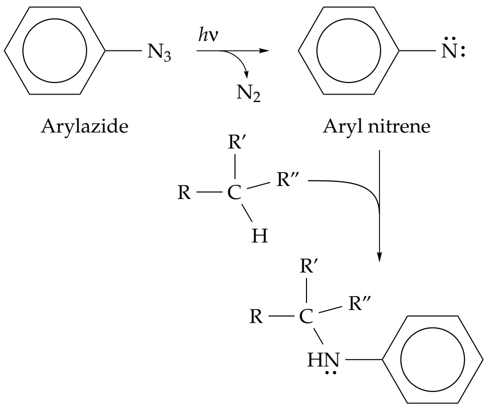
The UVB radiation causes most damage to skin. UVA radiation is at least an order of magnitude less damaging and is usually harmless to human skin. The UVC solar radiation is all absorbed by ozone and dioxygen of the atmosphere. UVC radiation produced by ultraviolet lamps is usually all absorbed in the epidermis.[5]

Ultraviolet light damages proteins as well as DNA. Residues of Trp, Tyr, His, Cys, and Met are especially susceptible to photolysis, or photooxidation by O_2, or by singlet oxygen. Also damaged are unsaturated lipids, porphyrins, flavins, etc. Kynurenic acid (Fig. 25-11) and urocanic acid (Eq. 14-44), an important ultraviolet filter in skin,[196] are also decomposed by light.

Repeated sunburn ages skin and may induce cancer.[4] However, light also has beneficial effects. It allows us to see, provides a source of vitamin D (Box 22-C), induces enzymatic repair of some DNA damage, and provides all of our food energy, directly or indirectly, by photosynthesis.

Light is used in **phototherapy**. This happens most frequently in the irradiation of newborn babies with white light to isomerize bilirubin (Fig. 24-24) from the 4*Z*, 15*Z* form to more readily excreted forms such as the 4*E*, 15*Z* isomer. About half of newborns have some **jaundice** (elevated bilirubin), and if it is severe it must be treated promptly to avoid neurological damage (see Chapter 24). There are sometimes complications, but the treatment is usually effective.[5] **Psoriasis** is frequently treated by irradiation with UVB light, which is thought to inhibit growth of the abnormal skin cells. Ten to 35 treatments are usually required, and the condition may return after some years. An alternative treatment is irradiation with UVA light after ingestion of 8-methoxypsoralen or other psoralen derivative.[5] Another skin condition that may respond to phototherapy is **vitiligo** (Box 25-A). Treatment with UVA and psoralen derivatives may stimulate repigmentation. If depigmentation is extreme, the remaining pigmentation may be reduced by bleaching with the monobenzyl ester of hydroquinone.

Photodynamic therapy is a cancer treatment that involves intravenous injection of a light-absorbing molecule such as a porphyrin, which may be taken up preferentially by cancer cells. Laser irradiation by deeply penetrating red light (650–800 nm wavelength) causes oxygen-dependent photosensitization.[5,197,198] Improvements in lasers, in fiber optics, and in photosensitizers may lead to widespread use of this type of therapy both for cancer and for some other conditions.[5,199]



(23-27)

4. Photoaffinity Labels

Photochemically reactive molecules have often been used as labels for specific sites in proteins and nucleic acids. Psoralen derivatives serve as relatively nonspecific photochemically activated crosslinking agents for DNA and double-stranded RNA.[195] **Aryl azides** are converted by light to aryl nitrenes, which react in a variety of ways including insertion into C–H bonds (Eq. 23-27).[200,201] In some cases UV irradiation can be used to join natural substrates to enzymes or hormones to receptors. For example, progesterone, testosterone, and other steroids have been used for direct photoaffinity labeling of their receptors.[202] Synthetic **benzophenones** have also been used widely as photoactivated probes.[203]

5. Microphotolysis and Ultrafast Light-Induced Reactions

Fluorescence microphotolysis, or photobleaching, has been widely used to study translational mobility of lipids and proteins in membranes. An attenuated laser beam may be focused down to the diameter of a cell or less. Then the intensity can be suddenly increased by several orders of magnitude, bleaching any fluorescent material present. The return of fluorescent material by free diffusion from a neighboring region (**fluorescence recovery after photobleaching**) or by diffusion through a membrane into a cell can then be observed.[204,205] Diffusion coefficients of labeled biopolymers or of components of cells can be evaluated, and translation and metabolism of lipids and other components can be followed.[206]

Laser-based techniques are being used for ultrafast observation of the results of a photochemical process, e.g., the light-induced dissociation of CO from the hemoglobin • CO complex. A dissociating laser pulse can be as short as 100 fs (0.1 ps) or less. This is shorter than the time of vibrational motion of nuclei in an electronically excited state (~0.3 ps). Using IR spectroscopy, events that follow can be observed at intervals as short as 0.1 ns.[207] X-ray diffraction measurements using 150 ps pulses have allowed direct observation of the CO dissociated from hemoglobin or myoglobin and its recombination with the same protein.[208–210] Femtosecond dynamics of electron transfer along a DNA helix is also being studied.[211,212]

6. Optical Tweezers, Light-Directed Synthesis, and Imaging

The radiation pressure exerted by light is very weak. A bright laser beam of several milliwatts of power can exert only a few piconewtons (pN) of force. However, a force of 10 pN is enough to pull a cell of *E. coli* through water ten times faster than it can swim.[213] In about 1986, it was found that a laser beam focused down to a spot of ~ one λ (~1 μm for an infrared laser) can trap and hold in its focus a refractile bead of ~1 μm diameter. This "optical tweezers" has become an important experimental tool with many uses.[213,214] For example, see Fig. 19-19. Not only are optical tweezers of utility in studying biological motors but also mechanical properties of all sorts of macromolecules can be examined. For example, DNA can be stretched and its extensibility measured.[215] Actin filaments have even been tied into knots![216]

Light-directed solid state synthesis of peptides[217] and oligonucleotides is another new tool.[218] Development of this technology may provide new advances in preparation of DNA chips (Chapter 26) with a higher density of components than are now available.

X-rays and more recently NMR (MRI)-imaging have become well known to us. It might seem impossible to use visible light for a similar type of imaging. However, a laser beam can pass through a person's head. Is it possible, using computer-based techniques, to create an image from the emergent light? Efforts are being made to do exactly this.[219,220] **Optical coherence tomography** using backscattered infrared light and related fast techniques have higher resolution than MRI, computerized tomography (CT), or ultrasound.[220–221a] Ultrabright synchrotron radiation is also being used in **infrared microspectrophotometry**.[222,223]

E. Photosynthesis

The photochemical reduction of CO_2 to organic materials[224–228] is the basic source of energy for the biosphere. Nevertheless, the process is limited to a few genera of photosynthetic bacteria (Table 1-1), eukaryotic algae, and higher green plants. Photosynthetic bacteria include the distinctly different purple, green, and bluegreen (cyanobacteria) groups, each of which has a different array of photosynthetic pigments. However, the basic mechanism of transduction of solar energy into chemical energy is the same in all of the bacteria and in green plants.

As discussed in Chapter 17, photosynthesis involves the incorporation of CO_2 into organic compounds by reduction with NADPH with coupled hydrolysis of ATP. This is most often via the Calvin–Benson cycle of Fig. 17-14. In a few organisms a reductive tricarboxylic acid cycle is employed. The idea that the chloroplasts of plants or the pigmented granules of photosynthetic bacteria generate NADPH or reduced ferredoxin plus ATP (Chapter 17) is now thoroughly accepted. However, it was not always obvious. Consider the overall equation (Eq. 23-28) for formation of glucose by photosynthesis in higher plants:

$$6\ CO_2 + 6\ H_2O \rightarrow 6\ O_2 + C_6H_{12}O_6 \qquad (23\text{-}28)$$

$$6\ CO_2 + 12\ H_2O^{*} \rightarrow 6\ O_2^{*} + C_6H_{12}O_6 + 6\ H_2O \qquad (23\text{-}29)$$

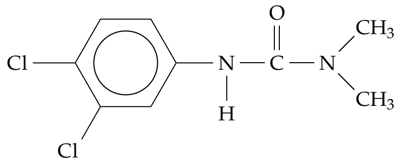
The stoichiometry of the reaction suggests that all 12 of the oxygen atoms of the evolved O_2 might come from CO_2 or that some might come from CO_2 and some from H_2O. In fact, water supplies both of the oxygen atoms needed for formation of O_2, as is indicated by the asterisks in Eq. 23-29. This possibility was suggested by van Niel[229,230] in 1931. He pointed out that in bacterial photosynthesis no O_2 is produced, and that bacteria must have access to a reducing agent to provide hydrogen for the reduction of CO_2 (Eq. 23-30).

$$H_2A \xrightarrow{hv} A + 2[H] \qquad (23\text{-}30)$$

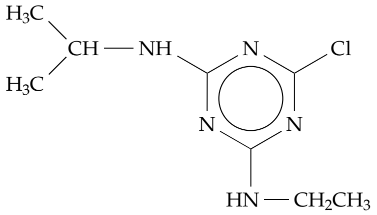
In this equation, H_2A might be H_2S (in the purple sulfur bacteria), elemental H_2, isopropanol, etc. From a consideration of these various reactions, van Niel concluded that in the O_2-producing cyanobacteria and eukaryotic plants water serves as the oxidizable substrate in Eq. 23-30 and is cleaved to form O_2 and to provide hydrogen atoms for reduction. This photochemical cleavage is the only known biological oxidation reaction of H_2O. No oxidizing agents present in living things are powerful enough to dehydrogenate water except for the photochemical **reaction centers** of photosynthetic organisms.

1. Two Photosystems, the Z Scheme, and Reaction Centers

It had long been known that for green plants light of wavelength 650 nm was much more efficient than that of 680 nm. However, Emerson and associates[34] in 1956 showed that a combination of light of 650 nm *plus* that of 680 nm gave a higher rate of photosynthesis than either kind of light alone. This result suggested that there might be two separate photosystems. What is now known as **photosystem I** (PSI) is excited by far red light (~700 nm), while **photosystem II** (PSII) depends upon the higher energy red light of 650 nm. Additional evidence supported the idea. Hill had shown many years before[231] that mild oxidizing agents such as ferricyanide and benzoquinone can serve as substrates for photoproduction of O_2, while Gaffron[232] found that some green algae could be adapted to photooxidize H_2 to protons (Eq. 23-30) and to use the electrons to reduce NADP. Thus, photosystem I could be disconnected from photosystem II. The powerful herbicide **dichlorophenyldimethylurea** (DCMU) was found to block electron transport between the two photosystems. In the presence of DCMU electrons from such artificial donors as ascorbic acid or an indophenol dye could be passed through photosystem I.



3-(3,4-Dichlorophenyl)-1,1-dimethylurea



Atrazine, another PSII inhibitor

The Z scheme. The result of these and other experiments was the development of the series formulation or zigzag scheme of photosynthesis[233] which is shown in Fig. 23-17. Passage of an electron through the system requires two quanta of light. Thus, four quanta are required for each NADPH formed and eight quanta for each CO_2 incorporated into carbohydrate.

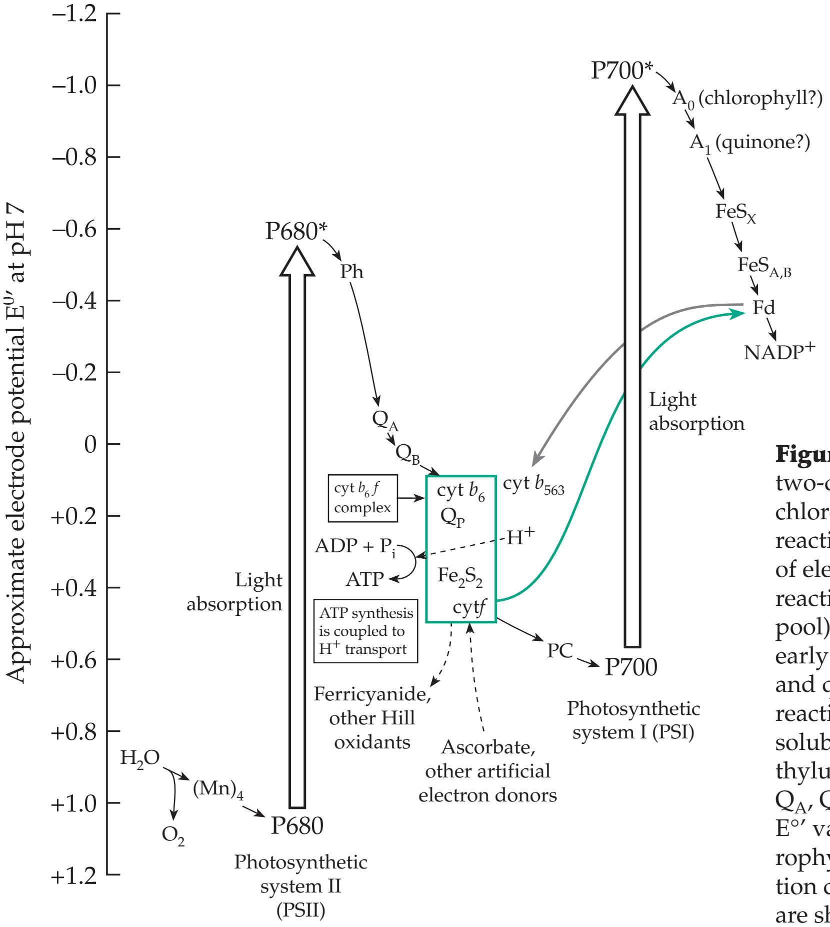

Approximate electrode potential E⁰′ at pH 7
−1.2
−1.0
−0.8
−0.6
−0.4
−0.2
0
+0.2
+0.4
+0.6
+0.8
+1.0
+1.2
P700*
A₀ (chlorophyll?)
A₁ (quinone?)
FeSₓ
FeS_A,B
Fd
NADP⁺
P680*
Ph
Q_A
Q_B
cyt b₆f complex
cyt b₆
Q_P
cyt b₅₆₃
H⁺
ADP + Pᵢ
ATP
Fe₂S₂
cytf
ATP synthesis is coupled to H⁺ transport
Light absorption
Light absorption
PC
P700
Photosynthetic system I (PSI)
Ferricyanide, other Hill oxidants
Ascorbate, other artificial electron donors
H₂O
(Mn)₄
O₂
P680
Photosynthetic system II (PSII)

Figure 23-17 The zigzag scheme (Z scheme) for a two-quantum per electron photoreduction system of chloroplasts. Abbreviations are P680 and P700, reaction center chlorophylls; Ph, pheophytin acceptor of electrons from PSII; Q_A, Q_B, quinones bound to reaction center proteins; PQ, plastoquinone (mobile pool); Cyt, cytochromes; PC, plastocyanin; A_0 and A_1, early electron acceptors for PSI, possibly chlorophyll and quinone, respectively; F_x, Fe_2S_2 center bound to reaction center proteins; F_A, F_B, Fe_4S_4 centers; Fd, soluble ferredoxin; and DCMU, dichlorophenyldimethylurea. Note that the positions of P682, P700, Ph, Q_A, Q_B, A_0, and A_1 on the $E^{\circ\prime}$ scale are uncertain. The $E^{\circ\prime}$ values for P682 and P700 should be for the (chlorophyll / chlorophyll cation radical) pair in the reaction center environment. These may be lower than are shown.

An important experiment of Emerson and Arnold[35] employed very short flashes of light and measurement of the quantum efficiency of photosynthesis during those flashes. A striking fact was observed. At most, during a single turnover of the photosynthetic apparatus of the leaf, one molecule of O_2 would be released for each 3000 chlorophyll molecules. However, it could be calculated that for each O_2 released only about eight quanta of light had been absorbed. It followed that about 400 chlorophyll molecules were involved in the uptake of one quantum of light. This finding suggested that a large number of chlorophyll molecules act as a single light receiving unit (usually called a light-harvesting or antenna complex) able to feed energy to one **reaction center**. The concept is now fully accepted.

Electron transport and photophosphorylation. Two molecules of NADPH are required to reduce one molecule of CO_2 via the Calvin–Benson cycle (Fig. 17-14), and three molecules of ATP are also needed. How are these formed? The Z scheme provides part of the answer. There is enough drop in potential between the upper end of PSI and the lower end of PSII to permit synthesis of ATP by electron transport. It is likely that only one molecule of ATP is formed for each pair of electrons passing through this chain. Since, according to Fig. 17-14, one and a half molecules of ATP are needed per NADPH, some other mechanism must exist for the synthesis of additional ATP. Furthermore, many other processes in chloroplasts depend upon ATP so that the actual need for photogenerated ATP may be larger than this.

Arnon[234,235] demonstrated that additional ATP can be formed in chloroplasts by means of **cyclic photophosphorylation**: Electrons from the top of PSI can be recycled according to the dashed lines in Fig. 23-17. An electron transport system, probably that of the Z scheme, is used to synthesize ATP. As isolation of proteins and cloning of their genes progressed, it became clear that a complex of proteins known as **cytochrome b_6f** is closely related to the cytochrome bc_1 of mitochondria (Fig. 18-8).[236–237a] As in that complex, cytochrome b_6 carries two hemes, designated b_h and b_l with E_m of −84 and −158 mv, respectively. E_m values are for *Chlamydomonas* (see Fig. 1-11). Heme b_l is closer to the positively charged membrane surface (lumen side) and heme b_h is closer to the negative surface (stroma side). In maize cytochrome b_6 is a 23-kDa subunit, and the c-type cytochrome *f* is a larger ~ 34-kDa subunit whose heme is close to the lumenal side. Its E_m value is +330 mv. A 20-kDa Rieske Fe–S protein (Chapter 16) and an additional 17-kDa subunit complete the core four-subunit complex, which is found in green plants, green algae,[236] and cyanobacteria.[238,239] Other smaller subunits are also present. It is usually assumed that a Q cycle equivalent to that of mitochondrial complex III (Fig. 18-9) operates in the pumping of protons across the thylakoid membrane (Fig. 23-18).[237b,c] However, at high rates of photosynthesis the electron transfer may bypass cytochrome *f*.[237d] The associated ATP synthase is also subject to complex regulatory mechanisms.[237e] The Rieske protein is encoded by a nuclear gene, but genes for other subunits are chloroplastic. Electrons may be carried from the cytochrome *f* subunit to PSI by plastocyanin[240] or, in many algae and cyanobacteria, by the small **cytochrome c_6**.[241–242b] It is often synthesized when copper is inadequate for synthesis of plastocyanin. Figure 23-18 is a schematic view of PSI, PSII, and the intermediate cyt b_6f complex in a thylakoid membrane.

In spite of the close similarities in structures and function, there are distinct differences between cyto-

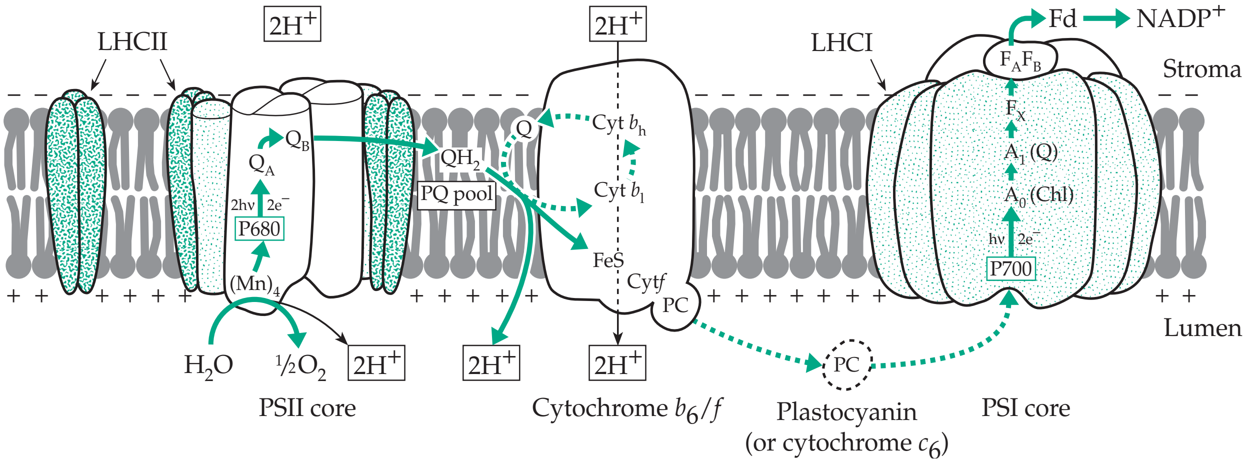


Figure 23-18 Schematic view of photosynthetic reaction centers and the cytochrome b_6f complex embedded in a thylakoid membrane. Plastocyanin (or cytochrome c_6 in some algae and cyanobacteria) carries electrons to the PSI core.

chrome b_6f and the cytochrome bc_1 of mitochondria.[243] Among these are the presence of stoichiometrically bound chlorophyll *a* and β-carotene[238,244] in the photosynthetic complex. The function of the chlorophyll is uncertain, but the carotene is probably there to quench the chlorophyll triplet state,[244] which would probably cause photodamage via formation of singlet O_2.

Bacterial photosynthesis. What is the relationship of the Z scheme of Fig. 23-17 to bacterial photosyntheses? In photoheterotrophs, such as the purple *Rhodospirillum*, organic compounds, e.g., succinate, serve as electron donors in Eq. 23-30. Because they can utilize organic compounds for growth, these bacteria have a relatively low requirement for NADPH or other photochemically generated reductants and a larger need for ATP. Their photosynthetic reaction centers receive electrons via cytochrome *c* from succinate ($E°'$ = + 0.03 V). The centers resemble PSII of chloroplasts and have a high midpoint electrode potential $E°'$ of 0.46 V. The initial electron acceptor is the Mg^{2+}-free bacteriopheophytin (see Fig. 23-20) whose midpoint potential is −0.7 V. Electrons flow from reduced bacteriopheophytin to menaquinone or ubiquinone or both via a cytochrome bc_1 complex, similar to that of mitochondria, then back to the reaction center P870. This is primarily a cyclic process coupled to ATP synthesis. Needed reducing equivalents can be formed by ATP-driven reverse electron transport involving electrons removed from succinate. Similarly, the purple sulfur bacteria can use electrons from H_2S.

In contrast, the reaction centers of green sulfur bacteria resemble PSI of chloroplasts. Their reaction centers also receive electrons from a reduced quinone via a cytochrome *bc* complex.[245] However, the reduced form of the reaction center bacteriochlorophyll donates electrons to iron−sulfur proteins as in PSI (Fig. 23-17). The latter can reduce a quinone to provide cyclic photophosphorylation. Cyanobacteria have a photosynthetic apparatus very similar to that of green algae and higher plants.

2. Chloroplast Structure

Chloroplasts come in various sizes and shapes, but all contain a small number of DNA molecules ranging in size from 120−160 kb. Complete sequences are known for DNA from chloroplasts of a liverwort (121,025 bp),[246] tobacco (155,844 bp),[247] maize (*Zea mays*),[248] and other plants. The 140,387 bp DNA from maize chloroplasts is a circular molecule containing the genes for 23S, 16S, 5S, and 4.5S RNA, for 30 species of tRNA, and for 70 different proteins. Among them are subunits of RNA polymerase, NADH dehydrogenase, subunits of both PSI and PSII, rubisco (large subunit), cytochromes *b* and *f*, six subunits of ATP synthase, and others. As with mitochondria, some subunits of the enzyme complexes that provide the cell with energy (e.g., ATP synthase) are encoded in the nucleus.

A characteristic of chloroplast genomes is the presence of a pair of large (in maize 22,348 bp) inverted repeat sequences. Since they can form a large hairpin structure with a very large loop, they may stabilize the gene sequence. The mutation rate within the repeat sequence is lower than in the single-stranded regions. The same genes are found in corresponding positions in both maize and rice. Both genomes have a number of sites of departure from the standard genetic code. These “editing sites” give rise to C→ U transitions in the RNA transcripts.

Most land plants have similar chloroplast DNA sequences, but considerable divergence is observed among algae.[249] For example, the red alga *Porphyra purpura* has 70 genes not found in chloroplasts of land plants. Each gene of the chloroplasts of the dinoflagellate *Heterocapsa triquetra* is carried on its own DNA minicircle.[250] However, ~2000 chloroplast proteins are encoded by nuclear DNA. The corresponding proteins are synthesized on cytoplasmic ribosomes and are transported into the chloroplasts.[249a] Some of these proteins must pass through both the double membrane of the envelope and the thylakoid membrane. As in mitochondria (Fig. 18-4) an array of different transport proteins are required. They are distinctly different from the mitochondrial transport proteins and involve their own unique targeting mechanisms.[251–253a]

Chloroplast membranes. Like the other energy-producing organelles, the mitochondria chloroplasts are surrounded by an outer double membrane or **envelope** and also contain an internal membrane system.[225–227,254–255a] Within the colorless **stroma** are stacks of flattened discs known as **grana** (Fig. 23-19). The discs themselves (the **thylakoids**) consist of pairs of closely spaced membranes 9 nm thick, each pair being separated by a thin internal space or loculus (Fig. 23-19). At least 75 different proteins are present in the isolated membranes. There is also a high content of **galactosyl diacylglycerol**, **digalactosyl diacylglycerol**, and **sulfolipid** (Chapter 8, Section A,4). Lipids account for half of the mass of thylakoid membranes.

Through the use of freeze-fracture and freeze-etching techniques of electron microscopy, it is possible to see, embeddded in the thylakoid membranes, particles which may represent individual **photosynthetic units** (also called quantosomes).[227,256–258] They are about 20 nm in diameter, and at least many of them presumably contain a reaction center surrounded by light-collecting chlorophyll−protein complexes. Others may represent the cytochrome b_6f complex and

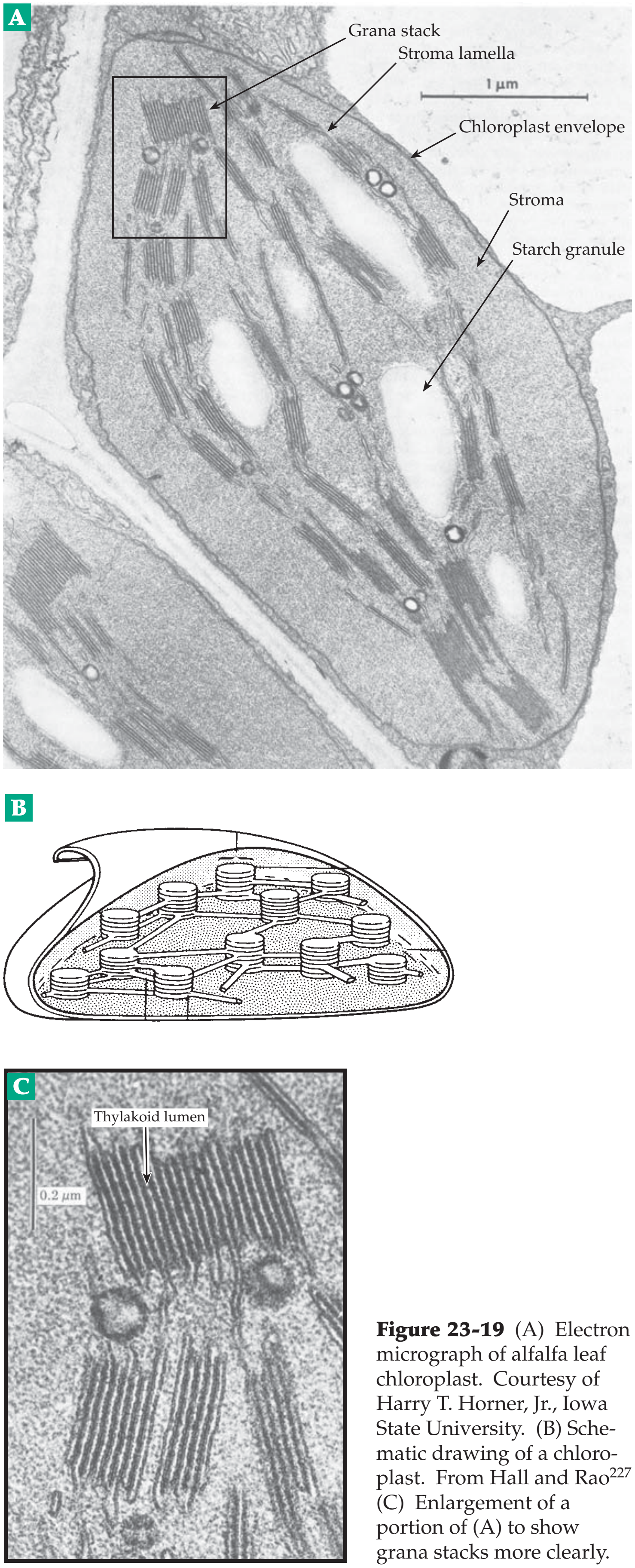
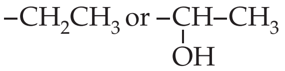


Figure 23-19 (A) Electron micrograph of alfalfa leaf chloroplast. Courtesy of Harry T. Horner, Jr., Iowa State University. (B) Schematic drawing of a chloroplast. From Hall and Rao[227] (C) Enlargement of a portion of (A) to show grana stacks more clearly.

ATP synthase, whose knobs also protrude into the stroma. A photosynthetic unit can also be defined chemically by the number of various types of molecules present in a chloroplast membrane for each four manganese atoms (Table 23-2). Separate units contain PSI and PSII. These reaction centers appear to have a different distribution within the thylakoids, the PSI units being located principally in the unstacked membranes and the PSII units in the grana stacks.[255,259]

Photosynthetic pigments and their environments. The chlorophylls (Fig. 23-20) are related in structure to the hemes (Figs. 16-5, 16-6), but ring IV (D) is not fully dehydrogenated as in the porphyrins. The **chlorin** ring system is further modified in chlorophyll by the addition of a fifth ring (V) containing an oxo group and a methyl ester. Ring V has been formed by crosslinking between the propionic acid side chain of ring III and a methine bridge carbon to give the parent compound **pheoporphyrin**. Chlorophylls contain constituents around the periphery that indicate a common origin with the porphyrins (Fig. 24-23). However, one of the carboxyethyl groups is esterified with the long-chain phytyl group in most of the chlorophylls. Chlorophyll *a* is the major pigment of chloroplasts and is a centrally important chromophore for photosynthesis in green plants. Most of the other chlorophylls, as well as the carotenoids and certain other pigments, are referred to as **accessory pigments**. Many of them have a light-receiving antenna function. Carotenoids are also photoprotectants. The relative numbers of pigment molecules in the photosynthetic units (average of PSI and PSII) of spinach chloroplasts are given in Table 23-2.

While the structure of chlorophyll *a* shown in Fig. 23-20 is the predominant one, other forms exist, e.g., with



$$-CH_2CH_3 \text{ or } -\underset{\displaystyle OH}{\underset{|}{CH}}-CH_3$$

replacing the vinyl group on ring I or with vinyl or hydroxyethyl replacing the ethyl group on ring II. The same kind of variation occurs for chlorophyll *b*.[260] In 80% acetone chlorophyll *a* has a sharp absorption band at 663 nm (15,100 cm^{-1}), but within chloroplasts the absorption maximum is shifted toward the red, the majority of the chlorophyll absorbing at 678 nm. Chlorophyll *b* (Fig. 23-20) is also nearly always present in green leaves.

The absorption peak in acetone is at 635 nm (15,800 cm^{-1}). Chlorophyll *c* found in diatoms, brown algae (Phaeophyta), and dinoflagellates (Fig. 1-9) lacks the phytyl group. Chlorophyll *d* contains a formyl group on ring I.[261]

Photosynthetic bacteria contain **bacteriochlorophylls** in which ring II is reduced (Fig. 23-20). The absorption band is shifted to the red from that of chlorophyll *a* to ~770 nm. The most abundant chlorophylls of green sulfur bacteria, **bacteriochlorophylls *c*, *d***, and ***e*** (or *Chlorobium* chlorophylls), contain a hydroxyethyl group on ring I; ethyl, *n*-propyl, or isobutyl groups on ring II; often an ethyl group instead of methyl on ring III; and a methyl group on the methine carbon linking rings I and IV. A variety of polyprenyl side chains can replace the phytyl group of the chlorophylls of higher plants.[262,263] The **pheophytins**, which are identical to the chlorophylls but lack the central magnesium ion, play an essential role in photosynthetic reaction centers. They can be formed in the laboratory by splitting the Mg^{2+} out from chlorophyll with a weak acid. Other derivatives are the **chlorophyllides** formed by hydrolysis of the methyl ester group and **chlorophyllins** formed by removal of both the methyl and phytyl groups.

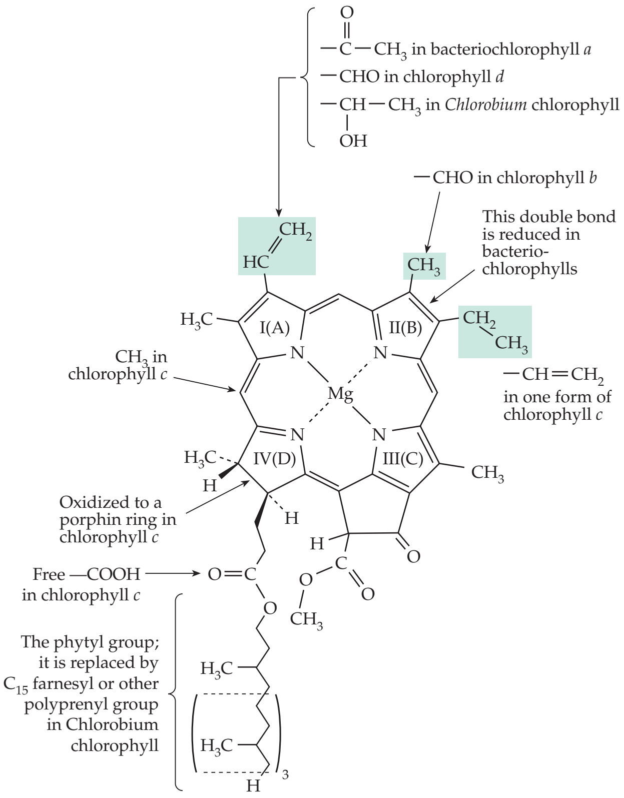
Since chlorophyll can be removed readily from chloroplasts by mild solvent extraction, it might appear that it is simply dissolved in the lipid portion of the membranes. However, from measurements of dichroism (Gregory,[226] p. 111) it was concluded that the chlorophyll molecules within the membranes have a definite orientation with respect to the planes of the thylakoids and are probably bound to fixed structures. The absorption spectrum of chlorophyll in leaves has bands that are shifted to the red by up to 900 cm^{-1} from the position of chlorophyll *a* in acetone. Most green plants contain at least four major chlorophyll bands at ~662, 670, 677, and 683 nm as well as other minor bands[264] (Fig. 23-21). This fact suggested that

TABLE 23-2
Approximate Composition of Photosynthetic Units in a Spinach Chloroplast[a]

Component	Number of molecules[b]
Chlorophyll *a*	160
Chlorophyll *b*	70
Carotenoids	48
Plastoquinone A	16
Plastoquinone B	8
Plastoquinone C	4
α-Tocopherol	10
α-Tocopherylquinone	4
Vitamin K_2	4
Phospholipids	116
Sulfolipids	48
Galactosylglycerides	490
Iron	12 atoms
Ferredoxin	5
Cytochrome b_{563}	1
Cytochrome b_{559}	
Cytochrome *f*	1
Copper	6 atoms
Plastocyanin	1
Manganese	2 atoms
Protein	928 kDa

[a] Averaged for PSI and PSII. After Gregory, R. P. F. (1971) *Biochemistry of Photosynthesis*, Wiley, New York [data of Luchtenthaler, H. K., and Park, R. B. (1963) *Nature* (*London*) **198**, 1070] and White, A., Handler, P., and Smith, E. L. (1973) *Principles of Biochemistry*, 5th ed., p. 528, McGraw-Hill, New York.

[b] Numbers of molecules assuming 2 Mn^{2+} ions per unit (4 for PSII and 0 for PSI).


O
—C—CH$_3$ in bacteriochlorophyll *a*
—CHO in chlorophyll *d*
—CH—CH$_3$ in *Chlorobium* chlorophyll
OH
—CHO in chlorophyll *b*
This double bond is reduced in bacterio-chlorophylls
CH$_2$
HC
CH$_3$
H$_3$C
I(A)
II(B)
CH$_2$
CH$_3$
—CH=CH$_2$ in one form of chlorophyll *c*
CH$_3$ in chlorophyll *c*
N
N
Mg
N
N
H$_3$C
IV(D)
III(C)
CH$_3$
H
H
Oxidized to a porphin ring in chlorophyll *c*
H
Free —COOH in chlorophyll *c*
O=C
O
C
O
O
CH$_3$
O
The phytyl group; it is replaced by C$_{15}$ farnesyl or other polyprenyl group in Chlorobium chlorophyll
H$_3$C
H$_3$C
3
H

Name of parent ring system:

without ring V closed: chlorin
with ring V closed: pheoporphyrin
with free — COOH on ring V: chlorophyllides
with free — COOH on both rings II and V: chlorophyllins
with Mg_{2+} replaced by 2 H^+: pheophytins

Figure 23-20 Structures of the chlorophylls.

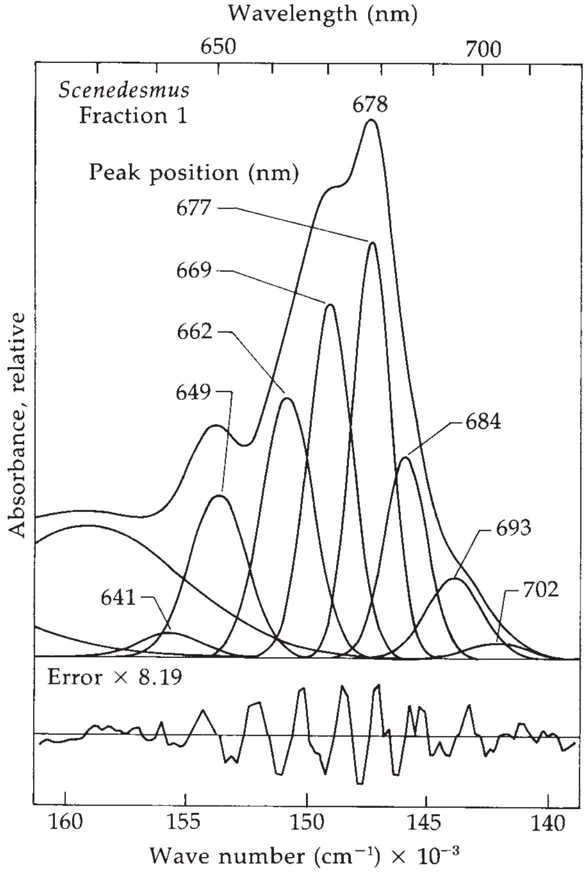


Figure 23-21 Absorption spectrum of chlorophyll in a suspension of chloroplast fragments from the green alga *Scenedesmus* showing the multicomponent nature of the chlorophyll environments. From French and Brown.[264]

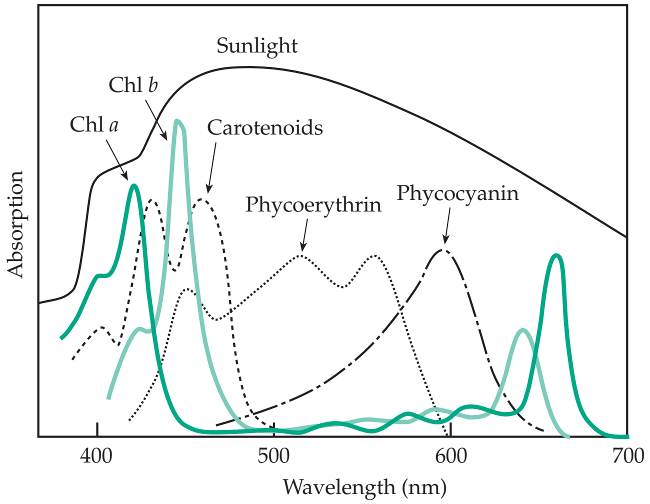


Figure 23-22 Absorption spectra of chlorophylls and accessory pigments compared. Redrawn from G. and R. Govindjee,[265] and from J. J. Wolken.[266]

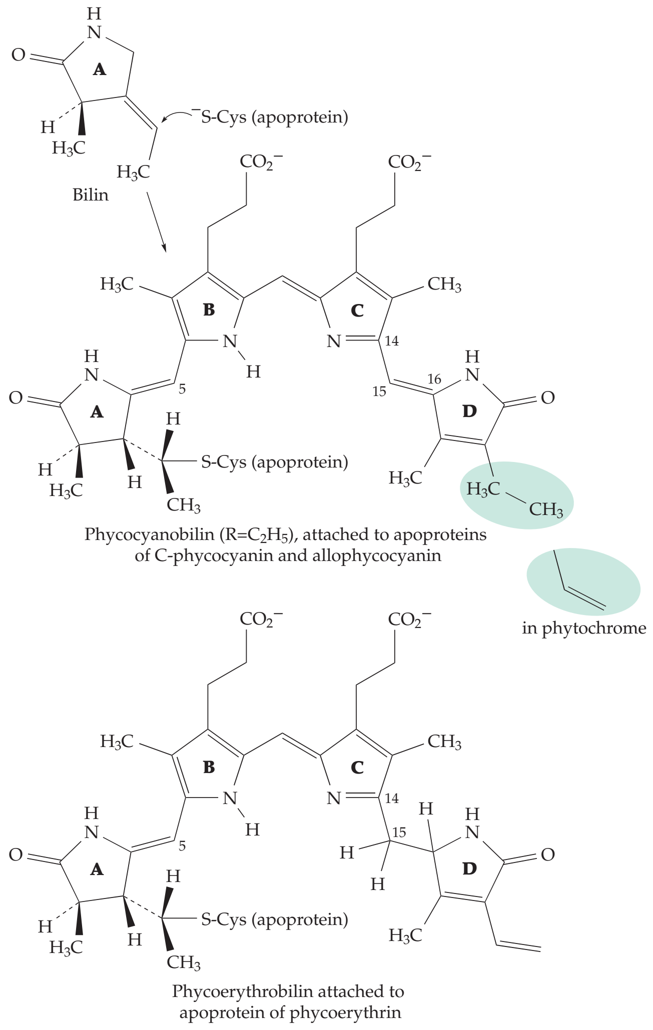
the chlorophyll exists in a number of different environments. As a result, the absorption is spread over a broader region leading to more efficient capture of light. Only a small fraction of the total chlorophyll is in the reaction centers; that for PSI absorbs at ~700 nm and that for PSII at ~682 nm.

Bacteriochlorophyll in *Chromatium* has three absorption bands with peak positions at 800, 850, and 890 nm. The last includes the reaction center bacteriochlorophyll and is the only form that fluoresces. Recent studies have established that most if not all chlorophyll is bound to specific proteins, a fact that can account for the various overlapping absorption bands.

The **carotenes** and **carotenoids** are very important accessory pigments (Fig. 23-22). The major component in most green plants is β-carotene. Green sulfur bacteria contain γ-carotene in which one end of the molecule has not undergone cyclization and resembles lycopene (Fig. 22-5). Chloroplasts also contain a large variety of oxygenated carotenoids (xanthophylls). Of these, neoxanthin, violaxanthin (Eq. 22-12), and lutein (p. 1240) predominate in higher plants and green algae. *Euglena* and related microorganisms contain much antheraxanthin (Eq. 22-10). A light-collecting protein from marine dinoflagellates contains both chlorophyll *a* and several molecules of the carotenoid peridinin. See Fig. 23-29.[267,268] Brown algae and diatoms contain mostly fucoxanthin and zeaxanthin (Fig. 22-5), while the bacterium *Rhodospirillum rubrum* synthesizes spirilloxanthin (p. 1240).

It is a striking fact that there are no naturally occurring green plants that lack carotenoid pigments.[269] Carotenoidless mutants are used in photosynthesis research, but they apparently cannot survive under natural conditions. Carotenoids not only participate as members of the light-receiving complex but also confer protection to chlorophyll against light-induced destruction by singlet oxygen. This accounts for the fact that carotenoids are usually intimately associated with chlorophyll in the pigment complexes. For example, see Figs. 23-29 and 23-30 and discussion on pp. 1308–1310.

A third class of accessory pigment of more limited distribution are the **open tetrapyrroles**, sometimes called "plant bile pigments" because of their relationship to the pigments of animal bile (Fig. 24-24). Among these are the **phycocyanins**, which provide the characteristic color to cyanobacteria. They are conjugated proteins (biliproteins) containing covalently bound phycocyanobilin (Fig. 23-24).[270] The red **phycoerythrins** of the Rhodophyta contain bound phycoerythrobilin (Fig. 23-23), an isomer of phycoerythrobilin. There are four common isomeric **bilins**, each having a different number of conjugated double bonds.[272,273] Together, they provide for a broad range



Phycocyanobilin (R=C_2H_5), attached to apoproteins of C-phycocyanin and allophycocyanin

Phycoerythrobilin attached to apoprotein of phycoerythrin

Figure 23-23 Structures of the open tetrapyrroles of plants. See also Fig. 24-24. After Szalontai *et al.*[271]

of colors from blue to red (Table 23-3). The bilins are attached to proteins via addition of a cysteine –SH group to the vinyl group of ring A of the bilin (Fig. 23-23). A double attachment may be formed by addition of –SH groups to both vinyl groups.[273–276] Isolated tetrapyrrole pigments tend to have a helical structure and to absorb light at lower wavelengths than do the protein-bound pigments which assume elongated conformations.

There are three major classes of conjugated phycobiliproteins,[273,277] all of which are αβ heterodimers often associated as $(\alpha\beta)_6$ (Fig. 23-24). The **allophycocyanins** carry one bilin per subunit, the phycocyanins carry one on the α and two on the β subunit, and the phycoerythrins carry two or three on the α subunit and three on the β (Fig. 23-24). Cysteine α-84 is one of the frequent attachment sites.[273] Three-dimensional structures are known for several of these proteins[278–281]

TABLE 23-3
The Common Bilin Pigments Present in Phycobiliniproteins[a]

Isomer	Number of conjugated double bonds	Absorption maximum (nm) when conjugated to proteins
Phycocyanobilin	8	~640
Phycobiliviolin	7	~590
Phycoerythrobilin	6	~550
Phycourobilin	5	~490

[a] Wedemayer, G. J., Kidd, D. G., Wemmer, D. E., and Glazer, A. N. (1992) *J. Biol. Chem.* **267**, 7315–7331.

The bilins are derived biosynthetically rather directly[282] from biliverdin IXα, whose formation is described in Fig. 24-24. The addition of an apoprotein –SH group to a carbon–carbon double bond of the bilin is catalyzed by a specific lyase.[283] **Phytochrome** (Section H) arises in a similar way[284] as does the blue biliprotein **insecticyanin** (Box 21-A).[285]

3. The Light-Receiving Complexes

Irradiation of chloroplasts leads to easily measurable fluorescence from chlorophyll *a*, but no fluorescence is observed from chlorophyll *b* or from other forms of chlorophyll, carotenoids, or other pigments. It appears that the latter all serve as light-collecting or antenna pigments that efficiently transfer their energy to chlorophyll *a* at the reaction centers.[286] As is evident from Fig. 23-22, the light-collecting pigments generally have higher energy absorption bands than do the reaction centers. Thus, a broad range of wavelengths of light are absorbed by an organism, and energy from all of them is funneled into the reaction centers. The light-collecting pigments are bound to specific proteins, which are located close to the reaction centers and are arranged to provide efficient energy-transfer. Distances between adjacent pigment molecules vary from 1 to 7 nm.[287]

Phycobilosomes. Algal and cyanobacterial phycocyanins and phycoerythrins are aggregated in special granules that are on the outsides of the photosynthetic membranes. The granules in the cyanobacteria are known as phycobilosomes (Fig. 23-24).[272,286,286a,288] The $(\alpha\beta)_6$ hexamers form the disks of the phycobilosomes. These are held together by linker proteins,[281,289] which fit asymmetrically into the central cavities. As is indicated in Fig. 23-24C, the disks and linker proteins are assembled into rods which are joined to form the phycobilosomes. The latter are organized into closely packed parallel arrays on the surface of the photosynthetic membranes.

Purple photosynthetic bacteria. The reaction centers of *Rhodobacter spheroides, Rhodospirillum rubrum*, and related purple bacteria are embedded in the plasma membrane. Each center is surrounded by a ring of bacteriochlorophyll *a* molecules bound noncovalently to heterodimeric (αβ) protein subunits made up of ~52- to 54-residue chains. Each αβ dimer binds two molecules of BChl *a*, whose central Mg^{2+} ions are coordinated by conserved histidine imidazole groups, as well as a molecule of spirilloxanthin. About 15–17 αβ subunits form the ring, which is designated LH1 (Fig. 23-25A).[291–293] Most of these bacteria also have smaller rings, designated LH2, floating in the membrane near the LH1 complex. The LH2 rings (Figs. 23-25B, C) consist of about nine αβ subunits with associated BChl *a* and carotenoid.[294–298] Under some conditions a third complex LH3 may be formed.[298a] In *Rhodopseudomonas acidophila* nine of the 27 BChl *a* molecules absorb light maximally at ~800 nm and are designated **B800**. The other 18, designated **B850**, absorb maximally at ~860 nm.[294,299] The B850 BChl *a* molecules have direct contact with the chromophores of neighboring molecules, allowing for easy energy transfer. The B800 chromophores are more isolated.

Low-temperature (1.2 K) single-molecule spectroscopic techniques have been used to obtain the fluorescence-excitation spectra shown in Fig. 23-26C. For an ensemble of LH2 complexes (upper trace) the spectral absorption bands are broad, but for individual LH2 complexes structure can be seen clearly for the B800 chromophores but not for the B850 chromophores. This difference has been interpreted to mean that the excitation energy of an electronically excited B850 molecule is delocalized over the whole ring of 18 BChl *a* molecules as an **exciton**. This permits both fast and efficient energy transfers from B800 to B850 and from B850 of one LH2 ring to another or to an LH1 ring and to the reaction center (Fig. 23-27).[299–300a] Energy transfer may occur by the Förster dipole–dipole mechanism (Section C,2).[298] Many of the antenna are supported by binding through their Mg^{2+} ion to an imidazole group of a protein as can also be seen in reaction center chlorophylls (Fig. 23-31C). Hydrogen-bonding to the C13-oxo groups of the chlorophylls may also be possible.[300b] The orientations of the transition dipole moments of the chlorophyll molecules may be arranged to facilitate rapid energy transfer.[300c]

Green sulfur and nonsulfur bacteria. In these organisms chlorophylls are present in rodlike particles

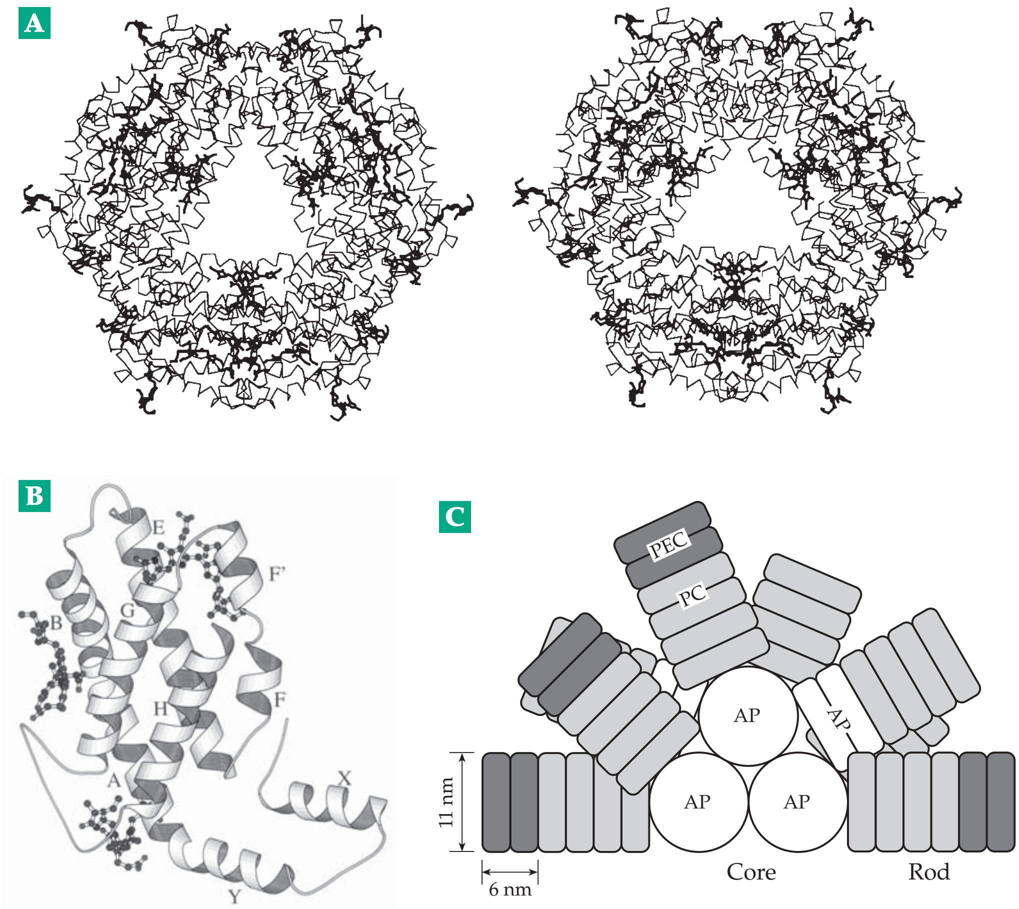


Figure 23-24 (A) Stereoscopic view of a hexameric $(\alpha\beta)_3$ phycobiliprotein. (B) The β subunit of the complex with two molecules of bound phycoerythrobilin and one of phycourobilin. From Chang *et al.*[279] (C) Schematic representation of a phycobilosome of a strain of the cyanobacterium *Anabaena*. Each disk in the structure contains an $(\alpha\beta)_3$ phycobiliprotein. The circles marked AP are cross-sections of rods, each one composed of about four disks of allophycocyanin (AP). The projecting rods contain C-phycocyanin (PC) and phycoerythrocyanin (PEC). From Lao and Glazer.[290]

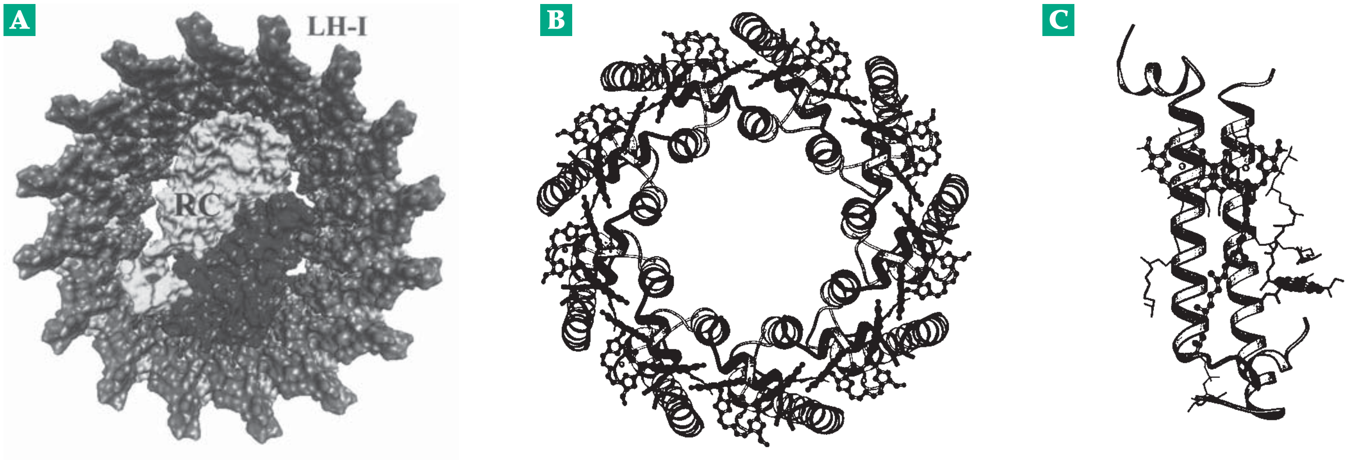


Figure 23-25 (A) The van der Waals contact surface of the periplasmic face of the reaction center and surrounding light-harvesting complex LH1 of *Rhodobacter sphaeroides*. Made with VMD by Theoretical Biophysics Group, UIUC. See also Hu and Schulten.[291] (B) Ribbon drawing of the structure of the circular light-harvesting complex LH2 of the purple photosynthetic bacterium *Rhodopseudomonas acidophila*. The tetrapyrrole rings of the 18 bacteriochlorophyll molecules are also shown. (C) Structure of one of the nine αβ protomers with three associated bacteriochlorophylls. One of these is near the top of the protein, and the other has its chromophoric group protruding at nearly a right angle on the right side. The complete phytyl side chains are also depicted in a stick representation. (B) and (C) are from Prince *et al.*[295] MolScript drawings courtesy of N. W. Isaacs.

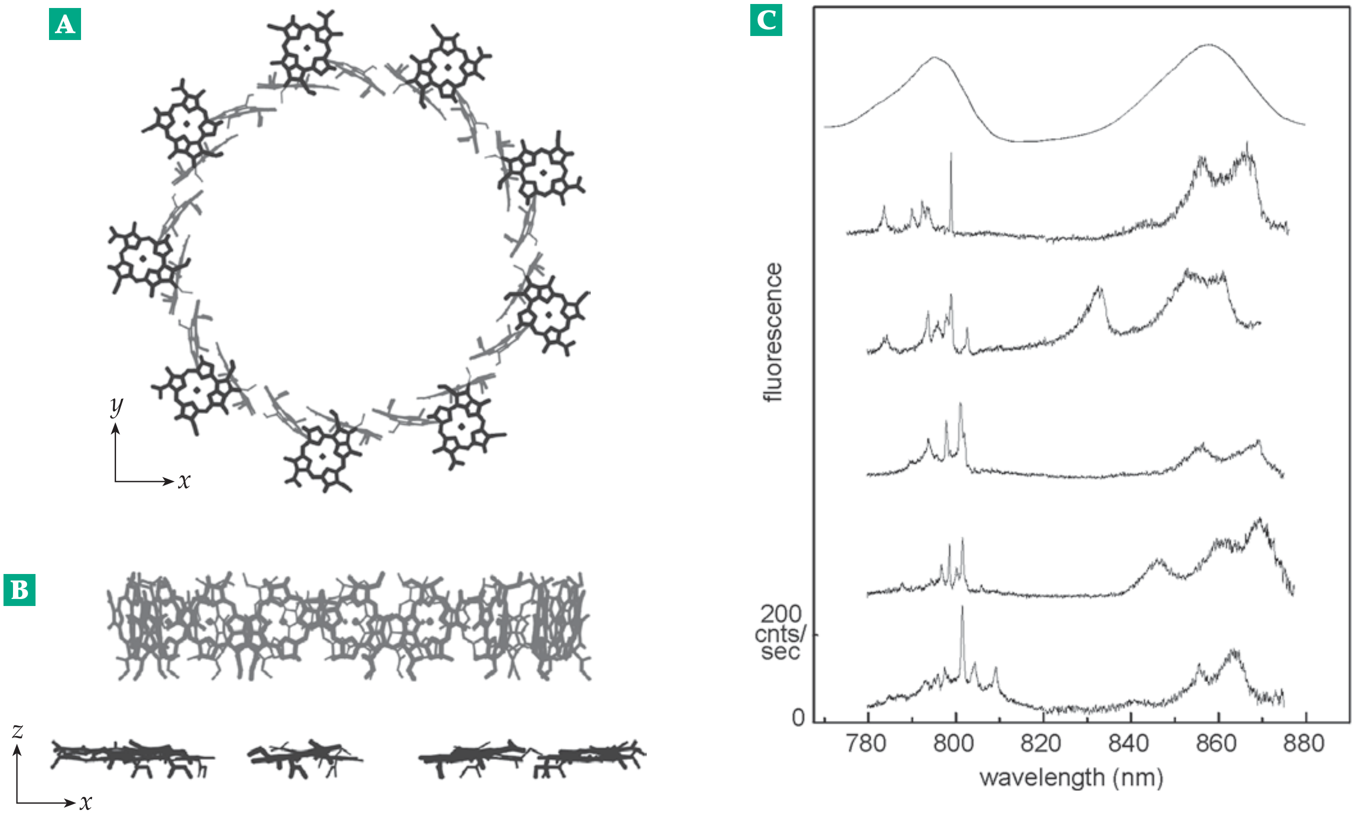


Figure 23-26 (A), (B) Arrangement of bacteriochlorophyll chromophores in the cyclic LH2 array of *Rhodopseudomonas acidophila*. The B850 subunits are gray while the B800 subunits are black. (C) Fluorescence-excitation spectra. Top trace, for an ensemble of LH2 complexes, other traces, for several individual LH2 complexes at 1.2K. Fine structure is evident for the B800 but not for the B850 chromophores. From van Oijen *et al.*[299] with permission.

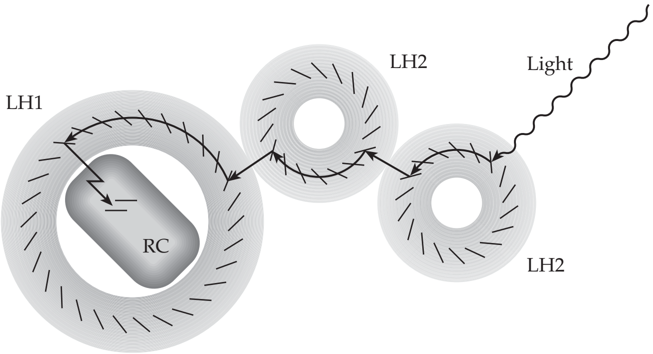


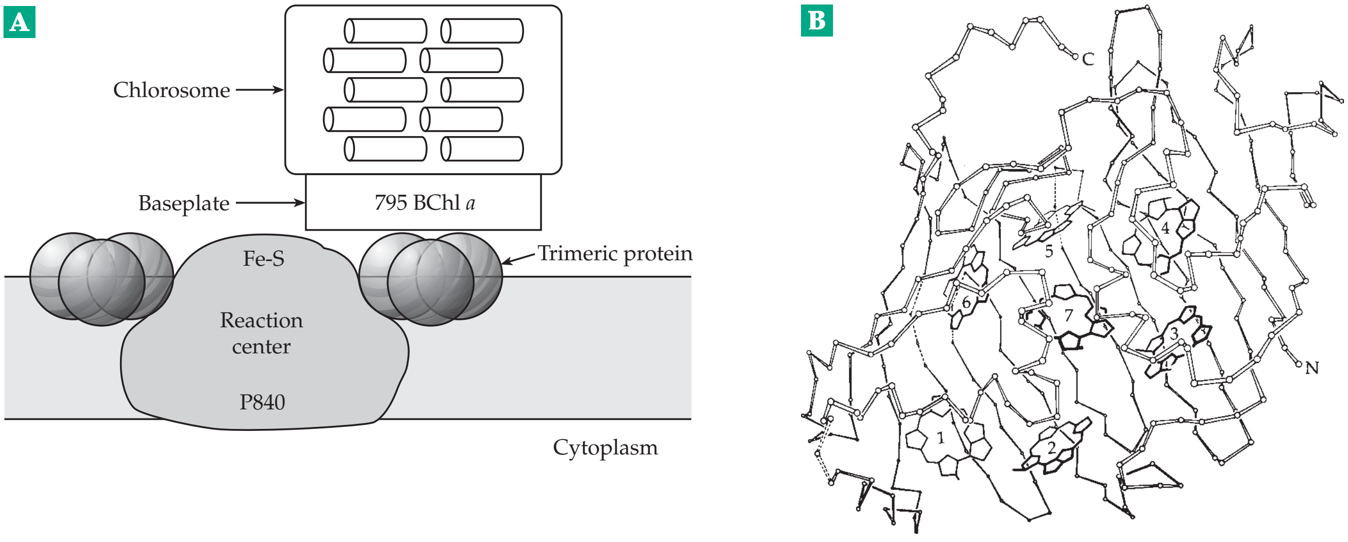
Figure 23-27 Illustration of proposed exciton transfer of the energy of light absorbed by bacteriochlorophyll *a* of purple bacteria. Energy absorbed by the light harvesting complex LH2 is transferred in steps to another LH2, to LH1 and to the reaction center. The short lines within the circles represent the edges of the BChl*a* chromophores. After Kühlbrandt[300] with permission.

of protein present within **chlorosomes**, baglike structures which may be as large as 100 x 260 nm and are attached to the inside of the cytoplasmic membranes,[301–302a] which contain the reaction centers (Fig. 23-28). The over 10,000 light-collecting chlorophylls *c*, *d*, or *e*, which may be present per reaction center, allow these bacteria to live in extremely weak light. The bacteria also contain a smaller "antenna" of ~5000 BChl *a* present as a complex with a water-soluble protein. The three-dimensional structure of this protein is also shown in Fig. 23-28. Each 45-kDa subunit of the trimeric protein contains seven embedded molecules of bacteriochlorophyll *a*.[302,303] Other light-collecting chlorophyll–protein complexes may contain an even higher ratio of chlorophyll to protein.[286]

Eukaryotic plants and cyanobacteria. Photosynthetic dinoflagellates, which make up much of the marine plankton, use both carotenoids and chlorophyll in light-harvesting complexes. The carotenoid **peridinin** (Fig. 23-29), which absorbs blue-green in the 470- to 550-nm range, predominates. The LH complex of *Amphidinium carterae* consists of a 30.2-kDA protein that forms a cavity into which eight molecules of peridinin but only two of chlorophyll a (Chl *a*) and two molecules of a galactolipid are bound (Fig. 23-29).[268]

The allenic carotenoid **fucoxanthin** (Fig. 22-5), which is absent in higher plants, predominates in brown algae, where it occurs in light-harvesting complexes along with Chl *a* and Chl *c*.[306,307]

A family of Chl *a*/*b* binding proteins are found in green plants.[308] These have apparently evolved inde-



Monomer of trimeric BChl protein

Figure 23-28 (A) Model of a light-harvesting chlorosome from green photosynthetic sulfur bacteria such as *Chlorobium tepidum* and species of *Prosthecochloris*. The chlorosome is attached to the cytoplasmic membrane via a baseplate, which contains the additional antenna bacteriochlorophylls (795 BChl *a*) and is adjacent to the trimeric BChl protein shown in (B) and near the reaction center. After Li *et al.*[302] and Rémigy *et al.*[304] (B) Alpha carbon diagram of the polypeptide backbone and seven bound BChl *a* molecules in one subunit of the trimeric protein from the green photosynthetic bacterium *Prosthecochloris*. For clarity, the magnesium atoms, the chlorophyll ring substituents, and the phytyl chains, except for the first bond, are omitted. The direction of view is from the three-fold axis, which is horizontal, toward the exterior of the molecule. From Fenna and Matthews.[305] See also Li *et al.*[302]

pendently of chlorophyll-binding proteins of green bacteria.[309] Quantitatively most important is the complex known as LHCII, the major Chl *a* / *b* protein associated with PSII and which may also provide energy to PSI. This one protein, whose structure is shown in Fig. 23-30,[310] is thought to bind half of all of the chlorophyll in green plants. The protein is organized as trimers.[311] Each 232-residue monomer binds 5–6 Chl *b*, 7–8 Chl *a*, ~ two molecules of lutein, and one of neoxanthin.[310,311] LHCII also carries all four characteristic thylakoid lipids: mono- and digalactosyl diacylglycerols, phosphatidylglycerol, and sulfoquinovosyl diacylglycerol.[312] The Chl *a* and Chl *b* molecules are in close contact (Fig. 23-30). Subpicosecond transient absorption spectroscopy[311] indicates that half of the Chl *b* to Chl *a* energy transfers occur in < 0.2 ps. Notice the close association of the two luteins in Fig. 23-30 with the chlorophyll rings. The carotenoids are thought to quench chlorophyll triplet states to prevent formation of singlet oxygen.

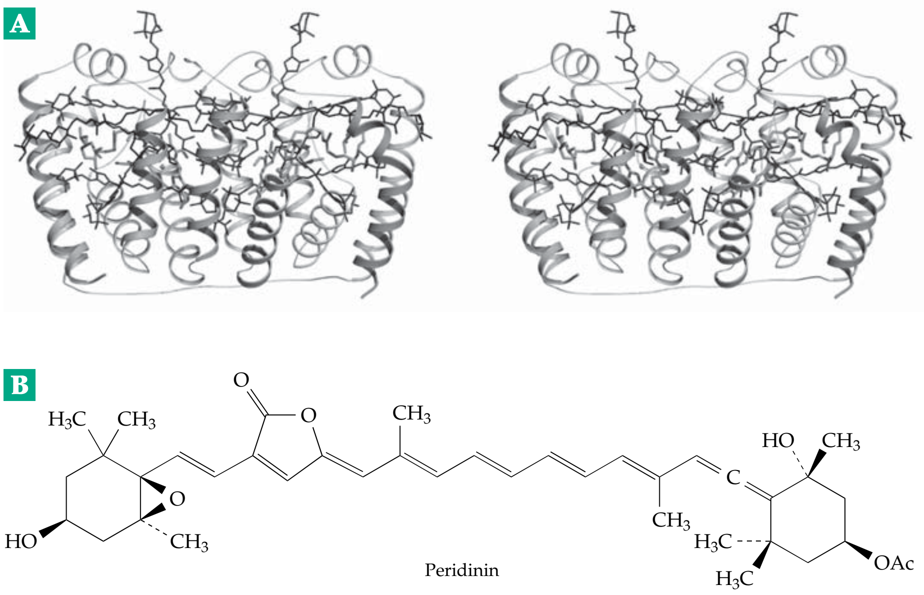


Figure 23-29 (A) Stereoscopic drawing of light-harvesting complex from the dinoflagellate protozoan *Amphidinium carterae*. The central cavity contains eight molecules of peridinin, two of which can be seen protruding from the top. Deeply buried toward the bottom are two molecules of Chl *a*. Also present are two molecules of digalactosyl diacylglycerol. From Hofmann *et al.*[268] Courtesy of Wolfram Welte. (B) Structure of peridinin.

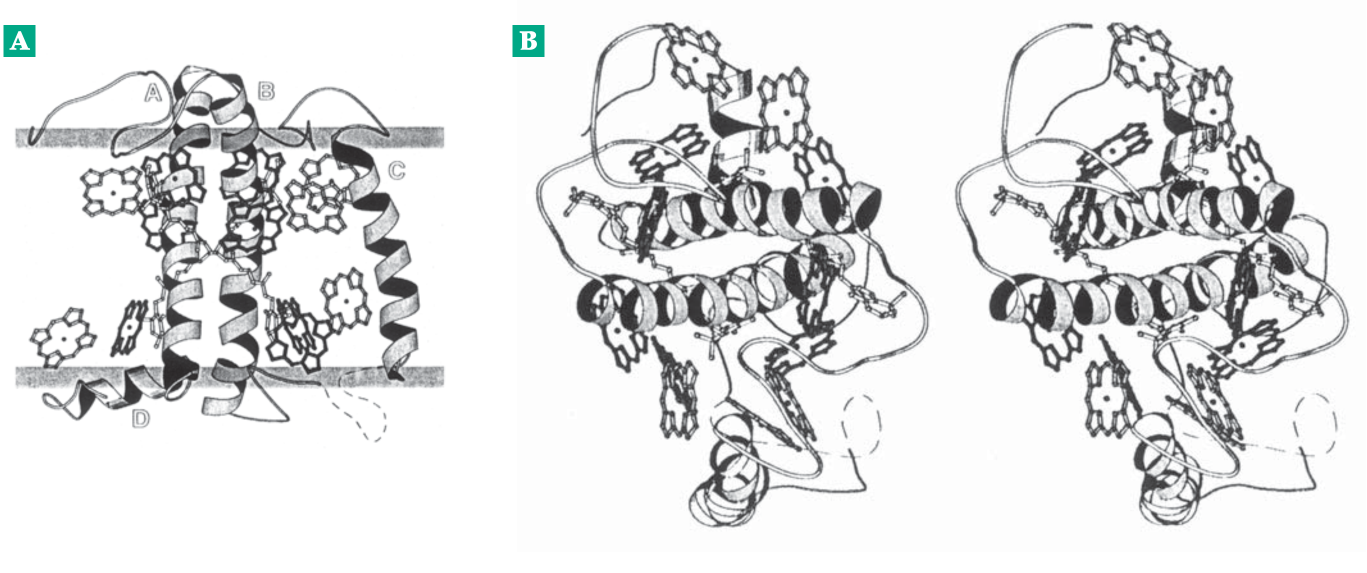


Figure 23-30 Views of light-harvesting protein LHCII of green plants. (A) Side view indicating the approximate position in the lipid bilayer of the thylakoid membrane. Helices are labeled A–D. (B) Stereoscopic top view from the stromal side of the membrane. The structure, at 0.34 nm resolution, was determined by electron crystallography on highly ordered two-dimensional crystals. MolScript drawings from Kühlbrandt *et al.*[310] Courtesy of Werner Kühlbrandt.

PSII also contains several additional chlorophyll-binding proteins, designated CP24, CP26, CP29, CP43, CP47, etc. These lie on both sides of a pair of reaction-center cores.[313–315a] A large fraction of the LHCII complexes are separate from the reaction centers and are mobile, while a smaller fraction are bound to the outer ends of the core complex.[308,314,315] The light-harvesting chlorophylls of PSI are arranged around the core. Some are bound in the N-terminal part of the major core subunits, the products of genes *PsaA* and *PsaB*.[316] They bind ~90 Chl *a* and 14 β-carotene molecules per reaction center. (See Fig. 23-33.) In addition, the peripheral LHCI, composed of four different proteins arranged around the core, binds ~110 Chl *a* and Chl *b* and ~70 molecules of xanthophyll.[317] Some species of cyanobacteria use antenna rings around their PSI trimers instead of phycobilinosomes.[317a]

In every case the light-harvesting complexes are arranged to allow rapid and very efficient transfer of electronic excitation energy from one chromophore to another and finally to the chromophores of the reaction center.[317b] The speed and efficiency appear to depend upon very rigid structures of the proteins and precise orientations of the bound chromophores to allow direct excitonic transfer of energy at distances of less than 2 nm or transfer by the Förster mechanism at distances not exceeding 10 nm.[302] An example of the precision of protein structures was observed when a posttranslationally modified asparagine N^5-methylasparagine at position 72 of the β subunit of many phycobiliniproteins was substituted by aspartate or glutamine. The fluorescence lifetime of the nearby bilin was reduced 7–10% in the mutants, an effect that could cut the >95% efficiency of energy transfer from the phycobilinisomes to the PSII reaction center.[318]

4. The Reaction Centers and Their Photochemistry

The initial or primary processes of photosynthesis occur in the reaction centers in which chlorophyll or bacteriochlorophyll absorbs a photon.[318a] Then, the chlorophyll, in its singlet excited state (Chl*), donates an electron to some acceptor A to form a radical A^- and to leave an oxidized chlorophyll Chl^+ radical (Eq. 23-31).

$$\mathrm{Chl} \xrightarrow{h\nu} \mathrm{Chl}^* \xrightarrow{\mathrm{A}} \mathrm{Chl}^+ + \mathrm{A}^- \qquad (23\text{-}31)$$

In the scheme of Fig. 23-18, acceptor A is Q_A for PSII and A_0 for PSI. The oxidized chlorophyll (Chl^+) quickly reacts further by receiving an electron from some donor.

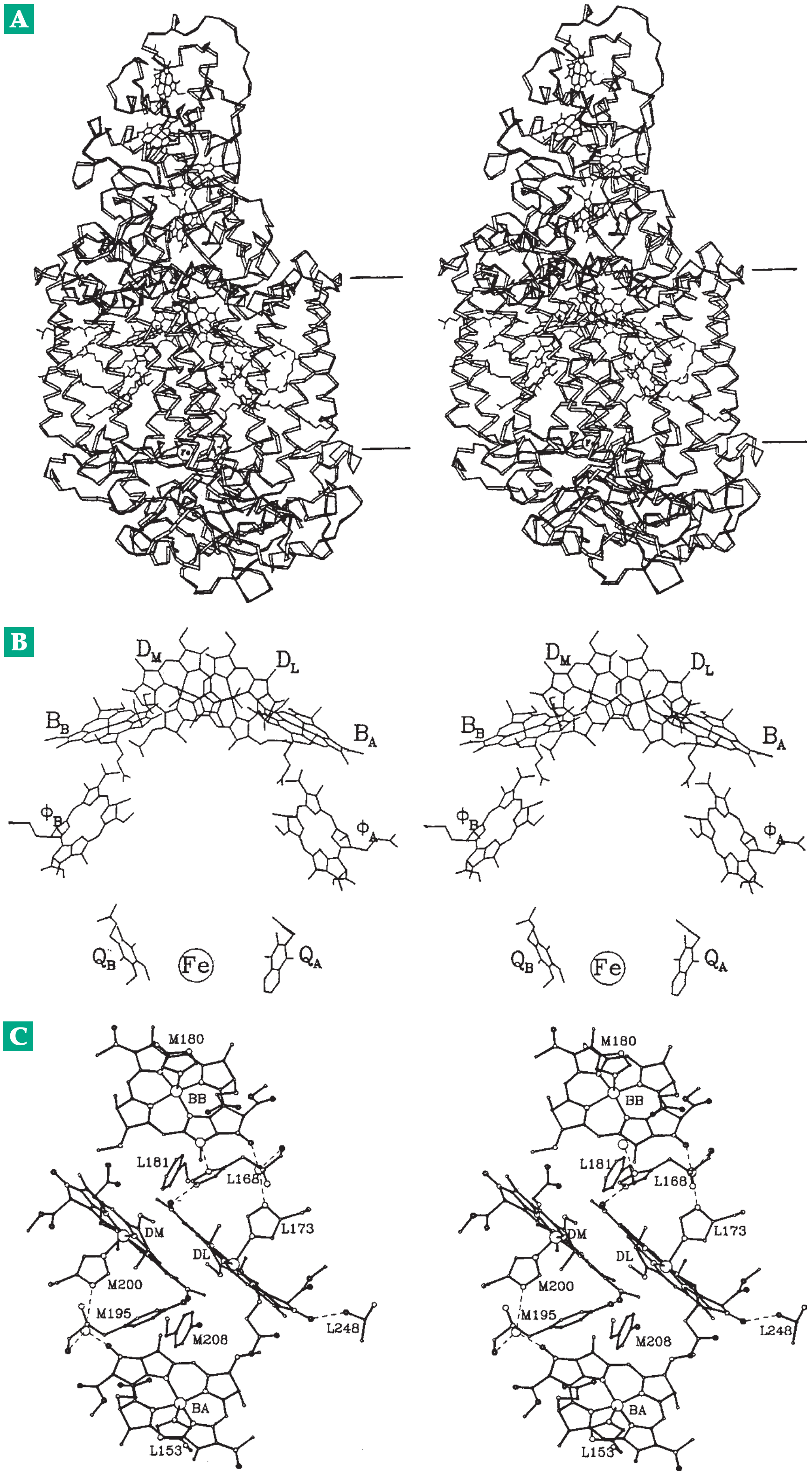
The photooxidation of chlorophyll indicated in Eq. 23-31 is accompanied by bleaching in the principal light absorption band. However, since there is so much light-gathering chlorophyll for each reaction center, the effect is small. The study of the process has been aided greatly by preparation of isolated bacterial photochemical reaction centers.

Reaction centers of purple bacteria. The exact composition varies, but the properties of reaction centers from several genera of purple bacteria are similar. In *Rhodopseudomonas viridis* there are three peptide chains designated H, M, and L (for heavy, medium and light) with molecular masses of 33, 28, and 24 kDa, respectively. Together with a 38-kDa tetraheme cytochrome (which is absent from isolated reaction centers of other species) they form a 1:1:1:1 complex. This constitutes reaction center P870. The three-dimensional structure of this entire complex has been determined to 0.23-nm resolution[288,319–323] (Fig. 23-31). In addition to the 1182 amino acid residues there are four molecules of bacteriochlorophyll (BChl), two of bacteriopheophytin (BPh), a molecule of menaquinone-9, an atom of nonheme iron, and four molecules of heme in the *c* type cytochrome. In 1984, when the structure was determined by Deisenhofer and Michel, this was the largest and most complex object whose atomic structure had been described. It was also one of the first known structures for a membrane protein. The accomplishment spurred an enormous rush of new photosynthesis research, only a tiny fraction of which can be mentioned here.

The reaction centers are embedded in the cytoplasmic membranes of the bacteria, with the bottom of the structure, as shown in Fig. 23-31, protruding into the cytoplasm and the heme protein at the top projecting out into the periplasm which lies within infoldings of the plasma membrane. Subunits L and M each contain five ~4.0 nm long roughly parallel helices, which span the cytoplasmic membrane. Another membrane-spanning helix is contributed by subunit H, which is located mainly on the cytoplasmic side. An approximate twofold axis of symmetry relates subunits L and M and the molecules of bound chlorophyll and pheophytin.

Spectral measurements suggesting exciton splitting were among early observations that led to the conclusion that the bacteriochlorophyll involved in the initial photochemical process exists as a dimer or **special pair** (Fig. 23-31),[319,324] a conclusion verified by the structure determination. The special pair of BChl *b* lies in the center of the helical bundle that is embedded in the membrane. Nearly perpendicular to the rings of the special pair are two more molecules of BChl *b*. The central magnesium atoms of all four bacteriochlorophylls are held by imidazole groups of histidine side chains.[319,325] Below the chlorophylls are

Figure 23-31 (A) Stereoscopic ribbon drawing of the photosynthetic reaction center proteins of *Rhodopseudomonas viridis*. Bound chromophores are drawn as wire models. The H subunit is at the bottom; the L and M subunits are in the center. The upper globule is the cytochrome *c*. The view is toward the flat side of the L, M module with the L subunit toward the observer. (B) Stereo view of only the bound chromophores. The four heme groups He1–He4, the bacteriochlorophylls (Bchl) and bacteriopheophytins (BPh), the quinones Q_A and Q_B, and iron (Fe) are shown. The four hemes of the cytochrome are not shown in (B). From Deisenhofer and Michel.[320] (C) Stereoscopic view of the Bchl *b* molecules along the local twofold axis. The special pair (D_M, D_L) is in the center with its tetrapyrrole rings almost perpendicular to the plane of the paper; the monomeric chlorophylls are labeled B_B and B_A. The four histidine ligands to the magnesium ions of the bacteriochlorophylls as well as two tyrosines (M195 and M208) and three water molecules (large circles) are also shown. From Deisenhofer *et al.*[321] with permission.



the two molecules of bacteriopheophytin and below them the nonheme iron and the menaquinone, the first quinone acceptor Q_A. It corresponds to Q_A of PSII shown in Fig. 23-18.

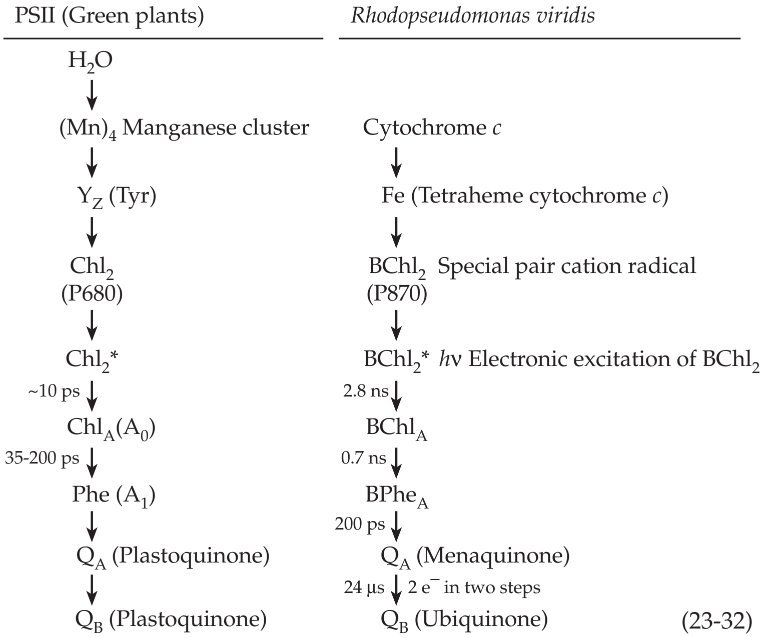
Isolated reaction centers usually contain or will bind a second quinone, which may be ubiquinone-10 (Q_{10})[325a] and which is usually designated Q_B. Its binding site is to the left of the nonheme iron in Fig. 23-31 in a position symmetrically related to that of Q_A. The reaction centers also contain a carotenoid 1,2-dihydro-neurosporene.[323] The reaction centers of purple bacterium, *Rhodospirillum rubrum,* each contain one molecule of spirilloxanthin; a variety of carotenoids are present in other species.[326] The reaction centers of a third purple bacterium, *Rhodobacter spheroides*, are closely similar in structure to that in Fig. 23-31 but lack the tetraheme.[327] Reaction centers of these bacteria accept electrons directly from a soluble cytochrome *c*.

Reaction center kinetics. After an 0.8-ps or shorter flash of light the decay of the singlet excited state of the bacteriochlorophyll dimer in isolated reaction centers can be followed by loss of its characteristic fluorescence.[328,329] The lifetime of this excited state in *R. sphaeroides* is only 4 ps indicating a rapid occurrence of the initial electron transfer of Eq. 23-31. A rise in absorbance at 1250 nm is interpreted as formation of the bacteriochlorophyll cation radical $BChl^+$ in the special pair. Other spectral changes support the formation of BPh^- as the first reduction product (A^- in Eq. 23-31). However, this is thought to occur in two steps[323] with the monomeric BChl (B_A in Fig. 23-31B) receiving the electron in ~2.8 ps and passing it to the pheophytin (Phe; ϕ_A in Fig. 23-31B) in ~0.7 ps (Eq. 23-32; corresponding steps for PSII of green plants are also shown).

The quinone Q_A (the secondary acceptor) is next reduced by the BPh^- radical in ~200 ps with development of a characteristic EPR signal[321,330] at $g = 1.82$. Over a much longer period of time (~320 ns) an electron passes from the tetraheme cytochrome subunit to the Chl^+ radical in the special pair.[323,323a] The relatively slow rate of this reaction may be related to the fact that the bacteriochlorophyll of the special pair is 2.1 nm (center-to-center) from the nearest heme in the cytochrome while BPh and Q_A are only 1.4 nm apart (see Fig. 23-31).[288,331] Over a period of ~24 µs after formation of the radical anion Q_A^- an electron from the Q_A^- radical is passed to Q_B, a weakly bound ubiquinone-9, to form the Q_B^- radical. Upon absorption of a second photon by the special BChl pair another electron is passed through the chain to form $Q_A^-Q_B^-$. Uptake of two protons with transfer of the second electron from Q_A^- to Q_B^- yields the ubiquinol QH_2, which dissociates from its binding site (Q_B) into the ubiquinone pool dissolved in the lipids of the membrane bilayer.[331a–g]

Why is this multistep sequence of electron transfers necessary? A variety of techniques such as femtosecond IR[332,333] and electronic[334–335a] spectroscopy, resonance Raman spectroscopy at low temperatures,[336] and study of many mutants[337–338c] have been directed toward an answer to this question. It has been generally accepted that light energy absorbed by any one of the "monomeric" Chl or pheophytins in the reaction centers is funneled "downhill" to the special pair within 0.1–0.2 ps to generate P*.[334,336] The ultrafast ~3 ps electron transfer from P* to the adjacent monomeric BChl or Chl is necessary to prevent loss of energy by fluorescence from P*. The subsequent energetically downhill transfer to a pheophytin and on to Q_A prevents reverse electron transfer, which could also lead to fluorescence. Both the efficiency and the quantum yield are very high.[339]

The rate of the ultrafast proton transfer becomes even higher at cryogenic temperature, suggesting quantum mechanical tunneling.[331,335,340–340b] The transfer is generally treated using Marcus theory (Chapter 16), which indicates a very small reorganization energy for the process. Another aspect of the process is a possible coupling of a vibrational mode of the protein matrix to the electron transfer. Femtosecond near-IR spectra show low-frequency vibrational modes of the excited-state reaction center chromophores which may facilitate electron transfer.[332,336] The transfer of an electron from the tetraheme cytochrome *c* of *R. viridis*[341] or from the small cytochrome c_2 of *Rhodobacter spheroides*[342] to Chl^+ of the special pair has similar characteristics but is slower than the initial electron transfer from P*. On the other hand, electron transfers from Q_A^- to Q_B^- involve two distinct steps and coupled uptake of two protons.[343] An unexplained fact is that photochemical electron transport through the reaction centers always occurs through the L-side to Q_A rather than the M-side.[343a–344] However, rapid electron transfer to the pheophytin on the B-side (M-side) has been observed following excitation with blue light. This may



(23-32)

TABLE 23-4
Properties of Various Reaction Centers

	Rhodopseudomonas viridis (now *Blasatochloris*)	Green plants and Cyanobacteria		Green sulfur bacteria, *heliobacteria*[a]
		PSII	PSI	
Subunits	L / M, H	D1 / D2, CP43, CP47	PsaA / PsaB nine others	$(PscA)_2$
Masses, kDa	24 / 28, 33	38 / 39.4	83 / 83	65 / 65
Input	Cyt *c* (4 Fe)	H_2O, $(Mn)_4$, Y_Z, Y_D	Cyt *f*, plastocyanin or cyt *c*	Cyt c_1, c_2
Special pair	P870	P680 (BChl $b)_2$	P700 (Chl $a)_2$	P840 (Chl $a)_2$
Monomeric chlorophyll	$BChl_L$, $BChl_M$		A_0 Both Chl *a*	A_0 Both Chl *a*-like
Pheophytin	$BPhe_L$, $BPhe_M$		A_0 (Phe, Phe)	
Quinone	Q_B, Q_A (Ubiquinone, menaquinone-9)	Q_B, Q_A (Plastoquinone)	Q_B, Q_A (A_1) (Both phylloquinone in cyanobacteria, plastoquinone in chloroplasts)	Q_B, Q_A (Both menaquinone-7)[b]
Iron	Fe^{3+}		F_X (Fe_4S_4)	F_X
Output	Ubiquinone Cyt bc_1, Cyt c_2	Ubiquinone Cyt b_6f	F_A, F_B, Fd $NADP^+$	F_A, F_B, Fd NAD^+

[a] Nitschke, W., and Rutherford, A. W. (1991) *Trends Biochem. Sci.* **16**, 241–245

[b] Kjear, B., Frigaard, N.-U., Yang, F., Zybailov, B., Miller, M., Golbeck, J. H., and Scheller, H. V. (1998) *Biochemistry* **37**, 3237–3242

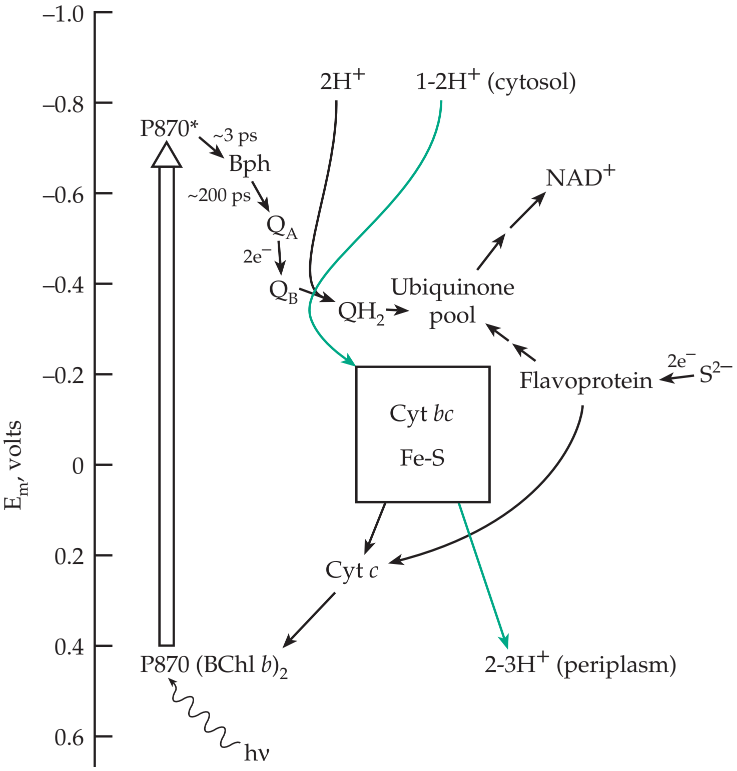


Figure 23-32 Simplified diagram of cyclic electron flow in purple bacteria. Two protons from the cytoplasm bind to Q_B^{2-} in the reaction center to form QH_2 (ubiquinol), which diffuses into the ubiquinone pool. From there it is dehydrogenated by the cytochrome bc_1 complex with expulsion of two protons into the periplasm. A third and possibly a fourth proton may be pumped (green arrows) across the membrane, e.g., via the Q cycle (Fig. 18-9). The protons are returned to the cytoplasm through ATP synthase with formation of ATP. Some electrons may flow to the reaction centers from such reduced substrates as S^{2-} and some electrons may be removed to generate NADPH using reverse electron transport.[345]

represent a photoprotective mechanism.[344a] In the PSI system, in which the two phylloquinones are tightly bound,[344b] both the A-side and the B-side seem to function in electron transfer.[344c]

Cyclic photophosphorylation in purple bacteria. QH_2 is eventually dehydrogenated in the cytochrome bc_1 complex, and the electrons can be returned to the reaction center by the small soluble cytochrome c_2, where it reduces the bound tetraheme cytochrome or reacts directly with the special pair in *Rhodobacter spheroides*. The overall reaction provides for a cyclic photophosphorylation (Fig. 23-32) that pumps 3–4 H^+ across the membrane into the periplasmic space utilizing the energy of the two photoexcited electrons. These protons can pass back into the cytoplasm via ATP synthases located in the same membrane with their catalytic centers in the cytosol (see Figs. 18-5 and 18-14).

Comparison with other reaction centers. Subunit L of the *R. viridis* reaction center was found to have a 25% sequence homology with a quinone-binding protein now known as D1, a component of the reaction center core of PSII of chloroplasts. This protein was identified as the specific target protein for inhibition by herbicides such as DCMU and atrazine (see Section 1 for structures). These compounds act as competitive inhibitors of quinone binding[346] and bind in the Q_B site in the *R. viridis* reaction center. This fact, together with the discovery that the core of PSII of green plants consists of a heterodimer of the related polypeptides D1 and D2, suggested that PSII is very similar to the bacterial center of Fig. 23-31. Both use a quinone as primary acceptor. However, P680 operates at a more positive potential (E_m ~1 V), consistent with the fact that it must provide an oxidizing agent able to oxidize H_2O to O_2 ($E°' = + 0.82$ V). Plastoquinone rather than menaquinone is the primary acceptor in PSII. A chlorophyll *a* dimer is apparently the initial electron donor. The methyl ester carbonyl groups on the edges of rings I (Fig. 23-20) of BChl may coordinate to other groups of the proteins.[288,346a] These ester groups are absent in Chl *a*.

The PSI reaction center (P700) of maize chloroplast was also found to contain a pair of homologous polypeptides with appropriately placed histidyl residues for chlorophyll binding.[347] FTIR spectra also indicate the presence of two Chl *a* molecules. Small shifts in IR frequencies upon electronic excitation suggest that one chlorophyll (P_1) is hydrogen-bonded through its 109-ester and 9-oxo groups (Fig. 23-20) while the chlorophyll (P_2) is free. This may account for the low value of E_m (Fig. 23-17). The charge on $P700^+$ appears to be carried in part on both chlorophylls. However, in the triplet state $^3P700^*$, which may be observed at low temperature, excitation is localized on P_1.[348] Although P700 operates at the low E_m of +0.49 V it produces a powerful reducing agent able to reduce ferredoxin. The first identified acceptors are other Fe–S centers present in integral membrane proteins.[345] The reaction centers of green bacteria[349] and PSI of cyanobacteria[245] have similar characteristics. As more genes have been sequenced and X-ray diffraction, electron crystallography, and electron microscopy with single-particle averaging have advanced, the fundamental similarity of all of the photosynthetic centers has been confirmed.[228,300,350–353] Gene sequences are often not highly conserved, but structures are more conserved. An evolutionary relationship of all of the reaction centers can be seen.[316,354]

PSI of cyanobacteria and green plants. The major reaction-center subunits **PsaA** and **PsaB** each have a C-terminal domain, resembling those of the L and M chains of purple bacteria, and a large N-terminal antenna-chlorophyll-binding domain. Cyanobacterial PSI contains ten other subunits, PsaC to PsaF, PsaI to PsaM, and PsaX. Thirty-one transmembrane helices have been assigned to the various subunits, several of which are in positions corresponding closely to those in the reaction centers of the purple bacteria (Fig. 23-33).[355–357] The PSI of higher plants is somewhat larger than that of cyanobacteria and contains somewhat different subunits.[356a,357,357a]

The electron donor to Chl^+ in PSI of chloroplasts is the copper protein plastocyanin (Fig. 2-16). However, in some algae either plastocyanin or a cytochrome *c* can serve, depending upon the availability of copper or iron.[345] Both Q_A and Q_B of PSI are phylloquinone in cyanobacteria but are plastoquinone-9 in chloroplasts. Mutant cyanobacteria, in which the pathway of phylloquinone synthesis is blocked, incorporate plastoquinone-9 into the A-site.[345a] Plastoquinone has the structure shown in Fig. 15-24 with nine isoprenoid units in the side chain. Spinach chloroplasts also contain at least six other plastoquinones. Plastoquinones C, which are hydroxylated in side-chain positions, are widely distributed. In plastoquinones B these hydroxyl groups are acylated. Many other modifications exist including variations in the number of isoprene units in the side chains.[358,359] There are about five molecules of plastoquinone for each reaction center, and plastoquinones may serve as a kind of electron buffer between the two photosynthetic systems.

Look at the Z scheme of Fig. 23-17. The lower end of each vertical arrow is located at an electrode midpoint potential E_m (or $E°'$) for the couple P^+/P, i.e., for a one-electron reduction of the Chl^+ or $BChl^+$ radical.[360] The top of the arrow is at the estimated value of E_m for the excited state P*. It is more negative than the ground-state value of E_m by the energy (in electron volts) of the light absorbed. This is a little misleading

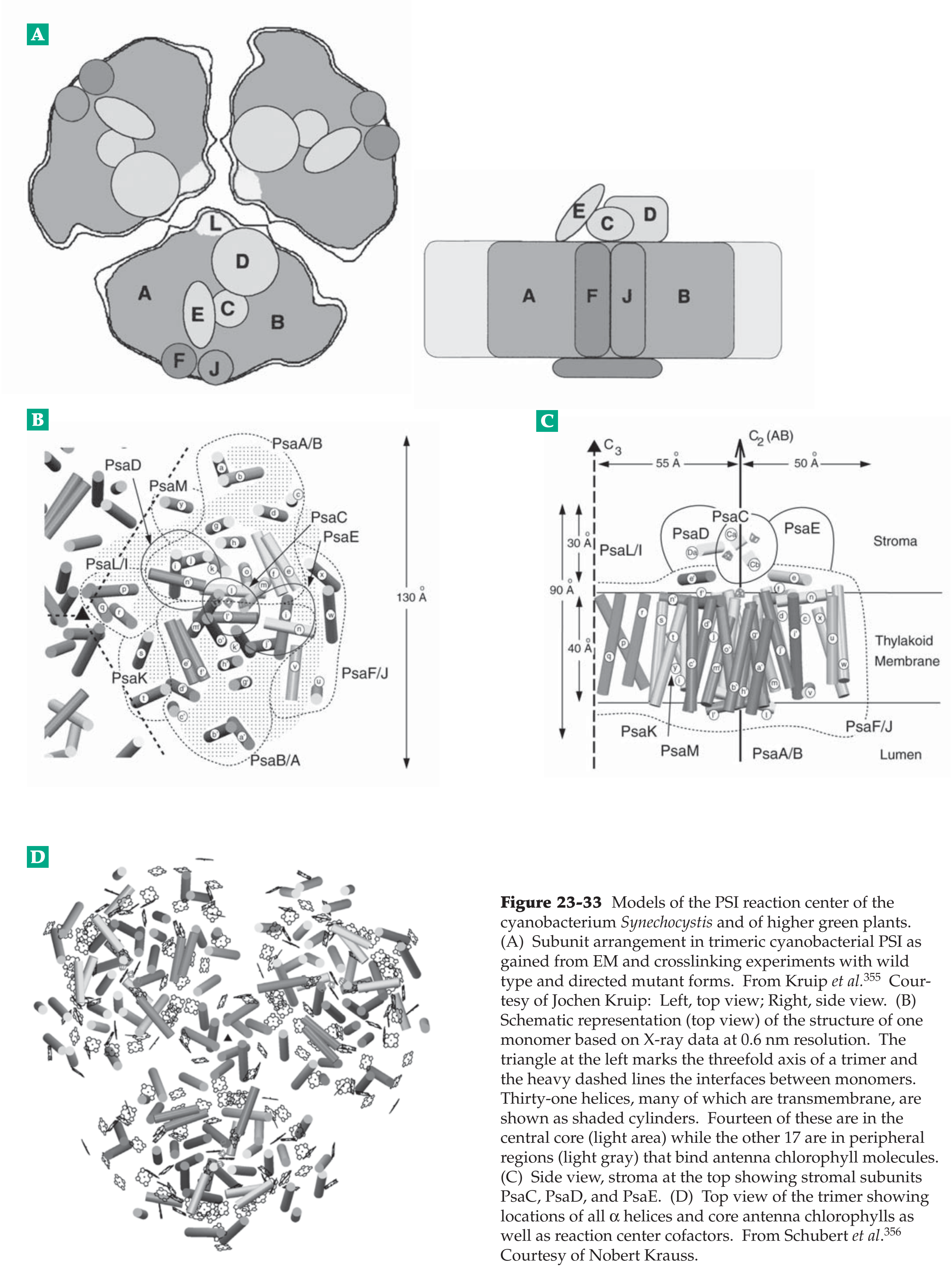

A
L
D
A
E
C
B
F
J
E
C
D
A
F
J
B
B
PsaA/B
PsaD
PsaM
PsaC
PsaE
PsaL/I
130 Å
PsaK
PsaF/J
PsaB/A
C
C_3
C_2 (AB)
55 Å
50 Å
PsaC
PsaD
PsaE
Stroma
30 Å
PsaL/I
90 Å
40 Å
Thylakoid
Membrane
PsaK
PsaM
PsaA/B
PsaF/J
Lumen
D

Figure 23-33 Models of the PSI reaction center of the cyanobacterium *Synechocystis* and of higher green plants. (A) Subunit arrangement in trimeric cyanobacterial PSI as gained from EM and crosslinking experiments with wild type and directed mutant forms. From Kruip *et al.*[355] Courtesy of Jochen Kruip: Left, top view; Right, side view. (B) Schematic representation (top view) of the structure of one monomer based on X-ray data at 0.6 nm resolution. The triangle at the left marks the threefold axis of a trimer and the heavy dashed lines the interfaces between monomers. Thirty-one helices, many of which are transmembrane, are shown as shaded cylinders. Fourteen of these are in the central core (light area) while the other 17 are in peripheral regions (light gray) that bind antenna chlorophyll molecules. (C) Side view, stroma at the top showing stromal subunits PsaC, PsaD, and PsaE. (D) Top view of the trimer showing locations of all α helices and core antenna chlorophylls as well as reaction center cofactors. From Schubert *et al.*[356] Courtesy of Nobert Krauss.

for it is not commonly appreciated that light carries entropy as well as energy. An important consequence of this fact is that not all of the energy of sunlight could be harnessed for chemical work. Knox[361] has calculated that at 700 nm at most 78% of the energy could be captured. See Parson[362] for further discussion. Nevertheless, the photoexcited P700* with E_m = – 1.26 V is able to reduce a series of membrane-bound Fe–S centers of $E°'$ ~ –0.5 to –0.6 V. There are three of these designated F_X, F_A, and F_B. Center F_X is an Fe_4S_4 cluster located at almost the same position as the single Fe^{3+} of the *R. viridis* reaction center (Fig. 23-31B). F_A and F_B are also Fe_4S_4 clusters both of which are carried on the small 79-residue PsaC.[363,364] This protein binds to the reaction center on the stromal side as shown in Fig. 23-33. Close to it are two other subunits, PsaD and PsaE, which appear to assist the docking of ferredoxin or flavodoxin to PsaC and cluster F_X.[365,366]

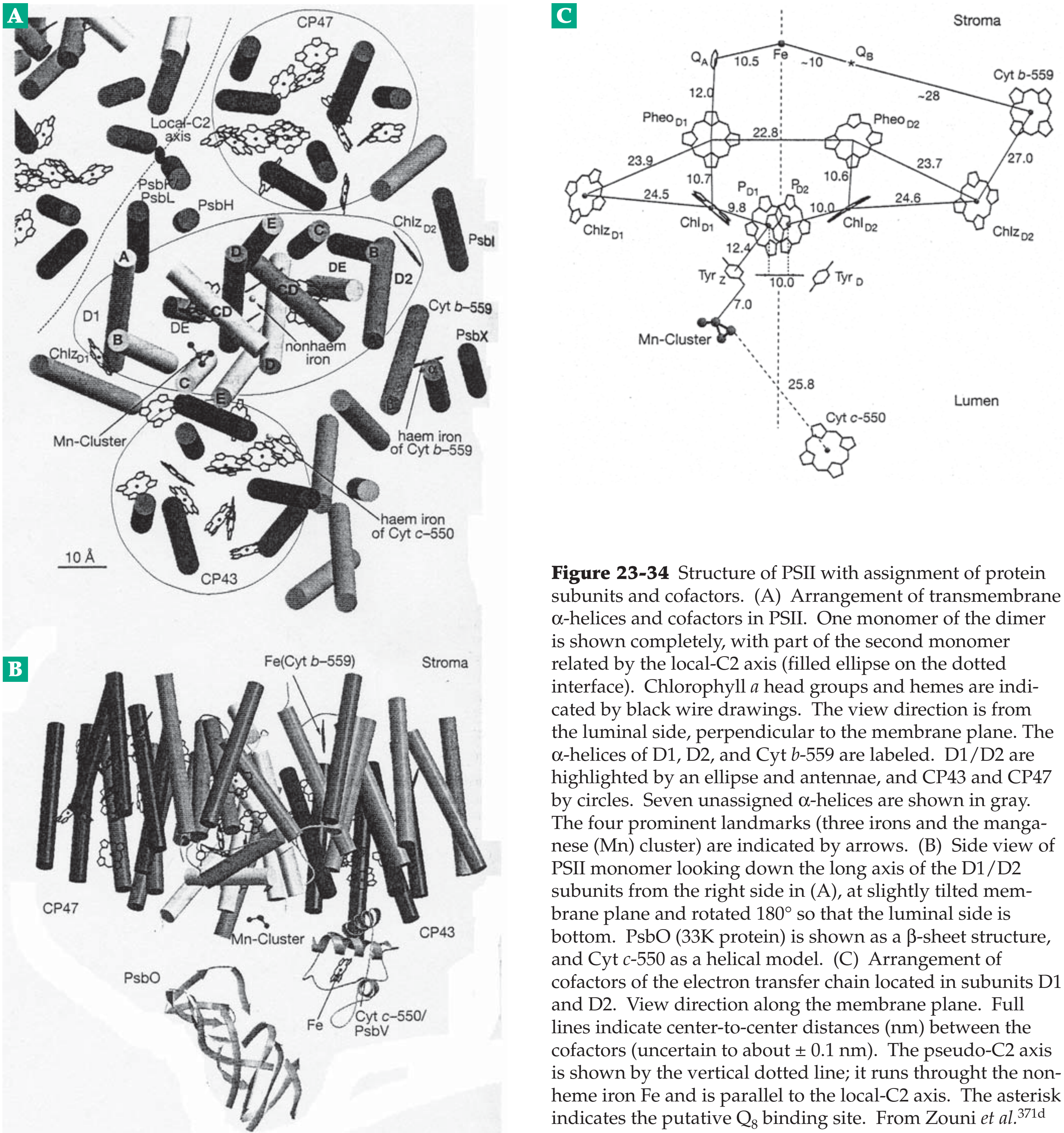

A
CP47
Local-C2
axis
PsbF/
PsbL
PsbH
A
E
C
D
B
DE
D2
CD
Chlz D2
PsbI
D1
B
DE
CD
nonhaem
iron
Cyt b-559
PsbX
Chlz D1
C
E
D
α
Mn-Cluster
haem iron
of Cyt b–559
haem iron
of Cyt c–550
10 Å
CP43
B
Fe(Cyt b–559)
Stroma
CP47
Mn-Cluster
CP43
PsbO
Fe
Cyt c–550/
PsbV
C
Stroma
Q A
10.5
Fe
~10
Q B
Cyt b-559
12.0
~28
Pheo D1
22.8
Pheo D2
23.9
23.7
27.0
10.7
P D1
P D2
10.6
24.5
9.8
10.0
24.6
Chlz D1
Chl D1
Chl D2
Chlz D2
12.4
Tyr Z
10.0
Tyr D
7.0
Mn-Cluster
25.8
Lumen
Cyt c-550

Figure 23-34 Structure of PSII with assignment of protein subunits and cofactors. (A) Arrangement of transmembrane α-helices and cofactors in PSII. One monomer of the dimer is shown completely, with part of the second monomer related by the local-C2 axis (filled ellipse on the dotted interface). Chlorophyll *a* head groups and hemes are indicated by black wire drawings. The view direction is from the luminal side, perpendicular to the membrane plane. The α-helices of D1, D2, and Cyt *b*-559 are labeled. D1/D2 are highlighted by an ellipse and antennae, and CP43 and CP47 by circles. Seven unassigned α-helices are shown in gray. The four prominent landmarks (three irons and the manganese (Mn) cluster) are indicated by arrows. (B) Side view of PSII monomer looking down the long axis of the D1/D2 subunits from the right side in (A), at slightly tilted membrane plane and rotated 180° so that the luminal side is bottom. PsbO (33K protein) is shown as a β-sheet structure, and Cyt *c*-550 as a helical model. (C) Arrangement of cofactors of the electron transfer chain located in subunits D1 and D2. View direction along the membrane plane. Full lines indicate center-to-center distances (nm) between the cofactors (uncertain to about ± 0.1 nm). The pseudo-C2 axis is shown by the vertical dotted line; it runs throught the nonheme iron Fe and is parallel to the local-C2 axis. The asterisk indicates the putative Q_8 binding site. From Zouni *et al.*[371d] Courtesy of Athina Zouni.

The photosynthetic centers of green photosynthetic sulfur bacteria also have centers F_X, F_A, and F_B.

The soluble electron carriers released from the reaction centers into the cytoplasm of bacteria or into the stroma of chloroplasts are reduced single-electron carriers. Bacterial ferredoxin with two Fe_4S_4 clusters is formed by bacteria if enough iron is present. In its absence flavodoxin (Chapter 15), which may carry either one or two electrons, is used. In chloroplasts the carrier is the soluble **chloroplast ferredoxin** (Fig. 16-16,C), which contains one Fe_2S_2 center. Reduced ferredoxin transfers electrons to $NADP^+$ (Eq. 15-28) via **ferredoxin:NADP⁺ oxidoreductase**, a flavoprotein of known three-dimensional structure.[367–369]

PSII and formation of oxygen. The structure of PSII has been difficult to determine directly, but its core has been modeled in atomic detail using bacterial reaction centers as a guide.[370,371] More recently electron crystallography provided a three-dimensional image at 0.8 nm resolution.[314,371a] The resolution has been extended down to 4 nm by X-ray crystallography.[371b−e] The structure of the cyanobacterial PSII (Fig. 23-34) is very similar to that of green plants.[371e,f] PSII contains at least 17 protein subunits, all of which are encoded by chloroplast genes. The large structurally similar D1 and D2 form the core. They are encoded by genes *PsbA* and *PsbD* and have molecular masses of 38.0 and 39.4 kDa, respectively.[315] Both Q_A and Q_B are plastoquinone. It is of historical interest that these cofactors were first designated simply as Q, not for quinone but for *quencher*. This is because Q_A apparently quenched the fluorescence of P680, the reaction-center chlorophyll *a*. If chloroplasts are irradiated with 650-nm light, PSII is activated but PSI is not. Under these conditions Q_A becomes reduced, and the fluorescence of Chl^+ increases, presumably because the electron acceptor Q_A is absent. If PSI is activated by addition of far-red light, Q_A remains more oxidized, and fluorescence is quenched by a mechanism that appears to depend upon one of the additional bound chlorophylls, as well as the chlorophyll pair Chl_2 and the 9-kDa cytochrome b_{559}. The latter is an essential PSII subunit (*PsbE* gene),[372,373] which forms a tight complex with the D1/D2 pair.

Other subunits in PII include the 56- and 50-kDa antenna proteins CP43 (*PsbC* gene) and CP47 (*PsbB* gene). Three **extrinsic proteins**, which bind to the lumenal side of the thylakoid membrane, are the 33-kDa **manganese stabilizing protein** (*PsbO* gene),[374,375] cytochrome c_{550}, and a 12-kDa subunit (Fig. 23-34). The position of the larger two of these proteins are shown in the model in Fig. 23-34, which is based on 0.8-nm resolution data.[315] These extrinsic proteins seem to function together to facilitate binding not only of Mn ions but also of Ca^{2+} and Cl^-, both of which are essential for O_2 evolution.[376] Other smaller subunits are also present.[376a] Together with its antenna complexes PSII may form large supercomplexes with as many as 25 subunits.[376b] The distribution of complexes varies in the different regions of the thylakoid, e.g., the stroma lamellae and grana stacks (Fig. 23-19).[376c]

The four-electron dehydrogenation of two water

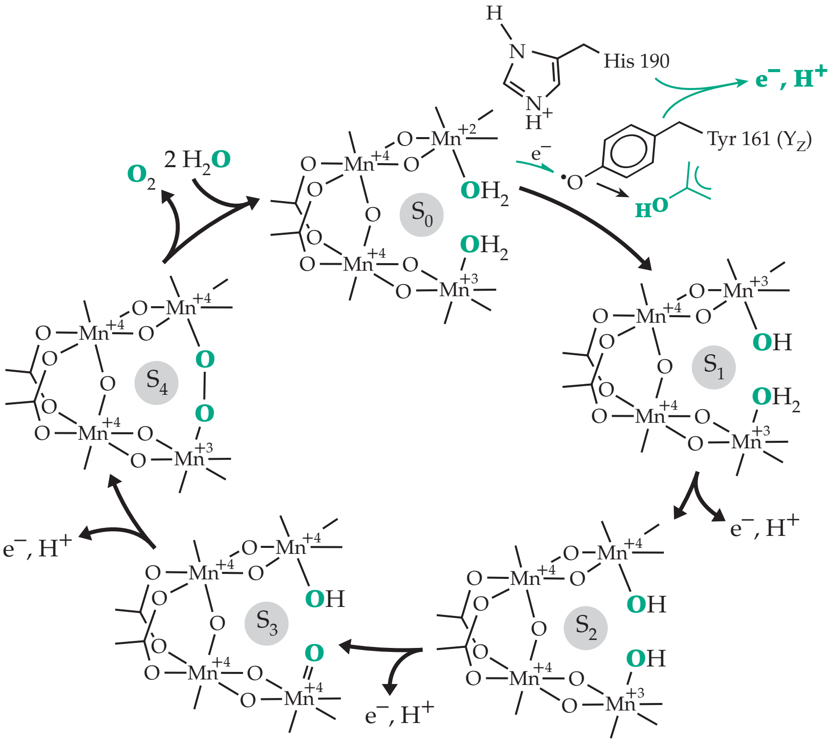


Figure 23-35 Proposed sequence of S-states of the manganese cluster of photosystem II. The successive states as two molecules of H_2O (green oxygen atoms) are converted to O_2 is shown with the successive states S_0–S_4 labeled. To save space and possible confusion tyrosine 161 (Y_Z) and the nearby His 190 are shown only by S_1. The Y_Z radical is thought to remove a hydrogen atom or H^+ from one bound H_2O and an electron from one Mn ion at each of the four S-states S_0–S_3 functioning in each case to eject a proton into the thylakoid lumen and to transfer an electron to P^+ of the reaction center. However, the exact sequence of e^- transfer and H^+ release may not be shown correctly. After Hoganson and Babcock.[392,392a]

molecules to give one molecule of O_2 by PSII is still not well understood. From experiments on oxygen evolution in the presence of repeated short flashes of light it was found that a four-quantum process is required.[376–380] There must be some way of storing oxidizing equivalents until enough are present to snap together an oxygen molecule. There is abundant evidence that manganese is required for this process and that the oxidation of H_2O occurs on a cluster of four atoms of manganese.[379]

The 33-kDa protein PsbO, which is present in all oxygen-forming photosynthetic organisms, is closely associated with the Mn_4 cluster. Removal of this protein leads to a gradual loss of two of the four Mn ions.[375] The structure of the Mn_4 cluster is not yet certain, but on the basis of EXAFS spectroscopy a pair of di-μ-oxo bridged Mn dimers linked by carboxylates and with a fifth μ-oxo bridge as in Fig. 23-35 has been proposed. The distance between Mn ions in the di-μ-oxo-Mn_2 groups is ~0.27 nm, and the planes of these groups are roughly parallel to the surface of the thylakoid membranes.[381,382] The protein groups that bind the Mn atoms include carboxylates, as shown in Fig. 23-35, and probably one or more histidine imidazole groups,[383] perhaps of His 332 and His 337 of the D1 chain.

The immediate donor of an electron to the reaction-center cation P^+ ($ChlZ^+$) of PSII was identified by EPR spectroscopy as a tyrosine radical.[384] On the basis of directed mutations this tyrosine, which is usually designated $\mathbf{Y_Z}$, is Tyr 161 of the D1 chain and is located ~1.2 nm from one of the chlorophylls of the $(Chl)_2$ pair.[371d,376,380,385–387] The two molecules in the pair are not in close contact, their central Mg atoms being ~1.0 nm apart. One of the two probably forms the $P680^{\bullet+}$ intermediate.[371d] Y_Z is also close to the Mn_4 cluster and to the imidazole group of His 190 of subunit D1 as is shown in Fig. 23-34C.[386,388] The $Y_Z^{\bullet}$ radical is able to accept an electron from the Mn_4 cluster within 30–1300 μs depending upon the oxidation state of the cluster (see Fig. 23-35).[388] If a proton is transferred synchronously from a bound H_2O, a neutral –OH group will be created on Y_Z. The proton may then be transferred to His 190, which can eject the proton on its other nitrogen atom into the lumen. Alternatively $Y_Z^{\bullet}$ may accept an electron to become tyrosinate $–O^-$, which then donates an electron to P^+, while His 190 accepts a proton directly from a bound H_2O.

In a mechanism proposed by Hoganson and Babcock (Fig. 23-35) four successive transfers, each of one H^+ + one e^-, leads to a three-electron oxidation of Mn ions, e.g., from the 2^+ and 3^+ oxidation states to all Mn^{4+}, and to joining of the two water oxygens to form a manganese peroxide linkage. Oxidation of the peroxide dianion to O_2 by the adjoining Mn^{4+} and Mn^{3+} ions completes the cycle. This mechanism is hypothetical, and various alternatives have been presented.[387–389b] Most assume a structure similar to that shown in Fig. 23-35. Some are based on nonenzymatic model reactions.[390,391] Chloride ions are essential to O_2 formation[393,394] especially in going from state S_2 to S_3 and S_3 to S_0. This suggests that Cl^- may function in passing electrons between Mn ions.[393,395] Calcium ions are also necessary, but it has been difficult to establish an exact function.[396–397a] A bicarbonate ion may also be an essential ligand in the Mn_4 cluster.[398,398a]

ATP synthesis in chloroplasts. The flow of electrons between PSII and PSI (Fig. 23-18) is of great importance for ATP formation. As previously mentioned, plastocyanin is usually the immediate donor to P700 and serves as a mobile carrier to bring electrons to this reaction center. In this function it is analogous to cytochrome *c* of mitochondrial membranes. The essentiality of plastocyanin was shown by study of copper-deficient *Scenedesmus* (Fig. 1-11). The photoreduction of CO_2 by H_2 is impaired in these cells, but the Hill reaction occurs at a normal rate.

Like mitochondria, chloroplasts (when illuminated) pump protons across their membranes (Fig. 23-18). However, while mitochondria pump protons to the outside, the protons accumulate on the inside of the thylakoids. The ATP synthase heads of coupling factor CF_1 are found on the outside of the thylakoids, facing the stromal matrix, while those of F_1 lie on the insides of mitochondrial membranes. However, the same mechanism of ATP formation is used in both chloroplasts and mitochondria (Chapter 18).

The cleavage of water at PSII also occurs on the inside of the thylakoids. The splitting of one water molecule leaves two protons (one per electron) inside the thylakoids, while the electrons are "photoejected" through the lipid bilayer to acceptor Q_A on the outside. The chlorophyll in PSI is likewise in contact with the inside of the bilayer with acceptor A_0 (Fig. 23-18) on the outside. Since the conversion of NAD^+ to NADH on the outside generates a proton, the overall reaction would be the pumping of one and a half protons per electron passing through the Z scheme.

The pathways involved in cyclic photophosphorylation in chloroplasts are not yet established. Electrons probably flow from the Fe–S centers Fd_x, Fd_a, or Fd_b back to cytochrome b_{563} or to the PQ pool as is indicated by the dashed line in Fig. 23-18. Cyclic flow around PSII is also possible. The photophosphorylation of inorganic phosphate to pyrophosphate (PP_i) occurs in the **chromatophores** (vesicles derived from fragments of infolded photosynthetic membranes) from *Rhodospirillum rubrum*. The PP_i formed in this way may be used in a variety of energy-requiring reactions in these bacteria.[399] An example is formation of NADH by reverse electron transport.

Protection of chloroplasts against radiation and oxygen. Carotenoids often act as accessory light-receiving pigments, but an additional function is protecting photosynthetic organisms against toxic effects of light.[400–403] Carotenoid photoprotection has been demonstrated in photosynthetic bacteria and in the reaction centers and light-harvesting complexes (LHC) of green plants. Excited chlorophyll molecules can pass from the singlet (^{1}Chl*) to the triplet (^{3}Chl*) state by intersystem crossing (Fig. 23-14). The triplet chlorophyll can then react with ordinary oxygen ($^3O_2$) to form singlet oxygen ($^1O_2$). The formation of ^{3}Chl* is favored when the intensity of sunlight is high and energy is absorbed in the LHCs faster than it can be utilized in the reaction centers. Carotenoids are able to quench excitation of both ^{3}Chl* and $^1O_2$. Strains of *Rhodobacter sphaeroides* that lack carotenoids rapidly form both ^{3}BChl* and BChl$^+$ cation radicals in their LH1 and LH2 complexes (Fig. 23-26). However, the presence of carotenoids quenches the triplet bacteriochlorophyll effectively and no formation of BChl$^+$ radicals is seen.[403] Another effect of excessive light energy is the reaction of O_2 with the highly reducing P700* of photoexcited PSI. This interaction can form triplet oxygen, which can react to generate superoxide anion radicals, H_2O_2, and hydroxyl radicals.[224,227,402]

In higher plants the quenching of both ^{3}Chl* and $^1O_2$ depends upon carotenoids[404] and also upon the large transthylakoid membrane potential that is generated by high light intensities.[405,406] High light intensity also induces a rapid reductive deoxygenation of epoxycarotenoids via the **xanthophyll** or **violaxanthin cycle** (Eq. 22-10).[401,407,407a,b] Epoxycarotenoids are found only in photosynthetic O_2 evolving organisms. Although occurring in response to light, the cyclic photodeoxygenation and reoxygenation is not a nonphotochemical process. Violaxanthin contains the epoxy structure at both ends of the molecule. Reduction of one end produces **antheroxanthin** and of both ends **zeaxanthin**. These three carotenoids are found in almost all higher green plants and algae. The deepoxidation is mediated by ascorbic acid, occurs in the lumen of the thylakoids, and is favored by the low pH developed during illumination. Epoxidation (Eq. 22-10) is catalyzed by a monooxygenase located on the stromal side. The significance of the xanthophyll cycle is puzzling. There may be specific binding sites, perhaps in the inner antenna complex on CP29 (Fig. 23-34), that bind zeaxanthin or antheroxanthin. This could alter the antenna structure to form an “exciton trap” in which the fluorescence lifetime would be decreased and excitation energy would be dissipated rapidly as heat.[408] Zeaxanthin is also found in the macular area of the primate retina.[401]

Recent studies, using an *Arabidopsis* mutant defective in the xanthophyll cycle, point to a chlorophyll-binding protein PsbS, which participates in nonphotochemical quenching at high light intensity.[401a,b] Another *Arabidopsis* protein, which is probably a blue light receptor, participates in an avoidance response by which chloroplasts move to the side wall to avoid strong light.[401c]

Yet another carotenoid function in PSII has been proposed. Under some conditions, when electron flow from Tyr$_z$ is blocked, the bound monomeric chlorophyll Z (Chlz) acts as a secondary electron donor to P680$^+$. The cytochrome b_{559} subunits may have a similar function. Both cyt b_{559} and the carotenoid are essential for assembly of PSII, and both may participate in a protective cycle.[409,410] Chloroplasts generate both O_2 and powerful reducing materials such as the membrane-bound FeS centers of PSI, which may form superoxide ions by single-electron donation to O_2. Probably for this reason, chloroplasts are rich in superoxide dismutase which converts superoxide to O_2 and H_2O_2. The latter can diffuse into peroxisomes and react with catalase and peroxidases. It can also be reduced to H_2O within the chloroplasts by ascorbic acid and ascorbate peroxidase.[224,227] The resulting dehydroascorbate (Box 18-D) can be reduced back to ascorbate by glutathione (Box 11-B) and dehydroascorbate reductase, in the following electron transfer sequence:

$$\text{NADPH} \rightarrow \text{G–S–S–G} \rightarrow \text{dehydroascorbate} \rightarrow H_2O_2 \tag{23-33}$$

Under extreme conditions of excess light energy **photoinhibition** is observed as a result of damage to the PSII structure.[411–415] The D1 polypeptide is cleaved, probably as a result of oxidation by $^1O_2$ and proteolysis. Damaged proteins are replaced and the PSII structure rebuilt, but the effect is a long-lasting decrease in photosynthetic efficiency. The cyanobacterium *Synechococcus* has three *PsbA* genes and resists UV-B radiation by exchanging a delicate D1 polypeptide with more resistant ones as necessary.[415] Other adaptations to varying light-intensity involve movement of light-harvesting complexes from the thylakoid stacks, which contain much PSII, to the stroma lamellae (Fig. 23-19), which contain more PSI. Some herbicides act by binding into the Q_B site in PSII. They may cause light-induced oxidative stress that kills the plant.[415a]

5. Control of Photosynthesis

The key reaction of the Calvin–Benson cycle of CO_2 reduction is the carboxylation of ribulose bisphosphate to form two molecules of 3-phosphoglycerate (Eq. 13-48). The properties of ribulose bisphosphate carboxylase (**rubisco**, Figs. 13-10 to 13-12), which catalyzes this reaction, are discussed in Chapter 13. It is controlled in part by CO_2 and by natural inhibitors,[416] but regulation of rubisco starts at the transcriptional level.

Light-induced transcription. Plants depend upon light both as a source of energy and also for control of development. Many genes are activated by light in response to at least three groups of photoreceptors. These are **phytochromes** (Section H) and the blue light responsive **cryptochromes** (Section I) and the ultraviolet light **UV-B photoreceptors**.[417] The synthesis of chlorophyll, of reaction center proteins, and of many enzymes are controlled by light-induced transcription.[418–420] Among these processes are synthesis of both the large and small subunits of rubisco (Fig. 13-10). The small subunits are synthesized in the cytoplasm in a precursor form. After illumination the concentration of the rubisco mRNA may be increased 100-fold.[421,422] On the other hand, the large subunit of the carboxylase is encoded in chloroplast DNA, and stimulation of its synthesis by light appears to be at the translational level.[423]

Light-induced control via the ferredoxin/thioredoxin system. Rubisco is activated by CO_2 (Chapter 13) and by fructose 6-*P* and is inhibited by fructose 1,6-P_2 (Fig. 23-36),[424] whose accumulation is a signal to turn off the carboxylase. Conversely, fructose 6-*P* in high concentrations turns on the Calvin–Benson cycle. Like the reactions of gluconeogenesis (Chapter 17), photosynthetic CO_2 incorporation is dependent on the highly regulated fructose-1,6-bisphosphatase. In chloroplasts it is activated by light through the mediation of reduced ferredoxin and thioredoxin.[424–427] The small mobile thioredoxin (Box 15-C) is reduced to its dithiol form by reduced ferredoxin[428–429a] and then reduces one or more disulfide linkages in the fructose 1,6-bisphosphatase to activate that enzyme (Fig. 23-36). Other light-activated enzymes of the Calvin–Benson cycle include sedoheptulose-1,7-bisphosphatase, the phosphoribulokinase that forms ribulose 1,5-bisphosphate and the $NADP^+$-dependent glyceraldehyde-3-phosphate dehydrogenase. $NADP^+$-dependent malate dehydrogenase, which has a major function in C_4 plants (see Fig. 23-38), is totally inactive in the dark but is activated by the ferredoxin–thioredoxin system in the light.[427] The activity of the LHCII complex is also affected.[429b]

Another aspect of chloroplast metabolism is synthesis of starch. Formation of ADP-glucose from glucose 1-phosphate is induced by 3-phosphoglycerate, a "feed-ahead" type of regulation (Fig. 23-36). Although fructose 2,6-bisphosphate is absent from chloroplasts, it has an important regulatory function in the cytoplasm of plants as it does in animals.[425,430] In the plant cytosol triose phosphates from the chloroplasts are converted to fructose 6-*P*, glucose 6-*P*, UDP-glucose, and sucrose. Inorganic phosphate P_i, which accumulates in plant vacuoles, also has a regulatory function.[431] It activates the kinase that converts fructose 6-*P* to fructose 2,6-P_2 and inhibits the phosphatase that converts the bisphosphate back to fructose 6-*P*. The accumulated fructose 2,6-P_2 inhibits fructose-1,6-bisphosphatase and slows the conversion of triose phosphates to sucrose (Eq. 23-34). Accumulation of fructose 6-*P* due to decreased utilization for sucrose formation will have a similar effect. However, both 3-phosphoglycerate and dihydroxyacetone phosphate have opposite effects and will act to remove the inhibition by lowering the fructose 2,6-P_2 level and to promote rapid sucrose formation (Eq. 23-34).[430]

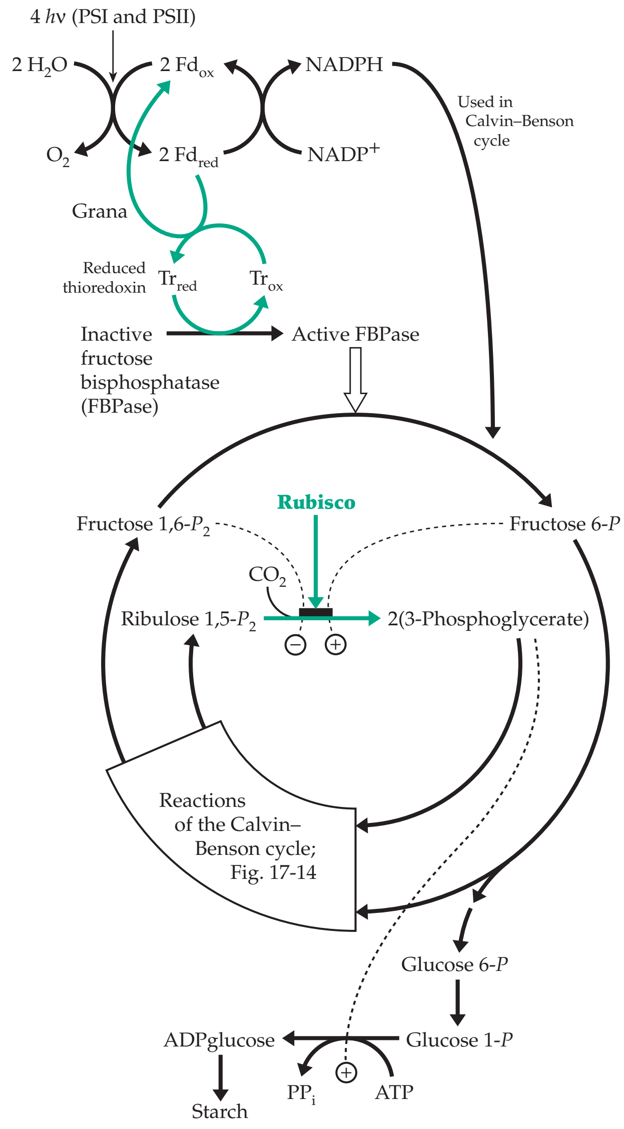
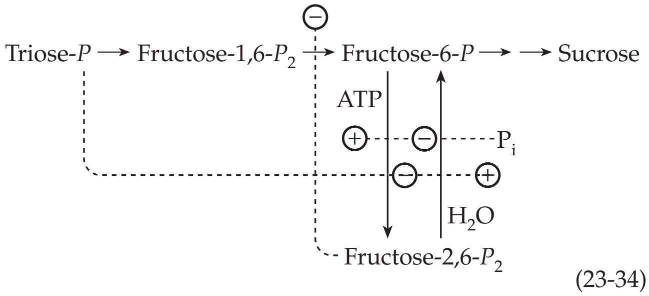


Figure 23-36 Some control mechanisms for photosynthetic assimilation of carbon dioxide. After Buchanan and Schurmann[424] with modifications.



(23-34)

6. Photorespiration; C_3 and C_4 Plants

The first product of incorporation of CO_2 via the Calvin–Benson cycle is 3-phosphoglyceric acid (Box 17-E). It was the rapid appearance of radioactivity from $^{14}CO_2$ in phosphoglycerate and other three-carbon (C_3) compounds that permitted Calvin and associates, using green algae, to work out the complex cycle as it is shown in Fig. 17-14. Green algae, spinach, and many common crop plants are often known as **C_3 plants**. Another group of plants, mostly of tropical origin and capable of extremely fast growth, behave differently.[224,227,432] In these plants, which include sugar cane, maize, and crabgrass, radioactivity from ^{14}C-containing CO_2 is found first in the C_4 compounds oxaloacetate, malate, and aspartate. These **C_4 plants** are characterized by high efficiencies in photosynthesis, a fact that explains the rapid growth of crabgrass and the high yield of corn. Maximum rates of CO_2 incorporation may attain 40–60 mg of CO_2 per square decimeter of leaf surface per hour (~0.3 mmol CO_2 $m^{-1}s^{-1}$ or ~0.10 mol CO_2 per mol of total chlorophyll per second), more than twice that for common crop plants.

Like all other organisms plants respire in the dark, but illumination of C_3 plants markedly increases the rate of oxygen utilization. This light-enhanced respiration (**photorespiration**) may attain 50% of the net rate of photosynthesis. Photorespiration prevents plants from achieving a maximum yield in photosynthesis. For this reason, its understanding and control assume great importance in agriculture. It is difficult to measure the rate of photorespiration, and the literature on the subject often refers instead to the **CO_2 compensation point**. This is the CO_2 concentration (at a given constant light intensity) at which photosynthetic assimilation and respiration balance. (Similarly, the **light compensation point** is the light intensity at which the rate of photosynthetic CO_2 incorporation and that of respiration exactly balance.) Normal air has a CO_2 content of ~0.03% or 300 ppm. For common C_3 crop plants the CO_2 compensation point is ~40–60 ppm at 25°C. The C_4 plants are characterized by a much lower CO_2 compensation point, often less than 10 ppm. In strong sunlight the CO_2 level of air in a field of growing plants drops. Furthermore, as the temperature rises on a hot day, the CO_2 compensation point rises. The result is a serious decrease in efficiency of photosynthesis for the C_3 plants but not for the C_4 plants.

Metabolism of glycolic acid. The 2-carbon glycolic acid is formed in large quantities in the chloroplasts of C_3 plants and moves out into the cytosol.[433,434] The major source of this acid is phosphoglycolate whose formation is catalyzed by rubisco in the chloroplasts through competition of O_2 for the CO_2 binding site of the enzyme (Eq. 13-50). It is easy to understand why an increase in the O_2 pressure in air increases the CO_2 compensation point for a plant. Another less important source of glycolate is transketolase, which may yield glycolaldehyde as a side product (Eq. 17-15). Glycolaldehyde can be oxidized readily to glycolate. Glycolate is metabolized rapidly, some in the chloroplasts,[435] but most in the peroxisomes. There the flavoprotein glycolate oxidase converts it to glyoxylate with formation of H_2O_2 (Fig. 23-37).[436] Some of the hydrogen peroxide formed may react nonenzymatically, decarboxylating glyoxylate to formate and CO_2, but most is probably destroyed by peroxidases or catalase. The latter enzyme is lacking in chloroplasts, one reason why oxidation of glycolate must occur in the peroxisomes.

Glyoxylate undergoes transamination to glycine, which can be oxidatively decarboxylated (Fig. 15-20) in the mitochondria. It can also be converted to serine,[437] some of which returns to the peroxisomes to be oxidized to hydroxypyruvic acid and glyceric acid (Fig. 23-37). The latter can be synthesized into glucose. The net result is the stimulation of a large amount of metabolism that ultimately produces CO_2 and apparently accounts for the light-induced respirations of plants. Because much of the glycine formed in the peroxisomes is oxidatively decarboxylated in the mitochondria, photorespiration also generates large amounts of NH_3. This is recycled into amino acids within the photosynthetic cells (green lines in Fig. 23-37), an energy-requiring process.[438]

Although much metabolism occurs as a result of photorespiration, it appears to waste energy rather than to provide energy to the plant. Why then don't plants avoid this process? Wouldn't a small change in the structure of ribulose bisphosphate carboxylase allow plants to avoid photorespiration and to grow more efficiently? The answer is not clear. It has been difficult to create such modified plants, and there is a possibility that they would not grow well. One theory is that photorespiration protects plants when the CO_2 pressure is low and the absorbed light would damage the chloroplasts if there were not a way to utilize the accumulating reduced Fe–S proteins generated by PSI. Photorespiration provides a mechanism.[434] Most

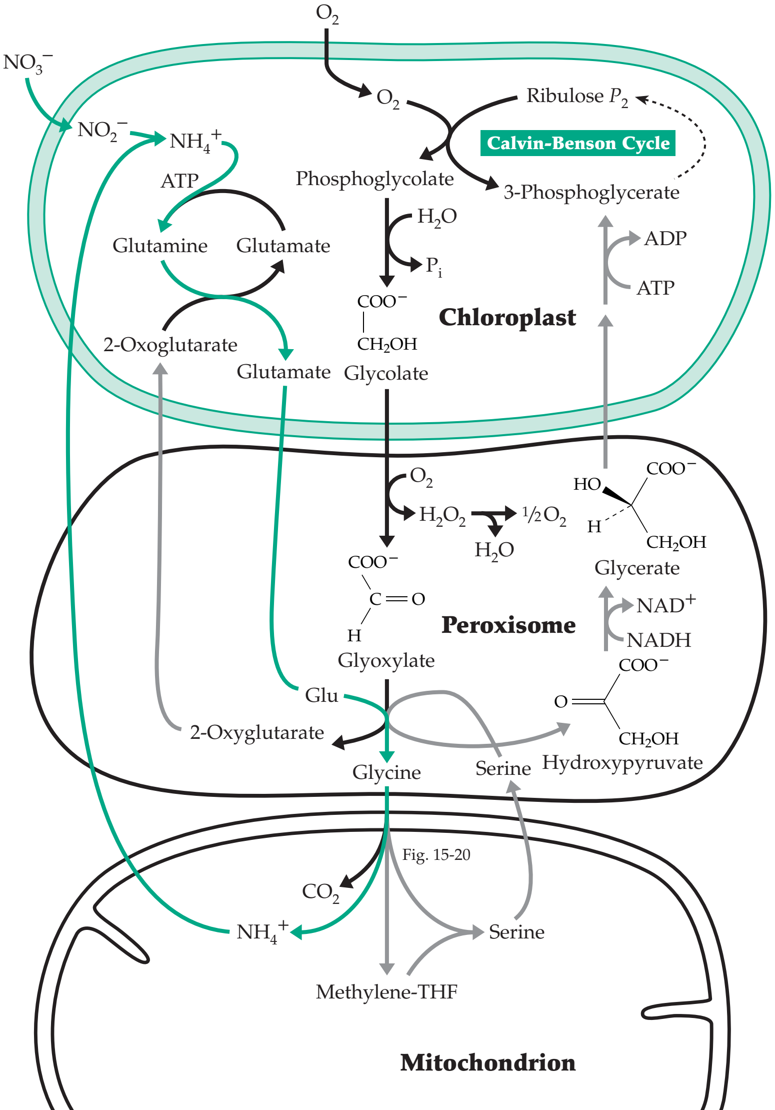
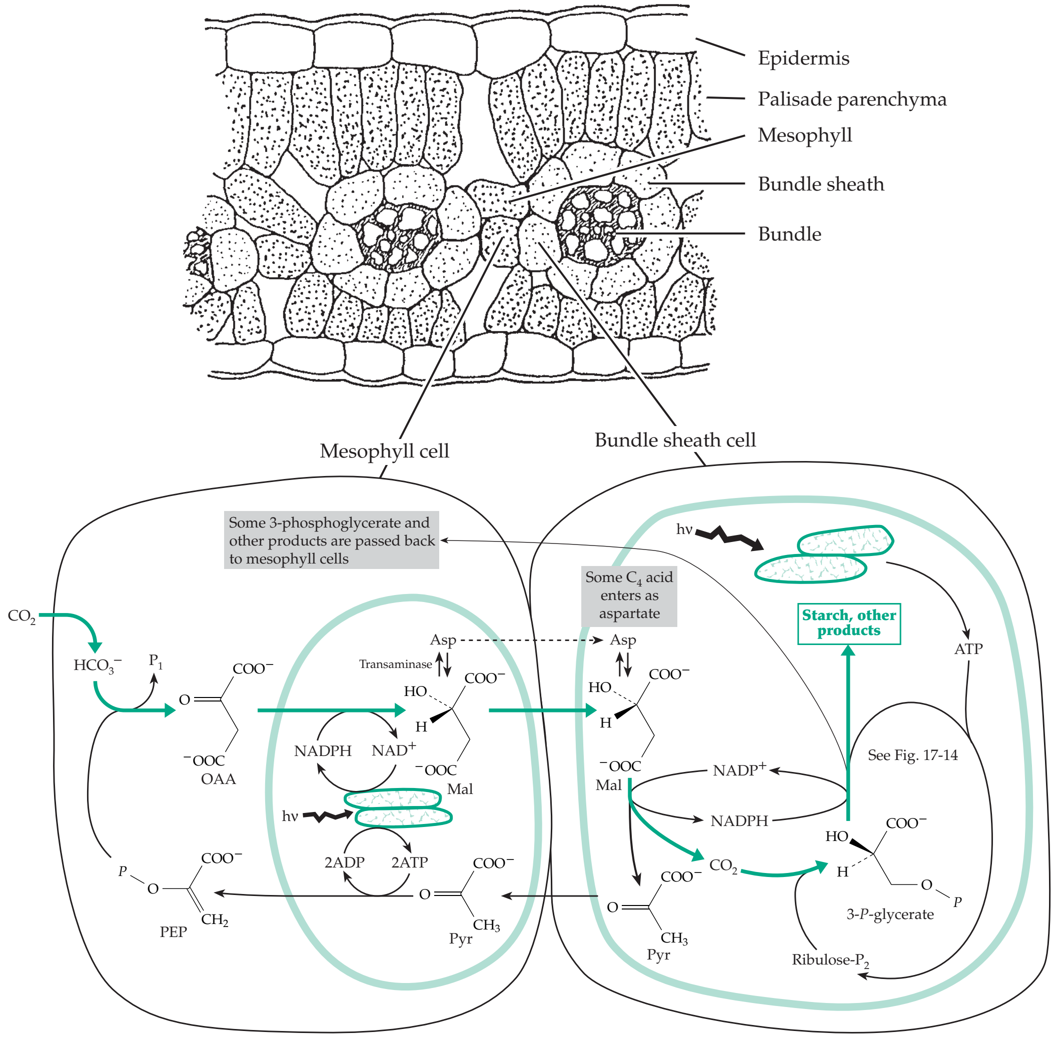

O_2
NO_3^-
O_2
Ribulose P_2
NO_2^-
NH_4^+
Calvin-Benson Cycle
ATP
Phosphoglycolate
3-Phosphoglycerate
H_2O
P_i
ADP
ATP
Glutamine
Glutamate
COO^-
CH_2OH
Chloroplast
2-Oxoglutarate
Glutamate
Glycolate
O_2
H_2O_2
½O_2
H_2O
HO
COO^-
H
CH_2OH
Glycerate
COO^-
C=O
H
Glyoxylate
Peroxisome
NAD^+
NADH
COO^-
O
CH_2OH
Glu
2-Oxyglutarate
Glycine
Serine
Hydroxypyruvate
Fig. 15-20
CO_2
NH_4^+
Serine
Methylene-THF
Mitochondrion

Figure 23-37 Production of glycolate by chloroplasts and some pathways of its metabolism in peroxisomes and in mitochondria. After Tolbert[436] and Givan *et al.*[438]

efforts to breed plants with lower photorespiration rates or to inhibit it chemically have failed.[433,439]

The C_4 cycle for concentration of carbon dioxide. The C_4 plants reduce their rate of photorespiration by using a CO_2 concentrating mechanism that enables them to avoid the competition from O_2. All species of C_4 plants have a characteristic internal leaf anatomy in which a single dense layer of dark green cells surrounds the vascular bundles in the leaves. This **bundle sheath** is surrounded by a loosely packed layer of cells, the **mesophyll**, an arrangement that is sometimes called the "Kranz anatomy." In C_4 plants there is a separation of the chemical reactions between the mesophyll and the bundle sheath cells. The incorporation of CO_2, as bicarbonate ion into oxaloacetate, occurs in the mesophyll cells, principally through the action of PEP carboxylase (Fig. 23-38). Oxaloacetate is reduced to malate by light-generated NADPH. Alternatively, it undergoes transamination to aspartate. Both malate and aspartate then diffuse out of the mesophyll cells and into bundle sheath cells where the malate undergoes oxidative decarboxylation via the malic enzyme (Eq. 14-42) to pyruvate (Fig. 23-38). Aspartate also can be converted to oxaloacetate, malate, and pyruvate in the same cells. The overall effect is to transport CO_2 from the mesophyll cells into the bundle sheath cells along with two reducing equivalents, which appear as NADPH following the action of the malic enzyme. The CO_2, the NADPH, and additional NADPH generated in the chloroplasts of the bundle sheath cells are then used in the Calvin–Benson cycle reactions to synthesize 3-phosphoglycerate and other materials. Of the CO_2 used in the bundle sheath cells, it is estimated that 85% comes via the C_4 cycle and only 15% enters by direct diffusion. The advantage to the cell is a higher CO_2 tension, less competition with O_2, and a marked reduction in photorespiration.

The pyruvate produced in the bundle sheath cell is largely returned to the mesophyll cells where it is acted upon by **pyruvate-phosphate dikinase**.[441] This unusual enzyme (Eq. 17-55) phosphorylates pyruvate to PEP while splitting ATP to AMP and PP_i. The latter is in turn degraded to P_i. In effect, two high-energy linkages are cleaved for each molecule of pyruvate phosphorylated. Because of this extra energy need, it is thought that cyclic photophosphorylation is probably more important in the chloroplasts of the mesophyll cells than in the bundle sheath cells. It also accounts for the fact that C_4 plants are less efficient than C_3 plants under cool or shaded conditions.[442] Other CO_2-concentrating mechanisms exist in plants.[442a] For example, cyanobacteria accumulate HCO_3^- ions in carboxysomes, polyhedral bodies to which rubisco adheres. An ABC type ATP-dependent transporter powers the bicarbonate accumulation.[442b]


Epidermis
Palisade parenchyma
Mesophyll
Bundle sheath
Bundle
Mesophyll cell
Bundle sheath cell
Some 3-phosphoglycerate and other products are passed back to mesophyll cells
Some C4 acid enters as aspartate
hv
Starch, other products
ATP
CO2
HCO3−
Pi
COO−
O
−OOC
OAA
Asp
Transaminase
HO
H
NADPH
NAD+
Mal
hv
2ADP
2ATP
P
PEP
CH2
Pyr
CH3
NADP+
NADPH
See Fig. 17-14
3-P-glycerate
Ribulose-P2

(23-35)

Figure 23-38 The C_4 cycle for concentrating CO_2 in the C_4 plants. From Haag and Renger with alterations.[440]

Metabolism in the family Crassulaceae. The crassulacean plants are a large group that includes many ornamental succulents such as *Sedum*. They have a remarkable metabolism by which large amounts of malic and isocitric acids are synthesized at night. During the day when photosynthesis occurs these acids disappear. The stomata in the leaves (Chapter 1) stay closed during the day and open only at night, an adaptation that permits the plants to live with little water. However, this requires that the plant accumulates carbon dioxide by night and incorporates it photosynthetically into organic compounds by day.[443] A possible mechanism is shown in Fig. 23-39. On the left side of the figure are reactions by which starch can be broken down at night to PEP. While it would also be possible to produce that compound by the glycolysis reactions, labeling studies have indicated that the pentose phosphate pathway is more important.[330] The PEP acts as the CO_2 acceptor to create oxaloacetate, which is then reduced to malic acid. A balanced fermentation reaction (Fig. 23-39; Eq. 23-35) can be written by using the NADPH formed in the conversion of glucose 6-*P* to ribulose 5-*P*. During the day when ATP and NADPH are available in abundance from photoreactions, the conversions on the right side of the figure can take place. The initial step, the release of CO_2 from malic acid by the malic enzyme, is the same as that employed by C_4 plants. In this case, it is used to release the CO_2 stored by night, making it available for incorporation via the Calvin–Benson cycle. The remaining pyruvate is reconverted to starch.

$$C_6H_{10}O_5\text{ (starch)} + 2\ CO_2\text{ (g)} \rightarrow 2\ C_4H_4O_5^{2-}\text{ (malate)} + 4\ H^+$$
$$\Delta G'\text{ (pH 7)} = -159\text{ kJ mol}^{-1} \qquad (23\text{-}35)$$

Many plants store substantial amounts of malate in their cytoplasm and in vacuoles. It apparently

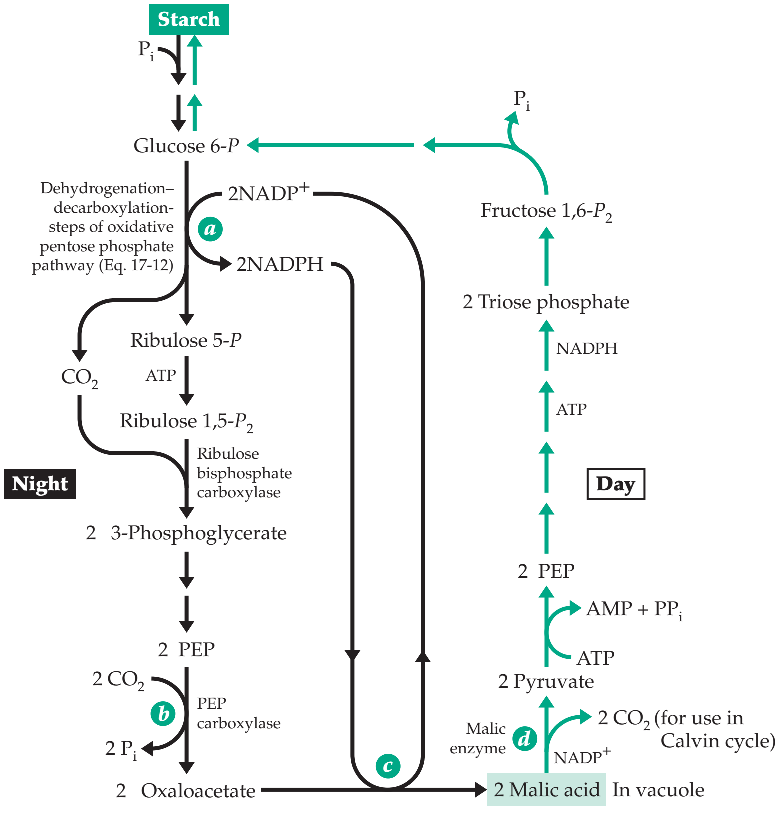

Starch
P_i
Glucose 6-P
P_i
Dehydrogenation–decarboxylation-steps of oxidative pentose phosphate pathway (Eq. 17-12)
2NADP$^+$
a
2NADPH
Fructose 1,6-P_2
2 Triose phosphate
Ribulose 5-P
CO_2
ATP
NADPH
ATP
Ribulose 1,5-P_2
Ribulose bisphosphate carboxylase
Night
Day
2 3-Phosphoglycerate
2 PEP
AMP + PP_i
ATP
2 PEP
2 CO_2
b
PEP carboxylase
2 P_i
2 Pyruvate
Malic enzyme
d
2 CO_2 (for use in Calvin cycle)
NADP$^+$
c
2 Oxaloacetate
2 Malic acid
In vacuole

Figure 23-39 A proposed night–day cycle of crassulacean acid metabolism.

serves as a ready reserve for carbohydrate synthesis.

7. Photosynthetic Formation of Hydrogen

A system consisting of chloroplasts, ferredoxin, and hydrogenase has been used to generate H_2 photosynthetically.[444] This may be a prototype of a method of solar energy generation for human use. Another photochemical hydrogen-generating system makes use of both the nitrogen-fixing heterocysts and photosynthetic vegative cells of the cyanobacterium *Anabaena cylindrica*.[445] In this instance hydrogen production is accomplished by nitrogenase (Eq. 24-6). Photogeneration of H_2 by bacteria is just one of many kinds of photometabolism observed among photosynthetic microorganisms.

F. Vision

The light receptors of the eye perform a very different function from those of chloroplasts. Visual receptors initiate nerve impulses, and their primary requirement is a high sensitivity. By the use of stacked membranes containing a high concentration of an intensely absorbing molecule[446,447] the most sensitive visual receptors are able to trap nearly every photon that strikes them. The retina of the human eye contains more than 10^8 tightly packed receptor cells of two types. The **rod cells** are extremely sensitive. Used for night-time vision, they give a "black and white picture" and are concentrated around the periphery of the retina. The retina works as a coincidence detector. An ensemble of ~500 rods must register 5–7 isomerizations within a few tenths of a second in order to trigger a nerve impulse.[448] The less sensitive **cone cells**, which are most abundant in the center of the retina, are of three types with different spectral sensitivities. They provide color vision.

The retinal receptors have a very active metabolism. Human rod cells (Fig. 23-40) may live and function for a hundred years.[190] A self-renewal process leads to a casting off of the older membranous discs from the end of the rod[194] and replacement by new discs at the end nearest to the nucleus. The rod outer segment is surrounded by a plasma membrane. Within the membrane but apparently not attached to it are ~500 stacked discs of ~2 μm diameter and with a repeat distance between centers of ~32 nm. Each disc is enclosed by a pair of membranes each ~7 nm thick with a very narrow space between them. From electron micrographs it appears that this space within the discs is sealed off at the edges. Somewhat larger spaces separate adjacent discs.

The membranes of the rod discs are ~60% protein and 40% lipid (Table 8-3). About 80% of the protein is **rhodopsin** (visual purple), a lipoprotein that is insoluble in water but soluble in detergent solutions. Digitonin is widely used to disperse rhodopsin molecules because it causes no change in optical properties. In addition to rhodopsin, in the outer segment discs of frog retinal rods, there are ~65 molecules of phospholipid and smaller amounts of other materials for each molecule of rhodopsin (Table 8-3). The cone cells have a similar architecture but have a different shape and contain different light receptors. The receptors in the cones are present in deep indentations of the plasma membrane rather than in discs within the cytoplasm.

1. Visual Pigments

The rod pigment rhodopsin is readily available from cattle retinas and has been studied for many

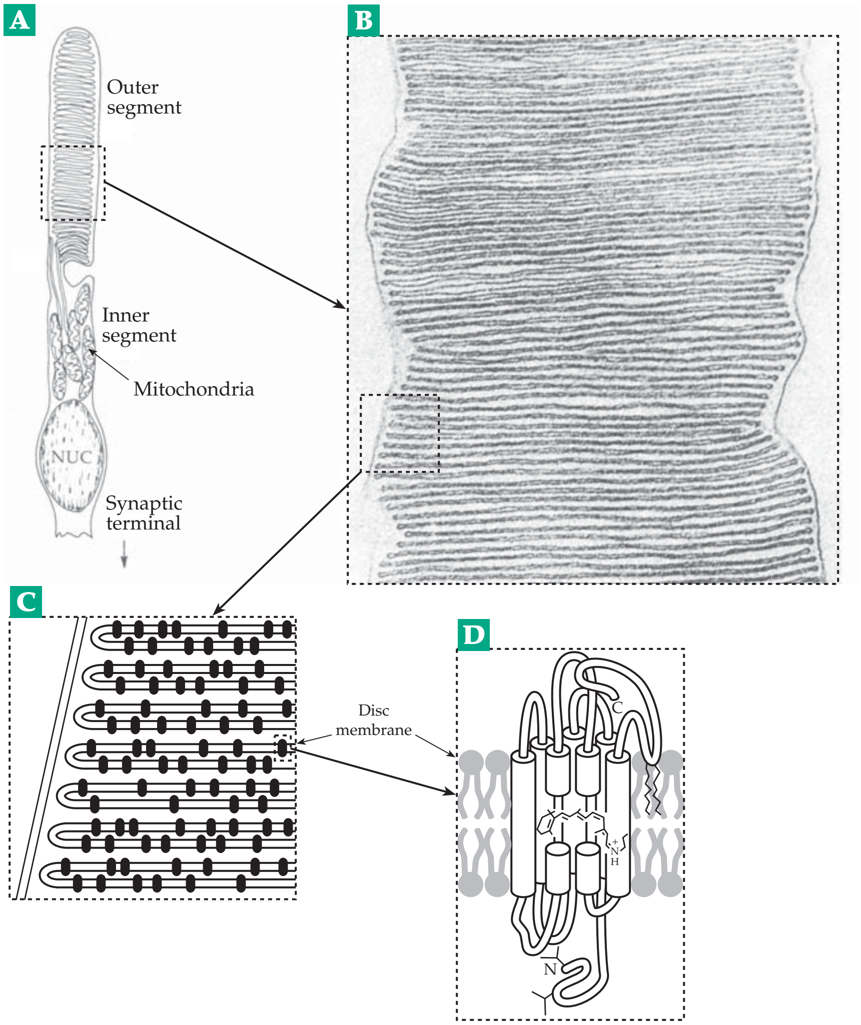
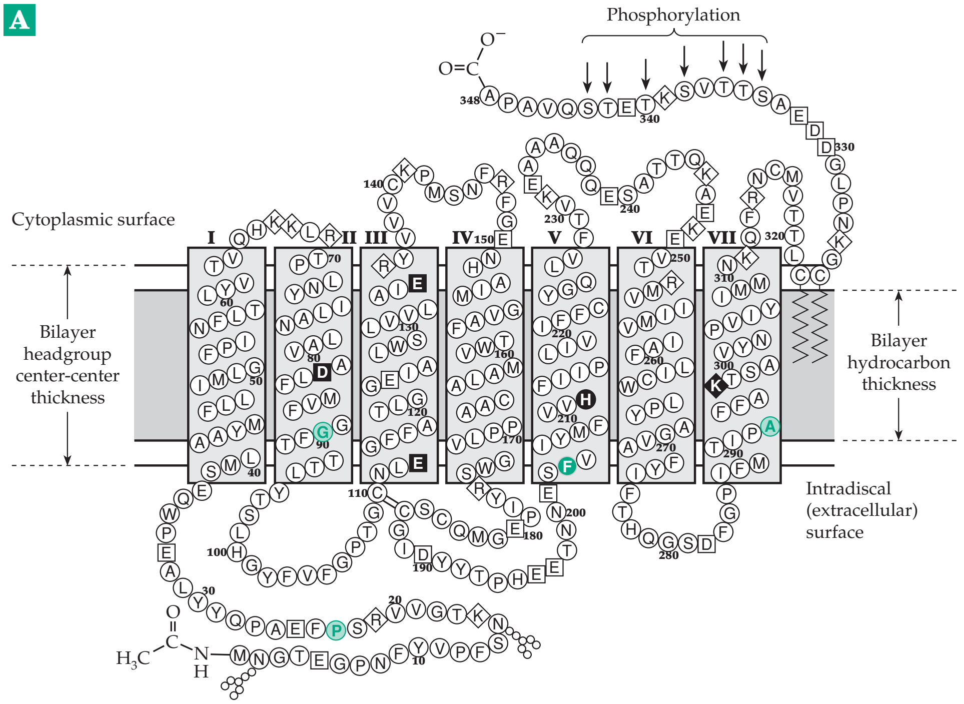


Figure 23-40 (A) Diagram of a vertebrate rod cell. From Abrahamson and Fager.[446] (B) Electron micrograph of a longitudinal section of the outer segment of a retinal rod of a rat. There are 600–2000 discs per rod and 2 x 10^4 and 8 x 10^5 rhodopsin molecules per disc. Courtesy of John E. Dowling. (C) Enlarged section; from Dratz and Hargrave.[451] (D) Schematic drawing of rhodopsin. The two α helices in the front have been partly cut away to reveal the 11-*cis* retinal in protonated Schiff base linkage to lysine 296. From Nathans.[448] Courtesy of Jeremy Nathans.

years. It has a molecular mass of ~41 kDa of which ~2 kDa is contributed by two asparagine linked oligosaccharides.[195] Both bovine and human rhodopsins consist of a 348-residue protein known as **opsin** to which is bound a molecule of vitamin A aldehyde, **retinal**. Human and bovine opsins are 93% homologous. A totally synthetic 1057 base pair gene for bovine rhodopsin was made by Khorana and associates.[449,450] The gene was constructed with 28 unique sites for cleavage by specific restriction endonucleases. These have allowed easy specific mutation of the gene and production of a wide variety of mutant forms of opsin. Similar synthetic genes have been constructed for the three human cone pigments[452] and for the related bacterial protein **bacteriorhodopsin**.[453]

Transmembrane structure.

From its circular dichroism rhodopsin was estimated to be 60% helical, and its amino acid sequence suggested that it contains seven parallel membrane-spanning helices (Fig. 23-41)[451] just as does bacteriorhodopsin (Section G; Fig. 23-45). Rhodopsin and other visual pigments are also members of the large family of **G-protein coupled receptors**, which includes the β_2 adrenergic receptor pictured in Fig. 11-6. It has been hard to determine the three-dimensional structure of rhodopsin or other receptors of this family. However, their relationship to bacteriorhodopsin, whose structure was obtained in 1975 by electron crystallography[454] and more recently by X-ray crystallography at 0.15 nm resolution,[455] permitted modeling based on similarities among the proteins.[456–458] New results of electron crystallography[459,460] and mass spectrometry[461] have been combined with studies of mutant forms of rhodopsin and other visual pigments to provide the picture given in Fig. 23-41A,B. Recently a higher resolution structure (Fig. 23-41D) has been obtained by X-ray diffraction.[461a–c] Mutations have been introduced in every part of the rhodopsin molecule, and the effects on photoreception, protein stability, and other properties have been observed.[448,450,462–465] Some of the essential residues identified are indicated in Fig. 23-41A, a schematic diagram showing the seven helices and connecting loops. Figure 23-41B shows an end view of the helix bundle with its retinal prosthetic group buried in the interior of the protein. Rhodopsin is roughly cylindrical with a length of 6–6.5 nm and a diameter of ~2.8 nm and is embedded in the phospholipid bilayer with its long axis perpendicular to the membrane surface (Fig. 23-40). The two oligosaccharide chains, which are attached near the N terminus, project into the intradiscal space on the side away from the cytoplasm. Palmitoyl groups on two cysteine side chains help to anchor the protein. Rhodopsin apparently exists in the mem-



● A few residues essential to the mechanism of vision.

○ Point mutations causing serious diseases. P23H is a frequent cause of retinitis pigmentosa. Mutations G90D and A292E cause congenital night blindness.

● The mutation C203R in the green-sensitive pigment causes red-green color blindness.

Figure 23-41 (A) Model of the topological organization of bovine rhodopsin according to the consensus analysis of Baldwin.[456] Oligosaccharide chains are portrayed on the glycosylation sites Asn 2 and Asn 15. Glutamate 113 provides the counter ion for the *N*-protonated retinal Schiff base. Cysteines 110 and 187 form an essential disulfide linkage. Histidine 211 modulates the interconversion of the metarhodopsin forms I and II. Lysine 296 forms a Schiff base with 11-*cis*-retinal. Cysteines 322 and 323 are sites of palmitoylation, and as many as six serine and threonine hydroxyls (indicated by green arrows) may become phosphorylated during desensitization of rhodopsin. Aspartate 83, glutamate 134, and histidine 211 may be essential for proton movements. The point mutation C203R in the green-sensitive pigment causes red-green color blindness. The mutation

---

branes as monomers. From the composition it can be calculated[451] that the average center-to-center distance of the cylindrical molecules must be ~5.6 nm.

The visual chromophores. Rhodopsin has been an object of scientific interest for over 100 years.[462] Wald and associates[469,470] established that rhodopsin contains 11-*cis*-retinal bound to the opsin in Schiff base linkage (Eq. 23-36). When native rhodopsin is treated with sodium borohydride, little reduction is observed. However, after the protein is bleached by light, reduction of the Schiff base linkage becomes rapid, and the retinal is incorporated into a secondary amine, which was identified as arising from Lys 296.

In a crystal structure[471] 11-*cis*-retinal has the 12-*s*-*cis* conformation shown at the top in Eq. 23-36 rather than the 12-*s*-*trans* conformation at the center and in which there is severe steric hindrance between the 10–H and 13–CH_3. Nevertheless, ^{1}H– and ^{13}C–NMR spectroscopy suggest that the retinal in rhodopsin is in a twisted 12-*s*-*trans* conformation.[472,472a] The Schiff base of 11-*cis*-retinal with *N*-butylamine has an absorption maximum at ~360 nm but *N*-protonation, as in the structure in Eq. 23-36, shifts the maximum to 440 nm with ε_{max} = 40,600 M^{-1} cm^{-1} (Fig. 23-42). This large shift in the wavelength of the absorption maximum (the *opsin shift*) indicates that binding to opsin stabilizes the light-excited state by almost 33 kJ / mol compared to that in the free *N*-protonated Schiff base. This is evidently the result of a fixed negative charge, that of Glu 113, which is near the polyene chain of the retinal (Fig. 23-41B) and probably separated from it by a hydrogen-bonded water molecule.[473–477]

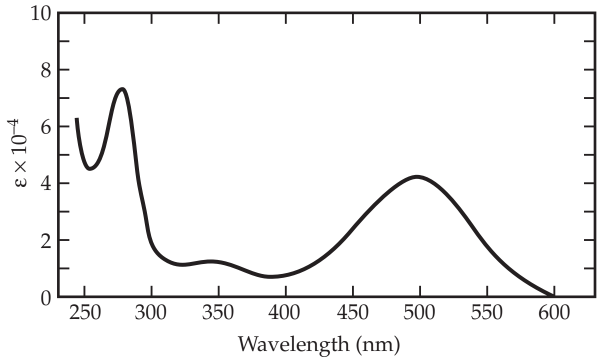
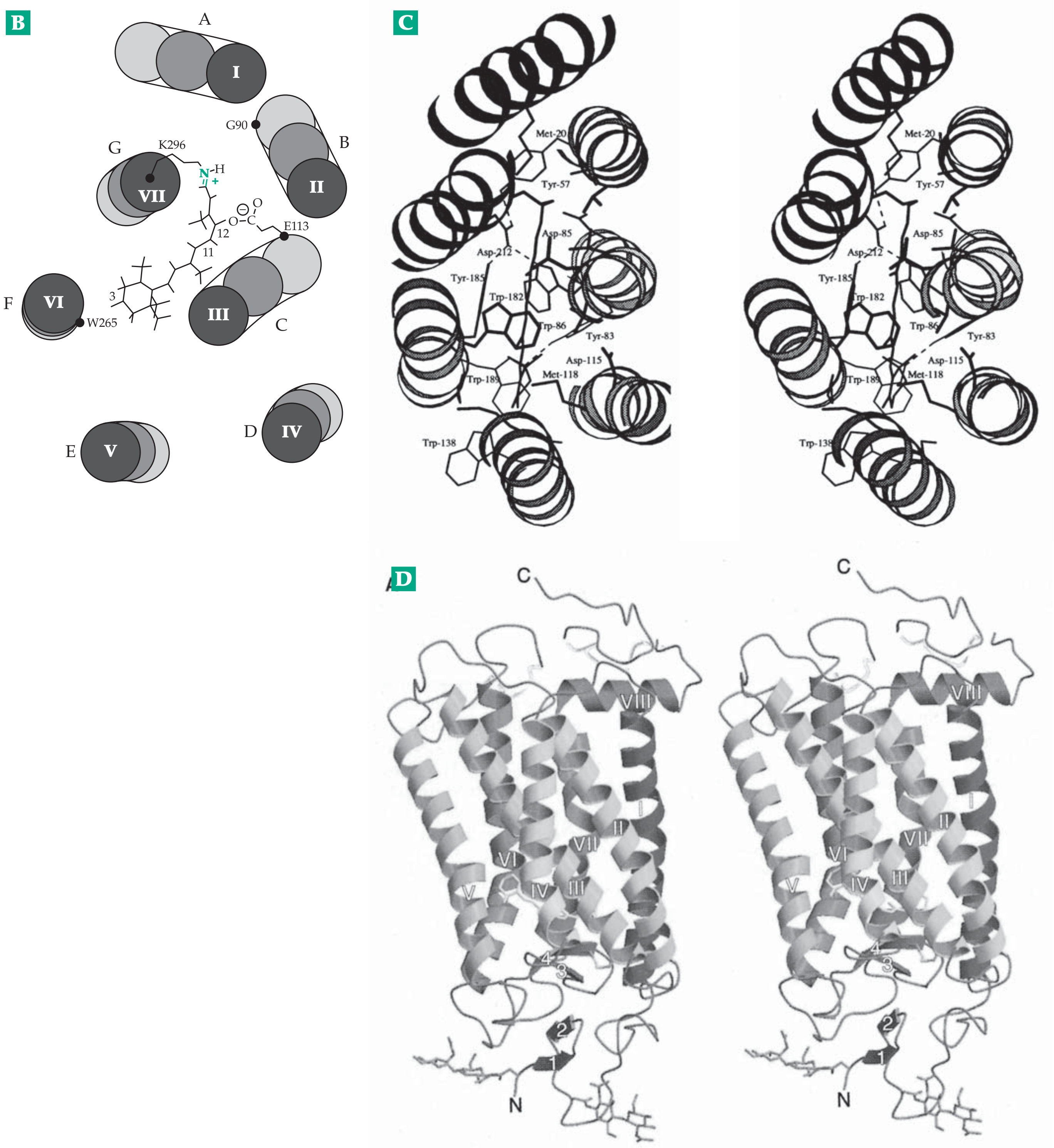
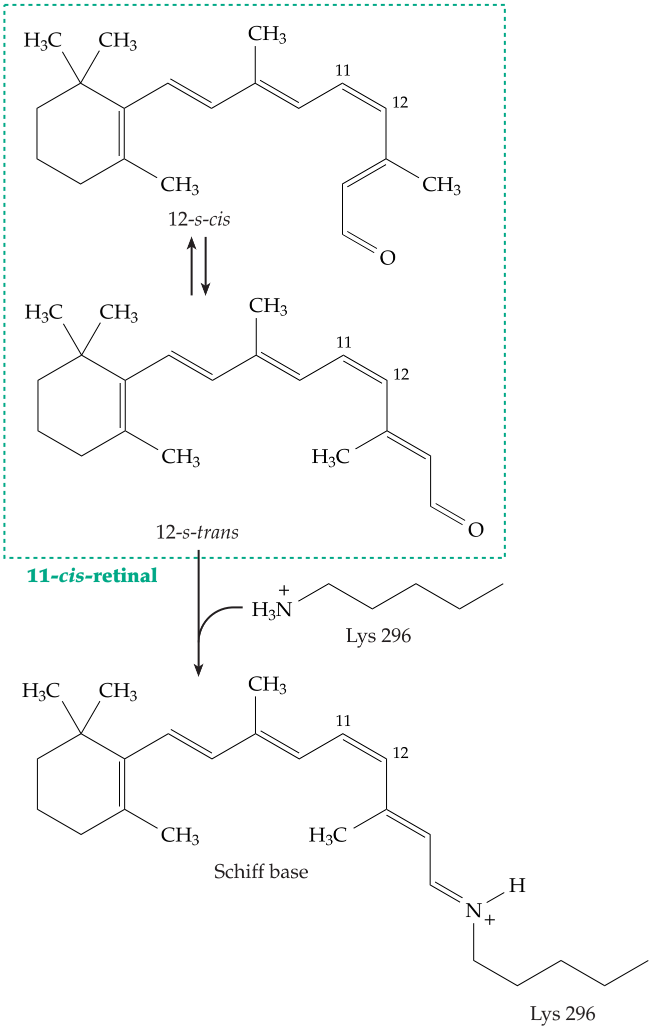


Figure 23-42 Absorption spectrum of cattle rhodopsin in aqueous dispersion with a nonionic detergent. From H. Shichi *et al.*[486,487]

P23H in rhodopsin is one of the most frequent causes of retinitis pigmentosa, and mutations G90D and A292E cause congenital night blindness. From Barnidge *et al.*[461] with modifications. (B) Structural model of rhodopsin based on the helix arrangement of Baldwin[456] and NMR constraints. The seven transmembrane helices, viewed from the cytoplasmic surface, are shown at three levels indicated by differences in shading. The α-carbons of residues of interest are shown in dots of various sizes, indicating the depth from the cytoplasmic domain. Gly 90, Glu 113, and Lys 296 are on the extracellular half of the transmembrane domain and close to each other in space. The 11-*cis*-retinal chromophore has been incorporated into the model using NMR constraints, which require a close interaction between Glu 113 and C_{12} of the chromophore. The relative position of the β-ionone ring and Trp 265 agrees well with crosslinking data of Zhang *et al*.[466] From Han and Smith.[467] (C) Stereoscopic view of the retinal-binding pocket of **bacteriorhodopsin** viewed from the cytoplasmic surface. The retinal, in Schiff base linkage to Lys 216, runs across the central cavity from top to bottom in this view. From Grigorieff *et al.*[468] (D) Ribbon drawing of bovine rhodopsin (stereoview). From Palczewski *et al.*[461a]


B
A
I
G90
B
II
G
K296
N-H
+
VII
O
O-C
E113
12
11
3
F
VI
W265
III
C
D
IV
E
V
C
Met-20
Tyr-57
Asp-85
Asp-212
Tyr-185
Trp-182
Trp-86
Tyr-83
Asp-115
Met-118
Trp-189
Trp-138
D
C
VIII
I
II
VII
VI
V
IV
III
4
3
2
1
N



11-*cis*-retinal

(23-36)

Three types of cone cells in the human retina are needed for color vision. Four genes specify the proteins for rhodopsin and for related cone photoreceptors absorbing blue (~425 nm), green (~530 nm), and red (~560 nm) light.[477–479b] All of the cone opsins also bind 11-*cis*-retinal. The rhodopsin gene is located on human chromosome 3, while that of the blue pigment is found on chromosome 7. However, the green and red sensitive pigment genes, whose sequences are 96% identical,[478] are close together on the q arm of the X-chromosome and near the gene for glucose-6-phosphate dehydrogenase. Examination of cloned DNA from persons with inherited red-green color-blindness shows that loss of a functional form of one of these genes is usually responsible for the problem.[480–482] Among Caucasians 8% of males and 1% of females differ from the normal in their color vision. About 30% of affected males are **dichromats** and lack either the red-sensitive pigment (they are protanopes) or the green-sensitive pigment (deuteranopes). They usually have a partial gene deletion. The other 70% often have hybrid genes created by errors in recombination during cell division.[483] A few deuteranopes have the point mutation C203R in the green pigment. John Dalton, of atomic theory fame, reasoned that his red-green confusion resulted from a blue tint in the vitreous humor of his eyes and ordered that they be dissected after his death (in 1884). There was no blue tint but DNA analysis performed more recently showed that Dalton was a deuteranope.[484] Defects in the blue-sensitive receptor are relatively rare affecting about 1 in 500 persons, while only one person in 100,000 has a total lack of color discrimination.[481,485]

All retinal-dependent visual pigments form Schiff bases with lysine side chains of the photoreceptor proteins. How can the same chromophore be "tuned" to absorb across the wavelength range of 360 to 635 nm? Modern techniques such as resonance Raman[477,479] and FTIR spectroscopies and study of mutant forms[488] have shown that interaction of the conjugated double bond system of the chromophores with immediately adjacent dipoles of side chain groups and peptide linkages is sufficient to account for the great variability in absorption maxima.

Visual pigments of many species have been investigated. Most vertebrate animals have a rhodopsinlike pigment plus a variable number of cone pigments. Mammals typically have only two, a short-wavelength pigment absorbing maximally in the ultraviolet, violet, or blue region[488a-c] and a long-wavelength pigment with maximum absorption in the green or red region.[489] The bottlenose dolphin has only a rod pigment of λ_{max} 524 nm.[489] In contrast, the chicken has four cone pigments with maximum absorption for violet, blue, green, and red.[490,490a] The red light receptor, called **iodopsin**, absorbs maximally at 571 nm. However, it binds chloride ions which induce an additional 40-nm red shift. The Cl^--binding site involves His 197 and Lys 200 which are present in an extracellular loop (Fig. 23-41A) and quite far from the bound retinal.[490] Human red and green color vision pigments, and also a green-sensitive pigment of the reptile *Gecko gecko*, undergo spectral shifts upon binding of Cl^- in the same site.[491] However, rhodopsin and most other visual pigments do not share this behavior.

Fishes live in a variety of environments and have a diversity of visual pigments. Goldfish have genes for five opsins, one of which gives rise to an ultraviolet light receptor. They are also sensitive to polarized light.[492] Related visual pigments occur throughout the animal kingdom. Even the eyespot of the alga *Chlamydomonas* (Fig. 1-11) contains rhodopsin with some sequence homology to invertebrate opsins.[493,494] The pineal glands of chickens and probably of reptiles[495–496a] as well as those of fish[497] also contain rhodopsinlike pigments. In a few freshwater marine species the visual pigments (**porphyropsins**) contain **3-dehydroretinal**. The peak positions of light absorption depend both upon the nature of the bound aldehyde and on the protein, the latter having the larger effect. Thus, retinal-based pigments absorb in the entire range 467–528 nm (18,900–21,400 cm^{-1}). The fruit fly, *Drosophila*, contains 3-hydroxy-11-*cis*-retinal in its rhodopsin and also contains other related photoreceptors.[498]

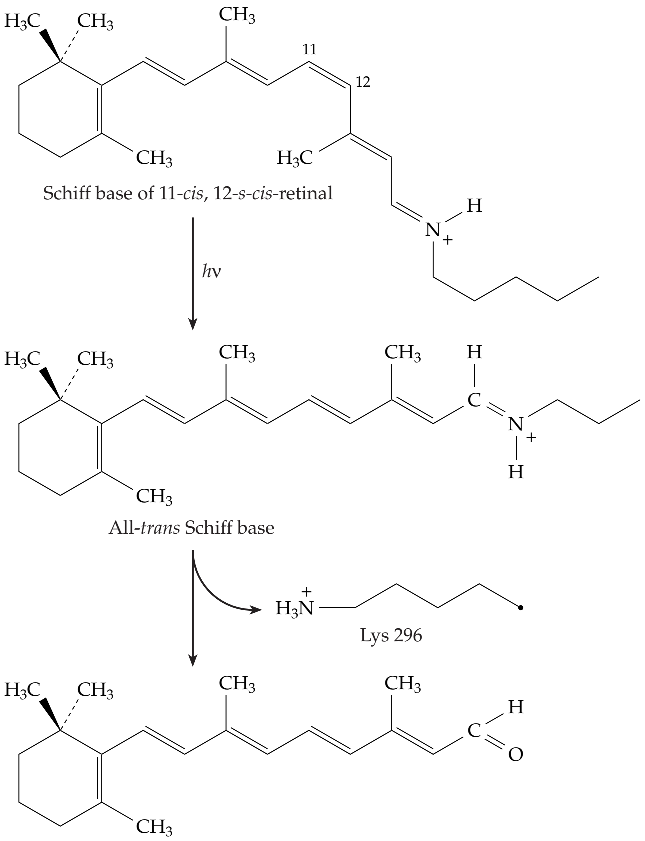
2. The Light-Induced Transformation

The retinal Schiff base chromophore is embedded in rhodopsin with its transition dipole moment parallel to the plane of the discs, i.e., perpendicular to the direction of travel of the incoming photons. Absorption of a photon leads to a sequence of readily detectable spectral changes.[37,461b,499,500] The relaxation times indicated in Eq. 23-37 are for 20°C.

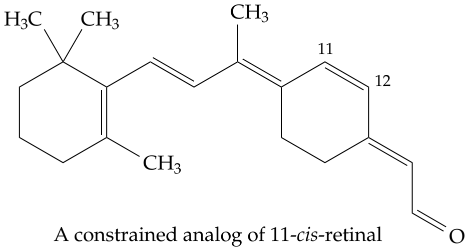
Rhodopsin (498 nm, 20.1 x 10^3 cm^{-1})
*h*ν ↓ $\tau < 6 \times 10^{-12}$s
Bathorhodopsin (prelumirhodopsin; 543 nm, 18.4 x 10^3 cm^{-1})
↓ τ ~50 ns
Lumirhodopsin (497 nm), 20.2 x 10^3 cm^{-1})
↓ τ ~ 50 µs
Metarhodopsin I (478 nm, 20.9 x 10^3 cm^{-1})
↓↑ milliseconds
Metarhodopsin II (380 nm, 26.2 x 10^3 cm^{-1})
↓ seconds
All-*trans*-retinal (387 nm, 25.8 x 10^3 cm^{-1}) + opsin (23-37)

The intermediate chemical species have been named bathorhodopsin, lumirhodopsin, and metarhodopsin I and II. At very low temperatures a transient form **photorhodopsin** with a wavelength maximum at ~580 nm may precede bathorhodopsin.[461b,501–502a] Furthermore, nanosecond photolysis of rhodopsin has revealed a **blue-shifted intermediate** that follows bathorhodopsin within ~40 ns and decays into lumirhodopsin.[500,503,504] The overall result is the light-induced isomerization of the bound 11-*cis*-retinal to all-*trans*-retinal (Eq. 23-38) and free opsin. The free opsin can then combine with a new molecule of 11-*cis*-retinal to complete the photochemical cycle.

What are the chemical structures of the intermediates in Eq. 23-37, and why are there so many of them? The answer to the last question is that the initial photochemical process is very fast. Subsequent conformational rearrangments and movement of protons are slower, occur in distinct steps, and give rise to the observed series of intermediates. To shed light on these processes many experiments have been done with analogs of retinal,[502,505–508] often using very rapid spectroscopic techniques.[37,508] These studies have shown that the isomerization of the Schiff base from 11-*cis* to all-*trans* occurs in the first very rapid step of Eq. 23-37.[499,509–510a] Constrained analogs of 11-*cis* retinal also combine with opsin to form rhodopsinlike molecules with absorption maxima near 500 nm.



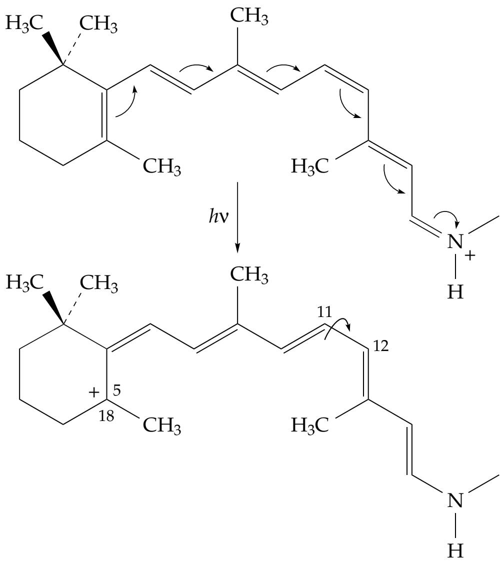
(23-38)



A constrained analog of 11-*cis*-retinal

However, most of these analogs cannot isomerize and illumination does not produce bathorhodopsinlike molecules.[504,509,511]

In the photoexcited state the positive charge on the Schiff base is delocalized. For example, with some of the positive charge located on C5, rotation could occur around the more nearly single C11–C12 bond to give the all-*trans* isomer (Eq. 23-39). However, it seems more likely that simultaneous rotation occurs around two connected single and double bonds.[511a] The mechanism of this photoisomerization, which is among the fastest known chemical reactions, is still being investigated.[511b] In native rhodopsin the conversion to bathorhodopsin occurs with a high quantum



(23-39)

yield of 0.67 within 0.2 ps, a time comparable to the period of torsional vibrations of the retinal. This extreme speed suggests that the isomerization is a concerted process that is **vibrationally coherent**.[511c] Vibrational motion in the electronically excited state is utilized in the isomerization process.[506,512]

The reaction sequence of Eq. 23-37 can be slowed by lowering the temperature. Thus, at 70K illumination of rhodopsin leads to a **photostationary state** in which only rhodopsin, bathorhodopsin, and a third form, **isorhodopsin**, are present in a constant ratio.[510] Isorhodopsin (maximum absorption at 483 nm)[513] contains 9-*cis*-retinal and is not on the pathway of Eq. 23-37. Resonance Raman spectroscopy at low temperature supports a distorted all-*trans* structure for the retinal Schiff base in bathorhodopsin.[510] The same technique suggests the *trans* geometry of the C=N bond shown in Eqs. 23-38 and 23-39. Simple Schiff bases of 11-*cis*-retinal undergo isomerization just as rapidly as does rhodopsin.[514]

Some step in the sequence of Eq. 23-37 must initiate a chemical cascade that sends a nerve impulse out of the rod cell. This is accomplished through a GTP-dependent G protein cascade as outlined in Fig. 23-43. Light-activated rhodopsin initiates the cascade by interacting with the G protein **transducin**. Energy for the activation comes from the quantum of absorbed light. While the primary chemical reaction has long been recognized as the isomerization of the retinal Schiff base,[515] it is not obvious how this generates the signal for transducin to bind and become activated.

The seven helices of rhodopsin form a “box” around the bound retinal. The environment of the retinal is largely hydrophobic. However, there are also buried polar groups, some of which lie in conserved positions in more than 200 G-protein-coupled receptors[458] and internal water molecules whose vibrational spectra can be detected.[516] As in bacteriorhodopsin (Fig. 23-41, C) the buried polar groups and water molecules are doubtless hydrogen-bonded in an internal network. We can anticipate two effects of the isomerization reaction: (1) It will distort the shape of the box in which it occurs. (2) It will break some hydrogen bonds and allow new ones to form and may affect the balance of electrical charges within the protein. This in turn can lead to proton movements and alterations in the internal hydrogen-bonded network. Both of these anticipated effects have been observed.

Conformational changes induced in the rhodopsin protein by the isomerization of the retinal Schiff base include significant movement of the end of helix VI(F) at the cytosolic surface as well as smaller movements of other helices.[517,518] Spectroscopic measurements indicate that the Schiff base nitrogen remains protonated in both the bathorhodopsin and metarhodopsin I forms and in metarhodopsin II, the first long-lived form in the sequence.[518a,b] It seems likely that the proton has jumped via a bridging water molecule and the E113 carboxylate to the external (intradiscal) surface of the molecule. At the same time one or more protons are apparently taken up on the cytoplasmic side.[519] Study of mutant forms suggests that glutamate 134, near the cytosolic surface, and histidine 211 may be involved in proton transport.[520] By analogy with bacteriorhodopsin, aspartate 83 is probably also involved. The combination of conformational change plus altered charge distribution may be needed to create a binding surface with a suitable shape and charge constellation to bind tightly to transducin for the next step.

3. The Nerve Impulse

Which of the intermediates in Eq. 23-37 is responsible for initiation of a nerve impulse? Some evidence favored metarhodopsin I,[326] but its lifetime may be too short. On the other hand, the transformation of metarhodopsin I to metarhodopsin II is the slowest step that could trigger a nerve impulse, which must travel the length of the rod to the synapse in about one ms,[521] and metarhodopsin II is generally believed to be the activated signaling form of rhodopsin.[521a–c]

Transducin, cyclic GMP, and phosphodiesterase. The essential consequence of light absorption is an alteration in the membrane potential in the vicinity of the absorbed photon with the resulting propagation of a nerve impulse down the plasma membrane to the synapse by cable conduction (Chapter 30). The type of potential change that is transmitted differs among vertebrates and invertebrates.[522] In the case of mammalian photoreceptors the rod outer segment is permeable to sodium ions so that a large

dark current of sodium ions flows in through the plasma membrane and is pumped out by sodium pumps in the inner portion of the cell. Visual stimulation causes this permeability to Na^+ to be decreased with an increase in polarization of the membrane. Absorption of a single photon by rhodopsin blocks the outflow of ~10^6 sodium ions.

At one time calcium ions seemed to be the logical internal messenger between rhodopsin and the plasma membrane. If light absorption opened channels from the internal space of the rod discs, calcium ions could be released and diffuse quickly to the plasma membrane and block the entrance of sodium ions.[523] However, light *does not* increase the free [Ca^{2+}] in the cytoplasm but may decrease it from 500 nM to as low as 50 nM.[524,525] Stryer suggested that the essential messenger is **cyclic GMP** (Chapter 11)[526–528] and that a *decrease* in cGMP concentration initiates the nerve response. Cyclic GMP is apparently responsible for keeping the sodium ion channels open. Absorption of a photon in the rod disc (Fig. 23-43) produces activated rhodopsin R* (metarhodopsin II), which acts as an allosteric effector for the heterotrimeric G protein transducin whose structure and properties have been discussed in Chapter 11. Like proteins G_s and G_i of the adenylate cyclase system, transducin contains three subunits: α, 40 kDa, 350 residues; β, 36 kDa; and γ, ~8 kDa.[529] In the resting state they are associated as $T_{\alpha\beta\gamma}$ with a molecule of GDP bound to the α subunit. When activated rhodopsin R* binds to transducin (step *a*, Fig. 23-43) it catalyzes a rapid exchange of GTP for GDP (step *b*). This is followed by dissociation of $T_{\beta\gamma}$ from T_α GTP. The latter serves as an allosteric effector for a **cGMP phosphodiesterase** bound to the disc surface converting it to an active form (step *c*).[529a,b] The activated phosphodiesterase, an $\alpha\beta\gamma_2$ oligomer,[530] hydrolyzes the cGMP (step *d*, Fig. 23-43), reducing its concentration and thereby inhibiting the Na^+ outflow.

Because one molecule of activated cGMP phosphodiesterase can hydrolyze more than 10^5 molecules of cGMP per second the light response is highly amplified. There is also an earlier stage of amplification. Each molecule of light-activated rhodopsin (R*) is able to catalyze the exchange of GTP for GDP on hundreds of molecules of $T_{\alpha\beta\gamma}$ before R* passes on to other intermediates and releases all-*trans*-retinal from opsin (light green lines, Fig. 23-43).

Rhodopsin kinase, recoverin, and arrestin. Metarhodopsin II (R*) can become phosphorylated by rhodopsin kinase on as many as seven serine and threonine side chains on its cytoplasmic surface (Fig. 23-41).[531,532] The 45-kDa protein arrestin binds to such phosphorylated R*,[533–535a] which is rapidly deactivated and desensitized so that it is less likely to be immediately reactivated. This is important in the adaptation of the eye to bright light. At the same time $T_\alpha 2GTP$ is hydrolyzed back to $T_\alpha 2GDP$ and reforms $T_{\alpha\beta\gamma}2GDP$, and guanylate cyclase regenerates the cGMP.[536] At least four different arrestins are known. Some function in nonvisual tissues. In all cases they seem to serve as "uncouplers" of G protein-coupled receptors.[536a]

Recovery of the inhibited rhodopsin, which occurs most rapidly in dim light, depends upon calcium ions. In dim light both Ca^{2+} and Na^+ enter the visual cells through the cGMP-controlled channels. At the same time Ca^{2+} flows out through a Na^+/Ca^{2+} exchanger. When the channels are blocked by cGMP formed in

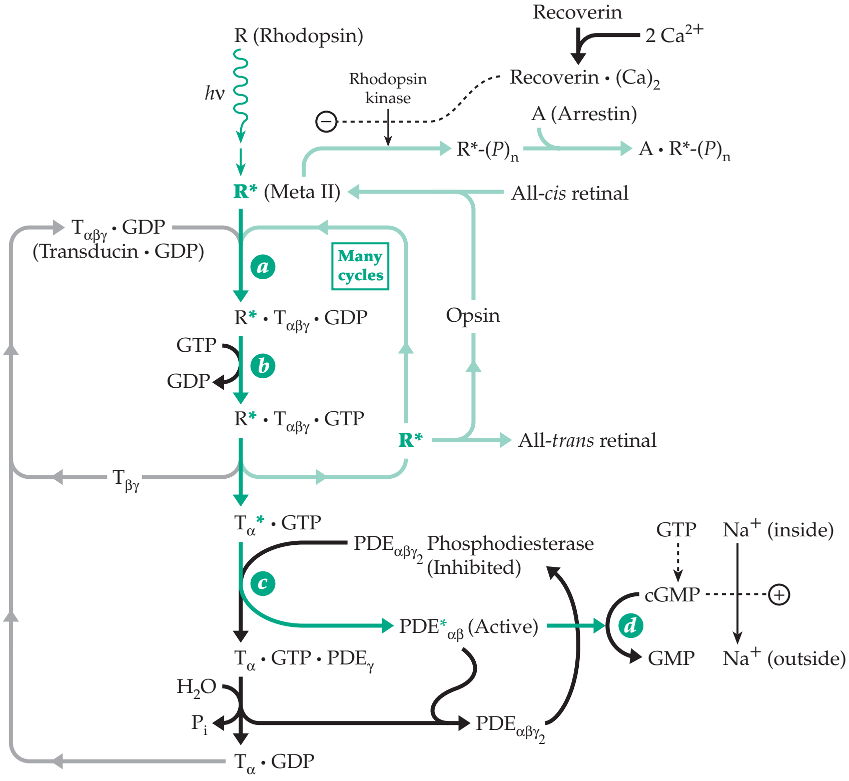


Figure 23-43 The light-activated transducin cycle. In step *a* photoexcited rhodopsin (R*) binds the GDP complex of the heterotrimeric transducin ($T_{\alpha\beta\gamma}$). After GDP–GTP exchange (step *b*) the activated transducin T*•GTP reacts with the inhibited phosphodiesterase ($PDE_{\alpha\beta\gamma 2}$) to release the activated phosphodiesterase ($PDE_{\alpha\beta}$). Based on scheme by Stryer[528] and other information.

response to light, [Ca^{2+}] falls as a result of continuing export by exchange with Na^+.[524,537] In the dark the [Ca^{2+}] concentration rises again and binds to a 23-kDa calcium sensor molecule called recoverin. Recoverin, like calmodulin (Fig. 6-8), has four EF-hand Ca^{2+}-binding motifs and also an unexpected feature. A **myristoyl group** attached at the N terminus is bound into a deep hydrophobic pocket in the protein. However, when two Ca^{2+} ions bind, recoverin undergoes a conformational change that allows the myristoyl group to be extruded and to bind to a membrane surface. This allows recoverin to move out of the cytosol to the surface of the disc membrane where it binds to and inhibits the activity of rhodopsin kinase,[538–541] increasing the sensitivity of photodetection.

Some details about cone cells and invertebrate vision. The biochemistry of retinal cones is less well known but is similar to that of rod cells. Cone pigments are present in the plasma membrane rather than in isolated discs (Fig. 23-40C). Different α, β, and γ subunits of transducin are formed in rods and cones.[522] Many differences are seen among various invertebrate visual systems. Inositol triphosphate (IP_3) and Ca^{2+} often serve as signals of photoexcitation. G proteins also play prominent roles.[522]

4. Regeneration of Visual Pigments; the Retinal Cycle

How is all-*trans*-retinal released from photobleached pigments and isomerized to 11-*cis*-retinal for the regeneration of the photopigments? Since new 11-*cis*-retinol is continuously brought in from the bloodstream and oxidized to retinal, isomerization can occur in other parts of the body. However, much of it takes place in the **pigment epithelium** of the retina, the layer of cells immediately behind the rod and cone cells. As indicated in Fig. 23-44, all-*trans*-retinal can leave the photoreceptor cells and, after reduction to retinal, be carried to the pigment epithelial cells by an **interphotoreceptor retinoid-binding** protein.[542,543] There it becomes esterified by the action of **lecithin:retinol acyltransferase**, an enzyme that transfers a fatty acyl group from lecithin to the retinol. The resulting retinyl esters are isomerized, and 11-*cis*-retinol is released.[543a] Some is stored as 11-*cis*-retinyl esters but enough is dehydrogenated to 11-*cis*-retinal to meet the needs of the photoreceptor cells and is transported back to them (Fig. 23-44). In the cephalopods the inner segment of the receptor cells contain a second pigment **retinochrome** that carries out a photochemical conversion of all-*trans*-retinal to 11-*cis*-retinal.[544,544a]

5. Diseases of the Retina

An important cause of blindness is **retinitis pigmentosa**, an inherited disease affecting about one in 3000 persons. Symptoms include progressive night blindness, degeneration of the rod cells, and gradual loss of cone cells and of nerve function in the retina. An autosomal dominant form of the disease arises from deletions or point mutations in the rhodopsin gene. In the United States 15% of cases arise from the mutation P23H.[545] By 1996 ~70 point mutations that cause retinitis pigmentosa had been discovered.[448,546–548] These mutations are found in all three of the rhodopsin domains: intradiscal, transmembrane, and cytosolic. Other rhodopsin point mutations such as G90D and A292E cause congenital night blindness.[549] Retinitis pigmentosa also arises from defects in **peripherin-RDS**, a structural component of rod cells identified originally from the gene *rds* (retinal degeneration slow) of the mouse.[545,550] Another form of congenital night blindness results from mutations in rhodopsin kinase.[551] A dominant **rod–cone dystrophy** is caused by a defect in the photoreceptor guanylate cyclase.[552] The most frequent cause of combined deafness and blindness in adults (**Usher syndrome**) is a defect in a cell adhesion molecule.[553] In **choroidemia**, another X-linked form of retinitis pigmentosa,

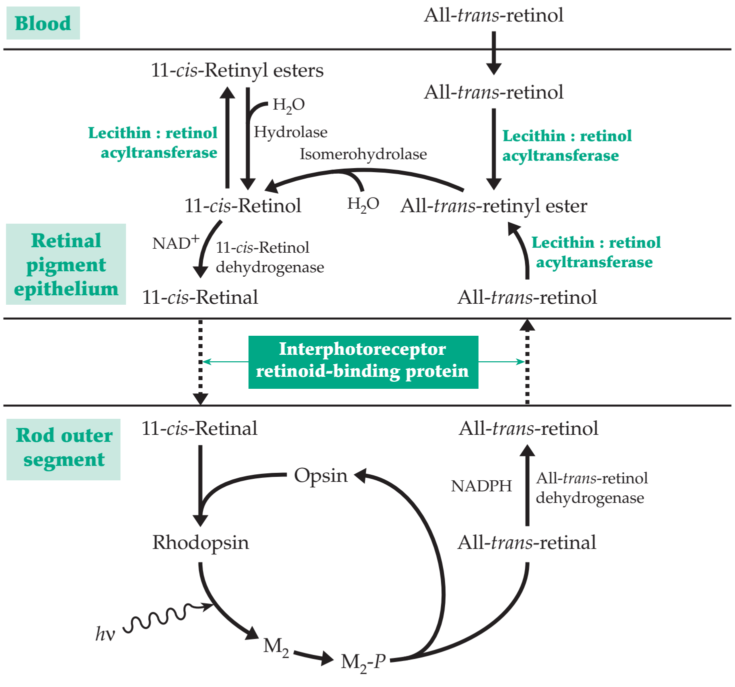

Blood
All-trans-retinol
11-cis-Retinyl esters
All-trans-retinol
H2O
Hydrolase
Lecithin : retinol acyltransferase
Lecithin : retinol acyltransferase
Isomerohydrolase
11-cis-Retinol
H2O
All-trans-retinyl ester
Retinal pigment epithelium
NAD+
11-cis-Retinol dehydrogenase
Lecithin : retinol acyltransferase
11-cis-Retinal
All-trans-retinol
Interphotoreceptor retinoid-binding protein
Rod outer segment
11-cis-Retinal
All-trans-retinol
Opsin
NADPH
All-trans-retinol dehydrogenase
Rhodopsin
All-trans-retinal
hv
M2
M2-P

Figure 23-44 Reactions of retinol and the retinal cycle of mammalian rod cells. After Palczewski *et al.*[543]

the choroid layer behind the pigment epithelium also deteriorates. A geranylgeranyl transferase specific for the Rab family of G proteins is defective.[554]

The most frequent cause of vision loss in the elderly is **macular degeneration**. Mild forms of the disease occur in nearly 30% of those over 75 years of age and more serious forms in 7% of that age group. There are many causes, some hereditary.[555–556a] Excessive accumulation of fluorescent lipofuchsin, perhaps arising in part from Schiff base formation between retinal and phosphatidylethanolamine, is sometimes observed.[557]

6. Proteins of the Lens

The lens of the eye encloses cells that cannot be replaced and contains proteins that don't turn over and must last a lifetime.[558] The fiber cells, which make up the bulk of the lens, have no nuclei. They elongate and stretch to cover the central nucleus, the original fetal lens, like the layers of an onion, the edges of the cells interdigitated with the next cell like a piece in a child's construction set.[559] These cells are tightly packed with proteins in aggregates whose size is on the order of the wavelengths of light. The high concentration of proteins is needed to provide transparency and also a high refractive index.[560–562] The membranes of the lens cells acquire increasing amounts of a 28-kDa **major intrinsic protein** as they age.[563] Three classes of soluble lens proteins, called **crystallins**, are found in virtually all lenses. Alpha crystallins, which account for ~40% of the total soluble protein, are heterodimers of ~20-kDa subunits that associate into ~800-kDa complexes.[564] They have a chaperoninlike activity.[565] Beta crystallin, which may account for ~35% of the protein, as well as the γ crystallins are β-sheet proteins with "Greek key" folding motifs.[558]

In addition to the α, β, and γ crystallins many animals have recruited additional "taxon-specific" crystallins δ, ε, λ, etc., that have evolved from preexisting enzymes, chaperonins, or other proteins.[561,566–568] For example, avian and reptilian lenses contain a δ crystallin homologous to argininosuccinate lyase.[569] Many crystallins are derived from dehydrogenases, e.g., for lactate dehydrogenase (duck),[570] hydroxyacyl-CoA dehydrogenase (rabbit),[571] and aldehyde dehydrogenase (squid and octopus).[572] A high concentration of NADH may be present.[568] A crystallin of a diurnal gecko is a retinol-binding protein with bound 3-dehydroretinal (vitamin A_2), which probably acts as an ultraviolet filter that improves visual acuity and protects against ultraviolet damage.[567,572a] Human lenses contain small molecules that act as UV filters, e.g., glucosides of **3-hydroxykynurenine** (Fig. 25-11) and 4-(2-amino-3-hydroxyphenyl)-4-oxobutanoic acid. Lenses tend to discolor and become fluorescent with age, in part by irreversible reactions of crystallins with these compounds.[573–574a]

A common problem with lenses is **cataract**, a term that describes any loss of opacity or excessive coloration. There are many kinds of cataract, most of which develop in older persons.[560,562] Since lens proteins are so long-lived deamidation of some asparagine occurs. However, the reactions are slow. One of the asparagines in α crystalline has a half-life of 15–20 years, and some glutamines are undamaged after 60 years.[575]

G. Bacteriorhodopsin and Related Ion Pumps and Sensors

Under certain conditions, the salt-loving *Halobacterium salinarum* forms a rhodopsinlike protein, which it inserts into patches of **purple membrane** in the surface of the cell. These membranes, which may constitute up to 50% of the cell surface, contain light-operated proton pumps that translocate protons from the inside to the outside of the cells.[454,576–578] In this manner they may provide energy for a variety of cell functions including ion transport and ATP synthesis. The 248-residue retinal-containing **bacteriorhodopsin** makes up 75% of the mass of the membrane. Its molecules aggregate into a two-dimensional crystalline array in the purple patches of the membrane. This allowed determination of the three-dimensional structure to 0.7 nm resolution in 1975 by electron microscopy and neutron diffraction.[454] More recently the structure has been established at progressively higher resolution by electron crystallography[579–581] and X-ray diffraction.[455,582] The most recent studies have been focused on determination of the structural alterations in the protein that accompany the proton pumping.[582a–e] A step-by-step picture is emerging.[582f] Internal changes in the retinal chromophore, movements of protons, and alterations in the shapes of some of the protein helices are involved. The surface loops have been studied both by electron crystallography and by atomic force microscopy.[583] Each bacteriorhodopsin molecule is folded into seven closely packed α-helical segments which extend roughly perpendicular to the membrane. Although 100 residues shorter than rhodopsin, the folding pattern is very similar (Figs. 23-41C; 23-45). The protein molecules form an extremely regular array with phospholipid molecules (mostly of phosphatidylglycerol) filling the spaces between them. The retinal is buried in the interior of the protein and is bound as an *N*-protonated Schiff base with the side chain of lysine 216.

The retinal in bacteriorhodopsin (bR) exists in both all-*trans* and 13-*cis* configurations. The all-*trans* form has an absorption maximum at 568 nm and the 13-*cis* at 548 nm. The two exist in a slow nonphotochemically mediated equilibrium in the dark.[584,584a] However, in the light the all-*trans* bR_{568} undergoes a rapid photochemical cycle of reactions, which is presented in simplified form in Fig. 23-46. The subscript numbers designate the wavelength of maximum absorption. Many efforts have been made to determine the structures of the intermediates K, L, M, and O and to relate them to a mechanism of proton pumping.[585–585c] Both K and L contain 13-*cis* retinal. Therefore, as with rhodopsin (Eq. 23-39) the very first step is a photochemical isomerization. Intermediate M appears to contain a deprotonated Schiff base, but O is both *N*-protonated and again all-*trans*. It follows logically that the proton bound initially to the Schiff base is pumped out of the cell and is then replaced by a new proton in the O form. However, several questions must be answered if we are to understand this proton pump. Where in the sequence do proton transfers occur? How is the sequence driven by the absorbed light energy? Protons must enter the pump from the cytosol and exit on the exterior (periplasmic) side, flowing out against a concentration gradient. There must be a "gating" or "switch" mechanism that ensures that protons enter and leave the pump in the correct direction.[586]

Some aspects of a possible mechanism for pumping the single proton bound to the retinal Schiff base are included in Fig. 23-46. In bR_{568} the Schiff base bound proton is H-bonded, via a water bridge, to the carboxylate group of Asp 85 (Fig. 23-45B). The charge constellation in the interior of the protein, part of which is shown in this figure, is such that protonation of the Schiff base is stable and the pK_a of the protonated Schiff base is high, with estimates of 16 or above.[587] One cause of the high pK_a is the presence of the nearby negative charges on D85 and D212. Absorption of light and isomerization of the retinal causes a downward movement of the $=NH^+-$ group of the Schiff base[588] and facilitates movement of the Schiff base proton via the water molecule to the D85 carboxylate as indicated by the green arrows in Fig. 23-45B. Loss of the positive charge will instantly substantially raise the pK_a of D85 from a low value, while the loss of the negative charge will lower the pK_a of the Schiff base to closer to 7. The electrostatic interactions of the D212 and E194 carboxylates with the positive charge of R82 may also be altered. At some point in the sequence the interaction of R82 with the E194 carboxylate could cause E204, which is known by spectroscopic measurements to be protonated in the intermediate, to lose its proton to the outside. At some other point, perhaps between M_{412} and another intermediate, M_{408} (Fig. 23-46), a conformational switch must occur to limit flow of a proton back to E204 and to allow a new proton to enter from the cytosol. The D96 carboxylate is thought to accept this proton and to transfer it via a chain of

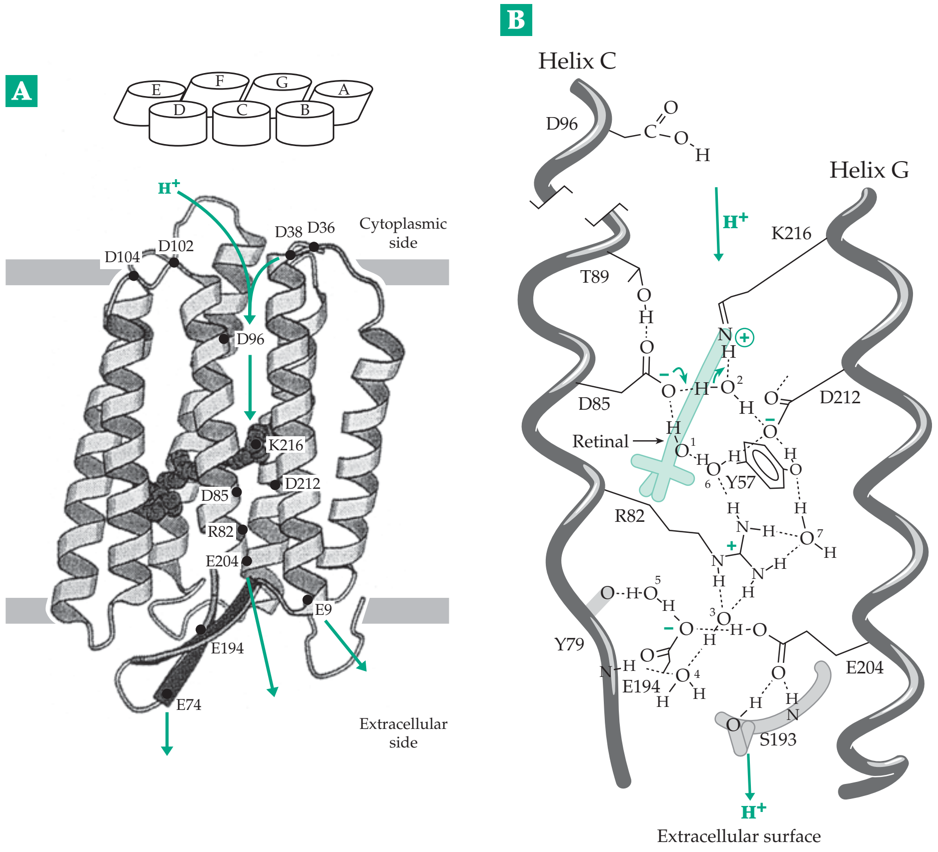


Figure 23-45 (A) Some aspects of the structure of bacteriorhodopsin. Ribbon diagram with the retinal Schiff base in ball-and-stick representation. At the top the helices are labeled as in Fig. 23-41. The locations of aspartate, glutamate, and arginine residues that might carry protons during the proton pumping action are indicated. Retinal is shown attached to lysine 216. From Kimura *et al*.[580] Courtesy of Yoshiaki Kimura. (B) Schematic drawing illustrating hydrogen-binding observed in the three-dimensional structure at 0.14 nm resolution. From Luecke *et al*.[455] The side chains shown are those thought to be involved in proton transport and in a hydrogen-bonded network with bound water molecules, principally between helices C and G. The positions of many of the hydrogen atoms in this network have not been established. They have been placed in reasonable positions in this figure but may be quite mobile. For another view of the hydrogen-bonded network see Fig. 23-41C.

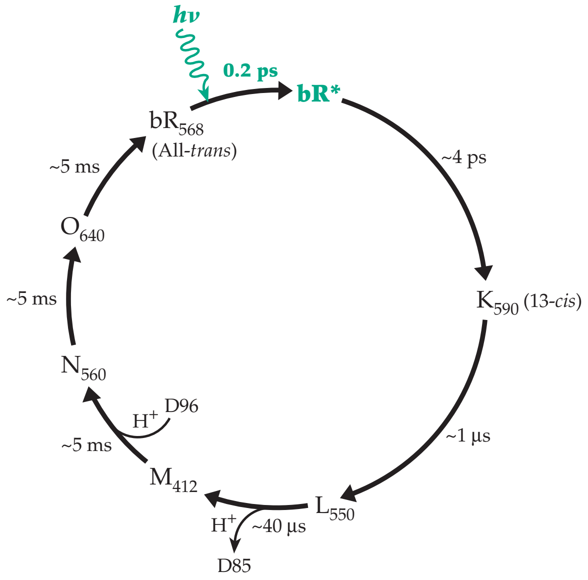
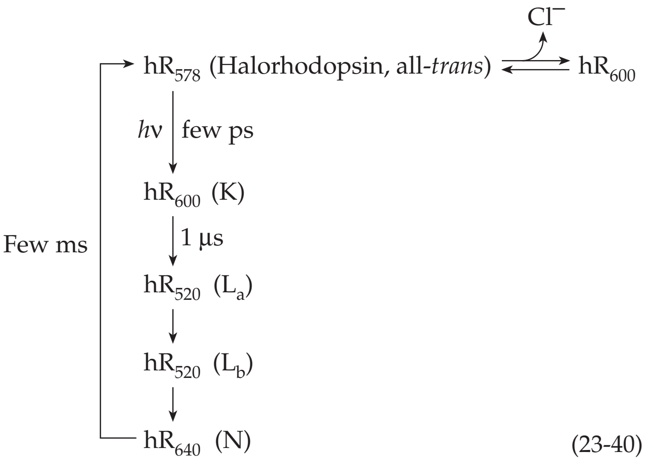


Figure 23-46 The photoreaction cycle of bacteriorhodopsin. After Bullough and Henderson.[585] The subscript numbers indicate the wavelengths of maximum absorption of each intermediate and the approximate lifetimes are given by the arrows. Resting bacteriorhodopsin as well as intermediates J and O have all-*trans* retinal but K through N are thought to all be 13-*cis*. A proton is transferred from L to aspartate 85 and then to the exterior surface of the membrane. A proton is taken up from the exterior surface via aspartate 96 to form N.

water molecules or water molecules plus protein groups to the Schiff base, which may isomerize back to all *trans* in going from intermediate N to O. Isomerization to all-*trans* O is a slow step. Both O_{640} and bR_{568} are all-*trans* and have a 15-*trans* Schiff base linkage. There may be a difference in protein conformation with the chromophore being more distorted in O than in bR_{568}. Many recent studies have provided additional information.[588a–j]

Halorhodopsin. In addition to bacteriorhodopsin there are three other retinal-containing proteins in membranes of halobacteria. From mutant strains lacking bacteriorhodopsin the second protein, **halorhodopsin**, has been isolated. It acts as a light-driven *chloride ion pump*, transporting Cl^- from outside to inside. Potassium ions follow, and the pump provides a means for these bacteria to accumulate KCl to balance the high external osmotic pressure of the environment in which they live.[578] The amino acid sequences of halorhodopsins from several species are very similar to those of bacteriorhodopsin as is the three-dimensional structure.[589] However, the important proton-carrying residues D85 and D96 of bacteriorhodopsin are replaced by threonine and alanine, respectively, in halorhodopsin.[590] Halorhodopsin (hR) undergoes a light-dependent cycle (Eq. 23-40) that involves an all-*trans* to 13-*cis* photoisomerization with some intermediates resembling those of the bacteriorhodopsin cycle.[590a]



(23-40)

Sensory rhodopsins. The third and fourth light-sensitive proteins of halobacteria are **sensory rhodopsins** (SR)[578,591,591a,b] that are used by the bacteria to control **phototaxis**. These bacteria swim toward long-wavelength light, the maximum in the action spectrum being at ~ 580 nm. They are repelled by blue or ultraviolet light, the maximum in the action spectrum being at ~ 370 nm. Evidently the bacteria can detect either a decrease with time in red light intensity or an increase with time in blue light intensity. Either is interpreted as unfavorable and causes the bacteria to tumble and move in a new direction (see Chapter 19). Sensory rhodopsin I (SRI) appears to be able to provide both light responses. Absorption of orange light by SRI_{587}, which contains all-*trans*-retinal, yields SRI_{373}, in which the retinal Schiff bases have been isomerized to 13-*cis* as in bacteriorhodopsin. The red light response is proportional to the fraction of SRI_{373} present. However, this is converted spontaneously back to SR_{587} within seconds. Nevertheless, photoexcitation of SRI_{373} with blue light causes a faster reconversion and induces swimming reversals, the repellent response.[592,593] SRI exists in the bacterial membranes in a complex with a 57-kDa protein designated **halobacterial transducer I** (HtrI), which resembles bacterial chemotaxis receptors (Figs. 11-8 and 19-5) and is modulated by action of a methyltransferase.[591b,c] Interaction of SRI with HtrI depends upon a histidine residue, H166 of SRI. It may be part of a proton transfer pathway.[593]

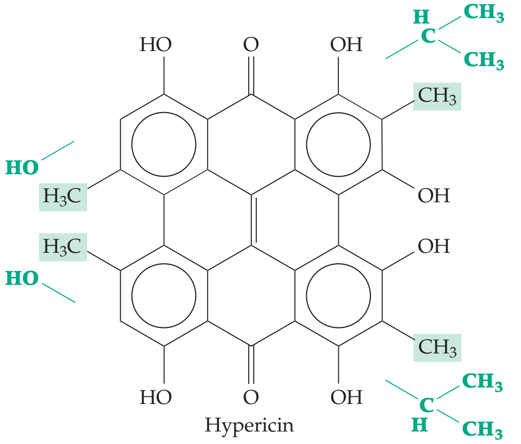
Sensory rhodopsin II (SRII, also called phoborhodopsin) is specialized for repellant phototaxis.[591a] Blue light converts $SRII_{487}$ in < 1 ms to UV-absorbing $SRII_{360}$. It decays in ~ 100 ms to $SRII_{540}$ which reverts to the initial $SRII_{487}$ in ~ 0.5 s. The cycle is accompanied by swimming reversals that result in a repellent effect of light.[594–596] The three-dimensional structure is known.[593a,b]

Whereas retinal-based proton pumps all have the conserved residues D85 and D96 of bacteriorhodopsin, only the aspartate correspnding to D85 is conserved (as D73) in the sensory rhodopsins. D96 is replaced by tyrosine or phenylalanine.[586] In SRI D73 appears to be protonated and, therefore, does not form a counterion for the Schiff base iminium ion.[597] However, in SRII D73 is apparently unprotonated and available to -serve as a counterion and as a proton acceptor as in bacteriorhodopsin.[597a,b] There is also a corresponding aspartate (Asp 83) in rhodopsin (Fig. 23-41). This suggests a common signaling mechanism for rhodopsin and the sensory rhodopsins. Finally, there are retinal-containing proteins in fungi and in algae. They may serve as blue light receptors.[598,598a]

The photoactive yellow proteins (xanthopsins). A 124-residue, 14-kDa yellow protein isolated from a halophilic phototrophic purple bacterium, *Ectothiorhodospira halophila*, was at first thought to be a rhodopsin-type pigment. However, this photoactive yellow protein (PYP) carries a covalently bonded **coumaroyl (4-hydroxycinnamoyl) group** in thioester linkage,[599–601c] which is completely surrounded by the small protein.[601] The coumaroyl group, which was probably derived from coumaroyl-CoA (Fig. 25-8), is bound as a phenolate anion by hydrogen bonds to tyrosine and glutamate side chains (Eq. 23-41). After a laser flash at room temperature a readily observed intermediate I_1 (also called pR) absorbing maximally at 460–465 nm appears within ~3 ns and decays within a few milliseconds to a bleached intermediate I_2 (also called pB or pM)[602,603] with maximum absorption at ~355 nm. This returns to the original 446-nm form within a few seconds. Earlier intermediates I_0 and $I_0^{\ddagger}$ have been identified by picosecond spectroscopy,[601c,604,605] and others have been identified at low temperatures.[602,606]

The structure of PYP is known to 0.1 nm resolution (Fig. 23-47).[601,607] Structures have also been determined for a very early intermediate by trapping at –100°C[607] and for I_1 (pR). The cofactor structures are shown in Eq. 23-41. The light-induced step is apparently the *cis–trans* isomerization,[608] and changes in hydrogen-bonding follow. The hydrogen bond between the phenolate ion of the coumaroyl group and glutamate 46 appears to break, and E46 may donate a proton to the phenolate group to form the 337-nm chromophore of I_2.[602,609,609a] The signaling mechanism may be similar to that in sensory rhodopsins.

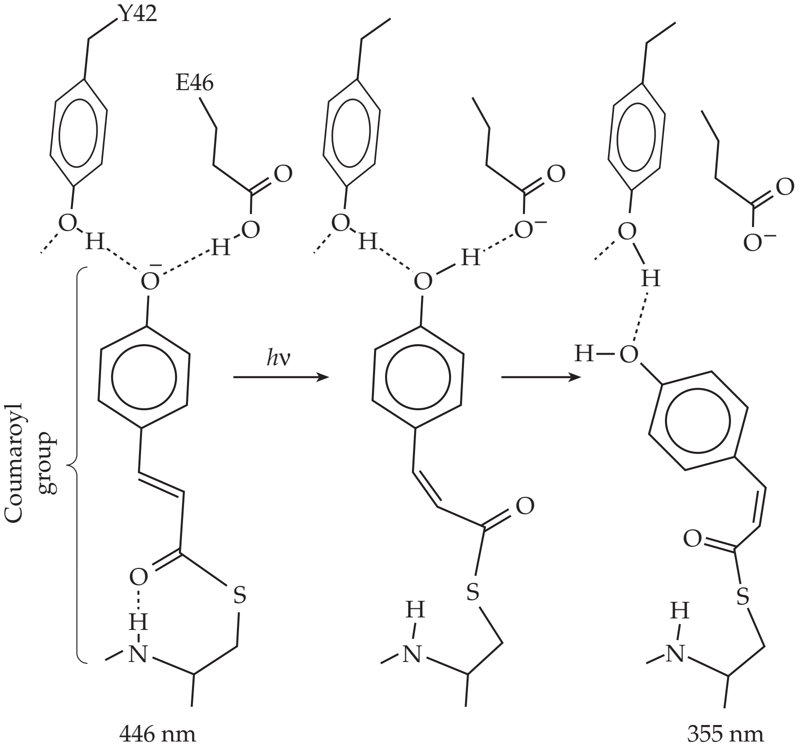
Stentorin. A protein with a bound chromophore called stentorin mediates the light-avoidance response of the protozoan *Stentor*. Stentorin,[610] which is found in pigment granules in the cell surface, is a derivative of **hypericin**, a plant compound with antidepressant activity and the active ingredient in the herb St. John's wort.



Hypericin

🟩 Replaced in stentorin by the green groups

Stentorin is covalently bonded to a 16-kDa protein in an acid-labile linkage. Its photocycle is not well investigated, but it is thought to initiate a response via



(23-41)

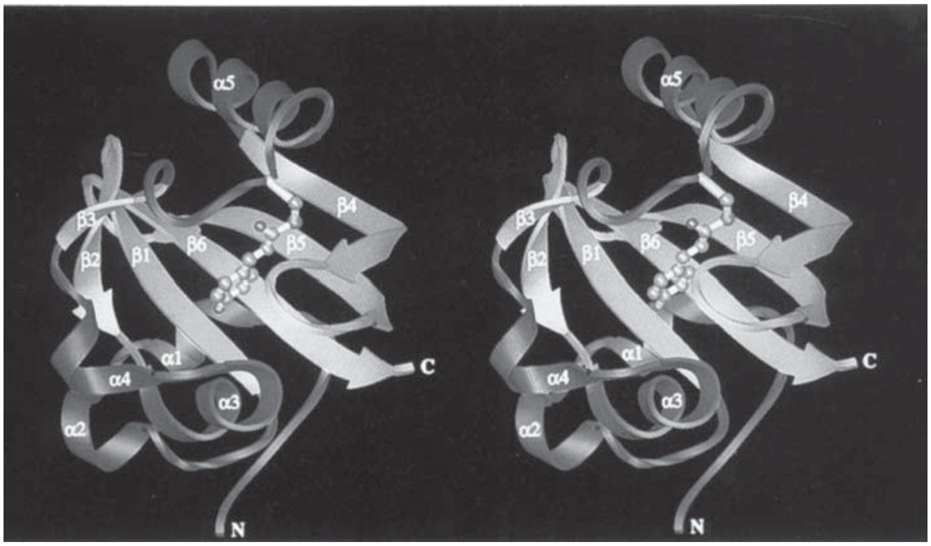

α5
β3
β2
β1
β6
β5
β4
α1
α4
α3
α2
C
N
α5
β3
β2
β1
β6
β5
β4
α1
α4
α3
α2
C
N

Figure 23-47 Ribbon drawing of the structure of the 125-residue yellow photoactive protein. The 4-hydroxycinnamoyl chromophore, which is attached to cysteine 69, is represented with balls and sticks. From Borgstahl *et al*.[601] Courtesy of Gloria Borgstahl.

---

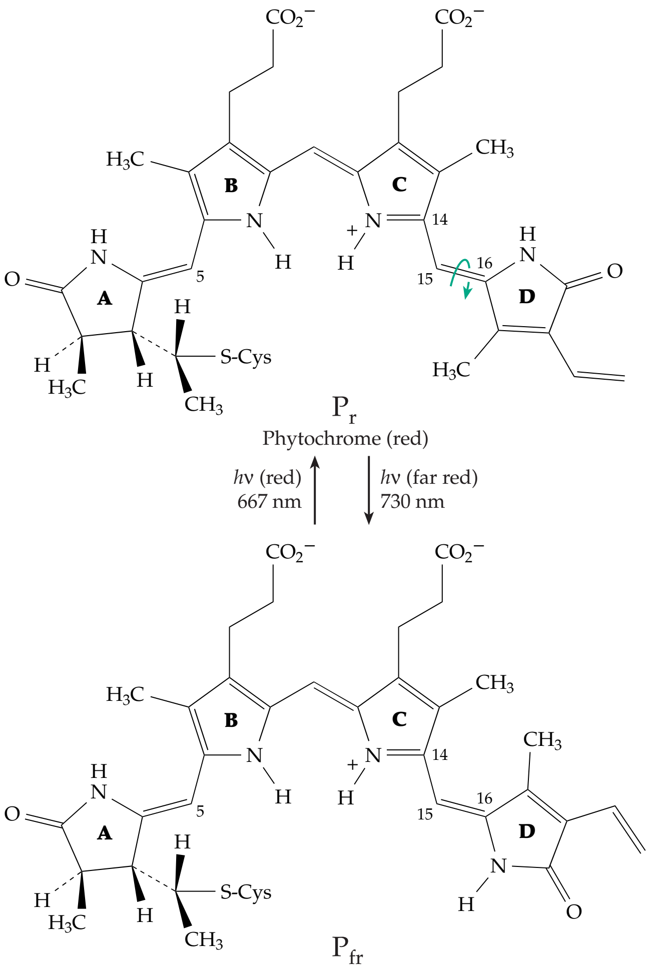
proton transfer.[611] However, stentorin proteins apparently do not belong to the bacteriorhodopsin family.

Hypericin and related compounds have also aroused interest because of their antiviral and antitumor activities.[612,613] Hypericin is a strong photoactivator which produces singlet oxygen with a quantum efficiency of 0.73. However, antiviral activity may involve a radical mechanism.[613] Hypericin is attractive as a possible agent for photodynamic therapy (Section D,3). It can also receive energy from photoexcited firefly luciferin (Section J). A proposed application is to incorporate the gene for the enzyme luciferase of the firefly luminescence system into DNA from the virus HIV. This DNA could be used to promote synthesis of luciferase only in virus-infected cells. Addition of the nontoxic hypericin would lead to photoactivation of hypericin only in virus-infected cells, where the luciferin–luciferase complex would act as a "molecular flashlight" to activate the hypericin and destroy the cell.[613]

H. Phytochrome

In 1951, it was discovered that a flash of red light (maximum activity at 660 nm) during an otherwise dark period promoted a variety of responses in plants.[614] These included flowering, germination of seeds (e.g., those of lettuce), and the expansion of leaves in dark-grown pea seedlings. Interestingly, the effect of the short flash of red light could be *completely reversed* if followed by a flash of *far-red light* (730 nm). This discovery led to the isolation, in 1959, of the chromoprotein phytochrome, a kind of molecular switch that initiates a whole series of far-reaching effects in plants. The phototransformation[615] is completely reversible (Eq. 23-42; Fig. 23-48), and the switch can be thrown in one direction or the other many times in rapid succession by light flashes.

Green plants have a family of phytochromes. There are five genes for the ~125-kDa chains of about 1100 residues each in *Arabidopsis*,[618–619c] and the corres-


CO_2^-
CO_2^-
H_3C
CH_3
B
C
N
H
+
N
H
14
H
N
O
A
5
15
16
H
N
D
O
H
S-Cys
H
H_3C
H
CH_3
H_3C
P_r
Phytochrome (red)
hν (red)
667 nm
hν (far red)
730 nm
CO_2^-
CO_2^-
H_3C
CH_3
B
C
N
H
+
N
H
14
CH_3
H
N
O
A
5
15
16
D
N
H
O
H
S-Cys
H
H_3C
H
CH_3
P_{fr}
Phytochrome (far red)

(23-42)

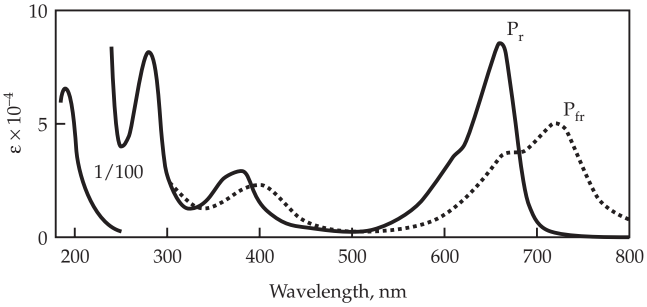


Figure 23-48 Absorption spectra of red (P_r) and far-red (P_{fr}) forms of purified oat phytochrome following saturating irradiations with red and far-red light. See Quail[616] and Anderson *et al*.[617]

sponding phytochromes A – E each appear to have distinct functions. The chromophore is an open tetrapyrrole closely related to phycocyanobilin and covalently attached to the peptide backbone near the N terminus through a cysteine side chain (top structure in Eq. 23-42). The initial photochemical reaction is thought to be the $Z \rightarrow E$ isomerizion around the C15 – C16 double bond, but there may also be rotation about the C14 – C15 single bond. The initial step occurs within a few microseconds and up to four intermediate species have been seen in the $P_r \rightarrow P_{fr}$ conversion and at least two other different ones in the $P_{fr} \rightarrow P_r$ conversion.[620–623c]

Phytochromes exist as two distinct domains, the N-terminal domain bearing the chromophore. However, the first 65 residues at the N terminus as well as the C-terminal domain probably interact with other proteins to transmit a signal. The slow responses to phytochrome are thought to involve regulation of transcription.[619b,623d] Thus, the synthesis of mRNA molecules specific for the small subunit of ribulose bisphosphate carboxylase and for the chlorophyll *a*/*b* binding protein increases in response to formation of P_{fr}. These responses are quite rapid occurring within 15–30 min.[614] Another response to P_{fr} is a decrease in the amount of the specific mRNA for phytochrome itself. That is, light induces a decrease in the concentration of this light-sensing molecule thereby decreasing the sensitivity of the system.[614]

Phytochrome is found not only in higher plants but also in algae, where it controls the movement of chloroplasts,[611] and also in cyanobacteria.[623e,f] Cyanobacterial phytochromes contain histidine kinase domains, which may function in a two-component system with a response regulator similar to protein CheY of the chemotaxis system in *E. coli* (Fig. 19-5).[624,625] Some nonphotosynthetic bacteria also use bacteriophytochromes for light sensing. In some cases biliverdin (Fig. 24-24) is the chromophore.[625a]

Phytochromes of higher plants also have histidine kinase-like N-terminal domains. Searches for associated signaling proteins have revealed a phytochrome-interacting factor in *Arabidopsis.* A possible partner for phytochrome B, it is a nuclear helix–loop–helix protein that may be a transcription factor.[626–627a] Phytochrome A may signal via a WD-repeat protein to control morphogenesis.[628]

One response under phytochrome control is the closing of leaflets of *Mimosa* at the onset of darkness. The response occurs within 5 min, too short a time to be the result of transcriptional control. This and the finding that some phytochrome is tightly bound to membranes have led to the proposal that one primary effect of phytochrome is to alter membrane properties. It is not certain whether it is P_r or P_{fr} that is active in causing a response, but P_{fr} seems to be the most likely candidate for the “active” form. According to one suggestion, phytochrome in plastid membranes may mediate the release of gibberelins stored within the plastids.[629]

I. Some Blue Light Responses

Numerous biological responses to light of wavelength 400–500 nm are known. These include phototropism in higher plants, the phototaxis of *Euglena*, and photorepair of DNA. On the basis of action spectra both carotenoids and flavins were long ago proposed as photoreceptors.[630–632] The action spectrum for opening of the stomates in coleoptiles matches the absorption spectrum of zeaxanthin.[631] On the other hand, genetic evidence[633,634] has strengthened the view that a flavin acts as the photoreceptor in the fungus *Phycomyces*. Recently compelling evidence for a flavoprotein receptor for phototropism in *Arabidopsis thaliana* has been obtained. Deficiency of a gene called *nph*1 (nonphototropic hypocotyl 1) is associated with loss of blue light-dependent phosphorylation of a 120-kDa protein. This protein was identified as the product of *nph*1 gene. The **nph1** protein is a soluble autophosphorylating Ser / Thr protein kinase with an N-terminal flavin-binding region. It apparently binds FMN and is a photoreceptor for phototropism in higher plants.[635–636b]

The complexity of the action spectra suggested the existence of more than one receptor.[634] In higher plants there are not only blue light receptors but also violet receptors and phytochrome. In addition to nph1 and a related protein **npl1**, *Arabidopsis* employs two cryptochromes (next section) and **phototropins**.[636c–e] These are also riboflavin 5'-phosphate (FMN)-dependent proteins. The action of light apparently causes addition of a highly conserved cysteinyl –SH group to the C4a position of the flavin.[636c] Phytochrome absorbs blue and ultraviolet light to some extent (Fig. 23-48) as well as red or far-red. This adds considerable complexity to the interpretation of light responses in plants.[637] The fern *Adiantum* contains a protein with a phytochrome photosensory domain fused to an NPH1 structure. It may mediate both real far-red and blue-light responses.[636] The protist *Euglena* (Fig. 1-9) makes use of a **photoactivated adenylate cyclase**, also a photoactivated enzyme, in a light avoidance response.[637a]

Fungi such as *Neurospora crassa* provide a simpler system for study of blue-light signaling than do green plants.[637] *Neurospora* contains no phytochromes. However, numerous genes including some involved in carotenoid biosynthesis and some that control the circadian cycle are regulated by blue light. Two mutants defective in riboflavin synthesis show a reduced sensitivity to blue light. A deficiency of either of two other genes *wc*-1 and *wc*-2 results in "blind" *Neurospora* unable to respond to light but able to grow. Proteins WC1 and WC2 are probably transcription factors, which act as a heterodimer. WC1, which contains bound FAD, is the photorecetor.[637,637b,c] Recently a rhodopsinlike protein NOP-1 of *Neurospora* has been identified.[637d]

1. Cryptochromes

The elusive nature of the principal blue-light receptor in plants gave rise to the name cryptochrome.[632] The gene for a cryptochrome in *Arabidopsis thaliana* was isolated by gene tagging and was cloned. It is surprisingly similar in sequence to the gene for the well-known **DNA photolyase** (Section 2).[638] It was soon recognized that cryptochromes, like photolyases, carry a bound flavin and also an antenna chromophore. The latter is probably a 5,10-methylenetetrahydrofolate, as in plant photolyases. It is a better light absorber than the flavin and passes electronic excitation to the flavin.

Two cryptochrome genes, *cry*1 and *cry*2, are present in *Arabidopsis*. The encoded proteins affect many aspects of plant growth. The cry-1 protein, together with NPH1, has a role in controlling phototropisin[639] while cryptochrome cry-2 affects flowering time, apparently via antagonistic signals from cry-2 and phytochrome B.[640–640b] cry-1 is also involved in controlling the daily rhythm of the plant, the **circadian cycle**. The circadian clock, which is discussed in Chapter 30, provides the organisms an oscillator with a period of about 24 hours. However, the oscillator must be **entrained** by the daylight cycle so that it remains in proper synchrony. The nature of the light signal and the mechanism of the entrainment are being investigated in many different organisms from fungi to human beings. In *Arabidopsis* the cycle is controlled by phytochromes A and B and by the cryptochrome cry-1.[641]

Cryptochrome genes have been found in many organisms. In the fly *Drosophila* cryptochrome appears to interact directly with the clock proteins that control the circadian cycle. Most important are products of two genes *per* (period) and *tim* (timeless). They are helix–loop–helix DNA binding proteins that form heterodimers, are translocated to the nucleus, and repress their own transcription. Morning light leads to a rapid disappearance of the TIM protein. The cryptochrome CRY appears to react directly with TIM to inactivate it. However, details remain to be learned.[642] The circadian clock mechanism appears to be universal and the cryptochrome-2 (*mcry*2 gene) appears to function in the mouse.[643,643a] A human cDNA clone was found to have a 48% identity with a relative of cryptochromes, the **(6–4) photolyase** of *Drosophila*. A second related human gene has also been found. The protein products of these two genes (*hcry*1, *hcry*2) lack photolyase activity. They too may encode cryptochromes.[644]

Where in the body is the light sensed for entraining the circadian cycle? Genes for CRY1 and CRY2 are specifically expressed in ganglion cells of the retina in mice. Severing of the optic nerve destroys both vision and light entrainment of mammals. However, in mice with the retinal degeneration (*rd*) syndrome all rod cells and virtually all cone cells are destroyed but the circadian rhythm is normal.[645] Furthermore, many blind persons with no conscious perception of light have normal light entrainment of their circadian cycle. For these reasons the ganglion cells of the retina, which are close to the location of the master circadian clock in the **suprachiasmatic nucleus** of the brain, are the most probable light sensory cells for the cycle[646,647] (see also Chapter 30). Recent evidence points to a retinal-based photoreceptor, **melanopsin**.[647a,b,c] However, vitamin A-deficient mice still display a normal circadian response.[647d]

2. Photolyases

A curious discovery was made many years ago. Bacteria given a lethal dose of ultraviolet radiation can often be saved by irradiating with visible or near ultraviolet light. This **photoreactivation**, which permits many bacteria to survive, results from the action of a **DNA photolyase**,[648,649] which often absorbs light maximally around 380 nm and carries out a photochemical reversal of Eq. 23-26, cutting the pyrimidine–pyrimidine covalent bonds of thymine dimers in DNA. The enzyme is present in cells in such small amounts, only 10–20 molecules per cell, that it was difficult to investigate until the gene had been cloned.[650,651] The significance cannot be doubted, for photoreactivation enzymes appear to be found in most organisms including some mammals. However, there is some doubt about the presence of a photolyase in the human body.

The *E. coli* DNA photolyase contains a blue flavin radical that arises from **FAD** and absorbs maximally at 580 nm (see also Chapter 15, Section B,6). The enzyme also contains a second chromophore in the form of bound 5,10-methenyltetrahydrofolylpolyglutamate with 3–6 γ-glutamyl residues.[652–653b] as shown in Fig. 23-49. The pterin coenzyme binds near the N terminus in a domain with an α/β folding pattern, while the FAD binds into a larger mostly helical domain. The pterin cofactor is not essential for repair activity, and it is generally agreed that because of its high molar extinction coefficient it acts as an effective **antenna**. It transfers energy in a nonradiative fashion to the $FADH^-$ anion located ~3 nm away.

The enzyme as isolated is in a stable blue radical form (Fig. 23-50; also Fig. 15-13) which must undergo a one-electron light-induced reduction to the anion $FADH^-$ before becoming active. A nearby indole ring of Trp may donate the electron and be reoxidized by a tyrosyl ring.[654] The $FADH^-$ donates an electron to the pyrimidine dimer, initiating the sequence of radical reactions[654a–c] which cleaves both pyrimidine–pyrimidine bonds in the photodimers (Fig. 23-50).

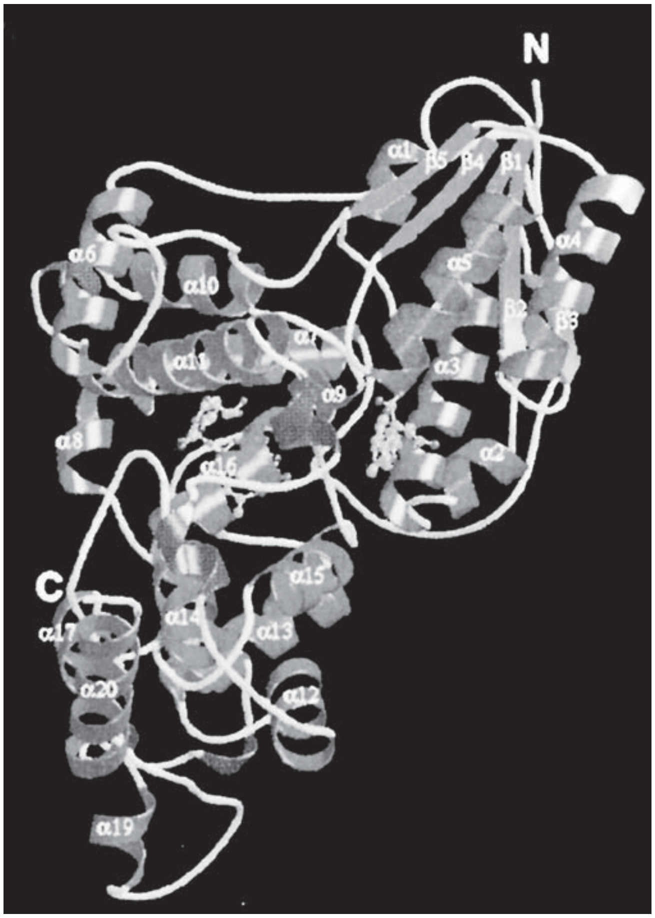


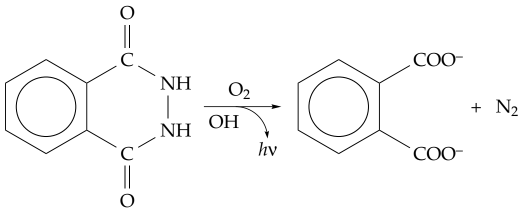
Figure 23-49 Overall view of the DNA photolyase structure from *E. coli*. The ribbon traces the 471-residue chain. The bound cofactors FAD (left) and 5,10-methenyltetrahydrofolate (right) are shown in ball-and-stick representation. From Park *et al*.[652] Courtesy of Johan Deisenhofer.

The structures of all of the photolyases are thought to resemble that in Fig. 23-49. However, in one large group, which includes methanogenic bacteria, **8-hydroxy-5-deazariboflavin** acts as the antenna chromophore.[649] Another light-induced defect in DNA is the so-called 6–4 photoproduct, a different pyrimidine dimer. The 6–4 dimers are normally removed in most organisms by efficient **excision repair** (Chapter 27). However, a 6–4 photolyase was discovered in both *Arabidopsis* and *Drosophila* and has also been found in *Xenopus* and the rattlesnake.[655,655a] It has a structure similar to that of the *E. coli* photolyase and presumably acts by a related mechanism[191,656–657a] that uses the light-excited reduced flavin in an electron donation and return cycle as in Fig. 23-50. A homolog of the *Drosophila* 6–4 photolyase gene has been found in human cells, but there is uncertainty about its function.[658] Is it really a photolyase or is it a cryptochrome involved in the circadian cycle?

J. Bioluminescence

The emission of visible light by living beings is one of the most fascinating of natural phenomena. Luminescent bacteria, glowing toadstools, protozoa that can light up ocean waves, luminous clams, fantastically illuminated railroad worms,[659] and fireflies[660–661a] have all been the objects of the biochemists' curiosity.[662–664] The chemical problem is an interesting one. The firefly's light with a wavelength of 560 nm (17,900 cm^{-1}) has an energy of 214 kJ/einstein. What kind of chemical reaction can lead to an energy yield that high? It is far too great to be provided by the splitting of ATP. Even the oxidation of NADH by oxygen would barely provide the necessary energy.

A clue comes from the fact that chemiluminescence is very common when O_2 is used as an oxidant in nonenzymatic processes. The slow oxidation of alcohols, aldehydes, and many nitrogen compounds (Eqs. 23-43, 23-44) is accompanied by emission of light



Luminol, a synthetic luminescent compound

(23-43)

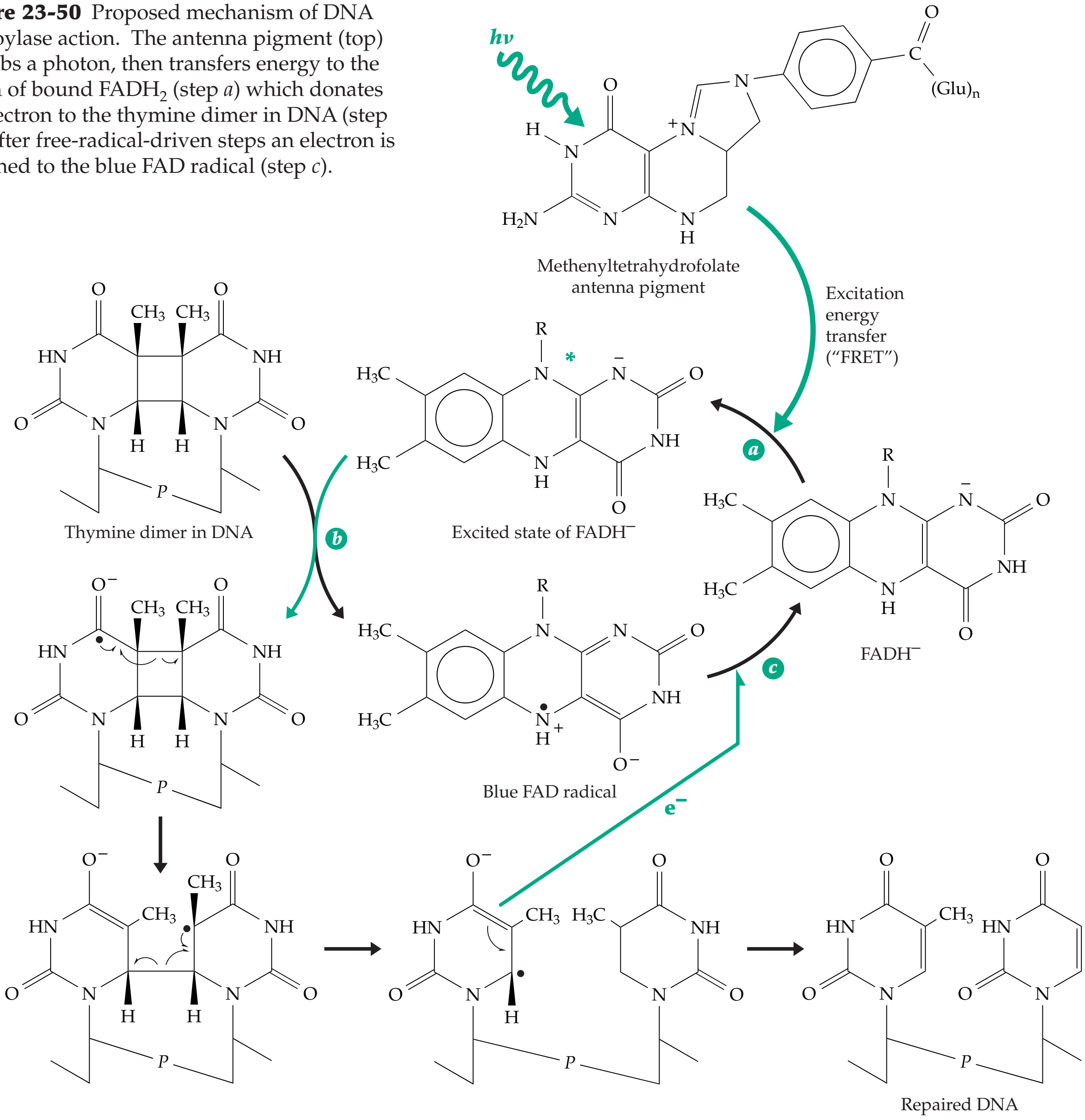
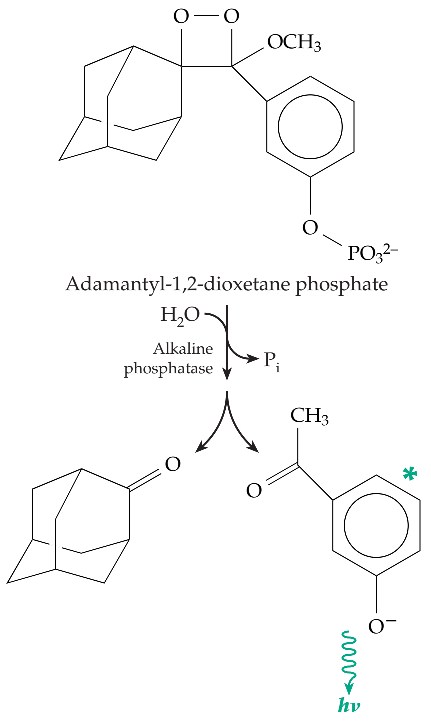


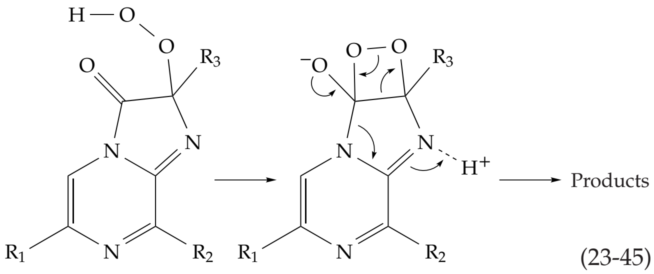
Figure 23-50 Proposed mechanism of DNA photoylase action. The antenna pigment (top) absorbs a photon, then transfers energy to the anion of bound $FADH_2$ (step *a*) which donates an electron to the thymine dimer in DNA (step *b*). After free-radical-driven steps an electron is returned to the blue FAD radical (step *c*).

visible to the eye. Chemiluminescence is especially pronounced in those reactions that are thought to occur by radical mechanisms. The recombination of free radicals provides enough energy to permit the release of visible light. Cleavage of a peroxide linkage, e.g., in a dioxetane (Eq. 23-44),[665] is often involved.[666] For example, the reaction of Eq. 23-44 is used in a sensitive light-detected assay for alkaline phosphatase.

In view of these facts it is perhaps not so surprising that many organisms have mastered the ability to channel the energy released in an oxygenation reaction into light emission. Attempts to extract luminous materials from organisms date from the last century when the French physiologist, DuBois, in 1887 prepared both a cold-water extract and a hot-water extract of luminous clams.[662] He showed that the material in the cold-water extract, which he named **luciferase**, caused emission of light when a heat-stable material (which he called **luciferin**) present in hot-water extract was added. These names have been retained and are now used in a general way. Thus, the luciferins are a family of compounds whose structures have been determined for a number of species (Fig. 23-51). Firefly luciferin is a carboxylic acid, but it must be activated in an ATP-requiring reaction to give **luciferyl adenylate**, whose structure is shown in the figure. The latter emits light in the presence of O_2 and luciferase. It can be seen that the original carboxyl group becomes CO_2, while the ring becomes oxidized. In addition, the acyl adenylate linkage is broken. In



(23-44)



(23-45)

the “sea pansy” *Renilla reniformis* (a coelenterate) the luciferin has quite a different structure.[667,668] However, the reaction with O_2 to produce CO_2 and an oxidized product causes the light emission, just as in the firefly. The luciferin of *Renilla*, which is called **coelenterazine** (Fig. 23-51), is also found in the jellyfish *Aequorea*, the shrimp *Oplaphorus*,[669] the “firefly squid” *Watasenia scintillans*,[670] and other luminous organisms.

In *Renilla* the coelenterazine is stored as a coelenterazine sulfate, possibly having the structure shown. To convert this storage form to the active luciferin the sulfo group is transferred onto adenosine 3',5'-bisphosphate to form 3'-phosphoadenosine 5'-phosphosulfate, the reverse of step *d* of Eq. 17-38. The luciferin of the ostracod crustacean *Vargula hilgendorfii* has a structure (Fig. 23-51) close to that from *Renilla*. In *Vargula* (formerly *Cypridina*) the luciferin and luciferase are produced in separate glands and are secreted into the surrounding water where they mix and produce a bright cloud of light.[671]

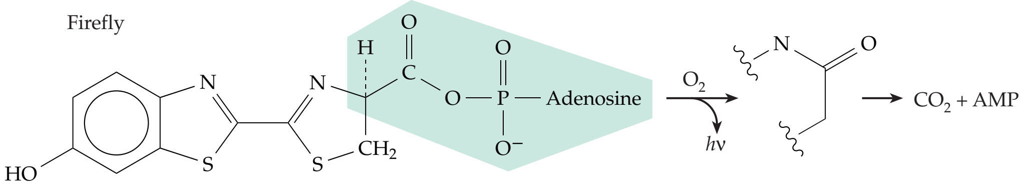
In most mechanisms suggested for luciferase action O_2 reacts at the carbon atom that becomes the carbonyl group in the product to form an intermediate peroxide. In the case of *Renilla* luciferin this can easily be visualized as a result of flow of electrons (perhaps one at a time) from the pyrazine nitrogen (at the bottom of the structure in Fig. 23-51) into the O_2. According to one proposal, the peroxide group that is formed adds to the carbonyl to form a four-membered dioxetane ring as shown in Eq. 23-45 for coelenterazine peroxide. The latter opens, as indicated by the arrows, to give the products. This theory was tested using $^{18}O_2$. In the case of *Vargula* luciferin the expected incorporation of one atom of ^{18}O into CO_2 was observed, but with firefly and *Renilla* luciferins no ^{18}O entered the CO_2. In these two cases, a somewhat different mechanism may hold.

The jellyfish *Aequorea* contains a **photoprotein**, which emits light only when calcium ions are present.[672,673] Since light emission can be measured with great sensitivity (modern photomultipliers can be used to count light quanta) the protein **aequorin** and related photoproteins[674a] are used as a sensitive indicator of calcium ion concentration.[674] (In a similar way the firefly luciferin–luciferase system, which requires ATP for activation, is widely used in an assay for ATP.)

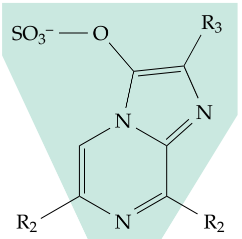
To identify the chromophore in aequorin over 4000 kg of jellyfish were used to obtain 125 mg of electrophoretically pure photoprotein.[674] From this one mg of a chromophoric substance AF-350 (Fig. 23-51) was isolated and characterized as a product. The close relationship to the *Renilla* and *Vargula* luciferins is obvious, and it is thought that coelenterazine is present in aequorin and other photoproteins as a peroxide (as in Eq. 23-45). For this reason no additional oxygen is needed to complete the reaction when Ca^{2+} acts to alter the conformation of the protein.[675–676a] The structure of a photoactive intermediate from the coelenterazine-containing protein obelin (from *Obelia longissima*), however, shows only one oxygen atom attached to C2 (Fig. 23-51)[676a] Although coelenterazine is utilized by many cnidarians they apparently cannot synthesize the compound but must obtain it through their diet. The source of biosynthesis is unknown.[676b] Some dinoflagellates emit light from a 137-kDa luciferin that contains three homologous domains each of which binds a molecule of a tetrapyrrole.[676c] From its structure the latter appears to have arisen from chlorophyll (Fig. 23-20), whose ring has been opened to give a structure somewhat similar to that of phytochrome (Fig. 23-23).[676d]

The first step in the formation of light in the firefly is a reaction with ATP to form luciferyl adenylate (Eq. 23-46, step *a*).[676e] The proton on the carbon may then be removed making use of the electron accepting properties of the adjacent ring system and carbonyl group before addition of the O_2. The reactions should be compared to those catalyzed by oxygenases, e.g., Eq. 18-42. The large 62-kDa firefly luciferase has a

Firefly

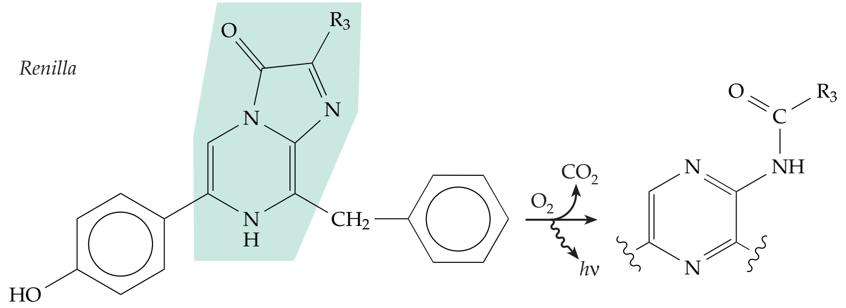

Adenosine
O_2
hν
CO_2 + AMP

Luciferyl adenylate

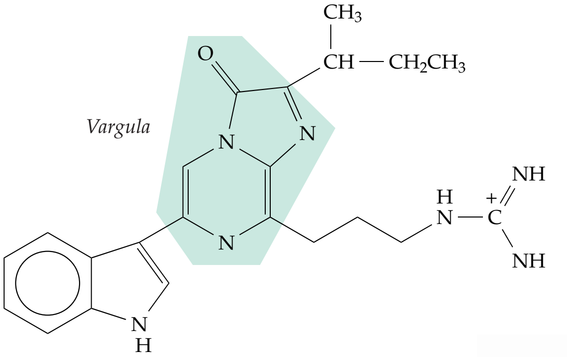


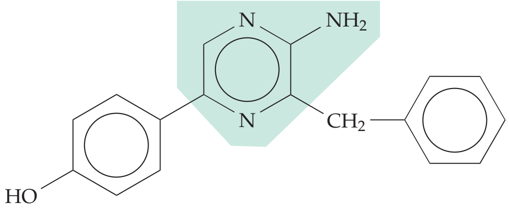
Luciferyl sulfate of *Renilla*

Renilla

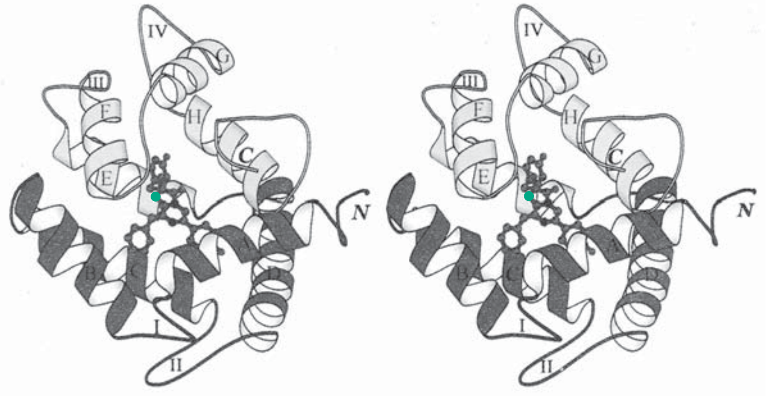


Vargula





Compound AF-350 from *Aequorea*


IV
G
III
F
H
C
E
N
B
C
A
D
I
II

Photoprotein **obelin** from *obelia* showing an oxy complex of bound coelenterazina. The oxygen atom is green. From Liu *et al.*[676a]

Latia

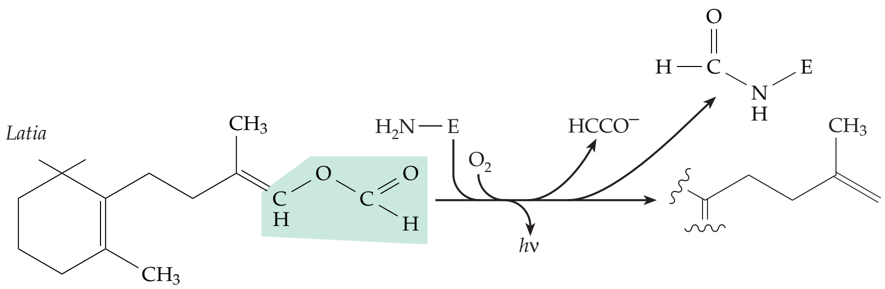
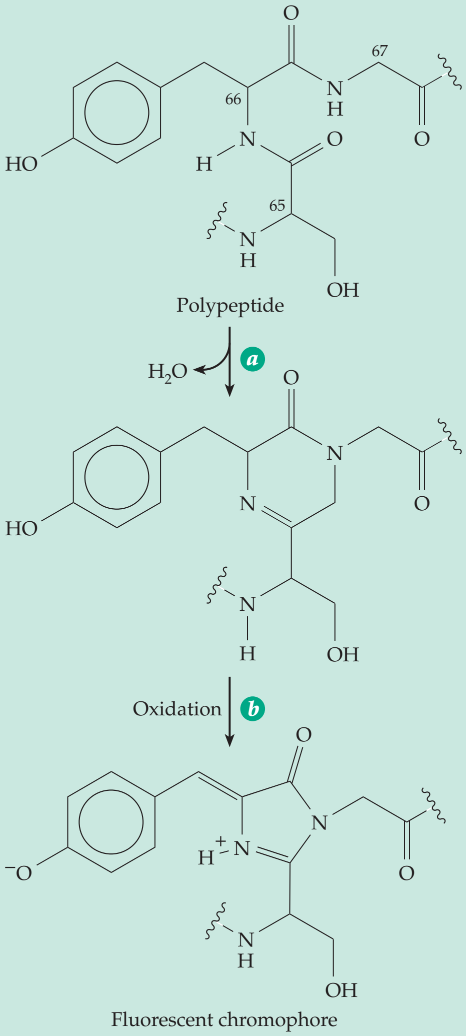

$H_2N—E$
O_2
hν
$HCCO^-$

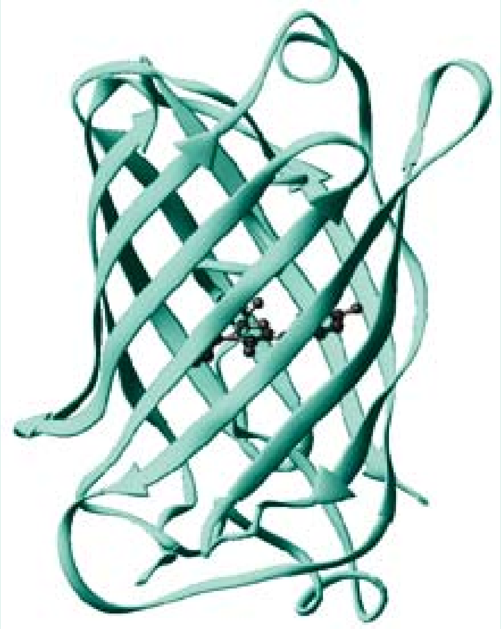
Figure 23-51 Structures of luciferins from several luminous organisms. The forms shown are the "activated" molecules ready to react with O_2. However, compound AF-350 is a breakdown product of the Ca^{2+}-activated luminous protein aequorin.

BOX 23-A THE GREEN FLUORESCENT PROTEIN AND OTHER LIGHT-EMITTING ANTENNAS

The Ca^{2+} -dependent luminescence of the jellyfish *Aquorea*, discussed in the main text, consists of blue light with a maximum intensity at 470 nm. However, the living organism has a more brilliant green luminescence. The excitation energy is transferred in a nonradiationless process to a green fluorescent protein with absorption maxima at 395 and 475 nm and an emission maximum for fluorescence at 508 nm.[a–c] A similar protein is used by *Renilla*.[d] The 238-residue green fluorescent protein (**GFP**) has a compact three-dimensional structure, an 11-stranded antiparallel β barrel with helices at one end and longer loops at the other, a "**β can**." The chromophore lies in the center of the cylinder. Numerous mutant forms have been made but only 15 residues in the terminal regions could be deleted without loss of fluorescence.[e]

The chromophore of GFP is formed spontaneously from -Ser-Tyr-Gly, residues 65–67 of the protein.[a,f,g] The entire protein has been synthesized chemically and forms the fluorescent chromophore just as in the protein produced biologically.[h] The reaction is autocatalytic, requiring only O_2 as the oxidant. Although in the living organism it accepts energy from the aqueorin chromophore a near ultraviolet lamp will elicit the fluorescence in the laboratory.





Ribbon drawing of the 238-residue green fluorescent protein showing the embedded chromophore as a ball-and-stick structure.[i] Courtesy of S. James Remington.

The green fluorescent protein is used widely in molecular biology as a fluorescent tag. Its rugged chemical nature, resistance to degradation by proteases and ability to form the chromophore autocatalytically from its own amino acids have permitted many applications. The entire GFP can be attached covalently to numerous cell components. Its gene can be spliced into the genome of an organism to form green-glowing worms, flies, and plants. Put behind a suitable promoter the fluorescent protein may be synthesized or not depending upon the control mechanism of a particular promoter (Chapter 28).[c,j–m]

The phenolic group of the GFP chromophore is apparently dissociated in the form absorbing at 395 nm and is in a tautomeric equilibrium with the other species. However, some histidine-containing replacement mutants have pH-dependent spectral changes in which the dipolar ionic form shown above, and absorbing at a longer wavelength, loses

BOX 23-A (continued)

a proton to form the anion. Observation of the excitation spectra for fluorescence of such mutant proteins within cells provides a new method for measuring the internal pH of cells and their organelles.[n–q] Some mutants emit blue or yellow light.[r,s] Two different color mutants have been fused with a molecule of calmodulin in such a way that the Ca^{2+}-induced conformational change in calmodulin allows fluorescence resonance energy transfer (FRET) between the two fluorophores. This provides a new type of calcium ion indicator.[s–u] A structurally similar red fluorescent protein, produced by a coral, extends the range of colors available as biological markers and may be useful in applications based on resonance energy transfer.[v,w,x]

Bioluminescent bacteria of the genus *Photobacterium* produce large amounts of a highly fluorescent 189-residue **lumazine protein** which contains bound 6,7-dimethyl-8-ribityllumazine (see Fig. 25-20).[y,z] Like the green fluorescent protein, it serves as a secondary light emitter receiving its energy by transfer from the flavin primary emitter. Its presence shifts the light-emission from the 495 nm of the luciferase to as low as 470 nm. *Vibrio fischeri* synthesizes a **yellow fluorescent protein** with either bound FMN or riboflavin. Its emission is at 542 nm, a longer wavelength than that of the luciferase emission. The value to the bacteria may be the higher quantum yield of fluorescence from the antenna emitter than from the luciferase. The luciferase fluorescence has a lifetime of 10 ns but on addition of the yellow fluorescent protein it is decreased to 0.25 ns with a greatly intensified emission.[u]

---

[a] Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A., and Tsien, R. Y. (1995) *Trends Biochem. Sci.* **20**, 448–455

[b] Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormö, M., and Remington, S. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2306–2311

[c] Chalfie, M., and Kain, S., eds. (1998) *Green Fluorescent Protein: Properties, Applications and Protocols*, Wiley-Liss, New York

[d] Hart, R. C., Matthews, J. C., Hori, K., and Cormier, M. J. (1979) *Biochemistry* **18**, 2204–2205

[e] Li, X., Zhang, G., Ngo, N., Zhao, X., Kain, S. R., and Huang, C.-C. (1997) *J. Biol. Chem.* **272**, 28545–28549

[f] Niwa, H., Inouye, S., Hirano, T., Matsuno, T., Kojima, S., Kubota, M., Ohashi, M., and Tsuji, F. I. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13617–13622

[g] Reid, B. G., and Flynn, G. C. (1997) *Biochemistry* **36**, 6786–6791

[h] Nishiuchi, Y., Inui, T., Nishio, H., Bódi, J., Kimura, T., Tsuji, F. I., and Sakakibara, S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13549–13554

[i] Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996) *Science* **273**, 1392–1395

[j] Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) *Science* **263**, 802–805

[k] Ohashi, T., Kiehart, D. P., and Erickson, H. P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2153–2158

[l] Yeh, E., Gustafson, K., and Boulianne, G. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7036–7040

[m] Hampton, R. Y., Koning, A., Wright, R., and Rine, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 828–833

[n] Elsliger, M.-A., Wachter, R. M., Hanson, G. T., Kallio, K., and Remington, S. J. (1999) *Biochemistry* **38**, 5296–5301

[o] Robey, R. B., Ruiz, O., Santos, A. V. P., Ma, J., Kear, F., Wang, L.-J., Li, C.-J., Bernardo, A. A., and Arruda, J. A. L. (1998) *Biochemistry* **37**, 9894–9901

[p] Tsien, R. Y. (1998) *Ann. Rev. Biochem.* **67**, 509–544

[q] Wachter, R. M., Yarbrough, D., Kallio, K., and Remington, S. J. (2000) *J. Mol. Biol.* **301**, 157–171

[r] Yang, T.-T., Sinai, P., Green, G., Kitts, P. A., Chen, Y.-T., Lybarger, L., Chervenak, R., Patterson, G. H., Piston, D. W., and Kain, S. R. (1998) *J. Biol. Chem.* **273**, 8212–8216

[s] Pozzan, T. (1997) *Nature (London)* **388**, 834–835

[t] Romoser, V. A., Hinkle, P. M., and Persechini, A. (1997) *J. Biol. Chem.* **272**, 13270–13274

[u] Nagai, T., Sawano, A., Park, E. S., and Miyawaki, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3197–3202

[v] Yarbrough, D., Wachter, R. M., Kallio, K., Matz, M. V., and Remington, S. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 462–467

[w] Garcia-Parajo, M. F., Koopman, M., van Dijk, E. M. H. P., Subramaniam, V., and van Hulst, N. F. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14392–14397

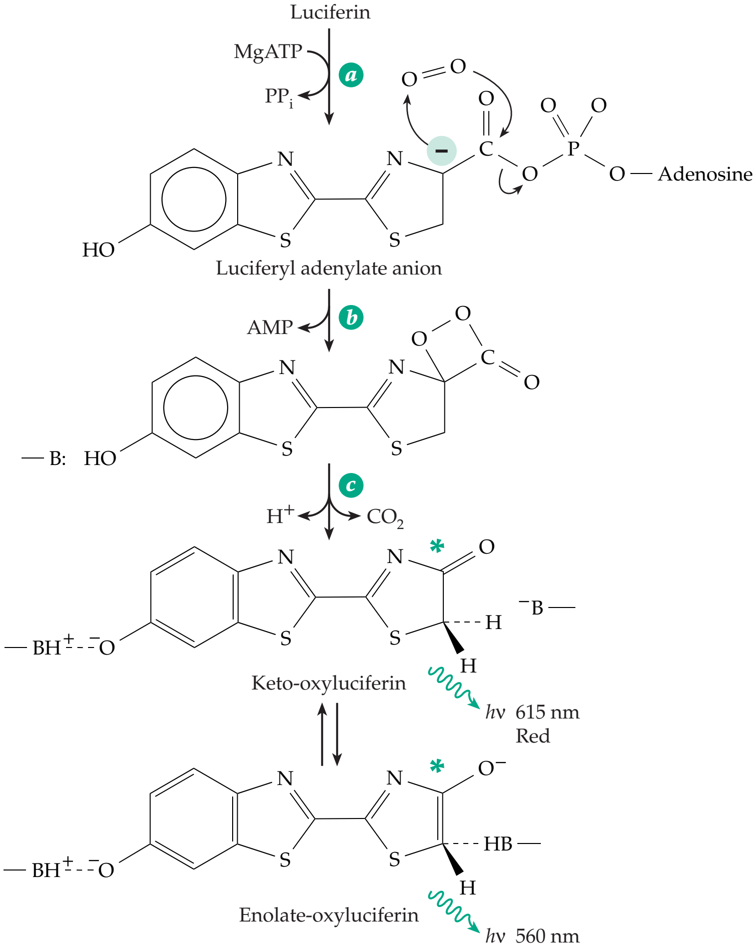
[x] Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7877–7882

[y] Hughes, R. E., Brzovic, P. S., Klevit, R. E., and Hurley, J. B. (1995) *Biochemistry* **34**, 11410–11416

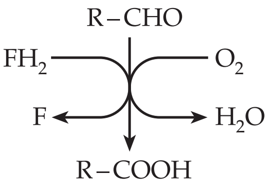
[z] Petushkov, V. N., Gibson, B. G., and Lee, J. (1996) *Biochemistry* **35**, 8413–8418

two-domain structure that suggests that domain movement may be essential to bring reactants together.[660] The structure is homologous to those of acyl-CoA ligases and peptide synthetases which share a similarity in step *a*. Formation of the dioxetane intermediate is assisted by the loss of AMP (Eq. 23-46, step *b*). The electronically excited decarboxylation product interacts with groups in the protein. It apparently exists as an anion bound to acidic and basic groups of the protein. An equilibrium between oxo- and enolate forms is thought to regulate the color of the emitted light which can vary from red to yellow and green in various fireflies, other beetles, and larvae.[661,661a,677,677a] Oxyluciferin can be reconverted to luciferin for the next flash.[677b]

A very different light-producing reaction is used by the limpet *Latia*. The luciferin is an unusual terpene derivative (Fig. 23-51) that lacks any chromophore suitable for light emission.[678] Evidently oxidation of this luciferin causes electronic excitation of some other molecule, presumably a “purple protein” which is also needed for luminescence. A complex of luciferin plus the purple protein is believed to react with the luciferase (abbreviated E-NH$_2$ in Fig. 23-51). It is


Luciferin
MgATP
PP$_i$
a
O=O
O
O
O
P
C
O
O — Adenosine
N
N
S
S
HO
Luciferyl adenylate anion
AMP
b
O
O
C
O
N
N
S
S
— B: HO
c
H$^+$
CO$_2$
N
N
*
O
S
S
H
$^-$B —
— BH$^+$- - $^-$O
H
Keto-oxyluciferin
hν 615 nm
Red
N
N
*
O$^-$
S
S
HB —
— BH$^+$- - $^-$O
H
Enolate-oxyluciferin
hν 560 nm

(23-46)


R–CHO
FH$_2$
O$_2$
F
H$_2$O
R–COOH

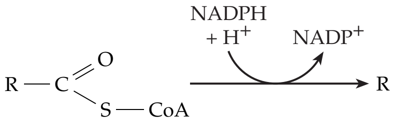
(23-47)

thought that the formyl group is released from its enolic ester linkage in the luciferin. A Schiff base of the resulting aldehyde may form with the enzyme and react with oxygen. Yet another type of luciferin is found in dinoflagellates (Fig. 23-51).[679]

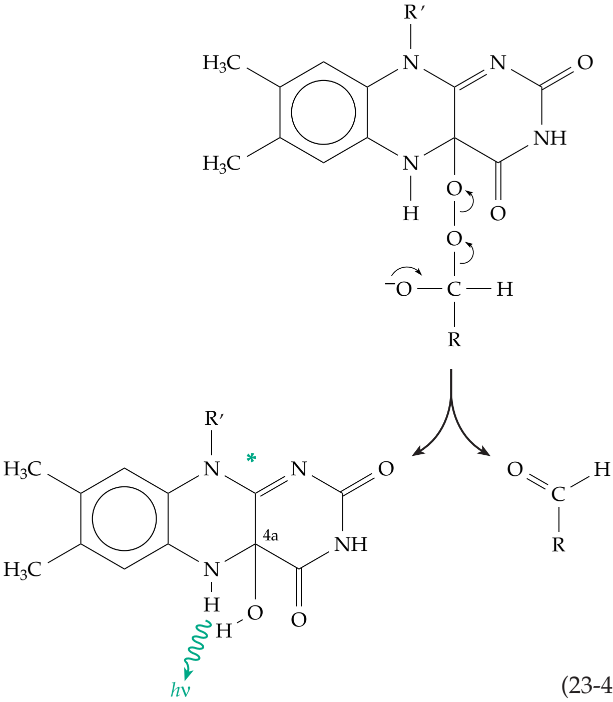
Luminescent bacteria all appear to obtain light from a riboflavin-5'-phosphate dependent oxygenase, which converts a long-chain aldehyde (usually *n*-tetradecanal) to a carboxylic acid (Eq. 23-47). Here FH_2 is the riboflavin 5'-*P*, which is thought to be supplied by a flavin reductase.[679a]

Bacterial luciferases are αβ heterodimers with subunit masses of ~40(α) and 35 (β) kDa.[664,680,681] In *Vibrio harveyi* these are encoded by the *lux A* and *lux B* genes. At least five other genes are essential for light production including two regulatory genes.[682,683] The tetradecanal and other long-chain aldehydes are supplied by reduction of the corresponding acyl-CoA (Eq. 23-48). A special thioesterase releases a myristoyl group from an acyl carrier protein, diverting it for luminescence in *V. harveyi*.[684] There is good evidence from ^{13}C NMR and electronic spectra for an enzyme-bound reduced flavin hydroperoxide as in Eq. 15-31. While this hydroperoxide can decompose slowly to flavin and H_2O_2 in the dark, it can also carry out the oxidation of the aldehyde with emission of light.[685,685a] The luminescent emission spectrum resembles the fluorescence spectrum of the 4a –OH adduct (Eq. 23-49), which is probably the light-emitting species.[686–688]

Cells of *Vibrio fischeri*, from the light organ of the fish *Monocentrus japonicus*, emit light only in dense cultures where a chemical inducer identified as N-(2-oxocaproyl)homoserine lactone[689,690] accumulates.


NADPH
+ H$^+$
NADP$^+$
R — C
O
S — CoA
R — C
O
H

(23-48)


R′
H$_3$C
H$_3$C
N
N
O
NH
N
H
O
O
O
$^-$O — C — H
R
R′
*
H$_3$C
H$_3$C
N
N
O
4a
NH
N
H
H
O
O
hν
O
C
H
R

(23-49)

References

1. Wald, G. (1959) *Sci. Am.* **201**(Oct), 92 – 108
2. Calvert, J. G., and Pitts, J. N., Jr. (1966) *Photochemistry*, Wiley, New York
3. Suzuki, H. (1967) *Electronic Absorption Spectra and Geometry of Organic Molecules*, Academic Press, New York
4. Kraemer, K. H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11 – 14
5. Kohen, E., Santus, R., and Hirschberg, J. G. (1995) *Photobiology*, Academic Press, San Diego
6. Murrell, J. N. (1967) *The Theory of the Electronic Spectra of Organic Molecules*, Academic Press, New York
7. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry Part II*, Freeman, San Francisco, California (pp. 349 – 480)
8. Herzberg, G. (1950) *Molecular Spectra and Molecular Structure*, 2nd ed., Vol. I, Van Nostrand-Reinhold, Princeton, New Jersey
9. Brown, S. B., ed. (1980) *An Introduction to Spectroscopy for Biochemists*, Academic Press, New York
10. Bell, J. E., ed. (1980, 1981) *Spectroscopy in Biochemistry*, CRC Press, Boca Raton, Florida
11. Campbell, I. D., and Dwek, R. A. (1983) *Biological Spectroscopy*, Addison-Wesley, Reading, Massachusetts

11a. Steinmeyer, G., Sutter, D. H., Gallmann, L., Matuschek, N., and Keller, U. (1999) *Science* **286**, 1507 – 1512

11b. Hopkins, J.-M., and Sibbett, W. (2000) *Sci. Am.* **283**(Sep), 72 – 79

11c. Service, R. F. (2001) *Science* **292**, 1627 – 1628

11d. Ihee, H., Lobastov, V. A., Gomez, U. M., Goodson, B. M., Srinivasan, R., Ruan, C.-Y., and Zewail, A. H. (2001) *Science* **291**, 458 – 462

11e. Lattman, E. E. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6535 – 6536

11f. Drescher, M., Hentschel, M., Kienberger, R., Tempea, G., Spielmann, C., Reider, G. A., Corkum, P. B., and Krausz, F. (2001) *Science* **291**, 1923 – 1927

11g. Bhattacharjee, Y. (2001) *Nature (London)* **412**, 474 – 476

12. Kochakian, C. D. (1988) *Trends Biochem. Sci.* **13**, 359 – 362

12a. Burns, D. A., and Ciurczak, E. W., eds. (2001) *Handbook of Near-Infrared Analysis*, 2nd ed., Dekker, New York

13. Parker, F. S. (1983) *Applications of Infrared, Raman and Resonance Raman Spectroscopy in Biochemistry, Biology and Medicine*, Plenum, New York
14. Haris, P. I., and Chapman, D. (1992) *Trends Biochem. Sci.* **17**, 328 – 333
15. Surewicz, W. K., Mantsch, H. H., and Chapman, D. (1993) *Biochemistry* **32**, 389 – 394

15a. Gremlich, H.-U., and Yan, B., eds. (2000) *Infrared and Raman Spectroscopy of Biological Materials*, Dekker, New York

16. Reisdorf, W. C., Jr., and Krimm, S. (1996) *Biochemistry* **35**, 1383 – 1386
17. Wright, W. W., Laberge, M., and Vanderkooi, J. M. (1997) *Biochemistry* **36**, 14724 – 14732
18. Dagneaux, C., Liquier, J., and Taillandier, E. (1995) *Biochemistry* **34**, 16618 – 16623
19. Sanchez-Ruiz, J. M., and Martinez-Carrion, M. (1988) *Biochemistry* **27**, 3338 – 3342

19a. Zhao, W., and Wright, J. C. (1999) *J. Am. Chem. Soc.* **121**, 10994 – 10998

20. Zscherp, C., Schlesinger, R., Tittor, J., Oesterhelt, D., and Heberle, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5498 – 5503
21. Meskers, S., Ruysschaert, J.-M., and Goormaghtigh, E. (1999) *J. Am. Chem. Soc.* **121**, 5115 – 5122
22. Fraser, R. D. B., and MacRae, T. P. (1973) *Conformation in Fibrous Proteins and Related Synthetic Polypeptides*, Academic Press, New York (pp. 95 – 106)
23. Grasselli, J. G., Snavely, M. K., and Balkin, B. J. (1981) *Chemical Applications of Raman Spectroscopy*, Wiley, New York

23a. Baumruk, V., Pancoska, P., and Keiderling, T. A. (1996) *J. Mol. Biol.* **259**, 774 – 791

24. Burke, M. J., and Rougvie, M. A. (1972) *Biochemistry* **11**, 2435 – 2439
25. Anderson, T. S., Hellgeth, J., and Lansbury, P. T., Jr. (1996) *J. Am. Chem. Soc.* **118**, 6540 – 6546
26. Reilly, K. E., and Thomas, G. J., Jr. (1994) *J. Mol. Biol.* **241**, 68 – 82
27. de Jongh, H. H. J., Goormaghtigh, E., and Ruysschaert, J.-M. (1997) *Biochemistry* **36**, 13603 – 13610
28. Miles, H. T., Lewis, T. P., Becker, E. D., and Frazier, J. (1973) *J. Biol. Chem.* **248**, 1115 – 1117
29. Hendra, P., Jones, C., and Warnes, G. (1991) *Fourier Transform Raman Spectroscopy*, Ellis Horwood, New York
30. Overman, S. A., and Thomas, G. J., Jr. (1998) *Biochemistry* **37**, 5654 – 5665
31. Lippert, J. L., Tyminski, D., and Desmeules, P. J. (1976) *J. Am. Chem. Soc.* **98**, 7075 – 7080
32. Craig, W. S., and Gaber, B. P. (1977) *J. Am. Chem. Soc.* **99**, 4130 – 4134
33. Tsuboi, M., Overman, S. A., and Thomas, G. J., Jr. (1996) *Biochemistry* **35**, 10403 – 10410

33a. Wen, Z. Q., Overman, S. A., Bondre, P., and Thomas, G. J., Jr. (2001) *Biochemistry* **40**, 449 – 458

34. Emerson, R., Chalmers, R., and Cederstrand, C. (1957) *Proc. Natl. Acad. Sci. U.S.A.* **43**, 133 – 143
35. Emerson, R., and Arnold, W. (1932) *J. Gen. Physiol.* **16**, 191 – 205
36. Carey, P. R. (1982) *Biochemical Applications of Raman and Resonance Raman Spectroscopies*, Academic Press, New York
37. Lugtenburg, J., Mathies, R. A., Griffin, R. G., and Herzfeld, J. (1988) *Trends Biochem. Sci.* **13**, 388 – 393
38. Wright, J. C. (1982) in *Applications of Lasers to Chemical Problems* (Evans, T. R., ed), pp. 35 – 179, Wiley, New York
39. Chen, X. G., Li, P., Holtz, J. S. W., Chi, Z., Pajcini, V., Asher, S. A., and Kelly, L. A. (1996) *J. Am. Chem. Soc.* **118**, 9705 – 9715
40. Chi, Z., Chen, X. G., Holtz, J. S. W., and Asher, S. A. (1998) *Biochemistry* **37**, 2854 – 2864
41. Overman, S. A., and Thomas, G. J., Jr. (1999) *Biochemistry* **38**, 4018 – 4027
42. Hu, X., and Spiro, T. G. (1997) *Biochemistry* **36**, 15701 – 15712
43. Miura, T., and Thomas, G. J., Jr. (1994) *Biochemistry* **33**, 7848 – 7856
44. Miura, T., and Thomas, G. J., Jr. (1995) *Biochemistry* **34**, 9645 – 9654
45. Chan, S. S., Austin, R. H., Mukerji, I., and Spiro, T. G. (1997) *Biophys. J.* **72**, 1512 – 1520
46. Hu, X., Rodgers, K. R., Mukerji, I., and Spiro, T. G. (1999) *Biochemistry* **38**, 3462 – 3467
47. Chi, Z., and Asher, S. A. (1999) *Biochemistry* **38**, 8196 – 8203
48. Macdonald, I. D. G., Sligar, S. G., Christian, J. F., Unno, M., and Champion, P. M. (1999) *J. Am. Chem. Soc.* **121**, 376 – 380
49. Dong, S., and Spiro, T. G. (1998) *J. Am. Chem. Soc.* **120**, 10434 – 10440
50. Kitagawa, T., Nishina, Y., Kyogoku, Y., Yamano, T., Ohishi, N., Takai-Suzuki, A., and Yagi, K. (1979) *Biochemistry* **18**, 1804 – 1808
51. Kitagawa, T., Nishina, Y., Kyogoky, Y., Yamano, T., Ohishi, N., Takai-Suzuki, A., and Yagi, K. (1979) *Biochemistry* **18**, 1804 – 1808
52. Clarkson, J., Palfey, B. A., and Carey, P. R. (1997) *Biochemistry* **36**, 12560 – 12566
53. Benecky, M. J., Copeland, R. A., Hays, T. R., Lobenstine, E. W., Rava, R. P., Pascal, R. A., Jr., and Spiro, T. G. (1985) *J. Biol. Chem.* **260**, 11663 – 11670
54. Benecky, M. J., Copeland, R. A., Rava, R. P., Feldhaus, R., Scott, R. D., Metzler, C. M., Metzler, D. E., and Spiro, T. G. (1985) *J. Biol. Chem.* **260**, 11671 – 11678

54a. Altose, M. D., Zheng, Y., Dong, J., Palfey, B. A., and Carey, P. R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3006 – 3011

55. Clark, R. J. H., D'Urso, N. R., and Zagalsky, P. F. (1980) *J. Am. Chem. Soc.* **102**, 6693 – 6698
56. Ozaki, Y., Pliura, D. H., Carey, P. R., and Storer, A. C. (1982) *Biochemistry* **21**, 3102 – 3108

56a. Vogel, K. M., Kozlowski, P. M., Zgierski, M. Z., and Spiro, T. G. (1999) *J. Am. Chem. Soc.* **121**, 9915 – 9921

56b. Franzen, S. (2001) *J. Am. Chem. Soc.* **123**, 12578 – 12589

57. Rospendowski, B. N., Kelly, K., Wolf, C. R., and Smith, W. E. (1991) *J. Am. Chem. Soc.* **113**, 1217 – 1225
58. Dähne, S. (1978) *Science* **199**, 1163 – 1167
59. Morley, J. O., Morley, R. M., and Fitton, A. L. (1998) *J. Am. Chem. Soc.* **120**, 11479 – 11488
60. Marder, S. R., Gorman, C. B., Meyers, F., Perry, J. W., Bourhill, G., Brédas, J.-L., and Pierce, B. M. (1994) *Science* **265**, 632 – 635
61. Horwitz, J., Strickland, E. H., and Billups, C. (1969) *J. Am. Chem. Soc.* **91**, 184 – 190
62. Siano, D. B., and Metzler, D. E. (1969) *J. Chem. Phys.* **51**, 1856 – 1861
63. Metzler, D. E., Harris, C. M., Johnson, R. J., Siano, D. B., and Thomson, J. A. (1973) *Biochemistry* **12**, 5377 – 5392
64. Metzler, C. M., Cahill, A. E., Petty, S., Metzler, D. E., and Lang, L. (1985) *Appl. Spectroscopy* **39**, 333 – 339
65. Horwitz, J., Strickland, E. H., and Billups, C. (1970) *J. Am. Chem. Soc.* **92**, 2119 – 2129
66. Horwitz, J., and Strickland, E. H. (1971) *J. Biol. Chem.* **246**, 3749 – 3752
67. Zucchelli, G., Dainese, P., Jennings, R. C., Breton, J., Garlaschi, F. M., and Bassi, R. (1994) *Biochemistry* **33**, 8982 – 8990
68. Tang, S. W., Coleman, J. E., and Myer, Y. P. (1968) *J. Biol. Chem.* **243**, 4286 – 4297
69. Hofrichter, J., and Eaton, W. A. (1976) *Annu Rev Biophys Bioeng.* **5**, 511 – 560
70. Arnone, A., Christen, P., Jansonias, J. N., and Metzler, D. E. (1985) in *Transaminases* (Christen, P., and Metzler, D. E., eds), pp. 349 – 357, Wiley, New York
71. Holmén, A., Broo, A., Albinsson, B., and Nordén, B. (1997) *J. Am. Chem. Soc.* **119**, 12240 – 12250
72. Peterson, J., Pearce, L. L., and Bominaar, E. L. (1999) *J. Am. Chem. Soc.* **121**, 5972 – 5980
73. Metzler, C. M., Mitra, J., Metzler, D. E., Makinen, M. W., Hyde, C. C., Rogers, P., and Arnone, A. (1988) *J. Mol. Biol.* **203**, 197 – 220
74. Morton, R. A. (1975) *Biochemical Spectroscopy*, Wiley, New York
75. Perkampus, H. H., Sandeman, I., and Timmons, C. J. (1966–1977) *UV Atlas of Organic Compounds*, Vol. 1 – 5, Plenum, New York
76. Lang, L., ed. (1961–1975) *Absorption Spectra in the Ultraviolet and Visible Region*, Academic Press, New York (A serial publication)
77. Petruska, J. (1961) *J. Chem. Phys.* **34**, 1120 – 1136
78. Stevenson, P. E. (1965) *J. Mol. Spectroscopy* **15**, 220 – 256
79. Strickland, E. H., Wilchek, M., Billups, C., and Horowitz, J. (1972) *J. Biol. Chem.* **247**, 572 – 580
80. Strickland, E. H., Billups, C., and Kay, E. (1972) *Biochemistry* **11**, 3657 – 3662

References

81. Pajcini, V., Chen, X. G., Bormett, R. W., Geib, S. J., Li, P., Asher, S. A., and Lidiak, E. G. (1996) *J. Am. Chem. Soc.* **118**, 9716–9726
82. Clark, L. B. (1995) *J. Am. Chem. Soc.* **117**, 7974–7986
83. Fülscher, M. P., and Roos, B. O. (1995) *J. Am. Chem. Soc.* **117**, 2089–2095
84. Fülscher, M. P., Serrano-Andrés, L., and Roos, B. O. (1997) *J. Am. Chem. Soc.* **119**, 6168–6176
85. Harders, H., Forster, S., Voelter, W., and Bacher, A. (1974) *Biochemistry* **13**, 3360–3364
86. Blackburn, G. M., and Gait, M. J., eds. (1996) *Nucleic Acids in Chemistry and Biology*, 2nd ed., Oxford Univ. Press, Oxford
87. Murrell, J. N. (1963) *The Theory of the Electronic Spectra of Organic Molecules*, Wiley, New York (Chapter 7)
88. Collins, K. D., and Stark, G. R. (1971) *J. Biol. Chem.* **246**, 6599–6605
89. Donovan, J. W. (1973) in *Methods in Enzymology*, Vol. 27, Part D (Hirs, C. H. W., and Timasheff, S. N., eds), pp. 497–525, Academic Press, New York
90. Metzler, D. E., Harris, C., Yang, I.-Y., Siano, D., and Thomson, J. A. (1972) *Biochem. Biophys. Res. Commun.* **46**, 1588–1597
91. Metzler, D. E., Metzler, C. M., and Mitra, J. (1986) *Trends Biochem. Sci.* **11**, 157–159
92. Kornblatt, J. A., Kornblatt, M. J., and Hoa, G. H. B. (1995) *Biochemistry* **34**, 1218–1223
93. Padrós, E., Dunach, M., Morros, A., Sobés, M., and Manosa, J. (1984) *Trends Biochem. Sci.* **9**, 508–510
94. Butler, W. L. (1979) *Methods Enzymol.* **56**, 501–515
95. Fasman, G. D., ed. (1996) *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum, New York
96. Foss, J. G. (1963) *J. Chem. Educ.* **40**, 592–597
97. Johnson, W. C., Jr. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 145–166
98. Strickland, E. H. (1974) *Crit. Revs. Biochem.* **2**, 113–175
99. Moffitt, W., Woodward, R. B., Moscowitz, A., Klyne, W., and Djerassi, C. (1961) *J. Am. Chem. Soc.* **83**, 4013–4018
100. Bayley, P. M. (1973) *Prog. Biophys. and Mol. Biol.* **27**, 1–76
101. Johnson, W. C., Jr., and Tinoco, I., Jr. (1972) *J. Am. Chem. Soc.* **94**, 4389–4390
102. Chen, Y.-H., Yang, J. T., and Chau, K. H. (1974) *Biochemistry* **13**, 3350–3359
103. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry, Part II*, Freeman, San Francisco, California (pp. 409–433)
104. Tinoco, I., Jr., Maestre, M. F., and Bustamante, C. (1983) *Trends Biochem. Sci.* **8**, 41–44
105. Applequist, J. (1973) *J. Am. Chem. Soc.* **95**, 8255–8262
106. Applequist, J. (1987) *Am. Scientist* **75**, 58–68
107. Applequist, J., and Bode, K. A. (1999) *J. Phy. Chem. B* **103**, 1767–1773
108. Bode, K. A., and Applequist, J. (1998) *J. Am. Chem. Soc.* **120**, 10938–10946 , Erratum 13545
109. Bode, K. A., and Applequist, J. (1997) *Biopolymers* **42**, 855–860
110. Paterlini, M. G., Freedman, T. B., and Nafie, L. A. (1986) *J. Am. Chem. Soc.* **108**, 1389–1397
111. Bose, P. K., and Polavarapu, P. L. (1999) *J. Am. Chem. Soc.* **121**, 6094–6095
112. Baumruk, V., and Keiderling, T. A. (1993) *J. Am. Chem. Soc.* **115**, 6939–6942
113. Self, B. D., and Moore, D. S. (1997) *Biophys. J.* **73**, 339–347
114. Bell, A. F., Hecht, L., and Barron, L. D. (1997) *J. Am. Chem. Soc.* **119**, 6006–6013
115. Peters, K. S., Watson, T., and Marr, K. (1991) *Ann. Rev. Biophys. Biophys. Chem.* **20**, 343–362
116. Hung, R. R., and Grabowski, J. J. (1999) *J. Am. Chem. Soc.* **121**, 1359–1364
117. Meisel, A., Leonhardt, G., and Szargan, R. (1989) *X-Ray Spectra and Chemical Binding*, Springer-Verlag, New York
118. Binsted, N., Strange, R. W., and Hasnain, S. S. (1992) *Biochemistry* **31**, 12117–12125
119. Ito, E., Oji, H., Araki, T., Oichi, K., Ishii, H., Ouchi, Y., Ohta, T., Kosugi, N., Maruyama, Y., Naito, T., Inabe, T., and Seki, K. (1997) *J. Am. Chem. Soc.* **119**, 6336–6344
120. Vértes, A., Korecz, L., and Burger, K. (1979) *Mössbauer Spectroscopy*, Elsevier, Amsterdam
121. Popescu, C. V., Bates, D. M., Beinert, H., Münck, E., and Kiley, P. J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13431–13435
122. Schünemann, V., Trautwein, A. X., Illerhaus, J., and Haehnel, W. (1999) *Biochemistry* **38**, 8981–8991
123. Clayton, R. K. (1970) *Light and Living Matter*, Vol. 1, McGraw-Hill, New York
124. Parker, C. A. (1968) *Photoluminescence of Solutions*, Elsevier, Amsterdam
125. Konev, S. V. (1967) *Fluorescence and Phosphorescence of Proteins and Nucleic Acids*, Plenum, New York
126. Lakowicz, J. R. (1985) *Principles of Fluorescence Spectroscopy*, Plenum, New York
127. Baeyens, W. R. G., De Keukeleire, D., and Korkidis, K., eds. (1991) *Luminescence Techniques in Chemical and Biochemical Analysis*, Dekker, New York
127a. Meyer-Almes, F.-J., and Auer, M. (2000) *Biochemistry* **39**, 13261–13268
128. Kotaki, A., and Yagi, K. (1970) *J. Biochem.* **68**, 509–516
129. Koziol, J., and Knobloch, E. (1965) *Biochim. Biophys. Acta.* **102**, 289–300
130. Chen, Y., and Barkley, M. D. (1998) *Biochemistry* **37**, 9976–9982
131. Beechem, J. M., and Brand, L. (1985) *Ann. Rev. Biochem.* **54**, 43–71
132. Anderson, S. R. (1991) *J. Biol. Chem.* **266**, 11405–11408
133. Hedstrom, J., Sedarous, S., and Prendergast, F. G. (1988) *Biochemistry* **27**, 6203–6208
134. Wahl, P., Auchet, J.-C., Visser, A. J. W. G., and Veeger, C. (1975) *Eur. J. Biochem.* **50**, 413–418
135. Vos, R., Engelborghs, Y., Izard, J., and Baty, D. (1995) *Biochemistry* **34**, 1734–1743
136. Broos, J., ter Veld, F., and Robillard, G. T. (1999) *Biochemistry* **38**, 9798–9803
137. Ross, J. B. A., Senear, D. F., Waxman, E., Kombo, B. B., Rusinova, E., Huang, Y. T., Laws, W. R., and Hasselbacher, C. A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 12023–12027
138. Farah, C. S., and Reinach, F. C. (1999) *Biochemistry* **38**, 10543–10551
139. Strambini, G. B., and Gonnelli, M. (1995) *J. Am. Chem. Soc.* **117**, 7646–7651
140. Zhou, J. S., and Hoffman, B. M. (1994) *Science* **265**, 1693–1696
141. Eftink, M. R., and Ghiron, C. A. (1976) *Biochemistry* **15**, 672–680
142. Calhoun, D. B., Vanderkooi, J. M., and Englander, S. W. (1983) *Biochemistry* **22**, 1533–1539
143. Wright, W. W., Owen, C. S., and Vanderkooi, J. M. (1992) *Biochemistry* **31**, 6538–6544
144. Vet, J. A. M., Majithia, A. R., Marras, S. A. E., Tyagi, S., Dube, S., Poiesz, B. J., and Kramer, F. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6394–6399
145. Hilinski, E. F., and Rentzepis, P. M. (1983) *Nature (London)* **302**, 481–487
146. Radda, G. K. (1971) *Curr. Top. Bioenerg.* **4**, 81–176
147. Lakowicz, J. R., Laczko, G., Gryczynski, I., and Cherek, H. (1986) *J. Biol. Chem.* **261**, 2240–2245
148. Ruggiero, A. J., Todd, D. C., and Fleming, G. R. (1990) *J. Am. Chem. Soc.* **112**, 1003–1014
149. Förster, T. (1948) *Ann Physik* **2**, 55–75
150. Stryer, L. (1968) *Science* **162**, 526–533
151. Wu, C.-W., and Stryer, L. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1104–1108
152. Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R., and Weiss, S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6264–6268
153. Weiss, S. (1999) *Science* **283**, 1676–1683
154. Ha, T., Zhuang, X., Kim, H. D., Orr, J. W., Williamson, J. R., and Chu, S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9077–9082
155. Fu, P. K.-L., and Turro, C. (1999) *J. Am. Chem. Soc.* **121**, 1–7
156. Hogue, C. W. V., MacManus, J. P., Banville, D., and Szabo, A. G. (1992) *J. Biol. Chem.* **267**, 13340–13347
157. Xiao, M., Li, H., Snyder, G. E., Cooke, R., Yount, R. G., and Selvin, P. R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15309–15314
158. Gordon, G. W., Berry, G., Liang, X. H., Levine, B., and Herman, B. (1998) *Biophys. J.* **74**, 2702–2713
159. Tsien, R. Y., and Miyawaki, A. (1998) *Science* **280**, 1954–1955
160. Wouters, F. S., Bastiaens, P. I. H., Wirtz, K. W. A., and Jovin, T. M. (1998) *EMBO J.* **17**, 7179–7189
161. Ward, L. D., Seckler, R., and Timasheff, S. N. (1994) *Biochemistry* **33**, 11900–11908
162. Farrar, S. J., Whiting, P. J., Bonnert, T. P., and McKernan, R. M. (1999) *J. Biol. Chem.* **274**, 10100–10104
163. Parkhurst, K. M., and Parkhurst, L. J. (1995) *Biochemistry* **34**, 293–300
164. Jares-Erijman, E. A., and Jovin, T. M. (1996) *J. Mol. Biol.* **257**, 597–617
164a. Norman, D. G., Grainger, R. J., Uhrín, D., and Lilley, D. M. J. (2000) *Biochemistry* **39**, 6317–6324
164b. Tóth, K., Brun, N., and Langowski, J. (2001) *Biochemistry* **40**, 6921–6928
165. Rhee, M.-J., Sudnick, D. R., Arkle, V. K., and Horrocks, W. DeW., Jr. (1981) *Biochemistry* **20**, 3328–3334
166. Szczesniak, M., Szczepaniak, K., Kwiatkowski, J. S., KuBulat, K., and Person, W. B. (1988) *J. Am. Chem. Soc.* **110**, 8319–8330
167. Vanderkooi, J. M., Kaposi, A., and Fidy, J. (1993) *Trends Biochem. Sci.* **18**, 71–76
168. Friedrich, J., Gafert, J., Zollfrank, J., Vanderkooi, J., and Fidy, J. (1994) *Proc. atl. Acad. Sci. U.S.A.* **91**, 1029–1033
168a. Bastiaens, P. I. H., and Pepperkok, R. (2000) *Trends Biochem. Sci.* **25**, 631–637
168b. Frohn, J. T., Knapp, H. F., and Stemmer, A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7232–7236
168c. Klar, T. A., Jakobs, S., Dyba, M., Egner, A., and Hell, S. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8206–8210
168d. Weiss, S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8747–8749
168e. Farber, S. A., Pack, M., Ho, S.-Y., Johnson, I. D., Wagner, D. S., Dosch, R., Mullins, M. C., Hendrickson, H. S., Hendrickson, E. K., and Halpern, M. E. (2001) *Science* **292**, 1385–1388
169. Ruth, J. L. (1984) *DNA* **3**, 123
170. Renz, M., and Kurz, C. (1984) *Nucleic Acids Res.* **12**, 3435–3444
171. Davenport, L., Dale, R. E., Bisby, R. H., and Cundall, R. B. (1985) *Biochemistry* **24**, 4097–4108
172. MacGregor, R. B., and Weber, G. (1986) *Nature (London)* **319**, 70–73

References

173. Tsien, R. Y., and Poenie, M. (1986) *Trends Biochem. Sci.* **11**, 450–455
174. Thomas, F., Serratrice, G., Béguin, C., Saint Aman, E., Pierre, J. L., Fontecave, M., and Laulhère, J. P. (1999) *J. Biol. Chem.* **274**, 13375–13383
175. Stewart, W. W. (1981) *Nature (London)* **292**, 17–21
176. Berland, K. M. (1997) *Biophys. J.* **72**, 1487–1488
177. Schwille, P., Meyer-Almes, F.-J., and Rigler, R. (1997) *Biophys. J.* **72**, 1878–1886
178. Korlach, J., Schwille, P., Webb, W. W., and Feigenson, G. W. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8461–8466
179. Marx, J. (1996) *Science* **273**, 430
180. Schröck, E., du Manoir, S., Veldman, T., Schoell, B., Wienberg, J., Ferguson-Smith, M. A., Ning, Y., Ledbetter, D. H., Bar-Am, I., Soenksen, D., Garini, Y., and Ried, T. (1996) *Science* **273**, 494–497

180a. Pathak, S., Cjoi, S.-K., Arnheim, N., and Thompson, M. E. (2001) *J. Am. Chem. Soc.* **123**, 4103–4104

181. Bruchez, M., Jr., Moronne, M., Gin, P., Weiss, S., and Alivisatos, A. P. (1998) *Science* **281**, 2013–2016
182. Chan, W. C. W., and Nie, S. (1998) *Science* **281**, 2016–2018
183. Turro, N. J., and Schuster, G. (1975) *Science* **187**, 303–312
184. Salem, L. (1976) *Science* **191**, 822–830
185. Bazhulina, N. P., Morozov, Y. V., Karpeisky, M. Y., Ivanov, V. I., and Kuklin, A. I. (1966) *Biofizika* **11**, 42–47
186. Bridges, J. W., Davies, D. S., and Williams, R. T. (1966) *Biochem. J.* **98**, 451–468
187. Johnson, G. F., Tu, J.-I., Bartlett, M. L. S., and Graves, D. J. (1970) *J. Biol. Chem.* **245**, 5560–5568
188. Shaltiel, S., and Cortijo, M. (1970) *Biochem. Biophys. Res. Commun.* **41**, 594–600
189. Loken, M. R., Hayes, J. W., Gohlke, J. R., and Brand, L. (1972) *Biochemistry* **11**, 4779–4786
190. Becker, M. M., and Wang, J. C. (1984) *Nature (London)* **309**, 682–687
191. Sancar, A. (1996) *Science* **272**, 48–49

191a. Lee, J.-H., Bae, S.-H., and Choi, B.-S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4591–4596

192. Shaw, A. A., Falick, A. M., and Shetlar, M. D. (1992) *Biochemistry* **31**, 10976–10983
193. Pashev, I. G., Dimitrov, S. I., and Angelov, D. (1991) *Trends Biochem. Sci.* **16**, 323–326
194. Pearlman, D. A., Holbrook, S. R., Pirkle, D. H., and Kim, S.-H. (1985) *Science* **227**, 1304–1305
195. Cimino, G. D., Gamper, H. B., Isaacs, S. T., and Hearst, J. E. (1985) *Ann. Rev. Biochem.* **54**, 1151–1193

195a. Yoon, J.-H., Lee, C.-S., O'Connor, T. R., Yasui, A., and Pfeifer, G. P. (2000) *J. Mol. Biol.* **299**, 681–693

196. Hanson, K. M., Li, B., and Simon, J. D. (1997) *J. Am. Chem. Soc.* **119**, 2715–2721
197. Edelson, R. L. (1988) *Sci. Am.* **259**(Aug), 68–75
198. Berns, M. W. (1991) *Sci. Am.* **264**(Jun), 84–90
199. Amato, I. (1993) *Science* **262**, 32–33
200. Staros, J. V. (1980) *Trends Biochem. Sci.* **5**, 320–322
201. Jo, E., Blazyk, J., and Boggs, J. M. (1998) *Biochemistry* **37**, 13791–13799
202. Gronemeyer, H. (1985) *Trends Biochem. Sci.* **10**, 264–267
203. Dormán, G., and Prestwich, G. D. (1994) *Biochemistry* **33**, 5661–5673
204. Gribbon, P., and Hardingham, T. E. (1998) *Biophys. J.* **75**, 1032–1039
205. Peters, R. (1985) *Trends Biochem. Sci.* **10**, 223–227
206. Pagano, R. E., and Longmuir, K. J. (1983) *Trends Biochem. Sci.* **8**, 157–161
207. Anfinrud, P., de Vivie-Riedle, R., and Engel, V. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8328–8329
208. Goldbeck, R. A., Paquette, S. J., Björling, S. C., and Kliger, D. S. (1996) *Biochemistry* **35**, 8628–8639
209. Service, R. F. (1997) *Science* **276**, 1986–1987
210. Brunori, M., Cutruzzola, F., Savino, C., Travaglini-Allocatelli, C., Vallone, B., and Gibson, Q. H. (1999) *Trends Biochem. Sci.* **24**, 253–255
211. Wan, C., Fiebig, T., Kelley, S. O., Treadway, C. R., Barton, J. K., and Zewail, A. H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6014–6019
212. Henderson, P. T., Jones, D., Hampikian, G., Kan, Y., and Schuster, G. B. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8353–8358
213. Block, S. M. (1992) *Nature (London)* **360**, 493–495
214. Svoboda, K., and Block, S. M. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 247–285
215. Wang, M. D., Yin, H., Landick, R., Gelles, J., and Block, S. M. (1997) *Biophys. J.* **72**, 1335–1346
216. Arai, Y., Yasuda, R., Akashi, K.-i, Harada, Y., Miyata, H., Kinosita, K., Jr., and Itoh, H. (1999) *Nature (London)* **399**, 446–448
217. Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T., and Solas, D. (1991) *Science* **251**, 767–773
218. McGall, G., Labadie, J., Brock, P., Wallraff, G., Nguyen, T., and Hinsberg, W. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13555–13560
219. Taubes, G. (1997) *Science* **276**, 1991–1993
220. Benaron, D. A., Cheong, W.-F., and Stevenson, D. K. (1997) *Science* **276**, 2002–2003
221. Bappart, S. A., Tearney, G. J., Bouma, B. E., Southern, J. F., Brezinski, M. E., and Fujimoto, J. G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4256–4261

221a. Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sorensen, J., Baldock, R., and Davidson, D. (2002) *Science* **296**, 541–545

222. Jamin, N., Dumas, P., Moncuit, J., Fridman, W.-H., Teillaud, J.-L., Carr, G. L., and Williams, G. P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4837–4840
223. Wetzel, D. L., and LeVine, S. M. (1999) *Science* **285**, 1224–1225
224. Foyer, C. H. (1984) *Photosynthesis*, Wiley, New York
225. Clayton, R. K. (1981) *Photosynthesis Physical Mechanisms and Chemical Patterns*, Cambridge Univ. Press, London
226. Gregory, R. P. F. (1989) *Biochemistry of Photosynthesis*, 3rd ed., Wiley, New York
227. Hall, D. O., and Rao, K. K. (1994) *Photosynthesis*, 5th ed., Cambridge Univ. Press, New York
228. Barber, J., and Andersson, B. (1994) *Nature (London)* **370**, 31–34
229. Kamminga, H. (1981) *Trends Biochem. Sci.* **6**, 164–165
230. van Niel, C. B. (1931) *Adv. Enzymol.* **1**, 263–328
231. Hill, R. (1937) *Nature (London)* **139**, 881–882
232. Gaffron, H. (1960) in *Plant Physiology*, Vol. 1B (Steward, F. C., ed), pp. 176–180, Academic Press, New York
233. Prince, R. C. (1996) *Trends Biochem. Sci.* **21**, 121–122
234. Arnon, D. I., Tsujimoto, H. Y., and McSwai, B. D. (1965) *Nature (London)* **207**, 1357–1372
235. Arnon, D. I. (1984) *Trends Biochem. Sci.* **9**, 258–262
236. Zito, F., Finazzi, G., Joliot, P., and Wollman, F.-A. (1998) *Biochemistry* **37**, 10395–10403

236a. Finazzi, G., Zito, F., Barbagallo, R. P., and Wollman, F.-A. (2001) *J. Biol. Chem.* **276**, 9770–9774

237. Haley, J., and Bogorad, L. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1534–1538

237a. Roberts, A. G., and Kramer, D. M. (2001) *Biochemistry* **40**, 13407–13412

237b. Deniau, C., and Rappaport, F. (2000) *Biochemistry* **39**, 3304–3310

237c. Sainz, G., Carrell, C. J., Ponamarev, M. V., Soriano, G. M., Cramer, W. A., and Smith, J. L. (2000) *Biochemistry* **39**, 9164–9173

237d. Fernández-Velasco, J. G., Jamshidi, A., Gong, X.-S., Zhou, J., and Ueng, R. Y. (2001) *J. Biol. Chem.* **276**, 30598–30607

237e. Bunney, T. D., van Walraven, H. S., and de Boer, A. H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4249–4254

238. Peterman, E. J. G., Wenk, S.-O., Pullerits, T., Pålsson, L.-O., van Grondelle, R., Dekker, J. P., Rögner, M., and van Amerongen, H. (1998) *Biophys. J.* **75**, 389–398
239. Carrell, C. J., Schlarb, B. G., Bendall, D. S., Howe, C. J., Cramer, W. A., and Smith, J. L. (1999) *Biochemistry* **38**, 9590–9599
240. Xue, Y., Ökvist, M., Hansson, Ö., and Young, S. (1998) *Protein Sci.* **7**, 2099–2105
241. Schnackenberg, J., Than, M. E., Mann, K., Wiegand, G., Huber, R., and Reuter, W. (1999) *J. Mol. Biol.* **290**, 1019–1030
242. De la Cerda, B., Díaz-Quintana, A., Navarro, J. A., Hervás, M., and De la Rosa, M. A. (1999) *J. Biol. Chem.* **274**, 13292–13297

242a. Molina-Heredia, F. P., Hervás, M., Navarro, J. A., and De la Rosa, M. A. (2001) *J. Biol. Chem.* **276**, 601–605

242b. Baymann, F., Rappaport, F., Joliot, P., and Kallas, T. (2001) *Biochemistry* **40**, 10570–10577

243. Kuras, R., Guergova-Kuras, M., and Crofts, A. R. (1998) *Biochemistry* **37**, 16280–16288
244. Zhang, H., Huang, D., and Cramer, W. A. (1999) *J. Biol. Chem.* **274**, 1581–1587
245. Oh-oka, H., Iwaki, M., and Itoh, S. (1998) *Biochemistry* **37**, 12293–12300
246. Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-i, Inokuchi, H., and Ozeki, H. (1986) *Nature (London)* **322**, 572–574
247. Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986) *EMBO J.* **5**, 2043–2049
248. Maier, R. M., Neckermann, K., Igloi, G. L., and Kössel, H. (1995) *J. Mol. Biol.* **251**, 614–628
249. Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S., Tsudzuki, J., Nakashima, K., Tsudzuki, T., Suzuki, Y., Hamada, A., Ohta, T., Inamura, A., Yoshinaga, K., and Sugiura, M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5967–5972

249a. Cline, K. (2000) *Nature (London)* **403**, 148–149

250. Zhang, Z., Green, B. R., and Cavalier-Smith, T. (1999) *Nature (London)* **400**, 155–159
251. Kouranov, A., and Schnell, D. J. (1996) *J. Biol. Chem.* **271**, 31009–31012
252. Asai, T., Shinoda, Y., Nohara, T., Yoshihisa, T., and Endo, T. (1999) *J. Biol. Chem.* **274**, 20075–20078
253. Bölter, B., Soll, J., Schulz, A., Hinnah, S., and Wagner, R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15831–15836

References

253a. Dabney-Smith, C., van den Wijngaard, P. W. J., Treece, Y., Vredenberg, W. J., and Bruce, B. D. (1999) *J. Biol. Chem.* **274**, 32351–32359
254. Douce, R., and Joyard, J. (1981) *Trends Biochem. Sci.* **6**, 237–240
255. Hoober, J. K. (1984) *Chloroplasts*, Plenum, New York
255a. von Wettstein, D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3633–3635
256. Muhlethaler, K. (1971) in *Structure and Function of Chloroplasts* (Gibbs, M., ed), pp. 7–34, Springer-Verlag, Berlin and New York
257. Miller, K. R. (2000) *Sci. Am.*, 102–113
258. Seibert, M., DeWit, M., and Staehelin, L. A. (1987) *J. Cell Biol.* **105**, 2257–2265
259. Anderson, J. M., and Anderson, B. (1982) *Trends Biochem. Sci.* **7**, 288–292
260. Rebeiz, C. A., and Lascelles, J. (1982) in *Photosynthesis*, Vol. I (Govindjee, ed), pp. 699–780, Academic Press, New York
261. Miyashita, H., Ikemoto, H., Kurano, N., Adachi, K., Chihara, M., and Miyachi, S. (1996) *Nature (London)* **383**, 402
262. Caple, M. B., Chow, H., and Strouse, C. E. (1978) *J. Biol. Chem.* **253**, 6730–6737
263. Smith, K. M., Kehres, L. A., and Tabba, H. D. (1980) *J. Am. Chem. Soc.* **102**, 7149–7151
264. French, C. S., and Brown, J. S. (1972) in *Photosynthesis. Two Centuries after Its Discovery by Joseph Priestly*, 2nd ed. (Forti, G., Avron, M., and Melandri, A., eds), pp. 291–306, Junk, The Hague
265. Govindjee, R. (1974) *Sci. Am.* **231**(Dec), 68–82
266. Wolken, J. J. (1975) *Photoprocesses, Photoreceptors, and Evolution*, Academic Press, New York
267. Song, P.-S. (1978) *Trends Biochem. Sci.* **3**, 25–27
268. Hofmann, E., Wrench, P. M., Sharples, F. P., Hiller, R. G., Welte, W., and Diederichs, K. (1996) *Science* **272**, 1788–1791
269. Goodwin, T. W., ed. (1971) *Aspects of Terpenoid Chemistry and Biochemistry*, Academic Press, New York (pp. 346–348)
270. Schirmer, W., Bode, R., Sidler, W., and Zuber, H. (1985) *J. Mol. Biol.* **184**, 257–277
271. Szalontai, B., Gombos, Z., Csizmadia, V., Bagyinka, C., and Lutz, M. (1994) *Biochemistry* **33**, 11823–11832
272. Glazer, A. N. (1989) *J. Biol. Chem.* **264**, 1–4
273. Jung, L. J., Chan, F. C., and Glazer, A. N. (1995) *J. Biol. Chem.* **270**, 12877–12884
274. Killilea, S. D., O'Carra, P., and Murphy, R. F. (1980) *Biochem. J.* **187**, 311–320
275. Bishop, J. E., Nagy, J. O., O'Connell, J. F., and Rapoport, H. (1991) *Chemical Society Special Publication, No. 2* **113**, 8024–8035
276. Fairchild, C. D., and Glazer, A. N. (1994) *J. Biol. Chem.* **269**, 28988–28996
277. MacColl, R., Guard-Friar, D., and Ryan, T. J. (1990) *Biochemistry* **29**, 430–435
278. Wilk, K. E., Harrop, S. J., Jankova, L., Edler, D., Keenan, G., Sharples, F., Hiller, R. G., and Curmi, P. M. G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8901–8906
279. Chang, W.-r, Jiang, T., Wan, Z.-l, Zhang, J.-p, Yang, Z.-x, and Liang, D.-c. (1996) *J. Mol. Biol.* **262**, 721–731
280. Liu, J.-Y., Jiang, T., Zhang, J.-P., and Liang, D.-C. (1999) *J. Biol. Chem.* **274**, 16945–16952
281. Reuter, W., Wiegand, G., Huber, R., and Than, M. E. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1363–1368
282. Rhie, G., and Beale, S. I. (1992) *J. Biol. Chem.* **267**, 16088–16093
283. Fairchild, C. D., and Glazer, A. N. (1994) *J. Biol. Chem.* **269**, 8686–8694
284. Terry, M. J., McDowell, M. T., and Lagarias, J. C. (1995) *J. Biol. Chem.* **270**, 11111–11118
285. Riley, C. T., Barbeau, B. K., Keim, P. S., Kézdy, F. J., Heinrikson, R. L., and Law, J. H. (1984) *J. Biol. Chem.* **259**, 13159–13165
286. Glazer, A. N. (1983) *Ann. Rev. Biochem.* **52**, 125–157
286a. Grossman, A. R., Bhaya, D., and He, Q. (2001) *J. Biol. Chem.* **276**, 11449–11452
287. Thornber, P. J., and Maxwell, J. P. (1981) *Trends Biochem. Sci.* **6**, 122–124
288. Deisenhofer, J., Michel, H., and Huber, R. (1985) *Trends Biochem. Sci.* **10**, 243–248
289. Brejc, K., Ficner, R., Huber, R., and Steinbacher, S. (1995) *J. Mol. Biol.* **249**, 424–440
290. Lao, K., and Glazer, A. N. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5258–5263
291. Hu, X., and Schulten, K. (1998) *Biophys. J.* **75**, 683–694
292. Karrasch, S., Bullough, P. A., and Ghosh, R. (1995) *EMBO J.* **14**, 631–638
293. Walz, T., Jamieson, S. J., Bowers, C. M., Bullough, P. A., and Hunter, C. N. (1998) *J. Mol. Biol.* **282**, 833–845
294. McDermott, G., Prince, S. M., Freer, A. A., Hawthornthwaite-Lawless, A. M., Papiz, M. Z., Cogdell, R. J., and Isaacs, N. W. (1995) *Nature (London)* **374**, 517–521
295. Prince, S. M., Papiz, M. Z., Freer, A. A., McDermott, G., Hawthornthwaite-Lawless, A. M., Cogdell, R. J., and Isaacs, N. W. (1997) *J. Mol. Biol.* **268**, 412–423
296. Hu, X., Damjanovic, A., Ritz, T., and Schulten, K. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5935–5941
297. Nagarajan, V., Alden, R. G., Williams, J. C., and Parson, W. W. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13774–13779
298. Fraser, N. J., Dominy, P. J., Ücker, B., Simonin, I., Scheer, H., and Cogdell, R. J. (1999) *Biochemistry* **38**, 9684–9692
298a. McLuskey, K., Prince, S. M., Cogdell, R. J., and Isaacs, N. W. (2001) *Biochemistry* **40**, 8783–8789
299. van Oijen, A. M., Ketelaars, M., Köhler, J., Aartsma, T. J., and Schmidt, J. (1999) *Science* **285**, 400–402
300. Kühlbrandt, W. (1995) *Nature (London)* **374**, 479–498
300a. van Grondelle, R., and Novoderezhkin, V. (2001) *Biochemistry* **40**, 15057–15068
300b. Alia, Matysik, J., Soede-Huijbregts, C., Baldus, M., Raap, J., Lugtenburg, J., Gast, P., van Gorkom, H. J., Hoff, A. J., and de Groot, H. J. M. (2001) *J. Am. Chem. Soc.* **123**, 4803–4809
300c. Simonetto, R., Crimi, M., Sandoná, D., Croce, R., Cinque, G., Breton, J., and Bassi, R. (1999) *Biochemistry* **38**, 12974–12983
301. Feick, R. G., and Fuller, R. C. (1984) *Biochemistry* **23**, 3693–3700
302. Li, Y.-F., Zhou, W., Blankenship, R. E., and Allen, J. P. (1997) *J. Mol. Biol.* **271**, 456–471
302a. Vassilieva, E. V., Antonkine, M. L., Zybailov, B. L., Yang, F., Jakobs, C. U., Golbeck, J. H., and Bryant, D. A. (2001) *Biochemistry* **40**, 464–473
303. Matthews, B. W., Fenna, R. E., Bolognesi, M. C., Schmid, M. F., and Olson, J. M. (1979) *J. Mol. Biol.* **131**, 259–285
304. Rémigy, H.-W., Stahlberg, H., Fotiadis, D., Müller, S. A., Wolpensinger, B., Engel, A., Hauska, G., and Tsiotis, G. (1999) *J. Mol. Biol.* **290**, 851–858
305. Fenna, R. E., and Matthews, B. W. (1975) *Nature (London)* **258**, 573–577
306. Douady, D., Rousseau, B., and Caron, L. (1994) *Biochemistry* **33**, 3165–3170
307. Pascal, A. A., Caron, L., Rousseau, B., Lapouge, K., Duval, J.-C., and Robert, B. (1998) *Biochemistry* **37**, 2450–2457
308. Green, B. R., Pichersky, E., and Kloppstech, K. (1991) *Trends Biochem. Sci.* **16**, 181–186
309. La Roche, J., van der Staay, G. W. M., Partensky, F., Ducret, A., Aebersold, R., Li, R., Golden, S. S., Hiller, R. G., Wrench, P. M., Larkum, A. W. D., and Green, B. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15244–15248
310. Kühlbrandt, W., Wang, D. N., and Fujiyoshi, Y. (1994) *Nature (London)* **367**, 614–621
311. Kleima, F. J., Gradinaru, C. C., Calkoen, F., van Stokkum, I. H. M., van Grendelle, R., and van Amerongen, H. (1997) *Biochemistry* **36**, 15262–15268
312. Simidjiev, I., Barzda, V., Mustárdy, L., and Garab, G. (1998) *Biochemistry* **37**, 4169–4173
313. Rhee, K.-H., Morris, E. P., Zheleva, D., Hankamer, B., Kühlbrandt, W., and Barber, J. (1997) *Nature (London)* **389**, 522–526
314. Rhee, K.-H., Morris, E. P., Barber, J., and Kühlbrandt, W. (1998) *Nature (London)* **396**, 283–286
315. Barber, J., Nield, J., Morris, E. P., and Hankamer, B. (1999) *Trends Biochem. Sci.* **24**, 43–45
315a. Voigt, B., Irrgang, K.-D., Ehlert, J., Beenken, W., Renger, G., Leupold, D., and Lokstein, H. (2002) *Biochemistry* **41**, 3049–3056
316. Schubert, W.-D., Klukas, O., Saenger, W., Witt, H. T., Fromme, P., and Krauss, N. (1998) *J. Mol. Biol.* **280**, 297–314
317. Croce, R., Zucchelli, G., Garlaschi, F. M., Bassi, R., and Jennings, R. C. (1996) *Biochemistry* **35**, 8572–8579
317a. Bibby, T. S., Nield, J., Partensky, F., and Barber, J. (2001) *Nature (London)* **413**, 590
317b. Rogl, H., Schödel, R., Lokstein, H., Kühlbrandt, W., and Schubert, A. (2002) *Biochemistry* **41**, 2281–2287
318. Thomas, B. A., McMahon, L. P., and Klotz, A. V. (1995) *Biochemistry* **34**, 3758–3770
318a. Heathcote, P., Fyfe, P. K., and Jones, M. R. (2002) *Trends Biochem. Sci.* **27**, 79–87
319. Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1985) *Nature (London)* **318**, 618–624
320. Deisenhofer, J., and Michel, H. (1991) *Ann. Rev. Biophys. Biophys. Chem.* **20**, 247–266
321. Deisenhofer, J., Epp, O., Sinning, I., and Michel, H. (1995) *J. Mol. Biol.* **246**, 429–457
322. Deisenhofer, J., and Michel, H. (1989) *Science* **245**, 1463–1473
323. Lancaster, C. R. D., and Michel, H. (1999) *J. Mol. Biol.* **286**, 883–898
323a. Chen, I.-P., Mathis, P., Koepke, J., and Michel, H. (2000) *Biochemistry* **39**, 3592–3602
324. Philipson, K. D., and Sauer, K. (1972) *Biochemistry* **11**, 1880–1885
325. Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1984) *J. Mol. Biol.* **180**, 385–398
325a. McAuley, K. E., Fyfe, P. K., Ridge, J. P., Cogdell, R. J., Isaacs, N. W., and Jones, M. R. (2000) *Biochemistry* **39**, 15032–15043
326. Siefermann-Harms, D. (1985) *Biochim. Biophys. Acta.* **811**, 325–355
327. El-Kabbani, O., Chang, C.-H., Tiede, D., Norris, J., and Schiffer, M. (1991) *Biochemistry* **30**, 5361–5369
328. Govindjee. (1978) *Photochemistry and Photobiology* **28**, 935–938
329. Blankenship, R. E., and Parson, W. W. (1978) *Ann. Rev. Biochem.* **47**, 635–653
330. Dutton, P. L., Leigh, J. S., Jr., and Reed, D. W. (1972) *Biochim. Biophys. Acta.* **292**, 654–664
331. Dutton, P. L., and Mosser, C. C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10247–10250
331a. Li, J., Takahashi, E., and Gunner, M. R. (2000) *Biochemistry* **39**, 7445–7454
331b. Rabenstein, B., Ullmann, G. M., and Knapp, E.-W. (2000) *Biochemistry* **39**, 10487–10496
331c. Paddock, M. L., Ädelroth, P., Chang, C., Abresch, E. C., Feher, G., and Okamura, M. Y. (2001) *Biochemistry* **40**, 6893–6902

331d. Tandori, J., Baciou, L., Alexov, E., Maróti, P., Schiffer, M., Hanson, D. K., and Sebban, P. (2001) *J. Biol. Chem.* **276**, 45513 – 45515

331e. Gerencsér, L., and Maróti, P. (2001) *Biochemistry* **40**, 1850 – 1860

331f. Ädelroth, P., Paddock, M. L., Tehrani, A., Beatty, J. T., Feher, G., and Okamura, M. Y. (2001) *Biochemistry* **40**, 14538 – 14546

331g. Xu, Q., and Gunner, M. R. (2002) *Biochemistry* **41**, 2694 – 2701

332. Vos, M. H., Rappaport, F., Lambry, J.-C., Breton, J., and Martin, J.-L. (1993) *Nature (London)* **363**, 320 – 325

333. Hamm, P., Zurek, M., Mäntele, W., Meyer, M., Scheer, H., and Zinth, W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1826 – 1830

334. Van Brederode, M. E., Jones, M. R., Van Mourik, F., Van Stokkum, I. H. M., and Van Grondelle, R. (1997) *Biochemistry* **36**, 6855 – 6861

335. Venturoli, G., Drepper, F., Williams, J. C., Allen, J. P., Lin, X., and Mathis, P. (1998) *Biophys. J.* **74**, 3226 – 3240

335a. van Brederode, M. E., van Stokkum, I. H. M., Katilius, E., van Mourik, F., Jones, M. R., and van Grondelle, R. (1999) *Biochemistry* **38**, 7545 – 7555

336. Cherepy, N. J., Shreve, A. P., Moore, L. J., Boxer, S. G., and Mathies, R. A. (1997) *Biochemistry* **36**, 8559 – 8566

337. Cua, A., Kirmaier, C., Holten, D., and Bocian, D. F. (1998) *Biochemistry* **37**, 6394 – 6401

338. Ivancich, A., Artz, K., Williams, J. C., Allen, J. P., and Mattioli, T. A. (1998) *Biochemistry* **37**, 11812 – 11820

338a. Takahashi, E., Wells, T. A., and Wraight, C. A. (2001) *Biochemistry* **40**, 1020 – 1028

338b. Kuglstatter, A., Ermler, U., Michel, H., Baciou, L., and Fritzsch, G. (2001) *Biochemistry* **40**, 4253 – 4260

338c. Eastman, J. E., Taguchi, A. K. W., Lin, S., Jackson, J. A., and Woodbury, N. W. (2000) *Biochemistry* **39**, 14787 – 14798

339. Tang, C.-K., Williams, J. C., Taguchi, A. K. W., Allen, J. P., and Woodbury, N. W. (1999) *Biochemistry* **38**, 8794 – 8799

340. Parson, W. W., Chu, Z. T., and Warshel, A. (1998) *Biophys. J.* **74**, 182 – 191

340a. Balabin, I. A., and Onuchic, J. N. (2000) *Science* **290**, 114 – 117

340b. Yakovlev, A. G., Shkuropatov, A. Y., and Shuvalov, V. A. (2002) *Biochemistry* **41**, 2667 – 2674

341. Dohse, B., Mathis, P., Wachtveitl, J., Laussermair, E., Iwata, S., Michel, H., and Oesterhelt, D. (1995) *Biochemistry* **34**, 11335 – 11343

342. Lin, X., Williams, J. C., Allen, J. P., and Mathis, P. (1994) *Biochemistry* **33**, 13517 – 13523

343. Stowell, M. H. B., McPhillips, T. M., Rees, D. C., Soltis, S. M., Abresch, E., and Feher, G. (1997) *Science* **276**, 812 – 816

343a. de Boer, A. L., Neerken, S., de Wijn, R., Permentier, H. P., Gast, P., Vijgenboom, E., and Hoff, A. J. (2002) *Biochemistry* **41**, 3081 – 3088

343b. Schulten, E. A. M., Matysik, J., Alia, Kiihne, S., Raap, J., Lugtenburg, J., Gast, P., Hoff, A. J., and de Groot, H. J. M. (2002) *Biochemistry* **41**, 8708–8717

344. Heller, B. A., Holten, D., and Kirmaier, C. (1995) *Science* **269**, 940 – 945

344a. Lin, S., Katilius, E., Haffa, A. L. M., Taguchi, A. K. W., and Woodbury, N. W. (2001) *Biochemistry* **40**, 13767 – 13773

344b. Purton, S., Stevens, D. R., Muhiuddin, I. P., Evans, M. C. W., Carter, S., Rigby, S. E. J., and Heathcote, P. (2001) *Biochemistry* **40**, 2167 – 2175

344c. Guergova-Kuras, M., Boudreaux, B., Joliot, A., Joliot, P., and Redding, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4437 – 4442

345. Nugent, J. H. A. (1984) *Trends Biochem. Sci.* **9**, 354 – 357

345a. Johnson, T. W., Zybailov, B., Jones, A. D., Bittl, R., Zech, S., Stehlik, D., Golbeck, J. H., and Chitnis, P. R. (2001) *J. Biol. Chem.* **276**, 39512 – 39521

346. Sinning, I. (1992) *Trends Biochem. Sci.* **17**, 150 – 154

346a. Witt, H., Schlodder, E., Teutloff, C., Niklas, J., Bordignon, E., Carbonera, D., Kohler, S., Labahn, A., and Lubitz, W. (2002) *Biochemistry* **41**, 8557–8569

347. Fish, L. E., Kuck, U., and Bogorad, L. (1985) *J. Biol. Chem.* **260**, 1413 – 1421

348. Breton, J., Nabedryk, E., and Leibl, W. (1999) *Biochemistry* **38**, 11585 – 11592

349. Francke, C., Permentier, H. P., Franken, E. M., Neerken, S., and Amesz, J. (1997) *Biochemistry* **36**, 14167 – 14172

350. Nitschke, W., and Rutherford, A. W. (1991) *Trends Biochem. Sci.* **16**, 241 – 245

351. Golbeck, J. H. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1642 – 1646

352. Karrasch, S., Typke, D., Walz, T., Miller, M., Tsiotis, G., and Engel, A. (1996) *J. Mol. Biol.* **262**, 336 – 348

353. Deisenhofer, J., and Norris, J. R., eds. (1993) *The Photosynthetic Reaction Center*, Vol. 1 and 2, Academic Press, San Diego, California

354. Blankenship, R. E., and Hartman, H. (1998) *Trends Biochem. Sci.* **23**, 94 – 97

355. Kruip, J., Chitnis, P. R., Lagoutte, B., Rögner, M., and Boekema, E. J. (1997) *J. Biol. Chem.* **272**, 17061 – 17069

356. Schubert, W.-D., Klukas, O., Krauss, N., Saenger, W., Fromme, P., and Witt, H. T. (1997) *J. Mol. Biol.* **272**, 741 – 769

356a. Jordan, P., Fromme, P., Witt, H. T., Klukas, O., Saenger, W., and Krauß, N. (2001) *Nature (London)* **411**, 909 – 916

356b. Kühlbrandt, W. (2001) *Nature (London)* **411**, 896 – 899

357. Kitmitto, A., Mustafa, A. O., Holzenburg, A., and Ford, R. C. (1998) *J. Biol. Chem.* **273**, 29592 – 29599

357a. Boekema, E. J., Jensen, P. E., Schlodder, E., van Breemen, J. F. L., van Room, H., Scheller, H. V., and Dekker, J. P. (2001) *Biochemistry* **40**, 1029 – 1036

358. Morton, R. A. (1971) *Biol. Rev. Cambridge Philos. Soc.* **46**, 47 – 96

359. Threlfall, D. R., and Whistance, G. R. (1971) in *Aspects of Terpenoid Chemistry and Biochemistry* (Goodwin, T. W., ed), pp. 372 – 374, Academic Press, New York

360. Blankenship, R. E., and Prince, R. C. (1985) *Trends Biochem. Sci.* **10**, 382 – 384

361. Knox, R. S. (1969) *Biophys. J.* **9**, 1351 – 1362

362. Parson, W. W. (1978) *Photochemistry and Photobiology* **28**, 389 – 393

363. Yu, L., Zhao, J., Lu, W., Bryant, D. A., and Golbeck, J. H. (1993) *Biochemistry* **32**, 8251 – 8258

364. Pandini, V., Aliverti, A., and Zanetti, G. (1999) *Biochemistry* **38**, 10707 – 10713

365. Barth, P., Lagoutte, B., and Sétif, P. (1998) *Biochemistry* **37**, 16233 – 16241

366. Meimberg, K., Lagoutte, B., Bottin, H., and Mühlenhoff, U. (1998) *Biochemistry* **37**, 9759 – 9767

367. Bruns, C. M., and Karplus, P. A. (1995) *J. Mol. Biol.* **247**, 125 – 145

368. Medina, M., Martínez-Júlvez, M., Hurley, J. K., Tollin, G., and Gómez-Moreno, C. (1998) *Biochemistry* **37**, 2715 – 2728

369. Martínez-Júlvez, M., Hermoso, J., Hurley, J. K., Mayoral, T., Sanz-Aparicio, J., Tollin, G., Gómez-Moreno, C., and Medina, M. (1998) *Biochemistry* **37**, 17680 – 17691

370. Xiong, J., Subramaniam, S., and Govindjee. (1996) *Protein Sci.* **5**, 2054 – 2073

371. Svensson, B., Etchebest, C., Tuffery, P., van Kan, P., Smith, J., and Styring, S. (1996) *Biochemistry* **35**, 14486 – 14502

371a. Büchel, C., Morris, E., Orlova, E., and Barber, J. (2001) *J. Mol. Biol.* **312**, 371 – 379

371b. Kuhl, H., Kruip, J., Seidler, A., Krieger-Liszkay, A., Bünker, M., Bald, D., Scheidig, A. J., and Rögner, M. (2000) *J. Biol. Chem.* **275**, 20652 – 20659

371c. Shen, J.-R., and Kamiya, N. (2000) *Biochemistry* **39**, 14739 – 14744

371d. Zouni, A., Witt, H.-T., Kern, J., Fromme, P., Krauß, N., Saenger, W., and Orth, P. (2001) *Nature (London)* **409**, 739 – 743

371e. Dismukes, G. C. (2001) *Science* **292**, 447 – 448

371f. Nield, J., Kruse, O., Ruprecht, J., da Fonseca, P., Büchel, C., and Barber, J. (2000) *J. Biol. Chem.* **275**, 27940 – 27946

372. Schweitzer, R. H., and Brudvig, G. W. (1997) *Biochemistry* **36**, 11351 – 11359

373. Merry, S. A. P., Nixon, P. J., Barter, L. M. C., Schilstra, M., Porter, G., Barber, J., Durrant, J. R., and Klug, D. R. (1998) *Biochemistry* **37**, 17439 – 17447

374. Betts, S. D., Ross, J. R., Pichersky, E., and Yocum, C. F. (1997) *Biochemistry* **36**, 4047 – 4053

375. Enami, I., Kamo, M., Ohta, H., Takahashi, S., Miura, T., Kusayanagi, M., Tanabe, S., Kamei, A., Motoki, A., Hirano, M., Tomo, T., and Satoh, K. (1998) *J. Biol. Chem.* **273**, 4629 – 4634

376. Babcock, G. T. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10893 – 10895

376a. Shi, L.-X., Lorkovic, Z. J., Oelmüller, R., and Schröder, W. P. (2000) *J. Biol. Chem.* **275**, 37945 – 37950

376b. Boekema, E. J., van Breemen, J. F. L., van Roon, H., and Dekker, J. P. (2000) *J. Mol. Biol.* **301**, 1123 – 1133

376c. Mamedov, F., Stefansson, H., Albertsson, P.-Å., and Styring, S. (2000) *Biochemistry* **39**, 10478 – 10486

377. Kok, B., Forbush, B., and McGloin, M. (1970) *Photochem. and Photobiol.* **11**, 457 – 475

378. Radmer, R., and Kok, B. (1975) *Ann. Rev. Biochem.* **44**, 409 – 433

379. Murata, N., and Miyao, M. (1985) *Trends Biochem. Sci.* **10**, 122 – 124

380. Lydakis-Simantiris, N., Dorlet, P., Ghanotakis, D. F., and Babcock, G. T. (1998) *Biochemistry* **37**, 6427 – 6435

381. Schiller, H., Dittmer, J., Iuzzolino, L., Dörner, W., Meyer-Klaucke, W., Solé, V. A., Nolting, H.-F., and Dau, H. (1998) *Biochemistry* **37**, 7340 – 7350

382. Chu, H.-A., Gardner, M. T., O'Brien, J. P., and Babcock, G. T. (1999) *Biochemistry* **38**, 4533 – 4541

383. Noguchi, T., Inoue, Y., and Tang, X.-S. (1999) *Biochemistry* **38**, 10187 – 10195

384. Barry, B. A., and Babcock, G. T. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7099 – 7103

385. Diner, B. A., Force, D. A., Randall, D. W., and Britt, R. D. (1998) *Biochemistry* **37**, 17931 – 17943

386. Hays, A.-M. A., Vassiliev, I. R., Golbeck, J. H., and Debus, R. J. (1998) *Biochemistry* **37**, 11352 – 11365

387. Ahlbrink, R., Haumann, M., Cherepanov, D., Bögershausen, O., Mulkidjanian, A., and Junge, W. (1998) *Biochemistry* **37**, 1131 – 1142

388. Mamedov, F., Sayre, R. T., and Styring, S. (1998) *Biochemistry* **37**, 14245 – 14256

References

388a. Bernát, G., Morvaridi, F., Feyziyev, Y., and Styring, S. (2002) *Biochemistry* **41**, 5830–5843
389. Siegbahn, P. E. M., and Crabtree, R. H. (1999) *J. Am. Chem. Soc.* **121**, 117–127
389a. Schlodder, E., and Witt, H. T. (1999) *J. Biol. Chem.* **274**, 30387–30392
389b. Geijer, P., Morvaridi, F., and Styring, S. (2001) *Biochemistry* **40**, 10881–10891
390. Baldwin, M. J., and Pecoraro, V. L. (1996) *J. Am. Chem. Soc.* **118**, 11325–11326
391. Limburg, J., Vrettos, J. S., Liable-Sands, L. M., Rheingold, A. L., Crabtree, R. H., and Brudvig, G. W. (1999) *Science* **283**, 1524–1527
392. Hoganson, C. W., and Babcock, G. T. (1997) *Science* **277**, 1953–1956
392a. Hillier, W., and Babcock, G. T. (2001) *Biochemistry* **40**, 1503–1509
393. Wincencjusz, H., Yocum, C. F., and van Gorkom, H. J. (1999) *Biochemistry* **38**, 3719–3725
394. Haddy, A., Hatchell, J. A., Kimel, R. A., and Thomas, R. (1999) *Biochemistry* **38**, 6104–6110
395. Kühne, H., Szalai, V. A., and Brudvig, G. W. (1999) *Biochemistry* **38**, 6604–6613
396. Seidler, A., and Rutherford, A. W. (1996) *Biochemistry* **35**, 12104–12110
397. Ädelroth, P., Lindberg, K., and Andréasson, L.-E. (1995) *Biochemistry* **34**, 9021–9027
397a. Vrettos, J. S., Stone, D. A., and Brudvig, G. W. (2001) *Biochemistry* **40**, 7937–7945
398. Klimov, V. V., Hulsebosch, R. J., Allakhverdiev, S. I., Wincencjusz, H., van Gorkom, H. J., and Hoff, A. J. (1997) *Biochemistry* **36**, 16277–16281
398a. Baranov, S. V., Ananyev, G. M., Klimov, V. V., and Dismukes, G. C. (2000) *Biochemistry* **39**, 6060–6065
399. Keister, D. L., and Raveed, N. J. (1974) *J. Biol. Chem.* **249**, 6454–6458
400. Schubert, H., Kroon, B. M. A., and Matthijs, H. C. P. (1994) *J. Biol. Chem.* **269**, 7267–7272
401. Demmig-Adams, B., Gilmore, A. M., and Adams, W. W., III. (1996) *FASEB J.* **10**, 403–412
401a. Demming-Adams, B., and Adams, W. W., III. (2000) *Nature (London)* **403**, 371–374
401b. Li, X.-P., Björkman, O., Shih, C., Grossman, A. R., Rosenquist, M., Jansson, S., and Niyogi, K. K. (2000) *Nature (London)* **403**, 391–395
401c. Kagawa, T., Sakai, T., Suetsugu, N., Oikawa, K., Ishiguro, S., Kato, T., Tabata, S., Okada, K., and Wada, M. (2001) *Science* **291**, 2138–2141
402. Havaux, M., and Niyogi, K. K. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8762–8767
403. Limantara, L., Fujii, R., Zhang, J.-P., Kakuno, T., Hara, H., Kawamori, A., Yagura, T., Cogdell, R. J., and Koyama, Y. (1998) *Biochemistry* **37**, 17469–17486
404. Telfer, A., Dhami, S., Bishop, S. M., Phillips, D., and Barber, J. (1994) *Biochemistry* **33**, 14469–14474
405. Bugos, R. C., and Yamamoto, H. Y. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6320–6325
406. Niyogi, K. K., Björkman, O., and Grossman, A. R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14162–14167
407. Ruban, A. V., Lee, P. J., Wentworth, M., Young, A. J., and Horton, P. (1999) *J. Biol. Chem.* **274**, 10458–10465
407a. Frank, H. A., Bautista, J. A., Josue, J. S., and Young, A. J. (2000) *Biochemistry* **39**, 2831–2837
407b. Jahns, P., Wehner, A., Paulsen, H., and Hobe, S. (2001) *J. Biol. Chem.* **276**, 22154–22159
408. Gilmore, A. M., Shinkarev, V. P., Hazlett, T. L., and Govindjee. (1998) *Biochemistry* **37**, 13582–13593
409. Hanley, J., Deligiannakis, Y., Pascal, A., Faller, P., and Rutherford, A. W. (1999) *Biochemistry* **38**, 8189–8195
410. Buser, C. A., Diner, B. A., and Brudvig, G. W. (1992) *Biochemistry* **31**, 11449–11459
411. Barber, J., and Andersson, B. (1992) *Trends Biochem. Sci.* **17**, 61–66
412. Sharma, J., Panico, M., Shipton, C. A., Nilsson, F., Morris, H. R., and Barber, J. (1997) *J. Biol. Chem.* **272**, 33158–33166
413. Hagman, Å., Shi, L.-X., Rintamäki, E., Andersson, B., and Schröder, W. P. (1997) *Biochemistry* **36**, 12666–12671
414. Krieger, A., Rutherford, A. W., Vass, I., and Hideg, É. (1998) *Biochemistry* **37**, 16262–16269
415. Campbell, D., Eriksson, M.-J., Öquist, G., Gustafsson, P., and Clarke, A. K. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 364–369
415a. Rutherford, A. W., and Krieger-Liszkay, A. (2001) *Trends Biochem. Sci.* **26**, 648–653
416. Kobza, J., and Seemann, J. R. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3815–3819
417. Puente, P., Wei, N., and Deng, X. W. (1996) *EMBO J.* **15**, 3732–3743
418. Fluhr, R., Kuhlemeier, C., Nagy, F., and Chua, N.-H. (1986) *Science* **232**, 1106–1112
419. Moses, P. B., and Chua, N.-H. (1988) *Sci. Am.* **258**(Apr), 88–93
420. Klein, R. R., and Mullet, J. E. (1990) *J. Biol. Chem.* **265**, 1895–1902
421. Morelli, G., Nagy, F., Fraley, R. T., Roger, S. G., and Chua, N.-H. (1985) *Nature (London)* **315**, 200–204
422. Giuliano, G., Pichersky, E., Malik, V. S., Timko, M. P., Scolnik, P. A., and Cashmore, A. R. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7089–7093
423. Inamine, G., Nash, B., Weissbach, H., and Brot, N. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5690–5694
424. Buchanan, B. B., and Schurmann, P. (1973) *Curr. Top. Cell. Regul.* **7**, 1–20
425. Wolosiuk, R. A., Ballicora, M. A., and Hagelin, K. (1993) *FASEB J.* **7**, 622–637
426. Kelly, G. (1982) *Trends Biochem. Sci.* **7**, 81–82
427. Johansson, K., Ramaswamy, S., Saarinen, M., Lemaire-Chamley, M., Issakidis-Bourguet, E., Miginiac-Maslow, M., and Eklund, H. (1999) *Biochemistry* **38**, 4319–4326
428. Hirasawa, M., Schürmann, P., Jacquot, J.-P., Manieri, W., Jacquot, P., Keryer, E., Hartman, F. C., and Knaff, D. B. (1999) *Biochemistry* **38**, 5200–5205
429. Staples, C. R., Gaymard, E., Stritt-Etter, A.-L., Telser, J., Hoffman, B. M., Schürmann, P., Knaff, D. B., and Johnson, M. K. (1998) *Biochemistry* **37**, 4612–4620
429a. Dai, S., Schwendtmayer, C., Schürmann, P., Ramaswamy, S., and Eklund, H. (2000) *Science* **287**, 655–658
429b. Rintamäki, E., Martinsuo, P., Pursiheimo, S., and Aro, E.-M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11644–11649
430. Cseke, C., Balogh, A., Wong, J. H., Buchanan, B. B., Stitt, M., Herzog, B., and Held, H. W. (1984) *Trends Biochem. Sci.* **9**, 533–535
431. Bligny, R., Gardestrom, P., Roby, C., and Douce, R. (1990) *J. Biol. Chem.* **265**, 1319–1326
432. Zelitch, I. (1975) *Science* **188**, 626–633
433. Sommerville, C. R., and Ogren, W. L. (1982) *Trends Biochem. Sci.* **7**, 171–174
434. Heber, V., and Krause, G. H. (1980) *Trends Biochem. Sci.* **5**, 32–34
435. Goyal, A., and Tolbert, N. E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3319–3324
436. Tolbert, N. E. (1973) *Curr. Top. Cell. Regul.* **7**, 21–50
437. Ho, C.-L., Noji, M., and Saito, K. (1999) *J. Biol. Chem.* **274**, 11007–11012
438. Givan, C. V., Joy, K. W., and Kleczkowski, L. A. (1988) *Trends Biochem. Sci.* **13**, 433–437
439. Kozaki, A., and Takeba, G. (1996) *Nature (London)* **384**, 557–560
440. Haag, E., and Renger, G. (1997) in *Bioenergetics* (Gräber, P., and Milazzo, G., eds), pp. 212–272, Birkhäuser Verlag, Basel
441. Burnell, J. N., and Hatch, M. D. (1985) *Trends Biochem. Sci.* **10**, 289–291
442. Moore, P. D. (1978) *Nature (London)* **272**, 400–401
442a. Voznesenskaya, E. V., Franceschi, V. R., Kiirats, O., Freitag, H., and Edwards, G. E. (2001) *Nature (London)* **414**, 543–546
442b. Maeda, S.-i, Price, G. D., Badger, M. R., Enomoto, C., and Omata, T. (2000) *J. Biol. Chem.* **275**, 20551–20555
443. Winter, K., and Smith, J. A. C., eds. (1996) *Crassulacean Acid Metabolism Biochemistry, Ecophysiology and Evolution*, Springer, Berlin
444. Benemann, J. R., Berenson, J. A., Kaplan, N. O., and Kamen, M. D. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2317–2320
445. Benemann, J. R., and Weare, N. M. (1974) *Science* **184**, 174–175
446. Abrahamson, E. W., and Fager, R. S. (1973) *Curr. Top. Bioenerg.* **5**, 125–200
447. Wald, G., and Brown, P. K. (1965) *Cold Spring Harb. Symp. on Quant. Biol.* **30**, 346
448. Nathans, J. (1992) *Biochemistry* **31**, 4923–4931
449. Ferretti, L., Karnik, S. S., Khorana, H. G., Nassal, M., and Oprian, D. D. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 599–603
450. Khorana, H. G. (1992) *J. Biol. Chem.* **267**, 1–4
451. Dratz, E. A., and Hargrave, P. A. (1983) *Trends Biochem. Sci.* **8**, 128–131
452. Oprian, D. D., Asenjo.AB, Lee, N., and Pelletier, S. L. (1991) *Biochemistry* **30**, 11367–11372
453. Nassal, M., Mogi, T., Karnik, S. S., and Khorana, H. G. (1987) *J. Biol. Chem.* **262**, 9264–9270
454. Henderson, R. (1975) *J. Mol. Biol.* **93**, 123–138
455. Luecke, H., Schobert, B., Richter, H.-T., Cartailler, J.-P., and Lanyi, J. K. (1999) *J. Mol. Biol.* **291**, 899–911
456. Baldwin, J. (1993) *EMBO J.* **12**, 1693–1703
457. Baldwin, J. M., Schertler, G. F. X., and Unger, V. M. (1997) *J. Mol. Biol.* **272**, 144–64
458. Pogozheva, I. D., Lomize, A. L., and Mosberg, H. I. (1997) *Biophys. J.* **72**, 1963–1985
459. Schertler, G. F. X., and Hargrave, P. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11578–11582
460. Herzyk, P., and Hubbard, R. E. (1998) *J. Mol. Biol.* **281**, 741–754
461. Barnidge, D. R., Dratz, E. A., Sunner, J., and Jesaitis, A. J. (1997) *Protein Sci.* **6**, 816–824
461a. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Trong, I. L., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) *Science* **289**, 739–745
461b. Okada, T., Ernst, O. P., Palczewski, K., and Hofmann, K. P. (2001) *Trends Biochem. Sci.* **26**, 318–331
461c. Teller, D. C., Okada, T., Behnke, C. A., Palczewski, K., and Stenkamp, R. E. (2001) *Biochemistry* **40**, 7761–7772
462. Khorana, H. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1166–1171
462a. Altenbach, C., Klein-Seetharaman, J., Cai, K., Khorana, H. G., and Hubbell, W. L. (2001) *Biochemistry* **40**, 15493–15500
463. Yeagle, P. L., Alderfer, J. L., and Albert, A. D. (1995) *Biochemistry* **34**, 14621–14625
464. Altenbach, C., Klein-Seetharaman, J., Hwa, J., Khorana, H. G., and Hubbell, W. L. (1999) *Biochemistry* **38**, 7945–7949
465. Struthers, M., Yu, H., Kono, M., and Oprian, D. D. (1999) *Biochemistry* **38**, 6597–6603
466. Zhang, H., Lerro, K. A., Yamamoto, T., Lien, T. H., Sastry, L., Gawinowicz, M. A., and Nakanishi, K. (1994) *J. Am. Chem. Soc.* **116**, 10165–10173

References

467. Han, M., and Smith, S. O. (1995) *Biochemistry* **34**, 1425 – 1432
468. Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M., and Henderson, R. (1996) *J. Mol. Biol.* **259**, 393 – 421
469. Wald, G. (1968) *Nature (London)* **219**, 800 – 807
470. Dowling, J. E. (1997) *Nature (London)* **387**, 356
471. Gilardi, R., Sperling, W., Karle, I. L., and Karle, J. (1971) *Nature (London)* **232**, 187 – 188
472. Shriver, J. W., Mateescu, G. D., and Abrahamson, E. W. (1979) *Biochemistry* **18**, 4785 – 4792
472a. Singh, D., Hudson, B. S., Middleton, C., and Birge, R. R. (2001) *Biochemistry* **40**, 4201 – 4204
473. Honig, B., Dinur, U., Nakanishi, K., Balogh-Nair, V., Gawinswicz, M. A., Arnaboldi, M., and Motto, M. G. (1979) *J. Am. Chem. Soc.* **101**, 7084 – 7086
474. Baasov, T., and Sheves, M. (1985) *J. Am. Chem. Soc.* **107**, 7524 – 7533
475. Sakmar, T. P., Franke, R. R., and Khorana, H. G. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8309 – 8313
475a. Lewis, J. W., Szundi, I., Fu, W.–Y., Sakmar, T. P., and Kliger, D. S. (2000) *Biochemistry* **39**, 599 – 606
476. Creemers, A. F. L., Klaassen, C. H. W., Bovee-Geurts, P. H. M., Kelle, R., Kragl, U., Raap, J., de Grip, W. J., Lugtenburg, J., and de Groot, H. J. M. (1999) *Biochemistry* **38**, 7195 – 7199
476a. Verhoeven, M. A., Creemers, A. F. L., Bovee-Geurts, P. H. M., De Grip, W. J., Lugtenburg, J., and de Groot, H. J. M. (2001) *Biochemistry* **40**, 3282 – 3288
477. Kochendoerfer, G. G., Lin, S. W., Sakmar, T. P., and Mathies, R. A. (1999) *Trends Biochem. Sci.* **24**, 300 – 305
478. Nathans, J., Thomas, D., and Hogness, D. S. (1986) *Science* **232**, 193 – 202
479. Kochendoerfer, G. G., Wang, Z., Oprian, D. D., and Mathies, R. A. (1997) *Biochemistry* **36**, 6577 – 6587
479a. Dukkipati, A., Vought, B. W., Singh, D., Birge, R. R., and Knox, B. E. (2001) *Biochemistry* **40**, 15098 – 15108
479b. Roorda, A., and Williams, D. R. (1999) *Nature (London)* **397**, 520–522
480. Nathans, J., Piantanida, T. D., Eddy, R. L., Shows, T. B., and Hogness, D. S. (1986) *Science* **232**, 203 – 210
481. Motulsky, A. G., and Deeb, S. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4275 – 4295, McGraw-Hill, New York
482. Nathans, J. (1989) *Sci. Am.* **260**(Feb), 42 – 49
483. Merbs, S. L., and Nathans, J. (1992) *Science* **258**, 464 – 466
484. Hunt, D. M., Dulai, K. S., Bowmaker, J. K., and Mollon, J. D. (1995) *Science* **267**, 984 – 988
485. Nathans, J., Davenport, C. M., Maumenee, I. H., Lewis, R. A., Hejtmancik, J. F., Litt, M., Lovrien, E., Weleber, R., Bachynski, B., Zwas, F., Klingaman, R., and Fishman, G. (1989) *Science* **245**, 831 – 838
486. Shichi, H., Lewis, M. S., Irreverre, F. and Stone, A. L. (1969) *J. Biol. Chem.* **244**, 529 – 536
487. Shichi, H., and Somers, R. L. (1974) *J. Biol. Chem.* **249**, 6570 – 6577
488. Fasick, J. I., Lee, N., and Oprian, D. D. (1999) *Biochemistry* **38**, 11593 – 11596
488a. Shi, Y., Radlwimmer, F. B., and Yokoyama, S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11731 – 11736
488b. Janz, J. M., and Farrens, D. L. (2001) *Biochemistry* **40**, 7219 – 7227
488c. Dukkipati, A., Kusnetzow, A., Babu, K. R., Ramos, L., Singh, D., Knox, B. E., and Birge, R. R. (2002) *Biochemistry* **41**, 9842–9851
489. Fasick, J. I., and Robinson, P. R. (1998) *Biochemistry* **37**, 433 – 438
490. Imamoto, Y., Hirano, T., Imai, H., Kandori, H., Maeda, A., Yoshizawa, T., Groesbeek, M., Lugtenburg, J., and Shichida, Y. (1999) *Biochemistry* **38**, 11749 – 11754
490a. Imai, H., Hirano, T., Kandori, H., Terakita, A., and Shichida, Y. (2001) *Biochemistry* **40**, 2879 – 2886
491. Wang, Z., Asenjo, A. B., and Oprian, D. D. (1993) *Biochemistry* **32**, 2125 – 2130
492. Johnson, R. L., Grant, K. B., Zankel, T. C., Boehm, M. F., Merbs, S. L., Nathans, J., and Nakanishi, K. (1993) *Biochemistry* **32**, 208 – 214
493. Foster, K. W., Saranak, J., Derguini, F., Zarrilli, G. R., Johnson, R., Okabe, M., and Nakanishi, K. (1989) *Biochemistry* **28**, 819 – 824
494. Deininger, W., Kröger, P., Hegemann, U., Lottspeich, F., and Hegemann, P. (1995) *EMBO J.* **14**, 5849 – 5858
495. Max, M., Surya, A., Takahashi, J. S., Margolskee, R. F., and Knox, B. E. (1998) *J. Biol. Chem.* **273**, 26820 – 26826
496. Max, M., McKinnon, P. J., Seidenman, K. J., Barrett, R. K., Applebury, M. L., Takahashi, J. S., and Margolskee, R. F. (1995) *Science* **267**, 1502 – 1506
496a. Nakamura, A., Kojima, D., Imai, H., Terakita, A., Okano, T., Shichida, Y., and Fukada, Y. (1999) *Biochemistry* **38**, 14738 – 14745
497. Soni, B. G., Philp, A. R., Foster, R. G., and Knox, B. E. (1998) *Nature (London)* **394**, 27 – 28
498. Zuker, C. S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 571 – 576
499. Peters, K., Applebury, M. L., and Rentzepis, P. M. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3119 – 3123
500. Strassburger, J. M., Gärtner, W., and Braslavsky, S. E. (1997) *Biophys. J.* **72**, 2294 – 2303
501. Zhu, Y., and Liu, R. S. H. (1993) *Biochemistry* **32**, 10233 – 10238
502. Liu, R. S. H., and Shichida, Y. (1991) in *Photochemistry in Organized and Constrained Media* (Ramamurthy, V., ed), VCH Publishers, New York (Chapter 18)
502a. Ishiguro, M. (2000) *J. Am. Chem. Soc.* **122**, 444 – 451
503. Hug, S. J., Lewis, W. J., Einterz, C. M., Thorgeirsson, T. E., and Kliger, D. S. (1990) *Biochemistry* **29**, 1475 – 1485
504. Lewis, J. W., Pinkas, I., Sheves, M., Ottolenghi, M., and Kliger, D. S. (1995) *J. Am. Chem. Soc.* **117**, 918 – 923
505. Shichida, Y., Nakamura, K., Yoshizawa, T., Trehan, A., Denny, M., and Liu, R. S. H. (1988) *Biochemistry* **27**, 6495 – 6499
506. Wang, Q., Schoenlein, R. W., Peteanu, L. A., Mathies, R. A., and Shank, C. V. (1994) *Science* **266**, 422 – 424
507. Jäger, F., Fahmy, K., Sakmar, T. P., and Siebert, F. (1994) *Biochemistry* **33**, 10878 – 10882
508. DeLange, F., Bovee-Geurts, P. H. M., VanOostrum, J., Portier, M. D., Verdegem, P. J. E., Lugtenburg, J., and DeGrip, W. J. (1998) *Biochemistry* **37**, 1411 – 1420
509. Fukada, Y., Shichida, Y., Yoshizawa, T., Ito, M., Kodama, A., and Tsukida, K. (1984) *Biochemistry* **23**, 5826 – 5832
510. Bagley, K. A., Balogh-Nair, V., Croteau, A. A., Dollinger, G., Ebrey, T. G., Eisenstein, L., Hong, M. K., Nakanishi, K., and Vittitow, J. (1985) *Biochemistry* **24**, 6055 – 6071
510a. Borhan, B., Souto, M. L., Imai, H., Shichida, Y., and Nakanishi, K. (2000) *Science* **288**, 2209 – 2212
511. Ridge, K. D., Bhattacharya, S., Nakayama, T. A., and Khorana, H. G. (1992) *J. Biol. Chem.* **267**, 6770 – 6775
511a. Liu, R. S. H., and Hammond, G. S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11153 – 11158
511b. González-Luque, R., Garavelli, M., Bernardi, F., Merchán, M., and Robb, M. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9379 – 9384
511c. Kim, J. E., Tauber, M. J., and Mathies, R. A. (2001) *Biochemistry* **40**, 13774 – 13778
512. Chosrowjan, H., Mataga, N., Shibata, Y., Tachibanaki, S., Kandori, H., Shichida, Y., Okada, T., and Kouyama, T. (1998) *J. Am. Chem. Soc.* **120**, 9706 – 9707
513. Loppnow, G. R., Miley, M. E., Mathies, R. A., Liu, R. S. H., Kandori, H., Shichida, Y., Fukada, Y., and Yoshizawa, T. (1990) *Biochemistry* **29**, 8985 – 8991
514. Freedman, K. A., and Becker, R. S. (1986) *J. Am. Chem. Soc.* **108**, 1245 – 1251
515. Yoshizawa, T., and Wald, G. (1963) *Nature (London)* **197**, 1279 – 1286
516. Nagata, T., Terakita, A., Kandori, H., Kojima, D., Shichida, Y., and Maeda, A. (1997) *Biochemistry* **36**, 6164 – 6170
517. Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L., and Khorana, H. G. (1996) *Science* **274**, 768 – 770
518. Dunham, T. D., and Farrens, D. L. (1999) *J. Biol. Chem.* **274**, 1683 – 1690
518a. Bartl, F. J., Ritter, E., and Hofmann, K. P. (2001) *J. Biol. Chem.* **276**, 30161 – 30166
518b. Vogel, R., Fan, G.-B., Siebert, F., and Sheves, M. (2001) *Biochemistry* **40**, 13342 – 13352
519. Dickopf, S., Mielke, T., and Heyn, M. P. (1998) *Biochemistry* **37**, 16888 – 16897
520. Jäger, S., Szundi, I., Lewis, J. W., Mah, T. L., and Kliger, D. S. (1998) *Biochemistry* **37**, 6998 – 7005
521. Parkes, J. H., and Liebman, P. A. (1984) *Biochemistry* **23**, 5054 – 5061
521a. Yeagle, P. L., Choi, G., and Albert, A. D. (2001) *Biochemistry* **40**, 11932 – 11937
521b. Vought, B. W., Salcedo, E., Chadwell, L. V., Britt, S. G., Birge, R. R., and Knox, B. E. (2000) *Biochemistry* **39**, 14128 – 14137
521c. Pan, D., and Mathies, R. A. (2001) *Biochemistry* **40**, 7929 – 7936
522. Yarfitz, S., and Hurley, J. B. (1994) *J. Biol. Chem.* **269**, 14329 – 14332
523. Puckett, K. L., Aronson, E. T., and Goldin, S. M. (1985) *Biochemistry* **24**, 390 – 400
524. Dizhoor, A. M., Olshevskaya, E. V., Henzel, W. J., Wong, S. C., Stults, J. T., Ankoudinova, I., and Hurley, J. B. (1995) *J. Biol. Chem.* **270**, 25200 – 25206
525. Yau, K.-W., and Nakatani, K. (1985) *Nature (London)* **313**, 579 – 583
526. Stryer, L., Hurley, J. B., and Fung, B. K. K. (1981) *Trends Biochem. Sci.* **6**, 245 – 247
527. Stryer, L. (1987) *Sci. Am.* **257**(Jul), 42 – 50
528. Stryer, L. (1991) *J. Biol. Chem.* **266**, 10711 – 10714
529. Lambright, D. G., Sondek, J., Bohm, A., Skiba, N. P., Hamm, H. E., and Sigler, P. B. (1996) *Nature (London)* **379**, 311 – 319
529a. Granovsky, A. E., and Artemyev, N. O. (2001) *Biochemistry* **40**, 13209 – 13215
529b. Norton, A. W., D'Amours, M. R., Grazio, H. J., Hebert, T. L., and Cote, R. H. (2000) *J. Biol. Chem.* **275**, 38611 – 38619
530. Shimoda, Y., Hurley, J. B., and Miller, W. H. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 616 – 619
531. Gibson, S. K., Parkes, J. H., and Liebman, P. A. (1998) *Biochemistry* **37**, 11393 – 11398
532. Palczewski, K., Buczylko, J., Lebioda, L., Crabb, J. W., and Polans, A. S. (1993) *J. Biol. Chem.* **268**, 6004 – 6013
533. Gurevich, V. V., Chen, C.-Y., Kim, C. M., and Benovic, J. L. (1994) *J. Biol. Chem.* **269**, 8721 – 8727

References

534. Smith, W. C., McDowell, J. H., Dugger, D. R., Miller, R., Arendt, A., Popp, M. P., and Hargrave, P. A. (1999) *Biochemistry* **38**, 2752–2761
535. Granzin, J., Wilden, U., Choe, H.-W., Labahn, J., Krafft, B., and Büldt, G. (1998) *Nature (London)* **391**, 918–921
535a. Oakley, R. H., Laporte, S. A., Holt, J. A., Caron, M. G., and Barak, L. S. (2000) *J. Biol. Chem.* **275**, 17201–17210
536. Baylor, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 560–565
536a. Penn, R. B., Pascual, R. M., Kim, Y.-M., Mundell, S. J., Krymskaya, V. P., Panettieri, R. A., Jr., and Benovic, J. L. (2001) *J. Biol. Chem.* **276**, 32648–32656
537. Laitko, U., and Hofmann, K. P. (1998) *Biophys. J.* **74**, 803–815
538. Klenchin, V. A., Calvert, P. D., and Bownds, M. D. (1995) *J. Biol. Chem.* **270**, 16147–16152
539. Ames, J. B., Ishima, R., Tanaka, T., Gordon, J. I., Stryer, L., and Ikura, M. (1997) *Nature (London)* **389**, 198–202
540. Ames, J. B., Tanaka, T., Stryer, L., and Ikura, M. (1996) *Curr. Opin. Struct. Biol.* **6**, 432–438
541. Lange, C., and Koch, K.-W. (1997) *Biochemistry* **36**, 12019–12026
542. Palczewski, K., Jäger, S., Buczylko, J., Crouch, R. K., Bredberg, D. L., Hofmann, K. P., Asson-Batres, M. A., and Saari, J. C. (1994) *Biochemistry* **33**, 13741–13750
543. Palczewski, K., Van Hooser, J. P., Garwin, G. G., Chen, J., Liou, G. I., and Saari, J. C. (1999) *Biochemistry* **38**, 12012–12019
543a. McBee, J. K., Kuksa, V., Alvarez, R., de Lera, A. R., Prezhdo, O., Haeseleer, F., Sokal, I., and Palczewski, K. (2000) *Biochemistry* **39**, 11370–11380
544. Hara, R., Hara, T., Tokunaga, F., and Yoshizawa, T. (1981) *Photochem. Photobiol.* **33**, 883–891
544a. Terakita, A., Yamashita, T., and Shichida, Y. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14263–14267
545. Humphries, P., Kenna, P., and Farrar, G. J. (1992) *Science* **256**, 804–808
546. Dryja, T. P. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4297–4309, McGraw-Hill, New York
546a. Hwa, J., Klein-Seetharaman, J., and Khorana, H. G. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4872–4876
547. Maniloff, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10004–10006
548. Liu, X., Garriga, P., and Khorana, H. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4554–4559
549. Rao, V. R., Cohen, G. B., and Oprian, D. D. (1994) *Nature (London)* **367**, 639–642
550. Kajiwara, K., Berson, E. L., and Dryja, T. P. (1994) *Science* **264**, 1604–1608
551. Khani, S. C., Nielsen, L., and Vogt, T. M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2824–2827
552. Tucker, C. L., Woodcock, S. C., Kelsell, R. E., Ramamurthy, V., Hunt, D. M., and Hurley, J. B. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9039–9044
553. Eudy, J. D., Weston, M. D., Yao, S., Hoover, D. M., Rehm, H. L., Ma-Edmonds, M., Yan, D., Ahmad, I., Cheng, J. J., Ayuso, C., Cremers, C., Davenport, S., Moller, C., Talmadge, C. B., Beisel, K. W., Tamayo, M., Morton, C. C., Swaroop, A., Kimberling, W. J., and Sumegi, J. (1998) *Science* **280**, 1753–1757
554. Seabra, M. C., Brown, M. S., and Goldstein, J. L. (1993) *Science* **259**, 377–381
555. Allikmets, R., Shroyer, N. F., Singh, N., Seddon, J. M., Lewis, R. A., Bernstein, P. S., Peiffer, A., Zabriskie, N. A., Li, Y., Hutchinson, A., Dean, M., Lupski, J. R., and Leppert, M. (1997) *Science* **277**, 1805–1807
556. Pennisi, E. (1998) *Science* **281**, 31
556a. Biswas, E. E. (2001) *Biochemistry* **40**, 8181–8187
557. Parish, C. A., Hashimoto, M., Nakanishi, K., Dillon, J., and Sparrow, J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14609–14613
558. Norledge, B. V., Trinkl, S., Jaenicke, R., and Slingsby, C. (1997) *Protein Sci.* **6**, 1612–1620
559. Koretz, J. F., and Handelman, G. H. (1988) *Sci. Am.* **259**(Jul), 92–99
560. Tardieu, A., and Delaye, M. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 47–70
561. Wistow, G. (1993) *Trends Biochem. Sci.* **18**, 301–306
562. Hejtmancik, J. F., Kaiser, M. I., and Piatigorsky, J. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4325–4349, McGraw-Hill, New York
563. König, N., Zampighi, G. A., and Butler, P. J. G. (1997) *J. Mol. Biol.* **265**, 590–602
564. Surewicz, W. K., and Olesen, P. R. (1995) *Biochemistry* **34**, 9655–9660
565. Clark, J. I., and Huang, Q.-L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15185–15189
566. Piatigorsky, J., and Wistow, G. (1991) *Science* **252**, 1078–1079
567. Röll, B., Amons, R., and de Jong, W. W. (1996) *J. Biol. Chem.* **271**, 10437–10440
568. Zigler, J. S., Jr., and Rao, P. V. (1991) *FASEB J.* **5**, 223–225
569. Abu-Abed, M., Turner, M. A., Vallée, F., Simpson, A., Slingsby, C., and Howell, P. L. (1997) *Biochemistry* **36**, 14012–14022
570. Kraft, H. J., Hendriks, W., de Jong, W. W., Lubsen, N. H., and Schoenmakers, J. G. G. (1993) *J. Mol. Biol.* **229**, 849–859
571. Mulders, J. W. M., Hendriks, W., Blankesteijn, W. M., Bloemendal, H., and de Jong, W. W. (1988) *J. Biol. Chem.* **263**, 15462–15466
572. Zinovieva, R. D., Tomarev, S. I., and Piatigorsky, J. (1993) *J. Biol. Chem.* **268**, 11449–11455
572a. Werten, P. J. L., Röll, B., van Aalten, D. M. F., and de Jong, W. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3282–3287
573. Garner, B., Vazquez, S., Griffith, R., Lindner, R. A., Carver, J. A., and Truscott, R. J. W. (1999) *J. Biol. Chem.* **274**, 20847–20854
574. Tessier, F., Obrenovich, M., and Monnier, V. M. (1999) *J. Biol. Chem.* **274**, 20796–20804
574a. Hood, B. D., Garner, B., and Truscott, R. J. W. (1999) *J. Biol. Chem.* **274**, 32547–32550
575. Takemoto, L., and Boyle, D. (1998) *Biochemistry* **37**, 13681–13685
576. Stoeckenius, W. (1999) *Protein Sci.* **8**, 447–459
577. Racker, E., and Stoeckenius, W. (1974) *J. Biol. Chem.* **249**, 662–663
578. Stoeckenius, W. (1985) *Trends Biochem. Sci.* **10**, 483–486
579. Mitsuoka, K., Hirai, T., Murata, K., Miyazawa, A., Kidera, A., Kimura, Y., and Fujiyoshi, Y. (1999) *J. Mol. Biol.* **286**, 861–882
580. Kimura, Y., Vassylyev, D. G., Miyazawa, A., Kidera, A., Matsushima, M., Mitsuoka, K., Murata, K., Hirai, T., and Fujiyoshi, Y. (1997) *Nature (London)* **389**, 206–211
581. Subramaniam, S., Lindahl, M., Bullough, P., Faruqi, A. R., Tittor, J., Oesterhelt, D., Brown, L., Lanyi, J., and Henderson, R. (1999) *J. Mol. Biol.* **287**, 145–161
582. Pebay-Peyroula, E., Rummel, G., Rosenbusch, J. P., and Landau, E. M. (1997) *Science* **277**, 1676–1681
582a. Subramaniam, S., and Henderson, R. (2000) *Nature (London)* **406**, 653–657
582b. Edman, K., Nollert, P., Royant, A., Belrhali, H., Pebay-Peyroula, E., Hajdu, J., Neutze, R., and Landau, E. M. (1999) *Nature (London)* **401**, 822–826
582c. Royant, A., Edman, K., Ursby, T., Pebay-Peyroula, E., Landau, E. M., and Neutze, R. (2000) *Nature (London)* **406**, 645–648
582d. Luecke, H., Schobert, B., Cartailler, J.-P., Richter, H.-T., Rosengarth, A., Needleman, R., and Lanyi, J. K. (2000) *J. Mol. Biol.* **300**, 1237–1255
582e. Xiao, W., Brown, L. S., Needleman, R., Lanyi, J. K., and Shin, Y.-K. (2000) *J. Mol. Biol.* **304**, 715–721
582f. Kühlbrandt, W. (2000) *Nature (London)* **406**, 569–570
583. Müller, D. J., Sass, H.-J., Müller, S. A., Büldt, G., and Engel, A. (1999) *J. Mol. Biol.* **285**, 1903–1909
584. Smith, S. O., Myers, A. B., Pardoen, J. A., Winkel, C., Mulder, P. P. J., Lugtenburg, J., and Mathies, R. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2055–2059
584a. Patzelt, H., Simon, B., terLaak, A., Kessler, B., Kühne, R., Schmieder, P., Oesterhelt, D., and Oschkinat, H. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9765–9770
585. Bullough, P. A., and Henderson, R. (1999) *J. Mol. Biol.* **286**, 1663–1671
585a. Herbst, J., Heyne, K., and Diller, R. (2002) *Science* **297**, 822–825
585b. Kandori, H., Belenky, M., and Herzfeld, J. (2002) *Biochemistry* **41**, 6026–6031
585c. Maeda, A., Balashov, S. P., Lugtenburg, J., Verhoeven, M. A., Herzfeld, J., Belenky, M., Gennis, R. B., Tomson, F. L., and Ebrey, T. G. (2002) *Biochemistry* **41**, 3803–3809
586. Haupts, U., Tittor, J., Bamberg, E., and Oesterhelt, D. (1997) *Biochemistry* **36**, 2–7
587. Gat, Y., and Sheves, M. (1993) *J. Am. Chem. Soc.* **115**, 3772–3773
588. Moltke, S., Wallat, I., Sakai, N., Nakanishi, K., Brown, M. F., and Heyn, M. P. (1999) *Biochemistry* **38**, 11762–11772
588a. Sass, H. J., Büldt, G., Gessenich, R., Hehn, D., Neff, D., Schlesinger, R., Berendzen, J., and Ormos, P. (2000) *Nature (London)* **406**, 649–653
588b. Kandori, H., Yamazaki, Y., Shichida, Y., Raap, J., Lugtenburg, J., Belenky, M., and Herzfeld, J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1571–1576
588c. Dioumaev, A. K., Brown, L. S., Needleman, R., and Lanyi, J. K. (2001) *Biochemistry* **40**, 11308–11317
588d. Oka, T., Yagi, N., Fujisawa, T., Kamikubo, H., Tokunaga, F., and Kataoka, M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14278–14282
588e. Vonck, J. (2000) *EMBO J.* **19**, 2152–2160
588f. Rouhani, S., Cartailler, J.-P., Facciotti, M. T., Walian, P., Needleman, R., Lanyi, J. K., Glaeser, R. M., and Luecke, H. (2001) *J. Mol. Biol.* **313**, 615–628
588g. Allen, S. J., Kim, J.-M., Khorana, H. G., Lu, H., and Booth, P. J. (2001) *J. Mol. Biol.* **308**, 423–435
588h. Aharoni, A., Weiner, L., Lewis, A., Ottolenghi, M., and Sheves, M. (2001) *J. Am. Chem. Soc.* **123**, 6612–6616
588i. Spassov, V. Z., Luecke, H., Gerwert, K., and Bashford, D. (2001) *J. Mol. Biol.* **312**, 203–219
588j. Imasheva, E. S., Lu, M., Balashov, S. P., Ebrey, T. G., Chen, Y., Ablonczy, Z., Menick, D. R., and Crouch, R. K. (2001) *Biochemistry* **40**, 13320–13330
589. Havelka, W. A., Henderson, R., and Oesterhelt, D. (1995) *J. Mol. Biol.* **247**, 726–738

References

590. Chon, Y.-S., Kandori, H., Sasaki, J., Lanyi, J. K., Needleman, R., and Maeda, A. (1999) *Biochemistry* **38**, 9449 – 9455

590a. Sato, M., Kanamori, T., Kamo, N., Demura, M., and Nitta, K. (2002) *Biochemistry* **41**, 2452 – 2458

591. Seiff, F., Wallat, I., Ermann, P., and Heyn, M. P. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3227 – 3231

591a. Kandori, H., Furutani, Y., Shimono, K., Shichida, Y., and Kamo, N. (2001) *Biochemistry* **40**, 15693 – 15698

591b. Swartz, T. E., Szundi, I., Spudich, J. L., and Bogomolni, R. A. (2000) *Biochemistry* **39**, 15101 – 15109

591c. Wegener, A.-A., Klare, J. P., Engelhard, M., and Steinhoff, H.-J. (2001) *EMBO J.* **20**, 5312 – 5319

592. Jung, K.-H., and Spudich, J. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6557 – 6561

593. Zhang, X.-N., and Spudich, J. L. (1997) *Biophys. J.* **73**, 1516 – 1523

593a. Luecke, H., Schobert, B., Lanyi, J. K., Spudich, E. N., and Spudich, J. L. (2001) *Science* **293**, 1499 – 1503

593b. Royant, A., Nollert, P., Edman, K., Neutze, R., Landau, E. M., Pebay-Peyroula, E., and Navarro, J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10131 – 10136

594. Hirayama, J., Imamoto, Y., Shichida, Y., Kamo, N., Tomioka, H., and Yoshizawa, T. (1992) *Biochemistry* **31**, 2093 – 2098

595. Spudich, E. N., Zhang, W., Alam, M., and Spudich, J. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4960 – 4965

596. Chizhov, I., Schmies, G., Seidel, R., Sydor, J. R., Lüttenberg, B., and Engelhard, M. (1998) *Biophys. J.* **75**, 999 – 1009

597. Rath, P., Olson, K. D., Spudich, J. L., and Rothschild, K. J. (1994) *Biochemistry* **33**, 5600 – 5606

597a. Bergo, V., Spudich, E. N., Scott, K. L., Spudich, J. L., and Rothschild, K. J. (2000) *Biochemistry* **39**, 2823 – 2830

597b. Ren, L., Martin, C. H., Wise, K. J., Gillespie, N. B., Luecke, H., Lanyi, J. K., Spudich, J. L., and Birge, R. R. (2001) *Biochemistry* **40**, 13906–13914

598. Bieszke, J. A., Braun, E. L., Bean, L. E., Kang, S., Natvig, D. O., and Borkovich, K. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8034 – 8039

598a. Sineshchekov, O. A., Jung, K.-H, and Spudich, J. L. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8689–8694

599. Hoff, W. D., Devreese, B., Fokkens, R., Nugteren-Roodzant, I. M., Van Beeumen, J., Nibbering, N., and Hellingwerf, K. J. (1996) *Biochemistry* **35**, 1274 – 1281

599a. Xie, A., Kelemen, L., Hendriks, J., White, B. J., Hellingwerf, K. J., and Hoff, W. D. (2001) *Biochemistry* **40**, 1510 – 1517

600. Kort, R., Hoff, W. D., Van West, M., Kroon, A. R., Hoffer, S. M., Vlieg, K. H., Crielaard, W., Van Beeumen, J. J., and Hellingwerf, K. J. (1996) *EMBO J.* **15**, 3209 – 3218

601. Borgstahl, G. E. O., Williams, D. R., and Getzoff, E. D. (1995) *Biochemistry* **34**, 6278 – 6287

601a. Brudler, R., Meyer, T. E., Genick, U. K., Devanathan, S., Woo, T. T., Millar, D. P., Gerwert, K., Cusanovich, M. A., Tollin, G., and Getzoff, E. D. (2000) *Biochemistry* **39**, 13478 – 13486

601b. Ren, Z., Perman, B., Srajer, V., Teng, T.-Y., Pradervand, C., Bourgeois, D., Schotte, F., Ursby, T., Kort, R., Wulff, M., and Moffat, K. (2001) *Biochemistry* **40**, 13788 – 13801

601c. Imamoto, Y., Kataoka, M., Tokunaga, F., Asahi, T., and Masuhara, H. (2001) *Biochemistry* **40**, 6047 – 6052

602. Imamoto, Y., Mihara, K., Hisatomi, O., Kataoka, M., Tokunaga, F., Bojkova, N., and Yoshihara, K. (1997) *J. Biol. Chem.* **272**, 12905 – 12908

603. Perman, B., Srajer, V., Ren, Z., Teng, T., Pradervand, C., Ursby, T., Bourgeois, D., Schotte, F., Wulff, M., Kort, R., Hellingwerf, K., and Moffat, K. (1998) *Science* **279**, 1946 – 1950

604. Genick, U. K., Devanathan, S., Meyer, T. E., Canestrelli, I. L., Williams, E., Cusanovich, M. A., Tollin, G., and Getzoff, E. D. (1997) *Biochemistry* **36**, 8 – 14

605. Ujj, L., Devanathan, S., Meyer, T. E., Cusanovich, M. A., Tollin, G., and Atkinson, G. H. (1998) *Biophys. J.* **75**, 406 – 412

606. Imamoto, Y., Kataoka, M., and Tokunaga, F. (1996) *Biochemistry* **35**, 14047 – 14053

607. Genick, U. K., Soltis, S. M., Kuhn, P., Canestrelli, I. L., and Getzoff, E. D. (1998) *Nature (London)* **392**, 206 – 209

608. Devanathan, S., Genick, U. K., Canestrelli, I. L., Meyer, T. E., Cusanovich, M. A., Getzoff, E. D., and Tollin, G. (1998) *Biochemistry* **37**, 11563 – 11568

609. Xie, A., Hoff, W. D., Kroon, A. R., and Hellingwerf, K. J. (1996) *Biochemistry* **35**, 14671 – 14678

609a. Unno, M., Kumauchi, M., Sasaki, J., Tokunaga, F., and Yamauchi, S. (2002) *Biochemistry* **41**, 5668–5674

610. Tao, N., Orlando, M., Hyon, J.-S., Gross, M., and Song, P.-S. (1993) *J. Am. Chem. Soc.* **115**, 2526 – 2528

611. Song, P.-S. (1983) *Annu Rev Biophys Bioeng.* **12**, 35 – 68

612. Smirnov, A., Fulton, D. B., Andreotti, A., and Petrich, J. W. (1999) *J. Am. Chem. Soc.* **121**, 7979 – 7988

613. Carpenter, S., Fehr, M. J., Kraus, G. A., and Petrich, J. W. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12273 – 12277

614. Quail, P. H. (1984) *Trends Biochem. Sci.* **9**, 450 – 453

615. Kelly, J. M., and Lagarias, J. C. (1985) *Biochemistry* **24**, 6003 – 6010

616. Quail, P. H. (1976) in *Plant Biochemistry*, 3rd ed. (Bonner, J., and Varner, J. E., eds), pp. 683 – 711, Academic Press, New York

617. Anderson, G. R., Jenner, E. L., and Mumford, F. E. (1970) *Biochim. Biophys. Acta.* **221**, 69 – 73

618. Quail, P. H., Boylan, M. T., Parks, B. M., Short, T. W., Xu, Y., and Wagner, D. (1995) *Science* **268**, 675 – 680

619. Ruddat, A., Schmidt, P., Gatz, C., Braslavsky, S. E., Gärtner, W., and Schaffner, K. (1997) *Biochemistry* **36**, 103 – 111

619a. Fankhauser, C. (2001) *J. Biol. Chem.* **276**, 11453 – 11456

619b. Nagy, F., and Schäfer, E. (2000) *EMBO J.* **19**, 157 – 163

619c. Smith, H. (2000) *Nature (London)* **407**, 585 – 591

620. Foerstendorf, H., Mummert, E., Schäfer, E., Scheer, H., and Siebert, F. (1996) *Biochemistry* **35**, 10793 – 10799

621. Andel, F., III, Lagarias, J. C., and Mathies, R. A. (1996) *Biochemistry* **35**, 15997 – 16008

622. Chen, E., Lapko, V. N., Lewis, J. W., Song, P.-S., and Kliger, D. S. (1996) *Biochemistry* **35**, 843 – 850

623. Kneip, C., Hildebrandt, P., Schlamann, W., Braslavsky, S. E., Mark, F., and Schaffner, K. (1999) *Biochemistry* **38**, 15185 – 15192

623a. Foerstendorf, H., Benda, C., Gärtner, W., Storf, M., Scheer, H., and Siebert, F. (2001) *Biochemistry* **40**, 14952 – 14959

623b. Andel, F., III, Murphy, J. T., Haas, J. A., McDowell, M. T., van der Hoef, I., Lugtenburg, J., Lagarias, J. C., and Mathies, R. A. (2000) *Biochemistry* **39**, 2667 – 2676

623c. Hennig, L., and Schäfer, E. (2001) *J. Biol. Chem.* **276**, 7913 – 7918

623d. Huq, E., Tepperman, J. M., and Quail, P. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9789 – 9794

623e. Park, C.-M., Kim, J.-I., Yang, S.-S., Kang, J.-G., Kang, J.-H., Shim, J.-Y., Chung, Y.-H., Park, Y.-M., and Song, P.-S. (2000) *Biochemistry* **39**, 10840 – 10847

623f. Schmitz, O., Katayama, M., Williams, S. B., Kondo, T., and Golden, S. S. (2000) *Science* **289**, 765 – 768

624. Yeh, K.-C., Wu, S.-H., Murphy, J. T., and Lagarias, J. C. (1997) *Science* **277**, 1505 – 1508

625. Jiang, Z., Swem, L. R., Rushing, B. G., Devanathan, S., Tollin, G., and Bauer, C. E. (1999) *Science* **285**, 406 – 409

625a. Bhoo, S.-H., Davis, S. J., Walker, J., Karniol, B., and Vierstra, R. D. (2001) *Nature (London)* **414**, 776 – 779

626. Smith, H. (1999) *Nature (London)* **400**, 710 – 712

627. Ni, M., Tepperman, J. M., and Quail, P. H. (1999) *Nature (London)* **400**, 781 – 784

627a. Sweere, U., Eichenberg, K., Lohrmann, J., Mira-Rodado, V., Bäurle, I., Kudla, J., Nagy, F., Schäfer, E., and Harter, K. (2001) *Science* **294**, 1108 – 1111

628. Hoecker, U., Tepperman, J. M., and Quail, P. H. (1999) *Science* **284**, 496 – 499

629. Evans, A., and Smith, H. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 138 – 142

630. Song, P.-S., and Moore, T. A. (1974) *Photochem. Photobiol.* **19**, 435 – 441

631. Quiñones, M. A., Lu, Z., and Zeiger, E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2224 – 2228

632. Cashmore, A. R., Jarillo, J. A., Wu, Y.-J., and Liu, D. (1999) *Science* **284**, 760 – 765

633. Presti, D., Hsu, W.-J., and Delbrück, M. (1976) *Photochem. Photobiol.* **26**, 403 – 405

634. Galland, P., and Lipson, E. D. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 104 – 108

635. Christie, J. M., Reymond, P., Powell, G. K., Bernasconi, P., Raibekas, A. A., Liscum, E., and Briggs, W. R. (1998) *Science* **282**, 1698 – 1701

636. Nozue, K., Kanegae, T., Imaizumi, T., Fukuda, S., Okamoto, H., Yeh, K.-C., Lagarias, J. C., and Wada, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15816 – 15830

636a. Sakai, T., Kagawa, T., Kasahara, M., Swartz, T. E., Christie, J. M., Briggs, W. R., Wada, M., and Okada, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6969 – 6974

636b. Motchoulski, A., and Liscum, E. (1999) *Science* **286**, 961 – 964

636c. Christie, J. M., and Briggs, W. R. (2001) *J. Biol. Chem.* **276**, 11457 – 11460

636d. Kinoshita, T., Doi, M., Suetsugu, N., Kagawa, T., Wada, M., and Shimazaki, K.-i. (2001) *Nature (London)* **414**, 656 – 660

636e. Swartz, T. E., Wenzel, P. J., Corchnoy, S. B., Briggs, W. R., and Bogomolni, R. A. (2002) *Biochemistry* **41**, 7183–7189

637. Linden, H., and Macino, G. (1997) *EMBO J.* **16**, 98 – 109

637a. Iseki, M., Matsunaga, S., Murakami, A., Ohno, K., Shiga, K., Yoshida, K., Sugai, M., Takahashi, T., Hori, T., and Watanabe, M. (2002) *Nature (London)* **415**, 1047 – 1048

637b. Froehlich, A. C., Liu, Y., Loros, J. J., and Dunlap, J. C. (2002) *Science* **297**, 815–819

637c. Linden, H. (2002) *Science* **297**, 777–778

637d. Bieszke, J. A., Spudich, E. N., Scott, K. L., Borkovich, K. A., and Spudich, J. L. (1999) *Biochemistry* **38**, 14138 – 14145

References

638. Ahmad, M., and Cashmore, A. R. (1993) *Nature (London)* **366**, 162–166
639. Ahmad, M., Jarillo, J. A., Smirnova, O., and Cashmore, A. R. (1998) *Nature (London)* **392**, 720–723
640. Guo, H., Yang, H., Mockler, T. C., and Lin, C. (1998) *Science* **279**, 1360–1363
640a. Guo, H., Mockler, T., Duong, H., and Lin, C. (2001) *Science* **291**, 487–489
640b. Wang, H., Ma, L.-G., Li, J.-M., Zhao, H.-Y., and Deng, X. W. (2001) *Science* **294**, 151–158
641. Somers, D. E., Devlin, P. F., and Kay, S. A. (1998) *Science* **282**, 1488–1490
642. Ceriani, M. F., Darlington, T. K., Staknis, D., Más, P., Petti, A. A., Weitz, C. J., and Kay, S. A. (1999) *Science* **285**, 553–556
643. Thresher, R. J., Vitaterna, M. H., Miyamoto, Y., Kazantsev, A., Hsu, D. S., Petit, C., Selby, C. P., Dawut, L., Smithies, O., Takahashi, J. S., and Sancar, A. (1998) *Science* **282**, 1490–1494
643a. Hardin, P. E., and Glossop, N. R. J. (1999) *Science* **286**, 2460–2461
644. Hsu, D. S., Zhao, X., Zhao, S., Kazantsev, A., Wang, R.-P., Todo, T., Wei, Y.-F., and Sancar, A. (1996) *Biochemistry* **35**, 13871–13877
645. Lucas, R. J., Freedman, M. S., Munoz, M., Garcia-Fernández, J.-M., and Foster, R. G. (1999) *Science* **284**, 505–507
646. Miyamoto, Y., and Sancar, A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6097–6102
647. Moore, R. Y. (1999) *Science* **284**, 2102–2103
647a. Hattar, S., Liao, H.-W., Takao, M., Berson, D. M., and Yau, K.-W. (2002) *Science* **295**, 1065–1070
647b. Berson, D. M., Dunn, F. A., and Takao, M. (2002) *Science* **295**, 1070–1073
647c. Barinaga, M. (2002) *Science* **295**, 955–957
647d. Thompson, C. L., Blaner, W. S., Van Gelder, R. N., Lai, K., Quadro, L., Colantunoni, V., Gottesman, M. E., and Sancar, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11708–11713
648. Hearst, J. E. (1995) *Science* **268**, 1858–1859
649. Sancar, A. (1994) *Biochemistry* **33**, 2–9
650. Snapka, R. M., and Sutherland, B. M. (1980) *Biochemistry* **19**, 4201–4208
651. Gindt, Y. M., Vollenbroek, E., Westphal, K., Sackett, H., Sancar, A., and Babcock, G. T. (1999) *Biochemistry* **38**, 3857–3866
652. Park, H.-W., Kim, S.-T., Sancar, A., and Deisenhofer, J. (1995) *Science* **268**, 1866–1872
653. Lipman, R. S. A., and Jorns, M. S. (1996) *Biochemistry* **35**, 7968–7973
653a. Durbeej, B., and Eriksson, L. A. (2000) *J. Am. Chem. Soc.* **122**, 10126–10132
653b. MacFarlane, A. W., IV, and Stanley, R. J. (2001) *Biochemistry* **40**, 15203–15214
654. Aubert, C., Mathis, P., Eker, A. P. M., and Brettel, K. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5423–5427
654a. Kay, C. W. M., Feicht, R., Schulz, K., Sadewater, P., Sancar, A., Bacher, A., Möbius, K., Richter, G., and Weber, S. (1999) *Biochemistry* **38**, 16740–16748
654b. Antony, J., Medvedev, D. M., and Stuchebrukhov, A. A. (2000) *J. Am. Chem. Soc.* **122**, 1057–1065
654c. Weber, S., Möbius, K., Richter, G., and Kay, C. W. M. (2001) *J. Am. Chem. Soc.* **123**, 3790–3798
655. Hitomi, K., Kim, S.-T., Iwai, S., Harima, N., Otoshi, E., Ikenaga, M., and Todo, T. (1997) *J. Biol. Chem.* **272**, 32591–32598
655a. Hitomi, K., Nakamura, H., Kim, S.-T., Mizukoshi, T., Ishikawa, T., Iwai, S., and Todo, T. (2001) *J. Biol. Chem.* **276**, 10103–10109
656. Kim, S.-T., Malhotra, K., Smith, C. A., Taylor, J.-S., and Sancar, A. (1994) *J. Biol. Chem.* **269**, 8535–8540
657. Zhao, X., Liu, J., Hsu, D. S., Zhao, S., Taylor, J.-S., and Sancar, A. (1997) *J. Biol. Chem.* **272**, 32580–32590
657a. Joseph, A., Prakash, G., and Falvey, D. E. (2000) *J. Am. Chem. Soc.* **122**, 11219–11225
658. Todo, T., Ryo, H., Yamamoto, K., Toh, H., Inui, T., Ayaki, H., Nomura, T., and Ikenaga, M. (1996) *Science* **272**, 109–112
659. Viviani, V. R., Bechara, E. J. H., and Ohmiya, Y. (1999) *Biochemistry* **38**, 8271–8279
660. Conti, E., Franks, N. P., and Brick, P. (1996) *Structure* **4**, 287–298
661. Branchini, B. R., Magyar, R. A., Murtiashaw, M. H., Anderson, S. M., and Zimmer, M. (1998) *Biochemistry* **37**, 15311–15319
661a. Branchini, B. R., Magyar, R. A., Murtiashaw, M. H., Anderson, S. M., Helgerson, L. C., and Zimmer, M. (1999) *Biochemistry* **38**, 13223–13230
662. McElroy, W. D., and Seliger, H. H. (1962) *Sci. Am.* **207**(Dec), 76–89
663. Campbell, A. K. (1986) *Trends Biochem. Sci.* **11**, 104–108
664. Fisher, A. J., Raushel, F. M., Baldwin, T. O., and Rayment, I. (1995) *Biochemistry* **34**, 6581–6586
665. Bronstein, I., and McGrath, P. (1989) *Nature (London)* **338**, 599–600
666. Adam, W., Bronstein, I., Edwards, B., Engel, T., Reinhardt, D., Schneider, F. W., Trofimov, A. V., and Vasil'ev, R. F. (1996) *J. Am. Chem. Soc.* **118**, 10400–10407
667. Hori, K., Wampler, J. E., Matthews, J. C., and Cormier, M. J. (1973) *Biochemistry* **12**, 4463–4468
668. Matthews, J. C., Hori, K., and Cormier, M. J. (1977) *Biochemistry* **16**, 85–91
669. Shimamura, O., and Johnson, F. H. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2611–2615
670. Tsuji, F. I. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4629–4632
671. Thompson, E. M., Nagata, S., and Tsuji, F. I. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6567–6571
672. Ray, B. D., Ho, S., Kemple, M. D., Prendergast, F. G., and Nageswara Rao, B. D. (1985) *Biochemistry* **24**, 4280–4287
673. Inouye, S., Noguchi, M., Sakaki, Y., Takagi, Y., Miyata, T., Iwanaga, S., Miyata, T., and Tsuji, F. I. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3154–3158
674. Shimomura, O., and Johnson, F. H. (1972) *Biochemistry* **11**, 1602–1608
674a. Markova, S. V., Vysotski, E. S., Blinks, J. R., Burakova, L. P., Wang, B.-C., and Lee, J. (2002) *Biochemistry* **41**, 2227–2236
675. Cormier, M. J., and Wampler, J. E. (1975) *Ann. Rev. Biochem.* **44**, 255–272
676. Hart, R. C., Matthews, J. C., Hori, K., and Cormier, M. J. (1979) *Biochemistry* **18**, 2204–2205
676a. Liu, Z.-J., Vysotski, E. S., Chen, C.-J., Rose, J. P., Lee, J., and Wang, B.-C. (2000) *Protein Sci.* **9**, 2085–2093
676b. Haddock, S. H. D., Rivers, T. J., and Robison, B. H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11148–11151
676c. Li, L., Liu, L., Hong, R., Robertson, D., and Hastings, J. W. (2001) *Biochemistry* **40**, 1844–1849
676d. Nakamura, H., Kishi, Y., Shimomura, O., Morse, D., and Hastings, J. W. (1989) *J. Am. Chem. Soc.* **111**, 7607–7611
676e. Branchini, B. R., Murtiashaw, M. H., Magyar, R. A., and Anderson, S. M. (2000) *Biochemistry* **39**, 5433–5440
677. Thompson, J. F., Geoghegan, K. F., Lloyd, D. B., Lanzetti, A. J., Magyar, R. A., Anderson, S. M., and Branchini, B. R. (1997) *J. Biol. Chem.* **272**, 18766–18771
677a. Branchini, B. R., Murtiashaw, M. H., Magyar, R. A., Portier, N. C., Ruggiero, M. C., and Stroh, J. G. (2002) *J. Am. Chem. Soc.* **124**, 2112–2113
677b. Gomi, K., and Kajiyama, N. (2001) *J. Biol. Chem.* **276**, 36508–36513
678. Shimomura, O., and Johnson, F. H. (1968) *Biochemistry* **7**, 2574–2580
679. Nakamura, H., Kishi, Y., Shimomura, O., Morse, D., and Hastings, J. W. (1989) *J. Am. Chem. Soc.* **111**, 7607–7611
679a. Jeffers, C. E., and Tu, S.-C. (2001) *Biochemistry* **40**, 1749–1754
680. Xin, X., Xi, L., and Tu, S.-C. (1994) *Biochemistry* **33**, 12194–12201
681. Tanner, J. J., Miller, M. D., Wilson, K. S., Tu, S.-C., and Krause, K. L. (1997) *Biochemistry* **36**, 665–672
682. Baldwin, T. O., Berends, T., Bunch, T. A., Holzman, T. F., Rausch, S. K., Shamansky, L., Treat, M. L., and Ziegler, M. M. (1984) *Biochemistry* **23**, 3663–3667
683. Cohn, D. H., Mileham, A. J., Simon, M. I., Nealson, K. H., Rausch, S. K., Bonam, D., and Baldwin, T. O. (1985) *J. Biol. Chem.* **260**, 6139–6146
684. Lawson, D. M., Derewenda, U., Serre, L., Ferri, S., Szittner, R., Wei, Y., Meighen, E. A., and Derewenda, Z. S. (1994) *Biochemistry* **33**, 9382–9388
685. Kurfürst, M., Ghisla, S., and Hastings, J. W. (1983) *Biochemistry* **22**, 1521–152
685a. Lin, L. Y.-C., Sulea, T., Szittner, R., Kor, C., Purisima, E. O., and Meighen, E. A. (2002) *Biochemistry* **41**, 9938–9945
686. Eckstein, J. W., Hastings, J. W., and Ghisla, S. (1993) *Biochemistry* **32**, 404–411
687. Francisco, W. A., Abu-Soud, H. M., DelMonte, A. J., Singleton, D. A., Baldwin, T. O., and Raushel, F. M. (1998) *Biochemistry* **37**, 2596–2606
688. Li, H., Ortego, B. C., Maillard, K. I., Willson, R. C., and Tu, S.-C. (1999) *Biochemistry* **38**, 4409–4415
689. Engebrecht, J., and Silverman, M. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4154–4158
690. Eberhard, A., Burlingame, A. L., Eberhard, C., Kenyon, G. L., Nealson, K. H., and Oppenheimer, N. J. (1981) *Biochemistry* **20**, 2444–2449

Study Questions

1. Why is the Emerson enhancement effect (i.e., light at 650 nm plus 680 nm gives a higher rate of photosynthesis than either one alone) not observed with photosynthetic bacteria?

2. Agents that uncouple oxidative phosphorylation in mitochondria uncouple photoelectron transport and ATP formation in photosynthesis. Explain.

3. The action spectrum of photosynthesis, which describes the efficiency of photosynthesis as a function of the wavelength of incident light, has a valley around 550 nm. Why?

4. Plants exposed to $C^{18}O_2$ will have the ^{18}O first appear in a) carbohydrate; b) water; c) oxygen gas. (More than one answer may be true.)

5. Plants exposed to $H_2{}^{18}O$ will have the label first appear in a) oxygen gas; b) carbohydrate; c) CO_2. (More than one may be correct.)

6. The general equation describing the photosynthesis of glucose in higher plants is:

   $$6\ CO_2 + 6\ H_2O \rightarrow C_6H_{12}O_6$$

   We know that the oxygen gas comes from water, yet the equation shows only six atoms of oxygen in water on the left-hand side vs 12 in oxygen gas on the right. Explain.

7. The fructose bisphosphatase of green plants has an amino acid sequence which is very similar to those of the corresponding enzymes isolated from other sources such as yeast or mammals, except that the plant enzyme has an additional sequence of 20 or so amino acids that has no counterpart in the enzymes found in the other species. What function might this additional sequence have in the plant enzyme?

8. The following substances are either inhibitors or activators of rubisco, the enzyme that catalyzes the condensation of CO_2 with ribulose bisphosphate to yield 3-phosphoglycerate. State whether the substance should be an activator or an inhibitor of the enzyme and succinctly provide the logic supporting your conclusion.
   a) Fructose 1,6-bisphosphate
   b) Visible light
   c) NADPH

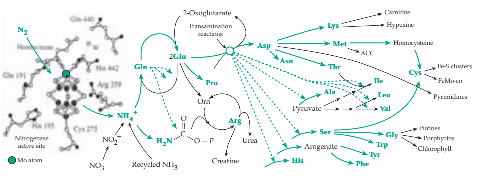
9. The reagent DCMU specifically inhibits electron transfer to plastoquinone in photosystem II. Discuss how the administration of this compound to a suspension of illuminated chloroplasts will affect the production of oxygen, ATP, and NADPH.

10. A chemical reagent is added to a solution of plant chloroplasts which immediately and specifically poisons photosystem II. What is the *short-term* effect of each of the following? Give a one-sentence defense for your conclusion.
    a) Cyclic photophosphorylation
    b) Noncyclic photophosphorylation
    c) Photorespiration
    d) NADPH production

11. If a C_3 and a C_4 plant are placed together in a sealed illuminated box, the C_3 plant withers and dies long before the C_4 plant. Explain.

12. What tricarboxylic acid cycle enzyme is analogous to the malate enzyme of bundle-sheath cells? What is the mechanism of the reaction?

13. There are two different forms of glyceraldehyde-3-phosphate dehydrogenase in higher plant cells.
    a) In which cell compartment is each one found?
    b) What are the reactions catalyzed by these two isozymes?
    c) Why are there two forms?



The air provides an abundant source of nitrogen for living organisms. Nitrogenase present in specialized bacteria utilizes the molybdenum- and iron-containing FeMo-co to reduce N_2 to two molecules of NH_3 (or NH_4^+). NH_3 is incorporated into the side chain of glutamine and much is transferred to the 5-carbon skeleton of 2-oxoglutarate to form glutamate. Nitrogen from glutamate and glutamine moves into the other amino acids via action of transaminases and glutamine amidotransferases. Thousands of compounds, a few of which are indicated here, are formed. (The 20 amino acid constituents of proteins are shown in green.) NH_3 from decaying materials is recycled, often after oxidation to nitrite or nitrate. Nitrates may also be formed by lightening and NH_3 industrially by catalytic reductions of N_2 by H_2 at high temperature and pressure (the Haber process).

Contents

1359 .... **A. Fixation of N_2 and the Nitrogen Cycle**
1359 .... 1. Reduction of Elemental Nitrogen
1360 .... 2. Nitrogenases
1362 .... *The mechanism of nitrogenase action*
1365 .... *Nitrogen fixation genes*
1365 .... *Legume nodules and cyanobacterial heterocysts*
1366 .... *Genetic engineering*
1366 .... *Other nitrogenases*
1366 .... 3. Interconversion of Nitrate, Nitrite, and Ammonium Ions
1367 .... **B. Incorporation of NH_3 into Amino Acids and Proteins**
1369 .... 1. Uptake of Amino Acids by Cells
1369 .... 2. Glutamate Dehydrogenase and Glutamate Synthase
1370 .... 3. Glutamine Synthetase
1371 .... 4. Catabolism of Glutamine, Glutamate, and Other Amino Acids
1371 .... *Catabolism initiated by decarboxylation*
1371 .... *Fermentation of glutamate*
1374 .... **C. Synthesis and Catabolism of Proline, Ornithine, Arginine, and Polyamines**
1374 .... 1. Synthesis and Catabolism of Proline
1374 .... 2. Synthesis of Arginine and Ornithine and the Urea Cycle
1376 .... *The urea cycle*
1376 .... *Carbamoyl phosphate synthetases*
1377 .... *Citrulline and argininosuccinate*
1378 .... *Excretion of ammonia*
1378 .... *Catabolism of arginine*
1379 .... *Insecticidal analogs of arginine*
1379 .... 3. Amidino Transfer and Creatine Synthesis
1380 .... 4. The Polyamines
1381 .... *Biosynthesis*
1382 .... *Breakdown*
1383 .... **D. Compounds Derived from Aspartate**
1383 .... 1. Control of Biosynthetic Reactions of Aspartate
1383 .... 2. Lysine Diaminopimelate, Dipicolinic Acid, and Carnitine
1386 .... 3. The Catabolism of Lysine
1388 .... 4. Metabolism of Homocysteine and Methionine
1389 .... *A salvage pathway*
1389 .... *The plant hormone ethylene*
1391 .... 5. Metabolism of Threonine
1391 .... **E. Alanine and the Branched-Chain Amino Acids**
1393 .... 1. Catabolism
1397 .... 2. Ketogenic and Glucogenic Amino Acids
1397 .... **F. Serine and Glycine**
1397 .... 1. Biosynthetic Pathways from Serine
1398 .... 2. Metabolism of Glycine
1399 .... 3. Porphobilinogen, Porphyrins, and Related Substances
1400 .... *Porphyrins*
1402 .... *Corrins*
1402 .... *Chlorophyll*
1403 .... *The porphyrias*
1404 .... *The bile pigments*
1406 .... **G. Cysteine and Sulfur Metabolism**
1407 .... 1. Synthesis and Catabolism of Cysteine
1407 .... 2. Cysteine Sulfinate and Taurine
1408 .... 3. Mercaptopyruvate, Thiosulfate, and Assembly of Iron–Sulfur Centers

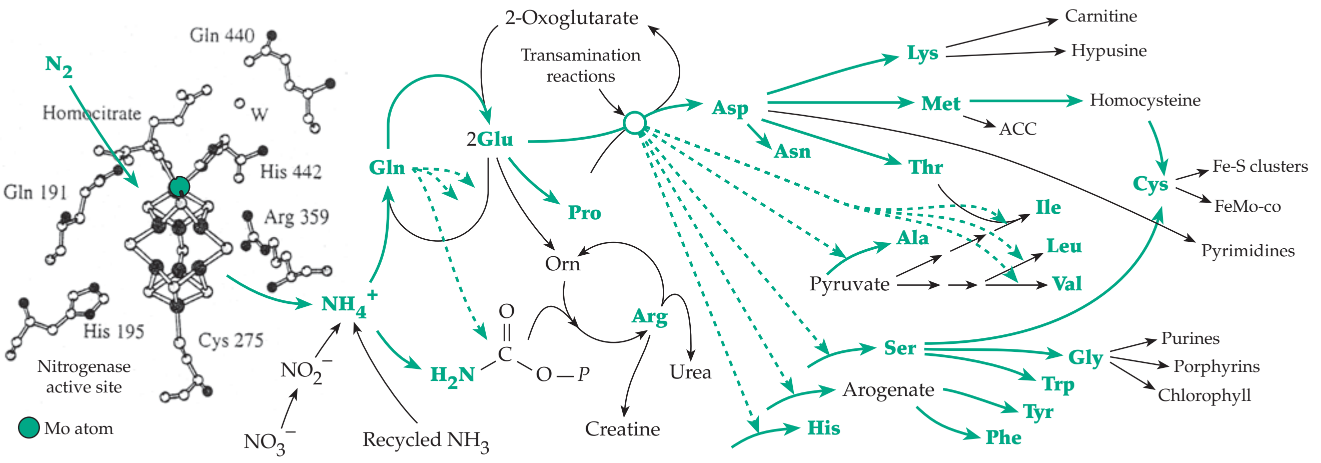
1411 .... **References**
1418 .... **Study Questions**

Boxes

1394 .... Box 24-A Maple Syrup Urine Disease and Jamaican Vomiting Sickness
1408 .... Box 24-B Sulfur Compounds of Garlic, Onions, Skunks, etc.

The Metabolism of Nitrogen and Amino Acids

24


N$_2$
Gln 440
Homocitrate
W
His 442
Gln 191
Arg 359
His 195
Cys 275
Nitrogenase active site
Mo atom
NH$_4^+$
NO$_2^-$
NO$_3^-$
Recycled NH$_3$
Gln
2-Oxoglutarate
Transamination reactions
2Glu
Pro
Orn
Arg
Urea
Creatine
H$_2$N–C(=O)–O–P
Asp
Asn
Thr
Lys
Carnitine
Hypusine
Met
Homocysteine
ACC
Cys
Fe-S clusters
FeMo-co
Pyrimidines
Ile
Ala
Leu
Pyruvate
Val
Ser
Gly
Purines
Porphyrins
Chlorophyll
Arogenate
Trp
Tyr
Phe
His

Because it is found in so many compounds and can exist in several oxidation states, nitrogen has a complex metabolism. The inorganic forms of nitrogen found in our surroundings range from the highly oxidized nitrate ion, in which N has an oxidation state of +5, to ammonia, in which the oxidation state is –3. Living cells both reduce and oxidize these inorganic forms. The organic forms of nitrogen are most often derived by incorporation of **ammonium ions** into amino groups or amide groups. Once it has been incorporated into an organic compound, nitrogen can be transferred into many other carbon compounds. Certain compounds including glutamic acid, aspartic acid, glutamine, asparagine, and carbamoyl phosphate are especially active in these transfer reactions. They constitute a **nitrogen pool** from which nitrogen can be withdrawn and to which it can be returned.

In addition to the pathways for synthesis and degradation of nitrogenous substances, many organisms have specialized metabolism for incorporation of excess nitrogen into relatively nontoxic excretion products. All of these aspects of nitrogen metabolism will be dealt with in this and the following chapter. We will look first at the reactions by which organic nitrogen compounds are formed from inorganic compounds, then at the reactions of the nitrogen pool. After that we will examine the specific reactions of synthesis and catabolism of individual nitrogenous compounds.

A. Fixation of N$_2$ and the Nitrogen Cycle

Most of the nitrogen of the biosphere exists as the unreactive N$_2$, which makes up 80% of the molecules of air. The “fixation” of N$_2$ occurs principally by the action of a group of bacteria known as **diazotrophs** and to a lesser extent by lightning, which forms oxides of nitrogen and eventually nitrate and nitrite. Human beings also contribute a smaller but significant share through production of chemical fertilizer by the Haber process. These reactions are an important part of the **nitrogen cycle**.[1,2] Quantitatively even more important are the biochemical processes of **nitrification**, by which ammonium ions from decaying organic materials are oxidized to NO$_2^-$ and NO$_3^-$ by soil bacteria (Fig. 24-1), and reactions of reduction and **assimilation** of nitrate and nitrite by bacteria, fungi, and green plants. Another reductive process catalyzed by **denitrifying bacteria** returns N$_2$ to the atmosphere (Fig. 24-1).

1. Reduction of Elemental Nitrogen

One of the most remarkable reactions of nitrogen metabolism is the conversion of dinitrogen (N$_2$) to ammonia. It was estimated that in 1974 this biological nitrogen fixation added 17 x 10^{10} kg of nitrogen to the earth (compared with 4 x 10^{10} kg fixed by chemical reactions).[3] The quantitative significance can be more easily appreciated by the realization that one square meter of land planted to nodulated legumes such as soybeans can fix 10–30 g of nitrogen per year.

Fixation of N$_2$ by *Clostridium pasteurianum* and a few other species was recognized by Winogradsky[4] in 1893. Subsequent nutritional studies indicated that both iron and molybdenum were required for the process. Inhibition by CO and N$_2$O was observed. While ammonia was the suggested product, the possibility remained that more oxidized compounds such as hydroxylamine were the ones first incorporated into organic substances. When cell-free preparations capable of fixing nitrogen were obtained in 1960 rapid progress became possible.[5] It was discovered that nitrogen-fixing bacteria are invariably able to reduce acetylene to ethylene, a catalytic ability that goes hand in hand with the ability to reduce N_2. A simple, sensitive **acetylene reduction test** permits easy measurement of the nitrogen-fixing potential of cells.

Application of this test revealed that nitrogen fixation is not restricted to a few species, but is a widespread ability of many prokaryotes. Most studied are *Azotobacter vinelandii*, Winogradsky's *C. pasteurianum*, *Klebsiella pneumoniae* (a close relative of *E. coli*), and several species of *Rhizobium,* the symbiotic bacterium of root nodules of legumes. The latter deserves special attention. Although some free-living rhizobia reduce N_2, the reaction usually takes place only in nodules developed by infected roots. Within these nodules the bacteria degenerate into **bacteroids**; and the special hemoglobin **leghemoglobin**,[6,7] whose sequence is specified by a plant gene,[8] is synthesized.

Legumes are not the only plants with nitrogen-fixing symbionts.[9] Some other angiosperms are hosts to nitrogen-fixing actinomycetes and some gymnosperms contain nitrogen-fixing blue-green algae. Leaf nodules of certain plants infected with *Klebsiella* fix nitrogen. While the nutritional significance is uncertain, nitrogen-fixing strains of *Klebsiella* have also been found in the intestinal tracts of humans in New Guinea. Of the free-living nitrogen-fixing organisms, cyanobacteria appear to be of most importance quantitatively. For example, in rice paddy fields cyanobacteria may fix from 2.4 to 10 g of nitrogen per square meter per year. Cyanobacteria in the oceans fix enormous amounts of nitrogen.[9a]

2. Nitrogenases

Cell-free nitrogenases have been isolated from a number of organisms. These enzymes all share the property of being inactivated by oxygen, a fact that impeded early work. Apparently nitrogen fixation occurs in anaerobic regions of cells. Leghemoglobin may protect the nitrogen-fixing enzymes in root nodules from oxygen. It probably also functions to deliver O_2 by facilitated diffusion to the aerobic mitochondria of the bacteroids at a stable, low partial pressure.[6,10] Some bacteria utilize protective proteins to shield the nitrogenase molecules when the O_2 pressure is too high.[10a]

Nitrogenases catalyze the six-electron reduction of N_2 to ammonia (Eq. 24-1) and are also able to reduce

$$N_2 + 6\,H^+ + 6\,e^- \rightarrow 2\,NH_3 \qquad (24\text{-}1)$$

many other compounds. For example, the reduction of acetylene to ethylene (Eq. 24-2) is a two-electron process. Azide is reduced to N_2 and NH_4^+ in another two-electron reduction (Eq. 24-3). Cyanide ions yield methane and ammonia (Eq. 24-4).[11] Alkyl nitriles as well as N_2O and carbonyl sulfide (COS) are also reduced. Carbon dioxide is reduced slowly to CO[12], and nitrogenases invariably catalyze reduction of protons to H_2 (Eq. 24-5).

$$HC{\equiv}CH + 2\,H^+ + 2\,e^- \rightarrow H_2C{=}CH_2 \qquad (24\text{-}2)$$

$$N{=}N{=}N^- + 4\,H^+ + 2\,e^- \rightarrow N_2 + NH_4^+ \qquad (24\text{-}3)$$

$$^-C{\equiv}N + 8\,H^+ + 6\,e^- \rightarrow CH_4 + NH_4^+ \qquad (24\text{-}4)$$

$$2\,H^+ + 2\,e^- \rightarrow H_2 \qquad (24\text{-}5)$$

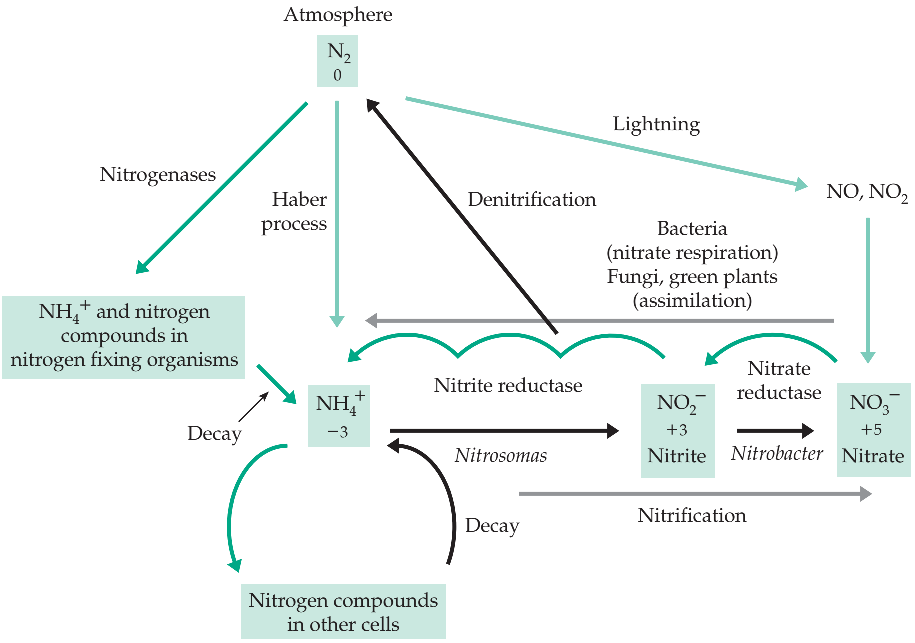


Figure 24-1 The nitrogen cycle. Conversion of N_2 (oxidation state 0) to NH_4^+ by nitrogen-fixing bacteria, assimilation of NH_4^+ by other organisms, decay of organic matter, oxidation of NH_4^+ by the nitrifying bacteria *Nitrosomas* and *Nitrobacter*, reduction of NO_3^- and NO_2^- back to NH_4^+, and release of nitrogen as N_2 by denitrifying bacteria are all part of this complex cycle.[1]

In early experiments it was found that sodium pyruvate was required for fixation of N_2 in cell-free extracts, and that large amounts of CO_2 and H_2 accumulated. Investigation showed that cleavage of pyruvate supplies cells with two important products: ATP and reduced ferredoxin. Pyruvate can be replaced by a mixture of ATP plus Mg^{2+} and reduced ferredoxin (Fd_{red}). Furthermore, the nonbiological reductant dithionite ($S_2O_4^{2-}$) can replace the reduced ferredoxin. Since ADP is inhibitory to the nitrogenase system, it is best in laboratory studies to supply ATP from an ATP-generating system such as a mixture of creatine phosphate, creatine kinase, and a small amount of ADP (Eqs. 6-65, 6-67).

The commonest type of nitrogenase can be separated easily into two components (Fig. 24-2). One of these, the **iron protein** (dinitrogenase reductase, azoferredoxin, or component II), is an extremely oxygen-sensitive iron-sulfur protein. It consists of two identical ~32-kDa peptide chains; those of *A. vinlandii* each contain 189 amino acid residues. The three-dimensional structure of the dimeric protein[13–16] shows that each subunit forms a nucleotide-binding domain with an ATP-binding site. About 2 nm away from this site is *a single Fe_4S_4 cluster which is shared symmetrically by the two subunits of the protein*. Each subunit contributes two thiolate groups from Cys 97 and Cys 132 as well as three N–H- -S hydrogen bonds from NH groups at helix ends.[13]

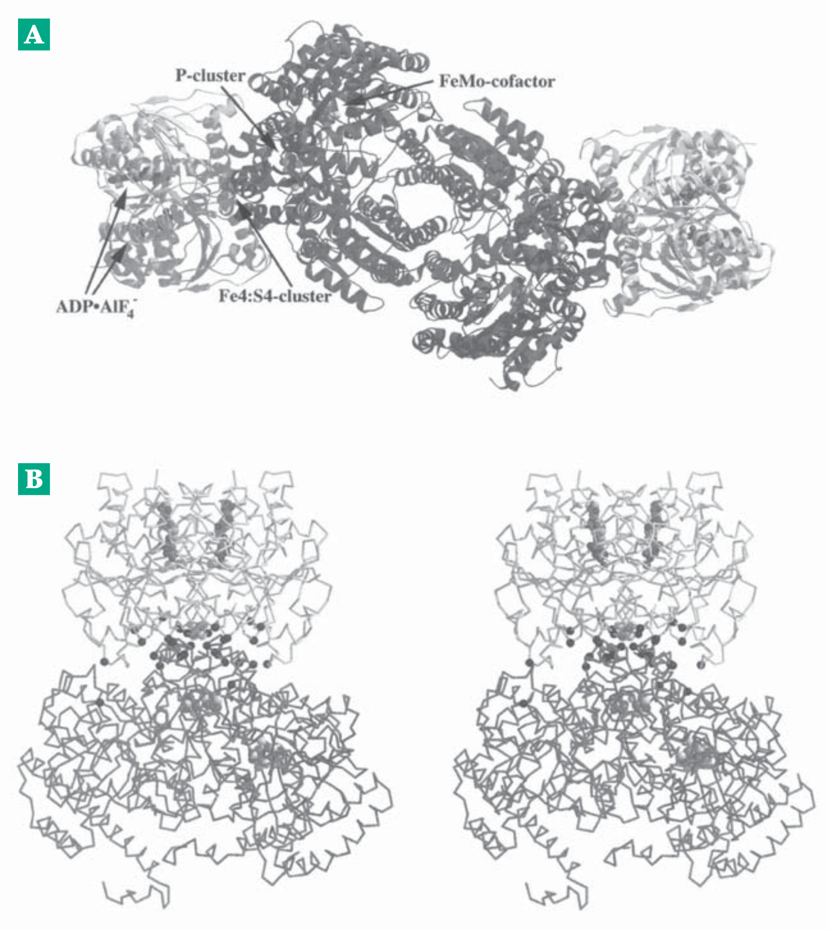
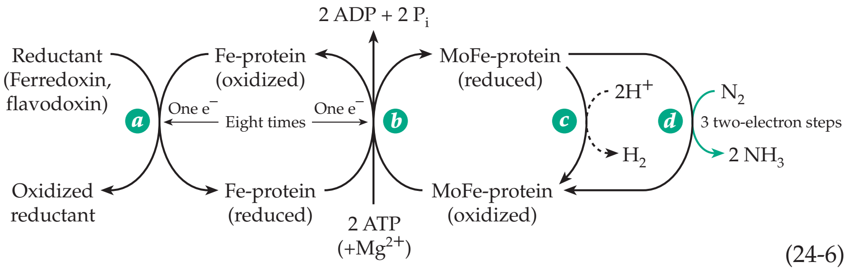


Figure 24-2 (A) Ribbon drawing of the three-dimensional structure of the nitrogenase from *Azotobacter vinlandii* viewed down the twofold axis of the molecule. The αβ subunits of the MoFe-protein are in the center while the Fe-protein subunits are at the outer ends. The Fe_4S_4 cluster of the Fe-protein and the FeMo-co and P-cluster of the MoFe-protein are marked for the left-hand complex. The site of binding of ATP is identified by the bound ADP•AlF_4^- complex. (B) Stereoscopic view of one complete half of the nitrogenase complex at a 90° angle to the view in (A). The Fe-protein is at the top, and the MoFe-protein is below. The ADP•AlF_4^- complex is visible in the two symmetrically located binding sites of the Fe-protein. The shared Fe_4S_4 cluster is in the center above the P-cluster. The small black spheres mark α-carbons of residues that interact in forming the complex of Fe-protein and MoFe-proteins. When the Fe_4S_4 cluster accepts one electron from a molecule of ferredoxin or flavodoxin, the Fe-protein binds to the MoFe-protein and donates an electron to one of the two nearby P-clusters (one of which is in each of the αβ subunits). At the same time both of the molecules of ATP bound to the Fe-protein are hydrolyzed. The oxidized Fe-protein then dissociates from the complex and is replaced by another reduced Fe-protein–ATP complex. The net result is that each electron is "pumped" from the Fe_4S_4 cluster of the Fe-protein into a P-cluster of one of the αβ units of the MoFe-protein. Electrons then move from the P-clusters into the FeMo-coenzyme. From Schindelin *et al.*[19] Courtesy of Douglas C. Rees.

The other component, the **molybdenum–iron protein** (dinitrogenase, molybdoferredoxin, or component I), contains both iron and molybdenum as well as labile sulfide. It is a mixed ($\alpha_2\beta_2$) tetramer of ~240-kDa mass and an analytical metal ion composition ~$Mo_2Fe_{30}S_{26}$. However, the X-ray structure[16–19a] suggests the composition $Mo_2Fe_{34}S_{36}$. The MoFe protein is a symmetric molecule in which each αβ subunit contains two types of complex metal clusters. The active sites for N_2 reduction, which are embedded in the α subunits, contain the **FeMo-coenzyme** molecules, each with the metal composition $MoFe_7S_9$ and also containing a molecule of **homocitrate**.[20–22a] The other clusters, known as **P-clusters**, are shared between the α and β subunits, which for *A. vinlandii* contain 491 and 522 amino acid residues, respectively. Each P-cluster is actually a *joined pair* of cubane-type clusters, one Fe_4S_4 and one Fe_4S_3 with two bridging cysteine –SH groups and one iron atom bonded to three sulfide sulfur atoms (Fig. 24-3).[17,23] The FeMo-coenzyme can be released from the MoFe-protein by acid denaturation followed by extraction with dimethylformamide.[24] While homocitrate was identified as a component of the isolated coenzyme, the three-dimensional structure of FeMo-co was deduced from X-ray crystallography of the intact molybdenum–iron protein.[14,17,18]



(24-6)

When the Fe-protein is reduced an EPR signal at $g = 1.94$, typical of iron–sulfur proteins (Fig. 16-17), is observed. This signal is altered by interaction with Mg and ATP, whereas ATP has no effect on the complex EPR signals produced upon reduction of MoFd. These are among the observations that led to the concept that the Fe-protein is an electron carrier responsible for reduction of the molybdenum in the MoFe-protein. The Mo(IV) or two atoms of Mo(III) formed in this way could then reduce N_2 in three two-electron steps with formation of Mo(VI) (Eq. 24-6). Three successive two-electron steps are required to completely reduce N_2 to two molecules of ammonia. An unexpected feature of nitrogenase action is that there is inevitably what was once regarded as a side reaction, the reduction of protons to H_2.[25,26] The amount of H_2 formed is variable and may be much greater than that of N_2 reduced. However, at high pressures of N_2 the ratio of H_2 formed to N_2 reduced is 1:1. This led to the suggestion[28] that H_2 formation is not a side reaction but an essential step in preparing the active site for the binding of N_2. Two protons that are bound somewhere on the reduced MoFe-protein could be reduced to H_2 in an obligatory step (Eq. 24-6, step *c*) that would, for example, cause a conformational change required for binding of N_2. If no reducible substrate (N_2, C_2H_2, etc.) is present, H_2 would still be formed slowly. Reducible substrates inhibit H_2 formation. However, addition of N_2 or any other reducible substrate causes an initial "burst" of H_2 to be released. This can be measured readily when a slow substrate such as CN^- is used as the inhibitor. The amount of H_2 released in the burst is stoichiometric with one H_2 per Mo being formed.[26] The overall stoichiometry for reduction of one N_2 becomes:

$$N_2 + 8\,e^- + 8\,H^+ + 16\,MgATP \rightarrow 2\,NH_3 + H_2 + 16\,MgADP + 16\,P_i \tag{24-7}$$

A second remarkable feature of nitrogenase is a requirement for hydrolysis of MgATP that is coupled to reduction of the MoFe-protein (Eq. 24-6). Two molecules of ATP are hydrolyzed to ADP and inorganic phosphate for each electron transferred. This large ATP requirement seems surprising in view of the fact that reduction of N_2 by reduced ferredoxin (Eq. 24-8) is thermodynamically spontaneous:

$$N_2 + 6\,Fd_{red} + 8\,H^+ \rightarrow 2\,NH_4^+ + 6\,Fd_{ox}^+$$
$$\Delta G'\ (pH\ 7) = -89.3\ kJ\ mol^{-1} \tag{24-8}$$

However, N_2 is exceedingly unreactive. In the commercial Haber process high pressure and temperature are needed to cause H_2 and N_2 to combine. Evidently cleavage of 16 molecules of ATP must be coupled to the nitrogenase reduction system to overcome the very high activation energy.

Not only are two molecules of ATP hydrolyzed to pump each electron, but the Fe-protein must receive electrons from a powerful (low $E^{\circ\prime}$) reductant such as reduced ferredoxin, reduced flavodoxin, or dithionite. *Klebsiella pneumoniae* contains a **pyruvate:flavodoxin oxidoreductase** (Eq. 15-35) that reduces either flavodoxin or ferredoxin to provide the low potential electron donor.[29,30] In some bacteria, e.g., the strictly aerobic *Azotobacter*, NADPH is the electron donor for reduction of N_2. The Fe-protein is thought to accept electrons from a chain that includes at least the ordinary bacterial ferredoxin (Fd) and a special one-electron-accepting **azotoflavin,** a flavoprotein that is somewhat larger than the flavodoxins (Chapter 15) and appears to play a specific role in N_2 fixation.[31] In *Clostridium* and *Rhizobium* reduced ferredoxins generated by cleavage of pyruvate reduce nitrogenase directly.[32]

The mechanism of nitrogenase action. The one-electron reduction of the Fe_4S_4 cluster of the Fe-protein (step *a* of Eq. 24-6) initiates the action. This reaction occurs before the Fe-protein forms a complex with the MoFe-protein. Following this initial reduction step the two molecules of ATP required for step *b* of Eq. 24-6 bind to the Fe-protein. One is bound to each subunit of this protein but neither is immediately adjacent to the shared Fe_4S_4 cluster, as can be seen from Fig. 14-3B. The binding to MgATP appears to induce a conformational change that permits the "docking" of the Fe-protein with the MoFe-protein to

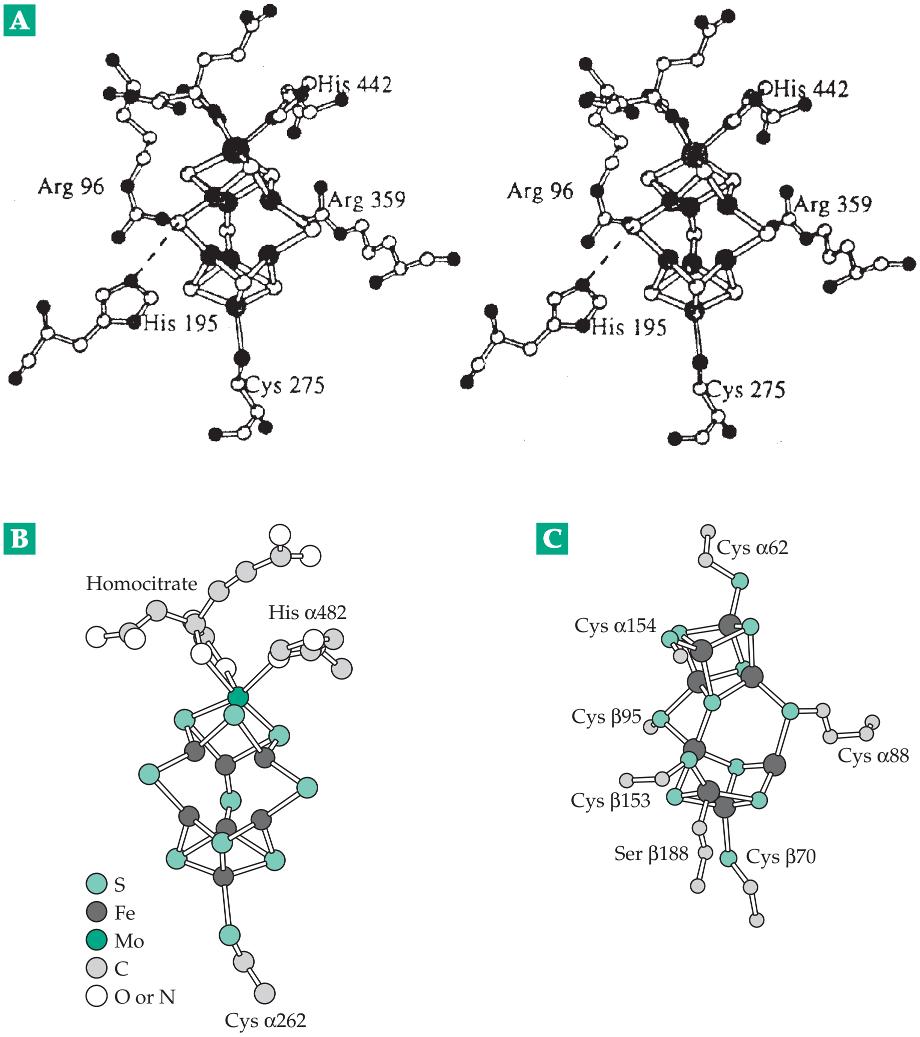


Figure 24-3 Structures of the metal-sulfide clusters of the MoFe-protein. (A) Stereoscopic view of the FeMo-co coenzyme with interacting side chains from the MoFe-protein of *A. vinlandii*. After Kim *et al.*[27] (B) FeMo-co with atom labels. From Kim *et al.*[18] (C) The structure of the oxidized form of the P-cluster. From Peters *et al.*[23] Recent studies by Einsle *et al.*[22a] indicate that the cluster probably also contains a nitrogen atom that is held within the cluster by coordination to six of the iron atoms.

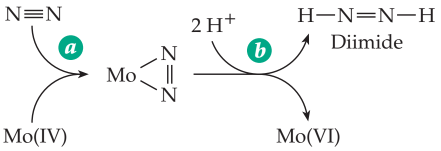
form the complex in which the electron transfer of step *b* (Eq. 24-6) occurs. Abundant evidence indicates that electron transfer does not occur without the binding of MgATP.[33–35d] The electron transfer is coupled to the hydrolysis of the ATP, but the two reactions appear to be consecutive events. In a deletion mutant of the Fe-protein (lacking Leu 127) the hydrolysis of ATP does not occur, but the complex between Fe-protein and MoFe-protein is formed and electron transfer to the MoFe-protein takes place.[36–37a] The binding of the MgATP causes the midpoint redox potential to drop from –0.42 V to –0.62 V, assisting the transfer.[38] X-ray crystallographic studies reveal a distinct conformational change similar to those observed with G-proteins (Chapter 11) and involving movement of the Fe_4S_4 center into a better position for electron transfer.[19,38a]

After electron transfer the complex of the two proteins is thought to be tightly bonded when unhydrolyzed ATP is present.[35] This has allowed the direct observation and imaging of the complex at low resolution (~1.5 nm) using rapid synchrotron X-ray scattering measurements.[39] The ATP is hydrolyzed, and the Fe-proton is released from the complex.

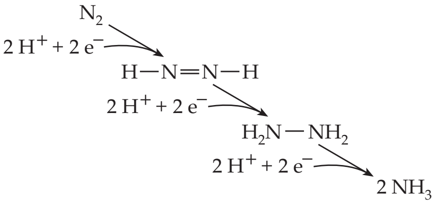
Only one electron is transferred to the MoFe-protein in each catalytic cycle of the Fe-protein. Thus, the cycle must be repeated eight times to accomplish the reduction of $N_2 + 2\ H^+$. Where in the MoFe-protein does a transferred electron go? EPR spectroscopic and other experiments with incomplete and catalytically inactive molybdenum coenzyme[40] have provided a clear answer. The electron is transferred first to one of the two P-clusters, both of which are close to the Fe_4S_4 cluster of the Fe-protein. The transfer causes an observable change both in the spectroscopic properties and in the three-dimensional structure of the P-cluster.[23,40a] Since protons are needed at the active site for the reduction reactions (the FeMo-coenzyme), it is probable that hydrolysis of ATP in the Fe-protein is accompanied by transport of protons across the interface with the MoFe-protein. The electron transfer from the P-cluster on to the FeMo-co center would be assisted by a protic force resulting from ATP cleavage.

With defective FeMo-co (apparently lacking homocitrate) no reduction of N_2, acetylene, or protons is observed.[40] If intact FeMo-co is present, reduction of the cofactor can be observed. An $s = 3/2$ EPR signal arising from the Mo is seen,[40] and EXAFS measurements reveal decreased Mo–Fe distances as the coenzyme is reduced.[41] The molybdenum is probably present as Mo(VI) in the oxidized state of nitrogenase,[42] but after reduction it isn't clear whether it is Mo(III) or Mo(IV). Isolated FeMo-co exists in three identified oxidation states related by $E°'$ values of −0.17 and −0.465 V.[43] Only the middle state is EPR-active, but it is the most reduced state that is involved in N_2 reduction.[42,43] With its P-cluster and FeMo-co center each αβ unit of the MoFe-protein could store several electrons. Two or more might be stored in a P-cluster, and Mo(VI) could, in principle, accept three electrons to form Mo(III). However, it is a little hard to imagine storage of the eight electrons needed to reduce both N_2 and H_2 (Eq. 24-8). The reduction of N_2 may begin before all eight electrons have been transferred into the MoFe-protein.

Another uncertainty lies in the mode of binding of N_2 and other substrates. Does N_2 bind end-on to Mo, does it slide between Fe atoms within the coenzyme, or does it bind in some other way? While N_2 is unreactive, it forms nitrides with metals and complexes with some metal chelates. These complexes are generally of an end-on nature, e.g., N≡N–Fe. Stiefel suggested that N_2 first forms a complex of this type with an iron atom of the MoFe-protein.[44] Then an atom of Mo(IV) could donate two electrons to the N_2 (Eq. 24-9, step *a*) to form a complex of N_2 and Mo(VI). Addition of two protons (Eq. 24-9, step *b*) would yield a molecule of **diimide**, which would stay bound at the iron site while the molybdenum underwent another round of reduction. The diimide could be reduced to hydrazine and finally to ammonia (Eq. 24-10):

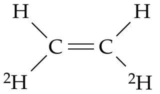


(24-9)



(24-10)

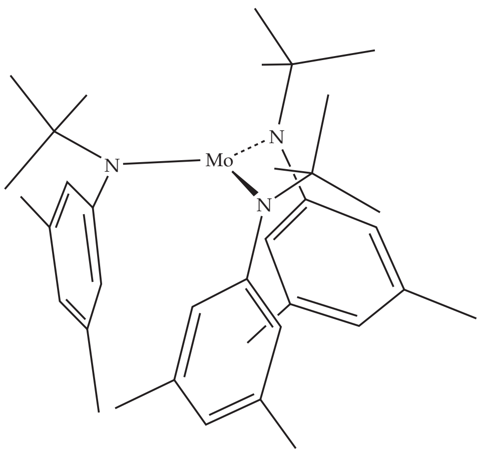
Mo(VI) attracts electrons sufficiently strongly that protons bound to surrounding ligands, such as H_2O, tend to dissociate completely. Thus, the molybdate ion MoO_4^{2-} is not protonated. The same would be true of nitrogenous ligands of a protein that might be coordinated with the bound molybdenum. On the other hand, reduction to Mo(IV) would tend to favor protonation of ligands such as the His 442 imidazole seen in Fig. 24-3A. Concurrently with the electron transfer from molybdenum to N_2 these protons could be transferred to the N_2 molecule (Eq. 24-10). The fact that strictly *cis*-dideuteroethylene is formed from acetylene in the presence of ^2H_2O is in accord with this idea.



However, looking at the FeMo-co molecule and the crowded surroundings of the Mo atom it may be more likely that reduction of N_2 occurs while it is bound to iron. Theoretical calculations as well as experimental data support this possibility.[44a] Recent crystallographic studies at a resolution of 0.12 nm revealed the presence of an atom, probably N, coordinated to six Fe atoms of FeMo-Co. This suggets, as previously proposed by Thorneley and Lowe,[44b,c] that a nitride ion (N_3^-) may be an intermediate in the formation of N_2.

Many mutant forms of nitrogenase have been investigated. Substitutions of His 195, Lys 191, and Gly 69 of the α chain affect reactions with various substrates.[45–45d] For example, the mutant obtained by substitution of His 195, whose imidazole forms an N–H- -S hydrogen bond to a central bridging sulfide atom of FeMo-co (Fig. 24-3A), with glutamine (H195Q mutant) reduces N_2 only very slowly.[45] However, it still reduces both acetylene and protons.[27,44a,45b] Thus, it may be that different modes of substrate binding are needed for the individual steps of Eq. 24-10.

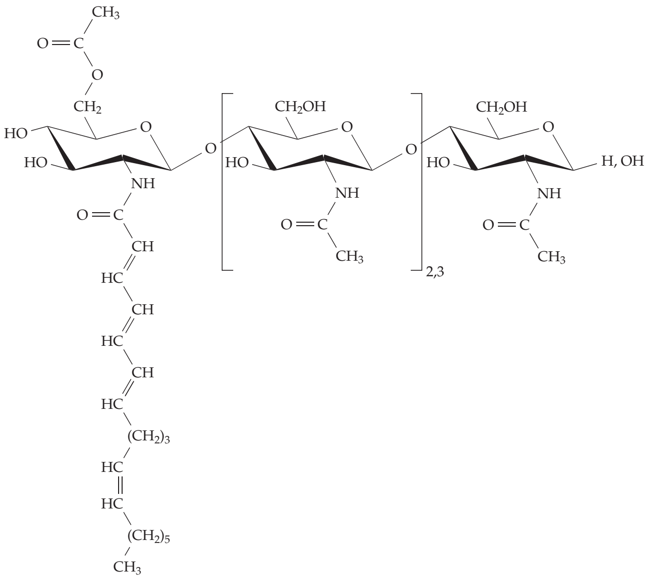
Because of the practical significance to agriculture there is interest in devising better nonenzymatic processes for fixing nitrogen using nitrogenase models that mimic the natural biological reaction.[42,46–49a] One interesting catalyst is the following molybdenum complex $Mo(III)(NRAr)_3$ where $R = C(C_2H_3)_2CH_3$ and Ar = 3,5-dimethylphenyl.[47,50,51]



Many other synthetic complexes have been studied including cubic $MoFe_3S_4$ clusters.[46] However, no exact chemical model for the FeMo-co coenzyme has been developed, and the rates of reaction for all of the model reactions are much slower than those of nitrogenases.

Nitrogen fixation genes. At least 17 genes needed for nitrogen fixation are present in the 23-kb *nif* region of the *Klebsiella* chromosome[52,53] (Fig. 24-4). A similar gene cluster in *A. vinlandii*[54] contains five polygenic transcriptional units and one monogenic unit. The nitrogenase structural genes are *nifK, D*, and *H* as is indicated in Fig. 24-4. The *nifF* and *J* genes encode associated electron-transport proteins. *NifM* is needed to activate the Fe-protein in an unknown fashion. *NifS* encodes a cysteine desulfhydrase needed for assembly of Fe–S clusters in the nitrogenase and the nifU and nifY proteins assist the assembly.[54a,b] The chaperone GroEL is also required.[54c] *Nif Q, B, V, X, N, E_1*, and *H* are needed for synthesis of FeMo-co and for its incorporation into the MoFe-protein.[55,55a] *NifA* is an activator gene for the whole cluster including the *nifL* gene product, which is altered by the presence of O_2 or of glutamine. Accumulation of the latter in cells (see Section B,2) strongly represses transcription of the nitrogenase genes.

Legume nodules and cyanobacterial heterocysts. Nitrogen fixation requires an anaerobic environment. Free-living bacteria fix nitrogen only when anaerobic. However, Rhizobia produce their own anaerobic environment by symbiotic association with the roots of legumes.[10,57–59] Formation of root nodules is a genetically determined process, several nodulation (*nod*) genes of the bacterium being required along with an unknown number of plant genes.[57,58] Initiation of nodulation results from a two-way molecular conversation between root hairs of the plant and bacterial cells.[60,61] The roots secrete **flavonoid compounds** (Chapter 21) which are recognized by bacterial sensors and induce transcription of the *nod* genes. Several of these genes encode enzymes required for synthesis of **Nod factors**,[62,63] small β-linked *N*-acetyl-D-glucosamine oligosaccharides containing 3–5 sugar residues and an *N*-linked long-chain fatty-acyl substituent at the nonreducing terminus (**lipochitooligosaccharides**). See also Box 20-E. Genes *nodA, B, C* specify enzymes needed for synthesis of the oligosaccharide core present in all Nod factors.



Structure of a Nod factor secreted by *Rhizobium leguminosarum*.[64]

Other Nod genes provide for modifications that restrict infection to specific species of legumes. For example, *nodS* encodes a methyltransferase and *nodU* a carbomoyltransferase.[65] NodH is a sulfotransferase.[66] NodD is a transcriptional activator that binds to DNA and induces the synthesis of the other Nod factors needed to initiate nodulation.[67] When an appropriate Nod factor is recognized, the root hairs on the legume

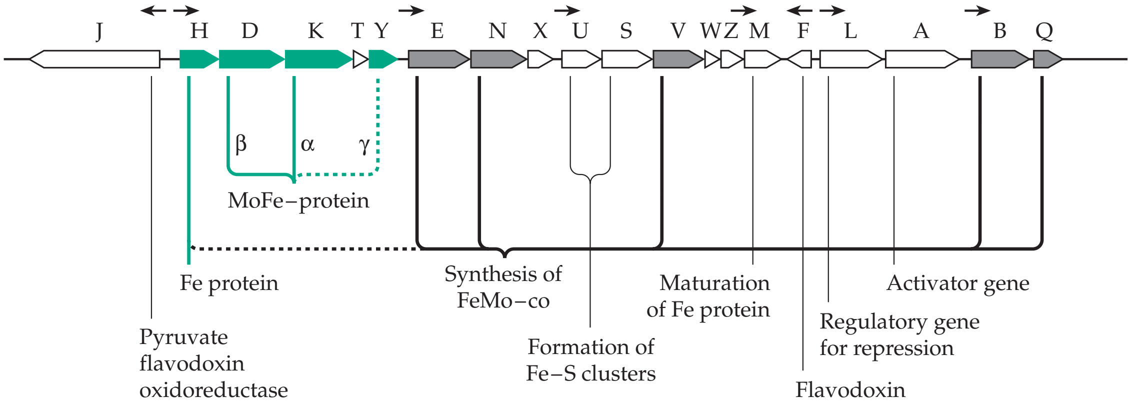


Figure 24-4 Sequence of *nif* genes of *Klebsiella pneumoniae*.[56] These precede the *his* operon directly at the right side. The nitrogenase structural genes are marked with green.

curl around the bacteria to initiate nodulation.[60,61] However, there are other factors. Infecting bacteria must reach a region of low oxygen in the plant. A hemoprotein **FixL** is a sensor kinase that regulates phosphorylation of transcription factor **FixJ**. This two-component system induces transcription of *nifA* (Fig. 24-4) and others.[68–69a] Nitrogen-fixing nodules, which are filled with the bacteroids derived from the infecting bacteria, synthesize leghemoglobin. The polypeptide chain of this protein is encoded by the plant, but its heme may be synthesized by bacteroid enzymes.[10,57] In at least one strain of *Rhizobium* the *nod* genes as well as the *fix* and *nif* genes are all carried on a 536-kb plasmid, which is almost as large as the whole 580-kb genome of *Mycoplasma genitalium* (Table 1-3).[70,71] This arrangement seems to have allowed these bacteria to form an unusually large number of Nod factors and to colonize a wider variety of hosts including a non-leguminous tree.

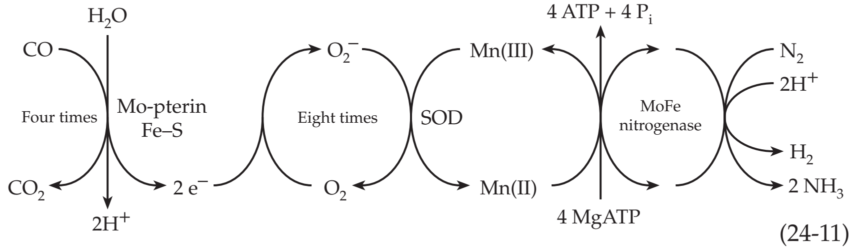
The H_2 that is produced in Eq. 24-6 (step *c*) may be used by bacteroids or by the plant cells. Some nodules evolve H_2, but in others it is utilized by hydrogenases as a source of energy.[57] From Eq. 24-6 it can be seen that up to 1/4 of the ATP utilized can, ideally, be recovered by use of the H_2 in this manner.

In cyanobacteria nitrogen fixation occurs in the **heterocysts**, specialized cells with thickened cell envelopes. They supply NH_4^+ to other cells in the filament of which they are a part. The cell envelopes prevent rapid diffusion of O_2 into the cells but do permit rapid enough entry of N_2 to maintain the observed rate of fixation of N_2.[72] In actinomycetes of the genus *Frankia*, which forms root nodules with woody plants, nitrogen fixation occurs in vesicles that are sheathed by multiple layers of **hopanoid lipids** (see Chapter 22).[73]

Genetic engineering. Because of the high cost of nitrogen fertilizers there is intense interest in improving biological nitrogen fixation. Ideas range from increasing the efficiency of nitrogenase by using fewer molecules of ATP, by limiting excessive evolution of H_2, or transferring the whole *nif* region of a bacterial genome into nonleguminous plants. The last proposal has generated much publicity, but it will probably be difficult because of the need to create an anaerobic environment suitable for nitrogen fixation. A crop plant engineered in this way might not resemble the hoped-for product. It would have an enormous energy requirement for nitrogen fixation, which would have to be met by photosynthesis. At present genetic engineering on *nif* genes to increase efficiency seems most likely to succeed.

Other nitrogenases. Although the well-characterized Mo-containing nitrogenase is responsible for most of their nitrogen fixation, bacteria often have alternative nitrogen fixation systems.[74] *Azotobacter vinlandii* produces three different nitrogenases in response to varying metal compositions in its surroundings.[75–76a] When the molybdenum level is adequate nitrogenase 1 is formed with its FeMo-coenzyme. In a low-molybdenum environment containing vanadium nitrogenase 2 is formed with an FeV-coenzyme.[75–77] If both molybdenum and vanadium are lacking, the bacteria form nitrogenase-3, which has an iron-only FeFe-coenzyme.

An unusual nitrogenase is formed by the chemolithotrophic *Streptomyces thermoautotrophicus*, which obtains energy from reduction of CO_2 or CO by H_2 (Eq. 17-50). These organisms form a MoFe nitrogenase that utilizes a manganese-containing superoxide dismutase to generate superoxide anion radicals. The latter transfer electrons to the MoFe protein in an ATP-dependent process. Electrons for generation of superoxide are formed using another molybdenum enzyme, a CO dehydrogenase containing molybdpterin cytosine dinucleotide (Fig. 16-31) and Fe–S centers.[78] The two systems function together as indicated by Eq. 24-11.



(24-11)

3. Interconversion of Nitrate, Nitrite, and Ammonium Ions

As is indicated in Fig. 24-1, the interconversions of nitrate and nitrite with ammonia and with organic nitrogen compounds are active biological processes. Two genera of nitrifying soil bacteria, which are discussed in Chapter 18, oxidize ammonium ions to nitrate. *Nitrosomas* carries out the six-electron oxidation to nitrite (Eq. 18-17) and *Nitrobacter* the two-electron oxidation of nitrite to nitrate (Eq. 18-18).[79]

The opposite sequence, reduction of nitrate and nitrite ions, provides a major route of acquisition of ammonia for incorporation into cells by bacteria, fungi, and green plants (Fig. 24-1). **Assimilatory** (biosynthetic) **nitrate reductases** catalyze the two-electron reduction of nitrate to nitrite (Eq. 16-61). This is thought to occur at the molybdenum atom of the large ~900-residue highly regulated[79a] molybdopterin-dependent enzyme. In green plants the reductant is usually NADH while in fungi it is more often NADPH.[80–82] In all cases the cofactors FAD, heme, and molybdopterin are bound to a single polypeptide chain with the molybdopterin domain near the N terminus, and the heme in the middle. The electron-accepting FAD domain is near the C terminus[83–85a] and is thought to transfer the two electrons through the following chain.

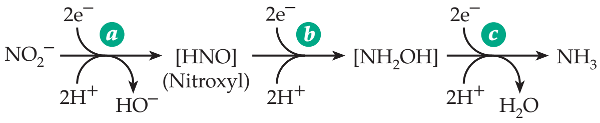
$$\text{NAD(P)H} \rightarrow \text{FAD} \rightarrow \text{Cyt } b \rightarrow \text{Mo-pterin} \rightarrow NO_3^- \qquad (24\text{-}12)$$

Bacterial assimilatory nitrate reductases have similar properties.[86,86a] In addition, many bacteria, including *E. coli*, are able to use nitrate ions as an oxidant for **nitrate respiration** under anaerobic conditions (Chapter 18). The **dissimilatory nitrate reductases** involved also contain molybdenum as well as Fe–S centers.[85] The *E. coli* enzyme receives electrons from reduced quinones in the plasma membrane, passing them through cytochrome *b*, Fe–S centers, and molybdopterin to nitrate. The three-subunit $\alpha\beta\gamma$ enzyme contains cytochrome *b* in one subunit, an Fe_3S_4 center as well as three Fe_4S_4 clusters in another, and the molybdenum cofactor in the third.[87] Nitrate reduction to nitrite is also on the pathway of denitrification, which can lead to release of nitrogen as NO, N_2O, and N_2 by the action of **dissimilatory nitrite reductases**. These enzymes[87a] have been discussed in Chapters 16 and 18.

Assimilatory nitrite reductases of plants, fungi, and bacteria carry out the six-electron reduction of nitrite to ammonium ions (Eq. 24-13) using electron donors such as reduced ferredoxins or NADPH.

$$NO_2^- + 6\,e^- + 8\,H^+ \rightarrow NH_4^+ + 2\,H_2O \qquad (24\text{-}13)$$

The enzymes from green plants and fungi are large multifunctional proteins,[80] which may resemble assimilatory sulfite reductases (Fig. 16-19). These contain **siroheme** (Fig. 16-6), which accepts electrons from either reduced ferredoxin (in photosynthetic organisms) or from NADH or NADPH. FAD acts as an intermediate carrier. It seems likely that the nitrite N binds to Fe of the siroheme and remains there during the entire six-electron reduction to NH_3. Nitroxyl (NOH) and hydroxylamine (NH_2OH) may be bound intermediates as is suggested in steps *a–c* of Eq. 24-14.



$$NO_2^- \xrightarrow[2H^+ \quad HO^-]{2e^- \quad a} \underset{\text{(Nitroxyl)}}{[HNO]} \xrightarrow[2H^+]{2e^- \quad b} [NH_2OH] \xrightarrow[2H^+ \quad H_2O]{2e^- \quad c} NH_3$$

(24-14)

B. Incorporation of NH_3 into Amino Acids and Proteins

Prior to 1935, amino acids were generally regarded as relatively stable nutrient building blocks. That concept was abandoned as a result of studies of the metabolism of $^{15}NH_3$ and of ^{15}N-containing amino acids by Schoenheimer and Rittenberg[88] and more recent studies using ^{13}N by Cooper *et al.*[89,90] These investigations showed that nitrogen could often be shifted rapidly between one carbon skeleton and another. This confirmed proposals put forth earlier by Braunstein, Meister, and others who had pointed out that the C_4 and C_5 amino acids, aspartate and glutamate, which are closely related to the tricarboxylic acid cycle, are able to exchange their amino groups rapidly with those of other amino acids via transamination (Fig. 24-5, step *d*). Since ammonia can be incorporated readily into glutamate (Fig. 24-5, step *a*; see next section), a general means is available for the biosynthesis of amino acids. The citric acid cycle is able to provide any needed amount of 2-oxoglutarate for the synthesis of both glutamate and glutamine.[91–94]

Glutamine, and to a lesser extent asparagine, act as soluble, nontoxic carriers of additional ammonia in the form of their amide groups. An active synthase converts glutamate and ammonia to glutamine (Fig. 24-5, step *d*), and another enzyme transfers the amide nitrogen into aspartate, in an ATP-dependent reaction, to form asparagine (Fig. 24-5, step *e*). The amide nitrogen of glutamine is incorporated in a similar way into a great variety of other biochemical compounds, including carbamoyl phosphate (Fig. 24-5, step *f*; Section C,2), glucosamine (Eq. 20-5), NAD^+, *p*-aminobenzoate,

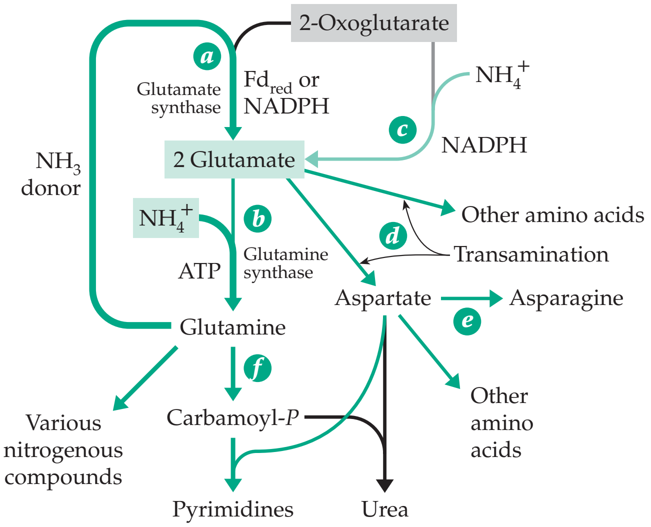


Figure 24-5 Major pathways of incorporation of nitrogen from ammonium ions into organic compounds, traced by green arrows.

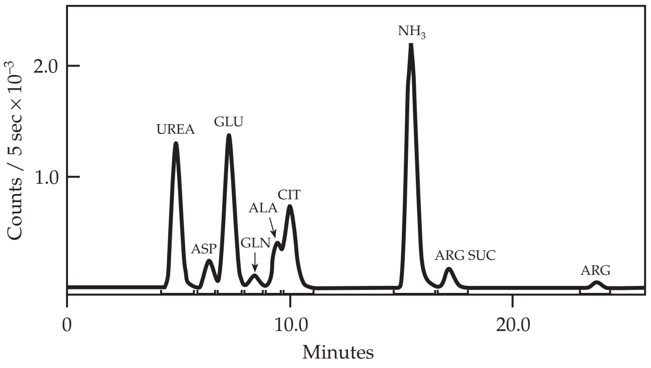
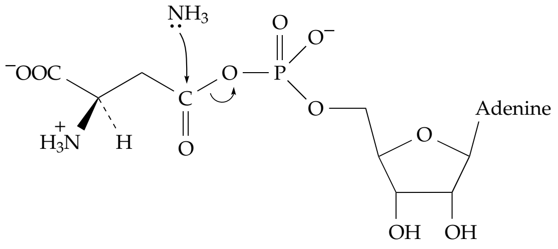


Figure 24-6 Elution profile of ^{13}N-containing metabolites extracted from liver 15 s after injection of $^{13}NH_3$ into the portal vein of an anesthetized adult male rat. CIT, citrulline; ARG SUC, argininosuccinate. From Cooper *et al.*[89]

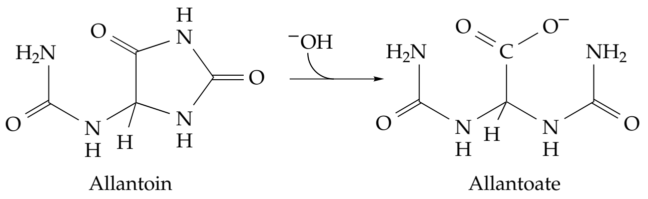
histidine, CTP, and purines (Chapter 25). These reactions are catalyzed by a family of **amidotransferases**,[95–97d] which hydrolyze the glutamine to glutamate and NH_3. The last is entrapped until it reacts with the second substrate. **Asparagine synthetase** apparently first binds ATP and aspartate, which probably react to form β-aspartyl adenylate (β-aspartyl-AMP). Glutamine then binds and is hydrolyzed.[98,98a]



β-Aspartyl adenylate

The liberated NH_3 can attack the β-aspartyl-AMP as indicated in the accompanying diagram to form asparagine. However, it is also possible that NH_3 is transferred via covalently bonded complexes[99] and is never free NH_3. An asparagine synthetase that utilizes free NH_3 as a nitrogen donor is also present in many organisms.

A third mechanism of synthesis, which was only recently recognized, appears to provide the sole source of asparagine for many bacteria.[98b] The asparagine-specific transfer RNA $tRNA^{Asn}$ is "mischarged" with aspartic acid to form $Asp\text{-}tRNA^{Asn}$. This compound is then converted to the properly aminoacylated $Asn\text{-}tRNA^{Asn}$ by a glutamine-dependent amidotransferase. (The entire ATP-dependent sequence is shown in Eq. 29-6.) The activated asparaginyl group is then transferred from $Asn\text{-}tRNA^{Asn}$ into proteins as they are synthesized.



(24-15)

Most green plants transport nitrogen from roots to growing shoots as asparagine. However, in peanuts **β-methyleneaspartate** is the major nitrogen carrier,[100] and in some legumes, including soybeans, **allantoin** and **allantoate** (Eq. 24-15) play this role. Allantoin arises from hydrolysis of purines (see Fig. 25-18), which are synthesized in root nodules of nitrogen-fixing plants.[101]

Glutamate, glutamine, and aspartate also play central roles in *removal* of nitrogen from organic compounds.[102] Transamination is reversible and is often the first step in catabolism of excess amino acids. 2-Oxoglutarate is the recipient of the nitrogen, and the glutamate that is formed can be deaminated to form ammonia which can then be incorporated into glutamine. Glutamate can also donate its nitrogen to form aspartate. In the brain glutamate is a major neurotransmitter but is toxic in excess. The astrocyte glial cells take up glutamate from the synaptic clefts between neurons, converting it to glutamine, which is then released into the extracellular space for reuptake by neurons.[103,104] In the animal body both aspartate and glutamine (via carbamoyl phosphate) are precursors of **urea**, the principal nitrogenous excretion product. These relationships are also summarized in Fig. 24-5, and details are provided in later sections.

While reductive amination of glutamate via glutamate synthase appears to be the major pathway for incorporation of nitrogen into amino groups, some direct amination of pyruvate and other 2-oxoacids in reactions analogous to that of glutamate dehydrogenase occurs in bacteria.[105,106] Another bacterial enzyme catalyzes reversible addition of ammonia to fumarate to form aspartate (p. 685).

An initially surprising conclusion drawn from the studies of Schoenheimer and Rittenberg was that proteins within cells are in a continuous steady state of synthesis and degradation. The initial biosynthesis, the processing, oxidative and hydrolytic degradative reactions of peptides, and further catabolism of amino acids all combine to form a series of metabolic loops as discussed in Chapter 17 and dealt with further in Chapters 12 and 29. Within cells some proteins are degraded much more rapidly than others, an important aspect of metabolic control. This is accomplished with the aid of the ubiquitin system (Box 10-C) and proteasomes (Box 7-A).[107] Proteins secreted into extracellular fluids often undergo more rapid turnover than do those that remain within cells.

1. Uptake of Amino Acids by Cells

While cells of autotrophic organisms can make all of their own amino acids, other organisms utilize many preformed amino acids. Human beings and other higher animals require several **essential amino acids** in their diets. Additional amounts of "nonessential" amino acids are also needed. It is true that amino groups can be transferred from one carbon skeleton to another among most of the amino acids. However, the body must take in enough amino groups to supply its need for all of the 20 amino acid components of proteins.[107a,b] Because of an unfavorable equilibrium constant, and the normally low concentration of NH_4^+, glutamate dehydrogenase (step *c* in Fig. 24-5) does not normally synthesize glutamate in the animal body. Its function is to deaminate excess glutamate. Furthermore, cells of some tissues take up amino acids that are made in other tissues. The active transport systems of bacteria have been described in Chapter 8. In mammals the absorption of amino acids takes place through epithelial cells of the intestinal tract, kidney tubules, and the brain (blood–brain barrier). Both Na^+-dependent transport (as for sugars; see Chapter 8) and Na^+-independent processes occur.[107c] Among the latter is the proposed **γ-glutamyl cycle,** which is described in Box 11-B. The cycle makes use of the γ-carboxyl group of glutamate, the same carboxyl that carries ammonia in the form of glutamine. Glutathione supplies the activated γ-glutamyl group. The amino acid to be transported reacts on the membrane surface by **transpeptidation**[108–109a] to form a **γ-glutamylamino acid** which enters the cytoplasm. It releases the free amino acid through an internal displacement by the free amino group of the glutamyl group. The natural tendency of the 5-carbon glutamate to undergo cyclization is used to provide the driving force for release of the bound amino acid. The cyclic product 5-oxoproline is then opened hydrolytically in an ATP-requiring reaction.[110] Cysteinylglycine formed in the initial transpeptidation is hydrolyzed by a peptidase, and glutathione is regenerated in two ATP-dependent steps as indicated in the scheme in Box 11-B.

The significance of the γ-glutamyl cycle is not fully understood. However, the finding of a mentally retarded individual who excretes 25–50 g / day of 5-oxoproline in the urine (possibly because of a defective 5-oxoprolinase) suggests that the pathway is a very active one.[111] A few persons deficient in γ-glutamyl transpeptidase have been found. They excrete glutathione and have a variety of medical problems.[109]

2. Glutamate Dehydrogenase and Glutamate Synthase

In animal tissues and in some bacteria the **glutamate dehydrogenase** reaction (Fig. 24-5, step *c*; see also Chapter 15)[112–115] provides a means of incorporating ammonia reversibly into glutamic acid. In eukaryotic cells the allosteric enzyme is found largely in the mitochondria.[115a] Glutamate dehydrogenase is also found in chloroplasts where it may function in glutamate synthesis when ammonia is present in excess.[116] The action of aminotransferases, both within and without mitochondria, distributes nitrogen from glutamate into most of the other amino acids. Especially active is aspartate aminotransferase (Eq. 14-24; Fig. 24-5, step *d*) which equilibrates aspartate and oxaloacetate with the 2-oxoglutarate–glutamate couple. However, the body obtains glutamate, as well as other amino acids, from foods, the initial source being largely green plants.

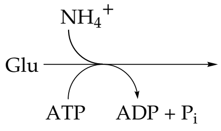
In plants as well as in *E. coli* and many other bacteria most glutamate is formed by **glutamate synthase**, which carries out reductive amination of 2-oxoglutarate (Fig. 24-5, reaction *a*). Glutamate synthase (also called **GOGAT**) utilizes the amide side chain of glutamine as the nitrogen donor. It is one of the previously mentioned amidotransferases in which glutamine is hydrolyzed to glutamate and NH_3 within the active site of the enzyme. Formation of a Schiff base and reduction probably occurs as in the reverse of reaction B of Table 15-1. However, one of the two glutamate molecules formed in reaction *a* of Fig. 24-5 must be reconverted by glutamine synthase to glutamine with the utilization of a molecule of ATP (Fig. 24-5, step *b*). Because of this coupling of ATP cleavage to the reaction the equilibrium in reaction *a* lies far toward the synthesis of glutamate. The low value of K_m for NH_4^+ that is characteristic of glutamine synthase favors glutamate synthesis even when little nitrogen is available.

Bacterial glutamate synthases are large oligomeric proteins containing flavin and Fe–S centers. That of *Azospirillum brasilense* consists of αβ units in which the 53-kDa β chains contain FAD and an NADPH binding site.[117–118a] The NADPH evidently transfers electrons to the FAD, which transfers them to an Fe_3S_4 center in the large 162-kDA α subunit. A molecule of bound FMN receives the electrons and reduces the iminoglutarate to glutamate. The site of binding and hydrolysis of glutamine is also present in subunit α. Chloroplasts of higher plants contain two glutamate synthases. One resembles the bacterial enzyme and utilizes NADPH as the reductant. The other requires reduced ferredoxin.[118–120]

Bacteria utilize both D-alanine and D-glutamate in the synthesis of their peptidoglycan layers (Fig. 8-29). Both D-amino acids are formed by racemases. That of alanine uses PLP (Chapter 15) but **glutamate racemase**[121–122a] does not. It may be able to remove the α-H of glutamate by utilizing the –COOH of the substrate, rather than the PLP ring, as an electron sink. Small amounts of D-amino acids occur also in animals.[123] Animal livers and kidneys contain **D-amino acid oxidase** and **D-aspartate oxidase,** which apparently function to metabolize D-amino acids from foods or those formed by brain activity (Chapter 30) or by aging.

3. Glutamine Synthetase

The formation of glutamine from glutamate (Eq. 24-16) also depends upon a coupled cleavage of ATP:



(24-16)

Glutamine synthase, as isolated from *E. coli*, contains 12 identical 51.6-kDa subunits arranged in the form of two rings of six subunits each with a center-to-center spacing of 4.5 nm. The units in one layer lie almost directly above those in the next,[104,124,125] the center-to-center spacing between the two layers is also 4.5 nm, and the array has 622 dihedral symmetry. The enzyme displays complex regulatory properties,[112,126–129] which are summarized in Fig. 24-7. The enzyme exists in two forms. **Active glutamine synthase** requires Mg^{2+} in addition to the three substrates glutamate, NH_4^+, and ATP. If the glutamate precursor, 2-oxoglutarate, is present in excess, the enzyme tends to remain in the active form because conversion to a modified form is inhibited; when the oxoglutarate concentration falls to a low value and glutamine accumulates, alteration is favored. The modifying enzyme **adenylyltransferase** (AT) in its active form AT_A transfers an adenylyl group from ATP to a tyrosine hydroxyl on glutamine synthase to give an adenylyl enzyme (GS-AMP). This **modified enzyme** requires Mn^{2+} instead of Mg^{2+} and is far more sensitive than the original enzyme to feedback inhibition by a series of end products of glutamine metabolism. All nine of the feedback inhibitors (serine, alanine, glycine, histidine, tryptophan, CTP, AMP, carbamoyl-*P*, and glucosamine 6-*P*) seem to bind to specific sites on the enzyme surface and to exert a cumulative inhibition. Serine, alanine, and glycine appear to be competitive inhibitors at the glutamate binding site.[130]

Relaxation of adenylylated glutamine synthase to the unmodified form is not catalyzed by a separate hydrolase but is promoted by a modified form of the adenylyltransferase AT_D. The active enzyme AT_A is

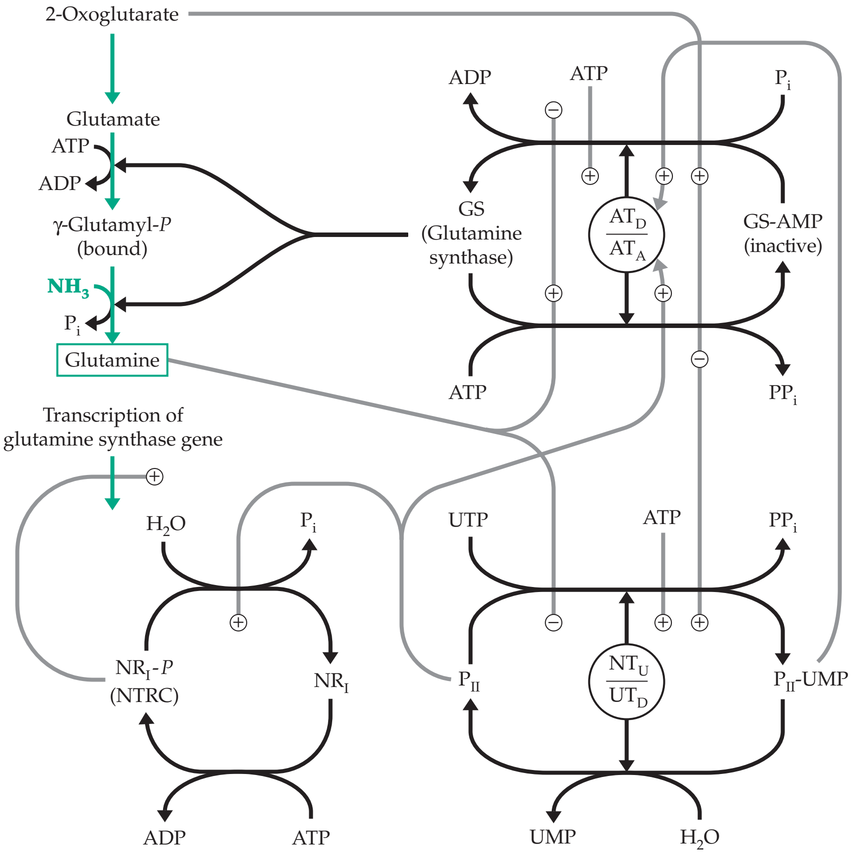
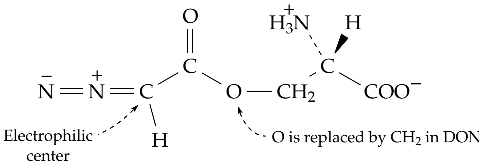


Figure 24-7 Regulation of glutamine synthase of *E. coli* using activation (+) and inhibition (–). Glutamine synthase (GS upper center) is converted by adenylylation of Tyr 397 into an inactive form GS-AMP by the action of an adenylyltransferase AT_A in complex with regulatory protein PII. PII is uridylylated at up to four sites by action of uridylyltransferase UT_U, which resides in the same polypeptide chain as a uridylyl removing enzyme UT_D (or UR). When PII carries a uridylyl group (PII-UMP), AT_A is transformed to AT_D, which reconverts the inactive GS-AMP to active GS by phosphorolytic removal of the adenylyl group. The ratios of AT_A / AT_D and UT_U / UT_D are controlled by the concentrations of the metabolites 2-oxoglutarate, a precursor, and glutamine, the immediate product. The amount of GS formed is controlled at the transcriptional level by an enhancer-binding transcription factor called NRI or NtrC (lower left). It is active when phosphorylated. Dephosphorylation of NRI-*P* is catalyzed by yet another protein and is stimulated by PII. Thus, PII both decreases synthesis of GS and promotes conversion of GS to its inactive form.

actually a complex AT • PII containing the regulatory protein PII. Subunit PII can be uridylylated on a tyrosine side chain by action of a 95-kDa **uridylyltransferase** (UT)[127,131] to form the modified glutamine synthetase AT • PII-UMP orAT_D. This form catalyzes phosphorolytic deadenylylation of glutamine synthetase, P_i displacing the adenylyl group to form ADP. Removal of the uridylyl group from PII-UMP is catalyzed by a fourth enzyme, UT_D (or UR), which is part of the same polypeptide chain as UT_U.[127] The cycle of interconversions of PII catalyzed by the UT_U and UT_D activities is shown at the lower right side of Fig. 24-7. From the allosteric modification reactions indicated by the gray lines, it is seen that glutamine not only promotes the adenylylation of glutamine synthetase but also inhibits the uridylylation of PII, thereby preventing AT_D from removing the adenylyl group from the synthetase. Furthermore, it allosterically inhibits the deadenylylation reaction itself. On the other hand, 2-oxoglutarate acts in the opposite way.

The glutamine synthase regulatory system has another important function. Protein PII stimulates the dephosphorylation of the enhancer-binding transcriptional regulator NRI-*P* (NtrC-*P*).[131,132] This slows transcription of the glutamine synthase gene (see Fig. 24-7) as well as a variety of other genes including those for the nitrogenase proteins in organisms that have them.[133] As a consequence, a deficiency of glutamine turns on a number of genes involved in nitrogen metabolism. Accumulation of glutamine promotes PII accumulation, modification of the synthase, and loss of gene activation.

Nitrogen can be transferred from glutamine into many other substrates.[102] Several antibiotic analogs of glutamine have been useful in studying these processes. Examples are the streptomyces antibiotics **L-azaserine** and 6-diazo-5-oxo-L-norleucine (DON).



L-Azaserine

These compounds act as alkylating agents; N_2 is released and a nucleophilic group from the enzyme becomes attached at the carbon atom indicated.[134] Other inhibitors bind noncovalently to form dead-end complexes.[134a]

4. Catabolism of Glutamine, Glutamate, and Other Amino Acids

Glutamine is hydrolyzed back to glutamate by glutaminases that are found both in eukaryotic tissues and in bacteria.[135,136] Liver contains an isozyme whose function appears to be to release NH_3 from glutamine for urea synthesis.[135,137] Glutamate dehydrogenase deaminates excess glutamate back to 2-oxoglutarate, which is degraded to succinyl-CoA and via β oxidation to malate, pyruvate, and acetyl-CoA. The last can reenter the citric acid cycle and be oxidized to CO_2 (Eq. 24-17). In fact, in mammalian tissues glutamate is essentially in equilibrium with 2-oxoglutarate and other citric acid cycle intermediates (see Box 17-C).

$$\begin{gathered}\text{Glu} \rightarrow \text{2-oxoglutarate} \rightarrow \text{succinyl-CoA} \rightarrow\ \rightarrow\ \rightarrow \\ \text{malate} \rightarrow \text{pyruvate} \rightarrow \text{acetyl-CoA} \rightarrow CO_2\end{gathered} \tag{24-17}$$

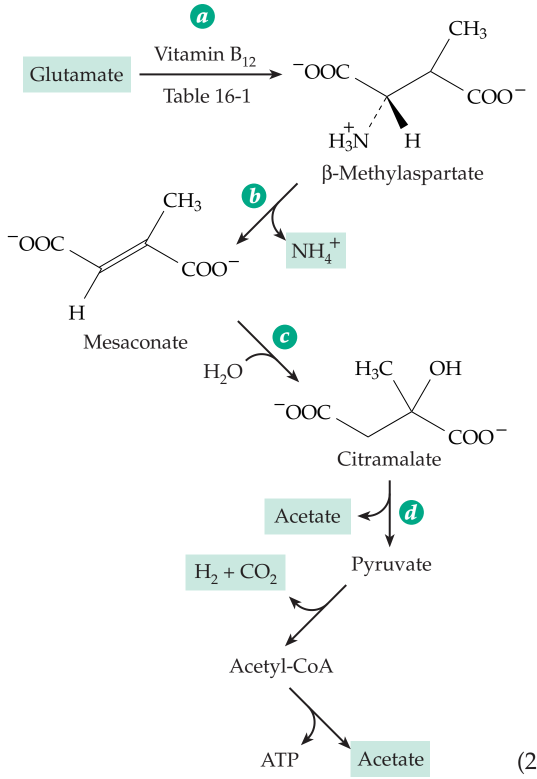
Many other amino acids are degraded in similar ways. In most cases the sequence is initiated by transamination to the corresponding 2-oxoacid. Beta oxidation and breakdown to such compounds as pyruvate and acetyl-CoA follows.

Catabolism initiated by decarboxylation. An alternative pathway for glutamate degradation is through the γ-aminobutyrate shunt (Fig. 17-5). This pathway is initiated by a PLP-dependent decarboxylation rather than by a deamination or transamination. Since decarboxylases are known for most amino acids, there are usually alternative breakdown pathways initiated by decarboxylation. In many cases these pathways lead to important products. For example, γ-aminobutyrate functions in the brain as an important neurotransmitter. Dihydroxyphenylalanine is converted to noradrenaline and adrenaline, tryptophan to serotonin, and histidine to histamine. All of these are neurotransmitters (Chapter 30) and/or have other hormonal functions. A calmodulin-dependent glutamate decarboxylase occurs in higher plants, which accumulate γ-aminobutyrate in response to a variety of stresses.[138] However, the significance of this accumulation is unclear.

Fermentation of glutamate. Special problems face anaerobic bacteria subsisting on amino acids. Their energy needs must be met by balanced fermentations.[138a] For example, glutamate may be converted to CO_2, ammonia, acetate$^-$, and butyrate$^-$ according to the reactions of Fig. 24-8. The end result is described by Eq. 24-18.

$$\begin{gathered}\text{2 Glutamate}^- + 2\ H_2O + H^+ \rightarrow \\ 2\ CO_2 + 2\ NH_4^+ + 2\ \text{acetate}^- + \text{butyrate}^- \\ \Delta G'\ (\text{pH } 7) = -131\ \text{kJ}\end{gathered} \tag{24-18}$$

The sequence begins with the γ-aminobutyrate shunt reactions (Fig. 24-8, steps *a* and *b*), but succinic semialdehyde is reduced to γ-hydroxybutyric acid using the NADH generated in the trans-deamination process of step *c*. With the aid of a CoA-transferase (step *d*) two molecules of the CoA ester of this hydroxy acid are formed at the expense of two molecules of acetyl-CoA. Use is then made of a β,γ elimination of water (step *e*), analogous to that involved in the formation of vaccenic acid (Eq. 21-2). Isomerization (perhaps by the same enzyme that catalyzes elimination) forms crotonyl-CoA (step *f*). The latter undergoes dismutation, one-half being reduced to butyryl-CoA and one-half being hydrated and oxidized to acetoacetyl-CoA in the standard β-oxidation sequence. Acetoacetyl-CoA is cleaved to regenerate the two molecules of acetyl-CoA. The organism can gain one molecule of ATP through cleavage of the butyryl-CoA. Perhaps a second can be gained by oxidative phosphorylation between the NADH produced in the formation of acetoacetyl-CoA and the reduction of crotonyl-CoA to butyryl-CoA. The two processes take place at sufficiently different redox potentials to permit this kind of coupling.


a
Glutamate
Vitamin B_{12}
Table 16-1
CH_3
^-OOC
COO^-
$H_3\overset{+}{N}$
H
β-Methylaspartate
b
CH_3
^-OOC
COO^-
H
NH_4^+
Mesaconate
c
H_2O
H_3C
OH
^-OOC
COO^-
Citramalate
Acetate
d
Pyruvate
$H_2 + CO_2$
Acetyl-CoA
ATP
Acetate

(24-19)

Another fermentation of glutamate is initiated by the vitamin B_{12}-dependent isomerization of glutamate to β-methylaspartate (Eq. 24-19, step *a*).[138b] This rearrangement of structure permits α,β elimination of ammonia (step *b*), a process not possible in the original glutamate. Hydration to **citramalate** (step *c*) and aldol cleavage yields acetate and pyruvate. Acetate is one of the usual end products of the fermentation. The pyruvate can be cleaved to H_2, CO_2, and acetyl-CoA by the pyruvate-formate-lyase system (Fig. 15-16; Eq. 17-25), and cleavage of the acetyl-CoA can provide ATP. Alternatively, two molecules of acetyl-CoA can be coupled and reduced to butyryl-CoA. The reducing power generated in the cleavage of pyruvate is used to reduce crotonyl-CoA rather than being released as H_2. The stoichiometry is identical to that in Fig. 24-8. Still other fermentation mechanisms are used by some Clostridia to degrade glutamate.[138a] See study question number 16 at the end of this chapter.

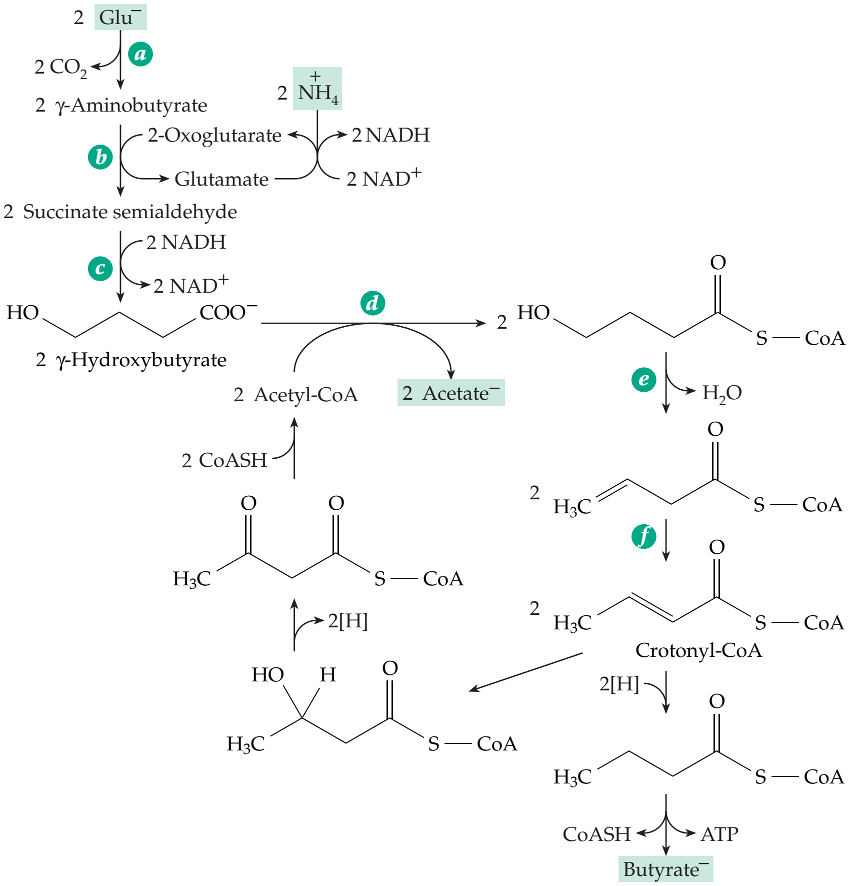

2 Glu^-
a
2 CO_2
2 γ-Aminobutyrate
2 $\overset{+}{NH_4}$
b
2-Oxoglutarate
Glutamate
2 NADH
2 NAD^+
2 Succinate semialdehyde
c
2 NADH
2 NAD^+
HO
COO^-
2 γ-Hydroxybutyrate
d
2 HO
O
S — CoA
2 Acetyl-CoA
2 $Acetate^-$
e
H_2O
2 CoASH
O
O
H_3C
S — CoA
2
H_3C
O
S — CoA
f
2[H]
2
H_3C
O
S — CoA
Crotonyl-CoA
HO
H
O
H_3C
S — CoA
2[H]
O
H_3C
S — CoA
CoASH
ATP
$Butyrate^-$

Figure 24-8 Fermentation of glutamate by *Clostridium aminobutylicum.*

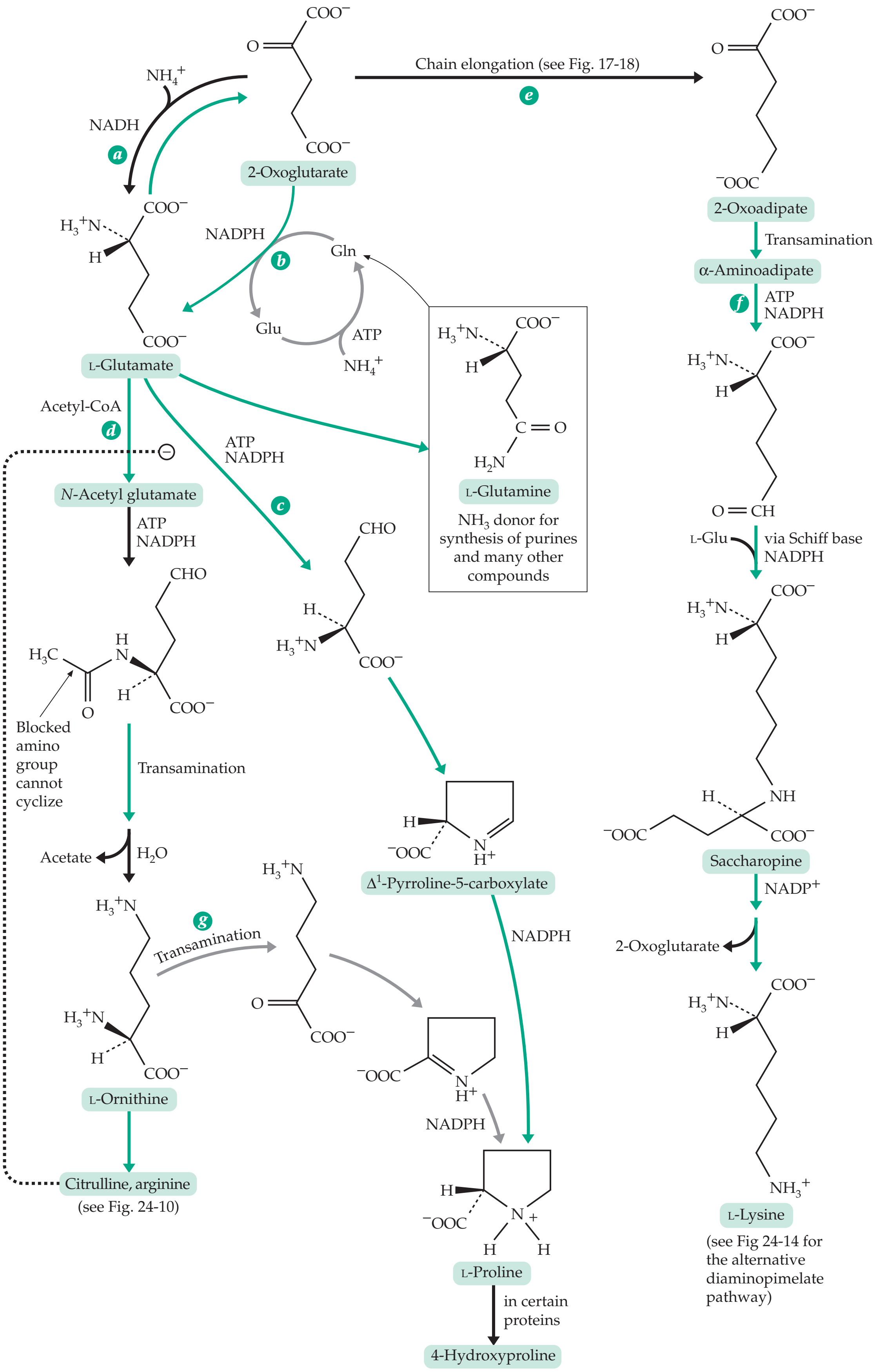

2-Oxoglutarate
Chain elongation (see Fig. 17-18)
e
2-Oxoadipate
NH4+
NADH
a
L-Glutamate
NADPH
b
Gln
Glu
ATP
NH4+
L-Glutamine
NH3 donor for synthesis of purines and many other compounds
Acetyl-CoA
d
ATP NADPH
c
N-Acetyl glutamate
ATP NADPH
Blocked amino group cannot cyclize
Transamination
Acetate
H2O
L-Ornithine
g
Transamination
Δ1-Pyrroline-5-carboxylate
NADPH
NADPH
L-Proline
in certain proteins
4-Hydroxyproline
Citrulline, arginine (see Fig. 24-10)
Transamination
α-Aminoadipate
f
ATP NADPH
L-Glu
via Schiff base NADPH
Saccharopine
NADP+
2-Oxoglutarate
L-Lysine
(see Fig 24-14 for the alternative diaminopimelate pathway)

Figure 24-9 Biosynthesis of glutamate, glutamine, proline, and lysine from 2-oxoglutarate.

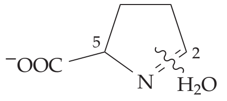
C. Synthesis and Catabolism of Proline, Ornithine, Arginine, and Polyamines

The 5-carbon skeleton of glutamic acid gives rise directly to those of proline, ornithine, and arginine. The reactions are outlined in Fig. 24-9. Arginine, in turn, is involved in the urea cycle, which is shown in detail in Fig. 24-10. Arginine is also a biosynthetic precursor of the polyamines. Another important biosynthetic product of glutamate metabolism is δ-aminolevulinate, a precursor to porphyrins (Eq. 24-44) in some organisms.[139]

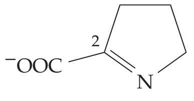
1. Synthesis and Catabolism of Proline

The ATP-dependent reduction of the γ-carboxyl group of glutamate to an aldehyde by NADPH (reaction *c*, Fig. 24-9) is of a standard biosynthetic reaction type, the opposite of the oxidation reaction of Fig. 15-6. Like the latter it is thought to occur via an acyl phosphate intermediate.[140,141] The oxidation product, **glutamate semialdehyde**, cyclizes and can be converted to proline by further reduction (Fig. 24-9, step *c* and subsequent reactions in center of scheme). The pathway has been well established in bacteria and yeast by both biochemical and genetic experiments. In plants both the initial reduction and the cyclization are catalyzed by a bifunctional enzyme.[142,143] An alternative pathway important in animals is initiated by transamination of ornithine to the corresponding 2-oxoacid, spontaneous cyclization, and reduction to proline (Fig. 24-9, step *g*).[140,144] Selected prolines in collagen and in plant glycoproteins[145] are oxygenated to form 4-hydroxyproline (Eqs. 8-6, 18-51).

One route of catabolism of proline is essentially the reverse of its formation from glutamate. **Proline oxidase** yields Δ^1-pyrroline 5-carboxylate.[145a,b]

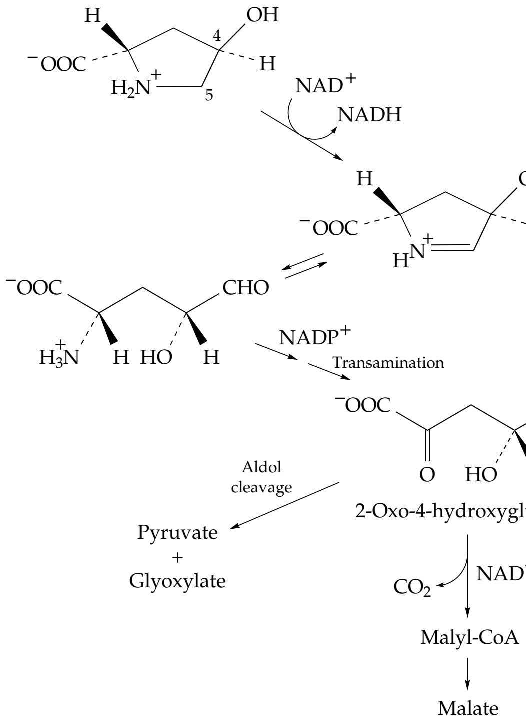


Δ^1-Pyrroline 5-carboxylate



Δ^1-Pyrroline 2-carboxylate

The corresponding open-chain aldehyde, formed by hydrolysis, can be oxidized back to glutamate by pyrroline 5-carboxylate dehydrogenase.[145a–147] Lack of this enzyme is associated with the human genetic deficiency causing **hyperprolinemia**.[147–148a]



(24-20)

Alternatively, degradation can be initiated by oxidation on the other side of the ring nitrogen to form Δ^1-pyrroline 2-carboxylate. The metabolic fate of this compound is uncertain. A corresponding pathway for breakdown of 4-hydroxy-L-proline of collagen yields glyoxylate and pyruvate or malate and CO_2 (Eq. 24-20).[149] Oxidation on the other side of the ring nitrogen of hydroxyproline is utilized by some pseudomonads to convert the amino acid into 2-oxoglutarate. Anaerobic bacteria may reduce proline to 5-aminovalerate and couple this reaction to the oxidative degradation of another amino acid (Stickland reaction).

2. Synthesis of Arginine and Ornithine and the Urea Cycle

If the amino group of glutamate is blocked by acetylation prior to the reduction to the semialdehyde (Fig. 24-9, step *d*) cyclization is prevented. The γ-aldehyde group can be transaminated to an amino group and the acetyl blocking group removed to form **ornithine**. Ornithine is not usually a constituent of proteins, but it is sometimes formed by hydrolytic modification of arginine at specific sites in a protein. A 67-kDa urate-binding glycoprotein of plasma is reported to contain 43 residues of ornithine.[150,151] It is postulated that a special arginase is needed to form these residues, and that it may be lacking in some cases of gout in which the urate-binding capacity of blood is impaired. Ornithine appears to be present in specific sites in a few other proteins as well.[151] *Neurospora* grown in a minimal medium accumulates large amounts of both ornithine and arginine, over 98% of which is sequestered in vesicles within the cytoplasm.[152,153] This appears to be a way of accumulating a store of arginine that is protected from the active catabolism of that amino acid by the fungus. However, accumulation of ornithine in the human body, as a result of lack of **ornithine aminotransferase**

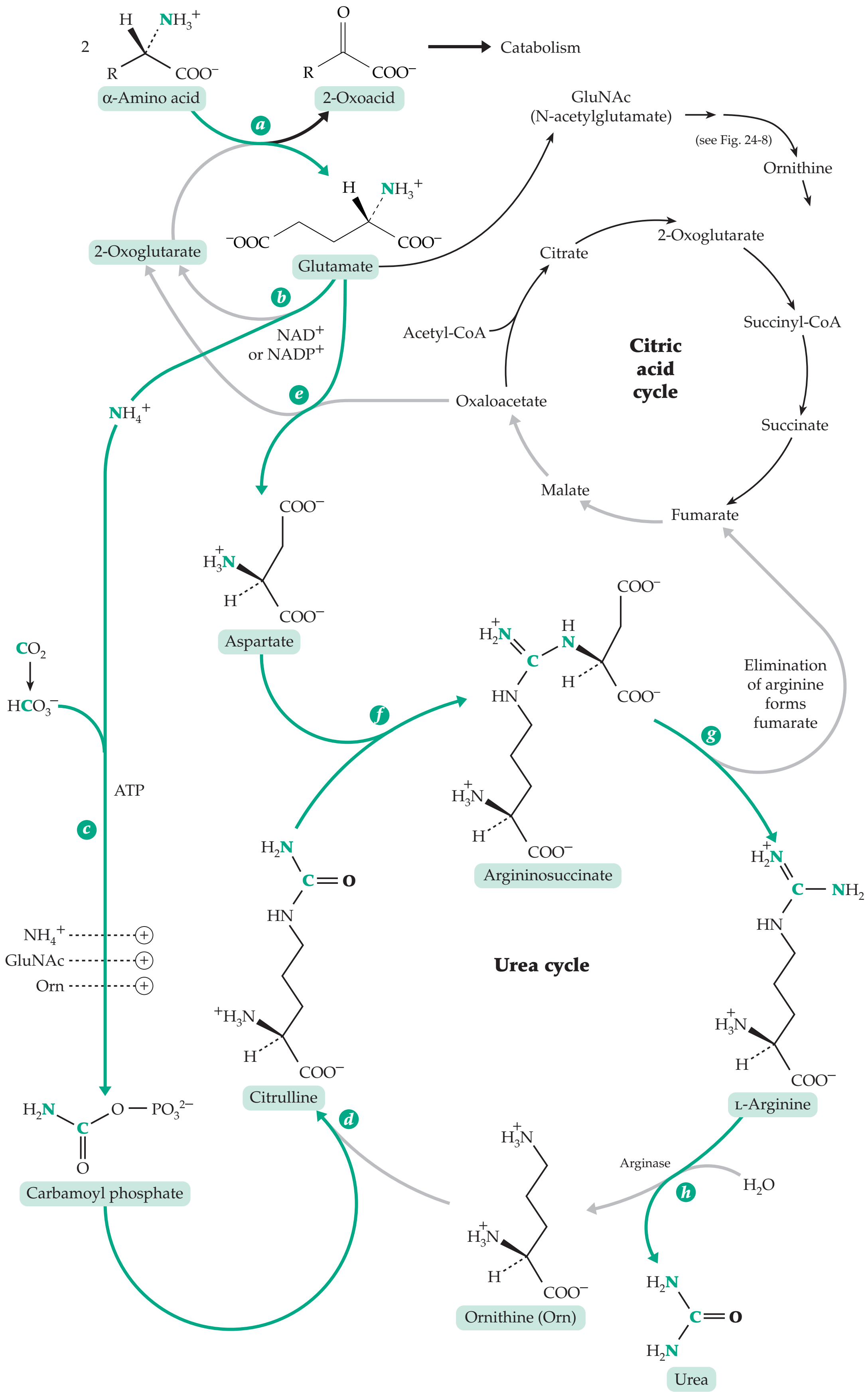

2
α-Amino acid
2-Oxoacid
Catabolism
a
2-Oxoglutarate
Glutamate
GluNAc
(N-acetylglutamate)
(see Fig. 24-8)
Ornithine
b
NAD+
or NADP+
e
NH_4^+
Citrate
2-Oxoglutarate
Succinyl-CoA
Acetyl-CoA
Citric acid cycle
Oxaloacetate
Succinate
Malate
Fumarate
Aspartate
CO_2
HCO_3^-
f
Argininosuccinate
g
Elimination of arginine forms fumarate
ATP
c
NH_4^+
GluNAc
Orn
Urea cycle
Citrulline
L-Arginine
Carbamoyl phosphate
d
Arginase
h
H_2O
Ornithine (Orn)
Urea

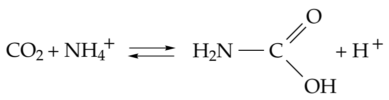
Figure 24-10 Biosynthesis of citrulline, arginine, and urea. The green arrows indicate reactions directly involved in deamination of amino acids and the synthesis of urea. N from amino acids and C from CO_2 are traced in green.

(Fig. 24-9, step *g*), causes gyrate atrophy of the choroid and retina, a disease that results in tunnel vision and blindness.[154,155] A major interest in arginine metabolism asises from its role in formation of urea in the human body. Study of arginine biosynthesis in bacteria has also been important in developing our understanding of regulation of gene expression.[156]

The urea cycle. In 1932, Krebs and Henseleit proposed that urea is formed in the liver by a cyclic process in which ornithine is converted first to **citrulline** and then to arginine.[157,158] The hydrolytic cleavage of arginine produces the urea and regenerates ornithine (Fig. 24-10, bottom). Subsequent experiments fully confirmed this proposal. Urea is the principal nitrogenous end product of metabolism in mammals and many other organisms, but the urea cycle reactions have other functions. As with the citric acid cycle, products other than urea can be withdrawn in any needed quantity. Most notably, the reactions of Fig. 24-10 provides for the biosynthesis of arginine in all organisms.[159,160] Also of physiological importance is the fact that the urea cycle involves both mitochondria and cytosolic enzymes.[161,162] This is illustrated in Fig. 24-11.

Let us trace the entire route of nitrogen removed by the liver from excess amino acids. Transaminases (step *a*, Fig. 24-10) transfer nitrogen to 2-oxoglutarate to form glutamate. Since urea contains two nitrogen atoms, two molecules of glutamate must donate their amino groups. One molecule is deaminated directly by glutamate dehydrogenase to form ammonia (step *b*). This ammonia is combined with bicarbonate (step *c*) to form carbamoyl phosphate, which transfers its carbamoyl group onto ornithine to form citrulline (step *d*). The second molecule of glutamate transfers its nitrogen by transamination to oxaloacetate (reaction *e*) to form aspartate. The aspartate molecule is incorporated intact into **argininosuccinate** by reaction with citrulline (reaction *f*). Undergoing a simple elimination reaction, the 4-carbon chain of argininosuccinate is converted to fumarate (step *g*) with arginine appearing as the elimination product. Finally, the hydrolysis of arginine (step *h*) yields urea and regenerates ornithine.

Carbamoyl phosphate synthetases. The first of the individual steps in the urea cycle is the formation of carbamoyl phosphate.[163] Carbon dioxide and ammonia equilibrate spontaneously with carbamic acid:

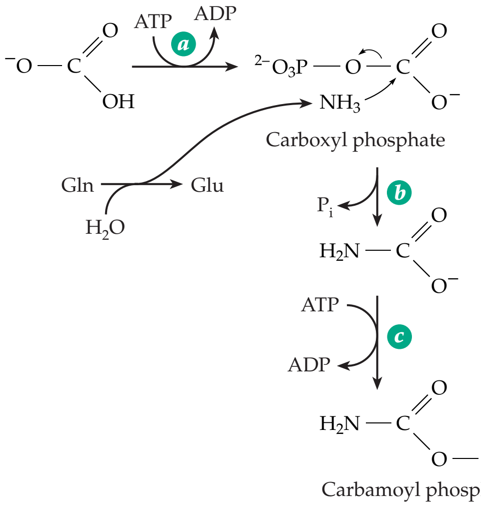


$$CO_2 + NH_4^+ \rightleftharpoons H_2N-C(=O)-OH + H^+$$

Carbamic acid

(24-21)

Some bacteria have a kinase able to convert carbamate into carbamoyl phosphate starting with step *a* of Eq. 24-22. However, the equilibrium constant is low (0.04 at pH 9, 10°C), and it is now believed that carbamate kinase functions in the opposite direction, providing a means of synthesis of ATP for bacteria degrading arginine (Section C,5,d). The biosynthetic carbamoyl phosphate synthases harness the cleavage of *two* molecules of ATP to formation of one molecule of carbamoyl phosphate (reaction *c*, Fig. 24-10).[163] In bacteria such as *E. coli*, a single synthase provides carbamoyl phosphate for biosynthesis of both arginine and pyrimidines (Fig. 25-14). However, fungi and higher animals have at least two carbamoyl-*P* synthases. Synthase I provides substrate for formation of citrulline from ornithine (Fig. 24-10), while carbamoyl-*P* synthase II, which is part of a larger multifunctional protein,[164] functions in pyrimidine synthesis. Synthase I is found in mitochondria and synthase II in the cytoplasm. Mammalian carbamoyl phosphate sythase I consists of a single 160-kDa peptide.[163] A powerful allosteric effector for the liver synthase is ***N*-acetylglutamate** (Fig. 24-10), a precursor of ornithine.[165] The enzyme from certain marine elasmobranchs, such as the spiny dogfish *Squalus acanthias*, have carbamoyl-*P* synthase III, an enzyme with somewhat different molecular properties.[166] It probably functions in synthesis of urea, which is used by these animals to regulate osmotic pressure.[167,168] Synthase I utilizes only free NH_3. The others are amidotransferases and prefer glutamine as the ammonia donor. Carbamoyl-*P* synthase from *E. coli* consists of two subunits (~42 and 118 kDa, respectively) and can utilize *either* free ammonia or glutamine.[169,170] The light subunit has **glutaminase** activity; i.e., it is able to hydrolyze glutamine to ammonia. All of these synthatases presumably act by first phosphorylating bicarbonate to an enzyme-bound carboxyl phosphate,[163,171,172] which can then undergo a displacement of phosphate by NH_3 to give enzyme-



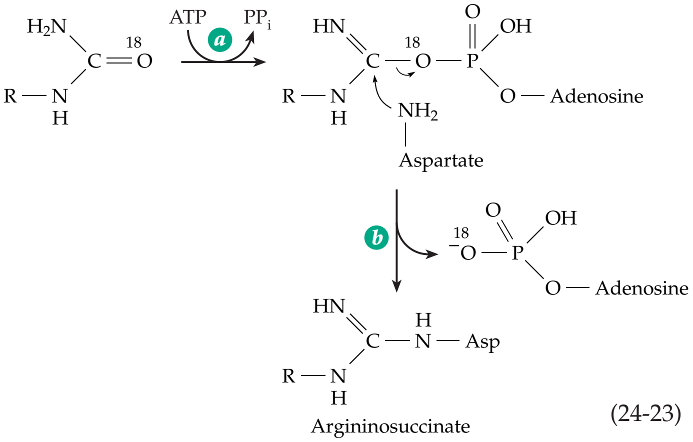
(24-22)

bound carbamate (Eq. 24-22, step *a*). Phosphorylation of the latter by ATP completes the reaction. In the single-chain enzymes the amidotransferase domain is at the N terminus.

Crystallographic study of a mutant form of the *E. coli* enzyme unable to act rapidly on glutamine showed that the latter released its ammonia to form a thioester with cysteine 269,[173,173a] suggesting a mechanism resembling that of serine proteases or papain (Chapter 12) for the glutaminase action. The X-ray crystallography also showed that the released NH_3 must travel 4.5 nm through the interior of the protein to the site of carbamate formation. The carbamate must travel another ~4.5 nm to the site from which carbamoyl phosphate is released.[170,172,173b] The C-terminal regions of the synthases undergo allosteric modification by a number of effectors.[163,173c] Both ornithine and IMP are activators for the *E. coli* enzyme, whereas UMP, a pyrimidine end product, exerts feedback inhibition. Phosphoribosyl pyrophosphate activates synthase II, and *N*-acetylglutamate activates the mammalian liver synthase I by binding near the C terminus.[165]

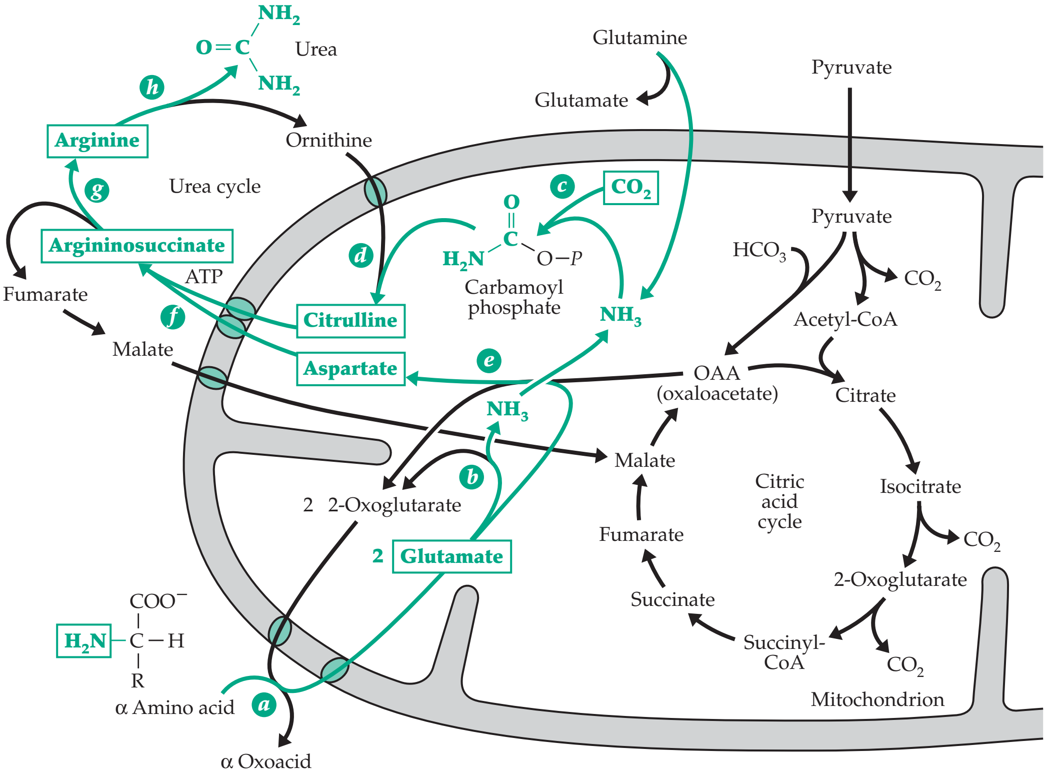
Citrulline and argininosuccinate. One NH_3 and one HCO_3^- for urea formation are provided by the carbamoyl group, which is transferred from carbamoyl-*P* to ornithine to form citrulline. The second nitrogen atom is transferred from glutamate to aspartate into argininosuccinate (steps *d* and *f*, Fig. 24-10). The equilibrium constant for ornithine transcarbamoylase (reaction *d*) is very high so that ornithine is completely converted to citrulline. The trimeric human enzyme is a trimer of 36-kDa subunits[174–175a] whose structural gene is on the X chromosome. Like many other mitochondrial matrix enzymes it is synthesized as a larger (40 kDa) pre-cursor, which enters the mitochondria in an energy-dependent process.[176] A genetic defect in this sex- linked gene is often lethal to boys, and even girls, heterozygous for the defect, sometimes have serious problems with accumulation of ammonia in the brain.[162,174,177]

The conversion of citrulline to argininosuccinate and the subsequent breakdown to fumarate and arginine take place in the cytosol (Fig. 24-11). The ureido



H₂N
18
C=O
R—N
H
ATP
PPᵢ
a
HN
C
O
OH
P
O — Adenosine
NH₂
Aspartate
b
¯O
C — N — Asp
Argininosuccinate

(24-23)



NH₂
O=C
Urea
NH₂
h
Arginine
Ornithine
g
Urea cycle
Argininosuccinate
Fumarate
ATP
f
Malate
d
Citrulline
Aspartate
c
CO₂
O
C
H₂N
O−P
Carbamoyl phosphate
NH₃
Glutamine
Glutamate
Pyruvate
Pyruvate
HCO₃
CO₂
Acetyl-CoA
e
OAA (oxaloacetate)
Citrate
NH₃
Malate
b
2 2-Oxoglutarate
2 Glutamate
Citric acid cycle
Isocitrate
CO₂
Fumarate
2-Oxoglutarate
Succinate
Succinyl-CoA
CO₂
Mitochondrion
COO⁻
H₂N
C−H
R
α Amino acid
a
α Oxoacid

Figure 24-11 Integration of the urea cycle with mitochondrial metabolism. Green lines trace the flow of nitrogen into urea upon deamination of amino acids or upon removal of nitrogen from the side chain of glutamine.

group of citrulline is activated by ATP for the argininosuccinate synthase reaction (Eq. 24-23, step *a*). Thus, ^{18}O present in this group is transferred into AMP. A citrulline adenylate intermediate (center) is likely. **Argininosuccinate lyase** (reaction *g*, Fig. 24-10)[177a] catalyzes the elimination of arginine with formation of fumarate. It is entirely analogous to the bacterial aspartase that eliminates ammonia from aspartate to form fumarate.[178] Like the latter enzyme and fumarate hydratase (Chapter 13), argininosuccinase promotes a trans elimination.[179] The fumarate produced can be reconverted through reactions of the citric acid cycle to oxaloacetate, which can be reaminated to aspartate (Fig. 24-11). Aspartate is used to introduce amino groups in an entirely similar way in other metabolic sequences such as in the formation of adenylic acid from inosinic acid (Fig. 25-16).

The cleavage of arginine to ornithine and urea by the Mn^{2+}-containing **arginase** (Chapter 16)[180,181] converts the biosynthetic route to arginine into a cycle for the synthesis of urea. This cyclic pathway is unique to organisms that excrete nitrogenous wastes as urea, but the biosynthetic path is nearly ubiquitous.[182] Human adults excrete approximately 20 g of urea nitrogen per day. If this rate decreases, ammonia accumulates in the blood to toxic levels. Normally, plasma contains 0.03 mM ammonia, and only 2–3 times this level is required to produce toxic symptoms. Therefore, it is not surprising that five different well-documented hereditary enzyme deficiencies affect the urea cycle.[162,183] One of the most common, **argininosuccinic aciduria**, is a deficiency of the breakdown of argininosuccinic acid.[178] Both lethal and nonlethal variants of this disease are known. Human argininosuccinate lyase consists of two subunits. Defects may occur in either subunit but considerable genetic heterogeneity exists and intragenic complementation between the two subunits accounts for many of the nonlethal forms of the disorder.[177a,178a] A common feature of all of the hereditary defects of the urea cycle is an intolerance to high protein intake and mental symptoms. Toxic accumulation of ammonia in blood is often seen also in **alcoholic liver cirrhosis** as a result of a decreased capacity of the liver for synthesis of urea.

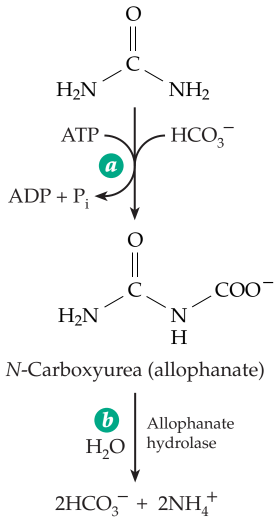
For some urea cycle defects a combination of a low-protein diet together with an arginine supplement prevents the ammonia intoxication while allowing normal growth. In other cases it is necessary to replace the natural dietary protein with a mixture of essential amino acids or with the corresponding 2-oxoacids, which can be converted to amino acids in the body with utilization of endogenous ammonia.[183] A specific treatment for lack of *N*-acetylglutamate synthetase, which forms the carbamoyl phosphate synthase activator *N*-acetylglutamate, is administration of the analog *N*-carbamoylglutamate. This also activates carbamoyl phosphate synthase and is not cleaved by acylases that would prevent the natural activator from being supplied artificially via the blood.[158]

Although the primary function of the urea cycle is usually regarded as the removal of NH_4^+ from the body, it also removes HCO_3^- in equal amounts (Eq. 24-24). This is essential for maintenance of neutral pH,

$$2\ HCO_3^- + 2\ NH_4^+ \longrightarrow H_2NCONH_2 + CO_2 + 3\ H_2O \tag{24-24}$$

and Atkinson and Bourke suggested that removal of HCO_3^- is as important a function of the cycle as removal of NH_4^+.[184] However, there are strong arguments against this concept.[162]

Excretion of ammonia. While mammals excrete urea, many invertebrate organisms that live in water as well as some fishes simply excrete NH_3. Other organisms hydrolyze urea to NH_3. Even green plants recycle nitrogen via urea and the Ni^{2+}-dependent urease (Eq. 16-47). Two compounds that can be hydrolyzed by cells to urea and glyoxylate are allantoin and allantoic acid (Eq. 24-15). If cells of *Saccharomyces cerevisiae* are grown on either of these compounds as a sole source of nitrogen, they make a biotin-dependent **urea carboxylase** (Eq. 24-25). This enzyme facilitates the hydrolysis of urea by conversion to the more easily degraded allophanate (Eq. 24-25).

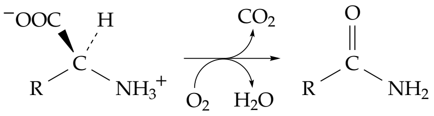


(24-25)

Catabolism of arginine. Arginine can also be converted back to glutamate and 2-oxoglutarate. The initial step is removal of the guanidino group to form ornithine. This occurs in the urea cycle and also in many bacteria[185] by the action of arginase (Fig. 24-5, step *h*). A parallel pathway involving conversion of arginine to N^1-succinylarginine, then on to succinylglutamate, and to free glutamate and succinate is used by some pseudomonads.[186] The alternative **arginine**

dihydrolase pathway, used by some bacteria and a few protozoa such as *Giardia*, is initiated by a different hydrolase that cleaves arginine to citrulline and ammonia.[187] Phosphorolysis of citrulline yields carbamoyl phosphate whose breakdown to CO_2 and ammonia (catalyzed by carbamate kinase, Eq. 24-22) can be utilized for generation of ATP by microorganisms that subsist on arginine.[188]

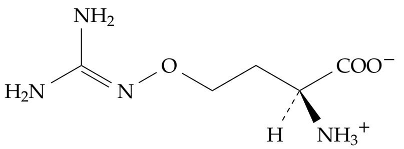
Degradation of L-arginine by *Streptomyces griseus* is initiated by a hydroxylase that causes decarboxylation and conversion of the amino acid into an amide (Eq. 24-26), a reaction analogous to that catalyzed by the flavin-dependent lysine oxygenase (Eq. 18-41). The product formed from arginine is γ-guanidinobutyramide, which is further degraded by the hydrolysis of the amide group and cleavage of the guanidino group to form urea and γ-aminobutyrate. *Pseudomonas putida* initiates degradation of arginine by decarboxylation to the corresponding 2-oxoacid and oxidative decarboxylation with a thiamin diphosphate-requiring enzyme to γ-guanidinobutyraldehyde. Dehydrogenation and hydrolysis lead, again, to γ-aminobutyrate.[189]



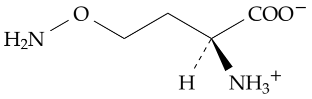
(24-26)

Specific arginine residues in proteins are methylated on their guanidino groups to give monomethylated and both symmetrically and asymmetrically dimethylated derivatives.[190,191] These methylated arginines also occur free in various mammalian tissues, where they may serve as endogenous regulators of nitric oxide synthases. A Zn^{2+}-containing dimethylarginase hydrolyzes the monomethyl and dimethyl arginines to citrulline and monomethyl or dimethyl amines.[191]

Insecticidal analogs of arginine. The toxic amino acid **L-canavanine** is synthesized by more than 1500 species of legumes including alfalfa and clover.[192–194] It is structurally similar to arginine, the 5-CH_2 group being replaced by O. However, the guanidino group is much less basic than in arginine. Canavanine is a natural insecticide, which in some plants accumulates to a level of 13% of the total dry matter.[193] Plants that store canavanine hydrolyze it to **canaline** and urea, which they use as a nitrogen source. Canaline is a toxic derivative of hydroxylamine and forms oximes with 2-oxoglutarate, other oxoacids, and PLP-containing enzymes. Although canavanine and canaline are effective insecticides, some beetles are adapted to these compounds to the extent that they feed exclusively on canavanine-containing seeds. The tobacco budworm is likewise resistant to these toxins and produces a **canavanine hydrolase** that converts canavanine to L-homoserine, a normal intermediate in the threonine, isoleucine, methionine biosynthetic pathway (Fig. 24-13), and **hydroxyguanidine**. The latter undergoes NADH-dependent reduction to guanidine[194] which can be catabolized.



L-Canavanine

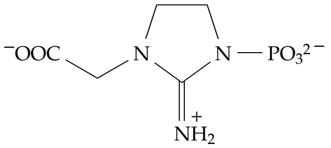


L-Canaline

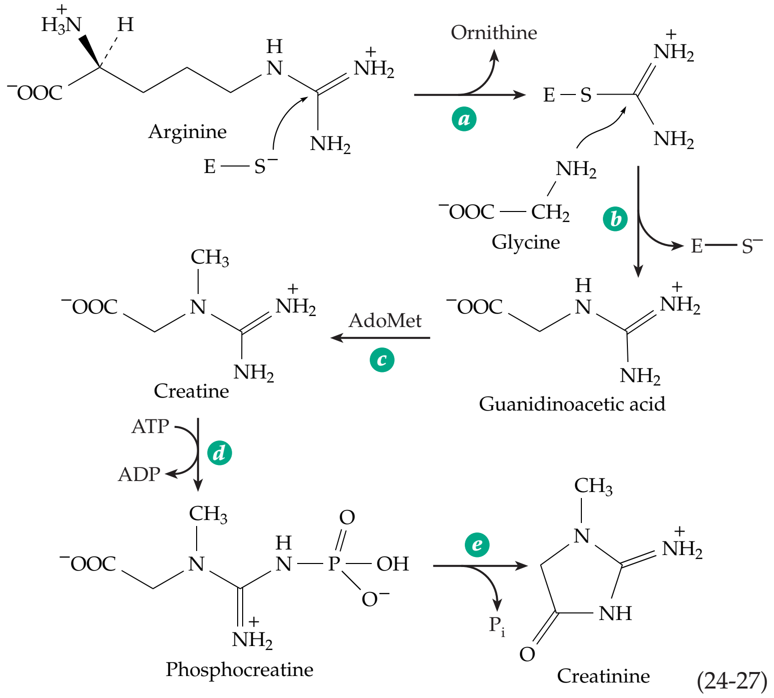
3. Amidino Transfer and Creatine Synthesis

The terminal amidino group of arginine is transferred intact to a number of other substances in simple displacement reactions. An example is the formation of **guanidinoacetic acid** (Eq. 24-27, steps *a* and *b*). The amidino group appears to be transferred first to the SH group of cysteine 407 then to glycine in a double displacement mechanism.[195–197] Transmethylation from *S*-adenosylmethionine (Eq. 24-27, step *c*) converts guanidinoacetic acid to **creatine**, a compound of special importance in muscle. Creatine kinase reversibly transfers the phospho group of ATP to creatine to form the *N*-phosphate (Eq. 24-27, step *d*). **Creatine phosphate**, and in some invertebrates phosphoarginine,[198] serves as an important "energy buffer" for muscular contraction (Chapter 19). Through the reversible action of creatine kinase it is able rapidly to transfer its phospho group back onto ADP as fast as the latter is formed during the hydrolysis of ATP in the contraction process. An end product of creatine phosphate metabolism is the anhydride **creatinine** formed from creatine phosphate as is indicated in Eq. 24-27, step *e* as well as directly from creatine. The urinary creatinine excretion for a given individual is extremely constant from day to day, the amount excreted apparently being directly related to the muscle mass of the person. Another example of the transfer of amidino groups from arginine is found in the synthesis of streptomycin (Box 20-B).

A cyclic analog of creatine, **cyclocreatine**, when fed to animals, accumulates in large amounts in muscle, heart, and brain and is a long-acting phosphagen.[199]



Cyclocreatine phosphate



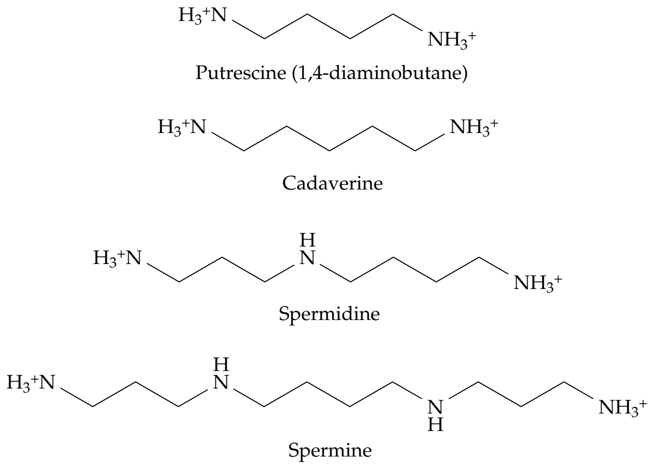
(24-27)

4. The Polyamines

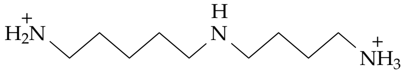
A series of related polyamino compounds, which are derived in part from arginine, are present in all cells in relatively high, often millimolar, concentrations.[200–203] The content of polyamines in cells tends to be stoichiometric with that of RNA, and the polyamines are concentrated in the ribosomes and also in the nucleus. Two moles of polyamine are usually present per mole of any isolated tRNA.[202,204] The first satisfactory crystals of tRNA for X-ray structure determination were obtained in the presence of spermine. Spermidine is associated with RNA in the turnip yellow mosaic virus.[205] The T-even bacteriophage and most bacteria contain polyamines in association with DNA. Polyamines are able to interact with double helical nucleic acids by bridging between strands, the positively charged amino groups interacting with the phosphates of the nucleic acid backbones. Tsuboi suggested that the tetramethylene portion of the polyamine lies in the minor groove bridging three base pairs, and the trimethylene portions (one in spermidine, two in spermine) bridge adjacent phosphate groups in one strand.[206] Polyamines may also stabilize supercoiled or folded DNA.

The structures of polyamines are shown here as di- and tri-cations, but it should be realized that there are multiple positions for protonation and therefore various tautomers. Also, polyamines show extreme anticooperativity in proton binding, i.e., successive pK_a values range from very low to very high for the last proton to leave. Polyamines are thought to have several functions. They can substitute to some extent for cellular K^+ and Mg^{2+}, and they may play essential controlling roles in nucleic acid and protein synthesis. A specific role of spermidine in cell division seems likely.[207,207a] An absolute requirement for polyamines has been demonstrated for some bacteria such as *Hemophilus parainfluenzae*[208] and for mutants of *Aspergillus* and *Neurospora*. Polyamines are also essential for mammalian cells. Polyamines activate some enzymes including the serine/threonine protein kinase CK2.[209]

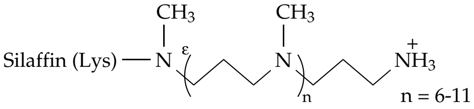
Mutants of *E. coli* have been constructed in which enzymes of all known biosynthetic pathways for polyamines are blocked by deletion of the genes for arginine decarboxylase (*SpeA*), agmatine ureahydrolase (*SpeB*), ornithine decarboxylase (*SpeC*), and adenosylmethionine decarboxylase (*SpeD*).[210] Even though polyamines cannot be detected in these cells they grow at one-third the normal rate. However, yeast cells require both putrescine and spermidine or spermine for growth.[211,211a] Another effect is seen in strains of yeast carrying the **"killer plasmid,"** a 1500-kDa double-stranded RNA plasmid that encodes a toxic protein, which is secreted and kills other susceptible strains of yeast. Yeast cells carrying the killer plasmid lose it when made deficient in polyamines.[212] The bacterial outer membrane porins OmF and OmC (Fig. 8-20) bind polyamines, especially spermine, and inhibit passage of ions. Polyamines may also modulate ion channels of heart, muscle, and neurons.[213] Both prokaryotic and eukaryotic cells have transporters that allow uptake of polyamines from their surroundings.[214,215]



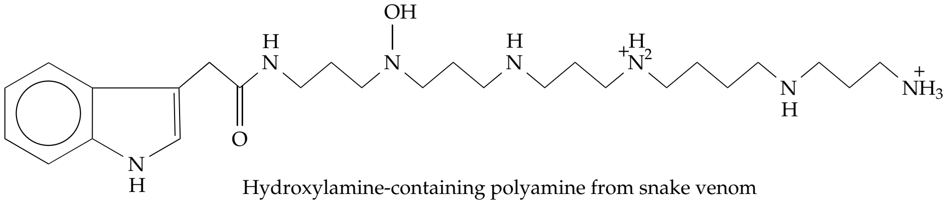
Major naturally occurring polyamines (as dications)



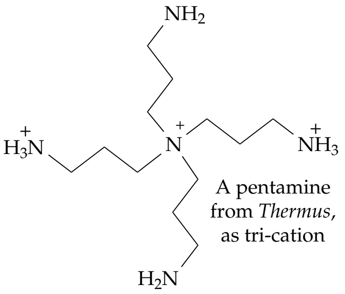
4-Aminobutylcadaverine



Polypeptide-bound polyamines of diatoms



Hydroxylamine-containing polyamine from snake venom



A pentamine from *Thermus*, as tri-cation

Biosynthesis. The 4-carbon putrescine arises most directly by decarboxylation of ornithine (Fig. 24-12, step *b*),[216] but it can also be formed by decarboxylation of arginine to agmatine followed by hydrolysis of the latter (Fig. 24-12, steps *c*,*d*). An alternative pathway utilizes an "agmatine cycle" in which agmatine is first hydrolyzed to ammonium ions and *N*-carbamoylputrescine. The latter transfers its ureido group to ornithine to form citrulline and releases free putrescine (Fig. 24-12, steps *f*,*g*). The citrulline is reconverted to arginine. This pathway appears to be important in plants.[217] Putrescine is normally present in all cells, and all cells are able to convert it on to spermidine. This is accomplished by decarboxylation of *S*-adenosylmethionine (Fig. 24-12, step *a*) and transfer of the propylamine group from the resulting decarboxylation product onto an amino group of putrescine (Fig. 24-12, step *h*).[218–221]:

The more complex spermine is found only in eukaryotes. It is formed by transfer of a second propylamine group onto spermidine (Fig. 24-12, step *i*). A historical note is that Anthony von Leeuwenhoek with one of his first microscopes observed crystals of the phosphate salt of spermine in human semen in 1678. The 5-carbon diamine cadaverine arises from decarboxylation of lysine (Fig. 24-12, step *d*). The extremely thermophilic bacterium *Thermus thermophilis* produces several additional polyamines including a pentamine (see above), a quaternary nitrogen compound.[222] Many other polyamines are known.[223–224c] Among these are a 4-aminobutylcadaverine isolated from root nodules of the adzuki bean[223] and very long partially aromatic hydroxylamine derivatives from venom of common funnel-web spiders (structures at top of page).[225] Cationic polypeptides called **silaffins**, with masses of ~3 kDa, apparently initiate the growth of the silica cell walls of diatoms (Box 4-B). These peptides contain polyamines consisting of 6 to 11 repeated *N*-methylpropylamine units covalently attached to lysine residues[224a,224b] and also many phosphoserines.[224d]

The synthesis of polyamines is tightly regulated. The PLP-dependent ornithine decarboxylase is present in very low concentrations[226] and apparently has the shortest half-life (~10 min) of any mammalian

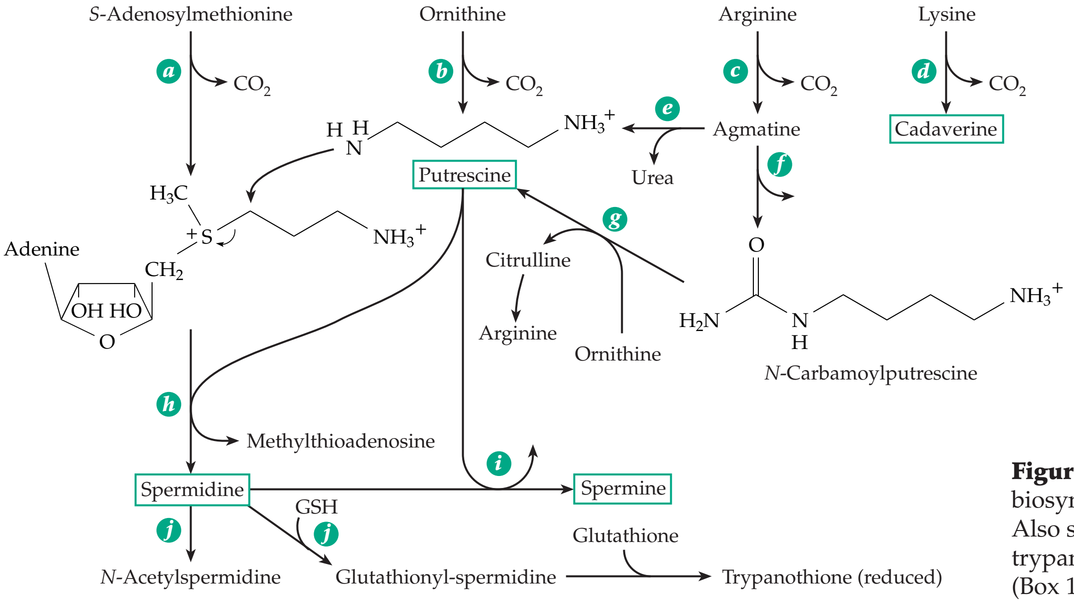
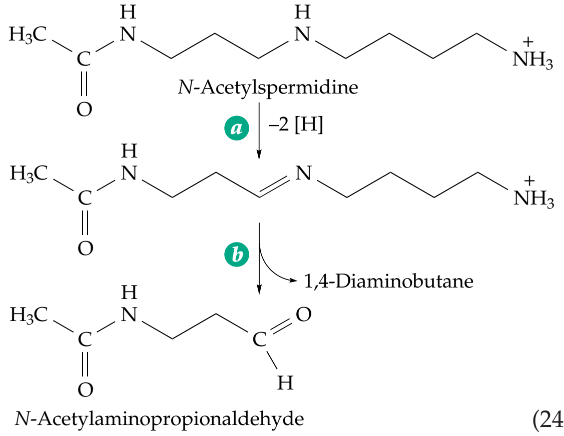

S-Adenosylmethionine
Ornithine
Arginine
Lysine
a
b
c
d
CO_2
CO_2
CO_2
CO_2
e
NH_3^+
Agmatine
Cadaverine
Urea
Putrescine
f
g
H_3C
S
NH_3^+
Adenine
CH_2
OH HO
O
Citrulline
Arginine
Ornithine
O
H_2N
N
H
NH_3^+
N-Carbamoylputrescine
h
Methylthioadenosine
i
Spermidine
Spermine
GSH
j
j
Glutathione
N-Acetylspermidine
Glutathionyl-spermidine
Trypanothione (reduced)

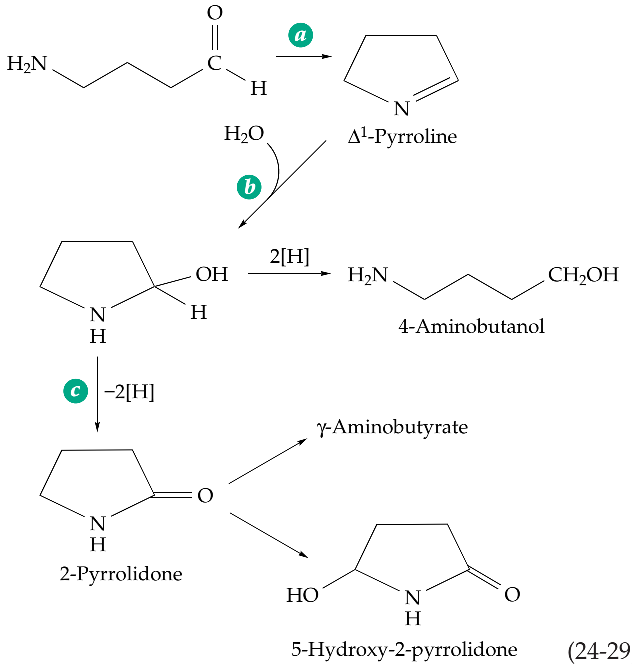
Figure 24-12 Pathways of biosynthesis of polyamines. Also shown is the formation by trypanosomes of trypanothione (Box 11-B).

enzyme.[227,228] Its concentration increases rapidly in most species with the onset of rapid growth, transformation to a neoplastic state, or initiation of cell differentiation. The rate of synthesis of the enzyme appears to be regulated by feedback repression by spermidine and by inactivation in response to a buildup of putrescine.[229] One mechanism of inactivation is the synthesis of a 26-kDa specific inhibitor called an **antizyme** in response to the presence of putrescine, spermidine, or spermine. The antizyme is ubiquitous in both prokaryotes and eukaryotes and keeps most of the ornithine decarboxylase bound and inactive and also promotes its degradation by 26S proteosomes.[230–230b] A polyamine-dependent protein kinase in *Physarum* phosphorylates the decarboxylase thereby inhibiting its activity.[231]

Breakdown. The catabolism of polyamines is less well understood than is their biosynthesis. Oxidative cleavages of spermine to spermidine and of the latter to 1,4-diaminobutane appear to occur in the animal body, and a substantial amount of this diamine is excreted in the urine.[201] Spermidine is acetylated on N^1 by acetyl-CoA and a spermidine *N*-acetyltransferase.[232,233] The resulting N^1-acetylspermidine is more readily cleaved by hepatic polyamine oxidase[233a] than is free spermine; again 1,4-diaminobutane is reformed together with an *N*-acetylaminopropionaldehyde (Eq. 24-28). This and other aldehydes formed from polyamines are very toxic but they may play essential roles in regulation of metabolism.[201] Transamination of 1,4-diaminobutane yields γ-aminobutyraldehyde which cyclizes (Eq. 24-29). Diamine oxidases of animal tissues oxidize 1,4-diaminobutane with formation of the same products.[234] Further metabolism of Δ^1-pyrroline yields γ-aminobutyrate, which can undergo transamination and oxidative metabolism as shown in Fig. 17-5. Other products[235] are also indicated in Eq. 24-29. Metabolism of other polyamines also begins by oxidation at the primary amino termini.[236] Formation of β-alanine, needed for synthesis of pantothenic acid, can also occur by oxidation of spermine.[236a]



(24-28)



(24-29)

When *E. coli* cells enter the stationary phase of the growth curve (Box 9-B), most of the spermidine is converted to **glutathionylspermidine** (γ-glutamylcysteinylglycylspermidine) in which glutathione and spermidine are joined by an amide linkage.[237–239] Trypanosomes join a second glutathione at the other end of the spermidine to form reduced **tryptathione**,[238] a compound also considered in Box 11-B. N^1-γ-Glutamylspermidine and related compounds have been found in proteolytic digests of certain proteins, suggesting that polyamines may be physiological substrates for transglutaminases.[240] Portions of polyamines are incorporated into a variety of products including **nicotine** (Fig. 30-22)[241] and the unusual amino acid **hypusine** (see p. 1386).[242]

Ornithine decarboxylase is specifically inhibited by the enzyme-activated inhibitor α-difluoromethylornithine, which can cure human infection with *Trypanosoma brucei* (African sleeping sickness) by interfering with polyamine synthesis.[243–244a] In combination with inhibitors of spermidine synthase or *S*-adenosylmethionine decarboxylase,[245] it can reduce polyamine levels and growth rates of cells. Another powerful inhibitor that acts on both ornithine and adenosylmethionine decarboxylases is the hydroxylamine derivative 1-aminooxy-3-aminopropane.[246]

$$H_2N{-}O{-}CH_2CH_2CH_2{-}NH_2$$

1-Aminooxy-3-aminopropane

Like difluoromethylornithine the compound at low concentrations is not toxic to cells but inhibits growth. It is hoped that adequate inhibition of growth of normal cells may allow more aggressive chemotherapeutic treatment of cancer.

D. Compounds Derived from Aspartate

The 4-carbon aspartate molecule is the starting point for synthesis of **pyrimidines** and of the amino acids **lysine**, **methionine**, **threonine**, **isoleucine**, and **asparagine**.[247,248] The pathways are summarized in Fig. 24-13. There are several branch points, and aspartate can be converted directly to asparagine, to carbamoylaspartate (the precursor of pyrimidines), or to β-aspartyl phosphate and aspartate semialdehyde. The latter can be converted in one pathway to lysine and in another to homoserine. Homoserine can yield either homocysteine and methionine or threonine. Although threonine is one of the end products and a constituent of proteins, it can also be converted further to 2-oxobutyrate, a precursor of isoleucine.

Most of the chemistry has been considered already. The reduction of aspartate via β-aspartyl phosphate[249,249a] and aspartate β-semialdehyde[250] is a standard one. Conversion to methionine can occur in two ways. In *E. coli* homoserine is succinylated with succinyl-CoA. The γ-succinyl group is then replaced by the cysteine molecule in a PLP-dependent γ-replacement reaction (Fig. 24-13). The product **cystathionine** (Eq. 14-33) undergoes elimination to form homocysteine. A similar pathway via *O*-phosphohomoserine occurs in chloroplasts of green plants.[251] A more direct γ replacement of the hydroxyl of homocysteine or *O*-phosphohomoserine by a sulfide ion has also been reported for both *Neurospora* and green plants.[252] Methylation of homocysteine to methionine (Fig. 24-13) has been considered previously, as has the conversion of homoserine to threonine by homoserine kinase[253] and the PLP-dependent **threonine synthase** (p. 746, Fig. 14-7).[254–255a] A standard PLP-requiring β elimination converts threonine to **2-oxobutyrate**, a precursor to isoleucine (Fig. 24-13).[256]

Formation of **asparagine** has been discussed in section B. Asparagine synthase of *E. coli*[98–99] cleaves ATP to AMP and PP_i rather than to ADP via an aspartyladenylate intermediate. In higher animals glutamine serves as the ammonia donor for synthesis of asparagine, but NH_4^+ can also function.[257] **L-Asparaginase**, a bacterial hydrolase, is an experimental antileukemic drug. It acts to deprive fast-growing tumor cells of the exogenous asparagine needed for rapid growth.[136,257a] Tissues with a low asparagine synthase activity are also damaged, limiting the clinical usefulness.

Aspartate can be decarboxylated either to α-alanine by a PLP-dependent enzyme[258] or to β-alanine by a pyruvoyl group-containing enzyme (Chapter 14). Beta-alanine is not only a component of the vitamin pantothenic acid but is found in the dipeptides carnosine (β-alanylhistidine) and anserine (β-alanyl-$N^δ$-methylhistidine) present in vertebrate muscles.[259] It is a crosslinking agent in insect cuticle.

Aspartate can be deaminated to fumarate by bacterial L-**aspartate oxidase**.[259a] This flavoprotein is structurally and mechanistically related to succinate dehydrogenase and can function as a soluble fumarate reductase (p. 1027). However, its main function appears to be to permit the intermediate iminoaspartate to react with dihydroxyacetone-*P* to form quinolinate, which can be converted to **NAD** (see Fig. 25-11).[259b]

1. Control of Biosynthetic Reactions of Aspartate

In *E. coli* there are three **aspartokinases** that catalyze the conversion of aspartate to β-aspartyl phosphate. All three catalyze the same reaction, but they have very different regulatory properties, as is indicated in Fig. 24-13. Each enzyme is responsive to a different set of end products.[247,260] The same is true for the two **aspartate semialdehyde reductases** which catalyze the third step. Both repression of transcription and feedback inhibition of the enzymes are involved. Two of the aspartokinases of *E. coli* are parts of bifunctional enzymes, which also contain the homoserine dehydrogenases that are needed to reduce aspartate semialdehyde in the third step. These aspartokinase–homoserine dehydrogenases I and II (Fig. 24-13) are encoded by *E. coli* genes *thrA* and *metL*, respectively, and have homologous sequences.[247,261–262a] The N-terminal portions are also homologous to the lysine-sensitive aspartokinase III which is encoded by the *lysC* gene.[263] In *Bacillus subtilis* the lysine-sensitive enzyme is known as aspartokinase II. It has an $α_2β_2$ oligomeric structure and both α and β chains are encoded within a single gene.[264] There is no associated homoserine dehydrogenase. Both genetic organization and processing of the synthesized protein are thus different in these two bacteria.

2. Lysine Diaminopimelate, Dipicolinic Acid, and Carnitine

Lysine cannot be made at all by animals but is nutritionally essential. There are two distinct pathways for its formation in other organisms. The **α-aminoadipate pathway** (shown in Fig. 24-9) occurs in a few lower fungi, the higher fungi, and euglenids. The 5-carbon 2-oxoglutarate is the starting compound. Bacteria, other lower fungi, and green plants all use the **diaminopimelate** pathway (Fig. 24-14) which originates with the 4-carbon aspartate.

The α-aminoadipate pathway (Fig. 24-9) parallels

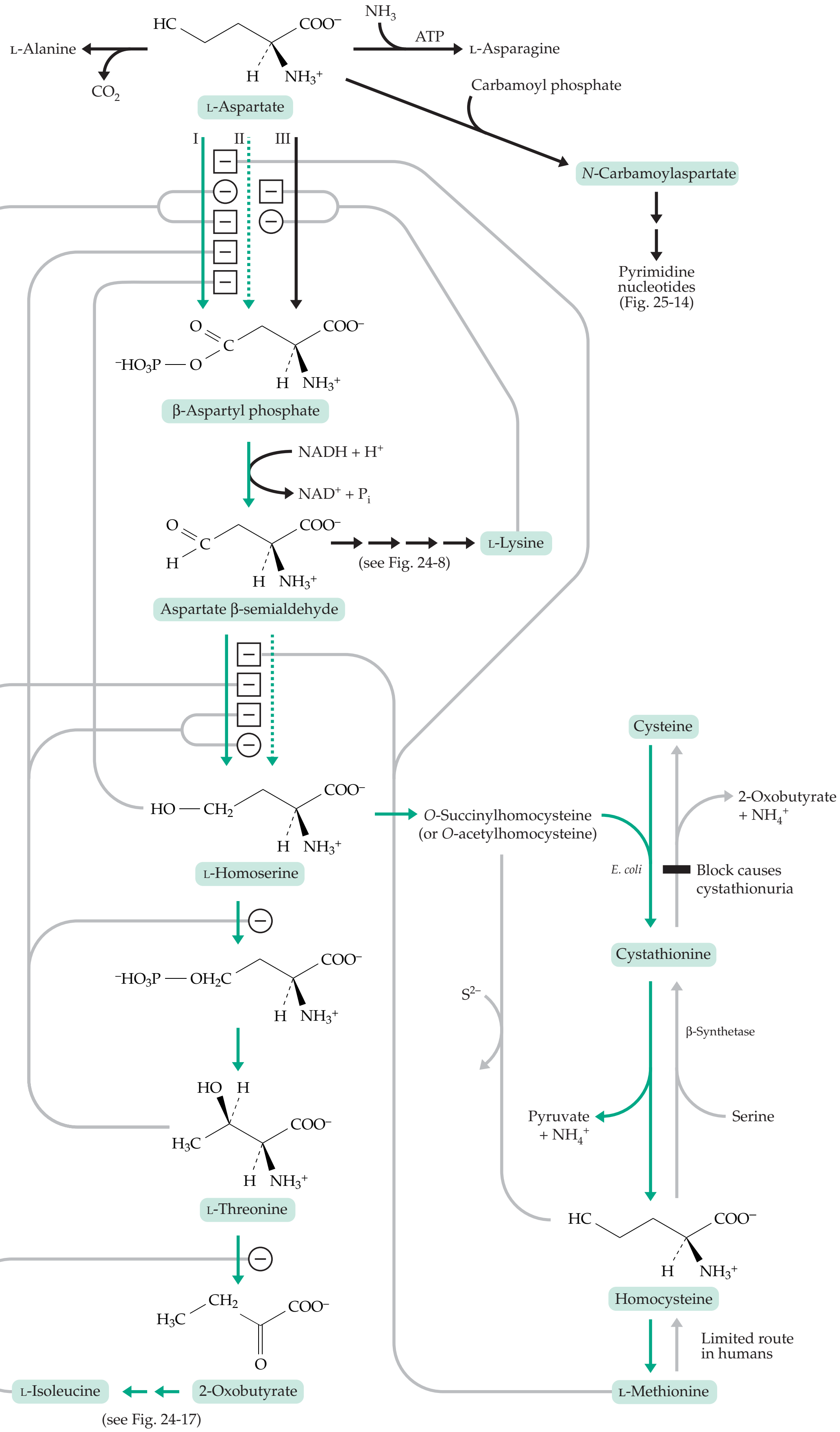
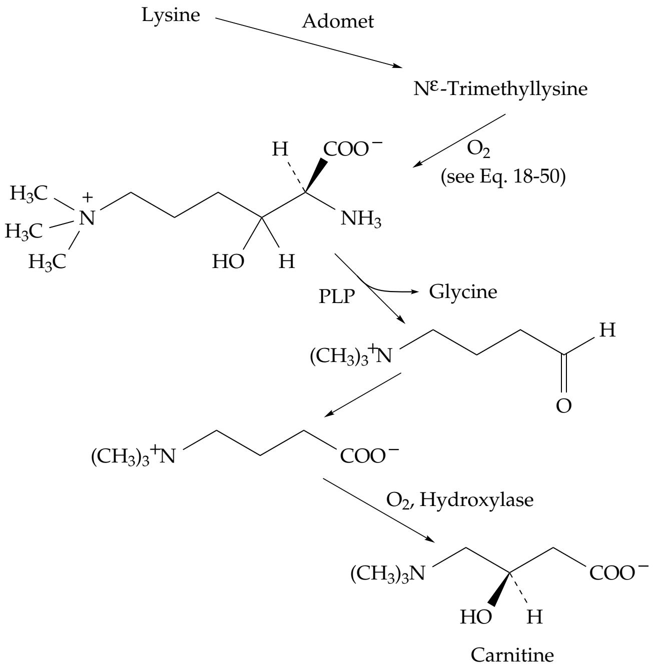

L-Alanine
CO2
NH3
ATP
L-Asparagine
Carbamoyl phosphate
L-Aspartate
N-Carbamoylaspartate
Pyrimidine nucleotides (Fig. 25-14)
I
II
III
β-Aspartyl phosphate
NADH + H+
NAD+ + Pi
L-Lysine
(see Fig. 24-8)
Aspartate β-semialdehyde
L-Homoserine
O-Succinylhomocysteine (or O-acetylhomocysteine)
Cysteine
2-Oxobutyrate + NH4+
E. coli
Block causes cystathionuria
Cystathionine
S2–
β-Synthetase
Pyruvate + NH4+
Serine
L-Threonine
Homocysteine
Limited route in humans
L-Isoleucine
2-Oxobutyrate
L-Methionine
(see Fig. 24-17)

Figure 24-13 Some biosynthetic reactions of aspartate: ⊖, feedback inhibition and ⊟, feedback repression.

that of ornithine biosynthesis, 2-oxoglutarate undergoing chain elongation (Fig. 17-18) to 2-oxoadipate followed by transamination to α-aminoadipate. This is followed by ATP-dependent reduction to the aldehyde.[264a] The final step of transamination is not accomplished in the usual way (with a PLP-dependent enzyme), but through formation of a Schiff base with glutamate and reduction to **saccharopine**.[265] Oxidation now produces the Schiff base of lysine with 2-oxoglutarate.

In the diaminopimelate pathway of lysine synthesis (Fig. 24-14) aspartate is converted to aspartate semialdehyde, and a two-carbon unit is added via aldol condensation with pyruvate.[266–269] Decarboxylation at the end of the sequence yields lysine. A series of cyclic intermediates exist, but it is noteworthy that the initial product of the aldol condensation (bracketed in Fig. 24-14) is converted to diaminopimelic acid by a simple sequence involving α,β elimination of the hydroxy group, reduction with NADPH, and transamination.[267,268,270,271] The process is complicated by the natural tendency for ring closure. All gram-negative and many gram-positive bacteria use the succinylase pathway shown in Fig. 24-14. Succinylation serves to shift the equilibrium back in favor of open-chain compounds.[272–274] Some species of *Bacillus* use acetylation in the same way, while a few bacteria manage to use a dehydrogenase to reductively aminate tetrahydropimelate to diaminopimelate.[275,276] The diaminopimelate pathway is of special significance to prokaryotic organisms for the reason that **dipicolinic acid** is formed as an important side product and because of formation of **diaminopimelic acids**. The cyclic


Lysine
Adomet
Nε-Trimethyllysine
O₂
(see Eq. 18-50)
H
COO⁻
H₃C
H₃C
H₃C
N
+
NH₃
HO
H
PLP
Glycine
(CH₃)₃⁺N
H
O
(CH₃)₃⁺N
COO⁻
O₂, Hydroxylase
(CH₃)₃N
COO⁻
HO
H
Carnitine

(24-30)

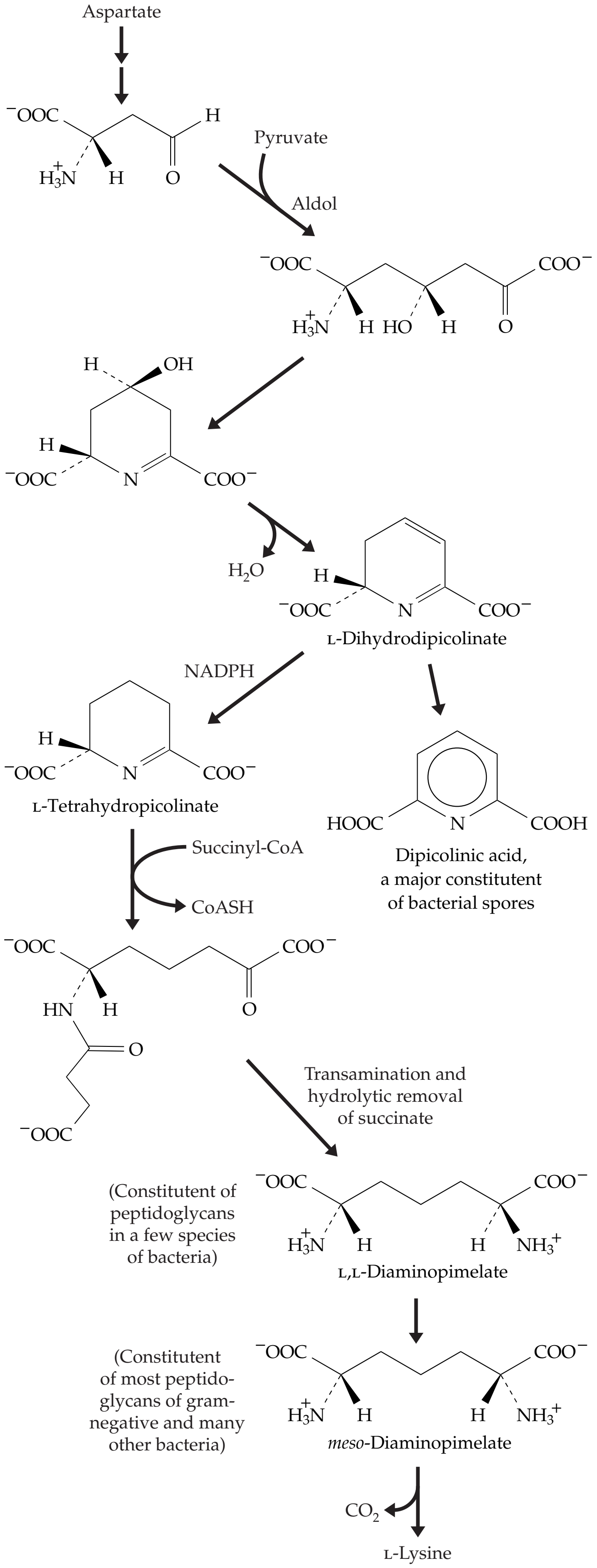
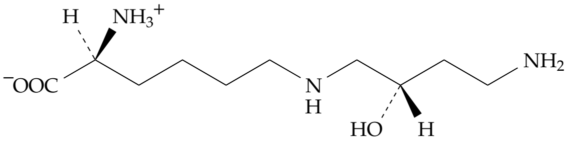

Aspartate
⁻OOC
H
H₃N⁺
H
O
Pyruvate
Aldol
⁻OOC
COO⁻
H₃N⁺
H
HO
H
O
H
OH
H
⁻OOC
N
COO⁻
H₂O
H
⁻OOC
N
COO⁻
L-Dihydrodipicolinate
NADPH
H
⁻OOC
N
COO⁻
L-Tetrahydropicolinate
HOOC
N
COOH
Dipicolinic acid,
a major constitutent
of bacterial spores
Succinyl-CoA
CoASH
⁻OOC
COO⁻
HN
H
O
O
⁻OOC
Transamination and
hydrolytic removal
of succinate
(Constituent of
peptidoglycans
in a few species
of bacteria)
⁻OOC
COO⁻
H₃N⁺
H
H
NH₃⁺
L,L-Diaminopimelate
(Constitutent
of most peptido-
glycans of gram-
negative and many
other bacteria)
⁻OOC
COO⁻
H₃N⁺
H
H
NH₃⁺
meso-Diaminopimelate
CO₂
L-Lysine

Figure 24-14 The biosynthesis of lysine by the diaminopimelate pathway.

dipicolinic acid is a major constituent of bacterial spores[277] but is rarely found elsewhere in nature. Both L, L- and meso-diaminopimelic acids[276,276a] are constituents of peptidoglycans of bacterial cell walls (Fig. 8-29).

Lysine is not only a constituent of proteins. It can also be trimethylated and converted to **carnitine** (p. 944). In mammals some specific lysyl side chains of proteins undergo *N*-trimethylation and proteolytic degradation with release of free trimethyllysine (Eq. 24-30).[278,279] The free trimethyllysine then undergoes hydroxylation by a 2-oxoglutarate–Fe^{2+}–ascorbate-dependent hydroxylase (Eq. 18-51) to form β-hydroxytrimethyllysine, which is cleaved by a PLP-dependent enzyme (Chapter 14). The resulting aldehyde is oxidized to the carboxylic acid and is converted by a second 2-oxoglutarate–Fe^{2+}–ascorbate-dependent hydroxylase to carnitine (Eq. 24-30; see also Eq. 18-50).

Hypusine (N^{ε}-(4-amino-2-hydroxybutyl)lysine)[242] occurs in mammalian initiation factor 4D, which is utilized in protein synthesis (Chapter 29) and is formed by transfer of the 4-carbon butylamine group from spermidine to a lysine side chain followed by hydroxylation.[280–282a] The lupine alkaloid lupinine[283] is formed from two C_5 units of cadaverine which arises by decarboxylation of lysine. Silaffins (pp. 178, 1381) also contain modified lysines.



Hypusine: N^{ε}-(4-amino-2-hydroxybutyl) lysine

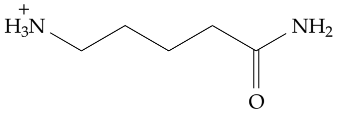
3. The Catabolism of Lysine

An unusual feature of lysine metabolism is that the α-amino group does not equilibrate with the "nitrogen pool." Catabolism is initiated by deamination and proceeds by β oxidation.[284] At least six variations of the β-oxidation process have been proposed. The evolutionary differences concern the manner in which the two amino groups are moved from the carbon skeleton. In the seemingly simplest pathway (A in Fig. 24-15), which is used by *Flavobacterium fuscum*,[285] the ε-amino group is removed in a direct (but atypical) transamination. The resulting α-aminoadipate semialdehyde is oxidized to α-aminoadipate, which *is degraded in a sequence characteristic for the catabolism of amino acids.* Transamination is followed by oxidative decarboxylation of the resulting 2-oxoacid and β oxidation of the coenzyme A derivative. A decarboxylation step by which the terminal carboxyl group is removed is interposed in the β-oxidation sequence for lysine degradation.

Perhaps the initial transamination in pathway A is chemically difficult, for most organisms use more complex sequences to form 2-oxoadipate. In pathway B (which takes place in liver mitochondria and is believed to be the predominant pathway in mammals),[286,287] the ε-amino group is reductively coupled with 2-oxoglutarate to form saccharopine. The latter is in turn oxidized on the opposite side of the bridge nitrogen to form glutamic acid and α-aminoadipate semialdehyde. The overall process is the same as direct transamination and just the opposite of that occurring in the aminoadipate pathway of biosynthesis (Fig. 24-9). Absence of one or both of these dehydrogenases causes familial hyperlysinemia.[287,288]

Pathway C has been established for *Pseudomonus putida*[289] and is also followed to some extent in both plants and animals. In most animal tissues it may be used principally for degradation of D-lysine.[290] However, it is the major L-lysine oxidation pathway in brain.[291] In a fungal parasitic species of *Rhizoctonia* L-lysine is converted to saccharopine via pathway B; then using an $NADP^+$-dependent saccharopine oxidase the sequence is shunted to pathway C.[292] L-Pipicolic acid formed in this way also gives rise to various alkaloids including the α-mannosidase inhibitor swainsonine (Fig. 20-7).[290] Pathway C, like pathway B, makes use of transamination via a reduction-oxidation sequence. It is strictly internal, the oxidizing carbonyl group being formed by transamination of the α-amino group of lysine. Pathway D, apparently used by yeasts,[293] avoids cyclic intermediates by acetylation of the ε-amino group prior to transamination. The 2-oxo group is then effectively blocked by reduction to an alcohol, the blocking group is removed from the ε-amino group, and that end of the molecule is oxidized in a straightforward way to a carboxyl group. Now the hydroxyl introduced at position 2 is presumably oxidized back to the ketone, which again can be converted to give 2-oxoadipate.

Some bacteria, e.g., *Pseudomonas putida*,[294] degrade L-lysine with a flavin-dependent oxygenase (Eq. 18-41) to δ-aminovaleramide:



The product is hydrolyzed and oxidized to **glutaryl-CoA**, rejoining the pathways shown in Fig. 24-15. A remarkable and very different approach to lysine breakdown has been developed by clostridia which obtain energy from the fermentation of Eq. 24-31:

$$\text{L-Lysine} + 2\,H_2O \longrightarrow \text{butyrate}^- + \text{acetate}^- + 2\,NH_4^+ \qquad (24\text{-}31)$$

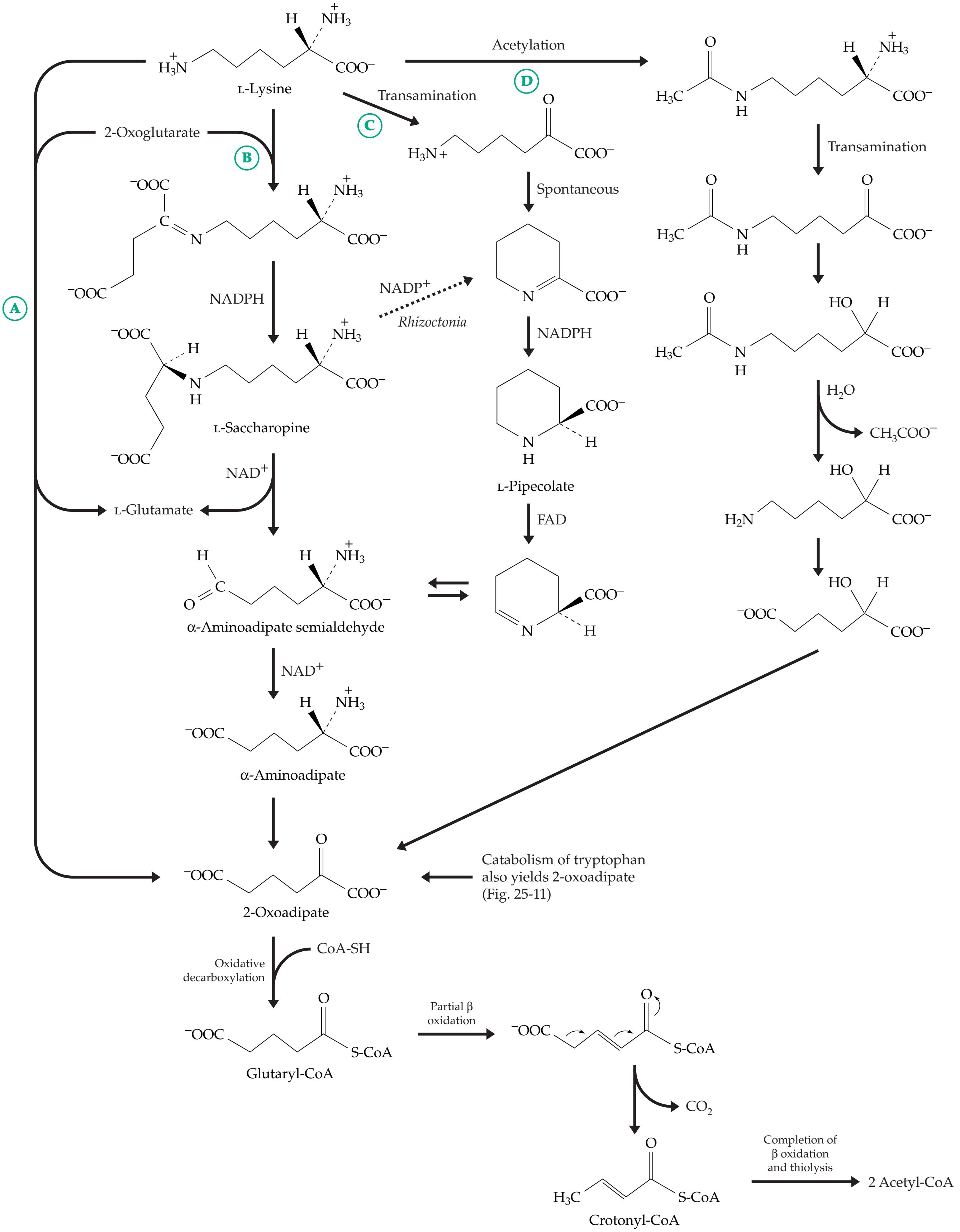
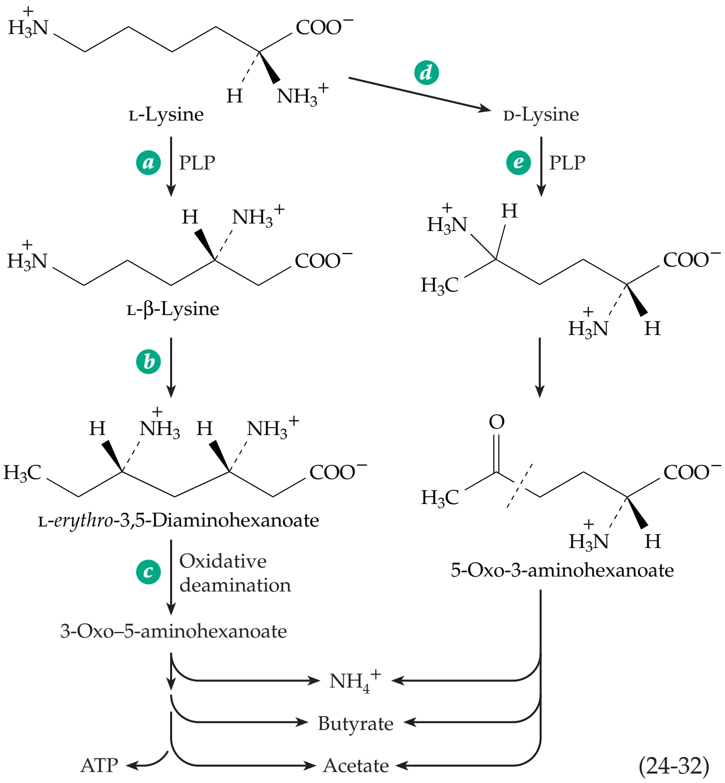

L-Lysine
Acetylation
D
Transamination
C
2-Oxoglutarate
B
A
Spontaneous
Transamination
NADPH
NADP+
Rhizoctonia
NADPH
L-Saccharopine
L-Pipecolate
H2O
CH3COO−
NAD+
L-Glutamate
FAD
α-Aminoadipate semialdehyde
NAD+
α-Aminoadipate
Catabolism of tryptophan also yields 2-oxoadipate (Fig. 25-11)
2-Oxoadipate
Oxidative decarboxylation
CoA-SH
Glutaryl-CoA
Partial β oxidation
CO2
Crotonyl-CoA
Completion of β oxidation and thiolysis
2 Acetyl-CoA

Figure 24-15 Catabolism of lysine.

The reaction is coupled to formation of one molecule of ATP from ADP and P$_i$. Two pathways have been worked out. In the first lysine is acted upon by a PLP-dependent **L-lysine 2,3-aminomutase** (Eq. 24-32, step *a*) to convert it to β-lysine (3,6-diaminohexanoate). The latter is further isomerized (Eq. 24-32, step *b*) by the vitamin B$_{12}$ and PLP-dependent β-lysine mutase. Oxidative deamination to a 3-oxo compound (Eq. 24-32, step *c*) permits chain cleavage. The reader can easily propose the remaining reactions of chain cleavage, ATP synthesis, elimination of ammonia, and balancing of the redox steps. An alternative pathway begins with a racemase (Eq. 24-32, step *d*) and isomerization of the resulting D-lysine by another B$_{12}$ and PLP-dependent enzyme (Eq. 24-32, step *e*).[294a] Oxidative deamination presumably occurs, but the mechanism for chain cleavage is not so obvious. It does occur between C-4 and C-5 as indicated by the dashed line in Eq. 24-32.

Another variation is used by *Pseudomonas* β4 (Eq. 24-31). Beta-lysine is acetylated on N-6, then undergoes transamination to a 2-oxo acid and removal of the first two carbons as acetyl-CoA. The resulting 4-aminobutyrate is then converted to succinate via succinate semialdehyde.[295]

Why are there so many pathways of lysine breakdown? The answer is probably related to the ease of spontaneous formation of cyclic intermediates as occurs in the pipecolate pathway (pathway C, Fig. 24-15). These intermediates may be too stable for efficient metabolism so the indirect pathways evolved. In the fermentation reactions additional constraints are imposed on the pathways by the need for balanced redox processes and a net Gibbs energy decrease.



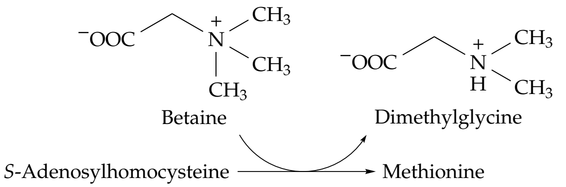
(24-32)

4. Metabolism of Homocysteine and Methionine

Autotrophic organisms synthesize methionine from aspartate as shown in the lower right side of Fig. 24-13. This involves transfer of a sulfur atom from cysteine into homocysteine, using the carbon skeleton of homoserine, the intermediate **cystathionine**, and two PLP-dependent enzymes, **cystathionine γ-synthase**[296,296a] and **cystathionine β-lyase**.[297] This **transsulfuration** sequence (Fig. 24-13, Eq. 14-33) is essentially irreversible because of the cleavage to pyruvate and NH$_4^+$ by the β-lyase. Nevertheless, this transsulfuration pathway operates in reverse in the animal body, which uses two different PLP enzymes, **cystathionine β-synthase** (which also contains a bound heme)[298–299c] and **cystathionine γ-lyase**[300] (Figs. 24-13, 24-16, steps *h* and *i*), in a pathway that metabolizes excess methionine.

For human beings methionine is nutritionally essential and comes entirely from the diet. However, the oxoacid analog of methionine can be used as a nutritional supplement. Dietary homocysteine can also be converted into methionine to a limited extent. Methionine is incorporated into proteins as such and as ***N*-formylmethionine** at the N-terminal ends of bacterial proteins (steps *a* and *b*, Fig. 24-16). In addition to its function in proteins methionine plays a major role in biological methylation reactions in all organisms. It is converted into ***S*-adenosylmethionine** (AdoMet or SAM; Fig. 24-16, step *e*; see also Eq. 17-37),[301–302b] which is the most widely used methyl group donor for numerous biological methylation reactions (Eq. 12-3). *S*-Adenosylmethionine is also the precursor of the special "wobble base" **queuine** (Fig. 5-33).[312]

The product of transmethylation, ***S*-adenosylhomocysteine**, is converted (step *g*) into homocysteine in an unusual NAD-dependent hydrolytic reaction (Eq. 15-14) by which adenosine is removed (step *g*).[302c] Homocysteine can be reconverted to methionine, as indicated by the dashed line in Fig. 24-16. This can be accomplished by the vitamin B$_{12}$- and tetrahydrofolate-dependent **methionine synthase**, (Eq. 16-43), which transfers a methyl group from methyl-tetrahydrofolate[303–303b]; by transfer of a methyl group from **betaine**, a trimethylated glycine (Eq. 24-33)[304], or by remethylation with AdoMet (Fig. 24-16).[304a]

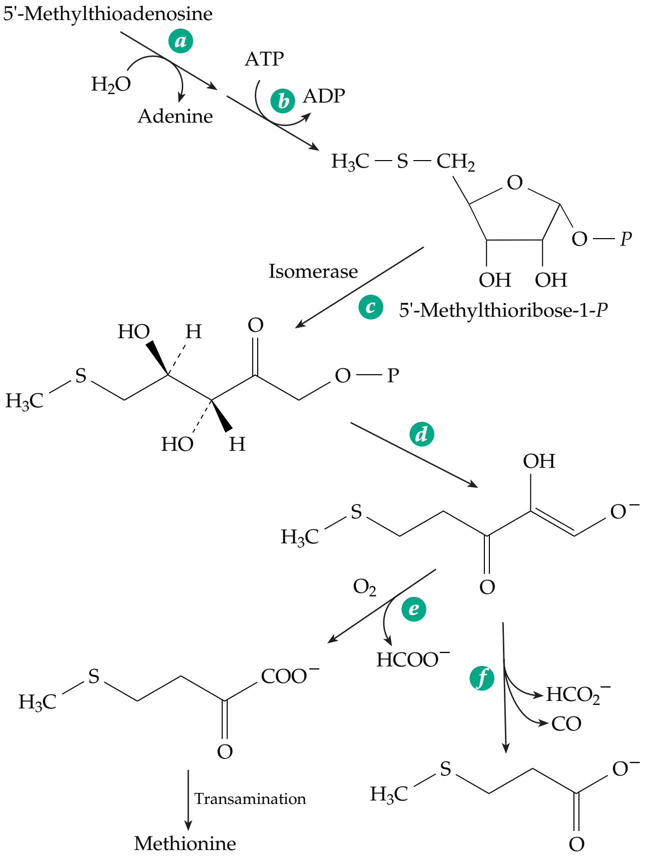


(24-33)

When present in excess methionine is toxic and must be removed. Transamination to the corresponding 2-oxoacid (Fig. 24-16, step *c*) occurs in both animals and plants. Oxidative decarboxylation of this oxoacid initiates a major catabolic pathway,[305] which probably involves β oxidation of the resulting acyl-CoA. In bacteria another catabolic reaction of methionine is γ-elimination of methanethiol and deamination to 2-oxobutyrate (reaction *d*, Fig. 24-16; Fig. 14-7).[306] Conversion to homocysteine, via the transmethylation pathway, is also a major catabolic route which is especially important because of the toxicity of excess homocysteine. A hereditary deficiency of cystathionine β-synthase is associated with greatly elevated homocysteine concentrations in blood and urine and often disastrous early cardiovascular disease.[299,307–309b] About 5–7% of the general population has an increased level of homocysteine and is also at increased risk of artery disease. An adequate intake of vitamin B_6 and especially of folic acid, which is needed for recycling of homocysteine to methionine, is helpful. However, if methionine is in excess it must be removed via the previously discussed transsulfuration pathway (Fig. 24-16, steps *h* and *i*).[310] The products are cysteine and 2-oxobutyrate. The latter can be oxidatively decarboxylated to propionyl-CoA and further metabolized, or it can be converted into leucine (Fig. 24-17) and cysteine may be converted to glutathione.[299a]

Methionine in plants can be converted to the sulfonium compound *S*-methyl-L-methionine, also called vitamin U. It has strong osmoprotectant activity and accumulates in many marine algae and some flowering plants.[311] Other organisms, including mammals, can use *S*-methylmethionine to methylate homocysteine, converting both reactants back to methionine[311a] enabling animals to meet some of their methionine need from this source.

A salvage pathway. Another product of *S*-adenosylmethionine is **5'-methylthioadenosine**, which can be formed by an internal displacement on the γ-methylene group by the carboxylate group (step *l*, Fig. 24-16). Methylthioadenosine also arises during formation of the compounds spermidine (Fig. 24-12) and ACC (Fig. 24-16). Mammalian tissues convert methylthioadenosine back to methionine by the sequence shown in Eq. 24-34. It undergoes phosphorolysis to 5'-methylthioribose whose ring is opened and converted to the 2-oxoacid analog of methionine.[313,314]



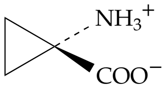
(24-34)

Step *c* of Eq. 24-34 may occur by ring opening to an enol phosphate which ketonizes to the observed product, but step *e* is a more complex multistep oxidative process.[314a,b] The last step is transamination to methionine with a glutamine-specific aminotransferase. Another enzyme from *Klebsiella* converts the same intermediate anion to methylthiopropionate, formate, and CO (Eq. 24-34, step *f*).[315]

The plant hormone ethylene. A major reaction of *S*-adenosylmethionine in plants is the formation of **ethylene**.[316,317] Ethylene has been recognized since 1858 as causing a thickening of stems of plants and a depression in the rate of elongation. In 1917, it was established that ethylene is formed in fruit and that addition of this gaseous compound hastened ripening. Ethylene is now an established plant hormone having a variety of effects including retardation of mitosis, inhibition of photosynthesis, and stimulation of respiration and of the enzyme phenylalanine ammonia-lyase (Eq. 14-45). These effects are indirectly a result of the action of ethylene on transcription of certain genes. In *Arabidopsis*, with which genetic studies are being made, ethylene binds to the N-terminal part of at least two receptor proteins, which have intracellular histidine kinase domains in the C-terminal parts.[318,319]

DNA-binding proteins specific for **ethylene-responsive elements** (EREs), having the conserved sequence AGCCGCC, are presumably phosphorylated by this kinase[320] and affect transcription of many genes. A protein homologous to one ethylene receptor of *Arabidopsis* has been identified in the tomato. A proline to leucine mutation at position 36, near the N terminus, destroys the sensitivity to ethylene and prevents ripening of this tomato. Another component in the ethylene signaling pathway in *Arabidopsis* is a protein serine / threonine kinase that resembles the mammalian raf kinase involved in the signaling cascade shown in Fig. 11-13.[318]

The formation of ethylene is often induced by the hormone **auxin** (Chapter 30), which stimulates activity of the synthase that forms **1-aminocyclopropane-1-carboxylate** (**ACC**) from *S*-adenosyl methionine (Eq. 14-27, step *j*: Fig. 24-16).[320a,b] Although ACC has been known as a minor plant product for over 25 years, it was much more recently identified as the immediate precursor of ethylene. ACC is often produced in response to stresses such as wounding, drought, or waterlogging of roots.[316,321] In the last of these cases the ACC is transferred through the xylem from the roots upward to shoots, which respond in characteristic ways to the ethylene that is released.



$$+ O_2 + \text{Ascorbate} \rightarrow H_2C = CH_2 + HCN + CO_2 + 2\,H_2O + \text{Dehydroascorbate}$$

(24-35)

The conversion of ACC to ethylene, HCN, and CO_2 is catalyzed by ACC oxidase, an Fe^{2+}-dependent enzyme of the isopenicillin-*N*-synthase (Eq. 18-52) subfamily of oxygenases. However, unlike most of

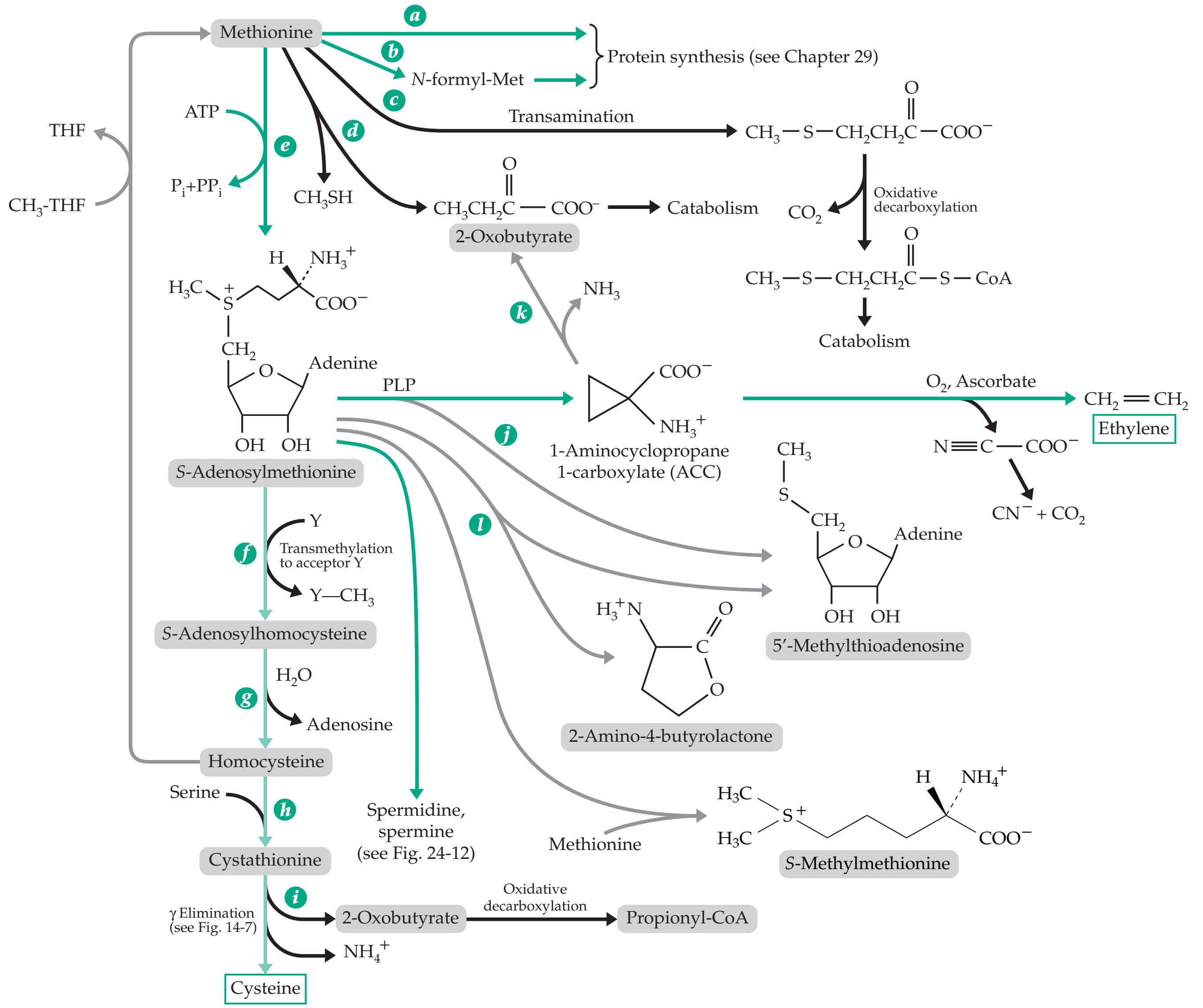
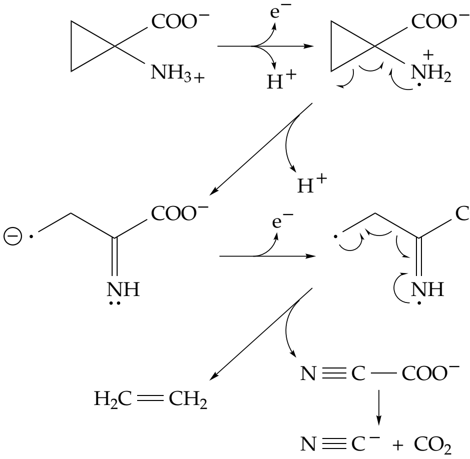

Methionine
a
b
c
d
e
N-formyl-Met
Protein synthesis (see Chapter 29)
Transamination
$CH_3-S-CH_2CH_2C(=O)-COO^-$
Oxidative decarboxylation
CO_2
$CH_3-S-CH_2CH_2C(=O)-S-CoA$
Catabolism
THF
CH_3-THF
ATP
P_i+PP_i
CH_3SH
$CH_3CH_2C(=O)-COO^-$
Catabolism
2-Oxobutyrate
k
NH_3
H
NH_3^+
H_3C
S^+
COO^-
CH_2
O
Adenine
OH OH
S-Adenosylmethionine
PLP
COO^-
NH_3^+
1-Aminocyclopropane 1-carboxylate (ACC)
O_2, Ascorbate
$CH_2=CH_2$
Ethylene
$N\equiv C-COO^-$
$CN^- + CO_2$
j
l
CH_3
S
CH_2
O
Adenine
OH OH
5′-Methylthioadenosine
f
Y
Transmethylation to acceptor Y
Y—CH_3
S-Adenosylhomocysteine
H_3^+N
C
O
O
2-Amino-4-butyrolactone
g
H_2O
Adenosine
Homocysteine
Serine
h
Spermidine, spermine (see Fig. 24-12)
Methionine
H_3C
S^+
H_3C
H
NH_4^+
COO^-
S-Methylmethionine
Cystathionine
i
γ Elimination (see Fig. 14-7)
2-Oxobutyrate
Oxidative decarboxylation
Propionyl-CoA
NH_4^+
Cysteine

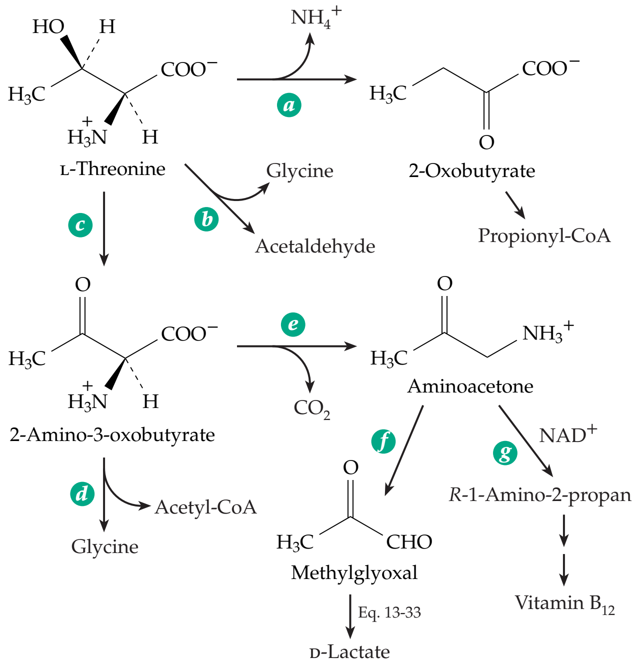
Figure 24-16 Some metabolic reactions of methionine. Biosynthetic reactions are indicated by green arrows.



(24-36)

these enzymes, which utilize 2-oxoglutarate as a co-substrate (Eq. 18-52), ACC oxidase employs ascorbate and forms HCN or cyanide ions.[322–324b] It also requires CO_2 or bicarbonate as an activator.[324,325] A radical mechanism (Eq. 24-36) is probable,[317] with two electrons from ACC and two from ascorbate being utilized to reduce O_2 to 2 H_2O.

Ethylene is rather inert, but it is metabolized slowly, some of it to ethylene glycol.[326] Plants store *N*-malonyl-ACC as a metabolically inert pool. Excess ACC can be deaminated in a PLP-dependent reaction to 2-oxobutyrate (step *k*, Fig. 24-16), a process that also occurs in bacteria able to subsist on ACC.[327,327a] There may also be other mechanisms for ethylene formation, e.g., peroxidation of lipids during scenescence of leaves.[328] See also Chapter 31, Section G.



(24-37)

5. Metabolism of Threonine

Excess threonine is degraded in several ways, one of which is a β elimination reaction catalyzed by L-threonine dehydratase (Eq. 24-37, step *a*). This PLP-requiring enzyme is produced in high amounts in *E. coli* grown on a medium devoid of glucose and oxygen. Under these circumstances the reaction provides a source of propionyl-CoA, which can be converted to propionate with generation of ATP. This **biodegradative threonine dehydratase** (threonine deaminase)[329,330] is allosterically activated by AMP, an appropriate behavior for a key enzyme in energy metabolism. A second **biosynthetic threonine dehydratase** is also produced by *E. coli*[331,332] and is specifically required for production of 2-oxobutyrate needed in the biosynthesis of isoleucine by bacteria, plants,[333] and other autotrophic organisms. In 1956, Umbarger[334] showed that this enzyme is inhibited by isoleucine, the end product of the synthetic pathway. This discovery was instrumental in establishing the concepts of feedback inhibition in metabolic regulation (Chapter 11) and of allostery.

A second catabolic reaction of L-threonine (Eq. 24-37, step *b*) is cleavage to glycine and acetaldehyde. The reaction is catalyzed by serine hydroxymethyltransferase (Eq. 14-30). Some bacteria have a very active D-threonine aldolase.[335] A quantitatively more important route of catabolism in most organisms is dehydrogenation (Eq. 24-37, step *c*)[336] to form 2-amino-3-oxobutyrate. This intermediate can be cleaved by another PLP-dependent enzyme to acetyl-CoA plus glycine (Eq. 24-38, step *d*). It can also be decarboxylated (Eq. 24-38, step *e*) to aminoacetone, a urinary excretion product, or oxidized by amine oxidases to **methylglyoxal** (Eq. 24-37, step *f*).[337] The latter can be converted to D-lactate through the action of glyoxalase (Eq. 13-33). Aminoacetone is also the source of 1-amino-2-propanol for the biosynthesis of vitamin B_{12} (Eq. 24-37, step *g*; Box 16-B).[338,338a]

E. Alanine and the Branched-Chain Amino Acids

As indicated in Fig. 24-17, pyruvate is the starting material for the formation of both L- and D-alanine and also the branched chain amino acids **valine**, **leucine**, and **isoleucine**.[339,340] The chemistry of the reactions has been discussed in the sections indicated in the figure. The first step is catalyzed by the thiamin diphosphate-dependent **acetohydroxyacid synthase** (acetolactate synthase), which joins two molecules of pyruvate or one of pyruvate and one of 2-oxobutyrate (Fig. 24-17; Fig. 14-3).[340a,b] In *E. coli* there are two isoenzymes encoded by genes *ilv B* and *ilv HI*. Both are regulated by feedback inhibition by valine, probably

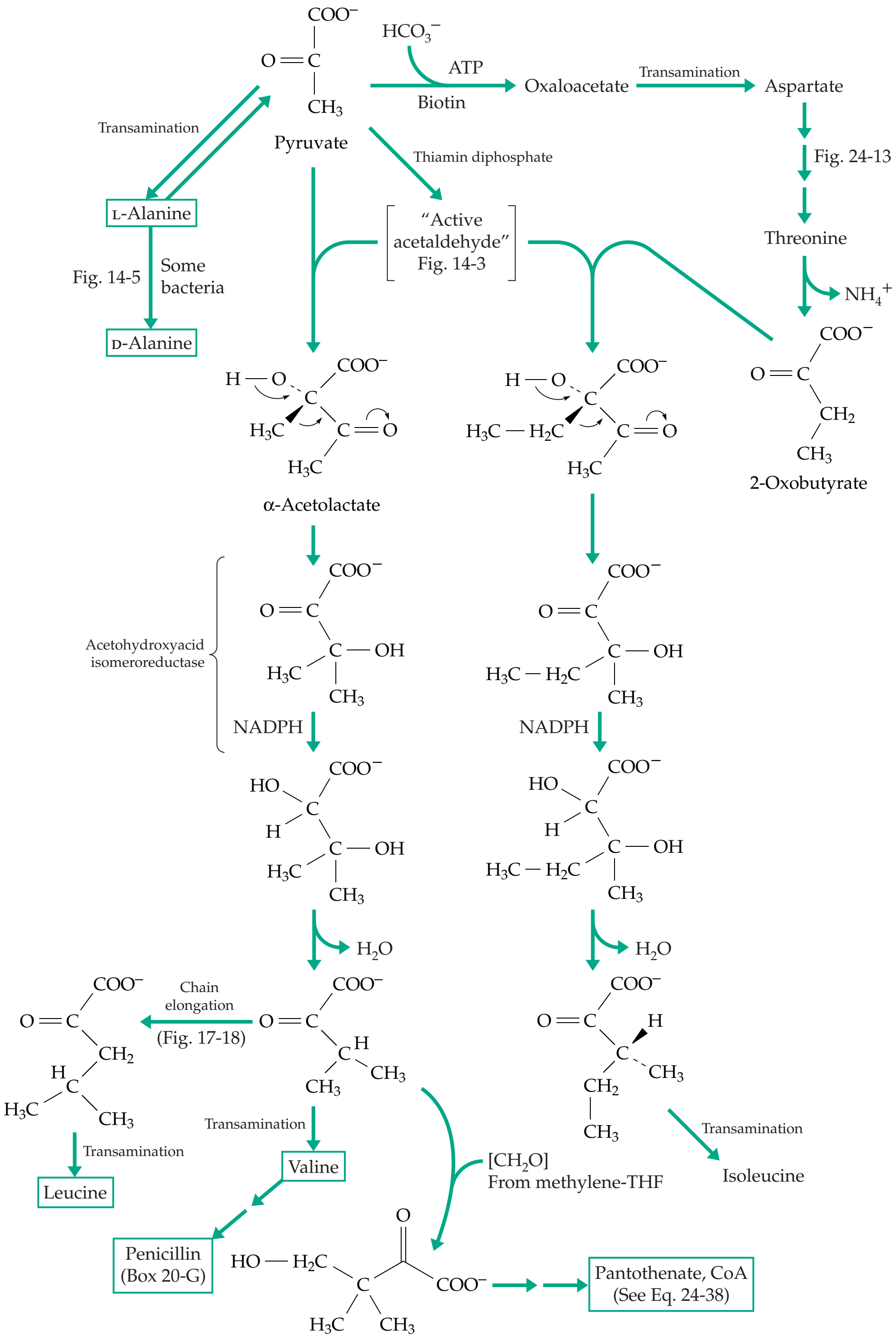
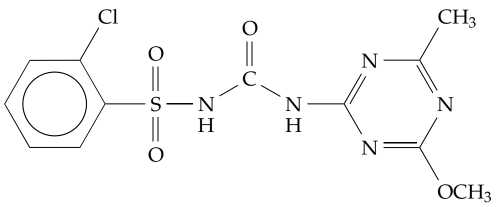
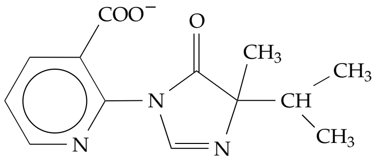

Pyruvate
HCO$_3^-$
ATP
Biotin
Oxaloacetate
Transamination
Aspartate
Fig. 24-13
Threonine
NH$_4^+$
2-Oxobutyrate
Transamination
L-Alanine
Fig. 14-5
Some bacteria
D-Alanine
Thiamin diphosphate
"Active acetaldehyde" Fig. 14-3
α-Acetolactate
Acetohydroxyacid isomeroreductase
NADPH
H$_2$O
Chain elongation
(Fig. 17-18)
Transamination
Leucine
Valine
Penicillin (Box 20-G)
[CH$_2$O] From methylene-THF
Isoleucine
Pantothenate, CoA (See Eq. 24-38)

Figure 24-17 Biosynthesis of leucine, isoleucine, valine, and coenzyme A.



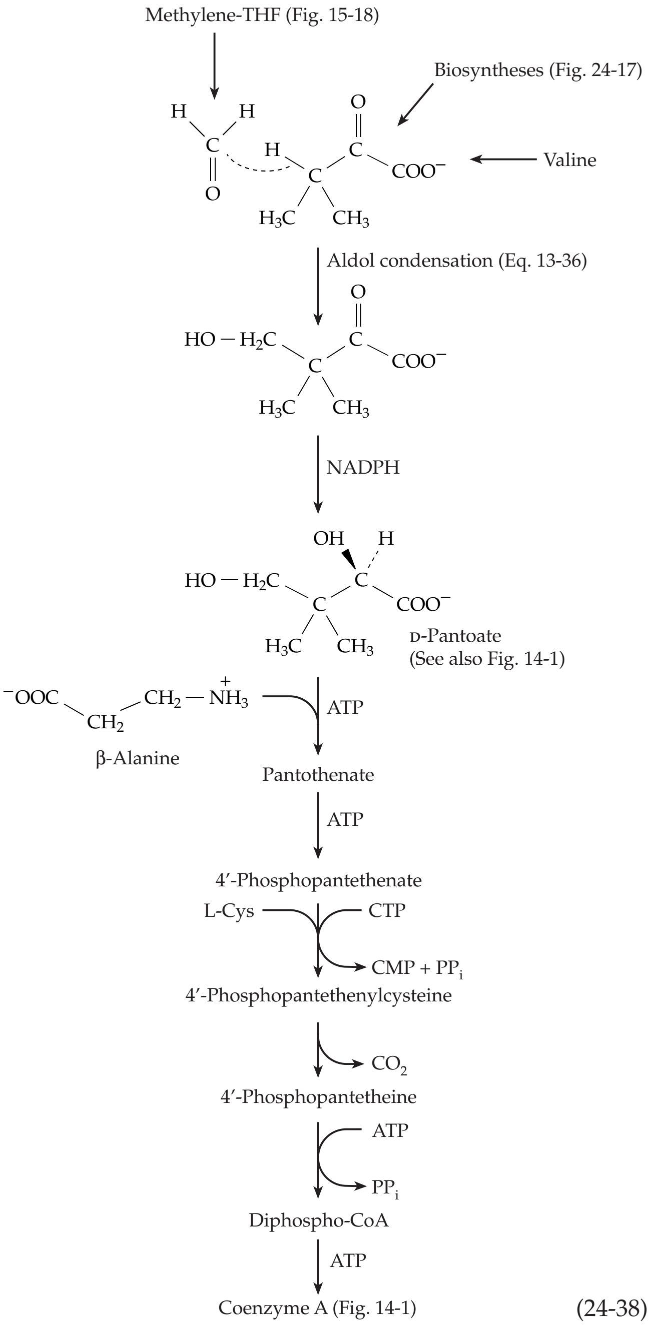
Chlorsulfuron



An imidazolinone herbicide

by an **attenuation** mechanism[341] (explained in Chapter 28). The enzymes are of some practical interest because they are specifically inhibited by two classes of herbicides, the **sulfonylureas**, of which chlorsulfuron is an example, and the **imidazolinones**.[342–345]

The second step in the synthesis, catalyzed by **acetohydroxyacid isomeroreductase**, involves shift of an alkyl group (Fig. 24-17). Neither this reaction nor the preceding one occurs in mammals. For this reason, the enzymes required are both attractive targets for herbicide design.[343,346] The third enzyme, **dihydroxy acid dehydratase**, catalyzes dehydration followed by tautomerization, resembling 6-phosphogluconate dehydratase (Eq. 13-32). The dihydroxyacid dehydratase from spinach contains an Fe_2S_2 cluster and may function by an aconitase type mechanism (Eq. 13-17).[347] In *Neurospora* isoleucine and valine are synthesized in the mitochondria.



(24-38)

While the 2-oxobutyrate needed for isoleucine formation is shown as originating from threonine in Fig. 24-17, bacteria can often make it in other ways,[348] e.g., from glutamate via β-methylaspartate (Fig. 24-8) and transamination to the corresponding 2-oxoacid. It can also be made from pyruvate by chain elongation using acetyl-CoA (Fig. 17-18); citramalate and mesaconate (Fig. 24-8) are intermediates. This latter pathway is used by some methanogens as are other alternative routes.[348] The first step unique to the biosynthetic pathway to leucine is the reaction of the 2-oxo analog of valine with acetyl-Co*A* to form **α-isopropylmalate**, the first step in a chain elongation sequence leading to the oxoacid precursor of leucine (Figs. 17-18; 24-17). The third enzyme required in the chain elongation is a decarboxylating dehydrogenase similar to isocitrate dehydrogenase.[349]

An additional series of reactions,[350] which are shown in Eq. 24-38, leads to **pantoic acid**, **pantetheine**, **coenzyme A**, and related cofactors.[350a–j] The initial reactions of the sequence do not occur in the animal body, explaining our need for pantothenic acid as a vitamin.

Alanine also gives rise to a precursor of the vitamin **biotin** (Eq. 24-39) after a PLP-dependent decarboxylative condensation with the 7-carbon dicarboxylic acid unit of pimeloyl-CoA in a reaction analogous to that of Eq. 14-32.[351] The resulting alcohol is reduced to 7-oxo-8-aminopelargonic acid which is converted by transamination, with *S*-adenosylmethionine as the nitrogen donor,[351a] to 7,8-diaminopelargonic acid. This compound undergoes a two-step ATP-dependent cyclization[352–355] to form **dethiobiotin**. The final step, insertion of sulfur into dethiobiotin, is catalyzed by **biotin synthase**, a free-radical-dependent enzyme related to pyruvate formate lyase (Fig. 15-16). It transfers the sulfur from cysteine via an Fe–S cluster.[355a–c] Biosynthesis of **lipoic acid** involves a similar insertion of two sulfur atoms into octanoic acid.[356] See also p. 1410.

1. Catabolism

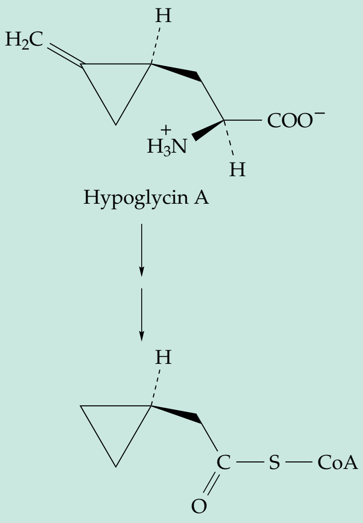
Degradation of amino acids most often begins with conversion, either by transamination[356a] or by NAD^+-dependent dehydrogenation,[357] to the corresponding 2-oxoacid and oxidative decarboxylation of the latter (Fig. 15-16). Alanine, valine, leucine, and isoleucine are all treated this way in the animal body. Alanine gives pyruvate and acetyl-CoA directly, but the others yield CoA derivatives that undergo

BOX 24-A MAPLE SYRUP URINE DISEASE AND JAMAICAN VOMITING SICKNESS

In a rare autosomal recessive condition (discovered in 1954) the urine and perspiration has a maple syrup odor.[a–c] High concentrations of the branched-chain 2-oxoacids formed by transamination of valine, leucine, and isoleucine are present, and the odor arises from decomposition products of these acids. The branched-chain amino acids as well as the related alcohols also accumulate in the blood and are found in the urine. The biochemical defect lies in the enzyme catalyzing oxidative decarboxylation of the oxoacids, as is indicated in Fig. 24-18. Insertions, deletions, and substitutions may be present in any of the subunits (Figs. 15-14, 15-15). The disease which may affect one person in ~200,000, is usually fatal in early childhood if untreated. Children suffer seizures, mental retardation, and coma. They may survive on a low-protein (gelatin) diet supplemented with essential amino acids, but treatment is difficult and a sudden relapse is apt to prove fatal. Some patients respond to administration of thiamin at 20 times the normal daily requirement. The branched-chain oxoacid dehydrogenase from some of these children shows a reduced affinity for the essential coenzyme thiamin diphosphate.[d]

Polled hereford calves in Australia develop maple syrup urine disease relatively often.[a,e] One cause was established as a mutation that introduces a stop codon that causes premature termination within the leader peptide during synthesis of the thiamin diphosphate-dependent E1 subunit. A similar biochemical defect in a mutant of *Bacillus subtilis*[f] causes difficulties for this bacterium, which requires branched-chain fatty acids in its membranes. Branched acyl-CoA derivatives are needed as starter pieces for their synthesis (Chapter 29). With the oxidative decarboxylation of the necessary oxoacids blocked, the mutant is unable to grow unless supplemented with branched-chain fatty acids.

Because persons may be born with defects in almost any gene, a variety of other problems leading to accumulation of organic acids are also known. **Methylmalonic aciduria** and propionic acidemia are discussed in Box 17-B. **Lactic acidemia** (Box 17-F) often results from a defect in pyruvate dehydrogenase. A rare defect of catabolism of leucine is **isovaleric acidemia**, a failure in oxidation of isovaleryl-CoA.[g] The symptoms of this disease are also present in the Jamaican vomiting sickness, caused by eating unripe ackee fruit. Although the ripe fruit is safe to eat, unripe fruit contains a toxin **hypoglycin A** with the following structure.[h–j] It is metabolized to an acyl-CoA derivative as shown.



Hypoglycin A

Toxic metabolite

This is an enzyme-activated inhibitor of the medium-chain fatty acyl-CoA dehydrogenase required for β oxidation of fatty acids.[j,k] The compound also inhibits isovaleryl-CoA dehydrogenase, causing an accumulation of isovaleric acid in the blood. Depression of the central nervous system by isovaleric acid in the blood could be responsible for some symptoms.[h,i] However, death from the highly fatal Jamaican vomiting sickness comes from the hypoglycemic effect. Blood glucose levels may fall as low as 0.5 mM, one-tenth the normal concentration and patients must be treated by infusion of glucose.

---

[a] Patel, M. S., and Harris, R. A. (1995) *FASEB J.* **9**, 1164–1172

[b] Chuang, D. T., and Shih, V. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1239–1278, McGraw-Hill, New York

[c] Mamer, O. A., and Reimer, M. L. J. (1992) *J. Biol. Chem.* **267**, 22141–22147

[d] Chuang, D. T., Ku, L. S., and Cox, R. P. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3300–3304

[e] Zhang, B., Healy, P. J., Zhao, Y., Crabb, D. W., and Harris, R. A. (1990) *J. Biol. Chem.* **265**, 2425–2427

[f] Willecke, K., and Pardee, A. B. (1971) *J. Biol. Chem.* **246**, 5264–5272

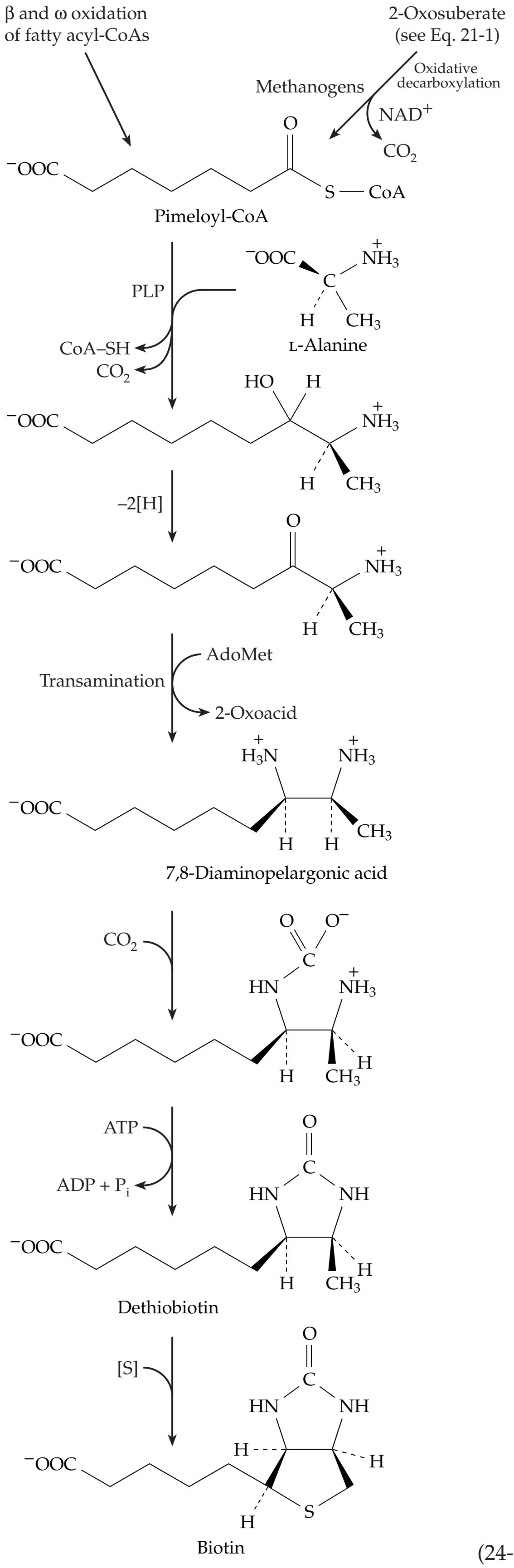
[g] Mohsen, A.-W. A., and Vockley, J. (1995) *Biochemistry* **34**, 10146–10152

[h] Tanaka, K., Isselbacher, K. J., and Shih, V. (1972) *Science* **175**, 69–71

[i] Tanaka, K. (1972) *J. Biol. Chem.* **247**, 7465–7478

[j] Lai, M.-t, Liu, L.-d, and Liu, H.-w. (1991) *J. Am. Chem. Soc.* **113**, 7388–7397

[k] Lai, M.-t, Li, D., Oh, E., and Liu, H.-w. (1993) *J. Am. Chem. Soc.* **115**, 1619–1628

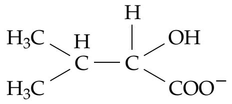


(24-39)

β-oxidation within the mitochondria[358] via the schemes shown in Fig. 24-18. There are some variations from the standard β oxidation sequence for fatty acids shown in Fig. 17-1. In the case of valine the sequence proceeds only to the stage of addition of water to form the β-hydroxy derivative. The latter is converted to free 3-hydroxyisobutyrate, and β oxidation is then completed by oxidation to methylmalonate semialdehyde.[359] The latter is oxidatively decarboxylated to form *S*-methylmalonyl-CoA.[360] Further metabolism of the latter is indicated in Fig. 17-3. However, some methylmalonate semialdehyde may be decarboxylated to propionaldehyde, which could be oxidized to propionate. Either of these compounds could then be metabolized to propionyl-CoA.[361]

In the degradation of isoleucine, β oxidation proceeds to completion in the normal way with generation of acetyl-CoA and propionyl-CoA. However, in the catabolism of leucine after the initial dehydrogenation in the β-oxidation sequence, carbon dioxide is added using a biotin enzyme (Chapter 14). The double bond conjugated with the carbonyl of the thioester makes this carboxylation analogous to a standard β-carboxylation reaction. Why add the extra CO$_2$? The methyl group in the β position blocks complete β oxidation, but an aldol cleavage would be possible to give acetyl-CoA and acetone. However, acetone is not readily metabolized further. By addition of CO$_2$ the product becomes acetoacetate, which can readily be completely metabolized through conversion to acetyl-CoA.

An alternative pathway of leucine degradation in the liver is oxidative decarboxylation by a cytosolic oxygenase to form α-hydroxyisovalerate.[362]



α-Hydroxyisovalerate

This compound may be metabolized via the valine catabolic pathway of Fig. 24-18. A third pathway, present in some bacteria, begins with the vitamin B$_{12}$-dependent isomerization of leucine to β-leucine (Chapter 16), which can undergo transamination to 3-oxoisocaproate. This can be converted to its CoA ester by a CoA transferase and can undergo β cleavage by free CoA-SH to form acetyl-CoA and isobutyryl-CoA. The latter may enter the valine catabolic pathway (Fig. 24-18). Leucine has long been known as a regulator of protein degradation in muscle.[362a–e] Dietary protein deficiency leads to especially rapid degradation of the branched-chain amino acids. The daily turnover of proteins for a 70-kg adult ingesting 70 g of protein per day has been estimated as 280 g, most of which must be reused.[362d] This large turnover can lead to

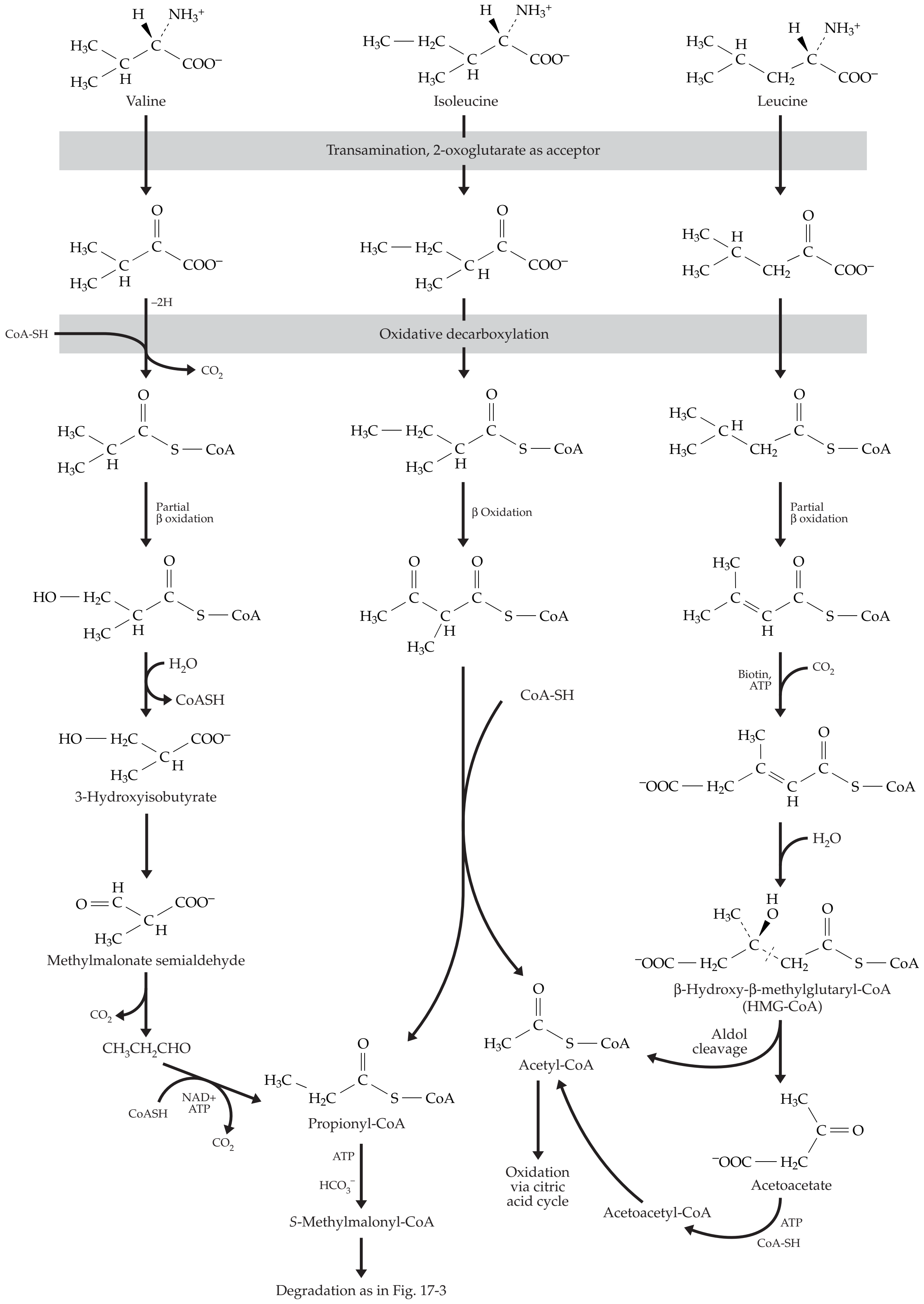

Valine
Isoleucine
Leucine
Transamination, 2-oxoglutarate as acceptor
–2H
CoA-SH
Oxidative decarboxylation
CO2
Partial β oxidation
β Oxidation
Partial β oxidation
H2O
CoASH
3-Hydroxyisobutyrate
CoA-SH
Biotin, ATP
CO2
H2O
Methylmalonate semialdehyde
β-Hydroxy-β-methylglutaryl-CoA (HMG-CoA)
CO2
CH3CH2CHO
Aldol cleavage
Acetyl-CoA
NAD+ ATP
CoASH
CO2
Propionyl-CoA
ATP
HCO3−
S-Methylmalonyl-CoA
Oxidation via citric acid cycle
Acetoacetate
Acetoacetyl-CoA
ATP
CoA-SH
Degradation as in Fig. 17-3

Figure 24-18 Catabolism of valine, leucine, and isoleucine.

excessive muscle wasting in disease states. A minor leucine metabolite found in muscle, β-hydroxy β-methylbutyrate has been proposed as a possible endogenous inhibitor of muscle breakdown.[362e,f] (See study question 17.)

Clostridium propionicum can use alanine as substrate for a balanced fermentation to form ammonium propionate, acetate, and CO_2 (Fig. 24-19).

2. Ketogenic and Glucogenic Amino Acids

According to a long-used classification amino acids are **ketogenic** if (like leucine) they are converted to acetyl-CoA (or acetyl-CoA and acetoacetate). When fed to a starved animal, ketogenic amino acids cause an increased concentration of acetoacetate and other ketone bodies in the blood and urine. On the other hand, **glucogenic** amino acids such as valine, when fed to a starved animal, promote the synthesis of glycogen (in the case of valine via methylmalonyl-CoA, succinate, and oxaloacetate). Examination of Fig. 24-18 shows that isoleucine is both ketogenic and glucogenic, a fact that was known long before the pathway of catabolism was worked out.

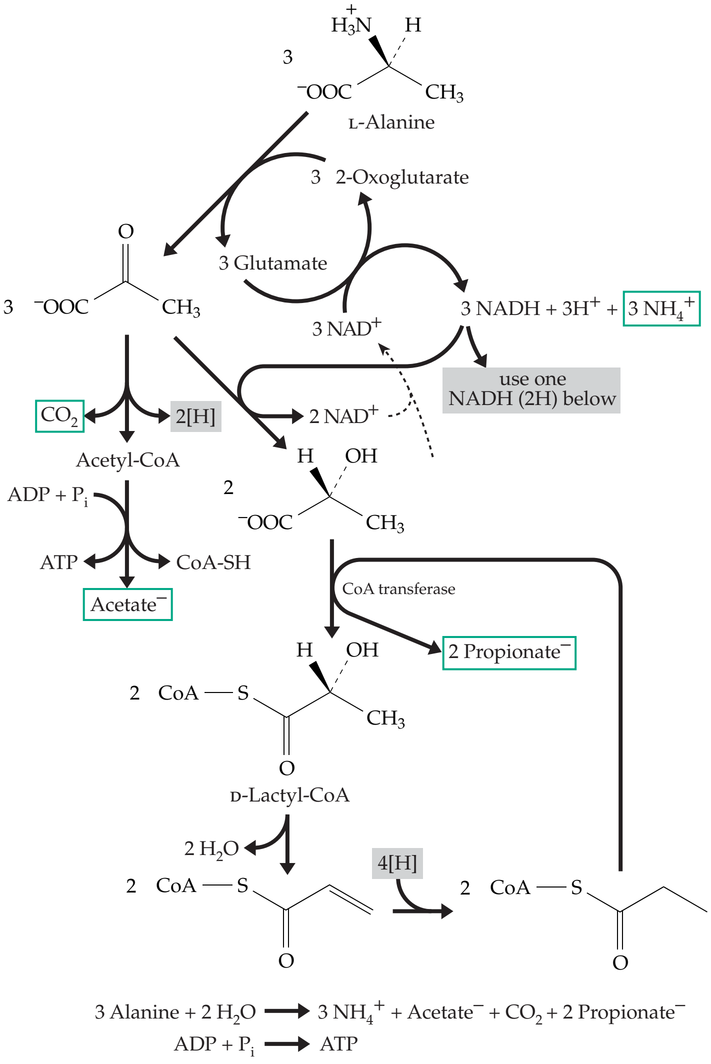


Figure 24-19 Fermentation of L-alanine by *Clostridium propionicum*. After Buckel.[363]

F. Serine and Glycine

Serine originates in a direct pathway from 3-phosphoglycerate (pathway *a*, Fig. 24-20) that involves dehydrogenation, transamination, and hydrolysis by a phosphatase. It can also be formed from glycine by the action of serine hydroxymethyltransferase (Eq. 14-30). This occurs in chloroplasts during photorespiration (Fig. 23-37)[364] and also with some methanogens and other autotropic bacteria and methylotrophs (Fig. 17-15). The glycine decarboxylase cycle shown in Fig. 15-20 provides another mechanism available in bacteria, plants, and animal mitochondria for reversible interconversion of glycine and serine. The principal route of catabolism of serine in many microorganisms is deamination to pyruvate (Fig. 24-20, step *b*),[364a] a reaction also discussed in Chapter 14 (Eq. 14-29). An alternative catabolic pathway is transamination to **hydroxypyruvate,** which as in plants (Fig. 23-37) can be reduced to D-glycerate and back to 3-phosphoglycerate.[365] That this pathway is important in human beings is suggested by the occurrence of a rare metabolic defect **L-glyceric aciduria** (or primary hyperoxaluria type II).[365–367] The biochemical defect may lie in the lack of reduction of hydroxypyruvate to D-glycerate. When hydroxypyruvate accumulates, lactate dehydrogenase effects its reduction to L-glycerate, which is excreted in large amounts (0.3–0.6 g / 24 h) in the urine. Surprisingly, the defect is accompanied by excessive production of oxalate from glyoxylate. This is apparently an indirect result of the primary defect in utilization of hydroxypyruvate. It has been suggested that oxidation of glyoxylate by NAD+ is coupled to the reduction of hydroxypyruvate by NADH.[366] This and other hyperoxalurias are very serious diseases characterized by the formation of calcium oxalate crystals in tissues and often death from kidney failure before the age of 20.

1. Biosynthetic Pathways from Serine

L-Serine gives rise to many other substances (Fig. 24-20) including **sphingosine** and the **phosphatides**. In many bacteria conversion to

O-acetyl-L-serine (step *c*, Fig. 24-20) provides for the formation of **cysteine** by a β-replacement reaction.[368–369a] Serine is also the major source of glycine (step *d*) and of the single-carbon units needed for the synthesis of methyl and formyl groups. The enzyme **serine hydroxymethyltranferase** (step *d*) also provides the principal route of formation of glycine from serine,[370,371] but a lesser portion comes via phosphatidylserine, phosphatidylethanolamine, phosphatidylcholine, and free choline (step *e*). This pathway includes decarboxylation of phosphatidylserine by a pyruvoyl group-dependent enzyme (pp. 753–755). In contrast, in green plants the major source of ethanolamine is a direct PLP-dependent decarboxylation of serine (Fig. 24-20).[371a] Because the body's capacity to generate methyl groups is limited, **choline** under many circumstances is a dietary essential and has been classified as a vitamin. However, in the presence of adequate amounts of folic acid and vitamin B_{12}, it is not absolutely required. Choline can be used to re-form phosphatidylcholine (Fig. 21-5), while an excess can be dehydrogenated to **glycine betaine,** which is one of the osmoprotectant substances in plants (Eq. 24-40).[372,373] This quaternary nitrogen compound is one of a small number of substances that, like methionine, are able to donate methyl groups to other compounds and which are also capable of methylating homocysteine to form methionine. However, the product of transmethylation from betaine, dimethylglycine, is no longer a methylating agent. The two methyl groups are removed oxidatively as formic acid to produce glycine (Eq. 24-40). A third source of glycine is transamination of glyoxylate (step *f*, Fig. 24-20). The equilibrium constant for the reaction favors glycine strongly for almost any amino donor.

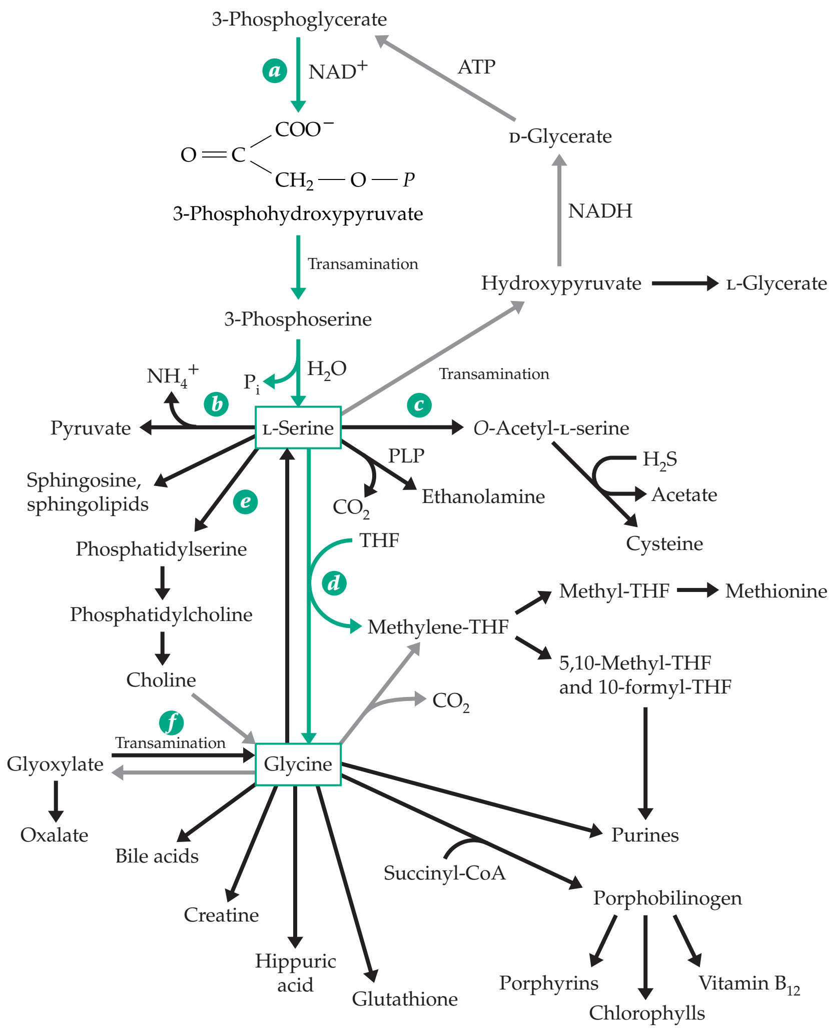
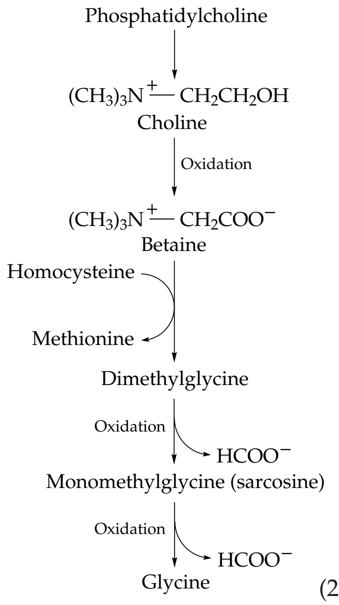

3-Phosphoglycerate
a
NAD+
ATP
D-Glycerate
COO−
O=C
CH2—O—P
3-Phosphohydroxypyruvate
NADH
Transamination
Hydroxypyruvate
L-Glycerate
3-Phosphoserine
NH4+
Pi
H2O
Transamination
b
c
Pyruvate
L-Serine
O-Acetyl-L-serine
PLP
H2S
Sphingosine, sphingolipids
Acetate
e
CO2
Ethanolamine
THF
Cysteine
Phosphatidylserine
d
Methyl-THF
Methionine
Phosphatidylcholine
Methylene-THF
5,10-Methyl-THF and 10-formyl-THF
Choline
CO2
f
Transamination
Glyoxylate
Glycine
Oxalate
Purines
Bile acids
Succinyl-CoA
Porphobilinogen
Creatine
Hippuric acid
Glutathione
Porphyrins
Vitamin B12
Chlorophylls

Figure 24-20 Metabolism of serine and glycine.

2. Metabolism of Glycine

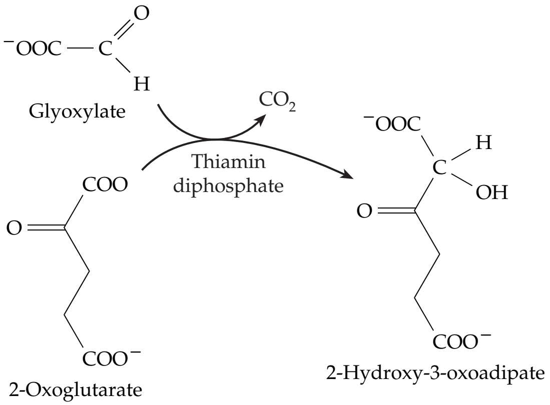
While glycine may be formed from glyoxylate by transamination, the oxidation of glycine by an amino acid oxidase (Table 15-2) permits excess glycine to be converted to glyoxylate. That this pathway, too, is quantitatively important in humans is suggested by the existence of **type 1 hyperoxaluria**.[366] It is thought that a normal pathway for utilization of glyoxylate is blocked in this condition leading to its oxidation to oxalate. The biochemical defect is frequently the absence of a liver-specific alanine:glyoxylate aminotransferase that efficiently converts accumulating glyoxylate back to glycine. In some cases the disease arises because, as a result of


Phosphatidylcholine
(CH3)3N+—CH2CH2OH
Choline
Oxidation
(CH3)3N+—CH2COO−
Betaine
Homocysteine
Methionine
Dimethylglycine
Oxidation
HCOO−
Monomethylglycine (sarcosine)
Oxidation
HCOO−
Glycine

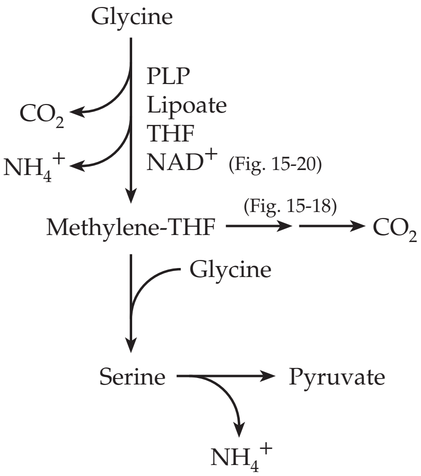
(24-40)

a mutation in its N-terminal targeting sequence the aminotransferase is targeted to mitochondria, where it functions less efficiently.[366a] Another possible defect lies in a thiamin-dependent enzyme that condenses glyoxylate with 2-oxoglutarate to form 2-hydroxy-3-oxoadipate (Eq. 24-41). The function of this reaction is uncertain, but the product could undergo decarboxylation and oxidation to regenerate 2-oxoglutarate. This would provide a cyclic pathway (closely paralleling the dicarboxylic acid cycle, Fig. 17-6) for oxidation of glyoxylate without formation of oxalate. Bear in mind that the demonstrated enzymatic condensation reactions of the rather toxic glyoxylate are numerous, and that its metabolism in most organisms is still not well understood.

An alternative route of catabolism is used by organisms such as *Diplococcus glycinophilus*, which are able to grow on glycine as a sole source of energy, of carbon, and of nitrogen,[374] and is also used in mitochondria of plants and animals.[374a] This glycine cleavage system, depicted in Fig. 15-20, involves decarboxylation, oxidation by NAD^+, release of ammonia, and transfer of the decarboxylated α-carbon of glycine to tetrahydrofolic acid (THF) to form methylene-THF. The C-1 methylene unit of the latter is used primarily for purine biosynthesis but can also be oxidized to CO_2 or can condense with another molecule of glycine (Fig. 15-18, step *c*, reverse) to form serine. This can in turn be converted to pyruvate and utilized for biosynthetic processes (Eq. 24-42).

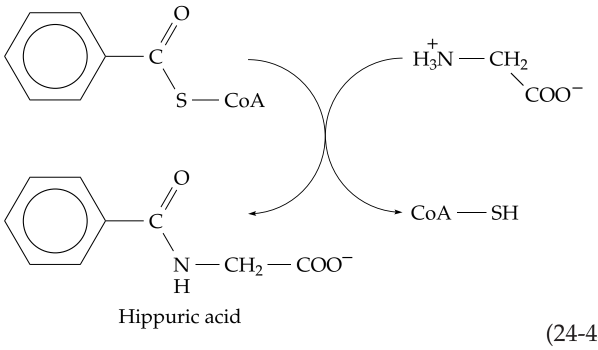


(24-41)



(24-42)

Glycine can be reduced to acetate and ammonia by the selenium-dependent clostridial glycine reductase system (Eq. 15-61). A variety of additional products can be formed from glycine as is indicated in Fig. 24-20. **Hippuric acid** (Box 10-A), the usual urinary excretion product in the "detoxication" of benzoic acid, is formed via benzoyl-CoA (Eq. 24-43):



(24-43)

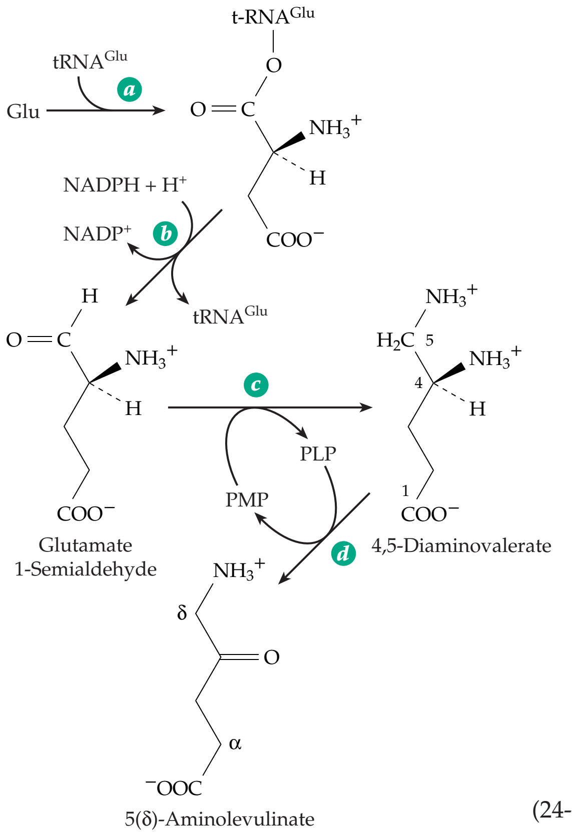
N-Methylation yields the monomethyl derivative **sarcosine**[375] and also dimethylglycine, compounds that may function as osmoprotectants (Box 20-C). Many bacteria produce **sarcosine oxidase**, a flavoprotein that oxidizes its substrate back to glycine and formaldehyde, which can react with tetrahydrofolate.[376–377a] The formation of porphobilinogen and the various pyrrole pigments derived from it and the synthesis of the purine ring (Chapter 25) represent two other major routes for glycine metabolism.

3. Porphobilinogen, Porphyrins, and Related Substances

In 1946, Shemin and Rittenberg[378] described one of the first successful uses of radiotracers in the study of metabolism. They reported that the atoms of the porphyrin ring in heme have their origins in the simple compounds acetate and glycine. As we now know, acetate is converted to succinyl-CoA in the citric acid cycle. Within the mitochondrial matrix of animal cells succinyl-CoA condenses with glycine to form **5(δ)-aminolevulinic acid** (Eq. 14-32),[379–381a] which is converted to **porphobilinogen** (Fig. 24-21), the immediate precursor to the porphyrins. The same pathways lead also to other tetrapyrroles including chlorophyll, the nickel-containing F_{430}, vitamin B_{12}, and other corrins.[382,383]

By degradation of ^{14}C-labeled porphyrins formed from labeled acetate and glycine molecules, Shemin and Rittenberg established the labeling pattern for the pyrrole ring that is indicated for porphobilinogen in Fig. 24-21. The solid circles mark those atoms that were found to be derived from methyl carbon atoms of acetate (bear in mind that acetyl groups of acetyl-CoA pass around the citric acid cycle more than once to introduce label from the methyl group of acetate into both the 2 and 3 positions of succinyl-CoA). Those atoms marked with open circles in Fig. 24-21 were found to be derived mainly from the methyl carbon of acetate and in small part from the carboxyl carbon. Atoms marked with asterisks came from glycine, while unmarked carbon atoms came from the carboxyl carbon of acetate.[384]

In cyanobacteria and in chloroplasts the intact 5-carbon skeleton of glutamate enters δ-aminolevulinate.[139,385,385a] A surprising finding was that the glutamate becomes coupled to one of the three known glutamate isoacceptor tRNAs that are utilized for protein synthesis. The aminoacyl-tRNA is formed in the usual manner with an ester linkage to the CCA end of the tRNA (Eq. 24-44; see also Eq. 17-36). This ester linkage can be reductively cleaved by NADPH to form glutamate 1-semialdehyde.[386] Isomerization of the glutamate semialdehyde to δ-aminolevulinate is accomplished by an aminomutase that is structurally and functionally related to aminotransferases.[139,387] The enzyme utilizes pyridoxamine phosphate (PMP) to transaminate the substrate carbonyl group to form 4,5-diaminovalerate plus bound PLP. A second transamination step yields the product and regenerates the PMP (Eq. 24-44, steps *c* and *d*).[387]



Glutamate 1-Semialdehyde

4,5-Diaminovalerate

5(δ)-Aminolevulinate

(24-44)

Porphyrins. As indicated in Fig. 24-21, the conversion of two molecules of 5-aminolevulinate into porphobilinogen is a multistep reaction initiated by **5-aminolevulinate dehydratase** (porphobilinogen synthase).[381,388–390b] The enzyme binds two molecules of substrate in distinct sites known as the A site and the P site (Fig. 24-21). The substrate in the P site forms a Schiff base with a lysine side chain (K247 in the *E. coli* enzyme), while a bound Zn^{2+} is thought to polarize the carbonyl of the substrate in the A site. An aldol condensation (Fig. 24-21, step *a*) ensues and is followed by dehydration to form a carbon–carbon double bond and ring closure (step *b*). Tautomerization step (*c*) leads to porphobilinogen. The enzyme is a sensitive target for poisoning by lead ions.[381,390c] Condensation to form porphyrins requires two enzymes, **porphobilinogen deaminase** (hydroxymethylbilane synthase) and **uroporphyrinogen III cosynthase**. Porphobilinogen deaminase has a bound coenzyme (prosthetic group) consisting of two linked pyrromethane groups, also derived from porphobilinogen.[391] The first step in assembling the porphyrin ring is condensation of porphobilinogen with this coenzyme (Fig. 24-21, step *d*). To initiate this step ammonia is eliminated, probably not by the direct displacements, but by electron flow from the adjacent nitrogen in the same pyrrole ring as indicated in the figure to give an exocyclic double bond. The terminal ring of the coenzyme then adds to the double bond. The condensation process is repeated four times to produce **preuroporphyrinogen** (hydroxymethylbilane).[392–394a] This intermediate is a precursor of the symmetric uroporphyrin I (Fig. 16-5). In the presence of the cosynthase a different ring-closure reaction takes place. The five-membered ring in porphobilinogen has a symmetric arrangement of double bonds. Thus, a condensation reaction can occur at either of the positions α to the ring nitrogen. A sequence of condensation, tautomerization, cleavage, and reformation of the ring as shown in steps *e* to *h* of Fig. 24-21 leads to the unsymmetric uroporphyrinogen III with its characteristic pattern of the carboxymethyl and carboxyethyl side chains. A series of decarboxylation and oxidation reactions then leads directly to protoporphyrin IX.

The first of these decarboxylations is catalyzed by the cytoplasmic **uroporphyrinogen decarboxylase,** which removes the carboxylate groups of the four acetate side chains sequentially from the D, A, B, and C rings.[395–396a] A possible mechanism, utilizing a tautomerized ring, is illustrated in the accompanying structural formula.

The decarboxylated product, coproporphyrinogen (Fig. 16-5), enters the mitochondria and is acted upon by **coproporphyrinogen oxidase,** which oxidatively

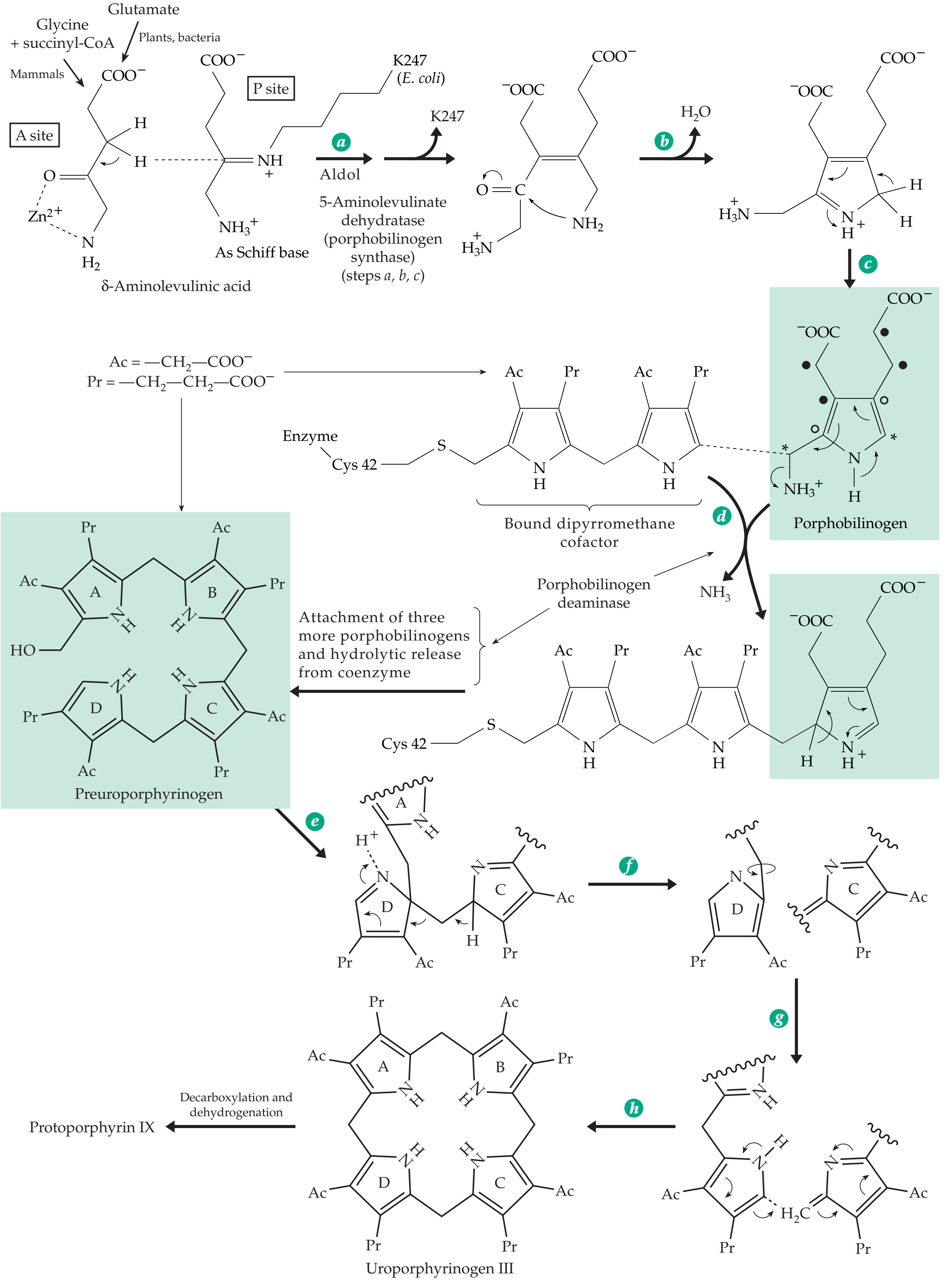
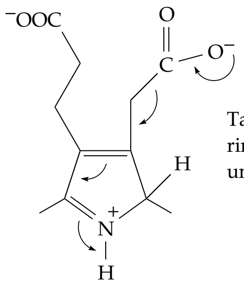

Glutamate
Glycine + succinyl-CoA
Plants, bacteria
Mammals
A site
P site
K247 (E. coli)
Zn2+
δ-Aminolevulinic acid
As Schiff base
Aldol
5-Aminolevulinate dehydratase (porphobilinogen synthase) (steps a, b, c)
H2O
a
b
c
Ac = —CH2—COO−
Pr = —CH2—CH2—COO−
Enzyme
Cys 42
Bound dipyrromethane cofactor
Porphobilinogen
d
NH3
Porphobilinogen deaminase
Attachment of three more porphobilinogens and hydrolytic release from coenzyme
Preuroporphyrinogen
e
f
g
h
Uroporphyrinogen III
Decarboxylation and dehydrogenation
Protoporphyrin IX

Figure 24-21 Biosynthesis of porphyrins, chlorins, and related compounds.



Tautomerized pyrrole ring of uroporphyrinogen undergoing decarboxylation

decarboxylates and oxidizes two of the propyl side chains to vinyl groups.[397] A flavoprotein, **protoporphyrinogen oxidase**, oxidizes the methylene bridges between the pyrrole rings[398,399] to form protoporphyrin IX. A somewhat different pathway from uroporphyrinogen is followed by sulfate-reducing bacteria.[400]

Ferrochelatase (protoheme ferro-lyase)[401–403] inserts Fe^{2+} into protoporphyrin IX to form heme. The enzyme is found firmly bound to the inner membrane of mitochondria of animal cells, chloroplasts of plants, and chromatophores of bacteria. While Fe^{2+} is apparently the only metallic ion ordinarily inserted into a porphyrin, the Zn^{2+} protoporphyrin chelate accumulates in substantial amounts in yeast, and Cu^{2+}–heme complexes are known (p. 843). Ferrochelatase, whose activity is stimulated by Ca^{2+}, appears to be inhibited by lead ions, a fact that may account for some of the acute toxicity of lead.[404]

Heme *b* is utilized for formation of hemoglobin, myoglobin, and many enzymes. It reacts with appropriate protein precursors to form the cytochromes *c*. Heme *b* is converted by prenylation to heme *o*[405] and by prenylation and oxidation to heme *a*.[405a] The porphyrin biosynthetic pathway also has a number of branches that lead to formation of corrins, chlorins, and chlorophylls as shown schematically in Fig. 24-22.

Corrins. The formation of vitamin B_{12}, other corrins, siroheme, and related chlorin chelates[406,407] requires a ring contraction with elimination of the methine bridge between rings A and D of the porphyrins (see Box 16-B). It is natural to assume that the methyl group at C-1 of the corrin ring might arise from the same precursor carbon atom as does the methine bridge in porphyrins, and it is easy to visualize a modified condensation reaction by which ring closure at step *e* in Fig. 24-21 occurs by nucleophilic addition to a C=N bond of ring A. However, ^{13}C-NMR data ruled out this possibility. When vitamin B_{12} was synthesized in the presence of ^{13}C-methyl-containing methionine, it was found that seven methyl groups contained ^{13}C. All of the "extra" methyl groups around the periphery of the molecule as well as the one at C-1 were labeled.[408] Other experiments established uroporphyrinogen III as a precursor of vitamin B_{12}. Therefore, it appeared that the ring first closed in a normal way and then reopened between rings A and D with removal of the carbon that forms the methylene bridge.[409] This turned out to be true, but with some surprises.

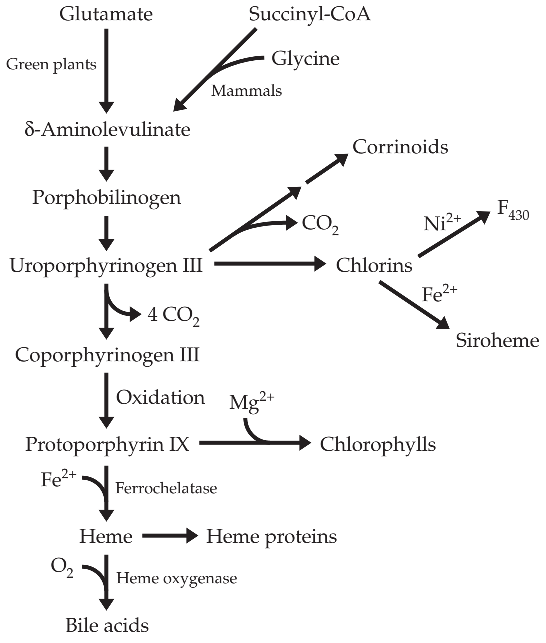


Figure 24-22 Abbreviated biosynthetic pathways from δ-aminolevulinate to heme proteins, corrins, chlorophylls, and related substances.

The complex pathways of corrin synthesis have been worked out in detail.[410–415a] This has been possible because of extensive use of ^{13}C and ^{1}H NMR and because the group of ~20 enzymes required has been produced in the laboratory from genes cloned from *Pseudomonas denitrificans*.[410] The first alterations of uroporphyrinogen are AdoMet-dependent methylations on carbon atoms. Surprisingly, one of these is on the bridge atom that is later removed. The details, including the insertion of Co^{2+} by a **cobaltochelatase**, are described by Battersby[410] and portrayed in Michal's *Biochemical Pathways*.[416]

Chlorophyll. The pathway of chlorophyll synthesis has been elucidated through biochemical genetic studies of *Rhodobacter spheroides*[417–418a] which produces bacteriochlorophyll, from studies of cyanobacteria,[419,420] and from investigations of green algae and higher plants,[421] which make chlorophyll *a*. The first step in the conversion of protoporphyrin IX into chlorophyll is the insertion of Mg^{2+} (Fig. 24-23, step *a*). This reaction does not occur readily spontaneously but is catalyzed by an ATP-dependent **magnesium protoporphyrin chelatase**.[419,422] Subsequently, the carboxyethyl side chain on ring C undergoes methylation (Fig. 24-23, step *b*) and β oxidation (step *c*).

Oxidative closure of ring E (step *d*) is followed by reduction of the vinyl group of ring B and of the double bond in ring D to form **chlorophyllide *a***. The latter is coupled with phytol, via phytyl diphosphate, to form chlorophyll *a*.[420] Chlorophyll *b* is derived from chlorophyll *a*, evidently by action of an as yet uncharacterized oxygenase, which converts the methyl group on ring B into a formyl group.[423,424] Bacteriochlorophylls also arise from chlorophyllide *a* and involve reduction of the double bond in ring B.[416,418,420] Most photosynthetic bacteria make bacteriochlorophylls esterified with the C_{20} phytol, but some substitute the unsaturated C_{20} geranylgeranyl group and a variety of other isoprenoid alcohols.

The porphyrias. The human body does not use all of the porphobilinogen produced, and a small amount is normally excreted in the urine, principally as coproporpyrins (Fig. 16-5). In a number of hereditary and acquired conditions blood porphyrin levels are elevated and enhanced urinary excretion (porphyria) is observed.[425–427] Porphyrias may be mild and almost without symptoms, but the intensely fluorescent free porphyrins are sometimes deposited under the skin and cause photosensitivity and ulceration. In extreme cases, in which the excreted porphyrins may color the urine a wine red, patients may have acute neurological attacks and a variety of other symptoms. Lucid accounts of such symptoms, experienced by King George III of England, have been written.[426,428] However, there are doubts about the conclusion that the king suffered from porphyria.[425]

Porphyria may result from several different enzyme deficiences in the porphyrin biosynthetic pathway. The condition is often hereditary but may be induced by drugs or other xenobiotic substances and may be continuous or intermittent.[425,426,429] In one type of congenital porphyria uroporphyrin I is excreted in large quantities. The biochemical defect appears to be a deficiency of the cosynthase that is required for formation of protoporphyrin IX. Another type of porphyria results from overproduction in the liver of

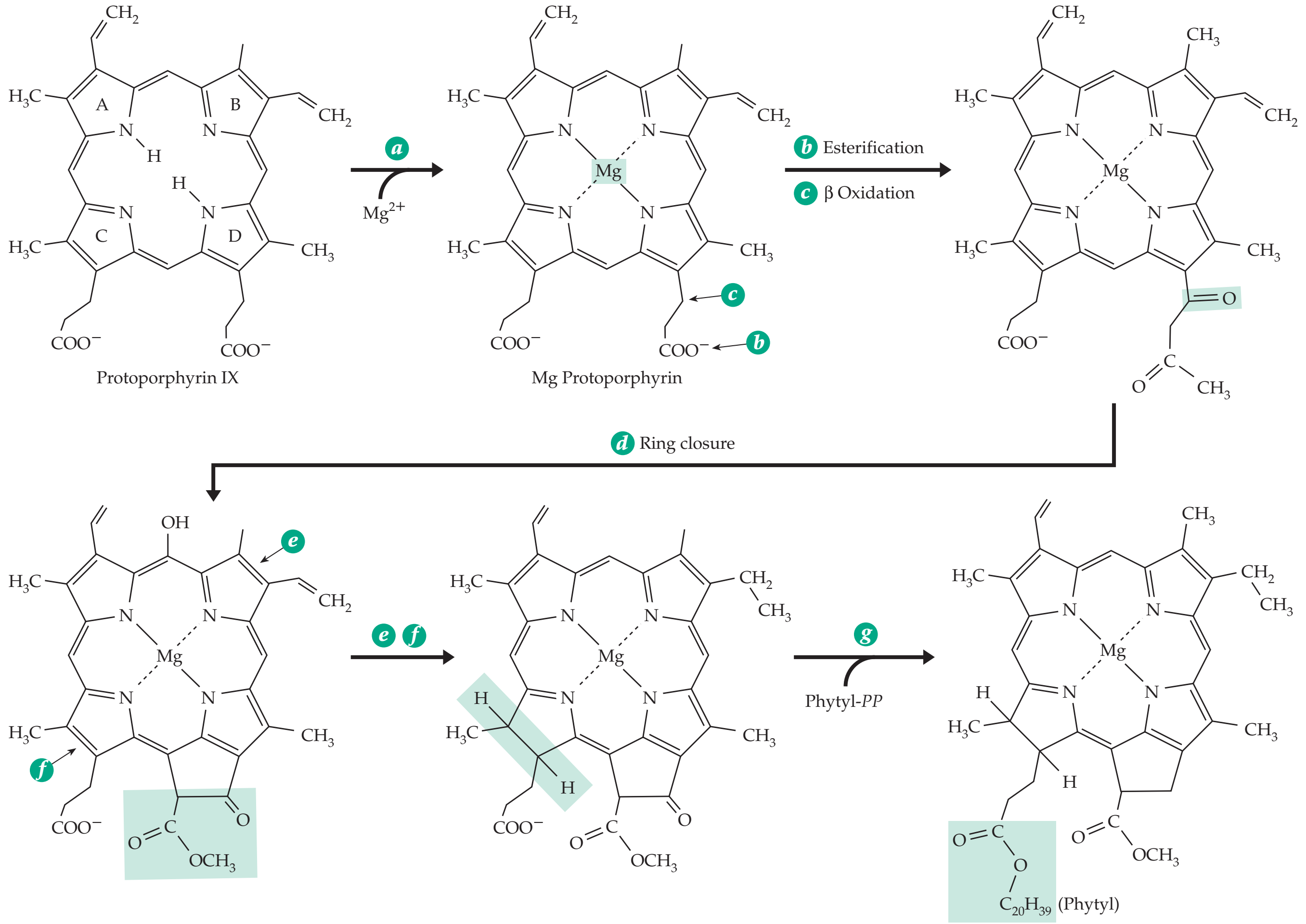
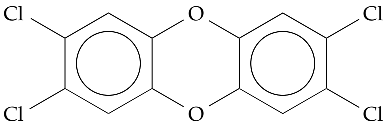


Figure 24-23 Outline of the biosynthetic pathways for conversion of protoporphyrin IX into the chlorophylls and bacteriochlorophylls. After Bollivar *et al.*[417]

δ-aminolevulinic acid, a compound with neurotoxic properties, possibly as a result of its similarity to the neurotransmitter γ-aminobutyrate.[426,427] This may account for some of the neurological symptoms of porphyria.

Some mild forms of intermittent porphyria may go unrecognized. However, ingestion of drugs can precipitate an acute attack, probably by inducing excessive synthesis of δ-aminolevulinate synthase. Among compounds having this effect are hexachlorobenzene and tetrachlorodibenzodioxin.



2,3,7,8-Tetrachlorodibenzo-*p*-dioxin

The latter is one of the most potent inducers of the synthase known.[430] The tendency for this dioxin to be present as an impurity in the herbicide 2,4,5-trichlorophenoxyacetic acid (2,4,5-T) has caused concern. For rodents this dioxin may be the most toxic small molecule known, the oral LD_{50} for guinea pigs being only 1 µg / kg body weight.[430] However, it is over 1000 times less toxic to humans.[430a] It is also a potent teratogenic agent. Synthesis of porphyrins in the liver is controlled by δ-aminolevulinate synthase. This key enzyme is sensitive to feedback inhibition by heme, but the increased synthesis of the enzyme induced by drugs can override the inhibition. Several times as much heme is synthesized in the erythroid cells of bone than in liver, but this is not subject to feedback regulation or to stimulation by drugs.[427] Heme is a potent and toxic regulator. Malaria mosquitos, which utilize blood for food but do not have a heme oxygenase (Fig. 24-24), detoxify heme by inducing its aggregation into an insoluble hydrogen-bonded solid known as β-hematin.[431]

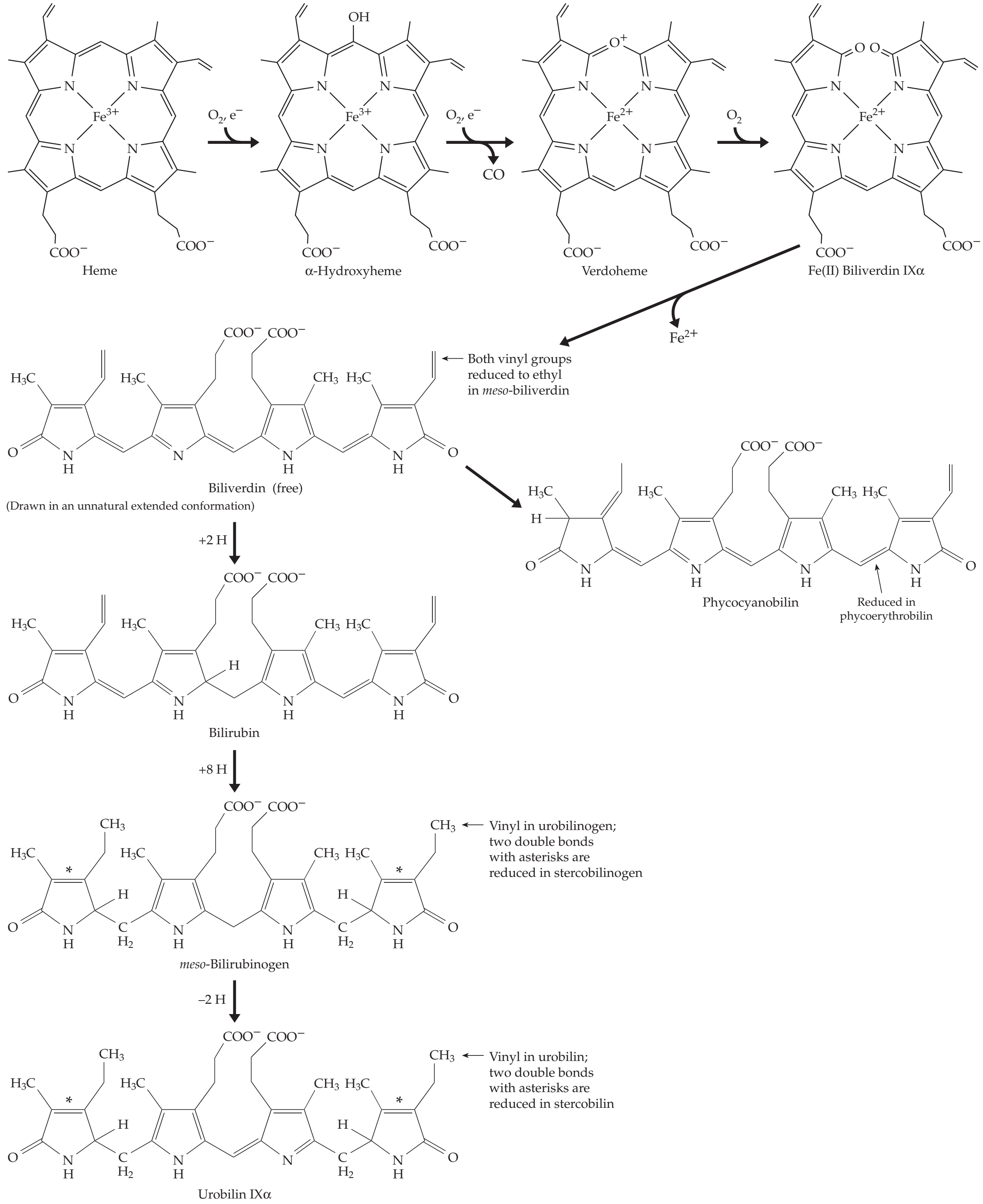
The bile pigments. The enzymatic degradation of heme is an important metabolic process if only because it releases iron to be reutilized by the body. Some of the pathways are illustrated in Fig. 24-24. The initial oxidative attack is by the microsomal **heme oxygenases,**[432–434a] which catalyze the uptake of three molecules of O_2, formation of CO, and release of the chelated Fe. The electron transport protein NADPH-cytochrome P450 reductase brings electrons from NADPH to the oxygenase. An enzyme–substrate heme complex is formed with the oxygenase. Then the Fe^{3+} is reduced to Fe^{2+} which binds O_2 as in myoglobin or hemoglobin. The complex hydroxylates its own heme α-carbon (Fig. 24-24), the other oxygen atom being reduced to OH^- by the Fe^{2+} and an additional electron from NADPH. The same enzyme catalyzes the next steps in which the α carbon is split out as CO by reaction with two molecules of O_2 to form the open tetrapyrrole dicarbonyl compound **biliverdin**, one of the bile pigments (Fig. 24-24).[435–435a] When $^{18}O_2$ was used, it was found that the biliverdin contains two atoms of ^{18}O, and that the CO contains one. Heme from the cytochromes *c* appears to be degraded by the same enzymes after proteolytic release from the proteins to which it is bound.[436]

There are two human heme oxygenases. The first (HO-1) is synthesized principally in the liver and spleen. Its formation is strongly induced by heme. The second heme oxygenase (HO-2) is distributed widely among tissues, but it is most abundant in certain neurons in the brain.[437,437a] Its major function may be to generate CO, which is now recognized as a probable neurohormone (Chapter 30). Bacteria, such as *Corynebacterium diphtheriae,* employ their own heme oxygenase as a means of recovering iron that they need for growth.[438]

A large number of other open tetrapyrroles can be formed from biliverdin by reduction or oxidation reactions. Within our bodies biliverdin is reduced to **bilirubin**, which is transported to the liver as a complex with serum albumin. In the liver bilirubin is converted into glucuronides (Eq. 20-16), glycosylation occurring on the propionic acid side chains.[439] A variety of these bilirubin conjugates are excreted into the bile. In the intestine they are hydrolyzed back to free bilirubin, which is reduced by the action of intestinal bacteria to urobilinogen, stercobilinogen, and *meso*-bilirubinogen. These compounds are colorless but are readily oxidized by oxygen to **urobilin** and **stercobilin**. Some of the urobilin and other bile pigments is reabsorbed into the blood and excreted into the urine where it provides the familiar yellow color.

The yellowing of the skin known as **jaundice** can occur if the heme degradation system is overburdened (e.g., from excessive hemolysis), if the liver fails to conjugate bilirubin, or if there is obstruction of the flow of heme breakdown products into the intestinal tract. Bilirubin is toxic, and continued exposure to excessive bilirubin levels can cause brain damage.[434,439] Bilirubin has a low water solubility and tends to form complexes with various proteins, perhaps partly because it assumes folded conformations rather than the linear one shown in Fig. 24-24.[440] These properties make it difficult to excrete. Thousands of newborn babies are treated for jaundice every year by prolonged irradiation with blue or white light which isomerizes 4*Z*,15*Z* bilirubin to forms that are more readily transported, metabolized, and excreted.[441] A more difficult problem is posed by the fatal deficiency of the glucuronosyl transferase responsible for formation of bilirubin glucuronide. Efforts are being made to develop a genetic therapy.[439]

The open tetrapyrroles of algae and the chromophore of phytochrome (Chapter 23) are all derived from **phycoerythrobilin**, which is formed from biliverdin, as indicated in Fig. 24-24. The animal bile pigments have not been found in prokaryotes. However, *Clostridium tetanomorphum*, which accumulates uroporphyrinogen III, a precursor to vitamin B_{12}, and does not synthesize protoporphyrin IX, makes a blue bile pigment **bactobilin**. This is a derivative of uroporphyrin rather than of protoporphyrin.[442]


Heme
O_2, e^-
α-Hydroxyheme
O_2, e^-
CO
Verdoheme
O_2
Fe(II) Biliverdin IXα
Fe^{2+}
Biliverdin (free)
+2 H
Bilirubin
+8 H
meso-Bilirubinogen
–2 H
Urobilin IXα
Phycocyanobilin

Both vinyl groups reduced to ethyl in *meso*-biliverdin

(Drawn in an unnatural extended conformation)

Reduced in phycoerythrobilin

Vinyl in urobilinogen; two double bonds with asterisks are reduced in stercobilinogen

Vinyl in urobilin; two double bonds with asterisks are reduced in stercobilin

Figure 24-24 The degradation of heme and the formation of open tetrapyrrole pigments.

G. Cysteine and Sulfur Metabolism

Cysteine not only is an essential constituent of proteins but also lies on the major route of incorporation of inorganic sulfur into organic compounds.[443] Autotrophic organisms carry out the stepwise reduction of sulfate to sulfite and sulfide (H_2S). These reduced sulfur compounds are the ones that are incorporated into organic substances. Animals make use of the organic sulfur compounds formed by the autotrophs and have an active oxidative metabolism by which the compounds can be decomposed and the sulfur reoxidized to sulfate. Several aspects of cysteine metabolism are summarized in Fig. 24-25. Some of the chemistry of inorganic sulfur metabolism has been discussed in earlier chapters. Sulfate is reduced to H_2S by sulfate-reducing bacteria (Chapter 18). The initial step in *assimilative* sulfate reduction, used by autotrophs including green plants and *E. coli*, is the formation of adenosine 5'-phosphosulfate (APS) (step *a*, Fig. 24-25; see also Eq. 17-38).[444–444c] The sulfate-reducing bacteria reduce adenylyl sulfate directly to sulfite (Eq. 18-32, step *b*), but the assimilative pathway of reduction in *E. coli* proceeds through 3'-phospho-5'-adenylyl sulfate (PAPS), a compound whose function as "active sulfate" has been considered in Chapter 17. Reduction of PAPS to sulfite (Fig. 24-25, step *d*) is accomplished by an NADPH-dependent enzyme.

The same pathway is found in the alga *Chlorella*, but a second route of sulfate reduction occuring in green plants may be more important.[445] Adenylyl sulfate transfers its sulfo group to a thiol group of a carrier (Eq. 24-43, step *a*). The resulting thiosulfonate is reduced by a ferredoxin-dependent reductase. Finally, a sulfide group is transferred from the $–S—S^–$ group of the reduced carrier directly into cysteine in a

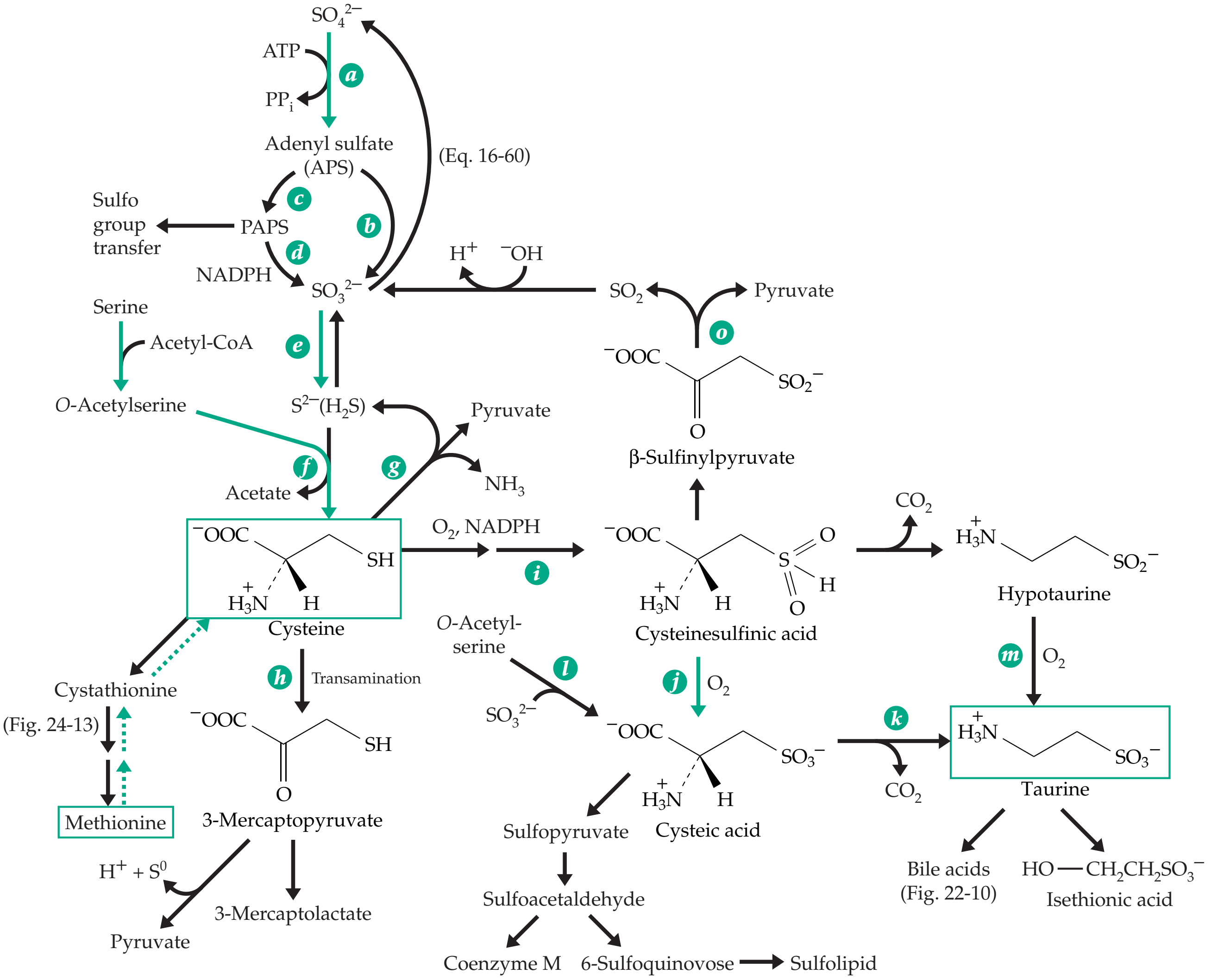
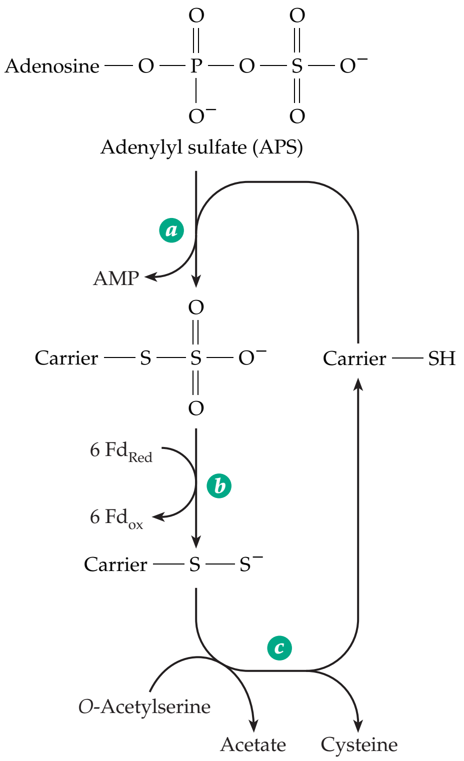

SO_4^{2-}
ATP
PP$_i$
a
Adenyl sulfate
(APS)
(Eq. 16-60)
Sulfo
group
transfer
PAPS
c
b
d
NADPH
SO_3^{2-}
H$^+$
$^-$OH
SO_2
Pyruvate
o
Serine
Acetyl-CoA
e
$^-$OOC
SO_2^-
O
β-Sulfinylpyruvate
O-Acetylserine
S^{2-}(H_2S)
Pyruvate
NH$_3$
g
f
Acetate
$^-$OOC
SH
$H_3\overset{+}{N}$
H
Cysteine
O_2, NADPH
i
$^-$OOC
S
O
H
O
$H_3\overset{+}{N}$
H
Cysteinesulfinic acid
CO_2
$H_3\overset{+}{N}$
SO_2^-
Hypotaurine
m
O_2
Cystathionine
(Fig. 24-13)
Methionine
h
Transamination
$^-$OOC
SH
O
3-Mercaptopyruvate
O-Acetyl-
serine
l
SO_3^{2-}
j
O_2
$^-$OOC
SO_3^-
$H_3\overset{+}{N}$
H
Cysteic acid
k
CO_2
$H_3\overset{+}{N}$
SO_3^-
Taurine
H$^+$ + S^0
Pyruvate
3-Mercaptolactate
Sulfopyruvate
Sulfoacetaldehyde
Coenzyme M
6-Sulfoquinovose
Sulfolipid
Bile acids
(Fig. 22-10)
HO—$CH_2CH_2SO_3^-$
Isethionic acid

Figure 24-25 Pathways of biosynthesis (green arrows) and catabolism of cysteine as well as other aspects of sulfur metabolism. Solid arrows are major biosynthetic pathways. The dashed arrows represent more specialized pathways; they also show processes occurring in the animal body to convert methionine to cysteine and to degrade the latter.



(24-43)

β-substitution reaction analogous to that described in the next paragraph.

1. Synthesis and Catabolism of Cysteine

Cysteine is formed in plants and in bacteria from sulfide and serine after the latter has been acetylated by transfer of an acetyl group from acetyl-CoA (Fig. 24-25, step *f*). This standard PLP-dependent β replacement (Chapter 14) is catalyzed by **cysteine synthase** (*O*-acetylserine sulfhydrase).[446,447] A similar enzyme is used by some cells to introduce sulfide ion directly into homocysteine, via either *O*-succinyl homoserine or *O*-acetyl homoserine (Fig. 24-13). In *E. coli* cysteine can be converted to methionine, as outlined in Eq. 16-22 and as indicated on the right side of Fig. 24-13 by the green arrows. In animals the converse process, the conversion of methionine to cysteine (gray arrows in Fig. 24-13, also Fig. 24-16), is important. Animals are unable to incorporate sulfide directly into cysteine, and this amino acid must be either provided in the diet or formed from dietary methionine. The latter process is limited, and cysteine is an essential dietary constituent for infants. The formation of cysteine from methionine occurs via the same transsulfuration pathway as in methionine synthesis in autotrophic organisms. However, the latter use cystathionine γ-synthase and β-lyase while cysteine synthesis in animals uses cystathionine β-synthase and γ-lyase.

Some bacteria degrade L-cysteine or D-cysteine[447a] via the PLP-dependent α,β elimination to form H_2S, pyruvate, and ammonia (reaction *g*, Fig. 24-25, Eq. 14-29). Another catabolic pathway is transamination (Fig. 24-25, rection *h*) to **3-mercaptopyruvate**.[448] The latter compound can be reductively cleaved to pyruvate and sulfide. Cysteine can also be oxidized by NAD^+ and lactate dehydrogenase to 3-mercaptopyruvate. An interesting PLP-dependent β-replacement reaction of cysteine leads to **β-cyanoalanine**, the lathyritic factor (Box 8-E) present in some plants.[449] This reaction also detoxifies the HCN produced during the biosynthesis of ethylene from ACC.

Cysteine and cystine are relatively insoluble and are toxic in excess.[450] Excretion is usually controlled carefully. However, in **cystinuria**, a disease recognized in the medical literature since 1810,[451] there is a greatly increased excretion of cystine and also of the dibasic amino acids.[451,452] As a consequence, stones of cystine develop in the kidneys and bladder. Patients may excrete more than 1 g of cystine in 24 h compared to a normal of 0.05 g, as well as excessive amounts of lysine, arginine, and ornithine. The defect can be fatal, but some persons with the condition remain healthy indefinitely. Cystinuria is one of several human diseases with altered membrane transport and faulty reabsorption of materials from kidney tubules or from the small intestine. Substances are taken up on one side of a cell (e.g., at the bottom of the cell in Fig. 1-6) and discharged into the bloodstream from the other side of the cell. In another rare hereditary condition, **cystinosis**, free cystine accumulates within lysosomes.[453]

2. Cysteine Sulfinate and Taurine

A quantitatively important pathway of cysteine catabolism in animals is oxidation to **cysteine sulfinate** (Fig. 24-25, reaction *i*),[450] a two-step hydroxylation requiring O_2, NADPH or NADH, and Fe^{2+}. Cysteine sulfinic acid can be further oxidized to **cysteic acid** (cysteine sulfonate),[454] which can be decarboxylated to **taurine**. The latter is a component of bile salts (Fig. 22-16) and is one of the most abundant free amino acids in human tissues.[455–457] Its concentration is high in excitable tissues, and it may be a neurotransmitter (Chapter 30). Taurine may have a special function in retinal photoreceptor cells. It is an essential dietary amino acid for cats, who may die of heart failure in its absence,[458] and under some conditions for humans.[459] In many marine invertebrates, teleosts, and amphibians taurine serves as a regulator of osmotic pressure, its concentration decreasing in fresh water and increasing in salt water. A similar role has been suggested for taurine in mammalian hearts. A chronically low concentration of Na^+ leads to increased taurine.[460] Taurine can be reduced to **isethionic acid**

(Fig. 24-25), another component of nervous tissue. Cysteic acid can arise in an alternative way from *O*-acetylserine and sulfite (reaction l, Fig. 24-25), and taurine can also be formed by decarboxylation of cysteine sulfinic acid to **hypotaurine** and oxidation of the latter (reaction *m*). Cysteic acid can be converted to the sulfolipid of chloroplasts (p. 387; Eq. 20-12).

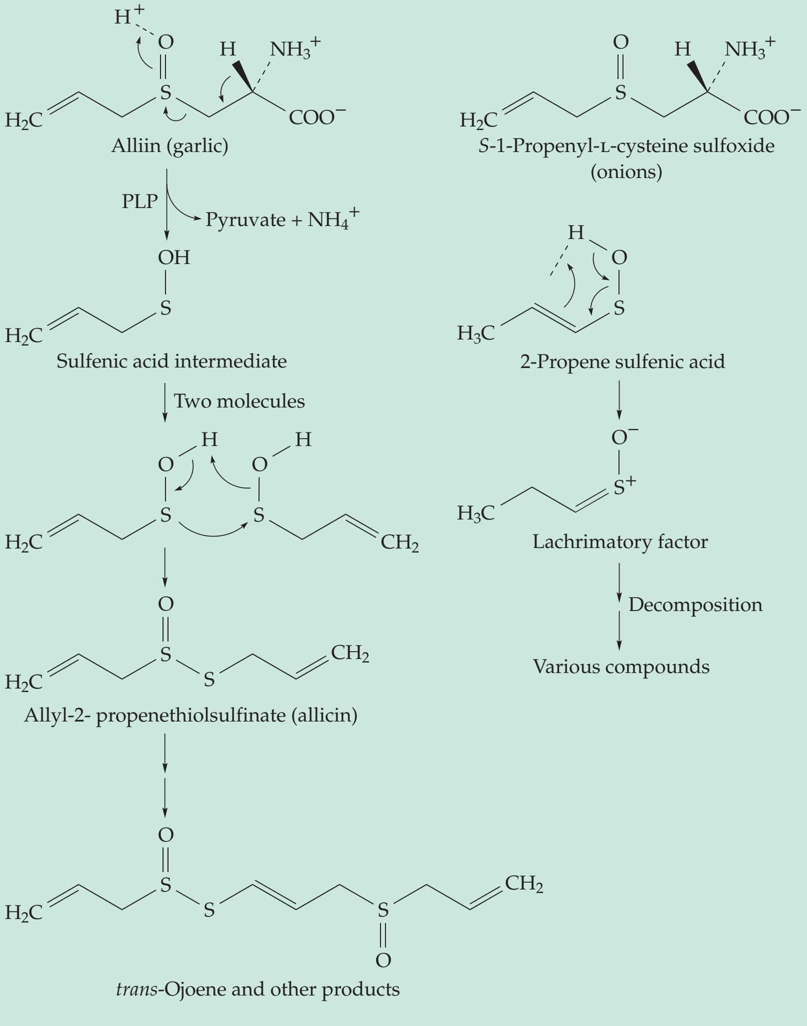
Another route of metabolism for cysteine sulfinic acid is transamination to 3-sulfinylpyruvate, a compound that undergoes ready loss of SO_2 in a reaction analogous to the decarboxylation of oxaloacetate (reaction *o*, Fig. 24-25). This probably represents one of the major routes by which sulfur is removed from organic compounds in the animal body. However, before being excreted the sulfite must be oxidized to sulfate by the Mo-containing sulfite oxidase. The essentiality of sulfite oxidase is evidenced by the severe neurological defect observed in its absence (Chapter 16).

Most of the sulfate generated in the body is excreted unchanged in the urine, but a significant fraction is esterified with oligosaccharides and phenolic compounds. These sulfate esters are formed by sulfo transfer from PAPS (Eq. 17-38).

3. Mercaptopyruvate, Thiosulfate, and Assembly of Iron–Sulfur Centers

An important property of 3-mercaptopyruvate arises from electron withdrawal by the carbonyl group. This makes the SH group electrophilic and able to be transferred as SH^+, S^0, to a variety of nucleophiles (Eq. 24-44). Thus sulfite yields thiosulfate ($S_2O_3^{2-} + H^+$, Eq. 24-45, step *a*), cyanide yields thiocyanate (Eq. 24-45, step *b*), and cysteine sulfinate yields alanine thiosulfonate.[448,461] The reactions are catalyzed by **mercaptopyruvate sulfurtransferase**, an enzyme very similar to **thiosulfate sulfurtransferase**. The latter is a liver enzyme often called by the traditional

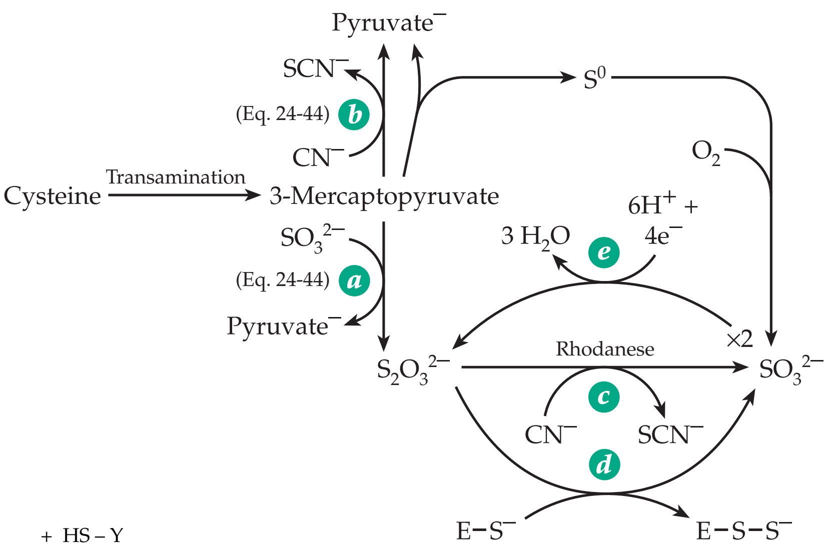
BOX 24-B SULFUR COMPOUNDS OF GARLIC, ONIONS, SKUNKS, ETC.


H+
O
H
NH3+
S
H2C
COO−
Alliin (garlic)
PLP
Pyruvate + NH4+
OH
S
H2C
Sulfenic acid intermediate
Two molecules
H
H
O
O
S
S
H2C
CH2
O
S
S
CH2
H2C
Allyl-2- propenethiolsulfinate (allicin)
O
S
S
S
CH2
H2C
O
trans-Ojoene and other products
O
H
NH3+
S
H2C
COO−
S-1-Propenyl-L-cysteine sulfoxide
(onions)
H
O
S
H3C
2-Propene sulfenic acid
O−
S+
H3C
Lachrimatory factor
Decomposition
Various compounds

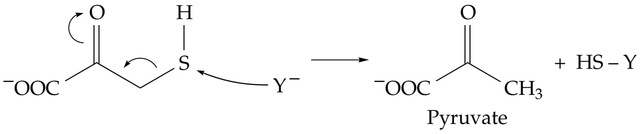
Many familiar odors and tastes come from sulfur-containing compounds. Crushing onions or garlic releases the pyridoxal phosphate-dependent enzyme **alliinase**. In garlic it acts upon the amino acid **alliin** (accompanying scheme) releasing, by β elimination, a sulfenic acid that dimerizes to form **allicin**, a chemically unstable molecule that accounts for the odor of garlic.[a,b] Among the breakdown products of allicin is the nonvolatile **ajoene**, a compound with anticoagulant activity and perhaps accounting for one aspect of the purported medical benefits of garlic. Another is an antibacterial activity.

Onions contain an amino acid that is a positional isomer of alliin. When acted upon by alliinase it produces 2-propene sulfenic acid, which isomerizes to the **lachrimatory factor** that brings tears to the eyes of onion cutters.[c] This, too, decomposes to form many other compounds.[a]

The defensive secretion of the striped skunk has intrigued chemists for over 100 years. The components were shown to contain name **rhodanese**.[462–463b] It acts by a double displacement mechanism, a thiolate anion in the active site of the enzyme serving as a carrier for the S atom being transferred. Equation 24-46 illustrates the transfer of S^0 from thiosulfate to CN^-, converting that ion to the less toxic thiocyanate (also shown as step *c* in Eq. 24-45). Crystallographic studies show that the negative charge on the thiol and dithiol anions of rhodanese is balanced by the partial


Pyruvate⁻
SCN⁻
(Eq. 24-44)
b
CN⁻
S⁰
O₂
Cysteine
Transamination
3-Mercaptopyruvate
6H⁺ +
4e⁻
3 H₂O
e
SO₃²⁻
(Eq. 24-44)
a
Pyruvate⁻
×2
S₂O₃²⁻
Rhodanese
SO₃²⁻
c
CN⁻
SCN⁻
d
E–S⁻
E–S–S⁻

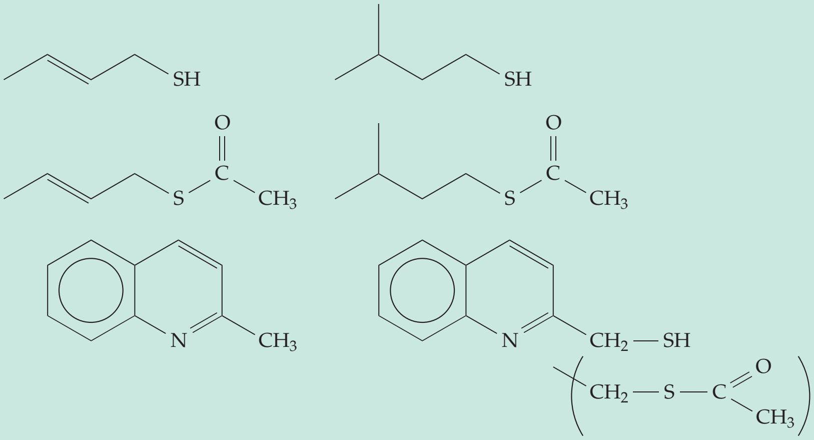
(24-45)


O
H
S
⁻OOC
Y⁻
O
⁻OOC
CH₃
+ HS – Y
Pyruvate

(24-44)

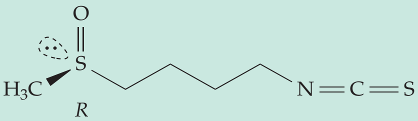
BOX 24-B (continued)

sulfur, and one was incorrectly identified and was long accepted as being butyl mercaptan. Modern capillary gas chromatography by Wood[d] has revealed the presence of seven major components, with the indicated structures. Two are simple volatile mercaptans, but three are thioacetates which hydrolyze in water only slowly, releasing smell for days or weeks from sprayed animals. Washing with mildly basic soap hastens the hydrolysis.


SH
SH
O
C
S
CH₃
O
C
S
CH₃
N
CH₃
N
CH₂ — SH
CH₂ — S — C
O
CH₃

Many readers (~40%) may be aware that after eating asparagus a strong odor appears in their urine. These genetic "stinkers" secrete *S*-methyl thioacrylate, and related compounds, derived from a plant constituent.[e]

The sulfur compound sulfuraphane, extracted from fresh broccoli, has received attention in recent years because of its strong action in inducing synthesis of quinone reductase and glutathione *S*-transferases that help detoxify xenobiotics and may have significant anticancer activity.[f]



Sulforaphane;
1-isothiocyanate-(4*R*)-(methylsulfinyl)butane

---

[a] Block, E. (1985) *Sci. Am.* **252**(Mar), 114–119

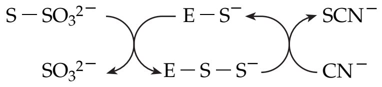
[b] Jain, M. K., and Apitz-Castro, R. (1987) *Trends Biochem. Sci.* **12**, 252–254

[c] Imai, S., Tsuge, N., Tomotake, M., Nagatome, Y., Sawada, H., Nagata, T., and Kumagai, H. (2002) *Nature (London)* **419**, 685

[d] Wood, W. F. (1990) *J. Chem. Ecol.* **16**, 2057

[e] White, R. H. (1975) *Science* **189**, 810

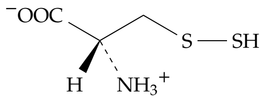
[f] Zhang, Y., Talalay, P., Cho, C.-G., and Posner, G. H. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2399–2403



(24-46)

positive charges at the N termini of two helices (see Fig. 2-20A) and by hydrogen bonds to protons of several side chains.[462] This evidently explains how the negative thiosulfate anion can react with another anion, E–S$^-$. Another interesting feature of this enzyme is that the monomer has a nearly perfect twofold axis of symmetry with respect to the protein folding pattern. However, the symmetry is lacking in the sequence and only one-half of the molecule contains an active site.[462]

Yet another enzyme able to release or transfer sulfur in the S^0 oxidation state is the PLP-dependent **cysteine desulfurase** that is encoded by the *nifS* gene of the nitrogenase gene cluster shown in Fig. 24-4. This enzyme releases S^0 from cysteine with formation of alanine[464–466] as is shown in Eq. 14-34 for release of Se^0 from selenocysteine. As with rhodanese an active site cysteine accepts the departing S^0 of cysteine to form an enzyme-bound persulfide. This protein may in turn transfer the sulfur into the forming Fe–S or Fe–S–Mo clusters.[464] Three PLP-dependent persulfide-forming sulfurtransferases related to the NifS protein have been found in *E. coli*. Similar enzymes are present in other organisms.[466a–d] A sulfur atom may be transferred from the bound persulfide anion to acceptor proteins involved in metal cluster formation. Some members of the nifS-like family act on cystine to release free thiocysteine (cysteine peroxide), which may also serve as a sulfur atom donor.[466e]

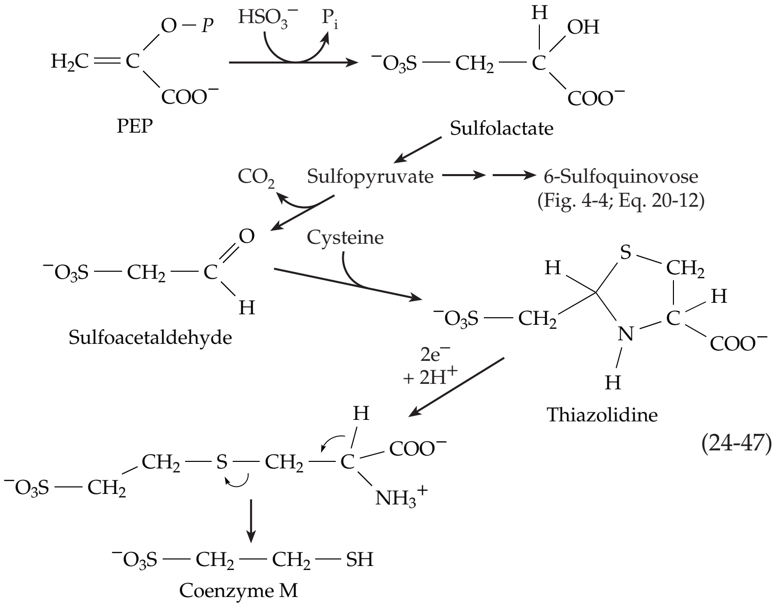


Thiocysteine (cysteine persulfide)

Thiocysteine can also arise in a similar manner through action of cystathionine β lyase on cystine. Thiocysteine is eliminated with production of pyruvate and ammonia from the rest of the cystine molecule.[467] One of the nifS-like proteins of *E. coli* is thought to transfer a selenium atom from selenocysteine (pp. 823–827) into **selenophosphate**.[466a,f] The latter can be formed by transfer of a phospho group from ATP to selenide HSe$^-$. The other products of ATP cleavage are AMP and P$_i$. Reduction of Se^0 to HSe$^-$ is presumably necessary.

Several additional proteins identified as necessary for metal-sulfide cluster formation are present in bacteria and in eukaryotes, both in the cytosol and in mitochondria. They may serve as intermediate sulfur carriers, as scaffolds or templates for cluster formation, or for insertion of intact Fe–S, Fe–S–Mo, or other types of clusters into proteins[468–473f] and into 2-selenouridine[473g] (see also p. 1617). Sulfurtransferases are also thought to be involved in insertion of sulfur atoms into organic molecules such as biotin, lipoic acid, or methanopterin.[474]

A reaction that is ordinarily of minor consequence in the animal body but which may be enhanced by a deficiency of sulfite oxidase is the reductive coupling of two molecules of sulfite to form thiosulfate (Eq. 24-45, step *e*). Several organic hydrodisulfide derivatives such as thiocysteine, thioglutathione, and thiotaurine occur in animals in small amounts. Another biosynthetic pathway, outlined in Eq. 24-47 converts sulfite and PEP into coenzyme M (Fig. 15-22).[475,475a] This cofactor is needed not only for methane formation (Fig. 15-2) but also for utilization of alkenes by soil bacteria.[475b]



(24-47)

References

1. Conn, E. E., Stumpf, P. K., Bruening, G., and Doi, R. H. (1987) *Outlines of Biochemistry*, 5th ed., Wiley, New York

2. Smil, V. (1997) *Sci. Am.* **277**(Jul), 76–81

3. Hardy, R. W. F., and Havelka, U. D. (1975) *Science* **188**, 633–643

4. Winogradsky, S. (1893) *C.R. Acad. Sci.* **116**, 1385–1388

5. Orme-Johnson, W. H. (1985) *Ann. Rev. Biophys. Biophys. Chem.* **14**, 419–459

6. Appleby, C. A., Nicola, N. A., Hurrell, J. G. R., and Leach, S. J. (1975) *Biochemistry* **14**, 4444–4450

7. Lee, H. C., Wittenberg, J. B., and Peisach, J. (1993) *Biochemistry* **32**, 11500–11506

8. Stouggard, J., Petersen, T. E., and Marcker, K. A. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5754–5757

9. Stacey, G., Burris, R. H., and Evans, H. J., eds. (1992) *Biological Nitrogen Fixation*, Chapman and Hall, London

9a. Capone, D. G., Zehr, J. P., Paerl, H. W., Bergman, B., and Carpenter, E. J. (1997) *Science* **276**, 1221–1229

10. Appleby, C. A. (1984) *Ann. Rev. Plant Physiol.* **35**, 443–478

10a. Lou, J., Moshiri, F., Johnson, M. K., Lafferty, M. E., Sorkin, D. L., Miller, A.-F., and Maier, R. J. (1999) *Biochemistry* **38**, 5563–5571

11. Lowe, D. J., Fisher, K., Thorneley, R. N. F., Vaughn, S. A., and Burgess, B. K. (1989) *Biochemistry* **28**, 8460–8466

12. Rasche, M. E., and Seefeldt, L. C. (1997) *Biochemistry* **36**, 8574–8585

13. Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J., and Rees, D. C. (1992) *Science* **257**, 1653–1659

14. Kim, J., and Rees, D. C. (1994) *Biochemistry* **33**, 389–397

15. Howard, J. B., and Rees, D. C. (1994) *Ann. Rev. Biochem.* **63**, 235–264

16. Peters, J. W., Fisher, K., and Dean, D. R. (1995) *Ann. Rev. Microbiol.* **49**, 335–366

17. Kim, J., and Rees, D. C. (1992) *Nature (London)* **360**, 553–560

18. Kim, J., Woo, D., and Rees, D. C. (1993) *Biochemistry* **32**, 7104–7115

19. Schindelin, H., Kisker, C., Schlessman, J. L., Howard, J. B., and Rees, D. C. (1997) *Nature (London)* **387**, 370–376

19a. Mayer, S. M., Lawson, D. M., Gormal, C. A., Roe, S. M., and Smith, B. E. (1999) *J. Mol. Biol.* **292**, 871–891

20. Hoover, T. R., Imperial, J., Ludden, P. W., and Shah, V. K. (1989) *Biochemistry* **28**, 2768–2771

21. Hoover, T. R., Imperial, J., Liang, J., Ludden, P. W., and Shah, V. K. (1988) *Biochemistry* **27**, 3647–3652

22. Imperial, J., Hoover, T. R., Madden, M. S., Ludden, P. W., and Shah, V. K. (1989) *Biochemistry* **28**, 7796–7799

22a. Einsle, O., Tezcan, F. A., Andrade, S. L. A., Schmid, B., Yoshida, M., Howard, J. B., and Rees, D. C. (2002) *Science* **297**, 1696 – 1700

23. Peters, J. W., Stowell, M. H. B., Soltis, S. M., Finnegan, M. G., Johnson, M. K., and Rees, D. C. (1997) *Biochemistry* **36**, 1181–1187

24. McLean, P. A., Wink, D. A., Chapman, S. K., Hickman, A. B., McKillop, D. M., and Orme-Johnson, W. H. (1989) *Biochemistry* **28**, 9402–9406

25. Jensen, B. B., and Burris, R. H. (1985) *Biochemistry* **24**, 1141–1147

26. Liang, J., and Burris, R. H. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9446–9450

27. Kim, C.-H., Newton, W. E., and Dean, D. R. (1995) *Biochemistry* **34**, 2798–2808

28. Cleland, W. W., Guth, J., and Burris, R. H. (1983) *Biochemistry* **22**, 5111–5122

29. Wahl, R. C., and Orme-Johnson, W. H. (1987) *J. Biol. Chem.* **262**, 10489–10496

30. Hoover, D. M., and Ludwig, M. L. (1997) *Protein Sci.* **6**, 2525–2537

31. Yoch, D. C. (1972) *Biochem. Biophys. Res. Commun.* **49**, 335–342

32. Carter, K. R., Rawlings, J., Orme-Johnson, W. H., Becker, R. R., and Evans, H. J. (1980) *J. Biol. Chem.* **255**, 4213–4223

33. Renner, K. A., and Howard, J. B. (1996) *Biochemistry* **35**, 5353–5358

34. Duyvis, M. G., Wassink, H., and Haaker, H. (1996) *J. Biol. Chem.* **271**, 29632–29636

35. Lanzilotta, W. N., Fisher, K., and Seefeldt, L. C. (1997) *J. Biol. Chem.* **272**, 4157–4165

35a. Chan, J. M., Wu, W., Dean, D. R., and Seefeldt, L. C. (2000) *Biochemistry* **39**, 7221–7228

35b. Jang, S. B., Seefeldt, L. C., and Peters, J. W. (2000) *Biochemistry* **39**, 14745–14752

35c. Clarke, T. A., Maritano, S., and Eady, R. R. (2000) *Biochemistry* **39**, 11434–11440

35d. Nyborg, A. C., Johnson, J. L., Gunn, A., and Watt, G. D. (2000) *J. Biol. Chem.* **275**, 39307–39312

36. Lanzilotta, W. N., Fisher, K., and Seefeldt, L. C. (1996) *Biochemistry* **35**, 7188–7196

37. Lanzilotta, W. N., and Seefeldt, L. C. (1996) *Biochemistry* **35**, 16770–16776

37a. Chiu, H.-J., Peters, J. W., Lanzilotta, W. N., Ryle, M. J., Seefeldt, L. C., Howard, J. B., and Rees, D. C. (2001) *Biochemistry* **40**, 641–650

38. Lanzilotta, W. N., and Seefeldt, L. C. (1997) *Biochemistry* **36**, 12976–12983

38a. Jang, S. B., Seefeldt, L. C., and Peters, J. W. (2000) *Biochemistry* **39**, 641–648

39. Grossman, J. G., Hasnain, S. S., Yousafzai, F. K., Smith, B. E., and Eady, R. R. (1997) *J. Mol. Biol.* **266**, 642–648

40. Ma, L., Brosius, M. A., and Burgess, B. K. (1996) *J. Biol. Chem.* **271**, 10528–10532

40a. Chan, J. M., Christiansen, J., Dean, D. R., and Seefeldt, L. C. (1999) *Biochemistry* **38**, 5779–5785

41. Christiansen, J., Tittsworth, R. C., Hales, B. J., and Cramer, S. P. (1995) *J. Am. Chem. Soc.* **117**, 10017–10024

42. Venters, R. A., Nelson, M. J., McLean, P. A., True, A. E., Levy, M. A., Hoffman, B. M., and Orme-Johnson, W. H. (1986) *J. Am. Chem. Soc.* **108**, 3487–3498

43. Shah, V. K., Ugalde, R. A., Imperial, J., and Brill, W. J. (1985) *J. Biol. Chem.* **260**, 3891–3894

44. Stiefel, E. I. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 988–992

44a. Rod, T. H., and Noskov, J. K. (2000) *J. Am. Chem. Soc.* **122**, 12751–12763

44b. Thorneley, R. N., and Lowe, D. J. ((1984) *Biochem. J.* **224**, 887 – 894

44c. Durrant, M. C. (2002) *Biochemistry* **41**, 13934 – 13945; 13946 – 13955

45. Dilworth, M. J., Fisher, K., Kim, C.-H., and Newton, W. E. (1998) *Biochemistry* **37**, 17495–17505

45a. Fisher, K., Dilworth, M. J., Kim, C.-H., and Newton, W. E. (2000) *Biochemistry* **39**, 2970–2979

45b. Fisher, K., Dilworth, M. J., Kim, C.-H., and Newton, W. E. (2000) *Biochemistry* **39**, 10855–10865

45c. Christiansen, J., Seefeldt, L. C., and Dean, D. R. (2000) *J. Biol. Chem.* **275**, 36104–36107

45d. Sorlie, M., Christiansen, J., Lemon, B. J., Peters, J. W., Dean, D. R., and Hales, B. J. (2001) *Biochemistry* **40**, 1540–1549

46. Palermo, R. E., Singh, R., Bashkin, J. K., and Holm, R. H. (1984) *J. Am. Chem. Soc.* **106**, 2600–2612

47. Leigh, G. J. (1995) *Science* **268**, 827–828

48. Ferguson, R., Solari, E., Floriani, C., Osella, D., Ravera, M., Re, N., Chiesi-Villa, A., and Rizzoli, C. (1997) *J. Am. Chem. Soc.* **119**, 10104–10115

49. Nishibayashi, Y., Iwai, S., and Hidai, M. (1998) *J. Am. Chem. Soc.* **120**, 10559–10560

49a. Verma, A. K., and Lee, S. C. (1999) *J. Am. Chem. Soc.* **121**, 10838–10839

50. Laplaza, C. E., Johnson, A. R., and Cummins, C. C. (1996) *J. Am. Chem. Soc.* **118**, 709–710

51. Laplaza, C. E., and Cummins, C. C. (1995) *Science* **268**, 861–863

52. Ow, D. W., and Ausubel, F. M. (1983) *Nature (London)* **301**, 307–313

53. Thiel, T., Lyons, E. M., Erker, J. C., and Ernst, A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9358–9362

54. Pulakat, L., Hausman, B. S., Lei, S., and Gavini, N. (1996) *J. Biol. Chem.* **271**, 1884–1889

54a. Yuvaniyama, P., Agar, J. N., Cash, V. L., Johnson, M. K., and Dean, D. R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 599–604

54b. Ribbe, M. W., Bursey, E. H., and Burgess, B. K. (2000) *J. Biol. Chem.* **275**, 17631–17638

54c. Ribbe, M. W., and Burgess, B. K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5521–5525

55. Shah, V. K., Imperial, J., Ugalde, R. A., Ludden, P. W., and Brill, W. J. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1636–1640

55a. Rangaraj, P., Rüttimann-Johnson, C., Shah, V. K., and Ludden, P. W. (2001) *J. Biol. Chem.* **276**, 15968–15974

56. Jacobson, M. R., Brigle, K. E., Bennett, L. T., Setterquist, R. A., Wilson, M. S., Cash, V. L., Beynon, J., Newton, W. E., and Dean, D. R. (1989) *J. Bacteriol.* **171**, 1017–1027

57. Dilworth, M., and Glenn, A. (1984) *Trends Biochem. Sci.* **9**, 519–523

58. Marx, J. L. (1985) *Science* **230**, 157–158

59. Albrecht, C., Geurts, R., and Bisseling, T. (1999) *EMBO J.* **18**, 281–288

60. Fisher, R. F., and Long, S. R. (1992) *Nature (London)* **357**, 655–660

61. Vijn, I., das Neves, L., van Kammen, A., Franssen, H., and Bisseling, T. (1993) *Science* **260**, 1764–1765

62. Cedergren, R. A., Lee, J., Ross, K. L., and Hollingsworth, R. I. (1995) *Biochemistry* **34**, 4467–4477

63. Dunn, S. M., Moody, P. C. E., Downie, J. A., and Shaw, W. V. (1996) *Protein Sci.* **5**, 538–541

64. Spaink, H. P., Sheeley, D. M., van Brussel, A. A. N., Glushka, J., York, W. S., Tak, T., Geiger, O., Kennedy, E. P., Reinhold, V. N., and Lugtenberg, B. J. J. (1991) *Nature (London)* **354**, 125–130

65. Mergaert, P., D'Haeze, W., Geelen, D., Promé, D., Van Montagu, M., Geremia, R., Promé, J.-C., and Holsters, M. (1995) *J. Biol. Chem.* **270**, 29217–29223

66. Schultze, M., Staehelin, C., Röhrig, H., John, M., Schmidt, J., Kondorosi, E., Schell, J., and Kondorosi, A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2706–2709

67. Fisher, R. F., and Long, S. R. (1993) *J. Mol. Biol.* **233**, 336–348

68. Monson, E. K., Ditta, G. S., and Helinski, D. R. (1995) *J. Biol. Chem.* **270**, 5243–5250

69. Soupène, E., Foussard, M., Boistard, P., Truchet, G., and Batut, J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3759–3763

69a. Miyatake, H., Mukai, M., Park, S.-Y., Adachi, S., Tamura, K., Nakamura, H., Nakamura, K., Tsuchiya, T., Iizuka, T., and Shiro, Y. (2000) *J. Mol. Biol.* **301**, 415–431

70. Freiberg, C., Fellay, R., Bairoch, A., Broughton, W. J., Rosenthal, A., and Perre, X. (1997) *Nature (London)* **387**, 394–401

References

71. Downie, A. (1997) *Nature (London)* **387**, 352–354
72. Minchin, F. (1986) *Nature (London)* **320**, 483–484
73. Berry, A. M., Harriott, O. T., Moreau, R. A., Osman, S. F., Benson, D. R., and Jones, A. D. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6091–6094
74. Bishop, P. E. (1986) *Trends Biochem. Sci.* **11**, 225–227
75. Blanchard, C. Z., and Hales, B. J. (1996) *Biochemistry* **35**, 472–478
76. Chatterjee, R., Ludden, P. W., and Shah, V. K. (1997) *J. Biol. Chem.* **272**, 3758–3765
76a. Rüttimann-Johnson, C., Rangaraj, P., Shah, V. K., and Ludden, P. W. (2001) *J. Biol. Chem.* **276**, 4522–4526
77. Pau, R. N. (1989) *Trends Biochem. Sci.* **14**, 183–186
78. Ribbe, M., Gadkari, D., and Meyer, O. (1997) *J. Biol. Chem.* **272**, 26627–26633
79. DiSpirito, A. A., and Hooper, A. B. (1986) *J. Biol. Chem.* **261**, 10534–10537
79a. Kanamaru, K., Wang, R., Su, W., and Crawford, N. M. (1999) *J. Biol. Chem.* **274**, 4160 – 4165
80. Campbell, W. H., and Kinghorn, J. R. (1990) *Trends Biochem. Sci.* **15**, 315–319
81. Lu, G., Lindqvist, Y., Schneider, G., Dwivedi, U., and Campbell, W. (1995) *J. Mol. Biol.* **248**, 931–948
82. Garde, J., Kinghorn, J. R., and Tomsett, A. B. (1995) *J. Biol. Chem.* **270**, 6644–6650
83. Hyde, G. E., Crawford, N. M., and Campbell, W. H. (1991) *J. Biol. Chem.* **266**, 23542–23547
84. Ratnam, K., Shiraishi, N., Campbell, W. H., and Hille, R. (1995) *J. Biol. Chem.* **270**, 24067–24072
85. Cramer, S. P., Solomonson, L. P., Adams, M. W. W., and Mortenson, L. E. (1984) *J. Am. Chem. Soc.* **106**, 1467–1471
85a. Skipper, L., Campbell, W. H., Mertens, J. A., and Lowe, D. J. (2001) *J. Biol. Chem.* **276**, 26995–27002
86. Lin, J. T., and Stewart, V. (1996) *J. Mol. Biol.* **256**, 423–435
86a. Butler, C. S., Charnock, J. M., Bennett, B., Sears, H. J., Reilly, A. J., Ferguson, S. J., Garner, C. D., Lowe, D. J., Thomson, A. J., Berks, B. C., and Richardson, D. J. (1999) *Biochemistry* **38**, 9000–9012
87. Augier, V., Guigliarelli, B., Asso, M., Bertrand, P., Frixon, C., Giordano, G., Chippaux, M., and Blasco, F. (1993) *Biochemistry* **32**, 2013–2023
87a. Zhao, Y., Lukoyanov, D. A., Toropov, Y. V., Wu, K., Shapleigh, J. P., and Scholes, C. P. (2002) *Biochemistry* **41**, 7464 – 7474
88. Bentley, R. (1985) *Trends Biochem. Sci.* **10**, 171–174
89. Cooper, A. J. L., Nieves, E., Coleman, A. E., Filc-DeRicco, S., and Gelbard, A. S. (1987) *J. Biol. Chem.* **262**, 1073–1080
90. Cooper, A. J. L., Nieves, E., Rosenspire, K. C., Filc-DeRicco, S., Gelbard, A. S., and Brusilow, S. W. (1988) *J. Biol. Chem.* **263**, 12268–12273
91. Chatham, J. C., Forder, J. R., Glickson, J. D., and Chance, E. M. (1995) *J. Biol. Chem.* **270**, 7999–8008
92. Martin, G., Chauvin, M.-F., and Baverel, G. (1997) *J. Biol. Chem.* **272**, 4717–4728
93. Chauvin, M.-F., Mégnin-Chanet, F., Martin, G., Lhoste, J.-M., and Baverel, G. (1994) *J. Biol. Chem.* **269**, 26025–26033
94. Ikeda, T. P., Shauger, A. E., and Kustu, S. (1996) *J. Mol. Biol.* **259**, 589–607
95. Zhalkin, H. (1993) *Adv. Enzymol.* **66**, 203–309
96. Boehlein, S. K., Richards, N. G. J., Walworth, E. S., and Schuster, S. M. (1994) *J. Biol. Chem.* **269**, 26789–26795
97. Muchmore, C. R. A., Krahn, J. M., Kim, J. H., Zalkin, H., and Smith, J. L. (1998) *Protein Sci.* **7**, 39–51
97a. Raushel, F. M., Thoden, J. B., and Holden, H. M. (1999) *Biochemistry* **38**, 7891–7899
97b. Bera, A. K., Smith, J. L., and Zalkin, H. (2000) *J. Biol. Chem.* **275**, 7975–7979
97c. Teplyakov, A., Obmolova, G., Badet, B., and Badet-Denisot, M.-A. (2001) *J. Mol. Biol.* **313**, 1093–1102
97d. Chittur, S. V., Klem, T. J., Shafer, C. M., and Davisson, V. J. (2001) *Biochemistry* **40**, 876–887
98. Boehlein, S. K., Stewart, J. D., Walworth, E. S., Thirumoorthy, R., Richards, N. G. J., and Schuster, S. M. (1998) *Biochemistry* **37**, 13230–13238
98a. Larsen, T. M., Boehlein, S. K., Schuster, S. M., Richards, N. G. J., Thoden, J. B., Holden, H. M., and Rayment, I. (1999) *Biochemistry* **38**, 16146–16157
98b. Min, B., Pelaschier, J. T., Graham, D. E., Tumbula-Hansen, D., and Söll, D. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2678–2683
99. Raghunathan, G., Miles, H. T., and Sasisekharan, V. (1993) *Biochemistry* **32**, 455–462
100. Winter, H. C., and Dekker, E. E. (1986) *J. Biol. Chem.* **261**, 11189–11193
101. Kohl, D. H., Schubert, K. R., Carter, M. B., Hagedorn, C. H., and Shearer, G. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2036–2040
102. Cooper, A. J. L., and Meister, A. (1985) in *Transaminases* (Christen, P., and Metzler, D. E., eds), pp. 533–563, Wiley, New York
103. Sibson, N. R., Dhankhar, A., Mason, G. F., Behar, K. L., Rothman, D. L., and Shulman, R. G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2699–2704
104. Liaw, S.-H., Kuo, I., and Eisenberg, D. (1995) *Protein Sci.* **4**, 2358–2365
105. Kagan, Z. A., Kretovich, V. L., and Polyakov, V. A. (1966) *Biokhim.* **31**, 355–364
106. Baker, P. J., Waugh, M. L., Wang, X.-G., Stillman, T. J., Turnbull, A. P., Engel, P. C., and Rice, D. W. (1997) *Biochemistry* **36**, 16109–16115
107. Peters, J.-M., Harris, J. R., and Finley, D., eds. (1998) *Ubiquitin and the Biology of the Cell*, Plenum, New York
107a. Katagiri, M., and Nakamura, M. (1999) *Biochem. Edu.* **27**, 83–85
107b. Katagiri, M., and Nakamura, M. (2002) *IUBMB Life* **53**, 125 – 129
107c. Wipf, D., Ludewig, U., Tegeder, M., Rentsch, D., Koch, W., and Frommer, W. B. (2002) *Trends Biochem. Sci.* **27**, 139–147
108. Rajagopalan, S., Wan, D.-F., Habib, G. M., Sepulveda, A. R., McLeod, M. R., Lebovitz, R. M., and Lieberman, M. W. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6179–6183
109. Harding, C. O., Williams, P., Wagner, E., Chang, D. S., Wild, K., Colwell, R. E., and Wolff, J. A. (1997) *J. Biol. Chem.* **272**, 12560–12567
109a. Stein, R. L., DeCicco, C., Nelson, D., and Thomas, B. (2001) *Biochemistry* **40**, 5804–5811
110. Guo-jie, Breslow, E., and Meister, A. (1996) *J. Biol. Chem.* **271**, 32293–32300
111. Van Der Werf, P., Griffith, O. W., and Meister, A. (1975) *J. Biol. Chem.* **250**, 6686–6692
112. Reitzer, L. J., and Magasanik, B. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 302–318, Am. Soc. for Microbiology, Washington,DC
113. Kanamori, K., Weiss, R. L., and Roberts, J. D. (1988) *J. Biol. Chem.* **263**, 2817–2823
114. Knapp, S., de Vos, W. M., Rice, D., and Ladenstein, R. (1997) *J. Mol. Biol.* **267**, 916–932
115. Wang, X.-G., and Engel, P. C. (1995) *Biochemistry* **34**, 11417–11422
115a. Herrero-Yraola, A., Bakhit, S. M. A., Franke, P., Weise, C., Schweiger, M., Jorcke, D., and Ziegler, M. (2001) *EMBO J.* **20**, 2404–2412
116. Melo-Oliveira, R., Oliveira, I. C., and Coruzzi, G. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4718–4723
117. Vanoni, M. A., Edmondson, D. E., Zanetti, G., and Curti, B. (1992) *Biochemistry* **31**, 4613–4623
118. Vanoni, M. A., Fischer, F., Ravasio, S., Verzotti, E., Edmondson, D. E., Hagen, W. R., Zanetti, G., and Curti, B. (1998) *Biochemistry* **37**, 1828–1838
118a. Morandi, P., Valzasina, B., Colombo, C., Curti, B., and Vanoni, M. A. (2000) *Biochemistry* **39**, 727–735
119. Sakakibara, H., Watanabe, M., Hase, T., and Sugiyama, T. (1991) *J. Biol. Chem.* **266**, 2028–2035
120. Knaff, D. B., Hirasawa, M., Ameyibor, E., Fu, W., and Johnson, M. K. (1991) *J. Biol. Chem.* **266**, 15080–15084
121. Ho, H.-T., Falk, P. J., Ervin, K. M., Krishnan, B. S., Discotto, L. F., Dougherty, T. J., and Pucci, M. J. (1995) *Biochemistry* **34**, 2464–2470
122. Gallo, K. A., and Knowles, J. R. (1993) *Biochemistry* **32**, 3981–3990
122a. Glavas, S., and Tanner, M. E. (2001) *Biochemistry* **40**, 6199–6204
123. D'Aniello, A., D'Onofrio, G., Pischetola, M., D'Aniello, G., Vetere, A., Petrucelli, L., and Fisher, G. H. (1993) *J. Biol. Chem.* **268**, 26941–26949
124. Almassy, R. J., Janson, C. A., Hamlin, R., Xuong, N.-H., and Eisenberg, D. (1986) *Nature (London)* **323**, 304–309
125. Liaw, S.-H., and Eisenberg, D. (1994) *Biochemistry* **33**, 675–681
126. Mura, U., Chock, P. B., and Stadtman, E. R. (1981) *J. Biol. Chem.* **256**, 13022–13029
126a. Stadtman, E. R. (2001) *J. Biol. Chem.* **276**, 44357–44364
127. Garcia, E., and Rhee, S. G. (1983) *J. Biol. Chem.* **258**, 2246–2253
128. Son, H. S., and Rhee, S. G. (1987) *J. Biol. Chem.* **262**, 8690–8695
129. Jiang, P., Peliska, J. A., and Ninfa, A. J. (1998) *Biochemistry* **37**, 12802–12810
130. Liaw, S.-H., Pan, C., and Eisenberg, D. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4996–5000
131. Jiang, P., Peliska, J. A., and Ninfa, A. J. (1998) *Biochemistry* **37**, 12795–12801
132. Flashner, Y., Weiss, D. S., Keener, J., and Kustu, S. (1995) *J. Mol. Biol.* **249**, 700–713
133. Rosenfeld, S. A., and Brenchley, J. E. (1983) in *Amino Acids, Biosynthesis and Genetic Regulation* (Herrmann, K. M., and Somerville, R. L., eds), pp. 1–17, Addison-Wesley, Reading, Massachusetts
134. Buchanan, J. M. (1973) *Adv. Enzymol.* **38**, 91–183
134a. Gill, H. S., and Eisenberg, D. (2001) *Biochemistry* **40**, 1903–1912
135. Neu, J., Shenoy, V., and Chakrabarti, R. (1996) *FASEB J.* **10**, 829–837
136. Jakob, C. G., Lewinski, K., LaCount, M. W., Roberts, J., and Lebioda, L. (1997) *Biochemistry* **36**, 923–931
137. Watford, M. (1993) *FASEB J.* **7**, 1468–1474
138. Snedden, W. A., Koutsia, N., Baum, G., and Fromm, H. (1996) *J. Biol. Chem.* **271**, 4148–4153
138a. Locher, K. P., Hans, M., Yeh, A. P., Schmid, B., Buckel, W., and Rees, D. C. (2001) *J. Mol. Biol.* **307**, 297–308
138b. Chih, H.-W., and Marsh, E. N. G. (2001) *Biochemistry* **40**, 13060–13067
139. Smith, M. A., King, P. J., and Grimm, B. (1998) *Biochemistry* **37**, 319–329
140. Adams, E., and Frank, L. (1980) *Ann. Rev. Biochem.* **49**, 1005–1061

References

141. Leisinger, T. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 345–351, Am. Soc. for Microbiology, Washington,DC
142. Zhang, C.-s, Lu, Q., and Verma, D. P. S. (1995) *J. Biol. Chem.* **270**, 20491–20496
143. García-Ríos, M., Fujita, T., LaRosa, P. C., Locy, R. D., Clithero, J. M., Bressan, R. A., and Csonka, L. N. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8249–8254
144. Mestichelli, L. J. J., Gupta, R. N., and Spenser, I. D. (1979) *J. Biol. Chem.* **254**, 640–647
145. Ashford, D., and Neuberger, A. (1980) *Trends Biochem. Sci.* **5**, 245–248
145a. Becker, D. F., and Thomas, E. A. (2001) *Biochemistry* **40**, 4714–4721
145b. Farrant, R. D., Walker, V., Mills, G. A., Mellor, J. M., and Langley, G. J. (2001) *J. Biol. Chem.* **276**, 15107–15116
146. Small, W. C., and Jones, M. E. (1990) *J. Biol. Chem.* **265**, 18668–18672
147. Hu, C.-aA., Lin, W.-W., and Valle, D. (1996) *J. Biol. Chem.* **271**, 9795–9800
148. Phang, J. M., Yeh, G. C., and Scriver, C. R. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1125–1146, McGraw-Hill, New York
148a. Farrant, R. D., Walker, V., Mills, G. A., Mellor, J. M., and Langley, G. J. (2000) *J. Biol. Chem.* **276**, 15107 – 15116
149. Gupta, S. C., and Dekker, E. E. (1980) *J. Biol. Chem.* **255**, 1107–1112
150. Sletten, K., Aakesson, I., and Alvsaker, J. O. (1971) *Nature New Biol.* **231**, 118–119
151. Akers, H. A., and Dromgoole, E. V. (1982) *Trends Biochem. Sci.* **7**, 156–157
152. Umbarger, H. E. (1978) *Ann. Rev. Biochem.* **47**, 533–606
153. Davis, R. H., and Weiss, R. L. (1988) *Trends Biochem. Sci.* **13**, 101–104
154. Shah, S. A., Shen, B. W., and Brünger, A. T. (1997) *Structure* **5**, 1067–1075
155. Valle, D., and Simell, O. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1147–1185, McGraw-Hill, New York
156. Glansdorff, N. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 321–339, Am. Soc. for Microbiology, Washington, DC
157. Krebs, H. A., and Henseleit, K. (1932) *Hoppe-Seyler's Z. Physiol. Chem.* **210**, 33–66
158. Krebs, H. A. (1982) *Trends Biochem. Sci.* **7**, 76–78
159. Wakabayashi, Y., Yamada, E., Yoshida, T., and Takahashi, H. (1994) *J. Biol. Chem.* **269**, 32667–32671
160. Gessert, S. F., Kim, J. H., Nargang, F. E., and Weiss, R. L. (1994) *J. Biol. Chem.* **269**, 8189–8203
161. Watford, M. (1989) *Trends Biochem. Sci.* **14**, 313–314
162. Brusilow, S. W., and Horwich, A. L. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1187–1232, McGraw-Hill, New York
163. Lim, A. L., and Powers-Lee, S. G. (1996) *J. Biol. Chem.* **271**, 11400–11409
164. Guy, H. I., and Evans, D. R. (1996) *J. Biol. Chem.* **271**, 13762–13769
165. McCudden, C. R., and Powers-Lee, S. G. (1996) *J. Biol. Chem.* **271**, 18285–18294
166. Hong, J., Salo, W. L., Lusty, C. J., and Anderson, P. M. (1994) *J. Mol. Biol.* **243**, 131–140
167. Devaney, M. A., and Powers-Lee, S. G. (1984) *J. Biol. Chem.* **259**, 703–706
168. Anderson, P. M. (1981) *J. Biol. Chem.* **256**, 12228–12238
169. Stapleton, M. A., Javid-Majd, F., Harmon, M. F., Hanks, B. A., Grahmann, J. L., Mullins, L. S., and Raushel, F. M. (1996) *Biochemistry* **35**, 14352–14361
170. Miles, B. W., Banzon, J. A., and Raushel, F. M. (1998) *Biochemistry* **37**, 16773–16779
171. Powers, S. G., and Meister, A. (1978) *J. Biol. Chem.* **253**, 1258–1265
172. Thoden, J. B., Holden, H. M., Wesenberg, G., Raushel, F. M., and Rayment, I. (1997) *Biochemistry* **36**, 6305–6316
173. Thoden, J. B., Miran, S. G., Phillips, J. C., Howard, A. J., Rauschel, F. M., and Holden, H. M. (1998) *Biochemistry* **37**, 8825–8831
173a. Thoden, J. B., Huang, X., Raushel, F. M., and Holden, H. M. (1999) *Biochemistry* **38**, 16158–16166
173b. Kim, J., Howell, S., Huang, X., and Raushel, F. M. (2002) *Biochemistry* **41**, 12575 – 12581
173c. Fresquet, V., Mora, P., Rochera, L., Ramón-Maiques, S., Rubio, V., and Cervera, J. (2000) *J. Mol. Biol.* **299**, 979–991
174. Shi, D., Morizono, H., Ha, Y., Aoyagi, M., Tuchman, M., and Allewell, N. M. (1998) *J. Biol. Chem.* **273**, 34247–34254
175. Tricot, C., Villeret, V., Sainz, G., Dideberg, O., and Stalon, V. (1998) *J. Mol. Biol.* **283**, 695–704
175a. Langley, D. B., Templeton, M. D., Fields, B. A., Mitchell, R. E., and Collyer, C. A. (2000) *J. Biol. Chem.* **275**, 20012–20019
176. Rosenberg, L. E., Kalousek, F., and Orsulak, M. D. (1983) *Science* **222**, 426–428
177. Maestri, N. E., Brusilow, S. W., Clissold, D. B., and Bassett, S. S. (1996) *N. Engl. J. Med.* **335**, 855–859
177a. Sampaleanu, L. M., Vallée, F., Thompson, G. D., and Howell, P. L. (2001) *Biochemistry* **40**, 15570–15580
178. Turner, M. A., Simpson, A., McInnes, R. R., and Howell, P. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9063–9068
178a. Yu, B., Thompson, G. D., Yip, P., Howell, P. L., and Davidson, A. R. (2001) *Biochemistry* **40**, 15581–15590
179. Garrard, L. J., Mathis, J. M., and Raushel, F. M. (1983) *Biochemistry* **22**, 3729–3735
180. Kanyo, Z. F., Scolnick, L. R., Ash, D. E., and Christianson, D. W. (1996) *Nature (London)* **383**, 554–557
181. Marathe, S., Yu, Y. G., Turner, G. E., Palmier, C., and Weiss, R. L. (1998) *J. Biol. Chem.* **273**, 29776–29785
182. Scolnick, L. R., Kanyo, Z. F., Cavalli, R. C., Ash, D. E., and Christianson, D. W. (1997) *Biochemistry* **36**, 10558–10565
183. Smith, I. (1981) *Nature (London)* **291**, 378–380
184. Atkinson, D. E., and Bourke, E. (1984) *Trends Biochem. Sci.* **9**, 297–300
185. Gargan, R., Rapoport, G., and Débarbouillé, M. (1995) *J. Mol. Biol.* **249**, 843–856
186. Jann, A., Stalon, V., Vander Wauven, C., Leisinger, T., and Haas, D. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4937–4941
187. Weickmann, J. L., Himmel, M. E., Squire, P. G., and Fahrney, D. E. (1978) *J. Biol. Chem.* **253**, 6010–6015
188. Abdelal, A. T. (1979) *Ann. Rev. Microbiol.* **33**, 139–168
189. Vanderbilt, A. S., Gaby, N. S., and Rodwell, V. W. (1975) *J. Biol. Chem.* **250**, 5322–5329
190. Gary, J. D., Lin, W.-J., Yang, M. C., Herschman, H. R., and Clarke, S. (1996) *J. Biol. Chem.* **271**, 12585–12594
191. Bogumil, R., Knipp, M., Fundel, S. M., and Vasák, M. (1998) *Biochemistry* **37**, 4791–4798
192. Rosenthal, G. A. (1984) *Sci. Am.* **250**(Jun), 164–171
193. Rosenthal, G. A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1780–1784
194. Melangeli, C., Rosenthal, G. A., and Dalman, D. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2255–2260
195. Fritsche, E., Humm, A., and Huber, R. (1999) *J. Biol. Chem.* **274**, 3026–3032
196. Takata, Y., Konishi, K., Gomi, T., and Fujioka, M. (1994) *J. Biol. Chem.* **269**, 5537–5542
197. Humm, A., Fritsche, E., Steinbacher, S., and Huber, R. (1997) *EMBO J.* **16**, 3373–3385
198. Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R., and Chapman, M. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8449–8454
199. Turner, D. M., and Walker, J. B. (1987) *J. Biol. Chem.* **262**, 6605–6609
200. Cohen, S. S. (1971) *Introduction to the Polyamines*, Prentice-Hall, Englewood Cliffs, New Jersey
201. Tabor, C. W., and Tabor, H. (1976) *Ann. Rev. Biochem.* **45**, 285–306
202. Cohen, S. S (1978) *Nature (London)* **274**, 209–210
203. Abraham, A. K., and Pihl, A. (1981) *Trends Biochem. Sci.* **6**, 106–107
204. Igarashi, K., Saisho, T., Yuguchi, M., and Kashiwagi, K. (1997) *J. Biol. Chem.* **272**, 4058–4064
205. Cohen, S. S., and Greenberg, M. L. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5470–5474
206. Tsuboi, M. (1964) *Bull. Chem. Soc. Jap.* **37**, 1514–1522
207. Balkema, G. W., Mangini, N. J., and Pinto, L. H. (1983) *Science* **219**, 1083–1087
207a. Chattopadhyay, M. K., Tabor, C. W., Tabor, H. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10330 — 10334
208. Herbst, E. J. et al. (1955) *J. Biol. Chem.* **214**, 175–
209. Leroy, D., Filhol, O., Delcros, J. G., Pares, S., Chambaz, E. M., and Cochet, C. (1997) *Biochemistry* **36**, 1242–1250
210. Hafner, E. W., Tabor, C. W., and Tabor, H. (1979) *J. Biol. Chem.* **254**, 12419–12426
211. Balasundaram, D., Tabor, C. W., and Tabor, H. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5872–5876
211a. Chattopadhyay, M. K., Tabor, C. W., and Tabor, H. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10330 – 10334
212. Cohn, M. S., Tabor, C. W., Tabor, H., and Wickner, R. B. (1978) *J. Biol. Chem.* **253**, 5225–5227
213. Iyer, R., and Delcour, A. H. (1997) *J. Biol. Chem.* **272**, 18595–18601
214. Kaouass, M., Gamache, I., Ramotar, D., Audette, M., and Poulin, R. (1998) *J. Biol. Chem.* **273**, 2109–2117
215. Antognoni, F., Del Duca, S., Kuraishi, A., Kawabe, E., Fukuchi-Shimogori, T., Kashiwagi, K., and Igarashi, K. (1999) *J. Biol. Chem.* **274**, 1942–1948
216. Momany, C., Ernst, S., Ghosh, R., Chang, N.-L., and Hackert, M. L. (1995) *J. Mol. Biol.* **252**, 643–655
217. Srivenugopal, K. S., and Adiga, P. R. (1981) *J. Biol. Chem.* **256**, 9532–9541
218. Bowman, W. H., Tabor, C. W., and Tabor, H. (1973) *J. Biol. Chem.* **248**, 2480–2486
219. Pegg, A. E., Wechter, R. S., Clark, R. S., Wiest, L., and Erwin, B. G. (1986) *Biochemistry* **25**, 379–384
219a. Tolbert, W. D., Ekstrom, J. L., Mathews, I. I., Secrist, J. A., III, Kapoor, P., Pegg, A. E., and Ealick, S. E. (2001) *Biochemistry* **40**, 9484–9494
220. Heby, O., and Persson, L. (1990) *Trends Biochem. Sci.* **15**, 153–158
221. Orr, G. R., Danz, D. W., Pontoni, G., Probhakaran, P. C., Gould, S. J., and Coward, J. K. (1988) *J. Am. Chem. Soc.* **110**, 5791–5799
222. Oshima, T., Hamasaki, N., Senshu, M., Kakinuma, K., and Kuwajima, I. (1987) *J. Biol. Chem.* **262**, 11979–11981

References

223. Fujihara, S., Abe, H., and Yoneyama, T. (1995) *J. Biol. Chem.* **270**, 9932–9938
224. Cacciapuoti, G., Porcelli, M., Bertoldo, C., De Rosa, M., and Zappia, V. (1994) *J. Biol. Chem.* **269**, 24762–24769
224a. Kröger, N., Deutzmann, R., and Sumper, M. (1999) *Science* **286**, 1129–1132
224b. Kröger, N., Deutzmann, R., Bergsdorf, C., and Sumper, M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14133–14138
224c. Ober, D., and Hartmann, T. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14777–14782
224d. Kröger, N., Lorenz, S., Brunner, B., and Sumper, M. (2002) *Science* **298**, 584 – 585
225. Jasys, V. J., Kelbaugh, P. R., Nason, D. M., Phillips, D., Rosnack, K. J., Saccomano, N. A., Stroh, J. G., and Volkmann, R. A. (1990) *J. Am. Chem. Soc.* **112**, 6696–6704
226. McConlogue, L., Gupta, M., Wu, L., and Coffino, P. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 540–544
227. Rom, E., and Kahana, C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3959–3963
228. Li, R.-S., Abrahamsen, M. S., Johnson, R. R., and Morris, D. R. (1994) *J. Biol. Chem.* **269**, 7941–7949
229. Pena, A., Reddy, C. D., Wu, S., Hickok, N. J., Reddy, E. P., Yumet, G., Soprano, D. R., and Soprano, K. J. (1993) *J. Biol. Chem.* **268**, 27277–27285
230. Hayashi, S.-i, Murakami, Y., and Matsufuji, S. (1996) *Trends Biochem. Sci.* **21**, 27–30
230a. Zhu, C., Lang, D. W., and Coffino, P. (1999) *J. Biol. Chem.* **274**, 26425–26430
230b. Coffino, P. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4421–4423
231. Atmar, V. J., and Kuehn, G. D. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5518–5522
232. Coleman, C. S., Huang, H., and Pegg, A. E. (1995) *Biochemistry* **34**, 13423–13430
233. Coleman, C. S., and Pegg, A. E. (1997) *J. Biol. Chem.* **272**, 12164–12169
233a. Bina, C., Angelini, R., Federico, R., Ascenzi, P., and Mattevi, A. (2001) *Biochemistry* **40**, 2766–2776
234. Richards, J. C., and Spenser, I. D. (1978) *J. Am. Chem. Soc.* **100**, 7402–7404
235. Lundgren, D. W., and Fales, H. M. (1980) *J. Biol. Chem.* **255**, 4481–4486
236. Lee, Y., and Sayre, L. M. (1998) *J. Biol. Chem.* **273**, 19490–19494
236a. White, W. H., Gunyuzlu, P. L., Toyn, J. H. (2001) *J. Biol. Chem.* **276**, 10794 – 10800
237. Tabor, H., and Tabor, C. W. (1975) *J. Biol. Chem.* **250**, 2648–2654
238. Smith, K., Nadeau, K., Bradley, M., Walsh, C., and Fairlamb, A. H. (1992) *Protein Sci.* **1**, 874–883
239. Bollinger, J. M., Jr., Kwon, D. S., Huisman, G. W., Kolter, R., and Walsh, C. T. (1995) *J. Biol. Chem.* **270**, 14031–14041
240. Folk, J. E., Park, M. H., Chung, S. I., Schrode, J., Lester, E. P., and Cooper, H. L. (1980) *J. Biol. Chem.* **255**, 3695–3700
241. Leete, E., and McDonnell, J. A. (1981) *J. Am. Chem. Soc.* **103**, 658–662
242. Park, M. H., Cooper, H. L., and Folk, J. E. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2869–2873
243. Osterman, A. L., Brooks, H. B., Rizo, J., and Phillips, M. A. (1997) *Biochemistry* **36**, 4558–4567
244. Coleman, C. S., Stanley, B. A., and Pegg, A. E. (1993) *J. Biol. Chem.* **268**, 24572–24579
244a. Marali, S. (1999) *J. Biol. Chem.* **274**, 21017–21022
245. Pegg, A. E., Jones, D. B., and Secrist, J. A., III. (1988) *Biochemistry* **27**, 1408–1415
246. Hyvönen, T., Alakuijala, L., Andersson, L., Khomutov, A. R., Khomutov, R. M., and Eloranta, T. O. (1988) *J. Biol. Chem.* **263**, 11138–11144
247. Cohen, G. N., and Saint-Girons, I. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 429–442, Am. Soc. for Microbiology, Washington, DC
248. Shames, S. L., Ash, D. E., Wedler, F. C., and Villafranca, J. J. (1984) *J. Biol. Chem.* **259**, 15331–15339
249. Angeles, T. S., Hunsley, J. R., and Viola, R. E. (1992) *Biochemistry* **31**, 799–805
249a. Hadfield, A., Kryger, G., Ouyang, J., Petsko, G. A., Ringe, D., and Viola, R. (1999) *J. Mol. Biol.* **289**, 991–1002
250. Ouyang, J., and Viola, R. E. (1995) *Biochemistry* **34**, 6394–6399
251. Ravanel, S., Gakière, B., Job, D., and Douce, R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7805–7812
252. Giovanelli, J., Mudd, S. H., and Datko, A. H. (1978) *J. Biol. Chem.* **253**, 5665–5677
253. Huo, X., and Viola, R. E. (1996) *Biochemistry* **35**, 16180–16185
254. Laber, B., Gerbling, K.-P., Harde, C., Neff, K.-H., Nordhoff, E., and Pohlenz, H.-D. (1994) *Biochemistry* **33**, 3413–3423
255. Curien, G., Job, D., Douce, R., and Dumas, R. (1998) *Biochemistry* **37**, 13212–13221
255a. Thomazeau, K., Curien, G., Dumas, R., and Biou, V. (2001) *Protein Sci.* **10**, 638–648
256. Chinchilla, D., Schwarz, F. P., and Eisenstein, E. (1998) *J. Biol. Chem.* **273**, 23219–23224
257. Horowitz, B., and Meister, A. (1972) *J. Biol. Chem.* **247**, 6708–6719
257a. Aghaiypour, K., Wlodawer, A., and Lubkowski, J. (2001) *Biochemistry* **40**, 5655–5664
258. Tate, S. S., and Meister, A. (1971) *Adv. Enzymol.* **35**, 503–543
259. Funkhouser, J. D., Abraham, A., Smith, V. A., and Smith, W. G. (1974) *J. Biol. Chem.* **249**, 5478–5484
259a. Tedeschi, G., Ronchi, S., Simonic, T., Treu, C., Mattevi, A., and Negri, A. (2001) *Biochemistry* **40**, 4738–4744
259b. Bossi, R. T., Negri, A., Tedeschi, G., and Mattevi, A. (2002) *Biochemistry* **41**, 3018–3024
260. Rafalski, J. A., and Falco, S. C. (1988) *J. Biol. Chem.* **263**, 2146–2151
261. Zakin, M. M., Duchange, N., Ferrara, P., and Cohen, G. N. (1983) *J. Biol. Chem.* **258**, 3028–3031
262. Wedler, F. C., and Ley, B. W. (1993) *J. Biol. Chem.* **268**, 4880–4888
262a. James, C. L., and Viola, R. E. (2002) *Biochemistry* **41**, 3726–3731
263. Cassan, M., Parsot, C., Cohen, G. N., and Patte, J.-C. (1986) *J. Biol. Chem.* **261**, 1052–1057
264. Chen, N.-Y., and Paulus, H. (1988) *J. Biol. Chem.* **263**, 9526–9532
264a. Ehmann, D. E., Gehring, A. M., and Walsh, C. T. (1999) *Biochemistry* **38**, 6171–6177
265. Fujioka, M., Takata, Y., Ogawa, H., and Okamoto, M. (1980) *J. Biol. Chem.* **255**, 937–942
266. Bhattacharjee, J. K. (1992) in *The Evolution of Metabolic Function* (Mortlock, R. P., ed), pp. 47–80, CRC Press, Boca Raton, Florida
267. Mirwaldt, C., Korndörfer, I., and Huber, R. (1995) *J. Mol. Biol.* **246**, 227–239
268. Blickling, S., Beisel, H.-G., Bozic, D., Knäblein, J., Laber, B., and Huber, R. (1997) *J. Mol. Biol.* **274**, 608–621
269. Blickling, S., Renner, C., Laber, B., Pohlenz, H.-D., Holak, T. A., and Huber, R. (1997) *Biochemistry* **36**, 24–33
270. Scapin, G., Blanchard, J. S., and Sacchettini, J. C. (1995) *Biochemistry* **34**, 3502–3512
271. Reddy, S. G., Scapin, G., and Blanchard, J. S. (1996) *Biochemistry* **35**, 13294–13302
272. Simms, S. A., Voige, W. H., and Gilvarg, C. (1984) *J. Biol. Chem.* **259**, 2734–2741
273. Beaman, T. W., Binder, D. A., Blanchard, J. S., and Roderick, S. L. (1997) *Biochemistry* **36**, 489–494
274. Beaman, T. W., Blanchard, J. S., and Roderick, S. L. (1998) *Biochemistry* **37**, 10363–10369
275. Scapin, G., Cirilli, M., Reddy, S. G., Gao, Y., Vederas, J. C., and Blanchard, J. S. (1998) *Biochemistry* **37**, 3278–3285
276. Cirilli, M., Zheng, R., Scapin, G., and Blanchard, J. S. (1998) *Biochemistry* **37**, 16452–16458
276a. Koo, C. W., Sutherland, A., Vederas, J. C., and Blanchard, J. S. (2000) *J. Am. Chem. Soc.* **122**, 6122–6123
277. Daniel, R. A., and Errington, J. (1993) *J. Mol. Biol.* **232**, 468–483
278. Henderson, L. M., Nelson, P. J., and Henderson, L. (1982) *Fed. Proc.* **41**, 2843–2847
279. Dunn, W. A., Rettura, G., Seifter, E., and Englard, S. (1984) *J. Biol. Chem.* **259**, 10764–10770
280. Golebiewski, W. M., and Spenser, I. D. (1984) *J. Am. Chem. Soc.* **106**, 1441–1442
281. Kang, K. R., Wolff, E. C., Park, M. H., Folk, J. E., and Chung, S. I. (1995) *J. Biol. Chem.* **270**, 18408–18412
282. Park, M. H., Joe, Y. A., and Kang, K. R. (1998) *J. Biol. Chem.* **273**, 1677–1683
282a. Wolff, E. C., Wolff, J., and Park, M. H. (2000) *J. Biol. Chem.* **275**, 9170–9177
283. Luckner, M. (1972) *Secondary Metabolism in Plants and Animals*, Academic Press, New York (p. 275)
284. Byron, C. M., Stankovich, M. T., and Husain, M. (1990) *Biochemistry* **29**, 3691–3700
285. Soda, K., Misono, H., and Yamamoto, T. (1968) *Biochemistry* **7**, 4102–4109
286. Grove, J. A., Linn, T. G., Willett, C. J., and Henderson, L. M. (1970) *Biochim. Biophys. Acta.* **215**, 191–194
287. Markovitz, P. J., and Chuang, D. T. (1987) *J. Biol. Chem.* **262**, 9353–9358
288. Markovitz, P. J., Chuang, D. T., and Cox, R. P. (1984) *J. Biol. Chem.* **259**, 11643–11646
289. Fangmeier, N., and Leistner, E. (1980) *J. Biol. Chem.* **255**, 10205–10209
290. Wickwire, B. M., Harris, C. M., Harris, T. M., and Broquist, H. P. (1990) *J. Biol. Chem.* **265**, 14742–14747
291. Mihalik, S. J., McGuinness, M., and Watkins, P. A. (1991) *J. Biol. Chem.* **266**, 4822–4830
292. Wickwire, B. M., Wagner, C., and Broquist, H. P. (1990) *J. Biol. Chem.* **265**, 14748–14753
293. Rothstein, M. (1965) *Arch. Biochem. Biophys.* **111**, 467–476
294. Chang, Y. F., and Adams, E. (1977) *J. Biol. Chem.* **252**, 7987–7991
294a. Tang, K.-H., Harms, A., and Frey, P. A. (2002) *Biochemistry* **41**, 8767 – 8776
295. Ohsugi, M., Kahn, J., Hensley, C., Chew, S., and Barker, H. A. (1981) *J. Biol. Chem.* **256**, 7642–7651
296. Brzovic, P., Holbrook, E. L., Greene, R. C., and Dunn, M. F. (1990) *Biochemistry* **29**, 442–451
296a. Steegborn, C., Laber, B., Messerschmidt, A., Huber, R., and Clausen, T. (2001) *J. Mol. Biol.* **311**, 789–801
297. Clausen, T., Huber, R., Laber, B., Pohlenz, H.-D., and Messerschmidt, A. (1996) *J. Mol. Biol.* **262**, 202–224
298. Kery, V., Bukovska, G., and Kraus, J. P. (1994) *J. Biol. Chem.* **269**, 25283–25288
299. Taoka, S., Ohja, S., Shan, X., Kruger, W. D., and Banerjee, R. (1998) *J. Biol. Chem.* **273**, 25179–25184

References

299a. Mosharov, E., Cranford, M. R., and Banerjee, R. (2000) *Biochemistry* **39**, 13005 – 13011

299b. Jhee, K.-H., Niks, D., McPhie, P., Dunn, M. F., and Miles, E. W. (2001) *Biochemistry* **40**, 10873 – 10880

299c. Taoka, S., Lepore, B. W., Kabil, Ö., Ojha, S., Ringe, D., and Banerjee, R. (2002) *Biochemistry* **41**, 10454 – 10461

300. Nishi, N., Tanabe, H., Oya, H., Urushihara, M., Miyanaka, H., and Wada, F. (1994) *J. Biol. Chem.* **269**, 1015 – 1019

301. Takusagawa, F., Kamitori, S., Misaki, S., and Markham, G. D. (1996) *J. Biol. Chem.* **271**, 136 – 147

302. Takusagawa, F., Kamitori, S., and Markham, G. D. (1996) *Biochemistry* **35**, 2586 – 2596

302a. McQueney, M. S., Anderson, K. S., and Markham, G. D. (2000) *Biochemistry* **39**, 4443 – 4454

302b. Sánchez del Pino, M. M., Corrales, F. J., and Mato, J. M. (2000) *J. Biol. Chem.* **275**, 23476 – 23482

302c. Elrod, P., Zhang, J., Yang, X., Yin, D., Hu, Y., Borchardt, R. T., and Schowen, R. L., (2002) *Biochemistry* **41**, 8134 – 8142

303. Chen, Z., Crippen, K., Gulati, S., and Banerjee, R. (1994) *J. Biol. Chem.* **269**, 27193 – 27197

303a. Bandarian, V., and Matthews, R. G. (2001) *Biochemistry* **40**, 5056 – 5064

303b. Olteanu, H., and Banerjee, R. (2001) *J. Biol. Chem.* **276**, 35558 – 35563

304. Garrow, T. A. (1996) *J. Biol. Chem.* **271**, 22831 – 22838

304a. Thomas, D., Becker, A., and Surdin-Kerjan, Y. (2000) *J. Biol. Chem.* **275**, 40718 – 40724

305. Livesey, G. (1984) *Trends Biochem. Sci.* **9**, 27– 29

306. Soda, K., Tanaka, H., and Esaki, N. (1983) *Trends Biochem. Sci.* **8**, 214 – 217

307. Welch, G. N., and Loscalzo, J. (1998) *N. Engl. J. Med.* **338**, 1042 – 1050

308. Durand, P., Lussier-Cacan, S., and Blache, D. (1997) *FASEB J.* **11**, 1157 – 1168

309. Watanabe, M., Osada, J., Aratani, Y., Kluckman, K., Reddick, R., Malinow, M. R., and Maeda, N. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1585 – 1589

309a. Jakubowski, H. (1999) *FASEB J.* **13**, 2277 – 2283

309b. Ragone, R. (2002) *FASEB J.* **16**, 401 – 404

310. Finkelstein, J. D., and Martin, J. J. (1986) *J. Biol. Chem.* **261**, 1582 – 1587

311. James, F., Nolte, K. D., and Hanson, A. D. (1995) *J. Biol. Chem.* **270**, 22344 – 22350

311a. Ranocha, P., Bourgis, F., Ziemak, M. J., Rhodes, D., Gage, D. A., and Hanson, A. D. (2000) *J. Biol. Chem.* **275**, 15962 – 15968

312. Slany, R. K., Bösl, M., Crain, P. F., and Kersten, H. (1993) *Biochemistry* **32**, 7811 – 7817

313. Furfine, E. S., and Abeles, R. H. (1988) *J. Biol. Chem.* **263**, 9598 – 9606

314. Wray, J. W., and Abeles, R. H. (1995) *J. Biol. Chem.* **270**, 3147 – 3153

314a. Dai, Y., Pochapsky, T. C., and Abeles, R. H. (2001) *Biochemistry* **40**, 6379 – 6387

314b. Al-Mjeni, F., Ju, T., Pochapsky, T. C., and Maroney, M. J. (2002) *Biochemistry* **41**, 6761 – 6769

315. Wray, J. W., and Abeles, R. H. (1993) *J. Biol. Chem.* **268**, 21466 – 21469

316. Adams, D. O., and Yang, S. F. (1981) *Trends Biochem. Sci.* **6**, 161 – 164

317. Pirrung, M. C. (1983) *J. Am. Chem. Soc.* **105**, 7207 – 7209

318. Chang, C. (1996) *Trends Biochem. Sci.* **21**, 129 – 133

319. Gamble, R. L., Coonfield, M. L., and Schaller, G. E. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7825 – 7829

320. Hao, D., Ohme-Takagi, M., and Sarai, A. (1998) *J. Biol. Chem.* **273**, 26857 – 26861

320a. Tatsuki, M., and Mori, H. (2001) *J. Biol. Chem.* **276**, 28051 – 28057

320b. McCarthy, D. L., Capitani, G., Feng, L., Gruetter, M. G., and Kirsch, J. F. (2001) *Biochemistry* **40**, 12276 – 12284

321. Boller, T., and Kende, H. (1980) *Nature (London)* **286**, 259 – 260

322. Peiser, G. D., Wang, T.-T., Hoffman, N. E., Yang, S. F., Liu, H.-w, and Walsh, C. T. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3059 – 3063

323. Barlow, J. N., Zhang, Z., John, P., Baldwin, J. E., and Schofield, C. J. (1997) *Biochemistry* **36**, 3563 – 3569

324. Zhang, Z., Barlow, J. N., Baldwin, J. E., and Schofield, C. J. (1997) *Biochemistry* **36**, 15999 – 16007

324a. Brunhuber, N. M. W., Mort, J. L., Christoffersen, R. E., and Reich, N. O. (2000) *Biochemistry* **39**, 10730 – 10738

324b. Thrower, J. S., Blalock, R., III, and Klinman, J. P. (2001) *Biochemistry* **40**, 9717 – 9724

325. Pirrung, M. C., Kaiser, L. M., and Chen, J. (1993) *Biochemistry* **32**, 7445 – 7450

326. Blomstrom, D. C., and Beyer, E. M., Jr. (1980) *Nature (London)* **283**, 66 – 68

327. Liu, H.-W., Auchus, R., and Walsh, C. T. (1984) *J. Am. Chem. Soc.* **106**, 5335 – 5348

327a. Yao, M., Ose, T., Sugimoto, H., Horiuchi, A., Nakagawa, A., Wakatsuki, S., Yokoi, D., Murakami, T., Honma, M., and Tanaka, I. (2000) *J. Biol. Chem.* **275**, 34557 – 34565

328. Bousquet, J.-F., and Thimann, K. V. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1724 – 1727

329. Rabinowitz, K. W., Niederman, R. A., and Wood, W. A. (1973) *J. Biol. Chem.* **248**, 8207 – 8215

330. Hirose, K., Kawata, Y., Yumoto, N., and Tokushige, M. (1991) *J. Biochem.* **110**, 971 – 975

331. Eisenstein, E. (1991) *J. Biol. Chem.* **266**, 5801 – 5807

332. Eisenstein, E., Yu, H. D., Fisher, K. E., Iacuzio, D. A., Ducote, K. R., and Schwarz, F. P. (1995) *Biochemistry* **34**, 9403 – 9412

333. Samach, A., Hareven, D., Gutfinger, T., Ken-Dror, S., and Lifschitz, E. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2678 – 2682

334. Umbarger, H. E. (1956) *Science* **123**, 848

335. Liu, J.-Q., Dairi, T., Itoh, N., Kataoka, M., Shimizu, S., and Yamada, H. (1998) *J. Biol. Chem.* **273**, 16678 – 16685

336. Epperly, B. R., and Dekker, E. E. (1991) *J. Biol. Chem.* **266**, 6086 – 6092

337. Ray, M., and Ray, S. (1985) *J. Biol. Chem.* **260**, 5913 – 5918

338. Kelley, J. J., and Dekker, E. E. (1984) *J. Biol. Chem.* **259**, 2124 – 2129

338a. Cheong, C.-G., Escalante-Semerena, J. C., and Rayment, I. (2002) *Biochemistry* **41**, 9079 – 9089

339. Barak, Z., Chipman, D. M., and Schloss, J. V., eds. (1990) *Biosynthesis of Branched Chain Amino Acids*, VCH Publ., Basel

340. Umbarger, H. E. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 352 – 366, Am. Soc. Microbiology, Washington, DC

340a. Lee, Y.-T., and Duggleby, R. G. (2001) *Biochemistry* **40**, 6836 – 6844

340b. Mendel, S., Elkayam, T., Sella, C., Vinogradov, V., Vyazmensky, M., Chipman, D. M., and Barak, Z. (2001) *J. Mol. Biol.* **307**, 465 – 477

341. Lawther, R. P., Calhoun, D. H., Adams, C. W., Hauser, C. A., Gray, J., and Hatfield, G. W. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 922 – 925

342. Ray, T. B. (1986) *Trends Biochem. Sci.* **11**, 180 – 183

343. Halgand, F., Vives, F., Dumas, R., Biou, V., Andersen, J., Andrieu, J.-P., Cantegril, R., Gagnon, J., Douce, R., Forest, E., and Job, D. (1998) *Biochemistry* **37**, 4773 – 4781

344. Roy, T. B. (1986) *Trends Biochem. Sci.* **11**, 180 – 183

345. Schloss, J. V., van Dyk, D. E., Vasta, J. F., and Kutny, R. M. (1985) *Biochemistry* **24**, 4952 – 4959

346. Biou, V., Dumas, R., Cohen-Addad, C., Douce, R., Job, D., and Pebay-Peyroula, E. (1997) *EMBO J.* **16**, 3405 – 3415

347. Flint, D. H., and Emptage, M. H. (1988) *J. Biol. Chem.* **263**, 3558 – 3564

348. Ekiel, I., Smith, I. C. P., and Sprout, G. D. (1984) *Biochemistry* **23**, 1683 – 1687

349. Wallon, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D., and Petsko, G. A. (1997) *J. Mol. Biol.* **266**, 1016 – 1031

350. Powers, S. G., and Snell, E. E. (1976) *J. Biol. Chem.* **251**, 3786 – 3793

350a. Stolz, J., and Sauer, N. (1999) *J. Biol. Chem.* **274**, 18747 – 18752

350b. Yun, M., Park, C.-G., Kim, J.-Y., Rock, C. O., Jackowski, S., and Park, H.-W. (2000) *J. Biol. Chem.* **275**, 28093 – 28099

350c. Matak-Vinkovic, D., Vinkovic, M., Saldanha, S. A., Ashurst, J. L., von Delft, F., Inoue, T., Miguel, R. N., Smith, A. G., Blundell, T. L., and Abell, C. (2001) *Biochemistry* **40**, 14493 – 14500

350d. Zheng, R., and Blanchard, J. S. (2000) *Biochemistry* **39**, 16244 – 16251

350e. Zheng, R., and Blanchard, J. S. (2001) *Biochemistry* **40**, 12904 – 12912

350f. Strauss, E., Kinsland, C., Ge, Y., McLafferty, F. W., and Begley, T. P. (2001) *J. Biol. Chem.* **276**, 13513 – 13516

350g. Strauss, E., and Begley, T. P. (2001) *J. Am. Chem. Soc.* **123**, 6449 – 6450

350h. Mootz, H. D., Finking, R., and Marahiel, M. A. (2001) *J. Biol. Chem.* **276**, 37289 – 37298

350i. Hoenke, S., Schmid, M., and Dimroth, P. (2000) *Biochemistry* **39**, 13233 – 13240

350j. Schneider, K., Dimroth, P., and Bott, M. (2000) *Biochemistry* **39**, 9438 – 9450

351. Alexeev, D., Alexeeva, M., Baxter, R. L., Campopiano, D. J., Webster, S. P., and Sawyer, L. (1998) *J. Mol. Biol.* **284**, 401 – 419

351a. Eliot, A. C., Sandmark, J., Schneider, G., and Kirsch, J. F. (2002) *Biochemistry* **41**, 12582 – 12589

352. Gibson, K. J., Lorimer, G. H., Rendina, A. R., Taylor, W. S., Cohen, G., Gatenby, A. A., Payne, W. G., Roe, D. C., Lockett, B. A., Nudelman, A., Marcovici, D., Nachum, A., Wexler, B. A., Marsilii, E. L., Turner, I. M., Sr., Howe, L. D., Kalbach, C. E., and Chi, H. (1995) *Biochemistry* **34**, 10976 – 10984

353. Käck, H., Sandmark, J., Gibson, K. J., Schneider, G., and Lindqvist, Y. (1998) *Protein Sci.* **7**, 2560 – 2566

354. Gibson, K. J. (1997) *Biochemistry* **36**, 8474 – 8478

355. Käck, H., Gibson, K. J., Lindqvist, Y., and Schneider, G. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5495 – 5500

355a. Birch, O. M., Hewitson, K. S., Fuhrmann, M., Burgdorf, K., Baldwin, J. E., Roach, P. L., and Shaw, N. M. (2000) *J. Biol. Chem.* **275**, 32277 – 32280

355b. Ugulava, N. B., Gibney, B. R., and Jarrett, J. T. (2001) *Biochemistry* **40**, 8343 – 8351

355c. Ollagnier-de-Choudens, S., Mulliez, E., Hewitson, K. S., and Fontecave, M. (2002) *Biochemistry* **41**, 9145 – 9152

356. Hayden, M. A., Huang, I., Bussiere, D. E., and Ashley, G. W. (1992) *J. Biol. Chem.* **267**, 9512 – 9515

356a. Yennawar, N. H., Conway, M. E., Yennawar, H. P., Farber, G. K., and Hutson, S. M. (2002) *Biochemistry* **41**, 11592 – 11601

357. Turnbull, A. P., Baker, P. J., and Rice, D. W. (1997) *J. Biol. Chem.* **272**, 25105 – 25111

358. Ikeda, Y., and Tanaka, K. (1983) *J. Biol. Chem.* **258**, 9477 – 9487

References

359. Hawes, J. W., Crabb, D. W., Chan, R. M., Rougraff, P. M., and Harris, R. A. (1995) *Biochemistry* **34**, 4231–4237
360. Kamoun, P. (1992) *Trends Biochem. Sci.* **17**, 175–176
361. Wolf, D. A., and Akers, H. A. (1986) *Trends Biochem. Sci.* **11**, 390–392
362. Sabourin, P. J., and Bieber, L. L. (1982) *J. Biol. Chem.* **257**, 7460–7467
362a. Hutson, S. M., Cree, T. C., and Harper, A. E. (1978) *J. Biol. Chem.* **253**, 8126–8133
362b. Tischler, M. E., Desautels, M., and Goldberg, A. L. (1982) *J. Biol. Chem.* **257**, 1613–1621
362c. Mordier, S., Deval, C., Béchet, D., Tassa, A., and Ferrara, M. (2000) *J. Biol. Chem.* **275**, 29900–29906
362d. Mitch, W. E., and Goldberg, A. L. (1996) *N. Engl. J. Med.* **335**, 1897–1905
362e. Nissen, S., Sharp, R., Ray, M., Rathmacher, J. A., Rice, D., Fuller, J. C., Jr., Connelly, A. S., and Abumrad, N. (1996) *J. Appl. Physiol.* **81**, 2095–2104
362f. Nissen, S. L., and Abumrad, N. N. (1997) *J. Nutr. Biochem.* **8**, 300–311
363. Buckel, W. (1992) *FEMS Microbiol. Rev.* **88**, 211–232
364. Ho, C.-L., Noji, M., Saito, M., and Saito, K. (1999) *J. Biol. Chem.* **274**, 397–402
364a. Xue, H.-H., Fujie, M., Sakaguchi, T., Oda, T., Ogawa, H., Kneer, N. M., Lardy, H. A., and Ichiyama, A. (1999) *J. Biol. Chem.* **274**, 16020–16027
365. Snell, K. (1986) *Trends Biochem. Sci.* **11**, 241–243
366. Williams, H. E., and Smith, L. H., Jr. (1971) *Science* **171**, 390–391
366a. Lumb, M. J., Drake, A. F., and Danpure, C. J. (1999) *J. Biol. Chem.* **274**, 20587–20596
367. Danpure, C. J., and Purdue, P. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2385–2424, McGraw-Hill, New York
368. Noji, M., Inoue, K., Kimura, N., Gouda, A., and Saito, K. (1998) *J. Biol. Chem.* **273**, 32739–32745
369. Tai, C.-H., Yoon, M.-Y., Kim, S.-K., Rege, V. D., Nalabolu, S. R., Kredich, N. M., Schnackerz, K. D., and Cook, P. F. (1998) *Biochemistry* **37**, 10597–10604
369a. Hindson, V. J., Moody, P. C. E., Rowe, A. J., and Shaw, W. V. (2000) *J. Biol. Chem.* **275**, 461–466
370. Renwick, S. B., Snell, K., and Baumann, U. (1998) *Structure* **6**, 1105–1116
371. Vatcher, G. P., Thacker, C. M., Kaletta, T., Schnabel, H., Schnabel, R., and Baillie, D. L. (1998) *J. Biol. Chem.* **273**, 6066–6073
371a. Rontein, D., Nishida, I., Tashiro, G., Yoshioka, K., Wu, W.-I., Voelker, D. R., Basset, G., and Hanson, A. D. (2001) *J. Biol. Chem.* **276**, 35523–35529
372. Rathinasabapathi, B., Burnet, M., Russell, B. L., Gage, D. A., Liao, P.-C., Nye, G. J., Scott, P., Golbeck, J. H., and Hanson, A. D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3454–3458
373. Glaasker, E., Konings, W. N., and Poolman, B. (1996) *J. Biol. Chem.* **271**, 10060–10065
374. Klein, S. M., and Sagers, R. D. (1967) *J. Biol. Chem.* **242**, 297–300, 301–305
374a. Piper, M. D., Hong, S.-P., Ball, G. E., and Dawes, I. W. (2000) *J. Biol. Chem.* **275**, 30987–30995
375. Fujioka, M., Takata, Y., Konishi, K., and Ogawa, H. (1987) *Biochemistry* **26**, 5696–5702
376. Chlumsky, L. J., Sturgess, A. W., Nieves, E., and Jorns, M. S. (1998) *Biochemistry* **37**, 2089–2095
376a. Eschenbrenner, M., Chlumsky, L. J., Khanna, P., Strasser, F., and Jorns, M. S. (2001) *Biochemistry* **40**, 5352–5367
376b. Zhao, G., and Jorns, M. S. (2002) *Biochemistry* **41**, 9747 – 9750
377. Reuber, B. E., Karl, C., Reimann, S. A., Mihalik, S. J., and Dodt, G. (1997) *J. Biol. Chem.* **272**, 6766–6776
377a. Harris, R. J., Meskys, R., Sutcliffe, M. J., and Scrutton, N. S. (2000) *Biochemistry* **39**, 1189–1198
378. Shemin, D., and Rittenberg, D. (1946) *J. Biol. Chem.* **166**, 621–625
379. Gong, J., Hunter, G. A., and Ferreira, G. C. (1998) *Biochemistry* **37**, 3509–3517
379a. Tan, D., Barber, M. J., and Ferreira, G. C. (1998) *Protein Sci.* **7**, 1208–1213
380. Whiting, M. J., and Granick, S. (1976) *J. Biol. Chem.* **251**, 1340–1346
381. Warren, M. J., Cooper, J. B., Wood, S. P., and Shoolingin-Jordan, P. M. (1998) *Trends Biochem. Sci.* **23**, 217–221
381a. Ruiz de Mena, I., Fernández-Moreno, M. A., Bornstein, B., Kaguni, L. S., and Garesse, R. (1999) *J. Biol. Chem.* **274**, 37321–37328
382. Warren, M. J., and Scott, A. I. (1990) *Trends Biochem. Sci.* **15**, 486–491
383. Jordan, P. M., ed. (1991) *Biosynthesis of Tetrapyrroles*, Elsevier Science Publ., New York
384. Bevan, D. R., Bodlaender, P., and Shemin, D. (1980) *J. Biol. Chem.* **255**, 2030–2035
385. Jahn, D., Verkamp, E., and Söll, D. (1992) *Trends Biochem. Sci.* **17**, 215–218
385a. Moser, J., Schubert, W.-D., Beier, V., Bringemeier, I., Jahn, D., and Heinz, D. W. (2001) *EMBO J.* **20**, 6583–6590
386. Rieble, S., and Beale, S. I. (1991) *J. Biol. Chem.* **266**, 9740–9745
387. Hennig, M., Grimm, B., Contestabile, R., John, R. A., and Jansonius, J. N. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4866–4871
388. Dent, A. J., Beyersmann, D., Block, C., and Hasnain, S. S. (1990) *Biochemistry* **29**, 7822–7828
389. Jaffe, E. K., Volin, M., and Myers, C. B. (1994) *Biochemistry* **33**, 11554–11562
390. Cheung, K.-M., Spencer, P., Timko, M. P., and Shoolingin-Jordan, P. M. (1997) *Biochemistry* **36**, 1148–1156
390a. Erskine, P. T., Newbold, R., Brindley, A. A., Wood, S. P., Shoolingin-Jordan, P. M., Warren, M. J., and Cooper, J. B. (2001) *J. Mol. Biol.* **312**, 133–141
390b. Kervinen, J., Jaffe, E. K., Stauffer, F., Neier, R., Wlodawer, A., and Zdanov, A. (2001) *Biochemistry* **40**, 8227–8236
390c. Jaffe, E. K., Volin, M., Bronson-Mullins, C. R., Dunbrack, R. L. J., Kervinen, J., Martins, J., Quinlan, J. F., Jr., Sazinsky, M. H., Steinhouse, E. M., and Yeung, A. T. (2000) *J. Biol. Chem.* **275**, 2619–2626
391. Louie, G. V., Brownlie, P. D., Lambert, R., Cooper, J. B., Blundell, T. L., Wood, S. P., Warren, M. J., Woodcock, S. C., and Jordan, P. M. (1992) *Nature (London)* **359**, 33–39
392. Woodcock, S. C., and Jordan, P. M. (1994) *Biochemistry* **33**, 2688–2695
393. Porcher, C., Picat, C., Daegelen, D., Beaumont, C., and Grandchamp, B. (1995) *J. Biol. Chem.* **270**, 17368–17374
394. Warren, M. J., Gul, S., Aplin, R. T., Scott, A. I., Roessner, C. A., O'Grady, P., and Shoolingin-Jordan, P. M. (1995) *Biochemistry* **34**, 11288–11295
394a. Mathews, M. A. A., Schubert, H. L., Whitby, F. G., Alexander, K. J., Schadick, K., Bergonia, H. A., Phillips, J. D., and Hill, C. P. (2001) *EMBO J.* **20**, 5832–5839
395. Roméo, P.-H., Raich, N., Dubart, A., Beaupain, D., Pryor, M., Kushner, J., Cohen-Solal, M., and Goossens, M. (1986) *J. Biol. Chem.* **261**, 9825–9831
396. Whitby, F. G., Phillips, J. D., Kushner, J. P., and Hill, C. P. (1998) *EMBO J.* **17**, 2463–2471
396a. Martins, B. M., Grimm, B., Mock, H.-P., Huber, R., and Messerschmidt, A. (2001) *J. Biol. Chem.* **276**, 44108–44116
397. Proulx, K. L., Woodard, S. I., and Dailey, H. A. (1993) *Protein Sci.* **2**, 1092–1098
398. Camadro, J.-M., Thome, F., Brouillet, N., and Labbe, P. (1994) *J. Biol. Chem.* **269**, 32085–32091
399. Dailey, T. A., and Dailey, H. A. (1996) *Protein Sci.* **5**, 98–105
400. Ishida, T., Yu, L., Akutsu, H., Ozawa, K., Kawanishi, S., Seto, A., Inubushi, T., and Sano, S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4853–4858
401. Dailey, H. A., Sellers, V. M., and Dailey, T. A. (1994) *J. Biol. Chem.* **269**, 390–395
402. Ferreira, G. C., Franco, R., Lloyd, S. G., Pereira, A. S., Moura, I., Moura, J. J. G., and Huynh, B. H. (1994) *J. Biol. Chem.* **269**, 7062–7065
402a. Ferreira, G. C., Franco, R., Mangravita, A., and George, G. N. (2002) *Biochemistry* **41**, 4809–4818
402b. Lu, Y., Sousa, A., Franco, R., Mangravita, A., Ferreira, G. C., Moura, I., and Shelnutt, J. A. (2002) *Biochemistry* **41**, 8253 – 8262
403. Smith, A. G., Santana, M. A., Wallace-Cook, A. D. M., Roper, J. M., and Labbe-Bois, R. (1994) *J. Biol. Chem.* **269**, 13405–13413
404. Daily, H. A., and Fleming, J. E. (1986) *J. Biol. Chem.* **261**, 7902–7905
405. Saiki, K., Mogi, T., Hori, H., Tsubaki, M., and Anraku, Y. (1993) *J. Biol. Chem.* **268**, 26927–26934
405a. Brown, K. R., Allan, B. M., Do, P., and Hegg, E. L. (2002) *Biochemistry* **41**, 10906 – 10913
406. Warren, M. J., Gonzalez, M. D., Williams, H. J., Stolowich, N. J., and Scott, A. I. (1990) *J. Am. Chem. Soc.* **112**, 5343–5345
407. Yap-Bondoc, F., Bondoc, L. L., Timkovich, R., Baker, D. C., and Hebbler, A. (1990) *J. Biol. Chem.* **265**, 13498–13500
408. Brown, C. E., Shemin, D., and Katz, J. J. (1973) *J. Biol. Chem.* **248**, 8015–8021
409. Scott, A. I., Townsend, C. A., Okada, K., and Kajiwara, M. (1974) *J. Am. Chem. Soc.* **96**, 8054–8069 and 8069–8080
410. Battersby, A. R. (1994) *Science* **264**, 1551–1557
411. Spencer, J. B., Stolowich, N. J., Santander, P. J., Pichon, C., Kajiwara, M., Tokiwa, S., Takatori, K., and Scott, A. I. (1994) *J. Am. Chem. Soc.* **116**, 4991–4992
412. Thibaut, D., Debussche, L., and Blanche, F. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8795–8799
413. Eisenreich, W., and Bacher, A. (1991) *J. Biol. Chem.* **266**, 23840–23849
414. Scott, A. I., Stolowich, N. J., Wang, J., Gawatz, O., Fridrich, E., and Müller, G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14316–14319
415. Wang, J., Stolowich, N. J., Santander, P. J., Park, J. H., and Scott, A. I. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14320–14322
415a. Thompson, T. B., Thomas, M. G., Escalante-Semerena, J. C., and Rayment, I. (1999) *Biochemistry* **38**, 12995–13005
416. Michal, G., ed. (1999) *Biochemical Pathways*, Wiley–Spektrum Academischer Verlag, New York–Heidelberg, Germany
417. Bollivar, D. W., Suzuki, J. Y., Beatty, J. T., Dobrowolski, J. M., and Bauer, C. E. (1994) *J. Mol. Biol.* **237**, 622–640
418. Bollivar, D. W., Wang, S., Allen, J. P., and Bauer, C. E. (1994) *Biochemistry* **33**, 12763–12768

References

418a. Fujita, Y., and Bauer, C. E. (2000) *J. Biol. Chem.* **275**, 23583–23588

419. Jensen, P. E., Gibson, L. C. D., Henningsen, K. W., and Hunter, C. N. (1996) *J. Biol. Chem.* **271**, 16662–16667

420. Oster, U., Bauer, C. E., and Rüdiger, W. (1997) *J. Biol. Chem.* **272**, 9671–9676

421. Porra, R. J., and Meisch, H.-U. (1984) *Trends Biochem. Sci.* **9**, 99–104

422. Gibson, L. C. D., Willows, R. D., Kannangara, C. G., von Wettstein, D., and Hunter, C. N. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1941–1944

423. Tanaka, A., Ito, H., Tanaka, R., Tanaka, N. K., Yoshida, K., and Okada, K. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12719–12723

424. Scheumann, V., Schoch, S., and Rüdiger, W. (1998) *J. Biol. Chem.* **273**, 35102–35108

425. Warren, M. J., Jay, M., Hunt, D. M., Elder, G. H., and Röhl, J. C. G. (1996) *Trends Biochem. Sci.* **21**, 229–234

426. Brownlie, P. D., Lambert, R., Louie, G. V., Jordan, P. M., Blundell, T. L., Warren, M. J., Cooper, J. B., and Wood, S. P. (1994) *Protein Sci.* **3**, 1644–1650

427. Kappas, A., Sassa, S., Galbraith, R. A., and Nordmann, Y. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2103–2159, McGraw-Hill, New York

428. Macalpine, I., and Hunter, R. (1969) *Sci. Am.* **221**(Jul), 38–46

429. Lee, J.-S., and Anvret, M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10912–10915

430. Poland, A., and Glover, E. (1973) *Science* **179**, 476–477

430a. Bate, R., ed. (1999) *What Risk*, Butterworth-Heinemann, Oxford; Boston

431. Bohle, D. S., Dinnebier, R. E., Madsen, S. K., and Stephens, P. W. (1997) *J. Biol. Chem.* **272**, 713–716

432. Maines, M. D. (1988) *FASEB J.* **2**, 2557–2568

433. Matera, K. M., Takahashi, S., Fujii, H., Zhou, H., Ishikawa, K., Yoshimura, T., Rousseau, D. L., Yoshida, T., and Ikeda-Saito, M. (1996) *J. Biol. Chem.* **271**, 6618–6624

434. Liu, Y., Moënne-Loccoz, P., Loehr, T. M., and Ortiz de Montellano, P. R. (1997) *J. Biol. Chem.* **272**, 6909–6917

434a. Sakamoto, H., Omata, Y., Palmer, G., and Noguchi, M. (1999) *J. Biol. Chem.* **274**, 18196–18200

435. Gray, C. H. (1983) *Trends Biochem. Sci.* **8**, 381–384

435a. Cunningham, O., Dunne, A., Sabido, P., Lightner, D., and Mantle, T. J. (2000) *J. Biol. Chem.* **275**, 19009–19017

436. Yoshinaga, T., Sassa, S., and Kappas, A. (1982) *J. Biol. Chem.* **257**, 7803–7807

437. Zakhary, R., Gaine, S. P., Dinerman, J. L., Ruat, M., Flavahan, N. A., and Snyder, S. H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 795–798

437a. Doré, S., Takahashi, M., Ferris, C. D., Hester, L. D., Guastella, D., and Snyder, S. H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2445–2450

438. Wilks, A., and Schmitt, M. P. (1998) *J. Biol. Chem.* **273**, 837–841

439. Takahashi, M., Ilan, Y., Chowdhury, N. R., Guida, J., Horwitz, M., and Chowdhury, J. R. (1996) *J. Biol. Chem.* **271**, 26536–26542

440. Nogales, D., and Lightner, D. A. (1995) *J. Biol. Chem.* **270**, 73–77

441. McDonagh, A. F., Palma, L. A., and Lightner, D. A. (1980) *Science* **208**, 145–151

442. Brumm, P. J., Fried, J., and Friedmann, H. C. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3943–3947

443. Cooper, A. J. L. (1983) *Ann. Rev. Biochem.* **52**, 187–222

444. Klaassen, C. D., and Boles, J. W. (1997) *FASEB J.* **11**, 404–418

444a. Ullrich, T. C., and Huber, R. (2001) *J. Mol. Biol.* **313**, 1117–1125

444b. MacRae, I. J., Segel, I. H., and Fisher, A. J. (2001) *Biochemistry* **40**, 6795–6804

444c. Beynon, J. D., MacRae, I. J., Huston, S. L., Nelson, D. C., Segel, I. H., and Fisher, A. J. (2001) *Biochemistry* **40**, 14509–14517

445. Schmidt, A., Abrams, W. R., and Schiff, J. A. (1974) *Eur. J. Biochem.* **47**, 423–434

446. Benci, S., Vaccari, S., Mozzarelli, A., and Cook, P. F. (1997) *Biochemistry* **36**, 15419–15427

447. Cook, P. F., and Wedding, R. T. (1976) *J. Biol. Chem.* **251**, 2023–2029

447a. Soutourina, J., Blanquet, S., and Plateau, P. (2001) *J. Biol. Chem.* **276**, 40864–40872

448. Cooper, A. J. L., Haber, M. T., and Meister, A. (1982) *J. Biol. Chem.* **257**, 816–826

449. Akopyan, T. N., Braunstein, A. E., and Goryachenkova, E. V. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1617–1621

450. Weinstein, C. L., Haschemeyer, R. H., and Griffith, O. W. (1988) *J. Biol. Chem.* **263**, 16568–16579

451. Scriver, C. R. (1986) *N. Engl. J. Med.* **315**, 1155–1156

452. Colonge, M. J., Volpini, V., Bisceglia, L., Rousaud, F., de Sanctis, L., Beccia, E., Zelante, L., Testar, X., Zorzano, A., Estivill, X., Gasparini, P., Nunes, V., and Palacín, M. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9667–9671

453. Schneider, J. A. (1985) *N. Engl. J. Med.* **313**, 1473–1474

454. Weinstein, C. L., and Griffith, O. W. (1988) *J. Biol. Chem.* **263**, 3735–3743

455. Wright, C. E., Tallan, H. H., and Lin, Y. Y. (1986) *Ann. Rev. Biochem.* **55**, 427–453

456. Uchida, S., Kwon, H. M., Preston, A. S., and Handler, J. S. (1991) *J. Biol. Chem.* **266**, 9605–9609

457. Lombardini, J. B., Schaffer, S. W., and Azuma, J., eds. (1992) *Taurine Nutritional Value and Mechanisms of Action*, Plenum, New York

458. Pion, P. D., Kittleson, M. D., Rogers, Q. R., and Morris, J. G. (1987) *Science* **237**, 764–768

459. Geggel, H. S., Ament, M. E., Heckenlively, J. R., Martin, D. A., and Kopple, J. D. (1985) *N. Engl. J. Med.* **312**, 142–146

460. Thurston, H. H., Hauhart, R. E., and Naccarato, E. F. (1981) *Science* **214**, 1373–1374

461. Jarabak, R., and Westley, J. (1980) *Biochemistry* **19**, 900–904

462. Falany, C. N. (1997) *FASEB J.* **11**, 206–216

463. Miller, D. M., Delgado, R., Chirgwin, J. M., Hardies, S. C., and Horowitz, P. M. (1991) *J. Biol. Chem.* **266**, 4686–4691

463a. Gliubich, F., Gazerro, M., Zanotti, G., Delbono, S., Bombieri, G., and Berni, R. (1996) *J. Biol. Chem.* **271**, 21054–21061

463b. Bordo, D., Deriu, D., Colnaghi, R., Carpen, A., Pagani, S., and Bolognesi, M. (2000) *J. Mol. Biol.* **298**, 691–704

464. Zheng, L., Cash, V. L., Flint, D. H., and Dean, D. R. (1998) *J. Biol. Chem.* **273**, 13264–13272

465. Mihara, H., Kurihara, T., Yoshimura, T., Soda, K., and Esaki, N. (1997) *J. Biol. Chem.* **272**, 22417–22424

466. Nakai, Y., Yoshihara, Y., Hayashi, H., and Kagamiyama, H. (1998) *FEBS Lett.* **433**, 143–148

466a. Lacourciere, G. M., Mihara, H., Kurihara, T., Esaki, N., and Stadtman, T. C. (2000) *J. Biol. Chem.* **275**, 23769–23773

466b. Lange, H., Kaut, A., Kispal, G., and Lill, R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1050–1055

466c. Clausen, T., Kaiser, J. T., Steegborn, C., Huber, R., and Kessler, D. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3856–3861

466d. Smith, A. D., Agar, J. N., Johnson, K. A., Frazzon, J., Amster, I. J., Dean, D. R., and Johnson, M. K. (2001) *J. Am. Chem. Soc.* **123**, 11103–11104

466e. Krupka, H. I., Huber, R., Holt, S. C., and Clausen, T. (2000) *EMBO J.* **19**, 3168–3178

466f. Ogasawara, Y., Lacourciere, G., and Stadtman, T. C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9494–9498

467. Gentry-Weeks, C. R., Spokes, J., and Thompson, J. (1995) *J. Biol. Chem.* **270**, 7695–7702

468. Urbina, H. D., Silberg, J. J., Hoff, K. G., and Vickery, L. E. (2001) *J. Biol. Chem.* **276**, 44521–44526

469. Krebs, C., Agar, J. N., Smith, A. D., Frazzon, J., Dean, D. R., Huynh, B. H., and Johnson, M. K. (2001) *Biochemistry* **40**, 14069–14080

470. Garland, S. A., Hoff, K., Vickery, L. E., and Culotta, V. C. (1999) *J. Mol. Biol.* **294**, 8–

471. Tong, W.-H., and Rouault, T. (2000) *EMBO J.* **19**, 5692–5700

472. Kaut, A., Lange, H., Diekert, K., Kispal, G., and Lill, R. (2000) *J. Biol. Chem.* **275**, 15955–15961

473. Lutz, T., Westermann, B., Neupert, W., and Herrmann, J. M. (2001) *J. Mol. Biol.* **307**, 815–825

473a. Nakai, Y., Nakai, M., Hayashi, H., and Kagamiyama, H. (2001) *J. Biol. Chem.* **276**, 8314–8320

473b. Schwartz, C. J., Djaman, O., Imlay, J. A., and Kiley, P. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9009–9014

473c. Voisine, C., Cheng, Y. C., Ohlson, M., Schilke, B., Hoff, K., Beinert, H., Marszalek, J., and Craig, E. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1483–1488

473d. Wu, G., Mansy, S. S., Wu, S.-P., Surerus, K. K., Foster, M. W., and Cowan, J. A. (2002) *Biochemistry* **41**, 5024–5032

473e. Wu, G., Mansy, S. S., Wu, S-p., Surerus, K. K., Foster, M. W., and Cowan, J. A. (2002) *Biochemistry* **41**, 5024 – 5032

473f. Nuth, M., Yoon, T., and Cowan, J. A. (2002) *J. Am. Chem. Soc.* **124**, 8774 – 8775

473g. Mihara, H., Kato, S-i, Lacourciere, G. M., Stadtman, T. C., Kennedy, R. A. J. D., Kurihara, T., Tokumoto, U., Takahashi, Y., and Esaki, N. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6679 – 6683

474. Leimkühler, S., and Rajagopalan, K. V. (2001) *J. Biol. Chem.* **276**, 22024–22031

475. White, R. H. (1988) *Biochemistry* **27**, 7458–7462

475a. Tallant, T. C., Paul, L., and Krzycki, J. A. (2001) *J. Biol. Chem.* **276**, 4485–4493

475b. Krum, J. G., Ellsworth, H., Sargeant, R. R., Rich, G., and Ensign, S. A. (2002) *Biochemistry* **41**, 5005 – 5014

476. Barrios, A. M., and Lippard, S. J. (2000) *J. Am. Chem. Soc.* **122**, 9172–9177

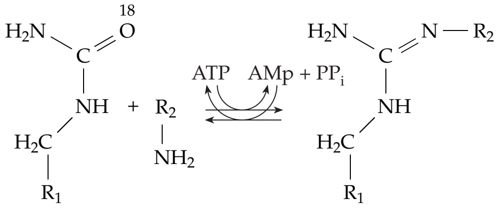
477. Rutherfurd, K. J., Rutherfurd, S. M., Moughan, P. J., and Hendriks, W. H. (2002) *J. Biol. Chem.* **277**, 114–119

478. Hans, M., Bill, E., Cirpus, I., Pierik, A. J., Hetzel, M., Alber, D., and Buckel, W. (2002) *Biochemistry* **41**, 5873 – 5882

Study Questions

1. Bacterial glutamine synthetase is feedback inhibited by serine, glycine, and alanine. Explain specifically the connection between these amino acids and glutamine that would account for the logic of this inhibition.

2. Argininosuccinate synthetase catalyzes the following reaction:



I II

a) The reaction as shown is reversible. What metabolic stratagem is employed to drive the reaction from left to right?
b) It has been shown that the oxygen-18 from I (see structure) is transferred to the phosphate group of AMP. Propose a biochemical reaction mechanism to account for the transfer.

3. An organism has been discovered whose urea cycle does not include any reaction with aspartate. Both of the urea nitrogen atoms come directly from ammonia. All other components of the cycle are present and the net reaction is

$$CO_2 + 2\ NH_3 + 3\ ATP \rightarrow \text{urea} + 2\ ADP + 2\ P_i + AMP + PP_i$$

Explain how this is accomplished. Include plausible reaction mechanisms for any *new* steps proposed.

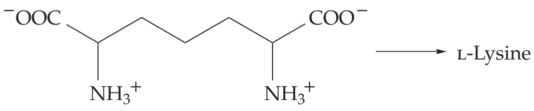
4. Which of the following compounds, if added to an active tissue preparation, might be expected to yield the greatest increase in urea production in terms of moles of urea produced per mole of added compound?

a) Ammonia
b) Bicarbonate
c) Aspartate
d) Ornithine

More than one answer may be correct. Explain how you decided.

5. On a given diet yielding 2500 kcal per day, a 70-kg man excretes 30.0 g of urea daily. What percentage of his daily energy requirement is met by protein? Assume that 1.0 g of protein yields 4.0 kcal and 0.16 g of nitrogen as urea.

6. In many organisms the immediate biosynthetic precursor of L-lysine is α,ε-diaminopimelic acid (structure below). What type of enzyme would catalyze this reaction; what coenzyme would be required; and what type of enzyme–substrate complex would be formed?



α,ε-Diaminopimelic acid

7. A possible mechanism for the action of urease is pictured in Fig. 16-25 and Eq. 16-47. Carbamate is thought to be one intermediate. Can you suggest an alternative possibility for the initial nickel ion-dependent steps. See Barrios and Lippard.[476]

8. Leucine is known as a "ketogenic" amino acid. Explain what this means.

9. In some organisms leucine is not ketogenic. Why?

10. Here is a possible metabolic reaction for a fungus.

$$\text{L-Leucine} + \text{2-oxoglutarate}^{2-} + 2\tfrac{1}{2}\ O_2 \rightarrow \text{L-glutamate}^{-} + \text{citrate}^{3-} + H_2O + 2\ H^+$$

$$\Delta G°' \text{ (pH 7)} = -1026 \text{ kJ/mol}$$

Suggest a metabolic pathway for this reaction. Is it thermodynamically feasible?

11. To be practical, the fungus should convert the L-glutamate back to 2-oxoglutarate using a glutamate dehydrogenase. Here are some values for Gibbs energies of formation from the elements under standard conditions (pH = 0).

2-Oxoglutarate^{2-}	−798.0 kJ/mol
L-Glutamate^{-}	−696.8 kJ/mol
NH_4^+	−79.5 kJ/mol

a) Calculate the apparent Gibbs energy change ΔG°′ (pH 7) for the following reaction:

$$\text{L-Leucine} + 3\ O_2 \rightarrow \text{citrate}^{3-} + NH_4^+ + H_2O + 2\ H^+$$

b) Could this fungal reaction be used as a commercial process for making citric acid?

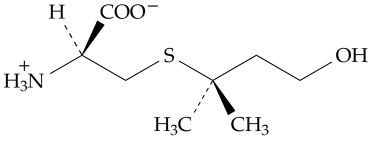
12. Two molecules of pyruvate can react to give a common precursor to valine, leucine, and pantoic acid. An isomerization step involving shift of a methyl group from one carbon to another is involved.

    a) Indicate as much as you can of the pathway for formation of valine showing coenzymes and mechanisms.
    b) Outline the reaction sequence by which the immediate oxoacid precursor of valine is converted to leucine.
    c) Outline a sequence by which the same oxoacid may be converted to pantoic acid using serine as an additional carbon atom source.

13. Explain how carboxylation can be coupled to cleavage of ATP and how this can be used to drive a metabolic sequence.

14. L-Serine is converted to pyruvate + NH_3 by serine dehydratase (deaminase) in a PLP-dependent reaction. However, using the same coenzyme selenocysteine is converted by selenocysteine lyase into L-alanine + elemental selenium Se^0. L-Cysteine may be converted by PLP-dependent enzymes into wither H_2S or into S^0 for transfer into metal clusters. Compare the chemical mechanisms.

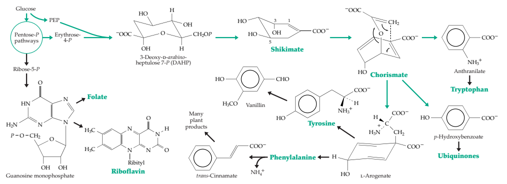
15. Felinine is found in urine of cats, the highest amounts in males. The compound arises from a reaction of glutathione. Propose a route of synthesis. See Rutherfurd *et al.*[477]



Felinine

16. Some clostridia ferment glutamate to ammonia, carbon dioxide, acetate, butyrate, and molecular hydrogen. Write a balanced equation and compare with Eq. 24-18 and Fig. 24-8. See Hans *et al.*[478]

17. How could β-hydroxy-β-methylbutyrate be formed in muscle? Could it be a physiologically important precursor to cholesterol?


Glucose
PEP
Pentose-P pathways
Erythrose-4-P
3-Deoxy-D-arabino-heptulose 7-P (DAHP)
Shikimate
Chorismate
Anthranilate
Tryptophan
Ribose-5-P
Folate
Guanosine monophosphate
Riboflavin
Ribityl
Vanillin
Tyrosine
Many plant products
trans-Cinnamate
Phenylalanine
L-Arogenate
p-Hydroxybenzoate
Ubiquinones

Aromatic compounds arise in several ways. The major route utilized by autotrophic organisms for synthesis of the aromatic amino acids, quinones, and tocopherols is the **shikimate pathway**. As outlined here, it starts with the glycolysis intermediate phosphoenolpyruvate (PEP) and erythrose 4-phosphate, a metabolite from the pentose phosphate pathway. Phenylalanine, tyrosine, and tryptophan are not only used for protein synthesis but are converted into a broad range of hormones, chromophores, alkaloids, and structural materials. In plants phenylalanine is deaminated to cinnamate which yields hundreds of secondary products. In another pathway ribose 5-phosphate is converted to pyrimidine and purine nucleotides and also to flavins, folates, molybdopterin, and many other pterin derivatives.

Contents

1421 .... **A. The Shikimate Pathway of Biosynthesis**
1423 .... 1. The Enzymes
1424 .... 2. From Chorismate to Phenylalanine and Tyrosine
1425 .... 3. Anthranilate, Tryptophan, *para*-Aminobenzoate, and Dihydroxybenzoate
1427 .... 4. Ubiquinones, Plastoquinones, Vitamin K, and Tocopherols
1428 .... **B. Metabolism of Phenylalanine and Tyrosine**
1428 .... 1. Catabolism of Tyrosine in Animals
1430 .... 2. The Thyroid Hormones
1432 .... 3. The Catecholamines
1434 .... 4. The Melanins
1434 .... 5. Microbial Catabolism of Phenylalanine, Tyrosine, and Other Aromatic Compounds
1438 .... 6. Quinic and Gallic Acids
1438 .... 7. The Metabolism of Phenylalanine and Tyrosine in Plants
1438 .... *Benzoic and salicylic acids*
1440 .... *Vanilla and other plant products*
1440 .... *Lignin, lignols, lignans, and phenolic coupling*
1443 .... 8. Alkaloids
1443 .... **C. Metabolism of Tryptophan and Histidine**
1443 .... 1. The Catabolism of Tryptophan
1446 .... *Formation of NAD*$^+$ *and NADP*$^+$
1446 .... *Auxin*
1447 .... *Alkaloids from tryptophan*
1449 .... 2. The Metabolism of Histidine
1449 .... *Regulation of histidine synthesis*
1450 .... *Catabolism of histidine*
1450 .... **D. Biosynthesis and Catabolism of Pyrimidines**
1451 .... 1. Synthesis of Pyrimidine Nucleotides
1452 .... *Deoxyribonucleotides*
1452 .... *Bacteriophage-induced alterations in metabolism*
1452 .... 2. Catabolism of Pyrimidine Nucleotides and Nucleosides
1453 .... 3. Reuse or Salvage of Nucleic Acid Bases
1454 .... **E. Biosynthesis and Metabolism of Purines**
1454 .... 1. The Enzymatic Reactions of Purine Synthesis
1456 .... 2. The Purine Nucleotide Cycle and Salvage Pathways for Purines
1459 .... 3. Oxidative Metabolism of Purines
1460 .... **F. Pterins, Flavins, Dimethylbenzimidazole, Thiamin, and Vitamin B_6**
1462 .... *Thiamin*
1463 .... *Pyridoxol (vitamin B_6)*

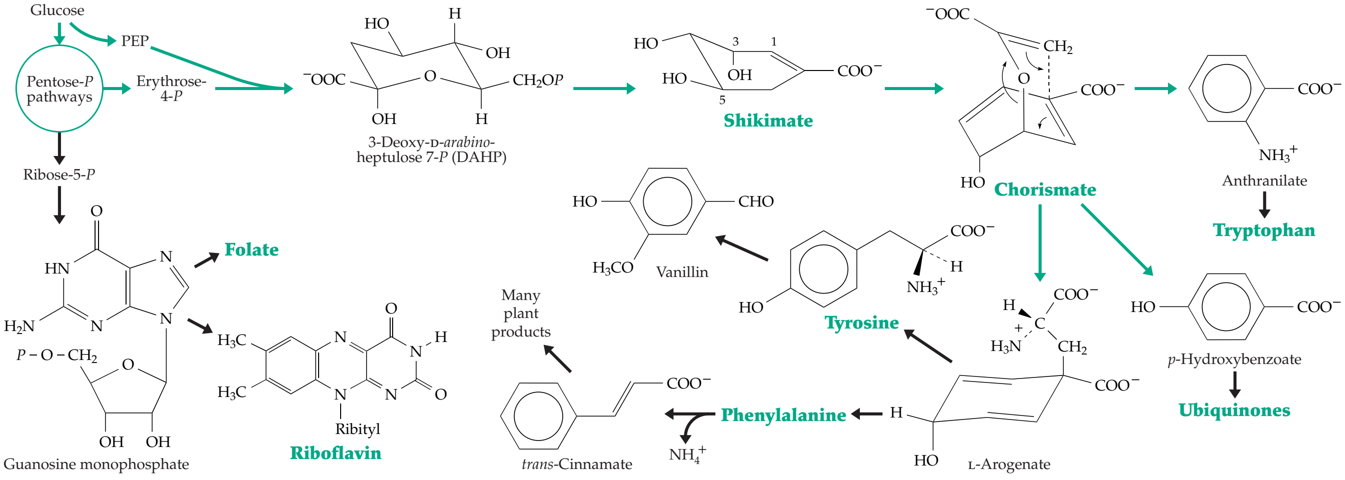
1465 .... **References**
1471 .... **Study Questions**

Boxes

1435 .... Box 25-A Skin Color
1442 .... Box 25-B The Cyanogenic Glycosides
1448 .... Box 25-C Royal Purple and Blue Denim
1459 .... Box 25-D Gout

Metabolism of Aromatic Compounds and Nucleic Acid Bases

25



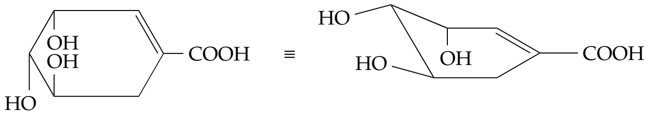
Aromatic rings are found in amino acids, purine and pyrimidine bases, vitamins, antibiotics, alkaloids, pigments of flowers and of the human skin, the lignin present in wood, and in many other substances. There are several biosynthetic pathways. One, which we met in Chapter 21, is the polyketide pathway. However, more important in most autotrophic organisms is the **shikimate pathway** which gives rise to phenylalanine, tyrosine, tryptophan, ubiquinone, plastoquinones, tocopherols, vitamin K, and other compounds.[1] The entire pathway, which is outlined in Fig. 25-1, is present in most bacteria and plants. However, animals are unable to synthesize the ring systems of the aromatic amino acids. Phenylalanine and tryptophan are dietary essentials. Tyrosine can also be formed in the animal body by the hydroxylation of phenylalanine. However, green plants lack phenylalanine hydroxylase and make tyrosine directly through the shikimate pathway (green arrows in Fig. 25-1).

There is interest in applying genetic engineering to increase the output of the shikimate pathway for production of industrially important aromatic compounds, e.g., the dye **indigo**, which is used in manufacture of blue denim (Box 25-C). Study of the enzymes involved has led to the development of potent inhibitors of the shikimate pathway which serve as widely used herbicides.[2,3]

A variety of pathways give rise to the nitrogen- and oxygen-containing heterocyclic rings of nature. All cells must be able to make pyrimidine and purine bases to be used in synthesis of nucleic acids and coenzymes. The pathway for synthesis of pyrimidine begins with aspartic acid and that for purines with glycine. In many organisms the pathway of purine formation is further enhanced because uric acid or a related substance is the major excretory product derived from excess nitrogen. This is true for both birds and reptiles, which excrete uric acid rather than urea, and for spiders which excrete guanine. In some plants, such as soy beans, the transport form of nitrogen is **allantoin** or **allantoic acid**, both of which are produced from uric acid.

A. The Shikimate Pathway of Biosynthesis

The shikimate pathway was identified through the study of ultraviolet light-induced mutants of *E. coli*, *Aerobacter aerogenes*, and *Neurospora*. In 1950, using the penicillin enrichment technique (Chapter 26), Davis obtained a series of mutants of *E. coli* that would not grow without the addition of aromatic substances.[4,5] A number of the mutants required five compounds: tyrosine, phenylalanine, tryptophan, *p*-aminobenzoic acid, and a trace of *p*-hydroxybenzoic acid. It was a surprise to find that the requirements for all five compounds could be met by the addition of shikimic acid, an aliphatic compound that was then regarded as a rare plant acid. Thus, shikimate was implicated as an intermediate in the biosynthesis of the three aromatic amino acids and of other essential aromatic substances.[6,7]



Shikimic acid

The mutants that grew in the presence of shikimic acid evidently had the biosynthetic pathway blocked

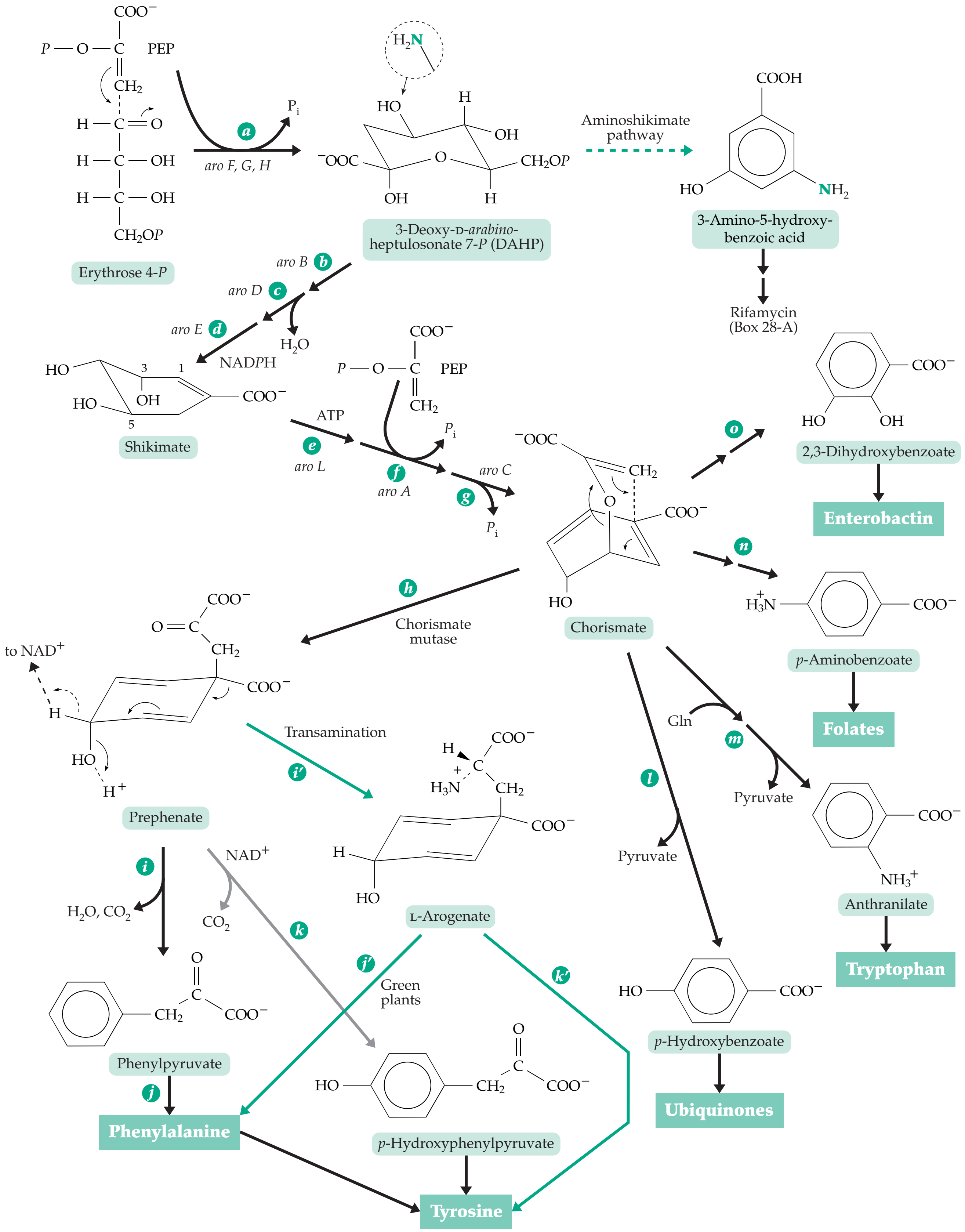
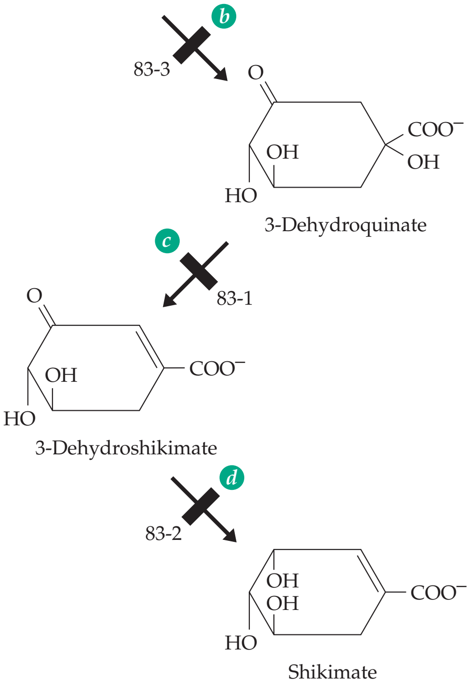

PEP
Erythrose 4-P
a
aro F, G, H
Pi
3-Deoxy-D-arabino-heptulosonate 7-P (DAHP)
Aminoshikimate pathway
3-Amino-5-hydroxy-benzoic acid
Rifamycin (Box 28-A)
aro B b
aro D c
H2O
aro E d
NADPH
Shikimate
ATP
e
aro L
PEP
Pi
f
aro A
aro C
g
Pi
Chorismate
o
2,3-Dihydroxybenzoate
Enterobactin
n
p-Aminobenzoate
Folates
h
Chorismate mutase
to NAD+
Prephenate
Transamination
i'
L-Arogenate
Gln
m
Pyruvate
Anthranilate
Tryptophan
l
Pyruvate
i
H2O, CO2
NAD+
CO2
k
j'
Green plants
k'
Phenylpyruvate
j
Phenylalanine
p-Hydroxyphenylpyruvate
Tyrosine
p-Hydroxybenzoate
Ubiquinones

Figure 25-1 Aromatic biosynthesis by the shikimate pathway. The symbols for several of the genes coding for the required enzymes are indicated. Their locations on the *E. coli* chromosome map are shown in Fig. 26-4. The aminoshikimate pathway which is initiated through 4-aminoDAHP leads to rifamycin and many other nitrogen-containing products.

at one or more earlier stages. Among these mutants, certain pairs were found that could not grow alone but that grew when plated together. The phenomenon is called **syntropism**. Mutant 83-2, which we now know to be blocked in the conversion of 5-dehydroshikimate to shikimate, accumulated dehydroshikimate and permitted mutant 83-1 or 83-3 to grow by providing it with a precursor that could be converted on to the end products (Eq. 25-1; the steps in this equation are lettered to correspond to those in Fig. 25-1). Eventually, the entire pathway was traced. The enzymes have all been isolated and studied[1] and the locations of the genes in the *E. coli* chromosome have been mapped[7–10] and are marked in Fig. 26-4.


b
83-3
O
COO⁻
OH
OH
HO
3-Dehydroquinate
c
83-1
O
OH
COO⁻
HO
3-Dehydroshikimate
d
83-2
OH
OH
COO⁻
HO
Shikimate

(25-1)

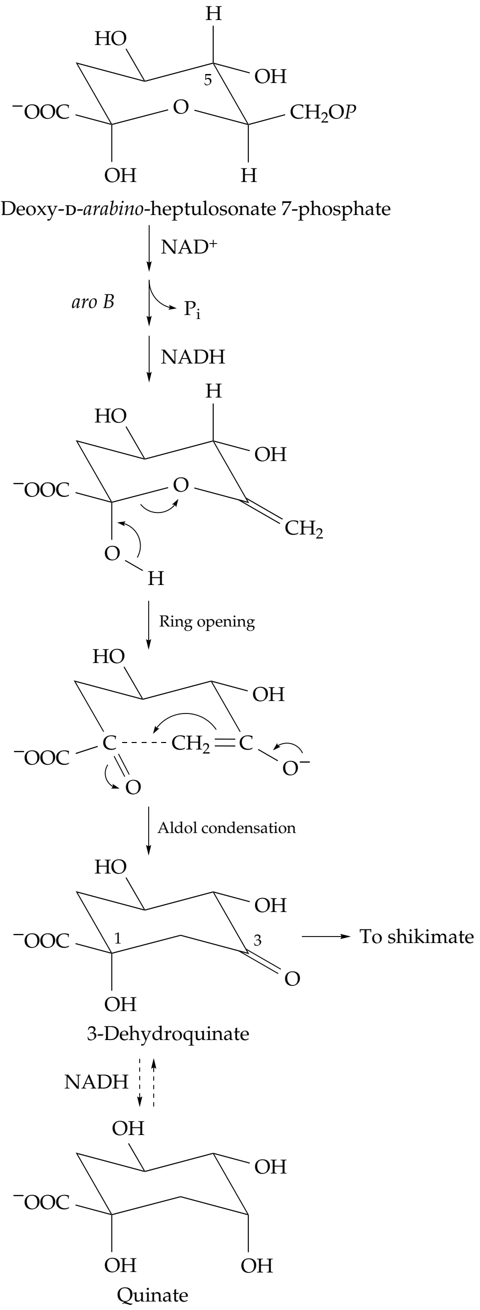
A variety of nitrogen-containing products, including rifamycin (Box 28-A), arise via the **aminoshikimate** pathway,[10a–d] which is also indicated in Fig. 25-1.

1. The Enzymes

The six carbons of the benzene ring of the aromatic amino acids are derived from the four carbons of erythrose 4-phosphate and two of the three carbons of phosphoenolpyruvate (PEP). The initial step in the pathway (Fig. 25-1, step *a*) is the condensation of erythrose 4-*P* with PEP and is catalyzed by 3-deoxy-D-*arabino*-heptulosonate-7-phosphate (DAHP) synthase. Closely analogous to an aldol condensation, the mechanism provides a surprise.[10] When PEP containing ^{18}O in the oxygen bridge to the phospho group reacts, the ^{18}O is retained in the eliminated phosphate; biochemical intuition would suggest that it should stay in the carbonyl group of the product. See Chapter 20, Section A,5.

Most bacteria and fungi have three isozymes of DAHP synthase, each controlled by feedback inhibition by one of the three products tyrosine, phenylalanine, or tryptophan. In *E. coli* these are encoded by genes *aro F*, *aro G*, and *aro H*, respectively.[11–12a] All of the enzymes contain one atom of iron per molecule and show spectral similarities to hemerythrin.[13]

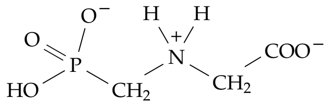
The product of the DAHP synthase, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate, is shown in its cyclic hemiacetal form at the beginning of Eq. 25-2. Its conversion to 3-dehydroquinate is a multistep process that is catalyzed by a single enzyme,[14,15] which is the product of *E. coli* gene *aro B*. The elimination of


H
HO
5
OH
⁻OOC
O
CH₂OP
OH
H
3-Deoxy-D-arabino-heptulosonate 7-phosphate
NAD⁺
aro B
Pᵢ
NADH
H
HO
OH
⁻OOC
O
CH₂
O
H
Ring opening
HO
OH
⁻OOC
C
CH₂=C
O⁻
O
Aldol condensation
HO
OH
⁻OOC
1
3
O
To shikimate
OH
3-Dehydroquinate
NADH
OH
OH
⁻OOC
OH
OH
Quinate

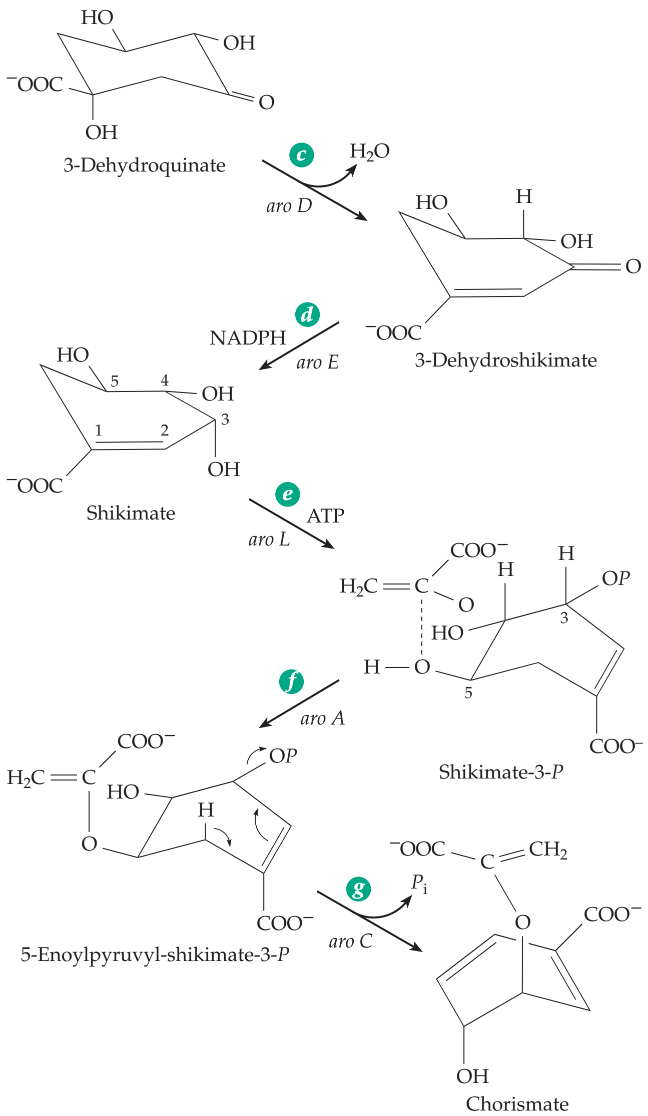
(25-2)

inorganic phosphate in the second step of the sequence is assisted by a transient oxidation of the hydroxyl group at C-5 to a carbonyl group.[6,16,17] The enzyme contains bound NAD$^+$ and is activated by Co^{2+}. The last step in the dehydroquinate synthase sequence is the addition of an enolic intermediate to a carbonyl group, an aldol condensation, which forms the 6-carbon ring (Eq. 25-2). Also indicated in this equation, with dashed lines, is the reversible conversion of dehydroquinate to quinate. Although it is a “side product” quinic acid accumulates to high concentrations in many plants.[18]

Dehydration of 3-dehydroquinate (step *c*), the first step in Eq. 25-3, is the first of three elimination reactions needed to generate the benzene ring of the end products. This dehydration is facilitated by the presence of the carbonyl group. After reduction of the product to shikimate (step *d*)[19] a phosphorylation reaction (step *e*)[20,21] sets the stage for the future elimination of P$_i$. In step *f*, condensation with PEP adds three carbon atoms that will become the α, β, and carboxyl carbon atoms of phenylalanine and tyrosine. The reaction occurs by displacement of P$_i$ from the α-carbon atom of PEP and resembles a reaction (Eq. 20-6, step *a*) in the synthesis of *N*-acetylmuramic acid.[22,23] When the reaction is carried out in ^{3}H-containing water, tritium enters the methylene group,[10,24] suggesting an addition–elimination mechanism (Eq. 25-4).[25] The enzyme 5-enoylpyruvylshikimate 3-phosphate synthase (**EPSP synthase**), which catalyzes this reaction, is strongly inhibited by the commercially important herbicide ***N*-(phosphonomethyl)glycine** (glyphosate).[3,26–27a]



N-(Phosphonomethyl) glycine (glyphosate)



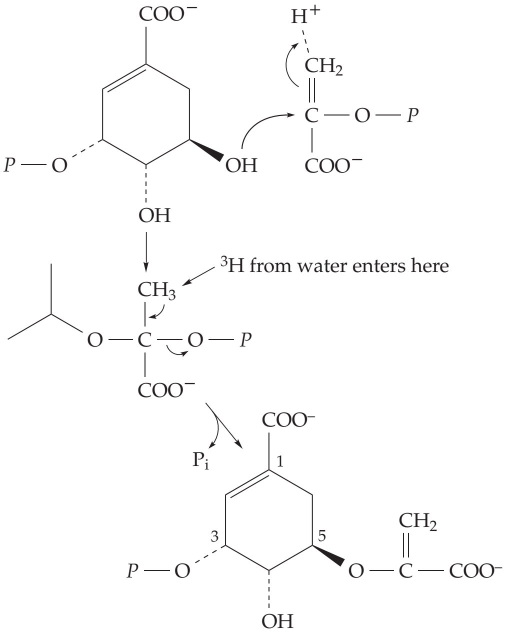
(25-3)

A single mutation (Pro 101 → Ser) in the 427-residue protein from *E. coli* makes the enzyme more resistant to the herbicide.[28] Other mutations affect binding and catalysis.[29]

Elimination of P$_i$ from 5-enolpyruvylshikimate 3-*P* (Eq. 25-3 and Fig. 25-1, step *g*) produces chorismate.[30] The 24-kDa chorismate synthase, which catalyzes this reaction, requires for activity a reduced flavin. Although there is no obvious need for an oxidation–reduction coenzyme, there is strong evidence that the flavin may play an essential role in catalysis, perhaps via a radical mechanism.[31–33b]

2. From Chorismate to Phenylalanine and Tyrosine

Chemical properties appropriate to a compound found at a branch point of metabolism are displayed by chorismic acid. Simply warming the compound in acidic aqueous solution yields a mixture of **prephenate** and ***para*-hydroxybenzoate** (corresponding to reactions *h* and *l* of Fig. 25-1). Note that the latter reaction is a simple elimination of the enolate anion of pyruvate. As indicated in Fig. 25-1, these reactions correspond to only two of several metabolic reactions of the chorismate ion. In *E. coli* the formation of **phenylpyruvate** (steps *h* and *i*, Fig. 25-1) is catalyzed by a single protein molecule with two distinctly different enzymatic activities: **chorismate mutase** and **prephenate dehydratase**.[34–36] However, in some organisms the enzymes are separate.[37] Both of the reactions catalyzed by these enzymes also occur spontaneously upon warming chorismic acid in acidic solution. The chorismate mutase reaction, which is unique in its mechanism,[37a] is discussed in Box 9-E. Stereochemical studies indicate that the formation of phenylpyruvate in Fig. 25-1, step *i*, occurs via a



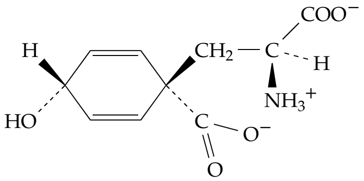
5-Enoylpyruvylshikimate 3-phosphate

(25-4)

chair-like transition state.[38] Phenylpyruvate is transaminated to phenylalanine to complete the biosynthesis of that amino acid. Regulation in *E. coli* is accomplished in part by feedback inhibition of the bifunctional chorismate mutase-prephenate dehydratase.[38a]

In *E. coli* and many other bacteria a second bifunctional enzyme, **chorismate mutase-prephenate dehydrogenase** causes the isomerization of chorismate and the oxidative decarboxylation of prephenate to *p*-hydroxyphenylpyruvate (steps *h* and *k*, Fig. 25-1).[39] The latter can be converted by transamination to tyrosine.[40–42]

A slightly different pathway for tyrosine formation was found initially in cyanobacteria but has now been identified in a variety of organisms including higher green plants. In this pathway prephenate undergoes transamination to **L-arogenate** (pretyrosine), step *i*, Fig. 25-1.[43–45] In bacteria L-arogenate is oxidatively decarboxylated to tyrosine (step *k*, Fig. 25-1). However, in green plants L-arogenate undergoes decarboxylative elimination (step *j*) to give L-phenylalanine. This is a major reaction in green plants, which cannot form tyrosine by hydroxylation of phenylalanine[46,47] but which form a variety of additional products from phenylalanine by a pathway characteristic for plants (Fig. 25-8).



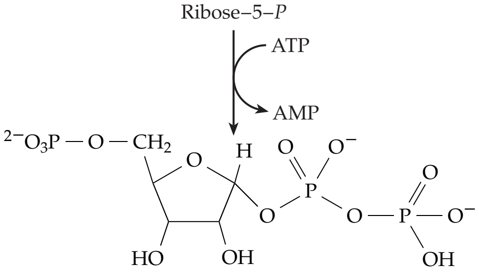
L-Arogenate (pretyrosine)

3. Anthranilate, Tryptophan, *para*-Aminobenzoate, and Dihydroxybenzoate

The three chemically similar reactions designated *m*, *n*, and *o* in Figs. 25-1 and 25-2 give rise to a variety of products. Step *m* leads to anthranilate and **L-tryptophan** and step *n* to the vitamin **folic acid.** Each of the three reactions *m*, *n*, and *o* involves addition of either NH_3 or HO^- at a position *ortho* or *para* to the carboxyl group of chorismate with elimination of the 5-OH group. The structures of the three intermediate products are shown in Fig. 25-2.[48–50] The three enzymes have significant similarities in amino acid sequence. Anthranilate synthase and *p*-aminobenzoate synthase are both two-subunit enzymes consisting of a 20-kDa subunit glutamine amidotransferase which is presumed to generate NH_3 (see Chapter 24,B).[49,51–53a] The second 50-kDa subunit of anthranilate synthase catalyzes the remaining steps in the reaction. However, *p*-aminobenzoate synthesis in *E. coli* requires an additional enzyme to catalyze the elimination of pyruvate in the final step[54] of synthesis.

The product of step *o* is known as **isochorismate**.[55,56] Isochorismate gives rise to a variety of products including **vitamin K**, salicylic acid,[56a] the iron chelator **enterobactin** (Fig. 16-1), and other siderophores. These are formed in *E. coli* via 2,3-dihydroxybenzoate as indicated in Fig. 25-2.[57–59a] The genes (*ent*) for the requisite enzymes are clustered at 14 min on the *E. coli* chromosome map (Fig. 26-4).

During the conversion of anthranilate to tryptophan, two additional carbon atoms must be incorporated to form the indole ring. These are derived from **phosphoribosyl pyrophosphate (PRPP)** which is formed from ribose 5-phosphate by transfer of a *pyrophospho* group from ATP.[60,61] The –OH group on the anomeric carbon of the ribose phosphate displaces AMP by attack on P_β of ATP (Eq. 25-5). In many organisms the enzyme that catalyzes this reaction is fused to subunit II of anthranilate synthase.[62] PRPP is also the donor of phosphoribosyl groups for biosynthesis of histidine (Fig. 25-13) and of nucleotides (Figs.



5-Phosphoribosyl 1-pyrophosphate (PRPP)

(25-5)

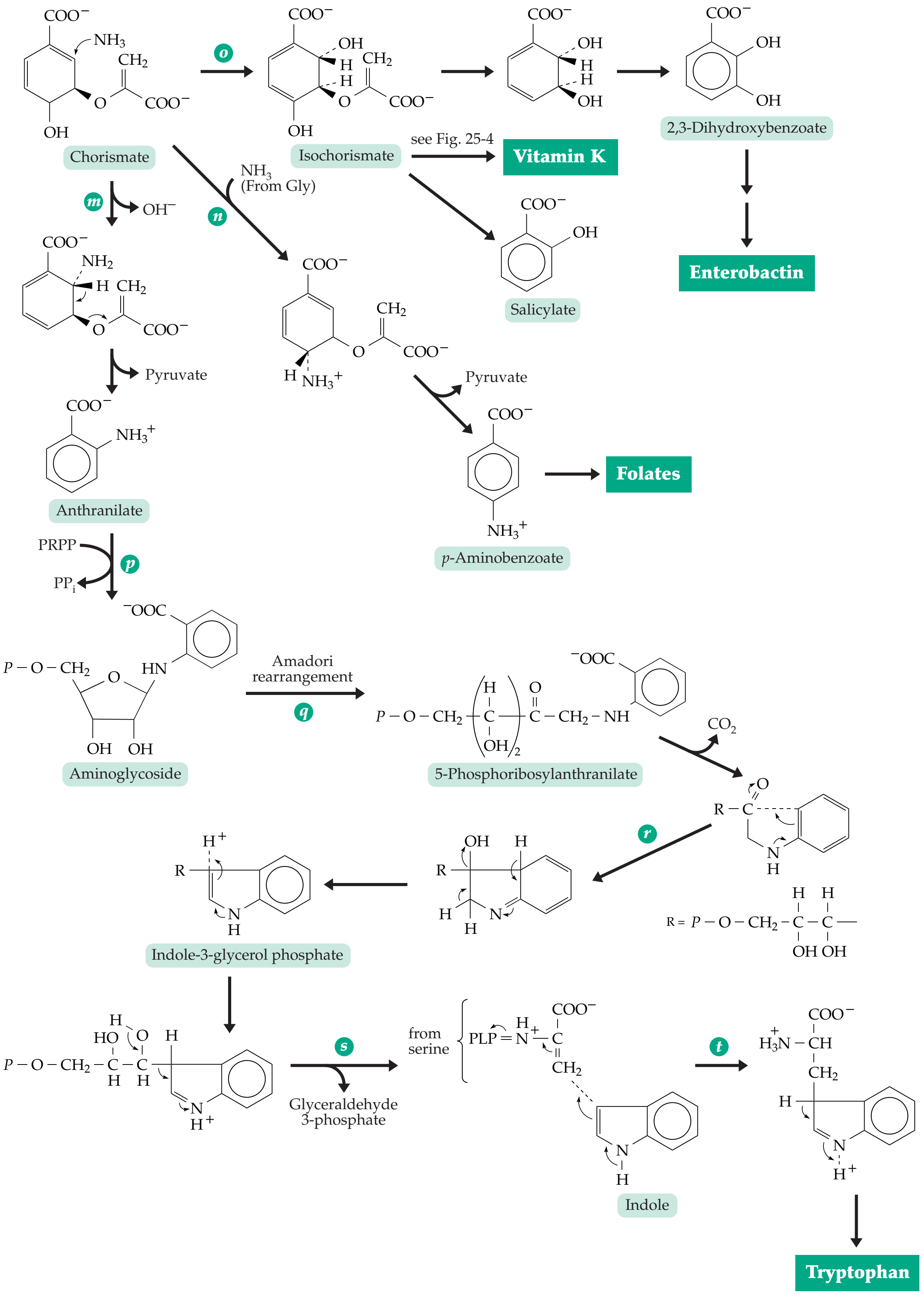

Chorismate
o
Isochorismate
see Fig. 25-4
Vitamin K
2,3-Dihydroxybenzoate
Enterobactin
m
OH⁻
n
NH3 (From Gly)
Salicylate
Pyruvate
Pyruvate
Anthranilate
p-Aminobenzoate
Folates
PRPP
PPi
p
Aminoglycoside
Amadori rearrangement
q
5-Phosphoribosylanthranilate
CO2
r
R = P – O – CH2 – C – C –
Indole-3-glycerol phosphate
s
Glyceraldehyde 3-phosphate
from serine
PLP = N
Indole
t
Tryptophan

Figure 25-2 The biosynthesis of tryptophan from chorismate and related synthetic reactions.

25-14, 15). In tryptophan biosynthesis PRPP is converted into an aminoglycoside of anthranilic acid by displacement of its pyrophospho group by the amino group of anthranilate (Fig. 25-2, step *p*). The aminoglycoside then undergoes an **Amadori rearrangement** (Eq. 4-8; Fig. 25-2, step *q*). The product has an open chain. Decarboxylation and ring closure, as indicated in this figure, yields **indoleglycerol phosphate**.[63,64]

A β replacement reaction catalyzed by the PLP-dependent **tryptophan synthase** converts indoleglycerol phosphate and serine to tryptophan. Tryptophan synthase from *E. coli* consists of two subunits associated as an $\alpha_2\beta_2$ tetramer (Fig. 25-3). The α subunit catalyzes the cleavage (essentially a reverse aldol) of indoleglycerol phosphate to glyceraldehyde 3-phosphate and free indole (Fig. 25-2, step *s*).[67] The β subunit contains PLP. It presumably generates, from serine, the Schiff base of aminoacrylate, as indicated in Fig. 25-2 (step *t*). The enzyme catalyzes the addition of the free indole to the Schiff base to form tryptophan. The indole must diffuse for a distance of 2.5 nm through a tunnel to the active site where it condenses with the aminoacrylate Schiff base.[65–68c]

The genes encoding the seven enzymes of the tryptophan biosynthetic pathway are organized as a single operon in some bacteria.[69] Its regulation in enteric bacteria is discussed in Chapter 28, Section A,5. The α subunit of tryptophan synthase and the enzymes catalyzing the preceding two steps in tryptophan synthesis are all $(\beta\alpha)_8$-barrel proteins similar to the one shown in Fig. 2-28.[69,70] The biosynthetic pathway for tryptophan in the green plant *Arabidopsis* is the same as in bacteria. The enzymes appear to be present in the chloroplasts.[71]

While the tryptophan synthase of *E. coli* is made up of two different subunits, that of *Neurospora* is a single polypeptide chain. This is one of the first proteins for which it was proposed that there were originally two separate genes, as in *E. coli*, but that they became fused during the course of evolution. After this proposal was made, gene fusion was demonstrated experimentally in *Salmonella* by introduction of two consecutive "frame shift mutations" between two genes of histidine biosynthesis (Chapter 26, Section B,1). Because of the frame shift, the stop signal for protein synthesis is no longer read, with the result that the organism makes a single long protein corresponding to both genes. Gene fusion evidently occurs in nature frequently.[63] There are many instances known in which the two distinctly different catalytic activities are possessed by the same protein in some organisms but by separate enzymes in others. The gene for the α subunit of trytophan synthase in *Salmonella* was of historical importance as it was used to establish the colinearity of the genetic code and its amino acid sequence (Chapter 26, Section B,5).

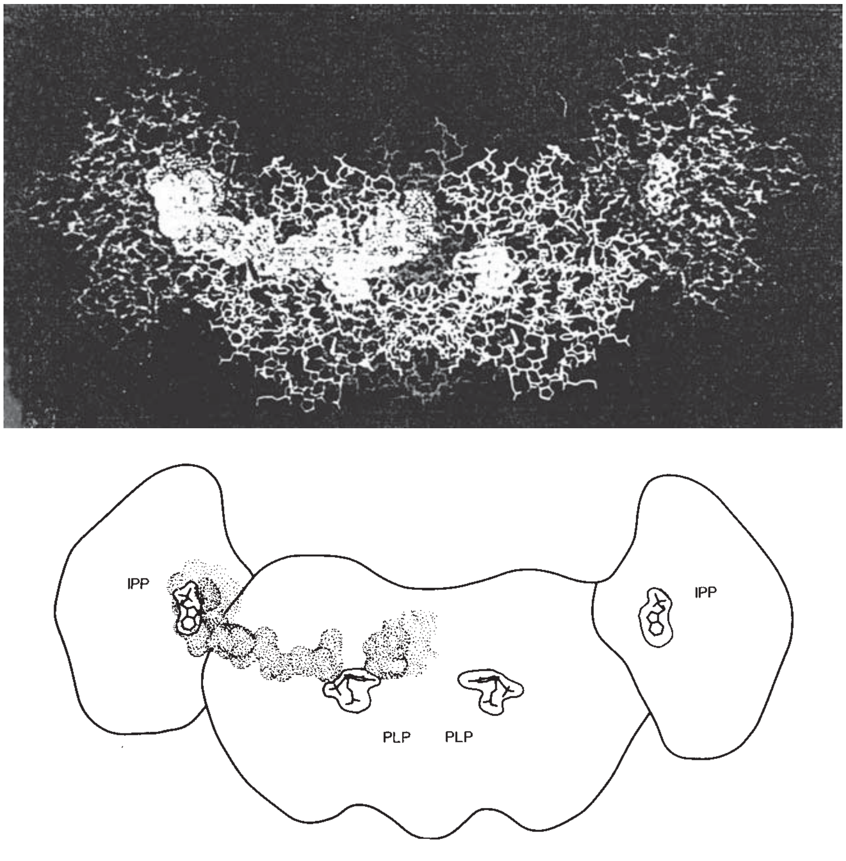


Figure 25-3 The structure of the two-enzyme $\alpha_2\beta_2$ complex tryptophan synthase.[65,66] The view is with the twofold axis of the $\alpha_2\beta_2$ complex vertical with the two α subunits at the ends and the β subunits in the center. The tunnel through which indole molecules released from indole propanol phosphate (IPP) in the α subunits to the pyridoxal phosphate (PLP) in the β subunits is shaded. Courtesy of C. Craig Hyde and Edith Wilson Miles.

4. Ubiquinones, Plastoquinones, Vitamin K, and Tocopherols

Radioactive carbon of [^{14}C]shikimate is efficiently incorporated into the quinones, vitamin K, and tocopherols. These chemically related redox agents (Fig. 15-24) also have a related biosynthetic origin, which has been elucidated in greatest detail for ubiquinone. In bacteria *p*-hydroxybenzoate is formed directly from chorismate (Fig. 25-1), but in plants it may originate from phenylalanine, *trans*-cinnamate, or *p*-coumarate as indicated in Fig. 25-4. The conversion of *p*-hydroxybenzoate on to the ubiquinones is also shown in this figure. A polyprenyl group is transferred onto a position ortho to the hydroxyl (see Chapter 22). Then a series of consecutive hydroxylation and *S*-adenosylmethionine-dependent transmethylation reactions lead directly to the ubiquinones.[72–73b] Several qui-nones that can serve as precursors to ubiquinones have been isolated from bacteria. Two of the corresponding quinols are shown as intermediates in Fig. 25-4. Chemical considerations suggest that both the methylations and hydroxylations occur on the reduced dihydroxy derivatives. A closely similar pathway is used for synthesis of ubiquinone[74–76] in mitochondria and in the membranes of the endoplasmic reticulum of fungal, plant, and animal cells.[77,78] In bacteria the decarboxylation step occurs early, as shown in Fig. 25-4, whereas in eukaryotes it occurs later.[75] Ubiquinone is poorly absorbed from the blood and it is apparently made in all aerobic tissues.[78] Ubiquinones are thought to serve as important antioxidant compounds in cell membranes. Dietary supplementation may be of value.[73a] Curiously, the nematode *C. elegans* slows its metabolism and lives longer if it has a defect in the hydroxylase catalyzing the next to last step in biosynthesis (Fig. 25-4).[73c] Mutants of *C. elegans* that cannot form their own ubiquinone-9 (containing nine prenyl units in the side chain) are unable to grow on bacteria that make ubiquinone-8. The worms appear to have both essential mitochondrial and nonmitochondrial requirements for ubiquinone.[73d] A ubiquinone deficiency with serious consequences can sometimes be caused in humans by inhibitors of 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase) taken to lower blood cholesterol.[79]

Labeling experiments have shown that the plastoquinones of chloroplasts as well as the tocopherols each bear one methyl group (marked with an asterisk in Fig. 25-4) that originates from chorismate. The dihydroxy compound **homogentisate** is probably an intermediate.[80–83] It is a normal catabolite of tyrosine in the animal body (Fig. 25-5, Eq. 18-49). Both prenylation and methylation by AdoMet are required to complete the synthesis of the plastoquinones and tocopherols. Possible biosynthetic intermediates with one or more double bonds in the polyprenyl side chain have been found in plants and also in fish oils.[83a]

The vitamins K and other naphthoquinones arise from ***O*-succinylbenzoate**[84–86] whose synthesis from chorismate and 2-oxoglutarate depends upon a thiamine diphosphate-bound intermediate, as indicated in Fig. 25-4. Elimination of pyruvate yields *O*-succinylbenzoate. The remaining reactions of decarboxylation, methylation, and prenylation (Fig. 25-4) resemble those of ubiquinone synthesis.

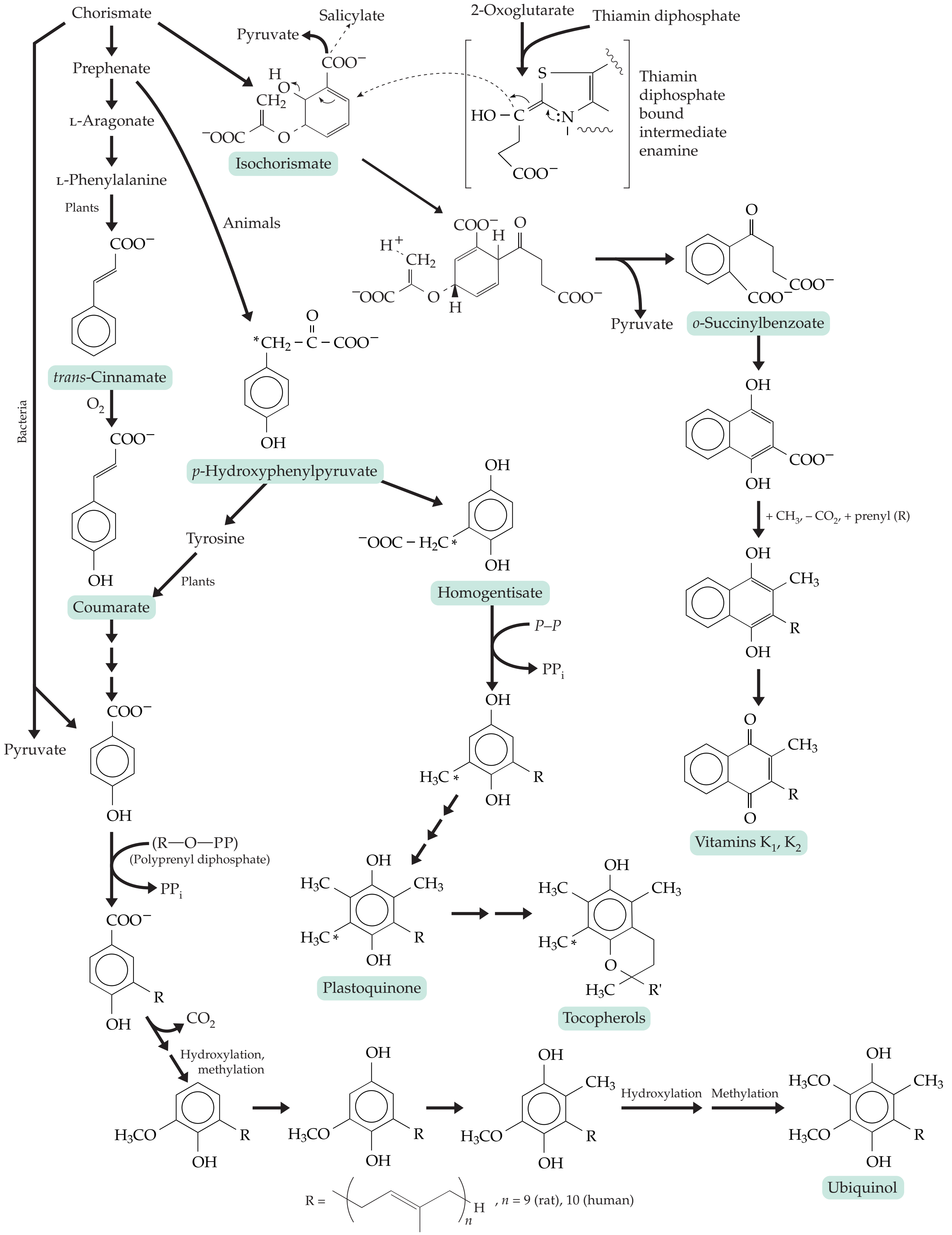
B. Metabolism of Phenylalanine and Tyrosine

Figure 25-5 shows the principal catabolic pathways, as well as a few biosynthetic reactions, of phenylalanine and tyrosine in animals. Transamination to phenylpyruvate (reaction *a*) occurs readily, and the product may be oxidatively decarboxylated to phenylacetate. The latter may be excreted after conjugation with glycine (as in Knoop's experiments in which phenylacetate was excreted by dogs after conjugation with glycine, Box 10-A). Although it does exist, this degradative pathway for phenylalanine must be of limited importance in humans, for an excess of phenylalanine is toxic unless it can be oxidized to tyrosine (reaction *b*, Fig. 25-5). Formation of phenylpyruvate may have some function in animals. The enzyme **phenylpyruvate tautomerase**, which catalyzes interconversion of enol and oxo isomers of its substrate, is also an important immunoregulatory cytokine known as **macrophage migration inhibitory factor**.[86a]

The pterin-dependent hydroxylation of phenylalanine to tyrosine (Eq. 18-45)[87,87a] has received a great deal of attention because of the occurrence of the metabolic disease **phenylketonuria (PKU)**,[88–91b] in which this reaction is blocked. Infants born with phenylketonuria appear normal but mental retardation sets in rapidly. However, if these infants are identified promptly and are reared on a low-phenylalanine diet which supplies only enough of the amino acid for essential protein synthesis, most brain damage can be prevented. Throughout most of the world every infant born is now given a simple urine test to identify phenylketonuria. Tolerance to phenylalanine increases with age, and adults may return to a near normal diet. However, there may still be problems with increased phenylalanine levels during fever and infections. A high phenylalanine level during pregnancy may damage the unborn child. Temporary insertion of multitubular reactors containing phenylalanine-ammonia lyase (Eq. 14-45) can be of value[92] as is administration of the enzyme in encapsulated form.[93] The mechanism by which phenylalanine damages the brain is uncertain.

1. Catabolism of Tyrosine in Animals

The major route of degradation of tyrosine in animals begins with transamination (Fig. 25-5, reaction *c*) to ***p*-hydroxyphenylpyruvate**. The enzyme tyrosine aminotransferase[94] is induced in the liver in response to the action of glucorticoid hormones (Chapter 22). The synthesis of the enzyme is also controlled at the translational level, release of the newly formed protein from liver ribosomes being stimulated by cyclic AMP. The enzyme is subject to posttranscriptional

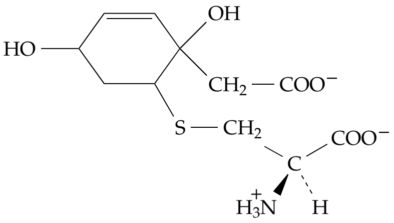


Chorismate
Prephenate
L-Aragonate
L-Phenylalanine
Plants
trans-Cinnamate
O_2
Coumarate
Bacteria
Pyruvate
Salicylate
Pyruvate
Isochorismate
2-Oxoglutarate
Thiamin diphosphate
Thiamin diphosphate bound intermediate enamine
Animals
p-Hydroxyphenylpyruvate
Tyrosine
Plants
Homogentisate
P–P
PP_i
Pyruvate
o-Succinylbenzoate
+ CH_3, – CO_2, + prenyl (R)
Vitamins K_1, K_2
(R—O—PP)
(Polyprenyl diphosphate)
PP_i
CO_2
Hydroxylation, methylation
Plastoquinone
Tocopherols
Hydroxylation
Methylation
Ubiquinol
R = , n = 9 (rat), 10 (human)

Figure 25-4 Pathways of biosynthesis of ubiquinones, plastoquinones, tocopherols, and vitamin K.

modification including phosphorylation and it undergoes unusually rapid turnover.[95,96]

The 2-oxoacid *p*-hydroxyphenylpyruvate is decarboxylated by the action of a dioxygenase (Eq. 18-49). The product **homogentisate** is acted on by a second dioxygenase, as indicated in Fig. 25-5, with eventual conversion to fumarate and acetoacetate. A rare metabolic defect in formation of homogentisate leads to tyrosinemia and excretion of **hawkinsin**[97] a compound postulated to arise from an epoxide (arene oxide) intermediate (see Eq. 18-47) which is detoxified by a glutathione transferase (Box 11-B).

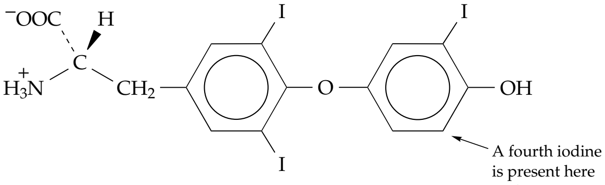


Hawkinsin

One of the first "inborn errors of metabolism" to be recognized was **alkaptonuria**, a lack of the oxygenase that cleaves the ring of homogentsic acid.[98] The condition is recognized by a darkening of the urine upon long standing (caused by oxidation of the homogentisate). Alkaptonuria was correctly characterized by Garrod (Box 1-D) in 1909 as a defect in the catabolism of tyrosine. Although relatively mild and not reducing the life span of individuals, it is nearly always accompanied by arthritis in later years and also by gray to bluish-black pigmentation of connective tissues, which may be visible through the shin or in the scleras (whites) of the eyes.[98] Absence of the next enzyme in the pathway, **maleylacetoacetate isomerase** causes one type of tyrosinemia.[99,99a] Absence of **fumarylacetoacetate hydrolase**, which acts on the product of the isomerase action causes the severe **type 1 hereditary tyrosinemia** which leads to accumulation of the toxic fumarylacetoacetate and its decarboxylation product succinylacetoacetate.[99,100]

2. The Thyroid Hormones

An important product of tyrosine metabolism in vetebrates is the thyroid hormone[101] of which the principal and most active forms are **thyroxine** (T_4) and **triiodothyronine** (T_3).[102] The thyroid gland is rich in iodide ion, which is actively concentrated from the plasma to ~1 µM free I^-.[103] This iodide reacts under the influence of a peroxidase (see Fig. 16-14 and accompanying discussion)[104] to iodinate tyrosyl residues of the very large ~660-kDa dimeric **thyroglobulin**, which is stored in large amounts in the lumen of the thyroid follicles.[105] Several of the tyrosine side chains (up to 15–25) are iodinated to form **mono-** and **diiodotyrosine** residues (Eq. 25-6), but only between four and eight of these, which are located at specific positions, are converted on to the hormones.[106,107]



Triiodothyronine

The coupling reaction by which the aromatic group from one residue of mono- or diiodotyrosine is joined in ether linkage with a second residue is also catalyzed readily by peroxidases. One dehydroalanine residue is formed for each molecule of hormone released.[108] A possible mechanism involves formation of an electron-deficient radical, which can undergo β elimination to produce a dehydroalanine residue and an aromatic radical. The latter could couple with a second radical to form triiodothyronine or thyroxine. However, as depicted in Eq. 25-6, the radical coupling may occur prior to chain cleavage. While β elimination (pathway A) has been favored, recent evidence suggests hydroxylation and cleavage to form a residue of aminomalonic semialdehyde in the thyroglobulin chain (pathway B).[108a] Alternatively, a PLP-dependent elimination of the radical could be used. Another possibility is oxidative attack on the 2-oxoacids derived from the iodotyrosines.

Thyroxine and triiodothyronine are released from thyroglobulin through the action of a series of proteases. Both the protease action and the release of the thyroid hormones into the bloodstream are stimulated by pituitary **thyrotropin (TSH)**.[109,110] Like glucagon thyrotropin is released from the pituitary in response to **thyrotropin-releasing hormone.**[111] Thyrotropin probably acts through cAMP-mediated mechanisms.[112] The hormones are carried throughout the body while bound to **thyroxine-binding globulin,** which serves as a carrier.[113] Some hormone is carried by other serum proteins such as **transthyretin** (thyroxine-binding prealbumin).[113,114] Both thyroxine and triiodothyronine have powerful hormonal effects on tissues, but the lag time for a response is shortest for triiodothyronine. Thus, it is thought that thyroxine undergoes loss of one iodine atom to form the more active triiodo form of the hormone within the target cells. Three **iodothyronine deiodinases**, all of which are selenoproteins, have been identified (Eq. 15-60).[115–116a]

Organically bound iodine is found in various invertebrates, but with one possible exception

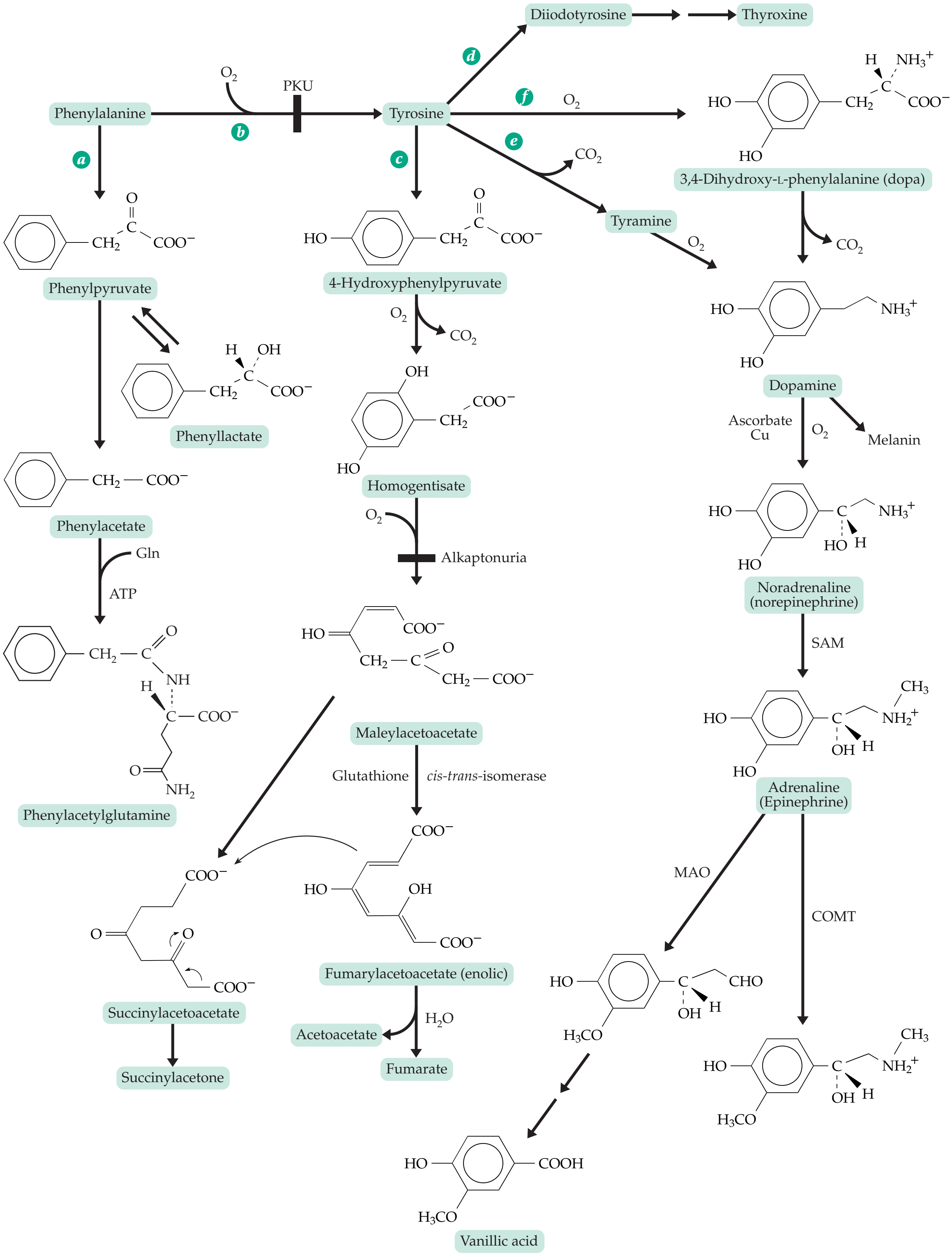
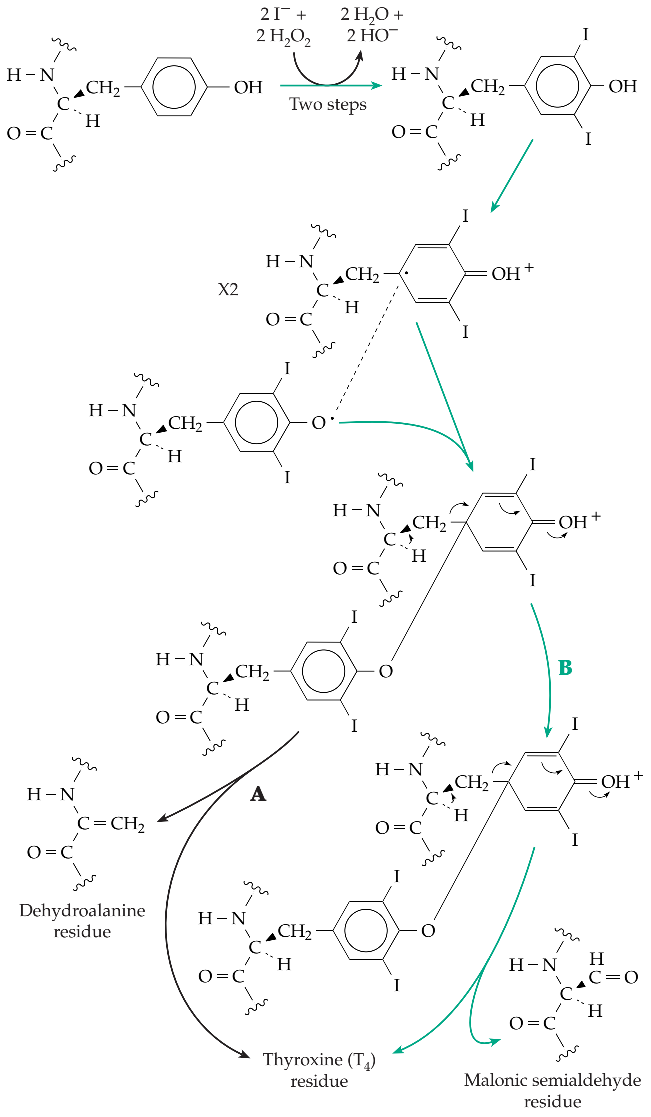

Phenylalanine
O_2
PKU
b
Tyrosine
a
c
d
e
f
Diiodotyrosine
Thyroxine
O_2
3,4-Dihydroxy-L-phenylalanine (dopa)
CO_2
CO_2
Tyramine
O_2
Phenylpyruvate
Phenyllactate
Phenylacetate
Gln
ATP
Phenylacetylglutamine
4-Hydroxyphenylpyruvate
O_2
CO_2
Homogentisate
O_2
Alkaptonuria
Maleylacetoacetate
Glutathione
cis-trans-isomerase
Fumarylacetoacetate (enolic)
H_2O
Acetoacetate
Fumarate
Succinylacetoacetate
Succinylacetone
Dopamine
Ascorbate
Cu
O_2
Melanin
Noradrenaline (norepinephrine)
SAM
Adrenaline (Epinephrine)
MAO
COMT
Vanillic acid

Figure 25-5 Some routes of metabolism of phenylalanine and tyrosine in animals.



(25-6)

thyroxine is present only in vertebrates.[101] Why do we need this iodinated hormone? Halogen-free analogs in which the iodine atoms of thyroxine and triiodothyronine have been replaced by methyl or other alkyl groups are biologically active. Frieden[101] concluded that the role of iodine is related more to the evolution of biosynthetic and catabolic pathways than to chemical properties of the hormones themselves.

Thyroxine and triiodothyronine have many effects, a major one in mammals and birds being stimulation of energy metabolism in tissues. It has long been recognized that a deficiency of thyroid hormone is reflected in an overall lower basal metabolic rate (Chapter 6). Maley and Lardy observed that thyroxine uncouples oxidative phosphorylation (Chapter 18) in isolated mitochondria.[117] When mitochondria from animals receiving extra thyroxine were compared with those from control animals, an increased rate of electron transport was observed. However, there was little or no change in the P / O ratio. Thus, the hormone apparently increased the rate of electron transport without decreasing the overall efficiency of ATP synthesis.

Thyroid hormones also have a general effect on growth and development in all vertebrates. This is especially striking in amphibia in which thyroid hormones control the metamorphosis from the tadpole to adult stages.[101,118] Zebrafish, and presumably also other fishes, require thyroid hormone to complete their life cycles.[119] At present it is thought that most, if not all, effects of thyroid hormones are a result of their action on the transcription of genes.[120–123] Nuclear **thyroid hormone receptors** belong to a family of ligand-regulated transcription factors that respond to steroid, retinoid, and thyroid hormones (Table 22-1).[123–124] These proteins control many metabolic functions, often forming heterodimers with other receptors and also being activated by coactivators[125] or corepressors.[124,125a,b] Transcription of the genes for the thyroxine-synthesis proteins **thyroperoxidase**, **thyroglobulin**, and **iodide transporter**[125c] is regulated by a **thyroid transcription factor**.[126]

A number of thyroid-related diseases are known. Thyroid deficiency is often evident by enlargement of the thyroid gland (**goiter**). The deficiencies may involve inadequacy in dietary intake of iodine, transport of iodide into the thyroid, poor formation of iodinated thyroglobulin, inefficient coupling to form the iodinated thyronine residues,[127] or mutations in thyroid hormone receptors.[122,128] A major cause of goiter is a deficiency in the content of iodine in soil, a condition affecting about one billion (10^9) persons. A more severe effect of thyroid deficiency is the fetal brain damage called **cretinism**.[125c,129] Victims are mentally retarded, deaf-mute, and often with motor rigidity. In Grave disease, the commonest type of **hyperthyroidism**, the blood contains specific thyroid-stimulating autoantibodies.[130,131] These bind to the thyrotropin (TSH) receptors of the thyroid plasma membrane and stimulate excessive formation of thyroid hormone.

3. The Catecholamines

A combination of decarboxylation and hydroxylation of the ring of tyrosine produces derivatives of *o*-dihydroxybenzene (catechol), which play important roles as neurotransmitters and are also precursors to **melanin**, the black pigment of skin and hair. Catecholamines may be formed by decarboxylation of tyrosine into tyramine (step *e*, Fig. 25-5) and subsequent oxidation. However, the quantitatively more important route is hydroxylation by the reduced pterin-dependent tyrosine hydroxylase (Chapter 18) to 3,4-dihydroxyphenylalanine, better known as **dopa**. The latter is decarboxylated to **dopamine**.[131a] Hydroxylation of dopamine by an ascorbic acid and copper-requiring enzyme (Eq. 18-53) produces the important hormone **noradrenaline** (norepinephrine), which is methylated to form **adrenaline** (epinephrine).

There are two principal catabolic routes for destruction of these catecholamines as is illustrated for adrenaline in Fig. 25-5. **Monoamine oxidase** (MAO)

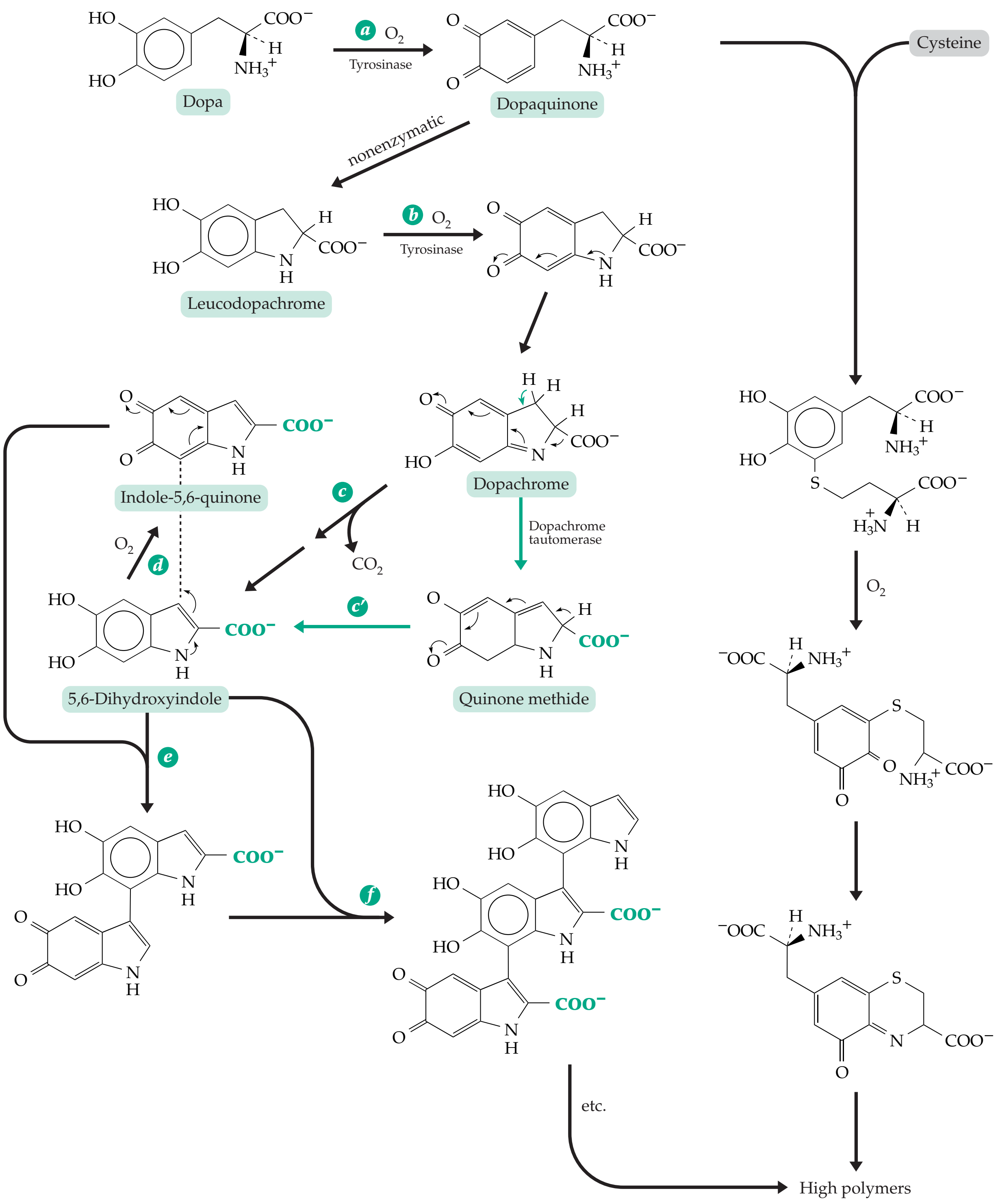

Dopa
a
O_2
Tyrosinase
Dopaquinone
Cysteine
nonenzymatic
Leucodopachrome
b
O_2
Tyrosinase
Indole-5,6-quinone
Dopachrome
c
CO_2
O_2
d
Dopachrome tautomerase
c′
5,6-Dihydroxyindole
Quinone methide
O_2
e
f
etc.
High polymers

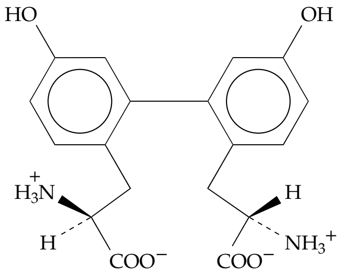
Figure 25-6 Postulated pathways for synthesis of the black pigment melanin and pigments (phaeomelanins) of reddish hair and feathers. Dopachrome reacts in two ways, with and without decarboxylation. The pathway without decarboxylation is indicated by green arrows. To the extent that this pathway is followed the green carboxylate groups will remain in the polymer. The black eumelanin is formed by reactions at the left and center while the reddish phaeomelanin is derived from polymers with cysteine incorporated by reactions at the right.

causes oxidative cleavage with deamination. Subsequent oxidative fission of the side chain together with methylation yields such end products as vanillic acid, which is excreted in the urine. The second catabolic route is immediate O-methylation by **catecholamine *O*-methyltransferase** (COMT), an active enzyme in neural tissues. The metabolites are relatively inactive physiologically and may be secreted as such or may undergo further oxidative degradation.

4. The Melanins

Dihydroxyphenylalanine (dopa) darkens rapidly when exposed to oxygen. The process is hastened greatly by tyrosinase (Chapter 16), which also catalyzes reaction *f* of Fig. 25-5, the oxidation of tyrosine to dopa. Tyrosinase is found in animals only in the organelles known as **melanosomes**, which are present in the melanin-producing **melanocytes** (Boxes 8-F; 25-A).[132] A series of enzymatic and nonenzymatic oxidation, decarboxylation, and coupling reactions forms the pigments. The initial steps for one pathway are indicated in Fig. 25-6. Oxidation of dopa to dopaquinone (step *a*)[133,134] is followed by an intramolecular addition reaction, together with tautomerization to the indole derivative, leucodopachrome. A second oxidation by tyrosinase (step *b*) is followed by decarboxylation and tautomerization to 5,6-dihydroxyindole (step *c*).[135] Alternatively, the tautomerization steps may take place without decarboxylation (green arrows, step *c′*).[133,136] In either case the dihydroxyindole that is formed can undergo a third oxidation step, also catalyzed by tyrosinase, to form indole-5,6-quinone (step *d*). Coupling of the last two products as indicated in step *e* yields a dimer, which is able to continue the addition of dihydroxyindole units (step *f*, etc.) with oxidation to form a high polymer of the black true melanin (**eumelanin**). However, the structure is not regular and is crosslinked. A related series of red polymers, the **phaeomelanins** found in red hair and feathers, are formed by addition of cysteine to dopaquinone.[137,138] Addition is possible at more than one position. The resulting adducts (only one is shown) can undergo oxidative ring closure in the manner indicated. Control of melanin formation is also complex. For example, more than 50 genetic loci affect the coat color of the house mouse.[139] Melanin in some fungi is formed by oxidative reactions of tetrahydroxynaphthalene formed via the polyketide pathways (Chapter 21).[140,140a] The melanin "inks" produced by cuttlefish and other cephalopods are formed in much the same way as melanins of skin.[140a]

Dopa is converted by at least some insects into *N*-β-alanyldopamine, which is a preferred substrate for the *o*-diphenol oxidase of the insect pupal cuticle. Oxidation of this substrate plays a crucial role in the hardening and darkening of the cuticle during pupal tanning.[141–142] There are many other oxidative reactions of tyrosine side chains within proteins. These include coupling of free radicals formed by peroxidases[143] or ultraviolet light[144,144a] to form dityrosines and other products. The walls of yeast ascospores,[145,146] the cements formed by reef-building annelids,[147,148] and adhesive plaques of marine mussels[149,149a] all contain polyphenolic proteins. The 120-kDa "foot protein" of the mussel *Mytilus edulis* consists of tandemly repeated decapeptides, each containing 2 residues of lysine, 1–2 residues of dopa, 1–2 residues of *trans*-4-hydroxyproline, and 1 residue of *trans*-2,3, *cis*-3,4-dihydroxyproline.[149]



Dityrosine
Other isomers are also formed

5. Microbial Catabolism of Phenylalanine, Tyrosine, and Other Aromatic Compounds

Bacteria and fungi play an essential role in the biosphere by breaking down the many aromatic products of plant metabolism.[150–153] These include vast amounts of lignin, alkaloids, flavonoid compounds, and other biochemically "inert" substances. Lignin is a major constituent of wood and a plant product second only to cellulose in abundance.

The chemical reactions used to degrade these aromatic compounds are numerous and complex. As was mentioned in Chapter 16, some fungi initiate the attack on lignin with peroxidases and produce soluble compounds that can be attacked by bacteria. In other cases elimination reactions may be used to initiate degradation. For example, some bacteria release phenol from tyrosine by β elimination (Fig. 14-5). However, more often hydroxylation and oxidative degradation of side chains lead to derivatives of benzoic acid or of the various hydroxybenzoic acids.[150,151,154–155a]

After the initial reactions many of the compounds are channeled into one of the major pathways illustrated in Fig. 25-7.[151,156,157] Dominant in aerobic bacteria is the conversion to **3-oxodipate** by one of the two convergent pathways shown. The products succinate and acetyl-CoA are readily oxidized by the citric

BOX 25-A SKIN COLOR

The principal pigment of human skin, hair, and eyes is **melanin**, which is synthesized in specialized cells, the **melanocytes**. They lie between the epidermis (outer layer) and the dermis (inner layer) as shown in Box 8-F. Melanocytes originate from embryonic nervous tissue and migrate into the skin by the third month of fetal life. They retain the highly branched morphology of neurons. Persons of different races all have the same numbers of melanocytes but the numbers and sizes of the pigmented melanosomes (Box 8-F) vary as does the content and chemical composition of the melanin.[a–d] Melanosomes not only are found in the dendrites of the melanocytes but are transferred from them into adjacent epithelial cells.[e,f]

Nevi (moles) are clusters of melanocytes that start to appear in the third year of life. They gradually increase in numbers but disappear in old age. **Freckles** appear beginning at about age six in genetically susceptible individuals. They are regions in which a higher concentration of melanin is formed.[a]

Both hair and the iris of the eye are also pigmented by melanin. Although dark eyes and dark hair are more prevalent among persons with dark skin there is no direct correlation. This is only one piece of evidence that the genetics of skin, eye, and hair coloration is complex. In mice over 150 different mutations occurring at more than 50 distinct genetic loci affect pigmentation.[f] Melanin formation begins with the action of tyrosinase. The human genome contains at least three genes for tyrosinase and related proteins.[b,g] The *Tyr* gene is absent in **oculocutaneous albinism**, the lack of pigment in eyes, hair, and skin. The tyrosinase-related protein 2 (TRP2), which has been identified as **dopachrome tautomerase** (see Fig. 25-6), is also a member of the tyrosinase family. Although a key enzyme in pigment synthesis, the amount of tyrosinase or of tyrosinase mRNA is the same in all skin types and colors.[h] Thus, differences in skin color must arise from differences in regulation.

Regulation of melanin formation is achieved in part by hormones, the **melanocyte-stimulating hormone** (MSH or melanotropin) being the most important.[a,f,g] The 13-residue pituitary hormone greatly increases pigmentation and stimulates differentiation of melanocytes. Other regulatory influences arise from interleukins, prostaglandins, interferons, tetrahydrobiopterins,[h] and protein kinase C.[i] Light also has a major effect, causing rapid tanning, especially in darker skin. Release of NO and cyclic GMP may be involved.[j]

Melanin and phaeomelanins have an important role in protecting skin from sunlight. This includes protection of light-sensitive vitamins, proteins, and DNA and RNA. The correlation of high pigmentation with the high intensity of light in tropical regions may reflect this property. Light-skinned persons of northern and southern latitudes, where light intensity is weaker, are less pigmented, allowing more adequate synthesis of vitamin D in the skin (Box 22-C).

A total lack of melanin as a result of a defective *Tyr* gene is seen in oculocutaneous albinism. Lacking protection from sunlight by melanin, albino individuals must shield their skin and eyes carefully. A second type of albinism results from mutations in the P gene, known in the mouse as the "pink-eyed dilution locus." In this condition synthesis of phaeomelanin is not impaired. Mutations in the *KIT* gene, which encodes a tyrosine kinase receptor lead to **piebaldism**, with white and dark splotched skin (or fur in animals).[b,k] While piebaldism is hereditary, **vitiligo** is an acquired autoimmune disease involving spotty loss of pigment and affecting 0.5 to 4% of the world's population. Melanocytes may be present in the affected areas but are unable to make melanin.[h,l]

---

[a] Lerner, A. B. (1961) *Sci. Am.* **205**(Jul), 98–108

[b] King, R. A., Hearing, V. J., Creel, D. J., and Oetting, W. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4353–4392, McGraw-Hill, New York

[c] Molnar, S. (1998) *Human Variation, Races, Types and Ethnic Groups*, 4th ed., Prentice Hall, Upper Saddle River, New Jersey (pp. 230–247)

[d] Robins, A. H. (1991) *Biological Perspectives on Human Pigmentation*, Cambridge Univ. Press, Cambridge

[e] Potterf, S. B., Muller, J., Bernardini, I., Tietze, F., Kobayashi, T., Hearing, V. J., and Gahl, W. A. (1996) *J. Biol. Chem.* **271**, 4002–4008

[f] Hearing, V. J., and Tsukamoto, K. (1991) *FASEB J.* **5**, 2902–2909

[g] Aroca, P., Urabe, K., Kobayashi, T., Tsukamoto, K., and Hearing, V. J. (1993) *J. Biol. Chem.* **268**, 25650–25655

[h] Schallreuter, K. U., Wood, J. M., Pittelkow, M. R., Gütlich, M., Lemke, K. R., Rödl, W., Swanson, N. N., Hitzemann, K., and Ziegler, I. (1994) *Science* **263**, 1444–1446

[i] Park, H.-Y., Russakovsky, V., Ohno, S., and Gilchrest, B. A. (1993) *J. Biol.Chem.* **268**, 11742–11749

[j] Roméro-Graillet, C., Aberdam, E., Biagoli, N., Massabni, W., Ortonne, J.-P., and Ballotti, R. (1996) *J. Biol. Chem.* **271**, 28052–28056

[k] Schmidt, A., and Beermann, F. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4756–4760

[l] Overwijk, W. W., Lee, D. S., Surman, D. R., Irvine, K. R., Touloukian, C. E., Chan, C.-C., Carroll, M. W., Moss, B., Rosenberg, S. A., and Restifo, N. P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2982–2987

acid cycle and associated reactions. Many different compounds can be converted into the starting compounds shown at the top of the figure. Both D- and L-mandelate, toluene, benzyl alcohol, L-tryptophan, phenanthrene, naphthalene, and benzene can be converted to catechol and be metabolized via the catechol branch of the pathway. Benzoate, *p*-toluate, shikimate, and quinate can be metabolized via the protocatechuate branch. Halogenated compounds, e.g., 3-chlorocatechol, may sometimes be degraded via

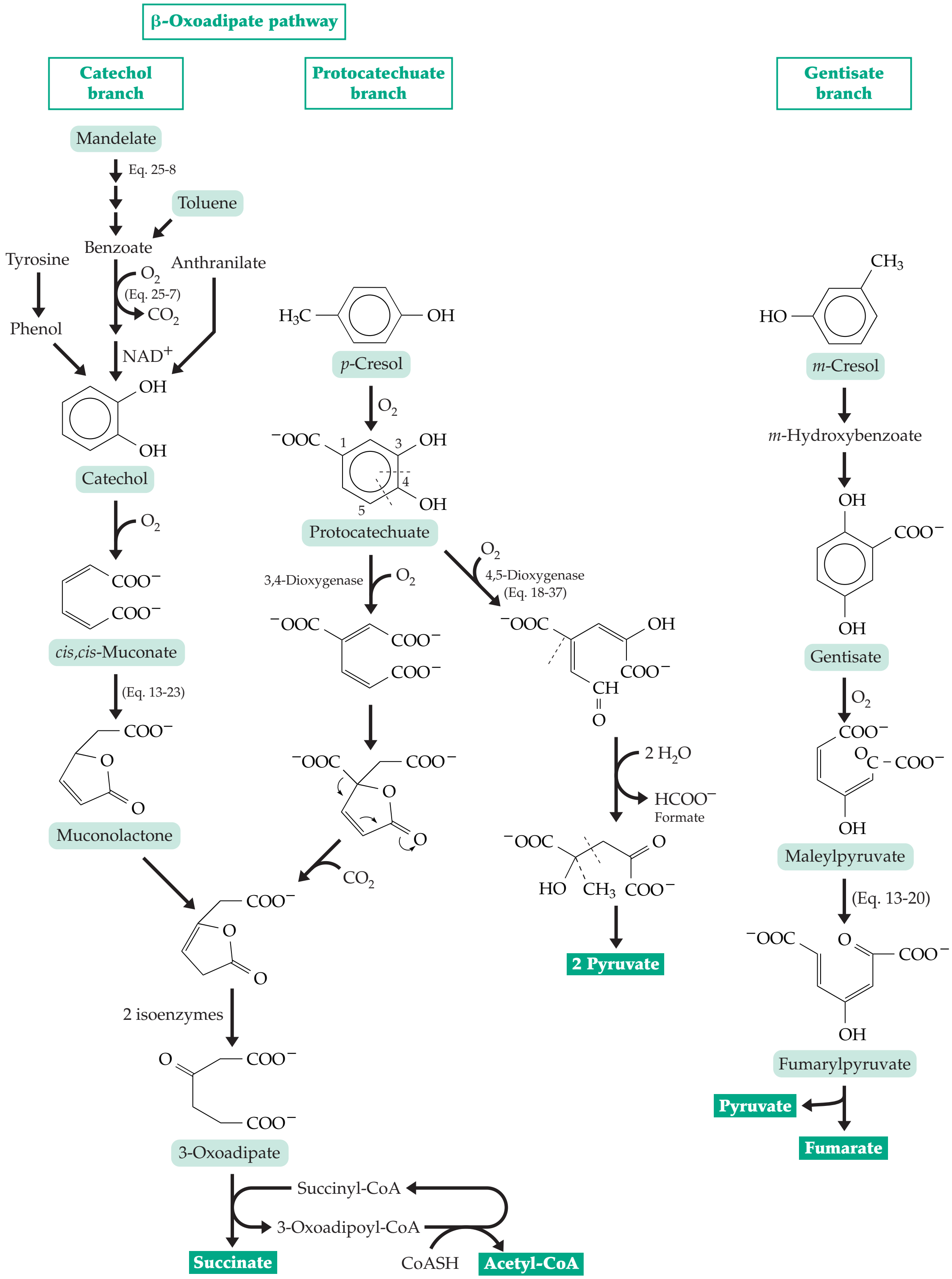
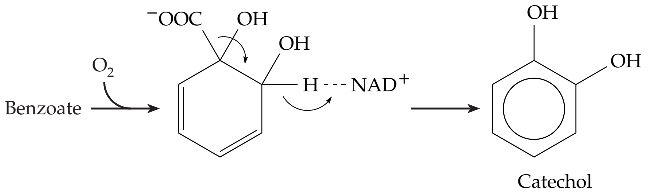

β-Oxoadipate pathway
Catechol branch
Protocatechuate branch
Gentisate branch
Mandelate
Eq. 25-8
Toluene
Benzoate
Tyrosine
Anthranilate
O_2
(Eq. 25-7)
CO_2
Phenol
NAD^+
OH
OH
Catechol
O_2
COO^-
COO^-
cis,cis-Muconate
(Eq. 13-23)
COO^-
O
O
Muconolactone
H_3C
OH
p-Cresol
O_2
^-OOC
1
3
OH
4
5
OH
Protocatechuate
3,4-Dioxygenase
O_2
^-OOC
COO^-
COO^-
^-OOC
COO^-
O
O
CO_2
COO^-
O
O
2 isoenzymes
O
COO^-
COO^-
3-Oxoadipate
Succinyl-CoA
3-Oxoadipoyl-CoA
Succinate
CoASH
Acetyl-CoA
O_2
4,5-Dioxygenase
(Eq. 18-37)
^-OOC
OH
COO^-
CH
O
2 H_2O
$HCOO^-$
Formate
^-OOC
O
HO
CH_3
COO^-
2 Pyruvate
CH_3
HO
m-Cresol
m-Hydroxybenzoate
OH
COO^-
OH
Gentisate
O_2
COO^-
O
C
COO^-
OH
Maleylpyruvate
(Eq. 13-20)
^-OOC
O
COO^-
OH
Fumarylpyruvate
Pyruvate
Fumarate

Figure 25-7 A few pathways of catabolism of aromatic substances by bacteria.

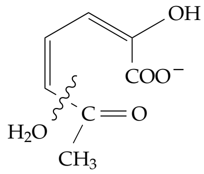
the same pathways[151,156] or in parallel or related pathways.[157–158c] However, chlorine atoms are sometimes eliminated as Cl⁻ at some point in the process.[159–160a] Other substrates, including *m*-hydroxybenzoate and sometimes anthranilate, are degraded via the gentisate pathway (Fig. 25-7).[157,161] Both benzoate and phenylacetate are sometimes degraded after conversion to coenzyme A thioesters.[161a,b]

Dioxygenases play a major role in all of these pathways of aromatic catabolism. In most cases a dioxygenase (Chapter 18) is required for the opening of the benzene ring. The pathways contain interesting isomerization steps, some of which have been discussed in Chapter 13, Section B. Microorganisms often have alternative choices in the chemistry of their attack. For example, tyrosine can be converted by one bacterium to homogentisate, as in animals (Fig. 25-5), or by other bacteria to protocatechuate, homoprotocatechuate, or gentisate (Fig. 25-7) before the ring is opened.[150] A single compound can be acted upon by more than one dioxygenase. Thus protocatechuate can be opened by a 3,4-dioxygenase or by a 4,5-dioxygenase (Fig. 18-22) leading to the branch point at protocatechuate in Fig. 25-7.

The initial hydroxylation of benzene, toluene, and other alkylbenzenes is accomplished by multicomponent aromatic ring dioxygenases that introduce two oxygen atoms to form diols.[158] Dioxygenation of benzoate yields a diol that can be **oxidatively decarboxylated** by reaction with NAD⁺ (Eq. 25-7) to form catechol.[157,162] Toluene gives 3-methylcatechol whose ring is, however, opened by an extradiol 2,3-dioxygenase, a so called *meta*-cleavage.[163,164] The product, 2-hydroxy-6-oxohepta-2,4-dienoate, is cleaved hydrolytically as indicated on the structure to give acetic acid and 2-hydroxypent-2,4-dienoate, which can be further metabolized. The hydrolytic cleavage is unusual[164] but is a β cleavage similar to the C–C bond cleavage by ribulose bisphosphate carboxylase (Eq. 13-48). Toluene can also be oxidized via benzoate through the β-oxoadipate pathway.



(25-7)

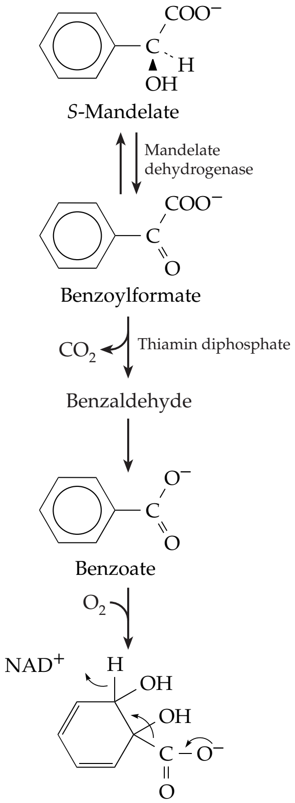


2-Hydroxy-6-oxohepta-2,4-dienoate

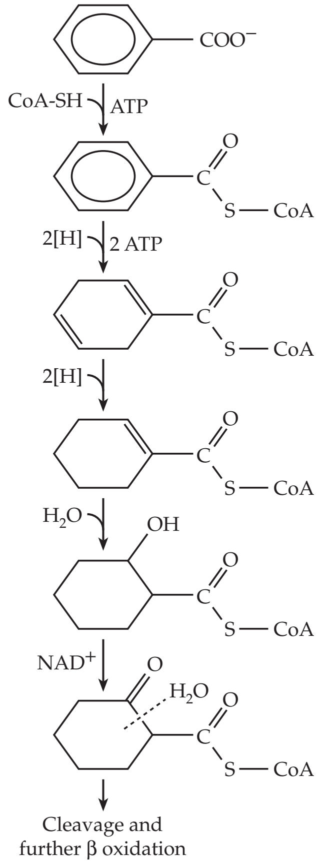
The plant acid *S*-mandelate must undergo conversion to *R*-mandelate by action of a racemase (Fig. 13-5) dehydrogenation, and side-chain cleavage as shown in Eq. 25-8 to form benzoate before it can be metabolized further.[165]

Although pseudomonads are well known for aerobic decomposition of aromatic compounds, some strains of *Pseudomonas*, as well as many other bacteria, are able to degrade aromatic compounds under completely anerobic conditions.[166,167] Benzoate can be converted to benzoyl-CoA and the ring can be partially reduced in two ATP- and NADH-dependent reactions (Eq. 25-9). The first of these reduction steps is unusual because ATP is apparently needed to drive the reaction.[166,166a,b] This is analogous to the need for ATP in nitrogen fixation (Eq. 24-6, step *b*).

Toluene, 3-chlorobenzoate, cinnamate, and 2-aminobenzoate can all be converted to benzoyl-CoA and be metabolized via the pathway of Eq. 25-9. Phenol, cresol, coumarate, protocatechuate, and vanillate can be converted to 4-hydroxybenzoyl-CoA and degraded in a similar fashion.[166] The breakdown of various forms of vitamin B_6 by bacteria is described in Section F (Eq. 25-24).



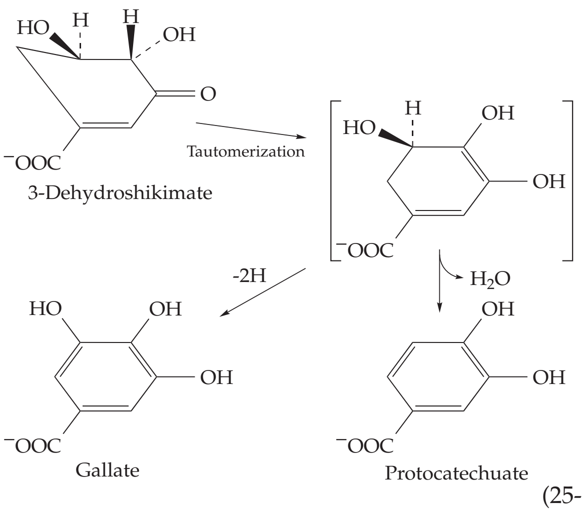
(25-8)



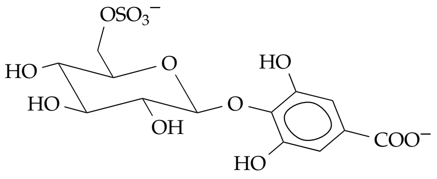
(25-9)

6. Quinic and Gallic Acids

Quinic acid, a compound accumulated by many green plants, can be formed by reduction of 3-dehydroquinate (Eq. 25-2) in both plants and bacteria. Quinic acid can be converted into useful industrial products such as benzoquinone and hydroquinone, and its production by bacteria provides a convenient route to these compounds.[168] In the main shikimate pathway 3-dehydroquinate is dehydrated to 3-dehydroshikimate (Eq. 25-3). The latter can be dehydrated further to **protocatechuate** (Eq. 25-10) either nonenzymatically[169] or by enzymatic action in bacteria.[170] Protocatechuate can be decarboxylated enzymatically to catechol, another compound of industrial value.[170] Nonenzymatic oxidation of 3-dehydroshikimate (Eq. 25-10) yields gallate.[169] Gallic acid derivatives are important plant constituents, but the biosynthetic origin has been obscure.[171] Gallate is probably formed from 3-dehydroshikimate as indicated in Eq. 25-10.[172] Esters and other derivatives of gallic acid constitute the **hydrolyzable tannins**. These materials accumulate in the vacuoles of the plants and are also deposited in the bark along with the **condensed tannins**, which are polymeric flavonoid compounds (Box 21-E).



(25-10)



Gallic acid 4-*O*-(β-D-glucopyranosyl-6'-sulfate),
the periodic leaf movement factor from *Mimosa*

7. The Metabolism of Phenylalanine and Tyrosine in Plants

Some of the pathways of animal and bacterial metabolism of aromatic amino acids also are used in plants. However, quantitatively more important are the reactions of the **phenylpropanoid pathway**,[173–174a] which is initiated by **phenylalanine ammonia-lyase** (Eq. 14-45).[175] As is shown at the top of Fig. 25-8, the initial product from phenylalanine is ***trans*-cinnamate**. After hydroxylation to 4-hydroxycinnamate (*p*-coumarate) and conversion to a coenzyme A ester,[175a] the resulting *p*-coumaryl-CoA is converted into mono-, di-, and trihydroxy derivatives including anthocyanins (Box 21-E) and other flavonoid compounds.[176] The dihydroxy and trihydroxy methylated products are the starting materials for formation of lignins and for a large series of other plant products, many of which impart characteristic fragrances. Some of these are illustrated in Fig. 25-8.

Benzoic and salicylic acids. Two of the simplest plant acids arising from *trans*-cinnamate are **benzoic acid**, accumulated in plums and cranberries, and **salicylic acid**, present in all green plants and accumulated as methyl esters or glycosides in some plants, e.g., those of the willow family. Salicylic acid is made by hydroxylation of benzoic acid,[177] which can be formed from *trans*-cinnamate by β oxidation as depicted in Fig. 25-8, but it may also arise from isochorismate as shown in Fig. 25-2.[178] Salicylic acid plays a central role in resistance of plants to a variety of

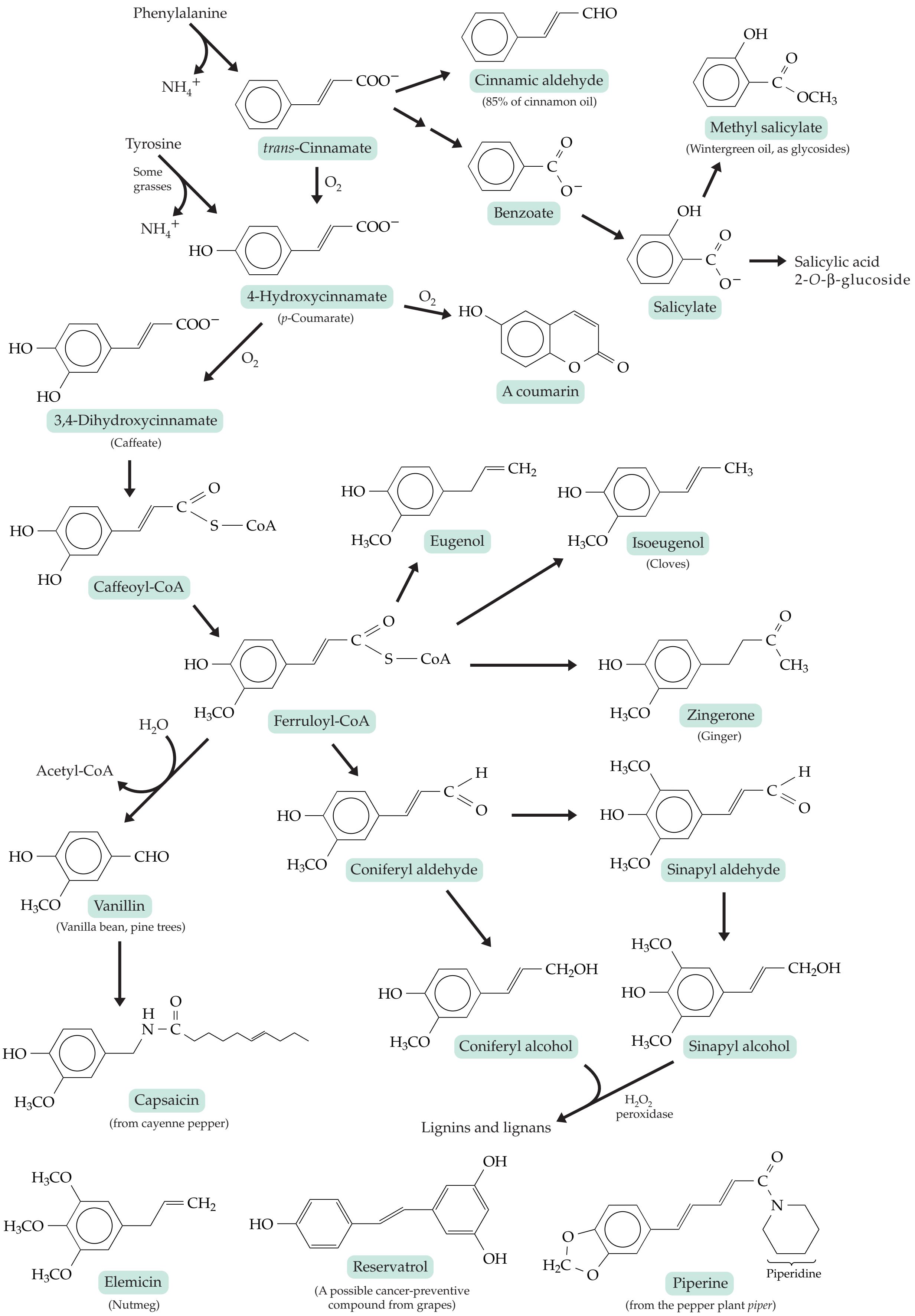
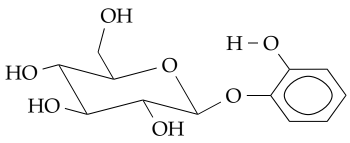

Phenylalanine
NH_4^+
COO^-
trans-Cinnamate
CHO
Cinnamic aldehyde
(85% of cinnamon oil)
OH
O
C
OCH_3
Methyl salicylate
(Wintergreen oil, as glycosides)
Tyrosine
Some grasses
NH_4^+
O_2
O
C
O^-
Benzoate
HO
COO^-
4-Hydroxycinnamate
(*p*-Coumarate)
OH
O
C
O^-
Salicylate
Salicylic acid 2-*O*-β-glucoside
O_2
HO
O
O
A coumarin
HO
COO^-
HO
O_2
3,4-Dihydroxycinnamate
(Caffeate)
HO
C
O
S — CoA
HO
Caffeoyl-CoA
HO
CH_2
H_3CO
Eugenol
HO
CH_3
H_3CO
Isoeugenol
(Cloves)
HO
C
O
S — CoA
H_3CO
Ferruloyl-CoA
HO
O
C
CH_3
H_3CO
Zingerone
(Ginger)
H_2O
Acetyl-CoA
HO
CHO
H_3CO
Vanillin
(Vanilla bean, pine trees)
H
C
O
HO
H_3CO
Coniferyl aldehyde
H_3CO
H
C
O
HO
H_3CO
Sinapyl aldehyde
HO
CH_2OH
H_3CO
Coniferyl alcohol
H_3CO
CH_2OH
HO
H_3CO
Sinapyl alcohol
H
O
N – C
HO
H_3CO
Capsaicin
(from cayenne pepper)
H_2O_2 peroxidase
Lignins and lignans
H_3CO
CH_2
H_3CO
H_3CO
Elemicin
(Nutmeg)
OH
HO
OH
Reservatrol
(A possible cancer-preventive compound from grapes)
O
C
N
O
H_2C
O
Piperidine
Piperine
(from the pepper plant *piper*)

Figure 25-8 Formation of some plant metabolites from phenylalanine and tyrosine via the phenylpropanoid pathway.

diseases, a phenomenon called **aquired systemic resistance**.[179–181] A large number of defense-related genes are induced by salicylate leading to increased synthesis of phytoalexins, proteins, and lignins. The mechanism may be to inhibit catalases allowing the level of H_2O_2 to rise. Hydrogen peroxide not only is a precursor to potent antimicrobial compounds as HOCl (Eqs. 16-12, 16-13) and a participant in lignin synthesis, but it may directly activate transcription of disease resistance genes.[182] This mode of action appears to parallel an **acute-phase response** of the vertebrate immune system through which H_2O_2 activates the transcription factor NF-kB (Fig. 5-40; Chapter 28).[182] It is possible that salicylate also has an effect on transcription in the human body.[183]

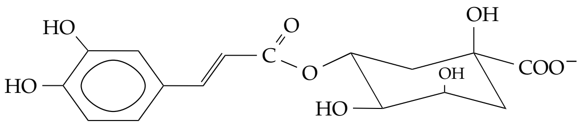
Salicylic acid derivatives, such as aspirin (acetylsalicylic acid), have long been known as pain relievers for the human body. The major effect is thought to be inhibition of a cyclooxygenase (Eq. 21-16). Willow bark has been known since the 18th century to contain a pain reliever, which was identified as salicylate esters and salicyl alcohol derivatives such as salicin.



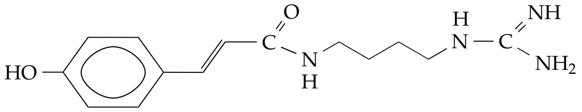
Salicin, a glucoside of salicyl alcohol

Vanilla and other plant products. One of the most widely used natural plant products is **vanilla extract**, which is obtained from cured, unripe fruit of the orchid *Vanilla planifolia*. The curing process releases **vanillin** and related compounds such as vanillic acid and 4-hydroxybenzaldehyde from glycosides. Because the flowers must be hand pollinated every day, natural vanilla extract is extremely expensive.[184,185] Most vanillin used in flavoring is obtained by hydrolysis of lignin, but production from glucose using bacterially produced enzyme reactors is possible.[185] Prince and Gunson, in an interesting article,[184] described the use of mass spectrometry and ^{13}C-enriched synthetic vanillin in the battles to distinguish natural vanilla extract from artificial mixtures of vanillin and other compounds and to camouflage the latter. Conversion of ferulic acid to vanillin in plants is apparently accomplished by β oxidation of acyl-CoA derivatives[186] as indicated in Fig. 25-8.

The ubiquitous plant compound **chlorogenic acid** (isolated from green coffee beans) is formed by transesterification with the glycoside cinnamoyl-glucose.[187] Coumaroyl-CoA is converted into monomeric and dimeric amides with **agmatine**, which provides barley plants with resistance to mildew.[188] Similar compounds with various polyamines and derived from *p*-coumaric, caffeic, ferulic, or sinapic acid appear to function in plant development. For example, caffeoylputrescine and caffeoyl-γ-aminobutyrate are accumulated specifically in sex organs of tobacco flowers.[189]

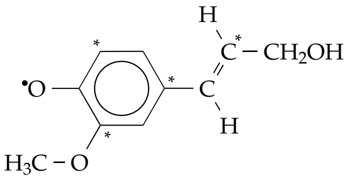


Chlorogenic acid
(3-*O*-caffeoylquinic acid)

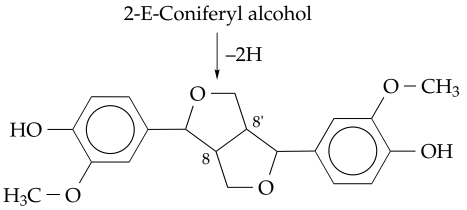


P-Coumaroylagmatine

Lignin, lignols, lignans, and phenolic coupling. Lignin is a complex material of relative molecular mass greater than 10,000. It is remarkably stable, being insoluble in hot 70% sulfuric acid. Lignin may be described as a "statistical polymer of oxyphenylpropane units." It arises from oxidative coupling of **coniferyl and sinapyl alcohols** (Figs. 25-8, 25-9) and related monomers known as **lignols**.[190–190c] The enzyme responsible for the polymerization may be a peroxidase, which catalyzes formation of lignin from the monomeric alcohols and H_2O_2. A radical generated by loss of an electron from a phenolate anion of coniferyl alcohol consists of a number of resonance forms in which the unpaired electron may be present not only on the oxygen but also at the positions marked by asterisks in the following structure:



Coupling of such radicals yields a great variety of products. One type of dimerization gives the stable ether linked **pinoresorcinol** (Eq. 25-11). Through a complex sequence of reactions, it can be converted into other plant compounds including the phytoalexin **plicatic acid**, a major component of western



(25-11)

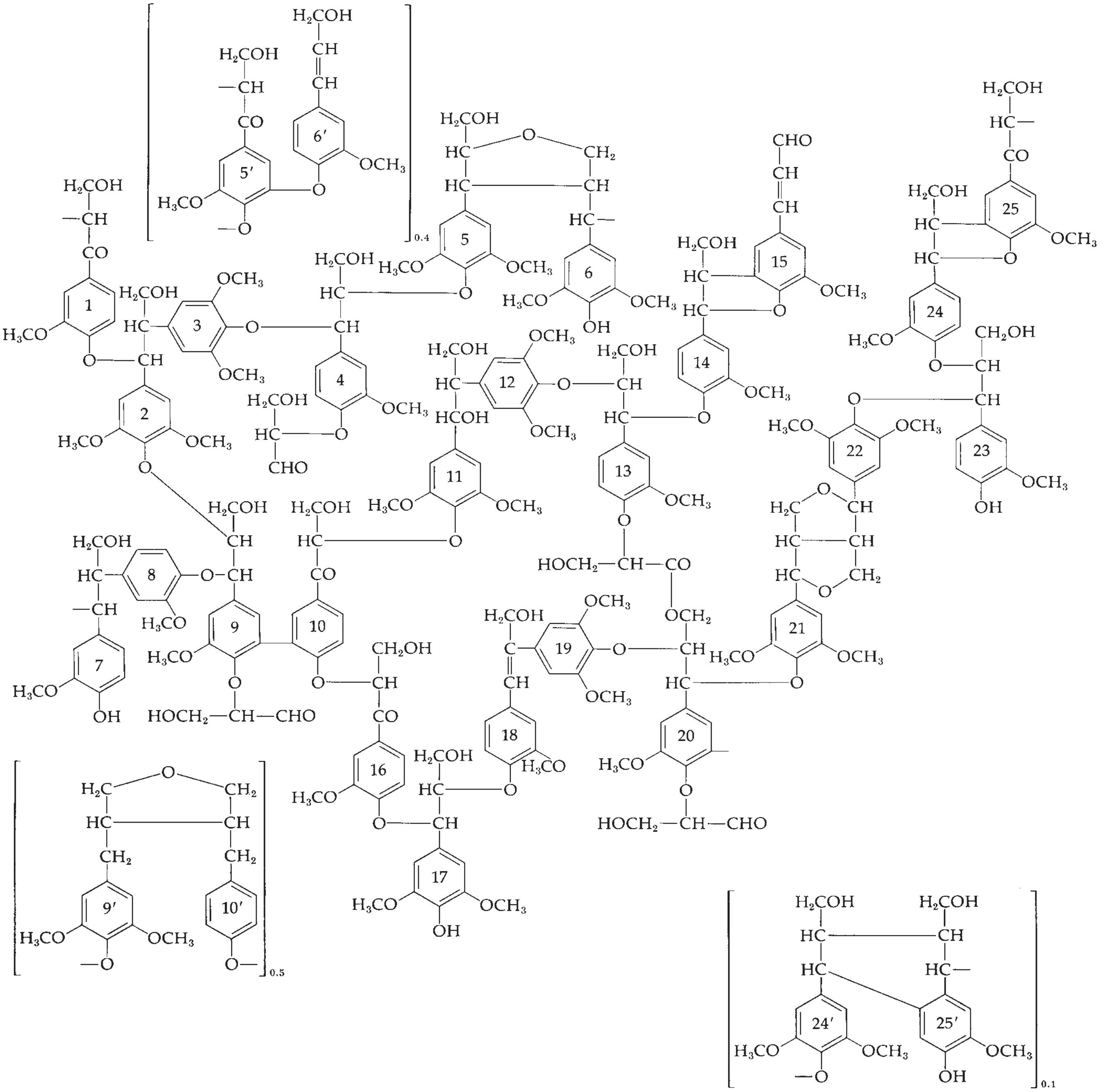
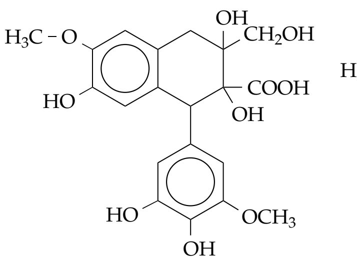
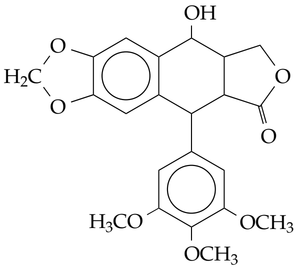


Figure 25-9 Proposed structure of beech lignin. There are 25 different C_9 units, of which several can, to some extent, be replaced by the three dimeric units in brackets. Redrawn from Nimz,[194] p. 317.

red cedar heartwood, and **podophyllotoxin**, found in the poisonous roots of the May apple (*Podophyllum peltatum*).[191] The little yellow "apple" is edible. Podophyllotoxin is used in cancer treatment. These dimeric 8,8'-carbon linked derivatives of lignols are called **lignans**[191a] while oligomers linked in other ways are **neolignans**.[192] The lignols are also incorporated covalently into **suberin**, a waxy layer of plant cell walls.[193]



Plicatic acid



Podophyllotoxin

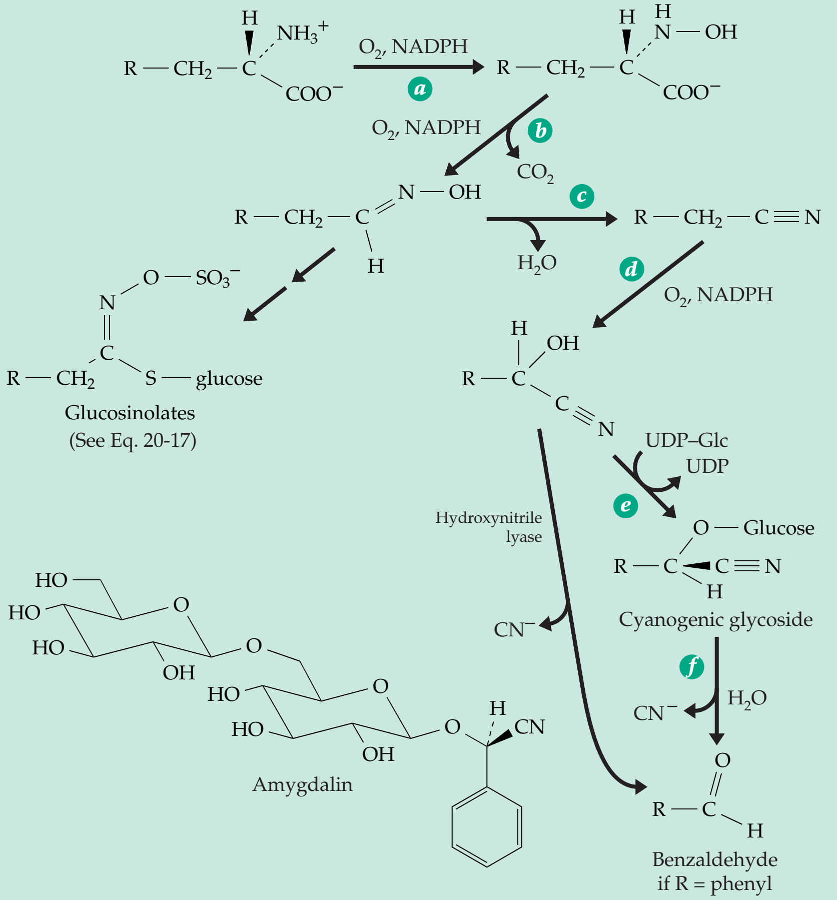
The lignols are synthesized within cells and are thought to move out into cell walls, possibly as phenolic glycosides.[190c,195] Cell wall peroxidases or laccases initiate polymerization.[173] The previously discussed lignans do not tend to polymerize, but other dimeric forms do. The dimers still contain hydroxyl groups capable of radical formation and addition to other units. At least ten types of intermonomer linkage other than that in the lignans are shown in Fig. 25-9. Lignin represents an enormous potentially valuable

BOX 25-B THE CYANOGENIC GLYCOSIDES

Cyanide-containing glycosides are synthesized by many higher plants including such crop plants as sorghum, cassava, and white clover.[a] The starting compounds are L-amino acids, most often phenylalanine, tyrosine, valine, or isoleucine.[b] The following sequence was proposed by Conn and others.[b–f] The conversion to an *N*-hydroxyamino acid in the first reaction (step *a*) is catalyzed by a cytochrome P450 hydroxylase system requiring NADPH and O_2. The same enzyme catalyzes a second hydroxylation that is followed by dehydration and decarboxylation (step *b*) to form an oxime.[f] The oxime is dehydrated to form a nitrile (step *c*) and a third hydroxylation (step *d*) produces an α-hydroxynitrile (cyanohydrin). Glycosylation by transfer from UDP-Glc or other sugar nucleotide (step *e*) forms the cyanogenic glycoside. If R = *p*-hydroxyphenyl in the foregoing equation and the sugar is glucose the product is **dhurrin**, present in **sorghum**. In **amygdalin**, present in bitter almonds and in pits of apricots, peaches, cherries, etc., two glucosyl units in β-1,6 linkage (gentiobiose) are attached to mandelonitrile.

Cyanogenic glycosides generate free cyanide by elimination when the glycosidic linkage is hydrolyzed.[a] This occurs with dhurrin at high pH and with others at pH=1, 70–100°C. Elimination of cyanide from the hydroxynitriles is catalyzed enzymatically.[g,h] Another cytochrome P450 dependent process utilizes oxidation to an oxime, as in the foregoing scheme, but converts the oxime to a **glucosinolate** in a two-step process.[i]

At one time amygdalin, sold as **Laetrile**, was promoted as a treatment for cancer, presumably based on the hope that the cancer cells would be poisoned by the released cyanide.[j] The tubers and leaves of the cassava plant provide a major source of food in many tropical countries. However, unless the cyanogenic glycosides are removed by boiling the tubers and pulping the leaves cassava is very toxic.[e,k]



---

[a] Vennesland, B., Castric, P. A., Conn, E. E., Solomonson, L. P., Volini, M., and Westley, J. (1982) *Fed. Proc.* **41**, 2639–

[b] Conn, E. E. (1979) *Nateuwissenschaften* **66**, 28–34

[c] Moller, B. L., and Conn, E. E. (1980) *J. Biol. Chem.* **255**, 3049–3056

[d] Moller, B. L., and Conn, E. E. (1979) *J. Biol. Chem.* **254**, 8575–8583

[e] Andersen, M. D., Busk, P. K., Svendsen, I., and Moller, B. L. (2000) *J. Biol. Chem.* **275**, 1966–1975

[f] Sibbesen, O., Koch, B., Halkier, B. A., and Moller, B. L. (1995) *J. Biol. Chem.* **270**, 3506–3511

[g] Lauble, H., Miehlich, B., Förster, S., Wajant, H., and Effenberger, F. (2001) *Protein Sci.* **10**, 1015–1022

[h] Lauble, H., Miehlich, B., Förster, S., Wajant, H., and Effenberger, F. (2002) *Biochemistry* **41**, 12043 – 12050

[i] Du, L., Lykkesfeld, J., Olsen, C. E., and Halkier, B. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12505–12509

[j] Newmark, J., Brady, R. O., Grimley, P. M., Gal, A. E., Waller, S. G., and Thistlethwaite, J. R. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6513–6516

[k] Ononogbu, I. C. (1980) *Trends Biochem. Sci.* **5**, X

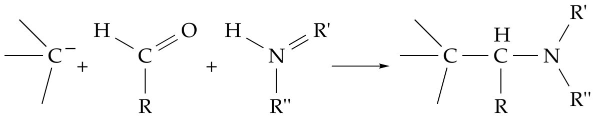
industrial source of aromatic raw materials, whose utilization has proved difficult. Oxidative degradation of lignin produces **humic acid**, an important organic constituent of soils.[196]

Oxidative coupling of radicals derived from phenols has a much wider role in nature than in lignin formation. Many alkaloids and other plant and fungal metabolites are synthesized using this reaction.[197] Tyrosine radicals are thought to be involved in formation of thyroxine (Eq. 25-6), melanin (Fig. 25-6), crosslinkages of ferulic acid with polysaccharides of plant cell walls,[197a] dityrosine, and other protein crosslinkages (Section B,4). Autocatalytic reaction of tyrosine-derived aminoquinone radicals in the Cu^{2+}-containing active site of amine oxidases apparently generates the mature prosthetic group topaquinone (Eq. 15-53).[198–198b] Oxidative coupling of tryptophanyl or cysteinyl side chains generates tryptophan tryptophylquinone (p. 817) or **cysteine tryptophylquinone** (CTA).[198b,c] As discussed on pp. 885–886 the tyrosine-cysteine thioether-bridged prosthetic groups of galactose oxidase (Fig. 16-29) and several other enzymes are also self-processing.[198d,e]

8. Alkaloids

More than 12,000 miscellaneous nitrogen-containing compounds, known as alkaloids, are produced by plants.[174a,199] Alkaloids are often thought of simply as end products of nitrogen metabolism in plants. However, most plants do not make alkaloids, whereas certain families of plants make many. There are probably ecological reasons.[200] Alkaloids often have potent physiological effects on animals, and many have been used as medicines from ancient times. Some have been prized through centuries as hallucinogens and intoxicants.

There are several classes of alkaloids. Among these are purines such as xanthine and caffeine, terpenes (Chapter 22), polyketides (Chapter 21), and alkaloids derived from amino acids. The basic amino acids ornithine, arginine, histidine, and lysine as well as the aromatic amino acids, anthranilate, and nicotinate are some of the starting materials.[199,201] Robinson[202,203] in 1917 recognized that many alkaloids are derived directly from aromatic amino acids. He proposed that alkaloids arise from **Mannich reactions** (Eq. 25-12) in which an amine and an aldehyde (probably through a Schiff base) react with a nucleophilic carbon such as that of an enolate anion. Many of the amines are formed by decarboxylation of amino acids, and the aldehydes may arise by oxidative decarboxylation (transamination and decarboxylation) of amino acids. Thus, amino acids can provide both of the major reactants for alkaloid synthesis. Furthermore, nucleophilic centers in the aromatic rings, e.g., in positions para to hydroxyl substituents, are frequent participants in the proposed Mannich condensations. While Robinson's ideas on alkaloid biosynthesis were initially speculative, they have been confirmed by isotopic labeling experiments and more recently by isolation of the enzymes involved. Nevertheless, many questions remain. The postulated aldehydes are not proved intermediates. The condensations with 2-oxo acids may occur prior to decarboxylation.



(25-12)

An example is shown in Fig. 25-10. Dopa is decarboxylated to dopamine and is oxidized to 3,4-dihydroxybenzaldehyde. A Mannich reaction (via the Schiff base as shown) leads to ring closure. Oxidation of the ring produces an **isoquinoline** ring, a structural characteristic of a large group of alkaloids. Methylation produces **papaverine**, found in the opium poppy. A related alkaloid **morphine** (Fig. 25-10), at first glance, appears dissimilar. However, the biosynthetic route is similar. The initial Schiff base is formed from tyramine. Closure of the third ring together with hydroxylation and methylation yield ***R*-reticuline**, a precursor to many alkaloids. Its two rings are then oxidatively coupled through a C–C bond and an ether linkage.[204,204a] *S*-Reticuline is the precursor to another large family of alkaloids.[205] The six-membered ring of another alkaloid, **colchicine** (Box 7-D), originates from phenylalanine, while the seven-membered tropolone ring is formed from tyrosine by ring expansion.

C. Metabolism of Tryptophan and Histidine

The biosynthesis of tryptophan is outlined in Fig. 25-2. This amino acid not only assumes great importance in the structure and functioning of proteins but is converted into hormones, both in animals and plants, and into alkaloids in some plants. Some of the pathways are indicated in Figs. 25-11 and 25-12.

1. The Catabolism of Tryptophan

The primary catabolic pathway for tryptophan in animal cells is initiated (step *a*, Fig. 25-11) by **tryptophan 2,3-dioxygenase** (tryptophan pyrrolase; Eq. 18-38).[206] The enzyme is induced both by glucocorticoids and by tryptophan.[207] The related **indolamine 2,3-dioxygenase** catalyzes the same reaction of L-tryptophan but also acts on D-tryptophan and other substrates. It has different tissue distribution and reg- ulatory properties[208] and may play a role in

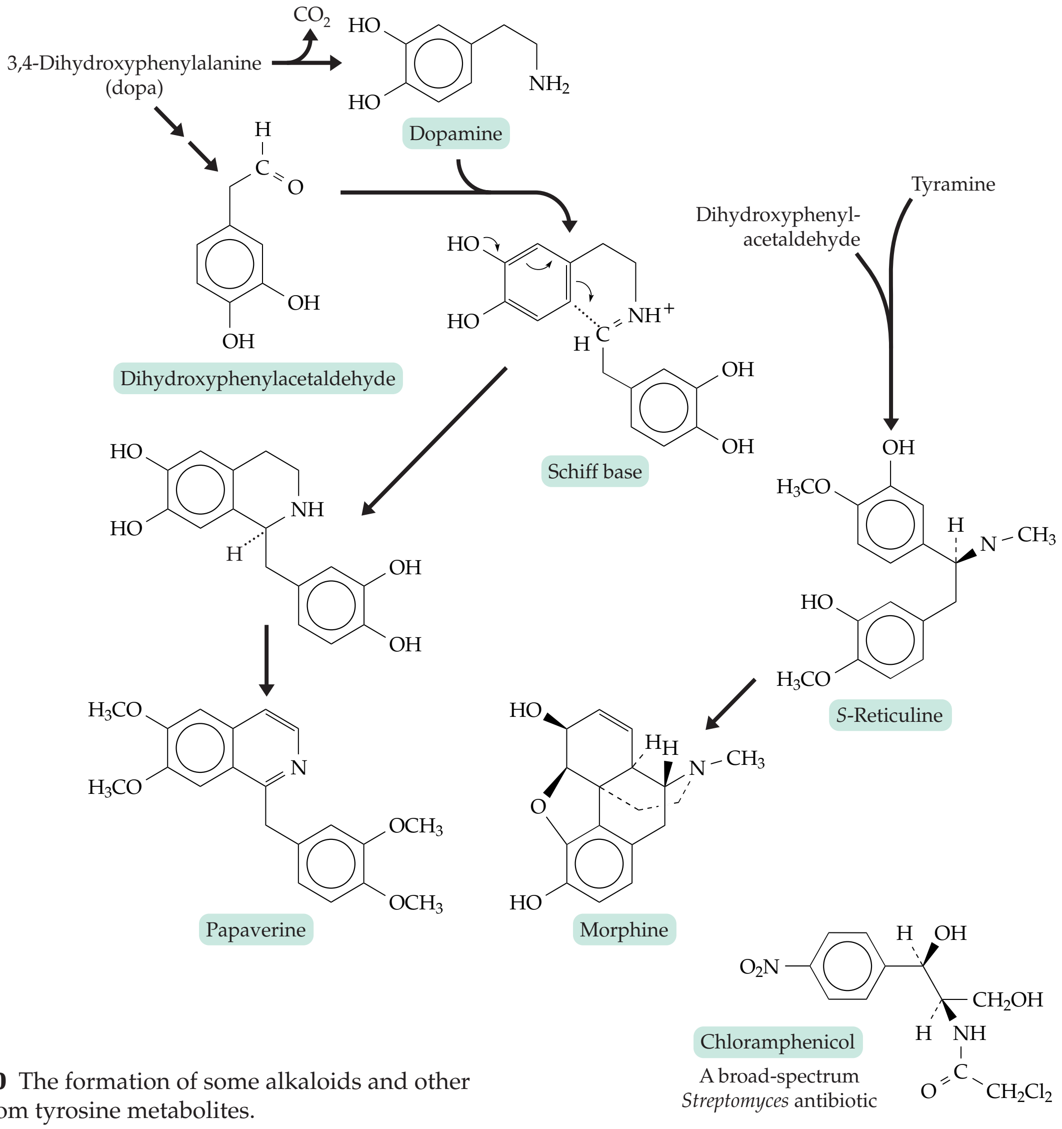

3,4-Dihydroxyphenylalanine (dopa)
CO_2
Dopamine
Dihydroxyphenylacetaldehyde
Schiff base
Tyramine
Dihydroxyphenyl-acetaldehyde
S-Reticuline
Papaverine
Morphine
Chloramphenicol
A broad-spectrum *Streptomyces* antibiotic

Figure 25-10 The formation of some alkaloids and other substances from tyrosine metabolites.

inflammatory responses.[206,209] An alternative pathway of tryptophan breakdown takes place in intestinal bacteria, which utilize tryptophan indolelyase (tryptophanase) to eliminate indole (step *b*, Fig. 25-11).[209a] The indole is hydroxylated to **indoxyl**, some of which is absorbed into the bloodstream and excreted in the urine as indoxyl sulfate.

Returning to the major tryptophan catabolic pathway, marked by green arrows in Fig. 25-11, formate is removed hydrolytically (step *c*) from the product of tryptophan dioxygenase action to form **kynurenine**, a compound that is acted upon by a number of enzymes. Kynureninase (Eq. 14-35) cleaves the compound to anthranilate and alanine (step *d*), while transamination leads to the cyclic **kynurenic acid** (step *e*). The latter is dehydroxylated in an unusual reaction to **quinaldic acid**, a prominent urinary excretion product.

Another major pathway of kynurenine metabolism (step *f*, Fig. 25-11) is hydroxylation to **3-hydroxykynurenine**, which in turn can undergo transamination to the cyclic **xanthurenic acid**. Xanthurenic acid is excreted from the human body, but in the malaria mosquito *Anopheles gambia* it acts as a mating factor for the malaria parasite *Plasmodium*.[210,211] In many insects, including *Anopheles*, 3-hydroxykynurenine is a precursor of insect eye pigments or "**omnochromes**."[210,212–213a] 3-Hydroxykynurenine also has neurotoxic properties.[213a] (See p. 1796.)

Cleavage of 3-hydroxykynurenine by kynureninase (step *g*, Fig. 25-11) forms 3-hydroxyanthranilate, which is opened under the action of another dioxygenase (step *h*) with eventual degradation to acetyl-CoA, as indicated. In insects the reactive 3-hydroxyanthranilate is utilized in "tanning" reactions, e.g., coupling to tyrosine residues to toughen insect cuticles and walls of cocoons.[214]

Tryptophan is hydroxylated to 5-hydroxytryptophan[213b] which is decarboxylated to **serotonin** (5-hydroxytryptamine), an important neurotransmitter

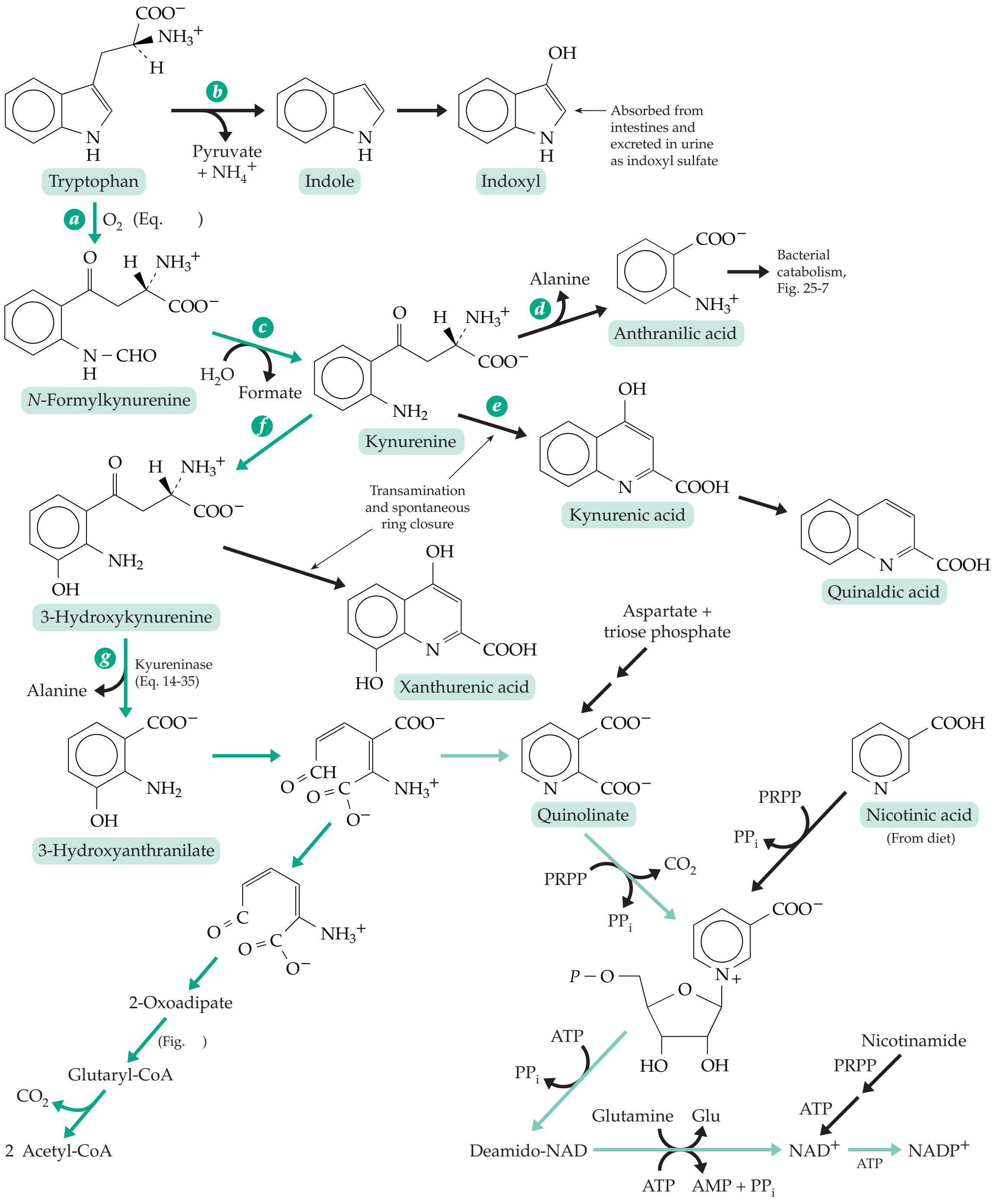
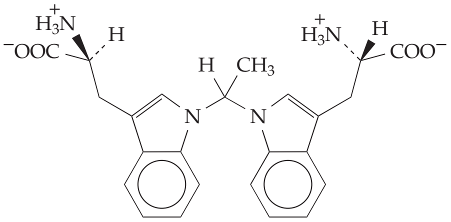

Tryptophan
b
Pyruvate + NH_4^+
Indole
Indoxyl
Absorbed from intestines and excreted in urine as indoxyl sulfate
a
O_2 (Eq.)
N-Formylkynurenine
c
H_2O
Formate
Kynurenine
d
Alanine
Anthranilic acid
Bacterial catabolism, Fig. 25-7
e
Kynurenic acid
Quinaldic acid
f
3-Hydroxykynurenine
Transamination and spontaneous ring closure
Xanthurenic acid
g
Kyureninase (Eq. 14-35)
Alanine
3-Hydroxyanthranilate
Aspartate + triose phosphate
Quinolinate
Nicotinic acid
(From diet)
PRPP
PP_i
PRPP
CO_2
PP_i
2-Oxoadipate
(Fig.)
Glutaryl-CoA
CO_2
2 Acetyl-CoA
ATP
PP_i
Nicotinamide
PRPP
ATP
Glutamine
Glu
Deamido-NAD
NAD^+
ATP
$NADP^+$
ATP
AMP + PP_i

Figure 25-11 Some catabolic reactions of tryptophan and synthetic reactions leading to NAD and NADP.

substance[215,215a] and a regulatory component of plants and animals alike.[216] In the pineal gland serotonin is methylated and acetylated to **melatonin**, the pineal hormone[217–221] (Fig. 25-12).

The following dangerous tryptophan derivative was evidently formed in a fermentation used to produce tryptophan sold as a food supplement in 1990. More than 1,500 persons became ill and 27 died, perhaps as a direct result of toxicity of this compound.[222,223]



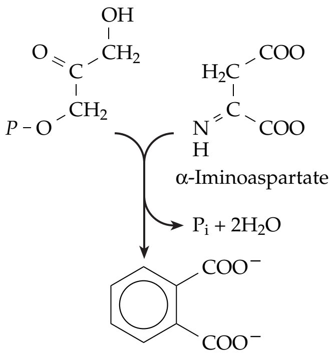
1,1'-Ethylidenebis[tryptophan]

Formation of NAD$^+$ and NADP$^+$. An alternative pathway, marked by shaded green arrows in Fig. 25-11, allows animals to form the nicotinamide ring of NAD$^+$ and NADP$^+$ from tryptophan.[224] The aldehyde produced by the ring opening reaction of step *h* can reclose (step *i*) to a pyridine ring in the form of **quinolinic acid**.[225] The latter, in a reaction that is also accompanied by decarboxylation, is coupled with a phosphoribosyl group of PRPP to form **nicotinate mononucleotide**.[225a] Adenylylation produces deamido NAD, which is converted to **NAD** by a glutamine- and ATP-dependent amination of the carboxyl group.[226]

As indicated in Fig. 25-11, free nicotinic acid can also be used to form NAD. Not surprisingly, nicotinic acid, an essential vitamin, is about 60 times more efficient than tryptophan as a source of NAD. Nevertheless, a high-tryptophan diet partially overcomes a deficiency in dietary intake of nicotinic acid. The effectiveness of a diet containing only maize as a source of protein in inducing the deficiency disease pellagra (Box 15-A) is in part a result of the low tryptophan content of maize protein. Nicotinic acid is rapidly converted in the liver to an amide with glycine, **nicotinuric acid**. Nicotinurate can be oxidatively cleaved by peptidylglycine monoxygenase to nicotinamide[227] in another alternative synthetic route to NAD.

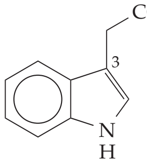
An alternative pathway for synthesis of quinolinate from aspartate and a triose phosphate exists in bacteria and in plants and provides the major route of nicotinic acid synthesis in nature. In *E. coli* the reaction is catalyzed by two enzymes, one an FAD-containing L-aspartate oxidase which oxidizes aspartate to α-iminoaspartate.[228] The latter condenses with dihydroxyacetone-*P* to form quinolinate (Eq. 25-13).[229] There are at least two other pathways for synthesis of quinolinic acid as well as five or more salvage pathways for resynthesis of degraded pyridine nucleotide coenzymes.[224,230,231]

Although quinolinic acid provides an important source of nicotinamide coenzymes, in excess it is a neurotoxic **excitotoxin** (Chapter 30) that has been associated with epilepsy and with inflammatory neuropathalogical conditions resulting from encephalitis.[213,232,233]

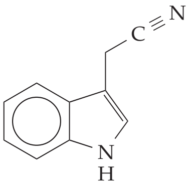


Quinolinate (25-13)

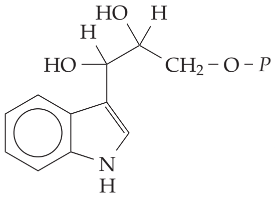
Auxin. The important plant hormone **indole-3-acetic acid** (IAA; often called by the more general name **auxin**) is partially derived by oxidative decarboxylation (Fig. 25-12, top) catalyzed by tryptophan-2-monooxygenase,[234–234b] a flavoprotein similar to lysine monooxygenase (Eq. 18-41). The reduction product indole-3-ethanol also occurs in plants and is metabolically active.[235] However, most IAA in plants is not formed from tryptophan but from some precursor,[236,236a] perhaps indole-3-glycerol phosphate, which immediately precedes tryptophan in its biosynthesis (Fig. 25-2). Routes of synthesis from indole-3-acetonitrile, indoleacetaldoxime,[236b] indole-3-glucosinolate (glucobrassicin), indole-3-pyruvate, and tryptamine have also been reported.[237–238a] Nitrilases are found in plants such as *Arabidopsis thaliana*[238] and also in symbiotic bacteria such as *Rhizobium*. Together with amidases they convert the acetonitrile, acetaldoximes, or acetamide derivatives to IAA,[237] which is transported throughout the plant.[238b]



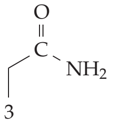
Indole-3-acetic acid



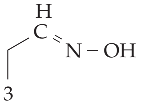
Indole-3-acetonitrile



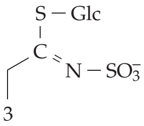
Indole-3-glycerol phosphate



Indole-3-acetamide



Indole-3-aldoxime



Indole-3-glucosinolate
(glucobrassicin)

About 95% of the IAA within plants is stored as conjugated forms that include amides with various amine acids and peptides[239] and glycosyl derivatives.[240] The gall-forming *Pseudomonas savastonoi* forms both IAA and conjugates such as N^{ε}-(indole-3-acetyl)-L-lysine, which aid these bacteria in colonizing olive and oleander plants.[241]

As a hormone IAA has a broad range of effects on plants, altering tissue differentiation, root growth, cell elongation, and cell division.[241a] The fastest observed response is an effect on cell elongation, which can be observed within 15–20 minutes.[242] In *A. thaliana* IAA causes very rapid transcription of at least five genes, one of which encodes 1-aminocyclopropanecarboxylase (ACC) synthase (Eq. 14-27).[243]

Alkaloids from tryptophan. The alkaloid **harmine**, which is found in several families of plants, can be formed from tryptophan and acetaldehyde (or pyruvate) in the same manner as is indicated for the formation of **papaverine** in Fig. 25-10. Some other characteristic plant metabolites such as **psilocybine**, an hallucinogenic material from the mushroom *Psilocybe aztecorum*, are formed directly from serotonin (Fig. 25-12). For many years **gramine** from barley was regarded as a curiosity because only one carbon atom separates the nitrogen atom from the indole ring. It is now believed that tryptophan is cleaved in a PLP-dependent reaction analogous to that of serine transhydroxymethylase (Eq. 14-30; Fig. 25-12). Other alkaloids arise in a more conventional fashion. Condensation of an isopentenyl group on the indole ring of tryptamine (Fig. 25-12) initiates the formation of **lysergic acid** and other ergot alkaloids.[244] The indole ring

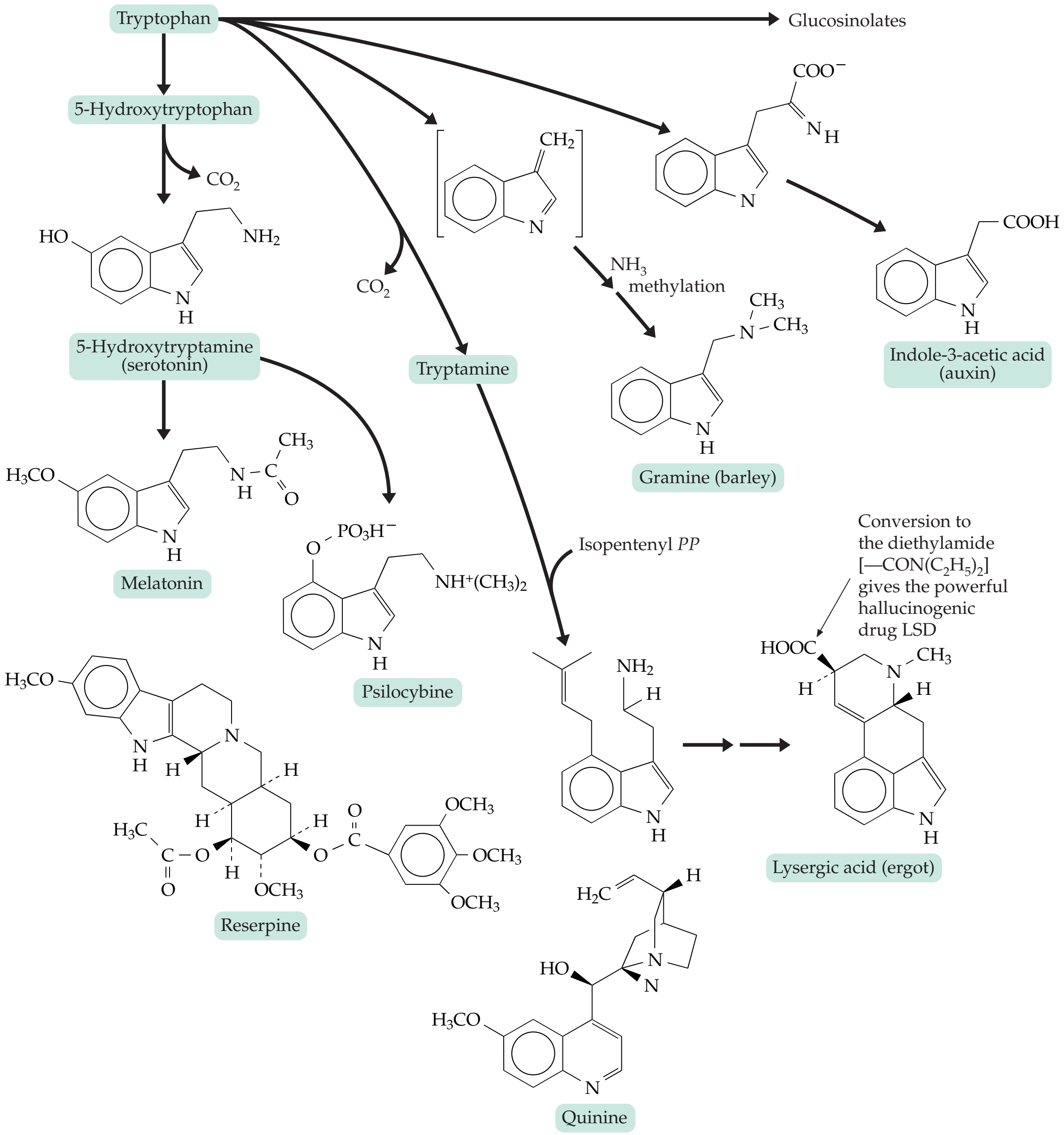
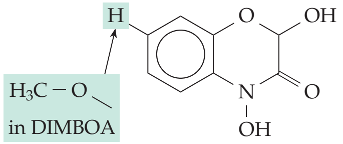

Tryptophan
Glucosinolates
5-Hydroxytryptophan
CO_2
HO
NH_2
N
H
5-Hydroxytryptamine (serotonin)
CO_2
Tryptamine
CH_2
N
NH_3
methylation
CH_3
N
CH_3
Gramine (barley)
COO^-
N
H
COOH
Indole-3-acetic acid (auxin)
H_3CO
CH_3
N
C
O
Melatonin
O
PO_3H^-
$NH^+(CH_3)_2$
Psilocybine
Isopentenyl *PP*
Conversion to the diethylamide $[—CON(C_2H_5)_2]$ gives the powerful hallucinogenic drug LSD
HOOC
CH_3
Lysergic acid (ergot)
H_3C
O
OCH_3
Reserpine
H_2C
HO
Quinine

Figure 25-12 Structures and some biosynthetic pathways for some hormones, indole alkaloids, and other metabolites of tryptophan.

of tryptophan is clearly visible in the structure of **reserpine** (Fig. 25-12). This compound from *Rauwolfia* is of medical interest because of its effect in lowering blood pressure and in depleting nervous tissues of serotonin, dopamine, and noradrenaline. Reserpine also contains a benzene ring, which is derived from tryptophan by a ring expansion. The periwinkle alkaloids, including the antitumor drug **vincristine** (see Box 7-D), are formed by condensation of tryptamine with the complex glycosidic aldehyde secologamin. Additional reactions form vincristine and more than 100 other indole alkaloids.[245,245a]

Another group of plant metabolites derived from tryptophan are cyclic hydroxamic acids whose names



DIBOA

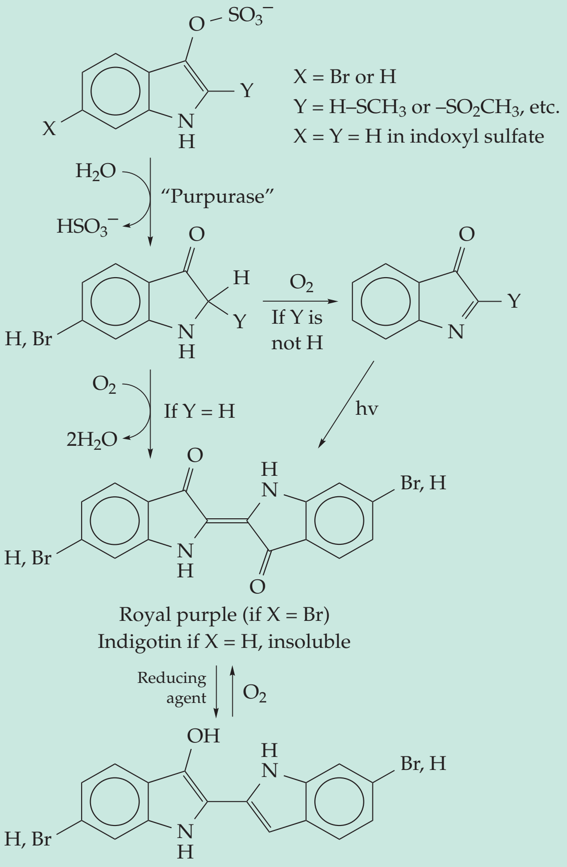
BOX 25-C ROYAL PURPLE AND BLUE DENIM

Ancient people, including the Phoenicians, Chinese, and Peruvians, discovered a dyeing process that utilized secretions of certain marine molluscs, animals that were also eaten as food.[a] In processes that were perhaps closely guarded secrets, the molluscan secretions were heated for days in vats with water, salt, and additional additives including human urine, honey, etc. When the mixture was right, wool was dipped and allowed to air dry in sunlight to give the famous royal purple colors. In this ancient process as many as 10,000 molluscs were used to produce one gram of the dye.[a] In other parts of the world blue dyes were generated by fermentation of plants of the genus *Indigofera* and also of the European woad plant.[b,c]

By 1909 the chemical nature of the royal purple pigment dibromoindogotin (see scheme) had been established and by 1897 synthetic indigo production had already begun. Ancient indigo dyeing utilized the precursors, sulfate esters of **indoxyl** or of substituted indoxyl, metabolites of tryptophan (see Fig. 25-11). The sulfate esters were hydrolyzed by the sulfatase "purpurase" to give the tautomer of indoxyl that is shown in the accompanying scheme. Atmospheric oxygen converts these compounds to the corresponding oxidized dyes, the **indigotins**. However, they are very insoluble and unsuitable for dyeing. In modern indigo dyeing to form such fabrics as blue denim, the indigotin is reduced with sodium dithionite, about 2 kg of the latter being used to reduce 1 kg of the dye to the reduced leuco form.[c] Either wool or cotton can by dyed with this reduced form, air oxidation returning the dye to the blue oxidized form.

Ancient dyers also had to maintain the dye in the reduced form. In fermentations of the woad plant a species of thermophilic *Clostridium* apparently supplied the reducing agent. Padden *et al.* suggested that such bacterial reduction might be used today to avoid pollution by the by-products of dithionite reduction.[c] Use of engineered bacteria to form high yields of indole and indoxyl as a source of indigoid precursors has also been suggested.[b]

One more complexity needs to be considered. Some of the precursors are adducts ($Y = -SCH_3$, $-SO_2CH_3$ in the structures) and cannot be oxidized directly to the indigotins. Use of sunlight in a photochemical process was required in these cases.[a]



Reduced, soluble, and nearly colorless leuco form.

---

[a] McGovern, P. E., and Michel, R. H. (1990) *Acc. Chem. Res.* **23**, 152

[b] Ensley, B. D., Ratzkin, B. J., Osslund, T. D., Simon, M. J., Wackett, L. P., and Gibson, D. T. (1983) *Science* **222**, 167–169

[c] Padden, A. N., Dillon, V. M., John, P., Edmonds, J., Collins, M. D., and Alvarez, N. (1998) *Nature (London)* **396**, 225

are often abbreviated to DIBOA and DIMBOA. They are part of the defense system of grasses against insects and fungi. DIBOA is the main hydroxamic acid in rye while DIMBOA predominates in wheat and maize. The compounds arise from indole generated from indole-3-glycerol-*P* followed by action of four cytochrome P450 enzymes.[246]

2. The Metabolism of Histidine

The biosynthesis of histidine, which might be regarded as the "super catalyst" of enzyme active centers, begins with a remarkable reaction of ATP, the "super coenzyme" of cells. The reaction is a displacement by N-1 of the adenine ring on C-1 of PRPP (step *a*, Fig. 25-13). The resulting product undergoes a ring opening reaction, step *b*, followed by an Amadori rearrangement (step *c*). The rearrangement product is cleaved via reaction with ammonia released from glutamine with formation of **5-aminoimidazole-4-carboxamide**, whose ribotide is an established intermediate in the synthesis of ATP and other purines. Here it is recycled via ATP (Fig. 25-13). The other product of the cleavage contains the five carbons of the original ribosyl group of PRPP, together with one nitrogen and one carbon split out from the ATP molecule and the nitrogen donated by glutamine.[246a] Ring closure (step *e*) forms the imidazole group, which is attached to a glycerol phosphate molecule. The glycerol-*P* end of the molecule undergoes dehydration[247] and ketonization of the resulting enol to a product, which can be transaminated[247a,b] and dephosphorylated to histidinol. Dehydrogenation of this alcohol forms histidine.[248–249a]

Regulation of histidine synthesis. In all, ten different genes code for the enzymes of histidine biosynthesis in *Salmonella typhimurium*. They are clustered as the **histidine operon**, a consecutive series of genes which are transcribed into messenger RNA as a unit.[250,251] The gene symbols *HisA*, *HisB*, etc., are indicated in Fig. 25-13, and their positions on the *E. coli* gene map are indicated in Fig. 26-4. The gene *HisB* codes for a complex protein with two different enzymatic activities as shown in Fig. 25-13.

The presence of an excess of histidine in a bacterial cell brings about repression of synthesis of all of the

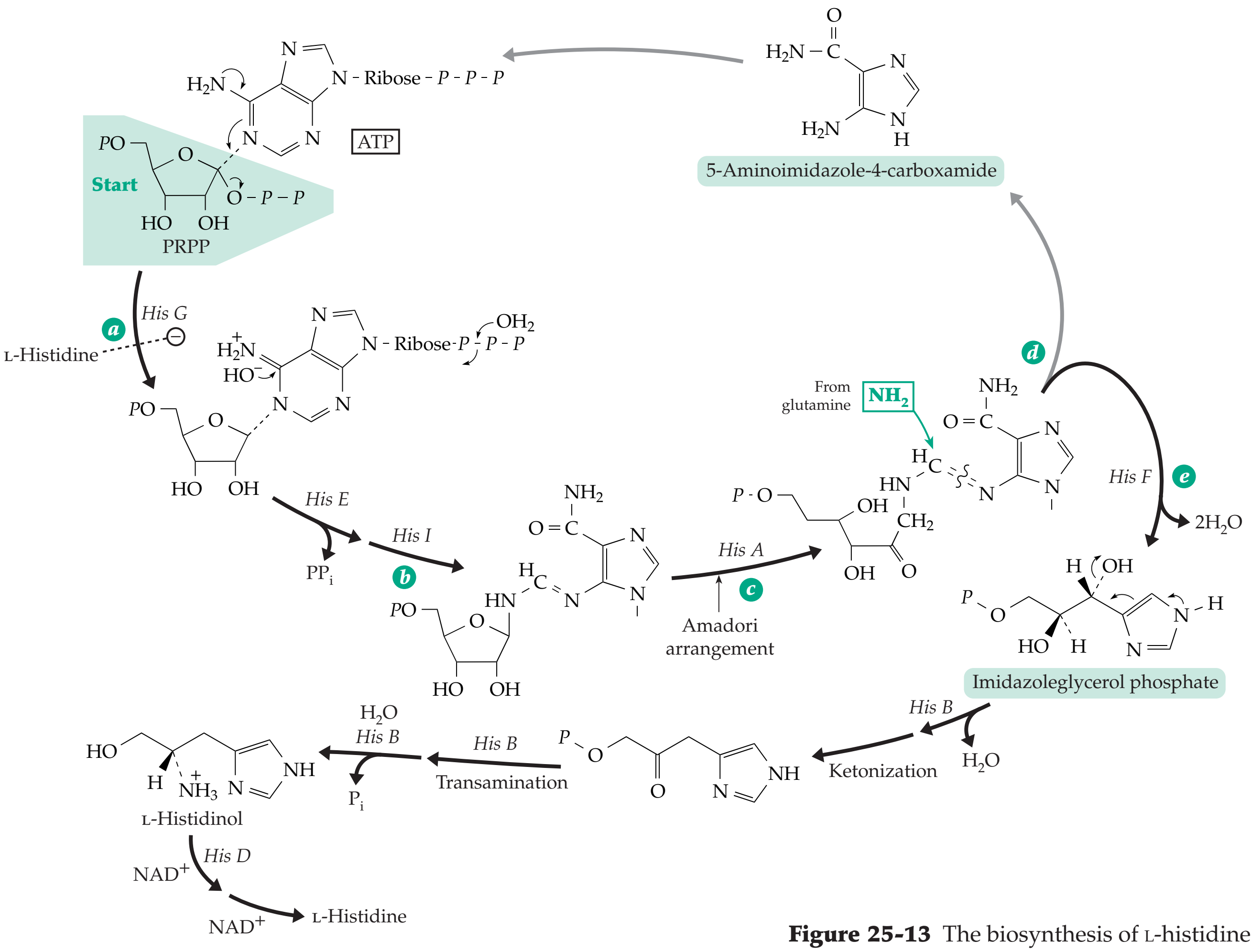
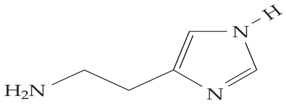

N - Ribose - P - P - P
ATP
PO
Start
O - P - P
HO OH
PRPP
5-Aminoimidazole-4-carboxamide
His G
a
L-Histidine
OH₂
N - Ribose - P - P - P
From glutamine
NH₂
d
His F
e
2H₂O
His E
PPᵢ
His I
b
His A
c
Amadori arrangement
Imidazoleglycerol phosphate
His B
H₂O
Ketonization
His B
Transamination
H₂O
His B
Pᵢ
L-Histidinol
His D
NAD⁺
NAD⁺
L-Histidine

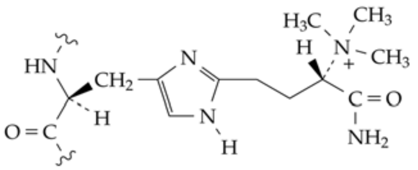
Figure 25-13 The biosynthesis of L-histidine.

enzymes encoded in the histidine operon. Details of the functioning of this and other operons are considered in Chapter 28. Histidine is also an allosteric inhibitor for the first enzyme of the biosynthetic sequence, i.e., step *a* of Fig. 25-13. Thus, instantaneous inhibition of the biosynthesis occurs if an excess of histidine accumulates. Similar patterns of both repression and feedback inhibition exist for many of the pathways of amino acid biosynthesis (Chapter 28).

Catabolism of histidine. The first steps of the major degradative pathway for histidine metabolism have already been discussed. Elimination of ammonia, followed by hydration and ring cleavage to **formiminoglutamate**, involves unusual reactions (Eq. 25-14)[252] which have been discussed earlier. Transfer of the formimino group to tetrahydrofolic acid and its further metabolism have also been considered (Chapter 15).

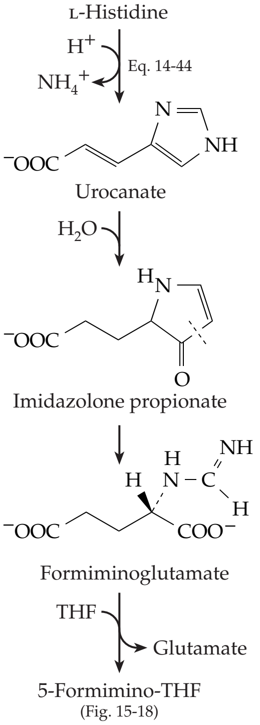


Histamine



Dipthamide residue within a peptide chain.

Other products from histidine include the hormonal substance **histamine** formed by decarboxylation, the oxidation product, imidazole acetic acid, and N^{δ}- and N^{ε}-methylhistidines. Histamine plays a role in release of gastric secretions and allergic responses (Chapter 5). Drugs (antihistamines) that inhibit its release are in widespread use. The unusual amino acid **diphthamide** has an unknown function in protein synthesis, occurring within the peptide chain of eukaryotic elongation factor 2 (Chapter 29).[253] Its biosynthesis from a specific histidine in EF-2 of all eukaryotes and archaebacteria requires four molecules of *S*-adenosylmethionine. The first transfers the four-carbon backbone of AdoMet to $C^{\varepsilon 1}$ of the histidyl group, a nucleophilic displacement resembling that of AdoMet-dependent C-methylation (Eq. 12-4). This is followed by transfer of three methyl groups, each from AdoMet, and finally an ATP-dependent amidation of the carboxyl group.[253] Diphthamide is the target for attack by diphtheria toxin (Box 29-A).



(25-14)

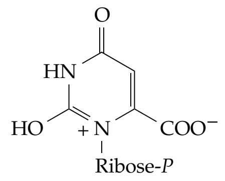
D. Biosynthesis and Catabolism of Pyrimidines

L-Aspartate contributes four of the six ring atoms of pyrimidines including the nitrogen. The α-carboxylate group is eventually lost as CO_2, the decarboxylation helping to drive the synthetic sequence. Six enzymatic steps are required to form the product uridine 5'-phosphate (UMP) as shown in Fig. 25-14, steps *a*–*f*. UMP is then converted on to the cytidine, uridine, and thymidine nucleotides as shown. This pathway of pyrimidine synthesis has been conserved throughout evolution and is used by all but a few specialized organisms.[254–258] The first step is synthesis of carbamoyl phosphate by the glutamine-dependent carbamoyl phosphate synthetase, an allosteric enzyme discussed in Chapter 24 (Eq. 24-22).[259–260a] The next step is transfer of the carbamoyl group to aspartate (Fig. 25-14, step *b*). The product is able to cyclize immediately (step *c*) by elimination of water to form **dihydroorotate**.[260a] The highly controlled aspartate carbamoyltransferase has been discussed in Chapters 7 and 11. Although this is a monofunctional enzyme in bacteria, it is fused with two other proteins in mammalian cells. The resulting multifunctional enzyme (with 240-kDa subunits) catalyzes three consecutive steps: the synthesis of carbamoyl phosphate, the carbamoyltransferase reaction, and the Zn^{2+}-dependent ring-closing reaction that forms dihydroorotate.[261,262] This cyclic product is oxidized to **orotate** (Fig. 25-14, step *d*) by a flavoprotein oxidase, which in mammals utilizes ubiquinones as electron acceptors.[263–265a] A displacement reaction with PRPP (Fig. 25-14, step *e*; see also Fig. 25-13, step *a*) converts it into **orotidine 5'-phosphate**[266,267] with release of PP_i.

1. Synthesis of Pyrimidine Nucleotides

Orotidine 5'-phosphate undergoes an unusual decarboxylation (Fig. 25-14, step *f*), which apparently is not assisted by any coenzyme or metal ion but is enhanced over the spontaneous decarboxylation rate 10^{17}-fold. No covalent bond formation with the enzyme has been detected.[268] It has been suggested that the enzyme stabilizes a dipolar ionic tautomer of the substrate. Decarboxylation to form an intermediate ylid would be assisted by the adjacent positive charge.[269,270] Alternatively, a concerted mechanism may be assisted by a nearby lysine side chain.[270a–d] Hereditary absence of this decarboxylase is one cause of orotic aciduria. Treatment with uridine is of some value.[271]

We see that **uridine 5'-phosphate** (UMP) is formed from aspartate in a relatively direct and simple way. Phosphorylation with ATP in two steps produces UDP and UTP. The **cytosine nucleotides** are formed from UTP, the initial step being amination to CTP (step *h*, Fig. 25-14). This reaction resembles in many respects the conversion of citrulline to arginine, which depends upon ATP and involves transfer of the nitrogen of aspartate (Eq. 24-23). However, in the formation of CTP glutamine serves as the nitrogen donor (NH_4^+ can substitute). Observation of positional


O
HN
HO
+ N
COO⁻
Ribose-P

Orotidine 5′-phosphate
(dipolar ionic tautomer)

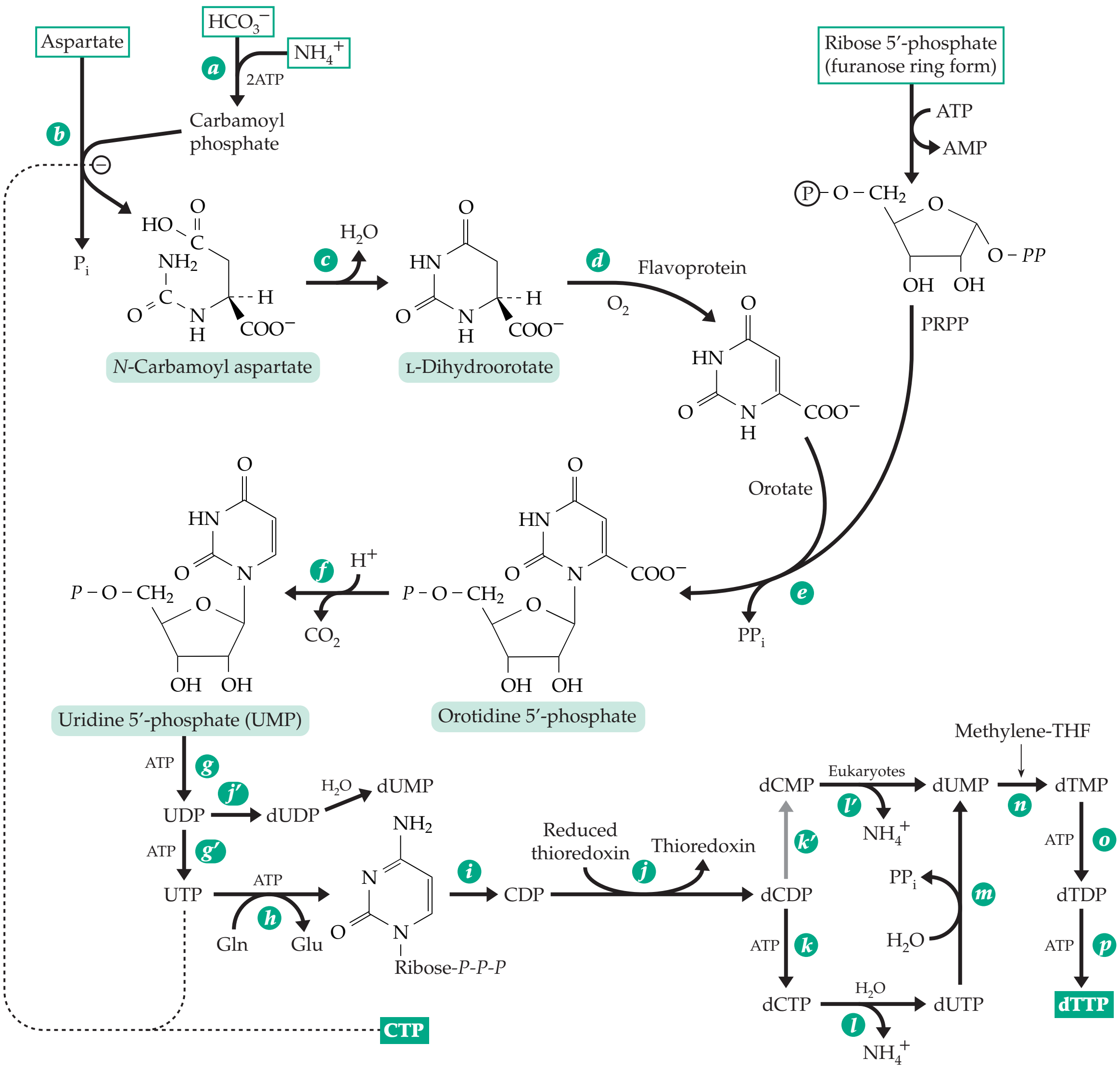
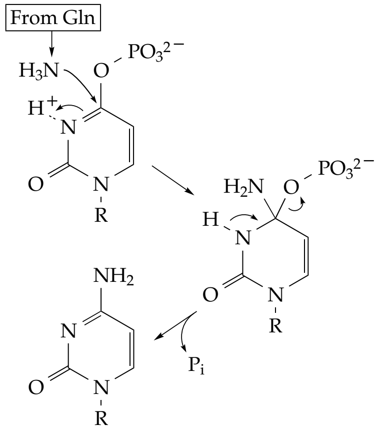

Aspartate
HCO3⁻
NH4⁺
a
2ATP
Carbamoyl phosphate
b
Pi
HO
O
C
NH2
C
O
N
H
H
COO⁻
N-Carbamoyl aspartate
c
H2O
HN
O
N
H
H
COO⁻
L-Dihydroorotate
d
Flavoprotein
O2
HN
O
N
H
COO⁻
Orotate
Ribose 5′-phosphate (furanose ring form)
ATP
AMP
P – O – CH2
O
O – PP
OH OH
PRPP
e
PPi
HN
O
N
COO⁻
P – O – CH2
O
OH OH
Orotidine 5′-phosphate
f
H⁺
CO2
HN
O
N
P – O – CH2
O
OH OH
Uridine 5′-phosphate (UMP)
ATP
g
UDP
j′
dUDP
H2O
dUMP
ATP
g′
UTP
ATP
h
Gln
Glu
NH2
N
O
N
Ribose-P-P-P
i
CDP
Reduced thioredoxin
Thioredoxin
j
dCDP
k′
dCMP
Eukaryotes
l′
NH4⁺
dUMP
Methylene-THF
n
dTMP
ATP
o
dTDP
ATP
p
dTTP
ATP
k
dCTP
H2O
l
NH4⁺
dUTP
PPi
m
H2O
CTP

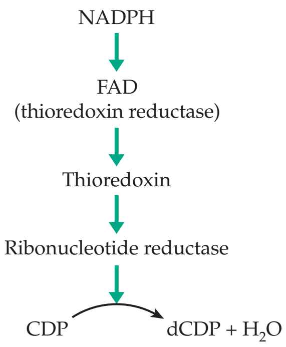
Figure 25-14 Assembly of the pyrimidine ring and biosynthesis of the pyrimidine nucleotide precursors of RNA and DNA.

isotope exchange with [γ-$^{18}O_4$]ATP suggested the occurrence of an enolic phosphate intermediate (Eq. 25-15)[272] as had also been suggested for adenylosuccinate synthase (Fig. 25-15, step *k*). CTP is incorporated directly into RNA and into such metabolic intermediates as CDP- choline, or it can be dephosphorylated to CDP (Fig. 25-14, step *i*). It is CDP that serves as the principal precursor for the deoxyribonucleotides **dCDP** and **thymidine diphosphate** (dTTP).



(25-15)

Deoxyribonucleotides. A chain involving NADPH, a flavoprotein, thioredoxin, and ribonucleotide reductase converts either the ribonucleoside diphosphates or triphosphates to the corresponding 2-deoxy forms (step *j*, Fig. 25-14) as indicated in Eq. 25-16.



(25-16)

Ribonucleotide reductases are discussed in Chapter 16. Some are iron-tyrosinate enzymes while others depend upon vitamin B_{12}, and reduction is at the nucleoside *triphosphate* level. Mammalian ribonucleotide reductase, which may be similar to that of *E. coli*, is regarded as an appropriate target for anticancer drugs. The enzyme is regulated by a complex set of feedback mechanisms, which apparently ensure that DNA precursors are synthesized only in amounts needed for DNA synthesis.[273] Because an excess of one deoxyribonucleotide can inhibit reduction of all ribonucleoside diphosphates, DNA synthesis can be inhibited by deoxyadenosine or by high levels of thymidine, despite the fact that both compounds are precursors of DNA.

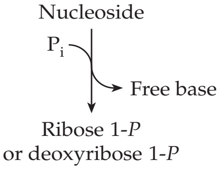
Phosphorylation of dCDP to dCTP (step *k*, Fig. 25-14) completes the biosynthesis of the first of the pyrimidine precursors of DNA. The uridine nucleotides arise in two ways. Reduction of UDP yields dUDP (step *j′*, Fig. 25-14). However, the deoxycytidine nucleotides are more often hydrolytically deaminated (reactions *l* and *l′*).[274] Methylation of dUMP to form **thymidylate**, dTMP (step *n*, Fig. 25-14), is catalyzed by thymidylate synthase. The reaction involves transfer of a 1-carbon unit from methylene tetrahydrofolic acid with subsequent reduction using THF as the electron donor. A probable mechanism is shown in Fig. 15-21. See also Box 15-E. Some bacterial transfer RNAs contain 4-thiouridine (Fig. 5-33). The sulfur atom is introduced by a sulfurtransferase (the *ThiI* gene product in *E. coli*). The same protein is essential for thiamin biosynthesis (Fig. 25-21).[274a]

Formation of dUMP in eukaryotes may occur by hydrolytic removal of phosphate from dUDP or from the conversions dCDP → dCMP → dUMP (steps *k′* and *l′*, Fig. 25-14). A more roundabout pathway is employed by *E. coli*: dCDP → dCTP → dUTP → dUMP (steps *k*, *l*, and *m*, Fig. 25-14). One of the intermediates is dUTP. DNA polymerases are able to incorporate dUMP from this compound into polynucleotides to form uracil-containing DNA. The only reason that this does not happen extensively within cells is that dUTP is rapidly converted to dUMP by a pyrophosphatase (step *m*, Fig. 25-14). The uracil that is incorporated into DNA is later removed by a repair enzyme (Chapter 27). The presence of dUTP in DNA provides the basis for one of the most widely used methods of directed mutation of DNA (Chapter 26).

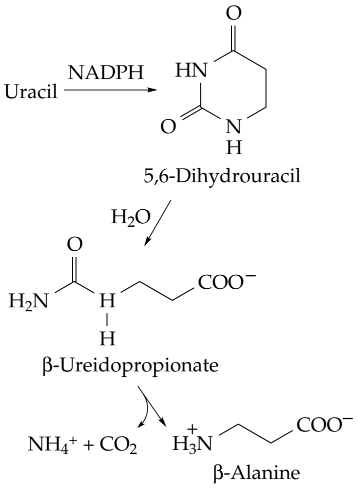
Bacteriophage-induced alterations in metabolism. Interesting alterations in nucleotide metabolism occur in cells of *E. coli* infected by T-even bacteriophage. Genes carried by the phage are transcribed, and the corresponding proteins are synthesized by the host cell.[273] A number of these viral gene products are enzymes affecting nucleotide metabolism. One enzyme catalyzes the hydrolytic conversion of dCTP to dCMP, and another promotes the synthesis of 4-hydroxymethyl-dCMP. Such virus-specified enzymes may be appropriate target sites for antiviral drugs.

2. Catabolism of Pyrimidine Nucleotides and Nucleosides

Nucleic acids within cells, as well as in the digestive tract, are continually under attack by many **nucleases**. Messenger RNA is degraded, often quite rapidly, as an essential part of the control of protein synthesis. Although DNA is very stable, nucleases are called upon to cut out damaged segments of single strands as part of essential repair processes (Chapter 27). Thus, there is an active breakdown of polynucleotides to mononucleotides, which are hydrolyzed to nucleosides by phosphatases. Nucleosides are converted to free bases by the action of **nucleoside phosphorylases** (Eq. 25-17). The further degradation of cytosine is initiated by deamination to uracil.[274b,c] Catabolism of uracil starts with reduction by NADPH according to Eq. 25-18 to form **β-alanine**.[275,275a] The latter can be oxidatively degraded to malonic semialdehyde and malonyl-CoA (see Fig. 17-3),[276] but it also serves as a biosynthetic precursor of pantothenic acid and coenzyme A (Eq. 24-38) and of the peptides carnosine and anserine.

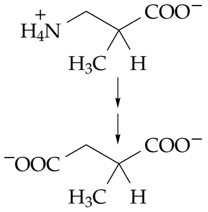


(25-17)



(25-18)

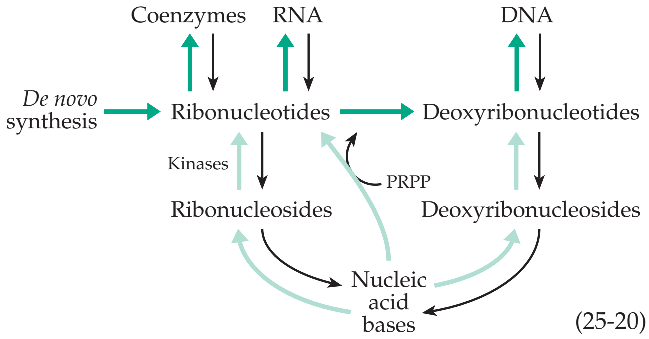
Thymine undergoes degradation in a pathway analogous to that of Eq. 25-18, but with the formation of **3-aminoisobutyrate**. The latter can be oxidatively converted to methylmalonate (Eq. 25-19), which can enter the methylmalonyl pathways (Fig. 17-3).



(25-19)

3. Reuse or Salvage of Nucleic Acid Bases

Almost all organisms except for protozoa synthesize nucleotides via the *de novo* pathways of Figs. 25-14 and 25-15. However, they also receive nucleotides, nucleosides, and free nucleic acid bases from catabolism of RNA and DNA. Both the synthetic and degradative pathways are carefully controlled by cells to ensure that they can grow and repair their nucleic acids and not be inhibited by accumulation of an excess of any component. Furthermore, animals receive an additional supply of preformed bases from their foods. Protozoa receive all of their bases in this way. A large set of enzymes is needed to break down the nucleotides and nucleosides. Another group of enzymes resynthesize nucleotides through **salvage pathways**.[273,277] This keeps levels of inhibitory compounds low, and at the same time produces a small constant pool of nucleotide triphosphates ready for use in nucleic acid synthesis. These pathways are summarized in a general way in Eq. 25-20, which uses dark green arrows for biosynthetic pathways, light green for salvage, and black for degradative reactions (after Kornberg and Baker[273]). The very active salvage pathways of protozoa have provided a variety of targets for inhibitors aimed at parasitic protozoa such as trypanosomes and *Toxoplasma gondii*. The target enzymes often have structures sufficiently different from the corresponding human enzymes to allow for design of selective inhibitors that can serve as drugs.[278,279]



(25-20)

Just as orotic acid is converted to a ribonucleotide in step *e* of Fig. 25-14, other free pyrimidine and purine bases can react with PRPP to give monoribonucleotides plus PP_i. The reversible reactions, which are catalyzed by **phosphoribosyltransferases** (ribonucleotide pyrophosphorylases), are important components of the salvage pathways by which purine and pyrimidine bases freed by the degradation of nucleic acids are recycled.[273] However, thymine is usually *not* reused. Thymine will react with deoxribose 1-*P* to form thymidine plus inorganic phosphate (thymidine phosphorylase), and thymidine is rapidly phosphorylated by the action of successive kinases to dTTP; a substrate for DNA polymerases.[273] This has allowed biochemists to introduce radioactive thymine or thymidine into the DNA of an organism, an important experimental tool. Another important reaction of the salvage pathways for pyrimidines is the conversion of cytosine to uracil, the same kind of hydrolytic deamination represented by step *l* in Fig. 25-14.

E. Biosynthesis and Metabolism of Purines

The first decisive experiments shedding light on the biosynthetic origins of purines were done with pigeons, which form large amounts of uric acid. Labeling experiments established the complex pattern indicated in the box in the upper left-hand corner of Fig. 25-15. Two carbon atoms were found derived from glycine, one from CO_2, and two from formate. One nitrogen came from glycine, two from glutamine, and one from aspartate. In the case of adenine, the 6-NH_2 group was also found derived from aspartate.

1. The Enzymatic Reactions of Purine Synthesis

The detailed biosynthetic pathway, for which enzymes have now been isolated and studied, is indicated in Fig. 25-15. The first “committed step” in purine synthesis is the reaction of PRPP with glutamine to form **phosphoribosylamine** (step *a*).[280,281] This is another glutamine-dependent amination, pyrophosphate being displaced by ammonia generated *in situ* from glutamine. The amino group of the intermediate so-formed is coupled with glycine in a standard manner (step *b*),[282] and the resulting product is formylated by 10-formyltetrahydrofolate (step *c*).[283–285b] The latter can be generated from free formate,[286] accounting for the labeling pattern indicated in the box in Fig. 25-15. For many years it was accepted, incorrectly, that 5,10-methenyl-THF was the formyl donor for this reaction.

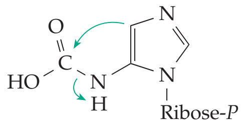
In step *d* of Fig. 25-15 a second glutamine-dependent amination takes place, possibly through aminolysis of an intermediate enol phosphate. An ATP-requiring ring closure and tautomerization (step *e*) serve to complete the formation of the imidazole ring.[287,287a] In many, perhaps all, eukaryotes a single multifunctional enzyme catalyzes steps *b*, *c*, and *e* of Fig. 25-15. The chicken enzyme has ~110-kDa subunits.[288] The product is a ribonucleotide of 5-amino-4-imidazole-4-carboxamide, (AIR), a compound that was isolated in 1945 from cultures of *E. coli* treated with sulfonamides. The latter are important drugs which are antagonists of *p*-aminobenzoate (Box 9-C) and interfere with completion of purine synthesis. This deprives the bacterial cells of essential folic acid derivatives. Its structure immediately suggested that 5-aminoimidazole-4-carboxamide might be a purine precursor. Later it was shown that it is actually the corresponding ribonucleotide (AIR or ZMP) that lies on the main route of purine synthesis. It also is an intermediate in the biosynthesis of thiamin (Fig. 25-21). Free 5-aminoimidazole-4-carboxamide participates in formation of histidine (Fig. 25-13).

The trinucleotide ZTP also accumulates, not only in bacteria but also in many eukaryotic cells. Bochner and Ames suggested it may be an **alarmone** signaling a deficit of folate coenzymes in the cell and causing a shutdown of protein synthesis. ZTP is synthesized by an unusual reaction, transfer of a pyrophosphate group from PRPP (phosphoribosyl pyrophosphate).

$$\text{ZMP + PRPP} \rightarrow \text{ZTP + Ribose 5'-}P \qquad (25\text{-}21)$$

This is similar to the reaction by which guanosine 5'-diphosphate 3'-diphosphate (ppGpp) is formed from GDP and ATP (Eq. 29-11).

In the next step of purine synthesis (Fig. 25-15, step *f*) a molecule of CO_2 is incorporated in an unusual type of carboxylation. It is shown in Fig. 25-15 as a single-step direct reaction of CO_2 with AIR. However, in many organisms it is a two-step ATP-dependent reaction to form a compound carboxylated on the 5-amino group. This rearranges[289–290a] to the product shown in Fig. 25-15.



While the arrows on the foregoing structure suggest a possible mechanism of rearrangement, the implied four-membered ring transition state makes it unlikely. The reaction resembles biotin-dependent carboxylations, suggesting the possibility that the carboxylate releases CO_2, which moves and rebinds while trapped within a closed active site.[290] In a two-step amination reaction (steps *g* and *h*) nitrogen is transferred from aspartate in a manner strictly comparable to that in urea synthesis in which argininosuccinic acid is an intermediate (Fig. 24-10). As in urea formation, the carbon skeleton of the aspartate molecule is eliminated as fumarate (step *h*), leaving the nitrogen in the purine precursor. The final carbon atom is added from 10-formyltetrahydrofolic acid (step *i*).[291–292b] Spontaneous ring closure is followed by dehydration to **inosine 5'-phosphate** (IMP, inosinic acid), step *j*. Steps *i* and *j* are catalyzed by a single bifunctional enzyme.

IMP is converted via two different pathways to either AMP or GTP. Conversion to AMP (Fig. 25-15, steps *k* and *l*) occurs via another two-step aspartate-dependent amination.[292c,d] The intermediate is

adenylosuccinate.[293–294b] Positional isotope exchange studies using [γ-^{18}O]GTP suggested that the enolic 6-phospho-IMP is an intermediate comparable to that in Eq. 24-23.[295,296] X-ray studies have confirmed the prediction.[297] An NAD^+-dependent oxidation converts IMP to the corresponding **xanthine** ribonucleotide

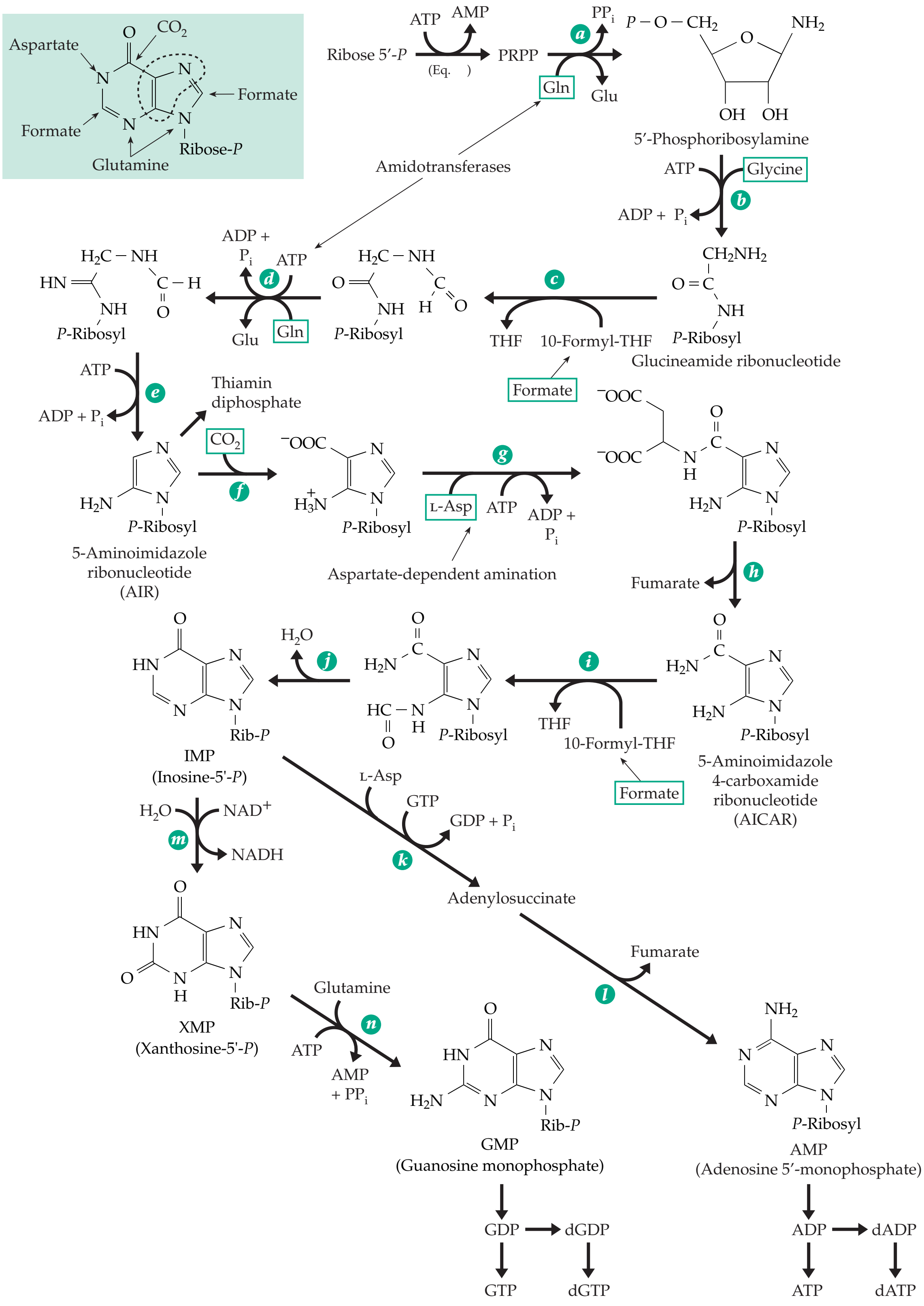

Aspartate
CO_2
Formate
Formate
Glutamine
Ribose-P
Ribose 5′-P
ATP
AMP
(Eq.)
PRPP
PP_i
a
Gln
Glu
Amidotransferases
5′-Phosphoribosylamine
ATP
Glycine
b
ADP + P_i
Glucineamide ribonucleotide
c
THF
10-Formyl-THF
Formate
ADP + P_i
ATP
d
Glu
Gln
P-Ribosyl
ATP
e
ADP + P_i
Thiamin diphosphate
CO_2
f
5-Aminoimidazole ribonucleotide (AIR)
g
L-Asp
ATP
ADP + P_i
Aspartate-dependent amination
Fumarate
h
i
THF
10-Formyl-THF
Formate
5-Aminoimidazole 4-carboxamide ribonucleotide (AICAR)
H_2O
j
IMP (Inosine-5′-P)
Rib-P
H_2O
NAD^+
m
NADH
L-Asp
GTP
GDP + P_i
k
Adenylosuccinate
Fumarate
l
XMP (Xanthosine-5′-P)
Glutamine
ATP
n
AMP + PP_i
GMP (Guanosine monophosphate)
AMP (Adenosine 5′-monophosphate)
GDP
dGDP
GTP
dGTP
ADP
dADP
ATP
dATP

Figure 25-15 Biosynthesis of purine nucleotides from ribose 5'-phosphate.

(step *m*),[297a–d] which is aminated in a glutamine-dependent process,[298–300] as indicated (step *n*).

Synthesis of purines is under complex control.[273] Some of the mechanisms found in bacteria are outlined in Fig. 25-16. Both feedback inhibition and activation are involved. Very important is the fact that GTP is needed in the synthesis of ATP, and that ATP is needed for synthesis of GTP. This kind of control ensures that an excess of either nucleotide will not be formed for long. In bacteria all of the final end product nucleotides inhibit the initial reaction of step *a* in Fig. 25-15.

Modified purine nucleosides are important constituents of transfer RNAs (Fig. 5-33; Chapter 28, Section A,7). Among them are the 7-deazaguanosines **queuosine** (Fig. 5-33) and **archaeosine**, which contains a 7-formamidino group. These nucleosides are incorporated into tRNAs by an exchange mechanism catalyzed by **tRNA-guanine** transglycosylase. This exchanges a precursor of queuine (7-aminomethyl-7-deazaguanine) for guanine in selected residues in the tRNAs. In eukaryotic tRNAs free queuine is exchanged into the tRNAs whereas in Archaea the archaeosine precursor is incorporated.[301a,b] The conversion of the precursors to the final modified bases occurs in the tRNAs (Eq. 28-4). Many other purine derivatives are found in nature, e.g., **puromycin** (Box 29-B), which is formed from adenosine.[301c]

2. The Purine Nucleotide Cycle and Salvage Pathways for Purines

Muscular work is accompanied by the production of ammonia, the immediate source of which is adenosine 5'-phosphate (AMP).[301,302] This fact led to the recognition of another substrate cycle (Chapter 11) that functions by virtue of the presence of a biosynthetic pathway and of a degradative enzyme in the same cells (cycle A, Fig. 25-17). This **purine nucleotide cycle** operates in the brain[303,304] as well as in muscle. The key enzyme 5'-AMP aminohydrolase (AMP deaminase; step *a*, Fig. 25-17) also occurs in erythrocytes and many other tissues.[304,305] Persons having normal erythrocyte levels but an absence of this enzyme in muscles suffer from muscular weakness and cramping after exercise.[306]

Purine bases from ingested foods, or formed by catabolism of nucleic acids, are able to react with PRPP under the influence of phosphoribosyltransferases.[306a] Two such enzymes are known to act on purines. One converts adenine to AMP (Fig. 25-17, step *b*) and also acts upon 5-aminoimidazole-4-carboxamide. This enzyme may be especially important to parasitic protozoa such as *Leishmania*, which lack the *de novo* pathway of purine synthesis (Fig. 25-15).[278,306b]

AMP can be converted by the action of AMP 5'-nucleotidase to adenosine (step *c*, Fig. 25-17), which is thought to be an important local hormone or second

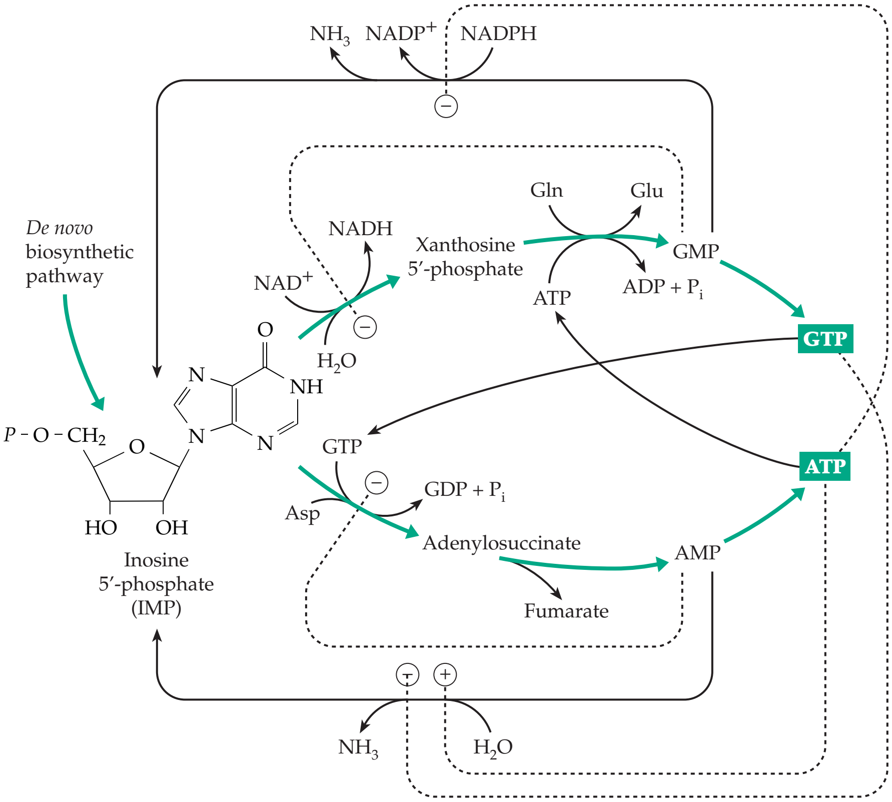


Figure 25-16 Control of the conversion of inosine 5'-phosphate to the adenine and guanine ribonucleotides and deoxyribonucleotides in bacteria by feedback inhibition and activation.

messenger.[307–309] Adenosine has a variety of effects on all organs of an animal. It affects heart rate, smooth muscle tone, and white blood cell function. It modulates the catabolic effects of hormones such as catecholamines and stimulation of the anabolic hormone insulin.[309] Adenosine receptors of at least three types are known.[310,311] Binding of adenosine to A_1 receptors inhibits adenylate cyclase, while binding to A_2 receptors stimulates this enzyme.[307] However, effects of adenosine on K^+ transport are probably more important.

Deamination of adenosine (step *d*) together with reconversion of the resulting inosine to IMP (steps *e* and *f*) completes a second cycle (cycle B, Fig. 25-17). Intense interest has been focused on adenosine deaminase because hereditary lack of this enzyme is linked to a severe immunodeficiency in which the numbers of B and T lymphocytes are inadequate to combat infection.[312] Until recently bone marrow transplantation in infancy was the only possible treatment for this otherwise fatal disease. Regular injections of adenosine deaminase covalently attached to polyethylene glycol (to delay removal from the bloodstream) have been used for some patients. Since 1990 **gene therapy**, transfer of an adenosine deaminase gene into white blood cells, has also been used with apparent success.[313–315] This topic is discussed in Chapter 31. Adenosine deaminase is a 40-kDa protein,[316,317] which exists as a complex with a large 200-kDa binding protein[315] which anchors the deaminase to cell membranes.

The basic cause of the severe immunodeficiency symptoms is uncertain. However, adenosine deaminase also catalyzes hydrolysis of 2'-deoxyadenosine, and in the absence of the enzyme both this compound and its trinucleotide precursor 2'-deoxy-ATP (dATP) accumulate in tissues.[312] Ribonucleotide reductase is allosterically inhibited by dATP, and this inhibition may interfere with DNA synthesis and with the rapid growth of lymphocytes needed in response to infections. Since T lymphocytes are more severely affected than B lymphocytes, it is necessary to postulate a difference in the extent to which these two cell types accumulate dATP.

The conversions of inosine to hypoxanthine (Fig. 25-17, step *e*), of guanosine to guanine (step *g*), and of other purine ribonucleosides and deoxyribonucleosides to free purine bases are catalyzed by **purine nucleoside phosphorylase**.[318–321b] Absence of this enzyme also causes a severe immune deficiency which involves the T cells. However, B cell function is not impaired.[312,315,322]

The last enzyme in cycle B of Fig. 25-17 (catalyzing step *f* and also step *b*) is the X-linked **hypoxanthine-guanine phosphoribosyltransferase** (HGPRT or HPRT).[322a,b] Its absence causes the **Lesch–Nyhan** syndrome characterized not only by overproduction of uric acid but by a serious disorder of the central nervous system. It causes both mental retardation and a compulsive form of self-mutilation of the gums and hands by biting.[323–325] The excessive production of uric acid is easy to understand because the accumulating hypoxanthine and guanine are both readily converted to uric acid by the reactions of Fig. 25-18. Patients with a partial deficiency in HGPRT escape the worst neurological symptoms but may have severe gouty arthritis (Box 25-D).[326] Efforts are being made to treat the disease by gene transfer.[327]

Trypanosomes and other parasitic protozoa are unable to synthesize purines and must obtain them from their hosts using salvage pathway. Selective inhibition of their HGPRT or of nucleoside hydrolases, which are absent from mammalian cells, are goals of drug development.[327a,b]

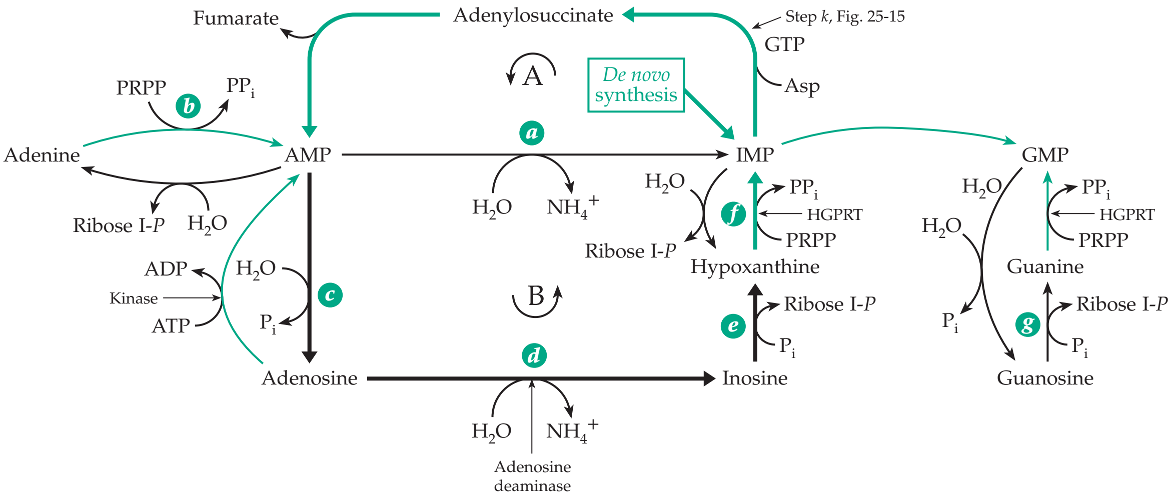


Figure 25-17 Some purine salvage pathways and related reactions. Green lines indicate biosynthetic pathways.

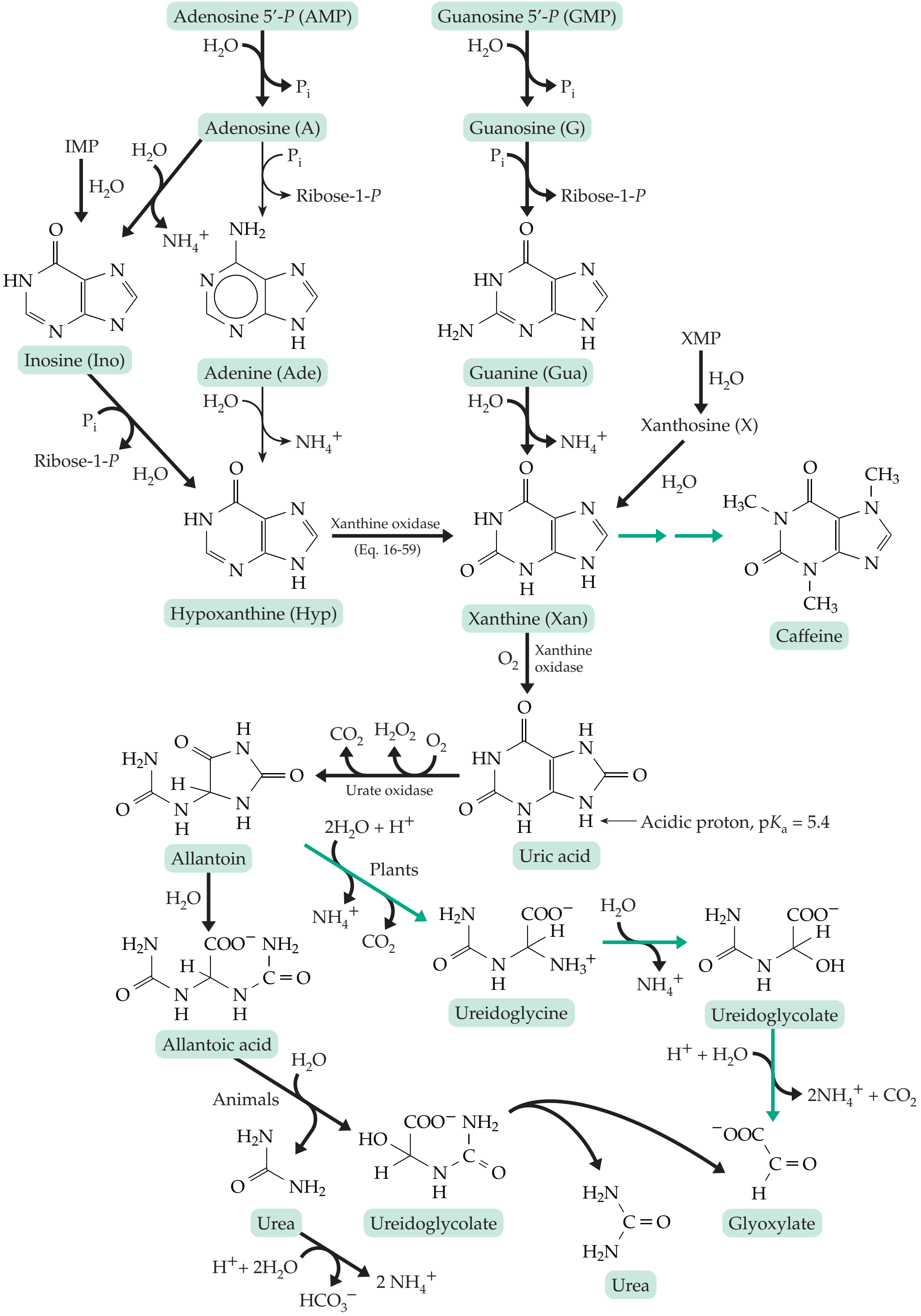

Adenosine 5′-P (AMP)
Guanosine 5′-P (GMP)
H2O
Pi
Adenosine (A)
Guanosine (G)
IMP
H2O
NH4+
Ribose-1-P
Inosine (Ino)
Adenine (Ade)
Guanine (Gua)
XMP
Xanthosine (X)
Ribose-1-P
Hypoxanthine (Hyp)
Xanthine oxidase
(Eq. 16-59)
Xanthine (Xan)
Caffeine
O2
Xanthine oxidase
CO2 H2O2 O2
Urate oxidase
Uric acid
Acidic proton, pKa = 5.4
Allantoin
2H2O + H+
Plants
NH4+
CO2
Ureidoglycine
Ureidoglycolate
Allantoic acid
H+ + H2O
2NH4+ + CO2
Animals
Urea
Ureidoglycolate
Urea
Glyoxylate
H+ + 2H2O
HCO3−
2 NH4+

Figure 25-18 Pathways of catabolism of purine nucleotides, nucleosides, and free bases. Spiders excrete xanthine while mammals and birds excrete uric acid. Spiders and birds convert all of their excess nitrogen via the *de novo* pathway of Fig. 25-15 into purines. Many animals excrete allantoin, urea, or NH_4^+. Some legumes utilize the pathway marked by green arrows in their nitrogen transport via ureides.

3. Oxidative Metabolism of Purines

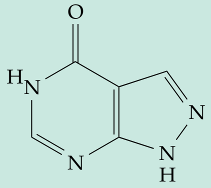
As indicated in Fig. 25-18, free adenine released from catabolism of nucleic acids can be deaminated hydrolytically to hypoxanthine, and guanine can be deaminated to xanthine.[328] The molybdenum-containing xanthine oxidase (Chapter 16) oxidizes hypoxanthine to xanthine and the latter on to uric acid. Some clostridia convert purine or hypoxanthine to xanthine by the action of a selenium-containing purine hydroxylase.[328a] Another reaction of xanthine occurring in some plants is conversion to the trimethylated derivative **caffeine**.[328b] One of the physiological effects of caffeine in animals is inhibition of pyrimidine synthesis.[329] However, the effect most sought by coffee drinkers may be an increase in blood pressure caused by occupancy of adenosine receptors by caffeine.[330]

Uric acid is the end product of purine metabolism in the human. Spiders excrete **xanthine**,[331,332] but in most animals **urate oxidase** converts uric acid to **allantoin**. Although urate oxidase contains no coenzymes or metal cofactors,[333] it catalyzes the reaction with O_2 to form a peroxide (Eq. 25-22),[334,335] a reaction resembling that of a reduced flavin (Eq. 15-31) or a reduced pterin. As in these other cases there may be an initial electron transfer between O_2 and urate to form a radical pair which couple. Elimination of H_2O_2, a feature also of flavoprotein oxidases, accomplishes oxidation of the urate ring. Hydration, ring opening, and decarboxylation complete the conversion to allantoin. The ease of formation of urate radicals permits uric acid to act as an effective oxidant. This may account for the fact that we long-lived human beings retain a high internal urate concentration.[335a,b]

BOX 25-D GOUT

A common metabolic derangement with an incidence of ~3 per 1000 persons is **hyperuricemia** or **gout**.[a,b] As with most metabolic defects, there is a family of diseases ranging from mild to severe. In acute gouty arthritis, a sudden attack occurs, usually in the night, when sodium urate crystals precipitate in one or more joints. In half the cases the victim is awakened by a terrible pain in the big toe. The disease most often strikes adult males. The heredity is apparently complex and not fully understood. The primary biochemical defect in gout is usually an overproduction of uric acid which, in some cases, may result from an overactive PRPP synthase.[c] In other cases a kidney defect interferes with excretion. The less severe saturnine gout, which occurs in relatively young persons of both sexes, is a result of chronic lead poisoning. It may involve deposition of guanine in the joints as a result of inhibition by Pb^{2+} of guanine aminohydrolase, the enzyme that hydrolyzes guanine to xanthine (Fig. 25-18).[d]

If properly controlled, simple gout may have few adverse effects. However, the severe neurological symptoms of Lesch–Nyhan syndrome (Section E,2 of text)[e] cannot be corrected by medication. Colchicine (Box 7-D), in a manner which is not understood, alleviates the painful symptoms of gout caused by the deposits of sodium urate in joints and tissues. It is also important to keep the dietary purine intake low and it is often necessary to inhibit xanthine oxidase. A widely used and effective inhibitor is the isomer of hypoxanthine known as **allopurinol**, which is taken daily in amounts of 100–600 mg or more.



Allopurinol

Allopurinol and its oxidation product **oxypurinol**, a xanthine analog, both inhibit xanthine oxidase and patients receiving allopurinol excrete much of their purines as xanthine and hypoxanthine. Nucleotide derivatives of oxypurinol also inhibit the *de novo* purine biosynthetic pathway. The accumulating hypoxanthine is reused to a greater extent than normal, decreasing the total purine excretion. A number of other drugs stimulate increased excretion of uric acid.[e] Although many patients tolerate allopurinol for many years, some experience dangerous side effects.

---

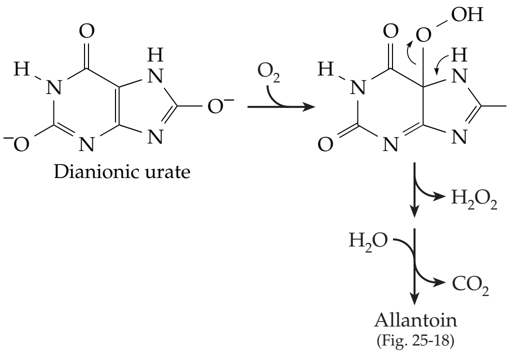
[a] Becker, M. A., and Roessler, B. J. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1655–1677, McGraw-Hill, New York

[b] Kelley, W. N., and Wyngaarden, J. B. (1972) in *The Metabolic Basis of Inherited Disease*, 3rd ed. (Stanbury, J. B., Wyngaarden, J. B., and Fredrickson, D. S., eds), pp. 969–1002, McGraw-Hill, New York

[c] Becker, M. A., Kostel, P. J., and Meyer, L. J. (1975) *J. Biol. Chem.* **250**, 6822–6830

[d] Farkas, W. T., Stanawitz, T., and Scheider, M. (1978) *Science* **199**, 786–787

[e] Rossiter, B. J. F., and Caskey, C. T. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1679–1706, McGraw-Hill, New York



(25-22)

Allantoin is the excretory product in most mammals other than primates. Most fish hydrolyze allantoin to **allantoic acid,** and some excrete that compound as an end product. However, most continue the hydrolysis to form urea and glyoxylate using peroxisomal enzymes.[336] In some invertebrates the urea may be hydrolyzed further to ammonia. In organisms that hydrolyze uric acid to urea or ammonia, this pathway is used only for degradation of purines from nucleotides. Excess nitrogen from catabolism of amino acids either is excreted directly as ammonia or is converted to urea by the urea cycle (Fig. 24-10).

Plants also form the **ureides** allantoin and allantoic acid, and in some legumes, such as soy beans, these compounds account for 70–80% of the organic nitrogen in the xylem. They appear to function in nitrogen transport.[337] As indicated in Fig. 25-18, the hydrolysis to glyoxylate, NH_4^+, and CO_2 follows a different pathway than in animals. See also Chapter 24, Section C.

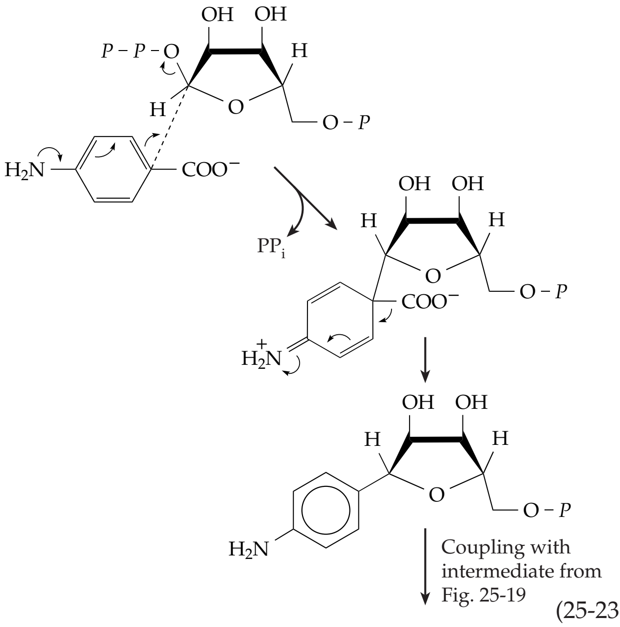
F. Pterins, Flavins, Dimethylbenzimidazole, Thiamin, and Vitamin B_6

Tracer studies have established that both folic acid and riboflavin originate from guanosine phosphates (Fig. 25-19).[338–341] All of the atoms of the purine ring are conserved in the products except for C-8 of the five-membered ring. The first step is opening of the 5-membered ring and the hydrolytic removal of formate (Fig. 25-19, step *a*). This is followed by an Amadori rearrangement (step *b*) and a simple ring closure between the resulting carbonyl and adjacent amino group (step c). The product is **7,8-dihydroneopterin triphosphate**. The single enzyme **GTP cyclohydrolase** catalyzes all three steps *a*–*c*.[342,342a–c] Dihydroneopterin is a central intermediate in pterin metabolism, being converted by one route into folate and methanopterin coenzymes and by another into biopterin, drosopterin, and others.[343]

An aldol cleavage (step *d*) followed by a series of reactions, shown at the left in in Fig. 25-19, leads to folate.[344–347] The reactions include the ATP-dependent conversion of 6-hydroxy-7,8-dihydropterin to its pyrophosphate ester.[347a] This is followed by coupling to *p*-aminobenzoate with elimination of PP_i to form dihydropteroic acid.[347b] ATP-dependent joining of glutamate and reduction yields **tetrahydrofolate**. Additional molecules of γ-linked glutamate are added to form the functional polyglutamate forms (p. 803).[347c,d] Gamma-glutamyl hydrolases provide essential turnover.[347e] Formation of biopterin is initiated by tautomerization of dihydroneopterin triphosphate (step *e*), a proton from the solvent binding to C-6 of the tetrahydropterin ring.[348] The single Zn^{2+}-dependent enzyme **6-pyruvoyltetrahydropterin synthase**[349–352] catalyzes the two consecutive tautomerization steps shown in Fig. 25-19 as well as elimination of the triphosphate group tripolyphosphate (step *f*), a reaction facilitated by the carbonyl group introduced in the preceding step.[349] The same enzyme promotes a final tautomerization to form 6-pyruvoyltetrahydropterin, a compound that is reduced by NADPH to tetrahydrobiopterin. The reaction is catalyzed by **sepiapterin reductase**.[353,354] Notice that biopterin, like the folates, is synthesized at the oxidation level of a tetrahydrobiopterin.

The *Drosophila* eye pigments **sepiapterin** and **drosopterins** (Figs. 15-17 and 25-19) arise from 6-pyruvoyltetrahydropterin.[355–357] Reduced glutathione appears to be the reducing agent needed to convert the 6-pyruvoyltetrahydropterin into the more reduced pyrimidinodiazepine (step *h*) with its 7-membered ring (Fig. 25-19).[357,358] Tetrahydrobiopterin can arise in mammalian cells, not only by the *de novo* pathway of Fig. 25-19 but also from salvage of sepiapterin.

Methanopterin (Fig. 15-17) is formed via a branch in the folate pathway. The 5-carbon chain that replaces the carboxyl group of *p*-aminobenzoate in the folates is derived by reaction of PRPP with *p*-aminobenzoate (Eq. 25-23).[359,360] White has proposed a



(25-23)

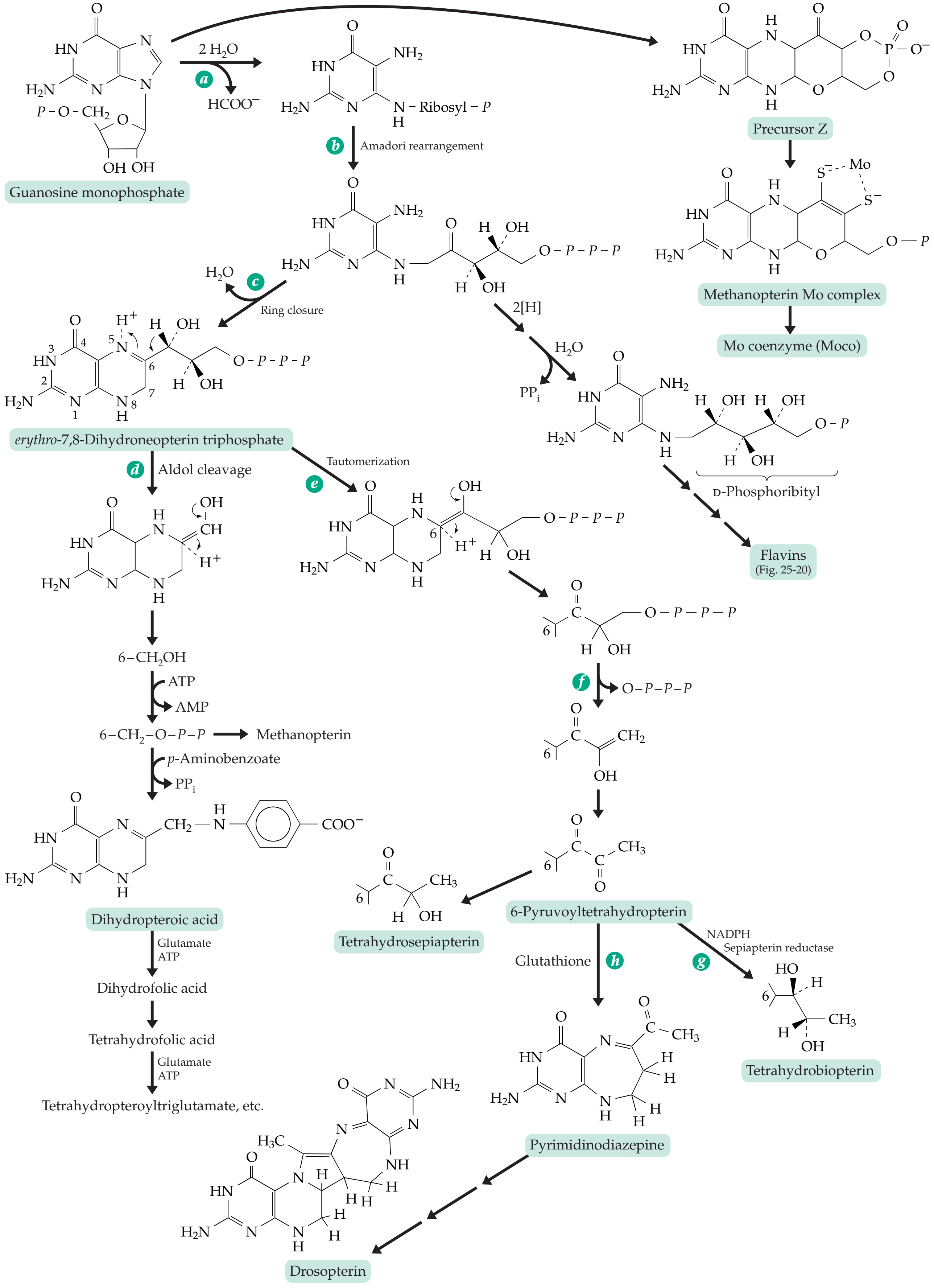

Guanosine monophosphate
2 H$_2$O
a
HCOO$^-$
N – Ribosyl – P
b
Amadori rearrangement
c
H$_2$O
Ring closure
O – P – P – P
erythro-7,8-Dihydroneopterin triphosphate
d
Aldol cleavage
Tautomerization
e
2[H]
H$_2$O
PP$_i$
O – P
D-Phosphoribityl
Flavins
(Fig. 25-20)
Precursor Z
Methanopterin Mo complex
Mo coenzyme (Moco)
6–CH$_2$OH
ATP
AMP
6–CH$_2$–O–P–P
Methanopterin
p-Aminobenzoate
PP$_i$
COO$^-$
Dihydropteroic acid
Glutamate
ATP
Dihydrofolic acid
Tetrahydrofolic acid
Glutamate
ATP
Tetrahydropteroyltriglutamate, etc.
f
O–P–P–P
6-Pyruvoyltetrahydropterin
Tetrahydrosepiapterin
Glutathione
h
g
NADPH
Sepiapterin reductase
Tetrahydrobiopterin
Pyrimidinodiazepine
Drosopterin

Figure 25-19 The biosynthesis of folic acid and other pterins.

detailed pathway for completion of the methanopterin synthesis.[359] After coupling with 6-hydroxymethyl-7,8-dihydropterin pyrophosphate (shown in abbreviated form in Fig. 25-19) with loss of PP_i, the ribose ring in the product of Eq. 25-23 is reductively opened to give the ribityl side chain of methanopterin (see Fig. 15-17). Hydrolytic removal of the phosphate is then followed by transfer of the α-linked ribose phosphate. The *S*-hydroxyglutarate (derived by reduction of 2-oxoglutarate) is added in an ATP-dependent reaction. The extra methyl groups at positions 7 and 9 are transferred from *S*-adenosylmethione and the compound is reduced to the tetrahydropterin state.

Biosynthesis of **molybdopterin** (Fig. 15-17) also begins with a guanosine derivative, quite possibly GMP as is shown in Fig. 25-19. However, the C-8 carbon atom is not lost as formate, as in step *a* of Fig. 25-19, but is incorporated into the side chain in the molybdopterin precursor Z. A pathway was proposed by Wuebbens and Rajagopalan.[361] The first stage in the reaction sequence,[361a–d] which is identical in most organisms, is formation of the metastable compound Z, a proposed structure of which is shown in Fig. 25-19. The conversion of this compound into molybdopterin requires opening of the cyclic phosphate ring and incorporation of two atoms of sulfur. These may both be released from cysteine as S^0 and carried by a sulfurtransferase as an enzyme-bound persulfide group (see Chapter 24, Section G,3).[361b] A thiocarboxylate group generated from a C-terminal carboxylate of a molybdopterin synthase subunit, as in the ThiS protein (p. 1463), may be the direct sulfur donor. It is probably formed in an ATP-dependent process.[361b] Incorporation of molybdenum, perhaps from MoO_4^{2-}, completes the synthesis of the molybdenum cofactor Moco.[361c,e] In some molybdoenzymes, e.g., xanthine oxidase, an additional sulfur atom is bound to the Mo atom (Fig. 16-32). This is also obtained from cysteine using a PLP-dependent sulfurtransferase similar to the NifS protein.[361f,g] In many bacteria molybdopterin is joined to GMP, AMP, IMP, or CMP to form a dinucleotide.[361h,i]

Both the fungus *Eremothecium* (Box 15-B) and mutants of *Saccharomyces* have been used to deduce the pathways of **riboflavin** synthesis outlined in Figure 25-20. The first reaction (step *a*) is identical to step *a* of Fig. 25-19 but is catalyzed by a different GTP cyclohydrolase.[362] Instead of an Amadori rearrangement it catalyzes the hydrolytic deamination and dephosphorylation (step *b*) to give the flavin precursor 4-ribitylamino-5-amino-2,6-dihydroxypyrimidine. Additional carbon atoms to build the benzene ring of riboflavin are supplied in two stages from ribulose 5-phosphate. Isotopic labeling showed that carbon atoms 1, 2, 3, and 5 of this compound are utilized as is marked in Fig. 25-20, while C-4 is eliminated as formic acid in a rearrangement (step *c*, Fig. 25-20). A plausible mechanism has been suggested.[363,363a,b] The product L-3,4-dihydroxy-2-butanone 4-phosphate condenses (step *d*) with the product formed from GMP, possibly via the Schiff base shown. Elimination of H_2O and inorganic phosphate followed by tautomerization gives **6,7-dimethyl-8-ribityllumazine**.[363a,364–364c] Completion of the flavin ring requires an additional four carbon atoms, which are supplied by a second molecule of 6,7-dimethyl-8-ribityllumazine, as indicated in Fig. 25-20.[364d] This disproportionation reaction appears remarkable but is less so when one considers that the bimolecular reaction to form riboflavin occurs spontaneously under mild conditions. The precursor 4-ribitylamino-5-amino-2,6-dihydroxypyrimidine is regenerated in this process (Fig. 25-20).

The enzyme complex that catalyses steps *d* to *f* of Fig. 25-20 has an unusual composition. An α_3 trimer of 23.5-kDa subunits is contained within an icosahedral shell of 60 16-kDa β subunits, similar to the protein coats of the icosahedral viruses (Chapter 7). The β subunits catalyze the formation of dimethylribityllumazine (steps *d*, *e*), while the α_3 trimer catalyzes the dismutation reaction of step *f*, the final step in riboflavin formation.[365] A separate bifunctional bacterial ATP-dependent synthetase phosphorylates riboflavin and adds the adenylyl group to form FAD.[366] Two separate mammalian enzymes are required.[367] Synthesis of **deazaflavins** of methanogens (Fig. 15-22) follows pathways similar to those of riboflavin. However, the phenolic ring of the deazaflavin originates from the shikimate pathway.[368]

Dimethylbenzimidazole, a constituent of vitamin B_{12} (Box 16-B), also arises from 6,7-dimethyl-8-ribityllumazine in a process resembling that of riboflavin synthesis, but in which the riboflavin formed is hydrolytically degraded to remove the pyrimidine ring and to form the imidazole ring.[369] Conversion to an α-ribazole and linkage to the aminopropyl group of the corrin ring is described by Thompson *et al*.[370] Various related cobamides are also formed by bacteria.[371] Synthesis of the corrin ring is described briefly in Chapter 24, and the chemistry of the ligands to cobalt at the "top" of the vitamin B_{12} molecules is considered in Chapter 16, Section B.

Thiamin. Investigation of the biosynthesis of thiamin has been difficult because only minute amounts are formed by microorganisms such as *E. coli* or yeast. Furthermore, significant differences in the routes of synthesis in different organisms have caused confusion.[372,372a] The pathways outlined in Fig. 25-21 are incomplete.

The pyrimidine portion of thiamin (Fig. 25-21) is distinct in structure from the pyrimidines of nucleic acids. In bacteria it originates from the purine precursor 5-aminoimidazole ribotide, which is converted into a hydroxymethylpyrimidine (Fig. 25-21)[373] which is

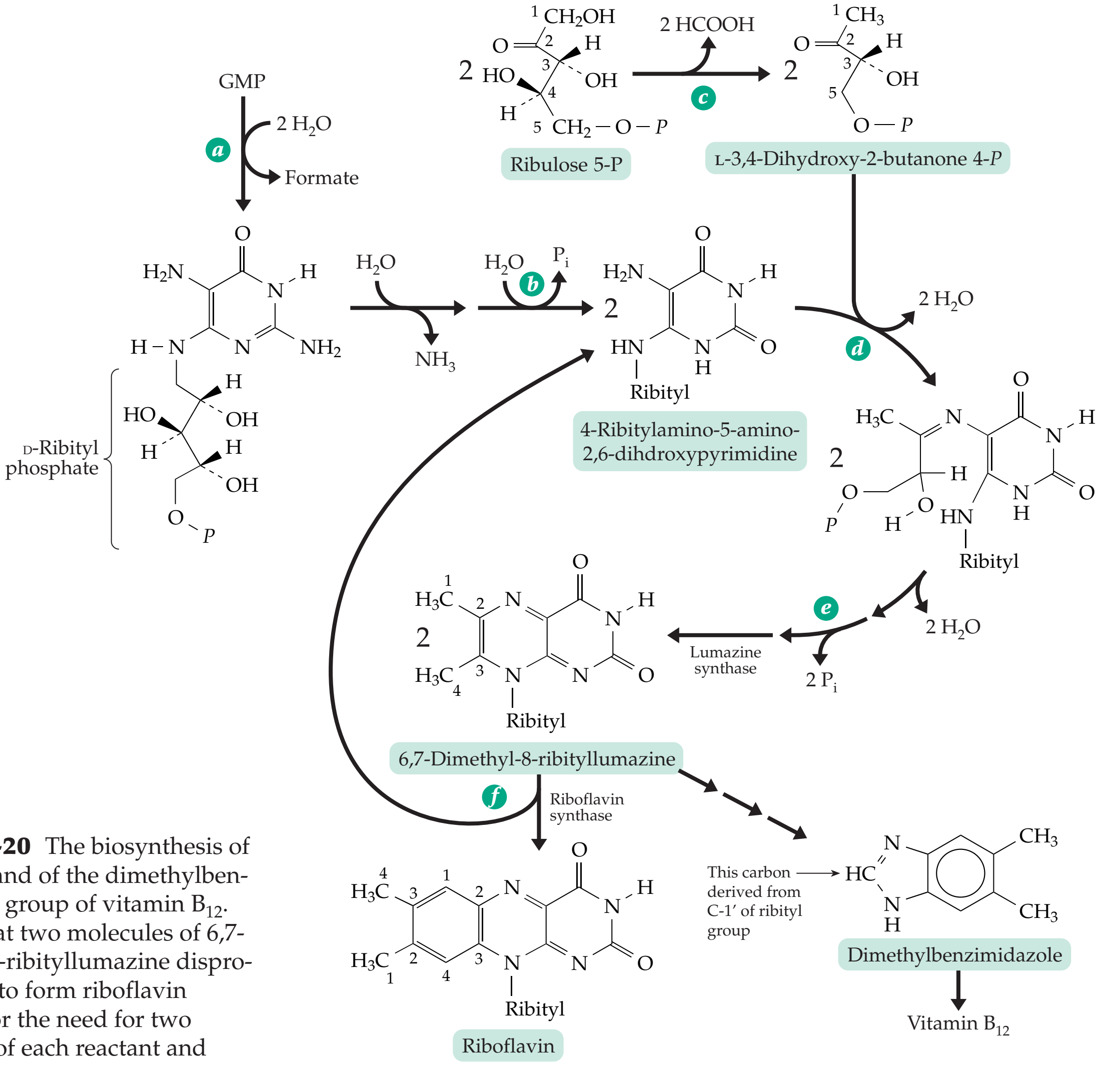
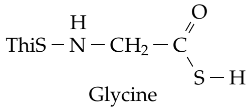

GMP
2 H_2O
a
Formate
D-Ribityl phosphate
H_2O
NH_3
H_2O
P_i
b
2
4-Ribitylamino-5-amino-2,6-dihdroxypyrimidine
Ribulose 5-P
2 HCOOH
c
L-3,4-Dihydroxy-2-butanone 4-P
2 H_2O
d
2
Ribityl
e
2 H_2O
2 P_i
Lumazine synthase
6,7-Dimethyl-8-ribityllumazine
f
Riboflavin synthase
Riboflavin
This carbon derived from C-1′ of ribityl group
Dimethylbenzimidazole
Vitamin B_{12}

Figure 25-20 The biosynthesis of riboflavin and of the dimethylbenzimidazole group of vitamin B_{12}. The fact that two molecules of 6,7-dimethyl-8-ribityllumazine disproportionate to form riboflavin accounts for the need for two molecules of each reactant and product in many steps.

coupled with the thiazolium ring to form the vitamin. All of the carbon atoms of the substituted pyrimidine can be derived from the 5-aminoimidazole ribotide, but the pathway is uncertain. Both glycine and formate enter the pyrimidine, but labeling patterns are different in *E. coli* and in yeast.[372]

The thiazole ring is assembled on the 5-carbon backbone of **1-deoxyxylulose 5-phosphate**, which is also an intermediate in the alternative biosynthetic pathway for terpenes (Fig. 22-2) and in synthesis of vitamin B_6 (Fig. 25-21). In *E. coli* the sulfur atom of the thiazole comes from cysteine and the nitrogen from tyrosine.[374] The same is true for chloroplasts,[375] whereas in yeast glycine appears to donate the nitrogen.[372] The thiamin biosynthetic operon of *E. coli* contains six genes,[372a,376] one of which (***ThiS***) encodes a protein that serves as a sulfur carrier from cysteine into the thiazole.[374] The C-terminal glycine is converted into a thiocarboxylate:



$$\text{ThiS}-\overset{\text{H}}{\text{N}}-\text{CH}_2-\text{C}(=\text{O})-\text{S}-\text{H}$$

Glycine

The *ThiI* gene, which encodes aother sulfurtransferase protein, is also needed.[374a] The enzymology of the insertion of this sulfur into the thiazole is uncertain but may resemble that involved in synthesis of biotin, lipoic acid, and molybdopterin.[374] Linkage of the two parts of the thiamin molecule (step *d*, Fig. 25-21) is catalyzed by thiamin phosphate synthase, evidently via an S_N2 type reaction.[377–377b]

Pyridoxol (vitamin B_6). Again 1-deoxyxylulose 5-*P* serves as a precursor.[378] In *E. coli* only two genes have been implicated in the condensation of this compound with 4-(phosphohydroxy)-L-threonine (Fig. 25-21, step *f*).[378a] One is an NAD$^+$-dependent dehydrogenase that acts on the second substrate prior to the condensation. Significant differences from the path-

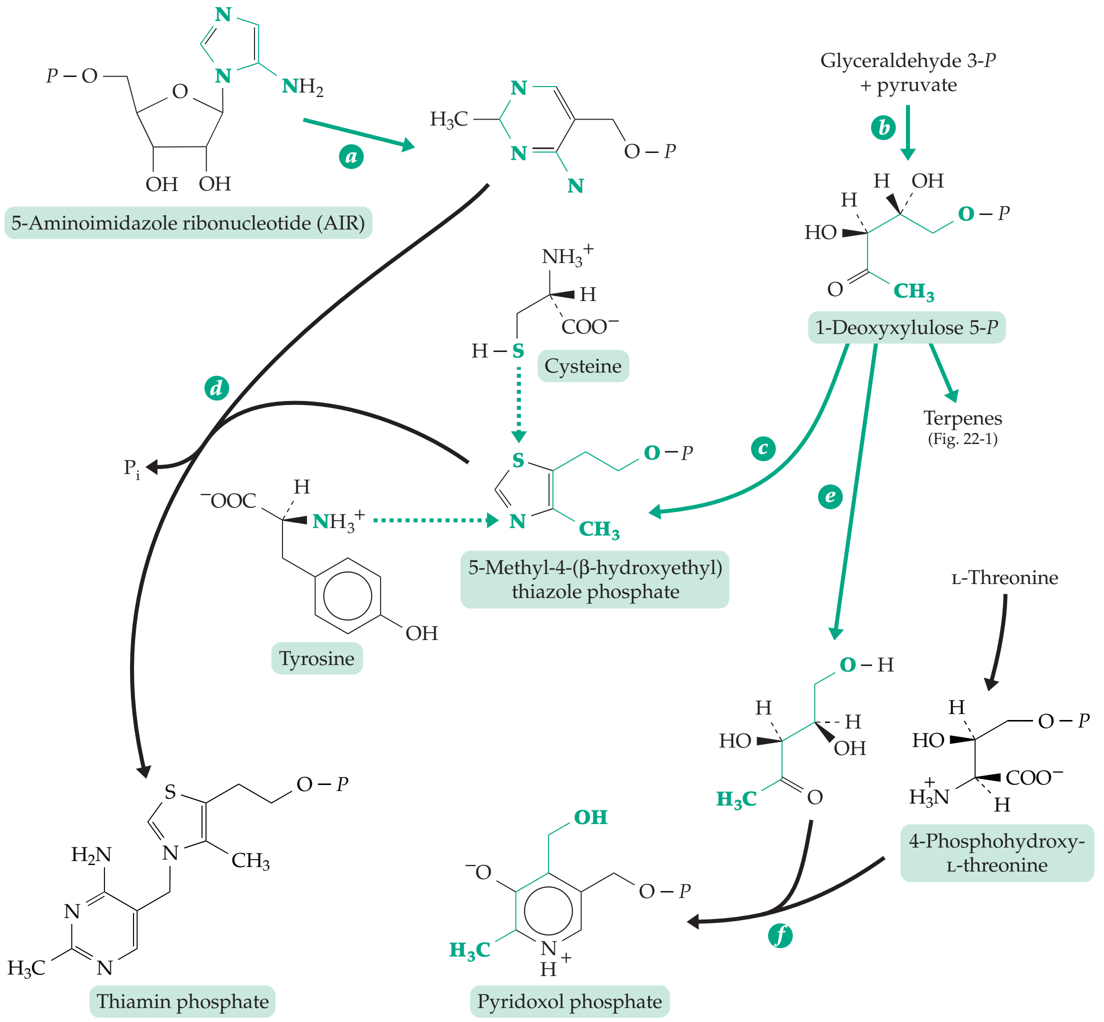
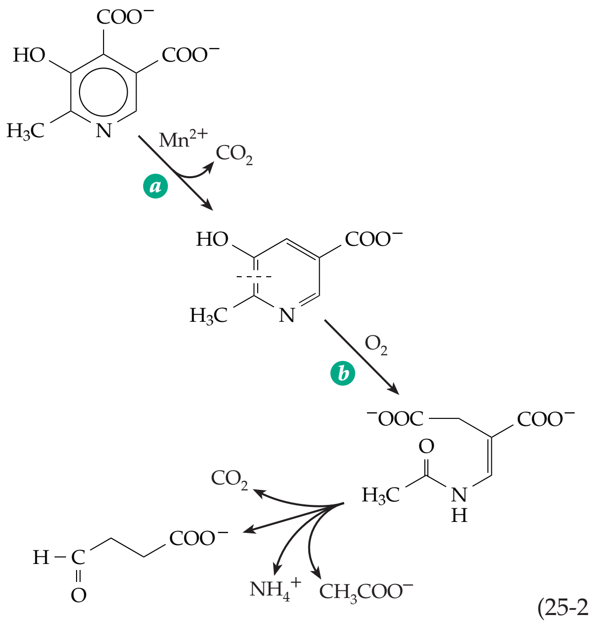

P – O
O
N
N
NH2
OH OH
5-Aminoimidazole ribonucleotide (AIR)
a
H3C
N
N
N
O – P
Glyceraldehyde 3-P
+ pyruvate
b
H OH
H
HO
O – P
O
CH3
1-Deoxyxylulose 5-P
Terpenes
(Fig. 22-1)
NH3+
H
COO−
H – S
Cysteine
d
Pi
S
O – P
N
CH3
c
e
−OOC
H
NH3+
5-Methyl-4-(β-hydroxyethyl)
thiazole phosphate
Tyrosine
OH
L-Threonine
O – H
H
HO
H
OH
H3C
O
H
HO
O – P
COO−
H3N+
H
4-Phosphohydroxy-
L-threonine
S
O – P
N
CH3
H2N
N
H3C
N
Thiamin phosphate
OH
−O
O – P
H3C
N+
H
f
Pyridoxol phosphate

Figure 25-21 Proposed pathways for biosynthesis of thiamin phosphate and pyridoxol phosphate.

way shown in Fig. 25-21 exist in yeast and other fungi.[378b,c] Interconversion of pyridoxal phosphate and other forms of vitamin B_6 is discussed in Chapter 14.

The degradation of pyridoxol by bacteria has been investigated in detail.[379–381] In one pathway the hydroxymethyl group in the 5 position and the substituent in the 4 position are both oxidized in the early steps to carboxylate groups. Then, as indicated in Eq. 25-24, a decarboxylation is followed by the action of an unusual dioxygenase.

Isolated from a strain of *Pseudomonas*, this enzyme contains bound FAD, which must be reduced by external NADH. Like a typical dioxygenase the enzyme introduces two atoms of oxygen into the product. However, it also uses the reduced FAD to reduce the double bond system (either before or after the attack by oxygen).[381] Another enzyme of the same bacterium is remarkable in hydrolyzing the product of the oxygenation reaction to four different products without the accumulation of intermediates.


COO−
HO
COO−
H3C
N
Mn2+
CO2
a
HO
COO−
H3C
N
O2
b
−OOC
COO−
O
H3C
N
H
CO2
H – C
O
COO−
NH4+
CH3COO−

(25-24)

References

1. Floss, H. G. (1986) in *Recent Advances in Phytochemistry*, Vol. 20 (Conn, E. E., ed), pp. 13–55, Plenum, New York
2. Amrhein, N. (1986) in *Recent Advances in Phytochemistry*, Vol. 20 (Conn, E. E., ed), pp. 83–117, Plenum, New York
3. Marzabadi, M. R., Gruys, K. J., Pansegrau, P. D., Walker, M. C., Yuen, H. K., and Sikorski, J. A. (1996) *Biochemistry* **35**, 4199–4210
4. Davis, B. D. (1950) *Experientia* **6**, 4–50
5. Davis, B. D. (1950) *J. Biol. Chem.* **191**, 315–325
6. Weiss, U., and Edwards, J. M. (1980) *The Biosynthesis of Aromatic Compounds*, Wiley, New York
7. Pittard, A. J. (1987) in *Escherichia coli and Salmonella typhimurium*, Vol. I (Neidhardt, F. C., ed), pp. 368–394, Am. Soc. for Microbiology, Washington, DC
8. Gollub, E., Zalkin, H., and Sprinson, D. B. (1967) *J. Biol. Chem.* **242**, 5323–5328
9. Shultz, J., Hermodson, M. A., Garner, C. C., and Herrmann, K. M. (1984) *J. Biol. Chem.* **259**, 9655–9661
10. DeLeo, A. B., Dayan, J., and Sprinson, D. B. (1973) *J. Biol. Chem.* **248**, 2344–2353

10a. Yu, T.-W., Müller, R., Müller, M., Zhang, X., Draeger, G., Kim, C.-G., Leistner, E., and Floss, H. G. (2001) *J. Biol. Chem.* **276**, 12546–12555

10b. Guo, J., and Frost, J. W. (2002) *J. Am. Chem. Soc.* **124**, 10642–10643

10c. Eads, J. C., Beeby, M., Scapin, G., Yu, T.-W., and Floss, H. G. (1999) *Biochemistry* **38**, 9840–9849

10d. Arakawa, K., Müller, R., Mahmud, T., Yu, T-W., and Floss, H. G. (2002) *J. Am. Chem. Soc.* **124**, 10644 – 10645

11. Shultz, J., Hermodson, M. A., Garner, C. C., and Harrmann, K. M. (1984) *J. Biol. Chem.* **259**, 9655–9661
12. Dyer, W. E., Weaver, L. M., Zhao, J., Kuhn, D. N., Weller, S. C., and Harrmann, K. M. (1990) *J. Biol. Chem.* **265**, 1608–1614

12a. Wagner, T., Shumilin, I. A., Bauerle, R., and Kretsinger, R. H. (2000) *J. Mol. Biol.* **301**, 389–399

13. Herrmann, K. M., Schultz, J., and Hermodson, M. A. (1980) *J. Biol. Chem.* **255**, 7079–7081
14. Montchamp, J.-L., and Frost, J. W. (1997) *J. Am. Chem. Soc.* **119**, 7645–7653
15. Carpenter, E. P., Hawkins, A. R., Frost, J. W., and Brown, K. A. (1998) *Nature (London)* **394**, 299–302
16. Maitra, U. S., and Sprinson, D. B. (1978) *J. Biol. Chem.* **253**, 5426–5430
17. Widlanski, T., Bender, S. L., and Knowles, J. R. (1989) *Biochemistry* **28**, 7572–7582
18. Ranjeva, R., Refeno, G., Boudet, A. M., and Marmé, D. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5222–5224
19. Chaudhuri, S., and Coggins, J. R. (1985) *Biochem. J.* **226**, 217–223
20. Huang, L., Montoya, A. L., and Nester, E. W. (1975) *J. Biol. Chem.* **250**, 7675–7681
21. Krell, T., Coggins, J. R., and Lapthorn, A. J. (1998) *J. Mol. Biol.* **278**, 983–997
22. Ramilo, C., Appleyard, R. J., Wanke, C., Krekel, F., Amrhein, N., and Evans, J. N. S. (1994) *Biochemistry* **33**, 15071–15079
23. Kim, D. H., Tucker-Kellogg, G. W., Lees, W. J., and Walsh, C. T. (1996) *Biochemistry* **35**, 5435–5440
24. Lee, J. J., Asano, Y., Shiek, T.-L., Spreafico, F., Lee, K., and Floss, H. G. (1984) *J. Am. Chem. Soc.* **106**, 3367–3368
25. Jakeman, D. L., Mitchell, D. J., Shuttleworth, W. A., and Evans, J. N. S. (1998) *Biochemistry* **37**, 12012–12019
26. Stallings, W. C., Abdel-Meguid, S. S., Lim, L. W., Shieh, H.-S., Dayringer, H. E., Leimgruber, N. K., Stegeman, R. A., Anderson, K. S., Sikorski, J. A., Radgette, S. R., and Kishore, G. M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5046–5050
27. Christensen, A. M., and Schaefer, J. (1993) *Biochemistry* **32**, 2868–2873

27a. Alibhai, M. F., and Stallings, W. C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2944–2946

28. Stalker, D. M., Hiah, W. R., and Comai, L. (1985) *J. Biol. Chem.* **260**, 4724–4728
29. Shuttleworth, W. A., Pohl, M. E., Helms, G. L., Jakeman, D. L., and Evans, J. N. S. (1999) *Biochemistry* **38**, 296–302
30. Morrell, H., Clark, M. J., Knowles, P. F., and Sprinson, D. B. (1967) *J. Biol. Chem.* **242**, 82–90
31. Balasubramanian, S., Coggins, J. R., and Abell, C. (1995) *Biochemistry* **34**, 341–348
32. Macheroux, P., Petersen, J., Bornemann, S., Lowe, D. J., and Thorneley, R. N. F. (1996) *Biochemistry* **35**, 1643–1652
33. Lauhon, C. T., and Bartlett, P. A. (1994) *Biochemistry* **33**, 14100–14108

33a. Osborne, A., Thorneley, R. N. F., Abell, C., and Bornemann, S. (2000) *J. Biol. Chem.* **275**, 35825–35830

33b. Kitzing, K., Macheroux, P., and Amrhein, N. (2001) *J. Biol. Chem.* **276**, 42658–42666

34. Hudson, G. S., Wong, V., and Davidson, B. E. (1984) *Biochemistry* **23**, 6240–6249
35. Hermes, J. D., Tipton, P. A., Fisher, M. A., O'Leary, M. H., Morrison, J. F., and Cleland, W. W. (1984) *Biochemistry* **23**, 6263–6275
36. Christendat, D., Saridakis, V. C., and Turnbull, J. L. (1998) *Biochemistry* **37**, 15703–15712
37. MacBeath, G., Kast, P., and Hilvert, D. (1998) *Biochemistry* **37**, 10062–10073

37a. Guo, H., Cui, Q., Lipscomb, W. N., and Karplus, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9032–9037

38. Sogo, S. G., Widlanski, T. S., Hoare, J. H., Grimshaw, C. E., Berchtold, G. A., and Knowles, J. R. (1984) *J. Am. Chem. Soc.* **106**, 2701–2703

38a. Pohnert, G., Zhang, S., Husain, A., Wilson, D. B., and Ganem, B. (1999) *Biochemistry* **38**, 12212–12217

39. Christendat, D., and Turnbull, J. L. (1999) *Biochemistry* **38**, 4782–4793
40. Christendat, D., and Turnbull, J. (1996) *Biochemistry* **35**, 4468–4479
41. Turnbull, J., Cleland, W. W., and Morrison, J. F. (1990) *Biochemistry* **29**, 10245–10254
42. Turnbull, J., Morrison, J. F., and Cleland, W. W. (1991) *Biochemistry* **30**, 7783–7788
43. Zamir, L. O., Jensen, R. A., Arison, B. H., Douglas, A. W., Albers-Schönberg, G., and Bowen, J. R. (1980) *J. Am. Chem. Soc.* **102**, 4499–4504
44. Zamir, L. O., Tiberio, R., Fiske, M., Berry, A., and Jensen, R. A. (1985) *Biochemistry* **24**, 1607–1612
45. Xia, T., and Jensen, R. A. (1990) *J. Biol. Chem.* **265**, 20033–20036
46. Jung, E., Zamir, L. O., and Jensen, R. A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7231–7235
47. Jensen, R. A. (1986) in *Recent Advances in Phytochemistry*, Vol. 20 (Conn, E. E., ed), pp. 57–81, Plenum, New York
48. Teng, C.-Y. P., and Ganem, B. (1984) *J. Am. Chem. Soc.* **106**, 2463–2464
49. Morollo, A. A., Finn, M. G., and Bauerle, R. (1993) *J. Am. Chem. Soc.* **115**, 816–817
50. Green, J. M., and Nichols, B. P. (1991) *J. Biol. Chem.* **266**, 12971–12975
51. Paluh, J. L., Zalkin, H., Betsch, D., and Weith, H. L. (1985) *J. Biol. Chem.* **260**, 1889–1894
52. Walsh, C. T., Erion, M. D., Walts, A. E., Delany III, J. J., and Berchtold, G. A. (1987) *Biochemistry* **26**, 4734–4745
53. Roux, B., and Walsh, C. T. (1992) *Biochemistry* **31**, 6904–6910

53a. Knöchel, T., Ivens, A., Hester, G., Gonzalez, A., Bauerle, R., Wilmanns, M., Kirschner, K., and Jansonius, J. N. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9479–9484

54. Liu, D. J., and Day, L. A. (1994) *Science* **265**, 671–674
55. Liu, J., Quinn, N., Berchtold, G. A., and Walsh, C. T. (1990) *Biochemistry* **29**, 1417–1425
56. Kozlowski, M. C., and Bartlett, P. A. (1991) *J. Am. Chem. Soc.* **113**, 5897–5898

56a. Wildermuth, M. C., Dewdney, J., Wu, G., and Ausubel, F. M. (2001) *Nature (London)* **414**, 562–565

57. Sakaitani, M., Rusnak, F., Quinn, N. R., Tu, C., Frigo, T. B., Berchtold, G. A., and Walsh, C. T. (1990) *Biochemistry* **29**, 6789–6798
58. Rusnak, F., Liu, J., Quinn, N., Berchtold, G. A., and Walsh, C. T. (1990) *Biochemistry* **29**, 1425–1435
59. Gehring, A. M., Mori, I., and Walsh, C. T. (1998) *Biochemistry* **37**, 2648–2659

59a. May, J. J., Wendrich, T. M., and Marahiel, M. A. (2001) *J. Biol. Chem.* **276**, 7209–7217

60. Schubert, K. R., Switzer, R. L., and Shelton, E. (1975) *J. Biol. Chem.* **250**, 7492–7500
61. Nosal, J. M., Switzer, R. L., and Becker, M. A. (1993) *J. Biol. Chem.* **268**, 10168–10175
62. Zalkin, H., Paluh, J. L., van Cleemput, M., Moye, W. S., and Yanofsky, C. (1984) *J. Biol. Chem.* **259**, 3985–3992
63. Eberhard, M., Tsai-Pflugfelder, M., Bolewska, K., Hommel, U., and Kirschner, K. (1995) *Biochemistry* **34**, 5419–5428
64. Knöchel, T. R., Hennig, M., Merz, A., Darimont, B., Kirschner, K., and Jansonius, J. N. (1996) *J. Mol. Biol.* **262**, 502–515
65. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., and Davies, D. R. (1988) *J. Biol. Chem.* **263**, 17857–17871
66. Schneider, T. R., Gerhardt, E., Lee, M., Liang, P.-H., Anderson, K. S., and Schlichting, I. (1998) *Biochemistry* **37**, 5394–5406
67. Rhee, S., Miles, E. W., and Davies, D. R. (1998) *J. Biol. Chem.* **273**, 8553–8555
68. Hyde, C. C., and Miles, E. W. (1990) *Bio/Technology* **8**, 27–32

68a. Miles, E. W. (2001) *Chem. Record* **1**, 140–151

68b. Weber-Ban, E., Hur, O., Bagwell, C., Banik, U., Yang, L.-H., Miles, E. W., and Dunn, M. F. (2001) *Biochemistry* **40**, 3497–3511

68c. Yamagata, Y., Ogasahara, K., Hioki, Y., Lee, S. J., Nakagawa, A., Nakamura, H., Ishida, M., Kuramitsu, S., and Yutani, K. (2001) *J. Biol. Chem.* **276**, 11062–11071

69. Sterner, R., Dahm, A., Darimont, B., Ivens, A., Liebl, W., and Kirschner, K. (1995) *EMBO J.* **14**, 4395–4402
70. Wilmanns, M., Hyde, C. C., Davies, D. R., Kirschner, K., and Jansonius, J. N. (1991) *Biochemistry* **30**, 9161–9169
71. Zhao, J., and Last, R. L. (1995) *J. Biol. Chem.* **270**, 6081–6087
72. Bentley, R., and Campbell, I. M. (1974) in *The Chemistry of the Quinonoid Compounds* (Patai, S., ed), pp. 683–736, Wiley, New York (Part 2)
73. Sippel, C. J., Goewert, R. R., Slachman, F. N., and Olson, R. E. (1983) *J. Biol. Chem.* **258**, 1057–1061

73a. Jonassen, T., and Clarke, C. F. (2000) *J. Biol. Chem.* **275**, 12381–12387

73b. Stenmark, P., Grünler, J., Mattsson, J., Sindelar, P. J., Nordlund, P., and Berthold, D. A. (2001) *J. Biol. Chem.* **276**, 33297–33300

References

73c. Miyadera, H., Amino, H., Hiraishi, A., Taka, H., Murayama, K., Miyoshi, H., Sakamoto, K., Ishii, N., Hekimi, S., and Kita, K. (2001) *J. Biol. Chem.* **276**, 7713–7716

73d. Hihi, A. K., Gao, Y., and Hekimi, S. (2002) *J. Biol. Chem.* **277**, 2202–2206

74. Marbois, B. N., and Clarke, C. F. (1996) *J. Biol. Chem.* **271**, 2995–3004

75. Hsu, A. Y., Poon, W. W., Shepherd, J. A., Myles, D. C., and Clarke, C. F. (1996) *Biochemistry* **35**, 9797–9806

76. Clarke, C. F., Williams, W., and Teruya, J. H. (1991) *J. Biol. Chem.* **266**, 16636 – 16644

77. Swiezewska, E., Dallner, G., Andersson, B., and Ernster, L. (1993) *J. Biol. Chem.* **268**, 1494–1499

78. Kalén, A., Appelkvist, E.-L., Chojnacki, T., and Dallner, G. (1990) *J. Biol. Chem.* **265**, 1158–1164

79. Folkers, K., Langsjoen, P., Willis, R., Richardson, P., Xia, L.-J., Ye, C.-Q., and Tamagawa, H. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8931–8934

80. Threfall, D. R. (1971) *Vitamins and Hormones* **29**, 153–200

81. Goodwin, T. W., and Merces, E. I. (1986) *Introduction to Plant Biochemistry*, 2nd ed., Pergamon, Oxford

82. Fernández-Cañón, J. M., and Peñalva, M. A. (1995) *J. Biol. Chem.* **270**, 21199–21205

83. Shintani, D., and DellaPenna, D. (1998) *Science* **282**, 2098–2100

83a. Yamamoto, Y., Fujisawa, A., Hara, A., and Dunlap, W. C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13144–13148

84. Bentley, R., and Meganathan, R. (1987) in *Escherichia coli and Salmonella Typhimurium* (Nedihardt, F. C., ed), pp. 512–520, Am. Soc. for Microbiology, Washington, D.C.

85. Koike-Takeshita, A., Koyama, T., and Ogura, K. (1997) *J. Biol. Chem.* **272**, 12380–12383

86. Palmer, D. R. J., Garrett, J. B., Sharma, V., Meganathan, R., Babbitt, P. C., and Gerlt, J. A. (1999) *Biochemistry* **38**, 4252–4258

86a. Stamps, S. L., Taylor, A. B., Wang, S. C., Hackert, M. L., and Whitman, C. P. (2000) *Biochemistry* **39**, 9671–9678

87. Daubner, S. C., and Fitzpatrick, P. F. (1999) *Biochemistry* **38**, 4448–4454

87a. Andersen, O. A., Flatmark, T., and Hough, E. (2001) *J. Mol. Biol.* **314**, 279 – 291

88. Nyhan, W. L. (1984) *Trends Biochem. Sci.* **9**, 71–72

89. Knox, W. E. (1972) in *The Metabolic Basis of Inherited Disease*, 3rd ed. (Standbury, J. B., Wyngaarden, J. B., and Fredrickson, D. S., eds), pp. 266–295, McGraw-Hill, New York

90. Woo, S. L. C. (1989) *Biochemistry* **28**, 1–7

91. Kaufman, S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3160–3164

91a. Levy, H. L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1811–1813

91b. Gámez, A., Pérez, B., Ugarte, M., and Desviat, L. R. (2000) *J. Biol. Chem.* **275**, 29737–29742

92. Ambrus, C. M., Ambrus, J. L., Horvath, C., Pederson, H., Skarma, S., Kant, C., Mirand, E., Guthrie, R., and Paul, T. (1978) *Science* **201**, 837–839

93. Scriver, C. R., Kaufman, S., Eisensmith, R. C., and Woo, S. L. C. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1015–1075, McGraw-Hill, New York

94. Onuffer, J. J., Ton, B. T., Klement, I., and Kirsch, J. F. (1995) *Protein Sci.* **4**, 1743–1749

95. Natt, E., Kida, K., Odievre, M., Di Rocco, M., and Scherer, G. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9297–9301

96. Hargrove, J. L., and Grammer, D. K. (1985) in *Transaminases* (Christen, P., and Metzler, D. E., eds), pp. 511–525, Wiley, New York

97. Wilcken, B., Hammond, J. W., Howard, N., Bohane, T., Hocart, C., and Halpern, B. (1981) *N. Engl. J. Med.* **305**, 865–869

98. La Du, B. N. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1371–1386, McGraw-Hill, New York

99. Fernández-Cañón, J. M., and Peñalva, M. A. (1998) *J. Biol. Chem.* **273**, 329–337

99a. Polekhina, G., Board, P. G., Blackburn, A. C., and Parker, M. W. (2001) *Biochemistry* **40**, 1567–1576

100. Kubo, S., Sun, M., Miyahara, M., Umeyama, K., Urakami, K.-i, Yamamoto, T., Jakobs, C., Matsuda, I., and Endo, F. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9552–9557

101. Frieden, E. (1981) *Trends Biochem. Sci.* **6**, 50–53

102. Tata, J. R. (1990) *Trends Biochem. Sci.* **15**, 282–284

103. Levy, O., De la Vieja, A., Ginter, C. S., Riedel, C., Dai, G., and Carrasco, N. (1998) *J. Biol. Chem.* **273**, 22657–22663

104. Sun, W., and Dunford, H. B. (1993) *Biochemistry* **32**, 1324–1331

105. Surks, M. I., and Sievert, R. (1995) *N. Engl. J. Med.* **333**, 1688–1693

106. Rawtich, A. B., Chernoff, S. B., Litwer, M. R., Rouse, J. B., and Hamilton, J. W. (1983) *J. Biol. Chem.* **258**, 2079–2082

107. Gentile, F., Ferranti, P., Mamone, G., Malorni, A., and Salvatore, G. (1997) *J. Biol. Chem.* **272**, 639–646

108. Gavaret, J.-M., Cahnmann, H. J., and Nunez, J. (1981) *J. Biol. Chem.* **256**, 9167–9173

108a. Ma, Y.-A., Sih, C. J., and Harms, A. (1999) *J. Am. Chem. Soc.* **121**, 8967–8968

109. Shupnik, M. A., Chin, W. W., Ross, D. S., Downing, M. F., Habener, J. F., and Ridgway, E. C. (1983) *J. Biol. Chem.* **258**, 15120–15124

110. Van Heuverswyn, B., Streydio, C., Brocas, H., Refetoff, S., Dumont, J., and Vassart, G. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5941–5945

111. Perlman, J. H., Colson, A.-O., Wang, W., Bence, K., Osman, R., and Gershengorn, M. C. (1997) *J. Biol. Chem.* **272**, 11937–11942

112. Kosugi, S., Okajima, F., Ban, T., Hidaka, A., Shenker, A., and Kohn, L. D. (1992) *J. Biol. Chem.* **267**, 24153–24156

113. Nettleton, E. J., Sunde, M., Lai, Z., Kelly, J. W., Dobson, C. M., and Robinson, C. V. (1998) *J. Mol. Biol.* **281**, 553–564

114. Sebastiao, M. P., Saraiva, M. J., and Damas, A. M. (1998) *J. Biol. Chem.* **273**, 24715–24722

115. Silva, J. E., and Larsen, P. R. (1983) *Nature (London)* **305**, 712–713

116. Pallud, S., Lennon, A.-M., Ramauge, M., Gavaret, J.-M., Croteau, W., Pierre, M., Courtin, F., and St. Germain, D. L. (1997) *J. Biol. Chem.* **272**, 18104–18110

116a. Leonard, J. L., Visser, T. J., and Leonard, D. M. (2001) *J. Biol. Chem.* **276**, 2600–2607

117. Maley, G. F., and Lardy, H. A. (1955) *J. Biol. Chem.* **215**, 377–388

118. Shi, Y.-B., and Brown, D. D. (1993) *J. Biol. Chem.* **268**, 20312–20317

119. Brown, D. D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13011–13016

120. Kumara-Siri, M. H., Shapiro, L. E., and Surks, M. I. (1986) *J. Biol. Chem.* **261**, 2844–2852

121. Apriletti, J. W., David-Inouye, Y., Eberhardt, N. L., and Baxter, J. D. (1984) *J. Biol. Chem.* **259**, 10941–10948

122. Brent, G. A. (1994) *N. Engl. J. Med.* **331**, 847–853

123. Wagner, R. L., Apriletti, J. W., McGrath, M. E., West, B. L., Baxter, J. D., and Fletterick, R. J. (1995) *Nature (London)* **378**, 690–697

123a. Wu, Y., Xu, B., and Koenig, R. J. (2001) *J. Biol. Chem.* **276**, 3929–3936

123b. Zhang, Y., Yin, L., and Hillgartner, F. B. (2001) *J. Biol. Chem.* **276**, 974–983

124. Hörlein, A. J., Näär, A. M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Söderström, M., Glass, C. K., and Rosenfeld, M. G. (1995) *Nature (London)* **377**, 397–404

125. Chen, Y., Chen, P.-L., Chen, C.-F., Sharp, Z. D., and Lee, W.-H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4443–4448

125a. Zhu, X.-G., Hanover, J. A., Hager, G. L., and Cheng, S. (1998) *J. Biol. Chem.* **273**, 27058–27063

125b. Tagami, T., Park, Y., and Jameson, J. L. (1999) *J. Biol. Chem.* **274**, 22345–22353

125c. Riedel, C., Dohán, O., De la Vieja, A., Ginter, C. S., and Carrasco, N. (2001) *Trends Biochem. Sci.* **26**, 490–496

126. Shaw-White, J. R., Bruno, M. D., and Whitsett, J. A. (1999) *J. Biol. Chem.* **274**, 2658–2664

127. Vassart, G., Dumont, J. E., and Refetoff, S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 2883–2928, McGraw-Hill, New York

128. Collingwood, T. N., and and 19 other authors. (1998) *EMBO J.* **17**, 4760–4770

129. Hetzel, B. S. (1994) *N. Engl. J. Med.* **331**, 1770–1771

130. Hearn, M. T. W. (1980) *Trends Biochem. Sci.* **5**, 75–79

131. Utiger, R. D. (1991) *N. Engl. J. Med.* **325**, 278–279

131a. Bertoldi, M., Frigeri, P., Paci, M., and Voltattorni, C. B. (1999) *J. Biol. Chem.* **274**, 5514 – 5521

132. Potterf, S. B., Muller, J., Bernardini, I., Tietze, F., Kobayashi, T., Hearing, V. J., and Gahl, W. A. (1996) *J. Biol. Chem.* **271**, 4002–4008

133. Jiménez-Cervantes, C., Solano, F., Kobayashi, T., Urabe, K., Hearing, V. J., Lozano, J. A., and García-Borrón, J. C. (1994) *J. Biol. Chem.* **269**, 17993–18001

134. Leonard, L. J., Townsend, D., and King, R. A. (1988) *Biochemistry* **27**, 6156–6159

135. Sugimoto, H., Taniguchi, M., Nakagawa, A., Tanaka, I., Suzuki, M., and Nishihira, J. (1999) *Biochemistry* **38**, 3268–3279

136. Rodríguez-López, J. N., Tudela, J., Varón, R., García-Carmona, F., and García-Cánovas, F. (1992) *J. Biol. Chem.* **267**, 3801–3810

137. Thomson, R. H. (1974) *Angew. Chem. Int. Ed. Engl.* **13**, 305–312

138. Deibel, R. M. B., and Chedekel, M. R. (1984) *J. Am. Chem. Soc.* **106**, 5884–5888

139. Pawelek, J., Körner, A., Bergstrom, A., and Bologna, J. (1980) *Nature (London)* **286**, 617–619

140. Chen, J. M., Xu, S. L., Wawrzak, Z., Basarab, G. S., and Jordan, D. B. (1998) *Biochemistry* **37**, 17735–17744

140a. Liao, D.-I., Thompson, J. E., Fahnestock, S., Valent, B., and Jordan, D. B. (2001) *Biochemistry* **40**, 8696–8704

141. Saul, S. J., and Sugumaran, M. (1990) *J. Biol. Chem.* **265**, 16992–16999

141a. Hall, M., Scott, T., Sugumaran, M., Söderhäll, K., and Law, J. H. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7764–7768

141b. Asano, T., and Ashida, M. (2001) *J. Biol. Chem.* **276**, 11100–11112

142. Sugumaran, M., Semensi, V., Kalyanaraman, B., Bruce, J. M., and Land, E. J. (1992) *J. Biol. Chem.* **267**, 10355–10361

References

143. McCormick, M. L., Gaut, J. P., Lin, T.-S., Britigan, B. E., Buettner, G. R., and Heinecke, J. W. (1998) *J. Biol. Chem.* **273**, 32030–32037

144. Malencik, D. A., and Anderson, S. R. (1994) *Biochemistry* **33**, 13363–13372

144a. Kanwar, R., and Balasubramanian, D. (2000) *Biochemistry* **39**, 14976–14983

145. Briza, P., Winkler, G., Kalchhauser, H., and Breitenbach, M. (1986) *J. Biol. Chem.* **261**, 4288–4294

146. Briza, P., Ellinger, A., Winkler, G., and Breitenbach, M. (1990) *J. Biol. Chem.* **265**, 15118–15123

147. Taylor, S. W., Waite, J. H., Ross, M. M., Shabanowitz, J., and Hunt, D. F. (1994) *J. Am. Chem. Soc.* **116**, 10803–10804

148. Waite, J. H., Jensen, R. A., and Morse, D. E. (1992) *Biochemistry* **31**, 5733–5738

149. Papov, V. V., Diamond, T. V., Biemann, K., and Waite, J. H. (1995) *J. Biol. Chem.* **270**, 20183–20192

149a. Yu, M., Hwang, J., and Deming, T. (1999) *J. Am. Chem. Soc.* **121**, 5825–5826

150. Dagley, S. (1978) *Naturwissenschaften* **65**, 85–95

151. Allewell, N. (1989) *Trends Biochem. Sci.* **14**, 473–474

152. Dagley, S., and Nicholson, D. E. (1970) *An Introduction to Metabolic Pathways*, Wiley, New York

153. Gibson, D. T., ed. (1984) *Microbial Degradation of Organic Compounds*, Dekker, New York

154. Nakai, C., Kagamiyama, H., Nozaki, M., Nakazawa, T., Inouye, S., Ebina, Y., and Nakazawa, A. (1983) *J. Biol. Chem.* **258**, 2923–2928

155. Correll, C. C., Batie, C. J., Ballou, D. P., and Ludwig, M. L. (1985) *J. Biol. Chem.* **260**, 14633–14635

155a. Palumbo, A., Poli, A., Cosmo, A. D., and d'Ischia, M. (2000) *J. Biol. Chem.* **275**, 16885–16890

156. Blasco, R., Wittich, R.-M., Mallavarapu, M., Timmis, K. N., and Pieper, D. H. (1995) *J. Biol. Chem.* **270**, 29229–29235

157. Doelle, H. W. (1975) *Bacterial Metabolism*, Academic Press, New York

158. Werlen, C., Kohler, H.-P. E., and van der Meer, J. R. (1996) *J. Biol. Chem.* **271**, 4009–4016

158a. Copley, S. D. (2000) *Trends Biochem. Sci.* **25**, 261–265

158b. Anandarajah, K., Kiefer, P. M., Donohoe, B. S., and Copley, S. D. (2000) *Biochemistry* **39**, 5303–5311

158c. Tame, J. R. H., Namba, K., Dodson, E. J., and Roper, D. I. (2002) *Biochemistry* **41**, 2982–2989

159. Seah, S. Y. K., Terracina, G., Bolin, J. T., Riebel, P., Snieckus, V., and Eltis, L. D. (1998) *J. Biol. Chem.* **273**, 22943–22949

160. Scholten, J. D., Chang, K.-H., Babbitt, P. C., Charest, H., Sylvestre, M., and Dunaway-Mariano, D. (1991) *Science* **253**, 182–185

160a. Zhang, W., Wei, Y., Luo, L., Taylor, K. L., Yang, G., and Dunaway-Mariano, D. (2001) *Biochemistry* **40**, 13474–13482

161. Harpel, M. R., and Lipscomb, J. D. (1990) *J. Biol. Chem.* **265**, 22187–22196

161a. Olivera, E. R., Minambres, B., Garcia, B., Muniz, C., Moreno, M. A., Ferrández, A., Díaz, E., García, J. L., and Luengo, J. M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6419–6424

161b. Zaar, A., Eisenreich, W., Bacher, A., and Fuchs, G. (2001) *J. Biol. Chem.* **276**, 24997–25004

162. Yamaguchi, M., and Fujisawa, H. (1982) *J. Biol. Chem.* **257**, 12497–12502

163. Zylstra, G. J., and Gibson, D. T. (1989) *J. Biol. Chem.* **264**, 14940–14946

164. Díaz, E., and Timmis, K. N. (1995) *J. Biol. Chem.* **270**, 6403–6411

165. Petsko, G. A., Kenyon, G. L., Gerlt, J. A., Ringe, D., and Kozarich, J. W. (1993) *Trends Biochem. Sci.* **18**, 372–376

166. Harwood, C. S., and Gibson, J. (1997) *J. Bacteriol.* **179**, 301–309

166a. Boll, M., Laempe, D., Eisenreich, W., Bacher, A., Mittelberger, T., Heinze, J., and Fuchs, G. (2000) *J. Biol. Chem.* **275**, 21889–21895

166b. Boll, M., Fuchs, G., and Lowe, D. J. (2001) *Biochemistry* **40**, 7612–7620

167. Härtel, U., Eckel, E., Koch, J., Fuchs, G., Linder, D., and Buckel, W. (1993) *Arch Microbiol* **159**, 174–181

168. Draths, K. M., Ward, T. L., and Frost, J. W. (1992) *J. Am. Chem. Soc.* **114**, 9725–9726

169. Richman, J. E., Chang, Y.-C., Kambourakis, S., Draths, K. M., Almy, E., Snell, K. D., Strasburg, G. M., and Frost, J. W. (1996) *J. Am. Chem. Soc.* **118**, 11587–11591

170. Draths, K. M., and Frost, J. W. (1991) *J. Am. Chem. Soc.* **113**, 9361–9363

171. Haslam, E. (1986) in *Recent Advances in Phytochemistry*, Vol. 20 (Conn, E. E., ed), pp. 163–200, Plenum, New York

172. Werner, I., Bacher, A., and Eisenreich, W. (1997) *J. Biol. Chem.* **272**, 25474–25482

173. Lauvergeat, V., Kennedy, K., Feuillet, C., McKie, J. H., Gorrichon, L., Baltas, M., Boudet, A. M., Grima-Pettenati, J., and Douglas, K. T. (1995) *Biochemistry* **34**, 12426–12434

174. Ellis, B. E., Kuroki, G. W., and Stafford, H. A., eds. (1994) *Genetic Engineering of Plant Secondary Metabolism*, Plenum, New York

174a. Crouteau, R., Kutchan, T. M., and Lewis, N. G. (2000) in *Biochemistry and Molecular Biology of Plants* (Buchanan, B., Gruissem, W., and Jones, R., eds), pp. 1250–1318, Am. Soc. Plant Physiologists, Rockville, Maryland

175. Schwede, T. F., Rétey, J., and Schulz, G. E. (1999) *Biochemistry* **38**, 5355–5361

175a. Stuible, H.-P., and Kombrink, E. (2001) *J. Biol. Chem.* **276**, 26893–26897

176. Becker-André, M., Schulze-Lefert, P., and Hahlbrock, K. (1991) *J. Biol. Chem.* **266**, 8551–8559

177. León, J., Shulaev, V., Yalpani, N., Lawton, M. A., and Raskin, I. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10413–10417

178. Lee, H.-i, León, J., and Raskin, I. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4076–4079

179. Delaney, T. P., Uknes, S., Vernooij, B., Friedrich, L., Weymann, K., Negrotto, D., Gaffney, T., Gut-Rella, M., Kessmann, H., Ward, E., and Ryals, J. (1994) *Science* **266**, 1247–1250

180. Chen, Z., Malamy, J., Henning, J., Conrath, U., Sánchez-Casas, P., Silva, H., Ricigliano, J., and Klessig, D. F. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4134–4137

181. Durner, J., and Klessig, D. F. (1996) *J. Biol. Chem.* **271**, 28492–28501

182. Durner, J., and Klessig, D. F. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11312–11316

183. Jurivich, D. A., Pachetti, C., Qiu, L., and Welk, J. F. (1995) *J. Biol. Chem.* **270**, 24489–24495

184. Prince, R. C., and Gunson, D. E. (1994) *Trends Biochem. Sci.* **19**, 521

185. Li, K., and Frost, J. W. (1998) *J. Am. Chem. Soc.* **120**, 10545–10546

186. Rosen, M. A., Shapiro, L., and Patel, D. J. (1992) *Biochemistry* **31**, 4015–4026

187. Villegas, R. J. A., and Kojima, M. (1986) *J. Biol. Chem.* **261**, 8729–8733

188. von Röpenack, E., Parr, A., and Schulze-Lefert, P. (1998) *J. Biol. Chem.* **273**, 9013–9022

189. Balint, R., Cooper, G., Staebell, M., and Filner, P. (1987) *J. Biol. Chem.* **262**, 11026–11031

190. Crawford, R. L. (1981) *Lignin Biodegradation and Transformation*, Wiley, New York

190a. Humphreys, J. M., Hemm, M. R., and Chapple, C. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10045–10050

190b. Li, L., Popko, J. L., Umezawa, T., and Chiang, V. L. (2000) *J. Biol. Chem.* **275**, 6537–6545

190c. Lim, E.-K., Li, Y., Parr, A., Jackson, R., Ashford, D. A., and Bowles, D. J. (2001) *J. Biol. Chem.* **276**, 4344–4349

191. Fujita, M., Gang, D. R., Davin, L. B., and Lewis, N. G. (1999) *J. Biol. Chem.* **274**, 618–627

191a. Xia, Z.-Q., Costa, M. A., Pélissier, H. C., Davin, L. B., and Lewis, N. G. (2001) *J. Biol. Chem.* **276**, 12614–12623

192. Lewis, N. G., and Davin, L. B. (1994) in *Isopentenoids and Other Natural Products* (Nes, W. D., ed), pp. 202–246, American Chemical Society, Washington D.C.

193. von Wettstein-Knowles, P. M. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 127–166, CRC Press, Boca Raton, Florida

194. Nimz, H. (1974) *Angew. Chem. Int. Ed. Engl.* **13**, 313–321

195. Ralph, J., Hatfield, R. D., Quideau, S., Helm, R. F., Grabber, J. H., and Jung, H.-J. G. (1994) *J. Am. Chem. Soc.* **116**, 9448–9456

196. Steelink, C. (1972) *Rec. Adv. Phytochem.* **4**, 239–271

197. Huang, K.-x, Fujii, I., Ebizuka, Y., Gomi, K., and Sankawa, U. (1995) *J. Biol. Chem.* **270**, 21495–21502

197a. Henriksen, A., Smith, A. T., and Gajhede, M. (1999) *J. Biol. Chem.* **274**, 35005–35011

198. Nakamura, N., Matsuzaki, R., Choi, Y.-H., Tanizawa, K., and Sanders-Loehr, J. (1996) *J. Biol. Chem.* **271**, 4718–4724

198a. Schwartz, B., Olgin, A. K., and Klinman, J. P. (2001) *Biochemistry* **40**, 2954–2963

198b. Klinman, J. P. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14766–14768

198c. Datta, S., Mori, Y., Takagi, K., Kawaguchi, K., Chen, Z.-W., Okajima, T., Kuroda, S., Ikeda, T., Kano, K., Tanizawa, K., and Mathews, F. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14268–14273

198d. Firbank, S. J., Rogers, M. S., Wilmot, C. M., Dooley, D. M., Halcrow, M. A., Knowles, P. F., McPherson, M. J., and Phillips, S. E. V. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12932–12937

198e. Xie, L., and van der Donk, W. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12863–12865

199. Roberts, M. F., and Wink, M., eds. (1998) *Alkaloids Biochemistry, Ecology, and Medicinal Applications*, Plenum, New York

200. Harborne, J. B. (1993) *Introduction to Ecological Biochemistry*, 4th ed., Academic Press, San Diego, California

201. Nakajima, K., Yamashita, A., Akama, H., Nakatsu, T., Kato, H., Hashimoto, T., Oda, J., and Yamada, Y. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4876–4881

202. Robinson, R. (1917) *J. Chem. Soc.* **111**, 876–899

203. Robinson, R. (1955) *The Structural Relations of Natural Products*, Oxford Univ. Press, London

204. Lenz, R., and Zenk, M. H. (1995) *J. Biol. Chem.* **270**, 31091–31096

204a. Grothe, T., Lenz, R., and Kutchan, T. M. (2001) *J. Biol. Chem.* **276**, 30717–30723

205. Kutchan, T. M., and Dittrich, H. (1995) *J. Biol. Chem.* **270**, 24475–24481

206. Hayaishi, O. (1993) *Protein Sci.* **2**, 472–475

207. Young, S. N., Oravec, M., and Sourkes, T. L. (1974) *J. Biol. Chem.* **249**, 3932–3936

208. Sono, M. (1989) *Biochemistry* **28**, 5400–5407

209. Taylor, M. W., and Feng, G. (1991) *FASEB J.* **5**, 2516–2522

References

209a. Phillips, R. S., Johnson, N., and Kamath, A. V. (2002) *Biochemistry* **41**, 4012 – 4019

210. Billker, O., Lindo, V., Panico, M., Etienne, A. E., Paxton, T., Dell, A., Rogers, M., Sinden, R. E., and Morris, H. R. (1998) *Nature (London)* **392**, 289 – 292

211. Garcia, G. E., Wirtz, R. A., Barr, J. R., Woolfitt, A., and Rosenberg, R. (1998) *J. Biol. Chem.* **273**, 12003 – 12005

212. Kayser, H. (1985) in *Comprehensive Insect Physiology, Biochemistry and Pharmacology*, Vol. 10 (Kerkut, G. A., and Gilbert, L. I., eds), pp. 368 – 416, Pergamon, Oxford

213. Stone, T. W., ed. (1988) *Quinolinic Acid and the Kynurenines*, CRC Press, Boca Raton, Florida

213a. Aquilina, J. A., Carver, J. A., and Truscott, R. J. W. (1999) *Biochemistry* **38**, 11455 – 11464

213b. Wang, L., Erlandsen, H., Haavik, J., Knappskog, P. M., and Stevens, R. C. (2002) *Biochemistry* **41**, 12569 – 12574

214. Manthey, M. K., Pyne, S. G., and Truscott, R. J. W. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1954 – 1957

215. Schmidt, A. W., and Peroutka, S. J. (1989) *FASEB J.* **3**, 2242 – 2249

215a. McKinney, J., Teigen, K., Froystein, N. Å., Salaün, C., Knappskog, P. M., Haavik, J., and Martínez, A. (2001) *Biochemistry* **40**, 15591 – 15601

216. Hajduch, E., Rencurel, F., Balendran, A., Batty, I. H., Downes, C. P., and Hundal, H. S. (1999) *J. Biol. Chem.* **274**, 13563 – 13568

217. Klein, D. C., and Namboodiri, M. A. A. (1982) *Trends Biochem. Sci.* **7**, 98 – 102

218. Utiger, R. D. (1992) *N. Engl. J. Med.* **327**, 1377 – 1379

219. Ishida, I., Obinata, M., and Deguchi, T. (1987) *J. Biol. Chem.* **262**, 2895 – 2899

220. Rodriguez, I. R., Mazuruk, K., Schoen, T. J., and Chader, G. J. (1994) *J. Biol. Chem.* **269**, 31969 – 31977

221. Slominski, A., Baker, J., Rosano, T. G., Guisti, L. W., Ermak, G., Grande, M., and Gaudet, S. J. (1996) *J. Biol. Chem.* **271**, 12281 – 12286

222. Mayeno, A. N., Lin, F., Foote, C. S., Loegering, D. A., Ames, M. M., Hedberg, C. W., and Gleich, G. J. (1990) *Science* **250**, 1707 – 1708

223. Aldhous, P. (1991) *Nature (London)* **353**, 490

224. Moat, A. G., and Foster, J. W. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical and Medical Aspects*, Vol. B (Dolphin, D., Avramovic, O., and Poulson, R., eds), pp. 1 – 24, Wiley (Interscience), New York

225. Keys, L. D., III, and Hamilton, G. A. (1987) *J. Am. Chem. Soc.* **109**, 2156 – 2163

225a. Saridakis, V., Christendat, D., Kimber, M. S., Dharamsi, A., Edwards, A. M., and Pai, E. F. (2001) *J. Biol. Chem.* **276**, 7225 – 7232

226. Rizzi, M., Nessi, C., Mattevi, A., Coda, A., Bolognesi, M., and Galizzi, A. (1996) *EMBO J.* **15**, 5125 – 5134

227. Merkler, D. J., Glufke, U., Ritenour-Rodgers, K. J., Baumgart, L. E., DeBlassio, J. L., Merkler, K. A., and Vederas, J. C. (1999) *J. Am. Chem. Soc.* **121**, 4904 – 4905

228. Tedeschi, G., Zetta, L., Negri, A., Mortarino, M., Ceciliani, F., and Ronchi, S. (1997) *Biochemistry* **36**, 16221 – 16230

229. Nasu, S., Wicks, F. D., and Gholson, R. K. (1982) *J. Biol. Chem.* **257**, 626 – 632

230. White, H. B., III. (1982) in *The Pyridine Nucleotide Coenzymes* (Everse, J., Anderson, B., and You, K.-S., eds), pp. 225 – 242, Academic Press, New York

231. Tritz, G. J. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 557 – 563, Am. Soc. for Microbiology, Washington,DC

232. Malherbe, P., Köhler, C., Da Prada, M., Lang, G., Kiefer, V., Schwarcz, R., Lahm, H.-W., and Cesura, A. M. (1994) *J. Biol. Chem.* **269**, 13792 – 13797

233. Heyes, M. P., Saito, K., Lackner, A., Wiley, C. A., Achim, C. L., and Markey, S. P. (1998) *FASEB J.* **12**, 881 – 896

234. Emanuele, J. J., and Pitzpatrick, P. F. (1995) *Biochemistry* **34**, 3716 – 3723

234a. Gadda, G., Dangott, L. J., Johnson, W. H., Jr., Whitman, C. P., and Fitzpatrick, P. F. (1999) *Biochemistry* **38**, 5822 – 5828

234b. Zhao, Y., Christensen, S. K., Fankhauser, C., Cashman, J. R., Cohen, J. D., Weigel, D., and Chory, J. (2001) *Science* **291**, 306 – 309

235. Brown, H. M., and Purves, W. K. (1976) *J. Biol. Chem.* **251**, 907 – 913

236. Normanly, J., Cohen, J. D., and Fink, G. R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10355 – 10359

236a. Mikkelsen, M. D., Hansen, C. H., Wittstock, U., and Halkier, B. A. (2000) *J. Biol. Chem.* **275**, 33712 – 33717

236b. Kato, Y., Nakamura, K., Sakiyama, H., Mayhew, S. G., and Asano, Y. (2000) *Biochemistry* **39**, 800 – 809

237. Kobayashi, M., Suzuki, T., Fujita, T., Masuda, M., and Shimizu, S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 714 – 718

238. Bartling, D., Seedorf, M., Schmidt, R. C., and Weiler, E. W. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6021 – 6025

238a. Kendrew, S. G. (2001) *Trends Biochem. Sci.* **26**, 218

238b. Jones, A. M. (1998) *Science* **282**, 2201 – 2202

239. Bartel, B., and Fink, G. R. (1995) *Science* **268**, 1745 – 1748

240. Szerszen, J. B., Szczyglowski, K., and Bandurski, R. S. (1994) *Science* **265**, 1699 – 1701

241. Roberto, F. F., Klee, H., White, F., Nordeen, R., and Kosuge, T. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5797 – 5801

241a. Crozier, A., Kamiya, Y., Bishop, G., and Yokota, T. (2000) in *Biochemistry and Molecular Biology of Plants* (Buchanan, B., Gruissem, W., and Jones, R., eds), pp. 850 – 929, Am. Soc. Plant Physiologists, Rockville, Maryland

242. Abel, S., Nguyen, M. D., and Theologis, A. (1995) *J. Mol. Biol.* **251**, 533 – 549

243. Abel, S., Nguyen, M. D., Chow, W., and Theologis, A. (1995) *J. Biol. Chem.* **270**, 19093 – 19099

244. Shibuya, M., Chou, H.-M., Fountoulakis, M., Hassam, S., Kim, S.-U., Kobayashi, K., Otsuka, H., Rogalska, E., Cassady, J. M., and Floss, H. G. (1990) *J. Am. Chem. Soc.* **112**, 297 – 304

245. Mizukami, H., Nordlöv, H., Lee, S.-L., and Scott, A. I. (1979) *Biochemistry* **18**, 3760 – 3763

245a. Stocking, E. M., Williams, R. M., and Sanz-Cervera, J. F. (2000) *J. Am. Chem. Soc.* **122**, 9089 – 9098

246. Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grün, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R. B., Briggs, S. P., Simcox, K., and Gierl, A. (1997) *Science* **277**, 696 – 699

246a. Beismann-Driemeyer, S., and Sterner, R. (2001) *J. Biol. Chem.* **276**, 20387 – 20396

247. Parker, A. R., Moore, J. A., Schwab, J. M., and Davisson, V. J. (1995) *J. Am. Chem. Soc.* **117**, 10605 – 10613

247a. Sivaraman, J., Li, Y., Larocque, R., Schrag, J. D., Cygler, M., and Mattte, A. (2001) *J. Mol. Biol.* **311**, 761 – 776

247b. Haruyama, K., Nakai, T., Miyahara, I., Hirotsu, K., Mizuguchi, H., Hayashi, H., and Kagamiyama, H. (2001) *Biochemistry* **40**, 4633 – 4644

248. Reizer, J., Michotey, V., Reizer, A., and Saier, M. H., Jr. (1994) *Protein Sci.* **3**, 440 – 450

249. Nagai, A., Ward, E., Beck, J., Tada, S., Chang, J.-Y., Scheidegger, A., and Ryals, J. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4133 – 4137

249a. Grubmeyer, C., and Teng, H. (1999) *Biochemistry* **38**, 7355 – 7362

250. Brenner, M., and Ames, B. N. (1971) in *Metabolic Pathways*, 3rd ed. (Vogel, H. J., ed), Academic Press, New York

251. Winkler, M. E. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 395 – 411, Am. Soc. for Microbiology, Washington,DC

252. Levy, H. L., Taylor, R. G., and McInnes, R. R. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1107 – 1123, McGraw-Hill, New York

253. Chen, J.-Y. C., and Bodley, J. W. (1988) *J. Biol. Chem.* **263**, 11692 – 11696

254. Jones, M. E. (1980) *Ann. Rev. Biochem.* **49**, 253 – 279

255. Gao, G., Nara, T., Nakajima-Shimada, J., and Aoki, T. (1999) *J. Mol. Biol.* **285**, 149 – 161

256. Serre, V., Guy, H., Liu, X., Penverne, B., Hervé, G., and Evans, D. (1998) *J. Mol. Biol.* **281**, 363 – 377

257. Neuhard, J., and Nygaard, P. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 446 – 473, Am. Soc. for Microbiology, Washington,DC

258. Switzer, R. L., and Quinn, C. L. (1993) in *Bacillus subtilis and Other Gram-Positive Bacteria* (Sonenshein, A. L., Hoch, J. A., and Losick, R., eds), pp. 343 – 358, American Society for Microbiology, Washington, D. C.

259. Guy, H. I., and Evans, D. R. (1996) *J. Biol. Chem.* **271**, 13762 – 13769

259a. Thoden, J. B., Raushel, F. M., Wesenberg, G., and Holden, H. M. (1999) *J. Biol. Chem.* **274**, 22502 – 22507

260. Braxton, B. L., Mullins, L. S., Rauschel, F. M., and Reinhart, G. D. (1999) *Biochemistry* **38**, 1394 – 1401

260a. Thoden, J. B., Phillips, G. N., Jr., Neal, T. M., Raushel, F. M., and Holden, H. M. (2001) *Biochemistry* **40**, 6989 – 6997

261. Williams, N. K., Manthey, M. K., Hambley, T. W., O'Donoghue, S. I., Keegan, M., Chapman, B. E., and Christopherson, R. I. (1995) *Biochemistry* **34**, 11344 – 11352

262. Zimmermann, B. H., Kemling, N. M., and Evans, D. R. (1995) *Biochemistry* **34**, 7038 – 7046

263. Hines, V., and Johnston, M. (1989) *Biochemistry* **28**, 1227 – 1234

264. Björnberg, O., Rowland, P., Larsen, S., and Jensen, K. F. (1997) *Biochemistry* **36**, 16197 – 16205

265. Rowland, P., Björnberg, O., Nielsen, F. S., Jensen, K. F., and Larsen, S. (1998) *Protein Sci.* **7**, 1269 – 1279

265a. Argyrou, A., Washabaugh, M. W., and Pickart, C. M. (2000) *Biochemistry* **39**, 10373 – 10384

266. Scapin, G., Ozturk, D. H., Grubmeyer, C., and Sacchettini, J. C. (1995) *Biochemistry* **34**, 10744 – 10754

267. Tao, W., Grubmeyer, C., and Blanchard, J. S. (1996) *Biochemistry* **35**, 14 – 21

268. Radzicka, A., and Wolfenden, R. (1995) *Science* **267**, 90 – 93

269. Beak, P., and Siegel, B. (1976) *J. Am. Chem. Soc.* **98**, 3601 – 3606

270. Smiley, J. A., and Jones, M. E. (1992) *Biochemistry* **31**, 12162 – 12168

270a. Appleby, T. C., Kinsland, C., Begley, T. P., and Ealick, S. E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2005 – 2010

270b. Harris, P., Poulsen, J.-C. N., Jensen, K. F., and Larsen, S. (2000) *Biochemistry* **39**, 4217–4224

270c. Miller, B. G., Snider, M. J., Short, S. A., and Wolfenden, R. (2000) *Biochemistry* **39**, 8113–8118

270d. Phillips, L. M., and Lee, J. K. (2001) *J. Am. Chem. Soc.* **123**, 12067–12073

271. Girot, R., Hamet, M., Perignon, J.-L., Guesnu, M., Fox, R. M., Cartier, P., Durandy, A., and Griscelli, C. (1983) *N. Engl. J. Med.* **308**, 700–704

272. Lewis, D. A., and Villafrance, J. J. (1989) *Biochemistry* **28**, 8454–8459

273. Kornberg, A., and Baker, T. A. (1992) *DNA Replication*, 2nd ed., Freeman, New York

274. Weiner, K. X. B., Ciesla, J., Jaffe, A. B., Ketring, R., Maley, F., and Maley, G. F. (1995) *J. Biol. Chem.* **270**, 18727–18729

274a. Kambampati, R., and Lauhon, C. T. (2000) *J. Biol. Chem.* **275**, 10727–10730

274b. Johansson, E., Mejlhede, N., Neuhard, J., and Larsen, S. (2002) *Biochemistry* **41**, 2563–2570

274c. Snider, M. J., Reinhardt, L., Wolfenden, R., and Cleland, W. W. (2002) *Biochemistry* **41**, 415–421

275. Kvalnes-Krick, K. L., and Traut, T. W. (1993) *J. Biol. Chem.* **268**, 5686–5693

275a. Dobritzsch, D., Schneider, G., Schnackerz, K. D., and Lindqvist, Y. (2001) *EMBO J.* **20**, 650–660

276. Kim, Y. S., and Bang, S. K. (1985) *J. Biol. Chem.* **260**, 5098–5104

277. Nygaard, P. (1993) in *Bacillus subtilis and Other Gram-Positive Bacteria* (Sonenshein, A. L., Hoch, J. A., and Losick, R., eds), pp. 359–378, American Society for Microbiology, Washington, D. C.

278. Iovannisci, D. M., Goebel, D., Allen, K., Kaur, K., and Ullman, B. (1984) *J. Biol. Chem.* **259**, 14617–14623

279. Schumacher, M. A., Carter, D., Scott, D. M., Roos, D. S., Ullman, B., and Brennan, R. G. (1998) *EMBO J.* **17**, 3219–3232

280. Muchmore, C. R. A., Krahn, J. M., Kim, J. H., Zalkin, H., and Smith, J. L. (1998) *Protein Sci.* **7**, 39–51

281. Smith, J. L., Zaluzec, E. J., Wery, J.-P., Niu, L., Switzer, R. L., Zalkin, H., and Satow, Y. (1994) *Science* **264**, 1427–1433

282. Wang, W., Kappock, T. J., Stubbe, J., and Ealick, S. E. (1998) *Biochemistry* **37**, 15647–15662

283. Benkovic, S. J. (1984) *Trends Biochem. Sci.* **9**, 320–322

284. Ononogbu, I. C. (1980) *Trends Biochem. Sci.* **5**, X

285. Shim, J. H., and Benkovic, S. J. (1998) *Biochemistry* **37**, 8776–8782

285a. Shim, J. H., and Benkovic, S. J. (1999) *Biochemistry* **38**, 10024–10031

285b. Gooljarsingh, L. T., Ramcharan, J., Gilroy, S., and Benkovic, S. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6565–6570

286. Kastanos, E. K., Woldman, Y. Y., and Appling, D. R. (1997) *Biochemistry* **36**, 14956–14964

287. Schrimsher, J. L., Schendel, F. J., Stubbe, J., and Smith, J. M. (1986) *Biochemistry* **25**, 4366–4371

287a. Mueller, E. J., Oh, S., Kavalerchik, E., Kappock, T. J., Meyer, E., Li, C., Ealick, S. E., and Stubbe, J. (1999) *Biochemistry* **38**, 9831–9839

288. Schrimsher, J. L., Schendel, F. J., and Stubbe, J. (1986) *Biochemistry* **25**, 4356–4365

289. Firestine, S. M., and Davisson, V. J. (1994) *Biochemistry* **33**, 11917–11926

290. Meyer, E., Kappock, T. J., Osuji, C., and Stubbe, J. (1999) *Biochemistry* **38**, 3012–3018

290a. Thoden, J. B., Kappock, T. J., Stubbe, J., and Holden, H. M. (1999) *Biochemistry* **38**, 15480–15492

291. Szabados, E., Hindmarsh, E. J., Phillips, L., Duggleby, R. G., and Christopherson, R. I. (1994) *Biochemistry* **33**, 14237–14245

292. Rayl, E. A., Moroson, B. A., and Beardsley, G. P. (1996) *J. Biol. Chem.* **271**, 2225–2233

292a. Wall, M., Shim, J. H., and Benkovic, S. J. (2000) *Biochemistry* **39**, 11303–11311

292b. Vergis, J. M., Bulock, K. G., Fleming, K. G., and Beardsley, G. P. (2001) *J. Biol. Chem.* **276**, 7727–7733

292c. Brosius, J. L., and Colman, R. F. (2000) *Biochemistry* **39**, 13336–13343

292d. Toth, E. A., Worby, C., Dixon, J. E., Goedken, E. R., Marqusee, S., and Yeates, T. O. (2000) *J. Mol. Biol.* **301**, 433–450

293. Silva, M. M., Poland, B. W., Hoffman, C. R., Fromm, H. J., and Honzatko, R. B. (1995) *J. Mol. Biol.* **254**, 431–446

293a. Poland, B. W., Lee, S.-F., Subramanian, M. V., Siehl, D. L., Anderson, R. J., Fromm, H. J., and Honzatko, R. B. (1996) *Biochemistry* **35**, 15753–15759

294. Lee, T. T., Worby, C., Bao, Z.-Q., Dixon, J. E., and Colman, R. F. (1999) *Biochemistry* **38**, 22–32

294a. Guicherit, O. M., Cooper, B. F., Rudolph, F. B., and Kellems, R. E. (1994) *J. Biol. Chem.* **269**, 4488–4496

294b. Iancu, C. V., Borza, T., Choe, J. Y., Fromm, H. J., and Honzatko, R. B. (2001) *J. Biol. Chem.* **276**, 42146–42152

295. Bass, M. B., Fromm, H. J., and Rudolph, F. B. (1984) *J. Biol. Chem.* **259**, 12330–12333

296. Webb, M. R., Reed, G. H., Cooper, B. F., and Rudolph, F. B. (1984) *J. Biol. Chem.* **259**, 3044–3046

297. Choe, J.-Y., Poland, B. W., Fromm, H. J., and Honzatko, R. B. (1999) *Biochemistry* **38**, 6953–6961

297a. Fleming, M. A., Chambers, S. P., Connelly, P. R., Nimmesgern, E., Fox, T., Bruzzese, F. J., Hoe, S. T., Fulghum, J. R., Livingston, D. J., Stuver, C. M., Sintchak, M. D., Wilson, K. P., and Thomson, J. A. (1996) *Biochemistry* **35**, 6990–6997

297b. Markham, G. D., Bock, C. L., and Schalk-Hihi, C. (1999) *Biochemistry* **38**, 4433–4440

297c. Barnes, B. J., Eakin, A. E., Izydore, R. A., and Hall, I. H. (2000) *Biochemistry* **39**, 13641–13650

297d. Kerr, K. M., Digits, J. A., Kuperwasser, N., and Hedstrom, L. (2000) *Biochemistry* **39**, 9804–9810

298. Zhang, R.-g, Evans, G., Rotella, F. J., Westbrook, E. M., Beno, D., Huberman, E., Joachimiak, A., and Collart, F. R. (1999) *Biochemistry* **38**, 4691–4700

299. Whitby, F. G., Luecke, H., Kuhn, P., Somoza, J. R., Huete-Perez, J. A., Phillips, J. D., Hill, C. P., Fletterick, R. J., and Wang, C. C. (1997) *Biochemistry* **36**, 10666–10674

300. Farazi, T., Leichman, J., Harris, T., Cahoon, M., and Hedstrom, L. (1997) *J. Biol. Chem.* **272**, 961–965

301. Lowenstein, J., and Tornheim, K. (1971) *Science* **171**, 397–400

301a. Bai, Y., Fox, D. T., Lacy, J. A., Van Lanen, S. G., and Iwata-Reuyl, D. (2000) *J. Biol. Chem.* **275**, 28731–28738

301b. Watanabe, M., Nameki, N., Matsuo-Takasaki, M., Nishimura, S., and Okada, N. (2001) *J. Biol. Chem.* **276**, 2387–2394

301c. Tercero, J. A., Espinosa, J. C., Lacalle, R. A., and Jiménez, A. (1996) *J. Biol. Chem.* **271**, 1579–1590

302. Coffee, C. J., and Kofke, W. A. (1975) *J. Biol. Chem.* **250**, 6653–6658

303. Schultz, V., and Lowenstein, J. M. (1978) *J. Biol. Chem.* **253**, 1938–1943

304. Merkler, D. J., and Schramm, V. L. (1993) *Biochemistry* **32**, 5792–5799

305. Sollitti, P., Merkler, D. J., Estupiñán, B., and Schramm, V. L. (1993) *J. Biol. Chem.* **268**, 4549–4555

306. Fishbein, W. N., Armbrustmacher, V. W., and Griffin, J. L. (1978) *Science* **200**, 545–548

306a. Craig, S. P., III, and Eakin, A. E. (2000) *J. Biol. Chem.* **275**, 20231–20234

306b. Phillips, C. L., Ullman, B., Brennan, R. G., and Hill, C. P. (1999) *EMBO J.* **18**, 3533–3545

307. Stiles, G. L. (1992) *J. Biol. Chem.* **267**, 6451–6454

308. Newby, A. C. (1984) *Trends Biochem. Sci.* **9**, 42–44

309. Linden, J. (1991) *FASEB J.* **5**, 2668–2676

310. Palmer, T. M., Benovic, J. L., and Stiles, G. L. (1995) *J. Biol. Chem.* **270**, 29607–29613

311. Erb, L., Garrad, R., Wang, Y., Quinn, T., Turner, J. T., and Weisman, G. A. (1995) *J. Biol. Chem.* **270**, 4185–4188

312. Osborne, W. R. A. (1981) *Trends Biochem. Sci.* **6**, 80–83

313. Anderson, W. F. (1995) *Sci. Am.* **273**(Sep), 124–128

314. Marshall, E. (1995) *Science* **269**, 1050–1055

315. Hershfield, M. S., and Mitchell, B. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1725–1768, McGraw-Hill, New York

316. Sideraki, V., Mohamedali, K. A., Wilson, D. K., Chang, Z., Kellems, R. E., Quiocho, F. A., and Rudolph, F. B. (1996) *Biochemistry* **35**, 7862–7872

317. Wang, Z., and Quiocho, F. A. (1998) *Biochemistry* **37**, 8314–8324

318. Mao.C, Cook, W. J., Zhou, M., Federov, A. A., Almo.SC, and Ealick, S. E. (1998) *Biochemistry* **37**, 7135–7146

319. Cook, W. J., Ealick, S. E., Krenitsky, T. A., Stoeckler, J. D., Helliwell, J. R., and Bugg, C. E. (1985) *J. Biol. Chem.* **260**, 12968–12969

320. Benveniste, P., and Cohen, A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8373–8377

321. Deng, H., Kurz, L. C., Rudolph, F. B., and Callender, R. (1998) *Biochemistry* **37**, 4968–4976

321a. Fedorov, A., Shi, W., Kicska, G., Fedorov, E., Tyler, P. C., Furneaux, R. H., Hanson, J. C., Gainsford, G. J., Larese, J. Z., Schramm, V. L., and Almo, S. C. (2001) *Biochemistry* **40**, 853–860

321b. Tebbe, J., Bzowska, A., Wielgus-Kutrowska, B., Schröder, W., Kazimierczuk, Z., Shugar, D., Saenger, W., and Koellner, G. (1999) *J. Mol. Biol.* **294**, 1239–1255

322. McRoberts, J. A., and Martin, D. W., Jr. (1980) *J. Biol. Chem.* **255**, 5605–5615

322a. Xu, Y., and Grubmeyer, C. (1998) *Biochemistry* **37**, 4114–4124

322b. Canyuk, B., Focia, P. J., and Eakin, A. E. (2001) *Biochemistry* **40**, 2754–2765

323. Focia, P. J., Craig, S. P., III, Nieves-Alicea, R., Fletterick, R. J., and Eakin, A. E. (1998) *Biochemistry* **37**, 15066–15075

324. Wilson, J. M., and Kelley, W. N. (1984) *J. Biol. Chem.* **259**, 27–30

325. Rossiter, B. J. F., and Caskey, C. T. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 2 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1679–1706, McGraw-Hill, New York

326. Davidson, B. L., Pashmforoush, M., Kelley, W. N., and Palella, T. D. (1989) *J. Biol. Chem.* **264**, 520–525

327. Willis, R. C., Jolly, D. J., Miller, A. D., Plent, M. M., Esty, A. C., Anderson, P. J., Chang, H.-C., Jones, O. W., Seegmiller, J. E., and Friedmann, T. (1984) *J. Biol. Chem.* **259**, 7842–7849

References

327a. Héroux, A., White, E. L., Ross, L. J., and Borhani, D. W. (1999) *Biochemistry* **38**, 14485–14494

327b. Versées, W., Decanniere, K., Pellé, R., Depoorter, J., Brosens, E., Parkin, D. W., and Steyaert, J. (2001) *J. Mol. Biol.* **307**, 1363–1379

328. Yuan, G., Bin, J. C., McKay, D. J., and Snyder, F. F. (1999) *J. Biol. Chem.* **274**, 8175–8180

328a. Self, W. T., and Stadtman, T. C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7208–7213

328b. Kato, M., Mizuno, K., Crozier, A., Fujimura, T., and Ashihara, H. (2000) *Nature (London)* **406**, 956–957

329. Rumsby, P. C., Kato, H., Waldren, C. A., and Patterson, D. (1982) *J. Biol. Chem.* **257**, 11364–11367

330. Onrot, J., Goldberg, M. R., Biaggioni, I., Hollister, A. S., Kincaid, D., and Robertson, D. (1985) *N. Engl. J. Med.* **313**, 549–554

331. Florkin, M. (1949) *Biochemical Evolution*, Academic Press, New York

332. Henderson, J. F., and Paterson, A. R. P. (1973) *Nucleotide Metabolism*, Academic Press, New York

333. Kahn, K., and Tipton, P. A. (1997) *Biochemistry* **36**, 4731–4738

334. Kahn, K., Serfozo, P., and Tipton, P. A. (1997) *J. Am. Chem. Soc.* **119**, 5435–5442

335. Kahn, K., and Tipton, P. A. (1998) *Biochemistry* **37**, 11651–11659

335a. Enomoto, A., and 18 other authors (2002) *Nature (London)* **417**, 447 – 452

335b. Hediger, M. A., (2002) *Nature (London)* **417**, 393, 395

336. Hayashi, S., Fujiwara, S., and Noguchi, T. (1989) *J. Biol. Chem.* **264**, 3211–3215

337. Reynolds, P. H. S., Boland, M. J., Blevins, D. G., Randall, D. D., and Schubert, K. R. (1982) *Trends Biochem. Sci.* **7**, 366–368

338. Burg, A. W., and Brown, G. M. (1968) *J. Biol. Chem.* **243**, 2349–2358

339. Shiota, T., Palumbo, M. P., and Tsai, L. (1967) *J. Biol. Chem.* **242**, 1961–1969

340. Brown, G. M., and Williamson, J. M. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 521–538, Am. Soc. for Microbiology, Washington,DC

341. Bracher, A., Eisenreich, W., Schramek, N., Ritz, H., Götze, E., Herrmann, A., Gütlich, M., and Bacher, A. (1998) *J. Biol. Chem.* **273**, 28132–28141

342. Nar, H., Huber, R., Auerbach, G., Fischer, M., Hösl, C., Ritz, H., Bracher, A., Meining, W., Eberhardt, S., and Bacher, A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12120–12125

342a. Bracher, A., Fischer, M., Eisenreich, W., Ritz, H., Schramek, N., Boyle, P., Gentili, P., Huber, R., Nar, H., Auerbach, G., and Bacher, A. (1999) *J. Biol. Chem.* **274**, 16727–16735

342b. Bracher, A., Schramek, N., and Bacher, A. (2001) *Biochemistry* **40**, 7896–7902

342c. Maita, N., Okada, K., Hatakeyama, K., and Hakoshima, T. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1212–1217

343. Ahn, C., Byun, J., and Yim, J. (1997) *J. Biol. Chem.* **272**, 15323–15328

344. Haussmann, C., Rohdich, F., Schmidt, E., Bacher, A., and Richter, G. (1998) *J. Biol. Chem.* **273**, 17418–17424

345. Hampele, I. C., D'Arcy, A., Dale, G. E., Kostrewa, D., Nielsen, J., Oefner, C., Page, M. G. P., Schönfeld, H.-J., Stüber, D., and Then, R. L. (1997) *J. Mol. Biol.* **268**, 21–30

346. Sun, X., Bognar, A. L., Baker, E. N., and Smith, C. A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6647–6652

347. Roy, K., Mitsugi, K., and Sirotnak, F. M. (1996) *J. Biol. Chem.* **271**, 23820–23827

347a. Bermingham, A., Bottomley, J. R., Primrose, W. U., and Derrick, J. P. (2000) *J. Biol. Chem.* **275**, 17962–17967

347b. Baca, A. M., Sirawaraporn, R., Turley, S., Sirawaraporn, W., and Hol, W. G. J. (2000) *J. Mol. Biol.* **302**, 1193–1212

347c. Sun, X., Cross, J. A., Bognar, A. L., Baker, E. N., and Smith, C. A. (2001) *J. Mol. Biol.* **310**, 1067–1078

347d. Ravanel, S., Cherest, H., Jabrin, S., Grunwald, D., Surdin-Kerjan, Y., Douce, R., and Rébeillé, F. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 15360–15365

347e. Chave, K. J., Auger, I. E., Galivan, J., and Ryan, T. J. (2000) *J. Biol. Chem.* **275**, 40365–40370

348. Smith, G. K., Cichetti, J. A., Chandrasurin, P., and Nichol, C. A. (1985) *J. Biol. Chem.* **260**, 5221–5224

349. Switchenko, A. C., and Brown, G. M. (1985) *J. Biol. Chem.* **260**, 2945–2951

350. Oppliger, T., Thöny, B., Nar, H., Bürgisser, D., Huber, R., Heizmann, C. W., and Blau, N. (1995) *J. Biol. Chem.* **270**, 29498–29506

351. Bürgisser, D. M., Thöny, B., Redweik, U., Hess, D., Heizmann, C. W., Huber, R., and Nar, H. (1995) *J. Mol. Biol.* **253**, 358–369

352. Ploom, T., Thöny, B., Yim, J., Lee, S., Nar, H., Leimbacher, W., Richardson, J., Huber, R., and Auerbach, G. (1999) *J. Mol. Biol.* **286**, 851–860

353. Citron, B. A., Milstien, S., Gutierrez, J. C., Levine, R. A., Yanak, B. L., and Kaufman, S. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6436–6440

354. Auerbach, G., Herrmann, A., Gütlich, M., Fischer, M., Jacob, U., Bacher, A., and Huber, R. (1997) *EMBO J.* **16**, 7219–7230

355. Nichol, C. A., Smith, G. K., and Duch, D. S. (1985) *Ann. Rev. Biochem.* **54**, 729–764

356. Tanaka, K., Akino, M., Hagi, Y., Doi, M., and Shiota, T. (1981) *J. Biol. Chem.* **256**, 2963–2972

357. Wiederrecht, G. J., and Brown, G. M. (1984) *J. Biol. Chem.* **259**, 14121–14127

358. Jacobson, K. B., Dorsett, D., Pfleiderer, W., McCloskey, J. A., Sethi, S. K., Buchanan, M. V., and Rubin, I. B. (1982) *Biochemistry* **21**, 1238–1243

359. White, R. H. (1996) *Biochemistry* **35**, 3447–3456

360. Rasche, M. E., and White, R. H. (1998) *Biochemistry* **37**, 11343–11351

361. Wuebbens, M. M., and Rajagopalan, K. V. (1995) *J. Biol. Chem.* **270**, 1082–1087

361a. Unkles, S. E., Heck, I. S., Appleyard, M. V. C. L., and Kinghorn, J. R. (1999) *J. Biol. Chem.* **274**, 19286–19293

361b. Leimkühler, S., Wuebbens, M. M., and Rajagopalan, K. V. (2001) *J. Biol. Chem.* **276**, 34695–34701

361c. Leimkühler, S., and Rajagopalan, K. V. (2001) *J. Biol. Chem.* **276**, 1837–1844

361d. Schrag, J. D., Huang, W., Sivaraman, J., Smith, C., Plamondon, J., Larocque, R., Matte, A., and Cygler, M. (2001) *J. Mol. Biol.* **310**, 419–431

361e. Kuper, J., Palmer, T., Mendel, R. R., and Schwarz, G. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6475–6480

361f. Watanabe, T., Ihara, N., Itoh, T., Fujita, T., and Sugimoto, Y. (2000) *J. Biol. Chem.* **275**, 21789–21792

361g. Bittner, F., Oreb, M., and Mendel, R. R. (2001) *J. Biol. Chem.* **276**, 40381–40384

361h. Temple, C. A., and Rajagopalan, K. V. (2000) *J. Biol. Chem.* **275**, 40202–40210

361i. Lake, M. W., Temple, C. A., Rajagopalan, K. V., and Schindelin, H. (2000) *J. Biol. Chem.* **275**, 40211–40217

362. Ritz, H., Schramek, N., Bracher, A., Herz, S., Eisenreich, W., Richter, G., and Bacher, A. (2001) *J. Biol. Chem.* **276**, 22273–22277

363. Kis, K., Volk, R., and Bacher, A. (1995) *Biochemistry* **34**, 2883–2892

363a. Kelly, M. J. S., Ball, L. J., Krieger, C., Yu, Y., Fischer, M., Schiffmann, S., Schmieder, P., Kühne, R., Bermel, W., Bacher, A., Richter, G., and Oschkinat, H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13025–13030

363b. Illarionov, B., Eisenreich, W., and Bacher, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98** 7224–7229

364. García-Ramírez, J. J., Santos, M. A., and Revuelta, J. L. (1995) *J. Biol. Chem.* **270**, 23801–23807

364a. Jordan, D. B., Bacot, K. O., Carlson, T. J., Kessel, M., and Viitanen, P. V. (1999) *J. Biol. Chem.* **274**, 22114–22121

364b. Braden, B. C., Velikovsky, C. A., Cauerhff, A. A., Polikarpov, I., and Goldbaum, F. A. (2000) *J. Mol. Biol.* **297**, 1031–1036

364c. Meining, W., Mörtl, S., Fischer, M., Cushman, M., Bacher, A., and Ladenstein, R. (2000) *J. Mol. Biol.* **299**, 181–197

364d. Illarionov, B., Kemter, K., Eberhardt, S., Richter, G., Cushman, M., and Bacher, A. (2001) *J. Biol. Chem.* **276**, 11524–11530

365. Ritsert, K., Huber, R., Turk, D., Ladenstein, R., Schmidt-Bäse, K., and Bacher, A. (1995) *J. Mol. Biol.* **253**, 151–167

366. Efimov, I., Kuusk, V., Zhang, X., and McIntire, W. S. (1998) *Biochemistry* **37**, 9716–9723

367. Bowers-Komro, D. M., Yamada, Y., and McCormick, D. B. (1989) *Biochemistry* **28**, 8439–8446

368. Eisenreich, W., Schwarzkopf, B., and Bacher, A. (1991) *J. Biol. Chem.* **266**, 9622–9631

369. Hörig, J. A., Renz, P., and Heckmann, G. (1978) *J. Biol. Chem.* **253**, 7410–7414

370. Thompson, T. B., Thomas, M. G., Escalante-Semerena, J. C., and Rayment, I. (1998) *Biochemistry* **37**, 7686–7695

371. Trzebiatowski, J. R., and Escalante-Semerena, J. C. (1997) *J. Biol. Chem.* **272**, 17662–17667

372. Himmeldirk, K., Sayer, B. G., and Spenser, I. D. (1998) *J. Am. Chem. Soc.* **120**, 3581–3589

372a. Miranda-Rios, J., Navarro, M., and Soberón, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9736–9741

373. Estramareix, B., and Thérisod, M. (1984) *J. Am. Chem. Soc.* **106**, 3857–3860

374. Taylor, S. V., Kelleher, N. L., Kinsland, C., Chiu, H.-J., Costello, C. A., Backstrom, A. D., McLafferty, F. W., and Begley, T. P. (1998) *J. Biol. Chem.* **273**, 16555–16560

374a. Mueller, E. G., Palenchar, P. M., and Buck, C. J. (2001) *J. Biol. Chem.* **276**, 33588–33595

375. Julliard, J.-H., and Douce, R. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2042–2045

376. Kelleher, N. L., Taylor, S. V., Grannis, D., Kinsland, C., Chiu, H.-J., Begley, T. P., and McLafferty, F. W. (1998) *Protein Sci.* **7**, 1796–1801

377. Chiu, H.-J., Reddick, J. J., Begley, T. P., and Ealick, S. E. (1999) *Biochemistry* **38**, 6460–6470

377a. Peapus, D. H., Chiu, H.-J., Campobasso, N., Reddick, J. J., Begley, T. P., and Ealick, S. E. (2001) *Biochemistry* **40**, 10103–10114

377b. Reddick, J. J., Nicewonger, R., and Begley, T. P. (2001) *Biochemistry* **40**, 10095–10102

378. Cane, D. E., Hsiung, Y., Cornish, J. A., Robinson, J. K., and Spenser, I. D. (1998) *J. Am. Chem. Soc.* **120**, 1936–1937

378a. Cane, D. E., Du, S., and Spenser, I. D. (2000) *J. Am. Chem. Soc.* **122**, 4213–4214

378b. Osmani, A. H., May, G. S., and Osmani, S. A. (1999) *J. Biol. Chem.* **274**, 23565–23569

378c. Gupta, R. N., Hemscheidt, T., Sayer, B. G., and Spenser, I. D. (2001) *J. Am. Chem. Soc.* **123**, 11353–11359

References

379. Huynh, M. S., and Snell, E. E. (1985) *J. Biol. Chem.* **260**, 2379–2383
380. Nelson, M. J. K., and Snell, E. E. (1986) *J. Biol. Chem.* **261**, 15115–15120
381. Chaiyen, P., Brissette, P., Ballou, D. P., and Massey, V. (1997) *Biochemistry* **36**, 13856–13864
382. Fleming, S. M., Robertson, T. A., Langley, G. J., and Bugg, T. D. H. (2000) *Biochemistry* **39**, 1522–1531
383. Krieger, C. J., Roseboom, W., Albracht, S. P. J., and Spormann, A. M. (2001) *J. Biol. Chem.* **276**, 12924–12927

Study Questions

1. The vitamin niacin (nicotinic acid) is converted to NAD through the intermediate, desamido NAD.

   a) Propose an enzymatic synthesis of desamido NAD beginning from niacin and ribose 5-phosphate
   b) Propose an enzymatic synthesis of NAD beginning from desamido NAD

2. AMP is formed from IMP in an ATP requiring reaction. The introduced nitrogen atom is derived directly from aspartate.

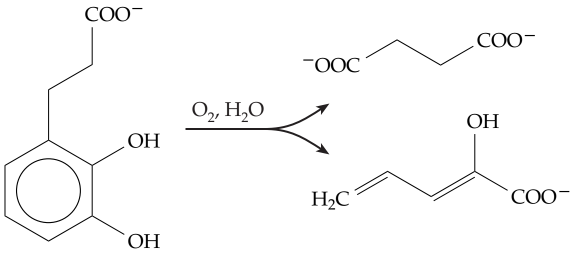
   a) Propose a reaction mechanism based on an analogous reaction sequence occurring during amino acid metabolism.
   b) Identify the analogous enzymes upon which you have based your answer by names of the enzymes or by the names of the substrates and products.

3. One of the mechanisms proposed for the decarboxylation of orotidine 5'-phosphate to UMP involves initial addition of an enzyme nucleophile to the pyrimidine ring. Describe and criticize this mechanism. Hint: The proposal has some similarity to the thymidylate synthase mechanism.

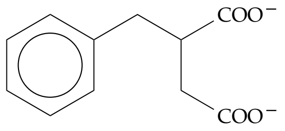
4. Name the enzymes and describe the chemical reaction that occurs for each of the four steps in the following pathway:

$$\mathrm{UDP} \xrightarrow{1} \mathrm{dUDP} \xrightarrow{2} \mathrm{dUTP} \xrightarrow{3} \mathrm{dUMP} \xrightarrow{4} \mathrm{dTMP}$$

5. The following reaction is catalyzed by a dioxygenase from *E. coli*. The dioxygenase reaction opens the ring and the intermediate is cleaved hydrolytically. Propose a structure for the intermediate and a mechanism for the hydrolytic cleavage. See Fleming *et al.*[382]

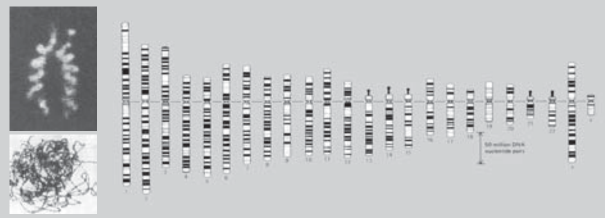


6. Anaerobic breakdown of toluene by denitrifying bacteria begins by the addition of toluene to fumarate to form benzylsuccinate. Benzylsuccinate synthase has an amino acid sequence homologous to that of pyruvate formate lyase (p. 799–801), contains a glycyl radical, and is activated in a manner similar to activation of pyruvate formate lyase. Propose a mechanism for formation of benzylsuccinate. See Krieger *et al.*[383]



Benzylsuccinate

7. Write out, using structural formulas, a step-by-step reaction sequence for the conversion of *O*-succinylbenzoate into isochorismate as indicated in Fig. 25-4.



Right: The human **karyotype**, a full set of chromosomes numbered according to size and content and showing characteristic banding patterns. The 22 autosomes plus either an X or Y chromosome make up the haploid set which, for the female, contains ~ 3,500 Mbp of DNA. Each diploid cell contains 46 chromosomes. From Alberts *et al.* "Molecular Biology of the Cell", Third edition, Garland, New York, 1994, p 356.

Upper Left: A pair of mitotic sister chromatids in a section stained with an antibody to topoisomerase II. Notice that the two chromatids are coiled with opposite helical handedness. Lower Left: Meiotic chromosomes of a lily at the pachytene stage in which sister chromatids are connected along their length in a **synaptonemal complex.** From Kleckner, N. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8167–8174

Contents

1473 .... **A. Historical Sketch**
1473 .... 1. DNA as the Genetic Material
1474 .... 2. The Double Helix
1474 .... 3. Ribonucleic Acids and Proteins
1474 .... *Messenger RNA*
1475 .... *Transfer RNA*
1475 .... 4. Deciphering the Genetic Code
1476 .... **B. Genetic Methods**
1476 .... 1. Mutations
1477 .... 2. Mapping the Chromosome of a Bacteriophage
1478 .... 3. The Cistron
1478 .... 4. Nutritional Auxotrophs
1479 .... 5. Establishing the Correspondence of a Genetic Map and an Amino Acid Sequence
1480 .... 6. Conditionally Lethal Mutations
1480 .... 7. The Nature of Suppressor Genes
1481 .... **C. Plasmids, Episomes, and Viruses**
1482 .... 1. Bacterial Sex Factors
1483 .... 2. Temperate Bacteriophage; Phage Lambda
1486 .... **D. Mapping of Chromosomes**
1486 .... 1. The Chromosome Map of *E. coli*
1486 .... 2. Restriction Endonucleases
1488 .... 3. Restriction Mapping
1489 .... 4. Electron Microscopy
1490 .... 5. Optical Mapping
1490 .... **E. Cloning, Modifying, and Transferring Genes**
1490 .... 1. Joining DNA Fragments
1491 .... *Some useful enzymes*
1491 .... *Forming homopolymeric tails*
1491 .... *Preparing material for cloning*
1492 .... 2. Cloning Vehicles (Vectors)
1493 .... *Plasmids related to ColE1*
1494 .... *Typical cloning procedure with pBR322*
1494 .... *Selecting clones using β-galactosidase in pUC cloning vehicles*
1494 .... *Filamentous bacteriophages*
1496 .... *Lambda cloning vehicles*
1496 .... *Cosmid vehicles and in vitro packaging*
1497 .... 3. Expression of Cloned Genes in Bacteria
1497 .... 4. Cloning and Transferring Eukaryotic Genes
1497 .... *Plant genes*
1498 .... *Transferring genes with engineered animal viruses*
1498 .... *Physical methods of gene transfer*
1499 .... 5. Genomic Libraries
1499 .... 6. Probes
1500 .... 7. Studies of Restriction Fragments
1500 .... *Locating mutations*
1500 .... *Positional cloning*
1500 .... 8. Directed Mutation
1501 .... *Targeting and replacing genes*
1501 .... *Knockout mice*
1501 .... **F. The Genetics of Eukaryotic Organisms**
1501 .... 1. Mitosis
1502 .... *Cyclin-dependent kinases*
1502 .... *Checking for completion of replication and for DNA damage*
1503 .... *Mitotic spindle formation and the spindle assembly checkpoint*
1503 .... *Anaphase*
1504 .... 2. Meiosis
1506 .... 3. Polytene Chromosomes
1507 .... 4. Cytoplasmic Inheritance
1507 .... **G. The Human Genome Project**
1507 .... 1. The Mammalian Genome and Human Health
1508 .... 2. Understanding Gene Sequences
1509 .... *Human variation*
1510 .... *Other evolutionary relationships*
1511 .... *Metabolic studies of eukaryotic cells*
1513 .... 3. Understanding Human Genetic Diseases
1516 .... *Triplet repeat diseases*
1517 .... *Cloned genes and diagnosis*
1517 .... *Vaccines, hormones, and other medicines*
1518 .... 4. Gene Therapy
1518 .... 5. Genetic Engineering of Bacteria, Plants, and Animals
1519 .... 6. Ethical Problems

1520 .... **References**
1527 .... **Study Questions**

Boxes

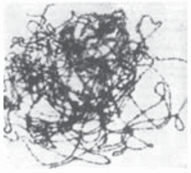
1513 .... Box 26-A Cystic Fibrosis

Tables

1485 .... Table 26-1 Some Genes of *E. coli*
1487 .... Table 26-2 Some Commonly Used Restriction Endonucleases, Their Sources, and Cleavage Sites
1492 .... Table 26-3 Some Enzymes Used in Molecular Cloning
1515 .... Table 26-4 Some Human Triple Repeat and Related Diseases

Biochemical Genetics

26



The most exciting biological discoveries of the 20th century include the unraveling of the genetic code (Chapter 5) and the understanding of the ways in which nucleic acids and proteins are synthesized. The biosynthesis of both nucleotides and amino acids has been considered in Chapters 24 and 25 and the basic chemistry of the polymerization processes in Chapters 12 and 17. Chapters 27–29 deal with the ways in which these polymerization reactions are controlled and by which the correct sequences of nucleotides or amino acids are obtained. The understanding of these matters is a development of genetics[1–2b] as well as of biochemistry; hence this introductory and historical chapter.

A. Historical Sketch

The discovery of deoxyribonucleic acid dates to 1869, when Miescher isolated a new chemical substance from white blood cells that he obtained from pus and later from sperm cells.[3] The material, which became known as nucleic acid, occurred in both plants and animals, thymus glands and yeast cells being among the best sources. Chemical studies indicated that the nucleic acids isolated from thymus glands and from yeast cells were different. As we now know, thymus nucleic acid was primarily DNA and yeast nucleic acid primarily RNA. For a while it was suspected that animals contained only DNA and plants only RNA, and it was not until the early 1940s that it was established that both substances were present in all organisms.[3–6]

1. DNA as the Genetic Material

In 1928, Griffith, using cells of *Diplococcus pneumoniae*, showed that genetic information that controls properties of the capsular polysaccharides could be transferred from one strain of bacteria to another. A material present in killed cells and in cell-free extracts permanently altered the capsular properties of cells exposed to the material.[7] This **transformation** of bacteria remained a mystery for many years. At the time of the experiments there was no hint of the genetic role of nucleic acids, which were generally regarded as strange materials. Furthermore, the covalent structure of nucleic acids was uncertain. A popular idea was that a tetranucleotide served as a repeating unit for some kind of regular polymer. Genes were most often thought to be protein in nature.

However, in 1944, Avery and associates showed that purified DNA extracted from pneumococci could carry out transformation.[8–12] The transforming principal appeared to contain little protein. It was not inactivated by proteolytic enzymes but was inactivated by deoxyribonuclease. Avery was 67 at the time of this discovery, refuting the popular contention that all important scientific discoveries are made by young people.

Other experiments also pointed to the conclusion that DNA was the genetic material. DNA was found localized in the nuclei of eukaryotic cells. The absolute amount per cell was constant for a given species. Studies of bacteriophage replication pointed strongly to DNA as the genetic material.[11] In 1952 Hershey and Chase showed that when a phage particle infects a cell the viral DNA enters the bacterium, but the protein "coat" remains outside.[13] This was demonstrated by preparing two types of labeled bacteriophage T2. One contained ^{32}P, which had been incorporated into the DNA, and the other ^{35}S incorporated into the proteins. Cells of *E. coli* were infected with the labeled phage preparations and were then agitated violently in a blender to shear the phage particles off of the bacteria. Over 80% of the ^{35}S was removed from the bacteria by this treatment, but most of the ^{32}P entered the bacteria and could be recovered in the next generation of progeny bacteriophage.[14]

2. The Double Helix

Development of newer methods of investigation of the chemical composition of nucleic acids led Chargaff to an important discovery. Although the base composition of DNA was extremely variable from organism to organism, the molar ratio of adenine to thymine was nearly always 1:1, as was that of cytosine to guanine![15] This observation provided the basis for the concept of base pairing in the structure of DNA. The final important information was supplied by X-ray diffraction studies of stretched fibers of DNA, which showed that molecules of DNA were almost certainly helical structures containing more than one chain. The crucial experiments were done by Rosalind Franklin[16–17b] and M. H. F. Wilkins, whose data were used by Watson and Crick in 1952 in constructing their model of the double helical structure[12,18–20] (Fig. 5-6). Once established, the structure of DNA itself suggested both the nature of the genetic code and a replication mechanism. *The genetic code had to lie in the sequence of nucleotides. Base pairing provided a mechanism by which the two mutually complementary strands could be separated. Biosynthesis of a new complementary strand alongside each one would result in a precise replication of each gene and of the entire genome.* In a similar fashion, RNA could be synthesized alongside a DNA "template" and could then be used to direct protein synthesis.

The presence of RNA in the cytoplasm had been linked to protein synthesis by experiments done in the early 1940s. After the discovery of the double helix, the concept followed quickly that DNA was the master "blueprint" from which secondary blueprints or **transcripts** of RNA could be copied. The RNA copies, later identified as **messenger RNA** (mRNA), provided the genetic information for specifying protein sequence. The flow of information from DNA to RNA to proteins could be symbolized as in Eq. 26-1.

$$\begin{tikzcd} \text{DNA} \arrow[r, "\text{Transcription}"] & \text{RNA} \arrow[r, "\text{Translation}"] \arrow[dr, dashed] & \text{Proteins} \arrow[d] \\ & & \text{Control of Metabolism} \end{tikzcd} \qquad (26\text{-}1)$$

Proteins, in one way or another, control nearly all of metabolism. This includes the reactions that form the nucleotide precursors of the nucleic acids and that lead to polymerization of the amino acids and nucleotides, reactions catalyzed by **protein enzymes** and **ribozymes**. Thus, the flow of information from DNA to proteins is only part of a larger loop of metabolic processes (p. 973). Genetic information flows from DNA out into the cell, and copies of the master blueprint are passed nearly unchanged from generation to generation. The simple concepts implied by Eq. 26-1 quickly caught the imagination of the entire community of scientists and led to a rapid blossoming of the field of biochemical genetics.[20a,b]

3. Ribonucleic Acids and Proteins

By 1942 it was clear from ultraviolet cytophotometry developed by Caspersson[21] and from cytochemical work of Brachet[22,23] that RNA had something to do with protein synthesis. Use of radioautography with ^{3}H-containing uridine showed that RNA was synthesized in the nucleus of eukaryotic cells and was transported out into the cytoplasm.[24,25]

Ribosomes were discovered by electron microscopists examining the structure of the endoplasmic reticulum using ultrathin sectioning techniques. Their presence in cells was established by 1956, and the name ribosome was proposed in 1957. At first it was difficult to study protein synthesis *in vitro* using isolated ribosomes. No *net* synthesis could be detected until Hoagland and associates measured the rate of incorporation of ^{14}C-labeled amino acids into protein.[26] This sensitive method permitted measurement of very small amounts of protein synthesis in cell-free preparations from rat liver and paved the way toward studies with ribosomes themselves.

Immediately after the Watson–Crick proposals were made in 1953, it was generally thought that ribosomal RNA (rRNA), which constitutes up to 90% of the total RNA of some cells, carried the genetic message from the nucleus to the cytoplasm. By 1960 it seemed unlikely. For one thing the size and composition of rRNA was similar for different bacteria, despite differences in base composition of the DNA (Chapter 5).[27] It had been concluded that a relatively unstable, short-lived form of RNA must carry the message. Ribosomal RNA, however, was quite stable.[28]

Messenger RNA. In 1956, Volkin and Astrachan[29,30] detected a rapidly labeled and labile RNA in phage-infected bacterial cells. Studies of enzyme induction also suggested the existence of mRNA. Many bacteria, including *E. coli*, when grown on glucose as the sole source of energy and then suddenly switched to lactose, are unable to utilize the new sugar immediately. However, transfer to lactose induces, within a period of two minutes, the synthesis of new proteins needed for the metabolism of this sugar. Among the new proteins is a **permease** for lactose and a **β-galactosidase** (Chapter 12)[31] that cleaves the disaccharide to glucose and galactose. When the lactose is exhausted, the level of the induced enzymes drops almost as quickly. These results suggested that the RNA that carries the genetic message for synthesis of the new enzymes must be unstable. It must be produced rapidly in response to the presence of the inducing sugar and must disappear rapidly in its absence.

In 1961, Jacob and Monod postulated messenger RNA (mRNA) as a short-lived polynucleotide.[30,32,33] An abundance of additional evidence supported the proposal. For example, RNA molecules produced after infection of *E. coli* by bacteriophage T4 underwent hybridization (Chapter 5) with denatured DNA of the bacteriophage. Furthermore, this virus-specific mRNA became associated with preexisting bacterial ribosomes and provided the template for synthesis of phage proteins.[34] The experiment provided direct evidence for transcription of mRNA from genes of the viral DNA.

Transfer RNA. Crick[35] suggested, in 1957, that special "adapter" molecules might be needed to align amino acids with their codons in the RNA transcript. He thought that the adapters might be polynucleotides. At the same time chemical studies of the RNA of cells revealed that a low-molecular-weight RNA made up 15% of the total RNA of *E. coli*. This RNA was recognized in the same year as constituting the needed adapters, when Hoagland demonstrated enzymatic "activation" of amino acids and their subsequent incorporation into protein. The name transfer RNA (tRNA; Figs. 5-30, 5-31) was proposed.[36]

4. Deciphering the Genetic Code

The genetic code consists of triplets of DNA base pairs (codons), each corresponding to a single amino acid. The triplets are consecutive, do not overlap, and are not separated by any "punctuation." Although this represented one of the simplest possibilities for a code, it required much effort more than a period of more than 10 years to prove it. Even after the triplet nature of the genetic code became evident, many questions remained. Were all of the 64 possible codons used by the living cell? If so, were they all used to code for amino acids or were some set aside for other purposes? How many codons were used for a single amino acid? Was the code universal or did different organisms use different codes? How could one decipher the code? Despite the complexity of these questions, they seem to have been almost completely answered.

In an important experiment[37] Nirenberg and Matthaei, in 1961, isolated ribosomes from *E. coli* and mixed them with crude extracts of soluble materials, also from *E. coli* cells. The extracts included tRNA molecules and aminocyl-tRNA synthases. The 20 amino acids, ATP, and an ATP-generating system (PEP + pyruvate kinase) were added. Nirenberg showed that under such conditions protein was synthesized by ribosomes in response to the presence of added RNA. For example, RNA from tobacco mosaic virus (Chapter 7) was very effective in stimulating protein synthesis. The crucial experiment, which was done originally simply as a "control," was one in which a synthetic polynucleotide consisting solely of uridylic acid units was substituted for mRNA. In effect, this was a synthetic mRNA containing only the codon UUU repeated over and over. The ribosomes read this code and synthesized a peptide containing only phenylalanine. Thus, poly(U) gave polyphenylalanine, and *UUU was identified as a codon specifying phenylalanine*. The first nucleotide triplet had been identified! In the same manner CCC was identified as a proline codon and AAA as a lysine codon. Study of mixed copolymers containing two different nucleotides in a random sequence suggested other codon assignments. A few years later, after Khorana had supplied the methods for synthesis of oligonucleotides and of regular alternating polymers of known sequence, the remaining codons were identified.

Another important technique was based on the observation that synthetic trinucleotides induced the binding to ribosomes of tRNA molecules that were "charged" with their specific amino acids.[38,39] For example, the trinucleotides UpUpU and ApApA stimulated the binding to ribosomes of ^{14}C-labeled phenylalanyl-tRNA and lysyl-tRNA, respectively. The corresponding dinucleotides had no effect, an observation that not only verified the two codons but also provided direct evidence for the triplet nature of the genetic code. Another powerful approach was the use of artificial RNA polymers, synthesized by combined chemical and enzymatic approaches.[40] For example, the polynucleotide CUCUCUCUCU · · · led to the synthesis by ribosomes of a regular alternating polypeptide of leucine and serine.

Table 5-5 shows the codon assignments, as we now know them, for each of the 20 amino acids. Table 5-6 shows the same 64 codons in a rectangular array. In addition to those codons assigned to specific amino acids, three are designated as **chain termination codons**: UAA, UAG, and UGA. These are frequently referred to as "nonsense" codons. The termination codons UAA and UAG are also known as *ochre* and *amber*, respectively, although these names have no scientific significance.[41] The codons AUG (methionine) and much less often GUG (valine) serve as the **initiation codons** in protein synthesis. Consequently, the N-terminal amino acid in most newly synthesized eukaryotic proteins is methionine, and in bacterial proteins it is *N*-formylmethionine. As explained in Chapter 29, *N*-formylmethionyl-tRNA is specifically bound to initiation sites containing the AUG codon in bacterial mRNA-ribosome complexes.

A number of studies suggested that the genetic code as worked out for *E. coli* might be universal. For example, in the laboratories of Wittman and of Fraenkel-Conrat, RNA extracted from tobacco mosaic virus was treated with nitrous acid, a procedure known to deaminate many cytosine residues to uracil (Eq. 5-12). Such treatment could change the codon UCU (serine) to UUU (phenylalanine) and the codon CCC (proline) to CUC (leucine). When the nitrous acid-treated RNA was used to infect tobacco plants and virus particles were prepared in quantity from the resultant mutant strains, it was found that the amino acid sequence of the virus coat protein had been altered,[42] and that many of the alterations were exactly those that would be predicted from Table 5-6. Likewise, the amino acid substitutions in known defects of hemoglobin (Fig. 7-23) could be accounted for, in most cases, by single base alterations. Thus, hemoglobin S arose as a result of the following change in the sixth codon of the globin β chain gene: GAG (Glu) → GTG (Val).[43] Another argument favoring a universal code was based on the observation that mRNA coding for a globin chain could be translated by ribosomes and tRNA molecules from *E. coli.* The resulting protein was authentic mammalian globin.[44]

As often happens, a well-established conclusion may have to be modified. There are exceptions to the universal genetic code in mitochondrial DNA and in some protozoa (Chapter 5).[45]

B. Genetic Methods

Our present knowledge of molecular biology has depended greatly on the methods of genetics. The following introduction begins with a consideration of mutations.

1. Mutations

Changes in the structure of DNA occur only rarely. The average gene may be duplicated 10^6 times before some mistake results in a single detectable mutation.[46] Nevertheless, by using bacteria or bacterial viruses, it is possible to screen enormous numbers of individuals for the occurrence of mutations. If one million virus particles are spread on an agar plate under conditions where mutation in a certain gene can be recognized, on the average one mutant is found. The most common mutations are **base-pair switches** or **point mutations** that result from incorporation of the wrong base during replication or repair. In these mutations one base of a triplet codon is replaced by another to form a different codon, causing the substitution of one amino acid by another in the corresponding protein as was seen for hemoglobin S. Changes involving replacement of one pyrimidine by another (C → T or T → C) or of one purine by another are sometimes called **transition mutations**, whereas if a pyrimidine is replaced by a purine, or vice versa, the mutation is known as a **transversion**. An example is the previously mentioned A → T in hemoglobin S. Transition mutations are by far the most common, one possible cause being pairing with a minor tautomer of one of the bases (Chapter 5). For example, A could pair with a minor tautomer of C, causing a mutation from T to C. Note that substitution of an incorrect base in one strand will lead, in the next round of replication, to correct pairing again but with an AT pair replaced by GC, or vice versa, in one of the daughter DNA duplex strands. A base substitution does not always cause an amino acid replacement because of the "degeneracy" of the code, i.e., the fact that more than one codon specifies a given amino acid.

From the observed rate of appearance of point mutations (one mutation per 10^6 gene duplications), we can estimate that one mutation occurs per 10^9 replications at a single nucleotide site. Point mutants tend to "back mutate," often at almost the same rate as is observed for the forward mutation. That is, one in 10^9 times a mutation of the same nucleotide will take place to return the code to its original form. The phenomenon is easy to understand. For example, if T should be replaced by C because the latter formed a minor tautomer and paired with A, the mutation would appear in progeny duplexes as a GC pair. When this pair was replicated, there would be a finite probability that the C of the parental DNA strand would again assume the minor tautomeric structure and pair with A instead of G, leading to a back mutation.

Although the rates of spontaneous mutation are low, they can be greatly increased by mutagenic chemicals (Chapter 27) or by irradiation. It is perfectly practical to measure the rates of both forward and back mutation. When this was done, it was found that certain chemicals, e.g., acridine dyes, induce mutations that undergo reverse mutation at a very much lower frequency than normal. It was eventually shown that these mutations resulted either from **deletions** of one or more nucleotides from the chain or from **insertions** of extra nucleotides. Deletion and insertion mutations often result from errors during genetic recombination and repair at times when the DNA chain is broken.

Mutations involving deletion or insertion of one or a few nucleotides are called **frame-shift mutations**.

Messenger RNA is read by the protein synthesizing machinery from some starting point. As is illustrated in Chapter 5, Section E,1 the codons are read three bases at a time, and the proper amino acid corresponding to each codon is inserted. *When a deletion or insertion in the mRNA is met, all subsequent codons may be misread because the reading frame has shifted forward or backward by one or two nucleotides.* The protein synthesized bears little resemblance to that formed by the nonmutant organism and is usually completely nonfunctional. Mutations are considered further in Chapter 27.

2. Mapping the Chromosome of a Bacteriophage

Intensive work on the "T-even" phage T2, T4, and T6 (Box 7-C) was begun in 1938 by Delbrück and associates. The genetic information for these viruses is carried in a single linear DNA molecule, which in the case of T4 contains ~1.7 x 10^5 base pairs (170 kb), enough for about 200 genes. Before the sequences of the viral DNA were known, the positions of more than 60 of these genes were mapped in the following way. When a bacteriophage infects a cell of *E. coli*, it injects its DNA through the cell wall and into the cytoplasm. About 20 minutes later the cell bursts, and ~100 fully formed replicas of the original virus particle are released. This rapid rate of production of progeny is so fast that it is possible to carry out in a test tube in 20 min a genetic experiment that would require the entire population of the earth if done with humans. The approach is explained nicely by Seymour Benzer, the man who first mapped the fine structure of a gene.[47] Bacteriophage particles, like bacteria, can be "plated out" on agar plates, which must contain a uniform suspension of bacteria susceptible to the virus. Wherever a virus particle lies, a bacterium is infected. Soon the infection spreads to neighboring bacteria with production of a transparent "plaque" (Fig. 26-1). The number of active virus particles present in a suspension can be determined easily by plating and counting of the plaques.

Mutant bacteriophages can be identified in various ways. Some biochemical traits affect the appearance of the plaque. Other easily detected traits include alteration in the specificity toward strains of the host bacterium. A key discovery that made genetic mapping possible for bacteriophage was that *genetic recombination between two phage particles can take place within a host bacterium*. When large numbers of bacteriophages of two different mutant strains were grown and were mixed together in excess with many bacteria, a few of the progeny phage were found to contain both mutant traits in the same virus and an equal number were "wild type." Although recombinations between mutations that are located close together in the DNA are rare, their frequency still greatly exceeds that of new mutations. While this type of experiment gave no hint about the nature of the events involved, it showed that recombination had occurred.

Study of recombination frequencies between different strains of phage soon revealed that some sites of mutation are **closely linked**. Recombination between these sites occurs only rarely. Other sites are weakly linked, and recombination occurs often. This behavior was reminiscent of that established many years earlier for genes of the fruit fly *Drosophila*, maize, and other higher organisms. Recombination by "crossing over" in the chromosomes of *Drosophila* was established by Morgan and associates in 1911.[1,2] The basic idea behind chromosome mapping in any organism is the assumption that *recombination frequencies between two mutations are directly proportional to the distance between them on the genetic map*. For the T4 phage a recombination frequency of 1% is taken as one unit. The total T4 map is 700 units long. The fact that this is greater than 100% means that if genes are located at opposite ends of the chromosome multiple recombination events can occur between them. However, a maximum of 50% crossing over is observed for distant gene pairs, and the approximate linearity of map distance and recombination frequency holds only for distances of 10 units or less.[48]

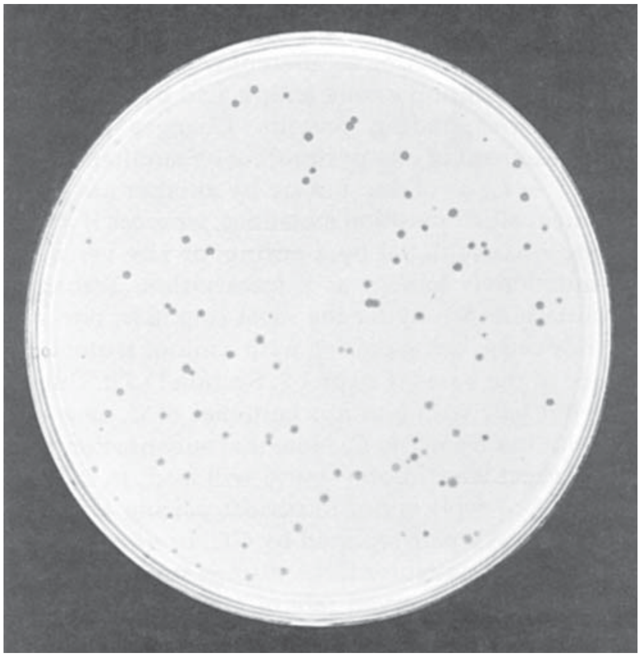


Figure 26-1 Plaques formed by bacteriophage ϕ11 growing on *Staphylococcus aureus*. Each transparent (dark) plaque is the result of lysis of bacteria by the progeny of a single bacteriophage particle. Courtesy of Peter Pattee.

How can recombinant bacteriophage be identified rapidly? Benzer used two strains of *E. coli*, the B strain and the K strain, as hosts. Mutants in gene *rII* form characteristic plaques on strain B but do not grow on strain K. To determine the recombination frequency between two different *rII* mutants, the viruses were added to liquid cultures of B cells (in which they replicate), and recombination was allowed to occur. Recombination permitted the emergence not only of a phage containing *both* mutations but also of a wild-type phage in which both mutations had been eliminated by the recombination process. Since only recombinants of the latter type grow in strain K, it was possible to detect a single recombinant among one billion progeny. Since the total DNA length in phage T4 is 166 kb, there are 237 base pairs for each of the 700 units of map length. Thus, a recombination frequency at 0.01% between two mutations meant that the two mutations were no more than three base pairs apart in the DNA. Benzer concluded that he had observed the expected recombination frequencies for mutations even of immediately adjacent bases in the DNA.

To make fine genetic mapping practical, a series of bacteriophage containing deletion mutations involving large segments of the *rII* gene were isolated. Using these, it was possible to establish in which segments of the gene a particular mutation lay. Then, recombination experiments with previously identified mutations in that same general region allowed the mutations to be pinpointed. In this way Benzer identified over 300 sites of mutations within the *rII* region. He concluded that the minimum distance between two mutable sites was compatible with the Watson–Crick structure of the gene.[47]

3. The Cistron

How can one tell whether two mutations are in the same gene or in nearby or adjacent genes? The answer can be supplied by a test of **complementation**. If two mutant bacteriophage are altered in different genes, they can often reproduce within a host if the bacterium is infected with both of the mutant phage. Since each one has a good gene for one of the two proteins involved, recombinant phage, in which all of the gene functions are fulfilled, will be formed. On the other hand, if both mutant phages are defective in the same gene (although at different locations), they usually cannot complement each other at high frequencies in a coinfection. The experiment is referred to as a ***cis*–*trans* comparison**. The coinfection with the two different mutants is the trans test. A control, the *cis* test, uses a recombinant containing both of the mutations in the same DNA and coinfection with a standard phage. Normal replication is expected in this instance.

When the complementation (or *cis*–*trans*) test was applied to various mutants in the *rII* region, it became clear that there are two genes, *rIIA* and *rIIB*. The name **cistron** was proposed by Benzer to represent that length of DNA identifiable in this fashion as a genetic unit. *For most purposes the terms gene and cistron are nearly synonymous.* When mapping of the *rII* region was done, there was no information about the functions of the proteins specified by these two cistrons. However, both the *rIIA* and *rIIB* proteins have since been shown to become incorporated in the membranes of phage-infected bacterial cells.[49,50] There they affect the ease of lysis of the infected cells and, in that manner, cause *rII* plaques to be larger and to have sharper edges than standard plaques.

4. Nutritional Auxotrophs

The beginning of biochemical genetics is often attributed to Beadle and Tatum, who, in 1940, discovered mutants of *Neurospora* with specific growth requirements. They X-rayed one parent strain to form mutants, then tested individual spores for their possible need of a specific nutrient for growth. The 299th spore tested required pyridoxine! Many other mutants requiring vitamins, amino acids, and nucleic acid bases were then discovered.[1,51] A few years later similar **nutritional auxotrophs**, as the mutants are called, were discovered for *E. coli*. Ordinary wild-type cells of *E. coli* can grow on a minimal medium containing a carbon compound as a source for energy together with inorganic nutrients. Irradiation with ultraviolet light or treatment with mutagenic chemicals produces many mutant cells that fail to grow on such a minimal medium. However, addition of one or more specific compounds, such as an amino acid or vitamin, usually permits growth. Selection of such nutritional auxotrophs can be accomplished by plating out large numbers of the irradiated or chemically treated cells on a solid, rich nutrient medium. Colonies (**clones**) are allowed to develop by multiplication of the individual bacteria. The auxotrophs are selected by **replica plating**.[2] In this procedure a sterile velveteen pad is pressed against a nutrient agar plate containing small colonies of bacteria and is used to “print” replica plates containing a minimal medium. The colonies on the initial and replica plates are compared and the colonies of auxotrophs (which do not grow on the minimal medium) are selected. In a second stage, the auxotrophs may be replica plated to minimal medium supplemented with various nutrients (amino acids, purines, pyrimidines, vitamins, etc.). Selection is made easier by pretreatment of the irradiated cells suspended in minimal medium with penicillin. Penicillin kills the growing cells, but the auxotrophs, which do not grow on the minimal medium, survive. The penicillin is then destroyed by adding penicillinase (a β-lactamase; Box 20-G) leaving a suspension much enriched in the percentage of auxotrophic mutants.[2]

A nutritional auxotroph of a bacterium often has a defect in a gene specifying a protein needed for the biosynthesis of the required nutrient. Individual genes recognized in this way are named with a genetic symbol. For example, gene *trpA* specifies one of the two protein subunits of tryptophan synthase. Other kinds of mutations, e.g., those affecting motility or other properties of the cells, can also be detected and are given appropriate symbols. A few of these genetic symbols are indicated on the *E. coli* chromosome map in Fig. 26-4, and many others are used throughout this book. On the basis of such nutritional experiments Beadle, by 1945, had proposed his famous one-gene-one-enzyme hypothesis.[1]

5. Establishing the Correspondence of a Genetic Map and an Amino Acid Sequence

Although the studies of the *rII* region of the T4 chromosome established that genetic mapping could be carried to the level of individual nucleotides in the DNA, it was still necessary to prove a linear correspondence between the nucleotide sequence in the DNA and the amino acid sequence in proteins. This was accomplished by Yanofsky[52,53] and associates through study of the enzyme tryptophan synthase of *E. coli.* Tryptophan synthase (Fig. 25-3) consists of two subunits, α and β, the former containing only 268 amino acids and encoded by the *trpA* gene. To obtain a fine structure map of the A gene, a large series of tryptophan auxotrophs unable to grow in the absence of added tryptophan were isolated. Genetic crosses were carried out with the aid of a **transducing bacteriophage** Plkc. Transducing bacteriophage (Section E,3), while multiplying in susceptible bacteria, sometimes incorporate a portion of the bacterial chromosome into their own DNA. Then, when the virus infects other bacteria, some of the genetic information can be transferred through recombination into the chromosome of bacteria that survive infection. Use of a series of deletion mutants, as in the *rII* mapping, permitted division of the A gene into a series of segments, and observation of recombination frequencies permitted fine structure mapping.

The second part of the proof of colinearity of DNA and protein sequences was the determination of the complete amino acid sequence of tryptophan synthase and peptide mapping (Chapter 3) of fragments of the mutant enzymes. From the peptide maps it was possible to identify altered peptides and to establish the exact nature of the amino acid substitutions present in a variety of different tryptophan auxotrophs. When this was done, it was found that those mutations that mapped very close together had amino acid substitutions at adjacent or nearly adjacent sites in the peptide chain.

The same problem was approached by Sarabhai and associates[54] through the nonsense mutations (Section 6), which lead to premature chain termination during protein synthesis. During late stages of the infection of *E. coli* by phage T4, most of the protein synthesis is that of a single protein of the virus head. Synthesis of protein by infected cells was allowed to proceed in the presence of specific ^{14}C-labeled amino acids. Then cell extracts were digested with trypsin or chymotrypsin, the head-protein peptides were separated by electrophoresis, and autoradiograms were prepared. A series of T4 nonsense mutants that mapped within the head-protein gene were shown to give rise to incomplete head-protein chains. The peptide fragments were of varying lengths. By examining the radioautograms prepared from the enzymatically fragmented peptides, it was possible to arrange the mutants in a sequence based on the length of peptide formed and to show that this was the same as that deduced by genetic mapping. More recently the colinearity of codon and amino acid sequences has been verified repeatedly by comparison of experimentally determined nucleotide sequences in RNA and DNA molecules with the corresponding amino acid sequences for thousands of proteins.

Before the triplet nature of codons had been established, Crick and associates used frame-shift mutations in a clever way to demonstrate that the genetic code did consist of triplets of nucleotides.[7,55,55a] Consider what will happen if two strains of bacteria, each containing a frame-shift mutation (e.g., a −1 deletion), are mated. Genetic recombination can occur to yield mutants containing *both* of the frame-shift mutations. It would be difficult to recognize such recombinants because, according to almost any theory of coding, they would still produce completely defective proteins. However, Crick *et al*. introduced a third frame-shift mutation of the same type into the same gene and observed that the recombinants containing all three deletions (or insertions) were able to synthesize at least partially active proteins. Thus, while introduction of one or two single nucleotide deletions completely inactivates a gene, deletion of three nucleotides close together within a gene shortens the total message by just three nucleotides. The gene will contain only a short region in which the codons are scrambled. The reading frame for the remainder of the protein will not be changed. The protein specified will often be functional because it has a normal sequence except for a small region where some amino acid substitutions will be found and where one amino acid will be completely missing.

6. Conditionally Lethal Mutations

Studies of plaque morphology and of nutritional auxotrophs are directed narrowly at one gene or group of genes. It is desirable to have a general means of detecting mutations in the many other genes present within cells. However, most mutations are **lethal,** and this effect cannot be overcome by adding any nutrient. Lethal mutations are very common in higher organisms, but since eukaryotic cells have pairs of homologous chromosomes, they can be carried in one chromosome and the individual survives. With bacteria and viruses there is only one chromosome, and lethal mutants cannot survive.

Nutritional auxotrophs can be described as **conditionally lethal mutants**; they survive only if the medium is supplemented with the nutrient, whose synthesis depends upon the missing enzyme. Other kinds of conditional lethal mutations permit study of almost every gene in an organism. For example, **temperature-sensitive** (*ts*) mutants grow perfectly well at a low temperature, e.g., 25°C, but do not grow at a higher temperature, e.g., 42°C.[41,56] Many temperature-sensitive mutations involve an amino acid replacement that causes the affected protein to be less stable to heat than is the wild type protein. Others involve a loss in protein-synthesizing ability for reasons that may be obscure. Temperature-sensitive mutations occur spontaneously in nature, an example being the gene that controls hair pigment in Siamese cats.[41] The gene (or gene product) is inactivated at body temperatures but is active in the cooler parts of the body, such as the paws, tail, and nose, with the result that the cat's hair is highly pigmented only in those regions.

Screening for conditionally lethal temperature-sensitive mutants of bacteriophage T4 permitted isolation of hundreds of mutants involving sites at random over the entire viral chromosome. Complementation studies permitted assignment of these to individual genes, which at first were identified only by number (Fig. 26-2). Later specific functions were associated with the genes.[57,58] For example, the product of gene 42 was identified as an enzyme required in the synthesis of hydroxymethyl-dCMP (Chapter 5). Genes 20–24, among others, must code for head proteins because mutants produce normal tails but no heads. Gene 23 codes for the major head subunits, while gene 20 has something to do with "capping" the end of the head. These mutants produce cylindrical "polyheads" in place of the normal heads. Mutants of genes 25–29 have defective base plates and do not form tails, while mutants 34–38 lack tail fibers. The specific ways in which some of these gene products are assembled to form the base plates and tails of the phage are indicated in Fig. 7-29. The positions of the *rIIA* and *rIIB* genes are also shown in Fig. 26-2.

A second type of conditionally lethal mutation leads to alteration of an amino acid codon to one of the three **chain termination codons** UAG, UAA, and UGA (Table 5-3).[60,61] These are often called **nonsense mutations** in contrast to **missense mutations** in which one amino acid is replaced by another. A chain termination mutant synthesizes only part of the product of the defective gene because of the presence of the termination codon. A remarkable aspect of chain termination mutations is that they can be **suppressed** by other mutations in distant parts of the virus or bacterial chromosome. Many otherwise lethal mutations of bacteriophage T4 were discovered by their ability to grow in mutant strains of *E. coli*, which contained **suppressor genes**,[62] and their inability to grow in the normal B strain. Three different suppressor genes *supD*, *supE*, and *supF* were found to suppress mutations that formed UAG. These are commonly known as *amber* suppressor genes. A second group of *ochre* suppressor genes, including *supB* and *supC*, suppressed mutations that formed codon UAA. Suppressors for mutations that form codon UGA have also been found.[63] Like the temperature-sensitive mutants, *amber* and *ochre* mutants can be obtained in almost any genes of a bacterial virus. Chain termination mutants of unessential genes in bacteria can be recognized by transferring the genes by conjugation or by viral transduction into a strain (sup^+) that contains a desired suppressor gene.

Conditionally lethal mutants have been of great value in developing our understanding of the genetics of bacterial viruses. They have also provided a powerful technique for approaching complex problems of bacterial physiology. For example, we may ask how many genes are required for a bacterium to sense the presence of a food and to swim toward it (Chapter 19). Even though few clues as to the basic chemistry underlying these phenomena can be obtained in this way, the use of temperature-sensitive mutants and complementation tests permits us to establish the total number of genes involved in these complex processes and to map their positions on a bacterial or viral chromosome. This is often an important step toward a more complete understanding of a biological phenomenon.

7. The Nature of Suppressor Genes

How can one mutation be suppressed by a second mutation at a different point in the chromosome? Rarely, a mutation is suppressed by a second mutation within the *same* gene. Such **intragenic complementation** sometimes occurs when a mutation leads to an amino acid replacement that disrupts the structural stability or function of a protein. Sometimes a mutation at another site involving a residue, which interacts with the first amino acid replaced, will alter the inter-

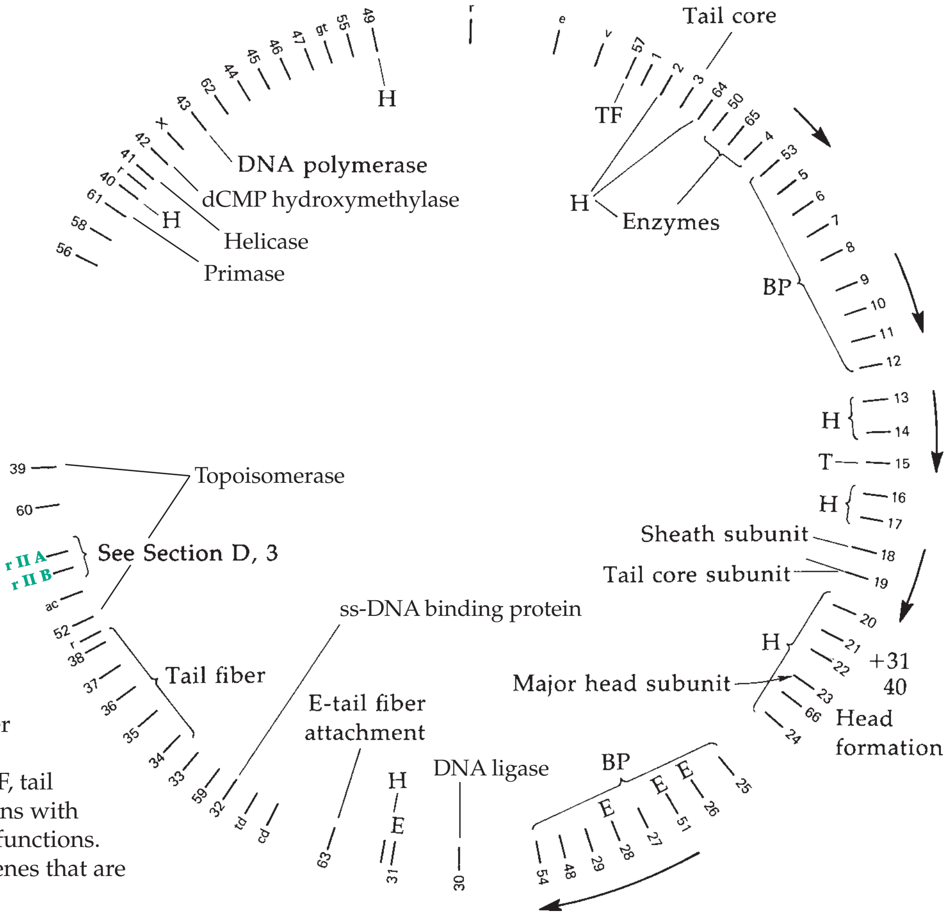

49
55
gt
47
46
45
44
62
43
X
42
41
r
40
61
58
56
H
H
DNA polymerase
dCMP hydroxymethylase
Helicase
Primase
r
e
v
57
1
2
3
64
50
65
4
53
5
6
7
8
9
10
11
12
Tail core
TF
H
Enzymes
BP
H
13
14
T
15
H
16
17
Sheath subunit
18
Tail core subunit
19
20
H
21
22
+31
40
Major head subunit
23
66
24
Head formation
25
26
E
E
51
27
E
28
BP
29
48
54
DNA ligase
30
H
E
31
E-tail fiber attachment
63
ss-DNA binding protein
cd
td
32
59
33
34
35
36
37
38
r
52
Tail fiber
ac
r II A
r II B
See Section D, 3
60
39
Topoisomerase

Figure 26-2 Linkage map for bacteriophage T4. See H. A. Sober and Eiserling and Dickson.[57,59] Abbreviations: H, head; T, tail; TF, tail fiber; BP, base plate; and E, proteins with enzymatic rather than structural functions. The bold arrows indicate some genes that are transcribed together (operons).

action between the two residues in a way that restores function to the protein. For example, if the first amino acid side chain is small and is replaced by mutation with a larger side chain, a second mutation leading to a decrease in the size of another side chain may permit the protein to fold and function properly. An example was found among mutants of tryptophan synthase.[64] Mutants in which Gly 211 of the α chain was replaced by Glu or Tyr 175 of the same chain by Cys both produced inactive enzymes. However, the double mutant with both replacements synthesized active tryptophan synthase. It is known now that these residues are adjacent to one another and form part of the binding site for the substrate indole-3-glycerol phosphate (Figs. 25-2; 25-3). Only the double mutant permits the substituted side chains to pack properly. In other cases, intragenic suppression involves changes in the subunit interactions in oligomeric proteins. These changes may affect the formation of correct quaternary structures of the proteins.

As discussed in the preceding section, the best known suppressor genes are those that suppress chain termination mutations (Section 6). These genes often encode mutant forms of tRNA molecules, which allow incorporation of an amino acid rather than chain termination to occur. They are discussed further in Chapter 29. Suppressor genes are not limited to bacteria. For example, the vermilion eye color mutation of *Drosophila* leads to a loss of brown eye pigments because of the inactivity of tryptophan 2,3-dioxygenase (Eq. 18-38). However, synthesis of the tryptophan dioxygenase from the *vermilion* mutant is inhibited by $tRNA_2^{Trp}$, one of the two tryptophanyl tRNAs. The suppressor mutation alters the tRNA in such a way that the inhibition is relieved.[61]

C. Plasmids, Episomes, and Viruses

The small pieces of DNA known as plasmids, which replicate independently of the chromosomes, have been discussed briefly in Chapter 5. Plasmids share a number of properties with viruses, and both are important to the techniques of contemporary molecular biology and genetic engineering. Bacterial plasmids may be present as one or several copies for each chromosome. Episomes are plasmids that are able to become integrated into the bacterial chromosome. Some extrachromosomal elements are episomes in one host and plasmids in another. Bacterial plasmids may be infectious (transferable) or noninfectious. In the former case, they are able to transfer their DNA into another cell and are known as **sex factors** (F agents or fertility factors). A sex factor is able to integrate into a chromosome and later to come out and transfer other genes with it. In this property it resembles a transducing phage.

Plasmids and episomes vary in size. The F sex factor is a 100-kb circular molecule of supercoiled DNA. Colicinogenic factors,[65] which may also be present in *E. coli* in as many as 10–15 copies per bacterial chromosome, are often much smaller (6–7.5 kb). Some larger colicinogenic plasmids are also sex factors. They carry genes for toxic protein antibiotics known as **colicins** (Box 8-D) which attack other strains of *E. coli*, providing a selective advantage for the strain producing the colicin. They also carry a gene or genes conferring on the host bacterium resistance to antibiotics such as penicillin and chloramphenicol. Penicillin is inactivated because the plasmids carry a gene encoding a penicillinase that hydrolytically cleaves the β-lactam ring (Box 20-G). Chloramphenicol (Fig. 25-10) is inactivated by the action of chloramphenicol *O*-acyltransferase.

1. Bacterial Sex Factors

Bacteria usually reproduce by a simple cell division. The DNA in the chromosome is doubled in quantity and the cell divides, each daughter cell receiving an identical chromosome. However, in 1946 Lederberg and Tatum showed that sexual reproduction is also possible.[66,67] They studied nutritional auxotrophs of *E. coli* strain K-12, which lacked the ability to synthesize amino acids or vitamins. When cells of two different mutants were mixed together and allowed to grow for a few generations, a few individual bacteria regained the ability to grow on a minimal medium. Since each of the two strains had one defective gene, the creation of an individual with neither of the two defects required combining of genetic traits from both strains. The existence of bacterial conjugation was recognized. Later it was established that true **genetic recombination** had occurred, i.e., genes from the two mating cells had been integrated into a single molecule of bacterial DNA.

This transfer of DNA between bacterial cells requires the presence of a plasmid sex factor (F agent), whose presence confers a male character to the individual cell. The F agent is large enough to contain about 90 genes and has a length of ~30 µm, ~2.5% that of the *E. coli* chromosome. Among other things, the *E. coli* F agent contains the genes needed to direct the synthesis of the **F pili** (sex pili). These tiny appendages, 8.5 nm in diameter (see Fig. 7-9), grow out quickly during a period of 4–5 min to a length of about 1.1 µm. The end of an F pilus becomes attached to a female cell (a bacterium lacking the F agent) and may induce the transfer of DNA into the female cell. The mechanism of transfer has been uncertain. It may involve formation of a cytoplasmic bridge between cells in close contact. The pilus may be retracted into the membrane of the male cell, pulling the two cells close together. The DNA probably flows through the pilus into the female.[68,69]

On rare occasions an F agent becomes integrated into the chromosome of a bacterium. Both the F agent and the chromosome have been shown by electron microscopy to be circular. The integration process requires the enzymatic cleavage of the DNA of both the chromosome and the F agent and the rejoining of the ends in such a way that a continuous circle is formed (Fig. 26-3). The enzymes that catalyze these reactions are considered in Chapter 27. Different F agents can be incorporated into the chromosome at different points around the circle. A strain of bacteria containing an integrated F agent is known as an *Hfr* (high frequency of recombination) strain.

When an Hfr strain conjugates with an F^- (female), replication of the entire male chromosome commences at some point near the end of the integrated F agent, and genes of the bacterial chromosome followed by those of the F factor are transferred into the female. Only a single strand of DNA (customarily referred to as the **plus strand**) is transferred from the donor cell and into the recipient cell (Fig. 26-3). There the complementary **minus strand** is synthesized to form a complete double-stranded DNA molecule bearing the genes from the Hfr cell. Only rarely does a copy of the entire chromosome of the donor cell enter the female cell. More often the DNA strand, or perhaps the pilus itself, breaks and only part of the chromosome is transferred.

Partial chromosome transfer from a male cell transforms the F^- cell into a partial diploid (**merozygote**) containing double the usual number of some of the genes. Within this partial diploid genetic recombination between the two chromosomes takes place (Fig. 26-3) by the mechanisms discussed in Chapter 27. The end result of the recombination process is that the daughter cells formed by subsequent division contain only single chromosomes with the usual number of genes. However, some genes come from each of the two parental strains. Thus, an F^- mutant unable to grow on a medium deficient in a certain nutrient may receive a gene from the male and now be able to grow on a minimal medium. Even though the number of such recombinants is small, they are easily selected from the very large number of mutant bacteria that are mixed together initially.

One result of DNA transfer from Hfr into F^- bacteria is sometimes the introduction of a complete copy of the F agent into the female bacterium. Since this con-

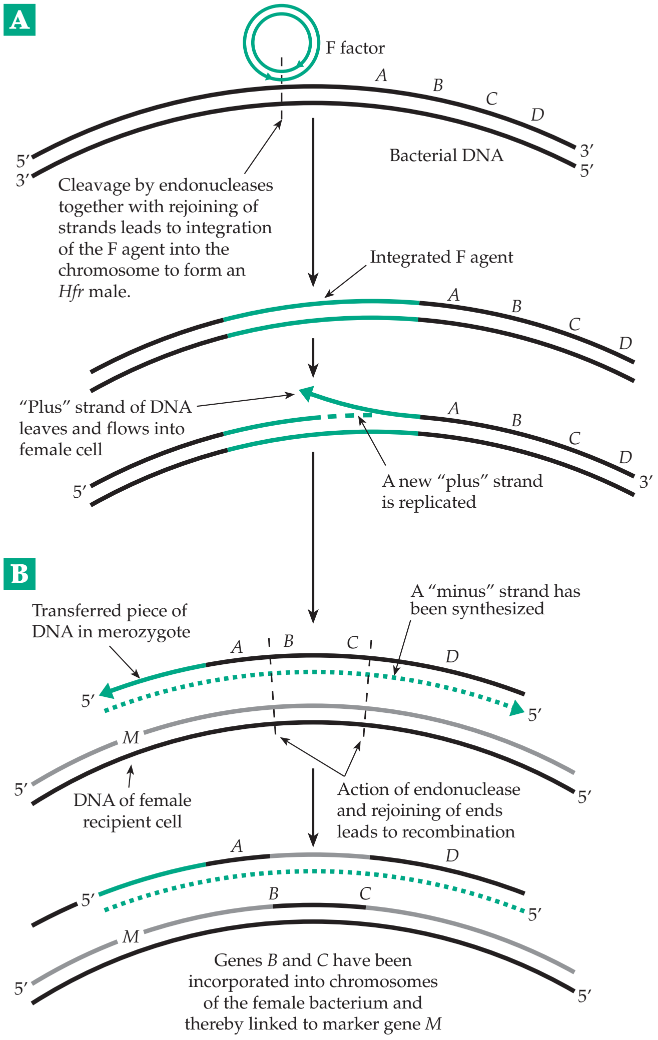

A
F factor
A
B
C
D
5′
3′
Bacterial DNA
3′
5′
Cleavage by endonucleases together with rejoining of strands leads to integration of the F agent into the chromosome to form an *Hfr* male.
Integrated F agent
A
B
C
D
"Plus" strand of DNA leaves and flows into female cell
A
B
C
D
5′
3′
A new "plus" strand is replicated
B
Transferred piece of DNA in merozygote
A "minus" strand has been synthesized
A
B
C
D
5′
5′
M
5′
5′
DNA of female recipient cell
Action of endonuclease and rejoining of ends leads to recombination
A
D
5′
B
C
5′
M
5′
5′
Genes *B* and *C* have been incorporated into chromosomes of the female bacterium and thereby linked to marker gene *M*

Figure 26-3 Integration of an F agent into a bacterial chromosome and transfer of some bacterial genes into another cell. (A) Incorporation of the F agent into *E. coli* genome and transfer of the "plus" strand of DNA out to a female recipient cell. (B) Genetic recombination between a piece of transferred DNA and the genome of the recipient cell.

verts the recipient into a male, Brinton referred to "bacterial sex as a virus disease." In fact, infectious plasmids and viruses display many similarities. For example, filamentous bacteriophages (Chapter 5; Fig. 7-7) adsorb to the F pili of male bacteria and the DNA, flowing in a direction opposite to that in bacterial conjugation, enters the cell.[70] The bacteriophage carry genes for the protein subunits of their coats (Figs. 26-2, 7-29), while F factors carry genes for synthesis of pilins. Pilins accumulate within the cell membrane and are extruded to generate F pili, just as viral subunits are extruded to form the virus coats. There is also a close similarity between episomes that can be integrated into bacterial chromosomes and the **temperate bacteriophages** considered in the next section.

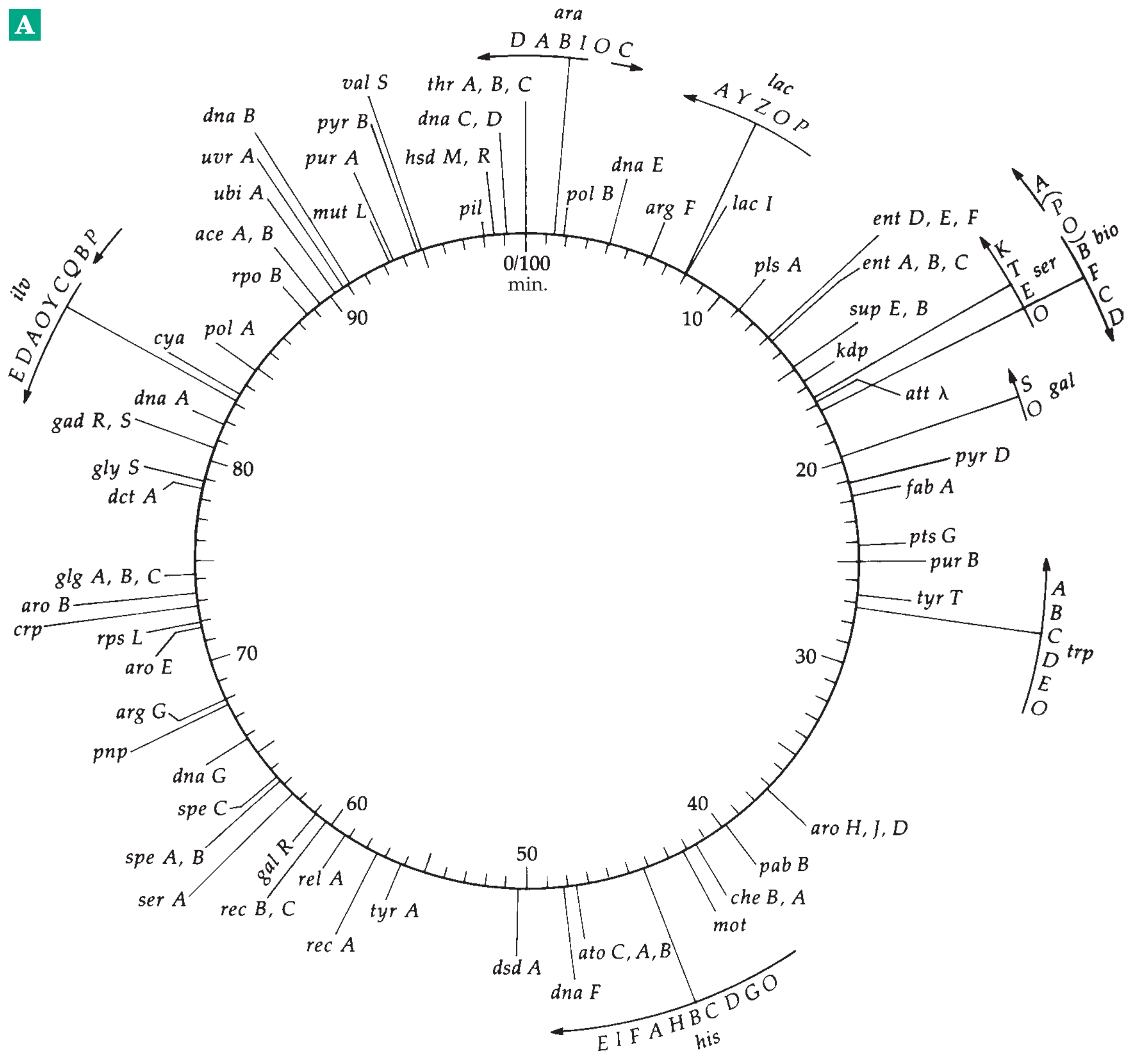
2. Temperate Bacteriophage; Phage Lambda

When DNA from a typical bacteriophage enters a bacterial cell, it seizes control of the metabolic machinery of the cell almost immediately and directs it entirely toward the production of new virus particles. This leads within a period of about 20 min to the production of one or two hundred progeny viruses and to the lysis and death of the cell. However, the DNA from a temperate phage may become repressed and, like an F factor, be integrated with the bacterial genome (Fig. 26-3). In the resulting **prophage** or **lysogenic** state, the repressed phage DNA is replicated as part of the bacterial genome but does no harm to the host cells unless some factor "activates" the incorporated genetic material by release of the repression. Replication of the phage and lysis of the bacterium then ensues. Temperate phage may also exist as plasmids (e.g., plasmid P1).

The decision for lysis or lysogeny, which is very important for the survival of the bacteriophage, is governed primarily by the nutritional status of the host. For a bacterium growing in a relatively rich environment such as the colon, lysis will increase the chances of daughter phage encountering host bacteria. However, in soils, *E. coli* grows very slowly, and a bacteriophage capable of entering a lysogenic state has an increased chance of survival until the host bacterium finds a richer growth medium.

The best known temperate phage is **phage lambda** of *E. coli*.[70–72] A tailed virus resembling the T-even phages (Box 7-C), phage λ has a smaller (~48.5 kb) DNA genome.[73] Within the bacterial cell the ends of the λ DNA may be joined to form a circular replicative form of the virus. In ~30% of the infected cells the λ DNA becomes integrated into the *E. coli* chromosome at the special site, *att* λ, which is located at 17 min on the *E. coli* chromosome map (Fig. 26-4). The incorporated phage DNA now occupies a linear segment amounting to about 1.2% of the total length of the *E. coli* chromosome. It is replicated along with the rest of the chromosome and for, the most part, goes unnoticed.

The host, *E. coli* K12, contains useful *amber* suppressors that make it easy to detect mutations in the bacteriophage. The integrated prophage can undergo mutations of almost any type, including large deletion mutations, and can still be investigated through complementation studies with other strains of virus. Thus, a family of modified **defective λ phage** was developed. When the λ prophage is excised from the bacterial chromosome, adjacent bacterial genes are occasionally carried with it. This allowed development of **λ transducing phage,** which can carry genes and transfer them into bacteria lacking these genes. More recently an important series of cloning vehicles have been derived from phage λ.


A
ara
D A B I O C
thr A, B, C
val S
pyr B
dna C, D
dna B
uvr A
pur A
hsd M, R
dna E
lac
A Y Z O P
ubi A
mut L
pil
pol B
arg F
lac I
ace A, B
rpo B
0/100
min.
90
10
pls A
ent D, E, F
ent A, B, C
ilv
E D A O Y C Q B P
cya
pol A
A P O B F C D
bio
K T E O
ser
sup E, B
kdp
att λ
S O
gal
dna A
gad R, S
gly S
dct A
80
20
pyr D
fab A
pts G
pur B
glg A, B, C
aro B
crp
tyr T
A B C D E O
trp
rps L
aro E
70
30
arg G
pnp
dna G
spe C
60
40
aro H, J, D
spe A, B
gal R
rel A
50
pab B
ser A
rec B, C
tyr A
che B, A
mot
rec A
dsd A
ato C, A, B
dna F
E I F A H B C D G O
his

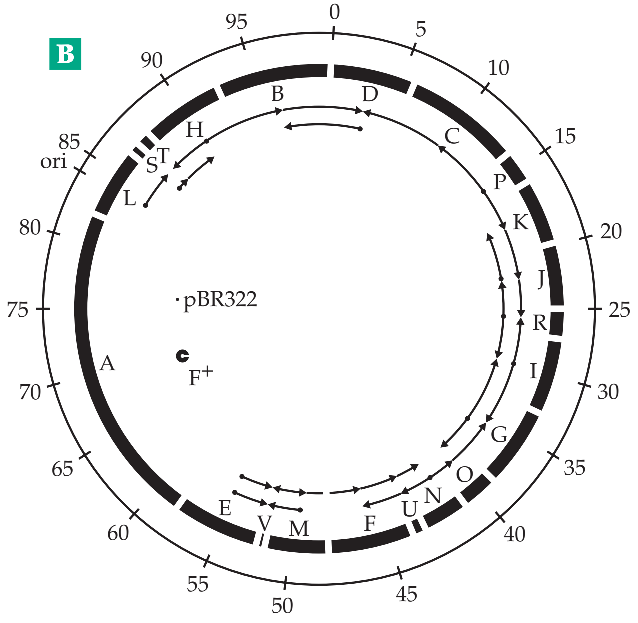

B
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
ori
90
95
B
D
H
C
S
T
L
P
K
J
pBR322
R
A
F+
I
G
O
E
V
M
F
U
N

Figure 26-4 (A) Genetic map of *E. coli*, strain K-12 published in 1976. Data from Bachmann, Low, and Taylor.[75] The scale in minutes is based on the results of interrupted conjugation experiments. The *thr* locus is arbitrarily taken as zero. The genetic symbols are defined in Table 26-1. Only a single strand of the DNA molecule is normally transcribed into RNA over any given region. Because of the antiparallel nature of the DNA helix, transcription from one strand proceeds clockwise and transcription from the other counterclockwise (the *r* strand is conventionally taken as the one that is transcribed in the clockwise direction when the chromosome is oriented as in the drawing). The directions of transcription of certain operons are indicated by the arrows. The 1983 version of this map[76] contained 1027 loci and we now know the location of all of the more than 4300 genes. (B) Physical map of the *E. coli* chromosome obtained by study of large 20–100 kb fragments obtained by digestion with restriction endonuclease *Not*I and separated by pulsed field gel electrophoresis. The fragments A–V were ordered using genetic information and overlapping fragments from partial *Not*I digests (arrows). On the outside the genetic map has been superimposed after distortion of the scale in minutes to make the two maps coincident. The 100-kb F$^+$ plasmid and the 4.5-kb plasmid pBR322 are also drawn to show relative sizes. From Smith *et al.*[77] A finer restriction map, based on 3400 cloned fragments, was prepared by Kohara, Akiyama, and Isono in 1987.[78]

TABLE 26-1
Some Genes of *E. coli*[a]

Gene symbol	Mnemonic	Map position (min)[a]	Phenotypic trait affected
aceA	Acetate	89	Isocitrate lyase
aceB	Acetate	89	Malate synthetase A
araA	Arabinose	1	L-Arabinose isomerase
araB	Arabinose	1	L-Ribulokinase
araC	Arabinose	1	Regulatory gene
araD	Arabinose	1	L-Ribulose-5-phosphate-4-epimerase
araI	Arabinose	1	Initiator locus
araO	Arabinose	1	Operator locus
argF	Arginine	6	Ornithine carbamoyltransferase
argG	Arginine	68	Argininosuccinic acid synthetase
aroB	Aromatic	73	Dehydroquinate synthetase
aroD	Aromatic	37	Dehydroquinate dehydratase
aroE	Aromatic	71	Dehydroshikimate reductase
aroH	Aromatic	37	DAHP synthetase (tryptophan-repressible isoenzyme)
aroJ	Aromatic	37	Probable operator locus for *aro*H
atoA	Acetoacetate	48	Coenzyme A transferase
atoB	Acetoacetate	48	Thiolase II
atoC	Acetoacetate	48	Regulatory gene
*att*λ	Attachment	17	Integration site for prophage λ
bioA	Biotin	17	Group II; 7-oxo-8-aminopelargonic acid (7 KAP) → 7,8-diaminopelargonic acid (DAPA)
bioB	Biotin	17	Conversion of dethiobiotin to biotin
bioC	Biotin	17	Block prior to pimeloyl-CoA
bioD	Biotin	17	Dethiobiotin synthetase
bioF	Biotin	17	Pimeloyl-CoA → 7 KAP
bioO	Biotin	17	Operator for genes *bio*B through *bio*D
bioP	Biotin	17	Promoter site for genes *bio*B through *bio*D
cheA	Chemotaxis	42	Chemotactic motility
cheB	Chemotaxis	42	Chemotactic motility
crp		73	Cyclic adenosine monophosphate receptor protein
cya		83	Adenylate cyclase
dctA		79	Uptake of C_4-dicarboxylic acids
dnaA	DNA	82	DNA synthesis; initiation defective
dnaB	DNA	91	DNA synthesis
dnaC	DNA	99	*dnaD*; DNA synthesis; initiation defective
dnaE	DNA	4	*polC*, DNA polymerase III and mutator activity
dnaF	DNA	48	*nrdA*; ribonucleoside diphosphate reductase
dnaG	DNA	66	DNA synthesis
dsdA	D-Serine	50	D-Serine deaminase
entA	Enterochelin	13	2,3-Dihydro-2,3-dihydroxybenzoate dehydrogenase
entB	Enterochelin	13	2,3-Dihydro-2,3-dihydroxybenzoate synthetase
entC	Enterochelin	13	Isochorismate synthetase
entD,E,F	Enterochelin	13	Unknown steps in conversion of 2,3-dihydroxybenzoate to enterochelin
fabA		22	β-Hydroxydecanoylthioester dehydratase
gadR		81	Regulatory gene for *gadS*
gadS		81	Glutamic acid decarboxylase
galE	Galactose	17	Uridine diphosphogalactose 4-epimerase
galK	Galactose	17	Galactokinase
galO	Galactose	17	Operator locus
galT	Galactose	17	Galactose 1-phosphate uridyltransferase
galR	Galactose	61	Regulatory gene
glgA	Glycogen	74	Glycogen synthetase
glgB	Glycogen	74	α-1,4-Glucan: α-1,4-glucan 6-glucosyltransferase
glgC	Glycogen	74	Adenosine diphosphate glucose pyrophosphorylase
glyS	Glycine	79	Glycyl-transfer RNA synthetase
hisA	Histidine	44	Isomerase
hisB	Histidine	44	Imidazole glycerol phosphate dehydrase: histidinol phosphatase
hisC	Histidine	44	Imidazole acetol phosphate aminotransferase
hisD	Histidine	44	Histidinol dehydrogenase
hisE	Histidine	44	Phosphoribosyl-adenosine triphosphate-pyrophospho-hydrolase
hisF	Histidine	44	Cyclase
hisG	Histidine	44	Phosphoribosyl-adenosine triphosphate-pyro-phosphorylase
hisH	Histidine	44	Amidotransferase
hisI	Histidine	44	Phosphoribosyl-adenosine monophosphate-hydrolase
hisO	Histidine	44	Operator locus
dsdM	Host specificity	98	Host modification activity: DNA methylase M
hsdR	Host specificity	98	Host restriction activity: endonuclease R
ilvA	Isoleucine–valine	83	Threonine deaminase (dehydratase)
ilvB	Isoleucine–valine	83	Acetohydroxy acid synthetase I
ilvC	Isoleucine–valine	83	α-Hydroxy-β-oxo acid reductoisomerase
ilvD	Isoleucine–valine	83	Dehydrase
ilvE	Isoleucine–valine	83	Aminotransferase B
ilvO	Isoleucine–valine	83	Operator locus for genes *ilv*A,D,E
ilvP	Isoleucine–valine	83	Operator locus for gene *ilv*B
ilvQ	Isoleucine–valine	83	Induction recognition site for *ilv*C
ilvY	Isoleucine–valine	83	Positive control element for *ilv*C induction
kdp	K accumulation	16	Defect in potassium ion uptake
lacA	Lactose	8	Thiogalactoside transacetylase
lacI	Lactose	8	Regulator gene
lacO	Lactose	8	Operator locus
lacP	Lactose	8	Promoter locus
lacY	Lactose	8	Galactoside permease (M protein)
lacZ	Lactose	8	β-Galactosidase
mot	Motility	42	Flagellar paralysis
mutL	Mutator	93	Generalized high mutability (AT → GC)
pabB	*p*-Aminobenzoate	40	Requirement
pil	Pili	98	Presence or absence of pili (fimbriae)
plsA	Phospholipid	11	Glycerol-3-phosphate acyltransferase
pnp		68	Polynucleotide phosphorylase
polA	Polymerase	85	DNA polymerase I
polB	Polymerase	2	DNA polymerase II
ptsG	Phosphotranserase	24	Catabolite repression system
purA	Purine	93	Adenylosuccinic acid synthetase
purB	Purine	25	Adenylosuccinase
pyrB	Pyrimidine	95	Aspartate carbamoyltransferase
pyrD	Pyrimidine	21	Dihydroorotic acid dehydrogenase
recA	Recombination	58	Ultraviolet sensitivity and competence for genetic recombination
recB	Recombination	60	Ultraviolet sensitivity, genetic recombination; exonuclease V subunit
recC	Recombination	60	Ultraviolet sensitivity, genetic recombination; exonuclease V subunit
relA	Relaxed	59	Regulation of RNA synthesis
rpoB	RNA polymerase	89	RNA polymerase: β subunit (*rif* gene)
rpsL	Ribosomal protein, small	72	Ribosomal protein S12 (*strA* gene, streptomycin resistance)
serA	Serine	62	3-Phosphoglyceric acid dehydrogenase
serO	Serine	20	Operator locus
serS	Serine	20	Seryl transfer RNA synthetase
speA	Spermidine	63	Arginine decarboxylase
speB	Spermidine	63	Agmatine ureohydrolase
speC	Spermidine	63	Ornithine decarboxylase
supB	Suppressor	15	Suppressor of *ochre* mutations
supE	Suppressor	15	Suppressor of *amber* mutations (*su*-2)
thrA	Threonine	0	Aspartokinase I-homoserine dehydrogenase I complex
thrB	Threonine	0	Homoserine kinase
thrC	Threonine	0	Threonine synthetase
trpA	Tryptophan	27	Tryptophan synthetase, A protein
trpB	Tryptophan	27	Tryptophan synthetase, B protein
trpC	Tryptophan	27	*N*-(5-Phosphoribosyl) anthranilate
trpD	Tryptophan	27	Phosphoribosyl anthranilatetransferase
trpE	Tryptophan	27	Anthranilate synthetase
trpO	Tryptophan	27	Operator locus
tyrA	Tyrosine	56	Chorismate mutase T-prephenate dehydrogenase
tyrT	Tyrosine	27	Tyrosine transfer RNA_1 (*su*-3 gene; amber suppressor)
ubiA	Ubiquinone	90	4-Hydroxybenzoate → 3-octaprenyl 4-hydroxybenzoate
uvrA	Ultraviolet	91	Repair of ultraviolet radiation damage to DNA, UV endo-nuclease
valS	Valine	95	Valyl-transfer RNA synthetase

---

[a] This list contains 126 of more than 1027 genes that had been mapped by 1983. (Bachmann, B. J. (1983) *Bacteriol. Rev.* **47**, 180–230). Their positions are shown diagrammatically in Fig. 26-4.

D. Mapping of Chromosomes

Let us now consider how knowledge of bacterial sex factors and of phage λ permitted the mapping of bacterial chromosomes. Together with the use of restriction endonucleases these techniques gave us the first precise physical maps of bacterial chromosomes and pointed the way toward the determination of complete genome sequences.

1. The Chromosome Map of *E. coli*

There are about 4,639,221 nucleotide pairs in the circular DNA molecule that is the chromosome of *E. coli* strain K-12. We now know the complete sequence, which includes all of the individual genes that are present.[74] However, our first knowledge of the location of these genes in the chromosome depended upon construction of a **linkage map** (Fig. 26-4A). Construction of this map, with 126 genes, began with the study of nutritional auxotrophs whose defective genes are located at many points on the chromosome. By 1983, 1027 genes had been mapped. In 1997, when the complete nucleotide sequence became known, 4288 protein coding genes could be recognized.[74] The map in Fig. 26-4 was established 30 years earlier, largely by use of interrupted bacterial mating.[79,80] In this procedure *Hfr* cells carrying specific mutations are mixed with wild-type F cells, and conjugation is allowed to proceed for a certain length of time. Then the cells are agitated violently, e.g., in a Waring blender. This breaks all of the conjugation bridges and interrupts the mating process. Mating is interrupted at different times, and the recipient bacteria are tested for the presence of genes transferred from the donor strain. Using this technique it was found that complete transfer of the chromosome takes ~100 min at 37°C, and that the approximate location of any gene on the chromosome can be determined by the length of time required for transfer of that gene into the recipient cell. It is a little more complex than this. Because complete chromosome transfer is rare, substrains of *E. coli* K-12 with an F agent integrated at different points were used. With certain F factors those genes lying clockwise around the circle in Fig. 26-4 immediately beyond the point of integration are transferred quickly and with high frequency.

The **time-of-entry map** in Fig. 26-4A is based not only on interrupted matings but also on the use of **transduction** by bacteriophage P1.[76,79] Transduction by phage permits the tranfer of a short fragment of DNA, about 2 min in length, on the *E. coli* map. Joint transduction, i.e., joint incorporation of two genes into the chromosome of the receptor, occurs with a frequency related to the map distance between these two genes. Thus, finer mapping was done within many segments of the *E. coli* chromosome. Meanings of the gene symbols used in the figure are given in Table 26-1. Similar maps were prepared for *Salmonella typhimurium* and *Bacillus subtilis*.

2. Restriction Endonucleases

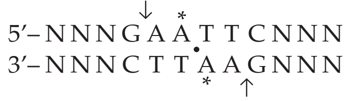
Many of the procedures for cloning genes, synthesizing more copies of a DNA (“amplifying” the DNA), making genetic maps, and generating mutants, are dependent upon **restriction endonucleases**. The name comes from a property of bacteria, which often can digest and destroy DNA of invading viruses or DNA that has been injected during mating with a bacterium of an incompatible strain. Investigation of this phenomenon, known as **restriction**, revealed that the DNA of viruses that are able to replicate within a particular host is *marked* in some fashion at specific sites in the molecule. The marking often consists of the presence of methyl groups. Properly methylated DNA is not degraded, but unmethylated DNA is cleaved by a highly specific endonuclease at the same sites that are normally methylated. Each species of bacteria (and often an individual strain within a species) has its own restriction enzymes. Restriction enzymes are very specific and cut DNA chains at unique base sequences. Three types are recognized.[80a]

Type I restriction enzymes, such as those encoded in the chromosome of *E. coli*, are large 300- to 400-kDa proteins composed of at least three kinds of polypeptide chain. They bind at specific sites of a foreign DNA and apparently cleave the chain randomly nearby. They require ATP, Mg^{2+}, and *S*-adenosylmethionine and have the unusual property of promoting the hydrolysis of large amounts of ATP.[81,82] The significance of these properties is still unknown.

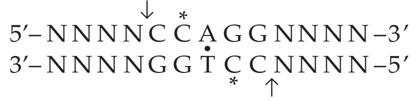
The type II restriction endonucleases, which are the ones most widely used in molecular biology, are relatively small 50- to 100-kDa monomeric or dimeric proteins. About 2400 different enzymes with 188 different specificities had been isolated by 1995.[83,84] The sites of attack, in most instances, are nucleotide sequences with a twofold axis of local symmetry.[85] For example, the following sites of cleavage have been identified for two restriction endonucleases encoded by the DNA of R-factor plasmids of *E. coli* and for a restriction enzyme from *Hemophilus influenzae*. In the diagrams ↓ are sites of cleavage, * are sites of methylation, and • are local twofold axes (centers of palindromes); N can be any nucleotide with a proper base pairing partner.

Restriction enzymes often create breaks in each of the two strands in positions symmetrically arranged around the local twofold axis. This is what we might expect of a dimeric enzyme that binds in the major or minor groove of the double helix, each active site

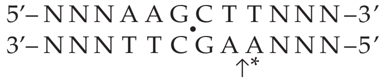
E. coli R factor (*Eco*RI)



E. coli R factor (*Eco*RII)



H. influenzae (*Hin*dIII)



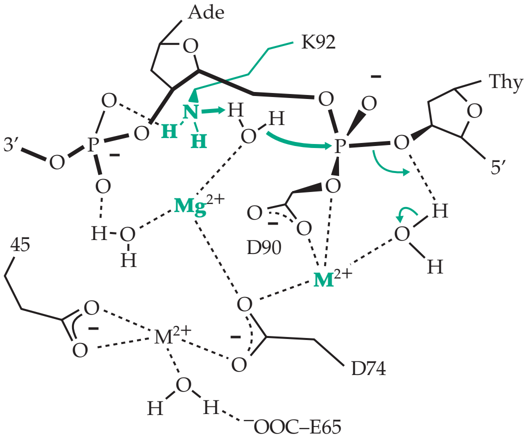
attacking one of the polynucleotide chains. In fact, the two 277-residue subunits of the *Eco*RI enzyme[86] bind primarily in the major grooves of the DNA, one active site on each strand. Each recognition unit makes 12 hydrogen bonds to the DNA. Each base pair forms two of these hydrogen bonds. Four arginines and two glutamates participate.[87] This provides a net charge of +2, which may be important for electrostatic interaction of the protein with backbone phosphate groups of the DNA. The binding of the protein affects the conformation of the DNA, widening the major groove from that in B DNA, and causing a torsional kink with some unwinding of the double helix.[88] It appears that the specificity for the GAATTC hexanucleotide is in part a result of direct complementary interactions between functional groups in the major groove (see Fig. 5-3), bound water molecules, and amino acid side chains from the enzyme (Fig. 26-5). Methylation of the 6-amino groups of the adenines in the center of the recognition sequence prevents cleavage by the *Eco*RI endonuclease, but appears to alter the interaction with the protein only slightly.[89]

Although they often share little sequence similarity and have quite different specificiities, many restriction enzymes have similar three-dimensional structures as well as mechanisms of action. This is true for the *Eco*RI, *Bam*HI (Fig. 26-5),[83,90] *Eco*RV,[91,91a] and *Cfr*10I enzymes,[84] and presumably many others. The specifically shaped and tightly packed active sites in the enzyme–substrate complexes ensure specificity. For example, the *Eco*RV endonuclease cleaves DNA at its recognition site at least a million times faster than at any other DNA sequence.[91] As mentioned in Chapter 12, restriction endonucleases require a metal ion, preferably Mg^{2+}, and probably act via a hydroxyl ion generated from $Mg^{2+}–OH_2$ at the active site. Three conserved active site residues, Asp 91, Glu 111, and Lys 113, in the *Eco*RI endonuclease interact with the DNA near the cleavage site. Lys 113 is replaced by Glu 113 in the *Bam*HI enzyme.[83,90]

The corresponding conserved residues in the smaller *Eco*RV enzyme are Asp 74, Asp 90, and Lys 92. They are shown in the following diagram that represents one of several possible metal-ion dependent mechanisms.[91–93] The metal-coordinated hydroxyl ion is generated by proton transfer to the $–NH_2$ group of Lys 92 and carries out an in-line attack on the backbone phospho group of thymidine at the cleavage point. At least two metal ions are needed, and three may be present, as shown in this diagram from Sam and Perona.[93]

TABLE 26-2
Some Commonly Used Restriction Endonucleases, Their Sources, and Cleavage Sites

Enzyme	Source	Cleavage Site
*Alu*I	*Arthrobacter luteus*	5' - - AG↓CT - - 3'
*Bam*HI	*Bacillus amyloliquefaciens H*	G↓GATCC
*Bcl*I	*Bacillus caldolyticus*	T↓GATCA
*Bgl*II	*Bacillus globigii*	A↓GATCT
*Cfr*10I	*Citrobacter freundii*	Pu↓CCGGPy
*Eco*RI	*Escherichia coli*	G↓AATTC
*Eco*RV	*Escherichia coli*	GAT↓ATC
*Hae*III	*Haemophilus aegypticus*	GG↓CC
*Hind*III	*Haemophilus influenzae*	A↓AGCTT
*Kpn*I	*Klebsiella pneumoniae*	GGTAC↓C
*Mbo*I	*Moraxella bovis*	↓GATC
*Pst*I	*Providencia stuartii*	CTGCA↓G
*Sal*I	*Streptomyces albus*	G↓TCGAC
*Sau*3AI	*Streptococcus aureus*	↓GATC
*Sfi*I	*Streptomyces fimbriatus*	GGCCNNNN↓NGGC



Restriction enzymes that cleave DNA at a large number of specific sequences are available commercially. A few are listed in Table 26-2. Another group of restriction enzymes have similar recognition sequences but cut the dsDNA that they recognize at a specific neighboring site rather than within the recognition sequence. An example is *Fok*I, which recognizes the nonpalindromic $\begin{matrix}\text{GGATG}\\\text{CCTAC}\end{matrix}$, but cuts the chains 9 and 13 base pairs to the right. This enzyme has been used by Szybalski and associates to devise a system for cutting ssDNA precisely at a desired point and converting it to ds fragments.[95]

3. Restriction Mapping

The calibration of the *E. coli* genetic map in minutes was a temporary expedient. It was followed by **physical maps** expressed directly as micrometers of DNA length (total length ~1.6 mm) or thousands of nucleotide units (kb). A physical map obtained by **restriction enzyme mapping** is shown in Fig. 26-4B. To obtain this map DNA fragments were prepared using specific restriction endonucleases (Section E, 1).

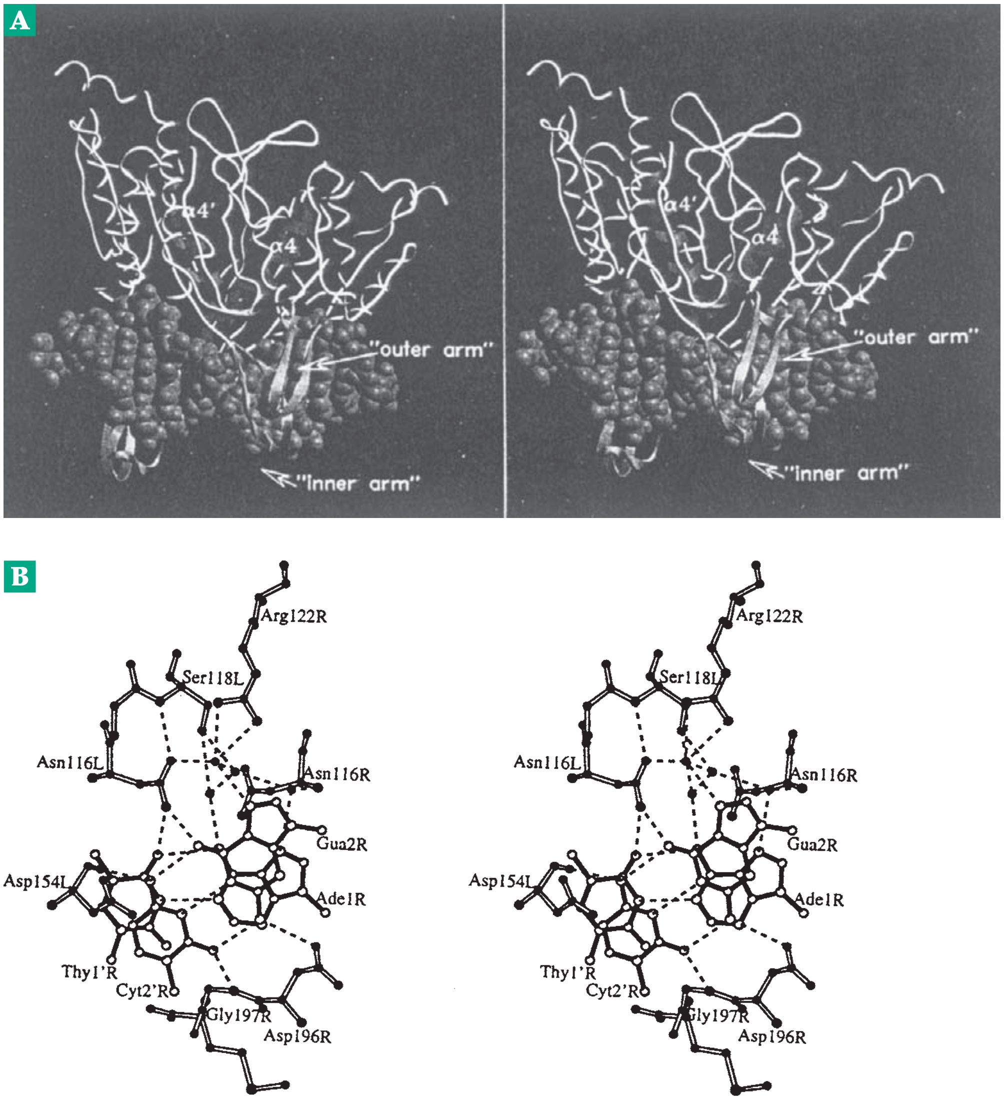


Figure 26-5 (A) Stereoscopic ribbon drawing of the dimeric *Eco*RI restriction endonuclease in a complex with DNA. The equivalent helices marked α4 and α4' point into the major groove of the DNA double helix while the inner and outer "arms" wrap around the DNA. From Bozic *et al.*[84] based on coordinates of Kim *et al.*[94] (B) Stereoscopic view of the two base pairs T•A and C•G of the right end of the recognition motif 5'-CGATCC-3' (Table 26-2) bound to the *Bam*HI restriction endonuclease. The third base pair C•G lies below the two that are shown. Notice the numerous hydrogen bonds, some of which bind atoms of the DNA directly to atoms of the protein and also hydrogen bonds to water molecules (filled circles). The tight packing of complementary charged and dipolar groups of protein, nucleic acid, water, and Mg^{2+} ions (not seen in this drawing) thoughout the complex accounts for the high specificity of these enzymes. From Newman *et al.*[83] Courtesy of Aneel Aggarwal.

The fragments were aligned using genetic markers, and their lengths were estimated by their electrophoretic mobilities.[77] The time-of-entry map has been added on a distorted scale in Fig. 26-4B. There are, on the average, 46.4 kb of DNA per minute, but this amount varies around the chromosome between 38 and 61 kb / min.

Mapping with restriction endonucleases was for many years an essential step in determination of the complete sequence of a piece of DNA. To make a restriction map the DNA, which may have been cut from a chromosome by a restriction endonuclease, is cloned. This permits isolation of a large amount of the DNA, which is then cut by other restriction endonucleases with differing specificities. Overlapping fragments resulting from the cleavages by single restriction enzymes are ordered to provide a map such as that of yeast mitochondrial DNA shown in Fig. 5-48. Any piece from the mapped DNA can now be cloned, and the exact sequence determined. The development of pulsed-field electrophoresis (Chapter 5), with its ability to separate DNA fragments 2000 kb or greater in length, allows restriction mapping with enzymes that cut at rare intervals to give very large fragments. For example, the *Not*I restriction endonuclease cuts the 4.7 Mb *E. coli* K12 genome into 22 fragments that were used to construct the complete restriction map of Fig. 26-4B.[77] Sequences of many viruses and plasmids, mitochondrial and plastid DNAs, and several bacterial genomes have been determined by use of restriction mapping and sequencing of the restriction fragments. Restriction fragment patterns have also been important to determine eukaryotic genome maps including the first genetic linkage map,[96] the first physical map of the human genome,[97] and the complete human genome sequences (Section G).

4. Electron Microscopy

Physical mapping by electron microscopy has been applied to bacteriophage, which can be obtained with large deletions in various parts of the genome.[98,99] The method can also be applied to cloned pieces of DNA. DNA might be isolated from two different phage strains, for example, from wild-type λ and from a mutant phage with a particular gene or genes deleted. The λ DNA can be denatured readily and separated into *r* strands and *l* strands by isopycnic centrifugation. If the isolated *l* strand of one strain is mixed with the *r* strand of another strain and annealed, a double-stranded DNA will be formed and, if there is a deletion in one strain, the homologous region in the normal λ DNA will form a single-stranded loop that can be visualized in the electron microscope. Figure 26-6 shows an example of a micrograph of such a **heteroduplex** molecule with a deletion loop and also a "bubble," where a segment of nonhomologous

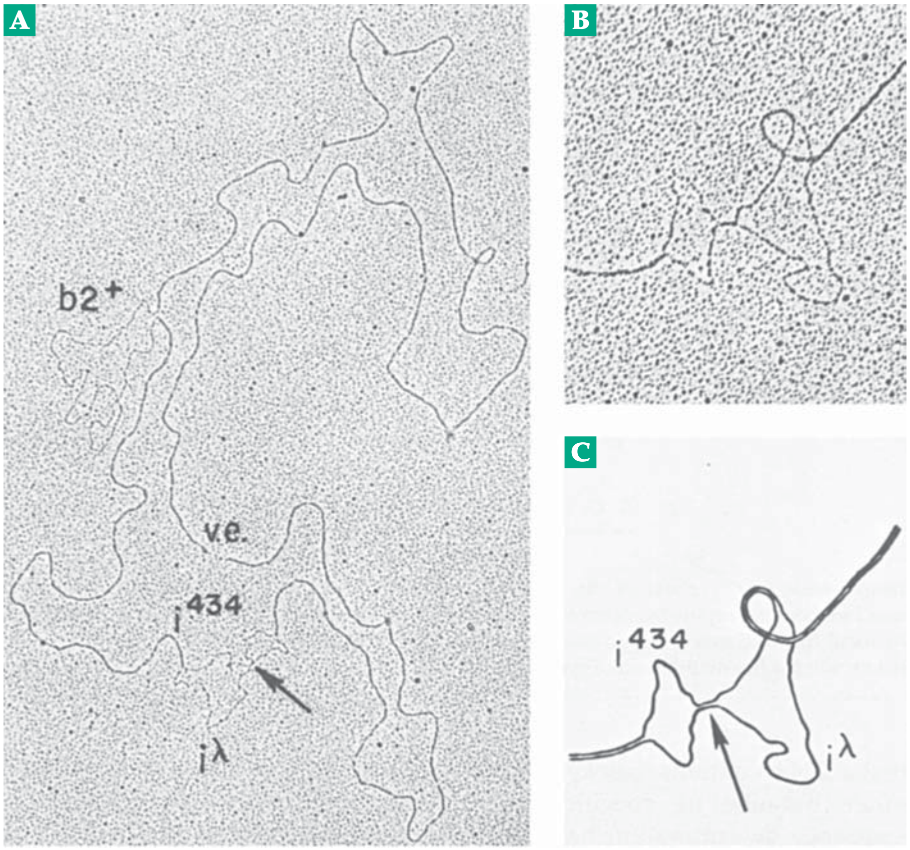


Figure 26-6 (A) Electron micrograph of a heteroduplex DNA molecule constructed from complementary strands of phages λb2 and λ*imm*434. In λb2 a segment of λ DNA has been deleted producing a deletion loop (labeled b2) and in λ*imm*434 a piece of DNA from phage 434 has been substituted for λ DNA resulting in a "nonhomology bubble" (labeled i^{434} / i^{λ}). The vegetative (cohesive) ends of the DNA are labeled v.e. (B) Enlargement of the nonhomology bubble. (C) Interpretative drawing of view in (B). Arrow marks a short (20–150 nucleotide) region of apparent homology. From Westmoreland, Szybalski, and Ris.[98]

DNA has been substituted in one strand.[98] Since distances can be measured accurately on the electron micrographs, rather precise (±50–100 bp) physical maps can be obtained. The chromosome map of phage λ was mapped in this way initially; now its complete nucleotide sequence is known. Another electron microscopic method is useful for location of AT-rich regions that denature readily. In a suitable concentration of formamide these regions melt to form visible single-stranded **denaturation loops** similar to the bubbles in Figure 26-6.

An important technique is to hybridize pieces of mRNA with DNA. If this is done with denatured (single-stranded) DNA and processed mRNA, which has been transcribed from genes with intervening sequences, the intervening sequences will form single-stranded loops in the DNA–RNA hybrid. A related technique depends upon the increased stability of DNA–RNA hybrids in high concentrations of formamide. Under these conditions an RNA segment will hybridize with its complementary strand of the DNA duplex displacing the other strand of DNA, which then appears as a visible **R-loop**. Intervening sequences appear as undisturbed DNA duplexes.

5. Optical Mapping

The long DNA molecules of bacterial or eukaryotic chromosomes are easily broken by vigorous stirring. However, if handled carefully very large fluorescently stained DNA molecules of 0.4–1.4 Mb lengths can be stretched out on a glass surface, and their lengths measured by optical microscopy. The technique depends upon binding of one of the ends of the DNA to the glass surface. If biotin is covalently attached to one end of the DNA, it will bind to a steptavidin-coated plate. However, unaltered DNA also binds to a vinyl-silane or a trichlorosilane coating under suitable conditions.[100,101] If DNA is incubated with a submerged silanized coverslip, which is then pulled out of the liquid mechanically at a constant speed, the DNA molecules are "dyamically combed" so that they are aligned for easy observation and measurement.[101] DNA can be transferred directly from agarose gels used for electrophoretic separation to the plates.[102] Used in combination with restriction enzymes these procedures allow rapid automated construction of physical maps. An example of whole-genome optical mapping is provided by the restriction map of the genome of the radiation-resistant bacterium *Deinococcus radiodurans*. The genome consists of two circular DNA molecules of 2.6 and 0.415 megabases and a smaller 176-kilobase DNA.[103] These were mapped without the laborious subcloning required by conventional restriction mapping. For example, the *E. coli* restriction map required analysis of 3400 phage clones. Mapping of eukaryotic chromosomes has involved additional methods which are discussed in Sections E and G,1. These incluse **radiation hybrid mapping**,[104] use of **meiotic recombination**, identification of **restriction fragment length polymorphisms** (**RFLPs**; described in Section E,7), and use of **expressed sequence tags** (**ESTs**), short DNA sequences deduced from mRNA molecules transcribed from the DNA.[105,106]

E. Cloning, Modifying, and Transferring Genes

A true revolution in biology and in medicine is in progress as a result of our ability to clone, sequence, mutate, and manipulate genes at will. Methods of sequence determination are discussed in Chapter 5 as is the laboratory synthesis of oligonucleotides and of complete genes. Both of these techniques are essential to present-day genetic engineering as are the techniques of cloning,[99,107–113] which are considered in this section.

A diploid cell contains only two copies of many genes, and these two copies are often not identical. In a gram of any tissue, which may contain 10^9 cells, we will have only ~1 ng of a 1 kb (1 kilobase) gene. To isolate this gene we would have to fish it out from a huge excess of other genes. Present-day cloning techniques offer a way to locate the gene, increase its quantity by many orders of magnitude, learn its sequence, induce any desired mutations at any points, and transfer the gene into other organisms in such a form that it can be expressed, i.e., be transcribed and induce synthesis of proteins. The methods have taken years to develop and continue to be improved. Only some basic procedures and concepts are described here. Numerous manuals[99,108,109,111,114–118] as well as commercial "kits" are available.

1. Joining DNA Fragments

The cloning and manipulation of genes usually depends upon the precise cutting of DNA into discrete fragments by restriction endonucleases. Many restriction enzymes generate **cohesive ends** (sticky ends). Thus, *Eco*RI produces DNA fragments with the single-stranded "tails" shown here at the 5'-ends of the cut duplexes:

$$\begin{array}{lr} 5'\text{-----}\,\underset{\bullet}{\mathrm{G}}\ 3' & 5'\mathrm{pA\,A\,T\,T}\,\underset{\bullet}{\mathrm{C}}\,\text{-----}\ 3' \\ 3'\text{-----}\,\overset{\bullet}{\mathrm{C}}\ \mathrm{T\,T\,A\,Ap}\ 5' & 3'\,\overset{\bullet}{\mathrm{G}}\,\text{-----}\ 5' \end{array}$$

These cohesive ends can be used to join together different restriction fragments. It is easy to see that the complementary single-stranded tails can form base pairs to regenerate the original hexanucleotide sequence cleaved by *Eco*RI. There will still be nicks between G and A at the specific cleavage points, but these nicks can be closed enzymatically using DNA ligase. Thus, the original DNA cleaved by *Eco*RI can be reformed, or another piece of DNA that also has tails generated by *Eco*RI can be grafted onto an end.

Many of the other enzymes in Table 26-2 also form cohesive ends. Five of them (*Bam*HI, *Bcl*I, *Bgl*II, *Mbo*I, and *Sau*3A) have at the center of their recognition sites the same tetranucleotide: GATC. Enzymes *Sau*3A and *Mbo* I are called **isoschizomers** because they have just the same 4-base recognition sequence and also yield the same restriction patterns. Notice that they will both cut all of the *Bam*HI, *Bcl*I, or *Bgl*II sites, but *Bam*HI and *Bgl*II will not cut all *Mbo*I or *Sau*3A sites. However, cohesive ends made by any of these enzymes can be joined. The gaps left during the joining of certain of the fragments can be ligated enzymatically. *Sau*3A will cut at either methylated or unmethylated sites but *Mbo*I will not cut at methylated sites.

Two enzymes (*Kpn*I and *Pst*I in the list in Table 26-2) form 3'-cohesive ends rather than 5'-cohesive ends. In addition, there are three (*Alu*I, *Eco*RV, and *Hae*III) that cut at the local twofold axis; they form no cohesive ends but leave **blunt ends** (flush ends). Blunt end fragments are also much used in genetic engineering. “Linkers” that provide cohesive ends can be added.[119] The *Sfi*I endonuclease cuts between two 4-bp palindromes in a 13-bp recognition sequence (Table 26-2).[120]

Some useful enzymes. Several enzymes of use in cloning[121] are listed in Table 26-3. The detailed chemistry of most of these is discussed in Chapter 27. Among these are the **ligases** that allow DNA fragments to be joined. They act on DNA strands with adjacent 3'-OH and 5'-phosphate termini. The *E. coli* ligase seals single stranded nicks using NAD^+ as an energy source (Eq. 27-5). It is therefore able to ligate DNA fragments with cohesive ends. The T4 DNA ligase, which is obtained from *E. coli* infected with phage T4, not only can seal nicks but can ligate pieces of DNA with blunt ends. Its activity is linked to cleavage of ATP. If two DNA strands are joined, but with gaps in one or both strands, the gaps can be filled efficiently by the 109-kDa **DNA polymerase I** from *E. coli*. Most often the 76-kDa **Klenow fragment**, which is lacking the 5' → 3' exonuclease activity, is used. **T4 DNA polymerase** has similar properties.

A problem with DNA fragments with cohesive ends is that they spontaneously form closed circles, a process that may compete with a desired joining to another piece of DNA. One solution to this problem is to hydrolyze off the 5'-phosphate groups with an **alkaline phosphatase** (Chapter 12). This prevents formation of covalently closed circles. However, pieces of DNA that retain their 5'-phosphate groups can be ligated to these dephosphorylated pieces. **T4 polynucleotide kinase**[122] can be used to put a phospho group back onto the 5' end of a chain. A ^{32}P end label can be added to such a polynucleotide using ^{32}P-labeled ATP.

Forming homopolymeric tails. Chromosomal DNA may be cleaved with restriction enzymes that leave blunt ends or it may be cleaved randomly by shearing. In either case the blunt ends can be treated first with **λ-exonuclease**, then with **terminal deoxynucleotidyl transferase**[123] isolated from calf thymus. The exonuclease treatment cuts off a few nucleotides from the 5' termini leaving short single-stranded 3'–OH termini. The terminal transferase, a polynucleotide polymerase that acts on ssDNA, is nonspecific and requires no template. Using an appropriate nucleotide triphosphate, it will add a single-stranded tail of either deoxyribonucleotides or ribonucleotides to the exposed 3' termini of a polynucleotide of three or more residues. If deoxyATP is used, a 3' poly(dA) tail will be added to each 5' terminus. Such a poly(dA)-tailed DNA fragment can be annealed and ligated to DNA carrying poly(dT) tails. This approach has been used widely to insert a piece of DNA into a cloning vehicle. For example, a circular plasmid (Fig. 26-7) can be opened by a single cleavage with *Eco*RI or other suitable restriction enzyme. The opened plasmid is treated with exonuclease, and poly(dA) tails are added. The piece of DNA to be cloned is tailed with poly(dT). After annealing and ligation recombinant plasmids carrying the **passenger DNA** will be formed.

If DNA is cleaved with *Pst*I or *Kpn*I (Table 26-2), the resulting 3' cohesive ends can be extended with a poly(dC) tail. If the cloning vehicle also has a site for *Pst*I or *Kpn*I, it can be opened and poly(dG) tails can be applied. A useful feature is that after annealing, filling in the gaps, and ligation the original *Pst*I or *Kpn*I sites are restored. This provides for easy recovery of the cloned fragments (Fig. 26-8).

Preparing material for cloning. DNA may be prepared for cloning by (1) random cleavage by shearing or by enzymatic attack, (2) cleavage by one or more restriction endonucleases, (3) preparation of cDNA from mRNA, or (4) nonenzymatic chemical synthesis of DNA segments. The use of random cleavage has largely been replaced by cleavage with restriction enzymes. A major problem is the separation of the very large number of different restriction fragments formed from a large piece of DNA or from an entire genome. The creation of “libraries” of such fragments is described in Section 5. Considerable simplification comes from separation of individual eukaryotic chromosomes before the library is prepared. Careful purification of DNA to be used in cloning is helpful. This may be done by electrophoresis in agarose or polyacrylamide gels or using HPLC. One technique is to embed cells directly in a gel, to diffuse in proteases and restriction enzymes that lyse the cells and release the DNA, and cleave it, and then to conduct the electrophoresis (see Chapter 5, Section H,1). DNA fragments of very large size can be separated. In addition to isolation of a fragment to be cloned the cloning vehicle must be prepared. This often involves release of a plasmid by lysis of bacteria that carry it and isolation using a suitable column. Likewise, after the DNA has been cloned and the content of DNA has been increased by growing a large bacterial culture, the plasmids must be released and purified, and the cloned DNA excised with a restriction enzyme. Alternatively, the cloning vehicle may be a virus, which must be isolated and disrupted to release the DNA. Use of the polymerase chain reaction (PCR; Fig. 5-47) allows cloning and amplification of DNA fragments with a minimum of purification.

2. Cloning Vehicles (Vectors)

Many cloning vehicles, more commonly referred to as **vectors**, originated with naturally occurring, independently replicating plasmids or viruses (replicons). More recently artificial chromosomes have been developed as cloning vehicles. Plasmids and viruses have been extensively engineered to provide convenience and safety. A large number of specialized

TABLE 26-3
Some Enzymes Used in Molecular Cloning

Name	Source	Reaction
T4 DNA polymerase	*E. coli* infected with bacteriophage T4	5'→ 3' chain growth 3'→ 5' exonuclease
E. coli DNA polymerase I and Klenow fragment	*E. coli*	5'→ 3' chain growth 3'→ 5' exonuclease 5'→ 3' exonuclease (lacking in Klenow fragment)
Reverse transcriptase	RNA tumor viruses, e.g., avian myoblastosis virus	5'→ 3' DNA chain growth
Ribonuclease H	*E. coli*	Cuts RNA in DNA-RNA hybrid
Lambda and T7 exonucleases	Bacteriophages	
Bal 31 nuclease	*Alteromonas espejiano*, a marine bacterium	Degrades both 3' and 5' termini of dsDNA
T7 RNA polymerase	Bacteriophage T7	DNA-dependent RNA polymerase
Terminal deoxyribonucleotide transferase	Thymus gland, plants	Limited 5'→ 3' chain growth; template independent addition of tails to DNA fragments
T4 DNA ligase	*E. coli* carrying an engineered λ phage	Ligation of DNA, either blunt or cohesive ends; uses ATP
E. coli DNA ligase	*E. coli*	Ligation of DNA with cohesive ends; uses NAD^+ as energy source
RNA ligase	Bacteriophage T4	Ligation of RNA and DNA
T4 polynucleotide kinase	Bacteriophage T4	Phosphorylation of 5′-OH terminus of a polynucleotide (DNA or RNA)
*Eco*RI methylase	*E. coli*	Transfer CH_3 from *S*-adenosylmethionine to adenines in *Eco*RI sites

See Chapter 12 for general discussion of nucleases.

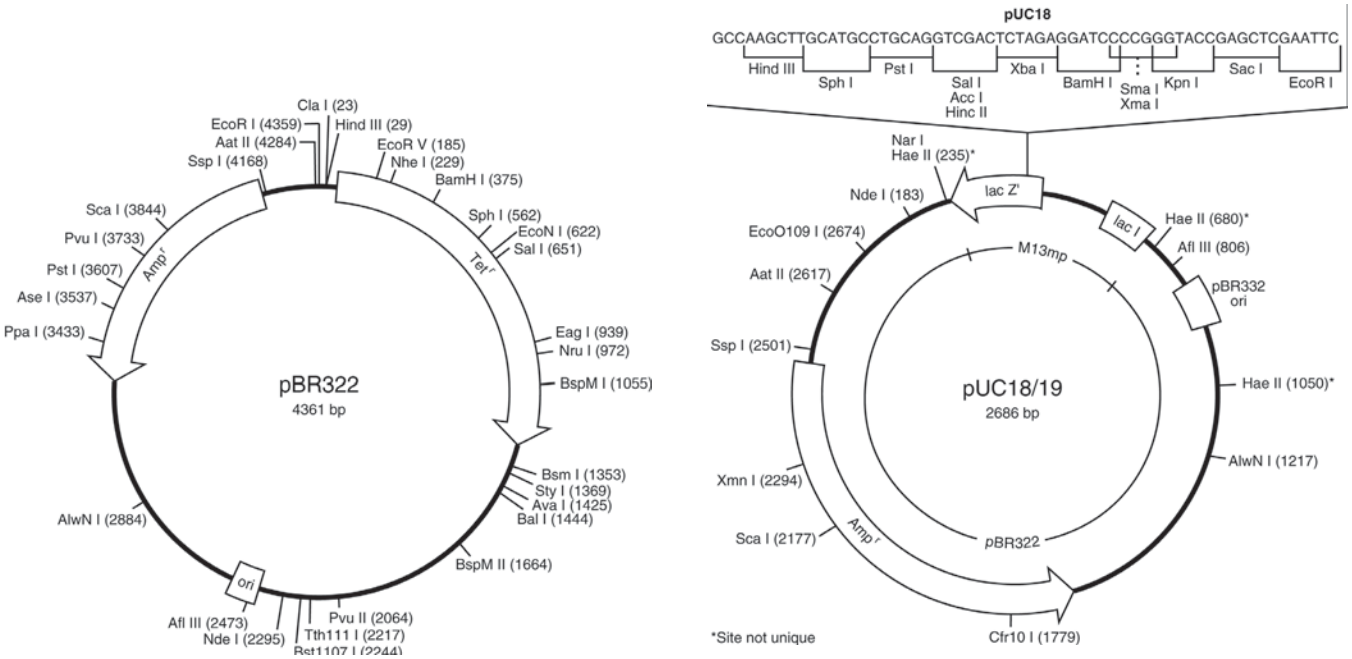

pUC18
GCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTC
Hind III
Sph I
Pst I
Sal I
Acc I
Hinc II
Xba I
BamH I
Sma I
Xma I
Kpn I
Sac I
EcoR I
Cla I (23)
EcoR I (4359)
Hind III (29)
Aat II (4284)
EcoR V (185)
Ssp I (4168)
Nhe I (229)
BamH I (375)
Sca I (3844)
Sph I (562)
Pvu I (3733)
EcoN I (622)
Sal I (651)
Pst I (3607)
Ase I (3537)
Ppa I (3433)
Amp^r
Tet^r
Eag I (939)
Nru I (972)
BspM I (1055)
pBR322
4361 bp
Bsm I (1353)
Sty I (1369)
Ava I (1425)
Bal I (1444)
AlwN I (2884)
BspM II (1664)
ori
Afl III (2473)
Pvu II (2064)
Nde I (2295)
Tth111 I (2217)
Bst1107 I (2244)
Nar I
Hae II (235)*
lac Z'
Nde I (183)
lac I
Hae II (680)*
EcoO109 I (2674)
Afl III (806)
M13mp
Aat II (2617)
pBR332
ori
Ssp I (2501)
pUC18/19
2686 bp
Hae II (1050)*
AlwN I (1217)
Xmn I (2294)
Sca I (2177)
Amp^r
pBR322
Cfr10 I (1779)
*Site not unique

Figure 26-7 Genetic map of cloning plasmids pBR322 and pUC18. Abbreviations: ori, origin of replication; Amp^r, ampicillin resistance gene; Tet^r, tetracycline resistance gene. Other abbreviations are for sites cleaved by specific restriction endonucleases, a few of which are defined in Table 26-2. The nucleotide sequence numbers and directions of transcription are also indicated. Reproduced by permission of Amersham Pharmacia Biotech Inc.

vehicles have been devised. Only a few will be described briefly here. Suitable books and manuals must be consulted for details.[5,99,108–110,114,115,121,123a]

Plasmids related to ColE1.

The small colicinogenic plasmid, ColE1 of *E. coli*, is attractive for cloning because there are 20 copies per bacterial cell. If chloramphenicol is added during the logarithmic phase of growth, the *E. coli* cells will make 1000–3000 copies. Thus, a relatively small culture of the bacteria containing the cloned DNA in a ColE1 plasmid will yield a large amount of the desired DNA. In the past one of the most widely used cloning vehicles[124] has been plasmid pBR322, which was derived from a close relative of ColEl. The genome size of the original plasmid was reduced by deletion of genes unnecessary for its successful replication. Transposons

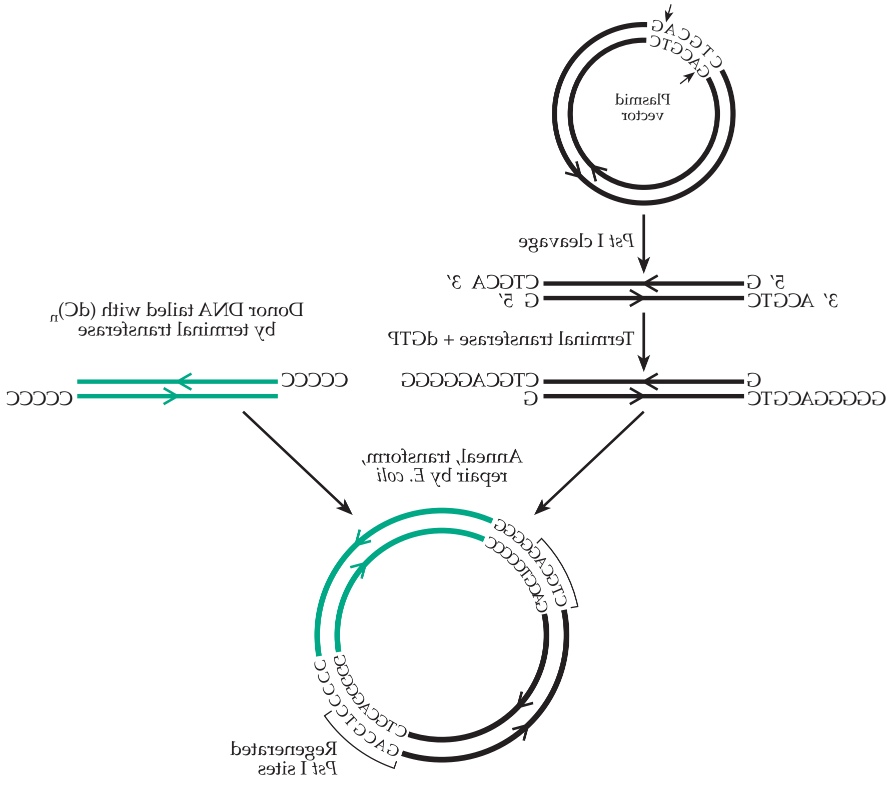

CTGCAG
GACGTC
Plasmid vector
Pst I cleavage
5' G
3' ACGTC
CTGCA 3'
G 5'
Terminal transferase + dGTP
G
GGGGGACGTC
CTGCAGGGGG
G
Donor DNA tailed with (dC)n by terminal transferase
CCCCC
CCCCC
Anneal, transform, repair by E. coli
Regenerated Pst I sites

Figure 26-8 Regeneration of a *Pst* I site by dG:dC tailing with terminal transferase. See Glover.[99] Arrowheads indicate 5′→3′ directions.

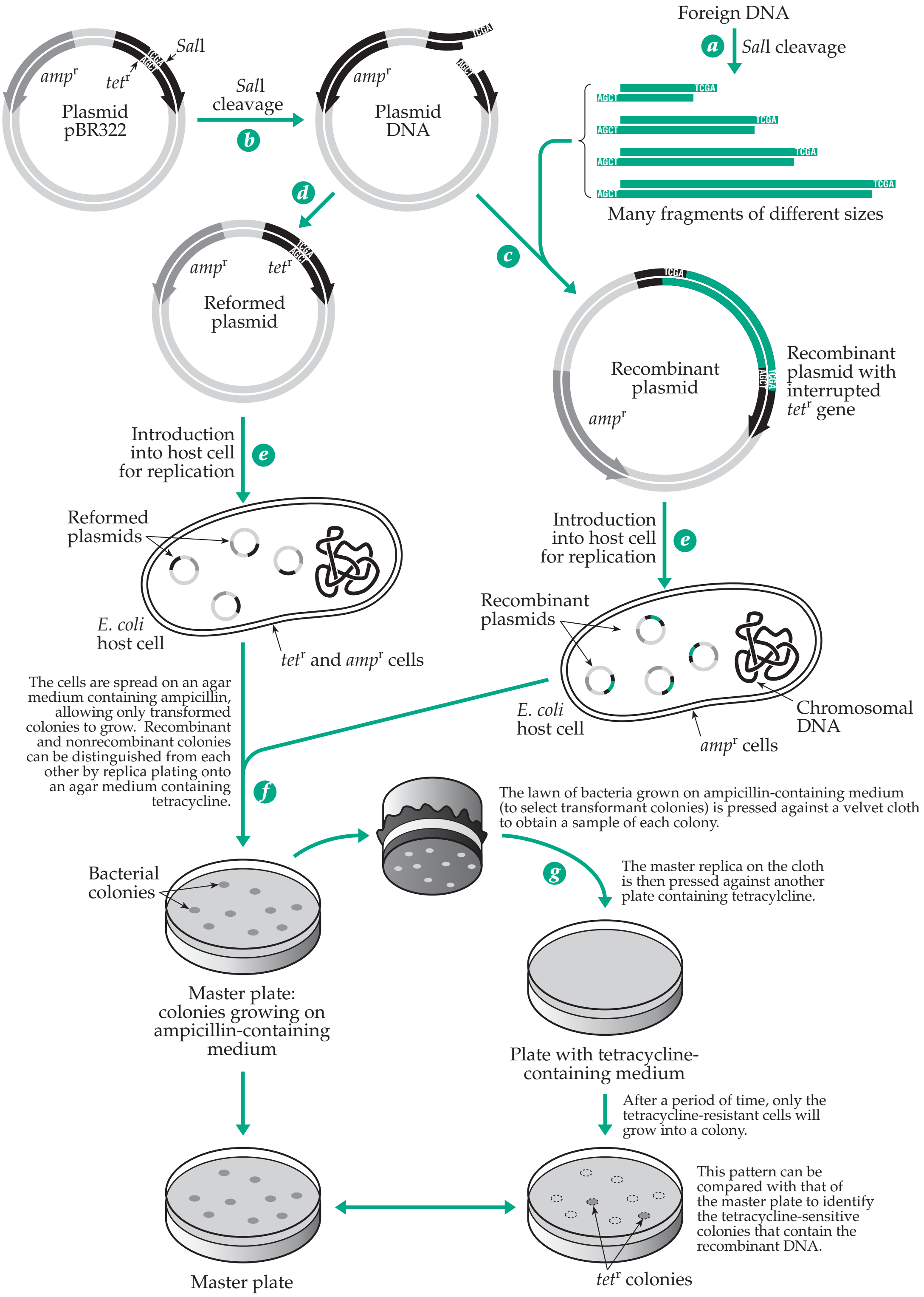
that might permit accidental transfer of DNA to other organisms were inactivated. Unneeded or undesirable restriction enzyme sites were eliminated, and useful restriction sites were introduced by point mutations. The resulting pBR322 has only one site of cleavage each for *Bam*HI, *Sal*I, *Pst*I, *Pvu*I, and *Eco*RI. These are at known positions in the 4363-nucleotide plasmid.[125]

Plasmid pBR322 contains two different antibiotic-resistance genes that were brought in from bacterial R-factors. These are used in selecting bacterial colonies that carry the desired recombinant plasmids. One of these is the β-lactamase gene, which confers resistance to ampicillin (Amp^r); the other provides resistance to tetracycline (Tet^r). Their positions are indicated in the pBR322 gene map (Fig. 26-7) as is the essential origin of replication (*ori*). The drug resistance genes are used as follows. If a unique restriction site such as that for *Bam*HI or *Sal*I that lies within the Tet^r gene is used to introduce the passenger DNA, the resistance to tetracycline is lost but that to ampicillin is retained. Thus, after incubation with the recombinant plasmids under conditions that favor their uptake by the host bacteria, the bacteria are plated onto an ampicillin-containing medium. Only those harboring the pBR322 plasmid with its Amp^r gene can grow. After these have produced small colonies, a replica plate is made on a tetracycline-containing medium. On this medium the desired recombinants do *not* grow because the Tet^r gene has been inactivated. This allows selection of colonies containing passenger DNA (Fig. 26-9). A further selection procedure is required to establish that the piece of DNA inserted into the recombinant plasmid is one that is desired.

Typical cloning procedure with pBR322. In much simplifed form the procedure might go as follows: (1) Purchase or isolate plasmid. (2) Cleave plasmid with *Bam*HI; heat at 70°C to inactivate the enzyme. (3) Treat with alkaline phosphatase to remove the 5'-phospho groups. (4) Mix with passenger DNA with cohesive ends generated by *Bam*HI, anneal, and join with DNA ligase. Although the resulting circular recombinant DNA contains a nick as a result of the missing 5'-phospho group, it will be taken up by bacteria and repaired. (5) Incubate joined DNA with cells of host *E. coli* that have been made permeable to DNA by treatment with Ca^{2+} ions. This type of transformation is called **transfection** and is widely used in cloning. (6) Plate transfected cells onto agar containing the first antibiotic, in this case ampicillin. (7) Make replica plate on medium containing second antibiotic, in this case tetracycline. (8) Screen selected colonies for desired DNA fragment. In one procedure a small sample from each of the selected colonies is placed onto spots on a nitrocellulose filter. Several colonies can be placed on one filter and the bacteria lysed, hybridized with a radioactive probe, and then viewed by autoradiography.

Selecting clones using β-galactosidase in pUC cloning vehicles. The newer **pUC** vehicles[126] contain the origin of replication and the ampicillin-resistance gene from pBR322. In addition, a segment of DNA from the *E. coli lac* operon (Fig. 28-2) has been grafted into an intergenic region (Fig. 26-7). It contains the lac control region as well as the coding sequence for the first 145 residues of β-galactosidase. Within bacteria containing the pUC DNA an N-terminal portion of β-galactosidase is synthesized. The specially designed host cell contains (in an episome) the gene for another defective β-galactosidase, one lacking the N-terminal portion. This defective enzyme, together with the N-terminal portions encoded in the pUC DNA, forms an active galactosidase. When the chromogenic substrate 5-chloro-4-bromo-3-indolyl-β-D-galactoside ("X-gal") and a suitable inducer, such as isopropylthio-β-galactoside (IPTG), are present, the unoccupied plasmid vehicles generate blue colonies. However, if passenger DNA is inserted within the galactosidase gene segment, formation of the enzyme will be disrupted and white colonies will appear. The restriction sites for cloning are placed in a **polylinker** near the 5' terminus of the galactosidase gene. The polylinker has been carefully engineered to maintain the correct reading frame and to avoid disruption of the galactosidase activity. It contains several different restriction sites for insertion of passenger DNA (Fig. 26-7). Such insertion does destroy the galactosidase activity allowing the user to detect the recombinant DNA from the white plaques.

Filamentous bacteriophages. An important series of cloning vehicles have been derived from the circular replicating forms of the filamentous bacteriophage M13 (Chapter 5; Fig. 7-7; Chapter 27).[127] Although the genome contains only a short intergenic region that can be deleted, up to 50 kb of passenger DNA can be inserted into these vehicles. Since long inserted sequences may be deleted spontaneously, M13 is most useful for cloning about 300- to 400-nucleotide chains. Many of the M13 vehicles also use the β-galactosidase blue-white screening technique. These modified viruses are highly infective, but the infected *E. coli* cells are not killed. Rather they produce large numbers of virus particles with single-stranded DNA representing one of the two DNA chains of the parental phage. These are widely used for sequencing by the chain-termination procedure of Sanger *et al.* (Chapter 5). The procedure requires a primer sequence. If M13 recombinants are sequenced, the primer consists of a synthetic oligonucleotide that can be annealed to the galactosidase gene fragment at its 3' end just in front of the DNA segment that is to be sequenced.


Foreign DNA
a
SalI cleavage
AGCT
TCGA
Many fragments of different sizes
ampr
tetr
SalI
Plasmid pBR322
SalI cleavage
b
ampr
Plasmid DNA
c
d
ampr
tetr
Reformed plasmid
Recombinant plasmid
ampr
Recombinant plasmid with interrupted tetr gene
Introduction into host cell for replication
e
Reformed plasmids
E. coli host cell
tetr and ampr cells
Introduction into host cell for replication
e
Recombinant plasmids
E. coli host cell
ampr cells
Chromosomal DNA
f
g
Bacterial colonies
Master plate: colonies growing on ampicillin-containing medium
Plate with tetracycline-containing medium
Master plate
tetr colonies

The cells are spread on an agar medium containing ampicillin, allowing only transformed colonies to grow. Recombinant and nonrecombinant colonies can be distinguished from each other by replica plating onto an agar medium containing tetracycline.

The lawn of bacteria grown on ampicillin-containing medium (to select transformant colonies) is pressed against a velvet cloth to obtain a sample of each colony.

The master replica on the cloth is then pressed against another plate containing tetracylcline.

After a period of time, only the tetracycline-resistant cells will grow into a colony.

This pattern can be compared with that of the master plate to identify the tetracycline-sensitive colonies that contain the recombinant DNA.

Figure 26-9 A classical scheme for cloning DNA in a pBR322 plasmid vehicle. A DNA sample is digested (step *a*) with one of the restriction endonucleases (e.g., *Sal*I), that cuts the cloning vehicle within the *Tet*^{r} gene (see Fig. 26-7). The plasmid is also cut with the *Sal*I restriction enzyme (step *b*). After mixing with the digested DNA sample and annealing, recombinant molecules are formed (step *c*). Some plasmids are reformed (step *d*). Both recombinant and reformed plasmids transform *E. coli* cells (step *e*). The transformed cells are plated on an agar medium containing ampicillin (step *f*). Only cells containing the ampicillin-resistance gene grow (step *g*). A replica plate is made and is pressed onto another plate containing tetracycline. The recombinant colonies do not grow on the medium because the tetracycline-resistance gene has been interrupted. By comparing the two plates recombinant colonies can be selected. These can be tested, using a suitable probe, to determine whether they carry a desired gene. After Atherly, Girton, and McDonald.[2]

Lambda cloning vehicles. Many cloning vehicles have been derived from the 48,502-bp[128] DNA from the temperate *E. coli* bacteriophage λ. The DNA from phage particles are taken up efficiently by *E. coli* cells, much more so than by transfection. The virus has a complex life cycle, which is discussed in Chapter 28. Within the phage head the λ DNA exists in a folded linear form with 12-base 5' cohesive ends (Fig. 28-11). After entrance into the bacterial cell the DNA cyclizes through its cohesive ends and is ligated by the *E. coli* ligase. Replication of the circular forms ensues. Later in the cycle, rolling circle replication (Eq. 27-7 and associated discussion) produces long concatamers with several phage genomes joined as a single chain. This DNA is "packaged" into new phage heads. As this is done, a nuclease cuts the concatamers at *cos* sites forming monomeric genomes with cohesive ends. This is the **lytic cycle** of the phage. In the alternative **lysogenic cycle** the DNA becomes integrated into the *E. coli* genome. Maintenance of the lysogenic state depends in part on gene *cI*, which encodes a repressor (see Chapter 28) that prevents expression of the genes required in the lytic pathway. Since only the lytic cycle is needed for cloning, it is convenient to place a cloning site within the *cI* gene. The screening of recombinant phage particles is done by examination of plaques. A phage without inserted DNA will be able to undergo both the lytic and lysogenic cycles and will form turbid plaques. However, if passenger DNA is inserted in the *cI* gene, the lysogenic cycle is prevented, and clear plaques are formed.

Of the ~50 genes present in native λ, only about half are necessary for replication in the lytic cycle. Thus, it is possible to delete about 1/3 of the genome to make room for more passenger DNA. However, to form mature phage particles the length of the DNA must be at least 75% of the native length. No more than 110% of the native amount may be present. The total DNA must fall between 38 and 53 kb in length. To accommodate these packaging requirements **replacement vectors** containing unnecessary "stuffer DNA" between two lambda "arms" are used. The unneeded stuffer piece has the same kind of restriction site or sites at each end so that it can easily be cut out and replaced by the passenger DNA. This permits cloning of DNA segments up to about 22 kb in length. Lambda vehicles have all been engineered to eliminate undesired restriction sites and to reduce the number of sites for *Eco*RI and other restriction enzymes commonly used for cloning. The widely used **Charon series**[129] have been further engineered so that they will grow only in strains of bacteria that cannot survive in the human intestinal tract. For example, amber mutations (Section B,6) are incorporated into genes needed for phage assembly, and the bacterial hosts must contain an amber suppressor gene. The bacteria are also nutritional auxotrophs with absolute requirements for thymidine and diaminopimelic acid in the medium. The latter compound is not found in the intestinal tract. The purpose of these alterations is to prevent the spread of recombinant DNA into the environment.

To sequence DNA carried in a lambda vehicle or to study it in other ways, it is often necessary to cut it with restriction enzymes, to prepare a restriction fragment map, and to subclone the fragments into a plasmid vehicle. Lambda vehicles, which will automatically transfer the passenger DNA into an M13 vehicle when propagated in a host carrying a special helper virus, have been devised.[130] The helper virus encodes proteins that recognize and cleave sequences that mark the initiation and termination of M13 DNA synthesis. These are used to mark the ends of the passenger DNA. As DNA synthesis occurs the displaced passenger DNA and M13 genes are excised, circularized, and converted into a replicating form of an M13 cloning vehicle.

Cosmid vehicles and in vitro packaging. Cosmids[131] are hybrids of a plasmid vehicle and phage λ. They contain the *cos* sites that are cleaved during packaging of λ DNA. A cosmid cleaved at a restriction site will form upon ligation a range of different sized DNA molecules that contain *cos* sites on both sides of a piece of passenger DNA, which may be up to 45 kb in length (Fig. 26-10). This can be cut at the *cos* sites and packed into heads using an *in vitro* packaging system. In this system the unassembled subunits of the phage particle are produced in special strains of bacteria and

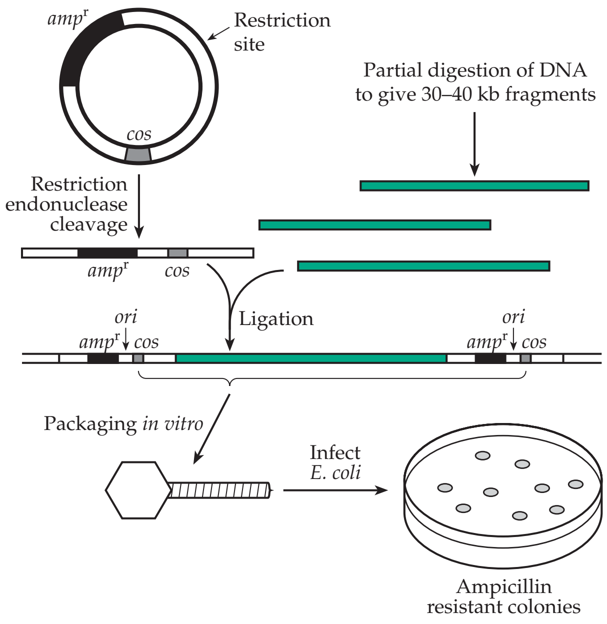

amp^r
Restriction site
cos
Partial digestion of DNA to give 30–40 kb fragments
Restriction endonuclease cleavage
amp^r
cos
Ligation
ori
amp^r
cos
ori
amp^r
cos
Packaging in vitro
Infect E. coli
Ampicillin resistant colonies

Figure 26-10 Cloning DNA in cosmids. See Glover.[99]

are allowed to assemble and package the cosmid DNA. Since cosmids contain relatively large pieces of DNA, they are useful for preparing sequentially overlapping clones that allow the investigator to "walk" along the DNA hunting for a target gene. Cosmid vehicles have been designed to allow both efficient genomic walking and restriction mapping.[132]

3. Expression of Cloned Genes in Bacteria

A major goal in recombinant DNA technology is the production of useful foreign proteins by bacteria, yeast, or other cultured cells. Protein synthesis depends upon both transcription and translation of the cloned genes and may also involve secretion of proteins from the host cells. The first step, transcription, is controlled to a major extent by the structures of promoters and other control elements in the DNA (Chapter 28). Since eukaryotic promoters often function poorly in bacteria, it is customary to put the cloned gene under the control of a strong bacterial or viral λ promoter. The latter include the λ promoter P_L (Fig. 28-8) and the *lac* (Fig. 28-2) and *trp* promoters of *E. coli*. These are all available in cloning vehicles.

It is often useful to create hybrid proteins fused to the *E. coli* β-galactosidase gene. If another gene is spliced in at either the N terminus or the C terminus of the galactosidase (*lac*Z) gene but is kept under control of the *lac* promoter, the resulting hybrid protein will have galactosidase activity, which can be used for screening. In addition, the hybrid protein will often react with antibodies directed against the protein whose gene is being cloned. Another kind of hybrid fuses the cloned gene to that of β-lactamase, for example at the *Pst I* site of plasmid pBR322 (Fig. 26-6). The β-lactamase activity will be gone, but the hybrid protein will be secreted because β-lactamase normally is secreted into the periplasmic space and its N-terminal signal sequence is now fused to the cloned protein. Engineering of a suitable site for cleavage by a protease can release the foreign protein in an active form. A variety of other **expression systems**, often using **reporter gene** products, have been developed.[133]

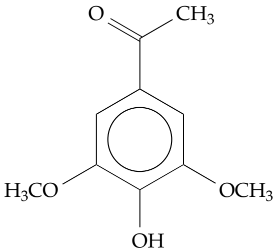
Bacteria often degrade foreign proteins by hydrolytic attack. One way in which such damage has been minimized is to clone multiple fused copies of the gene for the desired product. The resulting polyprotein may be resistant to degradation, and if the gene has been correctly engineered, may be cut apart by cyanogen bromide (Eq. 3-17) or by a specific protease.

For many years most cloning was done in *E. coli*, but cloning systems have now been developed for many other bacteria including *Bacillus* and other gram-positive bacteria, and also for yeast, insect cells, animals, and plants.[133–135]

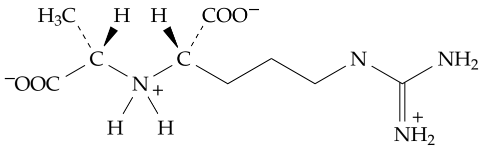
4. Cloning and Transferring Eukaryotic Genes

Eukaryotic genes cloned into bacterial plasmids are often poorly expressed. It is advantageous to clone such genes in eukaryotic cells, where the cutting and splicing of hnRNA to remove intervening sequences during formation of mRNA (Chapter 28) does occur. This permits expression of the cloned gene, something that is possible in bacteria only if cDNA that lacks the introns is cloned. The need for posttranslational modification of many proteins also interferes with expression in bacteria. Many methods of gene transfer and cloning have been developed.[136] The yeast *Saccharomyces cerevisiae* is often an ideal host for cloning. It grows rapidly in either its haploid or diploid stage (Chapter 1). Some strains carry a 2-µm 6.3-kb circular plasmid with 50–100 copies per cell.[109] This has been developed as a cloning vehicle. Recombinant plasmids can be used as **shuttle vehicles** for transferring genes cloned in *E. coli* into the yeast plasmid. Genes may also be cloned as minichromosomes, such as **yeast artificial chromosomes** (**YACs**). Artificial chromosomes contain origins of replication from yeast, human, or bacterial chromosomes as well as telomeres and centromeres[137] (see also Chapter 27). YACs have been widely used and became popular because they can accomodate 600 kbp or more of DNA.[137–139] However, their use in the human genome project resulted in serious problems of instability.[140] Bacterial artificial chromosomes (BACs), which accommodate only 200–300 kbp, **P1 artificial chromosomes** (**PACs**),[141,142] and human minochromosomes are more stable.[143] Another problem, which affects the use of yeast for production of eukaryotic proteins, is the tendency for poor removal of introns.

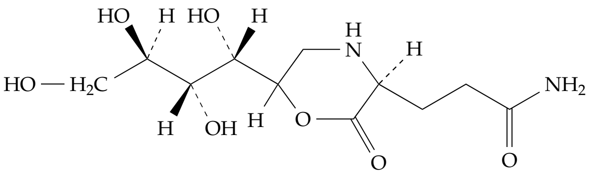
Plant genes. Much of the cloning in higher plant cells has made use of the **Ti plasmid** of *Agrobacterium tumefaciens*, a soil bacterium that enters wounds in dicotyledenous plants causing tumors known as **crown galls**.[144–146c] A related species *A. rhizogenes* harbors a similar plasmid that causes "hairy root" disease.[146c,147] The infecting bacteria respond to the synthesis of certain phenolic compounds such as **acetosyringone**, which are produced in plant wounds, by entering the plant cells.[148,149] Only bacteria carrying the Ti (tumor-inducing) plasmid cause tumors. The plasmid carries a 13-kb region



Acetosyringone, a compound that induces infection by *Agrobacterium* (transferred region) known as T-DNA that encodes enzymes for production of the auxin indoleacetic acid (Fig. 25-12), the cytokinin isopentenyl-AMP, and the compounds known as **opines**. These are reduction products of Schiff bases of amino acids and 2-oxoacids or sugars.



Octopine: derived from pyruvate and arginine
Histopine: derived from pyruvate and histidine[150]
Lysopine: derived from pyruvate and lysine
Nopaline: derived from 2-oxoglutarate and arginine
Leucopine: derived from 2-oxoglutarate and leucine



Agropine: derived from D-mannose and L-glutamine

The auxin and cytokinin, whose production is normally controlled, are now formed in large amounts and cause uncontrolled tumor growth. The opines are used by *Agrobacterium* as a unique source of energy and of metabolites for biosynthesis. The host plant cells, however, cannot catabolize the opines.

Upon entrance into a plant cell the T-region is excised from the Ti plasmid and can become integrated into the DNA of the host plant much as occurs during bacterial conjugation (Fig. 26-3).[151–154a] By deleting the genes for synthesis of auxin, cytokinin, and opines, the Ti plasmid loses its tumorigenic property but is still able to transfer genes into the plant genome. Fusion of an *E. coli* plasmid vehicle into the modified Ti plasmid creates a useful plant shuttle vehicle. Also see below under "physical methods."

Transferring genes with engineered animal viruses. There is special interest in transferring cloned DNA into human cells to correct genetic defects.[155] Transfer of human genes into other animals is also important for a variety of reasons. For example, human proteins of therapeutic value could be produced in animal milk. Animals can be engineered to have defects that mimic those in humans and which can then be studied in animals.[156]

A number of different viruses have been used to transfer foreign genes into eukaryotic cells and also to create stable plasmids for cloning. One of the first was the 5380 bp SV40. The relatively small size of SV40 limits the amount of DNA that can be incorporated. However, many of the functions of this virus can be performed by simultaneous infection with a **helper virus**, often an adenovirus which is itself defective. With the helper virus making proteins essential for SV40 replication, all but 85 base pairs at the origin of replication of SV40 can be deleted and replaced by other DNA. More recently engineered **adenoviruses** and **retroviruses** have been widely used. In both cases genes required for replication of the viruses have been deleted.[155] Adenovirus vehicles can carry 7 kbp or more of passenger DNA. They are efficiently taken up into endosomes and into the nucleus by both replicating and nonreplicating cells. The DNA is not incorporated into the genome of the host cell.[155,157,158] Replication-deficient avian **retroviruses**, or **retrotransposons** (Chapter 28), can carry up to 9 kbp of passenger DNA and are incorporated into the host genome by the process of homologous recombination (Chapter 27).[155,159,160] Both adenovirus and retroviral vehicles suffer from serious problems. The adenovirus proteins can induce fatal inflammatory reactions,[160a] and DNA transfer from the retroviral vehicles is inefficient. A nonpathogenic human parvovirus, the **adeno-associated virus** (AAV), is also being developed as a gene therapy vector.[161,161a,b] Another possibility is to engineer a **lentivirus**, even HIV, to provide efficient integration (Chapter 28).[162,162a] Safety concerns have delayed human testing of this possibility.

A group of insect viruses, the **baculoviruses**, are being used in cultured insect cells for large-scale production of proteins. The cloned genes are placed under the control of the promoter region of the gene that encodes the major viral coat protein.[135,163] The **baculovirus** vehicles can carry over 15 kbp of passenger DNA and may also be useful in human gene therapy.[164,165] Filamentous bacteriophages have been reengineered for the same purpose.[166]

Nonviral methods for DNA incorporation into the human genome may utilize **transposons** (Chapter 27)[166a] or mobile **group II introns** (Chapter 28).[166b,c]

Physical methods of gene transfer. Genes can often be transferred without the use of a cloning vehicle. This is especially important for certain plant cells, such as those of cereal grains, for which transfer of genes via the Ti plasmid has been difficult.[167] If DNA, which may be in a plasmid, is coprecipitated with calcium phosphate, it can often be taken up directly either by animal cells or by plant protoplasts.[168,169] Polycations also facilitate DNA uptake; cationic **liposomes** seem to be especially effective.[170] In the widely used **electroporation** technique a short electrical pulse of a few hundred volts / cm is applied to create transient pores in the plasma membrane through which the DNA can enter a cell.[108,171–175] Chromosomes can be transferred by cell fusion and either entire chromosomes, isolated DNA, or cloned DNA can be transferred into egg cells by microinjection.[176] Following transfer of DNA or of intact chromosomes, recombination sometimes leads to stable incorporation of some of the transferred genes into the host cell's genome.

A very important technique is the use of high-velocity microprojectiles shot from a particle gun. Spherical tungsten particles of ~0.5–1 µm diameter are coated with recombinant plasmids containing the genes to be transferred. The particles are then shot, using a gunpowder cartridge, into intact recipient cells. The particles penetrate cell membranes, mitochondria, and chloroplasts without serious damage if the number of particles is not too great.[177,178] The technique is very useful for genetic engineering of plants.[179]

5. Genomic Libraries

The human genome contains about 3 x 10^6 kb of DNA, about the average amount for a eukaryotic organism. If the entire genome is digested to completion with a restriction endonuclease, whose cleavage sites are distributed more or less randomly, the resulting restriction fragments constitute a "library" for that genome. If the average length of a fragment is 17 kb, about 1.8 x 10^5 unique fragments will be produced. To make a practical library these must be cloned into a suitable vehicle. Derivatives of phage λ or cosmids are most often used. The cloned fragments can then be packaged using the *in vitro* packaging system to form infectious phage particles, which can be propagated as plasmids in *E. coli* cells.

What is the probability that a given fragment among the total produced will be found in one of the recombinant phage in the library? From simple probability theory the number of clones that must be isolated and screened is given by Eq. 26-2.[180] Here N is the number of clones needed, p is the probability of

$$N = \ln(1-p)\ \ln(1-f) \qquad (26\text{-}2)$$

having the desired fragment in the library, and f is the fractional proportion of the genome represented by the fragment sought. For our example of a 17-kb fragment of the human genome this is 17 / 3000. From this equation one can calculate that to have a 99% chance of finding our fragment ($p = 0.99$) we need 8 x 10^5 clones.[99] What the equation does not show is that there will probably be some long fragments that cannot be cloned in the selected vehicle. These will be missing from the library.

Large numbers of clones obtained can be screened rapidly by colony hybridization using a labeled DNA probe. Thus, if it is desired to isolate a gene for a particular protein and some part of that protein has been sequenced, a synthetic DNA probe can be made. The phage containing the recombinant fragments can be plated, and after plaques form a nitrocellulose filter can be laid on the plate to form a replica. After release of DNA from the phage, denaturation by NaOH, and neutralization the single-stranded DNA fragments can be hybridized with the probe. Another screening method uses a probe carried in a plasmid that promotes homologous recombination between the probe sequence and restriction fragments with a similar sequence.[181] A problem that arises in preparing genomic libraries is that certain sequences, e.g., those involving highly polymorphic regions and inverted repetions, often cannot be propagated in most lambda cloning vehicles.[182]

In addition to genomic libraries, **cDNA libraries** can be prepared from mixed mRNAs. The total RNA of cells is isolated and passed through an affinity column containing oligo(dT) chains. These bind to the 3'-poly(A) tails of the mRNAs, allowing them to be isolated. The mixed mRNAs can then be cloned using a poly(AT)-tailed plasmid vehicle and a reverse transcriptase.[183,184]

6. Probes

The first step in screening the recombinant DNA in a library is use of some probe for detecting the desired DNA fragment. The most direct way is to synthesize a radioactive or fluorescent labeled oligonucleotide[185] complementary to a short known sequence in the protein. The number of codons for a single amino acid varies from 1 to 6 (Table 5-6). It is therefore desirable to prepare a probe complementary to segments of DNA containing a high proportion of codons for Trp and Met (1 codon each) and Asn, Asp, Cys, Glu, Gln, His, Lys, Phe, and Tyr (2 codons each). A popular procedure is to synthesize a mixture of probes containing all of the possible nucleotide sequences coding for the selected sequence of amino acids. The probe may be a mixture of more than 1024 different nucleotide sequences.[186] See also Chapter 5, Section H,4.

Antibodies are another popular type of probe. Antibodies to a specific protein may be utilized in isolating mRNA from ribosomes that are making that protein (Chapter 29). Thus one or more strongly binding antibodies may have already been obtained before the library clones are to be screened. To use this technique for screening recombinant DNA, the cloning must be done in a vehicle that causes expression of the gene, e.g., as proteins fused to *E. coli* galactosidase. One type of expression library is created by insertion of the cDNAs into copies of a bacteriophage gene that permits the expressed proteins to be displayed on surfaces of the phages[186a] (see Fig. 3-16).

7. Studies of Restriction Fragments

By cleavage with the correct restriction enzyme, cloned DNA fragments can be released from the vehicle in which they are carried in the library. What can be done with these fragments? The first obvious use is to sequence a gene that has been located with a probe. In many instances the gene will be longer than the cloned piece. However, the isolated restriction fragment, if labeled and denatured, becomes a highly specific probe for locating other restriction fragments that overlap it. For example, a fragment from an *Eco*RI library may bind to two or more fragments from a *Hin*dIII library. This permits "walking" along the chromosome to locate adjacent fragments. Cosmid vehicles that facilitate ^{32}P-labeling at the ends of the passenger DNA are useful.[187] A related approach called "jumping" depends upon converting very large DNA restriction fragments into circular molecules, digesting with restriction enzymes, and cloning the junction fragments of the circles. These fragments contain segments that may have been separated by as much as 100 kb in the genomic DNA, and enable the investigator to walk or jump again from a new location.[188]

Locating mutations. The study of restriction fragments provides a way of locating many mutations. For point mutations in genes for known proteins, sequencing reveals the exact defect. Some mutations, especially deletions, lead to changes in the lengths of restriction fragments. If the mutation causes loss of a restriction site, a longer piece of DNA will be present than in a digest of normal DNA. Such differences in length of restriction fragments are usually referred to as **restriction fragment length polymorphism** (RFLP, usually used in the plural as **RFLPs**). These polymorphisms are readily detectable by differences in mobility on gel electrophoresis.[96,189] They can often be mapped to particular chromosomes by hybridization *in situ* (Fig. 26-14),[190,191] by study of naturally occurring translocations of chromosome fragments, or other techniques.[192] An example is provided by the human hemoglobin abnormalities known as **thalassemias** (Chapter 32). Here, deletions remove certain restriction sites leading to observation of RFLP. Occasionally a point mutation is linked to RFLP at a nearby site. For example, in the United States most carriers of the sickle cell trait have a 13-kb *Hpa*I fragment that carries the globin S gene. However, noncarriers have their globin gene on a 7.6-kb *Hpa*I fragment.[99] Although the association is fortuitous, the linkage between the hemoglobin S gene and the mutated restriction site is broken only rarely by crossing-over during meiosis. RFLPs have been linked to many other human genetic defects and have also provided the basis for the first linkage maps of human chromosomes.[96,193]

Positional cloning. By using enzymes that cut at relatively rare sites, genomic DNA can be cut into very large restriction fragments. From studies of inheritance within families carrying specific genetic traits it is sometimes possible to find linkages between those traits and polymorphic restriction fragments.[194–196] This has been accomplished for a number of defective human genes including those responsible for sickle cell anemia,[197] cystic fibrosis (Box 26-A), Duchenne muscular dystrophy (Box 19-A), Huntington's disease,[198,199] X-linked chronic granulomatous disease[200] (p. 1072), neurofibromatosis (elephant man's disease),[201] the hereditary cancer retinoblastoma,[202] and others. By 1997 nearly 100 hereditary disease loci had been located by positional cloning.[203] These astonishing successes provided a major impetus for what became the Human Genome Project (Section G).[193]

Serious problems were met in actually locating these disease genes. Crossing-over is infrequent, occurring only at about 50 locations during each meiosis.[192] Therefore, linkage analysis does not tell us with any precision how close the linked gene is to a known DNA probe within a restriction fragment that may be up to 2000 kb in length. Finer restriction mapping or chromosome walking can be used to locate the precise piece of DNA that is defective.[204] This can still be a formidable problem. However, if the defective protein can be identified it can be sequenced. A specific oligonucleotide probe can be made for its gene and can be used to establish the exact chromosome location.

8. Directed Mutation

In addition to the developments of cloning and sequencing of DNA, a third technique is essential to the present revolution in molecular genetics. That is the ability to mutate any gene at any point in a specific way. Because of its precise nature the technique is called **directed mutation** in this book. However, the term **site-directed mutagenesis** is often used. Mutations can be introduced randomly in DNA in many ways including treatment with nitrous acid, bisulfite, formic acid, or hydrazine or by incorporation of nucleotide analogs.[205] Efficient procedures have been devised for isolating the mutants.[206] For many purposes **oligonucleotide-directed mutation** is the preferred technique.[207–210] An oligonucleotide of ~16–20 nucleotide length is synthesized with a sequence complementary to the coding strand containing the desired site of mutation. At that site the codon for the new amino acid is present. Despite this mismatch the oligonucleotide can be successfully hybridized with a single-stranded DNA such as that cloned in an M13 vehicle. Now the Klenow fragment of DNA polymerase I or a viral DNA polymerase (Chapter 27) is used to convert the single-stranded circular DNA into a double-stranded replicating form. Many of the single-stranded progeny will contain the mutated DNA. They can be screened with a labeled probe made from the oligonucleotide used to induce the mutation. It will hybridize most tightly to the correctly mutated gene.[207] Use of the PCR reaction simplifies the procedure.[210–212] In another screening procedure the template strand is synthesized in an M13 phage vehicle using uracil rather than thymine. The circular heteroduplex obtained after synthesis with T4 DNA polymerase and T4 DNA ligase is taken up by *E. coli* cells, which select against the uracil-containing strand and, therefore, in favor of the mutated strand.[208]

A third approach is to completely synthesize a gene for the protein under study. The sequence does not have to be exactly the natural one but can be made with restriction enzyme cleavage sites that permit easy excision and readdition of particular fragments. Synthetic fragments containing various mutations can then be grafted in at will. Genes of this type have been made for rhodopsin and related proteins (Chapter 23) and for numerous other proteins.

What are the uses of directed mutation? As we have seen in previous chapters, the technique is being used in every area of biochemistry to bring new understanding of protein functions and of the chemical basis of disease. Together with complete synthesis of genes it provides the basis for genetic engineering of specific proteins of plants, animals, and microorganisms. Many protein products can probably be improved. For example, enzymes can be made more stable.[213] Specificities can be changed, but this is difficult.

Targeting and replacing genes. One goal of human genetic therapy is to replace a defective gene in body cells with a good gene. Is this really possible? It is essential to engineer the DNA that is to be transferred, so that it contains all of the components needed for efficient expression in the host following its incorporation as a **transgene**. A transcription-initiation region with suitable promoter, both 5' and 3' untranslated regions, start and stop codons, and polyadenylation site (Chapter 28) must all be present. It is hoped that a correctly constructed promoter region will allow the transgene to be picked up by the machinery of homologous recombination (discussed in Chapter 27) and be incorporated into the host's DNA and expressed in the appropriate tissues.[214,215] Only some cells will take up the new gene and discard the old. However, it may happen enough to benefit a patient. Targeted gene replacement has been very successful in mice.[216]

Knockout mice. If targeted DNA is injected into a fertilized mouse egg, there is a chance that the mouse will have the targeted gene replaced in one chromosome, and that it will be stably transmitted to some of its progeny. If the transgene is totally nonfunctional, the mouse will be a "knockout mouse," suffering from a hereditary defect that can be transmitted through carriers such as its mother. The standard knockout technique is to inactivate the gene of choice in cultured embryonic stem cells and to inject these into mouse embryos. Some progeny will carry the inactivated gene in their germ cells. A refinement of the technique utilizes the **Cre recombinase** or related enzymes discussed in Chapter 27 to selectively remove pieces of DNA from genes in specialized tissues of mice.[217,218] By 1996 several hundred different knockout mice had been created.[215] Nevertheless, interpretations of results of gene knockout are sometimes complex.[219]

F. The Genetics of Eukaryotic Organisms

Whereas DNA synthesis takes place almost continuously in a rapidly growing bacterium, replication of DNA occupies a more limited part of the **cell cycle** of eukaryotes (Fig. 11-15). In a mammalian cell mitosis proper (Fig. 26-11) may require about one hour. It is followed by the "gap" period, G_1, whose length is variable and depends upon the cell type, the nutritional state of the cell, and other factors. About 10 h is typical. During the S phase (~9 h) active DNA replication takes place. This is followed by a second gap (G_2) that occupies 4 h in the 24 h cell cycle shown in Fig. 11-15. The length of the different segments of the cell cycle varies widely among different organisms. Indeed, the concept of a cell cycle can be criticized.[219a] It is only in a rapidly growing culture that all, or most, cells can follow the same cycle. In the adult body most cells are inhibited from division (or are not stimulated to divide) most of the time.

1. Mitosis

The distribution of chromosomes to daughter cells of somatic cells undergoing division is accomplished by mitosis whose successive phases are referred to as **prophase**, **metaphase**, **anaphase**, and **telophase** (Fig. 26-11). As the chromosomes condense during prophase, it is seen that each one actually consists of two separate entities coiled together. These are the identical **chromatids,** which are formed from the two identical double-stranded DNA molecules formed by replication of the DNA of the chromosome during the S phase of the cell cycle. As the folding of the chromosomes occurs (during prophase), the nuclear envelope completely fragments or dissolves in many species.

An important event that *precedes* the main stages of mitosis is the formation of **poles** in the cell. In animal cells, the poles are formed by the **centrioles,** which move apart and take up positions at opposite sides of the cell. Each of the centrioles is accompanied by a smaller "daughter" centriole lying at right angles to the larger parent. In plant cells, which lack centrioles, a more diffuse pole is formed. As the cell prepares for mitosis, fine microtubules (15 nm diameter) can be seen radiating from the poles. At the end of prophase the microtubules run from one pole to the other to form the **spindle**. Microtubules also become attached to the chromosomes at the **centromeres**.

At metaphase the chromosomes are precisely lined up in the center of the cell to form the **metaphase plate**. Now each centromere divides, permitting the sister chromatids to be completely separated. A protein complex **cohesin**, which holds the sister chromatids together, undergoes proteolysis by a **separase** at this stage.[220,220a,b] During anaphase the separated chromatids, now referred to as **daughter chromosomes**, move to opposite poles as if pulled by contraction of the spindle fibers. Telophase is the final stage in which new nuclear envelopes are formed around each set of daughter chromosomes. In humans and many other species the cell pinches in two. In plants and fungi new plasma membranes and cell walls are constructed through the center of the cell. A partitioning of mitochondria and of Golgi components between cells must also occur.[220c]

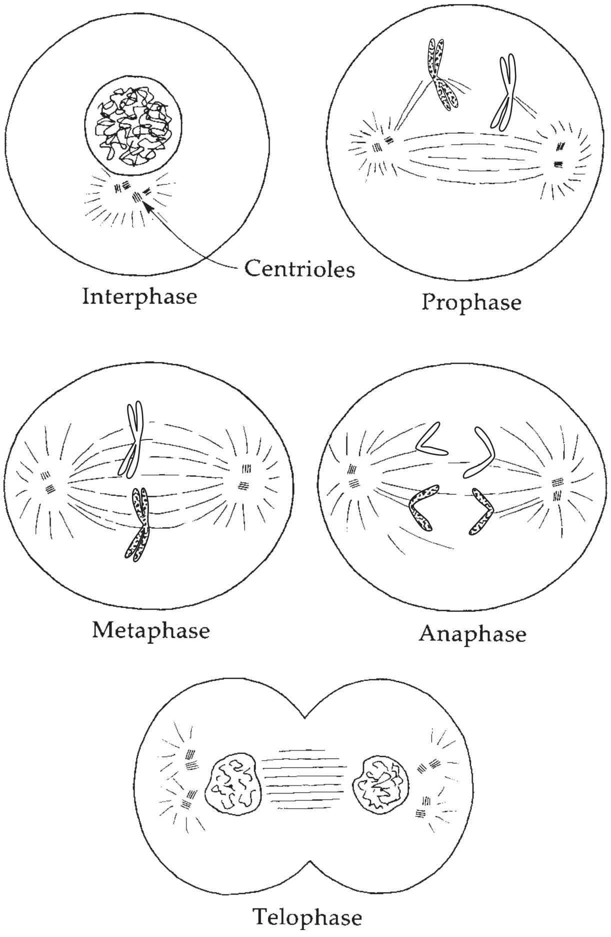


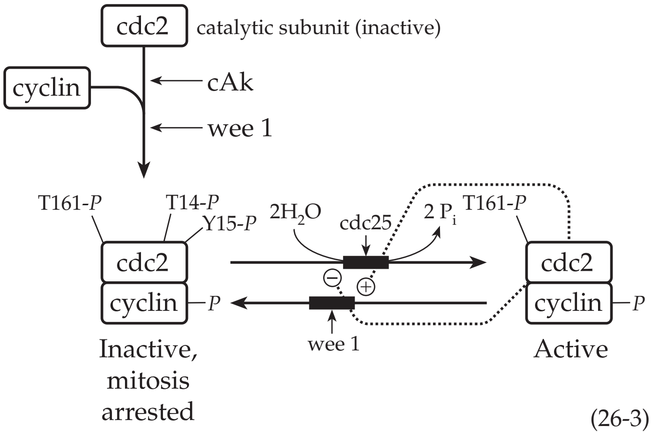
Figure 26-11 Mitosis. Illustrated for a cell with one homologous pair of chromosomes. After Mazia.[221]

The foregoing description overlooks the extreme complexity of mitosis, each stage of which must occur with precision and in the correct sequence.[222–225a] The replication of DNA, which takes place in the S phase of the cell cycle (and is discussed in Chapter 27) must be completed before mitosis begins. This is followed by condensation of the DNA into chromosomes (Chapter 27), breakdown of the nuclear membrane,[226,226a] assembly of the **kinetochores** by which the chromosomes attach to the spindle,[222] assembly of the spindle, attachment of chromosomes to the spindle, segregation of the chromosomes to opposite poles in anaphase, and finally the cleavage of the cell.

Cyclin-dependent kinases. As is shown in Fig. 11-15, the cell cycle is controlled by a series of complexes of the 30- to 45-kDa proteins called **cyclins** with **cyclin-dependent protein kinases** (**CDKs**). These kinases contain ~300-residue catalytic cores that resemble protein kinase A (Fig. 12-32). Like that kinase they transfer phospho groups from ATP to serine and threonine side chains of target proteins.[227–228] The kinases are inactive until a complex with the appropriate cyclin is formed and is activated by phosphorylation.[229] One or more additional kinases are required for this activation. Each stage in the cell cycle is controlled by one or more different cyclin–CDK complexes (Fig. 11-15). One of the best known CDKs is human **CDK2**, which functions in a complex with cyclin A during the S phase of the cycle.[228,230–232] Binding of the cyclin greatly alters the conformation of CDK2, opening the catalytic cleft and exposing threonine 160. Its hydroxyl group can be phosphorylated by the action of the **CDK-activating kinase** (**CAK**) with a 100-fold increase in catalytic activity.[228] Control of the CDKs also depends upon proteins that act as specific inhibitors and upon precise elimination of both cyclins and inhibitors via the ubiquitin system.[228a,b,c]

Checking for completion of replication and for DNA damage. The first checkpoint in the cell cycle is the **start** or **G_1 DNA damage checkpoint** (Fig. 11-15). Replication of DNA does not begin until the cell has had time to repair as much damage to DNA as is possible.[228d] As mentioned in Box 11-D, the cancer suppressor protein p53 is an essential component of the checking process.[224] A second checkpoint, the **G_2 checkpoint**, at the end of the G_2-phase (Fig. 11-15) requires verification that all DNA of all chromosomes has been replicated, checked for damage, and repaired if necessary. Control of these processes is accomplished by a mechanism first identified in

fission yeasts, such as *Schizosaccharomyces pombe*, and which has been conserved in metazoa.[223,233–236d] The CDK known as **cdc2** (fission yeasts), CDC28 in budding yeast, or CDC2 or CDK1 in mammals, functions from G_2 through the DNA damage checkpoint to mitosis (Fig. 11-15) and is sometimes described as the master cell-cycle switch.[236a] The catalytic subunit of cdc2, a serine / threonine protein kinase, is often called $p34^{cdc2}$. It is inactive unless complexed with a suitably phosphorylated cyclin (Eq. 26-3) and phosphorylated on Thr 161 by the action of kinase CAK (Eq. 26-3). However, when cdc2 becomes phosphorylated on Tyr 15 (in *Schizosaccharomyces pombe*) by the action of protein kinase wee 1 (Eq. 26-3) and on Thr 14 by another kinase, cdc2 is inhibited and mitosis is arrested at the G_2 checkpoint. This allows time to verify completion of replication as well as for repair before the replicated DNA strands are separated in mitosis. Hydrolytic removal of the phospho groups from Tyr 15 and Thr 14 of cdc2 by the action of phosphatase cdc25 then allows mitosis to begin.[223,237,237a] This phosphatase is also controlled by a phosphorylation-dephosphorylation cycle. The checkpoint kinase **Chk1** phosphorylates and inhibits cdc25.[238,238a]



(26-3)

When damaged DNA is present, protein p53 accumulates, just as at the G_1 checkpoint, and activates transcription of the seven phosphoserine-binding proteins of the 14-3-3 family.[224] These bind to the phosphorylated cdc25 phosphatase preventing activation of cdc2 by dephosphorylation.[224,238–240] In fission yeast one of the 14-3-3 proteins, **Rad 24**, apparently binds to the phosphorylated cdc25 and induces its export from the nucleus preventing mitosis.[241] Dephosphorylation of cdc25 and return to the nucleus allows mitosis to occur. A somewhat different regulatory mechanism is used by the budding yeast *Saccharomyces cerevisiae*.[242] In animal cells there are three cdc25 isoforms.[240,240a,b] In each case we probably have only a glimpse of a very complex process of checking for DNA damage and repair (see also Chapter 27).

Mitotic spindle formation and the spindle assembly checkpoint. Separation of the two copies of each replicated chromosome depends upon the spindle fibers. Their formation is preceded by replication of centrioles, if present, and formation of the two poles of the cell.[243–244b] Both γ-tubulin (Fig. 7-34) and acidic Ca^{2+}-binding proteins called **centrins** are involved.[244c–e] The microtubules appear to grow outward from the poles with their minus ends at the poles and their plus ends (Fig. 7-33) available for binding to the kinetochores, specialized protein complexes that assemble around the centromeric DNA (Chapter 27).[245–247] Each chromosome has two kinetochores, one for each daughter chromatid. These must be attached to spindle fibers coming from opposite poles. It has usually been assumed that random encounters of microtubule plus ends with kinetochores leads to correct linkage.[245,246] However, neither centrosomes or kinetochores are always essential to spindle assembly, and self-assembly occurs by motor-driven sorting according to polarity of the microtubules.[246,247] In addition to microtubules formation of spindles may require specialized matrix proteins.[247a]

The spindle isn't formed until replication is complete. A small G protein called **Ran**, a relative of Ras (Fig. 11-7), regulates spindle formation. Ran, in turn, depends upon a nucleotide exchange factor called **RCC1** (**regulator of chromatin condensation**).[248–250] RCC1 may signal that the replicated DNA is folding into chromosomes indicating that replication is complete. At the **spindle assembly checkpoint** (Fig. 11-15) the cell verifies that the metaphase spindle has been assembled correctly.[250a] All of the microtubules that pull the sister chromatids toward one pole or the other must be correctly attached to a kinetochore.[220,250b] In addition, there are interdigitated microtubules coming from both poles. Specific motor molecules (Chapter 19) then push the two poles apart. During the assembly and complex movements of the spindle both cytosolic dynein and four different types of kinesin-like motor molecules are required.[251,252] One of these is Kar3 (see Fig. 19-17). After assembly of the spindle is complete, a signal must be sent to the mitotic apparatus to move into anaphase. A clue to the nature of the signal has come from the observation that a single kinetochore lacking a spindle fiber connection causes arrest of mitosis.[253] Apparently unattached kinetochores send a "wait" signal, perhaps via the cytoskeleton.[253–255]

Anaphase. After the spindle has been checked, a sudden loss of cohesion between the sister chromatid pairs allows them to move toward the opposite poles. This process is catalyzed by the **anaphase-promoting complex** (**APC**, or cyclosome) and its activator protein **Cdc20**, a large multiprotein complex.[225] The APC also promotes proteolytic breakdown of cyclins and other proteins by a ubiquitin- and proteosome-dependent mechanism.[225,256–259c] Its E_3-ubiquitin ligase (Box 10-C) targets the mitotic cyclins and other proteins for destruction.[260] A specific E_2 ubiquitin-conjugating enzyme (Box 10-C) is required for degradation of cyclin B and exit from mitosis.[261] The centrosome also plays an active role in cytokinesis, the final step in cell division.[261a]

The cell cycle encompasses so many different processes that it is clearly impossible to describe it by the single diagram of Fig. 11-15 or by the text written here. The cycle is influenced by a host of growth factors and external stimuli, many of which act on transcription of cyclins and other essential proteins. Transcription factors such as those of the Fos / Jun (AP-1) family in response to the MAP cascade (Fig. 11-13) are among those that control the transcription of cyclins.[262–263c] However, during mitosis most transcription of any genes is repressed.[264]

Among other factors influencing the cell cycle is the size of the cell and the availability of nutrients including purine and pyrimidine nucleotides.[263b,c,264a] Lack of cholesterol decreases the cdc2 kinase activity and causes apoptosis.[265] A cell cycle regulator in *S. pombe* known as **suc1** is essential for cell cycle progression. Although its three-dimensional structure is known, its function (like that of its human homolog CksHs2) is uncertain.[266,267]

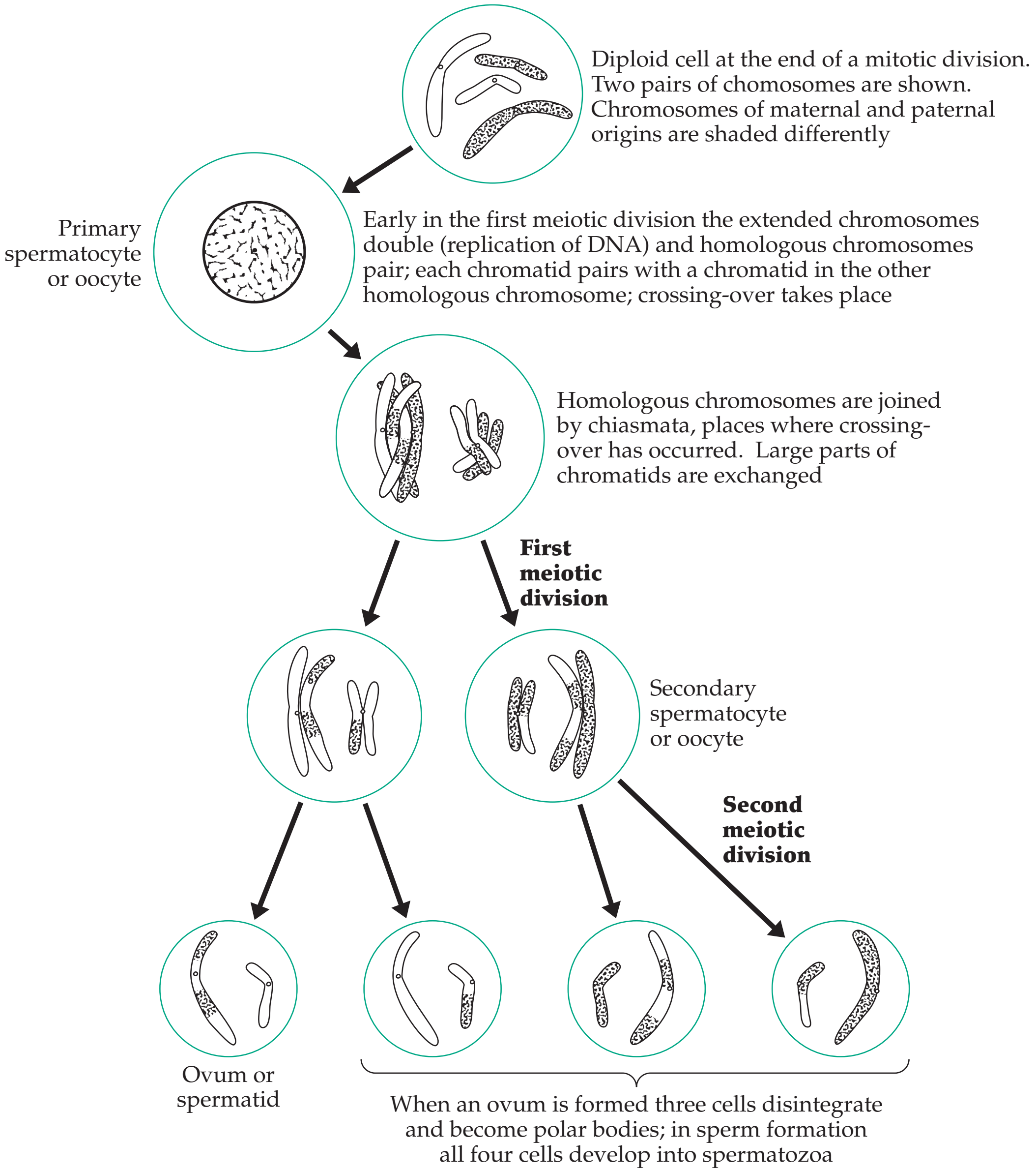
2. Meiosis

The mechanism by which chromosomes are distributed during the formation of **gametes** (egg and sperm cells) is known as meiosis (Chapter 1; Fig. 26-12). Formation of gametes involves a halving of the chromosome content of a cell, each gamete receiving only one chromosome of each homologous pair. Genes found in the same chromosome are said to be **linked** because of their tendency to be passed together to the offspring. Genes present in different chromosomes are not linked, and their inheritance follows the pattern of **random segregation** established in Mendel's famous studies.

The simple fact that the genetic material is put up in several different packages (chromosomes) is sufficient to provide for considerable mixing of genetic information between different individuals in sexual reproduction. However, it doesn't provide a means for exchanging genes on the same chromosomes. Mixing of genetic information within chromosomes occurs by genetic recombination occurring during **crossing-over**, an aspect of meiosis with an essential biological role. In the S phase preceding meiosis, DNA is duplicated just as it is prior to mitosis. This provides sufficient genetic material to produce *four haploid cells*. These are formed during meiosis by two consecutive cell divisions (Fig. 26-12). Crosssing-over occurs prior to the first of these divisions, at the four-strand stage. The two homologous chromosomes of a pair come together to form what is called a **bivalent** or **tetrad** made up of four chromatids. For each chromosome, at least one chromatid is seen to come into intimate contact with a chromatid in the other homologous chromosome at points known as **chiasmata** (Fig. 26-13). During metaphase of the first meiotic cell division the homologous chromosomes, each still containing two chromatids, separate. Each chromatid now carries with it some genetic information that was previously found in the other member of the homologous pair and vice versa (Fig. 26-12). Now, without further replication of DNA in the second meiotic cell division, the chromatids separate to form haploid cells.

The process of crossing-over provides a means by which genes that are linked on the same chromosome can be separated, providing offspring with mixtures of genetic traits other than those predicted by simple Mendelian theory. The effects of crossing-over were first studied extensively by T. H. Morgan with the fruit fly *Drosophila melanogaster*. Morgan discovered his first mutant, a white-eyed fly,[267a] in 1910. The first genetic maps were made by assuming a direct relationship between the frequency of crossing-over and the linear distance between genes in a chromosome. Thus, the same approach to genetic mapping that was used later with *E. coli*, i.e., the measurement of recombination frequencies, was applied much earlier to crossing-over in the chromosomes of *Drosophila*. Extensive genetic maps involving many mutations were obtained for the four chromosomes of this organism, and similar techniques have been applied to many other organisms. The unit of distance in these chromosome maps is the **morgan** (named for T. H. Morgan[2,267a]). One centimorgan is the distance that allows recovery of 1% of recombinant progeny.[2] In the human chromosomes, this is ~1000 kb (1 Mb).

Biochemical and genetic studies of meiosis have been conducted in many organisms including fission[268,269] and budding[270–271a] yeasts, *Drosophila*,[272,273] starfish,[274] *Xenopus*,[275] and the mouse.[276] Meiosis can be viewed as a modification of mitosis but with the added initial step of crossing over and recombination. In addition, the S-phase of the cell cycle is absent in the second meiotic division. As was mentioned in Chapter 1, meiosis may occur at different stages of the life cycle of organisms. An important advantage in using fungi for genetic studies is that, like prokaryotes, they are haploid during much of their life cycle. Biochemical defects such as the inability to synthesize a particular nutrient can be recognized readily at this stage. At the same time genetic crosses can be made, and crossing-over frequencies can be measured and used for genetic mapping. The onset of meiosis may



Diploid cell at the end of a mitotic division. Two pairs of chomosomes are shown. Chromosomes of maternal and paternal origins are shaded differently

Early in the first meiotic division the extended chromosomes double (replication of DNA) and homologous chromosomes pair; each chromatid pairs with a chromatid in the other homologous chromosome; crossing-over takes place

Homologous chromosomes are joined by chiasmata, places where crossing-over has occurred. Large parts of chromatids are exchanged

When an ovum is formed three cells disintegrate and become polar bodies; in sperm formation all four cells develop into spermatozoa

Figure 26-12 Meiosis. Cell division leading to formation of haploid gametes.

vary in time with diploid organisms. Meiosis in males may occur quickly. However, after chromatid pairing and recombination meiotic divisions of oocytes are arrested at the G$_2$ stage, in some species for many years.[274,275] The arrest is ended by hormonal stimulation, e.g., via progesterone acting through a cyclin B–cdc2 complex.[276] **Meiosis-activating sterols**, intermediates in lanosterol metabolism (Fig. 22-8) also accumulate and are thought to participate in control of meiosis.[276a] In haploid strains of fission yeasts sexual development is induced by starvation, especially of nitrogen. Cells of opposite mating types then fuse to form zygotes, which usually undergo meiosis immediately. Starvation is apparently signaled by the **cyclic AMP–protein kinase A cascade** (Fig. 11-13).[268]

During the prophase of the first meiotic division (meiosis I) two homologous pairs of partially "condensed" chromosomes must find each other and pair with appropriate orientation. A protein in the telomeres of the chromosomes seems to be involved.[269,277] The key structure in meiotic crossing-over is the ribbonlike **synaptonemal complex** formed by the pairs of homologous chromatids.[271,278–279b] This complex, in which a proteinaceous core or **axial element** separates the greatly extended chromatid pairs (Fig. 26-13), is fully formed in the **pachytene stage** of meiosis. Formation of the synaptonemal complex is preceded by development of a few double-stranded breaks in

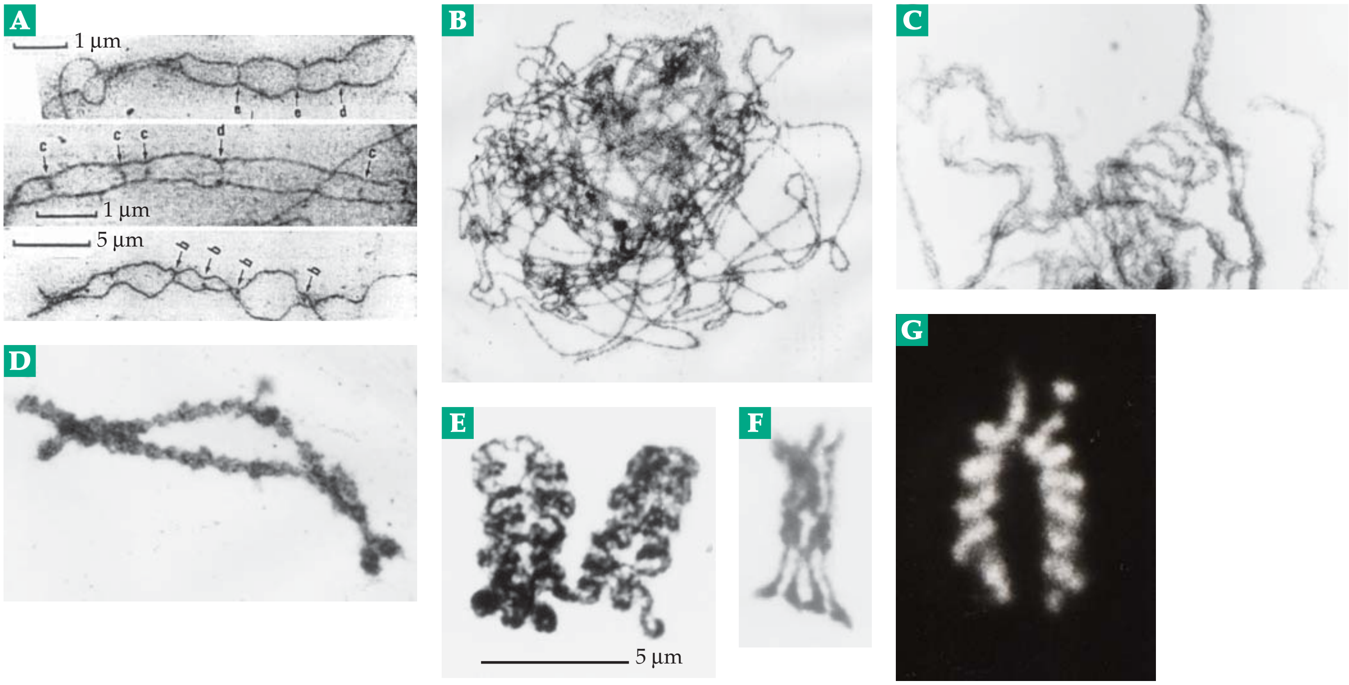


Figure 26-13 Synaptonemal complexes. (A) Aligned pairs of homologous chromatids lying ~0.4 µm apart in *Allium cepa*. Arrows indicate "recombination nodules" which may be involved in initiating formation of crossovers. Portions of meiotic chromosomes of lily are shown at successive stages: (B) Pachytene. (C) Portion of diplotene nucleus. (D) A bivalent at diplotene. (E) Two bivalents at diakinesis. Pairs of sister chromatids are coiled with appropriate handedness. (F) Sister chromatid cores are far apart in preparation for separation. A chiasma is present between the two central strands. (B) through (F) courtesy of Stephen Stack.[279,279d] (G) Pair of sister chromatids coiled with opposite handedness at metaphase. These are immunostained with anti-topoisomerase II antibodies. From Boy de la Tour and Laemmli.[280] Courtesy of U. K. Laemmli.

the DNA.[278,279c] The 3' ends of the DNA chains then invade the homologous chromatids to initiate the exchange via **Holliday junctions** (Fig. 5-28 and Fig. 27-11). The details of the recombination process are considered further in Chapter 27. The first meiotic division is a long process, for example, lasting ten days in mouse spermatocytes.[281] As cells pass to the metaphase stage of meiosis I, the chromosomes become much more compact, but the attachments between the homologous chromatids are still visible as chiasmata (Fig. 26-13). The chromatids then separate, the two homologs appearing as coils of opposite handedness. Within these coils the pairs of sister chromatids continue to be held together at their centromeres through metaphase and anaphase of meiosis I. One or more specific proteins, which must release their hold in meiosis II, are required.[270,277] A leading cause of embryonic and fetal death as well as severe mental and physical problems after birth is incorrect segregation of chromosomes during meiosis. If a gamete contains two copies of any chromosome the embryo will have three. Down syndrome, trisomy of chromosome 21, is the most frequent example. The spindle assembly checkpoint during meiosis I may sometimes be at fault.[281a,b]

3. Polytene Chromosomes

While most cells of higher organisms are normally diploid, the chromosome number may sometimes be doubled or increased even more. A cell with twice the diploid number of chromosomes is **tetraploid,** and with higher multiples of the haploid number it is **polyploid**. Plant breeders have succeeded in producing many tetraploid varieties of flowering plants often with increased size. One tetraploid mammal, the red viscacha rat, is known.[281] While most of our body cells are diploid, we, too, have polyploid cells. For example, some are always found in the liver. The most spectacular example of an increase in the normal DNA content of cells is provided by the giant **polytene chromosomes** of dipteran (fly) larvae. The DNA of cells in the salivary glands and some other parts of these organisms doubles as many as 11–14 times without cell division to give a several thousandfold (i.e.,

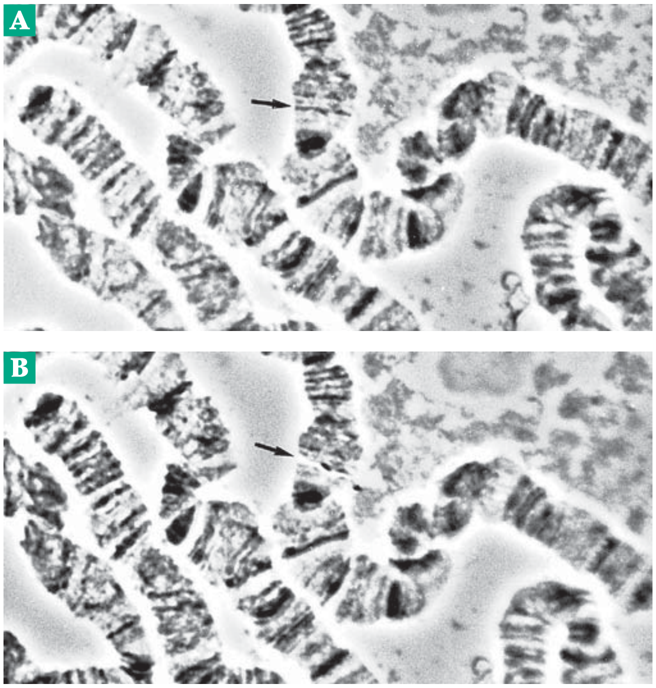


Figure 26-14 Microdissection of a *Drosophila* salivary gland chromosome. (A) Before the cut. (B) After cutting a medium size band. The arrow indicates the site of the cut. From Pirrotta.[282]

2^{14}-fold) increase. The supercoiled, duplicated DNA molecules all line up side by side in a much more extended form than in ordinary chromosomes. They can be seen readily by light microscopy. The total length of the four giant chromosomes of *Drosophila* is ~2 µm, compared to 7.5 cm in diploid cells. The giant chromosomes have a banded structure, ~5000 bands being visible along the length of the chromosome (Fig. 26-14). An average band contains ~36 kb of DNA in each of the strands. Since it has been possible to correlate visible changes in the appearance of these bands with particular mutations in the DNA, study of polytene chromosomes provided a second important method of mapping genes of the fruit fly. The maps produced by the two methods agree well.

Another use of polytene chromosomes is microdissection of DNA for cloning (Fig. 26-14). A piece of DNA containing 100–400 kb can be cut out of any desired spot, cleaved with restriction enzymes, and cloned.[282] Since it has been estimated that *Drosophila* may contain only ~9000 –17,000 genes there may be 2–3 genes per band in these chromosomes.[283] The technique has been extended to human and other mammalian chromosomes.[284]

4. Cytoplasmic Inheritance

Not all hereditary traits follow the Mendelian patterns expected for chromosomal genes. Some are inherited directly from the maternal cell because their genes are carried in the cytoplasm rather than the nucleus. There are three known locations for cytoplasmic genes: the mitochondria, the chloroplasts, and certain other membrane-associated sites.[285,286] An example of the last is found in "killer" strains of yeast. Cells with the killer trait release a toxin that kills sensitive cells but are themselves immune. The genes are carried in **double-stranded RNA** rather than DNA, but are otherwise somewhat analogous to the colicin factors of enteric bacteria (Box 8-D). Similar particles (κ factors) are found in *Paramecium*.[287]

Mitochondrial and chloroplast genes are discussed in Chapters 18 and 23, respectively.

G. The Human Genome Project

The discovery of site-specific restriction endonucleases in 1970 and the development of efficient DNA sequencing methods in 1977 sparked a revolution in biology. On Oct 1, 1990, the 15-year project to sequence the complete human genome was officially begun.[288–290] The project is far ahead of schedule. "First drafts" of the genome were published in 2001.[290a–e] A more complete version will include **annotation**, a listing of predicted exons, mRNA transcripts,[290f] functions of recognizable genes, etc.[290e] Such documentation was provided first for chromosome 22[290e] and is now available for other human genes.[290g] It was a surprise to discover, at least initially, that only 30,000–40,000 protein-coding genes could be recognized.[290a,290h] Earlier 50,000–150,000 genes had been predicted (Table 3-1). A nearly complete version of the human genome is anticipated for spring of 2003.[290i]

Included in the genome project are completion of the sequences for *Caenorhabditis elegans*,[290j] *Drosophila melanogaster*, and the mouse (*Mus musculus*).[291,291a–e] However, substantial difficulties remain in filling hard-to-sequence gaps, and some new approaches may need to be developed.[292] More recently the project has been extended to include many additional species. Among them are the rat,[292a] the zebrafish,[292b] the pufferfish *Fugu rubripes*,[292c] the human malaria parasite *Plasmodium falciparum*,[292d,e] the rodent malaria *Plasmodium yoelii, yoelli*,[292f] and the mosquito *Anopheles gambia*, which carries the parasite.[292g,h] Plant genome sequences are also being determined (p. 1511).

1. The Mammalian Genome and Human Health

Human beings and other mammals all have about

3500 Mbp of DNA containing perhaps 40,000 protein-coding genes. However, ~ 97% of the genome is repetitive DNA and other DNA that, at present, has no recognizable function. The genome project will reveal all of these sequences and will doubtless provide us with many surprises.

A primary goal of the genome project is to understand the relationships between gene sequences and human diseases and health. Until recently little was known about the locations or structures of genes in human chromosomes. Human genetic diseases provided the first clues. Although systematic genetic experiments cannot be done with human beings, almost the entire population is under some kind of medical care. Genetic defects are being detected more and more often, and the inheritance within families is traced with increasing frequency. Consequently, a huge body of knowledge of mutations to the human genome is available.[293] By 1995 ~4500 human genetic disorders had been discovered.[294] Many more have been found since then.

The first mapping of human genetic defects came with the recognition that sex-linked traits are encoded on the X-chromosome. Some linkage analysis was also possible from studies of inheritance within families. For example, among individuals who have two X-linked traits, e.g., color blindness plus one other, naturally occurring crossing-over occasionally breaks the linkage in some individuals within the family. **Somatic cell fusion**[191] provided an additional approach. Human lymphocytes can be fused with rodent cells under the influence of inactivated Sendai virus, which causes the cells to adhere and then to fuse. From such human–mouse or human–hamster hybrid cells, strains in which the nuclei have also fused can be selected. Although such cell lines can be propagated for many generations, they tend to lose chromosomes, especially those of human origin. By observing loss of particular human enzymes or other proteins (separable from the hamster enzymes by electrophoresis), it was possible to assign genes to specific chromosomes. This also required identification of the chromosomes lost at each stage in the experiment. New staining techniques made it possible to identify each of the 22 pairs of human autosomes as well as the X and Y chromosomes (Fig. 26-15).[295] Using a variety of techniques, 500 loci in human chromosomes had been mapped by 1983.[296]

Separation of individual human chromosomes with a fluorescence-activated cell sorter (Box 3-B)[299] permitted the preparation of libraries of cloned DNA from individual chromosomes. Fragments, of an average length of 4 kbp, from digestion by *Eco*R I or *Hin*dIII were packaged into the Charon 21A cloning vehicles and “amplified” by culturing infectious phage particles in *E. coli*.[282] However, this approach was inadequate for sequencing the entire genome. By 1987 the use of yeast artificial chromosomes (YACs)[300] and later bacterial artificial chromosomes (BACs)[301,302] and radiation hybrid mapping[104,303] provided the means for sequencing the 34- to 260-Mbp human chromosomes.[288,304] Also essential was the development of high-speed automatic sequencing machines[305–305b] and of computers adequate to assist in compiling the sequences.[302,306,306a] Sets of overlapping and redundant clones define a continuous sequenced segment that is called a **contig**. As the contigs are enlarged the gaps are filled.

In 1987, the first global genetic linkage map of the human genome was published.[288,307] It was based on 403 polymorphic loci, of which 393 were RFLPs studied in 21 three-generation families. By 1996 a linkage map containing 5264 loci, in the form of polymorphism in **short tandem** repeats $(AC/TG)_n$ represented by **microsatellite DNA**, was available.[304] Also mapped by 1998 were 2227 **single nucleotide polymorphisms** (SNPs), where the two alternative bases occur with a frequency >1%.[308] This was increased to 1.4 million by 2001.[308a] As many as 5.3 million common SNPs may occur, each with a frequency of 10–50%. On the average there may be one SNP for every 600 base pairs.[308b] These SNPs account for a large fraction of the diversity in human DNA. By 1995 a physical map with >15,000 **sequence tagged sites** (STSs) with an average spacing of 199 kb had been constructed.[303,309] An STS is defined as a cloned fragment in a YAC (or BAC) library that has been amplified by PCR and tested to establish that it contains a known locus. DNA sequences known as **expressed sequence tags** have also been mapped (see p. 1490). Using STSs, RFLPs, ESTs, and the growing contigs, more than 16,000 human genes had been mapped by 1996.[310] By 1998 more than 1,060,000 ESTs had been reported.[311] In 1999, the nucleotide sequence of the smallest human chromosome, the 33.4-Mbp chromosome 22, was completed. It contains at least 545 genes and 134 pseudogenes.[312] In the entire genome >1000 seed contigs had been assembled, and completion of the first phase of the Human Genome Project was in sight.[313] In 2001 a whole-genome clone-based physical map was published.[313a]

2. Understanding Gene Sequences

The vast array of sequence data coming from the human genome and from genes of other species are deposited, as they are reported in the scientific literature, in the Human Genome Central. It can be found on the Web at the National Center for Biotechnology Information (NCBI) at http://www.ncbi.nlm.nih.gov/genome/guide and European Bioinformatics Institute (EBI) at http://www.ensembl.org/genome/central/. The data have been doubling every 15 months. By

2000 almost 1.3 million human ESTs were included as were sequence data on more than 25,000 other species.[311,314] A problem is that there are many errors in the data with some seriously incorrect sequences.[315] It is hard to manage the mass of new data, but in time most of the errors will probably be corrected. Books[316–319] and computer programs[320,321–321c] are available to help understand genomes. The widely used programs **BLAST** and **FASTA**, available on the Web or in computational packages, routinely compare protein-coding genes with known genes in order to predict function.[322–325] An inexpensive high-powered desktop computer and an internet connection will enable a person to do complex biocomputing.[325a] See also Chapter 5, Section H,7.

Genes must be transcribed, and most transcripts must be spliced, modified, and translated by the ribosomal machinery. Genes cannot be fully understood without considering these processes, which are dealt with in Chapters 28 and 29.

Human variation. Between any two unrelated human beings there are on the average one base pair that is different out of every 500 – 1000 nucleotides. This amounts to ~ 4×10^6 differences in the whole genome.[305,326] In addition to these single nucleotide polymorphisms (SNPs) there are many differences in the >100,000 nearly randomly dispersed short tandem repeats of microsatellite DNA.[189,327] The latter form the basis of DNA fingerprinting (Box 5-D), and together with the SNPs are helpful in tracking genetic diseases. One of the more difficult goals is to identify genes that confer increased susceptibility to cancer and to the complex syndromes of diabetes and mental problems. Searching for correlations with SNPs has proved difficult.[327a,b] A new approach termed **haplotype mapping** may permit correlations of disease susceptibility with larger blocks of conserved DNA sequence known as haplotypes.[327c,d] Some controversy surrounds possible uses of these maps, as also is the case for proposals to conduct extensive genetic testing of populations including newborn infants.[327e,f,g] Another planned project is the sequencing of 1000 individual human genomes.[327h]

Studies of DNA have also shed light on human evolution.[328] Mitochondrial DNA,[329] as well as a variety of nuclear genes,[330] is being studied in attempts to establish approximate dates of evolutionary divergence. If we assume a constant rate of incorporation of nucleotide substitutions during evolution, we can use sequence data as a **molecular clock** to construct phylogenetic trees such as that of Fig. 1-5, which is based on the gene for 16S ribosomal RNA. A problem is that mutation rates of genes are not all the same, and there are not yet enough data to draw firm conclusions. Estimates based upon the maternally transmitted mitochondrial DNA (see Chapter 18) suggest that we are all descended from Africans who lived 100,000 to 200,000 years ago. Limited data from the Y chromosome, which is transmitted through males, agree with the mitochondrial DNA data.[328,328a] However, study of the pyruvate dehydrogenase gene has been interpreted to indicate a more ancient divergence of African and other peoples.[331] Analysis of DNA from a fossil bone shows that the mtDNA of Neandertal people differed from modern mtDNAs in an average of 26 positions. From this figure the mitochondrial DNA molecular clock predicts that the extinct Neandertal line diverged from ours ~ 550,000 to 690,000 years ago. The most ancient mtDNA found is Australian.[331a,b] Unfortunately, DNA in fossils is very unstable and has not been recovered from older fossils, e.g., from dinosaurs. However, the search for ancient DNA goes on.[331c,d]

We diverged from our closest ape relative, the chimpanzee, about 4 – 6 million years ago.[196,332] Chimpanzee and bonobo DNA sequences are ~98.8% identical to those of humans.[332a–e] If differences in inserted and deleted segments of genes are included, however, the identity drops to 95%.[332f] One of the longest DNA sequences to be compared among humans and apes is a 7-kb length around the pseudogene in the β globin cluster (Fig. 27-10). In this sequence chimpanzees are closest to humans with gorillas being the next closest.[195] We may well ask in what way we differ from these apes? Some specific differences have been found. Notably, human beings do not hydroxylate the glycolyl groups of sialic acids (Chapter 4) to form *N*-glycolylneuraminic acid residues on glycoproteins.[328,333] Could this really be the most important difference between us and the apes? More genomic analysis may tell.

What does DNA analysis tell us about race? Most investigators conclude that there is only one human race with no detectable boundaries between the group commonly referred to as races.[334–336] As Pääbo put it, "in terms of the variation at most loci, we all seem to be Africans, either living on that continent or in recent exile."[328] The differences in skin color seem to reflect adaptation to the environment in which people live.[336a] Variations in the level of melanocyte-stimulating hormone receptor, one regulator of skin color (Box 25-A), are especially high in Africans.[328]

DNA analysis has also been useful in tracing human migration.[337–338e] For example, a genetic marker in the Y chromosome is carried by 85% of native Americans, suggesting that they are all descended from a man who lived ~ 20,000 years ago, probably an immigrant from Siberia.[339] Contrary to usual assumptions women, more often than men, seem to have spread their DNA to new locations in the world.[327b] Studies of cattle and of the wild ox reveal information about domestification of these animals about 10,000 years ago in Europe, Asia, and Africa.[340]

Other evolutionary relationships. Studies of chromosome banding, chromosome maps, restriction fragments, and detailed sequences provide many insights into relationships among species. For example, the chromosome banding pattern and also DNA sequences show close similarities between human beings and the mouse.[194,312] The latter is often regarded as the premier organism for the study of mammalian genetics and development.[291a–c,341] Dense genetic maps are available for both the mouse and rat as are moderate-resolution maps for livestock, companion animals, and other mammals.[327,342,343] Comparative gene maps are being constructed for more than 40 mammals[344] and other species of animals, plants, and fungi. Comparisons of these genomes reveal much of interest. For example, the pufferfish *Fugu rupripes* has a genome only ⅑ the size of the human genome. However, both species seem to contain about the same number of genes. Many of them can be directly correlated and some human-disease-causing mutations have been identified first in the Fugu.[344a]

Evolutionary history is being rewritten in molecular terms. Comparison of sequences of individual proteins allows evolutionary relationships of their genes to be traced. Many families of **homologous** genes can be identified.[344b] These include both **orthologs**, genes in different species that have evolved from a common ancestor, and **paralogs**, genes related by gene duplication within a genome. Orthologs have the same function in different organisms, but paralogs have different functions within a single species.[345]

Sophisticated molecular clock studies suggest that the evolution of metazoan organisms began earlier than had been supposed. Ancestral primates appeared at least 65 million years ago.[344] Gene sequence data for many species suggest that a great variety of mammals lived 100 million years ago in the age of dinosaurs,[346] a view also supported by new fossil evidence.[344,347] Metazoans appeared earlier than the “Cambrian explosion” generally thought to have occurred ~550 million years ago.[348] New geochemical data suggest that cyanobacteria diverged from other bacteria as early as 2.1 x 10^9 years ago.[349] Gram-negative bacteria diverged from gram-positive microbes ~3.2 x 10^9 years ago,[350] *Salmonella* from *Escherichia* only 0.1 x 10^9 years ago. However, DNA analysis shows that within this latter time period many genes from other microorganisms have been inserted into the *E. coli* chromosome and into other bacterial chromosomes.[351,352] Some of these transfers have occurred with the help of bacteriophages.[353] A puzzle is the fact that among eukaryotic cells the enzymes catalyzing the genetic **information transfer** via transcription and protein synthesis resemble those of archaea. However, **operational enzymes** that catalyze other basic metabolic processes tend to be more similar to those of eubacteria such as *E. coli*.[353,354] So much gene transfer between organisms has occurred that it is difficult to establish the earliest parts of a phylogenetic tree of the type shown in Fig. 1-5. Another factor that confuses our study of bacteria is that less than 1% of all living microorganisms have been grown in pure cultures.[355,356]

We are still dangerously ignorant of the complexity of the microbial world, which both threatens us with diseases and sustains our environment. We do have complete genome sequences for more than 60 different bacteria with hundreds more expected within a few years.[356,356a] They contain from 0.58 Mbp (*Mycoplasma genitalium*) to 8.7 Mbp for the antibiotic-producing *Streptomyces coelicolor*,[356b] a size shared by the legume symbiont *Bradyzhizobium japonicum*.[356c] The latter is one of many bacteria that have genomes split into two or more parts, often a major chromosome plus one or more plasmids. For example, the 6.7-Mbp genome of *Sinorhizobium melitoti*, an alfalfa symbiont, comprises a 3.65-Mbp chromosome and 1.35- and 1.68-Mbp megaplasmids.[356d] The 5.67-Mbp genome of *Agrobacterium tumefaciens*, much used in genetic engineering (Section E,4), consists of a circular chromosome, a linear chromosome, and two plasmids.[146a–c] The genome sequence revealed a close similarity to those of the above-mentioned legume symbionts.

Sequences of many pathogenic bacteria are known.[356e–o] These include the causative agents of cholera,[356c] typhoid fever,[356f] plague,[356g] brucellosis,[356h] leprosy,[356i] tuberculosis,[356i] and anthrax.[356j] Also included is a virulent strain of *Streptococcus pneumonia* (respiratory infections, ear aches, meningitis),[356k] *Pseudomonas aeruginosa* (a common “opportunistic” pathogen),[356l] *Listeria monocytogenes*, which causes a severe food-borne disease (Box 19-C),[356m] the pathogenic *E. coli* 0157:H7 (see Fig. 1-2),[356n] and a tiny mucosal pathogen *Ureaplasma urealyticum*.[256o] Many surprises were found in the genome sequences. For example, *E. coli* 0157:H7 is related to the laboratory strain *E. coli* K-12, and the two strains share a ~4.1 Mbp common “backbone” sequence. However, the pathogenic strain has hundreds of “islands” of additional DNA spread throughout the genome and amounting to ~1.54 Mbp. Many of these carry genes associated with virulence.[356n]

Specialized features appear in virtually every genome.[356a] For example, the mycoplasma *U. urealyticum* contains genes for enzymes that allow the bacterium to obtain almost all of its ATP from hydrolysis of urea.[356o] Both *Caulobacter crescentus*[356p] and *E. coli* develop specialized structures involved in motility (Figs. 19-3, 32-1). *C. crescentus* and also spore-forming bacteria have alternative developmental plans (Fig. 32-1). At the low end of the genome size range are species of *Buchnera*, which are endocellular symbionts of aphids.[356q] Their genome is only slightly larger than that of *M. genitalium*. The extremely salt-tolerant *Halobacterium*[356r] and the heat-tolerant *Thermoplasma volcanium*[356s] are among the archaea for which complete genome sequences are known. Species of *Xanthomonas*, whose genomes have been sequenced, are economically important plant pathogens. *Xanthomonas campestris* is also grown commercially to produce xanthan gum (p. 179).[358a] Determination of the genomes of mycobacteria has been challenging.[356i,t] Among these slow-growing organisms is *Mycobacterium tuberculosis*, the causative agent of human tuberculosis. A large fraction of its ~4000 coding genes is devoted to metabolism of lipids and to the synthesis of unusual proteins and lipids of its cell wall.[356t]

The larger genomes code for many proteins of unknown function, but over 80% of the "ORFs" (presumed genes) of *Haemophilus influenzae* have been identified, as have their presumed functions.[357,358] The encoded proteins appear to catalyze 488 metabolic reactions on 343 different metabolites. Together these systems provide a **metabolic genotype**.[357] Results of such analyses are accumulating in **metabolic databases**.[359,359a] What is the minimum number of metabolic reactions needed for support of life? Transposon insertions can inactivate all but from 265 to 350 of the protein-encoding genes of *M. genitalium* without killing the bacteria.[360] From comparisons of a variety of bacterial chromosomes it seems likely that ~256 of these genes are truly essential.[350,361] This conclusion leads to an interesting question. Is it ethical to now try to generate such a minimal bacterium?[361a] Are there hazards, e.g., that its escape might endanger our health or the environment? On the other hand, genetic engineering of bacteria, which is already practiced, can provide useful improvements in bacteria used in foods and in industry.[362,362a]

Modeling with the aid of data available on the World Wide Web is leading to development of new mathematical descriptions of metabolic networks.[359a] An ambitious new project is to model the entire *E. coli* cell. Many experimental data will be required and it has been estimated that ten years will be needed. The effort involves investigators in many laboratories and will be at least ten times as complex as the determination of the human genome.[362b]

Metabolic studies of eukaryotic cells. The yeast, *Saccharomyces cerevisiae*, contains ~6200 genes of which, until recently, only 40% had been assigned a function. Now a variety of methods are being employed to understand this little fungus.[363–365] A useful approach is to systematically inactivate or "knock out" genes. Davis and associates[363] used a PCR-based strategy to delete one gene at a time of 2026 yeast genes. Of these genes 1620 were found not essential for growth in a rich medium. Ross-McDonald and coworkers engineered a transposon, a 274-bp derivative of the *E. coli* Tn3, and allowed it to be inserted into the genes of yeast cells by homologous recombination (Chapter 27, Section D). The transposon carried DNA for a short peptide tag in the form of a specific immunological epitope that could identify the transformed cells. More than 11,000 strains with disruptions in nearly 2000 genes were obtained. These and other deletion mutants are now available for study.[364a,365] A second yeast, *Schizosaccharomyces pombe*, has assumed major scientific importance in studies of the cell cycle and of metabolism.[365a] Its 13.8-Mbp genome is only a little smaller than that of *S. cerevisiae*, but it has ~1400 fewer recognized genes, a total of 4824.[365b] There are smaller eukaryotic genomes. That of the tiny marine chordate *Oikopleura dioica* may be only 51 Mbp. However, there may be a total of ~15,000 genes.[365c]

Transposon-induced mutations have also been created in nearly one-fourth of the ~12,000 genes of *Drosophila melanogaster*.[283] Studies of the expression of genes in both *Drosophila* and in the nematode *Caenorhabditis elegans* are directed toward understanding of development and differentiation. Of the nematode's predicted 19,293 genes, only 7% have been studied at the biochemical level. To understand what happens during development we need methods for studying simultaneously the expression of all of these genes. One approach is to look at messenger RNAs that are formed at different times during development. More than 9000 mRNAs have been identified in cells of *C. elegans*, and their patterns of expression have been observed using DNA microarrays.[366,367] A similar technique has been applied to yeast. DNA sequences of fragments of ~6400 yeast genes were amplified by PCR and printed onto a glass plate to form a "DNA chip." From mRNAs formed at different times during growth, fluorescent cDNA copies were made, and their amounts were checked by use of the DNA chip.[368] Another approach is to look directly at the proteins formed. Walhout and coworkers have devised a large-scale automated system for cloning all of the genes of *C. elegans*, expressing the protein products, and testing them in the yeast two-hybrid system (Box 29-F) for protein–protein interactions.[369] Another project is to use large-scale sequence comparisons between proteins of *C. elegans* and of other organisms to identify nematode genes that encode extracellular matrix proteins involved in cell adhesion and to trace their evolution.[370]

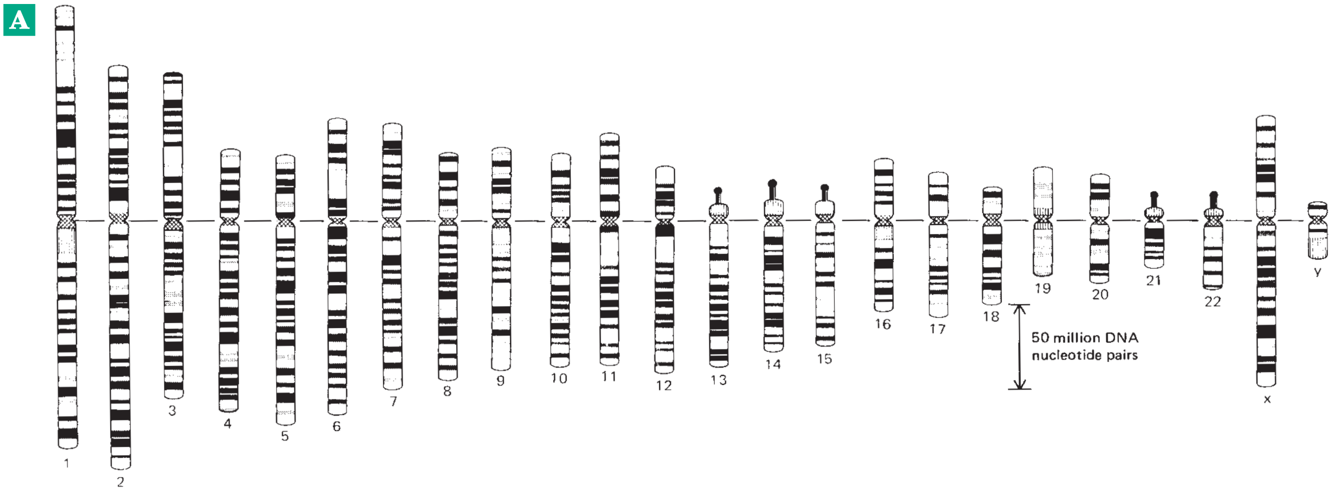
The genomes and the metabolism of the two insects *Drosophila melanogaster* and *Anopheles gambia*[370a] can now be compared. Many differences can be seen but almost half of the genes are orthologs. Many of these can be related also to those of pufferfish, mice,[370b] humans, and other species.

Among plant genomes that of *Arabidopsis thaliana* has been studied most. The sequences of its five chromosomes have been determined and analyzed.[371–375c]

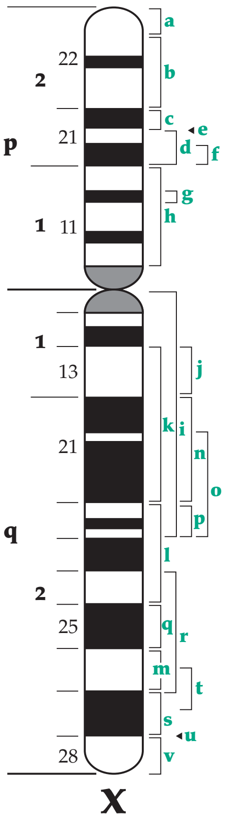
About 25,498 genes encoding proteins from 11 different families have been found.[375a] More than 14,000 ESTs were established from cDNAs.[375b] Many of the genes represent new families, some of which may be peculiar to plants. However, many others are homologous to those of *C. elegans* and *H. sapiens*. For example, developmentally important homeotic genes, marked by a **homeo sequence** as in animal genes, are present as are thousands of cell surface receptors. However, only Ser / Thr and histidine kinases are present in *Arabidopsis*. No tyrosine kinases have been identified.[375]

The ~125-Mbp *Arabidopsis* genome is tiny compared to the 3000 Mbp of DNA present in the genomes of maize and of many other plants.[374] However, the rice genome, only ~15% as large as that of maize,[376] has been chosen for complete sequencing.[375d,e] Draft sequences for the genomes of two subspecies of rice (*Oryza sativa L.*), ssp. *indica*[375f] and ssp. *japonica*,[375g] were published in 2002. The much larger genomes of other cereal grains (2500, 4900, and 16,000 for maize, barley, and wheat, respectively) will probably all be sequenced within a few years. The genes of maize are being mapped and studied using transposon-tagging

A



B



a Ichthyosis, X-linked
Placental steroid sulfatase deficiency
Kallmann syndrome
Chondrodysplasia punctata, X-linked recessive

b Hypophosphatemia
Aicardi syndrome
Hypomagnesemia, X-linked
Ocular albinism, X-linked
Retinoschisis

c Adrenal hypoplasia
Glycerol kinase deficiency

d Chronic granulomatous disease
Retinitis pigmentosa-3

e Duchenne muscular dystrophy
Becker muscular dystrophy

f Ornithine transcarbamylase deficiency

g Norrie disease
Retinitis pigmentosa-2

h Incontinentia pigmenti
Wiskott-Aldrich syndrome
Menkes syndrome

i Androgen insensitivity
Kennedy disease

j Sideroblastic anemia
Aarskog-Scott syndrome
PGK deficiency hemolytic anemia

k Charcot-Marie-Tooth disease
Choroideremia
Cleft palate, X-linked
Spastic paraplegia, X-linked, 2
Deafness with stapes fixation

l PRPS-related gout

m Lesch-Nyhan syndrome
HPRT-related gout

n Anhidrotic ectodermal dysplasia

o Agammaglobulinemia

p Pelizaeus–Merzbacher disease
Alport-like hereditary nephritis
Fabry disease

q Lowe syndrome

r Immunodeficiency, X-linked with hyper IgM
Lymphoproliterative syndrome

s Hemophilia B

t Albinism-deafness syndrome

u Fragile X syndrome

v Hunter syndrome
Hemophilia A
G6PD deficiency: Favism
Drug sensitive anemia
Chronic hemolytic anemia
MASA syndrome
Colorblindness (several forms)
Dyskeratosis congenita
TKCR syndrome
Adrenoleukodystrophy
Adrenomyeloneuropathy
Emery-Dreifuss muscular dystrophy
Diabetes insipidus, nephrogenic, X-linked
Myotubular myopathy, X-linked
Otopalatodigital syndrome, type I

Figure 26-15 (A) Numbering and staining patterns (G-banding) of human chromosomes. The horizontal line marks the centromeres. Chromosomes 13, 14, 15, 21, and 22 have nucleolar organizers located at the constrictions in their short arms. From Alberts *et al.*[297] Adapted from U. Franke.[298] (B) Map of genetic defects identified in X chromosomes by 1995. From McKusick and Amberger.[294]

methods.[377] The smaller, more compact genomes may have just as many genes as the longer ones but have less repetitive DNA.[378] Some ferns have 307,000 Mbp of DNA, nearly 100 times that of a human; bony fishes have ~307,000 Mbp and the ameba >200,000 Mbp. These organisms appear to lose unneeded repetitive DNA faster than those with smaller genomes.[378]

3. Understanding Human Genetic Diseases

Genetic diseases have always been with us, but it was not until 1949 that the first disease, sickle cell anemia (Box 7-B), was understood at the molecular level. A single base substitution in DNA and the resultant single amino acid substitution in hemoglobin causes this disastrous disease. It was soon recognized that defects in single proteins are causes of other inherited diseases. Many of the hundreds of other known genetic disorders[201,203,379] are discussed elsewhere in this book. Among them are muscular dystrophies and cardiomyopathies (Box 19-A),[380] lysosomal deficiency diseases (Chapter 22), problems with ion transporters[381–383] and channels (cystic fibrosis, Box 26-A), defective collagens (Box 8-E),[384] neurological disorders (Chapter 30), and defects in defense systems (X-linked granulomatous disease, Chapter 18, Section G). Many of these were first recognized by their frequent occurrence in boys. Some of these X-linked deficiencies are mapped in Fig. 26-15B.

One insight into molecular disease was the recognition that mutations that cause many diseases, e.g.,

BOX 26-A CYSTIC FIBROSIS

One of the commonest of genetic diseases, cystic fibrosis affects persons all over the world. The incidence is unusually high in persons of northern European descent, about one in 2500 children being born with the defect. The inheritance pattern showed that cystic fibrosis is recessive and is caused by a single-gene defect that is carried by almost 5% of white Americans. In the United States there are ~30,000 persons with the disease. Many die in early childhood and even with careful treatment only 50% live into their late twenties or beyond.[a,b] Through extensive linkage analysis the cystic fibrosis gene was mapped to chromosome 7, and its location was narrowed further to a 1600-kbp region between the oncogene *met* and another marker designated J3.11.[c] Random searching located closer markers, and "chromosome walking and jumping" led to identification and characterization of the gene in 1989.[c,d] The large 250-kbp gene contains 27 exons. The transcribed mRNA is 6129 bp in length, and the gene product is a 1480-residue amino acid protein,[e] which is known as the **cystic fibrosis transmembrane conductance regulator** (**CFTR**).[f]

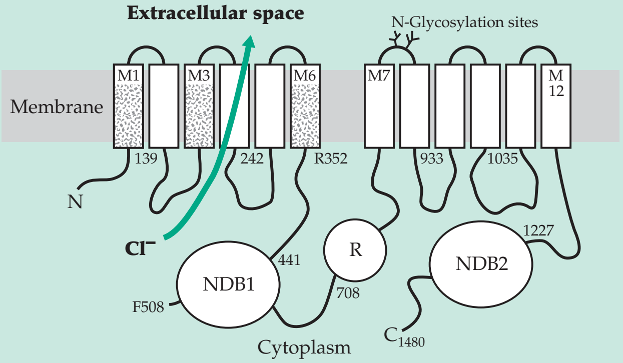
Children with cystic fibrosis lose excessive amounts of salt in perspiration and become dehydrated readily. A salty taste of the skin and an elevated chloride concentration of sweat are traditional diagnostic symptoms.[a] More serious problems arise from progressive respiratory failure and inadequate pancreatic secretion. Lung infections with *Pseudomonas aeruginosa* are the major cause of death.[g] The CFTR gene is expressed in many tissues, especially those of the mucous membranes. An alternatively spliced isoform may form chloride channels in heart.[h,i] As mentioned in Chapter 8, Section C,5 the CFTR protein is a member of the ATP-dependent ABC transporter family. However, it is atypical because it also contains a regulated chloride channel.[j] In secretory epithelia of intestines, pancreas, lungs, sweat glands, and kidneys Cl^- enters epithelial cells through their basolateral surfaces using an $Na^+ + K^+ + 2\ Cl^-$ cotransporter and exits the cells through their apical surfaces using the CFTR channel. Absorptive epithelia also contain both the cotransporter and the CFTR channel, but Cl^- flows into the cells from the exterior surface, and the distribution of the cotransporter and CFTR between basolateral and apical surfaces is opposite to that in secretory cells.[f]

From the amino acid sequence the CFTR protein is predicted to form two 6-helix transmembrane domains, two ~240-residue cytosolic ATP-binding domains, and a cytosolic regulatory (R) domain that contains at least five serine residues that may be phosphorylated by the cAMP-dependent **protein kinase A** (Chapter 11, Section C; Fig. 11-4).[k,l,m] See adjacent scheme. The two ATP-binding domains resemble those of myosin and other ATP-hydrolyzing proteins. The chloride channel, which is probably formed by helices M1, M3, and M6,[f] remains closed unless serine residues of the regulatory domain are phosphorylated. However, opening of the channels also requires binding of ATP in the nucleotide-binding domains. Binding of either vanadate or BeF_x stabilizes the open state of the channels.[k] There are four small 55- to 65-residue cytosolic loops, labeled 1–4 in the diagram. Transmembrane loop 3 also appears to function in regulation of the channel,[n] which involves a complex regulatory mechanism. Two sites of *N*-linked glycosylation are present in the short extracellular loop between helices M7 and M8.

cystic fibrosis, while affecting a single protein, occur at many places in the gene that encodes the protein. Not all mutations are caused by base substitutions; they are often a result of deletion or insertion of DNA. A whole group of diseases are caused by accumulation of repetitive DNA, often of nucleotide triplets, within genes or in control regions of genes. Another important insight is that understanding a newly discovered and very rare disease may help us to understand other related disorders. For example, after the Duchenne muscular dystrophy gene was located, the encoded protein dystrophin was found to have mutations resulting in other milder dystrophies (Box 19-A). Mutations are only rarely beneficial, but we know that many mutations alter the properties of proteins very little. We can anticipate that most genes may undergo mutations that cause some loss of good health and vitality without being diagnosed as causing disease.

We have also come to understand that many complex diseases such as diabetes, **polycystic ovary syndrome**,[385] **Crohn's disease** (inflammatory bowel disease),[386] and schizophrenia are in fact multiple diseases. Diabetes is a syndrome that can arise from causes such as defective insulin receptors or defective glucose transporters or from as yet unknown metabolic problems (Chapter 17).[387] Many cancers have a

BOX 26-A CYSTIC FIBROSIS (continued)



The CFTR protein can undergo endocytosis into clathrin-coated vesicles as part of its regulatory mechanism.[o] Since HCO_3^- is usually exchanged for Cl^- in epithelial ion transport, regulation of HCO_3^- uptake is also a significant aspect of CFTR function.[p]

Mutations that cause cystic fibrosis are found at many locations in the gene. However, ~70% of the mutations are caused by the absence of phenylalanine 508, as a result of a three-nucleotide deletion, in the first nucleotide-binding domain.[e,k,p] This deletion causes misfolding of the CFTR to give an inactive protein.[p] Hundreds of other mutations, some in the regulatory domain and some in the cytosolic loops,[q,r] also cause the disease. Cystic fibrosis is one of the diseases for which targeted gene transfer may become an effective treatment.[a,s] Cystic fibrosis induced in mice by targeted disruption of the CFTR gene has been successfully treated by gene therapy.[t] However, the gene is large and efficiency of transfer into animals is low. Nevertheless, human gene therapy for cystic fibrosis is being pursued cautiously.[s]

The CFTR protein has additional medical significance. Its stimulation by bacterial toxins is responsible for **secretory diarrhea**, which kills 3 million young children annually.[f]

---

[a] Welsh, M. J., and Smith, A. E. (1995) *Sci. Am.* **273**(Dec), 52–59

[b] Koh, J., Sferra, T. J., and Collins, F. S. (1993) *J. Biol. Chem.* **268**, 15912–15921

[c] Marx, J. L. (1989) *Science* **245**, 923–925

[d] Riordan, J. R., Rommens, J. M., Kerem, B.-s, Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.-L., Drumm, M. L., Iannuzzi, M. C., Collins, F. S., and Tsui, L.-C. (1989) *Science* **245**, 1066–1073

[e] Collins, F. S. (1992) *Science* **256**, 774–779

[f] Akabas, M. H. (2000) *J. Biol. Chem.* **275**, 3729–3732

[g] Ernst, R. K., Yi, E. C., Guo, L., Lim, K. B., Burns, J. L., Hackett, M., and Miller, S. I. (1999) *Science* **286**, 1561–1565

[h] Hart, P., Warth, J. D., Levesque, P. C., Collier, M. L., Geary, Y., Horowitz, B., and Hume, J. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6343–6348

[i] Howell, J. M., and Gawthorne, J. M., eds. (1987) *Copper in Animals and Man*, Vol. I and II, CRC Press, Boca, Raton,Florida

[j] Li, C., Ramjeesingh, M., Wang, W., Garami, E., Hewryk, M., Lee, D., Rommens, J. M., Galley, K., and Bear, C. E. (1996) *J. Biol. Chem.* **271**, 28463–28468

[k] Gadsby, D. C., and Nairn, A. C. (1994) *Trends Biochem. Sci.* **19**, 513–518

[l] Ikuma, M., and Welsh, M. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8675–8680

[m] Ostedgaard, L. S., Baldursson, O., and Welsh, M. J. (2001) *J. Biol. Chem.* **276**, 7689–7692

[n] Aleksandrov, L., Mengos, A., Chang, X.-b, Aleksandrov, A., and Riordan, J. R. (2001) *J. Biol. Chem.* **276**, 12918–12923

[o] Weixel, K. M., and Bradbury, N. A. (2000) *J. Biol. Chem.* **275**, 3655–3660

[p] Luo, X., Choi, J. Y., Ko, S. B. H., Pushkin, A., Kurtz, I., Ahn, W., Lee, M. G., and Muallem, S. (2001) *J. Biol. Chem.* **276**, 9808–9816

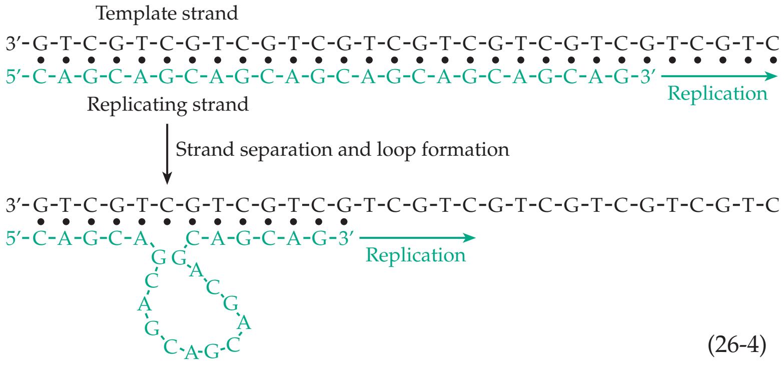
[q] Seibert, F. S., Jia, Y., Mathews, C. J., Hanrahan, J. W., Riordan, J. R., Loo, T. W., and Clarke, D. M. (1997) *Biochemistry* **36**, 11966–11974

[r] Cotten, J. F., Ostedgaard, L. S., Carson, M. R., and Welsh, M. J. (1996) *J. Biol. Chem.* **271**, 21279–21284

[s] Verma, I. M., and Somia, N. (1997) *Nature (London)* **389**, 239–242

[t] Hyde, S. C., Gill, D. R., Higgins, C. F., Trezise, A. E. O., MacVinish, L. J., Cuthbert, A. W., Ratcliff, R., Evans, M. J., and Colledge, W. H. (1993) *Nature (London)* **362**, 250–255

strong hereditary component, and many of these are being mapped to specific DNA locations.[388] Some specific proteins, such as **Ras** (Chapter 11) that are mutated in many cancers were first recognized as avian oncogenes. The tumor suppressors **Rb** (retinoblastoma protein, Box 11-D)[388,389] and **p53** (Fig. 11-15)[390,391] are major sites of mutation in cancer. See also Chapters 11 and 31. Since cancers contain multiple mutations, they are complex diseases. However, many specific susceptibility loci are being located, including some for breast cancer (Box 11-D),[392] prostate cancer,[393] and familial adenomatous polyposis, a hereditary disease leading to colon cancer.[388,394] Cancer has long been known to be associated with chromosome instability including deletion and insertion mutations at simple repeat sequences, frame-shift mutations,[395] DNA breakage, translocation,[396] and losses or gains of whole chromosomes.[397]


Template strand
3′-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C
5′-C-A-G-C-A-G-C-A-G-C-A-G-C-A-G-C-A-G-C-A-G-C-A-G-3′
Replication
Replicating strand
Strand separation and loop formation
3′-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C-G-T-C
5′-C-A-G-C-A
C-A-G-C-A-G-3′
Replication
G G A C G A C G A C A G

(26-4)

From a practical viewpoint the understanding that we are gaining will help us to provide better treatments of genetic diseases. At present almost every human genetic disease can be mimicked in a knockout mouse.[291a,398] A turnabout is that rare human hair-

TABLE 26-4
Some Human Triplet Repeat and Related Diseases[a]

Name	Repeat sequence (5' to 3')	Transmission[b]	Location
Fragile X syndrome (FRAXA)	$(CGG)_n$	XD	by 5' side of gene *FMR*-1 + C_pG island methylation
Type E (FRAXE)	$(CGG)_n$	XR	5' side of gene *FMR*-2
Synpolydactyly	$(GCG)_n$		HOXD13 gene (polyalanine)
Myotonic dystrophy (DM1)	$(CTG)_n$	AD	3' untranslated region of gene for cAMP-dependent protein kinase
DM2	$(CCTG)_n$	AD	Intron in a zinc finger protein gene
Friedreich ataxia	$(GAA)_n$	AR	Intron in frataxin, a mitochondrial protein
Huntington disease (HD)	$(CAG)_n$	AD	Huntingtin (polyglutamine)
Spinocerebellar ataxia, SCA-1	$(CAG)_n$	AD	Ataxin-1 (polyglutamine)
SCA-2	$(CAG)_n$	AD	Ataxin-2 (polyglutamine)
Machado–Joseph disease (SCA-3)	$(CAG)_n$	AD	Ataxin-3
SCA-6	$(CAG)_n$	AD	Calcium channel
Spinobulbar muscular atrophy (Kennedy disease; SBMA)	$(CAG)_n$	XR	Atrophin (androgen receptor)
Dentato-rubro-pallido-luysian atrophy (DRPLA)	$(CAG)_n$	AD	
Progressive myoclonus epilepsy	(G+C)-rich oligonucleotide repeat		Cystatin B

[a] See Mandel, J.-L. (1997) *Nature (London)* **386**, 767–769 and Richards, R. I., and Sutherland, G. R. (1997) *Trends Biochem. Sci.* **22**, 432–436

[b] XD, X dominant; XR, X recessive; AD, autosomal dominant; AR, autosomal recessive

less people have a mutation in a gene homologous to the well-known *hairless* gene of mice.[399] Genes homologous to those of many human genetic defects have also been identified in yeast.[400] Studies of both mutant mice and of mutant yeast cells can also help in understanding diseases and in devising therapies.

Triplet repeat diseases. With the exception of Down disease (extra chromosome 21, affecting 1 in 600 children) the most prevalent cause of mental retardation is the **fragile X syndrome,** which affects ~1 of 2000 newborn males. A fragile site,[401,402] where the X chromosome breaks easily, marks the location of the defect. Identification of nearby RFLPs led to cloning of the gene in 1991. The defect was found to be a repeated trinucleotide sequence 5'-$(CGG)_n$ in the DNA. The value of n varies, in most healthy individuals averaging ~30. However, for some normal individuals n may be 200–300. These persons may transmit the fragile X disease to their offspring in whom n increases from one generation to the next with increasing severity of the disease. There may be 2000 or more CGG triplets.[403] The defect lies at the 5' end in an untranslated part of the gene for the fragile X mental retardation protein (FMRP). A cytoplasmic RNA-binding protein, FMRP, may enter the nucleus and have an as yet uncertain function. In rare cases the fragile X syndrome arises from deletions or missense mutations in the *FMRP* gene.[404] There are actually two similar genes *FMRP*1 and *FMRP*2. Mutations in the latter are associated with a milder form of fragile X disease. FMRP genes with expanded $(CGG)_n$ tracts are not expressed, evidently because the mutation induces methylation of an adjacent "CpG island" (see Chapter 27) as well as of sites within the FMRP gene. Both FMRP proteins are apparently needed for normal brain function. The FMRP1 defect is genetically dominant, and female heterozygotes also suffer from the fragile X syndrome. However, FMRP2, also encoded on the X chromosome, is recessive.[405] FMRP appears to function in neuronal dendrites (Chapter 30) where it binds to polysomal aggregates and participates in regulation of translation of mRNA.[405a–c] Like DNA, mRNA may contain **G quartets** (Fig. 5-8, p. 227).[405d]

Twelve or more additional triplet repeat diseases, many with neurological symptoms, have been identified (Table 26-4).[405–407a] These involve other trinucleotide repeats 5'-$(GCG)_n$, 5'-$(CTG)_n$, 5'-$(GAA)_n$, and 5'-$(CAG)_n$. In **synpolydactyly**, an inherited developmental defect causing malformation of hands and feet, an expansion of a GCG trinucleotide occurs within the gene *HoxD*13. This results in incorporation of a polyalanine tract near the N terminus of the protein.[408] **Myotonic dystrophy** DM1 (Box 19-A) results from expansion of CTG to 6 kbp or more within the untranslated 3' region of a gene for cAMP-dependent protein kinase.[405] The mRNA transcripts accumulate in the nucleus and may bind to a CUG-binding protein that is involved in splicing other mRNAs, thereby poisoning the cell.[409] DM2 is caused by expansion of a CCTG quartet in an intron of a zinc-finger protein.[408a] Expansion of the GAA triplet is associated with the neurological disease **Friedreich ataxia**, which has a prevalence of ~1 in 50,000. The defect lies within an intron found in the gene for the 210-residue mitochondrial protein **frataxin**. The function of the protein is unknown.[409a] However, the defect in Freidrich ataxia leads to a deficit in mitochondrial ATP synthesis.[409b] Studies of a corresponding protein in yeast suggests that frataxin is an iron storage protein.[409c–e] Apparently the DNA defect in the intron interferes with splicing of the mRNA transcript.[410–413] The polyGAA strand in the triplet repeat region of the DNA is able to form various alternative structures including a parallel (GAA)•(TTC) duplex.[410a] Such structures in mRNA may interfere with proper splicing.

A defect in an α-tocopherol-transfer protein causes a similar set of neurological symptoms. Oxidative damage may therefore be a component of this disease.[410b]

Expansion of $(CAG)_n$ sequences causes a series of neurodegenerative diseases (Table 26-4),[413–416c] the commonest of which is **Huntington disease** (HD).[414–414c] In Huntington's disease the $(CAG)_n$ tract is found in the **huntington** gene. The protein is enormous, with over 3140 residues. It is essential for nerve development, but the function remains uncertain. The first exon at the HD gene encodes a polyglutamine tract of 6–35 residues, which when expanded to 36–100 or more causes the disease. A similar situation holds for the **spinocerebellar ataxias** (SCA) and other $(CAG)_n$ diseases (Table 26-4).[414,417–417b] The encoded proteins have a variety of functions. The SCA-1 protein ataxin-1 functions in the nuclear matrix.[418] The X-linked **spinobulbar muscular atrophy** (SBMA) gene encodes an androgen receptor protein.[414,417] The polyglutamine sequences inserted in those proteins seem to be toxic, but the mechanism of toxicity is uncertain. It may result from aggregation of the proteins within cells.[414] It has also been suggested that cleavage of these chains by **caspases** (cysteine proteases; Chapter 12) may produce truncated proteins, which induce apoptosis.[415]

A form of epilepsy (Table 27-4) appears to be a result of repeats of a (G + C)-rich sequence that may be a dodecamer.[405] Dinucleotide repeats and other "minisatellite" DNA sequences are also associated with instability of DNA and may undergo expansion.[419–421] A pentanucleotide repeat $(CCTTT)_n$ is associated with increased expression of the nitric oxide synthase gene NOS2A. Persons with $n = 14$ were found to have enhanced resistance to development of diabetic retinopathy. This seems to be a case of a beneficial "gain of function" mutation.[422]

How do repeat sequences expand from one generation to the next? There is probably more than one mechanism. One is **strand slippage** during DNA replication. If single-strand loops are present at any time one strand could be displaced (slipped) relative to the other. Replication could then either expand the repetitive sequence or cause a deletion (Eq. 26-4).[405,423–425] Expansion could also occur by gene conversion during homologous recombination (Chapter 27, Section D)[401,426] or during DNA repair.[421,425] Repeat sequences may also prevent proper formation of nucleosomes.[402]

Cloned genes and diagnosis. The first areas in which cloned DNA has affected medicine are in the diagnosis of genetic diseases and in the production of medicinally useful proteins in bacteria, yeast, or cultured cells. One of the first applications was for the diagnosis of the sickle-cell trait by use of PCR on DNA isolated from blood.[43] This was followed quickly by methods for recognition of other inherited diseases and by automated procedures. Intense efforts are now being made to develop DNA "chips" (usually small glass plates with an array of DNA fragments bound to the surface) that can recognize a great variety of defects. For example, DNA polymorphisms in the 16.6-kbp human mitochondrial genome can be recognized by a plate containing 135,000 oligonucleotide probes assembled in a regular grid by photolithography and solid-state synthesis (Chapter 3).[427] Cancers may be classified quickly.[428] Many systems are under development for binding and recognizing genomic DNA, cDNA, or mRNA using microchip arrays of up to 400,000 or more oligonucleotides on a 2 cm x 2 cm plate.[429–430b] One chip combines PCR with use of "zip-code primers" that direct the PCR products to specific zip-code addresses on the chip to give a universal array able to detect low abundance of mutations in any gene of interest.[431] Mass spectrometry is also being harnessed to identify oligonucleotides bound at any address on a chip.[432] Commercial chips are expensive, \$100 – 2000 apiece, and good for only one use. The price will fall. To build a machine to make your own chips go to http//cmgm.stanford.edu/pbrown/mguide/.[433,434]

An alternative to DNA chips is to miniaturize DNA sequencing and analysis machinery. Using nanoliter droplets of fluid passing through microchannels built by photolithographic techniques of computer chip construction, these nanolaboratories may be the size of a credit card but able to cleave DNA and conduct PCR reactions, gel electrophoresis, and sequence determinations.[435,436]

Vaccines, hormones, and other medicines. A myriad of products of recombinant DNA technology are already in use in medical diagnostics.[437] In the past problems have arisen because vaccines and hormones can cause allergic reactions and may harbor viruses. A small percentage of diabetics are allergic to animal insulins, but human insulin produced in bacteria is now available. A number of children receiving human growth hormone isolated from cadavers contracted the fatal Creutzfeld–Jakob disease,[438] a neurological disorder caused by a prion (Chapter 29). One way in which AIDS has been spread is through contamination of the blood-clotting factors VIII or IX needed by hemophiliacs. These sources of contamination are being eliminated by the use of bacterially produced proteins.[439]

Recombinant DNA techniques can be used in two ways in the production of vaccines. The first is to find a protein in the virus or other infective agent that is a good inducer of antibody formation, i.e., a good antigen. This protein, or even a fragment of it, can then be produced from its cloned gene or can be made synthetically. Since the cloned protein can be purified highly, it may make a superior vaccine to those made from killed cells, inactivated virus particles, or mixtures of proteins.[440] The first commercial vaccine of this type was against viral hepatitis B, a major cause of liver cancer.[441] Particles consisting of viral envelope proteins can be produced in yeast and be used for vaccination. A DNA encoding these proteins has been transferred, using the Ti plasmid (Section E, 4), into lettuce. Human volunteers produced anti-hepatitis antibodies after eating the lettuce.[442] This suggests that vaccination through ingestion of antigenic proteins in food crops may be feasible. Injection of a small piece of DNA carrying the gene may also lead to antibody formation.[443,444] Using either purified antigens, proteins, or DNA, it may be possible to develop effective vaccines against Rocky Mountain spotted fever,[445] a rickettsial disease, and against malaria, for which there are no satisfactory vaccines.

The vaccinia virus, formerly used to vaccinate against smallpox, has been engineered for use against other diseases.[446] Much of its 187-kb DNA can be excised and replaced with passenger DNA. The virus particles are stable and highly infectious. The vaccinia virus is unusual in carrying genes for both RNA and DNA polymerases and other proteins that permit it to undergo replication and transcription of its own genes in the cytoplasm of the infected cells. One application of a recombinant vaccinia virus is oral vaccination of wild foxes and racoons against rabies.[447]

A major advance based on cloned genes is the production of new medicines previously unavailable or available in only small amounts. Among these are the **interferons**[287] (Chapter 31) and many hormones such as the **interleukins** produced by lymphocytes[448] and the **atrial natriuretic hormones** (Chapter 23). Another candidate is the **α1-protease inhibitor** (Chapter 12). Perhaps better inhibitors than the natural one can be devised and produced in bacteria.[449]

4. Gene Therapy

A few years ago it seemed like fantasy, but there is little doubt that we will soon be able to routinely treat some genetic illnesses by introducing new copies of genes into the body. A current goal is to insert cloned genes into body cells (somatic cells) to correct specific hereditary defects. For example, juvenile diabetics would benefit from introduction of genes for insulin production into cells that could replace their atrophied pancreatic beta cells. At present we don't know how to do this. However, genes have been transferred into human beings lacking adenosine deaminase and showing severe **combined immunodeficiency** (Chapter 31) and those lacking hypoxanthine guanine phosphoribosyltransferase and displaying the **Lesch–Nyhan syndrome** (Chapter 25, Section E,2). Corrections of **Gaucher disease** and other deficiencies of lysosomal enzymes is also an early goal (Chapter 20, Section G,2).

By 1986 more than 5000 children had received bone marrow transplants from close relatives to correct severe combined immunodeficiency caused by a defective adenosine deaminase gene.[450] However, the patients must receive chemotherapy or irradiation to suppress their immune system before the transplantation. Hospitalization may last for 30–60 days. Using genetic therapy some bone marrow cells can be removed from the patient with a needle. The cells can be treated to introduce the corrected genes using a suitable retrovirus. Clones of corrected bone marrow stem cells, which will give rise to lymphocytes in the body, can be selected, cutured and reintroduced into the patient. As discussed on p. 1498, retrovirus vehicles that locate and become integrated at appropriate sites in the genome have been developed.[216,451] Use of homologous recombination (Chapter 27) introduces the cloned gene into its normal chromosomal location.[452] The first children were treated by transfer of the adenosine deaminase gene in 1991. The procedure has been partially successful.[453,454]

By 1999 more than 400 clinical gene therapy trials were planned or in progress.[455] Nevertheless, development of suitable vehicles for gene delivery has been slow.[456,456a] Uncertainty about the safety of adenovirus vectors is one problem.[160a,454,457,457a] Poor efficiency of gene transfer is another. A glycogen storage disease of knockout mice has been cured by transfer of human α-glucosidase (Box 20-D) using an adenovirus vector.[458] Mice have also been used in developing gene therapy for hemophilia,[456a,457a] sickle-cell disease,[458a] and aspartylglycosaminuria.[458b] However, gene targeting in animals other than the mouse has been difficult.[459] Genetic therapy may be most effective when the gene is transferred into stem cells, which can then take up residence within the body (Chapter 32). Young stem cells, which can be obtained from unbilical cord blood at birth, can be used.[460,461] For many diseases therapy will probably be best soon after birth, or even prior to birth during the second trimester of pregnancy.[455]

An alternative approach is to synthesize highly specific hydrolysis-resistant DNA analogs that can form triple helical structures with DNA (Chapter 5). If these can be made specific enough they might bind to a targeted DNA site, such as the sickle cell anemia locus and induce a "back mutation" from the faulty A•T base pair (see Fig. 7-23) to a T•A pair in at least some of the hematopoietic cells that give rise to hemoglobin.[462] Another possibility is to make oligonucleotide analogs that serve as mimics of **antisense RNA** (discussed in Chapter 28). This could impede translation of bad mRNAs, such as that giving rise to polyglutamine chains in the triple repeat diseases.

At this time no effort is being made to alter the DNA in human germ cells. It seems undesirable to experiment with such changes.[160a,462a] However, as methods are developed for genetic therapy of somatic cells, we will rear more and more healthy carriers of serious genetic defects. Eventually we may need to develop therapy for germ cells.

5. Genetic Engineering of Bacteria, Plants, and Animals

Many improvements in bacteria used in industrial fermentations have been made.[463] The number of copies of a useful gene may be increased, and repressed genes may be made more active by deletion mutations in regulatory genes (Chapter 28). New genes are being transferred between bacteria and into plants and animals. For example, *Bacillus thuringensis* produces crystalline protein toxins, which are harmless to mammals but active against many insects. The *Bt* gene for this toxin has already been transferred into many different plants and is having a major effect on agriculture in allowing for decreased use of insecticides.[464,465] Genetic engineering may allow the toxin to be made more specific for particular insect species.

Genetic engineering of plant genes[466] may improve the quality of storage proteins in cereal grains[467] and the flavors of fruits,[468] provide more drought-resistant plants,[469] and offer increased resistance of crop plants to particular herbicides[464] (Chapter 25, Section A,1) and to viruses.[470] Many things can be done to enhance the nutritional qualities of foods.[464] For example, rice has been engineered to produce and store β-carotene in the grain, a development that may benefit 400 million people in the world deficient in vitamin A and often suffering from infections and blindness.[471] (See also p. 1240.) Specialty products such as poly-β-hydroxybutyrate and many others can also be produced in plants.[472] Genetic engineering of animals has produced not only knockout mice but the possibility for nutritional improvement[156] and for production of useful products such as vaccines in milk.

6. Ethical Problems

With a host of new medicines and agricultural products coming and with the ability to alter genomes at will, we face new ethical problems. For example, who owns the human genome?[473–475] Should patenting of human genes be allowed?[476,477] How private are genetic data?[478] Who has access to human DNA data?[479] Insurance companies?[480,481] Should routine screening with new genetic tests be used (as is now the case for phenylketonuria, diabetes, and others), even if no effective treatment is available?[482,483]

Ethical questions arise in application of experimental forms of genetic therapy to patients dying of inherited disorders. What if the therapy simply provides a longer life of suffering to patients? Some biologists foresee a future in which human beings learn to control their own genes and prevent genetic deterioration resulting from the accumulation of harmful mutations. They find it exciting to think that we can, in an intelligent way, elect to continue our evolution in a desired direction. Others caution that our present knowledge is such that attempts to eliminate all "bad" genes from the population might be disastrous.[484] They point to the hemoglobin S gene (Box 7-B) and the role it once had in preserving life in a malaria-infested environment and urge that at this stage we allow maximum heterogeneity of genetic types. There are other dangers in allowing genetics to control human lives. In the past "eugenic" doctrines have been used to justify racist laws and (in Nazi Germany) genocide.

Problems have resulted from the availability of hormones and other products of genetic engineering. For example, should parents be allowed to request human growth hormone for short children who do not have a pituitary deficiency? Should it be available to athletes and to aging people? Use of stem cells in genetic therapy raises another set of questions. Present guidelines of the National Institutes of Health (NIH), which supports much of the gene therapy research, allows use of stem cells in therapy but doesn't allow their use in reproductive cloning or the combining of human stem cells with animal embryos.[485] Dangers from retroviruses may lurk in attempts to transplant animal organs (e.g., pig hearts) into people.[486–488]

In attempting to improve crop plants, proteins causing allergic reactions in some individuals may be tranferrred.[489] This is a compelling reason that mandatory monitoring and labeling of genetically modified foods is recommended.[490] The danger is emphasized by the highly popularized contamination of foods with "starlink" maize.

Should we rush commercial developments of genetic engineering to increase food supply at a time when there is an excess worldwide? Should we engineer increased herbicide-resistance into plants when drinking water supplies are being contaminated by herbicides already in use? On the other hand, less toxic herbicides effective at low levels can be used on plants engineered to resist them.[465] Will we generate superweeds resistant to herbicides by gene transfer? With so many genes being engineered, are there other dangers in their release into the environment?[465] A **terminator** technology that produces plants with sterile seeds has been rejected because it would force poor farmers to purchase seeds year after year from multinational corporations. Yet, it could eventually have beneficial consequences in preventing cross-fertilization with genetically modified plants.[491] Since plants with the *Bt* gene kill insects, will their widespread use decimate populations of butterflies or of various beneficial insects or adversely affect soil organisms in the **rhizosphere**?[492,493]

The darkest prospect may be the possibility that military organizations will develop new biological warfare weapons using recombinant DNA methods.[494] An international treaty bans such weapons. However, governments often respond with the argument "We must do it because we know that 'they' are doing it." Couldn't such an attitude bring on unparalleled disaster? Could new types of viruses spread throughout the world and literally tear the human genome to bits? Does any kind of imagined danger justify preparation to attack populations with new viruses or new toxins? What about the smallpox virus? Should remaining stocks in the United States and Russia be destroyed?[495,496]

References

1. Horowitz, N. H. (1995) *Protein Sci.* **4**, 1017–1019
2. Atherly, A. G., Girton, J. R., and McDonald, J. F. (1999) *The Science of Genetics*, Saunders, Philadelphia, Pennsylvania
2a. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002) *Molecular Biology of the Cell*, 4th ed., Garland Science, New York
2b. Lander, E. S., and Weinberg, R. A. (2000) *Science* **287**, 1777 – 1782
3. Mirsky, A. E. (1968) *Sci. Am.* **218**(Jun), 78–88
4. Portugal, F. H., and Cohen, J. S. (1977) *A Century of DNA: A History of the Discovery of the Structure and Function of the Genetic Substance*, MIT Press, Cambridge, Massachusetts
5. Adams, R. L. P., Knowler, J. T., and Leader, D. P. (1992) *The Biochemistry of the Nucleic Acids*, 11th ed., Chapman & Hall, London
6. Olby, R. (1975) *The Path to the Double Helix*, Univ. of Washington Press, Seattle, Washington
7. Stent, G. S., and Calendar, R. (1978) *Molecular Genetics*, 2nd ed., Freeman, San Francisco, California
8. Chargaff, E. (1971) *Science* **172**, 637–642
9. Olby, R. (1974) *Nature (London)* **248**, 782–785
10. McCarty, M. (1985) *The Transforming Principle*, Norton, New York
11. Cohen, S. S. (1984) *Trends Biochem. Sci.* **9**, 334–336
12. Chambers, D. A., Reid, K. B. M., and Cohen, R. L. (1994) *FASEB J.* **8**, 1219–1226
13. Hershey, A. D., and Chase, M. (1952) *J. Gen. Physiol.* **36**, 39–56
14. Hayes, W. (1968) *The Genetics of Bacteria and Their Viruses; Studies in Basic Genetics and Molecular Biology*, 2nd ed., Wiley, New York
15. Chargaff, E. (1950) *Experientia* **6**, 201–209
16. Brinton, C. C. (1971) *Crit. Rev. Microbiol* **1**, 105–160
17. Sayre, A. (1975) *Rosalind Franklin and DNA*, Norton, New York
17a. Maddox, B. (2002) *Rosalind Franklin: The Dark Lady of DNA*, HarperCollins, New York
17b. Blow, D. (2002) *Nature (London)* **418**, 725 – 726
18. Watson, J. D., and Crick, F. H. C. (1953) *Nature (London)* **171**, 737–738
19. Crick, F. (1974) *Nature (London)* **248**, 766–769
20. Watson, J. D. (1968) *The Double Helix*, Atheneum, New York
20a. Morange, M. (2000) *A History of Molecular Biology*, Harvard Univ. Press, Cambridge,MA (translated by Matthew Cobb)
20b. Echols, H. (2001) *Operators and Promoters: The Story of Molecular Biology and Its Creators*, Univ. of California Press, Berkeley, California (edited by Carol A. Gross)
21. Caspersson, T. (1941) *Naturwissenschaften* **29**, 33–43
22. Brachet, J. (1942) *Arch. Biol.* **53**, 207–257
23. Brachet, J. (1987) *Trends Biochem. Sci.* **12**, 244–246
24. Prescott, D. M. (1964) *Prog. Nucleic Acid Res. Mol. Biol.* **3**, 33–57
25. Sirlin, J. L. (1972) *Biology of RNA*, Academic Press, New York
26. Hoagland, M. B., Keller, E. B., and Zamecnik, P. (1956) *J. Biol. Chem.* **218**, 345–358
27. Belozersky, A. N., and Spirin, A. S. (1960) in *The Nucleic Acids*, Vol. 3 (Chargaff, E., and Davidson, J. N., eds), pp. 147–185, Academic Press, New York
28. Davern, C. I., and Meselson, M. (1960) *J. Mol. Biol.* **2**, 153–160
29. Volkin, E., and Astrachan, L. (1956) *Virology* **2**, 149–161
30. Volkin, E. (1995) *Trends Biochem. Sci.* **20**, 206–209
31. Rouwenhorst, R. J., Pronk, J. T., and van Dijken, J. P. (1989) *Trends Biochem. Sci.* **14**, 416–418
32. Jacob, F., and Monod, J. (1961) *J. Mol. Biol.* **3**, 318–356
33. Cohen, G. N. (1995) *FASEB J.* **9**, 981–982
34. Brenner, S., Jacob, F., and Meselson, M. (1961) *Nature (London)* **190**, 576–581
35. Crick, F. H. C. (1957) *Biochem. Soc. Symp.* **14**, 25–26
36. Hoagland, M. (1996) *Trends Biochem. Sci.* **21**, 77–80
37. Nirenberg, M. W., and Matthaei, J. H. (1961) *Proc. Natl. Acad. Sci. U.S.A.* **47**, 1588–1602
38. Nirenberg, M., and Leder, P. (1964) *Science* **145**, 1399–1407
39. Matthaei, J. H., Voigt, H. P., Heller, G., Neth, R., Schöch, G., Kübler, H., Amelunxen, F., Sander, G., and Parmeggiani, A. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 25–38
40. Khorana, H. G., Büchi, H., Ghosh, H., Gupta, N., Jacob, T. M., Kössel, H., Morgan, R., Narang, S. A., Ohtsuka, E., and Wells, R. D. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 39–49
41. Edgar, R. S., and Epstein, R. H. (1965) *Sci. Am.* **212**(Feb), 70–78
42. Fraenkel-Conrat, H. (1964) *Sci. Am.* **211**(Oct), 47–54
43. Wu, D. Y., Ugozzoli, L., Pal, B. K., and Wallace, R. B. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2757–2760
44. Lewin, B. (1974) *Gene Expression*, Vol. 2, Wiley, New York (p. 258)
45. Knight, R. D., Freeland, S. J., and Landweber, L. F. (1999) *Trends Biochem. Sci.* **24**, 241–247
46. Seid-Akhavan, M., Winter, W. P., Abramson, R. K., and Rucknagel, D. L. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 882–886
47. Benzer, S. (1962) *Sci. Am.* **206**(Jan), 70–84
48. Hayes, W. (1968) *The Genetics of Bacteria and Their Viruses*, 2nd ed., Wiley, New York (pp. 52–54)
49. Weintraub, S. B., and Frankel, F. R. (1972) *J. Mol. Biol.* **70**, 589–615
50. Ennis, H. L., and Kievitt, K. D. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1468–1472
51. Beadle, G. W., and Tatum, E. L. (1941) *Proc. Natl. Acad. Sci. U.S.A.* **27**, 499–
52. Yanofsky, C. (1967) *Sci. Am.* **216**(May), 80–94
53. Yanofsky, C. (1976) in *Reflections on Biochemistry* (Kornberg, A., ed), pp. 263–269, Pergamon, New York
54. Sarabhai, A. S., Stretton, A. O. W., Brenner, S., and Bolle, A. (1964) *Nature (London)* **210**, 13–17
55. Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961) *Nature (London)* **192**, 1227–1232
55a. Atkins, J. F., and Gesteland, R. F. (2001) *Nature (London)* **414**, 693
56. Watson, J. O. (1987) *Molecular Biology of the Gene*, 4th ed., Benjamin, New York
57. Eiserling, F. A., and Dickson, R. C. (1972) *Ann. Rev. Biochem.* **41**, 467–502
58. Wood, W. B., and Edgar, R. S. (1967) *Sci. Am.* **217**(Jul), 60–74
59. Sober, H. A., ed. (1968) *CRC Handbook of Biochemistry*, 1st ed., Chem. Rubber Publ. Co., Cleveland, Ohio (p. I25)
60. Garen, A. (1968) *Science* **160**, 149–159
61. Littauer, U. Z., and Inouye, H. (1973) *Ann. Rev. Biochem.* **42**, 439–470
62. Hartman, P. E., and Roth, J. R. (1973) *Adv. Genet.* **17**, 1–105
63. Lewin, B. (1974) *Gene Expression*, Vol. 1, Wiley, New York (p. 213)
64. Yanofsky, C., and Crawford, I. P. (1972) in *The Enzymes*, 3rd ed., Vol. 7 (Boyer, P. D., ed), Academic Press, New York (pp. 1–31)
65. Sherratt, D. (1975) *Nature (London)* **254**, 559–560
66. Lederberg, J., and Tatum, E. L. (1946) *Nature (London)* **158**, 558
67. Kornberg, A. (1974) *DNA Synthesis*, Freeman, San Francisco, California (p. 242–248)
68. Harrington, L. C., and Rogerson, A. C. (1990) *J. Bacteriol.* **172**, 7263–7264
69. Eisenbrandt, R., Kalkum, M., Lai, E.-M., Lurz, R., Kado, C. I., and Lanka, E. (1999) *J. Biol. Chem.* **274**, 22548–22555
70. Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R., eds. (1983) *Lambda II*, Cold spring Harbor Lab, Cold Spring Harbor, New York
71. Hershey, A. D., ed. (1971) *The Bacteriophage Lambda*, Cold Spring Harbor Lab, Cold Spring Harbor, New York
72. Echols, H. (1971) *Ann. Rev. Biochem.* **40**, 827–854
73. O'Brien, S. J., ed. (1993) *Genetic Maps*, 6th ed., Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York ((Books 1–6))
74. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277**, 1453–1462
75. Bachmann, B. J., Low, K. B., and Taylor, A. L. (1976) *Bacteriol. Rev.* **40**, 116–167
76. Bachmann, B. J. (1983) *Bacteriol. Rev.* **47**, 180–230
77. Smith, C. L., Econome, J. G., Schutt, A., Klco, S., and Cantor, C. R. (1987) *Science* **236**, 1448–1453
78. Kohara, Y., Akiyama, K., and Isono, K. (1987) *Cell* **50**, 495–508
79. Low, K. B. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), Am. Soc. for Microbiology, Washington, D. C.
80. Willetts, N., and Skurray, R. (1987) in *Escherichia coli and Salmonella typhimurium*, Vol. 2 (Neidhardt, F. C., ed), pp. 1110–1133, Am. Soc. for Microbiology, Washington, D. C.
80a. Bist, P., Sistla, S., Krishnamurthy, V., Acharya, A., Chandrakala, B., and Rao, D. N. (2001) *J. Mol. Biol.* **310**, 93 – 109
81. Davies, G. P., Martin, I., Sturrock, S. S., Cronshaw, A., Murray, N. E., and Dryden, D. T. F. (1999) *J. Mol. Biol.* **290**, 565–579
82. Janscak, P., MacWilliams, M. P., Sandmeier, U., Nagaraja, V., and Bickle, T. A. (1999) *EMBO J.* **18**, 2638–2647
83. Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1995) *Science* **269**, 656–663
84. Bozic, D., Grazulis, S., Siksnys, V., and Huber, R. (1996) *J. Mol. Biol.* **255**, 176–186
85. Luria, S. E. (1970) *Sci. Am.* **222**(Jan), 88–102
86. McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J., and Rosenberg, J. M. (1986) *Science* **234**, 1526–1541
87. Hager, P. W., Reich, N. O., Day, J. P., Coche, T. G., Boyer, H. W., Rosenberg, J. M., and Greene, P. J. (1990) *J. Biol. Chem.* **265**, 21520–21526
88. Frederick, C. A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.-C., Greene, P., Boyer, H. W., and Rosenberg, J. M. (1984) *Nature (London)* **309**, 327–331
89. Frederick, C. A., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., and Rich, A. (1988) *J. Biol. Chem.* **263**, 17872–17879

References

90. Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1994) *Nature (London)* **368**, 660 – 664
91. Kostrewa, D., and Winkler, F. K. (1995) *Biochemistry* **34**, 683 – 696
91a. Horton, N. C., Otey, C., Lusetti, S., Sam, M. D., Kohn, J., Martin, A. M., Ananthnarayan, V., and Perona, J. J. (2002) *Biochemistry* **41**, 10754 – 10763
92. Stanford, N. P., Halford, S. E., and Baldwin, G. S. (1999) *J. Mol. Biol.* **288**, 105 – 116
93. Sam, M. D., and Perona, J. J. (1999) *Biochemistry* **38**, 6576 – 6586
94. Kim, Y., Grable, J. C., Love, R., Greene, P. J., and Rosenberg, J. M. (1990) *Science* **249**, 1307 – 1309
95. Kim, S. C., Podhajska, A. J., and Szybalski, W. (1988) *Science* **240**, 504 – 506
96. Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980) *Am. J. Hum. Genet.* **32**, 314 – 331
97. Cohen, D., Chumakov, I., and Weissenbach, J. (1993) *Nature (London)* **366**, 698 – 701
98. Westmoreland, B. C., Szybalski, W., and Ris, H. (1969) *Science* **163**, 1343 – 1348
99. Glover, D. M., ed. (1987) *DNA Cloning: A Practical Approach*, Vol. III, IRL Press, Oxford
100. Allemand, J.-F., Bensimon, D., Jullien, L., Bensimon, A., and Croquette, V. (1997) *Biophys. J.* **73**, 2064 – 2070
101. Michalet, X., Ekong, R., Fougerousse, F., Rousseaux, S., Schurra, C., Hornigold, N., van Slegtenhorst, M., Wolfe, J., Povey, S., Beckmann, J. S., and Bensimon, A. (1997) *Science* **277**, 1518 – 1523
102. Samad, A. H., Cai, W. W., Hu, X., Irvin, B., Jing, J., Reed, J., Meng, X., Huang, J., Huff, E., Porter, B., Shenkar, A., Anantharaman, T., Mishra, B., Clarke, V., Dimalanta, E., Edington, J., Hiort, C., Rabbah, R., Skiada, J., and Schwartz, D. C. (1995) *Nature (London)* **378**, 516 – 517
103. Lin, J., Qi, R., Aston, C., Jing, J., Anantharaman, T. S., Mishra, B., White, O., Daly, M. J., Minton, K. W., Venter, J. C., and Schwartz, D. C. (1999) *Science* **285**, 1558 – 1561
104. Cox, D. R., Burmeister, M., Price, E. R., Kim, S., and Myers, R. M. (1990) *Science* **250**, 245 – 250
105. Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B., Moreno, R. F., Kerlavage, A. R., McCombie, W. R., and Venter, J. C. (1991) *Science* **252**, 1651 – 1656
106. Mann, M. (1996) *Trends Biochem. Sci.* **21**, 494 – 495
107. Watson, J. D., Gilman, M., Witkowski, J., and Zoller, M., eds. (1992) *Recombinant DNA*, 2nd ed., Scientific American Publ. Co., New York
108. Sambrook, J., and Russell, D. (2000) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Lab. Press, Plainview, New York
109. Berger, S. L., and Kimmel, A. R., eds. (1987) *Guide to Molecular Cloning Techniques*, Academic Press, San Diego
110. Perbal, B. (1988) *A Practical Guide to Molecular Cloning*, Wiley, New York
111. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry, Part II*, Freeman, San Francisco, California (pp. 409 – 433)
112. Greene, J. J., and Rao, V. B., eds. (1998) *Recombinant DNA Principles and Methodologies*, Dekker, New York
113. Miller, J. H. (1992) *A Short Course in Bacterial Genetics*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
114. Wu, R., ed. (1987) *Methods in Enzymology, "Recombinant DNA,"* Vols. 153 – 155, Academic Press, San Diego
115. Pouwels, P. H., Enger-Valk, B. E., and Brammar, W. J., eds. (1985) *Cloning Vectors: A Laboratory Manual*, Elsevier, Amsterdam
116. Adolph, K. W., ed. (1993) *Methods in Molecular Genetics: Gene and Chromosome Analysis*, Vols. 1 and 2, Academic Press, San Diego, California
117. Birren, B., Green, E. D., Klapholz, S., Myers, R. M., and Roskams, J., eds. (1997) *Analyzing DNA: A Laboratory Manual*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
118. Birren, B., Green, E. D., Klapholz, S., Myers, R. M., Riethman, H., and Roskams, J., eds. (1999) *Cloning Systems: A Laboratory Manual*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
119. Lathe, R., Kieny, M. P., Skory, S., and Lecocq, J. P. (1984) *DNA* **3**, 173 – 182
120. Nobbs, T. J., and Halford, S. E. (1995) *J. Mol. Biol.* **252**, 399 – 411
121. Winnacker, E. L. (1987) *From Genes to Clones*, VCH, New York
122. Midgley, C. A., and Murray, N. E. (1985) *EMBO J.* **4**, 2695 – 2703
123. Bollum, F. J. (1981) *Trends Biochem. Sci.* **6**, 41 – 43
123a. Shanks, O. C., Bissonette, L., and Ream, W. (2000) *Science* **289**, 413
124. Bolivar, F., Rodriguez, R. L., Coreene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H., and Falkow, S. (1977) *Gene* **2**, 95 – 113
125. Peden, K. W. C. (1983) *Gene* **22**, 277 – 280
126. Vieira, J., and Messing, J. (1982) *Gene* **19**, 259 – 268
127. Messing, J. (1981) *Nucleic Acids Res.* **9**, 309 – 321
128. Sanger, F., Coulson, A. R., Friedmann, T., Air, G. M., Barrell, B. G., Brown, N. L., Fiddes, J. C., Hutchison, C. A., III, Slocombe, P. M., and Smith, M. (1978) *J. Mol. Biol.* **125**, 225 – 246
129. Blattner, F. R., Williams, B. G., Blechi, A. E., Denniston-Thompson, K., Faber, H. E., Furlong, L.-A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L., and Smithies, O. (1977) *Science* **196**, 161 – 169
130. Short, J. M. (1988) *Nucleic Acids Res.* **15**
131. Collins, J., and Hohn, B. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4242 – 4246
132. Wahl, G. M., Lewis, K. A., Ruiz, J. C., Rothenberg, B., Zhao, J., and Evans, G. A. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2160 – 2164
133. Fernandez, J. M., and Hoeffler, J. P., eds. (1998) *Gene Expression Systems*, Academic Press, San Diego, California
134. Moss, B., Elroy-Stein, O., Mizukami, T., Alexander, W. A., and Fuerst, T. R. (1990) *Nature (London)* **348**, 91 – 92
135. O'Reilly, D. R., Miller, L. K., and Luckow, V. A. (1992) *Baculovirus Expression Vectors*, Freeman, New York
136. Kriegler, M. (1990) *Gene Transfer and Expression*, Stockton Press, New York
137. Burke, D. T., Carle, G. F., and Olson, M. V. (1987) *Science* **236**, 806 – 812
138. Murray, A. W., and Szostak, J. W. (1987) *Sci. Am.* **257**(Nov), 62 – 68
139. Forget, B. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7909 – 7911
140. Anderson, C. (1993) *Science* **259**, 1684 – 1687
141. Sternberg, N. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 103 – 107
142. Pierie, J. C., and Sternberg, N. L. (1992) *Methods Enzymol.* **216**, 549 –
143. Heller, R., Brown, K. E., Burgtorf, C., and Brown, W. R. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7125 – 7130
144. Chilton, M.-D. (1983) *Sci. Am.* **248**(Jun), 51 – 59
145. Weiler, E. W., and Schröder, J. (1987) *Trends Biochem. Sci.* **12**, 271 – 275
146. Zambryski, P. C. (1992) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43**, 465 – 490
146a. Wood, D. W., and 50 other authors. (2001) *Science* **294**, 2317 – 2323
146b. Goodner, B., and 30 other authors. (2001) *Science* **294**, 2323 – 2328
146c. Moriguchi, K., Maeda, Y., Satou, M., Hardayani, N. S. N., Kataoka, M., Tanaka, N., and Yoshida, K. (2001) *J. Mol. Biol.* **307**, 771 – 784
147. Slightom, J. L., Durand-Tardif, M., Jouanin, L., and Tepfer, D. (1986) *J. Biol. Chem.* **261**, 108 – 121
148. Hess, K. M., Dudley, M. W., Lynn, D. G., Joerger, R. D., and Binns, A. N. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7854 – 7858
149. Stachel, S. E., Messens, E., Van Montagu, M., and Zambryski, P. (1985) *Nature (London)* **318**, 624 – 629
150. Bates, H. A., Kaushal, A., Deng, P.-N., and Sciaky, D. (1984) *Biochemistry* **23**, 3287 – 3290
151. Tinland, B., Schoumacher, F., Gloeckler, V., Bravo-Angel, A. M., and Hohn, B. (1995) *EMBO J.* **14**, 3585 – 3595
152. Scheiffele, P., Pansegrau, W., and Lanka, E. (1995) *J. Biol. Chem.* **270**, 1269 – 1276
153. Bundock, P., den Dulk-Ras, A., Beijersbergen, A., and Hooykaas, P. J. J. (1995) *EMBO J.* **14**, 3206 – 3214
154. Fullner, K. J., Lara, J. C., and Nester, E. W. (1996) *Science* **273**, 1107 – 1109
154a. Ward, D. V., Draper, O., Zupan, J. R., and Zambryski, P. C. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11493 – 11500
155. Crystal, R. G. (1995) *Science* **270**, 404 – 410
156. Westphal, H. (1989) *FASEB J.* **3**, 117 – 120
157. Cotton, M., Wagner, E., Zatloukal, K., Phillips, S., Curiel, D. T., and Birnstiel, M. L. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6094 – 6098
158. Kojima, H., Ohishi, N., and Yagi, K. (1998) *Biochem. Biophys. Res. Commun.* **246**, 868 – 872
159. Yu, S.-F., von Rüden, T., Kantoff, P. W., Garber, C., Seiberg, M., Rüther, U., Anderson, W. F., Wagner, E. F., and Gilboa, E. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3194 – 3198
160. Chakraborty, A. K., Zink, M. A., Boman, B. M., and Hodgson, C. P. (1993) *FASEB J.* **7**, 971 – 977
160a. Ferber, D. (2001) *Science* **294**, 1638 – 1642
161. Daly, T. M., Vogler, C., Levy, B., Haskins, M. E., and Sands, M. S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2296 – 2300
161a. Yan, Z., Zhang, Y., Duan, D., and Engelhardt, J. F. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6716 – 6721
161b. Xie, Q., Bu, W., Bhatia, S., Hare, J., Somasundaram, T., Azzi, A., and Chapman, M. S. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10405 – 10410
162. Amado, R. G., and Chen, I. S. Y. (1999) *Science* **285**, 674 – 676
162a. Lois, C., Hong, E. J., Pease, S., Brown, E. J., and Baltimore, D. (2002) *Science* **295**, 868 – 872
163. Jasny, B. R. (1987) *Science* **238**, 1653
164. Hofmann, C., Sandig, V., Jennings, G., Rudolph, M., Schlag, P., and Strauss, M. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10099 – 10103
165. Yin, H. L., and Stull, J. T. (1999) *J. Biol. Chem.* **274**, 32529 – 32530
166. Poul, M.-A., and Marks, J. D. (1999) *J. Mol. Biol.* **288**, 203 – 211
166a. Kaminski, J. M., Huber, M. R., Summers, J. B., and Ward, M. B. (2002) *FASEB J.* **16**, 1242 – 1247
166b. Saldanha, R., Chen, B., Wank, H., Matsuura, M., Edwards, J., and Lambowitz, A. M. (1999) *Biochemistry* **38**, 9069 – 9083
166c. Guo, H., Karberg, M., Long, M., Jones, J. P., III, Sullenger, B., and Lambowitz, A. M. (2000) *Science* **289**, 452 – 457
167. Cocking, E. C., and Davey, M. R. (1987) *Science* **236**, 1259 – 1262

References

168. Smith, H. O., Danner, D. B., and Deich, R. A. (1981) *Ann. Rev. Biochem.* **50**, 41–68
169. Wahl, G. M., de Saint Vincent, B. R., and DeRose, M. L. (1984) *Nature (London)* **307**, 516–520
170. Felgner, P. L., and Ringold, G. M. (1989) *Nature (London)* **337**, 387–388
171. Miller, J. F., Dower, W. J., and Tompkins, L. S. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 856–860
172. Solioz, M., and Bienz, D. (1990) *Trends Biochem. Sci.* **15**, 175–177
173. Hengen, P. N. (1995) *Trends Biochem. Sci.* **20**, 248–249
174. Collombet, J.-M., Wheeler, V. C., Vogel, F., and Coutelle, C. (1997) *J. Biol. Chem.* **272**, 5342–5347
175. Chang, D. C., Hunt, J. R., Zheng, Q., and Gao, P. Q. (1992) in *Guide to Electroporation and Electrofusion* (Chang, D. C., ed), pp. 303–326, Academic Press, San Diego, California
176. Klobutcher, L. A., and Ruddle, F. H. (1981) *Ann. Rev. Biochem.* **50**, 533–554
177. Klein, T. M., Wolf, E. D., Wu, R., and Sanford, J. C. (1987) *Nature (London)* **327**, 70–73
178. Butow, R. A., and Fox, T. D. (1990) *Trends Biochem. Sci.* **15**, 465–468
179. Kohli, A., Leech, M., Vain, P., Laurie, D. A., and Christou, P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7203–7208
180. Hohn, B., and Murray, K. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3259–3263
181. Lutz, C. T., Hollifield, W. C., Seed, B., Davie, J. M., and Huang, H. V. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4379–4383
182. Wyman, A. R., Wolfe, L. B., and Botstein, D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2880–2884
183. Okayama, H., and Berg, P. (1982) *Mol. Cell. Biochem.* **2**, 161–170
184. Gubler, U., and Hoffman, B. J. (1983) *Gene* **25**, 263–269
185. Mason, W. T., ed. (1999) *Fluorescent and Luminescent Probes*, 2nd ed., Academic Press, San Diego, California
186. Keller, G. H., and Manak, M. M. (1989) *DNA Probes*, Stockton Press, New York
186a. Santi, E., Capone, S. Mennuni, C., Lahm, A., Tramontano, A, Luzzago, A., and Nicosia, A. (2000) *J. Mol. Biol.* **296**, 497 – 508
187. Little, P. F. R., and Cross, S. H. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3159–3163
188. Richards, J. E., Gilliam, T. C., Cole, J. L., Drumm, M. L., Wasmuth, J. J., Gusella, J. F., and Collins, F. S. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6437–6441
189. Housman, D. (1995) *N. Engl. J. Med.* **332**, 318–320
190. Viegas-Pequignot, E., Dutrilleax, B., Magdelenat, H., and Coppey-Moisam, M. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 582–586
191. Abbott, C., and Povey, S. (1995) *Somatic Cell Hybrids*, IRL Press, Oxford
192. Williamson, B. (1981) *Nature (London)* **293**, 10–11
193. White, R., and Lalouel, J.-M. (1988) *Sci. Am.* **258**(Feb), 40–48
194. Gusella, J. F. (1986) *Ann. Rev. Biochem.* **55**, 831–854
195. Davies, K. E., ed. (1986) *Human Genetic Diseases*, IRL Press, Washington, D.C.
196. Martin, J. B. (1987) *Science* **238**, 765–772
197. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) *Science* **230**, 1350–1354
198. Gusella, J. F., Tanzi, R. E., Bader, P. I., Phelan, M. C., Stevenson, R., Hayden, M. R., Hofman, K. J., Faryniarz, A. G., and Gibbons, K. (1985) *Nature (London)* **318**, 75–78
199. Gilliam, T. C., Bucan, M., MacDonald, M. E., Zimmer, M., Haines, J. L., Cheng, S. V., Pohl, T. M., Meyers, R. H., Whaley, W. L., Allitto, B. A., Faryniarz, A., Wasmuth, J. J., Frischauf, A.-M., Conneally, P. M., Lehrach, H., and Gusella, J. F. (1987) *Science* **238**, 950–952
200. Royer-Pokora, B., Kunkel, L. M., Monaco, A. P., Goff, S. C., Newburger, P. E., Baehner, R. L., Cole, F. S., Curnutte, J. T., and Orkin, S. H. (1986) *Nature (London)* **322**, 32–38
201. Reich, N. O., and Mashboon, N. (1993) *J. Biol. Chem.* **268**, 9191–9193
202. Lee, W.-H., Shew, J.-Y., Hong, F. D., Sery, T. W., Donoso, L. A., Young, L.-J., Bookstein, R., and Lee, E. Y.-H. P. (1987) *Nature (London)* **329**, 642–645
203. Thomson, G., and Esposito, M. S. (1999) *Trends Biochem. Sci.* **24**, M17–M20
204. Little, P. (1986) *Nature (London)* **321**, 558–559
205. Zaccolo, M., Williams, D. M., Brown, D. M., and Gherardi, E. (1996) *J. Mol. Biol.* **255**, 589–603
206. Myers, R. M., Lerman, L. S., and Maniatis, T. (1985) *Science* **229**, 242–247
207. Botstein, D., and Shortle, D. (1985) *Science* **229**, 1193–1201
208. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367–382
209. Zoller, M. J., and Smith, M. (1983) *Methods Enzymol.* **100**, 468–500
210. Higuchi, R., Krummel, B., and Saiki, R. (1988) *Nucleic Acids Res.* **16**, 7351–7367
211. Jones, D. H., and Howard, B. H. (1990) *BioTechniques* **8**, 178–183
212. Jones, D. H., Sakamoto, K., Vorce, R. L., and Howard, B. H. (1990) *Nature (London)* **344**, 793–794
213. Wells, J. A., and Powers, D. B. (1986) *J. Biol. Chem.* **261**, 6564–6570
214. Shuldiner, A. R. (1996) *N. Engl. J. Med.* **334**, 653–655
215. Jasin, M., Moynahan, M. E., and Richardson, C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8804–8808
216. Capecchi, M. R. (1994) *Sci. Am.* **270**(Mar), 52–59
217. Travis, J. (1992) *Science* **256**, 1392–1394
218. Barinaga, M. (1994) *Science* **265**, 26–28
219. Weissmann, C., and Aguzzi, A. (1999) *Science* **286**, 914–915
219a. Smith, J. A., and Martin, L. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1263–1267
220. Nasmyth, K. (2002) *Science* **297**, 559–565
220a. Hauf, S., Waizenegger, I. C., and Peters, J.-M. (2001) *Science* **293**, 1320–1323
220b. Campbell, J. L., and Cohen-Fix, O. (2002) *Trends Biochem. Sci.* **27**, 492–495
220c. Seemann, J., Pypaert, M., Taguchi, T., Malsam, J., and Warren, G. (2002) *Science* **295**, 848–851
221. Mazia, D. (1961) *Sci. Am.* **205**(Sep), 101–120
222. Earnshaw, W. C., and Mackay, A. M. (1994) *FASEB J.* **8**, 947–956
223. Russell, P. (1998) *Trends Biochem. Sci.* **23**, 399–402
224. Piwnica-Worms, H. (1999) *Nature (London)* **401**, 535–537
225. Nasmyth, K. (1999) *Trends Biochem. Sci.* **24**, 98–104
225a. Cortez, D., and Elledge, S. J. (2000) *Nature (London)* **406**, 354–356
226. Warren, G. (1985) *Trends Biochem. Sci.* **10**, 439–443
226a. Lippincott-Schwartz, J. (2002) *Nature (London)* **416**, 31–32
227. Morgan, D. O. (1995) *Nature (London)* **374**, 131–133
227a. Kong, M., Barnes, E. A., Ollendorff, V., and Donoghue, D. J. (2000) *EMBO J.* **19**, 1378–1388
227b. Healy, J. M. S., Menges, M., Doonan, J. H., and Murray, J. A. H. (2001) *J. Biol. Chem.* **276**, 7041–7047
228. Pavletich, N. P. (1999) *J. Mol. Biol.* **287**, 821–828
228a. Nash, P., Tang, X., Orlicky, S., Chen, Q., Gertler, F. B., Mendenhall, M. D., Sicheri, F., Pawson, T., and Tyers, M. (2001) *Nature (London)* **414**, 514–521
228b. Bartek, J., and Lukas, J. (2001) *Science* **294**, 66–67
228c. Schwab, M., and Tyers, M. (2001) *Nature (London)* **413**, 268–269
228d. Zhou, B.-B. S., and Elledge, S. J. (2000) *Nature (London)* **408**, 433–439
229. Fotedar, A., Cannella, D., Fitzgerald, P., Rousselle, T., Gupta, S., Dorée, M., and Fotedar, R. (1996) *J. Biol. Chem.* **271**, 31627–31637
230. De Bondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O., and Kim, S.-H. (1993) *Nature (London)* **363**, 595–602
231. Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massagué, J., and Pavletich, N. P. (1995) *Nature (London)* **376**, 313–320
232. Shiffman, D., Brooks, E. E., Brooks, A. R., Chan, S. C., and Milner, P. G. (1996) *J. Biol. Chem.* **271**, 12199–12204
233. Nurse, P. (1990) *Nature (London)* **344**, 503–507
234. Pines, J. (1993) *Trends Biochem. Sci.* **18**, 195–197
235. Solomon, M. J. (1994) *Trends Biochem. Sci.* **19**, 496–500
236. Lew, J., and Wang, J. H. (1995) *Trends Biochem. Sci.* **20**, 33–37
236a. Crawford, D. F., and Piwnica-Worms, H. (2001) *J. Biol. Chem.* **276**, 37166–37177
236b. Karsenti, E., and Vernos, I. (2001) *Science* **294**, 543–547
236c. Kastan, M. B. (2001) *Nature (London)* **410**, 766–767
236d. Davenport, R. J. (2001) *Science* **292**, 2415–2417
237. Morris, M. C., and Divita, G. (1999) *J. Mol. Biol.* **286**, 475–487
237a. Jacobs, H. W., Keidel, E., and Lehner, C. F. (2001) *EMBO J.* **20**, 2376–2386
238. Lopez-Girona, A., Furnari, B., Mondesert, O., and Russell, P. (1999) *Nature (London)* **397**, 172–175
238a. Yang, Q., Manicone, A., Coursen, J. D., Linke, S. P., Nagashima, M., Forgues, M., and Wang, X. W. (2000) *J. Biol. Chem.* **275**, 36892–36898
239. Reynolds, R. A., Yem, A. W., Wolfe, C. L., Deibel, M. R., Jr., Chidester, C. G., and Watenpaugh, K. D. (1999) *J. Mol. Biol.* **293**, 559–568
240. Hagting, A., Karlsson, C., Clute, P., Jackman, M., and Pines, J. (1998) *EMBO J.* **17**, 4127–4138
240a. Qian, Y.-W., Erikson, E., and Maller, J. L. (1998) *Science* **282**, 1701–1704
240b. Körner, K., Jérôme, V., Schmidt, T., and Müller, R. (2001) *J. Biol. Chem.* **276**, 9662–9669
241. Pines, J. (1999) *Nature (London)* **397**, 104 – 105
242. Sanchez, Y., Bachant, J., Wang, H., Hu, F., Liu, D., Tetzlaff, M., and Elledge, S. J. (1999) *Science* **286**, 1166–171
243. Kellogg, D. R., Moritz, M., and Alberts, B. M. (1994) *Ann. Rev. Biochem.* **63**, 639–674
244. Geier, B. M., Wiech, H., and Schiebel, E. (1996) *J. Biol. Chem.* **271**, 28366–28374
244a. Murray, A. W. (2001) *Science* **291**, 1499–1502
244b. Hinchcliffe, E. H., Miller, F. J., Cham, M., Khodjakov, A., and Sluder, G. (2001) *Science* **291**, 1547–1550
244c. Middendorp, S., Paoletti, A., Schiebel, E., and Bornens, M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9141–9146
244d. do Carmo Avides, M., and Glover, D. M. (1999) *Science* **283**, 1733–1735

References

244e. Flory, M. R., Moser, M. J., Monnat, R. J., Jr., and Davis, T. N. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5919–5923
245. Nicklas, R. B. (1997) *Science* **275**, 632–637
246. Heald, R., Tournebize, R., Blank, T., Sandaltzopoulos, R., Becker, P., Hyman, A., and Karsenti, E. (1996) *Nature (London)* **382**, 420–425
246a. Fraschini, R., Beretta, A., Sironi, L., Musacchio, A., Lucchini, G., and Piatti, S. (2001) *EMBO J.* **20**, 6648–6659
247. Hyams, J. (1996) *Nature (London)* **382**, 397–398
247a. Bloom, K. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4757–4759
248. Dasso, M. (1993) *Trends Biochem. Sci.* **18**, 96–101
249. Ohba, T., Nakamura, M., Nishitani, H., and Nishimoto, T. (1999) *Science* **284**, 1356–1358
250. Wilde, A., and Zheng, Y. (1999) *Science* **284**, 1359–1362
250a. Gruneberg, U., Campbell, K., Simpson, C., Grindlay, J., and Schiebel, E. (2000) *EMBO J.* **19**, 6475–6488
250b. Martin-Lluesma, S., Stucke, V. M., and Nigg, E. A. (2002) *Science* **297**, 2267–2270
251. Barton, N. R., and Goldstein, L. S. B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1735–1742
252. Lockhart, A., and Cross, R. A. (1996) *Biochemistry* **35**, 2365–2373
253. Pennisi, E. (1998) *Science* **279**, 477–478
254. Elledge, S. J. (1998) *Science* **279**, 999–1000
255. Hwang, L. H., Lau, L. F., Smith, D. L., Mistrot, C. A., Hardwick, K. G., Hwang, E. S., Amon, A., and Murray, A. W. (1998) *Science* **279**, 1041–1044
256. Yamano, H., Gannon, J., and Hunt, T. (1996) *EMBO J.* **15**, 5268–5279
257. Bachant, J. B., and Elledge, S. J. (1999) *Nature (London)* **398**, 757–758
258. Funabiki, H., Kumada, K., and Yanagida, M. (1996) *EMBO J.* **15**, 6617–6628
259. Page, A. M., and Hieter, P. (1999) *Ann. Rev. Biochem.* **68**, 583–609
259a. Prinz, S., and Amon, A. (1999) *Nature (London)* **402**, 133–135
259b. Nakaseko, Y., and Yanagida, M. (2001) *Nature (London)* **412**, 291–292
259c. Morgan, D. O., and Roberts, J. M. (2002) *Nature (London)* **418**, 495–496
260. Grossberger, R., Gieffers, C., Zachariae, W., Podtelejnikov, A. V., Schleiffer, A., Nasmyth, K., Mann, M., and Peters, J.-M. (1999) *J. Biol. Chem.* **274**, 14500–14507
261. Jiang, F., and Basavappa, R. (1999) *Biochemistry* **38**, 6471–6478
261a. Piel, M., Nordberg, J., Euteneuer, U., and Bornens, M. (2001) *Science* **291**, 1550–1553
262. Wisdom, R., Johnson, R. S., and Moore, C. (1999) *EMBO J.* **18**, 188–197
263. Takenaka, K., Moriguchi, T., and Nishida, E. (1998) *Science* **280**, 599–602
263a. Wilkinson, M. G., and Millar, J. B. A. (2000) *FASEB J.* **14**, 2147–2157
263b. Graves, L. M. (2000) *Nature (London)* **403**, 328–332
263c. Whitmarsh, A. J., and Davis, R. J. (2000) *Nature (London)* **403**, 255–256
264. Gottesfeld, J. M., and Forbes, D. J. (1997) *Trends Biochem. Sci.* **22**, 197–202
264a. Goldstein, L. S. B. (2001) *Science* **291**, 2102–2103
265. Martínez-Botas, J., Suárez, Y., Ferruelo, A. J., Gómez-Coronado, D., and Lasunción, M. A. (1999) *FASEB J.* **13**, 1359–1370
266. Bourne, Y., Arvai, A. S., Bernstein, S. L., Watson, M. H., Reed, S. I., Endicott, J. E., Noble, M. E., Johnson, L. N., and Tainer, J. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10232–10236
267. Parge, H. E., Arvai, A. S., Murtari, D. J., Reed, S. I., and Tainer, J. A. (1993) *Science* **262**, 387–395
267a. Rubin, G. M., and Lewis, E. B. (2000) *Science* **287**, 2216–2218
268. Yamamoto, M. (1996) *Trends Biochem. Sci.* **21**, 18–22
269. de Lange, T. (1998) *Nature (London)* **392**, 753–754
270. Conrad, M. N., Dominguez, A. M., and Dresser, M. E. (1997) *Science* **276**, 1252–1255
271. McKim, K. S., Green-Marroquin, B. L., Sekelsky, J. J., Chin, G., Steinberg, C., Khodosh, R., and Hawley, R. S. (1998) *Science* **279**, 876–878
271a. Lindgren, A., Bungard, D., Pierce, M., Xie, J., Vershon, A., and Winter, E. (2000) *EMBO J.* **19**, 6489–6497
272. Orr-Weaver, T. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10443–10449
273. McKim, K. S., and Hawley, R. S. (1995) *Science* **270**, 1595–1601
274. Picard, A., Galas, S., Peaucellier, G., and Dorée, M. (1996) *EMBO J.* **15**, 3590–3598
275. Lenormand, J.-L., Dellinger, R. W., Knudsen, K. E., Subramani, S., and Donoghue, D. J. (1999) *EMBO J.* **18**, 1869–1877
276. Sette, C., Barchi, M., Bianchini, A., Conti, M., Rossi, P., and Geremia, R. (1999) *J. Biol. Chem.* **274**, 33571–33579
276a. Xu, R., Wilson, W. K., and Matsuda, S. P. T. (2002) *J. Am. Chem. Soc.* **124**, 918–919
277. Watanabe, Y., and Nurse, P. (1999) *Nature (London)* **400**, 461–464
278. Roeder, G. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10450–10456
278a. Albini, S. M., and Jones, G. H. (1987) *Chromosoma* **95**, 324–328
279. Kleckner, N. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8167–8174
279a. Nabeshima, K., Kakihara, Y., Hiraoka, Y., and Nojima, H. (2001) *EMBO J.* **20**, 3871–3881
279b. Yuan, L., Liu, J.-G., Hoja, M.-R., Wilbertz, J., Nordqvist, K., and Höög, C. (2002) *Science* **296**, 1115–1118
279c. Gerton, J. L., DeRisi, J., Shroff, R., Lichten, M., Brown, P. O., and Petes, T. D. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11383–11390
279d. Stack, S. M. (1991) *Genome* **34**, 900–908
280. Boy de la Tour, E., and Laemmli, U. K. (1988) *Cell* **55**, 937–944
281. Gallardo, M. H., Bickham, J. W., Honeycutt, R. L., Ojeda, R. A., and Köhler, N. (1999) *Nature (London)* **401**, 341
281a. Shonn, M. A., McCarroll, R., and Murray, A. W. (2000) *Science* **289**, 300–303
281b. Sluder, G., and McCollum, D. (2000) *Science* **289**, 254–255
282. Pirrotta, V. (1984) *Trends Biochem. Sci.* **9**, 220–221
283. Burtis, K. C., and Hawley, R. S. (1999) *Nature (London)* **401**, 125–127
284. Kao, F.-t, and Yu, J.-w. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1844–1848
285. Sager, R. (1965) *Sci. Am.* **212**(Jan), 71–79
286. Goodenough, U. W., and Levine, R. P. (1970) *Sci. Am.* **223**(Nov), 22–29
287. Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsodi, J., Boll, W., Cantell, K., and Weissmann, C. (1980) *Nature (London)* **284**, 316–320
288. Olson, M. V. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4338–4344
289. Collins, F. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10821–10823
290. Guyer, M. S., and Collins, F. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10841–10848
290a. International Human Genome Sequencing Consortium. (2001) *Nature (London)* **409**, 860–921
290b. Pennisi, E. (2001) *Science* **291**, 1177–1180
290c. Venter, J. C., and many others. (2001) *Science* **291**, 1304–1351
290d. Olivier, M., and 54 other authors. (2001) *Science* **291**, 1298–1302
290e. Shoemaker, D. D., and 34 other authors. (2001) *Nature (London)* **409**, 922–927
290f. Caron, H., and 12 other authors. (2001) *Science* **291**, 1289–1292
290g. Deloukas, P., and 126 other authors. (2001) *Nature (London)* **414**, 865–871
290h. Claverie, J.-M. (2001) *Science* **291**, 1255–1257
290i. Pennisi, E. (2002) *Science* **296**, 1600–1601
290j. Kim, S. K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J. M., Eizinger, A., Wylie, B. N., and Davidson, G. S. (2001) *Science* **293**, 2087–2092
291. Collins, F. S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., and Walters, L. (1998) *Science* **282**, 682–689
291a. Malakoff, D. (2000) *Science* **288**, 248–253
291b. The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. (2001) *Nature (London)* **409**, 685–690
291c. The International Mouse Mutagenesis Consortium. (2001) *Science* **291**, 1251–1255
291d. Gregory, S. G., and 85 other authors (2002) *Nature (London)* **418**, 743–750
291e. Mural, R. J., and 176 other authors (2002) *Science* **296**, 1661–1671
292. Roach, J. C., Siegel, A. F., van den Engh, G., Trask, B., and Hood, L. (1999) *Nature (London)* **401**, 843–845
292a. Marshall, E. (2001) *Science* **291**, 1872
292b. Vogel, G. (2000) *Science* **290**, 1671
292c. Aparicio, S., and 40 other authors (2002) *Science* **297**, 1301–1310
292d. Gardner, M. J., and 44 other authors (2002) *Nature (London)* **419**, 498–511
292e. Doolittle, R. F. (2002) *Nature (London)* **419**, 493–494
292f. Carlton, J. M., and 43 other authors (2002) *Nature (London)* **419**, 512–519
292g. Holt, R. A., and 121 other authors (2002) *Science* **298**, 129–149
292h. Kaufman, T. C., Severson, D. W., and Robinson, G. E. (2002) *Science* **298**, 97–115
293. White, R., and Caskey, C. T. (1988) *Science* **240**, 1483–1488
294. McKusick, V. A., and Amberger, J. S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 119–125, McGraw-Hill, New York
295. Sawyer, J. R., and Hozier, J. C. (1986) *Science* **232**, 1632–1635
296. Tunnacliffe, A., Benham, F., and Goodfellow, P. (1984) *Trends Biochem. Sci.* **9**, 5–7
297. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) *Molecular Biology of the Cell*, 3rd ed., Garland, New York
298. Franke, U. (1981) *Cytogenet. Cell Genet.* **31**, 24–32
299. Gray, J. W., and Langlois, R. G. (1986) *Ann. Rev. Biophys. Biophys. Chem.* **15**, 195–235
300. Chumakov, I., and plus 35 other authors (1992) *Nature (London)* **359**, 380–387
301. Shizuya, H., Birren, B., Kim, U.-J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8794–8797
302. Venter, J. C., Adams, M. D., Sutton, G. G., Kerlavage, A. R., Smith, H. O., and Hunkapiller, M. (1998) *Science* **280**, 1540–1542
303. Marx, J. (1995) *Science* **270**, 1919–1920

References

304. Dib, C., Fauré, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., and Weissenbach, J. (1996) *Nature (London)* **380**, 152–154, iii–v + supplement
305. Rogers, J. (1999) *Science* **286**, 429–432
305a. Liu, S., Ren, H., Gao, Q., Roach, D. J., Loder, R. T., Jr., Armstrong, T. M., Mao Q, Blaga, I., Barker, D. L., and Jovanovich, S. B. (2000) *Proc. Natl. Acad. Sci., U.S.A.* **97**, 5369 – 5374
305b. Paegel, B. M., Emrich, C. A., Wedemayer, G. J., Scherer, J. R., and Mathies, R. A. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 574 – 579
306. Marshall, E. (1999) *Science* **284**, 1906–1909
306a. Green, P. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4143 – 4144
307. Donis-Keller, H., and 32 other authors (1987) *Cell* **51**, 319–337
308. Wang, D. G., and 26 other authors (1998) *Science* **280**, 1077–1082
308a. The International SNP Map Working Group. (2001) *Nature (London)* **409**, 928–933
308b. Patil, N., and 21 other authors (2001) *Science* **294**, 1719–1723
309. Hudson, T. J., and 50 other authors (1995) *Science* **270**, 1945–1954
310. Schuler, G. D., and 103 other authors (1996) *Science* **274**, 540–546
311. Wang, S. M., and Rowley, J. D. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11909–11914
312. Dunham, I., Shimizu, N., Roe, B. A., Chossoe, S., and 117 other authors (1999) *Nature (London)* **402**, 489–495
313. Butler, D. (1999) *Nature (London)* **402**, 447–448
313a. The International Human Genome Mapping Consortium. (2001) *Nature (London)* **409**, 934–941
314. Pandey, A., and Lewitter, F. (1999) *Trends Biochem. Sci.* **24**, 276–280
315. Pennisi, E. (1999) *Science* **286**, 447–450
316. Baxevanis, A. D., and Ouellette, B. F. F., eds. (1998) *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, Wiley, New York
317. Bishop, M. J., ed. (1999) *Genetics Databases*, Academic Press, San Diego, California
318. Durbin, R., Eddy, S., Krogh, A., and Mitchinson, G., eds. (1998) *Biological Sequence Analysis*, Cambridge Univ. Press., Cambridge
319. Bishop, M. J., ed. (1998) *Human Genome Computing*, 2nd ed., Academic Press, San Diego, California
320. Calvet, J. P. (1998) *Science* **282**, 1057–1058
321. Kereil, D. P., and Etzold, T. (1999) *Trends Biochem. Sci.* **24**, 155–157
321a. Tamames, J., and Tramontano, A. (2000) *Trends Biochem. Sci.* **25**, 402–403
321b. Jones, D. T., and Swindells, M. B. (2002) *Trends Biochem. Sci.* **27**, 161–164
321c. Zafar, N., Mazumder, R., and Seto, D. (2001) *Trends Biochem. Sci.* **26**, 514–516
322. Boguski, M. S. (1995) *N. Engl. J. Med.* **333**, 645–647
323. Borodovsky, M., Koonin, E. V., and Rudd, K. E. (1994) *Trends Biochem. Sci.* **19**, 309–313
324. Brown, N. P., Whittaker, A. J., Newell, W. R., Rawlings, C. J., and Beck, S. (1995) *J. Mol. Biol.* **249**, 342–359
325. Pennisi, E. (1998) *Science* **280**, 1692–1693
325a. Matthiessen, M. W. (2002) *Trends Biochem. Sci.* **27**, 586 – 588
326. Little, P. (1999) *Nature (London)* **402**, 467–468
327. O'Brien, S. J., Menotti-Raymond, M., Murphy, W. J., Nash, W. G., Wienberg, J., Stanyon, R., Copeland, N. G., Jenkins, N. A., Womack, J. E., and Graves, J. A. M. (1999) *Science* **286**, 458–481
327a. Gura, T. (2001) *Science* **293**, 593–595
327b. Buetow, K. H., Edmonson, M., MacDonald, R., Clifford, R., Yip, P., Kelley, J., Little, D. P., Strausberg, R., Koester, H., Cantor, C. R., and Braun, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 581–584
327c. Couzin, J. (2002) *Science* **296**, 1391,1393
327d. Adam, D. (2001) *Nature (London)* **412**, 105
327e. Yan, H., Kinzler, K. W., and Vogelstein, B. (2000) *Science* **289**, 1890–1892
327f. Wright, A. F., and Van Heyningen, V. (2001) *Nature (London)* **414**, 705–706
327g. Marshall, E. (2001) *Science* **294**, 2272–2274
327h. Trager, R. S. (2002) *Science* **298**, 947
328. Pääbo, S. (1999) *Trends Biochem. Sci.* **24**, M13–M16
328a. Templeton, A. R. (2002) *Nature (London)* **416**, 45–51
329. Olsen, G. J., and Woese, C. R. (1993) *FASEB J.* **7**, 113–123
330. Shenk, M. A., and Steele, R. E. (1993) *Trends Biochem. Sci.* **18**, 459–463
331. Harris, E. E., and Hey, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3320–3324
331a. Adcock, G. J., Dennis, E. S., Easteal, S., Huttley, G. A., Jermiin, L. S., Peacock, W. J., and Thorne, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 537–542
331b. Relethford, J. H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 390–391
331c. Jones, M. (2001) *The Molecule Hunt: Archaeology and the Search for Ancient DNA*, Allen Lane,
331d. Fish, S. A., Shepherd, T. J., McGenity, T. J., and Grant, W. D. (2002) *Nature (London)* **417**, 432–436
332. Kahn, P., and Gibbons, A. (1997) *Science* **277**, 176–178
332a. Gibbons, A. (2001) *Science* **292**, 627–629
332b. Fujiyama, A., Watanabe, H., Toyoda, A., Taylor, T. D., Itoh, T., Tsai, S.-F., and 11 other authors. (2002) *Science* **295**, 131–134
332c. Enard, W., Khaitovich, P., Klose, J., Zöllner, S., and 9 other authors. (2002) *Science* **296**, 340–343
332d. Cyranoski, D. (2002) *Nature (London)* **418**, 910–912
332e. Stone, A. C., Griffiths, R. C., Zegura, S. L., and Hammer, M. F. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 43–48
332f. Britten, R. J. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13633–13635
333. Gibbons, A. (1998) *Science* **281**, 1432–1434
334. Kennedy, E. P. (1992) *Ann. Rev. Biochem.* **61**, 1–28
335. Barbujani, G., Magagni, A., Minch, E., and Cavalli-Sforza, L. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4516–4519
336. Molnar, S. (1998) *Human Variation, Races, Types and Ethnic Groups*, 4th ed., Prentice Hall, Upper Saddle River, New Jersey (pp. 230–247)
336a. Jablonski, N. G., and Chaplin, G. (2002) *Sci. Am.* **287**(Oct), 74–81
337. Rannala, B., and Mountain, J. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9197–9201
338. Sugimoto, C., Kitamura, T., Guo, J., Al-Ahdal, M. N., Shchelkunov, S. N., Otova, B., Ondrejka, P., Chollet, J.-Y., El-Safi, S., Ettayebi, M., Grésenguet, G., Kocagöz, T., Chaiyarasamee, S., Thant, K. Z., Thein, S., Moe, K., Kobayashi, N., Taguchi, F., and Yogo, Y. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9191–9196
338a. Pääbo, S. (2001) *Science* **291**, 1219–1220
338b. Pennisi, E. (2001) *Science* **291**, 1733–1734
338c. Thomson, R., Pritchard, J. K., Shen, P., Oefner, P. J., and Feldman, M. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7360–7365
338d. Gibbons, A. (2000) *Science* **290**, 1080–1081
338e. Cann, R. L. (2001) *Science* **291**, 1742–1747
339. De Mendoza, D. H., and Braginski, R. (1999) *Science* **283**, 1439–1440
340. Bradley, D. G., MacHugh, D. E., Cunningham, P., and Loftus, R. T. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5131–5135
341. Makalowski, W., and Boguski, M. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9407–9412
342. Nadeau, J. H., Grant, P. L., Mankala, S., Reiner, A. H., Richardson, J. E., and Eppig, J. T. (1995) *Nature (London)* **373**, 363–365
343. Collins, F. S., and Jegalian, K. G. (1999) *Sci. Am.* **281**(Dec), 86–91
344. O'Brien, S. J., and Stanyon, R. (1999) *Nature (London)* **402**, 365–366
344a. Hedges, S. B., and Kumar, S. (2002) *Science* **297**, 1283 – 1285
344b. Knight, J. (2002) *Nature (London)* **417**, 374–376
345. Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997) *Science* **278**, 631–637
346. Gibbons, A. (1998) *Science* **280**, 675–676
347. Rowe, T. (1999) *Nature (London)* **398**, 283–284
348. Bromham, L., Rambaut, A., Fortey, R., Cooper, A., and Penny, D. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12386–12389
349. Knoll, A. H. (1999) *Science* **285**, 1025–1026
350. Maniloff, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10004–10006
351. Lawrence, J. G., and Ochman, H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9413–9417
352. Doolittle, W. F. (1999) *Trends Biochem. Sci.* **24**, M5–M8
353. Lake, J. A., Jain, R., and Rivera, M. C. (1999) *Science* **283**, 2027–2028
354. Pennisi, E. (1998) *Science* **280**, 672–674
355. Stein, J. L., and Simon, M. I. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6228–6230
356. Davies, J. (1999) *Trends Biochem. Sci.* **24**, M2–M5
356a. Doolittle, R. F. (2002) *Nature (London)* **416**, 697–700
356b. Bentley, S. D. and 42 other authors (2002) *Nature (London)* **417**, 141–147
356c. Barnett, M. J., and 25 other authors (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9883–9888
356d. Galibert, F., and 55 other authors (2001) *Science* **293**, 668–672
356e. Heidelberg, J. F., and 31 other authors (2000) *Nature (London)* **406**, 477–483
356f. Parkhill, J., and 40 other authors (2001) *Nature (London)* **413**, 848–852
356g. Parkhill, J., and 34 other authors (2001) *Nature (London)* **413**, 523–527
356h. Moreno, E., and Moriyón, I. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1–3
356i. Cole, S. T., and 43 other authors (2001) *Nature (London)* **409**, 1007–1011
356j. Enserink, M. (2002) *Science* **295**, 1442–1443
356k. Tettelin, H., and 38 other authors (2001) *Science* **293**, 498–506
356l. Stover, C. K., and 30 other authors (2000) *Nature (London)* **406**, 959–964
356m. Glaser, P., and 54 other authors (2001) *Science* **294**, 849–853
356n. Perna, N. T., and 27 other authors (2001) *Nature (London)* **409**, 529–533
356o. Glass, J. I., Lefkowitz, E. J., Glass, J. S., Heiner, C. R., Chen, E. Y., and Cassell, G. H. (2000) *Nature (London)* **407**, 757–761
356p. Nierman, W. C., and 36 other authors (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4136–4141
356q. Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. (2000) *Nature (London)* **407**, 81–86
356r. Ng, W. V., Kennedy, S. P., Mahairas, G. G., and 40 other authors. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12176 – 12181

References

356s. Kawashima, T., Amano, N., Koike, H., Makino, S.-i, Higuchi, S., and 10 other authors. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14257– 14262

356t. Cole, S. T., and 41 other authors. (1998) *Nature (London)* **393**, 537 – 544

357. Edwards, J. S., and Palsson, B. O. (1999) *J. Biol. Chem.* **274**, 17410 – 17416

358. Strauss, E. J., and Falkow, S. (1997) *Science* **276**, 707 – 711

358a. da Silva, A. C. R., and 64 other authors. (2002) *Nature (London)* **417**, 459 – 463

359. Karp, P. D. (1998) *Trends Biochem. Sci.* **23**, 114 – 116

359a. Covert, M. W., Schilling, C. H., Famili, I., Edwards, J. S., Goryanin, I. I., Selkov, E., and Palsson, B. O. (2001) *Trends Biochem. Sci* **26**, 179 – 186

360. Hutchison, C. A., III, Peterson, S. N., Gill, S. R., Cline, R. T., White, O., Fraser, C. M., Smith, H. O., and Venter, J. C. (1999) *Science* **286**, 2165 – 2169

361. Mushegian, A. R., and Koonin, E. V. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10268 – 10273

361a. Szostak, J. W., Bartel, D. P., and Luisi, P. L., (2001) *Nature (London)* **409**, 387 – 390

362. Cho, M. K., Magnus, D., Caplan, A. L., McGee, D., and Ethics of Genomics Group. (1999) *Science* **286**, 2087 – 2090

362a. Stephanopoulos, G., and Kelleher, J. (2001) *Science* **292**, 2024 – 2025

362b. Holden, C. (2002) *Science* **297**, 1459 – 1460

363. Winzeler, E. A., and 51 other authors (1999) *Science* **285**, 901 – 906

364. Hieter, P. (1999) *Nature (London)* **402**, 362 – 363

364a. Ross-Macdonald, P., and 17 other authors (1999) *Nature (London)* **402**, 413 – 418

365. Glaever, G., and 72 other authors. (2002) *Nature (London)* **418**, 387 – 391

365a. Eisen, J. A. (2002) *Nature (London)* **415**, 845 – 848

365b. Wood, V., and 132 other authors (2002) *Nature (London)* **415**, 871– 880

365c. Seo, H.-C., Kube, M., Edvardsen, R. B., Jensen, M. F., Beck, A., Spriet, E., Gorsky, G., Thompson, E. M., Leharch, H., Reinhardt, R., and Chourrout, D. (2001) *Science* **294**, 2506

366. Fields, S., Kohara, Y., and Lockhart, D. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8825 – 8826

367. Tabara, H., Motohashi, T., and Kohara, Y. (1996) *Nucleic Acids Res.* **24**, 2119 – 2124

368. DeRisi, J. L., Iyer, V. R., and Brown, P. O. (1997) *Science* **278**, 680 – 686

369. Walhout, A. J. M., Sordella, R., Lu, X., Hartley, J. L., Temple, G. F., Brasch, M. A., Thierry-Mieg, N., and Vidal, M. (2000) *Science* **287**, 116 – 122

370. Hutter, H., Vogel, B. E., Plenefisch, J. D., Norris, C. R., Proenca, R. B., Spieth, J., Guo, C., Mastwal, S., Zhu, X., Scheel, J., and Hedgecock, E. M. (2000) *Science* **287**, 989 – 994

370a. Zdobnov, E. M., and 35 other authors. (2002) *Science* **298**, 149 – 159

370b. Miki, R., and 29 other authors. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2199 – 2204

371. Lin, X., and 36 other authors (1999) *Nature (London)* **402**, 761 – 768

372. European Union Arabidopsis Genome Sequencing Consortium, and Cold Spring Harbor, W. U. in St. L. and P. E. B. A. S. C. (1999) *Nature (London)* **402**, 769 – 777

373. Mathé, C., Peresetsky, A., Déhais, P., Van Montagu, M., and Rouzé, P. (1999) *J. Mol. Biol.* **285**, 1977 – 1991

374. Gura, T. (2000) *Science* **287**, 412 – 414

375. Meyerowitz, E. M. (1999) *Nature (London)* **402**, 731 – 732

375a. The Arabidopsis Genome Initiative. (2000) *Nature (London)* **408**, 796 – 815

375b. Seki, M., and 19 other authors (2002) *Science* **296**, 141– 145

375c. Pennisi, E. (2000) *Science* **290**, 32 – 35

375d. Normile, D., and Pennisi, E. (2002) *Science* **296**, 32 – 39

375e. Bennetzen, J. (2002) *Science* **296**, 60 – 63

375f. Yu, J., and 99 other authors (2002) *Science* **296**, 79 – 92

375g. Goff, S. A., and 54 other authors (2002) *Science* **296**, 92 – 100

376. Somerville, C., and Somerville, S. (1999) *Science* **285**, 380 – 383

377. Pennisi, E. (1998) *Science* **282**, 652 – 654

378. Capy, P. (2000) *Science* **287**, 985 – 986

379. Scriver, C. R., Kaufman, S., Eisensmith, R. C., and Woo, S. L. C. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 1015 – 1075, McGraw-Hill, New York

380. Barinaga, M. (1998) *Science* **281**, 32 – 34

381. Schultheis, P. J., Lorenz, J. N., Meneton, P., Nieman, M. L., Riddle, T. M., Flagella, M., Duffy, J. J., Doetschman, T., Miller, M. L., and Shull, G. E. (1998) *J. Biol. Chem.* **273**, 29150 – 29155

382. Chairoungdua, A., Segawa, H., Kim, J. Y., Miyamoto, K.-i, Haga, H., Fukui, Y., Mizoguchi, K., Ito, H., Takeda, E., Endou, H., and Kanai, Y. (1999) *J. Biol. Chem.* **274**, 28845 – 28848

383. Chambers, E. J., Bloomberg, G. B., Ring, S. M., and Tanner, M. J. A. (1999) *J. Mol. Biol.* **285**, 1289 – 1307

384. Annunen, S., Paassilta, P., Lohiniva, J., Perälä, M., Pihlajamaa, T., Karppinen, J., Tervonen, O., Kröger, H., Lähde, S., Vanharanta, H., Ryhänen, L., Göring, H. H. H., Ott, J., Prockop, D. J., and Ala-Kokko, L. (1999) *Science* **285**, 409 – 412

385. Ward, C. J., Turley, H., Ong, A. C. M., Comley, M., Biddolph, S., Chetty, R., Ratcliffe, P. J., Gatter, K., and Harris, P. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1524 – 1528

386. Hugot, J.-P., 16 other authors, and Groupe d'Etude Thérapeutique des Affections Inflammatoires Digestives. (1996) *Nature (London)* **379**, 821 – 823

387. Ghosh, S., and 39 other authors (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2198 – 2203

388. Kinzler, K. W., and Vogelstein, B. (1993) *Nature (London)* **363**, 495

389. Brehm, A., and Kouzarides, T. (1999) *Trends Biochem. Sci.* **24**, 142 – 145

390. Cho, Y., Gorina, S., Jeffrey, P. D., and Pavletich, N. P. (1994) *Science* **265**, 346 – 355

391. Oren, M. (1999) *J. Biol. Chem.* **274**, 36031 – 36034

392. Venkitaraman, A. R. (1999) *Science* **286**, 1100 – 1102

393. Smith, J. R., and 22 other authors (1996) *Science* **274**, 1371 – 1374

394. Novelli, M. R., Williamson, J. A., Tomlinson, I. P. M., Elia, G., Hodgson, S. V., Talbot, I. C., Bodmer, W. F., and Wright, N. A. (1996) *Science* **272**, 1187 – 1190

395. Rampino, N., Yamamoto, H., Ionov, Y., Li, Y., Sawai, H., Reed, J. C., and Perucho, M. (1997) *Science* **275**, 967 – 969

396. Rabbitts, T. H. (1994) *Nature (London)* **372**, 143 – 149

397. Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1998) *Nature (London)* **396**, 643 – 649

398. Kappel, C. A., Bieberich, C. J., and Jay, G. (1994) *FASEB J.* **8**, 583 – 592

399. Ahmad, W., and 17 other authors (1998) *Science* **279**, 720 – 724

400. Bassett, D. E., Jr., Boguski, M. S., and Hieter, P. (1996) *Nature (London)* **379**, 589 – 590

401. Richards, R. I., and Sutherland, G. R. (1997) *Trends Biochem. Sci.* **22**, 432 – 436

402. Wang, Y.-H., Gellibolian, R., Shimizu, M., Wells, R. D., and Griffith, J. (1996) *J. Mol. Biol.* **263**, 511 – 516

403. Weisman-Shomer, P., Naot, Y., and Fry, M. (2000) *J. Biol. Chem.* **275**, 2231 – 2238

404. Fridell, R. A., Benson, R. E., Hua, J., Bogerd, H. P., and Cullen, B. R. (1996) *EMBO J.* **15**, 5408 – 5414

405. Mandel, J.-L. (1997) *Nature (London)* **386**, 767 – 769

405a. Moine, H., and Mandel, J.-L. (2001) *Science* **294**, 2487– 2488

405b. Greenough, W. T., Klintsova, A. Y., Irwin, S. A., Galvez, R., Bates, K. E., and Weiler, I. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7101– 7106

405c. Sohn, E. (2001) *Science* **294**, 1809

405d. Deng, J., Xiong, Y., and Sundaralingam, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13665 – 13670

406. Wells, R. D. (1996) *J. Biol. Chem.* **271**, 2875 – 2878

407. Timchenko, L. T., and Caskey, C. T. (1996) *FASEB J.* **10**, 1589 – 1597

407a. Sinden, R. R. (2001) *Nature (London)* **411**, 757– 758

408. Muragaki, Y., Mundlos, S., Upton, J., and Olsen, B. R. (1996) *Science* **272**, 548 – 551

408a. Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., and Ranum, L. P. W. (2001) *Science* **293**, 864 – 867

409. Singer, R. H. (1998) *Science* **280**, 696 – 697

409a. Dhe-Paganon, S., Shigeta, R., Chi, Y.-I., Ristow, M., and Shoelson, S. E. (2000) *J. Biol. Chem.* **275**, 30753 – 30756

409b. Lodi, R., Cooper, J. M., Bradley, J. L., Manners, D., Styles, P., Taylor, D. J., and Schapira, A. H. V. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11492 – 11495

409c. Chen, O. S., and Kaplan, J. (2000) *J. Biol. Chem.* **275**, 7626 – 7632

409d. Foury, F., and Talibi, D. (2001) *J. Biol. Chem.* **276**, 7762 – 7768

409e. Gakh, O., Adamec, J., Gacy, A. M., Twesten, R. D., Owen, W. G., and Isaya, G. (2002) *Biochemistry* **41**, 6798 – 6804

410. Campuzano, V., and 26 other authors (1996) *Science* **271**, 1423 – 1427

410a. LeProust, E. M., Pearso, C. E., Sinden, R. R., and Gao, X. (2000) *J. Mol. Biol.* **302**, 1063 – 1080

410b. Rosenberg, R. N. (1995) *N. Engl. J. Med.* **333**, 1351– 1353

411. Mariappan, S. V. S., Catasti, P., Silds, L. A., III, Bradbury, E. M., and Gupta, G. (1999) *J. Mol. Biol.* **285**, 2035 – 2052

412. Dürr, A., Cossee, M., Agid, Y., Campuzano, V., Mignard, C., Penet, C., Mandel, J.-L., Brice, A., and Koenig, M. (1996) *N. Engl. J. Med.* **335**, 1169 – 1175

413. Rosenberg, R. N. (1996) *N. Engl. J. Med.* **335**, 1222 – 1224

414. Perutz, M. F. (1999) *Trends Biochem. Sci.* **24**, 58 – 63

414a. Holbert, S., Denghien, I., Kiechle, T., Rosenblatt, A., Wellington, C., Hayden, M. R., Margolis, R. L., Ross, C. A., Dausset, J., Ferrante, R. J., and Néri, C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1811– 1816

414b. Peters, M. F., and Ross, C. A. (2001) *J. Biol. Chem.* **276**, 3188 – 3194

414c. Bates, G. P. (2001) *Nature (London)* **413**, 691– 694

415. Wellington, C. L., and 20 other authors (1998) *J. Biol. Chem.* **273**, 9158 – 9167

References

416. Karlin, S., and Burge, C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1560 – 1565
416a. Kaytor, M. D., and Warren, S. T. (1999) *J. Biol. Chem.* **274**, 37507– 37510
416b. Perutz, M. F., and Windle, A. H. (2001) *Nature (London)* **412**, 143 – 144
416c. Chai, Y., Wu, L., Griffin, J. D., and Paulson, H. L. (2001) *J. Biol. Chem.* **276**, 44889 – 44897
417. Diamond, M. L., Robinson, M. R., and Yamamoto, K. R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 657 – 661
417a. Fernandez-Funez, P., and 14 other authors (2000) *Nature (London)* **408**, 101– 106
417b. Mushegian, A. R., Vishnivetskiy, S. A., and Gurevich, V. V. (2000) *Biochemistry* **39**, 6809 – 6813
418. Skinner, P. J., Koshy, B. T., Cummings, C. J., Klement, I. A., Helin, K., Servadio, A., Zoghbi, H. Y., and Orr, H. T. (1997) *Nature (London)* **389**, 971 – 974
419. Collick, A., Norris, M. L., Allen, M. J., Bois, P., Barton, S. C., Surani, M. A., and Jeffreys, A. J. (1994) *EMBO J.* **13**, 5745 – 5753
420. Sutherland, G. R., and Richards, R. I. (1994) *N. Engl. J. Med.* **331**, 191 – 193
421. McMurray, C. T. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1823 – 1825
422. Warpeha, K. M., Xu, W., Liu, L., Charles, I. G., Patterson, C. C., Ah-Fat, F., Harding, S., Hart, P. M., Chakravarthy, U., and Hughes, A. E. (1999) *FASEB J.* **13**, 1825 – 1832
423. Petruska, J., Hartenstine, M. J., and Goodman, M. F. (1998) *J. Biol. Chem.* **273**, 5204 – 5210
424. Pearson, C. E., Eichler, E. E., Lorenzetti, D., Kramer, S. F., Zoghbi, H. Y., Nelson, D. L., and Sinden, R. R. (1998) *Biochemistry* **37**, 2701 – 2708
425. Iyer, R. R., Pluciennik, A., Rosche, W. A., Sinden, R. R., and Wells, R. D. (2000) *J. Biol. Chem.* **275**, 2174 – 2184
426. Jakupciak, J. P., and Wells, R. D. (1999) *J. Biol. Chem.* **274**, 23468 – 23479
427. Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X. C., Stern, D., Winkler, J., Lockhart, D. J., Morris, M. S., and Fodor, S. P. A. (1996) *Science* **274**, 610 – 614
428. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. (1999) *Science* **286**, 531 – 537
429. Service, R. F. (1998) *Science* **282**, 396 – 399
430. Gerhold, D., Rushmore, T., and Caskey, C. T. (1999) *Trends Biochem. Sci.* **24**, 168 – 173
430a. Pennisi, E. (2002) *Science* **297**, 1985, 1987
430b. Schena, M. (2002) *Microarray Analysis*, Wiley, New York
431. Gerry, N. P., Witowski, N. E., Day, J., Hammer, R. P., Barany, G., and Barany, F. (1999) *J. Mol. Biol.* **292**, 251 – 262
432. Tang, K., Fu, D.-J., Julien, D., Braun, A., Cantor, C. R., and Köster, H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10016 – 10020
433. Marshall, E. (1999) *Science* **286**, 444 – 447
434. Dalton, R. (2000) *Nature (London)* **403**, 234
435. Burns, M. A., Johnson, B. N., Brahmasandra, S. N., Handique, K., Webster, J. R., Krishnan, M., Sammarco, T. S., Man, P. M., Jones, D., Heldsinger, D., Mastrangelo, C. H., and Burke, D. T. (1998) *Science* **282**, 484 – 487
436. Service, R. F. (1998) *Science* **282**, 399 – 401
437. Landegren, U., Kaiser, R., Caskey, C. T., and Hood, L. (1988) *Science* **242**, 229 – 237
438. Kolata, G. (1986) *Science* **234**, 22 – 24
439. Schnieke, A. E., Kind, A. J., Ritchie, W. A., Mycock, K., Scott, A. R., Ritchie, M., Wilmut, I., Colman, A., and Campbell, K. H. S. (1997) *Science* **278**, 2130 – 2133
440. Arnon, R. (1986) *Trends Biochem. Sci.* **11**, 521 – 524
441. McAleer, W. J., Buynak, E. B., Maigetter, R. Z., Wampler, D. E., Miller, W. J., and Hilleman, M. R. (1984) *Nature (London)* **307**, 178 – 180
442. Kapusta, J., Modelska, A., Figlerowicz, M., Pniewski, T., Letellier, M., Lisowa, O., Yusibov, V., Koprowski, H., Plucienniczak, A., and Legocki, A. B. (1999) *FASEB J.* **13**, 1796 – 1799
443. Weiner, D. B., and Kennedy, R. C. (1999) *Sci. Am.* **281**(Jul), 50 – 57
444. Taubes, G. (1997) *Science* **278**, 1711 – 1714
445. McDonald, G. A., Anacker, R. L., and Garjian, K. (1987) *Science* **235**, 83 – 85
446. Brown, F., Schild, G. C., and Ada, G. L. (1986) *Nature (London)* **319**, 549 – 550
447. Rupprecht, C. E., Wiktor, T. J., Johnston, D. H., Hamir, A. N., Dietzschold, B., Wunner, W. H., Glickman, L. T., and Koprowski, H. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7947 – 7950
448. Waldmann, T. A. (1986) *Science* **232**, 727 – 732
449. Travis, J., Owen, M., George, P., Carrell, R., Rosenberg, S., Hallewell, R. A., and Barr, P. J. (1985) *J. Biol. Chem.* **260**, 4384 – 4389
450. Parkman, R. (1986) *Science* **232**, 1373 – 1378
451. Sandig, V., Youil, R., Bett, A. J., Franlin, L. L., Oshima, M., Maione, D., Wang, F., Metzker, M. L., Savino, R., and Caskey, C. T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1002 – 1007
452. Verma, I. M., and Somia, N. (1997) *Nature (London)* **389**, 239 – 242
453. Anderson, W. F. (1995) *Sci. Am.* **273**(Sep), 124 – 128
454. Marshall, E. (1995) *Science* **269**, 1050 – 1055
455. Zanjani, E. D., and Anderson, W. F. (1999) *Science* **285**, 2084 – 2088
456. Finkel, T., and Epstein, S. E. (1995) *FASEB J.* **9**, 843 – 851
456a. Miller, D. G., and Stamatoyannopoulos, G. (2001) *N. Engl. J. Med.* **344**, 1782 – 1783
457. Miller, H. I. (2000) *Science* **287**, 591 – 592
457a. Gura, T. (2001) *Science* **291**, 1692 – 1697
458. Amalfitano, A., McVie-Wylie, A. J., Hu, H., Dawson, T. L., Raben, N., Plotz, P., and Chen, Y. T. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8861 – 8866
458a. Pawliuk, R., Westerman, K. A., Fabry, M. E., Payen, E., Tighe, R., Bouhassira, E. E., Acharya, S. A., Ellis, J., London, I. M., Eaves, C. J., Humphries, R. K., Beuzard, Y., Nagel, R. L., and Leboulch, P. (2001) *Science* **294**, 2368 – 2371
458b. Dunder, U., Kaartinen, V., Valtonen, P., Väänänen, E., Kosma, V.-M., Heisterkamp, N., Groffen, J., and Mononen, I. (2000) *FASEB J.* **14**, 361– 367
459. Capecchi, M. R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 956 – 957
460. Marshall, E. (1996) *Science* **271**, 586 – 588
461. Kohn, D. B., and Parkman, R. (1997) *FASEB J.* **11**, 635 – 639
462. Broitman, S., Amosova, O., Dolinnaya, N. G., and Fresco, J. R. (1999) *J. Biol. Chem.* **274**, 21763 – 21768
462a. Knight, J. (2001) *Nature (London)* **413**, 12 – 15
463. Alper, J. (1999) *Science* **283**, 1625 – 1626
464. Kishore, G. M., and Shewmaker, C. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5968 – 5972
465. Trewavas, A. (1999) *Nature (London)* **402**, 231 – 232
466. Goodman, R. M., Hauptli, H., Crossway, A., and Knauf, V. C. (1987) *Science* **236**, 48 – 54
467. Jones, M. G. K. (1985) *Nature (London)* **317**, 579 – 580
468. Roberts, L. (1988) *Science* **241**, 1290
469. Frommer, W. B., Ludewig, U., and Rentsch, D. (1999) *Science* **285**, 1222 – 1223
470. Kling, J. (1996) *Science* **274**, 180 – 181
471. Gura, T. (1999) *Science* **285**, 994 – 995
472. Simon Moffat, A. (1992) *Science* **256**, 770 – 771
473. Thomas, S. M., Davies, A. R. W., Birtwistle, N. J., Crowther, S. M., and Burke, J. F. (1996) *Nature (London)* **380**, 387 – 388
474. Marshall, E. (1996) *Science* **273**, 1788 – 1789
475. Enserink, M. (1998) *Science* **281**, 890 – 891
476. Dickson, D. (1993) *Nature (London)* **366**, 391
477. Poste, G. (1995) *Nature (London)* **378**, 534 – 536
478. Fuller, B. P., Ellis Kahn, M. J., Barr, P. A., Biesecker, L., Crowley, E., Garber, J., Mansoura, M. K., Murphy, P., Murray, J., Phillips, J., Rothenberg, K., Rothstein, M., Stopfer, J., Swergold, G., Weber, B., Collins, F. S., and Hudson, K. L. (1999) *Science* **285**, 1359 – 1361
479. Butler, D. (1998) *Nature (London)* **391**, 727
480. Hudson, K. L., Rothenberg, K. H., Andrews, L. B., Kahn, M. J. E., and Collins, F. S. (1995) *Science* **270**, 391 – 393
481. Lapham, E. V., Kozma, C., and Weiss, J. O. (1996) *Science* **274**, 621 – 624
482. Fost, N. (1992) *FASEB J.* **6**, 2813 – 2817
483. Holtzman, N. A. (1999) *Science* **286**, 409
484. Wills, C. (1970) *Sci. Am.* **222**(Mar), 98 – 107
485. Vogel, G. (1999) *Science* **286**, 2050 – 2051
486. Stoye, J. P. (1997) *Nature (London)* **386**, 126 – 127
487. Weiss, R. A. (1999) *Science* **285**, 1221 – 1222
488. Paradis, K., Langford, G., Long, Z., Heneine, W., Sandstrom, P., Switzer, W. M., Chapman, L. E., Lockey, C., Onions, D., XEN 111 Study Group, and Otto, E. (1999) *Science* **285**, 1236 – 1241
489. Enserink, M. (1999) *Science* **286**, 1662 – 1668
490. Haslberger, A. G. (2000) *Science* **287**, 431 – 432
491. Jackson, A., and Inglehearn, C. (1999) *Nature (London)* **402**, 457
492. Crawley, M. J. (1999) *Nature (London)* **400**, 501 – 502
493. Saxena, D., Flores, S., and Stotzky, G. (1999) *Nature (London)* **402**, 480
494. Smith, R. J. (1984) *Science* **224**, 1215 – 1216
495. Many, B. W. J., Almond, J. W., Berns, K. I., Chanock, R. M., Lvov, D. K., Pettersson, R. F., Schatzmayr, H. G., and Fenner, F. (1993) *Science* **262**, 1223 – 1224
496. Joklik, W. K., Moss, B., Fields, B. N., Bishop, D. H. L., and Sandakhchiev, L. S. (1993) *Science* **262**, 1225 – 1226
497. Jesson, L. K., and Barrett, S. C. H. (2002) *Nature (London)* **417**, 707
498. Human Genome Organization Ethics Committee. (2000) *Science* **290**, 49
499. Williamson, R., and Duncan, R. (2002) *Nature (London)* **418**, 585 – 586
500. Adam, D. (2002) *Nature (London)* **417**, 370
501. Aldhous, P. (2002) *Nature (London)* **418**, 355 – 356
502. Rothenberg, K. H., and Terry, S. F. (2002) *Science* **297**, 196 – 197
503. Nowlan, W. (2002) *Science* **297**, 195 – 196
504. McDowell, N. (2002) *Nature (London)* **416**, 571
505. Stokstad, E. (2002) *Science* **297**, 1797, 1799
506. Kaiser, J. (2001) *Science* **292**, 34 – 36
507. Enserink, M. (2002) *Science* **297**, 30 – 31
508. Clarke, T. (2002) *Nature (London)* **419**, 429 – 430
509. Rieger, M. A., Lamond, M., Preston, C., Powles, S. B., and Roush, R. T. (2002) *Science* **296**, 2386 – 2388
510. Spurgeon, D. (2001) *Nature (London)* **409**, 749

Study Questions

1. The two chromatids in Fig. 26-13G are said to be coiled with opposite handedness. Can you draw this conclusion from Fig. 26 alone? What are the biological implications for mitosis?[280] Does the DNA have a differing chirality at the molecular level? Compare this observation with the existence of snail shells or flowers[497] with both right and left handedness within the same species.

2. Will the sequencing of the human genome ever be complete?

3. Should patenting of human genes be allowed? Under what circumstances should patents be allowed on genetically engineered genes?

4. Should the human genome be regarded as a **common heritage** such that there is a guarantee that the medical and other benefits arising from genetic research are available to all persons on earth? See Human Genome Organization Ethics Committee.[498]

5. DNA testing (Box 5-D) is widely used by police throughout the world. It has been estimated that if ten loci in the DNA are tested the chance of a random match between two people is one in a billion. In the United Kingdom it is planned to hold DNA profiles on record for one of every 15 people. Is this wise? Or should DNA profiles be recorded for all people?[499] If DNA profiles are on record how can we be sure that they are not used dishonestly? Should police have access to DNA data bases? See Adam.[500]

6. Should "race" be used as a variable in biomedical studies? See Aldhous.[501]

7. Should insurance companies be allowed to have access to genetic information about insured people? Companies usually obtain other medical information. See Rothenberg and Terry,[502] Adam,[500] and Nowlan.[503]

8. Should genetically engineered fish be allowed in "farms" that are set up in ocean waters? See McDowell[504] and Stokstad.[505]

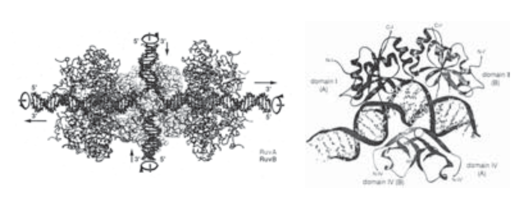
9. Do transgenic trees pose a threat to natural ecosystems? See Kaiser.[506]

10. Should we attempt to replace wild populations of mosquitoes with genetically engineered mosquitoes that can not transmit malaria? See Enserick[507] and Clarke.[508]

11. Could dispersal of pollen from genetically modified plants lead to undesirable "genetic pollution" of the environment? See Rieger *et al*.[509]

12. To what extent is genetic modification of plants and animals equivalent to changes made by conventional breeding? See Spurgeon.[510]

13. Choose one of the ethical questions that can be raised about application of our new knowledge about the genome (e.g., see pp. 1518, 1519, Chapter 32, and study questions on this page). Study literature available to you and prepare a recommendation to the public, Congress, or to local regulatory agencies. Follow a scientific approach. Try to find true facts that can be verified. Consider all viewpoints. State some of the uncertainties in your recommendation. Present your proposal to a class or to a friend for criticism. Then publish your view in a newspaper if you wish to.



Many proteins interact with DNA. These include polymerases that replicate DNA, helicases that unwind double helices, topoisomerases that cut and reseal DNA strands to avoid entangling, and enzymes that repair damage to DNA. At left is a motor complex consisting of the tetrameric binding protein RuvA (light shading) and the hexameric helicase RuvB (two copies in darker shading). See also Fig. 27-26B. From Putnam *et al.* (2001) *J. Mol. Biol.*, **311**, 297–310. Right bacterial protein (from *Thermus aquaticus*) MutS recognizes mispaired bases in DNA and initiates their removal. The DNA is bound in a bent conformation. In this duplex one strand contains an unpaired thymine (at top of bend), which would cause a mutation if not removed. From Obmolova *et al.* (2000) *Nature*, **407**, 703–7.

Contents

1529 .... **A. The Topology and Environment of DNA**
1529 .... 1. DNA in Viruses
1530 .... 2. Bacterial Chromosomes and Plasmids
1531 .... 3. Protamines, Histones, and Nucleosomes
1531 .... *Nucleosomes*
1533 .... *The linker histones*
1534 .... *Folding of nucleosome chains; chromosomes*
1535 .... 4. The Cell Nucleus
1535 .... *The nuclear matrix*
1535 .... *Other nonhistone nuclear proteins*
1536 .... *Nuclear membrane, pores, and lamina*
1537 .... *The nucleolus*
1537 .... **B. Organization of DNA**
1537 .... 1. Repetitive DNA
1537 .... *Centromeres*
1538 .... *Telomeres*
1538 .... *Short interspersed sequences (SINES)*
1539 .... *Long interspersed sequences (LINES)*
1539 .... 2. Genes for Ribosomal RNA and Small RNA Molecules
1539 .... 3. Other Gene Clusters and Pseudogenes
1540 .... 4. Introns, Exons, and Overlapping Genes
1540 .... 5. DNA of Organelles
1541 .... 6. Methylation of DNA
1541 .... *CG doublets*
1542 .... *Imprinting*
1542 .... **C. Replication**
1542 .... 1. Early Studies
1543 .... *Autoradiography*
1543 .... *The chemistry of DNA polymerization*
1544 .... *Discontinuous replication and RNA primers*
1544 .... 2. DNA Polymerases
1544 .... *Exonuclease activities, proofreading, and editing*
1547 .... *Other Class A polymerases*
1547 .... *Polymerases of Class B*
1548 .... *Other DNA polymerases*
1548 .... *DNA polymerase III*
1549 .... 3. Other Replication Proteins
1549 .... *DNA ligases*
1549 .... *Single-strand binding proteins*
1550 .... *Helicases and topoisomerases*
1552 .... *Primases, initiator proteins, and ribonucleases*
1553 .... 4. Replication of Bacterial DNA
1554 .... *Directions of replication*
1554 .... *Origins of replication*
1555 .... *Priming and initiation of DNA synthesis*
1556 .... *Elongation of DNA chains*
1557 .... *Termination of replication*
1557 .... 5. The Replication of Viral DNA
1559 .... 6. Packaging of Viral Genomes
1559 .... 7. Plasmids
1559 .... 8. Chromosome Ends
1561 .... 9. Mitochondrial and Chloroplast DNA
1561 .... 10. Replication of Eukaryotic Nuclear and Viral DNA
1562 .... *Eukaryotic viruses*
1562 .... *Artificial chromosomes*
1562 .... *Replication of nuclear DNA*
1564 .... **D. Integration, Excision, and Recombination of DNA**
1564 .... 1. Recombination Mechanisms
1565 .... *The Holliday recombination intermediate*
1566 .... *RecA and other strand-exchange proteins*
1567 .... *Processing the Holliday junction*
1568 .... 2. Nonreciprocal Recombination and Unequal Crossing-Over
1570 .... 3. Site-Specific Recombination and the Integration and Excision of DNA
1570 .... *The integrase (tyrosine recombinase) family*
1572 .... *The resolvase/invertase family and invertible DNA sequences*
1573 .... 4. Transposons and Insertion Sequences
1575 .... *Cut-and-paste (nonreplicative) transposons*
1575 .... *Replicative transposons*
1576 .... *The temperate bacteriophage Mu*
1576 .... *Some other transposons*
1577 .... 5. Other Causes of Genetic Recombination
1578 .... **E. Damage and Repair of DNA**
1578 .... 1. Causes of Mutations
1579 .... 2. Fidelity of Replication
1579 .... *The role of polymerases*
1580 .... *Editing (proofreading)*
1580 .... 3. Repair of Damaged DNA
1580 .... *Methyl-directed and other mismatch repair*
1581 .... *Excision repair*
1581 .... *Base excision repair (BER)*
1583 .... *Reactions of alkylated bases*
1583 .... *Repair of double-strand breaks*
1583 .... *The SOS response and translesion repair*
1584 .... *Poly (ADP-ribose)*
1584 .... **F. Mutagens in the Environment**

1590 .... **References**
1601 .... **Study Questions**

Boxes

1585 .... Box 27-A Deficiencies in Human DNA Repair

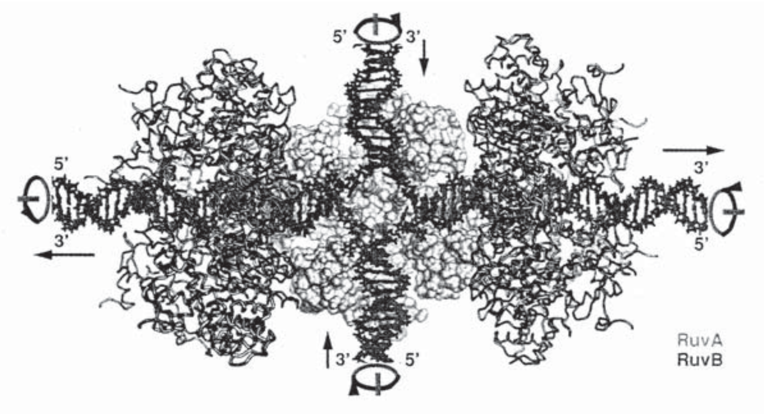
Tables

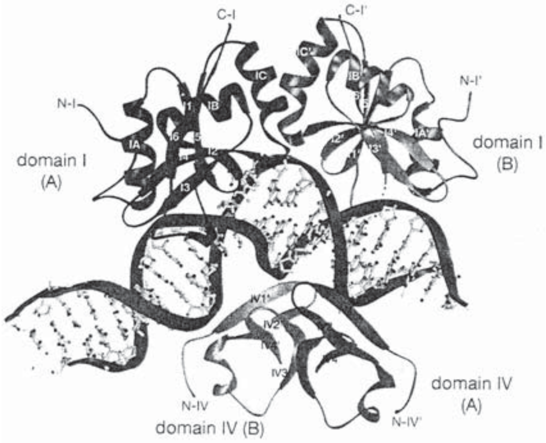
1545 .... Table 27-1 Families of DNA Polymerases

1551 .... Table 27-2 Some Proteins of DNA Replication in *E. coli*

Organization, Replication, Transposition, and Repair of DNA

27





Both the replication and transcription of DNA are complex processes. Although the basic chemistry is relatively simple many enzymes and other proteins are required. In part this reflects organizational and topological problems[1] associated with the huge amount of DNA present as a single molecule within a chromosome.

A. The Topology and Environment of DNA

Although we can isolate DNA in the form of simple double helical fragments, the topology of natural DNA is always more complex. Covalently closed circular DNA such as that in plasmids, mitochondria, and bacterial chromosomes is supercoiled (Chapter 5) and bound to proteins. The DNA of chromosomes and bacteriophage particles is folded further into more compact forms. For example, the chromosome of *E. coli* contains DNA about 1.5 mm in length folded within a cell that is only 2 µm long. The diploid length of DNA in a 20 µm cell of a human is about 1.5 meters. At the time of cell division human DNA must all be replicated and packaged into chromosomes, 23 pairs in each cell. The density of the compacted DNA varies. A bacterial nucleoid may contain 10–30 mg / ml of DNA.[2] In chromatin of a eukaryotic nucleus there may be 200 mg / ml of DNA and in nucleosomes 330–400 mg / ml.[2a] The tightly compacted head of the T4 bacteriophage (Box 7-C) contains 520 mg / ml.[3]

1. DNA in Viruses

In the simplest filamentous DNA viruses such as M13 the DNA is coated by a helical protein sheath (Fig. 7-7), as it is extruded from a cell. The sheath is peeled off as the virus enters another cell. However, in the large tailed phage (Box 7-C), which contains ~160 kb of polynucleotide chains, the DNA is closely packed within the heads. In a model that seems to accommodate most experimental results, the DNA rod bends sharply into a series of folds, which are laid down around the long axis of the head in spirally arranged shells (Fig. 27-1).[4,4a] The end of the DNA that enters the phage head first appears to be located in the center with successive shells of DNA around it.[4] In the large bacteriophage G the DNA appears to be folded to form 12 icosahedrally arranged pear-shaped rings in the corners of the capsid.[5] The 2.0 nm diameter double helical segments of DNA lie roughly parallel and are separated by only 0.5–1.0 nm of solvent,[6] which contains cations such as the polycation of spermidine. Another possibility is that the DNA may be wound as on a spool of thread.[7] The DNA chains have an external diameter of ~2 nm with 0.6 nm additional for a hydration layer. In a phage head the adjacent parallel chains are 2.6–2.7 nm apart. Thus, the packing is very tight. Even so, capsids tend to be only about half filled with DNA. An exception is provided by the tobacco mosaic virus (Fig. 7-8) in which the RNA genome is held precisely by protein subunits, which dissociate to release the RNA during infection.[3]

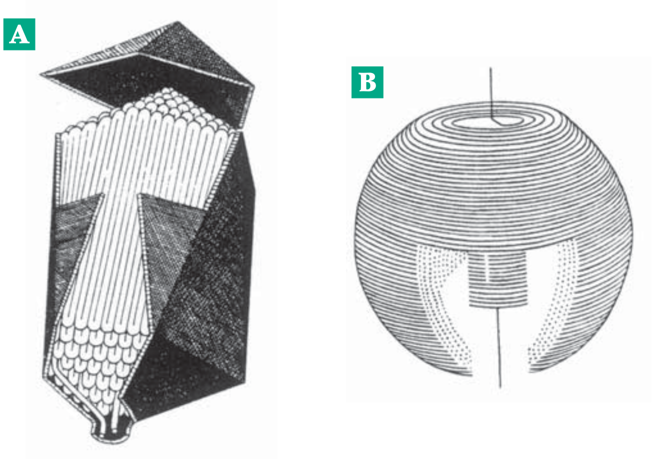


Figure 27-1 Possible ways of packing DNA into the heads of bacteriophage particles. (A) Spiral-fold. (B) Concentric shell model From Black *et al.*[4]

2. Bacterial Chromosomes and Plasmids

Most DNA in living organisms, whether bacteria or eukaryotes, is underwound. That is, the superhelix density (Chapter 5) is about −0.05 or one supercoil per 200 base pairs. In eukaryotes this negative supercoiling can be accounted for by the winding of DNA around the histones within the nucleosomes (Figs. 5-21, 27-4). The situation in bacteria is not as clear. There are many bacterial DNA-binding proteins,[8,9] but one of them known as HU is particularly abundant.[8–12] In *E. coli* it exists as mixed dimers and tetramers of 9.5 (α) and 9.2 (β) kDa subunits. Each HU tetramer can bind ~60 bp of DNA. There are about 60,000 HU monomers per cell, enough to coat ~20% of the genome. Possible modes of interaction with DNA have been proposed on the basis of the X-ray structure of HU.[8,13–14a] Binding to HU causes the DNA to be more tightly wound, introducing additional negative supercoils. The resulting unreleased torsional stress may be important in the functioning of the DNA.[15–17a] Binding is strongest to four-way junctions,[17a] and to DNA with nicks and gaps or to structures induced by supercoiling.[17b] Other basic histonelike proteins may also bind to the DNA. However, there are no structures that resemble eukaryotic nucleosomes.

If bacterial cells are lysed under certain conditions, e.g., in 1 M NaCl or in the presence of a "physiological" 5 mM spermidine, the entire bacterial chromosome can be isolated.[10] The DNA in these isolated chromosomes retains some torsional tension that, however, can be relaxed by nicking with nucleases or by γ-irradiation. However, a single nick relaxes the DNA very little. The explanation appears to be that the DNA is held by proteins of the nucleoid matrix in a series of loops (Fig. 27-2). A single nick relaxes just one loop. On this basis there are 43 ± 10 loops per genome with ~100 kb of DNA per loop.[18] A 136-residue protein **H-NS** is involved in condensation of bacterial DNA.[18a] It may act as a scaffolding protein, but it also functions in controlling transcription.[18b]

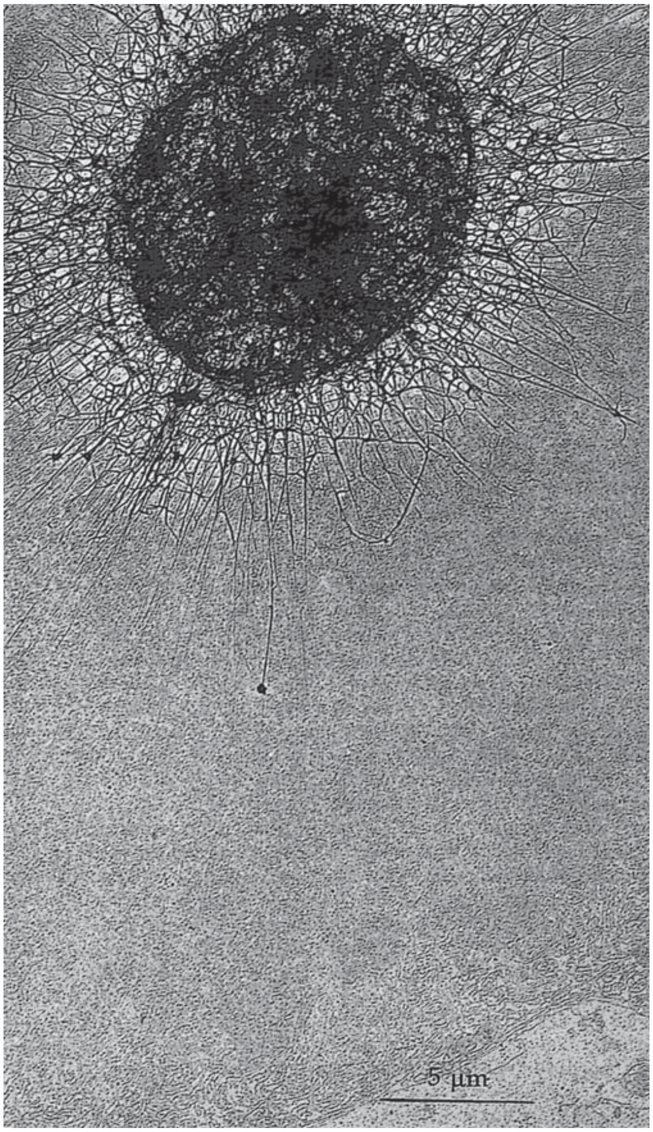


Figure 27-2 Electron micrograph of a bacterial nucleoid. The DNA is usually contained within the 'cage', but has been spread, using Kleinschmidt's procedure, to yield a surrounding skirt. The cage contains a protein network which includes elements of the cytoskeleton, enclosing the residual nuclear substructures. The denser fibrils radiating from this cage disappear to nuclease digestion and are probably aggregates of DNA fibers which merge with individual DNA strands at the extremities of the skirt. These strands are highly supercoiled, indicative of intact DNA. From Jackson and Patel, provided by Dr. S. J. McCready.[19]

3. Protamines, Histones, and Nucleosomes

Within bacterial cells the negatively charged phosphate groups of the DNA are neutralized to a large extent by the positively charged polyamines,[19a] by cations such as K^+ and Mg^{2+}, and by basic proteins such as HU. Within the mature heads of sperm cells of fish, the tightly packaged DNA is neutralized by the **protamines**, small ~5-kDa proteins rich in arginine.[4a,20–21a] Similar basic proteins are found in mammalian sperm.[22,23] However, within most eukaryotic cells, the charges on DNA are balanced principally by a group of basic proteins, first isolated and named **histone** by Kossel in 1884.[24]

There are five classes of histones, which range in molecular mass from ~11 to 21.5 kDa:[25,26]

$$\begin{array}{l} \text{H1 (including H1}^0\text{, H5; lysine-rich “linker” histone)} \\ \left.\begin{array}{l}\text{H2A, H2B (moderately lysine-rich)} \\ \text{H3, H4 (arginine-rich)}\end{array}\right\} \text{“core” histones} \end{array}$$

All of the core histones share a conserved 65-residue **histone fold.**[27,28] The arginine-rich histones have a strongly conserved amino acid sequence, histone H4 from pea seedlings differing from that of the bovine thymus by only two amino acids. On the other hand, the lysine-rich H1 is almost species-specific in its sequence. Differentiated tissues contain at least seven variant forms of histone H1 including proteins designated $H1^0$, H1t, and H5.[29–31]

The N-terminal 25–40 amino acids of the core histones are positively charged and highly conserved.[32] The 135-residue histone H3 of calf thymus carries a net charge of +18 within the first 53 residues. This is probably the portion that binds to DNA. On the other hand, the carboxyl terminal end is hydrophobic and only slightly basic.[33] Histones undergo substantial amounts of **micromodification** including phosphorylation, acetylation, and methylation.[33a,b] Mono-, di-, and tri-methyllysine residues may be present.[33c,d] The core histones all undergo acetylation on specific lysyl side chains. Nuclear histone acetyltransferases[34–38b] transfer acetyl groups from acetylCoA and hydrolytic deacetylases may remove them.[39–40b] The amount of acetylation varies during different stages of the cell cycle, suggesting a regulatory role.[41] Acetylation sites in H3 and H4 are highly conserved in all eukaryotes.[42]

A small fraction of histone H2A undergoes phosphorylation and dephosphorylation continuously,[43] but H1 and H3 are phosphorylated and dephosphorylated at specific stages of the cell cycle. Phosphorylation of H1 has been thought essential for “condensation” of chromatin,[44,45] the folding into the tightly packed chromosome structures. However, more recent experiments point to the N terminus of histone H2B as the required site of phosphorylation for chromosome condensation.[45a] In addition, histone H1 may interact with membrane lipids.[46] Histone H2A exhibits the greatest heterogeneity and appears to function in regulation of transcription, of gene silencing, and of repair of double-strand breaks in DNA.[46a] Each of the histones appears to be regulated separately. In animal chromatin ~10% of histone H2A and small fractions of H2B and of tissue-specific histones are covalently linked to ubiquitin (Box 10-C). However, this monoubiquitination may not be related to proteolytic degradation.[47] Archea contain histones that dimerize and bind DNA to form nucleosomes.[47a,b]

Nucleosomes. An early idea of the function of histones was that they serve as gene repressors. To some extent this view is still valid. However, the large quantity of histone and uniform distribution over the DNA suggested some other role. This was clarified when electron micrographs showed that chromatin fibers form **nucleosomes**,[48–50] regular repeating structures resembling beads on a string. The same structure is seen in the “minichromosomes” formed from virus SV40 (Fig. 27-3).[51–53] Two molecules each of histones H2A, H2B, H3, and H4 form the core of the nucleosome around which ~146 bp of dsDNA is coiled into approximately two negative, left-handed toroidal superhelical turns (Figs. 5-21, 27-4).

Digestion of chromatin by nucleases causes rapid cleavage into ~200-bp fragments and slower cleavage to 146 ± 20-bp fragments. This suggested that ~200-bp segments of DNA are folded around a histone octomer, contracting the 68 nm extended length of relaxed B-DNA into a 10-nm nucleosome. A short linker region of variable length, up to 80 bp, lies between the nucleosomes.[54] The fifth histone, H1 (or H5 in some species), may bind to this linker DNA. A nucleosome with bound H1 is sometimes called a **chromatosome**.[55]

The superhelix density of ~0.05 observed for DNA extracted from eukaryotic cells is just equal to one negative superhelix turn per nucleosome. For example, the number of nucleosomes seen in the minichromosome of Fig. 27-3 matches the numbers of supercoils in the SV40 DNA (Fig. 5-20). If there are two negative supercoils per nucleosome, as shown in Fig. 27-4, the DNA in the nucleosome must be wound more tightly than in relaxed DNA (10.0 bp per turn instead of the 10.6 of relaxed DNA).[56] NMR data suggest that within the nucleosome the regular base pairing in the DNA may be partially disrupted and that some parts of the histone have a high degree of mobility.[57]

Although nucleosomes are distributed rather evenly along the DNA of a cell, there are some DNA sequences that favor nucleosome formation. The resulting **positioned nucleosomes** are often found in the vicinity of gene promoters, enhancers, and other

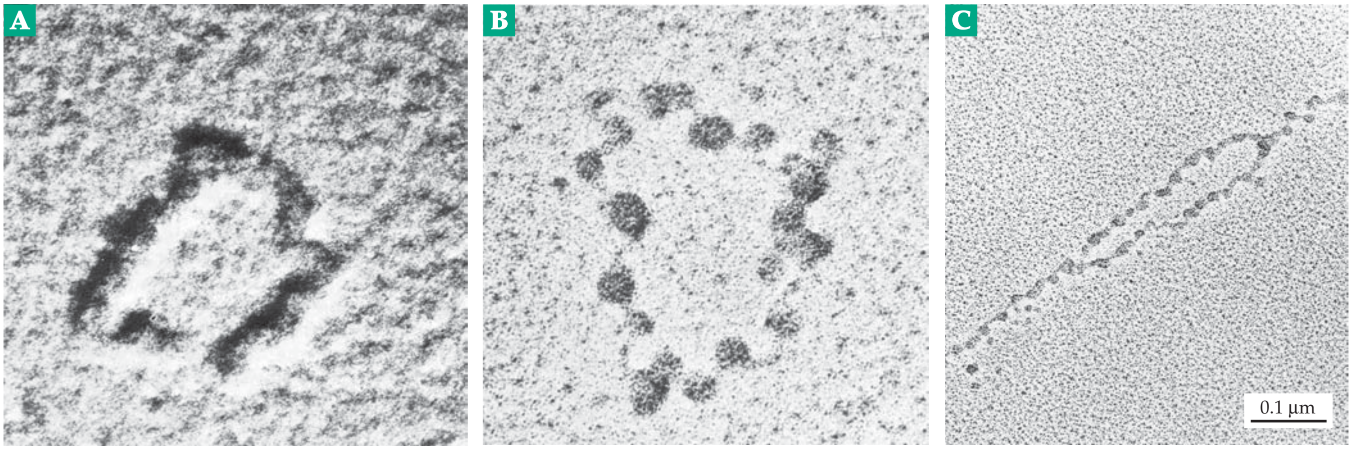


Figure 27-3 (A) Electron micrograph of "minichromosome" formed from virus SV40 growing in monkey cells in culture.[51] In this native form the nucleoprotein fiber is ~11 nm in diameter and ~210 nm in length. (B) Beaded form of minichromosome observed when the ionic strength was lowered. The 20 beads have diameters of ~11 nm and are joined by bridges roughly 2 nm in diameter and 13 nm long. Deproteinization and relaxation of the DNA revealed that the overall length of the DNA present is seven times the length of the native minichromosome. (C) Electron micrograph of chromatid of a blastoderm-stage embryo of *Drosophila melanogaster* in the process of replication. Nucleosomal particles are visible immediately adjacent to the replication forks. Courtesy of Steven L. McKnight and Oscar L. Miller, Jr.

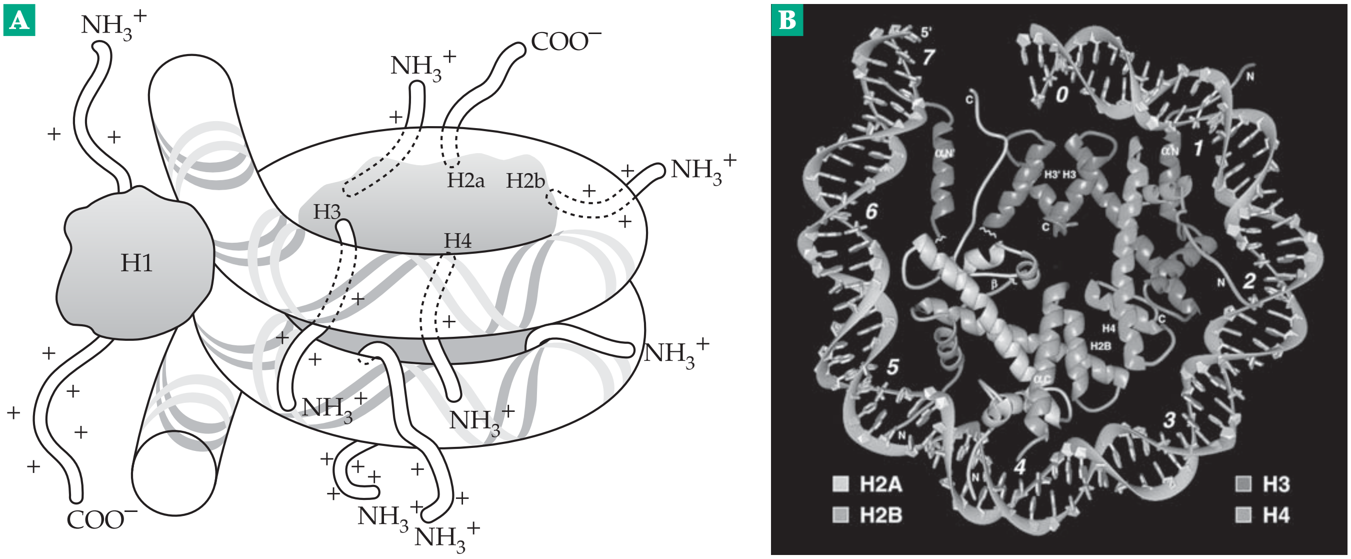


Figure 27-4 (A) A nucleosome is formed when dsDNA, shown schematically as a tube, wraps roughly twice around a histone octomer. This complex, or nucleosomal core particle, includes at its center two copies of histone H3, two of histone H4, and two H2A–H2B dimeric pairs, one of which is not visible. Ends of each histone molecule are thought to protrude like tails from the core, ready to interact with other molecules. In many organisms, histone H1, portrayed at left in one possible position, helps to anchor DNA to the core and promotes further compaction of the DNA into a 30-nanometer fiber. See Grunstein.[49] (B) Nucleosome core particle: 73-bp half. The view is down the superhelix axis with the pseudo dyad axis aligned vertically. The central base pair through which the dyad passes is above the superhelix axis location labeled 0. Each additional numerical label 1–7 represents one further DNA double helix turn. The complete histone proteins (except for the tail regions) that are primarily associated with the 73-bp superhelix half are shown. The two copies of each histone pair are distinguished as unprimed and primed, where the histone fold of the unprimed copy is primarily associated with the 73-bp DNA half and the primed copy with the 72-bp half. Four-helix bundles are labeled as H3', H3, and H2B, H4; histone-fold extensions of H3 and H2B are labeled as αN', αN, and αC, respectively; the interface between the H2A docking domain and the H4 C terminus as β; and N- and C-terminal tail regions as N or C. From Luger *et al*.[50]

control sequences.[58–61a] It seems likely that positioning of nucleosomes is related to control of transcription or of other activities. The same basic structure for chromatin has been found in animals, fungi, and green plants.

The linker histones. A nucleosome is pictured in Fig. 27-4A as if held in a compact configuration by the binding of histone H1 at a position that marks a pseudo twofold axis that lies in the plane of the nucleosome. However, this is only one of several possible locations for linker histones of the H1, H5 family.[62–67] The structure of the linker histones is somewhat different from that of core histones. They have an 80-residue globular domain with long N-terminal *and* C-terminal chains, both of which are rich in basic residues and evidently available for binding to DNA.[68] Perhaps they bind both to the DNA entering the nucleosome and to that leaving the nucleosome, reducing electrostatic repulsion of those two parts of the DNA superhelix.[69]

Another possible location for histone H1 or H5 is above the histone surface as shown in Fig. 27-4 and *inside* the DNA loop.[65] A third suggested location for the globular linker core is *between* the two turns of the DNA strand.[66] While one function of the linker histones may be to stabilize mononucleosomes, they may also play a role in compaction of the DNA into the 30-nm fibers universally seen in nuclei of cells.[62,64–64b]

Histone H1 can also be regarded as a general repressor, holding chromatin tightly folded and preventing transcription. The possible roles of acetylation, phosphorylation, methylation, ubiquitination, and other modifications of histones in controlling transcription, replication, and DNA repair are receiving increasing attention.[70–73] **Active chromatin**, where transcription is occurring, has an altered nucleosome structure and increased susceptibility to nuclease action. It appears to be less tightly packed than inactive chromatin and to contain regions called **hypersensitive sites** that are accessible to nucleases or chemical modification reagents.[74–76] There seems to be a direct link between increased acetylation of histones and enhanced initiation of transcription by RNA polymerase II.[35,38,77–82] Conversely, deacetylation is associated with repression of transcription. Both histone acetylase and deacetylase activities have been found in transcriptional regulators.[80–82]

The observation that H1 becomes phosphorylated during the initiation step of mitosis suggests another control mechanism for its repressor functions.[83] Several multiprotein complexes that "remodel" chromatin have been identified.[73,84] These complexes contain

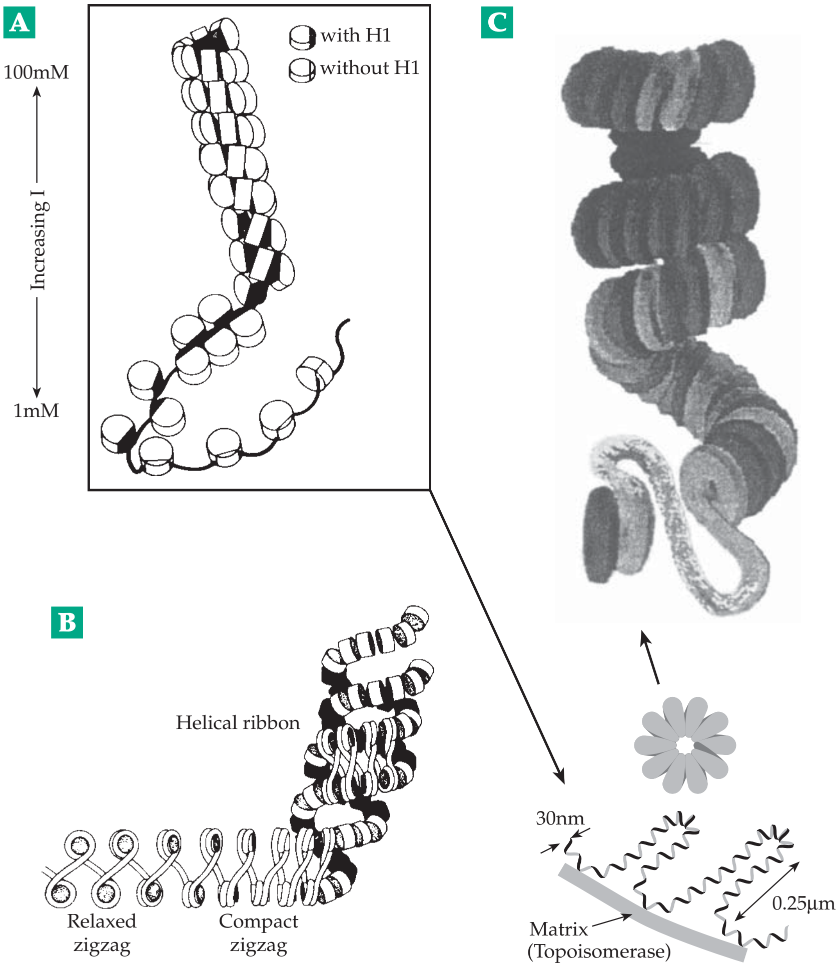


Figure 27-5 (A, B) Two possible models of the 30-nm chromatin fiber.[55] (A) Thoma *et al*.[85] (B) Woodcock *et al*.[64,87] The fully compacted structure is seen at the top of each figure. The bottom parts of the figures illustrate proposed intermediate steps in the ionic strength-induced compaction. (C) Possible organization of the DNA within a metaphase chromosome. Six nucleosomes form each turn of a solenoid in the 30-nm filament as in (A). The 30-nm filament forms ~30 kb-loop domains of DNA and some of these attach at the base to the nuclear matrix that contains topoisomerase II. About ten of the loops form a helical radial array of 250-nm diameter around the core of the chromosome. Further winding of this helix into a tight coil ~700 nm in diameter, as at the top in (C), forms a metaphase chromatid. From Manuelidis[91].

ATP-dependent **helicase** activities that open up DNA for transcription (see Chapter 28, Section B).

Folding of nucleosome chains; chromosomes. Electron microscopy shows that chromatin is packed into the nucleus largely as fibers of ~30- to 36-nm diameter.[55,84a] Thoma *et al.*[85] proposed that the fibers could be formed by winding the string of nucleosomes into a simple one-start helix with six nucleosomes per turn and a pitch of ~11 nm (Fig. 27-5A). The H1 molecules would be close together in the center. An alternative model, one of several suggested,[55,86] is shown in Fig. 27-5B. Its zigzag pattern of adjacent nucleosomes would generate a two-start helix.[64,87,88] Another alternative model envisions larger solenoids with interdigitated nucleosomes.[89]

A somewhat similar structure appears to be present in the specialized eukaryotic **lampbrush chromosomes** (Fig. 27-6), which are observed during the meiotic prophase of oocytes. They have been studied intensively in amphibians such as *Xenopus*. A lampbrush chromosome is actually a homologous pair of chromosomes, each one in turn consisting of two closely associated chromatids. The chromosomes are highly expanded, and about 5% of the DNA is extended in the form of ~4000 perfectly paired loops visible with an electron microscope. Each loop consists of ~50 µm or ~150 kb of extended DNA. No evidence of any breaks in the DNA is seen, a fact that supports the belief that a single DNA molecule extends from one end of the chromosome to the other through all of the loops.

Like the puffs of polytene chromosomes (Chapter 28), which may have a similar structure, lampbrush chromosomes appear to be actively engaged in transcription. Approximately 3% of the DNA may be functional in producing mRNA that is accumulated within the oocyte and is used as a template for protein synthesis during early embryonic development.[90]

A different arrangement is present in metaphase chromosomes, which appear as two dense parallel sister chromatids of ~700-nm diameter.[91] The DNA must be highly folded. In the model shown in Fig. 27-5C the 30-nm fiber is folded into ~30-kb loops, each one formed from ~25 turns of the 30-nm helix. The loops then form a helical array 250 nm in diameter with ~ ten loops (300 kb) per turn. This helix is further wound into a tight helix of ~700-nm diameter. A single turn of this helix may contain as much as 9 Mb of DNA.[91] About four hundred coils (an average of 18 coils per human haploid chromosome) could accommodate the entire genome. A group of five proteins, some of which are designated **SMC** (structural maintenance of chromosomes) proteins, form a complex called **cohesin**. SMC proteins, large multidomain proteins found in all eukaryotes, are also present in bacteria.[91a,b]

Interphase chromatin must be much less tightly

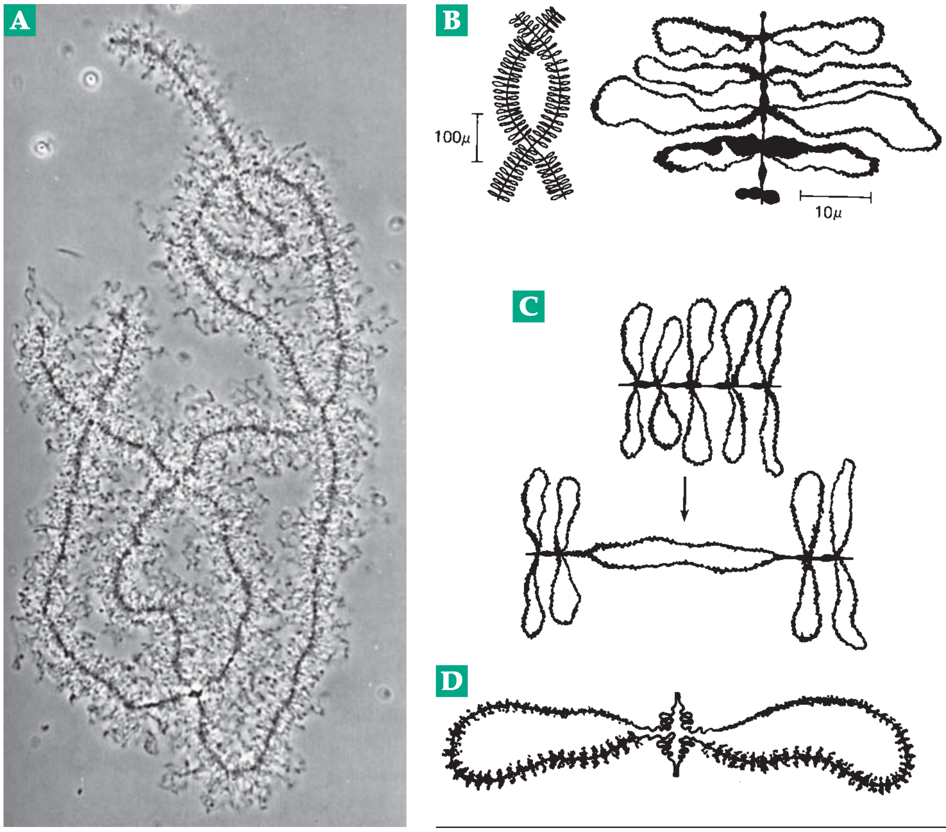


Figure 27-6 (A) Photomicrograph of a lampbrush chromosome from the nucleus of an oocyte of the newt *Triturus*. From L. M. Mays, *Genetics, A Molecular Approach*, Macmillan, New York, 1981, p. 227. (B–D) Diagrammatic views of lampbrush chromosomes. (B) The two homologous chromosomes (left) are held together by two chiasmata. A portion of the central chromosome axis (right) shows that two loops with identical morphology emerge at a given point, evidence that each chromosome has already split into two chromatids. (C) Accidental stretching of a chromosome reveals the continuity of the loop axis with the central axis. (D) A single loop pair, showing the single DNA molecules on which RNA chains (indicated by fuzzy shading) are being transcribed. From J. Gall, *Brookhaven Symp. Biol.* **8**, 17 (1955).

packed and contains regions in which large loops, e.g., of 20–120-kb size, are uncoiled enough to allow transcription factors and other proteins to locate their target sequences.[91c] Many models of interphase chromatin have been proposed.[92,93]

4. The Cell Nucleus

It has been clear for many years that the nucleus has a well-organized structure. However, techniques such as fluorescent labeling have only recently been used to provide important details. These studies show that individual interphase chromosomes occupy discrete territories within the nucleus.[94–96] Parts of the chromosomes may be unfolded and active in transcription, while others are more tightly coiled. Some parts, known as **heterochromatin**,[97–98a] are very tightly coiled and metabolically almost inert. These regions include the highly repetitive DNA of telomeres and centromeres as well as other regions and complete inactivated female X chromosomes. Replication, transcription, and RNA splicing complexes are found at distinct locations within cells.[99–102a] They are apparently fixed, perhaps attached to the inner nuclear membrane, while the DNA passes through the complexes.

The nuclear matrix. The lipid bilayers, the histones and other soluble proteins, and the DNA can all be removed from nuclei by extraction and enzymatic digestion. An insoluble residue, the **nuclear matrix,** is left.[103–107] Largely protein in nature, this matrix is spread throughout the nucleus. Remnants of the membranes remain in the form of proteins that were in or along the bilayer. The nucleolus is clearly defined. The DNA appears to be bound to the nuclear matrix proteins. A specific 320-kb piece of a *Drosophila* chromosome has been mapped and used to locate nontranscribed scaffold (or matrix) attachment regions of DNA bound to matrix proteins. These were found at intervals of 26–112 kb, the intervening loops containing up to five or more genes.[105,106,108,109] A 120-kDa protein together with topoisomerase II (Section C,2)[102,104] may be components of a **nuclear scaffold** that constrains the loops of DNA. The scaffold may also provide locations for the complexes of proteins involved in replication and other processes.[103] The **matrix attachment regions** (MARs) may also act as **insulators** that shield promoters for transcription within certain loops from control elements such as **enhancers** that may be present in adjacent loops.[91b,110] At least one nuclear matrix component becomes phosphorylated and moves to the nuclear poles during mitosis.[111]

Other nonhistone nuclear proteins. Polyacrylamide gel electrophoresis revealed more than 450 components in HeLa cell nuclei. Most are present in small amounts of <10,000 molecules per cell and are not detectable in cytoplasm.[112] Among the more acidic proteins are many enzymes including RNA polymerases. There are also gene repressors, hormone-binding proteins, protein kinases, and topoisomerases.[113] Among the six most abundant nonhistone nuclear proteins in the rat are the cytoskeletal proteins myosin, actin, tubulin, and tropomyosin.[114]

A group of small (<30 kDa) proteins, the **high mobility group (HMG)** proteins,[112,115] can be extracted from chromatin with 0.35 M NaCl. Two pairs, HMG-1 + HMG-2 R (renamed **HMGB**)[115a,b] and HMG-14 + HMG-17 (renamed **HMGN**),[115b,c] are present in nuclei of all mammals and birds. HMG-14

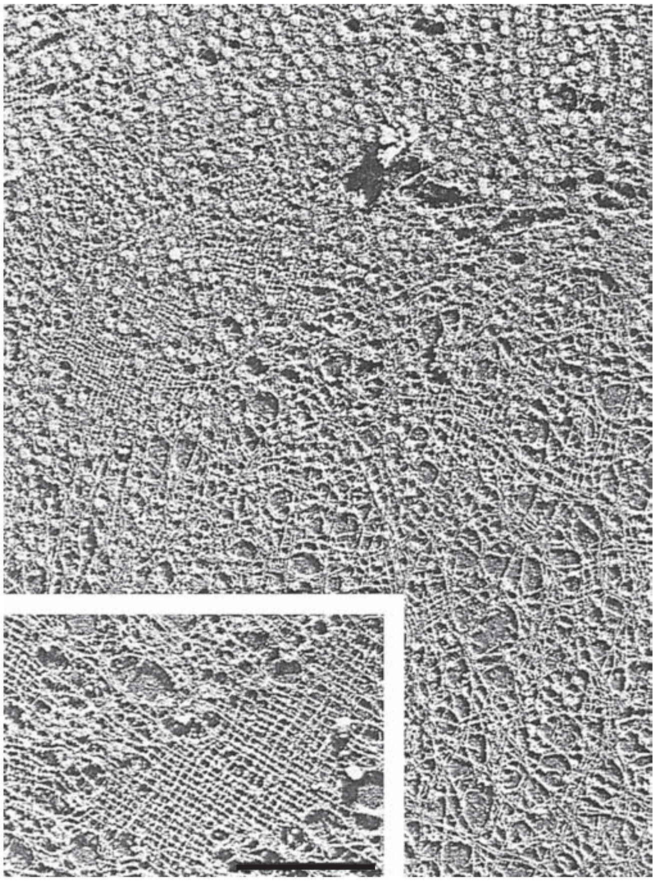


Figure 27-7 Native nuclear lamina of *Xenopus* oocytes. Freeze-dried metal-shadowed nuclear envelope extracted with Triton X-100, revealing the nuclear lamina meshwork partially covered with arrays of nuclear pore complexes. Inset, relatively well-preserved area of the meshwork of nearly orthogonal filaments from which pore complexes have been mechanically removed. Bar, 1 µm. From Aebi *et al.*[121]

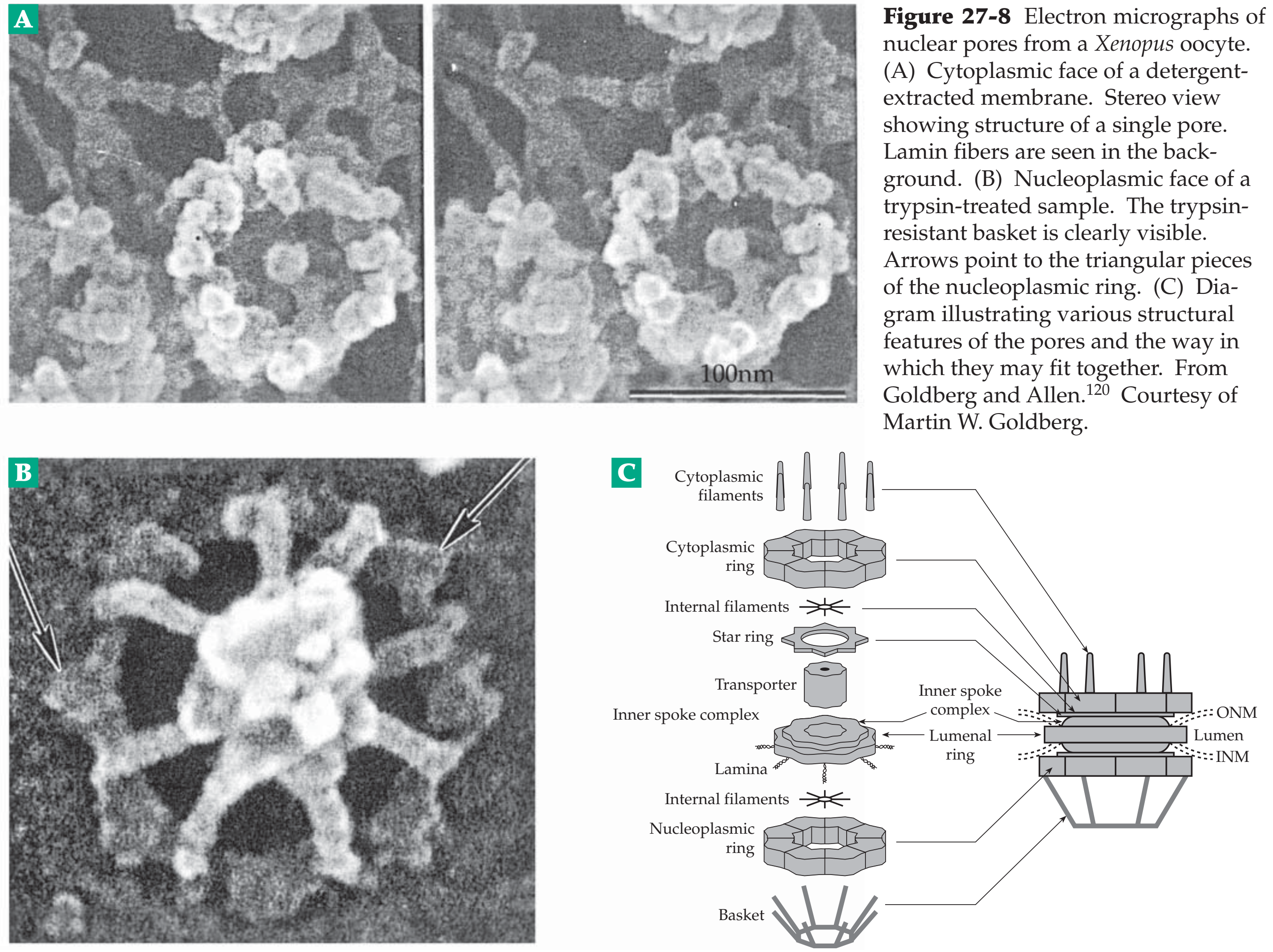


Figure 27-8 Electron micrographs of nuclear pores from a *Xenopus* oocyte. (A) Cytoplasmic face of a detergent-extracted membrane. Stereo view showing structure of a single pore. Lamin fibers are seen in the background. (B) Nucleoplasmic face of a trypsin-treated sample. The trypsin-resistant basket is clearly visible. Arrows point to the triangular pieces of the nucleoplasmic ring. (C) Diagram illustrating various structural features of the pores and the way in which they may fit together. From Goldberg and Allen.[120] Courtesy of Martin W. Goldberg.

and HMG-17 appear to be concentrated where active transcription is occurring. HMG-17 undergoes a variety of modifications including acetylation, methylation, phosphorylation, ubiquitination, glycosylation, and ADP-ribosylation, suggesting that it may assist in regulation of transcription. Other members of the HMG family bind more specifically to certain DNA sequences or structures.[115d,e]

Nuclear membrane, pores, and lamina. The nuclear membrane consists of two bilayer membranes separated by a 40- to 60-nm perinuclear space.[116,117] Both of the membranes and the intervening space are penetrated by large proteinaceous **nuclear pores**.[118–120b] The two membranes are fused together around the pores. On the inside of the innermost nuclear membrane is a relatively insoluble meshwork of intermediate filaments, the **nuclear lamina** (Fig. 27-7). It acts as a scaffold for the pores and may also interact with chromatin.[113,121–123a] Three proteins, the **lamins** A, B, and C, are the major components of the nuclear lamina. The A and C lamins, which differ only at their C termini, are homologous to cytoplasmic intermediate filament proteins and may exist as a network of coiled-coil polymers.[124] Lamins are phosphorylated, and the lamina is disassembled during the prophase of mitosis, apparently under the influence of the cdc2–cyclin B complex (Eq. 26-3),[125] and reappears in telophase at the end of mitosis.

Nuclear pores consist of octameric rings of protein subunits with a complex structure and an outside diameter of 120 nm and an inside diameter of ~80 nm (Fig. 27-8). A pore may consist of as many as 80 to 100 different proteins and have a mass of ~125 MDa.[119,126,126a,b] The pores are sometimes seen to be blocked by ~35 nm granules, perhaps pre-ribosomes. Transport through nuclear pores occurs in both directions. Numerous proteins enter the nucleus where many bind to RNA and are then exported as ribosomal subunits. Transfer RNAs and messenger RNAs must also be exported. A family of nuclear transporters known as **importins** and **exportins** mediate the movement of protein and RNA through the pores.[119,126b,127–130] They depend upon the G protein **Ran**, which also functions in spindle formation (p. 1503), and the hydrolysis of GTP.[131–132b] Large conformational changes may be linked to the transport cycle[133,134] and also have been observed in response to changes in the Ca^{2+} level.[135,136]

The nucleolus. This organelle is a dynamic structure, which breaks down and reforms in each mitotic cycle. It is organized around clusters of genes for the 28S, 18S, and 5.8S ribosomal RNA subunits. Both chromatin and ribonucleoproteins that transport the rRNA out to the cytoplasm are present.[95,137] The major nucleolar protein **nucleolin** is characterized by its glycine-rich C terminus, which contains seven repeats of GGRGG and also contains N^γ,N^γ-dimethylarginine.[138,139] Nucleolin may control transcription of the DNA that carries the rRNA genes. It appears to be required for ribosome synthesis (Chapter 29) and for attachment to the nuclear matrix. A variety of other proteins in smaller amounts are also needed in nucleoli. In addition to the true nucleoli other regions described as **speckles** and **coiled bodies** may develop around RNA splicing centers.[95,139a]

B. Organization of DNA

A vast amount of new information on DNA sequences and structures is available for yeast, *Arabidopsis*,[140] *Drosophila*, the nematode *Caenorhabditis elegans*,[141] human beings, and other species (see Chapter 26). Nevertheless, it may be worthwhile to consider older discoveries, some dating back 20 years or more.

When DNA is cut into ~10-kb fragments by shearing and the fragments are denatured by heat, the renaturation of the resulting single-stranded fragments upon cooling takes place in two or more steps. Some material reforms double helices rapidly, whereas other material is slow to renature (Fig. 5-46). At least four kinetically distinct fractions have been recognized: (1) About 70% of mammalian DNA appears to exist largely as single copies, i.e., with unique sequences. (2) About 20% of the total DNA is moderately repetitious, containing sequences that may be present 10^3–10^5 times. (3) About 5% reassociates very rapidly and is identified as highly repetitive **satellite DNA** of which there may be ~10^6 copies. (4) A smaller fraction contains long palindromic sequences. As a consequence, the single-stranded pieces can fold back almost instantaneously to form hairpin structures.[142,143] These differences also show up in density gradient centrifugation in CsCl, whether on large pieces of DNA[144] in the presence of ligands such as Ag^+ or on smaller restriction fragments which have often been studied by polyacrylamide gel electrophoresis.

If there are ~35,000 human genes, protein or RNA coding sequences must occupy only 2–3% of the genome.[145] The pufferfish *Fuga rubripes* probably has almost as many genes as we but only ~13% as much DNA. In further contrast the newt *Triturus cristatus* has six times as much DNA as a human.[146] The compact genome of the green plant *Arabidopsis thaliana* occupies only 120 Mb. In contrast are the 415-, 2500-, and 5300-Mb genomes of rice, maize, and barley, respectively.[147] What is the function of all of the apparently noncoding DNA present in some organisms? It is often viewed as "junk," whose only function may be to facilitate evolutionary changes in the genome. However, there is doubtless important undiscovered information in these regions.[145] It has been very hard to determine sequences, in part because of the large amount of highly repetitive DNA. Some regions have been "unclonable" in prokaryotic systems because of the presence of transposon-like sequences or "kinkable" elements (TG•CA steps), palindromes, etc.[145a,b]

1. Repetitive DNA

Rapidly renaturing DNA fragments often have a different base composition than the bulk of the DNA and, consequently, often separate as small satellite bands upon centrifugation in a CsCl gradient. Satellite DNA is usually associated with regions of the chromosome that do not unravel in telophase as does the bulk of the DNA. Satellite DNA usually consists of short highly repetitive sequences,[148] which occur in large clusters of up to 100 Mb of DNA, often near centromeres or telomeres or on a Y chromosome.[146] The DNA of a satellite band from the kangaroo rat contains the sequence 5'-GGACACAGCG-3' repeated so often that it accounts for 11% of the entire DNA of the cell. Longer repeating sequences ~170 bp are also often present as are **microsatellites** of 2- to 5-base-pair repeats. At least 30,000 microsatellite loci are present in the human genome.[149]

Centromeres. The attachment of spindle fibers to chromosomes depends upon the segments of DNA known as centromeres to organize the attached kinetochores. In the yeast *Saccharomyces cerevisiae* a 120-bp region containing three short conserved sequences is present in the centrosomes of all ten chromosomes.[150,151] This may fold into a distinctive looped structure.[152] Human centrosomes are large and complex,[153] but the DNA is highly repetitive, giving rise to **α-satellite DNA**.[154–157b] Sequences such as $(TGGAA)_n$ are repeated many times. Such sequences can form self-complementary looped structures containing some unpaired guanines that intercalate and stack between sheared G•A pairs.[158,158a] Complex regional centromeres involving kilobases or megabases of DNA have also been identified in fission yeasts, *Drosophila*, and green plants.[156,157] The great variability indicates that centromeric sequences undergo rapid evolution. This may be related to the fact that of the four cells formed by female meiosis only one becomes an egg.[158b] A series of unique **centromeric proteins** (CENP-A to CENP-G) bind to the DNA sequences of centromeres[159–162a] and direct the formation of the kinetochores. Even for the simpler centromere of budding yeasts, kinetochores have a complex structure.[150] The CENP proteins were first identified as autoantigens in sera of patients with the autoimmune disorder **scleroderma** (Chapter 31).[160,162]

Telomeres. The DNA sequences at the chromosome ends have a TG-rich strand, such as the $(TTGGGG)_{50-70}$ (*Tetrahymena*)[163] and the $(TTAGGG)_n$ of both human and trypanosome chromosomes.[164–166] The complementary DNA strand is CA-rich. The *S. cerevisiae* telomers have ~350 base pairs containing the sequences $(TG_{1-3} / C_{1-3}A)$ as well as one or more copies of a 6.7-kb nonrepetitive sequence and other elements.[166,167] In many species the repetitive telomeric sequences have 3' poly(G) tails at the ends of the DNA molecules. These tails are able to form G quartet structures (Fig. 5-26 and Chapter 5, Section C,4). A variety of telomere-binding proteins have been isolated.[167a–d] Some of these bind to G quartet structures[168–169a] and some, such as **RAP1** of yeast,[170–171a] to double-stranded telomere repeat regions. Special problems associated with replication of telomeres are discussed in Section C,8.

Short interspersed sequences (SINES). Much of the reiterated DNA is present in repeated segments 100–500 bp in length that lie between 1- to 2-kb segments of unique DNA.[172–174] The best known example is the human ***Alu*** family,[175–177b] so-called because it contains a site for cleavage by restriction enzyme *Alu*I. Sequences of the *Alu* family also exists in other primates and in rodents. The ~300-bp *Alu* sequence is reiterated over 500,000 times in the human genome with various sequence alterations, but an 80–90% homology. This sequence (Fig. 27-9) consists of two similar ~130-bp segments called the "left monomer" and "right monomer." The right monomer contains a 31-bp insertion and the left end carries a poly(dA) sequence. In addition there is a short 7- to 20-bp sequence, which is variable between different *Alu* sequences but is directly repeated at each end of a given *Alu* sequence.

The *Alu* sequence has strong homology with the 7S RNA (Table 5-4) that is part of the signal recognition particle involved in transport of newly synthesized peptide chains across the membranes of the ER (Chapter 10). *Alu* sequences are transcribed into hnRNA, the precursors to mRNA. Some *Alu* sequences are present in intervening sequences (introns) within genes and others are in noncoding sequences between genes. Sharp[176] suggested that specific proteins in the nucleus may bind to the *Alu* sequences, preventing hnRNA from leaving the nucleus before it has been processed to remove introns and other sequences absent in mRNA. However, the presence of the poly(dA) regions and the direct terminal repeats suggests that the whole *Alu* sequence is **pseudogene** derived from a **retroposon**, a type of transposon that originated from an RNA molecule. *Alu* and other SINES contain an RNA polymerase III promoter and are transcribed. Active retroposons form reverse transcripts (cDNAs) that can be integrated at various points in the genome (see Section D,4 and Chapter 28). One theory is that retroposons have no biological function but have invaded the genome at random locations.

Many SINES and other families of repetitive sequences have been characterized by the presence of a restriction enzyme cleavage site in each copy of the sequence. *Eco*RI cuts the previously mentioned α-satellite DNA.[177] A 319-bp reiterated sequence in the human genome surrounds a *Hin*f site,[178] and *Sau*3A cleaves a 849-bp sequence with ~1000 copies per haploid genome.[179] In the hermit crab 30% of the genome consists of repeated sequences, one 156-bp unit occuring ~7 million times. Many identical 14-bp GC-rich inverted repeats are present.[180] *Neurospora* and yeast both contain many copies of the GC-rich palindromic sequence 5'-CC$\underline{\text{CTGCAG}}TA\underline{\text{CTGCAG}}$GG-3', which contains the two underlined *Pst*I sites.[181] Some repetitive sequences such as the **CAT** family and the **homeodomain sequence** (Chapter 28, Section C,6) are found in the control regions preceding the 5' ends of groups of genes that are regulated coordinately.[182] Some repetitive sequences seem to be unstable in the genome and may be excised and lost or may increase in number during aging.[183] Repeated DNA

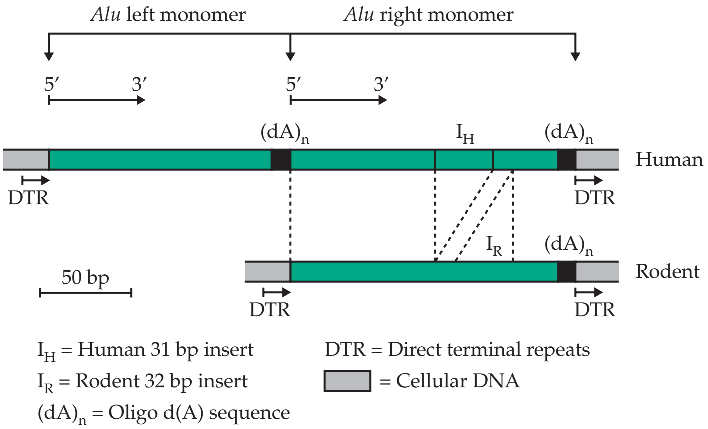


Figure 27-9 Structure of human and rodent *Alu* sequences in DNA. From Ullu.[175]

sequences have apparently originated not only with 7S RNA but also from mRNA or tRNA molecules.[172] A characteristic SINE in rodent DNA is known as the ID (identifier) sequence because it was once thought to be a marker for genes transcribed in neural tissues. Rat DNA contains ~130,000 copies of the ID sequence.[172a] Other **clustered repeats** are frequently found between genes that are present in large numbers. These include the genes for ribosomal RNA, tRNA, small nuclear RNAs, and histones. Triplet repeats are considered in Chapter 26, Section G,3.

Long interspersed repeat sequences (LINES). These moderately repetitive sequences may be several kb in length. Just one type seems to be abundant in each mammalian species.[122] The human L1 sequence is over 6 kb in length. Like the *Alu* sequence it has a poly(dA) sequence at the 3'-end, but it does not contain short terminal direct repeats. These interspersed repeat sequences usually contain genes which may be functional, but many of the copies contain pseudogenes or genes that are randomly truncated, having lost a segment from the 5' end. These sequences, too, seem to have been dispersed throughout the genome by retroposons.[122,184,185] Human chromosome 22, sequenced in 1999, contains within its 33.4 Mb of euchromatin at least 545 genes, 134 pseudogenes, and 8043 L1 sequences (9.7% of the DNA), as well as 20,188 *Alu* sequences (16.8% of the DNA) and many other interspersed repeats.[186] Repetitive DNA sequences are less common in prokaryotes, but they do exist. For example, many dispersed and clustered repeating units are present in *Halobacterium* DNA.[187]

2. Genes for Ribosomal RNA and Small RNA Molecules

Most genes are present as one copy each per haploid genome. However, there are many copies of the genes for ribosomal RNA and tRNA. In *Xenopus* DNA there are ~450 repeats of the 28 S and 18 S rRNA genes on one chromosome and ~24,000 copies of the 5 S RNA genes at the ends of the long arms of most of the chromosomes.[188] Nontranscribed spacer regions lie between the repeats of the 28S and 18S gene pairs, as can be seen with the electron microscope in Fig. 28-17. The gene for 5 S RNA has a high GC content. A denaturation map of the DNA shows easily melted regions separated by shorter 120-bp sequences, apparently of high GC content and presumably coding of the 5 S RNA. The easily denatured AT-rich spacers are ~630 bases long. Using restriction enzymes much of this DNA can be cut into segments that contain repeats within repeats. One 15-unit polynucleotide contains the sequence $A_4CUCA_3CU_3G$ repeated about 30 times.[189]

About 200 copies of rRNA genes per haploid genome are located at the constrictions in the short arms of human chromosomes 13, 14, 15, 21, and 22 (Chapter 26, banner). As in *Xenopus,* clusters of tandemly repeated 5S RNA genes are found at the ends of the long arms of most chromosomes. A similar organization of rRNA genes is found in the rat.[190] The ten different chromosome ends carrying rRNA genes in the human diploid nucleus come together to form the nucleolus, the site of synthesis of ribosomal subunits.[191] At first there are ten small nucleoli, but these fuse to form the single highly structured but membraneless nucleolus. The rRNA genes of the macronucleus of *Tetrahymena* have been “amplified” and are found on linear 21-kb palindromic molecules, about 10^4 copies being present per cell. Initiation of transcription begins near the center and proceeds outward in both directions.[192] These short chromosomes contain typical telomeric ends.

Genes for the tRNAs are spread throughout the genome of bacteria, mitochondria (Fig. 18-3), chloroplasts,[193] and eukaryotic nuclei. They sometimes occur in clusters but more often are far apart. In *Drosophila* there are probably at least 600 tRNA genes, many occurring in pairs of opposite polarity, i.e., as inverted repeat sequences. The genes for small nuclear RNAs U1–U6 (Chapter 28) are organized in a variety of ways. The ~100 human U1 genes occur on a single chromosome, perhaps organized in a tandem array and interspersed with as many as 10,000 defective pseudogenes.[194]

3. Other Gene Clusters and Pseudogenes

Closely related structural genes often occur in clusters.[195] Among these are clusters of **immunoglobulin** genes (Fig. 31-17) and clustered genes for the α and β **globins** (Fig. 27-10), which encode the protein sequences for the hemoglobins. The human α globin gene cluster occupies about 30 kb on chromosome 16 and the β globin genes 60 kb on chromosome 11.[196–197d] The α1 and α2 genes (Fig. 27-10) encode identical peptides, while the related ζ gene encodes the corresponding subunit of embryonic hemoglobin. The β cluster includes, in addition to the adult gene, a pair of fetal globin genes (γ^G and γ^A) differing by only one amino acid (Gly vs Ala) at position 136, the embryonic ε chain, and the minor adult δ chain.

Besides the functional genes the globin cluster contains **pseudogenes**, which are given the prefix ψ in Fig. 27-10. These are nonfunctional genes, which appear to encode peptides homologous to the known globins. However, they contain mutations that prevent expression. For example, deletion of a single nucleotide near the beginning of the pseudogene will scramble the genetic message by changing the reading

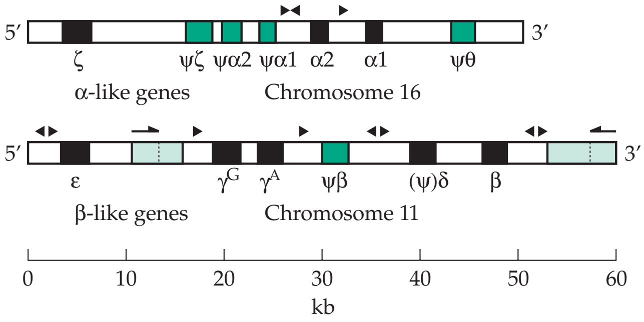


Figure 27-10 Organization of the globin genes on human chromosomes 11 and 16. The composition of the various embryonic, fetal, and adult hemoglobins is also indicated. Closed boxes indicate active genes and open boxes pseudogenes. The triangles (▸) indicate *Alu* repetitive sequences and their orientation. The shaded boxes indicate Kpn repeat sequences and the half-arrows their respective orientation. The Kpn sequence between the ε and γ^G genes in fact consists of two tandemly linked Kpn repeats. From Karlsson and Nienhuis[196] and Proudfoot.[197]

frame. Globin pseudogene ζψ has a nonsense mutation at codon 6, ψα1 contains a whole set of mutations that prevent transcription and translation, while ψα2 is mutated almost to the point of being unrecognizable as a globin relative.[198]

Do pseudogenes have a physiological function? The presence of pseudogenes in close proximity to functional genes suggests that there may be an as yet unrecognized controlling function.[198] On the other hand, they may be relics of evolution. If a gene family arises by duplication and mutation of duplicated copies, we may expect to find gene copies that have not been selected as useful but are still present in the genome.

There are regions of the genome that seem to be largely single copy DNA with few reiterated sequences or gene clusters.[144] While some gene families exist as clusters others are dispersed throughout the genome.

4. Introns, Exons, and Overlapping Genes

A major difference between prokaryotes and eukaryotes is the presence of intervening sequences (introns) between the coding sequences (exons) in eukaryotes. Introns are especially numerous in higher organisms. For example, the gene for the myosin heavy chain present in rat embryonic skeletal muscle[199] encodes a 1939-residue peptide but occupies a length of 22 kb of DNA. The gene is split into 41 exons, whose transcribed RNA must be cut and spliced at 40 places to form the mRNA. By comparison the corresponding gene from the nematode *Caenorhabdatis elegans* is far less fragmented. It is not clear why some genes have so many introns and others so few. The very compact chromosomes of viruses and most bacteria do not contain introns. However, they are sometimes present in mitochondrial and chloroplast genes.

A surprise, which was first recognized in viral RNA and DNA, is that genes sometimes overlap. For example, two proteins, one long and one shorter, are synthesized starting at the same point in the RNA genome of phage Qβ.[200] In the DNA of phage ϕX174 the third nucleotide of the stop signals for some genes are also the first nucleotides for the start signal for translation of the next gene. Pairs of genes have been found in which one of the genes of the pair is found completely inside the other gene but is translated in a different reading frame.

A four-base overlap between dihydrofolate reductase and thymidylate synthase has been found[201] in the DNA of phage T4. A transposable DNA insertion sequence (see Section D,5) in *E. coli* encodes two genes, one of which is contained within the other and which is transcribed from the opposite strand of DNA.[202] The double-stranded RNA of a reovirus produces two peptides from the same sequence using two different AUG initiation codons in different reading frames.[203]

Overlapping obviously limits severely the mutational alterations that are allowable. Perhaps it is for this reason that eukaryotic genes seem to be dispersed in widely separated locations and overlap is rare. However, the human type IV collagen genes for the α1 and α2 chains are encoded on opposite strands with their 5' flanking regions overlapping.[204] Some introns contain genes.[205,206] Chlorarachniophyte algae contain a multimembraned chloroplast thought to be a vestigial remnant of an endosymbiont. Its 380-kb genome consists of three short chromosomes that encode overlapping genes and contain the shortest introns (18–20 base pairs) known.[207]

5. DNA of Organelles

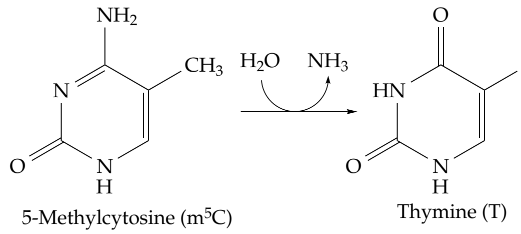
Mitochondrial DNA (mtDNA), discussed in Chapter 18, varies in size from <6 kb to 367 kb.[208] The small circular mtDNA of animals is extremely efficiently packed with genes for tRNAs, rRNA, and a small number of protein subunits (Fig. 18-3). However, the 78-kb yeast mtDNA contains many long AT-rich spacers as well as long introns, some of which contain genes for splicing enzymes. Otherwise, it is similar to animal mitochondrial DNA. Mitochondrial genomes of higher plants are much larger; that of maize is a 570-kb circle, which contains both direct and inverted repeat sequences. Recombination between these sequences is apparently the origin of smaller incomplete circular genomes that are also present. Plant mtDNA appears to contain the genes present in animal mitochondria plus additional genes. For example, one encodes a 5S RNA. The most unusual mtDNA is found in hemoflagellates such as *Trypanosoma* and *Leishmania*. The single large mitochondrion or **kinetoplast** located near the base of the flagellum contains a network of catenated circular DNA molecules (Fig. 5-16).[209,210] The genetic makeup is similar to that of other mtDNA, but the rRNA genes are unusually short. Chloroplast DNA, which varies in size from 120 to 160 kb, is discussed in Chapter 23.

6. Methylation of DNA

As mentioned in Chapter 5, a significant fraction of the pyrimidine and purine bases in DNA is methylated. One function of such methylation in bacteria, discussed in Chapter 26, is to protect against the action of restriction endonucleases.[211] For example, the gene for the well-known *Eco*RI endonuclease is carried in *E. coli* cells by an R factor. This plasmid also carries (just 29 base pairs away) the gene for a 326-residue **N^6-adenine methylase**.[212–213a] This enzyme uses *S*-adenosylmethionine (AdoMet) to methylate the two adenines (marked by asterisks) in the six-base-pair recognition sequence 5'-G A A* T T C / 3'-C T T A* AG converting them to N^6-methyladenines (m^6A). Other DNA methyltransferases place methyl groups on N-4 of cytosine or on C-5 of cytosine.[214–215a] The latter utilizes a mechanism illustrated in Eq. 12-4.[216,217] Such enzymes are components of both type II and type I restriction–methylation systems.[218] However, most of the m^6A in the *E. coli* chromosome arises from action of a different methylase, one that recognizes the palindromic sequence 5'-GATC and methylates adenines in both chains.[219] This **DNA adenine methylase**, a product of the *dam* gene, plays an important role in mismatch repair, transposition, regulation of transcription, and initiation of DNA replication.[220] The same methylase regulates at least 20 genes induced during infection by *Salmonella typhimurium*. Some of these genes are essential to virulence.[221] A similar methyltransferase appears to control differentiation of the stalked *Caulobacteria* cells.[220]

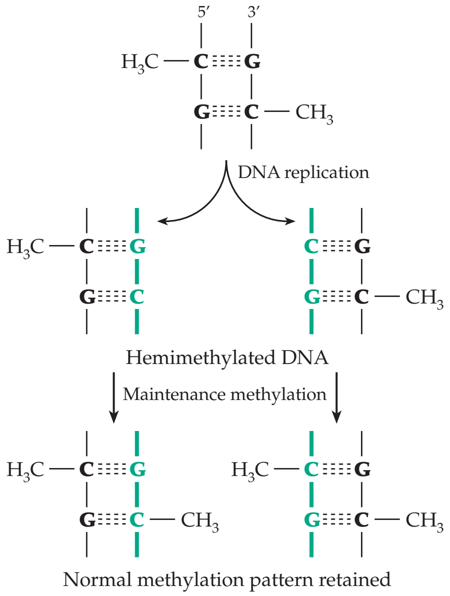
CG doublets. The only modified base commonly found in eukaryotes is 5-methylcytosine,[222,223] which upon deamination becomes thymine (Eq. 27-1). Most methylation occurs when C is followed by G. Usually 60–90% of all 5'-CG sequences (CpG sequences) in eukaryotic DNA are methylated. However, the fraction of methylated cytosine varies from almost zero for *Drosophilia*, *Caenorhabditis*, and *Saccharomyces* to as much as 30% in higher plants.[224]

When CG pairs are methylated on both cytosines, an interesting result arises upon replication of the DNA. The methyl groups don't prevent replication, but only one of the DNA strands in each daughter duplex is methylated (Eq. 27-2). However, additional methylation of DNA occurs within three hours of DNA replication by a **maintenance methyltransferase** that recognizes methylated CpG sequences in the old DNA strand and methylates the cytosine in the 3′-GpC of the newly synthesized strand.[220,225,226] Mammalian cells contain at least three different DNA (cytosine-5) methyltransferases.[227–229b] Enzymatic **demethylation**, which converts m^5C residues back to cytosine, may also occur.[230–233] It has been difficult to demonstrate[233a] but enzymatic demethylation of 5-methylcytosine has been reported.[233b,c] The observed loss of methylation during embryonic development (Chapter 32) may be a result of loss of the maintenance methylase. Most CpG doublets are found in large "islands" of several hundred bases to about 2 kb in length, which are unusually rich in G + C.[234,235] These islands lie near the 5' ends of ~60% of all human genes[236] and near origins of replication.[237] They are also found in regions of the DNA that are compacted into heterochromatin. The genes in methylated regions tend to be "silent," i.e., they are not actively transcribed.[236] Demethylation may permit transcription.



(27-1)

The degree of methylation of the CG doublets is variable both in position within a chromosome and



(27-2)

with stage of development. Cytosine methylation is essential to embryonic development,[238] and mice lacking the maintenance DNA methytransferase are developmentally retarded and die at mid-gestation.[239] Not all CG sequences are methylated and various patterns of DNA methylation are generated at different stages of development by rounds of methylation and by the action of demethylases.[232] The maintenance methylase (Eq. 27-2) ensures that a stable methylation pattern persists until altered by new rounds of methylase or demethylase action. A parallel is found in the prokaryotic *Caulobacter* in which three chromosomal sites are fully methylated in swarmer cells, become hemimethylated in stalked cells, and are fully methylated again just prior to cell division.[220] See Fig. 32-1.

DNA methylation affects transcription, either directly by preventing the binding of transcription factors or indirectly via a series of binding proteins specific for methylated CG doublets. In early stages of embryonic development there is very little methylation, but some genes are quickly silenced as methylation takes place. Heterochromatic regions including inactivated X chromosomes are heavily methylated. However, additional alteration in chromatin is required for complete silencing of genes.[236,240] Recent studies indicate that an abundant mammalian protein binds to the methylated DNA along with a **histone deacetylase**.[239,241–244] The latter acts on acetylated histones to free lysine side chains, which may interact in an inhibitory manner with the DNA.

In female mammalian cells most of the genes on one of the two X-chromosomes are completely inactivated. DNA methylation plays a major role in this process.[244,245] A perfect correlation has been observed between 5'-methylation of cytosines in CpG islands and inactivation of X-chromosome genes.[246] Methylation may also play a role in recombination and repair.[247] Methylation of DNA decreases with increasing age.[248] It increases as a result of oncogenic transformation of cells.[249] Some other modifications of DNA largely limited to bacteriophages are discussed on p. 234.[247,250]

Imprinting. With the exception of X-linked genes each person has two copies of each gene, one of maternal and one of paternal origin. Both copies of most of these genes are expressed. However, a few of the genes receive from one parent or the other an **imprint**, a mark that distinguishes the parental origin.[251–253] Such imprints are maintained in cells through embryonic development but are erased in embryonic gonads to allow for a new imprint in the germ cells. Imprinting depends upon DNA methylation, and all imprinted genes show the presence of differentially methylated regions.[253] See also Chapter 32, Section A,1.

C. Replication

Following the discovery of the double helix and the enthusiasm that it engendered many people thought that the synthesis of DNA was simple. The nucleotide precursors would align themselves along separated DNA template strands and perhaps spontaneously snap onto the growing chains. In fact, replication is a complex process that requires the cooperative action of many different gene products and perhaps an association with membrane sites. The matter is made more complex by the fact that some of the enzymes involved in replication are also required in the processes of genetic recombination, in repair of damaged DNA molecules, and in defensive systems of cells.

1. Early Studies

That the DNA content doubles prior to cell division was established by microspectrophotometry. It was clear that both daughter cells must receive one or more identical molecules of DNA. However, it was not known whether the original double-stranded DNA molecule was copied in such a way that an entirely new double-stranded DNA was formed or whether, as we now know to be the case, the two chains of the original molecule separated. The latter is called **semiconservative** replication, each of the separated strands having a new complementary strand synthesized along it to form the two identical double-stranded molecules.

The first definitive evidence for semiconservative replication was reported by Meselson and Stahl[254] in 1958. Cells of *E. coli* were grown on a medium containing isotopically pure $^{15}NH_4^+$ ions as the sole source of nitrogen. After a few generations of growth in this medium the DNA contained exclusively ^{15}N. Then the cells were transferred abruptly to a medium containing $^{14}NH_4^+$ and were allowed to grow and to double and quadruple in number. At various stages DNA was isolated and subjected to ultracentrifugation in a CsCl gradient. Small but easily detectable differences in density led to separation of dsDNA molecules containing only ^{15}N from those containing partly ^{15}N and from those containing only ^{14}N. At the beginning of the experiment only DNA containing entirely ^{15}N was present. However, after one generation of growth in the ^{14}N-containing medium, the density of *all* the DNA was such as to indicate a content of one-half ^{14}N and one-half ^{14}N. After a second generation of growth half of the DNA still contained both nitrogen isotopes in equal quantity, whereas half contained only ^{14}N, exactly the result expected for semiconservative replication. A similar experiment using 5-bromodeoxyuridine, a thymidine analog, is shown in Fig. 27-11.

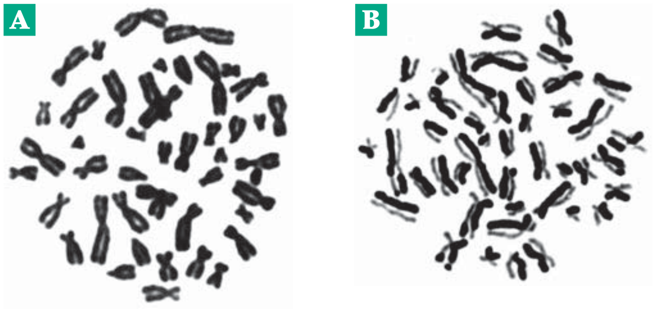
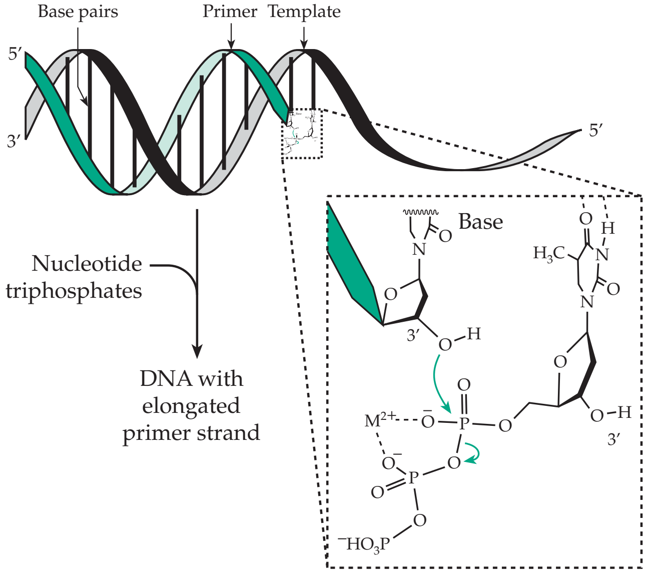


Figure 27-11 (A) Human chromosomes after one replication in the presence of 5-bromodeoxyuridine (BrdU). Both chromatids of each chromosome contain BrdU in one strand of the DNA duplex and normal thymidine in the other. (B) After two replications in the presence of BrdU one chromatid of each chromosome contains BrdU in both strands of the duplex and stains strongly with a special differential staining procedure. The other chromatid contains only normal thymidine in one strand of the duplex and is not stained. Courtesy of Carolina Biologicals.

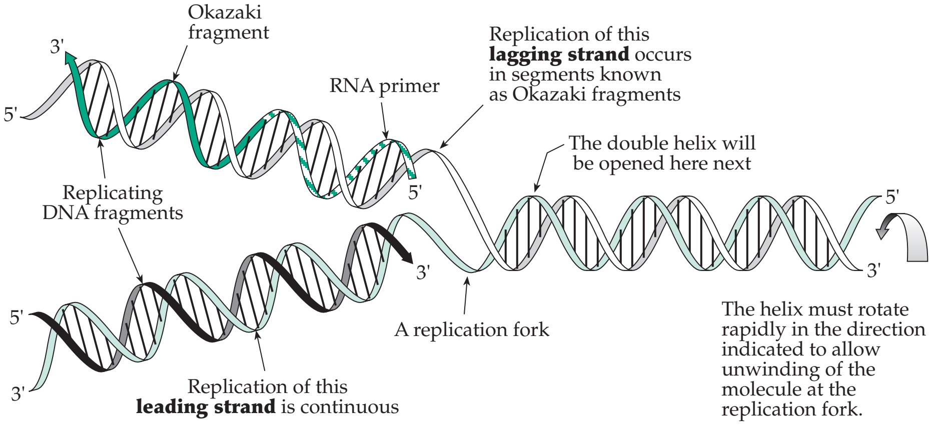
Autoradiography. Later a technique of direct autoradiography of DNA using ^{3}H-labeled thymidine[255] was applied by Cairns.[256,257] Cells of *E. coli* were grown on a medium containing the radioactive thymidine for various times but typically for 1 h (~2 generations). The cells were then ruptured, the DNA was spread on a thin membrane filter, and autoradiograms were prepared. When the DNA contained [^{3}H]thymidine the exposed trace in the autoradiogram could be followed around the entire 1.1–1.4 mm circumference of the spread DNA molecule. Molecules partially labeled and in the process of replication could also be identified. After two hours of growth in the presence of ^{3}H-labeled thymidine about half of the bacterial DNA was fully labeled, but half contained regions that were labeled only half as heavily. They presumably contained ^{3}H label in a single strand and, therefore, represented unreplicated regions. All of the molecules had undergone one round of replication with tritiated thymidine to yield lightly labeled molecules; parts of the molecules had not completed the second round. The more heavily labeled regions were interpreted as fully replicated. The shapes of the "replication figures" suggested that DNA is synthesized in a continuous manner starting from one point and continuing around the circular molecule at a constant rate. Although subsequent experiments (considered in Section 4) show that replication is usually **bidirectional**, the experiments of Cairns were important because they introduced a technique for direct visualization of replication *in vivo*.



(27-3)

The chemistry of DNA polymerization. What are precursors of DNA? Early experiments showed that the nucleoside [^{3}H]thymidine was efficiently incorporated into DNA, but for energetic reasons it seemed unlikely that thymidine was an immediate precursor. Evidence favoring the nucleoside triphosphates was provided in 1958 when Arthur Kornberg identified a DNA polymerase from *E. coli*. Kornberg's enzyme, now known as **DNA polymerase I**, was isolated in the amount of 600 mg from 90 kg of bacterial cells[258,259] (over 400 molecules of enzyme per cell). The 928-residue enzyme displayed many of the properties expected of a DNA-synthesizing enzyme. It requires a **template strand** of DNA as well as a shorter **primer strand**. As indicated in Eq. 27-3, the enzyme recognizes the 3' end of the primer strand and binds the proper nucleoside triphosphate to pair with the next base in the template strand. Then it catalyzes the displacement of pyrophosphate, at the same time linking the new nucleotide unit onto the 3' end of the primer strand. Continuing in this way, the enzyme is able to turn a single-stranded template DNA into a double-stranded DNA in which the newly synthesized strand contains, at each point, the base complementary to the one in the template strand.[259a]

Although the action of the DNA polymerase I, according to Eq. 27-3, provided a straightforward way to form a complementary strand of DNA, it did not explain how double-stranded DNA could be copied. One problem is that the two strands must be separated and unwound. If unwinding and replication occured at a single **replication fork** in the DNA, as indicated by Cairns' experiment, the entire molecule would have to spin at a speed of 300 revolutions per second to permit replication of the *E. coli* chromosome in 20 min. It also required that some kind of a **swivel**, or at least a nick in one chain, be present in the chromosome as indicated in the following drawing.


Okazaki fragment
3'
5'
RNA primer
Replication of this **lagging strand** occurs in segments known as Okazaki fragments
The double helix will be opened here next
5'
5'
3'
Replicating DNA fragments
5'
3'
3'
A replication fork
Replication of this **leading strand** is continuous
The helix must rotate rapidly in the direction indicated to allow unwinding of the molecule at the replication fork.

Another problem was posed by the fact that the two chains in DNA have opposite orientations. Thus, at the replication fork one of the new chains might be expected to grow by addition of a new nucleotide at the 3' end, while the other chain would grow at the 5' end. *If so, there should be two DNA polymerases*, one specific for polymerization at each of the two kinds of chain end. Nevertheless, despite intensive search the only DNA polymerases found added new units only at the 3' end.

Discontinuous replication and RNA primers. In 1968, Okazaki reported that during the time that replication of DNA is taking place bacterial cells contain short fragments of DNA. These are now called **Okazaki fragments** or **replication fragments**.[260] A second development was the discovery of the enzyme **DNA ligase**,[261,262] which is able to join two pieces of DNA to form a continuous chain. These two discoveries provided an explanation for the lack of a second kind of DNA polymerase. One strand, the **leading strand**, of the replicating DNA could be synthesized continuously in the 5' to 3' direction while the other strand, the **lagging strand**, would have to be synthesized in segments (the Okazaki fragments), which could then be joined by the DNA ligase. In 1971 Brutlag *et al.* reported that initiation of synthesis of the DNA of phage M13 in *E. coli* required formation of a short segment of RNA as a primer.[263] It was subsequently shown that *with few exceptions priming by RNA synthesis is always required for replication*.

Newer studies have revealed great complexity in the mechanisms of replication. As for synthesis of any polymer there are distinct steps of *initiation*, *elongation*, and *termination*. Topological problems associated with the unwinding and rewinding of the double helices and with disconnection of catenated circles and untying of knots are solved with the aid of special enzymes, the **helicases** and **topoisomerases**. Replication requires both DNA and RNA polymerases, a ligase, and ancillary proteins, some of whose functions aren't yet clear. Many of these proteins associate to form large multiprotein complexes, which are given names such as **primosome** (for priming) and **replisome**. Many bacteriophages and plasmids also replicate within cells of *E. coli* utilizing bacterial proteins as well as proteins encoded in the viral or plasmid genome.

2. DNA Polymerases

Immediately after its identification DNA polymerase I was generally assumed to be the principal chain elongation enzyme. However, an *amber* mutant of *E. coli* deficient in DNA polymerase I (gene *pol*A; Fig. 26-4A) synthesized DNA normally. This finding stimulated an intensive search for new polymerases. Two were found: DNA polymerases II (gene *pol B*)[264] and III. Both are present in amounts less than 25% of that of DNA polymerase I.[265,266] Both have properties similar to those of polymerase I, but there are important differences. By now DNA polymerases have been isolated from many organisms, many genes have been cloned and many sequences, both of bacterial and eukaryotic polymerases are known. Comparisons of both sequences and three-dimensional structures,[266a,b] a few of which are shown in Fig. 27-12, suggest that the polymerases belong to at least six families (Table 27-1). These include the RNA-dependent DNA polymerases known as **reverse transcriptases** as well as some **RNA polymerases**.[267–268b]

Some of the polymerases exist as single polypeptide chains, while others function only as large complexes. In every case a two-metal ion catalytic mechanism with in-line nucleotidyl transfer,[269] illustrated in Fig. 27-13, appears to be used by the enzymes.[267,270] Two-metal ion catalysis is also observed for phosphatases and ribozymes (Chapter 12).

Exonuclease activities, proofreading, and editing. DNA polymerase I not only catalyzes the growth of DNA chains at the 3' end of a primer strand but also, at about a 10-fold slower rate, the hydrolytic removal of nucleotides from the 3' end (**3'- 5' exonuclease activity**). The same enzyme also catalyzes hydrolytic removal of nucleotides from the 5' end of DNA chains. This latter **5'- 3' exonuclease activity**, the DNA polymerase activity, and the 3'-5' exonuclease activity all arise from separate active sites in the protein. DNA polymerases II and III do not catalyze

TABLE 27-1
Families of DNA Polymerases[a,]

Class	Name	Function	Molecular mass (kDa)
A	*E. coli* polymerase I (Pol I)	DNA excision repair	103
	Klenow fragment		68
	Bacillus subtilis Pol I[b]	DNA excision repair	
	Thermus aquaticus DNA polymerase (Taq)[c,d]	DNA excision repair	
	T7 DNA polymerase[e]	Virus replication	80
	T7 RNA polymerase[f,g]		99
	Eukaryotic Pol γ (gamma)	Mitochondrial replication	
	Eukaryotic Pol θ (theta)	DNA repair	
B	Eukaryotic Pol α (alpha)	DNA replication	180 (core)
	Eukaryotic Pol δ (delta)	DNA replication	
	Eukaryotic Pol ε (epsilon)	DNA replication	
	Eukaryotic Pol ζ (zeta)	Bypass synthesis	
	Bacteriophage T4 DNA Pol + accessory proteins[h,i]	DNA replication	43
	E. coli Pol II[j]		90
C	Bacterial DNA Pol III[p] + accessory proteins	DNA replication	~900 (holoenzyme)
D	Euryarchaeotic Pol II		
X	Eukaryotic DNA Pol β[n,o]	DNA repair	39 x 2
	Eukaryotic Pol λ (lambda)	Base excision repair	
	Eukaryotic Pol μ (mu)	Non-homologous end-joining	
	Eukaryotic Pol σ (sigma)	Sister chromatid cohesion	
Y	*E. coli* UmuC protein	SOS response	
	Eukaryotic Pol η (eta, XP-V, RAD 30)	Bypass synthesis	
	Eukaryotic Pol ι (iota)	Bypass synthesis	
	Eukarytoic Pol κ (kappa)	Bypass synthesis	
Reverse transcriptase family			
	HIV reverse transcriptase[k,l]		
	Telomerase[m]		
	RNA-dependent RNA polymerases		

---

[a] Based on reviews by Burgers, P. M. J. *et al.* (2001) *J. Biol. Chem.* **276**, 43487–43490 and Steitz, T. A. (1999) *J. Biol. Chem.* **274**, 17395–17398

[b] Kiefer, J. R., Mao, C., Braman, J. C., and Beese, L. S. (1998) *Nature (London)* **391**, 304–307

[c] Eom, S. H., Wang, J., and Steitz, T. A. (1996) *Nature (London)* **382**, 278–281

[d] Li, Y., Korolev, S., and Waksman, G. (1998) *EMBO J.* **17**, 7514–7525

[e] Doublié, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) *Nature (London)* **391**, 251–258

[f] Sousa, R., Chung, Y. J., Rose, J. P., and Wang, B.-C. (1993) *Nature (London)* **364**, 593–599

[g] Sastry, S., and Ross, B. M. (1999) *Biochemistry* **38**, 4972–4981

[h] Wang, J., Yu, P., Lin, T. C., Konigsberg, W. H., and Steitz, T. A. (1996) *Biochemistry* **35**, 8110–8119

[i] Jing, D. H., Dong, F., Latham, G. J., and von Hippel, P. H. (1999) *J. Biol. Chem.* **274**, 27287–27298

[j] Cai, H., Yu, H., McEntee, K., Kunkel, T. A., and Goodman, M. F. (1995) *J. Biol. Chem.* **270**, 15327–15335

[k] Morris, M. C., Berducou, C., Mery, J., Heitz, F., and Divita, G. (1999) *Biochemistry* **38**, 15097–15103

[l] Rodgers, D. W., Gamblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C., and Harrison, S. C. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1222–1226

[m] Lundblad, V. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8415–8416

[n] Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science* **264**, 1930–1935

[o] Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. (1997) *Biochemistry* **36**, 11205–11215

[p] Kornberg, A., and Baker, T. A. (1992) *DNA Replication*, 2nd ed., Freeman, New York

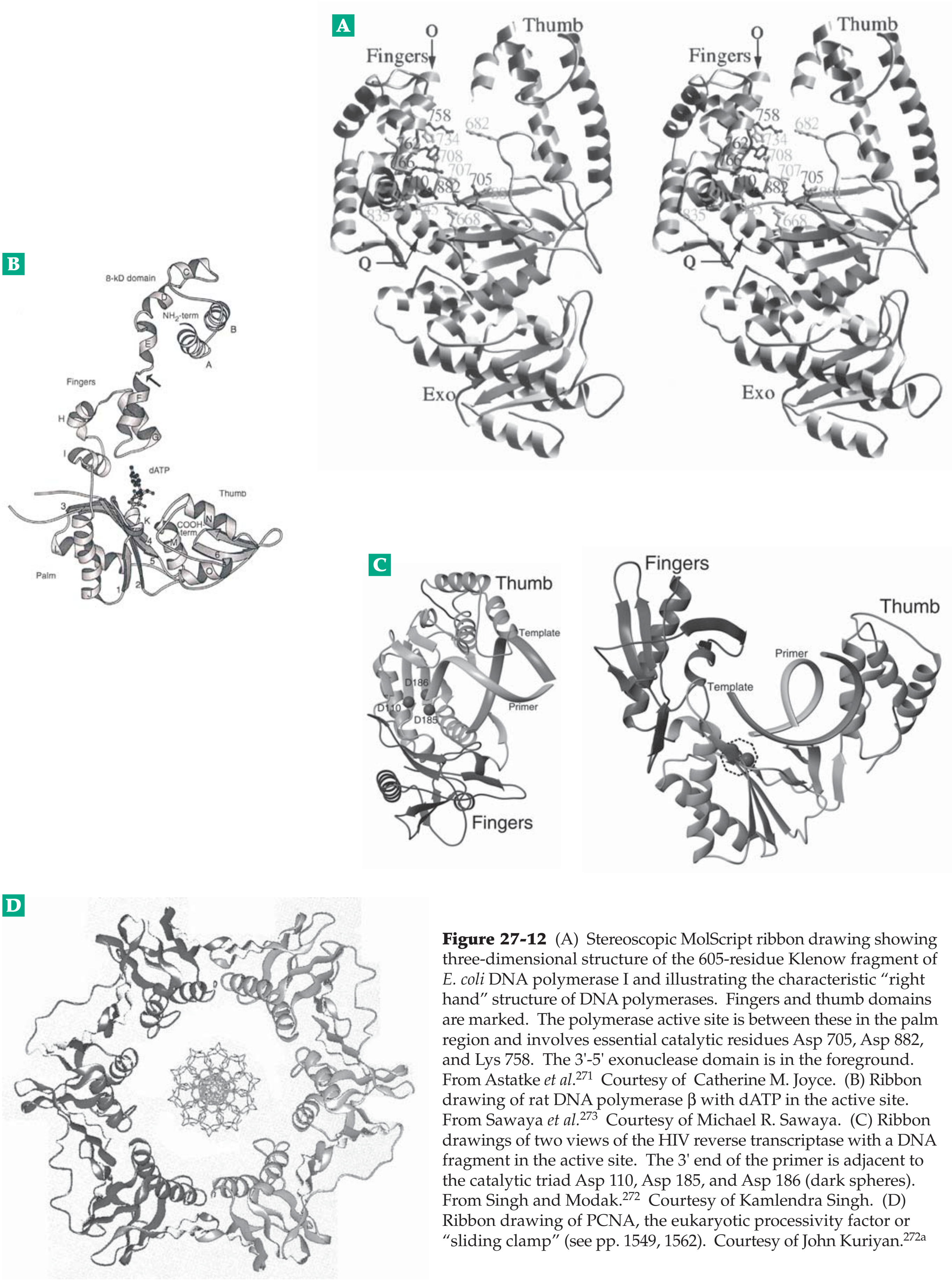

A
Thumb
O
Fingers
758
682
762
734
766
708
707
705
710
882
835
668
Q
Exo
B
8-kD domain
NH2-term
B
A
Fingers
H
I
dATP
Thumb
COOH-term
K
M
N
Palm
C
Thumb
Template
D186
D110
D185
Primer
Fingers
Fingers
Thumb
Primer
Template
D

Figure 27-12 (A) Stereoscopic MolScript ribbon drawing showing three-dimensional structure of the 605-residue Klenow fragment of *E. coli* DNA polymerase I and illustrating the characteristic "right hand" structure of DNA polymerases. Fingers and thumb domains are marked. The polymerase active site is between these in the palm region and involves essential catalytic residues Asp 705, Asp 882, and Lys 758. The 3'-5' exonuclease domain is in the foreground. From Astatke *et al.*[271] Courtesy of Catherine M. Joyce. (B) Ribbon drawing of rat DNA polymerase β with dATP in the active site. From Sawaya *et al.*[273] Courtesy of Michael R. Sawaya. (C) Ribbon drawings of two views of the HIV reverse transcriptase with a DNA fragment in the active site. The 3' end of the primer is adjacent to the catalytic triad Asp 110, Asp 185, and Asp 186 (dark spheres). From Singh and Modak.[272] Courtesy of Kamlendra Singh. (D) Ribbon drawing of PCNA, the eukaryotic processivity factor or "sliding clamp" (see pp. 1549, 1562). Courtesy of John Kuriyan.[272a]

hydrolysis from the 5' end.

Treatment with proteolytic enzymes cuts a 323-residue piece containing the 5'-3' exonuclease from DNA polymerase I leaving a larger C-terminal piece known as the **Klenow fragment**. This fragment retains the polymerase activity as well as the 3'-5' exonuclease activity and is widely used in genetic engineering. Its three-dimensional structure[271,275] is shown in Fig. 27-12A. In the Klenow fragment the large C-terminal domain contains the polymerase. The N-terminal domain contains the 3', 5'-exonuclease activity, which is thought to fulfill a **proofreading** and **editing** function.[275a] The polymerase acts at the 3' end of the growing DNA chain. Before moving on to the next position, the enzyme verifies that the correct base pair has been formed in the preceding polymerization event. If it has not, the exonuclease action removes the incorrect nucleotide and allows the polymerase to add the correct one. Thus, each base pair is checked twice, first before polymerization and then after polymerization.

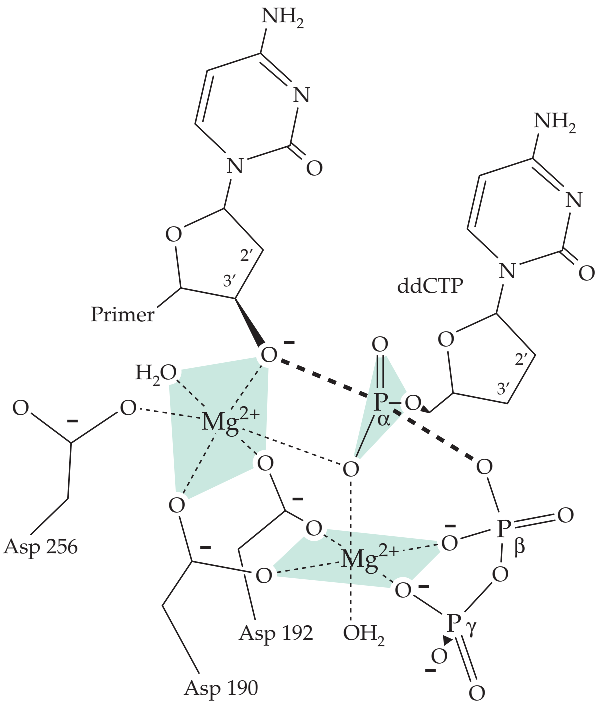


Figure 27-13 Proposed mechanism and transition state structure for the synthetic nucleotidyltransfer activity of DNA polymerase β (and other DNA polymerases). The chain-terminating inhibitor dideoxy CTP is reacting with the 3'-OH group of a growing polynucleotide primer chain. This –OH group (as –O$^-$) makes an in-line nucleophilic attack on P$_\alpha$ of the dideoxy-CTP. Notice the two metal ions, which interact with the phospho groups and which are held by three aspartate side chains. Two of the latter, Asp 190 and Asp 256, are present in similar positions in all of the polymerases. The active centers for the hydrolytic 3'-5' and 5'-3' exonuclease activities of some of the polymerases also appear to involve two-metal catalysis and in-line displacement. See Sawaya *et al.*[274]

A puzzle was the fact that structural studies indicate that the editing center is over 3 nm away from the catalytic center in Pol I.[267] The proposed solution to the puzzle is illustrated in Fig. 27-14. When the catalytic center "identifies" a nucleotide triphosphate as able to form a proper Watson–Crick nucleotide pair, it catalyzes the formation of the new nucleotide linkage. Then it releases the 3'-chain end, which sometimes "melts" and over a 10- to 100-ms time interval is able to reach over into the exonuclease site where the newly added nucleotide may be hydrolytically removed. However, if the newly formed nucleotide is properly paired, it will be less likely to melt, and the new nucleotide will be more likely to be retained.[259a,265,267] DNA polymerase I and other related polymerases utilize **processive mechanisms**, moving from one site to the next without diffusing away from the DNA. The schematic picture shown in Fig. 27-14 also indicates how the 5'-3' exonuclease activity can come into play, when the polymerase reaches the end of a single-stranded gap.

Other Class A polymerases. The *Thermus aquaticus* (***Taq***) polymerase is best known for its widespread use in the polymerase chain reaction (PCR; Fig. 5-47). Like *E. coli* I the enzyme is a large multidomain protein. The structure of the catalytic domains of the two enzymes are nearly identical, but the *Taq* polymerase has poor 3'-5' editing activity.[276] The enzyme has been carefully engineered to improve its characteristics for use in the PCR reaction.[277]

Some bacteriophage encode their own DNA polymerases. However, they usually rely on the host cell to provide accessory proteins. The sequence of the DNA polymerase from phage T7 is closely homologous to that of the Klenow fragment and the 3D structures are similar. The 80-kDa T7 polymerase requires the 12-kDa thioredoxin from the host cell as an additional subunit. It has been genetically engineered to improve its usefulness in DNA sequencing.[278]

About 45% of the sequence of the **RNA polymerase** encoded by phage T7, which transcribes RNA from the phage DNA, is also similar to that of the Klenow fragment. Sequences of these DNA polymerases are distantly related to those of reverse transcriptases.[279,280] The 136-kDa polymerase γ functions in mitochondria but is encoded in a nuclear gene. It is the only DNA polymerase that is inhibited by antiviral nucleotide analogs such as AZT (Box 28-C).[280a,b]

Polymerases of Class B. Although *E. coli* polymerase II is a member of this family, relatively little is known about its function. It may participate in DNA repair in the "SOS" response (Section E).[265]

The catalytic subunits of the major *eukaryotic* DNA polymerases as well as of archaeal DNA polymerases are members of the B family.[267,281–282a] Eukaryotic cells contain at least 13 DNA polymerases which are designated by Greek letters (Table 27-1). Polymerases α, δ, and ε are essential for nuclear DNA synthesis.[283–286b] They function together with accessory proteins in primase and replisome complexes considered in Section 8. Others participate in DNA repair (pp. 1583, 1584).

Phage T4 encodes a DNA polymerase much used in the laboratory because of its ability to polymerize using a long single-stranded template. It also depends upon accessory factors provided by the bacterial host (Section 5).[287,287a]

Other DNA polymerases. **Reverse transcriptases** synthesize DNA using an RNA template strand. They are best known for their function in retroviruses (Chapter 28). The HIV reverse transcriptase is a heterodimer of 51- and 66-kDa subunits. The larger subunit contains a **ribonuclease H** domain.[288–289a] The enzyme is a prime target for drugs such as AZT and others.[290,291] A different reverse transcriptase is found in all eukaryotic cells in **telomerase**, an enzyme essential for replication of chromosome ends. Reverse transcriptases have also been found in rare L1 sequences that are functioning retrotransposons (Section D).[292]

The ~335-residue catalytic subunit of **eukaryotic polymerase β**, which has a DNA repair function, is the simplest known DNA polymerase. The active enzyme (Fig. 27-12B) is as small as 38 kDa. It lacks proofreading and is less accurate than the eukaryotic replicative polymerases listed in Table 27-1.[293–293c] Structural analysis revealed a folding pattern (Class X) related to that of a **nucleotidyltransferase superfamily** that includes enzymes such as terminal deoxynucleotidyltransferase (Chapter 12) and the glutamine synthase adenylytransferase (Fig. 24-7).[294] However, its active site (Fig. 27-13) is similar to that of other DNA polymerases.

DNA polymerase III. This Class C enzyme is the major bacterial polymerase for DNA replication. In its complete holoenzyme form it can synthesize new DNA strands at rates as high as ~1 kilobase s^{-1} without dissociation from a template.[294a] A genetic approach has provided important information about DNA replication.[265,266] A series of temperature-sensitive mutants of *E. coli* unable to carry out DNA synthesis were obtained. From these mutations, genes *dnaA*, *F*, *G*, *I*, *J*, *K*, *L*, *P*, *I*, *X*, and *Z* were identified and located at various points on the chromosome map. Genes *C* and *D* map close together at 89 min, and it now appears that they are one gene. Gene *F* encodes a ribonucleotide reductase (Eq. 16-22). The functions of genes *A*–*E*, *G*, *X*, and *Z* are discussed in the following sections. None of these genes code for DNA polymerase I, but gene *dnaE* was identified as that of DNA polymerase III, which is now known to be the major polymerase in bacterial DNA replication.[265,295] To obtain rapid error-

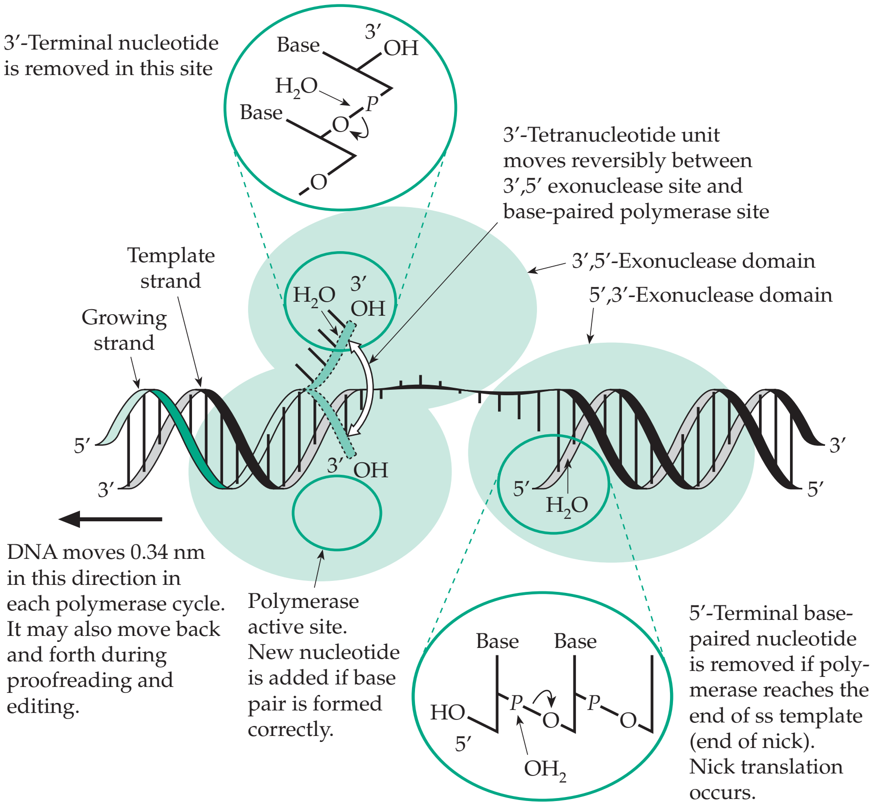

3′-Terminal nucleotide is removed in this site
Base
3′
OH
H_2O
P
Base
O
O
3′-Tetranucleotide unit moves reversibly between 3′,5′ exonuclease site and base-paired polymerase site
3′,5′-Exonuclease domain
5′,3′-Exonuclease domain
Template strand
Growing strand
H_2O
3′
OH
5′
3′
3′
OH
3′
5′
5′
H_2O
DNA moves 0.34 nm in this direction in each polymerase cycle. It may also move back and forth during proofreading and editing.
Polymerase active site. New nucleotide is added if base pair is formed correctly.
Base
Base
HO
5′
P
O
P
O
OH_2
5′-Terminal base-paired nucleotide is removed if polymerase reaches the end of ss template (end of nick). Nick translation occurs.

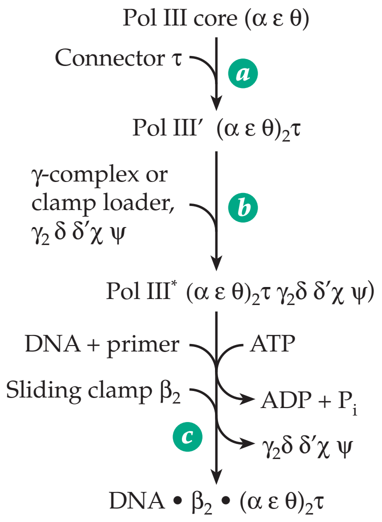
Figure 27-14 Schematic representation of DNA polymerase action on a nicked strand of DNA in which the nick has been enlarged. At the catalytic center new nucleotide units are added at the 3' end of a growing strand. At the 3'-5' exonuclease site the 3' terminal nucleotide may be removed hydrolytically. This will happen to the greatest extent if the nucleotide is poorly paired in the duplex. At the 5'-3' exonuclease site nucleotides are hydrolytically removed from the 5' end of a strand in the chain.[265,267]

free synthesis it must be combined with a number of other subunits to form **polymerase III holoenzyme** (or replisome). Ten different subunits, some as two or more copies, form the holoenzyme.[294a,296–299] Subunits are listed in Table 27-2. The polymerase catalytic center is in the 132-kDa α subunit. The 27.5-kDa ε subunit contains the 3'-5' exonuclease editing activity.[298,300,300a] Mutation in its gene (*dnaQ*) leads to a high spontaneous mutation rate in bacteria. Together with the θ subunit, α and ε form the polymerase III core. This complex has polymerase activity and improved proofreading ability but is still unable to act rapidly, accurately, and processively. Full catalytic activity requires at least the additional β, γ, δ, and τ subunits (Table 27-2). The presence of the τ subunit causes the core complex to dimerize to form Pol III′ (Eq. 27-4, step *a*). Pol III′ can add only about ten nucleotides to a growing DNA chain before it dissociates.[301] The presence of the β_2 dimer, known as the **processivity factor** or **sliding clamp**,[301a,b] is essential for highly processive polymerization. The β protein forms a ring around the duplex DNA and interacts with the polymerase clamping it tightly to the DNA. Putting the β_2 clamp onto the DNA is an ATP-dependent process that first involves binding of the **γ-complex** or **clamp loader** (Eq. 27-4, step *b*), ATP-dependent opening of the β_2 ring, and insertion of the DNA duplex (step *c*).[301a,c] This complex may form a replisome structure that acts simultaneously on the leading and lagging strands (see Fig. 27-19).[302,303] The *Bacillus subtilis* replisome appears to contain two different catalytic (α) subunits, perhaps one for each strand.[303a]

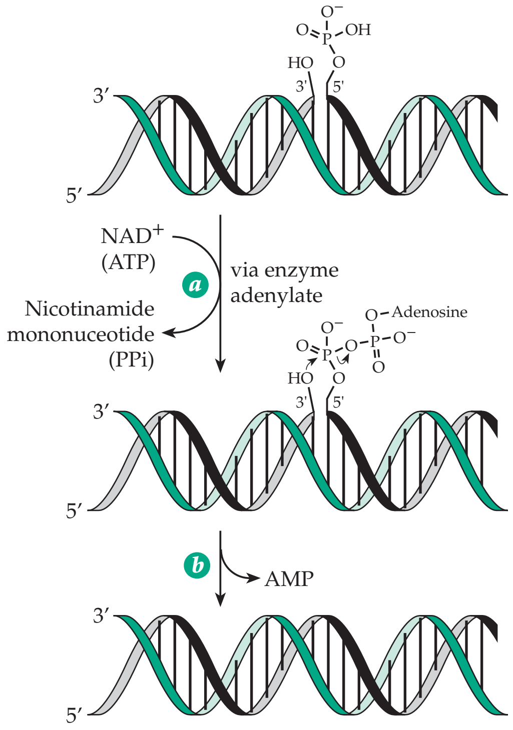
3. Other Replication Proteins

DNA ligases. These enzymes, which are essential to replication, have a specific function of repairing “nicked” DNA.[261,304–307a] Such DNA, as indicated in Eq. 27-5, has a break in one strand and contains a 3'-hydroxyl group and a 5'-phosphate group, which must be rejoined. The ligase from *E. coli* activates the phosphate group in an unusual way by transfer of an adenylyl group from NAD$^+$, with displacement of nicotinamide mononucleotide (Eq. 27-5, step *a*). The reaction is completed by displacement of AMP as indicated in Eq. 27-5, step *b*. Cells infected by bacteriophage T4 synthesize a virally encoded ligase, which utilizes ATP rather than NAD$^+$ as the activating reagent. The ~190-kDa mammalian DNA ligase I has been found deficient in some patients with the Bloom syndrome, a condition associated with poor DNA repair and a high incidence of cancer (see also p. 1550).[304]

Single-strand binding proteins. Genetic analysis of replication of the DNA of phage T4 within cells of *E. coli* revealed that at least five genes of the virus are required. One of these, gene 43 specifies the T4 DNA polymerase, while gene 32 codes for a single-strand binding protein, also known as the DNA unwinding, melting, or helix-destabilizing protein. It has a greater affinity for ssDNA than for dsDNA and binds to a length of ssDNA causing unwinding of the double helix and exposure of the purine and pyrimidine bases of the template strand.[308] The protein is required for replication, genetic recombination, and repair of DNA. Similar proteins are encoded in the genomes of many viruses.[309] The 87-residue single-strand binding protein encoded by gene 5 of phage M13 forms a dimer, which completely coats newly synthesized ssDNA preventing the DNA polymerase system of the host



(27-4)



(27-5)

bacteria from converting it into dsDNA. The polynucleotide chain binds into a groove in the protein with one tyrosine intercalated between the DNA bases.[310,311]

The *E. coli* single-strand binding protein, another helix-destabilizing protein that is usually called simply **SSB**, is a tetramer of 18.5-kDa subunits.[265,312,313] It is essential to DNA replication. About 35 nucleotides may bind to each tetramer.[314] The situation is not as clear in eukaryotes where DNA is largely coiled around histones in the nucleosomes. Several single-strand binding proteins have been identified,[315] but the need for SSB proteins in eukaryotic nuclear replication is uncertain.[316] A human mitochondrial SSB resembles that of *E. coli*.[317]

Helicases and topoisomerases. Cells of *E. coli* contain at least 12 DNA-dependent ATPases that cause unwinding of DNA at the expense of hydrolysis of ATP.[318] The activity of these **helicases** is essential to replication, repair, and recombination of DNA in all organisms.[319–320b] The primary replicative helicase of *E. coli*, which unwinds DNA ahead of the replication fork, is encoded by gene *dnaB*. The active form is a hexamer, which exists in at least two conformational states.[321,321b] A segment of ~20 nucleotide units of DNA binds to one hexamer.[322] Helicases have ATPase activity, and the dnaB hexamer contains six ATP-binding sites. However, only three of them may be occupied.[323,324] Many helicases have a hexameric ringlike structure; that of *Bacillus subtilis* is seen clearly in electron micrographs.[325] Three-dimensional structures are known for some.[321a,325a] Although these enzymes may bind to duplex DNA, they also bind to and move along single-stranded DNA in either the 5'→3' or 3'→5' direction. The directionality of a helicase can be determined by annealing two small pieces of ssDNA to the 5' and 3' ends of a longer strand of ssDNA. A 5'→3' helicase will translocate along the long ssDNA and displace the oligonucleotide annealed at the 3' end of the strand, while a 3'→5' helicase will displace the oligonucleotide annealed at the 5' end.[265]

The dnaB protein is a 5'→3' helicase. However, the first helicase identified, the product of the *E. coli rep* gene, is a 3'→5' helicase. It is one of the host proteins needed for the propagation of phages such as ϕX174 and M13. It catalyzes the unwinding of the double-stranded replication forms of these viruses. It binds to a stretch of ~20 nucleotides in a single-stranded region of nicked DNA. The hydrolysis of ATP moves the enzyme along the bound strand in a 3' to 5' direction opening up the DNA at a replication fork. Another *E. coli* 3'→5' helicase is protein *priA* (also called n'), a component of some primasome structures involved in replication of viruses.[318]

The bacteriophage T4 gene 41 protein, a 5'→3' helicase, functions together with the gene 61 primase in replication of that virus.[326–326b] The phage T7 gene 4 protein and virus SV40 large T antigen are also hexameric ringlike helicases. The *E. coli* protein **RecQ** is required for various aspects of recombination and is the prototype of a large group of helicases present in both prokaryotes and eukaryotes.[326c] The bacterial ruvB protein is a hexameric helicase that propels branch migration in Holliday junctions (Fig. 5-28) during genetic recombination,[327–327b] while helicase rho is required for termination of RNA synthesis.[328] Numerous eukaryotic helicases have been identified and purified.[329–331] Helicase DNA2 is needed for DNA replication in nuclear extracts from yeast.[332] Human 3'→5' DNA helicases, members of the RecQ family, are defective in some patients with **Bloom syndrome** and **Werner Syndrome**[332a] (Box 27-A). Bloom syndrome causes growth retardation, immunodeficiency, sensitivity to sunlight, and a predisposition to skin cancers and leukemias.[331a,b] The yeast (*S. cerevisiae*) genome contains genes for 134 different proteins that are probably helicases.[332b] RNA helicases are also known.[333]

A characteristic of all helicases is their ATPase activity, which apparently provides energy for “melting” the DNA. The mechanisms are not clear, but rapid separation of the stacked and hydrogen-bonded base pairs may be impossible without some assistance from an ATP-dependent process. In the case of the rep protein hydrolysis of two molecules of ATP seems to be required to melt one base pair.[323] It isn’t clear whether one strand of DNA passes through the center of the oligomeric ring, as shown in Fig. 27-15, or whether both strands pass through. Helicases vary in their amino acid sequences, but they all possess several characteristic **signature sequences** including the Walker A and B motifs found in other ATPases, in

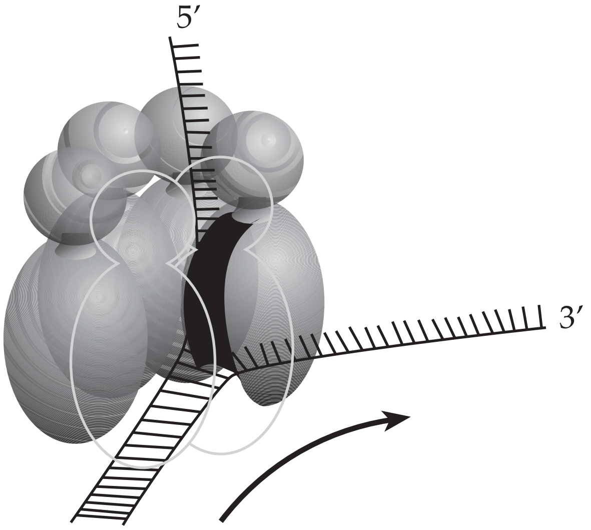
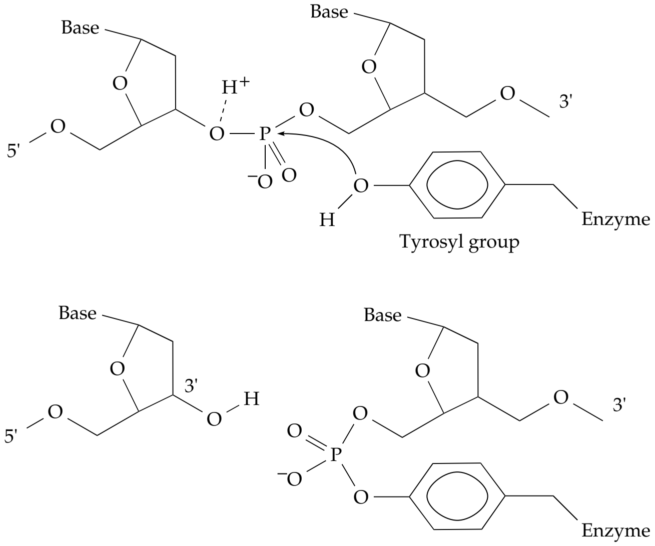


Figure 27-15 Model of DnaB hexamer proposed by Jezewska *et al*.[322] The arrow indicates the direction of movement of the DNA relative to the position of the helicase.

sequences related to the *E. coli* **recA** proteins (Section D) and in synthases (Chapter 18).[328,334]

While helicases can cause the unwinding of linear DNA duplexes, they cannot alter the linking number of covalently closed circular double-stranded DNA. The latter is the function of **topoisomerases**, which have been found in organisms from bacteria to higher plants and animals and may also be encoded by viruses.[265,335–337] There are two basic types of DNA topoisomerase. Those of type I change the linking number in steps of 1. One way that this might be accomplished is for an enzyme to nick one strand in the DNA allowing one of the cut ends to swivel around the unbroken strand, then to reseal the chain. However, it was found that topoisomerases can also cause catenation or decatenation of circular duplex DNA as long as at least one of the reacting DNA molecules is nicked. This observation suggested that a topoisomerase binds to a single-stranded region at a nick and cuts the chain but does not release the ends. This permits either a single strand or a duplex to pass through the broken strand, which is then resealed.

Topoisomerases of type I usually act most rapidly on negatively supercoiled DNA. They relax it by decreasing the number of negatively supercoiled turns one at a time. Negative supercoiling presumably facilitates binding of the enzyme to a single-stranded region by unwinding of the duplex. No ATP or other obvious source of energy is needed by type I topoisomerases. The chain cleavage involves a simple nucleophilic displacement by an –OH group of a tyrosine side chain (Tyr 319 of *E. coli* topoisomerase I),[338] which attacks a phosphorus atom in the DNA chain (Eq. 27-6). The result is covalent attachment of the enzyme to the 5'-end of the cut strand in type IA topoisomerases and at the 3'-end in type IB topoisomerases.[336,337,339] After the passage of the other



(27-6)

TABLE 27-2
Some Proteins of DNA Replication in *E. coli*[a]

Name	Gene	Mass of monomer (kDa)	Map location (min) Fig. 26-4
Polymerase I (also has 3'-5' and 5'-3' exonuclease, and RNase activities)	*polA*	103	87
Polymerase III			
Core			
α subunit (polymerase)	*dnaE*	130	4
ε subunit (3'-5' exonuclease)	*dnaQ*	27.5	5
θ subunit	*hol E*	10	
Sliding clamp, β_2	*dnaN*	37 × 2	83
Connector τ (ATPase)	*dnaX*	71	
Clamp loader (γ complex)			
γ subunit	*dnaX*	47.5 × 2	11
σ subunit		35	
σ' subunit		33	
χ subunit		15	
ψ subunit		12	
DNA binding proteins			
Single-strand, SSB	*ssb*	18.5 × 4	92
Double-strand, HU			
α subunit		9.5 × 2	
β subunit		9.5 × 2	
Helicases (ATP-dependent)			
Primary replicative	*dnaB*	52 × 6	92
Dna C protein	*dnaC*		99
PriA (n′), primosome	*priA*	76	88
Rep	*rep*	76.4	85
Initiation and priming proteins			
Dna A protein	*dnaA*	52	83
Primase (an RNA polymerase)	*dnaG*	60	67
PriB (n) primosome	*priB*		96
PriC (n′′) primosome	*priC*		
DnaT (primosome assembly)	*dnaT*		99
Ribonuclease HI	*rnhA*		
DNA ligase	*lig*	75	52
Topoisomerases			
Type I	*topA*		28
Type II, DNA gyrase ($\alpha_2\beta_2$)			
Subunit α	*gyrA*		97
Subunit β	*gyrB*		90

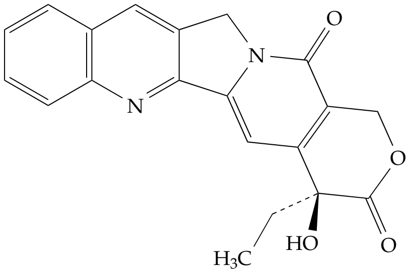
[a] In large part from Kornberg, A., and Baker, T. A. (1992) *DNA Replication*, 2nd ed., Freeman, New York

strand through the gap in the cut strand of a type IB topoisomerase (Fig. 27-16) the free 3'-OH oxygen atom (or 5'-OH) attacks the phosphorus atom in the phosphotyrosine diester to reform the chain and release the enzyme. X-ray diffraction studies show that both the Type IA *E. coli* topoisomerase I and the human type IB topoisomerase[339–341a] are large proteins with holes of appropriate diameter for a DNA double helix. As illustrated in Fig. 27-16B, the protein may open to allow a double helix to enter and occupy a suitable position for cleavage of one chain by the active-site Tyr 319. Topoisomerases are metalloenzymes, usually functioning best with Mg^{2+}. *E. coli* topoisomerase I also contains 3–4 tightly bound Zn^{2+} ions.

Topoisomerases of type II change the linking number by 2 in either the positive or negative direction and hydrolyze ATP in the process. The best known of these is the *E. coli* **DNA gyrase**, an $\alpha_2\beta_2$ dimer of 97-kDa (α) and 90-kDa (β) subunits.[343,344] The enzyme catalyzes the ATP-dependent introduction of negative supercoils into DNA. It also relaxes negatively supercoiled DNA slowly in the absence of ATP. Type II topoisomerases are found in all organisms.[335,345] They are encoded by some bacteriophage such as T4[346] and by plasmids.[347] However, most differ from bacterial DNA gyrase in not coupling DNA supercoiling to ATP hydrolysis. They require ATP but like topoisomerase I cause a relaxation of the supercoiling. Strands of one segment of DNA (called the "gate" or **G-segment**) are cleaved by the enzyme with staggered cuts four base pairs apart. Another segment of DNA (the "transport" or **T-segment**) is then passed through the gate and is thought to be released from a second gate in the complex (Fig. 27-16C).[342,348] The enzyme subunits bind through phosphotyrosine linkages as in Eq. 27-5 to the 5' phospho groups of the two cleaved chains, while the subunits bind ATP and may like tiny muscles twist the DNA.[349] Topoisomerases II are large dimeric proteins. The subunits of yeast topoisomerase II (Fig. 27-16C) are 1200-residue multidomain proteins.[342,350] Mechanisms of DNA cleavage by types I and II topoisomerases appear to be related.[351] However, the ATP-dependent conformational changes involved in a two-gate mechanism are unique to topoisomerases II.[348,348a] The bacterium *Sulfolobus* contains a type I topoisomerase that is called **reverse gyrase** because it utilizes ATP to introduce *positive* supercoils into DNA.[352] This is in contrast to gyrase, a type II topoisomerase that introduces negative supercoils.

Type II topoisomerases are essential and function in replication, DNA repair, transcription, and chromosome segregation at mitosis.[345,349] Yeast with a *top2* mutation dies during mitosis with hopelessly entangled daughter chromosomes.[353] A fluorescent antibody to eukaryotic topoisomerase II binds to chromosomes, probably at the bases of the radial loops present during mitosis,[353] and also to centrosomes[353a] Higher organisms contain more than one topoisomerase II.[354] Their specific functions are uncertain, but one appears to be to unknot entangled chromosomal DNA. In the crowded conditions of a cell nucleus topoisomerase can also cause inadvertent *formation* of knots.[355]

The functional role of topoisomerases of type I is less clear. Staining with fluorescent antibodies to the enzyme has revealed its presence in the transcriptionally active "puffs" of polytene chromosomes (p. 1635)[356] and in centromeres of mitotic cells.[357] A current hypothesis is that in *E. coli* class I topoisomerases act to relax negatively supercoiled strands of DNA behind transcription complexes, while gyrase acts to generate superhelical twists, which favor opening of the duplex ahead of transcription complexes.[354,358] Transcription of a supercoiled rRNA gene *in vitro* is diminished by the selective topoisomerase I inhibitor **camptothecin**, one of a group of antitumor drugs directed against topoisomerase of both types I and II.[349,354,359]



Camptothecin, a topoisomerase I inhibitor

In the autoimmune diseases scleroderma and systematic lupus erythematosus antigens to nuclear proteins or nucleic acids are present in the blood. Many patients with severe scleroderma have an antibody against topoisomerase I.[360]

Primases, initiator proteins, and ribonucleases. The priming segment needed for initiation of DNA replication is either a short segment of RNA or an oligonucleotide containing a mixture of ribonucleotide and deoxyribonucleotide units. The enzyme forming the primer is an RNA polymerase called **primase**. In *E. coli* it is encoded by gene *dna*G.[361–362a] Under some circumstances other RNA polymerases can act as the primase. Bacteriophages and plasmids may also encode primases (Table 27-1). For example, gene 61 of phage T4 (Fig. 26-2) encodes a primase, which together with the T4 helicase forms the priming particle or **primosome**.[363] The phage T7 gene 4 encodes a 63-kDa mutlifunctional protein that is both a primase and a helicase.[364–366b] The primase active site is on the outside of the hexameric ring. Additional proteins representing products of genes *dna*A and C are also required for initiation of replication in *E. coli*. Several molecules of the dnaA **initiator protein** bind to a specific DNA **origin** sequence and participate in assembling a primosome that also contains the hexameric dnaB helicase and, transiently, protein dnaC.[367–370] Replication of some single-stranded phages, such as ϕX174, in *E. coli* also require the *E. coli* priA, priB, and priC proteins (Table 27-2).[265,371,372] For successful completion of replication chaperonins the products of genes *dna*J and *grp* E are needed[373] as is a ribonuclease that digests the primer segments after they have been used as replication initiators.[374–376]

4. Replication of Bacterial DNA

The basic mechanisms of replication implied in Eq. 27-3 seems to be universal, but several questions had to be asked. “Is replication initiated at a fixed point or points in a chromosome?” and “Does replication occur in one direction only or do two forks form at the point

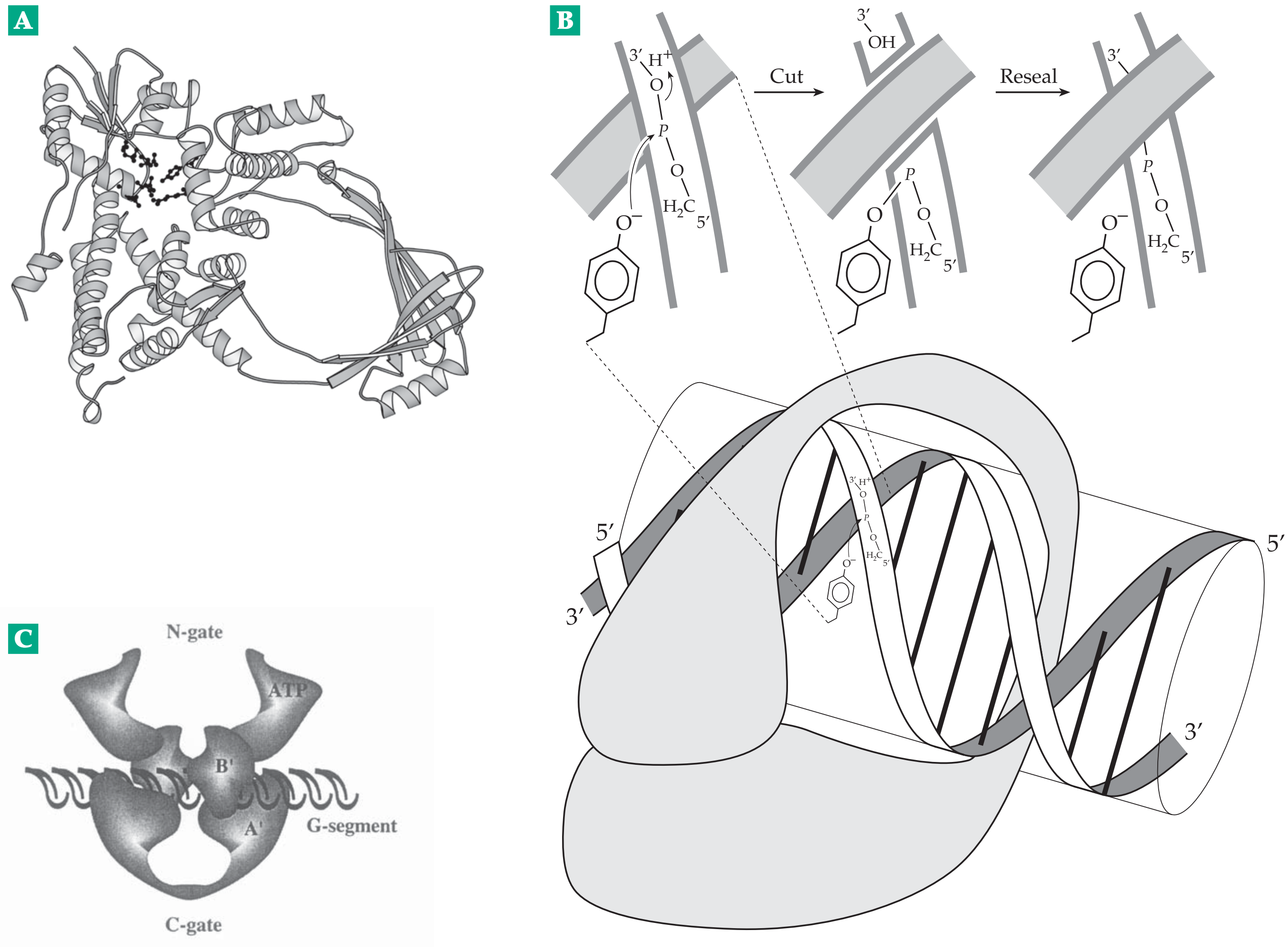


Figure 27-16 (A) Ribbon drawing of a large 67-kDa fragment of the 97-kDa (864-residue) *E. coli* topoisomerase I showing the position of the active-site tyrosine 319 and an adjacent arginine. From Lima *et al.*[341] Courtesy of Alfonso Mondragon. (B) Schematic diagram indicating a way in which topoisomerases of type 1 may pass one strand of DNA through another. The protein is shown binding to a single strand of a DNA duplex. This binding is facilitated by negative supercoiling. The enzyme then cuts the same strand by means of a nucleophilic displacement on a phosphorus atom using a tyrosinate side chain. The other cut end is held noncovalently by the enzyme, while the second strand passes through the gap. Then the gap is resealed by a reversal of the cleavage reaction. From Lima *et al.*[341] (C) Schematic model of a type II topoisomerase bound to a G-segment of DNA. This double helix is cut and another double helical strand, the T-strand, enters the N-gate. The gate then closes, and the central gate opens to allow the T-strand to pass through and exit through the C-gate. The shapes of the three domains are based on crystallographic data for the yeast enzyme. The ATPase, B′, and A′ domains consist of residues 1 to 409, 410 to 660, and 660 to 1200, approximately. From Olland and Wang.[342] Courtesy of James C. Wang. See also Champoux.[340c]

of origin and travel in opposite directions?" To answer these questions both genetic methods and electron microscopy have been employed.

Directions of replication. One technique for establishing the direction of replication in *E. coli* was to insert a λ prophage at the *att* site (Fig. 26-4, 17 min) and DNA from phage Mu-1 at a variety of other sites around the chromosome.[377] Phage Mu-1 was especially useful because it can be integrated at many different locations within the well-mapped genome. Integration within a gene inactivates that gene (an addition mutation) and allows the localization of the Mu prophage. Bacteria were prepared containing both λ and Mu-1 prophage, the latter at various sites. The bacteria were also auxotrophic for certain amino acids. Because of this, replication could be stopped by amino acid starvation. The bacteria usually completed any replication cycle in progress and then stopped. When the missing amino acids were added, replication began, again starting from the replication origin. Bromouracil, which enters DNA in place of thymine, was added at the same time. Consequently, the newly synthesized DNA strands were denser than the parent strands (see Fig. 27-11). After various times of replication the newly formed strands were separated by centrifugation in a CsCl gradient and were tested for hybridization with both λ and Mu-1 DNA. Since the cells did not all begin replication at the same time after addition of amino acids, a variety of lengths of newly replicated DNA were present. Nevertheless, from the observed ratios of Mu-1 DNA to λ DNA for the various strains it was possible to map the progress of replication beginning at an **origin** *ori*C near gene *ilv* at 74 min (Fig. 26-4). Replication was found to progress bidirectionally around the chromosome and to terminate between genes *trp* and *his* at ~25 min.

The use of autoradiographic methods confirmed bidirectional replication. Strains of amino acid auxotrophs with small nucleoside triphosphate pools were used. The addition of amino acids after starvation led to initiation of replication with only a 6-min lag. The cells were labeled with [^{3}H]thymidine, and after the replication forks had moved a short distance from the origin of replication the cells were given a pulse of "super-hot" [^{3}H]thymidine. Using autoradiography it was possible to observe the clearly bidirectional replication forks[378] (Fig. 27-17). Replication in other bacteria is also bidirectional.

Origins of replication. Replication of the *E. coli* chromosome begins and proceeds bidirectionally from its defined origin *ori*C. Replication of linear phage T7 is also bidirectional and begins at a point 17% of the way from one end.[265] In mammalian mitochondrial DNA the origin of replication for the H-strand is in the D-loop but that for the L-strand is 2 / 3 of the way around the circular chromosome within a cluster of tRNA genes (Fig. 18-3).[381,382] The single-stranded circular DNA genomes of Ff, ϕX174, and G4 phage also have distinct origins for initiation of replication to give RF circles.[265]

Most origins have quasi-palindromic nucleotide sequences, perhaps so that DNA can be looped out from the main duplex as is shown in Fig. 27-18A and B. The lengths of *ori* sequences vary, as does the complexity of their possible folding patterns. Plasmids have been constructed which not only contain the *E. coli* origin, but are dependent upon that origin for their own replication.[382a] Study of those plasmids indicate

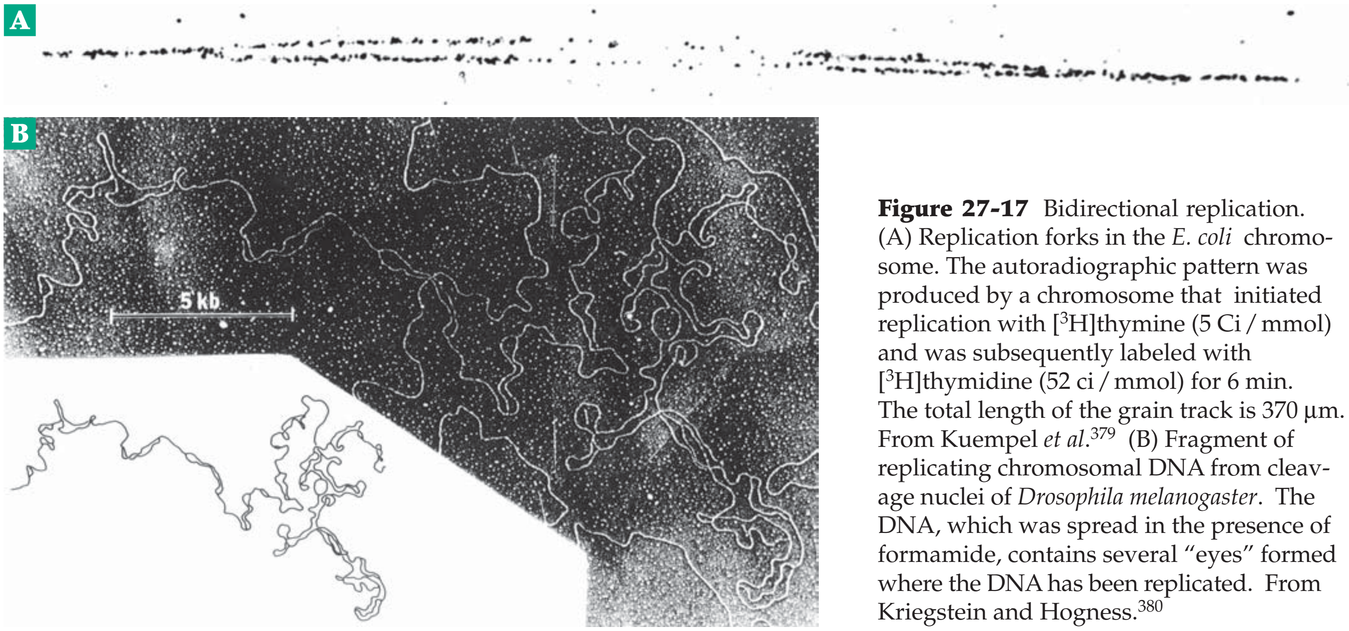


Figure 27-17 Bidirectional replication. (A) Replication forks in the *E. coli* chromosome. The autoradiographic pattern was produced by a chromosome that initiated replication with [^{3}H]thymine (5 Ci / mmol) and was subsequently labeled with [^{3}H]thymidine (52 ci / mmol) for 6 min. The total length of the grain track is 370 μm. From Kuempel *et al*.[379] (B) Fragment of replicating chromosomal DNA from cleavage nuclei of *Drosophila melanogaster*. The DNA, which was spread in the presence of formamide, contains several "eyes" formed where the DNA has been replicated. From Kriegstein and Hogness.[380]

that a 245-bp *ori*C sequence is essential.[265,383,383a] This sequence, which is shown in Fig. 27-18, contains several repeated oligonucleotides including 11 GATC sequences, which are sites of adenine N^6-methylation (see Section B, 6), and four "9-mers" (commonly known as dnaA "boxes") with the consensus sequence

$$\mathrm{TTAT}\overset{\mathrm{A}}{\underset{\mathrm{C}}{}}\mathrm{CA}\overset{\mathrm{A}}{\underset{\mathrm{C}}{}}\mathrm{C}$$

These appear in both 5' to 3' and 3' to 5' orientations, allowing them to form two base paired "stems." In addition, there are three direct repeats of a 13-residue consensus sequence 5'-GATCTNTTNTTTT (shaded in green in Fig. 27-18), which form an AT rich duplex. Other bacterial replication origins often follow a pattern similar to that in Fig. 27-18.[384] However, the origin for *Mycoplasma genitalium* has been hard to detect.[385] Replication origins of archaea have characteristics similar to those of bacteria and of organelles.[385a]

Priming and initiation of DNA synthesis. The first step in initiating a new round of replication and a new cell cycle in *E. coli* appears to be the binding of the initiator protein, dnaA, to the 9-mers in the negatively supercoiled origin.[295,383,386–388] The resulting complex is visible in an electron microscope and may consist of a core of up to 20–40 molecules of dnaA protein with the DNA wrapped around them (Fig. 27-19). ATP is also required and is bound to the protein where it hydrolyzes slowly. Similar initiator proteins are found

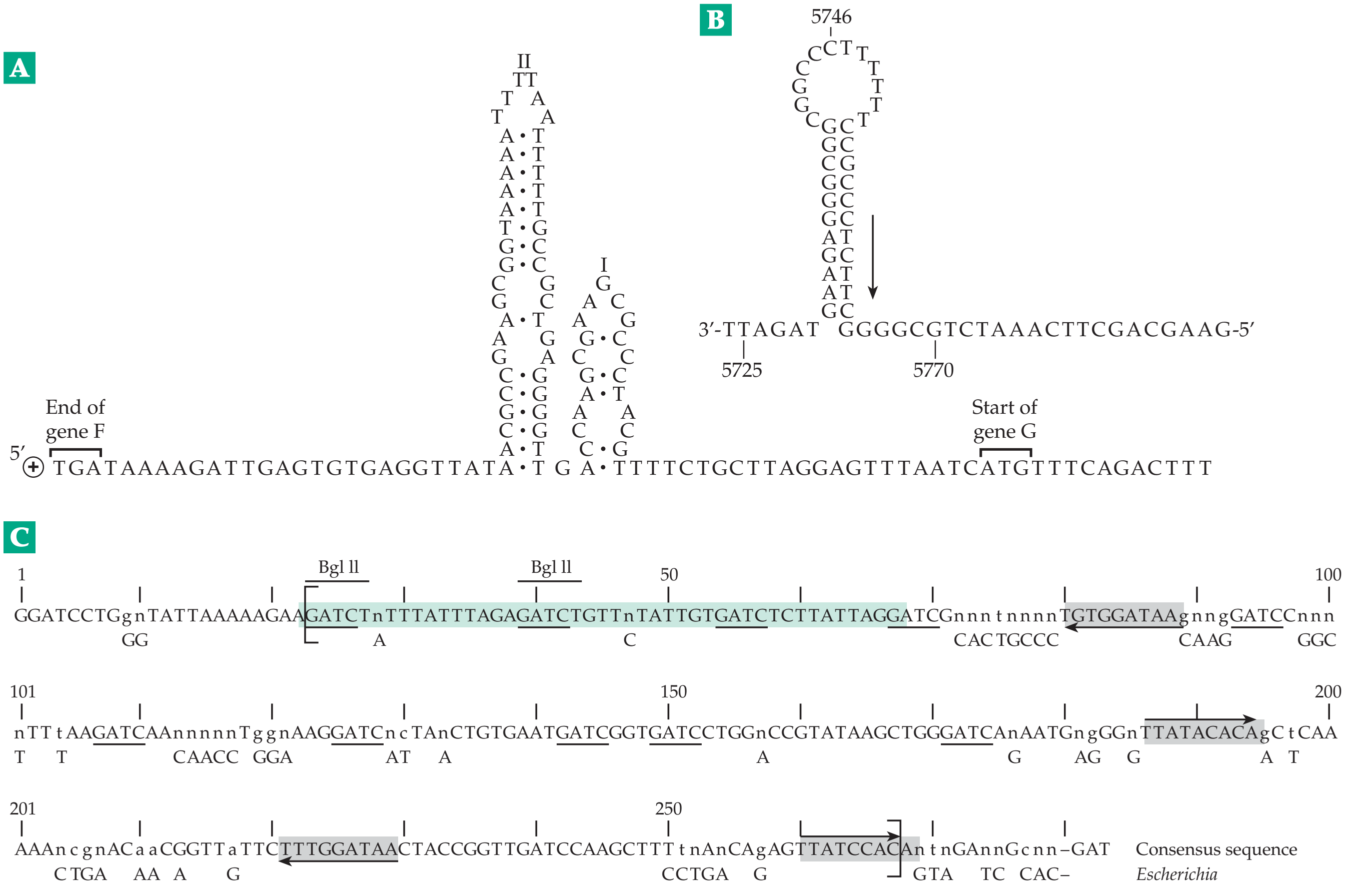

A
II
I
End of gene F
Start of gene G
5′ TGATAAAAGATTGAGTGTGAGGTTATA•T G A•TTTTCTGCTTAGGAGTTTAATCATGTTTCAGACTTT
B
5746
3′-TTAGAT GGGGCGTCTAAACTTCGACGAAG-5′
5725
5770
C
1
Bgl II
Bgl II
50
100
GGATCCTGGgnTATTAAAAAGAAGATCTnTTTATTTAGAGATCTGTTnTATTGTGATCTCTTATTAGGATCGnnntnnnnTGTGGATAAgnngGATCCnnn
101
150
200
nTTtAAGATCAAnnnnnTggnAAGGATCncTAnCTGTGAATGATCGGTGATCCTGGnCCGTATAAGCTGGGATCAnAATGngGGnTTATACACAgCtCAA
201
250
AAAncgnACaaCGGTTaTTCTTTGGATAACTACCGGTTGATCCAAGCTTTtnAnCAgAGTTATCCACAntnGAnnGcnn–GAT
Consensus sequence
Escherichia

Figure 27-18 Nucleotide sequences and proposed secondary structures for some origins of replication. (A) Bacteriophage ϕX174. The ends of the neighboring genes F and G are shown. From Kornberg and Baker.[265] (B) Human mitochondrial DNA; origin of replication of the L strand. The sequence shown is for the H strand. The arrow indicates the direction of replication. From Hixson *et al*.[381] (C) Bacterial *ori* region. The consensus sequence of origins of replication that will function in *E. coli*. Derived from sequences from *E. coli, Salmonella typhimurium, Enterobacter aerogenes, Klebsiella pneumoniae,* and *Erwinia carotovora*. In the consensus sequence a large capital letter means that the same nucleotide is present in all six origins; a lowercase letter means that the same nucleotide is present in three of the five origins; n means that any of the four nucleotides may be present; – represents a deletion, G being present here in the *Enterobacter* origin. Underlined are *dam* methylation sites. Green shaded sequences, three 13-mers in a region that undergoes easy opening of the duplex for insertion of replication proteins; gray boxes with arrows, four 9-mers that are specific sites for binding of replication protein A. Brackets enclose a 245-bp minimum origin. From Kornberg and Baker.[265] Data from Zyskind *et al*.[390]

in many bacteria and are encoded by some viruses and plasmids. In archaea, as well as in humans, the initiator proteins seem to combine functions of the *E. coli* dnaA protein with those of SSB.[391,391a]

Following the binding of the dnaA protein the hexameric helicase dnaB (Fig. 27-15) is loaded onto the adjacent DNA in the region of the 13-residue repeated sequences, which are labeled 1, 2, and 3 in Fig. 27-19. As shown in this figure introduction of the helicase is assisted by protein dnaC, which forms a complex with the helicase. Additional binding of protein HU and a temperature of >30°C are essential for tight binding of the proteins. Although the dnaC protein is needed for formation of the prepriming complex, it dissociates after the dnaB protein becomes firmly bound. One dnaB hexamer binds to each single strand of the DNA duplex with opposite orientations (Fig. 27-19).[383a] With the dnaB helicase in place on each strand, this ATP-driven enzyme processes along the DNA, unwinding the duplex in both di-rections. (Perhaps it may be more accurate to say that the *DNA moves through the helicase*.) The resulting "bubble" is held open by binding of SSB tetramers. The primase (dnaG protein) then binds adjacent to the helicase and synthesizes the RNA primer along each strand of DNA. As the helicase processes to the right in the fork shown in Fig. 27-19, the complex with the dnaA protein dissociates, permitting primer synthesis into the origin region. Alternatively, RNA polymerase (Chapter 28) can prime replication by initiating transcription on both strands of the DNA. Suitable promoters are present and oriented in opposite directions on the two strands.[392]

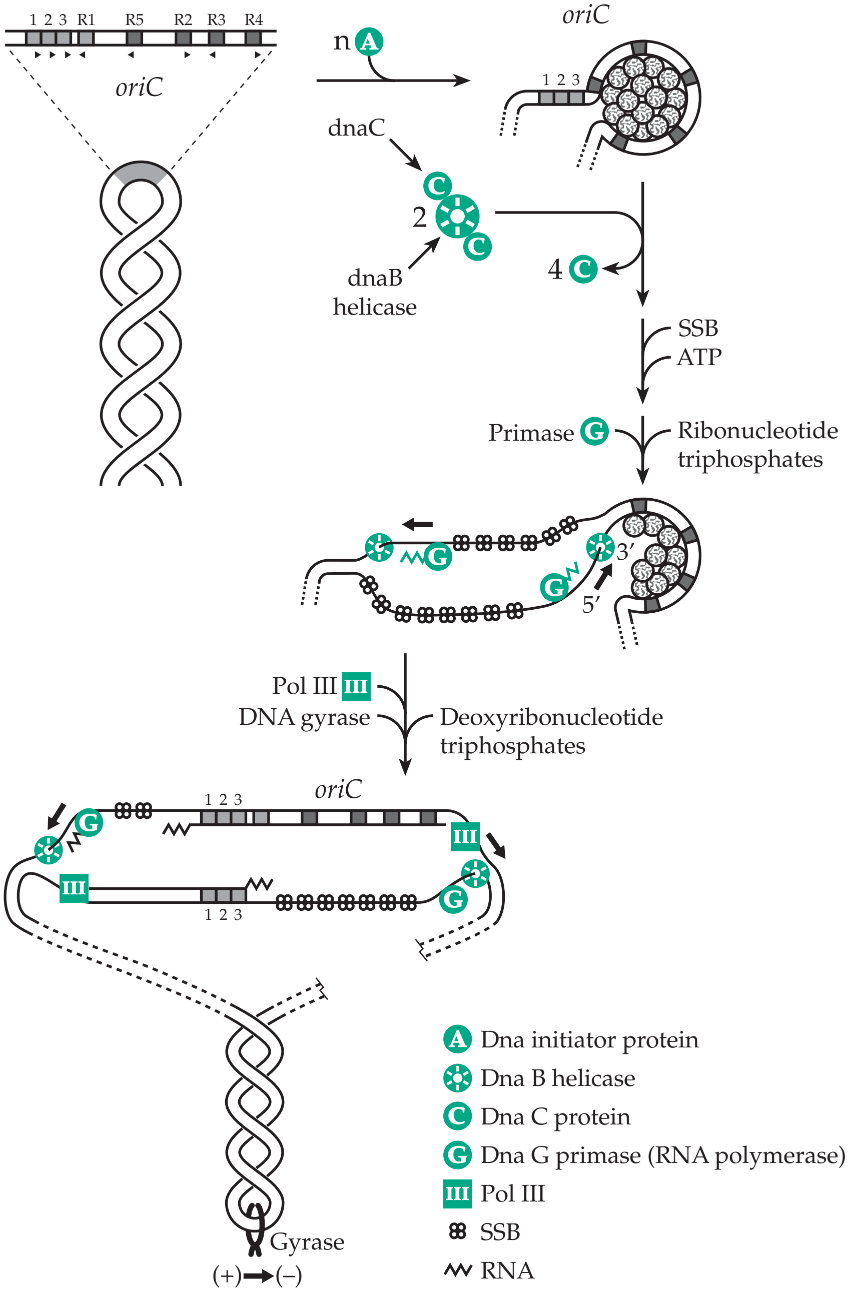


Figure 27-19 Hypothetical scheme for initiation of bidirectional replication in *E. coli*. The closed boxes R1 through R5 represent the 9-residue recognition sequences for the *E. coli* dnaA protein. The open boxes 1, 2, and 3 represent the three 13-residue repeats, possible sites for binding of the dnaB-dnaC protein complex. From McMacken *et al*.[383] Redrawn in simplified form.

It has long been postulated that the bacterial chromosome is attached to the plasma membrane. At least one such attachment site may be at or near the origin of replication.[393] Furthermore, the exchange of ADP for ATP in the dnaA protein is catalyzed specifically by cardiolipin and phosphatidylglycerol containing the unsaturated oleic acid.[386,393] Inositol polyphosphates may also play a role.[394]

Elongation of DNA chains.

DNA polymerase III in its holoenzyme form is the major polymerase for DNA replication. It elongates the primer chains rapidly and processively leaving only very small gaps at the ends of single-stranded regions. The rate of elongation, which is ~3 nucleotides / s for 8 kb *oriC* plasmids, may be determined by the rate of action of the dnaB helicase. A completely unwound *ori*C plasmid, bound to SSB, undergoes primer elongation ~ 10 times faster.[383] However, the rate for the intact *E. coli* replisome is nearly 1000 nucleotide s^{-1} with a rate of misincorporation of only one in 10^9 nucleotides.[394a]

Small *ori*C plasmids need to be primed at only one location, but the large bacterial chromosome must undergo priming at many sites on the lagging strand to permit DNA polymerase III to act on that strand with formation of the Okazaki fragments. The DNA polymerase complex may be a dimer that works on both strands at once. The lagging strand may be looped out to allow it to lie parallel with the leading strand (Fig. 27-20). The appearance of the electron micrograph in Fig. 27-20B supports this suggestion. However, the manner in which the lagging strand can be shifted to bring the next primed initiation site to the replication complex is not clear. DNA polymerase III holoenzyme itself may be organized as an asymmetric oligomer[395] that operates on both strands in a complex such as is shown in Fig. 27-20. The primase, either the dnaG protein[265,383] or the primosome used by phage ϕX174 (Section 5),[396] may synthesize the RNA primers on the lagging strand.

Termination of replication. As each Okazaki fragment is completed along the template for the lagging strand, the RNA primer piece is digested out, replaced by DNA, and the nick sealed by action of DNA ligase. Ribonuclease H, which is found in both bacteria and eukaryotic cells, specifically degrades the RNA component in these RNA–DNA hybrid regions.[396a] In bacteria another mechanism for primer removal is available. The 5'- 3' exonuclease activity of DNA polymerase I will cut out the RNA segment, while the 5'- 3' polymerase activity of the same enzyme will fill the gap.

Replication of *oriC* plasmids may occur by simply allowing replication of the leading strands at both replication forks to continue all the way around the circle.[397] However, in *E. coli* bidirectional replication continues only until the two replication forks converge. This can occur anywhere between two **terminators**, T1 and T2, located at 28.1 min and 35.6 min. The terminators slow replication in the counterclockwise and clockwise directions, respectively. A gene (*tus*) near T1 encodes a **terminator utilization substance**, a DNA-binding protein that associates with T1 and T2 and causes termination.[398–401] Another problem may be the separating of catenated DNA circles by action of a topoisomerase. Finally, it is essential to **partition** the original chromosome and its replica, one to each daughter cell. This requires at least three other gene products including one large 170-kDa protein.[402]

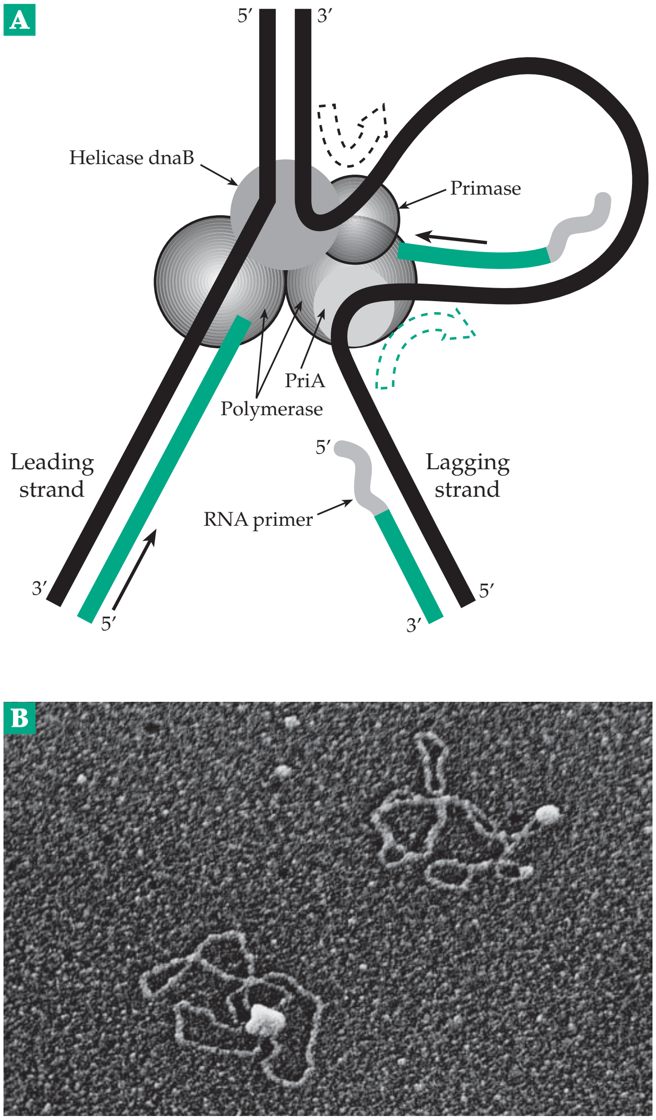


Figure 27-20 (A) Hypothetical replisome for concurrent replication of leading and lagging strands by a dimeric polymerase associated with helicase dnaB and a primosome. Open arrows indicate directions of movement of DNA, which is forming a loop as the polymerase fills a gap to complete an Okazaki fragment. The primase will then form a new primer and a new loop. From Kornberg and Baker.[265] (B) Electron micrograph of the primosome bound to covalently closed ϕX174 duplex replicative form. These enzymatically synthesized duplexes invariably contain a single primosome with one or two associated small DNA loops. From A. Kornberg in Hubscher and Spadari,[266] pp. 9,10.

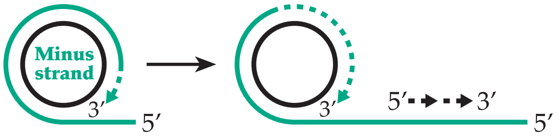
5. The Replication of Viral DNA

The replication of viral DNA usually depends upon the genes of both the host organism and the virus. For example, *ts* mutations in the *E. coli* genes *dnaB*, *D*, *E*, *F*, and *G* lead to a loss in ability to support growth of phage λ as well as loss of ability to reproduce under conditions where the *ts* gene products are inactivated. However, the phage can replicate in *E. coli* with mutated genes *dnaA* and *C* because phage λ encodes its own initiator proteins by genes *O* and *P* (marked on the gene map in Fig. 26-4). In addition to these two proteins, seven *E. coli* proteins are required to initiate replication at the lambda origin *ori* λ and to complete replication.[403–407] The *E. coli* dnaB helicase and the dnaC protein are needed, as in Fig. 27-19. The chaperonins dnaJ, dnaK, and GrpE are also necessary for replication of phage λ and other viruses.[408,409] As we have seen (Section 2) many viruses contain genes specifying their own DNA polymerases and primases, which function in cooperation with host proteins. For many dsDNA viruses the origins of replication, priming reactions, and chain elongation processes closely resemble those of *E. coli*.[363,410]

In contrast, the first step in replication of the filamentous Ff viruses (f1, fd, and M13) or the small icosahedral ϕX174 or G4 is conversion of the single-stranded closed circular DNA molecules of the infecting virus particles into circular double-stranded **replicative forms** (RF).[265,411,412] This occurs as the DNA enters the bacterial cell and is accomplished entirely by the enzymes of the host cell. Phage G4 DNA, whose replication has the simplest known requirements, contains a tight hairpin region at its origin. The rest of the DNA must be coated by SSB for replication to occur. The hairpin resists melting and serves as a binding site for *E. coli* primase. This is the only known case in which no other priming proteins are needed.[265] Primase synthesizes up to a 29-ribonucleotide primer after which DNA polymerase III holoenzyme copies the rest of the chain.

Replication of the closely related ϕX174 is more complex. It requires assembly of a **primosome** made up of at least seven host proteins[265,371,413] : dnaB, dnaC, primase (dnaG), and proteins priB, A, C (n, n′, n″), and dnaT (i), Table 27-2. The 76-kDa helicase priA (n′) may locate the **primosome assembly sequences**, which are ~70 nucleotides in length,[414] and displace SSB from them. These sequences can adopt secondary structures with a pair of hairpin loops.[265] Kornberg and associates suggested that the same kind of primosome formed at these sites in the ϕX DNA may participate in replication of the lagging strand of the chromosomal DNA. If so, helicase priA presumably functions in the replisome on one strand and dnaB helicase on the other as depicted in Fig. 27-20.

In the second and third stages of replication ϕX174 RF molecules are themselves replicated and are then used for synthesis of new viral (+) strands. At both stages a virally encoded **gene A protein**, which has endonuclease activity, nicks the duplex. Cutting the (+) strand it leaves a free 3'-OH on DNA residue 4305, while the 5'-phospho group of residue 4306 becomes covalently attached to a tyrosyl residue in the A protein.[411,415] The free 3'-OH serves as the primer for a **rolling-circle synthesis** (Eq. 27-7).[412,416,416a] As a new viral strand is synthesized along the complementary (-) strand as a template, the original viral DNA (+) strand is displaced (Eq. 27-7) as a single-stranded tail.



(27-7)

A strand complementary to the single-stranded tail is then formed in segments. A complete turn of the circle produces a viral strand twice the normal length. Cleavage by the endonuclease activity of the gene A protein and closure of the circle completes the replication. The displaced (+) strand can be cut off and either incorporated into a progeny phage or converted into another RF circle. The A protein, attached to the 5'-terminus of the (+) strand, is involved in either case. It can participate in repeated sequences of initiation and termination of viral (+) strand synthesis.[265,411] Once double-stranded circles are formed, they undergo several replications to give additional RF circles, which serve as templates for the synthesis of many single strands of viral (+) DNA, which are incorporated into the mature viruses. This synthesis of additional RF circles requires transcription of some viral RF genes.

In the final stage of replication the single-stranded (+) chains formed by the rolling-circle mechanism are packaged into phage particles. The gene 5 single-strand binding protein of M13 coats the DNA chains as they are formed, evidently preventing their conversion to RF circles. In the case of ϕX174 the new single-stranded DNA circles are packaged as they are synthesized (Fig. 7-28) to form complete icosahedral virus particles. See Kornberg and Baker[265] for details about these and many other virus replication systems.

Replication of the larger tailed viruses, which have many genes, is complex and varied. The lytic phage λ resembles the smaller viruses in using the host replication enzymes.[265,417,418] In the final stages a rolling-circle mechanism is utilized to form **concatemers** consisting of linear DNA duplexes with numerous successive copies of the viral DNA. The ssDNA that is formed in the rolling circle is converted to dsDNA as it is formed. Finally a **terminase** cleaves the DNA at specific *cos* sites, using staggered cuts, to form cohesive ends.[417,419–420a] However, there are uncertainties.[421] The linear dsDNA enters an empty preformed procapsid, apparently pumped in an ATP-dependent fashion, perhaps by a rotating portal ring.[421a]

6. Packaging of Viral Genomes

The construction of intact virus particles from the genomic DNA and protein subunits is often a complex process. It is simplest for the small filamentous ssDNA viruses (Fig. 7-7). The subunits are synthesized as soluble proteins, which enter the cell membrane, then lose their leader sequences. As the viral DNA coated by the viral gene 5 ssDNA-binding protein enters the membrane, the binding protein is replaced by the coat subunits.[422–424]

The process is somewhat more complex for the icosahedral viruses. In the ϕX, G4, α3 family the icosahedral procapsid is constructed with the aid of both internal and external **scaffolding proteins**[425,426] as is illustrated in Fig. 7-28. In phage that replicate via concatameric dsDNA the terminase that cleaves the DNA also interacts in a precise way with the packaging apparatus of the prohead.[427–429] For the tailed phage the ring-shaped oligomeric head–tail connector (Fig. 7-29), together with an ATPase, may function as a rotatory pump to feed the DNA into the prohead.[430–430c] This has been demonstrated for phage ϕ29.[430a–c] In some cases the terminase produces new phage DNA of unit size, but in other cases, e.g., with phage T4, the DNA may be cut more randomly when the head is full or when another piece is needed to fill it.[431] After the DNA is packaged the virus capsids usually expand and become stronger.[432,432a]

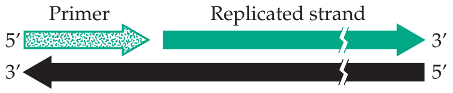
7. Plasmids

Most bacteria contain plasmids which are self-replicating but stably maintained at well-defined numbers of copies per cell.[265,433] They are usually not essential to the cell but may carry traits such as antibiotic resistance or toxin formation that benefit the bacterial host. A plasmid always carries in its DNA an orign (*ori*) of replication and a gene, usually designated *rep*, for an initiator protein. It usually encodes other proteins as well but may depend largely on host proteins for replication. Plasmids may use the *oriC* copied from the bacterial host's DNA, the origin from phage λ, yeast autonomously replicating sequences (ARSs, Section 10), or other origins. Replication of the small 6.6-kb plasmid ColE1, which is present at ~20 copies per cell, depends entirely on the host-cell replication machinery.[265] However, the control of copy number depends upon synthesis of **antisense RNA** and its reaction with the plasmid DNA (see also Chapter 28, p. 1615).[434,435] Similar copy number control is used by the larger ~100-kb resistance factor **R plasmids**.[436] Some plasmids use replication systems very similar to those of viruses such as ϕX174, often using rolling-circle replication.[437] However, the plasmids lack the proteins for virus coat formation and maturation.[438]

The F factor plasmids, discussed in Chapter 26, are large 100-kb circular DNA molecules containing ~60 genes, about 20 of which encode proteins involving transfer of DNA into another bacterial cell (Fig. 26-3).[265,439,440] F plasmids display strict copy number control with only 1–2 copies per host chromosome. The controls lie in a region known as the partition locus, which resembles regions of the host chromosomes that are involved in partition of the bacterial genome. They have repetitive sequences suggesting a similarity to centromeres of eukaryotic chromosomes.[441]

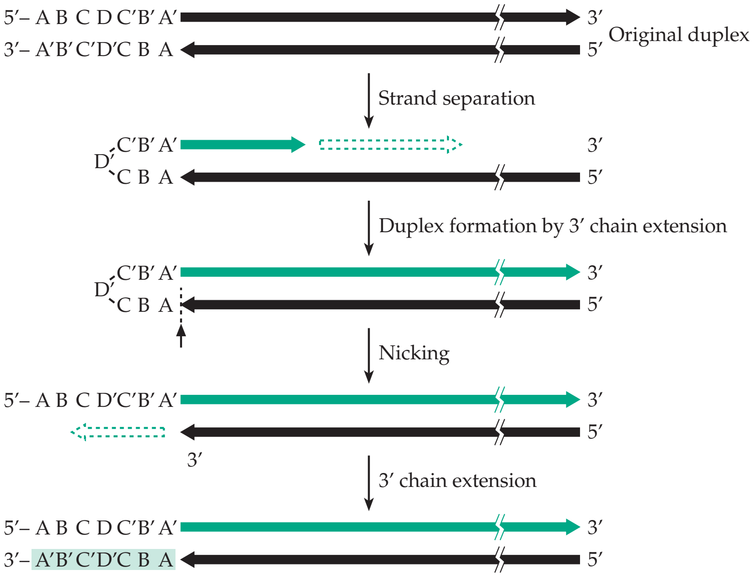
8. Chromosome Ends

The T-odd bacteriophages T1, T3, T5, and T7 are medium-sized phage with linear duplex DNA genomes. Replication of linear DNA in these and in many other genomes presents a problem. Even if the RNA primer segment is made at the very 3' end of the template strand, there will be a gap in the final replicated strand when the primer is digested out. Since there is no known enzyme that will add to the 3' end of a chain, this gap will remain unfilled. The problem is solved by **terminal redundancy**, the presence of a common 260-nucleotide sequence at both ends. Several daughter DNA molecules with gaps at the 5' ends can be joined by their cohesive ends to form a long **concatamer**. DNA polymerase I fills any gaps, and the chains can then be ligated and cleaved at different points to generate complete 5' ends.[442]



Another mechanism, which is utilized by some single-stranded parvoviruses,[416,443] obviates the need for an RNA primer by use of a palindromic sequence to form a hairpin loop (Eq. 27-8).

Yet another solution to this problem is used by some viruses. **Phage ϕ29** of *Bacillus subtilis* primes the replication of its 19,285 bp dsDNA at both ends by a **terminal protein,** which is linked covalently through its Ser 232 –OH group to dAMP.[443a] The 3'–OH of the deoxyadenosyl group primes the DNA replication. In a similar fashion replication of the eukaryotic **adenoviruses**, whose genome is a 35- to 36-kb linear DNA duplex, starts at the ends and is primed by one residue of dCMP covalently attached through a 5'-phosphodiester linkage to a serine side chain in a 80-kDa **preterminal protein**. It substitutes for the RNA oligonucleotides that prime most DNA synthesis.[444–447] The dCMP pairs with guanine at the 3' terminus of the template strand and provides the initiating 3'-OH group. During the replication the


5′– A B C D C′B′ A′
3′
3′– A′B′C′D′C B A
5′
Original duplex
Strand separation
C′B′ A′
D′
C B A
3′
5′
Duplex formation by 3′ chain extension
C′B′ A′
D′
C B A
3′
5′
Nicking
5′– A B C D′C′B′ A′
3′
5′
3′
3′ chain extension
5′– A B C D C′B′ A′
3′
3′– A′B′C′D′C B A
5′

(27-8)

preterminal protein is cleaved to the 55-kDa **terminal protein,** which remains covalently attached, one molecule at the 5' end of each strand. The genome can be replicated *in vitro* by five proteins: the virally encoded preterminal protein, DNA polymerase, DNA-binding protein, and two cellular transcription factors that bind in the adenovirus origin region.

The chromosome end problem is solved in another way in eukaryotes. As discussed in Section B,1, **telomeres** (chromosome ends) contain repeated sequences of variable length. One DNA strand is always G-rich. For example, in human cells the sequence 5'-(TTAGGG)$_n$ –3', where *n* may be ~20, occurs at the 3' ends of the G-rich strands. The other strand, whose 5' end is at the telomere end, is C-rich and has the complementary sequence 3'-(AACTCCC)$_n$ –5'. The 3' end of the G-rich strand is always longer by 12–16 nucleotides than the end of the C-rich strand. This 3' extension may fold back to form non-Watson–Crick structures that apparently involve G-quartets (Chapter 5).[447a,b] The shorter 5' end is thought to result from the need for a short RNA primer during replication. As shown in Fig. 27-21, when replicated in the normal fashion the full G-rich leading strand will be formed, but the C-rich lagging strand will be 8–12 nucleotides short, when the RNA primer is digested away. A result of this is that human somatic cells gradually lose telomeric repeats. However, in tumor cells, germline cells, and unicellular organisms the enzyme **telomerase** prevents this telomere loss.[448–449a]

Telomerase is a reverse transcriptase that copies the DNA sequence of the telomeric repeats from a small **guide RNA** that is part of the enzyme. The first telomerase studied was the relatively abundant enzyme from *Tetrahymena*. It contains a 159-nucleotide RNA with the sequence 5'-CAACCCCAA-3' at positions 43–51. This sequence is complementary to the 5'-TTGGGG-3' repeat sequence of the *Tetrahymena* telomeres.[450,451] A 127-kDa human protein contains a similar guide RNA with the sequence 5'-CUAACCCUACC-3', which is complementary to the human telomere repeat sequence as is illustrated in Fig. 27-21.[452–454] Telomerases[455,456] evidently allow the cell to elongate the telomere 5'-ends using the

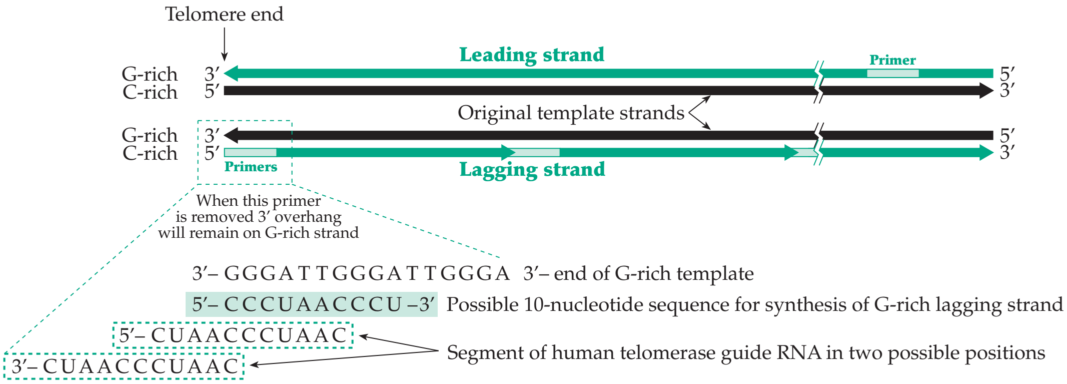

Telomere end
G-rich
C-rich
3′
5′
Leading strand
Primer
5′
3′
Original template strands
G-rich
C-rich
3′
5′
Primers
Lagging strand
5′
3′
When this primer
is removed 3′ overhang
will remain on G-rich strand
3′– GGGATTGGGATTGGGA 3′– end of G-rich template
5′– CCCUAACCCU –3′ Possible 10-nucleotide sequence for synthesis of G-rich lagging strand
5′– CUAACCCUAAC
3′– CUAACCCUAAC
Segment of human telomerase guide RNA in two possible positions

Figure 27-21 Aspects of telomere synthesis. The end of the chromosome and the 5' end of the C-rich strand is at the left. This is the template for replication of the leading G-rich strand (green). The primer lies far back in the chromosome. The C-rich strand is replicated in segments from several RNA oligonucleotide primers, one of which lies at the 5'-terminus. This first primer is removed by RNase activity leaving a 12–16 nucleotide 3'-overhang. The telomerase guide RNA can hybridize with the 3'-end of the G-rich strands providing a template that allows additional growth of the G-rich strand and extension of the C-rich strand also in the next replication.

guide sequence and the reverse transcriptase activity of the telomerase. Any number of additional repeats may be added to the 5' ends. The shortened 3' ends can also be lengthened in the next round of replication.

The control of telomerase must be important. The enzyme is active in early embryonic cells and some stem cells. However, most normal cells have little or no telomerase activity and lose telomere length throughout their lifespan with eventual growth arrest.[448,454,457] On the other hand, excessive telomerase activity may induce cancer.[451,458,459] Certain mutations in the telomerase guide RNA can cause greatly increased telomerase activity.[459] The control of telomerase is still poorly understood but involves specific telomere-binding proteins.[448,454,460–460b]

9. Mitochondrial and Chloroplast DNA

Replication of the ~16-kb mammalian mtDNA begins with RNA priming within a small **displacement loop** or D-loop. One daughter strand, the heavier or H strand, starts to grow on the primer. As it does, the parental H strand is displaced and the D-loop is enlarged. The H-strand grows until ~70% of the parental H strand has been displaced and the L-strand origin (Fig. 18-3) is uncovered. Then a new light L strand is laid down to form the second daughter duplex. The rate of formation of the new L strand is only 10 nucleotides / s, an hour being required to complete the process. The DNA formed is initially relaxed, another 40 min being needed to introduce the 100 superhelical turns present in the finished chromosome.[461]

The kinetoplast DNA of trypanosomes (Fig. 5-16) consists of thousands of catenated circular DNA molecules. Among these the smaller minicircles always contain the sequence GGGGTTGGTGTA at their origins of replication. The minicircles are individually removed from the mass prior to replication. The two progeny circles are then both recatenated into the mass.[462]

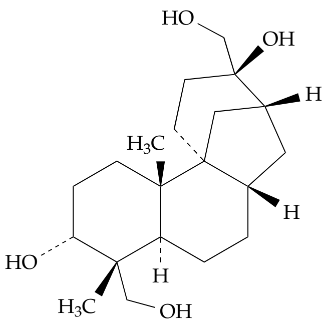
Chloroplasts contain large 120- to 169-kb circular genomes encoding about 100 proteins (Chapter 23). A characteristic feature of most chloroplast DNA is the presence of long inverted repeat sequences (10,058 bp in the liverwort, 25,339 bp in tobacco).[463,464] These are separated by 19,813 and 81,095 bp single copy regions in the liverwort and by similar sized regions in tobacco. Plastid DNA exists as a mixture of monomeric molecules with smaller amounts of dimers, trimers, and tetramers.[464]

Ethidium bromide inhibits the replication of chloroplast DNA and causes partial degradation of existing DNA in chloroplasts without interfering with replication of DNA in the nucleus. The effect is similar to that of the same drug on mitochondrial DNA. However, cells of *Chlamydomonas* treated with ethidium bromide are able later to regenerate their chloroplast DNA. This result has been interpreted to mean that there may be one or a few “master copies” of chloroplast DNA in specially protected locations. The result should also be considered in relationship to the following observation. Although nuclear and organelle DNA molecules replicate at different times in the cell cycle, constant proportions of the organelle and nuclear DNA tend to be maintained. Thus, there must be some kind of control mechanism leading to a coupling of DNA replication in nuclei, mitochondria, and chloroplasts.[465]

10. Replication of Eukaryotic Nuclear and Viral DNA

Replication in eukaryotes is similar in many ways to that in bacteria.[284,466–467a] However, the ~10^6 kb of DNA in a typical eukaryotic genome is divided into many **replicons**, segments of DNA 30–150 kb in length, each having its own origin of replication. In the relatively small ~14-Mb yeast genome there are ~400 replicons,[468,468a] but in mammalian DNA and also in plant DNA[464] there are probably thousands. DNA synthesis is initiated at different times during the S-phase of the cell cycle at the various origins in an ordered pattern.[469,469a,b] An important unanswered question is how the cell is able to replicate all segments of all of the chromosomes just once before entering mitosis.

Many of the proteins of eukaryotic replication are closely related in sequences and functions to those of bacteria. There are initiator proteins analogous to *E. coli* dnaA (Fig. 27-19). The DNA polymerases have been discussed in Section C, 2 (see Table 27-1). Eukaryotic polymerases α and δ and possibly ε are essential for replication.[470,471] Polymerase α, which is inhibited specifically by the fungal metabolite **aphidocolin**,[472] is a complex of a ~170-kDa DNA polymerase core, an RNA-synthesizing primase consisting



Aphidocolin

of 58- and 49-kDa subunits, and a 70-kDa subunit of uncertain function.[473–474a] The complex makes an RNA–DNA primer consisting of ~10 nucleotides of RNA and ~30 of DNA.[475] This pol α/primer is replaced early in replication by the highly processive polymerase δ and, perhaps, under some conditions by polymerase ε. The ringlike processivity factor or “clamp” that is provided in *E. coli* by protein dnaB is called the **proliferating cell nuclease antigen** (**PCNA**) in eukaryotes.[476] It is loaded onto the DNA by a **clamp loader**, the 5-subunit replication factor C (RFC).[476a] As in *E. coli* (Fig. 27-19) an SSB type protein known as **replication protein A** (**RPA**) is also essential.[467,477] PCNA is not only essential to eukaryotic replication but is also required for recombination and repair.[476] The polα/ primer primes leading strand synthesis initially but then switches to replication origins on the lagging strand where, together with other proteins, it primes the formation of the Okazaki fragments. Polα/ primer is also a logical participant in the control of the initiation of the S-phase of the cell cycle.[478]

Eukaryotic viruses. Investigation of viruses provided the first insights into eukaryotic DNA replication. Most of the factors needed for replication of the DNA of adenoviruses, simian virus 40 (SV40), and polyomavirus[447] within animal cells are supplied by the host. Replication of the 5-kb SV40 DNA, whose DNA forms typical nucleosomes (Fig. 27-3), appears to be an excellent model for eukaryotic replication in general.[479–482] The single SV40 origin of replication is a 64-bp sequence containing the 5-bp sequence GAGG C four times as pairs of inverted repeats. These are recognized by the 95-kDa virally encoded initiation protein which also has helicase activity and is known as the **T antigen**.[483,483a] A nearby 17-bp sequence containing only AT pairs is presumably the region of entry of the host cell’s polymerase α/ primer. Single-stranded DNA regions are coated with the replication protein A. After the primer is formed, the RFC complex loads the sliding clamp PCNA, and polymerase α is replaced by polymerase δ on both leading and lagging DNA strands permitting highly processive bidirectional chain elongation. Topoisomerase activity is required to decatenate the replicated chromosomes. Since SV40 DNA forms typical nucleosome (Fig. 27-3), its replication is thought to mimic chromosomal replication quite closely. The more complex herpes simplex virus HSV-1 has a 153-kb genome, a linear DNA duplex. It has ~75 genes and encodes its own DNA polymerase, origin-binding protein, SSB, and other proteins needed for replication within eukaryotic cells.[484,485]

Artificial chromosomes. Another approach to understanding eukaryotic replication, similar to the use of *ori*C plasmids in *E. coli*, is to study **autonomously replicating sequences** (**ARSs**)[469,486] and plasmids[468] and **artificial chromosomes**[487] made from them. ARS sequences were first found in the budding yeast *S. cerevisiae*. Plasmids containing an ARS, whose core consensus sequence is 5'-(A/T)TT TAT(A/G)TTT(A/T), replicate autonomously during S-phase. Such plasmids have been genetically engineered, providing them with telomeres and some kind of functional centromere, to form artificial chromosomes. **Yeast artificial chromosomes** (**YACs**) have become extremely important as cloning vehicles (Chapter 26), and they also serve as important tools for studying eukaryotic replication and its control. They can be cultured in yeast cells or can be transferred into animal cells, etc.

As mentioned in Section B, 1, human centromeres are rich in the repetitive α-satellite DNA. By joining α-satellite DNA-containing fragments of the X-chromosome to cloned telomeric DNA, human **minichromosomes** have been created.[488] These have been developed into **human artificial chromosomes**,[489] which may be practical vehicles for gene transfer in human therapy.

Replication of nuclear DNA. The budding yeast *Saccharomyces cerevisiae* has permitted the most detailed picture of DNA replication in a eukaryote. The complete genome sequences are known and the ARSs have been physically mapped.[468a] For example, in chromosome VI there are nine origins that differ in frequency of initiation and which replicate sequentially during the S-phase of the cell cycle.[490] The initiation (**replicator**) regions surround the 11-bp consensus sequence of the ARSs, each occupying at most ~150 bp. However, in metazoa and even in the fission yeast *Schizosaccharomyces pombe* the ARSs range from 500 to 1500 bp in length. These origin regions frequently overlap the **promoter** sequences, which control initiation of transcription (Chapter 28).[491] This association with transcription origins has also been observed in metazoan cells, where replication origins are often clustered.[492] However, there is no sequence homology between the ARSs of *S. cerevisiae* and replication origins in other species, even those of *S. pombe*.[490]

The study of replication in yeast ARSs and artificial chromosomes has revealed that initiation of replication requires not only an initiator protein but a complex of six proteins that form an **origin recognition complex** (**ORC**).[493–495a] This complex, which is essential to initiation of replication, may be joined by additional proteins in a **prereplication complex**. At least some of the ORC proteins have their homologous counterparts in metazoa, suggesting a highly conserved initiation machinery.[495a–c]

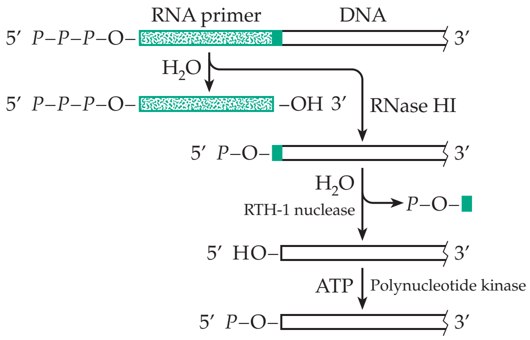
One or more of the proteins that bind to the ORC may constitute a **license** to replicate. The licensing concept states that when replication occurs the license is destroyed and the origin involved cannot initiate replication again without a new license. The **replication licensing factor** (**LRF**) is postulated to be unable to pass through the nuclear membrane.[494,496] It can only reach the replication origin after the S-phase has concluded and mitosis has taken place. At this time the membrane has been disrupted. A second signal, the **S-phase promoting factor** (**SPF**), cannot act without an intact nucleus and a license in place.[469,495,497–500] This system ensures that DNA is replicated only once per cell cycle. Among the proteins involved in the licensing is a group of **minichromosome maintenance** (**MCM**) **proteins**, so-named because of their importance to replication of ARSs and artificial chromosomes.[495] Six of these proteins (MCM2–MCM7) can form a hexameric complex with one subunit of each type as well as other complexes, e.g., $(MCM4,6,7)_2$.[500a–c] The latter acts as an ATP-dependent helicase. A somewhat simpler MCM complex is found in archaea.[500d] Some of the cell cycle proteins (Chapter 26, Section F,1), including Cdc6 and the protein kinases Cdc7 and Cdc28 as well as other proteins, are also required for regulation of replication origins. Proteins homologous to those of yeast have been identified in humans and other eukaryotes. Licensing of replication involves association of the MCM helicases with each ORC during the G1 phase of the cell cycle. Binding of the initiation factor Cdc6/18 and of a recently discovered loading factor **Cdt-1** apparently completes the licensing. Once licensing has occurred both cdc6/18 and Cdt-1 can dissociate from the DNA.[496] Removal of Cdt is facilitated by its binding to another protein **Geminen**, found first in the frog *X. laevis.*[500e–g] The ORC complexes may remain at the origins. It has been estimated that a yeast cell contains ~400–600 molecules of the very stable ORC, about one ORC per replication origin. However, a large excess of MCM proteins may be present. Their concentrations may regulate the number of ORC molecules that associate with DNA. Replication of DNA is not the only aspect of cell growth. For example, as DNA is replicated histones must be synthesized and assembled. This synthesis occurs during the S-phase and is tightly coupled to replication.[500h]

Initiation of replication in metazoans is still confusing. Almost any piece of DNA will be replicated if introduced into a *Xenopus* egg, where initiation appears to occur just once at a random position.[501,501a,b] However, in differentiated tissue the origins of replication seem to be fewer in number and more specifically located. A possible explanation is that high concentrations of ORC and MCM proteins in the embryo may lead to many relatively nonspecific origins and a replicon size of ~7 kb. The lower concentrations of these auxiliary factors in somatic cells may lead to fewer but more specific origins with a replicon size of ~170 kb.[495] Three distinct mammalian origins have been studied in detail. One is in the β globin locus (Fig. 27-10).[500] A second is near the dihydrofolate reductase gene.[502] A third, which is activated early in S-phase, is at the 3' end of the lamin B2 gene.[503] The latter has been localized to a 500-bp region. These findings suggest that the replicon concept, as developed for yeast, may be generally applicable.

Replication of the intact genome of *Drosophila* has been studied in rapidly dividing nuclei by electron microscopy.[504] The replication rate in these nuclei is ~300,000 bases / s, but it has been estimated that replication forks in animal chromosomes move no faster than ~50 bases / s. Thus, we would anticipate at least 6000 forks, or one fork per 10 kK bases. Indeed, this number of forks has been observed.[505] They occur in pairs with many short regions containing single-stranded DNA as if one strand at the fork is replicated more rapidly than the other as in mitochondrial DNA. The arrangement of the ssDNA regions at the two forks in a pair suggests bidirectional replication. However, replication forks are rarely seen in higher eukaryotes, but extensive regions of single-stranded DNA are often visible. Benbow and associates suggested that in higher eukaryotes the strands of duplex DNA may be separated throughout a whole looped domain of DNA. Replication could then occur with initiation at many points along each strand.[505]

Replication reactions are similar in bacteria and eukaryotes, but some details differ. In eukaryotes at least two DNA polymerases, α and δ, are required. In budding yeast polymerase ε is also essential.[506] Both polymerases δ and ε may replicate separate strands at the replication fork.[506a] Processing of Okazaki fragments also differs from that in bacteria, where either RNase H or the 5' to 3' exonuclease activity of DNA pol I removes the RNA primer (Fig. 27-14). This exonuclease activity is lacking in eukaryotic polymerases. Replication primers are removed in a two-step process by **RNase HI**, which makes an endonucleolytic attack that removes all but one nucleotide residue of the primer in a single piece, leaving a 5'-phospho group on the remaining ribonucleotides. That residue is removed by a 5' to 3' exonuclease designated RTH-1 nuclease (Eq. 27-9).[467,507,507a,b] This is a homolog of the yeast RAD27 protein. A polynucleotide kinase may then phosphorylate the 5' end of the DNA fragment.[508]

Another difference between bacterial and eukaryotic replication is the presence of nucleosomes in eukaryotes. Some evidence suggests that nucleosomes may open and close to allow replication forks to pass through.[509] Studies of SV40 minichromosomes indicate that passage of the replication machinery does destabilize nucleosomes, which must be partially reconstructed about 260 nucleotides past the elongation point.[510] Another factor is the variable extent and location of modifications to histones, in particular to



(27-9)

the H3 and H4 histone tails (Section A,3). A code has been proposed according to which certain modifications would favor transcription or mitosis, while lack of modification would silence the genes.[72]

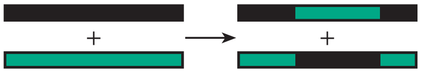
Much of the control of replication is at the initiation stage. Growth factors and other mitogenic stimuli acting at the plasma membrane can stimulate expression of such nuclear proteins as those encoded by the proto-oncogenes *c-myc*, *c-myb*, and *c-fos*. These may initiate a regulatory cascade (Fig. 11-13) and trigger mitosis.[511] As indicated in Chapter 26, the 34-kDa protein kinase encoded by fission yeast gene *cdc*2 (budding yeast CDC28) is essential for progression of the cell cycle through the G1 phase into mitosis (Eq. 26-3). A single oscillation in this kinase activity induced by a B-type cyclin can promote both replication and mitosis. However, in *S.cerevisiae* there are 14 different cyclinlike proteins, and their individual functions are not clear.[512] The signal that is sent to the ORCs is likewise unclear.[513] However, theoretical models involving Eq. 26-3 and many additional components have been proposed.[514] Multiple phosphorylations may occur, some on the RPA initiator protein.[515] Many proteins required for replication, including DNA polymerase and primase, are associated with the nuclear matrix.[516] The nuclear membrane may also be important in controlling replication.

D. Integration, Excision, and Recombination of DNA

The exchange of genetic information between chromosomes, plasmids, and viruses occurs in many ways, which are described collectively as recombination.[517–520] Mutants of *E. coli* deficient in recombinational ability often have defects in genes designated *recA, B, C,* etc. (for recombination), or *ruvA, B, C . . .* (for resistance to ultraviolet light). Some of these mutants are unusually sensitive to ultraviolet light because of their inability to repair damage to DNA. Several of the recombination enzymes are used for repair of ultraviolet damage and of double-strand breaks in DNA arising from other causes.[521,522] In eukaryotes recombination occurs during meiosis. Many viruses, including phage, also carry genes for their own general recombination systems. In addition, the DNA of some viruses, such as the temperate phage λ, undergoes recombination with the host DNA. This can happen during the processes of integration of the viral DNA into the host genome or excision of the viral DNA during the lysogenic cycle of replication (Chapter 28, Section B). Recombination occurs around specific sites in the chromosomes of both the virus and its host and is called **site-specific recombination**. Genetic recombination is essential to the development of genomic diversity, to the survival of a species, and to evolution.

1. Recombination Mechanisms

How can the homologous regions of two different DNA duplexes be brought together? As illustrated schematically in Eq. 27-10, the strand exchange must occur at exactly the same point in each duplex. An early attempt to explain this postulated a "copy choice" mechanism of replication. It was assumed that replication occurred along one DNA strand up to some random point at which the polymerase jumped and began to copy from the second of a pair of homologous chromosomes. The newly formed DNA molecule would be complementary to different parts of both parental DNA duplexes. To test the idea, Meselson and Weigle infected *E. coli* with two strains of phage λ containing ^{13}C- and ^{15}N-labeled DNA, respectively.[523] Recombinant DNA was found to contain some ^{13}C and some ^{15}N, as judged by density gradient centrifugation. It was clear that DNA from both parents was incorporated into DNA of recombinant progeny, a finding that ruled out the copy choice hypothesis.



(27-10)

If recombination occurs instead by enzymatic cutting of two homologous duplex DNA molecules followed by rejoining, how is it possible to avoid inactivation of genes by addition or deletion of genetic material? Recombination cannot depend upon the random action of a nonspecific enzyme with random rejoining. Yet, general recombination can occur at any point and with a roughly constant frequency throughout the DNA chain. The explanation of these facts lies in the occurrence of base pairing between at least some short homologous regions of strands of the two different DNA duplexes.

The Holliday recombination intermediate. In 1964, Holliday suggested a recombination process that would give rise to characteristic H-shaped intermediates.[524] Recombination could be initiated at special points on the duplexes, recognizable by a recombination enzyme (Fig. 27-22). A short amount of unraveling would be followed by strand exchange with the two broken strands being rejoined by a ligase as indicated in Fig. 27-22. The crossover points would then migrate up or down the chains as the two helices turned about their own axes. Long regions of heteroduplex DNA could be generated in this way, and the process could be terminated at a random distance from the starting point, accounting for the observed uniformity of genetic recombination events. Chain cleavage and rejoining of two of the strands would terminate the process. If these were the same strands broken in the initiation event (cleavage at points *aa′* in Fig. 27-22), genes lying outside the heteroduplex region would not be recombined, but cleavage of the other chains (at points *bb′*) would lead to their recombination. Intermediates of the type predicted by the Holliday model were soon observed by electron microscopy (Fig. 27-23).[525] Three-dimensional structures have been determined by X-ray crystallography[525a] and have been studied by atomic force microscopy.[525b]

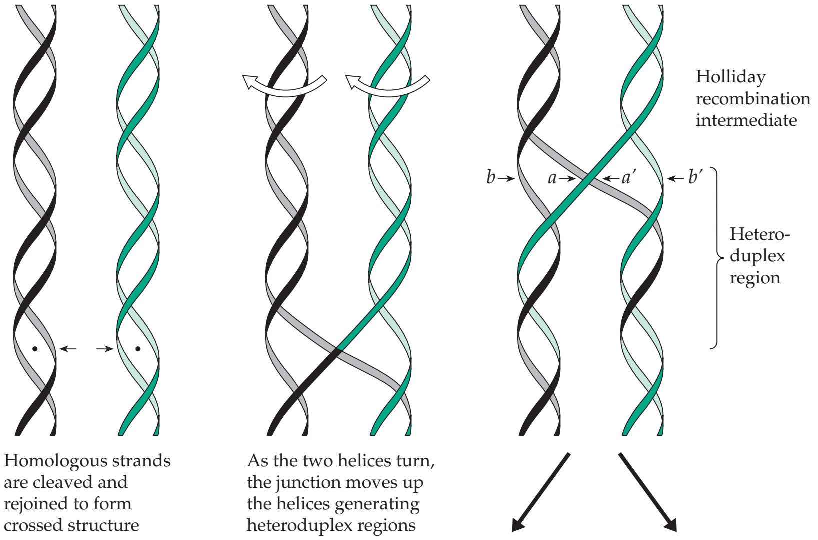

Holliday recombination intermediate
b
a
a′
b′
Hetero-duplex region
Homologous strands are cleaved and rejoined to form crossed structure
As the two helices turn, the junction moves up the helices generating heteroduplex regions
Cleavage at a, a′ and rejoining forms two parallel helices, each containing a heteroduplex region
Cleavage and rejoining at b, b′ produces two crossed duplexes

Figure 27-22 A recombination mechanism involving single-stranded exchanges. After Holliday.[524]

The cross-stranded structure shown in Fig. 27-22 can be formed with all base pairs in both duplexes intact.[526,527] All that is required is formation of a nick in each of the two polynucleotide chains and a rejoining of the backbones across the close gap between the duplexes. This model also accounts for the cutting of the two crossed strands at exactly equivalent points to terminate the process. Various mechanisms of recombination exist, and most make use of the key Holliday recombination intermediate (Fig. 27-22, Fig. 5-28, Eq. 27-11).[521,528–530a] Such four-way junctions can arise in several ways.[530a–d] For example, a 3' or 5' single-stranded tail in a piece of dsDNA can “invade”

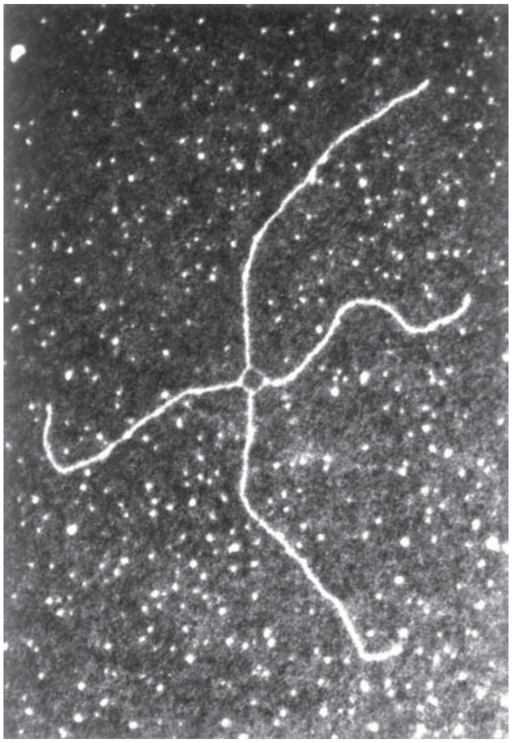
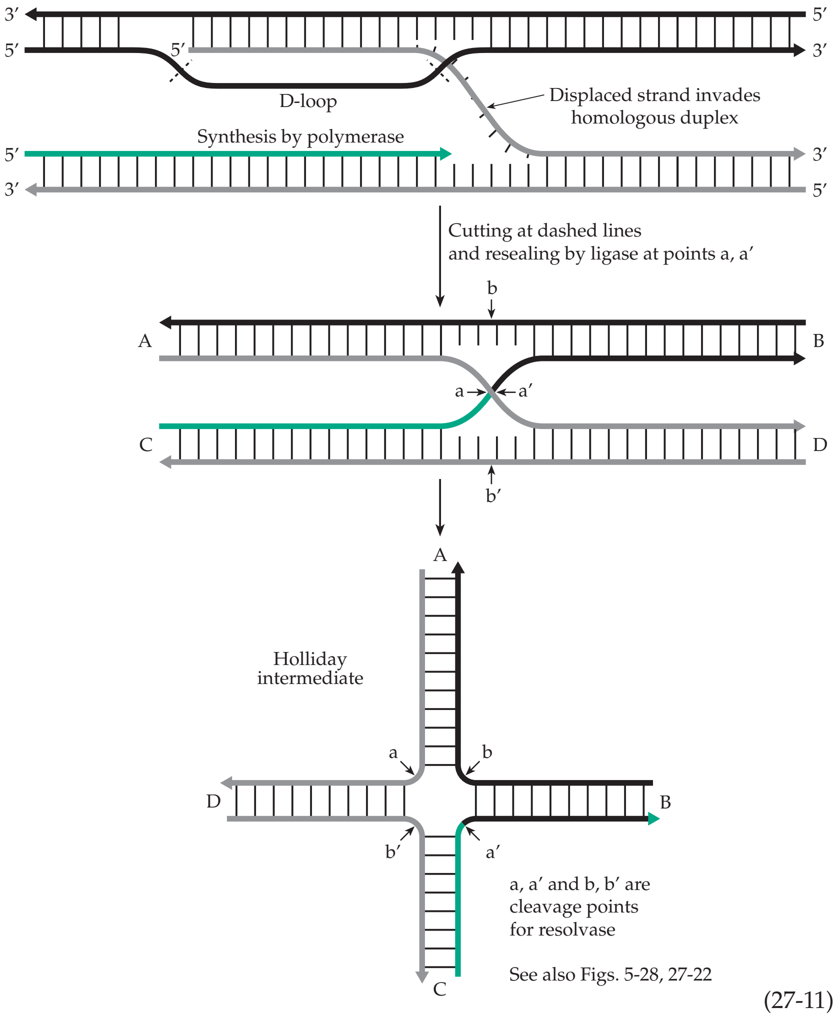


Figure 27-23 A chi form of DNA from the colicin E1 plasmid. These forms are thought to be derived from recombination intermediates of the Holliday type, which appear as “figure eight”-shaped molecules twice the length of the colicin genome. This figure eight form was cut at a specific site that occurs only once in the genome (twice in the figure eight) by restriction enzyme *Eco*R1 to give the chi form. The pairs of long and short arms are believed to represent homologous duplexes. The single strands in the crossover have pulled apart revealing the strand connections clearly. Such a structure would be expected from the Holliday intermediate (upper right corner of Fig. 27-22), e.g., if one of the two vertical duplexes were rotated end over end. From Potter and Dressler.[525]


3′
5′
5′
3′
5′
D-loop
Displaced strand invades homologous duplex
Synthesis by polymerase
5′
3′
3′
5′
Cutting at dashed lines and resealing by ligase at points a, a′
b
A
B
a
a′
C
D
b′
A
Holliday intermediate
a
b
D
B
b′
a′
a, a′ and b, b′ are cleavage points for resolvase
See also Figs. 5-28, 27-22
C

(27-11)

another dsDNA that has a homologous sequence as indicated in Eq. 27-11. In this drawing a 5' tail has been displaced during repair of a gap in one strand. The resulting D loop may be trimmed out and a new connection made to give the Holliday intermediate. The cleavage points a, a′, b, b′ marked in Eq. 27-11 correspond to those in Fig. 27-22. Holliday junctions may also be formed in stalled replication forks and must be removed to allow replication and transcription to continue.[530e] The existence of the Holliday intermediate has been supported not only by electron micrographs such as that of Fig. 27-23 but also by the identification of endonucleases that carry out the necessary cleavages of synthetic Holliday intermediates that have been made artificially (see Chapter 5.) Endonucleases with a high specificity for Holliday junctions have been found in bacteria, among proteins encoded by viruses, and in a wide variety of eukaryotic cells. Additional proteins including helicases, DNA-binding proteins, and specialized **strand exchange proteins** are also required to catalyze the individual steps in the recombination process.

The main **RecBCD pathway** of recombination in *E. coli* depends upon a dsDNA nuclease and an unwinding complex consisting of proteins RecB, C, and D.[531–533] The complex is a powerful *exonuclease*, which can digest the ends of a DNA duplex. It degrades the 3' ends most rapidly, leaving 5'-tails that can invade other homologous duplexes as in Eq. 27-11. The RecBCD complex is also an ATP-dependent helicase, which unwinds the DNA, preparing ssDNA for reaction with the strand-exchange protein **RecA**. The RecBCD complex also functions to completely degrade foreign dsDNA such as that from invading bacteriophages.[532] Why doesn't it also degrade the genomic DNA of the *E. coli* cell in which it functions? The answer lies in an eight-base DNA sequence, a recom-bination "hot spot" called **chi (χ)**: 5'–GCTGGTGG–3'. This χ sequence occurs 761 times in the leading strands for DNA replication in the *E. coli* genome.[533a–534a] When the RecBCD complex reaches a χ sequence, when approaching it from the 3' end, the enzyme stops its exonuclease action by inactivating the nucleolytic activity of the D subunit and promotes recombination about five- to ten-fold as fast as at other sites.[533,535]

RecA and other strand-exchange proteins. The 352-residue product of the *E. coli RecA* gene is a multifunctional **recombinase,** which is required both for recombination and also for DNA repair.[536–540a] In its repair function the RecA protein acts as a DNA-dependent protease that cleaves a number of repressors in response to damage to DNA. It has a quite different role in recombination where it (1) brings a piece of single-stranded DNA (an end or a gap) together with a duplex; (2) locates homologous sequences; and (3) forms a synaptic complex in which strand exchange can occur. Electron microscopic observations[536] show that the RecA protein binds to either single-stranded or duplex DNA in a cooperative manner to form long rodlike spiral filaments (Fig. 27-24). Measurement of the lengths of RecA protein-covered duplexes shows that the DNA is underwound and stretched by about 50%. It contains ~18 nucleotides per turn.[536,539,539a] Similar filaments are formed with single-stranded DNA, 3–4 nucleotides being bound per RecA protein monomer. Formation of this ssDNA complex, which may be regarded as an initiation

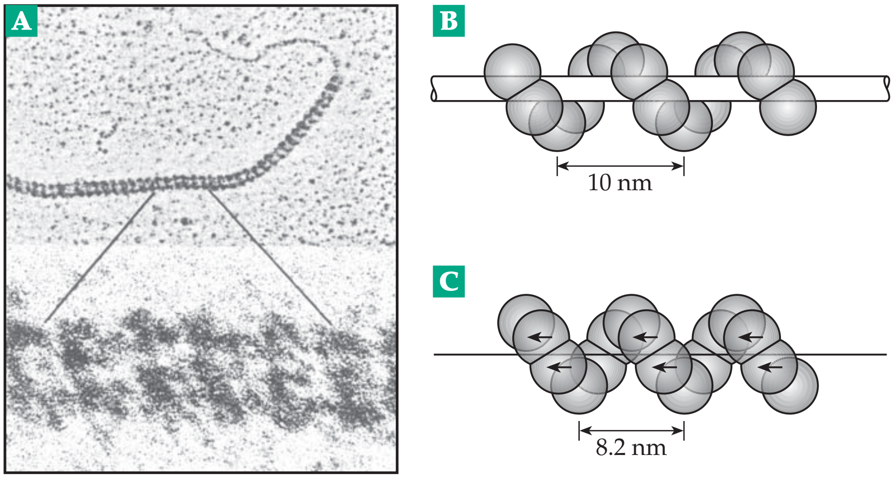
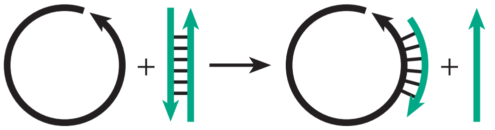


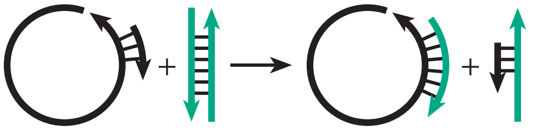
Figure 27-24 Structures of RecA protein spiral filaments. (A) RecA protein filament formed on circular duplex DNA in the presence of ATP(γ-S), shadowed with Pt and seen by electron microscopy. (B) Diagram of RecA bound to duplex DNA in the presence of ATP(γ-S), as determined by electron microscopy. RecA monomers are shown as spheres, but their exact shape is unknown. (C) Diagram of RecA spiral filament in crystals of RecA protein free of DNA, based on X-ray crystallographic data. Arrows indicated alignment of monomers. From Howard-Flanders, West, and Stasiak.[536]

complex for recombination, requires MgATP and is faciliated by prior coating of the DNA with SSB protein. The RecA protein subunits are added in the 5'- 3' direction of the DNA, and SSB is displaced in the process.[541]

The initiation complex binds to duplex DNA rapidly and more slowly promotes strand exchange. In related reactions a single-stranded SSB-coated circular DNA will bind to RecA protein, then exchange strands with a linear duplex (Eq. 27-12). The strand exchange requires ATP and advances in the 5' to 3' direction along the original single strand at the rate of a few bases / s. Strand exchange can also occur between two duplexes if there is a suitable gap in one strand, e.g., as is illustrated in Eq. 27-13.



(27-12)



(27-13)

A possible mechanism of strand exchange is illustrated in Figs. 27-25 and 27-26. The RecA protein has binding sites that can accommodate nucleotides from two DNA molecules, one single-stranded and the other a duplex. It may also accommodate two DNA duplexes.[542] As shown in Fig. 27-25, the RecA protein could test the hydrogen-bonding between many base pairs at once in a search of homologous regions. The two DNA chains would have to either slide past each other or repeatedly dissociate and reassociate[543] until a homologous region was found. Then strand exchange could occur. As is shown in Fig. 27-26, the single strand may be wound into the major groove of the duplex to form an interwound triplex. The matching of hydrogen-bonding atoms may be an attractive way of searching for homology, but the actual search seems to substitute speed for precision. Base substitutions are quite permissive. The need for precise hydrogen-bonding has not been demonstrated, and the exact recognition mechanisms in homologous recombination remain uncertain.[544–544b] Whole chromosomes must be aligned and checked rapidly in the homology search.[545]

Many proteins similar to the RecA protein and with similar functions have been found.[546,547] These include the products of gene *uvsX* of phage T4,[548,549] the β protein of phage lambda,[550] the yeast RAD51[551,551a] and human RAD51 proteins,[552,553] a meiosis-specific human homolog of the RecA protein,[554] and corresponding proteins from plastids of higher plants.[555] Both the UvsX protein of phage T4 and human RAD51 protein yield strands of coated DNA similar to those in Fig. 27-24.

Processing the Holliday junction. Completion of the recombination process requires “resolution” of the Holliday intermediate by endonuclease action followed by ligation and perhaps by gap repair. The major recombination pathway in *E. coli* employs a binding protein, a nuclease, and a helicase encoded by genes *RuvA*, *B,* and *C*.[528] **RuvA** is a DNA binding protein specific for symmetric Holliday junctions.[529] **RuvB** is a closely associated ATP-dependent helicase.[556–558] On the basis of genetic and X-ray crystallographic evidence it is now evident that some of the functions previously attributed to RecA are carried out by the RuvABC complex. As indicated in Fig. 27-26B, RuvA binds to the Holliday junctions, holding it in the symmetric square configuration in which branch

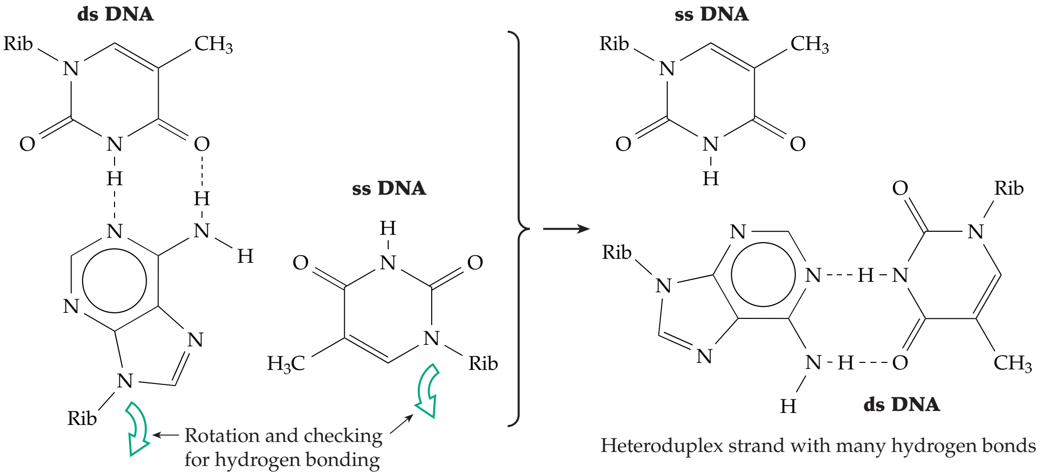


Figure 27-25 A possible mechanism for homologous pairing of an ssDNA with a duplex DNA and strand exchange. The ssDNA (right) binds together with a hydrogen-bonded duplex (left). The RecA protein rotates the bases into the heteroduplex configuration, where hydrogen bonds may be formed in many of the base pairs. After Howard-Flanders, West, and Stassiak.[536]

migration is possible. Two molecules of the oligomeric RuvB helicase apparently rotate the DNA, causing branch migration and movement of the DNA through the RuvAB complex.[559,559a,b] Under some circumstances a different helicase, encoded by *E. coli* gene **RecG,** moves Holliday junctions in the opposite direction.[560] In addition to RuvA a variety of other Holliday junction-binding proteins are known.[561] These include **p53** and the nuclear HMG proteins.[560,561]

RuvC is an endonuclease that is highly specific for Holliday junctions. It is a **resolvase** that cuts at either points a,a′ or b,b′ of Eq. 27-11 to form either "patched" or "spliced" recombinant DNA (Fig. 27-26C). Similar resolvases process bacteriophage DNA[562–564] and have also been found in yeasts and in mammals.[565,566] All are dimeric metal ion-dependent proteins.[567]

2. Nonreciprocal Recombination and Unequal Crossing-Over

The phenomenon of **gene conversion** or **nonreciprocal recombination**[568,569] was first recognized in genetic studies of fungi for which the four haploid miotic products can be examined individually (tetrad analysis; p. 20). Instead of the normal Mendelian ratio of 2:2 for the gene distribution in the progeny at a heterozygous locus a ratio of 3:1 is sometimes observed. One of the recombinant chromosomes appears to have been altered to a parental type. A reasonable mechanism by which this can occur arises from the fact that heteroduplex regions, which are present in recombination intermediates, contain defects in base pairing. One strand of the heteroduplex will have a base that does not properly pair with the base in the other strand, or will have an extra base that loops out from the heteroduplex. Since cells contain repair mechanisms that search for defects and carry out a repair process, there is a likelihood that one strand in the heteroduplex region will be altered to restore perfect base pairing, thus causing the observed gene conversion. "Flanking" genetic markers outside of the heteroduplex region are unaffected by gene conversion, and during meiosis crossing-over between these markers occurs in about 50% of gene conversion events as would be predicted from the model of Fig. 27-21. Data from yeast show that nonreciprocal recombination during meiosis may also result from double-strand breaks and gap formation followed by repair synthesis using both strands of the homologous chromosome as templates.[569]

Recombination is not limited to meiosis but can occur between homologous chromosomes during mitosis, during the G_1 period preceding mitosis, or even during the G_2 period.[570,571] Certain mutations in yeast abolish meiotic recombination but have much less effect on mitotic recombination.[572] Thus, the two processes are not identical. It has been suggested that mitotic recombination is utilized to maintain sequence homogeneity between repeated eukaryotic genes.[572,573]

Since DNA contains many repeated sequences, crossing-over sometimes occurs between locations that are not the same in the two duplexes. Such **unequal crossing-over** has the effect of lengthening one duplex and shortening the other. This may be very important in evolution. It may also, surprisingly, function to preserve homogeneity of chromosomes within a species.[574,575] For example, tandem arrays of ribosomal RNA genes (Section B,3) in yeast have 140 identical copies of their 9-kb repeat unit.[575] Unequal crossing-over between either sister chromatids or homologous chromosomes, when repeated often enough, can lead statistically to a highly homogeneous population.

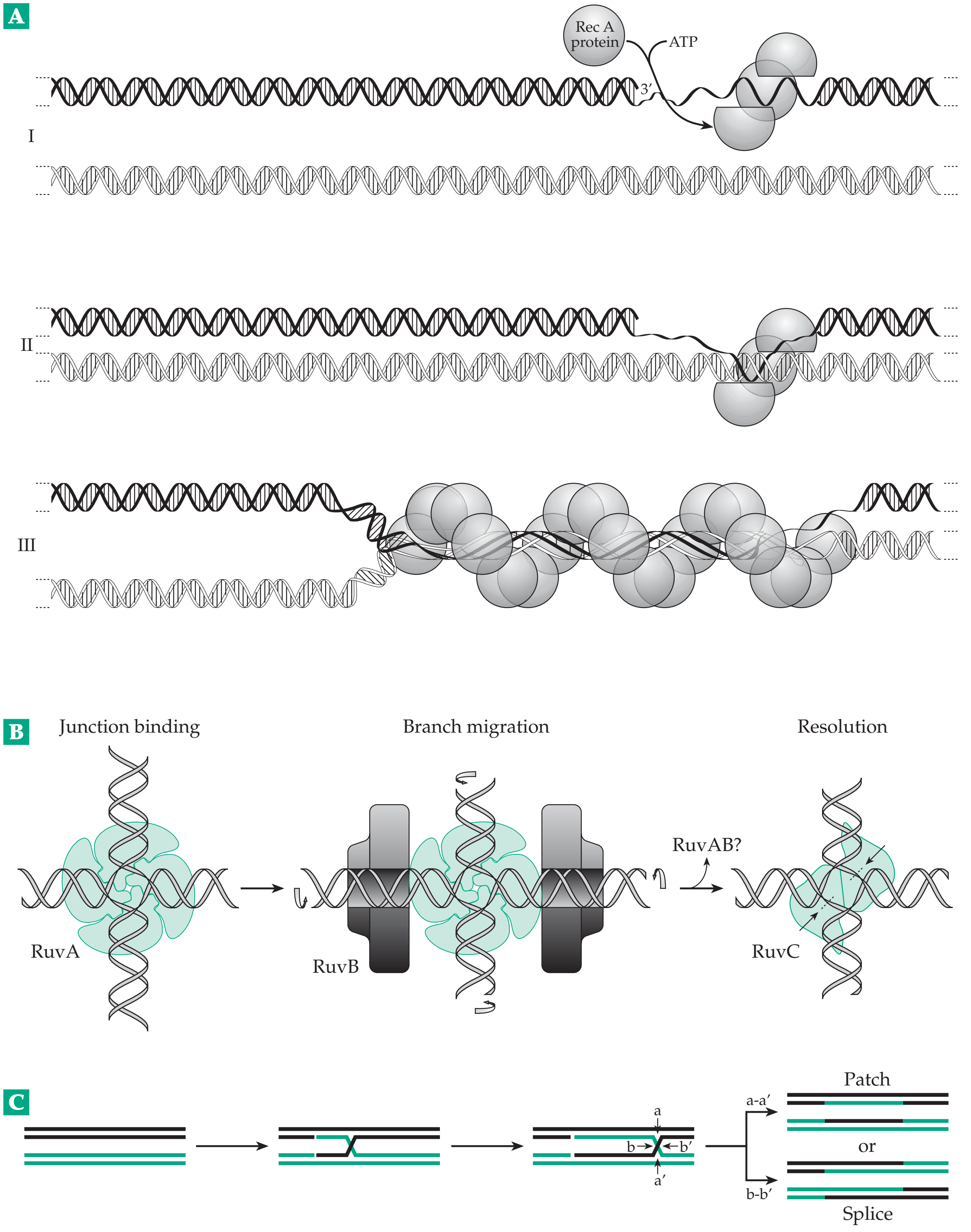


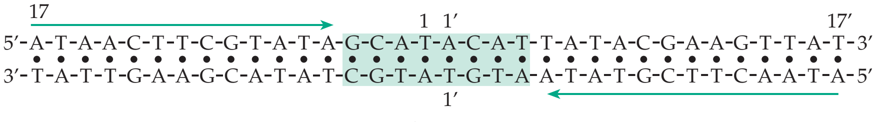
Figure 27-26 (A) Model for genetic recombination proposed by Howard-Flanders *et al.*[536] (I) RecA protein binds cooperatively to the single strand in a gapped duplex to form an initiation complex in preparation for pairing. (II) The initiation complex binds to the intact duplex, making transient contacts until a homologous site is reached. For clarity, the initiation complex is drawn with only a few protein monomers, but in reality it is likely to extend over hundreds or thousands of nucleotides. (III) When homologous contacts are made and the strands become paired locally, the initiation complex acts as nucleus for further cooperative binding, which extends the RecA spiral filament around all three or perhaps all four interacting strands. (B) Arrangement of the proteins and DNA during three stages of recombination catalyzed by the RuvABC system. The two RuvB hexameric rings are shown in cross section with the DNA passing through their centers. After Rafferty *et al.*[529] See also chapter banner, p. 1527. (C) Scheme of DNA rearrangement during homologous recombination in *E. coli*.

3. Site-Specific Recombination and the Integration and Excision of DNA

Recombination at specific sites in DNA is responsible for integration of DNA from viruses into the genome and for the cutting out of viral DNA and other pieces of DNA from the genome. The temperate bacteriophage λ and the F factors and R factors of bacteria can all be integrated into the genomic DNA of the host in this way. Genes encoded by the phage or plasmid are required. In the case of phage λ the viral genes *int* and *xis* are required for integration and excision, respectively.[576] These are not the same as the enzymes of the *rec* loci of the bacterium or the general recombination genes *exo* and *bet* of the phage. In addition, both integration and excision require an *E. coli* protein called integration host factor (**IHF**), a DNA bending protein resembling the DNA-binding HU.[18b,265,577,577a]

Integration of λ DNA (Fig. 27-27) occurs at the ~25-bp site *att B* in *E. coli* (Fig. 26-4) and the ~240-bp site *att P* in the λ chromosome (Fig. 28-11). These two sites contain identical 15-bp core sequences within which the recombination occurs. In a manner similar to that of the recA protein a homologous region is located by the complex of the Int (**integrase**) protein and IHF. Several molecules of Int protein bind and, together with the IHF protein, hold the phage DNA in a nucleosomelike structure (an **intasome**) in which the recombination occurs.[406,578–580b] Strand cleavage and rejoining occur within the short core sequence (Fig. 27-27). Parts of both *att P* and *att B* are recombined to give sites *att L* (left) and *att R* (right) in the DNA of the integrated prophage. In the integration complex the two core *att* sequences are aligned, and single-strand cuts are made at one of the points *a* or *b* that are indicated by the small arrows located on opposite strands and seven bp apart in the core sequence shown in Fig. 27-27. Rejoining of strands from the opposite duplex yields a Holliday intermediate. That this really occurs is shown by the fact that the Int protein cleaves synthetic Holliday intermediates derived from the *att* core and reseals the strands to give the expected products.[581] Cleavage of the Holliday intermediate at points a (Fig. 27-27) will lead to excision of the viral circle, but cutting at points b followed by resealing with opposite strands, as is observed, will yield integrated prophage. Although Int and IHF proteins are sufficient to promote integration, the **excisionase** encoded by phage λ gene *xis* is needed together with the Int protein for excision of the λ prophage.[576,582]

The integrase (tyrosine recombinase) family. The lambda integrase is the first recognized member of a family of a hundred or more closely related enzymes that are involved not only in integration and excision of phage DNA but also in converting multimeric forms of bacterial and plasmid chromosomes into monomers. One of the best known integrases is the 38.5-kDa **Cre recombinase**, which functions to keep the lysogenic phage P1 in a monomeric form by recombination between pairs of 34-bp core sequences designated *loxP*.


17
1 1′
17′
5′-A-T-A-A-C-T-T-C-G-T-A-T-A-G-C-A-T-A-C-A-T-T-A-T-A-C-G-A-A-G-T-T-A-T-3′
3′-T-A-T-T-G-A-A-G-C-A-T-A-T-C-G-T-A-T-G-T-A-A-T-A-T-G-C-T-T-C-A-A-T-A-5′
1′

A *loxP* site

Since the reaction doesn't require accessory protein factors and can be performed *in vitro* with a variety of DNA substrates, the *Cre–loxP* system is much used in genetic engineering.[583,584] A pair of related integrase subunits known as **XerC** and **XerD** perform a similar function for the *E. coli* chromosome as well as for multicopy plasmids.[585,586] The XerC / XerD system is

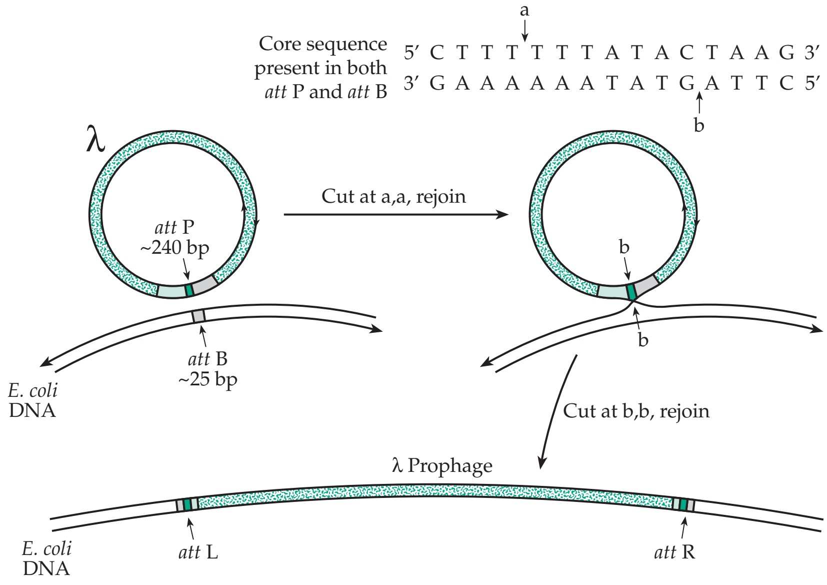

a
Core sequence present in both att P and att B
5′ C T T T T T T A T A C T A A G 3′
3′ G A A A A A A T A T G A T T C 5′
b
λ
att P
~240 bp
att B
~25 bp
E. coli DNA
Cut at a,a, rejoin
b
b
Cut at b,b, rejoin
λ Prophage
E. coli DNA
att L
att R

Figure 27-27 Integration of the phage λ genome into the *E. coli* chromosome at site *att*B. The same recognition sequences are present at *attB* and *attP*. These are cut at points *a* with rejoining to give the structure at center right. This is cleaved at points *b* with rejoining to give the integrated prophage.

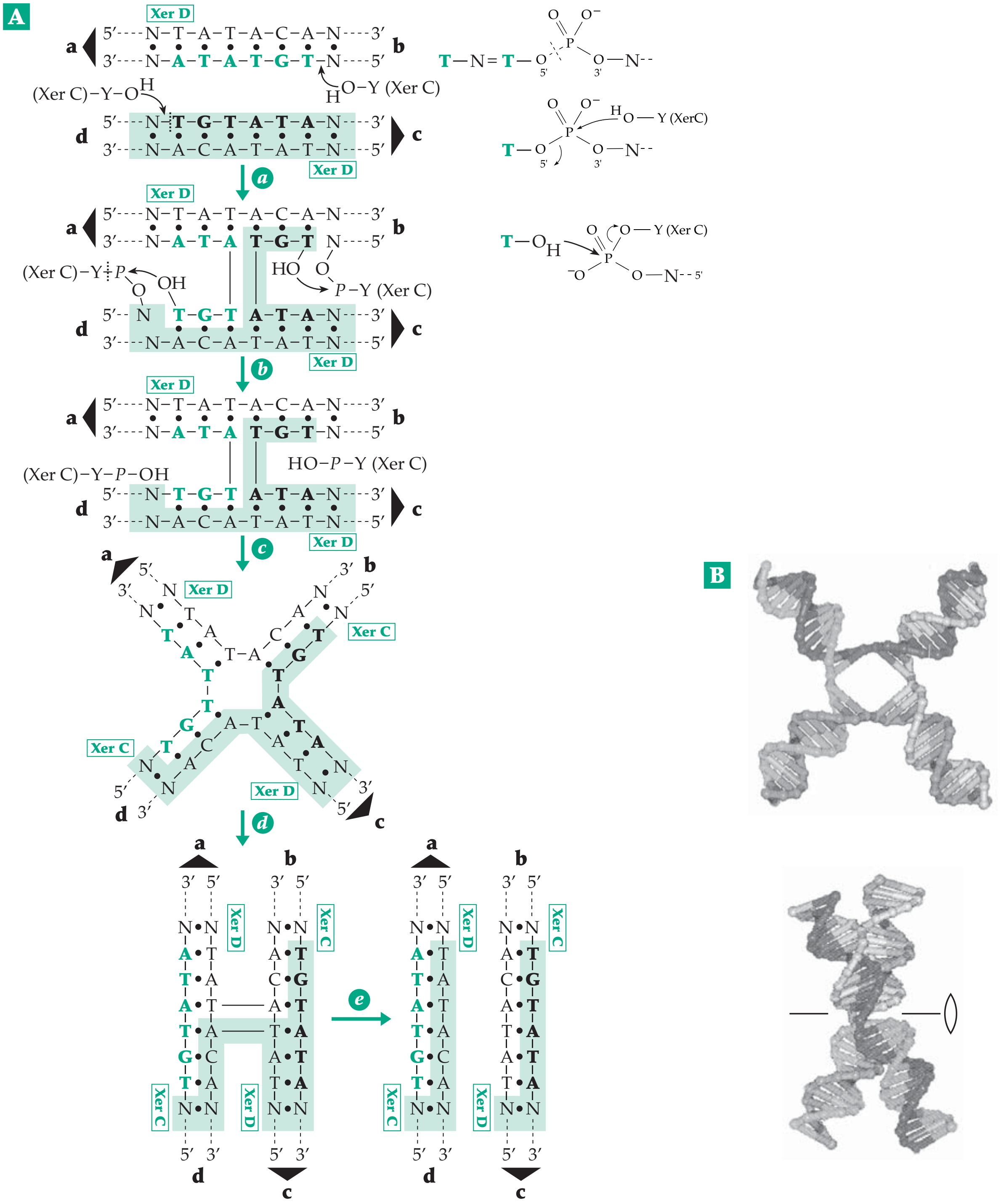
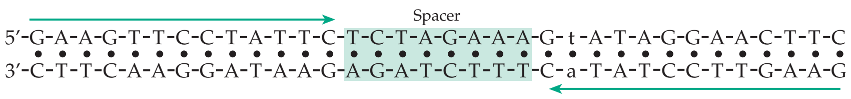

A
Xer D
5′----N-T-A-T-A-C-A-N----3′
3′----N-A-T-A-T-G-T-N----5′
a
b
(Xer C)-Y-OH
HO-Y (Xer C)
5′----N-T-G-T-A-T-A-N----3′
3′----N-A-C-A-T-A-T-N----5′
d
c
Xer D
a
(Xer C)-Y-P
OH
HO O
P-Y (Xer C)
b
(Xer C)-Y-P-OH
HO-P-Y (Xer C)
c
Xer C
Xer D
d
e
T—N=T—O
P
O—N--
HO—Y (XerC)
T—OH
O—Y (Xer C)
5′
3′
B

Figure 27-28 (A) Action of the integrase XerC / XerD on a pair of *E. coli dif* sites, containing the central six bp sequences TATACA / ATATGT, which are shown in an antiparallel orientation. In step *a* the active site tyrosine hydroxyl groups (Y-OH) of a pair of XerC subunits carry out transesterification reactions on the 5'-terminal thymidylate residues of the central hexanucleotide sequences to yield 3'-phosphodiester linkages to the XerC tyrosines. In step *b* the cut 5'-ends, containing free thymidylate 5'–OH groups, fold back to form new base pairs, and the strands are resealed in a second transesterification. This generates a Holliday junction, which *isomerizes* (steps *c* + *d*) to an isomeric species that is acted on by a second pair of transesterification steps (*e*) that are catalyzed by protein XerD, again with folding back of the central trinucleotides. After Arciszewska *et al.*[586] (B) The two isoforms of an antiparallel stacked X Holliday junction are shown. These can be reached from the symmetric square form shown in Fig. 5-28 and schematically in (A) by folding into X-conformations in which all base pairs are stacked either in pairs a,b and c,d or a,c and b,d. From Eichman *et al.*[525a]

atypical because it utilizes a pair of integrase subunits rather than just one. However, the basic chemistry (Fig. 27-28) is the same for the entire family.[587,587a] As with the λ integrase (Fig. 27-27) the XerC / XerD complex acts on a pair of identical core sequences that are aligned in an antiparallel fashion. Active sites in all of the integrases contain the conserved amino acid sequence Arg-His-Arg-Tyr. All of these residues are essential for catalysis.[585] Staggered cuts are made sequentially in the core sequences, e.g., at points a and b in Fig. 27-27 and adjacent to the 5'-terminal thymidylate residues in Fig. 27-28A. A transesterification reaction forms 3'-phosphotyrosyl linkages from the cut DNA to the integrase protein. The freed 5'-OH groups on the other cut end fold back and recombine with the bound 3' ends of the second duplex to generate a Holliday junction. In the mechanism proposed in this figure, three nucleotide units are involved in the folding back. The hydrogen bonds of their initial base pairs are broken, and new bonds are formed. This "base swapping" process is the equivalent of a short branch migration and verifies homology of the two recombination sites.

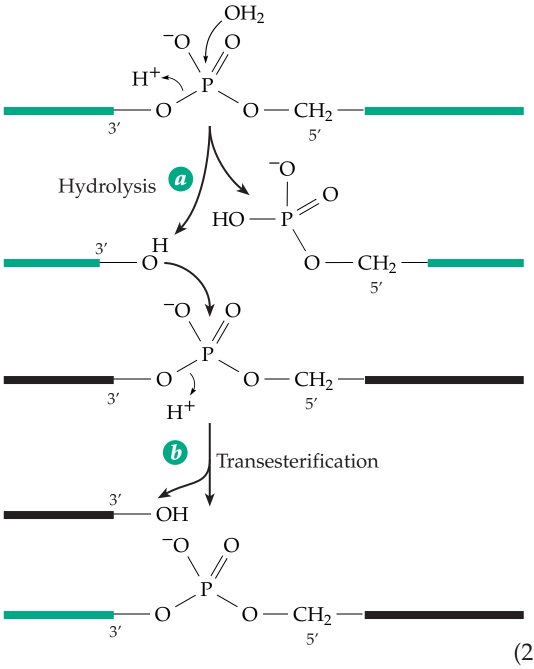
The Holliday junction can isomerize readily between two forms. In one a base-stacked double helix runs between ends a and b in Fig. 27-28B and another runs between ends c and d. The core hexanucleotide sequences lie between the marked XerC and XerD cleavage sites. In the other isomer (lower drawing) one helix has ends a and d and the other b and c. Following the isomerization (steps *c* and *d* in Fig. 27-28A) the XerD active sites act in two transesterification reactions (not shown but analogous to those in steps *a* and *b*) with base swapping of the trinucleotides at the cut ends. This generates the two separate recombinant duplexes. These might be two circular chromosomes or plasmids formed by recombination from a double length chromosome or plasmid. The previously mentioned λ amd Cre recombinases appear to act by closely similar mechanisms.

Tyrosine recombinases of the lambda family also function in eukaryotes. Best known is the **FLP (Flip) recombinase,** which is encoded by the 2-µm plasmid of *Saccharomyces cerevisiae* and is thought to function in amplifying the number of plasmid copies.[265] The 6.3-kbp plasmid contains a unique DNA sequence that lies between two 599-bp repeats in inverted orientation. Embedded in each repeat is an FLP recombination target (**FRT**) sequence, which is recognized by the plasmid recombinase. Each FRT segment includes inverted repeats 13 bp in length with an 8-bp spacer between them. As with other integrase systems the 8-bp spacer or **strand exchange region** is **asymmetric** and establishes the orientation of the recombination sites.[587–590] The role of the recombinase is to invert one of the 599-bp repeats with respect to the other (see Eq. 27-15). This switches replication of the plasmid to a rolling circle pattern.[265]


Spacer
5′-G-A-A-G-T-T-C-C-T-A-T-T-C-T-C-T-A-G-A-A-A-G-t-A-T-A-G-G-A-A-C-T-T-C
3′-C-T-T-C-A-A-G-G-A-T-A-A-G-A-G-A-T-C-T-T-T-C-a-T-A-T-C-C-T-T-G-A-A-G

FLP recombination target (FRT) sequence

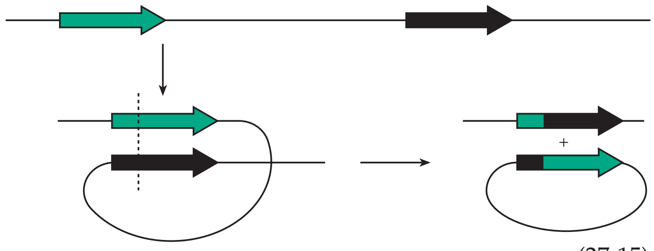
The resolvase / invertase family and invertible DNA sequences. A second large family of recombinases act by cleaving a target DNA sequence hydrolytically leaving a free 3'-OH end (Eq. 27-14, step *a*). This free end then attacks a phosphodiester linkage in a second strand of DNA, cleaving that strand with an in-line nucleophilic displacement (step *b*). Active sites usually contain a characteristic cluster of aspartate and glutamate (DDE) side chains, which probably act together with a metal ion, perhaps as in Fig. 27-13.[591] Enzymes in this resolvase / invertase family act either to resolve cointegrates in transposon action (next section) or to invert DNA sequences.



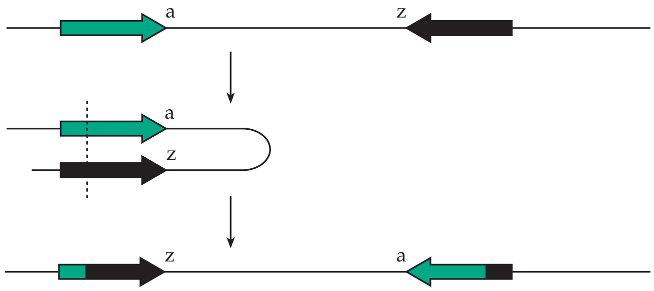
(27-14)

If recombination occurs within a piece of DNA at two homologous sites such as the *attL* and *attR* sites at the boundaries of the λ prophage, the intervening DNA will be excised as a circular particle (Eq. 27-15). In this instance the two homologous regions must be repeated in the same direction, as is indicated by the arrow structures in Eq. 27-15. If the homologous sequences are oriented in opposite directions, i.e., they are inverted repeats, excision will not occur but the piece of DNA between the repeats will be inverted (Eq. 27-16).

A number of such invertible



(27-15)



(27-16)

DNA recombination systems are known and are sometimes used to control specific genes.[592] For example, many strains of *Salmonella* have two types of flagella, which are composed of flagellin subunits encoded by genes H1 and H2 respectively.[592–594] On rare occasions an individual bacterium switches from one flagellar "phase" to the other. This occurs by recombination, as in Eq. 27-16, between two 26-recombination sites *hixL* and *hixR*, each of which contains a 14-bp inverted repeat. The 993-bp invertible segment encodes a recombinase gene called *hin* and a promoter, i.e., an mRNA initiation site, which has a specific orientation. In one orientation mRNA is transcribed from a short operon that includes the right inverted repeat IRR, the *H2* flagellin gene and gene *rh*1, which encodes a repressor for flagellin gene *H1*. Consequently, only gene *H2* is expressed. In the other orientation the RNA transcription is in the opposite direction so that neither *H2* nor *rh*1 is expressed. However, *H1*, which is located elsewhere, is expressed freely.

Two other proteins are required for efficient inversion by the Hin recombinase. A dimer of a 98-residue helix–turn–helix DNA binding protein called **Fis** (factor for inversion stimulation), a relative of protein HU,[18b] must bind to an enhancer, a 65-bp DNA segment. Binding of Fis to the enhancer helps to hold the supercoiled DNA and the recombinase in a correct orientation for reaction.[576,595,596] Protein HU is also needed. The same Fis protein binds to an enhancer for a 3-kbp invertible DNA sequence, which controls alternative host preferences for bacteriophage Mu (Fig. 27-29).[576,596] The chemistry of the inversion reaction is related to that of the replicative transposons discussed in Section 4.

Microorganisms sometimes control the synthesis of surface proteins using segments of invertible DNA. The pathogenic bacterium *Campylobacter fetus* utilizes DNA rearrangements to allow one of a large family of surface layer (S-layer) proteins to be formed.[597] The yeast FLP recombinase, mentioned in the preceding section, also inverts the sequence flanked by the 599-bp repeats.[589]

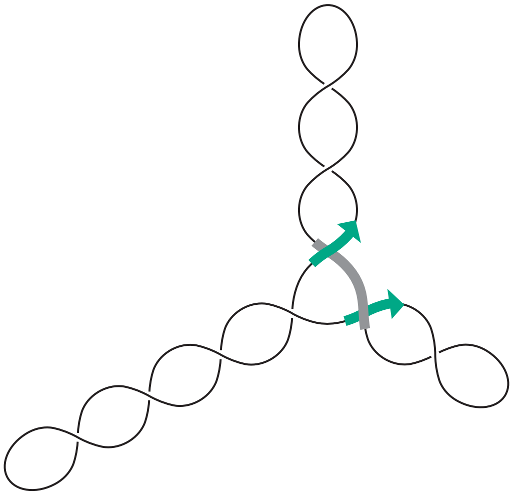
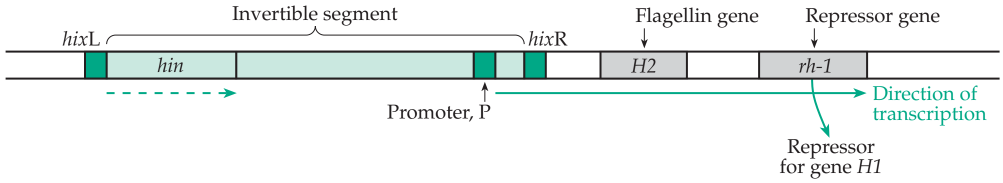


Figure 27-29 Formation of a synaptic complex of a supercoiled circular DNA containing the sites *gix* (green), which pass over and under the enhancer (gray). The recombinase Gin and the enhancer-binding Fis form a synaptic complex with DNA in this form as seen directly by electron microscopy. From Sadowski.[576]

4. Transposons and Insertion Sequences

The first evidence that some genes can move from one location to another within the genome came from studies of *Zea mays* by Barbara McClintock in the late 1940s.[598–602] She concluded that the variegated kernels found in some colored maize were a result of **controlling elements**, which could move from place to place turning on or inhibiting expression of various genes including some of those determining anthocyanin pigment formation. Two of these systems have been



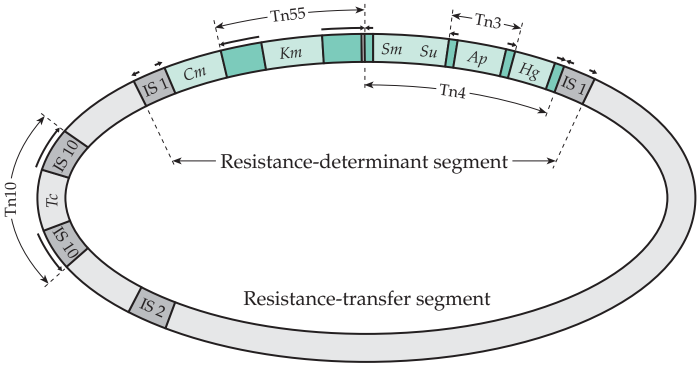


Figure 27-30 Transposons in an antibiotic-resistance plasmid. The plasmid appears to have been formed by the joining of a resistance-determinant segment and a resistance-transfer segment; there are insertion elements (*IS1*) at the junctions, where the two segments sometimes dissociate reversibly. Genes encoding resistance to the antibiotics chloramphenicol (*Cm*), kanamycin (*Km*), streptomycin (*Sm*), sulfonamide (*Su*), and ampicillin (*Ap*) and to mercury (*Hg*) are clustered on the resistance-determinant segment, which consists of multiple transposable elements; inverted-repeat termini are designated by arrows pointing outward from the element. A transposon encoding resistance to tetracycline (*Tc*) is on the resistance-transfer segment. Transposon Tn3 lies within Tn4. Each transposon can be transferred independently. From Cohen and Shapiro.[602]

studied especially intensively: the "**activator (Ac)–dissociation (Ds)**" system, discovered by McClintock and the "enhancer (En)–inhibitor (I)" system, discovered by Peterson[599,600] and independently by McClintock.[603] Each contains two genetic elements (segments of DNA) of which Ac and En are autonomous, i.e., they can move by themselves. Ds and I, however, cannot move unless the other element of the pair is also present. Both Ds and En have now been cloned and sequenced.

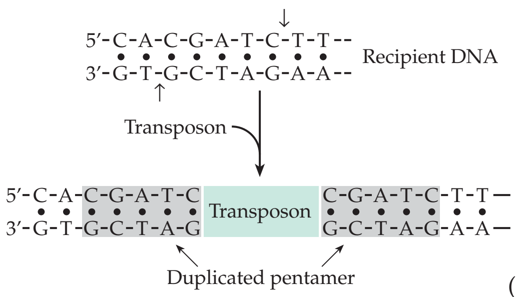
The general importance of transposable genetic elements was not appreciated by most molecular biologists until about 20 years after McClintock's discoveries. Then several moveable **insertion sequences** (Is elements) were found in enteric bacteria. Like the controlling elements of maize these small (0.8–1.4 kb) sequences can move and insert themselves at many points in the genome, often inactivating genes which they enter.[602–608] The *E. coli* genome contains eight copies of IS1, and five of IS2, as well as several others. Most species of *Shigella* contain more than 40 copies of IS1.[604] Mobile elements similar to those present in maize also exist in archaea.[604a]

Following the discovery of the IS elements it was found that transposable elements named **transposons** could transfer resistance to antibiotics between bacteria. All of these transposable elements have inverted repeat sequences at the ends. For example, IS1 contains the following sequence at both ends but with opposite orientation as if in a

$$\text{CAACTTA}\genfrac{}{}{0pt}{}{\text{C}}{\text{T}}\text{TGAT}$$

palindrome.[605] Some complex transposons have an IS sequence at each end. For example, Tn10 contains IS10, a relative of IS1, at both ends.[608] These provide the characteristic inverted repeat termini. Figure 27-30 shows a schematic drawing of a bacterial drug resistance plasmid containing IS1, IS2, and IS10 as well as transposons Tn3, Tn10, and Tn55. The two IS1 elements surround the large resistance determinant, which can be transferred as a block.[602] To be autonomously mobile a transposon must contain a **transposase** that enables it to be transferred. It usually carries other genes as well and may also contain one or more signals for transcriptional regulation such as promoters.

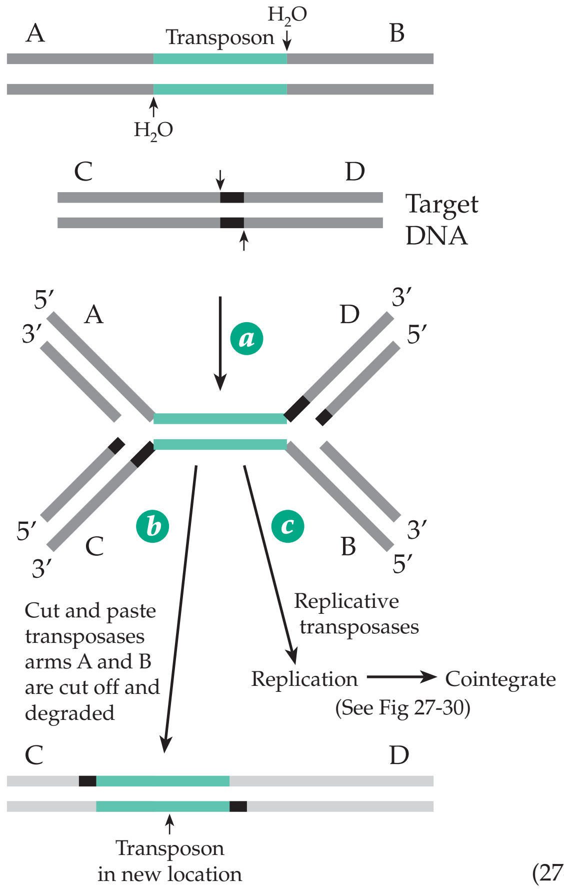
The chemistry of transposition is more complex than that of simple site-specific recombination. Transposition can occur at many sites in a genome, and no homology with the transposon termini is required. Transposition is accompanied by *duplication of a short sequence of the recipient DNA* exactly at the ends of the transposon. Usually 5, 9, or 11 base pairs are duplicated as in the hypothetical example of Eq. 27-17. This happens because staggered cuts 5, 9, or 11 bp apart in the recipient DNA are made during recombination. These are indicated by the small arrows in Eq. 27-17. Transposons causing 5-bp duplications are Tn3, Tn7,[608a] γδ, phage Mu of *E. coli*, and Ty1 of yeast[609]; IS1, Tn10,[609a] and Tn5 of *E. coli* cause 9-bp duplications.[610–611b] When a transposon of one major group moves to a new location, the original copy remains. In this case transposition involves a combination of replication and site-specific (for the transposon) recombination. As a consequence, a circular DNA molecule containing the transposon will often react with a second circular DNA to form a large circle, a **cointegrate**, which contains two copies of the transposon. However, another group of transposons utilize a "cut-and-paste" mechanism that doesn't require extensive DNA duplication.



(27-17)

Cut-and-paste (nonreplicative) transposons. The transposases of *E. coli* Tn5, Tn7, and Tn10 act by hydrolytically cutting both strands of duplex DNA at the transposon ends leaving the phospho groups attached to the 5' cut ends, as is depicted in detail in Eq. 27-14, step *a*. The two 3' ends then carry out transesterification reactions, as in Eq. 27-14, step *b*. These two steps are used to nick both strands of the DNA carrying the transposon and to join them to a target DNA sequence to give a branched intermediate (Eq. 27-18, step *a*). Nonreplicative transposons apparently cut off two arms, e.g., A and B, and heal up the small gaps by repair synthesis, leaving the transposon in a new location between C and D. The gap repair accounts for the duplication of the end sequences of the cut target DNA.[610–612a]

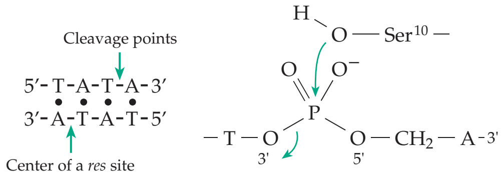
Replicative transposons. In 1979 Shapiro proposed the mechanism illustrated in Fig. 27-31 for replicative transposons. The two inversely repeated segments (green) at the ends of the transposon are aligned with the recipient DNA whose ends are labeled C and D. In fact, the recombining DNA molecules must be supercoiled.[1,613] Staggered cuts are made in the recipient DNA at points *a* and *b*, which are 5, 9, or 11 bp apart, depending upon the specific recombinase. Nicks are also made in the transposon ends. The 3' ends from the transposon are resealed with the 5' ends from the recipient DNA (step *a*) to give a structure that in effect has two replication forks. Replication (step *b*) yields the cointegrate, which contains two copies of the transposon as indicated. In a third step (step *c*) recombination between the two integrated transposons yields a copy of the original transposon-containing donor and the recipient DNA, which now also contains a copy of the transposon.



(27-18)

When a transposon reacts with another part of the same DNA circle there are two possibilities. The piece of DNA lying between the transposon and the recipient site may be excised as a circle containing a copy of the transposon. Alternatively, there will be inversion of that sequence as well as replication of the transposon (compare with Eqs. 27-15, 27-16).

The closely related transposons Tn3 and γδ are understood best. They contain not only a transposase gene but also a **resolvase** gene.[613–616] The transposase carries out the recombination reactions that yield the cointegrates, while the resolvase catalyzes site-specific recombination between the two transposons in the cointegrate to complete the transposition. Several subunits of the resolvases bind the two **resolution (res) sites** of the supercoiled DNA in a parallel orientation with the DNA supercoiled as in Fig. 27-29. However, no enhancer-binding protein is needed, and the two *res* sites must be supercoiled. Purified γδ resolvase uses hydroxyl groups of the serine-10 side chains as the nucleophiles to cleave the DNA by displacement at specific *res* sites to give transient enzyme-bound phosphodiester linkages.[615]



The synaptic complex contains 240 bp of DNA and at least two resolvase dimers. All four DNA chains are cut to give eight ends. Four of these are bound to the serine side chains in phosphodiester linkage. In the second step the freed 3'-OH groups react with the bound ends of the other duplex via a transesterification reaction to form the recombinant chains.

The resolvases act on supercoiled cointegrated DNA molecules that contain two directly repeated res sites to produce two singly linked circles (which are still supercoiled) each containing one res site as shown in Fig. 27-32. The two *res* (resolution) sites within the transposons are aligned, the open circle of DNA shown at the upper left being folded as shown in the lower part of the drawing. The DNA substrate is not knotted. However, after recombination it is catentated and will require action of a topoisomerase to separate

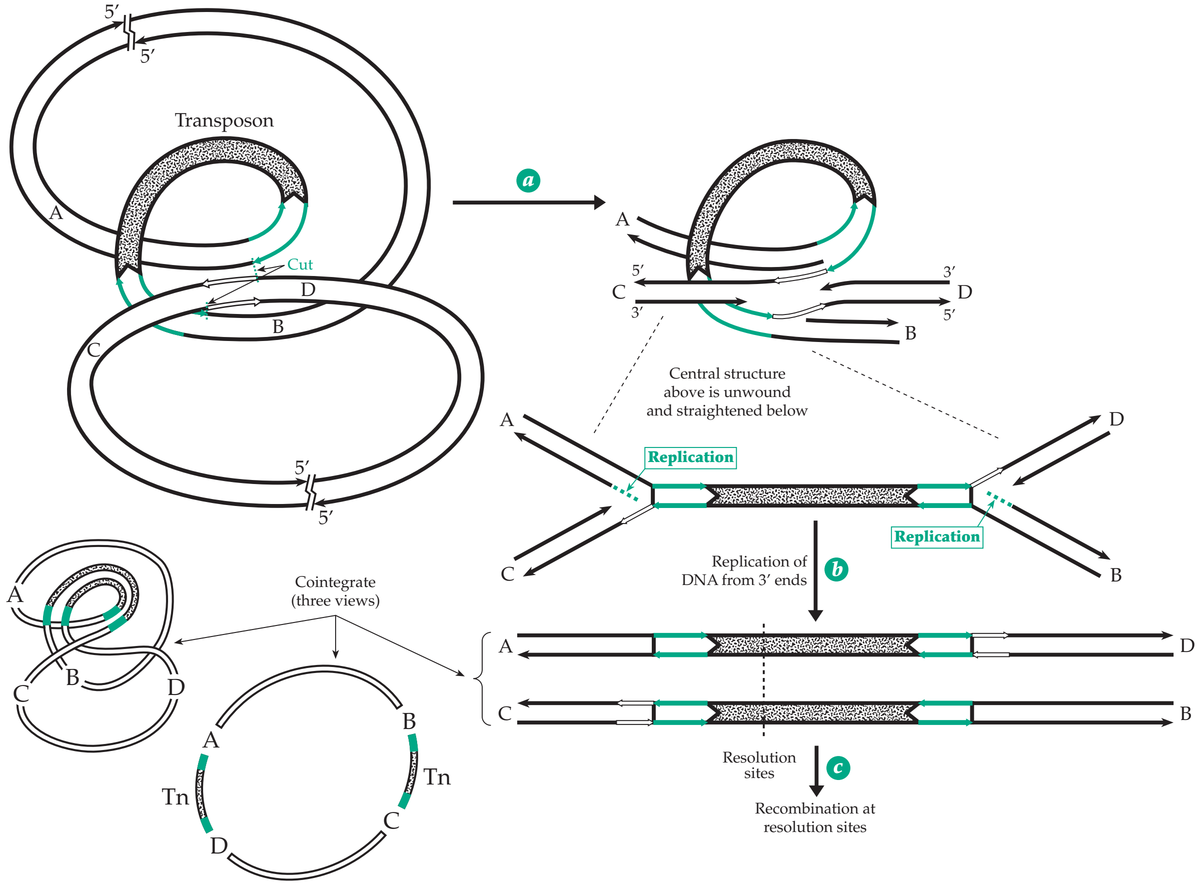

5′
5′
Transposon
A
Cut
D
B
C
5′
5′
a
A
5′
3′
C
3′
D
5′
B
Central structure above is unwound and straightened below
A
D
Replication
Replication
C
B
Replication of DNA from 3′ ends
b
A
D
C
B
Resolution sites
c
Recombination at resolution sites
Cointegrate (three views)
A
B
C
D
A
B
Tn
Tn
C
D

Figure 27-31 Scheme for integration of a transposon (stippled duplex) present in a piece of DNA with ends A and B into another piece of DNA with ends C and D and containing a suitable recognition sequence (open bars). Inversely repeated sequences in the transposon are shown as solid bars with a direction arrowhead. Arrowheads point toward 5′ strand ends. Cleavage and rejoining at points a,a and b,b yield an intermediate with two replication forks. Replication through the transposon yields one unchanged DNA segment with ends A and B and a transposon inserted into the other DNA segment. If A is continuous with B, a cointegrate structure is formed. See Cohen and Shapiro.[602]

the two products. Occasionally additional recombination events occur, perhaps processively along interwound double helices. This produces various knotted products.[617] An electron micrograph of one of these is shown in Fig. 5-17.

The temperate bacteriophage Mu. The most efficient transposon known is the 37-kb genome of the **mutator bacteriophage Mu**.[265,618–621] Once it becomes integrated into a bacterial chromosome, it replicates by repeated rounds of transposition within the host bacterium. During the lytic cycle some of the replicating DNA is excised as extrachromosomal circles of various sizes, which are packaged into virus particles by a headfull mechanism. The circles contain copies of some host DNA, but this is left behind when a virus particle infects a new cell.[265] The phage Mu DNA is integrated into host DNA by a cut-and-paste mechanism in a transpososome that contains a tetramer of the virally encoded transposase (MuA protein) bound to a supercoiled DNA. The transpososome resembles that in Fig. 27-29 and contains cleavage sites at the ends of the transposon and also an enhancer sequence.[620,621] Several steps involving conformational alterations occur in the transpososome. One is an ATP-dependent action of a second Mu-encoded protein MuB.[622] During the lytic cycle replicative transposition predominates.

Some other transposons. Transposons have a variety of biological functions. For example, haploid cells of the yeast *S. cerevisiae* exist as one of two mating types *a* or α. The mating type is established by transposition of one of two "cassettes" of genes from two different "silent" locations to a location from which they can be expressed.[623,624] See Chapter 28.

One of the best known eukaryotic transposons is the P element of *Drosophila*, which transposes only within the germ line cells of developing embryos, somatic cells being unaffected.[265,625,626] It belongs to

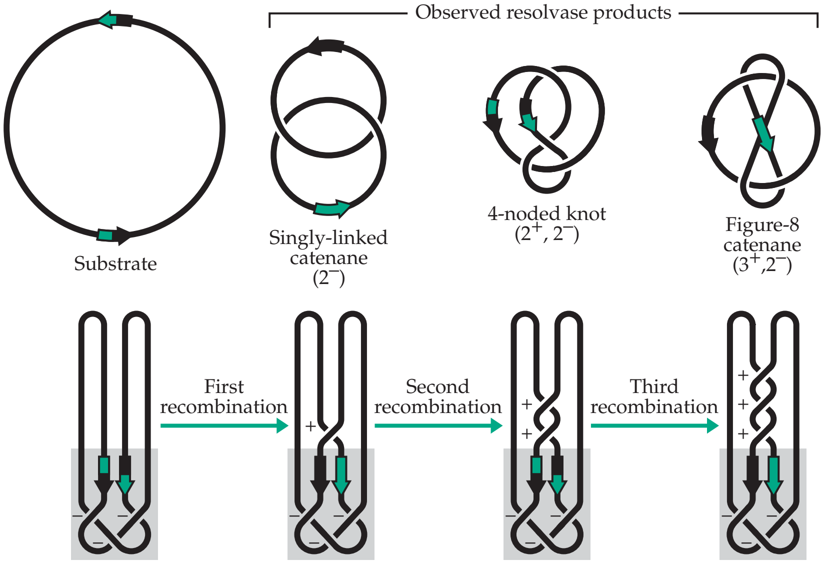

Observed resolvase products
Substrate
Singly-linked catenane (2^-)
4-noded knot $(2^+, 2^-)$
Figure-8 catenane $(3^+,2^-)$
First recombination
Second recombination
Third recombination
+
−

Figure 27-32 Scheme for resolution of an unknotted cointegrate molecule by a resolvase that cuts the transposons at resolution (*res*) sites and recombines them. The resolvase may act only once or repeatedly as shown. In the upper row, the duplex DNA substrate and products are represented in standard topological form as they might appear after nicking. In the lower row the DNAs are depicted as folded forms bound to the resolvase with the two directly repeated *res* sites (thick arrows) dividing the substrate into two domains (thick and thin regions). The substrate at synapsis has three (–) supercoils that entail crossing of the two domains. Successive rounds of recombination, each introducing a single (+) interdomainal node (see Fig. 5-17), are drawn in the lower row. Bound resolvase maintains the three synaptic supercoils. After dissociation from the resolvase at any stage, the product supercoil nodes either cancel with ones of opposite sign or are removed by subsequent nicking. The node composition is indicated in parentheses. From Wasserman *et al*.[617]

the same family as Barbara McClintock's Ac element of maize and the *Tc*1 family of nematodes.[627] P elements use a nonreplicative cut-and-paste method of transposition. The 87-kDa transposase protein requires GTP and Mg^{2+} for activity.[628] It was only in the past few decades that P elements have been found in *D. melanogaster*. They may have entered this fruit fly from another species, possibly transferred by a mite.[627,629]

A second *Drosophila* transposon called **mariner**[630] typifies the *mariner* / *Tc*1 transposon superfamily, which also contains members from nematodes,[631] other invertebrates, fishes,[632] amphibia,[633] and possibly human beings.[634] These transposons encode a transposase containing a D, D, D or D, D, E motif[630] but no other proteins. They contain short ~30-bp terminal inverted repeats and become inserted into host TA sequences.[631] Movement of some repetitive sequences of the LINE[635] and SINE[636] families within the human genome may be assisted by *mariner* transposons.[637]

The maize transposon *ac* is widely used as a means of inactivating genes and placing a "tag" that can be used to map the gene and to permit it to be cloned and sequenced.[598,638] Although initially of use only in maize the method has been extended to other plants,[639–641] and genetically engineered transposons have allowed it to be utilized in animals.[642]

A different kind of transposition controls self-sterility in maize. The cause of self-sterility in one strain has been traced to the presence of two linear episomes called S-1 (614 kb) and S-2 (514 kb) within mitochondria.[643] These have inverted terminal 208 bp repeats. On rare occasions they recombine with the circular mtDNA converting it to a linear form with the episomes covalently linked to one end. The change is accompanied by reversion to fertility.

As described in Section B,1, mammalian DNA contains many **retrotransposons** (retroposons) that lie within short direct repeats characteristic of transposons. However, they contain a poly(A) tail at the 3' end, an indication of their relationship to RNA transcripts, and are discussed in Chapter 28.

5. Other Causes of Genetic Recombination

Because their integration sites in host DNA do not depend upon homology with the transposon ends, transposition is sometimes called "**illegitimate recombination.**" Although no homology is required, there are prefered sites for integration. For example, Tn10 is transposed most readily into certain "hot spots" in the *Salmonella* DNA among which is the sequence 5'-GC^{m5}CAGGC.[608] Illegitimate recombination can also be induced by other processes that involve DNA chain cleavage, e.g., by topoisomerases.[644] Whenever a DNA chain breaks, it must be repaired, a process that often also involves recombination. Recombination is often observed to occur between direct repeat sequences, which are a major cause of instability in the genome.[645]

Under some circumstances selected segments of the genome are **amplified** by repeated replication of a gene or genes.[646,647] Amplication of specific genes occurs in viruses,[648,649] in bacteria where it may provide for adaptation to conditions of stress,[650] and in eukaryotes.[651] In oocytes of amphibia, such as *Xenopus*, excess DNA accumulates around the nucleoli and later breaks up to form 1000 or more separate nucleoli. As many as 3000 copies of the rDNA (which forms a distinct satellite band upon centrifugation) may be present. Much of this DNA exists as extrachromosomal rings containing 1–20 rDNA units. Using these genes as many as 10^{12} ribosomes per oocyte are synthesized.

Tetrahymena contains only one set of rRNA genes per haploid genome in its diploid **micronucleus**. Following sexual conjugation the chromosomes of the micronucleus undergo multiple replications to form polytene chromosomes containing thousands of copies. However, about 5% of the resulting DNA is excised as linear pieces with characteristic inverted repeat sequences and 3' single-stranded tails:

$$\begin{array}{l} 5'\text{--}\overrightarrow{C_4A_4C_4A_4C_4}\text{--}/\!/\text{--}G_4T_4G_4T_4G_4T_4G_4T_4G_4\text{--}3' \\ 3'\text{--}G_4T_4G_4T_4G_4T_4G_4T_4G_4\text{--}/\!/\text{--}\underleftarrow{C_4A_4C_4A_4C_4}\text{--}5' \end{array}$$

These tails are joined by a protein to form the circles, which are segregated in the **macronucleus**. The other 95% of the amplified DNA is degraded. At the next meiosis the macronucleus is discarded entirely, and a new one is formed at the next diploid stage.

Toxic drugs often cause cells to amplify genes that help resist the drug.[647] This can be a major problem in the chemotherapy of cancer. For example, a culture of human leukemia cells grown in the presence of increasing concentrations of methotrexate increased its level of dihydrofolate reductase 240-fold.[652] The cause is an increase in the number of copies of a chromosomal region containing the gene.[653] Cancer cells tend to amplify oncogenes such as c-*myc*.[654]

There are several mechanisms for gene amplification. The formation and breakup of polytene chromosomes in *Tetrahymena* is one. The circular copies of rDNA in *Xenopus* and many other species are generated by a rolling circle mechanism similar to that in Eq. 27-7. If each circle excised from the original chromosomes contains an origin of replication, many copies can be formed. Another possible mechanism is replication of a local region of DNA several times followed by excision of pieces of the DNA. A mechanism that may give rise to homogeneously staining regions is unequal crossing-over (Section 2) repeated several times within a gene cluster. Transposition can cause excision of DNA in a circular form, which can be amplified by a rolling-circle mechanism.

E. Damage and Repair of DNA

A characteristic of living things is their high degree of mutability. Harmful mutations take a toll of human life at an early age, and the very high incidence of cancer in older persons is largely a result of the accumulation of somatic mutations. Mutations are a major factor in aging and are continuously introducing new genetic defects into the population. Mutations can be described as **base substitutions**, **deletions**, or **additions**. Base substitution mutations are classified as **transitions**, in which a pyrimidine in one strand is replaced by a different pyrimidine. In the complementary strand a purine is replaced by the other purine, e.g.,

$$C\cdot G \rightarrow \colorbox{lightgreen}{T}*G \xrightarrow{\text{Replication}} \colorbox{lightgreen}{$T\cdot A$} + C\cdot G \qquad (27\text{-}19)$$

$$A\cdot T \rightarrow \underset{\text{Mispair}}{\colorbox{lightgreen}{G}*T} \longrightarrow \underset{\text{Mutation}}{\colorbox{lightgreen}{$G\cdot C$}} + A\cdot T \qquad (27\text{-}20)$$

In a *transversion* a purine in one chain is replaced by a pyrimidine, while the pyrimidine in the complementary chain is replaced by a purine:

$$C\cdot G \rightarrow \colorbox{lightgreen}{A}*G \longrightarrow \colorbox{lightgreen}{$A\cdot T$} + C\cdot G \qquad (27\text{-}21)$$

Here the central asterisks designate mispairs and the green shade marks mismatched bases and also the resulting mutant base pair formed in the next replication cycle.

1. Causes of Mutations

DNA can be damaged in many ways.[655] Spontaneous hydrolysis of the glycosidic bonds between nucleic acid bases and the deoxyribose to which it is connected cause the loss of $\sim 10^5$ purines and pyrimidine rings per day from the DNA in a mammalian cell.[656] About 100 residues per day of cytosine are deaminated by such agents as nitrite or bisulfite (Chapter 5, Section H,3) to form uracil. Like thymine, uracil will pair with adenine causing a $C \bullet G \rightarrow T \bullet G$ transition mutation as in Eq. 27-19. A few adenines per day are also deaminated to form hypoxanthine. Oxidized bases are formed by attack of $HO^\bullet$ radicals and other species of reduced oxygen.[657–659] Alkylating agents from the external environment as well as *S*-adenosylmethionine carry out slow, nonspecific alkylation of purine and pyrimidine bases. Polycyclic aromatic hydrocarbons and other carcinogens are converted to metabolites that alkylate DNA, and alkylated bases often mispair during replication.[660,661] Ultraviolet light induces formation of photohydrates (Eq. 23-25),[662] pyrimidine dimers (Eq. 23-26), and other photochemical products.[663] X-rays and gamma rays cleave nucleic acid bases and break chromosomes.[663a]

Natural radioactivity has a measurable effect.[663b] Mistakes caused by mispairing[664] or misalignment[665] are made during replication and recombination and even during repair of DNA. Errors in replication are especially numerous in highly repetitive DNA sequences. Some errors probably arise as a result of tautomerization and others from incorporation of uracil in place of thymine. Thus, to keep its DNA in repair a cell must continuously deal with missing bases, wrong bases, altered bases or sugars, pyrimidine dimers and other crosslinkages, deletions, and insertions.

2. Fidelity of Replication

During DNA replication in *E. coli* only one mistake is made on the average during polymerization of 10^9–10^{10} nucleotides.[666–668] The rate varies among different sites in the genome.[668a] In eukaryotes the error rate may be only 1 in 10^{-10} base pair or less per generation.[669,669a] Early workers often attributed the specificity in base pairing and the resultant high precision of replication entirely to the strength of the two or three hydrogen bonds formed together with the stabilization provided by the adjacent helix. However, the Gibbs energy of formation of the base pairs is small (Chapter 5), and the additional energy of binding to the end of an existing helix is insufficient to account for the specificity of pairing.[670,671] Thus, according to Eq. 6-30 a difference in ΔG of binding between the correct nucleotide and an incorrect one of 11 kJ / mol would give an error rate of 1 in 100.

The role of polymerases. The polymerase enzymes play a major role in ensuring correct pairing during replication, transcription, and protein synthesis. RNA and DNA polymerases are large molecules. The binding site on the enzyme can completely surround the double helix. Water will be excluded as the enzyme folds around the base pairs. This may have the effect of greatly increasing differences in the Gibbs energies of binding and thereby enhancing selectivity.[671] It has traditionally been assumed that formation of proper hydrogen bonds is essential for the base selection. A hypothetical active site is portrayed in Fig. 27-33. A guanine ring of the template strand of DNA is portrayed at the point where the complementary DNA strand is growing from the 3' end. The proper nucleoside triphosphate must be fitted in to form the correct GC base pair before the displacement reaction takes place to link the new nucleotide unit to the growing chain. Let us suppose that the enzyme possesses binding sites for the deoxyribose unit of the template nucleotide and for the sugar unit of the incoming nucleoside triphosphate and that the two binding sites are held at a fixed distance one from the other. As indicated in Fig. 27-33, some group H-Y might also be present at each binding site to hydrogen bond to the nitrogen or oxygen indicated by the heavy green arrows. All four of the bases could form hydrogen bonds of this type in the same position. Hydrophobic interactions could provide additional stabilization. With such an arrangement the correct nucleoside triphosphate could be selected no matter which one of the four bases occupied the binding site on the left side of the figure. (The outlines of the thymine and adenine rings have been drawn in with dotted and dashed lines, respectively.) If a purine is present on the left side, as is shown in the drawing, there is room on the right side only for a pyrimidine ring. Thus, A and G are excluded, and the choice is only between C and U (or T). However, U will be excluded because the dipoles needed to form the hydrogen bonds point in the wrong direction. These dipolar groups are hydrated in solution. They are unlikely to give up their associated water molecules unless hydrogen bonds can be formed within the base pair. Not only would a molecule of U (or T) be unable to form the stabilizing hydrogen bonds within the vacant site but also the electrostatic repulsion of the like ends of the dipoles would tend to prevent association. This would lower the affinity of the polymerase for mispaired bases. Verification that the proper base pair has been formed could be accomplished by using its tautomeric properties to sense an electronic

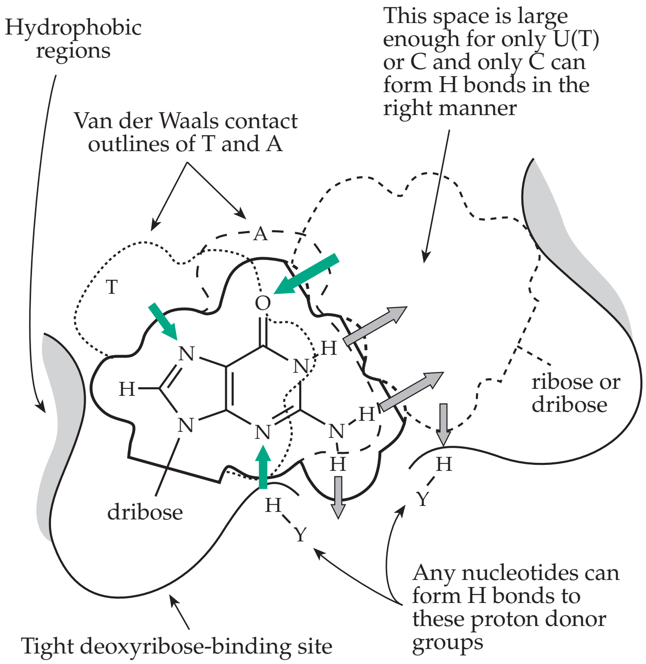
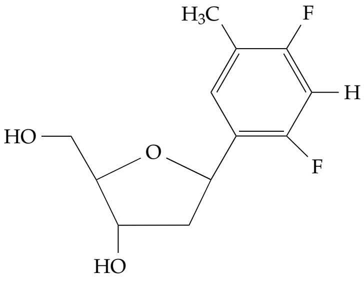

Hydrophobic regions
This space is large enough for only U(T) or C and only C can form H bonds in the right manner
Van der Waals contact outlines of T and A
A
T
O
N
H
N
H
N
N
N
H
H
H
ribose or dribose
dribose
H
Y
H
Y
Any nucleotides can form H bonds to these proton donor groups
Tight deoxyribose-binding site

Figure 27-33 Selecting the right nucleotide for the next unit in a growing RNA or DNA chain. A deoxyguanosine unit of the template chain is shown bound to a hypothetical site of a DNA polymerase.

displacement through the hydrogen-bonded network.[672,673] Considerable experimental evidence suggests, however, that hydrogen bonding may not be important to the selectivity of DNA polymerases.[666,674–676] For example, the triphosphate of the following analog of thymidine is efficiently incorporated by *E. coli* DNA polymerase I into DNA opposite the thymine base-pair partner adenine.[675]



Structural analog of thymidine

Similar results have been obtained for a variety of other analogs with poor hydrogen bonding characteristics. Apparently *shape* is most important.[666,674,676,676a] Binding strengths in transition state structures must also be considered.[676b]

Editing (proofreading). As is discussed in Section C,2, base-pairing is checked after a new monomer is added at the 3' end of a growing polynucleotide chain, as well as before polymerization occurs. If the wrong base pair has been formed, the newly created linkage is hydrolyzed, and the incorrect nucleotide is released. It has been estimated[677] that in *E. coli* the error rate for DNA polymerase III holoenzyme is $\sim 1 \times 10^{-7}$ per base pair of which proofreading by the ε subunit may provide $\sim 10^{-2}$. Additional mismatch repair reduces the error rate 200- to 300-fold to give the overall error rate of $<10^{-9}$. Fidelity of replication, which seems to be higher on the lagging strand than on the leading strand,[678] may also be improved by operation of other proteins, such as the sliding clamp, that are part of the replication apparatus.[679] Replicative proteins also appear to be designed to minimize **frameshift mutations** that could arise by slippage of the template versus the replicated strand in long runs of dA•dT pairs.[680]

3. Repair of Damaged DNA

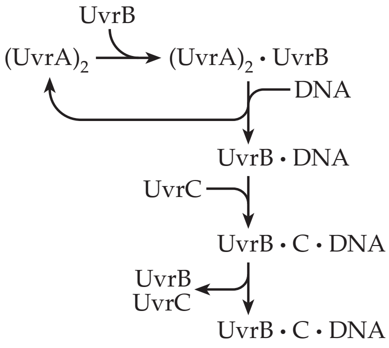
A final check of the fidelity of replication is made after a new strand has been formed. Mismatched base pairs are identified, and the incorrect nucleotides are cut out and replaced by correct ones.[655,670,681–683] Some of the thymine dimers created by the action of light are also repaired photochemically by photolyases (see Chapter 23). **Photoreactivation** was the first DNA repair process recognized.[684] However, most thymine dimers in human DNA must be excised and replaced. Loops, gaps, and double-stranded breaks are also sensed, and appropriate corrections are made. If necessary the sequence of a badly damaged segment of DNA may be copied from the DNA in a sister chromatid before mitosis is completed or from the homologous chromosome.[685]

Much of our understanding of DNA repair comes from investigations of mutant strains of *E. coli*, which display an elevated incidence of mutations. On this basis, the “mutator” genes *dam*, *mutD*, *mutH*, *mutL*, *mutS*, *mutU* (also called *uvrD*), and *mutY* were implicated in DNA repair.[677,686] An additional series of genes, *uvrA*, *B*, *C*, and *D*, were identified by their participation in resistance to ultraviolet radiation damage and were also recognized as being involved in DNA repair. Many corresponding genes were found later in the yeast, *Drosophila*, and human genomes. The study of human DNA repair has also been facilitated by the recognition of a group of human inherited defects (Box 27-A). There are 130 known human DNA repair genes.[686a]

Methyl-directed and other mismatch repair. In *E. coli* the methylase encoded by gene *dam* (Fig. 29-4), within ~7 min of replication, methylates adenine at N-6 in the sequence GATC, which occurs at many locations.[670,686,687] This methylation provides a label for the template strand in a newly replicated duplex. The GATC sites in the newly synthesized strand will not yet be methylated, and repair enzymes recognize it as the strand on which to carry out excision repair. Another system, which may function during recombination, acts on fully *dam* methylated DNA.[655] Since eukaryotes do not use methylation of GATC sites to distinguish the old template and newly replicated DNA strands, other mechanisms must operate.[669]

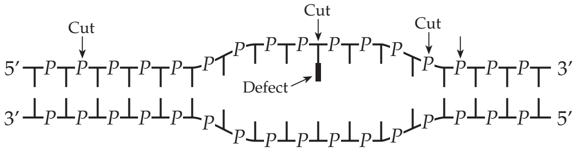
Ten proteins are required for methyl-directed mismatch correction. Of the required *mut* genes, *mut*D (also called *dnaQ*) was found to be the structural gene for the proofreading subunit ε of DNA polymerase III (Table 27-2).[688] *MutH*, *L*, and *S* as well as ATP are also needed, as is a helicase, an exonuclease, DNA pol III, and a ligase. Homodimers of proteins MutS and MutL preferentially bind to DNA with mismatched base pairs such as Pur-Pur (e.g., A•G) or Pyr•Pyr (less well recognized). Some of these mismatched pairs may contain modified bases. After **recognition** by the MutS•MutL complex (see banner, p. 1527)[688a–c] the endonuclease MutH is activated and cuts the DNA chain at the nearest unmethylated GATC sequence.[689] An exonuclease then degrades the chain past the mismatched base leaving a gap that must be filled by the action of DNA polymerase and DNA ligase. This **long-patch mismatch repair** (**MMR**) is utilized also by eukaryotes,[688a,b,690–692a] using proteins homologous to *E. coli* proteins MutS and MutL. However, methylation is not involved. In both yeast and human beings there are at least six MutS homologs, some of which have functions other than repair. For example, in yeast MutS- and MutL-related proteins are essential for normal levels of meiotic crossing-over.[692] Defects in human mismatch proteins may cause hereditary non-polyposis colorectal cancer (Box 27-A).[692b]

Excision repair. The *E. coli* mismatch repair is a type of excision repair. However, a different **nucleotide excision repair** system (**NER**) is utilized by all organisms from bacteria to human to remove a variety of defects. These include thymine dimers, photohydrates, oxidized bases, adducts of cisplatin (Box 5-B), mutagens derived from polycyclic aromatic compounds,[683] and poorly recognized C•C mismatched pairs.[692] In *E. coli* this excision repair process depends upon proteins encoded by genes ***UvrA, B, C,*** and ***D*** and also DNA polymerase I and DNA ligase.[693–695a] A dimer of protein UvrA forms a complex with helicase UvrB (Eq. 27-22).[696,696a]



(27-22)

The helicase, driven by ATP hydrolysis, may move along the DNA chain with its associated UvrA protein in search of defects.[696] When one is located UvrB binds tightly.[697] UvrA then dissociates, and the nuclease UvrC binds and cleaves the DNA chain in two places as in the following scheme. One is at the fourth or fifth phosphodiester linkage in the 3' direction from the defect. The other is at the eighth phosphodiester on the 5' side.[698] The resulting gap is filled by a DNA polymerase and DNA ligase. Another helicase, encoded by *UvrD*, is also required.[699]



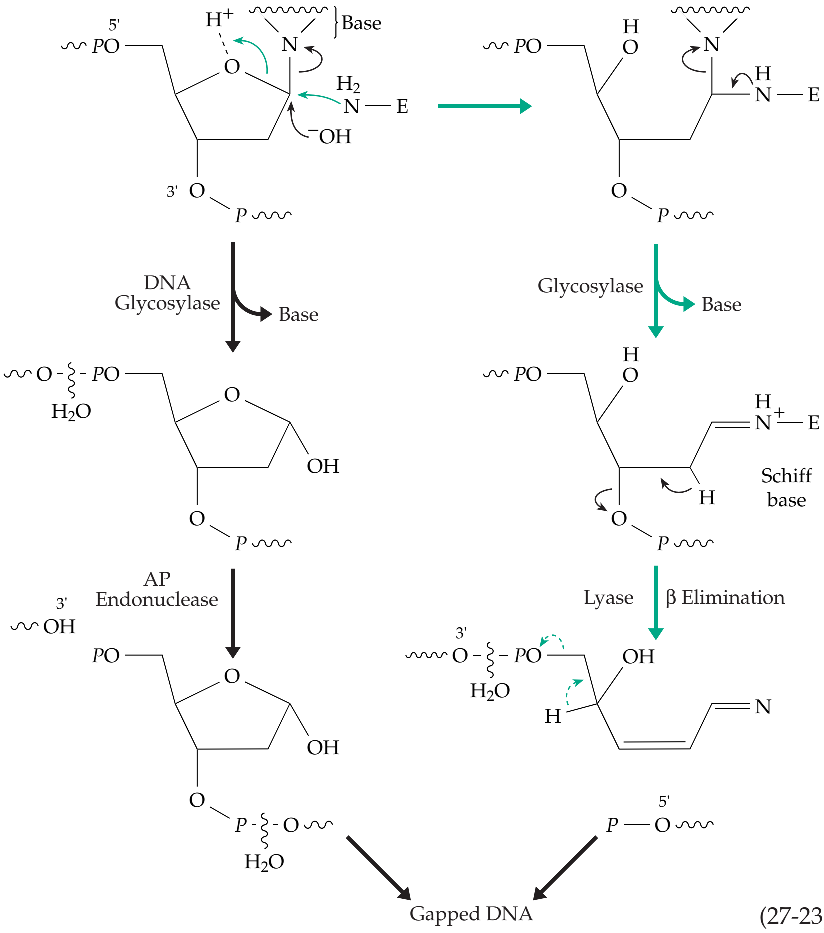
UvrBC preincision complex

Both yeast and human cells have similar but more complex systems of nucleotide excision repair.[700–703] The dual incision steps require at least six components. Several of these (XPA, C, F, G) have been found defective in various forms of the inherited disease **xeroderma pigmentosum** (Box 27-A), in which there is a high incidence of UV-induced skin cancer.[700,703a] Replication protein A (Section C,10), a three-subunit ssDNA binding protein, is also essential.[704] In eukaryotic excision repair the DNA is cleaved at the same position as in the bacterial preincision complex (preceding scheme), on the 3' side of the defect but further out on the 5' side, the excised polynucleotide being ~24–32 nucleotides in length. Repair of the resulting gap in the DNA duplex is accomplished by synthesis mediated by DNA polymerase δ or ε and action of a DNA ligase.

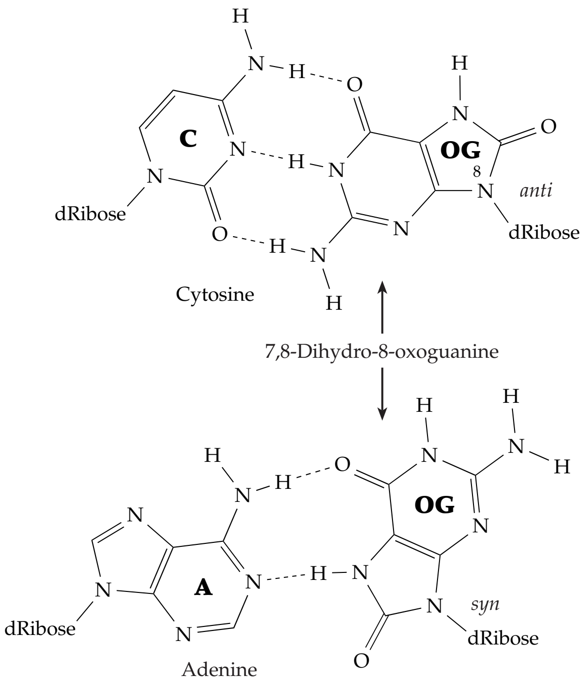
One of the lesions removed by nucleotide excision is the thymine photodimer.[705] In the fission yeast *S. pombe* an **alternative excision repair** system, specialized for removal of thymine dimers and 6–4 photoproducts (Chapter 23), produces two-nucleotide gaps with 3'-OH and 5'-phospho-group ends.[702,705a] Alternative NER pathways are also employed by bacteria.[705b]

Base excision repair (BER). The *N*-glycosyl linkages of the purine and pyrimidine bases to the deoxyribose residues of the sugar–phosphate backbone of DNA are subject to spontaneous hydrolysis, one important source of damage to DNA. Similar hydrolytic reactions are catalyzed by **DNA glycosylases**, which remove many mismatched or damaged bases.[706–708a] At least seven enzymes of this type are present in cells of *E. coli*. One of them, a **uracil-DNA glycosylase**, hydrolyzes the glycosyl linkage wherever uracil has been incorporated accidentally in place of thymine or has been produced by deamination of cytosine.[706,709–711d] Acting via a nucleophilic displacement mechanism (Eq. 27-23, black arrows), it removes the uracil, leaving the DNA backbone intact but with an **apyrimidinic site** in one chain. The resulting apyrimidinic (apurinic sites) are recognized and cleaved by **apurinic / apyrimidinic DNA endonucleases** (AP nucleases).[708,712–714] These enzymes cut the DNA backbone, some leaving a 3'-phosphate group and a 5'-OH terminus and others a 3'-OH and 5'-phosphate. The resulting gap can be filled, for example by the action of DNA polymerase I and a ligase. To prevent the eliminated uracil from being converted to dUTP and reincorporated into DNA a **deoxyUTPase**, essential to both *E. coli* and yeast, hydrolyzes dUTP to dUMP and inorganic phosphate, decreasing the concentration of dUTP.[715] Other enzymes, known as **DNA glycosylase/AP lyases,**[706,716–718] use an amino group of an enzyme side chain as a nucleophile in a ring-opening reaction that is followed by β elimination of the nucleotide base with formation of a Schiff base intermediate (green arrow, Eq. 27-23). An example is the bacteriophage T4 enzyme **T4 endonuclease V**.[706,719] Others include the *E. coli* **endonuclease III**, encoded by the **Nth** gene, its yeast homolog **Nth-Spo**,[717–718] and human DNA pol ι (iota).[718a] After the Schiff base is formed, the DNA backbone is cut on the 3' side of the abasic site by a second β elimination reaction. A δ elimination (green dashed arrows in Eq. 27-23) may also occur, cleaving the DNA chain on the 5' side of the lesion. Alternatively, a hydrolytic cleavage is possible as is also indicated in Eq. 27-23. Either type of DNA glycosylase forms a single-nucleotide gap or a gap missing just a few nucleotides. This gap can also be filled by polymerase and ligase action.

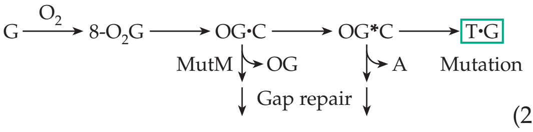
Both NER and BER forms of excision repair remove a great variety of defects, many of which are a result of oxidative damage.[657,720] Most prominent among these is **7,8-dihydro-8-oxoguanine** (8-OG), which is able to base pair with either cytosine (with normal Watson–Crick hydrogen bonding) or with adenine, which will yield a purine–purine mismatch and a C•G → A•T transversion mutation (Eq. 27-24), a frequent mutation in human cancers.[721,722]


5'
PO
H+
Base
N
O
H2
N — E
–OH
3'
O
P
H
O
DNA
Glycosylase
Base
Glycosylase
O-
PO
H2O
OH
Schiff
base
AP
Endonuclease
3'
OH
Lyase
β Elimination
H2O
H
N
P-
O
5'
Gapped DNA

(27-23)


C
OG
8
anti
dRibose
Cytosine
7,8-Dihydro-8-oxoguanine
A
OG
syn
dRibose
Adenine

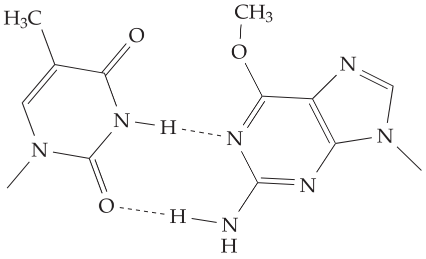
In *E. coli* three enzymes protect against 8-OG DNA mismatches. A glycosyltransferase encoded by the *mutM* (or *fpg*) gene removes 8-OG from DNA.[723,724] Because some A∗OG base pairs may remain, a second glycosyltransferase (**MutY**) removes adenine from them.[723,725–725b] Interestingly, MutY contains an Fe_4S_4 cluster, essential to its function.[725,726] Both yeast and human cells have a corresponding enzyme.[727,728] The third *E. coli* enzyme **(MutT)** is a nucleoside triphosphate pyrophosphohydrolase, similar to the previously mentioned dUTPase. It preferentially hydrolyzes free 8-oxo-dGTP, preventing its incorporation into DNA.[729]


G
O2
8-O2G
OG•C
OG*C
T•G
MutM
OG
A
Mutation
Gap repair

(27-24)

Other DNA glycosylases remove thymine rings that have been converted to saturated or fragmented forms by oxidizing agents or ionizing radiation.[720,730,731] Among these are 5,6-hydrated thymine and urea, which are still attached in ribosyl linkage. Purines such as hypoxanthine and 3-methyladenine can all be removed by glycosylases.

Reactions of alkylated bases. The O^6 alkylation of guanine rings in DNA is highly mutagenic, presumably because mispairing with thymine will cause C•G → T•A mutations.



Thymine–O^6-methylguanine pair

Both bacteria and higher eukaryotes have an enzyme-like **O^6-methylguanine–DNA methyltransferase,** whose synthesis may be induced by culture of cells in the presence of alkylating agents. This 354-residue Zn^{2+}-containing **Ada protein** acts as both acceptor and catalyst for transfer of the methyl group off from the O^6-methylguanine onto the sulfur atom of a cysteine side chain in the protein.[732–734b] The presence of alkylating agents also induces synthesis of DNA glycosylases that remove 3-methyladenine and other alkylated bases.[735,736] 5-Methylcytosine, present in CpG islands in human DNA, will occasionally be hydrolyzed to thymine, giving T∗G mismatches. The mismatched thymines are removed in cells of *E. coli* by a specialized endonuclease (**vsr** gene),[737] which hydrolyzes the phosphodiester linkage preceding the mismatch. In eukaryotic cells a **thymine–DNA glycosylase**[738,739] accomplishes the removal.

Repair of double-strand breaks. Following replication both in bacteria and in eukaryotes, recombinational exchanges may occur between sister chromatid duplexes or between homologous pairs of chromosomes.[740] A newly replicated chain may have gaps because of defects in the template strand from which it was copied. Copying from a sister duplex may permit assembly of a correct chromosome sequence and survival for the cell.[551a] Recombinational repair is also a major mechanism for preventing loss of genetic information from **double-strand breaks**. Such breaks are induced by ionizing radiation or chemical damage. They are often created at replication forks, stalling DNA synthesis with potentially disastrous consequences to the cell. Repair of these breaks, which is required for completion of replication, depends upon a large complex of replication and recombination proteins.[741–741b] In *E. coli* this includes protein RecF.[741,742] Repair of double-strand breaks by homologous recombination may be the most frequently employed type of DNA repair in bacteria, yeasts such as *S. cerevisiae*[530d] and also in mammals and other organisms.[742a–d]

A second repair pathway, called **nonhomologous end-joining** (**NHEJ**), is also utilized and may be relatively more important in higher organisms.[743–746b] The NHEJ mechanism repairs double-strand breaks caused by ionizing radiation and is also employed in specialized genomic rearrangements of developing B and T cells of the immune system. This V(D)J recombination (Chapter 31) is used to create the huge array of antibodies and antigen receptors required for immunological recognition and protection.[745,746a,747,748] NHEJ doesn't depend upon templates but simply rejoins duplex ends. Errors can be made if the ends have been damaged or if the wrong ends are joined. Several special proteins are required. One is a DNA-activated protein kinase (**DNA-PK**)[749] and a heterodimeric **end-binding protein** with subunits **KU70** and **KU80**.[745,750,750a] The serine / threonine protein kinase ATM (Box 27-A) is also activated and phosphorylates protein p53 in response to γ-irradiation of cells.[750] This is thought to be part of a signaling pathway for control of the cell cycle following DNA damage. Mitosis may be delayed while repair is completed. Alternatively, the damaged cell may be killed by apoptosis.

Subjects of current interest are the mechanisms by which completion of DNA synthesis is signaled at the various checkpoints in the cell cycle.[750b,c] These include points in addition to those marked in Fig. 11-15. At every point signals must accumulate to indicate that replication must be delayed to allow repair or that the cell must be allowed to die by apoptosis. Failure to accomplish repair may lead to cancer.[750d] Even one double-strand break will prevent the completion of mitosis.[750b] Among other problems faced by a dividing cell are slow replication and stalling at replication forks.[750c] A range of chemical alterations in chromatin and other proteins are associated with DNA damage and repair.[750b,e–h] Among these are conjugation of PCNA with ubiquitin and SUMO (Fig. 27-12D).[750i]

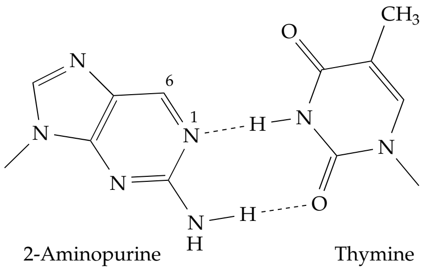
The SOS response and translesion repair. If cells of bacteria or eukaryotes are heavily damaged by UV, X-ray irradiation, or mutagenic chemicals, an emergency or **SOS response** is initiated.[711a,741,751–753a] In *E. coli* the two proteins specified by genes *lexA* and *recA* initiate the response. The *lexA* protein is a repressor that prevents transcription of a group of SOS genes (see Chapter 28). It is thought that some product from damaged DNA activates the RecA protease activity. The activated RecA protein then cleaves the lexA protein allowing transcription of the SOS genes. The SOS response is transient but complex. It includes increased recombinational activity, alterations in replication initiation, inhibition of nucleases, and induction of an **error-prone DNA synthesis**. The cell now replicates DNA more rapidly than normal but with an increased frequency of errors.[754,754a] For example, it will bypass thymine dimers and other defects, and it will continue the DNA synthesis even though incorrect bases have been put into the chain opposite the thymine dimer. Later the errors may be corrected by recombinational repair or by photoreactivation. In *E. coli* this translesional repair depends upon DNA polymerase III, the RecA protein, and two proteins encoded by genes *umuC* and *umuD*. These genes were recognized as providing resistance to mutations induced by UV radiation.[753–754b] Study of similar genes in yeast led to the discovery of two new and unusual DNA polymerases essential for translesional repair. DNA polymerase ζ (zeta), encoded by gene *REV*3 and *REV*7 subunits, bypasses abasic sites by inserting a dCMP residue into the growing DNA chain. Because C may be the wrong base the process is "error-prone." Polymerase η (eta) bypasses thymine photodimers by placing two consecutive dAMP residues in the growing strand. This is an error-free process.[755,756] Genes for human polymerases ζ and η have also been identified,[757–759b] and additional polymerases have been found in yeast and other organisms.[760–760e] The human XP-V gene for polymerase η is defective in xeroderma pigmentosa variant type (Box 27-A). Polymerase ι may act sequentially with pol η to bypass highly distorting lesions.[760b] The very imprecise pol θ may also be used to bypass hard-to-remove lesions.[760c] Polymerase λ is thought to play a role in DNA repair during meiosis,[760d] while pol κ is needed in some way to provide cohesion between sister chromatids.[760e] Some of these DNA polymerases are very inaccurate.[760a] It has been suggested that human pol ι may participate in hypermutation of immunoglobin variable genes.[760f] Polymerase ϕ of yeast may be required for synthesis of ribosomal RNA.[760g]

Poly (ADP-ribose). A eukaryotic peculiarity, which is not well understood, is the synthesis of poly(ADP-ribose) chains attached to many sites in nuclear proteins (see also Eq. 15-16). Increased synthesis is observed following damage to DNA.[761–763a] The poly (ADP-ribose) polymerase binds to DNA near strand breaks or nicks and, using NAD^+ as a substrate, synthesizes the highly branched polymer attached to a small number of nuclear target proteins.

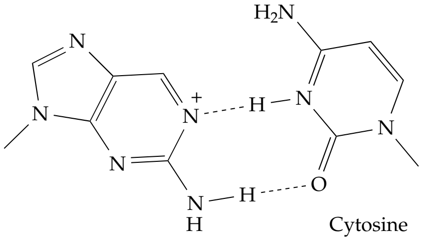
F. Mutagens in the Environment

More than 500 new chemicals are introduced into the environment industrially each year. Some widely used drugs, e.g., **hycanthone** (Fig. 5-22), are mutagenic. Powerful mutagens are present naturally in some foods.[764–766] Others have been added through ignorance. Although many of these have now been removed, the problem cannot be ignored.

One way in which chemical compounds can induce base substitution mutation is through their incorporation into the structure of DNA itself. Thus, 5-bromodeoxyuridine (or bromouracil) can replace thymidine in DNA, where it serves as an efficient mutagenic agent.[767] 2-Aminopurine, an analog of adenine, pairs with thymine, just as does adenine when incorporated into DNA.



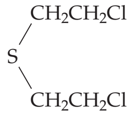
However, if it protonates on N-1, it can pair with cytosine, causing mutation.[768,769]



Another mutagenic base is N^6-hydroxylaminopurine.[770]

Many alkylating agents are powerful mutagens. Alkylation can occur at many places in DNA, but the N-7 position of guanine is especially susceptible (Eq. 5-18). The resultant positive charge on the imidazole ring portion of the alkylated guanine causes hydrolysis of the *N*-glycosyl linkage and depurination. However, this may be a lethal rather than a mutagenic event. The previously mentioned methylation on O-6 of guanine is probably more important in inducing mutations.[771]

Among the most biologically reactive alkylating agents are the nitrogen and sulfur "mustards" such as bis-(2-chloroethyl)sulfide. These toxic bifunctional compounds cause lethal crosslinking of DNA chains (see Eq. 5-20). The monofunctional half-mustards are mutagenic but are less acutely toxic. Another group of alkylating agents, the **nitrosamines** (Chapter 5), are



bis-(2-chloroethyl)sulfide

BOX 27-A DEFICIENCIES IN HUMAN DNA REPAIR

Problems with human repair systems account for several serious diseases. One is the rare skin condition **xeroderma pigmentosum** (**XP**), an autosomal recessive hereditary defect. Homozygous individuals are extremely sensitive to ultraviolet radiation and have a high incidence of multiple carcinomas.[a–e] Defects in at least seven human genes, listed in the accompanying table, can cause XP. Most of the genes have been cloned and found to be homologs of genes for nucleotide excision repair (NER). In rodents the *ERCC* genes have the same functions as do genes for resistance to ultraviolet light in yeast (*RAD* genes). This information has elucidated the functions of the XP genes and the basis for this group of human diseases.[a,b,f]

The XP-A protein (see table) is a zinc-finger-containing DNA-binding protein, which interacts with human replication protein A (Section C,10). It may also interact with proteins corresponding to the rodent ERCC1 and ERCC4 proteins to form a complex that recognizes DNA defects.[g–i] Proteins XPC and XPE may be specialized DNA binding proteins for cyclobutane dimers and other photoproducts.[j,k] The XPG protein is an endonuclease that probably cuts on the 3' side of the DNA defect.[a,l] In human cell lines with a variety of XP mutations the DNA problems have been corrected by transfer of the corresponding normal genes into the cells.[m,n]

While defects in protein XPD often cause typical XP symptoms, some defects in the same protein lead to **trichothiodystrophy** (TTD, brittle hair disease). The hair is sulfur deficient, and scaly skin (ichthyosis, Box 8-F), mental retardation, and other symptoms are observed.[o] Like their yeast counterparts (proteins RAD3 and RAD25), XPB and XPD are both DNA helicases.[o] They also constitute distinct subunits of the human transcription factor TFIIH[p], which is discussed in Chapter 28. It seems likely that XPD is involved in **transcription-coupled repair** (**TCR**) of DNA.[o,q–s] This is a subpathway of the nucleotide excision repair (NER) pathway, which allows for rapid repair of the transcribed strand of DNA. This is important in tissues such as skin, where the global NER process may be too slow to keep up with the need for rapid protein synthesis. Transcription-coupled repair also appears to depend upon proteins CSA and CSB, defects which may result in the rare **cockayne syndrome**.[b,o,t,u] Patients are not only photosensitive but have severe mental and physical retardation including skeletal defects and a wizened appearance.

In a variant form of XP, designated XP-V, nucleotide excision repair is normal, but DNA replication is very slow. Postreplicational translesional repair (bypass repair) is also slow, and patients are cancer-prone.[v,w] The recently discovered DNA polymerase η may be defective.[x]

Children with **ataxia telangiectasia** (**AT**) have progressive neurological problems, a weak immune system, premature aging, and a high incidence of cancer.[y] Their skin fibroblasts are deficient in the ability to repair X-ray damage, which causes many double-strand breaks. Apparently in this disease cells do not wait until DNA repair has been carried out after exposure to ionizing radiation but attempt to replicate the damaged DNA.[y] The defective protein **ATM** (ataxia telangiectasia, mutated) has been identified as a large 370-kDa Ser / Thr protein kinase with a carboxyl terminal domain similar to phosphatidylinositol 3-kinase.[z,aa] It appears to play a crucial role in the cell cycle DNA damage checkpoints (Fig. 11-15) by participating in the detection of double-strand breaks and in delay of replication, while they are repaired by homologous recombination.[y,bb,cc] Although AT is an autosomal recessive disease, women heterozygous for the ATM gene have an increased susceptibility to breast cancer. This observation led to the discovery that the proteins encoded by the well known breast cancer genes *Brca1* and *Brca2* form a complex with ATM. Phosphorylation of BRCA1 by ATM may initiate a signaling pathway through the p53, c-Abl, and Chk2 proteins that cause cell cycle arrest (Fig. 11-15). BRCA1 and BRCA2 also form a complex with protein RAD51, a RecA homolog necessary for homologous recombination. BRCA1 also may be essential to transcription-coupled repair.[dd,de]

There are two major forms of hereditary susceptibility to colon cancer.[ee] Familial adenomatous polyposis is caused by defects in the *APC* gene (see Chapter 32). The more common **hereditary non-polyposis colorectal cancer** (**HNPCC**), which includes many endometrial, stomach, and urinary tract tumors, results from defects in DNA mismatch repair.[ff–jj] The proteins hMSH2 and hMSL1 are homologs of the *E.coli* MutS and MutL (main text).

Cells of patients with **Bloom syndrome** (**BS**) have many chromosome breaks and a high frequency of sister chromatid exchanges, perhaps in an effort to correct these breaks. The body is small but well-proportioned.[kk] A somewhat similar disease, the **Werner syndrome** (**WS**), is associated with premature aging.[ll] The Bloom's protein **BLM** and the WS gene product **WRN** are both helicases related to *E.coli* RecQ. Protein BLM colocalizes with replication protein A as discrete foci in the meiotic synaptonemal complex.[mm] Protein WRN also seems to be associated with DNA replication. Defects

BOX 27-A DEFICIENCIES IN HUMAN DNA REPAIR (continued)

appear to increase homologous and illegitimate recombination.[nn] Both proteins may also function in transcription.[oo]

Many other diseases leading to a high incidence of cancer are known. Among them is the Nijmegan breakage syndrome, in which chromosomes are hypersensitive to breakage by ionizing radiation. The gene has been identified by positional cloning, and its protein is apparently involved in repair of double strand breaks.[pp] **Fanconi anemia**, **Gardner syndrome**, and hereditary **retinoblastoma** (Box 11-D) may also involve defects in DNA repair.

Some Human Hereditary Defects of DNA Repair

Human disease	Human gene involved		Yeast gene	Function
Xeroderma pigmentosum (XP)				
XP-A	*XPA*		*RAD*14	DNA-binding, damage recognition
XP-B	*XPB*	*ERCC*3	*RAD*25 (*SSL*2)	DNA helicase
XP-C	*XPC*		*RAD*4	DNA-binding, thymine dimers
XP-D	*XPD*	*ERCC*2	*RAD*3	DNA helicase
XP-E	*XPE*		*RAD16*	
XP-F	*XPF*		*RAD*10	DNA nuclease complex
XP-G	*XPG*		*RAD*2	DNA nuclease
XP-V	*XPV*		*RAD*30	DNA polymerase η
Cockayne syndrome (CS)				
CSA	*CSA*			Interaction with helicase CSB and with TFIIH
CSB	*CSB*			DNA helicase
Trichothiodystrophy (TID, brittle hair disease)				
	XPDE			
Ataxia telangiectasia (AT)				
	ATM			Cell cycle delay for repair of ds breaks
Human nonpolyposis colorectal cancer (HNPCC)				
	*hMSH*2		*MSH*2	Mismatch repair
	*hMSL*1		*MSL*1	
	GTBP			

---

[a] Hoeijmakers, J. H. J., and Bootsma, D. (1994) *Nature (London)* **371**, 654–655

[b] Tanaka, K., and Wood, R. D. (1994) *Trends Biochem. Sci.* **19**, 83–86

[c] de Vries, A., van Oostrom, C. T. M., Hofhuis, F. M. A., Dortant, P. M., Berg, R. J. W., de Gruijl, F. R., Wester, P. W., van Kreijl, C. F., Capel, P. J. A., van Steeg, H., and Verbeek, S. J. (1995) *Nature (London)* **377**, 169–173

[d] Nakane, H., and 19 other authors. (1995) *Nature (London)* **377**, 165–168

[e] Lehmann, A. R. (1995) *Trends Biochem. Sci.* **20**, 402–405

[f] Bootsma, D., and Hoeijmakers, J. H. J. (1993) *Nature (London)* **363**, 114–115

[g] Bankmann, M., Prakash, L., and Prakash, S. (1992) *Nature (London)* **355**, 555–558

[h] Li, L., Elledge, S. J., Peterson, C. A., Bales, E. S., and Legerski, R. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5012–5016

[i] Lee, S.-H., Kim, D.-K., and Drissi, R. (1995) *J. Biol. Chem.* **270**, 21800–21805

[j] Emmert, S., Kobayashi, N., Khan, S. G., and Kraemer, K. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2151–2156

[k] Hwang, B. J., and Chu, G. (1993) *Biochemistry* **32**, 1657–1666

[l] O'Donovan, A., Scherly, D., Clarkson, S. G., and Wood, R. D. (1994) *J. Biol. Chem.* **269**, 15965–15968

[m] Evans, M. K., Robbins, J. H., Ganges, M. B., Tarone, R. E., Nairn, R. S., and Bohr, V. A. (1993) *J. Biol. Chem.* **268**, 4839–4847

BOX 27-A (continued)

[n] Tebbs, R. S., Zhao, Y., Tucker, J. D., Scheerer, J. B., Siciliano, M. J., Hwang, M., Liu, N., Legerski, R. J., and Thompson, L. H. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6354–6358

[o] van Gool, A. J., van der Horst, G. T. J., Citterio, E., and Hoeijmakers, J. H. J. (1997) *EMBO J.* **16**, 4155–4162

[p] Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J. H. J., Chambon, P., and Egly, J.-M. (1993) *Science* **260**, 58–63

[q] Guzder, S. N., Sung, P., Prakash, S., and Prakash, L. (1995) *J. Biol. Chem.* **270**, 17660–17663

[r] Winkler, G. S., Araújo, S. J., Fiedler, U., Vermeulen, W., Coin, F., Egly, J.-M., Hoeijmakers, J. H. J., Wood, R. D., Timmers, H. T. M., and Weeda, G. (2000) *J. Biol. Chem.* **275**, 4258–4266

[s] Reardon, J. T., Bessho, T., Kung, H. C., Bolton, P. H., and Sancar, A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9463–9468

[t] Mu, D., and Sancar, A. (1997) *J. Biol. Chem.* **272**, 7570–7573

[u] Cooper, P. K., Nouspikel, T., Clarkson, S. G., and Leadon, S. A. (1997) *Science* **275**, 990–993

[v] Masutani, C., Araki, M., Yamada, A., Kusumoto, R., Nogimori, T., Maekawa, T., Iwai, S., and Hanaoka, F. (1999) *EMBO J.* **18**, 3491–3501

[w] Johnson, R. E., Kondratick, C. M., Prakash, S., and Prakash, L. (1999) *Science* **285**, 263–265

[x] Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanaoka, F. (1999) *Nature (London)* **399**, 700–704

[y] Beamish, H., Williams, R., Chen, P., and Lavin, M. F. (1996) *J. Biol. Chem.* **271**, 20486–20493

[z] Lavin, M. F., Khanna, K. K., Beamish, H., Spring, K., and Watters, D. (1995) *Trends Biochem. Sci.* **20**, 382–383

[aa] Chong, M. J., Murray, M. R., Gosink, E. C., Russell, H. R. C., Srinivasan, A., Kapsetaki, M., Korsmeyer, S. J., and McKinnon, P. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 889–894

[bb] Morrison, C., Sonoda, E., Takao, N., Shinohara, A., Yamamoto, K.-i, and Takeda, S. (2000) *EMBO J.* **19**, 463–471

[cc] Cortez, D., Wang, Y., Qin, J., and Elledge, S. J. (1999) *Science* **286**, 1162–1166

[dd] Gowen, L. C., Avrutskaya, A. V., Latour, A. M., Koller, B. H., and Leadon, S. A. (1998) *Science* **281**, 1009–1012

[de] Davies, A. A., Masson, J.-Y., McIlwraith, M. J., Stasiak, A. Z., Stasiak, A., Venkitaraman, A. R., and West, S. C. (2001) *Molecular Cell* **7**, 273–282

[ee] Peltomäki, P., Aaltonen, L. A., Sistonen, P., Pylkkänen, L., Mecklin, J.-P., Järvinen, H., Green, J. S., Jass, J. R., Weber, J. L., Leach, F. S., Petersen, G. M., Hamilton, S. R., de la Chapelle, A., and Vogelstein, B. (1993) *Science* **260**, 810–819

[ff] Papadopoulos, N., and 19 other authors. (1994) *Science* **18**, 1625–1629

[gg] Prolla, T. A., Pang, Q., Alani, E., Kolodner, R. D., and Liskay, R. M. (1994) *Science* **265**, 1091–1092

[hh] Kolodner, R. D. (1995) *Trends Biochem. Sci.* **20**, 397–401

[ii] Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'Arrigo, A., Truong, O., Hsuan, J. J., and Jiricny, J. (1995) *Science* **268**, 1912–1914

[jj] Mellon, I., Rajpal, D. K., Koi, M., Boland, C. R., and Champe, G. N. (1996) *Science* **272**, 557–560

[kk] German, J., Roe, A. M., Leppert, M. F., and Ellis, N. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6669–6673

[ll] Lebel, M., Spillare, E. A., Harris, C. C., and Leder, P. (1999) *J. Biol. Chem.* **274**, 37795–37799

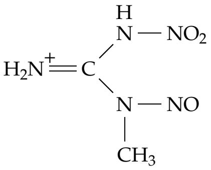
[mm] Walpita, D., Plug, A. W., Neff, N. F., German, J., and Ashley, T. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5622–5627

[nn] Yamagata, K., Kato, J.-I., Shimamoto, A., Goto, M., Furuichi, Y., and Ikeda, H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8733–8738

[oo] Lee, S.-K., Johnson, R. E., Yu, S.-L., Prakash, L., and Prakash, S. (1999) *Science* **286**, 2339–2342

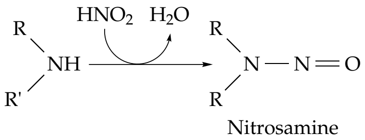
[pp] Dong, Z., Zhong, Q., and Chen, P.-L. (1999) *J. Biol. Chem.* **274**, 19513–19516

highly mutagenic.[772,773] Much used in the laboratory as a mutagen is ***N*-methyl-*N′*-nitro-*N*-nitrosoguanidine**:



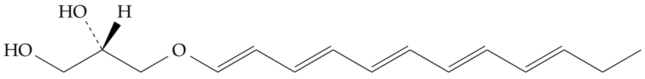
N-Ethyl-*N*-nitrosourea is one of the most potent carcinogens known.[773,774] These compounds, as well as diazomethane and other related substances, probably act via a common intermediate: $CH_3N_2^+$ or $C_2H_5N_2^+$.[775]

Any secondary amine will react with nitrous acid to form a nitrosamine (Eq. 27-25). Tertiary amines can also react with loss of one alkyl group. This can occur in the stomach, and the nitrosamines may be absorbed into the system. All plants contain some nitrate and some, such as spinach and beets, have large amounts. Bacon and other cured R′ R meats contain both nitrites and nitrates, and many drugs and natural food constituents are secondary amines. Cigarette smoke also contains a nitrosamine.[776] There is a possibility that these substances may induce human cancer.[772,777]



(27-25)

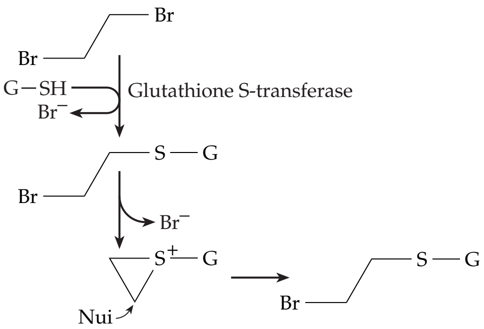
About 3% of residents of North America excrete feces that contain mutagenic unsaturated glyceryl ethers, which have been named **fecapentaenes**.[778,779]



A fecapentaene

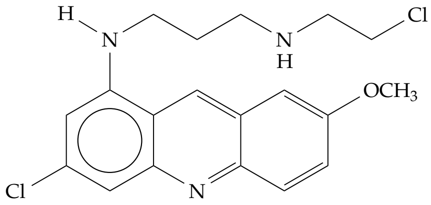
They could be among the causative agents of colorectal cancer. Protonation of a double bond will produce a reactive carbocation, which may be an active alkylating agent.[779]

Many halogenated compounds are carcinogenic. Among these is **1,2-dibromoethane,** which has been produced in the United States in quantities as great as 10^8 kg / year. It has been widely used as a fumigant for foods, as an industrial solvent, and as an additive to gasoline. It is a **procarcinogen**, whose carcinogenic properties are expressed only after **metabolic activation**. One pathway of activation is reaction with glutathione to form a thioether, which can cyclize to a sulfonium compound.[780] The latter alkylates nucleophilic groups in DNA as in Eq. 27-26. Other bifunctional electrophiles such as acrolein, malondialdehyde, vinyl chloride, and urethane are procarcinogens.[781–784] Styrene, another procarcinogen, is converted to the carcinogenic styrene oxide by action of a cytochrome P450.[785] Metabolites of glucose such as **pyruvaldehyde**, **methylglyoxal**, and other reactive aldehydes can also attack DNA.[786,787]



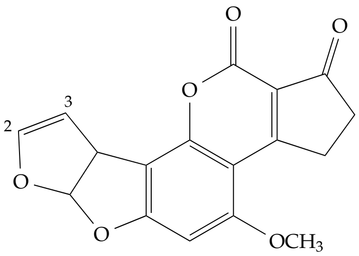
(27-26)

Less common than base pair exchanges are frame-shift mutations. A characteristic of such mutations is that they do not revert (back mutate) as readily as do base substitution mutations, and reversion is not induced by chemicals known to cause base substitutions. However, reversion of frame-shift mutations is induced by acridines and other flat molecules that are known to act as intercalating agents in DNA helices and which promote frame-shift mutations. They are especially effective in causing mutations of regions in which long repeated sequences of a single base such as $(A)_n$ or $(G)_n$ occur.[770,788] For example, deletion of two base pairs from a "hot spot" (site of frequent mutation) in the *Salmonella* histidine[789] operon with the following sequence is induced by 2-nitrosofluorene and causes reversion of a (–1) histidine-requiring mutant: 5'-CGCGCGCG. Whereas simple intercalating agents are often not very mutagenic, compounds that are both an intercalating agent and an alkylating agent are especially potent. An example is the following compound, which contains a half-mustard side chain:



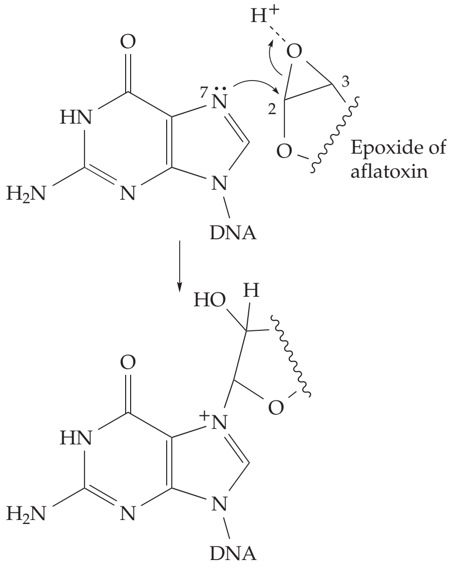
When the CH_2Cl group of the side chain is replaced by CH_2OH, the compound is 100 times less mutagenic.

The carcinogenic **aflatoxins**, which are produced by *Aspergillus flavus*, may be present in infected peanuts and other foodstuffs.[790] Like many other compounds that are carcinogenic or mutagenic, the aflatoxins are not unusually reactive chemically.



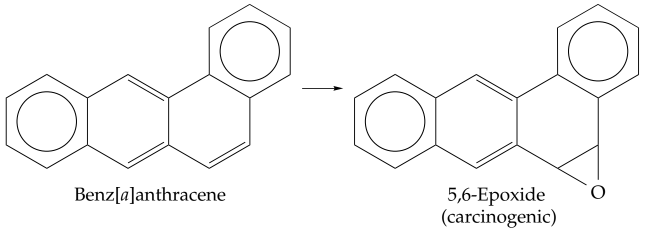
Aflatoxin B_1

However, they are activated in the animal body by hydroxylation and formation of 2,3-epoxides. The latter may react with N-7 of guanyl residues in DNA (Eq. 27-27).[791,792]

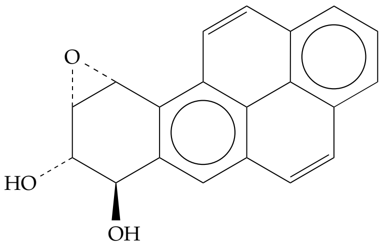


(27-27)

Benz(*a*)anthracene is also converted in the body to a carcinogenic epoxide (Eq. 27-28).[793] Benzo(*a*)pyrene was isolated from coal tar in 1929 and in 1930 provided the first demonstration of the carcinogenicity of a pure chemical compound.[794] It can also be activated by conversion in the ER to the 7,8-dihydrodiol 9,10-



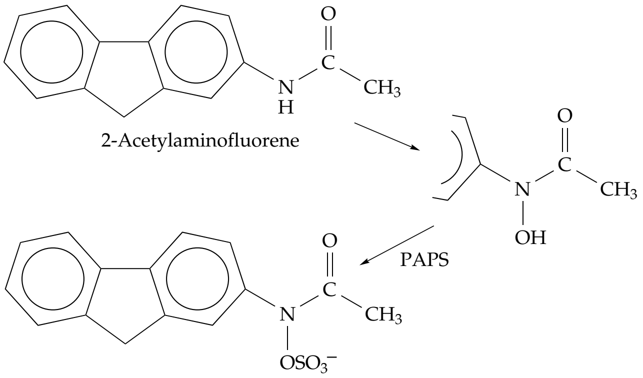
(27-28)



Benzo(*a*)pyrene diol epoxide

epoxide.[795–797] Hydroxylation activates 2-acetylaminofluorene to a carcinogenic *N*-sulfate (Eq. 27-29),[788,798] which reacts with guanine rings.[788,799]

Although it seems incongruous, many of the most potent antitumor drugs are powerful mutagens.



(27-29)

Among these are intercalating agents such as daunomycin (Figs. 5-22, 5-23), neocarzinostatin, and bleomycin (Box 5-B). These are alkylating reagents[800] or attack DNA in other ways. The fact that such compounds are in use for chemotherapy emphasizes the need for new approaches to cancer treatment.

How can compounds be recognized as mutagenic? This is an important question that is difficult to answer and which has led to many controversies. For example, how dangerous is formaldehyde in the environment?[801] Butadiene?[802] Dioxin?[790,803] Is fluoride a carcinogen?[804,805]

A quick test for mutagenic activity makes use of **tester strains** of bacteria developed by Ames and associates. These are *Salmonella* mutants that are unable to synthesize their own histidine, but which can grow when a mutagenic agent produces a back mutation.[806,807] One of the strains can be mutated by agents causing base exchanges, while the other three, which contain different types of frame-shift mutations, are affected differently by various mutagens. About 10^9 bacteria are spread on a Petri plate, and a small amount of the mutagenic chemical is introduced in the center of the plate. Where back mutation has occurred, a colony of the bacteria appears. The strains all carry a mutation in the main DNA excision-repair system so that most mutations are not repaired, and the test is very sensitive. Addition of a liver homogenate plus an NADPH-generating system to the sample tested allows activation of many aromatic chemicals by hydroxylation.[808] Feeding of mutagens to *Drosophila* eye color mutants permits testing for back mutation in a eukaryote.[809]

The bacterial tests have been widely used and have been of great value. For example, they revealed that certain chemicals that were being used as flame retardants in children's clothing are mutagens,[810] and that mutagens can be generated during cooking of meat and other foods.[811] However, there are good reasons for using other methods of identifying mutagens together with the bacterial tests.[812] Government regulatory agencies in the United States have relied largely on tests with rodents. Compounds are tested at very high doses with these short-lived animals. Extrapolation to human exposures at very low levels has been criticized.[764,765,813,814] However, rodent tests have identified many true carcinogens.[815] A broader range of tests are now in use.[816] Direct monitoring of the accumulation of defects in animal and human cells is now possible. For example, in the **^{32}P-postlabeling technique** DNA is enzymatically digested to the nucleotide level, and the adducts with mutagens are labeled with ^{32}P, separated, and their quantities measured.[817–819] Laser-excited fluorescence from such adducts, GC / mass spectrometry, and immunological methods can also be used to identify DNA adducts.[819] Careful vigilance is needed to keep our mutation rate at as low a level as possible.

References

1. Wasserman, S. A., and Cozzarelli, N. R. (1986) *Science* **232**, 951–960

2. Reich, Z., Wachtel, E. J., and Minsky, A. (1994) *Science* **264**, 1460–1463

2a. Daban, J.-R. (2000) *Biochemistry* **39**, 3862–3866

3. Moody, M. F. (1999) *J. Mol. Biol.* **293**, 401–433

4. Black, L. W., Newcomb, W. W., Boring, J. W., and Brown, J. C. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7960–7964

4a. Hud, N. V., and Downing, K. H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14925–14930

5. Sun, M., and Serwer, P. (1997) *Biophys. J.* **72**, 958–963

6. Marx, K. A., and Reynolds, T. C. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6484–6488

7. Widom, J., and Baldwin, R. L. (1983) *J. Mol. Biol.* **171**, 419–437

8. Kamashev, D., Balandina, A., and Rouviere-Yaniv, J. (1999) *EMBO J.* **18**, 5434–5444

9. Kobryn, K., Lavoie, B. D., and Chaconas, G. (1999) *J. Mol. Biol.* **289**, 777–784

10. Schmid, M. B. (1988) *Trends Biochem. Sci.* **13**, 131–135

11. Claret, L., and Rouviere-Yaniv, J. (1996) *J. Mol. Biol.* **263**, 126–139

12. Esser, D., Rudolph, R., Jaenicke, R., and Böhm, G. (1999) *J. Mol. Biol.* **291**, 1135–1146

13. Tanaka, I., Appelt, K., Dijk, J., White, S. W., and Wilson, K. S. (1984) *Nature (London)* **310**, 376–381

14. White, S. W., Wilson, K. S., Appelt, K., and Tanaka, I. (1999) *Acta Crystallogr. D* **55**, 801–809

14a. Grove, A., and Saavedra, T. C. (2002) *Biochemistry* **41**, 7597–7603

15. Ueguchi, C., Suzuki, T., Yoshida, T., Tanaka, K.-i, and Mizuno, T. (1996) *J. Mol. Biol.* **263**, 149–162

16. Starich, M. R., Sandman, K., Reeve, J. N., and Summers, M. F. (1996) *J. Mol. Biol.* **255**, 187–203

17. Krueger, J. K., McCrary, B. S., Wang, A. H.-J., Shriver, J. W., Trewhella, J., and Edmondson, S. P. (1999) *Biochemistry* **38**, 10247–10255

17a. Kamashev, D., and Rouviere-Yaniv, J. (2000) *EMBO J.* **19**, 6527–6535

17b. Grove, A., and Lim, L. (2001) *J. Mol. Biol.* **311**, 491–502

18. Pettijohn, D. E. (1988) *J. Biol. Chem.* **263**, 12793–12796

18a. Renzoni, D., Esposito, D., Pfuhl, M., Hinton, J. C. D., Higgins, C. F., Driscoll, P. C., and Ladbury, J. E. (2001) *J. Mol. Biol.* **306**, 1127–1137

18b. Schröder, O., and Wagner, R. (2000) *J. Mol. Biol.* **298**, 737–748

19. Jackson, D. A., and Patel, S. B. (1982) *Trends Biochem. Sci.* **7**, 272–274

19a. Korolev, N., Lyubartsev, A. P., Nordenskiöld, L., and Laaksonen, A. (2001) *J. Mol. Biol.* **308**, 907–917

20. Hunt, J. G., Kasinsky, H. E., Elsey, R. M., Wright, C. L., Rice, P., Bell, J. E., Sharp, D. J., Kiss, A. J., Hunt, D. F., Arnott, D. P., Russ, M. M., Shabanowitz, J., and Ausió, J. (1996) *J. Biol. Chem.* **271**, 23547–23557

21. Harborne, J. B. (1993) *Introduction to Ecological Biochemistry*, 4th ed., Academic Press, San Diego, California

21a. Prieto, C., Saperas, N., Arnan, C., Hills, M. H., Wang, X., Chiva, M., Aligué, R., Subirana, J. A., and Ausió, J. (2002) *Biochemistry* **41**, 7802–7810

22. Bellvé, A. R., McKay, D. J., Renaux, B. S., and Dixon, G. H. (1988) *Biochemistry* **27**, 2890–2897

23. de Yebra, L., Ballescà, J. L., Vanrell, J. A., Bassas, L., and Oliva, R. (1993) *J. Biol. Chem.* **268**, 10553–10557

24. Doenecke, D., and Karlson, P. (1984) *Trends Biochem. Sci.* **9**, 404–405

25. Kornberg, R. D., and Thomas, J. O. (1974) *Science* **184**, 865–868

26. Isenberg, I. (1979) *Ann. Rev. Biochem.* **48**, 159–191

27. Arents, G., and Moudrianakis, E. N. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11170–11174

28. Zemzoumi, K., Frontini, M., Bellorini, M., and Mantovani, R. (1999) *J. Mol. Biol.* **286**, 327–337

29. Wellman, S. E., Song, Y., and Mamoon, N. M. (1999) *Biochemistry* **38**, 13112–13118

30. Lindner, H., Sarg, B., Hoertnagl, B., and Helliger, W. (1998) *J. Biol. Chem.* **273**, 13324–13330

31. Zlatanova, J., and Doenecke, D. (1994) *FASEB J.* **8**, 1260–1268

32. Hansen, J. C., Tse, C., and Wolffe, A. P. (1998) *Biochemistry* **37**, 17637–17641

33. DeLange, R. J., Hooper, J. A., and Smith, E. L. (1973) *J. Biol. Chem.* **248**, 3261–3274

33a. Wu, J., and Grunstein, M. (2000) *Trends Biochem. Sci.* **25**, 619–623

33b. Jenuwein, T., and Allis, C. D. (2001) *Science* **293**, 1074–1080

33c. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z.-W., Schmid, M., Opravil, S., Mechtier, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000) *Nature (London)* **406**, 593–599

33d. Jacobs, S. A., and Khorasanizadeh, S. (2002) *Science* **295**, 2080–2083

34. Kelner, D. N., and McCarty, K. S., Sr. (1984) *J. Biol. Chem.* **259**, 3413–3418

35. Sternglanz, R. (1996) *Trends Biochem. Sci.* **21**, 357–358

36. Strenglanz, R., and Schindelin, H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8807–8808

37. Champagne, N., Bertos, N. R., Pelletier, N., Wang, A. H., Vezmar, M., Yang, Y., Heng, H. H., and Yang, X.-J. (1999) *J. Biol. Chem.* **274**, 28528–28536

38. Wade, P. A., Pruss, D., and Wolffe, A. P. (1997) *Trends Biochem. Sci.* **22**, 128–132

38a. Marmorstein, R. (2001) *J. Mol. Biol.* **311**, 433–444

38b. Kruhlak, M. J., Hendzel, M. J., Fischle, W., Bertos, N. R., Hameed, S., Yang, X.-J., Verdin, E., and Bazett-Jones, D. P. (2001) *J. Biol. Chem.* **276**, 38307–38319

39. Hay, C. W., and Candido, E. P. M. (1983) *J. Biol. Chem.* **258**, 3726–3734

40. Kölle, D., Brosch, G., Lechner, T., Pipal, A., Helliger, W., Taplick, J., and Loidl, P. (1999) *Biochemistry* **38**, 6769–6773

40a. Huang, X., and Kadonaga, J. T. (2001) *J. Biol. Chem.* **276**, 12497–12500

40b. Furumai, R., Komatsu, Y., Nishino, N., Khochbin, S., Yoshida, M., and Horinouchi, S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 87–92

41. Osley, M. A. (1991) *Ann. Rev. Biochem.* **60**, 827–861

42. Gorovsky, M. A. (1986) in *Molecular Biology of Ciliated Protozoas* (Gall, J. G., ed), Academic Press, Orlando, Florida

43. Prentice, D. A., Loechel, S. C., and Kitos, P. A. (1982) *Biochemistry* **21**, 2412–2420

44. Swank, R. A., Th'ng, J. P. H., Guo, X.-W., Valdez, J., Bradbury, E. M., and Gurley, L. R. (1997) *Biochemistry* **36**, 13761–13768

45. Allera, C., Lazzarini, G., and Patrone, E. (1995) *Biochemistry* **34**, 301–311

45a. de la Barre, A.-E., Angelov, D., Molla, A., and Dimitrov, S. (2001) *EMBO J.* **20**, 6383–6393

46. Koiv, A., Palvimo, J., and Kinnunen, P. K. J. (1995) *Biochemistry* **34**, 8018–8027

46a. Ausió, J., and Abbott, D. W. (2002) *Biochemistry* **41**, 5945–5949

47. Chen, H. Y., Sun, J.-M., Zhang, Y., Davie, J. R., and Meistrich, M. L. (1998) *J. Biol. Chem.* **273**, 13165–13169

47a. Soares, D. J., Sandman, K., and Reeve, J. N. (2000) *J. Mol. Biol.* **297**, 39–47

47b. Bell, S. D., Botting, C. H., Wardleworth, B. N., Jackson, S. P., and White, M. F. (2002) *Science* **296**, 148–151

48. Kornberg, R. D., and Klug, A. (1981) *Sci. Am.* **244**(Feb), 52–64

49. Grunstein, M. (1992) *Sci. Am.* **267**(Oct), 68–74B

50. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) *Nature (London)* **389**, 251–260

51. Griffith, J. D. (1975) *Science* **187**, 1202–1203

52. Krude, T., and Knippers, R. (1994) *J. Biol. Chem.* **269**, 21021–21029

53. Givens, R. M., Saavedra, R. A., and Huberman, J. A. (1996) *J. Mol. Biol.* **257**, 53–65

54. Sen, D., Mitra, S., and Crothers, D. M. (1986) *Biochemistry* **25**, 3441–3447

55. Felsenfeld, G., and McGhee, J. D. (1986) *Cell* **44**, 375–377

56. Wasserman, S. A., White, J. H., and Cozzarelli, N. R. (1988) *Nature (London)* **334**, 448–450

57. McMurray, C. T., van Holde, K. E., Jones, R. L., and Wilson, W. D. (1985) *Biochemistry* **24**, 7037–7044

58. Fitzgerald, D. J., and Anderson, J. N. (1999) *J. Mol. Biol.* **293**, 477–491

59. Lowary, P. T., and Widom, J. (1998) *J. Mol. Biol.* **276**, 19–42

59a. Thåström, A., Lowary, P. T., Widlund, H. R., Cao, H., Kubista, M., and Widom, J. (1999) *J. Mol. Biol.* **288**, 213–229

60. Flaus, A., and Richmond, T. J. (1998) *J. Mol. Biol.* **275**, 427–441

61. Widlund, H. R., Cao, H., Simonsson, S., Magnusson, E., Simonsson, T., Nielsen, P. E., Kahn, J. D., Crothers, D. M., and Kubista, M. (1997) *J. Mol. Biol.* **267**, 807–817

61a. Negri, R., Buttinelli, M., Panetta, G., De Arcangelis, V., Di Mauro, E., and Travers, A. (2001) *J. Mol. Biol.* **307**, 987–999

62. Leuba, S. H., Zlatanova, J., and van Holde, K. (1993) *J. Mol. Biol.* **229**, 917–929

63. Crane-Robinson, C. (1997) *Trends Biochem. Sci.* **22**, 75–77

64. Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J., and Woodcock, C. L. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14173–14178

64a. Vila, R., Ponte, I., Jiménez, M. A., Rico, M., and Suau, P. (2000) *Protein Sci.* **9**, 627–636

64b. Bharath, M. M. S., Ramesh, S., Chandra, N. R., and Rao, M. R. S. (2002) *Biochemistry* **41**, 7617–7627

65. Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E. N., and Wolffe, A. P. (1996) *Science* **274**, 614–617

66. Zhou, Y.-B., Gerchman, S. E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998) *Nature (London)* **395**, 402–405

67. Hayes, J. J. (1996) *Biochemistry* **35**, 11931–11937

68. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., and Sweet, R. M. (1993) *Nature (London)* **362**, 219–223

69. DeLucia, F., Alilat, M., Sivolob, A., and Prunell, A. (1999) *J. Mol. Biol.* **285**, 1101–1119

70. Pennisi, E. (1997) *Science* **275**, 155–157

71. Polach, K. J., and Widom, J. (1995) *J. Mol. Biol.* **254**, 130–149

72. Strahl, B. D., and Allis, C. D. (2000) *Nature (London)* **403**, 41–45

73. Cairns, B. R. (1998) *Trends Biochem. Sci.* **23**, 20–24

74. Kellenberger, E. (1987) *Trends Biochem. Sci.* **12**, 105–107

References

75. Gross, D. S., and Garrard, W. T. (1988) *Ann. Rev. Biochem.* **57**, 159–197

76. Felsenfeld, G., Boyes, J., Chung, J., Clark, D., and Studitsky, V. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9384–9388

77. Wu, C. (1997) *J. Biol. Chem.* **272**, 28171–28174

78. Grunstein, M. (1997) *Nature (London)* **389**, 349–352

79. Edmondson, D. G., and Roth, S. Y. (1996) *FASEB J.* **10**, 1173–1182

80. Nightingale, K. P., Wellinger, R. E., Sogo, J. M., and Becker, P. B. (1998) *EMBO J.* **17**, 2865–2876

81. Zhang, W., Bone, J. R., Edmondson, D. G., Turner, B. M., and Roth, S. Y. (1998) *EMBO J.* **17**, 3155–3167

82. Utley, R. T., Ikeda, K., Grant, P. A., Côté, J., Steger, D. J., Eberharter, A., John, S., and Workman, J. L. (1998) *Nature (London)* **394**, 498–502

83. Mueller, R. D., Yasuda, H., and Bradbury, E. M. (1985) *J. Biol. Chem.* **260**, 5081–5086

84. Muchardt, C., and Yaniv, M. (1999) *J. Mol. Biol.* **293**, 187–198

84a. Horn, P. J., and Peterson, C. L. (2002) *Science* **297**, 1824–1827

85. Thoma, F., Koller, T. H., and Klug, A. (1979) *J. Cell Biol.* **83**, 403–427

86. Walker, P. R., and Sikorska, M. (1987) *J. Biol. Chem.* **262**, 12218–12222

87. Woodcock, C. L. F., Frado, L.-L. Y., and Rattner, J. B. (1984) *J. Cell Biol.* **99**, 42–52

88. Rydberg, B., Holley, W. R., Mian, I. S., and Chatterjee, A. (1998) *J. Mol. Biol.* **284**, 71–84

89. Daban, J.-R., and Bermúdez, A. (1998) *Biochemistry* **37**, 4299–4304

90. Sirlin, J. L. (1972) *Biology of RNA*, Academic Press, New York (pp. 162–164)

91. Manuelidis, L. (1990) *Science* **250**, 1533–1540

91a. Uhlmann, F. (2002) *Nature (London)* **417**, 135–136

91b. Antes, T. J., Namciu, S. J., Fournier, R. E. K., and Levy-Wilson, B. (2001) *Biochemistry* **40**, 6731–6742

91c. Löwe, J., Cordell, S. C., and van den Ent, F. (2001) *J. Mol. Biol.* **306**, 25–35

92. Münkel, C., Eils, R., Dietzel, S., Zink, D., Mehring, C., Wedemann, G., Cremer, T., and Langowski, J. (1999) *J. Mol. Biol.* **285**, 1053–1065

93. van Holde, K., and Zlatanova, J. (1995) *J. Biol. Chem.* **270**, 8373–8376

94. Nagele, R., Freeman, T., McMorrow, L., and Lee, H.-y. (1995) *Science* **270**, 1831–1835

95. Lamond, A. I., and Earnshaw, W. C. (1998) *Science* **280**, 547–553

96. Wagner, R. P., Maguire, M. M., and Stallings, R. L. (1993) *Chromosomes: A Synthesis*, Wiley-Liss, New York

97. Cook, K. R., and Karpen, G. H. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5219–5221

98. Motta, M. C., Landsberger, N., Merli, C., and Badaracco, G. (1998) *J. Biol. Chem.* **273**, 18028–18039

98a. Ringrose, L., and Paro, R. (2001) *Nature (London)* **412**, 493–494

99. Wei, X., Samarabandu, J., Devdhar, R. S., Siegel, A. J., Acharya, R., and Berezney, R. (1998) *Science* **281**, 1502–1505

100. Baskin, Y. (1995) *Science* **268**, 1564–1565

101. Lemon, K. P., and Grossman, A. D. (1998) *Science* **282**, 1516–1519

102. Hirano, T. (1995) *Trends Biochem. Sci.* **20**, 357–361

102a. Misteli, T. (2001) *Science* **291**, 843–847

103. Tsutsui, K., Tsutsui, K., and Muller, M. T. (1988) *J. Biol. Chem.* **263**, 7235–7241

104. Nelson, W. G., Pienta, K. J., Barrack, E. R., and Coffey, D. S. (1986) *Ann. Rev. Biophys. Biophys. Chem.* **15**, 457–475

105. Pemov, A., Bavykin, S., and Hamlin, J. L. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14757–14762

106. Benham, C., Kohwi-Shigematsu, T., and Bode, J. (1997) *J. Mol. Biol.* **274**, 181–196

107. Pederson, T. (1998) *J. Mol. Biol.* **277**, 147–159

108. Mirkovitch, J., Spierer, P., and Laemmli, U. K. (1986) *J. Mol. Biol.* **190**, 255–258

109. Kay, V., and Bode, J. (1994) *Biochemistry* **33**, 367–374

110. Udvardy, A. (1999) *EMBO J.* **18**, 1–8

111. Harborth, J., Wang, J., Gueth-Hallonet, C., Weber, K., and Osborn, M. (1999) *EMBO J.* **18**, 1689–1700

112. Peterson, J. L., and McConkey, E. H. (1976) *J. Biol. Chem.* **251**, 548–554

113. Hnilica, L. S., ed. (1983) *Chromosomal Nonhistone Proteins*, Vol. III, CRC Press, Boca Raton, Florida

114. Quick, D. P., Orchard, P. J., and Duerre, J. A. (1981) *Biochemistry* **20**, 4724–4729

115. Landsman, D., McBride, O. W., Soares, N., Crippa, M. P., Srikantha, T., and Bustin, M. (1989) *J. Biol. Chem.* **264**, 3421–3427

115a. Müller, S., Scaffidi, P., Degryse, B., Bonaldi, T., Ronfani, L., Agresti, A., Beltrame, M., and Bianchi, M. E. (2001) *EMBO J.* **20**, 4337–4340

115b. Thomas, J. O., and Travers, A. A. (2001) *Trends Biochem. Sci.* **26**, 167–174

115c. Bustin, M. (2001) *Trends Biochem. Sci.* **26**, 431–437

115d. Banks, G. C., Li, Y., and Reeves, R. (2000) *Biochemistry* **39**, 8333–8346

115e. Webb, M., Payet, D., Lee, K.-B., Travers, A. A., and Thomas, J. O. (2001) *J. Mol. Biol.* **309**, 79–88

116. Dingwall, C., and Laskey, R. (1992) *Science* **258**, 942–947

117. Ashery-Padan, R., Weiss, A. M., Feinstein, N., and Gruenbaum, Y. (1997) *J. Biol. Chem.* **272**, 2493–2499

118. Davis, L. I. (1995) *Ann. Rev. Biochem.* **64**, 865–896

119. Görlich, D. (1998) *EMBO J.* **17**, 2721–2727

120. Goldberg, M. W., and Allen, T. D. (1996) *J. Mol. Biol.* **257**, 848–865

120a. Raska, I., Aebi, U., and Earnshaw, W. C. (2000) *EMBO J.* **19**, 3843–3848

120b. Shahin, V., Danker, T., Enss, K., Ossig, R., and Oberleithner, H. (2001) *FASEB J.* **15**, 1895–1901

121. Aebi, U., Cohn, J., Buhle, L., and Gerace, L. (1986) *Nature (London)* **323**, 560–564

122. Worman, H. J., Lazaridis, I., and Georgatos, S. D. (1988) *J. Biol. Chem.* **263**, 12135–12141

123. Gerace, L. (1986) *Trends Biochem. Sci.* **11**, 443–446

123a. Georgatos, S. D. (2001) *EMBO J.* **20**, 2989–2994

124. McKeon, F. D., Kirschner, M. W., and Caput, D. (1986) *Nature (London)* **319**, 463–468

125. Schneider, U., Mini, T., Jenö, P., Fisher, P. A., and Stuurman, N. (1999) *Biochemistry* **38**, 4620–4632

126. Favreau, C., Worman, H. J., Wozniak, R. W., Frappier, T., and Courvalin, J.-C. (1996) *Biochemistry* **35**, 8035–8044

126a. Rout, M. P., and Aitchison, J. D. (2001) *J. Biol. Chem.* **276**, 16593–16596

126b. Allen, N. P. C., Huang, L., Burlingame, A., and Rexach, M. (2001) *J. Biol. Chem.* **276**, 29268–29274

127. Weis, K. (1998) *Trends Biochem. Sci.* **23**, 185–189

128. Cingolani, G., Petosa, C., Weis, K., and Müller, C. W. (1999) *Nature (London)* **399**, 221–228

129. Chook, Y. M., and Blobel, G. (1999) *Nature (London)* **399**, 230–237

130. Panté, N., and Aebi, U. (1996) *Science* **273**, 1729–1732

131. Moore, M. S. (1998) *J. Biol. Chem.* **273**, 22857–22860

132. Vetter, I. R., Nowak, C., Nishimoto, T., Kuhlmann, J., and Wittinghofer, A. (1999) *Nature (London)* **398**, 39–46

132a. Geyer, M., Assheuer, R., Klebe, C., Kuhlmann, J., Becker, J., Wittinghofer, A., and Kalbitzer, H. R. (1999) *Biochemistry* **38**, 11250–11260

132b. Goldberg, M. W., Rutherford, S. A., Hughes, M., Cotter, L. A., Bagley, S., Kiseleva, E., Allen, T. D., and Clarke, P. R. (2000) *J. Mol. Biol.* **300**, 519–529

133. Akey, C. W. (1995) *J. Mol. Biol.* **248**, 273–293

134. Kiseleva, E., Goldberg, M. W., Daneholt, B., and Allen, T. D. (1996) *J. Mol. Biol.* **260**, 304–311

135. Perez-Terzic, C., Pyle, J., Jaconi, M., Stehno-Bittel, L., and Clapham, D. E. (1996) *Science* **273**, 1875–1877

136. Stoffler, D., Goldie, K. N., Feja, B., and Aebi, U. (1999) *J. Mol. Biol.* **287**, 741–752

137. Cook, P. R. (1999) *Science* **284**, 1790–1795

138. Chang, J.-H., Dumbar, T. S., and Olson, M. O. J. (1988) *J. Biol. Chem.* **263**, 12824–12827

139. Serin, G., Joseph, G., Ghisolfi, L., Bauzan, M., Erard, M., Amalric, F., and Bouvet, P. (1997) *J. Biol. Chem.* **272**, 13109–13116

139a. Wang, I.-F., Reddy, N. M., and Shen, C.-K. J. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13583–13588

140. Somerville, C., and Somerville, S. (1999) *Science* **285**, 380–383

141. C. elegans Sequencing Consortium. (1998) *Science* **282**, 2012–2018

142. Long, E. O., and Dawid, I. B. (1980) *Ann. Rev. Biochem.* **49**, 727–764

143. Jelinek, W. R., and Schmid, C. W. (1982) *Ann. Rev. Biochem.* **51**, 813–844

144. Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M., and Rodier, F. (1985) *Science* **228**, 953–956

145. Nowak, R. (1994) *Science* **263**, 608–610

145a. Razin, S. V., Ioudinkova, E. S., Trifonov, E. N., and Scherrer, K. (2001) *J. Mol. Biol.* **307**, 481–486

145b. Mashkova, T. D., Oparina, N. Y., Lacroix, M.-H., Fedorova, L. I., Tumeneva, I. G., Zinovieva, O. L., and Kisselev, L. L. (2001) *J. Mol. Biol.* **305**, 33–48

146. Charlesworth, B., Sniegowski, P., and Stephan, W. (1994) *Nature (London)* **371**, 215–220

147. Barakat, A., Matassi, G., and Bernardi, G. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10044–10049

148. Rosenberg, H., Singer, M., and Rosenberg, M. (1978) *Science* **200**, 394–402

149. Inoue, S., Kitajima, K., and Inoue, Y. (1996) *J. Biol. Chem.* **271**, 24341–24344

150. Pluta, A. F., Mackay, A. M., Ainsztein, A. M., Goldberg, I. G., and Earnshaw, W. C. (1995) *Science* **270**, 1591–1594

151. Xiao, Z.-X., and Fitzgerald-Hayes, M. (1995) *J. Mol. Biol.* **248**, 255–263

152. Tal, M., Shimron, F., and Yagil, G. (1994) *J. Mol. Biol.* **243**, 179–189

153. Ferrer, N., Azorín, F., Villasante, A., Gutiérrez, C., and Abad, J. P. (1995) *J. Mol. Biol.* **245**, 8–21

154. Catasti, P., Gupta, G., Garcia, A. E., Ratliff, R., Hong, L., Yau, P., Moyzis, R. K., and Bradbury, E. M. (1994) *Biochemistry* **33**, 3819–3830

155. Chou, S.-H., Zhu, L., and Reid, B. R. (1996) *J. Mol. Biol.* **259**, 445–457

156. Copenhaver, G. P., Nickel, K., Kuromori, T., Benito, M.-I., Kaul, S., Lin, X., Beven, M., Murphy, G., Harris, B., Parnell, L. D., McCombie, W. R., Martienssen, R. A., Marra, M., and Preuss, D. (1999) *Science* **286**, 2468–2474

References

157. Jiang, J., Nasuda, S., Dong, F., Scherrer, C. W., Woo, S.-S., Wing, R. A., Gill, B. S., and Ward, D. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14210–14213
157a. Schueler, M. G., Higgins, A. W., Rudd, M. K., Gustashaw, K., and Willard, H. F. (2001) *Science* **294**, 109–115
157b. Pennisi, E. (2001) *Science* **294**, 30–31
158. Zhu, L., Chou, S.-H., and Reid, B. R. (1995) *J. Mol. Biol.* **254**, 623–637
158a. Gao, Y.-G., Robinson, H., Sanishvili, R., Joachimiak, A., and Wang, A. H.-J. (1999) *Biochemistry* **38**, 16452–16460
158b. Henikoff, S., and Malik, H. S. (2002) *Nature (London)* **417**, 227
159. Pluta, A. F., Cooke, C. A., and Earnshaw, W. C. (1990) *Trends Biochem. Sci.* **15**, 181–185
160. Sugimoto, K., Hagishita, Y., and Himeno, M. (1994) *J. Biol. Chem.* **269**, 24271–24276
161. Pluta, A. F., and Earnshaw, W. C. (1996) *J. Biol. Chem.* **271**, 18767–18774
162. Fukagawa, T., Pendon, C., Morris, J., and Brown, W. (1999) *EMBO J.* **18**, 4196–4209
162a. Takahashi, K., Chen, E. S., and Yanagida, M. (2000) *Science* **288**, 2215–2219
163. Blackburn, E. H., and Szostak, J. W. (1984) *Ann. Rev. Biochem.* **53**, 163–194
164. Moyzis, R. K. (1991) *Sci. Am.* **265**(Aug), 48–55
165. Blackburn, E. H. (1991) *Trends Biochem. Sci.* **16**, 378–381
166. Zakian, V. A. (1995) *Science* **270**, 1601–1607
167. Vega-Palas, M. A., Venditti, S., and Di Mauro, E. (1998) *J. Biol. Chem.* **273**, 9388–9392
167a. de Lange, T. (2001) *Science* **292**, 1075–1076
167b. Stansel, R. M., de Lange, T., and Griffith, J. D. (2001) *EMBO J.* **20**, 5532–5540
167c. Tomaska, L., Makhov, A. M., Nosek, J., Kucejova, B., and Griffith, J. D. (2001) *J. Mol. Biol.* **305**, 61–69
167d. Mitton-Fry, R. M., Anderson, E. M., Hughes, T. R., Lundblad, V., and Wuttke, D. S. (2002) *Science* **296**, 145–147
168. Schierer, T., and Henderson, E. (1994) *Biochemistry* **33**, 2240–2246
169. Frantz, J. D., and Gilbert, W. (1995) *J. Biol. Chem.* **270**, 20692–20697
169a. Baumann, P., and Cech, T. R. (2001) *Science* **292**, 1171–1175
170. Shore, D. (1998) *Science* **281**, 1818–1819
171. König, P., and Rhodes, D. (1997) *Trends Biochem. Sci.* **22**, 43–47
171a. Rossetti, L., Cacchione, S., De Menna, A., Chapman, L., Rhodes, D., and Savino, M. (2001) *J. Mol. Biol.* **306**, 903–913
172. Ohshima, K., and Okada, N. (1994) *J. Mol. Biol.* **243**, 25–37
172a. Kim, J., Martignetti, J. A., Shen, M. R., Brosius, J., and Deininger, P. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3607–3611
173. Ünsal, K., and Morgan, G. T. (1995) *J. Mol. Biol.* **248**, 812–823
174. Izsvák, Z., Ivics, Z., Garcia-Estefania, D., Fahrenkrug, S. C., and Hackett, P. B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1077–1081
175. Ullu, E. (1982) *Trends Biochem. Sci.* **7**, 216–219
176. Sharp, P. A. (1983) *Nature (London)* **301**, 471–472
177. Adams, D. S., Eickbash, T. H., Herrea, R. S., and Lizandi, P. M. (1986) *J. Mol. Biol.* **187**, 465–478
177a. Carroll, M. L., Roy-Engel, A. M., Nguyen, S. V., Salem, A.-H., Vogel, E., Vincent, B., Myers, J., Ahmad, Z., Nuguyen, L., Sammarco, M., Watkins, W. S., Henke, J., Makalowski, W., Jorde, L. B., Dininger, P. L., and Batzer, M. A. (2001) *J. Mol. Biol.* **311**, 17–40
177b. Martinez, J., Dugaiczyk, L. J., Zielinski, R., and Dugaiczyk, A. (2001) *J. Mol. Biol.* **308**, 587–596
178. Shimizu, Y., Yoshida, K., Ren, C.-S., Fujinaga, K., Rejagopalan, S., and Chinnadurai, G. (1983) *Nature (London)* **302**, 587–590
179. Kiyama, R., Matsui, H., and Oishi, M. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4665–4669
180. Fowler, R. F., and Skinner, D. M. (1985) *J. Biol. Chem.* **260**, 1296–1303
181. Nargang, F. E., Bell, J. B., Stohl, L. L., and Lambowitz, A. M. (1983) *J. Biol. Chem.* **258**, 4257–4260
182. Wildeman, A. G., Rasquinha, I., and Nazar, R. N. (1986) *J. Biol. Chem.* **261**, 13401–13403
183. Reis, R. J. S., Lumpkin, C. K., Jr., McGill, J. R., Riabowol, K. T., and Goldstein, S. (1983) *Nature (London)* **301**, 394–398
184. Casavant, N. C., and Hardies, S. C. (1994) *J. Mol. Biol.* **241**, 390–397
185. Dhellin, O., Maestre, J., and Heidmann, T. (1997) *EMBO J.* **16**, 6590–6602
186. Dunham, I., Shimizu, N., Roe, B. A., Chossoe, S., and 117 other authors. (1999) *Nature (London)* **402**, 489–495
187. Sapienza, C., and Doolittle, W. F. (1982) *Nature (London)* **295**, 384–389
188. Wellauer, P. K., Reeder, R. H., Carroll, D., Brown, D. D., Deutch, A., Higashinakagawa, T., and Dawid, I. B. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 2823–2827
189. Carroll, D., and Brown, D. D. (1976) *Cell* **7**, 467–475
190. Mroczka, D. L., Cassidy, B., Busch, H., and Rothblum, L. E. (1984) *J. Mol. Biol.* **174**, 141–162
191. Sommerville, J. (1985) *Nature (London)* **318**, 410–411
192. Budarf, M. L., and Blackburn, E. H. (1986) *J. Biol. Chem.* **261**, 363–369
193. Karabin, G. D., and Hallick, R. B. (1983) *J. Biol. Chem.* **258**, 5512–5518
194. Mattaj, I. W. (1984) *Trends Biochem. Sci.* **9**, 435–437
195. Dandekar, T., Snel, B., Huynen, M., and Bork, P. (1998) *Trends Biochem. Sci.* **23**, 324–328
196. Karlsson, S., and Nienhuis, A. W. (1985) *Ann. Rev. Biochem.* **54**, 1071–1108
197. Proudfoot, N. (1986) *Nature (London)* **321**, 730–731
197a. Wang, Z., and Liebhaber, S. A. (1999) *EMBO J.* **18**, 2218–2228
197b. Li, J., Noguchi, C. T., Miller, W., Hardison, R., and Schechter, A. N. (1998) *J. Biol. Chem.* **273**, 10202–10209
197c. Filipe, A., Li, Q., Deveaux, S., Godin, I., Roméo, P.-H., Stamatoyannopoulos, G., and Mignotte, V. (1999) *EMBO J.* **18**, 687–697
197d. Ristaldi, M. S., Drabek, D., Gribnau, J., Poddie, D., Yannoutsous, N., Cao, A., Grosveld, F., and Imam, A. M. A. (2001) *EMBO J.* **20**, 5242–5249
198. Vanin, E. F., Goldberg, G. I., Tucker, P. W., and Smithies, O. (1980) *Nature (London)* **286**, 222–226
199. Strehler, E. E., Strehler-Page, M.-A., Prriard, J.-C., Periasamy, M., and Nadal-Ginard, B. (1986) *J. Mol. Biol.* **190**, 291–317
200. Kolata, G. B. (1977) *Science* **196**, 1187–1188
201. Purohie, S., and Mathews, C. K. (1984) *J. Biol. Chem.* **259**, 6261–6266
202. Rak, B., Lusky, M., and Hable, M. (1982) *Nature (London)* **297**, 124–128
203. Ernst, H., and Shatkin, A. J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 48–52
204. Soininen, R., Huotari, M., Hostikka, S. L., Prockop, D. J., and Tryggvason, K. (1988) *J. Biol. Chem.* **263**, 17217–17220
205. Henikoff, S., Keene, M. A., Fechtel, K., and Fristrom, J. W. (1986) *Cell* **44**, 33–42
206. Borst, P., and Grivell, L. A. (1981) *Nature (London)* **289**, 439–440
207. Gilson, P. R., and McFadden, G. I. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7737–7742
208. Gray, M. W., Burger, G., and Lang, B. F. (1999) *Science* **283**, 1476–1481
209. Chen, K. K., and Donelson, J. E. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2445–2449
210. de la Cruz, V. F., Lake, J. A., Simpson, A. M., and Simpson, L. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1401–1405
211. Jeltsch, A., Friedrich, T., and Roth, M. (1998) *J. Mol. Biol.* **275**, 747–758
212. Newman, A. K., Rubin, R. A., Kim, S.-H., and Modrich, P. (1981) *J. Biol. Chem.* **256**, 2131–2139
213. Surby, M. A., and Reich, N. O. (1996) *Biochemistry* **35**, 2201–2208
213a. Gowher, H., and Jeltsch, A. (2000) *J. Mol. Biol.* **303**, 93–110
214. Schumann, J., Walter, J., Willert, J., Wild, C., Koch, D., and Trautner, T. A. (1996) *J. Mol. Biol.* **257**, 949–959
215. Ahmad, I., and Rao, D. N. (1996) *J. Mol. Biol.* **259**, 229–240
215a. Pues, H., Bleimling, N., Holz, B., Wölcke, J., and Weinhold, E. (1999) *Biochemistry* **38**, 1426–1434
216. Erlanson, D. A., Chen, L., and Verdine, G. L. (1993) *J. Am. Chem. Soc.* **115**, 12583–12584
217. Lau, E. Y., and Bruice, T. C. (1999) *J. Mol. Biol.* **293**, 9–18
218. Powell, L. M., Connolly, B. A., and Dryden, D. T. F. (1998) *J. Mol. Biol.* **283**, 947–961
219. Herman, G. E., and Modrich, P. (1982) *J. Biol. Chem.* **257**, 2605–2612
220. Zweiger, G., Marczynski, G., and Shapiro, L. (1994) *J. Mol. Biol.* **235**, 472–485
221. Heithoff, D. M., Sinsheimer, R. L., Low, D. A., and Mahan, M. J. (1999) *Science* **284**, 967–970
222. Burdon, R. H., and Adams, R. L. P. (1980) *Trends Biochem. Sci.* **5**, 294–297
223. Doerfler, W. (1983) *Ann. Rev. Biochem.* **52**, 93–124
224. Antequera, F., Tamame, M., Villaneuva, J. R., and Santos, T. (1984) *J. Biol. Chem.* **259**, 8033–8036
225. Selker, E. U., Jensen, B. C., and Richardson, G. A. (1987) *Science* **238**, 48–53
226. Holliday, R. (1989) *Sci. Am.* **260**(Jun), 60–73
227. Hsu, D.-W., Lin, M.-J., Lee, T.-L., Wen, S.-C., Chen, X., and Shen, C.-K., J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9751–9756
228. Yoder, J. A., Soman, N. S., Verdine, G. L., and Bestor, T. H. (1997) *J. Mol. Biol.* **270**, 385–395
229. Pradhan, S., Bacolla, A., Wells, R. D., and Roberts, R. J. (1999) *J. Biol. Chem.* **274**, 33002–33010
229a. Gowher, H., and Jeltsch, A. (2001) *J. Mol. Biol.* **309**, 1201 – 1208
229b. Saito, Y., Kanai, Y., Sakamoto, M., Saito, H., Ishii, H., and Hirohashi, S. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10060 – 10065
230. Gjerset, R. A., and Martin, D. W., Jr. (1982) *J. Biol. Chem.* **257**, 8581–8583
231. Ramchandani, S., Bhattacharya, S. K., Cervoni, N., and Szyf, M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6107–6112
232. Cervoni, N., Bhattacharya, S., and Szyf, M. (1999) *J. Biol. Chem.* **274**, 8363–8366
233. Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giovannini, N., and Rungger, D. (1998) *EMBO J.* **17**, 1446–1453
233a. Smith, S. S. (2000) *J. Mol. Biol.* **302**, 1–7
233b. Bhattacharya, S. K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999) *Nature (London)* **397**, 579 – 583
233c. Cedar, H., and Verdine, G. L. (1999) *Nature (London)* **397**, 568 – 569
234. Bird, A. P. (1986) *Nature (London)* **321**, 209–213

235. Shiraishi, M., Lerman, L. S., and Sekiya, T. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4229–4233

236. Pieper, R. O., Patel, S., Ting, S. A., Futscher, B. W., and Costello, J. F. (1996) *J. Biol. Chem.* **271**, 13916–13924

237. Rein, T., Kobayashi, T., Malott, M., Leffak, M., and DePamphilis, M. L. (1999) *J. Biol. Chem.* **274**, 25792–25800

238. Mayer-Jung, C., Moras, D., and Timsit, Y. (1998) *EMBO J.* **17**, 2709–2718

239. Wakefield, R. I. D., Smith, B. O., Nan, X., Free, A., Soteriou, A., Uhrin, D., Bird, A. P., and Barlow, P. N. (1999) *J. Mol. Biol.* **291**, 1055–1065

240. Davey, C., Pennings, S., and Allan, J. (1997) *J. Mol. Biol.* **267**, 276–288

241. Bester, T. H. (1998) *Nature (London)* **393**, 311–312

242. Razin, A. (1998) *EMBO J.* **17**, 4905–4908

243. Naruse, Y., Aoki, T., Kojima, T., and Mori, N. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13691–13696

244. O'Neill, L. P., Keohane, A. M., Lavender, J. S., McCabe, V., Heard, E., Avner, P., Brockdorff, N., and Turner, B. M. (1999) *EMBO J.* **18**, 2897–2907

245. Herzing, L. B. K., Romer, J. T., Horn, J. M., and Ashworth, A. (1997) *Nature (London)* **386**, 272–275

246. Jegalian, K., and Page, D. C. (1998) *Nature (London)* **394**, 776–780

247. Gommers-Ampt, J. H., and Borst, P. (1995) *FASEB J.* **9**, 1034–1042

248. Wilson, V. L., Smith, R. A., Ma, S., and Cutler, R. G. (1987) *J. Biol. Chem.* **262**, 9948–9951

249. Bakin, A. V., and Curran, T. (1999) *Science* **283**, 387–390

250. Song, H. K., Sohn, S. H., and Suh, S. W. (1999) *EMBO J.* **18**, 1104–1113

251. Sapienza, C. (1990) *Sci. Am.* **263**(Oct), 52–60

252. Barlow, D. P. (1993) *Science* **260**, 309–310

253. Birger, Y., Shemer, R., Perk, J., and Razin, A. (1999) *Nature (London)* **397**, 84–88

254. Meselson, M., and Stahl, F. W. (1958) *Proc. Natl. Acad. Sci. U.S.A.* **44**, 671–682

255. Taylor, J. H. (1997) *Trends Biochem. Sci.* **22**, 447–450

256. Cairns, J. (1963) *J. Mol. Biol.* **6**, 208–213

257. Zimm, B. H. (1999) *Trends Biochem. Sci.* **24**, 121–123

258. Kornberg, A. (1969) *Science* **163**, 1410–1418

259. Lehman, I. R., Bessman, M. J., Simms, E. S., and Kornberg, A. (1958) *J. Biol. Chem.* **233**, 163–170

259a. Patel, P. H., Suzuki, M., Adman, E., Shinkai, A., and Loeb, L. A. (2001) *J. Mol. Biol.* **308**, 823–837

260. Sugimoto, K., Okazaki, T., and Okazaki, R. (1968) *Proc. Natl. Acad. Sci. U.S.A.* **60**, 1356–1362

261. Modrich, P., Anraku, Y., and Lehman, I. R. (1973) *J. Biol. Chem.* **248**, 7495–7501

262. Lehman, I. R. (1974) *Science* **186**, 790–797

263. Brutlag, D., Schekman, R., and Kornberg, A. (1971) *Proc. Natl. Acad. Sci. U.S.A.* **68**, 2826–2829

264. Cai, H., Yu, H., McEntee, K., Kunkel, T. A., and Goodman, M. F. (1995) *J. Biol. Chem.* **270**, 15327–15335

265. Kornberg, A., and Baker, T. A. (1992) *DNA Replication*, 2nd ed., Freeman, New York

266. Hubscher, U., and Spadari, S., eds. (1984) *Proteins Involved in DNA Replication*, Plenum, New York

266a. Friedberg, E. C., Feaver, W. J., and Gerlach, V. L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5681–5683

266b. Burgers, P. M. J., Koonin, E. V., Bruford, E., Blanco, L., Burtis, K. C., Christman, M. F., Copeland, W. C., Friedberg, E. C., and 13 other authors. (2001) *J. Biol. Chem.* **276**, 43487–43490

267. Steitz, T. A. (1999) *J. Biol. Chem.* **274**, 17395–17398

268. Sousa, R. (1996) *Trends Biochem. Sci.* **21**, 186–190

268a. Yang, G., Franklin, M., Li, J., Lin, T.-C., and Konigsberg, W. (2002) *Biochemistry* **41**, 10256–10261

268b. Hübscher, U., Maga, G., and Spadari, S. (2002) *Ann. Rev. Biochem.* **71**, 133–163

269. Brody, R. S., and Frey, P. A. (1981) *Biochemistry* **20**, 1245–1252

270. Steitz, T. A. (1998) *Nature (London)* **391**, 231–232

271. Astatke, M., Ng, K., Grindley, N. D. F., and Joyce, C. M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3402–3407

272. Singh, K., and Modak, M. J. (1998) *Trends Biochem. Sci.* **23**, 277–281

272a. Krishna, T. S. R., Kong, X-P, Gary, S., Burgers, P. M., and Kuriyan, J. (1994) *Cell* **79**, 1233–1243

273. Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science* **264**, 1930–1935

274. Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. (1997) *Biochemistry* **36**, 11205–11215

275. Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) *Science* **260**, 352–355

275a. Spratt, T. E. (2001) *Biochemistry* **40**, 2647–2652

276. Eom, S. H., Wang, J., and Steitz, T. A. (1996) *Nature (London)* **382**, 278–281

277. Li, Y., Mitaxov, V., and Waksman, G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9491–9496

278. Tabor, S., and Richardson, C. C. (1990) *J. Biol. Chem.* **265**, 8322–8328

279. Sousa, R., Chung, Y. J., Rose, J. P., and Wang, B.-C. (1993) *Nature (London)* **364**, 593–599

280. Osumi-Davis, P. A., Sreerama, N., Volkin, D. B., Middaugh, C. R., Woody, R. W., and Woody, A.-Y. M. (1994) *J. Mol. Biol.* **237**, 5–19

280a. Longley, M. J., Ropp, P. A., Lim, S. E., and Copeland, W. C. (1998) *Biochemistry* **37**, 10529–10539

280b. Fan, L., and Kaguni, L. S. (2001) *Biochemistry* **40**, 4780–4791

281. Kelman, Z., Pietrokovski, S., and Hurwitz, J. (1999) *J. Biol. Chem.* **274**, 28751–28761

282. Hopfner, K.-P., Eichinger, A., Engh, R. A., Laue, F., Ankenbauer, W., Huber, R., and Angerer, B. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3600–3605

282a. Rodriguez, A. C., Park, H.-W., Mao, C., and Beese, L. S. (2000) *J. Mol. Biol.* **299**, 447–462

283. Maga, G., and Hübscher, U. (1996) *Biochemistry* **35**, 5764–5777

284. Bambara, R. A., Murante, R. S., and Henricksen, L. A. (1997) *J. Biol. Chem.* **272**, 4647–4650

285. Mozzherin, D. J., and Fisher, P. A. (1996) *Biochemistry* **35**, 3572–3577

286. Dua, R., Levy, D. L., and Campbell, J. L. (1998) *J. Biol. Chem.* **273**, 30046–30055

286a. MacNeill, S. A., Baldacci, G., Burgers, P. M., and Hübscher, U. (2001) *Trends Biochem. Sci.* **26**, 16–17

286b. Xie, B., Mazloum, N., Liu, L., Rahmeh, A., Li, H., and Lee, M. Y. W. T. (2002) *Biochemistry* **41**, 13133–13142

287. Alley, S. C., Trakselis, M. A., Mayer, M. U., Ishmael, F. T., Jones, A. D., and Benkovic, S. J. (2001) *J. Biol. Chem.* **276**, 39340–39349

287a. Pietroni, P., Young, M. C., Latham, G. J., and von Hippel, P. H. (2001) *J. Mol. Biol.* **309**, 869–891

288. Davies, J. F., II, Hostomska, Z., Hostomsky, Z., Jordan, S. R., and Matthews, D. A. (1991) *Science* **252**, 88–95

289. Wang, J., Smerdon, S. J., Jäger, J., Kohlstaedt, L. A., Rice, P. A., Friedman, J. M., and Steitz, T. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7242–7246

289a. Sarafianos, S. G., Das, K., Tantillo, C., Clark, A. D., Jr., Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2001) *EMBO J.* **20**, 1449–1461

290. Tantillo, C., Ding, J., Jacobo-Molina, A., Nanni, R. G., Boyer, P. L., Hughes, S. H., Pauwels, R., Andries, K., Janssen, P. A. J., and Arnold, E. (1994) *J. Mol. Biol.* **243**, 369–387

291. Carroll, S. S., Olsen, D. B., Bennett, C. D., Gotlib, L., Graham, D. J., Condra, J. H., Stern, A. M., Shafer, J. A., and Kuo, L. C. (1993) *J. Biol. Chem.* **268**, 276–281

292. Mathias, S. L., Scott, A. F., Kazazian, H. H., Jr., Boeke, J. D., and Gabriel, A. (1991) *Science* **254**, 1808–1810

293. Pelletier, H., Sawaya, M. R., Wolfle, W., Wilson, S. H., and Kraut, J. (1996) *Biochemistry* **35**, 12742–12761

293a. Shah, A. M., Conn, D. A., Li, S.-X., Capaldi, A., Jäger, J., and Sweasy, J. B. (2001) *Biochemistry* **40**, 11372–11381

293b. Liu, J., and Tsai, M.-D. (2001) *Biochemistry* **40**, 9014–9022

293c. Eckert, K. A., Mowery, A., and Hile, S. E. (2002) *Biochemistry* **41**, 10490–10498

294. Holm, L., and Sander, C. (1995) *Trends Biochem. Sci.* **20**, 345–347

294a. Song, M.-S., Pham, P. T., Olson, M., Carter, J. R., Franden, M. A., Schaaper, R. M., and McHenry, C. S. (2001) *J. Biol. Chem.* **276**, 35165–35175

295. Kornberg, A. (1988) *J. Biol. Chem.* **263**, 1–4

296. Kelman, Z., and O'Donnell, M. (1995) *Ann. Rev. Biochem.* **64**, 171–200

297. Kim, D. R., and McHenry, C. S. (1996) *J. Biol. Chem.* **271**, 20699–20704

298. Perrino, F. W., Harvey, S., and McNeill, S. M. (1999) *Biochemistry* **38**, 16001–16009

299. Turner, J., Hingorani, M. M., Kelman, Z., and O'Donnell, M. (1999) *EMBO J.* **18**, 771–783

300. Brenowitz, S., Kwack, S., Goodman, M. F., O'Donnell, M., and Echols, H. (1991) *J. Biol. Chem.* **266**, 7888–7892

300a. DeRose, E. F., Li, D., Darden, T., Harvey, S., Perrino, F. W., Schaaper, R. M., and London, R. E. (2002) *Biochemistry* **41**, 94–110

301. LaDuca, R. J., Crute, J. J., McHenry, C. S., and Bambara, R. A. (1986) *J. Biol. Chem.* **261**, 7550–7557

301a. Leu, F. P., and O'Donnell, M. (2001) *J. Biol. Chem.* **276**, 47185–47194

301b. Matsumiya, S., Ishino, Y., and Morikawa, K. (2001) *Protein Sci.* **10**, 17–23

301c. Bertram, J. G., Bloom, L. B., Hingorani, M. M., Beechem, J. M., O'Donnell, M., and Goodman, M. F. (2000) *J. Biol. Chem.* **275**, 28413–28420

302. Hingorani, M. M., Bloom, L. B., Goodman, M. F., and O'Donnell, M. (1999) *EMBO J.* **18**, 5131–5144

303. Kuriyan, J., and O'Donnell, M. (1993) *J. Mol. Biol.* **234**, 915–925

303a. Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J., Jannière, L., and Ehrlich, S. D. (2001) *Science* **294**, 1716–1718

304. Lindahl, T., and Barnes, D. E. (1992) *Ann. Rev. Biochem.* **61**, 251–281

References

305. Husain, I., Tomkinson, A. E., Burkhart, W. A., Moyer, M. B., Ramos, W., Mackey, Z. B., Besterman, J. M., and Chen, J. (1995) *J. Biol. Chem.* **270**, 9683–9690
306. Arabshahi, A., and Frey, P. A. (1999) *J. Biol. Chem.* **274**, 8586–8588
307. Doherty, A. J., and Wigley, D. B. (1999) *J. Mol. Biol.* **285**, 63–71
307a. Doherty, A. J., and Dafforn, T. R. (2000) *J. Mol. Biol.* **296**, 43–56
308. Shamoo, Y., Friedman, A. M., Parsons, M. R., Konigsberg, W. H., and Steitz, T. A. (1995) *Nature (London)* **376**, 362–366
309. Folmer, R. H. A., Nilges, M., Konings, R. N. H., and Hilbers, C. W. (1995) *EMBO J.* **14**, 4132–4142
310. Guan, Y., Zhang, H., and Wang, A. H.-J. (1995) *Protein Sci.* **4**, 187–197
311. Olah, G. A., Gray, D. M., Gray, C. W., Kergil, D. L., Sosnick, T. R., Mark, B. L., Vaughan, M. R., and Trewhella, J. (1995) *J. Mol. Biol.* **249**, 576–594
312. Lohman, T. M., and Ferrari, M. E. (1994) *Ann. Rev. Biochem.* **63**, 527–570
313. Raghunathan, S., Ricard, C. S., Lohman, T. M., and Waksman, G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6652–6657
314. Lohman, T. M., Bujalowski, W., and Overman, L. B. (1988) *Trends Biochem. Sci.* **13**, 250–255
315. Cobianchi, F., SenGupta, D. N., Zmudzka, B. Z., and Wilson, S. H. (1986) *J. Biol. Chem.* **261**, 3536–3543
316. Richter, A., Sapp, M., and Knippers, R. (1986) *Trends Biochem. Sci.* **11**, 283
317. Kelly, T. J., Simancek, P., and Brush, G. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14634–14639
318. Subramanya, H. S., Bird, L. E., Brannigan, J. A., and Wigley, D. B. (1996) *Nature (London)* **384**, 379–383
319. Lohman, T. M., and Bjornson, K. P. (1996) *Ann. Rev. Biochem.* **65**, 169–214
320. Marians, K. J. (1997) *Structure* **5**, 1129–1134
320a. Soultanas, P., and Wigley, D. B. (2001) *Trends Biochem. Sci.* **26**, 47–54
320b. Soultanas, P., Dillingham, M. S., Wiley, P., Webb, M. R., and Wigley, D. B. (2000) *EMBO J.* **19**, 3799–3810
321. Yu, X., Jezewska, M. J., Bujalowski, W., and Egelman, E. H. (1996) *J. Mol. Biol.* **259**, 7–14
321a. Donate, L.-E., Llorca, O., Bárcena, M., Brown, S. E., Dixon, N. E., and Carazo, J.-M. (2000) *J. Mol. Biol.* **303**, 383–393
321b. Kaplan, D. L. (2000) *J. Mol. Biol.* **301**, 285–299
322. Jezewska, M. J., Rajendran, S., Bujalowska, D., and Bujalowski, W. (1998) *J. Biol. Chem.* **273**, 10515–10529
323. Hsieh, J., Moore, K. J. M., and Lohman, T. M. (1999) *J. Mol. Biol.* **288**, 255–274
324. Biswas, E. E., and Biswas, S. B. (1999) *Biochemistry* **38**, 10929–10939
325. Bárcena, M., San Martín, C., Weise, F., Ayora, S., Alonso, J. C., and Carazo, J. M. (1998) *J. Mol. Biol.* **283**, 809–819
325a. Niedenzu, T., Röleke, D., Bains, G., Scherzinger, E., and Saenger, W. (2001) *J. Mol. Biol.* **306**, 479–487
326. Richardson, R. W., and Nossal, N. G. (1989) *J. Biol. Chem.* **264**, 4725–4731
326a. Jones, C. E., Mueser, T. C., and Nossal, N. G. (2000) *J. Biol. Chem.* **275**, 27145–27154
326b. Delagoutte, E., and von Hippel, P. H. (2001) *Biochemistry* **40**, 4459–4477
326c. Harmon, F. G., and Kowalczykowski, S. C. (2001) *J. Biol. Chem.* **276**, 232–243
327. Stasiak, A., Tsaneva, I. R., West, S. C., Yu, X., and Engelman, E. H. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7618–7622
327a. Parsons, C. A., Stasiak, A., Bennet, R. J., and West, S. C. (1995) *Nature (London)* **374**, 375–378
327b. Yu, X., West, S. C., and Egelman, E. H. (1997) *J. Mol. Biol.* **266**, 217–222
328. West, S. C. (1996) *Nature (London)* **384**, 316–317
329. Seo, Y.-S., and Hurwitz, J. (1993) *J. Biol. Chem.* **268**, 10282–10295
330. Poll, E. H. A., Harrison, J., Umthun, A., Dobbs, D. L., and Benbow, R. M. (1994) *Biochemistry* **33**, 3841–3847
331. Karow, J. K., Chakraverty, R. K., and Hickson, I. D. (1997) *J. Biol. Chem.* **272**, 30611–30614
331a. Wang, W., Seki, M., Narita, Y., Sonoda, E., Takeda, S., Yamada, K., Masuko, T., Katada, T., and Enomoto, T. (2000) *EMBO J.* **19**, 3428–3435
331b. Wu, L., and Hickson, I. D. (2001) *Science* **292**, 229–230
332. Braguglia, D., Heun, P., Pasero, P., Duncker, B. P., and Gasser, S. M. (1998) *J. Mol. Biol.* **281**, 631–649
332a. Orren, D. K., Theodore, S., and Machwe, A. (2002) *Biochemistry* **41**, 13483 – 13488
332b. Egelman, E. H. (2001) *Nature (London)* **409**, 573,575
333. de la Cruz, J., Kressler, D., Tollervey, D., and Linder, P. (1998) *EMBO J.* **17**, 1128–1140
334. Gorbalenya, A. E., and Koonin, E. V. (1993) *Current Opinion in Structural Biology* **3**, 419–429
335. Wang, J. C. (1996) *Ann. Rev. Biochem.* **65**, 635–692
336. Roca, J. (1995) *Trends Biochem. Sci.* **20**, 156–160
337. Cheng, C., and Shuman, S. (1999) *Biochemistry* **38**, 16599–16612
338. Chen, S.-J., and Wang, J. C. (1998) *J. Biol. Chem.* **273**, 6050–6056
339. Stewart, L., Redinbo, M. R., Qiu, X., Hol, W. G. J., and Champoux, J. J. (1998) *Science* **279**, 1534–1541
340. Redinbo, M. R., Stewart, L., Champoux, J. J., and Hol, W. G. J. (1999) *J. Mol. Biol.* **292**, 685–696
340a. Redinbo, M. R., Champoux, J. J., and Hol, W. G. J. (2000) *Biochemistry* **39**, 6832–6840
340b. Dekker, N. H., Rybenkov, V. V., Duguet, M., Crisona, N. J., Cozzarelli, N. R., Bensimon, D., and Croquette, V. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12126 – 12131
340c. Champoux, J. J. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11998 – 12000
341. Lima, C. D., Wang, J. C., and Mondragón, A. (1994) *Nature (London)* **367**, 138–146
341a. Changela, A., DiGate, R. J., and Mondragón, A. (2001) *Nature (London)* **411**, 1077–1081
342. Olland, S., and Wang, J. C. (1999) *J. Biol. Chem.* **274**, 21688–21694
343. Smith, C. V., and Maxwell, A. (1998) *Biochemistry* **37**, 9658–9667
344. Williams, N. L., and Maxwell, A. (1999) *Biochemistry* **38**, 13502–13511
345. Nichols, M. D., DeAngelis, K., Keck, J. L., and Berger, J. M. (1999) *EMBO J.* **18**, 6177–6188
346. Huang, W. M., Wei, L. S., and Casjens, S. (1985) *J. Biol. Chem.* **260**, 8973–8977
347. Li, Z., Hiasa, H., Kumar, U., and DiGate, R. J. (1997) *J. Biol. Chem.* **272**, 19582–19587
348. Baird, C. L., Harkins, T. T., Morris, S. K., and Lindsley, J. E. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13685–13690
348a. Williams, N. L., Howells, A. J., and Maxwell, A. (2001) *J. Mol. Biol.* **306**, 969–984
349. Morris, S. K., and Lindsley, J. E. (1999) *J. Biol. Chem.* **274**, 30690–30696
350. Berger, J. M., Gamblin, S. J., Harrison, S. C., and Wang, J. C. (1996) *Nature (London)* **379**, 225–232
351. Berger, J. M., Fass, D., Wang, J. C., and Harrison, S. C. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7876–7881
352. Kikuchi, A., and Asai, K. (1984) *Nature (London)* **309**, 677–681
353. North, G. (1985) *Nature (London)* **316**, 394–395
353a. Barthelmes, H. U., Grue, P., Feineis, S., Straub, T., and Boege, F. (2000) *J. Biol. Chem.* **275**, 38823–38830
354. Froelich-Ammon, S. J., and Osheroff, N. (1995) *J. Biol. Chem.* **270**, 21429–21432
355. Sogo, J. M., Stasiak, A., Martínez-Robles, M. L., Krimer, D. B., Hernández, P., and Schvartzman, J. B. (1999) *J. Mol. Biol.* **286**, 637–643
356. Fleischmann, G., Pflugfelder, G., Steiner, E. K., Javaherian, K., Howard, G. C., Wang, J. C., and Elgin, S. C. R. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6958–6962
357. Brooks, M., and Dumas, L. B. (1989) *J. Biol. Chem.* **264**, 3602–3610
358. Massé, E., and Drolet, M. (1999) *J. Mol. Biol.* **294**, 321–332
359. Bailly, C., Carrasco, C., Hamy, F., Vezin, H., Prudhomme, M., Saleem, A., and Rubin, E. (1999) *Biochemistry* **38**, 8605–8611
360. Shero, J. H., Bordwell, B., Rothfield, N. F., and Earnshaw, W. C. (1986) *Science* **231**, 737–740
361. Hiasa, H., and Marians, K. J. (1999) *J. Biol. Chem.* **274**, 27244–27248
362. Sun, W., and Godson, G. N. (1998) *J. Mol. Biol.* **276**, 689–703
362a. Keck, J. L., Roche, D. D., Lynch, A. S., and Berger, J. M. (2000) *Science* **287**, 2482–2486
363. Jing, D. H., Dong, F., Latham, G. J., and von Hippel, P. H. (1999) *J. Biol. Chem.* **274**, 27287–27298
364. Frick, D. N., and Richardson, C. C. (1999) *J. Biol. Chem.* **274**, 35889–35898
365. Doublié, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) *Nature (London)* **391**, 251–258
366. Hacker, K. J., and Johnson, K. A. (1997) *Biochemistry* **36**, 14080–14087
366a. Picha, K. M., Ahnert, P., and Patel, S. S. (2000) *Biochemistry* **39**, 6401–6409
366b. VanLoock, M. S., Chen, Y.-J., Yu, X., Patel, S. S., and Egelman, E. H. (2001) *J. Mol. Biol.* **311**, 951–956
367. Skarstad, K., Thöny, B., Hwang, D. S., and Kornberg, A. (1993) *J. Biol. Chem.* **268**, 5365–5370
368. Tougu, K., Peng, H., and Marians, K. J. (1994) *J. Biol. Chem.* **269**, 4675–4682
369. Tougu, K., and Marians, K. J. (1996) *J. Biol. Chem.* **271**, 21398–21405
370. Komori, H., Matsunaga, F., Higuchi, Y., Ishiai, M., Wada, C., and Miki, K. (1999) *EMBO J.* **18**, 4597–4607
371. Ng, J. Y., and Marians, K. J. (1996) *J. Biol. Chem.* **271**, 15642–15648
372. Allen, G. C., Jr., and Kornberg, A. (1993) *J. Biol. Chem.* **268**, 19204–19209
373. Hupp, T. R., and Kaguni, J. M. (1993) *J. Biol. Chem.* **268**, 13143–13150
374. Katayanagi, K., Ishikawa, M., Okumura, M., Ariyoshi, M., Kanaya, S., Kawano, Y., Suzuki, M., Tanaka, I., and Morikawa, K. (1993) *J. Biol. Chem.* **268**, 22092–22099
375. Keck, J. L., Goedken, E. R., and Marqusee, S. (1998) *J. Biol. Chem.* **273**, 34128–34133
376. Ohtani, N., Haruki, M., Morikawa, M., Crouch, R. J., Itaya, M., and Kanaya, S. (1999) *Biochemistry* **38**, 605–618
377. Bird, R. E., Louarn, J., Martuscelli, J., and Caro, L. (1972) *J. Mol. Biol.* **70**, 549–566
378. Prescott, D. M., and Kuempel, P. L. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 2842–2845

References

379. Kuempel, P. L. et al. (1972) in *DNA Synthesis in Vitro* (Wells, R., and Inman, R., eds), pp. 463–472, University Park Press, Baltimore, Maryland
380. Kriegstein, H. J., and Hogness, D. S. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 135–139
381. Hixson, J. E., Wong, T. W., and Clayton, D. A. (1986) *J. Biol. Chem.* **261**, 2384–2390
382. Lee, D. Y., and Clayton, D. A. (1998) *J. Biol. Chem.* **273**, 30614–30621
382a. Krüger, R., Konieczny, I., and Filutowicz, M. (2001) *J. Mol. Biol.* **306**, 945–955
383. McMacken, R., Silver, L., and Georgopoulos, C. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 565–612, Am. Soc. Microbiol., Washington,DC
383a. Smelkova, N., and Marians, K. J. (2001) *J. Biol. Chem.* **276**, 39186–39191
384. Krause, M., Rückert, B., Lurz, R., and Messer, W. (1997) *J. Mol. Biol.* **274**, 365–380
385. Lobry, J. R. (1996) *Science* **272**, 745–746
385a. Kelman, Z. (2000) *Trends Biochem. Sci.* **25**, 521–523
386. Yung, B. Y.-M., and Kornberg, A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7202–7205
387. Sutton, M. D., Carr, K. M., Vicente, M., and Kaguni, J. M. (1998) *J. Biol. Chem.* **273**, 34255–34262
388. Sutton, M. D., and Kaguni, J. M. (1997) *J. Mol. Biol.* **274**, 546–561
389. Radzicka, A., and Wolfenden, R. (1988) *Biochemistry* **27**, 1664–1670
390. Zyskind, J. W., Cleary, J. M., Brusilow, W. S. A., Harding, N. E., and Smith, D. W. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1164–1168
391. Chédin, F., Seitz, E. M., and Kowalczykowski, S. C. (1998) *Trends Biochem. Sci.* **23**, 273–277
391a. Bochkareva, E., Belegu, V., Korolev, S., and Bochkarev, A. (2001) *EMBO J.* **20**, 612–618
392. Lother, H., and Messer, W. (1981) *Nature (London)* **294**, 376–378
393. Xia, W., and Dowhan, W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 783–787
394. York, J. D., Saffitz, J. E., and Majerus, P. W. (1994) *J. Biol. Chem.* **269**, 19992–19999
394a. Li, X., and Marians, K. J. (2000) *J. Biol. Chem.* **275**, 34757–34765
395. Maki, H., Maki, S., and Kornberg, A. (1988) *J. Biol. Chem.* **263**, 6570–6578
396. Lasken, R. S., and Kornberg, A. (1988) *J. Biol. Chem.* **263**, 5512–5518
396a. Chapados, B. R., Chai, Q., Hosfield, D. J., Qiu, J., Shen, B., and Tainer, J. A. (2001) *J. Mol. Biol.* **307**, 541–556
397. O'Donnell, M. E., and Kornberg, A. (1985) *J. Biol. Chem.* **260**, 12884–12889
398. Lee, E. H., and Kornberg, A. (1992) *J. Biol. Chem.* **267**, 8778–8784
399. Kamada, K., Horiuchi, T., Ohsumi, K., Shimamoto, N., and Morikawa, K. (1996) *Nature (London)* **383**, 598–603
400. Duggin, I. G., Andersen, P. A., Smith, M. T., Wilce, J. A., King, G. F., and Wake, R. G. (1999) *J. Mol. Biol.* **286**, 1325–1335
401. Manna, A. C., Pai, K. S., Bussiere, D. E., White, S. W., and Bastia, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3253–3258
402. Yamazoe, M., Onogi, T., Sunako, Y., Niki, H., Yamanaka, K., Ichimura, T., and Hiraga, S. (1999) *EMBO J.* **18**, 5873–5884
403. Dodson, M., McMacken, R., and Echols, H. (1989) *J. Biol. Chem.* **264**, 10719–10725
404. Learn, B. A., Um, S.-J., Huang, L., and McMacken, R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1154–1159
405. Wickner, S., and McKenney, K. (1987) *J. Biol. Chem.* **262**, 13163–13167
406. Echols, H. (1990) *J. Biol. Chem.* **265**, 14697–14700
407. Murialdo, H. (1991) *Ann. Rev. Biochem.* **60**, 125–153
408. Wickner, S. H. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2690–2694
409. Wickner, S., Skowyra, D., Hoskins, J., and McKenney, K. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10345–10349
410. Nossal, N. G. (1992) *FASEB J.* **6**, 871–878
411. Brown, D. R., Roth, M. J., Reinberg, D., and Hurwitz, J. (1984) *J. Biol. Chem.* **259**, 10545–10555
412. Higashitani, A., Greenstein, D., Hirokawa, H., Asano, S., and Horiuchi, K. (1994) *J. Mol. Biol.* **237**, 388–400
413. Masai, H., Nomura, N., Kubota, Y., and Arai, K.-i. (1990) *J. Biol. Chem.* **265**, 15124–15133
414. Abarzua, P., Soeller, W., and Marians, K. J. (1984) *J. Biol. Chem.* **259**, 14286–14292
415. Roth, M. J., Brown, D. R., and Hurwitz, J. (1984) *J. Biol. Chem.* **259**, 10556–10568
416. Mok, M., and Marians, K. J. (1987) *J. Biol. Chem.* **262**, 2304–2309
416a. Campos-Olivas, R., Louis, J. M., CLérot, D., Gronenborn, B., and Gronenborn, A. M. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10310–10315
417. Cue, D., and Feiss, M. (1998) *J. Mol. Biol.* **280**, 11–29
418. Yang, Q., Berton, N., Manning, M. C., and Catalano, C. E. (1999) *Biochemistry* **38**, 14238–14247
419. Higgins, R. R., and Becker, A. (1995) *J. Mol. Biol.* **252**, 31–46
420. Yeo, A., and Feiss, M. (1995) *J. Mol. Biol.* **245**, 126–140
420a. Cue, D., and Feiss, M. (2001) *J. Mol. Biol.* **311**, 233–240
421. Franklin, J. L., Haseltine, D., Davenport, L., and Mosig, G. (1998) *J. Mol. Biol.* **277**, 541–557
421a. Rao, V. B., and Mitchell, M. S. (2001) *J. Mol. Biol.* **314**, 401–411
422. Marvin, D. A. (1998) *Current Opinion in Structural Biology* **8**, 150–158
423. Papavoine, C. H. M., Christiaans, B. E. C., Folmer, R. H. A., Konings, R. N. H., and Hilbers, C. W. (1998) *J. Mol. Biol.* **282**, 401–419
424. Rakonjac, J., Feng, J.-n, and Model, P. (1999) *J. Mol. Biol.* **289**, 1253–1265
425. Burch, A. D., Ta, J., and Fane, B. A. (1999) *J. Mol. Biol.* **286**, 95–104
426. Dokland, T., Bernal, R. A., Burch, A., Pletnev, S., Fane, B. A., and Rossmann, M. G. (1999) *J. Mol. Biol.* **288**, 595–608
427. Chai, S., Lurz, R., and Alonso, J. C. (1995) *J. Mol. Biol.* **252**, 386–398
428. Valle, M., Kremer, L., Martínez-A, C., Roncal, F., Valpuesta, J. M., Albar, J. P., and Carrascosa, J. L. (1999) *J. Mol. Biol.* **288**, 899–909
429. Caetano de Sousa, P., Jr., Tuma, R., Prevelige, P. E., Jr., Silva, J. L., and Foguel, D. (1999) *J. Mol. Biol.* **287**, 527–538
430. Grimes, S., and Anderson, D. (1997) *J. Mol. Biol.* **266**, 901–914
430a. Rodríguez-Casado, A., Moore, S. D., Prevelige, P. E., Jr., and Thomas, G. J., Jr. (2001) *Biochemistry* **40**, 13583–13591
430b. Simpson, A. A., Tao, Y., Leiman, P. G., Badasso, M. O., He, Y., Jardine, P. J., Olson, N. H., Morais, M. C., Grimes, S., Anderson, D. L., Baker, T. S., and Rossmann, M. G. (2000) *Nature (London)* **408**, 745–750
430c. Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L., and Bustamante, C. (2001) *Nature (London)* **413**, 748–752
431. Leffers, G., and Rao, V. B. (1996) *J. Mol. Biol.* **258**, 839–850
432. Jardine, P. J., and Coombs, D. H. (1998) *J. Mol. Biol.* **284**, 661–672
432a. Conway, J. F., Wikoff, W. R., Cheng, N., Duda, R. L., Hendrix, R. W., Johnson, J. E., and Steven, A. C. (2001) *Science* **292**, 744–748
433. Burian, J., Stuchlík, S., and Kay, W. W. (1999) *J. Mol. Biol.* **294**, 49–65
434. Merlin, S., and Polisky, B. (1995) *J. Mol. Biol.* **248**, 211–219
435. Paulsson, J., and Ehrenberg, M. (1998) *J. Mol. Biol.* **279**, 73–88
436. Nordström, K., and Wagner, E. G. H. (1994) *Trends Biochem. Sci.* **19**, 294–300
437. Kramer, M. G., Khan, S. A., and Espinosa, M. (1997) *EMBO J.* **16**, 5784–5795
438. Novick, R. P. (1998) *Trends Biochem. Sci.* **23**, 434–438
439. Willetts, N., and Skurray, R. (1987) in *Escherichia coli and Salmonella typhimurium*, Vol. 2 (Neidhardt, F. C., ed), pp. 1110–1133, Am. Soc. for Microbiology, Washington, D. C.
440. Lanka, E., and Wilkins, B. M. (1995) *Ann. Rev. Biochem.* **64**, 141–169
441. Yates, P., Lane, D., and Biek, D. P. (1999) *J. Mol. Biol.* **290**, 627–638
442. White, J. H., and Richardson, C. C. (1987) *J. Biol. Chem.* **262**, 8851–8860
443. Willwand, K., Mumtsidu, E., Kuntz-Simon, G., and Rommelaere, J. (1998) *J. Biol. Chem.* **273**, 1165–1174
443a. Dufour, E., Méndez, J., Lázaro, J. M., de Vega, M., Blanco, L., and Salas, M. (2000) *J. Mol. Biol.* **304**, 289–300
444. King, A. J., and van der Vliet, P. C. (1994) *EMBO J.* **13**, 5786–5792
445. Pombo, A., Ferreira, J., Bridge, E., and Carmo-Fonseca, M. (1994) *EMBO J.* **13**, 5075–5085
446. Monaghan, A., and Hay, R. T. (1996) *J. Biol. Chem.* **271**, 24242–24248
447. Voyles, B. A. (1993) *The Biology of Viruses*, Mosby, St. Louis, Missouri
447a. Parkinson, G. N., Lee, M. P. H., and Neidle, S. (2002) *Nature (London)* **417**, 876–880
447b. Patel, D. J. (2002) *Nature (London)* **417**, 807–808
448. Greider, C. W. (1996) *Ann. Rev. Biochem.* **65**, 337–365
449. Kowald, A. (1997) *J. Mol. Biol.* **273**, 814–825
449a. Aigner, S., Lingner, J., Goodrich, K. J., Grosshans, C. A., Shevchenko, A., Mann, M., and Cech, T. R. (2000) *EMBO J.* **19**, 6230–6239
450. Shippen-Lentz, D., and Blackburn, E. H. (1990) *Science* **247**, 546–552
451. Greider, C. W., and Blackburn, E. H. (1996) *Sci. Am.* **274**(Feb), 92–97
452. Feng, J., Funk, W. D., Wang, S.-S., Weinrich, S. L., Avilion, A. A., Chiu, C.-P., Adams, R. R., Chang, E., Allsopp, R. C., Yu, J., Le, S., West, M. D., Harley, C. B., Andrews, W. H., Greider, C. W., and Villeponteau, B. (1995) *Science* **269**, 1236–1241
453. Bachand, F., and Autexier, C. (1999) *J. Biol. Chem.* **274**, 38027–38031
454. Shay, J. W., and Wright, W. E. (1999) *Science* **286**, 2284–2285
455. Singer, M. S., and Gottschling, D. E. (1994) *Science* **266**, 404–409
456. Lingner, J., Hughes, T. R., Shevchenko, A., Mann, M., Lundblad, V., and Cech, T. R. (1997) *Science* **276**, 561–567
457. Lee, H.-W., Blasco, M. A., Gottlieb, G. J., Horner, J. W., II, Greider, C. W., and DePinho, R. A. (1998) *Nature (London)* **392**, 569–574
458. Lustig, A. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3339–3341
459. McEachern, M. J., and Blackburn, E. H. (1995) *Nature (London)* **376**, 403–409
460. Smith, S., Giriat, I., Schmitt, A., and de Lange, T. (1998) *Science* **282**, 1484–1487
460a. Blackburn, E. H. (2000) *Nature (London)* **408**, 53–56

References

460b. Lundblad, V. (2000) *Science* **288**, 2141–2142
461. King, T. C., and Low, R. L. (1987) *J. Biol. Chem.* **262**, 6204–6213
462. Ntambi, J. M., Shapiro, T. A., Ryan, K. A., and Englund, P. T. (1986) *J. Biol. Chem.* **261**, 11890–11895
463. Palmer, J. D. (1983) *Nature (London)* **301**, 92–93
464. Dey, P. M., and Harborne, J. B., eds. (1997) *Plant Biochemistry*, Academic Press, San Diego, California
465. Klein, A., and Bonhoeffer, F. (1972) *Ann. Rev. Biochem.* **41**, 301–332
466. DePamphilis, M. L., ed. (1996) *DNA Replication in Eukaryotic Cells*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
466a. DePamphilis, M. L., ed. (1999) *Concepts of Eukaryotic DNA Replication*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
467. Waga, S., and Stillman, B. (1998) *Ann. Rev. Biochem.* **67**, 721–751
467a. Bell, S. P., and Dutta, A. (2002) *Ann. Rev. Biochem.* **71**, 333 – 374
468. Campbell, J. L. (1988) *Trends Biochem. Sci.* **13**, 212–217
468a. Wyrick, J. J., Aparicio, J. G., Chen, T., Barnett, J. D., Jennings, E. G., Young, R. A., Bell, S. P., and Aparicio, O. M. (2001) *Science* **294**, 2357–2360
469. Diller, J. D., and Raghuraman, M. K. (1994) *Trends Biochem. Sci.* **19**, 320–325
469a. Raghuraman, M. K., Winzeler, E. A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., Lockhart, D. J., Davis, R. W., Brewer, B. J., and Fangman, W. L. (2001) *Science* **294**, 115–121
469b. Stillman, B. (2001) *Science* **294**, 2301–2304
470. Weiser, T., Gassmann, M., Thömmes, P., Ferrari, E., and Hafkemeyer, P. (1991) *J. Biol. Chem.* **266**, 10420–10428
471. Sugino, A. (1995) *Trends Biochem. Sci.* **20**, 319–323
472. Spadari, S., Sala, F., and Pedrali-Noy, G. (1982) *Trends Biochem. Sci.* **7**, 29–32
473. Copeland, W. C., and Tan, X. (1995) *J. Biol. Chem.* **270**, 3905–3913
474. Arezi, B., Kirk, B. W., Copeland, W. C., and Kuchta, R. D. (1999) *Biochemistry* **38**, 12899–12907
474a. Arezi, B., and Kuchta, R. D. (2000) *Trends Biochem. Sci.* **25**, 572–576
475. Copeland, W. C., and Wang, T. S.-F. (1993) *J. Biol. Chem.* **268**, 11028–11040
476. Kelman, Z., and Hurwitz, J. (1998) *Trends Biochem. Sci.* **23**, 236–238
476a. Shiomi, Y., Usukura, J., Masamura, Y., Takeyasu, K., Nakayama, Y., Obuse, C., Yoshikawa, H., and Tsurimoto, T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14127–14132
477. Bochkarev, A., Pfuetzner, R. A., Edwards, A. M., and Frappier, L. (1997) *Nature (London)* **385**, 176–181
478. Foiani, M., Lucchini, G., and Plevani, P. (1997) *Trends Biochem. Sci.* **22**, 424–427
479. Kelly, T. J. (1988) *J. Biol. Chem.* **263**, 17889–17892
480. Hurwitz, J., Dean, F. B., Kwong, A. D., and Lee, S.-H. (1990) *J. Biol. Chem.* **265**, 18043–18046
481. Waga, S., and Stillman, B. (1994) *Nature (London)* **369**, 207–212
482. Challberg, M. D., and Kelly, T. J. (1989) *Ann. Rev. Biochem.* **58**, 671–717
483. SenGupta, D. J., and Borowiec, J. A. (1992) *Science* **256**, 1656–1661
483a. Weisshart, K., Förster, H., Kremmer, E., Schlott, B., Grosse, F., and Nasheuer, H.-P. (2000) *J. Biol. Chem.* **275**, 17328–17337
484. Lehman, I. R., and Boehmer, P. E. (1999) *J. Biol. Chem.* **274**, 28059–28062
485. Murata, L. B., and Dodson, M. S. (1999) *J. Biol. Chem.* **274**, 37079–37086
486. Toyn, J. H., Toone, W. M., Morgan, B. A., and Johnston, L. H. (1995) *Trends Biochem. Sci.* **20**, 70–73
487. Murray, A. W., and Szostak, J. W. (1987) *Sci. Am.* **257**(Nov), 62–68
488. Farr, C. J., Bayne, R. A. L., Kipling, D., Mills, W., Critcher, R., and Cooke, H. J. (1995) *EMBO J.* **14**, 5444–5454
489. Warburton, P. E., and Kipling, D. (1997) *Nature (London)* **386**, 553–555
490. Shirahige, K., Hori, Y., Shiraishi, K., Yamashita, M., Takahashi, K., Obuse, C., Tsurimoto, T., and Yoshikawa, H. (1998) *Nature (London)* **395**, 618–621
491. Gómez, M., and Antequera, F. (1999) *EMBO J.* **18**, 5683–5690
492. Kim, S.-M., and Huberman, J. A. (1999) *J. Mol. Biol.* **288**, 867–882
493. Gavin, K. A., Hidaka, M., and Stillman, B. (1995) *Science* **270**, 1667–1671
494. Marx, J. (1995) *Science* **270**, 1585–1587
495. Tye, B. K. (1999) *Ann. Rev. Biochem.* **68**, 649–686
495a. Dhar, S. K., Delmolino, L., and Dutta, A. (2001) *J. Biol. Chem.* **276**, 29067–29071
495b. Blow, J. J. (2001) *EMBO J.* **20**, 3293–3297
495c. Krude, T. (2000) *J. Biol. Chem.* **275**, 13699–13707
496. Prokhorova, T. A., and Blow, J. J. (2000) *J. Biol. Chem.* **275**, 2491–2498
497. Coverley, D., and Laskey, R. A. (1994) *Ann. Rev. Biochem.* **63**, 745–776
498. Chong, J. P. J., Thömmes, P., and Blow, J. J. (1996) *Trends Biochem. Sci.* **21**, 102–106
499. Stillman, B. (1994) *J. Biol. Chem.* **269**, 7047–7050
500. Aladjem, M. I., Rodewald, L. W., Kolman, J. L., and Wahl, G. M. (1998) *Science* **281**, 1005–1009
500a. Tye, B. K., and Sawyer, S. (2000) *J. Biol. Chem.* **275**, 34833–34836
500b. Sato, M., Gotow, T., You, Z., Komamura-Kohno, Y., Uchiyama, Y., Yabuta, N., Nojima, H., and Ishimi, Y. (2000) *J. Mol. Biol.* **300**, 421–431
500c. Lee, J.-K., and Hurwitz, J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 54–59
500d. Chong, J. P. J., Hayashi, M. K., Simon, M. N., Xu, R.-M., and Stillman, B. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1530–1535
500e. Wohlschlegel, J. A., Dwyer, B. T., Dhar, S. K., Cvetic, C., Walter, J. C., and Dutta, A. (2000) *Science* **290**, 2309–2312
500f. Lygerou, Z., and Nurse, P. (2000) *Science* **290**, 2271–2273
500g. Blow, J. J., and Tada, S. (2000) *Nature (London)* **404**, 560–561
500h. Keller, C., and Krude, T. (2000) *J. Biol. Chem.* **275**, 35512–35521
501. DePamphilis, M. L. (1993) *J. Biol. Chem.* **268**, 1–4
501a. Lucas, I., Chevrier-Miller, M., Sogo, J. M., and Hyrien, O. (2000) *J. Mol. Biol.* **296**, 769–786
501b. Gilbert, D. M. (2001) *Science* **294**, 96–100
502. Wu, J.-R., and Gilbert, D. M. (1996) *Science* **271**, 1270–1272
503. Dimitrova, D. S., Giacca, M., Demarchi, F., Biamonti, G., Riva, S., and Falaschi, A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1498–1503
504. Kasamatson, H., and Vinograd, J. (1973) *Nature (London)* **241**, 103–105
505. Gaudette, M. F., and Benbow, R. M. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5953–5957
506. Turchi, J. J., Huang, L., Murante, R. S., Kim, Y., and Bambara, R. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9803–9807
506a. Karthikeyan, R., Vonarx, E. J., Straffon, A. F. L., Simon, M., Faye, G., and Kunz, B. A. (2000) *J. Mol. Biol.* **299**, 405–419
507. Huang, L., Rumbaugh, J. A., Murante, R. S., Lin, R. J. R., Rust, L., and Bambara, R. A. (1996) *Biochemistry* **35**, 9266–9277
507a. Gomes, X. V., and Burgers, P. M. J. (2000) *EMBO J.* **19**, 3811–3821
507b. Bae, S.-H., and Seo, Y.-S. (2000) *J. Biol. Chem.* **275**, 38022–38031
508. Pohjanpelto, P., and Hölttä, E. (1996) *EMBO J.* **15**, 1193–1200
509. Bonne-Andrea, C., Wong, M. L., and Alberts, B. M. (1990) *Nature (London)* **343**, 719–726
510. Gasser, R., Koller, T., and Sogo, J. M. (1996) *J. Mol. Biol.* **258**, 224–239
511. Murray, A. W. (1987) *Nature (London)* **327**, 14–15
512. Fisher, D. L., and Nurse, P. (1996) *EMBO J.* **15**, 850–860
513. Santocanale, C., Neecke, H., Longhese, M. P., Lucchini, G., and Plevani, P. (1995) *J. Mol. Biol.* **254**, 595–607
514. Novak, B., and Tyson, J. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9147–9152
515. Brush, G. S., Morrow, D. M., Hieter, P., and Kelly, T. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15075–15080
516. Collins, J. M., and Chu, A. K. (1987) *Biochemistry* **26**, 5600–5607
517. Stahl, F. W. (1987) *Sci. Am.* **256**(Feb), 91–101
518. Kucherlapati, R., and Smith, G. R., eds. (1988) *Genetic Recombination*, Am. Soc. for Microbiology, Washington, DC
519. Kowalczykowski, S. C. (1991) *Ann. Rev. Biophys. Biophys. Chem.* **20**, 539–575
520. Edelmann, W., and Kucherlapati, R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6225–6227
521. Shinohara, A., and Ogawa, T. (1995) *Trends Biochem. Sci.* **20**, 387–391
522. Haber, J. E. (1999) *Nature (London)* **398**, 665–667
523. Meselson, M., and Weigle, J. J. (1961) *Proc. Natl. Acad. Sci. U.S.A.* **47**, 857–868
524. Holliday, R. (1964) *Genet. Res. Camb.* **5**, 282–304
525. Potter, H., and Dressler, D. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3000–3004
525a. Eichman, B. F., Vargason, J. M., Mooers, B. H. M., and Ho, P. S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3971–3976
525b. Sha, R., Liu, F., and Seeman, N. C. (2002) *Biochemistry* **41**, 5950–5955
526. Sigal, N., and Alberts, B. (1972) *J. Mol. Biol.* **71**, 789–793
527. Meselson, M. (1972) *J. Mol. Biol.* **71**, 795–798
528. Shinagawa, H., and Iwasaki, H. (1996) *Trends Biochem. Sci.* **21**, 107–111
529. Rafferty, J. B., Sedelnikova, S. E., Hargreaves, D., Artymiuk, P. J., Baker, P. J., Sharples, G. J., Mahdi, A. A., Lloyd, R. G., and Rice, D. W. (1996) *Science* **274**, 415–421
530. Ishimori, K., Sommer, S., Bailone, A., Takahashi, M., Cox, M. M., and Devoret, R. (1996) *J. Mol. Biol.* **264**, 696–712
530a. McIlwraith, M. J., and West, S. C. (2001) *J. Mol. Biol.* **305**, 23–31
530b. Postow, L., Ullsperger, C., Keller, R. W., Bustamante, C., Vologodskii, A. V., and Cozzarelli, N. R. (2001) *J. Biol. Chem.* **276**, 2790–2796
530c. Neschastnova, A. A., Markina, V. K., Popenko, V. I., Danilova, O. A., Sidorov, R. A., Belitsky, G. A., and Yakubovskaya, M. G. (2002) *Biochemistry* **41**, 7795–7801
530d. McIlwraith, M. J., Van Dyck, E., Masson, J.-Y., Stasiak, A. Z., Stasiak, A., and West, S. C. (2000) *J. Mol. Biol.* **304**, 151–164
530e. Bolt, E. L., Sharples, G. J., and Lloyd, R. G. (2000) *J. Mol. Biol.* **304**, 165–176

References

531. Farah, J. A., and Smith, G. R. (1997) *J. Mol. Biol.* **272**, 699–715
532. Wang, J., Chen, R., and Julin, D. A. (2000) *J. Biol. Chem.* **275**, 507–513
533. Anderson, D. G., and Kowalczykowski, S. C. (1998) *J. Mol. Biol.* **282**, 275–285
533a. Arnold, D. A., Handa, N., Kobayashi, I., and Kowalczykowski, S. C. (2000) *J. Mol. Biol.* **300**, 469–479
534. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277**, 1453–1462
534a. Flores, M.-J., Bierne, H., Ehrlich, S. D., and Michel, B. (2001) *EMBO J.* **20**, 619–629
535. Köppen, A., Krobitsch, S., Thoms, B., and Wackernagel, W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6249–6253
536. Howard-Flanders, P., West, S. C., and Stasiak, A. (1984) *Nature (London)* **309**, 215–220
537. Story, R. M., Weber, I. T., and Steitz, T. A. (1992) *Nature (London)* **355**, 318–325
538. Cui, X., Wise, R. P., and Schnable, P. S. (1996) *Science* **272**, 1334–1336
539. Arenson, T. A., Tsodikov, O. V., and Cox, M. M. (1999) *J. Mol. Biol.* **288**, 391–401
539a. Egelman, E. H. (2001) *J. Mol. Biol.* **309**, 539–542
540. Hegner, M., Smith, S. B., and Bustamante, C. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10109–10114
540a. Lusetti, S. L., and Cox, M. M. (2002) *Ann. Rev. Biochem.* **71**, 71 – 100
541. Register, J. C., III, and Griffith, J. (1985) *J. Biol. Chem.* **260**, 12308–12312
542. Zaitsev, E. N., and Kowalczykowski, S. C. (1999) *J. Mol. Biol.* **287**, 21–31
543. Adzuma, K. (1998) *J. Biol. Chem.* **273**, 31565–31573
544. Rao, B. J., Chiu, S., Bazemore, L. R., Reddy, G., and Radding, C. M. (1995) *Trends Biochem. Sci.* **20**, 109–113
544a. Voloshin, O. N., Wang, L., and Camerini-Otero, R. D. (2000) *J. Mol. Biol.* **303**, 709–720
544b. Rice, K. P., Eggler, A. L., Sung, P., and Cox, M. M. (2001) *J. Biol. Chem.* **276**, 38570–38581
545. Engels, W. R., Preston, C. R., and Johnson-Schlitz, D. M. (1994) *Science* **263**, 1623–1625
546. Kowalczykowski, S. C., and Eggleston, A. K. (1994) *Ann. Rev. Biochem.* **63**, 991–1043
547. Beernink, H. T. H., and Morrical, S. W. (1999) *Trends Biochem. Sci.* **24**, 385–389
548. Birkenkamp, K., and Kemper, B. (1996) *J. Mol. Biol.* **259**, 622–631
549. Ando, R. A., and Morrical, S. W. (1999) *Biochemistry* **38**, 16589–16598
550. Passy, S. I., Yu, X., Li, Z., Radding, C. M., and Egelman, E. H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4279–4284
551. Sung, P. (1994) *Science* **265**, 1241–1243
551a. Masson, J.-Y., and West, S. C. (2001) *Trends Biochem. Sci.* **26**, 131 – 136
552. Benson, F. E., Stasiak, A., and West, S. C. (1994) *EMBO J.* **13**, 5764–5771
553. Baumann, P., and West, S. C. (1998) *Trends Biochem. Sci.* **23**, 247–251
554. Li, Z., Golub, E. I., Gupta, R., and Radding, C. M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11221–11226
555. Cerutti, H., Osman, M., Grandoni, P., and Jagendorf, A. T. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8068–8072
556. Rafferty, J. B., Ingleston, S. M., Hargreaves, D., Artymiuk, P. J., Sharples, G. J., Lloyd, R. G., and Rice, D. W. (1998) *J. Mol. Biol.* **278**, 105–116
557. George, H., Mézard, C., Stasiak, A., and West, S. C. (1999) *J. Mol. Biol.* **293**, 505–519
558. Qiu, X.-B., Lin, Y.-L., Thome, K. C., Pian, P., Schlegel, B. P., Weremowicz, S., Parvin, J. D., and Dutta, A. (1998) *J. Biol. Chem.* **273**, 27786–27793
559. Rafferty, J. B., Sedelnikova, S. E., Hargreaves, D., Artymiuk, P. J., Baker, P. J., Sharples, G. J., Hahdi, A. A., Lloyd, R. G., and Rice, D. W. (1996) *Science* **274**, 415–421
559a. Putnam, C. D., Clancy, S. B., Tsuruta, H., Gonzalez, S., Wetmur, J. G., and Tainer, J. A. (2001) *J. Mol. Biol.* **311**, 297–310
559b. Yamada, K., Kunishima, N., Mayanagi, K., Ohnishi, T., Nishino, T., Iwasaki, H., Shinagawa, H., and Morikawa, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1442–1447
560. Lee, S., Cavallo, L., and Griffith, J. (1997) *J. Biol. Chem.* **272**, 7532–7539
561. Zlatanova, J., and van Holde, K. (1998) *FASEB J.* **12**, 421–431
562. Golz, S., and Kemper, B. (1999) *J. Mol. Biol.* **285**, 1131–1144
563. Raaijmakers, H., Vix, O., Töro, I., Golz, S., Kemper, B., and Suck, D. (1999) *EMBO J.* **18**, 1447–1458
564. Bolt, E. L., Sharples, G. J., and Lloyd, R. G. (1999) *J. Mol. Biol.* **286**, 403–415
565. Whitby, M. C., and Dixon, J. (1998) *J. Biol. Chem.* **273**, 35063–35073
566. Komori, K., Sakae, S., Shinagawa, H., Morikawa, K., and Ishino, Y. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8873–8878
567. Kvaratskhelia, M., George, S. J., Cooper, A., and White, M. F. (1999) *Biochemistry* **38**, 16613–16619
568. Fincham, J. R. S., and Oliver, P. (1989) *Nature (London)* **338**, 14–15
569. Nicolas, A., Treco, D., Schultes, N. P., and Szostak, J. W. (1989) *Nature (London)* **338**, 35–39
570. Fabre, F. (1978) *Nature (London)* **272**, 795–798
571. Groden, J., Nakamura, Y., and German, J. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4315–4319
572. Jackson, J. A., and Fink, G. R. (1981) *Nature (London)* **292**, 306–311
573. Klein, H. L. (1984) *Nature (London)* **310**, 748–753
574. Robertson, M. (1981) *Nature (London)* **293**, 333–334
575. Szostak, J. W., and Wu, R. (1980) *Nature (London)* **284**, 426–430
576. Sadowski, P. D. (1993) *FASEB J.* **7**, 760–767
577. Wang, S., Cosstick, R., Gardner, J. F., and Gumport, R. I. (1995) *Biochemistry* **34**, 13082–13090
577a. Read, E. K., Gumport, R. I., and Gardner, J. F. (2000) *J. Biol. Chem.* **275**, 33759–33764
578. Nash, H. A. (1990) *Trends Biochem. Sci.* **15**, 222–227
579. Tirumalai, R. S., Kwon, H. J., Cardente, E. H., Ellenberger, T., and Landy, A. (1998) *J. Mol. Biol.* **279**, 513–527
580. Crisona, N. J., Weinberg, R. L., Peter, B. J., Sumners, D. W., and Cozzarelli, N. R. (1999) *J. Mol. Biol.* **289**, 747–775
580a. Bankhead, T., and Segall, A. M. (2000) *J. Biol. Chem.* **275**, 36949–36956
580b. Lorbach, E., Christ, N., Schwikardi, M., and Dröge, P. (2000) *J. Mol. Biol.* **296**, 1175–1181
581. Hsu, P. L., and Landy, A. (1984) *Nature (London)* **311**, 721–726
582. Dorgai, L., Sloan, S., and Weisberg, R. A. (1998) *J. Mol. Biol.* **277**, 1059–1070
583. Guo, F., Gopaul, D. N., and Van Duyne, G. D. (1997) *Nature (London)* **389**, 40–46
584. Gopaul, D. N., Guo, F., and Van Duyne, G. D. (1998) *EMBO J.* **17**, 4175–4187
585. Subramanya, H. S., Arciszewska, L. K., Baker, R. A., Bird, L. E., Sherratt, D. J., and Wigley, D. B. (1997) *EMBO J.* **16**, 5178–5187
586. Arciszewska, L. K., Grainge, I., and Sherratt, D. J. (1997) *EMBO J.* **16**, 3731–3743
587. Jayaram, M. (1994) *Trends Biochem. Sci.* **19**, 78–82
587a. Ferreira, H., Sherratt, D., and Arciszewska, L. (2001) *J. Mol. Biol.* **312**, 45–57
588. Zhu, X.-D., Pan, G., Luetke, K., and Sadowski, P. D. (1995) *J. Biol. Chem.* **270**, 11646–11653
589. Huffman, K. E., and Levene, S. D. (1999) *J. Mol. Biol.* **286**, 1–13
590. Lee, J., Jayaram, M., and Grainge, I. (1999) *EMBO J.* **18**, 784–791
591. Allingham, J. S., Pribil, P. A., and Haniford, D. B. (1999) *J. Mol. Biol.* **289**, 1195–1206
592. Watson, M. (1984) *Trends Biochem. Sci.* **9**, 82–83
593. Lim, H. M., Hughes, K. T., and Simon, M. I. (1992) *J. Biol. Chem.* **267**, 11183–11190
594. Feng, J.-A., Johnson, R. C., and Dickerson, R. E. (1994) *Science* **263**, 348–355
595. Kostrewa, D., Granzin, J., Koch, C., Choe, H.-W., Raghunathan, S., Wolf, W., Labahn, J., Kahmann, R., and Saenger, W. (1991) *Nature (London)* **349**, 178–180
596. Safo, M. K., Yang, W.-Z., Corselli, L., Cramton, S. E., Yuan, H. S., and Johnson, R. C. (1997) *EMBO J.* **16**, 6860–6873
597. Dworkin, J., and Blaser, M. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 985–990
598. Fedoroff, N., ed. (1992) *The Dynamic Genome Barbara McClintock's Ideas in the Century of Genetics*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
599. Fedoroff, N. V. (1984) *Sci. Am.* **250**(Jun), 85–98
600. Burr, B., and Burr, F. A. (1983) *Trends Biochem. Sci.* **8**, 429–431
601. Kingsman, A. J., Chater, K. F., and Kingsman, S. M., eds. (1988) *Transposition*, Cambridge Univ. Press, London and New York
602. Cohen, S. N., and Shapiro, J. A. (1980) *Sci. Am.* **243**(Feb), 40–49
603. Watson, M. (1985) *Trends Biochem. Sci.* **10**, 178
604. Ohtsubo, H., Nyman, K., Doroszkiewicz, W., and Ohtsubo, E. (1981) *Nature (London)* **292**, 640–643
604a. Redder, P., She, Q., and Garrett, R. A. (2001) *J. Mol. Biol.* **306**, 1–6
605. Gamas, P., Galas, D., and Chandler, M. (1985) *Nature (London)* **317**, 458–460
606. Kröger, M., and Hobom, G. (1982) *Nature (London)* **297**, 159–162
607. Phadnis, S. H., and Berg, D. E. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 9118–9122
608. Lee, S. Y., Butler, D., and Kleckner, N. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7876–7880
608a. Rao, J. E., and Craig, N. L. (2001) *J. Mol. Biol.* **307**, 1161–1170
609. Farabaugh, P. J., and Fink, G. R. (1980) *Nature (London)* **286**, 352–356
609a. Pribil, P. A., and Haniford, D. B. (2000) *J. Mol. Biol.* **303**, 145–159
610. Sakai, J., Chalmers, R. M., and Kleckner, N. (1995) *EMBO J.* **14**, 4374–4383
611. Bhasin, A., Goryshin, I. Y., and Reznikoff, W. S. (1999) *J. Biol. Chem.* **274**, 37021–37029
611a. Davies, D. R., Goryshin, I. Y., Reznikoff, W. S., and Rayment, I. (2000) *Science* **289**, 77–85
611b. Bhasin, A., Goryshin, I. Y., Steiniger-White, M., York, D., and Reznikoff, W. S. (2000) *J. Mol. Biol.* **302**, 49–63
612. Gary, P. A., Biery, M. C., Bainton, R. J., and Craig, N. L. (1996) *J. Mol. Biol.* **257**, 301–316
612a. Williams, T. L., and Baker, T. A. (2000) *Science* **289**, 73 – 74
613. Benjamin, K. R., Abola, A. P., Kanaar, R., and Cozzarelli, N. R. (1996) *J. Mol. Biol.* **256**, 50–65
614. Arnold, P. H., Blake, D. G., Grindley, N. D. F., Boocock, M. R., and Stark, W. M. (1999) *EMBO J.* **18**, 1407–1414

References

615. Boocock, M. R., Zhu, X., and Grindley, N. D. F. (1995) *EMBO J.* **14**, 5129–5140
616. Watson, M. A., Boocock, M. R., and Stark, W. M. (1996) *J. Mol. Biol.* **257**, 317–329
617. Wasserman, S. A., Dungan, J. M., and Cozzarelli, N. R. (1985) *Science* **229**, 171–174
618. Savilahti, H., Rice, P. A., and Mizuuchi, K. (1995) *EMBO J.* **14**, 4893–4903
619. Krementsova, E., Giffin, M. J., Pincus, D., and Baker, T. A. (1998) *J. Biol. Chem.* **273**, 31358–31365
620. Naigamwalla, D. Z., Coros, C. J., Wu, Z., and Chaconas, G. (1998) *J. Mol. Biol.* **282**, 265–274
621. Jiang, H., Yang, J.-Y., and Harshey, R. M. (1999) *EMBO J.* **18**, 3845–3855
622. Stellwagen, A. E., and Craig, N. L. (1998) *Trends Biochem. Sci.* **23**, 486–490
623. Dranginis, A. M. (1986) *Trends Biochem. Sci.* **11**, 328–331
624. Sugawara, N., Ivanov, E. L., Fishman-Lobell, J., Ray, B. L., Wu, X., and Haber, J. E. (1995) *Nature (London)* **373**, 84–86
625. Flavell, A. (1986) *Nature (London)* **320**, 397
626. Sherratt, D. J., ed. (1995) *Mobile Genetic Elements*, Oxford Univ. Press, New York
627. Gloor, G. B., Nassif, N. A., Johnson-Schlitz, D. M., Preston, C. R., and Engels, W. R. (1991) *Science* **253**, 1110–1117
628. Beall, E. L., and Rio, D. C. (1998) *EMBO J.* **17**, 2122–2136
629. Marx, J. (1991) *Science* **253**, 1092–1093
630. Lohe, A. R., De Aguiar, D., and Hartl, D. L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1293–1297
631. van Pouderoyen, G., Ketting, R. F., Perrakis, A., Plasterk, R. H. A., and Sixma, T. K. (1997) *EMBO J.* **16**, 6044–6054
632. Goodier, J. L., and Davidson, W. S. (1994) *J. Mol. Biol.* **241**, 26–34
633. Lam, W. L., Seo, P., Robison, K., Virk, S., and Gilbert, W. (1996) *J. Mol. Biol.* **257**, 359–366
634. Oosumi, T., and Belknap, W. R. (1995) *Nature (London)* **378**, 672
635. Holzman, D. (1991) *Science* **254**, 1728–1729
636. Lykke-Andersen, J., Garrett, R. A., and Kjems, J. (1997) *EMBO J.* **16**, 3272–3281
637. Morgan, G. T. (1995) *J. Mol. Biol.* **254**, 1–5
638. Jones, J. D. G., Carland, F. M., Maliga, P., and Dooner, H. K. (1989) *Science* **244**, 204–207
639. Spaink, H. P. (1999) *Nature (London)* **402**, 135–136
640. Chopra, S., Brendel, V., Zhang, J., Axtell, J. D., and Peterson, T. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 15330–15335
641. Koes, R., Souer, E., van Houwelingen, A., Mur, L., Spelt, C., Quattrocchio, F., Wing, J., Oppedijk, B., Ahmed, S., Maes, T., Gerats, T., Hoogeveen, P., Meesters, M., Kloos, D., and Mol, J. N. M. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8149–8153
642. Morgan, B. A., Conlon, F. L., Manzanares, M., Millar, J. B. A., Kanuga, N., Sharpe, J., Krumlauf, R., Smith, J. C., and Sedgwick, S. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2801–2806
643. Schardl, C. L., Lonsdale, D. M., Pring, D. R., and Rose, K. R. (1984) *Nature (London)* **310**, 292–296
644. Ikeda, H. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 922–926
645. Bi, X., and Liu, L. F. (1996) *J. Mol. Biol.* **256**, 849–858
646. Stark, G. R., and Wahl, G. M. (1984) *Ann. Rev. Biochem.* **53**, 447–491
647. Schimke, R. T. (1988) *J. Biol. Chem.* **263**, 5989–5992
648. Syu, L.-J., and Fluck, M. M. (1997) *J. Mol. Biol.* **271**, 76–99
649. Wu, C. H. H., and Black, L. W. (1995) *J. Mol. Biol.* **247**, 604–617
650. Andersson, D. I., Slechta, E. S., and Roth, J. R. (1998) *Science* **282**, 1133–1135
651. Huang, T., and Campbell, J. L. (1995) *J. Biol. Chem.* **270**, 9607–9614
652. Sheriff, S., and Herriott, J. R. (1981) *J. Mol. Biol.* **145**, 441–451
653. Subramanian, V., Liu, T.-N., Yeh, W.-K., Narro, M., and Gibson, D. T. (1983) *J. Biol. Chem.* **256**, 2723–2730
654. Alitalo, K. (1985) *Trends Biochem. Sci.* **10**, 194–197
655. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) *DNA Repair and Mutagenesis*, ASM Press, Materials Park, Ohio
656. Kuhnlein, U. (1985) *J. Biol. Chem.* **260**, 14918–14924
657. Beckman, K. B., and Ames, B. N. (1997) *J. Biol. Chem.* **272**, 19633–19636
658. Purmal, A. A., Lampman, G. W., Bond, J. P., Hatahet, Z., and Wallace, S. S. (1998) *J. Biol. Chem.* **273**, 10026–10035
659. Jourdan, M., Garcia, J., Defrancq, E., Kotera, M., and Lhomme, J. (1999) *Biochemistry* **38**, 3985–3995
660. Huff, A. C., and Topal, M. D. (1987) *J. Biol. Chem.* **262**, 12843–12850
661. Rosenblatt, J., and Mitchison, T. J. (1998) *Nature (London)* **393**, 739–740
662. Boorstein, R. J., Hilbert, T. P., Cunningham, R. P., and Teebor, G. W. (1990) *Biochemistry* **29**, 10455–10460
663. Doetsch, P. W., Zastawny, T. H., Martin, A. M., and Dizdaroglu, M. (1995) *Biochemistry* **34**, 737–742
663a. Zhou, H., Suzuki, M., Randers-Pehrson, G., Vannais, D., Chen, G., Trosko, J. E., Waldren, C. A., and Hei, T. K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14410–14415
663b. Forster, L., Forster, P., Lutz-Bonengel, S., Willkomm, H., and Brinkmann, B. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13950 – 13954
664. Lai, M.-D., and Beattie, K. L. (1988) *Biochemistry* **27**, 1722–1728
665. Kunkel, T. A., and Soni, A. (1988) *J. Biol. Chem.* **263**, 14784–14789
666. Echols, H., and Goodman, M. F. (1991) *Ann. Rev. Biochem.* **60**, 477–511
667. Kunkel, T. A. (1992) *J. Biol. Chem.* **267**, 18251–18254
668. Timsit, Y. (1999) *J. Mol. Biol.* **293**, 835–853
668a. Viswanathan, M., Lacirignola, J. J., Hurley, R. L., and Lovett, S. T. (2000) *J. Mol. Biol.* **302**, 553–564
669. Radman, M., and Wagner, R. (1988) *Sci. Am.* **259**(Aug), 40–46
669a. Patel, P. H., Kawate, H., Adman, E., Ashbach, M., and Loeb, L. A. (2001) *J. Biol. Chem.* **276**, 5044–5051
670. Fersht, A. R. (1980) *Trends Biochem. Sci.* **5**, 262–265
671. Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., and Tinoco, I., Jr. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6252–6256
672. Metzler, D. E. (1977) *Biochemistry; The Chemical Reactions of Living Cells*, 1st ed., Academic Press, New York (pp. 912–915)
673. Metzler, D. E. (1979) *Adv. Enzymol.* **50**, 1–40
674. Goodman, M. F. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10493–10495
675. Moran, S., Ren, R. X.-F., and Kool, E. T. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10506–10511
676. Matray, T. J., and Kool, E. T. (1998) *J. Am. Chem. Soc.* **120**, 6191–6192
676a. Kool, E. T. (2002) *Ann. Rev. Biochem.* **71**, 191 – 219
676b. Showalter, A. K., and Tsai, M.-D. (2002) *Biochemistry* **41**, 10571 – 10576
677. Mo, J.-Y., and Schaaper, R. M. (1996) *J. Biol. Chem.* **271**, 18947–18953
678. Fijalkowska, I. J., Jonczyk, P., Tkaczyk, M. M., and Bialoskorska, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10020–10025
679. Kroutil, L. C., Frey, M. W., Kaboord, B. F., Kunkel, T. A., and Benkovic, S. J. (1998) *J. Mol. Biol.* **278**, 135–146
680. Seki, M., Akiyama, M., Sugaya, Y., Ohtsubo, E., and Maki, H. (1999) *J. Biol. Chem.* **274**, 33313–33319
681. Modrich, P. (1987) *Ann. Rev. Biochem.* **56**, 435–466
682. Bohr, V. A., and Wassermann, K. (1988) *Trends Biochem. Sci.* **13**, 429–433
683. Thoma, F. (1999) *EMBO J.* **18**, 6585–6598
684. Friedberg, E. C. (1995) *Trends Biochem. Sci.* **20**, 381
685. Arbel, A., Zenvirth, D., and Simchen, G. (1999) *EMBO J.* **18**, 2648–2658
686. Modrich, P. (1989) *J. Biol. Chem.* **264**, 6597–6600
686a. Wood, R. D., Mitchell, M., Sgouros, J., and Lindahl, T. (2001) *Science* **291**, 1284 – 1289
687. Modrich, P., and Lahue, R. (1996) *Ann. Rev. Biochem.* **65**, 101–133
688. Schaaper, R. M. (1993) *J. Biol. Chem.* **268**, 23762–23765
688a. Obmolova, G., Ban, C., Hsieh, P., and Yang, W. (2000) *Nature (London)* **407**, 703–711
688b. Lamers, M. H., Perrakis, A., Enzlin, J. H., Winterwerp, H. H. K., de Wind, N., and Sixma, T. K. (2000) *Nature (London)* **407**, 711–717
688c. Bowers, J., Tran, P. T., Joshi, A., Liskay, R. M., and Alani, E. (2001) *J. Mol. Biol.* **306**, 957 – 968
689. Ban, C., and Yang, W. (1998) *EMBO J.* **17**, 1526–1534
690. Modich, P. (1997) *J. Biol. Chem.* **272**, 24727–24730
691. Wang, H., Lawrence, C. W., Li, G.-M., and Hays, J. B. (1999) *J. Biol. Chem.* **274**, 16894–16900
692. Nakagawa, T., Datta, A., and Kolodner, R. D. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14186–14188
692a. Coïc, E., Gluck, L., and Fabre, F. (2000) *EMBO J.* **19**, 3408–3417
692b. Rich, T., Allen, R. L., and Wyllie, A. H. (2000) *Nature (London)* **407**, 777–783
693. Ahn, B., and Grossman, L. (1996) *J. Biol. Chem.* **271**, 21462–21470
694. Sancar, A. (1996) *Ann. Rev. Biochem.* **65**, 43–81
695. Zou, Y., and Van Houten, B. (1999) *EMBO J.* **18**, 4889–4901
695a. Verhoeven, E. E. A., Wyman, C., Moolenaar, G. F., Hoeijmakers, J. H. J., and Goosen, N. (2001) *EMBO J.* **20**, 601–611
696. Theis, K., Chen, P. J., Skorvaga, M., Van Houten, B., and Kisker, C. (1999) *EMBO J.* **18**, 6899–6907
696a. Villani, G., and Le Gac, N. T. (2000) *J. Biol. Chem.* **275**, 33185–33188
697. Shi, Q., Thresher, R., Sancar, A., and Griffith, J. (1992) *J. Mol. Biol.* **226**, 425–432
698. Verhoeven, E. E. A., van Kesteren, M., Moolenaar, G. F., Visse, R., and Goosen, N. (2000) *J. Mol. Biol.* **275**, 5120–5123
699. Ali, J. A., Maluf, N. K., and Lohman, T. M. (1999) *J. Mol. Biol.* **293**, 815–834
700. Wood, R. D. (1997) *J. Biol. Chem.* **272**, 23465–23468
701. Sancar, A. (1995) *J. Biol. Chem.* **270**, 15915–15918
702. Doetsch, P. W. (1995) *Trends Biochem. Sci.* **20**, 384–386
703. Guzder, S. N., Bailly, V., Sung, P., Prakash, L., and Prakash, S. (1995) *J. Biol. Chem.* **270**, 8385–8388
703a. Cappelli, E., Degan, P., Thompson, L. H., and Frosina, G. (2000) *Biochemistry* **39**, 10408–10412

References

704. Lao, Y., Gomes, X. V., Ren, Y., Taylor, J.-S., and Wold, M. S. (2000) *Biochemistry* **39**, 850–859
705. Zamble, D. B., Mu, D., Reardon, J. T., Sancar, A., and Lippard, S. J. (1996) *Biochemistry* **35**, 10004–10013
705a. Kunz, C., and Fleck, O. (2001) *J. Mol. Biol.* **313**, 241–253
705b. Van Houten, B., Eisen, J. A., and Hanawalt, P. C. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2581–2583
706. Seeberg, E., Eide, L., and Bjorås, M. (1995) *Trends Biochem. Sci.* **20**, 391–397
707. McCullough, A. K., Dodson, M. L., and Lloyd, R. S. (1999) *Ann. Rev. Biochem.* **68**, 255–285
708. Wilson, D. M., III, and Thompson, L. H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12754–12757
708a. Parker, A., Gu, Y., Mahoney, W., Lee, S.-H., Singh, K. K., and Lu, A.-L. (2001) *J. Biol. Chem.* **276**, 5547–5555
709. Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995) *Nature (London)* **373**, 487–493
710. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) *Nature (London)* **384**, 87–92
711. Drohat, A. C., Xiao, G., Tordova, M., Jagadeesh, J., Pankiewicz, K. W., Watanabe, K. A., Gilliland, G. L., and Stivers, J. T. (1999) *Biochemistry* **38**, 11876–11886
711a. Otterlei, M., Kavli, B., Standal, R., Skjelbred, C., Bharati, S., and Krokan, H. E. (2000) *EMBO J.* **19**, 5542–5551
711b. Garvish, J. F., and Lloyd, R. S. (2000) *J. Mol. Biol.* **295**, 479–488
711c. Werner, R. M., and Stivers, J. T. (2000) *Biochemistry* **39**, 14054–14064
711d. Dinner, A. R., Blackburn, G. M., and Karplus, M. (2001) *Nature (London)* **413**, 752–755
712. Mol, C. D., Kuo, C.-F., Thayer, M. M., Cunningham, R. P., and Tainer, J. A. (1995) *Nature (London)* **374**, 381–386
713. Otterlei, M., Warbrick, E., Nagelhus, T. A., Haug, T., Slupphaug, G., Akbari, M., Aas, P. A., Steinsbekk, K., Bakke, O., and Krokan, H. E. (1999) *EMBO J.* **18**, 3834–3844
714. Carey, D. C., and Strauss, P. R. (1999) *Biochemistry* **38**, 16553–16560
715. Harris, J. M., McIntosh, E. M., and Muscat, G. E. O. (1999) *J. Mol. Biol.* **288**, 275–287
716. Williams, S. D., and David, S. S. (1999) *Biochemistry* **38**, 15417–15424
717. Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P., and Tainer, J. A. (1995) *EMBO J.* **14**, 4108–4120
717a. Dizdaroglu, M., Bauche, C., Rodriguez, H., and Laval, J. (2000) *Biochemistry* **39**, 5586–5592
718. Karahalil, B., Roldán-Arjona, T., and Dizdaroglu, M. (1998) *Biochemistry* **37**, 590–595
718a. Bebenek, K., Tissier, A., Frank, E. G., McDonald, J. P., Prasad, R., Wilson, S. H., Woodgate, R., and Kunkel, T. A. (2001) *Science* **291**, 2156–2159
719. Fuxreiter, M., Warshel, A., and Osman, R. (1999) *Biochemistry* **38**, 9577–9589
720. Croteau, D. L., and Bohr, V. A. (1997) *J. Biol. Chem.* **272**, 25409–25412
721. Lipscomb, L. A., Peek, M. E., Morningstar, M. L., Verghis, S. M., Miller, E. M., Rich, A., Essigmann, J. M., and Williams, L. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 719–723
722. Bruner, S. D., Norman, D. P. G., and Verdine, G. L. (2000) *Nature (London)* **403**, 859–866
723. Hatahet, Z., Zhou, M., Reha-Krantz, L. J., Morrical, S. W., and Wallace, S. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8556–8561
724. Bhagwat, M., and Gerlt, J. A. (1996) *Biochemistry* **35**, 659–665
725. Porello, S. L., Cannon, M. J., and David, S. S. (1998) *Biochemistry* **37**, 6465–6475
725a. Miyako, K., Takamatsu, C., Umeda, S., Tajiri, T., Furuichi, M., Nakabeppu, Y., Sekiguchi, M., Hamasaki, N., Takeshige, K., and Kang, D. (2000) *J. Biol. Chem.* **275**, 12326–12330
725b. Zharkov, D. O., Gilboa, R., Yagil, I., Kycia, J. H., Gerchman, S. E., Shoham, G., and Grollman, A. P. (2000) *Biochemistry* **39**, 14768–14778
726. Manuel, R. C., Czerwinski, E. W., and Lloyd, R. S. (1996) *J. Biol. Chem.* **271**, 16218–16226
727. Guibourt, N., Castaing, B., Auffret Van Der Kemp, P., and Boiteux, S. (2000) *Biochemistry* **39**, 1716–1724
728. Kang, D., Nishida, J.-i, Iyama, A., Nakabeppu, Y., Furuichi, M., Fujiwara, T., Sekiguchi, M., and Takeshige, K. (1995) *J. Biol. Chem.* **270**, 14659–14665
729. Bessman, M. J., Frick, D. N., and O'Handley, S. F. (1996) *J. Biol. Chem.* **271**, 25059–25062
730. Masaoka, A., Terato, H., Kobayashi, M., Honsho, A., Ohyama, Y., and Ide, H. (1999) *J. Biol. Chem.* **274**, 25136–25143
731. Kung, H. C., and Bolton, P. H. (1997) *J. Biol. Chem.* **272**, 9227–9236
732. Demple, B., and Karran, P. (1983) *Trends Biochem. Sci.* **8**, 137–139
733. Spratt, T. E., Wu, J. D., Levy, D. E., Kanugula, S., and Pegg, A. E. (1999) *Biochemistry* **38**, 6801–6806
734. Myers, L. C., Verdine, G. L., and Wagner, G. (1993) *Biochemistry* **32**, 14089–14094
734a. Hashimoto, H., Inoue, T., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T., and Kai, Y. (1999) *J. Mol. Biol.* **292**, 707–716
734b. Lin, Y., Dötsch, V., Wintner, T., Peariso, K., Myers, L. C., Penner-Hahn, J. E., Verdine, G. L., and Wagner, G. (2001) *Biochemistry* **40**, 4261–4271
735. Asaeda, A., Ide, H., Asagoshi, K., Matsuyama, S., Tano, K., Murakami, A., Takamori, Y., and Kubo, K. (2000) *Biochemistry* **39**, 1959–1965
736. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) *EMBO J.* **19**, 758–766
737. Lutsenko, E., and Bhagwat, A. S. (1999) *J. Biol. Chem.* **274**, 31034–31038
738. Neddermann, P., Gallinari, P., Lettieri, T., Schmid, D., Truong, O., Hsuan, J. J., Wiebauer, K., and Jiricny, J. (1996) *J. Biol. Chem.* **271**, 12767–12774
739. Um, S., Harbers, M., Benecke, A., Pierrat, B., Losson, R., and Chambon, P. (1998) *J. Biol. Chem.* **273**, 20728–20736
740. Haber, J. E. (1999) *Trends Biochem. Sci.* **24**, 271–275
741. Cox, M. M., Goodman, M. F., Kreuzer, K. N., Sherratt, D. J., Sandler, S. J., and Marians, K. J. (2000) *Nature (London)* **404**, 37–41
741a. Connelly, J. C., and Leach, D. R. F. (2002) *Trends Biochem. Sci.*, **27**, 410 – 418
741b. Hopfner, K.-P., Craig, L., Moncalian, G., Zinkel, R. A., Usui, T., Owen, B. A. L., Karcher, A., Henderson, B., Bodmer, J.-L., McMurray, C. T., Carney, J. P., Petrini, J. H. J., and Tainer, J. A. (2002) *Nature (London)* **418**, 562 – 566
742. Webb, B. L., Cox, M. M., and Inman, R. B. (1999) *J. Biol. Chem.* **274**, 15367–15374
742a. Johnson, R. D., and Jasin, M. (2000) *EMBO J.* **19**, 3398–3407
742b. Arnaudeau, C., Lundin, C., and Helleday, T. (2001) *J. Mol. Biol.* **307**, 1235–1245
742c. Kondo, T., Wakayama, T., Naiki, T., Matsumoto, K., and Sugimoto, K. (2001) *Science* **294**, 867–870
742d. Hoege, C., Pfander, B., Moldovan, G.-L, Pyrowolakis, G., and Jentsch, S. (2002) *Nature (London)* **419**, 135 – 141
743. Yamaguchi-Iwai, Y., Sonoda, E., Sasaki, M. S., Morrison, C., Haraguchi, T., Hiraoka, Y., Yamashita, Y. M., Yagi, T., Takata, M., Price, C., Kakazu, N., and Takeda, S. (1999) *EMBO J.* **18**, 6619–6629
743a. Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., Devine, S. K., Day J. P., Wilkinson, A., d'Adda di Fagagna, F., Devine, K. M., Bowater, R. P., Jeggo, P. A., Jackson, S. P., and Doherty, A. J. (2002) *Science* **297**, 1686 – 1689
744. Hays, S. L., Firmenich, A. A., and Berg, P. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6925–6929
745. Critchlow, S. E., and Jackson, S. P. (1998) *Trends Biochem. Sci.* **23**, 394–398
746. Wilson, T. E., and Lieber, M. R. (1999) *J. Biol. Chem.* **274**, 23599–23609
746a. Fugmann, S. D. (2002) *Nature (London)* **416**, 691–694
746b. Ooi, S. L., Shoemaker, D. D., and Boeke, J. D. (2001) *Science* **294**, 2552–2556
747. Jackson, S. P., and Jeggo, P. A. (1995) *Trends Biochem. Sci.* **20**, 412–415
748. Chu, G. (1997) *J. Biol. Chem.* **272**, 24097– 24100
749. Hammarsten, O., DeFazio, L. G., and Chu, G. (2000) *J. Biol. Chem.* **275**, 1541–1550
750. Brown, K. D., Lataxes, T. A., Shangary, S., Mannino, J. L., Giardina, J. F., Chen, J., and Baskaran, R. (2000) *J. Biol. Chem.* **275**, 6651–6656
750a. Walker, J. R., Corpina, R. A., and Goldberg, J. (2001) *Nature (London)* **412**, 607–614
750b. Rouse, J., and Jackson, S. P. (2002) *Science* **297**, 547 – 551
750c. Kolodner, R. D., Putnam, C. D., and Myung, K. (2002) *Science* **297**, 552 –557
750d. Carr, A. M. (2002) *Science* **297**, 557 – 558
750e. Cha, R. S., and Kleckner, N. (2002) *Science* **297**, 602 – 606
750f. Sogo, J. M. Lopes, M., and Foiani, M. (2002) *Science* **297**, 599 – 602
750g. Celeste, A., and 20 other authors (2002) *Science* **296**, 922 – 927
750h. Bird, A. W., Yu, D. Y., Pray-Grant, M. G., Qiu, Q., Harmon, K. E., Megee, P. C., Grant, P. A., Smith, M. M., and Christman, M. F. (2002) *Nature (London)* **419**, 411 – 415
750i. Pickart, C. M. (2002) *Nature (London)* **419**, 120 – 121
751. Flanders, P. H. (1981) *Sci. Am.* **245**(Nov), 72–80
752. Kenyon, C. J. (1983) *Trends Biochem. Sci.* **8**, 84–87
753. Walker, G. C. (1995) *Trends Biochem. Sci.* **20**, 416–420
753a. Yasuda, T., Morimatsu, K., Kato, R., Usukura, J., Takahashi, M., and Ohmori, H. (2001) *EMBO J.* **20**, 1192–1202
754. McDonald, J. P., Peat, T. S., Levine, A. S., and Woodgate, R. (1999) *J. Mol. Biol.* **285**, 2199–2209
754a. Goodman, M. F. (2002) *Ann. Rev. Biochem.* **71**, 17 – 50
754b. Livneh, Z. (2001) *J. Biol. Chem.* **276**, 25639–25642
755. Johnson, R. E., Washington, M. T., Prakash, S., and Prakash, L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12224–12226
756. Baynton, K., and Fuchs, R. P. P. (2000) *Trends Biochem. Sci.* **25**, 74–79
757. Murakumo, Y., Roth, T., Ishii, H., Rasio, D., Numata, S.-i, Croce, C. M., and Fishel, R. (2000) *J. Biol. Chem.* **275**, 4391–4397
758. Wood, R. D. (1999) *Nature (London)* **399**, 639–640
759. Daube, S. S., Tomer, G., and Livneh, Z. (2000) *Biochemistry* **39**, 348–355
759a. Matsuda, T., Bebenek, K., Masutani, C., Rogozin, I. B., Hanaoka, F., and Kunkel, T. A. (2001) *J. Mol. Biol.* **312**, 335–346

References

759b. Kusumoto, R., Masutani, C., Iwai, S., and Hanaoka, F. (2002) *Biochemistry* **41**, 6090 – 6099

760. Hübscher, U., Nasheuer, H.-P., and Syväoja, J. E. (2000) *Trends Biochem. Sci.* **25**, 143 – 147

760a. Friedberg, E. C., Wagner, R., and Radman, M. (2002) *Science* **296**, 1627 – 1630

760b. Haracska, L., Johnson, R. E., Unk, I., Phillips, B. B., Hurwitz, J., Prakash, L., and Prokash, S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14256 – 14261

760c. Johnson, R. E., Prakash, S., and Prakash, L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3838 – 3843

760d. García-Díaz, M., Domínguez, O., López-Fernández, L. A., de Lera, L. T., Saníger, M. L., Ruiz, J. F., Párraga, M., García-Ortiz, M. J., Kirchhoff, T., del Mazo, J., Bernad, A., and Blanco, L. (2000) *J. Mol. Biol.* **301**, 851 – 867

760e. Wang, Z., Castano, I. B., De Las Penas, A., Adams, C., and Christman, M. F. (2000) *Science* **289**, 774 – 779

760f. Frank, E. G., Tissier, A., McDonald, J. P., Rapic-Otrin, V., Zeng, X., Gearhart, P. J., and Woodgate, R. (2001) *EMBO J.* **20**, 2914 – 2922

760g. Shimizu, K., Kawasaki, Y., Hiraga, S.-I., Tawaramoto, M., Nakashima, N., and Sugino, A. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9133 – 9138

761. de Murcia, G., and de Murcia, J. M. (1994) *Trends Biochem. Sci.* **19**, 172 – 176

761a. Jacobson, M. K., and Jacobson, E. L. (1999) *Trends Biochem. Sci.* **24**, 415 – 417

762. Lindahl, T., Satoh, M. S., Poirier, G. G., and Klungland, A. (1995) *Trends Biochem. Sci.* **20**, 405 – 411

763. Aoufouchi, S., Yélamos, J., and Milstein, C. (1999) *J. Mol. Biol.* **290**, 943 – 949

763a. Smith, S. (2001) *Trends Biochem. Sci.* **26**, 174 – 179

764. Ames, B. N., Magaw, R., and Gold, L. S. (1987) *Science* **236**, 271 – 280

765. Gold, L. S., Slone, T. H., Stern, B. R., Manley, N. B., and Ames, B. N. (1992) *Science* **258**, 261 – 265

766. Miller, J. A., and Miller, E. C. (1976) *Fed. Proc.* **35**, 1316 – 1321

767. Yu, H., Eritja, R., Bloom, L. B., and Goodman, M. F. (1993) *J. Biol. Chem.* **268**, 15935 – 15943

768. Goodman, M. F., and Ratliff, R. L. (1983) *J. Biol. Chem.* **258**, 12842 – 12846

769. Sowers, L. C., Fazakerley, G. V., Eritja, R., Kaplan, B. E., and Goodman, M. F. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5434 – 5438

770. Ripley, L. S., and Clark, A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6954 – 6958

771. Ginell, S. L., Kuzmich, S., Jones, R. A., and Berman, H. M. (1990) *Biochemistry* **29**, 10461 – 10465

772. Wolff, I. A., and Wasserman, A. E. (1972) *Science* **177**, 15 – 18

773. Russell, W. L., Kelly, E. M., Hunsicker, P. R., Bangham, J. W., Maddux, S. C., and Phipps, E. L. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5818 – 5819

774. Shaver-Walker, P. M., Urlando, C., Tao, K. S., Zhang, X. B., and Heddle, J. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11470 – 11474

775. Wurdeman, R. L., Church, K. M., and Gold, B. (1989) *J. Am. Chem. Soc.* **111**, 6408 – 6412

776. Klein-Szanto, A. J. P., Iizasa, T., Momiki, S., Garcia-Palazzo, I., Caamano, J., Metcalf, R., Welsh, J., and Harris, C. C. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6693 – 6697

777. Craddock, V. M. (1983) *Nature (London)* **306**, 638

778. Hirai, N., Kingston, D. G. I., Van Tassell, R. L., and Wilkins, T. D. (1982) *J. Am. Chem. Soc.* **104**, 6149 – 6150

779. Gupta, I., Suzuki, K., Bruce, W. R., Krepinsky, J. J., and Yates, P. (1984) *Science* **225**, 521 – 522

780. Ozawa, N., and Guengerich, F. P. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5266 – 5270

781. Langouët, S., Mican, A. N., Müller, M., Fink, S. P., Marnett, L. J., Muhle, S. A., and Guengerich, F. P. (1998) *Biochemistry* **37**, 5184 – 5193

782. Leonard, G. A., McAuley-Hecht, K. E., Gibson, N. J., Brown, T., Watson, W. P., and Hunter, W. N. (1994) *Biochemistry* **33**, 4755 – 4761

783. Weisenseel, J. P., Moe, J. G., Reddy, G. R., Marnett, L. J., and Stone, M. P. (1995) *Biochemistry* **34**, 50 – 64

784. Palejwala, V. A., Simha, D., and Humayun, M. Z. (1991) *Biochemistry* **30**, 8736 – 8743

785. Zegar, I. S., Setayesh, F. R., DeCorte, B. L., Harris, C. M., Harris, T. M., and Stone, M. P. (1996) *Biochemistry* **35**, 4334 – 4348

786. Bucala, R., Model, P., and Cerami, A. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 105 – 109

787. Shires, T. K., Tresnak, J., Kaminsky, M., Herzog, S. L., and Truc-pham, B. (1990) *FASEB J.* **4**, 3340 – 3346

788. Napolitano, R. L., Lambert, I. B., and Fuchs, R. P. P. (1994) *Biochemistry* **33**, 1311 – 1315

789. Isono, K., and Yourno, J. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1612 – 1617

790. Stoloff, L. (1987) *Science* **237**, 1283

791. Baertschi, S. W., Raney, K. D., Stone, M. P., and Harris, T. M. (1988) *J. Am. Chem. Soc.* **110**, 7929 – 7931

792. Iyer, R. S., Coles, B. F., Raney, K. D., Thier, R., Guengerich, F. P., and Harris, T. M. (1994) *J. Am. Chem. Soc.* **116**, 1603 – 1609

793. Heidelberger, C. (1975) *Ann. Rev. Biochem.* **44**, 79 – 121

794. Phillips, D. H. (1983) *Nature (London)* **303**, 468 – 472

795. Flowers, L., Bleczinski, W. F., Burczynski, M. E., Harvey, R. G., and Penning, T. M. (1996) *Biochemistry* **35**, 13664 – 13672

796. Schurter, E. J., Sayer, J. M., Oh-hara, T., Yeh, H. J. C., Yagi, H., Luxon, B. A., Jerina, D. M., and Gorenstein, D. G. (1995) *Biochemistry* **34**, 9009 – 9020

797. Mao, B., Gu, Z., Gorin, A., Chen, J., Hingerty, B. E., Amin, S., Broyde, S., Geacintov, N. E., and Patel, D. J. (1999) *Biochemistry* **38**, 10831 – 10842

798. Tang, M.-s, Bohr, V. A., Zhang, X.-s, Pierce, J., and Hanawalt, P. C. (1989) *J. Biol. Chem.* **264**, 14455 – 14462

799. Cho, B. P., and Zhou, L. (1999) *Biochemistry* **38**, 7572 – 7583

800. Eis, P. S., Smith, J. A., Rydzewski, J. M., Case, D. A., Boger, D. L., and Chazin, W. J. (1997) *J. Mol. Biol.* **272**, 237 – 252

801. Perera, F., and Petito, C. (1982) *Science* **216**, 1285 – 1291

802. Rall, D. P. (1992) *Science* **257**, 1330

803. Hanson, D. J. (1991) *Chem. Eng. News* **Aug 12**, 7 – 14

804. Marshall, E. (1990) *Science* **247**, 276 – 277

805. Hileman, B. (1990) *Chem. Eng. News* **May 7**, 4

806. Ames, B. N. (1979) *Science* **204**, 587 – 593 (see also discussion in Science 224, 659 – 670, 757 – 760 (1984))

807. Devoret, R. (1979) *Sci. Am.* **241**(Aug), 40 – 49

808. Muench, K. F., Misra, R. P., and Humayun, M. Z. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6 – 10

809. Green, M. M., Todo, T., Ryo, H., and Fujikawa, K. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6667 – 6671

810. Gold, M. O., Blum, A., and Ames, B. N. (1978) *Science* **200**, 785 – 787

811. Commoner, B., Vithayathil, A. J., Dolara, P., Nair, S., Madyastha, P., and Cuca, G. C. (1978) *Science* **201**, 913 – 916

812. Tennant, R. W., Margolin, B. H., Shelby, M. D., Zeiger, E., Haseman, J. K., Spalding, J., Caspary, W., Resnick, M., Stasiewicz, S., Anderson, B., and Minor, R. (1987) *Science* **236**, 933 – 941

813. Epstein, S. S., Ames, B. N., and Gold, L. S. (1988) *Science* **240**, 1043 – 1047

814. Ashby, J., and Morrod, R. S. (1991) *Nature (London)* **352**, 185 – 186

815. Weinstein, I. B. (1991) *Science* **251**, 387 – 388

816. Stone, R. (1995) *Science* **268**, 356 – 357

817. Randerath, K., Reddy, M. V., and Gupta, R. C. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6126 – 6129

818. Kadlubar, F. F. (1992) *Nature (London)* **360**, 189

819. Schut, H. A. J., and Shiverick, K. T. (1992) *FASEB J.* **6**, 2942 – 2951

Study Questions

1. Demethylation of 5-methylcytosine in DNA during early embryonic development has been proposed (see Chapter 32). Can you suggest one or more mechanisms by which such demethylation could occur?

2. Describe the structures and functions of histones and of nucleosomes in eukaryotic cells. Are there comparable proteins and structures in bacteria? Do you think that our knowledge of histones and nucleosomes is nearly complete?

3. List the major molecular components required for replication of DNA in *E. coli.* Describe briefly the functions of each protein or other component.

4. Compare replication in bacteria and in eukaryotes.

5. Compare systhesis of the leading and lagging strands in the elongation phase of DNA replication. Explain why DNA polymerases may have difficulty in replicating the 3'-end of the *lagging* strand of **linear** DNA. How has this problem been solved in many bacterial and viral systems? In eukaryotic cells?

6. The circular chromosome of an *E. coli* cell contains 4.6 x 10^6 base pairs. If a replication fork moves at a rate of ~1000 nucleotides per second, how much time will be required for replication of the DNA? Cells of *E. coli* can divide every 20 minutes under favorable conditions. How can you explain this rapid rate of growth?

7. DNA ligase, whose reaction is reversible, is able to relax supercoiled circular DNA in the presence of AMP but not in its absence. Outline the chemical mechanism of the ligase reaction. Why is it dependent on AMP? What other DNA ligase mechanism is known?

8. DNA polymerases involved in replication require a primer. Why? What is the nature of the primer?

9. Why is it essential for a cell to have several different mechanisms of DNA repair? Describe some of these mechanisms.

10. Why do cells use error-prone DNA polymerases under some circumstances?

11. Do you see a relationship of some types of DNA repair to the chemical events during meiosis?

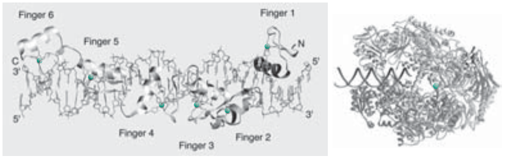
12. Is O^6-methyltransferase an enzyme?

13. Why are high rates of mutation observed in regions of DNA that contain 5-methylcytosine?

14. Why is uracil-DNA glycosylase important in DNA repair? Is it important for DNA replication?

15. Why do cells exposed to visible light following irradiation by ultraviolet light have a greater survival rate than cells kept in the dark after UV irradiation?

16. Can exposure of *E. coli* to nitrous acid (HNO_2) lead to mutation of a $tRNA^{GLY}$ to an amber suppressor? The Gly codons are GGX (where X = any nucleotide) and the amber codon is UAG.



Left. The N-terminal 190-residue fragment of transcription factor TFIIIA of *Xenopus laevis* bound to a 31 bp DNA segment of the promoter region for 5S ribosomal RNA. Six zinc finger motifs (zinc atoms are green) bind in several ways into the major groove of DNA and across the minor groove. From Nolte *et al*, *Proc. Natl. Acad. Sci. USA* **95**, 2938–2943, 1998. Courtesy of Raymond S. Brown. Right. Ribbon drawing of the three-dimensional structure of a 10-subunit form of yeast RNA polymerase II, which transcribes genes to form messenger RNA. A 20 base pair segment of B-DNA has been modeled but the transcription bubble is not shown. The active site Mg^{2+} is green. From Cramer *et al*, *Science* **288**, 640–649, 2000. Courtesy of Roger D. Kornberg.

Contents

1603 .... **A. Transcription and Processing of RNA in Bacteria**
1603 .... 1. The *lac* Operon
1604 .... *Repression and induction*
1605 .... *Nucleotide sequence of the lac control region*
1606 .... *The repressor structure*
1607 .... 2. Initiation of Transcription
1607 .... *Bacterial RNA polymerase*
1607 .... *Promoter sequences*
1608 .... *Control of stable RNA synthesis*
1609 .... *The initiation reaction*
1610 .... 3. Elongation of RNA Transcripts
1611 .... 4. Control of Various Bacterial Operons
1611 .... *Many repressors*
1612 .... *Inducible operons*
1612 .... *Feedback repression*
1612 .... *Positive control by activator proteins*
1613 .... *Control by looping*
1614 .... *Bacterial enhancers*
1614 .... *Other two-component control systems*
1614 .... *Antisense RNA*
1615 .... 5. Termination of Transcription in Bacteria
1615 .... *Attenuation*
1616 .... *Rho and other termination factors*
1616 .... 6. Effects of Antibiotics
1619 .... 7. Processing of Bacterial RNA
1619 .... *Ribosomal RNA*
1620 .... *Transfer RNA*
1622 .... **B. Viral RNA in Prokaryotes**
1622 .... 1. The Lysogenic State of Phage λ
1623 .... 2. Replication of RNA Bacteriophages
1624 .... **C. Transcription in Eukaryotic Cells and in Archaea**
1624 .... 1. Eukaryotic Nuclei and Transcription
1626 .... 2. RNA Polymerases
1628 .... 3. Transcriptional Units and Initiation of RNA Synthesis
1628 .... 4. Promoters, Transcription Factors, Enhancers, and Activators
1628 .... *The SV40 early promoter and transcription factor Sp1*
1628 .... *The TATA binding protein and general transcription initiation factors*
1630 .... *Transcriptional activators*
1630 .... *Mediators and coactivators*
1630 .... *Enhancers*
1632 .... 5. The Variety of DNA-Binding Proteins
1633 .... *Leucine zipper transcription factors*
1633 .... *Control of growth*
1634 .... *Response elements*
1634 .... *Zinc-containing transcription factors*
1634 .... *Winged helix transcription factors*
1634 .... *The NF-κB/Rel proteins*
1634 .... *HMG proteins as transcription factors*
1635 .... 6. The Variety of Transcriptional Responses
1635 .... *Nuclear hormone receptors*
1635 .... *Nutrient control*
1635 .... *Stress proteins*
1636 .... *Light-induced transcription*
1636 .... *Homeotic genes and homeodomain proteins*
1636 .... 7. Transcription by RNA Polymerases I and III
1637 .... 8. Elongation, Modification, and Termination of Transcription
1638 .... 9. Conformational Properties of RNA
1638 .... **D. Processing of Eukaryotic RNAs**
1638 .... 1. Ribosomal RNA
1640 .... 2. Small Nuclear, and Nucleolar, and Cytoplasmic RNAs
1640 .... *Processing of ribosomal RNA*
1641 .... *Modification guide RNAs*
1641 .... *Transcription and processing of snRNAs and snoRNAs*
1641 .... 3. Processing of 5S RNA and tRNAs
1642 .... 4. Messenger RNA, Caps, and Polyadenylate Tails
1642 .... 5. Splicing
1643 .... *Self-splicing RNA*
1646 .... *Relationship to viroids and virusoids*
1646 .... *Pre-tRNAs*
1647 .... *The spliceosome*
1647 .... *Alternative splicing pathways*
1647 .... *Trans splicing*
1648 .... 6. Modification and Editing of RNAs
1648 .... *Editing of RNA transcripts*
1648 .... *Finishing the transcripts*
1649 .... **E. Transcription of Mitochondrial, Chloroplast, and Viral Genes**
1649 .... 1. Viral Transcription and Replication
1650 .... 2. Replication of RNA Viruses
1650 .... *Small RNA viruses*
1650 .... *Influenza viruses*
1651 .... *HIV-1 and other retroviruses*
1651 .... *Accessory regulatory genes*
1657 .... 3. Retrotransposons

1657 .... **References**
1667 .... **Study Questions**

Boxes

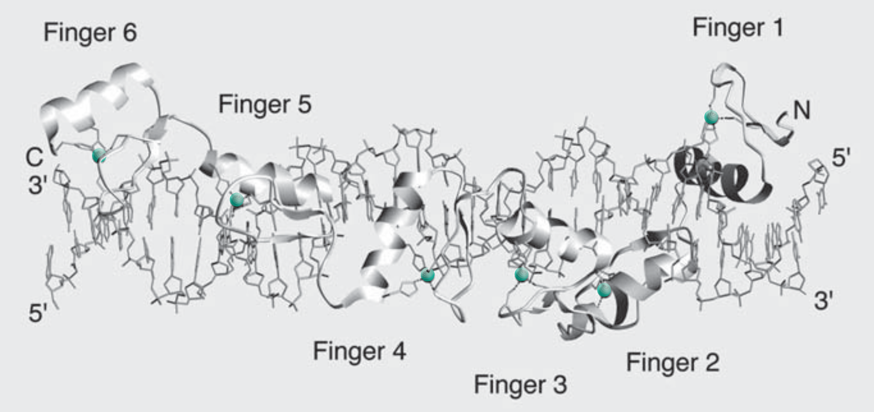
1618 .... Box 28-A The Antibiotics Rifamycin, Rifampicin, and Actinomycin D
1625 .... Box 28-B Powerful Poisons from Mushrooms
1654 .... Box 28-C Synthetic Antiviral Compounds

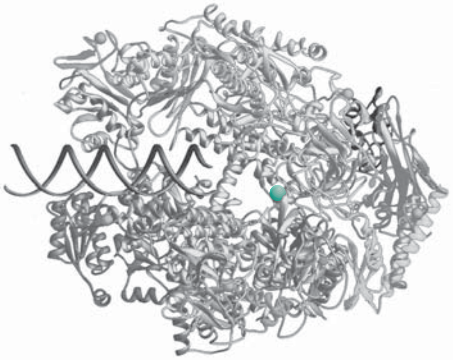
Tables

1631 .... Table 28-1 Nucleotide Sequences of Some Commonly Found Regulatory Elements in Promoters and Upstream Activator Sites
1632 .... Table 28-2 Functional Classification of Positive-Acting Eukaryotic Transcription Factors
1641 .... Table 28-3 Some Eukaryotic Small Nuclear and Nucleolar RNA Molecules

The Transcription of Genes

28





The copying of genetic information from DNA into messenger RNA is the initial step in the chain of reactions leading to synthesis of the multitude of proteins and specialized RNA molecules needed by cells. The requirement for these macromolecules varies with conditions, and in eukaryotic cells, with the stage of differentiation. Therefore, it is not surprising that transcription is highly controlled.

Cells make four principal kinds of RNA: ribosomal (rRNA), transfer (tRNA), messenger (mRNA), and a variety of small RNAs. The last, which range in length from a few up to several hundred nucleotide units,[1–1b] are designated variously as sRNAs, ncRNAs, miRNAs, siRNAs, snRNAs, and snoRNAs. The abbreviation *s*, *nc*, *mi*, *si*, *sn*, and *sno* stand for small, noncoding, micro, silencing, small nuclear, and small nucleolar, respectively. All of these RNAs are synthesized as larger transcripts, which undergo cleavage and other modifications within the cell. Therefore, a second major topic in this chapter is the processing of RNA precursors. We will also consider the fact that cells may be hosts for RNA viruses, may occasionally harbor RNA plasmids, and must sometimes transcribe viral DNA.

The absence of a nuclear membrane is a characteristic of bacteria that has a profound effect on transcription. Bacterial transcripts are processed rapidly, and their 5' ends often enter ribosomes and are directing protein synthesis, while the 3' ends of the genes are still being transcribed. In contrast, most eukaryotic RNA transcripts must be processed and transported out of the nucleus before they can function. As consequence, many aspects of the control of transcription differ between prokaryotes and eukaryotes.

A. Transcription and Processing of RNA in Bacteria

Even after the existence of mRNA had been recognized, it was not obvious how formation of single-stranded (ss) RNA would be accomplished, using a double-stranded (ds) template. The fact that purified RNA polymerases can synthesize RNA from the four ribonucleoside triphosphates using ssDNA as the template suggested that transcription, like DNA replication, involves base pairing. In line with this conclusion was the fact that the ssDNA obtained from bacteriophage ϕX174 was converted by RNA polymerases into a dsRNA–DNA hybrid. However, when dsDNA served as the template, free ssRNA was formed. Thus, it appeared likely that at the site of the polymerase action the dsDNA was momentarily pulled apart into single strands and that one of these was copied by the polymerase. More recent experiments have confirmed this view.

1. The *lac* Operon

Much of the terminology used to describe the control of transcription originated with Jacob and Monod. Based on studies of the induction of enzymes in bacteria they proposed the **operon model**.[1c–3] An operon is a regulated cluster of genes, one of which is shown diagrammatically in Fig. 28-1. This is the *lac* operon of *E. coli*. Found at position 8 min on the genetic map of Fig. 26-4, it is probably the most intensively studied group of *E. coli* genes. Three structural genes encode the amino acid sequences of β-galactosidase (*lacZ*),[3a] permease (*lacY*),[4] and a transacetylase (acetyltransferase, *lacA*), which transfers acetyl groups from acetyl-CoA to β galactosides. To account for the apparently synchronous control of these three genes, Jacob and Monod proposed that they function as a **transcriptional unit** the operon, which encodes a single molecule of mRNA. They proposed that each operon is controlled by a segment of the DNA molecule located at the beginning of the operon, i.e., at the 5' end of the coding chain or 3' end of the template chain. The first part of this **control region** they called the **promoter** (*P*). The promoter is the site of the initial binding of the RNA polymerase to the DNA, the binding constants for the association being very high. The rates of association and of initiation may be influenced strongly by various other proteins. One of these, the **catabolite gene activator protein** (**CAP**; also called cAMP receptor protein, CRP), is important to the *lac* operon. It also binds in the promoter region (Fig. 28-1) and stimulates transcription.

Repression and induction. Immediately adjacent to the promoter is the **operator** (*O*), which is a binding site for a **repressor** (R). When the operator is free, transcription is initiated and proceeds through the operator region and on to the genes coding for the three proteins. On the other hand, if the repressor is bound to the operator, transcription is blocked. When the operon model was first proposed, the chemical nature of the repressor was unknown. However, many repressors have been identified as oligomeric proteins able to undergo allosteric alteration. The *lac* repressor is made up of four identical 360-residue subunits. Each subunit has a helix–turn–helix binding domain that is specific for the DNA sequence of the operator and an allosteric binding site for an effector.[5,6] The drawing in Fig. 28-1 is simplified to show only two of the four subunits (see also Chapter 5, Section F,1).

The *lac* operon is ordinarily subject to repression and is activated by the presence of an **inducer**, now known to be **allolactose**, D-Gal*p*-β1→6-D-Glc. However, in experimental work artificial inducers such as **isopropyl-β-D-thiogalactoside** (**IPTG**) are most often used. Jacob and Monod postulated that the free repressor protein binds to the operator. In the presence of the inducer a conformational change takes place, destroying the affinity of the repressor protein for the operator site. Thus, in the presence of inducer the operator is not blocked, and the transcription takes place. Such an operon is said to be **negatively controlled** and **inducible**.

Important to the control of the operon is the **regulatory gene,** which codes for the synthesis of the repressor protein. In the case of the *lac* operon, the regulatory gene (the *I* gene) is located immediately preceding the operon itself (Fig. 28-1). However, for some operons the regulatory gene is located a considerable distance away. For example, the *gal* operon of *E. coli*, which codes for enzymes of galactose metabolism, is found at map position 17 min, while the regulatory gene is at 61 min.[7]

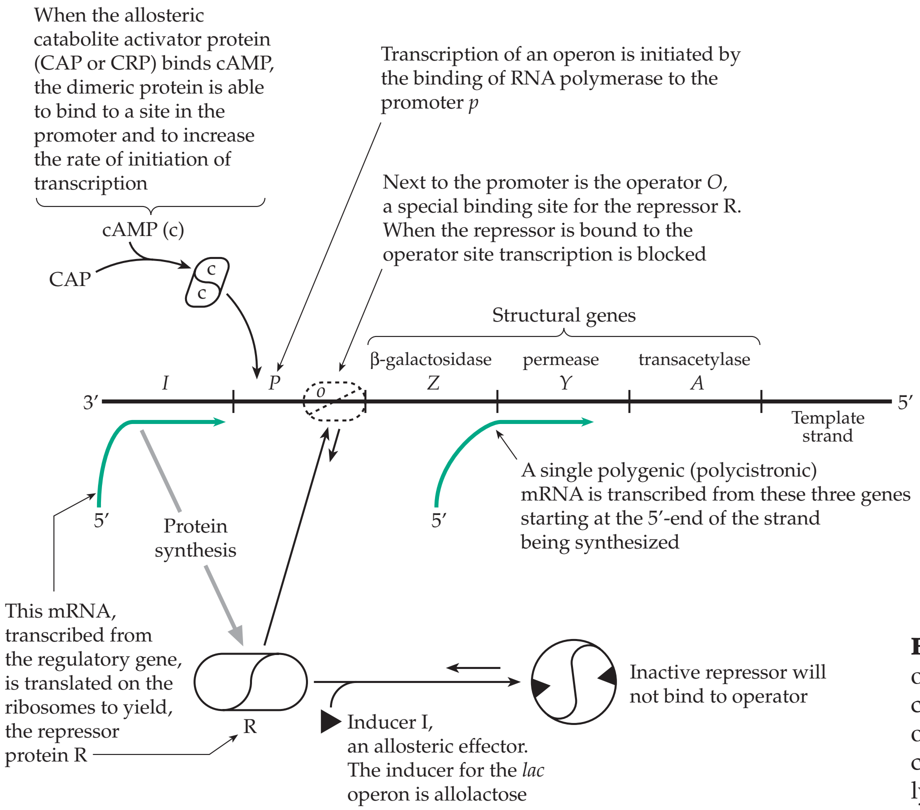

When the allosteric catabolite activator protein (CAP or CRP) binds cAMP, the dimeric protein is able to bind to a site in the promoter and to increase the rate of initiation of transcription
cAMP (c)
CAP
c
c
Transcription of an operon is initiated by the binding of RNA polymerase to the promoter *p*
Next to the promoter is the operator *O*, a special binding site for the repressor R. When the repressor is bound to the operator site transcription is blocked
Structural genes
β-galactosidase
permease
transacetylase
3′
I
P
o
Z
Y
A
5′
Template strand
5′
Protein synthesis
5′
A single polygenic (polycistronic) mRNA is transcribed from these three genes starting at the 5′-end of the strand being synthesized
This mRNA, transcribed from the regulatory gene, is translated on the ribosomes to yield, the repressor protein R
R
Inducer I, an allosteric effector. The inducer for the *lac* operon is allolactose
Inactive repressor will not bind to operator

Figure 28-1 Schematic representation of the *lac* operon of *E. coli* and of its control. Here only the template strand of the DNA is shown. However, the coding (nontranscribed) strand is usually the one labeled, as in Fig. 28-2.

Regulatory genes are normally transcribed at a slow but steady rate, presumably because RNA polymerase initiates RNA chains infrequently at the promoter sites of regulatory genes. Thus, each cell of *E. coli* normally contains only about ten molecules of the *lac* repressor protein. A mutation in a regulatory gene may lead to a defective repressor, which no longer binds at the operator. Then, transcription of the operon is uncontrolled, and mRNA is produced in greater amounts. In such a mutant strain (designated I^- in contrast to the normal I^+ strain) production of the enzyme representing the gene product becomes constitutive, just as is the formation of the enzymes of the central pathways of metabolism. The latter enzymes also appear to be produced regularly in large amounts without control by a repressor, establishing that *transcription* rather than translation regulates the expression of these genes. The operon can also become unregulated, i.e., **constitutively expressed**, if a mutation occurs in the operator site and the repressor can no longer bind because of the altered DNA sequence.

Nucleotide sequence of the lac control region. The sequence of the *E. coli* DNA representing the promoter-operator region of the *lac* operon is shown in Fig. 28-2. It includes the end of the *I* gene, at the left, and the beginning of the *Z* gene, at the right.[2] The series of codons representing the peptide sequence Glu-Ser-Gly-Gln-Stop at the left-hand end corresponds to the known C-terminal sequence of the repressor, while the three codons at the right are those of formyl-Met-Thr-Met, the known N-terminal sequence of the *Z* gene product β-galactosidase. Detailed genetic mapping of the region in 1973 made it possible to assign operator and promoter regions with confidence as indicated. The mRNA transcript begins in the operator region as shown in the figure. The initiation codon for the *Z* gene is 39 bases from the end of the transcript. In this figure the original numbering of the nucleotides is printed in black. Now it is customary to number from the transcription initiation position (+1). Green numbers are used in this way in Fig. 28-2. Positions to the left and toward the 5' end of the nontranscribed coding strand precede the initiation position. They are referred to as **upstream** and are numbered with negative integers. Positions to the right are **downstream** and numbered with positive integers.

The operator region was located by digesting the DNA with deoxyribonuclease in the presence of the repressor protein.[8] The bound repressor protected a region of 27 base pairs as indicated in the figure. The operator is centered on a region of local twofold rotational symmetry (Chapter 5; Fig. 5-34). The symmetry is not perfect, the sequence being **quasipalindromic**. The following precisely symmetric synthetic sequence, which contains an 11-bp inverted repeat of the left half of the *lac* operator sequence, binds *lac* repressor 8-fold more tightly than does the natural *E. coli* operator.[9]

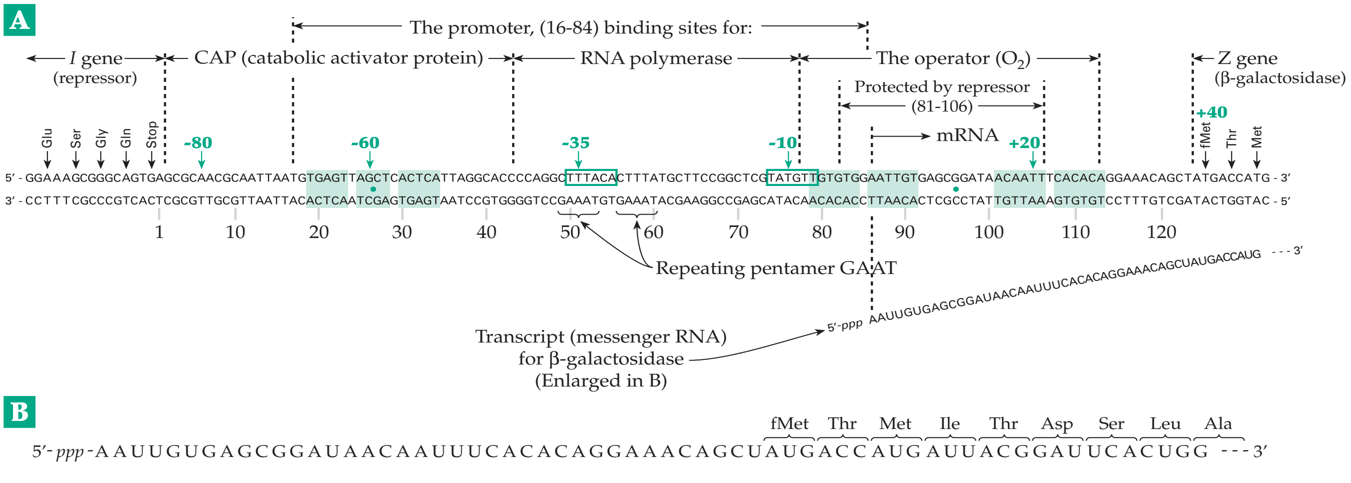

A
The promoter, (16-84) binding sites for:
I gene (repressor)
CAP (catabolic activator protein)
RNA polymerase
The operator (O₂)
Protected by repressor (81-106)
Z gene (β-galactosidase)
Glu
Ser
Gly
Gln
Stop
-80
-60
-35
-10
mRNA
+20
+40
fMet
Thr
Met
5′ - GGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATG - 3′
3′ - CCTTTCGCCCGTCACTCGCGTTGCGTTAATTACACTCAATCGAGTGAGTAATCCGTGGGGTCCGAAATGTGAAATACGAAGGCCGAGCATACAACACACCTTAACACTCGCCTATTGTTAAAGTGTGTCCTTTGTCGATACTGGTAC - 5′
1
10
20
30
40
50
60
70
80
90
100
110
120
Repeating pentamer GAAT
5′-ppp AAUUGUGAGCGGAUAACAAUUUCACACAGGAAACAGCUAUGACCAUG - - - 3′
Transcript (messenger RNA) for β-galactosidase (Enlarged in B)
B
fMet Thr Met Ile Thr Asp Ser Leu Ala
5′-ppp-AAUUGUGAGCGGAUAACAAUUUCACACAGGAAACAGCUAUGACCAUGAUUACGGAUUCACUGG - - - 3′

Figure 28-2 (A) Nucleotide sequence of the *lac* promoter-operator region of the *E. coli* chromosome.[7] The proposed locations of the *I* gene, the promoter (which contains CAP and RNA polymerase binding sites), the operator, and the beginning of the Z gene (β-galactosidase) are shown. Note the two regions of local 2-fold rotational symmetry, which are marked by bars and central dots and the repeating pentamer. Positions upstream (–) or downstream (+) from the +1 start position for transcription are marked in green. The –10 (Pribnow) and –35 promoter elements are boxed on the coding strand (top). Labels are usually applied to the coding strand. The mRNA is copied from the complementary template strand (bottom). (B) The sequence of an mRNA molecule initiated in the *lac* promoter-operator region of a mutant strain of *E. coli* with an altered promoter.[10] The peptide initiation amino acid is identified by the symbol fMet, and the successive amino acids from the known N-terminal sequence of β-galactosidase have been matched with the codons.

5′-G-A-A-T-T-G-T-G-A-G-C-G-C-T-C-A-C-A-A-T-T-C
3′-C-T-T-A-A-C-A-C-T-C-G-C-G-A-G-T-G-T-T-A-A-G

The dyad symmetry of the operator sequence is probably important in providing tight binding to two subunits of the symmetric tetrameric protein.[11–13] It is also possible that repressor molecules move along DNA chains in a one-dimensional diffusion process, and that the symmetry of the operator site facilitates recognition by a protein moving from either direction.[14,15]

The repressor structure. The *lac* repressor protein is a member of the **helix–turn–helix** family of DNA binding proteins (Fig. 28-3; see also Fig. 5-35). The first helix and the turn of this motif fit into the major groove of the DNA, the side chains from the helix interacting with specificity-determining groups in the major groove. The protein consists of three domains; the N-terminal DNA binding "head" (residues ~1– ~50), a core domain (residues 62–340), and a leucine heptad repeat domain (residues 340–360) that forms the dimer–dimer interface. The dimeric form of the repressor binds to the palindromic operator sequence. The tetramer can bind to two operator sequences.[16–20] The inducer IPTG binds to the core domain near the ONPF site shown in Fig. 28-3. The sequence, structure, and binding site of the core domain resemble those of sugar transport proteins such as the galactose-binding chemoreceptor protein (Fig. 4-18A). Binding of the inducer causes the conformational change that appears to disrupt the interactions of the "hinge helices," seen in the center of Fig. 28-3A, with the DNA. This causes the repressor to dissociate from the operator (Fig. 28-1) and allows RNA polymerase to bind and to initiate transcription.

The *E. coli lac* repressor is one of the most investigated of all proteins. For example, 4000 single-amino-acid mutants have been prepared and studied.[6,17,21] Suppressor mutations were used to determine the function of various portions of the protein.[22,23] Many of the mutant proteins were created using *amber* mutations that were induced in the gene at many positions. The mutated genes were transferred into plasmids for cloning. Each plasmid was used to infect five different strains of bacteria, each carrying a suppressor mutation that would introduce a different amino acid when the (termination) codon UAG was encountered (see Chapter 29, Section C,3). From these infected bacteria large quantities of the mutant forms of the *lac* repressor were isolated. It was found that many mutations near the N-terminal end interfered with binding of the repressor to DNA, whereas mutations near the center interfere with binding to the inducer.

In addition to the main *lac* operator O_1, which is marked in Fig. 28-2, there are two weaker auxiliary operator sequences designated O_2 and O_3 located 401

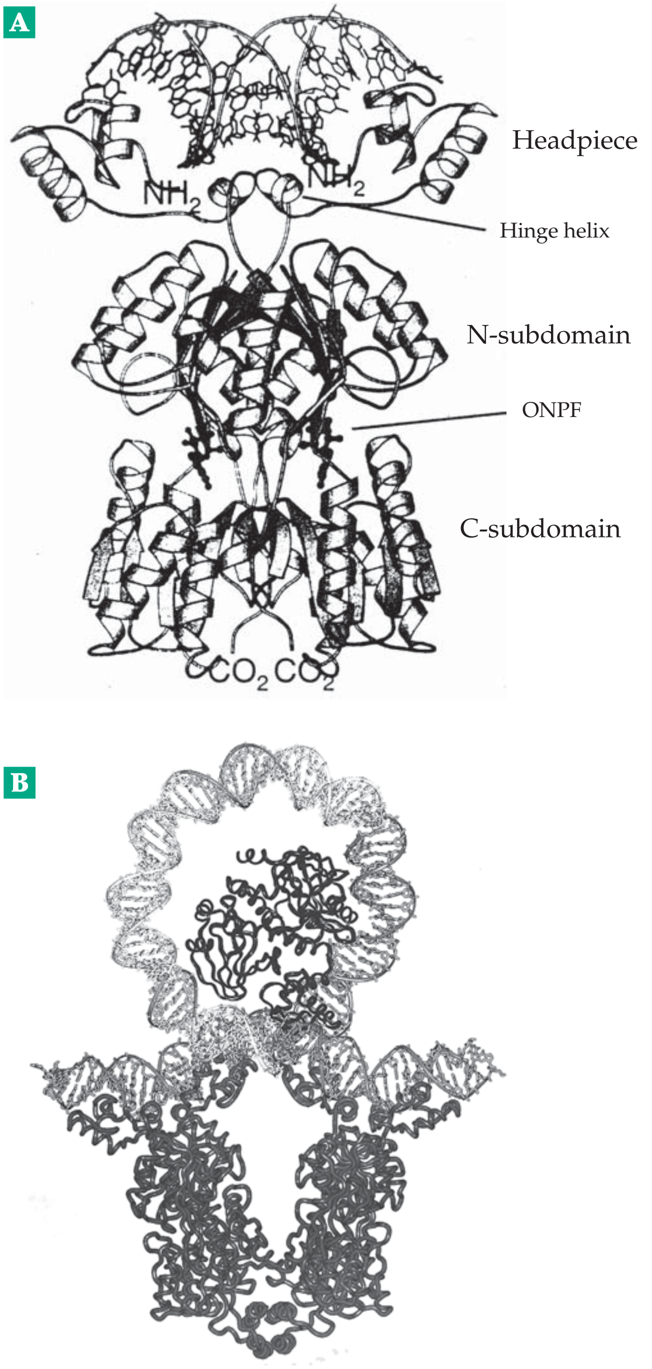


Figure 28-3 (A) Ribbon view of the dimeric *lac* repressor bound to a natural operator and to the anti-inducer *o*-nitrophenylfucoside (ONPF). The headpiece (residues 2–46) and the hinge helix (residues 50–58) form the DNA-binding domains. The core (residues 62–330), which is divided into N- and C-terminal subdomains, forms the binding site for ONPF. The C-terminal residues 334–360, which form a tetramerization domain, are absent from this MolScript drawing. Notice that the hinge helices bind to and widen the minor groove at the center of the operator. From Lewis *et al.*[5a] (B) Model of a 93-bp DNA loop corresponding to residues –82 to +11 of the *lac* operon (Fig. 28-2) bound to the tetrameric *lac* repressor. The active sites of the repressor are bound to the major operator O_1 and to the secondary operator O_3. From Lewis *et al.*[5]

bp downstream of O_1 in the lacZ coding region, and 92 bp upstream of O_1, respectively.[24] The DNA can bind to both O_1 and either O_2 or O_3 with a loop between them as in Fig. 28-3B. Binding of the 10–20 copies of the lac repressor present in a cell of *E. coli* to the operator sequence is so tight that expression of the genes controlled is reduced 1000-fold. However, when placed in front of other operators or in different positions relative to the Pribnow sequence (Fig. 28-2), its effectiveness varied greatly. The extent of repression appears to be affected both by thermodynamic factors[12,13] and by relative rates of repressor binding and of RNA polymerase movement (see also Eq. 28-1).[25]

2. Initiation of Transcription

The rate of RNA synthesis varies from one operon to another. Sequences of promoters, operators, and other control sequences as well as the state of repressors and activator proteins all affect these rates.[10,26,27] However, in every instance the first steps in transcription involve the binding of RNA polymerase to DNA.

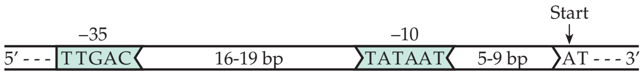
Bacterial RNA polymerase. Most RNA polymerases (**RNAPs**) are large multisubunit proteins. However, bacterial viruses sometimes induce their own RNA polymerases, and these may be monomeric. For example, the 99-kDa (883-residue) phage T7-encoded polymerase is a single peptide chain with a structure and two-metal-ion active site resembling those of *E. coli* DNA polymerase I.[28–30a] It is able to carry out all of the steps of the transcription cycle of the virus. In contrast, the most studied bacterial RNAP, that from *E. coli*, consists of five kinds of subunits[31–33d] with the composition $\alpha_2\beta\beta'\sigma\omega$. A similar composition has been found for RNAPs of other bacteria.[33e] Functions of the five subunits can be correlated directly with components of archaeal and eukaryotic RNAPs.[33f] However, the latter contain additional subunits. The two α subunits in the *E. coli* enzyme have identical sequences, but their locations and interactions are different.[33a–c]

Gene symbol	Subunit	Molecular mass (kDa)	Number of amino acid residues
α	*rpo*A	36.5	329
β	*rpo*B	150.6	1342
β′	*rpo*C	155.2	1407
σ	*rpo*D	70.2	
ω	*rpo*Z	6	

The three-dimensional structure of the *E. coli* RNAP bound to DNA in an initiation complex shows that the enzyme forms a groove into which the DNA can fit. It can then close to form a tunnel in which the template and nontemplate strands are separated (Fig. 28-4). The polymerase is present in large amounts, ~3000 molecules per cell in E. coli.[26] The β and β′ subunits of the *E. coli* enzyme each contain an essential zinc ion. The Zn^{2+} in the β′ subunit is present in a zinc finger motif near the N terminus.[34] It is thought to interact with the template strand of the DNA. The Zn^{2+} in the β strand is more loosely bound.[33,35] The active site is largely in the β subunit. In the assembly of the RNAP complex a dimer of the small α subunit binds to β, after which β′ is added.[36] The sequences of the β and β′ subunits have several highly conserved regions with homologous sequences from bacteria and from functionally equivalent regions in eukaryotic RNAPs. Three aspartate residues in the sequence NADFDGD may chelate two Mg^{2+} ions as in the active site of DNA polymerases (Fig. 27-13).[37] The basic chemistry of all of the polymerases may be similar, but the modular structure of the bacterial RNAP differs markedly from that of the DNA Pol I family.[38]

Of the RNA polymerase subunits σ (sigma) plays a unique role in initiation of transcription. It is required for the recognition of promoter sites.[39] However, it is not needed for elongation of an RNA chain and dissociates from the $\alpha_2\beta\beta'$ core complex soon after transcription is initiated. In a given bacterial species there is one predominant σ factor, but there are often smaller amounts of other σ factors with homologous sequences.[26,40,40a] In *E. coli* σ^{70} (where the superscript number is the molecular mass in kDa) is predominant, but other specialized σ subunits recognize different groups of promoters. For example, $\sigma^N(\sigma^{54})$ binds to promoters that allow transcription of genes involved in assimilation of nitrogen[41,41a] as well as in aromatic catabolism.[42] Synthesis of protein σ^S of *E. coli* is induced by stress such as carbon starvation.[43] At high temperatures (e.g., 40–49°C) σ^{32} is synthesized and permits transcription of genes for "heat shock proteins."[44] Actively growing cells of *Bacillus subtilis* contain at least five different sigma factors. An additional four control gene expression during spore formation.[45]

Promoter sequences. In 1975, Pribnow pointed out[46] that a series of six known promoters had a conserved 7-base sequence beginning six nucleotides upstream from the initiation site for transcription. Although this sequence varies somewhat from one promoter to another, it has been found in hundreds of *E. coli* promoters. This is called the **–10 region**, the **Pribnow sequence,** or Pribnow box (the last in recognition of the fact that people like to draw boxes around these special sequences). A typical 6-base consensus Pribnow sequence is 5'-TATAAT as written for the coding strand, whose sequence corresponds to that of the mRNA. Only three of these bases are highly conserved: 5'-TA_ _ _ T. For example, in the *lac* promoter (Fig. 28-2) the sequence is TATGTT. The nucleotides of the −10 consensus sequence are present with the following frequencies (as percentages): T(80)A(95)T(45)A(60)A(50)T(96).[47] The position of the −10 sequence is not exactly the same in all promoters but usually starts 5−9 bp upstream of the start position for transcription. About 16−19 bp upstream from the Pribnow sequence is another conserved sequence, the **−35 region**. A consensus sequence is TTGACA, the TTG sequence to the left being the most highly conserved. Both −10 and −35 regions are needed for efficient promoter activity. Lewin suggested the following "optimal" promoter sequence.[47] The location of these two regions in transcription initiation complex is shown in Fig. 28-4B.

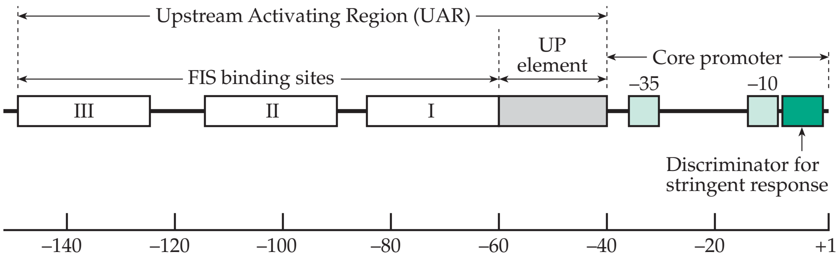


Naturally occurring promoters usually do not have the exact −10 and −35 consensus sequences, but artificially constructed promoters containing them are highly effective *in vivo*.[26] The fact that most promoters depart from these "ideal" sequences is not surprising because cells need varying amounts of different proteins. Promoter strengths vary over a range of at least 10^4. Much of this variation comes from variations in the specific −10 and −35 sequences, which appear to be specifically recognized by RNA polymerase. This variation includes an upward extension of the −10 region[48] for some promoters. Activator-binding sites are also often present in an **upstream activating region** (**UAR**; see following diagram). This may extend from the −35 sequence through the remainder of the promoter region and as far upstream as −200 to −500 bp. For example, the CAP-binding site in the *lac* operon DNA is centered at −60 (Fig. 28-2). In spite of the variation it was possible to locate 2584 operators (of which only 392 were previously known) and to predict the location of 2405 promoters, when the complete *E. coli* genome sequence became known.[49]

The binding of the various RNAP subunits, repressors, and activators has been studied using mutant promoter sequences,[31,37] antibiotic-binding sites,[50,51] chemical crosslinking reagents,[34,37] and a cysteine-tethered Fe-dependent DNA-cutting reagent.[32,52] Other "footprinting" techniques, e.g., observing cleavage of the DNA by hydroxyl radicals generated by reduction of H_2O_2 by Fe(II) (Fig. 5-50), have also been employed. It was shown that RNA polymerase binds to both the −10 and −35 sequences and also to sequences further upstream. The σ^{70} subunit associates with the DNA, principally the transcribed strand, along a region from about the −25 to the +12 position relative to the transcription start site.[32−32b] The α subunits bind to an UP element from ~−40 to −60 via their C-terminal domains (CTDs). See Fig. 28-4B.[53−54d]

Control of stable RNA synthesis. Whereas most mRNA has a relatively short lifetime, the stable ribosomal RNAs and transfer RNAs have much longer lives. Furthermore, in *E. coli* their transcription is coordinately controlled by seven *rrnB* P1 promoters.[55] The genes for stable RNA have promoters with the usual −10 and −35 sequences, but they contain a complex upstream activating sequence that includes the UP element and three binding sites for protein FIS (**the factor for inversion stimulation**).[56] This name reflects a second function, that of promoting inversion of a DNA segment in the Hin recombinase system (Chapter 27, Section D,3). FIS is a dsDNA-bending protein.[55,57] A dimer of 11.2-kDa subunits, it is an abundant protein. Like HU, IHF, H–NS, and Dps, it coats a significant fraction of the DNA in the *E. coli* chromosome.[58] It binds to the FIS sites using a helix−turn−helix motif. In addition, there is a GC-rich region at positions −7 to −1 with the consensus sequence 5'-GCGCC_C. It has been suggested that this **discriminator** is involved in the **stringent response**, the diminished rate of stable RNA synthesis observed during amino acid starvation.[59,59a]

The very complex stringent response, which involves ribosomes in the synthesis of guanosine 5'-triphosphate-3'-diphosphate (pppGpp) under some conditions,[60] is dealt with further in Chapter 29 (Section C,8). The regulator pppGpp, whose concentration may rise from ~50 µM to ~1 mM within minutes after deprivation of amino acids, may react directly with RNA polymerase at an allosteric site to inhibit transcription.[61−62a] Another possibility is that pppGpp acts on another protein that binds to the discriminator sequence. However, the fact that a phage T7 gene, which is also under stringent control, lacks the discriminator sequence argues against this.[63] pppGpp, whose concentration is usually high under conditions of slow growth, may be a major growth-controlling effector, which acts by inhibiting the replication of the rRNA needed for



ribosome formation and protein synthesis when nutrients are scarce. However, some operons are activated by the same effector.[62]

The initiation reaction. A promoter not only locates the site of initiation but also determines the direction of transcription and, therefore, the strand of the DNA duplex that is to serve as the template. The requirement for two specific recognition sequences ensures this directionality. The RNA polymerase may bind randomly to DNA, then move rapidly along the double helix until it locates a strong binding site[64–66] where it binds to the recognition sequences of the promoter through specific interactions in the major groove of the DNA helix (see Fig. 5-3).

A satisfactory mathematical model for initiation of transcription supposes that the polymerase and DNA bind reversibly to form a complex with formation constant K_f. This initial specific polymerase–promoter complex is referred to as a **closed complex** because it is thought that the bases in the DNA chain are all still paired. It is postulated that in a rate-determining step the closed complex is converted into an **open complex,** which is ready to initiate mRNA synthesis (Eq. 28-1).[26,67] In the open complex the hydrogen bonds holding together the base pairs have been broken, and the bases of the template chain are available for pairing with incoming ribonucleotide triphosphates.

$$\begin{array}{rl} \text{E(polymerase)} + \text{P(promoter)} \overset{K_f}{\rightleftharpoons} & \text{EP}_c\ \text{(closed)} \\ & \downarrow k \\ & \text{EP}_o\ \text{(open)} \\ & \downarrow \\ & \downarrow \\ & \text{Transcription} \end{array} \qquad (28\text{-}1)$$

It is clear from Eq. 28-1 that the efficiency of initiation depends upon both the affinity K_f and the rate constant k for opening of the double helix. Notice that the Pribnow sequence is AT-rich; therefore, opening of the helix at this point would be easier than in a GC-rich region. Thus, the Pribnow sequence may represent a point of entry of RNA polymerase to form the open complex.[67] Other upstream A•T tracts are often present frequently at about the −43 position in the UP element. They also seem to strengthen promoter activity.[68] The open complex is thought to undergo some kind of isomerization to form an **initial transcribing**

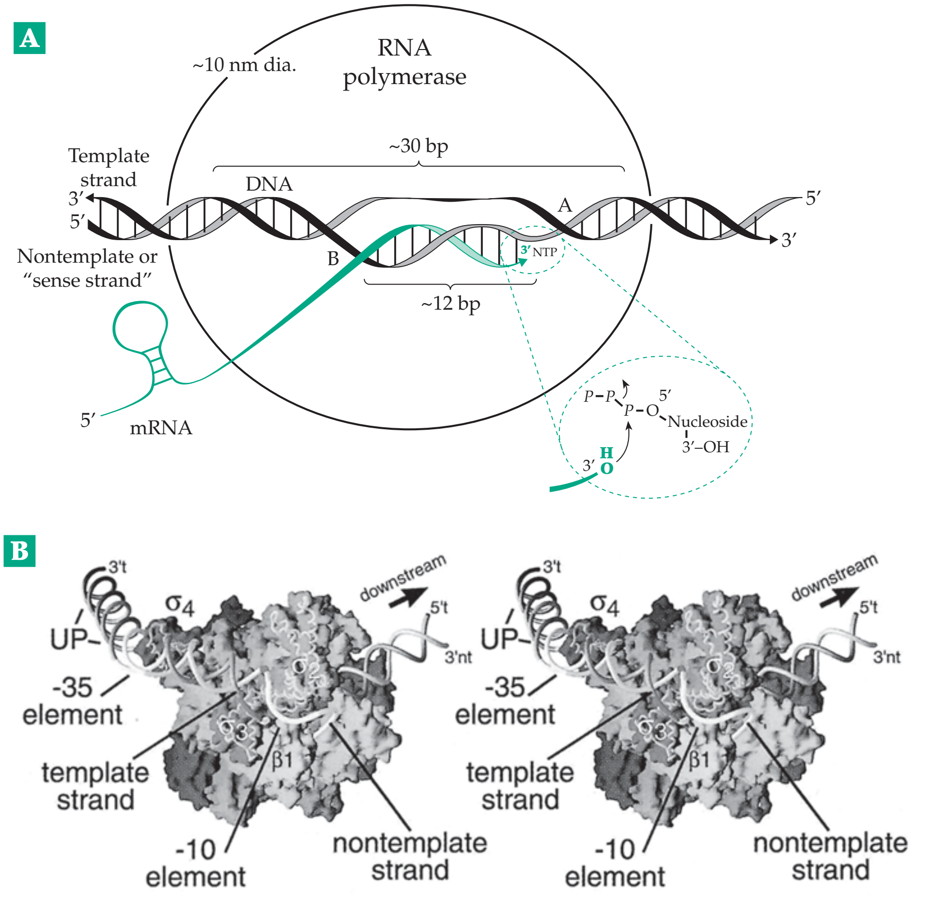


Figure 28-4 (A) Hypothetical structure of a "transcription bubble" formed by an RNA polymerase. Shown is a double-stranded length of DNA with the unwound bubble in the center. This contains a short DNA–RNA hybrid helix formed by the growing mRNA. The DNA double helix is undergoing separation at point A as is the hybrid helix at point B. NTP is the ribonucleotide triphosphate substrate. See Yager and von Hippel.[71] (B) Stereoscopic view of the structure of RNA polymerase from *Thermus aquaticus* in a complex with a promoter DNA. Included are the αI, αII, ω, β, β′, and σ subunits. However, the α C-terminal domains have been omitted. The template (t) strand passes through a tunnel, which is formed by the β and β′ subunits and two of the structural domains of the σ subunit. The nontemplate (nt) strand follows a different path. The position of the –10, –35, and UP elements of the DNA are marked. From Murakami *et al.*[33d] Courtesy of Seth A. Darst.

complex in which the first chemical steps in RNA formation occur.[69,70] These initial steps may involve rearrangements of subunit interactions and untwisting of DNA by torsional movements between subunits.[57]

Initiation of an RNA chain begins by reaction within the transcription bubble of either ATP or GTP with a second ribonucleotide triphosphate (Eq. 28-2) to form a dinucleotide still bearing a triphosphate at the 5' end. Further addition of nucleotide units at the 3' end by the same type

$$\text{ATP(GTP)} + \text{NTP} \rightarrow \text{PP}_i + ppp\text{Pu}p\text{N} \qquad (28\text{-}2)$$

of reaction leads to rapid transcription at a rate of ~50 nucleotides s^{-1} at 25°C. This is about one-thirtieth the rate of replication. The action of the RNA polymerase is apparently processive, a single molecule of the enzyme synthesizing the entire mRNA transcript.

3. Elongation of RNA Transcripts

After the newly initiated transcript has grown to 8–9 nucleotides, the sigma factor is lost from the RNA polymerase complex, the complex becomes very stable toward increased salt concentrations, and transcription proceeds processively in a **stably elongating mode**. As it does, the strands of the DNA duplex are pulled apart ahead of the polymerase and close up again behind the polymerase, the polymerase itself moving in a "transcription bubble" (Fig. 28-4). This is thought to lie within the ~10 nm diameter RNA polymerase complex and to encompass ~30 bp of DNA.[71] At the leading edge (point A in Fig. 28-4) a "separator" opens the DNA, which then closes behind the bubble. Behind the polymerase active site (next to substrate NTP in Fig. 28-4) the transcribed RNA forms a short hybrid helix with the DNA, presumably with an A type structure (Chapter 5). A severe topological problem is avoided if the transcribed RNA is separated at point B (Fig. 28-4) as the polymerase moves along the double helix. The driving force for the polymerization lies largely in the hydrolysis of the inorganic PP_i formed in the polymerization (Eq. 17-57). Stabilization of the RNA transcript by formation of loops and other secondary structure may also be a factor.

Although this picture seems clear and simple, many uncertainties remain. Transcription does not proceed evenly but by pauses and spurts. This has suggested the possibility of an "inchworm" type of movement of RNA polymerase.[72–75] However, the observations may also be explained by variations in the sequence. There are both pausing or stalling sites[76] and terminator sequences. The concentrations of the needed ribonucleotide triphosphate precursors will also affect the kinetics. In addition, defects in the DNA will be met. Transcriptionally linked repair (Box 27-A)[76a] may have to be called into play before transcription can continue.[77,78] Mispairing of bases can also occur in the growing RNA chain necessitating a pause during which the mismatch is recognized. Specialized proteins **GreA** and **GreB** participate in an editing step during which the RNAP backtracks for a few nucleotides, while a piece is hydrolyzed off from the 3' end before transcription can continue.[69,79,80]

If the transcription complex moves straight along a DNA double helix, separation of the strands will create positive supercoils (overwound DNA) in front of it and negative supercoils behind the bubble.[81] Experimental data support this prediction.[82] In *E. coli* the transcription of a plasmid generates positively supercoiled plasmid DNA when DNA gyrase (Chapter 27) is inhibited selectively.[83] A similar result was observed in yeast.[84] These and other data suggest that DNA gyrase may act to remove these positive supercoils, and that topoisomerase I may function in removing the negative supercoils generated behind the transcription bubble.

The electron micrograph in Fig. 28-5 shows RNA polymerase complexes apparently moving along a DNA strand with ribosomes assembled on the RNA and

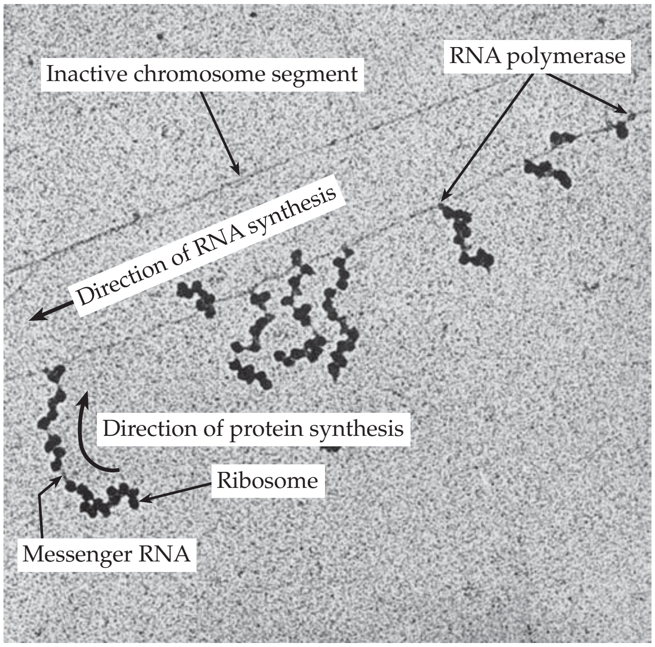

Inactive chromosome segment
RNA polymerase
Direction of RNA synthesis
Direction of protein synthesis
Ribosome
Messenger RNA

Figure 28-5 Electron micrograph showing transcription from an unidentified operon in *E. coli*. Note the DNA duplexes (horizontal) and the mRNA chains with ribosomes attached. The mRNA chains are shorter at the right side where transcription begins and larger to the left where transcription has proceeded for a longer time. From O. L. Miller, Jr.[85]

presumably synthesizing proteins. Actually, the DNA may be moving through the polymerase complexes.

4. Control of Various Bacterial Operons

With more than 2400 promoters an *E. coli* cell can be expected to utilize a great variety of control mechanisms. The best known of these involve the basic biosynthetic pathways and energy-yielding reactions. Here is a small sample.

Many repressors. Bacteria tend to keep most operons relatively inactive by use of repressors, most of which are proteins. Repressor proteins come in a variety of sizes and three-dimensional structures. Most are oligomers, and all have a DNA-binding motif, often at the N terminus but sometimes at the C terminus or elsewhere. The most studied and perhaps most abundant family[86] have the helix–turn–helix (HTH) motif. Within this family three-dimensional structures have been established for the *lac* repressors, the *trp* repressor (Fig. 5-35), the 66-residue *cro* repressor,[87–89] and a 92-residue DNA-binding "head piece" of the 236-residue lambda (cI) repressor.[90–91] The latter two proteins, both of which occur as symmetric dimers, are involved in preserving the lysogenic state of the λ prophage (Section B,1). The related 71-residue *cro* repressor from phage 434[92] and the head piece from the phage 434 repressor have similar structures.[93,94] All of these proteins have the N-terminal HTH structural motif, which binds to DNA and recognizes the appropriate sequence by interactions in the major groove of the DNA. Experiments in which the recognition helix of the 434 repressor was replaced with the corresponding helix from a phage P22 repressor supported the concept of an N-terminal HTH DNA-binding domain.[95,96] The hybrid repressor bound to the P22 operator rather than to that of phage 434. A similar HTH domain is present near the C terminus in the catabolite-activator protein (Fig. 28-6). The fit of the helix–turn–helix recognition corner varies from one protein to another. The small 53-residue **Arc** repressor of bacteriophage P22 of *Salmonella* has a very different interaction with DNA: the β sheet of an arc dimer fits into the major groove of the operator DNA making specific contacts.[97,98] A similar interaction characterizes the dimeric methionine repressor.[99]

Basic to the functions both of repressors and of activator proteins are allosterically induced conformational changes caused by the binding of inducers or corepressors. The changes begin at the binding sites of these small effector molecules but are transmitted to the DNA-binding heads. In several cases the conformational changes have been observed by X-ray crystallography and are seen to involve a movement of the recognition helices. This is also true for the *lacI* repressor (Fig. 28-3) and for the *TrpR* repressor (Fig. 5-35B) Notice that the tryptophan binds to the aporepressor immediately adjacent to the DNA-binding site, where it may control not only the shape but perhaps also the charge distribution within the recognition motif.

Another important factor in determining the strength of a repressor–operator interaction is the twist of the DNA or any other distortion of its regular helical structure. For example, in the center of the *cro* 434-operator complex the DNA is wound, while at the ends it is underwound.[92] Conformational changes in either the repressor or in the DNA or in both may be needed to provide optimal binding. As is seen in Fig. 28-3 the *lac* repressor causes a distinct bend in the DNA.

A theoretical possibility would be for the DNA in an operator site to be extruded as a cruciform structure like that in Fig. 5-34A. Such structures bind well to certain oligomeric proteins (e.g., see Fig. 27-26B), and they do appear to form in some promoters.[100] However, crystallographic structure determinations have ruled out cruciform structures for many repressors. A change from a linear helical duplex to cruciform structure would require a substantial unwinding of the helix and would mean that negatively superhelical DNA molecules would bind repressor much more tightly than does DNA without superhelical turns. Negative supercoiling does facilitate the *lac* operator–repressor interaction. However, there is only a 50–90° unwinding of the DNA.[101] Some AT-rich palindromes are readily converted to cruciform structures when negative supercoiling is increased. Placement of a promoter sequence within such a structure represses transcription unless the supercoiling is relaxed.[102]

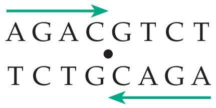
Binding of repressors and activator proteins to DNA control sequences is being studied in many ways. Among them are ^{1}H and ^{19}F NMR measurements on *lac* and *cro* repressors with specific tyrosine side chains replaced by deuterated tyrosine or 3-fluorotyrosine[103] or with 5-F-uracil replacing specific thymines in an operator sequence.[104] Footprinting shows that λ and *cro* repressors bind to only one side of the double helix, as is depicted in Fig. 5-50. Addition of dimethyl sulfate to growing *E. coli* causes methylation at many sites in the DNA and has been used to obtain footprints *in vivo*.[105] Not only are certain guanosines in a promoter protected from methylation when active transcription is occurring but also guanosines on the opposite side of the double helix become unusually reactive. This suggested that the DNA helix is bent in the transcriptional initiation complex and that the resulting distortion makes the bases on the outside of the bend more reactive.[105] There is also evidence that the repressor binds to the opposite face of the DNA that binds RNA polymerase in the –35 region of the promoter.[106]

Repressors may have similar recognition domains but may vary greatly both in size and in the functioning of their other domains, which may react both with small allosteric effectors and with other proteins. The repressor **BirA** of the *E. coli* biotin synthesis operon is an enzyme. The 321-residue protein activates biotin to form biotinyl 5'-adenylate and transfers the biotinyl group to proteins such as acetyl-CoA carboxylase[107–109a] and also represses transcription.

Inducible operons. The operon model as presented in Fig. 28-1 describes the negatively controlled inducible *lac* operon. There are many other examples of this type of control. Among them are control of the utilization of *N*-acetylglucosamine and xylose.[110] Ten genes for catabolism of inositol are encoded in two negatively controlled operons in *Bacillus subtilis*.[111] Similar controls are used by bacteria to protect against the antibiotic tetracycline. The **TetR** repressor controls a membrane associated protein TetA, which acts as an energy-dependent pump to remove tetracycline from the bacterial cell. Synthesis of TetA is normally repressed, but tetracycline, if present, acts as an inducer to cause TetR to dissociate from its operator site, allowing transcription of the transporter gene.[112–113a] The gene order, gene content, and regulatory mechanisms in an operon are often poorly conserved among related species of bacteria. However, functional and regulatory relationships may be maintained.[113b]

Feedback repression. A simple modification of the operon model accounts for **feedback repression** by end products of biosynthetic sequences (Fig. 11-1). The product, e.g., an amino acid, vitamin, purine, or pyrimidine base, acts as a corepressor binding to the aporepressor and causing an allosteric modification that inhibits transcription. It is not the aporepressor but the effector–repressor complex that binds to the operon and blocks transcription. An example is the tryptophan–repressor complex, whose 3-dimensional structure is shown in Fig. 5-35. If free tryptophan accumulates within a bacterial cell, it binds to the allosteric site in the *TrpR* repressor, inducing a conformational change that permits the tryptophan–repressor complex to bind to at least three 21-bp operator sites. Binding of the repressor–corepressor complex at the *aroH* operator locus represses the genes for the first three steps in the aromatic biosynthetic pathway (Fig. 25-1, step *a*). Binding at the *trpEDBCA* operator represses the genes for conversion of chorismate to tryptophan (Fig. 25-2), and binding at the *trpR* locus lowers the rate of synthesis of the trp repressor, providing a counteracting effect that may be important in controlling the growth rate. The *trp* operon is also well known for another type of control called attenuation (see Fig. 28-9 and associated text).

Feedback repression controls methionine synthesis in *E. coli*. The corepressor *S*-adenosylmethionine binds to an aporepressor that recognizes the following palindromic sequence in the DNA. This sequence occurs,



with minor variation, in front of four different operons encoding enzymes of methionine biosynthesis.[99,114] Together they form the methionine **regulon**. Other *E. coli* operons that are negatively regulated by operator–repressor interactions include those involved in biosynthesis of phenylalanine, tyrosine, arginine, threonine, and isoleucine. The *tyr* repressor modulates gene expression in at least eight operons, largely by repression.[115] The *ile* repressor acts on both the *ile* and *thr* operons.[116] The purine repressor, 341-residue PurR, belongs to the LacI family of repressors. Its C-terminal domain also has some sequence similarity to that of periplasmic sugar-binding proteins.[117] PurR represses several steps in the biosynthesis of IMP and in its conversion to GMP and AMP (Fig. 25-15).[118,119] In *E. coli* twelve biosynthetic genes, organized as nine transcriptional units, provide for the synthesis of arginine in eight enzymatic steps (Fig. 24-10).[120,121] As with the *trp* operons, synthesis of the repressor ArgR is autoregulated. The hexameric repressor has an N-terminal winged HTH DNA-binding domain.[121–123]

As is pointed out in Chapter 16, the acquisition of iron and control of its concentration is of crucial importance to bacteria. In *E. coli* the Fe^{2+}-binding protein **Fur** (ferric uptake regulator) represses promoters controlling siderophore biosynthesis as well as other responses. It is a **global regulator** that controls ~40 transcriptional units.[124] Similar proteins repress synthesis of the diphtheria toxin by *Corynebacterium diphtheriae*,[125,126] uptake of iron in these bacteria and in *Mycobacterium tuberculosis*,[127] and uptake of molybdate.[127a]

Positive control by activator proteins. Cyclic AMP in bacterial cells mediates the phenomenon of **catabolite repression**. This is the inhibition of the transcription of genes for enzymes needed in catabolism of lactose and other energy-yielding substrates, when the more efficient energy source glucose is present. Glucose causes a decrease in the concentration of cAMP by a complex mechanism[59,128] that may also cause a decrease in the concentration of inducer. When the glucose concentration decreases, the concentration of cAMP rises and stimulates the initiation of transcription in many operons. This is accomplished through the mediation of the 209-residue catabolite gene activator protein, CAP (also known as cyclic AMP receptor protein, CRP). This protein is a "global" regulator of gene expression that activates transcription at over 100 promoters in *E. coli*.[129] The CAP–cAMP complex binds to the *lac* promoter adjacent to the RNA polymerase site at the palindromic sequence in the DNA as is indicated in Fig. 28-2. The CAP molecule is a 45-kDa dimer of identical subunits, which resembles the repressors in having the HTH reading head that binds to the DNA.[130–132] However, the HTH motif is at the C terminus. The cAMP binds to two sites in each monomer (Fig. 28-6). Tightly bound cAMP molecules in an *anti* conformation bind in the center of the large N-terminal domains in the major allosteric sites. At higher cAMP concentration the second sites in the C-terminal DNA-binding domains are occupied.[131]

The CAP binds to DNA with the consensus sequence 5′-AAATGTGATCT/5′-AGATCACATTT, which may be located at variable distances from the promoter.[133] How does binding of the CAP–cAMP complex increase the rate of initiation of mRNA transcription? The anwer evidently lies in direct interaction between CAP and the N-terminal domain of the RNAP α subunit.[54d,129] Binding of CAP induces a 90° bend in the DNA, which may facilitate the protein–protein interaction and may lead to looping.[130,134]

The galactose (*gal*) operon of *E. coli* is negatively controlled and inducible by D-galactose or D-fucose, which bind to the *gal* repressor. There are two overlapping promoter sites, one of which is stimulated by adjacent binding of a CAP molecule.[135,136] A surprise came from the discovery that the operator was *upstream* from the promoter, that is, it comes before both the promoter and the structural genes to be transcribed. Later, a second operator sequence was found 90 bp away from the first and within the first structural gene. This suggested that the dimeric gal repressor binds the two operators to form a loop that blocks transcription.

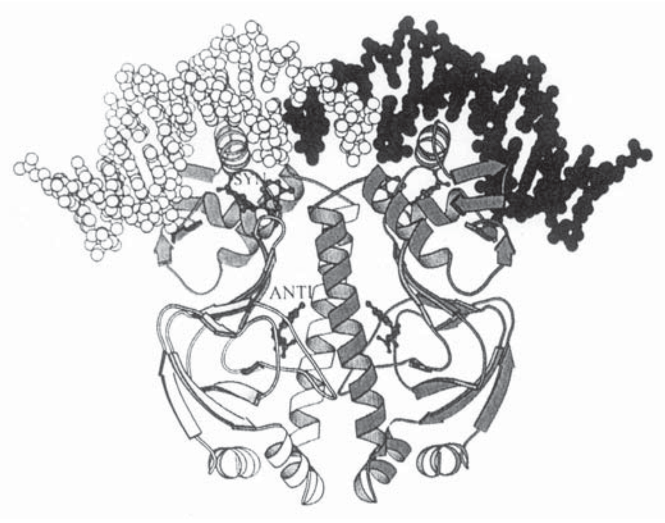


Figure 28-6 MolScript ribbon drawing of the CAP dimer bound to DNA with two molecules of the coactivator cAMP bound per monomer. A *syn*-cAMP molecule is bound to the HTH domain and a loop from the N-terminal domain, while the second *anti*-cAMP is bound more tightly in the center of the larger N-terminal domain. The DNA sequence for each half site is 5'-ATGTCACATTAATTGCGTTGCGC-3'. From Passner and Steitz.[131] Courtesy of Thomas A. Steitz.

Another example of positive regulation by CAP is provided by the seven proteins required for uptake of maltose and its catabolism by *E. coli*. These are encoded in two operons that are controlled as a single regulon. An apo-activator protein becomes an activator when it binds maltose.[137,138]

An *E. coli* protein known as **FRN** (for fumarate nitrate reduction) is a global transcription regulator homologous to CAP. It is active only under anaerobic conditions in which it controls more than 100 genes.[139–141] FRN contains an $[Fe_4S_4]^{2+}$ cluster, which is required for dimerization and binding to DNA. Exposure to O_2 converts the cluster into an $[Fe_2S_2]^{2+}$ cluster with loss of activator activity. The photosynthetic *Rhodospirillum rubrum* is able to adapt to growth on carbon monoxide as a carbon source. A CAP-like transcriptional activator **CooA** contains heme. It acts as a sensor for CO, which activates transcription, as does cyclic AMP with CAP.[142,142a] Sequence homologies suggest that several other bacterial activator proteins also have the HTH DNA-binding motif near their C termini.

There are other types of transcriptional activators in bacteria. One is transcription factor 1 (TF1) encoded by a *Bacillus subtilis* phage. It is a member of the protein HU family (Chapter 27). However, unlike the nonspecific HU it binds to some sites specifically and activates transcription.[143] The *E. coli* Ada protein is the acceptor protein in removal of methyl groups from DNA (Chapter 27). The same protein is an inducer of transcription of DNA repair enzymes in the large ada regulon. Methylation of Cys 69 of the Ada protein itself converts it into a gene activator.[144]

Control by looping. The arabinose utilization operon of *E. coli*, *araBAD*, encodes proteins needed for uptake of arabinose and conversion to D-xylulose 5-*P*. The repressor AraC in the absence of arabinose binds at operator 1 (O_1) to prevent further synthesis of repressor (autorepression) and also at the *araI* region to repress transcription of operon *araBAD*. The operator 2 (O_2) site, which is 211 bp upstream from *ara*I, is also needed for this repression.[145–147b] A loop is apparently formed by repressor binding (Fig. 28-7). Binding of arabinose to the repressor converts it into an activator, which stimulates initiation of transcription at the *BAD* promoter. Further stimulation is provided by the CAP-cAMP complex, which binds next at *araI*.

Looping is a recognized control mechanism for a number of other operons as well.[148]

Bacterial enhancers. Positive regulatory DNA sequences that are distant from the genes controlled are often called enhancers.[149,150] Their function is usually independent of position over a range of hundreds or more base pairs either upstream or downstream of the transcription initiation site. Quite common in eukaryotes (Section C,4), enhancers are less often found in bacteria. However, the binding sites for the **nitrogen regulatory protein C** (**NtrC** or Nr$_I$) of *E. coli* has the characteristics of an enhancer.[151,152] It functions with the rather complex glutamine synthetase (*glnALG*) operon in a major control point for nitrogen metabolism. The enzymology is illustrated in Fig. 24-7. When the supply of nitrogen from NH$_3$ is low the NtrC protein, a product of gene *glnG*, binds to the enhancer, which is located over 100 bp upstream and is thought to contact the σ^N (σ^{54}) subunit of the RNAP by formation of a loop.[153,153a] The process has been visualized by scanning force microscopy.[152] Another interesting aspect of this control system is the activation of Ntrc by phosphorylation of a specific aspartate side chain (Asp 54). The NtrC-*P* form is the active enhancer-binding protein. NtrC is a member of the family of two-component sensor-response regulator pairs, which frequently control bacterial metabolism and behavior (Chapter 11, Section C, 2; Fig. 19-5). The sensor protein is **NtrB** (NR$_{II}$), which is an autophosphorylating histidine kinase similar to the CheY protein of bacterial chemotaxis (Fig. 19-5).

NtrC-*P* dimerizes and binds to the enhancer sequence, where it appears to catalyze an ATP-dependent isomerization of the closed to open forms of the transcription initiation complex (Eq. 28-1).[153,154] The isomerization may depend upon looping.[152] Other operons that utilize the σ^N subunit of RNAP often also have upstream or downstream enhancers.[155,156]

Other two-component control systems. More than 60 different sensor–response regulator pairs have been discovered in bacteria. Many are associated with nutrition. For example, more than 30 genes of the phosphate (*pho*) regular are controlled by the sensor kinase **phoR**, which detects low phosphate ion concentrations and phosphorylates the response regulator **phoB**.[157–159] Protein **ArB** senses changes in environmental O$_2$ levels, and response regulator **ArcA** regulates ~30 operons in response.[159a] One effect is to activate ~30 genes needed for the conjugative transfer of DNA (Fig. 26-3).[160] Transcription of rhizobial *nif* and *fix* genes (Fig. 24-4) is controlled by the O$_2$ sensor **FixL** and its response regulator **FixJ**.[161] FixL is a heme protein whose kinase is active only when the heme is deoxygenated. The *E. coli* proteins discussed on p. 1075 mediate transcriptional responses to accumulation of superoxide anions or hydrogen peroxide.[161a]

In the absence of O$_2$ the *E. coli* FNR protein induces proteins of the anaerobic respiration pathways. Nitrate also has its own two-component system that senses nitrate availability and activates transcription of enzymes catalyzing nitrate respiration.[162] An expanded two-component system induces sporulation in *Bacillus subtilis* in response to poor growth conditions.[163,164] The crystal structure of one of two response regulators (**Spo0F**) has a structure closely related to that of CheY and the nitrate response regulator NarL.

Antisense RNA. Another mechanism of control of either transcription or of plasmid replication involves small molecules of RNA that are transcribed from the opposite strand than the template strand used for mRNA synthesis.[1,1a,165–166b] These **antisense** RNA molecules have at least some part of their sequence complementary to that of the mRNA and to the corresponding sequence in DNA. A well-studied example is control of the copy number of the colicin E1 and other plasmids of *E. coli*.[167–169] Two transcripts, RNAI and RNAII, are initiated upstream from the origin of replication (Fig. 28-8). RNA II is a 555-nucleotide primer of replication. It is synthesized as a longer transcript that is cut by RNase H at *ori*. This

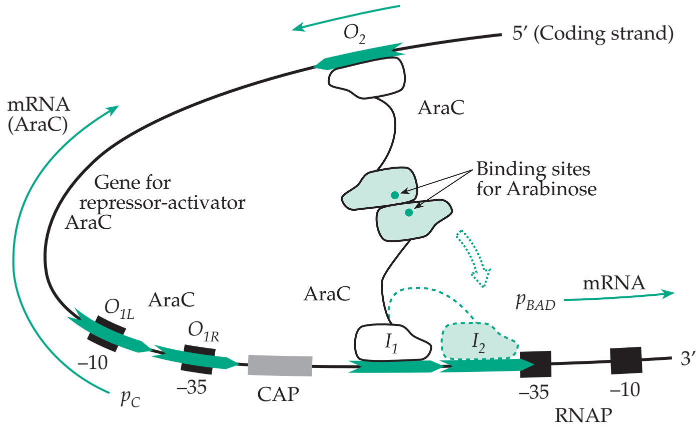


Figure 28-7 Scheme showing regulatory region of the *araCBAD* operon of *E. coli*. In the absence of arabinose the protein AraC acts as a repressor, which binds to upstream region I of the promoter and also to operator O_2, forming a loop of ~210 bp of DNA. Binding of arabinose to the N-terminal domains of the dimeric repressor allows the dimer to dissociate and the N-terminal domain of one subunit to bind to region I$_2$ of the promoter for genes *B,A,D* (dashed outline) activating their transcription. The AraC protein also binds to promoter P_c, repressing its own synthesis. After Zhang *et al*.[146]

cleavage is inhibited if the 108 nucleotide RNA I forms a duplex with the 5' end of the RNA II, which has a complementary base sequence. The process is more complex than this because both RNA II and RNA I have complex secondary structures and are brought together with the help of the small protein product of gene *gro*, which permits them to recognize each other prior to duplex formation.

The major outer membrane porins of *E. coli* (Fig. 8-20) are encoded by genes *ompC* and *ompF*. A small 174-nucleotide RNA called **mRNA-interfering complementary RNA** (mic RNA), whose gene is upstream of the *ompC* promoter, is transcribed in the opposite direction from that of the porin gene.[167] It is homologous with the 5' end of *ompF* mRNA, and its function is evidently to inhibit translation of the *ompF* mRNA. Since *ompC* and the *micF*RNA gene are apparently regulated coordinately, synthesis of the *ompF* product is inhibited if the *omp*C product is being synthesized, as happens when the bacteria are growing in a medium of high osmotic strength.[166b,170,171]

Many other examples of regulation by antisense RNA are being discovered. Small noncoding RNA molecules (**ncRNAs**) often serve as templates or **guides**[1,171a] in processes ranging from synthesis of telomere ends (Fig. 27-21) to editing,[172] modifying,[173] and splicing mRNA[174] (Section D). Both in bacteria[174a] and in eukaryotes (p. 1640) dsRNAs are often formed and subsequently cleaved to single-stranded antisense RNAs that act as guides to initiate hydrolytic destruction of defective, toxic, or unwanted mRNAs.[171a] The recognition of these natural regulatory mechanisms has led to keen interest in artifical regulation by antisense RNA. Synthetic antisense RNA injected into cells will inhibit expression of selected genes. There is the possibility of effective therapy against viruses, cancer, and inflammation if suitable antisense RNA could be generated within eukaryotic cells or introduced as drugs.[175–176b] Such drugs, which are typically ~15 nucleotides in length, are most satisfactory if they are stable, enter cells, and interact specifically with complementary sequences of DNA or RNA.[177] Stability can be improved by use of linkages other than the natural phosphodiester.[178–181] Phosphorothioates, in which sulfur replaces one nonbridging oxygen atom on phosphorus, is a favorite. Synthetic antisense oligonucleotide mimics may cause adverse reactions with proteins within cells.[178,182] Nevertheless, future successes seem likely.[182a] See also RNA interference (p. 1640).

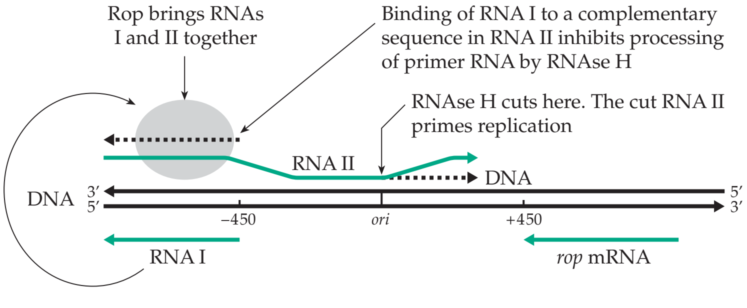

Rop brings RNAs I and II together
Binding of RNA I to a complementary sequence in RNA II inhibits processing of primer RNA by RNAse H
RNAse H cuts here. The cut RNA II primes replication
RNA II
DNA
DNA
3′
5′
5′
3′
−450
ori
+450
RNA I
rop mRNA

Figure 28-8 Simplified scheme for control of replication of the ColE1 type plasmid by antisense RNA. The primer for DNA synthesis is RNA II whose initial transcript extends past the replication *ori*. It is cut by RNase H at *ori* and then primes replication of the upper strand as shown in the figure. The antisense RNA is RNA I. It binds to protein Rop whose gene location is also indicated in the figure. Rop assists RNA I and RNA II in undergoing a complementary interaction. However, both RNAs apparently maintain a folded tertiary structure, and only some segments interact. The interaction with the Rop protein evidently in some way prevents initiation of replication until the Rop concentration falls because of replication of the host cell.[167,168]

5. Termination of Transcription in Bacteria

Encoded in DNA are not only the initiation signals for transcription but also termination signals or **terminators**.[183] Some of these are sensed by the bacterial RNA polymerase itself while the “reading” of others requires auxiliary proteins. Terminators can be either constitutive or regulatable. The simplest terminators result from GC-rich regions of dyad symmetry in the DNA. The RNA transcript is able to form a stable hairpin loop, possibly within the transcription bubble. If such a loop is followed closely by a series of uracils, the RNA and the polymerase will dissociate from the DNA template terminating transcription. The low stability of AU base pairs may facilitate dissociation, but RNA polymerase may also recognize the terminator loop. Sometimes a terminator will have a series of adenines preceding the loop. This is often a bidirectional terminator; the transcript from the other strand of DNA will have a loop followed by a series of U’s. There are many more complex termination mechanisms.

Attenuation. A major mechanism of feedback repression, known as attenuation, depends not upon a repressor protein but upon control of premature termination. It was first worked out in detail by Yanofsky *et al*. for the *trp* operon of *E. coli* and related bacteria.[184–186] Accumulation of tryptophan in the cell represses the *trp* biosynthetic operon by the action of accumulating tryptophanyl-tRNA^{Trp}, which specifically induces termination in the *trp* operon. Other specific “charged” aminoacyl-tRNA molecules induce termination at other amino acid synthesis operons. The first structural gene in the *trp* operon, *trp*E, is preceded at the 5' end by a 162-bp **leader sequence,** which is transcribed into mRNA. Within this RNA are two adjacent hairpin loops. The second of these loops (labeled **3:4** in Fig. 28-9) has a GC-rich stem and is followed by eight consecutive U's. It is a typical efficient terminator. An RNA polymerase, having just passed this sequence, will interact with the looped RNA formed behind it and will dissociate from the DNA to terminate transcription. However, if the terminator loop is prevented from forming, transcription will continue, and the structural genes of the operon will be expressed.

The *trp* operon contains a short gene for a **leader peptide** preceding the terminator. Its RNA transcript is shaded, and its initiation codon AUG and termination codon AAU are boxed in Fig. 28-9. In bacteria translation of mRNA begins while transcription is still in progress (Fig. 28-5). The 5' end of the mRNA may enter a ribosome before the RNA polymerase reaches the terminator loop, after which the leader peptide will be synthesized. The ninth and tenth codons of the leader peptide gene are of special importance. They lie at the beginning of the **1:2** or “protector” loop (Fig. 28-9) and code for tryptophan. If the level of tryptophan is high, tryptophanyl-tRNA^{Trp} will be formed and the mRNA will move rapidly through the ribosome, these tryptophan codons will pass through, and tryptophan will be incorporated into the peptide. The 1:2 loop will be opened, but the terminator loop will remain intact. The result will be termination of transcription. However, if the tryptophan concentration is low, there will be a shortage of charged tRNA^{Trp}, and peptide synthesis will be “stalled” with these tryptophan codons in the active sites of the ribosome. This will allow time for the attentuator region to assume the alternative secondary structure shown in Fig. 28-9B. Here the **1** limb of the protector loop is stalled in the ribosome allowing the **2** limb to form the altenative **2:3** loop. The terminator has been destroyed, and transcription continues through the rest of the operon. Thus a low level of tryptophan (and of tryptophanyl-tRNA) favors transcription of the *trp* synthetic operon.

Cells of *Bacillus subtilis* also synthesize a *trp* operon transcript that can form either an antiterminator or a terminator loop (Fig. 28-9C).[187,187a–c] Tryptophan, when present in a high enough concentration, binds to a *trp* RNA-binding attenuation protein (*TRAP*). This is an 11-subunit protein, which has 11 tryptophan-binding pockets and also 11 binding sites for GAG or UAG RNA triplets. When tryptophan accumulates within the cell it binds to TRAP, which then wraps ~53 residues of RNA transcript containing 11 GAG or UAG triplets around its perimeter (Fig. 28-9D). This prevents formation of the antiterminator loop but allows the terminator loop to form. At low tryptophan concentrations the antiterminator loop is formed, preventing formation of the terminator loop.[188]

Attenuation is also an important mechanism of control of transcription of biosynthetic operons for histidine, phenylalanine, leucine, isoleucine, and threonine.[189] Like the *trp* attenuator region, attenuators for these operons contain codons for the amino acid whose synthesis is being regulated: seven Phe codons in the *phe* attenuator, seven His codons in the *his* attenuator, four Leu codons in the *leu* attenuator. The *thr* operon, which is sensitive to both threonine and isoleucine, has eight Thr and four Ile codons, while the *ilv* attenuator has four Leu, five Ile, and six Val codons permitting feedback repressor by three kinds of charged tRNA.[190] The pyrimidine synthesis operon *pyr* has three attenuator sequences, one right after the promoter and two others, one just before each of the two genes in the operon. This permits partially independent control of the two genes.[191]

Rho and other termination factors. Termination proteins can also react with specific regions of DNA or of an RNA transcript to terminate transcription.[183] The best known termination factor is the rho protein; a hexamer of 45-kDa subunits. It interacts with transcripts at specific termination sequences, which are often C-rich, and in a process accompanied by hydrolysis of ATP causes release of both RNA and the polymerase from the DNA.[192,193] Additional *E. coli* proteins, products of genes ***nus A*** and *nus G*, cooperate with the rho factor at some termination sequences.[194–196c] The rho hexamer is a helicase that moves along the RNA transcript in the 5' → 3' direction driven by ATP hydrolysis. If it locates an appropriate termination signal, it may utilize its helicase activity to uncoil the DNA–RNA hybrid segment within the transcription bubble (Fig. 28-4).[197–198b]

Cells also contain **antitermination proteins,** which prevent termination of transcription of rRNA or tRNA genes at the many loops of secondary structure that are possible with these transcripts.[59,199–200b] These antitermination factors are also important in regulating transcription during the lytic phase of growth of phage λ (see Section B,1). Also important are rates of hydrolytic degradation of mRNA molecules.[201]

6. Effects of Antibiotics

The antibiotic **rifamycin** (Box 28-A) appears to interfere with initiation by competing for the binding of the initial purine nucleoside 5'-triphosphate. The same bacterial RNAP that synthesizes mRNA also transcribes both rRNA and the tRNAs. Thus, the synthesis of all forms of RNA is inhibited by rifamycin. When a population of bacteria is subjected to this antibiotic, a few individuals survive. These rifamycin-resistant mutants are no longer sensitive to the antibiotic. Among them are some mutants that produce an

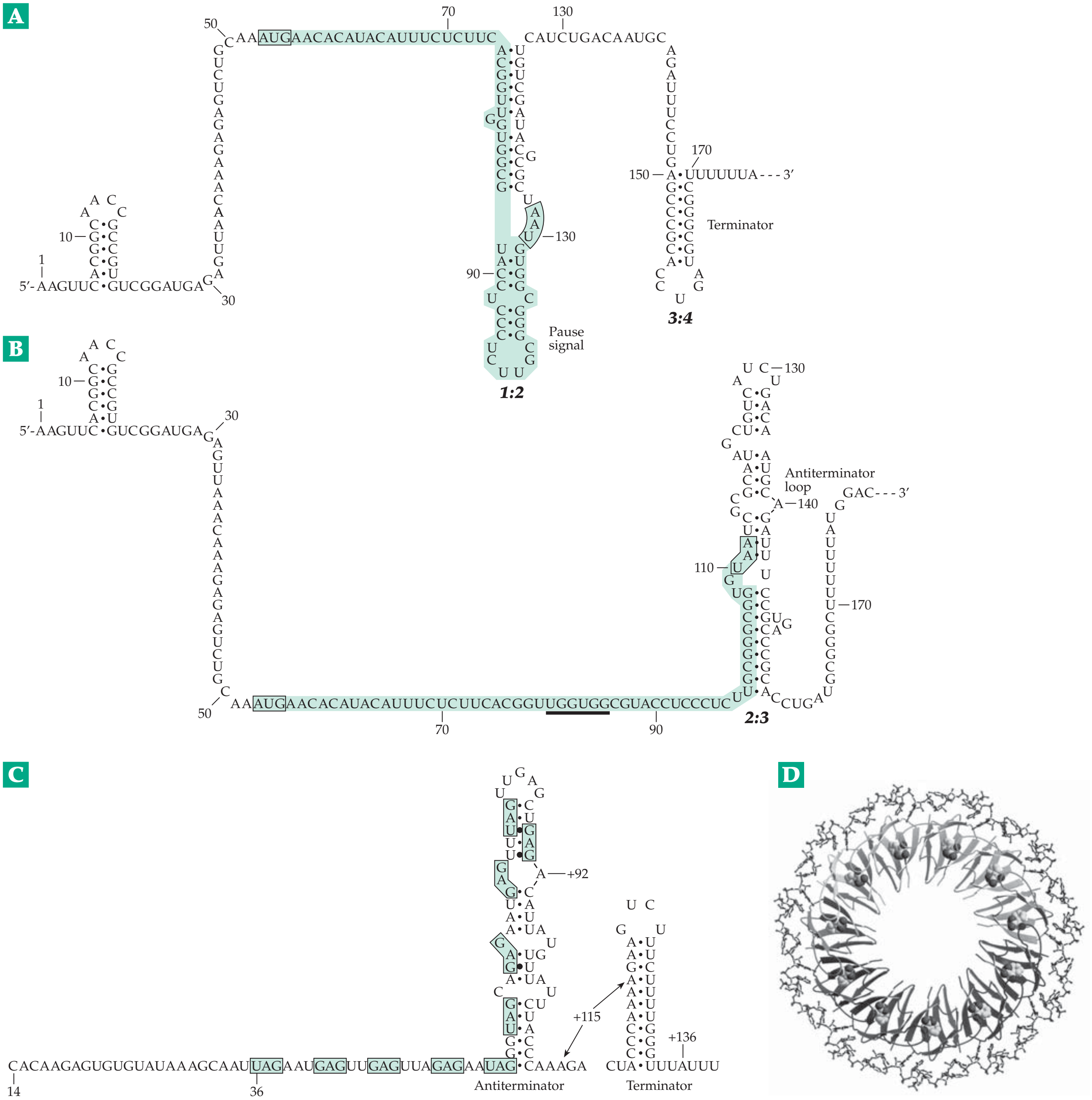

A
B
C
D
5′-AAGUUC•GUCGGAUGA
AAAUGAACACAUACAUUUCUCUUC
CAUCUGACAAUGC
A•UUUUUUA---3′
Pause signal
1:2
Terminator
3:4
Antiterminator loop
GAC---3′
2:3
AAAUGAACACAUACAUUUCUCUUCACGGUUGGUGGCGUACCUCCCUC
CACAAGAGUGUGUAUAAAGCAAUUAGAAUGAGUUGAGUUAGAGAAUAG•CAAAGA
CUA•UUUAUUU
Antiterminator
Terminator
1
10
30
50
70
90
110
130
140
150
170
14
36
+92
+115
+136

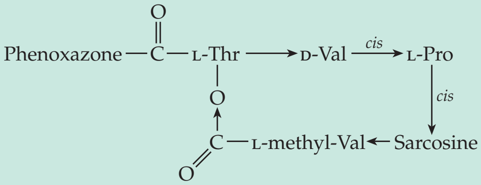
Figure 28-9 (A,B) Alternative leader RNA structures that mediate control of attenuation in the *trp* operon of *Serratia marcescens*. The leader peptide initiation and termination codons are *boxed*. Tryptophan codons are indicated by *underlining* in *B*. Formation of the 5' hairpin structure is predicted by computer analysis but is not implicated in control by attenuation. A structure **1:2** is thought to serve as a transcription pause signal. Structure **3:4** is predicted to form when there is an adequate supply of charged $tRNA^{Trp}$ and is thought to function as the transcription termination signal recognized by RNA polymerase. The **2:3** structure, or **antiterminator loop,** is predicted to form when charged $tRNA^{Trp}$ is unavailable. Its formation presumably precludes formation of structure **3:4**, thereby allowing RNA polymerase to continue transcription into the structural genes of the operon. From Kuroda and Yanofsky.[186] (C) Antiterminator and terminator loops, one of which may form from the leader sequence of the *Bacillus subtilis trp* operon mRNA. Numbering refers to the start point of transcription. The triplet repeats involved in attenuation are shaded. From Baumann *et al.*[187] (D) Structure of the 11-subunit tryptophan RNA-binding attenuation protein (TRAP) as a ribbon diagram with 11 molecules of L-tryptophan shown as van der Waals spheres. The apparently circular RNA structure reflects the fact that the gap between the beginning and end of the bound RNA segment is averaged (randomized) over eleven binding sites in the crystal structure. The 53-residue RNA containing 11 triplet repeats of GAG or UAG is bound around the perimeter and is shown as a ball-and stick model. From Antson *et al.*[188]

BOX 28-A THE ANTIBIOTICS RIFAMYCIN, RIFAMPICIN, AND ACTINOMYCIN D

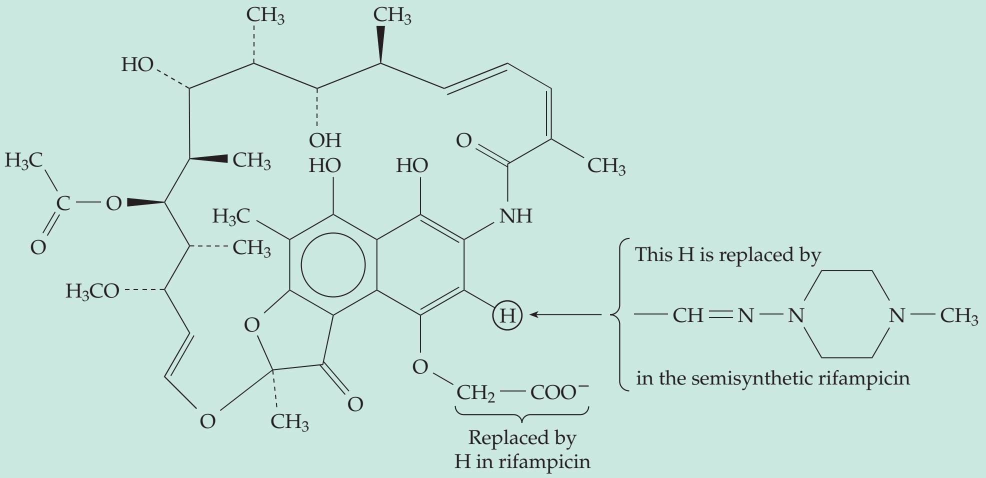
Rifamycin, produced by *Streptomyces mediterranei*, is of medical value because it affects **acid-fast** as well as gram-positive bacteria. The semisynthetic rifampicin has been especially useful in the treatment of tuberculosis. The ether linkage at the bottom of the ring at the right in the structural formula is cleaved, and the resulting hydroquinone is oxidized to a quinone within the bacteria.[a] At a concentration of 2 x 10^{-8} M rifampicin inhibits bacterial RNA polymerase 50%. It does not prevent the binding of polymerase to DNA but inhibits initiation of transcription. Mutants of *E. coli* resistant to rifampicin produce RNA polymerase whose β subunit has been altered, sometimes with a change in electrophoretic mobility. The related antibiotic **streptolydigin** also binds to the subunit of RNA polymerase and blocks elongation, resistant mutants mapping close to *rif* mutants.

The actinomycins, which are also produced by *Streptomyces*, not only kill bacteria but also are among the most potent antitumor agents known.[b] However, because of their extreme toxicity they are of little medicinal value. Actinomycin D, which is a specific inhibitor of RNA polymerase, contains a planar phenoxazone chromophore bearing two carboxyl groups, each one linked to an identical cyclic peptide made up of L-threonine, D-valine, L-proline, sarcosine (*N*-methylglycine), and L-methylvaline.

An ester linkage joins the methylvaline residue of the peptide to the side-chain hydroxyl of threonine. Two *cis* peptide linkages are present. Ignoring the obvious asymmetry of the phenoxazine ring, actinomycin possesses approximate twofold symmetry.



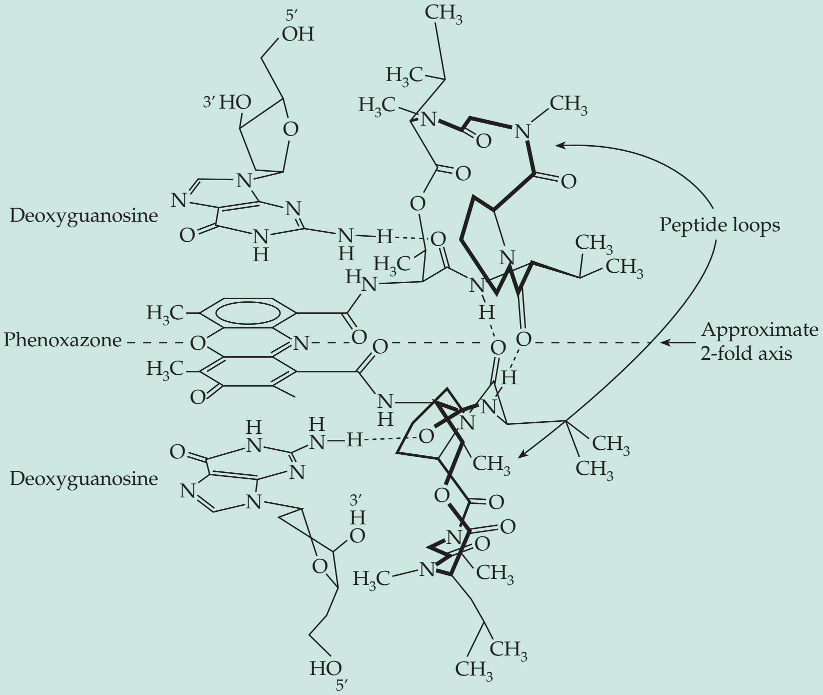
The antibiotic binds tightly to double-stranded DNA in regions containing guanine. A 2:1 deoxyguanosine–actinomycin complex has been crystallized, and the structure has been determined by X-ray diffraction.[c–e] The phenoxazine ring lies at the center of the complex, one peptide loop extending above it. The twofold symmetry is present in the dideoxyguanosine complex as well as in actinomycin itself. The phenoxazine ring lies between the two flat guanosine rings in van der Waals contact. The two amino groups of the guanine rings form strong hydrogen bonds with the carbonyl groups of the threonine residues. There are also weaker nonlinear hydrogen bonds from the N-3 atoms of the guanines to the NH groups of the same threonines.



RNAP with an altered β subunit. Since the mutant polymerases do not bind rifamycin, it was concluded that rifamycin binds to the β subunit and that the rifamycin-resistance gene *rpo*B or *rif* (which maps at 89 min) is the gene for the β subunits of RNA polymerase.

Streptolydigin inhibits both initiation and elongation. **Actinomycin D**, by binding to DNA, inhibits DNA polymerases as well as RNA polymerases, the latter at a concentration of only 10^{-6} M (Box 28-A). The eukaryotic RNA polymerases are not inhibited by rifamycin, but RNA polymerases II and III are completely inhibited by the mushroom poison **α-amanitin** (see Box 28-B). Inhibitors of DNA gyrase (Chapter 27) also interfere with transcription as do chain terminators such as cordycepin (3'-deoxyadenosine) and related nucleosides.

BOX 28-A (continued)



A symmetric pair of hydrogen bonds join the two carbonyl and NH groups of the D-valine residues in the peptide loops.

Model building studies show that a similar complex can be formed with double-stranded DNA.[f] While the amino groups of the guanine rings (see drawing) are hydrogen-bonded to the actinomycin, the other hydrogen atoms of the same amino groups as well as the N-1 hydrogen atoms and the carbonyl groups of the guanine ring are available for hydrogen bonding to form a GC base pair. Thus, the structure above can be modified readily into part of the double-stranded DNA molecule in which the phenoxazone ring of actinomycin is intercalated between two CG pairs. To do this the normal DNA structure has to be unwound by 18° at the point of insertion of the extra ring. Binding also occurs at other sites.[f] Sobell suggested that actinomycin binds to a premelted region of the DNA helix within the transcription bubble and immobilizes it. This interferes with the elongation of growing RNA chains.[g]

---

[a] Goldberg, I. H., and Friedman, P. A. (1971) *Ann. Rev. Biochem.* **40**, 775–810

[b] Perlman, D. (1970) in *Medicinal Chemistry*, 3rd ed. (Burger, A., ed), pp. 309–316, Wiley (Interscience), New York (Part 1)

[c] Sobell, H. M. (1973) *Prog. Nucleic Acid Res. Mol. Biol.* **13**, 153–190

[d] Sobell, H. M., Jain, S. C., Sakore, T. D., and Ponticello, G. (1971) *Cold Spring Harbor Symp. Quant. Biol.* **36**, 263–270

[e] Sobell, H. M. (1974) *Sci. Am.* **231**(Aug), 82–91

[f] Robinson, H., Gao, Y.-G., Yang, X.-l, Sanishvili, R., Joachimiak, A., and Wang, A. H.-J. (2001) *Biochemistry* **40**, 5587–5592

[g] Sobell, H. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5328–5331

7. Processing of Bacterial RNA

Newly formed rRNA and tRNA molecules are usually not functional but must undergo chain cleavage, methylation, and other alterations before they are "mature." However, most bacterial mRNA does not require processing. Bacteria produce a series of mRNA molecules of variable length, some corresponding to polycistronic (polygenic) and some to monocistronic operons. Most of the mRNA molecules produced are unstable with an average lifetime of about two min; however, some, such as those produced in bacteria about to undergo sporulation, survive much longer. Bacterial mRNA sometimes does undergo processing before it reaches the ribosomes. For example, following infection of *E. coli* cells by phage T7, ribonuclease RNase III cleaves the large 7000 nucleotide "early RNA" transcript from the virus DNA into five defined fragments.[202] Each fragment presumably carries the message for a single viral gene. Genes containing introns have been identified in several archaebacteria[203–205] and in certain phage.[206] The corresponding transcripts must be spliced, as are most eukaryotic transcripts.

Ribosomal RNA. Quantitatively the most im portant RNA, making up 90% of that present in cells, is ribosomal RNA. Synthesis of rRNA must be rapid, for an *E. coli* cell produces 5–10 new ribosomes per second, or 2 x 10^4 molecules of RNA per generation. Bacterial ribosomes contain three pieces of RNA. These are designated, according to their sedimentation constants, as 5S, 16S, and 23S and contain about 120, 1700, and 3300 nucleotides, respectively. All three pieces appear in cells as parts of larger **pre-rRNA** precursor molecules with extra nucleotide sequences at both the 3' and 5' ends.[207,208]

There are seven rRNA regions in the *E. coli* chromosome.[208a] Each region consists of a single transcriptional unit containing a gene each for 16S, 23S, and 5S rRNA with interspersed tRNA genes as follows: 16S, tRNA, 23S, tRNA, 5S, tRNA. A single transcript (which can, in certain mutant strains, appear as a 30S molecule) is cut by the endonuclease RNAse III into the smaller pre-rRNA molecules.[209,210] Other nucleases trim these to their final sizes, and methylases act to modify 24 residues in *E. coli* rRNAs.[208,211–213] Most RNAs of all organisms contain **pseudouridine** (ψ), which is formed by isomerization of specific uridines present in the RNAs (Eq. 28-3). In *E. coli* there is one pseudouridine in the 16S ribosome RNA and nine in the 23S RNA as well as one or more in most tRNAs.[214–217] The isomerization depends upon a carboxylate group of the enzyme, which evidently adds to the 6-position of the uracil ring to form a pivot around which the ring can rotate after it is eliminated from its attachment to the RNA and before it is reattached with a C–N linkage (Eq. 28-3).[217] Both bacteria and eukaryotes contain several pseudouridine synthases, which act to isomerize specific uridine residues in the precursor RNAs.[218,218a,b] In eukaryotes special **guide RNAs** direct the pseudouridine synthases to specific locations in their substrates.[174,219] The same thing is true for 2'-*O*-methylases that modify selected ribose rings in precursor RNAs.[219,220]

Transfer RNA. The genes for tRNA molecules in both bacteria and mammalian cells are grouped in clusters, which are transcribed as large precursors sometimes containing more than one kind of tRNA or containing tRNA fused to rRNA or mRNA sequences. At least three different nucleases are needed for cutting and trimming to form the mature tRNA. These enzymes may not always act in the same sequence. Thus, for some but not all tRNAs cleavage near the 3' end is needed before cleavage can take place at the 5' end.

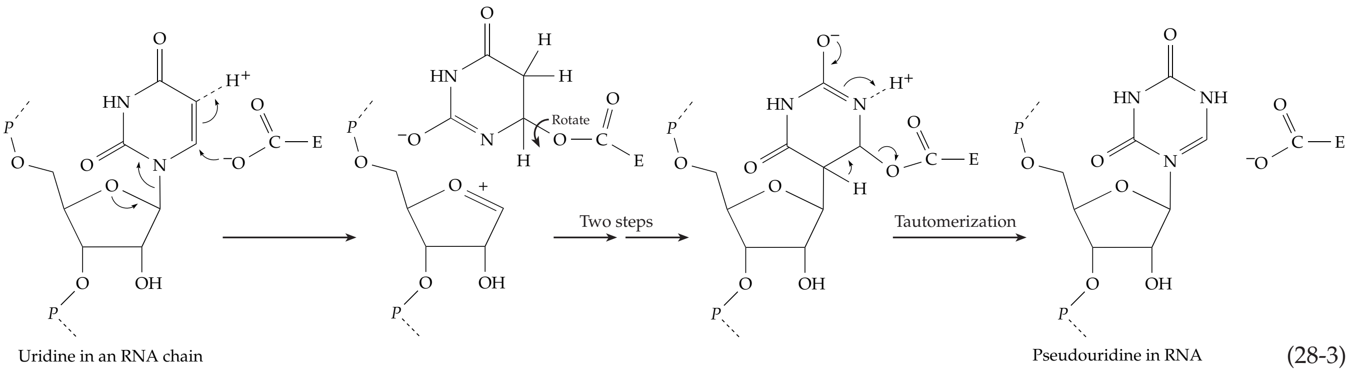
The best known processing nuclease is **RNase P,** which cleaves bacterial tRNA precursors to create the 5' ends of the mature tRNAs. All of the 64 tRNA precursors present in *E. coli* are cleaved by this unusual enzyme,[221–222c] which contains an essential piece of RNA (Chapter 12, Section D,6). Cleavage of polycistronic tRNA precursors by RNase P or of the previously mentioned rRNA and tRNA pecursor by RNAses P and III releases mature tRNAs carrying extra nucleotides at the 3' ends are removed by exoribonucleases, a surprising number of which are present in cells. In *E. coli* they include RNases II, D, BN, and T,[223,224] as well as polynucleotide phosphorylase.[225]

The structure of the precursor to the minor *E. coli* tyrosine $tRNA_1$ is shown in Fig. 28-10. This is encoded by the *amber* suppressor gene *Sup*F (see Chapter 29). Transcription of its gene is initiated by GTP 43 bp upstream of the 5' end of the mature tRNA and usually terminates at a ρ-dependent signal 225 bp beyond the CCA terminus of the tRNA. An endonuclease cuts the transcript a few nucleotides beyond the CCA end. It is then trimmed to an ~130-nucleotide piece still containing 2–3 extra nucleotides at the 3' end. This intermediate is cut by RNase P at the 5' end after which final trimming is done at the 3' end (Fig. 28-10).[226,227]

An important chemical achievement was the synthesis by H. G. Khorana and associates of the double-stranded DNA segment coding for this *E. coli* tyrosine tRNA[228] and its precursor.[229] This was one of the first synthetic genes (Chapter 5). The synthesis was extended to include the gene termination region, which lies beyond the CCA end of the tRNA. Two noteworthy features appeared. There is a local center of dyad symmetry (indicated by vertical bars and a central dot in Fig. 28-10), which may serve as a termination signal. The operator is located in the 29-nucleotide sequence *preceding* the tyrosine tRNA gene.[230–232]

The 3'-terminal group of three nucleotides, CCA, is invariant among all tRNA molecules and is labile, undergoing active removal and resynthesis. The rate of this turnover is sufficient to involve about 20% of the tRNA molecules of a cell per generation, but it is very much slower than the rate of participation of the tRNA molecules in protein synthesis. The physiological significance of end turnover is unknown.[233] While this CCA sequence is encoded in bacterial tRNA genes, it is added in a separate reaction in eukaryotes.[234]

In addition to the cutting and trimming of precursors by nucleases, extensive modification of purine and pyrimidine bases is required to generate mature tRNAs.[235] Some of these modification reactions are



(28-3)

discussed in Chapter 5 (see Fig. 5-33). In Fig. 28-10 the modifications in the mature tyrosine suppressor tRNA are shown, and in Fig. 5-30 those in *E. coli* phenylalanine tRNA are indicated. Modification usually begins with tRNA precursors. For example, the precursor in Fig. 28-10 is methylated to form ribothymidine at position 63.[236] Pseudouridine is then introduced at positions 40 and 64 by isomerization of the uridines present in the initial transcript (Eq. 28-3).[237] The Tψ pair at position 63 and 64 of Fig. 28-10 is almost universally found in tRNA,[238] but the positions are usually designated as 54 and 55 as in Fig. 5-30. Position 8 in most tRNAs is occupied by 4-thiouridine. The sulfur atom is transferred from cysteine as S^0 using a PLP-dependent mechanism similar to that in Eq. 14-34 and involving an ***S*-sulfonylcysteine (persulfide) intermediate**.[239,240] Some modifications are completed on the mature tRNA.[241] Some tRNAs require RNA splicing for maturation (see Section D,5).[205]

The hypermodified nucleoside **queuosine** is found in the first (wobble) position of anticodons of most eukaryotic tRNAs for Asn, Asp, His, and Tyr and also in most bacteria.[242–244] Bacteria apparently make aminomethyl-7-deazaguanine from guanine (Eq. 28-4, step *a*) and transfer this compound into the appropriate position in tRNA (step *b*) by a tRNA-guanine transglycosylase.[245–246a] The incorporated nucleoside is then converted to queuine by incorporation of the 5-carbons of the ribosyl group in *S*-adenosylmethionine (Eq. 28-4, step *c*) to form an intermediate epoxide.[247] This is converted to queuine in the vitamin B_{12}-dependent step *d*. Eukaryotes are unable to form queuine and must obtain it as a nutrient factor or from intestinal flora.[244] It is exchanged into tRNA by action of the transglycosylase.[245] Queuine might be considered a

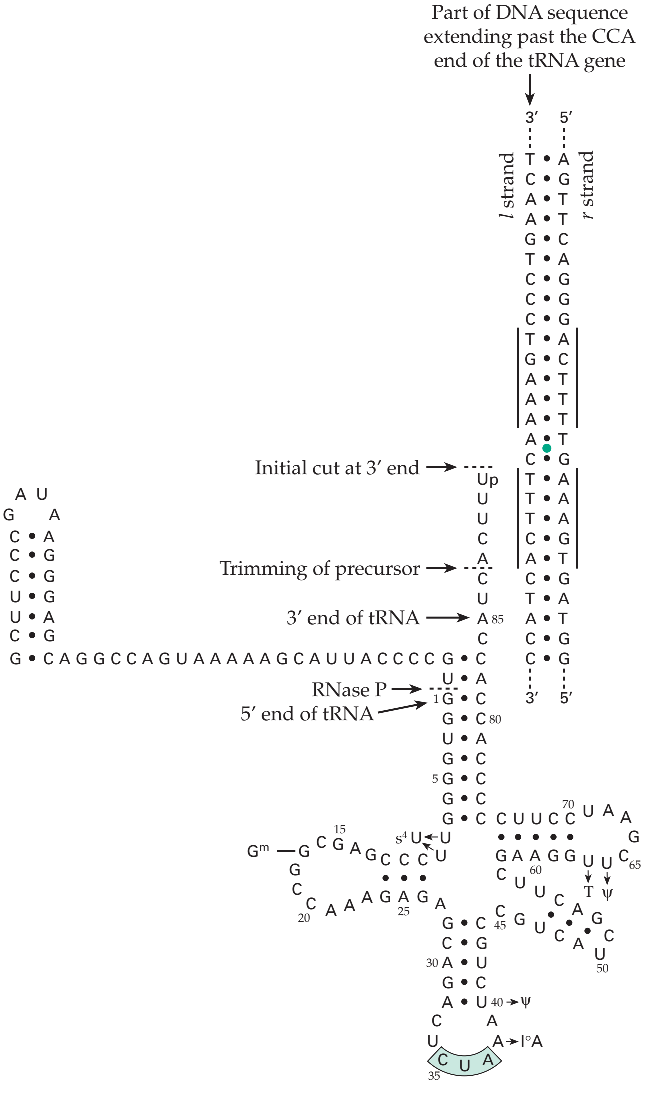
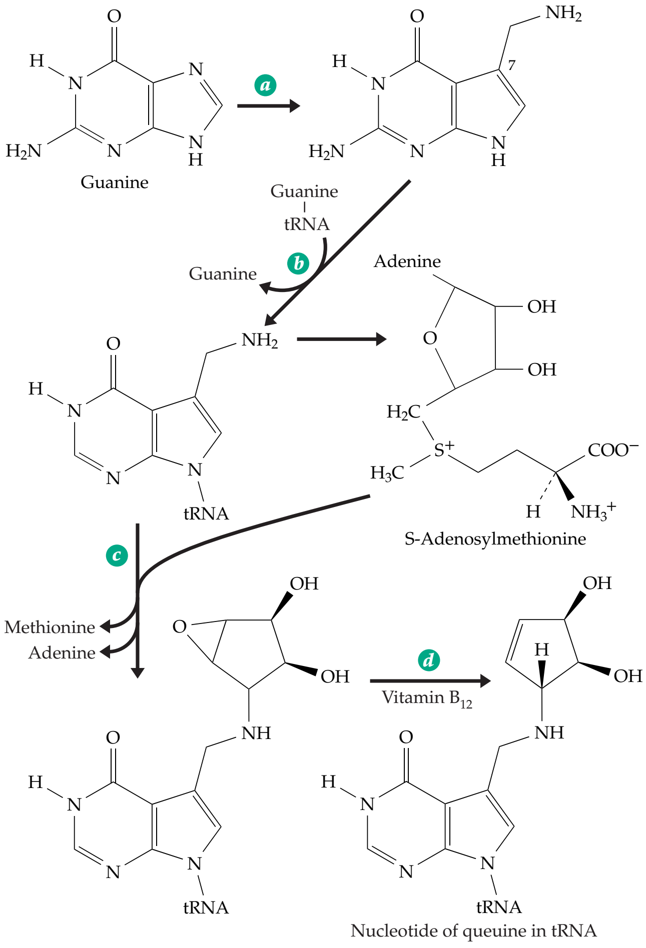

Part of DNA sequence extending past the CCA end of the tRNA gene
3′
5′
l strand
r strand
Initial cut at 3′ end
Trimming of precursor
3′ end of tRNA
RNase P
5′ end of tRNA
Up
A 85
C 80
1
5
15
20
25
30
35
40→ψ
45
50
60
65
70
Gm
s⁴U
A→I°A
T
ψ

Figure 28-10 Sequence of an *E. coli* tyrosine tRNA precursor drawn in a hypothetical secondary structure. Nucleotides found modified in the mature tRNA are indicated with their modifications (S^4, 4-thiouridine; G^m, 2'-*O*-methylguanosine; I^0, N^6-isopentenyladenosine; ψ, pseudouridine; T, ribothymidine; see also Fig. 5-33).[241] A partial sequence of the tRNA gene past the CCA end is also shown. Note the region of local 2-fold rotational symmetry (indicated by the bars and the dot). The anticodon 3'-CUA (shaded) of this suppressor tRNA pairs with termination codon 5'-UAG.


Guanine
a
7
Guanine
tRNA
b
Guanine
Adenine
S-Adenosylmethionine
tRNA
c
Methionine
Adenine
d
Vitamin B12
tRNA
tRNA
Nucleotide of queuine in tRNA

(28-4)

vitamin. However, germ-free mice reared on a queuine-deficient diet seemed normal. Therefore, the essentiality of queuine in the human diet is in doubt.[242,243]

B. Viral RNA in Prokaryotes

Bacteria not only transcribe their own genes but sometimes transcribe, or assist in transcribing, genes of invading DNA viruses or of integrated proviruses. In addition, they assist in replication of RNA viruses, another process that requires RNA synthesis. Viruses sometimes make use of host RNA polymerases but often synthesize their own catalytic subunits. Bacteriophage T4 uses the *E. coli* RNA polymerase and σ factors but modifies their action through the binding of several phage-encoded proteins.[248] In contrast, phage T7 encodes its own relatively simple RNAP whose initiation complex (Section A,2)[29] and elongation complexes have been studied.[249–249b]

1. The Lysogenic State of Phage λ

The study of bacteriophage lambda has provided many insights into biological process.[250] As we have seen (Chapter 27), the DNA of phage λ can become incorporated into the genomic DNA of *E. coli*. The resulting prophage contains many genes (Fig. 28-11), but they remain largely unexpressed until the SOS signal (Chapter 27) is generated. Certain prophage genes are then expressed with the result that the λ DNA is excised as a replicating virus.[251] How can the λ genes remain unexpressed in the prophage but be expressed rapidly at the time of excision? Part of the answer has been found[252–254] in the *cI* and *Cro* repressors. The short L1 operon (Fig. 28-11) of the λ prophage is transcribed continuously by the *E. coli* RNA polymerase. This operon contains genes *cI* and *rex*, which are transcribed from the *l* strands of the prophage DNA as indicated in Fig. 28-11. The protein C_I (or CI) specified by gene *cI* is the **lambda repressor,** which binds to two operator sites in the prophage DNA. One operator (o_L) is to the left and the other (o_R) to the right of the *cI* gene. From a study of fragments of DNA protected by the repressor, it was concluded that each operator has three subsites, which are filled from left to right at o_L and from right to left at o_R successively by up to six repressor monomers. Each presumed subsite has a similar 17-bp quasipalindromic sequence to which a dimeric repressor can bind. The binding is cooperative, probably because the repressor molecules contact each other, apparently binding the DNA into a loop.

The right operator o_R controls not only the R1 operon but also the L1 operon, which encodes the λ repressor (*cI* gene). The first of the three subsites in the operator is adjacent to the L1 promoter P_{RM}, and binding of λ repressor activates that promoter at the same time that binding to the adjacent subsites blocks transcription of the R1 operon. Thus, the λ repressor positively controls its own synthesis. At the same time, blocking of promoters P_R and P_L prevents synthesis of virally encoded enzymes that catalyze excision of the λ DNA and replication and transcription of the rest of the genes.

The SOS signal causes rapid hydrolytic cleavage of the λ repressor and transcription of the other λ operons. The matter is more complex than this. Gene products *cIII* and *cII*, from the **early left** and **early right operons**, respectively, stimulate the transcription of *cI* and are needed for establishing the lysogenic state initially.[256] Once established these genes do not function since they are never transcribed. There are only a few molecules of the λ repressor present in a cell, but this is ordinarily sufficient to maintain the prophage state. On the other hand, irradiation of the bacterium with ultraviolet light activates the SOS response and results in rapid hydrolytic cleavage of the λ repressor and transcription of other phage operons.

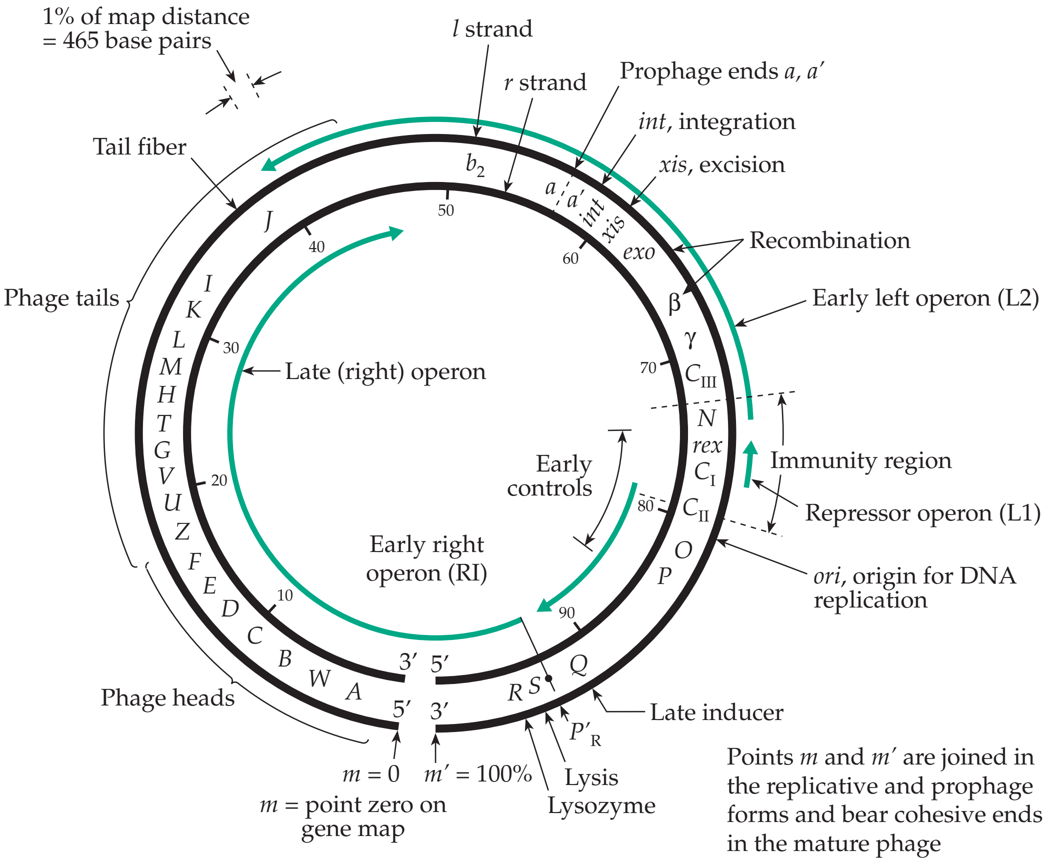
Of special significance to the lytic cycle is the Cro repressor gene *cro*, found at the beginning of operon R1. Although it binds to the same operator sequences as does the λ repressor, the Cro repressor has opposite effects.[254] It represses transcription of operon L2 and hence synthesis of λ repressor, but it positively activates P_R and P_L. The earliest proteins synthesized during lytic development are the Cro protein and the product of the first gene *N* in the left operon L2. The N protein, an antiterminator that permits transcription to continue on past points t_L and t_R, is an unstable, short-lived molecule of $t_{1/2}$ ~2 min.[257] Leftward transcription proceeds through genes *exo* and β, which are involved with recombination, and *xis*, which is required for excision. When the DNA is integrated into the *E. coli* chromosome, it is cut at points *aa′* (Fig. 28-11) and is inserted just to the right of the *gal* operon (Fig. 26-4). Prophage transcription can now continue past point a′ and into the genes of the bacterium. Translation of the mRNA formed from this early left operon generates the enzymes needed to free the prophage and to permit reformation of the circular replicative form of the phage DNA. The excision is also made near point a′, and it is easy to see how the nearby *gal* genes can sometimes be included in the excised λ genome.

The product of gene *N* also permits rightward transcription through genes *O*, *P*, and *Q* and at a slower rate on along the rest of the chromosome to point a. Genes *O* and *P* code for proteins that permit the host replication system to initiate formation of new λ DNA molecules. Replication begins at the point *ori* and occurs in both directions. Gene *Q* codes for a protein that activates transcription of the **late genes** beginning at promoter P_R.

As indicated in Fig. 28-11, the chromosome can be divided into four major operons, the short one that produces repressor, and the early left, early right, and late operons. The early operons code largely for replication and recombination enzymes and control proteins. The late operon is concerned with production of proteins needed for assembly of the virus particles and must be transcribed at an even higher rate; hence the need for the product of gene *Q*. Within the late operon, genes *A* to *F* are involved in packaging of λ DNA and in formation of heads, while genes *S* to *J* are concerned with the production and assembly of tails. Genes *S* and *R* produce proteins that lead to destruction of the host membrane and to lysis of the cell. During the late stages of lytic growth the early genes are largely shut off by the Cro repressor. We can see that even in a virus the control of transcription can be a complex process.

2. Replication of RNA Bacteriophages

The small icosahedral RNA-containing bacteriophage are of interest because of the small number of


1% of map distance
= 465 base pairs
l strand
r strand
Prophage ends a, a′
int, integration
xis, excision
Recombination
Early left operon (L2)
Tail fiber
Phage tails
Late (right) operon
b2
a
a′
int
xis
exo
β
γ
CIII
N
rex
CI
CII
O
P
Q
S
R
A
W
B
C
D
E
F
Z
U
V
G
T
H
M
L
K
I
J
10
20
30
40
50
60
70
80
90
Early controls
Early right operon (RI)
Immunity region
Repressor operon (L1)
ori, origin for DNA replication
3′ 5′
5′ 3′
Phage heads
Late inducer
P′R
Lysis
Lysozyme
m = 0
m′ = 100%
m = point zero on gene map
Points m and m′ are joined in the replicative and prophage forms and bear cohesive ends in the mature phage

Expanded diagram of immunity region

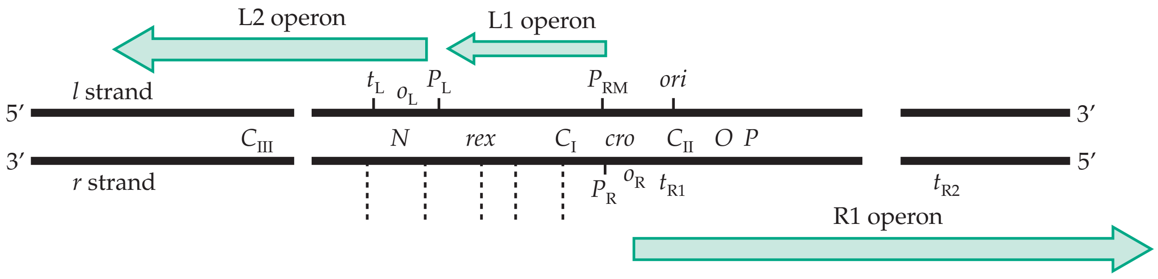

L2 operon
L1 operon
l strand
tL
oL
PL
PRM
ori
5′
3′
CIII
N
rex
CI
cro
CII
O
P
3′
5′
r strand
PR
oR
tR1
tR2
R1 operon

Figure 28-11 Genetic and physical map of the λ phage genome. After Szybalski. See Honigman *et al.*[255] for a more detailed diagram of the immunity region. The gene for the lambda repressor is labeled C_I.

genes and the possibility of obtaining a detailed understanding of their replication. The four genes in the 3569-nucleotide MS2 RNA (Fig. 29-17) code for the A protein (maturation protein), the coat protein, a replicase (RNA polymerase) subunit, and a protein needed for lysis of the host cell. The last gene overlaps both the replicase gene and that of the coat protein.[258,259] The somewhat larger phage Qβ contains a 4.2-kb RNA genome. One subunit of the Qβ replicase is encoded by the virus, but three bacterial proteins are needed to form the complete replicase.[260] They are ribosomal protein S1 and elongation factors EF-Tu and EF-Ts, proteins that normally function within *E. coli* in translation of mRNA (Chapter 29). Their ability to associate with RNA has been exploited by the phage for a quite different purpose.

Replication of a single-stranded virus must take place in two steps. From the (+) strand present in the virus a complementary (–) strand is first formed. Initiation of this step requires another bacterial host factor Hfq[260,261] and GTP. The (–) strands formed do not associate with the (+) strands. They are apparently released from the replicase in a single-stranded form and presumably fold into highly structured molecules with many hairpin loops (as for the (+) strand of MS2 RNA shown in Fig. 29-17). The (–) strands are then copied (Hfq is not needed for this) to make a large number of new (+) strands for incorporation into the finished virus particles. The Qβ replicase is able to synthesize *in vitro* complete complementary strands to either (+) or (–) viral RNA molecules. However, the system is specific for the viral RNA and will not copy any arbitrary nucleotide sequence, certain sequences at the 3' end being essential for initiation of replication. During replication in the test tube mistakes are made including premature termination and mispairing of bases.[262,263] Thus mutation takes place, and it is possible to select RNA molecules much smaller than the original viral RNA that will be replicated readily by the Qβ replicase system. One such fragment contains only 114 nucleotides in a known sequence.[264]

C. Transcription in Eukaryotic Cells and in Archaea

There are three primary domains of life, represented by the bacteria, archaea, and eukaryotes. Some of the clearest evidence for the independent evolution of these three groups of organisms is found in the transcriptional apparatus. While the basic chemistry is the same, the details of initiation and control of transcription in bacteria and in eukaryotes are very different.[264a] The archaea share characteristics of both bacteria and eukaryotes. Archaeal RNA polymerases have a complexity similar to that of eukaryotes and also share a similar mechanism of initiation of transcription.[265–266b] Several of the protein transcription factors of archaea also resemble those of eukaryotes.[267,268] However, in a comparison of DNA sequences from the complete genomes of four archaeal species, it was found that of 280 predicted transcription factors or transcription-associated proteins 168 were homologous to bacterial proteins and only 51 to eukaryotic proteins.[268] This tends to confirm the ancient divergence of the three primary domains of life.

In bacteria transcription and translation are closely linked. Polyribosomes may assemble on single DNA strands as shown in Fig. 28-5. It has often been assumed that RNA synthesis occurs on loops of DNA that extend out into the cytosol. However, recent studies indicate that most transcription occurs in the dense nucleoid and that assembly of ribosomes takes place in the cytosol.[268a] In a similar way eukaryotic transcription occurs in the nucleus and protein synthesis in the cytosol. Nevertheless, some active ribosomes are present in the nucleus.[268b]

1. Eukaryotic Nuclei and Transcription

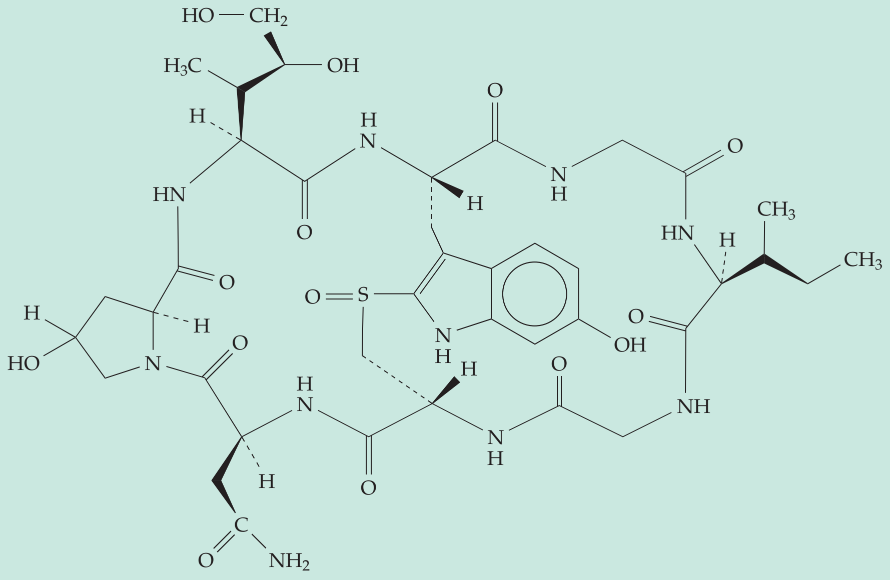
In cells with true membrane-enclosed nuclei the messenger RNA molecules are relatively long lived. They must move out from the nucleus to the sites of protein synthesis in the cytoplasm. In addition to the need for eukaryotic mRNA to travel further and to last longer than that of bacteria, a number of other differences are evident. Eukaryotic mRNAs are usually transcribed from single genes. Polygenic operons are uncommon in most animals but are numerous in *C. elegans*.[268c] Eukaryotic cells appear to rely less on negative control through specialized protein repressors than do bacteria but use a greater variety of positive control mechanisms. However, most genes are repressed by being held in a **silent state**.

Another characteristic of eukaryotes is the extensive processing of transcripts. Most primary transcripts that give rise to mRNA appear first in the nucleus as **heterogeneous nuclear RNA** (**hnRNA**). Like mRNA it has a base composition resembling that of DNA. The molecular size varies from 1.5 to 30 kb or more. It turns over rapidly, most of it having a half-life of ~10 min. However, some may last as long as 20 h. Only about 5% of the hnRNA ever leaves the nucleus as mRNA, most being degraded without export to the cytoplasm.[269] The processing consists of **capping** at the 5' ends, removal of introns (splicing), **cleavage** by nucleases, **polyadenylation** at the 3' ends, **methylation**, formation of pseudouridines, other covalent base modifications, and sometimes **editing**. Because of the complexity of eukaryotic transcription there are many points at which control can be exerted during initiation of transcription, termination of transcription, splicing and methylation, transport of mRNA out of the nucleus, and the degradation of mRNA. However, the first points of control are found in the nucleosomes and in the structure of the chromatin.

It has been recognized for many years that genes in heterochromatin (Chapter 27) are usually not transcribed. In most chromatin the great majority of the genes are silent most of the time. Genes are repressed by the folding of the DNA into nucleosomes and by the further folding into higher order folds or coils (Fig. 27-5).[270–272] Various activating transcription factors as well as RNA polymerases must bind to the DNA, displacing it from the histones around which it is wrapped in the nucleosomes. The processes by which inactive nucleosomal DNA becomes active in response to external signals are beginning to be understood. Chemical alterations in the histones in the nuclear matrix and in other nuclear proteins and also in the DNA itself may be involved.[273–275] As pointed out in Chapter 27, the CpG "islands" that lie upstream of many genes are heavily methylated in the silent heterochromatin. Repressor proteins may bind to methyl-CpG groups.[276,276a]

In recent years attention has been focused on the N-terminal "tails" of histones H3 and H4 (Fig. 27-4) in which lysine side chains undergo reversible acetylation and which may also be phosphorylated and

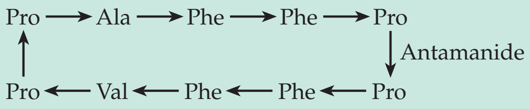
BOX 28-B POWERFUL POISONS FROM MUSHROOMS



Several deadly species of the genus *Amanita* produce colorless toxic octapeptides, the **amanitins**.[a,b] Two residues of glycine, one of L-isoleucine, one of the unusual L-dihydroxyisoleucine, one of L-asparagine, and one of L-hydroxyproline are present in α-amanitin. In the center a modified tryptophan residue has been combined oxidatively with an SH group of a cysteine residue. If the dihydroxyisoleucine residue of α-amanitin is replaced with unhydroxylated leucine, the resulting compound, known as amanullin, is nontoxic. The LD_{50} for mice is 0.3 mg kg^{-1} and 50 g of fresh *Amanita phalloides* may be sufficient to kill a person. Amanitins act slowly, and it is impossible to kill mice in less than 15 h, no matter how high the dose.

α-Amanitin completely blocks transcription by eukaryotic RNA polymerases II and III. Polymerase II is the major nuclear RNA polymerase, and its inhibition prevents almost all protein synthesis by the cell. Note that the amanitin molecule is semisymmetric overall, much as is a actinomycin (Box 28-A), with an aromatic group protruding from behind in the center.

The same mushrooms contain several fast-acting toxic heptapeptides, the **phalloidins**, whose structures are similar to those of the amanitins. However, they contain a reduced sulfur atom (—S—) in the cross-bridge. They are specifically toxic to the liver.[c] The same mushrooms also contain an antidote to the phalloidins, **antamanide**. This cyclic decapeptide, like the toxins, is made up entirely of L-amino acids, and it apparently competes for the binding site of the phalloidins. Unfortunately, it is of little value in treating cases of mushroom poisoning. Antamanide is a specific sodium-binding ionophore.



---

[a] Wieland, T., and Wieland, O. (1972) *Microb. Toxins* **8**, 249–280

[b] Wieland, T., and Faulstich, H. (1983) *Handbook of Natural Toxins*, Marcel Dekker, New York

[c] Wieland, T., Nassal, M., Kramer, W., Fricker, G., Bickel, U., and Kurz, G. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5232–5236

methylated.[33b,271,275a–c] The resulting modifications in shape and electrical charge can affect the ways in which the histones interact with the DNA, with each other, and with other proteins in transcription complexes.[271a,275a]The histone tails may interact with adjacent molecules either to pack the chromatin more tightly or to loosen it and allow transcription to take place. Active chromatin has long been known to be highly acetylated, while silent chromatin has a low degree of acetylation, but a high degree of methylation of both histones and DNA (pp. 1541–1542). Methylation of histone H3 on lysine 9 (H3-Lys^9 or H3-K9) is especially significant.[275d,e] See Chapter 32, Section A,1 and C,1 for further discussion.

Several transcriptional activators form complexes with acetylating enzymes, the **histone acetyltransferases** (**HATs**), while transcriptional repressor proteins often associate with **histone deacetylases**.[277–279c] The deacetylases are often found in very large complexes. For example, the mammalian complex **Sin3** contains two histone deacetylases plus at least five other subunits, some of which evidently bind to histones.[277–278a] Sodium butyrate in millimolar concentrations is a powerful inhibitor of these deacetylases.[280] Special **silencer** sequences in the DNA are sometimes present and provide sites for binding of transcriptional repressors. Among these are **silent information repressors** (**Sir** proteins). They regulate regions of DNA that can be converted to a heterochromatin-like state. They have been located in silenced mating type loci in yeast (see p. 1880), in telomeres, and in DNA containing ribosomal RNA genes.[280a,b] Sir proteins have an unusual histone deacetylase activity. The acetyl groups removed from histones are transferred by reaction with NAD$^+$ to ADP ribose (see Eq. 15-16).[280c,d] Regions of silenced DNA are often set apart by **insulator** or bounding regions.[280e–g]

An important mechanism of silencing some genes is the use of an antisense RNA strand, as is illustrated for a bacterial system in Fig. 28-8. This **RNA interference** is also used in animals and plants, often as a way of blocking replication of viruses.[1,1a,280h–j] The small 20–25 nucleotide siRNAs that function in this way are abundant in *C. elegans* and in *Drosophila*.

Histone acetyltransferases also form large complexes that acetylate not only histones but also other nuclear proteins.[281–283] The **SAGA HAT** complex of *S. cerevisiae* has a molecular mass of ~2.0 MDa and contains at least 14 subunits.[283a] Large acetylating complexes have also been identified in *Tetrahymena*, *Drosophila*, *Arabidopsis*, and mammalian species.[283] Some subunits of these complexes have been identified as previously known transcription factors. Such multiprotein complexes are sometimes described as **cis-regulatory elements** (CREs).[283b,c]

Changes in the structural properties of chromatin observed during silencing of genes or during their activation are often described as **chromatin remodeling**. Large multisubunit complexes are involved. Their action is characteristically dependent upon ATP hydrolysis.[284–284b] Complexes **SWI/SNF** and **RSC**, first found in yeast but also present in human cells, appear to participate in the disruption of nucleosomes needed for initiation of transcription.[284–287c] Among other distinctly different remodeling complexes are **ISWI** of *Drosophila*, its human homolog, and the human Williams syndrome transcription factor **WCRF**.[284] The ATP-dependent component in these complexes has a conserved sequence that is shared with DNA helicases. RNA helicases are also required for all processes that form, modify, or utilize RNA.[288,288a–c] However, the chromatin remodeling complexes appear not to unwind DNA but to open the nucleosomal DNA for initiation of transcription.[272,284] They may act in a processive fashion and be coupled to transcription. **Peptidyl-prolyl isomerases**, such as the cyclophilins (Box 9-F) may also be essential components of chromatin-remodeling complexes.[289]

Just as they participate in driving the cell cycle, ubiquitin and proteasomes also function in the control of transcription. In many cases specific transcription factors are targeted for destruction after they are used to activate or repress a gene.[289a–c] However, a 19S regulatory complex, which consists of a base and a lid of the proteasome (Box 7-A), may participate directly in control of transcription rather than in mediating proteolysis.[289a–e] In addition to ubiquitin a 97-residue relative designated SUMO-1 is linked to proteins by enzymes resembling E_1 and E_2 of the ubiquitin system (Box 10-C). Conjugation with **SUOMO-1** regulates some transcription factors and has other functions,[289f,g] e.g., participation in control of nuclear pores. The 81-residue protein **NEDD8**, which is 80% homologous with ubiquitin, controls some transcriptional processes in heart and skeletal muscle.[289h,i] Furthermore, ubiquitin-like sequences (UBX domain) are present in the C-terminal ends of a variety of specific proteins.[289j]

2. RNA Polymerases

Eukaryotic nuclei contain at least three RNA polymerases[269,290–292] which have the following functions:

Polymerase I	Formation of large pre-rRNAs
Polymerase II	Transcription of most genes to give precursors to mRNA and most small nuclear RNAs (snRNAs and small nucleolar RNAs (snoRNAs)
Polymerase III	Formation of 5S rRNA, tRNAs, and small RNA U6

Polymerase I is localized in the nucleolus.[293,294] Mitochondria contain a fourth RNA polymerase[295,295a] and chloroplasts a fifth.[296]

Like the bacterial polymerase, eukaryotic RNA polymerases are large 500–600 kDa aggregates of 9–14 subunits each. Yeast and human RNA Pol IIs each contain twelve subunits (Fig. 28-12).[290] There are two large nonidentical subunits, which in mammalian cells have masses of 214 kDa and 140 kDa and are homologous to the β′ and β core subunits of the *E. coli* polymerase, respectively.[291,297–299c] The active site contains one or two catalytic Mg^{2+} ions.[290] The largest subunit has an unusual singly glycosylated C-terminal domain (known as the **CTD**). It contains the repeating sequence $(YSPTSPS)_n$. The number of repeats varies: $n = 18$ in plasmodia, 27 in yeast, 45 in *Drosophila*, and 52 in mammals.[300] The numerous serine side chains in this tail domain undergo phosphorylation and dephosphorylation to varying extents during each catalytic cycle.[301] This may be a way of easing the transcriptional complex through nucleosomes,[302] but its most important function appears to be the linking of transcription to pre-mRNA processing.[303,304]

RNA polymerases I and III have properties similar to those of Pol II.[304a,b] Polymerases II and III are very sensitive to inhibition by the lethal mushroom poison **amanitin** (Box 28-B). However, both RNA polymerase I and RNA polymerases of mitochondria resemble the bacterial enzyme in being resistant. Genes transcribed by polymerase I, II, and III are often referred to as genes of classes I, II, and III, respectively. While mRNA is transcribed from class II genes, rRNAs, tRNAs, and some small RNAs, which must undergo processing but are not polyadenylated, are transcribed from genes of classes I and III. Each type of nuclear RNA functions in its own sites as independent "factories."[305,306]

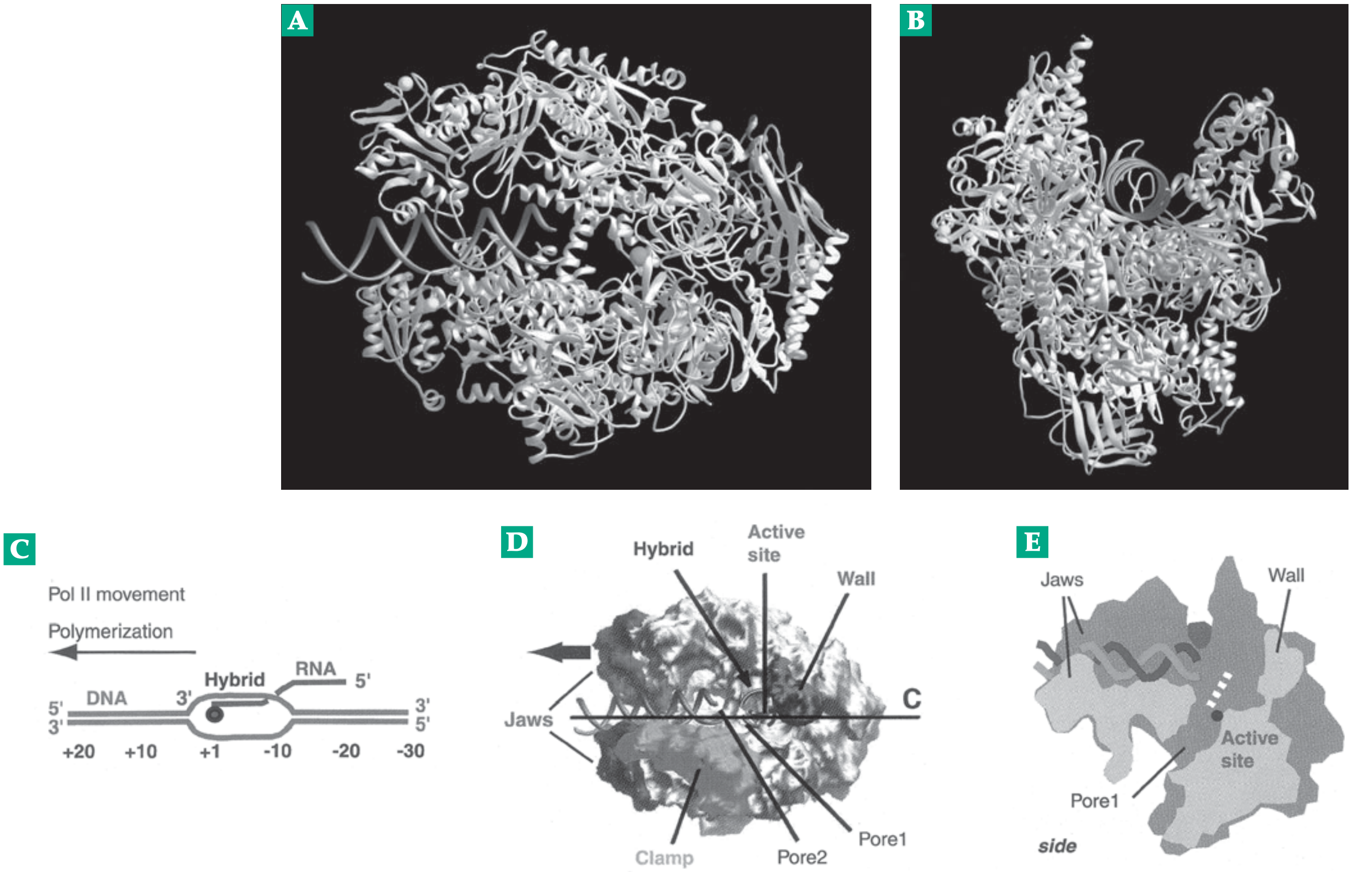
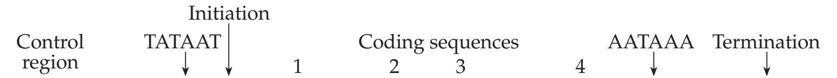


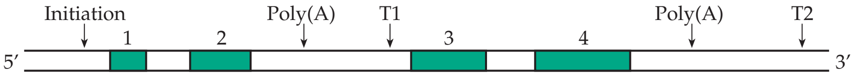
Figure 28-12 Three-dimensional structure of yeast RNA polymerase II. (A) View from the "top," with backbones of ten subunits (of 12) in the 514-kDa protein shown as ribbon drawings. A 20-base-pair segment of B-DNA has been modeled in a location indicated by electron crystallography. (B) Side view looking toward the end of the DNA. Eight zinc atoms as well as an active-site magnesium (green) are visible. (C) Schematic drawing showing the transcription bubble as proposed for a transcribing polymerase. (D) Surface representation of the polymerase viewed as in (A). (E) Side view of a section cut along the line marked C in D. The dashed white line represents the axis of the DNA-RNA hybrid segment. The hybrid axis must lie at an anlge with respect to the axis of the incoming DNA. Pore 1 may be a route for exit of RNA during "backtracking." The nucleotide triphosphate substrates may also enter via pore 1. From Cramer *et al.*[290] Courtesy of Roger D. Kornberg.

3. Transcriptional Units and Initiation of RNA Synthesis

Typical simple **transcriptional units** for class II eukaryotic genes contain the following elements: (a) site of initiation, (b) **TATA sequence** (Goldberg–Hogness sequence) at position ~–30 bp, (c) **upstream regulatory elements**, (d) **enhancers**, (e) a series of **coding sequences** or exons separated by introns, an **AATAAA sequence** that in the RNA transcript may establish the 3' polyadenylation site, (f) a termination region.[307]



Simple transcriptional unit



Complex transcriptional unit

The preceding paragraph describes a **simple transcriptional unit**. There are also **complex transcriptional units**.[308] For example, the terminator sequence T1 may be followed by additional exons and a second polyadenylation signal and second termination sequence T2. Termination may sometimes occur at T1 and sometimes at T2 resulting, after splicing, in two mRNAs, one containing only exons 1 and 2 and the other all four exons. In some cases two or more different modes of splicing may occur with one or more exons omitted from the final processed mRNA. Thus, a single transcriptional unit can give rise to two or more different proteins which share some common sequences.[309–312] Multiple initiation sites sometimes exist as well.

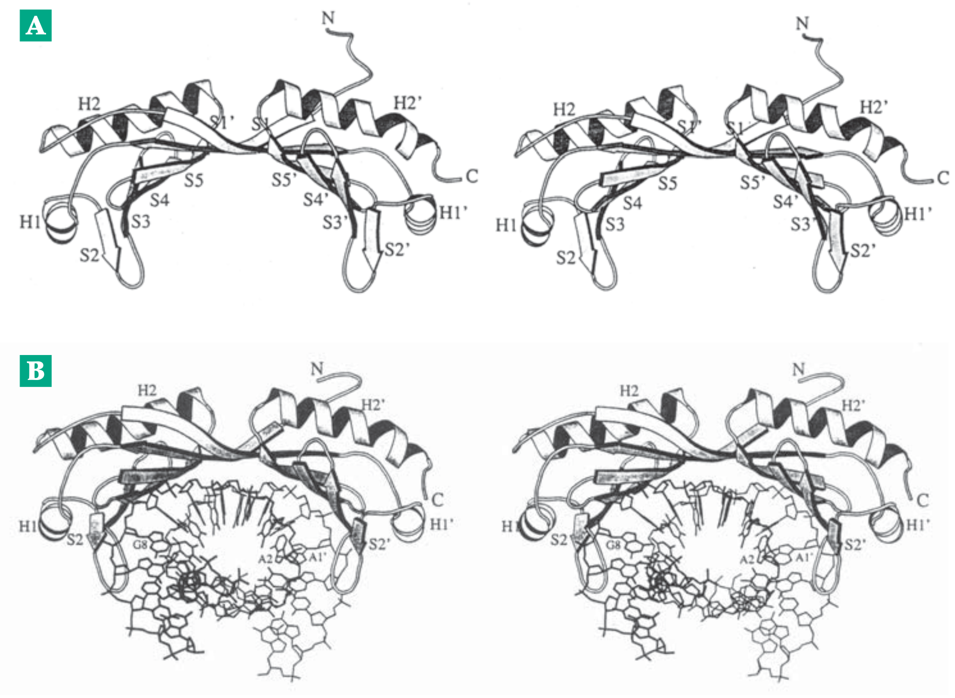
Initiation of RNA synthesis is a complex process, which is summarized diagrammatically in Eq. 28-5. Some details are given in Sections 4–7. Elongation of the RNA being formed is also complex, often requiring splicing and other processing steps before synthesis can be terminated. These steps, which are discussed in Section 8 and in Section D, also depend upon large complexes of proteins, which are directly coupled to the RNA polymerase.[312a–c]

4. Promoters, Transcription Factors, Enhancers, and Activators

Eukaryotic promoter sequences are less well defined than are those of bacteria, and the initiation points for transcription are more variable. Most promoters contain the TATA sequence, which is usually required for binding of polymerase II. In addition, there are upstream regulatory elements (Table 28-1). Many promoters contain the sequence 5'- **CCAAT** at about –75, and others have 5'- **GGGCGG** or similar sequences at close to –100.[273, 273a,b] A sequence found upstream from many yeast genes is 5'- **TGACTC** or the longer semipalindromic 5'-**ATGACTCAT**.[313,314] Some sequences are unique to small sets of genes such as the “heat-shock” genes[315] considered in Section 6. Most upstream sequences are not polymerase-binding sites but attachment sites for additional protein transcription factors. Computer programs that help locate pro- moter sequences have been devised.[316,317]

The SV40 early promoter and transcription factor Sp1. The study of transcription in eukaryotes has been difficult because purified polymerases do not initiate transcription at most promoter sites. As a consequence, much of the early work was done with viruses such as adenovirus and SV40. Their genes are transcribed by RNA polymerase II and have unusually effective promoters. A protein known as **Sp1**, isolated from human cultured cells, protects an SV40 promoter from digestion by DNase. Sp1, which is now known as an **acessory factor** or **coactivator** for transcription,[318,318a] protects a region that extends from about –45 to –104 bp and contains the hexanucleotide GGGCGG sequence repeated six times. The Sp1 protein was isolated as a mixture of related 95- and 105-kDa peptides[319] and was found to contain two DNA-binding zinc-finger domains (Figs. 5-37, 5-38).[320] The Sp1 protein is synthesized in most cells and binds not only to SV40 promoters but also to many promoters of host cells[321] (an example is provided by the mouse mitochondrial aspartate aminotransferase gene whose sequence is shown in Fig. 5-4). In genes that lack a TATA sequence the binding of Sp1 or the related Sp2, 3, and 4 to GGGCGG or similar GC-rich sequences is essential to initiation of transcription.[322,323] A possible role is to assist in nucleosome remodeling.[322] Sp1 also binds, together with other transcription factors, to certain enhancer sequences.[324,324a,b] Its effects are modulated by posttranscriptional phosphorylation and glycosylation.[321]

The TATA binding protein and general transcription initiation factors. A slow basal level of transcription can be observed when all but a small part of the control region at the 5' end of a gene is deleted.[325] This minimum promoter, which includes the TATA sequence, is the binding site of both the RNA



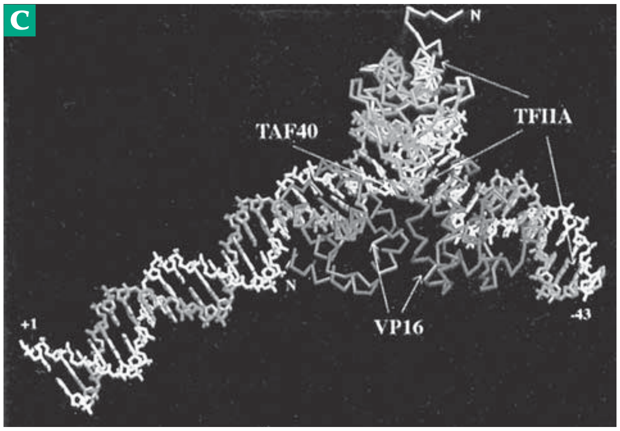
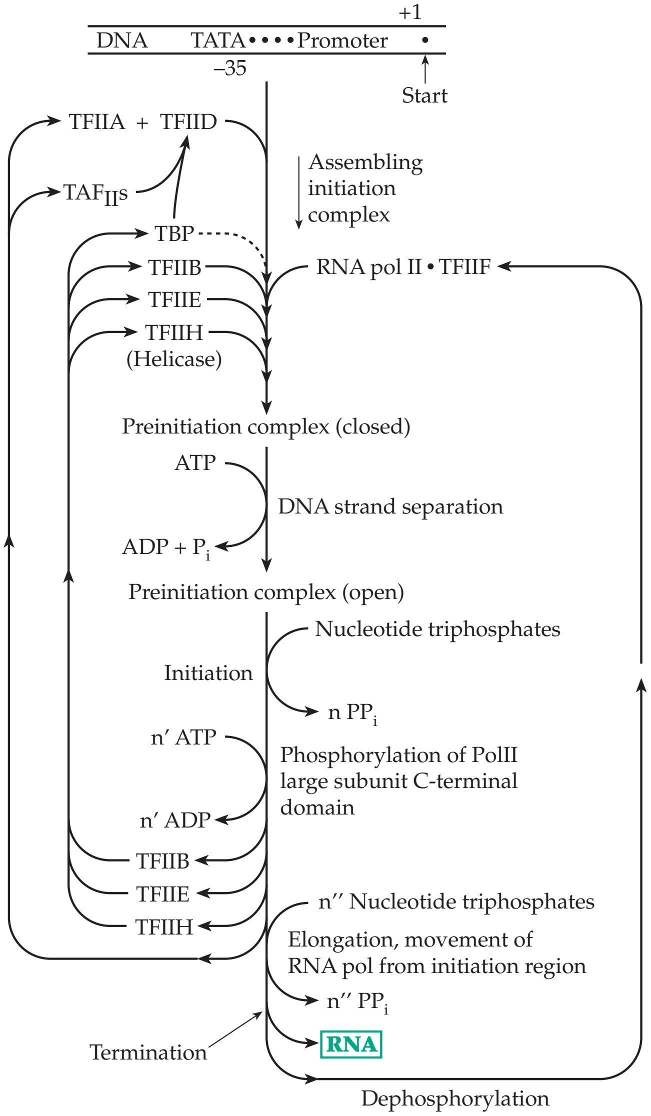


Figure 28-13 (A) Stereoscopic ribbon drawing of the phylogenetically conserved 180-residue C-terminal portion of the TATA-binding protein (TBP) from *Arabidopsis thaliana*. The sequence consists of two direct repeats, giving the protein an approximate twofold symmetry. From Nikolov *et al.*[337] (B) Structure of the corresponding C-terminal core (residues 155–335) of the human TATA-binding protein (TBP) bound to the TATA sequence of a promoter in adenovirus DNA. From Nikolov *et al.*[327] (C) Structure of human transcription factor IIB bound to a TBP from *Arabidopsis thaliana*, which, in turn, is bound to an adenovirus TATA sequence. Hypothetical B DNA extensions have been modeled at both ends of the DNA segment. The +1 at the left end is the transcription start site and the –43 upstream end is to the right. From Nikolov *et al*.[338] Courtesy of Stephen K. Burley.

polymerase and of transcription initiation factors that are designated TFII-A, -B, -D, -E, -F, and H. Because they affect many genes these are called **general transcription factors**.[326–329] TFIID is a large complex of a DNA-binding subunit known as the **TATA-binding protein (TBP)** together with 8–12 additional tightly bound subunits known as **TBP-associated factors** (**TAF$_{II}$s**). Many of the TAF$_{II}$s have histone-fold structures. Some possess histone acetyltransferase and other enzymatic activities.[329a] They may stabilize initiation complexes on specific gene promoters.[330–332a] TBP binds specifically to the TATA sequence (Table 28-1),[333] which is found in most promoters for RNA polymerases I, II, and III. Its three-dimensional structure resembles a saddle, which sits astride the TATA sequence (Fig. 28-13). The DNA is bent, untwisted by ~117°, and the minor groove broadened to allow a good fit.[327,328,334] TFIIB is thought to bind first to the DNA–TBP complex, after which the RNA polymerase II complex binds and becomes positioned on a promoter site. Other factors, including the ATP-dependent bidirectional helicase TFIIH,[335,335a] also add (Eq. 28-5). ATP may be needed for more than one step in initiation.[336] TFIID contains TAF$_{II}$ subunits. They may bind along with TBP[336a] as indicated in Eq. 28-5, or they may bind at a later point in an assembly pathway. It is often stated in current literature that the growing initiation complex "recruits" the next subunit. It is important to realize that this simply means that the next subunit that strikes the complex by diffusion sticks, perhaps with cooperativity in binding. The word recruit doesn't mean that the complex advertises a vacancy for the next subunit. The very large complexes (**transcriptosomes**) that are formed vary in composition and in assembly pathways. Proteins related to TBP bind to different promoters.[310] It is also


+1
DNA
TATA••••Promoter
–35
Start
TFIIA + TFIID
TAF$_{II}$s
Assembling initiation complex
TBP
TFIIB
TFIIE
TFIIH (Helicase)
RNA pol II • TFIIF
Preinitiation complex (closed)
ATP
DNA strand separation
ADP + P$_i$
Preinitiation complex (open)
Nucleotide triphosphates
Initiation
n PP$_i$
n′ ATP
Phosphorylation of PolII large subunit C-terminal domain
n′ ADP
TFIIB
TFIIE
TFIIH
n′′ Nucleotide triphosphates
Elongation, movement of RNA pol from initiation region
n′′ PP$_i$
Termination
RNA
Dephosphorylation

(28-5)

important to recognize that some promoters bind transcription factors tightly while others do so only weakly. Some are constitutive, always functioning, but most are inducible, acting only upon appropriate stimulation. The many operators present in a genome and the many species of eukaryotes present on earth ensure a vast variety of detailed pathways and control mechanisms.

Transcriptional activators. Many proteins serve as activators of transcription, causing larger increases in rate over those observed with TBP alone. Some of these are listed in Table 28-2.[338a] The table also lists two proteins (Sp1 and NF1), and the DNA sequence CCAAT, which control constitutive or continuously active genes. A large group of transcription factors are active in development. Receptors may be resident in cytoplasm, cell membrane or nucleus, as indicated in Table 28-2. Some cytoplasmic factors are *latent*, becoming active only following stimulation by an external signal. All of the factors in Table 28-2 are positive-acting.[338a] However, some negative-acting factors are known. One, designated **Ssn6-Tup1** in *S. cerevisiae* is a global repressor, affecting many genes.[338b] For example, it opposes the activator GAL4. Proteins related to Ssn6-Tup1 are found in flies, worms, and mammals.

Genetic studies indicate that gene activator proteins often bind to TFIIB, TFIID, and TFIIH.[272] The coactivator **TafII130** (which binds to TFIID), and the bound transcriptional activator Sp1 apparently interact with the protein **huntingtin** in regulation of transcription in the brain.[338c] A defect in huntingtin leads to the fatal neurodegenerative Huntington disease (pp. 1516, Chapter 30). Other activator proteins bind to upstream activator sequences, as in prokaryotes. Among the most studied of these is GAL4, an 881-residue yeast protein that binds to a specific 17-bp palindromic upstream site near the TATA sequence. It activates transcription of genes needed for galactose metabolism.[339–343] The GAL4 protein contains a binuclear metal cluster composed of two Zn^{2+} ions and six cysteine side chains, two of which bridge the pair of metal ions.[341,343] GAL4 is able to activate genes of *Drosophila* and of human cells. The specific GAL4-binding sequence has been introduced into 5' control regions of genes in various positions.[344] It was found that neither an exact distance nor alignment between the GAL4 binding site and the TATA sequence is required, but activation is best when this distance is not too large. The explanation for the lack of a requirement for alignment seems to lie in the flexibility of the C-terminal segment of GAL4, which carries a large negative charge and may bind to the repeated C-terminal sequence of RNA polymerase II to activate it.[345] Many other promoter-specific activators are known.

Mediators and coactivators. Transcriptional activators that act in a crude cell-free system often do not function with purified DNA, RNA polymerase, and the basal transcription factors as indicated in Eq. 28-5. Studies with yeast, *Drosophila*, and human cells revealed that additional large multisubunit complexes known as mediators are needed.[272,346–348] A yeast mediator complex consists of 20 subunits.[349–350b] Many activator proteins bind to the DNA sequences known as enchancers, discussed in the next section. Mediator complexes may also interact with enhancer-bound activators. Individual proteins, such as the TAF subunits, that bind to and cooperate with activator proteins are often called coactivators.[351]

Enhancers. Complex DNA sequences called enhancers help to regulate transcription of many eukaryotic genes. The first of these was discovered in an upstream control region of the virus SV40 DNA and consists of two repeats of a 72-bp sequence.[47,352,353] The presence of an enhancer sequence may cause as much as 100- to 1000-fold increase in the rate of transcription as compared with the same transcriptional unit from which the enhancer has been deleted. A surprising fact is that enhancers as far as 1–2 kbp upstream or even far downstream of the promoter and in either of the two possible orientations are effective. This finding suggested that enhancers induce long-range conformational alterations in DNA. Alternatively, they might contain points of entry for RNA polymerase or for an initiation factor that could move along the DNA to the promoter region. However, the synthetic DNA molecule shown in Fig. 28-14 contains two copies of an enhancer in opposite orientations in one strand but none in the other strand.[354] The

TABLE 28-1
Nucleotide Sequences of Some Commonly Found Regulatory Elements in Promoters and Upstream Activator Sites

Sequence[a] transcription	*Description. Positions are relative to the start site*
RNA polymerase binding sites:	
5'-**TAT**AA**T**[b]	Bacterial –10 or Pribnow promoter sequence
5'-**TT**GACA	Bacterial –35 region promoter sequence
5'-(C/T)TTA(A/T)Ann	Archaeal –30 region, TBP binding site[c]
5'-TATA(A/T)A(A/T) or 5'-TATA@A@n[d]	Eukaryotic –30 region; yeast –60 to –120 region
Upstream eukaryotic promoter sites:	
5'-GGGCGG	–100 region, Sp1
5'-CCAAT	–75 region, CTF
Small GC clusters ~5 bp apart	Binding site for TFIIIA
5'-GCGCC-C	~ –5; “discriminator” sequence: inhibition of gene expression by ppGpp
Enhancer elements and transcription factors:	
5'-ATGA(C/G)TCAT	AP-1, cJun, GCN4 (yeast)
5'-CCCCAGGC	AP-2
5'-CAC(G/T)	Myc / Max heterodimer
5'-ATGACGTCAT	CRE (cAMP responsive element)
5'-GGTCAnnnTGACC	Estrogen-responsive element
5'-GGGTGAnnnGGGTGA	Vitamin D-responsive element; direct repeats
5'-CC(A/T)$_6$GC	SRE (serum response element)
5'-ATGCAAAT	Homeotic genes; Oct-2
5'-GGTCAnnnTGACC	ERE (erythroid responsive element)

---

[a] The sequences are all for the sense strand of the DNA, n = any base. See Keller, W., König, P., and Richmond, T. J. (1995) *J. Mol. Biol.* **254**, 657–667; Adams, R. L. P., Knowler, J. T., and Leader, D. P. (1992) *The Biochemistry of the Nucleic Acids*, 11th ed., Chapman & Hall, London; Lewin, B. (2000) *Genes VII*, Oxford Univ. Press, New York.

[b] Consensus sequence. The bases in boldface are the most highly conserved.

[c] From DeDecker, B. S., O'Brien, R., Fleming, P. J., Geiger, J. H., Jackson, S. P., and Sigler, P. B. (1996) *J. Mol. Biol.* **264**, 1072–1084.

[d] The symbol @ refers to either A/T or T/A. See Juo, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J., and Dickerson, R. E. (1996) *J. Mol. Biol.* **261**, 239–254.

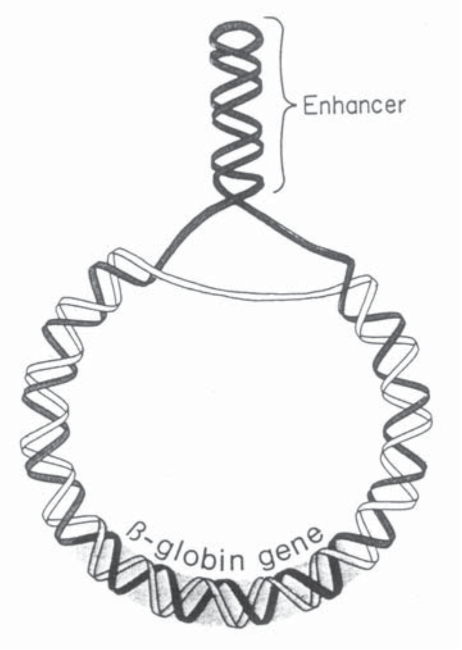


Figure 28-14 A "tailed circle" consisting of an enhancer linked to, but topologically separated from, a gene. One of the DNA strands of this plasmid bears two copies of an SV40 enhancer sequence, one copy inverted with respect to the other. This extra region protrudes from the circle and self-pairs to form a functional enhancer. The main body of the circle contains the β-globin gene, transcription of which is increased by the enhancer. Twisting of the enhancer has no effect on the winding of the strands on the main body of the circle; nevertheless, the enhancer efficiently increases β-globin transcription. From Ptashne.[355]

TABLE 28-2
Functional Classification of Positive-Acting Eukaryotic Transcription Factors[a]

Positive-acting transcription factors

- **I** Constitutive
  - Sp1
  - CCAAT
  - NF1
- **II** Regulatory (conditional)
  - **A** Developmental "cell specific"
    - GATA
    - HNFs
    - Pit1
    - MyoD
    - Myf5
    - Bicoid
    - Hox
    - Winged helix (forkhead)*
  - **B** Signal dependent
    - **1** Steroid receptor superfamily
      - GR
      - ER
      - PR
      - TR
      - RARs
      - RXRs
      - PPARs
    - **2** Internal signals
      - SREBP
      - p53
      - Steroid 'orphans'
    - **3** Cell surface receptor-ligand
      - **a** Resident nuclear factors
        - ETS
        - CREBs
        - ATMs
        - SRF (MADS box)
        - FOS-JUN
        - MEF-2*
      - **b** Latent cytoplasmic factors
        - STATs
        - SMADs
        - NFκB/Rel
        - CI/GLI (HH)
        - NOTCH (NICD)
        - CATEN INS (Wnt)
        - TUBBY
        - NFAT

[a] Major groups are labeled in large type and specific examples are listed below (green). Some bacterial proteins are included. Proteins designated by asterisks can be trapped in the cytoplasm by phosphorylation of serine site chains. From Brivanlou and Darnell.[338a]

enhancer is thus topologically separated from the globin gene present in the DNA. Nevertheless, the enhancer functioned efficiently. This suggested that enhancers are protein-binding sequences and that the bound proteins may, perhaps, close a loop to affect the transcription initiation complex.[149,355,356] An enhancer may control a whole loop of DNA in a chromosome affecting many genes.[357] Enhancer DNA sequences are often complex or *modular*, consisting of several shorter elements. For example, the SV40 enhancer contains 8- to 10-bp sequences that are repeated with minor variations and are also present in other enhancers. These are sites for the binding of activator proteins, some of which carry messages from signaling cascades such as the MAP kinase pathway of Fig. 11-13. These pathways are central to the control of cell growth. Their activation is recognized as a mitogenic response. Some enhancers (*nonmodular*) are more compact and directly tied to the RNA pol complex. An example is the human interferon-β (IFN-β) enhancer.[357a,b] It responds cooperatively to three separate activator proteins: NF-κB, interferon regulatory factors, and ATF-2/c-Jun together with an architectural HMG protein. These form an **enhanceosome**, which interacts directly with the transcriptosome complex. The whole complex is sometimes referred to as a cis-regulatory module (CRM). The DNA domains affected by enhancers may be separated by insulator or boundary regions (p. 1626).

5. The Variety of DNA-Binding Proteins

A large number of proteins, often present in very small amounts, bind to DNA and may affect transcription. Some, such as the histones, are relatively nonselective. The more specific transcription factors are often capable of binding tightly, but may do so only upon allosteric activation. Architectures of transcription factors vary. Some place an α helix in the major groove of B-DNA while for others β strands or peptide loops may interact with the DNA in the wide groove. Many transcription factors bind only to a bent or distorted DNA helix, often with a broadened groove. Intercalation of groups from other domains of a transcription factor within the minor groove may help to bring about necessary distortion.[358] Transcription factors sometimes, perhaps often, have other roles, e.g., as enzymes. For example, a pterin dehydratase acts as a transcriptional cofactor in liver cells.[359]

While architectures vary greatly, the DNA-binding domain or domains are usually rich in positively charged side chains, which may interact directly by hydrogen bonding with the charged phosphate groups of the DNA backbone or indirectly via bound water molecules (see Fig. 5-36).

Leucine zipper transcription factors. A large family of DNA enhancer-binding proteins, which are involved in regulating cell growth in specialized cells of differentiated tissues and also in yeast cells, have related structure domains with a characteristic coiled coil that holds two subunits together.[360,361] An example is the Max structure shown in Fig. 2-21. One of the first proteins of this type to be recognized was the mammalian phorbol ester-induced **cell activator protein AP-1**, which is a heterodimer of the protein **cJun**, encoded by cellular protooncogene *c-jun*, and **cFos**, encoded by protooncogene *c-fos*.[362] The heterodimer binds to the palindromic sequence 5'-TGACTCA. Each of the monomeric proteins c-jun and c-fos, as well as other members of the leucine zipper family, has an N-terminal DNA-binding domain rich in positively charged basic amino acid side chains, an **activation domain** that can interact with other proteins in the initiation complex, and the leucine-rich dimerization domain.[363] The parallel coiled-coil structure (Fig. 2-21) allows for formation of either homodimers or heterodimers. However, cFos alone does not bind to DNA significantly and the cJun/cFos heterodimer binds much more tightly than does cJun alone.[364] The yeast transcriptional activator protein GCN4 binds to the same 5'-TGACTCA sequence as does the mammalian AP-1 and also has a leucine zipper structure.[360,364,365]

Several mammalian leucine zipper proteins bind to the CCAAT sequence (Table 28-1) and are, therefore, as a family designated C/EBP.[361,366,367] A 30-residue segment of C/EBP contains four leucine residues at 7-residue intervals. When plotted as a helical wheel (Fig. 2-20) the four leucines are aligned on one side.[368] Similar sequences are present in the proteins cMyc, cJun, and cFos and in GCN4. These observations suggested that if the peptide sequence forms an α helix, the leucine side chain from two identical subunits or closely related proteins might interdigitate in a knobs-in-holes fashion to form the leucine zipper (Fig. 2-21).[368] The structures of these leucine zipper proteins have now been thoroughly investigated and verified by X-ray and NMR methods.[360,369–371] Mutational alterations in the zipper regions of these proteins decreases both activation[372] and dimerization.[373] The CCAAT sequence is found in many enhancers[366,374,375] and is present in ~25% of all eukaryotic promoters that function in differentiated tissues. A trimeric protein known as NF-Y binds specifically to this sequence.[366] Its subunits contain glutamine-rich domains and histone-fold domains that suggest formation of nucleosome-like structures.[366,376]

Another transcriptional activator of the leucine zipper class is **Myc**, a product of the *c-myc* oncogene (Chapter 11) and a key regulator of both cell growth and programmed cell death (apoptosis).[377–379a] It binds to the sequence 5'-CACGTG, and its binding is greatly enhanced by formation of the heterodimer **Myc/Max** (Fig. 2-21). Max is not an activator and may dimerize with certain other proteins to become an inhibitor of transcription. However, it is ubiquitously present and ready to join with Myc to activate an appropriate series of genes. A related enhancer sequence 5'-TGACGTCA is a **cyclic AMP response element** (CRE) that functions within hormone-responsive tissues that use cAMP as a second messenger.[364,380] Cyclic AMP activates protein kinase A (Fig. 11-4), whose catalytic subunit diffuses into the nucleus and phosphorylates the cyclic AMP-response element-binding protein **CREB**, a coactivator that binds to the CRE and which also contacts the general transcriptional factor complex.[365,381–382b] Transcriptional responses to cAMP are quite complex. For example, activation of the phosphoenolpyruvate carboxykinase (PEPCK) gene is maximal only when CREB, C/EBP, and AP-1 are all bound at adjacent DNA sites.[382c] CREB (also known as ATF) is a family of proteins that control the activities of hundreds of genes. Participating in this control is a coactivator, the CREB-binding protein (**CBP**),[382d] which is, in turn, subject to control by methylation.[383]

Control of growth. A large variety of transcription factors control cell growth. Some of these are indicated in Figs. 11-14 and 11-15 and Table 28-2. Since growth in cell numbers requires completion of the cell cycle, the specialized transcription factors involved are necessary. As is indicated in Fig. 11-5, factor E2F is of central importance. In fact, there are at least six mammalian E2F proteins, five of which have both a conserved DNA-binding domain and conserved activation domains.[383a,b] E2F1, 2, and 3 can all induce the S phase of the cycle. E2F6, in contrast, appears to bind to EF2-binding sites in DNA and also to Myc-binding sites to silence these genes and to help keep the cell in the G_0 state.[383c,d] As shown in Fig. 11-5, the retinoblastoma protein Rb also binds to EF2 and represses transcription. However, it allows transcription when phosphorylated.[383e]

Among the many other proteins that influence growth are cFos and cJun, which may be activated by the MAP kinase pathway (Fig. 11-13).[383f,g] Binding of cytokines (pp. 1571, 1845) activates signaling pathways from cell membrane receptors to two other families of transcription factors. These are the **STATS** (p. 1845)[383h–j] and **SMADS**.[383h,k,l] Upon activation STATS and SMADS move from the cytosol into the nucleus where they find their binding sites on DNA.

Response elements. DNA binding sites for activator proteins are often described as "response elements."[47] Thus, the site for the cAMP-responsive protein CREB is the response element CRE. The binding site for AP-1 is **TRE**, named after the phorbol ester TPA. HSE is the heat shock response element, GRE the glucocorticoid response element (or the glucose response element), and **SRE** the sterol regulatory element.[384,384a] Response elements tend to be present in many enhancers and cooperate with other enhancer-binding proteins to activate groups of genes. An especially large group of genes respond to the **serum-response element** (also SRE), which is found within the *c-fos* promoter region.[384b,c] It is the DNA binding site for the ubiquitous serum protein SRF (serum response factor),[385–387] which is involved in growth control, cell cycle progression, and wound repair.

Zinc-containing transcription factors. The zinc finger domain (Fig. 5-37), which is also designated Krüppel-like finger in reference to a *Drosophila* protein,[388,388a] is a repeated motif present in many transcription factors (see Fig. 5-38).[389–391] In each finger a Zn^{2+} ion is coordinated by two –SH groups and two imidazole groups to form the Cys_2His_2Zn domain. One of the first proteins in which zinc fingers were recognized is TFIIIA, one of the factors that controls transcription of 5S RNA genes. A large 30-kDa N-terminal domain contains nine ~30-residue repeats of the sequence $XF/YXCX_{2-4}CX_3FX_{4-5}LX_{2-3}HX_{3-4}HX_{2-6}$ where X can be any amino acid. The fact that TFIIIA binds 7–11 Zn^{2+} ions per polypeptide chain suggested that the repeated sequences might be Zn^{2+}-binding domains in which each Zn^{2+} is held by two cysteine and two histidine side chains. Each of the nine Zn^{2+}-binding domains might constitute a metal-binding "finger" able to interact with about five bases in the DNA.[391–394] In agreement with this idea was the observation of an ~5-bp repeat of guanine clusters in the DNA.[395] The three-dimensional structures of a large fragment of the TFIIIA N-terminal domain (see title page banner for this chapter)[396] and of numerous other zinc finger proteins are known.

Yeast proteins often contain a pair of zinc fingers, but in the nematode *Caenorhabditis elegans* and in *Drosophila* there are more proteins with three or more zinc fingers.[391] The previously discussed Sp1 has three.[390] In *C. elegans* there are more than 100 genes that encode proteins with the Cys_2His_2 zinc-binding motif.[391] In addition, there are many proteins with four-cysteine zinc-binding motifs. These include Cys_4 Zn proteins of the **GATA family** of transcription factors, which are found in fungi, plants, and animals.[397,397a] GATA-1 is a specific transcription factor for regulation of erythroid genes. It binds to the consensus sequence 5'(T/A)GATA(A/G) found in globin genes.[398] In fungi members of the GATA family regulate nitrogen metabolism, biosynthesis of siderophores, and uptake of iron.[397,399] Another family (LIM) has $Cys_2HisCysZn$ domains.[391] A widely distributed motif in transcriptional repressors is a $Cys_3HisCys_4Zn_2$ or RING finger domain.[400]

If zinc-containing domains lose their Zn^{2+} they do not bind tightly to DNA. Regulation of the flow of zinc ions from storage sites in metallothioneins (Box 6-E) into transcription factors as well as into more than 300 enzymes poses interesting mechanistic questions.[401]

Winged helix transcription factors. Liver-specific expression of certain genes in rats depends upon **hepatocyte nuclear factor-3** (HNF-3). Related proteins are encoded by the **forkhead family** of genes in *Drosophila*. These proteins have characteristic C-terminal DNA-binding domains, each consisting of three helices, one of which fits into the major groove of DNA. Also present is a twisted three-strand β structure and two flexible loops or "wings."[402,403] The structure of the DNA-binding domain is similar to that of histone H5 and also resembles the HTH domains of prokaryotic repressors (Figs. 5-35, 28-3) and of CAP (Fig. 28-6).[404] There are many members of the HNF-3/Forkhead family of proteins and of the related **Ets-domain** transcription factors.[405–406a]

The NF-κB/Rel proteins. Nuclear factor NF-κB plays a crucial role in cellular immune responses and in inflammatory disease.[407–408c] This transcription factor was first recognized for its function in regulating transcription of the κ light chains of immunoglobulins. It is a member of the larger NF-κB/Rel family, which act in concert with a group of DNA-binding inhibitors of the IκB family.[409] The structure of an NF-κB dimer bound to its DNA target, whose consensus sequence is 5'-GGGRNYYYCC, is shown in Fig. 5-40. Its architecture[410] is quite unlike that of other transcription factors discussed in this book.

HMG proteins as transcription factors. The abundant high mobility group (HMG) nuclear proteins (Chapter 27) bind to DNA, some of them to four-way junctions. The latter may be present in cruciform structures thought to play a role in regulation of transcription.[411] They are often regarded as modulators of chromatin structure. The ~80-residue HMG domain contains three helices and binds into a flattened, underwound, and bent DNA minor groove.[412,413] HMG proteins also act as transcription factors, which may interact directly with TBP, p53, steroid hormone receptors, and enhancers.[414] Cooperative binding with other DNA-binding proteins is characteristic of the effects of HMG proteins.[414a] Such interactions may be affected by acetylation. Members of the enhancer-binding HMG-14 / -17 family undergo acetylation at seven specific sites.[415]

6. The Variety of Transcriptional Responses

Every protein has specialized functions, and specific regulatory mechanisms often control transcription of its genes. A cell must respond to a large number of stimuli, and responses often include activation or repression of transcription.[415a] In some cases an internal signal, such as a change in concentration of a nutrient or a key metabolite, provides the stimulus. In other cases an external stimulus such as heat or light is the inducer. A few examples follow. Others are mentioned throughout the book.

Many types of hormonal response, including those of insulin[415b,416] (Chapter 11), are transcriptionally mediated. This is also true for plants (Chapter 30). Defensive responses of both animals and plants (Chapter 31) are mediated in part by transcriptional responses.

Nuclear hormone receptors. Among the best known transcription factors are a large family of hormone receptors, which not only bind specific hormones but also contain in a central domain a pair of Cys_4Zn fingers that interact with response elements (discussed in Chapter 22).[416a–c] A subfamily of these binding proteins includes receptors for glucocorticoids, progesterone,[417] androgens, and mineralocorticoids. Another member binds the insect hormone 20-hydroxyecdysone (Fig. 22-12) and regulates the puffing seen in giant salivary gland chromosomes of *Drosophila*.[418,418a] A larger subfamily binds estrogens, vitamin D_3, thyroid hormone,[419] and retinoic acid.[420] The vitamin D receptor appears to serve also as a **bile acid sensor**.[420a] The same subfamily also includes many "orphan receptors."[421–424] The latter have been discovered by DNA sequence comparisons and have led through "endocrinology in reverse" to discovery of new hormonal signaling pathways.[425] The binding of a nuclear receptor to its response element in DNA is well illustrated by the estrogen receptor,[426–427b] which binds to a palindromic **estrogen response element** with the consensus sequence 5'-GGTCAnnnTCACC. It is regulated both by hormone-binding and by phosphorylation. The latter is catalyzed by a cyclinA-CDK2 complex in response to cell cycle alterations.[428] The glucocorticoid receptor protein forms a complex with a second protein, which has been identified as the 90-kDa chaperonin hsp90.[429] A **sterol regulatory element-binding protein** functions in a more general way to activate over 20 different genes that encode enzymes needed for synthesis of cholesterol and unsaturated fatty acids by animal cells.[429a,b] The steroid hormones often bind to their receptors in the cytosol, and the resulting complex is translocated into the nucleus (p. 1264).

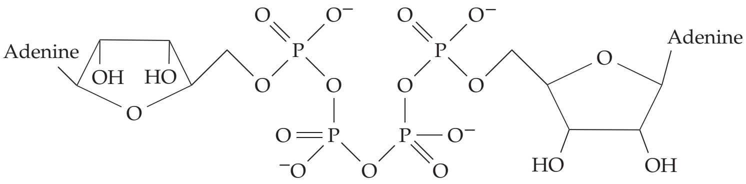
Nutrient control. In addition to the sterols and fat-soluble vitamins, other dietary constituents are also recognized by transcriptional activators or repressors.[430] These include glucose,[431,432] amino acids,[433] phosphate ions,[434,435] and various metal ions. Best known among the latter is iron. The mammalian **iron response element** (IRE) is a hairpin loop RNA structure, which like the bacterial attenuator system (Fig. 28-9) functions posttranscriptionally.[436–438a] Iron regulatory proteins (IRP1 and IRP2), which contain Fe_4S_4 clusters, bind to the IRE sequences and inhibit translation. IRP1 is identical to cytosolic aconitase. A high intracellular iron concentration promotes assembly of the Fe_4S_4 cluster and binding to the IRE (see also Chapter 16). In the green algae *Chlamydomonas* a **copper response element** (GTAC) in DNA induces expression of genes important to copper uptake.[439] Many nutritional response systems have been recognized first in bacteria. For example, *E. coli* controls uptake of molybdate[127a,438,440] and of phosphate[437,441] as well as of sugars and ammonia (discussed in Section A,4).

Stress proteins. Ritossa reported in 1962 that when *Drosophila* were suddenly warmed from 25°C to 36–37°C, a series of new puffs could be observed within the polytene chromosomes of the salivary glands.[442] These same puffs, which were also seen after other stresses, appeared within a few minutes and were associated with formation of new mRNAs. This **heat-shock** phenomenon was subsequently found to be universal. It is observed in all species of animals as well as in plants and bacteria.[443–448]

The principal **heat-shock proteins** (hsp), encoded by the new mRNAs, belong to five conserved classes: hsp100, hsp90, hsp70, hsp60, and small heat shock proteins. The function of some proteins as chaperonins has been discussed in Chapter 10. The *E. coli* chaperonin GroEL, a member of the hsp60 group, forms cylindrical aggregates with chambers in which proteins may fold (Box 7-A).[449] In a similar way a 16.5-kDa small heat-shock protein of *Methanococcus* *jannaschii* forms 24-subunit hollow spheres with octahedral geometry.[450] The structure of the 68- to 70-kDa hsp 70 (dnaK) in *E. coli* has been conserved with high homology throughout evolution. Using DNA chip technology 77 heat-shock genes have been identified recently in *E. coli*.[451]

Induction of heat-shock proteins depends upon a heat-shock promoter element (**HSE**) that binds an activating transcription factor **HSF**.[452–455] An increase in temperature not only induces synthesis of heat-shock proteins but represses synthesis of most other proteins. Thus, in *E. coli* or *Salmonella* a shift from 30°C to 42°C causes the appearance of 13 heat-shock proteins. At 50°C synthesis of almost all other proteins stops. In *E. coli* transcription of heat-shock genes is controlled by alternative factors, σ^{32} and σ^{E}.[456,456a]

In *S. typhimurium* a series of unusual nucleotides such as P^1, P^4-diadenosine 5'-tetraphosphate (ApppA or Ap_4A) accumulate. The related compounds ApppGpp (with a 3' pyrophosphate), ApppG, AppG, and ApppA also accumulate and appear not only in bacteria but in eukaryotes as well.[457] They are formed as a side product in aminoacylation of transfer RNAs (Chapter 29).[458,459] Lee *et al*.[460] proposed that Ap_4A and related nucleotides are formed in response to oxidative stress and serve as **alarmones** that signal the need to reduce transcription of most genes and to increase transcription of genes for protective proteins. However, there is little correlation between the Ap_4A concentration and the heat shock response.[461] See also p. 1715.



P^1,P^4-Diadenosine 5'-tetraphosphate (A*pppp*A or AP_4A)

Another type of **stress response protein** is related to glutathione *S*-transferase.[462] Cells of *E. coli* also accumulate a 15.8-kDa **universal stress protein** in response to most types of stress. It is an autophosphorylating phosphoprotein, a member of a little-studied group of phosphoproteins normally present in very small amounts.[463] Cells also respond to various other types of stress such as deprivation of glucose, hypoxia,[464] ultraviolet irradiation, presence of hydrogen peroxide,[465,465a] or change in osmotic pressure.[466] Salt-tolerant plants synthesize new proteins in response to increased salinity of water.[467]

Responses to low oxygen tension in tissues (**hypoxia**) are important to all aerobic organisms.[464,467a–d] In mammals transcription of hypoxia-responsive genes is regulated by **hypoxia inducible factors** HIF-1 and HIF-2. A subunit of HIF-1 undergoes 2-oxoglutarate-dependent hydroxylation on proline and asparagine residues. This may be a step in induction of ubiquitination and destruction of this component of the transcription factor complex. The human **von Hippel-Lindau** (VHL) tumor suppressor is a ubiquitin E3 ligase, which is also present in this transcription factor complex.[467b,d–h] One subunit of HIF-1 also interacts with the tumor suppressor p53.[467i] Together with the VHL protein and the elongation factor elongin (p. 1637)[467j] HIF participates in controlling both production of red blood cells and growth of new blood vessels (angiogenesis, Chapter 32, Section D).[467k]

Light-induced transcription. Light has a strong effect on transcription, especially in plants and photosynthetic bacteria. The photosystem II subunits D1, D2, CP47, and CP43 (see Fig. 23-34) are encoded in the chloroplast genome. D1 and D2 are unstable in light, and their rate of synthesis is increased as a result of elevated levels of transcription that are induced by a blue light response.[468] The light-induced conversion of phytochrome to its far-red absorbing form **Pfr** (Eq. 23-42) causes increased transcription of a variety of plant genes.[469] See also Chapter 23.

Homeotic genes and homeodomain proteins. Geneticists discovered in *D. melanogaster* and other species genes that establish the placement of antennae and legs on particular segments and in general to specify the body plan.[470,471] These homeotic genes encode a series of proteins containing a 60-residue **homeodomain**, a DNA-binding domain of the helix–turn–helix class.[470,472] Some homeodomain-containing proteins bind to DNA containing the octameric sequence shown in Table 28-1 and are known as **octamer-binding transcription factors** (Oct). One of these (Oct-2) is specifically needed for activation of immunoglobulin genes,[473,474] while Oct-1 binds to promoters of various other genes including that of histone H2B, U1, and U2 snRNAs.[475,475a] Another transcription factor, **Pit-1**, which activates genes for growth hormone and for prolactin in the pituitary, binds to the same octamer.[476,476a] Homeotic genes are considered further in Chapter 32 and immunoglobulin genes in Chapter 31.

7. Transcription by RNA Polymerases I and III

Promoters for RNA Pol I, like those of Pol II, lie upstream of the initiation site for transcription. At least two transcription factors have been identified[47,477–478a] and vary among species. The human factors bind to a G•C-rich DNA sequence in the –45 to

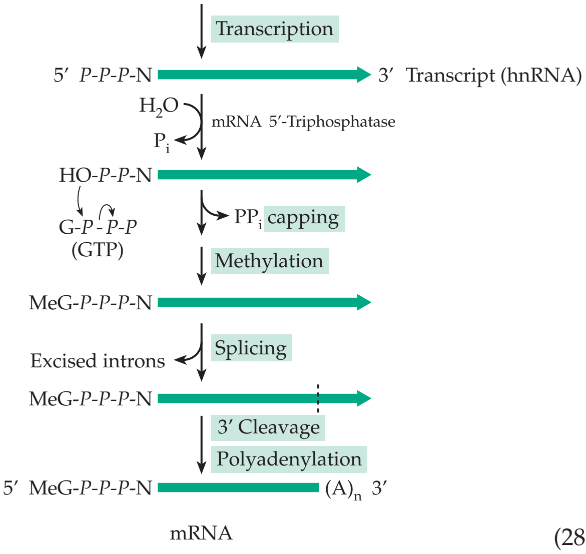
+20 region and to a related upstream control element **UCE** at −180 to −107.

In vertebrate cells transcription by RNA pol III is controlled by three types of promoter: class 1 (5S RNA), class 2 (tRNAs), and class 3 (U6 snRNA).[47,479,480] Initiation of transcription of class 1 genes requires the 40-kDa **transcription factor TF IIIA**.[479] This was the first eukaryotic transcription factor to be purified and characterized,[481,482] and one of the first of the zinc finger proteins to be recognized. It was a surprise to find that TFIIIA does not bind to the promoter region but to a DNA sequence in the center of the 5S RNA gene between positions +55 and +80.[47] The presence of the TFIIIA binding site in the middle of the gene it controls suggests that TFIIIA interacts with other proteins that bind in the promoter region to form a loop. TFIIIA is involved in initiation of transcription of the ~24,000 oocyte type 5S RNA genes of *Xenopus*, but the ~400 somatic type 5S RNA genes are not activated in the same way. TFIIIA is also unusual in that it binds not only to DNA but also to a specific CCUGG sequence in the transcribed 5S rRNA. This RNA is stored as a 7S ribonucleoprotein particle until it is needed after the oocyte is fertilized and begins rapid protein synthesis.[483]

After TFIIIA binds, proteins TFIIIC and then TFIIIB also bind. Although promoters of classes 2 and 3 do not require TFIIIA, all three classes depend upon TFIIIB and TFIIIC.[484] The TATA-binding protein TBF is one of three components present in TFIIIB, which may be regarded as the true initiation factor.[47,485] Both TFIIIA and TFIIIC can be described as **assembly factors**.[47] A silkworm RNA pol III has been reported to require a transcription factor consisting of RNA.[486]

8. Elongation, Modification, and Termination of Transcription

As in prokaryotic transcription (Fig. 28-4) elongation by RNA polymerase II occurs within a transcription bubble of ~20–30 nucleotides in length.[487] Most transcriptionally active DNA is still in the form of nucleosomes, which must be unwound as the transcription bubble moves. Details are still uncertain.[269,488] All of the major steps in processing of the pre-mRNA transcripts, which include capping, splicing, 3′-end cleavage, and polyadenylation (Eq. 28-6), are coupled to transcription. This is apparently accomplished, in part, by physical connections of the necessary proteins to the CTD domain of RNA polymerase II.[304,312a,b] While pre-mRNA usually undergoes all of the steps of Eq. 28-6, rRNA and tRNAs are not capped or polyadenylated and often are not spliced.



(28-6)

Elongation of the RNA chain depends upon five **general elongation factors**, designated **P-TEFb**, **SII** (TFIIS), **TFIIF**, **Elongin** (SIII), and **ELL**.[489–492] Many of these subunits of the transcriptosome function in the suppression of pausing in the uneven movement of the template DNA through the complex. Some may have to deal with torsional strain induced in the transcribed DNA.[493] Elongin is a heterotrimer of subunits A, B, and C.[494] Most of the other elongation factors are heterodimers,[489] but SII is a monomer that contains two conserved Zn^{2+}-binding **zinc ribbon** motifs.[492] It promotes cleavage of the growing transcript within the transcription bubble at stalled sites, allowing transcription to be restarted from a fresh 3' end. Defects in elongation factors P-TEFb, ELL, and **CSB** (Cockayne syndrome complementation group) have been correlated with human diseases including cancer.[495]

Despite the complexity of the processes represented by Eq. 28-6, a yeast cell is able to transcribe genes at rates of about one in every 6–8 s.[496] This can be compared with a rate of about once in 2–3 s for RNA polymerase of *E. coli*.

The least well defined step in eukaryotic transcription is termination, which follows the various steps of processing, discussed in Section D. The final 3' end of processed transcripts of RNA pol II action in mammals is marked by the sequence AAUAAA, which is found about 10–30 nucleotides upstream from the end.[497] This is usually followed by a polyU or GU-rich sequence.[498] In yeast the termination and polyadenylation signals are less clear.[499,500] The initial transcript almost always continues beyond the AAUAAA signal, sometimes for hundreds of nucleotides. However, the excess RNA is rapidly degraded by a large complex of proteins.[499–502] The precise 3' end cleavage is energy-dependent, requiring creatine phosphate rather than ATP or GTP.[500] Transcription termination by RNA polymerases I and III is more like that of bacteria. Terminator sequences are present in the DNA, and terminator proteins interact with them.[480,503–504b]

9. Conformational Properties of RNA

Newly formed RNA transcripts fold quickly into structures of complex shapes,[505–508b] the folding being influenced by interactions with proteins and with other RNA molecules. RNA chains are flexible, with many sterically allowed conformations.[509] As with proteins folding probably begins with a nucleation event, perhaps involving monovalent or divalent metal ions[510–512] and continues rapidly.[513] Folding is affected by hydrogen bonding,[514] base stacking, and binding of ions, and by formation of pseudoknots (Fig. 5-29).[515]

As is apparent from the structures of tRNAs (Figs. 5-30, 5-31, 28-20), the *Tetrahymena* self-splicing ribozyme (Fig. 12-26), and ribosomal RNA structures (Fig. 29-2), a large fraction of a folded RNA exists as hairpin or **stem –loop** structures. These are A-type structures with largely Watson–Crick base pairs. However, mismatched pairs, triples, and quadruples of bases are also formed. Recently discovered RNA structural elements include **base platforms**, formed by pairing of adjacent bases,[516] interdigitation of unpaired bases (also seen in DNA; Fig. 5-27), and wobble pairs (Chapter 5).[517] Hydrogen bonding between riboses of consecutive nucleotides in two strands may help to form a **ribose zipper**.[437,507] Guanine-rich tetraplexes (pp. 208, 227), cytosine-rich i-motif structures (p. 228), and water-mediated U•C base pairs also arise in RNAs.[517a–c]

The terminal loops, which usually contain the consensus sequence **GNRA**, may constitute up to one-third of the entire molecule.[518] These loops interact with many binding proteins, such as those in the snRNA-protein particles.[519] GNRA loops may also dock into the shallow groove of RNA helices.[518] Adenosines that are not paired in double helices, e.g., those in GNRA loops, are able to interact in a variety of ways with other parts of an RNA molecule or with other molecules. They are involved in helix packing interactions in virtually every RNA studied.[517d] Although examples are still rare, specific mRNA molecules may provide binding pockets for small regulatory molecules, e.g., amino acids and thiamin.[517e]

D. Processing of Eukaryotic RNAs

All RNA found in eukaryotes undergoes major alterations prior to functioning. The cutting out of rRNA and tRNA molecules from larger precursors resembles that in bacteria, but subsequent processing is much more complex, as is that of mRNA.

1. Ribosomal RNA

Eukaryotic ribosomes contain four pieces of RNA (Tables 5-4 and 29-1), which are usually designated by their sedimentation coefficients. The 18S, 5.8S, and 28S RNAs are encoded as single transcriptional units with spacers separating the sequences that encode the mature RNAs. A typical animal cell contains several hundred copies of this transcriptional unit, all located in the DNA in the nucleolus (Fig. 28-15), and each having its own set of promoter sequences, enhancers, and transcription factors.[47,520–522] The promoter sequences vary substantially among different species.[523] The primary transcripts from these units are the sole product of RNA polymerase I.

Electron micrographs of portions of unwound cores of nucleoli have revealed fibrils of RNA coated with protein growing from the DNA strands of the pre-rRNA genes (Fig. 28-16), ~80–100 RNA chains of different length being transcribed concurrently from a single gene. The overall gene length in the electron microscope is 2.3 µm, only a little less than the calculated length for a fully extended DNA molecule in

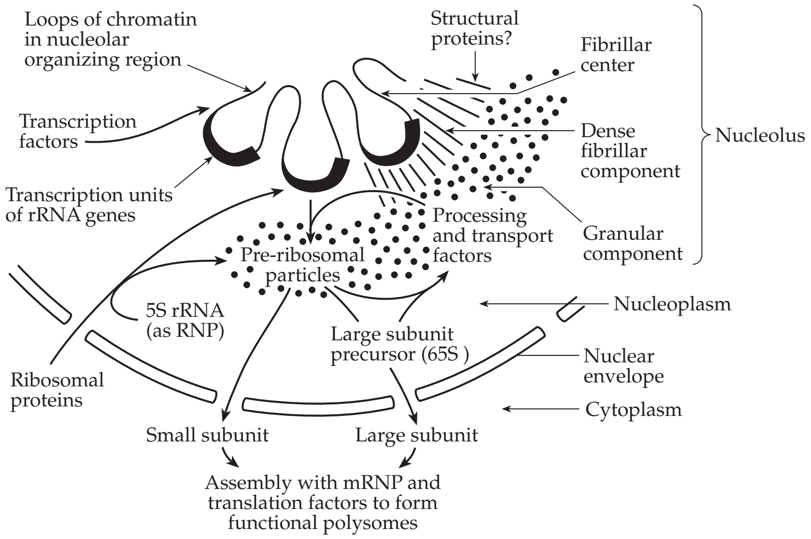

Loops of chromatin in nucleolar organizing region
Structural proteins?
Fibrillar center
Transcription factors
Dense fibrillar component
Nucleolus
Transcription units of rRNA genes
Processing and transport factors
Granular component
Pre-ribosomal particles
5S rRNA (as RNP)
Nucleoplasm
Large subunit precursor (65S)
Nuclear envelope
Ribosomal proteins
Cytoplasm
Small subunit
Large subunit
Assembly with mRNP and translation factors to form functional polysomes

Figure 28-15 Steps in ribosome formation in the nucleolus. From Sommerville.[524] Abbreviations are NOR, nucleolar organizing regions; RNP, ribonucleoproteins.

the B form. However, judging by the lengths of the transcripts formed, the pre-rRNA chains are folded extensively.

The primary eukaryotic rRNA transcripts extend several hundred nucleotides past the 3' termini of the mature 28S RNA molecules. As the primary transcripts, which are formed in the core region of the nucleolus, move away into the outer cortex of the nucleolus, cleavage occurs in a number of steps (Eq. 28-7).[525,526] Electron microscopy provided the first direct confirmation of the relationship of one pre-rRNA molecule to another (Fig. 28-17).[527,528] The 18S portion of the 45S RNA seen in Fig. 28-17 lies nearest to the 5'-end just as does the 16S rRNA in the large transcript of the prokaryotic rRNA genes (Section A,7).

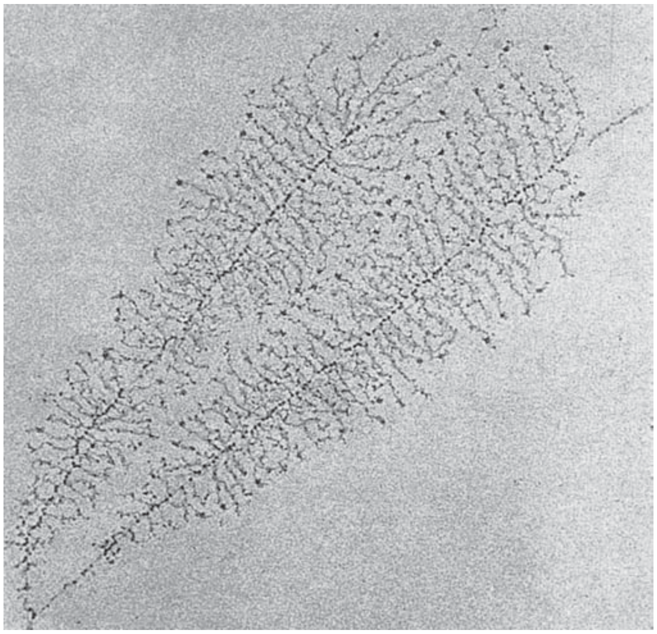
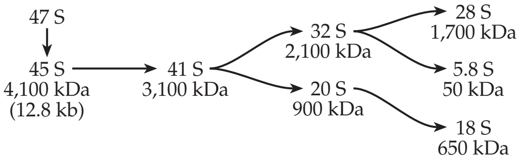


Figure 28-16 Ribosomal RNA genes from an embryo of *Drosophila melanogaster* in the process of transcription. The densely packed ribonucleoprotein strands are shorter where transcription begins and contain increasing lengths of rRNA where transcription has proceeded for a longer time. Also note the characteristic granular knobs at the tips of the strands. From Miller.[85]



(28-7)

As is indicated in Fig. 28-15, transcription is thought to occur from the loops of DNA that form the nucleolar organizing region. The 100-kDa **nucleolin**, the major protein of the nucleolus, binds to the nontranscribed spacer sequences in the DNA.[529–530] It also binds to the newly formed transcripts, as do various proteins that enter the nucleus from the cytoplasm.[524,531] More than 270 proteins, many of which participate in synthesis of ribosomes, have been detected in the nucleolus.[531a] Some of these proteins, acting together with the snoRNAs discussed in the next section, catalyze hydrolytic cleavage of the pre-rRNA molecules. For completion of pre-ribosomal particles additional protein molecules enter the nucleolus and associate with the pre-rRNA particles, then diffuse out of the nucleus.

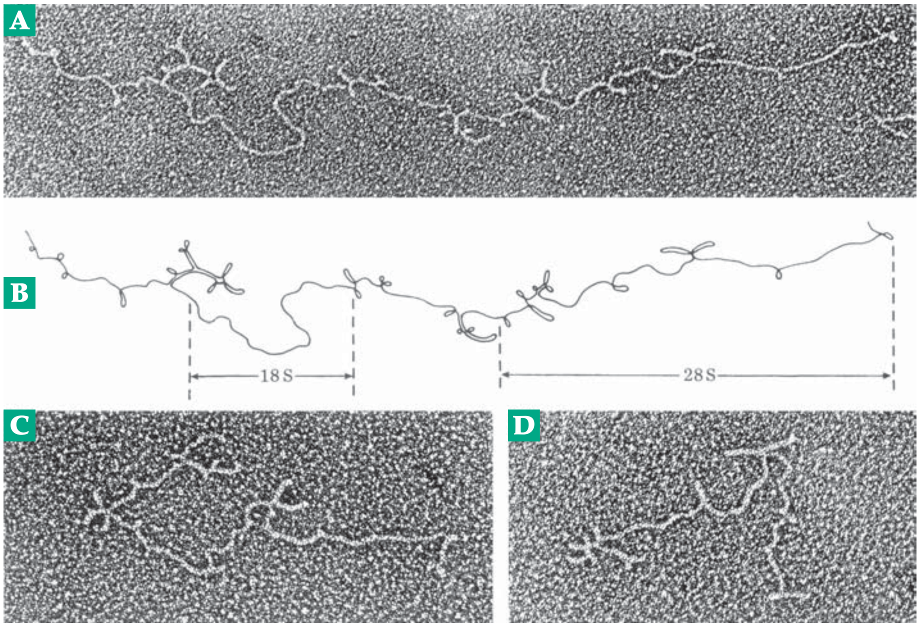


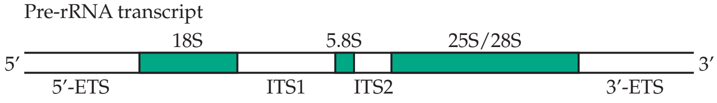
Figure 28-17 (A) Electron micrograph of the 45S precursor of rRNA from HeLa cells after spreading from 80% formamide and 4 M urea. The molecule is shown in reverse contrast. (B) Tracing of molecule in (A) showing several regions of secondary structure as hairpin loops. The 28S and 18S rRNA regions are indicated. (C) 32S rRNA. (D) 28S rRNA. Notice that the same secondary structure can be seen in the 28S RNA as in its 32S and 45S precursors. From Wellauer and Dawid.[527]

2. Small Nuclear, and Nucleolar, and Cytoplasmic RNAs

The nuclei of all eukaryotic cells contain a group of small nuclear RNAs (**snRNAs**), some of which (**snoRNAs**) are localized to the nucleolus and function there.[532–535e] At least ten of these are always present (Table 28-3), and yeast appears to contain more than 100. Some are present in small amounts, and it may be that a large number of snRNAs will also be found in other eukaryotes. The uridine-rich or U series of snRNA are especially abundant (10^4–10^6 molecules per nucleus). Several of them (U1,2,4–6,11,12,16,18) function in RNA splicing. Species U1–U10 contain from ~60 to ~215 nucleotides but some snRNAs from yeast have over 300 nucleotides and one over 1000.[532] The most abundant snRNAs, the metabolically stable U1, U2, U4, U5, and U6, are exported from the nucleus. In the cytoplasm each snRNA associates with a complex of several proteins to form ribonucleoprotein particles (**snRNP particles**). The proteins, known as **Sm proteins**, range in mass from 11 to 70 kDa[536] and are designated B, B′, D1, D2, D3, E, F, and G.[535,537,537a] The proteins may associate to a complex B_2D_2EFG, which then binds the snRNAs. Whereas many proteins are found in most of the snRNP particles, some are associated with specific RNAs. Thus, mature U1 snRNP may have a partial stoichiometry $U1A_2B_2C_2D_4E$.[536,538] The snRNA particle, known as U4 / U6snRNP, contains two snRNAs. After returning to the nucleus the four snRNP complexes U1, U2, U5, and U4 / U6, together with the pre-mRNA molecules, associate via an ordered assembly pathway that gives rise to the large **spliceosomes** in which the removal of introns takes place.[539] Some of the snRNP complexes, e.g. the spliceosomal U6 snRNP and snRNPs involved in mRNA degradation, contain Sm-like (Lsm) proteins.[537a]

More recently small 20–25 nucleotide antisense micro RNAs (**miRNAs**) have become recognized as important in control of the breakdown of mRNA, in formation of heterochromatin, and in control of specific stages in development (Chapter 32), and in cellular defense mechanisms (Chapter 31, Section G).[1a,537b] In animals the 21–23 nt miRNAs are formed by enzymatic cleavage of double-stranded dsRNA by a ribonuclease III-type enzyme called **Dicer**.[171a,537c] In the resulting fragments, which have 5'-phosphate termini, one strand is antisense to a "target" sequence in mRNA. **Small temporal RNAs** (stRNAs) that guide development are cut by Dicer from RNA stem-loop structures. In other cases, e.g., in formation of the siRNAs that silence individual genes, an RNA-dependent RNA polymerase acts on an RNA transcript to form a long dsRNA that can be cleaved by Dicer. This often happens with foreign RNAs, e.g., from viruses or trans-genes. Unwinding of the small fragments formed by Dicer, in a process that may require ATP, provides the single-stranded antisense siRNA molecules. Such antisense molecules bind to their target RNAs together with a group of proteins to form **RNA-induced silencing complexes** (RISCs). These complexes contain an RNAse (different from Dicer) that cuts the dsRNA of an siRNA-target complex, initiating destruction of the mRNA and silencing of its gene.[1a,171a,539a] This **RNA-interference** is widely utilized by plants and animals. It has also provided the basis for development of many practical tools for understanding gene sequences, for genetic engineering, and for design of new drugs.

Processing of ribosomal RNA. Transcripts of rRNA genes vary in size from ~35 to 47S (6–15 kb) and often contain spacer regions at both ends as well as between the 18S, 5.8S, and 28S sequences. For example, human 47S transcripts have 414 extra nucleotides at the 5' ends.[540] One group of snoRNAs participates in the hydrolytic cleavage of pre-rRNA. RNAs U3, U14, snR10, snR30, as well as MRP RNA (Table 28-3) are always required.[541–545] Also needed is U22, an intron-encoded RNA.[531] The reactions represented by Eq. 28-7 are best known for *Saccharomyces cerevisiae*.[525,541–541c] Four spacers, the **5'-external transcribed** spacer (5'-ETS), the first and second **internal transcribed spacers** (ITS1 and ITS2), and the 3'-external transcribed sequence (3'-ETS), must be removed.[525,541,546] Removal of the 5'-ETS depends upon the snoRNA U3, which contains two highly conserved sequences able to form base-paired structures with the 5' end of the 5.8S rRNA region of the pre-rRNA gene.[542–544,546a] Both U8 and U22 are also needed for cleavage of the pre-rRNA.[546b] Although the exact functions of snoRNAs and their associated proteins in the cleavage of pre-rRNA are still uncertain, these RNAs probably act as guide molecules for the cleavage reactions. They may have ribozyme activity and may perhaps be chaperones.



Cleavage at the 5' end of the 5.8S region requires RNase MRP, a relative of the RNase P that cleaves at the 5' ends of tRNAs (Fig. 28-10).[525,547] MRP (**mitochondrial processing protein**) also cleaves primers for mitochondrial DNA replication. The importance of the enzyme is emphasized by the existence of a hereditary defect in the MRP RNA (Table 28-3) that causes abnormalities in bone, cartilage, hair, and the immune system.[547a] Most bacterial rRNA genes have a tRNA gene in the position corresponding to that of 5.8S RNA in eukaryotes. This provides an RNase P-dependent cleavage mechanism, which is alternative to action of other nucleases.[525]

The most abundant protein in the fibrillar regions of the nucleus, where the early stages of pre-rRNA processing occur, is **fibrillarin**.[541,548,549] Many of the snoRNAs are closely associated with this protein. Fibrillarin is also well known as an autoantigen, which can induce formation of destructive antibodies that cause **scleroderma** (Chapter 31, Section F).

TABLE 28-3
Some Eukaryotic Small Nuclear and Nucleolar RNA Molecules

Designation	Number of nucleotides Vertebrate	Yeast
U1[a]	164	568
U2[a]	188	1175
U3[b]	206 – 228	333
U4[a]	142 – 146	160
$\mathrm{U4_{atac}}$[c]	131	
U5[a]	116 – 118	183 or 196
U6[a,d]	116	214
$\mathrm{U6_{atac}}$[c]	125	
U7[e]	57 – 58	
U8[b]	136 – 140	
U9	130	
U10	60	
U11[c]	131 – 135	
U12[c]	150	
U13[b,f]	105	
U14[b,f]	87– 96	125 – 128
U18[b,f]	67– 70	102
X[b]	150	
U20[b,f]	80	
U21[b,f]	93	
U22(Y)[b]	125	
U24[b,f,g]	77	
U32 – U40[b,f,h]		
SnR10[b]		245
SnR30[b]		605
SnR38[f]		93
SnR39[f]		85
SnR40[f]		96
SnR41[f]		
MRP RNA (RNA 7– 2)	260 – 280	339

[a] Major spliceosomal RNAs.

[b] Fibrillarin-associated SnoRNAs that function in pre-ribosomal RNA processing. See Morrissey, J. P., and Tollervey, D. (1995) *Trends Biochem. Sci.* **20**, 78–82.

[c] Function in AT–AC spliceosomes. See Tarn, W.-Y., and Steitz, J. A. (1997) *Trends Biochem. Sci.* **22**, 132–137 and Fournier, M. J., and Maxwell, E. S. (1993) *Trends Biochem. Sci.* **18**, 131–135.

[d] γ-Monomethyl cap.

[e] Required for 3'-end formation in histone mRNAs.

[f] SnoRNAs with long complementarities to rRNA. C and D sequences are present. See Bachellerie, J.-P., Michot, B., Nicoloso, M., Balakin, A., Ni, J., and Fournier, M. J. (1995) *Trends Biochem. Sci.* **20**, 261–264 and Nicoloso, M., Qu, L.-H., Michot, B., and Bachellerie, J.-P. (1996) *J. Mol. Biol.* **260**, 178–195.

[g] Polyadenylated, noncoding.

[h] Participate in 2'-*O*-ribose methylation.

Modification guide RNAs. A second group of snoRNAs function in methylation, pseudouridine formation, and other RNA modifications (Section 6).[174,525,541,548] These snoRNAs have long sequences complementary to highly conserved regions of pre-rRNA, enabling them to form helical regions that may guide the docking with modification enzymes. Many of them also contain characteristic conserved sequences: **C**, 5'-UGAUGA; **D**, 5'-CUGA; **H**, 5'-AnAnnA; and 5'-ACA. Sequences C and D are present in snoRNAs that act as methylation guides, while the H and ACA sequences characterize guide RNAs for pseudouridine formation.[174]

Transcription and processing of snRNAs and snoRNAs. In higher organisms each of the snRNAs has several genes,[538] e.g., there are 50 –60 U1 genes in the human haploid genome. However, in yeast there are often single copies.[541] All of the snRNA genes, except for that of U6,[550] are transcribed by RNA polymerase II. The transcripts, which are capped at the 5'-end but are not polyadenylated, pass into the cytoplasm, where they undergo further processing and become associated with proteins. The 3' ends are trimmed, the 7-methylguanosine of the cap is methylated further, and methylation may occur on other bases as well.[551] SnoRNAs are not capped. It was a great surprise to discover that many of the snoRNA genes lie within introns that occur in abundantly expressed genes for functionally unrelated proteins.[173,541,548]

Patients with the autoimmune disease systemic **lupus erythematosus** make autoantibodies directed against the Sm proteins of snRNP particles.[552,553] Antibodies from different patients vary in their specificities, making these antibodies a useful tool in the isolation and study of snRNAs and their protein complexes.[552,554]

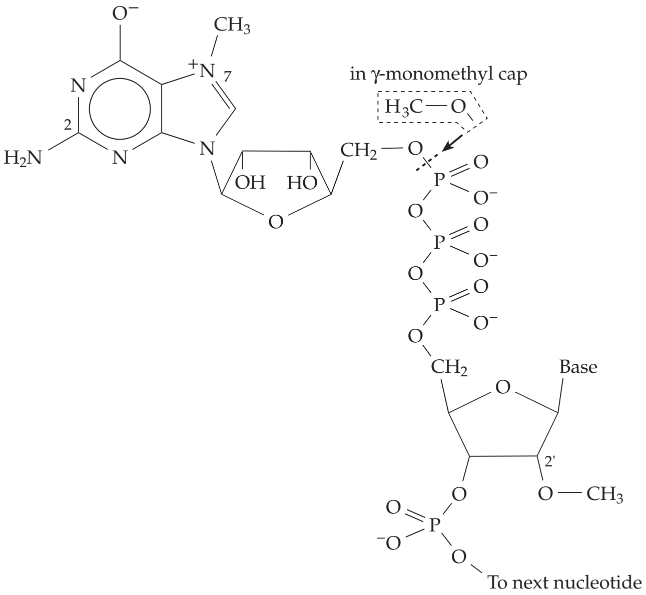
3. Processing of 5S RNA and tRNAs

The genes for 5S ribosomal RNA and all of the tRNAs are transcribed by RNA polymerase III. In the yeast genome the 5S RNA genes are located in the spacers between the transcriptional units containing the other rRNAs. However, in animals the 5S RNA genes are separate from the other rRNA genes and are not located in the nucleolus. In *Drosophila* ~500 copies of the 5S RNA gene are located in the right arm of chromosome 2. In *Xenopus* ~400 genes are active in somatic cells and another 24,000 only in oocytes. These are arranged as large transcriptional units, each containing several thousand copies of the 120-bp 5S RNA gene separated by 720-bp spacers. Cleavage and trimming are required to form the final product, but there are usually no introns to be removed.

Most eukaryotic tRNAs are formed from monomeric precursors, each gene acting as a transcriptional unit. Processing is similar to that in bacteria (Section A, 7). Eukaryotic RNase P usually cleaves the 5' end,[555–558c] and another enzyme cuts at the 3' end.[542,556] The 3' CCA sequence of the mature tRNA is usually not present in the primary transcript but is added.[559,559a,b] As in bacteria (p. 1620) extensive modification of bases also occurs in the tRNA precursors of eukaryotes.[235,560–562] Many tRNA genes contain introns, which must be removed by splicing (Section 5).

4. Messenger RNA, Caps, and Polyadenylate Tails

The first processing event (Eq. 28-6) for most of the pre-mRNA and snRNA transcripts made by RNA polymerase II is addition to the 5' end of a "cap," a terminal structure containing 7-methylguanosine from which a proton has dissociated to form a dipolar ion.[563–565] The cap structure may be abbreviated 5′-m^{7}G(5')pppNm —. The 5' terminal ribose is often methylated on O2', as shown below. More complex caps are methylated at additional sites, e.g., the guanine may be dimethylated on the 2-NH$_2$ group.[551] Most snRNAs, including the U1–U5 and U7–U13 snRNAs, have such 2,2,7-trimethylguanosine caps.[551,566] Many viral transcripts, including those of the much-studied vaccinia virus, have similar caps.[564] However, U6 and some other snRNAs, which are transcribed by RNA polymerase III, have γ-monomethyl (me-ppp) caps[567] or undergo a series of additions and deletions of uridylate residues at the 3' ends.[566,568]



Cap synthesis occurs as follows. The 5' end of an RNA transcript initially contains a triphosphate group arising from the fact that a nucleotide triphosphate serves as the primer in initiating transcription. The terminal phospho group is removed by a triphosphatase leaving a diphosphate, which is then guanylated by GTP (Eq. 28-6).[569] The capped transcripts are exported from the nucleus, after which additional methylation may follow.[570]

The cap structure affects several processes.[565,571] A family of cap-binding proteins recognize the structure and may facilitate splicing as well as export from the nucleus.[571a] The cap is very important for ribosome binding and initiation of translation (Chapter 29). The trimethylated caps of snRNAs, on the other hand, may be signals for retention in the nucleus where they function.[563] Following capping is the often elaborate process of splicing to remove introns (Section 5).

A poly(A) "tail" consisting of ~250 residues of adenylic acid is added next by poly(A) polymerase, a component of an enzyme complex that also cleaves the RNA chains.[545,571b] Most eukaryotic mRNA is polyadenylated with the exception of that encoding histones. The function of the poly(A) is unclear. It is needed for transport of mRNA out of the nucleus, but it does confer a greatly increased stability to the mRNA in the cytoplasm where the adenylate units are gradually removed.[307,308] In contrast, in chloroplasts and plant mitochondria polyadenylation is required for rapid degradation of mRNA.[571c,d] Polyadenylation may also increase the efficiency of translation.[572] Polyadenylation occurs rapidly within ~1 min after transcription is completed.

Mature mRNA molecules vary in lifetime.[573,574] Some last for hours or days. Among the latter are mRNAs of maternal origin that accumulate in oocytes and are utilized during the early stages of embryonic development.[575,575a,b] Other mRNAs, e.g., transcripts of the *c-fos* and *c-myc* proto-oncogene products, have half-lives of 30 min.[573,576] Some mRNA molecules are degraded while attached to ribosomes in response to recognition of the synthesized peptide (Chapter 29). Longer lived mRNA molecules may be protected by RNase inhibitors.[577] Hydrolytic removal of caps often initiates degradation, and Sm-like protein complexes participate.[578–579a]

5. Splicing

An essential modification of the precursor forms of large rRNA, tRNA, and mRNA molecules is the splicing out of intervening sequences. This occurs prior to polyadenylation of mRNA and is usually slow, the half-life of introns varying from a few seconds to 10–20 min.[551] Splicing occurs by at least four distinctly different pathways.[47,580]

Self-splicing RNA. The precursor to the 26S rRNA of *Tetrahymena* contains a 413-nucleotide intron, which was shown by Cech and coworkers to be self-splicing, i.e., not to require a protein catalyst for maturation.[581,582] This pre-rRNA is a ribozyme with true catalytic properties (Chapter 12). It folds into a complex three-dimensional structure which provides a binding site for free guanosine whose 3'–OH attacks the phosphorus at the 5' end of the intron as shown in Fig. 28-18A, step *a*. The reaction is a simple displacement on phosphorus, a transesterification similar to that in the first step of pancreatic ribonuclease action (Eq. 12-25). The resulting free 3'–OH then attacks the phosphorus atom at the other end of the intron (step *b*) to accomplish the splicing and to release the intron as a linear polynucleotide. The excised intron undergoes a third transesterification reaction, of uncertain significance (step *c*), to form a circular polynucleotide and a short displaced 15-residue oligonucleotide. The *Tetrahymena* pre-rRNA intron is a member of a group of similar **Group I introns**, many of which are found in fungal mitochondrial pre-mRNA and pre-rRNA.[583] All are excised by a similar mechanism. Many are self-splicing, but others require a protein catalyst.[584–585a] A similar splicing sequence is involved in removal of a 1017-nucleotide intron from the thymidylate kinase gene of phage T4 and other introns in T-even phage. The later are among the relatively rare introns in prokaryotic systems.[583]

All Group I introns have several small conserved sequences, which suggest a common folded tertiary structure as is indicated in Fig. 28-19A. The conserved sequences are labeled **A**, **B**, **9L**, **2**, **9R**, and **9R′**. **A** is paired with **B**, **9R** with **9R′,** and **9L** with **2**. The sites of chain cleavage at the 3' and 5' ends of the intron are indicated by the heavy arrows. They are evidently selected by formation of the double-stranded regions.[586–588]

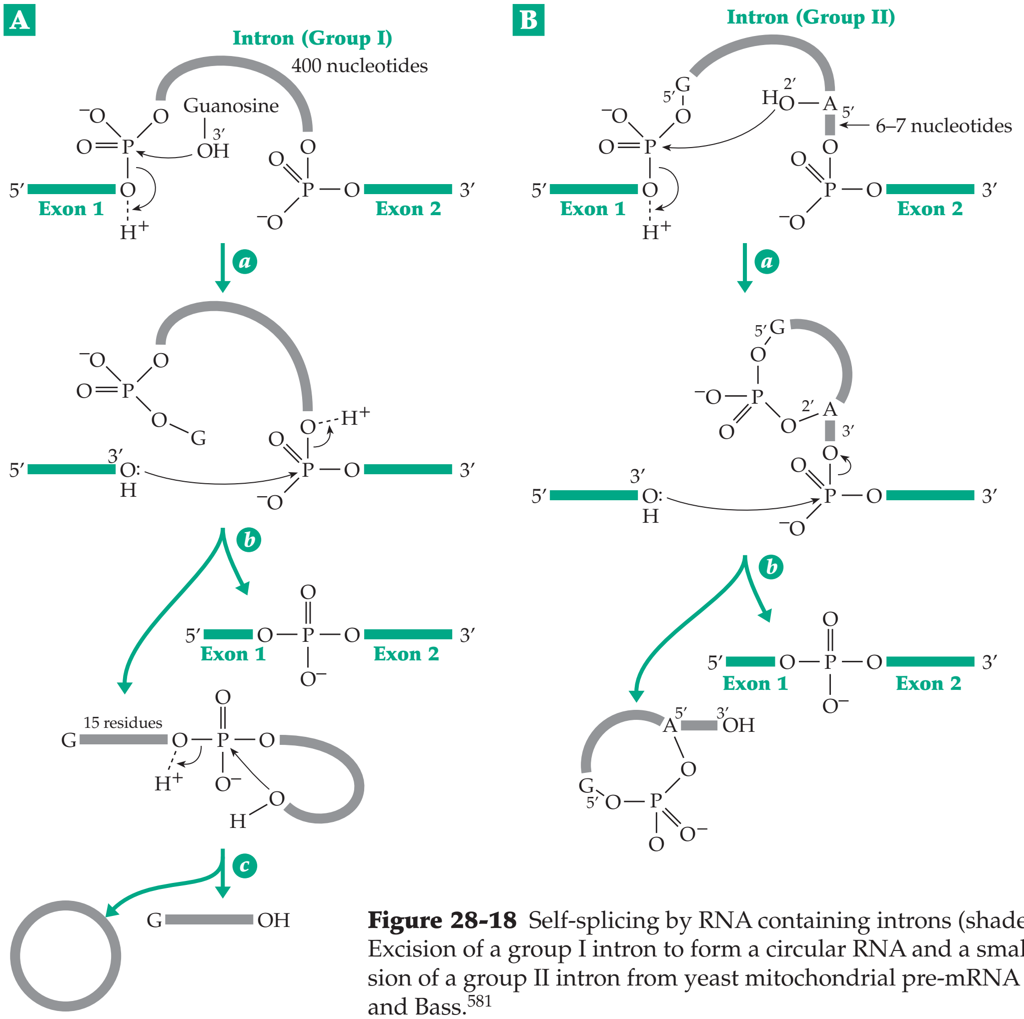
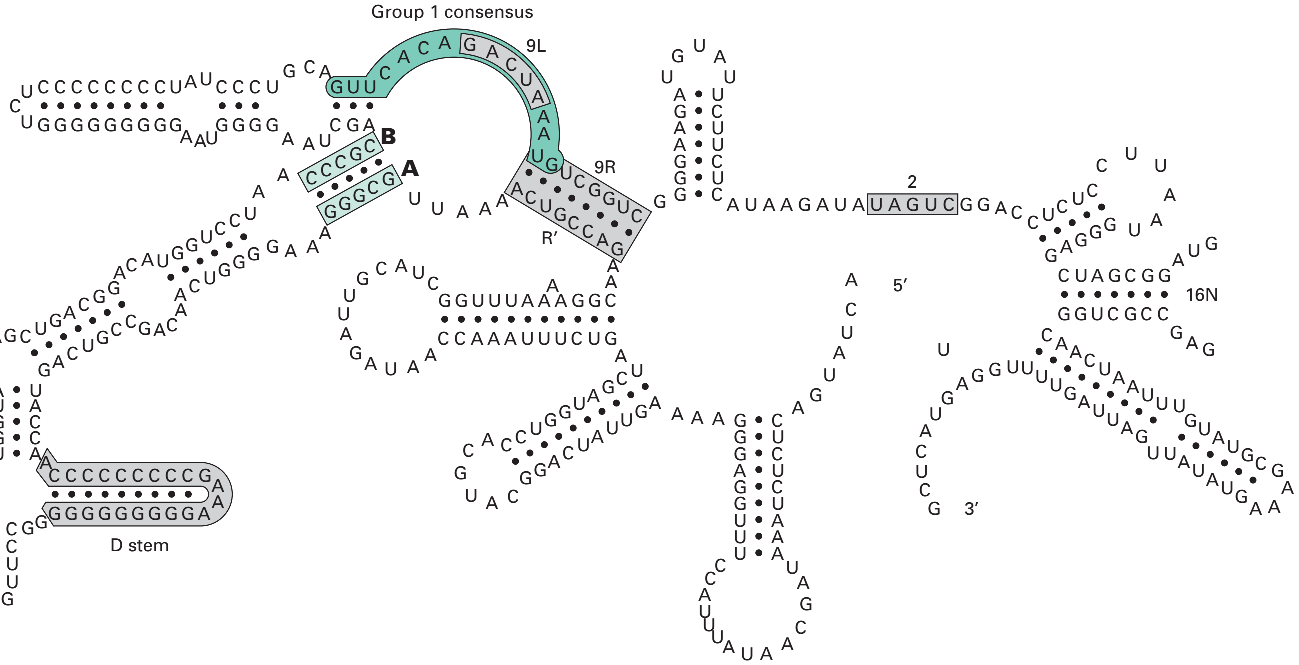


Figure 28-18 Self-splicing by RNA containing introns (shaded) of groups I or II. (A) Excision of a group I intron to form a circular RNA and a small oligonucleotide. (B) Excision of a group II intron from yeast mitochondrial pre-mRNA as a circular RNA. See Cech and Bass.[581]

Another type of intron (Group II) also undergoes self-splicing.[589–590d] The best known example is the last intron in the yeast mitochondrial pre-mRNA. The splicing pathway shown in Fig. 28-18B is similar chemically to that of the group I introns. However, the initial attack is not by free guanosine, but by the 2' OH of an internal adenosine, the intermediate product having a **lariat structure** with a loop at the end. Otherwise, processing is similar to that of group I introns. The same pathway with lariat formation is followed by the more widely used removal of introns from pre-mRNA in spliceosomes (see Fig. 28-22).

A



B

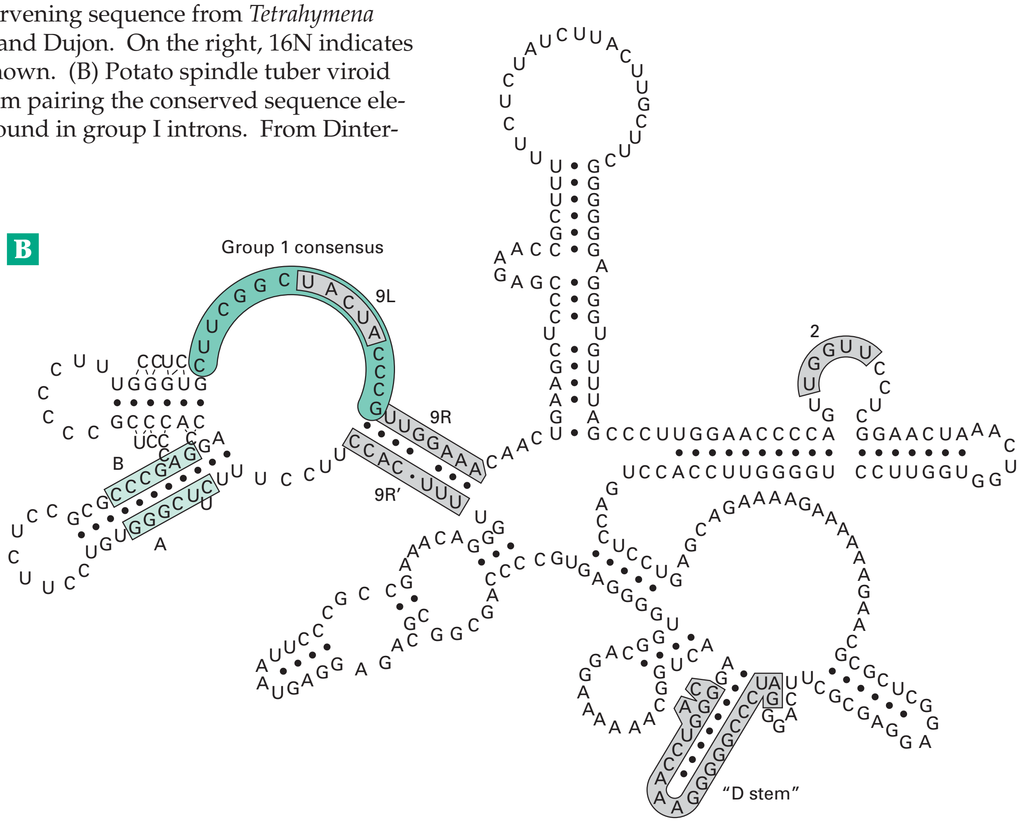


Figure 28-19 (A) Proposed secondary structures of a self-splicing Group I intervening sequence from *Tetrahymena* according to Michel and Dujon. On the right, 16N indicates 16 nucleotides not shown. (B) Potato spindle tuber viroid structure derived from pairing the conserved sequence elements that are also found in group I introns. From Dinter-Gottlieb.[596]

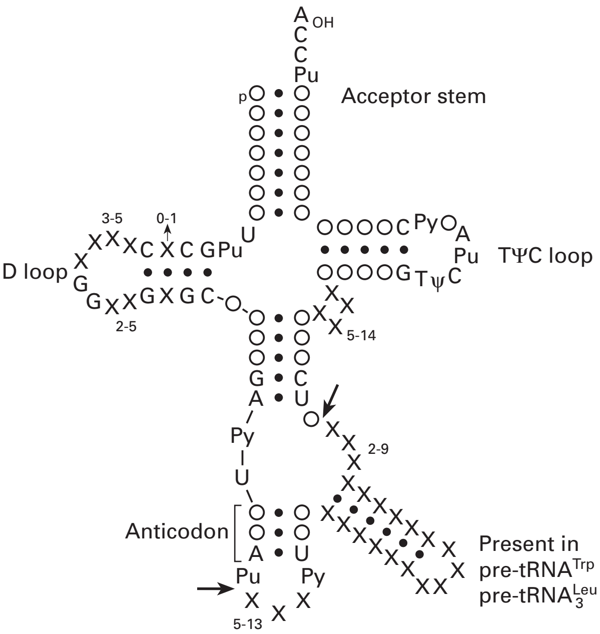

A OH
C
C
Pu
p
Acceptor stem
3-5
0-1
D loop
2-5
U
GPu
Py
A
Pu
TΨC loop
TψC
5-14
2-9
Py
U
Anticodon
Pu
Py
5-13
Present in
pre-tRNA^{Trp}
pre-tRNA$_3^{Leu}$

Figure 28-20 Composite structure representing several tRNA precursors arranged in a similar secondary structure (see also Fig. 5-30). The arrows indicate splice points. Variable positions are designated (○) for the mature tRNA and (X) for the intervening sequence and also in loops where insertions or deletions occur. From Ogden *et al.*[603]

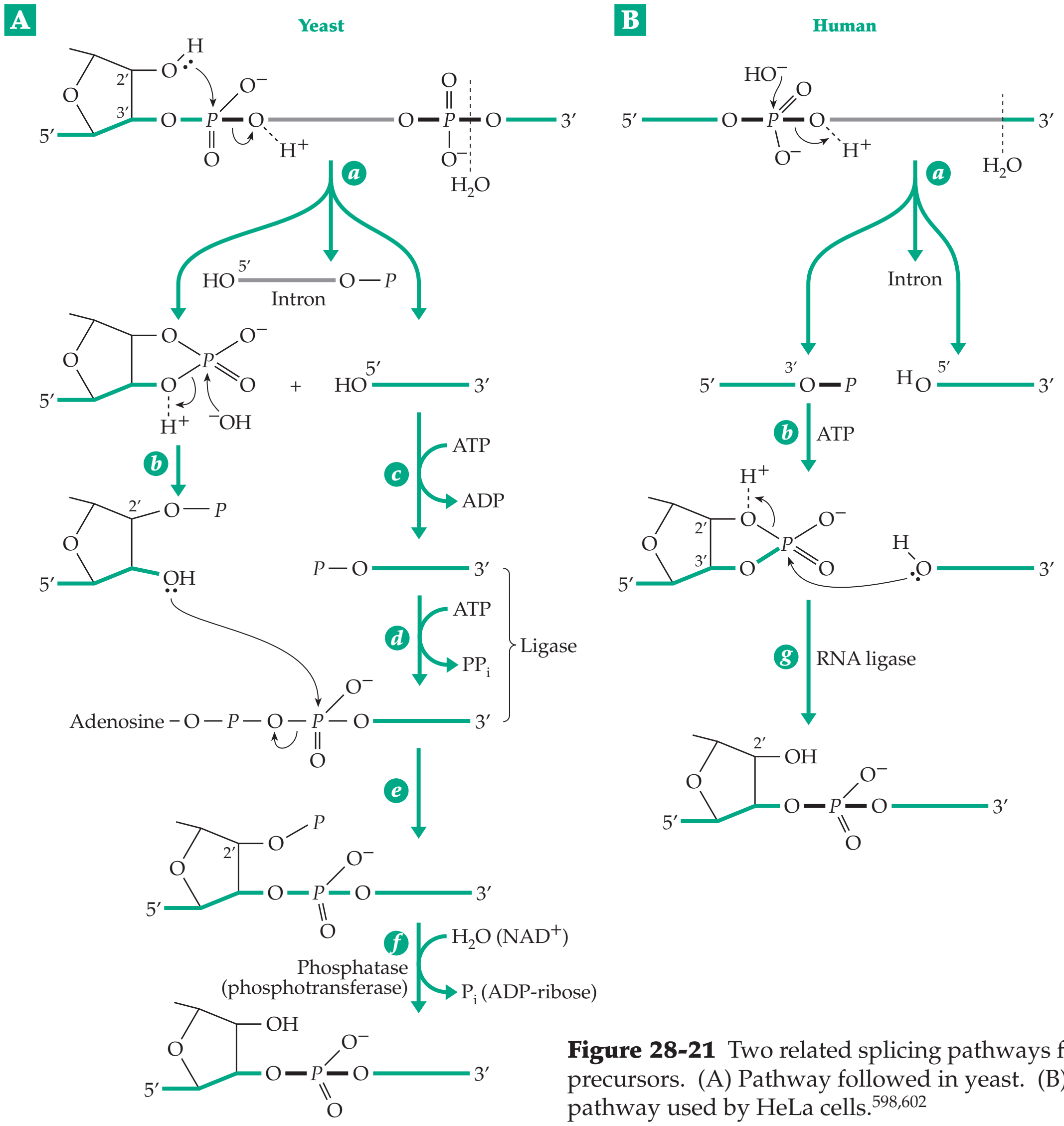

A
Yeast
B
Human
a
Intron
b
c
ATP
ADP
d
ATP
PP_i
Ligase
Adenosine
e
f
H_2O (NAD^+)
Phosphatase
(phosphotransferase)
P_i (ADP-ribose)
g
RNA ligase
H_2O
H^+
5′
3′

Figure 28-21 Two related splicing pathways for tRNA precursors. (A) Pathway followed in yeast. (B) Related pathway used by HeLa cells.[598,602]

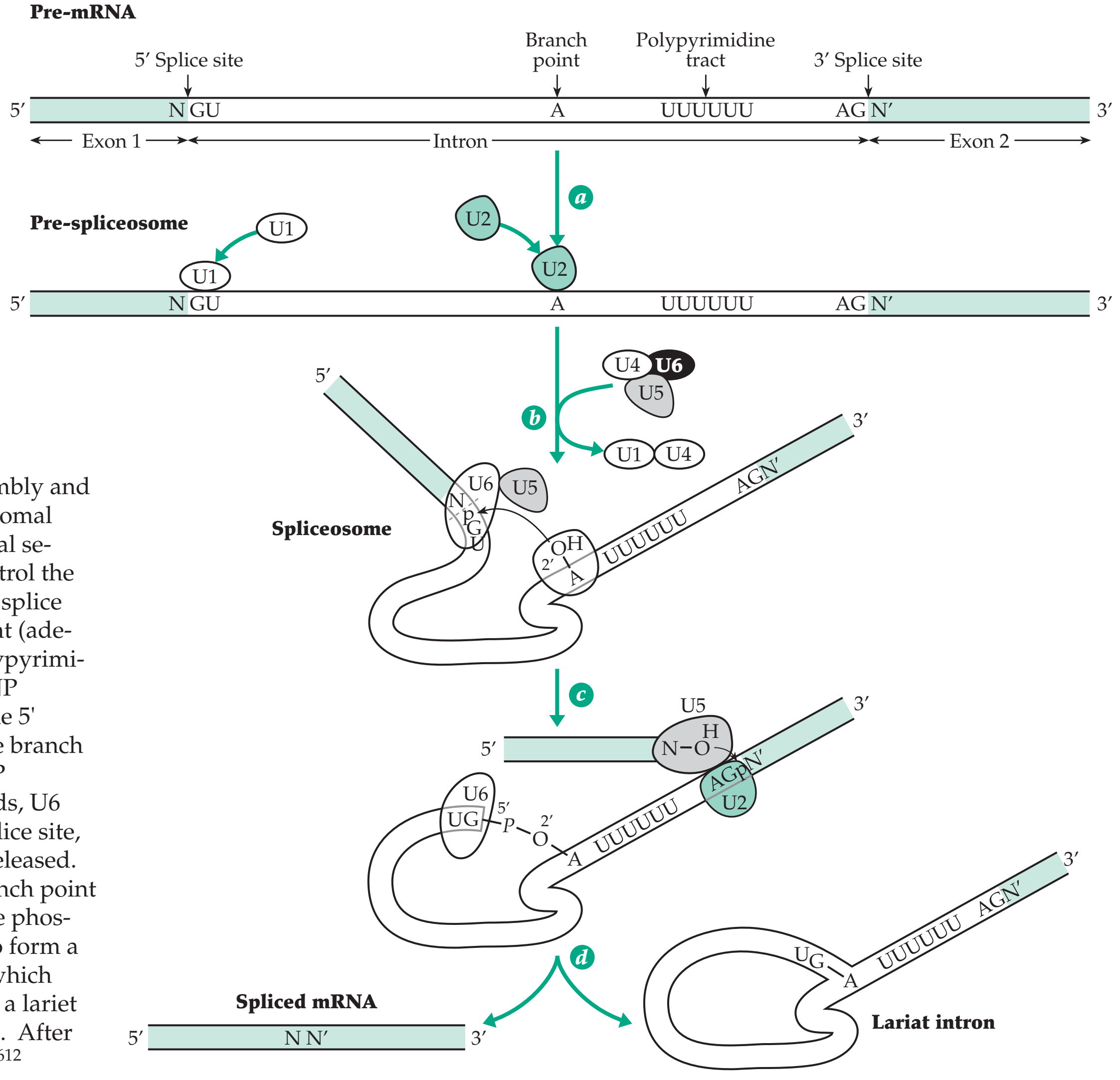

Pre-mRNA
5′ Splice site
Branch point
Polypyrimidine tract
3′ Splice site
5′
N GU
A
UUUUUU
AG N′
3′
Exon 1
Intron
Exon 2
a
Pre-spliceosome
U1
U1
U2
U2
5′
N GU
A
UUUUUU
AG N′
3′
U4
U6
U5
b
U1
U4
5′
3′
Spliceosome
U6
U5
N
p
G
U
OH
2′
A
UUUUUU
AGN′
c
U5
H
N–O
5′
3′
AGpN′
U2
U6
UG
5′
P
2′
O
A
UUUUUU
d
3′
UG
A
UUUUUU
AGN′
Spliced mRNA
5′
N N′
3′
Lariat intron

Figure 28-22 Assembly and action of the spliceosomal complex. Four special sequence elements control the process: the 5' and 3' splice sites, the branch point (adenosine A), and a polypyrimidine tract. The snRNP particle U1 locates the 5' splice site and U2 the branch point. The tri-snRNP U4•U6•U5 then binds, U6 recognizing the 5' splice site, and U1 and U4 are released. The 2'-OH of the branch point adenosine attacks the phosphodiester linkage to form a lariet intermediate, which releases the intron in a lariet form in the final step. After Valcárcel and Green.[612]

Relationship to viroids and virusoids. The smallest of viruses are the naked ~250–460 nucleotide single-stranded circular RNA molecules called viroids (Chapter 5).[591–593] Viroids are closely related to group I introns. The conserved paired sequence characteristic of the group I introns are also present in viroids (Fig. 28-19B), and it looks as if viroids may be “escaped introns.” Another group of “plant satellite RNAs” or virusoids replicate only with the help of larger RNA viruses. These satellite RNAs are replicated by a rolling-circle mechanism.[593a] The resulting long RNA molecules are self-cleaving, cutting themselves to form the unit length satellite RNAs.[594,595] These are the simplest known self-cleaving RNA molecules. They have been discussed in Chapter 12.

Pre-tRNAs. In the removal of type I introns the formation of specific stem and loop structures directs the splicing reactions (Fig. 28-18).[47,597] Stems and loop structures already exist in tRNA precursors. Cleavage sites are usually located just to the 3' side of the anticodon as is seen in the pre-tRNA molecule and in the composite structure shown in Fig. 28-20. The chemistry of the splicing process[205,598–601] is shown in Fig. 28-21. Cleavage at the two splice sites excises the intron. Then the two ends are rejoined. In yeast cleavage at the 5' end of the intron is by a multimeric enzyme with a pancreatic ribonuclease-type of action that leaves a 2',3'-cyclic phosphodiester (Fig. 28-21A, step *a*) that is opened hydrolytically (step *b*). Another hydrolytic cleavage, catalyzed by a different subunit of the enzyme complex, occurs at the 3' end of the intron. The resulting 5' OH is phosphorylated by GTP (step *c*),[601] and the resulting 5'-phospho group reacts with ATP to displace PP_i and to form a transient adenosine-5'-diphosphate terminus (step *d*). This reacts with the 3' OH formed at the 5' splice site (step *e*) to create a phosphodiester linkage between the two pieces of RNA. A phosphotransferase then removes the superfluous 2' phosphate (step *f*).[601a] Steps *d* and *e* are catalyzed by an **RNA ligase**.[598,599]

In HeLa cells and presumably in normal mammalian nuclei, the initial cleavage at the 5' end of the intron leaves a 3' phospho group (Fig. 28-18B, step *a*), which is cyclized, probably in an ATP-dependent process, to the 2',3'-cyclic phosphate (step *b*). This is ligated to the other piece of RNA by a direct displacement on the cyclic phospho group (step *c*).[602]

The spliceosome. The hnRNA of nuclei, which includes all of the pre-mRNA, is associated with proteins, which sometimes form very large 200S particles.[604] After limited cleavage with nucleases they tend to sediment in the 30S–40S range and to contain a variety of proteins.[605,606] Some of the proteins may have been involved in control of transcription.[606] Others participate in splicing. The smaller snRNP particles then appear to come into the nucleus and displace much, but not all, of the protein present in the pre-mRNA ribonucleoprotein particles.[605]

The 50–60S spliceosome complexes, with their protein and RNA components, are reminiscent of ribosomes. Although smaller in size, they can be vizualized by electron microscopy.[607] Each spliceosome is constructed from the four snRNP particles and additional proteins (Fig. 28-22).[539,608–612] Each spliceosome can accommodate ~500 nucleotides of pre-mRNA. As the pre-mRNA moves through the spliceosome the splice sites, which have only weakly conserved consensus sequences, must be located. The two ends of the introns, which may be much longer than the average 137-nucleotide exon, must be brought close together in the spliceosome.[613] The exact splice sites are usually located by an invariant **GU** at the 5' end and **AG** at the 3' end. The sequence of the first 18 nucleotides of U1 snRNA is largely complementary to that of the 5' splice site, which has the consensus sequence AG:**GU**RAGU, where the colon marks the junction. The **GU** is invariant.

After the U1 snRNP binds to the pre-mRNA (step *a*, Fig. 28-22)[614] the U2 snRNP binds to another almost invariant sequence CUR**A**CU found 20 to 55 nucleotides upstream of the 3' junction.[608,615–617] The A in this sequence becomes a branch point. It is brought close to the 5' splice site with the aid of a preassembled complex of snRNPs U4, U6, and U5. In this complex U4 and U6 are tightly paired, additional proteins are also present,[618–621] and enhancers may be located in adjacent exons.[617] Upon binding of U6 to the 5' splice site, the U1 and U4 snRNPs are released (step *b*, Fig. 28-22) and the 2'-OH of the branch point adenosine attacks the backbone phosphorus atom (step *c*) at the 5' splice junction forming a lariat intermediate. The 3' end created at the 5' junction must now be held and brought close to the 3' splice junction, which is located with the aid of U5 snRNP.[622] The 3' splice junction, utilized in the second splicing step (step *d*, Fig. 28-22) has the consensus sequence (T/C)N(C/T)AG:G.

The first splicing step is dependent upon a divalent metal ion, but the second is not.[623] Both steps appear to be in-line nucleophilic displacement reactions.[624] Additional **splicing factors** are needed for formation of the U4, U6, U5 complex and its function in the second splicing step, which also appears to require ATP.[621,622]

A small fraction of eukaryotic mRNA introns are characterized by **AU** and **AC** (rather than GU and AG) ends. The spliceosomes that act on these introns contain modified snRNAs U4 and U6, which are designated $U4_{atac}$ and $U6_{atac}$. They also require U11 and U12 snRNPs.[534,625,626]

Since pre-mRNAs usually contain many introns, a series of splicing events must occur. These apparently take place consecutively beginning at the 5' end. Similar splicing pathways are followed in yeast, higher plants, insects,[627] and mammals.

Alternative splicing pathways. RNA that contains many introns can undergo splicing in more than one way. Many examples of alternative splicing have been discovered.[612] The mammalian isoenzyme forms of pyruvate kinase called M_1, M_2, L, and R are all tetramers of 60-kDa subunits. The M_1 and M_2 forms are encoded by a single gene. The two mRNAs contain 1593-nucleotide coding regions, which are identical except for a 160-nucleotide sequence that determines the amino acid sequence in a region responsible for intersubunit contact. The difference between the M_1 and M_2 forms involves a choice of two alternative exon regions, one or the other of which is omitted during splicing.[628] Other examples have been found in human collagen,[629] in fibronectin,[630] in neuropeptide formation (Chapter 30),[631] and among human proline-rich salivary proteins,[632] cytoskeletal tropomyosin,[633] platelet-derived growth factor,[634] coagulation factor X,[635] and porphobilinogen deaminase.[636] Alternative splicing is very common in transcripts of viral DNA (Section E).

Alternative splicing could have arisen accidentally, but it is controlled by proteins. Best known is the **alternative splicing factor** (**ASF** or SF2).[612,637] It was first recognized by its function in *Drosophila melanogaster*, where the sex of individuals is determined by alternative splicing of an mRNA.[638] In addition to ASF other serine- and arginine-rich **SR proteins** participate in alternative splice site selection.[612]

Trans splicing. Every mRNA in trypanosomes has, at its 5' end, a short 35-nucleotide sequence that is not encoded in the transcribed gene. It was found that for each mRNA molecule two transcripts are formed and are spliced together, always with the 5' piece from a short SL (spliced leader) gene being joined to each of the others.[639] This trans splicing has since been observed in many plants, animals, and protists.[640]

Among these are the nematode *Caenorhabditis*,[639,641] flatworms,[642] yeast,[643] plant chloroplasts and mitochondria,[640,644] and mammalian cells.[645]

6. Modification and Editing of RNAs

Both mRNA and rRNA undergo rapid methylation of selected residues. About 1–3 internal adenylate residues per kb are methylated at their N^6 positions.[646] These are usually the central adenylates in the sequences GAC or AAC. The methylated sites are not uniformly distributed but are clustered, sometimes in the 3' untranslated ends of the RNA. Many more residues (55 in yeast and ~100 in vertebrates) in rRNA are methylated on selected ribose 2'-OH groups.[173, 535c–e,647–648a] About an equal number of uridine residues are converted to pseudouridines (Eq. 28-3). Methylation sites are apparently selected by the fibrillarin-associated snoRNAs U32–U40 (Table 28-3). These **methylation-guide** snoRNAs contain 10- to 14-nucleotide sequences that are complementary to segments of the rRNA that contain the methylation sites, and evidently provide rigid helical regions that are targets for the methylase action. The snoRNAs contain the previously mentioned C and D sequences. Methylation occurs on the ribose of the nucleotide that is base-paired with the fifth nucleotide that is upstream of the D sequence in the snoRNA. Each snoRNA directs methylation of a different ribose.[173] In a similar manner the snoRNAs containing the ACA motif located three nucleotides upstream of their 3' ends appear to direct the conversion of uridines to **pseudouridines**, as in Eq. 28-3.

Editing of RNA transcripts. Interpretation of the genetic code utilized by the mitochondria of trypanosomes and other kinetoplastid protozoa was confounded by the discovery that the DNA sequences of many genes, including the COIII cytochrome oxidase gene present in the kinetoplast maxi circle DNA, do not appear to encode the correct amino acid sequence. In fact, the RNA transcripts are extensively edited, mostly by insertion of multiple uridine residues at many positions[649,650] and by occasional deletions at others. Editing of some mRNAs causes 45% of the message to be rewritten.[651] Several additional kinds of editing were soon discovered, not only in protozoa but also in mammals, plants, and archaea.[652,652a]

A frequent editing change is the hydrolytic deamination of a C to a U residue.[652b] For example, human apolipoprotein B is synthesized in two forms: apoB100, a full-length 512-kDa protein made in the liver and used for transport of cholesterol and triglycerides, and a shorter 241-kDa form, apoB48, used in absorption of dietary lipids (Chapter 21, Section A1). ApoB100 is synthesized from a full-length mRNA, but apoB48 is made according to a shortened mRNA in which a glutamine codon (CAA) has been converted by editing to the translation stop codon UAA.[652,653] A special enzyme deaminates only cytidine 6666 of the mRNA.[654] C to U editing occurs in chloroplasts and mitochondria of plants.[655–657] In *Arabidopsis* mitochondria 456 different C to U conversions have been identified in mRNAs.[656]

Deamination of adenosine residues produces inosine, which occurs in brain mRNA once in ~17,000 ribonucleotides.[658] Some ionotropic glutamate receptors in the brain have subunits translated from inosine-containing mRNAs. A glutamine codon CAG is edited to CIG, an arginine codon. The arginine codon AGA is converted to the glycine codon IGA at another site, ATT is converted to ITT and TAC to TIC. All of these changes affect the properties of the glutamate-activated ion channel.[659] The adenosine deaminases involved in these editing events are usually specific for double-stranded helical segments of RNA, e.g., for the stems in stem–loop structures.[660–662] At least one human adenosine deaminase not only binds to RNA but also contains a DNA-binding domain specific for Z-DNA.[663]

Returning to the trypanosomes and their relatives, mitochondrial RNAs undergo extensive insertion and deletion of U's. The editing site is located by a **guide RNA** (gRNA), which directs the hydrolytic cleavage of the chain and either the addition of U's to the 3' cut end by transfer from UTP or hydrolytic deletion of U's from the 3' cut end. The chain ends are then rejoined by an RNA ligase.[172,664–667c] The functional significance of the editing of kinetoplast mRNA is uncertain. However, at least some of the edited mRNA is translated to give proteins that are presumably used.[668] While trypanosomes usually insert only U's, the slime mold *Physarum polycephalum* may insert dinucleotides such as AA, AU, CU, or GU[669] and may also add nucleotides at the 3'-ends of RNAs.[669a]

A to G editing occurs in RNA of *Drosophila*.[670] Yet another type of editing has been observed in viral RNA from paramyxovirus. The virally encoded RNA polymerase sometimes "stutters" reading the same template base two or more times, with a resulting insertion of a base.[671] Editing of transcripts usually serves an essential biological function, creating the correct sequence for translation of the mRNA and often generating multiple isoforms of proteins.

Finishing the transcripts. Additional modifications must be made to some mRNAs, and there will doubtless be many surprises as the details are worked out. One detail, which was discovered in the 1980s, is the specific function of snRNA U7 in recognition of the 3' end of pre-mRNAs for histones. The U7 RNA apparently base-pairs with a sequence near the 3' end cleavage site, acting as a cutting guide.[47,672,673]

Newly synthesized mRNA emerges from the nuclear pores as nucleoprotein complexes containing as many as ten different proteins. A major component is a 78-kDa polypeptide thought to be associated with the poly(A) tail. These ribonucleoproteins are sometimes stored for long periods of time, for example, in mature seed embryos and in amphibian oocytes.[575a,674] They may also travel rapidly for long distances, e.g., down nerve axons[675] or from cell to cell in plants via transport in the phloem.[676]

E. Transcription of Mitochondrial, Chloroplast, and Viral Genes

In the compact 16-kb chromosomes of mammalian mitochondria, the genes are tightly packed against one another (Fig. 18-3).[677,678] Most genes are transcribed using heavy H strands as templates and specialized bacteriophage-type RNA polymerases encoded in nuclear DNA.[679] A single promoter in the D loop region (Fig. 18-3) is used to make a long transcript from the entire H strand.[680] The transcript is then cut precisely by mitochondrial RNase P at the 5' ends of the tRNAs.[681] Similarly precise cleavage must occur at the 3' ends because there are often no nucleotides or only one to a few nucleotides separating adjacent genes. The nucleases involved differ from those used in nuclear tRNA 3' processing.[682] In animal mitochondria tRNA genes sometimes overlap by one nucleotide. Polyadenylation of the transcripts provides a form of editing that is required to create a UAA translational termination signal to which termination proteins bind.[683] The tRNA genes lack the 3' terminal CCA, which must be added. The lighter L strand has its own promoter, also located in the D loop. Both promoters contain the conserved nonanucleotide sequence 5'-ATATAAGTA. The pre-mRNAs created by these cleavages are not capped but are polyadenylated. Since the promoters are simple and the transcription factors few, mitochondrial transcription is controlled largely by mRNA stability, translation, and posttranslational events.[684]

The 70- to 100-kb circular mitochondrial DNA molecules of yeast and of higher plants contain more genes than do animal mitochondria, but most of the increased size is accounted for by intergenic spacers and by a few long introns. All yeast mitochondrial genes except for that of tRNA^{Thr} are transcribed from a single strand. About 20 different primary transcripts have been identified in *Saccharomyces cerevisiae*. These originate at several points in the genome but always at the sequence 5-ATATAAGTA, the 3′A corresponding to the 5' nucleotide of the transcript.[678] One of these sequences is located at the origin of replication, suggesting the possibility that a normal RNA transcript provides the primer for DNA replication in yeast mitochondria.[677]

The 100- to 160-kb chloroplast genomes (Chapter 23, Section E,2) also have many prokaryotic features. They encode ~50 proteins as well as the tRNAs and rRNAs. Promoter and terminator sequences resemble those of bacteria and protein sequences are often homologous to those in bacteria. This applies, for example, to the α, β, and β′ subunits of RNA polymerase.[685]

1. Viral Transcription and Replication

Because viruses contain small genomes, study of transcription of viral DNA and of replication of RNA viruses has played an important role in helping us to understand transcription in eukaryotes.[47,686–688] An example is the discovery of the virus SV40 enhancer, which has been discussed in Section C,4. Study of viral life cycles is also essential to future progress in fighting viral diseases. Each of the many different viruses has its own often very complex life cycle. Only a few details can be given here. For lucid summaries see Voyles.[259]

Eukaryotic DNA viral genomes, like that of phage λ, usually contain early and late transcriptional units. The small papovaviruses, such as SV40 and polyoma virus, have 5.2 kb genomes. Like the small RNA viruses, they make use of overlapping genes and alternative RNA processing. In SV40 DNA there are two overlapping promoters called **early-early** and **late-early**. The first of these contains a TATA sequence, and both promoters also depend upon a 21-bp repeat segment as well as the SV40 enhancer. At least two proteins, one that binds to the enhancer and one that binds to the 21-bp repeat, are needed for initiation of early transcription. The early-late promoter lacks the TATA sequence but requires the 21-bp repeat and enhancer. A 94-kDa encoded protein called the **large T-antigen** (Chapter 27, Section C,10) is one of the regulators of transcription as well as of DNA replication.[689,690] This protein is also sufficient to transform rodent cells in culture. Although predominantly nuclear it is also inserted into the membrane where it acts as an antigen.

The large icosahedral adenoviruses cause respiratory infections in humans and attack and may cause cancer in many other vertebrate species including birds and amphibians. The 35.9-kb genome of human adenovirus-2 encodes at least 30 proteins, 10 of which appear in the virion. One of these is covalently linked to the 5' end of the DNA. As with smaller DNA viruses extensive use is made of alternative splicing of the transcribed RNA. There are at least six early transcriptional units, each with its own promoter. A variety of mRNAs are created using the various coding segments and a complex array of regulatory sequences control transcription. A 32-kDa phosphoprotein transcription factor designated E1A is encoded by a “pre-early” gene. It is required along with host-encoded transcription factors for viral transcription.[691,692]

The late region of the adenovirus genome encodes structural proteins for the virus coat. Most of its transcripts begin about 16.5% of the way along the 36.5-kb dsDNA. However, the initial transcripts are cleaved at several different positions to yield a series of different 3'-poly(A)-terminated transcripts. In an exceedingly complex process the transcripts undergo splicing out of genes at their 5'-ends so that the final mRNAs typically code for single proteins. At the 5' ends the cap is joined to short segments from the original mRNA 5' end.[687,688,693]

Viruses SV40, polyoma, and some strains of adenoviruses are oncogenic in some species and cause transformations of cells in culture. Transformed cells always contain integrated viral DNA. That of SV40 can be incorporated at many different sites in the host genome. The integrated DNA does not always include the complete SV40 genome, and parts of the DNA may be inverted, deleted, or scrambled. Integration is not an essential part of the viral life cycle and has no effect on the infective properties of the viruses.[688] Cells transformed by adenoviruses usually also contain only a fragment of the viral genome in their DNA. However, one small set of genes from early region 1A is present in all transformed cells. The encoded proteins appear to be modulators of transcription and may cause cancer by promoting uncontrolled transcription of certain genes.[694]

2. Replication of RNA Viruses

The RNA (+) strands present in many RNA viruses often serve immediately after infection as a messenger RNA. However, replication requires formation of (–) strands of viral RNA from which new (+) strands can be transcribed for assembly into new virus particles. Other RNA viruses contain (–) strands of RNA or double-stranded RNA and, therefore, have significantly different life cycles.[259]

Small RNA viruses. The human polio virus, the common cold virus (rhinoviruses, Fig. 7-15), and other picorna viruses have 7.2- to 7.5-kb genomes with considerable homology (50% between the polio and rhinoviruses) and similar overall structures.[695] The polio genome encodes eight different proteins, one being a small 22-residue peptide that becomes covalently linked to the 5'-end of the RNA through a phosphodiester linkage to the side chain of a tyrosine. Cleavage of this linkage by a host enzyme allows the viral RNA (+) strand, which is polyadenylated at the 3' end, to serve as an mRNA for synthesis of a single large 220-kDa polyprotein. This is cleaved by a host protease at several Gln-Gly bonds to form several proteins. These include the 22-residue RNA-linked peptide, two capsid proteins, and a capsid precursor protein. The latter is cleaved during capsid assembly by a viral protease at an Asn-Ser bond to give two more capsid proteins. These four proteins have masses of 7, 26, 29, and 32 kDa. A viral protease, a large 58-kDa replicase, and a 37-kDa protein of unknown function are also cut from the polyprotein. A host protein initiates cleavage of the polyprotein, but the virally encoded protease later takes over this function.

One of the best understood of the many viral pathogens of plants is the tobacco mosaic virus (Fig. 7-8). Its 6.7-kb positive strand RNA encodes a replicase, coat protein, and at least one other protein.[696]

Influenza viruses. These negative-stranded viruses are classified into types A, B, and C, but it is only type A that infects nonhuman species including birds, horses, pigs, seals, mink, and whales.[697–699] Type A influenza viruses have also caused the great pandemics such as those in 1918–1919 and in 1968. Influenza viruses are surrounded by a lipid bilayer in which the virally encoded **hemagglutinin** and a **neuraminidase** (p. 186) are embedded. The inside of the bilayer is coated with a matrix protein and within this coat eight pieces of RNA of total length 13.6 kb are coiled together with a basic nucleoprotein. Also present are ten molecules each of three other proteins. The eight pieces of RNA vary in length from 900 to 2500 nucleotides. Seven of them encode one each of the seven virion proteins. One encodes an additional nonstructural protein while the smallest piece, using overlapping nucleotide sequences, encodes two nonstructural proteins.[697]

The existence of a fragmented genome evidently underlies the ability of influenza A viruses to undergo rapid changes in antigenic behavior. If a cell is coinfected with two strains of virus, the eight fragments act as independent chromosomes, which can be reassorted into new combinations in the progeny viruses. As a consequence, it is difficult to develop safe, live virus vaccines. A large reservoir of infection among migratory water birds and other animals facilitates the appearance of new strains and their rapid spread throughout the world.[698]

The first step in the replication of influenza viruses, which takes place in the cytoplasm, is the synthesis of (+) strands that can serve both as mRNA for synthesis of proteins and as templates for synthesis of new (–) strands. Three of the capsid proteins form the required RNA polymerase. This “transcriptase” is primed preferentially by 5'-capped 10- to 13-nucleotide segments of RNA that have been cut by a viral nuclease from host mRNAs.[700] The mRNAs made from viral RNA are polyadenylated and are translated by the host cell’s ribosomes. However, some transcripts are used as templates to form viral (–) strands, which are not polyadenylated and which contain uncapped pppA at the 5'-ends.

HIV-1 and other retroviruses. Because of their association with viral oncogenes (Chapter 11) and because of the **human immunodeficiency virus** (HIV-1) and the AIDS epidemic a great deal of attention is focused on retroviruses.[701–701b] Each retrovirus particle contains *two* identical single-stranded (+) RNA molecules, which may be as long as 10 kb. Their unique characteristic is that they induce synthesis of DNA, which must be integrated into the host genome before new viral (+) strands are transcribed. Retroviruses may sometimes cause cancer and may carry oncogenes (Chapter 11). Study of the **Rous sarcoma virus** (RSV), which infects chickens, and of the related **avian myeloblastosis virus** (AMV) and of HIV has revealed a common structure and a complex life style that are largely shared by all known retroviruses.

The organization of retroviruses[687,688,702] always includes a sequence of genes designated *gag* (glycoprotein antigen core proteins), *pol* (polymerase), and *env* (envelope) (Fig. 28-23). These are often followed by an oncogene.[703] In RSV this is the *src* gene (Chapter 11). At each end of the retrovirus gene sequence is a short direct repeat labeled R in Fig. 28-23. In RSV the R sequence is 21 nucleotides in length[688] and in HIV (see Fig. 28-23)[704] it is 98 nucleotides long.[705,706] The 5' end of the viral RNA is capped, and the 3' end is polyadenylated. The dsDNA of the integrated form of the virus (Fig. 28-23) is longer and at each end is bounded by **long terminal repeats** (LTRs). Each LTR consists of a sequence, designated U3, that is present next to R at the 3' end of the viral RNA. In the LTR this is followed by sequence R and then by U5, a unique sequence that came from the 5′ end of the viral RNA. Each ds-LTR begins and ends with a short inversely repeated segment:

5'-TGT —— ACA in RSV
5'-CTG —— CAG in HIV

The integrated provirus is always bounded by a sequence of host DNA that is repeated without inversion at the opposite end. For RSV this is a 5-bp sequence.

The LTRs in RSV are 569 bp in length, and those in HIV are 634 bp in length, 83 bp coming from U5, 98 from R, and 453 from U3.[706] The LTRs themselves often contain promoter and other control elements and even entire genes. The organization of a retrovirus (Fig. 28-23) reflects the complex mode of replication, which is presented in simplified form in Fig. 28-24. The key enzyme is the RNA-directed **reverse transcriptase** (Fig. 27-12).[707–709] The initial synthesis of DNA by this enzyme is primed by a tRNA. RSV uses $tRNA^{Trp}$ and HIV $tRNA^{Lys}$ for this purpose.[710–712] The 3' end of the tRNA, including the nucleotides forming both the acceptor stem and the stem of the TΨC loop (Fig. 5-30), must relinguish its normal base-pairing to form ~18 Watson–Crick base pairs with a **primer binding site** (labeled PB in Fig. 28-24) near the 5' end of the retroviral DNA. Because synthesis of the (–) strand of the retroviral DNA begins so close to the 5' end of the template (Fig. 28-24), only a short piece of DNA, including sequences U5' and R′ complementary to U5 and R, can be formed. For replication to continue RNA must be removed from hybrid regions. This is accomplished by the **RNaseH** activity of the reverse transcriptase.[708,709,713,713a] After removal of the RNA the primer tRNA must undergo a **strand transfer**, in which it shifts from the 5' end of the viral RNA template to the 3' end (Fig. 28-24), utilizing pairing between the right-hand R sequence of the template and the R′ sequence of the growing cDNA copy.[703,711,714] This transfer is sometimes to the second of the pair of identical RNA molecules in the virus, providing a way of increasing diversity by recombination. A second strand transfer of the growing (+) strand is needed to complete the dsDNA, which now contains the two identical LTRs. The 5' and 3' ends of the template RNA are doubtless held close together to facilitate strand transfer.

Integration of the dsDNA into the host DNA can occur at many places. The mechanism of integration probably resembles that used by phage λ (Fig. 27-27) and accounts for the duplication of host sequences at the two ends of the integrated virus. A virally encoded integrase catalyzes the process (see also Chapter 27, Section D,3).[715–717] It is the integrated virus that is transcribed to form new (+) viral RNA strands.

Integrated retroviruses are usually transcribed as full-length RNA copies, which may or may not have introns spliced out. The smaller spliced pieces encode the *env* and other genes such as *src* (Fig. 28-25). The *gag-pol* region is translated as a polyprotein that is cleaved into a number of pieces. These include four proteins of the virus core (encoded by *gag*), the reverse transcriptase with its associated RNaseH, and an **integrase**[718,718a] (all encoded by *pol*).[259] There is also an **aspartic protease** only 99 residues in length within *pol* (Fig. 28-25; Chapter 12, Box 12-C). The promoter and control region for transcription is located in the U3 region and is placed into a position where it can function only upon synthesis of the first LTR. The gene *env* encodes the major viral envelope protein and is translated from a spliced mRNA (Fig. 28-25).

Accessory regulatory genes. HIV and some related retroviruses such as HTLV-1 (which causes rare T-cell leukemias)[719] are distinguished from other retroviruses by a marked increase in the rate of DNA transcription within infected cells as compared with uninfected cells. This is thought to be a result of synthesis of virally encoded proteins that are trans-acting

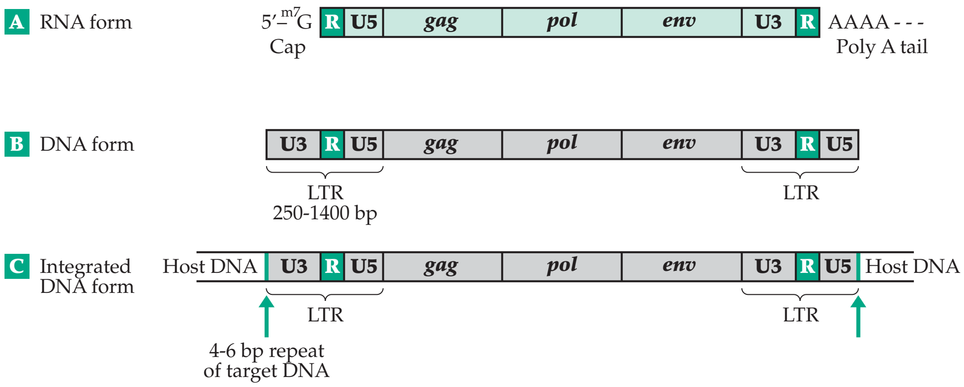
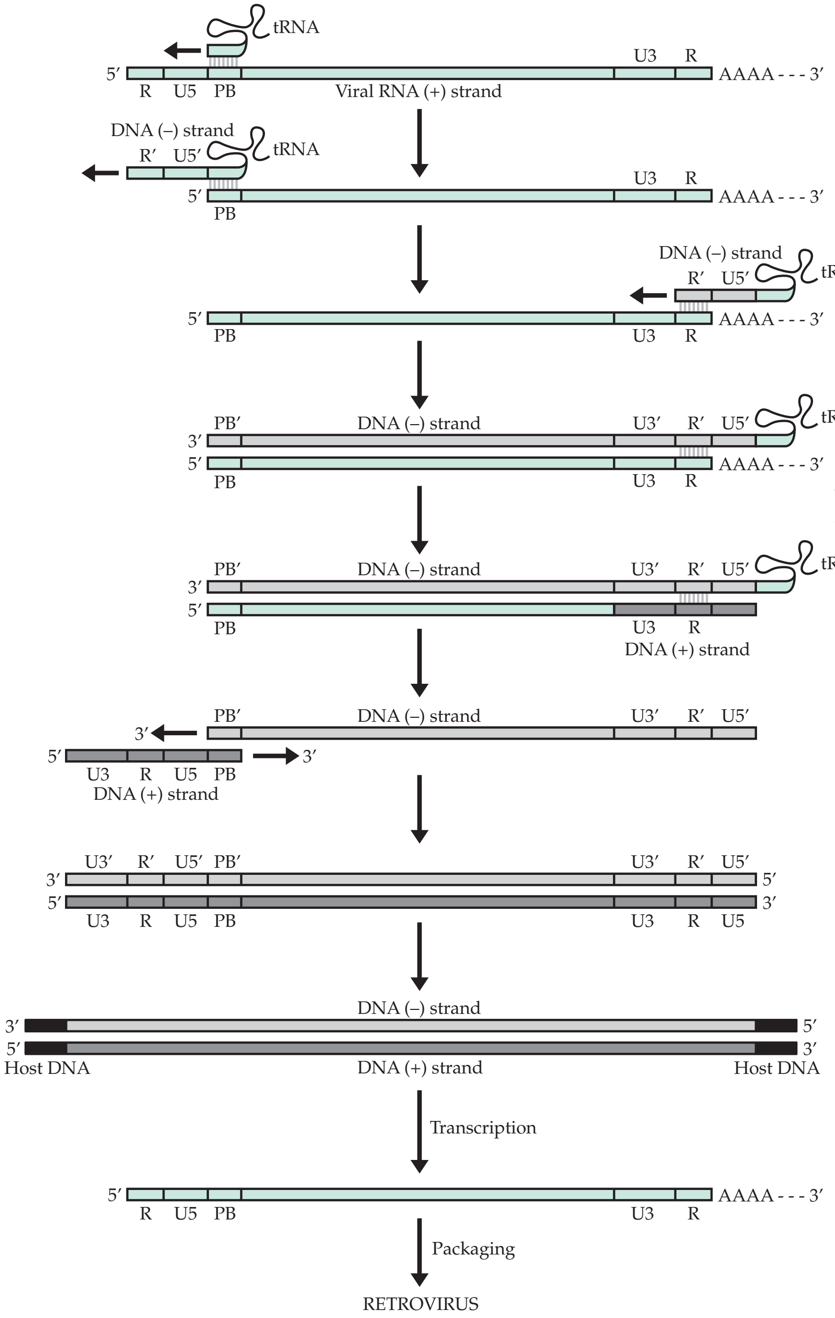

A RNA form
5′–m7G
Cap
R U5 gag pol env U3 R
AAAA - - -
Poly A tail
B DNA form
U3 R U5 gag pol env U3 R U5
LTR
250-1400 bp
LTR
C Integrated DNA form
Host DNA
U3 R U5 gag pol env U3 R U5
Host DNA
LTR
LTR
4-6 bp repeat
of target DNA

Figure 28-23 Comparison of the forms of a retrovirus. (A) Infective RNA (+) strand. (B) The double-stranded DNA form. (C) The DNA form integrated into the host DNA. LTR, long terminal repeats.


tRNA
5′ R U5 PB U3 R AAAA - - - 3′
Viral RNA (+) strand
DNA (–) strand
R′ U5′
tRNA
5′ PB U3 R AAAA - - - 3′
DNA (–) strand
R′ U5′
tRNA
5′ PB U3 R AAAA - - - 3′
PB′ DNA (–) strand U3′ R′ U5′
tRNA
3′
5′ PB U3 R AAAA - - - 3′
PB′ DNA (–) strand U3′ R′ U5′
tRNA
3′
5′ PB U3 R
DNA (+) strand
PB′ DNA (–) strand U3′ R′ U5′
3′
5′ U3 R U5 PB 3′
DNA (+) strand
U3′ R′ U5′ PB′ U3′ R′ U5′
3′ 5′
5′ 3′
U3 R U5 PB U3 R U5
DNA (–) strand
3′ 5′
5′ 3′
Host DNA
DNA (+) strand
Host DNA
Transcription
5′ R U5 PB U3 R AAAA - - - 3′
Packaging
RETROVIRUS

Viral (+) RNA strand. Start of DNA synthesis is primed by tRNA.

DNA (–) strand synthesis has reached end of R. RNA has been removed in hybrid regions by RNase H.

Minus strand has been repositioned and poly(A) tail has been removed from (+) strand. DNA (–) strand synthesis continues.

U3 region of RNA strand is removed and synthesisof DNA (+) strand begins.

All RNA is removed and DNA (+) strand is repositioned to pair PB and PB′.

DNA synthesis is completed for both strands.

DNA provirus is integrated into host genomes.

Integrated viral DNA is transcribed by RNA polymerase II and transcript is polyadenylated.

Two of these RNA molecules are packaged together with proteins to form mature retrovirus.

Figure 28-24 Simplified scheme for replication of the RNA genome of a retrovirus. See Sugden.[703] PB, Primer-binding site.

regulators of transcription. The HIV genome map (Fig. 28-26) shows the positions of the nine recognized genes marked in the three reading frames. In addition to the *gag*, *pol*, and *env* genes there are genes for six accessary proteins: Tat, Rev, Vif, Vpr, Vpu, and Nef.[720–722]

Transcription is initiated at the promoter in the 3' LTR. This contains a TATAA sequence, an SP1 binding site, and an enhancer that binds transcription factor NF-κB (Fig. 5-40). The full-length 9-kb transcript contains, according to Frankel and Young,[722] the following

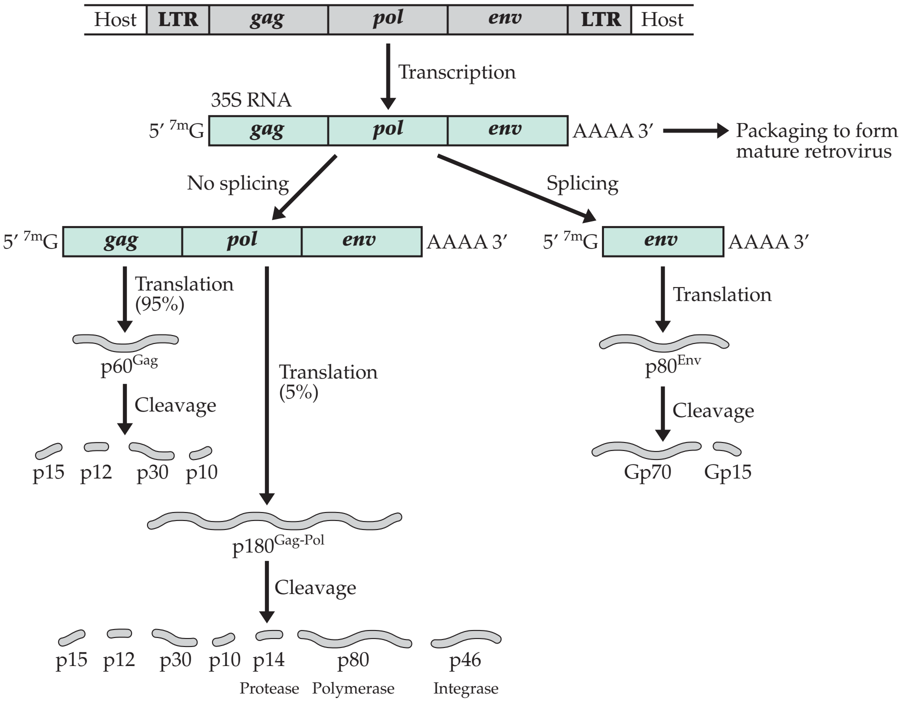

Host
LTR
gag
pol
env
LTR
Host
Transcription
35S RNA
5′ 7mG
gag
pol
env
AAAA 3′
Packaging to form mature retrovirus
No splicing
Splicing
5′ 7mG
gag
pol
env
AAAA 3′
5′ 7mG
env
AAAA 3′
Translation (95%)
p60Gag
Cleavage
p15
p12
p30
p10
Translation (5%)
p180Gag-Pol
Cleavage
p15
p12
p30
p10
p14
Protease
p80
Polymerase
p46
Integrase
Translation
p80Env
Cleavage
Gp70
Gp15

Figure 28-25 Gene expression from a typical retrovirus that has been integrated into a host's genome. This figure illustrates how a variety of proteins are encoded by a single rather short piece of DNA. After Voyles.[259]

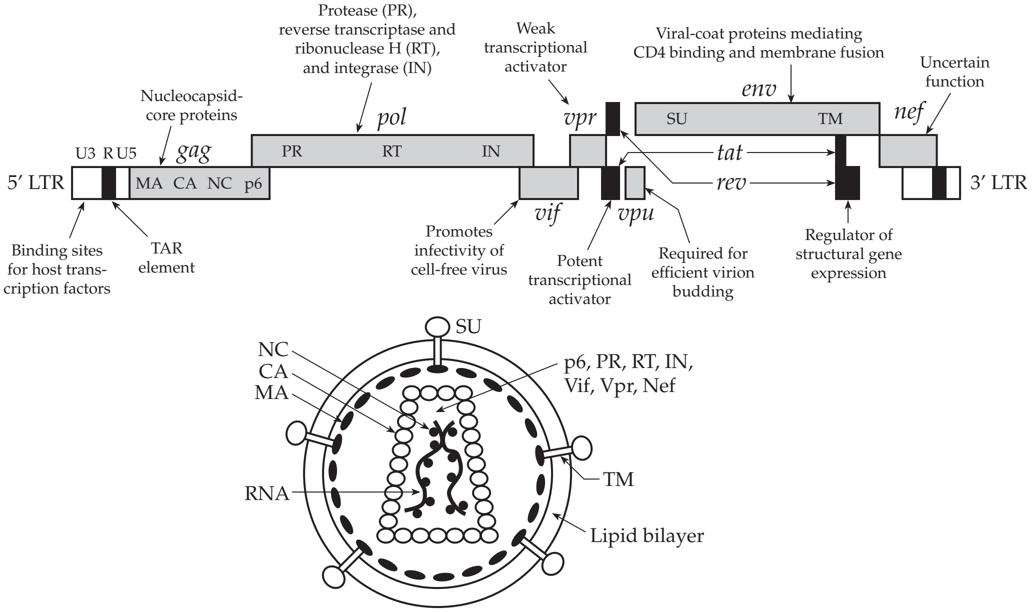

Protease (PR), reverse transcriptase and ribonuclease H (RT), and integrase (IN)
Weak transcriptional activator
Viral-coat proteins mediating CD4 binding and membrane fusion
Uncertain function
Nucleocapsid-core proteins
pol
vpr
env
nef
SU
TM
U3 R U5
gag
PR
RT
IN
tat
rev
5′ LTR
MA CA NC p6
vif
vpu
3′ LTR
Binding sites for host transcription factors
TAR element
Promotes infectivity of cell-free virus
Potent transcriptional activator
Required for efficient virion budding
Regulator of structural gene expression
SU
NC
CA
MA
p6, PR, RT, IN, Vif, Vpr, Nef
TM
RNA
Lipid bilayer

Figure 28-26 Simplified genetic map of the AIDS virus HIV-1. All three reading frames are utilized to encode nine genes, which give rise to 15 proteins. After Frankel and Young.[722]

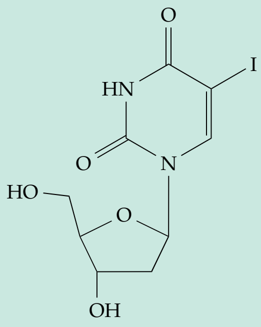
BOX 28-C SYNTHETIC ANTIVIRAL COMPOUNDS

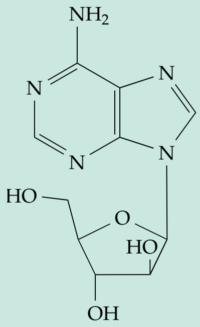
Most bacterial infections can be treated successfully with antibiotics, but the development of satisfactory antiviral agents has been slow. Yet we are susceptible to many dangerous virus diseases, and viruses also take a huge toll among domestic animals and plants.[a] The first antiviral drug, 5-iodo-2'-deoxyuridine (idoxuridine), was introduced in 1962 and was used for 20 years by ophthalmologists to treat serious eye infections by the herpes simplex virus (HSV).

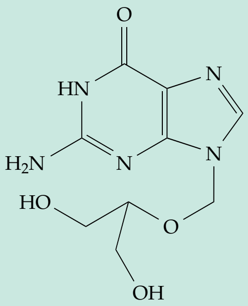
More recently 9-β-D-arabinofuranosyladenine (Ara-A) has become a preferred drug in treatment of ocular herpes infections. This compound, which is a naturally occurring antibiotic, can also be administered intravenously for life-threatening infections such as herpes encephalitis.[a] Ara-A is quite toxic but the guanine derivative 9-(2-hydroxyethoxymethyl) guanine (acyclovir) is less so. Another acyclic 2'-deoxyguanine analog, 9-(1,3-dihydroxypropoxymethyl)-guanine (DHPG), is more soluble, more potent, and has a broader range of effectiveness.[a–c]

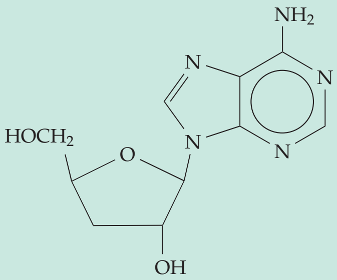
One of the first effective drugs against RNA viruses was ribavirin (1-β-D-ribofuranosyl-1,2,4-triazole-3-carboxamide), an analog of guanosine. It has a broad range of action and is used to treat severe viral pneumonia and bronchitis caused by respiratory syncytial virus (RSV). This common virus usually is mild but may cause death in infants and children. Ribavirin may also be of some value in the treatment of AIDS.

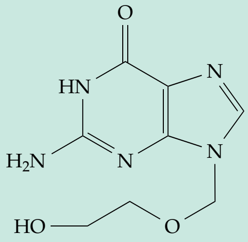
How do these compounds exert their antiviral effects? The nucleoside analogs are usually phosphorylated to the corresponding mono-, di-, and triphosphate derivatives by cellular enzymes. Thus, Ara-A yields Ara-ATP, which inhibits the herpes virus-encoded DNA polymerase. Ara-A may also enter the viral DNA. In addition Ara-A may inhibit

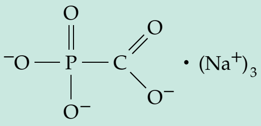

Idoxuridine

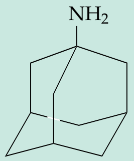

Ara-A

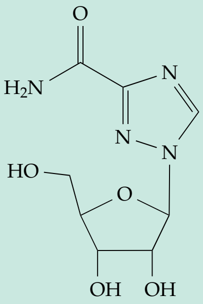

DHPG

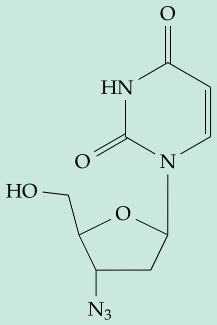

Cordycepin
(3'-deoxyadenosine)
preferentially inhibits
synthesis of ribosomal
and tRNA

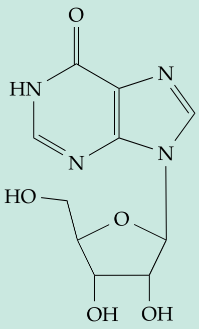

Acyclovir

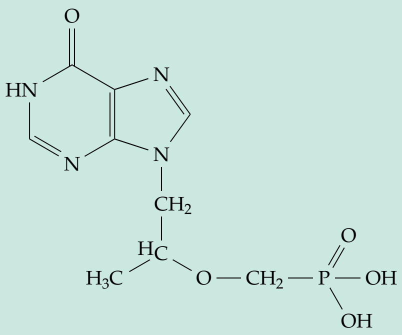

Trisodium
phosphonoformate
(Foscarnet)


Amantadine


Ribavirin


3′-Deoxy-3′-azidothymidine (AZT)


Dideoxyadenosine


(*R*)-9-(2-Phosphonylmethoxypropyl) adenine

BOX 28-C (continued)

polyadenylation of virally induced mRNA. Some analogs, such as acyclovir, undergo little conversion to the monophosphate by cellular kinases but are efficiently phosphorylated by herpes virus-encoded thymidine kinase. Thus, acyclovir does little damage to uninfected cells.[a] Ribavirin 3'-monophosphate may inhibit IMP dehydrogenase (Fig. 25-16, left) thereby interfering with GTP production. At the same time ribavirin triphosphate competes with GTP to inhibit virally encoded RNA polymerase.[a]

Phosphonoformate is a pyrophosphate analog and inhibits both DNA polymerases and reverse transcriptase. However, toxicity may prevent long-term treatment of AIDS patients. Amantadine has a narrow antiviral specificity. It specifically inhibits initiation of the replication of influenza virus RNA of type A (but not of type B). Active only against retroviruses, 3'-azidothymidine is a reverse transcriptase inhibitor, which acts by a chain termination mechanism. It was synthesized in the early 1960s but only recently has been used in treatment of AIDS victims. More recently a series of 2',3'-dideoxynucleosides, such as dideoxyinosine, have also been used.[d] Acyclic phosphonates, such as phosphonylmethoxypropyladenine, avoid the need for metabolic phosphorylation of the drug.[e]

Development of synthetic antiviral compounds is hardly beyond its infancy.[f] Serious problems must be overcome with most of these compounds. Toxicity (sometimes carcinogenicity), development of resistance by viruses, and enzymatic destruction limit the utility of most drugs. For example, adenosine deaminase destroys Ara-A quite rapidly. With our rapidly advancing knowledge of viral life cycles and protein and nucleic acid structures many new drug targets have been identified.[g] Among the targets for HIV are the reverse transcriptase,[h] protease,[i] and integrase.[j,k] Computer-assisted design, as well as new techniques of synthesis and screening, have allowed development of many non-nucleoside inhibitors.

Oligonucleotide phosphoramidates and other triplex-forming compounds may be designed to bind to specific DNA targets.[l]

---

[a] Robins, R. K. (1986) *Chem. Eng. News* **64**, 28–40

[b] Cheng, Y.-C., Grill, S. P., Dutschman, G. E., Nakayama, K., and Bastow, K. F. (1983) *J. Biol. Chem.* **258**, 12460–12464

[c] Biron, K. K., Fyfe, J. A., Stanat, S. C., Leslie, L. K., Sorrell, J. B., Lambe, C. U., and Coen, D. M. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8769–8773

[d] Sandberg, J. A., and Slikker, W., Jr. (1995) *FASEB J.* **9**, 1157–1163

[e] Tsai, C.-C., Follis, K. E., Sabo, A., Beck, T. W., Grant, R. F., Bischofberger, N., Benveniste, R. E., and Black, R. (1995) *Science* **270**, 1197–1199

[f] Mitsuya, H., Yarchoan, R., and Broder, S. (1990) *Science* **249**, 1533–1544

[g] Richman, D. D. (1996) *Science* **272**, 1886–1888

[h] Althaus, I. W., Chou, J. J., Gonzales, A. J., Deibel, M. R., Chou, K.-C., Kezdy, F. J., Romero, D. L., Palmer, J. R., Thomas, R. C., Aristoff, P. A., Tarpley, W. G., and Reusser, F. (1993) *Biochemistry* **32**, 6548–6554

[i] Rosin, C. D., Belew, R. K., Walker, W. L., Morris, G. M., Olson, A. J., and Goodsell, D. S. (1999) *J. Mol. Biol.* **287**, 77–92

[j] Robinson, W. E., Jr., Reinecke, M. G., Abdel-Malek, S., Jia, Q., and Chow, S. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6326–6331

[k] Goldgur, Y., Craigie, R., Cohen, G. H., Fujiwara, T., Yoshinaga, T., Fujishita, T., Sugimoto, H., Endo, T., Murai, H., and Davies, D. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13040–13043

[l] Giovannangeli, C., Perrouault, L., Escudé, C., Gryaznov, S., and Hélène, C. (1996) *J. Mol. Biol.* **261**, 386–398

---

essential regions, whose positions may vary somewhat from one isolate of the virus to another:

A complex pattern of splicing produces more than 30 mRNAs.[723] When viral RNAs are first produced most are doubly spliced, allowing the split genes *tat* and *rev* to be expressed by synthesis of Tat and Rev. The *trans*-activator protein Tat is a small 86-residue cysteine-rich protein, which binds the Tar stem–loop structure and greatly stimulates transcription.[724–725a]

The 116-residue *rev* (regulator of expression of virion genes) gene product is also a transactivator, which is needed for rapid production of singly spliced (4 kb) or unspliced (9kb) *gag-pol* mRNA required for formation of virus structural proteins.[726,727] The effect of Rev is probably on transport from the nucleus rather than on splicing.

Transcription is repressed by the 206-residue N-terminal myristoylated protein, Nef, a phosphoprotein that associates with cytoplasmic membranes.[728–729a] It has been difficult to learn its exact function, but it seems to be required for maintenance of the integrated provirus for long periods of time without extensive replication. Mutations in gene *nef* do not eliminate the ability of the virus to replicate in T lymphocytes and to kill them.

The 23-kDa protein Vif (**viral infectivity factor**) is not needed for growth but is essential for infectivity.[729b] Other genes in HIV are *vpu*, which encodes an 81-residue integral membrane protein (**virion protein U**), and *vpr*, which encodes the 96-kDa **virion protein R**. Several possible functions have been proposed for these small proteins.[722,729b,c]

HIV-1 is a member of the group of slow viruses or **lentiviruses**.[730] Other lentiviruses include the human HIV-2,[731] an immunodeficiency virus that attacks cats causing leukemia,[732] and the human leukemia virus

Position from 5' end	Description
1–55	TAR, a 59-residue stem-loop structure; binding site for Tat
182–199	PB, primer binding site
240–350	Packaging signal. Binds envelope protein NC
248–271	Dimerization site with “kissing loop”
290	Major splice donor site, used to form all spliced mRNAs
1631–1673	Gag-Pol frameshifting region where –1 ribosomal frameshifting occurs to allow 5–10% synthesis of Gag-Pol polyprotein
7362–7596	Rev response element. Binding site for Rev
5358 and 7971	Two major splice acceptor sites; other minor sites are also used
9205–9210	Polyadenylation signal

HTLV-1.[733] Another lentivirus causes two diseases of sheep, **maedi**, a pulmonary disease, and **visna**, a paralytic condition somewhat similar to multiple sclerosis.[734] Because of its slow development there has been doubt as to the cause of AIDS, but there is now little doubt that HIV-1 is the true culprit.

Success in treating AIDS may depend upon better understanding of the complex life cycle of HIV-1,[722,730,735] which is summarized in Fig. 28-27. The cycle begins with the binding of the virion envelope protein to the immunoglobulin-like surface protein **CD4**, which is found principally on the type T4 helper T cells (Chapter 31). Binding of CD4 to the HIV envelope proteins appears to activate the T cells to enter the cell cycle and to take up and integrate the virus. The virus infection destroys these CD4+ lymphocytes with a half-life of less than two days.[735]

A major effort is being made to devise a vaccine against HIV. However, rapid mutation of the *env* gene makes it difficult to accomplish.[736,736a] This high rate of mutation appears to be a result of a high frequency of errors by the HIV reverse transcriptase.[737] There is

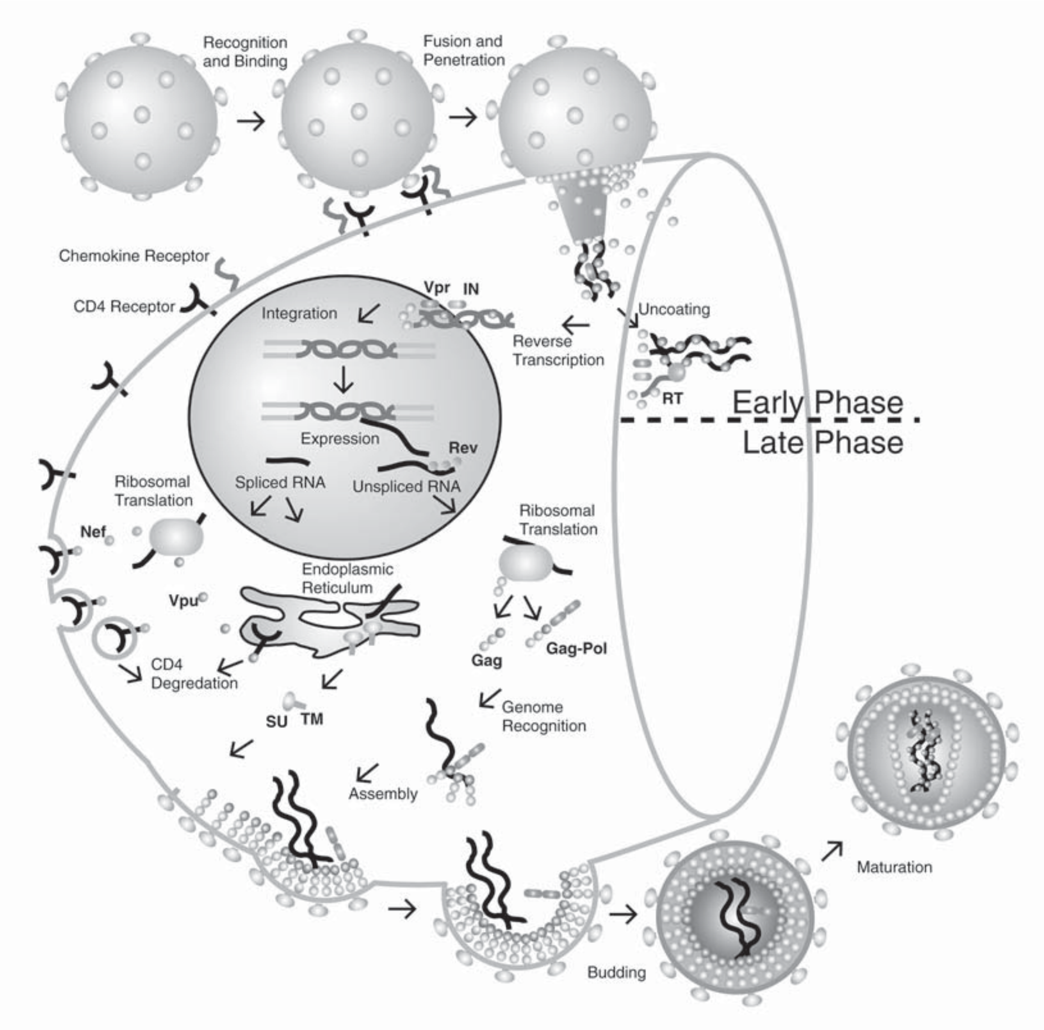

Recognition and Binding
Fusion and Penetration
Chemokine Receptor
CD4 Receptor
Integration
Vpr
IN
Reverse Transcription
Uncoating
RT
Early Phase
Late Phase
Expression
Rev
Spliced RNA
Unspliced RNA
Ribosomal Translation
Nef
Ribosomal Translation
Endoplasmic Reticulum
Vpu
Gag
Gag-Pol
CD4 Degredation
SU
TM
Genome Recognition
Assembly
Maturation
Budding

Figure 28-27 General features of the HIV-1 replication cycle. The early phase (upper portion of the diagram) begins with CD4 recognition and involves events up to and including integration of the proviral DNA, and the late phase includes all events from transcription of the integrated DNA to virus budding and maturation. From Turner and Summers.[735]

hope from the fact that some individuals are naturally resistant to HIV infection.[738,739]

3. Retrotransposons

Transposition of DNA, which is discussed in Chapter 27, Section D,4, may seem to be a rare and relatively unimportant event in our body cells. However, transposon DNA accounts for 35% or more of the human genome[740] and apparently plays a major role in evolution. Like other transposons, the DNA sequences known as retrotransposons also move about within DNA. However, they use an indirect mechanism that involves synthesis of mRNA and reverse transcription.[740,741] The reverse transcribed complementary DNA may be inserted back into the genome at new locations. The necessary chemical reactions parallel those involved in the replication of retroviruses (Fig. 28-23, 28-24). Retrotransposens, truncated retrotransposons, and related sequences constitute as much as 16% of the human genome.[741]

There are two classes of retrotrotransposons: those with long terminal repeats (LTRs) and those without (LTRs). The first group is closely related to retroviruses, but its members lack genes for envelope proteins. They do carry *gag* and *pol* genes similar to those of retroviruses (Fig. 28-3). Most retrotransposons are defective and do not move. Over evolutionary time they accumulate in the genome, sometimes to the extent that the genome size grows enormously. This has happened often during the evolution of plants, some of which (e.g., certain lilies) have 40 times more DNA per cell than do humans (Table 1-3).[742] Although most retrotransposons are inactive, some of them occasionally jump to new locations where they may mutate a gene and may sometimes cause disease. However, their major significance is probably in facilitating evolution, perhaps including the formation of new species.[743]

The non-LTR transposons are exemplified by the 6–7 kbp LINES (p. 1539)[741,743] and the short 90–400 bp SINES (p. 1538, Fig. 27-9).[744] Mammalian genomes contain ~50,000 truncated members of the LINE-1 (L1) family and 3000–5000 full-length L1s. Only a few of these are active in our present population. The RNA intermediates that participate in retroposition of LINES are generated by RNA polymerase II, while RNA polymerase III forms the RNA intermediates for propagation of SINES. Participation of these RNAs in trans splicing processes can modify existing genes, contributing to the remodeling of the genome.[744]

Yet another group of mobile elements in the genome are **intein genes**, which encode protein-splicing polypeptides (see Box 29-D). Many inteins also have **homing endonuclease** activity and cleave DNA at specific insertion sequences, initiating incorporation of intein DNA into new locations in the genome.[745] Group II introns, which are found in bacteria and in organelles of fungi and plants, may also act as mobile DNA elements.[746]

References

1. Storz, G. (2002) *Science* **296**, 1260 – 1262
1a. Zamore, P. D. (2002) *Science* **296**, 1265 – 1269
1b. Ruvkun, G. (2001) *Science* **294**, 797 – 799
1c. Cohen, G. N. (1995) *FASEB J.* **9**, 981–982
2. Dickson, R. C., Abelson, J., Barnes, W. M., and Reznikoff, W. S. (1975) *Science* **187**, 27 – 35
3. Müller-Hill, B. (1996) *The lac Operon. A Short History of a Genetic Paradigm*, de Gruyter, Berlin
3a. Juers, D. H., Heightman, T. D., Vasella, A., McCarter, J. D., Mackenzie, L., Withers, S. G., and Matthews, B. W. (2001) *Biochemistry* **40**, 14781 – 14794
4. Page, M. G. P., and Rosenbusch, J. P. (1988) *J. Biol. Chem.* **263**, 15906 – 15914
5. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) *Science* **271**, 1247 – 1254
5a. Bell, C. E., and Lewis, M. (2001) *J. Mol. Biol.* **312**, 921 – 926
6. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., and Miller, J. H. (1994) *J. Mol. Biol.* **240**, 421 – 433
7. Nakanishi, S., Adhya, S., Gottesman, M., and Pastan, I. (1973) *J. Biol. Chem.* **248**, 5937 – 5942
8. Gilbert, W., and Maxam, A. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3581 – 3584
9. Simons, A., Tils, D., van Wilcken-Bergmann, B., and Muller-Hill, B. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1624 – 1628
10. von Hippel, P. H., Bear, D. G., Morgan, W. D., and McSwiggen, J. A. (1984) *Rev. Biochem.* **53**, 389 – 446
11. Rastinejad, F., Artz, P., and Lu, P. (1993) *J. Mol. Biol.* **233**, 389 – 399
12. Levandoski, M. M., Tsodikov, O. V., Frank, D. E., Melcher, S. E., Saecker, R. M., and Record, M. T., Jr. (1996) *J. Mol. Biol.* **260**, 697 – 717
13. Frank, D. E., Saecker, R. M., Bond, J. P., Capp, M. W., Tsodikov, O. V., Melcher, S. E., Levandoski, M. M., and Record, M. T., Jr. (1997) *J. Mol. Biol.* **267**, 1186 – 1206
14. von Hippel, P. H., and Berg, O. G. (1989) *J. Biol. Chem.* **264**, 675 – 678
15. Shimamoto, N. (1999) *J. Biol. Chem.* **274**, 15293 – 15296
16. Harada, K., and Frankel, A. D. (1995) *EMBO J.* **14**, 5798 – 5811
17. Pace, H. C., Kercher, M. A., Lu, P., Markiewicz, P., Miller, J. H., Chang, G., and Lewis, M. (1997) *Trends Biochem. Sci.* **22**, 334 – 339
18. Horton, N., Lewis, M., and Lu, P. (1997) *J. Mol. Biol.* **265**, 1 – 7
19. Slijper, M., Bonvin, A. M. J. J., Boelens, R., and Kaptein, R. (1996) *J. Mol. Biol.* **259**, 761 – 773
20. Barry, J. K., and Matthews, K. S. (1999) *Biochemistry* **38**, 3579 – 3590
21. Suckow, J., Markiewicz, P., Kleina, L. G., Miller, J., Kisters-Woike, B., and Müller-Hill, B. (1996) *J. Mol. Biol.* **261**, 509 – 523
22. Kolata, G. B. (1976) *Science* **191**, 373
23. Sommer, H., Lu, P., and Miller, J. H. (1976) *J. Biol. Chem.* **251**, 3774 – 3779
24. Müller, J., Oehler, S., and Müller-Hill, B. (1996) *J. Mol. Biol.* **257**, 21 – 29
25. Lanzer, M., and Bujard, H. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8973 – 8977
26. McClure, W. R. (1985) *Ann. Rev. Biochem.* **54**, 171 – 204
27. Auble, D. T., Allen, T. L., and deHaseth, P. L. (1986) *J. Biol. Chem.* **261**, 11202 – 11206
28. Place, C., Oddos, J., Buc, H., McAllister, W. T., and Buckle, M. (1999) *Biochemistry* **38**, 4948 – 4957
29. Cheetham, G. M. T., and Steitz, T. A. (1999) *Science* **286**, 2305 – 2309
30. Brieba, L. G., and Sousa, R. (2000) *Biochemistry* **39**, 919 – 923
30a. Brieba, L. G., Gopal, V., and Sousa, R. (2001) *J. Biol. Chem.* **276**, 10306 – 10313
31. Opalka, N., Mooney, R. A., Richter, C., Severinov, K., Landick, R., and Darst, S. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 617 – 622

References

32. Owens, J. T., Chmura, A. J., Murakami, K., Fujita, N., Ishihama, A., and Meares, C. F. (1998) *Biochemistry* **37**, 7670–7675

32a. Vuthoori, S., Bowers, C. W., McCracken, A., Dombroski, A. J., and Hinton, D. M. (2001) *J. Mol. Biol.* **309**, 561–572

32b. Kuznedelov, K., Minakhin, L., Niedziela-Majka, A., Dove, S. L., Rogulja, D., Nickels, B. E., Hochschild, A., Heyduk, T., and Severinov, K. (2002) *Science* **295**, 855–857

33. Brodolin, K., Mustaev, A., Severinov, K., and Nikiforov, V. (2000) *J. Biol. Chem.* **275**, 3661–3666

33a. Ebright, R. H. (2000) *J. Mol. Biol.* **304**, 687–698

33b. Minakhin, L., Bhagat, S., Brunning, A., Campbell, E. A., Darst, S. A., Ebright, R. H., and Severinov, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 892–897

33c. Murakami, K. S., Masuda, S., and Darst, S. A. (2002) *Science* **296**, 1280–1284

33d. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002) *Science* **296**, 1285–1290

33e. Vassylyev, D. G., Sekine, S.-i, Laptenko, O., Lee, J., Vassylyeva, M. N., and Borukhov, S. (2002) *Nature (London)* **417**, 712–718

33f. Lin, S., Katilius, E., Haffa, A. L. M., Taguchi, A. K. W., and Woodbury, N. W. (2001) *Biochemistry* **40**, 13767–13773

34. Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. (1996) *Science* **273**, 211–217

35. Wu, F. Y. H., Huang, W.-J., Sinclair, R. B., and Powers, L. (1992) *J. Biol. Chem.* **267**, 25560–25567

36. Katayama, A., Fujita, N., and Ishihama, A. (2000) *J. Biol. Chem.* **275**, 3583–3592

37. Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Kashlev, M., Heumann, H., Nikiforov, V., Goldfarb, A., and Mustaev, A. (1996) *Science* **273**, 107–109

38. Mustaev, A., Kozlov, M., Markovtsov, V., Zaychikov, E., Denissova, L., and Goldfarb, A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6641–6645

39. Buckle, M., Pemberton, I. K., Jacquet, M.-A., and Buc, H. (1999) *J. Mol. Biol.* **285**, 955–964

40. Helmann, J. D., and Chamberlin, M. J. (1988) *Ann. Rev. Biochem.* **57**, 839–872

40a. Chadsey, M. S., and Hughes, K. T. (2001) *J. Mol. Biol.* **306**, 915–929

41. Carmona, M., Claverie-Martin, F., and Magasanik, B. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9568–9572

41a. Wigneshweraraj, S. R., Chaney, M. K., Ishihama, A., and Buck, M. (2001) *J. Mol. Biol.* **306**, 681–701

42. Pérez-Martín, J., and de Lorenzo, V. (1996) *J. Mol. Biol.* **258**, 562–574

43. Marschall, C., Labrousse, V., Kreimer, M., Weichart, D., Kolb, A., and Hengge-Aronis, R. (1998) *J. Mol. Biol.* **276**, 339–353

44. Blaszczak, A., Zylicz, M., Georgopoulos, C., and Liberek, K. (1995) *EMBO J.* **14**, 5085 – 5093

45. Huang, X., and Helmann, J. D. (1998) *J. Mol. Biol.* **279**, 165–173

46. Pribnow, D. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 784–788

47. Lewin, B. (2000) *Genes VII*, Oxford Univ. Press, New York

48. Bown, J. A., Owens, J. T., Meares, C. F., Fujita, N., Ishihama, A., Busby, S. J. W., and Minchin, S. D. (1999) *J. Biol. Chem.* **274**, 2263–2270

49. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277**, 1453–1462

50. Severinov, K., Mustaev, A., Severinova, E., Kozlov, M., Darst, S. A., and Goldfarb, A. (1995) *J. Biol. Chem.* **270**, 29428–29432

51. Yang, X., and Price, C. W. (1995) *J. Biol. Chem.* **270**, 23930–23933

52. Traviglia, S. L., Datwyler, S. A., and Meares, C. F. (1999) *Biochemistry* **38**, 4259–4265

53. Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. L. (1993) *Science* **262**, 1407–1413

54. Ho Jeon, Y., Negishi, T., Shirakawa, M., Yamazaki, T., Fujita, N., Ishihama, A., and Kyogoku, Y. (1995) *Science* **270**, 1495–1497

54a. Zhang, G., and Darst, S. A. (1998) *Science* **281**, 262–266

54b. Wada, T., Yamazaki, T., and Kyogoku, Y. (2000) *J. Biol. Chem.* **275**, 16057–16063

54c. Calles, B., Monsalve, M., Rojo, F., and Salas, M. (2001) *J. Mol. Biol.* **307**, 487–497

54d. Benoff, B., Yang, H., Lawson, C. L., Parkinson, G., Liu, J., Blatter, E., Ebright, Y. W., Berman, H. M., and Ebright, R. H. (2002) *Science* **297**, 1562–1566

55. Bokal, A. J., IV, Ross, W., and Gourse, R. L. (1995) *J. Mol. Biol.* **245**, 197–207

56. Muskhelishvili, G., Buckle, M., Heumann, H., Kahmann, R., and Travers, A. A. (1997) *EMBO J.* **16**, 3655–3665

57. Travers, A., and Muskhelishvili, G. (1998) *J. Mol. Biol.* **279**, 1027–1043

58. Pan, C. Q., Finkel, S. E., Cramton, S. E., Feng, J.-A., Sigman, D. S., and Johnson, R. C. (1996) *J. Mol. Biol.* **264**, 675–695

59. Lamond, A. I. (1985) *Trends Biochem. Sci.* **10**, 271–274

59a. Pemberton, I. K., Muskhelishvili, G., Travers, A. A., and Buckle, M. (2000) *J. Mol. Biol.* **299**, 859–864

60. Cashel, M., and Rudd, K. E. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 1410–1438, Am. Soc. for Microbiology, Washington, DC

61. Baracchi, E., and Bremer, H. (1988) *J. Biol. Chem.* **263**, 2597–2602

62. Choy, H. E. (2000) *J. Biol. Chem.* **275**, 6783–6789

62a. Barker, M. M., Gaal, T., and Gourse, R. L. (2001) *J. Mol. Biol.* **305**, 689–702

63. Yamagishi, M., Cole, J. R., Nomura, M., Studier, F. W., and Dunn, J. J. (1987) *J. Biol. Chem.* **262**, 3940–3943

64. Singer, P. T., and Wu, C.-W. (1988) *J. Biol. Chem.* **263**, 4208–4214

65. Gelles, J., and Landick, R. (1998) *Cell* **93**, 13–16

66. Bustamante, C., Guthold, M., Zhu, X., and Yang, G. (1999) *J. Biol. Chem.* **274**, 16665–16668

67. Lefèvre, J.-F., Lane, A. N., and Jardetzky, O. (1988) *Biochemistry* **27**, 1086–1094

68. Ellinger, T., Behnke, D., Knaus, R., Bujard, H., and Gralla, J. D. (1994) *J. Mol. Biol.* **239**, 466–475

69. von Hippel, P. H. (1998) *Science* **281**, 660–665

70. Levin, J. R., Blake, J. J., Ganunis, R. A., and Tullius, T. D. (2000) *J. Biol. Chem.* **275**, 6885–6893

71. Yager, T. D., and von Hippel, P. H. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed), pp. 1241–1275, Am. Soc. for Microbiology, Washington, DC

72. Barile, M. F., Razin, S., Tully, J. G., and Whitcomb, R. F., eds. (1979) *The Mycoplasmas*, Vol. I, Academic Press, New York

73. Harris, E. H. (1989) *The Chlamydomonas Sourcebook*, Academic Press, San Diego, California

74. Milan, S., D'Ari, L., and Chamberlin, M. J. (1999) *Biochemistry* **38**, 218–225

75. Davenport, R. J., Wuite, G. J. L., Landick, R., and Bustamante, C. (2000) *Science* **287**, 2497–2500

76. Ellinger, T., Behnke, D., Bujard, H., and Gralla, J. D. (1994) *J. Mol. Biol.* **239**, 455–465

76a. Lee, K.-B., Wang, D., Lippard, S. J., and Sharp, P. A. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4239–4244

77. Selby, C. P., and Sancar, A. (1993) *Science* **260**, 53–58

78. Viswanathan, A., You, H. J., and Doetsch, P. W. (1999) *Science* **284**, 159–162

79. Nudler, E. (1999) *J. Mol. Biol.* **288**, 1–12

80. Erie, D. A., Hajiseyedjavadi, O., Young, M. C., and von Hippel, P. H. (1993) *Science* **262**, 867–873

81. Frank-Kamenetskii, M. (1989) *Nature (London)* **337**, 206

82. Krasilnikov, A. S., Podtelezhnikov, A., Vologodskii, A., and Mirkin, S. M. (1999) *J. Mol. Biol.* **292**, 1149–1160

83. Wu, H. (1988) *Cell* **53**, 433–440

84. Giaever, G. N., and Wang, J. C. (1988) *Cell* **55**, 849–856

85. Miller, O. L., Jr. (1973) *Sci. Am.* **228**(March), 34–42

86. Weickert, M. J., and Adhya, S. (1992) *J. Biol. Chem.* **267**, 15869–15874

87. Takeda, Y., Kim, J. G., Caday, C. G., Steers, E., Jr., Ohlendorf, D. H., Anderson, W. F., and Matthews, B. W. (1986) *J. Biol. Chem.* **261**, 8608–8616

88. Matthews, B. W., Ohlendorf, D. H., Anderson, W. F., Fisher, R. G., and Takedo, Y. (1983) *Trends Biochem. Sci.* **8**, 25–29

89. Jana, R., Hazbun, T. R., Fields, J. D., and Mossing, M. C. (1998) *Biochemistry* **37**, 6446–6455

90. Pabo, C. O., and Lewis, M. (1982) *Nature (London)* **298**, 443–447

90a. Bell, C. E., and Lewis, M. (2001) *J. Mol. Biol.* **314**, 1127–1136

91. Jordan, S. R., and Pabo, C. O. (1988) *Science* **242**, 893–899

92. Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M., and Harrison, S. C. (1988) *Science* **242**, 899–907

93. Wolberger, C., Dong, Y., Ptashne, M., and Harrison, S. C. (1988) *Nature (London)* **335**, 789–795

94. Padmanabhan, S., Jiménez, M. A., González, C., Sanz, J. M., Giménez-Gallego, G., and Rico, M. (1997) *Biochemistry* **36**, 6424–6436

95. Wharton, R. P., and Ptashne, M. (1985) *Nature (London)* **316**, 601–605

96. Wharton, R. P., and Ptashne, M. (1986) *Trends Biochem. Sci.* **11**, 71–73

97. Sauer, R. T., Milla, M. E., Waldburger, C. D., Brown, B. M., and Schildbach, J. F. (1996) *FASEB J.* **10**, 42–48

98. Nooren, I. M. A., Rietveld, A. W. M., Melacini, G., Sauer, R. T., Kaptein, R., and Boelens, R. (1999) *Biochemistry* **38**, 6035–6042

99. He, Y.-y, Stockley, P. G., and Gold, L. (1996) *J. Mol. Biol.* **255**, 55–66

100. Dai, X., Kloster, M., and Rothman-Denes, L. B. (1998) *J. Mol. Biol.* **283**, 43–58

101. Wang, J. C., Barkley, M. D., and Bourgeois, S. (1974) *Nature (London)* **251**, 247–249

102. Horwitz, M. S. Z., and Loeb, L. A. (1988) *Science* **241**, 703–705

103. Arndt, K. T., Boschelli, F., Cook, J., Takeda, Y., Tecza, E., and Lu, P. (1983) *J. Biol. Chem.* **258**, 4177–4183

104. Metzler, W. J., and Lu, P. (1989) *J. Mol. Biol.* **205**, 149–164

105. Borowiec, J. A., and Gralla, J. D. (1986) *Biochemistry* **25**, 5051–5057

106. Ho, Y.-S., Wulff, D. L., and Rosenberg, M. (1983) *Nature (London)* **304**, 703–708

107. Brennan, R. G., Vasu, S., Matthews, B. W., and Otsuka, A. J. (1989) *J. Biol. Chem.* **264**, 5

References

108. Eisenstein, E., and Beckett, D. (1999) *Biochemistry* **38**, 13077–13084
109. Streaker, E. D., and Beckett, D. (1999) *J. Mol. Biol.* **292**, 619–632
109a. Weaver, L. H., Kwon, K., Beckett, D., and Matthews, B. W. (2001) *Protein Sci.* **10**, 2618–2622
110. Dahl, M. K., Degenkolb, J., and Hillen, W. (1994) *J. Mol. Biol.* **243**, 413–424
111. Yoshida, K.-I., Shibayama, T., Aoyama, D., and Fujita, Y. (1999) *J. Mol. Biol.* **285**, 917–929
112. Hinrichs, W., Kisker, C., Düvel, M., Müller, A., Tovar, K., Hillen, W., and Saenger, W. (1994) *Science* **264**, 418–420
113. Kisker, C., Hinrichs, W., Tovar, K., Hillen, W., and Saenger, W. (1995) *J. Mol. Biol.* **247**, 260–280
113a. Schubert, P., Schnappinger, D., Pfleiderer, K., and Hillen, W. (2001) *Biochemistry* **40**, 3257–3263
113b. Lathe, W. C., III, Snel, B., and Bork, P. (2000) *Trends Biochem. Sci.* **25**, 474–479
114. Cohen, G. N. (1986) in *Regulation of Gene Expression* (Booth, I. R., and Higgins, C. F., eds), pp. 1–20, Cambridge Univ. Press, London
115. Cornish, E. C., Argyropoulos, V. P., Pittard, J., and Davidson, B. E. (1986) *J. Biol. Chem.* **261**, 403–410
116. Weiss, D. L., Johnson, D. I., Weith, H. L., and Somerville, R. L. (1986) *J. Biol. Chem.* **261**, 9966–9971
117. Choi, K. Y., Lu, F., and Zalkin, H. (1994) *J. Biol. Chem.* **269**, 24066–24072
118. Xu, H., Moraitis, M., Reedstrom, R. J., and Matthews, K. S. (1998) *J. Biol. Chem.* **273**, 8958–8964
119. Glasfeld, A., Koehler, A. N., Schumacher, M. A., and Brennan, R. G. (1999) *J. Mol. Biol.* **291**, 347–361
120. Charlier, D., Roovers, M., Van Vliet, F., Boyen, A., Cunin, R., Nakamura, Y., Glansdorff, N., and Piérard, A. (1992) *J. Mol. Biol.* **226**, 367–386
121. Van Duyne, G. D., Ghosh, G., Maas, W. K., and Sigler, P. B. (1996) *J. Mol. Biol.* **256**, 377–391
122. Wang, H., Glansdorff, N., and Charlier, D. (1998) *J. Mol. Biol.* **277**, 805–824
123. Ni, J., Sakanyan, V., Charlier, D., Glansdorff, N., and Van Duyne, G. D. (1999) *Nature Struct. Biol.* **6**, 427–432
124. Escolar, L., Pérez-Martín, J., and de Lorenzo, V. (1998) *J. Mol. Biol.* **283**, 537–547
125. White, A., Ding, X., vanderSpek, J. C., Murphy, J. R., and Ringe, D. (1998) *Nature (London)* **394**, 502–506
126. Pohl, E., Holmes, R. K., and Hol, W. G. J. (1998) *J. Biol. Chem.* **273**, 22420–22427
127. Pohl, E., Holmes, R. K., and Hol, W. G. J. (1999) *J. Mol. Biol.* **285**, 1145–1156
127a. Gourley, D. G., Schüttelkopf, A. W., Anderson, L. A., Price, N. C., Boxer, D. H., and Hunter, W. N. (2001) *J. Biol. Chem.* **276**, 20641–20647
128. Postma, P. W. (1986) in *Regulation of Gene Expression* (Booth, I. R., and Higgins, C. F., eds), Cambridge Univ. Press, London
129. Busby, S., and Ebright, R. H. (1999) *J. Mol. Biol.* **293**, 199–213
130. Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991) *Science* **253**, 1001–1007
131. Passner, J. M., and Steitz, T. A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2843–2847
132. Parkinson, G., Wilson, C., Gunasekera, A., Ebright, Y. W., Ebright, R. E., and Berman, H. M. (1996) *J. Mol. Biol.* **260**, 395–408
133. Gunasekera, A., Ebright, Y. W., and Ebright, R. H. (1992) *J. Biol. Chem.* **267**, 14713–14720
134. Vossen, K. M., Stickle, D. F., and Fried, M. G. (1996) *J. Mol. Biol.* **255**, 44–54
135. Dalma-Weiszhausz, D. D., and Brenowitz, M. (1996) *Biochemistry* **35**, 3735–3745
136. Roy, S., Garges, S., and Adhya, S. (1998) *J. Biol. Chem.* **273**, 14059–14062
137. Danot, O., Vidal-Ingigliardi, D., and Raibaud, O. (1996) *J. Mol. Biol.* **262**, 1–11
138. Richet, E., and Sogaard-Andersen, L. (1994) *EMBO J.* **13**, 4558–4567
139. Ziegelhoffer, E. C., and Kiley, P. J. (1995) *J. Mol. Biol.* **245**, 351–361
140. Lazazzera, B. A., Beinert, H., Khoroshilova, N., Kennedy, M. C., and Kiley, P. J. (1996) *J. Biol. Chem.* **271**, 2762–2768
141. Popescu, C. V., Bates, D. M., Beinert, H., Münck, E., and Kiley, P. J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13431–13435
142. Shelver, D., Thorsteinsson, M. V., Kerby, R. L., Chung, S.-Y., Roberts, G. P., Reynolds, M. F., Parks, R. B., and Burstyn, J. N. (1999) *Biochemistry* **38**, 2669–2678
142a. Thorsteinsson, M. V., Kerby, R. L., Youn, H., Conrad, M., Serate, J., Staples, C. R., and Roberts, G. P. (2001) *J. Biol. Chem.* **276**, 26807–26813
143. Greene, J. R., Morrissey, L. M., Foster, L. M., and Geiduschek, E. P. (1986) *J. Biol. Chem.* **261**, 12820–12827
144. Takahashi, K., Kawazoe, Y., Sakumi, K., Nakabeppu, Y., and Sekiguchi, M. (1988) *J. Biol. Chem.* **263**, 13490–13492
145. Lobell, R. B., and Schleif, R. F. (1990) *Science* **250**, 528–532
146. Zhang, X., Reeder, T., and Schleif, R. (1996) *J. Mol. Biol.* **258**, 14–24
147. Niland, P., Hühne, R., and Müller-Hill, B. (1996) *J. Mol. Biol.* **264**, 667–674
147a. Wu, M., and Schleif, R. (2001) *J. Mol. Biol.* **307**, 1001–1009
147b. LaRonde-LeBlanc, N., and Wolberger, C. (2000) *Biochemistry* **39**, 11593–11601
148. Echols, H. (1990) *J. Biol. Chem.* **265**, 14697–14700
149. Rippe, K., von Hippel, P. H., and Langowski, J. (1995) *Trends Biochem. Sci.* **20**, 500–506
150. Blackwood, E. M., and Kadonaga, J. T. (1998) *Science* **281**, 60–63
151. Magasanik, B. (1988) *Trends Biochem. Sci.* **13**, 475–479
152. Rippe, K., Guthold, M., von Hippel, P. H., and Bustamante, C. (1997) *J. Mol. Biol.* **270**, 125–138
153. Flashner, Y., Weiss, D. S., Keener, J., and Kustu, S. (1995) *J. Mol. Biol.* **249**, 700–713
153a. Schulz, A., Langowski, J., and Rippe, K. (2000) *J. Mol. Biol.* **300**, 709–725
154. Brahms, G., Brahms, S., and Magasanik, B. (1995) *J. Mol. Biol.* **246**, 35–42
155. Cullen, P. J., Bowman, W. C., Hartnett, D.-F., Reilly, S. C., and Kranz, R. G. (1998) *J. Mol. Biol.* **278**, 903–914
156. Dworkin, J., Jovanovic, G., and Model, P. (1997) *J. Mol. Biol.* **273**, 377–388
157. Solà, M., Gomis-Rüth, F. X., Serrano, L., González, A., and Coll, M. (1999) *J. Mol. Biol.* **285**, 675–687
158. Makino, K., Amemura, M., Kawamoto, T., Kimura, S., Shinagawa, H., Nakata, A., and Suzuki, M. (1996) *J. Mol. Biol.* **259**, 15–26
159. Shi, L., Liu, W., and Hulett, F. M. (1999) *Biochemistry* **38**, 10119–10125
159a. Georgellis, D., Kwon, O., and Lin, E. C. C. (2001) *Science* **292**, 2314–2316
160. Strohmaier, H., Noiges, R., Kotschan, S., Sawers, G., Högenauer, G., Zechner, E. L., and Koraimann, G. (1998) *J. Mol. Biol.* **277**, 309–316
161. Miyatake, H., Mukai, M., Adachi, S.-i, Nakamura, H., Tamura, K., Iizuka, T., Shiro, Y., Strange, R. W., and Hasnain, S. S. (1999) *J. Biol. Chem.* **274**, 23176–23184
161a. Rusnak, F., and Reiter, T. (2000) *Trends Biochem. Sci.* **25**, 527–529
162. Li, J., Kustu, S., and Stewart, V. (1994) *J. Mol. Biol.* **241**, 150–165
163. Asayama, M., Yamamoto, A., and Kobayashi, Y. (1995) *J. Mol. Biol.* **250**, 11–23
164. Madhusudan, Zapf, J., Hoch, J. A., Whiteley, J. M., Xuong, N. H., and Varughese, K. I. (1997) *Biochemistry* **36**, 12739–12745
165. Eguchi, Y., Itoh, T., and Tomizawa, J.-i. (1991) *Ann. Rev. Biochem.* **60**, 631–652
166. Weintraub, H. M. (1990) *Sci. Am.* **262**(Jan), 40–46
166a. Nellen, W., and Lichtenstein, C. (1993) *Trends Biochem. Sci.* **18**, 419–423
166b. Delihas, N., and Forst, S. (2001) *J. Mol. Biol.* **313**, 1–12
167. Green, P. J., Pines, O., and Inouye, M. (1986) *Ann. Rev. Biochem.* **55**, 569–597
168. Cesareni, G., and Banner, D. W. (1985) *Trends Biochem. Sci.* **10**, 303–306
169. Asano, K., Niimi, T., Yokoyama, S., and Mizobuchi, K. (1998) *J. Biol. Chem.* **273**, 11826–11838
170. Ikenaka, K., Ramakrishnan, G., Inouye, M., Tsung, K., and Inouye, M. (1986) *J. Biol. Chem.* **261**, 9316–9320
171. Andersen, J., Forst, S. A., Zhao, K., Inouye, M., and Delihas, N. (1989) *J. Biol. Chem.* **264**, 17961–17970
171a. Matzke, M., Matzke, A. J. M., and Kooter, J. M. (2001) *Science* **293**, 1080–1083
172. Madison-Antenucci, S., Sabatini, R. S., Pollard, V. W., and Hajduk, S. L. (1998) *EMBO J.* **17**, 6368–6376
173. Nicoloso, M., Qu, L.-H., Michot, B., and Bachellerie, J.-P. (1996) *J. Mol. Biol.* **260**, 178–195
174. Lafontaine, D. L. J., and Tollervey, D. (1998) *Trends Biochem. Sci.* **23**, 383–388
174a. Gerdes, K., Nielsen, A., Thorsted, P., and Wagner, E. G. H. (1992) *J. Mol. Biol.* **226**, 637–649
175. Cohen, J. S., and Hogan, M. E. (1994) *Sci. Am.* **271**(Jun), 76–82
176. Askari, F. K., and McDonnell, W. M. (1996) *N. Engl. J. Med.* **334**, 316–318
176a. Sullenger, B. A., and Gilboa, E. (2002) *Nature (London)* **418**, 252–258
176b. Gitlin, L., Karelsky, S., and Andino, R. (2002) *Nature (London)* **418**, 430–434
177. Wittung-Stafshede, P. (1998) *Science* **281**, 657–658
178. Wagner, R. W. (1994) *Nature (London)* **372**, 333–335
179. Tereshko, V., Gryaznov, S., and Egli, M. (1998) *J. Am. Chem. Soc.* **120**, 269–283
180. Aramini, J. M., Kalisch, B. W., Pon, R. T., van de Sande, J. H., and Germann, M. W. (1996) *Biochemistry* **35**, 9355–9365
181. Mujeeb, A., Reynolds, M. A., and James, T. L. (1997) *Biochemistry* **36**, 2371–2379
182. Branch, A. D. (1998) *Trends Biochem. Sci.* **23**, 45–50
182a. Braasch, D. A., and Corey, D. R. (2002) *Biochemistry* **41**, 4503–4510
183. Platt, T. (1986) *Ann. Rev. Biochem.* **55**, 339–372
184. Yanofsky, C. (1981) *Nature (London)* **289**, 751–758
185. Yanofsky, C. (1988) *J. Biol. Chem.* **263**, 609–612
186. Kuroda, M. I., and Yanofsky, C. (1984) *J. Biol. Chem.* **259**, 12838–12843
187. Baumann, C., Otridge, J., and Gollnick, P. (1996) *J. Biol. Chem.* **271**, 12269–12274
187a. Losick, R., and Sonenshein, A. L. (2001) *Science* **293**, 2018–2019
187b. Valbuzzi, A., and Yanofsky, C. (2001) *Science* **293**, 2057–2059
187c. Yakhnin, A. V., and Babitzke, P. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11067–11072

References

188. Antson, A. A., Dodson, E. J., Dodson, G., Greaves, R. B., Chen, X.-p, and Gollnick, P. (1999) *Nature (London)* **401**, 235 – 242
189. Jensen, K. F., Bonekamp, F., and Poulsen, P. (1986) *Trends Biochem. Sci.* **11**, 362 – 365
190. Hauser, C. A., Sharp, J. A., Hatfield, L. K., and Hatfield, G. W. (1985) *J. Biol. Chem.* **260**, 1765 – 1770
191. Watson, M. D. (1981) *Trends Biochem. Sci.* **6**, 180 – 182
192. Stitt, B. L. (1988) *J. Biol. Chem.* **263**, 11130 – 11137
193. Bear, D. G., and Peabody, D. S. (1988) *Trends Biochem. Sci.* **13**, 343 – 347
194. Nehrke, K. W., and Platt, T. (1994) *J. Mol. Biol.* **243**, 830 – 839
195. Liu, K., and Hanna, M. M. (1995) *J. Mol. Biol.* **247**, 547 – 558
196. Greenblatt, J. (1991) *Trends Biochem. Sci.* **16**, 408 – 411
196a. Toulokhonov, I., Artsimovitch, I., and Landick, R. (2001) *Science* **292**, 730 – 733
196b. Gopal, B., Haire, L. F., Gamblin, S. J., Dodson, E. J., Lane, A. N., Papavinasasundaram, K. G., Colston, M. J., and Dodson, G. (2001) *J. Mol. Biol.* **314**, 1087 – 1095
196c. Zhou, Y., Mah, T.-F., Yu, Y.-T. N., Mogridge, J., Olson, E. R., Greenblatt, J., and Friedman, D. I. (2001) *J. Mol. Biol.* **310**, 33 – 49
197. Zhu, A. Q., and von Hippel, P. H. (1998) *Biochemistry* **37**, 11202 – 11214
198. Kim, D.-E., and Patel, S. S. (1999) *J. Biol. Chem.* **274**, 32667 – 32671
198a. Vincent, F., Openshaw, M., Trautwein, M., Gaskell, S. J., Kohn, H., and Widger, W. R. (2000) *Biochemistry* **39**, 9077 – 9083
198b. Yu, X., Horiguchi, T., Shigesada, K., and Egelman, E. H. (2000) *J. Mol. Biol.* **299**, 1279 – 1287
199. Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993) *Nature (London)* **364**, 401 – 406
200. O'Hara, B. P., Norman, R. A., Wan, P. T. C., Roe, S. M., Barrett, T. E., Drew, R. E., and Pearl, L. H. (1999) *EMBO J.* **18**, 5175 – 5186
200a. Gopal, B., Papavinasasundaram, K. G., Dodson, G., Colston, M. J., Major, S. A., and Lane, A. N. (2001) *Biochemistry* **40**, 920 – 928
200b. Carlomagno, M. S., and Nappo, A. (2001) *J. Mol. Biol.* **309**, 19 – 28
201. Sozhamannan, S., and Stitt, B. L. (1997) *J. Mol. Biol.* **268**, 689 – 703
202. Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol.* **166**, 477 – 535
203. Daniels, C. J., Gupta, R., and Doolittle, W. F. (1985) *J. Biol. Chem.* **260**, 3132 – 3134
204. Lykke-Andersen, J., Aagaard, C., Semionenkov, M., and Garrett, R. A. (1997) *Trends Biochem. Sci.* **22**, 326 – 331
205. Abelson, J., Trotta, C. R., and Li, H. (1998) *J. Biol. Chem.* **273**, 12685 – 12688
206. Starzyk, R. M. (1986) *Trends Biochem. Sci.* **11**, 60
207. Sollner-Webb, B., and Mougey, E. B. (1991) *Trends Biochem. Sci.* **16**, 58 – 62
208. Li, Z., Pandit, S., and Deutscher, M. P. (1999) *EMBO J.* **18**, 2878 – 2885
208a. Nomura, M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1820 – 1822
209. Perry, R. P. (1976) *Ann. Rev. Biochem.* **45**, 605 – 629
210. King, T. C., Sirdeshmukh, R., and Schlessinger, D. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 185 – 188
211. Denoya, C., and Dubnau, D. (1989) *J. Biol. Chem.* **264**, 2615 – 2624
212. Tscherne, J. S., Nurse, K., Popienick, P., Michel, H., Sochacki, M., and Ofengand, J. (1999) *Biochemistry* **38**, 1884 – 1892
213. Stahl, D. A., Pace, B., Marsh, T., and Pace, N. R. (1984) *J. Biol. Chem.* **259**, 11448 – 11453
214. Raychaudhuri, S., Niu, L., Conrad, J., Lane, B. G., and Ofengand, J. (1999) *J. Biol. Chem.* **274**, 18880 – 18886
215. Niu, L., and Ofengand, J. (1999) *Biochemistry* **38**, 629 – 635
216. Ramamurthy, V., Swann, S. L., Spedaliere, C. J., and Mueller, E. G. (1999) *Biochemistry* **38**, 13106 – 13111
217. Gu, X., Liu, Y., and Santi, D. V. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14270 – 14275
218. Conrad, J., Sun, D., Englund, N., and Ofengand, J. (1998) *J. Biol. Chem.* **273**, 18562 – 18566
218a. Spedaliere, C. J., Hamilton, C. S., and Mueller, E. G. (2000) *Biochemistry* **39**, 9459 – 9465
218b. Arluison, V., Buckle, M., and Grosjean, H. (1999) *J. Mol. Biol.* **289**, 491 – 502
219. Kiss-László, Z., Henry, Y., and Kiss, T. (1998) *EMBO J.* **17**, 797 – 807
220. Segal, D. M., and Eichler, D. C. (1991) *J. Biol. Chem.* **266**, 24385 – 24389
221. Frank, D. N., and Pace, N. R. (1998) *Ann. Rev. Biochem.* **67**, 153 – 180
222. Christian, E. L., and Harris, M. E. (1999) *Biochemistry* **38**, 12629 – 12638
222a. Zuleeg, T., Hansen, A., Pfeiffer, T., Schübel, H., Kreutzer, R., Hartmann, R. K., and Limmer, S. (2001) *Biochemistry* **40**, 3363 – 3369
222b. Cole, K. B., and Dorit, R. L. (2001) *J. Mol. Biol.* **307**, 1181 – 1193
222c. Mikkelsen, N. E., Brännvall, M., Virtanen, A., and Kirsebom, L. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6155 – 6160
223. Misra, T. K., and Apirion, D. (1978) *J. Biol. Chem.* **253**, 5594 – 5599
224. Nashimoto, M., Tamura, M., and Kaspar, R. L. (1999) *J. Mol. Biol.* **287**, 727 – 740
225. Deutscher, M. P., and Marlor, C. W. (1985) *J. Biol. Chem.* **260**, 7067 – 7071
226. Robertson, H. D., Altman, S., and Smith, J. D. (1972) *J. Biol. Chem.* **247**, 5243 – 5251
227. Reilly, R. M., and RajBhandary, U. L. (1986) *J. Biol. Chem.* **261**, 2928 – 2935
228. Khorana, H. G., and 19 other authors. (1976) *J. Biol. Chem.* **251**, 565 – 570
229. Ramamoorthy, B., Lees, R. G., Kleid, D. G., and Khorana, H. G. (1976) *J. Biol. Chem.* **251**, 676 – 694
230. Sekiya, T., Contreras, R., Kupper, H., Landy, A., and Khorana, H. G. (1976) *J. Biol. Chem.* **251**, 5124 – 5140
231. Gommers-Ampt, J. H., and Borst, P. (1995) *FASEB J.* **9**, 1034 – 1042
232. Söll, D., and RajBhandary, U. L., eds. (1995) *tRNA Structure, Biosynthesis, and Function*, Am. Soc. for Microbiology, Washington, D.C.
233. Deutscher, M. P., Marlor, C. W., and Zaniewski, R. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6429 – 6430
234. Yue, D., Weiner, A. M., and Maizels, N. (1998) *J. Biol. Chem.* **273**, 29693 – 29700
235. Kline, L. K., and Söll, D. (1982) in *The Enzymes*, Vol. 15 (Boyer, P. D., ed), pp. 567 – ___, Academic Press, New York
236. Greenberg, R., and Dudock, B. (1980) *J. Biol. Chem.* **255**, 8296 – 8302
237. Kwong, L. K., Moore, V. G., and Kaiser, I. I. (1977) *J. Biol. Chem.* **252**, 6310 – 6315
238. Becker, H. F., Motorin, Y., Sissler, M., Florentz, C., and Grosjean, H. (1997) *J. Mol. Biol.* **274**, 505 – 518
239. Kambampati, R., and Lauhon, C. T. (2000) *J. Biol. Chem.* **275**, 10727 – 10730
240. Kaiser, J. T., Clausen, T., Bourenkow, G. P., Bartunik, H.-D., Steinbacher, S., and Huber, R. (2000) *J. Mol. Biol.* **297**, 451 – 464
241. Schaeffer, K. P., Altman, S., and Soll, D. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3626 – 3630
242. Reyniers, J. P., Pleasants, J. R., Wostmann, B. S., Katze, J. R., and Farkas, W. R. (1981) *J. Biol. Chem.* **256**, 11591 – 11594
243. Katze, J. R., Gündüz, U., Smith, D. L., Cheng, C. S., and McCloskey, J. A. (1984) *Biochemistry* **23**, 1171 – 1176
244. Slany, R. K., Bösl.M, Crain, P. F., and Kersten, H. (1993) *Biochemistry* **32**, 7811 – 7817
245. Romier, C., Reuter, K., Suck, D., and Ficner, R. (1996) *Biochemistry* **35**, 15734 – 15739
246. Curnow, A. W., and Garcia, G. A. (1995) *J. Biol. Chem.* **270**, 17264 – 17267
246a. Kittendorf, J. D., Barcomb, L. M., Nonekowski, S. T., and Garcia, G. A. (2001) *Biochemistry* **40**, 14123 – 14133
247. Phillipson, D. W., Edmonds, C. G., Crain, P. F., Smith, D. L., Davis, D. R., and McCloskey, J. A. (1987) *J. Biol. Chem.* **262**, 3462 – 3471
248. Kolesky, S., Ouhammouch, M., Brody, E. N., and Geiduschek, E. P. (1999) *J. Mol. Biol.* **291**, 267 – 281
249. Gopal, V., Brieba, L. G., Guajardo, R., McAllister, W. T., and Sousa, R. (1999) *J. Mol. Biol.* **290**, 411 – 431
249a. Liu, C., and Martin, C. T. (2001) *J. Mol. Biol.* **308**, 465 – 475
249b. Severinov, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5 – 7
250. Gottesman, M. (1999) *J. Mol. Biol.* **293**, 177 – 180
251. Gottesman, S. (1988) in *Biochemistry*, 2nd ed. (Zubay, G., ed), pp. 1000 – 1007, Macmillan, New York
252. Johnson, A. D., Poteete, A. R., Lauer, G., Sauer, R. T., Ackers, G. K., and Ptashne, M. (1981) *Nature (London)* **294**, 217 – 223
253. Ptashne, M., Johnson, A. D., and Pabo, C. O. (1982) *Sci. Am.* **247**(Nov), 128 – 140
254. Albright, R. A., and Matthews, B. W. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3431 – 3436
255. Honigman, A., Hu, S.-L., Chase, R., and Szbalski, W. (1976) *Nature (London)* **262**, 112 – 116
256. Shimatake, H., and Rosenberg, M. (1981) *Nature (London)* **292**, 128 – 132
257. Maurizi, M. R. (1987) *J. Biol. Chem.* **262**, 2696 – 2793
258. Kastelein, R. A., Remaut, E., Fiers, W., and van Duin, J. (1982) *Nature (London)* **295**, 35 – 41
259. Voyles, B. A. (1993) *The Biology of Viruses*, Mosby, St. Louis, Missouri
260. Brown, D., and Gold, L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11558 – 11562
261. Schuppli, D., Georgijevic, J., and Weber, H. (2000) *J. Mol. Biol.* **295**, 149 – 154
262. Blumenthal, T., and Carmichael, G. G. (1979) *Ann. Rev. Biochem.* **48**, 525 – 548
263. Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and VandePol, S. (1982) *Science* **215**, 1577 – 1585
264. Mills, D. R., Kramer, F. R., Dobkin, C., Nishihara, T., and Spiegelman, S. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4252 – 4256
264a. Weinzierl, R. O. J. (1999) *Mechanisms of Gene Expression.Structure,Function and Evolution of the Basal Transcriptional Machinery*, Imperial College Press,
265. Langer, D., Hain, J., Thuriaux, P., and Zillig, W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5768 – 5772
266. Bell, S. D., Kosa, P. L., Sigler, P. B., and Jackson, S. P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13662 – 13667
266a. Bell, S. D., and Jackson, S. P. (2000) *J. Biol. Chem.* **275**, 31624 – 31629
266b. Thomsen, J., De Biase, A., Kaczanowski, S., Macario, A. J. L., Thomm, M., Zielenkiewicz, P., MacColl, R., and de Macario, E. C. (2001) *J. Mol. Biol.* **309**, 589 – 603
267. Krüger, K., Hermann, T., Armbruster, V., and Pfeifer, F. (1998) *J. Mol. Biol.* **279**, 761 – 771

References

268. Kyrpides, N. C., and Ouzounis, C. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8545 – 8550

268a. Lewis, P. J., Thaker, S. D., and Errington, J. (2000) *EMBO J.* **19**, 710 – 718

268b. Iborra, F. J., Jackson, D. A., and Cook, P. R. (2001) *Science* **293**, 1139 – 1142

268c. von Mering, C., and Bork, P. (2002) *Nature (London)* **417**, 797 – 798

269. Jackson, D. A., Pombo, A., and Iborra, F. (2000) *FASEB J.* **14**, 242 – 254

270. Workman, J. L., and Kingston, R. E. (1998) *Ann. Rev. Biochem.* **67**, 545 – 579

271. Hagmann, M. (1999) *Science* **285**, 1200 – 1203

271a. Jenuwein, T., and Allis, C. D. (2001) *Science* **293**, 1074 – 1080

271b. Bird, A. (2001) *Science* **294**, 2113 – 2115

271c. Gamble, M. J., and Freedman, L. P. (2002) *Trends Biochem. Sci.* **27**, 165 – 167

271d. Burke, L. J., and Baniahmad, A. (2000) *FASEB J.* **14**, 1876 – 1888

272. Kornberg, R. D. (1999) *Trends Biochem. Sci.* **24**, M46 – M49

273. Carey, M., and Smale, S. T. (2000) *Transcriptional Regulation in Eukaryotes*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York

273a. Steidl, S., Hynes, M. J., and Brakhage, A. A. (2001) *J. Mol. Biol.* **306**, 643 – 653

273b. Niehof, M., Kubicka, S., Zender, L., Manns, M. P., and Trautwein, C. (2001) *J. Mol. Biol.* **309**, 855 – 868

274. Vandromme, M., Gauthier-Rouvière, C., Lamb, N., and Fernandez, A. (1996) *Trends Biochem. Sci.* **21**, 59 – 64

275. Pennisi, E. (2000) *Science* **288**, 1372 – 1373

275a. Lo, W.-S., Duggan, L., Emre, N. C. T., Belotserkovskya, R., Lane, W. S., Shiekhattar, R., and Berger, S. L. (2001) *Science* **293**, 1142 – 1146

275b. Wang, H., Huang, Z.-Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P., and Zhang, Y. (2001) *Science* **293**, 853 – 857

275c. Nakayama, J.-i, Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. S. (2001) *Science* **292**, 110 – 113

275d. Jenuwein, T. (2002) *Science* **297**, 2215 – 2218

275e. Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. S., and Martienssen, R. A. (2002) *Science* **297**, 1833 – 1837

276. Chandler, S. P., Guschin, D., Landsberger, N., and Wolffe, A. P. (1999) *Biochemistry* **38**, 7008 – 7018

276a. Burnett, E., Christensen, J., and Tattersall, P. (2001) *J. Mol. Biol.* **314**, 1029 – 1039

277. Ng, H. H., and Bird, A. (2000) *Trends Biochem. Sci.* **25**, 121 – 126

278. Alland, L., Muhle, R., Hou, H., Jr., Potes, J., Chin, L., Schreiber-Agus, N., and DePinho, R. A. (1997) *Nature (London)* **387**, 49 – 55

278a. Arévalo-Rodríguez, M., Cardenas, M. E., Wu, X., Hanes, S. D., and Heitman, J. (2000) *EMBO J.* **19**, 3739 – 3749

279. Heinzel, T., Lavinsky, R. M., Mullen, T.-M., Söderström, M., Laherty, C. D., Torchia, J., Yang, W.-M., Brard, G., Ngo, S. D., Davie, J. R., Seto, E., Eisenman, R. N., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. (1997) *Nature (London)* **387**, 43 – 48

279a. Lechner, T., Lusser, A., Pipal, A., Brosch, G., Loidl, A., Goralik-Schramel, M., Sendra, R., Wegener, S., Walton, J. D., and Loidl, P. (2000) *Biochemistry* **39**, 1683 – 1692

279b. Bernstein, B. E., Tong, J. K., and Schreiber, S. L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13708 – 13713

279c. Johnson, C. A., Padget, K., Austin, C. A., and Turner, B. M. (2001) *J. Biol. Chem.* **276**, 4539 – 4542

280. Bernhard, D., Ausserlechner, M. J., Tonko, M., Löffler, M., Hartmann, B. L., Csordas, A., and Kofler, R. (1999) *FASEB J.* **13**, 1991 – 2001

280a. Gross, D. S. (2001) *Trends Biochem. Sci.* **26**, 685 – 686

280b. Smith, J. S., and Boeke, J. D. (2001) *Science* **291**, 608 – 609

280c. Tanny, J. C., and Moazed, D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 415 – 420

280d. Sauve, A. A., Celic, I., Avalos, J., Deng, H., Boeke, J. D., and Schramm, V. L. (2001) *Biochemistry* **40**, 15456 – 15463

280e. Ghosh, D., Gerasimova, T. I., and Corces, V. G. (2001) *EMBO J.* **20**, 2518 – 2527

280f. Donze, D., and Kamakaka, R. T. (2001) *EMBO J.* **20**, 520 – 531

280g. Bell, A. C., West, A. G., and Felsenfeld, G. (2001) *Science* **291**, 447 – 450

280h. Tijsterman, M., Ketting, R. F., Okihara, K. L., Sijen, T., and Plasterk, R. H. A. (2002) *Science* **295**, 694 – 697

280i. Ahlquist, P. (2002) *Science* **296**, 1270 – 1273

280j. Jorgensen, R. A. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11561 – 11563

281. Kouzarides, T. (2000) *EMBO J.* **19**, 1176 – 1179

282. Brown, C. E., Lechner, T., Howe, L., and Workman, J. L. (2000) *Trends Biochem. Sci.* **25**, 15 – 19

283. Trievel, R. C., Rojas, J. R., Sterner, D. E., Venkataramani, R. N., Wang, L., Zhou, J., Allis, C. D., Berger, S. L., and Marmorstein, R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8931 – 8936

283a. Sterner, D. E., Belotserkovskaya, R., and Berger, S. L. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11622 – 11627

283b. Hughes, J. D., Estep, P. W., Tavazoie, S., and Church, G. M. (2000) *J. Mol. Biol.* **296**, 1205 – 1214

283c. Chen, J., Rappsilber, J., Chiang, Y.-C., Russell, P., Mann, M., and Denis, C. L. (2001) *J. Mol. Biol.* **314**, 683 – 694

284. Bochar, D. A., Savard, J., Wang, W., Lafleur, D. W., Moore, P., Côté, J., and Shiekhattar, R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1038 – 1043

284a. Fry, C. J., and Peterson, C. L. (2002) *Science* **295**, 1847 – 1848

284b. Aalfs, J. D., and Kingston, R. E. (2000) *Trends Biochem. Sci.* **25**, 548 – 555

284c. Becker, P. B., and Hörz, W. (2002) *Ann. Rev. Biochem.* **71**, 247 – 273

285. Kwon, H., Imbalzano, A. N., Khavarl, P. A., Kingston, R. E., and Green, M. R. (1994) *Nature (London)* **370**, 477 – 481

286. Richard-Foy, H. (1994) *Nature (London)* **370**, 417 – 418

287. Gavin, I. M., and Simpson, R. T. (1997) *EMBO J.* **16**, 6263 – 6271

287a. Xue, Y., Canman, J. C., Lee, C. S., Nie, Z., Yang, D., Moreno, G. T., Young, M. K., Salmon, E. D., and Wang, W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13015 – 13020

287b. Olave, I. A., Reck-Peterson, S. L., and Crabtree, G. R. (2002) *Ann. Rev. Biochem.* **71**, 755 – 781

287c. Asturias, F. J., Chung, W.-H., Kornberg, R. D., and Lorch, Y. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13477 – 13480

288. de la Cruz, J., Kressler, D., and Linder, P. (1999) *Trends Biochem. Sci.* **24**, 192 – 198

288a. Linder, P., Tanner, N. K., and Banroques, J. (2001) *Trends Biochem. Sci.* **26**, 339 – 341

288b. Henn, A., Medalia, O., Shi, S.-P., Steinberg, M., Franceschi, F., and Sagi, I. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5007 – 5012

288c. Story, R. M., Li, H., and Abelson, J. N. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1465 – 1470

289. Wu, X., Wilcox, C. B., Devasahayam, G., Hackett, R. L., Arévalo-Rodríguez, M., Cardenas, M. E., Heitman, J., and Hanes, S. D. (2000) *EMBO J.* **19**, 3727 – 3738

289a. Conaway, R. C., Brower, C. S., and Conaway, J. W. (2002) *Science* **296**, 1254 – 1258

289b. Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W., and Pavietich, N. P. (2002) *Nature (London)* **416**, 703 – 709

289c. Ottosen, S., Herrera, F. J., and Triezenberg, S. J. (2002) *Science* **296**, 479 – 481

289d. Gonzalez, F., Delahodde, A., Kodadek, T., and Johnston, S. A. (2002) *Science* **296**, 548 – 550

289e. Tsuge, T., Matsui, M., and Wei, N. (2001) *J. Mol. Biol.* **305**, 1 – 9

289f. Goodson, M. L., Hong, Y., Rogers, R., Matunis, M. J., Park-Sarge, O.-K., and Sarge, K. D. (2001) *J. Biol. Chem.* **276**, 18513 – 18518

289g. Sampson, D. A., Wang, M., and Matunis, M. J. (2001) *J. Biol. Chem.* **276**, 21664 – 21669

289h. Kito, K., Yeh, E. T. H., and Kamitani, T. (2001) *J. Biol. Chem.* **276**, 20603 – 20609

289i. Marx, J. (2001) *Science* **292**, 838 – 839

289j. Buchberger, A., Howard, M. J., Proctor, M., and Bycroft, M. (2001) *J. Mol. Biol.* **307**, 17 – 24

290. Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R., and Kornberg, R. D. (2000) *Science* **288**, 640 – 649

291. Kim, W.-Y., and Dahmus, M. E. (1989) *J. Biol. Chem.* **264**, 3169 – 3176

292. Geiduschek, E. P., and Tocchini-Valentini, G. P. (1988) *Ann. Rev. Biochem.* **57**, 873 – 914

293. Shaw, P. J., Highett, M. I., Beven, A. F., and Jordan, E. G. (1995) *EMBO J.* **14**, 2896 – 2906

294. Seither, P., Iben, S., and Grummt, I. (1998) *J. Mol. Biol.* **275**, 43 – 53

295. Kelly, J. L., and Lehman, I. R. (1986) *J. Biol. Chem.* **261**, 10340 – 10347

295a. Masters, B. S., Stohl, L. L., and Clayton, D. A. (1987) *Cell* **51**, 89 – 99

296. Hanley–Bowdoin, L., and Chua, N.-H. (1987) *Trends Biochem. Sci.* **12**, 67 – 70

297. Myer, V. E., and Young, R. A. (1998) *J. Biol. Chem.* **273**, 27757 – 27760

298. Fu, J., Gerstein, M., David, P. R., Gnatt, A. L., Bushnell, D. A., Edwards, A. M., and Kornberg, R. D. (1998) *J. Mol. Biol.* **280**, 317 – 322

299. Jensen, G. J., Meredith, G., Bushnell, D. A., and Kornberg, R. D. (1998) *EMBO J.* **17**, 2353 – 2358

299a. Todone, F., Weinzierl, R. O. J., Brick, P., and Onesti, S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6306 – 6310

299b. Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) *Science* **292**, 1863 – 1876

299c. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001) *Science* **292**, 1876 – 1882

300. Simanek, E. E., Huang, D.-H., Pasternack, L., Machajewski, T. D., Seitz, O., Millar, D. S., Dyson, H. J., and Wong, C.-H. (1998) *J. Am. Chem. Soc.* **120**, 11567 – 11575

301. Adams, R. L. P., Knowler, J. T., and Leader, D. P. (1992) *The Biochemistry of the Nucleic Acids*, 11th ed., Chapman & Hall, London

302. Dahmus, M. E. (1996) *J. Biol. Chem.* **271**, 19009 – 19012

303. Corden, J. L., and Patturajan, M. (1997) *Trends Biochem. Sci.* **22**, 413 – 416

304. Proudfoot, N. (2000) *Trends Biochem. Sci.* **25**, 290 – 293

304a. Geiduschek, E. P., and Kassavetis, G. A. (2001) *J. Mol. Biol.* **310**, 1 – 26

304b. Kassavetis, G. A., Letts, G. A., and Geiduschek, E. P. (2001) *EMBO J.* **20**, 2823 – 2834

References

305. Pombo, A., Jackson, D. A., Hollinshead, M., Wang, Z., Roeder, R. G., and Cook, P. R. (1999) *EMBO J.* **18**, 2241–2253
306. Carter, K. C., Bowman, D., Carrington, W., Fogarty, K., McNeil, J. A., Fay, F. S., and Lawrence, J. B. (1993) *Science* **259**, 1330–1332
307. Nevins, J. R. (1983) *Ann. Rev. Biochem.* **52**, 441–466
308. Leff, S. E., Rosenfeld, M. G., and Evans, R. M. (1986) *Ann. Rev. Biochem.* **55**, 1091–1117
309. Ayoubi, T. A. Y., and Van de Ven, W. J. M. (1996) *FASEB J.* **10**, 453–460
310. Dantonel, J.-C., Wurtz, J.-M., Poch, O., Moras, D., and Tora, L. (1999) *Trends Biochem. Sci.* **24**, 335–339
311. Munoz-Sanjuan, I., Smallwood, P. M., and Nathans, J. (2000) *J. Biol. Chem.* **275**, 2589–2597
312. Ince, T. A., and Scotto, K. W. (1995) *J. Biol. Chem.* **270**, 30249–30252
312a. Conaway, J. W., Shilatifard, A., Dvir, A., and Conaway, R. C. (2000) *Trends Biochem. Sci.* **25**, 375–380
312b. Maniatis, T., and Reed, R. (2002) *Nature (London)* **416**, 499–506
312c. Sousa, R. (2001) *Trends Biochem. Sci.* **26**, 695–697
313. Arndt, K. T., Styles, C., and Fink, G. R. (1987) *Science* **237**, 874–880
314. Moye, W. S., and Zalkin, H. (1987) *J. Biol. Chem.* **262**, 3609–3624
315. Shuey, D. J., and Parker, C. S. (1986) *Nature (London)* **323**, 459–461
316. Prestridge, D. S. (1995) *J. Mol. Biol.* **249**, 923–932
317. Scherf, M., Klingenhoff, A., and Werner, T. (2000) *J. Mol. Biol.* **297**, 599–606
318. Kaiser, K., and Meisterernst, M. (1996) *Trends Biochem. Sci.* **21**, 342–345
318a. Tsuji, Y., Torti, S. V., and Torti, F. M. (1998) *J. Biol. Chem.* **273**, 2984–2992
319. Kadonaga, J. T., Jones, K. A., and Tjian, R. (1986) *Trends Biochem. Sci.* **11**, 20–23
320. Kadonaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. (1987) *Cell* **51**, 1079–1090
321. Garcia, A., Cereghini, S., and Sontag, E. (2000) *J. Biol. Chem.* **275**, 9385–9389
322. Li, B., Adams, C. C., and Workman, J. L. (1994) *J. Biol. Chem.* **269**, 7756–7763
323. Fry, C. J., and Farnham, P. J. (1999) *J. Biol. Chem.* **274**, 29583–29586
324. Daniel, S., Zhang, S., DePaoli-Roach, A. A., and Kim, K.-H. (1996) *J. Biol. Chem.* **271**, 14692–14697
324a. Johnson-Pais, T., Degnin, C., and Thayer, M. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2211–2216
324b. Goodwin, A. J., McInerney, J. M., Glander, M. A., Pomerantz, O., and Lowrey, C. H. (2001) *J. Biol. Chem.* **276**, 26883–26892
325. Van Dyke, M. W., Roeder, R. G., and Sawadogo, M. (1988) *Science* **241**, 1335–1338
326. Roeder, R. G. (1996) *Trends Biochem. Sci.* **21**, 327–335
327. Nikolov, D. B., Chen, H., Halay, E. D., Hoffmann, A., Roeder, R. G., and Burley, S. K. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4862–4867
328. Nikolov, D. B., and Burley, S. K. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 15–22
328a. Sachs, A. B., and Buratowski, S. (1997) *Trends Biochem. Sci.* **22**, 189–192
329. Andel, F., III, Ladurner, A. G., Inouye, C., Tjian, R., and Nogales, E. (1999) *Science* **286**, 2153–2156
329a. Gangloff, Y.-G., Romier, C., Thuault, S., Werten, S., and Davidson, I. (2001) *Trends Biochem. Sci.* **26**, 250–257
330. Burley, S. K., and Roeder, R. G. (1996) *Ann. Rev. Biochem.* **65**, 769–799
331. Green, M. R. (2000) *Trends Biochem. Sci.* **25**, 59–63
332. Verrijzer, C. P., and Tjian, R. (1996) *Trends Biochem. Sci.* **21**, 338–342
332a. Metzger, D., Scheer, E., Soldatov, A., and Tora, L. (1999) *EMBO J.* **18**, 4823–4834
333. Juo, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J., and Dickerson, R. E. (1996) *J. Mol. Biol.* **261**, 239–254
334. Parkhurst, K. M., Richards, R. M., Brenowitz, M., and Parkhurst, L. J. (1999) *J. Mol. Biol.* **289**, 1327–1341
335. Svejstrup, J. Q., Vichi, P., and Egly, J.-M. (1996) *Trends Biochem. Sci.* **21**, 346–350
335a. Holstege, F. C. P., and Young, R. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2–4
336. Yan, M., and Gralla, J. D. (1997) *EMBO J.* **16**, 7457–7467
336a. Mizzen, C. A., and Allis, C. D. (2000) *Science* **289**, 2290–2291
337. Nikolov, D. B., Hu, S.-H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N.-H., Roeder, R. G., and Burley, S. K. (1992) *Nature (London)* **360**, 40–46
338. Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G., and Burley, S. K. (1995) *Nature (London)* **377**, 119–128
338a. Brivanlou, A. H., and Darnell, J. E., Jr. (2002) *Science* **295**, 813–818
338b. Smith, R. L., and Johnson, A. D. (2000) *Trends Biochem. Sci.* **25**, 325–330
338c. Dunah, A. W., Jeong, H., Griffin, A., Kim, Y.-M., Standaert, D. G., Hersch, S. M., Mouradian, M. M., Young, A. B., Tanese, N., and Krainc, D. (2002) *Science* **296**, 2238–2243
339. Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S. C. (1992) *Nature (London)* **356**, 408–414
340. Ptashne, M., and Gann, A. (1997) *Nature (London)* **386**, 569–577
341. Rodgers, K. K., and Coleman, J. E. (1994) *Protein Sci.* **3**, 608–619
342. Lohr, D., Venkov, P., and Zlatanova, J. (1995) *FASEB J.* **9**, 777–787
343. Corton, J. C., Moreno, E., and Johnston, S. A. (1998) *J. Biol. Chem.* **273**, 13776–13780
344. Ruden, D. M., Ma, J., and Ptashne, M. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4262–4266
345. Sigler, P. B. (1988) *Nature (London)* **333**, 210–212
346. Ryu, S., and Tjian, R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7137–7142
347. Zawel, L., and Reinberg, D. (1995) *Ann. Rev. Biochem.* **64**, 533–561
348. Malik, S., and Roeder, R. G. (2000) *Trends Biochem. Sci.* **25**, 277–283
349. Svejstrup, J. Q., Li, Y., Fellows, J., Gnatt, A., Bjorklund, S., and Kornberg, R. D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6075–6078
350. Spåhr, H., Bève, J., Larsson, T., Bergström, J., Karlsson, K.-A., and Gustafsson, C. M. (2000) *J. Biol. Chem.* **275**, 1351–1356
350a. Dotson, M. R., Yuan, C. X., Roeder, R. G., Myers, L. C., Gustafsson, C. M., Jiang, Y. W., Li, Y., Kornberg, R. D., and Asturias, F. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14307–14310
350b. Liu, Y., Ranish, J. A., Aebersold, R., and Hahn, S. (2001) *J. Biol. Chem.* **276**, 7169–7175
351. Tan, R. C., Truong, T. N., McCammon, J. A., and Sussman, J. L. (1993) *Biochemistry* **32**, 401–403
352. Clarke, M. F., FitzGerald, P. C., Brubaker, J. M., and Simpson, R. T. (1985) *J. Biol. Chem.* **260**, 12394–12397
353. Ondek, B., Gloss, L., and Herr, W. (1988) *Nature (London)* **333**, 40–45
354. Plon, S., and Wang, J. (1986) *Cell* **45**, 575–580
355. Ptashne, M. (1986) *Nature (London)* **322**, 697–701
356. Ross, E. D., Keating, A. M., and Maher, L. J., III. (2000) *J. Mol. Biol.* **297**, 321–334
357. Udvardy, A. (1999) *EMBO J.* **18**, 1–8
357a. Struhl, K. (2001) *Science* **293**, 1054–1055
357b. Yie, J., Senger, K., and Thanos, D. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13108–13113
358. Werner, M. H., Gronenborn, A. M., and Clore, G. M. (1996) *Science* **271**, 778–784
359. Cronk, J. D., Endrizzi, J. A., and Alber, T. (1996) *Protein Sci.* **5**, 1963–1972
360. Keller, W., König, P., and Richmond, T. J. (1995) *J. Mol. Biol.* **254**, 657–667
361. Hanson, R. W. (1998) *J. Biol. Chem.* **273**, 28543
362. Gardner, K., Moore, T. C., Davis-Smyth, T., Krutzsch, H., and Levens, D. (1994) *J. Biol. Chem.* **269**, 32963–32971
363. Lekstrom-Himes, J., and Xanthopoulos, K. G. (1998) *J. Biol. Chem.* **273**, 28545–28548
364. Turner, R., and Tjian, R. (1989) *Science* **243**, 1689–1694
365. Okahata, Y., Niikura, K., Sugiura, Y., Sawada, M., and Morii, T. (1998) *Biochemistry* **37**, 5666–5672
366. Maity, S. N., and de Crombrugghe, B. (1998) *Trends Biochem. Sci.* **23**, 174–178
367. Croniger, C., Leahy, P., Reshef, L., and Hanson, R. W. (1998) *J. Biol. Chem.* **273**, 31629–31632
368. Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988) *Science* **240**, 1759–1764
369. Ferré-D'Amaré, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) *Nature (London)* **363**, 38–44
370. Chen, L., Glover, J. N. M., Hogan, P. G., Rao, A., and Harrison, S. C. (1998) *Nature (London)* **392**, 42–48
371. Mackay, J. P., Shaw, G. L., and King, G. F. (1996) *Biochemistry* **35**, 4867–4877
372. Gentz, R., Rauscher, F. J., III, Abate, C., and Curran, T. (1989) *Science* **243**, 1695–1699
373. Dang, C. V., McGuire, M., Buckmire, M., and Lee, W. M. F. (1989) *Nature (London)* **337**, 664–666
374. Garlatti, M., Tchesnokov, V., Daheshia, M., Feilleux-Duché, S., Hanoune, J., Aggerbeck, M., and Barouki, R. (1993) *J. Biol. Chem.* **268**, 6567–6574
375. Umek, R. M., Friedman, A. D., and McKnight, S. L. (1991) *Science* **251**, 288–292
376. Liberati, C., di Silvio, A., Ottolenghi, S., and Mantovani, R. (1999) *J. Mol. Biol.* **285**, 1441–1455
377. Grandori, C., and Eisenman, R. N. (1997) *Trends Biochem. Sci.* **22**, 177–181
378. Noguchi, K., Kitanaka, C., Yamana, H., Kokubu, A., Mochizuki, T., and Kuchino, Y. (1999) *J. Biol. Chem.* **274**, 32580–32587
379. Grigorieva, I., Grigoriev, V. G., Rowney, M. K., and Hoover, R. G. (2000) *J. Biol. Chem.* **275**, 7343–7350
379a. Otsuki, Y., Tanaka, M., Kamo, T., Kitanaka, C., Kuchino, Y. and Sugimura, H. (2002) *J. Biol. Chem.* In press 206733200
380. Lalli, E., and Sassone-Corsi, P. (1994) *J. Biol. Chem.* **269**, 17359–17362
381. Shaywitz, A. J., and Greenberg, M. E. (1999) *Ann. Rev. Biochem.* **68**, 821–861
382. De Cesare, D., Fimia, G. M., and Sassone-Corsi, P. (1999) *Trends Biochem. Sci.* **24**, 281–284
382a. Goren, I., Tavor, E., Goldblum, A., and Honigman, A. (2001) *J. Mol. Biol.* **313**, 695–709
382b. Nagadoi, A., Nakazawa, K.-i, Uda, H., Okuno, K., Maekawa, T., Ishii, S., and Nishimura, Y. (1999) *J. Mol. Biol.* **287**, 593–607
382c. Yeagley, D., Agati, J. M., and Quinn, P. G. (1998) *J. Biol. Chem.* **273**, 18743–18750

References

382d. Vo, N., and Goodman, R. H. (2001) *J. Biol. Chem.* **276**, 13505 – 13508

383. Xu, W., Chen, H., Du, K., Asahara, H., Tini, M., Emerson, B. M., Montminy, M., and Evans, R. M. (2001) *Science* **294**, 2507 – 2511

383a. Martínez-Balbás, M. A., Bauer, U.-M., Nielsen, S., Brehm, A., and Kouzarides, T. (2000) *EMBO J.* **19**, 662 – 671

383b. Classon, M., Salama, S., Gorka, C., Mulloy, R., Braun, P., and Harlow, E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10820 – 10825

383c. Ogawa, H., Ishiguro, K.-i, Gaubatz, S., Livingston, D. M., and Nakatani, Y. (2002) *Science* **296**, 1132 – 1136

383d. La Thangue, N. B. (2002) *Science* **296**, 1034 – 1035

383e. Chin, L., Pomerantz, J., and DePinho, R. A. (1998) *Trends Biochem. Sci.* **23**, 291 – 296

383f. Chang, L., and Karin, M. (2001) *Nature (London)* **410**, 37 – 40

383g. Weston, C. R., Lambright, D. G., and Davis, R. J. (2002) *Science* **296**, 2345, 2347

383h. Cantley, L. C. (2001) *Science* **292**, 2019 – 2021

383i. Janknecht, R., and Hunter, T. (1999) *Science* **284**, 443 – 444

383j. Horvath, C. M. (2000) *Trends Biochem. Sci.* **25**, 496 – 502

383k. ten Dijke, P., Miyazono, K., and Helden, C.-H. (2000) *Trends Biochem. Sci.* **25**, 64 – 70

383l. Wu, G., Chen, Y.-G., Ozdamar, B., Gyuricza, C. A., Chong, P. A., Wrana, J. L., Massagué, J., and Shi, Y. (2000) *Science* **287**, 92 – 97

384. Towle, H. C. (1995) *J. Biol. Chem.* **270**, 23235 – 23238

384a. Osborne, T. F. (2000) *J. Biol. Chem.* **275**, 32379 – 32382

384b. Duan, R., Xie, W., Burghardt, R. C., and Safe, S. (2001) *J. Biol. Chem.* **276**, 11590 – 11598

384c. Hassler, M., and Richmond, T. J. (2001) *EMBO J.* **20**, 3018 – 3028

385. Treisman, R. (1992) *Trends Biochem. Sci.* **17**, 423 – 426

386. Ling, Y., West, A. G., Roberts, E. C., Lakey, J. H., and Sharrocks, A. D. (1998) *J. Biol. Chem.* **273**, 10506 – 10514

387. Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C. F., Trent, J. M., Staudt, L. M., Hudson, J., Jr., Boguski, M. S., Lashkari, D., Shalon, D., Botstein, D., and Brown, P. O. (1999) *Science* **283**, 83 – 87

388. Kerrigan, L. A., Croston, G. E., Lira, L. M., and Kadonaga, J. T. (1991) *J. Biol. Chem.* **266**, 574 – 582

388a. Bieker, J. J. (2001) *J. Biol. Chem.* **276**, 34355 – 34358

389. Mackay, J. P., and Crossley, M. (1998) *Trends Biochem. Sci.* **23**, 1 – 4

390. Turner, J., and Crossley, M. (1999) *Trends Biochem. Sci.* **24**, 236 – 241

391. Clarke, N. D., and Berg, J. M. (1998) *Science* **282**, 2018 – 2022

392. Naltner, A., Ghaffari, M., Whitsett, J. A., and Yan, C. (2000) *J. Biol. Chem.* **275**, 56 – 62

393. Pizzi, S., Dieci, G., Frigeri, P., Piccoli, G., Stocchi, V., and Ottonello, S. (1999) *J. Biol. Chem.* **274**, 2539 – 2548

394. Laity, J. H., Dyson, H. J., and Wright, P. E. (2000) *J. Mol. Biol.* **295**, 719 – 727

395. Klug, A., and Rhodes, D. (1987) *Trends Biochem. Sci.* **12**, 464 – 469

396. Nolte, R. T., Conlin, R. M., Harrison, S. C., and Brown, R. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2938 – 2943

397. Muro-Pastor, M. I., Gonzalez, R., Strauss, J., Narendja, F., and Scazzocchio, C. (1999) *EMBO J.* **18**, 1584 – 1597

397a. Molkentin, J. D. (2000) *J. Biol. Chem.* **275**, 38949 – 38952

398. Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., and Gronenborn, A. M. (1993) *Science* **261**, 438 – 446

399. Haas, H., Zadra, I., Stöffler, G., and Angermayr, K. (1999) *J. Biol. Chem.* **274**, 4613 – 4619

400. Van Winkle, L. J. (1999) *Biomembrane Transport*, Academic Press, San Diego, California

401. Zeng, J., Vallee, B. L., and Kägi, J. H. R. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9984 – 9988

402. Lawson, C. L., and Carey, J. (1993) *Nature (London)* **366**, 178 – 182

403. Marsden, I., Jin, C., and Liao, X. (1998) *J. Mol. Biol.* **278**, 293 – 299

404. Clark, K. L., Halay, E. D., Lai, E., and Burley, S. K. (1993) *Nature (London)* **364**, 412 – 420

405. Pio, F., Kodandapani, R., Ni, C.-Z., Shepard, W., Klemsz, M., McKercher, S. R., Maki, R. A., and Ely, K. R. (1996) *J. Biol. Chem.* **271**, 23329 – 23337

406. Wasylyk, B., Hagman, J., and Gutierrez-Hartmann, A. (1998) *Trends Biochem. Sci.* **23**, 213 – 216

406a. Alvarez, B., Martínez-A, C., Burgering, B. M. T., and Carrera, A. C. (2001) *Nature (London)* **413**, 744 – 747

407. Tisné, C., Delepierre, M., and Hartmann, B. (1999) *J. Mol. Biol.* **293**, 139 – 150

408. Barnes, P. J., and Karin, M. (1997) *N. Engl. J. Med.* **336**, 1066 – 1071

408a. Nourbakhsh, M., Kälble, S., Dörrie, A., Hauser, H., Resch, K., and Kracht, M. (2001) *J. Biol. Chem.* **276**, 4501 – 4508

408b. Tam, W. F., and Sen, R. (2001) *J. Biol. Chem.* **276**, 7701 – 7704

408c. Schmitz, M. L., Bacher, S., and Kracht, M. (2001) *Trends Biochem. Sci.* **26**, 186 – 190

409. Malek, S., Huxford, T., and Ghosh, G. (1998) *J. Biol. Chem.* **273**, 25427 – 25435

410. Chen, F. E., Huang, D.-B., Chen, Y.-Q., and Ghosh, G. (1998) *Nature (London)* **391**, 410 – 413

411. Shlyakhtenko, L. S., Hsieh, P., Grigoriev, M., Potaman, V. N., Sinden, R. R., and Lyubchenko, Y. L. (2000) *J. Mol. Biol.* **296**, 1169 – 1173

412. Murphy, F. V., IV, Sweet, R. M., and Churchill, M. E. A. (1999) *EMBO J.* **18**, 6610 – 6618

413. Sutrias-Grau, M., Bianchi, M. E., and Bernusés, J. (1999) *J. Biol. Chem.* **274**, 1628 – 1634

414. Yie, J., Merika, M., Munshi, N., Chen, G., and Thanos, D. (1999) *EMBO J.* **18**, 3074 – 3089

414a. Lewis, R. T., Andreucci, A., and Nikolajczyk, B. S. (2001) *J. Biol. Chem.* **276**, 9550 – 9557

415. Bergel, M., Herrera, J. E., Thatcher, B. J., Prymakowska-Bosak, M., Vassilev, A., Nakatani, Y., Martin, B., and Bustin, M. (2000) *J. Biol. Chem.* **275**, 11514 – 11520

415a. Cowell, I. G. (1994) *Trends Biochem. Sci.* **19**, 38 – 42

415b. Harrington, R. H., and Sharma, A. (2001) *J. Biol. Chem.* **276**, 104 – 113

416. Zhang, J., Ou, J., Bashmakov, Y., Horton, J. D., Brown, M. S., and Goldstein, J. L., (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3756 – 3761

416a. Olefsky, J. M. (2001) *J. Biol. Chem.* **276**, 36863 – 36864

416b. Khorasanizadeh, S., and Rastinejad, F. (2001) *Trends Biochem. Sci.* **26**, 384 – 390

416c. Valverde, M. A., and Parker, M. G. (2002) *Trends Biochem. Sci.* **27**, 172 – 173

417. Vázquez, F., Rodríguez-Manzaneque, J. C., Lydon, J. P., Edwards, D. P., O'Mally, B. W., and Iruela-Arispe, M. L. (1999) *J. Biol. Chem.* **274**, 2185 – 2192

418. Lehmann, M., and Korge, G. (1995) *EMBO J.* **14**, 716 – 726

418a. Wurtz, J.-M., Guillot, B., Fagart, J., Moras, D., Tietjen, K., and Schindler, M. (2000) *Protein Sci.* **9**, 1073 – 1084

419. Guschin, D., Chandler, S., and Wolffe, A. P. (1998) *Biochemistry* **37**, 8629 – 8636

420. Baudino, T. A., Kraichely, D. M., Jefcoat, S. C., Jr., Winchester, S. K., Partridge, N. C., and MacDonald, P. N. (1998) *J. Biol. Chem.* **273**, 16434 – 16441

420a. Makishima, M., Lu, T. T., Xie, W., Whitfield, G. K., Domoto, H., Evans, R. M., Haussler, M. R., and Mangelsdorf, D. J. (2002) *Science* **296**, 1313 – 1316

421. Beato, M. (1991) *FASEB J.* **5**, 2044 – 2051

422. Weatherman, R. V., Fletterick, R. J., and Scanlan, T. S. (1999) *Ann. Rev.Biochem.* **68**, 559 – 581

423. Di Croce, L., Okret, S., Kersten, S., Gustafsson, J.-Å., Parker, M., Wahli, W., and Beato, M. (1999) *EMBO J.* **18**, 6201 – 6210

424. Avram, D., Fields, A., Pretty On Top, K., Nevrivy, D. J., Ishmael, J. E., and Leid, M. (2000) *J. Biol. Chem.* **275**, 10315 – 10322

425. Kliewer, S. A., Lehmann, J. M., and Willson, T. M. (1999) *Science* **284**, 757 – 760

426. Kosztin, D., Bishop, T. C., and Schulten, K. (1997) *Biophys. J.* **73**, 557 – 570

427. Driscoll, M. D., Sathya, G., Muyan, M., Klinge, C. M., Hilf, R., and Bambara, R. A. (1998) *J. Biol. Chem.* **273**, 29321 – 29330

427a. Campbell, R. A., Bhat-Nakshatri, P., Patel, N. M., Constantinidou, D., Ali, S., and Nakshatri, H. (2001) *J. Biol. Chem.* **276**, 9817 – 9824

427b. An, J., Tzagarakis-Foster, C., Scharaschmidt, T. C., Lomri, N., and Leitman, D. C. (2001) *J. Biol. Chem.* **276**, 17808 – 17814

428. Rogatsky, I., Trowbridge, J. M., and Garabedian, M. J. (1999) *J. Biol. Chem.* **274**, 22296 – 22302

429. Kosano, H., Stensgard, B., Charlesworth, M. C., McMahon, N., and Toft, D. (1998) *J. Biol. Chem.* **273**, 32973 – 32979

429a. Bennett, M. K., and Osborne, T. F. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6340 – 6344

429b. Luong, A., Hannah, V. C., Brown, M. S., and Goldstein, J. L. (2000) *J. Biol. Chem.* **275**, 26458 – 26466

430. Clarke, S. D., and Abraham, S. (1992) *FASEB J.* **6**, 3146 – 3152

431. Lu, C.-A., Lim, E.-K., and Yu, S.-M. (1998) *J. Biol. Chem.* **273**, 10120 – 10131

432. Cereghino, G. P., and Scheffler, I. E. (1996) *EMBO J.* **15**, 363 – 374

433. Hong, S.-P., Piper, M. D., Sinclair, D. A., and Dawes, I. W. (1999) *J. Biol. Chem.* **274**, 10523 – 10532

434. Svaren, J., and Hörz, W. (1997) *Trends Biochem. Sci.* **22**, 93 – 97

435. Shimizu, T., Toumoto, A., Ihara, K., Shimizu, M., Kyogoku, Y., Ogawa, N., Oshima, Y., and Hakoshima, T. (1997) *EMBO J.* **16**, 4689 – 4697

436. Addess, K. J., Basilion, J. P., Klausner, R. D., Rouault, T. A., and Pardi, A. (1997) *J. Mol. Biol.* **274**, 72 – 83

437. Wardrop, S. L., Watts, R. N., and Richardson, D. R. (2000) *Biochemistry* **39**, 2748 – 2758

438. Ke, Y., Sierzputowska-Gracz, H., Gdaniec, Z., and Theil, E. C. (2000) *Biochemistry* **39**, 6235 – 6242

438a. Meehan, H. A., and Connell, G. J. (2001) *J. Biol. Chem.* **276**, 14791 – 14796

439. Quinn, J. M., Barraco, P., Eriksson, M., and Merchant, S. (2000) *J. Biol. Chem.* **275**, 6080 – 6089

440. Hall, D. R., Gourley, D. G., Leonard, G. A., Duke, E. M. H., Anderson, L. A., Boxer, D. H., and Hunter, W. N. (1999) *EMBO J.* **18**, 1435 – 1446

441. Okamura, H., Hanaoka, S., Nagadoi, A., Makino, K., and Nishimura, Y. (2000) *J. Mol. Biol.* **295**, 1225 – 1236

442. Ritossa, F. (1962) *Experientia* **18**, 571 – 573

References

443. Schlesinger, M. J. (1990) *J. Biol. Chem.* **265**, 12111 – 12114
444. Ang, D., Liberek, K., Skowyra, D., Zylicz, M., and Georgopoulos, C. (1991) *J. Biol. Chem.* **266**, 24233 – 24236
445. Morimoto, R. I. (1993) *Science* **259**, 1409 – 1410
446. Giardina, C., and Lis, J. T. (1995) *J. Biol. Chem.* **270**, 10369 – 10372
447. Kimpel, J. A., and Key, J. L. (1985) *Trends Biochem. Sci.* **10**, 353 – 357
448. Craig, E. A., and Gross, C. A. (1991) *Trends Biochem. Sci.* **16**, 135 – 140
449. Welch, W. J. (1993) *Sci. Am.* **268**(May), 57 – 64
450. Kim, K. K., Kim, R., and Kim, S.-H. (1998) *Nature (London)* **394**, 595 – 599
451. Korber, P., Stahl, J. M., Nierhaus, K. H., and Bardwell, J. C. A. (2000) *EMBO J.* **19**, 741 – 748
452. Goldenberg, C. J., Luo, Y., Fenna, M., Baler, R., Weinmann, R., and Voellmy, R. (1988) *J. Biol. Chem.* **263**, 19734 – 19739
453. Harrison, C. J., Bohm, A. A., and Nelson, H. C. M. (1994) *Science* **263**, 224 – 227
454. Larson, J. S., Schuetz, T. J., and Kingston, R. E. (1995) *Biochemistry* **34**, 1902 – 1911
455. Wyman, C., Grotkopp, E., Bustamante, C., and Nelson, H. C. M. (1995) *EMBO J.* **14**, 117 – 123
456. Rouvière, P. E., De Las Penas, A., Mecsas, J., Lu, C. Z., Rudd, K. E., and Gross, C. A. (1995) *EMBO J.* **14**, 1032 – 1042
456a. Dartigalongue, C., Missiakas, D., and Raina, S. (2001) *J. Biol. Chem.* **276**, 20866 – 20875
457. Brevet, A., Plateau, P., Best-Belpomme, M., and Blanquet, S. (1985) *J. Biol. Chem.* **260**, 15566 – 15570
458. Wahab, S. Z., and Yang, D. C. H. (1985) *J. Biol. Chem.* **260**, 5286 – 5289
459. Harnett, S. P., Lowe, G., and Tansley, G. (1985) *Biochemistry* **24**, 2908 – 2915
460. Lee, P. C., Bochner, B. R., and Ames, B. N. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7496 – 7500
461. Guédon, G. F., Gilson, G. J. P., Ebel, J. P., Befort, N. M.-T., and Remy, P. M. (1986) *J. Biol. Chem.* **261**, 16459 – 16465
462. Kodym, R., Calkins, P., and Story, M. (1999) *J. Biol. Chem.* **274**, 5131 – 5137
463. Freestone, P., Nyström, T., Trinei, M., and Norris, V. (1997) *J. Mol. Biol.* **274**, 318 – 324
464. Discher, D. J., Bishopric, N. H., Wu, X., Peterson, C. A., and Webster, K. A. (1998) *J. Biol. Chem.* **273**, 26087 – 26093
465. Sundaresan, M., Yu, Z.-X., Ferrans, V. J., Irani, K., and Finkel, T. (1995) *Science* **270**, 296 – 299
465a. Thannickal, V. J., and Fanburg, B. L. (1995) *J. Biol. Chem.* **270**, 30334 – 30338
466. Burg, M. B., Kwon, E. D., and Kültz, D. (1996) *FASEB J.* **10**, 1598 – 1606
467. Ramagopal, S. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 94 – 98
467a. Dagsgaard, C., Taylor, L. E., O'Brien, K. M., and Poyton, R. O. (2001) *J. Biol. Chem.* **276**, 7593 – 7601
467b. Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J. M., Lane, W. S., and Kaelin, W. G., Jr. (2001) *Science* **292**, 464 – 468
467c. Zhu, H., and Bunn, H. F. (2001) *Science* **292**, 449,451
467d. Lando, D., Peet, D. J., Whelan, D. A., Gorman, J. J., and Whitelaw, M. L. (2002) *Science* **295**, 858 – 861
467e. Srinivas, V., Leshchinsky, I., Sang, N., King, M. P., Minchenko, A., and Caro, J. (2001) *J. Biol. Chem.* **276**, 21995 – 21998
467f. Jaakkola, P., Mole, D. R., Tian, Y.-M., Wilson, M. I., Gielbert, J., Gaskell, S. J., von Kriegsheim, A., Hebestreit, H. F., Mukherji, M., Schofield, C. J., Maxwell, P. H., Pugh, C. W., and Ratcliffe, P. J. (2001) *Science* **292**, 468 – 472
467g. Min, J.-H., Yang, H., Ivan, M., Gertler, F., Kaelin, W. G., Jr., and Pavletich, N. P. (2002) *Science* **296**, 1886 – 1889
467h. Semenza, G. L. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11570 – 11572
467i. Hansson, L. O., Friedler, A., Freund, S., Rüdiger, S., and Fersht, A. R. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10305 – 10309
467j. Stebbins, C. E., Kaelin, W. G., Jr., and Pavletich, N. P. (1999) *Science* **284**, 455 – 461
467k. Baysal, B. E., Ferrell, R. E., Willett-Brozick, J. E., Lawrence, E. C., Myssiorek, D., Bosch, A., van der Mey, A., Taschner, P. E. M., Rubinstein, W. S., Myers, E. N., Richard, C. W., III, Cornelisse, C. J., Devilee, P., and Devlin, B. (2000) *Science* **287**, 848 – 851
468. Kim, M., Thum, K. E., Morishige, D. T., and Mullet, J. E. (1999) *J. Biol. Chem.* **274**, 4684 – 4692
469. Martínez-García, J. F., Huq, E., and Quail, P. H. (2000) *Science* **288**, 859 – 863
470. Wright, C. V. E., Cho, K. W. Y., Oliver, G., and De Robertis, E. M. (1989) *Trends Biochem. Sci.* **14**, 52 – 56
471. Scott, M. P. (1987) *Ann. Rev. Biochem.* **56**, 195 – 227
472. Mihara, H., and Kaiser, E. T. (1988) *Science* **242**, 925 – 927
473. Ford, E., and Hernandez, N. (1997) *J. Biol. Chem.* **272**, 16048 – 16055
474. Gstaiger, M., Georgiev, O., van Leeuwen, H., van der Vliet, P., and Schaffner, W. (1996) *EMBO J.* **15**, 2781 – 2790
475. Lee, L., Stollar, E., Chang, J., Grossmann, J. G., O'Brien, R., Ladbury, J., Carpenter, B., Roberts, S. and Luisi, B. (2001) *Biochemistry* **40**, 6580 – 6588
475a. Lundbäck, T., Chang, J.-F., Phillips, K., Luisi, B., and Ladbury, J. E. (2000) *Biochemistry* **39**, 7570 – 7579
476. Robertson, M. (1988) *Nature (London)* **336**, 522 – 524
476a. Diamond, S. E., and Gutierrez-Hartmann, A. (2000) *J. Biol. Chem.* **275**, 30977 – 30986
477. Tanaka, N., Kato, H., Ishikawa, Y., Hisatake, K., Tashiro, K., Kominami, R., and Muramatsu, M. (1990) *J. Biol. Chem.* **265**, 13836 – 13842
478. Kahl, B. F., Li, H., and Paule, M. R. (2000) *J. Mol. Biol.* **299**, 75 – 89
478a. Miller, G., Panov, K. I., Friedrich, J. K., Trinkle-Mulcahy, L., Lamond, A. I., and Zomerdijk, J. C. B. M. (2001) *EMBO J.* **20**, 1373 – 1382
479. McBryant, S. J., Kassavetis, G. A., and Gottesfeld, J. M. (1995) *J. Mol. Biol.* **250**, 315 – 326
480. Gunnery, S., Ma, Y., and Mathews, M. B. (1999) *J. Mol. Biol.* **286**, 745 – 757
481. Pieler, T., and Theunissen, O. (1993) *Trends Biochem. Sci.* **18**, 226 – 230
482. Miller, J., McLachlan, A. D., and Klug, A. (1985) *EMBO J.* **4**, 1609 – 1614
483. Veldhoen, N., You, Q., Setzer, D. R., and Romaniuk, P. J. (1994) *Biochemistry* **33**, 7568 – 7575
484. Gabrielsen, O. S., and Sentenac, A. (1991) *Trends Biochem. Sci.* **16**, 412 – 416
485. Andrau, J.-C., Sentenac, A., and Werner, M. (1999) *J. Mol. Biol.* **288**, 511 – 520
486. Young, L. S., Dunstan, H. M., Witte, P. R., Smith, T. P., Ottonello, S., and Sprague, K. U. (1991) *Science* **252**, 542 – 546
487. Choder, M., and Aloni, Y. (1988) *J. Biol. Chem.* **263**, 12994 – 13002
488. Lewis, J. D., and Tollervey, D. (2000) *Science* **288**, 1385 – 1389
489. Reines, D., Conaway, J. W., and Conaway, R. C. (1996) *Trends Biochem. Sci.* **21**, 351 – 355
490. Gnatt, A., Fu, J., and Kornberg, R. D. (1997) *J. Biol. Chem.* **272**, 30799 – 30805
491. Izban, M. G., Parsons, M. A., and Sinden, R. R. (1998) *J. Biol. Chem.* **273**, 27009 – 27016
492. Hemming, S. A., and Edwards, A. M. (2000) *J. Biol. Chem.* **275**, 2288 – 2294
493. Nelson, P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14342 – 14347
494. Botuyan, M. V., Koth, C. M., Mer, G., Chakrabartty, A., Conaway, J. W., Conaway, R. C., Edwards, A. M., Arrowsmith, C. H., and Chazin, W. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9033 – 9038
495. Conaway, J. W., and Conaway, R. C. (1999) *Ann. Rev. Biochem.* **68**, 301 – 319
496. Iyer, V., and Struhl, K. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5208 – 5212
497. Guo, Z., and Sherman, F. (1996) *Trends Biochem. Sci.* **21**, 477 – 481
498. Wahle, E., and Keller, W. (1996) *Trends Biochem. Sci.* **21**, 247 – 250
499. Chanfreau, G., Noble, S. M., and Guthrie, C. (1996) *Science* **274**, 1511 – 1514
500. Hirose, Y., and Manley, J. L. (1997) *J. Biol. Chem.* **272**, 29636 – 29642
501. Ohnacker, M., Barabino, S. M. L., Preker, P. J., and Keller, W. (2000) *EMBO J.* **19**, 37 – 47
502. Keller, R. W., Kühn, U., Aragón, M., Bornikova, L., Wahle, E., and Bear, D. G. (2000) *J. Mol. Biol.* **297**, 569 – 583
503. Reeder, R. H., and Lang, W. H. (1997) *Trends Biochem. Sci.* **22**, 473 – 477
504. Syroid, D. E., and Capone, J. P. (1994) *J. Mol. Biol.* **244**, 482 – 493
504a. Ohndorf, U.-M., Steegborn, C., Knijff, R., and Sondermann, P. (2001) *J. Biol. Chem.* **276**, 27188 – 27196
504b. Wang, Z., Bai, L., Hsieh, Y.-J., and Roeder, R. G. (2000) *EMBO J.* **19**, 6823 – 6832
505. Draper, D. E. (1996) *Trends Biochem. Sci.* **21**, 145 – 149
506. Moore, P. B. (1999) *Ann. Rev. Biochem.* **67**, 287 – 300
507. Hermann, T., and Patel, D. J. (1999) *J. Mol. Biol.* **294**, 829 – 849
508. Pan, J., and Woodson, S. A. (1999) *J. Mol. Biol.* **294**, 955 – 965
508a. Zhuang, X., Bartley, L. E., Babcock, H. P., Russell, R., Ha, T., Herschlag, D., and Chu, S. (2000) *Science* **288**, 2048 – 2051
508b. Kent, O., Chaulk, S. G., and MacMillan, A. M. (2000) *J. Mol. Biol.* **304**, 699 – 705
509. Murthy, V. L., Srinivasan, R., Draper, D. E., and Rose, G. D. (1999) *J. Mol. Biol.* **291**, 313 – 327
510. Michel, F., and Westhof, E. (1996) *Science* **273**, 1676 – 1677
511. Gultyaev, A. P., van Batenburg, F. H. D., and Pleij, C. W. A. (1995) *J. Mol. Biol.* **250**, 37 – 51
512. Narlikar, G. J., Bartley, L. E., Khosla, M., and Herschlag, D. (1999) *Biochemistry* **38**, 14192 – 14204
513. Sclavi, B., Sullivan, M., Chance, M. R., Brenowitz, M., and Woodson, S. A. (1998) *Science* **279**, 1940 – 1943
514. Silverman, S. K., and Cech, T. R. (1999) *Biochemistry* **38**, 8691 – 8702
515. Rivas, E., and Eddy, S. R. (1999) *J. Mol. Biol.* **285**, 2053 – 2068
516. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science* **273**, 1696 – 1699
517. Biswas, R., Wahl, M. C., Ban, C., and Sundaralingam, M. (1997) *J. Mol. Biol.* **267**, 1149 – 1156
517a. Deng, J., Xiong, Y., and Sundaralingam, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13665 – 13670
517b. Snoussi, K., Nonin-Lecomte, S., and Leroy, J.-L. (2001) *J. Mol. Biol.* **309**, 139 – 153
517c. Schneider, C., Brandl, M., and Sühnel, J. (2001) *J. Mol. Biol.* **305**, 659 – 667

References

517d. Soukup, J. K., Minakawa, N., Matsuda, A., and Strobel, S. A. (2002) *Biochemistry* **41**, 10426 – 10438
517e. Stormo, G. D., and Ji, Y. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9465 – 9467
518. Costa, M., and Michel, F. (1997) *EMBO J.* **16**, 3289 – 3302
519. Nagai, K., Oubridge, C., Ito, N., Avis, J., and Evans, P. (1995) *Trends Biochem. Sci.* **20**, 235 – 240
520. Bell, S. P., Learned, R. M., Jantzen, H.–M., and Tjian, R. (1988) *Science* **241**, 1192 – 1197
521. Schmitz, M. L., Maier, U.-G., Brown, J. W. S., and Feix, G. (1989) *J. Biol. Chem.* **264**, 1467 – 1472
522. Baker, S. M., and Platt, T. (1986) *Cell* **47**, 839 – 840
523. Read, C., Larose, A.-M., Leblanc, B., Bannister, A. J., Firek, S., Smith, D. R., and Moss, T. (1992) *J. Biol. Chem.* **267**, 10961 – 10967
524. Sommerville, J. (1985) *Nature (London)* **318**, 410 – 411
525. Morrissey, J. P., and Tollervey, D. (1995) *Trends Biochem. Sci.* **20**, 78 – 82
526. Borovjagin, A. V., and Gerbi, S. A. (1999) *J. Mol. Biol.* **286**, 1347 – 1363
527. Wellauer, P. K., and Dawid, I. B. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2827 – 2831
528. Dawid, I. B., and Wellauer, P. K. (1976) *Cell* **8**, 443 – 448
529. Ghisolfi, L., Joseph, G., Amalric, F., and Erard, M. (1992) *J. Biol. Chem.* **267**, 2955 – 2959
529a. Bouvet, P., Allain, F. H.-T., Finger, L. D., Dieckmann, T., and Feigon, J. (2001) *J. Mol. Biol.* **309**, 763 – 775
529b. Allain, F. H.-T., Bouvet, P., Dieckmann, T., and Feigon, J. (2000) *EMBO J.* **19**, 6870 – 6881
530. Srivastava, M., McBride, O. W., Fleming, P. J., Pollard, H. B., and Burns, A. L. (1990) *J. Biol. Chem.* **265**, 14922 – 14931
531. Tycowaki, K. T., Shu, M.–D., and Steitz, J. A. (1994) *Science* **266**, 1558 – 1561
531a. Pederson, T. (2002) *Trends Biochem. Sci.* **27**, 111 – 112
532. Guthrie, C. (1986) *Trends Biochem. Sci.* **11**, 430 – 434
533. Birnstiel, M. L., ed. (1988) *Small Nuclear Ribonucleoprotein Particles*, Springer, Vienna/New York
534. Tarn, W.-Y., and Steitz, J. A. (1997) *Trends Biochem. Sci.* **22**, 132 – 137
535. Séraphin, B. (1995) *EMBO J.* **14**, 2089 – 2098
535a. Filipowicz, W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14035 – 14037
535b. Hüttenhofer, A., Kiefmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., Bachellerie, J.-P., and Brosius, J. (2001) *EMBO J.* **20**, 2943 – 2953
535c. Lowe, T. M., and Eddy, S. R. (1999) *Science* **283**, 1168 – 1171
535d. Xu, Y.-x, Liu, L., Lopez-Estrano, C., and Michaeli, S. (2001) *J. Biol. Chem.* **276**, 14289 – 14298
535e. Barneche, F., Gaspin, C., Guyot, R., and Echeverría, M. (2001) *J. Mol. Biol.* **311**, 57 – 73
536. Feeney, R. J., Sauterer, R. A., Feeney, J. L., and Zieve, G. W. (1989) *J. Biol. Chem.* **264**, 5776 – 5783
537. Hartmuth, K., Raker, V. A., Huber, J., Branlant, C., and Lührmann, R. (1999) *J. Mol. Biol.* **285**, 133 – 147
537a. Collins, B. M., Harrop, S. J., Kornfeld, G. D., Dawes, I. W., Curmi, P. M. G., and Mabbutt, B. C. (2001) *J. Mol. Biol.* **309**, 915 – 923
537b. Baulcombe, D. (2002) *Science* **297**, 2002 – 2003
537c. Hutvágner, G., and Zamore, P. D. (2002) *Science* **297**, 2056 – 2060
538. Parry, H. D., Scherly, D., and Mattaj, I. W. (1989) *Trends Biochem. Sci.* **14**, 15 – 19
539. Grabowski, P. J., and Sharp, P. A. (1986) *Science* **233**, 1294 – 1299
539a. Hannon, G. J. (2002) *Nature (London)* **418**, 244 – 251
540. Sollner-Webb, B., and Tower, J. (1986) *Ann. Rev. Biochem.* **55**, 801 – 830
541. Fournier, M. J., and Maxwell, E. S. (1993) *Trends Biochem. Sci.* **18**, 131 – 135
541a. Saveanu, C., Bienvenu, D., Namane, A., Gleizes, P.-E., Gas, N., Jacquier, A., and Fromont-Racine, M. (2001) *EMBO J.* **20**, 6475 – 6484
541b. Gadal, O., Strauss, D., Braspenning, J., Hoepfner, D., Petfalski, E., Philippsen, P., Tollervey, D., and Hurt, E. (2001) *EMBO J.* **20**, 3695 – 3704
541c. Jäkel, S., and Görlich, D. (1998) *EMBO J.* **17**, 4491 – 4502
542. Beltrame, M., and Tollervey, D. (1995) *EMBO J.* **14**, 4350 – 4356
543. Hughes, J. M. X. (1996) *J. Mol. Biol.* **259**, 645 – 654
544. Méreau, A., Fournier, R., Grégoire, A., Mougin, A., Fabrizio, P., Lührmann, R., and Branlant, C. (1997) *J. Mol. Biol.* **273**, 552 – 571
545. Hengst, J. A., Georgoff, I., Isom, H. C., and Jacob, S. T. (1988) *J. Biol. Chem.* **263**, 19270 – 19273
546. van Nues, R. W., Rientjes, J. M. J., Morré, S. A., Mollee, E., Planta, R. J., Vanema, J., and Raué, H. A. (1995) *J. Mol. Biol.* **250**, 24 – 36
546a. Borovjagin, A. V., and Gerbi, S. A. (2000) *J. Mol. Biol.* **300**, 57 – 74
546b. Peculis, B. A. (2001) *RNA* **7**, 207 – 219
547. Lygerou, Z., Allmang, C., Tollervey, D., and Séraphin, B. (1996) *Science* **272**, 268 – 270
547a. Clayton, D. A. (2001) *Nature (London)* **410**, 29, 31
548. Bachellerie, J.–P., Michot, B., Nicoloso, M., Balakin, A., Ni, J., and Fournier, M. J. (1995) *Trends Biochem. Sci.* **20**, 261 – 264
549. Wang, H., Boisvert, D., Kim, K. K., Kim, R., and Kim, S.-H. (2000) *EMBO J.* **19**, 317 – 323
550. Reddy, R., Henning, D., Das, G., Harless, M., and Wright, D. (1987) *J. Biol. Chem.* **262**, 75 – 81
551. Hernandez, N., and Weiner, A. M. (1986) *Cell* **47**, 249 – 258
552. Mimori, T., Hinterberger, M., Pettersson, I., and Steitz, J. A. (1984) *J. Biol. Chem.* **259**, 560 – 565
553. Rokeach, L. A., Haselby, J. A., and Hoch, S. O. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4832 – 4836
554. Deutscher, S. L., and Keene, J. D. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3299 – 3303
555. Frendeway, D., Dingermann, T., Cooley, L., and Söll, D. (1985) *J. Biol. Chem.* **260**, 449 – 454
556. Levinger, L., Vasisht, V., Greene, V., Bourne, R., Birk, A., and Kolla, S. (1995) *J. Biol. Chem.* **270**, 18903 – 18909
557. Pfeiffer, T., Tekos, A., Warnecke, J. M., Drainas, D., Engelke, D. R., Séraphin, B., and Hartmann, R. K. (2000) *J. Mol. Biol.* **298**, 559 – 565
558. Rossmanith, W. (1997) *J. Mol. Biol.* **265**, 365 – 371
558a. Doudna, J. A., and Cech, T. R. (2002) *Nature (London)* **418**, 222 – 228
558b. Xiao, S., Scott, F., Fierke, C. A., and Engelke, D. R. (2002) *Ann. Rev. Biochem.* **71**, 165 – 189
558c. Houser-Scott, F., Xiao, S., Millikin, C. E., Zengel, J. M., Lindahl, L., and Engelke, D. R. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2684 – 2689
559. Deutscher, M. D. (1984) *Crit. Revs. Biochem.* **17**, 45 – 71
559a. Li, F., Wang, J., and Steitz, T. A. (2000) *J. Mol. Biol.* **304**, 483 – 492
559b. Seth, M., Thurlow, D. L., and Hou, Y.-M. (2002) *Biochemistry* **41**, 4521 – 4532
560. Grosjean, H., Edqvist, J., Stråby, K. B., and Giegé, R. (1996) *J. Mol. Biol.* **255**, 67 – 85
561. Constantinesco, F., Motorin, Y., and Grosjean, H. (1999) *J. Mol. Biol.* **291**, 375 – 392
562. House, C. H., and Miller, S. L. (1996) *Biochemistry* **35**, 315 – 320
563. Lamond, A. I. (1990) *Trends Biochem. Sci.* **15**, 451 – 452
564. Schnierle, B. S., Gershon, P. D., and Moss, B. (1994) *J. Biol. Chem.* **269**, 20700 – 20706
565. Pillutla, R. C., Yue, Z., Maldonado, E., and Shatkin, A. J. (1998) *J. Biol. Chem.* **273**, 21443 – 21446
566. Terns, M. P., Grimm, C., Lund, E., and Dahlberg, J. E. (1995) *EMBO J.* **14**, 4860 – 4871
567. Shimba, S., and Reddy, R. (1994) *J. Biol. Chem.* **269**, 12419 – 12423
568. Booth, B. L., Jr., and Pugh, B. F. (1997) *J. Biol. Chem.* **272**, 984 – 991
569. Wang, S. P., Deng, L., Ho, C. K., and Shuman, S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9573 – 9578
570. Kuge, H., and Richter, J. D. (1995) *EMBO J.* **14**, 6301 – 6310
571. Hu, G., Gershon, P. D., Hodel, A. E., and Quiocho, F. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7149 – 7154
571a. Hsu, P.-C., Hodel, M. R., Thomas, J. W., Taylor, L. J., Hagedorn, C. H., and Hodel, A. E. (2000) *Biochemistry* **39**, 13730 – 13736
571b. Martin, G., Keller, W., and Doublié, S. (2000) *EMBO J.* **19**, 4193 – 4203
571c. Hayes, R., Kudla, J., and Gruissem, W. (1999) *Trends Biochem. Sci.* **24**, 199 – 202
571d. Gagliardi, D., Perrin, R., Maréchal-Drouard, L., Grienenberger, J.-M., and Leaver, C. J. (2001) *J. Biol. Chem.* **276**, 43541 – 43547
572. Galili, G., Kawata, E. E., Smith, L. D., and Larkins, B. A. (1988) *J. Biol. Chem.* **263**, 5764 – 5770
573. Chen, C.-Y. A., and Shyu, A.-B. (1995) *Trends Biochem. Sci.* **20**, 465 – 470
574. Jacobson, A., and Peltz, S. W. (1996) *Ann. Rev. Biochem.* **65**, 693 – 739
575. Bashirullah, A., Halsell, S. R., Copperstock, R. L., Kloc, M., Karaiskakis, A., Fisher, W. W., Fu, W., Hamilton, J. K., Etkin, L. D., and Lipshitz, H. D. (1999) *EMBO J.* **18**, 2610 – 2620
575a. Matsumoto, K., Wassarman, K. M., and Wolffe, A. P. (1998) *EMBO J.* **17**, 2107 – 2121
575b. Bashirullah, A., Cooperstock, R. L., and Lipshitz, H. D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7025 – 7028
576. Raghow, R. (1987) *Trends Biochem. Sci.* **12**, 358 – 360
577. Hunt, T. (1988) *Nature (London)* **334**, 567 – 568
578. Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Séraphin, B. (2000) *EMBO J.* **19**, 1661 – 1671
579. Tharun, S., He, W., Mayes, A. E., Lennertz, P., Beggs, J. D., and Parker, R. (2000) *Nature (London)* **404**, 515 – 518
579a. Gao, M., Wilusz, C. J., Peltz, S. W., and Wilusz, J. (2001) *EMBO J.* **20**, 1134 – 1143
580. Phizicky, E. M., and Greer, C. L. (1993) *Trends Biochem. Sci.* **18**, 31 – 34
581. Cech, T. R., and Bass, B. L. (1986) *Ann. Rev. Biochem.* **55**, 599 – 629
582. Zaug, A. J., Grosshans, C. A., and Cech, T. R. (1988) *Biochemistry* **27**, 8924 – 8931
583. Saldanha, R., Mohr, G., Belfort, M., and Lambowitz, A. M. (1993) *FASEB J.* **7**, 15 – 24
584. Weeks, K. M., and Cech, T. R. (1995) *Biochemistry* **34**, 7728 – 7738
585. Ho, Y., and Waring, R. B. (1999) *J. Mol. Biol.* **292**, 987 – 1001
585a. Strauss-Soukup, J. K., and Strobel, S. A. (2000) *J. Mol. Biol.* **302**, 339 – 358
586. Bevilacqua, P. C., Sugimoto, N., and Turner, D. H. (1996) *Biochemistry* **35**, 648 – 658

References

587. Golden, B. L., and Cech, T. R. (1996) *Biochemistry* **35**, 3754–3763
588. Nowakowski, J., and Tinoco, I., Jr. (1996) *Biochemistry* **35**, 2577–2585
589. Michel, F., and Ferat, J.-L. (1995) *Ann. Rev. Biochem.* **64**, 435–461
590. Sharp, P. A., and Eisenberg, D. (1987) *Science* **238**, 729–730
590a. Chanfreau, G., and Jacquier, A. (1996) *EMBO J.* **15**, 3466–3476
590b. Swisher, J., Duarte, C. M., Su, L. J., and Pyle, A. M. (2001) *EMBO J.* **20**, 2051–2061
590c. Su, L. J., Qin, P. Z., Michels, W. J., and Pyle, A. M. (2001) *J. Mol. Biol.* **306**, 655–668
590d. Zhang, L., and Doudna, J. A. (2002) *Science* **295**, 2084–2088
591. Diener, T. O. (1984) *Trends Biochem. Sci.* **9**, 133–136
592. Daròs, J. A., Marcos, J. F., Hernández, C., and Flores, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12813–12817
593. Feldstein, P. A., Hu, Y., and Owens, R. A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6560–6565
593a. Navarro, J.-A., and Flores, R. (2000) *EMBO J.* **19**, 2662–2670
594. Forster, A. C., Davies, C., Sheldon, C. C., Jeffries, A. C., and Symons, R. H. (1988) *Nature (London)* **334**, 265–267
595. Haseloff, J., and Gerlach, W. L. (1988) *Nature (London)* **334**, 585–591
596. Dinter-Gottlieb, G. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6250–6254
597. Greer, C. L., and Abelson, J. (1984) *Trends Biochem. Sci.* **9**, 139–141
598. Laski, F. A., Fire, A. Z., RajBhandary, U. L., and Sharp, P. A. (1983) *J. Biol. Chem.* **258**, 11974–11980
599. Phizicky, E. M., Schwartz, R. C., and Abelson, J. (1986) *J. Biol. Chem.* **261**, 2978–2986
600. McCraith, S. M., and Phizicky, E. M. (1991) *J. Biol. Chem.* **266**, 11986–11992
601. Belford, H. G., Westaway, S. K., Abelson, J., and Greer, C. L. (1993) *J. Biol. Chem.* **268**, 2444–2450
601a. Steiger, M. A., Kierzek, R., Turner, D. H., and Phizicky, E. M. (2001) *Biochemistry* **40**, 14098–14105
602. Filipowicz, W., and Gross, H. J. (1984) *Trends Biochem. Sci.* **9**, 68–71
603. Ogden, R. C., Knapp, G., Peebles, C. L., Johnson, J., and Abelson, J. (1981) *Trends Biochem. Sci.* **6**, 154–158
604. Skoglund, U., Andersson, K., Strandberg, B., and Daneholt, B. (1986) *Nature (London)* **319**, 560–564
605. Dreyfuss, G., Swanson, M. S., and Pinol-Roma, S. (1988) *Trends Biochem. Sci.* **13**, 86–91
606. Richter, J. D. (1988) *Trends Biochem. Sci.* **13**, 483–486
607. Müller, S., Wolpensinger, B., Angenitzki, M., Engel, A., Sperling, J., and Sperling, R. (1998) *J. Mol. Biol.* **283**, 383–394
608. Sharp, P. A. (1987) *Science* **235**, 766–771
609. Guthrie, C. (1991) *Science* **253**, 157–163
610. Newman, A. J. (1997) *EMBO J.* **16**, 5797–5800
611. Plessel, G., Lührmann, R., and Kastner, B. (1997) *J. Mol. Biol.* **265**, 87–94
612. Valcárcel, J., and Green, M. R. (1996) *Trends Biochem. Sci.* **21**, 296–301
613. Berget, S. M. (1995) *J. Biol. Chem.* **270**, 2411–2414
614. Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., and Nagai, K. (1994) *Nature (London)* **372**, 432–438
615. Padgett, R. A., Grabkowski, P. J., Konarska, M. M., and Sharp, P. A. (1985) *Trends Biochem. Sci.* **10**, 154–157
616. Deirdre, A., Scadden, J., and Smith, C. W. J. (1995) *EMBO J.* **14**, 3236–3246
617. Blencowe, B. J. (2000) *Trends Biochem. Sci.* **25**, 106–110
618. Reuter, K., Nottrott, S., Fabrizio, P., Lührmann, R., and Ficner, R. (1999) *J. Mol. Biol.* **294**, 515–525
619. Xie, J., Beickman, K., Otte, E., and Rymond, B. C. (1998) *EMBO J.* **17**, 2938–2946
620. Furman, E., and Glitz, D. G. (1995) *J. Biol. Chem.* **270**, 15515–15522
621. Makarov, E. M., Makarova, O. V., Achsel, T., and Lührmann, R. (2000) *J. Mol. Biol.* **298**, 567–575
622. Jiang, J., Horowitz, D. S., and Xu, R.-M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3022–3027
623. Sontheimer, E. J., Sun, S., and Piccirilli, J. A. (1997) *Nature (London)* **388**, 801–805
624. Moore, M. J., and Sharp, P. A. (1993) *Nature (London)* **365**, 364–368
625. Tarn, W.-Y., and Steitz, J. A. (1996) *Science* **273**, 1824–1832
626. Will, C. L., Schneider, C., Reed, R., and Lührmann, R. (1999) *Science* **284**, 2003–2005
627. Bailey-Serres, J., Rochaix, J.-D., Wassenegger, M., and Filipowicz, W. (1999) *EMBO J.* **18**, 5153–5158
628. Noguchi, T., Inoue, H., and Tanaka, T. (1986) *J. Biol. Chem.* **261**, 13807–13812
629. Saitta, B., Stokes, D. G., Vissing, H., Timpl, R., and Chu, M.-L. (1990) *J. Biol. Chem.* **265**, 6473–6480
630. Magnuson, V. L., Young, M., Schattenberg, D. G., Mancini, M. A., Chen, D., Steffensen, B., and Klebe, R. J. (1991) *J. Biol. Chem.* **266**, 14654–14662
631. Nawa, H., Kotani, H., and Nakanishi, S. (1984) *Nature (London)* **312**, 729–734
632. Maeda, N., Kim, N.-S., Azen, E. A., and Smithies, O. (1985) *J. Biol. Chem.* **260**, 11123–11130
633. MacLeod, A. R., Houlker, C., Reinach, F. C., Smillie, L. B., Talbot, K., Modi, G., and Walsh, F. S. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7835–7839
634. Collins, T., Bonthron, D. T., and Orkin, S. H. (1987) *Nature (London)* **328**, 621–624
635. Hsu, T.-C., Shore, S. K., Seshsmma, T., Bagasra, O., and Walsh, P. N. (1998) *J. Biol. Chem.* **273**, 13787–13793
636. Chretien, S., Dubart, A., Beaupain, D., Raich, N., Grandchamp, B., Rosa, J., Goossens, M., and Romeo, P.-H. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6–10
637. Lamond, A. I. (1991) *Trends Biochem. Sci.* **16**, 452–453
638. Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999) *Nature (London)* **398**, 579–585
639. Bruzik, J. P., Van Doren, K., Hirsch, D., and Steitz, J. A. (1988) *Nature (London)* **335**, 559–562
640. Bonen, L. (1993) *FASEB J.* **7**, 40–46
641. Hannon, G. J., Maroney, P. A., Yu, Y.-T., Hannon, G. E., and Nilsen, T. W. (1992) *Science* **258**, 1775–1780
642. Davis, R. E., Hardwick, C., Tavernier, P., Hodgson, S., and Singh, H. (1995) *J. Biol. Chem.* **270**, 21813–21819
643. Ghetti, A., and Abelson, J. N. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11461–11464
644. Herrin, D. L., and Schmidt, G. W. (1988) *J. Biol. Chem.* **263**, 14601–14604
645. Eul, J., Graessmann, M., and Graessmann, A. (1995) *EMBO J.* **14**, 3226–3235
646. Vester, B., Nielsen, A. K., Hansen, L. H., and Douthwaite, S. (1998) *J. Mol. Biol.* **282**, 255–264
647. Bachellerie, J.-P., and Cavaillé, J. (1997) *Trends Biochem. Sci.* **22**, 257–261
648. Gaspin, C., Cavaillé, J., Erauso, G., and Bachellerie, J.-P. (2000) *J. Mol. Biol.* **297**, 895–906
648a. Jády, B. E., and Kiss, T. (2001) *EMBO J.* **20**, 541–551
649. Blum, B., Bakalara, N., and Simpson, L. (1990) *Cell* **60**, 189
650. Wilson, M. A., and Pohorille, A. (1996) *J. Am. Chem. Soc.* **118**, 6580–6587
651. Riley, G. R., Corell, R. A., and Stuart, K. (1994) *J. Biol. Chem.* **269**, 6101–6108
652. Hodges, P., and Scott, J. (1992) *Trends Biochem. Sci.* **17**, 77–81
652a. Keegan, L. P., Gallo, A., and O'Connell, M. A. (2000) *Science* **290**, 1707–1709
652b. Gerber, A. P., and Keller, W. (2001) *Trends Biochem. Sci.* **26**, 376–384
653. Navaratnam, N., Fujino, T., Bayliss, J., Jarmuz, A., How, A., Richardson, N., Somasekaram, A., Bhattacharya, S., Carter, C., and Scott, J. (1998) *J. Mol. Biol.* **275**, 695–714
654. Hersberger, M., and Innerarity, T. L. (1998) *J. Biol. Chem.* **273**, 9435–9442
655. Gray, M. W. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8157–8159
656. Giegé, P., and Brennicke, A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 15324–15329
657. Hirose, T., and Sugiura, M. (1997) *EMBO J.* **16**, 6804–6811
658. Paul, M. S., and Bass, B. L. (1998) *EMBO J.* **17**, 1120–1127
659. Maas, S., Melcher, T., Herb, A., Seeburg, P. H., Keller, W., Krause, S., Higuchi, M., and O'Connell, M. A. (1996) *J. Biol. Chem.* **271**, 12221–12226
660. Bass, B. L. (1997) *Trends Biochem. Sci.* **22**, 157–162
660a. Bass, B. L. (2002) *Ann. Rev. Biochem.* **71**, 817–846
661. Lehmann, K. A., and Bass, B. L. (1999) *J. Mol. Biol.* **291**, 1–13
662. Gerber, A., Grosjean, H., Melcher, T., and Keller, W. (1998) *EMBO J.* **17**, 4780–4789
663. Schwartz, T., Rould, M. A., Lowenhaupt, K., Herbert, A., and Rich, A. (1999) *Science* **284**, 1841–1845
664. Kable, M. L., Seiwert, S. D., Heidmann, S., and Stuart, K. (1996) *Science* **273**, 1189–1195
665. Kable, M. L., Heidmann, S., and Stuart, K. D. (1997) *Trends Biochem. Sci.* **22**, 162–166
666. Rusché, L. N., Cruz-Reyes, J., Piller, K. J., and Sollner-Webb, B. (1997) *EMBO J.* **16**, 4069–4081
667. Frech, G. C., Bakalara, N., Simpson, L., and Simpson, A. (1995) *EMBO J.* **14**, 178–187
667a. Müller, U. F., Lambert, L., and Göringer, H. U. (2001) *EMBO J.* **20**, 1394–1404
667b. Grams, J., McManus, M. T., and Hajduk, S. L. (2000) *EMBO J.* **19**, 5525–5532
667c. Schnaufer, A., Panigrahi, A. K., Panicucci, B., Igo, R. P., Jr., Salavati, R., and Stuart, K. (2001) *Science* **291**, 2159–2162
668. Horváth, A., Berry, E. A., and Maslov, D. A. (2000) *Science* **287**, 1639–1640
669. Wang, S. S., Mahendran, R., and Miller, D. L. (1999) *J. Biol. Chem.* **274**, 2725–2731
669a. Cheng, Y.-W., Visomirski-Robic, L. M., and Gott, J. M. (2001) *EMBO J.* **20**, 1405–1414
670. Petschek, J. P., Mermer, M. J., Scheckelhoff, M. R., Simone, A. A., and Vaughn, J. C. (1996) *J. Mol. Biol.* **259**, 885–890
671. Jacques, J.-P., Hausmann, S., and Kolakofsky, D. (1994) *EMBO J.* **13**, 5496–5503
672. Turner, P. (1985) *Nature (London)* **316**, 105–106
673. Mowry, K. L., and Steitz, J. A. (1987) *Science* **238**, 1682–1687
674. Spirin, A. S., and Ajtkhozhin, M. A. (1985) *Trends Biochem. Sci.* **10**, 162–165
675. Tiedge, H., Bloom, F. E., and Richter, D. (1999) *Science* **283**, 186–187
676. Strauss, E. (1999) *Science* **283**, 12–13
677. Tzagoloff, A., and Myers, A. M. (1986) *Ann. Rev. Biochem.* **55**, 249–285

References

678. Mueller, D. M., and Getz, G. S. (1986) *J. Biol. Chem.* **261**, 11756 – 11764
679. Hedtke, B., Börner, T., and Weihe, A. (1997) *Science* **277**, 809 – 811
680. Shadel, G. S., and Clayton, D. A. (1993) *J. Biol. Chem.* **268**, 16083 – 16086
681. Rossmanith, W., Tullo, A., Potuschak, T., Karwan, R., and Sbisà, E. (1995) *J. Biol. Chem.* **270**, 12885 – 12891
682. Mayer, M., Schiffer, S., and Marchfelder, A. (2000) *Biochemistry* **39**, 2096 – 2105
683. Yokobori, S.-i, and Pääbo, S. (1997) *J. Mol. Biol.* **265**, 95 – 99
684. Margossian, S. P., and Butow, R. A. (1996) *Trends Biochem. Sci.* **21**, 392 – 396
685. Ohyama, K., Fukazawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986) *Nature (London)* **322**, 572 – 574
686. Lewin, R. (1985) *Science* **230**, 55
687. Sambrook, J. (1983) in *Biochemistry* (Zubay, G., ed), Addison-Wesley, Reading, Massachusetts (Chapter 28)
688. Freifelder, D. (1987) *Molecular Biology*, 2nd ed., Jones and Bartlett, Boston, Massachusetts (Chapter 23)
689. Wildeman, A. G. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2123 – 2127
690. Fanning, E., and Knippers, R. (1992) *Ann. Rev. Biochem.* **61**, 55 – 85
691. Nevins, J. R. (1991) *Trends Biochem. Sci.* **16**, 435 – 439
692. Swaminathan, S., and Thimmapaya, B. (1996) *J. Mol. Biol.* **258**, 736 – 746
693. Berget, S. M., Moore, C., and Sharp, P. A. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3171 – 3175
694. Bos, J. L., and van der Eb, A. J. (1985) *Trends Biochem. Sci.* **10**, 310 – 313
695. Hogle, J. M., Chow, M., and Filman, D. J. (1985) *Science* **229**, 1358 – 1365
696. Saito, T., Meshi, T., Takamatsu, N., and Okada, Y. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6074 – 6077
697. Lamb, R. A., and Choppin, P. W. (1983) *Ann. Rev. Biochem.* **52**, 467 – 506
698. Scholtissek, C., and Naylor, E. (1988) *Nature (London)* **331**, 215
699. Pekosz, A., He, B., and Lamb, R. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8804 – 8806
700. Ishihama, A., Mizumoto, K., Kawakami, K., Kato, A., and Honda, A. (1986) *J. Biol. Chem.* **261**, 10417 – 10421
701. Coffin, J. M., Hughes, S. H., and Varmus, H. E., eds. (1998) *Retroviruses*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York

701a. Piot, P., Bartos, M., Ghys, P. D., Walker, N., and Schwartländer, B. (2001) *Nature (London)* **410**, 968 – 973

701b. Cohen, J. (2002) *Science* **296**, 2320 – 2324

702. Varmus, H. (1988) *Science* **240**, 1427 – 1434
703. Sugden, B. (1993) *Trends Biochem. Sci.* **18**, 233 – 235
704. Gallo, R. C., and Montagnier, L. (1988) *Sci. Am.* **259**(Oct), 41 – 48
705. Sanchez-Pescador, R., Power, M. D., Barr, P. J., Steimer, K. S., Stempien, M. M., Brown-Shimer, S. L., Gee, W. W., Renard, A., Randolph, A., Levy, J. A., Dina, D., and Luciw, P. A. (1985) *Science* **227**, 484 – 492
706. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Jr., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., and Wong-Staal, F. (1985) *Nature (London)* **313**, 277 – 284
707. Varmus, H. E. (1985) *Nature (London)* **314**, 583 – 584
708. Vaccaro, J. A., Singh, H. A., and Anderson, K. S. (1999) *Biochemistry* **38**, 15978 – 15985
709. Gabbara, S., Davis, W. R., Hupe, L., Hupe, D., and Peliska, J. A. (1999) *Biochemistry* **38**, 13070 – 13076
710. Litvak, S., Sarih-Cottin, L., Fournier, M., Andreola, M., and Tarrago-Litvak, L. (1994) *Trends Biochem. Sci.* **19**, 114 – 118
711. Litvak, S., and Araya, A. (1982) *Trends Biochem. Sci.* **7**, 361 – 364
712. Fossé, P., Mougel, M., Keith, G., Westhof, E., Ehresmann, B., and Ehresmann, C. (1998) *J. Mol. Biol.* **275**, 731 – 746
713. Gao, H.-Q., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (1999) *J. Mol. Biol.* **294**, 1097 – 1113

713a. Sevilya, Z., Loya, S., Hughes, S. H., and Hizi, A. (2001) *J. Mol. Biol.* **311**, 957 – 971

714. Yu, H., Jetzt, A. E., Ron, Y., Preston, B. D., and Dougherty, J. P. (1998) *J. Biol. Chem.* **273**, 28384 – 28391
715. Bujacz, G., Jaskólski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1995) *J. Mol. Biol.* **253**, 333 – 346
716. Andrake, M. D., and Skalka, A. M. (1996) *J. Biol. Chem.* **271**, 19633 – 19636
717. Greenwald, J., Le, V., Butler, S. L., Bushman, F. D., and Choe, S. (1999) *Biochemistry* **38**, 8892 – 8898
718. Maignan, S., Guilloteau, J.-P., Zhou-Liu, Q., Clément-Mella, C., and Mikol, V. (1998) *J. Mol. Biol.* **282**, 359 – 368

718a. Craigie, R. (2001) *J. Biol. Chem.* **276**, 23213 – 23216

719. Gallo, R. C. (1986) *Sci. Am.* **255**(Dec), 88 – 98
720. Gallo, R. C. (1987) *Sci. Am.* **256**(Jan), 47 – 56
721. Emerman, M., and Malim, M. H. (1998) *Science* **280**, 1880 – 1884
722. Frankel, A. D., and Young, J. A. T. (1998) *Ann. Rev. Biochem.* **67**, 1 – 25
723. Mikaélian, I., Krieg, M., Gait, M. J., and Karn, J. (1996) *J. Mol. Biol.* **257**, 246 – 264
724. Wang, Z., and Rana, T. M. (1998) *Biochemistry* **37**, 4235 – 4243
725. Ivanov, D., Kwak, Y. T., Nee, E., Guo, J., García-Martínez, L. F., and Gaynor, R. B. (1999) *J. Mol. Biol.* **288**, 41 – 56

725a. Jeang, K.-T., Xiao, H., and Rich, E. A. (1999) *J. Biol. Chem.* **274**, 28837 – 28840

726. Cullen, B. R., and Malim, M. H. (1991) *Trends Biochem. Sci.* **16**, 346 – 350
727. Charpentier, B., Stutz, F., and Rosbash, M. (1997) *J. Mol. Biol.* **266**, 950 – 962
728. Barnham, K. J., Monks, S. A., Hinds, M. G., Azad, A. A., and Norton, R. S. (1997) *Biochemistry* **36**, 5970 – 5980
729. Echarri, A., González, M. E., and Carrasco, L. (1996) *J. Mol. Biol.* **262**, 640 – 651

729a. Arold, S. T., and Baur, A. S. (2001) *Trends Biochem. Sci.* **26**, 356 – 363

729b. Pomerantz, R. J. (2002) *Nature (London)* **418**, 594 – 595

729c. Segura-Totten, M., and Wilson, K. L. (2001) *Science* **294**, 1016 – 1017

730. Greene, W. C. (1991) *N. Engl. J. Med.* **324**, 308 – 317
731. Guyadar, M., Emerman, M., Sonigo, P., Clavel, F., Montagnier, L., and Alizon, M. (1987) *Nature (London)* **326**, 662 – 669
732. Overbaugh, J., Donahue, P. R., Quackenbush, S. L., Hoover, E. A., and Mullins, J. I. (1988) *Science* **239**, 906 – 910
733. Nerenberg, M., Hinrides, S. H., Reynolds, R. K., Khoury, G., and Jay, G. (1987) *Science* **237**, 1324 – 1329
734. Haase, A. T. (1986) *Nature (London)* **322**, 130 – 136
735. Turner, B. G., and Summers, M. F. (1999) *J. Mol. Biol.* **285**, 1 – 32
736. Miedema, F., and Klein, M. R. (1996) *Science* **272**, 505 – 506

736a. Ezzell, C. (2002) *Sci. Am.* **286**(Jun), 40 – 45

737. Roberts, J. D., Bebenek, K., and Kunkel, T. A. (1988) *Science* **242**, 1171 – 1173
738. Hill, C. M., and Littman, D. R. (1996) *Nature (London)* **382**, 668 – 669
739. Cohen, J. (1995) *Science* **270**, 917 – 918
740. Kidwell, M. G., and Lisch, D. R. (1998) *Nature (London)* **393**, 22 – 23
741. Kazazian, H. H., Jr. (2000) *Science* **289**, 1152 – 1153
742. Moffat, A. S. (2000) *Science* **289**, 1455, 1457
743. Ovchinnikov, I., Rubin, A., and Swergold, G. D. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10522 – 10527
744. Perl, A., Colombo, E., Samoilova, E., Butler, M. C., and Banki, K. (2000) *J. Biol. Chem.* **275**, 7261 – 7272
745. Southworth, M. W., Adam, E., Panne, D., Byer, R., Kautz, R., and Perler, F. B. (1998) *EMBO J.* **17**, 918 – 926
746. Strauss, E. (2000) *Science* **289**, 374

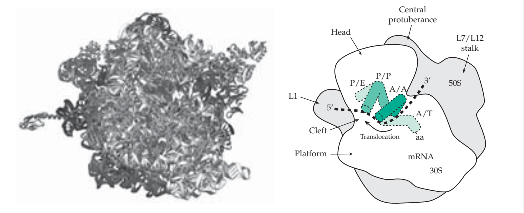
Study Questions

(Additional questions are located on p. 1738.)

1. Describe the sequence of events involved in the initiation of transcription in *E. coli.* As part of your answer, describe those features that must be present in a gene for proper recognition and transcription by the RNA polymerase.

2. How does transcription in eukaryotes differ from that in prokaryotes?

3. In *E. coli* precise spacing between the conserved -35 and -10 (Pribnow) promoter elements has been found to be a critical determinant of promoter strength. What does this suggest about the interaction between RNA polymerase and these conserved sequences in the DNA?



The ~15,000 ribosomes in a bacterial cell synthesize over 4000 proteins following the genetic code in messenger RNA molecules. A ribosome (left) consists of two large subunits, each composed largely of ribosomal RNA, whose folded chains can be seen. About 100 proteins are bound, largely to solvent-exposed surfaces, but with extended "tails" protruding into the ribosome. The messenger RNA (mRNA; right) moves through the ribosome between the large subunits. Amino acids, activated for reaction, are carried into the ribosomes by transfer RNAs (green) which move consecutively from A/T to A/A, A/P, P/P, and P/E sites. They insert their activated amino acids into the growing polypeptide chain in the P site of the 50S subunit. Image of ribosome[33a] courtesy of the authors.

Contents

1673 .... **A. The Architecture of Ribosomes**
1673 .... 1. Ribosomal RNA
1673 .... *Chemical modification and crosslinking*
1676 .... *Phylogenetic comparison*
1676 .... *Structural domains in 16S ribosomal RNA*
1676 .... *23S rRNA*
1677 .... *5S rRNA*
1677 .... 2. Ribosomal Proteins
1681 .... 3. RNA–Protein Interactions and Assembly of Ribosomes
1681 .... *5′-Domain of 16S RNA*
1681 .... *Central domain of 16S RNA*
1684 .... *3′-Domain of 16S RNA*
1684 .... *23S and 5S RNAs*
1684 .... *Eukaryotic ribosomal proteins*
1686 .... 4. Locating Active Sites in Ribosomes
1687 .... *The tRNA-binding sites*
1687 .... *The peptidyltransferase site*
1687 .... *The GTPase-activating center*
1687 .... **B. Transfer RNAs**
1687 .... 1. Structures of Transfer RNAs
1688 .... *Cloverleaf and L forms*
1693 .... *Initiator tRNAs*
1693 .... 2. Pairing of Codon and Anticodon
1694 .... 3. Aminoacylation of tRNAs
1694 .... *Structures*
1694 .... *Recognition of cognate tRNAs*
1695 .... *Mechanisms of reaction*
1695 .... *Correcting errors*
1697 .... *Essential modification reactions of aminoacyl-tRNAs*
1698 .... *Additional functions of aminoacyl-tRNA synthetases*
1698 .... **C. Protein Synthesis: The Ribosome Cycle**
1698 .... 1. Initiation
1700 .... *Prokaryotic initiation factors*
1700 .... *Initiation of protein synthesis in eukaryotes*
1702 .... 2. Elongation of Polypeptide Chains
1702 .... *Codon-specific binding of an aminoacyl-tRNA (decoding)*
1704 .... *The peptidyltransferase reaction*
1705 .... *Elongation factor EF-G and translocation*
1708 .... *Polyribosomes*
1708 .... *Rates of synthesis of ribosomes and of proteins*
1709 .... 3. Termination of Polypeptide Synthesis
1709 .... *Recycling factors*
1709 .... 4. Preventing and Correcting Errors in Translation on the Ribosome
1710 .... *Codon usage*
1710 .... *Dealing with lost peptidyl-tRNAs and broken transcripts*
1711 .... 5. Suppressor Genes
1711 .... *Selenocysteine (Sec)*
1712 .... *Expanding the genetic code*
1712 .... 6. Read-Through Translation and Ribosomal Frameshifting
1714 .... 7. RNA Viruses
1715 .... 8. Other Functions of Ribosomes
1720 .... **D. Processing, Secretion, and Turnover of Proteins**
1721 .... 1. Cotranslational and Posttranslational Processing
1722 .... 2. Forming Functional Proteins
1722 .... 3. Translocation through Membranes
1723 .... 4. Translocation into Organelles
1723 .... 5. Membrane Proteins
1724 .... 6. Secretion of Proteins
1727 .... 7. Protein Folding
1728 .... 8. Completing the Cycle: Proteolytic Degradation of Proteins
1728 .... **E. Proteomics**

1729 .... **References**
1738 .... **Study Questions**

Boxes

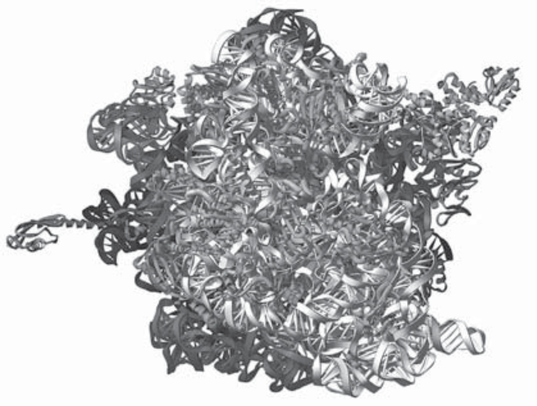
1685 .... Box 29-A The Diphtheria Toxin and Other Ribosome-Inactivating Proteins
1689 .... Box 29-B Antibiotics That Inhibit Protein Synthesis
1713 .... Box 29-C Nonribosomal Peptide Synthesis
1716 .... Box 29-D Protein Splicing, Inteins, and Homing Endonucleases
1718 .... Box 29-E Prions and Amyloid Diseases
1725 .... Box 29-F The Yeast Two-Hybrid System for Identifying Protein–Protein Interactions

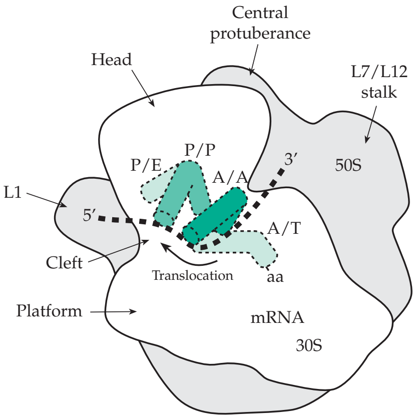
Tables

1672 .... Table 29-1 The Composition of Ribosomes
1677 .... Table 29-2 Ribosomal Proteins from *E. coli*

Ribosomes and the Synthesis of Proteins

29





The many thousands of proteins present in each cell are made within the ribosomes, which are able to read each specific mRNA that comes their way. While ribosomes appear as little more than blurred dots in most electron micrographs, the 15,000 ribosomes of one cell of *E. coli* represent one-fourth of the total mass of the cell. Eukaryotic cells contain many times more of these little molecular machines. When ribosomes were first observed in the early 1950s,[1–3] nobody could imagine either their composition or their function. Less than 50 years later (1999) their complete three-dimensional structure was known at nearly atomic resolution, and the function of ribosomes in protein synthesis was quite well understood. However, the structure could not have been obtained without the development of a whole range of new methods.

As electron microscopy developed, the fuzzy granules in micrographs assumed well-recognized forms. By the 1970s, the approximate shapes of the two ribosomal subunits were known, locations of several ribosomal proteins had been established, and binding sites of transfer RNAs and other features of ribosomes were being mapped.[2,4–7] The resulting picture of the ribosome structure, which is shown schematically in Fig. 29-1, is quite similar to the present-day view. Later, three-dimensional images were reconstructed from electron micrographs (electron tomography),[8,9] and cryo-electron microscopy provided detailed images at a resolution of ~2 nm.[10]

In 1950, when the study of ribosomes began, no methods for determining the sequences of amino acids in proteins or of nucleotides in nucleic acids existed.[11] Sanger published the sequences of the two short chains of insulin in 1953, and the first transfer RNA sequence was published by Holley in 1965.[21] Nevertheless, by 1980 the Wittmanns and coworkers in Berlin had sequenced 53 of the *E. coli* ribosomal proteins[4,22] (Table 29-2), and the three rRNA molecules had also been sequenced.[22a,b,c] In 1950, X-ray crystallography of proteins was still in its infancy; the structure of myoglobin was not determined until 1960. Ribosomal proteins proved hard to crystallize, the first structure being solved in 1980.[23] NMR structural analysis yielded several structures including that of L30 (Fig. 3-25A). In recent years, high-resolution structures of many additional ribosomal proteins in their free forms have been established as have the structures of most of the proteins bound into ribosomes.[24]

The first crystals of bacterial ribosomes in a three-dimensional lattice suitable for study by X-ray diffraction at a resolution of ~1 nm[27] were obtained by Yonath in about 1980.[6,25,26] Now atomic structures are being established at a resolution of 0.3–0.1 nm, or less.[17–19,28–33g] However, such progress would have been impossible without information about ribosomes obtained from improved cryo-electron microscopy,[10,20,33e,34–37a] phylogenetic analysis of ribosomal RNAs,[38–39b] mutational analysis,[40,41] neutron scattering,[42–47] chemical and photochemical crosslinking,[48–53] photoaffinity labeling,[54] immunological labeling,[55] chemical footprinting,[56,57] fluorescence resonance energy transfer (FRET),[58] mass spectrometry,[59,60] and study of the effects of toxic proteins (Box 29-A)[61] and antibiotics (Box 29-B).

Why is the ribosome so large and complex? Aside from the fact that it must form the peptide linkages, it must translate the genetic code in the mRNA into the correct amino acid sequence for each of the thousands of proteins present in the cell. The process takes place

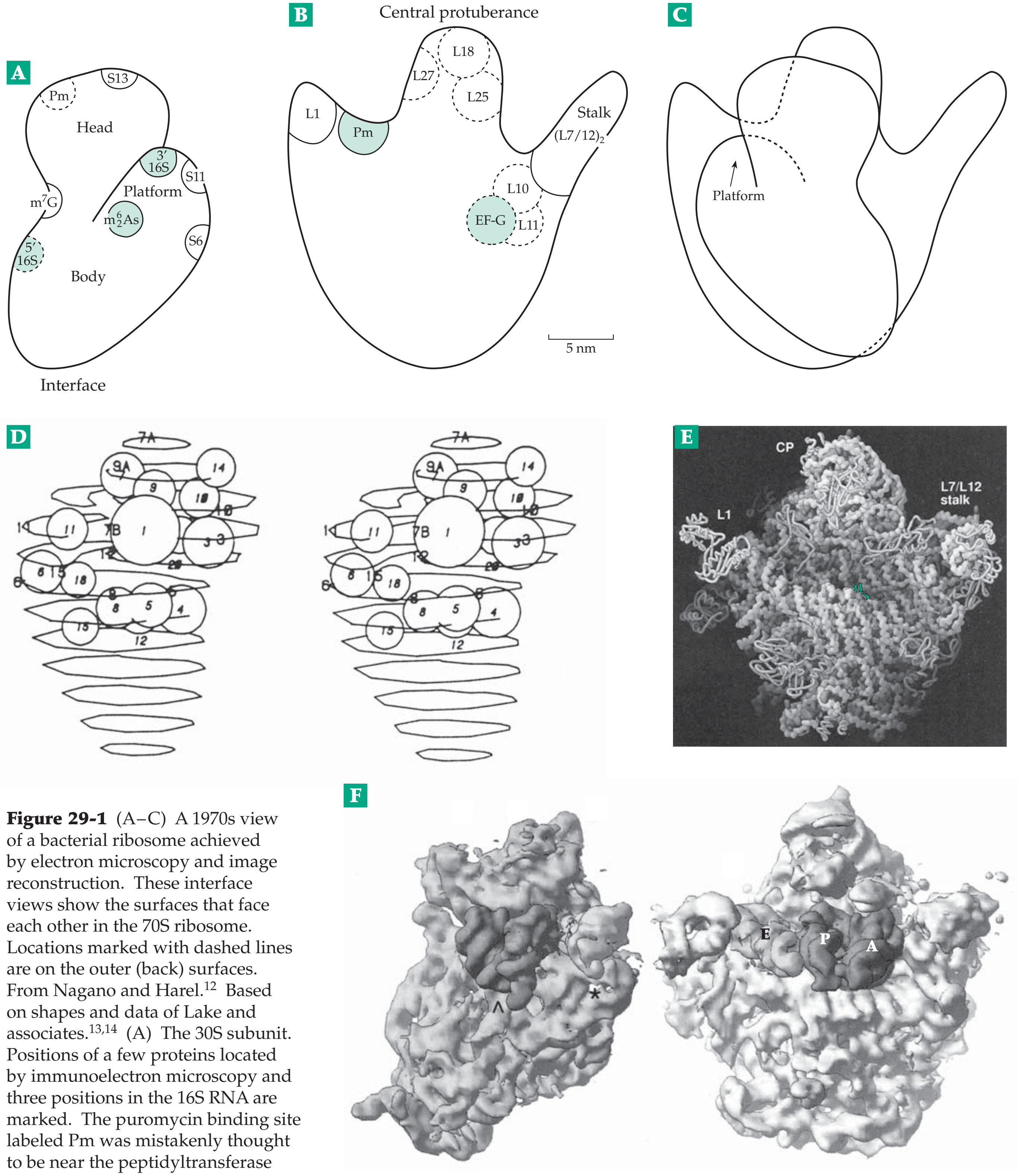
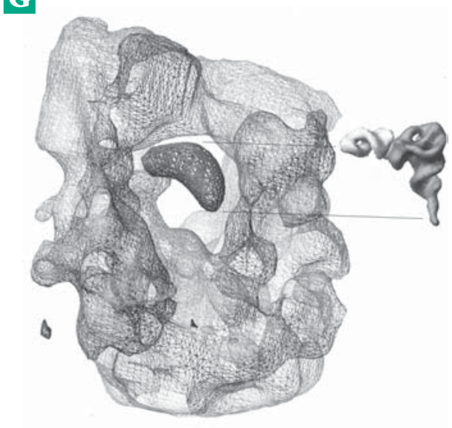
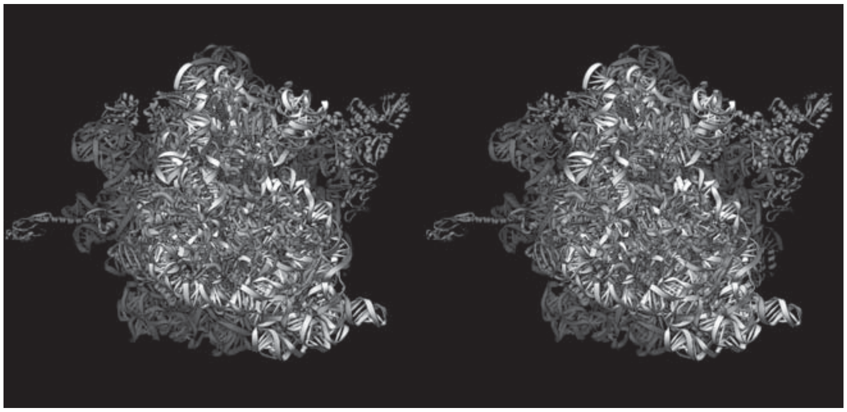
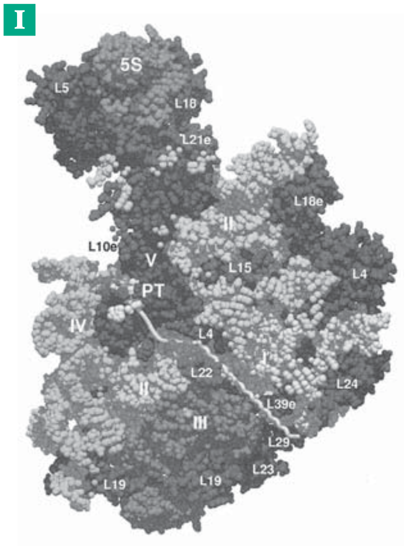
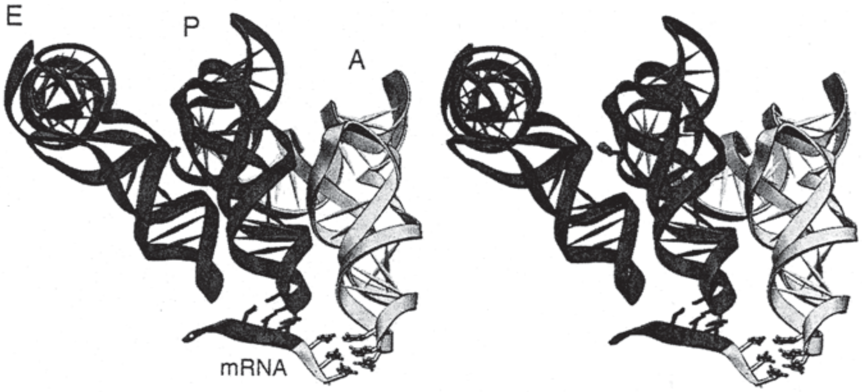

A
Pm
S13
Head
3′ 16S
S11
Platform
m7G
m$^6_2$As
5′ 16S
S6
Body
Interface
B
Central protuberance
L18
L27
L25
L1
Pm
Stalk
(L7/12)$_2$
L10
EF-G
L11
5 nm
C
Platform
D
7A
9A
14
9
10
11
7B
1
3
12
20
6
15
18
8
5
4
13
E
CP
L7/L12 stalk
L1
F
E
P
A
*

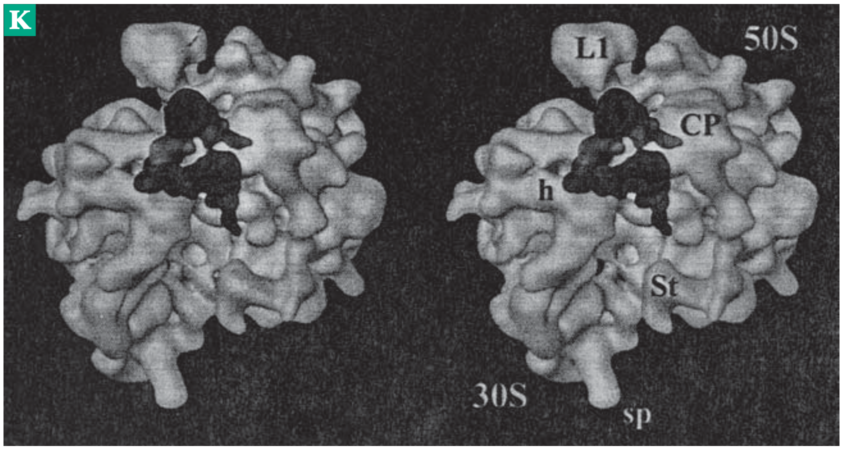
Figure 29-1 (A–C) A 1970s view of a bacterial ribosome achieved by electron microscopy and image reconstruction. These interface views show the surfaces that face each other in the 70S ribosome. Locations marked with dashed lines are on the outer (back) surfaces. From Nagano and Harel.[12] Based on shapes and data of Lake and associates.[13,14] (A) The 30S subunit. Positions of a few proteins located by immunoelectron microscopy and three positions in the 16S RNA are marked. The puromycin binding site labeled Pm was mistakenly thought to be near the peptidyltransferase center. (B) The 50S subunit. Only a few positions of many located by a variety of techniques are marked. EF-G marks the site at which elongation factor G (see Fig. 29-12) binds. A prominent feature is the "stalk", designated St in some of the drawings. It was early shown to be formed by two copies each of the nearly identical L7 and L12. The stalk is flexible and in many circumstances may be folded down or not visible as in the X-ray structures of F–H. (C) The 70S ribosome. The 50S subunit is oriented as in (A) while the 30S subunit has its outer face toward the viewer. (D) Stereoscopic view of a neutron scattering map of the 30S subunit of an *E. coli* ribosome. The proteins studied are represented as spheres with volumes corresponding to those of the anhydrous proteins. Also marked on the map are positions of proteins located by immunoelectron microscopy as mapped by Kahan *et al.*[15] Figure from Ramakrishnan *et al.*[16] Courtesy of V. Ramakrishnan. (E) Model of the 50S subunit from *Haloarcula marismortui.* From











Ban *et al.*[17] Courtesy of T. A. Steitz. The peptidyltransferase center is marked by the green image of the transition state inhibitor shown in Fig. 29-13. (F) Model of three tRNAs bound to a ribosome from *Thermus thermophilus* in the A (aminoacyl), P (peptidyl), and E (exit) sites. These are based on 0.75-nm X-ray data and a number of difference electron density maps. The 3'-CCA end of the A-site tRNA is not modeled but is marked "∧". Views are left, facing the inner surface of the 30S subunit; right, facing the inner surface of the 50S subunit. (G) Schematic side view of a ribosome showing a molecule of tRNA bound in the A site between the 30S and 50S subunits. The anticodon of the tRNA is base-paired with mRNA in the "decoding site" on the 30S subunit. The 3'-CCA end with attached aminoacyl group lies in the peptidyltransferase site in the 50S subunit. (F) and (G) are courtesy of Cate *et al.*[18] (H) Stereoscopic view of a model of the 70S ribosome from *T. thermophilus*. The 30S subunit (lighter) is toward the viewer. Courtesy of Harry F. Noller and Albion Baucom. (I) Section through the 0.24 nm-resolution model of the 50S subunit shown in (E). The modeled path of the polypeptide chain through the exit tunnel is marked. Courtesy of Nissen *et al.*[19] (J) Stereo diagram of the relative orientations of the A-, P-, and E-tRNAs and mRNA showing codon–anticodon interactions and the kink between the A and P codons. (H) and (J) courtesy of Yusupov *et al.*[33a] (K) Stereoscopic view of tRNAs in the P site and in an overlapping P/E site as observed by cryo-electron microscopy at a resolution of 0.5 nm in an image of the 70S ribosome at 1.5 nm resolution. The anticodon arms are to the left. Two tRNA molecules are not present simultaneously but their images have been presented together. From Agrawal *et al.*[20] Courtesy of Rajendra Agrawal.

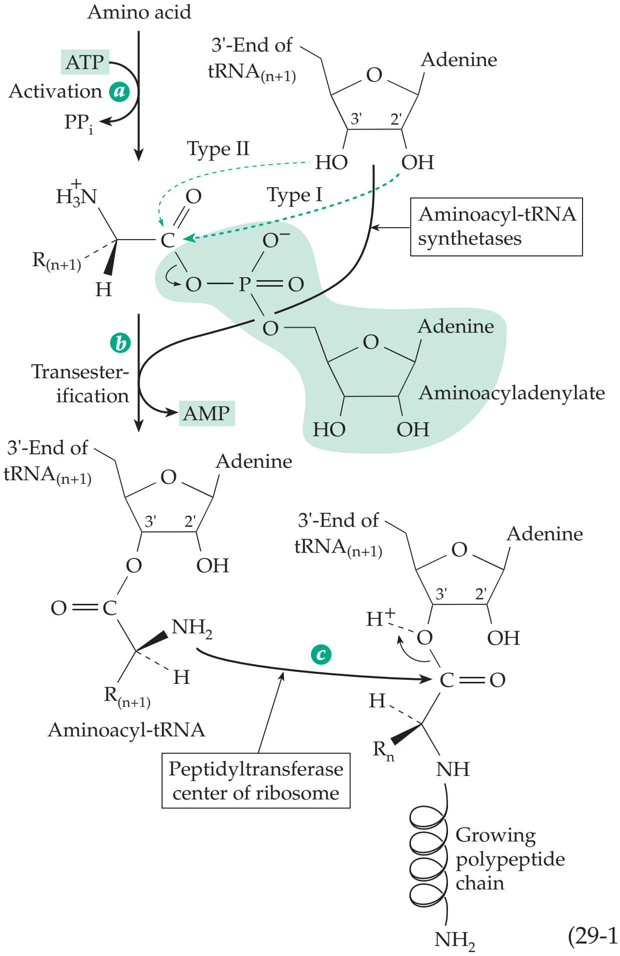
TABLE 29-1
The Composition of Ribosomes

Prokaryotic[a]		Eukaryotic[b]	
Component	Mass, kDa	Component	Mass, kDa
Small (30S) subunit	850	Small (40S) subunit	1440
16S RNA	500	18S RNA	700
Proteins (21)	350 (total)	Proteins (~30)	740
Large (50S) subunit	1450	Large (60S) subunit	2800
23S RNA	950	28S RNA	1700
5S RNA	40	5.8S RNA	51
		5S RNA	39
Proteins (32–34)	460	Proteins (~46)	1010
Complete (70S) ribosome	2300	Complete (80S) ribosome	4240

[a] Data from Wittmann, H. G. (1982) *Ann. Rev. Biochem.* **51**, 155–183. Based on sequences of all components. Presence of spermine, K^+, etc., may add 10%.

[b] Data from Freifelder, D. (1983) *Molecular Biology*, 2nd ed., Jones and Bartlett, Boston, Massachusetts (p. 419) and Mao, H., and Williamson, J. R. (1999) *J. Mol. Biol.* **292**, 345–349.

in several steps: (1) **Initiation** of protein synthesis in bacteria requires binding of the 30S ribosomal subunit to an mRNA molecule, location of the start signal (initiation codon and nearby Shine–Dalgarno sequence, shown in Fig. 29-2 and Eq. 29-8), and binding of the initiator tRNA carrying formylmethionine. The 30S complex must then bind to the 50S ribosomal subunit. (2) **Elongation** of the polypeptide chain in the resulting complete 70S ribosome ensues with binding of the appropriate aminoacyl tRNA to the next codon in the 5' → 3' direction. Base pairs form between the anticodon of the tRNA and the mRNA codon that lies in the **aminoacyl (A) site**; the peptide bond is then formed by the **peptidyltransferase** reaction. This reaction is followed by **translocation**, movement of the initiator tRNA into an exit site at the same time that the second tRNA (together with its mRNA codon and the attached growing peptide chain) moves into the **peptidyl (P) site**. The elongation cycle is repeated until the peptide chain is complete. (3) **Termination** of translation involves release of the completed protein and preparation of the ribosomal subunits for another cycle. The entire process is powered by the hydrolysis of ATP and GTP. The ATP is utilized in a three-step process for the activation of the amino acids, which become linked to the appropriate tRNAs (Eqs. 17-36 and 29-1).[61a] If the inorganic pyrophosphate that is formed is hydrolyzed, two molecules of ATP are required for activation of each amino acid molecule. In addition, at least two molecules of GTP are hydrolyzed to GDP and inorganic phosphate within the ribosome for each peptide linkage formed.


Amino acid
ATP
Activation
a
PP$_i$
3'-End of tRNA$_{(n+1)}$
Adenine
O
3'
2'
Type II
HO
OH
$H_3\overset{+}{N}$
O
Type I
Aminoacyl-tRNA synthetases
C
O$^-$
$R_{(n+1)}$
O
P
O
H
O
Adenine
O
b
Transesterification
Aminoacyladenylate
AMP
HO
OH
3'-End of tRNA$_{(n+1)}$
Adenine
O
3'
2'
3'-End of tRNA$_{(n+1)}$
Adenine
O
OH
O
3'
2'
O=C
NH$_2$
H$^+$
O
OH
c
H
C=O
$R_{(n+1)}$
H
Aminoacyl-tRNA
R_n
Peptidyltransferase center of ribosome
NH
Growing polypeptide chain
NH$_2$

(29-1)

The pairing of codons and anticodons required for insertion of the correct amino acid into the growing polypeptide chain is often referred to as **decoding** of the gene sequence. However, an equally important part of the decoding is the attachment of the correct amino acid to its corresponding **cognate tRNA**. This occurs in the cytoplasm and also in the nucleus.[62] The base pairing of tRNAs and mRNA, which follows, occurs in the **decoding center** on the 30S ribosomal subunit. Both the A and P sites of tRNA-binding and the decoding center are formed by folds of the 16S RNA. The peptide bond formation takes place at the opposite ends of the tRNA molecules in the **peptidyltransferase center** of the 50S subunit (see Figs. 29-1, 29-14). As pointed out in Chapter 12 (top of p. 650), peptidyltransferase is a **ribozyme**. Its active site consists entirely of segments of the 23S RNA (see Fig. 29-14). Another important site is the **GTPase activating center**, at which the 23S

RNA interacts with specific G proteins known as **initiation, elongation,** and **termination factors**.

The proteins in a ribosome may help to hold the RNA into conformations that are correct for its functions. They may also catalyze conformational alterations during the various steps of the translation process. In addition, the proteins may help provide binding sites for substrate molecules and participate in regulatory activities. Both the tRNA **exit (E) site** and the **tunnel** through which the polypeptide chain leaves the ribosome are composed, in part, of ribosomal proteins.

A. The Architecture of Ribosomes

Ribosomes of *E. coli* each have a mass of ~ 2.3 x 10^6 daltons and are ~ 65% RNA and 35% protein. Ribosomes of eukaryotic organisms are larger (~4.3 x 10^6 daltons) and consist of ~50% RNA and 50% protein. Under some conditions such as a low Mg^{2+} concentration complete bacterial ribosomes, called **70S ribosomes**, dissociate into two subunits of unequal size, which are known as **30S** and **50S ribosomal subunits**. The larger 50S subunit is about twice the size of the smaller one (Table 29-1). The small 30S ribosomal subunit contains the 16S rRNA, a chain of ~1500–1700 nucleotides (nt) that, if fully extended, would stretch to a length of over 500 nm. In addition to the highly folded RNA molecule, the 30S subunit contains 21 proteins, each one unique in its amino acid composition and sequence (Table 29-2). Many of these proteins, which are designated S1, S2, S3, etc., are of relatively low molecular mass. Many are strongly basic. They contain numerous lysine and arginine residues, many of which are able to interact with RNA in the ribosome. However, neutral and acidic proteins are also present. The 50S ribosomal subunit contains the ~ 2900 nt 23S rRNA, the ~120 nt 5S RNA, and about 31–34 proteins, two of which (L7 and L12) are present as two copies each. The composition of ribosomes is variable, but most proteins are present in a strict 1:1 ratio. Others may be lacking in some of the ribosomes. Some proteins bind to the ribosomes transiently during their function in protein synthesis as do certain proteins with functions other than protein synthesis. In both subunits the RNA molecules form the internal core. Proteins are largely found on the solvent-exposed surfaces. Some of them form the stalk and other features.[17] They often have globular domains with extended tails that interact with the ribosomal RNA.

Eukaryotic ribosomes are not only larger but also (Table 29-1) contain more protein subunits than do those of bacteria: ~30 for the small subunit and 49 for the large subunit.[63] However, the number of essential proteins may be the same. Both eukaryotic ribosomal proteins and rRNA molecules are larger than those of bacteria. Bacterial ribosomes are ~22 nm in diameter and ~30 nm in the third dimension. Eukaryotic ribosomes are of the order of 1.17 times larger in linear dimensions. Ribosomes of chloroplasts resemble those of eubacteria such as *E. coli* but contain a few more proteins.[64] Mammalian mitochondrial ribosomes also resemble those of bacteria in many respects.[65] However, their RNA chains are shorter and they contain more proteins.[66,66a] The protein content is ~66% compared with ~35% for *E. coli* ribosomes.

1. Ribosomal RNA

The sequences of all three pieces of RNA in the *E. coli* ribosomes are known as are those from many other species. These include eukaryotic mitochondrial, plastid, and cytosolic rRNA. From the sequences alone, it was clear that these long molecules could fold into a complex series of hairpin loops resembling those in tRNA. For example, the 16S rRNA of *E. coli* can fold as in Fig. 29-2A and eukaryotic 18S RNA in a similar way (Fig. 29-4).[38,39,67–69] The actual secondary structures of 16S and 18S RNAs, within the folded molecules revealed by X-ray crystallography, are very similar to that shown in Fig. 29-2A. Ribosomal RNAs undergo many posttranscriptional alterations. Methylation of 2’-hydroxyls and of the nucleic acid bases as well as conversion to pseudouridines (pp. 1638–1641) predominate over 200 modifications, principally in functionally important locations that have been found in human rRNA.[69a]

Chemical modification and crosslinking. Before high-resolution X-ray data were available, two major biochemical approaches were used to deduce the secondary structures of ribosomal RNAs.[38,39] The first was the application of chemical reagents and enzymes that modify the RNA. Crosslinking reagents were used to establish pairs of nucleotides that lie close together in the three-dimensional structure. Cleavage by specific endonucleases was used to establish whether a region of the molecule is double-helical or single-stranded.[68] Nucleases were also used to clip out base-paired fragments, which were separated, denatured, and sequenced. This revealed both hairpin loops and pairings between regions that are far apart in the primary sequence. The ability of nucleic acid bases to undergo specific chemical reactions at positions not involved in base pairing was used to establish whether or not a given base was actually paired.[67,69] Thus, every position in *E. coli* 16S RNA was probed by reactions of dimethylsulfate with adenine at N1 and cytosine at N3, reaction of kethoxal (Eq. 5-16) with guanine at N1 and N2', and by reaction of a carbodiimide with uracil at N3 and with guanine at N1.[67]

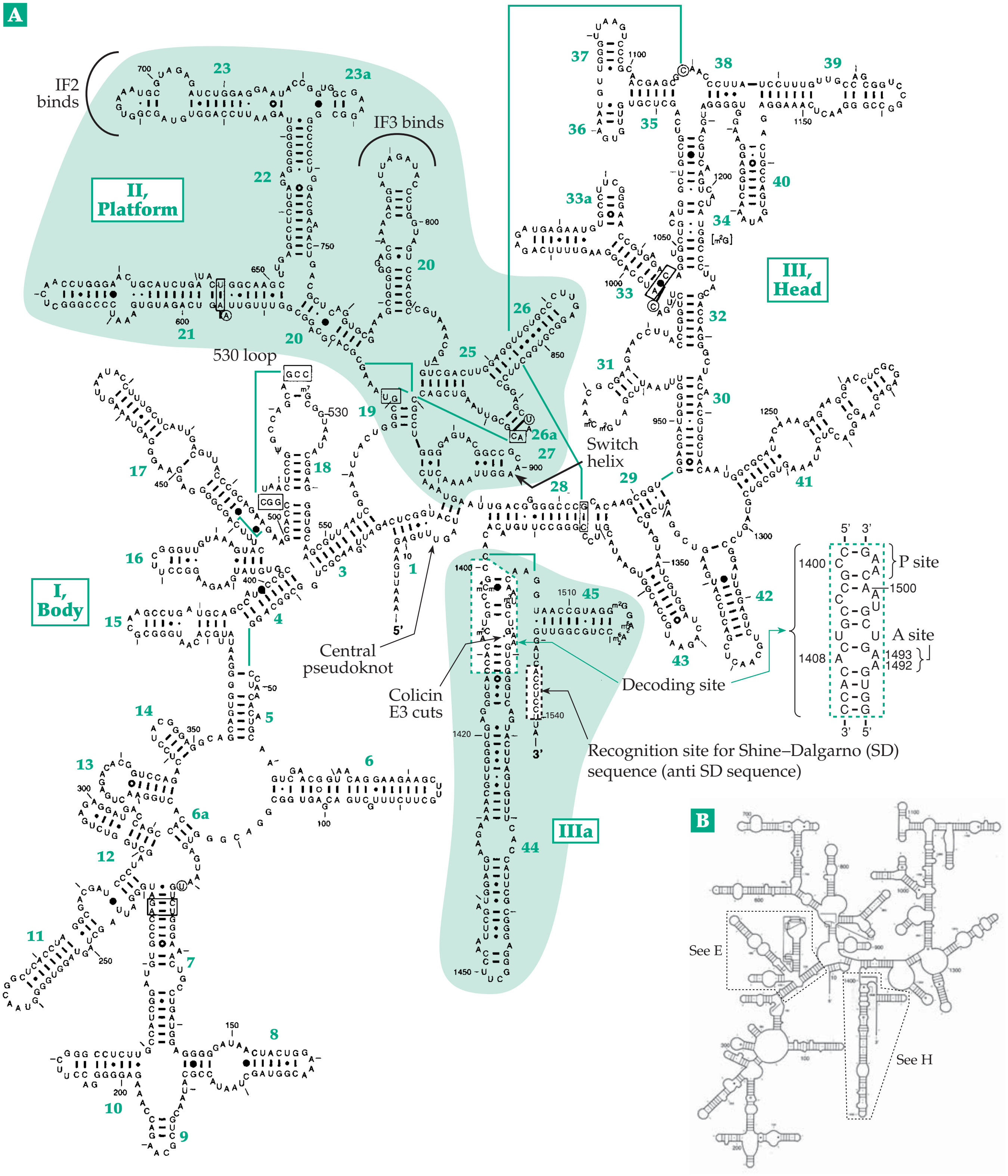
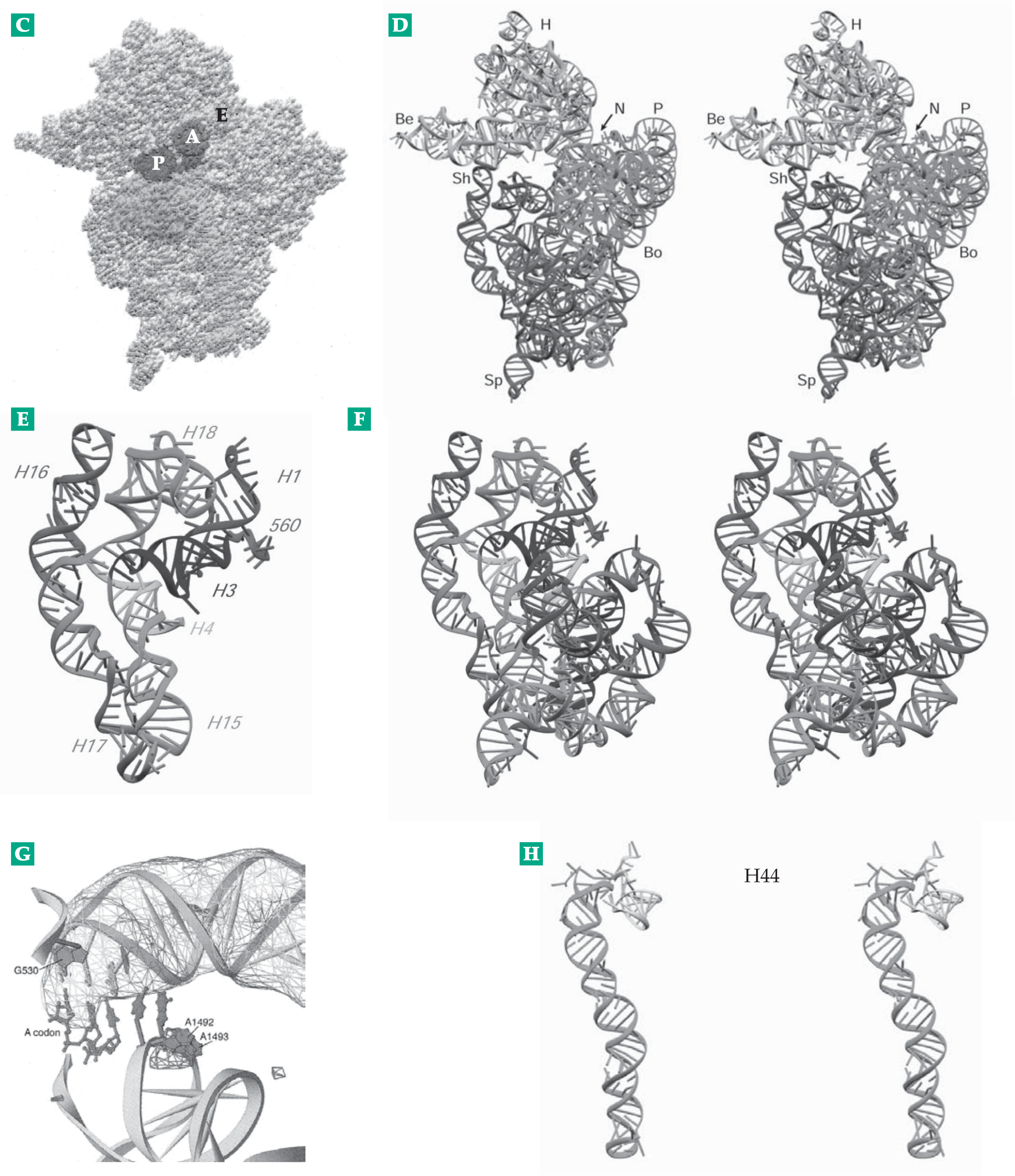

A
IF2 binds
23
23a
IF3 binds
II, Platform
22
21
20
530 loop
19
25
26
26a
27
28
Switch helix
17
18
16
I, Body
15
4
3
1
5'
Central pseudoknot
Colicin E3 cuts
14
5
13
6
6a
12
11
7
8
9
10
37
38
39
36
35
40
33a
34
III, Head
33
32
31
30
29
41
42
43
45
44
IIIa
Decoding site
Recognition site for Shine–Dalgarno (SD) sequence (anti SD sequence)
P site
A site
1400
1500
1408
1493
1492
3'
B
See E
See H

Figure 29-2 (A) Secondary structure model for the 1542-residue *E. coli* 16S rRNA based on comparative sequence analysis.[73a] Dots indicate G•U or A•G pairs; dashes indicate G•C or A•U pairs. Strongly implied tertiary interactions are shown by solid green lines. Helix numbering according to Brimacombe. Courtesy of Robin Gutell. (B) Simplified schematic drawing of type often used. (C) Positions of the A, P, and E sites on the 30S ribosomal subunit from Carter *et al.*[70] (D) Stereoscopic view of the three-dimensional fold of the 16S RNA from *Thermus thermophilus* as revealed by X-ray structural analysis at 0.3 nm resolution. Features labeled are the head (H), beak (Be), neck (N), platform (P), shoulder (Sh), spur (Sp), and body (Bo). (E-H) Selected parts of the 16S RNA. In (E) and (F) the helices are numbered as in (A). (F) and (H) are stereoscopic views. The decoding site


C
E
A
P
D
H
Be
N
P
Sh
Bo
Sp
H
Be
N
P
Sh
Bo
Sp
E
H18
H16
H1
560
H3
H4
H15
H17
F
G
G530
A codon
A1492
A1493
H
H44

is located at the upper end of helix 44. (G), (H). In (G) the electron density difference observed upon binding of tRNA into the A site is displayed as a Fourier difference map (at 0.7-nm resolution). The molecular model of the tRNA with its anticodon paired with a codon from mRNA is superimposed. Two positions of bases A1492 and A1493 are shown as they are found in the presence and absence of paromomycin. A patch of negative density can be seen near the A1492 and A1493 labels, indicating that these groups may rearrange to interact with the minor groove of the codon–anticodon helix when the A-tRNA is bound. See also p. 1690. Courtesy of Yusupov *et al.*[33a] (D) through (F) and (H) are from Wimberly *et al.*[33] Courtesy of Venki Ramakrishnan.

Phylogenetic comparison. This technique, also called comparative sequence analysis, has proved very powerful.[38,39,39a,b,71,72] An example is illustrated in Fig. 29-3. Here a loop from 23S RNA of *E. coli* is shown and is compared with the sequence of 26S RNA from the fungus *Physarum polycephalum*.[38] Wherever the latter differs from the *E. coli* sequence, the substituted base is indicated in a box. These square boxes, which are concentrated in base-paired regions, indicate compensatory changes for which there is usually a second change that preserves base pairing in a double helical region. The studies also showed clearly that bacterial 16S RNA is homologous with eukaryotic 18S RNA,[68,69] with 17S RNA of dinoflagellates,[73] and also with 12S RNA of human mitochondria. Likewise, 23S RNA of bacteria corresponds to 28S RNA of eukaryotes.

Structural domains in 16S ribosomal RNA. Three major compact structural domains, 5', central, and 3', can be distinguished in 16S RNA.[33] An extended subdomain is also present at the 3' end. These are indicated on Fig. 29-2 as I, II, III, and IIIa. The double helical segments are also numbered. Ribosomal RNA molecules must be folded into compact forms to fit into the envelope of the ribosomal subunits. The individual structural domains form independent globular cores to which several proteins apiece are bound. Domains I, II, and III form the body, platform, and head, respectively (Fig. 29-1A). Relatively minor changes in conformation accompany the incorporation of the rRNA molecules into the ribosomes.[67] On this basis, and taking account of all available data, attempts were made for many years to predict a three-dimensional structure.[12,74–76] One of these[77] is portrayed in Fig. 5-32A as a series of cylinders representing the 45 double-helical segments suggested by the structure of Fig. 29-2A. This can be compared with the X-ray based model shown in Fig. 29-2B.

23S rRNA. The large RNA of the 50S subunit consists of six structural domains.[5] Its secondary structure is shown in Fig. 29-4. As with 16S RNA each domain is tightly folded. However, the domains are interdigitated in such a way that they form a single monolithic structural unit.[17] Nevertheless, there are distinct catalytic sites, as described in Section 4. Like proteins, which are able to undergo conformational alterations that usually involve some rearrangement in their internal hydrogen-bonding patterns, these large RNA molecules may also assume alternative conformations. Conformational changes may involve not only alternative hydrogen bonding patterns but also alternative base-pairing.[77a] Such changes may be essential to the functioning of ribosomes[86] and may also accompany maturation of pre-rRNAs.[87] Eukaryotic 28S RNAs have basically the same structures as the

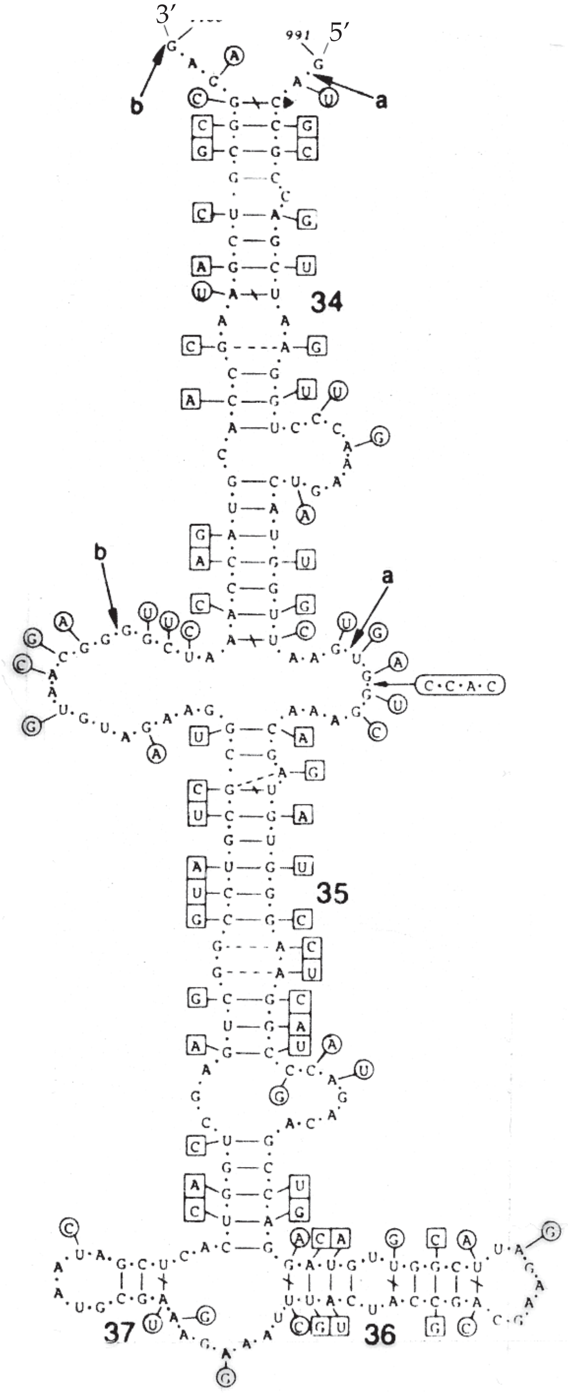


Figure 29-3 Example of phylogenetic comparisons in ribosomal RNA. The diagram shows helices 34–37 of *E. coli* 23S RNA, compared with the corresponding region of *Physarum polycephalum* 26S RNA. The diagram depicts the *E. coli* helices, with changes in the *P. polycephalum* sequences denoted by symbols in boxes on the side. Bases in square boxes are compensating; those in round boxes are mismatching or in single-stranded regions. Solid triangles denote deletions, while bases with arrows indicate insertions. Dotted lines or "crossed-out" base pairs denote modified base-pairing in *P. polycephalum*. The letters "a" and "b" indicate the termini of RNA fragments isolated as a base-paired complex. From Brimacombe.[38]

23S RNAs of bacteria but have been expanded by insertion of additional nucleotides at many places.[10,36,79]

Most of the chemical activity of ribosomes occurs in the interface between the 30S and 50S subunits. Entrance and exit tunnels for both mRNA and the aminoacylated tRNAs are formed between these subunits. The mRNA apparently moves across the platform as the tRNAs move from A to P to E sites experiencing codon selection (decoding) and peptidyltransferase activity. Many loop ends from 16S RNA interact with those of 23S RNA.[41,88]

5S rRNA. This ~120-nucleotide molecule organizes one domain of the 50S ribosomal subunit.[89] Extensive phylogenetic comparisons of 5S RNA sequences led to the secondary structure shown in Fig. 29-5.[90,91] The three-dimensional structure, as seen in a ribosome, is also shown in this figure. Study of base-pairing possibilities suggests that 5S RNA can exist in more than one conformation.[90,92] In a possible second conformation the sequence GUGUGGGG (residues 79–86) pairs in an antiparallel fashion with the sequence CCCCAUGC (residues 35–42), with loss of base pairing in stem 4 (Fig. 29-5). A structure somewhat similar to that of 5S RNA is probable for eukaryotic 5.8S RNA.[94] Nearly a thousand different prokaryotic and eukaryotic 5S RNA sequences have been compared.[95] From them **phylogenetic trees**, which suggest evolutionary pathways between species, have been constructed.[96] Sequences of 16S RNA have been used in a similar way (Fig. 1-5).[97]

2. Ribosomal Proteins

Ribosomal proteins are soluble in concentrated salt solutions. Most of them can be dissolved without damage by buffers containing 2 M LiCl and can then be separated by electrophoresis or ion exchange chromatography and gel filtration (molecular sieving).[98] Although many of them are quite insoluble and are often unstable, all ribosomal proteins of *E. coli* (Table 29-2) have been separated and sequenced, mainly by Wittmann-Liebold and coworkers.[22] The ribosomal proteins of other bacteria usually resemble those of *E. coli*.[99] The more numerous eukaryotic ribosomal proteins have also been isolated and studied individually.[100] Many of these 84 proteins appear to correspond directly in properties and functions to those of *E. coli*.[100a] As with ribosomal RNAs, the sizes of the eukaryotic proteins have been expanded.[101] Mitochondria have their own set of ribosomal proteins, which are more numerous than those of either *E. coli* or yeast.[65–66a,102–102c] Pure individual ribosomal proteins are now produced from the cloned genes as are 16S and 23S ribosomal RNAs.

Most ribosomal proteins are folded into compact forms, much of whose surfaces are accessible to added reagents. However, X-ray structures have revealed that parts of some proteins penetrate deeply into the RNA core.[17] Much of the RNA is also accessible from the outside, and the ribosome contains ~50% of its mass as internal hydration. A ribosome usually contains only one molecule of each kind of protein with the exception of proteins L7 and L12 of the large subunit. There are two of each. Sequencing of the 120-residue proteins from *E. coli* shows that L7 is

TABLE 29-2
Ribosomal Proteins from *E. coli*[a]

Proteins of 30 S Ribosomal subunits			Proteins of 30 S Ribosomal subunits		
Designation	Mass, kDa	Binding[b]	Designation	Mass, kDa	Binding[b]
S1	61.2		L1	24.6	
S2	26.6		L2	29.4	+
S3	25.8		L3	22.3	
S4	23.1	+	L4	22.1	
S5	17.5		L5	20.2	
S6	15.7		L6	18.8	+
S7					
(strain K)	19.7	+	L7	12.2	
(strain B)	17.1	+			
S8	14.0	+	L8		
S9	14.6		L9	15.5	
S10	11.7		L10	17.7	
S11	13.7		L11	14.9	
S12	13.6		L12	12.2	
S13	13.0		L13	16.0	
S14	11.1		L14	13.5	
S15	10.0	+	L15	15.0	
S16	9.2		L16	15.3	+
S17	9.6	+	L17	14.4	+
S18	8.9		L18	12.8	+
S19	10.3		L19	13.0	+
S20	9.6	+	L20	13.4	+
S21	8.4		L21	11.6	
			L22	12.2	
Total mass 350 (strain K)			L23	11.0	+
			L24	11.2	+
			L25	10.7	+
			L26 = S20	9.6	
			L27	9.0	
			L28	8.9	
			L29	7.3	
			L30	6.4	
			L31	7.0	
			L32	6.3	
			L33	6.3	
			L34	5.4	
			Total mass 460[c]		

[a] Molecular masses from Wittmann, H. G. (1982) *Ann. Rev. Biochem.* **51**, 155–183

[b] A plus sign indicates direct binding to ribosomal RNA.

[c] Four copies of L7/L12 are assumed.

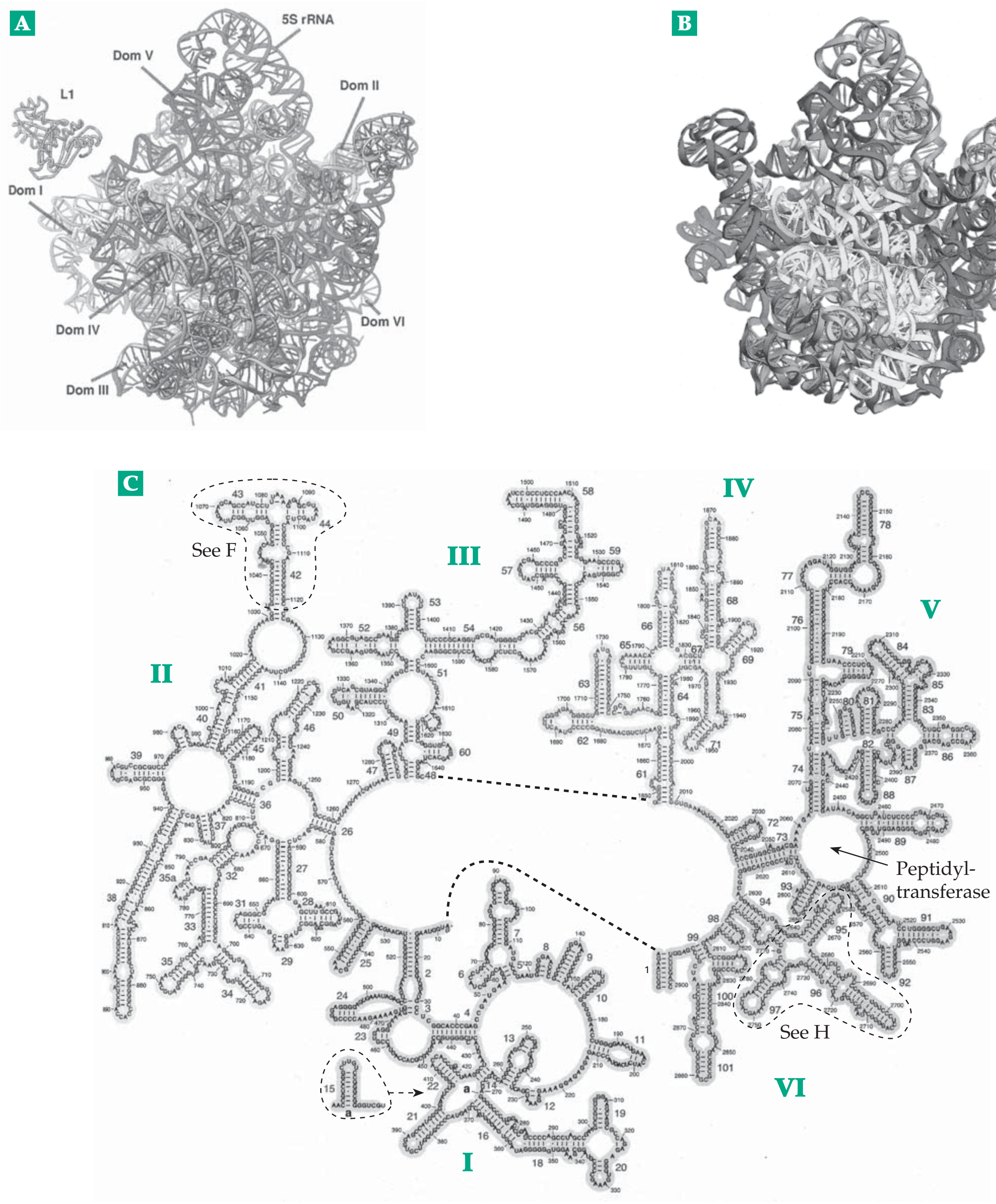
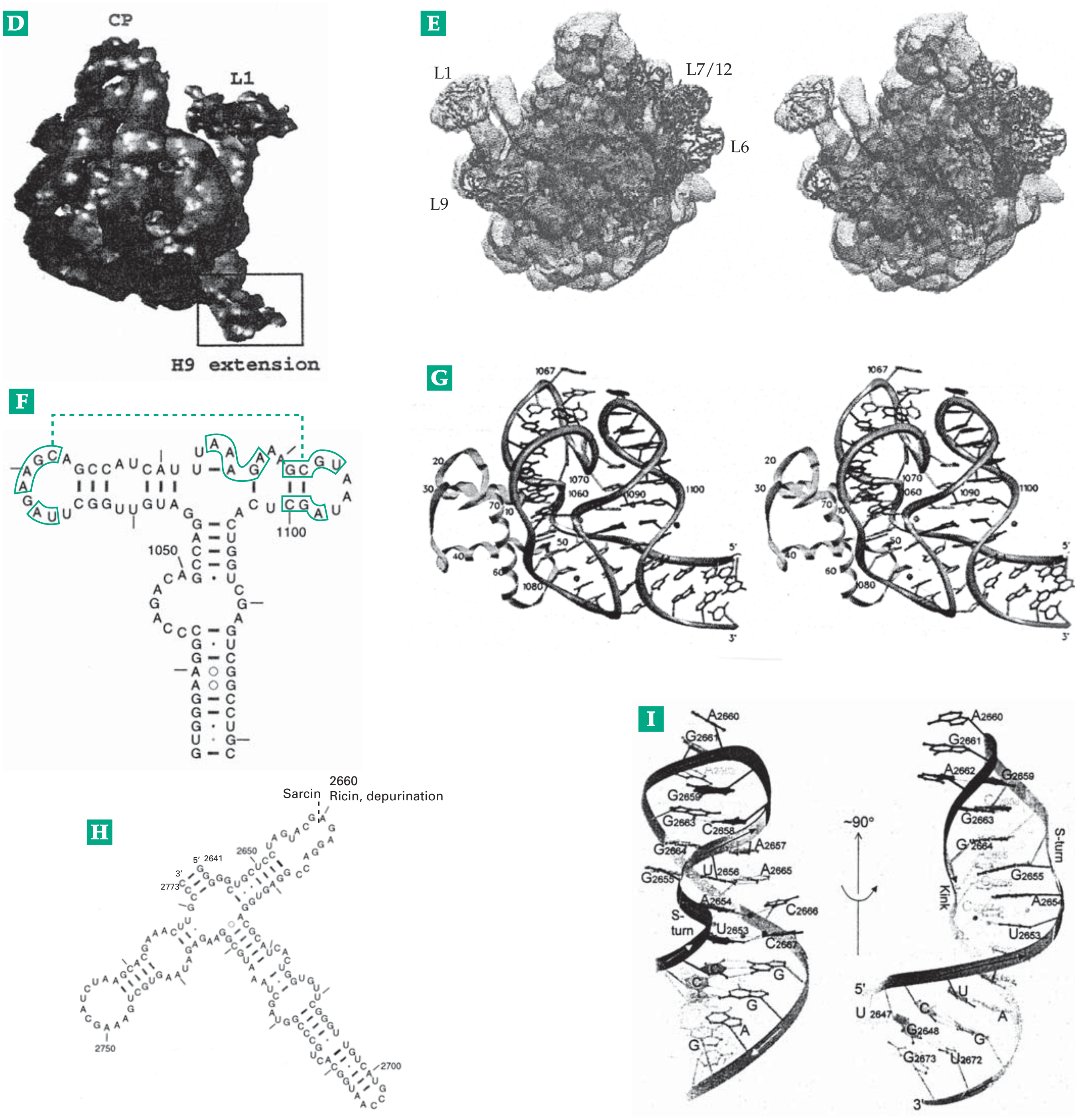


Figure 29-4 Structure of 23S–28S ribosomal RNAs. (A) The three-dimensional structure of RNA from the 50S subunit of ribosomes of *Haloarcula marismortui*. Both the 5S RNA and the six structural domains of the 23S RNA are labeled. Also shown is the backbone structure of protein L1. From Ban *et al.*[17] Courtesy of Thomas A. Steitz. (B) The corresponding structure of the 23S RNA from *Thermus thermophilus*. Courtesy of Yusupov *et al.*[33a] (C) Simplified drawing of the secondary structure of *E. coli* 23S RNA showing the six domains. The peptidyltransferase loop (see also Fig. 29-14) is labeled. This diagram is customarily presented in two halves, which are here connected by dashed lines. Stem-loop 1, which contains both residues 1 and 2000, is often shown in both halves but here only once. From Merryman *et al.*[78] Similar diagrams for *Haloarcula marismortui*[17] and for the mouse[79] reveal a largely conserved structure with nearly identical active sites. (D) Cryo-electron microscopic (Cryo-EM) reconstruction of a 50S subunit of a modified *E. coli* ribosome. The RNA has been modified genetically to have an


D
CP
L1
H9 extension
E
L1
L7/12
L6
L9
F
1050
1100
G
1067
20
30
40
60
70
10
50
1070
1060
1090
1100
1080
5'
3'
H
2660
Sarcin
Ricin, depurination
2650
5' 2641
3'
2773
2750
2700
I
A2660
G2661
A2662
G2659
G2663
C2658
G2664
A2657
G2655
U2656
A2665
A2654
C2666
S-turn
U2653
C2667
~90°
Kink
5'
U 2647
G2648
G2673
U2672
3'

approximately 34 nt predicted extension of helix 9 of the 16S RNA (see Fig. 29-2). The helix 9 extension, clearly visible in this image, locates that helix in *E. coli* ribosomes, which have not yet given crystals satisfactory for X-ray investigation. From Matadeen *et al*.[79a] Courtesy of Richard Brimacombe. (E) Stereoscopic interface view of the 50S subunit of an *E. coli* ribosome with atomic structures of ribosomal proteins fitted to the cryo-EM density (semitransparent) of the 50S subunit. Protein structures are displayed as backbone tubes, and rRNA fragments in ball-and-stick format. Courtesy of Mueller *et al*.[37a] (F) The GTPase-activating loop of 23S RNA of *E. coli*. This loop, from domain II, binds to protein L11, which shields nucleotide A1067 from methylation and prevents the binding of the antibiotics thiostrepton (Box 29-B) and micrococcin. Green nucleotides are highly conserved in bacterial, chloroplast, and mitochondrial RNAs. The small loop (1054–1081) containing the thiostrepton-binding site is also part of the binding site for elongation factors EF-Tu and EF-G.[80,81] (G) Stereoscopic view of the 58-nucleotide loop shown in (E) with the associated protein L11. Courtesy of Conn *et al*.[82] (H) Secondary structure of the sarcin/ricin (SR domain) of the *E. coli* 23S RNA.[83–85] The site of hydrolytic cleavage by the ribonuclease sarcin (Box 29-A) is indicated as is the site of depurination catalyzed by the plant toxin ricin (Box 29-A). (I) Three-dimensional structure of the sarcin–ricin loop. The two views are from directions 90° apart. The sites of attack by ribotoxins are at the top. Courtesy of Correll *et al*.[83]

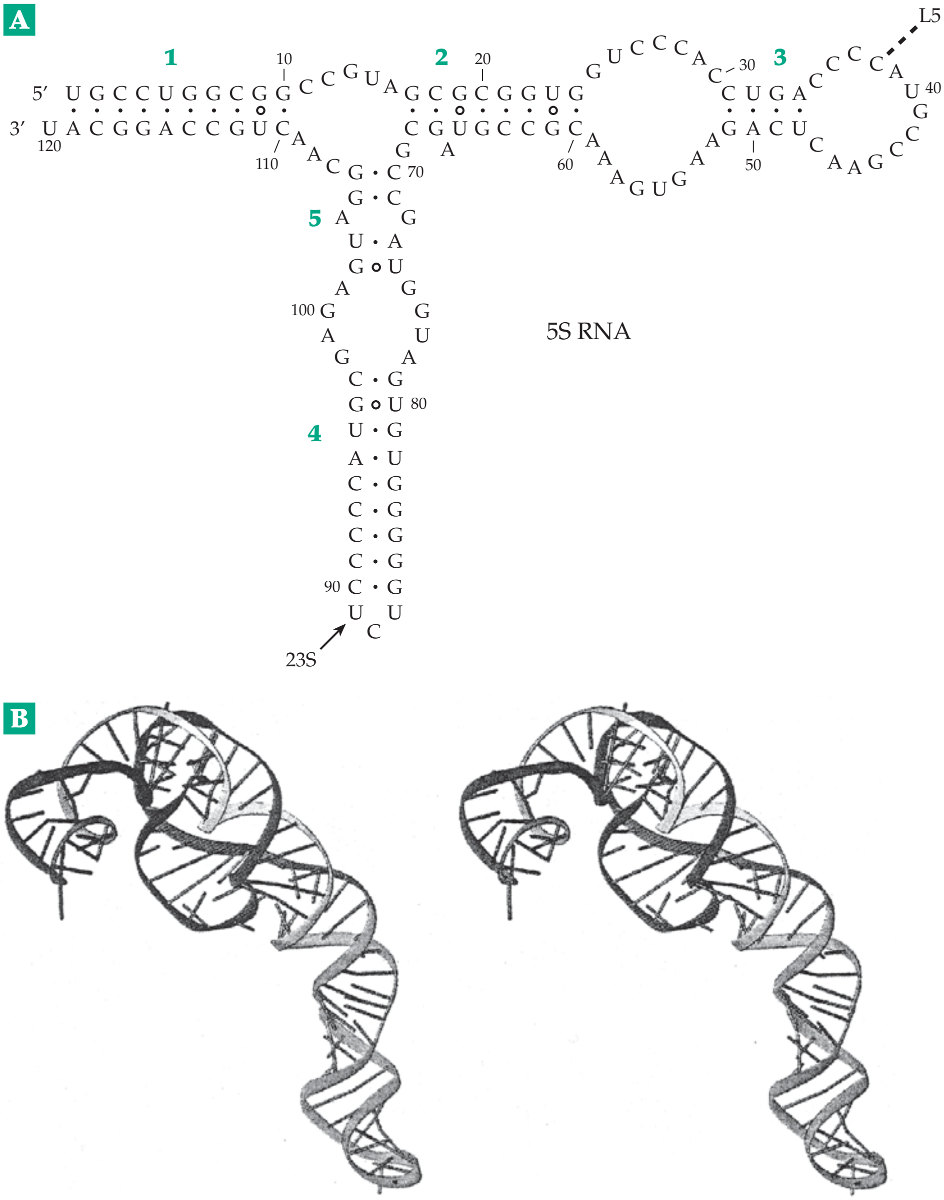

A
1
2
3
4
5
L5
5′ U G C C U G G C G G C C G U A G C G C G G U G G U C C C A C C U G A C C C C A U G C C
3′ U A C G G A C C G U C A A C G C G U G C C G C A A G U G A A G A C U C A A G C C
10
20
30
40
50
60
70
80
90
100
110
120
G · C
G · C
A G
U · A
G ∘ U
A G
G G
A U
G A
C · G
G ∘ U
U · G
A · U
C · G
C · G
C · G
C · G
C · G
U U
C
23S
5S RNA
B

Figure 29-5 (A) Secondary structure of *E. coli* 5S RNA with five universal helical stems (labeled 1–5). This small RNA is found in the central protuberance of the 50S ribosomal subunit. See Fig 29-4A. Photocrosslinking using thiouridine-containing 5S RNA suggested a close proximity of U89 (marked by arrow) with nucleotide 2477 of the 23S RNA in the loop end of helix 89 (Fig. 29-4).[93] (B) Stereoscopic view of the 5S RNA as observed in ribosomes of *Haloarcula marismortui*. From Ban *et al*.[17] Courtesy of Thomas A. Steitz.

N-acetylated L12. Thus, the 50S ribosomal subunit is often described as containing four copies of protein L7/12. They form the flexible stalk seen in Fig. 19-1.

Most ribosomal proteins are rich in lysine and arginine and, therefore, carry a substantial net positive charge. Proteins S20, L7/12, and L10 have over 20% alanine, while L29 is almost as rich in leucine. Proteins S10, S13, L7/L12, L27, L29, and L30 are surprisingly low (<2 mol %) in aromatic amino acids. Proteins S5, S18, and L7 have acetylated N termini while L11, L3, L7/12, L11, L16, and L33 contain methylated amino acids. L11 contains nine methyl groups.[22] Protein S6 is the major phosphoprotein of eukaryotic ribosomes.[103,104] Most ribosomal proteins have no known enzymatic activity. Although often difficult to crystallize, high-resolution three-dimensional structures are known for many free ribosomal proteins.[24] Most of them have shapes resembling those previously found in globular proteins, including DNA-binding proteins. Many have extended "tails" that reach into the interior of the ribosome.[33b] A few seem to assume a defined shape only when packed into a suitable niche in the ribosome. Proteins L7/L12 and the 60-residue L30[105,105a] have similar folding patterns with 2–3 helices and a 3-strand β sheet. The structure of L30 of *E. coli* was deduced[106] by NMR methods (Fig. 3-25) and resembles that from the *Bacillus stearothermophilus* determined by X-ray diffraction.[105,107]

Many specific parts of ribosomal RNA molecules and specific proteins within the intact ribosome were located prior to the determination of high resolution crystal structures. One major approach was the use of **immunoelectron microscopy**. Antibodies to specific ribosomal proteins or to special sites in the RNA were prepared, and electron microscopy was used to map the binding sites of the antibodies on the ribosomal subunit surfaces.[108,109] In this manner, the locations of numerous proteins in both the 30S and 50S subunits were identified. A few of these are indicated in Fig. 29-1A,B.[5] In several instances more than one distinct antibody binding site was found for a given protein. Pairs of sites were sometimes 8–19 nm apart, suggesting that these proteins assumed an elongated or fibrous conformation. However, X-ray studies have established more compact structures for many of the proteins. Perhaps the ease of denaturation of the proteins led to some errors in localization with antibodies. The X-ray studies have now established exact locations for almost all of the ribosomal proteins. However, the correct identification of each protein involved extensive measurements, many of which were done prior to the availability of the X-ray structures.

A variety of crosslinking reagents have been used to locate the positions of specific proteins within ribosomes. For example, bifunctional compounds may bind covalently to two different SH groups or NH_2 groups.[110,111] Among the many crosslinked protein pairs identified in this way are S5-S8, S7-S9, S6-S18, and S13-S19.[112] Crosslinking experiments on both small and large ribosomal subunits have yielded complex distance maps that helped to establish the packing relationships.[113]

Another important approach has been to isolate ribosomal proteins from bacteria grown in D_2O and then to reconstitute ribosomal subunits with pairs of deuterated proteins. By studying **neutron scattering** the distances between the centers of mass of these pairs could be measured. By triangulation the three-dimensional relationship of the entire group of proteins could be determined. The results of such studies[43,47,114] for the 30S subunit are shown in Fig. 29-1D. Most of the results are in agreement with those obtained by other methods. Neutron scattering from the 50S subunit was investigated by using pairs of protonated proteins in a subunit consisting of otherwise deuterated components. This gives an increase in sensitivity.[47]

3. RNA–Protein Interactions and Assembly of Ribosomes

Within bacterial cells the assembly of ribosomes is coupled to rRNA synthesis and requires only 1–2 minutes.[115] In the laboratory both the 30S ribosomal subunits[116] and the 50S subunits[117–121] of *E. coli* can be completely dissociated into individual protein and RNA molecules and can be reconstituted in a functional form. This is true for both natural 16S or 23S RNA or for RNAs prepared by *in vitro* transcription. In these reassembly experiments, which were pioneered by Nomura,[116] it was found that the order of addition of the protein is important. Some proteins bind directly to ribosomal RNA. For example, S4, S7, S8, S15, S17, and S20 bind directly to 16S RNA.[31] Other proteins bind only after one or more proteins have already bound and the RNA has folded properly to form a structural core (Fig. 29-6A). Domains I, II, and III each form an independent RNA-protein assembly. The lower half of domain I of the 16S RNA, from positions ~60–300, is unreactive toward single-stranded probes and may serve as one core for assembly of the ribosome.[67] Protein S20 binds to the 240–286 stem, which is in this core. Protein S4 also binds directly to 16S RNA in the 5' domain. Proteins S8 and S15 bind in the central domain and S7, which is structurally related to the DNA-binding proteins HU and IHF (Chapter 27),[24,122] binds near the 3' end.

5′- Domain of 16S RNA. The 23-kDa protein S4, one of the largest ribosomal proteins, appears to have an important organizing role for the 5' domain.[123,124] It binds in such a way as to protect sequences 27–47 and 394–556 of the RNA (Fig. 29-2) from chemical modification. The small loop at positions 323–330 is protected in the 30S subunit, and the residues A325, A327, A379, and G331 are universally conserved.[67] The same is true of bulge loop 505–510 and the loop sequence 518–533, which contains 7-methylguanine (m^7G) at position 526. Reconstitution experiments also suggested that S16 binds to S4 as well as to S20. Some mutations in proteins S4 and S5 are associated with reduced fidelity of translation, while others lead to spectinomycin resistance.[134]

Central domain of 16S RNA. Proteins S6, S8, S15, and S18 bind to the central domain II of 16S RNA (Fig. 29-2)[31,31a,67,125–129] and organize the platform region (Fig. 29-6). Protein S8 binds with high affinity to regions 588–606 and 632–651 of helix 21 and plays a key role in ribosomal assembly.[126,130,130a] S15 protects residues 655–672 and 734–751 of helix 22. The region contains functionally important conserved loops at positions 570–571, 766–768, and 811–820 as well as many individual adenines in other locations. S15 binds not only to the 16S RNA but also to the 715 loop of 23S RNA in the large subunit and to its own mRNA.[129] A Mg^{2+}-dependent conformational change in the RNA seems to be important in the assembly of the central domain.[131] S6 and S18 bind to 16S RNA after S15 has bound (Fig. 29-6). Proteins S11 and S21 also bind after S15.[129] S11 binds to the 690 loop of the RNA, as is illustrated in Fig. 29-6B. This loop is conserved in all three phylogenetic domains. Located in the platform of the small subunit, it protrudes into the interface to interact with domain IV of the 23S RNA and is also a site of binding of initiation factor IF3.[133] The mutant A649G in 16S RNA confers resistance to **pactamycin** in *E. coli*. Protein S8 is not only an important structural protein in the central domain but also

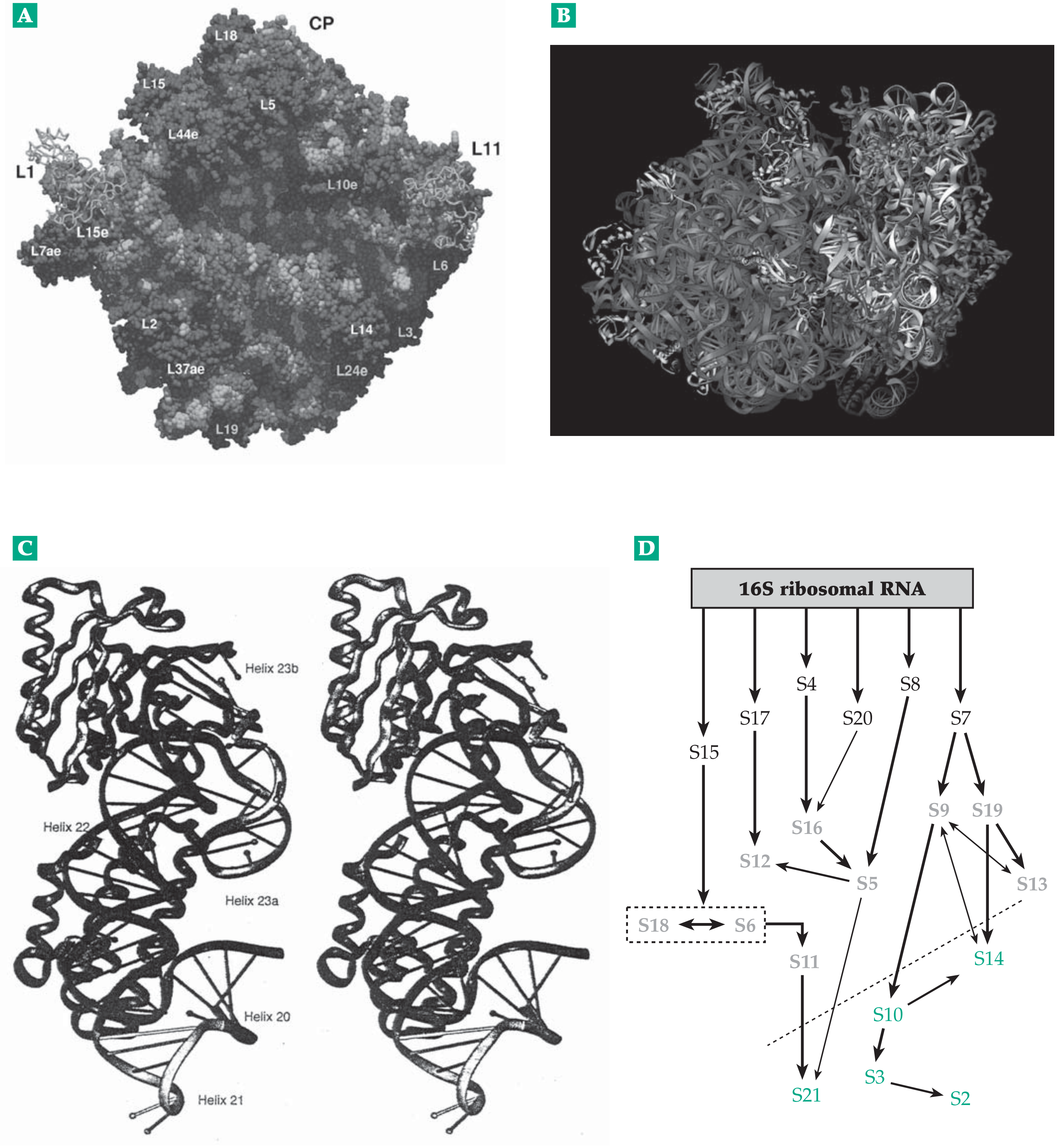
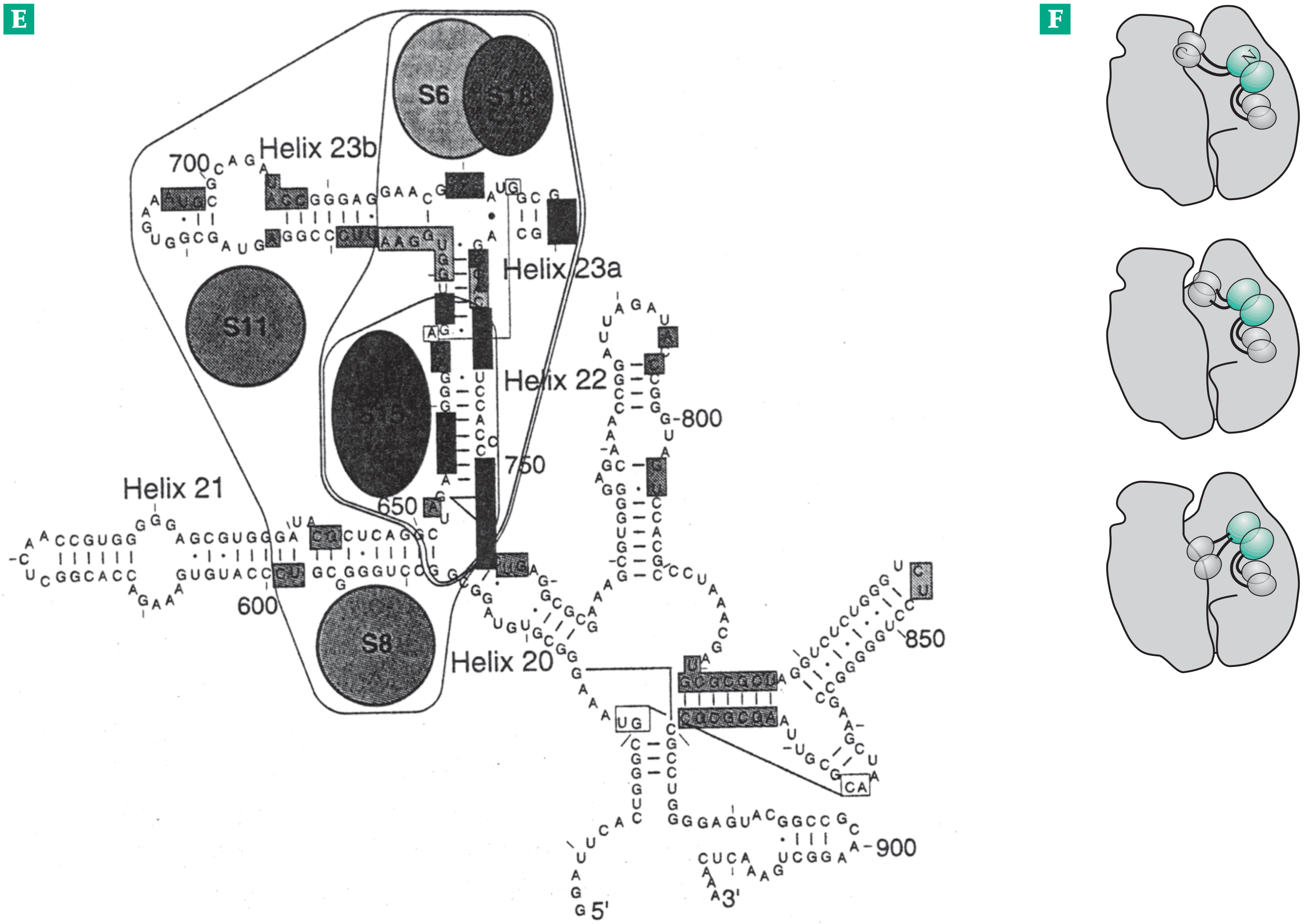
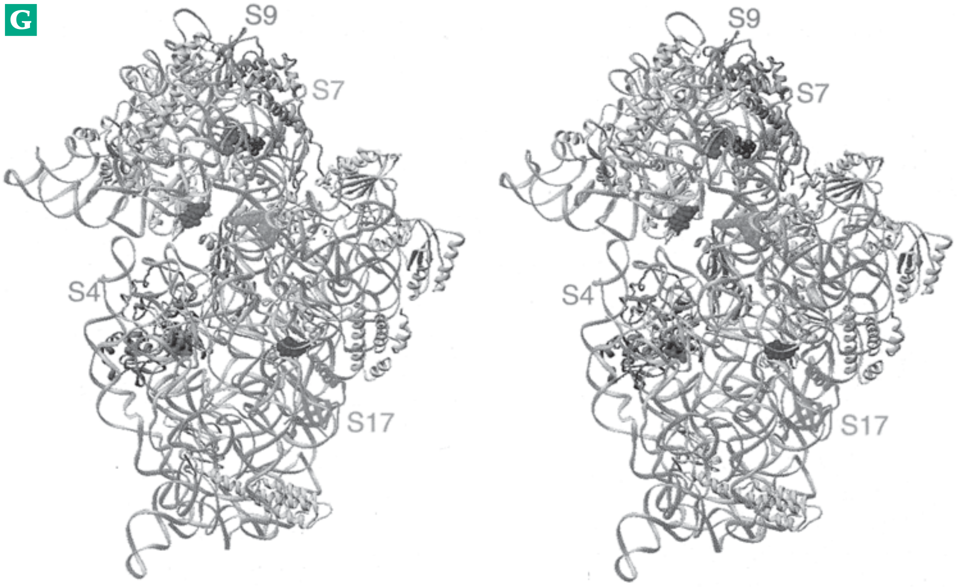

A
CP
L18
L15
L5
L44e
L11
L1
L10e
L15e
L7ae
L6
L2
L14
L3
L37ae
L24e
L19
B
C
Helix 23b
Helix 22
Helix 23a
Helix 20
Helix 21
D
16S ribosomal RNA
S4
S8
S17
S20
S7
S15
S9
S19
S16
S12
S5
S13
S18
S6
S11
S14
S10
S3
S21
S2

Figure 29-6 Some protein–RNA interactions within the ribosome. (A) A space-filling model of the 23S and 5S RNA with associated proteins from the ribosome of *Haloarcula marismortui*. The CCA ends of bound tRNA molecules in the A, P, and E sites are also included. The view is looking into the active site cleft. The proteins with e after the number are related to eukaryotic ribosomal proteins more closely than to those of *E. coli*.[17] Courtesy of T. A. Steitz. (B) Three-dimensional structure of a 70S ribosome from *Thermus thermophilus*. The 30S subunit is to the right of the 50S subunit. Courtesy of Yusupov *et al.*[33a] (C) Stereoscopic view of the helix 21 to helix 23b region of the 16S RNA with associated proteins S6 (upper left), S18 (upper center, front), and S15 (lower back) from *T. thermophilus*. Courtesy of Agalarov *et al.*[31] (D) Simplified *in vitro* assembly map of the central domain of the 30S bacterial ribosome. Courtesy of Gloria Culver. (E) Contacts of proteins with the central (platform) domain of the 16S RNA component. The sequence shown is that of *Thermus thermophilus*. Courtesy of Agalarov *et al.* (F) Three drawings showing alternative location of the four copies of protein L7/L12. The N-terminal and C-terminal


E
F
Helix 23b
700
S6
S18
Helix 23a
S11
Helix 22
S15
800
750
Helix 21
650
600
S8
Helix 20
850
900
5'
3'
C
N


G
S9
S7
S4
S17
S9
S7
S4
S17

globular domains (labeled on one of the four molecules) are connected by a flexible region. One pair of L7/L12 molecules retains a fixed position toward the center of the 50S subunit but the C-terminal domains of the other pair are seen to occupy three different positions. Courtesy of Montesano-Roditis *et al*.[146a] (G) Stereoscopic view of the 30S subunit of the *T. thermophilus* ribosome with six positions at which tetracycline binds and one at which the peptide-like antibiotic **edeine** binds. From Pioletti *et al*.[146b]

acts as a translational repressor of its own gene and of other genes of the spectinomycin-resistance operon (see Box 29-B). This operon encodes genes for ten ribosomal proteins, of both the large and small subunits.[126] S8 is a neighbor to proteins S2, S4, S5, S12, S15, and S17.

3′- Domain of 16S RNA. Domain III of 16S RNA binds proteins S2, S3, S7, S9, S10, S13, S14, and S19. Proteins S3 and S14 may be necessary for assembly of ribosomes but may no longer be needed once the 30S subunit has been correctly formed.[135] The largest of the *E. coli* ribosomal proteins is the 556-residue S1. It does not bind tightly and has sometimes been regarded as nonessential. However, mutations in the S1 gene can be lethal, and the protein seems to be essential for both initiation of translation and for elongation of polypeptide chains *in vivo*.[137] S1 behaves as an elongated molecule 22 nm in length[135] and is unusual in having an ~86-residue RNA-binding sequence repeated four times in the central and N-terminal regions. The protein possesses RNA-unwinding activity[137,138] and may employ these four motifs in unwinding mRNA as it enters the ribosome.

23S and 5S RNAs. Reconstitution of the large ribosomal subunit reveals that proteins L3 and L24 act as assembly initiators.[115,118] L1, L9, L20, and several other proteins (Table 29-2) also bind directly and independently to the 23S RNA. Assembly maps similar to that in Fig. 19-6A have been prepared for the 50S subunit.[117]

One of the most prominent features of the 50S subunit is the L1 protuberance, seen on the left side in Fig. 29-6A. This protuberance is formed almost entirely by protein L1, which is one of the largest ribosomal proteins. It binds to the 2105–2184 loop in domain V of the 23S RNA (see Fig. 29-14).[139] L1 has an important regulatory role in bacteria in which it represses translation of its own structural gene by binding to a region in its mRNA close to the Shine–Dalgarno sequence. The polygenic mRNA also carries the code for protein L11.[139] This is one of several examples of such autogenous regulation of translation of ribosomal proteins.[139a] L1 also interacts in the ribosome with the 5S rRNA.[140] The 272-residue L2 also associates directly with 23S RNA and assists in ribosome assembly.[141] Protein L2 is one of the structurally most highly conserved of the ribosomal proteins.[46,142] It binds to the 1794–1865 region of domain IV of 23S RNA. Histidine 229 of this protein may play a functional role in the ribosome. The protein is elongated, and one end contacts 16S RNA.[46,33a] Protein L9 binds to domain V of 23S RNA in the 2100–2190 region. It is an elongated molecule with two globular α/β domains separated by an α helix. This enables it to bind also to domain III, acting as a rigid strut.[143–145]

On the right side of the 50S subunit, as viewed in Fig. 29-1, is the stalk, a pentameric protein complex consisting of two L7/L12 (*E. coli*) or $(L12)_2$ dimers bound to one molecule of L10.[24,146–147] The stalk is not always seen in X-ray structures, e.g., in Fig. 29-6A, and appears to be flexible. In crosslinking experiments the N-terminal domains of L7/L12 can be linked to L10 and also to its neighbor, L11,[82,148,149] which lies in the GTPase-activating center (Fig. 29-4F) at the base of the stalk. However, the C-terminal domains can be crosslinked to three distinctly different locations: to L11 on the platform surface, to L2 and L5 near the peptidyltransferase center, and to S2, S3, and S14 of the head and neck of the 30S subunit.[146] Domain I of 23S RNA, near the 5' end, binds to protein L20.[153]

An independent and essential structural domain of the ribosome is formed around the 5S RNA.[5,108,154–156] Proteins L5, L18, and L25, whose structure is similar to that of glutaminyl-tRNA synthetase,[154] bind specifically to one loop of the 5S RNA.[156a] Furthermore, the L5-L18-L25-5S RNA complex binds the oligonucleotide TCC. This suggests an interaction between the 5S RNA and the TΨC arm of a tRNA molecule bound to the ribosome. In addition, it has been observed that L18 + either L5 or L25 cause 5S RNA to bind to 23S RNA.

Eukaryotic ribosomal proteins. The functions of the 70–80 different eukaryotic ribosomal proteins are less well known than those of *E. coli*. In eukaryotes the assembly of ribosomes begins in the nucleus with binding of proteins to the individual ribosomal RNA precursors (Chapter 28).[121,156b] Significant functional properties that are peculiar to eukaryotic ribosomal proteins include the following: S6 is the site of multiple phosphorylation reactions, which control initiation of protein synthesis.[132,132a,132b] Mammalian S3 may function in the nucleus in DNA repair.[136] Eukaryotic proteins P0, P1, and P2 are homologous to *E. coli* stalk proteins L10, L7, and L12, respectively. Higher eukaryotes possess only one type of P1 and P2,[150] but yeast,[150a,b] maize,[151] and other species have multiple forms. An L7-related protein is also required for a nucleolar function in ribosomal protein synthesis, perhaps as a component of a snoRNP complex (Chapter 28).[152] Rat liver L37 is involved in peptidyltransferase, but sequencing of the 111-residue protein reveals homology with *E. coli* L34 rather than with L16.[157] Proteins L14, L21, L24, L27, L29, and L30 bind to the 5.8S RNA of the large subunit of yeast ribosomes.[158]

Yeast protein L30, which is not homologous to any bacterial protein, controls its own synthesis by a feedback inhibition at the mRNA splicing step. L30 binds to its own pre-mRNA near the 5' splice site, blocking completion of the spliceosome assembly (Chapter 28).[159]

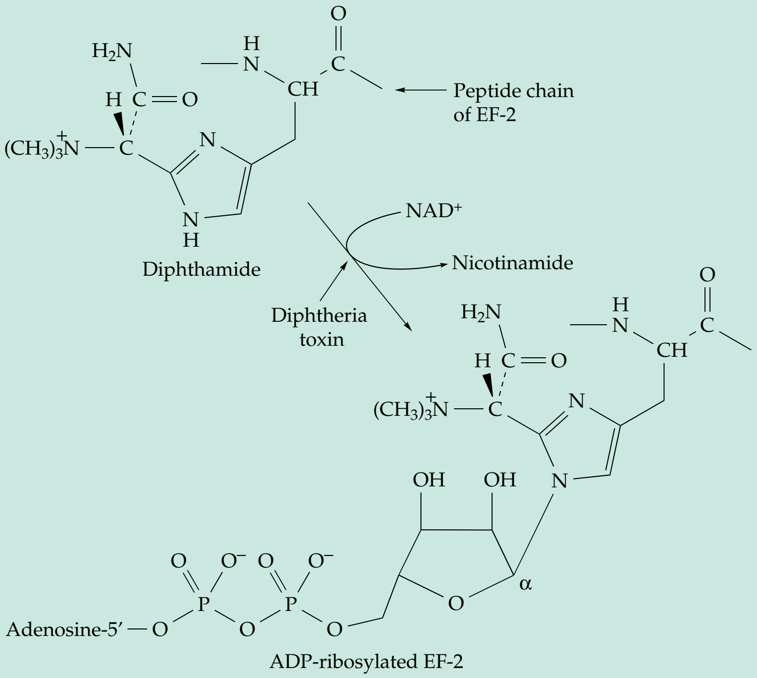
BOX 29-A THE DIPHTHERIA TOXIN AND OTHER RIBOSOME-INACTIVATING PROTEINS

Until a suitable vaccine was developed, an infection by *Corynebacterium diphtheriae* was one of the dread diseases of childhood. Despite the fact that the bacteria caused only superficial membranous lesions in the throat, the patient often died with evident damage to many organs. The cause is a potent heat-labile toxic protein,[a–d] which the bacterium produces when infected by a temperate bacteriophage carrying the *tox* gene and when the inorganic iron of the surroundings has been largely depleted. Diphtheria toxin is a 535-residue protein with a minimum lethal dose (LD_{50}) of only 0.16 mg kg^{-1} for the guinea pig. Tests in cell culture show that the toxin blocks incorporation of amino acids into proteins by inactivation of the eukaryotic elongation factor EF2, which is required for **translocation**, an essential step in protein synthesis in mammalian ribosomes. The toxin acts as an enzyme that transfers (with inversion at the ribose C1) an ADP-ribosyl group from NAD^+ to a side-chain ring nitrogen of the single residue of **diphthamide** in EF2. This modified histidine is found in EF2 and, apparently, in no other protein.[d]

The modified elongation factor reacts normally with GTP, but the complex so formed is unable to participate in translocation. A concentration of the toxin in the cytoplasm of 10^{-8} M is sufficient to promote the fatal reaction. The reaction with diphthamide parallels that of cholera toxin (Box 11-A).

The diphtheria toxin molecule[e,f] consists of three domains, an N-terminal catalytic (C) domain (residues 1–193), a central, largely α-helical, transmembrane (T) domain (residues 205–378), and the C-terminal receptor-binding (R) domain (residues 386–535). Before it enters a cell, the toxin molecule is "nicked" by protease activity between the catalytic and transmembrane domains, a step that is apparently necessary for binding to its receptor, which has been identified as a heparin-binding EGF-like growth factor precursor.[g] The catalytic domain (also called the A fragment) enters the cell through endocytosis from coated pits. Additional proteolytic cleavage, as well as reduction of a disulfide linkage, are required for activity.[f,h] After entering the cytosol the toxin fragment catalyzes inactivation of EF2.[f] A single molecule can kill a cell. The larger 613-residue exotoxin A of *Pseudomonas aeruginosa* catalyzes the same reaction as does diphtheria toxin. It also consists of three domains: a large β-sheet-containing N-terminal region, a central α-helical domain, and a C-terminal domain. The last contains the ADP-ribosyltransferase active site.[a,i]

What is the origin of the *tox* gene, and why is it carried by a virus? Cells do normally contain ADP-ribosyltransferases.[j] The genes for such a protein may have become incorporated into a virus and, after a period of evolution, came to specify the toxic protein.



Another family of toxins attacks ribosomes in a very different way, cleaving ribosomal RNA at specific sites. One of the best known of these is the neurotoxin from *Shigella dysentariae* (Shiga toxin). Like the cholera toxin (Box 11-A) it consists of a single catalytic A subunit and a pentameric ring of B subunits,[k] which binds to specific surface glycolipids.[l] **Verotoxin**, another poison from certain strains of *E. coli*, has a similar structure. Although they have very different effects and there is no detectable similarity in their amino acid sequences, the *structures* of the B pentamers of verotoxin and of the cholera toxin-like heat-labile enterotoxin of *E. coli* are similar.[m,n] The A subunit of Shiga toxin hydrolyzes the *N*-glycosyl linkage of adenine to the ribose ring at position 4324 of 28S ribosomal RNA.[k] A number of plants form very toxic lectins: **ricin** (from castor bean),[o] **viscumin** (from mistletoe), **modeccin**, **abrin**,[p] gelonin,[q] and **volkensin**.[o] The names are derived from the genus names of the plants. All appear to be glycoproteins consisting of two disulfide-linked chains, one of which is a lectin. The

BOX 29-A THE DIPHTHERIA TOXIN AND OTHER RIBOSOME-INACTIVATING PROTEINS (continued)

lectin subunits of ricin[r] and of volkensin[s] bind to galactose residues. The A chains are cytotoxins, which enter cells and, like Shiga toxin, inactivate 60S ribosomal subunits. The 267-residue A chain of ricin is similar to that of a pokeweed viral antigen[t] and of Shiga toxin. It catalyzes the same reaction,[u] the depurination of adenosine 4324. (The pokeweed toxin also catalyzes the corresponding reaction with A2660 of the *E. coli* 23S RNA.[v]) Like the diphtheria toxin these toxic proteins bind to cell surface receptors, are taken up by endocytosis, and are transported through the Golgi to the endoplasmic reticulum. Their structures facilitate uptake but allow them to escape degradation in proteasomes.[w]

Ricin is one of the most toxic substances known. A single molecule can inactivate over 1700 ribosomes per minute and kill the cell.[u] With an LD$_{50}$ of only 1 µg / kg of body weight for many animals, ricin has been used as a poison by assassins. Of more importance is the attempt to couple ricin and related toxins to immunoglobulins to produce **immunotoxins** that will attack cancer cells (Box 31-A). A related goal is to design a potent inhibitor that could serve as an antidote.[x] It is fortunate that most plant seeds do not contain toxins like ricin. Many plants, including such important food grains as wheat and barley, do contain ribosome-inactivating proteins similar to the A chain of ricin. However, the plants lack the B (lectin) subunits and do not enter animal cells.

A group of unusual fungal ribonucleases, which includes **α-sarcin** and **restrictocin**, are produced by *Aspergillus*. The cytotoxic nucleases enter animal cells, where they cut the 28S RNA of ribosomes, specifically on the 3' side of guanosine 4325 in the sarcin / ricin domain (see Fig. 29-4), thereby blocking protein synthesis.[u,y,z] *Staphylococcus aureus* produces a 22-kDa toxic protein thought to be responsible for **toxic shock syndrome**.[aa] Another toxic ribonuclease is **colicin E3** (Box 8-D), which cuts the 16S RNA of *E. coli* after nucleotide 1493 (see Fig. 29-1A).[bb] Colicin D stops protein synthesis by cleavage of four isoaccepting tRNA^{Arg} molecules between positions 38 and 39 in the anticodon loop.[cc]

[a] Han, X. Y., and Galloway, D. R. (1995) *J. Biol. Chem.* **270**, 679–684

[b] Pappenheimer, A. M., Jr. (1977) *Ann. Rev. Biochem.* **46**, 69–94

[c] Pappenheimer, A. M. J. (1993) *Protein Sci.* **2**, 292–298

[d] Ward, W. H. J. (1987) *Trends Biochem. Sci.* **12**, 28–31

[e] Choe, S., Bennett, M. J., Fujii, G., Curmi, P. M. G., Kantardjieff, K. A., Collier, R. J., and Eisenberg, D. (1992) *Nature (London)* **357**, 216–222

[f] Weiss, M. S., Blanke, S. R., Collier, R. J., and Eisenberg, D. (1995) *Biochemistry* **34**, 773–781

[g] Shen, W. H., Choe, S., Eisenberg, D., and Collier, R. J. (1994) *J. Biol. Chem.* **269**, 29077–29084

[h] Blanke, S. R., Huang, K., Wilson, B. A., Papini, E., Covacci, A., and Collier, R. J. (1994) *Biochemistry* **33**, 5155–5161

[i] Douglas, C. M., and Collier, R. J. (1990) *Biochemistry* **29**, 5043–5049

[j] Eide, B., Gierschik, P., and Spiegel, A. (1986) *Biochemistry* **25**, 6711–6715

[k] Kozlov, Y. V., Chernaia, M. M., Fraser, M. E., and James, M. N. G. (1993) *J. Mol. Biol.* **232**, 704–706

[l] Saleh, M. T., and Gariépy, J. (1993) *Biochemistry* **32**, 918–922

[m] Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L., and Read, R. J. (1992) *Nature (London)* **355**, 748–750

[n] Sixma, T. K., Stein, P. E., Hol, W. G. J., and Read, R. J. (1993) *Biochemistry* **32**, 191–198

[o] Lord, J. M., Roberts, L. M., and Robertus, J. D. (1994) *FASEB J.* **8**, 201–208

[p] Tahirov, T. H., Lu, T.-H., Liaw, Y.-C., Chen, Y.-L., and Lin, J.-Y. (1995) *J. Mol. Biol.* **250**, 354–367

[q] Hosur, M. V., Nair, B., Satyamurthy, P., Misquith, S., Surolia, A., and Kannan, K. K. (1995) *J. Mol. Biol.* **250**, 368–380

[r] Weston, S. A., Tucker, A. D., Thatcher, D. R., Derbyshire, D. J., and Pauptit, R. A. (1994) *J. Mol. Biol.* **244**, 410–422

[s] Stirpe, F., Barbieri, L., Abbondanza, A., Falasca, A. I., Brown, A. N. F., Sandvig, K., Olsnes, S., and Pihl, A. (1985) *J. Biol. Chem.* **260**, 14589–14595

[t] Marchant, A., and Hartley, M. R. (1995) *J. Mol. Biol.* **254**, 848–855

[u] Glück, A., and Wool, I. G. (1996) *J. Mol. Biol.* **256**, 838–848

[v] Chan, Y.-L., Sitikov, A. S., and Wool, I. G. (2000) *J. Mol. Biol.* **298**, 795–805

[w] Sandvig, K., and van Deurs, B. (2000) *EMBO J.* **19**, 5943–5950

[x] Yan, X., Hollis, T., Svinth, M., Day, P., Monzingo, A. F., Milne, G. W. A., and Robertus, J. D. (1997) *J. Mol. Biol.* **266**, 1043–1049

[y] Endo, Y., Chan, Y.-L., Lin, A., Tsurugi, K., and Wool, I. G. (1988) *J. Biol. Chem.* **263**, 7917–7920

[z] Nayak, S. K., Bagga, S., Gaur, D., Nair, D. T., Salunke, D. M., and Batra, J. K. (2001) *Biochemistry* **40**, 9115 – 9124

[aa] Blomster-Hautamaa, D. A., Kreiswirth, B. N., Kornblum, J. S., Novick, R. P., and Schlievert, P. M. (1986) *J. Biol. Chem.* **261**, 15783–15786

[bb] Lasater, L. S., Cann, P. A., and Glitz, D. G. (1989) *J. Biol. Chem.* **264**, 21798–21805

[cc] Tomita, K., Ogawa, T., Uozumi, T., Watanabe, K., and Masaki, H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8278–8283

4. Locating Active Sites in Ribosomes

In early studies, antibodies against haptens covalently linked to ends of the 16S RNA were used to locate the 3' end of 16S RNA on the upper “platform” of the 30S subunit (Fig. 29-1A).[5,160] The 5' terminus was found in the lower body. Two N^6,N^6-dimethyladenosines occur at positions 1518 and 1519, about 25 residues from the 3' end of the 16S RNA (Fig. 29-2). Antibodies were used to locate this position on the lower platform. Likewise, the m^{7}G at position 526 lies in the “neck” as shown in Fig. 29-1A.[161] Taking into account known protein–RNA interactions, domain II of the 16S RNA was located in the “platform” on the upper left side of the "body" (viewed from the "outside" as in Fig. 29-1C), while domain III is in the "head." Recent structural studies have confirmed these biochemical localizations.[19,29,33,33a]

The tRNA-binding sites. During protein synthesis tRNA is bound sequentially in at least three places located between the 30S and 50S subunits. These are known as the **A** (**aminoacyl**), **P** (**peptidyl**), and **E** (**exit**) sites. The latter binds deacylated tRNA before it is released from the ribosome.[162,163] Because tRNA is such a large molecule, these sites have subsites in both 30S and 50S ribosomal subunits. When in the P site, a tRNA has its anticodon held firmly and base-paired with a codon in the mRNA in the decoding site of the 30S subunit. The CCA 3'-end with its attached peptidyl chain lies in the 50S subunit at the peptidyltranferase center. A "charged" aminoacyl-tRNA enters the A site, which is close to the 5S RNA in the central protuberance of the large ribosomal subunit, with its anticodon in the decoding site and its aminoacyl group at the peptidyltransferase site.

One end of the P site must be close to the 3' end of the 16S RNA near the two $m_2{}^6A$ residues (Fig. 29-1A). This conclusion, which was based on photochemical linking of a hypermodified base at position 34 (see Fig. 29-7) in $tRNA^{Val}$ with C-1400 of 16S RNA by cyclobutane dimer formation (Eq. 23-26),[164] has been confirmed by structural studies.[29,33,33d] Investigation of tRNA binding, effects of mutations in ribosomal RNA, and effects of antibiotics pointed to locations of the P and A sites in both ribosomal subunits. These have been located precisely by crystallography. See Fig. 29-1F,J; 29-2C. Residue 6530 together with nucleotides 921–927, 1390–1394, and 1491–1505 of 16S RNA participate in forming the form A and P sites in the decoding center.[33c,d,164a,b,378] The two adjacent adenine rings of A1492 and A1493 swing out from helix 44 (Fig. 29-2; see also Fig. 29-14) to form a major part of the A site. In the 50S subunit the adjacent cytosines C74 and C75 of the CCA 3'-ends of the tRNAs in the A and P sites interact respectively with G2553 of the A loop and G2252 of the P loop (Fig. 29-14B,E).[164b–e] Tetracycline (Fig. 22-7) also binds into the A site (see Box 29-B). It can be photochemically crosslinked to proteins S18 and S4.[167]

The peptidyltransferase site. The position was located by binding of derivatives of the antibiotic **puromycin** (Fig. 29-13). An arylazide derivative of puromycin was photochemically linked (Eq. 23-27) to proteins L23, L18/22, and L15; immunoelectron microscopy, using antibodies to the N^6-dimethyladenosine of puromycin,[165,166] located the binding site adjacent to the central protuberance between the 50S subunit and 30S subunit near S14.[5] 4-Thio-dT-p-C-p-puromycin was photochemically crosslinked to G2553 of the peptidyltransferase A site (see Fig. 29-14). X-ray data provided a precise structure of the peptidyltransferase site (see pp. 1702–1704).[166a] Studies of mutant ribosomes together with affinity labeling and crosslinking experiments pointed to the **peptidyltransferase loop** marked on Fig. 29-4 and further illustrated in Fig. 29-14.[164a,167a,b]

The GTPase-activating center. Also shown in Fig. 29-1C is a site for binding of the elongation factor EF-G (Section C,2). This was located, in part, because the antibiotic **thiostrepton** prevents EF-G from binding to the ribosome. Thiostrepton binds to a complex of protein L11 and a 61-fragment of the 23S RNA (positions 1052–1112; see Fig. 29-4F).[80] Another elongation factor, EF-Tu, also binds at the same site or adjacent to the EF-G site by the head of the small subunit.[168] An additional location of interest is the **polypeptide exit tunnel**, which brings the growing protein chain from the peptidyltransferase site out of the ribosome (Fig. 29-1I).[5,6,33f]

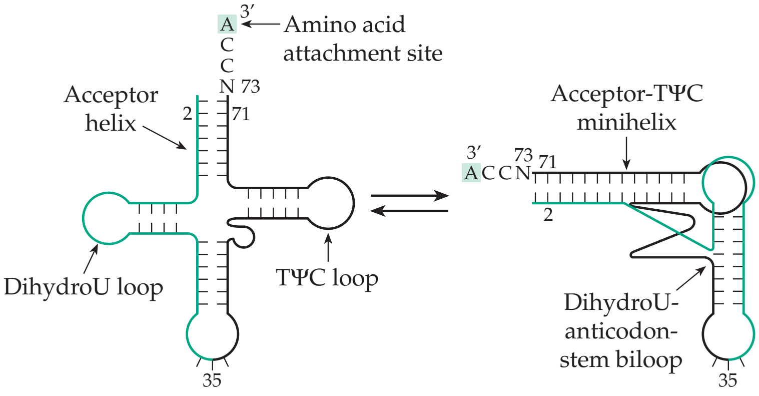
B. Transfer RNAs

The small 4S tRNA molecules have masses of ~26 kDa and consist of 75 ± 5 nucleotides (Figs. 5-30, 5-31, and 29-6). The basic structures are similar in bacteria and eukaryotic cells. The need for "adapters," to carry amino acids to the proper positions along the mRNA template, had been predicted prior to the discovery of tRNA.[169,170] It had been expected that there would be a base sequence constituting an **anticodon**, which would fit against the proper codon at some binding site on the protein-synthesizing machinery. This is just what tRNA molecules do, but their chemistry contained many surprises.

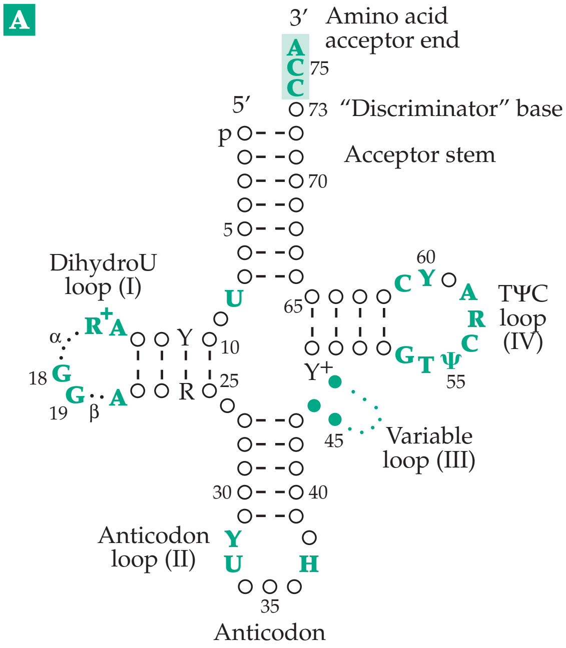
1. Structures of Transfer RNAs

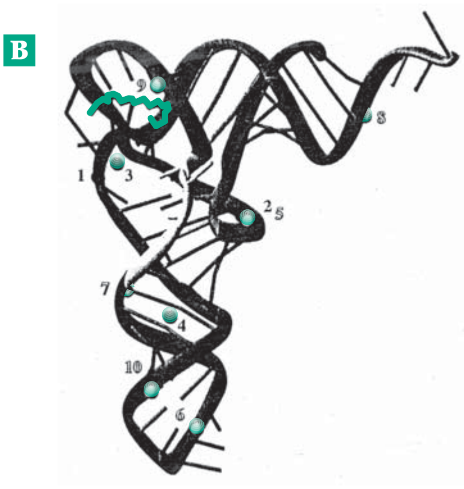
The first surprise was that these molecules are much longer than seems necessary for the formation of adapters. In addition, 10–20% of their bases are modified greatly from their original form.[171] Another surprise was that the anticodons are not all made up of "standard" bases. Thus, hypoxanthine (whose nucleoside is inosine) occurs in some anticodons. Conventional "cloverleaf" representations of tRNA, which display their secondary structures, are shown in Figs. 5-30 and 29-7. However, the molecules usually have an L shape rather than a cloverleaf form (Figs. 5-31 and 29-6),[172] and the L form is essential for functioning in protein synthesis as indicated by X-ray and other data.[173] Three-dimensional structures, now determined for several different tRNAs,[174,175] are all very similar. Structures in solution are also thought to be similar for the various tRNA molecules.[176,177] One of the four hydrogen-bonded “stems” of a tRNA in the cloverleaf form terminates in the universally conserved CCA-3′ **amino acid acceptor end** (Fig. 29-7), which can carry an esterified amino acid generated as in Eq. 29-1, steps *a*, *b*. The other three stems terminate in loops, which usually contain a large number of modified bases. The modifications may serve to optimize the interaction of the tRNA with other components of the protein-synthesizing machinery.[178] The **dihydroU loop** (loop I) contains 5,6-dihydrouridine in various amounts and in varying positions. The **anticodon loop** (loop II) always contains the anticodon directly opposite the amino acid acceptor end in the cloverleaf drawing. On the 5' side of the anticodon at position 33 there is almost always a U (shaded in Fig. 29-7A) preceded by another pyrimidine. A hydrogen bond from the N3 proton of U-33 and a phosphate oxygen of residue 36 stabilize the U-turn that precedes the anticodon triplet (Fig. 5-31).[177] Next to the 3' side of the anticodon there is usually a **hypermodified** base, such as N^6-(Δ^2-isopentenyl) adenosine (Fig. 5-33) or a more complex derivative.[178] The **variable loop** (loop III) can range between 5 and 21 nt in length.[177,181] The **TψC loop** (loop IV) contains the specific nucleotide sequence for which the loop is named.

Cloverleaf and L forms. Interconversion between the cloverleaf and L forms of tRNA molecules can be pictured as in Eq. 29-2. Notice that in the L


3′
A
C
C
N 73
Amino acid attachment site
Acceptor helix
2
71
DihydroU loop
TΨC loop
35
Anticodon
Acceptor-TΨC minihelix
3′
ACCN
73
71
2
DihydroU-anticodon-stem biloop
35

(29-2)


A
3′ Amino acid acceptor end
A
C 75
C
5′
73 “Discriminator” base
p
Acceptor stem
70
5
60
DihydroU loop (I)
C Y A
TΨC loop (IV)
U
65
R
α
R A
Y
10
C
18 G
G T Ψ
Y+
55
25
G
A
R
19 β
45
Variable loop (III)
30
40
Anticodon loop (II)
Y
U
H
35
Anticodon


B
9
8
1
3
2
5
7
4
10
6

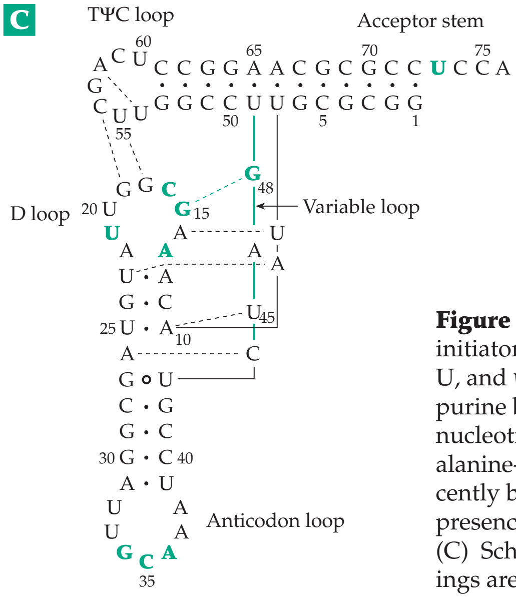
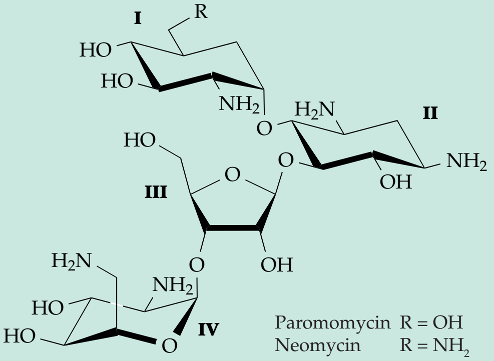

C
TΨC loop
Acceptor stem
60
65
70
75
A C U C C G G A A C G C G C C U C C A
G
C U U G G C C U U G C G C G G
55
50
5
1
G G C
G
48
D loop
20 U
G 15
Variable loop
U
A
U
A A
A
A
U A
G C
25 U A
U 45
10
A
C
G U
C G
G C
30 G C 40
A U
U A
U A
Anticodon loop
G C A
35

Figure 29-7 (A) Generalized cloverleaf diagram of all tRNA sequences except for initiator tRNAs numbered as in yeast $tRNA^{Phe}$ (Fig. 5-30). Invariant bases: A, C, G, T, U, and ψ; semivariant bases: Y (pyrimidine base), R (purine base), H (hypermodified purine base). The dotted regions (α, β, variable loop) contain different numbers of nucleotides in various tRNA sequences. See Rich.[179] (B) L form of the yeast phenylalanine-specific $tRNA^{Phe}$. The structure is the same as that in Fig. 5-31 but has recently been redetermined at a resolution of 0.20 nm.[175] The new data revealed the presence of ten bound Mg^{2+} ions (green circles) as well as bound spermine (green). (C) Schematic representation of L form of *E. coli* $tRNA^{Cys}$. Some tertiary base pairings are indicated by dashed lines. No modified bases are shown. See Hou *et al.*[180]

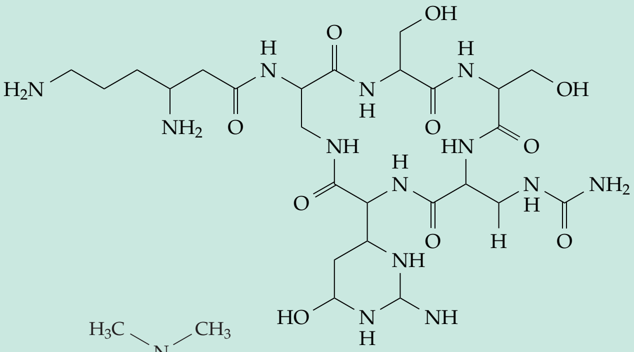
BOX 29-B ANTIBIOTICS THAT INHIBIT PROTEIN SYNTHESIS

Many of our most effective antibiotics act by blocking protein synthesis on ribosomes.[a–de] The usefulness of many of these remarkable drugs in human medicine depends upon the fact that they inhibit protein synthesis on bacterial 70S ribosomes but do not seriously affect eukaryotic ribosomes. Antibiotics act at a variety of sites involving both ribosomal proteins and rRNA. Some of the resulting points of inhibition of the ribosomal elongation cycle[e] are marked on Fig. 29-12. Binding sites for many antibiotics have been located by footprinting experiments, by studies of mutants, and more recently by X-ray crystallography. Exposure of bacteria to antibiotics often gives rise to antibiotic-resistant mutants. In some of these mutants a specific ribosomal protein has been altered, but in others a specific RNA base has been changed.

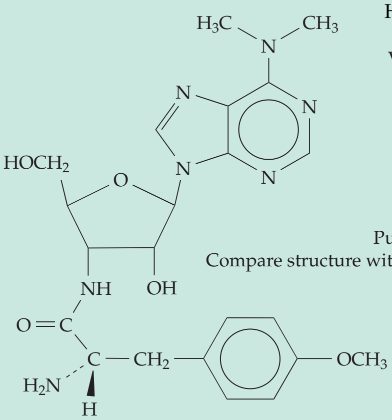
One of the first antibiotics to be studied was **puromycin** (Fig. 29-13), which binds to the 50S subunit and causes premature termination of peptide synthesis. A glance at its structure reveals how it can do this. It resembles in fine detail the 3' end of a tRNA molecule bearing an aminoacyl group. However, it is not an aminoacyl group, and once the growing peptide chain has been transferred onto the puromycin, further chain elongation is impossible. It was shown in 1975 that puromycin could be crosslinked to several different proteins in the ribosome by ultraviolet irradiation.[f,g] L23 of the 50S subunit and S14 of the 30S subunit were labeled most heavily. More recently puromycin, acting on 70S ribosomes, has been shown to label protein S7 and several large subunit proteins near the peptidyltransferase center in the central domain V of the 23S RNA.[h] A puromycin-derived transition state inhibitor has permitted precise identification of the peptidyltransferase site (Fig. 29-13).[i,j] In earlier work immunoelectron microscopy on *N*-bromoacetylpuromycin-labeled 50S subunits had located the site marked Pm in Fig. 29-1A. However, this is quite far from the recently determined location of the peptidyltransferase center,[k]



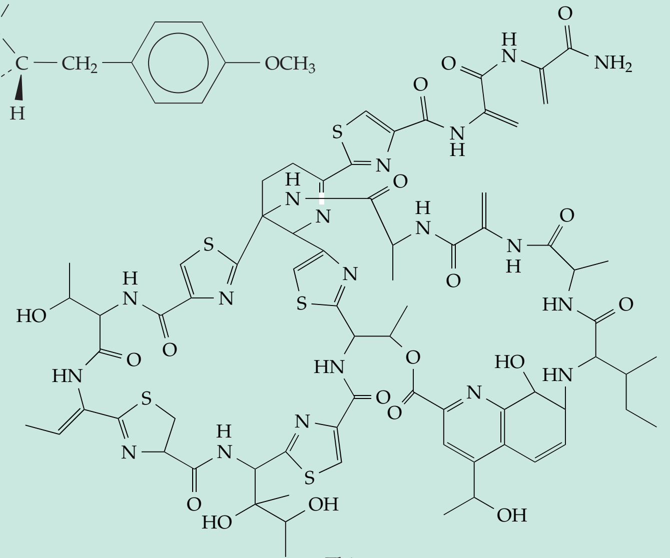
Paromomycin R = OH
Neomycin R = NH_2



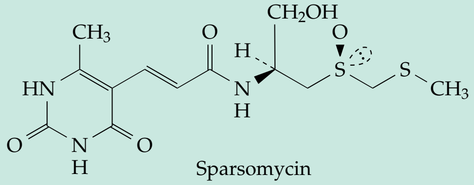
Viomycin



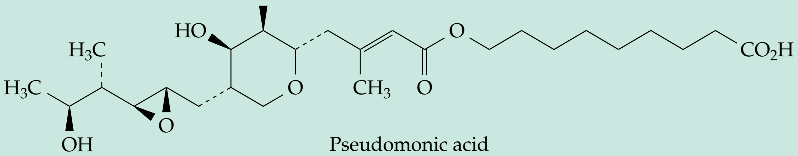
Puromycin
Compare structure with that of an aminoacyl-tRNA



Thiostrepton



Sparsomycin



Pseudomonic acid

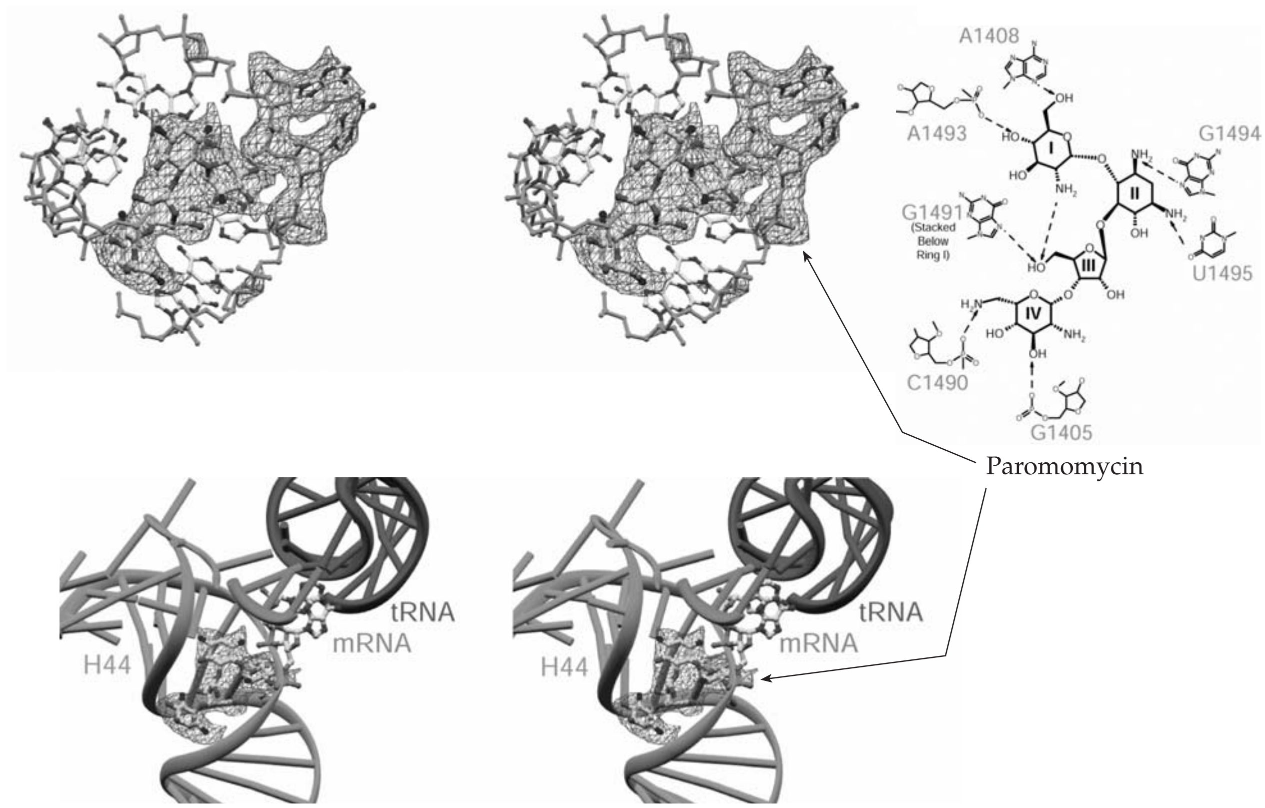
BOX 29-B ANTIBIOTICS THAT INHIBIT PROTEIN SYNTHESIS (continued)

illustrating the difficulty in use of active-site labeling techniques.

A large number of other antibiotics also bind in the vicinity of the peptidyltransferase center (Fig. 29-14).[l,m] Among them are macrolide antibiotics such as erythromycin (Fig. 21-1)[n] and spiramycin,[o,p] chloramphenicol (Fig. 25-10), griseofulvin (Fig. 21-10), streptogramins,[q] oxazolidinones such as linezolid,[r] sparsomycin,[s] and lincomycin. Erythromycin has been very useful in locating the peptidyltransferase center. For example, 23S RNA mutant G2057A, in which A has replaced the normal G, and mutants G2058G and G2058U are resistant to erythromycin.[l,n,t] Mutant C2518U in *Halobacterium halobium* 23S RNA (C2499 in *E. coli*) is among mutants resistant to sparsomycin.[s] Chloramphenicol not only blocks the peptidyltransferase but also causes an accumulation of the compound ppGpp (p. 1715).

The aminoglycoside antibiotics **streptomycin** (Box 20-B),[c,d,u,v] the neomycins,[w] **paromomycin** (see drawing below),[c,x–z] gentamycin,[aa] and kanamycin have one structural unit in common. They often bind to 16S ribosomal RNA in the decoding center. However, they bind in distinctly different ways. Streptomycin causes ribosomes to misread the genetic code[bb,cc] primarily at the first base of a codon. Thus, when poly(U) serves as a messenger RNA, the expected polyphenylalanine product contains 40% isoleucine. When a bacterial population is subjected to the action of any antibiotic, a few mutants are able to grow and survive in the presence of the antibiotic. Streptomycin-resistant mutants of *E. coli* arise at the very low frequency of $\sim 10^{-12}$. One of the genes affected (*rpsL*) was mapped at 72 min. Subsequently, it was shown that streptomycin binds to ribosomal protein S12, and that *rpsL* is the gene for this protein. Mutations in the universally conserved 2660 loop of 23S RNA in the sarcin/ricin domain lead to blockage of the elongation cycle. Bacteria containing both a G2661C mutation in their 23S RNA and also a streptomycin resistance mutation in protein S12 lose efficiency in the action of EF-Tu and die. However, they survive in the presence of streptomycin.[dd]

Streptomycin can also be chemically crosslinked to 16S RNA,[ee] and several aminoglycoside antibiotics including streptomycin and spectinomycin bind



Figures are from Carter *et al.*[c]

BOX 29-B (continued)

to and, in footprinting experiments, protect specific nucleotides in 16S RNA.[ff] Streptomycin binds tightly to the upper part of helix 44 of bacterial 16S RNA (Fig. 29-2). This part of the helix is in the region that binds messenger RNA. It also contains parts of the A and P sites of the decoding region of the 30S ribosomal subunit. The same antibiotic binds less tightly to the 915 region in the center of the 16S RNA.[gg,hh] As mentioned above, some streptomycin-resistant mutants become dependent upon the antibiotic and will not grow in its absence. This streptomycin dependence sometimes results from modification in ribosomal protein S4, but the dependence can be suppressed by specific mutations in S5.[ii] It is clear that a single point mutation altering one amino acid is all that is necessary to enormously change the sensitivity of a living organism to a particular toxin, or even to make the organism dependent upon that toxin. Paromomycin also binds to the upper end of helix 44 in the major groove of the RNA (see figure in this box) and close to the streptomycin-binding site.[c,z,jj] Binding distorts the structure of the bulge loop containing adenosines A1492 and A1493, which are markers for the A site of the decoding region. Messenger RNA is also bound at this site as shown in the accompanying figure. Gentamycin also binds in the A site.[aa]

Hygromycin B also binds at the very top of helix 44 blocking the translocation step in the ribosomal cycle.[kk] **Spectinomycin** binds not only to RNA but also to protein S5, as indicated by analysis of resistant mutants. The S5 structural gene *spcA* maps at 64 min, a position in a ribosomal protein operon of the *E. coli* chromosome. The 16S RNA binding site at one end of helix 34 (with protection of G1064 and C1192) is adjacent to S5 as shown by X-ray structural analysis and directed hydroxyl radical probing (see Fig. 29-2).[c] The antibiotic also interferes with the translocation step of polypeptide elongation. **Kasugamycin** inhibits the binding of fMet-tRNA (initiation). In this case, resistant mutants appear in which it is not a protein subunit that has been modified but the 16S RNA. In resistant strains there is less methylation of adenosines 1518 and 1519 (Fig. 29-2) than in normal strains.[ll]

The **tetracyclines** (Fig. 21-10) inhibit the binding of aminoacyl-tRNA at the A site in the 30S ribosomal subunit.[kk] However, this doesn't appear to be a direct effect. Tetracyclines bind to the 16S RNA at two sites. A major site is on helix 34 near the spectinomycin site in the platform region. A second site is on helix 27, the switch helix, which plays a direct role in translocation (see Eq. 29-9).[kk] Although the basis of the inhibition is not clear, there are distinct differences in binding to the 16S bacterial and 18S eukaryotic RNAs that explain the high specificity of the antibiotic toward bacteria.

Another site of antibiotic action is the GTPase-activating center. This center contains a double hairpin structure in the 23S RNA, which binds to protein L11 and the L10•(L12)$_4$ stalk complex. Several proteins, including initiation factor IF1 and the elongation factors EF-Tu, EF-G (bacteria)/EF-2 (eukaryotes), bind to this part of the 50S ribosome (Fig. 29-4). The thiopeptide antibiotics **thiostrepton**, **micrococcin**, and siomycin[t,mm,nn] also bind in this region. Thiostrepton acts by preventing association with the ribosome of an incoming aminoacyl-tRNA as the EF-Tu•GTP complex. Its binding site is primarily in the 23S RNA, but it probably interferes with peptide elongation by interfering with a conformational change in protein L11.[t] A related cyclic peptide (GE2270A) binds to EF-Tu•GDP competing for binding of an aminoacyl-tRNA and blocking the GDP-GTP exchange.[nn] Certain mutations in the EF-Tu protein confer resistance to this antibiotic.[oo] In a similar way kirromycin prevents release of EF-Tu from the ribosome after GTP hydrolysis.[nn,pp,qq]

The binding site of initiation factor IF1 involves both the 30S and 50S ribosomal subunits. The large oligosaccharide antibiotic **evenimicin** protects a specific set of nucleotides in two loops near the peptidyltransferase center (Fig. 29-14).[rr] Erythromycin,[b] other macrolide antibiotics, cycloheximide (Fig. 21-10), and fusidic acid (p. 1266) all prevent translocation by stabilizing the pre-translocation complex.[ss,tt] Fusidic acid may bind to EF-G on the ribosome, preventing an essential conformational change in this G protein.[tt] Fusidic acid also inhibits accumulation of ppGpp. Figure 29-14 shows the locations of some mutations in *E. coli* 23S rRNA that confer resistance to erythromycins and chloramphenicol. Notice that both domains II and V are involved.[n] **Pactamycin** binds to helices 23b and 24a, a binding site for initiation factor IF3.[kk]

Many antibiotics, which inhibit protein synthesis, do not bind to ribosomes but block any of a variety of vital chemical processes needed for growth. Among them are **pseudomonic acid**, which inhibits isoleucyl-tRNA synthetase from many gram-positive bacteria.[uu,vv] **Rapamycin**, best known as an immunosuppressant (Box 9-F), inhibits phosphoinositide-3-kinase and also phosphorylation of the cap-binding protein 4G, a component of the eukaryotic initiation factor complex (Fig. 29-11).[ww] The bacterial enzyme peptide deformylase, which is absent from the human body, has been suggested as a target for design of synthetic antibiotics.[xx]

BOX 29-B ANTIBIOTICS THAT INHIBIT PROTEIN SYNTHESIS (continued)

[a] Gorini, L. (1966) *Sci. Am.* **214**(Apr), 102–109

[b] Narayanan, C. S., and Dubnau, D. (1987) *J. Biol. Chem.* **262**, 1766–1771

[c] Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000) *Nature (London)* **407**, 340–348

[d] Schroeder, R., Waldsich, C., and Wank, H. (2000) *EMBO J.* **19**, 1–9

[de] Schlünzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001) *Nature (London)* **413**, 814 –821

[e] Nierhaus, K. H. (1990) *Biochemistry* **29**, 4997–5008

[f] Olson, H. M., Nicholson, A. W., Cooperman, B. S., and Glitz, D. G. (1985) *J. Biol. Chem.* **260**, 10326–10331

[g] Weitzmann, C. J., and Cooperman, B. S. (1990) *Biochemistry* **29**, 3458–3465

[h] Bischof, O., Kruft, V., and Wittmann-Liebold, B. (1994) *J. Biol. Chem.* **269**, 18315–18319

[i] Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 920–930

[j] Welch, M., Chastang, J., and Yarus, M. (1995) *Biochemistry* **34**, 385–390

[k] Lührmann, R., Bald, R., Stöffler-Meilicke, M., and Stöffler, G. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7276–7280

[l] Garrett, R. (1983) *Trends Biochem. Sci.* **8**, 189–190

[m] Rodriguez-Fonseca, C., Amils, R., and Garrett, R. A. (1995) *J. Mol. Biol.* **247**, 224–235

[n] Douthwaite, S., and Aagaard, C. (1993) *J. Mol. Biol.* **232**, 725–731

[o] Bischof, O., Urlaub, H., Kruft, V., and Wittmann-Liebold, B. (1995) *J. Biol. Chem.* **270**, 23060–23064

[p] Poulsen, S. M., Kofoed, C., and Vester, B. (2000) *J. Mol. Biol.* **304**, 471–481

[q] Porse, B. T., and Garrett, R. A. (1999) *J. Mol. Biol.* **286**, 375–387

[r] Kloss, P., Xiong, L., Shinabarger, D. L., and Mankin, A. S. (1999) *J. Mol. Biol.* **294**, 93–101

[s] Tan, G. T., DeBlasio, A., and Mankin, A. S. (1996) *J. Mol. Biol.* **261**, 222–230

[t] Porse, B. T., Leviev, I., Mankin, A. S., and Garrett, R. A. (1998) *J. Mol. Biol.* **276**, 391–404

[u] Davies, C., Bussiere, D. E., Golden, B. L., Porter, S. J., Ramakrishnan, V., and White, S. W. (1998) *J. Mol. Biol.* **279**, 873–888

[v] Recht, M. I., Douthwaite, S., and Puglisi, J. D. (1999) *EMBO J.* **18**, 3133–3138

[w] Fourmy, D., Recht, M. I., and Puglisi, J. D. (1998) *J. Mol. Biol.* **277**, 347–362

[x] Fourmy, D., Recht, M. I., Blanchard, S. C., and Puglisi, J. D. (1996) *Science* **274**, 1367–1371

[y] Recht, M. I., Douthwaite, S., Dahlquist, K. D., and Puglisi, J. D. (1999) *J. Mol. Biol.* **286**, 33–43

[z] Lynch, S. R., and Puglisi, J. D. (2001) *J. Mol. Biol.* **306**, 1037–1058

[aa] Yoshizawa, S., Fourmy, D., and Puglisi, J. D. (1998) *EMBO J.* **17**, 6437–6448

[bb] Browning, K. S., Maia, D. M., Lax, S. R., and Ravel, J. M. (1987) *J. Biol. Chem.* **262**, 538–541

[cc] Tai, P.-C., Wallace, B. J., and Davis, B. D. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 275–279

[dd] Bilgin, N., and Ehrenberg, M. (1994) *J. Mol. Biol.* **235**, 813–824

[ee] Gravel, M., Melancon, P., and Brakier-Gingras, L. (1987) *Biochemistry* **26**, 6227–6232

[ff] Moazed, D., and Noller, H. F. (1987) *Nature (London)* **327**, 389–394

[gg] Pinard, R., Payant, C., Melancon, P., and Brakier-Gingras, L. (1993) *FASEB J.* **7**, 173–176

[hh] Spickler, C., Brunelle, M.-N., and Brakier-Gingras, L. (1997) *J. Mol. Biol.* **273**, 586–599

[ii] Culver, G. M., Heilek, G. M., and Noller, H. F. (1999) *J. Mol. Biol.* **286**, 355–364

[jj] VanLoock, M. S., Easterwood, T. R., and Harvey, S. C. (1999) *J. Mol. Biol.* **285**, 2069–2078

[kk] Brodersen, D. E., Clemons, W. M., Jr., Carter, A. P., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000) *Cell* **103**, 1143–1154

[ll] Vila-Sanjurjo, A., Squires, C. L., and Dahlberg, A. E. (1999) *J. Mol. Biol.* **293**, 1–8

[mm] Porse, B. T., Cundliffe, E., and Garrett, R. A. (1999) *J. Mol. Biol.* **287**, 33–45

[nn] Heffron, S. E., and Jurnak, F. (2000) *Biochemistry* **39**, 37–45

[oo] Zuurmond, A.-M., de Graff, J. M., Olsthoorn-Tieleman, L. N., van Duyl, B. Y., Mörhle, V. G., Jurnak, F., Mesters, J. R., Hilgenfeld, R., and Kraal, B. (2000) *J. Mol. Biol.* **304**, 995–1005

[pp] Mesters, J. R., Zeef, L. A. H., Hilgenfeld, R., de Graaf, J. M., Kraal, B., and Bosch, L. (1994) *EMBO J.* **13**, 4877–4885

[qq] Alexander, C., Bilgin, N., Lindschau, C., Mesters, J. R., Kraal, B., Hilgenfeld, R., Erdmann, V. A., and Lippmann, C. (1995) *J. Biol. Chem.* **270**, 14541–14547

[rr] Belova, L., Tenson, T., Xiong, L., McNicholas, P. M., and Mankin, A. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3726–3731

[ss] Johanson, U., AEvarsson, A., Liljas, A., and Hughes, D. (1996) *J. Mol. Biol.* **258**, 420–432

[tt] Laurberg, M., Kristensen, O., Martemyanov, K., Gudkov, A. T., Nagaev, I., Hughes, D., and Liljas, A. (2000) *J. Mol. Biol.* **303**, 593–603

[uu] Yanagisawa, T., Lee, J. T., Wu, H. C., and Kawakami, M. (1994) *J. Biol. Chem.* **269**, 24304–24309

[vv] Sassanfar, M., Kranz, J. E., Gallant, P., Schimmel, P., and Shiba, K. (1996) *Biochemistry* **35**, 9995–10003

[ww] Raught, B., Gingras, A.-C., Gygi, S. P., Imataka, H., Morino, S., Gradi, A., Aebersold, R., and Sonenberg, N. (2000) *EMBO J.* **19**, 434–444

[xx] Hao, B., Gong, W., Rajagopalan, P. T. R., Zhou, Y., Pei, D., and Chan, M. K. (1999) *Biochemistry* **38**, 4712–4719

form the acceptor stem and the TψC arm form a single **acceptor-TψC-minihelix**, while the other two domains fold together to create an **anticodon-dihydrouridine** stem loop.[172] New tertiary interactions, some of which are indicated in Fig. 29-7C, are formed. Mitochondrial tRNAs of metazoa often lack some elements of the cloverleaf. An extreme example is the bovine mtRNA^{Ser}, which recognizes AGY codons and completely lacks the dihydrouridine loop. This fact suggests that the L shape of tRNAs cannot be completely invariant.[174,176] As is shown in Fig. 29-7B, divalent metal ions such as Mg^{2+} are bound at discrete sites in tRNA molecules.[171] The tertiary interactions in the "core" of the L form contains several stacked layers of base pairs and triplets (Fig. 29-7C). The top layer is usually the single base 59; below it in succession are the 15:48 pair (see Fig. 29-7), the 21:8:14 triplet, the 13:22 pair (see also p. 231), and then base pairs present in the dihydroU loop. Considerable variation is observed among the different tRNAs.[182,183] The structural features of this core may also be utilized for recognition by aminoacyl tRNA synthetases.

Initiator tRNAs. While the TψC sequence has been found in all bacterial and most eukaryotic tRNAs examined, it is replaced by UCG in eukaryotic initiator tRNAs. In these tRNAs the preceding two nucleotides, beginning in the stem of loop IV, are also conserved; the complete conserved sequence being GAUCG.[184] Other characteristics of initiator tRNAs are the absence of base-pairing between residues 1, and 72, and the presence of C rather than G at position 1, A rather than G at position 72, and CCU in place of the two dihydroU residues in loop I.[185] Initiator tRNAs of chloroplasts resemble those of bacteria,[186] whereas archaeobacteria have their own unique peculiarities.[187] These include the presence of a hypermodified base known as **archaeosine** (p. 1456) in position 15 of the dihydroU loop.[188,189]

2. Pairing of Codon and Anticodon

Accurate protein synthesis depends upon both correct charging of the tRNAs and correct recognition by an anticodon in the tRNA of the complementary codon in the mRNA. A surprise was the discovery of inosine (I) in anticodons of yeast tRNA (but not in most *E. coli* tRNAs). Another unexpected finding was that fewer than 61 kinds of tRNA exist in a given cell (61 = 64 codons minus three stop codons). Consideration of these matters led Crick, in 1966, to propose the **wobble hypothesis**.[190] According to this proposal the first two bases at the 5' end of the codon (and at the 3' end of the anticodon) must pair in the same ways as do the bases in DNA. However, the third base pair (3' end of the codon and 5' end of the anticodon) is under a less severe steric restriction. That is, there may be some “wobble.” Crick suggested the accompanying rule for pairing of the third base. All of the observed deviations from the AU, CG pairing of a Watson–Crick helix can be explained in this way. An anticodon with G at the 5' end can pair with codons with either C or U at their 3′ end. Anticodons with C or A at the 5′ end pair strictly. Anticodons ending with U at the 5′ end can pair with codons containing either A or G in the 3' position. Anticodons with I in the 5' position can recognize codons with any of the three bases in the third position. Comparison with Table 5-5 makes it immediately clear why fewer than 61 anticodons are needed. Many codons represent the same amino acid, and frequently the nature of the base in the 3' position of the codon is immaterial to the meaning of that codon. Thus, there is an economy in using less than the full array of anticodons. Crick showed that his proposal was chemically feasible if the spatial relationships for the wobble pair were allowed to vary from the usual ones in Watson–Crick base pairs. This is illustrated in Fig. 29-8 for binding of inosine to C (a normal Watson–Crick base pair) and to A and to U. Although the word wobble does not convey an exact meaning, the hypothesis has predicted many things correctly. For example, only three tRNAs are required to recognize the six serine codons. In fact, only three are found in *E. coli*.

5′-Base in anticodon	Paired 3′-base in codon
G	C or U
C	G
A	U
U	A or G
I	C, A, or U

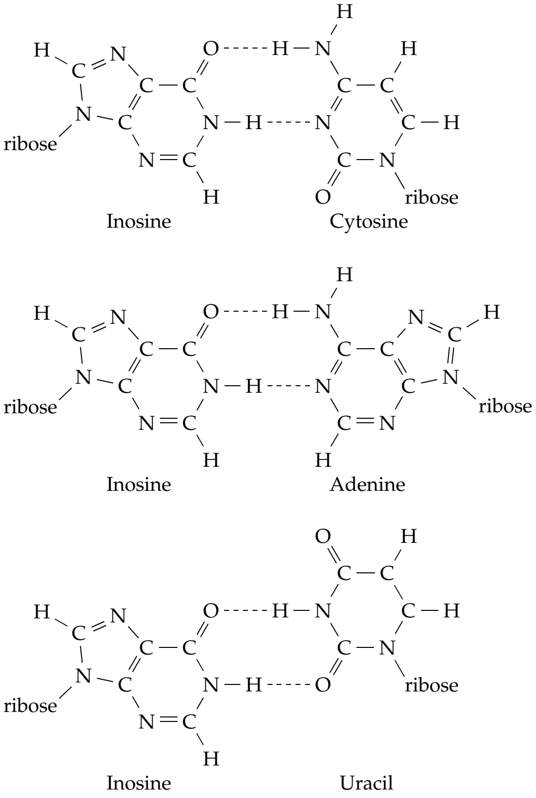


Figure 29-8 Pairing of inosine with cytosine (a Watson–Crick pair) and of inosine with adenine and uracil (wobble pairs).

The matter is made more complex by the fact that U34 in the first (5') anticodon position of tRNA is usually modified postranscriptionally.[190a] For example, it is usually converted to the 2-thio-5-CH_3 or 2-thio-5-CH_2-NH_2^+-CH_3 derivative in anticodons recognizing A in the 3' position of a codon. For recognition of other bases the 5–OH, –OCH_3, or –OCH_2COO^- derivative is usually present. Yokoyama *et al*. attributed the selectivity to the stabilization of the C2' endo form of the ring in the former group.[191]

3. Aminoacylation of tRNAs

Discrimination between some pairs of tRNAs depends entirely on the anticodon sequence. For example, $tRNA^{Met}$ contains the anticodon CAU. That for a minor $tRNA^{Ile}$ is the same except that the cytosine has been posttranscriptionally modified by covalent linkage of a molecule of lysine via its ε-amino group to C2 of the cytosine. The latter base (**lysidine**) is correctly recognized by *E. coli* isoleucyl-tRNA synthetase; but, if the cytosine is unmodified, it is aminoacylated by methionyl-tRNA synthetase.[192] In most instances the acceptor specificity, or **tRNA identity**, is not determined solely by the anticodon sequence. Thus, when a methionine initiator tRNA was modified to contain a tryptophan anticodon, it was only partially charged with tryptophan *in vivo*. However, when A73 of the methionine tRNA was also converted to G73, only tryptophan was inserted.[193] Nucleotide 73 (Fig. 29-7) is sometimes called the **discriminator nucleotide**.[194–196] It is A in methionine and leucine tRNAs,[197] G in tryptophan tRNAs, and C in histidyl RNAs.[198] The tRNA features needed to establish its identity are sometimes referred to as its **identity-determinant set**.[196,198a,b] This includes the anticodon and other features needed for recognition by the aminoacyl-tRNA synthetases that "charge" the tRNAs with aminoacyl groups.[199–204] For example, for *E. coli* $tRNA^{Val}$ the recognition determinants are A35 and C36 of the anticodon, A73, G20, G45, and a regular A-RNA acceptor helix.[205] All known mature tRNAs contain a 3'-CCA end on which the aminoacylation occurs. Nevertheless, alterations in this sequence still allow correct aminoacylation of some tRNAs.[206]

The aminoacyl-tRNA synthetases (amino acid: tRNA ligases) join amino acids to their appropriate transfer RNA molecules for protein synthesis. They have the very important task of selecting both a specific amino acid and a specific tRNA and joining them according to Eq. 29-1.[175,195,207] These reactions represent the first step in the decoding of mRNAs. Organisms usually contain one aminoacyl-tRNA synthetase for each of the 20 amino acids. Each synthetase must select a specific amino acid and a correct tRNA for that amino acid. The same enzyme transfers an activated amino acid to all of the **isoacceptor tRNAs** specific for a given amino acid. Some aminoacyl-tRNA synthetases attach the aminoacyl group to the 2'–OH of the tRNA substrate, some to the 3'–OH. The chemical mechanism is the same in both cases.

Structures. Aminoacyl-tRNA synthetases vary in size, the subunit masses ranging from 37- to 110-kDa (329–951 residues). There are monomeric species, dimers, tetramers, and $(\alpha\beta)_2$ mixed tetramers.[207,208] Sequence comparisons, together with X-ray structural investigations, have shown that the enzymes can be classified into two groups, each containing ten enzymes.[207,209,209–209b] **Class I aminoacyl-tRNA synthetases** share two consensus sequence motifs: HIGH and KMSKS. Their ATP-binding active sites are in a Rossman fold nucleotide-binding domain (Fig. 2-13).[210] The KMSKS sequence parallels the Walker sequence found in various nucleotide-binding proteins including ATP synthase.[211–213] The actual sequences vary considerably, e.g., the KMSKS sequence for a tyrosyl-tRNA synthetase is actually KFGKT.[211]

Class II aminoacyl-tRNA synthetases contain a different set of three "signature sequences," two of which form an ATP-binding catalytic domain. The active site structure is built on an antiparallel β sheet and is surrounded by two helices (Fig. 29-9). Each class contains subgroups with inserted loops that form other domains. In the following tabulation the reference numbers refer to three-dimensional structural studies.

Class I
Glu,[209] Gln,[218–220] Arg,[221]
Tyr,[222] Trp,[212]
Ile,[223] Leu,[224] Val, Cys, Met[210,225,226]
Class II
His,[227–229] Pro, Ser,[230,231] Thr
Asp,[232,233] Asn,[234]Lys,[217,235]
Phe,[236] Ala, Gly,[204,237]

The 37-kDa 334-residue subunits of the dimeric type I tryptophanyl-tRNA synthetase[238] are the smallest known; the largest bacterial synthetase is an alanine-specific type II tetramer with 95-kDa 875-residue subunits.[239] Gene deletions show that a much smaller core, comparable in size to that of the tryptophanyl-tRNA synthetase, is needed for amino acid activation. The synthetases share little sequence homology except for a short 11-residue part of the adenylate binding site near the N terminus.[240,241] Some of the synthetases contain bound zinc ions.[225,242]

Recognition of cognate tRNAs. Many attempts have been made to learn what part or parts of tRNA molecules are involved in recognition by aminoacyl-tRNA synthetases. Nucleotide sequences of isoacceptor tRNAs have been compared. Chemically modified and fragmented tRNA molecules have been studied, and many mutant tRNAs have been made. These have often been mutants of suppressor tRNAs that place specific amino acids such as phenylalanine or alanine into a peptide at a termination codon, often the termination codon UAG (see Section C4). An alternative approach is to synthesize DNA templates, which can be transcribed *in vitro* by phage T7 RNA polymerase to give mutant tRNAs,[243] or to make such tRNAs by solid-phase chemical synthesis.[244] Although these contain no modified bases, they serve as substrates for the aminoacyl-tRNA synthetases.

The results of these efforts show that no method of tRNA recognition is universal.[244a] In some cases, e.g., for methionine- or valine-specific tRNAs, the synthetase does not aminoacylate a modified tRNA if the anticodon structure is incorrect. Although the anticodon is 7.5 nm away from the CCA end of the tRNA, the synthetases are large enzymes. Many of them are able to accommodate this large distance between a recognition site and the active site (Fig. 29-9A). For some other tRNAs the anticodon is not involved in recognition.[245] For yeast tRNA^{Phe} residues in the stem of the dihydrouridine loop and at the upper end of the amino acid acceptor stem seem to be critical.[241]

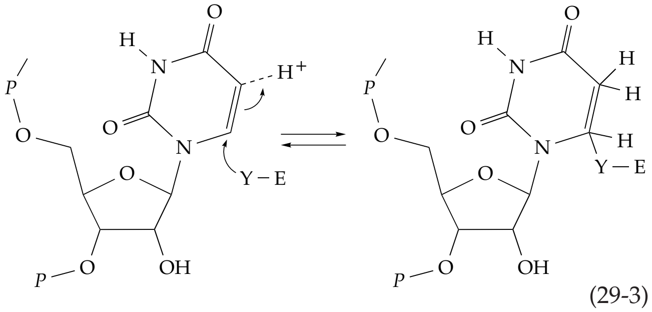
For some other tRNAs only the acceptor helix is essential for recognition. Change of one base-pair, the pair G3 • U70 (a "wobble" pair) of an *E. coli* tRNA^{Ala} • mRNA complex to the unnatural A3 • U70, prevents aminoacylation. Conversely, a G3 • U70 pair formed with tRNAs specific for other amino acids causes them to become substrates for the alanyl-tRNA synthetase.[241,246] Even a shortened tRNA minihelix consisting of a 7-bp acceptor stem, 6-nucleotide loop, and ACCA 3' end is a substrate for this enzyme.[247,247a] A seryl-tRNA synthetase depends upon recognition of two base pairs in the acceptor stem.[248] Synthetic DNA oligomers with sequences corresponding to those of *E. coli* tRNA^{Phe} or tRNA^{Lys}, and with either deoxythymidine or deoxyuridine in the positions occupied by ribouridine in the tRNAs, are also substrates for the synthetases. The affinity and reaction rates are somewhat decreased, but the ribose 2'-OH is not essential for recognition.[249]

Mechanisms of reaction. Activation of an amino acid occurs by a direct in-line nucleophilic displacement by a carboxylate oxygen atom of the amino acid on the α phosphorus atom of MgATP to form the aminoacyl adenylate (Eq. 29-1, step *a*). For yeast phenylalanyl-tRNA synthetases the preferred form of MgATP appears to be the β,γ-bidentate (Λ screw sense) complex (p. 643).[250] This is followed by a second nucleophilic displacement, this one on the C=O group of the aminoacyl adenylate by the –OH group of the tRNA (Eq. 29-1, step *b*; Fig. 29-9C). A conformational change in the protein may be required to permit dissociation of the product, the aminoacyl-tRNA. In the complex of a class I synthetase with aminoacyl adenylate and tRNA the 3' CCA acceptor end of the tRNA is straight, but in a class II synthetase it is bent. In the two classes of synthetase the tRNAs approach the enzyme in a mirror-symmetric fashion. The 2'-OH of the terminal ribose is positioned to attack the carbonyl of the aminoacyl adenylate in class I enzymes, while the 3'-OH is positional for the attack in class II enzymes.[207]

The three-dimensional structure of *E. coli* glutaminyl-tRNA synthetase is shown in Fig. 29-9A. The active site lies against a β sheet in a structure similar to the nucleotide binding domain of a dehydrogenase (Fig. 2-13). The site of binding of a ATP is marked in Fig. 29-9A. The details of this binding to tyrosine-tRNA synthetase have been studied intensively.[251–255] Binding of the tRNA substrates is less well understood. A large series of mutants involving 40 basic residues were prepared by Bedouelle and Winter[256]; study of these mutant enzymes, together with computer-assisted modeling, led to a proposed structure for a transition state for a complex with tRNA as is shown in Fig. 29-9B. Kinetic studies of heterodimers prepared from mutant and normal enzyme confirm that both subunits of the dimeric enzyme interact with the tRNA.

The active site of a type II synthetase is shown in Fig. 29-9C,D.[217,217a] The expected movement of electrons in the reaction with ATP is illustrated by the green arrow in D. Both metal ions and active-site protein groups may participate as is also proposed for another type II enzyme.[229]

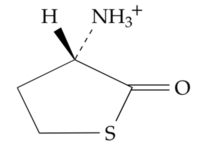
Some data suggested that a transient covalent linkage of tRNA to the synthetases may form through addition of a nucleophilic group of the enzyme to the 6' position of the uracil (or 4-thiouracil) present in position 8 of all tRNAs (Eq. 29-3).[257] The two isoacceptors tRNA^{Tyr} species in *E. coli* contain 4-thiouracil at this position. The C=C bond in this base can be saturated by sodium borohydride reduction, which was found not only to prevent the covalent interaction with the enzyme but also to prevent aminoacylation of the tRNA. However, Eq. 29-3 probably describes a side reaction irrelevant to tRNA function.



(29-3)

Correcting errors. Much attention has been devoted to "proofreading" or "editing" activities of amino acid-tRNA synthetases (see p. 482). For the majority of the enzymes proofreading is not needed.[209a] Thus, tyrosyl-tRNA synthetase (Fig. 29-9B) mistakenly chooses phenylalanine instead of tyrosine only 5 in 10^4 times, apparently a tolerable rate of error. This enzyme, as well as a tryptophan-specific synthetase,[238] depends largely upon differences in the Gibbs energy of binding to select the correct substrate. However, the discrimination between valine and isoleucine by isoleucyl-tRNA synthetase poses a more difficult problem. It is apparently solved, in part, by a "double sieve" editing mechanism,[223,224,258,259] which is described briefly on p. 482. In the first sieve competitors that are larger than the substrate or are differently shaped are excluded by steric repulsion from binding in the active site. Isoleucyl-tRNA synthetase doesn't convert leucine into an aminoacyladenylate, but it does act on the smaller valine. However, most of the resulting enzyme-bound valyl-adenylate is hydrolyzed to valine and AMP before it can be transferred to $\text{tRNA}^{\text{Ileu}}$. It shifts into an editing site, which is too small for the isoleucyl-adenylate, in effect passing through a second sieve.[259a] Some of the activated valine is transferred to tRNA and is removed in a second editing reaction.[223] Misactivation of threonine or some other amino acids by valyl-tRNA synthetase is corrected in an analogous fashion.[259b,c]

A similar editing process prevents isoleucyl-, leucyl-, and methionyl-tRNA synthetases from attaching L-homocysteine to tRNAs.[260–263] In this case, instead of hydrolysis the editing site catalyzes conversion of the homocysteinyl-adenylate into homocysteine lactone. Naturally occurring mutations in tRNA molecules can sometimes have serious consequences. For example, a human mutation is responsible for a fragile mitochondrial isoleucine tRNA and serious cardiomyopathy and opthalmophegia (see also Box 18-B).[263a]



L-Homocysteine lactone

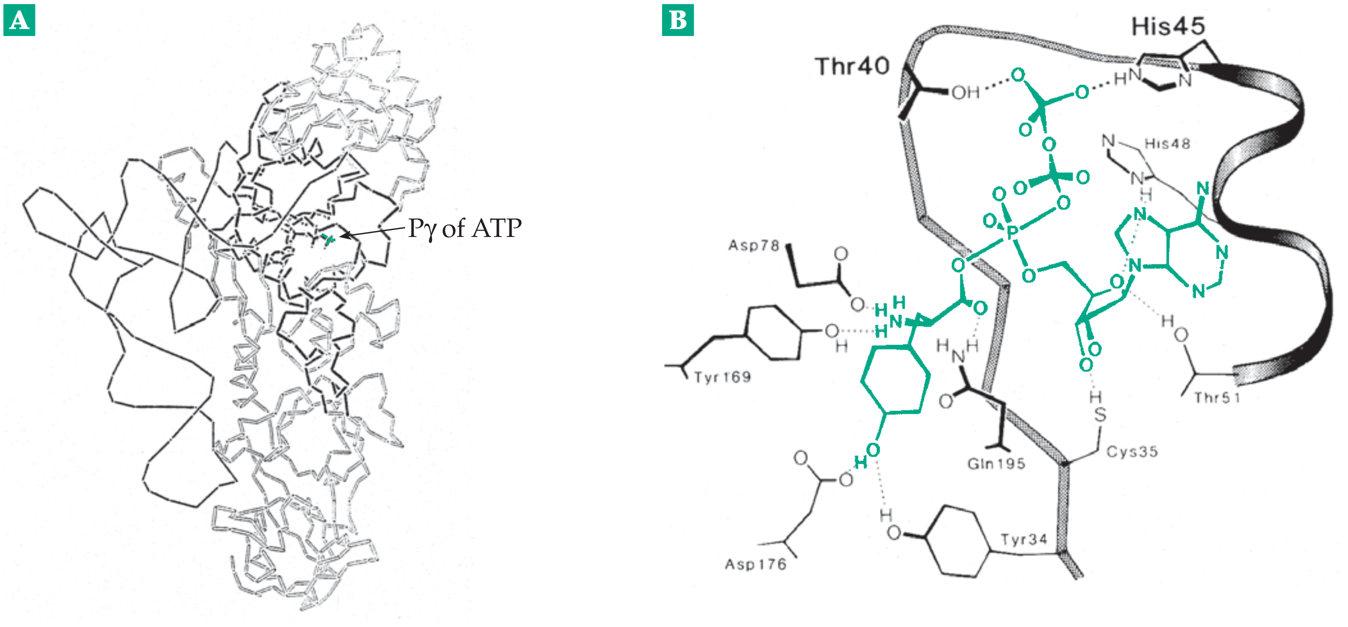


Figure 29-9 Selected views of aminoacyl-tRNA synthetase structure and action. (A) Alpha-carbon trace of the type I *E. coli* glutaminyl-tRNA synthetase. The phosphate backbone of tRNA^{Gln} is shown in black; ATP is shown in the active-site cleft. The canonical dinucleotide fold domain near the N terminus is shaded. Two structural motifs (black), proposed to link the active site with regions of the protein–RNA interface involved in tRNA discrimination, are indicated. The α helix (top) connects tRNA recognition in the minor groove of the acceptor stem with binding of the ribose group of ATP. The large loop (center) connects anticodon recognition by the two β-barrel domains (bottom) with sequences flanking the MSK sequence motif, which interacts with the phosphates of ATP. From Perona *et al.*[214] Courtesy of Thomas A. Steitz. (B) The active site structure of tyrosyl-tRNA synthetase from *Bacillus stearothermophilus* with a penta-coordinate transition state structure modeled.[215] From Leatherbarrow *et al.*[216] (C) Schematic representation of the active site of the lysyl-tRNA synthetase showing potential hydrogen bonding interaction in the ternary complex with lysine and ATP. The invariant motif 2 Arg 262 plays a key role in the recognition of the lysine carboxylate and the ATP α phosphate, while the invariat motif 2 Arg 480 binds the

Proofreading involves kinetic as well as thermodynamic considerations.[264–266] The relative rates of a hydrolytic reaction and the competing activating reaction must always be considered. These ratios can be strongly affected by conformational changes, which may occur in several steps (see also Section C,2).

Essential modification reactions of aminoacyl-tRNAs. In bacteria the initiator tRNA needed to start the synthesis of a polypeptide is initially aminoacylated by methionine, but the methionyl-tRNA^{fMet} must then be *N*-formylated by transfer of a formyl group from N^{10}-formyltetrahydrofolate (Fig. 15-18; Eq. 29-4).[267,268] In gram-positive bacteria and in archaea, mitochondria, and chloroplasts the glutamine-specific tRNA^{Gln} is charged with glutamate to form Glu-tRNA^{Gln}. The latter is converted by action of an ATP-dependent amidotransferase (see Eq. 24-22) to the necessary Gln-tRNA^{Gln} (Eq. 29-5).[268a] In a similar way, RNA^{Asn} in some organisms is charged with aspartate, then converted by transamidation to Asp-tRNA^{Asn} (Eq. 29-6).[207,267,269] An important reaction, that occurs in all kingdoms of life, is the charging of the special tRNA^{Sec} with serine and conversion of the product into selenocysteinyl-tRNA^{Sec} (p. 827; Eq 29-7).[267,270]

$$\text{tRNA}^{\text{fMet}} \xrightarrow[\text{ATP}]{\text{Met}} \text{Met-tRNA}^{\text{fMet}} \xrightarrow{\text{Methyl-THF} \quad \text{THF}} \text{fMet-tRNA}^{\text{fMet}} \tag{29-4}$$

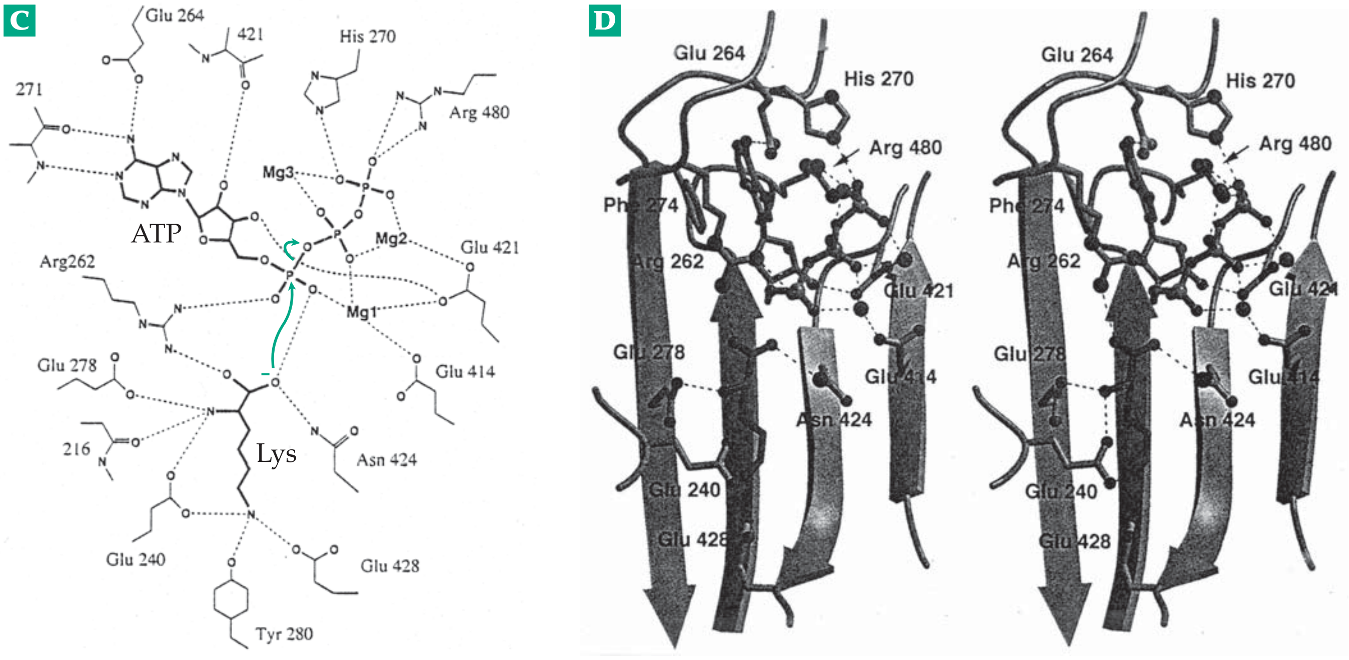
$$\text{tRNA}^{\text{Gln}} \xrightarrow[\text{ATP}]{\text{Glu}} \text{Glu-tRNA}^{\text{Gln}} \xrightarrow[\text{ATP}]{\text{Gln} \quad \text{Glu}} \text{Gln-tRNA}^{\text{Gln}} \tag{29-5}$$

$$\text{tRNA}^{\text{Asn}} \xrightarrow[\text{ATP}]{\text{Asp}} \text{Asp-tRNA}^{\text{Asn}} \xrightarrow[\text{ATP}]{\text{Gln} \quad \text{Glu}} \text{Asn-tRNA}^{\text{Asn}} \tag{29-6}$$

$$\text{tRNA}^{\text{Sec}} \xrightarrow[\text{ATP}]{\text{Ser}} \text{Ser-tRNA}^{\text{Sec}} \xrightarrow[\text{(p. 827)}]{\text{Se-}P \quad P_i} \text{Seleno-cysteinyl-tRNA}^{\text{Sec}} \tag{29-7}$$

The introduction of selenocysteine into proteins at selected stop codons using this tRNA is described in Section C,5.

Examination of the complete genome sequences of methanogens revealed an apparent lack of cysteinyl-tRNA synthetase. However, prolyl-tRNA synthetase does correctly aminoacylate the tRNAs for both proline and cysteine in these archaeobacteria.[271–272a]



γ phosphate of the ATP. A number of conserved residues in the motif 2 loop (residues 264–271) assume an ordered conformation only upon ATP binding. The positions of the Mg^{2+} sites are indicated. (D) View of the active site of the type II lysyl-tRNA synthetase showing the conformations of the substrates lysine and ATP before the first step of the reaction takes place. The ATP molecule is located on one side of the central β sheet of the C-terminal domain, with the adenine ring sandwiched between a conserved phenylalanine (Phe 274) and the motif 3 arginine residue (Arg 480). The pyrophosphate moiety is bent toward the adenine placing the α phosphate in the correct position for nucleophilic attack of the lysine carboxylate oxygen. The hydrogen bonding and electrostatic interactions between the substrates and some of the key residues, including the invariant motif 2 arginine (Arg 262), are shown. The three Mg^{2+} ions (green) involved in catalysis are included. (C) and (D) are from Desogus *et al.*[217]

Additional functions of aminoacyl-tRNA synthetases. The primary function of these enzymes in protein synthesis is well known, but they have a whole range of other activities.[273,274] In *E. coli* the large alanyl-tRNA synthetase can repress transcription of its own gene by binding to a palindromic sequence in the control region of the gene.[275] Expression of some genes, such as that for threonyl-tRNA synthetase, is regulated at the translational level.[273,274,276] In mammalian cells the formation of the threonine-specific synthetase appears to be regulated by a phosphorylation–dephosphorylation mechanism.[277] Other synthetases participate in mitochondrial RNA splicing[278] and in aminoacylation of tRNA-like 3' ends of viral genomes (see Fig. 28-24) and of N termini of certain proteins.[279] For example, an arginyl group may be transferred onto the N terminus of a protein, marking it for rapid degradation.[280] Under conditions of apoptosis, tyrosyl tRNA synthetase is hydrolytically cleaved to form two different cytokines.[274] Phenylalanyl-tRNA synthetase is a DNA-binding protein.[281] Within the nucleus newly synthesized tRNAs are checked before being exported to the cytoplasm. Only tRNAs with mature 5' and 3' ends are exported. In both *Xenopus* oocytes[281a] and in *S. cere- visiae*[281b] the tRNAs are also tested prior to export, using aminoacyl-tRNA synthetases, to ensure that they are functional.

Many proteins have structures related to those of aminoacyl-tRNA synthetases.[282,283] For example, asparagine synthetase A functions via an aspartyl-adenylate intermediate (Chapter 24, Section B), and its structure resembls that of aspartyl-tRNA synthetase.[284] The *his G* gene of histidine biosynthesis (Fig. 25-13) encodes an ATP phosphoribosyltransferase with structural homology to the catalytic domain of histidyl-tRNA synthetase.[284] The reason is not clear, but some aminoacyl-tRNA synthetases, especially the histidyl-tRNA synthetase, are common autoantigens for the inflammatory disease **polymyositis**.[285,286]

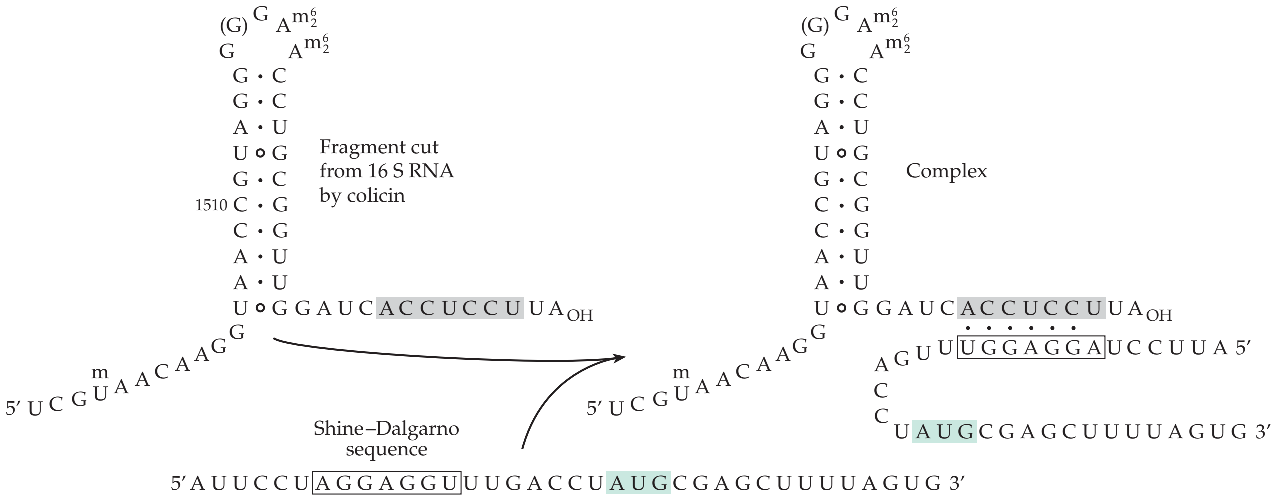
C. Protein Synthesis: The Ribosome Cycle

Initiation (Figs. 29-10 and 29-11), elongation (Fig. 29-12), and termination are three distinct steps in the synthesis of a protein. A variety of specialized proteins are required for each stage of synthesis. Their sequential interaction with ribosomes can be viewed as a means of ensuring an orderly sequence of steps in the synthesis cycle. The rate of protein formation will depend upon the concentrations of amino acids, tRNAs, protein factors, numbers of ribosomes, and kinetic constants. The formation of specific proteins can also be inhibited by **translational repressors**, proteins that compete with ribosomes for binding to target mRNAs.[287]

1. Initiation

For most polypeptide chains initiation begins with one of the three **initiation codons**, most commonly the methionine codon AUG. When properly placed in an mRNA chain, GUG may also serve as a bacterial initiation codon. In such cases, it codes for methionine rather than for valine. Occasionally UUG, AUU, ACG, and perhaps other codons can initiate translation.[288,289] This is less frequent in eukaryotes than in bacteria. The sequence of bases preceding the initiation codon must also be important for recognition of the "start" signal.

In *E. coli* polypeptide chains are always initiated with the amino acid ***N*-formylmethionine**. Some bacteria can apparently live without the ability to formylate methionyl-tRNA,[290] but most eubacteria as well as mitochondria and chloroplasts use formylmethionine for initiation. In a few cases, both among bacteria and eukaryotes, initiation can sometimes occur with other amino acids.[291] The first step is the alignment of the proper initiation codon correctly on


Fragment cut from 16 S RNA by colicin
1510
5′ U C G U A A C A A G G
U ∘ G G A U C A C C U C C U U A OH
Complex
A G U U U G G A G G A U C C U U A 5′
U A U G C G A G C U U U U A G U G 3′
Shine–Dalgarno sequence
5′ A U U C C U A G G A G G U U U G A C C U A U G C G A G C U U U U A G U G 3′

R 17, a protein initiator region

(29-8)

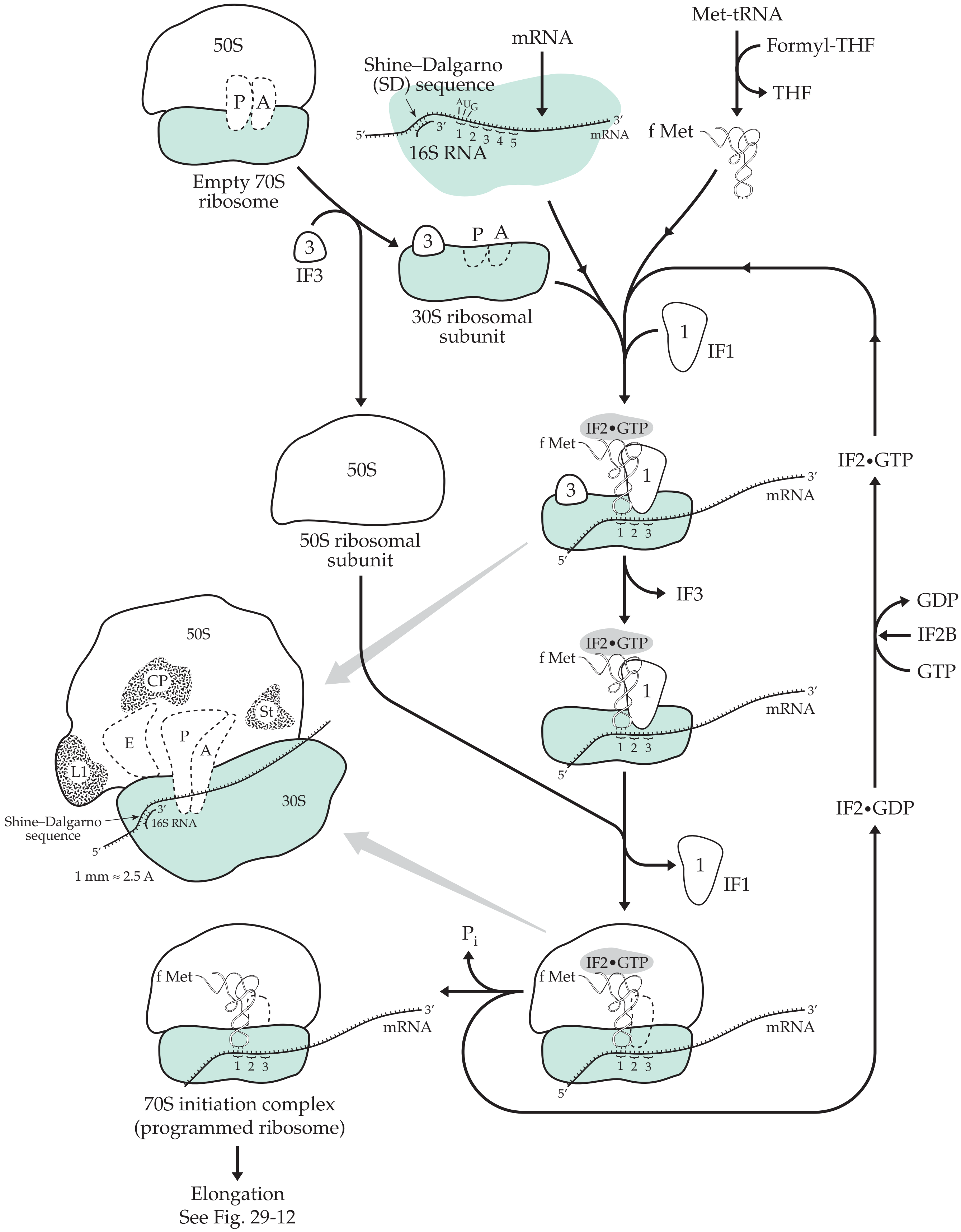

50S
P
A
Empty 70S ribosome
mRNA
Shine–Dalgarno (SD) sequence
AUG
5′
3′
1 2 3 4 5
16S RNA
mRNA
Met-tRNA
Formyl-THF
THF
f Met
3
IF3
P A
30S ribosomal subunit
1
IF1
50S
50S ribosomal subunit
IF2•GTP
f Met
mRNA
IF3
GDP
IF2B
GTP
IF2•GDP
50S
CP
St
E
P
A
L1
30S
Shine–Dalgarno sequence
16S RNA
1 mm ≈ 2.5 Å
IF1
P_i
70S initiation complex (programmed ribosome)
Elongation
See Fig. 29-12

Figure 29-10 Initiation of protein synthesis on bacterial ribosomes. Images are not drawn to scale. Some details are indicated on the larger scale image at the left.

the ribosome and the binding to it of a molecule of initiator tRNA carrying *N*-formylmethionine.[268,292,293] The process by which this occurs is relatively complex, partly because it is essential for the ribosomes to distinguish the true initiation codon from the many AUG codons in internal positions in the message. In bacteria recognition of the initiation codon is assisted by base pairing between the conserved sequence ACCUCCU at positions 1534–1540 at the 3' end of the 16S RNA (Fig. 29-2A) and the complementary **Shine–Dalgarno sequence** AGGAGGU, which is found near the 5' end of most mRNA molecules.[53] This is illustrated in Eq. 29-8 for a messenger RNA in the form of the A protein initiator region from the R17 phage RNA.[294,295] Ribosomal protein S1 also seems to be required for this binding.[135,295a]

Prokaryotic initiation factors. In addition to the ribosomal proteins, the initiation factors **IF1**, **IF2**, and **IF3**, whose molecular masses are 9.5, 9.7, and 19.7 kDa, respectively,[70,296] are essential. They coordinate a sequence of reactions that begins with the dissociation of 70S ribosomes into their 30S and 50S subunits. Then, as is shown in Fig. 29-10, the mRNA, the initiator tRNA charged with formylmethionine, the three initiation factors, and the ribosomal subunits react to form 70S **programmed ribosomes**, which carry the bound mRNA and are ready to initiate protein synthesis. IF2 is a specialized G protein (Chapter 11), which binds and hydrolyzes GTP. It resembles the better known elongation factor EF-Tu (Section 2). The ~172-residue IF3 consists of two compact α/β domains linked by a flexible sequence, which may exist as an α helix.[296a–298] Its C-terminal domain binds to the central domain of the 16S RNA near nucleotides 819–859 (Fig. 29-2). When bound it protects nucleotides in the 690 loop from chemical modification[297] and induces a conformational change in the loop.[297a] Binding of IF3 prevents association of the 30S and 50S subunits, assuring the cell of a supply of free 30S subunits for translational initiation. It also promotes the binding of the other two factors: IF1 and IF2.[299] Binding of IF2, as its GTP complex, stimulates the binding of fMet-tRNA in the adjacent P site.[300] Another function of IF2, in cooperation with IF1, may be to remove peptidyl-tRNAs with short polypeptide chains under conditions in which such peptidyl-tRNAs accumulate to abnormal levels.[301] However, the order of binding, which is implied in Fig. 29-10, has been hard to establish.

IF1, which is essential to the viability of bacteria, binds and partially occludes the A site of the ribosome, preventing the initiator fMet-tRNA from incorrectly occupying the 30S A site.[70,296a,302] Binding of IF1 also causes the functionally important bases A1492 and A1493 of 16S RNA (Fig. 29-2) to be flipped out of helix 44 and to bind to pockets in IF1. This induces further long-range conformational changes.[70] It has also been hard to establish whether the charged initiator tRNA binds into the P site before or after the mRNA binds to the 30S subunit. Some evidence supports the latter possibility,[296] which is indicated in Fig. 29-10. In any case, an important step is the specific base-pairing of the initiator tRNA with the first AUG start codon. IF3 seems to be essential for this pairing process, which establishes the correct reading frame for translation of the genetic message.[303] A proofreading function at this step is often attributed to IF3.[304] Intact ribosomes bind charged tRNAs tightly in the P site. Perhaps the initial binding to the 30S subunit is loose enough to allow the mRNA, which ties itself to the 3' end of the 16S RNA via Eq. 29-3, to move back and forth until the correct reading frame is located. Then a conformational change occurs and locks the initiator tRNA in place. This change also weakens the binding of IF3, which dissociates from the complex, allowing the 50S subunits to rejoin the complex. The ribosome-binding domain of IF3 is homologous to spliceosome protein U1A (Chapter 28).[304]

The hydrolysis of GTP during initiation is essential as is shown by the fact that 5'-guanylmethylene diphosphonate, a GTP analog containing a methylene bridge between the terminal and central phosphorus atoms (see p. 558), can substitute for GTP in all steps up to and including the binding of the 50S ribosome. However, it cannot function in the final step because it cannot undergo hydrolysis. Why is GTP hydrolysis needed? It may provide energy for the conformational rearrangement of ribosomal components, or it may simply be required for release of the IF2•GDP complex. For example, IF2•GTP may bind to the ribosome with a high affinity, but IF2•GDP only weakly. Remember that G-proteins exist in at least two conformations, one stabilized by GDP and another by GTP (Chapter 11). When the hydrolysis of the bound GTP is incorporated into a reaction sequence, it provides a Gibbs energy change that may be needed to drive the reactions. In this case, it ensures that the charged initiator tRNA is firmly bound and ready to initiate translation.

Some information about spatial arrangements of the ribosomal proteins involved in initiation was provided by the fact that antibodies against proteins S19 and S21 block the formation of a complex with fMet-tRNA, while antibodies against S2, S18, and S20 block the binding of IF3. Crosslinking experiments showed that IF2 and S19 are close together and that IF3 is close to S12 (Fig. 29-1A).

Initiation of protein synthesis in eukaryotes. Most eukaryotic mRNAs have a 5' cap (p. 1642) and lack a Shine–Dalgarno sequence. Otherwise, initiation follows a pattern similar to that in bacteria but more complex.[305–308] There are at least ten eukaryotic

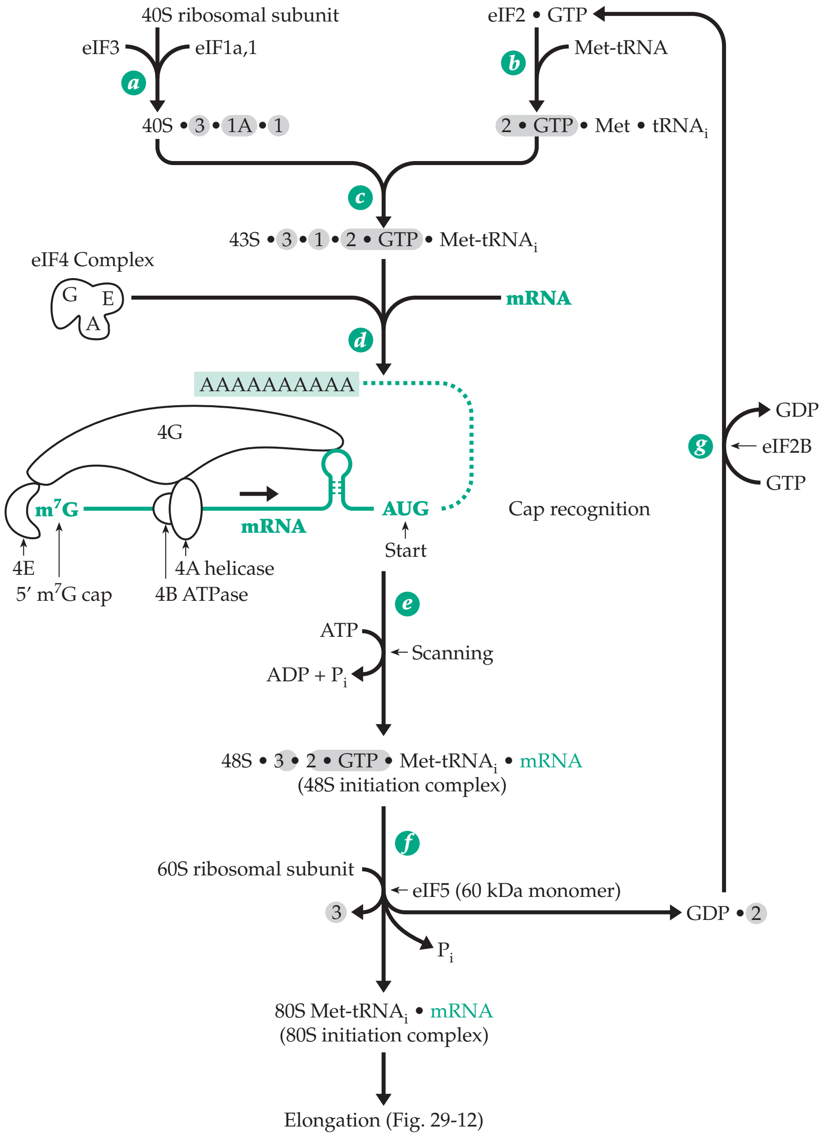

40S ribosomal subunit
eIF3
eIF1a,1
a
40S • 3 • 1A • 1
eIF2 • GTP
b
Met-tRNA
2 • GTP • Met • tRNA$_i$
c
43S • 3 • 1 • 2 • GTP • Met-tRNA$_i$
eIF4 Complex
G
E
A
mRNA
d
AAAAAAAAAA
4G
m^{7}G
mRNA
AUG
Cap recognition
4E
5′ m^{7}G cap
4A helicase
4B ATPase
Start
e
ATP
ADP + P$_i$
Scanning
48S • 3 • 2 • GTP • Met-tRNA$_i$ • mRNA
(48S initiation complex)
f
60S ribosomal subunit
eIF5 (60 kDa monomer)
3
P$_i$
GDP • 2
g
GDP
eIF2B
GTP
80S Met-tRNA$_i$ • mRNA
(80S initiation complex)
Elongation (Fig. 29-12)

Figure 29-11 Initiation of eukaryotic protein synthesis. ①, ②, ③ = eIF1,2,3

initiation factors (eIFs), some composed of several peptides.[309] Hydrolysis of both ATP and GTP is required to form the initiation complex. Cap-binding proteins help to locate the 5' end, but the first initiation codons occur at greatly varying distances from the cap. Ribosomes apparently conduct a systematic scanning beginning at the 5' end cap of the mRNA to locate the first initiation codon.[305,310,311] Initiation of translation in eukaryotes is also often subject to controls that are more complex than those in bacteria.[308,312–314a] At least 25 polypeptides are involved.[315] Specific functions of only a few of these are described here. Functions of some are unknown or uncertain, and new proteins such as the stimulatory factor 4H[316,317] have been discovered recently.

The first initiation step is the dissociation of idle 80S ribosomes into their 40S and 60S subunits. This depends upon the ~700-kDa eIF3, a complex of 5–11 peptides of mass 30 to 170 kDa each, which binds to the 40S subunit (Fig. 29-11, step *a*).[306,311,318–320] In a separate reaction (step *b*) the charged initiator tRNA (Met-tRNA$_i$) is bound by the G protein eIF2,[321–324] an αβγ mixed trimer whose α subunit not only binds GTP but is also the site of regulation by a phosphorylation–dephosphorylation mechanism.[325,326] As is indicated in Fig. 29-11, GDP-containing eIF2, released from ribosomes upon formation of the initiation complex, does not bind the charged initiator tRNA. The bound GDP must first be exchanged for GTP, a process that requires a five-subunit guanine nucleotide exchange factor eIF2B (Fig. 29-11, step *g*).[327–328a] However, if the α subunit of eIF2 becomes phosphorylated, the nucleotide exchange and consequently the initiation of protein synthesis, is retarded.[329–330c] In reticulocytes protein synthesis stops rapidly if there is a deficiency of heme. This appears to occur via a **heme-sensitive eIF2α kinase** whose catalytic activity is inhibited by heme.[322,327,331] The [NAD$^+$]/[NADH] ratio may also be a factor in controlling the nucleotide-exchange GEF-catalyzed reaction.[332]

The ternary complex of eIF2 with GTP and Met-tRNA$_i$ binds to the 40S complex (Fig 29-11, step *c*). If the ribo-trinucleotide AUG, the initiation codon, is added to this complex, it is converted to a 43S initiation complex.[321] However, with natural mRNAs several additonal protein factors as well as ATP are needed (Fig. 29-11, step *d*). Factor 4F (eIF4F) is a large complex of several components known as factors 4A, 4B, 4E, and the large 220-kDa 4G (formerly EIF4γ or p220).[333–334a] Factor 4G is a multifunctional adapter or scaffold that apparently organizes the complex and coordinates various control mechanisms.[335–338] Factor 4E is a cap-binding protein, which recognizes and binds tightly to the ^{7}mG cap present on most mRNAs.[306,339–341a] It also binds to 4G. Factor 4A, an ATPase, acts together with 4B as an RNA helicase to unwind the mRNA and remove hairpin loops[342,343] during the scanning to locate the initiation codon (step *e*). Kozak suggested that the 43S ribosome scans from the cap at the 5' end and stops at the first initiation codon, which is usually the AUG found within the sequence (A/G)NNAUGG.

AUG codons in other positions, known as **internal ribosome entry sites** (IRES),[311,344–347] and, more rarely, non-AUG codons can also initiate translation with lower efficiency.[348] Thus, mechanisms exist for synthesis of small amounts of proteins of varying lengths and of proteins that are encoded in any one of the three reading frames.[305,349–350a] Even circular RNAs can serve as mRNAs by this mechanism.[351] It is significant that, as shown in Fig. 19-11, factor 4G, the large subunit of eIF4, also binds to the poly(A) tail present on the 3' terminus of most mRNAs. This binding, which seems to be essential for rapid initiation,[352–354] is mediated by yet another protein, the **poly(A)-binding protein**. The importance of this protein in the human body is emphasized by its identification as **ataxin-2**, the protein defective in type 2 spinocerebellar ataxia (see Table 26-4).[354] The significance to the regulation of initiation is not clear, but the poly(A) binding may favor reuse of the mRNA, which may be translated repeatedly under conditions of rapid growth.

The last initiation step (step *f*, Fig. 29-11) is the reaction of the 60S ribosomal subunit with the 48S initiation complex to form the 80S initiation complex. Initiation factors 3, 4C, the eIF2•GDP complex, and inorganic phosphate are all released in this process, which is promoted by IF5. This monomeric ~60-kDa protein[355,356] also stimulates conversion of the GTP bound to IF2 into GDP and P_i. IF5 is unique as the only known protein containing **hypusine**, N^ε-(4-amino-2-hydroxybutyl)lysine, a posttranslationally modified lysine. It occurs only at position 50 in the 17-kDa protein.[356–358] Hypusine is not present in eubacteria but is essential for viability of both eukaryotes and archaeobacteria[358] and is present within an invariant 12-residue sequence.

2. Elongation of Polypeptide Chains

Once the initiating fMet-tRNA of bacteria or the eukaryotic Met-tRNA_i is in place in the P site of a ribosome and is paired with the initiation codon in the mRNA, peptide chain growth can commence. Amino acid residues are added in turn by insertion at the C-terminal end of the growing peptide chain. Elongation requires three processes repeated over and over until the entire peptide is formed.

1. Codon-specific binding of a charged tRNA bearing the next amino acid at the A site (decoding).
2. Formation of the peptide bond. This process transfers the growing peptide chain from the tRNA in the P site onto the aminoacyl-tRNA in the A site.
3. Translocation of the peptidyl tRNA from the A site to the P site. This process also involves movement of the used tRNA from the P site into the exit site and simultaneous movement of the mRNA to bring the next codon into place in the A site. Both the release from the A site and translocation require energy. This is provided by the hydrolysis of GTP, one molecule for each of the two processes.[359,360]

The elongation cycle for *E. coli* is shown in Fig. 29-12. That for eukaryotic ribosomes is similar except that 40S and 60S subunits are involved in formation of the complete 80S ribosome.

Codon-specific binding of an aminoacyl-tRNA (decoding). The binding of an aminoacyl-tRNA to the A site of the 70S or 80S initiation complex depends upon a protein called **elongation factor Tu** (**EF-Tu** or **eIF1** in eukaryotes), which is present as a mixed dimer with a second protein, **EF-Ts**. In *E. coli* EF-Ts is a stable 35-kDa protein, while Tu is a 43-kDa soluble protein present in a large excess over Ts. Tu is one of the most abundant soluble proteins in bacterial cells and accounts for about 5% of the total protein. Most of the tRNAs in a bacterial cell are present as complexes with Tu. Tu may also have functions other than in protein synthesis and is found associated with the plasma membrane as well as with ribosomes.

Factor Tu is a G protein. It not only carries the aminoacyl-tRNAs into the A site on ribosomes, as shown in Fig. 29-12, but also binds and hydrolyzes GTP during the elongation cycle.[361–368] Factor Ts is a nucleotide exchange factor that catalyzes the exchange of GDP bound to Tu for GTP.[369] This is shown in Fig. 29-12 (steps *a* and *b*). The GTP/GDP-binding site of EF-Tu is located in the N-terminal portion. Eukaryotic eEF-Tu is also called EF1α or EF-TA. As isolated from various sources it has a molecular mass ranging from 50 to 53 kDa. Like the bacterial counterpart, it is abundant.[368,370–373] Like bacterial EF-Tu it exists largely as a complex with a more abundant nucleotide exchange factor EF1β. The complex tends to be bound to actin filaments.[373,374] Fungal eEF-Tu contains mono-, di-, and trimethylated lysine at up to 16 positions.[371]

EF-Tu will bind to any aminoacylated tRNA other than $\text{tRNA}^{\text{f-Met}}$, the initiator tRNA[374a] (step *c*, Fig. 29-12), and carry it to the ribosome (step *d*), where it binds into the A site. There it is selected if it forms a proper base pair with the mRNA codon in the A site or is rejected if it does not. This decoding process involves both an initial step and a proofreading step. The aminoacyl-tRNA binds both to the decoding site in the 16S RNA and to the peptidyltransferase site in the 23S RNA. (See discussions on p. 1687.) The decoding site is on the platform at the upper end of helix 44 (Fig. 29-2). Nucleotide G1401 plays a crucial role.[375] When one of the isoacceptor species of *E. coli* tRNA^{Val} is irradiated with ultraviolet light, the

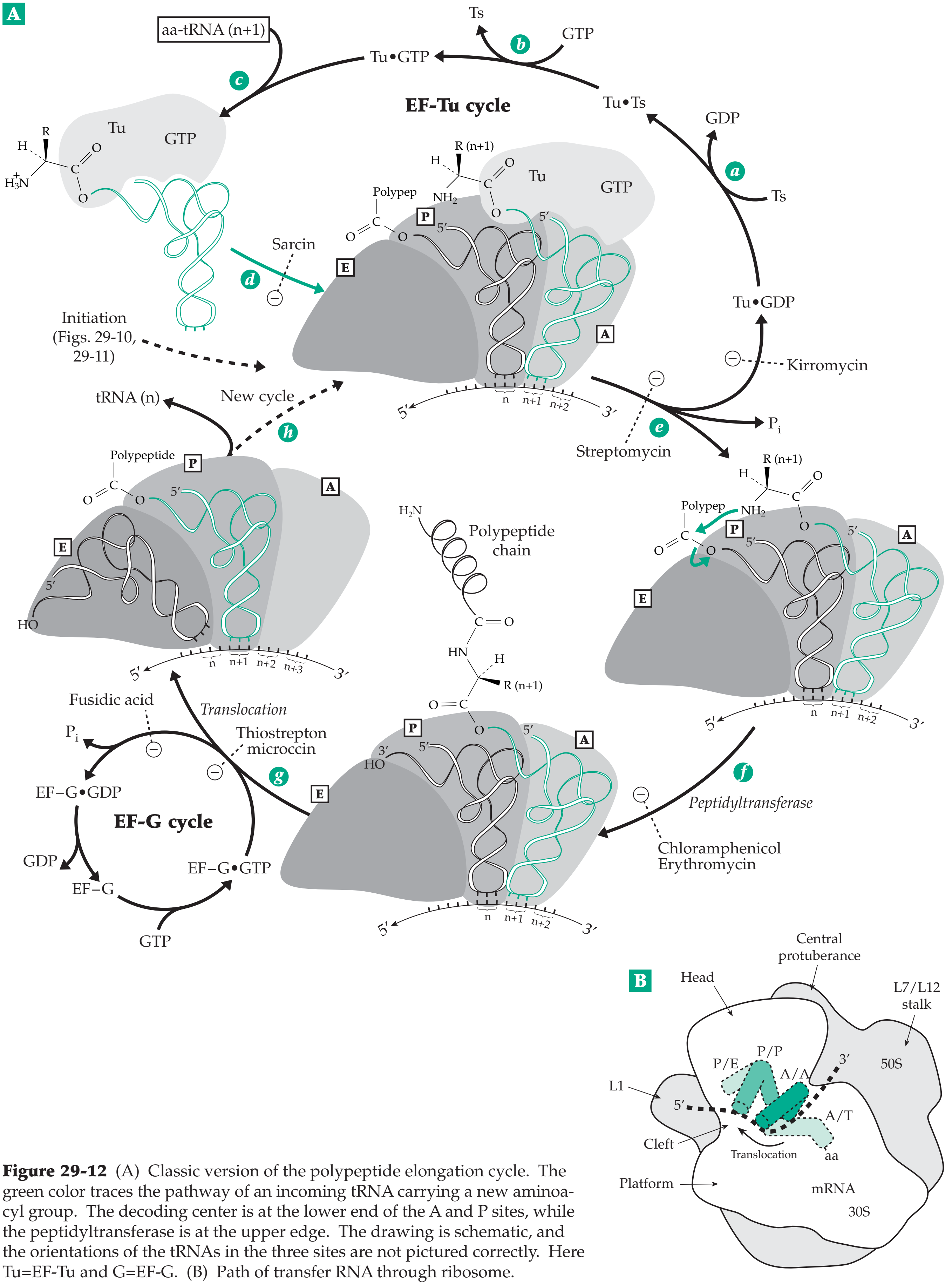
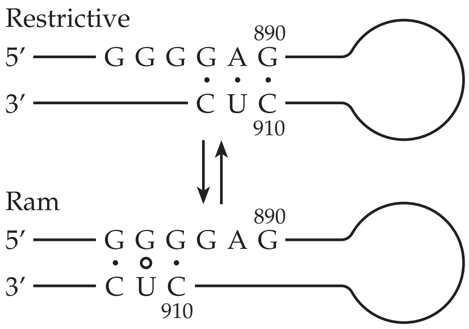

A
aa-tRNA (n+1)
Ts
b
GTP
Tu•GTP
c
EF-Tu cycle
Tu•Ts
GDP
a
Ts
Tu
GTP
R
H
H3N
C
O
O
R (n+1)
H
NH2
Polypep
Tu
GTP
P
5′
O
E
Sarcin
d
A
Tu•GDP
Initiation (Figs. 29-10, 29-11)
Kirromycin
tRNA (n)
New cycle
h
n
n+1
n+2
5′
3′
e
P_i
Streptomycin
Polypeptide
P
A
E
5′
HO
H2N
Polypeptide chain
C=O
HN
H
R (n+1)
O=C
R (n+1)
Polypep
NH2
P
A
E
n
n+1
n+2
n+3
5′
3′
Fusidic acid
Translocation
Thiostrepton microccin
P_i
g
EF–G•GDP
EF-G cycle
GDP
EF–G
EF–G•GTP
GTP
P
A
E
3′
HO
f
Peptidyltransferase
Chloramphenicol
Erythromycin
B
Head
Central protuberance
L7/L12 stalk
50S
P/E
P/P
A/A
3′
L1
5′
A/T
Cleft
Translocation
aa
Platform
mRNA
30S

Figure 29-12 (A) Classic version of the polypeptide elongation cycle. The green color traces the pathway of an incoming tRNA carrying a new aminoacyl group. The decoding center is at the lower end of the A and P sites, while the peptidyltransferase is at the upper edge. The drawing is schematic, and the orientations of the tRNAs in the three sites are not pictured correctly. Here Tu=EF-Tu and G=EF-G. (B) Path of transfer RNA through ribosome.

5-(carboxymethoxy)uridine at position 34 in the anticodon becomes crosslinked specifically with C-1400 of the 16S rRNA in the 30S subunit.[5,376] This nucleotide lies in the deep cleft, in the decoding region, between the neck and the platform of this ribosomal subunit (Fig. 29-1). Various crosslinking and protection experiments[377,378] show that other helix 44 residues bind the tRNAs in both the P and A sites. A1492 and A1493 form part of the A site, while C1400 is in the P site.[378,379] Also strongly affecting tRNA binding and decoding is the nearby **switch helix** in the 900 region of the 16S RNA. This helix readily undergoes a shift between two hydrogen-bonded configurations (Eq. 29-9).[378,378a,380]



(29-9)

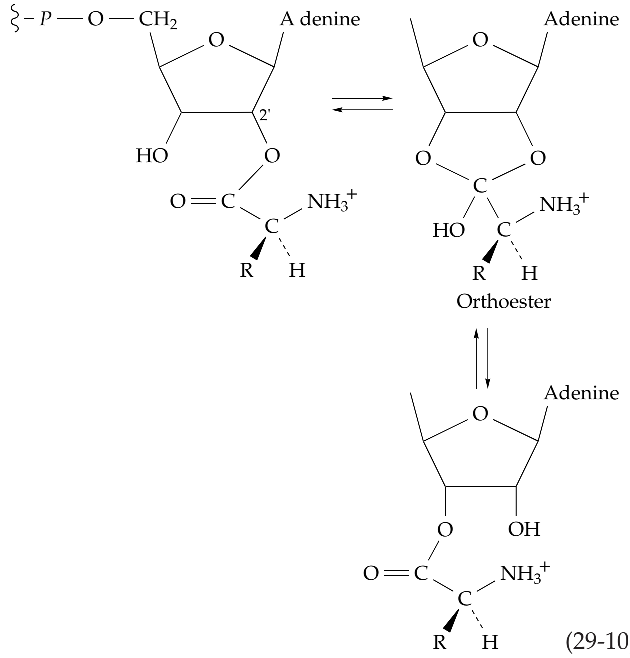
Judging by the effects of mutations in 16S RNA or in proteins S5 and S12 that favor one or the other conformation, the restrictive conformation gives a greater fidelity in translation than the “ram” (ribosomal ambiguity) conformation.[380] This loop is near the central pseudoknot in the 16S RNA and is involved in binding S5 and S12 as well as streptomycin (Box 29-B), all of which affect fidelity of protein synthesis. As mentioned on p. 1687 the adenine rings of residues A1492 and A1493 move out to interact with the CCA-3' ends of the tRNA (Fig. 29-14).

The location of binding of EF-Tu on ribosomes has been established directly by cryo-EM. It binds both to the L7/L12 stalk and to the body of the 50S ribosomal subunit.[381] The other end of the P site is at the peptidyltransferase locus and has been photochemically labeled by azide derivatives of aminoacyl groups bound to a tRNA.[382] The labeling is primarily in the 50S subunit of *E. coli* ribosomes and involves the central loop of domain V (residues 2043–2625) of the 23S RNA. Residues U2584 and U2585 are major sites of crosslinking (see Fig. 29-14). The presence of nearby sites of mutation leading to resistance to chloramphenicol or erythromycin[383,384] (Box 29-B) served to confirm the central loop as part of the peptidyltransferase. Domain II of 23S RNA is also involved, and there is evidence that the unique sequence UGG at positions 807–809 may also interact with the CCA end of tRNA in the P site.[361]

Bound Tu leaves a “footprint” at positions 2655 and 2661 in the sarcin/ricin loop of domain VI (Fig. 29-4H) when chemical probes are applied.[385] From a thermodynamic viewpoint, the hydrolysis of GTP to GDP and P_i during the functioning of EF-Tu is unnecessary, but it appears to drive a conformational change needed to bring the reacting groups together or as part of a proofreading mechanism.[385-385d] The hydrolysis of GTP appears to follow codon–anticodon recognition between the tRNA and mRNA in the A site (see Figs. 29-2G and 29-14). EF-Tu • GDP has a greatly reduced affinity for an aminoacyl-tRNA and dissociates, leaving the latter firmly bound into the A site.[385b–d] Simonson and Lake proposed that binding of a tRNA into the A site is *preceded* by binding into a **D site**. After the initial binding the anticodon of the tRNA wing from the D site into the A site as a result of a conformational rearrangement of the base stacking within the tRNA.[385e]

The peptidyltransferase reaction. It has been difficult to establish whether the 2' or the 3' of the terminal adenosine of tRNA carries the activated aminoacyl or peptidyl group. Rapid equilibration between the two via an orthoester may occur (Eq. 29-10), and EF-Tu of *E. coli* binds to either the 2' or 3' isomer.[386,386a] However, reaction of the 3'-aminoacyl-tRNA at the peptidyltransferase site is probable.[387,387a,b]

The peptidyltransferase reaction resembles that of the proteases (Chapter 12, pp. 649, 650), with a tetrahedrally bonded intermediate probable (Fig 29-13A). As is shown on pp. 649–650, the catalytic acid has been proposed to be the N3 atom of adenosine 2486 (2451 in *E. coli*) in the *H. marismortui* 23S RNA. This is in the central loop of domain V (Fig. 29-14). However, replacement of A2451 with G, U, or C did



(29-10)

not totally destroy peptidyltransferase activity in *Thermus aquaticus*, nor did replacement of the essential G2447 with A, U, or C. Polacek *et al.* suggest that the ribosome may need only to hold the reacting aminoacyl and peptidyl groups attached to tRNA in the correct positions for reaction and that no other catalysis is necessary.[387c] However, A2451G, U, or C mutant ribosomes have very low activity and organisms with such mutations are often not viable.[33f,167b] The pH rate profile for peptidyltransferase activity indicates a pK_a of ~7.5 in the RNA.[33f,387d] This is consistent with the view that A2451 may serve as a catalytic base. There has been considerable discussion about the pK_a. Can it be assigned to A2451? As mentioned on pp. 751–753, many enzymes have a broad pH region of maximum velocity over which catalytic groups of quite different microscopic pK_as (pp. 305–307) may function. For ribosomal RNAs, as for proteins, tight bonding between ionized groups in a substrate–catalyst complex may lock in an overall protonation state of the macromolecule. However, a proton may jump from one group to another within the complex (e.g., as in Fig. 29-14D) to provide a set of tautomeric species in a pH-independent equilibrium. Among these some will be on the catalytic pathway. One may arise by deprotonation of the reacting $-NH_2$ group of the aminoacyl-tRNA (Eq. 29-1, step *c*). Conformational changes,[387e] which may be induced by proton movements, may also be encompassed within the array of pH-independent equilibria.

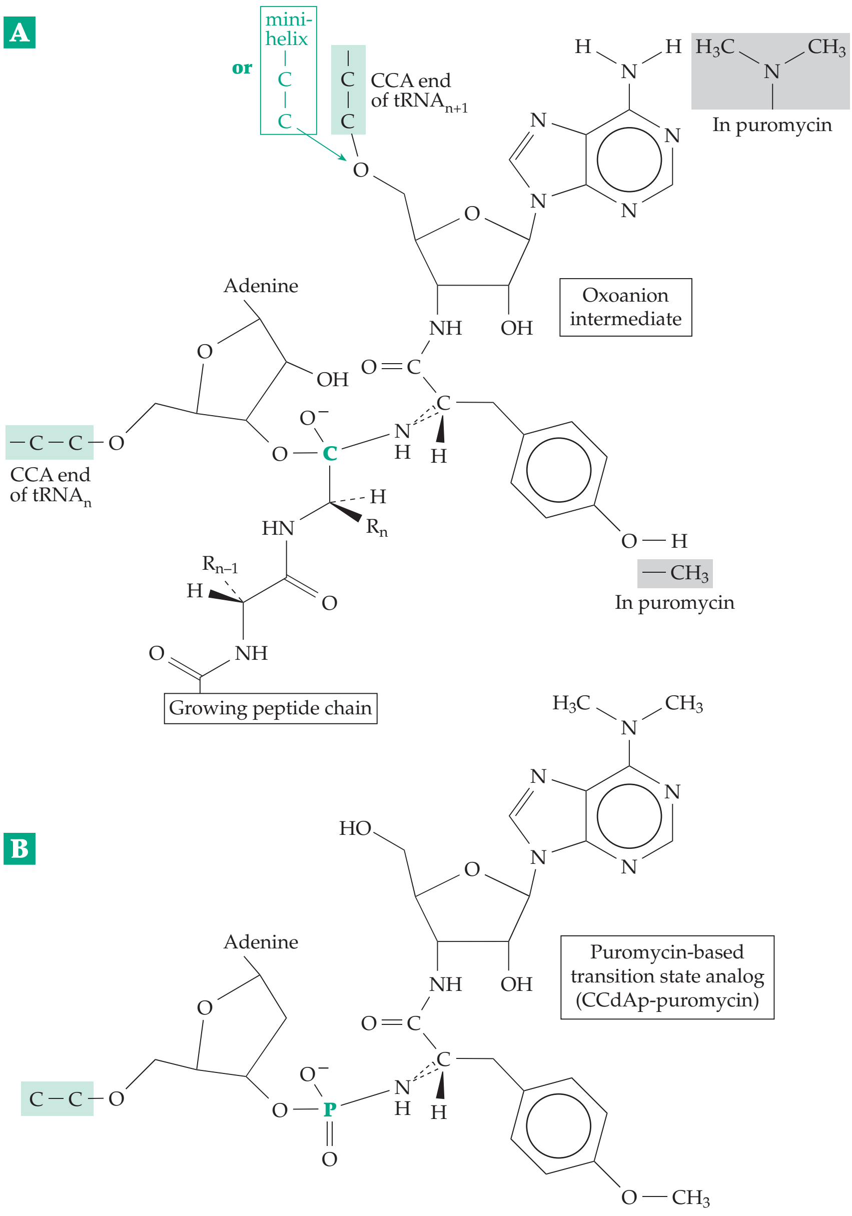


Figure 29-13 (A) Structure of expected intermediate with tetrahedral C-atom in peptidyltransferase reaction with a tRNA, with a minihelix analog, or with the antibiotic puromycin. (B) Transition-state (or bisubstrate) analog formed with puromycin and a mimic of the CCA end of a tRNA. See Box 29-B.

A careful stereochemical analysis has led to the conclusion that for all of the different aminoacyl groups to be able to react in the same way at the peptidyltransferase site and to all generate trans amide linkages, the torsion angles ϕ and ψ of the resulting peptide must be approximately those of an α helix.[388] Thus, the peptide emerging from the ribosome exit tunnel may be largely helical.

Elongation factor EF-G and translocation. The third step in the elongation sequence on ribosomes (Fig. 29-12, step *g*) depends upon **EF-G**, a monomeric GTP-binding protein with a sequence homologous with that of other members of the G protein family. It apparently utilizes the Gibbs energy of hydrolysis of GTP to GDP to drive translocation of the peptidyl-tRNA from the A site to the P site (Fig. 29-12) and of the previously utilized (deacylated) tRNA to the exit site. EF-G binds to the 50S ribosomal subunit at the base of the L7 / 12 stalk as indicated in Fig. 29-1.[392,393] It competes with EF-Tu, which binds in nearly the same location.[5] EF-G is a large five-domain GTPase. Domain 1 contains the GTPase site and resembles other G proteins, and domain 2 has some similarity to the

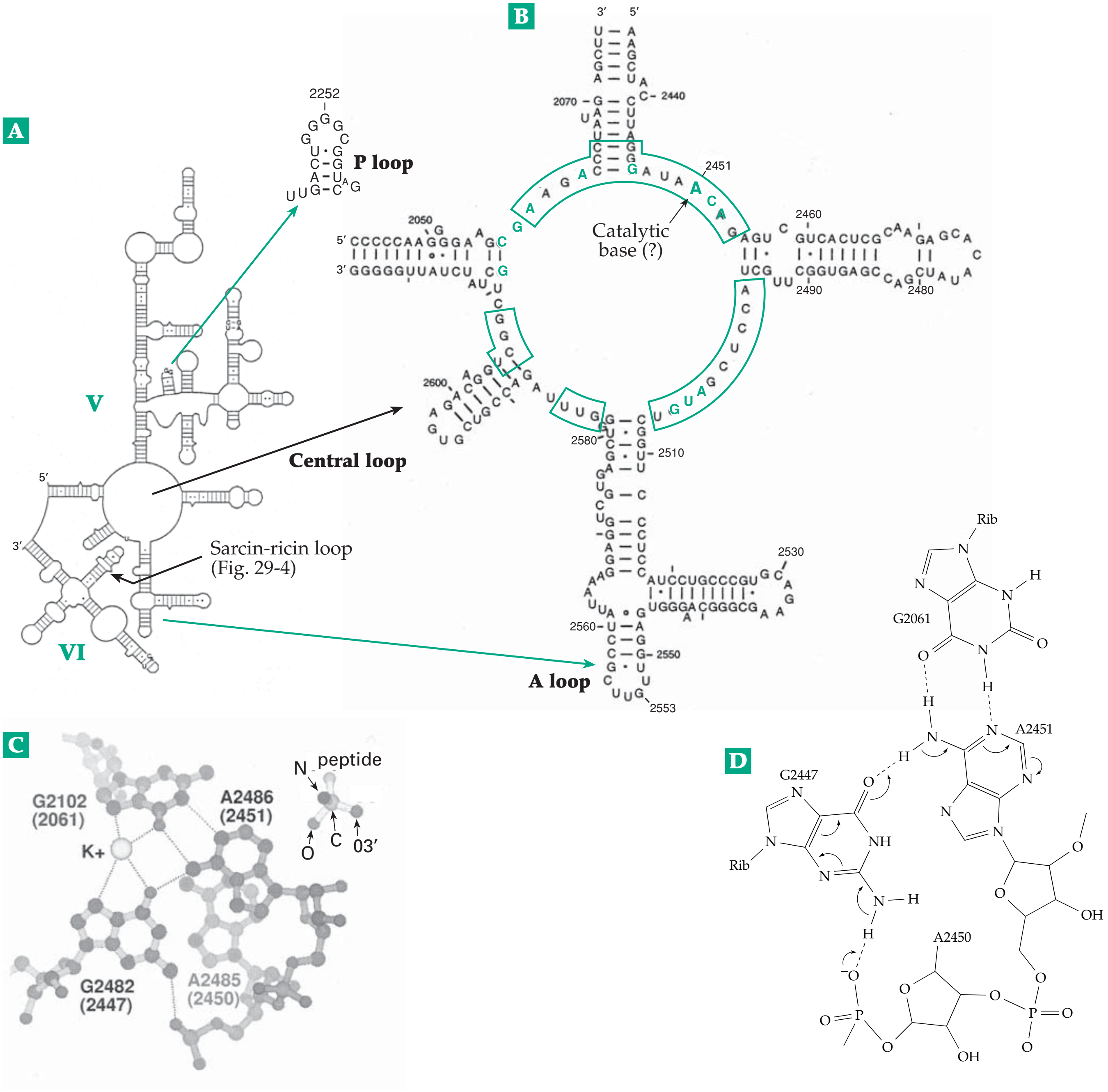
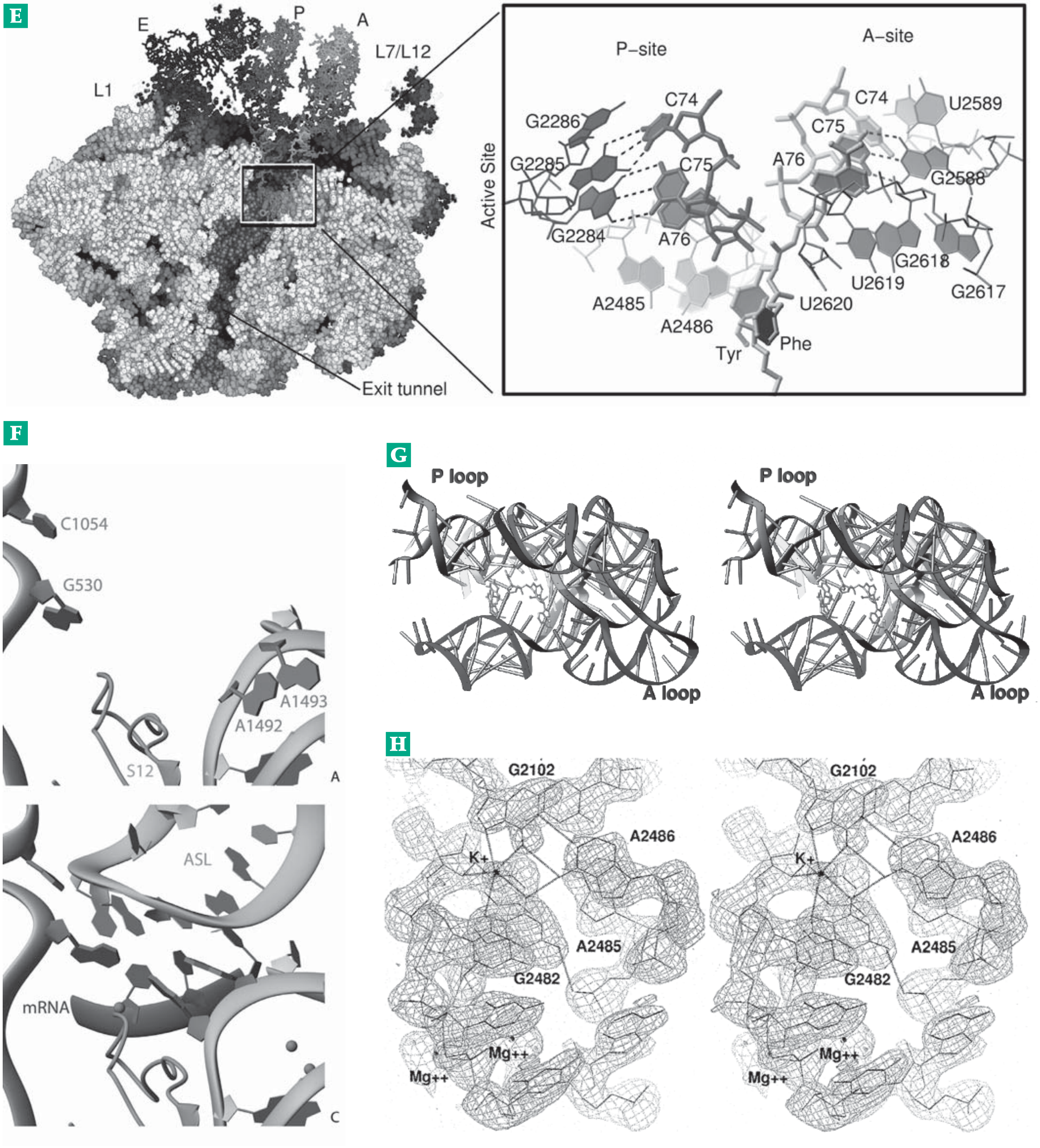

A
B
C
D
2252
P loop
V
VI
5′
3′
Central loop
Sarcin-ricin loop
(Fig. 29-4)
A loop
3′ 5′
2070
2440
2451
2050
2460
2490
2480
Catalytic
base (?)
2600
2580
2510
2530
2560
2550
2553
G2102
(2061)
A2486
(2451)
K+
G2482
(2447)
A2485
(2450)
N peptide
O
C
03′
Rib
G2061
A2451
G2447
A2450
NH
OH

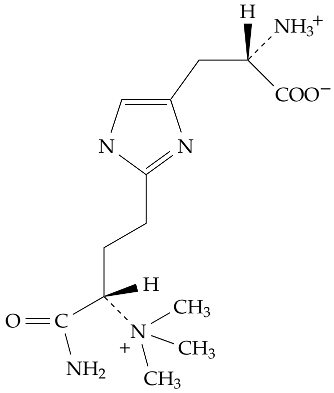
Figure 29-14 The ribosomal peptidyltransferase center in the 23S RNA. (A) Secondary structure map for domains V and VI of *E. coli* 23S RNA. From Samaha *et al.*[389] (B) Sequences of the P loop and A loops and of the central loop of domain V and of the 23S RNA of *Halobacterium halobium* with numbering according to the *E. coli* sequence. Sequences within the green boxes are highly conserved in eubacteria, chloroplast, and mitochondrial RNAs. See Kloss *et al.*[390] and Garret.[80] Sites of mutations that confer resistance to erythromycin (G2057, A2058, and C2611) and chloramphenicol (G2057, G2447, A2451, C2452, A2503, and U2504) are indicated. See Douthwaite and Aagaard.[383] (C) A skeletal representation of the peptidyltransferase active site which is depicted more completely in the stereoscopic view in (H). Included is the peptidyl group on the 2'-end of a tRNA. (D) Schematic region of domain V showing the hydrogen-bonding interactions of the catalytic base A2486 (2451, *E. coli*) with neighboring bases and also locations of nearby K^+ and Mg^{2+} ions. (C, D, G, H) are from Nissen *et al.*[19] (E) Interaction of the CCA 3'-ends of ribosome-bound tRNAs (at top) with the large ribosomal subunit. This is a cutaway view with tRNAs in the A, D, and E sites. The ribosome is sliced to show the acceptor ends of the tRNAs in the entrance to the peptide tunnel. Enlarged section shows interactions of the CCA 3'-ends in the P and A sites. The CCA end of a deacylated tRNA is shown in the P site, while a peptidyl-tRNA analog (CCA puromycin–phenylalanine–caproic acid–biotin) is shown in the A site. Bases of the 23S rRNA are numbered as in *H. marismortui*. Bases A2486 and U2620 (corresponding to A2451 and U2585) are closest to the newly formed peptide bond. From Schmeing *et al.*[33f] (F) Discrete states of the A site of the 30S subunit, as deduced from four different crystal structures. The tRNA, anticodon stem-loop (ASL), A-site mRNA codon, P-site mRNA, protein S12, and


E
E
P
A
L7/L12
L1
Exit tunnel
P–site
A–site
Active Site
C74
G2286
G2285
C75
G2284
A76
A2485
A2486
Tyr
Phe
U2620
U2619
G2618
G2617
G2588
U2589
C74
C75
A76
F
C1054
G530
A1493
A1492
S12
A
ASL
mRNA
C
G
P loop
P loop
A loop
A loop
H
G2102
A2486
K+
A2485
G2482
Mg++
Mg++
G2102
A2486
K+
A2485
G2482
Mg++
Mg++

important bases involved in conformational changes are shown. A few elements of the 16S RNA such as helix 44 (lower right), 530 loop (mid and lower left), and helix 34 (upper left) are also shown. At the top, the native 30S subunit. A1492 and A1493 have been stacked in the interior of H44 and G530 is in the *syn* conformation. C1054 is shown in the upper left corner. Below, when the codon and cognate tRNA-ASL bind in the A site, A1492 and A1493 flip out to monitor the codon-anticodon interaction, and G530 switches to the *anti* conformation to interact with A1492, the anticodon in the second position, and the codon in the third. Two Mg^{2+} ions are visible near the region vacated by A1492 and A1493 in the interior of helix 44, and one is located close to the ribose of the codon in the wobble position. From Ogle *et al.*[33d] Courtesy of Venki Ramakrishnan. (G) Three-dimensional structure of the active-site region of domain V including the P loop[391] and A loop and active-site region. (H) Three-dimensional structure of the active-site region showing the hydrogen-bonding interactions of the catalytic base A 2486 (2451, *E. coli*) with neighboring bases and also locations of nearby K^+ and Mg^{2+} ions. (C, G, H) are from Nissen *et al.*[19]

corresponding domain in EF-Tu. GTP hydrolysis appears to induce within EF-G a conformational change that is coupled directly to a rapid mechanical movement within the ribosome.[394,394a] The large domain 4 protrudes from the molecule and by its shape mimics a molecule of tRNA, and the complete EF-G molecule bears a striking resemblance to an EF-Tu•aminoacyl-tRNA complex.[395,396] It leaves chemical footprints around position 1067 (the thiostrepton-binding site) in domain II of the 23S RNA and in the universally conserved loop around position 2660.[385]

Factor EF-G from eukaryotes (eEF2) is similar to the bacterial protein, but its interaction with the larger eukaryotic ribosomes seems to be more complex. For example, interaction with the ribosomal stalk is more extensive.[37] EF2 contains a single modified histidine called **diphthamide**.[397] This amino acid is not found in other proteins but is always present in eukaryotic EF2 and also in EF-G from archaeobacteria. It is the site of modification by diphtheria toxin (Box 29-A).



Diphthamide

The ribosomal translocation process is quite complex. As the tRNAs move from A to P to E sites on the 16S RNA platform, the mRNA must also move in discrete single-codon steps. The acceptor stems of the tRNAs in the A and P sites must react at the appropriate times in the peptidyltransferase center. Study of protection from chemical probes suggests that tRNAs sometimes lie with the anticodon loop in the A site of the small ribosomal subunit, while the acceptor stem is in the P site of the large subunit (an A/P site as illustrated in Fig. 29-12B). Each aminoacyl-tRNA enters as a complex with EF-Tu and may initially bind with its anticodon in the A site and the acceptor stem with attached EF-Tu in a transient T site, the composite state being A/T. After loss of EF-Tu the acceptor stem can move into the A site to give an A/A state. The peptidyltransferase reaction itself necessarily involves movement at the acceptor stems by 0.1 nm or more. However, additional movement of ~1 nm is needed to move the two tRNAs into states A/P and P/E, respectively. Movement of the mRNA then moves the anticodon loops of the peptidyl-tRNA into the P/P state and of the deacylated tRNA completely into the exit site (Fig. 29-12B).[86,397a] Translocation may occur at different times in the 30S and 50S subunits. The pathway of the mRNA through the ribosome is known[397b,c] and is indicated approximately in Fig. 29-12B.

EF-G seems to be the motor protein that drives translocation in the 30S subunit. When it is not attached to a ribosome, the EF-G•GTP complex is very stable, but in its functioning location GTP is hydrolyzed rapidly. This occurs prior to translocation[398,398a] and presumably causes an internal alteration in the ribosome that energizes it for the translocation step. G proteins usually undergo large changes in conformation when GTP is hydrolyzed (Chapter 11). A very large change is observed for EF-Tu, but such a change has not been seen for EF-G. However, large conformational changes in the ribosome do evidently accompany translocation.[399] The hydrolytic activity of EF-G is stimulated strongly by its binding to the L7/L12 stalk proteins.[400] Eukaryotic EF2 like EF-G binds to the stalk proteins (P1, P2, P0 complex) and also to domains of 28S rRNA equivalent to the bacterial 1067 and sarcin / ricin loops.[401] However, EF-G and EF2 are not functionally interchangeable.

Translocation occurs slowly even in the absence of GTP. However, it is greatly enhanced by GTP hydrolysis.[394a,402] Even empty ribosome, without tRNAs, but in the presence of GTP and EF-Tu and EF-G, hydrolyze GTP. The ribosome may sequentially bind EF-Tu, then EF-G, oscillating between two differing states.[403] The movement of mRNA through the ribosome has been plotted using a variety of immunochemical, crosslinking, and chemical footprinting methods.[52,404–407]

A third elongation factor, eEF3, which is an ATPase, is required by yeast and fungi.[408–410] The 1044-residue yeast protein may be required for ATP-dependent release of deacylated tRNA from the exit site.

Polyribosomes. Under suitable conditions ribosomes isolated from cells are found to sediment together in clusters, often of five or more. These **polyribosomes** (or **polysomes**), which can be seen in electron micrographs (Fig. 28-5), are held together by chains of mRNA. Polyribosomes arise because a single mRNA molecule is being translated by several ribosomes at once. As the 5' terminus of the mRNA emerges from one ribosome, it may soon combine with another and initiate translation of a second peptide chain, etc. The length of the mRNA determines how many ribosomes are likely to be associated in a polyribosome.

Rates of synthesis of ribosomes and of proteins. In a rapidly growing yeast cell with a generation time of ~100 min there are nearly 200,000 ribosomes. Almost 200 new ribosomes must be formed in one minute. Each of the 150 tandemly repeated ribosomal RNA genes must be transcribed into the 4560 nucleotides of one ribosome in less than one minute. The ~150 nuclear pores must import nearly 1000 ribosomal proteins per minute and must export ~25 ribosomal subunits per minute.[410a] The ribosomes that are formed can at 37°C add 14–17 amino acids per second to a growing polypeptide chain,[410b,c] while eukaryotic ribosomes can add 2–4 amino acids per second.[410d,e]

3. Termination of Polypeptide Synthesis

A ribosome faithfully translates the genetic message, adding amino acids to the polypeptide chain until a stop codon is reached. Then a **termination** or **release factor** acts, probably by binding directly to the stop codon on the mRNA in the A site.[411–413b] In *E. coli* termination factor **RF1**, a 47-kDa protein, recognizes UAA or UAG, while **RF2**, a very similar protein,[414,415] recognizes UAA or UGA. There are several hundred molecules per cell of these release factors. They not only recognize the stop codons but also catalyze the hydrolytic removal of the peptidyl chain from the tRNA in the ribosomal P site. They bind into the A site, where they may interact with mRNA bases in addition to those of the stop codon.[415,416] Hydrolytic release of the polypeptide chain from the tRNA in the P site may represent a change in specificity of the peptidyltransferase center induced by binding of a release factor. Genes are often terminated by a succession of two stop codons. Thus, there is a safety factor that prevents translation from continuing in case the first stop codon is missed. An example is provided by the *I* gene of the *lac* operon of *E. coli* (Fig. 28-2), which has a second stop codon in phase with the TGA codon marked in the figure and located five codons further "downstream" (to the right). A third release factor **RF3**[416] is a GTP-binding protein resembling EF-G. It is not essential to life for *E. coli*, but it accelerates the release of RF1 or RF2 and is needed for rapid growth.[417,418] Eukaryotes contain one release factor **eRF1**, which recognizes all three termination codons, and a second release factor **eRF3**, which binds and hydrolyzes GTP.[413,413b,418a,419]

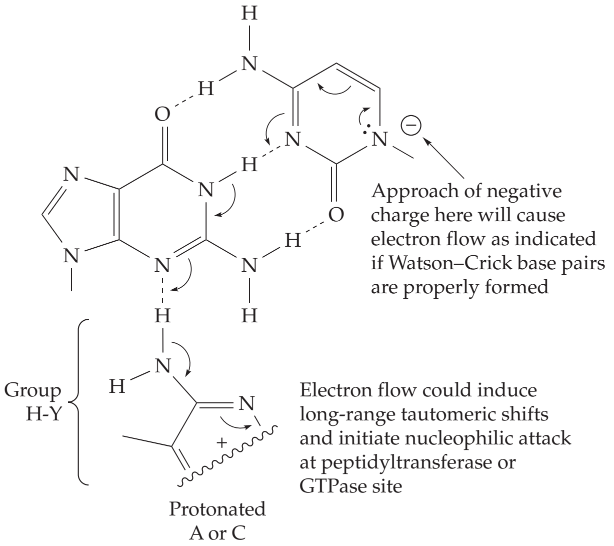
Just as elongation factor EF-G mimicks the aminoacyl-tRNA•Tu complex, release factors RF1 and RF2, in their shapes, mimick molecules of tRNA.[419a] One domain of human eRF1 has an anticodon-recognition domain and a conserved GGQ sequence in a second domain, which mimicks the amino-acceptor arm of tRNA.[419,419b] Mutations in either eRF1 or eRF3 affect translational accuracy and may allow "read-through" of stop codons. In yeast (*S. cerevisiae*) a 685-residue subunit of eRF3 has an N-terminal domain, that like the human **prion protein** (Box 29-E) is capable of being transformed into a self-seeding amyloid-like conformation. In the yeast the formation of amyloid aggregate leads to depletion of the termination factor and increased readthrough.[420–422]

Recycling factors. Even though release factors remove the completed polypeptide chain, a ribosome is not ready for reuse until the deacylated tRNA in the P site is removed and the mRNA is released. This depends upon **ribosome recycling factor** (**RRF**) together with EF-G. The recycling factor is also a tRNA mimic.[419a,423-424b] It may bind into the empty A site, and in an action similar to that of the translocation step of elongation remove the P site tRNA.[417,423–425] However, probing with hydroxyl radicals indicates a different mode of binding.[425a]

4. Preventing and Correcting Errors in Translation on the Ribosome

The wrong amino acid is inserted into most positions in a protein about one time in ~10^4, a frequency[361,426,427] of ~10^{-4}. However, in *E. coli* misreading of certain codons is observed more often. For example, AAU (Asn) is read as AAA (Lys) with a frequency[428] of ~5×10^{-3}. Misreading also depends upon adjacent codons, i.e., the codon context.[429] Having all of the tRNAs charged with the correct amino acids, as discussed in Section B3, is a first essential for accurate translation. A second is finding the correct location of the initiation codon and binding of the aminoacyl-initiator tRNA into the P site. The decoding process by which the correct aminoacyl-tRNA is brought into the A site is still not fully understood. It has often been proposed that (as in DNA replication; Chapter 27, Section C,2) the fidelity of this process depends upon two consecutive recognition steps.[265,266] The first is the binding of the complex of EF-Tu•GTP and the charged tRNA to the ribosome. The second may be associated with the conformational change that locks the aminoacyl-tRNA into the A site and perhaps sends to the peptidyltransferase center a signal that the correct codon-anticodon pairing has been achieved. Some checking is done in the first step. For example, many of the 380 possible mischarged forms of aminoacyl-tRNAs that may have escaped previous proofreading steps (pp. 1695–1696) are rejected because they bind too loosely or too tightly to EF-Tu.[429a,b] Codon-anticodon base pairing may also be checked in the P site[429c] after translocation. The P site is buried deep in a cleft in the RNA of the large subunit. It is designed to hold the mRNAs in a kinked conformation (Fig. 29-1E) with the codons in the A and P sites oriented differently. Some tRNA residues required for high-fidelity participate in imposing this geometry. Mutants in either of the major rRNAs or in tRNAs can lead to loss of fidelity in base pairing and sometimes to excessive frame-shifting.[429c] During the proofreading process a mispaired aminoacyl-tRNA may be allowed to dissociate and be replaced by a new one. Certain mutations, such as those in ribosomal protein S12 that lead to streptomycin resistance, cause greatly increased fidelity of protein synthesis. However, these mutations slow bacterial growth,[426,430] perhaps because some misreading is necessary for synthesis of minor essential proteins.

A strictly hypothetical way in which an alteration in hydrogen bonding could be used to signal the achievement of correct base pairing is illustrated in Fig. 29-15. As indicated by the curved arrows, the approach of a negatively charged group could induce an electron pair to move from the ring nitrogen on the right side. If the pairs of hydrogen bonds were correct, a concerted flow of electrons could take place across the base pair and out into group H-Y and beyond through the postulated tautomeric chain. If the base pair were not correctly formed, the signal could not be transmitted, except during an occasional mispairing with a minor tautomer. Note that another reciprocal electron transfer in the opposite direction to that shown in the figure is also possible through the same base pair. Similar tautomeric shifts are possible for all legitimate base pairs. Initiation of a signal of the type shown could also occur by the addition of some nucleophile to a purine or pyrimidine ring, e.g., to C-6 of the cytosine ring as in Eq. 29-3. In ribosomes such electronic signals could be passed in turn through each of the base pairs involved in codon–anticodon recognition and also through other base pairs formed within loops of ribosomal RNA. If group H-Y is connected by a suitable chain of hydrogen bonds that passes through the active site of the peptidyltransferase, coupling between the recognition signal and the formation of the transition state might be accomplished. Since changes in hydrogen bonding can trigger conformational alterations, the sensing of correct hydrogen bonding could increase the rate of the peptidyltransferase reaction as has been observed experimentally.[266] Base pairing in both the A and P sites may be sensed in similar ways. The observation that rRNA residues hydrogen-bond with groups in the minor grooves of base pairs[33d] seems to be consistent with the proposal of Fig. 29-15.



Approach of negative charge here will cause electron flow as indicated if Watson–Crick base pairs are properly formed

Electron flow could induce long-range tautomeric shifts and initiate nucleophilic attack at peptidyltransferase or GTPase site

Figure 29-15 Hypothetical scheme by which an electronic signal might be sent through a base pair to initiate the peptidyltransferase reaction. See also Metzler.[431]

Codon usage. The usage of codons in specific mRNAs is not random.[432] For example, in a glyceraldehyde-3-phosphate dehydrogenase gene of yeast >96% of the 1004 codons make use of only 25 of the 61 possible coding triplets. Genes undergoing most rapid transcription are more highly biased toward these 25 than are other genes.[433] Many other evolutionary factors have affected usage. These include the need for translational accuracy.[432]

Dealing with lost peptidyl-tRNAs and broken transcripts. Many problems arise during protein synthesis. For example, a peptidyl-tRNA may become detached from a ribosome. In *E. coli* this seems to happen most frequently with peptidyl-$tRNA^{Lys}$. A 193-residue **peptidyl-tRNA hydrolase** is essential for life![434–436] It releases the tRNA for reuse, recycling all peptidyl-tRNAs other than formylmethionyl-tRNA. Perhaps the enzyme is essential because detached peptidyl-tRNAs are toxic, but it is more likely to be to avoid a shortage of free $tRNA^{Lys}$.[434]

If a ribosome starts to synthesize a protein using a broken mRNA, it will reach the end of the mRNA but will not find a termination codon. The peptidyl-tRNA will eventually fall off, but the ribosome will be stalled temporarily. Eukaryotes try to prevent this problem by sending only intact mRNAs to ribosomes,[436a,b] but bacteria have a ribosome rescue system that also tags the partially formed protein on the stalled ribosome for rapid proteolytic degradation.[437–440c] Bacteria synthesize a special 362-residue RNA that resembles a tRNA but also contains a short mRNA-like module that codes for the 11-residue peptide tag AANDENYALAA. This hybrid tRNA-mRNA, which is designated tmRNA (or *ssrA* RNA), mimicks $tRNA^{Ala}$ and is recognized and charged by alanyl-tRNA synthetase. The resulting $tmRNA^{Ala}$ binds into the A site of the stalled ribosome, undergoes the peptidyltransferase and translocation steps (Fig. 29-16). The old mRNA is released, the mRNA-like sequence of tmRNA becomes seated, and translation of the new tail sequence follows. The tail sequence is similar to C-terminal sequences that are known to mark other proteins for rapid proteolytic degradation. An associated protein SmpB is also required for functioning of the tmRNA system.[438a,b]

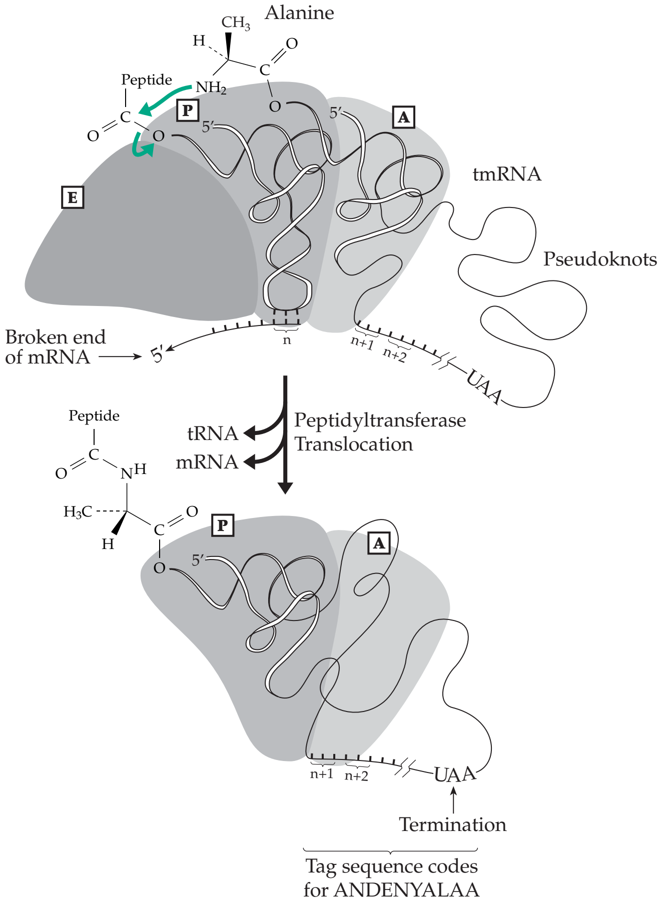


Figure 29-16 Schematic diagram of the tmRNA structure and its function in the rescue of ribosomes stalled at the end of a messenger RNA that has been broken and has lost its in-frame termination codon. After it binds into the ribosomal A site the tmRNA, which has been charged with alanine, undergoes the peptidyltransferase reaction and translocation to the P site. Then it lays down its mRNA-like coding sequence, which is used by the ribosome to add ten more amino acids to form the 11-residue C-terminal degradation signal AANDENYALAA. This induces rapid degradation of the imperfect protein that has been formed.[436a]

5. Suppressor Genes

The suppression of nonsense mutations by suppressor genes has been discussed in Chapter 26. The chemical nature of these genes was discerned, in part, from experiments involving transfer of suppressor gene *supF(su3)* into the DNA of a bacteriophage. This DNA was found to specifically hybridize with a minor transfer RNA, $\text{tRNA}_1^{\text{Tyr}}$. Subsequent investigation showed that *sup F* is a structural gene for this tRNA, and that in it the normal 5'-GUA-3' (Tyr) anticodon has been replaced with CUA. The latter can pair with the chain termination codon 5'-UAG-3' (the *amber* codon) permitting the ribosome to insert tyrosine at the site of chain termination signals introduced in *amber* mutations. It may seem puzzling that a tRNA, which prevents chain termination, does not prevent synthesis of other essential proteins within the bacterium. However, suppression is typically less than 30% efficient. Hence, many protein chains terminate normally. Since two chain termination signals are often present in a gene, most protein synthesis in the presence of the small amount of suppressor tRNA present is concluded normally. Premature chain termination caused by selected *amber* mutations will be partially inhibited, permitting the cell to make enough of the missing proteins to survive. The nucleotide sequence of a further mutated *supF* tRNA and of its longer precursor is shown in Fig. 28-10. Several other suppressor genes have also been identified as specific tRNA structural genes.[441]

A suppressor of frame-shift mutations in *Salmonella* is a tRNA containing at the anticodon position the nucleotide quartet CCCC instead of the usual CCC triplet anticodon.[442,443] It has eight unpaired bases in the anticodon loop instead of the usual seven. Other frame-shift repressor tRNAs have been identified in *E. coli*,[444] *Salmonella*, and yeast.[445] Not all suppressor genes encode tRNAs. For example, a UGA suppressor from *E. coli* is a mutant 16S rRNA from which C1054 has been deleted.[446] A general nonsense suppressor in yeast is homologous to yeast elongation factor EF-1α as well as to *E. coli* EF-Tu.[447]

Among other suppressor genes present in eukaryotic organisms[448] are mammalian genes encoding serine tRNAs that are *opal* (UGA) suppressors. These and other eukaryotic suppressor tRNAs have specific and important normal functions in cells. For example, a specific kinase phosphorylates the *opal* suppressor seryl-tRNA to its phosphoseryl derivative.[448] This suppressor tRNAs may sometimes be responsible for introducing phosphoserine at specific positions in proteins. An *opal* suppressor is also used for the introduction of selenocysteine. An *amber* suppressor is used by some methane-forming Archaea to introduce **pyrrolysine** into specific sites in methyltransferases. In pyrrolysine the epsilon amino group of lysine is joined by an amide linkage to a derivative of pyrroline-5-carboxylate (p. 1374).[448a,b]

Selenocysteine (Sec) Selenocysteine is incorporated into a small number of proteins in species from all three kingdoms of life by a suppressor tRNA^{Sec} that reads certain UGA codons, which are marked as representing selenocysteine.[449,450] The selenocysteinyl-tRNA is made from a seryl-tRNA (Eq. 29-7) as described further in Chapters 16 and 24. In *E. coli* selenocysteine is present in three proteins, all formate dehydrogenases. The archaeon *Methanococcus jannaschii* contains genes for seven selenocysteine-containing proteins. Only one Sec-containing protein has been found in the nematode *Caenorhabditis elegans* and none in the yeast

Saccharomyces cerevisiae. However, there are at least 14 in the human body.[451,452] One of these, selenoprotein P, contains ten selenocysteine residues.[452,453] Products of four special genes are needed for incorporation of selenocysteine into *E. coli* proteins.[454,455] *Sel C* encodes the special $\mathit{tRNA}^{\mathrm{Sec}}$, which becomes charged with selenocysteine.[456] *Sel D* encodes selenophosphate synthetase and *Sel A* selenocysteine synthetase (Eq. 29-7). *Sel B* encodes a special elongation factor, which resembles EF-Tu but has an extra domain that binds to an mRNA segment known as the **SECIS** (selenocysteine insertion sequence).[457–461a] The SECIS sequence follows the 3' end of the UGA termination codon. It is a 40-nt segment that is able to form a stem-loop structure. However, in archaea and in eukaryotes the SECIS sequence lies at the end of the selenoprotein gene in the 3' nontranslated region. It may be some distance away and may function by a foldback mechanism. It recodes the entire message, acting on any in-frame UGA codon.[461] In mammals a special SECIS-binding protein SBP2 is also required.[462]

Expanding the genetic code. Suppressor tRNAs can also be created artificially and are being used in protein engineering. *Amber*, *ochre*, or *opal* chain termination mutations can be introduced readily at many points in a protein (Chapter 26). Suppressor tRNAs can be made that will then place any one of the possible amino acids into most of the mutated positions.[463] Synthetic amino acids not normally found in proteins can also be incorporated using such tRNAs.[464–467] The TAG(UAG) *amber* stop codon is often used together with a genetically engineered tRNA. In early experiments these techniques were used to create hundreds of mutant forms of the *lac* repressor protein (see Chapter 28, p. 1606). Since then a variety of additional approaches have been explored. Transfer RNAs have been engineered to recognize four-base codons such as AGGU and CGGG.[468,469] Organisms such as *Micrococcus luteus*, in which not all of the available triplet codons have been utilized, allows development of a mutation system using an unassigned codon rather than a stop codon.[468] A general method for site-specific incorporation of any amino acid or amino acid analog requires a suppressor tRNA that is not aminoacylated by any aminoacyl-tRNA synthetase present within the host cell, and also an aminoacyl-tRNA synthetase that acts only on the suppressor tRNA and no other tRNA in the cell.[470] Several such systems are being developed.[470–472] Another idea is to utilize a 65th codon–anticodon pair, one depending upon a new synthetic nucleoside that can be incorporated into mRNA.[473]

Another possible application of suppressor genes is *in vivo* suppression of undesirable termination codons. An example comes from a β^{o} thalassemia caused by mutation of lysine codon CAG to UAG. By changing the anticodon of a human tRNA$^{\mathrm{Lys}}$ gene to an *amber* suppressor, premature termination of globin chains was suppressed in an *in vivo* test.[474,475]

6. Read-Through Translation and Ribosomal Frameshifting

If termination codons are not recognized efficiently by termination factors, synthesis continues past the termination codons and new longer protein chains are made. This **read-through translation**[429,476] may sometimes be accidental, but it is also used by cells to form several important proteins. For example, the 14-kDa coat protein of bacteriophage Qβ is elongated by read-through during translation of the RNA about 4% of the time. This produces a 38-kDa protein known as A_1, which has an extra 200 amino acid residues at the C terminus and is essential for formation of infectious virus particles.

A similar situation is met with retroviruses (Chapter 28) whose coat protein gene *gag* is fused to the reverse transcription polyprotein precursor gene *pol* (Fig. 28-26).[477] In fact, the polyprotein gene overlaps the 3' end of the *gag* gene. Read-through translation allows synthesis of the polyprotein about 5% of the time. However, the *pol* gene is written in a different reading frame: –1 with respect to the *gag* gene. For example, in the HIV genome (Fig. 28-26) the *pol* gene begins at nucleotide 1638 in the –1 reading frame with respect to the *gag* gene. In both the Rous sarcoma virus and HIV a polyprotein that is formed contains peptide sequences from both *gag* and *pol* genes.[478,479] This fusion of the two proteins is accomplished by a frameshift, which occurs on the ribosome as it operates in the region of overlap of the *gag* and *pol* genes. This mechanism allows synthesis of relatively small amounts of the enzymes encoded by *pol* but large amounts of structural proteins encoded by *gag*.[480] Many other examples of essential ribosomal frameshifting are known.[429,477,481–483] For example, the gene for *E. coli* ribosomal release factor RF2 has a UGA termination codon at position 26, but the coding sequence for the protein continues in the +1 frame.[484]

Ribosomal frameshifting can be accounted for by more than one mechanism. It can occur when a four-base anticodon is present in a suppressor tRNA. It can result from incorrect base pairing. If a tRNA slips over by one nucleotide, a single base in the mRNA can be left unpaired with the reading frame being shifted +1. However, most frameshifts are in the –1 direction and occur at specific locations in the mRNA, i.e., they are **programmed frameshifts**.[484a] These often occur at "slippery sites"[485] including the following mRNA sequences in which three codons are marked:

$$\overbrace{\;X}^{1}\;\overbrace{X\;X\;Y}^{2}\;\overbrace{Y\;Y\;Z}^{3}$$

This sequence is followed closely by an element of secondary structure, most often a pseudoknot.[486–492] Eukaryotic frameshifts are almost always in the –1 direction, the exception being found in the mammalian mRNA for **antizyme**, a negative regulator of ornithine decarboxylase (Chapter 24, p. 1382).[493] The frame-shift occurs at an initially in-frame termination codon (UGA), which is followed by a pseudoknot. Most translation is terminated at this stop codon, but frameshifting, which is induced by a high polyamine concentration, allows read-through and synthesis of the antizyme protein. In rare cases frameshifting may lead to **translational bypass** of some codons on the mRNA. Such a case is found in a bacteriophage T4 mRNA for which the *E. coli* ribosomes bypass 50 nucleotides in order to complete the synthesis of a

BOX 29-C NONRIBOSOMAL PEPTIDE SYNTHESIS

Many small biologically active peptides, including hormones and some antibiotics, are synthesized on ribosomes as precursor proteins, which are cut into small pieces and may then be modified in a variety of ways. However, many other peptides including many antibiotics are made without use of ribosomes by large polyfunctional synthetases. The first of these, gramicidin S synthetase, was described by Lipmann and coworkers in 1971.[a] It is discussed on p. 994 as is the mechanism of synthesis. It is now recognized that these enzymes are modular and have much in common with fatty acid synthetases (Fig. 17-12 and p. 1186) and polyketide synthetases (Fig. 21-11). They are able to link not only the amino acids found in proteins but also modified and unusual amino acids. They may also join one or more α-hydroxy acids to a peptide to form a depsipeptide, and they may contain modules that carry out modification reactions such as methylation, acylation, or glycosylation.[b–e] Because of their modular nature they are attractive proteins for genetic engineering.[b,f,g]

Each synthetase module contains three active site domains: The **A domain** catalyzes activation of the amino acid (or hydroxyacid) by formation of an aminoacyl- or hydroxyacyl-adenylate, just as occurs with aminoacyl-tRNA synthetases. However, in three-dimensional structure the A domains do not resemble either of the classes of aminoacyl-tRNA synthetases but are similar to luciferyl adenylate (Eq. 23-46) and acyl-CoA synthetases.[h] The **T-domain** or **peptidyl carrier protein domain** resembles the acyl carrier domains of fatty acid and polyketide synthetases in containing bound phosphopantetheine (Fig. 14-1). Its –SH group, like the CCA-terminal ribosyl –OH group of a tRNA, displaces AMP, transferring the activated amino acid or hydroxy acid to the thiol sulfur of phosphopantetheine. The **C-domain** catalyzes condensation (peptidyl transfer). The first or **initiation module** lacks a C-domain, and the final **termination module** contains an extra termination domain. The process parallels that outlined in Fig. 21-11.[i]

A few of the products of nonribosomal peptide synthesis are gramicidin S (Fig. 2-4), enniatins, bacitracins, and tyrocidines (p. 994),[b,e] vancamycin (Box 20-H),[j] actinomycin (Box 28-A),[k] the siderophore yersiniabactin,[l] surfactin (Fig. 2-4),[m,n] and cyclosporin (Box 9-F).[o,p] The δ -(L-α-aminoadipoyl)-L-cysteinyl-D-valine synthetase,[q,r] which forms the precursor to penicillin and cephalosporins (Box 20-G), also belongs to this group of enzymes as do synthetases that make cyclooctadepsipeptides with antihelminthic activity[s] and many other compounds.[t–v] topoisomerase subunit.[494–495a] Ribosomal protein L9 may play a role in the bypass process.

---

[a] Lipmann, F. (1971) *Science* **173**, 875–884

[b] Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5848–5853

[c] Linne, U., and Marahiel, M. A. (2000) *Biochemistry* **39**, 10439–10447

[d] Guenzi, E., Galli, G., Grgurina, I., Pace, E., Ferranti, P., and Grandi, G. (1998) *J. Biol. Chem.* **273**, 14403–14410

[e] Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999) *Science* **284**, 486–489

[f] Symmank, H., Saenger, W., and Bernhard, F. (1999) *J. Biol. Chem.* **274**, 21581–21588

[g] de Ferra, F., Rodriguez, F., Tortora, O., Tosi, C., and Grandi, G. (1997) *J. Biol. Chem.* **272**, 25304–25309

[h] Pfeifer, E., Pavela-Vrancic, M., von Döhren, H., and Kleinkauf, H. (1995) *Biochemistry* **34**, 7450–7459

[i] Cane, D. E., Walsh, C. T., and Khosla, C. (1998) *Science* **282**, 63–68

[j] Trauger, J. W., and Walsh, C. T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3112–3117

[k] Pfennig, F., Schauwecker, F., and Keller, U. (1999) *J. Biol. Chem.* **274**, 12508–12516

[l] Miller, D. A., and Walsh, C. T. (2001) *Biochemistry* **40**, 5313–5321

[m] Weinreb, P. H., Quadri, L. E. N., Walsh, C. T., and Zuber, P. (1998) *Biochemistry* **37**, 1575–1584

[n] Reuter, K., Mofid, M. R., Marahiel, M. A., and Ficner, R. (1999) *EMBO J.* **18**, 6823–6831

[o] Lawen, A., and Traber, R. (1993) *J. Biol. Chem.* **268**, 20452–20465

[p] Hoffmann, K., Schneider-Scherzer, E., Kleinkauf, H., and Zocher, R. (1994) *J. Biol. Chem.* **269**, 12710–12714

[q] Shiau, C.-Y., Byford, M. F., Aplin, R. T., Baldwin, J. E., and Schofield, C. J. (1997) *Biochemistry* **36**, 8798–8806

[r] Kallow, W., Kennedy, J., Arezi, B., Turner, G., and von Döhren, H. (2000) *J. Mol. Biol.* **297**, 395–408

[s] Weckwerth, W., Miyamoto, K., Iinuma, K., Krause, M., Glinski, M., Storm, T., Bonse, G., Kleinkauf, H., and Zocher, R. (2000) *J. Biol. Chem.* **275**, 17909–17915

[t] Milne, J. C., Roy, R. S., Eliot, A. C., Kelleher, N. L., Wokhlu, A., Nickels, B., and Walsh, C. T. (1999) *Biochemistry* **38**, 4768–4781

[u] Gaitatzis, N., Kunze, B., and Müller, R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11136–11141

[v] Gewolb, J. (2002) *Science* **295**, 2205–2207

7. RNA Viruses

The RNA-containing bacteriophages are convenient sources of relatively simple mRNA molecules, whose sequences can be studied.[496] The genetic information for these viruses is carried by RNA molecules consisting of only 3500–4500 nucleotides and which may contain only four genes (p. 247). The RNA from phages f2, R17, MS2, and the more distant Qβ have been studied intensively.[483,497]

Parts of the 2569-nucleotide sequence for the RNA of phage MS2[498] are shown in Fig. 29-17. The 5' end (upper left center) still bears the triphosphate group of the initiating GTP. Following a number of hairpin loops there is a ribosome-protected region, which begins with the initiation codon GUG for the A protein

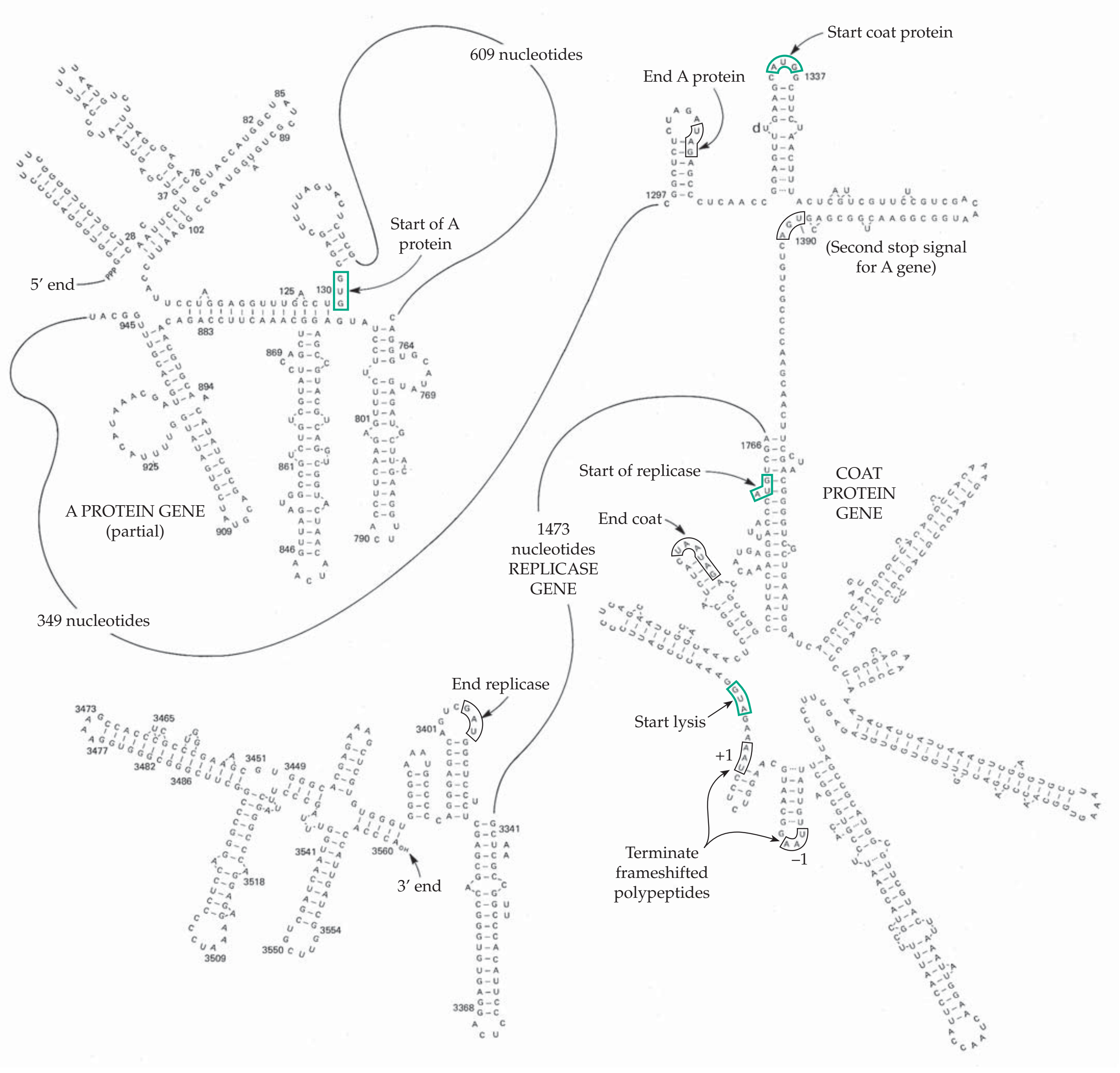

Start coat protein
1337
End A protein
609 nucleotides
1297
5′ end
Start of A protein
125
130
945
883
869
764
769
894
801
925
861
790
909
846
1390
(Second stop signal for A gene)
1766
Start of replicase
COAT PROTEIN GENE
End coat
1473 nucleotides REPLICASE GENE
A PROTEIN GENE (partial)
349 nucleotides
End replicase
Start lysis
+1
−1
Terminate frameshifted polypeptides
3473
3465
3477
3482
3486
3451
3449
3401
3341
3541
3560
3′ end
3518
3554
3550
3509
3368

Figure 29-17 Partial sequence and secondary structure model of RNA of bacteriophage MS2. The initiation and termination codons for each of the three genes (A protein, coat protein, and replicase) are enclosed in boxes as is the second stop signal that is in-frame for the A protein gene but out-of-frame for the coat protein gene. The entire coat protein gene is shown but less than one-third of the entire sequence is given. From W. Fiers and associates.[499–501]

(enclosed in a box). Here is some of the first direct evidence ever obtained that GUG as well as AUG is a biologically important initiation codon. Following the initiation codon the nucleotide sequence codes exactly for the established amino acid sequence of the protein. The termination codon UAG is also enclosed in a box in the figure. Following this is a short intergenic region, which includes one side of a hairpin loop with the initiator codon AUG for the next gene at the end. The nucleotide sequence following this codes exactly for the experimentally established sequence end of the coat protein.[499] One other feature of the sequence shown is the UGA termination codon in a box shortly after the beginning of the coat protein gene (at position 1390). This termination signal is out of phase with the initiator codon AUG; hence, it does not represent a termination point for the coat gene. However, it is in phase with the UAG termination codon for the A protein. In the presence of various host amber (UAG) suppressor genes, the A protein is elongated and terminated at this UGA codon.

The coat gene, containing only 390 nucleotides, is shown in its entirety. The secondary structure proposed resembles a flower.[499] The gene ends with a double stop signal UAAUAG. Following an intergenic sequence of 36 nucleotides the long replicase gene starts with an AUG codon. It ends at position 3395 leaving an untranslated segment of 174 nucleotides at the 3' end.

Initially it was thought that MS2 RNA contained only the three genes mentioned in the preceding paragraphs, but later it was found to have an additional gene required for lysis of the host cell.[483] The initiation codon for this gene begins at position 1678 (Fig. 29-17) in the +1 reading frame. There is a UAA stop just two codons before this in the same frame and another UAA stop codon in the −1 frame beginning at position 1652. As a result of these stop codons any reading frameshift during synthesis of the coat protein yields mistranslated proteins that are terminated at these codons. There is no Shine–Dalgarno sequence to bind ribosomes for initiation in this region, but because the initiation codon for the lysis (L) protein is nearby, reinitiation occurs and the L protein is made in the relatively small amounts needed. This arrangement permits efficient use of the RNA by making use of overlapping genes. It also ensures that enough coat protein has been synthesized to make new virus particles before the L protein accumulation causes lysis.[483]

Many viral RNAs that are formed within eukaryotic cells lack a 5' cap. They depend upon internal ribosomal entry sites (IRESs). This has been studied most with picorna viruses.[338,502,503] These viruses not only initiate translation at discrete sites in uncapped RNA but carry out a proteolytic cleavage of initiation factor 4G (Fig. 29-11), which seems to be necessary for initiation of viral-RNA translation.[338,504] The IRES region of hepatitis C viral RNA contains a complex pseudoknotted secondary structure that is necessary for initiation.[346,505,506] Cryo-EM reveals a pronounced change in the 40S ribosomal subunit structure when the viral IRES binds.[346] Some RNA viruses of plants have complex secondary structures in the untranslated 3' region that promote efficient initiation of translation.[506a]

8. Other Functions of Ribosomes

In addition to making proteins, ribosomes also participate in regulatory mechanisms that influence the entire cell. One such mechanism is seen in the **stringent response**.[507–510] Many amino acid-requiring auxotrophs of *E. coli* and other bacteria, when deprived of an essential amino acid, respond by decreasing their production of ribosomal RNA, ribosomal proteins, purine nucleoside triphosphates, lipids, and other essential materials. However, mutations in the gene *rel* (relaxed) lead to continued production of rRNA even in the absence of an essential amino acid. (The stringent response is "relaxed.") It was observed that the **guanosine polyphosphates ppGpp** and **pppGpp**, originally termed MS or "magic spot" compounds, accumulate in stringent (rel^+) strains to a concentration of ~1 mM but not in relaxed (rel^-) strains. Guanosine polyphosphates are synthesized on the ribosomes by transfer of a pyrophospho group from ATP (Eq. 29-11):

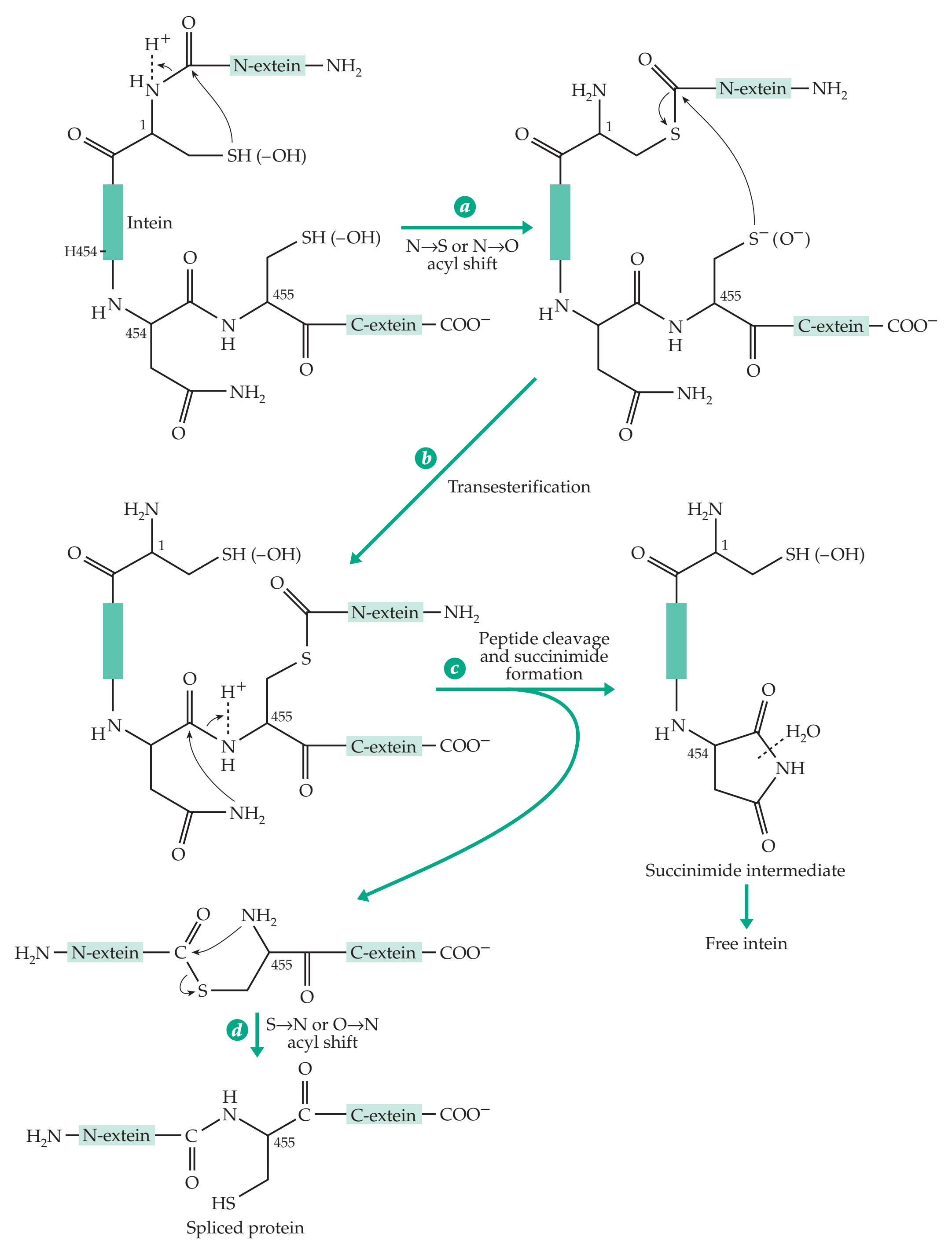
$$\text{ATP} + \text{GDP(GTP)} \longrightarrow \text{ppGpp(pppGpp)} + \text{AMP} \tag{29-11}$$

The reaction is catalyzed by the 84-kDa ppGpp synthetase (**stringent factor**), which is encoded by the *rel* gene and is present only in stringent strains.[511–512a] It binds to ribosomes and becomes active only if mRNA is bound to the ribosomes and if codon-selected uncharged tRNA is present in the A sites. A second ppGpp synthetase (PSII) is encoded by gene **spoT**, which also codes for a ppGpp hydrolase.[510] The presence of an uncharged tRNA in the ribosomal A site is expected during amino acid starvation. The stringent factor competes with elongation factor EF-G for its ribosomal site.[513]

The most important effect of accumulating ppGpp may be to bind to an allosteric site on RNA polymerase.[509] The ppGpp-polymerase complex appears to be inefficient in initiating transcription of genes for rRNA, other stable RNAs, and ribosomal proteins. However, it stimulates expression of various amino acid biosynthetic genes and catabolic genes, perhaps via the "discriminator sequence" (Chapter 28; p. 1608). This is not the only effect of ppGpp. The fidelity of translation is decreased when amino acid concentrations fall and

BOX 29-D PROTEIN SPLICING, INTEINS, AND HOMING ENDONUCLEASES

Like self-splicing RNAs, which excise introns from their chains, a few proteins are able to splice out segments of their own chains as **inteins**. The surrounding protein sequences can be referred to as **exteins**. Over 100 self-splicing proteins are known. They are found in all kingdoms of life.[a–d] The inteins, which are excised, are typically 50 kDa in size but range from ~360 to over 500 residues. The mechanism of splicing is related to the chemistry of pyruvoyl enzyme activation (Eq. 14-41), succinimide formation from asparagine residues (Eq. 2-24), and protein carboxymethylation (Box 12-A). The intein always contains serine or cysteine in its N-terminal (1)-position and asparagine in its C-terminal position. The latter is always followed by cysteine, serine, or threonine in the N-terminal



Intein
H454
N-extein
C-extein
SH (–OH)
S^- (O^-)
a
N→S or N→O acyl shift
b
Transesterification
c
Peptide cleavage and succinimide formation
d
S→N or O→N acyl shift
H_2O
Succinimide intermediate
Free intein
Spliced protein

BOX 29-D (continued)

position of the C-extein. The penultinate residue in the intein is usually (~90 %) histidine, which is thought to play a catalytic role. Other residues in the catalytic domains, which form the ends of the inteins, may also participate in catalysis.

One of the first inteins discovered was found in the 119-kDa precursor to a subunit of a vacuolar ATPase of yeast.[a,c] In this 50-kDa intein Thr 72, His 75, and His 197 may have catalytic functions.[d] The intein is spliced out to form the 69-kDa subunit. The splicing mechanism, which is illustrated for this intein, is shown in the accompanying equations.[b,d–g] Step *a* is an N→S or N→O acyl shift. This is followed by transesterification (step *b*) which involves either thioesters (as illustrated) or oxygen esters. Formation of a succinmide intermediate (step *c*) releases the intein and the spliced protein. The latter must undergo an S→N or O→N acyl shift (step *d*), and the succinimide in the extein must be hydrolyzed to complete the process.

Why do cells ever splice proteins? It isn't clear. However, a curious fact is that many inteins are **homing endonucleases**.[h–k] The genes for these nucleases are often present in introns in mRNA, and the homing endonuclease often cuts DNA in such a way as to initiate movement of its own gene (Chapter 27). The endonuclease itself is found in the center of the intein between the two end domains, which contain the catalytic centers for the splicing reaction.

A few cases are known in which proteins undergo *trans* splicing. For example, the *dnaE* gene of *Synechocystis*, which codes for DNA polymerase III, is actually two partial genes that are 745 kb apart and on opposite strands of the DNA. One of the partial genes codes for a protein containing the N-terminal splice site for an intein, and the other gene codes for a polypeptide containing the C-terminal splice site. Evidently the two splicing domains associate and then catalyze the splicing sequence in the usual way. Split inteins have become very useful in protein engineering because they can be used to join various polypeptide sequences.[k–m] They have also provided an efficient system for purification of specific proteins.[b,n]

---

[a] Cooper, A. A., and Stevens, T. H. (1995) *Trends Biochem. Sci.* **20**, 351–356

[b] Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B., and Xu, M.-Q. (1996) *J. Biol. Chem.* **271**, 22159–22168

[c] Chong, S., Williams, K. S., Wotkowicz, C., and Xu, M.-Q. (1998) *J. Biol. Chem.* **273**, 10567–10577

[d] Poland, B. W., Xu, M.-Q., and Quiocho, F. A. (2000) *J. Biol. Chem.* **275**, 16408–16413

[e] Clarke, N. D. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11084–11088

[f] Xu, M.-Q., and Perler, F. B. (1996) *EMBO J.* **15**, 5146–5153

[g] Shao, Y., Xu, M.-Q., and Paulus, H. (1996) *Biochemistry* **35**, 3810–3815

[h] Chuprina, V. P., Heinemann, U., Nurislamov, A. A., Zielenkiewicz, P., and Dickerson, R. E. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 593–597

[i] Gimble, F. S., and Wang, J. (1996) *J. Mol. Biol.* **263**, 163 – 180

[j] Pietrokovski, S. (1998) *Protein Sci.* **7**, 64–71

[k] Perler, F. B. (1999) *Trends Biochem. Sci.* **24**, 209–211

[l] Martin, D. D., Xu, M.-Q., and Evans, T. C., Jr. (2001) *Biochemistry* **40**, 1393–1402

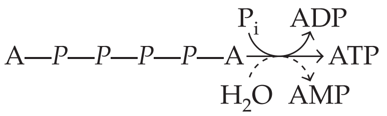
[m] Otomo, T., Ito, N., Kyogoku, Y., and Yamazaki, T. (1999) *Biochemistry* **38**, 16040–16044

[n] Evans, T. C., Jr., Martin, D., Kolly, R., Panne, D., Sun, L., Ghosh, I., Chen, L., Benner, J., Liu, X.-Q., and Xu, M.-Q. (2000) *J. Biol. Chem.* **275**, 9091–9094

---

ribosomal action slows. However, ppGpp apparently binds to the ribosome and slows the binding of the aminoacyl-tRNA•EF-Tu•GTP complex. This allows more time for rejection of mispaired tRNAs and increases the fidelity of translation.[514] Under conditions of nutrient starvation the accumulating ppGpp may promote enzymatic degradation of unneeded proteins[512a,b] and may also induce programmed cell death.[515]

Another "alarmone" that regulates both transcription and DNA replication and other cell functions is diadenosine tetraphosphate (Ap_4A). Effects of Ap_4A and related compounds have been discussed in Chapter 28 (p. 1635). These compounds affect many biological events including replication, growth, and differentiation.[516] However, the synthesis of Ap_4A is a reaction not of ribosomes but of an aminoacyl-tRNA synthetase. An enzyme-bound aminoacyl adenylate carries out adenylylation of ATP rather than aminoacylation of tRNA, especially when Zn^{2+} is present. Ap_4A is abundant in blood platelets, where it is stored in dense granules.[517] Both Ap_2A and AP_3A accumulate as gran- ules in myocardial tissues,[518] and Ap_5A and Ap_6A are also present in adrenal chromaffin cells, in blood plate- lets, and in synaptic vesicles.[519] These compounds are catabolized by hydrolases or in lower eukaryotes by phosphorylases. For example, Ap_4A may be converted into ATP + AMP or converted into ATP and ADP (Eq. 29-12).[516,520]



$$\mathrm{A{-}\mathit{P}{-}\mathit{P}{-}\mathit{P}{-}\mathit{P}{-}A} \xrightarrow[\mathrm{H_2O}\ \ \mathrm{AMP}]{\mathrm{P_i}\ \ \mathrm{ADP}} \mathrm{ATP}$$

(29-12)

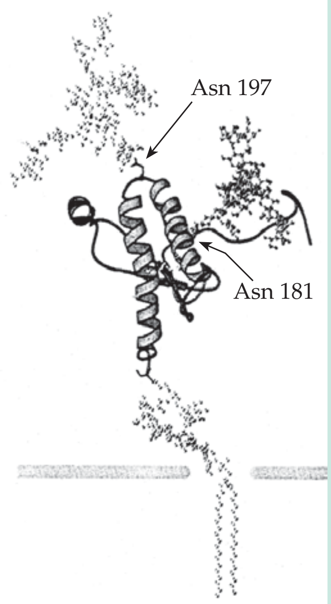
A quite different role of ribosomes is to regulate the life span of certain mRNA molecules. The best studied example is the mRNA for the microtubule

BOX 29-E PRIONS AND AMYLOID DISEASES

The mysterious prions (proteinaceous infective agents), which are described briefly on p. 248, are under intensive investigation. Prion diseases affect fewer than one in 100,000 persons, but there is fear of a possible epidemic. Furthermore, there is a close relationship of prions to a large family of **amyloid diseases**. The most frequent of these is Alzheimer disease, which is estimated to affect one-third of people over 85 years of age in the United States.[a,b]

Prion diseases include **scrapie** of sheep and goats, **bovine spongiform encephalopathy** (**BSE** or mad cow disease), **chronic wasting disease** (**CWD**) of deer and elk, and the human diseases **kuru**, **Creutzfeldt-Jakob disease** (**CJD**), **Gerstmann–Sträussler–Scheinker syndrome** (**GSS**),[c] and **fatal familial insomnia** (**FFI**).[a,d–f] The diseases have a variety of symptoms that include dementia, ataxia (loss of muscular coordination), insomnia, and behavioral problems. All involve some loss of neurons, which may or may not be indicated by a sponge-like appearance of the brain. A characteristic feature of prion disease is the appearance of **amyloid (starch-like) plaques**, which consist of fibrils of insoluble protein.[a] Exhaustive attempts failed to identify a virus particle or an associated DNA or RNA. On this basis, Stanley Prusiner suggested that the diseases are transmitted by pure proteins.[a] All of the diseases seem to involve the same protein, which is known as the **prion protein** (**PrP**). It is encoded by a single-copy gene on human chromosome 20.[g] The amino acid sequence of the C-terminal region of PrP is highly conserved among all animals. However, there are more than 20 known human genetic variants, and a second prion protein has been found in mice.[h] The function of the normal cellular prion protein (**PrP^C**) is unknown, but it appears to be a copper ion carrier, which may be essential to proper synaptic function.[i,j] "Knockout mice" lacking PrP are resistant to prion disease[j,k] but may not be completely healthy.

If it were not for the diseases, PrP^C might be viewed as just another cell surface glycoprotein. Determination of its three-dimensional structure has been difficult, but use of NMR spectroscopy and modeling has given a nearly complete picture, which is shown in the accompanying drawing.[d,l–n] The 250-residue (~220 residues after removal of N- and C-terminal signal sequences) has a long N-terminal tail, a glycosylated globular domain, and a C terminus that is anchored in the outer membrane of neurons by a glycosylphosphatidylinositol (GPI) anchor similar to that shown in Fig. 8-13. The globular domain contains three α helices, a small β sheet, and two glycosylation sites. These last carry



typical N-linked, branched sialic acid-containing oligosaccharides with a total of 52 or more sugar residues. The N-terminal 120 amino acid residue "tail" appears to be largely unstructured. However, it contains five octapeptide repeats with the consensus sequence PHGGGWGQ, each able to bind one Cu^{2+} or Mn^{2+} ion.[d,i,o,p]

How can this ordinarily harmless protein become a killer? The prion is a 20- to 30-kDa hydrophobic particle, which is thought to arise from PrP^C by a conformational alteration in which the α helices are largely changed into a β structure. The new conformer is often designated PrP^{Sc} or PrP-res. The latter abbreviation arises from the fact that native PrP^C can be completely hydrolyzed to small fragments by proteinase K, but PrP^{Sc} contains a 142-residue extremely resistant core (residues 90–231), which is not hydrolyzed and is over 80% β sheet.[q] Evidently the PrP^{Sc} form is able to associate to form a "seed" that, when conditions are favorable, can induce the conformational change in other molecules spreading the PrP^{Sc} form throughout the brain and even into tissues of the immune system.[r] With prion diseases and other amyloid diseases the body may be able to fight off the process by normal proteolytic turnover of the prion protein.

About 85% of all cases of prion disease are **sporadic CJD**. These are thought to arise by spontaneous conversion of PrP^C to PrP^{Sc}. Inherited (familial) forms of CJD, GSS, and FFI are also known. A series of point mutations as well as expansion of the octapeptide repeats[s] account for the various diseases, which have an autosomal dominant inheritance. At least 23 pathogenic mutations have been

BOX 29-E (continued)

reported.[t,u] The point mutations occur at several locations, some of them adjacent to the glycosylation sites.[m] These mutant proteins may be more readily converted to the less soluble PrP^{Sc} type structure, initiating the disease process. However, a mutant with a stop codon (TAG) in the place of the tyrosine 145 codon loses its C-terminal anchor and is degraded rapidly in the proteasomal pathway.[t]

The infectious forms of prion diseases are more puzzling. They account for less than one percent of all cases. Attention was first focused on kuru, a disease of the Fore people of New Guinea. In earlier times they practiced a ritualistic cannibalism of brain tissue that apparently propagated the disease, which is now nearly extinct. Of present concern are over 100 cases of a "new variant" form of CJD, some involving teenage persons and young adults, which have been reported in Europe.[u,v] This disease may have originated in sheep, then jumped to cattle, where it was spread by the ingestion of prion-contaminated meat and bone meal.[a] In addition, more than 120 cases of CJD have arisen from injection of prion-contaminated human growth hormone. Other cases have been traced to contaminated surgical instruments, to tissue grafts, and to use of contaminated human pituitary gonadotrophin.[a]

A hard-to-understand aspect of the "protein-only" theory of prion diseases is the existence of various "strains" of prion proteins. These do not involve differences in amino acid sequence but differences in the conformations of the PrP^{Sc} forms and in the glycosylation patterns.[d,m,w] How can there be several different conformations of the same protein, all of which seed the conversion of normal PrP into differing insoluble forms? In spite of this puzzle, support for the explanation of strain differences comes from a yeast prion system, which involves transcription termination factor eRF3.[x–z] In this system, which involves a prion whose insoluble form can be redissolved by guanidine hydrochloride,[aa] differing strains have also been described.[y,bb,cc] Nevertheless, the presence of the various strains of animal prions, as well as observed vaccination of inbred mice against specific strains,[dd] may be more readily understood if the disease is transmitted by an unidentified virus rather than by a pure protein.[r,u,ee,ff] In fact, the diseases have not been successfully transmitted by truly virus-free proteins synthesized from recombinant DNA.[ee]

What are the prospects for a cure for prion diseases? Several compounds show some effect in slowing accumulation of amyloid plaques,[d,v,gg,hh] but suitable drugs have not been developed. Prevention is the best cure, but more needs to be known about the basic biology of the disease transmission before effective strategies for prevention can be developed.[u]

What is the nature of the insoluble forms of the prion protein? They are hard to study because of the extreme insolubility, but the conversion of α helix to β sheet seems to be fundamental to the process and has been confirmed for the yeast prion by X-ray diffraction.[ii] It has been known since the 1950s that many soluble α-helix-rich proteins can be transformed easily into a fibrillar form in which the polypeptide chains are thought to form a β sheet. The chains are probably folded into hairpin loops that form an antiparallel β sheet (see Fig. 2-11).[jj–ll] For example, by heating at pH 2 insulin can be converted to fibrils, whose polarized infrared spectrum (Fig. 23-3A) indicates a **cross-β structure** with strands lying perpendicular to the fibril axis.[jj,mm] Many other proteins are also able to undergo similar transformation. Most biophysical evidence is consistent with the cross-β structure for the fibrils, which typically have diameters of 7–12 nm.[ii,ll,nn] These may be formed by association of thinner 2 to 5 nm fibrils.[oo] However, β-helical structures have been proposed for some amyloid fibrils[pp] and polyproline II helices for others.[qq]

A wide range of human diseases involving amyloid deposits are known. These include not only the prion diseases and the neurodegenerative diseases, Alzheimer, Parkinson, and the polyglutamine repeat diseases (Table 26-4)[,rr,ss] but also **systemic amyloidoses**.[tt] Among the latter are deposits of transthyretin,[uu] the 37-residue **amylin** that develops in the β cells of the pancreas in type II diabetes,[vv] mutant forms of lysozyme,[ww] and of β2 microglobulin,[xx] and gelsolin.[yy] A serum protein amyloid P, a calcium-binding protein, is usually also a component of amyloid deposits.[zz]

---

[a] Prusiner, S. B. (2001) *N. Engl. J. Med.* **344**, 1516–1526

[b] Manuelidis, L., Fritch, W., and Xi, Y.-G. (1997) *Science* **277**, 94–98

[c] Tagliavini, F., Lievans, P. M.-J., Tranchant, C., Warter, J.-M., Mohr, M., Giaccone, G., Perini, F., Rossi, G., Salmona, M., Piccardo, P., Ghetti, B., Beavis, R. C., Bugiani, O., Frangione, B., and Prelli, F. (2001) *J. Biol. Chem.* **276**, 6009–6015

[d] Caughey, B. (2001) *Trends Biochem. Sci.* **26**, 235–242

[e] Prusiner, S. B., ed. (1999) *Prion Biology and Diseases*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York

[f] Prusiner, S. B. (1996) *Trends Biochem. Sci.* **21**, 482–487

[g] Prusiner, S. B. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13363–13383

[h] Moore, R. C., and 20 other authors. (1999) *J. Mol. Biol.* **292**, 797–817

BOX 29-E PRIONS AND AMYLOID DISEASES (continued)

[i] Kramer, M. L., Kratzin, H. D., Schmidt, B., Römer, A., Windl, O., Liemann, S., Hornemann, S., and Kretzschmar, H. (2001) *J. Biol. Chem.* **276**, 16711–16719

[j] Yokoyama, T., Kimura, K. M., Ushiki, Y., Yamada, S., Morooka, A., Nakashiba, T., Sassa, T., and Itohara, S. (2001) *J. Biol. Chem.* **276**, 11265–11271

[k] Weissmann, C., and Aguzzi, A. (1999) *Science* **286**, 914–915

[l] Zahn, R., Liu, A., Lührs, T., Riek, R., von Schroetter, C., Garcia, F. L., Billeter, M., Calzolai, L., Wider, G., and Wüthrich, K. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 145–150

[m] Rudd, P. M., Wormald, M. R., Wing, D. R., Prusiner, S. B., and Dwek, R. A. (2001) *Biochemistry* **40**, 3759–3766

[n] Viles, J. H., Donne, D., Kroon, G., Prusiner, S. B., Cohen, F. E., Dyson, H. J., and Wright, P. E. (2001) *Biochemistry* **40**, 2743–2753

[o] Prince, R. C., and Gunson, D. E. (1998) *Trends Biochem. Sci.* **23**, 197–198

[p] Brown, D. R., Hafiz, F., Glasssmith, L. L., Wong, B.-S., Jones, I. M., Clive, C., and Haswell, S. J. (2000) *EMBO J.* **19**, 1180–1186

[q] Baskakov, I. V., Aagaard, C., Mehlhorn, I., Wille, H., Groth, D., Baldwin, M. A., Prusiner, S. B., and Cohen, F. E. (2000) *Biochemistry* **39**, 2792–2804

[r] Balter, M. (1999) *Science* **286**, 660–662

[s] Narwa, R., and Harris, D. A. (1999) *Biochemistry* **38**, 8770–8777

[t] Zanusso, G., Petersen, R. B., Jin, T., Jing, Y., Kanoush, R., Ferrari, S., Gambetti, P., and Singh, N. (1999) *J. Biol. Chem.* **274**, 23396–23404

[u] Almond, J., and Pattison, J. (1997) *Nature (London)* **389**, 437–438

[v] Thompson, C. (2001) *Nature (London)* **409**, 660–661

[w] Hill, A. F., Desbruslais, M., Joiner, S., Sidle, K. C. L., Gowland, I., Collinge, J., Doey, L. J., and Lantos, P. (1997) *Nature (London)* **389**, 448–450

[x] Patino, M. M., Liu, J.-J., Glover, J. R., and Lindquist, S. (1996) *Science* **273**, 622–626

[y] Liebman, S. W., and Derkatch, I. L. (1999) *J. Biol. Chem.* **274**, 1181–1184

[z] Zhou, P., Derkatch, I. L., Uptain, S. M., Patino, M. M., Lindquist, S., and Liebman, S. W. (1999) *EMBO J.* **18**, 1182–1191

[aa] Wickner, R. B., Edskes, H. K., Maddelein, M.-L., Taylor, K. L., and Moriyama, H. (1999) *J. Biol. Chem.* **274**, 555–558

[bb] Sparrer, H. E., Santoso, A., Szoka, F. C., Jr., and Weissman, J. S. (2000) *Science* **289**, 595–599

[cc] Chien, P., and Weissman, J. S. (2001) *Nature (London)* **410**, 223–227

[dd] Manuelidis, L. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2520–2525

[ee] Manuelidis, L., Sklaviadis, T., Akowitz, A., and Fritch, W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5124–5128

[ff] Manuelidis, L. (1997) *Annales De L'Institut Pasteur* **8**, 311–326

[gg] Tagliavini, F., and 20 other authors. (1997) *Science* **276**, 1119–1122

[hh] Supattapone, S., Nguyen, H.-O. B., Cohen, F. E., Prusiner, S. B., and Scott, M. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14529–14534

[ii] Balbirnie, M., Grothe, R., and Eisenberg, D. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2375–2380

[jj] Burke, M. J., and Rougvie, M. A. (1972) *Biochemistry* **11**, 2435–2439

[kk] Bradbury, E. M., Brown, L., Downie, A. R., Elliott, A., Fraser, R. D. B., Hanby, W. E., and Macdonald, T. R. R. (1960) *J. Mol. Biol.* **2**, 276–286

[ll] Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. F. (1997) *J. Mol. Biol.* **273**, 729–739

[mm] Bouchard, M., Zurdo, J., Nettleton, E. J., Dobson, C. M., and Robinson, C. V. (2000) *Protein Sci.* **9**, 1960–1967

[nn] Hörnberg, A., Eneqvist, T., Olofsson, A., Lundgren, E., and Sauer-Eriksson, A. E. (2000) *J. Mol. Biol.* **302**, 649–669

[oo] Chiti, F., Taddei, N., Bucciantini, M., White, P., Ramponi, G., and Dobson, C. M. (2000) *EMBO J.* **19**, 1441–1449

[pp] Lazo, N. D., and Downing, D. T. (1998) *Biochemistry* **37**, 1731–1736

[qq] Blanch, E. W., Morozova-Roche, L. A., Cochran, D. A. E., Doig, A. J., Hecht, L., and Barron, L. D. (2000) *J. Mol. Biol.* **301**, 553–563

[rr] Koo, E. H., Lansbury, P. T., Jr., and Kelly, J. W. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9989–9990

[ss] Kaytor, M. D., and Warren, S. T. (1999) *J. Biol. Chem.* **274**, 37507–37510

[tt] Hardy, J., and Gwinn-Hardy, K. (1998) *Science* **282**, 1075–1079

[uu] Sebastiao, M. P., Lamzin, V., Saraiva, M. J., and Damas, A. M. (2001) *J. Mol. Biol.* **306**, 733–744

[vv] Moriarty, D. F., and Raleigh, D. P. (1999) *Biochemistry* **38**, 1811–1818

[ww] Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C. F., and Pepys, M. B. (1997) *Nature (London)* **385**, 787–793

[xx] Chiti, F., Mangione, P., Andreola, A., Giorgetti, S., Stefani, M., Dobson, C. M., Bellotti, V., and Taddei, N. (2001) *J. Mol. Biol.* **307**, 379–391

[yy] Robinson, R. C., Choe, S., and Burtnick, L. D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2117–2118

[zz] Lambowitz, A. M., and Perlman, P. S. (1990) *Trends Biochem. Sci.* **15**, 440–444

proteins tubulin.[521–523] Accumulating α and β tubulin subunits act in a feedback loop to induce the degradation of their mRNA. They do this by binding to the N-terminal sequence MREI of β tubulin as it is emerging from ribosomes. This binding allosterically activates an endonuclease that cuts the polysome-bound mRNA. A similar mechanism for tubulin mRNA may involve the MREC N-terminal sequence of that protein.

An unexpected finding was that **phosphatidylserine synthase** of *E. coli* is tightly bound to ribosomes.[524] This enzyme, which incorporates serine into phospholipids according to step *h* of Fig. 21-4, is responsible for synthesis of the principal membrane lipid of *E. coli*. Its localization on ribosomes may be linked to the joint regulation of the synthesis of proteins and lipids.

D. Processing, Secretion, and Turnover of Proteins

The concepts of processing and turnover of proteins have been introduced in Chapter 10, and many details have been presented in other chapters. However, as we complete our discussion of protein synthesis, it is appropriate to discuss processing further. As polypeptide chains leave the ribosomes via the exit channels, they may follow several different paths. They may enter the cytosol and fold quickly into a compact form. This may require only a few seconds, whereas the translation process in the ribosome may take many seconds. The folding will therefore be **cotranslational**.[525] Depending upon the N-terminal signal peptide the protein may later unfold and pass through a membrane pore or **translocon** into the endoplasmic reticulum (ER), a mitochondrion, chloroplast, or peroxisome. Wherever it is, it will be crowded together with thousands of other proteins. It will interact with many of these, and evolution will have enabled some of these to become chaperones (discussed in Chapter 10).[526]

A single rapidly growing cell of *E. coli* may contain ~2.3 x 10^6 soluble polypeptide chains of ~2600 different types with an average length of ~317 residues and average mass of ~35 kDa.[527] These are chaperoned in a variety of ways,[528,529] not only in the cytoplasm but in the periplasm (see p. 364).[530,531] The three chaperones **trigger factor** (TF), DnaK, and GroEL participate in folding newly formed proteins.[525a,b] TF is a prolyl-*cis–trans* isomerase (Box 9-F), which associates with the large ribosomal subunit with a 1:1 stoichiometry.[525b–d] DnaK and related chaperones hold and protect newly formed polypeptides in extended conformations, while the GroES-GroEL chaperonins assist folding within their internal cavities (Box 7-A).[525b,e–g] **Prefoldin**, a 90-kDa complex,[525b] has a special function in chaperoning microtubule subunits. A Type II chaperonin also assists the folding of actin and tubulin.[525h] Chaperones assist not only in the folding of proteins but also in translocation into the ER (e.g., by the Hsp 70 homolog BiP), and into mitochondria and other organelles.[532,533] **Co-chaperones** are additional proteins that act as selective agents to direct proteins to a particular chaperone. For example, the DnaJ protein is a scanning factor for the Hsp 70 chaperone DnaK. As is described on p. 518, it catalyzes ATP-dependent association of a substrate with the binding cavity of DnaK.[533] A chaperone whose function has long remained elusive is an abundant secreted glycoprotein known as **clusterin.**[534] It seems to have a protective function in protection against stress.

How are the possible choices for newly formed proteins made? Much seems to depend upon the amino acid sequences at the ends of the polypeptide chains. As they emerge from a ribosome, some N-terminal signal sequences bind to recognition proteins. One such protein labels the ends of proteins destined for secretion into the vesicles of the ER. This protein ensures that the protein end binds to the **signal recognition particle** (SRP), enters a translocon pore, and undergoes cotranslational passage into the periplasmic space in bacteria or the ER in eukaryotes. Cotranslational modification reactions also occur both in the cytosol and in the ER vesicles. These too influence the choice of destinations as do additional signal sequences in the polypeptide chains. Proteins may be directed to the various organelles, to residence in membranes,[535] or to secretion into the external medium. It was somewhat surprising to discover that under some circumstances most newly synthesized proteins are degraded in proteasomes.[536] Cotranslational degradation of proteins with imperfect ends may account for some of this.[537–539] In addition, imperfect proteins that are retained in the ER may be sent back into the cytoplasm for degradation in proteasomes.[540]

1. Cotranslational and Posttranslational Processing

The modifications that lead to the presence of fully functional proteins in their proper locations begin while peptide chains are still emerging from the ribosomes.[541] In bacteria and in eukaryotic organelles the first of these modifications is hydrolytic removal of the *N*-formyl group by an Fe^{2+}-dependent **deformylase** leaving the N-terminal methionine.[542–543] Deformylases are present in eukaryotes as well as in bacteria, making the deformylase a less attractive target for antibiotic design than has sometimes been proposed.[544,544a] When the chain is only 20–30 residues long, the terminal methionine that remains after deformylation may be removed by a ribosome-associated methionine aminopeptidase.[545] The methionine is usually removed if followed by P, G, A, S, or T and is retained if followed by K, R, L, I, F, or N. With other amino acids removal is variable.[546] A ribosome-bound *N*-acetyltransferase may acetylate the N terminus either before or after removal of Met.[547] Approximate rules for eukaryotic cells are[541]:

N-terminal D, E, N	Acetylation without removal of Met
N-terminal P, V, C	Removal of Met; no acetylation
N-terminal G, A, S, T	Removal of Met followed by acetylation
Other N termini	No modification

An example is provided by actin, which contains acetyl-Met-Asp, acetyl-Met-Gln or acetyl-Met-Cys-Asp at the N terminus immediately after synthesis. Then, within ~15 min the acetyl-Met is cleaved off, and the next terminal residue is acetylated.[548] *N*-Acylation of nascent peptides by fatty acyl groups can also occur cotranslationally. For example, 14-carbon myristoyl groups are added in amide linkage to the N-terminal glycines of many cellular and virally encoded proteins.[535,549,550] This may take place on the ribosomes,[551] but it is often not clear whether the modification is cotranslational or posttranslational. The same may be said of many other "posttranslational" alterations, many of which may begin on a nascent polypeptide chain. Fatty acyl groups (mainly palmitoyl) may form thioester linkages with cysteine side chains.[552] This often occurs near the C terminus (see p. 559). Other C-terminal modifications include prenylation (p. 559)[553] and attachment to diacylglycerols via thioester linkages to cysteine (pp. 402, 428)[553a] or to glycosylphosphatidylinositol glycan anchors (Fig. 8-13; Eq. 20-23).

The addition of an N-terminal myristoyl group to a protein causes a relatively permanent alteration as does methylation of histidine, lysine, or arginine side chains.[554,555] So do hydroxylation, vitamin K-dependent carboxylation (Eq. 15-55), and many other alterations. In contrast, glycosylation, phosphorylation, and sulfation produce reversible alterations. Sometimes, as in the conversion of proenzymes to active enzymes, a modification step is used to generate a catalytic activity. In other instances, as in the processing of glycoproteins in the Golgi, the major function of the modification reaction seems to be one of directing a protein to the correct intracellular location.

2. Forming Functional Proteins

Proenzymes and other precursor proteins are often almost totally inactive until they are activated by some alteration that occurs when they reach their destination in a cell or in the body. Cleavage of the polypeptide chains of proenzymes, covalent attachment of coenzymes,[556,557] oxidation (Eq. 16-57) or halogenation (Eq. 25-6) of tyrosine or tryptophan[558] side chains, and oxidation of cysteine in the sequence CTPSR to formylglycine (in sulfatase formation; Eq. 12-44)[559] are only a few of many modifications needed to form functional proteins. Sometimes, as in activation of chymotrypsinogen, a single simple modification creates the active protein. In other cases modification may be quite complex. For example, although many polypeptide antibiotics are formed by nonribosomal synthesis (Box 29-C), some are created on ribosomes and may require extensive subsequent alteration. An example is **microcin**, a 69-residue peptide antibiotic formed by some strains of *E. coli*. Eight Gly-Cys and Gly-Ser pairs in a pre-microcin chain are cyclized to thiazole and oxazole rings. Then the 69-residue antibiotic is cut out from the precursor and secreted into the medium.[560,561] The 22-residue antibiotic **epidermin** is one of a family of **lantibiotics** that contain lanthionine as a characteristic component. Biosynthesis involves dehydration of serine and threonine residues, sulfide (thioether) bridge formation, oxidative decarboxylation, and removal of a leader peptide.[562]

3. Translocation through Membranes

The processes by which proteins are selected for secretion into the periplasmic space of bacteria or into the vesicles of the endoplasmic reticulum of eukaryotic cells are similar and have been discussed in Chapter 10 (pp. 519–521). However, some details are still being worked out. The first step in translocation is binding of the N terminus of a protein that is emerging from a ribosome to the **signal recognition particle**.[563,564] The core of this particle has a universally conserved structure consisting of two proteins and an RNA molecule.[565–571b]

Bacteria	Eukaryotes
4.5S RNA (~114 nt)	7S RNA (~295 nt)
Protein Ffh	Protein SRP54
Protein Ftsy (SRP receptor)[572]	Protein SRα

All of these proteins are GTPases, and in eukaryotes SRα is associated with a third protein SRβ, which is also a GTPase.[573] Either protein Ffh or SRP54 recognizes the N terminus of the protein that is to be translocated, chaperoning it to the receptor Ftsy[574] or SRα, where it may be anchored to the translocating pore (**translocon**). Hydrolysis of GTP by both proteins accompanies the recognition process. The eukaryotic SRP is more complex, containing six proteins and a larger RNA than in bacteria.[568,575–576] One domain of the 7S RNA is homologous to the bacterial 4.5S RNA, while an additional domain is closely related in its sequence to that of the highly repetitive *Alu* sequences in DNA (Fig. 27-9). However, the significance of this similarity is unclear.

In eukaryotic cells binding of SRP54 induces a transient retardation of translocation, an **elongation arrest**, while the SRP complex binds to its receptor SRα. This 70-kDa peripheral membrane protein is tightly associated with the 30-kDa integral membrane protein SRβ. Binding to this receptor leads the nascent polypeptide chain from the ribosome directly into the Sec61 translocon,[33f,576a] which consists of α, β, and γ subunits and a central aqueous pore. The ribosome apparently also becomes bonded firmly to the translocon until synthesis of the polypeptide chain has been completed.[573,577,578] The translocon complex also contains additional components[563,577,579] including the **leader peptidase** (signal peptidase, p. 620)[563,580–582a] and the **oligosaccharyltransferase** of ER membranes (Eq. 20-21).[563,583] The latter transfers an oligosaccharide from a lipid carrier onto certain asparagine side chains of polypeptides entering the ER. This and other glycosylation reactions help to keep the polypeptide moving to its correct destination, whether it be in some membrane surface, a lysosome or other organelle, or a secretion vesicle (Chapter 20). Furthermore, some proteins are translocated by a mechanism that doesn't depend upon SRP but utilizes a different complex, which consists in yeast of proteins Sec62, Sec63, Sec71, Sec72, and Kar2 (the chaperone BiP present in the ER lumen).[577]

Translocation of most bacterial proteins occurs posttranslationally rather than cotranslationally.[584] After recognition by SRP the polypeptide chains are transferred to chaperone complexes.[585] Some proteins are escorted to the folding compartment of the GroEL-GroES chaperonin (Box 7-A). Those that are to be secreted are often chaperoned by the protein **SecB**.[586] Genetic analysis shows that for *E. coli* the secretion of many proteins requires the products of genes *secA*, *secB, secD, secE, secF, secG,* and *secY*. Gene *secA* encodes a 92-kDa cytoplasmic ATPase protein (**SecA**), SecB is a 64-kDa homotetramic chaperone that prevents folding of preproteins prior to export, and SecY is a 42-kDa integral membrane protein.[587,588]

A complex of the three transmembrane proteins **SecYEG** forms the translocation channel in *E. coli* membranes. **SecY** and **SecE** are essential components in most bacteria and are homologous to components of the eukaryotic **Sec61p** translocon complex.[589] From their sequences SecY, SecE, and SecG are predicted to have ten, three, and two transmembrane sequences, respectively. Additional accessory proteins in the complex are designated SecD, SecF, and yajC. Their functions are uncertain, and they are not essential for transport. The driving force for translocation is provided by the peripheral ATPase Sec A.[588,590] For many proteins the signal sequence, which is usually positively charged, stays on the negatively charged cytoplasmic surface of the membrane, while SecA in an ATP-dependent process pushes a loop of protein through the membrane. Hydrolysis of a second ATP molecule seems to be required to release SecA, allowing it to reload with ATP and to assist the next 20–30 residue polypeptide sequence to be translocated.[591,592] Surprisingly, SecA is also an ATP-dependent RNA helicase.[590] The significance of its apparent ability to translocate along either an RNA or a polypeptide chain is unclear. The protonmotive force provided by the membrane potential is another important factor in the translocation of many proteins.[591,593–596] Yet another factor is the lipid composition of the membrane. Non-bilayer lipids seem to be required for efficient transport.[597]

As is mentioned in Chapter 10, bacteria have additional mechanisms of polypeptide transport. A recently recognized Sec-independent pathway is used by *E. coli* and many other bacteria to secrete proteins that contain the twin-arginine motif RRX$\phi\phi$, where ϕ is a hydrophobic amino acid, in their N-terminal signal sequences. Proteins encoded by genes *tatABC* are required by this **Tat pathway**.[594,598,599] Related pathways have been identified both in mitochondria and in chloroplasts. Small peptides may pass out of the periplasmic space through the porins in the outer membrane of gram-negative bacteria. However, larger proteins require conduit molecules such as the **TolC** channel-tunnel, which directly connects an inner membrane translocon with a channel in the outer membrane of *E. coli* cells.[600,601]

Eukaryotic cells also have additional transport mechanisms. One of these is an ABC transporter (p. 417) known as the **transporter associated with antigen processing** (**TAP**). It carries small polypeptides generated by proteasomes from the cytosol into the ER for export and binding to **MHC Class I** molecules and subsequent presentation to the immune system (Fig. 31-15).[540,602]

4. Translocation into Organelles

Most of the proteins of mitochondria are encoded in nuclear DNA and are synthesized on cytoplasmic ribosomes. Mitochondria do not utilize proteins homologous to those of the bacterial Sec system but have their own set of transport proteins.[603–605] These proteins, which include an outer membrane complex (Tom) and an inner membrane complex (Tim), are discussed in Chapter 18 (see Fig. 18-4). Perhaps these specialized mitochondrial proteins are needed because transport into the mitochondrial matrix is in an opposite direction to the transport out through bacterial membranes.

The transport of proteins into chloroplasts also occurs by more than one mechanism. An SRP-dependent pathway may be needed only for insertion of proteins into membranes.[594] Other proteins, among which are the 23-kDa and 16-kDa photosystem II proteins (Chapter 23), enter by a pathway related to the Tat pathway of bacteria. In thylakoids this pathway is directly dependent upon the large pH difference (Δ pH) across the thylakoid membrane. In contrast to the bacterial Sec pathway, the Δ pH pathway seems to be able to transport completely folded proteins.

Proteins destined for peroxisomes have their own targeting signals. One of these (**PTS1**) is the sequence SKL at the C terminus. A second signal (**PTS2**) is an N-terminal nonapeptide (R / K)(L / V / I) X_5(H / Q) (L / A).[606,607]

5. Membrane Proteins

Some proteins enter membranes immediately after synthesis. The translocon channel is not required. However, in *E. coli* an additional protein **YidC** is needed.[603] Homologs of this protein are found in mitochondria (**Oxa1** protein) and in thylakoid membranes of chloroplasts (**Alb3** protein).[608] These proteins may function in cotranslational insertion. If a protein carries a positively charged N-terminal region, it will tend to stick to the negatively charged cytoplasmic surface of a cell membrane. This "positive inside" rule (p. 401)[609] is strong for bacterial proteins but somewhat weaker for eukaryotic cells. A second topological rule is that hydrophobic segments of proteins will be attracted to membrane surfaces and can enter the membrane (perhaps via translocon pores) as loops (Fig. 29-18).[596] Passage of the loop out through the membrane will be facilitated if negatively charged groups are present in the loop and are acted upon by the membrane potential (pp. 401,402).[595,610,611] If the entire polypeptide chain follows the loop out through the membrane, the protein will be anchored to the inside of the cell membrane with its C terminus outside. On the other hand, if the C terminus also has a positively charged cluster nearby, a membrane associated **leader peptidase** (or signal peptidase) may cut the loop past the signal sequence on the outside of the membrane leaving the bulk of the protein with its N terminus outside (Fig. 29-18B). How are **polytopic** integral membrane proteins with multiple cytoplasmic and external loops formed? Hydrophobic signal sequences are not always at the N terminus of a polypeptide chain. Suitable **internal signal sequences** may be found in the sequences that form the transmembrane helices, e.g., those present in the many 7-helix receptors found in a membrane. This suggests the possibility that successive loops may be translocated. If the N terminus is allowed to pass through a translocon in one of the steps, the topology of Fig. 11-6 or of Fig. 23-41 will result.[611]

Genetic methods have also been applied to study the insertion of coat subunits of phage M13 into the plasma membrane of *E. coli*.[612–614] The subunits are stored in the plasma membrane waiting to form a cylindrical shell about a viral DNA molecule as it is extruded from the bacterium.[612] The rod-like subunits (Fig. 7-7) have their N termini in the periplasmic space and their C termini in the cytoplasm. Each end carries a cluster of electrically charged residues, mostly negative at the N terminus and positive at the C terminus. Insertion into the membrane occurs only if the membrane has its normal membrane potential with a positive external surface charge and a negative internal charge, complementary to the charges on the coat subunit. Insertion does not occur unless the leader peptide with its positively charged N-terminal cluster and the C-terminal positive cluster are both present.[610,615] This suggests insertion by the loop mechanism of Fig. 29-18A. Genetic studies of the *E. coli* leader peptidase revealed that this protein also has an internal signal sequence, which becomes inserted into the membrane and which is not cleaved.[612,616,617] The final orientation of the mature enzyme is indicated in Fig. 29-18C.

Targeting of proteins to specialized domains of a membrane are less well understood. These include caveolae and lipid rafts, domains that are high in cholesterol and sphingolipids and which function in endocytosis and in cell signaling. A recent proposal is that proteins with hydrophobic surfaces needed in these domains become coated with a lipid "shell" before entering the membrane.[617a]

6. Secretion of Proteins

Cells continuously secrete materials via small cytoplasmic vesicles, which in eukaryotes arise largely from the Golgi apparatus (pp. 425–427; Fig. 20-8). The vesicles of this **constitutive pathway** may have diameters of ~50 nm. They carry phospholipids, proteins, and other constituents for incorporation into the plasma membrane of the cell.[618,619] In addition, there are

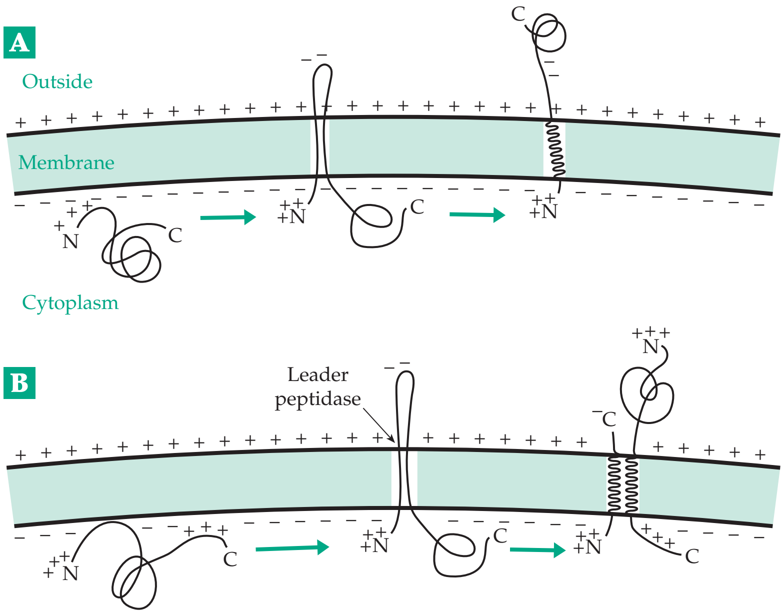
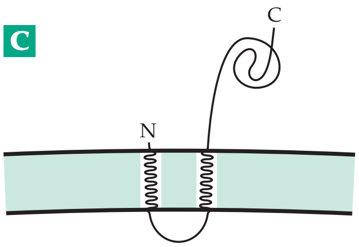


Figure 29-18 (A) Proposed mechanism for insertion of a loop of polypeptide chain through a translocon pore in a membrane with the positively charged N terminus anchored to the negatively charged inner membrane surface. (B) Cleavage of a polypeptide loop formed as in (A) by a leader peptidase to give a polypeptide chain anchored by a positively charged cluster near its C terminus. (C) Membrane topology of the *E. coli* leader peptidase. The active site is in the periplasmic domain. See Tschantz *et al.*[580]

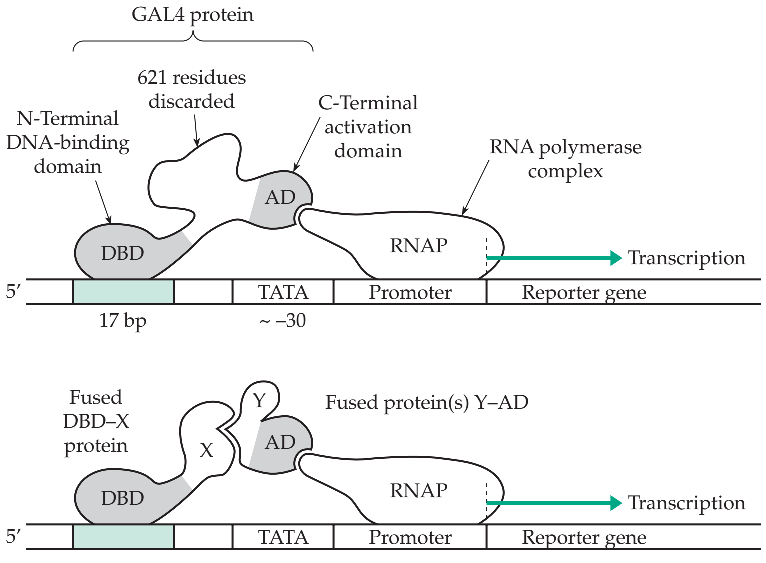


BOX 29-F THE YEAST TWO-HYBRID SYSTEM FOR IDENTIFYING PROTEIN–PROTEIN INTERACTIONS

Many techniques including ultracentrifugation, chemical crosslinking, and X-ray crystallography are used to identify interactions between proteins. However, for study of the entire proteome new approaches are needed. One new technique that has already been widely applied is the yeast two-hybrid system.[a–f] In its original form[a,b] a transcriptional activator is utilized together with a reporter gene, e.g., a fused *GAL*1– *lac2* gene that when transcribed yields β-galactosidase. This enzyme then cleaves a chromogenic substrate (see p. 1494) to give a blue color. The 881-residue transcription factor GAL4 (p. 1630) is often used in two-hybrid systems. It binds upstream of the TATA sequence in the promoter regions of genes that code for enzymes of galactose catabolism. The N-terminal **DNA-binding domain** of *GAL4* binds to a specific 17-bp palindromic sequence in the DNA, while the acidic C-terminal **activation domain** activates transcription by interacting with the RNA polymerase complex bound to the promoter (see figure). The GAL4 activator seems to be quite flexible and is able to activate transcription, even if the distance from its binding site on DNA to the transcription initiation site is varied considerably. The two-hybrid system was constructed by cloning separately the pieces of DNA that code for residues 1–147 of the DNA-binding domain of the GAL4 protein (DBD in the figure) and for residues 768–881 of the activation domain (AD). The intervening nucleotides coding for the remaining 620 residues of GAL4 are discarded. The two-hybrid system tests whether a protein X, sometimes called the **bait**, binds or otherwise interacts strongly with another protein (Y, often called the **prey**) or with a series of other proteins (Y_1, $Y_2 \cdots Y_n$). To carry out the test the gene for protein X is fused with that for the DNA-binding domain of GAL4. When expressed in a living yeast cell the hybrid protein DBD–X will be formed. The gene fusion must be in-frame to ensure a correct structure for the X portion. Likewise, genes for protein Y, or for a series Y_1, $Y_2 \cdots Y_n$, are fused in-frame to the gene fragment carrying the GAL4 activation domain. Y–AD hybrids will be formed. The test is made using a strain of yeast in which the GAL4 gene has been replaced with a hybrid *GAL*1– *lacZ* reporter gene. If both DBD–X and Y–AD are present, and if they interact strongly (bind tightly), transcription of the reporter gene will be activated, and a blue colony will grow from the yeast cell. It is useful to create the hybrids with protein X, or proteins X_1 $X_2 \cdots$ (the baits) in a haploid strain of one of the two yeast mating types *MAT*a or *MAT*α (pp. 20, 1574) and the proteins Y (the prey) in a strain of the other mating type. Mating of the two strains will produce diploid cells that express both the DBD–X and Y–AD hybrids.

Many variants of the two-hybrid system have been devised.[d] For example, a green fluorescent protein reporter can be used.[g] Because significant biological protein–protein interactions often require that three or more proteins interact,[d] hybrid systems involving more than two proteins have been developed. Two-hybrid systems for bacteria have also been devised.[h] A virtue of the two-hybrid methods is that they work with undenatured, if not totally natural, proteins. This is in contrast to widely used methods that involve separation of denatured proteins on gels or columns.

The most popular two-hybrid systems utilize microarrays.[d,f] In the simplest approach hybrid DBD–X is tested against a library of Y–ADs



Top: The yeast GAL4 protein interacts with the RNA polymerase complex to activate transcription of a suitable reporter gene. Bottom: Two hybrid proteins, one containing the DNA-binding domain of GAL4, fused to protein X and the other containing protein Y fused to the activation domain of GAL4, are present in a cell. If X and Y bind strongly to each other, activation domain AD will be held close to the RNA polymerase and will activate transcription.

BOX 29-F THE YEAST TWO-HYBRID SYSTEM FOR IDENTIFYING PROTEIN–PROTEIN INTERACTIONS (continued)

prepared by random cleavage of DNA of known sequence. For example, the entire genome of the gastric pathogenic bacterium *Helicobacter pylori* was cut into ~1000 nucleotide pieces. These were cloned into plasmids in *E. coli*, then into yeast. Over 10 million *E. coli* clones provided a final two million independent yeast colonies, which carried the Y–ADs (prey). The genomic DNA fragments were also used to prepare 285 DBD–X (baits) from 261 genes. In a series of two-hybrid screening tests more than 1200 different protein–protein interactions connecting 47% of the proteome were detected.[i]

The complete yeast (*S. cerevisiae*) has been probed using at least two large-scale two-hybrid investigations. Uetz *et al.*[j] generated a large set of ~6000 genetically engineered yeast colonies, each one expressing just one of the possible Y–AD hybrid proteins (prey) derived from the ~6000 gene products identified in the yeast genome. These strains were distributed into microtiter plates and were individually crossed with 192 strains of yeast, each of which expressed a single BDB–X hybrid. This simple automated array screening identified 281 interacting pairs. In a second approach, the cells producing the Y–AD prey hybrids were mixed to give a single library. This was then screened against nearly all of the possible DBD–X hybrids in a large-scale automated procedure. The two approaches together detected 957 probable interactions involving 1004 different proteins.[j] In an independent study, using similar approaches but different cloning vehicles, Ito *et al.*[k,l] identified 4549 two-hybrid interactions among 3278 proteins. Of these 841 interactions were judged to be most relevant (core). Surprisingly, only 135 were identical to those found by Uetz *et al.* The significance and possible reasons for this disparity have been discussed.[f,j,l]

Interpretation of results of these studies is still difficult. Results of two-hybrid methods become more useful if they can be coordinated with other approaches. For example, computational methods can predict interactions from genome sequences alone.[m,n,o] More than 45,000 interactions have been predicted among yeast proteins.[m] Reliable identification of such motifs as DNA-binding domains and Ca^{2+}- binding domains can complement two-hybrid analysis.[n] The yeast genome is predicted to contain 162 coiled-coil sequences and at least 213 unique interactions between them.[o] Examination of sequences of protein families in the Protein Data Bank (PDB) led to prediction of 8151 interactions of 664 types between protein families in yeast.[p]

Improved experimental procedures of other types can also complement two-hybrid methods. Among these are formaldehyde crosslinking with immunoprecipitation,[q] methods that couple mass spectrometry and crosslinking,[r] and detection of intermolecular nuclear Overhauser enhancements in protein–protein complexes.[s] Phage display methods (see Fig. 3-16) have been developed as another method of detecting protein–protein interactions[e] as has fluorescence resonance energy transfer (FRET; p. 1291).[d] Evanescent wave methods, e.g., surface plasmon resonance (Box 3-F) are increasingly being used to quantify protein–protein interactions. These may be combined with single-hybrid methods in inexpensive and rapid micro-devices.[d]

---

[a] Fields, S., and Song, O.-K. (1989) *Nature (London)* **340**, 245–246

[b] Chien, C.-t, Bartel, P. L., Sternglanz, R., and Fields, S. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9578–9582

[c] Finley, R. L., Jr., and Brent, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12980–12984

[d] Mendelsohn, A. R., and Brent, R. (1999) *Science* **284**, 1948–1950

[e] Allen, J. B., Walberg, M. W., Edwards, M. C., and Elledge, S. J. (1995) *Trends Biochem. Sci.* **20**, 511–516

[f] Oliver, S. (2000) *Nature (London)* **403**, 601–603

[g] Shioda, T., Andriole, S., Yahata, T., and Isselbacher, K. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5220–5224

[h] Joung, J. K., Ramm, E. I., and Pabo, C. O. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7382–7387

[i] Rain, J.-C., and 12 other authors. (2001) *Nature (London)* **409**, 211–215

[j] Uetz, P., and 19 other authors. (2000) *Nature (London)* **403**, 623–627

[k] Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S., and Sakaki, Y. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1143–1147

[l] Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4569–4574

[m] Marcotte, E. M., Pellegrini, M., Ng, H.-L., Rice, D. W., Yeates, T. O., and Eisenberg, D. (1999) *Science* **285**, 751–753

[n] Gallet, X., Charloteaux, B., Thomas, A., and Brasseur, R. (2000) *J. Mol. Biol.* **302**, 917–926

[o] Newman, J. R. S., Wolf, E., and Kim, P. S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13203–13208

[p] Park, J., Lappe, M., and Teichmann, S. A. (2001) *J. Mol. Biol.* **307**, 929–938

[q] Orlando, V. (2000) *Trends Biochem. Sci.* **25**, 99–104

[r] Bennett, K. L., Kussmann, M., Björk, P., Godzwon, M., Mikkelsen, M., Sorensen, P., and Roepstorff, P. (2000) *Protein Sci.* **9**, 1503–1518

[s] Clore, G. M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9021–9025

regulated pathways for storage and release by exocytosis of hormones, neuropeptides, and neurotransmitters. The last are secreted from both small (~ 50 nm diam.) synaptic vesicles and larger **dense core vesicles** (>100 nm diam.),[620,621] which are discussed in Chapter 30, Section B,8. In every case specialized proteins (discussed on pp. 427 and 521 and in Chapters 20 and 30) are involved.[622–624]

7. Protein Folding

Before a protein can function its polypeptide chain must fold into its own native tertiary structure.[624a] This folding is influenced by many surrounding proteins, by the state of glycosylation of side chains,[625] and by other posttranslational modifications, by the presence of cis amide linkages in unfolded or folded forms (pp. 82, 83; Box 9-F),[626] and by possibility for formation of disulfide bridges (pp. 521, 522).[626a–628] The prediction of the folding pattern of proteins from the amino acid sequence remains a major goal of protein chemistry. In principle, a protein fold can be predicted from the DNA sequence of the genes, with proper allowance for effects of posttranscriptional modification. This goal once seemed intractable, but two things have provided new hope. (1) The speed and power of computers is still increasing. This not only has allowed more rapid calculations but also has led to improvement in experimental methods. (2) Methods for studying folding, which include mass spectroscopy,[629] NMR spectroscopy,[630–633] and optical methods,[634,635] have become more rapid and more sensitive.[636] As a consequence, we have an abundance of new data.

Anfinsen, in 1963, proposed that the three-dimensional structure of a protein is in its lowest Gibbs energy state when present in its natural environment.[637,638] However, there is a problem with this suggestion (the “Levinthal paradox”; p. 82). Even if a polypeptide chain occupies only two of the lowest energy regions of the Ramachandran diagram (Fig. 2-9), a 100-residue protein would have ~10^{30} possible conformations. If a folding protein checked all of these conformations at a realistic rate of ~10^{11} s^{-1}, it would take ~10^{11} years to fold. Furthermore, in a test tube of protein, which would contain at most 10^{18} molecules, each of the molecules would probably have a different conformation.[639] In fact, most proteins fold reliably to the same final structure in less than a second, and some in a millisecond.[639–643] It is also true that proteins, under altered solvent conditions, can misfold into totally “incorrect” structures.[644] Most can assume an amyloid structure under some conditions (Box 29-E). One clear conclusion is that folding is not totally random but follows a **folding pathway,** which is dictated by the sequence. Nevertheless, experimental data indicate that there is an ensemble of related structures at each stage of folding.[631]

We cannot answer the question posed by Anfinsen’s hypothesis. Does the native state have a minimum value of the Gibbs energy? Nevertheless, it is observed that proteins usually behave as if folded, unfolded forms are in a true thermodynamic equilibrium, and that this equilibrium is attained rapidly. The difference ΔG between a folded and a denatured protein is only 21–63 kJ mol^{-1}, which shows that folded proteins are only marginally more stable than are unfolded polypeptide chains.[645] The value of ΔG of unfolding as a function of temperature T is given by Eq. 29-13, where $\Delta H(T)$ and ΔC_p are the changes in enthalpy and heat capacity upon unfolding.[645,646]

$$\Delta G\,(T) = \Delta H_m\,(1 - T/T_m) - \Delta C_p\,[(T_m - T) - T\ln(T/T_m)] \tag{29-13}$$

ΔH_m is the enthalpy change at T_m, the midpoint of the thermal unfolding curve (the “melting” temperature). The temperature of maximum stability T_s occurs when $\Delta S = 0$ (Eq. 29-14). T_s is usually between –10°C and

$$T_s = T_m \exp(-\Delta H_m / [T_m \Delta C_p]) \tag{29-14}$$

35°C. For staphylococcal nuclease (Fig. 12-29) T_s = 18°C and T_m = –19°C and +57°C, i.e., *the protein is denatured by either cooling below 18°C or heating above 57°C*, a behavior that is common for many proteins. Cold denaturation is observed whenever the unfolded state has a higher heat capacity than does the folded state.[647]

Can we predict the Gibbs energy of unfolding from the protein sequence? To do this it is necessary to utilize experimental data on known proteins to obtain a series of terms (Eq. 29-15) that can be summed to give the $\Delta G°_u$, the standard Gibbs energy of unfolding.

$$\Delta G°_u = \Delta G°_{charge} + \Delta G°_{hyd} + \Delta G°_{conf} + \Delta G°_{vW} + \Delta G°_{Hbond} \tag{29-15}$$

The terms refer to summations of all of the charge–charge, hydrophobic, configurational, van der Waals, and hydrogen-bonding interactions between both main-chain and side-chain atoms.[638,646,648] Such computations are formidable and are uncertain, especially for electrostatic (charge–charge and hydrogen-bonding) interactions.[649–652] Both the folded and denatured state must be considered,[653] as must the heat capacities[654] and configurational entropies.[655]

While we tend to think of proteins as having fixed structures, conformational changes are basic to life. Many proteins are very flexible and in part disordered.[655a] At the same time proteins can be misfolded leading to amyloid formation (Box 29-E) and to other diseases.[655b–e]

8. Completing the Cycle: Proteolytic Degradation of Proteins

Like all other body constituents proteins must undergo breakdown as well as synthesis. Regular turnover of all proteins is essential, and defects in the process may lead to amyloid deposits (Box 29-E). Turnover and degradation of proteins depends upon a variety of proteases, many of which are discussed in Chapter 12, Section C. Because of the changing needs of body cells, specific proteins turn over at widely varying rates. Some enzymes, hormones, and regulatory proteins have half-lives of only a few minutes while others may function for months or years. How can the regulation of the breakdown of thousands of different proteins be controlled? The answer seems to lie in the amino acid sequence. Just as the sequence determines the location that a protein occupies in the cell and its folding pattern, it also determines the turnover rate.

Much of the breakdown takes place in the cytosol in proteasomes (Box 7-A) and is controlled by the ubiquitin system (Box 10-C), which selects the proteins for degradation; control of the system is quite complex.[656–661] One aspect depends upon the N-terminal amino acid of the substrate protein. Defective proteins often have N-terminal destabilizing amino acids such as phenylalanine, leucine, aspartic acid, lysine, and arginine (p. 527). If an internal lysine is also present, the protein may be conjugated with ubiquitin and degraded rapidly. Many metabolic processes, such as the cell cycle (Fig. 11-15), are controlled by protein degradation.[538,662] In some cases an arginine residue is transferred onto the N terminus of a protein by an **arginyl-tRNA protein transferase**. This creates a better substrate for ubiquitination and rapid degradation.[280,663] In other cases proteolytic cleavage uncovers an arginine or other destabilizing residue and speeds hydrolysis.[538] While ubiquitination often initiates the degradation of proteins, it also helps to direct proteins to specific locations within a cell.[663a] Rapid degradation of a ubiquitinated protein may require hydrolytic **deubiquitination** by a metalloprotease, which is a subunit of the 26S proteasome lid (Box 7-A). This allows the ubiquitin to be recycled and also directs the deubiquitinated protein into the proteasome.[663b]

As mentioned on p. 1854, an important function of proteasomes is formation of short antigenic peptides for use by the immune system.[664] Inhibition of proteasome activity reduces or prevents antigen presentation (Chapter 31).[665,666] In this immune surveillance system mature proteins of host cells are cut up and checked for self-identity. The checking also includes the rapidly degraded imperfect proteins and foreign proteins from invading organisms or viruses.[667,667a]

Lysosomes, which contain more than 50 proteases, lipases, glycosidases, and other hydrolases, also play a major role in protein degradation.[668] Their importance is emphasized by the range of lysosomal deficiency diseases (Table 20-1).[668] Lysosomes also function in the process of **autophagy**, by which cells can sacrifice a whole section, organelles and all, by walling off a large vesicle or **autophagosome** and fusing it with a lysosome.[669] In such a way a tadpole can resorb its tail while becoming a frog. We have now come full circle: our proteins have been converted back to the amino acids and other small molecules derived from them. The amino acids can be reutilized or can be catabolized, depending upon the needs of the organism.

E. Proteomics

The vast amount of data on protein structures and improved methods of predicting structures[670–672] have led to development of new areas of science variously designated as genomics, proteomics, transcriptomics,[673,674] and bioinformatics.[675–677] These fields encompass all of the methods for sequence determination, observation of gene expression, protein synthesis by cells, and mathematical analysis of resulting data. Proteomics includes new approaches to polypeptide separation[678] and identification,[678,679] sequencing at the attomole level,[680–681] and comparison of sequences between species.[682–684] Protein separation by liquid chromatography,[685] capillary electrophoresis,[686] or two-dimensional gel electrophoresis[687,688] can be followed by mass spectrometry of intact proteins or of proteolytic fragments.[689,690] Microarrays on proteome chips can be used to observe production of thousands of proteins simultaneously.[691–694] Structural genomics centers have been established for rapid determination of protein structures using NMR[695,696] and X-ray methods.[697–699] If each center determines 200 or more structures per year there will soon be 16,000 new structures, enough to allow us to predict much about all the rest.[699] Then we can study all the important remaining details for millions and millions of proteins.

Current efforts to understand the structures, conformational movements, and functions of these molecules range from the classification of nearly 10,000 different protein folds[700] to investigation of the dynamics of single protein molecules.[701] Well known motifs such as β sheets and α helices are studied with the goal of more accurate predictions of structure and better understanding of interactions between proteins in solution and in membranes. For example, one natural topology is the β barrel, which may contain 8 to 22 strands (e.g., see Fig. 8-20). These cylindrical proteins are abundant in outer membranes of gram-negative bacteria.[702–704] The partial electrical changes at the edges of the β sheets (see Fig. 2-11,B) may interact to help stabilize the barrels. In contrast, soluble β barrel-containing proteins are designed to avoid edge-to-edge interactions, which could cause aggregation of the proteins.[704] A recently discovered membrane-protein motif is an α **barrel**, which is composed of 12 α helices stacked side-by-side with side-chain groups fitting together in a knobs-in-holes fashion.[705] An example is the TolC protein of *E. coli*. A trimer of 428-residue subunits forms a long cylinder, which is a 12-stranded β barrel at one end and an α barrel at the other.[705] More common are transmembrane α helices, many of which are present in 7-helix receptors (e.g., Fig. 11-6). Relatively accurate prediction methods are now available for these structures,[706,707] but there are still uncertainties about mechanisms of transmission of signals across the membranes.

Predictions of structures of more complex proteins from their amino acid sequences presents a major challenge.[708] Assignment of domains within the protein is a first step.[709–712] Regions of probable helix or β-strand structure can be recognized but it is difficult to predict the exact lengths of the helices and the structures of connecting loops and strands. These depend upon many factors including the possible formation of ion pairs[713] and of locks at the ends of strands created by van der Waals interactions.[714] There are also circular proteins.[715] Composite structures such as those of silks (Box 2-B) have surprising properties. Both silkworm,[716,717] and spider silks[718,719] undergo marked changes in properties upon spinning of the random coil forms of the proteins found in silk glands into the drawn fibers.

Whether we discuss silk, proteins embedded in membranes, or soluble complexes of cytosolic proteins, we must ask questions about interactions. A first step is to identify interactions[720–730] among proteins either *in vitro* or in living cells.[731] Proteomic methods, which include the yeast two-hybrid method (Box 29-F), are widely used for this purpose. It is possible to identify large sets of interacting proteins, to identify disease states, to observe effects of drugs, and to compare metabolism among species.

References

1. Nomura, M. (1990) *Trends Biochem. Sci.* **15**, 244–247
2. Lake, J. A. (1981) *Sci. Am.* **245**(Aug), 84–97
3. Schachman, H. D., Pardee, A. B., and Stanier, R. Y. (1952) *Arch. Biochem. Biophys.* **38**, 245–260
4. Wittmann, H. G. (1983) *Ann. Rev. Biochem.* **52**, 35–65
5. Prince, J. B., Gutell, R. R., and Garrett, R. A. (1983) *Trends Biochem. Sci.* **8**, 359–363
6. Yonath, A., Leonard, K. R., and Wittmann, H. G. (1987) *Science* **236**, 813–816
7. Lake, J. A. (1985) *Ann. Rev. Biochem.* **54**, 507–530
8. Frank, J., Verschoor, A., Wagenknecht, T., Radermacher, M., and Carazo, J.-M. (1988) *Trends Biochem. Sci.* **13**, 123–127
9. Baumeister, W., and Steven, A. C. (2000) *Trends Biochem. Sci.* **25**, 624–631
10. Morgan, D. G., Ménétret, J.-F., Radermacher, M., Neuhof, A., Akey, I. V., Rapoport, T. A., and Akey, C. W. (2000) *J. Mol. Biol.* **301**, 301–321
11. De Chadarevian, S. (1999) *Trends Biochem. Sci.* **24**, 203–206
12. Nagano, K., and Harel, M. (1987) *Prog. Biophys. and Mol. Biol.* **48**, 67–101
13. Lake, J. A. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 1903–1907
14. Lake, J. A., and Strycharz, W. A. (1981) *J. Mol. Biol.* **153**, 979–992
15. Kahan, L., Winkelmann, D. A., and Lake, J. A. (1981) *J. Mol. Biol.* **145**, 193–214
16. Ramakrishnan, V., Capel, M., Kjeldgaard, M., Engelman, D. M., and Moore, P. B. (1984) *J. Mol. Biol.* **174**, 265–284
17. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 905–920
18. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., and Noller, H. F. (1999) *Science* **285**, 2095–2104
19. Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 920–930
20. Agrawal, R. K., Penczek, P., Grassucci, R. A., Burkhardt, N., Nierhaus, K. H., and Frank, J. (1999) *J. Biol. Chem.* **274**, 8723–8729
21. Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., and Zamir, A. (1965) *Science* **147**, 1462–1465
22. Wittmann, H. G. (1982) *Ann. Rev. Biochem.* **51**, 155–183

22a. Brownlee, G. G., Sanger, F., and Burell, B. G. (1967) *Nature (London)* **215**, 735–736

22b. Brosius, J., Palmer, M. L., Kennedy, P. J., and Noller, H. F. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4801–4805

22c. Brosius, J., Dull, T. J., and Noller, H. F. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 201–204

23. Leijonmarck, M., Eriksson, S., and Liljas, A. (1980) *Nature (London)* **286**, 824–826
24. Ramakrishnan, V., and White, S. W. (1998) *Trends Biochem. Sci.* **23**, 208–212
25. Yonath, A. (1984) *Trends Biochem. Sci.* **9**, 227–230
26. Yonath, A., and Wittmann, H. G. (1989) *Trends Biochem. Sci.* **14**, 329–335
27. Yonath, A., Glotz, C., Gewitz, H. S., Bartels, K. S., von Böhlen, K., Makowski, I., and Wittman, H. G. (1988) *J. Mol. Biol.* **203**, 831–834
28. Moore, P. B. (2001) *Biochemistry* **40**, 3243–3250
29. Schluenzen, F., Tociij, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000) *Cell* **102**, 615–623
30. Culver, G. M., Cate, J. H., Yusupova, G. Z., Yusupov, M. M., and Noller, H. F. (1999) *Science* **285**, 2133–2135
31. Agalarov, S. C., Prasad, G. S., Funke, P. M., Stout, C. D., and Williamson, J. R. (2000) *Science* **288**, 107–112

31a. Recht, M. I., and Williamson, J. R. (2001) *J. Mol. Biol.* **313**, 35–48

32. Pennisi, E. (1999) *Science* **285**, 2048–2051
33. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) *Nature (London)* **407**, 327–339

33a. Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. D., and Noller, H. F. (2001) *Science* **292**, 883–896

33b. Ramakrishnan, V., and Moore, P. B. (2001) *Curr. Opin. in Struct.Biol.* **11**, 144–154

33c. Nissen, P., Ippolito, J. A., Ban, N., Moore, P. B., and Steitz, T. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4899–4903

33d. Ogle, J. M., Brodersen, D. E., Clemons, W. M. J., Tarry, M. J., Carter, A. P., and Ramakrishnan, V. (2001) *Science* **292**, 897–902

33e. Mathews, M. B., and Pe'ery, T. (2001) *Trends Biochem. Sci.* **26**, 585–587

33f. Moore, P. B., and Steitz, T. A. (2002) *Nature (London)* **418**, 229–235

33g. Pioletti, M., Schlünzen, F., Harms, J., Zarivach, R., Glühmann, M., Avila, H., Bashan, A., Bartels, H., Auerbach, T., Jacobi, C., Hartsch, T., Yonath, A., and Franceschi, F. (2001) *EMBO J.* **20**, 1829–1839

34. Malhotra, A., Penczek, P., Agrawal, R. K., Gabashvili, I. S., Grassucci, R. A., Jünemann, R., Burkhardt, N., Nierhaus, K. H., and Frank, J. (1998) *J. Mol. Biol.* **280**, 103–116
35. Liljas, A. (1999) *Science* **285**, 2077–2078
36. Dube, P., Bacher, G., Stark, H., Mueller, F., Zemlin, F., van Heel, M., and Brimacombe, R. (1998) *J. Mol. Biol.* **279**, 403–421
37. Gomez-Lorenzo, M. G., Spahn, C. M. T., Agrawal, R. K., Grassucci, R. A., Penczek, P., Chakraburtty, K., Ballesta, J. P. G., Lavandera, J. L., Garcia-Bustos, J. F., and Frank, J. (2000) *EMBO J.* **19**, 2710–2718

37a. Mueller, F., Sommer, I., Baranov, P., Matadeen, R., Stoldt, M., Wöhnert, J., Görlach, M., van Heel, M., and Brimacombe, R. (2000) *J. Mol. Biol.* **298**, 35–59

38. Brimacombe, R. (1984) *Trends Biochem. Sci.* **9**, 273–277
39. Noller, H. F. (1984) *Ann. Rev. Biochem.* **53**, 119–162

39a. Gutell, R. R., Noller, H. F., and Woese, C. R. (1986) *EMBO J.* **5**, 1111–1113

39b. Pace, N. R., Olsen, G. J., and Woese, C. R. (1986) *Cell* **45**, 325–326

References

40. Dragon, F., Payant, C., and Brakier-Gingras, L. (1994) *J. Mol. Biol.* **244**, 74–85
41. Lee, K., Varma, S., SantaLucia, J., Jr., and Cunningham, P. R. (1997) *J. Mol. Biol.* **269**, 732–743
42. Capel, M. S., Kjeldgaard, M., Engelman, D. M., and Moore, P. B. (1988) *J. Mol. Biol.* **200**, 65–87
43. Moore, P. B. (1988) *Nature (London)* **331**, 223–227
44. Svergun, D. I., Koch, M. H. J., Pedersen, J. S., and Serdyuk, I. N. (1994) *J. Mol. Biol.* **240**, 78–86
45. Wadzack, J., Burkhardt, N., Jünemann, R., Diedrich, G., Nierhaus, K. H., Frank, J., Penczek, P., Meerwinck, W., Schmitt, M., Willumeit, R., and Stuhrmann, H. B. (1997) *J. Mol. Biol.* **266**, 343–356
46. Willumeit, R., Forthmann, S., Beckmann, J., Diedrich, G., Ratering, R., Stuhrmann, H. B., and Nierhaus, K. H. (2001) *J. Mol. Biol.* **305**, 167–177
47. Svergun, D. I., and Nierhaus, K. H. (2000) *J. Biol. Chem.* **275**, 14432–14439
48. Herwig, S., Kruft, V., Eckart, K., and Wittmann-Liebold, B. (1993) *J. Biol. Chem.* **268**, 4643–4650
49. Urlaub, H., Kruft, V., Bischof, O., Müller, E.-C., and Wittmann-Liebold, B. (1995) *EMBO J.* **14**, 4578–4588
50. Wang, R., Alexander, R. W., VanLoock, M., Vladimirov, S., Bukhtiyarov, Y., Harvey, S. C., and Cooperman, B. S. (1999) *J. Mol. Biol.* **286**, 521–540
51. Culver, G. M., Heilek, G. M., and Noller, H. F. (1999) *J. Mol. Biol.* **286**, 355–364
52. Wower, J., Kirillov, S. V., Wower, I. K., Guven, S., Hixson, S. S., and Zimmermann, R. A. (2000) *J. Biol. Chem.* **275**, 37887–37894
53. Shapkina, T. G., Dolan, M. A., Babin, P., and Wollenzien, P. (2000) *J. Mol. Biol.* **299**, 615–628
54. Juzumiene, D. I., Shapkina, T. G., and Wollenzien, P. (1995) *J. Biol. Chem.* **270**, 12794–12800
55. Montesano-Roditis, L., Glitz, D. G., Perrault, A. R., and Cooperman, B. S. (1997) *J. Biol. Chem.* **272**, 8695–8703
56. Laughrea, M., and Tam, J. (1992) *Biochemistry* **31**, 12035–12041
57. Hüttenhofer, A., and Noller, H. F. (1994) *EMBO J.* **13**, 3892–3901
58. Czworkowski, J., Odom, O. W., and Hardesty, B. (1991) *Biochemistry* **30**, 4821–4830
59. Urlaub, H., Thiede, B., Müller, E.-C., Brimacombe, R., and Wittmann-Liebold, B. (1997) *J. Biol. Chem.* **272**, 14547–14555
60. Benjamin, D. R., Robinson, C. V., Hendrick, J. P., Hartl, F. U., and Dobson, C. M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7391–7395
61. Glück, A., and Wool, I. G. (1996) *J. Mol. Biol.* **256**, 838–848
61a. Zamecnik, P. (1984) *Trends Biochem. Sci.* **9**, 464–466
62. Schimmel, P., and Wang, C.-C. (1999) *Trends Biochem. Sci.* **24**, 127–128
63. Freifelder, D. (1983) *Molecular Biology*, 2nd ed., Jones and Bartlett, Boston, Massachusetts (p. 419)
64. Yamaguchi, K., and Subramanian, A. R. (2000) *J. Biol. Chem.* **275**, 28466–28482
65. O'Brien, T. W., Fiesler, S. E., Denslow, N. D., Thiede, B., Wittmann-Liebold, B., Mougey, E. B., Sylvester, J. E., and Graack, H.-R. (1999) *J. Biol. Chem.* **274**, 36043–36051
66. Koc, E. C., Burkhart, W., Blackburn, K., Koc, H., Moseley, A., and Spremulli, L. L. (2001) *Protein Sci.* **10**, 471–481
66a. Koc, E. C., Burkhart, W., Blackburn, K., Moyer, M. B., Schlatzer, D. M., Moseley, A., and Spremulli, L. L. (2001) *J. Biol. Chem.* **276**, 43958–43969
67. Moazed, D., Stern, S., and Noller, H. F. (1986) *J. Mol. Biol.* **187**, 399–416
68. Choi, Y. C. (1985) *J. Biol. Chem.* **260**, 12769–12772
69. Rairkar, A., Rubino, H. M., and Lockard, R. E. (1988) *Biochemistry* **27**, 582–592
69a. Decatur, W. A., and Fournier, M. J. (2002) *Trends Biochem. Sci.* **27**, 344–351
70. Carter, A. P., Clemons, W. M., Jr., Brodersen, D. E., Morgan-Warren, R. J., Hartsch, T., Wimberly, B. T., and Ramakrishnan, V. (2001) *Science* **291**, 498–501
71. Leontis, N. B., and Westhof, E. (1998) *J. Mol. Biol.* **283**, 571–583
72. Schnare, M. N., Damberger, S. H., Gray, M. W., and Gutell, R. R. (1996) *J. Mol. Biol.* **256**, 701–719
73. Herzog, M., and Maroteaux, L. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8644–8648
73a. Gutell, R. R. (1994) *Nucleic Acids Res.* **22**, 3502–3507
74. Brimacombe, R., Atmadja, J., Stiege, W., and Schüler, D. (1988) *J. Mol. Biol.* **199**, 115–136
75. Malhotra, A., and Harvey, S. C. (1994) *J. Mol. Biol.* **240**, 308–340
76. Stern, S., Weiser, B., and Noller, H. F. (1988) *J. Biochem. Biophys. Methods* **204**, 447–481
77. Mueller, F., and Brimacombe, R. (1997) *J. Mol. Biol.* **271**, 524–544
77a. Lodmell, J. S., and Dahlberg, A. E. (1997) *Science* **277**, 1262–1267
78. Merryman, C., Moazed, D., Daubresse, G., and Noller, H. F. (1999) *J. Mol. Biol.* **285**, 107–113
79. Larsson, S. L., and Nygård, O. (2001) *Biochemistry* **40**, 3222–3231
79a. Matadeen, R., Sergiev, P., Leonov, A., Pape, T., van der Sluis, E., Mueller, F., Osswald, M., von Knoblauch, K., Brimacombe, R., Bogdanov, A., van Heel, M., and Dontsova, O. (2001) *J. Mol. Biol.* **307**, 1341–1349
80. Garrett, R. (1983) *Trends Biochem. Sci.* **8**, 189–190
81. Porse, B. T., Cundliffe, E., and Garrett, R. A. (1999) *J. Mol. Biol.* **287**, 33–45
82. Conn, G. L., Draper, D. E., Lattman, E. E., and Gittis, A. G. (1999) *Science* **284**, 1171–1174
83. Correll, C. C., Wool, I. G., and Munishkin, A. (1999) *J. Mol. Biol.* **292**, 275–287
84. Macbeth, M. R., and Wool, I. G. (1999) *J. Mol. Biol.* **285**, 965–975
85. Chan, Y.-L., Sitikov, A. S., and Wool, I. G. (2000) *J. Mol. Biol.* **298**, 795–805
86. Wilson, K. S., and Noller, H. F. (1998) *Cell* **92**, 337–349
87. Klein, B. K., Staden, A., and Schlessinger, D. (1985) *J. Biol. Chem.* **260**, 8114–8120
88. Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. D., and Noller, H. F. (2001) *Science* **1**, 1–20
89. Khaitovich, P., and Mankin, A. S. (1999) *J. Mol. Biol.* **291**, 1025–1034
90. Christensen, A., Mathiesen, M., Peattie, D., and Garrett, R. A. (1985) *Biochemistry* **24**, 2284–2291
91. Romby, P., Westhof, E., Toukifimpa, R., Mache, R., Ebel, J.-P., Ehresmann, C., and Ehresmann, B. (1988) *Biochemistry* **27**, 4721–4730
92. Leontis, N. B., and Moore, P. B. (1986) *Biochemistry* **25**, 3916–3925
93. Dontsova, O., Tishkov, V., Dokudovskaya, S., Bogdanov, A., Döring, T., Rinke-Appel, J., Thamm, S., Greuer, B., and Brimacombe, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4125–4129
94. Lee, K. M., and Marshall, A. G. (1986) *Biochemistry* **25**, 8245–8252
95. Funari, S. S., Rapp, G., Perbandt, M., Dierks, K., Vallazza, M., Betzel, C., Erdmann, V. A., and Svergun, D. I. (2000) *J. Biol. Chem.* **275**, 31283–31288
96. Garrett, R. A. (1985) *Nature (London)* **318**, 233–235
97. Olsen, G. J., and Woese, C. R. (1993) *FASEB J.* **7**, 113–123
98. Dijk, J., and Littlechild, J. A. (1979) *Methods Enzymol.* **59**, 481–502
99. Kruft, V., and Wittmann-Liebold, B. (1991) *Biochemistry* **30**, 11781–11787
100. Bielka, H., ed. (1982) *The Eukaryotic Ribosome*, Springer, Berlin
100a. Wool, I. G., Chan, Y. L., and Gluck, A. (1995) *Biochem. Cell Biol.* **73**, 933–947
101. van Beekvelt, C. A., Kooi, E. A., de Graaff-Vincent, M., van't Riet, J., Venema, J., and Raué, H. A. (2000) *J. Mol. Biol.* **296**, 7–17
102. Goldschmidt-Reisin, S., Kitakawa, M., Herfurth, E., Wittmann-Liebold, B., Grohmann, L., and Graack, H.-R. (1998) *J. Biol. Chem.* **273**, 34828–34836
102a. Saveanu, C., Fromont-Racine, M., Harington, A., Ricard, F., Namane, A., and Jacquier, A. (2001) *J. Biol. Chem.* **276**, 15861–15867
102b. Graack, H.-R., Bryant, M. L., and O'Brien, T. W. (1999) *Biochemistry* **38**, 16569–16577
102c. Suzuki, T., Terasaki, M., Takemoto-Hori, C., Hanada, T., Ueda, T., Wada, A., and Watanabe, K. (2001) *J. Biol. Chem.* **276**, 33181–33195
103. Chan, Y.-L., and Wool, I. G. (1988) *J. Biol. Chem.* **263**, 2891–2896
104. Andres, J. L., Johansen, J. W., and Maller, J. L. (1987) *J. Biol. Chem.* **262**, 14389–14393
105. Wilson, K. S., Appelt, K., Badger, J., Tanaka, I., and White, S. W. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7251–7255
105a. Leijonmarck, M., Appelt, K., Badger, J., Liljas, A., Wilson, K. S., and White, S. W. (1988) *Proteins* **3**, 243–251
106. Stern, S., Wilson, R. C., and Noller, H. F. (1986) *J. Mol. Biol.* **192**, 101–110
107. van de Ven, F. J. M., and Hilbers, C. W. (1986) *J. Mol. Biol.* **192**, 389–417
108. Nag, B., Tewari, D. S., Sommer, A., Olson, H. M., Glitz, D. G., and Traut, R. R. (1987) *J. Biol. Chem.* **262**, 9681–9687
109. Wittmann, H.-G. (1976) *Eur. J. Biochem.* **61**, 1–13
110. Brockmöller, J., and Kamp, R. M. (1988) *Biochemistry* **27**, 3372–3381
111. Xiang, R. H., and Lee, J. C. (1989) *J. Biol. Chem.* **264**, 10542–10546
112. Pohl, T., and Wittmann-Liebold, B. (1988) *J. Biol. Chem.* **263**, 4293–4301
113. Yeh, Y.-C., Traut, R. R., and Lee, J. C. (1986) *J. Biol. Chem.* **261**, 14148–14153
114. Capel, M. S., Engelman, D. M., Freeborn, B. R., Kjeldgaard, M., Langer, J. A., Ramakrishnan, V., Schindler, D. G., Schneider, D. K., Schoenborn, B. P., Sillers, I.-Y., Yabuki, S., and Moore, P. B. (1987) *Science* **238**, 1403–1406
115. Liiv, A., Tenson, T., and Remme, J. (1996) *J. Mol. Biol.* **263**, 396–410
116. Nomura, M. (1973) *Science* **179**, 864–873
117. Herold, M., and Nierhaus, K. H. (1987) *J. Biol. Chem.* **262**, 8826–8833
118. Liiv, A., and Remme, J. (1998) *J. Mol. Biol.* **276**, 537–545
119. Green, R., and Noller, H. F. (1999) *Biochemistry* **38**, 1772–1779
120. Khaitovich, P., Tenson, T., Kloss, P., and Mankin, A. S. (1999) *Biochemistry* **38**, 1780–1788
121. Michael, W. M., and Dreyfuss, G. (1996) *J. Biol. Chem.* **271**, 11571–11574
122. Wimberly, B. T., White, S. W., and Ramakrishnan, V. (1997) *Structure* **5**, 1187–1198

References

123. Davies, C., Gerstner, R. B., Draper, D. E., Ramakrishnan, V., and White, S. W. (1998) *EMBO J.* **17**, 4545–4558
124. Markus, M. A., Gerstner, R. B., Draper, D. E., and Torchia, D. A. (1998) *EMBO J.* **17**, 4559–4571
125. Agalarov, S. C., Zheleznyakova, E. N., Selivanova, O. M., Zheleznaya, L. A., Matvienko, N. I., Vasiliev, V. D., and Spirin, A. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 999–1003
126. Nevskaya, N., Tishchenko, S., Nikulin, A., Al-Karadaghi, S., Liljas, A., Ehresmann, B., Ehresmann, C., Garber, M., and Nikonov, S. (1998) *J. Mol. Biol.* **279**, 233–244
127. Kean, J. M., and Draper, D. E. (1985) *Biochemistry* **24**, 5052–5061
128. Svensson, P., Changchien, L., Craven, G. R., and Noller, H. F. (1988) *J. Mol. Biol.* **200**, 301–308
129. Serganov, A., Bénard, L., Portier, C., Ennifar, E., Garber, M., Ehresmann, B., and Ehresmann, C. (2001) *J. Mol. Biol.* **305**, 785–803
130. Kalurachchi, K., and Nikonowicz, E. P. (1998) *J. Mol. Biol.* **280**, 639–654
130a. Tishchenko, S., Nikulin, A., Fomenkova, N., Nevskaya, N., Nikonov, O., Dumas, P., Moine, H., Ehresmann, B., Ehresmann, C., Piendl, W., Lamzin, V., Garber, M., and Nikonov, S. (2001) *J. Mol. Biol.* **311**, 311–324
131. Orr, J. W., Hagerman, P. J., and Williamson, J. R. (1998) *J. Mol. Biol.* **275**, 453–464
132. Hei, Y.-j, Pelech, S. L., Chen, X., Diamond, J., and McNeill, J. H. (1994) *J. Biol. Chem.* **269**, 7816–7823
132a. Martin, K. A., Schalm, S. S., Richardson, C., Romanelli, A., Keon, K. L., and Blenis, J. (2001) *J. Biol. Chem.* **276**, 7884–7891
132b. Martin, K. A., Schalm, S. S., Romanelli, A., Keon, K. L., and Blenis, J. (2001) *J. Biol. Chem.* **276**, 7892–7898
133. Morosyuk, S. V., Cunningham, P. R., and SantaLucia, J., Jr. (2001) *J. Mol. Biol.* **307**, 197–211
134. Davies, C., Bussiere, D. E., Golden, B. L., Porter, S. J., Ramakrishnan, V., and White, S. W. (1998) *J. Mol. Biol.* **279**, 873–888
135. Subramanian, A. R. (1984) *Trends Biochem. Sci.* **9**, 491–494
136. Kim, J., Chubatsu, L. S., Admon, A., Stahl, J., Fellous, R., and Linn, S. (1995) *J. Biol. Chem.* **270**, 13620–13629
137. Sorensen, M. A., Fricke, J., and Pedersen, S. (1998) *J. Mol. Biol.* **280**, 561–569
138. Laughrea, M., and Tam, J. (1991) *Biochemistry* **30**, 11412–11420
139. Nikonov, S., Nevskaya, N., Eliseikina, I., Fomenkova, N., Nikulin, A., Ossina, N., Garber, M., Jonsson, B.-H., Briand, C., Al-Karadaghi, S., Svensson, A., AEvarsson, A., and Liljas, A. (1996) *EMBO J.* **15**, 1350–1359
139a. Zengel, J. M., and Lindahl, L. (1994) *Prog. Nucleic Acid Res. Mol. Biol.* **47**, 331–370
140. Yeh, L.-C. C., and Lee, J. C. (1995) *J. Mol. Biol.* **246**, 295–307
141. Nag, B., Tewari, D. S., Etchison, J. R., Sommer, A., and Traut, R. R. (1986) *J. Biol. Chem.* **261**, 13892–13897
142. Nakagawa, A., Nakashima, T., Taniguchi, M., Hosaka, H., Kimura, M., and Tanaka, I. (1999) *EMBO J.* **18**, 1459–1467
143. Hoffman, D. W., Cameron, C. S., Davies, C., White, S. W., and Ramakrishnan, V. (1996) *J. Mol. Biol.* **264**, 1058–1071
144. Kuhlman, B., Yang, H. Y., Boice, J. A., Fairman, R., and Raleigh, D. P. (1997) *J. Mol. Biol.* **270**, 640–647
145. Lieberman, K. R., Firpo, M. A., Herr, A. J., Nguyenle, T., Atkins, J. F., Gesteland, R. F., and Noller, H. F. (2000) *J. Mol. Biol.* **297**, 1129–1143
146. Dey, D., Bochkariov, D. E., Jokhadze, G. G., and Traut, R. R. (1998) *J. Biol. Chem.* **273**, 1670–1676
146a. Montesano-Roditis, L., Glitz, D. G., Traut, R. R., and Stewart, P. L. (2001) *J. Biol. Chem.* **276**, 14117–14123

146a. Montesano-Roditis, L., Glitz, D. G., Traut, R. R., and Stewart, P. L. (2001) *J. Biol. Chem.* **276**, 14117–14123

146b. Pioletti, M., Schlünzen, F., Harms, J., Zarivach, R., Glühmann, M., Avila, H., Bashan, A., Bartels, H., Auerbach, T., Jacobi, C., Hartsch, T., Yonath, A., and Franceschi, F. (2001) *EMBO J.* **20**, 1829–1839
147. Wahl, M. C., Bourenkov, G. P., Bartunik, H. D., and Huber, R. (2000) *EMBO J.* **19**, 174–186
148. Porse, B. T., Leviev, I., Mankin, A. S., and Garrett, R. A. (1998) *J. Mol. Biol.* **276**, 391–404
149. Holmberg, L., and Noller, H. F. (1999) *J. Mol. Biol.* **289**, 223–233
150. Gonzalo, P., Lavergne, J.-P., and Reboud, J.-P. (2001) *J. Biol. Chem.* **276**, 19762–19769
150a. Zurdo, J., González, C., Sanz, J. M., Rico, M., Remacha, M., and Ballesta, J. P. G. (2000) *Biochemistry* **39**, 8935–8943
150b. Guarinos, E., Remacha, M., and Ballesta, J. P. G. (2001) *J. Biol. Chem.* **276**, 32474–32479
151. Szick-Miranda, K., and Bailey-Serres, J. (2001) *J. Biol. Chem.* **276**, 10921–10928
152. Dunbar, D. A., Gragon, F., Lee, S. J., and Baserga, S. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13027–13032
153. Rohl, R., and Nierhaus, K. H. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 729–733
154. Stoldt, M., Wöhnert, J., Görlach, M., and Brown, L. R. (1998) *EMBO J.* **17**, 6377–6384
155. Lu, M., and Steitz, T. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2023–2028
156. Chan, Y. L., Lin, A., McNally, J., and Wool, I. G. (1987) *J. Biol. Chem.* **262**, 12879–12886
156a. Stoldt, M., Wöhnert, J., Ohlenschläger, O., Görlach, M., and Brown, L. R. (1999) *EMBO J.* **18**, 6508–6521
156b. Johnson, A. W., Lund, E., and Dahlberg, J. (2002) *Trends Biochem. Sci.* **27**, 580 – 585
157. Lin, A., McNally, J., and Wool, I. G. (1983) *J. Biol. Chem.* **258**, 10664–10671
158. Lee, J. C., Henry, B., and Yeh, Y.-C. (1983) *J. Biol. Chem.* **258**, 854–858
159. Mao, H., and Williamson, J. R. (1999) *J. Mol. Biol.* **292**, 345–359
160. Stoffler-Meilicke, M., Stoffler, G., Odom, O. W., Zinn, A., Kramer, G., and Hardesty, B. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5538–5542
161. Trempe, M. R., Ohgi, K., and Glitz, D. G. (1982) *J. Biol. Chem.* **257**, 9822–9829
162. Hausner, T.-P., Geigenmüller, U., and Nierhaus, K. N. (1988) *J. Biol. Chem.* **263**, 13103–13111
163. Prince, J. B., and Garrett, R. A. (1982) *Trends Biochem. Sci.* **7**, 79
164. Denman, R., Nègre, D., Cunningham, P. R., Nurse, K., Colgan, J., Weitzmann, C., and Ofengand, J. (1989) *Biochemistry* **28**, 1012–1019
164a. Noller, H. F. (1993) *FASEB J.* **7**, 87–89
164b. Vila-Sanjurjo, A., and Dahlberg, A. E. (2001) *J. Mol. Biol.* **308**, 457–463
164c. Samaha, R. R., Green, R., and Noller, H. F. (1995) *Nature (London)* **377**, 309–314
164d. Porse, B. T., Thi-Ngoc, H. P., and Garrett, R. A. (1996) *J. Mol. Biol.* **264**, 472–483
164e. Green, R., Samaha, R. R., and Noller, H. F. (1997) *J. Mol. Biol.* **266**, 40–50
165. Lührmann, R., Bald, R., Stöffler-Meilicke, M., and Stöffler, G. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7276–7280
166. Olson, H. M., Nicholson, A. W., Cooperman, B. S., and Glitz, D. G. (1985) *J. Biol. Chem.* **260**, 10326–10331
166a. Green, R., Switzer, C., and Noller, H. F. (1998) *Science* **280**, 286–289
167. Goldman, R. A., Hasan, T., Hall, C. C., Strycharz, W. A., and Cooperman, B. S. (1983) *Biochemistry* **22**, 359–368
167a. Porse, B. T., and Garrett, R. A. (1995) *J. Mol. Biol.* **249**, 1–10
167b. Thompson, J., Kim, D. F., O'Connor, M., Lieberman, K. R., Bayfield, M. A., Gregory, S. T., Green, R., Noller, H. F., and Dahlberg, A. E. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9002–9007
168. Langer, J. A., and Lake, J. A. (1986) *J. Mol. Biol.* **187**, 617–621
169. Ibba, M., Becker, H. D., Stathopoulos, C., Tumbula, D. L., and Söll, D. (2000) *Trends Biochem. Sci.* **25**, 311–316
170. Weisblum, B. (1999) *Trends Biochem. Sci.* **24**, 247–250
171. Yue, D., Kintanar, A., and Horowitz, J. (1994) *Biochemistry* **33**, 8905–8911
172. Schimmel, P. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4521–4522
173. Senger, B., Aphasizhev, R., Walter, P., and Fasiolo, F. (1995) *J. Mol. Biol.* **249**, 45–58
174. Frazer-Abel, A. A., and Hagerman, P. J. (1999) *J. Mol. Biol.* **285**, 581–593
175. Jovine, L., Djordjevic, S., and Rhodes, D. (2000) *J. Mol. Biol.* **301**, 401–414
176. Auffinger, P., and Westhof, E. (1999) *J. Mol. Biol.* **292**, 467–483
177. Perreau, V. M., Keith, G., Holmes, W. M., Przykorska, A., Santos, M. A. S., and Tuite, M. F. (1999) *J. Mol. Biol.* **293**, 1039–1053
178. Wilson, R. K., and Roe, B. A. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 409–413
179. Rich, A. (1978) *Trends Biochem. Sci.* **3**, 34–37
180. Hou, Y.-M., Westhof, E., and Giegé, R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6776–6780
181. Himeno, H., Yoshida, S., Soma, A., and Nishikawa, K. (1997) *J. Mol. Biol.* **268**, 704–711
182. Christian, T., Lipman, R. S. A., Evilia, C., and Hou, Y.-M. (2000) *J. Mol. Biol.* **303**, 503–514
183. Hamann, C. S., and Hou, Y.-M. (2000) *J. Mol. Biol.* **295**, 777–789
184. Drabkin, H. J., and RajBhandary, U. L. (1985) *J. Biol. Chem.* **260**, 5580–5587
185. Woo, N. H., Roe, B. A., and Rich, A. (1980) *Nature (London)* **286**, 346–351
186. Calagan, J. L., Pirtle, R. M., Pirtle, I. L., Kashdan, M. A., Vreman, H. J., and Dudock, B. S. (1980) *J. Biol. Chem.* **255**, 9981–9984
187. Kuchino, Y., Ihara, M., Yabusaki, Y., and Nishimura, S. (1982) *Nature (London)* **298**, 684–685
188. Bai, Y., Fox, D. T., Lacy, J. A., Van Lanen, S. G., and Iwata-Reuyl, D. (2000) *J. Biol. Chem.* **275**, 28731–28738
189. Watanabe, M., Nameki, N., Matsuo-Takasaki, M., Nishimura, S., and Okada, N. (2001) *J. Biol. Chem.* **276**, 2387–2394
190. Crick, F. H. C. (1966) *J. Mol. Biol.* **19**, 548–555
190a. Inagaki, Y., Kojima, A., Bessho, Y., Hori, H., Ohama, T., and Osawa, S. (1995) *J. Mol. Biol.* **251**, 486–492
191. Booth, V., Keizer, D. W., Kamphuis, M. B., Clark-Lewis, I., and Sykes, B. D. (2002) *Biochemistry* **41**, 10418 – 10425
192. Auld, D. S., and Schimmel, P. (1995) *Science* **267**, 1994–1996
193. Pak, M., Willis, I. M., and Schulman, L. H. (1994) *J. Biol. Chem.* **269**, 2277–2282
194. Li, S., Pelka, H., and Schulman, L. H. (1993) *J. Biol. Chem.* **268**, 18335–18339
195. Hong, K.-W., Ibba, M., Weygand-Durasevic, I., Rogers, M. J., Thomann, H.-U., and Söll, D. (1996) *EMBO J.* **15**, 1983–1991

References

196. Nureki, O., Niimi, T., Muramatsu, T., Kanno, H., Kohno, T., Florentz, C., Giegé, R., and Yokoyama, S. (1994) *J. Mol. Biol.* **236**, 710–724
197. Tocchini-Valentini, G., Saks, M. E., and Abelson, J. (2000) *J. Mol. Biol.* **298**, 779–793
198. Yan, W., and Francklyn, C. (1994) *J. Biol. Chem.* **269**, 10022–10027
198a. Shimada, A., Nureki, O., Goto, M., Takahashi, S., and Yokoyama, S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13537–13542
198b. Xu, F., Jiang, G., Li, W., He, X., Jin, Y., and Wang, D. (2002) *Biochemistry* **41**, 8087–8092
199. Normanly, J., and Abelson, J. (1989) *Ann. Rev. Biochem.* **58**, 1029–1049
199a. Sampson, J. R., DiRenzo, A. B., Behlen, L. S., and Uhlenbeck, O. C. (1989) *Science* **243**, 1363–1366
200. Saks, M. E., Sampson, J. R., and Abelson, J. N. (1994) *Science* **263**, 191–197
201. McClain, W. H. (1993) *FASEB J.* **7**, 72–78
202. Ribas de Pouplana, L., and Schimmel, P. (2001) *J. Biol. Chem.* **276**, 6881–6884
203. Sherman, J. M., Thomann, H.-U., and Söll, D. (1996) *J. Mol. Biol.* **256**, 818–828
204. Arnez, J. G., Dock-Bregeon, A.-C., and Moras, D. (1999) *J. Mol. Biol.* **286**, 1449–1459
205. Horowitz, J., Chu, W.-C., Derrick, W. B., Liu, J. C.-H., Liu, M., and Yue, D. (1999) *Biochemistry* **38**, 7737–7746
206. Liu, M., and Horowitz, J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10389–10393
207. Arnez, J. G., and Moras, D. (1997) *Trends Biochem. Sci.* **22**, 211–216
208. Schimmel, P. (1991) *Trends Biochem. Sci.* **16**, 1–3
209. Nureki, O., Vassylyev, D. G., Katayanagi, K., Shimizu, T., Sekine, S.-i, Kigawa, T., Miyazawa, T., Yokoyama, S., and Morikawa, K. (1995) *Science* **267**, 1958–1965
209a. Liu, J., Ibba, M., Hong, K.-W., and Söll, D. (1998) *Biochemistry* **37**, 9836–9842
209b. Ribas de Pouplana, L., and Schimmel, P. (2001) *Trends Biochem. Sci.* **26**, 591–596
210. Zelwer, C., Risler, J. L., and Brunie, S. (1982) *J. Mol. Biol.* **155**, 63–81
211. Xin, Y., Li, W., and First, E. A. (2000) *Biochemistry* **39**, 340–347
212. Ilyin, V. A., Temple, B., Hu, M., Li, G., Yin, Y., Vachette, P., and Carter, C. W., Jr. (2000) *Protein Sci.* **9**, 218–231
213. First, E. A., and Fersht, A. R. (1995) *Biochemistry* **34**, 5030–5043
214. Perona, J. J., Rould, M. A., and Steitz, T. A. (1993) *Biochemistry* **32**, 8758–8771
215. Fersht, A. R., Leatherbarrow, R. J., and Wells, T. N. C. (1986) *Trends Biochem. Sci.* **11**, 321–325
216. Leatherbarrow, R. J., Fersht, A. R., and Winter, G. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7840–7844
217. Desogus, G., Todone, F., Brick, P., and Onesti, S. (2000) *Biochemistry* **39**, 8418–8425
217a. Onesti, S., Desogus, G., Brevet, A., Chen, J., Plateau, P., Blanquet, S., and Brick, P. (2000) *Biochemistry* **39**, 12853–12861
218. Rould, M. A., Perona, J. J., Söll, D., and Steitz, T. A. (1989) *Science* **246**, 1135–1142
219. Rould, M. A., Perona, J. J., and Steitz, T. A. (1991) *Nature (London)* **352**, 213–218
220. Arnez, J. G., and Steitz, T. A. (1994) *Biochemistry* **33**, 7560–7567
221. Cavarelli, J., Delagoutte, B., Eriani, G., Gangloff, J., and Moras, D. (1998) *EMBO J.* **17**, 5438–5448
222. Brick, P., Bhat, T. N., and Blow, D. M. (1988) *J. Mol. Biol.* **208**, 83–98
223. Nureki, O., Vassylyev, D. G., Tateno, M., Shimada, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T. L., Schimmel, P., and Yokoyama, S. (1998) *Science* **280**, 578–582
224. Cusack, S., Yaremchuk, A., and Tukalo, M. (2000) *EMBO J.* **19**, 2351–2361
225. Fourmy, D., Mechulam, Y., and Blanquet, S. (1995) *Biochemistry* **34**, 15681–15688
226. Mechulam, Y., Schmitt, E., Maveyraud, L., Zelwer, C., Nureki, O., Yokoyama, S., Konno, M., and Blanquet, S. (1999) *J. Mol. Biol.* **294**, 1287–1297
227. Arnez, J. G., Harris, D. C., Mitschler, A., Rees, B., Francklyn, C. S., and Moras, D. (1995) *EMBO J.* **14**, 4143–4155
228. Qiu, X., Janson, C. A., Blackburn, M. N., Chhohan, I. K., Hibbs, M., and Abdel-Meguid, S. S. (1999) *Biochemistry* **38**, 12296–12304
229. Arnez, J. G., Augustine, J. G., Moras, D., and Francklyn, C. S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7144–7149
230. Biou, V., Yaremchuk, A., Tukalo, M., and Cusack, S. (1994) *Science* **263**, 1404–1410
231. Cusack, S., Yaremchuk, A., and Tukalo, M. (1996) *EMBO J.* **15**, 2834–2842
232. Eiler, S., Dock-Bregeon, A.-C., Moulinier, L., Thierry, J.-C., and Moras, D. (1999) *EMBO J.* **18**, 6532–6541
233. Archontis, G., Simonson, T., and Karplus, M. (2001) *J. Mol. Biol.* **306**, 307–327
234. Berthet-Colominas, C., Seignovert, L., Härtlein, M., Grotli, M., Cusack, S., and Leberman, R. (1998) *EMBO J.* **17**, 2947–2960
235. Commans, S., Lazard, M., Delort, F., Blanquet, S., and Plateau, P. (1998) *J. Mol. Biol.* **278**, 801–813
236. Reshetnikova, L., Moor, N., Lavrik, O., and Vassylyev, D. G. (1999) *J. Mol. Biol.* **287**, 555–568
237. Logan, D. T., Mazauric, M.-H., Kern, D., and Moras, D. (1995) *EMBO J.* **14**, 4156–4167
238. Xu, Z.-J., Love, M. L., Ma, L. Y. Y., Blum, M., Bronskill, P. M., Bernstein, J., Grey, A. A., Hofmann, T., Camerman, N., and Wong, J. T.-F. (1989) *J. Biol. Chem.* **264**, 4304–4311
239. Jasin, M., Regan, L., and Schimmel, P. (1983) *Nature (London)* **306**, 441–447
240. Webster, T. A., Lathrop, R. H., and Smith, T. F. (1987) *Biochemistry* **26**, 6950–6957
241. Schimmel, P. (1989) *Biochemistry* **28**, 2747–2759
242. Liu, J., Lin, S.-x, Blochet, J.-E., Pézolet, M., and Lapointe, J. (1993) *Biochemistry* **32**, 11390–11396
243. Sampson, J., and Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1033–1037
244. Ogilvie, K. K., Usman, N., Nicoghosian, K., and Cedergren, R. J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5764–5768
244a. Hendrickson, T. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13473–13475
245. Alexander, R. W., and Schimmel, P. (1999) *Biochemistry* **38**, 16359–16365
246. McClain, W. H., Chen, Y.-M., Foss, K., and Schneider, J. (1988) *Science* **242**, 1681–1684
247. Francklyn, C., and Schimmel, P. (1989) *Nature (London)* **337**, 478–481
247a. Sardesai, N. Y., Green, R., and Schimmel, P. (1999) *Biochemistry* **38**, 12080–12088
248. Rogers, M. J., and Söll, D. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6627–6631
249. Khan, A. S., and Roe, B. A. (1988) *Science* **241**, 74–79
250. Connolly, B. A., Von der Haar, F., and Eckstein, F. (1980) *J. Biol. Chem.* **255**, 11301–11307
251. Ferscht, A. R. (1987) *Biochemistry* **26**, 8031–8037
252. Leatherbarrow, R. J., and Fersht, A. R. (1987) *Biochemistry* **26**, 8524–8528
253. Ferscht, A. R. (1988) *Biochemistry* **27**, 1577–1580
254. Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., Freeman, New York
255. Xin, Y., Li, W., and First, E. A. (2000) *J. Mol. Biol.* **303**, 299–310
256. Bedouelle, H., and Winter, G. (1986) *Nature (London)* **320**, 371–373
257. Starzyk, R. M., Koontz, S. W., and Schimmel, P. (1982) *Nature (London)* **298**, 136–140
258. Fersht, A. R. (1977) *Enzyme Structure and Mechanism*, 1st ed., Freeman, San Francisco, California (p. 283)
259. Fersht, A. R. (1998) *Science* **280**, 541
259a. Mursinna, R. S., Lincecum, T. L., Jr., and Martinis, S. A. (2001) *Biochemistry* **40**, 5376–5381
259b. Tardif, K. D., Liu, M., Vitseva, O., Hou, Y.-M., and Horowitz, J. (2001) *Biochemistry* **40**, 8118–8125
259c. Tardif, K. D., and Horowitz, J. (2002) *Nucleic Acids Res.* **30**, 2538–2545
260. Gao, W., Goldman, E., and Jakubowski, H. (1994) *Biochemistry* **33**, 11528–11535
261. Jakubowski, H. (1995) *J. Biol. Chem.* **270**, 17672–17673
262. Jakubowski, H. (1996) *Biochemistry* **35**, 8252–8259
263. Serre, L., Verdon, G., Choinowski, T., Hervouet, N., Risler, J.-L., and Zelwer, C. (2001) *J. Mol. Biol.* **306**, 863–876
263a. Kelley, S. O., Steinberg, S. V., and Schimmel, P. (2001) *J. Biol. Chem.* **276**, 10607–10611
264. Yarus, M. (1992) *Trends Biochem. Sci.* **17**, 171–174
265. Farabaugh, P. J., and Björk, G. R. (1999) *EMBO J.* **18**, 1427–1434
266. Rodnina, M. V., and Wintermeyer, W. (2001) *Trends Biochem. Sci.* **26**, 124–130
267. Ibba, M., Curnow, A. W., and Söll, D. (1997) *Trends Biochem. Sci.* **22**, 39–42
268. Wallis, N. G., Dardel, F., and Blanquet, S. (1995) *Biochemistry* **34**, 7668–7677
268a. Horiuchi, K. Y., Harpel, M. R., Shen, L., Luo, Y., Rogers, K. C., and Copeland, R. A. (2001) *Biochemistry* **40**, 6450–6457
269. Curnow, A. W., Hong,, K-w., Yuan, R., Kim, S.-i, Martins, O., Winkler, W., Henkin, T. M., and Söll, D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11819–11826
270. Böck, A., Forchhammer, K., Heider, J., and Baron, C. (1991) *Trends Biochem. Sci.* **16**, 463–467
271. Yarus, M. (2000) *Science* **287**, 440–441
272. Stathopoulos, C., Li, T., Longman, R., Vothknecht, U. C., Becker, H. D., Ibba, M., and Söll, D. (2000) *Science* **287**, 479–482
272a. Stathopoulos, C., Kim, W., Li, T., Anderson, I., Deutsch, B., Palioura, S., Whitman, W., and Söll, D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14292–14297
273. Wool, I. G. (1996) *Trends Biochem. Sci.* **21**, 164–165
274. Martinis, S. A., Plateau, P., Cavarelli, J., and Florentz, C. (1999) *EMBO J.* **18**, 4591–4596
275. Putney, S. D., and Schimmel, P. (1981) *Nature (London)* **291**, 632–635
276. Romby, P., Caillet, J., Ebel, C., Sacerdot, C., Graffe, M., Eyermann, F., Brunel, C., Moine, H., Ehresmann, C., Ehresmann, B., and Springer, M. (1996) *EMBO J.* **15**, 5976–5987
277. Gerken, S. C., and Arfin, S. M. (1984) *J. Biol. Chem.* **259**, 11160–11161
278. Lambowitz, A. M., and Perlman, P. S. (1990) *Trends Biochem. Sci.* **15**, 440–444
279. Abramochkin, G., and Shrader, T. E. (1996) *J. Biol. Chem.* **271**, 22901–22907
280. Li, J., and Pickart, C. M. (1995) *Biochemistry* **34**, 139–147
281. Dou, X., Limmer, S., and Kreutzer, R. (2001) *J. Mol. Biol.* **305**, 451–458
281a. Lund, E., and Dahlberg, J. E. (1998) *Science* **282**, 2003–2004
281b. Sarkar, S., Azad, A. K., and Hopper, A. K. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14366–14371

References

282. Schimmel, P., and Ribas De Pouplana, L. (2000) *Trends Biochem. Sci.* **25**, 207–209

283. Kisselev, L., Frolova, L., and Haenni, A.-L. (1993) *Trends Biochem. Sci.* **18**, 263–267

284. Sissler, M., Delorme, C., Bond, J., Ehrlich, S. D., Renault, P., and Francklyn, C. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8985–8990

285. Raben, N., Nichols, R., Dohlman, J., McPhie, P., Sridhar, V., Hyde, C., Leff, R., and Plotz, P. (1994) *J. Biol. Chem.* **269**, 24277–24283

286. Ge, Q., Trieu, E. P., and Targoff, I. N. (1994) *J. Biol. Chem.* **269**, 28790–28797

287. Winter, R. B., Morrissey, L., Gauss, P., Gold, L., Hsu, T., and Karam, J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7822–7826

288. Peabody, D. S. (1989) *J. Biol. Chem.* **264**, 5031–5035

289. Gupta, K. C., and Patwardhan, S. (1988) *J. Biol. Chem.* **263**, 8553–8556

290. Newton, D. T., Creuzenet, C., and Mangroo, D. (1999) *J. Biol. Chem.* **274**, 22143–22146

291. RajBhandary, U. L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1325–1327

292. Li, S., Kumar, N. V., Varshney, U., and RajBhandary, U. L. (1996) *J. Biol. Chem.* **271**, 1022–1028

293. Takeuchi, N., Kawakami, M., Omori, A., Ueda, T., Spremulli, L. L., and Watanabe, K. (1998) *J. Biol. Chem.* **273**, 15085–15090

294. Shine, J., and Dalgarno, L. (1975) *Nature (London)* **254**, 34–38

295. Steitz, J. A., and Jakes, K. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4734–4738

295a. van Dieijen, G., Zipori, P., van Prooijen, W., and van Duin, J. (1978) *Eur. J. Biochem.* **90**, 571–580

296. Wu, X.-Q., Iyengar, P., and RajBhandary, U. L. (1996) *EMBO J.* **15**, 4734–4739

296a. Roll-Mecak, A., Shin, B.-S., Dever, T. E., and Burley, S. K. (2001) *Trends Biochem. Sci.* **26**, 705–709

297. Biou, V., Shu, F., and Ramakrishnan, V. (1995) *EMBO J.* **14**, 4056–4064

297a. Morosyuk, S. V., Cunningham, P. R., SantaLucia, J., Jr. (2001) *J. Mol. Biol.* **307**, 197–211

298. Moreau, M., de Cock, E., Fortier, P.-L., Garcia, C., Albaret, zC., Blanquet, S., Lallemand, J.-Y., and Dardel, F. (1997) *J. Mol. Biol.* **266**, 15–22

299. Yu, N.-J., and Spremulli, L. L. (1997) *Biochemistry* **36**, 14827–14835

300. La Teana, A., Pon, C. L., and Gualerzi, C. O. (1996) *J. Mol. Biol.* **256**, 667–675

301. Karimi, R., Pavlov, M. Y., Heurgué-Hamard, V., Buckingham, R. H., and Ehrenberg, M. (1998) *J. Mol. Biol.* **281**, 241–252

302. Dahlquist, K. D., and Puglisi, J. D. (2000) *J. Mol. Biol.* **299**, 1–15

303. Meinnel, T., Sacerdot, C., Graffe, M., Blanquet, S., and Springer, M. (1999) *J. Mol. Biol.* **290**, 825–837

304. Garcia, C., Fortier, P.-L., Blanquet, S., Lallemand, J.-Y., and Dardel, F. (1995) *J. Mol. Biol.* **254**, 247–259

305. Kozak, M. (1983) *Microbiol. Rev.* **47**, 1–45

306. Moldave, R. (1985) *Ann. Rev. Biochem.* **54**, 1109–1149

307. Lorsch, J. R., and Herschlag, D. (1999) *EMBO J.* **18**, 6705–6717

308. Rhoads, R. E. (1999) *J. Biol. Chem.* **274**, 30337–30340

309. Rhoads, R. E. (1993) *J. Biol. Chem.* **268**, 3017–3020

310. Browning, K. S., Maia, D. M., Lax, S. R., and Ravel, J. M. (1987) *J. Biol. Chem.* **262**, 538–541

311. Pestova, T. V., and Hellen, C. U. T. (1999) *Trends Biochem. Sci.* **24**, 85–87

312. Altmann, M., and Trachsel, H. (1993) *Trends Biochem. Sci.* **18**, 429–432

313. Sonenberg, N., Hershey, J. W. B., and Mathews, M. B., eds. (2000) *Translational Control of Gene Expression*, Cold Spring Harbor Lab. Press, Plainview, New York

314. Carey, M., and Smale, S. T. (2000) *Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques*, Cold Spring Harbor Lab. Press, Plainview, New York

314a. Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I., and Hellen, C. U. T. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7029–7036

315. Jackson, R. J. (1998) *Nature (London)* **394**, 829–831

316. Richter-Cook, N. J., Dever, T. E., Hensold, J. O., and Merrick, W. C. (1998) *J. Biol. Chem.* **273**, 7579–7587

317. Richter, N. J., Rogers, G. W., Jr., Hensold, J. O., and Merrick, W. C. (1999) *J. Biol. Chem.* **274**, 35415–35424

318. Melander, Y., Holmberg, L., and Nygård, O. (1997) *J. Biol. Chem.* **272**, 3254–3258

319. Korneeva, N. L., Lamphear, B. J., Hennigan, F. L. C., and Rhoads, R. E. (2000) *J. Biol. Chem.* **275**, 41369–41376

319a. Browning, K. S., Gallie, D. R., Hershey, J. W. B., Hinnebusch, A. G., Maitra, U., Merrick, W. C., and Norbury, C. (2001) *Trends Biochem. Sci.* **26**, 284

320. Valásek, L., Phan, L., Schoenfeld, L. W., Valásková, V., and Hinnebusch, A. G. (2001) *EMBO J.* **20**, 891–904

321. Raychaudhuri, P., and Maitra, U. (1986) *J. Biol. Chem.* **261**, 7723–7728

322. Proud, C. G. (1986) *Trends Biochem. Sci.* **11**, 73–77

323. Gil, J., Esteban, M., and Roth, D. (2000) *Biochemistry* **39**, 7521–7530

324. Sudhakar, A., Krishnamoorthy, T., Jain, A., Chatterjee, U., Hasnain, S. E., Kaufman, R. J., and Ramaiah, K. V. A. (1999) *Biochemistry* **38**, 15398–15405

325. Kimball, S. R., Fabian, J. R., Pavitt, G. D., Hinnebusch, A. G., and Jefferson, L. S. (1998) *J. Biol. Chem.* **273**, 12841–12845

326. De Haro, C., Méndez, R., and Santoyo, J. (1996) *FASEB J.* **10**, 1378–1387

327. Rose, D. W., Welch, W. J., Kramer, G., and Hardesty, B. (1989) *J. Biol. Chem.* **264**, 6239–6244

328. Williams, D. D., Pavitt, G. D., and Proud, C. G. (2001) *J. Biol. Chem.* **276**, 3733–3742

328a. Asano, K., Krishnamoorthy, T., Phan, L., Pavitt, G. D., Hinnebusch, A. G. (1999) *EMBO J.* **18**, 1673–1688

329. Mueller, P. P., Grueter, P., Hinnebusch, A. G., and Trachsel, H. (1998) *J. Biol. Chem.* **273**, 32870–32877

330. Qiu, H., Dong, J., Hu, C., Francklyn, C. S., and Hinnebusch, A. G. (2001) *EMBO J.* **20**, 1425–1438

330a. Wang, X., Paulin, F. E. M., Campbell, L. E., Gomez, E., O'Brien, K., Morrice, N., and Proud, C. G. (2001) *EMBO J.* **20**, 4349–4359

330b. Wang, X., Li, W., Williams, M., Terada, N., Alessi, D. R., and Proud, C. G. (2001) *EMBO J.* **20**, 4370–4379

330c. Cho, S., and Hoffman, D. W. (2002) *Biochemistry* **41**, 5730–5742

331. Chen, J.-J., and London, I. M. (1995) *Trends Biochem. Sci.* **20**, 105–108

332. Dholakia, J. N., Mueser, T. C., Woodley, C. L., Parkhurst, L. J., and Wahba, A. J. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6746–6750

333. Gingras, A.-C., Raught, B., and Sonenberg, N. (1999) *Ann. Rev. Biochem.* **68**, 913–963

334. Herbert, T. P., Kilhams, G. R., Batty, I. H., and Proud, C. G. (2000) *J. Biol. Chem.* **275**, 11249–11256

334a. Karim, M. M., Hughes, J. M. X., Warwicker, J., Scheper, G. C., Proud, C. G., and McCarthy, J. E. G. (2001) *J. Biol. Chem.* **276**, 20750–20757

335. Dever, T. E. (1999) *Trends Biochem. Sci.* **24**, 398–403

336. De Gregorio, E., Preiss, T., and Hentze, M. W. (1999) *EMBO J.* **18**, 4865–4874

337. Korneeva, N. L., Lamphear, B. J., Hennigan, F. L. C., Merrick, W. C., and Rhoads, R. E. (2001) *J. Biol. Chem.* **276**, 2872–2879

338. Lamphear, B. J., Kirchweger, R., Skern, T., and Rhoads, R. E. (1995) *J. Biol. Chem.* **270**, 21975–21983

339. McCubbin, W. D., Edery, I., Altmann, M., Sonenberg, N., and Kay, C. M. (1988) *J. Biol. Chem.* **263**, 17663–17671

340. Keiper, B. D., Lamphear, B. J., Deshpande, A. M., Jankowska-Anyszka, M., Aamodt, E. J., Blumenthal, T., and Rhoads, R. E. (2000) *J. Biol. Chem.* **275**, 10590–10596

341. Rao, G. N. (2000) *J. Biol. Chem.* **275**, 16993–16999

341a. Niedzwiecka, A., Stepinski, J., Darzynkiewicz, E., Sonenberg, N., and Stolarski, R. (2002) *Biochemistry* **41**, 12140–12148

342. Naranda, T., Strong, W. B., Menaya, J., Fabbri, B. J., and Hershey, J. W. B. (1994) *J. Biol. Chem.* **269**, 14465–14472

343. Dominguez, D., Altmann, M., Benz, J., Baumann, U., and Trachsel, H. (1999) *J. Biol. Chem.* **274**, 26720–26726

344. Chappell, S. A., Edelman, G. M., and Mauro, V. P. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1536–1541

345. Hentze, M. W. (1997) *Science* **275**, 500–501

345a. Fernandez, J., Yaman, I., Mishra, R., Merrick, W. C., Snider, M. D., Lamers, W. H., and Hatzoglou, M. (2001) *J. Biol. Chem.* **276**, 12285–12291

346. Spahn, C. M. T., Kieft, J. S., Grassucci, R. A., Penczek, P. A., Zhou, K., Doudna, J. A., and Frank, J. (2001) *Science* **291**, 1959–1962

347. Hudder, A., and Werner, R. (2000) *J. Biol. Chem.* **275**, 34586–34591

348. Kozak, M. (1997) *EMBO J.* **16**, 2482–2492

349. Geballe, A. P., and Morris, D. R. (1994) *Trends Biochem. Sci.* **19**, 159–164

350. Cigan, A. M., Feng, L., and Donahue, T. F. (1988) *Science* **242**, 93–97

350a. Galy, B. (2001) *Trends Biochem. Sci.* **26**, 220

351. Chen, C.-y, and Sarnow, P. (1995) *Science* **268**, 415–417

352. Otero, L. J., Ashe, M. P., and Sachs, A. B. (1999) *EMBO J.* **18**, 3153–3163

353. Le, H., Browning, K. S., and Gallie, D. R. (2000) *J. Biol. Chem.* **275**, 17452–17462

354. Kozlov, G., Trempe, J.-F., Khaleghpour, K., Kahvejian, A., Ekiel, I., and Gehring, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4409–4413

355. Das, S., Maiti, T., Das, K., and Maitra, U. (1997) *J. Biol. Chem.* **272**, 31712–31718

356. Joe, Y. A., and Park, M. H. (1994) *J. Biol. Chem.* **269**, 25916–25921

357. Park, M. H., Wolff, E. C., and Folk, J. E. (1993) *Trends Biochem. Sci.* **18**, 475–479

358. Joao, H. C., Csonga, R., Klier, H., Koettnitz, K., Auer, M., and Eder, J. (1995) *Biochemistry* **34**, 14703–14711

359. Chetverin, A. B., and Spirin, A. S. (1982) *Biochim. Biophys. Acta.* **683**, 153–179

360. Gnirke, A., Geigenmüller, U., Rheinberger, H.-J., and Nierhaus, K. H. (1989) *J. Biol. Chem.* **264**, 7291–7301

361. Thompson, R. C. (1988) *Trends Biochem. Sci.* **13**, 91–93

362. Abel, K., and Jurnak, F. (1996) *Structure* **4**, 229–238

363. Jurnak, F. (1985) *Science* **230**, 32–36

References

364. Sprinzl, M. (1994) *Trends Biochem. Sci.* **19**, 245 – 250
365. Krásny, L., Mesters, J. R., Tieleman, L. N., Kraal, B., Fucík, V., Hilgenfeld, R., and Jonák, J. (1998) *J. Mol. Biol.* **283**, 371 – 381
366. Song, H., Parsons, M. R., Rowsell, S., Leonard, G., and Phillips, S. E. V. (1999) *J. Mol. Biol.* **285**, 1245 – 1256
367. Krab, I. M., and Parmeggiani, A. (1999) *Biochemistry* **38**, 13035 – 13041
368. Andersen, G. R., Thirup, S., Spremulli, L. L., and Nyborg, J. (2000) *J. Mol. Biol.* **297**, 421 – 436
369. Jiang, Y., Nock, S., Nesper, M., Sprinzl, M., and Sigler, P. B. (1996) *Biochemistry* **35**, 10269 – 10278
370. Uetsuki, T., Naito, A., Nagata, S., and Kaziro, Y. (1989) *J. Biol. Chem.* **264**, 5791 – 5798
371. Riis, B., Rattan, S. I. S., Clark, B. F. C., and Merrick, W. C. (1990) *Trends Biochem. Sci.* **15**, 420 – 424
372. Cavallius, J., and Merrick, W. C. (1998) *J. Biol. Chem.* **273**, 28752 – 28758
373. Edmonds, B. T., Bell, A., Wyckoff, J., Condeelis, J., and Leyh, T. S. (1998) *J. Biol. Chem.* **273**, 10288 – 10295
374. Condeelis, J. (1995) *Trends Biochem. Sci.* **20**, 169 – 170
374a. Pleiss, J. A., and Uhlenbeck, O. C. (2001) *J. Mol. Biol.* **308**, 895 – 905
375. Cunningham, P. R., Nurse, K., Weitzmann, C. J., Nègre, D., and Ofengand, J. (1992) *Biochemistry* **31**, 7629 – 7637
376. Gornicki, P., Nurse, K., Hellmann, W., Boublik, M., and Ofengand, J. (1984) *J. Biol. Chem.* **259**, 10493 – 10498
377. Weller, J., and Hill, W. E. (1994) *J. Biol. Chem.* **269**, 19369 – 19374
378. Lynch, S. R., and Puglisi, J. D. (2001) *J. Mol. Biol.* **306**, 1023 – 1035
378a. Ericson, G., Minchew, P., and Wollenzien, P. (1995) *J. Mol. Biol.* **250**, 407 – 419
379. VanLoock, M. S., Easterwood, T. R., and Harvey, S. C. (1999) *J. Mol. Biol.* **285**, 2069 – 2078
380. Lodmell, J. S., and Dahlberg, A. E. (1997) *Science* **277**, 1262 – 1267
381. Stark, H., Rodnina, M. V., Rinke-Appel, J., Brimacombe, R., Wintermeyer, W., and van Heel, M. (1997) *Nature (London)* **389**, 403 – 406
382. Hall, C. C., Smith, J. E., and Cooperman, B. S. (1985) *Biochemistry* **24**, 5702 – 5711
383. Douthwaite, S., and Aagaard, C. (1993) *J. Mol. Biol.* **232**, 725 – 731
384. Villsen, I. D., Vester, B., and Douthwaite, S. (1999) *J. Mol. Biol.* **286**, 365 – 374
385. Moazed, D., Robertson, J. M., and Noller, H. F. (1988) *Nature (London)* **334**, 362 – 364
385a. Rodnina, M. V., Pape, T., Fricke, R., Kuhn, L., and Wintermeyer, W. (1996) *J. Biol. Chem.* **271**, 646 – 652
385b. Piepenburg, O., Pape, T., Pleiss, J. A., Wintermeyer, W., Uhlenbeck, O. C., and Rodnina, M. V. (2000) *Biochemistry* **39**, 1734 – 1738
385c. Knudsen, C., Wieden, H.-J., and Rodnina, M. V. (2001) *J. Biol. Chem.* **276**, 22183 – 22190
385d. Vogeley, L., Palm, G. J., Mesters, J. R., and Hilgenfeld, R. (2001) *J. Biol. Chem.* **276**, 17149 – 17155
385e. Simonson, A. B., and Lake, J. A. (2002) *Nature (London)* **416**, 281 – 285
386. Hecht, S. M., Tan, K. H., Chinault, A. C., and Arcari, P. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 437 – 441
386a. Sprinzl, M., Kucharzewski, M., Hobbs, J. B., and Cramer, F. (1977) *Eur. J. Biochem.* **78**, 55 – 61
387. Taiji, M., Yokoyama, S., and Miyazawa, T. (1985) *Biochemistry* **24**, 5776 – 5780
387a. Wagner, T., and Sprinzl, M. (1983) *Biochemistry* **22**, 94 – 98
387b. Limmer, St., Vogtherr, M., Nawrot, B., Hillenbrand, R., and Sprinzl, M. (1997) *Angew. Chem. Int. Ed. Engl.* **36**, 2485 – 2489
387c. Polacek, N., Gaynor, M., Yassin, A., and Mankin, A. S. (2001) *Nature (London)* **411**, 498 – 501
387d. Katunin, V. I., Muth, G. W., Strobel, S. A., Wintermeyer, W., and Rodnina, M. V. (2002) *Molecular Cell* **10**, 339 – 346
387e. Bayfield, M. A., Dahlberg, A. E., Schulmeister, U., Dorner, S., and Barta, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10096 – 10101
388. Lim, V. I., and Spirin, A. S. (1986) *J. Mol. Biol.* **188**, 565 – 577
389. Samaha, R. R., Green, R., and Noller, H. F. (1995) *Nature (London)* **377**, 309 – 314
390. Kloss, P., Xiong, L., Shinabarger, D. L., and Mankin, A. S. (1999) *J. Mol. Biol.* **294**, 93 – 101
391. Gregory, S. T., and Dahlberg, A. E. (1999) *J. Mol. Biol.* **285**, 1475 – 1483
392. Agrawal, R. K., Penczek, P., Grassucci, R. A., and Frank, J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6134 – 6138
393. Wilson, K. S., and Noller, H. F. (1998) *Cell* **92**, 131 – 139
394. Savelsbergh, A., Matassova, N. B., Rodnina, M. V., and Wintermeyer, W. (2000) *J. Mol. Biol.* **300**, 951 – 961
394a. Katunin, V. I., Savelsbergh, A., Rodnina, M. V., and Wintermeyer, W. (2002) *Biochemistry* **41**, 12806 – 12812
395. Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B. F. C., and Nyborg, J. (1995) *Science* **270**, 1464 – 1472
396. Nyborg, J., Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Clark, B. F. C., and Reshetnikova, L. (1996) *Trends Biochem. Sci.* **21**, 81 – 82
397. Van Ness, B. G., Howard, J. B., and Bodley, J. W. (1980) *J. Biol. Chem.* **255**, 10710 – 10716
397a. Nierhaus, K. H., Wadzack, J., Burkhardt, N., Jünemann, R., Meerwinck, W., Willumeit, R., and Stuhrmann, H. B. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 945 – 950
397b. Frank, J., and Agrawal, R. K. (1998) *Biophys. J.* **74**, 589 – 594
397c. Culver, G. M. (2001) *Structure* **9**, 751 – 758
398. Rodnina, M. V., Savelsbergh, A., Katunin, V. I., and Wintermeyer, W. (1997) *Nature (London)* **385**, 37 – 41
398a. Agrawal, R. K., Linde, J., Sengupta, J., Nierhaus, K. H., and Frank, J. (2001) *J. Mol. Biol.* **311**, 777 – 787
399. Gabashvili, I. S., Agrawal, R. K., Grassucci, R., Squires, C. L., Dahlberg, A. E., and Frank, J. (1999) *EMBO J.* **18**, 6501 – 6507
400. Mohr, D., Wintermeyer, W., and Rodnina, M. V. (2000) *EMBO J.* **19**, 3458 – 3464
401. Uchiumi, T., Hori, K., Nomura, T., and Hachimori, A. (1999) *J. Biol. Chem.* **274**, 27578 – 27582
402. VanLoock, M. S., Agrawal, R. K., Gabashvili, I. S., Qi, L., Frank, J., and Harvey, S. C. (2000) *J. Mol. Biol.* **304**, 507 – 515
403. Mesters, J. R., Potapov, A. P., de Graaf, J. M., and Kraal, B. (1994) *J. Mol. Biol.* **242**, 644 – 654
404. Montesano-Roditis, L., and Glitz, D. G. (1994) *J. Biol. Chem.* **269**, 6458 – 6470
405. Bhangu, R., and Wollenzien, P. (1992) *Biochemistry* **31**, 5937 – 5944
406. Graifer, D. M., Juzumiene, D. I., Karpova, G. G., and Wollenzien, P. (1994) *Biochemistry* **33**, 6201 – 6206
407. Sergiev, P. V., Lavrik, I. N., Wlasoff, V. A., Dokudovskaya, S. S., Dontsova, O. A., Bogdanov, A. A., and Brimacombe, R. (1997) *RNA* **3**, 464 – 475
408. Qin, S., Moldave, K., and McLaughlin, C. S. (1987) *J. Biol. Chem.* **262**, 7802 – 7807
409. Triana-Alonso, F. J., Chakraburtty, K., and Nierhaus, K. H. (1995) *J. Biol. Chem.* **270**, 20473 – 20478
410. Kambampati, R., Pellegrino, C., Paiva, A., Huang, L., Mende-Mueller, L., and Chakraburtty, K. (2000) *J. Biol. Chem.* **275**, 16963 – 16968
410a. Warner, J. R. (1999) *Trends Biochem. Sci.* **24**, 437 – 440
410b. Schleif, R., Hess, W., Finkelstein, S., and Ellis, D. (1973) *J. Bacteriol.* **115**, 9 – 14
410c. Andersson, D. I., Bohman, K., Isaksson, L. A., and Kurland, C. G. (1982) *Mol. Gen. Genet.* **187**, 467 – 472
410d. Dintzis, H. M. (1961) *Proc. Natl. Acad. Sci. U.S.A.* **47**, 247 – 261
410e. Goustin, A. S., and Wilt, F. H. (1982) *Biochim. Biophys. Acta.* **699**, 22 – 27
411. Tate, W. P., and Brown, C. M. (1992) *Biochemistry* **31**, 2443 – 2450
412. Caskey, C. T. (1980) *Trends Biochem. Sci.* **5**, 234 – 237
413. Stansfield, I., Jones, K. M., and Tuite, M. F. (1995) *Trends Biochem. Sci.* **20**, 489 – 491
413a. Wilson, K. S., Ito, K., Noller, H. F., and Nakamura, Y. (2000) *Nature Struct. Biol.* **7**, 866 – 870
413b. Kisselev, L. L., and Buckingham, R. H. (2000) *Trends Biochem. Sci.* **25**, 561 – 566
414. Tate, W. P., Hornig, H., and Luhrmann, R. (1983) *J. Biol. Chem.* **258**, 10360 – 10365
415. Poole, E. S., Brown, C. M., and Tate, W. P. (1995) *EMBO J.* **14**, 151 – 158
416. Zhang, S., Rydén-Aulin, M., and Isaksson, L. A. (1998) *J. Mol. Biol.* **284**, 1243 – 1246
417. Pavlov, M. Y., Freistroffer, D. V., MacDougall, J., Buckingham, R. H., and Ehrenberg, M. (1997) *EMBO J.* **16**, 4134 – 4141
418. Pavlov, M. Y., Freistroffer, D. V., Heurguè-Hamard, V., Buckingham, R. H., and Ehrenberg, M. (1997) *J. Mol. Biol.* **273**, 389 – 401
418a. Ito, K., Frolova, L., Seit-Nebi, A., Karamyshev, A., Kisselev, L., and Nakamura, Y. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8494 – 8499
419. Velichutina, I. V., Hong, J. Y., Mesecar, A. D., Chernoff, Y. O., and Liebman, S. W. (2001) *J. Mol. Biol.* **305**, 715 – 727
419a. Nakamura, Y., Ito, K., and Ehrenberg, M. (2000) *Cell* **101**, 349 – 352
419b. Song, H., Mugnier, P., Das, A. K., Webb, H. M., Evans, D. R., Tuite, M. F., Hemmings, B. A., and Barford, D. (2000) *Cell* **100**, 311 – 321
420. Wickner, R. B., Edskes, H. K., Maddelein, M.-L., Taylor, K. L., and Moriyama, H. (1999) *J. Biol. Chem.* **274**, 555 – 558
421. Sparrer, H. E., Santoso, A., Szoka, F. C., Jr., and Weissman, J. S. (2000) *Science* **289**, 595 – 599
422. King, C.-Y. (2001) *J. Mol. Biol.* **307**, 1247 – 1260
423. Selmer, M., Al-Karadaghi, S., Hirokawa, G., Kaji, A., and Liljas, A. (1999) *Science* **286**, 2349 – 2352
424. Yoshida, T., Uchiyama, S., Nakano, H., Kashimori, H., Kijima, H., Ohshima, T., Saihara, Y., Ishino, T., Shimahara, H., Yoshida, T., Yokose, K., Ohkubo, T., Kaji, A., and Kobayashi, Y. (2001) *Biochemistry* **40**, 2387 – 2396
424a. Kim, K. K., Min, K., and Suh, S. W. (2000) *EMBO J.* **19**, 2362 – 2370
424b. Karimi, R., Pavlov, M. Y., Buckingham, R. H., and Ehrenberg, M. (1999) *Mol. Cell. Biol.* **3**, 601 – 609
425. Rolland, N., Janosi, L., Block, M. A., Shuda, M., Teyssier, E., Miège, C., Chéniclet, C., Carde, J.-P., Kaji, A., and Joyard, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5464 – 5469
425a. Lancaster, L., Kiel, M. C., Kaji, A., and Noller, H. F. (2002) *Cell* **111**, 129 – 140

References

426. Gast, F.-U., Peters, F., and Pingoud, A. (1987) *J. Biol. Chem.* **262**, 11920 – 11926
427. Ibba, M., and Söll, D. (1999) *Science* **286**, 1893 – 1897
428. Parker, J., Johnston, T. C., Borgia, P. T., Holtz, G., Remaut, E., and Fiers, W. (1983) *J. Biol. Chem.* **258**, 10007 – 10012
429. Precup, J., and Parker, J. (1987) *J. Biol. Chem.* **262**, 11351 – 11355
429a. LaRiviere, F. J., Wolfson, A. D., and Uhlenbeck, O. C. (2001) *Science* **294**, 165 – 168
429b. Ibba, M. (2001) *Science* **294**, 70 – 71
429c. Stahl, G., McCarty, G. P., and Farabaug, P. J. (2002) *Trends Biochem. Sci.* **27**, 178 – 183
430. Kurland, C. G. (1987) *Trends Biochem. Sci.* **12**, 169 – 171
431. Metzler, D. E. (1977) *Biochemistry; The Chemical Reactions of Living Cells*, Academic Press, New York (pp. 914 – 915)
432. Oresic, M., and Shalloway, D. (1998) *J. Mol. Biol.* **281**, 31 – 48
433. Bennetzen, J. L., and Hall, B. D. (1982) *J. Biol. Chem.* **257**, 3026 – 3031
434. Heurgué-Hamard, V., Mora, L., Guarneros, G., and Buckingham, R. H. (1996) *EMBO J.* **15**, 2826 – 2833
435. Schmitt, E., Mechulam, Y., Fromant, M., Plateau, P., and Blanquet, S. (1997) *EMBO J.* **16**, 4760 – 4769
436. Fromant, M., Plateau, P., Schmitt, E., Mechulam, Y., and Blanquet, S. (1999) *Biochemistry* **38**, 4982 – 4987
436a. van Hoof, A., Frischmeyer, P. A., Dietz, H. C., and Parker, R. (2002) *Science* **295**, 2262 – 2264
436b. Maquat, L. E. (2002) *Science* **295**, 2221 – 2222
437. Keiler, K. C., Waller, P. R. H., and Sauer, R. T. (1996) *Science* **271**, 990 – 994
438. Karzai, A. W., and Sauer, R. T. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3040 – 3044
438a. Roche, E. D., and Sauer, R. T. (2001) *J. Biol. Chem.* **276**, 28509 – 28515
438b. Gillet, R., and Felden, B. (2001) *EMBO J.* **20**, 2966 – 2976
439. Nameki, N., Tadaki, T., Muto, A., and Himeno, H. (1999) *J. Mol. Biol.* **289**, 1 – 7
440. Barends, S., Wower, J., and Kraal, B. (2000) *Biochemistry* **39**, 2652 – 2658
440a. Barends, S., Karzai, A. W., Sauer, R. T., and Wower, J. (2001) *J. Mol. Biol.* **314**, 9 – 21
440b. Zwieb, C., Guven, S. A., Wower, I. K., and Wower, J. (2001) *Biochemistry* **40**, 9587 – 9595
440c. Stagg, S. M., Frazer-Abel, A. A., Hagerman, P. J., and Harvey, S. C. (2001) *J. Mol. Biol.* **309**, 727 – 735
441. Littauer, U. Z., and Inouye, H. (1973) *Ann. Rev. Biochem.* **42**, 439 – 470
442. Riddle, D. L., and Carbon, J. (1973) *Nature New Biol.* **242**, 230 – 234
443. Magliery, T. J., Anderson, J. C., and Schultz, P. G. (2001) *J. Mol. Biol.* **307**, 755 – 769
444. Murgola, E. J., Prather, N. E., Mims, B. H., Pagel, F. T., and Hijazi, K. A. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4936 – 4939
445. Bossi, L., and Smith, D. M. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6105 – 6109
446. Murgola, E. J., Hijazi, K. A., Göringer, H. U., and Dahlberg, A. E. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4162 – 4165
447. Surguchov, A. P. (1988) *Trends Biochem. Sci.* **13**, 120 – 123
448. Hatfield, D. (1985) *Trends Biochem. Sci.* **10**, 201 – 204
448a. Srinivasan, G., James, C. M., and Krzycki, J. A. (2002) *Science* **296**, 1459 – 1462
448b. Hao, B., Gong, W., Ferguson, T. K., James, C. M., Krzycki, J. A., and Chan, M. K. (2002) *Science* **296**, 1462 – 1466
449. Engelberg-Kulka, H., and Schoulaker-Schwarz, R. (1988) *Trends Biochem. Sci.* **13**, 419 – 421
450. Stadtman, T. C. (1996) *Ann. Rev. Biochem.* **65**, 83 – 100
451. Kryukov, G. V., Kryukov, V. M., and Gladyshev, V. N. (1999) *J. Biol. Chem.* **274**, 33888 – 33897
452. Lescure, A., Gautheret, D., Carbon, P., and Krol, A. (1999) *J. Biol. Chem.* **274**, 38147 – 38154
453. Saito, Y., Hayashi, T., Tanaka, A., Watanabe, Y., Suzuki, M., Saito, E., and Takahashi, K. (1999) *J. Biol. Chem.* **274**, 2866 – 2871
454. Low, S. C., and Berry, M. J. (1996) *Trends Biochem. Sci.* **21**, 203 – 208
455. Suppmann, S., Persson, B. C., and Böck, A. (1999) *EMBO J.* **18**, 2284 – 2293
456. Ioudovitch, A., and Steinberg, S. V. (1999) *J. Mol. Biol.* **290**, 365 – 371
457. Forchhammer, K., Leinfelder, W., and Böck, A. (1989) *Nature (London)* **342**, 453 – 456
458. Kromayer, M., Wilting, R., Tormay, P., and Böck, A. (1996) *J. Mol. Biol.* **262**, 413 – 420
459. Liu, Z., Reches, M., and Engelberg-Kulka, H. (1999) *J. Mol. Biol.* **294**, 1073 – 1086
460. Rother, M., Wilting, R., Commans, S., and Böck, A. (2000) *J. Mol. Biol.* **299**, 351 – 358
461. Low, S. C., Grundner-Culemann, E., Harney, J. W., and Berry, M. J. (2000) *EMBO J.* **19**, 6882 – 6890
461a. Nasim, M. T., Jaenecke, S., Belduz, A., Kollmus, H., Flohé, L., and McCarthy, J. E. G. (2000) *J. Biol. Chem.* **275**, 14846 – 14852
462. Copeland, P. R., Fletcher, J. E., Carlson, B. A., Hatfield, D. L., and Driscoll, D. M. (2000) *EMBO J.* **19**, 306 – 314
463. Normanly, J., Masson, J.-M., Kleina, L. G., Abelson, J., and Miller, J. H. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6548 – 6552
464. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) *Science* **244**, 182 – 188
465. Anthony-Cahill, S. J., Griffith, M. C., Noren, C. J., Suich, D. J., and Schultz, P. G. (1989) *Trends Biochem. Sci.* **14**, 400 – 403
466. Mendel, D., Ellman, J. A., Chang, Z., Veenstra, D. L., Kollman, P. A., and Schultz, P. G. (1992) *Science* **266**, 1798 – 1802
467. Chung, H.-H., Benson, D. R., and Schultz, P. G. (1993) *Science* **259**, 806 – 809
468. Moore, B., Persson, B. C., Nelson, C. C., Gesteland, R. F., and Atkins, J. F. (2000) *J. Mol. Biol.* **298**, 195 – 209
469. Hohsaka, T., Ashizuka, Y., Sasaki, H., Murakami, H., and Sisido, M. (1999) *J. Am. Chem. Soc.* **121**, 12194 – 12195
470. Kowal, A. K., Köhrer, C., and RajBhandary, U. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2268 – 2273
471. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) *Science* **292**, 498 – 500
472. Döring, V., Mootz, H. D., Nangle, L. A., Hendrickson, T. L., de Crécy-Lagard, V., Schimmel, P., and Marlière, P. (2001) *Science* **292**, 501 – 504
473. Bain, J. D., Switzer, C., Chamberlin, A. R., and Benner, S. A. (1992) *Nature (London)* **356**, 537 – 539
474. Temple, G. F., Dozy, A. M., Roy, K. L., and Kan, Y. W. (1982) *Nature (London)* **296**, 537 – 540
475. Ho, Y.-S., and Kan, Y. W. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2185 – 2188
476. Ryoji, M., Hsia, K., and Kaji, A. (1983) *Trends Biochem. Sci.* **8**, 88 – 90
477. Hatfield, D., and Oroszlan, S. (1990) *Trends Biochem. Sci.* **15**, 186 – 190
478. Jacks, T., and Varmus, H. E. (1985) *Science* **230**, 1237 – 1242
479. Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J., and Varmus, H. E. (1988) *Nature (London)* **331**, 280 – 283
480. Varmus, H. (1988) *Science* **240**, 1427 – 1434
481. Alam, S. L., Atkins, J. F., and Gesteland, R. F. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14177 – 14179
482. Gesteland, R. F., and Atkins, J. F. (1996) *Ann. Rev. Biochem.* **65**, 741 – 768
483. Kastelein, R. A., Remaut, E., Fiers, W., and van Duin, J. (1982) *Nature (London)* **295**, 35 – 41
484. Dayhuff, T. J., Atkins, J. F., and Gesteland, R. F. (1986) *J. Biol. Chem.* **261**, 7491 – 7500
484a. Harger, J. W., Meskauskas, A., and Dinman, J. D. (2002) *Trends Biochem. Sci.* **27**, 448 – 454
485. Spanjaard, R. A., and van Duin, J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7967 – 7971
486. Klovins, J., and van Duin, J. (1999) *J. Mol. Biol.* **294**, 875 – 884
487. Barak, Z., Lindsley, D., and Gallant, J. (1996) *J. Mol. Biol.* **256**, 676 – 684
487a. Chamorro, M., Parkin, N., and Varmus, H. E. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 713 – 717
487b. Chen, X., Chamorro, M., Lee, S. I., Shen, L. X., Hines, J. V., Tinoco, I., Jr., and Varmus, H. E. (1995) *EMBO J.* **14**, 842 – 852
487c. Chen, X., Kang, H., Shen, L. X., Chamorro, M., Varmus, H. E., and Tinoco, I., Jr. (1996) *J. Mol. Biol.* **260**, 479 – 483
488. Giedroc, D. P., Theimer, C. A., and Nixon, P. L. (2000) *J. Mol. Biol.* **298**, 167 – 185
489. Marczinke, B., Fisher, R., Vidakovic, M., Bloys, A. J., and Brierley, I. (1998) *J. Mol. Biol.* **284**, 205 – 225
490. Marczinke, B., Hagervall, T., and Brierley, I. (2000) *J. Mol. Biol.* **295**, 179 – 191
491. Liphardt, J., Napthine, S., Kontos, H., and Brierley, I. (1999) *J. Mol. Biol.* **288**, 321 – 335
492. Kim, Y.-G., Su, L., Maas, S., O'Neill, A., and Rich, A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14234 – 14239
493. Matsufuji, S., Matsufuji, T., Wills, N. M., Gesteland, R. F., and Atkins, J. F. (1996) *EMBO J.* **15**, 1360 – 1370
494. Adamski, F. M., Atkins, J. F., and Gesteland, R. F. (1996) *J. Mol. Biol.* **261**, 357 – 371
495. Herr, A. J., Gesteland, R. F., and Atkins, J. F. (2000) *EMBO J.* **19**, 2671 – 2680
495a. Herr, A. J., Wills, N. M., Nelson, C. C., Gesteland, R. F., and Atkins, J. F. (2001) *J. Mol. Biol.* **311**, 445 – 452
496. Weissmann, C., Billeter, M. A., Goodman, H. M., Hindley, J., and Weber, H. (1973) *Ann. Rev. Biochem.* **42**, 303 – 328
497. Beekwilder, J., Nieuwenhuizen, R., Poot, R., and van Duin, J. (1996) *J. Mol. Biol.* **256**, 8 – 19
498. Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volckaert, G., and Ysebaert, M. (1976) *Nature (London)* **260**, 500 – 507
499. Min Jou, W., Haegeman, G., Ysebaert, M., and Fiers, W. (1972) *Nature (London)* **237**, 82 – 88
500. Fiers, W., Contreras, R., Duerinck, F., Haegeman, G. H., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volckaert, C., and Ysebaert, M. (1976) *Nature (London)* **260**, 500 – 507
501. Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Merregaert, J., Min Jou, W., Raeymakers, A., Volckaert, G., Ysebaert, M., Van de Kerckhove, J., Nolf, F., and Van Montagu, M. (1975) *Nature (London)* **256**, 273 – 278
502. Jackson, R. J., Howell, M. T., and Kaminski, A. (1990) *Trends Biochem. Sci.* **15**, 477 – 483
503. Pilipenko, E. V., Gmyl, A. P., Maslova, S. V., Belov, G. A., Sinyakov, A. N., Huang, M., Brown, T. D. K., and Agol, V. I. (1994) *J. Mol. Biol.* **241**, 398 – 414
504. Gan, W., and Rhoads, R. E. (1996) *J. Biol. Chem.* **271**, 623 – 626

References

505. Kieft, J. S., Zhou, K., Jubin, R., Murray, M. G., Lau, J. Y. N., and Doudna, J. A. (1999) *J. Mol. Biol.* **292**, 513–529

506. Lott, W. B., Takyar, S. S., Tuppen, J., Crawford, D. H. G., Harrison, M., Sloots, T. P., and Gowans, E. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4916–4921

506a. Guo, L., Allen, E., and Miller, W. A. (2000) *RNA* **6**, 1808–1820

507. Kaplan, S., Atherly, A. G., and Barrett, A. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 689–692

508. Cozzone, A. J. (1981) *Trends Biochem. Sci.* **6**, 108–110

509. Baracchini, E., and Bremer, H. (1988) *J. Biol. Chem.* **263**, 2597–2602

510. Murray, K. D., and Bremer, H. (1996) *J. Mol. Biol.* **259**, 41–57

511. Metzger, S., Dror,, IB, Aizenman, E., Schreiber, G., Toone, M., Friesen, J. D., Cashel, M., and Glaser, G. (1988) *J. Biol. Chem.* **263**, 15699–15704

512. Hernandez, V. J., and Cashel, M. (1995) *J. Mol. Biol.* **252**, 536–549

512a. Barker, M. M., Gaal, T., Josaitis, C. A., and Gourse, R. L. (2001) *J. Mol. Biol.* **305**, 673–688

512b. Gottesman, S., and Maurizi, M. R. (2001) *Science* **293**, 614 – 615

513. Wagner, E. G. H., and Kurland, C. G. (1980) *Biochemistry* **19**, 1234–1240

514. Nègre, D., Cortay, J.-C., Donini, P., and Cozzone, A. J. (1989) *Biochemistry* **28**, 1814–1819

515. Marianovsky, I., Aizenman, E., Engelberg-Kulka, H., and Glaser, G. (2001) *J. Biol. Chem.* **276**, 5975–5984

516. Swarbrick, J. D., Bashtannyk, T., Maksel, D., Zhang, X.-R., Blackburn, G. M., Gayler, K. R., and Gooley, P. R. (2000) *J. Mol. Biol.* **302**, 1165–1177

517. Kim, B. K., Zamecnik, P., Taylor, G., Guo, M. J., and Blackburn, G. M. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11056–11058

518. Luo, J., Jankowski, J., Knobloch, M., van der Giet, M., Gardanis, K., Russ, T., Vahlensieck, U., Neumann, J., Schmitz, W., Tepel, M., Deng, M. C., Zidek, W., and Schlüter, H. (1999) *FASEB J.* **13**, 695–705

519. Ingram, S. W., Stratemann, S. A., and Barnes, L. D. (1999) *Biochemistry* **38**, 3649–3655

520. Cartwright, J. L., and McLennan, A. G. (1999) *J. Biol. Chem.* **274**, 8604–8610

521. Yen, T. J., Machlin, P. S., and Cleveland, D. W. (1988) *Nature (London)* **334**, 580–585

522. Hunt, T. (1988) *Nature (London)* **334**, 567–568

523. Cleveland, D. W. (1988) *Trends Biochem. Sci.* **13**, 339–343

524. Louie, K., and Dowhan, W. (1980) *J. Biol. Chem.* **255**, 1124–1127

525. Fedorov, A. N., and Baldwin, T. O. (1997) *J. Biol. Chem.* **272**, 32715–32718

525a. Patzelt, H., Rüdiger, S., Brehmer, D., Kramer, G., Vorderwülbecke, S., Schaffitzel, E., Waitz, A., Hesterkamp, T., Dong, L., Schneider-Mergener, J., Bukau, B., and Deuerling, E. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14244–14249

525b. Hartl, F. U., and Hayer-Hartl, M. (2002) *Science* **295**, 1852–1858

525c. Stoller, G., Rücknagel, K. P., Nierhaus, K. H., Schmid, F. X., Fischer, G., and Rahfeld, J.-U. (1995) *EMBO J.* **14**, 4939–4948

525d. Maier, R., Scholz, C., and Schmid, F. X. (2001) *J. Mol. Biol.* **314**, 1181–1190

525e. Richardson, A., Landry, S. J., and Georgopoulos, C. (1998) *Trends Biochem. Sci.* **23**, 138–143

525f. Keskin, O., Bahar, I., Flatow, D., Covell, D. G., and Jernigan, R. L. (2002) *Biochemistry* **41**, 491–501

525g. Flaux, J., Bertelsen, E. B., Horwich, A. L., and Wüthrich, K. (2002) *Nature (London)* **418**, 207–211

525h. Llorca, O., Martín-Benito, J., Grantham, J., Ritco-Vonsovici, M., Willison, K. R., Carrascosa, J. L., and Valpuesta, J. M. (2001) *EMBO J.* **20**, 4065–4075

526. Ruddon, R. W., and Bedows, E. (1997) *J. Biol. Chem.* **272**, 3125–3128

527. Ellis, R. J. (2000) *Trends Biochem. Sci.* **25**, 210–212

528. Sakikawa, C., Taguchi, H., Makino, Y., and Yoshida, M. (1999) *J. Biol. Chem.* **274**, 21251–21256

529. Ma, J., Sigler, P. B., Xu, Z., and Karplus, M. (2000) *J. Mol. Biol.* **302**, 303–313

530. Schäfer, U., Beck, K., and Müller, M. (1999) *J. Biol. Chem.* **274**, 24567–24574

531. Jones, C. H., Danese, P. N., Pinkner, J. S., Silhavy, T. J., and Hultgren, S. J. (1997) *EMBO J.* **16**, 6394–6406

532. Frydman, J., and Höhfeld, J. (1997) *Trends Biochem. Sci.* **22**, 87–92

533. Rüdiger, S., Schneider-Mergener, J., and Bukau, B. (2001) *EMBO J.* **20**, 1042–1050

534. Wilson, M. R., and Easterbrook-Smith, S. B. (2000) *Trends Biochem. Sci.* **25**, 95–98

535. Zha, J., Weiler, S., Oh, K. J., Wei, M. C., and Korsmeyer, S. J. (2000) *Science* **290**, 1761–1765

536. Schubert, U., Antón, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W., and Bennink, J. R. (2000) *Nature (London)* **404**, 770–774

537. Turner, G. C., and Varshavsky, A. (2000) *Science* **289**, 2117–2120

538. Scheffner, M., and Whitaker, N. J. (2001) *Nature (London)* **410**, 882–883

539. Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) *Science* **286**, 1888–1893

540. Bonifacino, J. S. (1996) *Nature (London)* **384**, 405–406

541. Arfin, S. M., and Bradshaw, R. A. (1988) *Biochemistry* **27**, 7979–7984

542. Rajagopalan, P. T. R., Grimme, S., and Pei, D. (2000) *Biochemistry* **39**, 779–790

542a. Deng, H., Callender, R., Zhu, J., Nguyen, K. T., and Pei, D. (2002) *Biochemistry* **41**, 10563–10569

543. Solbiati, J., Chapman-Smith, A., Miller, J. L., Miller, C. G., and Cronan, J. E., JR. (1999) *J. Mol. Biol.* **290**, 607–614

544. Giglione, C., Serero, A., Pierre, M., Boisson, B., and Meinnel, T. (2000) *EMBO J.* **19**, 5916–5929

544a. Serero, A., Giglione, C., and Meinnel, T. (2001) *J. Mol. Biol.* **314**, 695–708

545. Boissel, J.-P., Kasper, T. J., and Bunn, H. F. (1988) *J. Biol. Chem.* **263**, 8443–8449

546. Ben-Bassat, A., and Bauer, K. (1987) *Nature (London)* **326**, 315

547. Polevoda, B., and Sherman, F. (2000) *J. Biol. Chem.* **275**, 36479–36482

548. Strauch, A. R., and Rubenstein, P. A. (1984) *J. Biol. Chem.* **259**, 7224–7229

549. Farazi, T. A., Waksman, G., and Gordon, J. I. (2001) *Biochemistry* **40**, 6335–6343

550. Utsumi, T., Sato, M., Nakano, K., Takemura, D., Iwata, H., and Ishisaka, R. (2001) *J. Biol. Chem.* **276**, 10505–10513

551. Wilcox, C., Hu, J.-S., and Olson, E. N. (1987) *Science* **238**, 1275–1278

552. Towler, D. A., Gordon, J. I., Adams, S. P., and Glaser, L. (1988) *Ann. Rev. Biochem.* **57**, 69–99

553. Wu, Z., Demma, M., Strickland, C. L., Radisky, E. S., Poulter, C. D., Le, H. V., and Windsor, W. T. (1999) *Biochemistry* **38**, 11239–11249

553a. Matsuyama, S., Yokota, N., and Tokuda, H. (1997) *EMBO J.* **16**, 6947 – 6955

554. Aletta, J. M., Cimato, T. R., and Ettinger, M. J. (1998) *Trends Biochem. Sci.* **23**, 89–91

555. Zobel-Thropp, P., Gary, J. D., and Clarke, S. (1998) *J. Biol. Chem.* **273**, 29283–29286

556. Chapman-Smith, A., and Cronan, J. E., Jr. (1999) *Trends Biochem. Sci.* **24**, 359–363

557. Campeau, E., and Gravel, R. A. (2001) *J. Biol. Chem.* **276**, 12310–12316

558. Craig, A. G., Jimenez, E. C., Dykert, J., Nielsen, D. B., Gulyas, J., Abogadie, F. C., Porter, J., Rivier, J. E., Cruz, L. J., Olivera, B. M., and McIntosh, J. M. (1997) *J. Biol. Chem.* **272**, 4689–4698

559. Dierks, T., Lecca, M. R., Schlotterhose, P., Schmidt, B., and von Figura, K. (1999) *EMBO J.* **18**, 2084–2091

560. Stubbe, J. (1996) *Science* **274**, 1152–1153

561. Kelleher, N. L., Hendrickson, C. L., and Walsh, C. T. (1999) *Biochemistry* **38**, 15623–15630

562. Kupke, T., Kempter, C., Gnau, V., Jung, G., and Götz, F. (1994) *J. Biol. Chem.* **269**, 5653–5659

563. Rapoport, T. A. (1992) *Science* **258**, 931–935

564. Corsi, A. K., and Schekman, R. (1996) *J. Biol. Chem.* **271**, 30299–30302

565. Freymann, D. M., Keenan, R. J., Stroud, R. M., and Walter, P. (1997) *Nature (London)* **385**, 361–364

566. Montoya, G., Svensson, C., Luirink, J., and Sinning, I. (1997) *Nature (London)* **385**, 365–368

567. Peluso, P., Herschlag, D., Nock, S., Freymann, D. M., Johnson, A. E., and Walter, P. (2000) *Science* **288**, 1640–1643

568. Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., and Doudna, J. A. (2000) *Science* **287**, 1232–1239

569. Batey, R. T., Sagar, M. B., and Doudna, J. A. (2001) *J. Mol. Biol.* **307**, 229–246

570. Diener, J. L., and Wilson, C. (2000) *Biochemistry* **39**, 12862–12874

571. Walter, P., Keenan, R., and Schmitz, U. (2000) *Science* **287**, 1212–1213

571a. Peluso, P., Shan, S.-o, Nock, S., Herschlag, D., and Walter, P. (2001) *Biochemistry* **40**, 15224–15233

571b. Pool, M. R., Stumm, J., Fulga, T. A., Sinning, I., and Dobberstein, B. (2002) *Science* **297**, 1345–1348

572. de Leeuw, E., te Kaat, K., Moser, C., Menestrina, G., Demel, R., de Kruijff, B., Oudega, B., Luirink, J., and Sinning, I. (2000) *EMBO J.* **19**, 531–541

573. Fulga, T. A., Sinning, I., Dobberstein, B., and Pool, M. R. (2001) *EMBO J.* **20**, 2338–2347

574. Bibi, E., Herskovits, A. A., Bochkareva, E. S., and Zelazny, A. (2001) *Trends Biochem. Sci.* **26**, 15–16

575. Lütcke, H. (1995) *Eur. J. Biochem.* **228**, 531–550

575a. Huang, Q., Abdulrahman, S., Yin, J., and Zwieb, C. (2002) *Biochemistry* **41**, 11362 – 11371

575b. Hainzl, T., Huang, S., and Sauer-Eriksson, A. E. (2002) *Nature (London)* **417**, 767 – 771

576. Weichenrieder, O., Wild, K., Strub, K., and Cusack, S. (2000) *Nature (London)* **408**, 167–173

576a. Beckman, R. (2001) *Cell* **107**, 361–72

577. Young, B. P., Craven, R. A., Reid, P. J., Willer, M., and Stirling, C. J. (2001) *EMBO J.* **20**, 262–271

578. Römisch, K. (2001) *Trends Biochem. Sci.* **26**, 13

579. Beswick, V., Baleux, F., Huynh-Dinh, T., Képès, F., Neumann, J.-M., and Sanson, A. (1996) *Biochemistry* **35**, 14717–14724

580. Tschantz, W. R., Paetzel, M., Cao, G., Suciu, D., Inouye, M., and Dalbey, R. E. (1995) *Biochemistry* **34**, 3935–3941

581. Chatterjee, S., Suciu, D., Dalbey, R. E., Kahn, P. C., and Inouye, M. (1995) *J. Mol. Biol.* **245**, 311–314

582. VanValkenburgh, C., Chen, X., Mullins, C., Fang, H., and Green, N. (1999) *J. Biol. Chem.* **274**, 11519 – 11525

582a. Weihofen, A., Binns, K., Lemberg, M. K., Ashman, K., and Martoglio, B. (2002) *Science* **296**, 2215 – 2218

583. Suzuki, T., Yan, Q., and Lennarz, W. J. (1998) *J. Biol. Chem.* **273**, 10083 – 10086

584. Netzer, W. J., and Hartl, F. U. (1997) *Nature (London)* **388**, 343 – 349

585. Wickner, W., and Leonard, M. R. (1996) *J. Biol. Chem.* **271**, 29514 – 29516

586. Kim, J., Miller, A., Wang, L., Müller, J. P., and Kendall, D. A. (2001) *Biochemistry* **40**, 3674 – 3680

587. Verner, K., and Schatz, G. (1988) *Science* **241**, 1307 – 1313

588. van der Does, C., Manting, E. H., Kaufmann, A., Lutz, M., and Driessen, A. J. M. (1998) *Biochemistry* **37**, 201 – 210

589. Collinson, I., Breyton, C., Duong, F., Tziatzios, C., Schubert, D., Or, E., Rapoport, T., and Kühlbrandt, W. (2001) *EMBO J.* **20**, 2462 – 2471

590. Sianidis, G., Karamanou, S., Vrontou, E., Boulias, K., Repanas, K., Kyrpides, N., Politou, A. S., and Economou, A. (2001) *EMBO J.* **20**, 961 – 970

591. Kluger, R., and Smyth, T. (1981) *J. Am. Chem. Soc.* **103**, 1216 – 1218

592. Yahr, T. L., and Wickner, W. T. (2000) *EMBO J.* **19**, 4393 – 4401

593. Driessen, A. J. M. (1992) *Trends Biochem. Sci.* **17**, 219 – 223

594. Dalbey, R. E., and Robinson, C. (1999) *Trends Biochem. Sci.* **24**, 17 – 22

595. Delgado-Partin, V. M., and Dalbey, R. E. (1998) *J. Biol. Chem.* **273**, 9927 – 9934

596. Gafvelin, G., Sakaguchi, M., Andersson, H., and von Heijne, G. (1997) *J. Biol. Chem.* **272**, 6119 – 6127

597. Rietveld, A. G., Koorengevel, M. C., and de Kruijff, B. (1995) *EMBO J.* **14**, 5506 – 5513

598. Jongbloed, J. D. H., Martin, U., Antelmann, H., Hecker, M., Tjalsma, H., Venema, G., Bron, S., van Dijl, J. M., and Müller, J. (2000) *J. Biol. Chem.* **275**, 41350 – 41357

599. Sambasivarao, D., Turner, R. J., Simala-Grant, J. L., Shaw, G., Hu, J., and Weiner, J. H. (2000) *J. Biol. Chem.* **275**, 22526 – 22531

600. Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., and Hughes, C. (2000) *Nature (London)* **405**, 914 – 919

601. Buchanan, S. K. (2001) *Trends Biochem. Sci.* **26**, 3 – 6

602. Reits, E. A. J., Vos, J. C., Grommé, M., and Neefjes, J. (2000) *Nature (London)* **404**, 774 – 778

603. Samuelson, J. C., Chen, M., Jiang, F., Möller, I., Wiedmann, M., Kuhn, A., Phillips, G. J., and Dalbey, R. E. (2000) *Nature (London)* **406**, 637 – 640

604. Tokatlidis, K., and Schatz, G. (1999) *J. Biol. Chem.* **274**, 35285 – 35288

605. Schatz, G. (1998) *Nature (London)* **395**, 439 – 440

606. McNew, J. A., and Goodman, J. M. (1996) *Trends Biochem. Sci.* **21**, 54 – 58

607. Subramani, S. (1996) *J. Biol. Chem.* **271**, 32483 – 32486

608. Hell, K., Neupert, W., and Stuart, R. A. (2001) *EMBO J.* **20**, 1281 – 1288

609. von Heijne, G. (1989) *Nature (London)* **341**, 456 – 458

610. Kuhn, A., Wickner, W., and Kreil, G. (1986) *Nature (London)* **322**, 335 – 339

611. Bibi, E. (1998) *Trends Biochem. Sci.* **23**, 51 – 55

612. Dalbey, R. E., Kahn, A., and Wickner, W. (1987) *J. Biol. Chem.* **262**, 13241 – 13245

613. Yamane, K., Ichehara, S., and Mizushima, S. (1987) *J. Biol. Chem.* **263**, 2358 – 2362

614. Meijer, A. B., Spruijt, R. B., Wolfs, C. J. A. M., and Hemminga, M. A. (2001) *Biochemistry* **40**, 5081 – 5086

615. Soekarjo, M., Eisenhawer, M., Kuhn, A., and Vogel, H. (1996) *Biochemistry* **35**, 1232 – 1241

616. Moore, K. E., and Miura, S. (1988) *J. Biol. Chem.* **263**, 11575 – 11583

617. von Heijne, G., Wickner, W., and Dalbey, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3363 – 3366

617a. Anderson, R. G. W., and Jacobson, K. (2002) *Science* **296**, 1821– 1825

618. Thomas-Reetz, A. C., and De Camilli, P. (1994) *FASEB J.* **8**, 209 – 216

619. Rothman, J. E., and Wieland, F. T. (1996) *Science* **272**, 227 – 234

620. Nicholls, D. G. (1994) *Proteins, Transmitters and Synapses*, Blackwell Scientific Publications, Oxford

621. Bean, A. J., Zhang, X., and Hökfelt, T. (1994) *FASEB J.* **8**, 630 – 638

622. Huttner, W. B., Gerdes, H.-H., and Rosa, P. (1991) *Trends Biochem. Sci.* **16**, 27 – 30

623. Guo, W., Grant, A., and Novick, P. (1999) *J. Biol. Chem.* **274**, 23558 – 23564

624. Peters, C., Bayer, M. J., Bühler, S., Andersen, J. S., Mann, M., and Mayer, A. (2001) *Nature (London)* **409**, 581 – 588

624a. Pain, R., ed. (2000) *Mechanisms of Protein Folding*, 2nd ed., Oxford Univ. Press, London

625. Holst, B., Bruun, A. W., Kielland-Brandt, M. C., and Winther, J. R. (1996) *EMBO J.* **15**, 3538 – 3546

626. Pirkl, F., and Buchner, J. (2001) *J. Mol. Biol.* **308**, 795 – 806

626a. Gilbert, H. F. (1997) *J. Biol. Chem.* **272**, 29399 – 29402

627. Qiao, Z.-S., Guo, Z.-Y., and Feng, Y.-M. (2001) *Biochemistry* **40**, 2662 – 2668

628. Tu, B. P., Ho-Schleyer, S. C., Travers, K. J., and Weissman, J. S. (2000) *Science* **290**, 1571 – 1573

629. Miranker, A., Robinson, C. V., Radford, S. E., and Dobson, C. M. (1996) *FASEB J.* **10**, 93 – 101

630. Balbach, J., Steegborn, C., Schindler, T., and Schmid, F. X. (1999) *J. Mol. Biol.* **285**, 829 – 842

631. Arrington, C. B., Teesch, L. M., and Robertson, A. D. (1999) *J. Mol. Biol.* **285**, 1265 – 1275

632. Rumbley, J., Hoang, L., Mayne, L., and Englander, S. W. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 105 – 112

633. Steegborn, C., Schneider-Hassloff, H., Zeeb, M., and Balbach, J. (2000) *Biochemistry* **39**, 7910 – 7919

634. Arai, M., Ikura, T., Semisotnov, G. V., Kihara, H., Amemiya, Y., and Kuwajima, K. (1998) *J. Mol. Biol.* **275**, 149 – 162

635. Panick, G., Malessa, R., Winter, R., Rapp, G., Frye, K. J., and Royer, C. A. (1998) *J. Mol. Biol.* **275**, 389 – 402

636. Radford, S. E. (2000) *Trends Biochem. Sci.* **25**, 611 – 618

637. Anfinsen, C. B. (1973) *Science* **181**, 223 – 230

638. Dill, K. A. (1990) *Biochemistry* **29**, 7133 – 7155

639. Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000) *Trends Biochem. Sci.* **25**, 331 – 339

640. Baker, D. (2000) *Nature (London)* **405**, 39 – 42

641. Clarke, D. T., Doig, A. J., Stapley, B. J., and Jones, G. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7232 – 7237

642. Nölting, B., Golbik, R., Neira, J. L., Soler-Gonzalez, A. S., Schreiber, G., and Fersht, A. R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 826 – 830

643. Zhou, Y., and Karplus, M. (1999) *J. Mol. Biol.* **293**, 917 – 951

644. Damaschun, G., Damaschun, H., Gast, K., and Zirwer, D. (1999) *J. Mol. Biol.* **291**, 715 – 725

645. Pace, C. N. (1990) *Trends Biochem. Sci.* **15**, 14 – 17

646. Rees, D. C., and Robertson, A. D. (2001) *Protein Sci.* **10**, 1187– 1194

647. Andersen, N. H., Cort, J. R., Liu, Z., Sjoberg, S. J., and Tong, H. (1996) *J. Am. Chem. Soc.* **118**, 10309 – 10310

648. Swint-Kruse, L., and Robertson, A. D. (1995) *Biochemistry* **34**, 4724 – 4732

649. Yang, A.-S., and Honig, B. (1993) *J. Mol. Biol.* **231**, 459 – 474

650. Yang, A.-S., and Honig, B. (1994) *J. Mol. Biol.* **237**, 602 – 614

651. Pace, C. N. (2001) *Biochemistry* **40**, 310 – 313

652. Sippl, M. J. (1996) *J. Mol. Biol.* **260**, 644 – 648

653. Shortle, D. (1996) *FASEB J.* **10**, 27 – 34

654. Häckel, M., Hinz, H.-J., and Hedwig, G. R. (1999) *J. Mol. Biol.* **291**, 197 – 213

655. Makhatadze, G. I., and Privalov, P. L. (1996) *Protein Sci.* **5**, 507 – 510

655a. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) *Biochemistry* **41**, 6575 – 6582

655b. Dobson, C. M. (1999) *Trends Biochem. Sci.* **24**, 329 – 332

655c. Stevens, F. J., Pokkuluri, P. R., and Schiffer, M. (2000) *Biochemistry* **39**, 15291 – 15296

655d. Ellis, R. J., and Pinheiro, T. J. T. (2002) *Nature (London)* **416**, 483 – 484

655e. Oxenoid, K., Sönnichsen, F. D., and Sanders, C. R. (2001) *Biochemistry* **40**, 5111 – 5118

656. Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P. M., Huibregtse, J. M., and Pavletich, N. P. (1999) *Science* **286**, 1321 – 1326

657. Reyes, J. C. (2001) *Trends Biochem. Sci.* **26**, 18 – 20

658. Johnston, S. C., Riddle, S. M., Cohen, R. E., and Hill, C. P. (1999) *EMBO J.* **18**, 3877 – 3887

659. Whitby, F. G., Masters, E. I., Kramer, L., Knowlton, J. R., Yao, Y., Wang, C. C., and Hill, C. P. (2000) *Nature (London)* **408**, 115 – 120

660. Joazeiro, C. A. P., and Hunter, T. (2000) *Science* **289**, 2061 – 2062

661. Satoh, K., Sasajima, H., Nyoumura, K.-i, Yokosawa, H., and Sawada, H. (2001) *Biochemistry* **40**, 314 – 319

662. Byrd, C., Turner, G. C., and Varshavsky, A. (1998) *EMBO J.* **17**, 269 – 277

663. Davydov, I. V., and Varchavsky, A. (2000) *J. Biol. Chem.* **275**, 22931 – 22941

663a. Marx, J. (2002) *Science* **297**, 1792 – 1794

663b. Verma, R., Aravind, L., Oania, R., McDonald, W. H., Yates, J. R., III, Koonin, E. V., and Deshaies, R. J. (2002) *Science* **298**, 611 – 615

664. Kuttler, C., Nussbaum, A. K., Dick, T. P., Rammensee, H.-G., Schild, H., and Hadeler, K.-P. (2000) *J. Mol. Biol.* **298**, 417 – 429

665. Fineschi, B., and Miller, J. (1997) *Trends Biochem. Sci.* **22**, 377 – 382

666. Cascio, P., Hilton, C., Kisselev, A. F., Rock, K. L., and Goldberg, A. L. (2001) *EMBO J.* **20**, 2357 – 2366

667. Schild, H., and Rammensee, H.-G. (2000) *Nature (London)* **404**, 709 – 710

667a. Cyr, D. M., Höhfeld, J., and Patterson, C. (2002) *Trends Biochem. Sci.* **27**, 368 – 375

668. Dell’Angelica, E. C., Mullins, C., Caplan, S., and Bonifacino, J. S. (2000) *FASEB J.* **14**, 1265 – 1278

669. Jentsch, S., and Ulrich, H. D. (1998) *Nature (London)* **395**, 321 – 323

670. Wolf, Y. I., Grishin, N. V., and Koonin, E. V. (2000) *J. Mol. Biol.* **299**, 897 – 905

671. Simons, K. T., Strauss, C., and Baker, D. (2001) *J. Mol. Biol.* **306**, 1191 – 1199

672. Takada, S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11698 – 11700

673. Abbott, A. (1999) *Nature (London)* **402**, 715 – 720

References

674. Pradet-Balade, B., Boulmé, F., Beug, H., Müllner, E. W., and Garcia-Sanz, J. A. (2001) *Trends Biochem. Sci.* **26**, 225–229
675. Attwood, T. K., and Parry-Smith, D. J. (1999) *Introduction to Bioinformatics*, Longman, Green, New York
676. Spengler, S. J. (2000) *Science* **287**, 1221–1223
677. Attwood, T. K. (2000) *Science* **290**, 471–473
678. Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9390–9395
679. MacBeath, G., and Schreiber, S. L. (2000) *Science* **289**, 1760–1763
680. Horn, D. M., Zubarev, R. A., and McLafferty, F. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10313–10317
680a. Rajagopol, I., and Ahern, K. (2001) *Science* **294**, 2571 – 2573
681. Miyashita, M., Presley, J. M., Buchholz, B. A., Lam, K. S., Lee, Y. M., and Vogel, J. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4403–4408
682. Mann, M., and Pandey, A. (2001) *Trends Biochem. Sci.* **26**, 54–61
683. Baldauf, S. L., Roger, A. J., Wenk-Siefert, I., and Doolittle, W. F. (2000) *Science* **290**, 972–977
684. Gerstein, M. (1997) *J. Mol. Biol.* **274**, 562–576
685. Gómez, S. M., Nishio, J. N., Faull, K. F., and Whitelegge, J. P. (2002) *Molecular & Cellular Proteomics* **1**, 46 – 59
686. Michels, D. A., Hu, S., Schoenherr, R. M., Eggertson, M. J., and Dovichi, N. J. (2002) *Molecular & Cellular Proteomics* **1**, 69 – 74
687. Smolka, M., Zhou, H., and Aebersold, R. (2002) *Molecular & Cellular Proteomics* **1**, 19 – 29
688. Herbert, B. R., Harry, J. L., Packer, N. H., Gooley, A. A., Pedersen, S. K., and Williams, K. L. (2001) *Trends in Biotechnology* **19**, S3 – S9
689. Ge, Y., Lawhorn, B. G., ElNaggar, M., Strauss, E., Park, J.-H., Begley, T. P., and McLafferty, F. W. (2002) *J. Am. Chem. Soc.* **124**, 672 – 678
690. Griffin, T. J., and Aebersold, R. (2001) *J. Biol. Chem.* **276**, 45497 – 45500
691. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) *Science* **293**, 2101 – 2105
692. Kodadek, T. (2002) *Trends Biochem. Sci.* **27**, 295 – 300
693. Smith, G. P., Patel, S. U., Windass, J. D., Thornton, J. M., Winter, G., and Griffiths, A. D. (1998) *J. Mol. Biol.* **277**, 317 – 332
694. Blagoev, B., Pandey, A. (2001) *Trends Biochem. Sci.* **26**, 639 – 641
695. Yee, A., and 23 other authors. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1825 – 1830
696. Riek, R., Fiauz, J., Bertelsen, E. B., Horwich, A. L., and Wüthrich, K. (2002) *J. Am. Chem. Soc.* **124**, 12144 – 12153
697. Tugarinov, V., Muhandiram, R., Ayed, A., and Kay, L. E. (2002) *J. Am. Chem. Soc.* **124**, 10025 – 10035
698. Jhoti, H. (2001) *Trends in Biotechnology* **19**, S67 – S71
699. Thornton, J. (2001) *Trends Biochem. Sci.* **26**, 88–89
700. Koonin, E. V., Wolf, Y. I., and Karev, G. P. (2002) *Nature (London)* **420**, 218 – 223
701. Mallis, R. J., Brazin, K., Jez, J. M., Wilson, E. K., Dieckmann, G. R., Robic, S., and Harrahy, J. (2001) *Trends Biochem. Sci.* **26**, 642 – 643
702. Zhang, C., and Kim, S.-H. (2000) *J. Mol. Biol.* **299**, 1075 – 1089
703. Wimley, W. C. (2002) ***Protein Sci.*** 11, 301 – 312
704. Richardson, J. S., and Richardson, D. C. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2754 – 2759
705. Calladine, C. R., Sharff, A., and Luisi, B. (2001) *J. Mol. Biol.* **305**, 603 – 618
706. Chen, C. P., Kernytsky, A., and Rost, B. (2002) *Protein Sci.* **11**, 2774 – 2791
707. Chen, C. P., and Rost, B. (2002) *Protein Sci.* **11**, 2766 – 2773
708. Baker, D., and Sali, A. (2001) *Science* **294**, 93 – 96
709. Marsden, R. L., McGuffin, L. J., and Jones, D. T. (2002) *Protein Sci.* **11**, 2814 – 2824
710. Jiang, W. Baker, M. L., Ludtke, S. J., and Chiu, W. (2001) *J. Mol. Biol.* **308**, 1033 – 1044
711. Taylor, W. R. (2002) *Nature (London)* **416**, 657 – 660
712. Apic, G., Gough, J., and Teichmann, S. A. (2001) *J. Mol. Biol.* **310**, 311 – 325
713. Lee, K. K., Fitch, C. A., and Garcia-Moreno, E. B. (2002) *Protein Sci.* **11**, 1004 – 1016
714. Berezovsky, I. N., and Trifonov, E. N. (2002) *J. Mol. Biol.* **307**, 1419 – 1426
715. Trabi, M., and Craik, D. J. (2002) *Trends Biochem. Sci.* **27**, 132 – 138
716. Shao, Z., and Vollrath, F. (2002) *Nature (London)* **418**, 741
717. Asakura, T., Ashida, J., Yamane, T., Kameda, T., Nakazawa, Y., Ohgo, K., and Komatsu, K. (2001) *J. Mol. Biol.* **306**, 291 – 305
718. van Beek, J. D., Hess, S., Vollrath, F., and Meier, B. H. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10266 – 10271
719. Vollrath, F., and Knight, D. P. (2001) *Nature (London)* **410**, 541 – 548
720. Sprinzak, E., and Margalit, H. (2001) *J. Mol. Biol.* **311**, 681 – 692
721. von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., and Bork, P. (2002) *Nature (London)* **417**, 399 – 403
722. Gavin, A.-C., and 37 other authors. (2002) *Nature (London)* **415**, 141 – 147
723. Stagljar, I., and Fields, S. (2002) *Trends Biochem. Sci.* **27**, 559 – 563
724. Tong, A. H. Y., Drees, B., Nardelli, G., Bader, G. D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., Quondam, M., Zucconi, A., Hogue, C. W. V., Fields, S., Boone, C., and Cesareni, G. (2002) *Science* **295**, 321 – 324
725. Cornish-Bowden, A., and Cárdenas, M. L. (2001) *Trends Biochem. Sci.* **26**, 463 – 465
726. Grant, S. G. N., and Husi, H. (2001) *Trends in Biotechnology* **19**, S49 – S54
727. Natsume, T., Nakayama, H., and Isobe, T. (2001) *Trends in Biotechnology* **19**, S28 – S33
728. Ray, P., Pimenta, H., Paulmurugan, R., Berger, F., Phelps, M. E. Iyer, M., and Gambhir, S. S. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3105 – 3110
729. Florens, L., and 17 other authors. (2002) *Nature (London)* **419**, 520 – 526
730. Koller, A., Washburn, M. P., Lange, B. M., Andon, N. L., Deciu, C., Haynes, P. A., Hays, L. Schieltz, D., Ulaszek, R., Wei, J., Wolters, D., and Yates, J. R., III. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11969 – 11974
731. Remy, I., and Michnick, S. W. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7678 – 7683

Study Questions
for chapters 28 and 29

1. Describe the role of sigma factors (σ) in transcription by prokaryotic RNA polymerases. What is the effect of the release of σ from the holoenzyme once transcription has been initiated. How would a mutation that prevents a σ factor from dissociating from core RNA polymerase affect the rate of transcription?

2. Explain how histidine biosynthesis is controlled in *E. coli*, a bacterium that has no *his* repressor.

3. One mechanism of transcriptional control in prokaryotes, especially of several operons controlling the biosynthesis of amino acids, is **attenuation**. Briefly describe the mechanism of attenuation. How does the supply of amino acid in the cell affect the process?

4. Is transcription attenuation likely to be an important mechanism of transcriptional regulation in eukaryotic cells?

5. Discuss two main DNA-recognition motifs found in eukaryotic transcription factors. Describe their structures, indicate how they bind to DNA, and discuss how each specifically recognizes its DNA binding site.

Study Questions

6. How can a DNA enhancer sequence located as many as several thousand base pairs from a gene transcription start site influence transcription even if its orientation is reversed?

7. Some eukaryotic DNA viruses code for two or more mRNA transcripts of differing lengths from the same region on the DNA. Suggest an explanation. How do you expect the two translation products of these mRNAs to differ?

8. High salt concentrations weaken the interaction of histones with DNA but have little effect on the binding of many regulatory proteins. Explain this observation in terms of the modes of interaction of the two types of protein.

9. Discuss the changes that must be made in a typical eukaryotic structural gene to allow its protein product to be synthesized in bacteria.

10. List the different types of covalent modification that may be made to tRNA. To ribosomal RNA. To messenger RNA.

11. List various small RNAs and their functions within cells.

12. Some amino acids utilize only one codon of the 64 in the genetic code. Other amino acids use as many as six codons (Tables 5-5, 5-6). What advantages to a cell is provided by utilization of several codons for a single amino acid?

13. In what ways is the genetic code not quite "universal?" What is meant by "editing" of mRNA?

14. Why is it necessary to have "adapters" in the form of tRNAs to read the genetic code during translation?

15. Most nonsense suppressor genes are mutants of tRNA genes. In view of this fact, how can cells survive the presence of such mutations?

16. Explain how the protein synthesizing machinery is able to differentiate the initiation AUG codon from an internal AUG (methionine) codon in prokaryotes. How is this accomplished in eukaryotes?

17. The amino acid sequence of a mature protein sometimes differs from that deduced from the DNA nucleotide sequence of the structural gene for that protein. Discuss three ways by which this may occur.

18. Write out in detail, using structural formulas, the chemical mechanism of synthesis of an aminoacyl-tRNA and of incorporation of the aminoacyl group into a peptide chain being formed by a ribosome.

19. a) Calculate the minimum number of ATP equivalents consumed in the biosynthesis of a 300-amino acid *E. coli* protein, having the N-terminal sequence Ala-Ser-Val-Tyr, from the free amino acids.

    b) Much of this energy involves hydrolysis of GTP. What is the role of this GTP hydrolysis in protein synthesis?

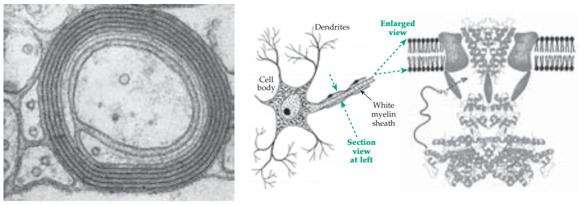
20. How do the polypeptide products produced in the presence of (a) puromycin and (b) streptomycin differ from polypeptides synthesized in the absence of these inhibitors? Explain your answer.

21. What is the significance to protein synthesis of each of the following?
    Shine–Dalgarno sequence
    Signal recognition particle
    proteasome

22. How can useful antibiotics that act on ribosomes kill bacteria but not people?

23. Compare termination of translation in bacteria and in eukaryotes.

24. List some types of error that are likely to be made during protein synthesis. What mechanisms have cells developed to deal with these?



Center: Diagram of the cell body of a neuron with dendrites and a short section of its long myelinated axon (see Fig. 30-8). Left: Electron micrograph of a thin section through an axon showing the myelin sheath formed by the wrapping of the plasma membrane of a neuroglial cell around the axon (see p. 390 and Fig. 30-9). Right: Model of a voltage-regulated K^+ channel in the cell membrane of an axon. The pore, which is formed from four α-subunits, is represented by that of the bacterial pore shown in Fig. 8-21. Also shown are an inner cytoplasmic activation gate consisting of four β-subunits, which are proposed to form ball-and-chain devices that can close the pores in response to voltage changes. From Zhou *et al.* See Fig. 30-18.

Contents

1741 .... **A. The Hormones**
1741 .... 1. Receptors, Feedback Loops, and Cascades
1741 .... *Receptor types*
1742 .... *Feedback loops*
1743 .... *Signaling cascades*
1743 .... *Second messengers as hormones*
1743 .... *The vertebrate hormones*
1743 .... 2. Hormones of the Pituitary Gland (Hypophysis) and Hypothalamus
1745 .... *Pituitary growth hormone and related hormones*
1745 .... *The pituitary glycoprotein hormones*
1745 .... *Hypothalamic releasing hormones*
1747 .... *Vasopressin and oxytocin*
1748 .... *Melanocortins*
1748 .... 3. Pancreatic and Related Hormones
1749 .... 4. Gastrointestinal and Brain Peptides
1750 .... 5. Other Mammalian Peptide Hormones
1752 .... *Endogenous opioid peptides*
1752 .... *ATP, ADP, and adenosine*
1752 .... *Kinins*
1752 .... *Endothelins*
1753 .... *Peptides as attractants*
1753 .... 6. Protein Growth Factors and Cytokines
1756 .... 7. Nonpeptide Mammalian Hormones
1756 .... *The volatile hormones nitric oxide (NO) and carbon monoxide (CO)*
1757 .... *Hormonal lipids*
1758 .... 8. Nonvertebrate Hormones and Pheromones
1758 .... *Neurohormones of invertebrate animals*
1760 .... *Hormones of insects and crustaceans*
1760 .... 9. Plant Hormones
1762 .... 10. Secretion of Hormones
1762 .... **B. Neurochemistry**
1762 .... 1. The Anatomy and Functions of Neurons
1765 .... 2. Organization of the Brain
1766 .... 3. Neuronal Pathways and Systems
1767 .... 4. The Propagation of Nerve Impulses
1769 .... 5. Ion Conducting Channels
1769 .... *The sodium ion channel of the electric eel*
1772 .... *Calcium ion channels*
1772 .... *Potassium ion channels*
1774 .... *Chloride channels and the ionic environment of neurons*
1774 .... *Receptor-associated ion channels*
1775 .... 6. A Plethora of Neurotoxins
1776 .... 7. Neuronal Metabolism
1777 .... 8. Synapses and Gap Junctions
1777 .... *A synaptic vesicle cycle*
1781 .... *Astrocyctes and other glia*
1781 .... *Gap junctions in synapses*
1782 .... 9. Neurotransmitters
1784 .... *Cholinergic receptors and their agonists and antagonists*
1785 .... *Amino acids as neurotransmitters*
1787 .... *Excitotoxicity*
1787 .... *The inhibitory neurotransmitter gamma-aminobutyrate (GABA)*
1789 .... *Glycine*
1789 .... *Anesthetics*
1789 .... *Adrenergic synapses: the catecholamines*
1792 .... *Parkinson disease*
1792 .... *Serotonin and melatonin*
1793 .... *Other neurotransmitters*
1795 .... *Lipid mediators in the brain*
1795 .... *Nitric oxide and carbon monoxide*
1796 .... 10. Some Addictive, Psychotropic, and Toxic Drugs
1797 .... *Opioid receptors*
1797 .... *Ethanol*
1798 .... *Psychotropic or mind-changing drugs*
1798 .... *Neurotoxins produced by the body*
1798 .... 11. The Senses: Sight, Smell, Taste, Hearing, Touch, and Others
1798 .... *Odor*
1799 .... *Taste*
1800 .... *Pain*
1801 .... *Mechanoreceptors*
1801 .... *Hearing*
1801 .... 12. The Chemistry of Learning, Memory, and Thinking
1801 .... *Memory systems*
1802 .... *Brain rhythms*
1803 .... *Chemical changes in synapses*
1804 .... *The complexity of the brain*
1804 .... *Intelligence*
1805 .... *Behavior*
1805 .... 13. Circadian Cycles and Sleep
1808 .... 14. Mental Illness
1808 .... *Depression*
1810 .... *Schizophrenia*
1811 .... *Epilepsy*
1811 .... *Neurodegenerative diseases*

1815 .... **References**
1829 .... **Study Questions**

Boxes

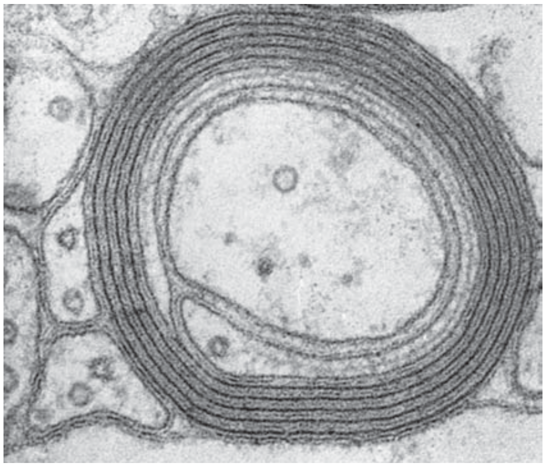
1806 .... Box 30-A Positron Emission Tomography (PET), Functional Magnetic Resonance (FMRI), and Other Imaging Techniques

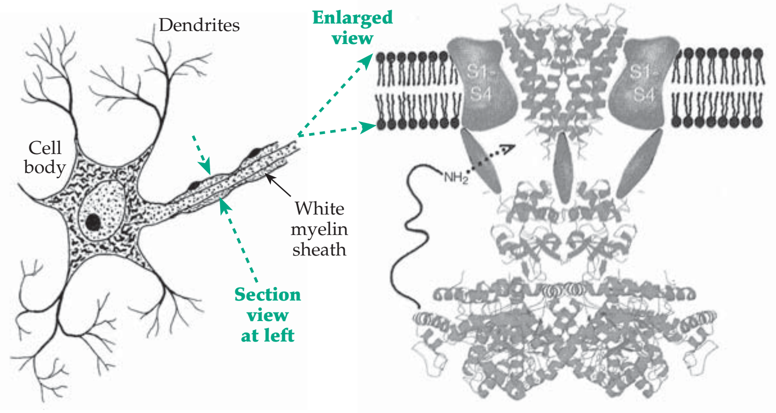
Tables

1744 .... Table 30-1 Peptide and Protein Hormones of Vertebrates
1746 .... Table 30-2 Nonpeptide Vertebrate Hormones
1747 .... Table 30-3 Releasing and Inhibiting Hormones from the Hypothalamus
1751 .... Table 30-4 Some Pancreatic and Gastrointestinal Hormones and Neurohormones and Their Sequences
1759 .... Table 30-5 Some Microbial and Invertebrate Peptide Hormones
1780 .... Table 30-6 Some Proteins Important to the Formation and Functioning of Synaptic Vesicles

Chemical Communication Between Cells

30





The regulation of growth and metabolism of complex multicellular organisms depends heavily upon chemical messages sent between cells. This includes secretion of hormones into the circulatory system,[1–3] chemical transfer of information through communicating cell junctions, and passage of signals between neurons in the brain. This chapter deals with these matters and also with communication between organisms, i.e., with the biochemistry of ecological relationships. Embryonic development and differentiation of tissues also require communication between cells as does the functioning of the immune system. These topics are considered in Chapters 31 and 32.

A. The Hormones

The term hormone has traditionally been applied to substances synthesized in and secreted by one tissue and which act to influence distant target organs or tissues. However, many peptide hormones also act as neurotransmitters, passing across very short gaps between cells. In addition, many chemical messengers, including the peptide growth factors, act more locally. Looking at lower invertebrates as simple as *Hydra*, we find peptides resembling our own hormones and neurotransmitters. These are secreted by neuroendocrine cells of *Hydra* and diffuse throughout the body. In higher animals hormones regulate the concentrations of nutrients such as glucose and of ions such as Ca^{2+} and phosphate in the blood. They control the volume and osmotic pressure of body fluids, as well as digestion, growth, reproduction, and responses to stress.

1. Receptors, Feedback Loops, and Cascades

Every hormone must have one or more receptors, most of which are proteins. These may be found embedded in the outer surface of the plasma membrane, in the cytoplasm, or in the cell nucleus. Binding of a hormone to its receptor often elicits both a rapid response and a slower one. For example, we have seen that glucagon, adrenaline, and vasopressin bind to cell surface receptors and promote the synthesis of cyclic AMP (Fig. 11-4). The cAMP induces rapid chemical modifications of many proteins. Some of these may diffuse into the nucleus and affect transcription of genes, a slower response. Insulin (Chapter 11, Section G) also exerts both rapid and slower responses.

Receptor types. Many different kinds of protein can serve as hormone receptors. Some of these are discussed in Chapter 11. The most abundant are the G protein-coupled 7-helix receptors[4–5c] such as that of a β adrenergic receptor pictured in Fig. 11-6. Glucagon, adrenaline, ACTH, and gastrin are a few of the hormones that bind to receptors of this type. Similar receptors respond to light (rhodopsin; Chapter 23) and over 1000 different 7-helix receptors respond to smell and taste. The G proteins and their controlling cycles, Eq. 11-10,[5–7] have also been considered in Chapter 11. The reality of the dissociation and reassociation of the α and βγ subunits in response to binding of a hormone has been demonstrated in living cells by the use of fluorescence resonance energy transfer (FRET).[8] Not all receptors activate G proteins. One large group of membrane-associated receptors have single transmembrane helices but require dimerization to be effective.

The bacterial chemoreceptor (Figs. 11-8 and 19-5) has a very small ligand-binding domain and a larger internal domain that activates a histidine kinase. Many growth-factor receptors, including the insulin receptor (Figs. 11-11, 11-12), have internal domains with protein tyrosine kinase activity.

In contrast, steroid hormones, thyroxine, and retinoids bind to internal receptors. In 1968, Gorski *et al*.[9] and Jensen *et al.*[10,11] proposed independently that steroid hormone receptors in the cytoplasm bind incoming steroid molecules and after an "activation" step carry the hormone into the nucleus, where the hormone–receptor complex would bind at many sites in the chromatin inducing transcription of selected genes.[12] Doubt has been cast on the assumption that the steroid hormone receptors must bind hormone initially in the cytoplasm. However, the role of steroid receptors in regulating transcription is well established (see discussion in Chapter 22, Section E,5; Chapter 28, Section C,6).

Feedback loops. Maintenance of a steady state within an organism depends upon numerous negative-feedback loops. Hormones assist in adjusting reaction rates to maintain a steady state when conditions are changed. For example, blood glucose rises after a meal. This increase is sensed in the pancreatic beta cells (pp. 998, 999), which release insulin. The released insulin promotes uptake of glucose by cells and its conversion into glycogen and lipid stores. When the glucose level falls, inhibitory mechanisms that decrease insulin release are allowed to operate.

Similar regulatory loops can be traced for nearly all hormones. Sometimes they involve several stages and involve sensing devices in the central nervous system. In such cases neural impulses stimulate the **hypothalamus** of the brain (Fig. 30-1) to release **neurohormones**, which travel to the anterior lobe of the pituitary gland. The pituitary, in turn, releases hormones such as **corticotropin** (adrenocorticotropic hormone, **ACTH**), which stimulate the adrenal cortex to release its hormones. The latter exerts feedback inhibition upon the hypothalamus to decrease the secretion of ACTH by the pituitary. Steroids also participate in feedback loops to the hypothalamus.[13] Using ^{3}H-labeled hormones or fluorescent analogs, it has been possible to locate specific brain cells sensitive to a given hormone by autoradiography.[14]

A characteristic of hormonal effects is that they are seldom unique, and are often balanced by counteracting effects of other hormones. For example, both glucagon and adrenaline promote the release of glucose from liver glycogen into the bloodstream. The glucocorticoids stimulate the rate of production of glucose from other body constituents (Chapter 11). Growth hormone tends to increase glucose levels by inhibiting utilization of sugar by tissues. On the other hand, insulin acts to promote uptake of glucose by tissues and a more efficient utilization. The thyroid

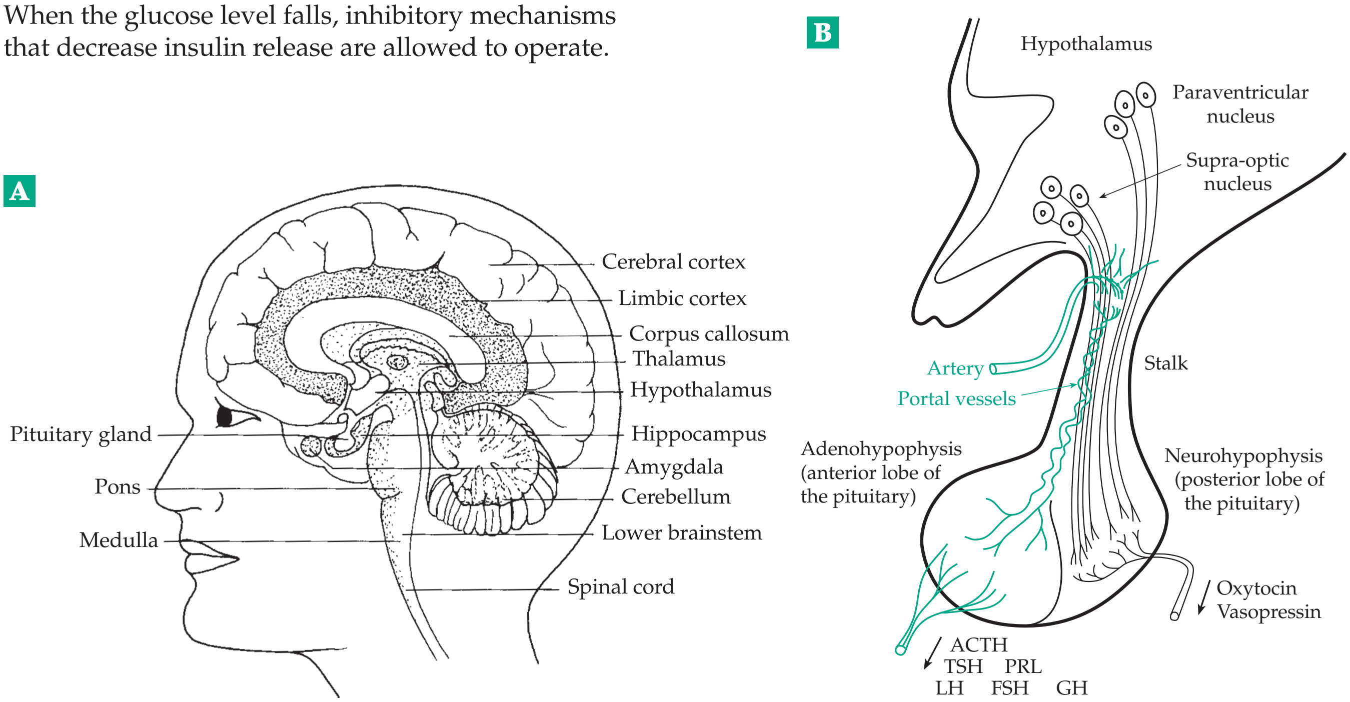

A
Cerebral cortex
Limbic cortex
Corpus callosum
Thalamus
Hypothalamus
Pituitary gland
Hippocampus
Amygdala
Pons
Cerebellum
Medulla
Lower brainstem
Spinal cord
B
Hypothalamus
Paraventricular nucleus
Supra-optic nucleus
Artery
Portal vessels
Stalk
Adenohypophysis (anterior lobe of the pituitary)
Neurohypophysis (posterior lobe of the pituitary)
Oxytocin
Vasopressin
ACTH
TSH
PRL
LH
FSH
GH

Figure 30-1 (A) Median sagittal section of the human brain. From Maya Pines.[15] (B) Drawing illustrating the synthesis of peptide hormones in the hypothalamus and transport via portal blood vessels into the anterior lobe of the pituitary gland or via nerve tracts into the posterior lobe.[16]

hormone increases the overall rate of metabolism of cells and also tends to promote a decrease in blood glucose.

Signaling cascades. As we have seen in Chapter 11, hormones frequently elicit the synthesis of second messengers such as cAMP, inositol phosphates, or diacyglycerol. This not only provides amplification of the initial hormonal signal but also allows a single hormone to control a "domain" of many metabolic processes. Each of these processes, in turn, can influence others. Many processes are affected by several different hormones and by more than one second messenger. Since we are far from knowing how many hormones exist and how many second messengers are released, the network of regulatory interactions within cells may be one of overwhelming complexity. An abbreviated version of a mitogen-activated kinase (MAP kinase) cascade[4,17] is shown in Fig. 11-13. These cascades are not only initiated but also are propagated by a series of phosphorylation reactions catalyzed by more than 1000 protein kinases encoded by the human genome.[18] Together with more than 500 protein phosphatases, which are often joined together as a bifunctional protein (p. 545),[19,20] they form a complex branching network of interactions.[21,22] These help to control responses not only to hormones but also to varying metabolite concentrations and physical stimuli.

Second messengers as hormones. The same compounds that serve as intracellular second messengers sometimes act as hormones. Tomkins suggested[23] that cAMP and some other small molecules serve as "symbols" indicating a metabolic need. For example, in bacteria ppGpp (p. 1715) serves as a symbol of nitrogen or amino acid deficiency. In cells ranging from those of bacteria to animals, cAMP is a symbol for carbon-source starvation. In *E. coli* cAMP levels increase during carbon-source starvation and stimulate the initiation of transcription of many bacterial operons (Chapter 28). In *Dictyostelium discoideum* (Box 11-C) cAMP is released by cells, when substrate depletion occurs. In this instance, the cyclic nucleotide acts as a hormone transmitting a signal to other cells.

Whereas cAMP is sometimes used by lower organisms as a hormone, its metabolic lability makes it unsuitable for higher animals. Thus, in our bodies the hormones glucagon and adrenaline carry a message to cell surfaces, where binding to receptors stimulates cAMP production. This, in turn, leads to mobilization of metabolic stores such as those of glycogen and triglycerides, just as if these cells had also been subjected to acute starvation. Tomkins suggested that hormones are produced by "sensor" cells in direct contact with environmental signals and travel to and activate more sequestered "responder" cells. The picture can be generalized further by realizing that neurotransmitters are largely derivatives of amino acids. These amino acids may have originally served as intracellular symbols reflecting changes in environmental amino acid concentration but were later utilized in short-range intercellular communication within the nervous system.

The vertebrate hormones. The principal established vertebrate hormones are listed in Tables 30-1 and 30-2. Also given are references to other parts of the text, where specific hormones are discussed. The hormones can be divided into four groups on the basis of chemical structure: (1) peptides and proteins, (2) derivatives of the aromatic amino acids, (3) steroids and prostaglandins, and (4) volatile compounds such as NO and CO. The most numerous are the peptide hormones, many of which also act as neurotransmitters. Peptide hormones, e.g., those with insulin-like effects, function in all phyla of the metazoa, and hormone-like molecules are found in bacteria.[24]

2. Hormones of the Pituitary Gland (Hypophysis) and Hypothalamus

Connected to the brain by a stalk (Fig. 30-1), the pituitary gland releases at least ten peptide or protein hormones that regulate the activity of other **endocrine** (hormone-producing) glands in distant parts of the body. The pituitary is composed of several distinct parts: the anterior lobe (**adenohypophysis**), a thin intermediate portion (**pars intermedia**), and a posterior lobe (**neurohypophysis**). Each has its own characteristic endocrine functions.

The anterior lobe of the pituitary secretes a series of ten or more peptide hormones ranging in size from the ~20-residue β-melanotropin to the ~200-residue growth hormone (somatotropin). Several of these contain a common heptapeptide unit, which is marked in green in the following structure:

Ac • Ser • Tyr • Ser • Met • Glu • His • Phe • Arg • Trp • Gly • Lys • Pro • ValNH_2

Structure of α-melanotropin from pig, beef, and horse

Not only this heptapeptide but also the entire amino acid sequence of **α-melanotropin** is found within the sequence of **corticotropin** (Fig. 30-2), which has an additional 29 amino acids at the C-terminal end.[25] The same heptapeptide was also found in the **lipotropins**. The explanation is, in part, that several of these hormones arise from a single 31-kDa precursor protein called **prepro-opiomelanocortin**.[25,26] It contains an N-terminal signal sequence that is removed shortly after synthesis, as well as pairs of adjacent basic residues (Arg-Arg, Arg-Lys, Lys-Arg and Lys-Lys) at a number of places (Fig. 30-2). After removal of the signal sequence, further cleavage is thought to occur within the secretory vesicles by proteases, which cut either on the carboxyl side of these basic pairs or between them.[25,27,28] The same precursor is made in both anterior and intermediate lobes and is rapidly cut to ACTH, β-lipotropin, and an N-terminal part. In the intermediate lobe the ACTH is then cleaved at the Lys-Lys and Arg-Arg pairs to form αMSH and another peptide called corticotropinlike intermediate lobe peptide (CLIP). Beta lipotropin is degraded rapidly in the intermediate lobe and more slowly in the anterior lobe to γ-lipotropin and the opioid peptide β-endorphin

TABLE 30-1
Peptide and Protein Hormones of Vertebrates

Source and name of hormone	No. residues	Principal site of action	References Chapter, Section
A. Pituitary gland (hypophysis)			
1. Adenohypophysis (anterior portion)			
Corticotropin (ACTH)[a]	39	Adrenal cortex, adipose tissue	Fig. 2-4
β-Melanotropin (β melanocyte-stimulating hormone, β-MSH)[a]	18–22	Skin	Fig. 30-2
β-Lipotropin (β-LPH)[a]	91	Precursor of β-MSH and β-endorphin	Fig. 30-2
γ-Lipotropin (γ-LPH)[a]	58	Precursor of β-MSH	Fig. 30-2
β Endorphin[a]	31	Brain	
Somatotropin (growth hormone, GH)	~200	All tissues	
Prolactin (mammotropin)	~200		
Thyrotropin (thyroid-stimulating hormone, TSH)[b]		Thyroid	Ch 25, B2
Follitropin (follicle-stimulating hormone, FH)[b]		Ovaries, testes	
Lutropin (luteinizing hormone, ICSH or LH)[b]		Ovaries, testes	
2. Pars intermedia (intermediate portion)			
α-Melanotropin (α-melanocyte-stimulating hormone, αMSH)[a]	13	Skin	Fig. 30-2, pp. 1742, 1748
3. Neurohypophysis (posterior portion)			
Oxytocin (ocytocin)	9	Uterus, mammary glands	Fig. 2-4
Vasopressin (antidiuretic hormone)	9		Fig. 2-4
B. Pancreas			
Insulin	51	All cells	Fig. 7-17, Ch 11, G
Glucagon	29	Liver, adipose tissue	Ch 11, D
C. Ovary (corpus luteum)			
Relaxin	—	Pelvic ligaments	p. 1746
D. Thyroid			Ch 25, B2
Calcitonin (thyrocalcitonin)	32	Bones, kidney	Box 6-D
E. Parathyroid			
Parathyrin (parathyroid hormone)	84	Bones, kidney	Box 6-D
F. Kidney			
Erythropoietin		Bone marrow	
Renin (an enzyme)		Adrenal cortex	p. 621; Box 22-D

[a] Arise by cleavage of pro-opiomelanocortin.

[b] Related two-subunit (αβ) proteins with a common β subunit for these three hormones, for human chorionic gonadotropin (hCGH), and for mitogen-regulated protein (proliferin).

(Section B,10). Precursor proteins have been identified for many other peptide hormones, even those with very short chains.[29–30a] Proteolytic cleavages and other processing reactions occur within the secretory pathways of organisms from yeast to humans.[30b,c]

Many pituitary hormones have a pyroglutamate (5-oxoproline) residue at the N terminus (e.g., see Fig. 2-4). This presumably arises by attack of the terminal $-NH_2$ group on the amide carbon of an N-terminal glutamine side chain with displacement of NH_3 (Eq. 10-10).[29] The C terminus is often an amide of the carboxyl group with ammonia, which usually arises from a peptide chain containing one additional glycine residue at the C terminus (Eq. 10-11). The processing of peptide hormones doesn't end with their synthesis. They are usually degraded quickly or are converted into derivatives with weaker hormonal activity.

Pituitary growth hormone and related hormones. The pituitary growth hormone (**somatotropin**)[31] and **prolactin** are 22- to 23-kDa proteins, which share homology also with human **placental lactogen**, a lactogenic hormone secreted by the placenta,[32] and with a growth factor called **mitogen-regulated protein** (or proliferin).[33–35] The polypeptide chain of the 191-residue porcine[36] and human[37] growth hormone folds into an antiparallel four helix bundle (similar to that in Fig. 2-22) but with two long irregular connecting strands. The high degree of homology among somatotropins of many other species indicates a near identity of three-dimensional structures. However, biological function is species-specific. Humans and monkeys respond only to growth hormone from primates. The interaction of an aspartate side chain at position 171 with arginine 43 of the receptor protein may account for some of this specificity.[38] The receptor is a member of a large superfamily of receptor proteins with single transmembrane helices and extracellular domains similar to that of tissue factor (Fig. 12-18) and also, in some respects, to immunoglobulin domains.[39,40] Receptors for growth hormones bind in specific ways to two molecules of receptor protein.[41,42]

Human growth hormone produced in bacteria is used to help very short children to grow. Bovine growth hormone produced in bacteria is used to increase milk production from cows.[43] However, this use may be damaging to the cows.[44] Some humans produce too much growth hormone, often as a result of tumors in the pituitary. The resulting condition of **acromegaly**[45] causes excessive bone growth and many other problems. Growth hormone has a broad range of other effects, e.g., mimicking the action of insulin.[46]

Lactogenic hormones also have 4-helix bundle structures, and the prolactin receptor structure resembles that of growth hormone as well as those of a large cytokine family.[47,48] Prolactin affects placental development during pregnancy. However, during the latter half of pregnancy the placenta has a dominant endocrine effect, synthesizing both progesterone and the placental lactogen.[32,49]

The pituitary glycoprotein hormones. The thyroid-stimulating hormone **thyrotropin** (TSH), together with **folitropin** (FH) and **lutropin** (LH; Table 30-1), form a family of related ~28-kDa dimeric glycoproteins in which each subunit has a three-loop structure stabilized by a characteristic "cystine knot."[50] Also included in the family is the placental **chorionic gonadotropin**,[51,52] which is found only in human beings and a few other species. LH has a central role in promoting both spermatogenesis and ovulation by stimulating synthesis of steroid hormones in the testes and ovary, respectively.[53] Human chorionic gonadotropin (hCG) is also essential for maintenance of pregnancy and acts by stimulating the ovaries to secrete required steroid hormones.

All of these glycoprotein hormones are αβ dimers, and within a single species the subunits of TSH, FH, LH, and CG are identical. However, the β subunits are all different.[54,55] In the human there are at least six genes or pseudogenes for the hCG β chain in a cluster that also contains a single LH β chain gene.[54] The hormones undergo glycosylation and sulfation in the Golgi before secretion.[56] The hormones bind to 7-helix receptors, which are coupled to formation of cAMP or inositol trisphosphate.[57] Mutations in LH may cause male infertility,[58] while mutations in the corresponding receptor may cause male precocious puberty.[59]

Hypothalamic releasing hormones. As was mentioned previously, the anterior lobe of the pituitary

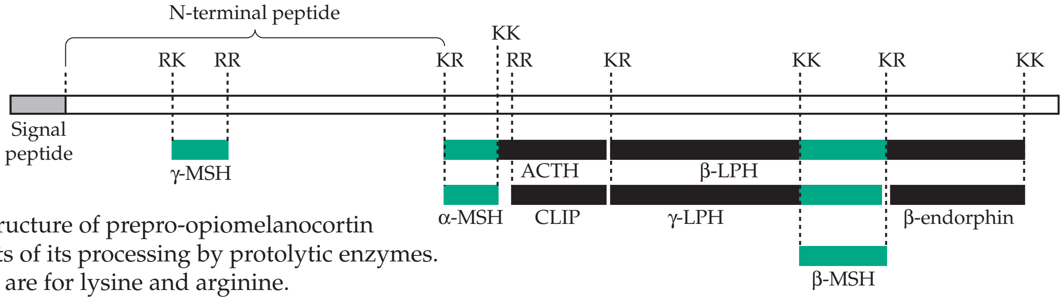


Figure 30-2 Schematic structure of prepro-opiomelanocortin and of some of the products of its processing by protolytic enzymes. The abbreviations K and R are for lysine and arginine.

releases its hormones in response to at least nine neurohormones known as **releasing hormones** or releasing factors.[16,60,61] They are secreted in minute quantities by the hypothalamus into a special portal vein that carries them directly to the pituitary where they exert their effects (Fig. 30-1B). As is indicated in Fig. 2-4 and in Table 30-3, several releasing factors are small peptides, but others are quite large. **Thyrotropin-releasing hormone** (THR, thyroliberin)[62] is a tripeptide; but human **growth hormone-releasing hormone** (somatoliberin) is a 44-residue peptide. Both are synthesized as larger proteins, which are cleaved and processed to form the mature C-amidated hormones.[63,64] **Corticotropin-releasing hormone** (CRH; CRF; corticoliberin), the 41-residue ACTH-releasing factor, is also cut from the much larger prepro-CRF.[65,66] Release of both LH and FSH is stimulated by a single **gonadotropin-releasing hormone** (GnRH).[5a,10,67,68] The releasing factors bind to 7-helix G-protein coupled receptor.[68] Both the releasing factor and gonadotropin are released into the appropriate parts of the bloodstream in a pulsatile fashion emphasizing the neural origin of their release.[69,70]

The hypothalamus also synthesizes **release-inhibiting factors**.[67] One of these, **somatostatin** (Table 30-3), inhibits release of somatotropin, thus counteracting the effect of the growth hormone releasing hormone. Somatostatin acts both in the pituitary and also in the pancreas, where it inhibits the release of both insulin and glucagon.[71,72] The result is a lowering of blood glucose. This suggested a new approach to the treatment of diabetes. However, because of the many other effects of somatostatin[73] and its rapid degradation it has not been useful clinically. Nevertheless, hundreds of analogs of somatostatin have been synthesized, some of which may be of practical

TABLE 30-2
Nonpeptide Vertebrate Hormones

Type, source, and name of hormone	Principal site of action	References Chapter, Section
A. Amino acid derivatives		
1. Thyroid		Ch 25,B,2
Thyroxine and triiodothyronine	Most cells	
2. Adrenal medulla		Ch 11
Adrenaline, noradrenaline (epinephrine, norepinephrine)	Most cells	
3. Pineal gland		
Melatonin	Melanophores	Fig. 25-12
4. Nerves and other cells		
Serotonin (5-hydroxytryptamine)	Arterioles, central nervous system	Fig. 25-12
B. Steroids and prostaglandins		Ch 22, F
1. Testes		
Testosterone	Most cells	Fig. 22-11
2. Ovaries		
Estrogen (estradiol-17β)	Most cells	Fig. 22-11
3. Corpus luteum		
Progesterone	Uterus, mammary glands	Ch 12; Ch 28
4. Adrenal cortex		Fig. 22-11
Corticosterone, cortisol	Most cells	Fig. 22-11
Aldosterone	Kidney	Fig. 22-11
5. Various tissues		
Prostaglandins	Smooth muscle	Ch 21, D
C. Volatile hormones		Fig. 21-8
1. Nitric oxide, NO	Endothelium, brain	
2. Carbon monoxide, CO	Brain	

value.[67] The biological activity of somatostatin resides largely in the sequence FYKT at positions 6–10, a sequence that is thought to form a beta turn (Fig. 30-3). Much of the rest of the molecule can be left off and the disulfide bridge moved up as far as positions 6–11 with retention of high potency. Human somatostatin is synthesized initially as a 116-residue precursor.[74]

A 56-residue peptide, which is formed from the 10-kDa precursor to GnRH, inhibits secretion of prolactin.[75] Inhibition of FSH release is accomplished by feedback inhibition. Hormones known as **inhibins** are produced in the gonads and act to inhibit release of FSH from the pituitary.[76]

Vasopressin and oxytocin. In contrast to the large peptide hormones made in the anterior lobe of the pituitary are **vasopressin** and **oxytocin**, which are secreted from the neurohypophysis, the posterior lobe.[60] The neurohypophysis consists of neural tissue, whose secretions are directly controlled by the central nervous system. In fact, the cell bodies of the secretory neurons are located in specific nuclei of the hypothalamus (Fig. 30-1B). About 4000 vasopressin-secretory neurons and a similar number of oxytocin neurons are present in the neurohypophysis of the rat.[16] Vasopressin is a major regulator of blood volume and pressure,[77] and its secretion is influenced by stress. It increases the water permeability of the kidney collecting duct cells by inducing translocation of aquaporin proteins from intracellular storage vesicles into the apical plasma membrane.[78] Vasopressin binds to

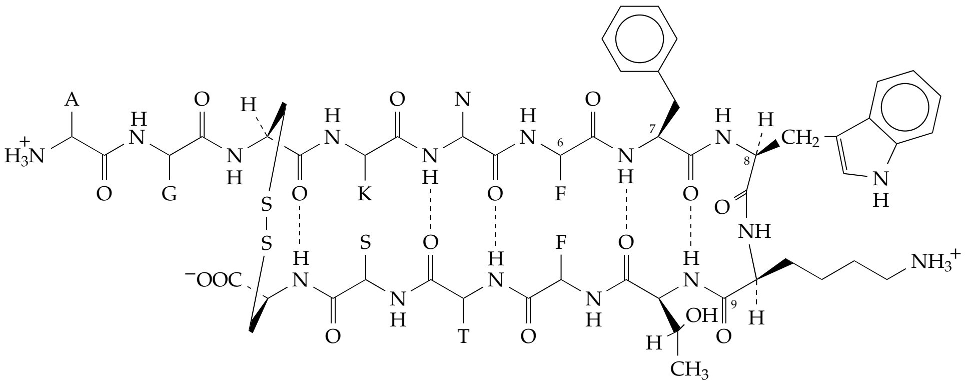


Figure 30-3 Possible secondary structure of somatostatin with a beta turn at residues 7–10 and a disulfide bond between positions 3 and 14. The true conformation would have the plane of the beta sheet puckered and twisted.

TABLE 30-3
Releasing and Inhibiting Hormones from the Hypothalamus

Name	Number of amino acid residues	Sequence[a]
Thyrotropin-releasing hormone (thyroliberin, TRH)	3	*p*EHP–NH_2
Gonadotropin-releasing hormone (GnRH, LH- and FSH-releasing hormone)	10	*p*EHWSYGLRPG–NH_2
GH-releasing hormone (somatoliberin)	44	
Corticotropin-releasing hormone (corticoliberin, CRH)	41	
MSH-releasing factor (melanoliberin)[b]	5	CYIQNC └S—S┘
Somatostatin (GH-release-inhibiting hormone)	14	AGCKNFFWKTFTSC └——S—S——┘
MSH-release-inhibiting factor[b]	3	PLG
Prolactin-releasing factor		
Dopamine (prolactin-release-inhibiting factor)		

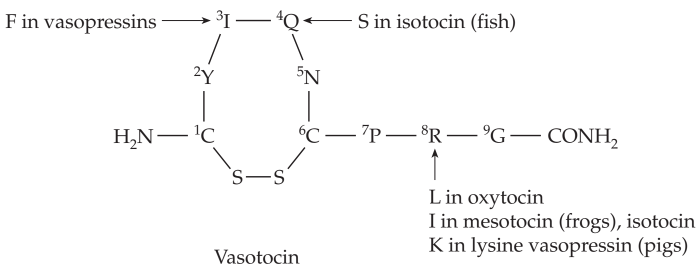
[a] Standard one-letter abbreviations are used. *p*E is pyroglutamyl (5-oxoprolyl) and –NH_2 at the right indicates a C-terminal *carboxamide*.
[b] Ring and tail fragments of oxytocin.

G-protein-coupled receptors.[79] A defect in the type 2 vasopressin receptor leads to the condition of nephrogenic **diabetes insipidus** in which the body fails to concentrate the urine.[77,80] Oxytocin acts on smooth muscles of the uterus during childbirth and triggers the release of milk from the mammary glands.[81] The latter response is partially controlled by the suckling of the infant, which induces the nervous system to release oxytocin into the bloodstream.

Hormones related to oxytocin and vasopressin occur in most vertebrates, the compound **vasotocin** shown in Fig. 30-4 being the most common. Substitution of phenylalanine for isoleucine at position 3 gives arginine vasopressin, the vasopressin found in our bodies. Structure of oxytocin and related hormones[82] are also shown in Fig. 30-4. Like somatostatin, vasopressin and oxytocin may also form antiparallel pleated sheet structures with β turns. The structural requirements for hormone activity have been studied intensively. Both the macrocyclic hexapeptide ring and the tripeptide side chains are necessary for maximal activity.[83]

The gene for arginine vasopressin is that of a 166-residue precursor protein carrying a 19-residue signal sequence at the N terminus.[84] This sequence is followed by that of vasopressin, then after a GKR linker by the 95-residue **neurophysin II**. Finally, after one additional arginine there is a 39-residue glycopeptide. Oxytocin originates in a parallel way from its own precursor.[85]

The 93- to 95-residue neurophysins act as carriers for vasopressin and ocytocin, forming specific complexes with them. Neurophysins contain 14 cysteine residues, which form seven disulfide bonds. There is a striking similarity in sequence between the neurophysins, snake venom toxins, a wheat germ lectin (agglutinin), a ragweed pollen allergen, and a small plant protein called hevein. On the basis of the alignment of cysteine residues, Drenth proposed[86] that all of these proteins have a disulfide-linked core whose structure is shown in Fig. 30-16.



Vasotocin
Hormone from nonmammalian vertebrates

Figure 30-4 Structure of the nonmammalian hormone vasotocin and of related hormones including ocytocin and vasopressin.

Melanocortins. The melanocortin peptides, which are derived from pro-opiomelanocortin as indicated in Fig. 30-2, are formed in varying amounts in the pituitary, in two brain nuclei, and in some peripheral tissues.[25] In some animals α-MSH arises primarily in an intermediate part of the pituitary. The hormone has a direct effect on the melanocytes (Box 8-F) causing darkening of the skin. In addition, the various melanocortin peptides (ACTH, α-, β-, and γ-MSH) bind to five different types of receptors. These have been linked to the control of energy homeostasis, appetite, and obesity in both mice and humans.[87–89] The sequence His-Phe-Arg-Trp (of the green-shaded sequence in Fig. 30-2) is essential for binding. In keratinocytes α-MSH may form a 1:1 complex with tetrahydrobiopterin,[90] the coenzyme for tyrosine hydroxylase, and a regulator of tyrosinase, an essential enzyme for melanin formation (Fig. 25-6).

3. Pancreatic and Related Hormones

The functions of the 51-residue insulin (Figs. 7-17 and 7-18) are discussed in Chapter 11. Its actions begin early in life. Mammalian preimplantation blastocysts already show a response to insulin. The glucose transporter GLUT1 is present at the earliest stages; synthesis of GLUT2 and GLUT3 (p. 416) begins at the eight-cell stage. However, the insulin-regulated transporter GLUT4 is not present in the blastocyst. A newly discovered insulin-regulated GLUT8 may function during preimplantation development.[91] Although the secretion of insulin is of primary importance to the regulation of the glucose concentration in mammals, it is still not clearly understood.[92] The beta cells have insulin receptors and other components of the insulin signaling system such as the insulin receptor substrates IRS-1 and IRS-2 and phosphatidylinositol 3-kinase (PI3-K). The sensing of glucose by the beta cells is also not yet well understood.[93] This lack of knowledge has made it difficult to improve the treatment of diabetes. A new approach is to engineer non-beta cells to secrete a steady supply of insulin. Such a possibility has been demonstrated in mice using gut K-cells.[94]

The **insulinlike growth factors** (IGF-I and IGF-II) are produced in many different tissues and promote growth of other cells (see Section 6). **Relaxin**,[94a] which is produced in the corpus luteum of ovaries during pregnancy, is responsible for inducing widening of the birth canal during the late stages of pregnancy and inhibits contraction of uterine muscle, perhaps by decreasing the activity of the kinase that phosphorylates the 20-kDa light chains of myosin.[95] Relaxin is found throughout the animal kingdom, even in the protozoan *Tetrahymena*.[96] Its structure is apparently identical in pigs, whales, and in a primitive tunicate.[97] In human males relaxin is apparently produced in the prostate, where it may function as a sperm motility factor.[97] Relaxin, IGF-I, and IGF-II are all structurally homologous to proinsulin and contain the characteristic 3-disulfide structure of insulin. The IGF-I receptor structure resembles that of the insulin receptor (Fig. 11-11) and also that of the epidermal growth factor (EGF) receptor.[98]

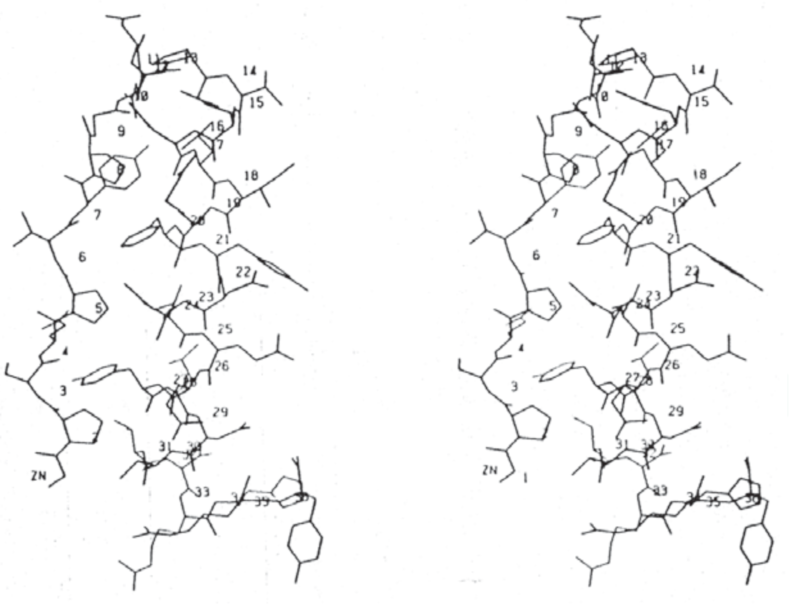


Figure 30-5 Structure of the avian pancreatic polypeptide, a small globular protein. From Blundell *et al*.[107]

Glucagon belongs to a family that also includes the gastrointestinal hormones **secretin**, **gastrointestinal inhibitory peptide** (**GIP**), **vasoactive intestinal peptide** (**VIP**), and **glicentin** (Table 30-4). The function of glucagon in regulation of the blood glucose level is considered in Chapter 17, but the hormone may have other effects. A complex processing pathway converts 14- to 16-kDa preproglucagons into the active hormone.[99,100] Proglucagon is processed to glucagon in the pancreas, but in the endocrine L cells of intestinal mucosa it yields glicentin, a polypeptide containing the entire glucagon sequence, and other products.[101–102a] Glucagon receptors generate both cAMP and Ca^{2+} as second messengers.[103]

The 27-residue secretin stimulates secretion of bicarbonate into the pancreatic juice and inhibits gastric secretion of acid. The 28-residue VIP is found throughout the gastrointestinal tract of mammals and birds as well as in the brain and the lungs. It is a potent vasodilator and may be the major relaxant of pulmonary smooth muscle.[104] It has been reported totally absent from lungs of asthma patients.[105] The gastrointestinal inhibitory peptide (GIP) is larger than VIP but also has a close homology with glucagon (Table 30-4).[106]

The 36-residue **pancreatic polypeptide** is a hormone of uncertain functions. The crystalline polypeptide has at the N terminus an 8-residue collagenlike helix that lies parallel to a C-terminal α helix (Fig. 30-5). The overall shape resembles that of both insulin and glucagon.[107,108] This PP-fold includes also neuropeptide Y, which is considered in the next section,[109] and neuropeptide YY.[110]

4. Gastrointestinal and Brain Peptides

The largest endocrine gland in the body is the gastointestinal tract, which produces a profusion of peptide hormones, many of which are also found in the brain.[111,112] Indeed, a majority of the known vertebrate peptide hormones occur in the brain.[82,112,113] For example, glucagon has been found in the brainstem and hypothalamus.[114] Many of these peptides, or closely related ones, are also found in lower invertebrates.[115] For example, the 10-residue *Hydra* **head activator** is also present in mammalian brain.[115,116] The concentrations of these peptides in the brain is very low (10^{-12} 10^{-15} M).

Gastrin is produced in the lower portion of the stomach and regulates the secretion of acid as well as growth of the gastrointestinal mucosa.[112,117] It may also function as an Fe^{3+} carrier.[117a] The shorter gastrin 17 as well as the longer 34-residue gastrin 34 are both active as is a synthetic pentapeptide with the hormone's C-terminal sequence.[13] The family of **pancreozymin-cholecystokinins** (CCK) are 8- to 58-residue peptides produced in the upper intestinal tract. They have a 4-residue amidated C-terminal sequence in common with gastrin. This tetrapeptide has some biological activity, but eight residues are required for full activity as is conversion of the tyrosine at position seven from the C terminus to an *O*-sulfate ester. Both gastrin and CCK molecules are partially converted to sulfate esters.[118,119] Most regions of the brain contain CCK peptides in amounts exceeding those of other neurotransmitters. These arise (in pigs) from a 114-residue preproCCK.[120–122] The sulfated insect neuropeptide, **leukosulfakinin** (Table 30-4), is homologous to gastrin and CCK.[123] Another hormone **gastrotropin**[124] is produced by cells of the intestinal mucosa in the distal ileum and stimulates gastric secretion.

Motilin, a 22-residue intestinal neuropeptide, stimulates motor activity of the gastrointestinal tract.[125,126] **Bombesin** was first isolated from frog skin but probably also functions in both the intestinal tract and the brain. It has a powerful hypothermic effect.[127,128] A mammalian homolog of bombesin, the 27-residue **gastrin-releasing peptide** (GRP), is found throughout the gastrointestinal and pulmonary tracts as well as the central nervous system.[120,129] Bombesin-like material, possibly GRP, is produced by some cancers and may serve as an autocrine growth factor.[130] The 29-residue **galanin** was originally isolated from procine intestine but is found throughout the central nervous system. It may function as a neurotransmitter or modulator.[131] The 15-residue **guanylin** is an important regulator of epithelial transport in the intestine and probably in other tissues.[132,133] The active hormone, which is cut from the 99-residue proguanylin, contains two disulfide bonds. Guanylin receptors activate guanylate cyclase with production of cyclic GMP, which functions in the regulation of intestinal fluid and electrolyte absorption. The 18- or 19-residue heat stable enterotoxins of some strains of *E. coli* bind to and activate the guanylin receptors. The resulting overproduction of cGMP causes severe diarrhea (see also Box 11-A).[133]

Neuropeptides Y (NPY) and YY are 36-residue amidated peptides that are members of the pancreatic polypeptide (PP) family (Fig. 30-5). NPY is produced both in the peripheral nervous system and in the brain,[110,134] where it is one of the most abundant neuropeptides. Another member of the PP family is **seminalplasmin**, a regulator of calcium ion transport in bovine sperm.[135] NPY is best known for its stimulation of appetite. It also inhibits anxiety and increases memory retention. It has a vasoconstrictive effect on blood vessels, participating in cardiovascular regulation.[136,137] Peptide YY is formed in endocrine cells of the intestine, while NPY is formed in neurons of the parasympathetic system.[138] Both participate in regulation of fluid and electrolyte secretion. Both are found in other vertebrate species.[139]

NPY is one of the most important of several regulators of feeding behavior of animals. PYY_{3-36}, another member of the neuropeptide Y family, suppresses appetite by antagonizing the action of NPY.[139a,b] A large variety of hormonal effects seem to be involved in control of appetite.[139b,140,141] There are both short-term and long-term mechanisms. For example, when introduced into the gut of rats prior to feeding, CCK and various other gastrointestinal peptides decrease the amount of food eaten.[140]

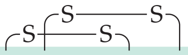
Much attention has been focused on the 146-residue cytokine **leptin**, a hormone produced by adipose tissue.[141–144] Leptin, which is sometimes described as the antiobesity hormone, was recognized by mutations of the *obese* gene (OB) or of the OB receptor in genetically obese mice. When food is scarce, the fat cells shrink and decrease their secretion of leptin. The decrease is sensed by receptors in the hypothalamus, which signal for increased NPY secretion and decreased secretion of αMSH. NPY increases appetite, while αMSH has an opposing role of blocking feelings of hunger .[144–145a] Nevertheless, there are doubts that leptin's primary role is control of obesity.[146,147]

The 13-residue **neurotensin** was first isolated from the hypothalamus but is more abundant in cells of the ileum.[13] It induces gut contraction, lowers blood pressure, and has a variety of other effects.[127,148] **Substance P** (SP; Table 30-4) has been regarded as a possible neurotransmitter for some time[127] but is also found in the digestive tract. It is the most abundant of a family of five neurokinins (or tachykinins). Others include neurokinin A (substance K), neurokinin B, neuropeptide K, and neuropeptide γ. They have a common C-terminal sequence $FXGLM\text{-}NH_2$.[149,150] Substance P is thought to be involved in the perception of pain, and mice lacking a substance P receptor appear to have reduced sensitivity to pain.[151,152] Substance P as well as the related substance K are derived from two large precursor proteins, which appear to arise as a result of alternative modes of splicing of mRNA.[29]

5. Other Mammalian Peptide Hormones

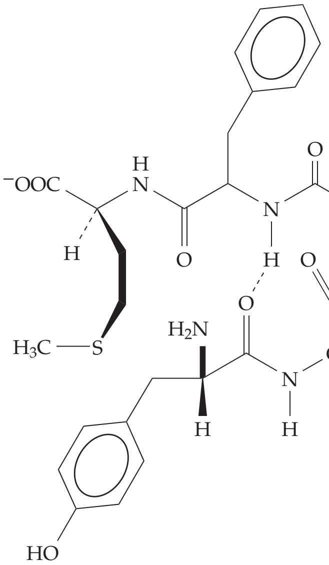
The action of the 32-residue thyroid hormone **calcitonin**[153] has been described in Box 22-C. This calcium-regulatory hormone is produced in the thyroid C cells from a precursor having an extra 82 residues at the N terminus and 16 residues at the C terminus. The same gene gives rise in neural tissues to a neuropeptide, possibly a neurotransmitter, called **calcitonin gene-related polypeptide** (CGRPP).[29,154] The 84-residue **parathyrin** (parathyroid hormone) is present in secretion granules as a 90-residue prohormone containing six extra residues at the N terminus. The primary biosynthetic product **preproparathyrin** contains an additional 25 residues.[155] An N-terminal 34-residue fragment of the hormone, when injected subcutaneously daily, causes an increase in bone density in persons with osteoporosis.[156–157] The hormone acts via a G-protein-coupled receptor in bone and kidney (see Box 22-C).[158,159] A calcium ion receptor, which binds Ca^{2+} cooperatively, acts as a sensor that regulates release of the parathyroid hormone to regulate the serum Ca^{2+} concentration.[160] A 141-residue **parathyroid hormone-related protein** has an N-terminal sequence homologous with that of parathyroid hormone, eight of the first 13 residues being identical. It is secreted by a variety of cells and serves as a growth factor.[161]

TABLE 30-4
Some Pancreatic and Gastrointestinal Hormones and Neurohormones and Their Sequences

Name and source	No. of residues	Sequence[a]
Glucagon	29	HSQGTFTSDYSKYLDSRRAQDFVQWLMNT
Secretin (pancreas)	27	HSDGTFTSELSRLRDSARLQRLLQGLV–NH_2
Vasoactive intestinal peptide	28	HSDAVFTDNYTRLRKQMARKKYLNSILN–NH_2
Gastrointestinal inhibitory peptide (GIP)	43	YAEGTFISDYSIAMDKIRQQDFVNWLLAQ–Q^{43}
Glicentin	100	A proglucagon containing the entire glucagon sequence in residues 64–92
Pancreatic polypeptide	36	
Neuropeptides Y (NPY) and YY	36	
Gastrin (stomach)		
Gastrin-17	17	pEGPWLEEEEEAYGWMDF–NH_2[b]
Cholecystokinin, CCK or pancreozymin (gallbladder, pancreas), many forms exist		
CCK 58	58	
CCK 8	8	DYMGWMDF–NH_2
Motilin (porcine)	23	FVPIFTYGELQRMQEKERNKGQ
Bombesin	14	pEQRLGNQWAVGHLM–NH_2
Gastrin-releasing peptide	27	
Galanin	29	
Guanylin	15	 PNTCEICAYAACTGC
Neurotensin	13	pELYENKPRRPYIL
Substance P	11	RPKPQQFFGLM–NH_2
Physaelemin (frog skin)	11	pEADPNKFYGLM
Neurophysins	93 – 95	
β-Endorphin	31	YGGFMTSEKSQTPLVTLFKNAIIKNAHKKGQ
Dynorphin	17	YGGFLRRIRPKLKWDNQ
Met-enkephalin	5	YGGFM
Leu-eukephalin	5	YGGFL
Angiotensin II	8	DRVYIHPF
Bradykinin (BK)	9	RPPGFSPFR
Lys-bradykinin (kallidin)	10	KRPPGFSPFR
Sleep peptide	9	WAGGDASGE
Atrial natriuretic hormone	28	
Chemotactic factors		
for neutrophils	3	f–MLF
for phagocytes	4	TKPR
Speract	10	GFDLNGGGVG

[a] Standard one-letter abbreviations; pE, 5-oxoprolyl; f–, formyl; –NH_2, C-terminal carboxamide.
[b] Y-12 may be sulfated.

Endogenous opioid peptides. Extensive processing is also involved in formation of analgesic opioid peptides, which are present naturally in the brain (see also Section B). The formation of β-endorphin in the hypothalamus from prepro-opiomelanocortin (Fig. 30-2) has already been mentioned. Prior to the discovery of β-endorphin, the pentapeptides **Met-enkephalin** and **Leu-enkephalin** (Table 30-4) were discovered and were found to compete with opiate drugs for receptors in the brain. The larger β-endorphin, which contains the Met-enkephalin sequence at its N terminus, is a far more potent opiate antagonist than are the enkephalins. Since the Met-enkephalin sequence within β-endorphin is not flanked by basic residues, it apparently is normally not released. Two other recently discovered brain peptides are **endomorphin-1** (YPWF–NH_2) and **endomorphin-2** (YPFF–NH_2). They are also potent agonists for the opioid receptors, especially the μ receptor (see Section B,10).[161a,161b]



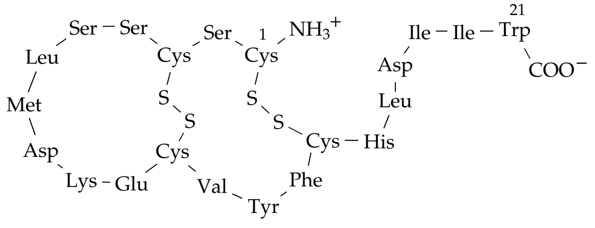
Tyr-Gly-Gly-Phe-Met
(Met-enkephalin)

Both Met-enkephalin and Leu-enkephalin have their own pro- and prepro forms.[29] Bovine **preproenkephalin A** is a 268-residue protein containing a 20-residue signal sequence and four sequences of Met-enkephalin and one of Leu-enkephalin, each flanked by pairs of basic residues. There are also Met-enkephalin-Arg-Gly-Leu (YGGFMRGL) and Met-enkephalin-Arg-Phe sequences. Not all of these are cut out cleanly, and other peptides such as Met-enkephalin-Arg-Arg-Val-NH_2 are also found in brain. **Proenkephalin B** contains three copies of Leu-enkephalin contained within longer peptides. One of these, **β-dynorphin** (Table 30-4), is also a potent opioid compound. The enkephalins are thought to act as neurotransmitters, which are rapidly degraded after their release by two or three membrane-bound peptidases.[29,162] Attempts are being made to design inhibitors that might inactivate these enzymes allowing buildup of enkephalin concentrations with a resultant analgesic effect.

ATP, ADP, and adenosine. Usually regarded as a strictly intracellular compound, ATP is also released into extracellular space. There the ATP, as well as ADP and adenosine, have a variety of local hormonal functions. ATP receptors are found in many tissues and are present in some nerve synapses.[163–166] ATP is one of the substances that induces sensations of pain.[164,167,168] It may affect secretion of saliva,[163] signal a full urinary bladder, induce a feeling of warmth,[167] and have functions in the immune system, in platelet clotting,[166] and as a neurotransmitter. Adenosine has been recognized for many years as an extracellular signaling molecule, a local hormone that can arise by breakdown of ATP or by secretion from cells.[169,170] At least four types of receptor are present in the human body.[170–174] Adenosine is thought to modulate neural responses in many tissues. It may be involved in sleep,[172] in regulation of serotonin transport,[171] and in control of appetite.[171,174] Extracellular ADP appears to have a role in controlling bone osteoclasts (p. 441).[175]

Kinins. These hormones are small peptides that induce contraction of smooth muscles, lower blood pressure (Box 22-D), and increase vascular permeability.[176] They also have a function in contact-activated blood coagulation. The most important human kinins are the nonapeptide **bradykinin**[177,178] and the related decapeptide **lysine-bradykinin** (Table 30-4). Other forms such as Met-Lys-bradykinin and Ile-Ser-bradykinin (T-kinin) are also known. The precursors to the kinins, the **kininogens**,[176] are cleaved by the protease **kallikrein** (Fig. 12-17) or by kallikreinlike enzymes to form the kinins. Kinins are suspected of being important producers of pain in inflammatory conditions such as arthritis.[176a]

Endothelins. Endothelial cells of blood vessels produce **endothelins** that cause vascular smooth muscle contraction and a rise in blood pressure.[179–183] Three human genes code for the closely related endothelins-1, -2, and -3. A 203-residue preproendothelin-1 is processed to form the 39-residue prohormone called **big endothelin-1**. Some of this peptide is secreted and circulates in plasma, where it may have various hormonal functions. Cleavage of the prohormone by a cellular metalloprotease yields endothelin-1, a 21-residue peptide held in a looped configuration by two disulfide bridges. It is homologous to a group of neurotoxins that includes the α-scorpion toxins and ω-conotoxin.[181] These toxins act on voltage-dependent ion channels. Endothelin-2 is produced largely in the kidneys and intestine, while endothelin-3 is found in



Endothelin-1

high concentrations in the brain. Type A endothelin receptors are 7-helix G-protein-coupled proteins, which activate phospholipase C with generation of inositol 1,4,5-trisphosphate and diacylglycerol (Ins-P_3; Figs. 11-4, 11-9). The Ins-P_3 causes release of Ca^{2+}, while diacylglycerol mediates mitogenic responses.

Opposing the effects of the endothelins, which act slowly, is a fast-acting endothelium-derived **relaxing factor**, which has been identified as nitric oxide, NO. It is discussed in Chapter 18, Section F, and in Section 7 of this chapter. Also affecting blood pressure is the potent vasorelaxant **atrial natriuretic factor**. This 28-residue peptide, which is discussed in Box 22-D, is produced by the cardiac atria and stimulates the excretion of Na^+ and of water by the kidneys.[184] It also promotes hydrolysis of lipids within human adipocytes.[185]

Peptides as attractants. Small peptides as well as larger polypeptides serve to attract cells within the human body and other multicelled organisms. Both unicellular and multicellular organisms also use peptides as pheromones. The human immune system depends upon hormonelike **chemotactic factors**. Neutrophils are attracted by such peptides as formyl-Met-Leu-Phe,[186–188] which have a bacterial origin, while the basic tetrapeptide Thr-Lys-Pro-Arg activates the phagocytic polymorphonuclear leukocytes and macrophages.[189] Larger 8- to 10-kDa proteins known as **chemokines** (chemotactic cytokines) attract leukocytes to sites of inflammation (Fig. 30-6).[190–192] Some proteins serve as pheromones. Examples range from the 40-residue mating pheromones of protozoa of the genus *Euplotes*[193,194] to the 17-kDa sex pheromone of the female hamster.[195] The decapeptide **speract** (Table 30-4) is produced by sea urchin eggs and stimulates the respiration of spermatozoa.[196] Similar factors probably function in fertilization of human ova.

6. Protein Growth Factors and Cytokines

The pituitary growth hormone is only one of a large family of protein growth factors that are secreted by cells and which promote the growth of other cells.[200] Many of the growth factors are also described as **cytokines**, local protein hormones that conduct cell-to-cell communication to regulate growth, development, and differentiation.[197,201–203] Among the first growth factors to be recognized were the **insulinlike growth factors** (IGF or somatomedins), mitogenic peptides isolated from plasma. They share some of the metabolic effects of insulin but are less active. On the other hand, they are much more active than insulin[204] in promoting cell growth and proliferation of cells.[205] The abundant IGF-I (somatomedin C), a 70-residue single-chain basic peptide with a sequence and three-dimensional structure homologous to that of proinsulin,[206,207] is considered a major mediator of the action of the pituitary growth hormone (GH, somatotropin). Studies in cell culture suggest that GH may induce differentiation of cells, and that IGF-I may then cause a rapid proliferation of the newly differentiated cells.[208] The homologous 67-residue IGF-II may have a similar function in fetal development.[209] The cell surface receptor for IGF-I is similar to the insulin receptor, but IGF-II receptor is structurally different. It is a monomeric 250-kDa protein; and although it is a substrate for a tyrosine kinase, it has no kinase activity of its own.[210]

The 53-residue **epidermal growth factor** (EGF or **urogastrone**) is found in human urine and in very high concentration in the submaxillary salivary glands of male mice. Like the pancreas these glands contain both endocrine and exocrine tissues. EGF is synthesized in mice as a 1217-residue precursor, which contains not only the EGF sequence but also seven other related sequences.[211] Related growth factors include transforming growth factor-α (TGF-α), neuregulins,[212–214] betacelulin, and epiregulin, all of which promote growth of epithelial cells and are involved in wound healing.[215] The EGF and related peptide chains are each crosslinked by three disulfide bridges. The three-dimensional structure of EGF, deduced from NMR measurements, contains largely β structure and loops and is organized into two domains in a "mitten shape."[216] The receptor for EGF is a 1186-residue transmembrane glycoprotein. Its relationship to cellular oncogene c-*erbB* has been discussed in Chapter 11, Section H. The extracellular glycosylated N-terminal region of the receptor contains the EGF-binding site.[217–220] It also contains two cysteine-rich repeat sequences homologous to one of those in the insulin receptor A chain (Fig. 11-11). The cytoplasmic C-terminal part of the EGF receptor contains a 250-residue tyrosine-specific protein kinase sequence. Following dimerization the EGF receptor phosphorylates tyrosine residues in various proteins including itself (autophosphorylation).[212,219,221] The receptor is also phosphorylated on Thr 654 and other residues through the action of the Ca^{2+}- and phospholipid-dependent diacylglycerol-activated **protein kinase C** (Fig. 11-9).[222] Serines 1002, 1046, and 1047 may also become phosphorylated, perhaps resulting in desensitization of the receptor.[223]

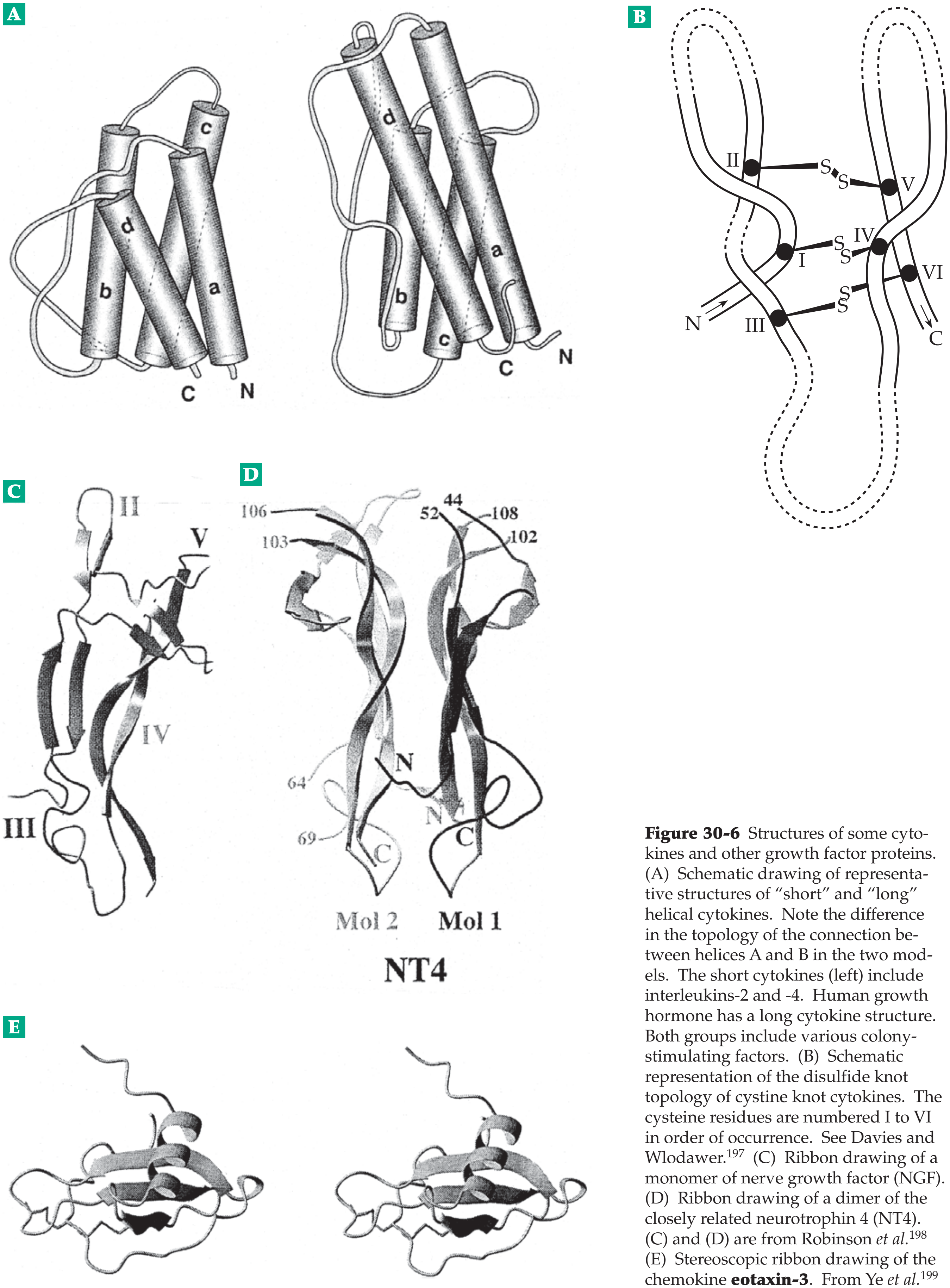


Figure 30-6 Structures of some cytokines and other growth factor proteins. (A) Schematic drawing of representative structures of "short" and "long" helical cytokines. Note the difference in the topology of the connection between helices A and B in the two models. The short cytokines (left) include interleukins-2 and -4. Human growth hormone has a long cytokine structure. Both groups include various colony-stimulating factors. (B) Schematic representation of the disulfide knot topology of cystine knot cytokines. The cysteine residues are numbered I to VI in order of occurrence. See Davies and Wlodawer.[197] (C) Ribbon drawing of a monomer of nerve growth factor (NGF). (D) Ribbon drawing of a dimer of the closely related neurotrophin 4 (NT4). (C) and (D) are from Robinson *et al.*[198] (E) Stereoscopic ribbon drawing of the chemokine **eotaxin-3**. From Ye *et al.*[199]

Binding of EGF to its receptor produces within minutes an increased transcription rate for the prolactin gene and other nearby genes.[221] The urinary form of EGF, urogastrone, is an inhibitor of ulcer formation. It is found in relatively large amounts in the urine of pregnant women (who tend not to develop ulcers).

Platelet-derived growth factor (PDGF)[224] is released from the α-granules of blood platelets during clot formation and is thought to stimulate the growth and mitosis in fibroblasts that is necessary for wound healing.[225] It consists of two chains, A and B. The 31-kDa precursor of the A chain is encoded by the cellular oncogene *c-sis* (p. 571).[224] The **PDGF receptor** is another transmembrane glycoprotein with a C-terminal tyrosine kinase domain. However, its construction differs from that of the insulin or EGF receptors. The external part of the single-chain receptor appears to contain five immunoglobulinlike domains (see Figs. 2-16 and 12-18).[226] Binding of PDGF to the receptor causes responses within minutes.[217,226] These include activation of the tyrosine kinase, hydrolysis of phosphatidylinositides, increases in the levels of cAMP and of Ca^{2+}, and increased transcription of a group of genes. The last include the proto-oncogenes c-*myc* and c-*fos*, which encode proteins that regulate transcription. The PDGF receptor is part of a recognized autocrine stimulatory loop in cells infected with a virus carrying the v-*sis* oncogene.[217] The oncogene product resembles PDGF and binds to the PDGF receptors of the cell producing the v-*sis* product. In this way the cancer cell stimulates its own growth.

Transformation of kidney fibroblasts into cancerlike cells can be induced by the concerted action of PDGF, an analog of EGF, and the **transforming growth factor** (TGF-β).[225,227,228] The latter is one of 30 or more related growth factors that have numerous functions in normal tissues.[229] Platelets produce a relatively large amount of the 25-kDa TGF-β, and it too may be involved in wound healing. TGF-α is a smaller protein with a structure resembling that of EGF.[227,228] While TGF-β inhibits epidermal cell growth, TGF-α stimulates growth. It is found in elevated levels in the skin lesions of **psoriasis** (Box 8-F) and may be the cause of the excessive epithelial growth in that disease.[230]

There are at least nine **fibroblast growth factors** (FGFs). Originally found in brain, they act on many cells including the endothelial cells that line blood vessels.[231] Basic FGF[232] and acidic FGFs[233] have homologous sequences[234] and are also related to the lymphokine interleukin-1. **Vascular endothelial cell growth factor** (VEGF), which is similar to PDGF, is essential for maintenance of the endothelium. The FGFs and VEGF as well as TFG are potent **angiogenic factors** needed for growth of blood vessels.[235–238] These proteins are important not only to normal blood vessels but also to invasive tumors that must develop blood vessels in order to grow. Excessive production of angiogenic factors may also be a factor in eye diseases including the retinal deterioration caused by diabetes.[239] Another protein, **angiogenin**, is a ribonuclease,[240] which is discussed on p. 648.

There are four closely related transmembrane FGF receptors and subforms that arise by alternative mRNA splicing.[241–243] The receptor structures include three external immunoglobulinlike domains and an internal tyrosine kinase domain at the C terminus. Mutations in FGF receptors are associated with a variety of skeletal defects and other hereditary problems.[241,244] For example, the Gly380Arg substitution in the transmembrane segment of FGF receptor 3 is the major cause of **human dwarfism** (achondroplasia).[245] The fibroblast growth factors, as well as other proteins such as the IGFs, HGF, and TGF-β, bind not only to their receptors but also to heparan and heparin. This binding appears to be a major factor in controlling the availability of the growth factors.[242,246,247]

The **nerve growth factor** (NGF) was identified over 40 years ago by Rita Levi-Montalcini[248] on the basis of its activity in promoting the profuse outgrowth of neurites from embryonic neurons (Fig. 30-7). The 118-residue monomer consists largely of three β-hairpin loops, which are held together by three disulfide bridges that form a “cystine knot.” The C15–C80 disulfide passes through a ring formed by the C58–C108 and C68–C110 disulfide bridges (Fig. 30-6B,C).[249] A similar folding pattern and disulfide core are found in TGF-β2 and also in several other **neurotrophins**, growth factors involved in the development and survival of neurons (Fig. 30-6D).[198,250–251a] NGF may also have a more general function in promoting tissue repair.[252] Like EGF nerve growth factor is most abundant in the submaxillary glands of male mice. Larger oligomers containing bound Zn^{2+} are present in mouse submaxillary glands. Two different receptor proteins, one of which is a tyrosine kinase, are present on many cell surfaces.[198,251,253] The glial cells, which lie between the neurons, have their own growth factors.[254]

Bone formation and resorption are influenced by several protein factors. For example, IGF-I stimulates formation of bone, but EGF promotes breakdown.[256] Additional **bone-derived growth factors** and **morphogenetic factors** also have been described.[256,257] A **cartilage-inducing** factor has been identified as TGF-β.[258]

A group of glycoproteins function as hematopoietic growth regulators in the development of blood cells.[259–264] The 166-residue cytokine **erythropoietin** is the primary regulator of red blood cell formation in mammals.[260,264,265] At least four glycoprotein **colony-stimulating factors** (CSF) promote proliferation of granulocytes and macrophages.[259,266,267] The lymphocyte-produced **lymphokines** include the **interleukins** and other proteins. Two species of

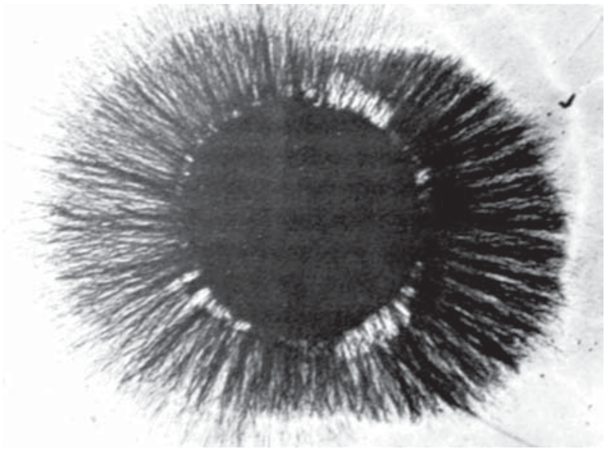


Figure 30-7 Effect of one ng of nerve growth factor in promoting the production of neurites in a chick embryonic sensory ganglion. From Frazier *et al*.[255]

interleukin-1 (IL-1) serve as mediators of inflammation.[267,268] They induce proliferation of T lymphocytes and fibroblasts, bone resorption, release of acute phase proteins (Section C), breakdown of cartilage, and fever.

Interleukin-2 (T-cell growth factor; Fig. 30-6A) is secreted by some activated T-lymphocytes. This 133-residue largely helical protein is involved in generation of cytotoxic T-cells, stimulation of interferon release, and of release of a B-cell growth factor.[269] Considerable excitement has accompanied the possibility of activating lymphocytes with IL-2 produced from cloned genes in bacteria to increase their ability to kill cancer cells. However, IL-2 is toxic, and this is limiting its use. See also Chapter 31, Section C.

IL-3 is one of the colony-stimulating factors, which stimulates the growth of many types of blood cells.[270] Other lymphokines include one derived from T helper cells, which activates resting T lymphocytes thus amplifying an immune response. Others (Chapter 31) are **α-interferon** and the neurotrophic factor (autocrine motility factor) **neuroleukin**.[271] It acts in monomeric form, but as a dimer it seems to be identical to the enzyme phosphoglucose isomerase.[272] While most hormones regulating growth and differentiation seem to be large peptides or proteins, **bursin**, which induces differentiation of lymphocytes, is the amidated tripeptide Lys-His-Gly-NH_2.[273] The corresponding differentiation hormone for T lymphocytes is the 49-residue **thymopoietin**, a hormone of the thymus gland.[274]

Tumor necrosis factor (TNF, also called cachetin) is a 157-residue hormone secreted by macrophages. It is a mediator of inflammatory responses including fever, shock, and **cachexia**, the wasting of the body during chronic diseases including cancer. TNF was isolated as the causative agent of cachexia and also as a factor produced in acute bacterial infections, which sometimes caused death of tumor cells and spontaneous recovery from cancer. In the latter case, it is the lipopolysaccharide (Fig. 8-30) and other bacterial endotoxins that induce the release of TNF by macrophages. Its extreme toxicity has prevented immediate harnessing of the tumor-killing potential of TNF. One function of TNF is regulation of transcription factor NF-κB (Fig. 5-40) in neutrophils and macrophages,[275] a key part of the inflammatory response. TNF also mediates programmed cell death (apoptosis)[276] and has been linked to obesity-induced insulin resistance.[277] The cell surface TNF receptors have a variety of modular structures consisting of various disulfide-linked subdomains.[276]

This long list of vertebrate peptide growth and regulatory hormones is not complete. The biological actions of these hormones are also complex. Growth factors usually have pleiotropic effects, which may involve many tissues as well as many regulatory systems. Are there any simplifying generalizations? Loret *et al*.[278] point out that some growth factors such as IFG-1 and EGF are ubiquitous, affecting virtually all tissues. Others, such as PDGF and thrombin (Fig. 12-17), are more localized in their effects. Some, such as the lymphokines, are more specialized. For one group of hormone receptors the effects are mediated by tyrosine kinases and internalization of the receptors. Another group of receptors activate G proteins and, in turn, adenylate cyclase or phospholipase C. The regulatory domains of the various receptors overlap, a property that allows different tissues to respond differently to hormonal stimuli. The result is the network of interactions that makes the body so sensitive and responsive.

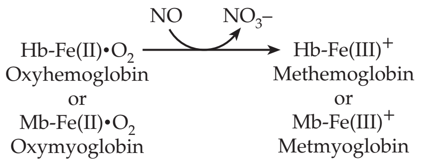
7. Nonpeptide Mammalian Hormones

Most nonpeptide hormones have been considered in other places in the book as indicated in Table 30-2. Because of their importance in the brain adrenaline, noradrenaline, serotonin, and melatonin are also dealt with in Section B,9 of this chapter.

The volatile hormones nitric oxide (NO) and carbon monoxide (CO). The free radical molecule nitric oxide, commonly abbreviated as either •NO or simply NO, is formed by hydroxylation of guanidine groups of arginine (Eq. 18-65). First recognized as the **endothelium-derived relaxing factor**,[279] NO has received increasing attention because of its involvement in a broad range of physiological processes. These include regulation of blood pressure through effects on smooth muscles of the vascular endothelium, regulation of several aspects of the innate immune system (Chapter 31), and neurotransmitter functions both in the brain and in the peripheral nervous system.[279–281] Roles for NO in bacteria, other microorganisms, and plants have also been discovered.[281,282] These often involve regulation of transcription.[283] As mentioned in Chapter 18, Section F,2, many of the effects of NO are a result of activation of soluble guanylate cyclase (p. 561).[283–285a] In the endothelium other hormones, such as the endothelins (p. 1750), atrial natriuretic factor, and bradykinin (Box 22-D), cooperate in the regulation of NO synthase.[279] Neuronal NO synthase functions in the brain in olfaction and in formation of memory. In the peripheral system it mediates penile erection[284,285–286a] and plays a variety of roles in the enteric nervous system.[287] Neuronal NO synthase is often localized to synaptic regions by binding to tissue-specific proteins.[288] NO may also regulate cellular respiration by inhibition of cytochrome *c* oxidase.[285]

In high enough concentrations NO is toxic. It is formed of phagocytic cells and utilized in the killing of ingested pathogens.[279] It also contributes to the inflammatory response of tissues.[289,290] Even the firefly's flash is triggered by a pulse of NO.[291] The dangerous **stonefish**, whose sting causes death within six hours, apparently utilizes NO to kill its victim. A 148-kDa lethal protein (stonustoxin) in its venom induces rapid formation of NO, which causes a fatal drop in blood pressure.[291a]

Like carbon monoxide, NO binds tightly to many metal centers within a cell.[292,293] This has added greatly to the problem of understanding the mechanisms of its action. NO also reacts rapidly with thiol groups of proteins and of small molecules such as glutathione.[294,295] Because of its importance in the regulation of blood pressure, reactions of NO with hemoglobin and the related myoglobin have been studied intensively.[296–300a] NO binds to hemoglobin 1000 times more tightly than either O_2 or CO, preferentially occupying the hemes of the α subunits.[298] Because there is so much hemoglobin in red blood cells, at most one NO per hemoglobin molecule can react. This allows as much as one NO to be carried to tissues along with three O_2 molecules. If the NO could be released in the capillaries, it would activate guanylate cyclase. The resulting cGMP would induce relaxation of smooth muscles and reduce blood pressure.[279] However, tight bonding of NO to deoxyhemoglobin would prevent this release. A plausible possibility (with experimental support) is that NO is not bound to Fe but to the SH group of the conserved cysteine 93 of a β subunit of hemoglobin as SNO (*S*-nitrosothiol) hemoglobin. The NO may bind initially to the iron atom of an α subunit, but then be transferred to the nearby β Cys 93 (p. 359) to form the SNO-Hb.[296] NO may then move from SNO-Hb to thiol groups in the tissues. Recent evidence suggests that the transfer occurs first to an SH group in the anion exchange **AE1** (p. 420).[299] An alternative explanation, which does not involve SNO-Hb, is that hemoglobin Fe-NO is converted to **nitrite** via oxidation of the iron to form a methemoglobin subunit (Eq. 30-1), and that it is nitrite which serves as the endothelial relaxing agent.[300]



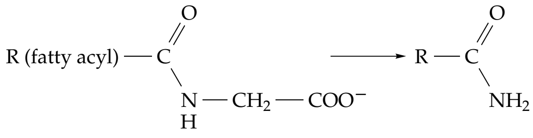
$$\underset{\substack{\text{Oxyhemoglobin}\\ \text{or}\\ \text{Mb-Fe(II)}\bullet\text{O}_2\\ \text{Oxymyoglobin}}}{\text{Hb-Fe(II)}\bullet\text{O}_2} \xrightarrow{\text{NO} \quad \text{NO}_3^-} \underset{\substack{\text{Methemoglobin}\\ \text{or}\\ \text{Mb-Fe(III)}^+\\ \text{Metmyoglobin}}}{\text{Hb-Fe(III)}^+} \qquad (30\text{-}1)$$

The function of the monomeric myoglobin has often been assumed to be participation in facilitated diffusion of O_2. Although this is an important function under some circumstances, an additional role for myoglobin may be to scavenge NO, via Eq. 30-1, preventing its buildup to dangerous levels. The metmyoglobin produced can be reduced by methemoglobin reductase.[301,302] A different situation is met by the parasitic nematode *Ascaris*, whose hemoglobin binds O_2 so tightly that it can't serve as an O_2 carrier. It may serve as an **NO-activated deoxygenase**, again using Eq. 30-1 to remove O_2, which can be toxic to the nematode.[303] Free myoglobin can also react with NO to form heme–NO and heme–nitroxyl complexes.[304,305]

Like NO, CO also binds tightly to heme iron and is able to activate guanylate cyclase.[306] CO is formed in the human body by the action of heme oxygenases (Fig. 24-24). Synthesis of heme oxygenase-1 (HO-1) in smooth muscle is induced by a low oxygen tension (hypoxia). The resulting elevated level of CO not only may produce increased vasodilation, but also may inhibit synthesis of vascular smooth muscle cells.[307] Heme oxygenase-2 (HO-2) is found in the brain, where it is colocalized with soluble guanylate cyclase.[308] Some other organisms have a more active CO metabolism. The CO oxidation system of *Rhodospirillum rubrum* is activated by a CO-sensing heme protein, which acts as a transcriptional regulator. The CO binds to the heme iron, apparently inducing a conformational change that allows the protein to bind to its target DNA sequence.[309,310]

Hormonal lipids. We have already considered a number of hormones that are not water-soluble but may have to be transported by carrier proteins to their sites of action. These include retinoic acid (Box 22-A), metabolites of vitamin D (Box 22-C), and the platelet-activating factor (Box 8-A). The last functions in the brain[311] as well as in blood. Hormonal lipids also include the prostaglandins (Fig. 21-7), leukotrienes, and lipoxins (Fig. 21-8). These are products of the eicosenoid cascade or network, which is activated by receptors linked to phospholipase C (Fig. 11-9). Ceramide formed by hydrolysis of sphingomyelin initiates additional responses.[312,313] Sphingolipids may also be important mediators of apoptosis.[314]

The sleep-inducing oleamide (p. 382) modulates signaling by serotonin-dependent and Gaba-dependent neurons and blocks gap junction signaling in brain glial cells.[315–318] Oleamide is one of a family of fatty acid amides found in human plasma. One of these, found also in the brain, is **anandamide** (arachidonoylethanolamide). It is an endogenous activator of the brain cannabinoid receptors.[318–320] The 22-carbon **erucamide** (*cis*-13-docosenamide) stimulates growth of blood vessels.[316] The fatty acid amides are apparently synthesized from corresponding acylglycines (Eq. 30-2) by the action of the peptidylglycine α-amidating enzyme using the mechanism of Eq. 10-11. See also pp. 1792, 1793.

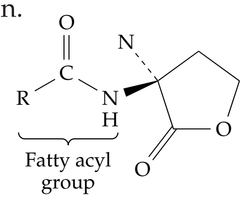


$$\mathrm{R\ (fatty\ acyl)-\overset{O}{\overset{\|}{C}}-\underset{H}{N}-CH_2-COO^-} \longrightarrow \mathrm{R-\overset{O}{\overset{\|}{C}}-NH_2} \qquad (30\text{-}2)$$

The fatty acid amides are destroyed by an integral membrane protein, a **fatty acid amide hydrolase**.[321,322]

8. Nonvertebrate Hormones and Pheromones

The chemical signals that are passed between bacteria and other microbial cells often resemble hormones of vertebrates. Thus, some bacteria secrete peptide mating pheromones. The sequence of an octapeptide[270] of this type from *Streptococcus faecalis* is given in Table 30-5. Many bacteria utilize "quorum signaling." They do not secrete signaling molecules until they sense that there are enough of them to be effective if they act in unison. Then they all secrete an inducer. The best-known example is the induction of bioluminescence of *Vibrio fischeri* (Eq. 23-49). Long-chain fatty acyl derivatives of L-homoserine lactone act as secreted inducers.[323,324] A lactonase hydrolytically inactivates the inducer to avoid excessive accumulation.

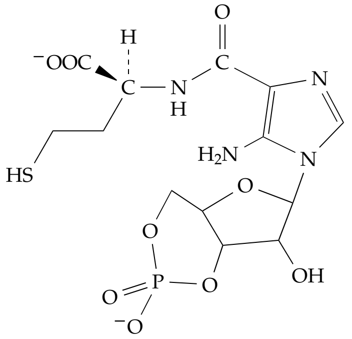


N-Acylhomoserine lactones

Depending upon the types of bacteria and the specific response a variety of different fatty acyl groups may be present in the inducer. Other responses include the formation of bacterial film (biofilms) on a surface and release of virulence factors that induce attack on a host.

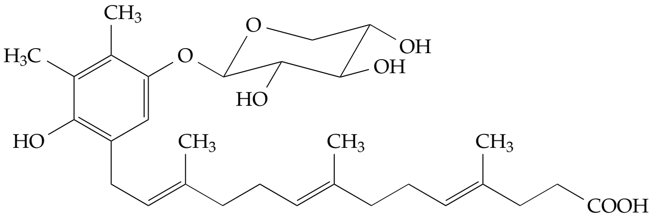
Sexual conjugation in yeast is also induced by pheromones (mating factors).[325–327] Yeast cells of mating type **a** synthesize the 12-residue mating factor **a** which contains a C-terminal cysteine methyl ester *S*-alkylated with a *trans*,*trans*-farnesyl group (Table 30-5). Cells of type **α** synthesize a 13-residue factor **α**.[327a] Cells are attracted to the pheromone produced by cells of the opposite type. The **tremerogens**, sex hormones of certain basidiomycetes, have related structures (Table 30-5).[328]

Peptides are not the only fungal hormones. The water mold *Blastocladiella* releases a **zoospore maintenance factor**, a cyclic phosphate derivative of 5'-phosphoribosyl-5-aminoimidazolecarboxamide.[329] It is similar to the succinocarboxamide, which is an intermediate in *de novo* synthesis of purines (Fig. 25-15). Substitution of homocysteine for L-aspartate in step *g* of that sequence could generate a precursor to the zoospore maintenance factor.



Proposed structure of zoospore maintenance factor of *Blastocladiella.*

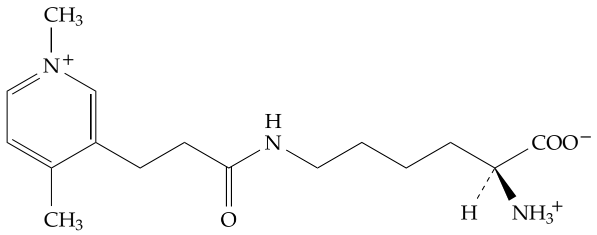
Male sperm cells of the alga *Chlamydomonas allensworthii* (Fig. 1-11) are attracted to female gametes by a pentosylated isoprenoid quinone.[330]



Compound 1

Neurohormones of invertebrate animals. The Cnidaria (coelenterates) have the simplest known nervous system. Simple amidated tetrapeptides, some of which are also found in molluscs, are among their neurotransmitters.[331] EGRFamide and L-3- phenyllactyl-LRNamide are found in some sea anemones.[332] When a hydra (Fig. 1-13) is cut into two pieces, one containing a head and one a foot, each piece reforms the missing end. The decapeptide **head activator** (Table 30-5) diffuses upward from the foot end and induces formation of a head. Similarly, a hormone produced by head cells may induce growth of a new foot end. Glutathione, flowing out from the hydra's prey after wounding by a nematocyst, is the feeding attractant for *Hydra vulgaris*.[282,333] Termination of the response is dependent upon nitric oxide in this primitive invertebrate.

Certain large anemones enter into a symbiotic relationship with fishes, which recognize chemical signals from the anemones and are also chemically protected from the anemones' stings. One of the fish attractants, **amphikuemin**,[334] is effective at a concentration of 10^{-10} M.



Amphikuemin, a powerful fish attractant

A variety of peptide neurohormones are produced by molluscs. Among these are the sea snail *Aplysia*, which is studied because of its simple nervous system and giant neurons.[335,336] Proteolytic processing of precursors within single cells often yields neurohormones specific to those cells. The structures of two **small cardioactive peptides** secreted from single neurons[337] are shown in Table 30-5. A 4.4-kDa egg-laying hormone is formed in at least eight processing steps.[338–341] Among the peptides identified in the freshwater snail *Lymnaea* are many that are characteristic of mammalian pituitary gland, pancreas, brain, and intestinal tract. These include TRH, ACTH, αMSH, arginine vasopressin, ocytocin, calcitonin, gastrin, gastrointestinal peptide, glucogen, insulin, Met-enkephalin, pancreatic poly- peptide, secretin, somatostatin, substance P, and vasoactive intestinal peptide. Also present are FMRF amide and arginine vasotocin.[342]

The cardioacceleratory peptide FMRFamide (Table 30-5), which was discovered in 1977, was the first in a large series of related neuropeptides that are found in organisms ranging from the nematode *Caenorhobditis*

TABLE 30-5
Some Microbial and Invertebrate Peptide Hormones

Source name	Number of residues	Sequence[a]
Streptococcus faecalis sex hormone	8	FLVMFLSG
Yeast mating factor a	12	farnesyl[b] YIIKGV(L)FWDPAC–OCH$_3$
Tremerogen A-10	10	farnesyl[b] EHDPSAPGNGYC–OCH$_3$
Hydra head activator	10	pEPPGGSKVIF
Antho-RF amide (sea anemone)	4	pEGRF–NH$_2$
Small cardiovascular peptides (*Aplysia*)		
SCP-A	11	ARPGYLAFPRM–NH$_2$
SCP-B	9	MNYLAFPRM–NH$_2$
FMRFamide (coelenterates, molluscs)	4	FMRF–NH$_2$
(octopus)	7	YGGFMRF–NH$_2$
Shrimp blanching hormone	8	pELNFSPGW–NH$_2$
Fidler crab pigment-dispersing hormone	18	NSELINSILGLPKVMNDA–NH$_2$
Proctolin (cockroaches)	5	RYLPT
Myotropic neuropeptide (cockroaches)	11	EQFEDYGHMRF–NH$_2$ ↑ sulfate ester in leukosulfakinin
Adipokinetic hormone (locust)	10	PELNFTPNWGT–NH$_2$
Crustacean cardioactive peptide	9	┌S–S┐ EPFCNAFTGC–NH$_2$

[a] One-letter abbreviations, pE, 5-oxoproline; –NH$_2$, C-terminal carboxamide.
[b] Alkylated on S.

elegans to vertebrate animals. At least 18 genes in *C. elegans* encode 53 distinct FMRFamide-related peptides. Disruption of one of these genes causes hyperactivity, uncoordination, and other behavioral difficulties in the nematodes.[343] Several groups of FMRFamide-related peptides have been found in *Drosophilia*[344] and other insects.[345] One group has the C-terminal FLRFamide and another HMRFamide. Among the latter are sulfate esters such as the cockroach neuropeptide shown in Table 30-5. The sequence of the sea urchin sperm chemoattractant **speract** (sperm attractant peptide-1; SAP-1) is shown in Table 30-4. This is one of a family of egg-associated peptides that stimulate sperm metabolism and mobility. The DNA sequence that codes for speract predicts formation of a 296-residue protein that contains four speract sequences plus six related decapeptide sequences, each separated by a single lysine residue.[346] Many other SAP peptides, some containing the unusual amino acid *o*-bromo-L-phenylalanine,[347] are formed.

Hormones of insects and crustaceans. Peptide neurohormones of insect brains[348] include the pentapeptide **proctolin** (Table 30-5), which was first iso-lated from the cockroach and has since been found in crustaceans and in mammalian brain. It has been traced to specific insect neurons.[349] A nonapeptide neurohormone from the shore crab does not resemble any other known vertebrate or invertebrate hormone.[350]

The prawn *Pandalus borealis* changes its body color by means of movable pigment granules. The neurosecretory octapeptide **blanching hormone**[351] (Table 30-5) controls the process. This is a member of a larger family of peptide hormones with related functions such as the 18-residue **pigment-dispersing hormone** of the fiddler crab[352] and similar hormones in insects.[353]

One of the insect neurohormones, the **activation hormone**, controls the secretion of the corpora allata, paired glands that synthesize the **juvenile hormone** (Fig. 22-4) in insect larvae. While the structure of the juvenile hormone varies somewhat with species, it is usually a polyprenyl ester. A specific binding protein provides the hormone with protection from degradative enzymes. However, in the tobacco hornworm an esterase, able to hydrolyze the protein-bound juvenile hormone, is produced at the start of pupal differentiation.[354] The exact mechanism of action of juvenile hormones has been difficult to determine. However, it affects polyamine synthesis.[355,356]

The **mandibular organ-inhibiting hormone** of the crab *Cancer pagurus* produces two neurohormones that inhibit the secretion of methyl farnesoate, which is thought to function as the juvenile hormone in crustaceans.[357]

A related role of the insect corpora allata is to store and release the **prothoracicotropic hormone**, a peptide neurohormone formed in a single neurosecretory cell of the brain.[358,359] The steroid hormone **ecdysone** (Fig. 22-12) is secreted by the insect's prothoracic gland. Also known as the **molting hormone**, ecdysone is required for the periodic replacement of the exoskeleton of the larvae.[359a] It induces molting in crayfish and other arthropods and appears to be needed by such members of lower phyla as schistosomes and nematodes. It also controls the biting behavior of mosquitos.[360] In addition to α-ecdysone, the 20- and 26-hydroxyecdysones and 20,26-dihydroxyecdysone have been identified in insects.[348] It has been suggested that different ecdysones may function at different stages of insect development.

Ecdysone stimulates the synthesis of RNA in tissues. Visual demonstration of the effect is provided by its action on polytene chromosomes of fly larvae (Fig. 26-14).[361] Fifteen minutes after the application of ecdysone, a puff is induced on one band of the chromosome; a second puff forms at a later time while a preexisting puff diminishes. Thus, like steroid hormones in mammals, ecdysone appears to have a direct controlling effect on transcription. The cuticle-shedding process (**ecdysis**) is initiated by the brain peptide **eclosian**. However, the brain may be responding to the **ecdysis-triggering hormone**, a peptide that is secreted by a series of epitracheal glands located in various segments of the body.[362]

Adipokinetic hormones control metabolism of insects during long-distance flight.[359,363] In the migratory locust these hormones consist of a pair of related octapeptides and a decapeptide (Table 30-5). The hormones stimulate triacylglycerol lipase in the insects' fat bodies, induce release of carbohydrates from body stores, and affect many other aspects of metabolism.[363] Insects also have hormones of the insulin family, proteins consisting of disulfide-linked A and B chains as in insulin. The silkworm *Bombyx mori* has 38 genes for the insulinlike **bombyxins**, which are synthesized in the brain.[364]

Insects produce many different types of sex attractant pheromones (e.g., see p. 382). By 1995 more than 300 structures had been determined for pheromones from >1600 insect species.[365]

9. Plant Hormones

Plants possess a kind of circulatory system by which fluids are transported from the roots upward in the xylem and downward from the leaves through the phloem. Many compounds are carried between cells in this manner, while others are transported across cell membranes and against concentration gradients by active transport. A number of compounds that move between cells in either of these two manners have been classified as hormones.[366–369] The major plant hormones consist of five compounds or groups of compounds: **auxins** (p. 1446), **gibberellins** (Eq. 22-5), **cytokinins** (Fig. 5-33), **abscisic acid** (Fig. 22-4), and **ethylene** (Fig. 24-16). A number of other plant regulators, some involved in defensive reactions, are sometimes also described as hormone. These include the **brassinosterols** (Fig. 22-9) and related compounds,[366,367,369,370] **jasmonic acid** (Eq. 21-18), **salicylic acid** (Chapter 25, Section B,7), bacterially produced **lipooligosaccharides** such as the NOD factors (Box 20-E), and polypeptides such as **systemin**. The sugars glucose and sucrose have hormonelike functions as does light, which controls many plant functions.[368]

Plant hormones have multiple and overlapping functions, which are exerted predominantly by repression of gene expression. This makes it difficult to discuss their functions briefly. Most studied are the auxins, of which the principal member is **indole-3-acetic acid** (Fig. 25-12). This compound, whose biosynthesis is discussed on p. 1446, has been implicated as a controlling agent for cell division and cell elongation. In this capacity auxin influences a great variety of plant processes. Produced principally by growing shoots, auxin diffuses down the stem aided by special **efflux carriers**[371,371a] and inhibits the growth of lateral buds. However, the hormone stimulates the growth of stems, thus establishing the apical dominance of the tip of a plant. Other hormones also have an influence. Auxin is well established as the controlling agent in phototropism, the tendency of a plant to bend toward the sun.

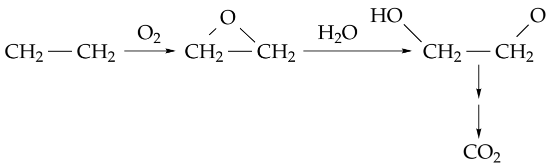
A sensitive test for auxin, which is dependent on the bending of the coleoptile of *Avena sativa* (the common oat) in response to the hormone, allows detection of as little as three pmol of auxin. Using this assay, it was shown that auxin is transported laterally away from the illuminated side of plants, causing the darker side to elongate more rapidly. Both membrane-associated and soluble binding sites, which may represent natural auxin receptors, have been identified, and auxin response elements have been located in DNA. The membrane-bound receptors may regulate an ATP-dependent proton pump, while the soluble receptors may act to regulate gene transcription.

The gibberellins together with brassinosterols are active in helping to determine the *form* of plants. There are 66 different known gibberellins, which are synthesized (Eq. 22-5) in mature leaves and are transported downward. Their very active effect in stimulating RNA synthesis in dwarf varieties of vegetables suggests that gibberellins also serve as gene activators to promote RNA synthesis. A possible function in the **geotropic** response of plant roots is suggested by the presence of higher concentrations of glibberellins in the upper half of horizontal roots than in the lower half.[372] On the other hand, auxin has long been known to have a higher concentration on the lower side of the root,[371a] and it has been assumed to inhibit elongation (in contrast to stimulation in stems). Because they are structurally and in physical properties somewhat similar to sterols, gibberellins have been assumed to act by a steroid hormonelike mechanism. The brassinosterols, which are true sterols, may also be expected to bind to soluble receptors and to regulate transcription.

The cytokinins are isopentenyladenosine derivatives (Fig. 5-33), which may be hydroxylated or substituted in the 2 position by a methylthio group.[373,374] Cytokinins are synthesized in roots and translocated upward to other parts of the plant. They may originate in part from degradation of cytokinin-containing tRNA, but there is evidence that they may also be synthesized independently.[374–376] The N^6-isopentenyladenine in tRNAs is generated by transfer of the isopentenyl group from Δ^2-isopentenyl pyrophosphate. The role of cytokinins may be at the level of gene transcription, but it has been difficult to identify signaling sequences.[368] Their hormonal influence on plants appears to be independent of their function in tRNA. The most striking effect of cytokinins in solution is on differentiation of plant cells (Chapter 32).

The terpenoid abscisic acid (Fig. 22-4) is synthesized by degradation of a carotenoid precursor. It is formed by plants in response to *stress* from low temperature, high salinity, or drought. Abscisic acid appears to block the growth-promoting effects of hormones such as the gibberellins and cytokinins. It is sometimes regarded as a general gene repressor, which prepares plants for dormancy. Synthesis of abscisic acid occurs in response to the short day and long night pattern of the fall. It often opposes the action of gibberellins. The signaling pathway for abscisic acid apparently involves release of Ca^{2+} ions and formation of cyclic ADP-ribose (cADPR; p. 564). Among other effects this induces the closing of stomata in leaf surfaces.[368]

Ethylene not only hastens the ripening of fruit but also tends to promote senescence in all parts of plants. Its signaling mechanisms are the best-known for any plant hormone.[368,369] The synthesis and action of ethylene are discussed in Chapter 24, Section D,4. Ethylene is metabolized slowly in plants by oxidation to ethylene oxide. The latter is hydrolyzed to form ethylene glycol, which is metabolized further to CO_2 (Eq. 30-3).



(30-3)

A postulated flowering hormone, **florigen**, has not been isolated.[366] Flowering seems to be controlled by a variety of different hormonal effects.[369] Jasmonic acid (jasmonate) and salicylic acid act as plant defense signals.[377,378] Salicylic acid activates a large number of transcription factors, which induce resistance to a variety of pathogenic organisms, a response referred to as **systemic acquired resistance**. See also Chapter 31, Section G. Among other compounds synthesized as part of the systemic acquired resistance of plants are proteins known as **systemins**.[382] Initially discovered in tomatoes, systemins have been discovered in more than 100 other species of plants. Jasmonic acid emitted from tomato plants acts as a pheromone that attracts wasps to attack caterpillars that feed on the tomato plants.[379] In fact, wounding by herbivores may stimulate emission of a variety of volatile compounds that may attract predators to the attacking herbivores.[380,381]

Many other compounds influence plant growth. Among them are the vitamins, thiamine, pyridoxine, and nicotinic acid, which are synthesized in the leaves and transported downward to the roots. Since they promote growth of roots, they are sometimes referred to as **root growth hormones**. However, they are nutrients universally needed by cells. Various compounds secreted by other organisms can either stimulate or inhibit growth of a given plant. Some are powerful toxins. Others, such as the previously mentioned NOD factors, and evidently also the riboflavin degradation product **lumichrome** (Box 15-B), are beneficial.[383] These **plant growth regulators** may be produced by other plants, by microorganisms, or by fungi.[384,385] Much use is made in agriculture of synthetic growth regulators.

10. Secretion of Hormones

In Chapter 11 the effects of binding of hormones to cell surface receptors have been emphasized. Equally important are the mechanisms that control the secretion of hormones. The topic of exocytosis has been considered briefly in Chapter 8, Section C,6 and aspects of the Golgi in Fig. 20-8 and associated text. Both hormones and neurotransmitters are secreted by exocytosis of vesicles. Cells have two pathways for secretion.[386,387] The **constitutive pathway** is utilized for continuous secretion of membrane constituents, enzymes, growth factors, viral proteins, and components of the extracellular matrix. This pathway carries small vesicles that originate in the trans-Golgi network (TGN; Fig. 20-8). The **regulated pathway** is utilized for secretion of hormones and neurotransmitters in response to chemical, electrical, or other stimuli.

Many neurotransmitters are packaged into **small synaptic vesicles** ~50 nm in diameter. These may originate from large endosomes rather than from the Golgi. They are usually recycled and refilled repeatedly.[386] Secretion of hormones, and of some neurotransmitters, occurs via **large dense-core vesicles** of ~100 nm diameter. These originate from the TGN and are not recycled. They are prominent in chromaffin cells and other cells that secrete large amounts of a signaling molecule. Secretion of hormones and that of neurotransmitters have several common features. Indeed, hormones of the hypothalamus, neurohypophysis, and the adrenal medulla are secreted by specialized neurons. However, while hormones are often carried in the bloodstream, neurotransmitters are most often secreted into the very small volume of a single synapse. The exocytosis must occur very rapidly from a small number of SSVs.

A common feature, and also a puzzle, of vesicular signaling is the nearly universal response to calcium ions. Exocytosis is usually triggered by a rise in the concentration of Ca^{2+}, and most receptor signaling also leads to an increase in cytosolic Ca^{2+}.[388–391] The puzzle lies in the ability of cells to use a common mechanism for so many specific purposes. This topic is considered further in Section B,8. There are also many other factors that can control exocytosis. Recent evidence suggests that NO may play a role.[392]

B. Neurochemistry

The nervous system, a network of neurons in active communication, reaches its ultimate development in the 1.5 kg human brain.[149,393–396] Many invertebrates, such as leeches,[396a] crayfish, insects, and snails, have brains containing no more than 10^4 to 10^5 neurons,[396b,397,398] but the human brain contains ~10^{11}. Each of these neurons interconnects through **synapses** with hundreds or thousands of other neurons. The number of connections is estimated to be as many as 60,000 with each Purkinje cell of the human cerebellum. There may be many more than 10^{14} synapses in the human brain.[399,400]

In addition to neurons, the brain contains 5–10 times as many **glial** cells of several types. The neuroglia occupy 40% of the volume of brain and spinal cord in the human. Some glial cells seem to bridge the space between neurons and bloodcarrying capillaries. Others synthesize myelin. Some are very irregular in shape.

1. The Anatomy and Functions of Neurons

Although neurons have many shapes and forms, a common pattern is evident.[400a] At one end of the elongated cell (Fig. 30-8) is a series of **dendrites**, thin fibers often less than 1 µm in diameter. The ends of the dendrites form synapses with other neurons and act as receivers of incoming messages. Additional messages come into synapses on the **cell body**, while the **axon** serves as the output end of the cell. The axon, a long fiber of diameter 1–20 µm, is also branched. As a consequence, the nervous system contains both highly branching and highly converging pathways. Many of the axons are wrapped in a myelin sheath (Fig. 30-9; pp. 390 and 1767).

The ends of the fine nerve fibers are thickened to form the **synaptic knobs**, which make synaptic contacts with dendrites on cell bodies of other neurons. In most instances the arrival of a nerve signal at the **presynaptic** end of a neuron causes the release of a transmitter substance (neurohormone). The transmitter passes across the 10–50 nm (typically 20 nm) **synaptic cleft** between the two cells and induces a change in the electrical potential of the **postsynaptic** membrane of the next neuron (Fig. 30-10).[149,401] Excitatory transmitters usually cause **depolarization** of the membrane. By this we mean that the membrane potential, which in a resting neuron is –50 to –70 mv (Chapter 8), falls to nearly zero often as a consequence of an increased permeability to Na^+ and a resultant inflow of sodium ions. The resulting **postsynaptic potential** (really a drop in the potential difference) is propagated to the cell body and axon and under appropriate circumstances may initiate an **action potential**. This is a narrow spikelike region of depolarization that travels down the axon at a constant velocity and with undiminished intensity (Fig. 30-11).

A characteristic of many neurons is an *all-or-none response* or firing. An action potential passes down the axon only if there is sufficient depolarization. In general, a stimulus must reach a neuron through *more than one synapse* before the neuron will fire. Furthermore, neurons are often *inhibitory*, releasing transmitters that counter the excitatory synapses and tend to prevent firing. Inhibition is important in damping out small excitations; thus sharpening the response of the nervous system toward strong stimuli. Another characteristic of basic importance to the operation of the brain is that neurons fire at longer or shorter intervals depending upon the strength and duration of the stimulus. The stronger the stimulus to a given neuron, the more rapid the train of spikes that passes down the axon. Thus, the brain functions to a large extent in

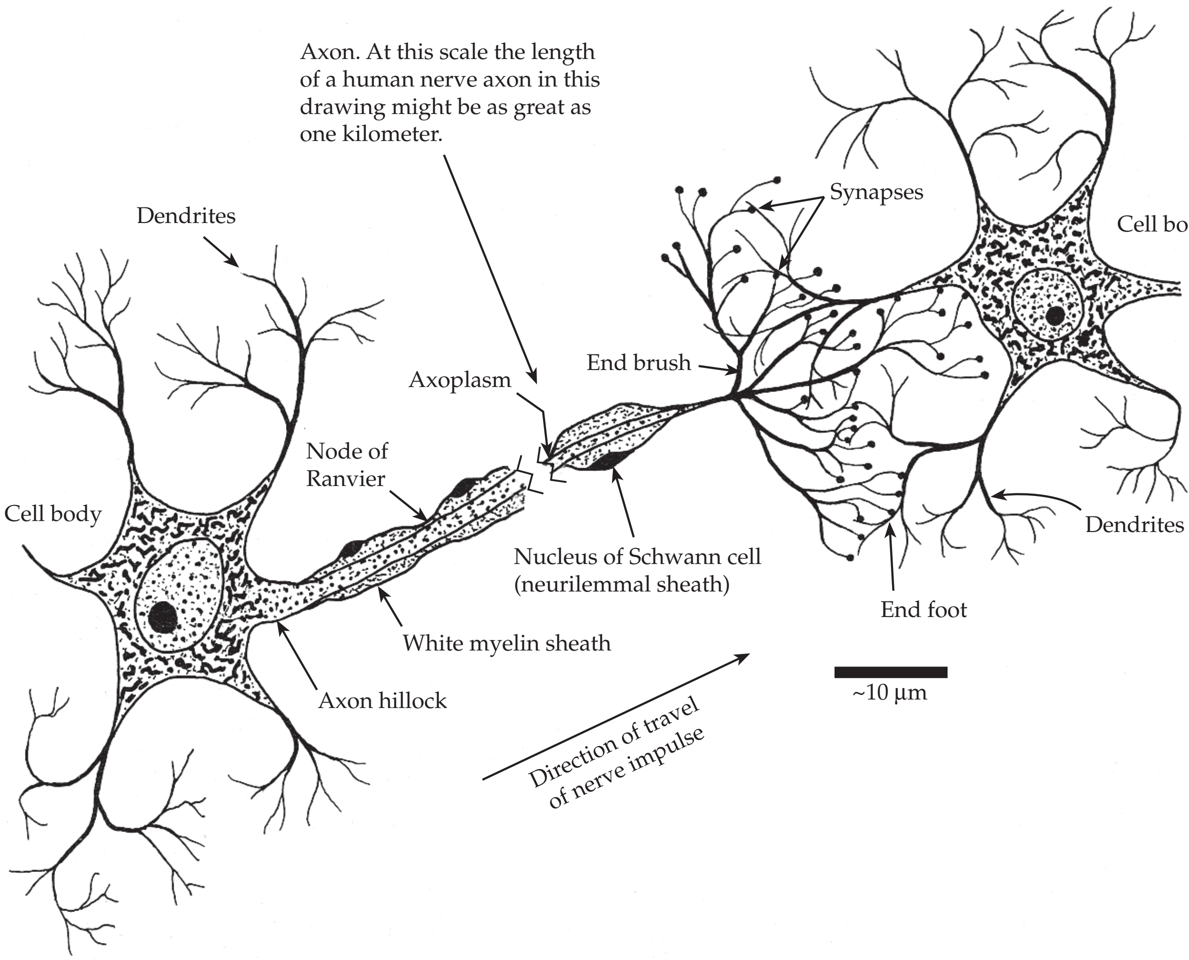

Axon. At this scale the length of a human nerve axon in this drawing might be as great as one kilometer.
Dendrites
Synapses
Cell body
End brush
Axoplasm
Node of Ranvier
Cell body
Nucleus of Schwann cell (neurilemmal sheath)
Dendrites
End foot
White myelin sheath
Axon hillock
~10 µm
Direction of travel of nerve impulse

Figure 30-8 Schematic drawing of a neuron (after Brand and Westfall,[401] p. 1192).

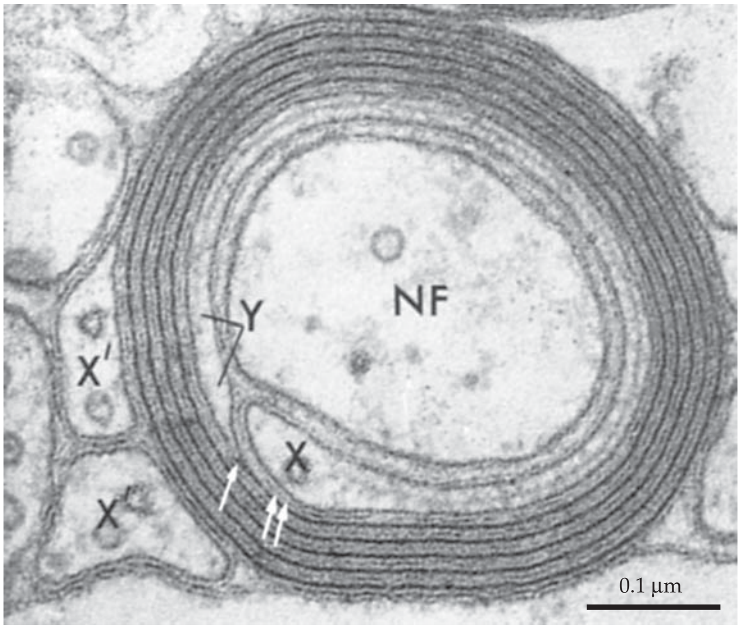


Figure 30-9 Micrograph of a section through an axon of a neuron from rat brain. The structure of the myelin sheath can be seen clearly. The growing lip of cytoplasm (X) from a neuroglial cell is advancing around the axon process (NF) and insinuating itself into the space between the plasma membrane of the axon and the membrane that limits the thin layer of cytoplasm (Y) left behind by the growing lip during its previous turn. This cytoplasmic layer disappears as the inner leaflets of its plasma membrane fuse to form the major dense line of the myelin sheath. This process is occurring at the point indicated by the single arrow. The outer leaflet of the plasma membrane surrounding the lip fuses with its own outer leaflet laid down on the previous turn. The two outer leaflets thus give rise to the less dense intermediate line of the sheath (double arrow). The cell body from which the investing cytoplasmic sheet originated cannot be seen in this micrograph, but cytoplasm within the lateral margins of the sheet does appear (X′). The micrograph, by A. Hirano and H. M. Dembritzer, originally appeared in *J. Cell Biol.* **34**, 555 (1967), where a more complete explanation of myelin sheath formation is provided. Figure copied from Porter and Bonneville.[402] Courtesy of Mary Bonneville.

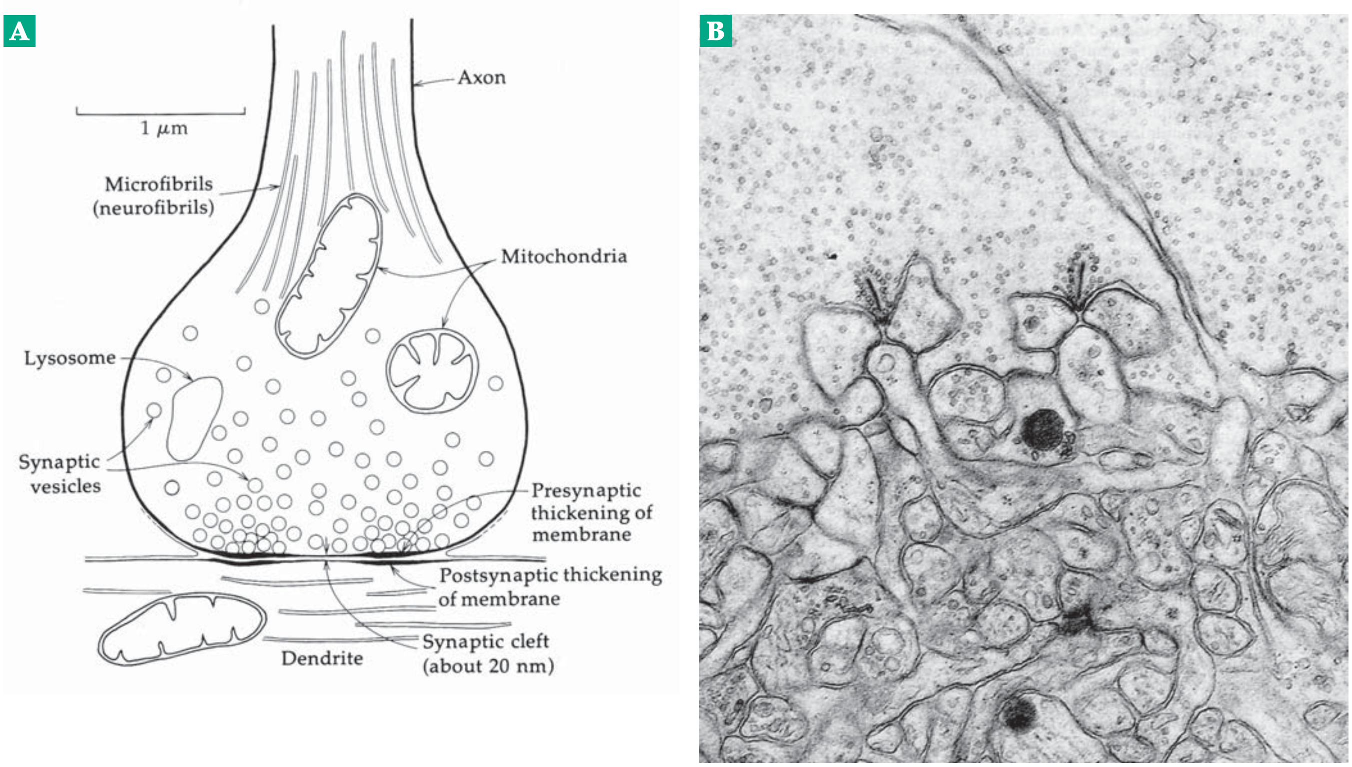

A
B
Axon
1 µm
Microfibrils (neurofibrils)
Mitochondria
Lysosome
Synaptic vesicles
Presynaptic thickening of membrane
Postsynaptic thickening of membrane
Synaptic cleft (about 20 nm)
Dendrite

Figure 30-10 (A) Schematic drawing of a synapse. (B) Electron micrograph showing the synaptic junctions in the basal part (pedicle) of a retinal cone cell of a monkey.[403] Each pedicle contains synaptic contacts with 12 triads, each made up of processes from a bipolar cell center that carries the principal output signal and processes from two horizontal cells that also synapse with other cones. A ribbon structure within the pedicle is characteristic of these synapses. Note the numerous synaptic vesicles in the pedicle, some arranged around the ribbon, the synaptic clefts, and the characteristic thickening of the membranes surrounding the cleft (below the ribbons). Micrograph courtesy of John Dowling.

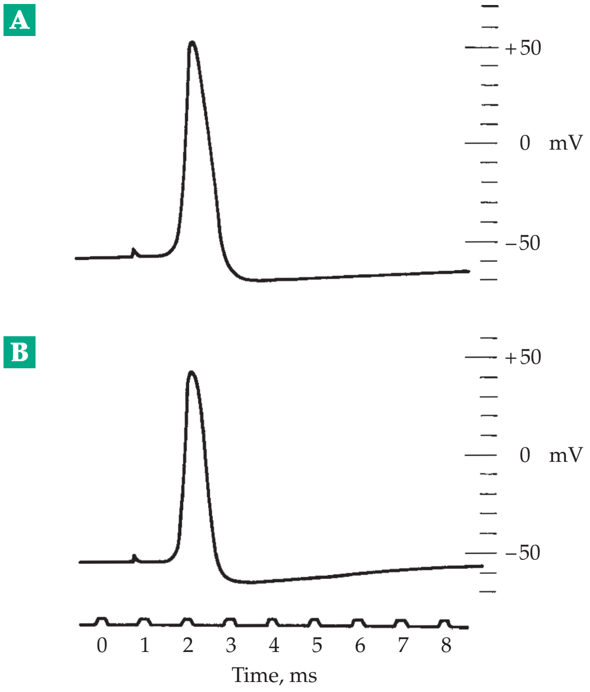


Figure 30-11 (A) Action potential recorded with internal electrode from extruded axon filled with potassium sulfate (16°C). (B) Action potential of an intact axon, with same amplification and time scale (18°C). The voltage scale gives the potential of the internal electrode relative to its potential in the external solution with no correction for junction potential. From A. Hodgkin, *Conduction of Nervous Impulses*, 1964. Courtesy of Charles C. Thomas, Publisher, Springfield, Illinois.

decoding trains of impulses. The frequency of the impulses from neurons varies from a few per second to a maximum of about 200 s^{-1} in most nerves (up to 1600 s^{-1} in the Renshaw cells of the spinal cord). The maximum frequency is dictated by the refractory period of ~1 ms (Section B,3).

Although the concepts of neuronal function outlined in the preceding paragraphs have been accepted for many years, more recent discoveries require that they be modified somewhat. Dendrites seem to be able to transmit information as well as to receive it. Furthermore, while information is certainly transmitted long distances by spike action potentials, shorter neurons and dendrites may communicate extensively by exchange of chemicals through low resistance gap junctions, also called **electric synapses** (Chapter 1). Small changes in membrane potential transmitted through these junctions may alter the behavior of adjacent neurons. Chemical transmitters do not always have an electrical effect on postsynaptic neurons but may influence metabolism or gene transcription.

2. Organization of the Brain

The anatomy of the brain is quite complex, and only a few terms will be defined here. The **cerebrum**, which is made up of two hemispheres, accounts for the largest part of the brain. The deeply folded outermost layer, the **cerebral cortex**, consists of **gray matter**, a mass of cell bodies, and fine unmyelinated nerve fibers. Beneath this lies a layer of **white matter** made up of myelin-covered axons connecting the cerebral cortex with other parts of the brain. The two cerebral hemispheres are connected by the **corpus callosum**, a band of ~2 x 10^8 nerve fibers. Remarkably, these fibers can be completely severed with a relatively minimal disruption of the nervous system. In the past the corpus callosum was sometimes cut to control almost incessant epileptic seizures that could not be prevented by drugs. The "split-brain" patients suffered relatively little disability as long as both eyes functioned normally. Studies of these patients provided some insights into the differing functions of the two hemispheres of the cerebrum.[395]

Deeper in the cerebrum lie the **basal ganglia,** which include the caudate, lenticular, and amygdaloid nuclei. The lenticular nuclei are further divided into putamen (an outer portion) and the globus pallidus. The putamen and caudate nuclei together are known as the **striatum** (Fig. 30-12). The lower lying subthalamic nuclei and substantia nigra are sometimes also included in the basal ganglia.

The outer parts of the cerebrum, including the basal ganglia, make up the telencephalon. Deep in the center of the brain is the diencephalon consisting of the **thalamus** (actually two thalami), **hypothalamus**, **hypophysis** (Figs. 30-1, 30-13), and other attached regions. A major structure at the back of the brain is the **cerebellum**. Like the cerebrum, its cortex is highly folded. The 30 billion neurons of the cerebellum are organized in a highly regular fashion.[393,404] The interconnections of the seven types of neurons present in this part of the brain have been worked out in fine detail.

The basal part of the brain or **brain stem** consists of the medulla oblongata and the pons. While the bulk of the tissue consists of myelinated nerve tracts passing into the spinal cord, synaptic regions such as the olivary nucleus are also present.

The brain, which must function in a chemically stable environment, is protected by a tough outer covering, the **arachnoid membrane**, and by the **blood–brain barrier**[406,407] and the **blood–cerebrospinal barrier**. Both of these barriers consist of tight junctions similar to those seen in Fig. 1-15A. They are formed between the endothelial cells of the cerebral capillaries and between the epithelial cells that surround the capillaries of the **choroid plexus**. The choroid plexus consists of capillary beds around portions

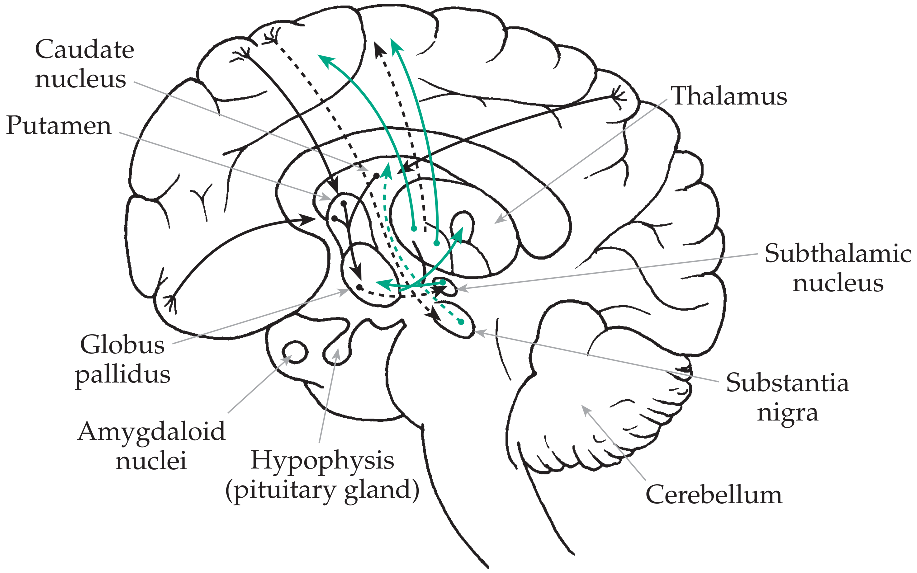

Caudate nucleus
Putamen
Thalamus
Subthalamic nucleus
Globus pallidus
Substantia nigra
Amygdaloid nuclei
Hypophysis (pituitary gland)
Cerebellum

Figure 30-12 Diagram illustrating some of the major interconnections of the “extrapyramidal system” of the brain. Arrows indicate major direction of projections. The nigrostriatal (substantia nigra to striatum) and related neuronal pathways are indicated with dashed lines. After Noback and Demarest,[405] pp. 182 and 183.

of the fluid-filled **ventricles** deep in the interior of the brain. They serve as a kind of “kidney” for the brain assisting in bringing nutrients in from the blood and helping to keep dangerous compounds out.[406]

3. Neuronal Pathways and Systems

Consider a message originating with a nerve receptor in the skin or in another sense organ. A nerve signal passes via a **sensory neuron (afferent fiber)** upward toward the brain. It may pass through two or more synapses (often through one in the spinal cord and one in the thalamus) finally reaching a spot in the sensory region of the cerebral cortex. From there the signal in modified form spreads through the **interneurons** of virtually the entire cortex. In each synapse, as well as in the cortex, the impulse excites inhibitory fibers that dampen impulses flowing through adjacent fibers. Likewise, if a given impulse is not strong enough, it will itself be inhibited before reaching the cortex. Among the important sensory neurons are those from the seven million cone cells and 100 million rod cells of the eye. The nerve signals pass out of the retina by way of a million axons from retinal ganglion cells reaching, among other parts of the brain, the **visual cortex** (Fig. 30-14).[408]

The neuronal events that occur within the cerebral cortex are extraordinarily complex and little understood.[409] In what way the brain is able to initiate voluntary movement of muscles is obscure. However, it is established that the signals that travel out of the brain down the **efferent fibers** to the muscles arise from large **motor neurons** of the **motor cortex**,[410] a region that extends in a band across the brain and adjacent to the sensory cortex (Fig. 30-14). The axons of these cells form the **pyramidal tract** that carries impulses downward to synapses in the spinal cord and from there to the **neuromuscular junctions**. These are specialized synapses at which acetycholine is released, carrying the signal to the muscle fibers themselves. Passing over the cell surface and into the

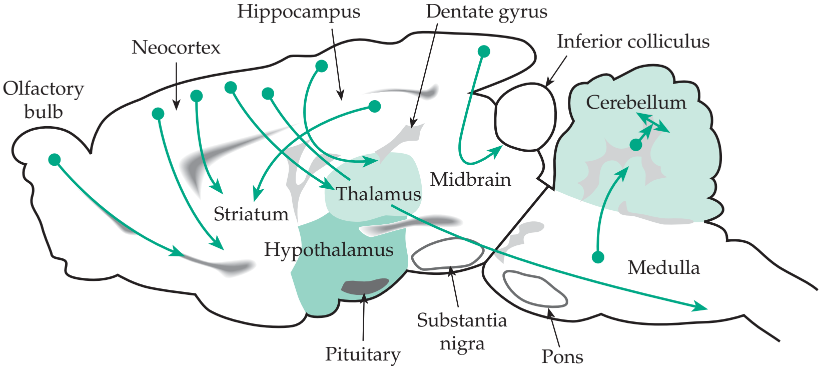

Hippocampus
Dentate gyrus
Neocortex
Inferior colliculus
Olfactory bulb
Cerebellum
Midbrain
Thalamus
Striatum
Hypothalamus
Medulla
Substantia nigra
Pituitary
Pons

Figure 30-13 Section through a rat brain. This brain, which has been very widely used in neurochemical studies, appears superficially to be quite different from the human brain (Fig. 30-1), which is characterized by its large cerebral cortex. However, basic pathways are the same. Some major pathways for glutamate-secreting (glutamatergic) neurons are marked by arrows. Most of these originate in the neocortex (outer layers of the cerebral cortex) and the hippocampus. From Nicholls.[149] Courtesy of David G. Nicholls.

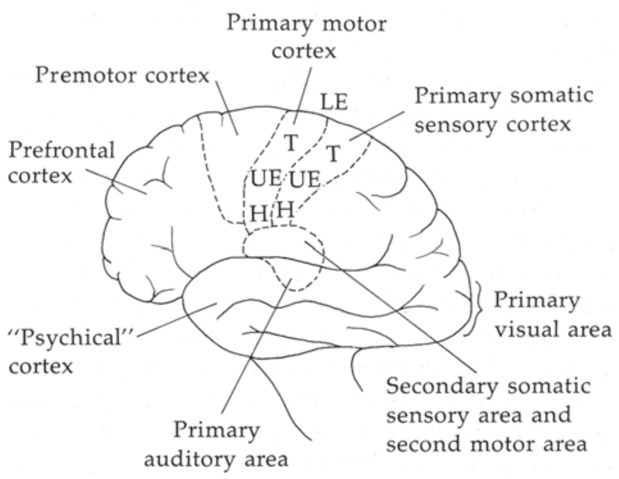


Figure 30-14 The location of several functional areas of the cerebral cortex. The representation of body parts on the primary motor and somatic sensory cortices include the head (H), upper extremity (UE), trunk (T), and lower extremity (LE). After Noback and Demarest,[405] p. 193.

T tubules (Chapter 19, Section B,4; Fig. 19-21), a wave of depolarization initiates the release of calcium and muscular contraction.

At the same time that the motor neurons send signals to the muscles, branches travel into other parts of the brain including the olivary nuclei, which send neurons into the cerebellum. The cerebellum acts as a kind of computer needed for fine tuning of the impulses to the muscles. Injury to the cerebellum leads to difficulty in finely coordinated motions. Input to the Purkinje cells arises from the climbing fibers, which originate in the inferior olive of the brain stem. Each climbing fiber activates a single Purkinje cell, but the dendrites of each Purkinje cell also form as many as 200,000 different synapses with parallel fibers that run across the cortex of the cerebellum (Fig. 30-15). The parallel fibers receive input from many sources via a complex series of mossy fibers and granule cells and influence the firing of the Punkinje cells. The output from the Purkinje cells is entirely inhibitory. It is transmitted via synapses in the cerebellar nuclei to neurons that lead back to the cerebral cortex, into the thalamus, and down the spinal cord.[411] The pathway to the cortex completes an inhibitory feedback loop, of which there are many in the nervous system. For details see Llinás[404] and Nicholls.[149]

In addition to the **somatic motor system** that operates the voluntary (striated) muscles via the pyramidal tract, there is the **autonomic system,** which controls the involuntary (smooth) muscles, glands, heartbeat, blood pressure, and body temperature. This system has its origins in both the cerebral cortex and hypothalamus. It is subdivided into two systems, the **sympathetic** and **parasympathetic** systems, which are anatomically distinct. The sympathetic system is geared to the fight and fright reactions. Its **postganglionic fibers** (those below the ganglia in the spinal cord) liberate norephinephrine (noradrenaline) and include the adrenal medulla, which consists of specialized neurons, the **chromaffin** cells. The parasympathetic system has to do more with homeostasis and maintenance of body systems. Biochemically it is characterized by the release of acetylcholine as a transmitter substance.

The hypothalamus, a four gram portion of the brain, receives a great deal of biochemical attention because of its function in the autonomic nervous system, in homeostasis, and in endocrine secretion. Its liberation of neurohormones that stimulate the hypophysis has already been considered in Section A,3. The hypothalamus is also involved in the regulation of the body temperature, of water balance, and possibly of glucose concentration.

Two other systems of importance in the brain are the **reticular system** and the **limbic system**. The former is the mediator of the sleep–wake cycle and is responsible for characteristic waves in the electroencephalogram. The limbic system is the mediator of **affect** or mood and of **instincts**. It is anatomically complex with centers in the amygdala, other subcortical nuclei, and the limbic lobe of the cortex. The limbic cortex forms a ring lying largely within the longitudinal fissure between the two hemispheres. It includes the olfactory cortex, the **hippocampus**, a region associated with formation of conscious memories, and other evolutionarily older regions of the cerebral cortex. Within the limbic lobe are the **pleasure centers**. When electrodes are implanted in these regions, animals will repeatedly push levers that are designed to electrically stimulate these centers. There are also **punishing centers,** whose stimulation causes animals to avoid further stimulation.

4. The Propagation of Nerve Impulses

Although the chemical basis of the conduction of nerve impulses via an action potential is not entirely clear, the electrical events have been described with precision. If the permeability of a membrane toward sodium ions is increased in a local region, sodium ions flow through the membrane into the cell neutralizing the negative charge inside and depolarizing the membrane. Such depolarization leads to propagation of an electrical signal of diminishing intensity over the surface of the membrane in a manner analogous to the flow of electrical current along a coaxial cable. It is thought that local increases in Na^+ permeability of the plasma membrane often trigger nerve impulses. Other

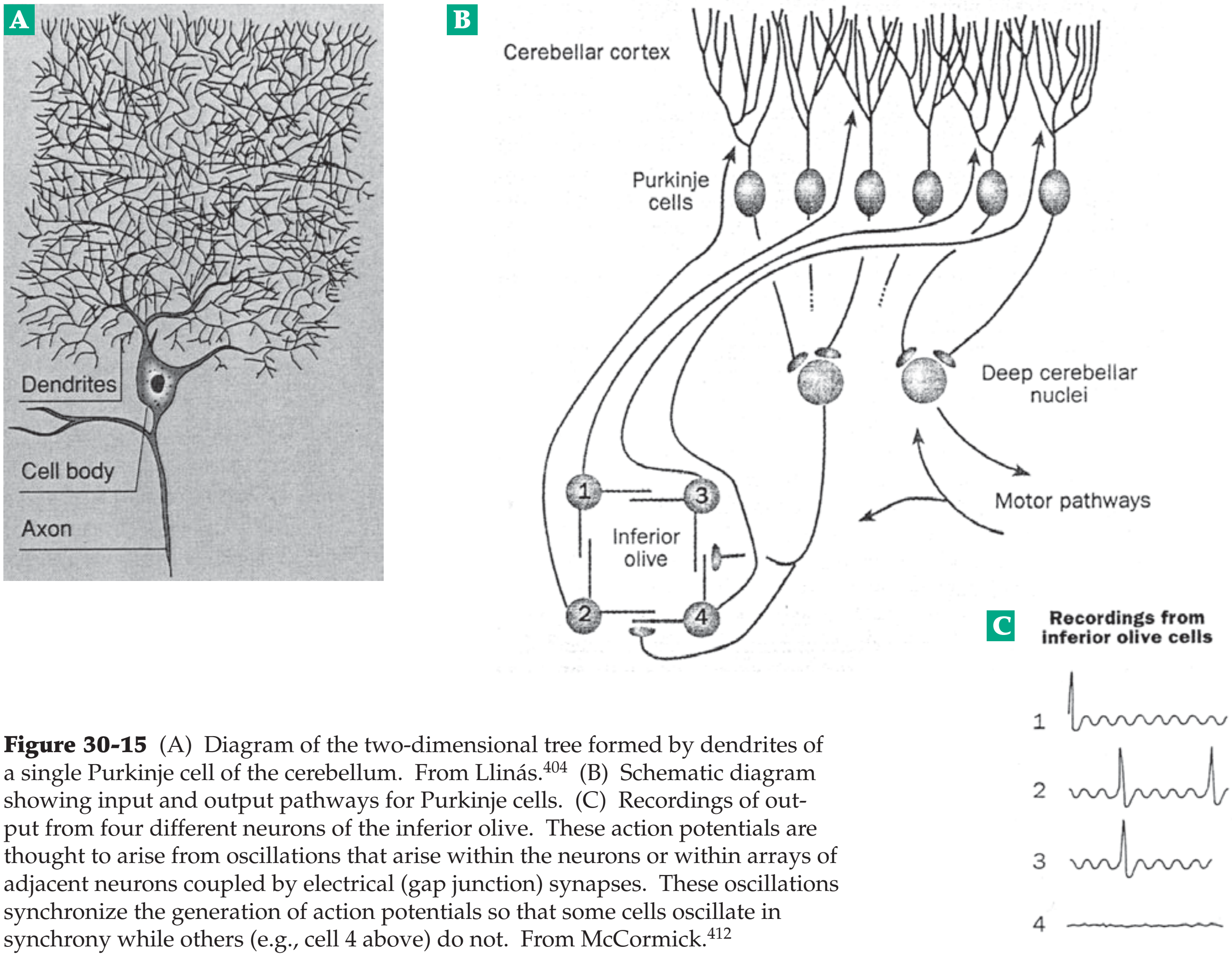


Figure 30-15 (A) Diagram of the two-dimensional tree formed by dendrites of a single Purkinje cell of the cerebellum. From Llinás.[404] (B) Schematic diagram showing input and output pathways for Purkinje cells. (C) Recordings of output from four different neurons of the inferior olive. These action potentials are thought to arise from oscillations that arise within the neurons or within arrays of adjacent neurons coupled by electrical (gap junction) synapses. These oscillations synchronize the generation of action potentials so that some cells oscillate in synchrony while others (e.g., cell 4 above) do not. From McCormick.[412]

ions such as Ca^{2+} may also play a role. While the kind of passive transmission of electrical signals that results from a local depolarization of the membrane is suitable for very short nerve cells, it cannot be used to send signals for long distances. Most nerve axons employ the more efficient action potential. This is an impulse that passes along the axon and for a short fraction of a second (~0.5 ms in mammalian nerves) changes the membrane potential in the characteristic way shown in Fig. 30-11. Initially, the negative potential of 50–70 mV drops rapidly to zero and then becomes positive by as much as 40–50 mV, after which it returns to the resting potential. The remarkable thing about the action potential is that it is propagated down the axons at velocities of 1–100 m / s without loss of intensity.

To establish the chemical basis of the action potential, A. L. Hodgkin and A. F. Huxley in the 1950s devised the **voltage clamp**, a sophisticated device by which the transmembrane current can be measured while using a feedback mechanism to fix the membrane potential at a preselected value.[413–417] Using the voltage clamp the membrane conductance could be measured as a function of the membrane potential and of time. It was found that immediately after a decrease in membrane potential was imposed with the voltage clamp, the permeability of the membrane toward sodium ions rose rapidly. Since an increased sodium ion permeability automatically leads to depolarization in an adjacent region of the membrane, a self-propagating wave is established and moves down the axon. The voltage clamp studies also revealed that after a fraction of a millisecond the permeability to potassium ions also increases. At the same time the sodium ion permeability decreases again, and the normal membrane potential is soon reestablished. However, during an **absolute refractory period** of ~0.5 ms no other nerve impulse can be passed. The sequence of events during passage of the nerve impulse can be described as the opening of sodium channels followed by the opening of potassium channels, and then by a closing of the channels in the same sequence. The results of these investigations led Hodgkin and Huxley to propose equations that quantitatively describe the action potential and that predict the observed conduction velocities and other features of nerve impulses.

A special feature of nerves that are designed to transmit impulses very rapidly is the presence of the wrapping of **myelin** (Fig. 30-9). As can be seen in this figure, the extracellular surfaces of the consecutive wraps bind tightly together, and the cytoplasm of the cell interior is squeezed out to form the compact myelin sheath.[418] Mutations in the integral membrane proteolipid protein (p. 401) are associated with a variety of defects in myelin formation. Some of these are severe, for example, leading to loosely wrapped myelin.[419,420] The proteolipid protein is encoded by an X-linked gene. The most abundant protein in peripheral nerve myelin is the integral membrane **peripheral myelin glycoprotein P_0**. It is encoded by an autosomal gene for which 29 known defects account for a variety of human diseases,[421–422a] including an autoimmune inner ear disease.[423] The extracellular domain of P_0, like many other cell adhesion molecules (p. 407), has a structure related to that of immunoglobulins. Four molecules of P_0, each of which carries a single immunoglobulin domain, associate via these domains in a kind of square donut that protrudes from the outer cell surface. There it can interact with four similar donuts from the apposed cell surface, zipping up the cell–cell interface by a kind of Velcro action.[422,424,425] Protein P_0 accounts for 50% of the total protein of peripheral myelin, but the **myelin basic protein**, which constitutes 20% of the total protein, is also essential.[426] This protein exists as a variety of forms that arise from differential splicing of its mRNA and extensive posttranslational modification. Deimination of argine side chains to form citrulline residues has been associated with development of the autoimmune disease **multiple sclerosis**.[427,428] **Peripheral myosin protein 22** is a 160-residue polypeptide with four membrane-spanning helices. It accounts for 2–5% of the myelin protein and is the site of defects that cause the demyelinating **Charcot–Marie–Tooth disease** and other serious human diseases.[428a,b]

The axon is effectively insulated from the surrounding medium by the myelin sheets except for special regions, the **nodes of Ranvier**, which lie at 1- to 2-mm intervals along the nerve. The nerve impulse in effect jumps from one nerve to the next. This **saltatory conduction** occurs much more rapidly (up to 100 m / s) than conduction in unmyelinated axons. It depends upon Na^+ and K^+ channels that are concentrated in the nodes of Ranvier.

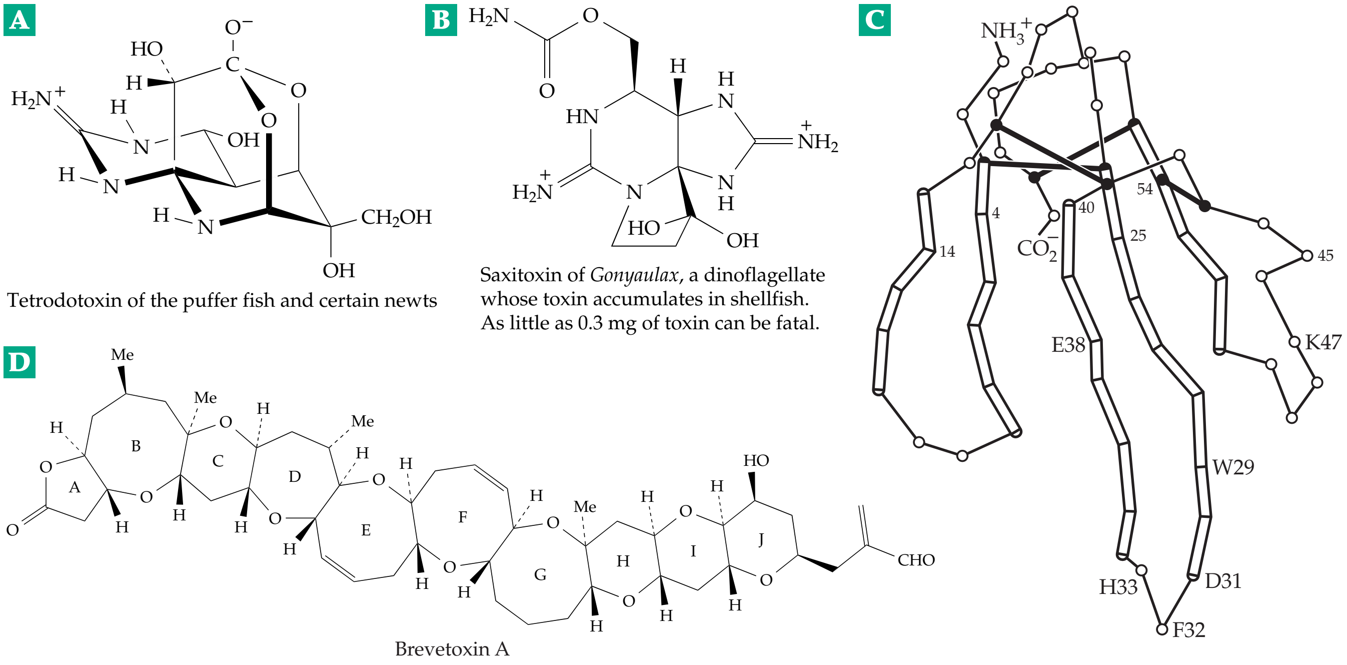
5. Ion Conducting Channels

What is known about the channels through which Na^+ and K^+ flow during nerve excitation? That the channels for the two ions are separate was shown by the fact that **tetrodotoxin** (found in the puffer fish)[429,429a] and **saxitoxin** of dinoflagellates, as well as scorpion toxins (see Fig. 30-16), exert their toxic action by blocking the Na^+ channels while having no effect upon conductance for K^+. At the same time the K^+ channels can be blocked by certain quaternary ammonium salts. Since the binding constants for the toxins are high ($K_f \sim 3 \times 10^8$ M^{-1} for tetrodotoxin), it is possible to titrate the sodium channels. The number is usually quite small, about 10–400 Na^+ channels / μm^2 of surface[430] (the same surface area contains 2×10^6 phopholipid molecules). However, membranes in the nodes of Ranvier of mammalian nerve fibers[431] contain ~12,000 channels / μm^2. Note that the ion channels described here are not the same as those in the ion pump, i.e., the Na^+,K^+-ATPase (Fig. 8-25). In some neurons the number of conduction channels for Na^+ appears to be ten times less than the number of pumping channels, i.e., of Na^+,K^+-ATPase.[432,432a]

Since the number of ion-conducting channels is small, the rate of sodium passage through the open channels must be extremely rapid and has been estimated as ~10^8 ions / s.[433] This is within an order of magnitude of the diffusion-limited rate (Eq. 9-30). On this basis it is clear that the channels cannot act by means of ionophoric carriers but form pores that can be opened and closed (**gated**) in response to changes in the membrane potential. They are **voltage-sensitive ion channels**.[433,434] The channels are selective for specific ions and the selectivity parallels that of sites in some cation-exchange resins such as those containing carboxylate groups. This suggested that the inside surface of the channel might contain one or more carboxylate groups from protein side chains as well as other polar groups. A Na^+ ion approaching the channel entrance might exchange some of its hydration sphere for ligands from the channel surface. The differing affinity of the “ion exchange” sites for various cations could ensure that it is predominately Na^+ that passes through the channel. Anions could be excluded by electrostatic repulsion. Recent structural studies have allowed these speculations to be replaced with experimental findings as described in the following paragraphs. They have revealed that the selectivity mechanism are similar for Na^+ and Ca^{2+} channels.

The sodium ion channel of the electric eel.

Making use of the binding of radioactively labeled specific toxins to identify them, the subunits of the sodium channel proteins were purified from several sources including the electrical tissue of the electric eel *Electrophorus electricus*,[437–439] heart and skeletal muscle, and brain.[440–441b] In all cases a large ~260-kDa glycoprotein, which may be 30% carbohydrate, is present. The saxitoxin-binding protein from rat brain has two additional 33–36 kDa subunits with a stoichiometry of $\alpha\beta_1\beta_2$. The *Electrophorus* α subunit consists of 1820 residues,[437] while rat brain contains α proteins of 2009

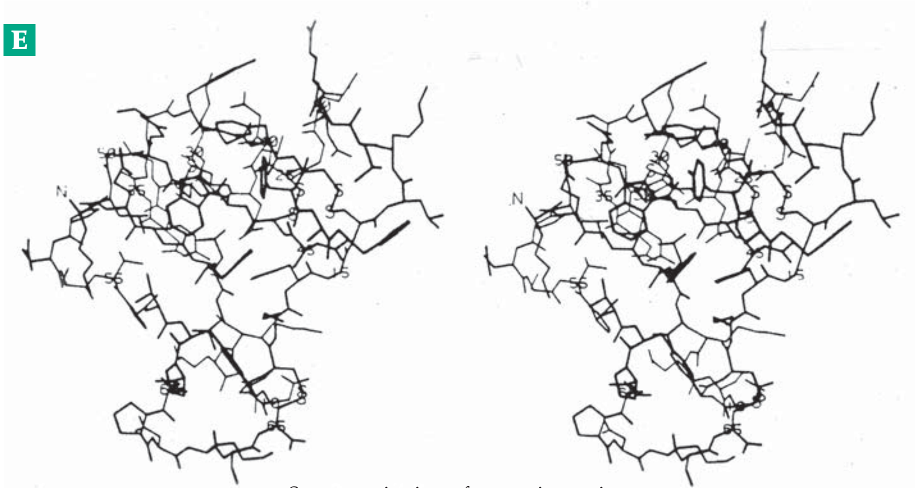


Tetrodotoxin of the puffer fish and certain newts

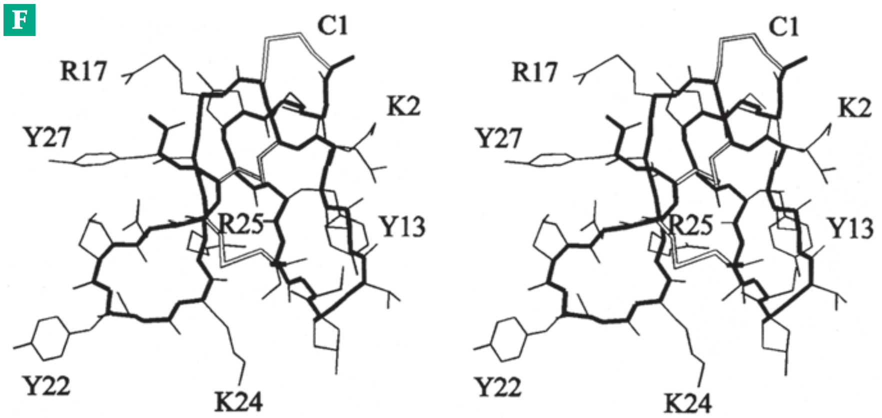
Saxitoxin of *Gonyaulax*, a dinoflagellate whose toxin accumulates in shellfish. As little as 0.3 mg of toxin can be fatal.

Brevetoxin A

Structure of erabutoxin b. Many snake venom neurotoxins have similar structures with conserved disulfide bridge positions. See Hatanaka, H. *et. al.*[513]



Stereoscopic view of a scorpion toxin



ω Conotoxin, blocks presynaptic voltage-regulated Ca^{2+} channels

Figure 30-16 Structures of some neurotoxins that affect ion channels. Other neurotoxins include the Na^+, K^+-ATPase inhibitor ouabain (Fig. 22-12), batrachotoxin (Fig. 22-12), and picrotoxin (Fig. 22-4). The structure of a scorpion toxin is from Almassy *et al.*,[494a] that of ω conotoxin is from Pallaghy *et al.*,[435] and that of brevetoxin is redrawn after Shimizu *et al.*[436]

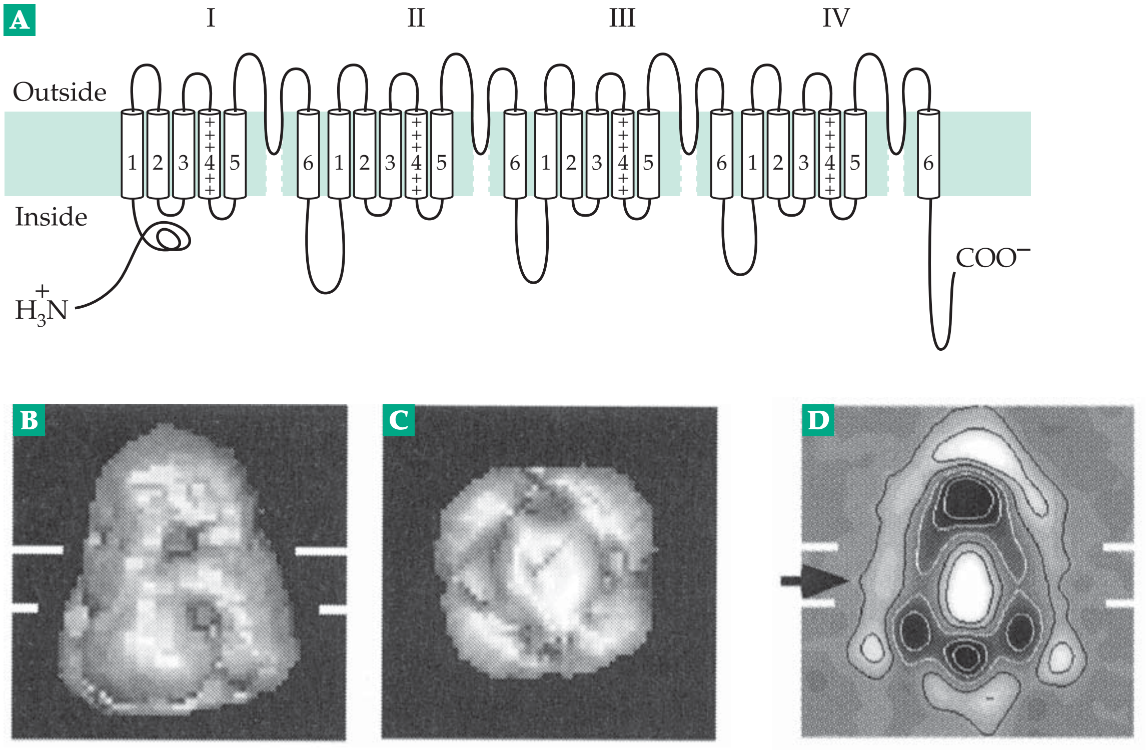


Figure 30-17 (A) Two-dimensional map of the ~260-kDa α subunit of the voltage-gated Na^+ channel from the electric eel *Electrophorus electris*.[438,441] (B) Image of the sodium channel protein obtained by cryo-electron microscopy and image analysis at 1.9 nm resolution. In this side view the protein appears to be bell-shaped with a height of ~13.5 nm, a square bottom (cytoplasmic surface) ~10 nm on a side, and a hemispherical top with a diameter of ~6.5 nm. (C) Bottom view of the protein. (D) Axial section which cuts the bottom, as viewed in (C), approximately along a diagonal. From Sato *et al.*[438] Notice the cavities (dark) and domain structures (light). The black arrow marks a constriction between upper (extracelllar) and lower (cytoplasmic) cavities. White lines indicate approximate position of the lipid bilayer. From Sato *et al.*[438] Courtesy of Chikara Sato.

---

and 2005 residues, respectively, for Na^+ channels designated I and II.[441] In fact, mammals contain ten distinct Na^+ channel genes.[442] In every case the channel proteins contain four consecutive homologous sequences of about 300 residues apiece. Within these the hydro-pathy plots (see Fig. 2-30) suggest that each homology region forms six membrane-spanning helices as shown in Fig. 30-17A.[441] The four sets may then fold together into a square arrangement that provides a pore somewhat familar to that of the voltage-gated K^+ channel (see Fig. 30-18). The three-dimensional structure of the sodium channel protein, based on cry-electron microscopy, appears to be complex. The central channel may resemble that of Fig. 30-18, but there also seem to be smaller peripheral channels (Fig. 30-17).[438] Bacteria also contain Na^+ channels but they are tetramers of smaller subunits, resembling in this respect bacterial K^+ channels (Figs. 30-18).[442a,b]

How do the "gates" to ion channels open? Presumably some part of the channel protein senses the change in potential and undergoes an appropriate alteration in conformation that opens the gate.[434] The current carried by the ions flowing out through a small number or even a single channel can be measured with tiny **patch electrodes** having openings ~1 μm^2 in area. These are pressed against the nerve membrane, where they form a tight seal. With such a small patch of membrane surface the electrical noise level is low, and it is possible to measure the conductance of the pore.[149,443] From such measurements it was found that a single pore can allow $>10^8$ ions to pass through in one second. Another thing that is apparently measured with patch electrodes is a small **gating current**, which precedes the opening of the channels by ~0.1 ms. This has been interpreted as a flow of ~6 charges across the membrane or the movement of a larger number of dipoles needed to open the gate. One possibility is that a loss of the electrical field from the surface charges on the bilayer induces a rearrangement of charges on protein side chains within the bilayer or induces changes in interactions between two or more dipoles. Such changes could trigger conformation alterations within the proteins, allowing the channel to switch from open to closed.

Recordings with single channels indicate that after a sodium channel is open for a random length of time it spontaneously closes and passes into a third state, an "inactive" state from which it cannot reopen during the refractory period. After the membrane is repolarized it can function again.[444a]

Calcium ion channels. Immediately after the Na^+ pores open as a result of membrane depolarization, voltage-sensitive Ca^{2+} channels also open. These allow a rapid influx of Ca^{2+}, which can trigger many processes including the secretion of neurotransmitters within the synapses.[434,444] There are several types of voltage-sensitive Ca^{2+} channels.[444a,b] The most abundant type are specifically inhibited by dihydropyridines and are called **dihydropyridine-sensitive** or L-type channels.[434,445–445b] They are most numerous in the transverse tubular membranes of skeletal muscle where they appear to form a complex with the very large calcium release channels, the **ryanodine receptors** (Fig. 19-21 and associated discussion).[446,446a] These channels appear to have a structure similar to that of the Na^+ channels.[434] Calcium channels are also discussed on p. 422 and on pp. 1114-1115. Calcium ions play a central role in cell signaling and there are a large number of different calcium channels in bacteria, plants, and animals. Many of these are coupled to specific receptors.[445b,447] Some are involved in controlling intracellular stores.[447–449] Some release Ca^{2+} in response to mechanical movement and function in feeling, hearing, maintaining balance, and cardiovascular regulation. Plants sense wind and gravity, and microorganisms sense changes in osmotic pressure with the aid of these channel proteins.[450–452]

Potassium ion channels. Several types of K^+-selective cation channels have been recognized on the basis of electrophysiological and pharmacological studies.[149] More recently, the cloning of channel genes has permitted the study of the proteins by X-ray crystallography. The first structure determined[452a] was that of the *Streptococcus lividans* K^+ channel (designated KcsA; Fig. 8-21). There are three large structural families of K^+ channel proteins.[453–455a] One group consists of voltage-regulated (K_v) channels, such as those involved in the action potential of neurons. Like the *S. lividans* channel, they are tetramers whose predicted structure contains six transmembrane helices per subunit with a pore-forming loop (P region) between helices 5 and 6. This is just what is seen in the *S. lividans* channel and in one-fourth of the much larger Na^+ channel protein (Fig. 30-17A). Furthermore, all known potassium channels, from bacteria to human beings, have the conserved sequence GYGD in the C-terminal half of the P region.[453] A great variety of K_v channels are known. There are ~70 genes for these channels in the *Caenorhabditis elegans* genome.[456] One of the first K_v channel genes to be cloned was from a *Drosophilia* mutant known for its neurological defect as *shaker*. Its structure (Fig. 30-18), which is based in part on modeling from the KcsA channel, has the ion selective filter with the conserved sequence **TVGYG** in the expected location. At the cytoplasmic end of the pore is an additional structure not found in the KcsA channel. This is the **inactivation gate**, so called because it accomplished the rapid self-inactivation of the K^+ channels during passage of the action potential (Fig. 30-18A). This is one of the factors necessary for recovery and repolarization of the axon membrane. The inactivation gate is composed of N-terminal ~130 residue "T1" domains of the α subunits together with parts of the β subunits, which are associated as a tetramer beneath the channel in the cytoplasm (Fig. 30-18B). Various experimental data including mutational analysis suggest that small ball-like domains at the N termini of the β subunits block the channel.[456–458] Zhou *et al*. propose that the N termini unfold into an extended conformation, passing through "windows" between the T1 domains and the channel and allowing the $–NH_3^+$ ends to bind into the central cavity in the channel.[459] The same site can be blocked by well-known quaternary amine inhibitors such as tetraethylammonium, tetrabutylammonium ions, or tetrabutylantimony, an analog used for X-ray crystallography.

The T1 domain of the channel not only participates in control of the ion flux but also stabilizes the pore complex.[459a] Among the various K^+, Ca^{2+}, and Na^+ channels the regulatory β subunits are quite variable in their structures and mechanisms of gating.[459b] Some β subunits have bound NADH. A speculative possibility (p. 737) is that the rapid interconversion of the positively charged thiazolium ion and negatively charged thiolate ion forms of thiamin (Eq. 7-19) plays some role in nerve conduction, e.g., voltage sensing.

Some questions about ion channels have been hard to answer. For example, how are small cations allowed to flow rapidly through a very small opening in a 2–3 nm thick nonpolar core of a membrane?[460,461] From basic electrostatic principles ΔG for transfer of an ion to the center of a membrane has been estimated as ~160 kJ/mol, a high thermodynamic barrier to transport. A solution to this problem apparently lies partly in the fact that at the center of the lipid bilayer the ion channel contains a cavity large enough (~0.5 nm diameter) to hold about 50 water molecules. Cations tend to enter this cavity, and X-ray studies have shown that the electron-dense Rb^+ does occupy the cavity. A second stabilizing factor is that four helices have their negative (C-terminal) ends pointing toward the cavity. Although the electrostatic effect of these helix dipoles (Fig. 2-20A) might be regarded as negligible, computations indicate that within the low dielectric bilayer the stabilizing effect of the helices becomes significant.[460,461]

How are the pores in these channels opened and closed? Different channels are gated in different ways.

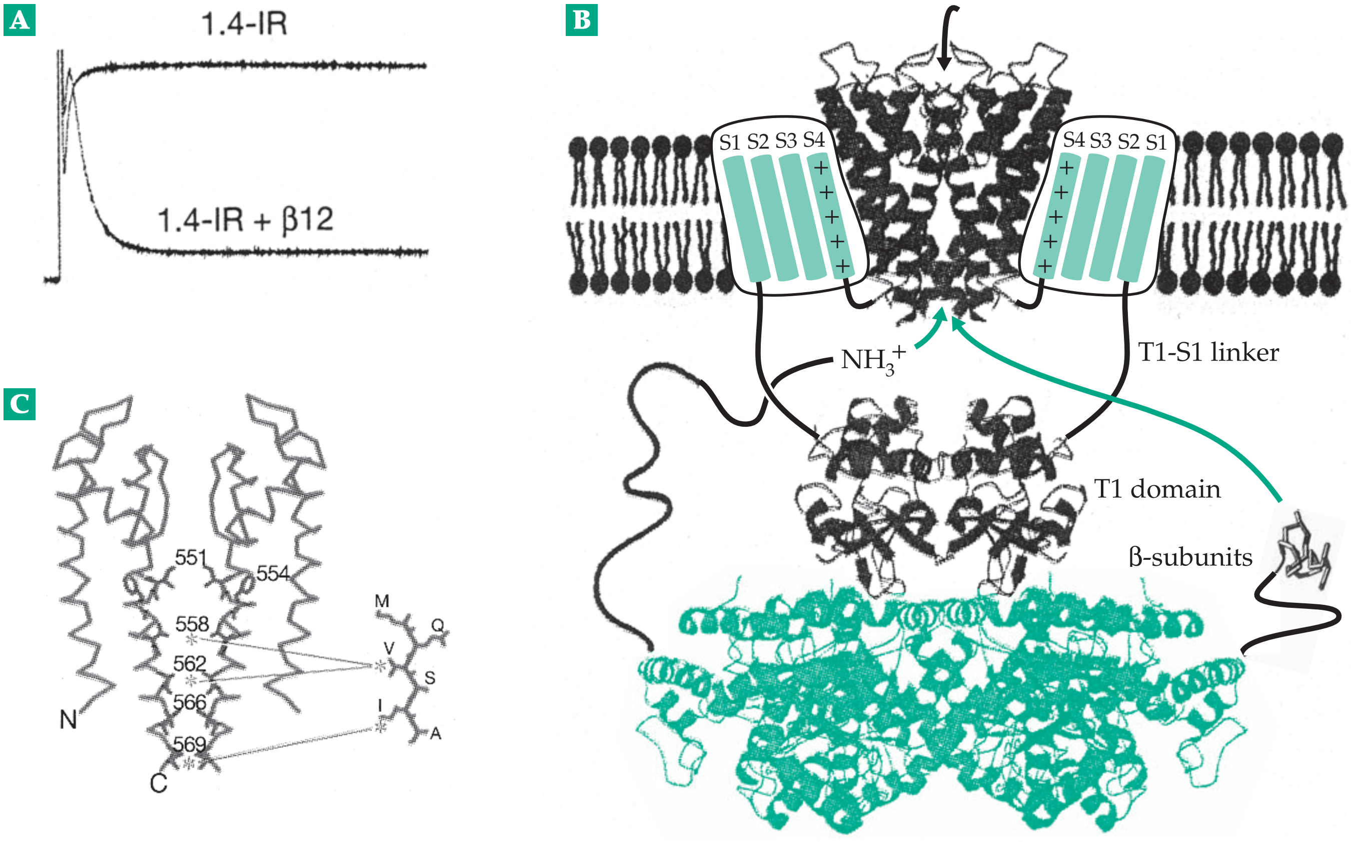


Figure 30-18 (A) K^+ currents recorded from *Xenopus laevus* oocytes carrying cloned genes of *Drosophila* shaker K^+ channels under two-electrode voltage-clamp conditions. Trace 1.4-IR was obtained from a cell expressing channels that lack the inactivation gate. Trace 1.4-IR + β_{12}, obtained from a cell expressing β subunits as well, shows rapid self-inactivation. (From Zhou *et al.*[459]) (B) Composite model of a voltage-dependent K^+ channel. The pore structure in the α subunit is represented by the KcsA channel (Fig. 8-21). The structure of the T1-β complex is from Gulbis *et al*.[458a] The drawing is modified from that of Zhou.[459] (C) Ball-and-stick view of the selectivity filter showing positions of four bound K^+ ions. Two of the four TVGYG peptide strands of the conduction pore are shown. Courtesy of Roderick MacKinnon.

The KcsA channel, which is mostly closed at neutral pH, responds by opening at a low external pH.[462] Using methods of spin labeling and EPR spectroscopy, Perozo *et al.* found small translational and rotational movements of the helices that form the pore (Fig. 30-18). These may alter the diameter of the pore, opening or closing it.[463] How do the electrostatic sensors control the process? The details are uncertain, but the sensor is thought to lie in a conserved sequence of arginine and lysine residues interspersed with hydrophobic amino acids in transmembrane helix 4 of the channel protein (Fig. 30-18; see also Fig. 30-17).[456] How do potassium pores select K^+ over Na^+ or Ca^{2+}? One factor is that Na^+ is more heavily hydrated than K^+ (p. 311). This allows K^+ to pass through the channel more readily than Na^+.[464] Potassium ions travel through the 1.2-nm-long selectivity filter at a rate of $\sim 10^8\ s^{-1}$ in consecutive steps of dehydration, movement, and rehydration occurring in ~10 ns.[464a–d] The process is catalyzed by polypeptides and may depend upon competition between a state in which a ring of four hydrogen-bonded peptide groups is formed and a state in which the four carbonyl groups coordinate a K^+ ion.[464d]

Belonging to the same structural group as the K_v channels are Ca^{2+}-regulated K^+ channels[465,466] Some bacterial channels are controlled by binding of Ca^{2+} ions to a “gating ring” on the intracellular membrane surface.[466a] A mammalian channel is controlled by a complex of calmodulin with the intracellular end of the α subunits of the channel[466b] and others.[453,467] A second large group of K^+ channels, containing seven subfamilies, are the **inward rectifying** (Kir) channels.[455,468] They are tetramers of 360- to 500-residue polypeptide chains, each chain forming two transmembrane helices with a P region between them.[453,469] These channels support a large conductance when K^+ ions flow out from a cell but only a small conductance when they flow in.[470] Kir channels are subject to a variety of controls, which include effects of pH.[471,472]

Some are inhibited by ATP,[473–474b] and others by eicosanoids[475] or inositol hexaphosphate.[476] Some of the ATP-sensitive channels contain an ABC transporter subunit and are binding sites for sulfonylureas and other drugs. See discussion on p. 421. A number of human disorders in Kir channels have been identified.[468] The human Kir channels participate in regulation of resting membrane potentials in K^+ homeostasis, control of heart rate, and hormone secretion.[468] A third group of K^+ channels are dimeric, but each subunit contains two tandem P regions and 4–8 transmembrane helices.[455]

Chloride channels and the ionic environment of neurons. All cells contain voltage-gated chloride channels, which are encoded by the *Clc* genes mentioned on pp. 420, 421.[477,477a] Recently crystal structures[477a–c] have revealed chloride channels formed in single polypeptide chains arranged as dimers. The selectivity filter involves stabilization by the positive ends of α-helix dipoles. The importance of the corresponding proteins to the human body is shown by the existence of several specific diseases arising from mutations in their genes (p. 420).[478,479] A calcium-regulated Cl^- channel is also present[480] as is the ATP-gated CFTR channel (Box 26-A).[480a] In addition, other ligand-gated Cl^- channels, such as γ-aminobutyrate receptor channels (Section B,9), are found in the central nervous system.[481] A glutamate-gated chloride channel in invertebrate organisms is the site of action of the antihelminthic and insecticidal compound **ivermectin**.[481a]

The significance of ion channels can be better appreciated by considering the ionic environment of nerve axons.[149] Mammalian neurons have roughly the following millimolar concentrations of ions in the cytosol and in the external medium. (The concentration gradients for the much-studied squid axon are substantially higher.[149,482]) The membrane potentials that could arise from each one of these concentrations, according to Eq. 8-2, are also given.[149] In a resting neuron the K^+ potential dominates with an observed membrane potential of ~–80 mV. Some K^+ channels are open and the K^+ and Cl^- concentrations are nearly in Donnan equilibrium across the membrane. The Na^+ and Ca^{2+} channels are closed, and the sodium and calcium pumps keep the internal concentrations of these ions low.

	Cytosol	Extracellular	E_m (mV)
K^+	150	5.5	–90
Na^+	15	150	+60
Ca^{2+}	10^{-4}	1.5	+270
Cl^-	9	12.5	–70

When an action potential is propagated, a wave of depolarization moves along the axon, changing the membrane potential suddenly to a less negative value. When it reaches ~50 mV the Na^+ channels open, allowing sodium ions to flow into the cell causing further propagation of the wave of depolarization. After ~1–2 ms the Na^+ channels begin to deactivate. At the same time the slower K^+ channels open allowing potassium ions to flow out and to repolarize the membrane, the membrane potential sometimes transiently reaching more negative values (hyperpolarization) than the ~80 mV resting potential. Action of the Na^+,K^+-ATPase then restores the original state. The finely tuned properties and sequential opening and closing of the channel proteins are essential to the conduction of nerve impulses.

The existence of voltage-gated ion channels in bilayers are not limited to nerve membranes. They are present to some extent in all cell membranes. Even the paramecium has at least seven kinds of Na^+, K^+, and Ca^{2+} channels.[483] Channels may also be formed by many peptide antibiotics. Among them are the human defensins (Chapter 31) and the ~20-residue **alamethicin**. Six to eleven of the mostly helical monomers of that antibiotic assemble to form a single voltage-dependent channel.[484,484a] The bacterial toxin colicin E1 (Box 8-D) forms voltage-dependent channels within bacterial membranes.[485]

Receptor-associated ion channels. Many neurotransmitters, including acetylcholine and glutamate, act to open ion channels that are part of the receptor protein or of a tight complex of proteins.[149,486] Such **ionotropic receptors** are responsible for most rapid neuronal action. For example, binding of acetylcholine to its receptor in the neuromuscular junction causes the release of Ca^{2+} ions from the exterior into the muscle fibers. Binding of glutamate to its ionotropic receptor in a synaptic ending of a dendrite causes an influx of ions into the cytoplasm, initiating an action potential in the dendrite. In most instances the properties of the receptor channel favor the rapid flow of Ca^{2+} ions into the cytoplasm.

Many other receptors are 7-helix transmembrane proteins, which activate guanine nucleotide G proteins (Chapter 11, Section D, 3). The G proteins couple some receptors directly to Ca^{2+} channels; they couple other receptors to adenylate cyclase and cyclic AMP-activated channels and yet others via phospholipase C to K^+ channels and indirectly to Ca^{2+} channels (Fig. 30-19). All of these G protein coupled receptors are referred to as **metabotropic receptors**. A single synapse often contains both ionotropic receptors and metabotropic receptors. The ionotropic receptors induce a rapid (<1 ms) response, while the metabotropic receptors act more slowly. However, in most cases the final effect is the release of calcium ions into the cytoplasm

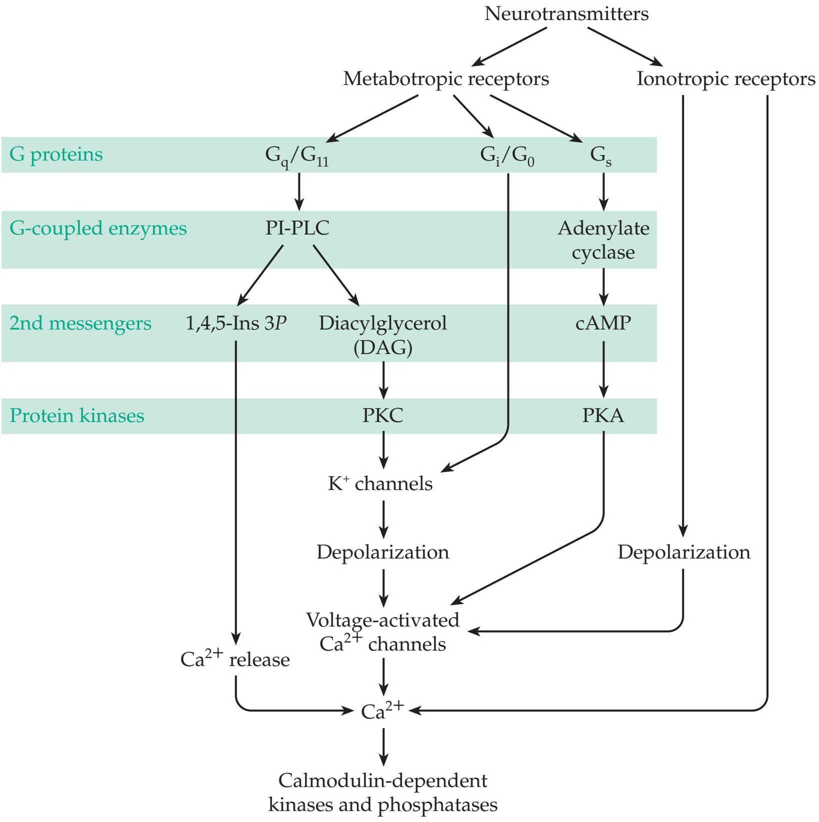


Figure 30-19 Major signaling pathways from metabotropic and ionotropic receptors in neurons. Various G proteins control the signaling from mutabotropic receptors using phosphatidylinisitol-specific phospholipase C (PI-PLC) and adenylate cyclase or acting directly on K^+ ion channels. Adapted from Fig. 5.1 of Nicholls' *Proteins, Transmitters, and Synapses*.[149]

(Fig. 30-19). The rapid response may be initiation of an action potential, while the slow response may be activation of calmodulin-dependent kinases and phosphatases.[149]

6. A Plethora of Neurotoxins

Bacteria, protozoa, and venomous animals synthesize numerous toxins that are used to kill their prey or to defend themselves. Sea anemones, jellyfish, cone snails, insects, spiders, scorpions, and snakes all make potent and highly specific neurotoxins. Plants form a host of alkaloids and other specialized products, some of which are specifically neurotoxic and able to deter predators. More than 500 species of marine cone snails of the genus *Conus* synthesize a vast array of polypeptide toxins (**conotoxins**),[487–489] some with unusual posttranslational modifications.[490,491] The slow-moving snails are voracious predators that use their toxins, which they inject with a disposible harpoon-like tooth,[492] to paralyze fish, molluscs, or worms.[493]

The targets for natural biological toxins include ion channels and receptors for transmitters. At least four parts of the voltage-gated sodium channels are binding sites for extremely toxic natural products.[494–499] **Tetrodotoxin** (Fig. 30-16),[496,497] which is found in the puffer fish, certain newts,[429a] and venom of the blue-ringed octopus, and also the shellfish poison **saxitoxin** (Fig. 30-16) block the entry of sodium ions into the channels.[498] **Bactrachotoxin** (Fig. 22-12) and related lipophilic compounds such as **veratridine** *increase* sodium permeability by blocking the channels permanently open. **Pyrethroid insecticides** (p. 1237) prolong the time that the sodium channels stay open after excitation. Some **scorpion toxins** (Fig. 30-16),[494,499] which all have a hydrophobic core made from a short α helix and a three-strand antiparallel β sheet,[500–502] and **sea anemone toxins**[495,503–505] also stabilize the open conformation of the Na^+ channels. Other smaller ~4-kDa scorpion toxins block K^+ or Cl^- channels or other receptors.[500,506,507] Some are most toxic to insects and others to mammals.[500] Although their three-dimensional structures resemble those of scorpion toxins, the amino acid sequences of anemone toxins show no homology.[505] The most potent poison produced by the red tide organism, the dinoflagellate *Gymnodinium breve* (Fig. 1-9), is **brevetoxin A** (Fig. 30-16).[436,508] It selectively opens one class of sodium channels.[495]

Venoms of **cobras**, **sea snakes**, and pit vipers contain several 6- to 7-kDa proteins that bind to acetylcholine receptors (Fig. 30-23) of the postsynaptic neurons, preventing binding of the neurotransmitter and opening of the ion channels.[509,510] All of these toxins contain four disulfide bridges and share with certain plant proteins a folding pattern that has been called the **toxin-agglutinin fold**[511,512] (Fig. 30-16). These toxins include **erabutoxin a** (Fig. 30-16) from a sea snake[513,514] as well as the 74-residue toxin **bungaratoxin a** (from the banded Krait). This toxin, which has been used to titrate acetylcholine receptors in neuromuscular junctions, is a member of the *long neurotoxin* group, which contains 71–74 residues and five disulfide bonds.[515] Other *short neurotoxins* are 60–62 residues in length with four disulfide bridges.[516] Cobra toxins contain both neurotoxins and **cardiotoxins**, which have somewhat similar structures but quite different modes of action.[517,518] In contrast, **crotoxin** from the venom of a South American rattlesnake[510,510a] and **β-bungarotoxin**[519] consist of 13-kDa phospholipases A_2 complexed with smaller 7.5-kDa proteins. They act at the presynaptic membranes of selected neurons by blocking neurotransmitter release.[520]

The seven types of **botulinum toxin**[521–523a] and the **tetanus toxin**[524] are the most neurotoxic substances known. Only 10^8 molecules are sufficient to kill a mouse. Both toxins are zinc proteases, which block presynaptic transmitter release by cleaving specific synaptic vesicles proteins (see p. 1780 and Fig. 30-20).[522,523,525–528] They bind initially to ganglioside in the neuromuscular junction, one subunit then being internalized as with the diphtheria toxin (Box 29-A). Botulinum toxins specifically enter motor neurons,[521,528a] while tetanus toxin is taken up via synaptic vesicle endocytosis[529] by both peripheral and central neurons. Retrograde axonal transport carries the toxin into the central nervous system and across synaptic clefts into cholinergic interneurons, which are poisoned.

The black widow spider produces the 130-kDa **α-latrotoxin,** which causes massive release of acetylcholine, norepinephrine, dopamine, and GABA from synaptosomal endings.[530,531] The small **anatoxin-a** or "very fast death factor" (Fig. 30-22), which is synthesized by various cyanobacteria, antagonizes both muscarinic and nicotinic acetylcholine receptors.[532] Cone snails synthesize mixtures of the 13- to 17-residue conotoxins (Fig. 30-16).[493] They cause rapid paralysis of fish permitting the snails to prey on the much faster fish. They bind to a variety of targets, which include Na^+, K^+, and Ca^{2+} channels,[435,492] and acetylcholine,[533,534] and glutamate[490] receptors. One of the toxins is a 17-residue peptide containing five residues of γ-carboxyglutamate and is also notable for the fact that intercerebral injection of less than one microgram of the toxin induces a prolonged sleeplike state in mice.[490,493] The venom of *Conus geographicus* is so toxic that two-thirds of human stinging cases are fatal.

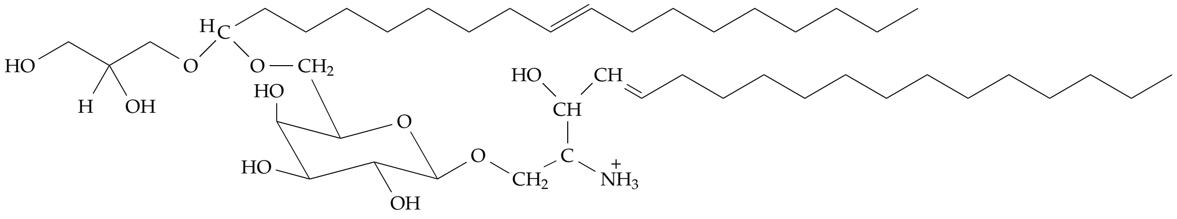
The most deadly nonproteinaceous toxin known, **palytoxin**, is also the most complex structure ever established without the aid of X-ray crystallography.[535,536] It is produced by marine zoanthids of the genus *Palythoa* and has the molecular formula $C_{129}H_{223}N_3O_{54}$.

7. Neuronal Metabolism

The brain has a very high rate of metabolism. Although accounting for only 1/50 of the body mass its utilization of energy amounts to 1/5 of the basal metabolism. This is ~20 watts and is nearly constant day and night. It reflects the unusually active metabolism of neurons, a major part of which can be attributed to the sodium–potassium ion pumps in the membranes and to the maintenance of the excitable state.[536a] The source of energy for these processes is the ATP that is utilized to drive the ion pumps and thereby to maintain the membrane potential needed to drive the action potentials. The ATP is formed largely by oxidative metabolism of glucose and, to a lesser extent, of acetoacetate. The large surface area of the axons as well as the frequency with which they transmit nerve impulses accounts for the high rate of metabolism.

Another factor peculiar to neurons doubtless contributes also to their rapid metabolism. The nucleus and most of the ribosomes are found in the cell body. Although few ribosomes are seen in axons and dendrites[536b], many proteins are needed in high concentrations within the axons and synaptic endings. Among these are enzymes catalyzing synthesis and catabolism of neurotransmitters and membrane proteins. If an axon is cut, the separated synaptic endings soon atrophy, an observation that long ago suggested that essential materials, which may include mRNAs,[537] might flow from the cell body. It has now been established experimentally that many materials do move at the rate of 0.3–3 mm / day from the cell body down the axon.[538] More remarkable is **fast axonal transport** by which proteins and other materials move at rates of up to 5 µm / s (0.4 m / day). This transport is specifically blocked by vinblastine (Box 7-D) and batrachotoxin (Fig. 22-12). As has been pointed out in Chapter 20, an ATP-hydrolyzing protein chemically related to the myosin heads functions together with microtubules to provide a kind of miniature railway that moves materials along the microtubules. Transport is sometimes in the opposite direction, i.e., from the synaptic endings to the cell body. This **retrograde axonal transport** may be of importance in altering neuronal properties in response to electrical activity at synaptic endings. It also provides a means of recycling materials originally sent in the other direction.

Brain cells appear to transcribe an unusually large fraction of the genome.[539–541] About 20% of the DNA of human brain was found to hybridize with mRNA formed by brain cells. In other tissues about half this amount of DNA appears to be transcribed. A related observation that seems surprising is the absence of common electrophoretic variants of enzymes in the brain.[539] However, brain cells synthesize specialized isoforms of many proteins, e.g., of the G proteins (p. 558), the cytoskeletal protein 4.1 (Fig. 8-14),[542] and transglutaminase.[543] Unusual lipids, such as the cationic acetal of a galactosylcerebroside shown above,[544] are also formed. Adult rat brain contains about 30,000 different kinds of polyadenylated messenger RNA,[540] much of which lacks the poly(A) tail.[513] Many of these mRNAs contain a specific 82-nucleotide sequence within at least one of their introns. Sutcliffe *et al*. suggest that this is an **identifier sequence** instructing brain cells to express these genes.[540,545] However, the sequence is also found in genes transcribed in other tissues, and its significance is not clear.[546,547]



8. Synapses and Gap Junctions

Like the micro-transistors in a computer chip, synapses are the devices by which the brain operates. Synapses process and integrate information from many input channels, send signals on to other neurons, and store information. The information is not stored in digital form, but as chemical alterations in the synapses themselves.[482,548,549] Synapses are formed when axons, growing in response to a chemical trail, reach their destinations and send out branches, each with a bulbous terminal knob (**bouton**). When these boutons meet receptive regions on dendrites of another axon, synapses are formed.[550] The synapse is a very firm connection with a thin, tight synaptic cleft through which signaling takes place. It is surrounded in part by astrocytes or other glial cells (Fig. 30-20A,C).

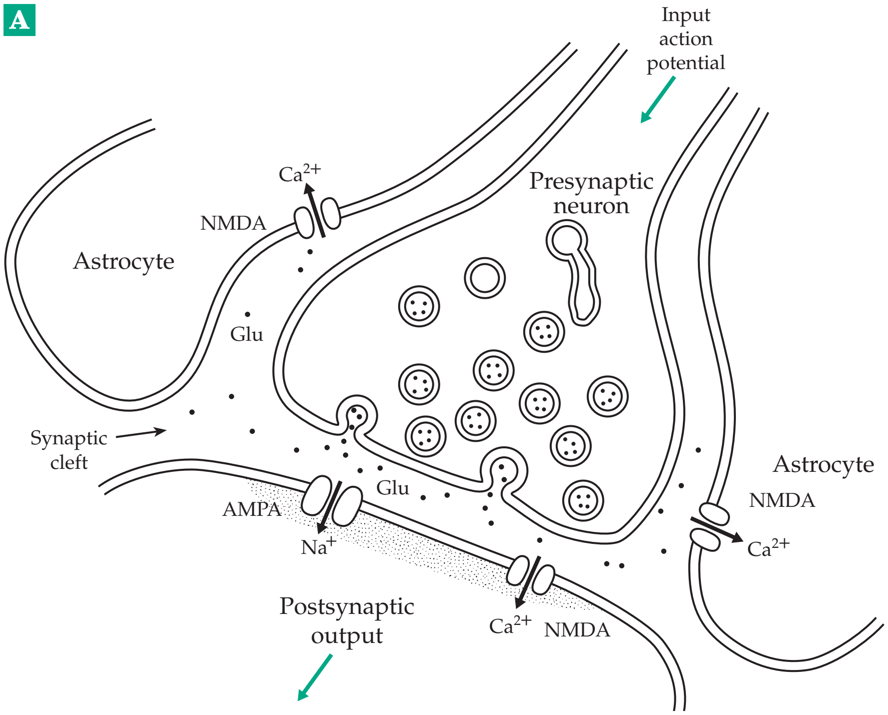
With the advent of electron microscopy, the fine structure of synaptic contacts became evident. The synaptic knobs were often found to contain vesicles of ~30−80 nm diameter, which were later shown by chemical analysis and staining procedures to contain the neurotransmitters (Fig. 30-10). In the case of the acetylcholine-releasing synapses (**cholinergic synapses**) each 80-nm vesicle contains ~40,000 molecules of acetylcholine,[551] the concentration in the vesicle being of the order of 0.5 M. To show that the acetylcholine released at a synapse stimulated the postsynaptic membrane to initiate an impulse, the technique of **electrophoretic injection** or **microiontophoresis** was developed.[552] By using ultramicrocapillaries a small pulse of current, e.g., 3×10^8 amp for 1 ms, can be used to inject electrically a compound directly into a synaptic cleft. The results may be observed with separate recording electrodes, one of which is inserted into an axon or a muscle fiber. By this means it was shown that amounts of acetylcholine comparable to those released at the large synapses of the neuromuscular junction do cause muscles to contract.

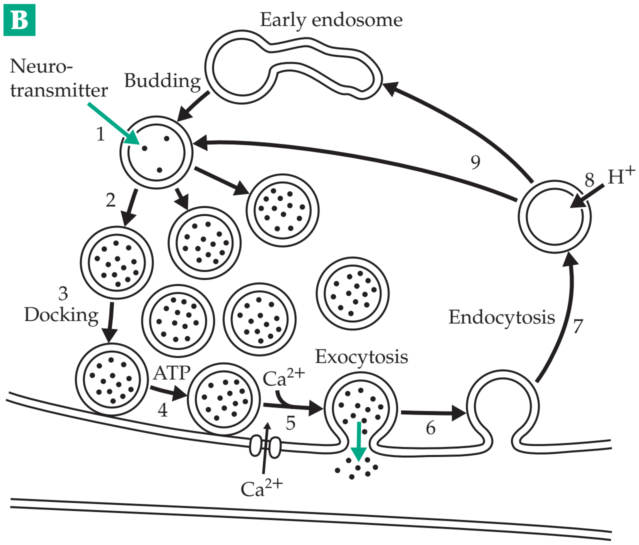
How does the release of neurotransmitter occur? That the release is "quantal," i.e., involving the entire content of a vesicle, was established from the observation of **miniature end-plate potentials**. These are fluctuations in the postsynaptic potential observed under conditions of weak stimulation of the presynaptic neuron. They reflect the randon release of neurotransmitter from individual vesicles.[553] Normally, a strong impulse will release on the order of 100−200 quanta of transmitter, enough to initiate an action potential in the postsynaptic neuron.

A synaptic vesicle cycle. The number of synaptic vesicles in a single synapse in the brain varies from fewer than 100 to several hundred. In specialized synapses there may be thousands. However, at any moment only a fraction of the total are in the "active zone," often aligned along the presynaptic membrane (Fig. 30-20A) or in specialized ribbons such as those in Fig. 30-10B. The vesicles are normally reused repeatedly, undergoing a cycle of filling with neurotransmitter, translocation to the active zone, ATP-dependent priming, exocytosis with release of the neurotransmitter into the synaptic cleft, coating with clathrin, endocytosis, and acidification as outlined in Fig. 30-20B.[554–557] The entire cycle may be completed within 40–60 s to avoid depletion of active vesicles.[558,559] A key event in the cycle is the arrival of an action potential at the presynaptic neuron end.

The accompanying depolarization of the membrane at the synaptic ending permits a rapid inflow of calcium ions through a voltage-gated calcium channel.[444,560] Within less than 0.1 ms the transient increase in intracellular $[Ca^{2+}]$ triggers the release of the contents of the vesicles. About four calcium ions are needed to release one clathrin-coated vesicle (Fig. 30-20A,B). The membrane fusion required for transmitter release involves cytoskeletal proteins of the synaptic endings as well as specialized proteins that are present in the membranes of the synaptic vesicles (Table 30-6). In fact, every step in the cycle depends upon specialized proteins.[387]

Synaptic vesicles can be isolated in large quantities. Their composition is well known, and the proteins have been studied intensively. Indeed, much of what we know about exocytosis and vescular transport has been learned from investigation of synaptic vesicles.[554,561,562] A small synaptic vesicle of 35 nm diameter will contain ~10,000 phospholipid molecules in its membrane and only about 200 protein molecules, at least one of which must be a 13-subunit vacuolar type proton pump (Fig. 18-14). This pump acidifies the vacuole, allowing uptake of a neurotransmitter. Although many different proteins may be found in synaptic membranes, only about 15, which are listed in Table 30-6, are found in all synaptic vesicles and appear essential to function.


A
Input
action
potential
Presynaptic
neuron
Ca^{2+}
NMDA
Astrocyte
Glu
Synaptic
cleft
Glu
AMPA
Na^{+}
Astrocyte
NMDA
Ca^{2+}
Postsynaptic
output
Ca^{2+}
NMDA


B
Early endosome
Neuro-
transmitter
Budding
1
9
8
H^{+}
2
3
Docking
Endocytosis
7
ATP
Ca^{2+}
Exocytosis
4
5
6
Ca^{2+}

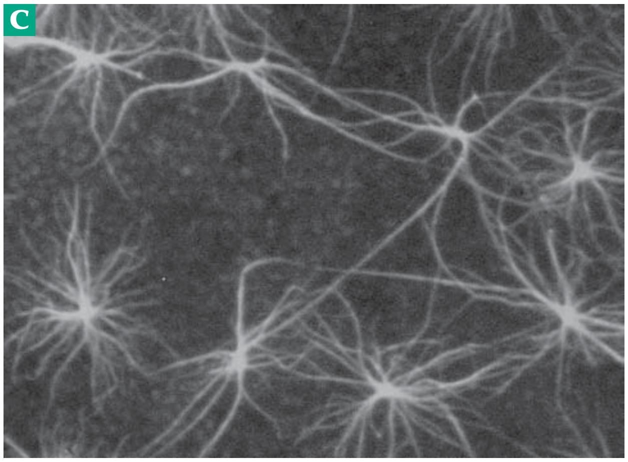
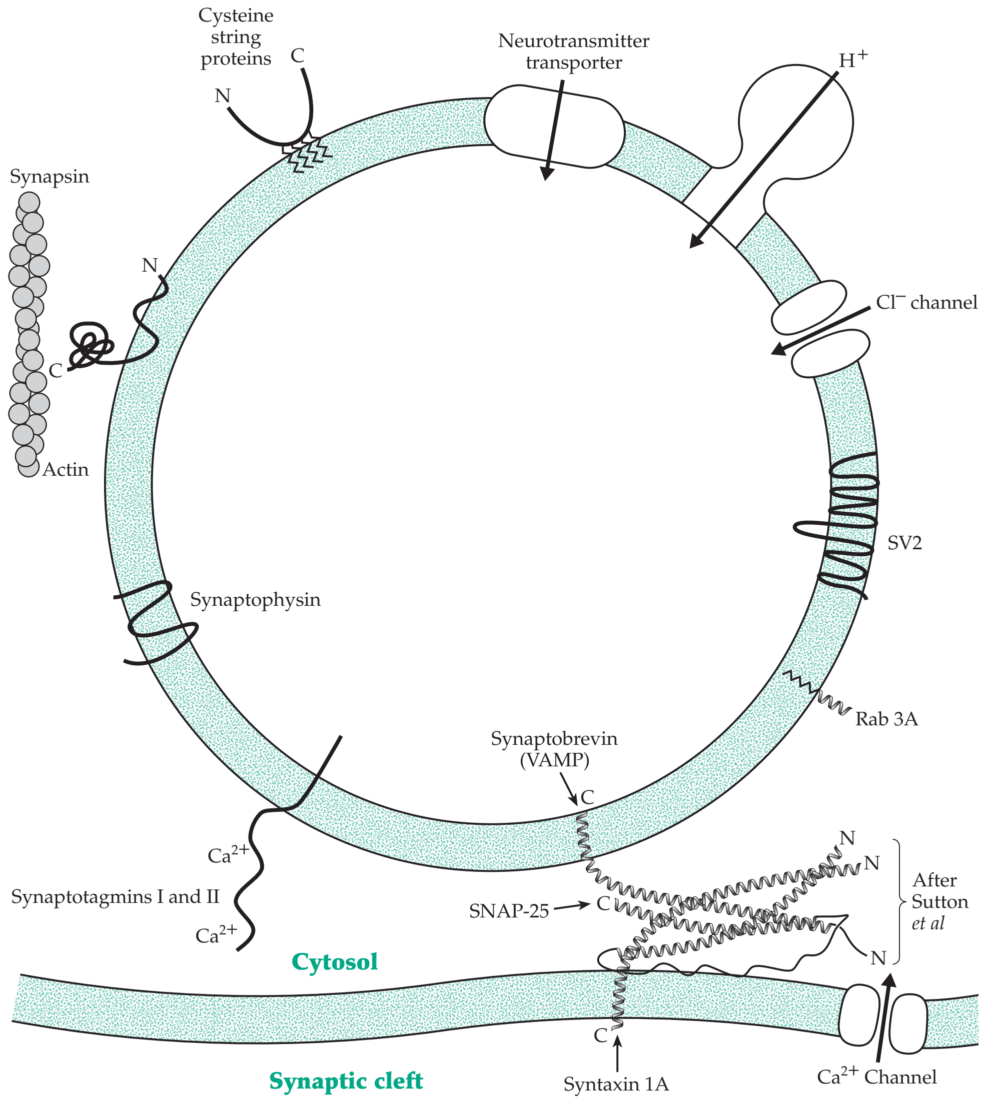


Figure 30-20 (A) Schematic drawing of a fast glutamatergic synapse. An action potential arrives at the synapse, depolarizing the presynaptic membrane and allowing calcium ions to enter the cytoplasm via voltage-gated Ca^{2+} channels. The Ca^{2+} ions induce exocytosis of small synaptic vesicles from the "active zone" near the membrane, releasing glutamate into the synaptic cleft. After diffusing rapidly across the narrow ~50-nm synaptic gap the glutamate binds to its receptors on the ending on a dendrite from a second (postsynaptic) neuron. Glutamatergic synapses usually have two types of receptor, NMDA and AMPA (see Fig. 30-24 and text). Both are ligand-gated ion channels, which release Ca^{2+} and Na^{+} into the cytosol of the postsynaptic ending depolarizing its membrane and possibly initiating an action potential. (B) The synaptic vesicle cycle. The synaptic vesicles, which are formed by budding from an early endosome, are filled with neurotransmitter (1). The filled vesicles are then transported to the active zone near the presynaptic membrane (2), are "docked" on the membrane surface (3), and undergo ATP-dependent priming (4). Binding of four Ca^{2+} ions induces exocytosis and rapid release of the neurotransmitter (5). The empty vesicles receive a clathrin coat (6) and undergo endocytosis (7) and uptake of protons (8) to acidify the content in preparation for a second round of neurotransmitter uptake. Alternatively the vesicle can fuse with an endosome as part of the cycle. After Südhof and Scheller.[554] (C) Small section of brain stained to reveal the astrocytes whose extensions form synapses not only with neurons, as in (A), but also with capillary blood vessels.[149] From Kimelberg and Norenberg.[564] Micrograph from Andreas Karschin, Heinz Wässle, and Jutta Schnitzer. (D) Illustration of some proteins essential to the synaptic vesicle cycle. Several are integral membrane proteins. Synaptotagmins contain Ca^{2+}-binding domains and may serve as calcium sensors. The vesicle is portrayed as if docked to the presynaptic membrane by interaction of the SNARE proteins synaptobrevin, syntaxin, and synaptotagmin. The 4-helix bundle is as portrayed by Sutton *et al.*[563]

D


Cysteine string proteins
C
N
Neurotransmitter transporter
H^+
Synapsin
N
C
Actin
Cl^- channel
SV2
Synaptophysin
Rab 3A
Synaptobrevin (VAMP)
C
N
N
After Sutton *et al*
Ca^{2+}
Synaptotagmins I and II
Ca^{2+}
SNAP-25
C
N
Cytosol
Synaptic cleft
C
Syntaxin 1A
Ca^{2+} Channel

The synaptic vesicles, which are formed by budding from early endosomes, take up neurotransmitters using one of the transporters (step 1 in Fig. 30-20B). Transmitter uptake is G-protein dependent[565] and is driven by the proton electrochemical gradient generated by a vacuolar type (V-type) ATPase (Chapter 18).[149,566] The filled vesicles move into the active zone where they undergo an ATP-dependent priming of uncertain nature.[555,567] Exocytosis (step 5 in Fig. 30-20B) requires membrane fusion, and it is possible that partial fusion occurs during the priming steps. Priming is also thought to involve interaction between vesicle-associated v-SNARES and synaptic membrane-associated t-SNARES (p. 521).[556,563] A major v-SNARE has been identified as **synaptobrevin**, which is also known as **VAMP** (vesicle-associated membrane protein).[563,568,568a] The C-terminal-anchored synaptobrevin is inserted into the plasma membrane of neuronal and neuroendocrine cells prior to endocytosis and budding of the synaptic vesicles.[568] The target t-SNARES have been identified as the synaptic plasma membrane proteins **syntaxin**[568b] and **SNAP-25**.[569–573] Syntaxin is an integral membrane protein, whereas SNAP-25 is anchored by palmitoylation.[571] These proteins bind together to form a synaptobrevin•syntaxin•SNAP-25 complex, which forms a four-helix bundle as shown in Fig. 30-20. Synaptobrevin and syntaxin each contribute one helix, while SNAP-25 provides two; all four have a mutually parallel orientation.[563,574,574a] The helix bundle is so tight that it has a high melting temperature and is resistant to proteolytic cleavage. Nevertheless, the helical domains of both synaptobrevin and syntaxin are sites of very specific cleavage by the zinc proteases of tetanus and botulinin toxins.[527,563,570] Cutting of the protein chains by these toxins prevents proper formation of the four-helix bundle and prevents release of neurotransmitter. It is thought that the complex, which probably forms at several points on the periphery of the docked synaptic vesicle, is essential for membrane fusion.

Other proteins are also needed. All cell fusion processes seem to require regulatory proteins that are essential to neurotransmission in the nematode *C. elegans*. Two of these are encoded by the nematode genes *unc*-13 and *unc*-18. The corresponding mammalian proteins **munc-13** and **munc-18** interact with syntaxin and are essential for exocytosis of synaptic vesicles.[572,575] An ATPase is also needed for correct functioning of the SNARE complex[574] as are other additional proteins.[570]

Details of the control of exocytosis are also uncertain. **Synaptotagmin I**, which contains two Ca^{2+}-binding domains, is probably the sensor that detects the rapid influx of Ca^{2+} that initiates exocytosis.[576–578b] It binds several Ca^{2+} ions via a β-sandwich motif that contains five aspartate side chains at its tip. This motif is conserved in a large family of synaptotagmins. A possibility is that Ca^{2+}–synaptotagmin complexes may self-associate to form a protein ring around the site where the fusion pore forms.[576] Synaptotagmin I also interacts with both syntaxin and with **neurexins**, proteins related to laminin (Fig. 8-33) and present in numerous variant forms in nerve endings. Neurexins are also targets for the α-lathrotoxin of the black widow spider.[531,579] Other proteins that may participate in membrane fusion include the unique **cysteine string proteins**, which in *Drosophila* contain 13 cysteine residues, 11 of which are palmitoylated.[580,581] Nitric

TABLE 30-6
Some Proteins Important to the Formation and Functioning of Synaptic Vesicles[a]

1. Synaptic vesicle proteins	
Synapsins Ia, Ib, IIa, IIb	Peripheral, abundant
Rab3, rabphilin	Rab 3 has lipid anchor
Cysteine string proteins (CSP)	Ca^{2+}-binding
Synaptotagmins	Single transmembrane helix; Ca^{2+} receptor N terminus in vesicle
Synaptobrevins (VAMPs)[b]	SNARE proteins, C termini in vesicle
Synaptophysins, synaptogyrin	Integral membrane protein
SV2 A, B, C	Integral membrane protein, Cl^- transporter
SCAMPS 1 and 4	Integral membrane protein
SVOP	Integral membrane protein
Vacuolar H^+ pump	13 subunits
Cytochrome 561	H^+ generator
Neurotransmitter transporters	For acetylcholine, glutamate, GABA/glycine, catecholamines, ATP
Ancillary transporters	Zn^{2+}, Cl^-
2. Presynaptic membrane proteins	
Syntaxin [b]	t-SNARE
SNAP-25[b]	t-SNARE
Munc-13	
Ca^{2+} channel	
Agrin	
Neurexin	
Actin and microtubules	In dendrites
3. Postsynaptic specializations	
Receptors	e.g., NMDA, AMPA

[a] Based on data of Südhof and Scheller: Südhof, T. C., and Scheller, R. H. (2000) in *Synapses* (Cowan, W. M., Südhof, T. C., and Stevens, C. F., eds), pp. 177–215, Johns Hopkins Univ. Press, Baltimore, Maryland and Südhof, T. C. (1995) *Nature (London)* **375**, 645–653.

[b] Targets for clostridial toxins, tetanus, botulinin.

oxide NO may be involved in a late stage of exocytosis,[392] and phospholipase D1 may also be required.[582]

Presynaptic nerve terminals may contain as few as a hundred vesicles, which must be recycled rapidly after exocytosis in order to allow for repetitive firing.[558,559,583] Several proteins are needed for endocytosis (step 7 in Fig. 30-20). These include **endophilin I**,[584] the vesicle transport ATPase **NSF**,[574] GTPases,[565] and the soluble NSF attachment protein α-SNAP (which is not related to SNAP-25).[585]

Functions of some other abundant proteins of synaptic vesicles have not yet been accurately defined. The **synapsins** are abundant peripheral membrane ATP-binding proteins with multiple phosphorylation sites and variable C-terminal domains that interact with cytoskeletal proteins such as actin microtubules, microfilaments, and spectrin.[554,561,586,587] Another abundant protein is **synaptophysin**, an integral membrane protein found in all synaptic vesicles.[554,561,588] Other proteins are discussed by Südhof and Scheller.[554] The small G protein **rab 3** together with the Ca^{2+}-binding protein **rabphilin** participate in a G-protein cycle that helps to drive exocytosis.[554] Synaptotagmin, as well as clathrin assembly proteins bind inositol hexaphosphate ($InsP_6$; Fig. 11-9), which undergoes active turnover in synapses. This suggests a role for $InsP_6$ in the endocytosis steps of the synaptic vesicle cycle.[589] The brain is rich in zinc ions. Much of the Zn^{2+} is bound into zinc finger domains of transcriptional regulators, but much is also present in a relatively free form within synapses of the hippocampus, cerebral cortex, and other regions.[590,591] Zinc ions may function as a neuromodulator in glutamatergic synapses.[591]

What does a neurotransmitter do at the postsynaptic membrane? In the case of acetylcholine in neuromuscular junctions the principal action appears to be one of opening sodium channels and thereby depolarizing the postsynaptic membrane. If enough nerve impulses arrive, an action potential will be initiated in the postsynaptic neuron. In other cases, the first response may be activation of a protein kinase either directly or by opening a channel for Ca^{2+}, which indirectly regulates protein kinases and phosphatases.[592] Thus, a complex cascade may be activated. See also Fig. 30-19.

The postsynaptic nerve ending, which is usually the tip of an axonal dendrite, has its own set of proteins, which varies to some extent with the nature of the neurotransmitter. In excitatory cells the plasma membrane of the postsynaptic neuron is thickened to ~30–40 nm to form the "**postsynaptic density**," a disc-like structure of clustered receptors of two types, which extends ~ 30 nm into the cytosol.[593,594] Only single receptor channels are indicated in Fig. 30-20, but many receptors are present in the clusters[594,595] as are other specialized proteins. One of these, designated PSD-95, was found to associate with the NMDA receptor using the yeast two-hybrid system (Box 29-F).[594] Neuronal nitric oxide synthase may also be present.

The large neuromuscular junctions, which contain clusters of acetylcholine receptors, have wider synaptic clefts (> 40 nm), which contain basal lamina, a dense network of collagen fibrils together with the heparan sulfate proteoglycan **agrin** (p. 437). Agrin activates a muscle-specific kinase MusK, which phosphorylates the acetylcholine receptors inducing clustering of the receptors together with other proteins embedded in the plasma membrane and binding to the cytosolic protein **rapsyn** (see Fig. 30-23B).[596,597] Agrin is also a component of **immunological synapses,** which are important in lymphocyte development (Chapter 31).[596,598,599] The neuromuscular junction is formed between two cell types, a neuron and a muscle myotube. Both contribute proteins, which include a muscle-specific laminin.[600]

Astrocytes and other glia. Although the glial cells greatly outnumber neurons, they were long regarded simply as glue, as implied by the name glia. We know now that the several types of glial cells have functions in many different aspects of brain chemistry.[149,564,601–605] The oligodendrocytes generate myelin sheaths around many brain neurons. Macrophages that invade the brain differentiate into **microglia** that serve as part of the innate immune system (Chapter 31). **Bergmann glia** of the cerebellum help guide axons during brain development. The astrocytes have many processes, which not only contact synapses directly (Fig. 30-20A,D) but also form contacts with capillary blood vessels. They often contain receptor ion channels of the same types as are found in postsynaptic membranes (see Fig. 30-20A) and respond to Ca^{2+} influx as do neurons.[602–603a] Glia often take up neurotransmitters and ions from synapses in order to prepare for consecutive nerve impulses. Glia may also control the number of synapses formed,[604–604b] and they may have other roles in brain development. For example, an iodothyronine deiodinase (Eq. 15-60) is expressed primarily in neonatal brain, where it supplies thyroid hormone essential to brain development.[605]

Gap junctions in synapses. Not all neurons communicate via chemical synapses. Gap junctions, which are found in both neurons, astrocytes, and other cells, serve as **electrical synapses**. Thus, heart cells are all electrically coupled together by gap junctions.[606 – 608] Gap junctions are formed with the aid of hexameric **connexons,** which are present in each of the opposed membranes and are aligned one with the other (Fig. 1-15F,G).[607,609,610] There may be thousands of connexons in a single gap junction, which resemble ion channels in appearance but contain pores ~1.5 nm in diameter. They are formed from 26- to 43- kDa

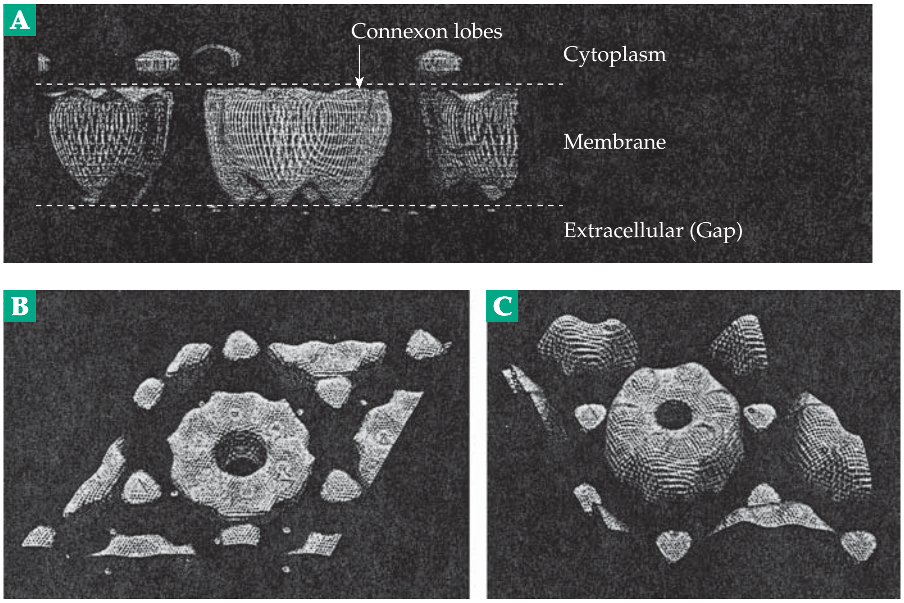


Figure 30-21 Images of gap junction connexins obtained by electron crystallographic methods at a resolution of 1.6 nm. (A) Cross-section. The thinkness of the (43 x 6) kDa hexameric connexin is 5.0 mm. (B) View of the connexin from the cytoplasmic side. (C) View from the extracellular side. From Perkins, Goodenough, and Sosinsky.[609] Courtesy of Guy Perkins.

protein subunits of the multigene family of **connexins**.[610–611a] Each gap junction consists of a pair of hexameric rings of connexins (Fig. 30-21), one ring from each of the two juxtaposed membrane surfaces.[609] Defects in connexins cause inherited deafness, neuropathy, malignancy, and cataract formation.[612–613a] The connexin subunits each contain four transmembrane helices and are related structurally to the peripheral myosin protein 22, the myelin proteolipid (p. 1767), and the protein stargazin (p. 1901), which is involved in synapse formation in the brain.[428a]

Another type of channel has been recognized quite recently. An ion channel, which regulates Mg^{2+} ion transport in kidney tubules, forms within the tight junctions that seal the extracellular space between cells (Fig. 1-15B). A protein **paracellin** forms channels through the tight junction protein complexes that surround the cells.[614,615]

9. Neurotransmitters

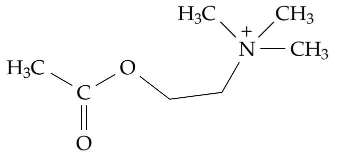
Studies of neuromuscular junctions of the autonomic nervous system as early as 1904 led to the suggestion that adrenaline might be released at the nerve endings. Later it was shown that, while adrenaline does serve as a transmitter at neuromuscular junctions in amphibians, it is primarily a hormone in mammals. Nevertheless, it was through this proposal that the concept of chemical communication in synapses was formulated. By 1921, it was shown that acetylcholine is released at nerve endings of the parasympathetic system, and it later became clear the motor nerve endings of the somatic system also release acetylcholine.

Acetylcholine is an established neurotransmitter because it meets five important criteria: (1) a synthetic mechanism exists within the presynaptic neuron; (2) a mechanism of storage (in vesicles) is evident; (3) the transmitter is released in proportion to the strength of the stimulus (frequency of firing); (4) postsynaptic action of the transmitter has been demonstrated directly by microiontophoresis; and (5) an efficient means for inactivation of the transmitter is present. The same five criteria must be met by other compounds if they are to be considered as transmitters.

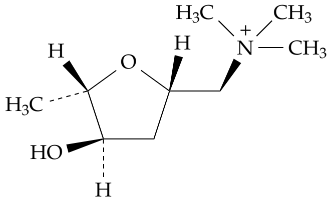
At present, in addition to acetylcholine, glutamate, and γ-aminobutyrate (GABA), glycine, noradrenaline (norepinephrine), and dopamine and 5-hydroxytryptamine (serotonin) are regarded as established transmitters. Other probable (**putative**) or possible **candidate transmitters** are also known. Aspartate, taurine, and a large number of peptides (Tables 30-1, 30-4) are under consideration.

Some transmitters, including noradrenaline, dopamine, serotonin, and various neuropeptides, are sometimes called **neuromodulators** rather than neurotransmitters. These compounds may not initiate a nerve impulse but may act on adenylate cyclase to increase or decrease cAMP levels and protein kinase activity. They may also diffuse through the extracellular space to influence a region of the brain greater than a single synaptic cleft. However, the distinction between transmitters and modulators is not exact.

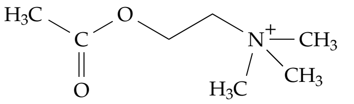
For many years it was assumed that a single neuron released only a single transmitter. We know now that this is incorrect.[616] For example, enzymes in neuromuscular junctions synthesize not only acetylcholine but also catecholamines, taurine, and GABA.[617] Some synapses in the central nervous system release both glycine and GABA.[618]


H3C
CH3
N
CH3
H3C
O
C
O

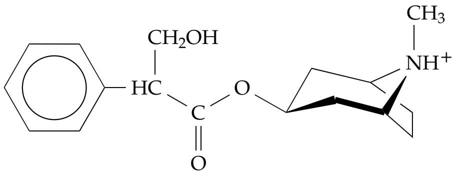
Acetylcholine in a skewed conformation


H3C
CH3
H
N
CH3
H
O
H3C
HO
H

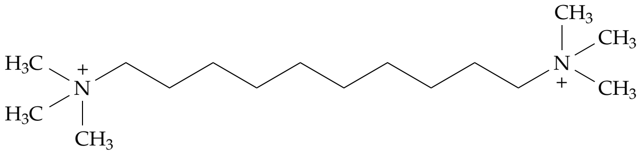
L(+)-Muscarine, a cholinergic agonist from the mushroom *Amanita muscaria*


H3C
O
C
O
N
CH3
CH3
H3C

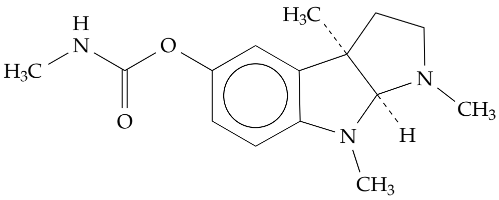
Acetylcholine in extended conformation


CH3
CH2OH
NH+
HC
O
C
O

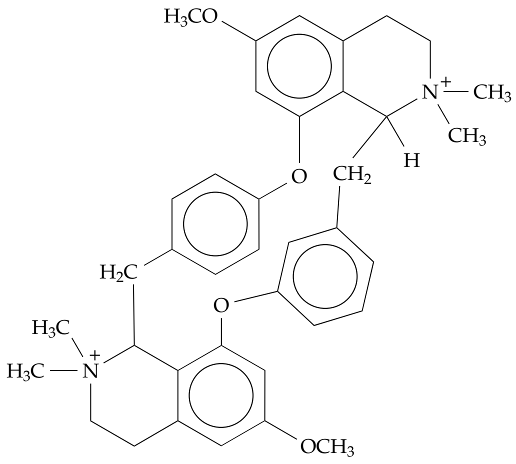
Atropine


CH3
CH3
H3C
N
N
H3C
CH3
CH3

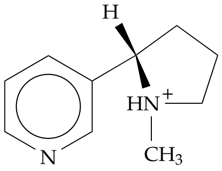
Decamethonium, a synthetic drug with potent curare-like activity


H
N
H3C
O
H3C
N
CH3
O
N
H
CH3

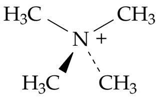
Physostigmine, an acetylcholinesterase inhibitor widely used in treatment of glaucoma. Compare with other carbamate esters (Box 12-E)


H3CO
N
CH3
CH3
H
O
CH2
H2C
O
H3C
H3C
N
OCH3

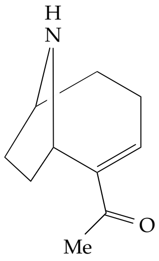
D-Tubocurarine, the principal ingredient of South American arrow poisons. Blocks cholinergic receptors in skeletal muscle



Nicotine (protonated)


H3C
CH3
N+
H3C
CH3

Tetramethylammonium ion



Anatoxin-a (very fast death factor) from the cyanobacterium *Anabaena*

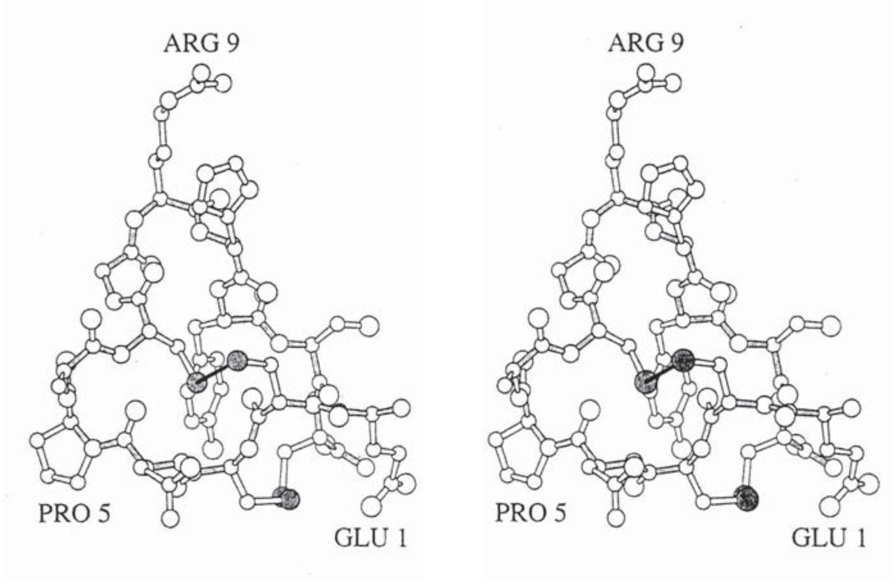

ARG 9
PRO 5
GLU 1
ARG 9
PRO 5
GLU 1

Figure 30-22 Some inhibitors of cholinergic synapses. The structure of conotoxin GI is from Guddat *et al.*[533] Courtesy of A. B. Admundson.

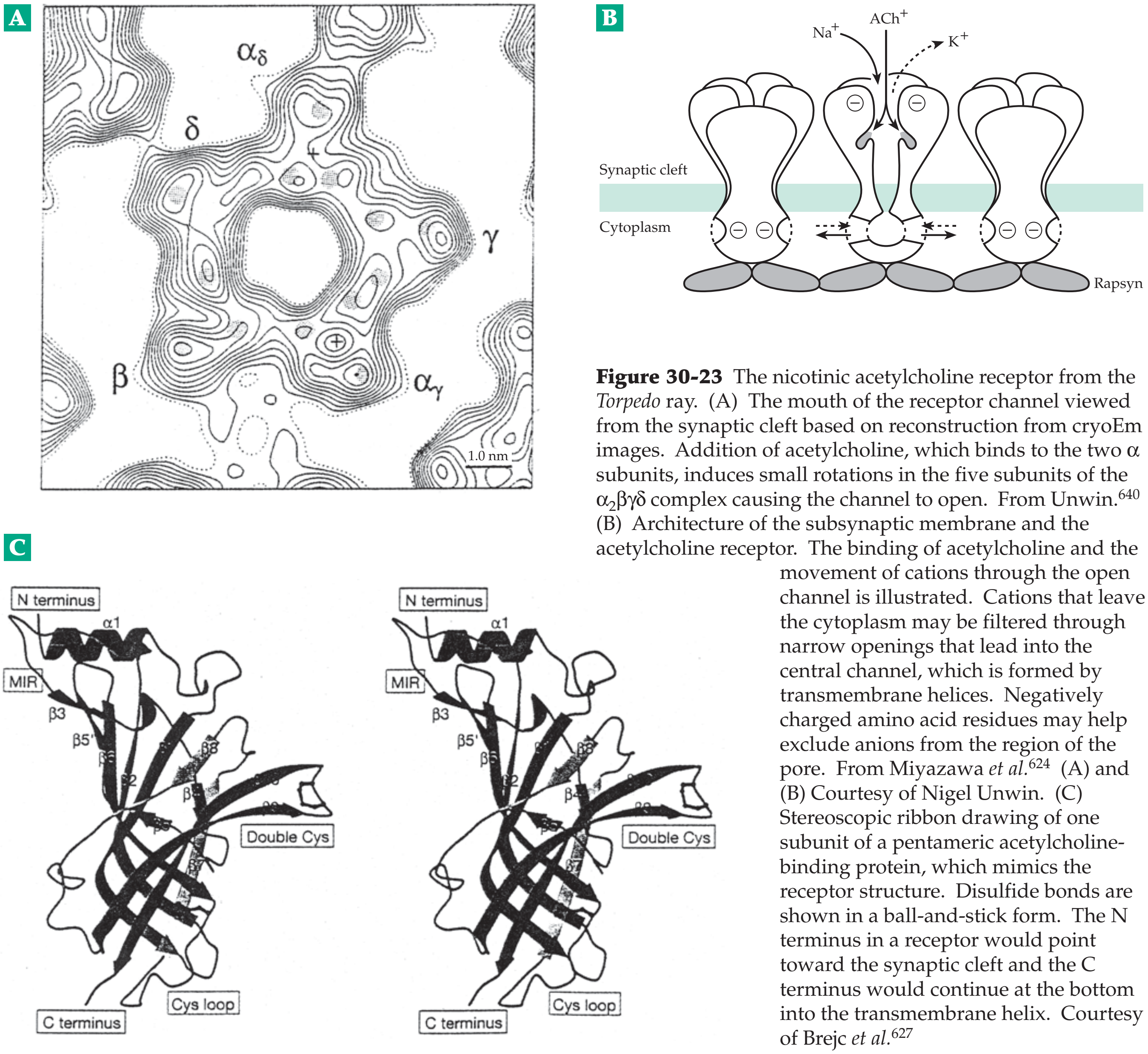


Figure 30-23 The nicotinic acetylcholine receptor from the *Torpedo* ray. (A) The mouth of the receptor channel viewed from the synaptic cleft based on reconstruction from cryoEm images. Addition of acetylcholine, which binds to the two α subunits, induces small rotations in the five subunits of the $\alpha_2\beta\gamma\delta$ complex causing the channel to open. From Unwin.[640] (B) Architecture of the subsynaptic membrane and the acetylcholine receptor. The binding of acetylcholine and the movement of cations through the open channel is illustrated. Cations that leave the cytoplasm may be filtered through narrow openings that lead into the central channel, which is formed by transmembrane helices. Negatively charged amino acid residues may help exclude anions from the region of the pore. From Miyazawa *et al.*[624] (A) and (B) Courtesy of Nigel Unwin. (C) Stereoscopic ribbon drawing of one subunit of a pentameric acetylcholine-binding protein, which mimics the receptor structure. Disulfide bonds are shown in a ball-and-stick form. The N terminus in a receptor would point toward the synaptic cleft and the C terminus would continue at the bottom into the transmembrane helix. Courtesy of Brejc *et al.*[627]

Cholinergic receptors and their agonists and antagonists. Among the acetylcholine-releasing (cholinergic) neurons are the motor neurons that form synapses at neuromuscular junctions, the preganglionic neurons of the entire autonomic system, and the postganglionic neurons of the parasympathetic system. There are also many cholinergic synapses within the brain. In contrast, in insects neuromuscular transmission is mediated by glutamate while acetylcholine is the principal neurotransmitter in the central nervous system.[619]

Important in the study of neurotransmitters is the identification of specific agonists, which mimic the action of a transmitter, and of antagonists, which block the action of the transmitter. Two groups of compounds influence acetylcholine-secreting neurons, leading to the classification of these neurons either as **muscarinic** (activated by muscarine; Fig. 30-22) or **nicotinic** (stimulated by nicotine). The muscarinic receptors, which are found in many autonomic neurons, are specifically inhibited by **atropine** and **decamethonium** (Fig. 30-22). The nicotinic synapses occur in ganglia and skeletal muscle. They are inhibited by curare and its active ingredient **D-tubocurarine** (Fig. 30-22) and by the protein snake venom **α-bungarotoxin**. This toxin has been used to titrate the number of acetylcholine receptors in the motor end plate of the rat diaphragm. About 4 x 10^7 receptors per end plate (or 13,000 / μm^2) were found.[620]

Nicotinic receptors (nAChRs; Fig. 30-23) of the type found in neuromuscular junctions are most frequently isolated from the electric organs of the electric eel *Electrophorus* or from electric fish of the genus *Torpedo*. They have been studied more intensively than any other receptor.[621–626a] They contain four kinds of subunit with a stoichiometry $\alpha_2\beta\gamma\sigma$ and molecular masses of 39, 48, 58, and 64 kDa respectively. The amino acid sequences of the four proteins contain homologous regions, some of which are thought to represent membrane-spanning segments of the peptides. These receptors are ligand-gated ion channels and are closely similar to $GABA_A$ and $GABA_C$ receptors, to glycine receptors, and to 5-hydroxytryptamine (serotonin) receptors of the $5\text{-}HT_3$ type. Parts of their amino acid sequences are also homologous to those of both the voltage-gated Na^+ channels and gap junctions,[433,627] suggesting that the transmembrane domain may resemble that of Fig. 30-18.[628] However, notice the difference in symmetry. Acetylcholine binds to the two α subunits (Fig. 30-23). Neurotoxins may bind at several sites.[629] Some indication of the possible function of the various subunits comes from studies of the neuromuscular junction in which the different subunits are degraded at different rates with half-lives of from one to ten days. During development fetal ε subunits are replaced by adult γ subunits. Perhaps more rapid changes in receptor composition are sometimes needed.[630]

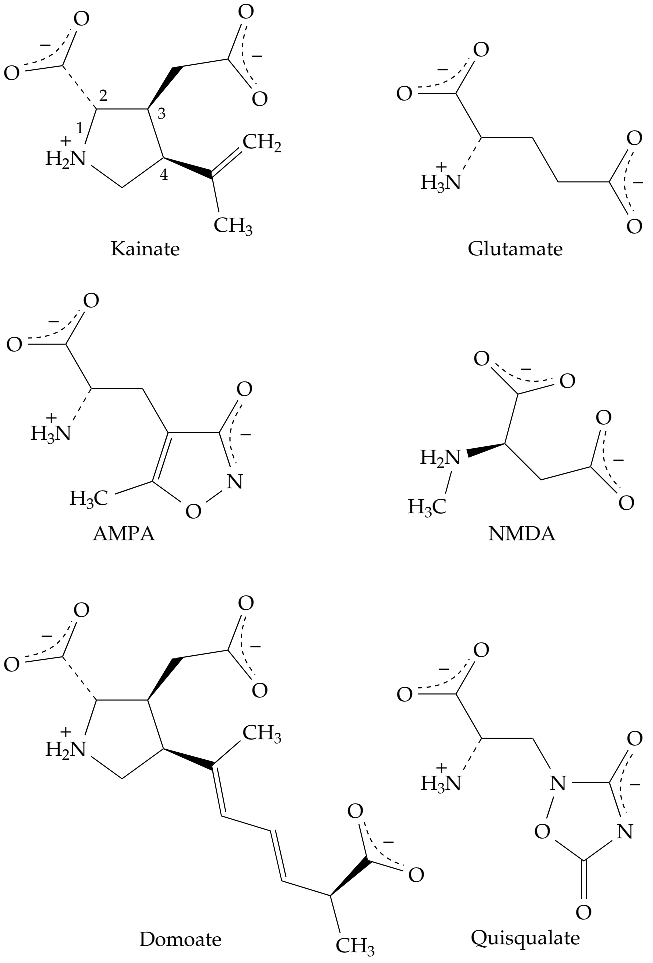
Similar nAChRs are also found in the brain.[621,631,632] However, they are not identical but have at least 17 differing amino acid sequences (α1–α10, β1–β4, γ, δ, and ε). The neuromuscular junction receptor (muscle type) from fish is described as $(\alpha1)_2 \bullet \beta1 \bullet \gamma/\varepsilon \bullet \delta$.[626] The brain contains homopentamers of subunits α7, α8, and α9 as well as various heteropentamers. The various forms possess different affinities for acetylcholine and for antagonists such as nicotine.[633,634] In the brain the highest affinity for nicotine is shown by an α4β2 form, which represents over 80% of the nAChR in mammalian brain.[634,635] Knockout mice in which the β2 subunit gene has been deleted lose their sensitivity to nicotine.

Conductance measurements showed that the nicotinic receptors contain channels permeable to Na^+ and other cations and that they are acetylcholine-gated ion channels. Construction of a three-dimensional image from electron micrographs at various angles of tilt shows a tube with approximate pentagonal symmetry and a narrow channel through the center (Fig. 30-23).[622,624,636] Acetylcholine binds to sites on the two α subunits ~3 nm away from the ion channel. An allosteric change opens the channel, allowing cations (largely Na^+) to flow out, depolarizing the membrane. There are at least four structural states in the channel opening-and-closing cycle.[637,638] The three-dimensional structure has been modeled using an acetylcholine-binding protein of known structure from a snail[626,627,639] as a mimic of the cytoplasmic nicotine-binding domain of the receptor. The structure of one subunit of the binding protein is shown in Fig. 30-23C. This protein, which is secreted into synapses by glial cells, may provide a buffering action by binding the acetylcholine. Although the most rapid effect of acetylcholine binding to the nicotinic receptor is depolarization of the postsynaptic membrane, other slower effects follow. Thus, protein kinases are activated and phosphorylate the receptor as well as other proteins.[641]

After a pulse of transmitter is released, it must be removed or inactivated quickly to prepare the synapse for arrival of a new nerve impulse. This is accomplished in two ways in cholinergic synapses. The first is via hydrolytic destruction by acetylcholinesterase[642–645] (pp. 634–637; Eq. 12-25). This esterase and the related butyrylcholinesterase[646] are present in the synaptic membrane itself. The second mechanism is energy-dependent transport of acetylcholine into the neuron for reuse. Since much of the transmitter is hydrolyzed, new acetylcholine is synthesized by transfer of an acetyl group of acetyl-CoA to choline.[647]

In the central nervous system muscarinic acetylcholine receptors are more abundant than nicotinic receptors. They consist of single-chain proteins of mass ~70 kDa. They are not ion channels but are 7-helix receptors homologous in sequence with β-adrenergic receptors (Fig. 11-6) and with rhodopsin.[648] Five different subtypes (M1–M5) have been characterized. The M1, M3, and M5 receptors are coupled to the G_q/G_{11} family of G proteins (pp. 557–558), and M2 and M4 are coupled to G_i/G_o proteins.[649–651] Their effects are slower and of longer duration than those of the nicotinic receptors. It has been difficult to assign functions to the individual types. Most regions of the brain contain more than one type, but they are thought to be involved in locomotion, learning, memory, thermoregulation, and cardiac and pulmonary functions. Many drugs, some of which are used in treatment of Parkinson and Alzheimer diseases, epilepsy, and asthma, affect muscarinic receptors. The M2 receptors predominate in the heart where they help to regulate the beating frequency and atrial contractility. Sudden infant death may sometimes result from a defect in muscarinic receptors.[652] Knockout mice lacking M2 receptors also have problems with movement control, body temperature, and pain responses.[651] Mice lacking M3 receptors are lean with very low levels of serum leptin and insulin.[653] Many of the muscarinic receptors activate adenylate cyclase, while others are coupled to the phosphoinositide cascade. Some indirectly activate K^+ channels.[654] Muscarinic receptors are also studied in insects, but it is difficult to correlate the insect and mammalian receptors.[655]

Amino acids as neurotransmitters. The concentrations of **glutamate** and of its decarboxylation product **γ-aminobutyrate** (GABA) are high in all regions of the brain. The two compounds are generated sequentially in the γ-aminobutyrate shunt, a pathway that accounts for a quantitatively significant part of the total metabolism of the brain (Fig. 17-5). Because they are present in all parts of the brain in high concentrations, there was initially reluctance to accept glutamate and GABA as neurotransmitters. However, it is now accepted that L-glutamate is the major excitatory transmitter in the central nervous system.[656–658] It seems to be responsible for nearly all of the very fast acting nerve impulses in the brain. At the same time GABA is recognized as the most important inhibitory transmitter. The role of glutamate as an excitatory transmitter was first established for the neuromuscular junction of arthropods.[659] Although it is a constituent of all animal tissues, the concentration of glutamate is much higher in brain than in other tissues, and it is higher in neurons than in glia. Microiontophoretic application of either glutamate or aspartate to the brain cortex leads to very strong excitatory responses.



Kainate

Glutamate

AMPA

NMDA

Domoate

Quisqualate

Figure 30-24 Chemical structures of some agonists of ionotropic glutamate receptors (iGluR).

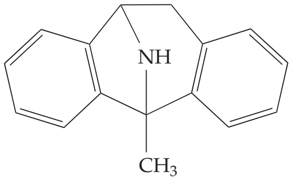
Three subtypes of ionotropic glutamate receptors (iGluR) are named for the specific agonists **α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid** (AMPA), ***N*-methyl-D-aspartate** (NMDA), and **kainate**. The receptors resemble the acetylcholine receptor in containing a cation channel.[149,660–662] In addition, there are 7-helix **metabotropic glutamate receptors,** which are coupled to G proteins.[663,664] The AMPA receptors were in the recent past called **quisqualate** receptors, because they are also activated by the agonist with that name. The toxic domoate (Fig. 30-24) also binds to kainate receptors. Both domoic acid and kainic acid are terrible convulsant toxins. They are formed by two different red algae. Domoic acid accumulates in contaminated mussels and causes shellfish poisoning. The ionotropic glutamate receptors, which may be stimulated by either glutamate or aspartate, are directly linked to the opening of cation channels. Their activation may also induce the inositol phosphate cascade and slower Ca^{2+}-dependent changes. A peculiarity of the high-conductance NMDA channels is that they are blocked by Mg^{2+} in a voltage-dependent manner. They do not open unless the frequency of nerve impulses is high or some other factor causes membrane depolarization.[656]

The AMPA receptors, which are thought to be the predominant mediators of fast excitatory transmission in the brain,[665] are oligomers (probably tetramers[666,666a]) of 950- to 1500-residue protein subunits. These subunits have large N-terminal domains in the synaptic cleft. There are probably three transmembrane helices and a membrane-associated loop similar to those depicted in Fig. 30-17A. A long C-terminal tail protrudes into the cytosol, while a large loop between transmembrane regions extends from the outer membrane surface, joining with the N-terminal domain to form the ligand-binding site, the structure of which resembles those of bacterial periplasmic binding proteins.[661,665,667] Four related AMP receptors, designated GluR1, 2, 3, and 4, have been identified. Related kainate receptors, whose properties overlap those of AMPA receptors, are designated GluR5, 6, and 7.[662] Although AMPA receptors are essential for fast signal transmission they lose sensitivity rapidly (on a millisecond time scale) as a result of conformational alterations.[667a] Many factors, including inhibition by polyamines,[667b] affect these receptors. However, brief high-frequency activation of some AMP receptors leads to a long-lasting increase in efficiency, termed LTP, which is important to learning (see p. 1801).[666a]

The NMDA receptors are heterooligomers with two type of subunits. The NR1 (or ζ) subunits exist as a series of at least eight splice variants. The NR2A, B, C, and D (ε series) are encoded by four different genes.[668,669] NR1 is regarded as the principal subunit ard NR2 as a regulatory subunit. As with the AMPA receptors[670] the oligomeric NMDA receptors are anchored at appropriate locations in the postsynaptic membrane by scaffolding proteins containing PDZ domains (Table 7-3).[671] The C-terminal domains of the ε subunits are unusually long and participate in anchoring. NMDA receptors are found not only in neurons but also in astrocytes (Fig. 30-20), where they are thought to have important signaling functions.[672,673] These include regulation of Ca^{2+} flow, in part via gap junctions.[603a]

The N-terminal domain of the NR1 subunit of the NMDA receptor contains a glycine-binding site.[674] Full activity of the receptor requires a **coagonist** bound in this site. Surprisingly, **D-serine** seems to be the normal coagonist, at least in some sites.[675,676] This newly recognized neurotransmitter is synthesized from L-serine by a pyridoxal phosphate-dependent recemase and is destroyed by the flavoprotein D-amino acid oxidase. Associated with NMDA receptors are clusters of **ephrin receptors**, proteins that bind the glycosylphosphatidylinositol (GPI)-anchored proteins known as ephrins in presynaptic membranes. Binding of ephrins to their postsynaptic receptors activates tyrosine kinases and enhances the influx of Ca^{2+} ions.[676a,b]

Specific inhibitors of NMDA channels include a 27-residue "spasmotic" conotoxin,[490] 2-amino-4-phosphonobutyrate, related longer chain aminophosphonates, and the following potent anticonvulsant drug, which is able to penetrate the blood–brain barrier.[677]



(+)5-Methyl-10,11-dihydro-5*H*-dibenzo[*a*,*d*]cyclohepten-5,10-imine

Metabotropic glutamate receptors have been classified into eight types (mGluRs1–8).[678–680a] Group I (mGluRs1–5) are selectively activated by 3,5-dihydroxyphenylglycine; Group II (mGluR2 and mGluR3) are activated by L-2-(carboxycyclopropyl)glycine; and Group III (mRluR4 and mGluR 6–8) are activated by L-2-aminophosphonobutyrate. They are all 7-helix G-protein-coupled receptors with external ligand-binding domains that resemble those of bacterial periplasmic binding proteins.[680] Splice variants for at least mGlnR1 are known.[678] Metabotropic glutamate receptors are neuromodulary but nevertheless play essential roles in the cerebellum and other parts of the brain. For example, mice deficient in the mGluR1 protein have severe problems with motor coordination and learning.[681,682] Metabotropic glutamate receptors may participate in calcium sensing and signaling.[683,684]

Synaptosomal particles have a high-affinity proton-dependent uptake system for glutamate.[685] Glutamate and aspartate may also be taken up from the synaptic cleft by neurons or by glial cells, which then transfer the glutamate into neurons for reuse.[686,687] Five distinct mammalian transporter genes have been cloned.[688] They are driven by concentration gradients of Na^+ and K^+ across the membrane.[689,690] However, some serve as glutamate-gated chloride ion channels.[691,691a]

Excitotoxicity. As essential as glutamate is for brain function it is toxic in excess. Excessive stimulation of the NMDA receptors, which occurs during convulsions, strokes, or traumatic injury and which can accompany anoxia or hypoglycemia, causes neuronal death.[660,692–694] Blocking these receptors with the above-mentioned anticonvulsant drug or aminophosphonates has a remarkable protective effect against the neurotoxicity of the accumulating glutamate.[658,677] Vitamin E and **tocotrienols** (Fig. 15-24) may also be protective.[695]

The inhibitory neurotransmitter gamma-aminobutyrate (GABA). Glutamate, aspartate, and cysteic acid are all potent excitors, but their decarboxylation products γ-aminobutyrate (**GABA**), β-alanine, and taurine are inhibitors as is also glycine. Of these GABA is the most important.[696] Its concentration in the brain is high and varies at least threefold in different parts of the brain. It is hardly present elsewhere in the body. GABA and GABA-binding sites are found in 30–50% of the nerve endings. The function as an inhibitory transmitter has also been demonstrated in inhibitory neurons present in the peripheral nervous system of arthropods. Virtually every neuron in the brain is to some extent subject to inhibition by GABA.[697,698] Glial cells also have GABA receptors.

The receptors for GABA are divided into type A, which are blocked by **bicuculline**,[699] and type B, which are stimulated by **baclofen** (Fig. 30-25).[698] The $GABA_A$ receptors are the major sites of fast synaptic inhibition in the central nervous system.[700] They are structurally related to the nicotinic acetylcholine, glycine, and serotonin type 3 ($5\text{-}HT_3$) receptors. Cloning has revealed 16 different mammalian subunits: α1–α4, β1–β3, γ1–γ3, δ, ε, π, and Φ.[701–704a] The oligomeric receptors are ligand-gated chloride ion channels[481,705] as are also glycine receptors. These receptors are clustered in synaptic membranes, apparently anchored in part by their β subunits[706] and scaffold proteins such as the microtubule-binding **gephyrin** (from the Greek word for bridge)[701,707] and a small ~14-kDa GABA receptor-associated protein.[708] A novel serine protein kinase is also associated with GABA receptors.[703]

Whereas excitatory transmitters lead to depolarization of the postsynaptic membrane, inhibitory transmitters cause **hyperpolarization**, apparently by increasing the conductance of K^+ and Cl^-. The result is

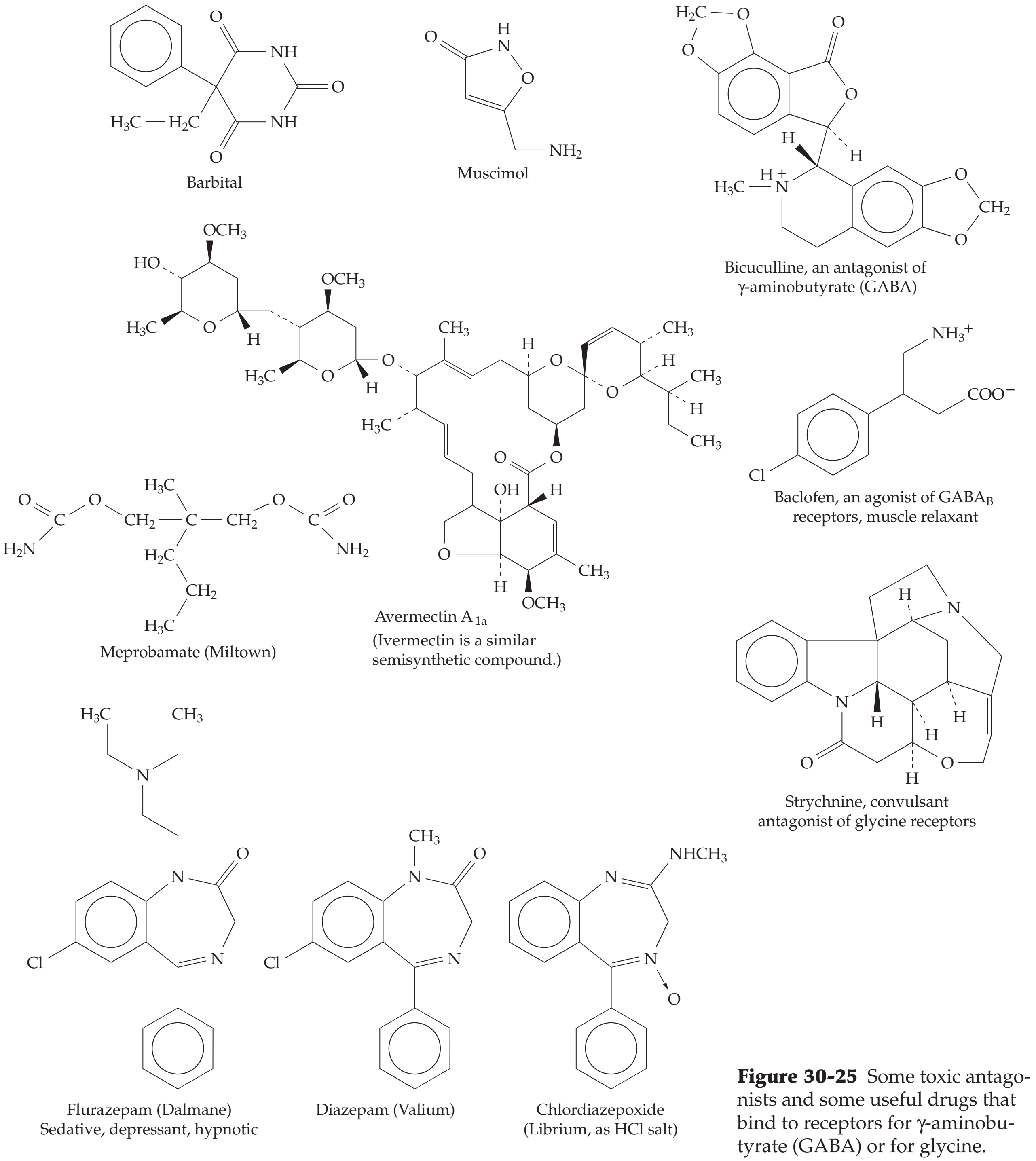

Barbital
Muscimol
Bicuculline, an antagonist of γ-aminobutyrate (GABA)
Avermectin A 1a (Ivermectin is a similar semisynthetic compound.)
Baclofen, an agonist of GABAB receptors, muscle relaxant
Meprobamate (Miltown)
Strychnine, convulsant antagonist of glycine receptors
Flurazepam (Dalmane) Sedative, depressant, hypnotic
Diazepam (Valium)
Chlordiazepoxide (Librium, as HCl salt)

Figure 30-25 Some toxic antagonists and some useful drugs that bind to receptors for γ-aminobutyrate (GABA) or for glycine.

that it is more difficult to excite the postsynaptic membrane in the presence of, than in the absence of, these transmitters. GABA-dependent interneurons also contain the calcium-binding **parvalbumin** (Fig. 6-7), which suggests that a Ca^{2+}-dependent process is involved.[709]

The $GABA_B$ receptors resemble metabotropic glutamate receptors.[710,711] They are 7-helix G-protein coupled proteins, which activate adenylate cyclase. They tend to dimerize, and maximum activity is observed for heterodimers of $GABA_B1$ and $GABA_B2$ receptors.[712,713] They are often coupled to inward rectifying K^+ channels.[714]

The GABA receptors provide binding sites for a great variety of toxins and drugs.[481] These include barbiturates, anesthetics, antianxiety drugs, and the insecticides such as toxaphine, cyclodienes, and pyrethroids.[481] **Diazepam**, **chlordiazepoxide**, and flurazepam[700,702,715–717] (Fig. 30-25) are antianxiety drugs and muscle relaxants, which, during the 1970s, were the most frequently prescribed drugs in the United States.[716] Binding of benzodiazepines to GABA receptor-chloride channels enhances the effect of GABA. The drugs induce relaxation but can interfere with memory, reduce concentration, and cause physical clumsiness. They may also intensify the effects of alcohol and can be addictive.[718]

Specific antagonists for $GABA_A$ receptors include the alkaloid convulsants bicuculline (Fig. 30-25)[699] and **picrotoxin** (Fig. 22-4) and the convulsant terpenoid compound **thujone** (Fig. 22-3), which is present in the wormwood plant *Artemesia absinthium*. Thujone is present in the liqueur absinthe, which was the national drink of France in the late 19th century but, because of its toxicity, has been illegal in most countries since ~1915.[719]

GABA enters synaptic vesicles via a vesicular GABA transporter, an integral membrane protein whose gene has been found in *Caenorhabditis elegans*.[720] Termination of GABA neurotransmission is accomplished by rapid Na^+-dependent uptake into neurons for reuse and uptake into glial cells.[721,722] Excess GABA is continuously oxidized to succinic semialdehyde by GABA aminotransferase[723] in the GABA cycle of Fig. 17-4. Notice the manner in which this cycle incorporates synthesis of both of the neurotransmitters glutamate and GABA. Glutamine also functions in neurons, perhaps serving as a buffer for glutamate.

The hereditary triple-repeat disease Huntington's chorea (**Huntington disease**), with an incidence of 5–10 per 100,000 persons, affects principally persons of age over 40 and is associated with a deficiency of GABA in basal ganglia.[724] The cortex is also affected. Severe neurologic symptoms arise as a result of premature death of neurons in the basal ganglia. Convulsions may also arise because of a deficiency of GABA in the brain.

Glycine. Glycine appears to be the most important neuroinhibitor in the spinal cord and brainstem. It is present at concentration of 3–5 mM in the spinal cord and in the medulla but is low in the cerebral cortex. **Strychnine** (Fig. 30-25) is a specific antagonist of glycine receptors in spinal synapses.[725] Ivermectin (Fig. 30-25) also blocks glycine Cl^- channels.[726] A mutant mouse called *spastic* is deficient in glycine receptor function. A small dose of strychnine produces an effect on a normal mouse that resembles the effect of this mutation.[727,728] A similar disorder affects some Hereford calves.[729] Strychnine-binding studies have suggested a deficit of glycine receptors in human spasticity and in the loss of motor control associated with **Parkinson disease** and **amyotrophic lateral sclerosis**.[725] A human **startle disease**, which causes an exaggerated muscular response to unexpected stimuli, also results from reduced glycinergic neurotransmission.[730]

Most glycine receptors are Cl^- ion channels that open in response to transmitter binding.[725] The strychnine-binding subunit shows significant homology with the nAChR proteins,[725] and the overall structures resemble those of GABA receptors and of nAChRs.[731,732,732a] Human $\alpha1$–$\alpha4$ and β subunits have been identified.[733,734] Two integral membrane glycine transporters are known.[735–737]

Anesthetics. Several types of neurotransmitter receptors provide binding sites for anesthetics. Some anesthetics are molecules of moderate size, e.g., **barbiturate** derivatives, while others, such as **diethyl ether** or **halothane** ($CF_3CHClBr$), are very small. The latter is one of the most widely used inhalation anesthetics. Both Mg^{2+} and Mn^{2+} are also powerful CNS depressants and can cause general anesthesia. It has often been proposed that the effectiveness of anesthetics is related to solubility in lipids, but it has been difficult to pinpoint a site of action. Now it is clear that specific synaptic proteins often provide the binding sites for anesthetics. Important among these are the glycine receptors.[715,738,739] GABA receptors[740,740a] and kainate glutamate receptors may also bind anesthetics.[741]

Adrenergic synapses: the catecholamines. The three closely related tyrosine metabolites, **dopamine**, **noradrenaline**, and **adrenaline**, known collectively as catecholamines, are important products of neuronal metabolism.[149,393] Dopamine and noradrenaline serve as neurotransmitters. Catecholamine-containing neurons are found throughout the brain, including the cortex and cerebellum regions. Very large dopamine-containing neurons are present in the brains of gastropod molluscs.[742] In the human brain a prominent series of dopamine neurons run from the substantia nigra to the caudate nuclei and putamen of the striatum, the **nigrostriatal** pathway (Fig. 30-12).[149,743,743a] In many invertebrates **octopamine**,[744–746] which is synthesized via tyramine (Fig. 30-26), apparently functions in place of noradrenaline. Note the precursor–product relationship between dopamine, noradrenaline, and adrenaline. The synthetic pathways to these neurotransmitters involve decarboxylation and hydroxylation, types of reaction important in formation of other transmitters as well. The most important process for terminating the action of released catecholamine transmitters is reuptake by the neurons. High-affinity uptake systems transport the catecholamine molecules back into the neurons and then into the synaptic vesicles. The uptake is specifically blocked by the drug **reserpine** (Fig. 25-12).[746a] The dopamine transporter is a major binding site for cocaine (see Fig. 30-28).[747–751] Catecholamine transmitters are catabolized by two enzymes. One is the

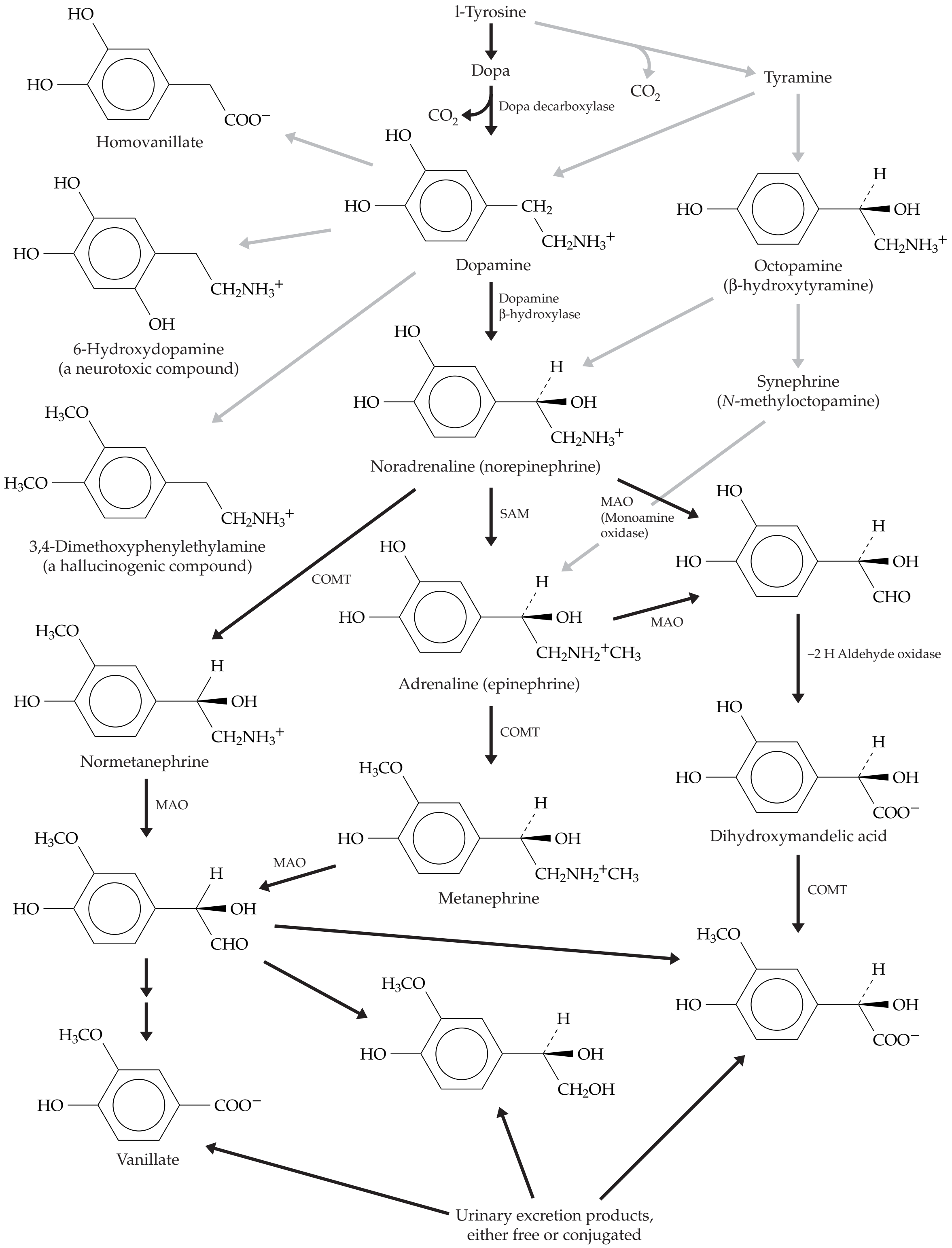
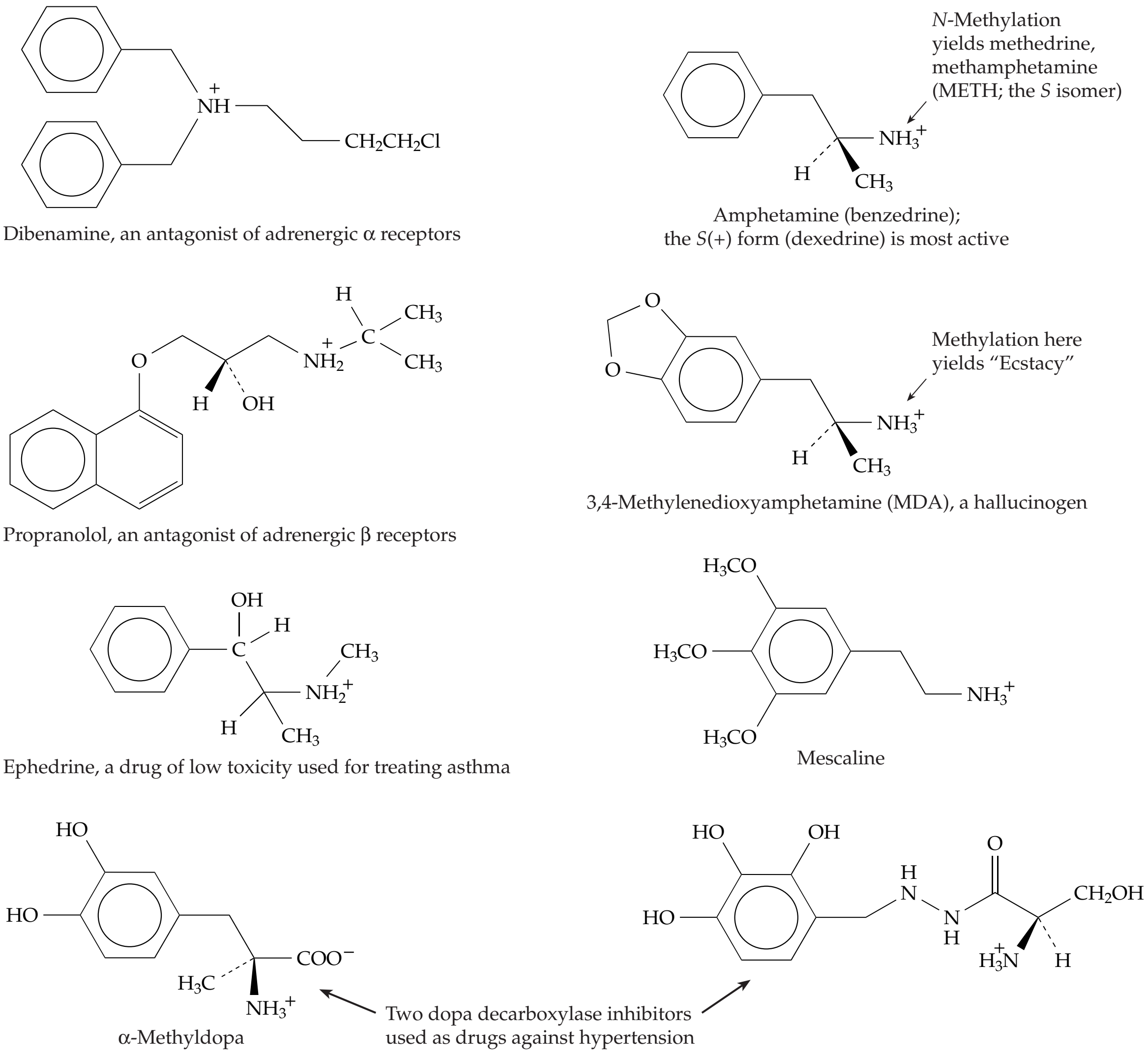

l-Tyrosine
Dopa
CO2
Dopa decarboxylase
CO2
Tyramine
HO
HO
COO−
Homovanillate
HO
HO
CH2
CH2NH3+
Dopamine
HO
H
OH
CH2NH3+
Octopamine
(β-hydroxytyramine)
HO
HO
CH2NH3+
OH
6-Hydroxydopamine
(a neurotoxic compound)
Dopamine
β-hydroxylase
HO
HO
H
OH
CH2NH3+
Noradrenaline (norepinephrine)
Synephrine
(N-methyloctopamine)
H3CO
H3CO
CH2NH3+
3,4-Dimethoxyphenylethylamine
(a hallucinogenic compound)
SAM
MAO
(Monoamine
oxidase)
COMT
HO
HO
H
OH
CHO
HO
HO
H
OH
CH2NH2+CH3
Adrenaline (epinephrine)
MAO
–2 H Aldehyde oxidase
H3CO
HO
H
OH
CH2NH3+
Normetanephrine
COMT
MAO
HO
HO
H
OH
COO−
Dihydroxymandelic acid
H3CO
HO
H
OH
CH2NH2+CH3
Metanephrine
H3CO
HO
H
OH
CHO
MAO
COMT
H3CO
HO
H
OH
COO−
H3CO
HO
COO−
Vanillate
H3CO
HO
H
OH
CH2OH
Urinary excretion products,
either free or conjugated

Figure 30-26 Some pathways of metabolism of the catecholamines. See also Fig. 25-5.

flavoprotein **monoamine oxidase** (**MAO**; Chapter 16), an enzyme present within the mitochondria of neurons as well as in other cells in all parts of the body.[752] The second enzyme is **catechol-*O*-methyltransferase** (**COMT**; Eq. 12-3), which is found in postsynaptic membranes as well as in liver, kidney, and other tissues. It apparently provides the principal means of inactivating circulating catecholamines. In a process that occurs in all organisms sulfo groups are transferred from PAPS (Eq. 17-38) onto hydroxyl groups of catecholamines, steroid compounds, and proteins (p. 659). Sulfation of catecholamines is relatively specific to humans.[752a,b]

Both adrenaline and noradrenaline stimulate smooth muscles throughout the body and have a hypertensive effect. Their postsynaptic receptors are 7-helix transmembrane proteins (Fig. 11-6). A comparison of the effects of various analogs led to the classification of these receptors into classes α, α_2, β, and β_2, which are discussed briefly on pp. 553–555. The α receptors, which are structurally closely related to rhodopsin,[753,754] are coupled via Gq / 11 proteins to a phosphoinositide-activated phospholipase C (Figs. 11-9, 30-19).[755] They usually provoke an excitatory response. However, in intestinal smooth muscles they are inhibitory. Adrenaline is usually more active at α receptors than is noradrenaline. A specific antagonist is **dibenane** (Fig. 30-27). The β receptors usually induce muscular relaxation but cause myocardial stimulation. Noradrenaline is usually more active than adrenaline. In most cases the β receptors of the postsynaptic membrane respond to the neurotransmitter by causing a hyperpolarization of the cell membrane and inhibition of nerve impulses. A specific antagonist is **propranolol** (Fig. 30-27). The β receptors are coupled via proteins of the G_s family (pp. 557, 558). The β2 receptors have received special attention because of their importance to heart and pulmonary functions. Both heart failure and asthma are associated with poor β2 receptor function.[756,757] The β2 receptors affect many other processes including insulin action.[758] Intense efforts are being made to understand them at the structural level.[757,759,760] Of special interest are the mechanisms by which receptors are densensitized after passage of impulses, a process that often involves multiple phosphorylation reactions[761]



Dibenamine, an antagonist of adrenergic α receptors

N-Methylation yields methedrine, methamphetamine (METH; the *S* isomer)

Amphetamine (benzedrine); the *S*(+) form (dexedrine) is most active

Propranolol, an antagonist of adrenergic β receptors

Methylation here yields "Ecstacy"

3,4-Methylenedioxyamphetamine (MDA), a hallucinogen

Ephedrine, a drug of low toxicity used for treating asthma

Mescaline

α-Methyldopa

Two dopa decarboxylase inhibitors used as drugs against hypertension

Figure 30-27 Some agonists and antagonists of adrenergic synapses (shown as cations in most cases).

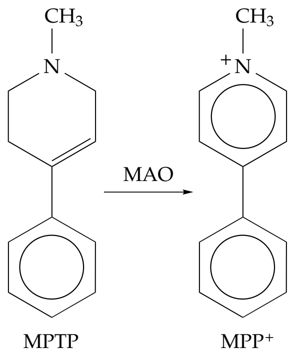
as well as interaction with **arrestin** (Fig. 23-43) and receptor internalization.[762]

Attention has been focused on dopamine because of its relationship to neurological diseases and to addiction (discussed in Section 10). Dopamine receptors constitute a large family, which are classified into two main subfamilies. The D_1 subfamily consists of D_{1A} and D_{1B} (D_5) receptors and the D_2 subfamily of D_2, D_3, and D_4 receptors.[763,764] The D_1 receptors, which are prominent in the prefrontal cortex and also in the striatum, are more abundant than the D_2 receptors, which are also present in the striatum and the pituitary and are targets for antipsychotic drugs such as **haloperidol** (Fig. 30-33).[765] The recently discovered and less numerous D_3 receptors are present in only a few regions of the brain. However, a deficiency of D_3 receptors may also be involved in addiction, schizophrenia, and Parkinson disease.[766,767]

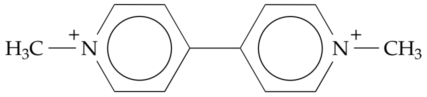
The role of the catecholamines as transmitters in the sympathetic nervous system and in the peripheral ganglia has been well established, but the function in the central nervous system is less clear. Catecholamines are present in varying quantities throughout the brain, and histochemical techniques[149,768] have made it possible to visualize both dopamine and noradrenaline-containing neurons by the green fluorescence produced from reaction with formaldehyde or glyoxylate.[769] The reactions are presumably analogous to those in Fig. 25-10. Another method for tracing dopamine receptors in the central nervous sytem is through labeling with specific antibodies to dopamine-β-hydroxylase (Eq. 18-53), the enzyme that converts dopamine to noradrenaline, to tyrosine hydroxylase, or to other specific neuronal enzymes.[770]

Parkinson disease. Neurons of the nigrostriatal pathway degenerate in Parkinson disease, a condition accompanied by severe tremors and rigidity. The significance of dopamine was illustrated by the finding that the precursor amino acid **L-dopa** caused dramatic improvement in many persons with Parkinson disease.[771] Dopamine and other catecholamines do not cross the blood–brain barrier but L-dopa does. This leads to an increase in the dopamine level in the basal ganglia of the brain, which apparently compensates for the deficiency resulting from the neuronal degeneration.

In 1982 a number of young people in California injected themselves with an illegally manufactured opiate drug that was subsequently found contaminated with *N*-methyl-4-phenyltetrahydropyridine (MPTP). Within a few days they developed irreversible symptoms of Parkinson disease. Subsequent investigation revealed that MPTP itself is not toxic but that it is oxidized by monoamine oxidase B (MAO-B) to the corresponding pyridinium derivative MPP^+ (Eq. 30-4). It is this pyridinium derivative, or perhaps related free radicals, that is toxic.[772] MPP^+ is readily taken up by mitochondria and is apparently concentrated in the mitochondria of the nigrostriatal cells to a toxic level.[773] The MAO inhibitor pargyline (Fig. 30-33) interferes with the oxidation of Eq. 30-4 and prevents development of Parkinson disease in squirrel monkeys exposed to MPTP.[774] These results suggested possible environmental causes for Parkinson disease and also a new approach to treatment.[775,776] MPP^+ has been marketed as a herbicide, and it has a close structural relationship to another herbicide, **paraquat**.



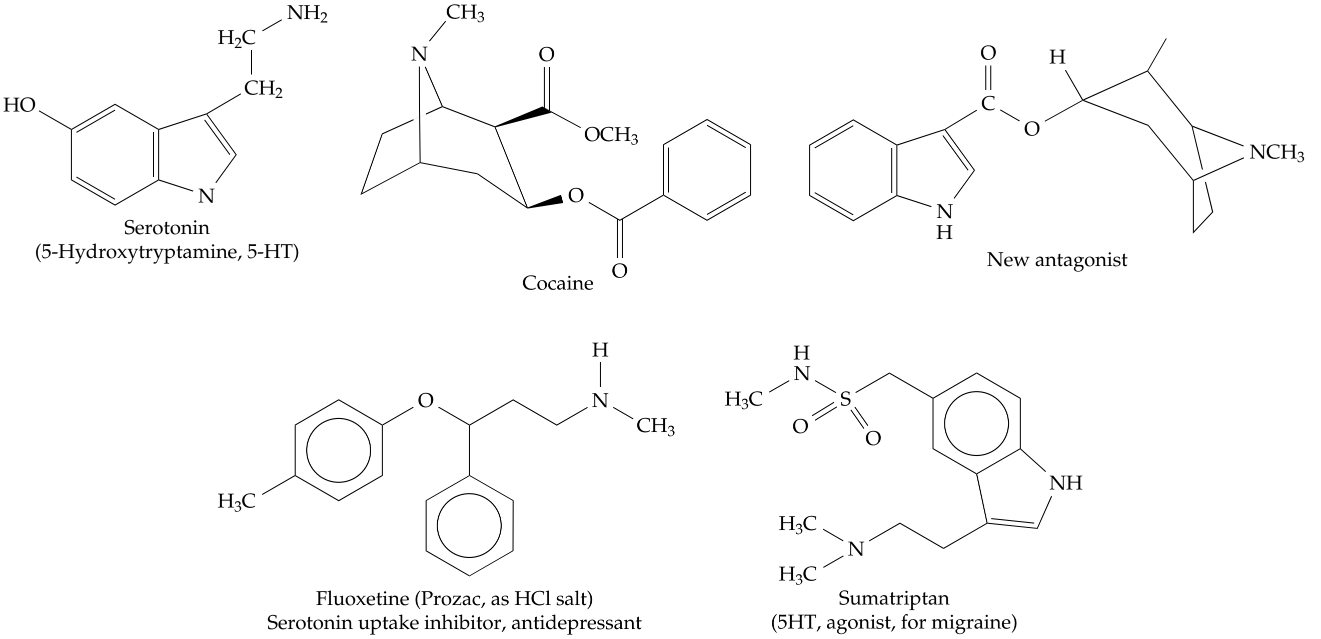
(30-4)



Paraquat

Many food constituents including peppermint, spearmint, and tea contain 4-phenylpyridine, another close relative.[775] While administration of L-dopa to replace the deficit in the basal ganglia seemed the ideal treatment for Parkinson disease, mental deterioration is not stopped, and for some patients the drug loses its effectiveness in about three years. Based on the new information about MPTP, treatment with extra vitamin E as an antioxidant along with an MAO inhibitor is being tested as a way to prevent further damage from environmental toxins.[776]

Serotonin and melatonin. The indolealkyl amine serotonin (5-hydroxytryptamine, 5-HT; Fig. 30-28), is found in all mammalian brains and in invertebrates as well. Its distribution in the brain is limited, serotonin-containing neurons being found in the raphe nuclei of the brainstem from which they ascend into the brain and down the spinal cord. Serotonin-containing neurons have been traced within brains of snails using ^3H-labeled serotonin.[777] Studies with these simpler brains have revealed both inhibitory and excitatory responses to these neurons. Serotonin-accumulating neurons are also found in the retina[778] and are widely distributed in the peripheral nervous



Serotonin
(5-Hydroxytryptamine, 5-HT)

Cocaine

New antagonist

Fluoxetine (Prozac, as HCl salt)
Serotonin uptake inhibitor, antidepressant

Sumatriptan
(5HT, agonist, for migraine)

Figure 30-28 Serotonin (5-hydroxytryptamine) and some drugs that affect receptors and transporters.

system.[779] Serotonin-containing granules are present in blood platelets.[780]

Serotonin appears to be involved in activation of pain fibers, when tissues are injured. **Cocaine** (Fig. 30-28) is a powerful pain killer and a weak antagonist of responses to serotonin, a fact that has led to the synthesis of new antagonists, such as the one in Fig. 30-28 whose structure encompasses that of both cocaine and serotonin.[779,781] It is active at a concentration as low as 10^{-14} M and is among the most potent known drugs of any type.

Serotonin is synthesized via tryptophan and 5-hydroxytryptophan with decarboxylation of the latter (Fig. 25-12). Within the **pineal body** of the brain and in the retina, serotonin is acetylated to *N*-acetylserotonin,[782,783] which is then *O*-methylated to **melatonin**, the pineal hormone (Fig. 25-12). A specific inhibitor of serotonin synthesis is *p*-chlorophenylalanine, and studies with this and other inhibitors suggest that serotonin is required for sleep.[784]

At least 14 distinct types of serotonin receptors (5-HT_{1A}, 5-HT_{1B}, 5-HT_{2A}, etc.) have been identified.[785,786] They are present in the heart, in gastrointestinal tissues, adrenal and other glands,[787] and bone[788] as well as in the brain. Drugs, such as **sumatriptan** (Fig. 30-28), which activate serotonin receptors are important in the treatment of **migraine**. This common disorder of serotonin metabolism is characterized by severe or moderately severe headache and a variety of other symptoms, which are frequently preceded by a visual aura.[789] Serotonin is removed from synapses via a transporter, which also contains the binding site of the widely used antidepressant **Prozac** (Fig. 30-28) and related drugs.[790–791b]

Serotonin and melatonin are evidently involved in maintenance of the 24-h circadian rhythm of the body (see Section 13).[792,792a] Melatonin regulates the sexual cycle in photoperiodic animals and influences the onset of puberty.[792–794] The serotonin content of the brain is influenced by the diet, being higher after a meal rich in carbohydrates. Serotonin may serve as a chemical message sent from one set of neurons to the rest of the brain, reporting on the nature of dietary intake.[784] Melatonin, which can readily form free radicals, may function as part of the body's antioxidant system.[795,796]

Other neurotransmitters. The abundant glutamate, GABA, and glycine are major neurotransmitters. Do other amino acids also function in the brain? Roles for L-aspartate and D-serine (p. 1785) have been identified, but it is very difficult either to discover or to disprove a neurotransmitter function for other amino acids. It is even more difficult for small amounts of various amines and small peptides that are present in the brain. **Taurine** (Fig. 24-25) is one of the most abundant free amino acids in animals and meets several criteria for consideration as both an inhibitory and an excitatory transmitter.[797,798] However, its function is still uncertain (see Chapter 24). **Homocysteic acid**, formed by oxidation of homocysteine, is a powerful neuroexcitatory substance, but its concentration in the brain is very low.[149] D-Aspartate is also present at high concentrations in the cerebellum, pituitary, pineal gland, and adrenal chromaffin cells. It appears to be a modulator of melatonin synthesis.[799,800]

Receptors for **histamine**, which probably acts as a neuromodulator,[801] occur in the brain.[802] Histamine is formed by decarboxylation of histidine (p. 745)[803] and is inactivated by histidine *N*-methyltransferase. Histamine is best known for its presence in mast cells,[804] components of the immune system that release histamine during inflammatory and allergic reactions (Chapter 31). However, histaminergic neurons of the hypothalamus extend throughout the whole forebrain,[805] and specific receptors have been found both in the brain and in peripheral tissues.[806] Several other amines that are formed by decarboxylation of amino acids are present in trace amounts but may have important functions, some of which may be related to psychiatric disorders. These include tyramine (from tyrosine), β-phenylethylamine (from phenylalanine), and tryptamine (from tryptophan). As previously mentioned, octopamine is also present in trace amounts in mammalian brains.[807]

ATP, **ADP**, and **adenosine** are among the purines that are present in some synapses and activate a variety of receptors. Adenosine receptors are blocked specifically by methylated xanthines such as caffeine (Fig. 25-18) and theophylline.[808–808b] A drug almost 10^5 times as potent as theophylline is 1,3-dipropyl-8-(2-amino-4-chlorophenyl)xanthine.[809] Adenosine receptors, which are present in large numbers in the hippocampus,[149] form functional complexes with metabotropic glutamate receptors.[678] Adenosine

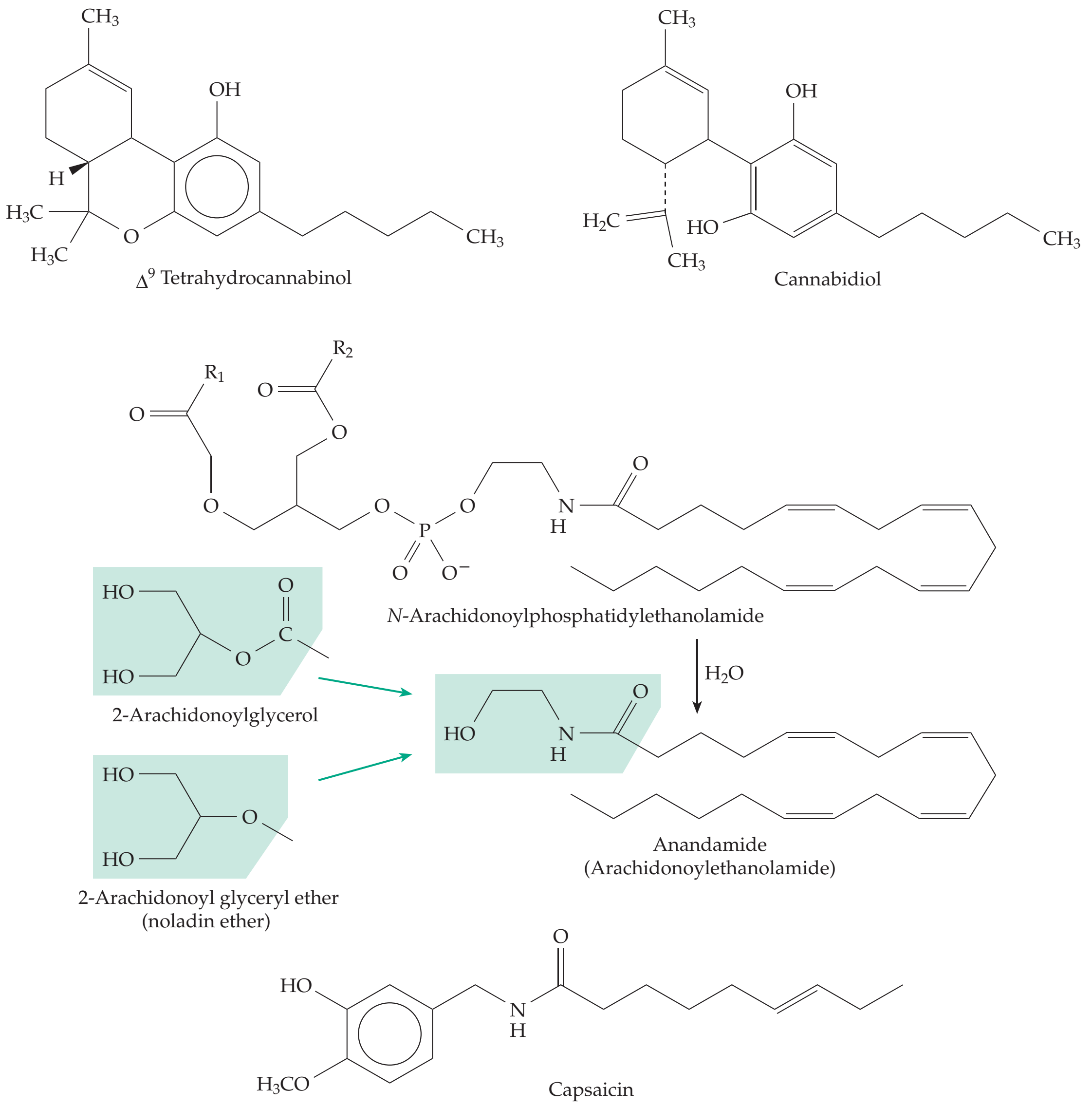

Δ9 Tetrahydrocannabinol
Cannabidiol
N-Arachidonoylphosphatidylethanolamide
H2O
2-Arachidonoylglycerol
2-Arachidonoyl glyceryl ether
(noladin ether)
Anandamide
(Arachidonoylethanolamide)
Capsaicin

Figure 30-29 Structures of the active components of cannabis, tetrahydrocannabinol, and cannabidiol, and structures of endogenous cannabinoids and of the vanilloid lipid capsaicin.

usually has a depressive effect. Craving for chocolate is often attributed to the methylxanthines present, but it may be a result of anandamide and related compounds.[810]

The occurrence of a variety of **neuropeptides** in the brain has been discussed in Section A. The first of these to be discovered[811] was the 11-residue **substance P** (Table 30-4), which was isolated in 1931. Like other neuropeptides it may function either as a transmitter or neuromodulator or perhaps both. Substance P, as well as many other neuropeptides, has been localized to specific neurons. Along with somatostatin, CCK, and enkephalins, it is found in high concentrations in the basal ganglia. Enkephalin and substance P are also found in specific neural elements in the visual system of lobsters.[812] In some cases a neuron contains both synaptic vesicles containing a major neurotransmitter and also vesicles containing a peptide or other cotransmittor. The peptide pituitary hormones ACTH, MSH, and vasopressin as well as the hypothalamic neurohormones may have effects on learning and behavior.[813]

Lipid mediators in the brain. The brain is rich in phospholipids, glycolipids, and long-chain unsaturated fatty acids. Many signaling functions seem likely, and some are discussed in Section A,7. Prostaglandin D_2 is a major prostanoid in the brain, which induces both hypothermia and sleep.[814] As mentioned in Section A,7, **oleamide** also induces sleep, perhaps by modulating the effects of 5-HT receptors.[815,816] **Anandamide** is a lipid derived by hydrolysis of the unusual phospholipid *N*-arachidonoylphosphatidylethanolamide. This is one of a recently discovered series of amides, esters, and ethers derived from arachidonic acid (Fig. 30-29).[816a,b] They have been identified as endogenous ligands of the abundant **cannabinoid receptors**.[816] The latter were identified as binding sites of Δ^9-tetrahydrocannabinol and cannabadiol (Fig. 30-29), both of which are constituents of **marijuana**. Anandamide was the first of the endogenous cannabinoids to be isolated.[817] However, the monoglyceride **2-arachidonoylglycerol** (Fig. 30-29) is much more abundant in brain and also activates cannabinoid receptors.[818] It arises by hydrolysis of a diglyceride.[819,820] Recently 2-arachidonoyl glyceryl ether (**noladin ether**; Fig. 30-29) has been identified as another endogenous agonist of the CB_1 cannabinoid receptors.[821] A possible alternative pathway for anandamide synthesis is via an energy-dependent coupling of arachidonic acid with ethanolamine.[822,823] The two known types of cannabinoid receptors are both 7-helix proteins coupled by G_i or G_o proteins to adenylate cyclase and to Ca^{2+} and K^+ channels.[824,825] The CB_1 receptors are found largely in the brain and are responsible for the psychoactive effects of cannabis, while the CB_2 receptors are more widely distributed. They seem to have a special role in cells of the immune system, e.g., in macrophages and B cells.[818,820,825–828] Palmitoylethanolanide has been proposed as an additional endogenous ligand for CB_2 receptors.[820,829] Cannabinoid receptors of invertebrate immune system cells and of human monocytes have been found coupled to NO release.[830]

Cannabinoid receptors are present at extremely high levels in the basal ganglia of the brain,[831,832] but they do not appear to be essential. Knockout mice lacking the CB_1 receptors appear normal in most respects. However, they do not respond to cannabinoid drugs and, curiously, do not become addicted to morphine as normal mice and have less severe withdrawal symptoms than normal after morphine addiction.[826] The CB_1 receptors in the basal ganglia modulate GABA neurons that have outputs to the substantia nigra and the globus pallidus (Fig. 30-30B). The nigrostriatal neurons also secrete substance P and dynorphin, while those extending to the globus pallidus generally contain enkephalin as a cotransmitter.[832] These interconnections affect the dopaminergic neurons. Cannabinoids also have pain supressing and neuro-protective effects. They may have many possible medicinal uses, which are being explored.[833–837]

The endogenous cannabinoid compounds are lipids and are not stored in synaptic vesicles but are presumably released by enzymatic action following passage of a nerve impulse. Recent evidence suggests that the endocannabinoids are released at a postsynaptic membrane and then diffuse back to a presynaptic surface and outward to other cell surfaces where they affect signaling.[838–840] This **retrograde signaling** in synapses of the hippocampus is thought to be involved in **long-term potentiation** (**LTP**), the changes in synaptic properties that occur during learning and in the formation of memories (Section 12). A monoglyceride lipase participates in inactivation of endocannabinoids.[840a] Anandamide is also a substrate for cyclooxygenase-2 (Eq. 21-16), whose action may lead to formation of additional immunomodulatory compounds.[841,842] Long-chain relatives of arachidonic acid such as docosohexaenoic acid (DHA; Box 21-B) are especially high in brain lipids.[843,843a]

Nitric oxide and carbon monoxide. The gaseous molecules NO and CO have both been found in the brain, and neuronal NO synthase (nNOS or NOS I) has been studied intensively.[844–847] NO synthases and the functions of NO and CO are discussed in Section A7 and in Chapter 18. Complexity in understanding the role of NO in the brain arises from the fact that different isoenzyme forms of NO synthase occur in three different types of cell: nNOS in neurons, iNOS from microglial immune system cells, and eNOS from endothelial cells of capillary blood vessels.[846] All three types of cells are so tightly intermingled in the brain that it is hard to interpret observed experimental effects. Elevated Ca^{2+} concentrations that can arise from stimulation of NMDA receptors in the hippocampus seem particularly effective in activating the calmodulin-dependent nNOS. This suggests that, like the endogenous cannabinoids, NO may be a retrograde messenger in LTP.[149] The possibility that CO may function in a similar way also remains uncertain, as does any pathway for metabolism of CO. Certainly NO and CO generated in the brain will have some effects that arise from their very tight binding to heme groups. An example is the observed inhibition of dopamine β-hydroxylase by N_2O_3 with a resulting decrease in noradrenaline synthesis.[848]

10. Some Addictive, Psychotropic, and Toxic Drugs

Humans have a long history of use of stimulant and mind-altering substances. Tea, coffee, alcohol, tobacco, opium, cocaine, marijuana, and a host of modern synthetic compounds have been used as stimulants, as medications, and for pleasurable experiences.[849–851] Many are also addictive and sometimes lethal. Stimulant drugs such as nicotine, cocaine, methamphetamine (METH), and other amphetamines (Fig. 30-27)[849,852–854b] can give users feelings of increased energy, well-being, and self-confidence. Nicotine enhances fast excitatory transmission[855] and may sharpen memory.[856] However, all are acutely toxic and are highly addictive. Amphetamines and cocaine act directly to increase the brain dopamine level causing euphoria. However, in response the dopamine receptors rapidly decrease their sensitivity. This leads to mental depression and the desire for more drug. Nicotine appears to indirectly affect the same dopamine neurons.[857] The wisdom and ethics of giving hypoactive children the addictive stimulant **methylphenidate** (Ritalin; see Fig. 30-33) have been questioned.[858,859] The depressive drugs, including **morphine** and other narcotics (Fig. 30-30), barbiturates (Fig. 30-22), and ethanol, are all strongly addic-

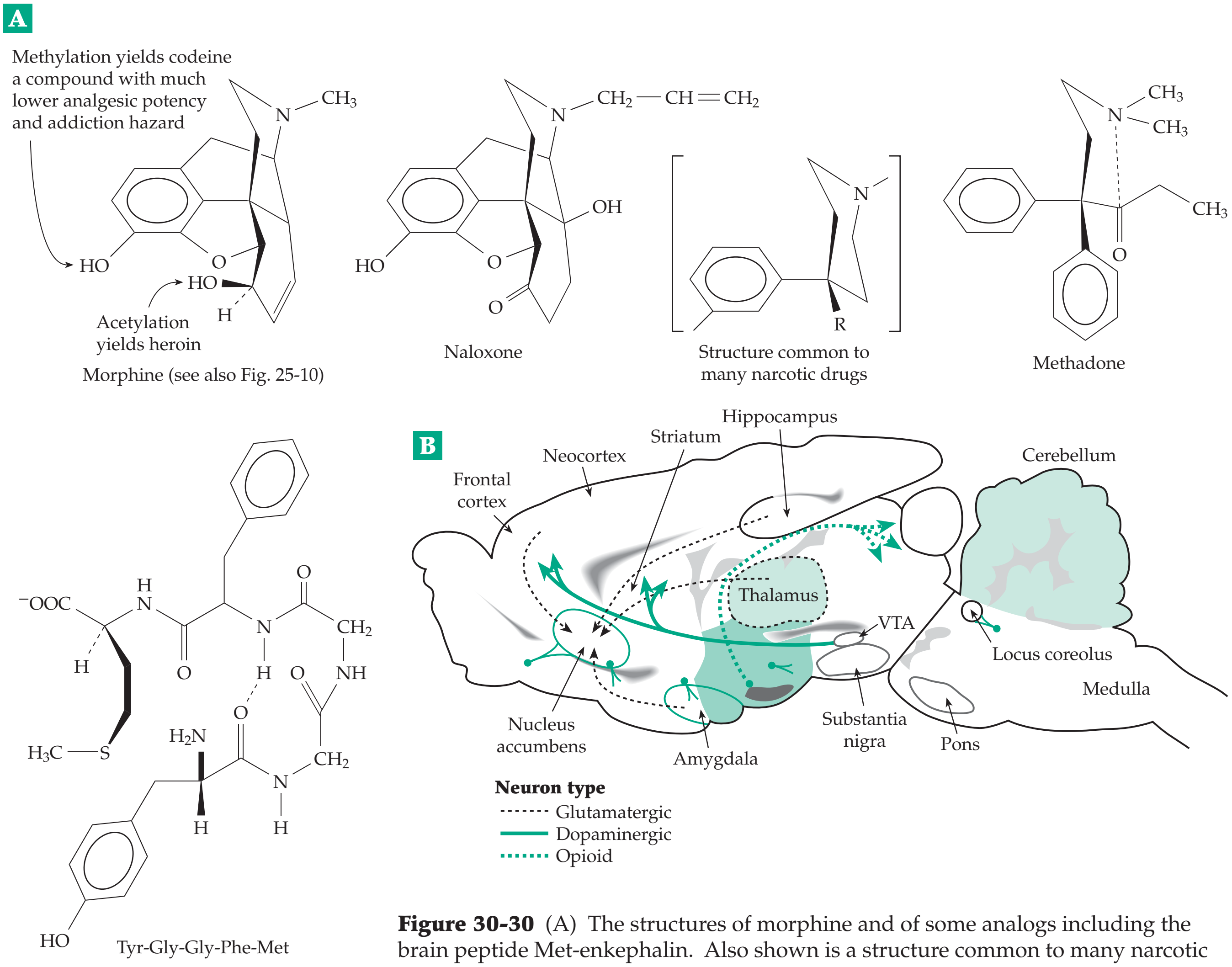

A
Methylation yields codeine a compound with much lower analgesic potency and addiction hazard
N
CH3
HO
O
HO
H
Acetylation yields heroin
Morphine (see also Fig. 25-10)
CH2—CH=CH2
OH
HO
O
O
Naloxone
N
R
Structure common to many narcotic drugs
CH3
CH3
O
CH3
Methadone
−OOC
H
N
H
O
N
H
O
CH2
NH
H3C—S
H2N
O
CH2
N
H
H
HO
Tyr-Gly-Gly-Phe-Met
B
Frontal cortex
Neocortex
Striatum
Hippocampus
Cerebellum
Thalamus
VTA
Locus coreolus
Medulla
Nucleus accumbens
Amygdala
Substantia nigra
Pons
Neuron type
Glutamatergic
Dopaminergic
Opioid

Figure 30-30 (A) The structures of morphine and of some analogs including the brain peptide Met-enkephalin. Also shown is a structure common to many narcotic drugs. (B) Diagram of a rat brain as shown in Fig. 30-13 with some aspects of the mesolimbic dopamine system emphasized. See Shulteis and Koob.[869]

tive for susceptible individuals. The phenomenon is most striking in the case of the opiates. Addiction leads to physical dependence, a situation in which painful withdrawal symptoms occur in the absence of the drug. At the same time a striking tolerance to the drug is developed. The addicted individual can survive what would otherwise be a fatal dose without ill effect. Aside from the pathological hunger for the drug, an addict can function normally in almost every respect.[860] Dependence develops only from frequent doses of drug over a long period of time and is not observed with cocaine or amphetamines.[861] Marijuana is only mildly addictive, according to some data about the same as caffeine.[862] However, this conclusion is controversial.

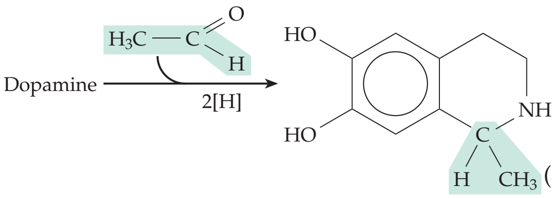
Opioid receptors. Direct binding of highly radioactive opiates has permitted localization of specific opiate receptors of several types.[863–866] The three major types (μ, δ, κ) are all 7-helix receptors coupled to adenylate cyclase, K^+ and Ca^{2+} channels, and the MAP kinase cascade.[866] The μ receptors bind morphine most tightly.[867,867a] These receptors are found in various cortical and subcortical regions of the brain. Most narcotics are polycyclic in nature and share the grouping indicated in Fig. 30-30. However, the flexible molecule **methadone** binds to the same receptors.[868] Among antagonists that block the euphoric effects of opiates the most effective is **naloxone** (Fig. 30-30).

What is the natural function of opiate receptors? Opiates are the most powerful **analgesic agents** known. The existence of the **enkephalins**, **endorphins**, and **endomorphins** (Section A,5; Table 30-4) in the brain suggests that opiate drugs mimic the normal action of these peptides, which may function in controlling pain. Although opiates are powerful drugs, their efficiency in diminishing pain is directly related to their addiction potential. To date, it has not been possible to design a nonaddictive analgesic drug of the potency of morphine.

Addiction seems to follow compensatory changes in the receptor–agonist system that result from the occupation of the receptor sites by the drug. For example, studies of opiate receptors indicate that morphine acts in an inhibitory fashion, lowering the internal level of cAMP.[861,870] The neuron then compensates by increasing the number or activity of adenylate cyclase molecules restoring the internal cAMP level. This leads to dependence upon morphine because in its absence the cAMP level rises too high. The increased number of adenylate cyclase molecules and associated receptors also accounts for the observed tolerance. It is now clear that this adaptation is complex. The properties of many synapses in various parts of the brain are altered by phosphorylation or dephosphorylation or other reactions of receptors and other synaptic proteins. Some changes are rapid, but others are slower and involve alterations in transcriptional patterns within neurons. These changes occur in three different neuronal systems: (1) physical control systems, in which changes lead to physical dependence; (2) motivational control systems; and (3) associative memory systems.[861]

The **mesolimbic dopamine system** is thought to be involved either directly or indirectly in addiction to many drugs. The dopaminergic neurons of this system have cell bodies in the **ventral tegmental area** (VTA) of the brain (Fig. 30-30B) and extend into the **nucleus accumbens**, a region at the base of the striatum that is thought to provide the "rewarding effects," i.e., pleasure from drugs such as cocaine or amphetamines. There is direct experimental support for this conclusion.[871] Less certain is the proposal that opiates and other depressive drugs indirectly cause a similar effect in the nucleus accumbens.[869,870,872] A more recent view is to regard addiction as an aberrant form of learning.[861,871,873] This concept is applicable also to "behavioral addictions."[873a]

Ethanol. As with morphine addiction, tolerance to alcohol is developed, and a lack of ethanol produces withdrawal symptoms. The principal route of metabolism of ethanol (both ingested and the small amount of endogenous alcohol) is believed to be oxidation in the liver to the chemically reactive acetaldehyde (p. 774),[874,875] which is further oxidized to acetate. Some theories of alcoholism assume that addiction, and possibly also the euphoric feeling experienced by some drinkers, results from a metabolite of ethanol in the brain. For example, acetaldehyde could form alkaloids (Eq. 30-5).[876]

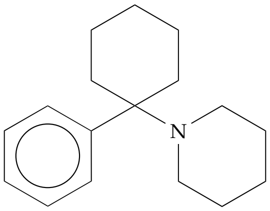


(30-5)

In fact, small amounts of morphine, 6-acetylmorphine, codeine, and thebaine, all opiate compounds, have been found in mammalian brain[877,878] and have presumably arisen by the same pathway observed in plants (Fig. 25-10). However, there is no cross reactivity between morphine and alcohol in addicted mice,[879] and acetaldehyde is probably not the addictive agent. Acetaldehyde is very reactive and may be responsible for much of the damage caused by ethanol.[880] At a blood ethanol concentration of 20 mm a person is legally intoxicated, and large amounts of acetaldehyde may be formed and react with many amines, nucleotides, proteins, etc. Ethanol blocks glutamatergic NMDA receptors and activates GABA receptors.[740a] These effects may be involved in the neurodegeneration of fetal alcohol syndrome.[881,882] Ecitotoxicity may also be a factor in alcohol damage.[692] Alcoholic liver disease may involve malnutrition as well as direct damage.[883]

Experiments with mice and rats have established a genetic propensity toward addiction to alcohol. Animals from some strains shun alcohol and become addicted only if force-fed for prolonged periods. Others, which may have low levels of neuropeptide Y in the brain,[884] accept the alcohol readily and become addicted quickly. That a similar situation holds for humans is quite possible. However, a specific “alcoholism gene” has not been found.

Psychotropic or mind-changing drugs. Hallucinogenic compounds have long been a source of special fascination to many people. The presence of the indole ring in the powerful hallucinogen **lysergic acid diethylamide** (LSD; Fig. 25-12) suggests that this compound may mimic the action of serotonin. However, other experiments suggest antagonism of dopamine receptors in the striatum.[885] Other hallucinogens include 3,4-methylenedioxyamphetamine (MDA; Fig. 30-27),[886] a compound that damages serotonergic neurons, its derivative “ecstasy” (Fig. 30-27),[886a] mescaline (Fig. 30-27), and phencyclidine (angel dust), a compound introduced in the late 1950s as a general anesthetic. Unfortunately, it produces a long-lasting condition resembling schizophrenia.[887] A common site of action for a large variety of hallucinogens has been suggested.[888]



Phencyclidine (Angel dust)
a discredited anesthetic

Neurotoxins produced by the body. Some normal body constituents are neurotoxic in excess. These incluse **quinolinic acid** (Fig. 25-11),[889] **3-hydroxykynurenine** (Fig. 25-11; p. 1444),[890] and homocysteine.[891] Elevated levels of homocysteine are also associated with vascular disease and stroke (Chapter 24). 3-Hydroxykynurenine is a precursor to ommochrome pigments of insects and an intermediate in conversion of tryptophan into the nicotinamide ring of NAD in humans (Fig. 25-11). 6-Hydroxydopamine (Fig. 30-26), which may be formed in the body, is severely toxic to catecholaminergic neurons.[892]

Other neurotoxins can be formed from environmental pollutants. The solvent 1,4-butanediol is converted to **γ-hydroxybutyrate**, which is also a drug of abuse.[893] Many compounds in commercial use have not been adequately tested as neurotoxins.[894]

11. The Senses: Sight, Smell, Taste, Hearing, Touch, and Others

Our brains receive a continuous stream of impulses from receptors that sense light, taste and odor molecules, sound waves, touch, pain, gravitational pull, etc. Of these receptors those of vision, which are discussed in Chapter 23, may be the best known. The photoreceptors consist of rhodopsin and related 7-helix proteins embedded in membranes of the rod and cone cells (Fig. 23-40). A complex series of control mechanisms, some of which are outlined in Fig. 23-43, permit enormous amplification of the initial signal generated by a G protein and a cGMP-gated ion channel. The array of rods and cones in the retina send messages via the optic nerve to the **visual cortex**, an area of ~15 cm^2 on the cerebral cortex surface at the back of the brain.[149,895] The visual cortex is divided into two halves, but curiously, the right eye sends its signals to the left brain and vice versa. The image viewed by the retina can be mapped to the visual cortex. There it may reside in the form of chemical alterations in the ~40,000 neurons thought to be present in the visual cortex[409,896–898] for a short time until it is stored in short-term working memory locations.

Receptors for the other senses, like those for sight, also consist of clusters, often in regular arrays, of 7-helix receptors. Most of these are also G protein-coupled ion channels that are controlled by cAMP or cGMP.[899]

Odor. Even bacteria possess something akin to our ability to taste and smell. As is discussed in Chapter 19, Section A, many bacteria are attracted to L-serine or D-ribose and are repelled by phenol. Receptor proteins in the plasma membrane are involved in sensing these compounds and in allowing bacteria to move toward food and away from danger. Many other examples of chemotaxis are known among the lower invertebrates such as *Euglena*. Chemoreceptors in *Hydra* sense glutathione that flows from the broken tissue of their prey and control the animal’s feeding behavior. Related organisms respond to proline. Asparagine induces the bending of the tentacles of the sea anemone *Anthopleura*, while glutathione induces swallowing.[900] Salmon return to their home streams using a memory of specific odors.[901]

Thoughout the animal kingdom the sense of smell is essential for survival. Perhaps it is not surprising that from the nematode *C. elegans* to human beings there is a largely conserved mechanism for sensing odors.[902] A large array of 7-helix G protein-coupled olfactory receptors embedded in an epithelial membrane carry signals directly into the nervous system.

In *C. elegans*, which has only 302 neurons, there are 32 chemosensory neurons and more than 100 genes for 7-helix receptors that are expressed in these neurons.[903–905] The fruit fly *Drosophila melanogaster* has at least 59 genes for olfactory receptors.[906,906a] Zebrafish and catfish have ~100.[905,907] Mice and rats have ~1000 olfactory receptor genes and human beings at least 500, which account for about 1–4 % of the genome.[905,908,909] In higher animals most receptors are coupled via G proteins, adenylate cyclase, and cAMP to ion channels in the membrane.[905,908] Insects utilize both cAMP and Ins3-*P* in their chemosensory receptors.[906] The signaling pathways parallel those of the visual receptors (Figs. 23-40, 23-43), which, however, utilize cGMP. Each gene is thought to give rise to a receptor of a specific **type** able to respond to specific structural features in an odor molecule.

Human olfactory cells are located in the **olfactory epithelium** on the upper surface of the back portion of the nasal cavity. They are neurons with chemosensory cilia similar to the rods and cones of the retina (Fig. 23-40). The cilia, which can be detached and isolated from the olfactory epithelium, contain the odorant-stimulated G-protein-dependent adenylate cyclase.[910,911] There are ~10 million receptor cells of at least 500–1000 different types. The 10 million axons form bundles of ~5000 axons each and pass through small perforations in the skull directly into the **olfactory bulb** (at the front of the brain before the pituitary, Fig. 30-13), a distance of 3–4 cm. The cortex of the olfactory bulb is lined with ~1800 **glomeruli**. Each glomerulus is a bundle, ~0.1–0.2 mm in diameter, of synaptic endings of the neurosensory nerves coming from the olfactory epithelium with dendrites of neurons that run to the **olfactory cortex** and other regions of the brain.[909] Each sensory receptor sends signals to a single glomerulus, but the glomerulus receives signals from 500 or more sensory neurons, which are not all of the same types. The glomerular cortex of the mouse is divided into four zones, each of which contains only some of the types of receptor. It seems that the cortex contains a crude “map” that relates position to the type of smell.[912] The neural processing involved in the discrimination of odors is not yet clear.[912,913] Interneurons of the olfactory bulb are unusual, being continuously discarded and replaced by new neurons that arise from neural stem cells.[908,914] This process seems to be essential for odor discrimination but not for the sensitivity of odor detection.

Most mammals have a second olfactory apparatus, the **vomeronasal organ** (VNO) or “sexual nose,” which is located on the lower surface of the nasal cavity. It is a fluid-filled cavity containing chemosensory receptors through which nasal fluid is literally pumped, when the animal seeks to maximize the sensitivity of detection.[908,915] The VNO is especially important to reproduction, defense, and food-seeking. A specialized set of olfactory sensory neurons that project to atypical glomeruli in the olfactory bulb utilize cGMP signaling and may also function in reproductive behavior.[916]

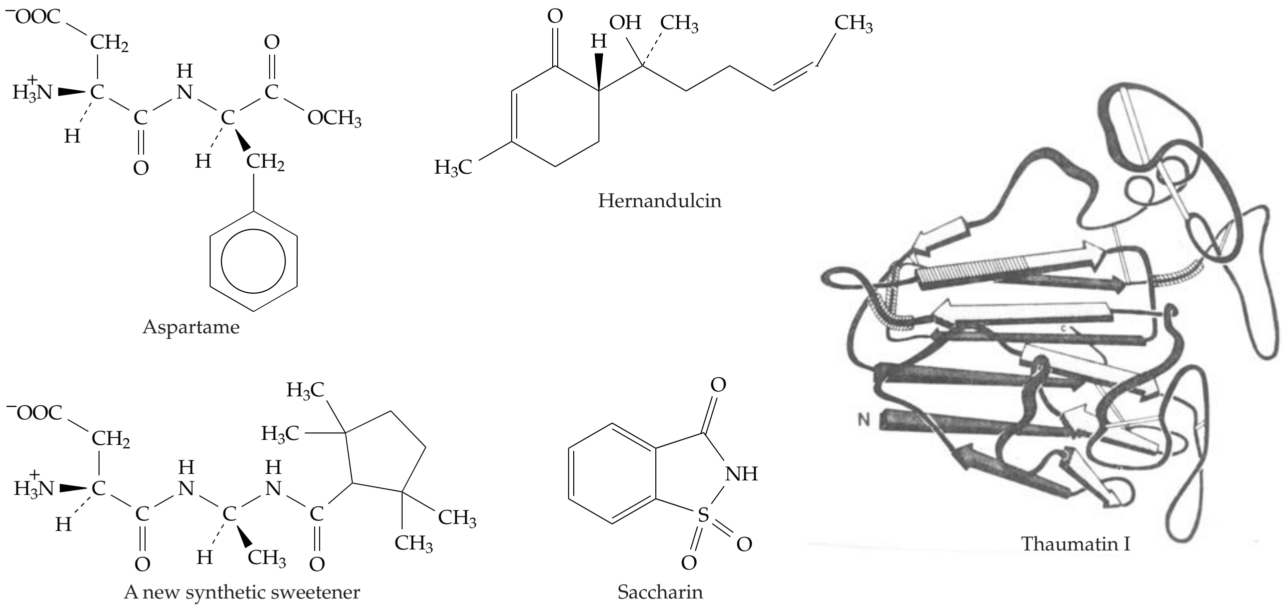
The olfactory epithelia are bathed in an aqueous mucus through which odorant molecules must pass. A number of specialized proteins, including **odorant-binding proteins,** are secreted in this fluid.[917–919] Many odorant-binding proteins are **lipocalins** (Box 21-A) and presumably assist in transporting lipophilic odorant molecules to the olfactory receptors. They tend to have a low specificity for the odorant and a weak binding affinity, properties that are consistent with this function. Pheromone-binding lipocalins encoded by ~30 genes are also found in rodent urine,[920] where they play a similar role. In contrast, the pheromone-binding proteins of some male moths are largely α helical.[920a] Although pheromones are not as important to human physiology, axillary odors from both males and females do apparently carry chemical signals. One well-established effect is the synchronization of menstrual cycles of women living in the same house or dormitory. Alipoprotein D apparently serves as a binding protein that carries odorant precursors that are acted on by bacteria to produce the pheromones.[921]

Virtually all people lack the ability to detect some specific odors. A striking example of such an **anosmia** is the inability to smell the volatile steroid **androstenone** (5α-androst-16-en-3-one), a constituent of perspiration, of some pork products, truffles, and celery.[922]

Taste. Less is known about the biochemistry of taste. The taste that we perceive is affected by odor, temperature, and physical contact. However, five primary tastes are recognized.[923,923a]

Salty: apparently perceived by an ion-channel-linked receptor
Sour: apparently linked to an H^+ channel
Bitter: perceived by bitter-sweet G protein-coupled receptors
Sweet: also perceived by bitter-sweet receptors
Umami: a recently recognized taste, that of glutamate

An experimental difficulty lies in the fact that there are only a few thousand taste buds in the tongue, with only 50–100 cells in a bud. They age rapidly, having a lifespan of only about ten days.[924] There may be only 30,000–50,000 hard-to-isolate taste receptor cells on the tongue’s surface.[923] However, very recently published reports describe a large family of bitter and sweet receptors in mice and humans[924–928] and in *Drosophila*.[929,930] The sweet-sour receptors are thought to activate a G protein called **gustducin**,[931,932] which plays a role similar to that of transducin in vision and



Aspartame

Hernandulcin

A new synthetic sweetener

Saccharin

Thaumatin I

Figure 30-31 Structures of some very sweet compounds. The backbone structure of the protein thaumatin I is included. The main body of this structure consists of two β sheets forming a flattened β barrel. β Strands in the top sheet are shaded light, and those in the bottom sheet are darker. Open bars represent disulfide bonds, and the regions with sequences homologous to monellin are indicated by the hatched marks. From de Vos *et al.*[940]

---

also activates ion channels.[933] Like the odor receptors, taste buds are also bathed in a special fluid. The **von Ebner's glands** in the tongue contain binding proteins,[934,935] at least some of which are lipocalins.

The relationship between structures and sweet taste in humans has been investigated intensively, but no simple rules have been discovered (see Robyt[936] for a discussion). Sucrose is usually perceived as very sweet. D-Fructose and D-xylose are nearly twice as sweet, but D-glucose is less sweet than sucrose. D-Galactose is usually perceived as not sweet and D-mannose as bitter. Many sucrose derivatives, in which hydroxyl groups have been replaced with Cl or other halogen, are very sweet. One tetrachloro derivative of this type is 7000 times sweeter than sucrose.[936] Some especially sweet materials are depicted in Fig. 30-31. These include peptide derivatives[937,938] such as Asp-Phe-OCH_3 (**aspartame**), the sesquiterpene **hernandulcin**,[939] and **chemostimulatory proteins**. Among these are some of the sweetest substances known, the 207-residue **thaumatins**[940,941] and **monellin**[942,943] present in certain tropical berries. Thaumatins are ~3000 times sweeter than sucrose on a weight basis and 10^5 times sweeter on a molar basis. Thus, sucrose tastes sweet at a concentration of 10^{-3} M or higher but thaumatin[944] at 3×10^{-8} M.

The proteins **miraculin** and **circulin** from tropical fruits modify taste. Acids taste sweet rather than sour after the tongue has been treated with either protein.[945,946] Exposure of the tongue to artichokes often makes water taste sweet.[947] Thus, the response of taste receptors can be temporarily altered by binding of other substances, perhaps at adjacent sites on a receptor.

Pain. Receptors for pain (**nociceptors**) are spread over the body in nerve endings found in the skin, muscle, joints, and internal organs. There are several types of receptors, most of which are present in excitatory glutamatergic neurons.[948] Some release substance P. Some activate tyrosine kinases and others ATP-gated ion channels. Some pain receptors are also activated by intense heat or pressure or by irritant compounds. Among the latter is capsaicin (Fig. 30-29), the active ingredient in chili peppers, and an ultrapotent compound, **resiniferatoxin**. Both capsaicin, which is 10,000 times more potent than jalapeño peppers,[949] and resiniferatoxin, which is 20-fold more potent than capsaicin, bind to **vanilloid receptors**. These are ligand-gated ion channels related to the Shaker K^+ channel (Fig. 30-18). They are nonselective but with a high permeability to Ca^{2+} and are members of the **transient receptor potential** (TRP) family.[948,950–954] Pain seems to stimulate an increase in anandamide (Fig. 30-29), which has an analgesic effect. Nevertheless, anandamide and *N*-vanillyloleamide activate capsaicin receptors.[955] Because the activated receptors become desensitized rapidly, capsaicin has been used in a paradoxical manner as an analgesic agent.[951] Sensing of temperature changes also depends upon TRP channels.[955a,b]

Mechanoreceptors. The transduction of mechanical force into a chemical signal provides the basis for the senses of touch and hearing. Plants detect wind and gravitational force,[956] and many organisms, even bacteria, respond to changes in osmotic pressure using mechanoreceptors.[957] One of the best known mechanoreceptors is from *Mycobacterium tuberculosis*. It is a homopentamer whose three-dimensional structure[450–452,958] resembles that of the nicotinic acetylcholine receptor (Fig. 30-23). A second type of mechanoreceptor is found in the inner membranes of *E. coli* and in plasma membranes of many other bacteria, archaea, and some eukaryotes.[957a] These receptors, which are also sensitive to voltage changes, are heptamers of a 282-residue protein that forms a symmetric ion channel in the center. There is also a large cytoplasmic domain consisting largely of β structure.[957b] How do such receptors sense mechanical stress? In bacteria they respond to stretch in the membrane induced by an increase in osmotic pressure. One suggestion is that the membrane expansion pulls apart the radially symmetric ion channel in the receptor.[958,958a] In higher organisms transmembrane adhesion receptors and their linkage to the internal cytoskeleton provide a framework for detection of mechanical forces and linkage to mechanoreceptors.[956,957,959]

Hearing. Movement of the **stereocilia** of the hair cells of the inner ear activates mechanoreceptors. Each stereocilium contains a core of crosslinked actin filaments, and tens to hundreds of these cilia are connected in hair bundles, which move in response to arrival of sound waves of appropriate frequencies. The movement of the stereocilia induces the opening of receptor ion channels in the hair cell membrane allowing K^+ and other ions to flow inward.[960] The matter is much more complex than this because of the tuning and amplification mechanisms in the cochlea of the inner ear.[960–962] These mechanisms allow receptors in hair cells to respond to very weak vibrations of specific frequencies. Both mechanical and biochemical mechanisms are involved. A number of specific proteins participate. Among these is a motor protein called **prestin**, which seems to be involved in the rapid changes in length and stiffness of some hair cells in the cochlea.[962,963]

Other sets of hair cells are formed in specialized parts of the inner ear.[964] The three semicircular canals detect angular acceleration in three directions, while the sac-like utricle and saccule detect linear acceleration including gravitational attraction. These two organs each contain a patch of hair cells whose tips project into a gelatinous layer, which is overlain by a field of small crystals of calcium carbonate. These little stones (**otoliths**) provide an inertial mass, which resists movement causing the hair cell tips to bend and activate mechanoreceptors to send information about balance and orientation to the brain.

While discussing vibrations we may ask whether 60 cycle electromagnetic field fluctuations caused by electrical power transmission can affect the human body? Considerable effort has been expended in addressing this question. The tentative conclusion is that such effects, if they exist, are extremely difficult to detect. However, the possibility has not been disproven.[965]

12. The Chemistry of Learning, Memory, and Thinking

What is known about the chemistry underlying memory, thinking, and the generation of the stream of consciousness within the brain? Nerve impulses originating in sensory receptors are sent to several regions of the brain, among which are the sensory regions of the cerebral cortex (Fig. 30-14). Memory also depends upon other regions of the brain including the hippocampus, amygdala, and cerebellum (Fig. 30-1). Learning, remembering, and thinking all require transfers of information between various neurons and between different parts of the brain. These tranfers may perhaps be coordinated via endogenous electrical rhythms (brain waves).

Memory stystems. Memories exist in several forms and are found in various regions that are reached by several pathways.[966–967a] Two major forms of memory are:

Explicit (declarative, episodic):
Conscious recall of facts and events involving people, places, and things
Implicit (associative):
Nonconscious recall of motor skills, conditioned responses, etc.

Explicit memory depends upon the **temporal lobe** of the midbrain, an area that includes the hippocampus and the nearby subiculum and entorhinal cortex.[966,968–971] Implicit associative learning and memory involve the cerebellum, amygdala, and other regions.[972,972a]

Both types of memory possess both **short-term** and **long-term** components. Short-term memory lasts only minutes to hours, but long-term memory lasts days, weeks, and sometimes a lifetime. The difference between the two is clearly seen in individuals who have damage to the hippocampus and impairment of short-term memory. A blow to the head may cause total loss of short-term memory of associations (**amnesia**).[969,973,974] Some persons with damage to the hippocampus may never regain their temporary memory, but long-term memories are intact, and new long-term memories may still be formed. An increasingly important tool for study of memory is brain imaging using

PET or **fMRI** (Box 30-A). These tools have become rapid and sensitive with the ability to observe regions of the brain that become activated by visual, auditory, or other stimuli.[897,898]

The brain often needs to store information for a short period of time. For example, one can recall many details of a visual image after closing one's eyes or shifting one's gaze. The sensory images may be stored in **working memory**.[896,975] Similarly, if one mentally multiplies two 2-digit numbers the partial product obtained by multiplying the two right most digits is temporarily stored in working memory until the next arithmetic operation is completed, etc. PET and fMRI tomography indicates that regions in the prefrontal cortex may be involved.[976,977]

Some short-term memory appears to be stored by neurons that continue to fire after a stimulus has stopped. It has been proposed that such memory consists of reverberations of electrical activity in loops of coupled axons.[978,979]

Implicit memory can be studied in animals. Much has been learned from the large marine snails *Aplysia* and *Hermissendra* whose simple nervous systems and large neurons have been investigated for over 40 years.[967a,980–984] The basic chemical mechanism associated with learning in these creatures seems to be similar to those in our own brain. Olfactory memory can be studied in *Drosophila*, even though the organization of the fly's brain differs from ours.[985–987b]

To be useful for more than a few minutes stored information must be transferred from the temporary to more permanent forms. We know that even temporary memory depends upon chemical changes in synapses. Long-term memory involves both stable chemical changes and also changes in the physical connections between neurons. Before discussing these changes let us consider briefly the waves of nerve impulses that drive the necessary alterations.

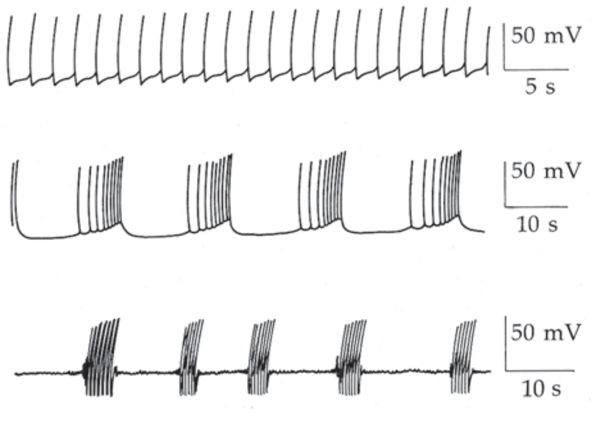


Figure 30-32 Intracellular recordings from isolated neurons of the mollusc *Aplysia*: (A) beating pacemaker, (B) bursting pacemaker, and (C) oscillating pacemaker. From Chen *et al*.[1002]

Brain rhythms. A live brain displays characteristic oscillatory activity. Using electrodes placed on the scalp of an awake but relaxed individual, a rhythmic change in the recorded voltage with a frequency of ~10 Hz (**alpha waves**) can be detected. Such **electroencephalograms** (EEGs) contain other rhythms at ~5–6 (theta), ~40 (gamma), and ~200 (high frequency) Hz.[978,988,989] More recent studies employ microelectrodes placed on individual neurons. Some cells generate spikes at frequencies as high as 800 Hz. However, the significance is uncertain.[990] The 40-Hz frequency, which is prominent in the hippocampus, has aroused the most interest[991–995] because of its probable relationship to learning and memory. Psychophysical experiments have suggested that humans can store only 7 ± 2 items, such as digits in a telephone number. To remember more digits usually requires a conscious effort to place them in longer-term memory. One proposal is that the seven items in temporary memory are stored as 40-Hz oscillations and that ~7 such items can be stored within a single 5-Hz theta oscillation.[978,979] Thinking rates have also been estimated as 7±2 thoughts per second. This is also the same as the syllable rate in speech. This allows us to speak at the same rate that we think.[996] Stuttering may be a result of lack of synchronization of thinking and speaking.

Individual cells or groups of cells are able to initiate rhythms.[997] Examples are provided by the slower Ca^{2+} oscillations shown in Box 6-D,[998–1001] by the periodic release (at ~155 intervals) of cAMP by cells of *Dictyostelium* (p. 20), and by the 24-h circadian cycle observed for virtually all living cells (Section 13). In simple invertebrates the source of neural rhythms appears to reside in **pacemaker neurons** that fire spontaneously at regular intervals. Their cell membranes apparently undergo a cyclic series of changes in ionic permeabilities sufficient to initiate action potentials. Three types of pacemaker output from molluscan neurons[1002] are illustrated in Fig. 30-32. In lobsters three-neuron **pacemaker groups** provide a pyloric rhythm. In these groups the oscillation period of a pacemaker neuron is adjusted from its intrinsic value by feedback through inhibitory and electrical connections to the other cells.[1003] Electrical coupling seems to be basic to oscillatory cell networks.[1004] Individual neurons or small groups of neurons in our own bodies act as pacemakers for the heartbeat rhythm.[705] The slow 3- to 8-Hz rhythm observed in the EEG apparently originates in pacemaker bursts from the basal ganglia.[993] Rhythms from these endogenous pacemakers may combine with pulses from sensory neurons to evoke **conscious thought**. However, the basis of consciousness is still poorly understood.[1005,1006]

As can be seen from Fig. 30-32, neurons send "trains" of spikes down their axons. These form synapses with dendrites, usually on dendritic spikes, of a postsynaptic cell.[593,1007–1009] However, each such cell typically receives input from thousands of other neurons. At any moment most of these are probably "silent," but others are sending trains of impulses. Among the important questions are "How does the postsynaptic neuron know whether to fire or not?" and "What kinds of information, if any, are encoded in the trains of impulses both in the presynaptic inputs and in the output of the postsynaptic neuron?"[1010,1011] Part of the answer to the first question is probably that firing occurs if two or more input impulses arrive synchronously,[1010,1012–1014] and if there are not too many inhibitory impulses that damp the response. In the hippocampus a network of neurons electrically coupled via gap junctions may be synchronized to the theta and gamma brain rhythms by high-frequency (150–200 Hz) oscillations.[988] See also Fig. 30-15.

Chemical changes in synapses. It has long been recognized that the synapses are the probable sites of alterations that lead to memory, whether long-term or short-term. Study of individual synapses has demonstrated the phenomena of **potentiation** (facilitation) and **depression** (habituation). Potentiation refers to the fact that a second impulse will often be transmitted through a synapse more effectively than the first, while depression refers to a decreased response to repeated stimuli. Memory may consist of potentiation and depression at specific synapses. The underlying chemical changes in the synapses are referred to collectively as **synaptic plasticity**. Chemical changes associated with short-term memory are often transient. Those associated with long-term memory are described as **long-term potentiation** (**LTP**) and **long-term depression** (**LTD**).

Many experimental results have confirmed the chemical basis of memory. For example, learning is facilitated by administration to animals of small doses of strychnine.[1015] Puromycin and other inhibitors of protein synthesis disrupt the transfer of information into long-term memory. They are especially effective during the first hour after the initial learning event.[1016] Increased synthesis both of mRNA and of proteins within the cell bodies of neurons is observed.

Short-term memory is not affected by inhibitors of protein synthesis, but alteration of synaptic proteins and membranes may be induced by covalent modification of existing macromolecules.[1016] One way in which this happens has been described for *Aplysia*. As the snail learns a simple gill-withdrawal reflex, the duration of the action potentials in sensory neurons is increased, and there is a greater release of transmitters. This change comes about because stimulation of the sensory neuron causes simultaneous activation of interneurons that synapse with the sensory neurons. The interneurons release the neuromodulator serotonin, which binds to receptors in the membrane of the sensory neurons. This activates adenylate cyclase, which in turn activates a protein kinase that phosphorylates a class of open K^+ channels. Phosphorylation causes the channels to close with a consequent strengthening of the action potential. Thus serotonin brings about presynaptic facilitation.[1017] The peptide FMRF amide (Table 30-5) has the opposite effect. It causes hyperpolarization and a decrease in the duration of the action potential. It also binds to a receptor on the neuronal membrane, and presumably via a different second messenger than cAMP causes the K^+ channels to stay open a longer fraction of the time.

Evidence that LTP is essential to learning in rats was provided by the observation that the antagonist 2-aminophosphonovalerate, which blocks the NMDA class of glutamate receptors (Fig. 30-20A), impairs both LTP and learning.[1018] The potentiation is thought to result, in part, from Ca^{2+} influx through the nonselective NMDA cation channels. The increased intracellular calcium may then induce phosphorylation of various proteins with associated long-lasting changes in the postsynaptic endings.[1019] A large amount of evidence favors this interpretation of LTP.[1020] However, it is a great oversimplification. Most studies of LTP in mammals (usually rodents) have focused on the CA1 region of the hippocampus and nearby brain regions. The excitatory axons in this organ are largely glutamatergic, and, as shown in Fig. 30-20A, the postsynaptic (dendritic) membranes contain both fast AMPA receptors and the slower NMDA receptors. Both are ionotropic. The AMPA receptor channels allow mainly K^+ and Na^+ to pass and are responsible for most of the nerve transmission. However, the NMDA receptors have an important controlling influence.

A generally accepted theory is that no LTP arises unless both the presynaptic and postsynaptic neurons are activated. This can happen if a presynaptic action potential activates many AMPA receptors in a synapse allowing enough flow of $Na^+ + K^+$ to depolarize the postsynaptic membrane and possibly to initiate an action potential in the postsynaptic neuron. (However, many factors, probably including influences from neighboring neurons,[1021] will affect this outcome.) The NMDA receptors are usually blocked by extracellular Mg^{2+} ions, and their ion channels remain closed. However, when the postsynaptic membrane becomes depolarized, the Mg^{2+} dissociates, and if the NMDA receptors are also occupied by glutamate, their channels will open, permitting Ca^{2+} to enter the neuron (Fig. 30-20A). This not only enhances the probability of developing a postsynaptic action potential but is also the trigger for LTP.[1022] Ca^{2+} ions have a variety of effects, one of which is to bind to calmodulin. This activates a calcium–calmodulin-dependent protein kinase, which phosphorylates postsynaptic structural and signaling proteins to increase the synapse strength.[641,1023–1025]

Modifications in existing proteins, such as are induced by Ca^{2+} and calmodulin, can provide LTP for a few hours, but other mechanisms must provide for longer-term effects. These require transcription of genes and protein synthesis, processes that occur in the cell bodies of neurons and may depend upon axonal transport mechanisms.[1026,1027] Among the experimentally observed results of LTP are ultrastructural changes in synapses and in dendrites.[1009,1028] Long-term memory is also thought to involve changes in the neocortex. Again, NMDA receptor activation seems to be involved.[1029,1030]

LTP has been demonstrated experimentally, but does it really influence memory? Evidence that it does has been provided by clever experiments with transgenic mice. Using the Cre recombinase (Chapter 27) the NR2B subunit of the NMDA receptor was overexpressed in the hippocampus[1031] and in the forebrain of mice.[1032] This was expected to provide better synaptic strengthening than for receptors with the similar NR2A subunit. It was found experimentally that these transgenic mice were more intelligent than normal mice.

LTP is also thought to affect the presynaptic as well as the postsynaptic neuron. One way in which this may happen is for a **retrograde messenger** to pass across the synapse and induce alterations in the presynaptic cell. One proposed retrograde messenger is nitric oxide, **NO**.[1033,1034] Neuronal NO synthase (nNOS; NOS1) contains a calmodulin binding site and is activated by Ca^{2+}. However, other substances as simple as K^+ might also be the messenger.

Long-term depression (LTD) is the *loss* of synaptic strength after passage of an impulse. There is evidence that during brain activity, including that in the hippocampus, both LTP and LTD are essential.[1035,1036] LTD may depend upon cyclic ADP-ribose (p. 564).[1037] LTD, like LTP, may also spread via retrograde signaling.[1038]

Does learning affect a few specific neurons or a large number of neurons? The rate of glucose utilization in different parts of the brain can be estimated from the rate at which labeled 2-deoxyglucose is taken up (see Box 30-A). From changes in this rate (obtained using ^{14}C-labeling) in brains of split-brain cats performing visual tasks it was estimated that from 10^{10} to 10^{11} neurons are activated.[1039] This supports the idea that memory is distributed over a large area of the brain, just as information about an image is stored in all parts of a hologram.

An alternative to the idea that synaptic potentiation and depression provide the chemical basis for learning is **molecular coding**. Thus, it was reported that a 15-amino acid peptide isolated from rats trained to avoid the dark carries behavioral information. When this peptide was injected into brains of untrained rats, they also avoided the dark.[1040] This was one of several reports of transfer of learned behavior through chemical substances extracted from the brain. These ideas are hard to accept in the light of our present knowledge of the brain. However, in view of the large number of different neuropeptides known (Table 30-4) the possibility that some aspects of long-term memory may be associated with transcription of specific amino acid sequences within specific neurons should perhaps still be considered.

The complexity of the brain. A major obstacle to our understanding of the human brain is its enormous complexity.[400] This problem can be appreciated if we consider the small nematode *Caenorhabditis elegans*. All of the synaptic connections among its 302 neurons had been mapped by 1986.[400,1041,1042] There are 5000 chemical synapses and 600 electrical (gap junction) connections. There are 80 different types of K^+-selective channels, 90 types of ligand-gated receptors, and ~1000 G-protein-linked receptors. Twenty-six of the neurons are GABAergic and are involved in three distinct behavioral motions that involve muscular contractions. Despite intensive efforts the system has been hard to understand. The brain of the macaque monkey has been described in great detail.[409] Over half of its cerebral cortex is devoted to vision, and this can be subdivided into 20 functional areas. However, the human brain, with its extremely large cerebral cortex, cannot be compared accurately with the monkey brain. Whereas anatomical studies are done on postmortem human tissues, *in vivo* studies rely largely on fMRI and PET imaging (Box 30-A). The resolution of these images is now less than 1 mm, but 1 mm^3 of human visual cortex contains more than 40,000 neurons![409] The microcircuits in the neocortex are still largely unknown and the tissue is of “apparently impenetrable complexity.” There may be several hundred different classes of neocortical neurons.[1043] The tissue is rich in GABAergic interneurons.[1044,1045] Some fast-spiking GABAergic neurons are also connected by electrical synapses and may be involved in detecting and promoting synchronous activity.[1046]

Intelligence. We must all agree that there is such a thing as intelligence, but can it be measured? In 1904, Spearman proposed the existence of a general intelligence factor g that could be measured as the IQ (intelligence quotient). Since then various tests have been devised that attempt to measure IQ.[1047,1048] Most recently use of PET scan data has indicated that various types of analytical analysis lead to brain activity in the lateral frontal cortex in one or both cerebral hemispheres[1049] suggesting that this is a region important to

IQ. A question that has been raised is whether analytical intelligence, creative intelligence, and practical intelligence are correlated?[1048]

Is intelligence hereditary? Both logic and observation say that heredity must be a major factor. However, it is hard to know how to measure the hereditary component.[1050–1052] Also hard to understand is why IQ scores have been increasing about one standard deviation unit per generation.[1047] Is this really true?[843] Does environment also influence IQ? The fact that new hippocampal nerve cells are formed continuously provides one mechanism by which learning, nutrition, and other influences may alter intelligence.

A difficult-to-explain aspect of the brain is the existence of rare **savants**, persons with amazing mental abilites in music, art, or computation but who are unable to communicate (autistic) and mentally retarded.[1053–1055] One boy at age four could play Mozart piano sonatas flawlessly after a single hearing. A three-year old girl drew horses with lifelike perspective from memory but was unable to communicate. Some mathematical savants can instantly state the day of the week for any arbitrary date such as June 12, 1929; others rapidly identify prime numbers. They evidently use the same strategies as mathematically trained persons. Do we all have these abilities but can't have access to them? How can we explain the fact that rarely a blow to the head will convert a person into a savant?

Behavior. It may seem impossible to interpret complex behavioral patterns at the molecular level. However, the genetics of behavior is a well recognized field of investigation, and some behavioral traits have been linked to single genes. If a gene can be located cloning, sequencing, and biochemical studies may follow quickly. The behavioral genetics of lower organisms, e.g., of *Drosophila*, have provided many insights.[1056] Recently, however, the mouse has become a major object of behavioral studies.[1057] Its genome is well known, and a very large number of mutations have been mapped. The ability to prepare **knockout mice** (p. 1501) and to carry out gene transfer experiments on such animals makes them very attractive for study.

Some behavioral traits are based on simple alterations, often defects, in motor skills. For example, the following traits in mutant mice have been traced to specific brain structures and often to specific biochemical alterations.

Staggerer	Purkinje cell defect
Vibrator	Phosphatidylinositol transfer protein gene
Tottering	Mutation in voltage-gated Ca^{2+} channel
Lurcher	Abnormality in cerebellum
Weaver	Gly → Ser mutation in K^+ channel

Knockout mice lacking oxytocin or vasopressin have altered social behavior toward other mice. Those lacking galanin seem less intelligent than normal mice, as if they had Alzheimer disease. Mice lacking neuronal NO synthase became aggressive.[1057] Human personality,[1058] language abilities,[1059] and sexual behavior all have a genetic component. However, claims that a "gay gene" has been found are not generally accepted.[1060]

13. Circadian Cycles and Sleep

In mammals an approximately 24-h (**circadian**) rhythm controls behavior and affects many physiological functions. As previously mentioned (p. 1800), the brain has its own rhythms, which originate with pacemaker neurons. The heart beats with another neurally established rhythm. The circadian rhythm has a much longer period and, therefore, seems more mysterious. It is observable, even with single cells and for virtually all organisms.[1061–1062a] In most instances the cycle becomes synchronized with the daily light−dark cycle with the aid of suitable light-absorbing pigments often crytochromes (see Chapter 23, Section I,1). However, the cycle can be observed in various ways under conditions of constant light intensity and temperature. For example, the unicellular marine alga *Gonyaulax* undergoes dramatic circadian changes in the intensity of its bioluminescence. Over one 10-day period the luminescence peaked every 22.99 ± 0.01 hours.[1063,1064] It is more difficult to measure the period for human beings (see Czeisler *et al.*[1065] for a discussion), but under suitable conditions during which time cues were missing a precise period of 24.18 hours was observed for the level of melatonin in the blood, the body temperature, and other quantities.[1065] From cyanobacteria,[1066,1067] fungi (*Neurospora*),[1068] insects (*Drosophila*),[1069–1071b] and frogs[1072] to mice and people,[1073,1074] the circadian cycle affects the organism's chemistry and behavior. Green plants likewise observe a circadian cycle.[1075]

The cycle is thought to originate in feedback loops that control transcription of a small set of genes. In *Drosophila* the set includes seven genes: *period* (*per*), *timeless* (*tim*), *clock* (*clk*), *cycle* (*cyc*), *double-time* (*dbt*), *vrille*, and *cryptochrome* (*cry*).[1070] Many corresponding genes have been found in mammals. For example, the mouse NPAS2 is a close relative of the *Drosophila* CLOCK protein, and the period proteins PER1 and PER2 and the cryptochromes CRY1 and CRY2 are also related to the *Drosophila* proteins.[1076,1077] In *Drosophila* the heterodimers PER•TIM and CYC•CLK are thought to serve as DNA-binding transcription factors that repress transcription of their own genes when they reach a high enough concentration in the nucleus.[1073,1076,1077] Because some time is required for transcription and

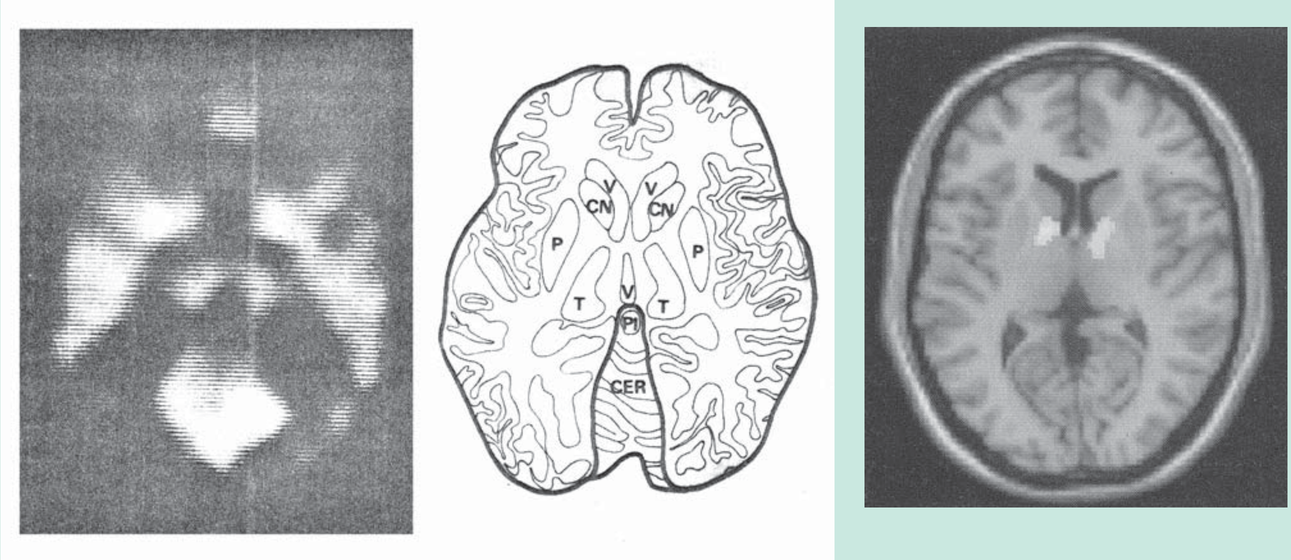
BOX 30-A POSITRON EMISSION TOMOGRAPHY (PET), FUNCTIONAL MAGNETIC RESONANCE (FMRI), AND OTHER IMAGING TECHNIQUES

In the widely used technique of transmission computerized tomography (CT) an image of a slice through the body of a patient is obtained using X-rays. An X-ray source moves in a ring around the patient while detectors measure the intensity of the transmitted radiation and send it in digital form to a computer. which generates the desired image. A chemically more sophisticated view of the body can be obtained by positron emission tomography (PET). This technique makes use of a metabolite or drug labeled with a short-lived radioisotope that decays by emission of positrons (antielectrons). Among these are ^{11}C, ^{13}N, ^{15}O, and ^{18}F with half lives of 20 min, 10 min, 2 min, and 110 min, respectively. The isotopes are produced in a cyclotron, and are rapidly introduced into suitable compounds, which can be injected into a bloodstream.[a–d] An emitted positron travels only a few millimeters before undergoing annihilation with an electron to produce two high-energy (50 keV) photons (γ-rays) that travel in opposite directions and are detected by an array of scintillation detectors.

$$\beta^+ + e^- \rightarrow 2\,h\nu$$

Present-day PET technology allows images to be formed in a few seconds, and in some cases in a fraction of a second. Among the useful compounds for PET imaging is [^{18}F]2-fluoro-2-deoxy-D-glucose (see figure). This compound, which contains the longer-lived ^{18}F, is phosphorylated by hexokinase, and the resulting phosphate ester is effectively trapped in the brain. 3-Deoxy-3-fluoro-D-glucose is another useful tracer.[e] One of the most useful PET measurements has been blood-flow monitored by ^{15}O-containing H_2O, which is administered into a vein in the arm.[b] The ^{15}O has a half-life of only two minutes and is almost completely gone in ten minutes. However, very low doses of radioactivity are used, and several images can be obtained before the radioactivity has decayed. A common practice is to subtract images obtained after the isotope has decayed from those obtained at various times while it was still present. The technique is also useful for study of the binding and transport of hormones,[f] other metabolites, drugs, and other inhibitors.[c]

The NMR technique **magnetic resonance imaging** (**MRI**),[g–i] so called to avoid the word nuclear, is rapidly displacing many applications of PET scanning. MRI uses proton NMR spectroscopy to generate very sharp images based largely on the water present in tissues. These images can be made



Left: PET image of a human brain obtained using 2-[^{18}F]fluoro-2-deoxyglucose. This tomographic brain slice at the level of the basal ganglia shows the cortical gray matter and subcortical white matter. As marked on the drawing on the right: V, ventricles; CN, caudate nucleus; P, putamen; T, thalamus; PI, pineal gland; CER, cerebellum. From Rottenberg and Cooper.[p] Right: fMRI image illustrating modulation of neural activity in the ventral striatum, an area of the brain associated with reward, when eye contact was made with an attractive face. The activation map shown is derived in a complex manner and is based on recorded brain activity of persons viewing images of a series of faces. It portrays the differences in neural activity when viewing images of attractive faces of either sex with the eye gaze directed at the subject and with the eyes averted. See the report of Kampe *et al.* for details.[q]

BOX 30-A (continued)

to depend upon variations in T_1 and T_2 (Chapter 3) as well as upon differences in the water content. The first MRI scans required 20 minutes, but the use of more powerful magnets and more sensitive instruments has reduced the acquisition times in ultrafast MRI to ~0.1 s. The decay of the NMR signal from a single RF pulse is observed at several different times.[h] The dynamics of blood flow and neural activity can be followed. Every technique has disadvantages as well as advantages. MRI does not use radioisotopes, but overheating of the brain must be carefully avoided. In addition, patients may suffer from uncomfortably loud noises generated by rapidly changing magnetic gradients.[h] As with PET scans isotopic tracers may be used. However, most MRI scanning is done with ^{1}H from the solvent water. As with PET, MRI is often used to measure blood flow but with an indirect method. The Fe of deoxyhemoglobin (Hb) is paramagnetic, but upon oxygenation to HbO_2 it becomes diamagnetic (pp. 850–851), and the ^{1}H signal of the solvent H_2O becomes sharper. In metabolically active regions of the brain the demand for oxygenated blood is greatly increased. Perhaps surprisingly, the ratio $[HbO_2]/[Hb]$ is greater in these areas than in less active areas where a greater fraction of the hemoglobin remains unoxygenated.[g] However, the exact interpretation of the ultrafast MRI images is uncertain. In **functional MRI** (**fMRI**), differences in images acquired after some physiological change are recorded. For example, after a visual or other sensory stimulus a change in the MRI image of some region of the cortex will be observed (see figure).[i–l] The technique is allowing many deductions about learning, memory, and communication pathways in the brain[j,m] and is being used to investigate many aspects of brain disease.

Another brain imaging technique is **magnetoencephalography** (**MEG**).[c,n] It has been uniquely valuable in mapping the sensory regions of the human cerebral cortex. Looking ahead, optical methods, which include use of infrared radiation, are also under development.[o] They may not be adequate for study of the human brain but can be used for smaller animals, for studies of embryonic development, etc.

---

[a] Ter-Pogossian, M. M., Raichle, M. E., and Sobel, B. E. (1980) *Sci. Am.* **243**(Oct), 171–181

[b] Raichle, M. E. (1994) *Sci. Am.* **270** (Apr), 58–64

[c] Volkow, N. D., Rosen, B., and Farde, L. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2787–2788

[d] Phelps, M. E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9226–9233

[e] Berkowitz, B. A., Moriyama, T., Fales, H. M., Byrd, R. A., and Balaban, R. S. (1990) *J. Biol. Chem.* **265**, 12417–12423

[f] Berman, K. F., Schmidt, P. J., Rubinow, D. R., Danaceau, M. A., Van Horn, J. D., Esposito, G., Ostrem, J. L., and Weinberger, D. R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8836–8841

[g] Shulman, R. G., Rothman, D. L., and Blamire, A. M. (1994) *Trends Biochem. Sci.* **19**, 522–526

[h] McKinstry, R. C., and Feinberg, D. A. (1998) *Science* **279**, 1965–1966

[i] Disbrow, E. A., Slutsky, D. A., Roberts, T. P. L., and Krubitzer, L. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9718–9723

[j] Ungerleider, L. G. (1995) *Science* **270**, 769–775

[k] McCarthy, G., Blamire, A. M., Puce, A., Nobre, A. C., Bloch, G., Hyder, F., Goldman-Rakic, P., and Shulman, R. G. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8690–8694

[l] Chen, W., Zhu, X.-H., Thulborn, K. R., and Ugurbil, K. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2430–2434

[m] Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., and Buckner, R. L. (1998) *Science* **281**, 1188–1191

[n] Yang, T. T., Gallen, C. C., Schwartz, B. J., and Bloom, F. E. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3098–3102

[o] Lok, C. (2001) *Nature (London)* **412**, 372–374

[p] Rottenberg, D. A., and Cooper, A. J. L. (1981) *Trends Biochem. Sci.* **6**, 120–122

[q] Kampe, K. K. W., Frith, C. D., Dolan, R. J., and Frith, U. (2001) *Nature (London)* **413**, 589

---

protein synthesis, this inhibitory feedback can lead to oscillations in the concentrations of the circadian clock proteins. Proteosomal degradation of the TIM protein may also be a factor.[1071] The need for proteins encoded by other genes indicates that the matter is more complex. Individual cells or individual tissues, e.g., mammalian retinas,[1074,1078] may independently set up circadian cycles. However, these normally become **entrained** by the daylight cycle and are reset daily as discussed in Chapter 23, Section I,1. Other factors such as temperature, activity, and food may also affect the resetting. One factor, which may be influenced by food, is the NAD^+/NADH and $NADP^+$/NADPH ratios within cells.[1077] The circadian cycles for mammalian tissues are synchronized by a **master clock** that originates in neural tissues[1062a] and specifically in a region of the hypothalamus containing the **suprachiasmatic nuclei**.[1079–1080b]

The pineal gland appears also to play a role in maintaining the mammalian circadian cycle.[1081–1083] The concentration of the pineal hormone melatonin (Fig. 27-11) as well as its precursor *N*-acetylserotonin and the enzyme serotonin *N*-acetyltransferase (Eq. 30-4) all fluctuate far more than do the concentrations of other metabolites during the 24-h cycle. These metabolites increase over 10-fold concentration at night and decrease by day. During the daytime the serotonin *N*-acyltransferase, which forms the precursor, is rapidly and apparently irreversibly inactivated, perhaps through a disulfide exchange reaction.[1081] Bright light will reset the circadian cycle[1084] keeping it approximately (circa) daily. The effect of light is apparently a result of signals sent to the hypothalamus from the optic nerves. In chickens and in lower vertebrates the pineal gland may directly sense light passing through the skull.

The circadian cycle is not the only timing device used by animals. A short-term interval timer helps male doves to know how long to sit on a nest[1085] and helps all of us in timing everyday tasks.

We spend a third of our lives asleep, but our understanding of sleep from a molecular viewpoint is minimal. Sleep is essential for the life of mammals, which die if completely deprived of sleep. It has been shown that during prolonged sleep deprivation sleep-inducing material accumulates in the brain. One such substance, isolated from human urine, appears to be a peptide containing glutamate, alanine, diaminopimedic acid, and muramic acid.[1086] Thus, it resembles a fragment of bacterial peptidoglycan. Prostaglandin D_2 also induces sleep.[1087] Hayashi proposed that a balance between this substance and prostaglandin E_2, which induces wakefulness, is in part responsible for the sleep–wake cycle.[1087] More recently oleic acid amide (**oleamide**; p. 382) was identified as a sleep-inducing compound. A fact observed by everyone is that the longer one is awake the higher the probability of going to sleep. The accumulation of sleep inducers is part of a homeostatic mechanism. On the other hand, the circadian cycle probably provides the signal to awake and tends to consolidate our sleep into the characteristic 8-hour period.[1088] The melatonin level, which drops in daylight, plays a role.[1089] Release of adrenocorticotropin (ACTH) one hour before waking may also be important.[1090]

During much of the night's sleep the EEG is characterized by the slow 5- to 6-Hz waves.[989,993] However, after ~90 min there is an ~10-min period of **rapid eye movement** (**REM**) **sleep** during which the EEG resembles that of an awake person and dreaming occurs. The closed eyes move rapidly in unison, breathing is irregular, and the heart rate increases. Motor neurons are inhibited allowing only minimal body movement. Three more periods of slow-wave sleep, each shorter than the preceding one, are followed by REM sleep. The REM sleep periods become successively longer. The fourth period lasts 20–30 min and is followed by awakening. All placental and marsupial mammals follow a similar sleep pattern and all dream.[989] The importance of dreaming is not obvious[1088a] but is often thought of as a reprocessing of memory, a means of ridding the mind of unneeded memories, a process of **unlearning**. However, this is uncertain as is the relationship of sleep to learning and memory.[1088a,b]

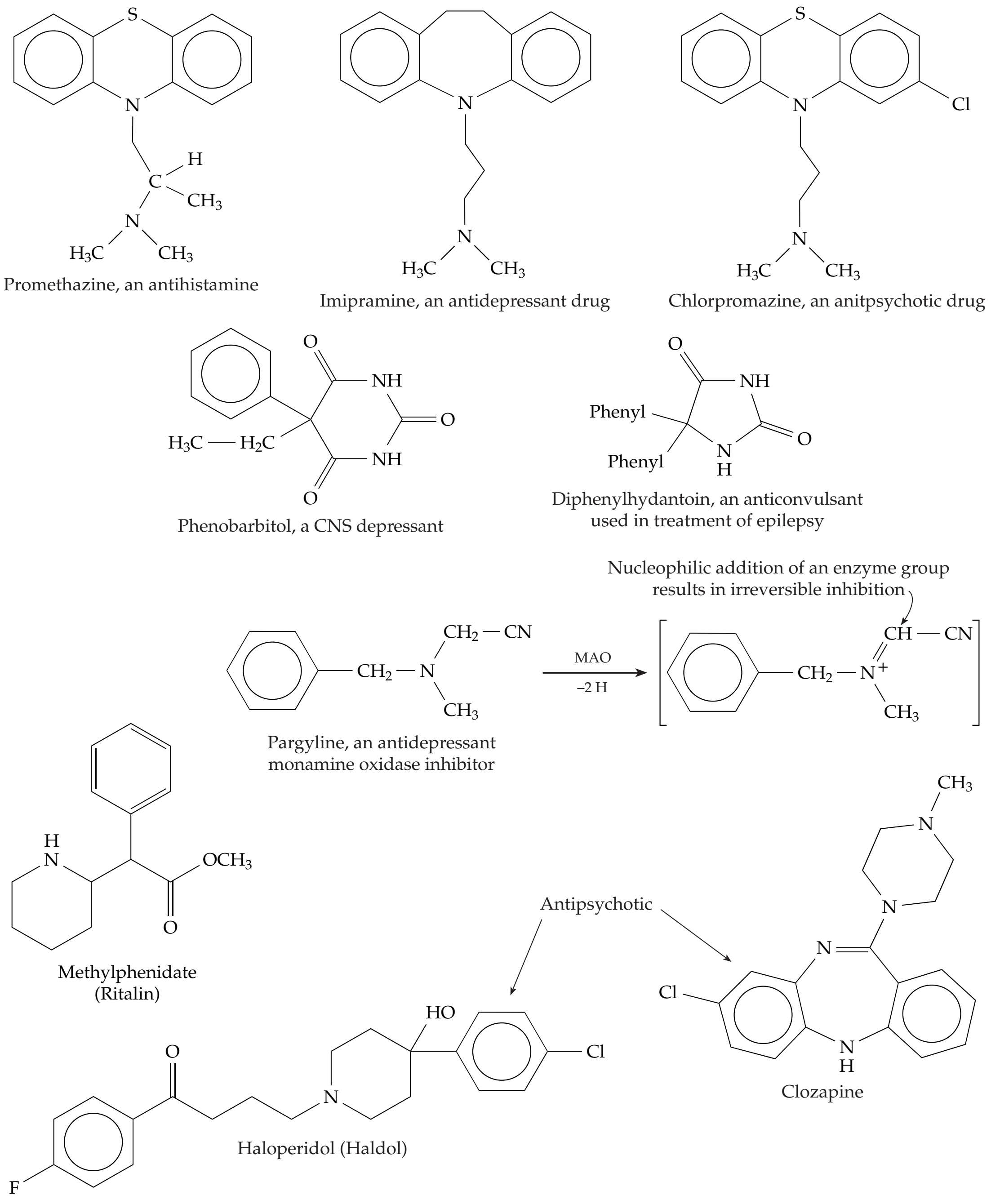
A number of disorders of sleep are known. Among these is **narcolepsy**, uncontrollable, sudden daytime sleepiness. It affects 1 in 2000 individuals.[1088] The same occurs in dogs.[1091] After a 10-year effort at great expense the narcolepsy gene of dogs (*canarc*-1) was located by positional cloning.[1088,1092] The corresponding human (and rat) gene was independently discovered by other investigators. It encodes a receptor for neuropeptides produced by the hypothalamus and named **hypocretins** or **orexins** for their stimulation of appetite. It seems probable that the hypocretin / orexin neuropeptides are involved in promoting wakefulness. Another sleep disorder is **familial advanced sleep phase syndrome**. Persons with this trait are "morning larks" who tend to fall asleep at ~7:30 p.m. and awake suddenly at ~4:30 a.m., about four hours in advance of a typical sleep period. A missence Ser → Gly mutation in the human period gene (hper2) has been found.[1093]

Some mammals hibernate. Special blood proteins that induce hibernation apparently control the process.[1094]

14. Mental Illness

Whereas many metabolic defects affect only a small number of individuals, emotional illnesses including depression, **schizophrenia**, and other **affective disorders** at one time or another afflict a large fraction of the population. Autism affects thousands of children.[1055] Parkinson disease and **Alzheimer disease** are just two of a number of degenerative neural diseases attacking older people. Less commonly, young persons contact **multiple sclerosis** and **muscular dystrophy,** which is often a disease of neuromuscular junctions.

Depression. Depression is our most common mental problem. One in four women and one in ten men will have a major depression during their lifetime.[1095] More than 15 million people in the United States are affected by severe depression in any given year and more than 30,000 may commit suicide.[1096,1097] Worldwide psychiatric problems, mostly depression, account for 28% of all disabilities.[1098] The **biogenic amine hypothesis** states that depression results from the depletion of neurotransmitters in the areas of the brain involved in sleep, arousal, appetite, sex drive, and psychomotor activity. An excess of transmitters is proposed to give rise to the manic phase of the bipolar (manic–depressive) cycle that is sometimes observed. In support of this hypothesis is the observation that administration of reserpine precipitates depression, which may be serious in 15–20% of hypertensive patients receiving the drug. Similar effects are observed with the dopa decarboxylase inhibitor **α-methyldopa**



Promethazine, an antihistamine

Imipramine, an antidepressant drug

Chlorpromazine, an anitpsychotic drug

Phenobarbitol, a CNS depressant

Diphenylhydantoin, an anticonvulsant used in treatment of epilepsy

Nucleophilic addition of an enzyme group results in irreversible inhibition

Pargyline, an antidepressant monamine oxidase inhibitor

Methylphenidate (Ritalin)

Antipsychotic

Haloperidol (Haldol)

Clozapine

Figure 30-33 Some drugs used to treat psychiatric disorders. See also Figs. 30-25 and 30-28.

(Fig. 30-27). The fact that L-tryptophan has some antidepressant activity, but L-dopa does not, was one clue that a low concentration of serotonin (5-hydroxytryptamine) might be responsible for depression. Excessive formation of histamine[1099] and decreased formation of tyramine and octopamine[1100] have also been suggested as causes of depression.

Strong support for the biogenic amine theory of depression is provided by the powerful antidepressant effect of inhibitors of monoamine oxidase. An example is **pargyline** (Fig. 30-33), which forms a covalent adduct with the flavin of MAO.[1101] Although effective, this drug is somewhat dangerous. Because their monoamine oxidase activity is so low, patients taking pargyline have been killed by ingesting compounds such as tyramine, which occurs in cheese. Less easy to understand but clinically more important are tricyclic antidepressants such as **imipramine** (Fig. 30-33),[746a] whose antidepressant action was discovered accidentally. Notice the close similarity to chlorpromazine but the greater flexibility of the central ring.[1102] Imipramine was found to block transporters of both noradrenaline and serotonin. In 1986, the less toxic serotonin reuptake inhibitor fluoxetine (Prozac; Fig. 30-28) was introduced and is now used by many millions of people.[1103,1104] Nevertheless, its mode of action is not entirely clear. For example, it blocks nicotinic acetylcholine receptors and may have many other effects. Interestingly, depression sometimes responds to a placebo just as well as to an antidepressant drug.[1105] In addition to newer drugs related to Prozac, antagonists of substance P are also effective antidepressants.[1106] MRI images of brains of depressed patients show that hippocampal volume has decreased and suggest that formation of new neurons is inhibited.[1106a] Antidepressants seem to stimulate growth of new cells as does exercise, which also has an antidepressant effect.[1107] Dietary treatment can also help.[843] Among older people depression may be caused by deficiency of vitamin B_{12} and can be treated by injection of the vitamin.[1108] An **anxiety peptide** that may be the natural ligand for benzodiazepine receptors has been reported.[1109]

Another recognized type of depression is **seasonal affective disorder** (**SAD**). People in far northern or southern latitudes develop this condition in the winter, apparently from lack of sunshine needed to lower the melatonin level in the morning (see Section 13). Light therapy is beneficial.[1110] Persons with the SAD syndrome also tend to crave carbohydrates and to stay in bed for 9–10 hours.

An effective treatment for **bipolar disorder** (manic–depressive illness) is the administration of lithium salts.[445,1111–1113a] Inhibition of the hydrolysis of inositol phosphate by Li^+ (Fig. 11-9) may be related to its therapeutic effect. Reduced phosphatidylinositol turnover may dampen responses to neurotransmitters.[1114] Li^+ may affect gene expression in neuropeptide-secreting neurons.[1115] Bipolar disorder apparently has more than one cause. There are strong indications of genetic susceptibility,[1116] and genes that increase susceptibility have been located on chromosomes 4, 12, 13, 18, 21, and X.[1117]

Schizophrenia. Among the most baffling of mental illnesses are the group of diseases known as schizophrenia. They involve thought disorder, disturbance of the affect, and withdrawal from interactions with other people. Hallucinations and paranoid feelings are common.[813] In some cases a striking loss of gray matter in some areas of the brain is revealed by MRI scans.[1117a,b] The schizophrenias are of varying degrees of severity and shade continuously into the affective or mood disorders, which include manic-depression and depression. As many as one person in a hundred is affected by schizophrenia.[1118–1119a] There is a complex genetic susceptibility.[1119,1120] One theory about the persistence of the genes favoring schizophrenia is that they are also associated with creativity.[1121]

A revolution in the treatment of the schizophrenias, as well as in thinking about mental illnesses, took place following the synthesis, in 1950, of the antipsychotic drug **chlorpromazine** (Fig. 30-33). At about the same time the effect of the *Rauwolfia* alkaloid reserpine (Fig. 25-12) in calming mentally disturbed persons was rediscovered. The Indian plant *Rauwolfia* had been used for centuries in Hindu medicine for the same purpose. The tricyclic phenothiazines such as promethazine (Fig. 30-33) earlier had been found to have powerful **antihistamine** activity. It was the search for better antihistamine drugs that led to the synthesis of chlorpromazine.[1122] As many as 250 million people throughout the world were treated with chlorpromazine and related drugs in the 20 years following its discovery before newer and safer drugs (e.g., **clozapine**; Fig. 30-33) were developed.[1123] What does chlorpromazine do? A possible clue comes from the fact that it sometimes induces serious "extrapyramidal" side effects including tremors and other symptoms of Parkinson disease. This suggested that chlorpromazine may block dopamine receptors in the corpus striatum, thereby precipitating a functional deficiency of dopamine.[1124] If so, it is possible that schizophrenia may result from an overactivity of dopamine neurons, perhaps including some of the same neurons that are hypoactive in Parkinson disease. Supporting this view is the observation that amphetamines (Fig. 30-27), which may substitute for dopamine, worsen the symptoms of schizophrenia and in very high doses induce striking schizophrenialike symptoms in normal individuals.

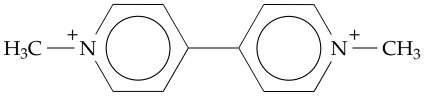
A stereotyped compulsive behavior is induced both in humans and in laboratory animals by amphetamines. This provided the basis for a method that has been used to measure the action of drugs on amphetamine-sensitive centers of the brain. A lesion in the nigrostriatal bundle on one side of a rat brain was made by injection of a neurotoxic compound such as 6-hydroxydopamine. This caused degeneration of dopamine-containing neurons on one side of the brain. When rats that had been injured in this way were given amphetamines, they developed a compulsive rotational behavior. Administration of chlorpromazine and several other antipsychotic drugs neutralized this behavior and in direct proportion to the efficacy in clinical use, an observation that also supports the theory that schizophrenia involves overactivity of dopamine neurons.

If schizophrenia results from an elevated dopamine content of the brain, the fault may lie with either an oversupply or a reduced rate of metabolism of dopamine. The possibilities of reduced activity of monoamine oxidase or of dopamine β-hydroxylase have both been suggested. The plasma level of the dopamine metabolite **homovanillic acid** (Fig. 30-26) is elevated in schizophrenia and is correlated with the severity of the illness,[1125] suggesting the hypothesis of a decreased rate of metabolism. Possible defects in dopamine receptors may be at fault.[1126]

Chlorpromazine may also act on brain cholinergic neurons.[1127] Blockage of muscarinic acetylcholine receptors in the brain by belladonna alkaloids such as atropine (Fig. 30-22) has often been used in treatment of Parkinson disease. Apparently antagonizing acetylcholine action is to some extent functionally equivalent to increasing dopamine concentrations. There is evidence that suggests a role for cholecystokinin (CCK) in development of schizophrenia. CCK-containing neurons interact with dopaminergic neurons in the midbrain.[1128] GABA neurons in the prefrontal cortex may be faulty.[1129] Excessive glutamate may also induce schizophrenia. The schizophrenia-like symptoms induced by phenylcyclidine (p. 1796) are eased by antagonists of metabotropic glutamate receptors. This suggests another possible therapy.[1130] Among other suggested causes of schizophrenia are dysregulation by **retinoids**[1131] and action of retroviruses.[1132] Demyelination in portions of the prefrontal cortex may disrupt neural connectivity.[1132a] Recent genetic evidence points to a possible defect in proline dehydrogenase which reduces Δ^1-pyrroline-5-carboxylate to L-proline (Fig. 24-9).[1132b]

Numerous theories of mental illness have embodied proposals that a toxic metabolite is produced in abnormal quantities. An example is **6-hydroxydopamine** (Fig. 30-26), which is known to damage dopamine-containing neurons.[1133] Overactive methylation of catecholamines has also been suggested as a cause of mental disorders.[1134] The hallucinogen **3,4-dimethoxyphenylethylamine** (Fig. 30-26) has been identified in urine during acute schizophrenic attacks, but the variability is so high that no definite conclusion has been reached. *N*-Methylation of serotonin yields **bufotenin** (*N*-methylserotonin) and ***N*-dimethylserotonin**, known hallucinogenic agents. Enzymatic synthesis of the latter by human brain and other tissues has been demonstrated,[1135] and administration of tryptophan and methionine to schizophrenic patients exacerbates their illness.

Another theory of mental illness postulates endogenous alkaloid formation. Aldehydes formed by oxidation of catecholamines as well as formaldehyde and acetaldehyde are present in tissues in small amounts. Condensation with amines could generate Schiff bases and alkaloids as in Fig. 25-10. This "plant chemistry" is spontaneous and can apparently take place in the brain, where it may have a potent effect. Incubation of tryptamine derivatives with 5-methyltetrahydrofolic acid and an enzyme preparation from brain gives **tryptolines**. Dopamine and its derivatives form related tetrahydroisoquinolines such as the product that arises from reaction with acetaldehyde (see Eq. 30-5). This product has been found in elevated amounts in alcoholics (who synthesize excess acetaldehyde), in phenylketonurics, and in L-dopa-treated patients with Parkinson disease.[1136]



Paraquat

Epilepsy. The brain disorders known as epilepsies affect 1–2 % of the population worldwide. Characteristic of epilepsies are recurrent **seizures**, sudden brief changes in behavior caused by the simultaneous, disordered firing of large numbers of neurons in the brain. Many seizures are thought to be initiated in specific areas of the cerebral cortex. For example, seizure-induced firing of neurons in the thumb area of the right motor cortex will be accompanied by rhythmic jerking in the left thumb. More than 40 different types of epilepsy are known.[1137–1138a]

GABA is the principal inhibitor neurotransmitter, and one cause of epilepsy may be a deficiency in GABA formation from glutamate. The brain contains two isoforms of glutamate decarboxylase, designated GAD65 and GAD67, in accordance with their molecular masses in kDa. They are encoded by separate genes.[1139,1140] GAD67 is formed mainly in cell bodies of neurons, binds its cofactor PLP tightly, and is essential to survival of young mice. GAD65 is associated mainly with nerve termini, where it is anchored, apparently by association with other proteins to the membranes of synaptic vesicles.[1141] It binds PLP weakly. Some convulsive agents such as 1,1-dimethylhydrazine are thought to act by interfering with PLP-dependent enzymes (Box 14-C) among which is GAD.[1140,1142,1143] Convulsions are one of the most striking symptoms of a severe vitamin B_6 deficiency. A zinc deficiency can also cause convulsions, apparently because pyridoxine kinase is a Zn•ATP-requiring protein and the rate of synthesis of PLP is too slow to supply apo-GAD with the PLP needed for GABA synthesis. The PLP in GAD65 undergoes rapid substrate-dependent transamination to pyridoxamine phosphate (see Chapter 14), which must be replaced by new PLP.[1143,1144]

Epilepsy may arise also from defects in a GABA transporter[1145] or receptor.[1146] One form of epilepsy is a triple-repeat disease of cystatin B (Table 26-4). Mutation in potassium channels,[1147] glutamate receptors,[1148] absence of neuropeptide Y,[1149] and absence of L-isoaspartyl / D-aspartyl *O*-methyltransferase (Box 12-A)[1150] have all been associated with epilepsy.

Neurodegenerative diseases. As many as 5% of persons of age 65 and 20% of those of age 80 are afflicted with the progressive senile dementia known as **Alzheimer disease**. The condition is characterized by a gradual loss of memory and of the abilities to speak, think, or take care of one's self. Histologically Alzheimer disease is marked by the accumulation within neurons of **paired helical filaments**. These filaments, of 10 nm diameter, twist about each other to form a helix with an 80-nm pitch. The helices aggregate to create **neurofibrillary tangles**. The tangles are composed largely of a highly phosphorylated form of the microtubule-associated protein **tau** (p. 372)[1151–1155a,b] together with phosphorylated neurofilaments, apolipoprotein E (p. 1183), and other materials. The tangles are found in the cell bodies, axons, and dendrites of neurons in the hippocampus, amygdala, cerebral cortex, and other areas of the brain. Tangles may also be present in Parkinson disease, in the nearly extinct **Guam disease**,[1156,1157] and in some types of prion disease.[1158] Outside of the diseased neurons are numerous, spherical **amyloid plaques**. Their principal component is a 40- to 43-residue fragment called **amyloid β-protein** (**Aβ**), which appears to be toxic to neurons. Aβ is cut from a larger **amyloid precursor protein** (APP).[1154,1159–1160a] The APP gene is a member of a family of 16 related genes found in many organisms including nematodes, flies, and mammals. In humans the APP gene is found on chromosome 21, the chromosome that is present in three copies in **Down syndrome**. People with Down syndrome who live into their late thirties or beyond develop Alzheimer disease,[1153] presumably from excessive synthesis of APP. Both APP and its cleavage product Aβ are formed by nonneuronal cells throughout the body. However, the Aβ plaques form only in the brain, and the APP gene is essential for life. The rare **familial British dementia** resembles Alzheimer disease in producing amyloid plaques and neurofibrillary tangles. They appear principally in the cerebellum and arise from a different precursor protein.[1161,1162]

There are many other neurodegenerative diseases, some with a high incidence, and others rare. They include **Parkinson disease** (p. 1790), **Huntington disease** (Table 26-4), **spinal muscular atrophy** (SMA; a leading hereditary cause of infant mortality),[1162a,b] amyotrophic lateral sclerosis (**ALS**), prion diseases (Box 29-E), **ataxias**, and other diseases caused by triple-repeat DNA sequences (Table 26-4) and X-linked adrenoleukodystrophy (ALD; p. 945).[1163] In the last, membrane function is disrupted. Although these diseases arise from a variety of causes many of them have in common amyloidosis, the deposition of insoluble proteins in or around neurons.[1163a]

Parkinson disease, some cases of Alzheimer disease, and some types of prion disease are accompanied by the presence of **Lewy bodies** within the cytoplasm of neurons and also in nearby glia. These deposits consist largely of a dense core of fibrils of **α-synuclein**, a small 140-residue protein abundant in various parts of the brain.[1164,1164a,b] Mutations in the α-synuclein gene are associated with autosomal-dominant inheritance of early-onset Parkinson disease.[1165,1166] Just as tau tends to be associated with microtubules, α-synuclein may function in cooperation with microfilaments.[1158] Studies of an autosomal-recessive form of inherited juvenile Parkinson disease led to mutations in a large (>1 Mbp) gene on chromosome 6. It encodes the 465-residue **Parkin.**[1167,1168] Parkin is an E3 ubiquitin ligase (p. 524), which ubiquitinates α-synuclein.[1168,1169] This finding suggests that abnormally slow degradation of synuclein may be an important cause of Parkinson disease.

One of the triple-repeat polyglutamine diseases discussed in Chapter 26 (Table 26-4) is Huntington disease. The defective **Huntingtin** is a cytosolic protein that normally protects neurons but fails when the polygultamine sequence becomes too long.[1170,1171] Neurons of the cerebral cortex and striatum die, apparently by apoptosis.[1172] Huntingtin interacts with p53, with a CREB-binding protein, and with an EGF receptor suggesting that it functions in regulation of transcription.[1172–1173a] One of the genes whose transcription is regulated is that of the neurotrophin known as **brain-derived neurotrophic factor** (**BDNF**).[1170]

Many approaches have been taken in therapy of Parkinson disease. As mentioned on p. 1790 enhancing dopamine production by administration of L-dopa or by use of MAO inhibitors is a standard treattment. Experimental gene therapy with a glial cell line–derived neurotrophic factor also appears promising.[1174,1175]

Another aspect of neurodegeneration involves oxidative damage. A clue comes from amyotrophic lateral sclerosis (**ALS**), which struck down the New York Yankees baseball player Lou Gehrig, after he had started 2130 consecutive games over a 15-year period. ALS (Lou Gehrig disease) is the most prevalent of more than 70 diseases that cause loss of motor neurons.[1176] As pointed out on p.1075, the cause of a rare hereditary form of ALS is a defect in superoxide dismutase, which appears to promote excessive formation of free radicals.[1177] However, this interpretation is uncertain.[1178] Parkinson disease induced by the compound MPTP (Eq. 30-4) may also arise as a result of free radical damage.[1179] Among possible effects, MPTP may induce apoptosis.[1180] Both oxidative damage and apoptosis may be factors in Alzheimer and other neurodegenerative diseases as well.[1181,1182]

In every disease in which an abnormal protein is found there must be pathways of processing the protein to generate its functional form and pathways for degradation. These pathways are being investi gated for all of the neurodegenerative diseases and none more intensively than for Alzheimer disease. The amyloid precursor protein APP is an integral membrane glycoprotein with a large ~687-residue extracellular N-terminal portion, which resembles a cell surface receptor. It contains both a protease inhibitor-like domain and a zinc-binding region, which can be phosphor-ylated. It binds heparin and collagen as well as other proteins.[1183,1184] Rare mutations that cause early-onset familial Alzheimer disease are found in the APP gene. Some of these mutations alter the regulation of pre-mRNA splicing. Splicing generates eight different APP isoforms vaying in length from 677 to 770 residues. The properties of the isoforms vary. For example, if the 18 residues of exon 15 are spliced out a new motif for posttranslational modification is created by fusion of exons 14 and 16. The newly created sequence ENEGSG is recognized by a xylosyltransferase, which initiates formation of the terminal unit for glycosaminoglycan formation. The resulting proteoglycan is known as **appican** (p. 1154).[1185] The precursor protein APP is transported down axons to the nerve endings and is proteolytically cleaved to form the insoluble amyloid deposits. Alzheimer disease may occur when there is excessively rapid proteolysis of the precursor or if there is a failure to metabolize the amyloid protein.[1186] The folding and glycosylation reactions of APP occur in the ER and the Golgi, but the major problem in Alzheimer disease appears to be in subsequent proteolytic processing, some of which may occur as the APP is being transported through the Golgi to the cytoplasmic membrane. The protein may be cleaved at three sites by enzymes known as α-, β-, and γ-**secretases** as is indicated in Fig. 30-34, in which the protein is represented as an unfolded "stick."[1154,1187–1189a] Most of the mutations in APP that cause Alzheimer disease are near these three cleavage sites. As indicated in Fig. 30-34, cutting at the α site liberates into the extracellular space the large N-terminal portion as a soluble protein called APPsα.[1190,1191] It is thought to have a protective effect on neurons. If this cleavage occurs the protein is not cut at the β site and fragment Aβ is not formed. However, if cleavage by β-secretase (beta-site APP-cleaving enzyme or BACE) occurs first and is accompanied by or followed by cleavage at the γ-site, Aβ is liberated (Fig. 30-34). The β-secretase is an integral membrane protein, which carries a pepsinlike domain in its lumenal (or extracellular) part.[1192–1194] Since the Aβ peptide in an aggregated form appears to be toxic to neurons, a logical therapy for Alzheimer disease may be to block either the β- or γ-secretase.[1195,1196]

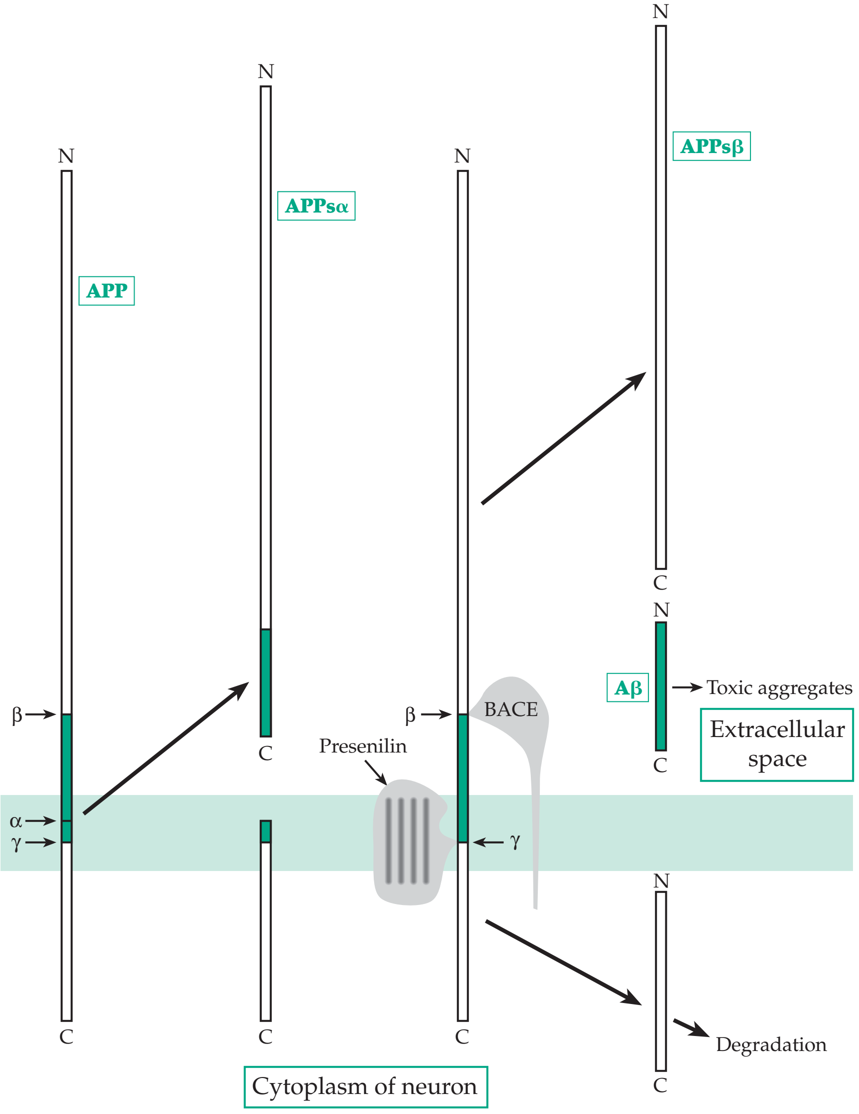


Figure 30-34 Cleavage of the amyloid precursor protein APP with liberation of amyloid Aβ protein. The proteins are represented as sticks (not to scale) but in reality contain both intracellular and extracellular globular domains.

The γ-secretase has been difficult to locate but has been identified as a result of other rare familial forms of Alzheimer disease. These are caused by mutations in genes for proteins known as **presenilin-1** (on chromosome 14) and **presenilin-2** (on chromosome 1).[1188,1197,1198] The presenilins are integral membrane proteins with multiple transmembrane helices.

They have been regarded as regulators of γ-secretase, but there is much evidence that the presenilin molecules may be cleaved proteolytically and that the C- and N-terminal domains formed in this way may associate to form an unusual aspartyl protease. It, too, is a target for inhibitors.[1198] The picture is made more complex by the fact that presenilins form complexes with other proteins. These include a newly discovered protein **nicastrin**,[1199–1200a] a large 709-residue transmembrane glycoprotein. Nicastrin not only seems to modulate presenilin action but also participates in an important developmental process via the highly conserved **Notch pathway** (Chapter 32).[1200] Many other proteins are found in the amyloid plaques of Alzheimer disease. Among them are acetylcholinesterase,[1201] proteoglycans,[1202] hydroxyacyl-CoA dehydrogenase,[1203] GM1 ganglioside,[1204] apolipoprotein A-1,[1187] and lithostatine.[1205]

What are the possible adverse consequences of accumulation of the Aβ protein? It may cause inflammation by activation of **microglia**,[1157] which may cause damage by release of NO.[1206] Aβ may induce death of neurons by apoptosis.[1201,1207–1209] A defect in protesomal degradation may be a factor.[1208] Both Aβ and the prion protein may promote oxidative damage. The brain derives most of its energy from oxidative metabolism, a major source of damaging radicals. Mitochondria are found in dendrites as well as cell bodies.[1210] Methionine residues in glycine-rich parts of the Aβ and prion proteins are suspected as centers of free radical formation.[1202,1211]

Both amyloid plaques and the tangles of protein tau-containing paired helical filaments are typically present in Alzheimer disease. Which comes first? Some hereditary neurodegenerative diseases are known in which tau filaments are present in neurons and sometimes also in glia.[1203,1212,1213] Since mutations in tau don't lead to Alzheimer disease whereas mutations affecting APP do, it is often assumed that the primary defect in the disease is with APP and that accumulating Aβ induces the observed changes in tau. However, this is by no means certain. Six different isoforms of tau (the longest with 441 residues) are created by alternative splicing of the mRNA. During its normal functioning tau is phosphorylated and carries an average of 2–3 phospho groups. In Alzheimer disease the level of tau is greatly increased (4- to 8-fold) and the molecules carry 5–9 phospho groups.[1204,1213a] It is this hyperphosphorylated tau that forms the paired helical filaments and tau tangles, which appear to clog the slender neurons.

What does tau do normally? Although it has been studied for many years, its exact functions are elusive. However, the role of the microtubules in axonal transport is well established. The tau isoforms may play a functional role in this process. The hyperphosphorylated tau of Alzheimer disease doesn't promote proper assembly of microtubules and may interfere with axonal transport of materials along the microtubules (see p. 1119).[1214,1215] Alzheimer disease may reflect an imbalance between the phosphorylation and dephosphorylation processes. Another possible problem with tau may be slow isomerization of prolyl linkages because of a deficiency of a prolyl *cis–trans* isomerase (Box 9-F).[1216]

Until 1993 **apolipoprotein E** was best known for its central role in plasma lipoproteins and cholesterol transport (Fig. 21-1). However, one of the three common alleles of the apoE gene confers a significant risk of development of Alzheimer disease.[1217,1218] A high blood cholesterol level is also correlated with increased risk.[1219,1220] Membrane abnormalities in mitochondria have been associated with Alzheimer disease.[1221] Also related to membranes and lipid metabolism, **vitamin E** appears to combat Alzheimer disease.[843,1218]

Environmental and nutritional factors may also affect the development of Alzheimer disease and other mental illness. Aluminum frequently accumulates in the neurons containing neurofibrillary tangles.[1206,1222] Copper and zinc ions can cause the amyloid Aβ to aggregate. However, Zn^{2+} may actually protect against neurotoxicity.[1223] The amino acid β-*N*-methylamino-L-alanine, a constituent of the toxic seeds of a type of palm (*Cycas circinalis* L.), may have induced both ALS and Guam disease, a condition resembling Parkinson disease, in a population in Guam that traditionally used these seeds as food.[1157]

Can neurodegenerative diseases be prevented or delayed? Much evidence suggests that the answer is yes. Rare early onset forms pose a special problem, but for most of us maintaining an active life style, using our minds, and choosing a good diet with adequate amounts of vitamins and essential ω3 fatty acids (Box 21-B) may be very helpful.[843] New methods of treatment are being tested. Antiinflammatory drugs are helpful,[1218,1218a] and even vaccination against Aβ and other amyloid proteins appears possible.[1224] Is it possible that antibodies and phagocytic cells can clear the cobwebs from our brains?

References

1. Norman, A., and Litwack, G. (1987) *Hormones*, Harcourt Brace Jovanovich, Orlando, Florida
2. Baulieu, E.-E., and Kelly, P. A., eds. (1990) *Hormones*, Hermann, Paris, France
3. Weintraub, B. D., ed. (1994) *Molecular Endocrinology*, Raven Press, New York
4. Gutkind, J. S. (1998) *J. Biol. Chem.* **273**, 1839–1842
5. Lefkowitz, R. J. (1998) *J. Biol. Chem.* **273**, 18677–18680

5a. Benard, O., Naor, Z., and Seger, R. (2001) *J. Biol. Chem.* **276**, 4554–4563

5b. Rockman, H. A., Koch, W. J., and Lefkowitz, R. J. (2002) *Nature (London)* **415**, 206–212

5c. Neves, S. R., Ram, P. T., and Iyengar, R. (2002) *Science* **296**, 1636–1639

6. IIri, T., Farfel, Z., and Bourne, H. R. (1998) *Nature (London)* **394**, 35–38
7. Zheng, B., De Vries, L., and Farquhar, M. G. (1999) *Trends Biochem. Sci.* **24**, 411–414
8. Janetopoulos, C., Jin, T., and Devreotes, P. (2001) *Science* **291**, 2408–2411
9. Gorski, J., Toft, D., Shyamala, G., Smith, D., and Notides, A. (1968) *Rec. Progr. in Horm. Res.* **24**, 45–80
10. Jensen, E. V., Suzuki, T., Kawashima, T., Stumpf, W. E., Jungblut, P. W., and DeSombre, E. R. (1968) *Proc. Natl. Acad. Sci. U.S.A.* **59**, 632–638
11. Sluyser, M., ed. (1985) *Interaction of Steroid Hormone Receptors with DNA*, Ellis Horwood, Chichester
12. O'Malley, B. W., and Schrader, W. T. (1976) *Sci. Am.* **234**(Feb), 32–43
13. Wallis, M., Howell, S. L., and Taylor, K. W. (1985) *The Biochemistry of the Polypeptide Hormones*, Wiley, Chichester
14. McEwen, B. S. (1976) *Sci. Am.* **235**(Jul), 48–58
15. Pines, M. (1975) *Saturday Rev.* **Aug. 9**, 14
16. Guillemin, R., and Burgus, R. (1972) *Sci. Am.* **227**(Nov), 24–33
17. Sharrocks, A. D., Yang, S.-H., and Galanis, A. (2000) *Trends Biochem. Sci.* **25**, 448–453
18. Cohen, P. (2000) *Trends Biochem. Sci.* **25**, 596–601
19. Hafen, E. (1998) *Science* **280**, 1212–1213
20. Peles, E., Schlessinger, J., and Grumet, M. (1998) *Trends Biochem. Sci.* **23**, 121–124
21. Weng, G., Bhalla, U. S., and Lyengar, R. (1999) *Science* **284**, 92–96
22. Roberts, C. J., Nelson, B., Marton, M. J., Stoughton, R., Meyer, M. R., Bennett, H. A., He, Y. D., Dai, H., Walker, W. L., Hughes, T. R., Tyers, M., Boone, C., and Friend, S. H. (2000) *Science* **287**, 873–880
23. Tomkins, G. M. (1975) *Science* **189**, 760–763
24. Lenard, J. (1992) *Trends Biochem. Sci.* **17**, 147–150
25. Oosterom, J., Nijenhuis, W. A. J., Schaaper, W. M. M., Slootstra, J., Meloen, R. H., Gispen, W. H. H., Burbach, J. P. H., and Adan, R. A. H. (1999) *J. Biol. Chem.* **274**, 16853–16860
26. Muller, L., Zhu, P., Juliano, M. A., Juliano, L., and Lindberg, I. (1999) *J. Biol. Chem.* **274**, 21471–21477
27. Jutras, I., Seidah, N. G., Reudelhuber, T. L., and Brechler, V. (1997) *J. Biol. Chem.* **272**, 15184–15188
28. Feliciangeli, S., Kitabgi, P., and Bidard, J.-N. (2001) *J. Biol. Chem.* **276**, 6140–6150
29. Lynch, D. R., and Snyder, S. H. (1986) *Ann. Rev. Biochem.* **55**, 773–799

29a. Bateman, R. C., Jr., Temple, J. S., Misquitta, S. A., and Booth, R. E. (2001) *Biochemistry* **40**, 11246–11250

30. Loh, Y. P., Parish, D. C., and Tuteja, R. (1985) *J. Biol. Chem.* **260**, 7194–7205

30a. Yasothornsrikul, S., Aaron, W., Toneff, T., and Hook, V. Y. H. (1999) *Biochemistry* **38**, 7421–7430

30b. Jutras, I., Seidah, N. G., and Reudelhuber, T. L. (2000) *J. Biol. Chem.* **275**, 40337–40343

30c. Rockwell, N. C., and Fuller, R. S. (2001) *Biochemistry* **40**, 3657–3665

31. Paladini, A. C., Pena, C., and Retegui, L. A. (1979) *Trends Biochem. Sci.* **4**, 256–260
32. Lowman, H. B., Cunningham, B. C., and Wells, J. A. (1991) *J. Biol. Chem.* **266**, 10982–10988
33. Nilsen-Hamilton, M., Shapiro, J. M., Massoglia, S. L., and Hamilton, R. T. (1980) *Cell* **20**, 19–28
34. Parfett, C. L. J., Hamilton, R. T., Howell, B. W., Edwards, D. R., Nilsen-Hamilton, M., and Denhardt, D. T. (1985) *Mol. Cell. Biol.* **5**, 3289–3292
35. Lee, S., and Nathans, D. (1988) *J. Biol. Chem.* **263**, 3521–3527
36. Abdul-Meguid, S. S., Shieh, H., Smith, W. W., Dayringer, H. E., Violang, B. N., and Bentle, L. A. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6434–6437
37. Ultsch, M. H., Somers, W., Kossiakoff, A. A., and de Vos, A. M. (1994) *J. Mol. Biol.* **236**, 286–299
38. Souza, S. C., Frick, G. P., Wang, X., Kopchick, J. J., Lobo, R. B., and Goodman, H. M. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 959–963
39. Sundström, M., Lundqvist, T., Rödin, J., Giebel, L. B., Milligan, D., and Norstedt, G. (1996) *J. Biol. Chem.* **271**, 32197–32203
40. de Vos, A. M., Ultsch, M., and Kossiakoff, A. A. (1992) *Science* **255**, 306–312
41. Wells, J. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1–6
42. Clackson, T., Ultsch, M. H., Wells, J. A., and de Vos, A. M. (1998) *J. Mol. Biol.* **277**, 1111–1128
43. Talmadge, K., Vamvakopoulos, N. C., and Fiddes, J. C. (1984) *Nature (London)* **307**, 37–40
44. Millstone, E., Brunner, E., and White, I. (1994) *Nature (London)* **371**, 647–648
45. Melmed, S. (1990) *N. Engl. J. Med.* **322**, 966–977
46. Souza, S. C., Frick, G. P., Yip, R., Lobo, R. B., Tai, L.-R., and Goodman, H. M. (1994) *J. Biol. Chem.* **269**, 30085–30088
47. Gertler, A., Grosclaude, J., Strasburger, C. J., Nir, S., and Djiane, J. (1996) *J. Biol. Chem.* **271**, 24482–24491
48. Herman, A., Helman, D., Livnah, O., and Gertler, A. (1999) *J. Biol. Chem.* **274**, 7631–7639
49. Cross, J. C., Werb, Z., and Fisher, S. J. (1994) *Science* **266**, 1508–1517
50. Darling, R. J., Wilken, J. A., Miller-Lindholm, A. K., Urlacher, T. M., Ruddon, R. W., Sherman, S. A., and Bedows, E. (2001) *J. Biol. Chem.* **276**, 10692–10699
51. Weller, C. T., Lustbader, J., Seshadri, K., Brown, J. M., Chadwick, C. A., Kolthoff, C. E., Ramnarain, S., Pollak, S., Canfield, R., and Homans, S. W. (1996) *Biochemistry* **35**, 8815–8823
52. Moyle, W. R., Campbell, R. K., Rao, S. N. V., Ayad, N. G., Bernard, M. P., Han, Y., and Wang, Y. (1995) *J. Biol. Chem.* **270**, 20020–20031
53. Seger, R., Hanoch, T., Rosenberg, R., Dantes, A., Merz, W. E., Strauss, J. F., III, and Amsterdam, A. (2001) *J. Biol. Chem.* **276**, 13957–13964
54. Policastro, P. F., Daniels-McQueen, S., Carle, G., and Boime, I. (1986) *J. Biol. Chem.* **261**, 5907–5916
55. Sairam, M. R. (1989) *FASEB J.* **3**, 1915–1926
56. Kaetzel, D. M., and Nilson, J. H. (1988) *J. Biol. Chem.* **263**, 6344–6351
57. Fernandez, L. M., and Puett, D. (1996) *J. Biol. Chem.* **271**, 925–930
58. Weiss, J., Axelrod, L., Whitcomb, R. W., Harris, P. E., Crowley, W. F., and Jameson, J. L. (1992) *N. Engl. J. Med.* **326**, 179–183
59. Laue, L., Chan, W.-Y., Hsueh, A. J. W., Kudo, M., Hsu, S. Y., Wu, S.-M., Blomberg, L., and Cutler, G. B., Jr. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1906–1910
60. Schally, A. V., Arimura, A., and Kastin, A. J. (1973) *Science* **179**, 341–350
61. Ling, N., Zeytin, F., Böhlen, P., Esch, F., Brazeau, P., Wehrenberg, W. B., Baird, A., and Guillemin, R. (1985) *Ann. Rev. Biochem.* **54**, 403–423
62. Laakkonen, L. J., Guarnieri, F., Perlman, J. H., Gershengorn, M. C., and Osman, R. (1996) *Biochemistry* **35**, 7651–7663
63. Bulant, M., Delfour, A., Vaudry, H., and Nicolas, P. (1988) *J. Biol. Chem.* **263**, 17189–17196
64. Perez de la Cruz, I., and Nillni, E. A. (1996) *J. Biol. Chem.* **271**, 22736–22745
65. Qi, L. J., Leung, A. T., Xiong, Y., Marx, K. A., and Abou-Samra, A.-B. (1997) *Biochemistry* **36**, 12442–12448
66. Taylor, A. L., and Fishman, L. M. (1988) *N. Engl. J. Med.* **319**, 213–222
67. Strulovici, B., Tahilramani, R., and Nestor, J. J., Jr. (1987) *Biochemistry* **26**, 6005–6011
68. Flanagan, C. A., Rodic, V., Konvicka, K., Yuen, T., Chi, L., Rivier, J. E., Millar, R. P., Weinstein, H., and Sealfon, S. C. (2000) *Biochemistry* **39**, 8133–8141
69. Marshall, J. C., and Kelch, R. P. (1986) *N. Engl. J. Med.* **315**, 1459–1468
70. Pincus, S. M., Mulligan, T., Iranmanesh, A., Gheorghiu, S., Godschalk, M., and Veldhuis, J. D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14100–14105
71. Koch, B. D., Dorflinger, L. J., and Schonbrunn, A. (1985) *J. Biol. Chem.* **260**, 13138–13145
72. Reichlin, S. (1983) *N. Engl. J. Med.* **309**, 1556–1564
73. Pfeiffer, M., Koch, T., Schröder, H., Klutzny, M., Kirscht, S., Kreienkamp, H.-J., Höllt, V., and Schulz, S. (2001) *J. Biol. Chem.* **276**, 14027–14036
74. Cai, R.-Z., Szoke, B., Lu, R., Fu, D., Redding, T. W., and Schally, A. V. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1896–1900
75. Nikolics, K., Mason, A. J., Szönyi, E., Ramachandran, J., and Seeburg, P. H. (1985) *Nature (London)* **316**, 511–517
76. Tsonis, C. G., and Sharpe, R. M. (1986) *Nature (London)* **321**, 724–725
77. Sharif, M., and Hanley, M. R. (1992) *Nature (London)* **357**, 279–280
78. Nielsen, S., Chou, C.-L., Marples, D., Christensen, E. I., Kishore, B. K., and Knepper, M. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1013–1017
79. Innamorati, G., Le Gouill, C., Balamotis, M., and Birnbaumer, M. (2001) *J. Biol. Chem.* **276**, 13096–13103
80. Rosenthal, W., Antaramian, A., Gilbert, S., and Birnbaumer, M. (1993) *J. Biol. Chem.* **268**, 13030–13033
81. Nishimori, K., Young, L. J., Guo, Q., Wang, Z., Insel, T. R., and Matzuk, M. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11699–11704
82. Bloom, F. E. (1981) *Sci. Am.* **245**(Oct), 148–168
83. Larive, C. K., Guerra, L., and Rabenstein, D. L. (1992) *J. Am. Chem. Soc.* **114**, 7331–7337
84. Land, H., Schutz, G., Schmale, H., and Richter, D. (1982) *Nature (London)* **295**, 299–303
85. Brownstein, M. J., Russell, J. T., and Gainer, H. (1980) *Science* **207**, 373–378
86. Drenth, J. (1981) *J. Biol. Chem.* **256**, 2601–2602
87. Kiefer, L. L., Veal, J. M., Mountjoy, K. G., and Wilkison, W. O. (1998) *Biochemistry* **37**, 991–997
88. Haskell-Luevano, C., Cone, R. D., Monck, E. K., and Wan, Y.-P. (2001) *Biochemistry* **40**, 6164–6179
89. Seeley, R. J., Yagaloff, K. A., Fisher, S. L., Burn, P., Thiele, T. E., van Dijk, G., Baskin, D. G., and Schwartz, M. W. (1997) *Nature (London)* **390**, 349
90. Moore, J., Wood, J. M., and Schallreuter, K. U. (1999) *Biochemistry* **38**, 15317–15324

References

91. Carayannopoulos, M. O., Chi, M. M.-Y., Cui, Y., Pingsterhaus, J. M., McKnight, R. A., Mueckler, M., Devaskar, S. U., and Moley, K. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7313 – 7318
92. Aspinwall, C. A., Qian, W.-J., Roper, M. G., Kulkarni, R. N., Kahn, C. R., and Kennedy, R. T. (2000) *J. Biol. Chem.* **275**, 22331 – 22338
93. Rolland, F., Winderickx, J., and Thevelein, J. M. (2001) *Trends Biochem. Sci.* **26**, 310 – 317
94. Cheung, A. T., Dayanandan, B., Lewis, J. T., Korbutt, G. S., Rajotte, R. V., Bryer-Ash, M., Boylan, M. O., Wolfe, M. M., and Kieffer, T. J. (2000) *Science* **290**, 1959 – 1962
94a. Hsu, S. Y., Nakabayashi, K., Nishi, S., Kumagai, J., Kudo, M., Sherwood, O. D., and Hsueh, A. J. W. (2002) *Science* **295**, 671 – 674
95. Nishikori, K., Weisbrodt, N. W., Sherwood, O. D., and Sanborn, B. M. (1983) *J. Biol. Chem.* **258**, 2468 – 2474
96. Schwabe, C., LeRoith, D., Thompson, R. P., Shiloach, J., and Roth, J. (1983) *J. Biol. Chem.* **258**, 2778 – 2781
97. Georges, D., and Schwabe, C. (1999) *FASEB J.* **13**, 1269 – 1275
98. Garrett, T. P. J., McKern, N. M., Lou, M., Frenkel, M. J., Bentley, J. D., Lovrecz, G. O., Elleman, T. C., Cosgrove, L. J., and Ward, C. W. (1998) *Nature (London)* **394**, 395 – 399
99. Andrews, P. C., Hawke, D. H., Lee, T. D., Legesse, K., Noe, B. D., and Shively, J. E. (1986) *J. Biol. Chem.* **261**, 8128 – 8133
100. Rouillé, Y., Martin, S., and Steiner, D. F. (1995) *J. Biol. Chem.* **270**, 26488 – 26496
101. Holst, J. J. (1980) *Biochem. J.* **187**, 337 – 343
102. Rouillé, Y., Kantengwa, S., Irminger, J.-C., and Halban, P. A. (1997) *J. Biol. Chem.* **272**, 32810 – 32816
102a. Lovshin, J., Estall, J., Yusta, B., Brown, T. J., and Drucker, D. J. (2001) *J. Biol. Chem.* **276**, 21489 – 21499
103. Jelinek, L. J., Lok, S., Rosenberg, G. B., Smith, R. A., Grant, F. J., Biggs, S., Bensch, P. A., Kuijper, J. L., Sheppard, P. O., Sprecher, C. A., O'Hara, P. J., Foster, D., Walker, K. M., Chen, L. H. J., McKernan, P. A., and Kindsvogel, W. (1993) *Science* **259**, 1614 – 1616
104. Patthi, S., Simerson, S., and Velicelebi, G. (1988) *J. Biol. Chem.* **263**, 19363 – 19369
105. Ollerenshaw, S., Jarvis, D., Woolcock, A., Sullivan, C., and Scheibner, T. (1989) *N. Engl. J. Med.* **320**, 1244 – 1248
106. Miyawaki, K., and 16 other authors. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14843 – 14847
107. Blundell, T. L., Pitts, J. E., Tickle, I. J., Wood, S. P., and Wu, C.-W. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4175 – 4179
108. Tonan, K., Kawata, Y., and Hamaguchi, K. (1990) *Biochemistry* **29**, 4424 – 4429
109. Bjornholm, B., Jorgensen, F. S., and Schwartz, T. W. (1993) *Biochemistry* **32**, 2954 – 2959
110. Cabrele, C., Langer, M., Bader, R., Wieland, H. A., Doods, H. N., Zerbe, O., and Beck-Sickinger, A. G. (2000) *J. Biol. Chem.* **275**, 36043 – 36048
111. Turner, A. J., ed. (1987) *Neuropeptides and Their Peptidases*, Ellis Horwood, Chichester
112. Uvnäs-Moberg, K. (1989) *Sci. Am.* **261**(Jul), 78 – 83
113. Krieger, D. T. (1983) *Science* **222**, 975 – 985
114. Drucker, D. J., and Asa, S. (1988) *J. Biol. Chem.* **263**, 13475 – 13478
115. Hoffman, J., and Porchet, M., eds. (1984) *Biosynthesis, Metabolism and Mode of Action of Invertebrate Hormones*, Springer, Berlin
116. Schaller, H. C., and Bodemüller, H. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7000 – 7004
117. Kowalski-Chauvel, A., Pradayrol, L., Vaysse, N., and Seva, C. (1996) *J. Biol. Chem.* **271**, 26356 – 26361
117a. Pannequin, J., Barnham, K. J., Hollande, F., Shulkes, A., Norton, R. S., and Baldwin, G. S. (2002) *J. Biol. Chem.* **277**, 48602 – 48609
118. Gigoux, V., Escrieut, C., Fehrentz, J.-A., Poirot, S., Maigret, B., Moroder, L., Gully, D., Martinez, J., Vaysse, N., and Fourmy, D. (1999) *J. Biol. Chem.* **274**, 20457 – 20464
119. Ding, X.-Q., Dolu, V., Hadac, E. M., Holicky, E. L., Pinon, D. I., Lybrand, T. P., and Miller, L. J. (2001) *J. Biol. Chem.* **276**, 4236 – 4244
120. Rehfeld, J. F. (1986) *J. Biol. Chem.* **261**, 5841 – 5847
121. Giragossian, C., and Mierke, D. F. (2001) *Biochemistry* **40**, 3804 – 3809
122. Anders, J., Blüggel, M., Meyer, H. E., Kühne, R., ter Laak, A. M., Kojro, E., and Fahrenholz, F. (1999) *Biochemistry* **38**, 6043 – 6055
123. Nachman, R. J., Holman, G. M., Haddon, W. F., and Ling, N. (1986) *Science* **234**, 71 – 73
124. Walz, D. A., Wider, M. D., Snow, J. W., Dass, C., and Desiderio, D. M. (1988) *J. Biol. Chem.* **263**, 14189 – 14195
125. Edmondson, S., Khan, N., Shriver, J., Zdunek, J., and Gräslund, A. (1991) *Biochemistry* **30**, 11271 – 11279
126. Feighner, S. D., Tan, C. P., McKee, K. K., Palyha, O. C., Hreniuk, D. L., Pong, S.-S., Austin, C. P., Figueroa, D., MacNeil, D., Cascieri, M. A., Nargund, R., Bakshi, R., Abramovitz, M., Stocco, R., Kargman, S., O'Neill, G., Van Der Ploeg, L. H. T., Evans, J., Patchett, A. A., Smith, R. G., and Howard, A. D. (1999) *Science* **284**, 2184 – 2188
127. Brown, M., Rivier, J., and Vale, W. (1977) *Science* **196**, 998 – 1000
128. Erne, D., and Schwyzer, R. (1987) *Biochemistry* **26**, 6316 – 6319
129. Nagalla, S. R., Gibson, B. W., Tang, D., Reeve, J. R., Jr., and Spindel, E. R. (1992) *J. Biol. Chem.* **267**, 6916 – 6922
130. Takuwa, N., Takuwa, Y., Bollag, W. E., and Rasmussen, H. (1987) *J. Biol. Chem.* **262**, 182 – 188
131. Xu, Z.-Q., Shi, T.-J., and Hökfelt, T. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14901 – 14905
132. Forte, L. R., and Currie, M. G. (1995) *FASEB J.* **9**, 643 – 650
133. Skelton, N. J., Garcia, K. C., Goeddel, D. V., Quan, C., and Burnier, J. P. (1994) *Biochemistry* **33**, 13581 – 13592
134. Bader, R., Bettio, A., Beck-Sickinger, A. G., and Zerbe, O. (2001) *J. Mol. Biol.* **305**, 307 – 329
135. Herzog, H., Hort, Y., Schneider, R., and Shine, J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 594 – 598
136. Pellieux, C., Sauthier, T., Domenighetti, A., Marsh, D. J., Palmiter, R. D., Brunner, H.-R., and Pedrazzini, T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1595 – 1600
137. Larhammar, D., Blomqvist, A. G., Yee, F., Jazin, E., Yoo, H., and Wahlestedt, C. (1992) *J. Biol. Chem.* **267**, 10935 – 10938
138. Voisin, T., Couvineau, A., Rouyer-Fessard, C., and Laburthe, M. (1991) *J. Biol. Chem.* **266**, 10762 – 10767
139. Blomqvist, A. G., Söderberg, C., Lundell, I., Milner, R. J., and Larhammar, D. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2350 – 2354
139a. Batterham, R. L., Cowley, M. A., Small, C. J., Herzog, H., Cohen, M. A., Dakin, C. L., Wren, A. M., Brynes, A. E., Low, M. J., Ghatel, M. A., Cone, R. D., and Bloom, S. R. (2002) *Nature (London)* **418**, 650 – 654
139b. Schwartz, M. W., and Morton, G. J. (2002) *Nature (London)* **418**, 595 – 597
140. Woods, S. C., Seeley, R. J., Porte, D., Jr., and Schwartz, M. W. (1998) *Science* **280**, 1378 – 1383
141. Friedman, J. M., and Halaas, J. L. (1998) *Nature (London)* **395**, 763 – 770
142. Leibel, R. L., Chung, W. K., and Chua, S. C., Jr. (1997) *J. Biol. Chem.* **272**, 31937 – 31940
143. Zhang, F., Basinski, M. B., Beals, J. M., Briggs, S. L., Churgay, L. M., Clawson, D. K., DiMarchi, R. D., Furman, T. C., Hale, J. E., Hsiung, H. M., Schoner, B. E., Smith, D. P., Zhang, X. Y., Wery, J.-P., and Schevitz, R. W. (1997) *Nature (London)* **387**, 207 – 209
144. Chicurel, M. (2000) *Nature (London)* **404**, 538 – 540
145. Cowley, M. A., Smart, J. L., Rubinstein, M., Cerdán, M. G., Diano, S., Horvath, T. L., Cone, R. D., and Low, M. J. (2001) *Nature (London)* **411**, 480 – 484
145a. Bednarek, M. A., Hreniuk, D. L., Tan, C., Palyha, O. C., MacNeil, D. J., Van der Ploeg, L. H. Y., Howard, A. D., and Feighner, S. D. (2002) *Biochemistry* **41**, 6383 – 6390
146. Flier, J. S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4242 – 4245
147. Unger, R. H., Zhou, Y.-T., and Orci, L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2327 – 2332
148. Wahlestedt, C. (1998) *Science* **281**, 1624 – 1625
149. Nicholls, D. G. (1994) *Proteins, Transmitters and Synapses*, Blackwell Scientific Publications, Oxford
150. Li, Y.-M., Marnerakis, M., Stimson, E. R., and Maggio, J. E. (1995) *J. Biol. Chem.* **270**, 1213 – 1220
151. Iversen, L. (1998) *Nature (London)* **392**, 334 – 335
152. De Felipe, C., Herrero, J. F., O'Brien, J. A., Palmer, J. A., Doyle, C. A., Smith, A. J. H., Laird, J. M. A., Belmonte, C., Cervero, F., and Hunt, S. P. (1998) *Nature (London)* **392**, 394 – 397
153. Motta, A., Temussi, P. A., Wünsch, E., and Bovermann, G. (1991) *Biochemistry* **30**, 2364 – 2371
154. Rosenfeld, M. G., Amara, S. G., and Evans, R. M. (1984) *Science* **225**, 1315 – 1320
155. Habener, J. E., Potts, J. T., Jr., and Rich, A. (1976) *J. Biol. Chem.* **251**, 3893 – 3899
156. Jin, L., Briggs, S. L., Chandrasekhar, S., Chirgadze, N. Y., Clawson, D. K., Schevitz, R. W., Smiley, D. L., Tashjian, A. H., and Zhang, F. (2000) *J. Biol. Chem.* **275**, 27238 – 27244
156a. Shimizu, M., Potts, J. T., Jr., and Gardella, T. J. (2000) *J. Biol. Chem.* **275**, 21836 – 21843
157. Neer, R. M., and 11 other authors. (2001) *N. Engl. J. Med.* **344**, 1434 – 1441
158. Greenberg, Z., Bisello, A., Mierke, D. F., Rosenblatt, M., and Chorev, M. (2000) *Biochemistry* **39**, 8142 – 8152
159. Rölz, C., Pellegrini, M., and Mierke, D. F. (1999) *Biochemistry* **38**, 6397 – 6405
160. Ruat, M., Snowman, A. M., Hester, L. D., and Snyder, S. H. (1996) *J. Biol. Chem.* **271**, 5972 – 5975
161. Mierke, D. F., Maretto, S., Schievano, E., DeLuca, D., Bisello, A., Mammi, S., Rosenblatt, M., Peggion, E., and Chorev, M. (1997) *Biochemistry* **36**, 10372 – 10383
161a. Julius, D. (1997) *Nature (London)* **386**, 442
161b. Fiori, S., Renner, C., Cramer, J., Pegoraro, S., and Moroder, L. (1999) *J. Mol. Biol.* **291**, 163 – 175
162. Dickenson, A. H. (1986) *Nature (London)* **320**, 681 – 682
163. Tenneti, L., Gibbons, S. J., and Talamo, B. R. (1998) *J. Biol. Chem.* **273**, 26799 – 26808
164. Cockayne, D. A., Hamilton, S. G., Zhu, Q.-M., Dunn, P. M., Zhong, Y., Novakovic, S., Malmberg, A. B., Cain, G., Berson, A., Kassotakis, L., Hedley, L., Lachnit, W. G., Burnstock, G., McMahon, S. B., and Ford, A. P. D. W. (2000) *Nature (London)* **407**, 1011 – 1015
165. Surprenant, A., and Evans, R. J. (1993) *Nature (London)* **362**, 211 – 212
166. Grinthal, A., and Guidotti, G. (2000) *Biochemistry* **39**, 9 – 16
167. Cook, S. P., and McCleskey, E. W. (2000) *Nature (London)* **407**, 951 – 952

References

168. Goettsch, S., and Bayer, P. (2001) *Trends Biochem. Sci.* **26**, 12
169. Newby, A. C. (1984) *Trends Biochem. Sci.* **9**, 42–44
170. Huang, N.-K., Lin, Y.-W., Huang, C.-L., Messing, R. O., and Chern, Y. (2001) *J. Biol. Chem.* **276**, 13838–13846
171. Miller, K. J., and Hoffman, B. J. (1994) *J. Biol. Chem.* **269**, 27351–27356
172. Satoh, S., Matsumura, H., Suzuki, F., and Hayaishi, O. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5980–5984
173. Porkka-Heiskanen, T., Strecker, R. E., Thakkar, M., Bjorkum, A. A., Greene, R. W., and McCarley, R. W. (1997) *Science* **276**, 1265–1268
174. LaNoue, K. F., and Martin, L. F. (1994) *FASEB J.* **8**, 72–80
175. Hoebertz, A., Meghji, S., Burnstock, G., and Arnett, T. R. (2001) *FASEB J.* **15**, 1139–1148
176. Müller-Esterl, W., Iwanaga, S., and Nakanishi, S. (1986) *Trends Biochem. Sci.* **11**, 336–339

176a. Chuang, H.-h, Prescott, E. D., Kong, H., Shields, S., Jordt, S.-E., Basbaum, A. I., Chao, M. V., and Julius, D. (2001) *Nature (London)* **411**, 957–962

177. Steranka, L. R., Farmer, S. G., and Burch, R. M. (1989) *FASEB J.* **3**, 2019–2025
178. Jarnagin, K., Bhakta, S., Zuppan, P., Yee, C., Ho, T., Phan, T., Tahilramani, R., Pease, J. H. B., Miller, A., and Freedman, R. (1996) *J. Biol. Chem.* **271**, 28277–28286
179. Yanagisawa, M., Kurihara, H., Kimura, S., Tomobe, Y., Kobayashi, M., Mitsui, Y., Yazaki, Y., Goto, K., and Masaki, T. (1988) *Nature (London)* **332**, 411–415
180. Takuwa, N., Takuwa, Y., Yanagisawa, M., Yamashita, K., and Masaki, T. (1989) *J. Biol. Chem.* **264**, 7856–7861
181. Levin, E. R. (1995) *N. Engl. J. Med.* **333**, 356–363
182. Sokolovsky, M. (1991) *Trends Biochem. Sci.* **16**, 261–264
183. Pollock, D. M., Keith, T. L., and Highsmith, R. F. (1995) *FASEB J.* **9**, 1196–1204
184. Inagami, T. (1989) *J. Biol. Chem.* **264**, 3043–3046
185. Sengenès, C., Berlan, M., De Glisezinski, I., Lafontan, M., and Galitzky, J. (2000) *FASEB J.* **14**, 1345–1351
186. Volpi, M., Naccache, P. H., Molski, T. F. P., Shefcyk, J., Huang, C.-K., Marsh, M. L., Munoz, J., Becker, E. L., and Sha'afi, R. I. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2708–2712
187. Mills, J. S., Miettinen, H. M., Barnidge, D., Vlases, M. J., Wimer-Mackin, S., Dratz, E. A., Sunner, J., and Jesaitis, A. J. (1998) *J. Biol. Chem.* **273**, 10428–10435
188. Miettinen, H. M., Gripentrog, J. M., Mason, M. M., and Jesaitis, A. J. (1999) *J. Biol. Chem.* **274**, 27934–27942
189. Tzehoval, E., Segal, S., Stabinsky, Y., Fridkin, M., Spirer, Z., and Feldman, M. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3400–3404
190. Handel, T. M., and Domaille, P. J. (1996) *Biochemistry* **35**, 6569–6584
191. Clore, G. M., and Gronenborn, A. M. (1995) *FASEB J.* **9**, 57–62
192. Sticht, H., Escher, S. E., Schweimer, K., Forssmann, W.-G., Rösch, P., and Adermann, K. (1999) *Biochemistry* **38**, 5995–6002
193. Raffioni, S., Miceli, C., Vallesi, A., Chowdhury, S. K., Chait, B. T., Luporini, P., and Bradshaw, R. A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2071–2075
194. Anderson, D. H., Weiss, M. S., and Eisenberg, D. (1997) *J. Mol. Biol.* **273**, 479–500
195. Vincent, F., Löbel, D., Brown, K., Spinelli, S., Grote, P., Breer, H., Cambillau, C., and Tegoni, M. (2001) *J. Mol. Biol.* **305**, 459–469
196. Dangott, L. J., and Garbers, D. L. (1984) *J. Biol. Chem.* **259**, 13712–13716
197. Davies, D. R., and Wlodawer, A. (1995) *FASEB J.* **9**, 50–56
198. Robinson, R. C., Radziejewski, C., Spraggon, G., Greenwald, J., Kostura, M. R., Burtnick, L. D., Stuart, D. I., Choe, S., and Jones, E. Y. (1999) *Protein Sci.* **8**, 2589–2597
199. Ye, J., Mayer, K. L., Mayer, M. R., and Stone, M. J. (2001) *Biochemistry* **40**, 7820–7831
200. Heath, J. K. (1994) *Growth Factors*, IRL Press, Oxford
201. Taniguchi, T. (1995) *Science* **268**, 251–255
202. Wells, J. A., ed. (1998) *Advances in Protein Chemistry, Cytokines*, Vol. 52, Academic Press, San Diego, California
203. Oppenheim, J. J., Feldmann, M., Durum, S. K., Hirano, T., Vilcek, J., and Nicola, N. A., eds. (2000) *Cytokine Reference*, Academic Press, San Diego
204. Harada, S., Smith, R. M., Smith, J. A., White, M. F., and Jarett, L. (1996) *J. Biol. Chem.* **271**, 30222–30226
205. Jackson, J. G., White, M. F., and Yee, D. (1998) *J. Biol. Chem.* **273**, 9994–10003
206. Rotwein, P., Pollock, K. M., Didier, D. K., and Krivi, G. G. (1986) *J. Biol. Chem.* **261**, 4828–4832
207. Dubaquié, Y., and Lowman, H. B. (1999) *Biochemistry* **38**, 6386–6396
208. Zezulak, K. M., and Green, H. (1986) *Science* **233**, 551–553
209. Frunzio, R., Chiariotti, L., Brown, A. L., Graham, D. E., Rechler, M. M., and Bruni, C. B. (1986) *J. Biol. Chem.* **261**, 17138–17149
210. Corvera, S., Whitehead, R. E., Mottola, C., and Czech, M. P. (1986) *J. Biol. Chem.* **261**, 7675–7679
211. Urdea, M. S., Merryweather, J. P., Mullenbach, G. T., Coit, D., Heberlein, U., Valenzuela, P., and Barr, P. J. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7461–7465
212. Abe, Y., Odaka, M., Inagaki, F., Lax, I., Schlessinger, J., and Kohda, D. (1998) *J. Biol. Chem.* **273**, 11150–11157
213. Marchionni, M. A. (1995) *Nature (London)* **378**, 334–335
214. Chang, H., Riese, D. J., II, Gilbert, W., Stern, D. F., and McMahan, U. J. (1997) *Nature (London)* **387**, 509–512
215. Strachan, L., Murison, J. G., Prestidge, R. L., Sleeman, M. A., Watson, J. D., and Kumble, K. D. (2001) *J. Biol. Chem.* **276**, 18265–18271
216. Kohda, D., and Imagaki, F. (1988) *J. Biochem.* **103**, 554–571
217. Yarden, Y., and Ullrich, A. (1988) *Ann. Rev. Biochem.* **57**, 443–478
218. Kumagai, T., Davis, J. G., Horie, T., O'Rourke, D. M., and Greene, M. I. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5526–5531
219. Schlessinger, J. (1988) *Trends Biochem. Sci.* **13**, 443–447
220. Morrison, P., Chung, K.-C., and Rosner, M. R. (1996) *Biochemistry* **35**, 14618–14624
221. Lin, C. R., Chen, W. S., Kruiger, W., Stolarsky, L. S., Weber, W., Evans, R. M., Verma, I. M., Gill, G. N., and Rosenfeld, M. G. (1984) *Science* **224**, 843–848
222. Hunter, T., Ling, N., and Cooper, J. A. (1984) *Nature (London)* **311**, 480–483
223. Kuppuswamy, D., Dalton, M., and Pike, L. J. (1993) *J. Biol. Chem.* **268**, 19134–19142
224. Betsholtz, C., Johnsson, A., Heldin, C.-H., Westermark, B., Lind, P., Urdea, M. S., Eddy, R., Shows, T. B., Philpott, K., Mellor, A. L., Knott, T. J., and Scott, J. (1986) *Nature (London)* **320**, 695–699
225. Assoian, R. K., Grotendorst, G. R., Miller, D. M., and Sporn, M. B. (1984) *Nature (London)* **309**, 804–806
226. Williams, L. T. (1989) *Science* **243**, 1564–1570
227. Massagué, J. (1985) *Trends Biochem. Sci.* **10**, 237–240
228. Kohda, D., Shimada, I., Miyake, T., Fuwa, T., and Inagaki, F. (1989) *Biochemistry* **28**, 953–958
229. Piek, E., Heldin, C.-H., and ten Dijke, P. (1999) *FASEB J.* **13**, 2105–2124
230. Elder, J. T., Fisher, G. J., Lindquist, P. B., Bennett, G. L., Pittelkow, M. R., Coffey, R. J., Jr., Ellingsworth, L., Derynck, R., and Voorhees, J. J. (1989) *Science* **243**, 811–814
231. Thomas, K. A., and Gimenez-Gallego, G. (1986) *Trends Biochem. Sci.* **11**, 81–84
232. Riboni, L., Viani, P., Bassi, R., Giussani, P., and Tettamanti, G. (2001) *J. Biol. Chem.* **276**, 12797–12804
233. Blaber, M., DiSalvo, J., and Thomas, K. A. (1996) *Biochemistry* **35**, 2086–2094
234. Thomas, K. A. (1988) *Trends Biochem. Sci.* **13**, 327–328
235. Schreiber, A. B., Winkler, M. E., and Derynck, R. (1986) *Science* **232**, 1250–1253
236. Folkman, J., and Klagsbrun, M. (1987) *Science* **235**, 442–447
237. Thomas, K. A. (1996) *J. Biol. Chem.* **271**, 603–606
238. Radisavljevic, Z., Avraham, H., and Avraham, S. (2000) *J. Biol. Chem.* **275**, 20770–20774
239. Antonetti, D. A., Barber, A. J., Hollinger, L. A., Wolpert, E. B., and Gardner, T. W. (1999) *J. Biol. Chem.* **274**, 23463–23467
240. Leonidas, D. D., Shapiro, R., Allen, S. C., Subbarao, G. V., Veluraja, K., and Acharya, K. R. (1999) *J. Mol. Biol.* **285**, 1209–1233
241. Burke, D., Wilkes, D., Blundell, T. L., and Malcolm, S. (1998) *Trends Biochem. Sci.* **23**, 59–62
242. Richard, C., Liuzzo, J. P., and Moscatelli, D. (1995) *J. Biol. Chem.* **270**, 24188–24196
243. Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B., and Blundell, T. L. (2000) *Nature (London)* **407**, 1029–1034
244. Ibrahimi, O. A., Eliseenkova, A. V., Plotnikov, A. N., Yu, K., Ornitz, D. M., and Mohammadi, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7182–7187
245. Webster, M. K., and Donoghue, D. J. (1996) *EMBO J.* **15**, 520–527
246. Faham, S., Hileman, R. E., Fromm, J. R., Linhardt, R. J., and Rees, D. C. (1996) *Science* **271**, 1116–1120
247. Taipale, J., and Keski-Oja, J. (1997) *FASEB J.* **11**, 51–59
248. Levi-Montalcini, R. (1987) *Science* **237**, 1154–1162
249. Bradshaw, R. A., Blundell, T. L., Lapatto, R., McDonald, N. Q., and Murray-Rust, J. (1993) *Trends Biochem. Sci.* **18**, 48–52
250. Thoenen, H. (1995) *Science* **270**, 593–598
251. Wiesmann, C., Ultsch, M. H., Bass, S. H., and de Vos, A. M. (1999) *Nature (London)* **401**, 184–188

251a. Tartaglia, N., Du, J., Tyler, W. J., Neale, E., Pozzo-Miller, L., and Lu, B. (2001) *J. Biol. Chem.* **276**, 37585–37593

252. Micera, A., Vigneti, E., Pickholtz, D., Reich, R., Pappo, O., Bonini, S., Maquart, F. X., Aloe, L., and Levi-Schaffer, F. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6162–6167
253. Wooten, M. W., Seibenhener, M. L., Mamidipudi, V., Diaz-Meco, M. T., Barker, P. A., and Moscat, J. (2001) *J. Biol. Chem.* **276**, 7709–7712
254. Brockes, J. P., Lemke, G. E., and Balzer, D. R., Jr. (1980) *J. Biol. Chem.* **255**, 8374–8377
255. Frazier, W. A., Ohléndorf, C. E., Boyd, L. F., Aloe, L., Johnson, E. M., Ferrendelli, J. A., and Bradshaw, R. A. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2448–2452
256. Simpson, E. (1984) *Trends Biochem. Sci.* **9**, 527–530

References

257. Wang, E. A., Rosen, V., Cordes, P., Hewick, R. M., Kriz, M. J., Luxenberg, D. P., Sibley, B. S., and Wozney, J. M. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9484 – 9488
258. Seyedin, S. M., Thompson, A. Y., Bentz, H., Rosen, D. M., McPherson, J. M., Conti, A., Siegel, N. R., Galluppi, G. R., and Piez, K. A. (1986) *J. Biol. Chem.* **261**, 5693 – 5695
259. Whetton, A. D., and Dexter, T. M. (1986) *Trends Biochem. Sci.* **11**, 207 – 211
260. Lai, P.-H., Everett, R., Wang, F.-F., Arakawa, T., and Goldwasser, E. (1986) *J. Biol. Chem.* **261**, 3116 – 3121
261. Goldberg, M. A., Dunning, S. P., and Bunn, H. F. (1988) *Science* **242**, 1412 – 1415
262. D'Andrea, A. D. (1994) *N. Engl. J. Med.* **330**, 839 – 845
263. Mufson, R. A. (1997) *FASEB J.* **11**, 37 – 44
264. Livnah, O., Stura, E. A., Johnson, D. L., Middleton, S. A., Mulcahy, L. S., Wrighton, N. C., Dower, W. J., Jolliffe, L. K., and Wilson, I. A. (1996) *Science* **273**, 464 – 471
265. Wells, J. A. (1996) *Science* **273**, 449 – 450
266. Kawasaki, E. S., Ladner, M. B., Wang, A. M., Van Arsdell, J., Warren, M. K., Coyne, M. Y., Schweickart, V. L., Lee, M.-T., Wilson, K. J., Boosman, A., Stanley, E. R., Ralph, P., and Mark, D. F. (1985) *Science* **230**, 291 – 296
267. Taylor, E. W., Fear, A. L., Bohm, A., Kim, S.-H., and Koths, K. (1994) *J. Biol. Chem.* **269**, 31171 – 31177
268. Dinarello, C. A., and Wolff, S. M. (1993) *N. Engl. J. Med.* **328**, 106 – 112
269. Cohen, F. E., Kosen, P. A., Kuntz, I. D., Epstein, L. B., Ciardelli, T. L., and Smith, K. A. (1986) *Science* **234**, 349 – 355
270. Clark-Lewis, I., Aebersold, R., Ziltener, H., Schrader, J. W., Hood, L. E., and Kent, S. B. H. (1986) *Science* **231**, 134 – 139
271. Gurney, M. E., Heinrich, S. P., Lee, M. R., and Yin, H.-S. (1986) *Science* **234**, 566 – 574
272. Sun, Y.-J., Chou, C.-C., Chen, W.-S., Wu, R.-T., Meng, M., and Hsiao, C.-D. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5412 – 5417
273. Audhya, T., Kroon, D., Heavner, G., Viamontes, G., and Goldstein, G. (1986) *Science* **231**, 997 – 999
274. Audhya, T., Schlesinger, D. H., and Goldstein, G. (1981) *Biochemistry* **20**, 6195 – 6200
275. Vancurova, I., Miskolci, V., and Davidson, D. (2001) *J. Biol. Chem.* **276**, 19746 – 19752
276. Naismith, J. H., and Sprang, S. R. (1998) *Trends Biochem. Sci.* **23**, 74 – 79
277. Kanety, H., Feinstein, R., Papa, M. Z., Hemi, R., and Karasik, A. (1995) *J. Biol. Chem.* **270**, 23780 – 23784
278. Loret, C., Sensenbremmer, M., and Labourdette, G. (1989) *J. Biol. Chem.* **264**, 8319 – 8327
279. Moncada, S., and Higgs, A. (1993) *N. Engl. J. Med.* **329**, 2002 – 2010
280. Mayer, B., and Hemmens, B. (1997) *Trends Biochem. Sci.* **22**, 477 – 481
281. Durner, J., Gow, A. J., Stamler, J. S., and Glazebrook, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14206 – 14207
282. Colasanti, M., Lauro, G. M., and Venturini, G. (1995) *Nature (London)* **374**, 505
283. Lawson, D. M., Stevenson, C. E. M., Andrew, C. R., and Eady, R. R. (2000) *EMBO J.* **19**, 5661 – 5671
284. Champion, H. C., Bivalacqua, T. J., Hyman, A. L., Ignarro, L. J., Hellstrom, W. J. G., and Kadowitz, P. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11648 – 11652
285. Clementi, E., Brown, G. C., Foxwell, N., and Moncada, S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1559 – 1562
285a. Corbin, J. D., and Francis, S. H. (1999) *J. Biol. Chem.* **274**, 13729 – 13732
286. Kim, N. N., Cox, J. D., Baggio, R. F., Emig, F. A., Mistry, S. K., Harper, S. L., Speicher, D. W., Morris, S. M., Jr., Ash, D. E., Traish, A., and Christianson, D. W. (2001) *Biochemistry* **40**, 2678 – 2688
286a. Lue, T. F. (2000) *N. Engl. J. Med.* **342**, 1802 – 1805
287. Xue, L., Farrugia, G., Miller, S. M., Ferris, C. D., Snyder, S. H., and Szurszewski, J. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1851 – 1855
288. Hillier, B. J., Christopherson, K. S., Prehoda, K. E., Bredt, D. S., and Lim, W. A. (1999) *Science* **284**, 812 – 815
289. Denlinger, L. C., Fisette, P. L., Garis, K. A., Kwon, G., Vazquez-Torres, A., Simon, A. D., Nguyen, B., Proctor, R. A., Bertics, P. J., and Corbett, J. A. (1996) *J. Biol. Chem.* **271**, 337 – 342
290. Hall, J. P., Merithew, E., and Davis, R. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14022 – 14024
291. Trimmer, B. A., Aprille, J. R., Dudzinski, D. M., Lagace, C. J., Lewis, S. M., Michel, T., Qazi, S., and Zayas, R. M. (2001) *Science* **292**, 2486 – 2488
291a Yew, W. S., Kolatkar, P. R., Kuhn, P., and Khoo, H. E. (1999) *J Struct. Biol.* **128**, 216 – 218
292. Inoue, K., Akaike, T., Miyamoto, Y., Okamoto, T., Sawa, T., Otagiri, M., Suzuki, S., Yoshimura, T., and Maeda, H. (1999) *J. Biol. Chem.* **274**, 27069 – 27075
293. Andersen, J. F., Champagne, D. E., Weichsel, A., Ribeiro, J. M. C., Balfour, C. A., Dress, V., and Montfort, W. R. (1997) *Biochemistry* **36**, 4423 – 4428
294. Spencer, N. Y., Zeng, H., Patel, R. P., and Hogg, N. (2000) *J. Biol. Chem.* **275**, 36562 – 36567
295. Mayer, B., Pfeiffer, S., Schrammel, A., Koesling, D., Schmidt, K., and Brunner, F. (1998) *J. Biol. Chem.* **273**, 3264 – 3270
296. Gross, S. S., and Lane, P. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9967 – 9969
297. Herold, S., Exner, M., and Nauser, T. (2001) *Biochemistry* **40**, 3385 – 3395
298. Nagatomo, S., Nagai, M., Tsuneshige, A., Yonetani, T., and Kitagawa, T. (1999) *Biochemistry* **38**, 9659 – 9666
299. Pawloski, J. R., Hess, D. T., and Stamler, J. S. (2001) *Nature (London)* **409**, 622 – 626
300. Gladwin, M. T., Shelhamer, J. H., Schechter, A. N., Pease-Fye, M. E., Waclawiw, M. A., Panza, J. A., Ognibene, F. P., and Cannon, R. O., III. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11482 – 11487
300a. Ruschitzka, F. T., Wenger, R. H., Stallmach, T., Quaschning, T., Wit, C., Wagner, K., Labugger, R., Kelm, M., Noll, G., Rülicke, T., Shaw, S., Lindberg, R. L. P., Rodenwaldt, B., Lutz, H., Bauer, C., Lüscher, T. F., and Gassmann, M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11609 – 11613
301. Brunori, M. (2001) *Trends Biochem. Sci.* **26**, 209 – 210
302. Flögel, U., Merx, M. W., Gödecke, A., Decking, U. K. M., and Schrader, J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 735 – 740
303. Minning, D. M., Gow, A. J., Bonaventura, J., Braun, R., Dewhirst, M., Goldberg, D. E., and Stamler, J. S. (1999) *Nature (London)* **401**, 497 – 502
304. Lin, R., and Farmer, P. J. (2000) *J. Am. Chem. Soc.* **122**, 2393 – 2394
305. Witting, P. K., Douglas, D. J., and Mauk, A. G. (2001) *J. Biol. Chem.* **276**, 3991 – 3998
306. Friebe, A., Schultz, G., and Koesling, D. (1996) *EMBO J.* **15**, 6863 – 6868
307. Morita, T., Mitsialis, S. A., Koike, H., Liu, Y., and Kourembanas, S. (1997) *J. Biol. Chem.* **272**, 32804 – 32809
308. Arakawa, T., and Timasheff, S. N. (1982) *Biochemistry* **21**, 6545 – 6552
309. Uchida, T., Ishikawa, H., Ishimori, K., Morishima, I., Nakajima, H., Aono, S., Mizutani, Y., and Kitagawa, T. (2000) *Biochemistry* **39**, 12747 – 12752
310. Nakajima, H., Honma, Y., Tawara, T., Kato, T., Park, S.–Y., Miyatake, H., Shiro, Y., and Aono, S. (2001) *J. Biol. Chem.* **276**, 7055 – 7061
311. Ho, Y. S., Swenson, L., Dwrewenda, U., Serre, L., Wei, Y., Dauter, Z., Hattori, M., Adachi, T., Aoki, J., Arai, H., Inoue, K., and Derewenda, Z. S. (1997) *Nature (London)* **385**, 89 – 93
312. Serhan, C. N., Haeggström, J. Z., and Leslie, C. C. (1996) *FASEB J.* **10**, 1147 – 1158
313. Ruvolo, P. P., Gao, F., Blalock, W. L., Deng, X., and May, W. S. (2001) *J. Biol. Chem.* **276**, 11754 – 11758
314. Lehtonen, J. Y. A., Horiuchi, M., Daviet, L., Akishita, M., and Dzau, V. J. (1999) *J. Biol. Chem.* **274**, 16901 – 16906
315. Cravatt, B. F., Prospero–Garcia, O., Siuzdak, G., Gilula, N. B., Henriksen, S. J., Boger, D. L., and Lerner, R. A. (1995) *Science* **268**, 1506 – 1509
316. Wilcox, B. J., Ritenour–Rogers, K. J., Asser, A. S., Baumgart, L. E., Baumgart, M. A., Boger, D. L., DeBlassio, J. L., deLong, M. A., Glufke, U., Henz, M. E., King, L., III, Merkler, K. A., Patterson, J. E., Robleski, J. J., Vederas, J. C., and Merkler, D. J. (1999) *Biochemistry* **38**, 3235 – 3245
317. Boger, D. L., Patterson, J. E., Guan, X., Cravatt, B. F., Lerner, R. A., and Gilula, N. B. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4810 – 4815
318. Patricelli, M. P., Lovato, M. A., and Cravatt, B. F. (1999) *Biochemistry* **38**, 9804 – 9812
319. Maccarrone, M., Bari, M., Lorenzon, T., Bisogno, T., Di Marzo, V., and Finazzi-Agró, A. (2000) *J. Biol. Chem.* **275**, 13484 – 13492
320. Zygmunt, P. M., Petersson, J., Andersson, D. A., Chaunag, H.-H., Sorgard, M., Di Marzo, V., Julius, D., and Högestätt, E. D. (1999) *Nature (London)* **400**, 452 – 456
321. Patricelli, M. P., Lashuel, H. A., Giang, D. K., Kelly, J. W., and Cravatt, B. F. (1998) *Biochemistry* **37**, 15177 – 15787
322. Patterson, J. E., Ollmann, I. R., Cravatt, B. F., Boger, D. L., Wong, C.-H., and Lerner, R. A. (1996) *J. Am. Chem. Soc.* **118**, 5938 – 5945
323. Strauss, E. (1999) *Science* **284**, 1302 – 1304
324. Dong, Y.-H., Wang, L.-H., Xu, J.-L., Zhang, H.-B., Zhang, X.-F., and Zhang, L.-H. (2001) *Nature (London)* **411**, 813 – 817
325. Anderegg, R. J., Betz, R., Carr, S. A., Crabb, J. W., and Duntze, W. (1987) *J. Biol. Chem.* **263**, 18236 – 18240
326. Kurjan, J. (1992) *Ann. Rev. Biochem.* **61**, 1097 – 1129
327. Song, J., and Dohlman, H. G. (1996) *Biochemistry* **35**, 14806 – 14817
327a. Mentesana, P. E., and Konopka, J. B. (2001) *Biochemistry* **40**, 9685 – 9694
328. Sakagami, Y., Yoshida, M., Isogai, A., and Suzuki, A. (1981) *Science* **212**, 1525 – 1527
329. Gottschalk, W. K., and Sonneborn, D. R. (1985) *J. Biol. Chem.* **260**, 6592 – 6599
330. Starr, R. C., Marner, F. J., and Jaenicke, L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 641 – 645
331. Grimmelikhuijzen, C. J. P., and Graff, D. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9817 – 9821
332. Grimmelikhuijzen, C. J. P., Rinehart, K. L., Jacob, E., Graff, D., Reinscheid, R. K., Nothacker, H.-P., and Staley, A. L. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5410 – 5414
333. Bellis, S. L., Kass-Simon, G., and Rhoads, D. E. (1992) *Biochemistry* **31**, 9838 – 9843
334. Murata, M., Miyagawa-Kohshima, K., Nakanishi, K., and Naya, Y. (1986) *Science* **234**, 585 – 587

References

335. Scheller, R. H., Kaldany, R.-R., Kreiner, T., Mahon, A. C., Nambu, J. R., Schaefer, M., and Taussig, R. (1984) *Science* **225**, 1300–1308

336. Chiu, D. T., and Zare, R. N. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3338–3340

337. Lloyd, P. E., Schacher, S., Kupfermann, I., and Weiss, K. R. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9794–9798

338. Berry, R. W. (1981) *Biochemistry* **20**, 6200–6205

339. Kaldany, R.-R. J., Campanelli, J. T., Makk, G., Evans, C. J., and Scheller, R. H. (1986) *J. Biol. Chem.* **261**, 5751–5757

340. Newcomb, R., Fisher, J. M., and Scheller, R. H. (1988) *J. Biol. Chem.* **263**, 12514–12521

341. Garden, R. W., Shippy, S. A., Li, L., Moroz, T. P., and Sweedler, J. V. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3972–3977

342. Joose, J. (1984) in *Biosynthesis, Metabolism and Mode of Action of Invertebrate Hormones* (Hoffmann, J., and Porchet, M., eds), pp. 17–35, Springer, Berlin

343. Nelson, L. S., Rosoff, M. L., and Li, C. (1998) *Science* **281**, 1686–1690

344. Duve, H., Johnsen, A. H., Sewell, J. C., Scott, A. G., Orchard, I., Rehfeld, J. F., and Thorpe, A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2326–2330

345. Lange, A. B., Orchard, I., Wang, Z., and Nachman, R. J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9250–9253

346. Ramarao, C. S., Burks, D. J., and Garbers, D. L. (1990) *Biochemistry* **29**, 3383–3388

347. Yoshino, K.-i, Takao, T., Suhara, M., Kitai, T., Hori, H., Nomura, K., Yamaguchi, M., Shimonishi, Y., and Suzuki, N. (1991) *Biochemistry* **30**, 6203–6209

348. Downer, R. G. H., and Laufer, H., eds. (1983) *Endocrinology of Insects*, Liss, New York

349. O'Shea, M., Adams, M. E., and Bishop, C. A. (1982) *Fed. Proc.* **41**, 2940–2947

350. Stangier, J., Hilbich, C., Beyreuther, K., and Keller, R. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 575–579

351. Fernlund, P., and Josefsson, L. (1972) *Science* **177**, 173–175

352. Rao, K. R., Riehm, J. P., Zahnow, C. A., Kleinholz, L. H., Tarr, G. E., Johnson, L., Norton, S., Landau, M., Semmes, O. J., Sattelberg, R. M., Jorenby, W. H., and Hintz, M. F. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5319–5322

353. Rao, K. R., Mohrherr, C. J., Riehm, J. P., Zahnow, C. A., Norton, S., Johnson, L., and Tarr, G. E. (1987) *J. Biol. Chem.* **262**, 2672–2675

354. Sanburg, L. L., Kramer, K. J., Kezdy, F. J., Law, J. H., and Oberlander, H. (1975) *Nature (London)* **253**, 266–267

355. Touhara, K., Lerro, K. A., Bonning, B. C., Hammock, B. D., and Prestwich, G. D. (1993) *Biochemistry* **32**, 2068–2075

356. Cayre, M., Strambi, C., Charpin, P., Augier, R., and Strambi, A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8238–8242

357. Wainwright, G., Webster, S. G., Wilkinson, M. C., Chung, J. S., and Rees, H. H. (1996) *J. Biol. Chem.* **271**, 12749–12754

358. Agui, N., Bollenbacher, W. E., Granger, N. A., and Gilbert, L. I. (1980) *Nature (London)* **285**, 669–670

359. Coudron, T. A., Law, J. H., and Koeppe, J. K. (1981) *Trends Biochem. Sci.* **6**, 248–252

359a. Takeuchi, H., Chen, J.-H., O'Reilly, D. R., Turner, P. C., and Rees, H. H. (2001) *J. Biol. Chem.* **276**, 26819–26828

360. Beach, R. (1979) *Science* **205**, 829–831

361. Pongs, O. (1985) in *Interaction of Steroid Hormone Receptors with DNA* (Sluyser, M., ed), pp. 226–240, Ellis Horwood, Chichester

362. Zitnan, D., Kingan, T. G., Hermesman, J. L., and Adams, M. E. (1996) *Science* **271**, 88–91

363. Oudejans, R. C. H. M., Vroemen, S. F., Jansen, R. F. R., and Van der Horst, D. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8654–8659

364. Kondo, H., Ino, M., Suzuki, A., Ishizaki, H., and Iwami, M. (1996) *J. Mol. Biol.* **259**, 926–937

365. Roelofs, W. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 44–49

366. Takahashi, N., ed. (1986) *Chemistry of Plant Hormones*, CRC Press, Boca Raton, Florida

367. Crozier, A., and Hillman, J. R., eds. (1984) *The Biosynthesis and Metabolism of Plant Hormones*, Cambridge Univ. Press, London

368. Moller, S. G., and Chua, N.-H. (1999) *J. Mol. Biol.* **293**, 219–234

369. Buchanan, B. B., Gruissem, W., and Jones, R. L. (2000) *Biochemistry and Molecular Biology of Plants*, American Society of Plant Physiologists, Rockville, Maryland

370. Busby, S., and Ebright, R. H. (1999) *J. Mol. Biol.* **293**, 199–213

371. Steinmann, T., Geldner, N., Grebe, M., Mangold, S., Jackson, C. L., Paris, S., Gälweiler, L., Palme, K., and Jürgens, G. (1999) *Science* **286**, 316–318

371a. Friml, J., Wisniewska, J., Benková, E., Mendgen, K., and Palme, K. (2002) *Nature (London)* **415**, 806–809

372. El-Antably, H. M. M., and Larsen, P. (1974) *Nature (London)* **250**, 76–77

373. Shaw, G. (1994) in *Cytokinins. Chemistry, Activity and Function* (Mok, D. W. S., and Mok, M. C., eds), CRC Press, Boca Raton, Florida

374. Martin, R. C., Mok, M. C., Habben, J. E., and Mok, D. W. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5922–5926

375. Koshimiza, K., and Iwamura, H. (1986) *Chemistry of Plant Hormones*, CRC Press, Boca Raton, Florida (pp. 153–199)

376. Hall, M. A. (1986) *Hormones, Receptors and Cellular Interactions in Plants*, Cambridge Univ. Press, London (pp. 69–89)

377. Menke, F. L. H., Champion, A., Kijne, J. W., and Memelink, J. (1999) *EMBO J.* **18**, 4455–4463

378. Niggeweg, R., Thurow, C., Kegler, C., and Gatz, C. (2000) *J. Biol. Chem.* **275**, 19897–19905

379. Whitfield, J. (2001) *Nature (London)* **410**, 736–737

380. Kessler, A., and Baldwin, I. T. (2001) *Science* **291**, 2141–2144

381. Farmer, E. E. (2001) *Nature (London)* **411**, 854–856

382. Pearce, G., Moura, D. S., Stratmann, J., and Ryan, C. A. (2001) *Nature (London)* **411**, 817–820

383. Phillips, D. A., Joseph, C. M., Yang, G.-P., Martinez-Romero, E., Sanborn, J. R., and Volpin, H. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12275–12280

384. Wain, R. L. (1977) *Chem. Soc. Rev.* **6**, 261–276

385. Nickell, L. G. (1978) *Chem. Eng. News* **Oct. 9**, 18–34

386. Thomas-Reetz, A. C., and De Camilli, P. (1994) *FASEB J.* **8**, 209–216

387. Bean, A. J., Zhang, X., and Hökfelt, T. (1994) *FASEB J.* **8**, 630–638

388. Berridge, M. J., Bootman, M. D., and Lipp, P. (1998) *Nature (London)* **395**, 645–648

389. Cancela, J. M., Gerasimenko, O. V., Gerasimenko, J. V., Tepikin, A. V., and Petersen, O. H. (2000) *EMBO J.* **19**, 2549–2557

390. Putney, J. W., Jr. (2001) *Nature (London)* **410**, 648–649

391. Ma, H.-T., Patterson, R. L., van Rossum, D. B., Birnbaumer, L., Mikoshiba, K., and Gill, D. L. (2000) *Science* **287**, 1647–1651

392. Machado, J. D., Segura, F., Brioso, M. A., and Borges, R. (2000) *J. Biol. Chem.* **275**, 20274–20279

393. Siegel, G. J., Albers, R. W., Agronoff, B. W., and Katzman, R. (1981) *Basic Neurochemistry*, 3rd ed., Little, Brown, Boston, Massachusetts

394. entire issue. (1979) *Sci. Am.* **241**(Sept)

395. Eccles, J. C. (1973) *The Understanding of the Brain*, McGraw-Hill, New York

396. McCool, B. A., Plonk, S. G., Martin, P. R., and Singleton, C. K. (1993) *J. Biol. Chem.* **268**, 1397–1404

396a. Johansen, K. M., and Johansen, J. (1995) *J. Neurobiol.* **27**, 227–239

396b. Nicholls, J. G., and Van Essen, D. (1974) *Sci. Am.* **230**(Jan), 38–48

397. Nicholls, J. G., and Van Essen, D. (1974) *Sci. Am.* **230**(Jan), 38–48

398. McKay, R. D. G., Hockfield, S., Johansen, J., Thompson, I., and Frederiksen, K. (1983) *Science* **222**, 788–799

399. Barnes, D. M. (1986) *Science* **233**, 155–156

400. Koch, C., and Laurent, G. (1999) *Science* **284**, 96–98

400a. Kandel, E. R., and Squire, L. R. (2000) *Science* **290**, 1113–1120

401. Brand, E. D., and Westfall, T. C. (1970) in *Medicinal Chemistry*, 3rd ed. (Burger, A., ed), pp. 1190 –1234, Wiley (Interscience), New York (Part II)

402. Porter, K. R., and Bonneville, M. A. (1973) *Fine Structure of Cells and Tissues*, 4th ed., Lea & Febiger, Philadelphia, Pennsylvania

403. Dowling, J. E. (1965) *Science* **147**, 57–59

404. Llinás, R. R. (1975) *Sci. Am.* **232**(Jan), 56–71

405. Noback, C. R., and Demarest, R. J. (1972) *The Nervous System; Introduction and Review*, McGraw–Hill, New York

406. Spector, R., and Johanson, C. E. (1989) *Sci. Am.* **261**(Nov), 68–74

407. Boado, R., Li, J. Y., Nagaya, M., Zhang, C., and Pardridge, W. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12079–12084

408. Glickstein, M., and Gibson, A. R. (1976) *Sci. Am.* **235**(Nov), 90–98

409. Crick, F., and Jones, E. (1993) *Nature (London)* **361**, 109–110

410. Barinaga, M. (1995) *Science* **268**, 1696–1698

411. Vranesic, I., Iijima, T., Ichikawa, M., Matsumoto, G., and Knöpfel, T. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 13014–13017

412. McCormick, D. A. (1995) *Nature (London)* **374**, 412–413

413. Aidley, D. J. (1971) *The Physiology of Excitable Cells*, Cambridge Univ. Press, London and New York

414. Adelman, W. J., Jr., ed. (1971) *Biophysics and Physiology of Excitable Membranes*, Van Nostrand–Reinhold, New York

415. Nystrom, R. A. (1973) *Membrane Physiology*, Prentice-Hall, Englewood Cliffs, New Jersey

416. Black, S. (1973) *Adv. Enzymol.* **38**, 193–234

417. Hodgkin, A. L. (1964) *The Conduction of the Nervous Impulse*, Thomas, Springfield, Illinois

418. Hirano, A., and Dembitzer, H. M. A. (1967) *J. Cell Biol.* **34**, 555

419. Weimbs, T., and Stoffel, W. (1994) *Biochemistry* **33**, 10408–10415

420. Griffiths, I., Klugmann, M., Anderson, T., Yool, D., Thomson, C., Schwab, M. H., Schneider, A., Zimmermann, F., McCulloch, M., Nadon, N., and Nave, K.-A. (1998) *Science* **280**, 1610–1613

421. Su, Y., Brooks, D. G., Li, L., Lepercq, J., Trofatter, J. A., Ravetch, J. V., and Lebo, R. V. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10856–10860

422. Barinaga, M. (1996) *Science* **273**, 1657–1658

422a. Gallego, R. G., Blanco, J. L. J., Thijssen-van Zuylen, C. W. E. M., Gotfredsen, C. H., Voshol, H., Duus, J. O., Schachner, M., and Vliegenthart, J. F. G. (2001) *J. Biol. Chem.* **276**, 30834–30844

References

423. Cao, M. Y., Dupriez, V. J., Rider, M. H., Deggouj, N., Gersdorff, M. C. H., Rousseau, G. G., and Tomasi, J.-P. (1996) *FASEB J.* **10**, 1635 – 1640
424. Shapiro, L., Doyle, J. P., Hensley, P., Colman, D. R., and Hendrickson, W. (1996) *Neuron* **17**, 435 – 440
425. Lemke, G. (1996) *Nature (London)* **383**, 395 – 396
426. Beniac, D. R., Luckevich, M. D., Czarnota, G. J., Tompkins, T. A., Ridsdale, R. A., Ottensmeyer, F. P., Moscarello, M. A., and Harauz, G. (1997) *J. Biol. Chem.* **272**, 4261 – 4268
427. Pritzker, L. B., Joshi, S., Harauz, G., and Moscarello, M. A. (2000) *Biochemistry* **39**, 5382 – 5388
428. Li, Y., Li, H., Martin, R., and Mariuzza, R. A. (2000) *J. Mol. Biol.* **304**, 177 – 188
428a. Sanders, C. R., Ismail-Beigi, F., and McEnery, M. W. (2001) *Biochemistry* **40**, 9453 – 9459
428b. Notterpek, L., Roux, K. J., Amici, S. A., Yazdanpour, A., Rahner, C., and Fletcher, B. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14404 – 14409
429. Fuhrman, F. A. (1967) *Sci. Am.* **217**(Aug), 61 – 71
429a. Huey, R. B., and Moody, W. J. (2002) *Science* **297**, 1289 – 1290
430. Spiro, T. G. (1974) *Acc. Chem. Res.* **7**, 339 – 344
431. Ritchie, J. M., and Bogart, R. B. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 211 – 215
432. Ritchie, J. M. (1973) *Prog. Biophys. and Mol. Biol.* **26**, 149 – 187
432a. Kaplan, J. H. (2002) *Ann. Rev. Biochem.* **71**, 511 – 535
433. Catterall, W. A. (1986) *Ann. Rev. Biochem.* **55**, 953 – 985
434. Catterall, W. A. (1988) *Science* **242**, 50 – 61
435. Pallaghy, P. K., Duggan, B. M., Pennington, M. W., and Norton, R. S. (1993) *J. Mol. Biol.* **234**, 405 – 420
436. Shimizu, Y., Chou, H.-N., Bando, H., Van Duyne, G., and Clardy, J. C. (1986) *J. Am. Chem. Soc.* **108**, 514 – 515
437. Noda, M., Shimizu, S., Tanabe, T., Takai, T., Kayano, T., Ikeda, T., Takahashi, H., Nakayama, H., Kanaoka, Y., Minamino, N., Kangawa, K., Matsuo, H., Raftery, M. A., Hirose, T., Inayama, S., Hayashida, H., Miyata, T., and Numa, S. (1984) *Nature (London)* **312**, 121 – 127
438. Sato, C., Ueno, Y., Asai, K., Takahashi, K., Sato, M., Engel, A., and Fujiyoshi, Y. (2001) *Nature (London)* **409**, 1047 – 1051
439. Catterall, W. A. (2001) *Nature (London)* **409**, 988 – 991
440. Messner, D. J., and Catterall, W. A. (1986) *J. Biol. Chem.* **261**, 211 – 215
441. Noda, M., Ikeda, T., Kayano, T., Suzuki, H., Takeshima, H., Kurasaki, M., Takahashi, H., and Numa, S. (1986) *Nature (London)* **320**, 188 – 192
441a. Hilber, K., Sandtner, W., Kudlacek, O., Glaaser, I. W., Weisz, E., Kyle, J. W., French, R. J., Fozzard, H. A., Dudley, S. C., and Todt, H. (2001) *J. Biol. Chem.* **276**, 27831 – 27839
441b. Tan, H. L., Bink-Boelkens, M. T. E., Bezzina, C. R., Viswanathan, P. C., Beaufort-Krol, G. C. M., van Tintelen, P. J., van den Berg, M. P., Wilde, A. A. M., and Balser, J. R. (2001) *Nature (London)* **409**, 1043 – 1047
442. Lopreato, G. F., Lu, Y., Southwell, A., Atkinson, N. S., Hillis, D. M., Wilcox, T. P., and Zakon, H. H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7588 – 7592
442a. Ren, D., Navarro, B., Xu, H., Yue, L., Shi, Q., and Clapham, D. E. (2001) *Science* **294**, 2372 – 2375
442b. Catterall, W. A. (2001) *Science* **294**, 2306 – 2308
443. Reichardt, L. F. (1984) *Trends Biochem. Sci.* **9**, 173 – 176
444. Akerman, K. E. O., and Nicholls, D. G. (1983) *Trends Biochem. Sci.* **8**, 63 – 64
444a. Kinoshita, M., Nukada, T., Asano, T., Mori, Y., Akaike, A., Satoh, M., and Kaneko, S. (2001) *J. Biol. Chem.* **276**, 28731 – 28738
444b. Saegusa, H., Kurihara, T., Zong, S., Kazuno, A.-a, Matsuda, Y., Nonaka, T., Han, W., Toriyama, H., and Tanabe, T. (2001) *EMBO J.* **20**, 2349 – 2356
445. Ellis, S. B., Williams, M. E., Ways, N. R., Brenner, R., Sharp, A. H., Leung, A. T., Campbell, K. P., McKenna, E., Koch, W. J., Hui, A., Schwartz, A., and Harpold, M. M. (1988) *Science* **241**, 1661 – 1664
445a. Serysheva, I. I., Ludtke, S. J., Baker, M. R., Chiu, W., and Hamilton, S. L. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10370 – 10375
445b. Yang, S.-N., Yu, J., Mayr, G. W., Hofmann, F., Larsson, O., and Berggren, P.-O. (2001) *FASEB J.* **15**, 1753 – 1763
446. Sharp, A. H., and Campbell, K. P. (1989) *J. Biol. Chem.* **264**, 2816 – 2825
446a. Zhao, M., Li, P., Li, X., Zhang, L., Winkfein, R. J., and Chen, S. R. W. (1999) *J. Biol. Chem.* **274**, 25971 – 25974
447. Putney, J. W., Jr. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14669 – 14671
448. Zhao, X.-S., Shin, D. M., Liu, L. H., Shull, G. E., and Muallem, S. (2001) *EMBO J.* **20**, 2680 – 2689
449. Churchill, G. C., and Galione, A. (2001) *EMBO J.* **20**, 2666 – 2671
450. Blount, P., Sukharev, S. I., Moe, P. C., Schroeder, M. J., Guy, H. R., and Kung, C. (1996) *EMBO J.* **15**, 4798 – 4805
451. Oakley, A. J., Martinac, B., and Wilce, M. C. J. (1999) *Protein Sci.* **8**, 1915 – 1921
452. Chang, G., Spencer, R. H., Lee, A. T., Barclay, M. T., and Rees, D. C. (1998) *Science* **282**, 2220 – 2226
452a. Abbott, A. (2002) *Nature (London)* **418**, 268 – 269
453. Li, H. L., Sui, H. X., Ghanshani, S., Lee, S., Walian, P. J., Wu, C. L., Chandy, K. G., and Jap, B. K. (1998) *J. Mol. Biol.* **282**, 211 – 216
454. Caprini, M., Ferroni, S., Planells-Cases, R., Rueda, J., Rapisarda, C., Ferrer-Montiel, A., and Montal, M. (2001) *J. Biol. Chem.* **276**, 21070 – 21076
455. Katz, A. M. (1993) *N. Engl. J. Med.* **328**, 1244 – 1251
455a. Yellen, G. (2002) *Nature (London)* **419**, 35 – 42
456. Choe, S., Kreusch, A., and Pfaffinger, P. J. (1999) *Trends Biochem. Sci.* **24**, 345 – 349
457. Jing, J., Chikvashvili, D., Dinger-Lahat, D., Thornhill, W. B., Reuveny, E., and Lotan, I. (1999) *EMBO J.* **18**, 1245 – 1256
458. Aldrich, R. W. (2001) *Nature (London)* **411**, 643 – 644
458a. Gulbis, J. M., Zhou, M., Mann, S., and MacKinnon, R. (2000) *Science* 289, 123 – 127
459. Zhou, M., Morals-Cabral, J. H., Mann, S., and MacKinnon, R. (2001) *Nature (London)* **411**, 657 – 661
459a. Strang, C., Cushman, S. J., DeRubeis, D., Peterson, D., and Pfaffinger, P. J. (2001) *J. Biol. Chem.* **276**, 28493 – 28502
459b. Hanlon, M. R., and Wallace, B. A. (2002) *Biochemistry* **41**, 2886 – 2894
460. Roux, B., Bernéche, S., and Im, W. (2000) *Biochemistry* **39**, 13295 – 13306
461. Roux, B., and MacKinnon, R. (1999) *Science* **285**, 100 – 102
462. Zagrovic, B., and Aldrich, R. (1999) *Science* **285**, 59 – 61
463. Perozo, E., Cortes, D. M., and Cuello, L. G. (1999) *Science* **285**, 73 – 78
464. Guidoni, L., Torre, V., and Carloni, P. (1999) *Biochemistry* **38**, 8599 – 8604
464a. Zhou, Y., Morais-Cabral, J. H., Kaufman, A., and MacKinnon, R. (2001) *Nature (London)* **414**, 43 – 48
464b. Morais-Cabral, J. H., Zhou, Y., and MacKinnon, R. (2001) *Nature (London)* **414**, 37 – 42
464c. Bernèche, S., and Roux, B. (2001) *Nature (London)* **414**, 73 – 76
464d. Rivas, J. C. M., Schwalbe, H., and Lippard, S. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9478 – 9483
465. Brenner, R., Peréz, G. J., Bonev, A. D., Eckman, D. M., Kosek, J. C., Wiler, S. W., Patterson, A. J., Nelson, M. T., and Aldrich, R. W. (2000) *Nature (London)* **407**, 870 – 871
466. Tian, L., Duncan, R. R., Hammond, M. S. L., Coghill, L. S., Wen, H., Rusinova, R., Clark, A. G., Levitan, I. B., and Shipston, M. J. (2001) *J. Biol. Chem.* **276**, 7717 – 7720
466a. Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T., and MacKinnon, R. (2002) *Nature (London)* **417**, 515 – 522
466b. Schumacher, M. A., Rivard, A. F., Bächinger, H. P., and Adelman, J. P. (2001) *Nature (London)* **410**, 1120 – 1124
467. Wang, H.-S., Pan, Z., Shi, W., Brown, B. S., Wymore, R. S., Cohen, I. S., Dixon, J. E., and McKinnon, D. (1998) *Science* **282**, 1890 – 1893
468. Abraham, M. R., Jahangir, A., Alekseev, A. E., and Terzic, A. (1999) *FASEB J.* **13**, 1901 – 1910
469. Maingret, F., Patel, A. J., Lesage, F., Lazdunski, M., and Honoré, E. (1999) *J. Biol. Chem.* **274**, 26691 – 26696
470. Lu, Z., and MacKinnon, R. (1995) *Biochemistry* **34**, 13133 – 13138
471. Qu, Z., Yang, Z., Cui, N., Zhu, G., Liu, C., Xu, H., Chanchevalap, S., Shen, W., Wu, J., Li, Y., and Jiang, C. (2000) *J. Biol. Chem.* **275**, 31573 – 31580
472. Repunte, V. P., Nakamura, H., Fujita, A., Horio, Y., Findlay, I., Pott, L., and Kurachi, Y. (1999) *EMBO J.* **18**, 3317 – 3324
473. Tucker, S. J., Gribble, F. M., Proks, P., Trapp, S., Ryder, T. J., Haug, T., Reimman, F., and Ashcroft, F. M. (1998) *EMBO J.* **17**, 3290 – 3296
474. Carrasco, A. J., Dzeja, P. P., Alekseev, A. E., Pucar, D., Zingman, L. V., Abraham, M. R., Hodgson, D., Bienengraeber, M., Puceat, M., Janssen, E., Wieringa, B., and Terzic, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7623 – 7628
474a. Loussouarn, G., Pike, L. J., Ashcroft, F. M., Makhina, E. N., and Nichols, C. G. (2001) *J. Biol. Chem.* **276**, 29098 – 29103
474b. Moreau, C., Jacquet, H., Prost, A.-L., D'hahan, N., and Vivaudou, M. (2000) *EMBO J.* **19**, 6644 – 6651
475. Rogalski, S. L., and Chavkin, C. (2001) *J. Biol. Chem.* **276**, 14855 – 14860
476. Lemtiri-Chlieh, F., MacRobbie, E. A. C., and Brearley, C. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8687 – 8692
477. Petalcorin, M. I. R., Oka, T., Koga, M., Ogura, K.-i, Wada, Y., Ohshima, Y., and Futai, M. (1999) *J. Mol. Biol.* **294**, 347 – 355
477a. Harrop, S. J., DeMaere, M. Z., Fairlie, W. D., Reztsova, T., Valenzuela, S. M., Mazzanti, M., Tonini, R., Qiu, M. R., Jankova, L., Warton, K., Bauskin, A. R., Wu, W. M., Pankhurst, S., Campbell, T. J., Breit, S. N., and Curmi, P. M. G. (2001) *J. Biol. Chem.* **276**, 44993 – 45000
477b. Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T., and MacKinnon, R. (2002) *Nature (London)* **415**, 287– 294
477c. Jentsch, T. J. (2002) *Nature (London)* **415**, 276 – 277
478. George, A. L., Jr. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7843 – 7845
479. Cooper, E. C., and Jan, L. Y. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4759 – 4766
480. Ho, M. W. Y., Kaetzel, M. A., Armstrong, D. L., and Shears, S. B. (2001) *J. Biol. Chem.* **276**, 18673 – 18680
480a. Zou, X., and Hwang, T.-C. (2001) *Biochemistry* **40**, 5579 – 5586
481. Eldefrawi, A. T., and Eldefrawi, M. E. (1987) *FASEB J.* **1**, 262 – 271

References

481a. Smith, M. M., Warren, V. A., Thomas, B. S., Brochu, R. M., Ertel, E. A., Rohrer, S., Schaeffer, J., Schmatz, D., Petuch, B. R., Tang, Y. S., Meinke, P. T., Kaczorowski, G. J., and Cohen, C. J. (2000) *Biochemistry* **39**, 5543 – 5554

482. Llinas, R. R. (2000) *The Squid Synapse: A Model for Chemical Transmission*, Oxford Univ. Press, London

483. Saimi, Y., Hinrichsen, R. D., Forte, M., and Kung, C. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5112 – 5116

484. Woolley, G. A., and Wallace, B. A. (1993) *Biochemistry* **32**, 9819 – 9825

484a. Cafiso, D. S. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 141 – 165

485. Cleveland, MvB., Slatin, S., Finkelstein, A., and Levinthal, C. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3706 – 3710

486. Nicoll, R. A. (1988) *Science* **241**, 545 – 551

487. Olivera, B. M., Rivier, J., Clark, C., Ramilo, C. A., Corpuz, G. P., Abogadie, F. C., Mena, E. E., Woodward, S. R., Hillyard, D. R., and Cruz, L. J. (1990) *Science* **249**, 257 – 263

488. McIntosh, J. M., Santos, A. D., and Olivera, B. M. (1999) *Ann. Rev. Biochem.* **68**, 59 – 88

489. Fainzilber, M., Nakamura, T., Lodder, J. C., Zlotkin, E., Kits, K. S., and Burlingame, A. L. (1998) *Biochemistry* **37**, 1470 – 1477

490. Lirazan, M. B., Hooper, D., Corpuz, G. P., Ramilo, C. A., Bandyopadhyay, P., Cruz, L. J., and Olivera, B. M. (2000) *Biochemistry* **39**, 1583 – 1588

491. Rigby, A. C., Lucas-Meunier, E., Kalume, D. E., Czerwiec, E., Hambe, B., Dahlqvist, I., Fossier, P., Baux, G., Roepstorff, P., Baleja, J. D., Furie, B. C., Furie, B., and Stenflo, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5758 – 5763

492. Shon, K.-J., Grilley, M. M., Marsh, M., Yoshikami, D., Hall, A. R., Kurz, B., Gray, W. R., Imperial, J. S., Hillyard, D. R., and Olivera, B. M. (1995) *Biochemistry* **34**, 4913 – 4918

493. McIntosh, J. M., Olivera, B. M., Cruz, L. J., and Gray, W. R. (1984) *J. Biol. Chem.* **259**, 14343 – 14346

494. Jover, E., Bablito, J., and Couraud, F. (1984) *Biochemistry* **23**, 1147 – 1152

494a. Almassy, R. J., Fontecilla-Camps, J. C., Suddath, F. L., Bugg, C. E. (1983) *J. Mol. Biol.* **170**, 497 – 527

495. Schweitz, H., Bidard, J.-N., Frelin, C., Pauron, D., Vijverberg, H. P. M., Mahasneh, D. M., and Lazdunski, M. (1985) *Biochemistry* **24**, 3554 – 3561

496. Shimizu, Y., Hsu, C.-P., and Genenah, A. (1981) *J. Am. Chem. Soc.* **103**, 605 – 609

497. Kao, C. Y., and Levinson, S. R. (1986) *Ann. N.Y. Acad. Sci.* **479**, entire volume

498. Morabito, M. A., and Moczydlowski, E. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2478 – 2482

499. Ovchinnikov, Y. A., and Grishin, E. V. (1982) *Trends Biochem. Sci.* **7**, 26 – 28

500. Jablonsky, M. J., Jackson, P. L., and Krishna, N. R. (2001) *Biochemistry* **40**, 8273 – 8282

501. Tugarinov, V., Kustanovich, I., Zilberberg, N., Gurevitz, M., and Anglister, J. (1997) *Biochemistry* **36**, 2414 – 2424

502. Gurevitz, M., Gordon, D., Ben-Natan, S., Turkov, M., and Froy, O. (2001) *FASEB J.* **15**, 1201 – 1205

503. Fontecilla-Camps, J. C., Almassy, R. J., Ealick, S. E., Suddath, F. L., Watt, D. D., Feldmann, R. J., and Bugg, C. E. (1981) *Trends Biochem. Sci.* **6**, 291 – 296

504. Fogh, R. H., Kem, W. R., and Norton, R. S. (1990) *J. Biol. Chem.* **265**, 13016 – 13028

505. Loret, E. P., Menendez Soto del Valle, R., Mansuelle, P., Sampieri, F., and Rochat, H. (1994) *J. Biol. Chem.* **269**, 16785 – 16788

506. Ellis, K. C., Tenenholz, T. C., Jerng, H., Hayhurst, M., Dudlak, C. S., Gilly, W. F., Blaustein, M. P., and Weber, D. J. (2001) *Biochemistry* **40**, 5942 – 5953

507. Delepierre, M., Prochnicka-Chalufour, A., Boisbouvier, J., and Possani, L. D. (1999) *Biochemistry* **38**, 16756 – 16765

508. Baden, D. G. (1989) *FASEB J.* **3**, 1807 – 1817

509. Duffton, M. J., and Hider, R. C. (1980) *Trends Biochem. Sci.* **5**, 53 – 56

510. Achari, A., Radvanyi, F. R., Scott, D., Bon, C., and Sigler, P. B. (1985) *J. Biol. Chem.* **260**, 9385 – 9387

510a. Délot, E., and Bon, C. (1993) *Biochemistry* **32**, 10708 – 10713

511. Juillerat, M. A., Schwendimann, B., Hauert, J., Fulpius, B. W., and Bargetzi, J. P. (1982) *J. Biol. Chem.* **257**, 2901 – 2907

512. Drenth, J., Low, B. W., Richardson, J. S., and Wright, C. S. (1980) *J. Biol. Chem.* **255**, 2652 – 2655

513. Hatanaka, H., Oka, M., Kohda, D., Tate, S-i, Suda, A., Tamiya, N., and Inagaki, F. (1994) *J. Mol. Biol.* **240**, 155 – 166

514. Corfield, P. W. R., Lee, T.-J., and Low, B. W. (1989) *J. Biol. Chem.* **264**, 9239 – 9242

515. Basus, V. J., Billeter, M., Love, R. A., Stroud, R. M., and Kuntz, I. D. (1988) *Biochemistry* **27**, 2763 – 2771

516. Ruoppolo, M., Moutiez, M., Mazzeo, M. F., Pucci, P., Ménez, A., Marino, G., and Quéméneur, E. (1998) *Biochemistry* **37**, 16060 – 16068

517. Bilwes, A., Rees, B., Moras, D., Ménez, R., and Ménez, A. (1994) *J. Mol. Biol.* **239**, 122 – 136

518. Sun, Y.-J., Wu, W.-g, Chiang, C.-M., Hsin, A.-Y., and Hsiao, C.-D. (1997) *Biochemistry* **36**, 2403 – 2413

519. Rehm, H., and Betz, H. (1982) *J. Biol. Chem.* **257**, 10015 – 10022

520. Montecucco, C., and Rossetto, O. (2000) *Trends Biochem. Sci.* **25**, 266 – 270

521. Montecucco, C. (1986) *Trends Biochem. Sci.* **11**, 314 – 317

522. Montecucco, C., and Schiavo, G. (1993) *Trends Biochem. Sci.* **18**, 324 – 327

523. Lacy, D. B., and Stevens, R. C. (1999) *J. Mol. Biol.* **291**, 1091 – 1104

523a. Eswaramoorthy, S., Kumaran, D., and Swaminathan, S. (2002) *Biochemistry* **41**, 9795 – 9802

524. Eisel, U., Jarausch, W., Goretzki, K., Henschen, A., Engels, J., Weller, U., Hudel, M., Habermann, E., and Niemann, H. (1986) *EMBO J.* **5**, 2495 – 2502

525. Fu, F.-N., Lomneth, R. B., Cai, S., and Singh, B. R. (1998) *Biochemistry* **37**, 5267 – 5278

526. Keller, J. E., and Neale, E. A. (2001) *J. Biol. Chem.* **276**, 13476 – 13482

527. Turton, K., Chaddock, J. A., and Acharya, K. R. (2002) *Trends Biochem. Sci.* **27**, 552 – 558

528. Blasi, J., Chapman, E. R., Link, E., Binz, T., Yamasaki, S., De Camilli, P., Südhof, T. C., Niemann, H., and Jahn, R. (1993) *Nature (London)* **365**, 160 – 163

528a. Li, Y., Foran, P., Lawrence, G., Mohammed, N., Chan-Kwo-Chion, C.-K.-N., Lisk, G., Aoki, R., and Dolly, O. (2001) *J. Biol. Chem.* **276**, 31394 – 31401

529. Matteoli, M., Verderio, C., Rossetto, O., Iezzi, N., Coco, S., Schiavo, G., and Montecucco, C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13310 – 13315

530. Nicholls, D. G., Rugolo, M., Scott, I. G., and Meldolesi, J. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7924 – 7928

531. Khvotchev, M., and Südhof, T. C. (2000) *EMBO J.* **19**, 3250 – 3262

532. Tufariello, J. J., Meckler, H., Pushpananda, K., and Senartne, A. (1984) *J. Am. Chem. Soc.* **106**, 7979 – 7980

533. Guddat, L. W., Martin, J. A., Shan, L., Edmundson, A. B., and Gray, W. R. (1996) *Biochemistry* **35**, 11329 – 11335

534. Hu, S.-H., Loughnan, M., Miller, R., Weeks, C. M., Blessing, R. H., Alewood, P. F., Lewis, R. J., and Martin, J. L. (1998) *Biochemistry* **37**, 11425 – 11433

535. Klein, L. L., McWhorter, W. W., Jr., Ko, S. S., Pfaff, K.-P., and Kishi, Y. (1982) *J. Am. Chem. Soc.* **104**, 7362 – 7364

536. Shimizu, Y. (1983) *Nature (London)* **302**, 212

536a. Raichle, M. E., and Gusnard, D. A. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10237 – 10239

536b. Job, C., and Eberwine, J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13037 – 13042

537. Mohr, E., Prakash, N., Vieluf, K., Fuhrmann, C., Buck, F., and Richter, D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7072 – 7079

538. Terada, S., Nakata, T., Peterson, A. C., and Hirokawa, N. (1996) *Science* **273**, 784 – 788

539. Caplan, R., Cheung, S. C.-Y., and Omenn, G. S. (1974) *J. Neurochemistry* **22**, 517 – 520

540. Sutcliffe, J. G., and Milner, R. J. (1984) *Trends Biochem. Sci.* **9**, 95 – 99

541. Geschwind, D. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10676 – 10678

542. Parra, M., Gascard, P., Walensky, L. D., Gimm, J. A., Blackshaw, S., Chan, N., Takakuwa, Y., Berger, T., Lee, G., Chasis, J. A., Snyder, S. H., Mohandas, N., and Conboy, J. G. (2000) *J. Biol. Chem.* **275**, 3247 – 3255

543. Kim, S.-Y., Grant, P., Lee, J.-H., Pant, H. C., and Steinert, P. M. (1999) *J. Biol. Chem.* **274**, 30715 – 30721

544. Hikita, T., Tadano-Aritomi, K., Iida-Tanaka, N., Anand, J. K., Ishizuka, I., and Hakomori, S.-i. (2001) *J. Biol. Chem.* **276**, 23084 – 23091

545. Sutcliffe, J. G., Milner, R. J., Gottesfeld, J. M., and Reynolds, W. (1984) *Science* **225**, 1308 – 1315

546. Owens, G. P., Chaudhari, N., and Hahn, W. E. (1985) *Science* **229**, 1263 – 1265

547. Lone, Y.-C., Simon, M.-P., Kahn, A., and Marie, J. (1986) *J. Biol. Chem.* **261**, 1499 – 1502

548. Eccles, J. (1965) *Sci. Am.* **212**(Jan), 56 – 66

549. Cowan, W. M., Südhof, T. C., and Stevens, C. F., eds. (2001) *Synapses*, The Johns Hopkins Univ. Press, Baltimore

550. Kalil, R. E. (1989) *Sci. Am.* **261**(Dec), 76 – 85

551. Hall, Z. W. (1972) *Ann. Rev. Biochem.* **41**, 925 – 952

552. Krnjevic, K. (1971) *Methods Neurochem.* **1**, 129 – 172

553. Katz, B. (1971) *Science* **173**, 123 – 126

554. Südhof, T. C., and Scheller, R. H. (2000) in *Synapses* (Cowan, W. M., Südhof, T. C., and Stevens, C. F., eds), pp. 177 – 215, Johns Hopkins Univ. Press, Baltimore, Maryland

555. Südhof, T. C. (1995) *Nature (London)* **375**, 645 – 653

556. Bajjalieh, S. M., and Scheller, R. H. (1995) *J. Biol. Chem.* **270**, 1971 – 1974

557. Bennett, M. K., and Scheller, R. H. (1994) *Ann. Rev. Biochem.* **63**, 63 – 100

558. Klingauf, J., Kavalali, E. T., and Tsien, R. W. (1998) *Nature (London)* **394**, 581 – 585

559. Wang, L.-Y., and Kaczmarek, L. K. (1998) *Nature (London)* **394**, 384 – 388

560. Zucker, R. S., and Landó, L. (1986) *Science* **231**, 574 – 579

561. Greengard, P., Valtorta, F., Czernik, A. J., and Benfenati, F. (1993) *Science* **259**, 780 – 785

562. Jahn, R., and Südhof, T. C. (1999) *Ann. Rev. Biochem.* **68**, 863 – 911

563. Sutton, R. B., Fasshauer, D., Jahn, R., and Brunger, A. T. (1998) *Nature (London)* **395**, 347 – 353

564. Kimelberg, H. K., and Norenberg, M. D. (1989) *Sci. Am.* **260**(Apr), 66 – 76

References

565. Takahashi, T., Hori, T., Kajikawa, Y., and Tsujimoto, T. (2000) *Science* **289**, 460–463

566. Erickson, J. D., Varoqui, H., Schäfer, M. K.-H., Modi, W., Diebler, M.-F., Weihe, E., Rand, J., Eiden, L. E., Bonner, T. I., and Usdin, T. B. (1994) *J. Biol. Chem.* **269**, 21929–21932

567. Hartwell, L. H., and Kastan, M. B. (1994) *Science* **266**, 1821–1828

568. Whitley, P., Grahn, E., Kutay, U., Rapoport, T. A., and von Heijne, G. (1996) *J. Biol. Chem.* **271**, 7583–7586

568a. Schoch, S., Deák, F., Königstorfer, A., Mozhayeva, M., Sara, Y., Südhof, T. C., and Kavalali, E. T. (2001) *Science* **294**, 1117–1122

568b. Misura, K. M. S., Bock, J. B., Gonzalez, L. C., Jr., Scheller, R. H., and Weis, W. I. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9184–9189

569. Weis, W. I., and Scheller, R. H. (1998) *Nature (London)* **395**, 328–329

570. Tsujimoto, S., and Bean, A. J. (2000) *J. Biol. Chem.* **275**, 2938–2942

571. Vogel, K., Cabaniols, J.-P., and Roche, P. A. (2000) *J. Biol. Chem.* **275**, 2959–2965

572. Betz, A., Okamoto, M., Benseler, F., and Brose, N. (1997) *J. Biol. Chem.* **272**, 2520–2526

572a. Augustin, I., Rosenmund, C., Südhof, T. C., and Brose, N. (1999) *Nature (London)* **400**, 457–460

573. Puffer, E. B., Lomneth, R. B., Sarkar, H. K., and Singh, B. R. (2001) *Biochemistry* **40**, 9374–9378

574. May, A. P., Whiteheart, S. W., and Weis, W. I. (2001) *J. Biol. Chem.* **276**, 21991–21994

574a. Kim, C. S., Kweon, D.-H., and Shin, Y.-K. (2002) *Biochemistry* **41**, 10928–10933

575. Dulubova, I., Sugita, S., Hill, S., Hosaka, M., Fernandez, I., Südhof, T. C., and Rizo, J. (1999) *EMBO J.* **18**, 4372–4382

576. Ubach, J., Lao, Y., Fernandez, I., Arac, D., Südhof, T. C., and Rizo, J. (2001) *Biochemistry* **40**, 5854–5860

577. Ubach, J., Zhang, X., Shao, X., Südhof, T. C., and Rizo, J. (1998) *EMBO J.* **17**, 3921–3930

578. Fukuda, M., Kojima, T., and Mikoshiba, K. (1996) *J. Biol. Chem.* **271**, 8430–8434

578a. Fernández-Chacón, R., Königstorfer, A., Gerber, S. H., Garcia, J., Matos, M. F., Stevens, C. F., Brose, N., Rizo, J., Rosenmund, C., and Südhof, T. C. (2001) *Nature (London)* **410**, 41–49

578b. Coppola, T., Magnin-Lüthi, S., Perret-Menoud, V., Gattesco, S., Schiavo, G., and Regazzi, R. (2001) *J. Biol. Chem.* **276**, 32756–32762

579. Ushkaryov, Y. A., Petrenko, A. G., Geppert, M., and Südhof, T. C. (1992) *Science* **257**, 50–56

580. Gundersen, C. B., Mastrogiacomo, A., Faull, K., and Umbach, J. A. (1994) *J. Biol. Chem.* **269**, 19197–19199

581. Chamberlain, L. H., and Burgoyne, R. D. (1996) *J. Biol. Chem.* **271**, 7320–7323

582. Vitale, N., Caumont, A.-S., Chasserot-Golaz, S., Du, G., Wu, S., Sciorra, V. A., Morris, A. J., Frohman, M. A., and Bader, M.-F. (2001) *EMBO J.* **20**, 2424–2434

583. Ryan, T. A., Smith, S. J., and Reuter, H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5567–5571

584. Schmidt, A., Wolde, M., Thiele, C., Fest, W., Kratzin, H., Podtelejnikov, A. V., Witke, W., Huttner, W. B., and Söling, H.-D. (1999) *Nature (London)* **401**, 133–141

585. Hayashi, T., Yamasaki, S., Nauenburg, S., Binz, T., and Niemann, H. (1995) *EMBO J.* **14**, 2317–2325

586. Südhof, T. C., Czernik, A. J., Kao, H.-T., Takei, K., Johnston, P. A., Horiuchi, A., Kanazir, S. D., Wagner, M. A., Perin, M. S., De Camilli, P., and Greengard, P. (1989) *Science* **245**, 1474–1480

587. Hosaka, M., and Südhof, T. C. (1999) *J. Biol. Chem.* **274**, 16747–16753

588. Johnston, P. A., and Südhof, T. C. (1990) *J. Biol. Chem.* **265**, 8869–8873

589. Hilton, J. M., Plomann, M., Ritter, B., Modregger, J., Freeman, H. N., Falck, J. R., Krishna, U. M., and Tobin, A. B. (2001) *J. Biol. Chem.* **276**, 16341–16347

590. Huang, E. P. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13386–13387

591. Palmiter, R. D., Cole, T. B., Quaife, C. J., and Findley, S. D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14934–14939

592. Hemmings, H. C., Jr., Nairn, A. C., McGuinness, T. L., Huganir, R. L., and Greengard, P. (1989) *FASEB J.* **3**, 1583–1592

593. Kennedy, M. B. (2000) *Science* **290**, 750–754

594. Sheng, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7058–7061

595. Malenka, R. C., and Nicoll, R. A. (1998) *Nature (London)* **396**, 414–415

596. Trautmann, A., and Vivier, E. (2001) *Science* **292**, 1667–1668

597. Lin, W., Burgess, R. W., Dominguez, B., Pfaff, S. L., Sanes, J. R., and Lee, K.-F. (2001) *Nature (London)* **410**, 1057–1064

598. Khan, A. A., Bose, C., Yam, L. S., Soloski, M. J., and Rupp, F. (2001) *Science* **292**, 1681–1686

599. Qi, S. Y., Groves, J. T., and Chakraborty, A. K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6548–6553

600. Hall, Z. W. (1995) *Science* **269**, 362–363

601. Travis, J. (1994) *Science* **266**, 970–972

602. Temburni, M. K., and Jacob, M. H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3631–3632

603. Lino, M., Goto, K., Kakegawa, W., Okado, H., Sudo, M., Ishiuchi, S., Miwa, A., Yakaysau, Y., Saito, I., Tsuzuki, K., and Ozawa, S. (2001) *Science* **292**, 926–927

603a. Kast, B. (2001) *Nature (London)* **412**, 674–676

604. Ullian, E. M., Sapperstein, S. K., Christopherson, K. S., and Barres, B. A. (2001) *Science* **291**, 657–661

604a. Song, H., Stevens, C. F., and Gage, F. H. (2002) *Nature (London)* **417**, 39–44

604b. Svendsen, C. N. (2002) *Nature (London)* **417**, 29–32

605. Guadaño-Ferraz, A., Obregón, M. J., St. Germain, D. L., and Bernal, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10391–10396

606. Veenstra, R. D., and DeHaan, R. L. (1986) *Science* **233**, 972–974

607. Revel, J. P., Yancey, S. B., and Nicholson, B. J. (1986) *Trends Biochem. Sci.* **11**, 375–377

608. Zimmer, D. B., Green, C. R., Evans, W. H., and Gilula, N. B. (1987) *J. Biol. Chem.* **262**, 7751–7763

609. Perkins, G., Goodenough, D., and Sosinsky, G. (1997) *Biophys. J.* **72**, 533–544

610. Sosinsky, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9210–9214

611. Goodenough, D. A., Goliger, J. A., and Paul, D. L. (1996) *Ann. Rev. Biochem.* **65**, 475–502

611a. Unger, V. M., Kumar, N. M., Gilula, N. B., and Yeager, M. (1999) *Science* **283**, 1176–1180

612. Bergoffen, J., Scherer, S. S., Wang, S., Scott, M. O., Bone, L. J., Paul, D. L., Chen, K., Lensch, M. W., Chance, P. F., and Fischbeck, K. H. (1993) *Science* **262**, 2039–2042

613. George, C. H., Kendall, J. M., Campbell, A. K., and Evans, W. H. (1998) *J. Biol. Chem.* **273**, 29822–29829

613a. del Castillo, I., Villamar, M., Moreno-Pelayo, M. A., del Castillo, F. J., Alvarez, A., Telleria, D., Menendez, I., and Moreno, F. (2002) *N. Engl. J. Med.* **346**, 243–249

614. Wong, V., and Goodenough, D. A. (1999) *Science* **285**, 62

615. Simon, D. B., Lu, Y., Choate, K. A., Velazquez, H., Al-Sabban, E., Praga, M., Casari, G., Bettinelli, A., Colussi, G., Rodriguez-Soriano, J., McCredie, D., Milford, D., Sanjad, S., and Lifton, R. P. (1999) *Science* **285**, 103–106

616. Osborne, N. N. (1977) *Nature (London)* **270**, 622–623

617. Chan-Palay, V., Engel, A. G., Wu, J.-Y., and Palay, S. L. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7027–7030

618. Geha, R. S., and Rosen, F. S. (1994) *N. Engl. J. Med.* **330**, 1008–1009

619. Hermsen, B., Stetzer, E., Thees, R., Heiermann, R., Schrattenholz, A., Ebbinghaus, U., Kretschmer, A., Methfessel, C., Reinhardt, S., and Maelicke, A. (1998) *J. Biol. Chem.* **273**, 18394–18404

620. Fambrough, D. M., and Hartzell, H. C. (1972) *Science* **176**, 189–191

621. Changeux, J.-P. (1993) *Sci. Am.* **269**(Nov), 58–62

622. Brisson, A., and Unwin, P. N. T. (1985) *Nature (London)* **315**, 474–477

623. Sakmann, B., Methfessel, C., Mishina, M., Takahashi, T., Takai, T., Kurasaki, M., Fukuda, K., and Numa, S. (1985) *Nature (London)* **318**, 538–543

624. Miyazawa, A., Fujiyoshi, Y., Stowell, M., and Unwin, N. (1999) *J. Mol. Biol.* **288**, 765–786

625. Arias, H. R. (1998) *Biochim. Biophys. Acta.* **1376**, 173–220

626. Grutter, T., and Changeux, J.-P. (2001) *Trends Biochem. Sci.* **26**, 459–463

626a. Bezakova, G., Rabben, I., Sefland, I., Fumagalli, G., and Lømo, T. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9924–9929

627. Brejc, K., van Dijk, W. J., Klaassen, R. V., Schuurmans, M., van der Oost, J., Smit, A. B., and Sixma, T. K. (2001) *Nature (London)* **411**, 269–276

628. Barrantes, F. J., Antollini, S. S., Blanton, M. P., and Prieto, M. (2000) *J. Biol. Chem.* **275**, 37333–37339

629. Osaka, H., Malany, S., Molles, B. E., Sine, S. M., and Taylor, P. (2000) *J. Biol. Chem.* **275**, 5478–5484

630. Salpeter, M. M. (1999) *Science* **286**, 424–425

631. Boulter, J., Evans, K., Goldman, D., Martin, G., Treco, D., Heinemann, S., and Patrick, J. (1986) *Nature (London)* **319**, 368–374

632. Vicente-Agullo, F., Rovira, J. C., Sala, S., Sala, F., Rodriguez-Ferrer, C., Campos-Caro, A., Criado, M., and Ballesta, J. J. (2001) *Biochemistry* **40**, 8300–8306

633. Osaka, H., Sugiyama, N., and Taylor, P. (1998) *J. Biol. Chem.* **273**, 12758–12765

634. Labarca, C., Schwarz, J., Deshpande, P., Schwarz, S., Nowak, M. W., Fonck, C., Nashmi, R., Kofuji, P., Dang, H., Shi, W., Fidan, M., Khakh, B. S., Chen, Z., Bowers, B. J., Boulter, J., Wehner, J. M., and Lester, H. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2786–2791

635. Jeanclos, E. M., Lin, L., Treuil, M. W., Rao, J., DeCoster, M. A., and Anand, R. (2001) *J. Biol. Chem.* **276**, 28281–28290

636. Tierney, M. L., and Unwin, N. (2000) *J. Mol. Biol.* **303**, 185–196

637. Avizonis, D. Z., Farr-Jones, S., Kosen, P. A., and Basus, V. J. (1996) *J. Am. Chem. Soc.* **118**, 13031–13039

638. Grutter, T., Goeldner, M., and Kotzyba-Hibert, F. (1999) *Biochemistry* **38**, 7476–7484

639. Dougherty, D. A., and Lester, H. A. (2001) *Nature (London)* **411**, 252–255

640. Unwin, N. (1995) *Nature (London)* **373**, 37–43

641. Swope, S. L., Moss, S. J., Blackstone, C. D., and Huganir, R. L. (1992) *FASEB J.* **6**, 2514–2523

642. Hass, R., Marshall, T. L., and Rosenberry, T. L. (1988) *Biochemistry* **27**, 6453–6457

643. Arnon, R., Silman, I., and Tarrab-Hazdai, R. (1999) *Protein Sci.* **8**, 2553–2561

644. Perrier, A. L., Cousin, X., Boschetti, N., Haas, R., Chatel, J.-M., Bon, S., Roberts, W. L., Pickett, S. R., Massoulié, J., Rosenberry, T. L., and Krejci, E. (2000) *J. Biol. Chem.* **275**, 34260–34265

645. Simon, S., Krejci, E., and Massoulié, J. (1998) *EMBO J.* **17**, 6178–6187

646. Kaplan, D., Ordentlich, A., Barak, D., Ariel, N., Kronman, C., Velan, B., and Shafferman, A. (2001) *Biochemistry* **40**, 7433 – 7445
647. Blusztajn, J. K., and Wurtman, R. J. (1983) *Science* **221**, 614 – 620
648. Kubo, T., Fukuda, K., Mikami, A., Maeda, A., Takahashi, H., Mishina, M., Haga, T., Haga, K., Ichiyama, A., Kangawa, K., Kojima, M., Matsuo, H., Hirose, T., and Numa, S. (1986) *Nature (London)* **323**, 411 – 416
649. Gainetdinov, R. R., and Caron, M. G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12222 – 12223
650. Nathanson, N. M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6245 – 6247
651. Gomeza, J., Shannon, H., Kostenis, E., Felder, C., Zhang, L., Brodkin, J., Grinberg, A., Sheng, H., and Wess, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1692 – 1697
652. Kinney, H. C., Filiano, J. J., Sleeper, L. A., Mandell, F., Valdes-Dapena, M., and White, W. F. (1995) *Science* **269**, 1446 – 1449
653. Yamada, M., Miyakawa, T., Duttaroy, A., Yamanaka, A., Moriguchi, T., Makita, R., Ogawa, M., Chou, C. J., Xia, B., Crawley, J. N., Felder, C. C., Deng, C.-X., and Wess, J. (2001) *Nature (London)* **410**, 207 – 212
654. Hill, J. J., and Peralta, E. G. (2001) *J. Biol. Chem.* **276**, 5505 – 5510
655. Heinrich, R., Wenzel, B., and Elsner, N. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9919 – 9923
656. Baudry, M. (1986) in *Excitatory Amino Acids* (Roberts, P. J., Storm-Mathisen, J., and Bradford, H. F., eds), pp. 301 – 321, Macmillian, London
657. Sladeczek, F., Pin, J.-P., Récasens, M., Bockaert, J., and Weiss, S. (1985) *Nature (London)* **317**, 717 – 719
658. Wieloch, T. (1985) *Science* **230**, 681 – 683
659. Johnson, J. L. (1972) *Brain Res.* **37**, 1 – 19
660. Choi, D. W. (1992) *Science* **258**, 241 – 243
661. Armstrong, N., Sun, Y., Chen, G.-Q., and Gouaux, E. (1998) *Nature (London)* **395**, 913 – 917
662. Bortolotto, Z. A., Clarke, V. R. J., Delany, C. M., Parry, M. C., Smolders, I., Vignes, M., Ho, K. H., Miu, P., Brinton, B. T., Fantaske, R., Ogden, A., Gates, M., Ornstein, P. L., Lodge, D., Bleakman, D., and Collingridge, G. L. (1999) *Nature (London)* **402**, 297 – 301
663. Nakanishi, S., and Masu, M. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 319 – 348
664. Nakanishi, S. (1992) *Science* **258**, 597 – 603
665. Abele, R., Keinänen, K., and Madden, D. R. (2000) *J. Biol. Chem.* **275**, 21355 – 21363
666. Rosenmund, C., Stern-Bach, Y., and Stevens, C. F. (1998) *Science* **280**, 1596 – 1599
666a. Zamanillo, D., Sprengel, R., Hvalby, O., Jensen, V., Burnashev, N., Rozov, A., Kaiser, K. M. M., Köster, H. J., Borchardt, T., Worley, P., Lübke, J., Frotscher, M., Kelly, P. H., Sommer, B., Andersen, P., Seeburg, P. H., and Sakmann, B. (1999) *Science* **284**, 1805 – 1811
667. Jayaraman, V., Keesey, R., and Madden, D. R. (2000) *Biochemistry* **39**, 8693 – 8697
667a. Sun, Y., Olson, R., Horning, M., Armstrong, N., Mayer, M., and Gouaux, E. (2002) *Nature (London)* **417**, 245 – 253
667b. Rozov, A., and Burnashev, N. (1999) *Nature (London)* **401**, 594 – 598
668. Meddows, E., Le Bourdellés, B., Grimwood, S., Wafford, K., Sandhu, S., Whiting, P., and McIlhinney, R. A. J. (2001) *J. Biol. Chem.* **276**, 18795 – 18803
669. Nakazawa, T., Komai, S., Tezuka, T., Hisatsune, C., Umemori, H., Semba, K., Mishina, M., Manabe, T., and Yamamoto, T. (2001) *J. Biol. Chem.* **276**, 693 – 699
670. Wong, R. W. C., and Hirokawa, N. (2001) *Trends Biochem. Sci.* **26**, 410 – 411
671. Sheng, M. (1997) *Nature (London)* **386**, 221 – 223
672. Schipke, C. G., Ohlemeyer, C., Matyash, M., Nolte, C., Kettenmann, H., and Kirchhoff, F. (2001) *FASEB J.* **15**, 1270 – 1272
673. Hansson, E., and Rönnbäck, L. (1995) *FASEB J.* **9**, 343 – 350
674. Ivanovic, A., Reiländer, H., Laube, B., and Kuhse, J. (1998) *J. Biol. Chem.* **273**, 19933 – 19937
675. Mothet, J.-P., Parent, A. T., Wolosker, H., Brady, R. O., Jr., Linden, D. J., Ferris, C. D., Rogawski, M. A., and Snyder, S. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4926 – 4931
676. Snyder, S. H., and Ferris, C. D. (2001) in *Synapses* (Cowan, W. M., Südhof, T. C., and Stevens, C. F., eds), pp. 651 – 680, Johns Hopkins Univ. Press, Baltimore
676a. Takasu, M. A., Dalva, M. B., Zigmond, R. E., and Greenberg, M. E. (2002) *Science* **295**, 491 – 495
676b. Ghosh, A. (2002) *Science* **295**, 449 – 451
677. Wong, E. H. F., Kemp, J. A., Priestley, T., Knight, A. R., Woodruff, G. N., and Iversen, L. L. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7104 – 7108
678. Ciruela, F., Escriche, M., Burgueño, J., Angulo, E., Casadó, V., Soloviev, M. M., Canela, E. I., Mallol, J., Chan, W.-Y., Lluis, C., McIlhinney, R. A. J., and Franco, R. (2001) *J. Biol. Chem.* **276**, 18345 – 18351
679. Herrero, I., Miras-Portugal, and Sánchez-Prieto, J. (1998) *J. Biol. Chem.* **273**, 1951 – 1958
680. Kunishima, N., Shimada, Y., Tsuji, Y., Sato, T., Yamamoto, M., Kumaska, T., Nakaniashi, S., Jingami, H., and Morikawa, K. (2000) *Nature (London)* **407**, 971 – 977
680a. Tsuchiya, D., Kunishima, N., Kamiya, N., Jingami, H., and Morikawa, K. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2660 – 2665
681. Conquet, F., Bashir, Z. I., Davies, C. H., Daniel, H., Ferraguti, F., Bordi, F., Franz-Bacon, K., Reggiani, A., Matarese, V., Condé, F., Collingridge, G. L., and Crépel, F. (1994) *Nature (London)* **372**, 237 – 247
682. Ichise, T., Kano, M., Hashimoto, K., Yanagihara, D., Nakao, K., Shigemoto, R., Katsuki, M., and Aiba, A. (2000) *Science* **288**, 1832 – 1835
683. Kubo, Y., Miyashita, T., and Murata, Y. (1998) *Science* **279**, 1722 – 1725
684. Kawabata, S., Kohara, A., Tsutsumi, R., Itahana, H., Hayashibe, S., Yamaguchi, T., and Okada, M. (1998) *J. Biol. Chem.* **273**, 17381 – 17385
685. Bellocchio, E. E., Reimer, R. J., Fremeau, R. T., Jr., and Edwards, R. H. (2000) *Science* **289**, 957 – 960
686. Barbour, B., Brew, H., and Attwell, D. (1988) *Nature (London)* **335**, 433 – 435
687. Mitrovic, A. D., Amara, S. G., Johnston, G. A. R., and Vandenberg, R. J. (1998) *J. Biol. Chem.* **273**, 14698 – 14706
688. Grewer, C., Watzke, N., Wiessner, M., and Rauen, T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9706 – 9711
689. Eskandari, S., Kreman, M., Kavanaugh, M. P., Wright, E. M., and Zampighi, G. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8641 – 8646
690. Grewer, C., Mobarekeh, S. A. M., Watzke, N., Rauen, T., and Schaper, K. (2001) *Biochemistry* **40**, 232 – 240
691. Smith, M. M., Warren, V. A., Thomas, B. S., Brochu, R. M., Ertel, E. A., Rohrer, S., Schaeffer, J., Schmatz, D., Petuch, B. R., Tang, Y. S., Meinke, P. T., Kaczorowski, G. J., and Cohen, C. J. (2000) *Biochemistry* **39**, 5543 – 5554
691a. Slotboom, D. J., Konings, W. N., and Lolkema, J. S. (2001) *Trends Biochem. Sci.* **26**, 534 – 539
692. Cavalheiro, E. A., and Olney, J. W. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5947 – 5948
693. Mukherjee, P. K., DeCoster, M. A., Campbell, F. Z., Davis, R. J., and Bazan, N. G. (1999) *J. Biol. Chem.* **274**, 6493 – 6498
694. Lipton, S. A., and Rosenberg, P. A. (1994) *N. Engl. J. Med.* **330**, 613 – 621
695. Sen, C. K., Khanna, S., Roy, S., and Packer, L. (2000) *J. Biol. Chem.* **275**, 13049 – 13055
696. Olsen, R. W., Wong, E. H. F., Stauber, G. B., and King, R. G. (1984) *Fed. Proc.* **43**, 2773 – 2778
697. Dingledine, R., Myers, S. J., and Nicholas, R. A. (1990) *FASEB J.* **4**, 2636 – 2645
698. Burt, D. R., and Kamatchi, G. L. (1991) *FASEB J.* **5**, 2916 – 2923
699. Gilardi, R. D. (1973) *Nature (London)* **245**, 86 – 88
700. Rudolph, U., Crestani, F., Benke, D., Brünig, I., Benson, J. A., Fritschy, J.-M., Martin, J. R., Bluethmann, H., and Möhler, H. (1999) *Nature (London)* **401**, 796 – 800
701. Schaerer, M. T., Kannenberg, K., Hunziker, P., Baumann, S. W., and Sigel, E. (2001) *J. Biol. Chem.* **276**, 26597 – 26604
702. Renard, S., Olivier, A., Granger, P., Avenet, P., Graham, D., Sevrin, M., George, P., and Besnard, F. (1999) *J. Biol. Chem.* **274**, 13370 – 13374
703. Kannenberg, K., Schaerer, M. T., Fuchs, K., Sieghart, W., and Sigel, E. (1999) *J. Biol. Chem.* **274**, 21257 – 21264
704. O'Shea, S. M., and Harrison, N. L. (2000) *J. Biol. Chem.* **275**, 22764 – 22768
704a. Cromer, B. A., Morton, C. J., and Parker, M. W. (2002) *Trends Biochem. Sci.* **27**, 280 – 287
705. Jayaraman, V., Thiran, S., and Hess, G. P. (1999) *Biochemistry* **38**, 11372 – 11378
706. Connolly, C. N., Wooltorton, J. R. A., Smart, T. G., and Moss, S. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9899 – 9904
707. Kennedy, M. B. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11135 – 11136
708. Chen, L., Wang, H., Vicini, S., and Olsen, R. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11557 – 11562
709. Celio, M. R. (1986) *Science* **231**, 995 – 997
710. Kaupmann, K., Huggel, K., Heid, J., Flor, P. J., Bischoff, S., Mickel, S. J., McMaster, G., Angst, C., Bittiger, H., Froestl, W., and Bettler, B. (1997) *Nature (London)* **386**, 239 – 246
711. Galvez, T., Prézeau, L., Milioti, G., Franek, M., Joly, C., Froestl, W., Bettler, B., Bertrand, H.-O., Blahos, J., and Pin, J.-P. (2000) *J. Biol. Chem.* **275**, 41166 – 41174
712. Kuner, R., Köhr, G., Grünewald, S., Eisenhardt, G., Bach, A., and Kornau, H.-C. (1999) *Science* **283**, 74 – 77
713. Galvez, T., Duthey, B., Kniazeff, J., Blahos, J., Rovelli, G., Bettler, B., Prézeau, L., and Pin, J.-P. (2001) *EMBO J.* **20**, 2152 – 2159
714. Slesinger, P. A., Stoffel, M., Jan, Y. N., and Jan, L. Y. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12210 – 12217
715. Yamakura, T., Mihic, S. J., and Harris, R. A. (1999) *J. Biol. Chem.* **274**, 23006 – 23012
716. Tallman, J. F., Paul, S. M., Skolnick, P., and Gallager, D. W. (1980) *Science* **207**, 274 – 284
717. Löw, K., Crestani, F., Keist, R., Benke, D., Brünig, I., Benson, J. A., Fritschy, J.-M., Rülicke, T., Bluethmann, H., Möhler, H., and Rudolph, U. (2000) *Science* **290**, 131 – 134
718. Wisden, W., and Stephens, D. N. (1999) *Nature (London)* **401**, 751 – 752
719. Olsen, R. W. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4417 – 4418
720. McIntire, S. L., Reimer, R. J., Schuske, K., Edwards, R. H., and Jorgensen, E. M. (1997) *Nature (London)* **389**, 870 – 876
721. Tamura, S., Nelson, H., Tamura, A., and Nelson, N. (1995) *J. Biol. Chem.* **270**, 28712 – 28715
722. Bismuth, Y., Kavanaugh, M. P., and Kanner, B. I. (1997) *J. Biol. Chem.* **272**, 16096 – 16102
723. Storici, P., Capitani, G., De Biase, D., Moser, M., John, R. A., Jansonius, J. N., and Schirmer, T. (1999) *Biochemistry* **38**, 8628 – 8634
724. Gray, T. M., and Matthews, B. W. (1984) *J. Mol. Biol.* **175**, 75 – 81

References

725. Grenningloh, G., Rienitz, A., Schmitt, B., Methfessel, C., Zensen, M., Beyreuther, K., Gundelfinger, E. D., and Betz, H. (1987) *Nature (London)* **328**, 215–220
726. Shan, Q., Haddrill, J. L., and Lynch, J. W. (2001) *J. Biol. Chem.* **276**, 12556–12564
727. White, W. F., and Heller, A. H. (1982) *Nature (London)* **298**, 655–657
728. Becker, C.-M. (1990) *FASEB J.* **4**, 2767–2774
729. Gundlach, A. L. (1990) *FASEB J.* **4**, 2761–2766
730. Rajendra, S., Lynch, J. W., Pierce, K. D., French, C. R., Barry, P. H., and Schofield, P. R. (1994) *J. Biol. Chem.* **269**, 18739–18742
731. Lynch, J. W., Rajendra, S., Pierce, K. D., Handford, C. A., Barry, P. H., and Schofield, P. R. (1997) *EMBO J.* **16**, 110–120
732. Griffon, N., Büttner, C., Nicke, A., Kuhse, J., Schmalzing, G., and Betz, H. (1999) *EMBO J.* **18**, 4711–4721
732a. Leite, J. F., and Cascio, M. (2002) *Biochemistry* **41**, 6140–6148
733. Matzenbach, B., Maulet, Y., Sefton, L., Courtier, B., Avner, P., Guénet, J.-L., and Betz, H. (1994) *J. Biol. Chem.* **269**, 2607–2612
734. García-Alcocer, G., García-Colunga, J., Martínez-Torres, A., and Miledi, R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2781–2785
735. Martínez-Maza, R., Poyatos, I., López-Corcuera, B., Núñez, E., Giménez, C., Zafra, F., and Aragón, C. (2001) *J. Biol. Chem.* **276**, 2168–2173
736. Horiuchi, M., Nicke, A., Gomeza, J., Aschrafi, A., Schmalzing, G., and Betz, H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1448–1453
737. Geerlings, A., Núnez, E., López-Corcuera, B., and Aragón, C. (2001) *J. Biol. Chem.* **276**, 17584–17590
738. Franks, N. P., and Lieb, W. R. (1997) *Nature (London)* **389**, 334–335
739. Beckstead, M. J., Phelan, R., and Mihic, S. J. (2001) *J. Biol. Chem.* **276**, 24959–24964
740. Davies, P. A., Hanna, M. C., Hales, T. G., and Kirkness, E. F. (1997) *Nature (London)* **385**, 820–823
740a. Mihic, S. J., Ye, Q., Wick, M. J., Koltchine, V. V., Krasowski, M. D., Finn, S. E., Mascia, M. P., Valenzuela, C. F., Hanson, K. K., Greenblatt, E. P., Harris, R. A., and Harrison, N. L. (1997) *Nature (London)* **389**, 385–389
741. Minami, K., Wick, M. J., Stern-Bach, Y., Dildy-Mayfield, J. E., Brozowski, S. J., Gonzales, E. L., Trudell, J. R., and Harris, R. A. (1998) *J. Biol. Chem.* **273**, 8248–8255
742. Berry, M. S., and Cottrell, G. A. (1973) *Nature New Biol.* **242**, 250–253
743. Palacios, J. M. (1986) *Nature (London)* **323**, 205
743a. Nishi, A., Bibb, J. A., Snyder, G. L., Higashi, H., Nairn, A. C., and Greengard, P. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12840–12845
744. Axelrod, J., and Saavedra, J. M. (1977) *Nature (London)* **265**, 501–504
745. Gerhardt, C. C., Lodder, H. C., Vincent, M., Bakker, R. A., Planta, R. J., Vreugdenhil, E., Kits, K. S., and van Heerikhuizen, H. (1997) *J. Biol. Chem.* **272**, 6201–6207
746. Chang, D.-J., Li, X.-C., Lee, Y.-S., Kim, H.-K., Kim, U. S., Cho, N. J., Lo, X., Weiss, K. R., Kandel, E. R., and Kaang, B.-K. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1829–1834
746a. Carlsson, A. (2001) *Science* **294**, 1021–1024
747. Balter, M. (1996) *Science* **271**, 909
748. White, F. J. (1998) *Nature (London)* **393**, 118–119
749. Daniels, G. M., and Amara, S. G. (1999) *J. Biol. Chem.* **274**, 35794–35801
750. Lin, Z., Itokawa, M., and Uhl, G. R. (2000) *FASEB J.* **14**, 715–728
751. Reith, M. E. A., Berfield, J. L., Wang, L. C., Ferrer, J. V., and Javitch, J. A. (2001) *J. Biol. Chem.* **276**, 29012–29018
752. Mitchell, D. J., Nikolic, D., Rivera, E., Sablin, S. O., Choi, S., van Breemen, R. B., Singer, T. P., and Silverman, R. B. (2001) *Biochemistry* **40**, 5447–5456
752a. Dajani, R., Cleasby, A., Neu, M., Wonacott, A. J., Jhoti, H., Hood, A. M., Modi, S., Hersey, A., Taskinen, J., Cooke, R. M., Manchee, G. R., and Coughtrie, M. W. H. (1999) *J. Biol. Chem.* **274**, 37862–37868
752b. Bidwell, L. M., McManus, M. E., Gaedigk, A., Kakuta, Y., Negishi, M., Pedersen, L., and Martin, J. L. (1999) *J. Mol. Biol.* **293**, 521–530
753. Chen, S., Xu, M., Lin, F., Lee, D., Riek, P., and Graham, R. M. (1999) *J. Biol. Chem.* **274**, 16320–16330
754. Salminen, T., Varis, M., Nyrönen, T., Pihlavisto, M., Hoffrén, A.-M., Lönnberg, T., Marjamäki, A., Frang, H., Savola, J.-M., Scheinin, M., and Johnson, M. S. (1999) *J. Biol. Chem.* **274**, 23405–23413
755. Wu, D., Jiang, H., and Simon, M. I. (1995) *J. Biol. Chem.* **270**, 9828–9832
756. Turki, J., Lorenz, J. N., Green, S. A., Donnelly, E. T., Jacinto, M., and Liggett, S. B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10483–10488
757. Small, K. M., Brown, K. M., Forbes, S. L., and Liggett, S. B. (2001) *J. Biol. Chem.* **276**, 31596–31601
758. Wang, H.-y, Doronin, S., and Malbon, C. C. (2000) *J. Biol. Chem.* **275**, 36086–36093
759. Ghanouni, P., Steenhuis, J. J., Farrens, D. L., and Kobilka, B. K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5997–6002
760. Liapakis, G., Ballesteros, J. A., Papachristou, S., Chan, W. C., Chen, X., and Javitch, J. A. (2000) *J. Biol. Chem.* **275**, 37779–37788
761. Choi, D.-J., Koch, W. J., Hunter, J. J., and Rockman, H. A. (1997) *J. Biol. Chem.* **272**, 17223–17229
762. Shiina, T., Kawasaki, A., Nagao, T., and Kurose, H. (2000) *J. Biol. Chem.* **275**, 29082–29090
763. Lee, S.-H., Minowa, M. T., and Mouradian, M. M. (1996) *J. Biol. Chem.* **271**, 25292–25299
764. Demchyshyn, L. L., McConkey, F., and Niznik, H. B. (2000) *J. Biol. Chem.* **275**, 23446–23455
765. Castner, S. A., Williams, G. V., and Goldman-Rakic, P. S. (2000) *Science* **287**, 2020–2022
766. Guillin, O., Diaz, J., Carroll, P., Griffon, N., Schwartz, J.-C., and Sokoloff, P. (2001) *Nature (London)* **411**, 86–89
767. White, F. J. (2001) *Nature (London)* **411**, 35–37
768. Iversen, L. L. (1974) *Nature (London)* **250**, 700–701
769. Kater, S. B., and Nicholson, C. (1973) *Intracellular Staining in Neurobiology*, Springer-Verlag., New York
770. Hökfelt, T., Johansson, O., and Goldstein, M. (1984) *Science* **225**, 1326–1334
771. Antelman, S. M., and Caggiula, A. R. (1977) *Science* **195**, 646–653
772. Javitch, J. A., Uhl, G. R., and Snyder, S. H. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4591–4595
773. Ramsay, R. R., and Singer, T. P. (1986) *J. Biol. Chem.* **261**, 7585–7587
774. Langston, J. W., Irwin, I., Langston, E. B., and Forno, L. S. (1984) *Science* **225**, 1480–1482
775. Snyder, S. H., and D'Amato, R. J. (1985) *Nature (London)* **317**, 198–199
776. Lewin, R. (1985) *Science* **230**, 527–528
777. Pentreath, V. W., and Cottrell, G. A. (1974) *Nature (London)* **250**, 655–658
778. Vaney, D. I. (1986) *Science* **233**, 444–446
779. Richardson, B. P., Engel, G., Donatsch, P., and Stadler, P. A. (1985) *Nature (London)* **316**, 126–131
780. Johnson, R. G., and Scarpa, A. (1981) *J. Biol. Chem.* **256**, 11966–11969
781. Iversen, L. L. (1985) *Nature (London)* **316**, 107–108
782. Chong, N. W., Bernard, M., and Klein, D. C. (2000) *J. Biol. Chem.* **275**, 32991–32998
783. Ganguly, S., Gastel, J. A., Weller, J. L., Schwartz, C., Jaffe, H., Namboodiri, M. A. A., Coon, S. L., Hickman, A. B., Rollag, M., Obsil, T., Beauverger, P., Ferry, G., Boutin, J. A., and Klein, D. C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8083–8088
784. Kolata, G. B. (1976) *Science* **192**, 41–42
785. Schmidt, A. W., and Peroutka, S. J. (1989) *FASEB J.* **3**, 2242–2249
786. Julius, D. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15153–15154
787. Hegde, S. S., and Eglen, R. M. (1996) *FASEB J.* **10**, 1398–1407
788. Westbroek, I., van der Plas, A., de Rooij, K. E., Klein-Nulend, J., and Nijweide, P. J. (2001) *J. Biol. Chem.* **276**, 28961–28968
789. The Subcutaneous Sumatriptan International Study Group. (1991) *N. Engl. J. Med.* **325**, 316–321
790. Tang, H., Braun, T. F., and Blair, D. F. (1996) *J. Mol. Biol.* **261**, 209–221
791. Penado, K. M. Y., Rudnick, G., and Stephan, M. M. (1998) *J. Biol. Chem.* **273**, 28098–28106
791a. Scanlon, S. M., Williams, D. C., and Schloss, P. (2001) *Biochemistry* **40**, 10507–10513
791b. Ni, Y. G., Chen, J.-G., Androutsellis-Theotokis, A., Huang, C.-J., Moczydlowski, E., and Rudnick, G. (2001) *J. Biol. Chem.* **276**, 30942–30947
792. Brzezinski, A. (1997) *N. Engl. J. Med.* **336**, 186–195
792a. Roy, D., and Belsham, D. D. (2002) *J. Biol. Chem.* **277**, 251–258
793. Tamarkin, L., Baird, C. J., and Almeida, O. F. X. (1985) *Science* **227**, 714–720
794. Dubocovich, M. L. (1988) *FASEB J.* **2**, 2765–2773
795. Turjanski, A. G., Leonik, F., Estrin, D. A., Rosenstein, R. E., and Doctorovich, F. (2000) *J. Am. Chem. Soc.* **122**, 10468–10469
796. Martín, M., Macías, M., Escames, G., León, J., and Acuña-Castroviejo, D. (2000) *FASEB J.* **14**, 1677–1679
797. Chan-Palay, V., Lin, C.-T., Palay, S., Yamamoto, M., and Wu, J.-Y. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2695–2699
798. McBride, W. J., and Frederickson, R. C. A. (1980) *Fed. Proc.* **39**, 2701–2705
799. Schnell, M. J., Cooper, O. B., and Snyder, S. H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2013–2018
800. Nakatsuka, S., Hayashi, M., Muroyama, A., Otsuka, M., Kozaki, S., Yamada, H., and Moriyama, Y. (2001) *J. Biol. Chem.* **276**, 26589–26596
801. Bowsher, R. R., Verburg, K. M., and Henry, D. P. (1983) *J. Biol. Chem.* **258**, 12215–12220
802. Steinberg, G. H., Kandel, M., Kandel, S. I., and Wells, J. W. (1985) *Biochemistry* **24**, 6107–6115
803. Tanaka, S., Nemoto, K.-i, Yamamura, E., and Ichikawa, A. (1998) *J. Biol. Chem.* **273**, 8177–8182
804. Chuang, W.-L., Christ, M. D., Peng, J., and Rabenstein, D. L. (2000) *Biochemistry* **39**, 3542–3555
805. Chiang, P. K., and Sacktor, B. (1975) *J. Biol. Chem.* **250**, 3399–3408
806. Arrang, J.-M., Garbarg, M., Lancelot, J.-C., Lecomte, J.-M., Pollard, H., Robba, M., Schunack, W., and Schwartz, J.-C. (1987) *Nature (London)* **327**, 117–123
807. Borowsky, B., and 16 other authors. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8966–8971
808. Snyder, S. H., Katims, J. J., Annau, Z., Bruns, R. F., and Daly, J. W. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3260–3264

References

808a. Lindskog, M., Svenningsson, P., Pozzi, L., Kim, Y., Fienberg, A. A., Bibb, J. A., Fredholm, B. B., Nairn, A. C., Greengard, P., and Fisone, G. (2002) *Nature (London)* **418**, 774–778

808b. Vaugeois, J.-M. (2002) *Nature (London)* **418**, 734–736

809. Bruns, R. F., Daly, J. W., and Snyder, S. H. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2077–2080

810. di Tomaso, E., Beltramo, M., and Piomelli, D. (1996) *Nature (London)* **382**, 677–678

811. Jan, L. (1983) *Science* **220**, 64–65

812. Mancillas, J. R., McGinty, J. F., Selverston, A. I., Karten, H., and Bloom, F. E. (1981) *Nature (London)* **293**, 576–578

813. Snyder, S. H., Banerjee, S. P., Yamamura, H. I., and Greenberg, D. (1974) *Science* **184**, 1243–1253

814. Urade, Y., Fujimoto, N., and Hayaishi, O. (1985) *J. Biol. Chem.* **260**, 12410–12415

815. Huidobro-Toro, J. P., and Harris, R. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8078–8082

816. Boger, D. L., Patterson, J. E., and Jin, Q. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4102–4107

816a. Huang, S. M., Bisogno, T., Petros, T. J., Chang, S. Y., Zavitsanos, P. A., Zipkin, R. E., Sivakumar, R., Coop, A., Maeda, D. Y., De Petrocellis, L., Burstein, S., Di Marzo, V., and Walker, J. M. (2001) *J. Biol. Chem.* **276**, 42639–42644

816b. Wilson, R. I., and Nicoll, R. A. (2002) *Science* **296**, 678–682

817. Devane, W. A., Hanus, L., Breuer, A., Pertwee, R. G., Stevenson, L. A., Griffin, G., Gibson, D., Mandelbaum, A., Etinger, A., and Mechoulam, R. (1992) *Science* **258**, 1946–1949

818. Sugiura, T., Kodaka, T., Nakane, S., Miyashita, T., Kondo, S., Suhara, Y., Takayama, H., Waku, K., Seki, C., Baba, N., and Ishima, Y. (1999) *J. Biol. Chem.* **274**, 2794–2801

819. Kozak, K. R., Rowlinson, S. W., and Marnett, L. J. (2000) *J. Biol. Chem.* **275**, 33744–33749

820. Bisogno, T., Maurelli, S., Melck, D., De Petrocellis, L., and Di Marzo, V. (1997) *J. Biol. Chem.* **272**, 3315–3323

821. Hanus, L., Abu-Lafi, S., Fride, E., Breuer, A., Vogel, Z., Shalev, D. E., Kustanovich, I., and Mechoulam, R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3662–3665

822. Kruszka, K. K., and Gross, R. W. (1994) *J. Biol. Chem.* **269**, 14345–14348

823. Iversen, L. (1994) *Nature (London)* **372**, 619

824. Bayewitch, M., Rhee, M.-H., Avidor-Reiss, T., Breuer, A., Mechoulam, R., and Vogel, Z. (1996) *J. Biol. Chem.* **271**, 9902–9905

825. Bouaboula, M., Dussossoy, D., and Casellas, P. (1999) *J. Biol. Chem.* **274**, 20397–20405

826. Ledent, C., Valverde, O., Cossu, G., Petitet, F., Aubert, J.-F., Beslot, F., Böhme, G. A., Imperato, A., Pedrazzini, T., Roques, B. P., Vassart, G., Fratta, W., and Parmentier, M. (1999) *Science* **283**, 401–404

827. Xie, X.-Q., Melvin, L. S., and Makriyannis, A. (1996) *J. Biol. Chem.* **271**, 10640–10647

828. Hanus, L., Breuer, A., Tchilibon, S., Shiloah, S., Goldenberg, D., Horowitz, M., Pertwee, R. G., Ross, R. A., Mechoulam, R., and Fride, E. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14228–14233

829. Calignano, A., La Rana, G., Giuffrida, A., and Piomelli, D. (1998) *Nature (London)* **394**, 277–281

830. Stefano, G. B., Liu, Y., and Goligorsky, M. S. (1996) *J. Biol. Chem.* **271**, 19238–19242

831. Iversen, L. L. (2000) *The Science of Marijuana*, Oxford Univ. Press, Oxford

832. Steiner, H., Bonner, T. I., Zimmer, A. M., Kitai, S. T., and Zimmer, A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5786–5790

833. Meng, I. D., Manning, B. H., Martin, W. J., and Fields, H. L. (1998) *Nature (London)* **395**, 381–383

833a. Panikashvili, D., Simeonidou, C., Ben-Shabat, S., Hanus, L., Breuer, A., Mechoulam, R., and Shohami, E. (2001) *Nature (London)* **413**, 527–531

834. Straus, S. E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9363–9364

835. Baker, D., Pryce, G., Croxford, J. L., Brown, P., Pertwee, R. G., Huffman, J. W., and Layward, L. (2000) *Nature (London)* **404**, 84–87

836. editorial. (2001) *Nature (London)* **410**, 613

837. De Petrocellis, L., Melck, D., Palmisano, A., Bisogno, T., Laezza, C., Bifulco, M., and Di Marzo, V. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8375–8380

838. Wilson, R. I., and Nicoll, R. A. (2001) *Nature (London)* **410**, 588–592

839. Christie, M. J., and Vaughan, C. W. (2001) *Nature (London)* **410**, 527–530

840. Barinaga, M. (2001) *Science* **291**, 2530–2531

840a. Dinh, T. P., Carpenter, D., Leslie, F. M., Freund, T. F., Katona, I., Sensi, S. L., Kathuria, S., and Piomelli, D. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10819–10824

841. Yu, M., Ives, D., and Ramesha, C. S. (1997) *J. Biol. Chem.* **272**, 21181–21186

842. Kozak, K. R., Prusakiewicz, J. J., Rowlinson, S. W., Schneider, C., and Marnett, L. J. (2001) *J. Biol. Chem.* **276**, 30072–30077

843. Carper, J. (2000) *Your Miracle Brain*, Harper Collins Publ., New York

843a. Kitajka, K., Puskás, L. G., Zvara, A., Hackler, L., Jr., Barceló-Coblijn, G., Yeo, Y. K., and Farkas, T. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2619–2624

844. Stuehr, D., Pou, S., and Rosen, G. M. (2001) *J. Biol. Chem.* **276**, 14533–14536

845. Wolff, D. J., Mialkowski, K., Richardson, C. F., and Wilson, S. R. (2001) *Biochemistry* **40**, 37–45

846. Pou, S., Keaton, L., Surichamorn, W., and Rosen, G. M. (1999) *J. Biol. Chem.* **274**, 9573–9580

847. Tochio, H., Zhang, Q., Mandal, P., Li, M., and Zhang, M. (1999) *Nature Struct. Biol.* **6**, 417–421

848. Zhou, X., Espey, M. G., Chen, J. X., Hofseth, L. J., Miranda, K. M., Hussain, S. P., Wink, D. A., and Harris, C. C. (2000) *J. Biol. Chem.* **275**, 21241–21246

849. Cho, A. K. (1990) *Science* **249**, 631–634

850. Musto, D. F. (1991) *Sci. Am.* **265**(Jul), 40–47

851. Courtwright, D. T. (2001) *Forces of Habit Drugs and the Making of the Modern World*, Harvard Univ. Press, Cambridge, Massachusetts

852. Jayanthi, S., Deng, X., Bordelon, M., McCoy, M. T., and Cadt, J. L. (2001) *FASEB J.* **15**, 1745–1752

853. Billman, G. E. (1990) *FASEB J.* **4**, 2469–2475

854. Leshner, A. I. (1996) *N. Engl. J. Med.* **335**, 128–129

854a. Regan, C. (2001) *Intoxicating Minds*, Weidenfeld & Nicolson,

854b. Helmuth, L. (2001) *Science* **294**, 983–984

855. McGehee, D. S., Heath, M. J. S., Gelber, S., Devay, P., and Role, L. W. (1995) *Science* **269**, 1692–1696

856. McGehee, D. S., and Role, L. W. (1996) *Nature (London)* **383**, 670–671

857. Pich, E. M., Pagliusi, S. R., Tessari, M., Talabot-Ayer, D., Hooft van Huijsduijnen, R., and Chiamulera, C. (1997) *Science* **275**, 83–86

858. Breggin, P. R., and Baughman, F. A., Jr. (2001) *Science* **291**, 595

859. Marshall, E. (2000) *Science* **290**, 1280–1282

860. Dole, V. P. (1970) *Ann. Rev. Biochem.* **39**, 821–840

861. Hyman, S. E. (1996) *Science* **273**, 611–612

862. Grinspoon, L., Bakalar, J. B., Zimmer, L., and Morgan, J. P. (1997) *Science* **277**, 748

863. Knapp, R. J., Malatynska, E., Collins, N., Fang, L., Wang, J. Y., Hruby, V. J., Roeske, W. R., and Yamamura, H. I. (1995) *FASEB J.* **9**, 516–525

864. Pak, Y., O'Dowd, B. F., Wang, J. B., and George, S. R. (1999) *J. Biol. Chem.* **274**, 27610–27616

865. Befort, K., Zilliox, C., Filliol, D., Yue, S., and Kieffer, B. L. (1999) *J. Biol. Chem.* **274**, 18574–18581

866. Xu, W., Li, J., Chen, C., Huang, P., Weinstein, H., Javitch, J. A., Shi, L., de Riel, J. K., and Liu-Chen, L.-Y. (2001) *Biochemistry* **40**, 8018–8029

867. Zubieta, J.-K., Smith, Y. R., Bueller, J. A., Xu, Y., Kilbourn, M. R., Jewett, D. M., Meyer, C. R., Koeppe, R. A., and Stohler, C. S. (2001) *Science* **293**, 311–315

867a. Zhang, L., DeHaven, R. N., and Goodman, M. (2002) *Biochemistry* **41**, 61–68

868. Bürgi, H. B., Dunitz, J. D., and Shefter, E. (1973) *Nature (London), New Biology* **244**, 186–188

869. Schulteis, G., and Koob, G. (1994) *Nature (London)* **371**, 108–109

870. Nestler, E. J., and Aghajanian, G. K. (1997) *Science* **278**, 58–63

871. Robbins, T. W., and Everitt, B. J. (1999) *Nature (London)* **398**, 567–570

872. Wickelgren, I. (1997) *Science* **278**, 36–37

873. Nestler, E. J. (2001) *Science* **292**, 2266–2267

873a. Holden, C. (2001) *Science* **294**, 980–982

874. Lieber, C. S. (1988) *N. Engl. J. Med.* **319**, 1639–1650

875. Hoek, J. B., and Taraschi, T. F. (1988) *Trends Biochem. Sci.* **13**, 269–274

876. Pari, K., Sundari, C. S., Chandani, S., and Balasubramanian, D. (2000) *J. Biol. Chem.* **275**, 2455–2462

877. Kodaira, H., Lisek, C. A., Jardine, I., Arimura, A., and Spector, S. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 716–719

878. Weitz, C. J., Lowney, L. I., Faull, K. F., and Feistner, G. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5335–5338

879. Goldstein, A., and Judson, B. A. (1971) *Science* **172**, 290–292

880. Braun, K. P., Cody, R. B., Jr., Jones, D. R., and Peterson, C. M. (1995) *J. Biol. Chem.* **270**, 11263–11266

881. Kumari, M. (2001) *J. Biol. Chem.* **276**, 29764–29771

882. Ikonomidou, C., Bittigau, P., Ishimaru, M. J., Wozniak, D. F., Koch, C., Genz, K., Price, M. T., Stefovska, V., Hörster, F., Tenkova, T., Dikranian, K., and Olney, J. W. (2000) *Science* **287**, 1056–1060

883. Tsukamoto, H., and Lu, S. C. (2001) *FASEB J.* **15**, 1335–1349

884. Thiele, T. E., Marsh, D. J., Marie, L. S., Bernstein, I. L., and Palmiter, R. D. (1998) *Nature (London)* **396**, 366–369

885. von Hungen, K., Roberts, S., and Hill, D. F. (1974) *Nature (London)* **252**, 588–589

886. Ricaurte, G., Bryan, G., Strauss, L., Seiden, L., and Schuster, C. (1985) *Science* **229**, 986–988

886a. Simantov, R., and Tauber, M. (1997) *FASEB J.* **11**, 141–146

887. Vincent, J. P., Kartalovski, B., Geneste, P., Kamenka, J. M., and Lazdunski, M. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4678–4682

888. Keup, W., ed. (1970) *Origin and Mechanisms of Hallucinations*, Plenum, New York

889. Holzman, D. (1993) *Science* **259**, 25–26

890. Aquilina, J. A., Carver, J. A., and Truscott, R. J. W. (1999) *Biochemistry* **38**, 11455–11464

891. Lipton, S. A., Kim, W.-K., Choi, Y.-B., Kumar, S., D'Emilia, D. M., Rayudu, P. V., Arnelle, D. R., and Stamler, J. S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5923–5928

892. Borisenko, G. G., Kagan, V. E., Hsia, C. J. C., and Schor, N. F. (2000) *Biochemistry* **39**, 3392–3400

893. Zvosec, D. L., Smith, S. W., McCutcheon, J. R., Spillane, J., Hall, B. J., and Peacock, E. A. (2001) *N. Engl. J. Med.* **344**, 87–94

894. Williams, S. (1990) *Science* **248**, 958

References

895. Glickstein, M. (1988) *Sci. Am.* **259**(Sep), 118–127
896. Supèr, H., Spekreijse, H., and Lamme, V. A. F. (2001) *Science* **293**, 120–124
897. Miyashita, Y. (1995) *Science* **268**, 1719–1720
898. Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. E. (1998) *Science* **281**, 1185–1187
899. Gold, G. H., and Pugh, E. N., Jr. (1997) *Nature (London)* **385**, 677–679
900. Lindstedt, K. J. (1971) *Science* **173**, 333–334
901. Nevitt, G. A., Dittman, A. H., Quinn, T. P., and Moody, W. J., Jr. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4288–4292
902. Krieger, J., and Breer, H. (1999) *Science* **286**, 720–723
903. Thomas, J. H. (1994) *Science* **264**, 1698–1699
904. Zhang, Y., Chou, J. H., Bradley, J., Bargmann, C. I., and Zinn, K. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12162–12167
905. Mombaerts, P. (1999) *Science* **286**, 707–711
906. Störtkuhl, K. F., and Kettler, R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9381–9385
906a. Pennisi, E. (1999) *Science* **283**, 1239
907. Yoshihara, Y., Nagao, H., and Mori, K. (2001) *Science* **291**, 835,837
908. Axel, R. (1995) *Sci. Am.* **273**(Oct), 154–159
909. Firestein, S. (2001) *Nature (London)* **413**, 211–218
910. Snyder, S. H., Sklar, P. B., and Pevsner, J. (1988) *J. Biol. Chem.* **263**, 13971–13974
911. Sklar, P. B., Anholt, R. R. H., and Snyder, S. H. (1986) *J. Biol. Chem.* **261**, 15538–15543
912. Mori, K., Nagao, H., and Yoshihara, Y. (1999) *Science* **286**, 711–715
913. Floriano, W. B., Vaidehi, N., Goddard, W. A., III, Singer, M. S., and Shepherd, G. M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10712–10716
914. Gheusi, G., Cremer, H., McLean, H., Chazal, G., Vincent, J.-D., and Lledo, P.-M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1823–1828
915. Keverne, E. B. (1999) *Science* **286**, 716–720
916. Meyer, M. R., Angele, A., Kremmer, E., Kaupp, U. B., and Müller, F. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10595–10600
917. Bal, R. S., and Anholt, R. R. H. (1993) *Biochemistry* **32**, 1047–1053
918. Spinelli, S., Ramoni, R., Grolli, S., Bonicel, J., Cambillau, C., and Tagoni, M. (1998) *Biochemistry* **37**, 7913–7918
919. Vincent, F., Spinelli, S., Ramoni, R., Grolli, S., Pelosi, P., Cambillau, C., and Tegoni, M. (2000) *J. Mol. Biol.* **300**, 127–139
920. Timm, D. E., Baker, L. J., Mueller, H., Zidek, L., and Novotny, M. V. (2001) *Protein Sci.* **10**, 997–1004
920a. Horst, R., Damberger, F., Luginbühl, P., Güntert, P., Peng, G., Nikonova, L., Leal, W. S., and Wüthrich, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14374–14379
921. Zeng, C., Spielman, A. I., Vowels, B. R., Leyden, J. J., Biemann, K., and Preti, G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6626–6630
922. Wysocki, C. J., Dorries, K. M., and Beauchamp, G. K. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7976–7978
923. Firestein, S. (2000) *Nature (London)* **404**, 552–553
923a. Margolskee, R. F. (2002) *J. Biol. Chem.* **277**, 1–4
924. Lindemann, B. (2001) *Nature (London)* **413**, 219–225
925. Matsunami, H., Montmayeur, J.-P., and Buck, L. B. (2000) *Nature (London)* **404**, 601–604
926. Adler, E., Hoon, M. A., Mueller, K. L., Chandrashekar, J., Ryba, N. J. P., and Zuker, C. S. (2000) *Cell* **100**, 693–702
927. Chandrashekar, J. (2000) *Cell* **100**, 703–711
928. Caicedo, A., and Roper, S. D. (2001) *Science* **291**, 1557–1560
929. Clyne, P. J., Warr, C. G., and Carlson, J. R. (2000) *Science* **287**, 1830–1834
930. Ishimoto, H., Matsumoto, A., and Tanimura, T. (2000) *Science* **289**, 116–119
931. McLaughlin, S. K., McKinnon, P. J., and Margolskee, R. F. (1992) *Nature (London)* **357**, 563–569
932. Ruiz-Avila, L., Wong, G. T., Damak, S., and Margolskee, R. F. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8868–8873
933. Misaka, T., Kusakabe, Y., Emori, Y., Gonoi, T., Arai, S., and Abe, K. (1997) *J. Biol. Chem.* **272**, 22623–22629
934. Li, X.-J., and Snyder, S. H. (1995) *J. Biol. Chem.* **270**, 17674–17679
935. Gachon, A. M. F. (1993) *Trends Biochem. Sci.* **18**, 206–207
936. Robyt, J. F. (1998) *Essentials of Carbohydrate Chemistry*, Springer, New York
937. Stegink, L. D., and Filer, L. J., Jr. (1984) *Aspartame*, Dekker, New York
938. Fuller, W. D., Goodman, M., and Verlander, M. S. (1985) *J. Am. Chem. Soc.* **107**, 5821–5822
939. Compadre, C. M., Pezzuto, J. M., Kinghorn, A. D., and Kamath, S. K. (1985) *Science* **227**, 417–418
940. de Vos, A. M., Hatada, M., Van Der Wel, H., Krabbendam, H., Peerdeman, A. F., and Kim, S.-H. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1406–1409
941. Kim, S.-H., de Vos, A., and Ogata, C. (1988) *Trends Biochem. Sci.* **13**, 13–15
942. Somoza, J. R., Jiang, F., Tong, L., Kang, C.-H., Cho, J. M., and Kim, S.-H. (1993) *J. Mol. Biol.* **234**, 390–404
943. Spadaccini, R., Crescenzi, O., Tancredi, T., De Casamassimi, N., Saviano, G., Scognamiglio, R., Di Donato, A., and Temussi, P. A. (2001) *J. Mol. Biol.* **305**, 505–514
944. Van Der Wel, H. (1980) *Trends Biochem. Sci.* **5**, 122–123
945. Theerasilp, S., and Kurihara, Y. (1988) *J. Biol. Chem.* **263**, 11536–11539
946. Harada, S., Otani, H., Maeda, S., Kai, Y., Kasai, N., and Kurihara, Y. (1994) *J. Mol. Biol.* **238**, 286–287
947. Bartoshuk, L. M., Lee, C.-H., and Scarpellino, R. (1972) *Science* **178**, 988–990
948. Julius, D., and Basbaum, A. I. (2001) *Nature (London)* **413**, 203–210
949. Clapham, D. E. (1997) *Nature (London)* **389**, 783–784
950. Vogel, G. (2000) *Science* **288**, 241–242
951. Kedei, N., Szabo, T., Lile, J. D., Treanor, J. J., Olah, Z., Iadarola, M. J., and Blumberg, P. M. (2001) *J. Biol. Chem.* **276**, 28613–28619
952. Caterina, M. J., Schumacher, M. A., Tominaga, M., Rosen, T. A., Levine, J. D., and Julius, D. (1997) *Nature (London)* **389**, 816–824
953. Levitan, I. B., and Cibulsky, S. M. (2001) *Science* **293**, 1270–1271
954. Cahalan, M. D. (2001) *Nature (London)* **411**, 542–544
955. Boger, D. L., Sato, H., Lerner, A. E., Hedrick, M. P., Fecik, R. A., Miyauchi, H., Wilkie, G. D., Austin, B. J., Patricelli, M. P., and Cravatt, B. F. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5044–5049
955a. McKemy, D. D., Neuhausser, W. M., and Julius, D. (2002) *Nature (London)* **416**, 52–58
955b. Smith, G. D., Gunthorpe, M. J., Kelsell, R. E., Hayes, P. D., Reilly, P., Facer, P., Wright, J. E., Jerman, J. C., Walhin, J.-P., Ooi, L., Egerton, J., Charles, K. J., Smart, D., Randall, A. D., Anand, P., and Davis, J. B. (2002) *Nature (London)* **418**, 186–190
956. Ingber, D. (1999) *FASEB J.* **13**, S3–S15
957. Gillespie, P. G., and Walker, R. G. (2001) *Nature (London)* **413**, 194–202
957a. Bezanilla, F., and Perozo, E. (2002) *Science* **298**, 1562–1563
957b. Bass, R. B., Strop, P., Barclay, M., and Rees, D. C. (2002) *Science* **298**, 1582–1587
958. Sukharev, S., Betanzos, M., Chiang, C.-S., and Guy, H. R. (2001) *Nature (London)* **409**, 720–724
958a. Perozo, E., Cortes, D. M., Sompornpisut, P., Kloda, A., and Martinac, B. (2002) *Nature (London)* **418**, 942–948
959. Watson, P. A. (1991) *FASEB J.* **5**, 2013–2019
960. Steel, K. P. (1999) *Science* **285**, 1363–1364
961. Cho, A. (2000) *Science* **288**, 1954–1955
962. Zheng, J., Shen, W., He, D. Z. Z., Long, K. B., Madison, L. D., and Dallos, P. (2000) *Nature (London)* **405**, 149–155
963. Oliver, D., He, D. Z. Z., Klöcker, N., Ludwig, J., Schulte, U., Waldegger, S., Ruppersberg, J. P., Dallos, P., and Fakler, B. (2001) *Science* **292**, 2340–2343
964. Parker, D. E. (1980) *Sci. Am.* **243**(Nov), 118–132
965. Lacy-Hulbert, A., Metcalfe, J. C., and Hesketh, R. (1998) *FASEB J.* **12**, 395–420
966. Bailey, C. H., Bartsch, D., and Kandel, E. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13445–13452
967. Thompson, R. F., and Kim, J. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13438–13444
967a. Kandel, E. R. (2001) *Science* **294**, 1030–1038
968. Squire, L. R., and Zola-Morgan, S. (1991) *Science* **253**, 1380–1386
969. Henke, K., Weber, B., Kneifel, S., Wieser, H. G., and Buck, A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5884–5889
970. Biegler, R., McGregor, A., Krebs, J. R., and Healy, S. D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6941–6944
971. Frey, S., and Petrides, M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8723–8727
972. McGaugh, J. L., Cahill, L., and Roozendaal, B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13508–13514
972a. Nakazawa, K., Quirk, M. C., Chitwood, R. A., Watanabe, M., Yeckel, M. F., Sun, L. D., Kato, A., Carr, C. A., Johnston, D., Wilson, M. A., and Tonegawa, S. (2002) *Science* **297**, 211–218
973. Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J., and Dolan, R. J. (1994) *Nature (London)* **368**, 633–635
974. Teng, E., and Squire, L. R. (1999) *Nature (London)* **400**, 675–677
975. Jiang, Y., Haxby, J. V., Martin, A., Ungerleider, L. G., and Parasuraman, R. (2000) *Science* **287**, 643–646
976. Smith, E. E., Geva, A., Jonides, J., Miller, A., Reuter-Lorenz, P., and Koeppe, R. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2095–2100
977. Courtney, S. M., Ungerleider, L. G., Keil, K., and Haxby, J. V. (1997) *Nature (London)* **386**, 608–611
978. Lisman, J. (1998) *Nature (London)* **394**, 132–133
979. Lisman, J. E., and Idiart, M. A. P. (1995) *Science* **267**, 1512–1514
980. Alkon, D. L. (1983) *Sci. Am.* **249**(Jul), 70–84
981. Bailey, C. H., Chen, M., Keller, F., and Kandel, E. R. (1992) *Science* **256**, 645–649
982. Johnston, D. (1997) *Science* **278**, 401–402
983. Sherff, C. M., and Carew, T. J. (1999) *Science* **285**, 1911–1914
984. Levenson, J., Endo, S., Kategaya, L. S., Fernandez, R. I., Brabham, D. G., Chin, J., Byrne, J. H., and Eskin, A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12858–12863
985. Menzel, R., and Müller, U. (2001) *Nature (London)* **411**, 433–434
986. Zars, T., Fischer, M., Schulz, R., and Heisenberg, M. (2000) *Science* **288**, 672–675
987. Grotewiel, M. S., Beck, C. D. O., Wu, K. H., Zhu, X.-R., and Davis, R. L. (1998) *Nature (London)* **391**, 455–460
987a. Pascual, A., and Préat, T. (2001) *Science* **294**, 1115–1117
987b. Perez-Orive, J., Mazor, O., Turner, G. C., Cassenaer, S., Wilson, R. I., and Laurent, G. (2002) *Science* **297**, 359–365

References

988. Draguhn, A., Traub, R. D., Schmitz, D., and Jefferys, J. G. R. (1998) *Nature (London)* **394**, 189–192

989. Winson, J. (1990) *Sci. Am.* **263**(Nov), 86–96

990. Schechter, B. (1996) *Science* **274**, 339–340

991. Joliot, M., Ribary, U., and Llinás, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11748–11751

992. Fisahn, A., Pike, F. G., Buhl, E. H., and Paulsen, O. (1998) *Nature (London)* **394**, 186–189

993. Wichmann, T., and DeLong, M. R. (1999) *Nature (London)* **400**, 621–622

994. Plenz, D., and Kital, S. T. (1999) *Nature (London)* **400**, 677–682

995. Whittington, M. A., Traub, R. D., and Jefferys, J. G. R. (1995) *Nature (London)* **373**, 612–615

996. Perkins, W. H. (2001) *Science* **294**, 786

997. Goldbeter, A. (1996) *Biochemical Oscillations and Cellular Rhythms: The Molecular Basis of Periodic and Chaotic Behavior*, Cambridge Univ. Press, London and New York

998. Vergara, L. A., Stojikovic, S. S., and Rojas, E. (1995) *Biophys. J.* **69**, 1606–1614

999. Thomas, A. P., Bird, G. S. J., Hajnóczky, G., Robb-Gaspers, L. D., and Putney, J. W., Jr. (1996) *FASEB J.* **10**, 1505–1517

1000. Hoyer, J., Köhler, R., and Distler, A. (1998) *FASEB J.* **12**, 359–366

1001. Marchant, J. S., and Parker, I. (2001) *EMBO J.* **20**, 65–76

1002. Chen, C. F., von Baumgarten, R., and Takeda, R. (1971) *Nature New Biol.* **233**, 27–29

1003. Nadim, F., Manor, Y., Kopell, N., and Marder, E. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8206–8211

1004. Loewenstein, Y., Yarom, Y., and Sompolinsky, H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8095–8100

1005. Posner, M. I. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7398–7403

1006. Crick, F., and Koch, C. (1995) *Nature (London)* **375**, 121–123

1007. Koch, C., Zador, A., and Brown, T. H. (1992) *Science* **256**, 973–974

1008. Häusser, M., Spruston, N., and Stuart, G. J. (2000) *Science* **290**, 739–744

1009. Kaech, S., Parmar, H., Roelandse, M., Bornmann, C., and Matus, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7086–7092

1010. Gerstner, W., Kreiter, A. K., Markram, H., and Herz, A. V. M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12740–12741

1011. Fairhall, A. L., Lewen, G. D., Bialek, W., and de Ruyter van Steveninck, R. R. (2001) *Nature (London)* **412**, 787–792

1012. Salinas, E., and Romo, R. (2000) *Nature (London)* **404**, 131–133

1013. MacLeod, K., and Laurent, G. (1996) *Science* **274**, 976–979

1014. Barinaga, M. (1998) *Science* **280**, 376–378

1015. Alpern, H. P., and Crabbe, J. C. (1972) *Science* **177**, 722–724

1016. Goelet, P., Castellucci, V. F., Schacher, S., and Kandel, E. R. (1986) *Nature (London)* **322**, 419–422

1017. Belardetti, F., Kandel, E. R., and Siegelbaum, S. A. (1987) *Nature (London)* **325**, 153–156

1018. Morris, R. G. M., Anderson, E., Lynch, G. S., and Baudry, M. (1986) *Nature (London)* **319**, 774–776

1019. Kennedy, M. B. (1988) *Nature (London)* **335**, 770–772

1020. Bliss, T. V. P., and Collingridge, G. L. (1993) *Nature (London)* **361**, 31–39

1021. Schuman, E. M., and Madison, D. V. (1994) *Science* **263**, 532–536

1022. Malenka, R. C., and Nicoll, R. A. (1999) *Science* **285**, 1870–1874

1023. Shen, K., and Meyer, T. (1999) *Science* **284**, 162–166

1024. Silva, A. J., Paylor, R., Wehner, J. M., and Tonegawa, S. (1992) *Science* **257**, 206–211

1025. Dosemeci, A., Tao-Cheng, J.-H., Vinade, L., Winters, C. A., Pozzo-Miller, L., and Reese, T. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10428–10432

1026. Frey, U., and Morris, R. G. M. (1997) *Nature (London)* **385**, 533–536

1027. Barinaga, M. (2000) *Science* **290**, 736–738

1028. Buchs, P.-A., and Muller, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8040–8045

1029. Bear, M. F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13453–13459

1030. Rioult-Pedotti, M.-S., Friedman, D., and Donoghue, J. P. (2000) *Science* **290**, 533–536

1031. Tsien, J. Z. (2000) *Sci. Am.* **282**(Apr), 62–68

1032. Tang, Y.-P., Shimizu, E., Dube, G. R., Rampon, C., Kerchner, G. A., Zhuo, M., Liu, G., and Tsien, J. Z. (1999) *Nature (London)* **401**, 63–69

1033. Schuman, E. M., and Madison, D. V. (1991) *Science* **254**, 1503–1506

1034. Montague, P. R., Gancayco, C. D., Winn, M. J., Marchase, R. B., and Friedlander, M. J. (1994) *Science* **263**, 973–977

1035. Bear, M. F. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9457–9458

1036. Mulkey, R. M., Herron, C. E., and Malenka, R. C. (1993) *Science* **261**, 1051–1055

1037. Reyes-Harde, M., Empson, R., Potter, B. V. L., Galione, A., and Stanton, P. K. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4061–4066

1038. Nicoll, R. A., and Malenka, R. C. (1997) *Nature (London)* **388**, 427–428

1039. John, E. R., Tang, Y., Brill, A. B., Young, R., and Ono, K. (1986) *Science* **233**, 1167–1175

1040. Ungar, G. (1972) *Naturwissenschaften* **59**, 85–91

1041. White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986) *Phil. Trans. Roy. Soc. London B* **314**, 1–

1042. Herman, R. K. (1993) *Nature (London)* **364**, 282–283

1043. Kozloski, J., Hamzei-Sichani, F., and Yuste, R. (2001) *Science* **293**, 868–872

1044. Gupta, A., Wang, Y., and Markram, H. (2000) *Science* **287**, 273–278

1045. Steriade, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3625–3627

1046. Galarreta, M., and Hestrin, S. (2001) *Science* **292**, 2295–2299

1047. Mackintosh, N. J. (1998) *IQ and Human Intelligence*, Oxford Univ. Press, Oxford

1048. Sternberg, R. J. (2000) *Science* **289**, 399–401

1049. Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F. N., and Emslie, H. (2000) *Science* **289**, 457–460

1050. Devlin, B., Daniels, M., and Roeder, K. (1997) *Nature (London)* **388**, 468–471

1051. Herrnstein, R. J., and Murray, C. (1994) *The Bell Curve: Intelligence and Class Structure in American Life*, Free Press, New York

1052. McClearn, G. E., Johansson, B., Berg, S., Pedersen, N. L., Ahern, F., Petrill, S. A., and Plomin, R. (1997) *Science* **276**, 1560–1563

1053. Hermelin, B. (2001) *Bright Splinters of the Mind: A Personal Story of Research with Autistic Savants*, Jessica Kingsley Publishers, London

1054. Snyder, A. (2001) *Nature (London)* **413**, 251–252

1055. Stokstad, E. (2001) *Science* **294**, 34–37

1056. Greenspan, R. J. (1995) *Sci. Am.* **272**(Apr), 72–78

1057. Pfaff, D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5957–5960

1058. Bouchard, T. J., Jr. (1994) *Science* **264**, 1700–1701

1059. Pinker, S. (2001) *Nature (London)* **413**, 465–466

1060. Wickelgren, I. (1999) *Science* **284**, 571

1061. Young, M. W. (2000) *Sci. Am.* **282**(Mar), 64–71

1062. Sassone-Corsi, P. (1998) *Nature (London)* **392**, 871–874

1062a. Reppert, S. M., and Weaver, D. R. (2002) *Nature (London)* **418**, 935–941

1063. Morse, D. S., Fritz, L., and Hastings, J. W. (1990) *Trends Biochem. Sci.* **15**, 262–265

1064. Roenneberg, T., and Rehman, J. (1996) *FASEB J.* **10**, 1443–1447

1065. Czeisler, C. A., Duffy, J. F., Shanahan, T. L., Brown, E. N., Mitchell, J. F., Rimmer, D. W., Ronda, J. M., Silva, E. J., Allan, J. S., Emens, J. S., Dijk, D.-J., and Kronauer, R. E. (1999) *Science* **284**, 2177–2181

1066. Kondo, T., Mori, T., Lebedeva, N. V., Aoki, S., Ishiura, M., and Golden, S. S. (1997) *Science* **275**, 224–227

1067. Iwasaki, H., Taniguchi, Y., Ishiura, M., and Kondo, T. (1999) *EMBO J.* **18**, 1137–1145

1068. Crosthwaite, S. K., Dunlap, J. C., and Loros, J. J. (1997) *Science* **276**, 763–769

1069. Lee, C., Parikh, V., Itsukaichi, T., Bae, K., and Edery, I. (1996) *Science* **271**, 1740–1744

1070. Young, M. W. (2000) *Science* **288**, 451–453

1071. Naidoo, N., Song, W., Hunter–Ensor, M., and Sehgal, A. (1999) *Science* **285**, 1737–1741

1071a. Williams, J. A., Su, H. S., Bernards, A., Field, J., and Sehgal, A. (2001) *Science* **293**, 2251–2256

1071b. Panda, S., Hogenesch, J. B., and Kay, S. A. (2002) *Nature (London)* **417**, 329–335

1072. Green, C. B., and Besharse, J. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14884–14888

1073. Okamura, H., Miyake, S., Sumi, Y., Yamaguchi, S., Yasui, A., Muijtjens, M., Hoeijmakers, J. H. J., and van der Horst, G. T. J. (1999) *Science* **286**, 2531–2534

1074. Gekakis, N., Staknis, D., Nguyen, H. B., Davis, F. C., Wilsbacher, L. D., King, D. P., Takahashi, J. S., and Weitz, C. J. (1998) *Science* **280**, 1564–1569

1075. Suárez-López, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001) *Nature (London)* **410**, 1116–1120

1076. Schibler, U., Ripperger, J. A., and Brown, S. A. (2001) *Science* **293**, 437–438

1077. Rutter, J., Reick, M., Wu, L. C., and McKnight, S. L. (2001) *Science* **293**, 510–514

1078. Tosini, G., and Menaker, M. (1996) *Science* **272**, 419–421

1079. Morell, V. (1995) *Science* **272**, 349

1080. Turek, F. W. (1981) *Nature (London)* **292**, 289–290

1080a. Pennartz, C. M. A., de Jeu, M. T. G., Bos, N. P. A., Schaap, J., and Geurtsen, A. M. S. (2002) *Nature (London)* **416**, 286–290

1080b. Cheng, M. Y., Bullock, C. M., Li, C., Lee, A. G., Bermak, J. C., Belluzzi, J., Weaver, D. R., Leslie, F. M., and Zhou, Q.-Y. (2002) *Nature (London)* **417**, 405–410

1081. Klein, D. C., and Namboodiri, M. A. A. (1982) *Trends Biochem. Sci.* **7**, 98–102

1082. Reppert, S. M., Weaver, D. R., Rivkees, S. A., and Stopa, E. G. (1988) *Science* **242**, 78–81

1083. Baler, R., and Klein, D. C. (1995) *J. Biol. Chem.* **270**, 27319–27325

1084. Czeisler, C. A., Allan, J. S., Strogatz, S. H., Ronda, J. M., Sánchez, R., Ríos, C. D., Freitag, W. O., Richardson, G. S., and Kronauer, R. E. (1986) *Science* **233**, 667–671

1085. Morell, V. (1996) *Science* **271**, 905–906

1086. Krueger, J. M., Pappenheimer, J. R., and Karnovsky, M. L. (1982) *J. Biol. Chem.* **257**, 1664–1669

1087. Hayaishi, O. (1988) *J. Biol. Chem.* **263**, 14593–14596

1088. Takahashi, J. S. (1999) *Science* **285**, 2076–2077

1088a. Siegel, J. M. (2001) *Science* **294**, 1058–1063

1088b. Maquet, P. (2001) *Science* **294**, 1048–1052

1089. Roush, W. (1995) *Science* **269**, 1220–1221

References

1090. Born, J., Hansen, K., Marshall, L., Mölle, M., and Fehm, H. L. (1999) *Nature (London)* **397**, 29–30
1091. Mefford, I. N., Baker, T. L., Boehme, R., Foutz, A. S., Ciaranello, R. D., Barchas, J. D., and Dement, W. C. (1983) *Science* **220**, 629–632
1092. Chicurel, M. (2000) *Nature (London)* **407**, 554–556
1093. Toh, K. L., Jones, C. R., He, Y., Eide, E. J., Hinz, W. A., Virshup, D. M., Ptácek, L. J., and Fu, Y.-H. (2001) *Science* **291**, 1040–1043
1094. Kondo, N., and Kondo, J. (1992) *J. Biol. Chem.* **267**, 473–478
1095. Holden, C. (1986) *Science* **233**, 723–726
1096. Holden, C. (1991) *Science* **254**, 1450–1452
1097. Stinson, S. C. (1990) *Chem. Eng. News* **Oct 15**, 33–68
1098. Holden, C. (2000) *Science* **288**, 39–40
1099. Kanof, P. D., and Greengard, P. (1978) *Nature (London)* **272**, 329–333
1100. Sandler, M., Ruthven, C. R. J., Goodwin, B. L., Reynolds, G. P., Rao, V. A. R., and Cooper, A. (1979) *Nature (London)* **278**, 357–358
1101. Maycock, A. L., Abeles, R. H., Salach, J. I., and Singer, T. P. (1976) *Biochemistry* **15**, 114–125
1102. Post, M. L., Kennard, O., and Horn, A. S. (1974) *Nature (London)* **252**, 492–495
1103. Barondes, S. H. (1994) *Science* **263**, 1102–1103
1104. García-Colunga, J., Awad, J. N., and Miledi, R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2041–2044
1105. Enserink, M. (1999) *Science* **284**, 238–240
1106. Kramer, M. S., and 26 other authors. (1998) *Science* **281**, 1640–1645
1106a. Sapolsky, R. M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12320–12322
1107. Vogel, G. (2000) *Science* **290**, 258–259
1108. Lindenbaum, J., Healton, E. B., Savage, D. G., Brust, J. C. M., Garrett, T. J., Podell, E. R., Marcell, P. D., Stabler, S. P., and Allen, R. H. (1988) *N. Engl. J. Med.* **318**, 1720–1728
1109. Marx, J. L. (1985) *Science* **227**, 934
1110. Wurtman, R. J., and Wurtman, J. J. (1989) *Sci. Am.* **260**(Jan), 68–75
1111. Segal, D. S., Callaghan, M., and Mandell, A. J. (1975) *Nature (London)* **254**, 58–59
1112. Price, L. H., and Heninger, G. R. (1994) *N. Engl. J. Med.* **331**, 591–598
1113. Dixon, J. F., Los, G. V., and Hokin, L. E. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8358–8362
1113a. Dichtl, B., Stevens, A., and Tollervey, D. (1997) *EMBO J.* **16**, 7184–7195
1114. Menkes, H. A., Baraban, J. M., Freed, A. N., and Snyder, S. H. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5727–5730
1115. Dobner, P. R., Tischler, A. S., Lee, Y. C., Bloom, S. R., and Donahue, S. R. (1988) *J. Biol. Chem.* **263**, 13983–13986
1116. Morell, V. (1996) *Science* **272**, 31–32
1117. Kelsoe, J. R., Spence, M. A., Loetscher, E., Foguet, M., Sadovnick, A. D., Remick, R. A., Flodman, P., Khristich, J., Mroczkowski-Parker, Z., Brown, J. L., Masser, D., Ungerleider, S., Rapaport, M. H., Wishart, W. L., and Luebbert, H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 585–590
1117a. Thompson, P. M., Vidal, C., Giedd, J. N., Gochman, P., Blumenthal, J., Nicolson, R., Toga, A. W., and Rapoport, J. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11650–11655
1117b. Cannon, T. D., Thompson, P. M., van Erp, T. G. M., Toga, A. W., Poutanen, V.-P., Huttunen, M., Lonnqvist, J., Standerskjold-Nordenstam, C.-G., Narr, K. L., Khaledy, M., Zoumalan, C. I., Dail, R., and Kaprio, J. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3228–3233
1118. Barnes, D. H. (1987) *Science* **235**, 430–433
1119. Brzustowicz, L. M., Hodgkinson, K. A., Chow, E. W. C., Honer, W. G., and Bassett, A. S. (2000) *Science* **288**, 678–682
1119a. Sawa, A., and Snyder, S. H. (2002) *Science* **296**, 692–695
1120. Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., Goldman, D., and Weinberger, D. R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6917–6922
1121. Nettle, D. (2001) *Nature (London)* **412**, 119
1122. Zirkle, C. L., and Kaiser, C. (1970) in *Medicinal Chemistry*, 3rd ed., Vol. Part II (Burger, A., ed), pp. 1410–1469, Wiley (Interscience), New York
1123. Meltzer, H. Y., Park, S., and Kessler, R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13591–13593
1124. Seeman, P., and Lee, T. (1975) *Science* **188**, 1217–1219
1125. Davis, K. L., Davidson, M., Mohs, R. C., Kendler, K. S., Davis, B. M., Johns, C. A., DeNigris, Y., and Horvath, T. B. (1985) *Science* **227**, 1601–1602
1126. Taubes, G. (1994) *Science* **265**, 1034–1035
1127. Hornykiewicz, O. (1982) *Nature (London)* **299**, 484–486
1128. Wang, R. Y., and Schoenfeld, R., eds. (1987) *Cholecystokinin Antagonists*, Liss, New York
1129. Woo, T.-U., Whitehead, R. E., Melchitzky, D. S., and Lewis, D. A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5341–5346
1130. Moghaddam, B., and Adams, B. W. (1998) *Science* **281**, 1349–1351
1131. Goodman, A. B. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7240–7244
1132. Lewis, D. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4293–4294
1132a. Hakak, Y., Walker, J. R., Li, C., Wong, W. H., Davis, K. L., Buxbaum, J. D., Haroutunian, V., and Fienberg, A. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4746–4751
1132b. Chakravarti, A. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4755–4756
1133. Stein, L., and Wise, C. D. (1971) *Science* **171**, 1032–1036
1134. Friedhoff, A. J. (1973) in *Biological Psychiatry* (Mendels, J., ed), pp. 113–129, Wiley, New York
1135. Wyatt, R. J., Erdelyi, E., Do Amaral, J. R., Elliott, G. R., Renson, J., and Barchas, J. D. (1975) *Science* **187**, 853–855
1136. Shen, R.-S., Smith, R. V., Davis, P. J., Brubaker, A., and Abell, C. W. (1982) *J. Biol. Chem.* **257**, 7294–7297
1137. Dichter, M. A., and Ayala, G. F. (1987) *Science* **237**, 157–163
1138. McNamara, J. O. (1999) *Nature (London)* **399**, A15–A22
1138a. Browne, T. R., and Holmes, G. L. (2001) *N. Engl. J. Med.* **344**, 1145–1151
1139. Kash, S. F., Tecott, L. H., Hodge, C., and Baekkeskov, S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1698–1703
1140. Kanaani, J., Lissin, D., Kash, S. F., and Baekkeskov, S. (1999) *J. Biol. Chem.* **274**, 37200–37209
1141. Hsu, C.-C., Davis, K. M., Jin, H., Foos, T., Floor, E., Chen, W., Tyburski, J. B., Yang, C.-Y., Schloss, J. V., and Wu, J.-Y. (2000) *J. Biol. Chem.* **275**, 20822–20828
1142. Erlander, M. G., Tillakaratne, N. J. K., Feldblum, S., Patel, N., and Tobin, A. J. (1991) *Neuron* **7**, 91–100
1143. Qu, K., Martin, D. L., and Lawrence, C. E. (1998) *Protein Sci.* **7**, 1092–1105
1144. Porter, T. G., and Martin, D. L. (1988) *Biochim. Biophys. Acta.* **874**, 235–244
1145. During, M. J., Ryder, K. M., and Spencer, D. D. (1995) *Nature (London)* **376**, 174–177
1146. Rice, A., Rafiq, A., Shapiro, S. M., Jakoi, E. R., Coulter, D. A., and DeLorenzo, R. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9665–9669
1147. Dedek, K., Kunath, B., Kananura, C., Reuner, U., Jentsch, T. J., and Steinlein, O. K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12272–12277
1148. Brusa, R., Zimmermann, F., Koh, D.-S., Feldmeyer, D., Gass, P., Seeburg, P. H., and Sprengel, R. (1995) *Science* **270**, 1677–1680
1149. Erickson, J. C., Clegg, K. E., and Palmiter, R. D. (1996) *Nature (London)* **381**, 415–418
1150. Kim, E., Lowenson, J. D., Clarke, S., and Young, S. G. (1999) *J. Biol. Chem.* **274**, 20671–20678
1151. Mandelkow, E.-M., and Mandelkow, E. (1993) *Trends Biochem. Sci.* **18**, 480–483
1152. Huang, Y., Liu, X. Q., Wyss-Coray, T., Brecht, W. J., Sanan, D. A., and Mahley, R. W. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8838–8843
1153. Ashall, F., and Goate, A. M. (1994) *Trends Biochem. Sci.* **19**, 42–46
1154. Selkoe, D. J. (1996) *J. Biol. Chem.* **271**, 18295–18298
1155. von Bergen, M., Friedhoff, P., Biernat, J., Heberle, J., Mandelkow, E.-M., and Mandelkow, E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5129–5134
1155a. Götz, J., Chen, F., van Dorpe, J., and Nitsch, R. M. (2001) *Science* **293**, 1491–1495
1155b. Lewis, J., Dickson, D. W., Lin, W.-L., Chisholm, L., Corral, A., Jones, G., Yen, S.-H., Sahara, N., Skipper, L., Yager, D., Eckman, C., Hardy, J., Hutton, M., and McGowan, E. (2001) *Science* **293**, 1487–1491
1156. Stone, R. (1993) *Science* **261**, 424–426
1157. Spencer, P. S., Nunn, P. B., Hugon, J., Ludolph, A. C., Ross, S. M., Roy, D. N., and Robertson, R. C. (1987) *Science* **237**, 517–522
1158. Hardy, J., and Gwinn–Hardy, K. (1998) *Science* **282**, 1075–1079
1159. Selkoe, D. J. (1991) *Sci. Am.* **265**(Nov), 68–71
1160. Suzuki, T., Ando, K., Isohara, T., Oishi, M., Lim, G. S., Satoh, Y., Wasco, W., Tanzi, R. E., Nairn, A. C., Greengard, P., Gandy, S. E., and Kirino, Y. (1997) *Biochemistry* **36**, 4643–4649
1160a. Hardy, J., and Selkoe, D. J. (2002) *Science* **297**, 353–356
1161. El-Agnaf, O. M. A., Sheridan, J. M., Sidera, C., Siligardi, G., Hussain, R., Haris, P. I., and Austen, B. M. (2001) *Biochemistry* **40**, 3449–3457
1162. El-Agnaf, O. M. A., Nagala, S., Patel, B. P., and Austen, B. M. (2001) *J. Mol. Biol.* **310**, 157–168
1162a. Paushkin, S., Charroux, B., Abel, L., Perkinson, R. A., Pellizzoni, L., and Dreyfuss, G. (2000) *J. Biol. Chem.* **275**, 23841–23846
1162b. Chang, J.-G., Hsieh-Li, H.-M., Jong, Y.-J., Wang, N. M., Tsai, C.-H., and Li, H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9808–9813
1163. Min, K.-T., and Benzer, S. (1999) *Science* **284**, 1985–1988
1163a. Taylor, J. P., Hardy, J., and Fischbeck, K. H. (2002) *Science* **296**, 1991–1995
1164. Masliah, E., Rockenstein, E., Veinbergs, I., Mallory, M., Hashimoto, M., Takeda, A., Sagara, Y., Sisk, A., and Mucke, L. (2000) *Science* **287**, 1265–1269
1164a. Manning, B., AB, McCormack, A. L., Li, J., Uversky, V. N., Fink, A. L., and Di Monte, D. A. (2002) *J. Biol. Chem.* **277**, 1641–1644
1164b. Auluck, P. K., Chan, H. Y. E., Trojanowski, J. Q., Lee, V. M.-Y., and Bonini, N. M. (2002) *Science* **295**, 865–868
1165. Li, J., Uversky, V. N., and Fink, A. L. (2001) *Biochemistry* **40**, 11604–11613
1166. Sung, J. Y., Kim, J., Paik, S. R., Park, J. H., Ahn, Y. S., and Chung, K. C. (2001) *J. Biol. Chem.* **276**, 27441–27448
1167. Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., Yokochi, M., Mizuno, Y., and Shimizu, N. (1998) *Nature (London)* **392**, 605–608

References

1168. Kahle, P. J., Leimer, U., and Haass, C. (2000) *Trends Biochem. Sci.* **25**, 524–527
1169. Shimura, H., Schlossmacher, M. G., Hattori, N., Frosch, M. P., Trockenbacher, A., Schneider, R., Mizuno, Y., Kosik, K. S., and Selkoe, D. J. (2001) *Science* **293**, 263–269
1170. Trottier, Y., and Mandel, J. L. (2001) *Science* **293**, 445–446
1171. Zuccato, C., Ciammola, A., Rigamonti, D., Leavitt, B. R., Goffredo, D., Conti, L., MacDonald, M. E., Friedlander, R. M., Silani, V., Hayden, M. R., Timmusk, T., Sipione, S., and Cattaneo, E. (2001) *Science* **293**, 493–498
1172. Steffan, J. S., Kazantsev, A., Spasic-Boskovic, O., Greenwald, M., Zhu, Y.-Z., Gohler, H., Wanker, E. E., Bates, G. P., Housman, D. E., and Thompson, L. M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6763–6768
1173. Liu, Y. F., Deth, R. C., and Devys, D. (1997) *J. Biol. Chem.* **272**, 8121–8124
1173a. Freiman, R. N., and Tjian, R. (2002) *Science* **296**, 2149–2150
1174. Choi-Lundberg, D. L., Lin, Q., Chang, Y.-N., Chiang, Y. L., Hay, C. M., Mohajeri, H., Davidson, B. L., and Bohn, M. C. (1997) *Science* **275**, 838–841
1175. Olson, L. (2000) *Science* **290**, 721–724
1176. Brady, S. (1995) *Nature (London)* **375**, 12–13
1177. Liu, H., Zhu, H., Eggers, D. K., Nersissian, A. M., Faull, K. F., Goto, J. J., Ai, J., Sanders-Loehr, J., Gralla, E. B., and Valentine, J. S. (2000) *Biochemistry* **39**, 8125–8132
1178. Singh, R. J., Karoui, H., Gunther, M. R., Beckman, J. S., Mason, R. P., and Kalyanaraman, B. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6675–6680
1179. Muralikrishnan, D., and Mohanakumar, K. P. (1998) *FASEB J.* **12**, 905–912
1180. Mochizuki, H., Hayakawa, H., Migita, M., Shibata, M., Tanaka, R., Suzuki, A., Shimo-Nakanishi, Y., Urabe, T., Yamada, M., Tamayose, K., Shimada, T., Miura, M., and Mizuno, Y. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10918–10923
1181. Giasson, B. I., Duda, J. E., Murray, I. V. J., Chen, Q., Souza, J. M., Hurtig, H. I., Ischiropoulos, H., Trojanowski, J. Q., and Lee, V. M.-Y. (2000) *Science* **290**, 985–989
1182. Utton, M. A., Gibb, G. M., Burdett, I. D. J., Anderton, B. H., and Vandecandelaere, A. (2001) *J. Biol. Chem.* **276**, 34288–34297
1183. Scheuermann, S., Hambsch, B., Hesse, L., Stumm, J., Schmidt, C., Beher, D., Bayer, T. A., Beyreuther, K., and Multhaup, G. (2001) *J. Biol. Chem.* **276**, 33923–33929
1184. Minopoli, G., de Candia, P., Bonetti, A., Faraonio, R., Zambrano, N., and Russo, T. (2001) *J. Biol. Chem.* **276**, 6545–6550
1185. Bergsdorf, C., Paliga, K., Kreger, S., Masters, C. L., and Beyreuther, K. (2000) *J. Biol. Chem.* **275**, 2046–2056
1186. Iwata, N., Tsubuki, S., Takaki, Y., Shirotani, K., Lu, B., Gerard, N. P., Gerard, C., Hama, E., Lee, H.-J., and Saido, T. C. (2001) *Science* **292**, 1550–1552
1187. Selkoe, D. J. (1999) *Nature (London)* **399**, A23–A31
1188. Wolfe, M. S., and Haass, C. (2001) *J. Biol. Chem.* **276**, 5413–5416
1189. Hong, L., Koelsch, G., Lin, X., Wu, S., Terzyan, S., Ghosh, A. K., Zhang, X. C., and Tang, J. (2000) *Science* **290**, 150–153
1189a. Esler, W. P., and Wolfe, M. S. (2001) *Science* **293**, 1449–1454
1190. Haass, C., Hung, A. Y., Selkoe, D. J., and Teplow, D. B. (1994) *J. Biol. Chem.* **269**, 17741–17748
1191. Tomita, S., Kirino, Y., and Suzuki, T. (1998) *J. Biol. Chem.* **273**, 19304–19310
1192. Yan, R., and 14 other authors. (1999) *Nature (London)* **402**, 533–537
1193. Vassar, R., and 23 other authors. (1999) *Science* **286**, 735–741
1194. Creemers, J. W. M., Dominguez, D. I., Plets, E., Serneels, L., Taylor, N. A., Multhaup, G., Craessaerts, K., Annaert, W., and De Strooper, B. (2001) *J. Biol. Chem.* **276**, 4211–4217
1195. Pennisi, E. (1999) *Science* **286**, 650–651
1196. Steinhilb, M. L., Turner, R. S., and Gaut, J. R. (2001) *J. Biol. Chem.* **276**, 4476–4484
1197. Haass, C., and De Strooper, B. (1999) *Science* **286**, 916–919
1198. Zhang, L., Song, L., Terracina, G., Liu, Y., Pramanik, B., and Parker, E. (2001) *Biochemistry* **40**, 5049–5055
1199. Fagan, R., Swindells, M., Overington, J., and Weir, M. (2001) *Trends Biochem. Sci.* **26**, 213–214
1200. Xia, X., Qian, S., Soriano, S., Wu, Y., Fletcher, A. M., Wang, X.-J., Koo, E. H., Wu, X., and Zheng, H. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10863–10868
1200a. Esler, W. P., Kimberly, W. T., Ostaszewski, B. L., Ye, W., Diehl, T. S., Selkoe, D. J., and Wolfe, M. S. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2720–2725
1201. Hashimoto, Y., Niikura, T., Ito, Y., and Nishimoto, I. (2000) *J. Biol. Chem.* **275**, 34541–34551
1202. Varadarajan, S., Kanski, J., Aksenova, M., Lauderback, C., and Butterfield, D. A. (2001) *J. Am. Chem. Soc.* **123**, 5625–5631
1203. Evans, D. B., Rank, K. B., Bhattacharya, K., Thomsen, D. R., Gurney, M. E., and Sharma, S. K. (2000) *J. Biol. Chem.* **275**, 24977–24983
1204. del C. Alonso, A., Zaidi, T., Novak, M., Grundke-Iqbal, I., and Iqbal, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6923–6928
1205. Grégoire, C., Marco, S., Thimonier, J., Duplan, L., Laurine, E., Chauvin, J.-P., Michel, B., Peyrot, V., and Verdier, J.-M. (2001) *EMBO J.* **20**, 3313–3321
1206. Perl, D. P., and Brody, A. R. (1980) *Science* **208**, 297–299
1207. Zhou, Y., Morais-Cabral, J. H., Kaufman, A., and MacKinnon, R. (2001) *Nature (London)* **414**, 43–48
1208. Lam, Y. A., Pickart, C. M., Alban, A., Landon, M., Jamieson, C., Ramage, R., Mayer, R. J., and Layfield, R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9902–9906
1209. Marx, J. (2001) *Science* **293**, 2192–2194
1210. Coyle, J. T., and Puttfarcken, P. (1993) *Science* **262**, 689–695
1211. Rauk, A., Armstrong, D. A., and Fairlie, D. P. (2000) *J. Am. Chem. Soc.* **122**, 9761–9767
1212. Spillantini, M. G., Goedert, M., Crowther, R. A., Murrell, J. R., Farlow, M. R., and Ghetti, B. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4113–4118
1213. Pérez, M., Arrasate, M., de Garcini, E. M., Munoz, V., and Avila, J. (2001) *Biochemistry* **40**, 5983–5991
1213a. Alonso, AdC., Zaidi, T., Novak, M., Barra, H. S., Grundke-Iqbal, I., and Iqbal, K. (2001) *J. Biol. Chem.* **276**, 37967–37973
1214. Goldstein, L. S. B. (2001) *Science* **291**, 2102–2103
1215. Planel, E., Yasutake, K., Fujita, S. C., and Ishiguro, K. (2001) *J. Biol. Chem.* **276**, 34298–34306
1216. Lu, P.-J., Wulf, G., Zhou, X. Z., Davies, P., and Lu, K. P. (1999) *Nature (London)* **399**, 784–788
1217. Weisgraber, K. H., and Mahley, R. W. (1996) *FASEB J.* **10**, 1485–1494
1218. Reiman, E. M., Caselli, R. J., Chen, K., Alexander, G. E., Bandy, D., and Frost, J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3334–3339
1218a. Helmuth, L. (2002) *Science* **297**, 1260–1262
1219. Wolozin, B. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5371–5373
1220. Yip, C. M., Elton, E. A., Darabie, A. A., Morrison, M. R., and McLaurin, J. (2001) *J. Mol. Biol.* **311**, 723 – 734
1221. Farber, S. A., Slack, B. E., and Blusztajn, J. K. (2000) *FASEB J.* **14**, 2198 – 2206
1222. Abdel-Ghany, M., El-Sebae, A. K., and Shalloway, D. (1993) *J. Biol. Chem.* **268**, 11976–11981
1223. Yoshiike, Y., Tanemura, K., Murayama, O., Akagi, T., Murayama, M., Sato, S., Sun, X., Tanaka, N., and Takashima, A. (2001) *J. Biol. Chem.* **276**, 32293–32299
1224. Lee, VM-Y (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8931–8932
1225. Watanabe, C. M. H., Wolfram, S., Ader, P., Rimbach, G., Packer, L., Maguire, J. J., Schultz, P. G., and Gohil, K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6577–6580

Study Questions

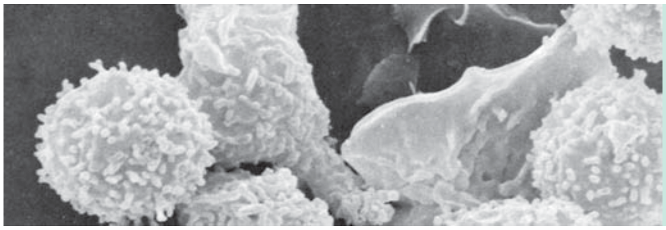
1. Compare the sensing of signals and responses to signals in liver, muscle, or other tissue with signaling in the brain.

2. Compare signaling by ionotropic receptors, metabotropic receptors and gap junctions.

3. List major neurotransmitters in the brain. In what other locations do these compounds function?

4. Compare addiction to gambling with addiction to cocaine. Are they similar on a biochemical basis? What about addiction to Internet games, chatting, pornography, compulsive overeating, etc.?[873a]

5. Can extracts of leaves of *Gingko biloba* counteract age-related neurological disorders?[1225]



The large flat cell, a portion of which is seen here, is a macrophage which has ingested bacterial proteins and is displaying peptide fragments on its surface. Some of the small, spherical T lymphocytes (T cells) seen here, interact with the macrophage, recognize an antigen, and respond by becoming helper T cells. They can then stimulate B lymphocytes (B cells) to multiply and produce antibodies. See Fig. 31-11 for an enlarged view. Scanning electron micrograph courtesy of Morton H. Nielsen and Ole Werdelin, University of Copenhagen.

Contents

1831 .... **A. Locations and Organization of the Immune System**
1833 .... 1. Development of Lymphocytes and Other Specialized Cells
1834 .... 2. Triggering an Immune Response
1835 .... **B. The Immunoglobulins (Antibodies)**
1835 .... 1. Molecular Structures
1839 .... 2. Antigenicity
1839 .... 3. Responses to Antibody Binding
1840 .... 4. Clonal Expansion of B Cells; Plasma Cells
1840 .... 5. Help from T Cells
1840 .... **C. Some Specialized Proteins of the Immune System**
1842 .... 1. Defensins and Other Antibacterial Polypeptides
1844 .... 2. Complement
1846 .... 3. Cytokines, Interferons, and the Acute-Phase Response
1847 .... *Interferons*
1849 .... *Inflammatory influences*
1850 .... **D. Organizing the Immune Response**
1850 .... 1. Coreceptors and the B-Cell Response
1850 .... 2. The Leukocyte Differentiation Antigens
1851 .... 3. Functions of T Cells
1852 .... 4. Natural Killer Cells
1852 .... 5. Identifying Self: The Major Histocompatibility Complex
1855 .... 6. Antigen Presentation and MHC Restriction
1856 .... 7. T-Cell Receptors
1858 .... 8. Self-Tolerance
1859 .... 9. Immunologic Memory and Vaccination
1859 .... **E. The Rearranging Genes for Immunoglobulins and T-Cell Receptors**
1859 .... 1. Rearrangements of Germline DNA
1861 .... 2. Somatic Hypermutation and Affinity Maturation
1862 .... 3. Immunoglobulin Class (Isotype) Switching
1862 .... **F. Disorders of the Immune System**
1862 .... 1. Allergy
1864 .... 2. Autoimmune Diseases
1867 .... 3. Immunodeficiencies
1868 .... 4. Cancers of the Immune System
1869 .... **G. Defense Mechanisms of Plants**

1870 .... **References**
1877 .... **Study Questions**

Boxes

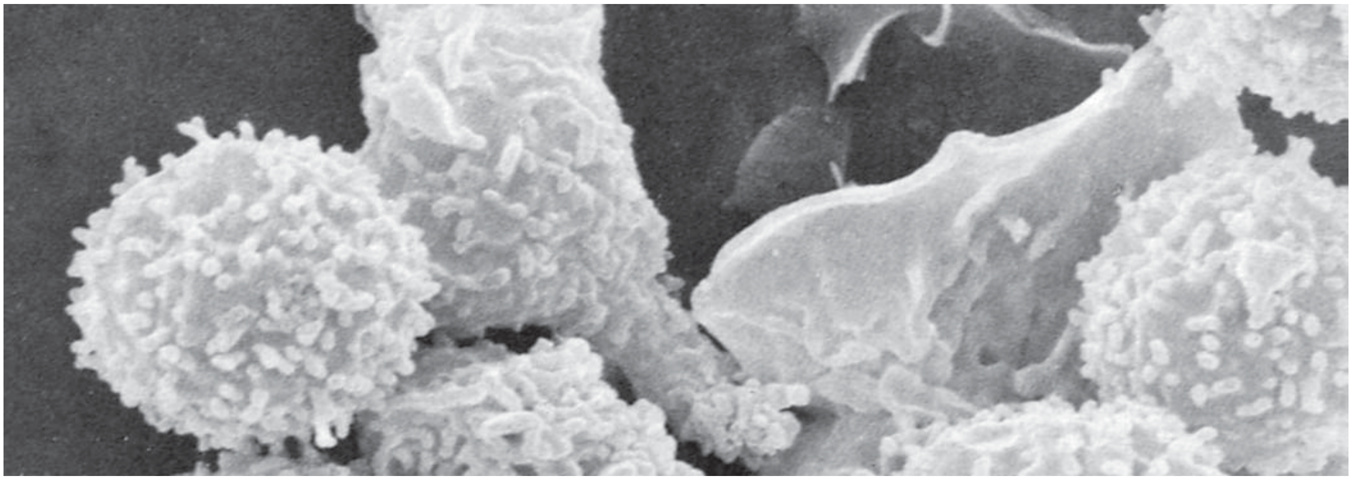
1841 .... Box 31-A Monoclonal Antibodies
1842 .... Box 31-B Catalytic Antibodies
1848 .... Box 31-C Immunoassays
1864 .... Box 31-D Myasthenia Gravis
1866 .... Box 31-E Evading the Immune System
1868 .... Box 31-F An Insecticidal Protein

Tables

1833 .... Table 31-1 The Two Major Branches of the Immune System
1834 .... Table 31-2 Cells of the Immune System
1865 .... Table 31-3 Some Autoimmune Diseases

Biochemical Defense Mechanisms

31



Our bodies are under constant attack by viruses, bacteria, protozoa, and metazoan parasites. Persons born without an immune system adequate to fight off these invaders die very quickly unless heroic measures are taken. We have learned to cooperate with our immune systems by immunizing ourselves against some bacteria and viruses. At other times we may fight a stubborn battle with our own defense systems against allergic reactions and a variety of autoimmune responses.[1]

A. Locations and Organization of the Immune System

The immune system has many components, many of which are dispersed throughout the human body (Fig. 31-1). Lymphocytes, which are a foundation of the immune system, constitute little more than 1% of the body's mass. However, this represents ~ 10^{12} cells of several types, about 10 times more than there are neurons in the brain. These 10^{12} cells make antibodies and T-cell receptors, both of which are thought to have ~ 10^{15} different peptide sequences.[2] And this is only the beginning of the complexities.[3–8]

Immune responses have often been described in terms of **humoral** and **cellular** components. The humoral response involves the small circulating **B lymphocytes** (B cells), the **antibodies** (immunoglobulins), and proteins of the **complement** system. The cellular response is mediated by another group of small lymphocytes, the **T lymphocytes** (T cells). They resemble B cells in appearance but have quite different functions. However, newer knowledge has provided a somewhat different description of the body's defense systems, which can be classified into three levels. (1) The skin and internal mucous membranes, which are resistant to infection and have antibacterial properties, provide the first level of defense.[2,9] (2) A fast-acting **innate immune system** can respond within a few minutes to breaches in the barriers provided by the tough outer skin and the glycoproteins of mucous surfaces and provides a second level.[2,10–14] (3) A slower **adaptive** (acquired) part of the immune system leads to synthesis of antibodies and to long-term immunity, providing the third level (Table 31-1). Both B and T lymphocytes together with **antigen-presenting cells** (**APCs**) are necessary for the selection and development of immunoglobulin structures appropriate for attack on an invading organism.

The innate immune system utilizes **phagocytic cells** including neutrophils, monocytes, and macrophages[15] to ingest and kill invading organisms. Basophils, mast cells,[16] eosinophils, and other cells release inflammatory mediators, which attract additional lymphocytes and affect their development.[17] Specialized T lymphocytes called **natural killer (NK) cells**[18] may also attack foreign cells (Table 31-2). The innate system is ancient and has apparently evolved to recognize molecular structures that are foreign to the host but are characteristic of pathogens. These structures, which are described as **pathogen-associated molecular patterns** (**PAMPs**), include those of lipopolysaccharides of bacterial cell walls (Fig. 8-30), mannans (p. 175),[19] other carbohydrates of surface layers,[20] oxidized phosphatidylcholines,[20a] bacterial flagellins,[21] various posttranscriptionally modified proteins, teichoic acids (p. 431), etc. However, they do not include patterns characteristic of host cells. Those have been avoided during evolution of the system. Although

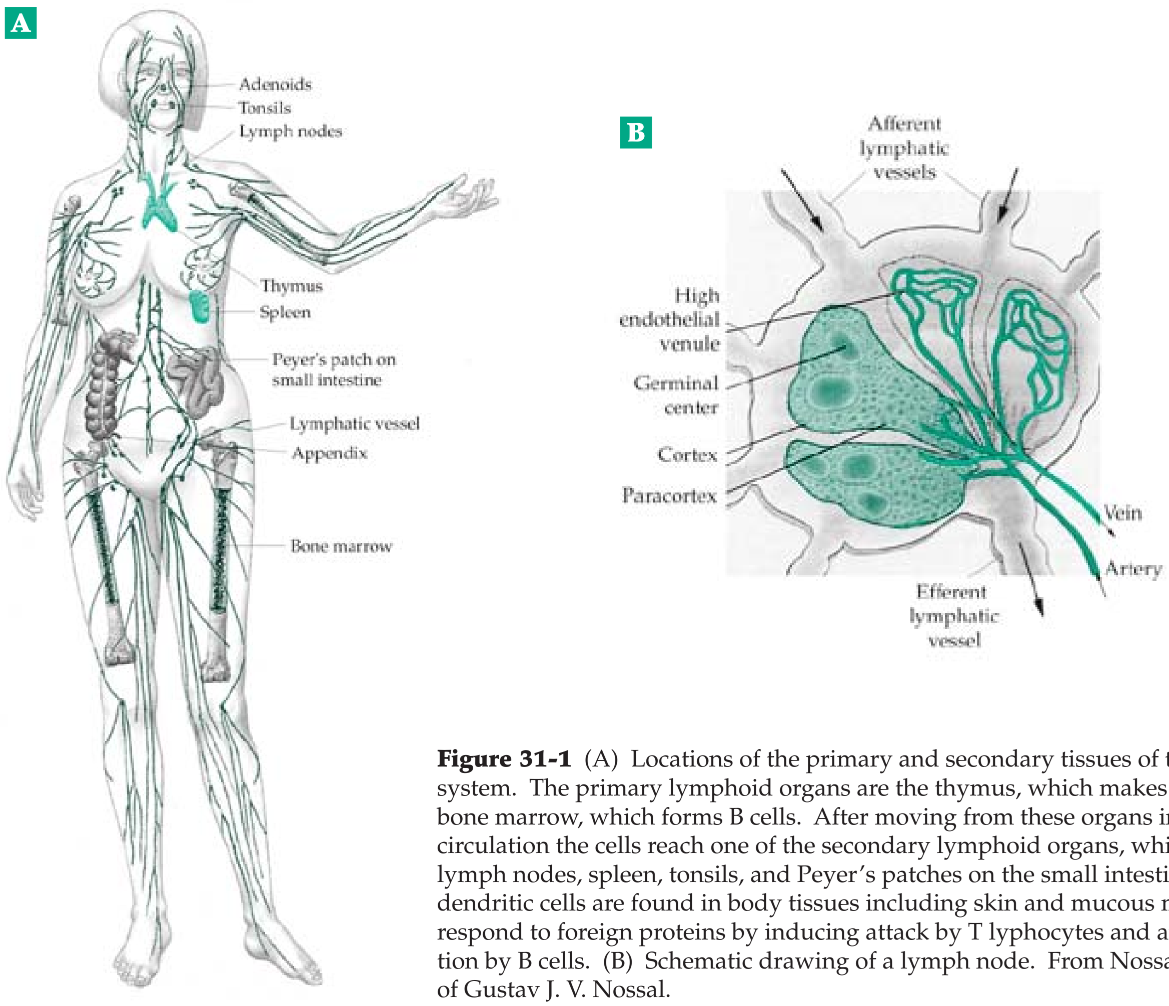


Figure 31-1 (A) Locations of the primary and secondary tissues of the immune system. The primary lymphoid organs are the thymus, which makes T cells, and the bone marrow, which forms B cells. After moving from these organs into the blood circulation the cells reach one of the secondary lymphoid organs, which include lymph nodes, spleen, tonsils, and Peyer's patches on the small intestine. Immature dendritic cells are found in body tissues including skin and mucous membranes and respond to foreign proteins by inducing attack by T lyphocytes and antibody formation by B cells. (B) Schematic drawing of a lymph node. From Nossal.[1] Courtesy of Gustav J. V. Nossal.

T lymphocytes are major mediators of the innate immune response they are under control of the **dendritic cells** (**DCs**), which are found in "immature" forms in tissues throughout the body.[9,22–26a]

The immature DCs are phagocytic cells that act as "immunological sensors." They recognize various PAMPs, which act as **danger signals**,[11b] using what are known as **toll-like receptors** (**TLRs**).[23,24,27,27a,b] They are also the most active APCs. Their proteasomes cleave proteins, both of the host and of invading organisms, into short peptides. These peptide fragments are displayed on the APC surfaces for recognition by T lymphocytes and for activation of adaptive immune responses. Some "autoreactive" B cells are also part of the innate immune system[28,29] as are the IgA antibodies present in mucous membranes.[30] The innate immune system of insects resembles that of vertebrates. The toll-like receptors of the latter are named for their resemblance to the Toll receptors of *Drosophila*, which are utilized in resistance to fungi. In both mammals and insects the innate immune system activates responses via the transcription factor NF-κB. However, many details of the signaling pathways differ.[30a–c]

The innate system also provides for synthesis of small antibiotic peptides called **defensins**[12,31–33] as well as larger proteins. Some of these proteins constitute the **complement system**, while others are described as **acute-phase reactants**. Some defensins are also **cytokines**, which attract lymphocytes.

The innate system is of special importance during early infancy. Prior to birth and for at least 4–12 months after birth a child's immune system is poorly developed. It may not become fully competent until age ~5.[34,35] During the prenatal period maternal antibodies are transferred to the child. IgG crosses the placenta and enters the fetal circulation. Breast milk provides IgA, which remains largely in the child's gut, as well as other protective proteins. UNICEF and the World Health Organization recommend breast-feeding to two years or beyond.[34]

While the innate immune system provides for immediate and direct attack on invaders, it also provides information to the slower adaptive system. Genes both for immunoglobulins and for the T-cell receptors of the adaptive system undergo extensive rearrangement during development of an individual.

TABLE 31-1
The Two Major Branches of the Immune System

	Innate (natural)	Adaptive (acquired)
Cells	Dendritic cells Phagocytic cells (neutrophils, monocytes, macrophages) Cells that release inflammatory mediators (basophils, mast cells, eosinophils) Natural killer (NK) cells	Dendritic cells B lymphocytes (B cells) T lymphocytes (T cells) Other antigen-presenting cells, e.g., macrophages
Molecular components	Antibacterial peptides (defensins, complement, acute-phase proteins)	Immunoglobulins
Receptor genes	Fixed in genome	Complement proteins Encoded in gene segments; rearrangement necessary
Recognition	Conserved molecular patterns	Small molecular groups (epitopes)
Immunogenic memory	Absent	Present
Self–nonself discrimination	Perfect	Imperfect
Action time	Immediate	Delayed

This provides potential defensive proteins directed at almost every imaginable invader. It also ensures that every individual has a set of proteins that labels its own cells as "self," and that virtually every individual on earth has cell surface proteins different from those of every other person. In both the innate and adaptive responses the immune system must carefully distinguish "self" from "nonself."[36,37] In the innate system this discrimination developed during evolution of the host and its pathogens. In the adaptive system it depends upon interaction of the T cells with surface molecules, primarily those of the **major histocompatibility complex** (MHC).

Another basic characteristic of immune responses is the development of **immunologic memory**.[38–40] This is exemplified by the fact that vaccination can sometimes impart immunity for a person's lifetime. If a foreign protein is injected into an animal, after a lag period of 2–5 days the animal will synthesize antibodies against this foreign antigen. This is called a **primary adaptive immune response**. If after a few days or weeks a second injection of the same protein is made, a much more rapid synthesis of additional antibodies occurs. This **secondary immune response** may take place within hours and will last longer than the primary response. It is a manifestation of immunologic memory.

1. Development of Lymphocytes and Other Specialized Cells

Both the B cells and T cells arise in the fetal liver or bone marrow (Fig. 31-1) from pluripotent stem cells. In birds the B cells develop in a special organ, the bursa of Fabricius. Mammalian B cells complete their differentiation into mature **lymphocytes** within the bone marrow. However, the T cells must travel to the **thymus,** where they complete their maturation. The T lymphocytes include the previously mentioned NK cells as well as the somewhat similar **cytolytic T cells** and **immunoregulatory T cells**. The latter are further characterized as **helper T cells**[41] or **suppressor T cells**. The adaptive response requires cooperation of helper T cells in many instances. The mature B and T cells leave the bone marrow and thymus, which are known as the **primary lymphoid tissues,** and enter the blood circulation. Following "homing" signals[42] they take up residence in a variety of locations in the lymph nodes, spleen, adenoids, tonsils, and Peyer's patches. The last are small clusters of lymphoid cells in the wall of the intestine. All of these tissues, which are referred to as **secondary lymphoid tissues**, are the sites in which the adaptive immune system is developed.

TABLE 31-2
Cells of the Immune System[a]

Type	Functions
B lymphocytes	
Plasma cells	Antibody synthesis
Memory B cells	Immunologic memory
T lymphocytes	
Cytolytic	Destroy infected and malignant cells
Helper cells	
Type 1 (T_H1)	Participate in activation of B cells
Type 2 (T_H2)	
Memory T cells	Immunologic memory
Natural killer (NK) cells	Destroy infected and malignant cells with pore-forming protein perforin and cytotoxic granules
Dendritic cells	
Interdigitating	Antigen recognition and processing
Follicular	Antigen presentation
Microglia	Defensive network in brain[b]
Cells that release inflammatory mediators	
Mast cells and basophils	Posses high-affinity receptors for IgE May secrete histamine, prostaglandins, leukotrienes Important in allergies
Eosinophils	Weakly phagocytic, secrete cationic proteins, reactive reduced oxygen species, leukotrienes, prostaglandins, cytokinins
Phagocytic cells	
Neutrophils	Acute inflammatory response
Monocytes	
Macrophages	Carry receptors for carbohydrates not normally exposed on surfaces of cells in vertebrates e.g., mannose; kill engulfed organisms with $^\bullet O_2^-$, HOCl, NO, cationic proteins and peptides, lysozyme Antigen processing and presentation

[a] General reference: Delves, P. J., and Roitt, I. M. (2000) *N. Engl. J. Med.* **343**, 37–49.

[b] Streit, W. J., and Kincaid-Colton, C. A. (1995) *Sci. Am.* **273** (Nov), 54–61.

An important component of the immune system that was neglected until recently is located in the mucous membranes and the skin.[9,30,43,43a] The mucosal surfaces of airways and the gastrointestinal tract provide the point of entry for many diseases. Both internal and external body surfaces are protected by dendritic cells, whose immature forms in skin are called Langerhans cells. See figure in Box 8-F.[22,26a]

Eight different types of cells of the immune system (Table 31-2) develop by differentiation of pluripotent stem cells[44–48] as indicated in Fig. 31-2. Dendritic cells[26a] (which are not shown in this figure) may be formed from monocytes but may also arise by other routes.[22] The development of the various cells takes place under the influence of a number of **hemopoietic regulators**. Among these are the protein hormone **erythropoietin** and various interleukins and colony-stimulating factors.[44,45,49]

2. Triggering an Immune Response

When a foreign antigen enters the body the B cells, with receptors of appropriate specificity and present in the lymph nodes, are stimulated to divide repeatedly and to produce a large clone of **plasma cells**. These contain a highly developed ER and actively synthesize and secrete immunoglobulins. One activated B cell may produce 10 million antibody molecules per hour.[1] This B-cell response occurs within a network of **follicular dendritic cells** in the **germinal centers** of the lymph nodes (Fig. 31-1B).[2,50,51] The antibodies are "adaptor" molecules, which bind to antigenic proteins often on surfaces of invading microorganisms. Another part of the antibody binds to one of several **effectors** systems. These immobilize microorganisms, induce phagocytosis, activate the complement system, carry antibodies across placental membranes, etc.[52,53]

Induction of a T-cell response is more complex and is very demanding.[22,47,54] Antigenic peptide frag-

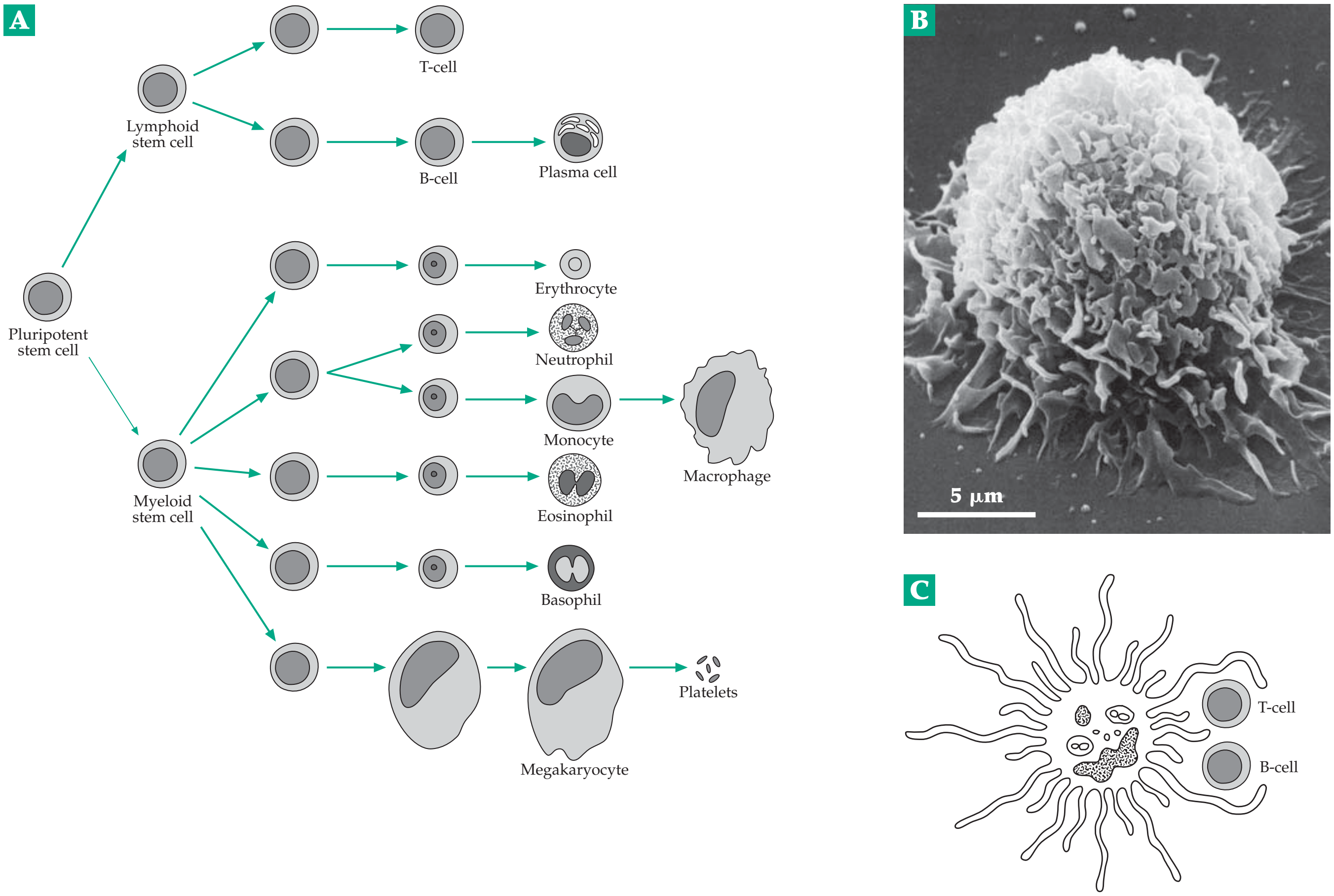


Figure 31-2 (A) Development of eight types of blood cells including those of the immune system from pluripotent (multipotential) stem cells. The cells develop under the influence of a variety of protein growth factors. Some steps, e.g., maturation of B cells, involve complex rearrangements in the DNA of the cell. (B) Scanning electron micrograph of a macrophage, a large motile cell that plays a key role in the immune system. It moves by means of its surface "ruffles." It actively phagocytoses both pathogens and waste materials and is also one of the cells that releases the hormones known as colony-stimulating factors. Micrograph courtesy of Shirley G. Quam. Drawings courtesy of David W. Golde and Judith C. Gasson.[44] (C) Schematic drawing of a dendritic cell. Redrawn from Banchereau and Steinman.[22] Both macrophages and dendritic cells present antigens for recognition by T cells and synthesize cytokines, which affect lymphocyte development.

ments from infected or malignant cells anywhere in the body must be recognized by an appropriate receptor on a T cell that is circulating in the blood. Only a few T cells with receptors for any given antigen specificity exist.[54,55] The foreign antigen fragment must bind to a protein of the **major histocompatibility complex** (**MHC**) while present within a dendritic cell or other APC. The resulting MHC•antigen complexes pass through the ER and the Golgi to the outer cell surface in a rather complex process. Recognition of the antigen that is "presented" in this manner on the APC surface is accomplished with the aid of ~10^{15} different receptor proteins (**T-cell receptors**) on the T-cell surfaces. When a T-cell receptor is occupied by an HMC•antigen complex of appropriate specificity the T cell is activated to participate in adaptive immunity. However, some T cells, notably γδ T cells, like antibodies bind to antigen directly.[56,56a,b] The recognition process occurs in an **immunological synapse**, which has elements of similarity to neurological synapses.[54,57–58a]

B. The Immunoglobulins (Antibodies)

1. Molecular Structures

There are five classes of antibodies or immunoglobulins.[59,60] The first three, IgG, IgM, and IgA, are quantitatively the most significant, but IgD and IgE are also important. For example, the content of IgE is elevated in allergic persons. The basic structure of all of the immunoglobulins is that of a quasi-symmetric dimer composed of a pair of light chains and a pair of heavy chains whose lengths vary among the different

Symbol	Mass (kDa)	Formula
IgG	150	$\kappa_2\gamma_2$ or $\lambda_2\gamma_2$
IgM	950	$(\kappa_2\mu_2)_{2,5}$ or $(\lambda_2\mu_2)_{2,5}$
IgA	320	$(\kappa_2\alpha_2)_n$ or $(\lambda_2\alpha_2)_n$
IgD	180	$\kappa_2\delta_2$ or $\lambda_2\delta_2$
IgE	190	$\kappa_2\varepsilon_2$ or $\lambda_2\varepsilon_2$

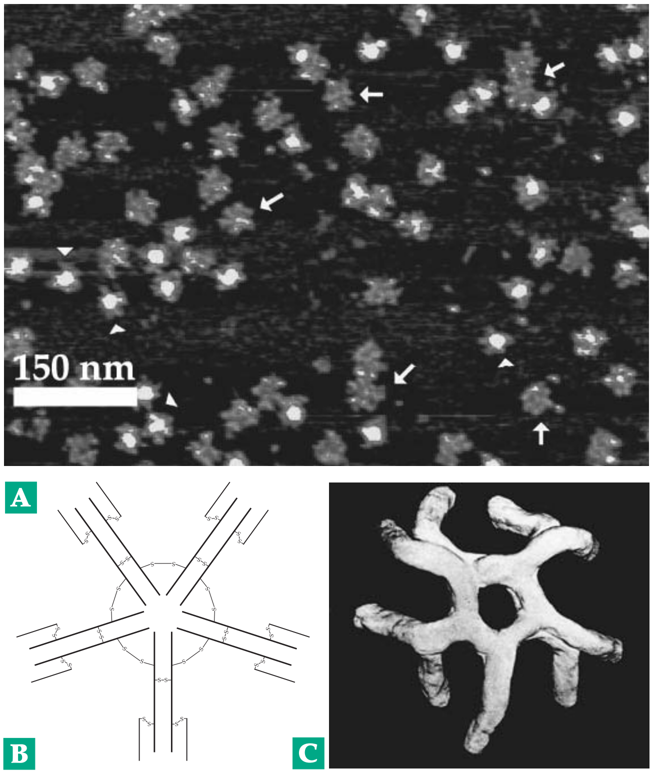


Figure 31-3 (A) Cryo atomic force (AFM) micrograph of molecules of the human immunoglobulin IgM. Courtesy of Zhifeng Shao, University of Virginia. (B) Schematic diagram. One-fifth of this structure is shown in greater detail in Fig. 31-4A. (C) Model based on earlier electron microscopic images. From Feinstein and Munn. [64c]

classes of immunoglobulins. Two classes of **light chains,** κ and λ, are found in human antibodies. The **heavy chains** are designated γ, μ, α, δ, and ε (see accompanying tabulation). Both IgM and IgA contain an additional J chain.

Treatment with mercaptoethanol splits the disulfide linkages holding the chains together, permitting preparation of monomeric light and heavy chains. When peptide chains of the immunoglobulins were hydrolyzed enzymatically, the resulting peptide fragments were found to be extremely heterogeneous. They were mixtures of many different kinds of peptides. The result was not unexpected, for it had long been recognized that the body contains millions of different antibodies, with binding sites specific for different antigenic determinants. It had been unclear how different binding sites could be formed, but the heterogeneity in amino acid sequence suggested the correct answer: Each antibody has its own sequence.

Progress toward understanding of the detailed structure of antibodies came when it was recognized that patients with tumors of the lymphatic system, e.g., the bone marrow tumors **multiple myeloma**, produced tremendous quantities of homogeneous immunoglobulins or parts thereof. Similar tumors were soon discovered in mice and provide a ready source of experimental material. The **Bence-Jones proteins** that are secreted in the urine of myeloma patients were found to be light chains of immunoglobulins. Sequence determinations showed that each Bence-Jones protein was homogeneous, even though no two patients secreted the same protein.[61,62] Later, intact myeloma globulins and macroglobulins (IgM) of a homogeneous kind were also obtained.

The first complete amino acid sequence of an IgG molecule was announced in 1969.[63] The protein contained 446 amino acids in each heavy chain and 214 in each light chain. The longer heavy chains of IgM molecules contain 576 amino acids.[64] In all of the immunoglobulins the heavy and light chains are held together by disulfide linkages, and the chains are folded into loops to form compact domains. The IgM molecule is polymerized through additional disulfide linkages to form a pentamer readily visible with the electron microscope (Fig. 31-3). The heavy chains also carry oligosaccharide units. In IgM there are five of these, as indicated in Fig. 31-4A. They contain mannose and *N*-acetylglucosamine units linked to asparagine. Other immunoglobulins (IgA, IgE, and IgG) contain fucose, galactose, and *N*-acetylneuraminic acid as well. In fact, almost all of the most important macromolecules that participate in innate and adaptive immune responses are glycoproteins.[64b]

Digestion of an intact molecule of IgG with papain cleaves both heavy chains in the hinge region near the interchain disulfide bridge. This splits the molecule into three parts; two **Fab** (antibody-binding) **fragments**, each containing the N-terminal end of a heavy chain together with a linked light chain, and an **Fc fragment**. Even before it was known that IgG could be split into two Fab fragments, the antibody was known to be divalent, i.e., capable of binding with two different antigens (Fig. 31-4). The shape and overall structure of IgG molecules have been verified by electron microscopy and numerous X-ray diffraction studies.

Sequence determinations showed that in some regions of immunoglobulin molecules there is extreme variation in the amino acid sequence between one homogeneous antibody and the next; other regions have a constant sequence. The molecule can also be divided into domains. The **variable regions**, which occupy the N-terminal ends of the chains, are designated

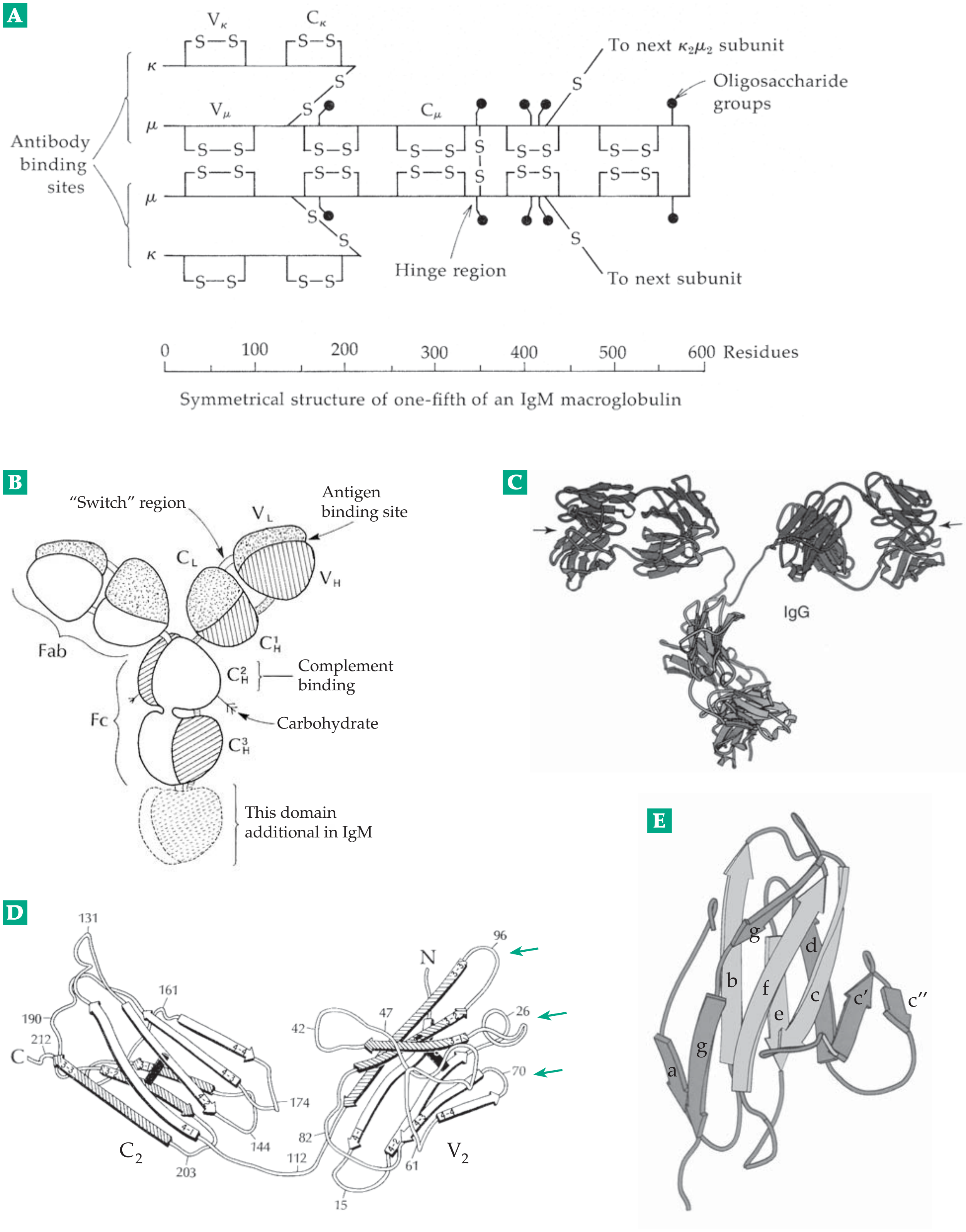

A
Vκ
Cκ
κ
μ
Vμ
Cμ
S—S
To next κ₂μ₂ subunit
Oligosaccharide groups
Antibody binding sites
Hinge region
To next subunit
0 100 200 300 400 500 600 Residues
Symmetrical structure of one-fifth of an IgM macroglobulin
B
"Switch" region
Antigen binding site
VL
VH
CL
C¹H
C²H
C³H
Complement binding
Carbohydrate
Fab
Fc
This domain additional in IgM
C
IgG
D
131
161
190
212
C
C2
203
144
174
112
82
15
61
V2
N
96
47
42
26
70
E
a
b
c
c′
c″
d
e
f
g

Figure 31-4 Schematic structure of one-fifth of an IgM molecule. From Putnam *et al.*[64] (A) Covalent structure. (B) Schematic three-dimensional representation. (C) Ribbon diagram of an IgG molecule. From Cochran *et al.*[64a] (D) Folding patterns of one chain in a constant and a variable domain of a Bence-Jones protein. From Schiffer *et al.*[66] Green arrows indicate hypervariable regions. (E) MolScript drawing of the common core structure of Ig-like domains. The lighter shaded strands (b, c, e, f) form the core common to all Ig-like domains, which is surrounded by structurally more varied additional strands (darker). The front sheet has up to five strands (a, f, c, e, c″) and the back sheet up to four (a, b, e, d). Strand c″ is very flexible and is not always a part of the β sheet. From Bork, Holm, and Sander.[65] See also Fig. 2-16.

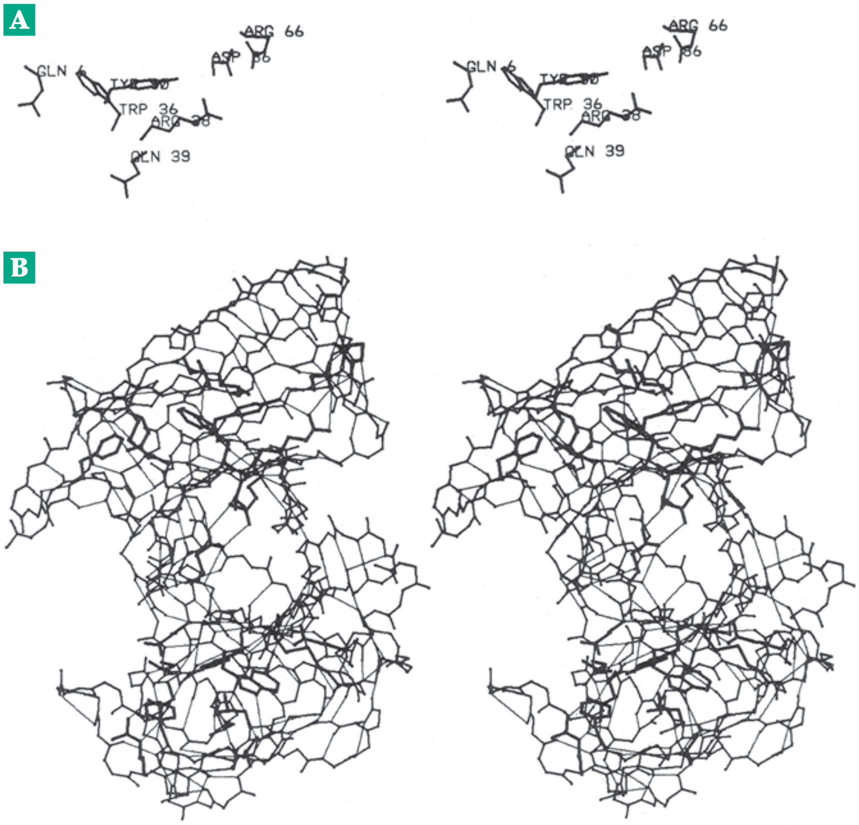


Figure 31-5 The extensive conserved hydrogen-bonding pattern in an immunoglobulin variable domain provided by polar residues buried inside the V_L and V_H domains. (A) To facilitate orientation, prominent side chains are displayed here and identified by names and numbers in the same orientation as in (B). (B) Polypeptide chain backbones of both domains are denoted by heavy lines and hydrogen bonds by light lines. In addition to the regular interbackbone hydrogen-bonding network characteristic of antiparallel β-sheets, there are hydrogen bonds provided by side-chain atoms. Note the two hydrogen bonds of Gln-38 (V_L) and Gln-39 (V_H) that span the domain–domain interface. From Novotny and Haber.[75]

V_L and V_H for the light and heavy chains, respectively. The **constant regions** are C_L and C_H. Examination of the C_H region showed that much of the sequence is repeated after ~110 residues. In the IgG molecule the constant region of the heavy chains is made up of three such homologous domains (C_H1, C_H2, and C_H3). A fourth C_H domain is present in IgM. These facts suggest that duplication of a smaller gene coding for about 110 amino acids took place in the evolution of the immunoglobulins. Within the variable regions of immunoglobulin chains are **hypervariable regions** that form the antigen binding sites. These regions are located at the ends of the Fab fragments and involve both the light and heavy chains.[66a]

Within all of the domains each of the two peptide chains is folded in a similar way. Seven extended lengths of chain form two mostly antiparallel sheets between which hydrophobic side chains are packed. The overall size of the unit is ~4.0 x 2.5 x 2.5 nm. An S–S bridge links the two sheets in the center of each domain. The folding patterns in the variable domains are somewhat more complex. Different domains are linked by segments of extended peptide chain known as hinge or switch regions.[67–69] These impart a segmental mobility, which seems to be important for functioning of the molecules.[70]

The exact mode of binding to Fab fragments has been established for several specific **haptens**. Haptens are small molecules having the binding properties of antigenic determinants but unable by themselves to induce formation of antibodies when injected into animals. Binding of the hapten phosphocholine to one Fab fragment and of vitamin K to another[67] involves the hypervariable regions of both the heavy and light chains. The same is true for the binding of lysozyme[71–73] and a bacterial oligosaccharide.[74] The binding sites for the haptens or for the antigenic determinants on larger antigens are largely within the nine-stranded elliptical barrel or β sandwich formed from the two β sheets (Fig. 31-5).[69,75] Four strands come from the V_L domain and five from the V_H domain. The barrel forms the bottom and sides of the antibody-binding site, which can also be viewed as consisting of six separate loops of peptide chain.[76] IgA molecules are similar to those of IgG but have structurally different hinge regions as well as an extra 18-residue tailpiece at the C terminus of each heavy chain.[77]

As shown in Fig. 31-5, which provides a three-dimensional view of the variable domain of an Fab fragment of an immunoglobulin, there is a conserved hydrogen-bonded network even in this region. There are also "framework residues," which are highly conserved.[78] The antibody-binding site is provided largely by the three hypervariable regions present in each of the V_L and V_H domains. These are usually referred to in current literature as **complementarity-determining regions** (**CDRs**).[78–81] Each pair of heavy and light chains is held together by a conserved disulfide bridge.[82] Three-dimensional structures of a substantial number of different Fab fragments have provided precise knowledge about the antibody-binding cavities and about the forces involved in binding.[78,79,81,83] Among the established structures are those of Fab fragments specific for the following antigens: the haptens *p*-azophenylarsonate[84] and phencyclidine (p. 1798),[85] a sweet-tasting hapten,[86] triple-stranded DNA,[87] a DNA photoproduct,[88] creatine kinase,[89] staphylococcal nuclease,[90] an HIV capsid protein,[91] and an EGF receptor.[92] Structures are also known for single-domain antibodies from camels and llamas. These antibodies are naturally lacking in heavy chains but have single chains that fold back to mimic two-chain Fab fragments.[93,94] Similar single-chain antibody domains have also been created artifically.[95]

Not all proteins bind to antibodies in the usual binding cavity. **Protein G**, a cell surface protein from *Streptococcus* bonds to IgG molecules from many different species. Its binding site is on the outer surface of the heavy chain C_H1 domain.[96]

2. Antigenicity

Antibodies often bind haptens or complete antigens very tightly. The association constants K_f observed for monoclonal antibodies (Box 31-A) range from 10^6 to 10^{12} M^{-1}.[97] However, most natural antibodies have a lower affinity for their antigens. When protein antigens are denatured, the binding constants often decrease by 10^{-4} to 10^{-5}. This suggested that only antigenic determinants of relatively rigid structures serve as good antigens. However, when the reaction of antibodies with proteins of well-established three-dimensional structure were studied, it was found that the best antigenic determinants are those with some segmental mobility.[98,99] Furthermore, while some small peptides are good antigens, peptides are most highly antigenic when they can readily fold into a bend or other definite conformation.[100] Good antigenicity apparently requires some segmental flexibility as well as a definite conformation for the antigenic determinant.

3. Responses to Antibody Binding

Both B cells and T cells circulate throughout the body, spending only about 30 min during each cycle. They may meet and bind to an antigen in one of several different places.[50] Lymphocytes, which encounter blood-borne pathogens, usually initiate an immune response in the spleen. Responses to microorganisms in tissues are usually generated in lymph nodes. Ingested pathogens activate lymphocytes in specialized epithelial **microfold** (**M**) **cells** from which the antigen is transported to the Peyer's patches. Responses to inhaled or intranasal pathogens arise in the tonsils and adenoids. In every case one major aspect of the immune response results from binding of antibodies to antigens.

Antibodies by themselves do not destroy bacteria or viruses, but they induce responses that do. One immediate effect of antibodies is to remove offending materials or cells from circulation. When multivalent antibodies each combine with two different cells **agglutination** occurs. The agglutinated cells or multicellular organisms can then be destroyed by phagocytes. The coating of a cell surface by IgG is one form of a process called **opsonization**, a process that marks the cell as foreign and a target for phagocytosis.[3,53] Antibody–antigen interactions trigger several other responses as well. One of these results from the binding of protein **C1q**, a component of complement. Complement consists of a series of blood proteins that is poised to respond and to *complement* the action of antibodies in a variety of ways that are described in Section C,2 (see also Figs. 31-8 and 31-9). It has been established that it is the C_H2 domain of the Fc region of IgG that binds to C1q.[101] The binding occurs only after antigen (but not a small hapten) binds to the immunoglobulin.

Complement C1q is only one of several types of **Fc receptor**.[53,102,103] Others are involved in antigenic stimulation of B and T lymphocytes, macrophages, polymorphonuclear lymphocytes, and mast cells. Binding of the antibody–antigen complex to the receptors on phagocytic cells induces phagocytosis and release of oxygen metabolites, leukotrienes, prostaglandins, and other mediators of inflammation. The Fc domain mediates the uptake of antibodies from the mother's milk by young rats.[104,105] It also is the binding site of antibodies to **protein A**, a constituent of the cell wall of *Staphylococcus aureus*,[106] which is also widely used as a tool in immunological studies (Box 31-C). The neonatal Fc receptor, which is related structurally to Class I MHC antigens (Section D,5), is one of three major types of Fc receptor. The other two are the receptors for Fcγ (of IgG) and Fcε (of IgE). They (like their ligands) are members of the immunoglobulin superfamily. An exception is FcεRII (also called CD23), which resembles a C-type lectin. Some Fc receptors, e.g., FcγRI (CD64) and FcεRI, have a high affinity for their ligands with K_d ~10^{-8} to 10^{-10} M. Others, such as FcγRII (CD32) and FcγRIII (CD16), have lower affinities with K_d ~10^{-5} to 10^{-7} M.[103] Three-dimensional structures of several Fc receptor fragments, some in complexes with Fc fragments (Fig. 31-6), are known.[53,103,107–109] These include both IgG and IgE receptors.

It may be worthwhile to recall that many quite different proteins are members of the immunoglobulin structural family (Fig. 2-16). These include proteins

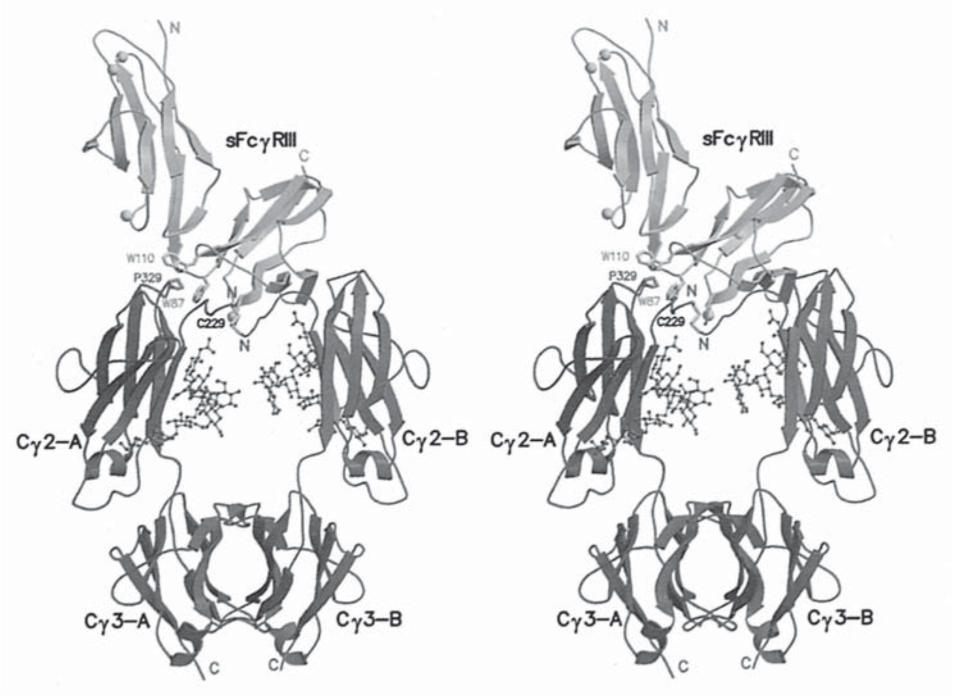


Figure 31-6 Three-dimensional ribbon representation of the structure of a complex of a soluble Fc fragment of a human IgG1 molecule. Pro 329 of the IgG and Trp 87 and Trp 110 of the Fc-receptor fragment form a "proline sandwich," which is shown in ball-and-stick form. The oligosaccharide attached to the Fc fragment of the antibody and the disulfide bridge between the two Cys 229 residues (at the N termini of the C2 domains of the heavy γ chains) are also shown. The small spheres on the Fc receptor fragment are potential sites for *N*-glycosylation. From Sondermann *et al.*[107] Courtesy of Uwe Jacob.

encoded by 64 genes of *E. elegans*, an organism that doesn't form antibodies.[110]

4. Clonal Expansion of B Cells; Plasma Cells

The immunoglobulins are synthesized both by plasma cells and by their precursors, the B cells. Each B lymphocyte makes antibodies of specific sequence in two forms, secreted and membrane-bound or **antigen receptor**. Mature "virgin" B cells, which are responsible for the primary immune response, make largely monomeric IgM and some IgD.[111] It isn't clear why IgD should be the predominant surface immunoglobulin on most B cells.[112] The secreted and membrane-bound (receptor) antibodies differ in their C-terminal sequences but are otherwise the same. A B cell responds to the binding of an antigen with a shape complementary to that of its IgM antigen receptor by multiplying and differentiating. Some of the progeny B cells start to divide and begin to differentiate into clones of plasma cells that secrete IgG or into cells of the gut that secrete IgA. Some of the B cells give rise to **memory cells**, long-lived lymphocytes that can be triggered into rapid proliferation many years later if the same antigen is encountered. The B cells also undergo a shift to secretion of pentameric IgM rather than to synthesis of membrane-bound antibody.

5. Help from T Cells

The maturation of B cells is a complex process that requires the cooperation of helper T cells. The B cells must process some antigen and present the peptide fragments for recognition by the MHC complex and T-cell receptors as described in Section D. If the antigen is recognized as foreign, the differentiation goes forward with the activated T cells secreting lymphokines that promote B-cell growth.

Before activated lymphocytes start to divide, interesting surface phenomena occur. If fluorescent antigens are allowed to bind to a lymphocyte, the cell surface is seen to be relatively evenly covered with the antibody–antigen complexes. Then after a short time the antibodies aggregate to form "patches" and begin to migrate to one side of the cell, where they eventually form a "cap." At still longer times the cap material is engulfed by the lymphocytes. Perhaps this phenomenon simply reflects the oriented flow of liquid within membranes (Chapter 8). On the other hand, the membrane-bound immunoglobulins, like other cell surface receptors, are integral membrane proteins whose cytoplasmic C termini may be attached on the inside of the cell to the cytoskeleton, which may control the capping process. The binding of lectins sometimes triggers lymphocytes into antibody synthesis, but it is not clear how the binding of a lectin to a carbohydrate receptor can have the same effect as binding of an antigen to a surface IgM.[113]

C. Some Specialized Proteins of the Immune System

The immune response depends not only upon recognition of foreign antigens but also upon an extensive signaling network and upon a series of specialized

BOX 31-A MONOCLONAL ANTIBODIES

A mouse may make over 10 million different antibodies. Because of this heterogeneity it was impossible to learn antibody structures until the discovery of the myeloma proteins (Fig. 31-4D). These were produced in the bone marrow by clones of specific immunoglobulin-forming malignant cells. However, it was still not possible to obtain homogeneous antibodies to any desired antigenic determinant. The discovery of a method of forming such **monoclonal antibodies** by Milstein and Köhler[a–c] in 1975 provided a new tool with many biochemical and medical applications.[d–f] What Milstein and Kohler did was to immunize mice against an antigen of interest. They then fused B cells from the spleen of the immunized mouse with cultured myeloma cells. The resulting **hybridomas** grow vigorously and produce antibodies of the type dictated by the B cells. Since each hybridoma cell is derived from a single B cell, it makes a single kind of antibody. By plating out and selecting clones of hybridoma cells it is often possible to find a monoclonal antibody that binds well to a specific antigenic determinant. The hybridoma can be cultured indefinitely, producing its monoclonal antibody in any desired quantity.

A major application of monoclonal antibodies is in clinical assays for drugs, bacterial and viral products, tumor antigens, hormones, and other circulating proteins. Their use in conjunction with immunoassays (Box 31-C) has provided increased specificity and sensitivity. Another major application is to observe binding of antibodies to specific proteins by electron microscopy. The location of specific receptor proteins can be established[g–j] as can the locations of ribosomal proteins and many other cellular components (Fig. 29-1). Monoclonal antibodies to acetylcholine receptors have been shown to induce symptoms of myasthenia gravis (Box 31-D), supporting the autoimmune origin of that disease.[h] Monoclonal antibodies specific for such a small hapten as mercuric ion have been isolated.[k]

Several problems have limited the wider use of monoclonal antibodies created by the hydridoma method. The antibodies are those of a mouse and are antigenic to humans.[f,l–n] This long prevented many medicinal uses. Years of effort have gone into attempts to “humanize” the antibodies. One approach is to introduce human immunoglobin genes into mice. Another is to use recombinant DNA techniques to clone genes for immunoglobulin fragments and to introduce these into cells of *E. coli* in which additional genetic diversity in the antibodies arises.[l,o] Selection of antibody fragments is often accomplished using bacteriophage display systems (Fig. 3-16).[f,l] After selection gene fragments can be reassembled into a final form. Recently, using cloning of large pieces of the several Mbp of human immunoglobin genes into yeast artificial chromosomes (p. 1497), it has been possible to prepare purely human monoclonal antibodies.[f,n]

Many attempts have been made to link monoclonal antibodies specific for antigenic determinants on cancer cells to protein toxins such as ricin (Box 29-A). It is hoped that this may provide an effective way of carrying toxins into cancer cells.[f,p–r] Therapeutic human monoclonal antibodies are already in use as antirejection drugs for kidney transplantation, for treatment of rheumatoid arthritis, Crohn disease, and for some types of cancer.[f]

---

[a] Milstein, C. (1980) *Sci. Am.* **243**(Oct), 66–74

[b] Milstein, C. (1986) *Science* **231**, 1261–1268

[c] Kohler, G. (1986) *Science* **233**, 1281–1286

[d] Yelton, D. E., and Scharff, M. D. (1981) *Ann. Rev. Biochem.* **50**, 657–680

[e] Birch, J. R., and Lennox, E. S., eds. (1994) *Monoclonal Antibodies*, Wiley-Liss, New York

[f] Ezzell, C. (2001) *Sci. Am.* **285**(Oct), 36–41

[g] Greaves, M. F., ed. (1984) *Monoclonal Antibodies of Receptors: Probes for Receptor Structure*, Chapman & Hall, London

[h] Tzartos, S. J. (1984) *Trends Biochem. Sci.* **9**, 63–67

[i] Harlow, E., and Lane, D. (1999) *Using Antibodies: A Laboratory Manual*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York

[j] Goldman, R. D. (2000) *Trends Biochem. Sci.* **25**, 593–595

[k] Wylie, D. E., Lu, D., Carlson, L. D., Carlson, R., Babacan, K. F., Schuster, S. M., and Wagner, F. W. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4104–4108

[l] Marks, C., and Marks, J. D. (1996) *N. Engl. J. Med.* **335**, 730–733

[m] Neuberger, M. S. (1985) *Trends Biochem. Sci.* **10**, 347– 349

[n] Neuberger, M., and Brüggemann, M. (1997) *Nature (London)* **386**, 25–26

[o] Plückthun, A. (1990) *Nature (London)* **347**, 497–498

[p] Collier, R. J., and Kaplan, D. A. (1984) *Sci. Am.* **251**(Jul), 56–64

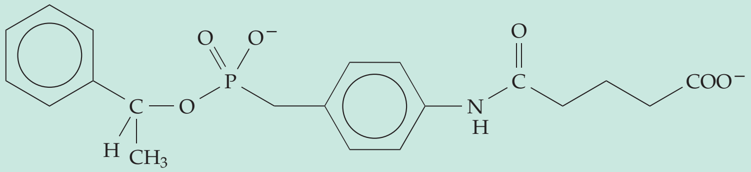
[q] Pastan, I., and FitzGerald, D. (1991) *Science* **254**, 1173–1177

[r] Oeltmann, T. N., and Frankel, A. E. (1991) *Trends Biochem. Sci.* **5**, 2334–2337

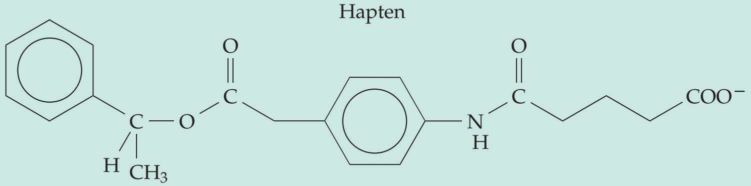
antibiotics, toxins, and hormones. Some of these, such as the defensins, are major, rapid-acting components of the innate system. Both the innate and adaptive systems utilize the complement proteins, and both employ numerous cytokines and other signaling proteins. Plants, and perhaps also other organisms, employ gene silencing by small RNA molecules as part of their defense against viruses.[113a,b]

BOX 31-B CATALYTIC ANTIBODIES

Both enzymes and antibodies are proteins. Antibodies consist of subunits with multiple domains, just as do some enzymes. Both enzymes and antibodies have binding sites for small molecules between domains or subunits. In view of such similarities it isn't surprising that some antibodies have catalytic properties. The possibility was suggested in 1969 by Jencks.[a] He also proposed that injection of a mouse with a hapten, that resembled a transition state for an enzyme, might induce formation of antibodies complementary to the transition-state structure. These might be catalytic. By the early 1980s such antibodies were discovered.[b–d] Some of the first catalytic antibodies (also dubbed **abzymes**) had esterase activity. The haptens used to induce antibody formation were phosphonates such as the following.[e,f]



Hapten



Substrate

Using the transition-state analog shown on p. 485 a catalytic antibody with chorismate mutase activity was isolated.[g] Many antibodies catalyzing additional reactions have also been found. Although they are usually less active than natural enzymes, in some cases they approach enzymatic rates. Furthermore, they may catalyze reactions for which no known enzymes exist.[h]

Catalytic antibodies, like enzymes, must be isolated and purified to homogeneity before they can be studied. Initially this was done by using the hybridoma technique for isolation of monoclonal antibodies (Box 31-A). After induction of antibody formation by injecting a selected hapten into a mouse, large numbers of monoclonal antibodies had to be tested for catalytic activity. Even if several thousand different monoclonal antibodies were tested, only a few with catalytic properties could be found.[i] Newer methods have incorporated recombinant DNA techniques (Box 31-A) and use of combinatorial libraries and phage display.[j–m] Incorporation of acidic or basic groups into the haptens used to induce antibody formation may yield antibodies capable of mimicking the acid–base catalysis employed by natural enzymes.[n,o]

A sample of the types of reaction for which catalytic antibodies have been discovered or designed include the following: ester hydrolysis,[e] transesterification,[p] amide hydrolysis,[q] serine protease-like hydrolysis,[r] elimination,[h,s] aldol cleavage,[t] decarboxylation,[u,v] deiodination by a selenium-containing antibody,[w] pericyclic rearrangements,[g,x] and the Diels–Alder reaction.[y,z] Like natural enzymes catalytic enzymes can be mutated and engineered and can be used to study fundamental aspects of catalysis.[aa] Fluorescent probes incorporate near active sites may provide information about mechanisms or may signal information of diagnostic significance.[bb]

In science we must always expect the unexpected. Do antibodies all catalyze the reaction of singlet molecular oxygen $^1O_2^*$ with H_2O to form H_2O_3 and H_2O_2? How?[cc]

1. Defensins and Other Antibacterial Polypeptides

Only higher vertebrates have an adaptive immune system with circulating antibodies. However, from bacteria to higher plants and human beings all of us utilize defensive polypeptides for protection. More than 500 have been identified.[114] Many have a broad specificity, attacking both bacteria and other pathogens. Among these peptides are more than 200 bacterially produced antibiotics such as gramicidin, tyrocidines, and colicins (Boxes 20-G, 8-D). More recently discovered are the 37- to 59-residue **bacteriocins**, formed by lactic acid bacteria.[115] Like colicin E1 (Box 8-D) and alamethicin (p. 1774) they disrupt cytoplasmic membranes of some other groups of bacteria.

Helicobacter pylori, which is associated with stomach ulcers, forms a 38-residue antibiotic that may help protect infected persons from other bacteria.[116] This peptide forms a simple two-helix structure and is one of a large number of simple helical antimicrobial polypeptides 40 residues or less in length. Among them

BOX 31-B (continued)

[a] Jencks, W. P. (1969) *Catalysis in Chemistry and Enzymology*, McGraw-Hill, New York (p. 288)

[b] Lerner, R. A., and Tramontano, A. (1987) *Trends Biochem. Sci.* **12**, 427–430

[c] Benkovic, S. J. (1992) *Ann. Rev. Biochem.* **61**, 29–54

[d] Lerner, R. A., Benkovic, S. J., and Schultz, P. G. (1991) *Science* **252**, 659–667

[e] Wedemayer, G. J., Wang, L. H., Patten, P. A., Schultz, P. G., and Stevens, R. C. (1997) *J. Mol. Biol.* **268**, 390–400

[f] Janda, K. D., Benkovic, S. J., and Lerner, R. A. (1989) *Science* **244**, 437–440

[g] Haynes, M. R., Stura, E. A., Hilvert, D., and Wilson, I. A. (1994) *Science* **263**, 646–652

[h] Larsen, N. A., Heine, A., Crane, L., Cravatt, B. F., Lerner, R. A., and Wilson, I. A. (2001) *J. Mol. Biol.* **314**, 93–102

[i] Tawfik, D. S., Zemel, R. R., Arad-Yellin, R., Green, B. S., and Eshhar, Z. (1990) *Biochemistry* **29**, 9916–9921

[j] Posner, B., Smiley, J., Lee, I., and Benkovic, S. (1994) *Trends Biochem. Sci.* **19**, 145–150

[k] Janda, K. D., Lo, L.-C., Lo, C.-H. L., Sim, M.-M., Wang, R., Wong, C.-H., and Lerner, R. A. (1997) *Science* **275**, 945–948

[l] Gao, C., Lavey, B. J., Lo, C.-H. L., Datta, A., Wentworth, P., Jr., and Janda, K. D. (1998) *J. Am. Chem. Soc.* **120**, 2211–2217

[m] Baca, M., Scanlan, T. S., Stephenson, R. C., and Wells, J. A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10063–10068

[n] Kemp, D. S. (1995) *Nature (London)* **373**, 196–197

[o] Thorn, S. N., Daniels, R. G., Auditor, M.-T. M., and Hilvert, D. (1995) *Nature (London)* **373**, 228–230

[p] Wirsching, P., Ashley, J. A., Benkovic, S. J., Janda, K. D., and Lerner, R. A. (1991) *Science* **252**, 680–685

[q] Thayer, M. M., Olender, E. H., Arvai, A. S., Koike, C. K., Canestrelli, I. L., Stewart, J. D., Benkovic, S. J., Getzoff, E. D., and Roberts, V. A. (1999) *J. Mol. Biol.* **291**, 329–345

[r] Zhou, G. W., Guo, J., Huang, W., Fletterick, R. J., and Scanlan, T. S. (1994) *Science* **265**, 1059–1064

[s] Romesberg, F. E., Flanagan, M. E., Uno, T., and Schultz, P. G. (1998) *J. Am. Chem. Soc.* **120**, 5160–5167

[t] Karlstrom, A., Zhong, G., Rader, C., Larsen, N. A., Heine, A., Fuller, R., List, B., Tanaka, F., Wilson, I. A., Barbas, C. F., III, and Lerner, R. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3878–3883

[u] Hotta, K., Lange, H., Tantillo, D. J., Houk, K. N., Hilvert, D., and Wilson, I. A. (2000) *J. Mol. Biol.* **302**, 1213–1225

[v] Smiley, J. A., and Benkovic, S. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8319–8323

[w] Lian, G., Ding, L., Chen, M., Liu, Z., Zhao, D., and Ni, J. (2001) *J. Biol. Chem.* **276**, 28037–28041

[x] Driggers, E. M., Cho, H. S., Liu, C. W., Katzka, C. P., Braisted, A. C., Ulrich, H. D., Wemmer, D. E., and Schultz, P. G. (1998) *J. Am. Chem. Soc.* **120**, 1945–1958

[y] Romesberg, F. E., Spiller, B., Schultz, P. G., and Stevens, R. C. (1998) *Science* **279**, 1929–1933

[z] Heine, A., Stura, E. A., Yli-Kauhaluoma, J. T., Gao, C., Deng, Q., Beno, B. R., Houk, K. N., Janda, K. D., and Wilson, I. A. (1998) *Science* **279**, 1934–1940

[aa] Romesberg, F. E., Santarsiero, B. D., Spiller, B., Yin, J., Barnes, D., Schultz, P. G., and Stevens, R. C. (1998) *Biochemistry* **37**, 14404–14409

[bb] Simeonov, A., and 14 other authors. (2000) *Science* **290**, 307–313

[cc] Wentworth, P., Jr., Jones, L. H., Wentworth, A. D., Zhu, X., Larsen, N. A., Wilson, I. A., Xu, X., Goddard, W. A., III, Janda, K. D., Eschenmoser, A., and Lerner, R. A. (2001) *Science* **293**, 1806–1811

are the **cecropins** of insects and **magainins** and **buforins** of amphibians.[117,117a] Many of these kill by disrupting membranes or by forming pores in membranes. However, others enter bacteria and disrupt functions of nucleic acids, enzymes, etc.[117,118] Many antibacterial peptides have been isolated from insects,[12,119] scorpions,[120] spiders and horseshoe crabs,[121,122] and amphibians.[123] All of these organisms lack adaptive immunity but have strong innate immunity.

The human body is protected by two groups of defensins formed in the skin, in mucous membranes, in secretions of neutrophils, and other phagocytic cells. The α-defensins (Fig. 31-7) are 29–35 residues in length and are active against both gram-positive and gram-negative bacteria as well as fungi and enveloped viruses including HIV.[12,124–125a] The β-defensins are mainly active against gram-negative bacteria and yeast. They also possess immunostimulatory activity that is important in activating the adaptive immune response.[32,125,126] Various tissue-specific defensins have been discovered.[127] Peptides of the **trefoil family** protect the gastrointestinal tract.[128–130] Eosinophils,[131] leukocytes, and neutrophils make additional protective proteins. One leukocyte defensin is a macrocyclic peptide, whose gene may have arisen by fusion of two segments encoding nonapeptide segments of α-defensins.[132] Neutrophils form, in addition to defensins, α-helical peptides called **cathelicidins**, which protect skin from invasive bacterial infection.[133,133a] Their synthesis is greatly increased after wounding. They may be among the proteins whose absence after severe burning is likely to be fatal.

Both α and β-defensins consist largely of β strands (Fig. 31-7) and are linked by three disulfide bridges. Some scorpion and insect defensins resemble scorpion toxins (Fig. 30-16) and have four S–S bridges. Fungi and green plants[135,136] also form antimicrobial peptides. A 30-residue fungal protein is highly knotted and contains four S–S bridges.[137] Some polypeptides from the oleander and related plants are 29- to 31-residue macrocyclic structures with two S–S bridges in a **cysteine knot** structure[138–140] (Fig. 31-7). They are exceptionally stable and protease-resistant and may have defensive activity against insects. Defensins are small polypeptides, but larger proteins are also part of the innate defense system. For example, a

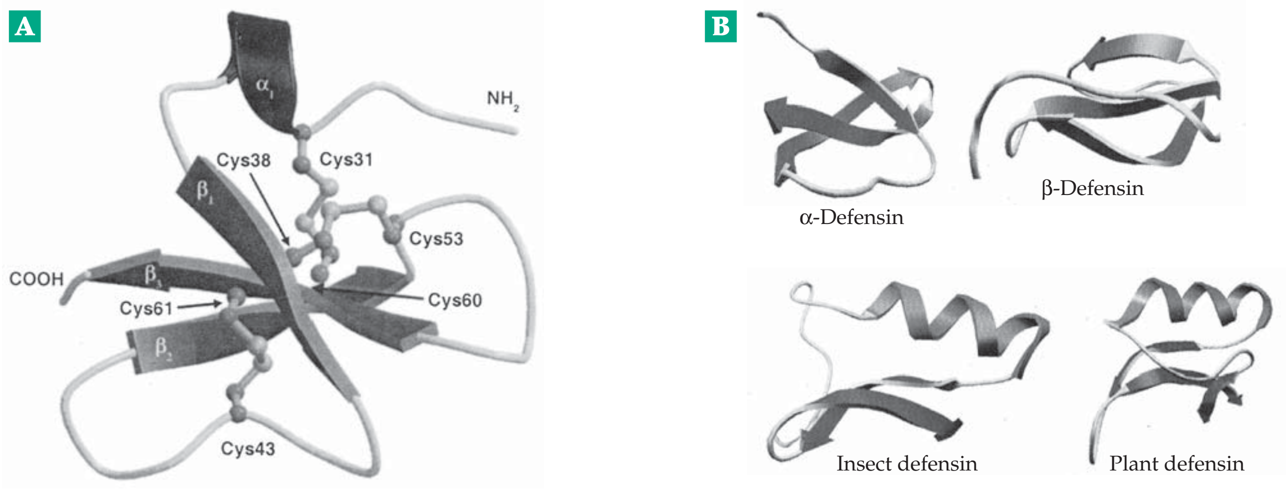


Figure 31-7 Ribbon structures of some defensins. (A) Structure of a human β-defensin showing the three disulfide bonds. From Bauer *et al.*[134] Courtesy of Heinrich Sticht. (B) Comparison of the folding patterns of four types of defensins. Mammalian α- and β-defensins are all β sheets with somewhat different arrangements of disulfide bridges. Insect and plant defensins have an α helix joined to the β sheet. Mammalian and insect defensins have three disulfide bridges, while plant defensins have four. From Hoffmann *et al.*[12] Courtesy of Jules A. Hoffmann.

93-residue protein from onion seeds resembles plant in lipid-transfer proteins.[135] Some frog skins contain a 60-residue trypsin inhibitor.[141] Ribosome-inactivating proteins are well known (Box 29-A).

2. Complement

Complement is a group of more than 30 proteins found in blood serum, which are activated in a cascade mechanism when antibody and antigen combine[10,111,142–148] (Eq. 31-1). This **classical pathway** for activation of complement is outlined in Fig. 31-8. The proteins involved in the cascade are designated C1 to C9. Many of them undergo proteolytic cleavage, the cleavage products being designated by a or b, e.g., C3a and C3b. The b fragment is usually the larger of the two. There is also an **alternative pathway** that is part of the innate system. It is activated by such foreign surfaces as lipopolysaccharides of bacterial cell walls. Its special proteins are called **factors**, e.g., factor B, factor D. A third pathway, the **lectin pathway**, is activated by microbial surface mannans, which bind to a serum **mannan-binding lectin** (**MBL**). This protein, a so-called defense collagen, resembles protein C1q (next paragraph).[149] It activates two associated serine proteases (MASP1 and MASP2), which are able to cause cleavage of proteins C4 and C2 and possibly C3 in the classical pathway (Fig. 31-8). The ultimate effects of the action of complement include destruction of cells by lysis and activation of leukocytes, which engulf foreign cells by phagocytosis. Complement also induces the release of **chemotactic factors** that attract polymorphonuclear leukocytes and monocytes to the site involved.[150]

$$\begin{tikzcd}
\text{Antigen + antibody} \arrow[d] & \\
\text{Complex} \arrow[r, "\text{Complement proteins}"] & \text{Cell lysis, phagocytosis, etc.}
\end{tikzcd}$$

(31-1)

The classical pathway begins with the **recognition component C1** of complement. This is a complex of three proteins, C1q, C1r, and C1s. Proteins C1r and C1s form a mixed tetramer Clr_2s_2, while C1q binds to the C_H2 domain of "activated" antibodies, that is, with immunoglobulins that have combined with an antigen. It takes at least a dimer or larger aggregate of IgG to activate C1q, whereas a single molecule of the naturally pentameric IgM suffices. The mechanism by which this activation occurs is uncertain. Perhaps a change of conformation within the immunoglobulin acompanies antibody binding and is responsible for generation of a binding site for C1q. It may seem strange that haptens cannot cause complement binding, and that they do not cause detectable conformational alterations in Fab. Only multivalent antigens able to bind to more than one antibody induce complement binding. However, as we have learned in recent years, many biological responses involve transient assembly of large aggregates of different protein components. In this context, the requirement for two or more antibody molecules doesn't seem so strange.

The 400-kDa C1q consists of a central portion of diameter 3–6 nm and length 10–12 nm to which are attached six very thin connecting strands. These are

~14.5 nm long and ~1.5 nm in diameter and terminate ~135-residue globular ends of ~6 nm diameter,[151–153] which are thought to be the sites of combination with the immunoglobulins. The thin connecting strands have, for most of their length, a collagenlike structure with a high content of hydroxyproline and hydroxylysine. The latter is glycosylated by glucosylgalactosyl disaccharides as in collagen itself (pp. 181, 432, 433). The reason for this unusual structure is not obvious. We do know that the binding of antigens activates the complement-binding regions of antibodies, and that the activated antibodies then bind C1q. This binding in some manner activates C1q, which in turn activates C1r subunits of the Clr_2s_2 tetramer.[154] The latter is thought to bind at the center of C1q, while the antibodies bind at the outer ends. We don't know how the activation message is carried from the outer arms to the center. C1q is a member of a group of collagen-like proteins that includes protein MBL (also designated MBP) and surfactant protein A (SP-A; p. 436).[155]

Activated C1r (often designated $\overline{\text{C1r}}$ but here and in Fig. 31-8 as C1**r**) is one of five different serine proteases involved in activation of complement.[156] The substrate for the trypsinlike C1**r** is C1s, a proenzyme which is converted by the action of C1**r** into another trypsinlike serine protease C1**s**.[157–160] Through a rather elaborate cascade mechanism, depicted in Fig. 31-8, the important proenzyme C2 is activated.[161] Its active form C2**a** is a serine protease, which cleaves proteins C3 and C5 to the active forms C3b and C5b. Protein C4 is also cleaved to C4b by activated C1. C4 and C3 are also activated, and protein C5 is cleaved

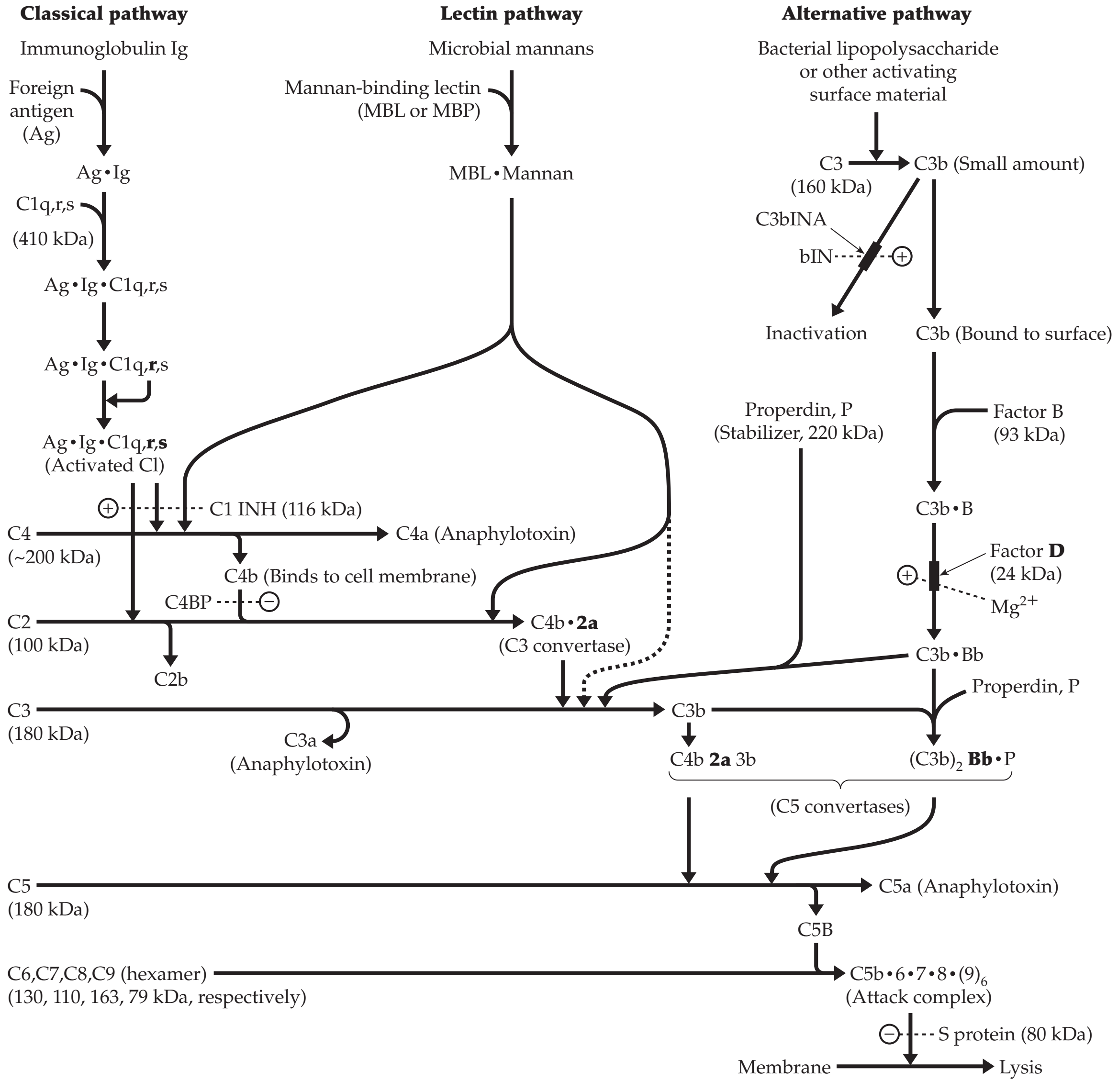

Classical pathway
Immunoglobulin Ig
Foreign antigen (Ag)
Ag•Ig
C1q,r,s
(410 kDa)
Ag•Ig•C1q,r,s
Ag•Ig•C1q,**r**,s
Ag•Ig•C1q,**r**,**s**
(Activated Cl)
+
C1 INH (116 kDa)
C4
(~200 kDa)
C4a (Anaphylotoxin)
C4b (Binds to cell membrane)
C4BP
−
C2
(100 kDa)
C2b
C4b•**2a**
(C3 convertase)
C3
(180 kDa)
C3a
(Anaphylotoxin)
Lectin pathway
Microbial mannans
Mannan-binding lectin
(MBL or MBP)
MBL•Mannan
Alternative pathway
Bacterial lipopolysaccharide or other activating surface material
C3
(160 kDa)
C3b (Small amount)
C3bINA
bIN
+
Inactivation
C3b (Bound to surface)
Properdin, P
(Stabilizer, 220 kDa)
Factor B
(93 kDa)
C3b•B
+
Factor **D**
(24 kDa)
Mg2+
C3b•Bb
Properdin, P
C3b
C4b **2a** 3b
(C3b)2 **Bb**•P
(C5 convertases)
C5
(180 kDa)
C5a (Anaphylotoxin)
C5B
C6,C7,C8,C9 (hexamer)
(130, 110, 163, 79 kDa, respectively)
C5b•6•7•8•(9)6
(Attack complex)
−
S protein (80 kDa)
Membrane
Lysis

Figure 31-8 Pathways for activation of the complement system. Active proteases are designated by abbreviations in boldface.

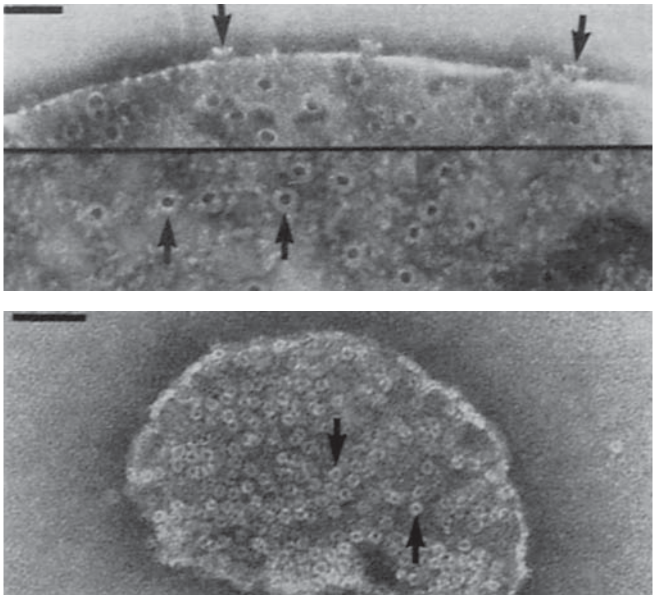


Figure 31-9 Electron micrograph of a negatively stained sheep erythrocyte lysed with human complement. The cylindrical "attack complex" embedded in the membrane is seen in the upper left frame in side projection and in the lower frames in axial projection. The top views are of a proteolytically "stripped" ghost, the lower view from a freshly lysed ghost. The inner diameter of the cylinders is 10 nm; scale bars 50 nm. From Bhakdi and Tranum-Jensen.[169]

to C5b. The 200-kDa C4 consists of three chains, all derived by proteolytic cleavage of a precursor. It is also glycosylated and sulfated. Both C4 and C3 contain internal thiol ester linkages and act as "molecular mousetraps" (Box 12-D).[162] They react to fix the proteins covalently to the assembling complement complex.[163] Protein C5b interacts with C6, C7, C8; and six molecules of C9 to generate an "attack complex," which inserts donutlike rings into the cell membrane being attacked (Fig. 31-9).[164,164a] Although there has been some uncertainty about the mechanism of lysis, it seems likely that it is at least partly a result of loss of ions through the holes in the donut.

In the alternative pathway of activation a small amount of C3b is formed and becomes bound to the cell surface. This binds another proenzyme factor B,[165] which is converted by protease factor D[166] to active protease **Bb**. The latter in its complex with C3b is the enzyme that cleaves C3 in large amounts and permits a rapid formation of more **Bb** and also of the complex (C3b)$_2$ **Bb•**P, which attacks C5 (Fig. 31-8). These complexes are stabilized by the abundant serum protein **properdin** (P).[167]

Other components of complement are the plasma C1 protease inhibitor,[168] which prevents accidental activation of the system, and protease C3bINA, which inactivates C3b. The latter depends upon accessory protein bIN. Another component, serum carboxypeptidase B (SCPB), inactivates anaphylotoxins C3a, C4a, and C5a. These small ~80-residue pieces have a variety of powerful biological activities.[170–174] They are chemotactic factors for leukocytes and induce release of histamine from mast cells. In excess, they can cause anaphylaxis; hence their rapid degradation is essential. An excess of C5a may be present both in asthma and in rheumatoid arthritis.

The ninth component of complement, C9, is a 70-kDa 537-residue water-soluble glycoprotein, which contains a hydrophobic domain that aggregates to form the ion conducting channels.[175,176] Proteins that closely resemble C9 and are called **perforins** or **cytolysins** are found in cytoplasmic granules of cytotoxic T-lymphocytes and natural killer cells. These ~66-kDa proteins are assembled into rings similar to those formed by C9 and may be involved in the killing action of these cells (Fig. 31-9).[177–179] Certain pathogenic amebas, which may cause a fatal infection, also utilize a similar pore-forming protein.[180]

Every regulatory system in the body must be prevented from overactivity or activity that is unnecessarily prolonged. This can help us understand that, just as with blood clotting (Fig. 12-17), a network of regulatory factors controls the complement system. Among these are an inhibitory C4b-binding protein (C4BP),[181] which acts to prevent excessive formation of the C4b•C2**a** complex (Fig. 31-8). Complement **cofactor I** is a serine protease that cleaves both C3b and C4b into smaller pieces in the presence of **cofactor H**[181a] or of C4BP. Its absence leads to excessive consumption of C3 and recurrent pyogenic infections.[182] The **membrane cofactor protein** (MCP) stimulates this action of cofactor I in inhibiting attacks of complement on the cells that carry MCP.[183] Acting in the opposite direction is **complement receptor 2** (CR2 or CD21), which acts as a receptor for proteolytic fragment C3d. This fragment binds CR2-bearing cells to the B cell receptor, amplifying the B cell response to foreign antigens.[184]

Complement is involved not only in attacking foreign cells but in inflammation. Unfortunately, this is sometimes accompanied by serious problems. Human diseases in which complement is thought to be involved include glomerulonephritis, rheumatoid arthritis, myasthenia gravis, and lupus erythematosus.

3. Cytokines, Interferons, and the Acute-Phase Response

The body responds in many ways to infection, injury, or cancer. These include the secretion of cytokines, interferons, and proteins of the acute phase response. These proteins, many of which are quite small, are involved in communication with other cells, often with specialized cells of the immune system. **Cytokines**, some of which are considered in Chapter 30 (Section A,6; Fig. 30-6), are small hormonelike molecules. They may stimulate, inhibit, or exhibit other effects on cells of the immune system. They often have pleiotropic effects, not acting in the same way on all types of cells.[185] The cytokines known as **interleukins** (IL-1, etc.) are produced by leukocytes. **Lymphokines** are formed by lymphocytes and **monokines** by monocytes. Based on their functions there are four categories of cytokines.[111]

(1) Mediation of natural immunity: type I interferons, IL-1, IL-6, and more than 40 **chemokines** (small highly basic chemotactic proteins).
(2) Regulation of lymphocytes, activation, growth, and differentiation of B and T cells: IL-2, IL-4, IL-21, TGF-β.
(3) Regulation of immune-mediated inflammation: Interferon-γ, tumor necrosis factor (TNF), IL-5, IL-10, IL-12, and migration inhibition factor (MIF).
(4) Stimulation of hematopoeiesis (IL-3, IL-7), colony-stimulating factors (CSF; see also Chapter 32).

Cytokines all function using a group of transmembrane receptors embedded in the plasma membranes of target cells. The receptors have no tyrosine kinase activity but associate with and activate kinases known as **Janus kinases (JAKs)**. These kinases phosphorylate tyrosine side chains in their receptors, and the phosphorylated receptors activate transcription factors of the **STAT** (signal transducer–activators of transcription) group.[186–195] The specificity of cytokine action results from a combination of receptor recognition and recognition of the various STAT molecules by different JAKs.[111] Cytokines have a variety of structures. Many are helix bundles or have β sheet structures (Fig. 30-6).

Interferons. The interferons (IFNs),[196,197] which were discovered in 1957, are proteins secreted by leukocytes, fibroblasts, and activated lymphocytes. They inhibit replication of viruses as well as the growth of host cells and also have antitumor activity. Interferons are classified as α (from leukocytes), β (from fibroblasts), and γ (from lymphocytes). According to their affinities for the two types of known interferon receptors, interferons IFN-α, IFN-β, and the less well known IFN-ω and IFN-τ are designated type I,[198–201] while interferon γ (IFN-γ) is type II. At least 15 homologous 166-residue human α interferons are known.

The binding of interferons to their receptors induces a rapid increase in the transcription of particular genes and synthesis of corresponding proteins.[196,202] One of the proteins induced is a **double-stranded RNA-activated 2'-5'A synthase**, which polymerizes ATP to a series of 2'-5' linked oligonucleotides containing triphosphates at the 5' termini.[202–204] Double-stranded RNA is uncommon except in replicating viruses, and it is thought that the activation by dsRNA is related to establishment of an antiviral state. Another interferon-induced enzyme is the small subunit of eukaryotic protein synthesis initiation factor eIF-2. This is converted to an inactive phosphorylated form by a dsRNA-dependent protein kinase[205] (Fig. 31-10). The protein kinase also appears to be an interferon-induced protein[206] as is the oligo(2'-5' A)-activated RNAse indicated in Fig. 31-10.[207] Interferons have effects other than inducing the antiviral state. Thus, human IFN-β_2 is identical to a B-cell differentiation factor.[208] Both IFN-α and IFN-β have antigrowth activity and are currently in use for treatment of some forms of cancer as well as for viral infections.[209]

Interleukin-1 (Il-1) plays a key role in the body's response to microbes and to tissue injury.[210,211] It actually consists of three similar proteins, **Il-1α**, **IL-1β**, and **Il-1 receptor antagonist**. The first two are the active cytokines with a wide range of effects among

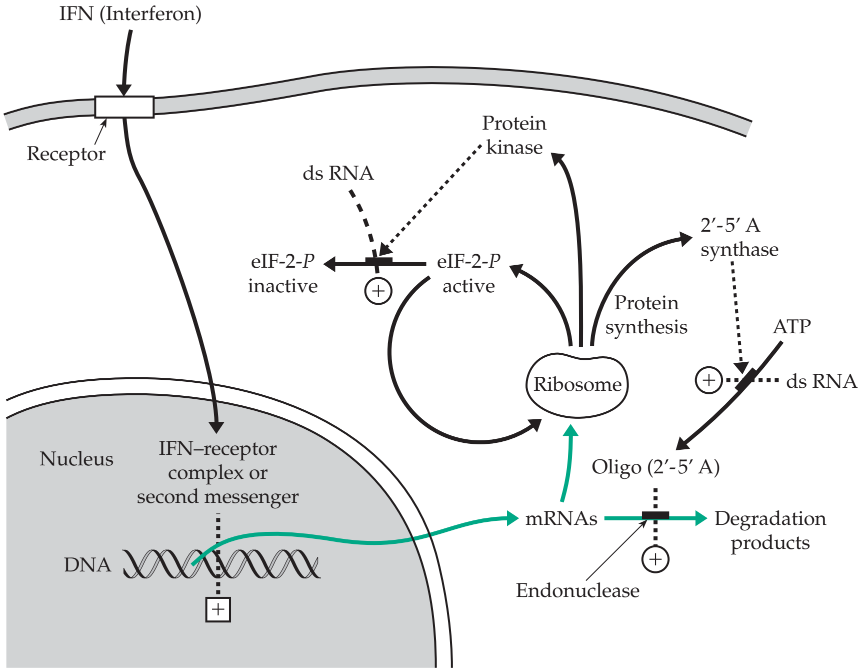


Figure 31-10 Some responses to the binding of an interferon to a cell surface receptor.

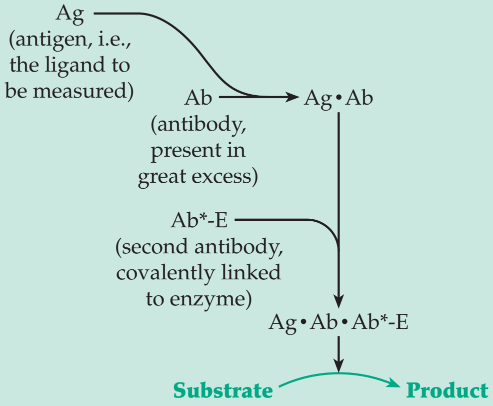
which are induction of inflammation and pain.[212] Il-1β is thought to be most active in promoting inflammation but only after it is cleaved by **interleukin-1β-converting enzyme** (see p. 619).[213] Blocking of Il-1 receptors provides a potential new method for control of pain.[214] **Interleukin-6** (IL-6) is also needed for an optimal immune system. Its effects overlap those of Il-1, and it has a potent activity in inducing the acute-phase response.[215] Like Il-1, it is a four-helix cytokine.

Also required by the immune response are the numerous **chemokines**. Chemoattractant molecules provide concentration gradients that direct the movement of B and T cells and other leukocytes.[216–219] Chemokines bind to seven-helix receptors, often

BOX 31-C IMMUNOASSAYS

Among the important techniques that have permitted rapid progress in studies of hormone action is the use of specific antibodies formed against hormones, hormone-protein conjugates, or other molecules.[a–c] The first of these techniques to come into general use was the **radioimmunoassay** (RIA),[d–f] which was devised by Yalow and Berson.[d] In one form of RIA various amounts of a sample containing an unknown quantity of hormone, e.g., insulin, are placed in a series of tubes. Additional tubes containing known amounts of the hormone are also prepared. Then a standard quantity of radiolabeled hormone (often iodinated with a γ emitter such as ^{125}I) is added to each tube together with a standard quantity of the specific antibody to the hormone. The solution is incubated for minutes or hours to obtain equilibrium between hormone (the antigen) and antibody-hormone complex. The antibody-hormone complex is then separated, e.g., by gel filtration or ammonium sulfate precipitation, and the radioactivity of the complex is measured. In the tubes containing higher concentrations of hormone, the labeled hormone has been diluted more, and the amount bound to antibody is less than in tubes with lower concentrations of hormone. The tubes of known concentrations are used to construct a standard curve from which the unknown concentrations can be read. As little as a femptomole of hormone (i.e., the amount present in 1 ml of a 10^{-12} M solution can be detected).[f] Methods are available for virtually every pure hormone.[c]

The RIA methods were made more convenient by adsorbing either the antibody of antigen to the plastic surface of a tube or depression plate. This facilitates separation of the antibody-ligand complex and washing. A variety of other immunoassays techniques have been devised. For example, in **enzyme-linked immunoabsorbent assays** (ELISA),[c] the amount of adsorbed antibody-ligand complex is measured by treating the washed surface with a second antibody, which is directed against the first. The second antibody is linked covalently to an enzyme, whose activity can then be measured by a suitable colorimetric procedure. The reactions involved are as follows.[g,h]


Ag
(antigen, i.e., the ligand to be measured)
Ab
(antibody, present in great excess)
Ag • Ab
Ab*-E
(second antibody, covalently linked to enzyme)
Ag • Ab • Ab*-E
Substrate
Product

Variations, which avoid the use of radioisotopes, are replacing RIA. Some utilize stable isotopes. However, ^{14}C at such low levels that there is no radioactive waste can be coupled with accelerator mass spectrometry to provide very sensitive immunoassays.[i] A great variety of other procedures are available. Some involve coupling to antibodies that carry fluorescent labels. Many are now automated. Often protein A from *Staphylococcus aureus* is utilized in various ways that take advantage of its ability to bind to the Fc portion of IgG from virtually all mammals. For example, it may fix antibodies to a surface or to a label.[j]

---

[a] Price, C. P., and Newman, D. J., eds. (1991) *Principles and Practice of Immunoassay*, Stockton Press, New York

[b] Lindbladh, C., Mosbach, K., and Bülow, L. (1993) *Trends Biochem. Sci.* **18**, 279–283

[c] Crowther, J. R. (1995) *ELISA: Theory and Practice*, Humana Press, Totowa, New Jersey

[d] Yalow, R. S. (1978) *Science* **200**, 1236–1245

[e] Brooker, B., Terasake, W. L., and Price, M. G. (1976) *Science* **194**, 270–276

[f] Jaffe, B. M., and Behrmann, H. R., eds. (1974) *Methods of Hormone Radioimmunoassay*, Academic Press, New York

[g] van Vunakis, H. and Langone, J. J.(1980) *Methods Enzymol.* **70**, entire volume

[h] Langone, J. J., and van Vunakis, H..eds. (1983) *Methods of Enzymology* **92**, entire volume

[i] Shan, G., Huang, W., Gee, S. J., Buchholz, B. A., Vogel, J. S., and Hammock, B. D. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2445–2449

[j] Surolia, A., Pain, D., and Khan, M. I. (1982) *Trends Biochem. Sci.* **7**, 74–76

without rigid specificies. However, some such as **eotaxin** (Fig. 30-6E) are more specific. Eotaxin attracts primarily eosinophils and basophils during allergic reactions.[220] Interleukin-8 (IL-8) is a proinflammatory cytokine and a powerful attractant for neutrophils. Neutrophils are attracted into affected tissues, where they undergo a respiratory burst and generate toxic compounds from O_2 (pp. 1072–1074).[221–222c]

The second group of cytokines regulate B and T lymphocytes. Among them interleukin-2 (IL-2) stands out as the major promoter of growth and differentiation of T cells. It was the first hormone of the immune system to be recognized.[223] Both Il-2 and Il-4 have short four-helix structures (Fig. 30-6A). Il-2 is synthesized by activated T cells and binds to a multisubunit receptor complex. The latter associates with tyrosine kinases of both the Src family (p. 572) and with Janus kinases and also activates phosphatidylinositol 3-kinase.[224–227] Il-2 promotes growth and differentiation, and clonal expansion of T cells, a key aspect of the cellular immune system. It also acts as an immunomodulator of B cells, macrophages, and NK cells. Considerable excitement has accompanied the possibility of activating lymphocytes with IL-2 produced from cloned genes in bacteria to increase their ability to kill cancer cells. However, IL-2 is toxic, and this is limiting its use.

Interleukin-4 (IL-4), cooperating with IL-21,[227a] stimulates growth of activated B cells, T lymphocytes, and mast cells, induces formation of cytotoxic $CD8^+$ T cells, and enhances formation of IgG.[228,229] The **transforming growth factor-β** (TGF-β) is another cytokine that modulates the development of the immune system. It affects a very broad range of tissues and is discussed in Chapter 32.

Inflammatory influences. Inflammation (p. 1211), which usually accompanies infection and can also arise from allergic responses, is affected by many substances.[229a–e] These include chemotactic factors that attract neutrophils and monocytes[222b,229d] and the adhesion molecules that assist in the movement of lymphocytes.[229e,f] Some epithelial tissues, such as the mucosal surfaces of the gastrointestinal tract, are maintained in a continuous very low level of inflammation. This reflects the balance between activation of the immune system and inhibition of the system by signals from microorganisms both pathogenic and commensal.[229g,h]

The third group of cytokines[229b,c] are among the molecules that regulate inflammation. One of these is **interferon-γ** (IFN-γ). Like the type I interferons it induces an antiviral state. However, its most important biological function lies in modulation of the immune system. It induces synthesis of both class I and class II (HLA-DR) antigens (see Section D), activates macrophages, and regulates synthesis and activity of other lymphokines.[230–231b] One mechanism of immunomodulation may involve induction of an aminopeptidase that participates in "trimming" of antigenic peptides that participate in antigen presentation (Section D,6). This may directly alter the T-cell response.[232] IFN-γ has a major effect on skin cell-mediated immune responses.[233] INF-γ exists in solution as a symmetric dimer, which binds to two molecules of its major surface receptor.[231,234] The antiviral activity of IFN-γ is largely a result of induction of large amounts of **guanylate-binding proteins**, large 60- to 100-kDa GTPases.[235,236]

Occasionally a well-established cancer regresses and disappears spontaneously. In the late 1800s it was observed that this sometimes happened, when a person had a concurrent severe bacterial infection. W. B. Coley pursued this lead for many years, treating cancer patients with extracts of killed bacteria, which, although highly toxic were safer than live bacteria. In the laboratory filtrates from cultures of gram-negative bacteria were shown to kill some experimental mouse tumors. The active ingredient was identified as a highly toxic and pyrogenic lipopolysaccharide (Fig. 8-28; Chapter 20, Section E).[237,238] This lipopolysaccharide has a powerful activating effect on macrophages. More recently it was found that the activated macrophages produce a protein known as **tumor necrosis factor** (TNF-α) that can destroy tumor cells and also acts together with interferons in inducing resistance to viruses.[239,240] A similar **lymphotoxin** (TNF-β) is secreted by lymphocytes.[241] Although highly toxic there has been hope of obtaining engineered forms of these proteins more specifically toxic to tumors.

TNF-α is identical to **cachetin**, a protein that suppresses completely the lipoprotein lipase of adipose tissue and is believed to be responsible for **cachexia**, a condition of general ill health, malnutrition, weight loss, and wasting of muscle that accompanies cancer and other chronic diseases. Nevertheless, TNF-α may be overproduced in obesity as well. It has been suggested that abnormal production of TNF-α may induce cachexia while abnormal action of the cytokine may cause obesity.[233] Some TNF receptors have "death domains" and trigger apoptosis, while other receptors promote proliferation and differentiation via transcription factor NF-κB.[242]

Other cytokines with lymphocyte-regulatory functions are IL-5, IL-10, IL-12, and the **macrophage migration-inhibition factor** (MIF). IL-10 is secreted by B cells, T cells, keratinocytes, monocytes, and macrophages. It suppresses synthesis of many cytokines but stimulates growth and activity of activated B cells.[243] IL-12 is formed by monocytes, macrophages, neutrophils, and dendritic cells. It activates T cells and NK cells, is a very potent stimulator of INF-γ formation, and also inhibits angiogenesis in tumor cells. It stimulates defenses against a wide range of infectious diseases caused by bacteria, fungi, protozoa, and worms.[244,245] The 115-residue MIF is formed not only by the immune system but also by many other tissues. The first lymphokine to be discovered, MIF, inhibits migration of macrophages and is also a mediator of toxic shock.[246,247] MIF is also an enzyme, a **phenylpyruvate tautomerase** (p. 692).[247]

The fourth group of cytokines are involved in hematopoiesis and control the developmental steps portrayed in Fig. 31-2. They are discussed in Chapter 32.

The **acute-phase response** consists of increased production of a group of plasma proteins in response to tissue injury or inflammation.[229c] Important acute-phase reactants are the **C-reactive protein**,[229a,248,249] **serum amyloid A**,[250,251] **haptoglobin**, **hemopexin**, **α1-acid glycoprotein**,[252] and **α2-macroglobulin**.[253] The C-reactive protein precipitates pneumoccocal polysaccharides in the presence of Ca^{2+}. It is present in primitive invertebrates and may serve as a rudimentary immunoglobulin.[248] Serum amyloid A is one of the apolipoproteins associated with high-density lipoproteins (Chapter 21, Section A). Its concentration may increase as much as 1000-fold during the acute-phase response, and during prolonged stress it may precipitate as extracellular amyloid fibers. This secondary amyloidosis is sometimes a severe pathological problem. The level of the general protease inhibitor α_2-macroglobulin can increase several hundredfold.

D. Organizing the Immune Response

A person’s immune system must be able to respond to a large variety of foreign antigens without reacting against the individual’s own tissues. The huge variety of antibodies that can be formed arise from the existence of B cells with millions of different sequences in their antibody genes. When an immune response occurs only a few B cells are stimulated to proliferate, and it is these selected clones that provide the needed specific antibodies and memory cells. However, it is not immediately obvious how we avoid a disastrous attack of the immune system triggered by the many antigenic determinants (**epitopes**) present in our own cell surfaces and macromolecules. Part of the answer is that the immune system “learns” early in life what is self and what is nonself. Thus, while foreign tissues cannot usually be grafted without rejection, cells of two immunologically incompatible embryos can be mixed at a very early stage of development, and an animal tolerant to both types of cell will develop.

A full understanding of self-discrimination is not yet available.[253a] The adaptive system, as generally understood, is outlined in the following pages. A current view of the innate system is presented by Medzhitov and Janeway.[11a] However, an alternative description, the **Danger model**, is being developed by Matzinger. Her view is that the immune system is designed not so much to recognize *nonself* as to send *alarm signals* from injured tissues.[11b] Most of the basic mechanisms of the adaptive immune system are not in dispute, but many hard-to-explain phenomena remain uncertain.

1. Coreceptors and the B-Cell Response

Early in life most B cells that would produce antibodies directed against a person’s own tissues (autoreactive B cells) are eliminated or altered to reduce their reactivity.[254] When functional B-cell receptors do bind an antigen, the B cell will not be activated unless **coreceptors** also bind to the antigen-bearing particle. The transmembrane glycoproteins known as CD22, CD21, CD72, and FcγRIIb are among the many coreceptor molecules. Coreceptors often induce tyrosine phosphorylation of internal receptor domains and attract other molecules to form a signaling complex that may release cytokines.[255–256] The coreceptors ensure that an immune response doesn’t take place without at least two signals. They also help to localize the immune response.

Among the most important factors in B-cell activation are the effects of T cells. B cells can independently mount an attack using IgMs against surface antigens. However, B-cell responses to many antigens, e.g., those present on flagella or inserted into membranes, are also dependent upon assistance from helper T cells (T_H cells).[257] These cells also have a major role in determining the longer term fate of B cells. Upon activation B cells may survive or die via apoptosis. They may proliferate (clonal expansion) and differentiate into plasma cells or may become unreactive (**anergy**). They may become long-lived memory cells.

2. The Leukocyte Differentiation Antigens

Before discussing T-cell responses it seems appropriate to mention the nomenclature of molecules (largely glycoproteins) that have been recognized as antigens present on leukocyte surfaces. These same molecules are found on other cells, but the designation of the antigens by a **cluster of differentiation** number, such as CD1, CD4, or CD8, has provided a convenient way of distinguishing different types of leukocytes.[258–261] For example, helper T_H cells are usually $CD4^+$, carrying predominantly CD4. Cytotoxic T cells are predominantly $CD8^+$.[262,263] Both CD4 and CD8 consist largely of Ig-like domains. CD4 is a 55-kDa transmembrane protein with tyrosine kinase activity.[264] It is a monomer containing four Ig-like domains, but CD8 is a disulfide-linked $\alpha\beta$ dimer.[263]

Not all CD molecules are related to IgG. Proteins are often designated by a specific name followed by a CD number, e.g., FcγRII / CD32, ICAM-1 (CD54).

3. Functions of T Cells

T cells carry the responsibility of identifying antigens as foreign or as belonging to self. They do this in immunological synapses (Fig. 31-11) in conjunction with the major histocompatibility complex MHC (Section 5). T cells circulate through the body searching for antigens that indicate danger to the body. To avoid being swept through the bloodstream too rapidly and to be able to enter the lymphoid organs lymphocytes form tethers with adhesion molecules such as the **selectins** (p. 188).[266] They then roll more slowly to their destination. Within the lymphoid tissues the T cells may form synapses with activated B cells, dendritic cells, and macrophages. Within these cells proteosomes generate a stream of peptide fragments, some of which arise from phagocytosed pathogens. These foreign peptide fragments are displayed on the cell surfaces as complexes with type II proteins of the major histocompatibility complex (MHC; see Fig. 31-13). The complexes are checked by CD4$^+$ T cells, some of whose **T-cell receptors** (TCRs) will probably be complementary to the surfaces of the complex of the class II MHC protein and the foreign peptide. The T cell will recognize two things about this complex: the MHC protein is of *self* origin but the antigen is *foreign*. The CD4 on the T cell must also bind to the MHC on the surface of the antigen-presenting cell. Other costimulatory interactions may be needed as well.[50,267] Both CD4$^+$ and CD8$^+$ T cells tend to bind to oligomeric **activation clusters** of receptors within the immunological synapses.[267a] Other proteins also participate in assembly of these activation complexes.[267b] Of

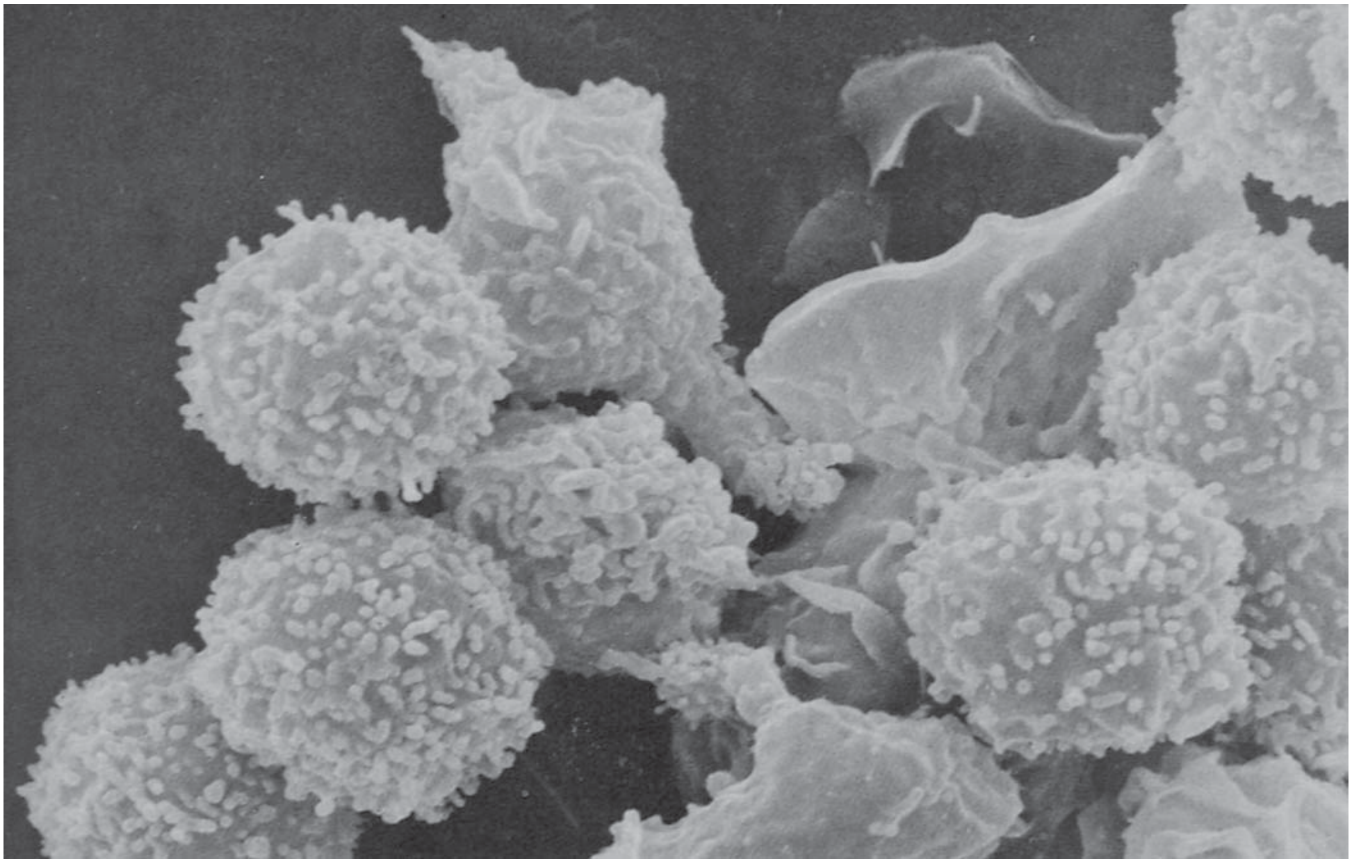


Figure 31-11 The T lymphocytes seen here are forming synapses with the large flat macrophage in the center. The macrophage is displaying antigenic peptide fragments bound to molecules of the major histocompatibility complex (MHC). Most T cells carry their own specific type of receptor. If it is complementary to a displayed antigen fragment it will bind, and the T cell will respond. Depending upon what other coreceptors are activated, it may become a T_H helper cell, or a cytotoxic T cell, or it may become inactive. Micrograph from Grey *et al.*[265] Scanning electron micrograph courtesy of Morton H. Nielsen and Ole Werdelin, University of Copenhagen.

particular interest is the recognition of an antigenic peptide produced by a B cell. The B cell has probably already recognized and phagocytized a foreign protein and is displaying peptides from that protein on its MHC I molecules. Recognition of this peptide complex by a $CD4^+$ T cell will stimulate the cell to become a T_H helper cell, which will in turn stimulate the B cell to proliferate and differentiate into a clone of as many antibody-forming plasma cells. The essential nature of the costimulation by CD4 is emphasized by the fact that infection by HIV-1, which is mediated by CD4, leads to loss of CD4 from plasma membranes and to the weakening of the immune response toward various pathogens that is characteristic of AIDS.[268–270]

In a similar manner $CD8^+$ T cells recognize peptide fragments displayed on MHC class I molecules. These fragments arise via a somewhat different pathway that forms fragments of viral proteins or proteins of other intracellular pathogens. Recognition by a $CD8^+$ T cell usually converts it into a **cytotoxic (killer) T cell**, which will kill the infected cell.[270a,b] This type of immune reaction was first recognized by the phenomena of **delayed hypersensitivity** and of **transplantation immunity**, i.e., the rejection of transplanted tissues. Both phenomena are caused by cytotoxic T cells. In delayed hypersensitivity they appear to be confused and to attack host cells.

Some very hydrophobic antigens are presented by neither a class I nor a class II MHC molecule but by members of the CD1 family, leukocyte surface proteins that are not encoded in the MHC gene region.[266,271,272]

4. Natural Killer Cells

An additional line of defense is provided by **natural killer cells** (NK cells), a type of circulating lymphoid cell able to kill cancer cells, to participate in antiviral defenses, and to help control immune responses.[273–276] NK cells, which utilize their own signaling pathways, are also able to use MHC class I molecules to recognize and to spare the lives of normal, healthy cells.[277,277a,b] Partial deprivation of a night's sleep can reduce NK cell activity, damaging the cellular immune response.[278]

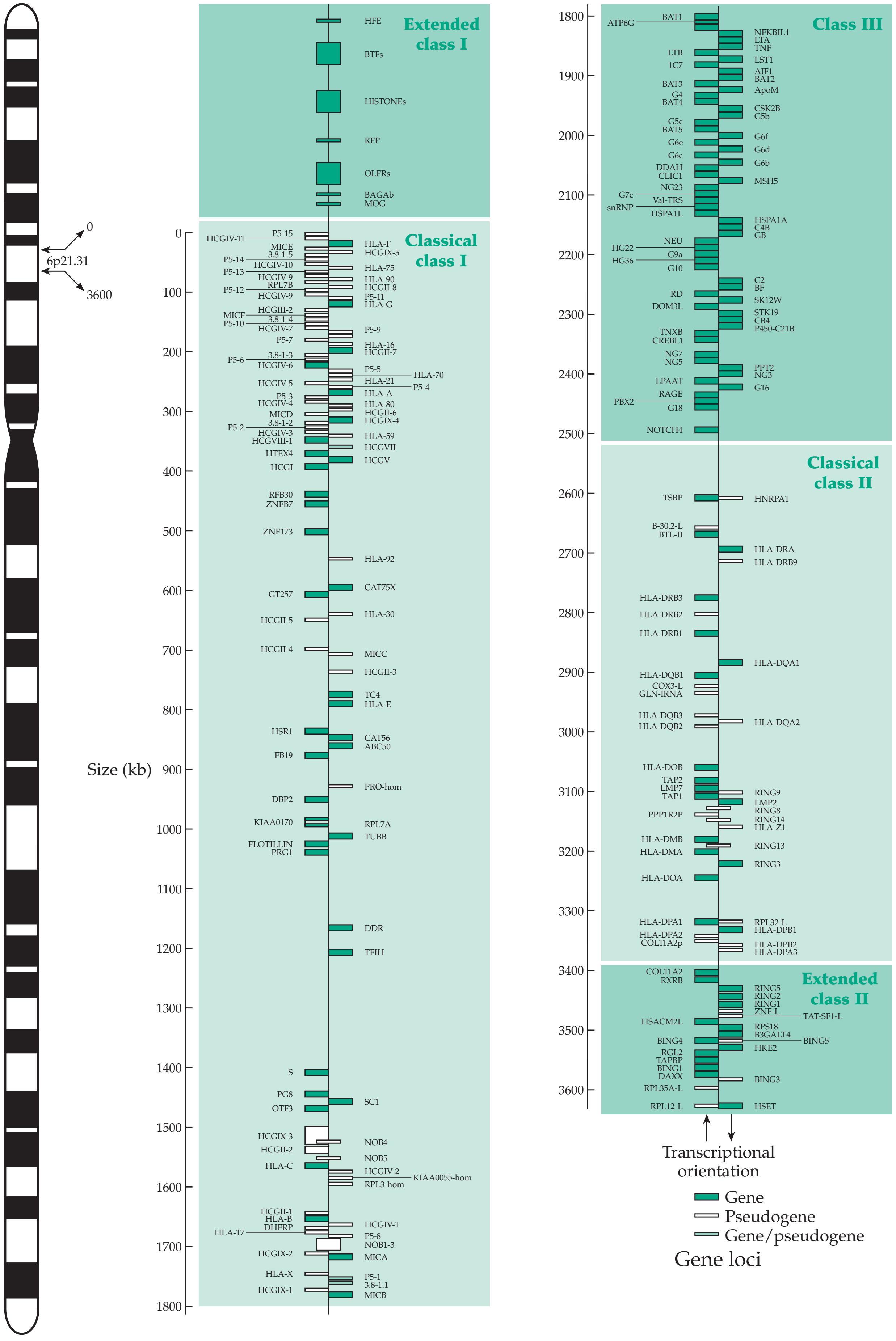
5. Identifying Self: The Major Histocompatibility Complex

Proteins encoded by a single cluster of genes are known as the **major histocompatibility complex** (MHC).[279] These proteins, which are essential to T-cell function, were first recognized as the primary determinants of the compatibility of grafted tissues with the host's immune system.[280] A lack of histocompatibility can be disastrous. Not only are grafted tissues rejected but T lympocytes from the grafted tissues sometimes proliferate, attack, and kill the host. The MHC of mice is usually referred to as the **H-2 complex**[281,282] and that of humans as **leukocyte locus A (HLA)**.[283,284] Although the MHC is the most important determinant of histocompatibility, differences in other genes may also lead to a slow rejection of transplanted tissues. Since there are many different MHC genes, transplantation is successful only within inbred lines.

Some of the MHC genes have a large number (50–100) of alleles. So great is this genetic polymophism that it is extremely unlikely that two individuals will have an identical set of histocompatibility genes. The MHC (HLA) genes are located in a 2-centimorgan (~3.6 kb) region of the short arm of human chromosome 6 (Fig. 31-12)[284,285] and on chromosome 17 of mice. These genes are of at least three classes. **Class I genes** (called HLA-A, HLA-B, HLA-C, HLA-E, HLA-F, and HLA-G in humans; see Fig. 31-12) encode the major transplantation antigens, which are found on the surfaces of nearly all cells of the body. **Class II genes** encode proteins found largely on the membranes of B lymphocytes, macrophages, and dendritic cells. They are designated HLA-DP, HLA-DN, HLA-DM, HLA-DQ, HLA-DR, and HLA-DO.[286] **Class III genes** encode several components of the complement system. Many other genes and pseudogenes are interspersed with those of the MHC.

All type I MHC molecules are integral membrane glycoproteins each of which is composed of a 45-kDa heavy chain of about 350 residues together with a noncovalently linked 17-kDa light chain. The genetic variation occurs in the heavy chain between residues 43 and 195 in the human proteins. This chain appears

Figure 31-12 (Opposite page) Arrangement of genes of the human major histocompatibility complex (MHC). Left: Banding pattern of a stained chromosome 6 with the MHC region marked. Center and right: locations of all genes and pseudogenes in this region. The MHC molecules can be divided into three classes on the basis of their structure and function. The class I antigens constitute a single class structurally but fall into two functional groups. The first of these contains the "classical" class I antigens, first discovered as the transplantation antigens and now known to function as target antigens in the recognition and destruction of virus-infected cells by cytotoxic T lymphocytes. They are expressed on virtually all somatic cells. The class II antigens are expressed largely on B lymphocytes and macrophages of the immune system and are essential for presenting antigen to the helper and suppressor T cells that regulate the immune response. Many class III products are components of the complement system. These maps are based on serological and biochemical data, together with complete sequences. From the MHC sequencing consortium.[284]


0
6p21.31
3600
Size (kb)
Extended class I
HFE
BTFs
HISTONEs
RFP
OLFRs
BAGAb
MOG
Classical class I
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
HCGIV-11
P5-15
MICE
3.8-1-5
P5-14
HCGIV-10
P5-13
HCGIV-9
RPL7B
P5-12
HCGIV-9
HLA-F
HCGIX-5
HLA-75
HLA-90
HCGII-8
P5-11
HLA-G
MICF
HCGIII-2
3.8-1-4
P5-10
HCGIV-7
P5-7
P5-9
HLA-16
HCGII-7
3.8-1-3
P5-6
HCGIV-6
P5-5
HLA-70
HLA-21
P5-4
HCGIV-5
HLA-A
P5-3
HCGIV-4
HLA-80
HCGII-6
MICD
HCGIX-4
3.8-1-2
P5-2
HCGIV-3
HCGVIII-1
HLA-59
HCGVII
HTEX4
HCGV
HCGI
RFB30
ZNFB7
ZNF173
HLA-92
CAT75X
GT257
HLA-30
HCGII-5
HCGII-4
MICC
HCGII-3
TC4
HLA-E
HSR1
CAT56
ABC50
FB19
PRO-hom
DBP2
KIAA0170
RPL7A
TUBB
FLOTILLIN
PRG1
DDR
TFIH
S
PG8
SC1
OTF3
HCGIX-3
NOB4
HCGII-2
NOB5
HLA-C
HCGIV-2
KIAA0055-hom
RPL3-hom
HCGII-1
HLA-B
HCGIV-1
DHFRP
HLA-17
P5-8
NOB1-3
HCGIX-2
MICA
HLA-X
P5-1
3.8-1.1
HCGIX-1
MICB
Class III
1800
1900
2000
2100
2200
2300
2400
2500
2600
2700
2800
2900
3000
3100
3200
3300
3400
3500
3600
BAT1
ATP6G
NFKBIL1
LTA
TNF
LTB
LST1
1C7
AIF1
BAT2
BAT3
ApoM
G4
BAT4
CSK2B
G5b
G5c
BAT5
G6f
G6e
G6d
G6c
G6b
DDAH
CLIC1
MSH5
NG23
G7c
Val-TRS
snRNP
HSPA1L
HSPA1A
C4B
GB
NEU
HG22
G9a
HG36
G10
C2
BF
RD
SK12W
DOM3L
STK19
CB4
P450-C21B
TNXB
CREBL1
NG7
NG5
PPT2
NG3
LPAAT
G16
RAGE
PBX2
G18
NOTCH4
Classical class II
TSBP
HNRPA1
B-30.2-L
BTL-II
HLA-DRA
HLA-DRB9
HLA-DRB3
HLA-DRB2
HLA-DRB1
HLA-DQA1
HLA-DQB1
COX3-L
GLN-IRNA
HLA-DQB3
HLA-DQA2
HLA-DQB2
HLA-DOB
TAP2
RING9
LMP7
TAP1
LMP2
RING8
PPP1R2P
RING14
HLA-Z1
HLA-DMB
RING13
HLA-DMA
RING3
HLA-DOA
HLA-DPA1
RPL32-L
HLA-DPB1
HLA-DPA2
COL11A2p
HLA-DPB2
HLA-DPA3
Extended class II
COL11A2
RXRB
RING5
RING2
RING1
ZNF-L
TAT-SF1-L
HSACM2L
RPS18
B3GALT4
BING4
BING5
HKE2
RGL2
TAPBP
BING1
DAXX
BING3
RPL35A-L
RPL12-L
HSET
Transcriptional orientation
Gene
Pseudogene
Gene/pseudogene
Gene loci

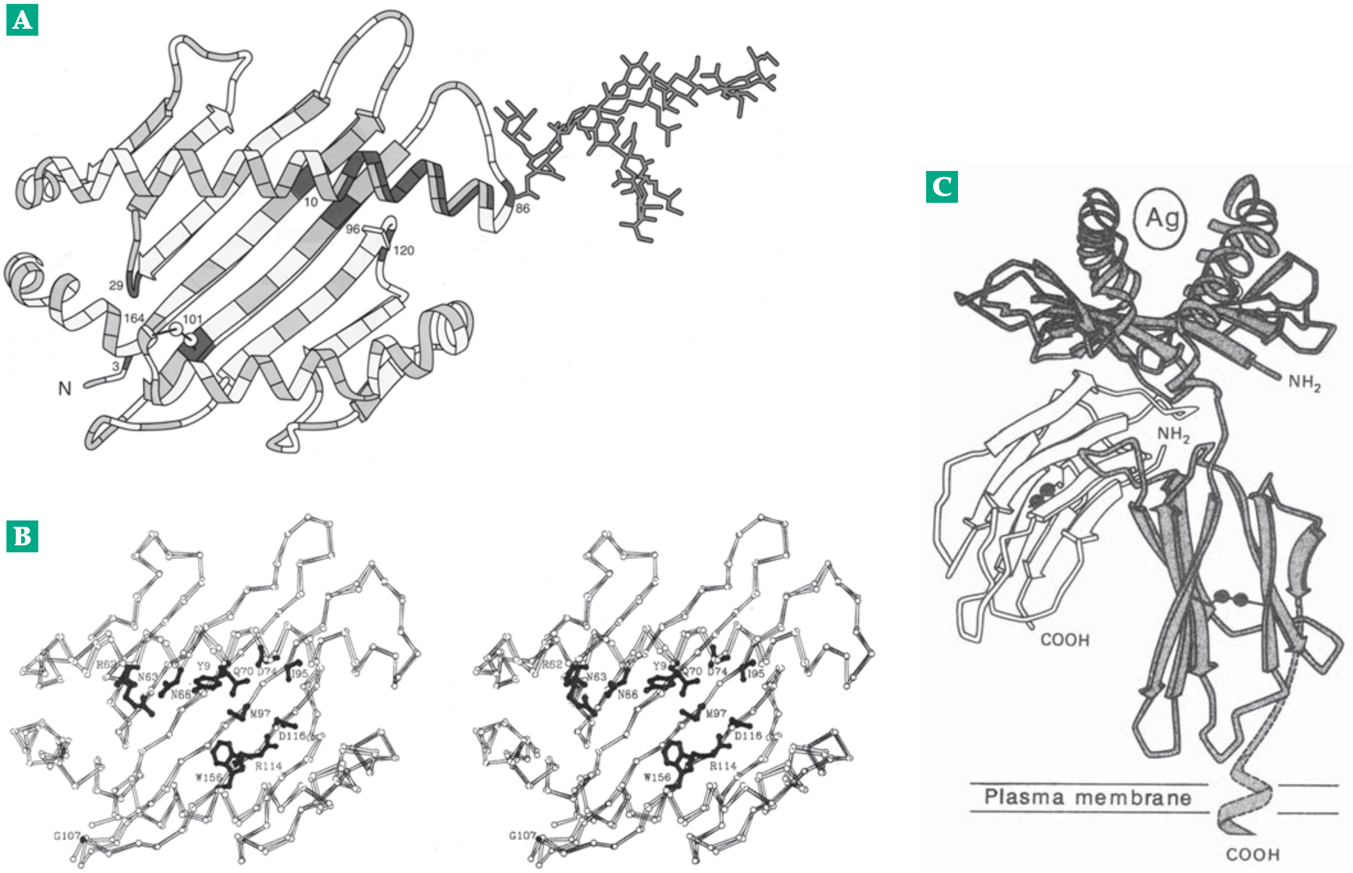


Figure 31-13 The structure of Class I MHC molecules. (A) The specificity pocket in the N-terminal part of the ~360-residue α chain. The numbered residues are invariant in all of the ~20 different Class I molecules. An oligosaccharide is shown on the invariant Asn 86. (B) A stereoscopic view of a similar MHC molecule showing some of the polar residues that protrude into the peptide-binding groove and may form hydrogen bonds with the peptide. From Garrett *et al.*[291] Courtesy of Don C. Wiley. (C) Side view of a complete MHC molecule with an antigenic peptide (Ag) bound into the peptide-binding groove. The C terminus of the long ~360-residue α chain is in the cytoplasm of the displaying cell. The small 99-residue β chain (unshaded) is a molecule of β microglobulin, which is also a constituent of blood plasma.[292] Courtesy of Peter Parham.

to consist of three ~90-residue domains protruding from the outside of the cell, about 25 residues embedded in the membrane, and a short C-terminal tail in the cytoplasm (Fig. 31-13).[287] The light chain has an invariant composition and is identical to the plasma protein β_2-microglobulin, whose gene is located on a different chromosome. Its structure closely resembles that of a single immunoglobulin domain.[288] The MHC Class II antigens (Fig. 31-14) are also αβ dimers, the α chains being 34-kDa glycoproteins and the 28-kDa β chains being larger than in the type I antigen.[283,289] While the MHC of humans and mice have been studied the most, all vertebrates possess similar self-identification systems. Although both the sequences and the folding patterns of the MHC antigens (Figs. 31-13, 31-14) are somewhat similar to those of immunoglobulins, there are many differences. Furthermore, the cause and significance of the polymorphism is quite different in the two cases. Each individual has millions of antibodies with different variable regions but only one set of HLA antigens, which are largely the same on germ, embryonic cells, and adult cells.

Serological tests allow tissue types to be defined by the HLA genes.[290] Thus the commonest HLA type in Caucasian populations is HLA-Al / B8 / Dw3, whereas A1 / B17 is common among Asian Indians. In every case subtypes can be defined, and this fact together with the polymorphism in other genes leads to a unique HLA type for nearly every individual. As is indicated in Fig. 31-12, complete nucleotide sequences are known for typical alleles of all of these genes.[284] It is of medical interest that the susceptibility of an individual to many degenerative diseases is determined in part by the HLA type.[289,295] Thus among patients with a kind of arthritis, **ankylosing spondylitis** that affects 1 or 2 per 1000 men of Caucasian origin, 96% have the HLA-B27 antigen. Of patients with **celiac disease**, a type of intolerance to gluten, 60% have the HLA-B8 antigen. Persons with

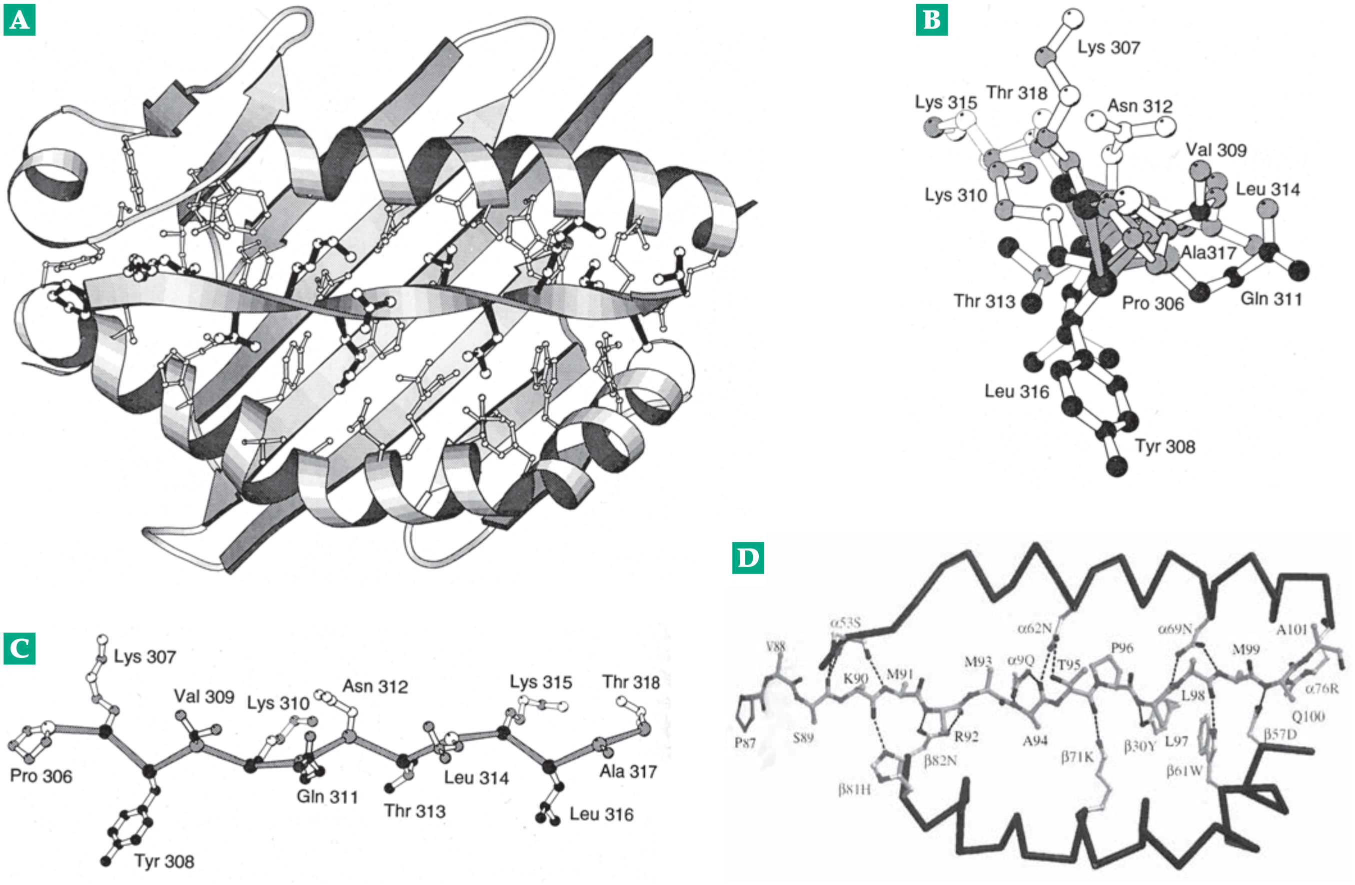

A
B
Lys 307
Lys 315
Thr 318
Asn 312
Val 309
Leu 314
Lys 310
Ala317
Thr 313
Pro 306
Gln 311
Leu 316
Tyr 308
C
Lys 307
Val 309
Lys 310
Asn 312
Lys 315
Thr 318
Pro 306
Gln 311
Thr 313
Leu 314
Ala 317
Leu 316
Tyr 308
D
α53S
V88
K90
M91
M93
α9Q
α62N
T95
P96
α69N
M99
A101
α76R
L98
Q100
β57D
P87
S89
R92
A94
β82N
β81H
β71K
β30Y
L97
β61W

Figure 31-14 Illustration of the binding of a short peptide fragment (central ribbon) into the antigen-binding groove of an MHC type II molecule. (A) The binding groove of molecule HLA-DR1 with a bound peptide (HA) derived from an influenza virus. (B) End view of the bound peptide. (C) Side view of the same peptide. From Stern *et al*.[293] (D) A similar HLA molecule (HLA-DR3) with the peptide CLIP (class II associated invariant chain peptide) bound into the antigen-binding groove. The binding is almost identical to that in (A). Notice the specific hydrogen bonding to side chains of the HLA molecule. From Ghosh *et al*.[294] Courtesy of Don C. Wiley.

HLA-Bw17 and B13 have an increased susceptibility to **psoriasis** and those with HLA-DRw4 an increased tendency to develop **rheumatoid arthritis**. Susceptibility to **autoimmune insulin-dependent (Type I) diabetes** is strongly correlated with the presence of the neutral residues Ala, Val, or Ser at position 57 of the HLA-DQ β chain.[296,297] However, aspartate in position 57 protects against the disease. It may prevent development of a dangerous autoantibody to this cell surface protein. The presence in populations of both humans and apes of a balanced polymorphism among the residues Ala, Val, Ser, and Asp at this position suggests an essential evolutionary origin to this disease susceptibility.[297] HLA-B53 protects against severe malaria in Africa. Other diseases with a strong association with HLA type include multiple sclerosis, **Crohn disease** (inflammatory bowel disease),[298] and several diseases induced by infections with viruses, bacteria, trypanosomes, etc. For example, arthritis can follow infection by *Salmonella*. This suggests that the killer T cells can be confused when stimulated by foreign antigens, which are too closely related to the HLA antigens of the host.

6. Antigen Presentation and MHC Restriction

T cells usually do not respond to intact antigens on cell surfaces but only to partially degraded antigens. Antigen-presenting cells (APCs) of various types must process the antigen through endocytosis and partial digestion before the foreign antigen can bind as an MHC complex to a T-cell receptor. Apparently the processed foreign antigen must lie in the binding site of an MHC Class I or Class II chain (Figs. 31-13C, 31-14) before the T cell will recognize the complex and respond.

Antigen processing begins with cytosolic proteasomes that are present in all cells (Box 7-A; Chapter 29, Section D,8). They cleave proteins of the cell and of intracellular parasites into short peptide fragments,[55,299–301] which may need to be trimmed to shorter 8- to 11-residue peptides suitable for binding into the groove in an MHC Class I molecule.[302,303] The peptides are carried into the ER with the aid of the **TAP** (transporter associated with antigen processing) complex,[304–305a] which is discussed briefly in Chapter 29, Section D,3. Chaperones, such as hsp70, may also participate in the transport. In contrast, MHC Class II proteins receive their antigenic peptides via an endosomal–lysosomal pathway. Proteins from phagocytized pathogens are cleaved by proteases in an endosome or lysosome into fragments that tend to be longer (13–25 residues) than the 9- to 11-residue peptides generated by proteasomes.[299,306]

Peptides bind into the groove in a Class I MHC molecule in a manner similar to that illustrated in Fig. 31-14 for a Class II MHC molecule. However, in the Class II complex the longer peptides extend from the two ends of the binding groove. The peptide, which assumes a polyproline II helical conformation, is held by hydrogen bonds from the Class II MHC molecule to the peptide backbone.[307–311] A single peptide may shift and bind in a different register with the possibility of being recognized by a different T cell receptor, when it is displayed.[308] In contrast, MHC Class I molecules bind best to 8- or 9-residue peptides, which are held by an array of hydrogen bonds to the $-NH_3^+$ and $-COO^-$ termini (not shown in Fig. 31-13, which displays the empty binding groove). The central part of this groove contains a deep pocket, which together with smaller pockets near the ends provides specificity.[312–318] However, there is a puzzle. Because of the great genetic variability in the MHC genes there will be great differences (polymorphism) in the shapes of the binding pockets among different people. However, an individual has at most six different kinds of MHC molecules. Yet, a single MHC molecule has been estimated to be able to bind more than 10,000 different peptides.[314] Essential to this process is a final trimming of the peptides at their N termini to provide better fits.[318a]

Peptides are loaded onto Class I MHC molecules, while they are together in the ER. They move as tightly bound complexes through the Golgi and into the external plasma membrane, where they remain tethered via the MHC molecule (Fig. 31-13C). In contrast, binding to Class II MHC molecules occurs in the endosomes or lysosomes. The process is somewhat complex. Class II MHC molecules are chaperoned from the cytosol into late endosomal/lysosomal organelles with their antigen-binding grooves occupied by a peptide fragment known as the Class II-associated invariant chain peptide (CLIP; Fig. 31-14D). This is cut from the end of an invariant chain known as Ii.[319–321]

The MHC proteins HLA-DO and HLA-DM (Fig. 31-12) are resident in the lysosome-like organelles and chaperone the class II molecules, until they are ready for loading with peptides. HLA-DM assists in removal of the CLIP peptide when loading occurs.[322] An asparaginyl endopeptidase may also be required[323] as well as a disulfide reductase.[324] After loading the Class II MHC•peptide complexes, like the Class I MHC•peptide complexes, are exported to the plasma membrane. There they may be recognized by a T cell, which utilizes its T-cell receptor to recognize an antigen and its CD4 or CD8 proteins to distinguish Class I from Class II MHC complexes. The binding of CD4 and CD8 to their cognate MHC molecules has also been described at the molecular level.[310,325] As mentioned previously, some hydrophobic antigens are presented by CD1 molecules. They also have an MHC-like fold with a large hydrophobic binding groove.[272]

An interesting approach to the treatment of autoimmune diseases is design of peptide mimics that bind into the antigen-binding groove of specific MHC proteins. For example, a protease-resistant pyrrolinone–peptide hybrid has been designed to bind to the rheumatoid-arthritis-associated HLA-DR1.[326]

An important distinction between B- and T-cell responses is that T cells recognize a foreign antigen only when associated with an MHC antigen of the same type as is carried by the T cell. This "MHC restriction" limits the actions of T lymphocytes. The function of cytotoxic T cells appears to be primarily one of killing virus-infected cells and perhaps cancer cells. MHC restriction ensures that the T cell is attached by its rereceptor to a Class I MHC molecule belonging to self. The dual recognition ensures that the cell probably contains a foreign antigen and should be killed. The Class I MHC antigens are found on almost all body cells. Therefore, cytotoxic $CD8^+$ T cells can attack infected cells of all types. They may kill their target cells by injecting their cell membranes with the complement C9-like cytolysins.

Regulatory $CD4^+$ T cells recognize the Class II MHC molecules, which are found primarily on lymphocytes. Thus, the attention of regulatory T cells is directed towards other lymphocytes. In this case, too, MHC restriction enables helper T cells to recognize B lymphocytes as self. If foreign antigen is present so that B lymphocytes have been activated by the binding of a foreign antigen, they will be stimulated by the T_H cells to proliferate and make and release immunoglobulins. This is accomplished in part by secretion of the lymphokine interleukin-2 (Fig. 30-6) and B-cell growth factors. Some T cells become suppressor T cells.

7. T-Cell Receptors

T cells both mediate the recognition of self and also participate in the immunologic response to foreign antigens. Their surface receptors function much like the immunoglobulins that are attached to the surfaces of B cells. The T-cell receptors are αβ disulfide-linked heterodimeric glycoproteins (Fig. 31-15) consisting of 40- to 45-kDa α subunits and 42- to 44-kDa β subunits.[267,327–330a] They are associated in the T-cell membrane with a larger complex called CD3, which contains additional 26- to 28-kDa γ, δ, ε, and ζ chains.[267,331,332] The polypeptides of the CD3 complex have C-terminal cytoplasmic tails that contain tyrosine residues within several immune system tyrosine-based activation motifs (ITAMs; Fig. 31-15). As is illustrated in this figure, the activating antigen is cradled in the binding groove of an MHC molecule attached to the APC (top) with some side-chain and backbone atoms of the peptide available for bonding to the T-cell receptor.[333,334] Notice that in the synapse the T-cell receptor also makes direct contact with the MHC molecule (Fig. 31-15).

Signaling by an activated T-cell receptor is quite complex. The ITAMs are sites of tyrosine phosphorylation by kinases of the Src family.[335,336] Another tyrosine kinase, Zap-70 (Fig. 31-15), associates with the C-terminal tails of the disulfide-linked dimer of subunits ζ. It recognizes the phosphotyrosines groups via its SH2 domains (see Fig. 7-30). Zap-70 appears to act synergistically with the Src kinases.[332,337,338] The nature of the APC is also of importance. For example, some dendritic cells secrete Il-12, which favors formation of T_H1 helper cells. A second type of dendritic cell induces formation of T_H2 helper cells.[338a] Dendritic cells may also control growth and proliferation of T cells by regulating the availability of cysteine, which is a nutritional essential for lymphocytes.[338b] Other effects may result from endocytosis by T cells of occupied T-cell receptor•MHC complexes.[338c]

A second type of T-cell receptor, the γδ receptor, is carried by a small subgroup of T cells. It may have a distinct role in generating an immune response to certain microorganisms including *Mycobacterium tuberculosis*.[339–342a] Like immunoglobulins, T-cell receptors have a great variety of amino acid sequences. They have C-terminal constant domains and N-terminal variable and hypervariable regions. Thus, T cells can bind to a variety of foreign antigens. However, except as a result of autoimmune diseases, they do not attack cells recognized as self unless these cells are infected with a virus or for some other reason carry foreign surface antigens.

Certain bacterial immunostimulatory molecules known as **superantigens** are able to stimulate MHC Class II molecules to activate large numbers of T cells without any assistance of an antigenic peptide, sometimes with disastrous results.[343,344] Superantigens are a group of related proteins that includes enterotoxins from species of *Staphylococcus*[344–346] and *Streptococcus*,[347] a staphylococcal exfoliative toxin,[348] and **toxic shock syndrome** toxins.[349] Superantigen molecules don't occupy the peptide-binding groove in the MHC molecule but bind as intact proteins at an external site. They also bind to the variable region of the TCR in the MHC-TCR complex. Attempts are being made to design decoy molecules that prevent binding of a particular superantigen.[350]

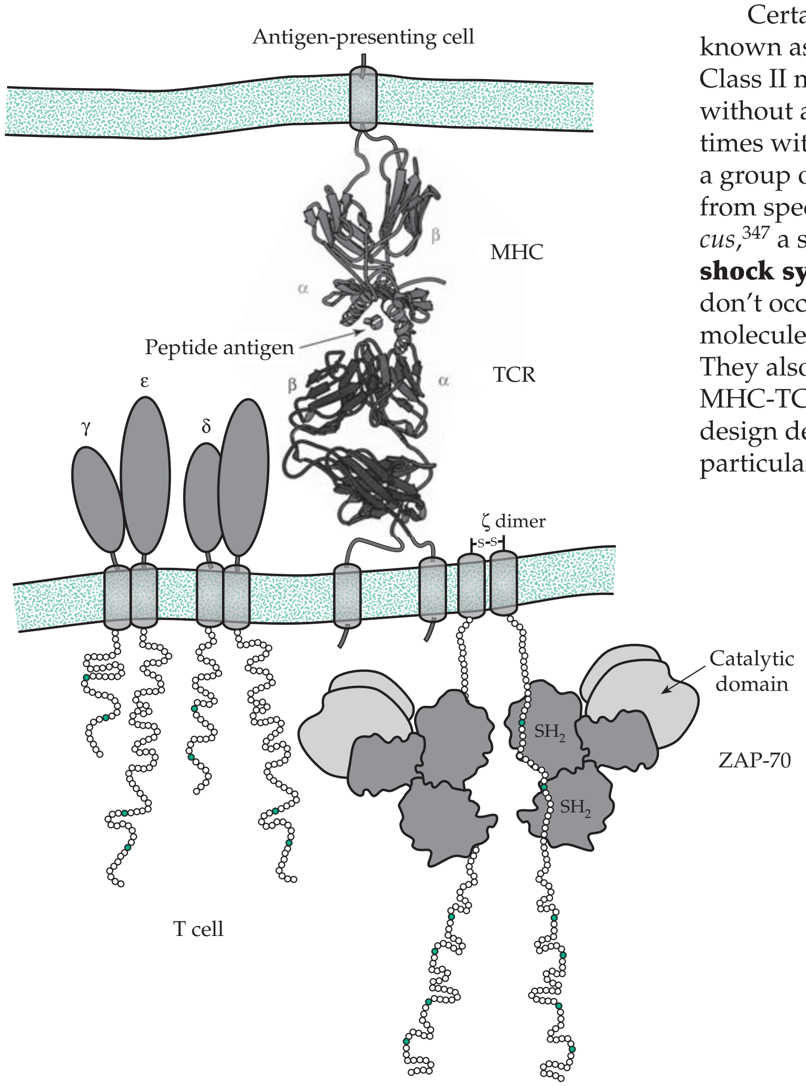


Figure 31-15 Interaction between an MHC•peptide complex on an antigen-presenting cell (APC) with a T-cell antigen receptor (TCR) that is attached to the plasma membrane of a T cell. Structures of the α- and β-subunits of the MHC molecule and of the T-cell receptor are based on crystallographic data. The detailed structures of the disulfide-linked σε and γε modules of the T-cell receptor are not shown. The dimeric ζ_2 subunit has large cytoplasmic domains that are thought to be involved in signaling the protein ZAP-70 (zeta-associated protein of M_r 70,000). The σ, ε, and γ subunits may also undergo phosphorylation of their tyrosyl groups (gray spheres), which are found in immune system tyrosine-based activation motifs (ITAMS). Drawing modified from those of Cochran *et al.*[267] and Hatada *et al.*[332]

8. Self-Tolerance

The immune system is flexible enough and powerful enough to protect us from a great variety of dangers, even from viruses and organisms that may be entirely new. However, it is hard to understand how the immune system completely avoids fatal damage to our own bodies. The answer is complex. It has baffled generations of investigators[350a,b] and is still not fully understood.

Since the discovery of vaccination in 1796 immunology has claimed the attention of many scientists. However, it was not until 1891 that the German bacteriologist Emil von Behring proposed the term antibody for the protective materials in blood.[351] By about 1900 Paul Ehrlich and Svante Tiselius, who wrote the first immunochemistry book,[352] initiated serious investigations. Ehrlich proposed that binding of an antigen to a surface receptor would induce the cell to make additional identical surface receptors, which would be released to become antibodies.[351,353] The concept was correct, but it would be many years before knowledge of the structures and biosynthesis of antibodies became available.

Nevertheless, immunological tolerance interested Ehrlich and other immunologists.[353a] One proposal, offered by Niels Jerne in 1974, was that self-tolerance depends upon **immunological networks**.[354–356] Consider a lymphocyte bearing a bound immunoglobulin receptor or a bound T cell receptor. It will be specific for some epitope E. The receptors are shown in Fig. 31-16 as having V-shaped antigen-binding sites. Jerne pointed out that the variable region of this receptor will itself carry epitopes that can be recognized by other appropriate antibodies. These epitopes on the receptor are called **idiotopes** and as a group define the **idiotype** of that receptor or immunoglobulin.[356a] There will usually be other lymphocytes with receptors that recognize the idiotype of the first lymphocyte (see Fig. 31-16). These constitute an **antiidiotype set**. In addition there will be lymphocytes, whose idiotopes resemble those of epitope E and which will therefore be recognized by the first lymphocyte as foreign. These lymphocytes constitute an **internal image set**. There will be other sets of lymphocytes that recognize the lymphocytes of the antiidiotype set or of the internal image set. Thus, there will be an elaborate network of clones of interacting lymphocytes. When an immune response occurs many members of this network will respond. A B lymphocyte will recognize a particular antigen and gives rise to a clone of plasma cells making antibodies against that antigen. The body will then make new antibodies against the first antibodies formed, etc. A whole segment of the network will respond in this fashion. Jerne suggested that the overall effect would be to limit and suppress the immune response. Anti-idiotype antibodies as well as anti-anti-idiotypic antibodies have been prepared[356] and have been used in studies of receptors.[357] However, Jerne's theory is generally regarded as incorrect or at least a great oversimplification.[358]

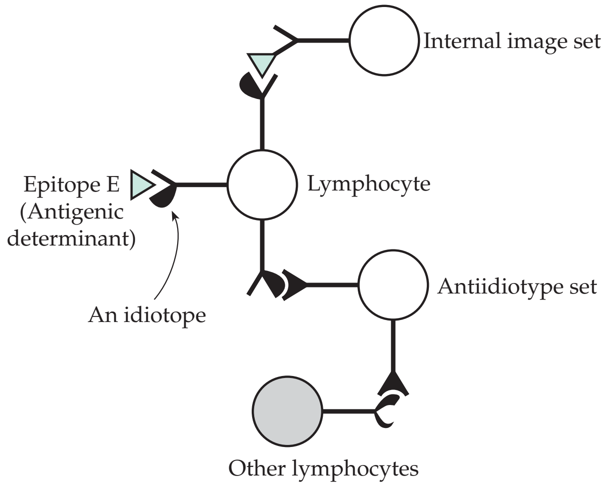


Figure 31-16 Schematic depiction of lymphocyte receptors forming anti-idiotype and internal image sets as proposed by Jerne.

A process of **clonal selection**[359] is now thought to be basic to self-tolerance. Credit for the theory, developed in the 1950s, is usually given to F. M. Burnet. However, Ehrlich, Jerne, and David Talmage were also prominent contributors.[37,351,359,360] An essential postulate of the clonal selection theory is that each B lymphocyte is predetermined to make antibodies of only a single specificity. The mechanism of **allelic exclusion**,[359] which makes this possible, is described in Section E. Clonal selection can occur because the B cells carry their antibodies as surface receptors. Binding of an antigen provides a signal for clonal expansion. However, during development in the thymus the progenitor B cells carrying self-reactive antibodies are killed by apoptosis. Later, peripheral B lymphocytes also undergo selection by a complex network of signaling and apoptosis.[361–363] In a similar manner excess lymphocytes that build up during an immune response must be removed.[364]

The learning of self by T lymphocytes also happens in the thymus early in development, during the first three weeks of life in mice.[365,366] This again involves selection against potentially autoreactive lymphocytes carrying idiotypes that are also present on the body's own tissues and which have a high affinity for self peptide•MHC complexes. However, T cells with a weaker affinity for a self MHC molecule but a high potential affinity for a nonself peptide are allowed to develop.[366,367] Only about 1% of the lymphocytes that develop in the thymus emerge as mature T cells.[368,369] Others appear to be killed (clonal deletion) or to become unreactive toward antigen.[353a] The latter enter a state referred to as **anergy**.[370–373]

As with every aspect of metabolism, homeostasis is essential to the immune system, which must be able to both grow and shrink rapidly.[373a–c] Antigens, cytokines, apoptosis-inducing signals, immune inhibitor receptors,[229h] and receptor tyrosine kinases[373d] all participate in preserving the delicate balance that is required.

9. Immunologic Memory and Vaccination

In 1796, Edward Jenner carried out the first human vaccination. Attempting to protect a teenaged boy from disfigurement and possible death from smallpox, he vaccinated him with material from a cowpox lesion on the hand of a milkmaid. (She had contracted the disease from a cow named Blossom, whose hide hangs in St. George's Hospital in London.)[353,374,375] Six weeks later he inoculated the boy with virulent smallpox. Fortunately, the boy didn't contract the disease. Today vaccination is in use for more than 70 bacteria, viruses, parasites, and fungi, and the results have been impressive.[376] Poliomyelitis has been almost eliminated.[377] Smallpox has not been seen for many years, and the decreases in diphtheria, measles, mumps, whooping cough, and rubella have been impressive.

A nagging question is "How long will vaccination last?" One unplanned experiment resulted from two epidemics of measles in the remote Faroe Islands. The first outbreak was in 1781 after which the islands remained free of measles for 65 years. The second outbreak in 1846 affected 75–95 % of the population, but according to a physician who investigated the epidemic not a single one of the many aged people living who had had measles in 1781 contracted the disease a second time.[38] This bit of history confirms that immunological memory is sometimes very long-lived (although it doesn't prove that the smallpox vaccination older people received is still good!).[377a]

Vaccines have been prepared traditionally by use of viruses or organisms killed by compounds such as formaldehyde or by attenuated viruses or live organisms. These are selected for a low degree of virulence after repeated passages through live animals or cell cultures.[376] Newer methods utilize purified viral proteins, bacterial capsular polysaccharides, or DNA.[378–379a] In the future edible vaccines may be produced in plants.[380] Nevertheless, it is often very difficult to devise effective vaccines. In spite of 80 years of effort better vaccines against tuberculosis are needed.[381,382] All efforts to produce an AIDS vaccine have failed.[375,379] A satisfactory vaccine must activate both B cells and T cells. Activation of the latter may be especially difficult. Continuous development of new strains of bacteria is a problem for vaccination against tuberculosis and has been an insurmountable barrier to vaccination against AIDS. However, in the latter case it may be possible to use vaccination to prevent an HIV infection from progressing to AIDS with complete destruction of $CD4^+$ T cells.[379] As the immune system and also diseases become better understood, it is possible that new strategies for induction of specific cytolytic T cells can be devised,[383] e.g., for AIDS[384] and even for cancer.[385] A current obstacle to development of new vaccines is that pharmaceutical companies view vaccines as unprofitable.[385a]

We need a better understanding of how memory B and T cells are formed and selected for long-term survival.[386] After differentiation and selection in the bone marrow, B cells move to the spleen. These **transitional B cells** within germinal centers undergo further selection to become mature B cells and B memory cells.[256,387,388] Interactions with cytokines and with coreceptors play important roles. A subset of B lymphocytes may remain in the germinal centers, serving as a kind of stem cell providing new memory B cells continuously.[39] **Naive T cells**, which have not yet encountered antigen, travel throughout the body but apparently don't enter nonlymphoid tissues. However, after being presented with antigen by dendritic cells within lymph nodes, some T cells move to the skin and other peripheral locations.[389,390] After a pathogen is destroyed most of these T cells die, but a few remain as long-lived memory cells. These are able to respond to a second encounter with a pathogen.[40,383,391] Apparently continuous new exposure to antigen is not needed for long-term immunity, but the slowly dividing $CD8^+$ memory T cells may require continuous stimulation by Il-15 to counteract inhibition by Il-2.[392]

E. The Rearranging Genes for Immunoglobulins and T-Cell Receptors

An impressive example of the kind of permanent changes that can occur in the genome of specialized cells is provided by the genes of immunoglobulins and T-cell receptors. B lymphocytes make tens of millions of antibodies of differing sequence, and T lymphocytes make a similar number of different T-cell receptors. This diversity is established in major part in the DNA but also by alternative splicing and editing of RNA.

1. Rearrangements of Germline DNA

Each ~110-residue domain of an immunoglobulin is encoded by a single exon, but the exon for the amino-terminal or variable domain is assembled from two or three small genes (or segments) selected from a large family of such genes present in the germ cells and in the lymphoid progenitor cells.[393–396] Within the

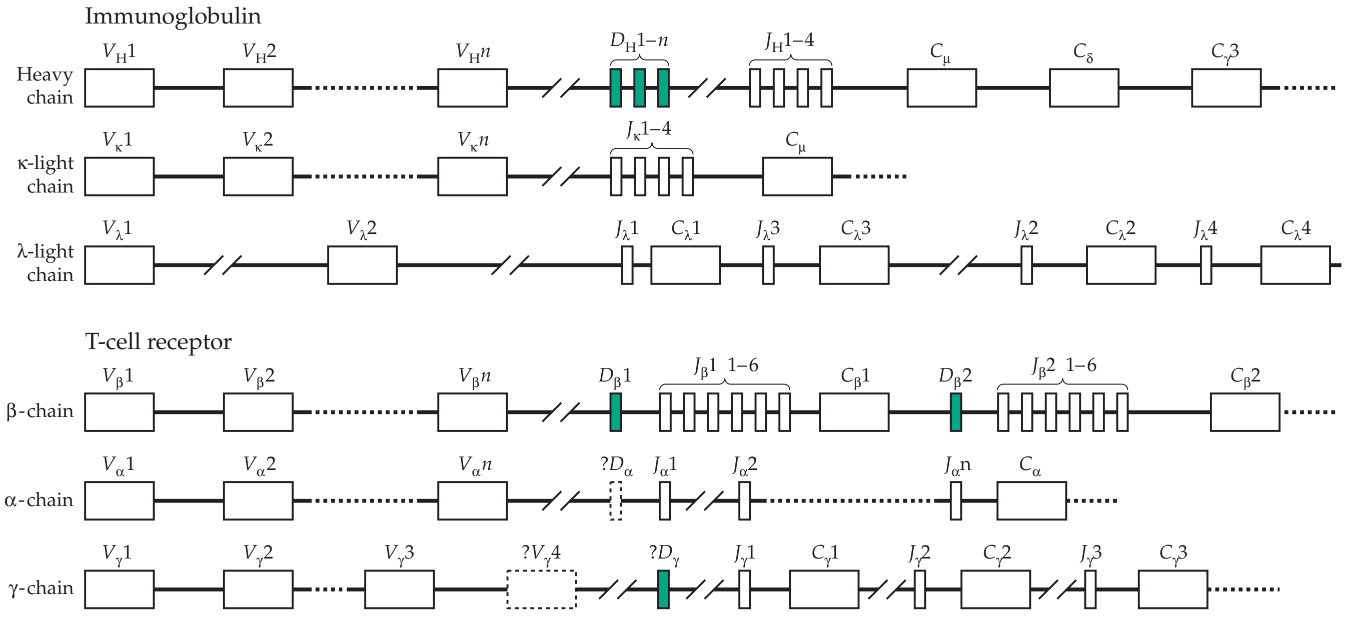

Immunoglobulin
Heavy chain
V_H1
V_H2
V_Hn
$D_H1–n$
$J_H1–4$
C_μ
C_δ
$C_\gamma3$
κ-light chain
$V_\kappa1$
$V_\kappa2$
$V_\kappa n$
$J_\kappa1–4$
C_μ
λ-light chain
$V_\lambda1$
$V_\lambda2$
$J_\lambda1$
$C_\lambda1$
$J_\lambda3$
$C_\lambda3$
$J_\lambda2$
$C_\lambda2$
$J_\lambda4$
$C_\lambda4$
T-cell receptor
β-chain
$V_\beta1$
$V_\beta2$
$V_\beta n$
$D_\beta1$
$J_\beta1$ 1–6
$C_\beta1$
$D_\beta2$
$J_\beta2$ 1–6
$C_\beta2$
α-chain
$V_\alpha1$
$V_\alpha2$
$V_\alpha n$
$?D_\alpha$
$J_\alpha1$
$J_\alpha2$
$J_\alpha n$
C_α
γ-chain
$V_\gamma1$
$V_\gamma2$
$V_\gamma3$
$?V_\gamma4$
$?D_\gamma$
$J_\gamma1$
$C_\gamma1$
$J_\gamma2$
$C_\gamma2$
$J_\gamma3$
$C_\gamma3$

Figure 31-17 Organization of the immunoglobulin and T-cell receptor gene families of the mouse. The human γ-gene pool is larger as is the λ light-chain immunoglobin gene pool. All six gene pools contain separate gene segments encoding the variable and constant regions of the antigen receptors of lymphocytes. In the course of lymphocyte development, one of the V segments is juxtaposed by chromosomal rearrangement with one of the J segments and, where applicable, a D segment to form a complete variable-region gene. Each V segment has two regions of hypervariability, which are known in the case of the immunoglobulins to contribute to the antigen-binding site in the folded molecule. A third hypervariable region, which also contributes to the antigen-binding site, is generated by the junction of the V segment with the J or the D and J segment(s) at recombination. There is more germline diversity in the T-cell receptor than in the immunoglobulin gene pools in the J segments. The β-gene pool is organized in a way that also allows more combinatorial diversity. The β-gene pool contains fewer J regions than the gene pool, but it has two distinctive features that allow for exceptional diversification during somatic rearrangement. First, the ‘rules’ for recombination allow in principle the joining of both D to D segments and V segments directly to J segments, neither of which is possible in the immunoglobulin heavy-chain pool. The D segments of the β genes can be read in any of the three possible reading frames, so that varying the site of the V – D junction alone can make a substantial contribution to the diversity of the third hypervariable region. From Robertson.[397]

V region the three short hypervariable segments alternate with four framework segments that have a more nearly constant structure. The V region of the light-chain genes (of either the κ or λ type) is put together in part from a V gene that encodes an approximately 95-residue sequence making up the first three framework regions plus two hypervariable regions and part of the third hypervariable segment. There are ~100 different V_κ genes and ~30 V_λ genes[398–400] in the light-chain family. The similar arrangement of genes in the mouse is indicated in Fig. 31-17. These V genes are spaced at intervals of 14–30 kb within the DNA.

The rest of the third hypervariable region and the fourth framework region of the κ and λ chains are encoded by short **J (joining) genes** that specify ~15 residues. There are about five J_κ and three J_λ genes in the mouse. The first mechanism for creating antibody diversity lies in the large number of V genes, which are especially diverse in their hypervariable regions. The second mechanism is the joining of any one of these V genes with any one of the J genes, the joining taking place within the third hypervariable region. There are also at least six different constant (C_λ) genes in the human genome.

The heavy-chain genes are more complex. There are over 200 V_H genes located at 14- to 16-kbp intervals, which are followed by ten ~15–17 bp **D (diversity) genes** and 4–6 J_H genes. During the differentiation of the lymphoid stem cells pre-pro-B cells can be identified in which the V, D, and J genes are still separate. Later there are pro-B cells containing joined DJ_H segments, then pre-B cells with joined VDJ_H segments, then B cells with a VJ_L (either κ or λ) segment also joined. The overall process of gene rearrangement, mRNA splicing, and immunoglobulin synthesis is outlined in Fig. 31-18. The joining of the gene segments occurs by recombinational mechanisms that involve 7-bp and 9-bp recognition sequences known as **recombination signal sequences** (RSSs)[401–402b]:

$$\begin{array}{cc} 5'\text{-CAC}\underset{\bullet}{\text{T}}\text{GTG} & 5'\text{-GGTTT}\underset{\bullet}{\text{T}}\text{TGT} \\ 3'\text{-GTG}\overset{\bullet}{\text{A}}\text{CAC} & 3'\text{-CCAAA}\overset{\bullet}{\text{A}}\text{ACA} \\ \text{heptamer} & \text{nonamer} \end{array}$$

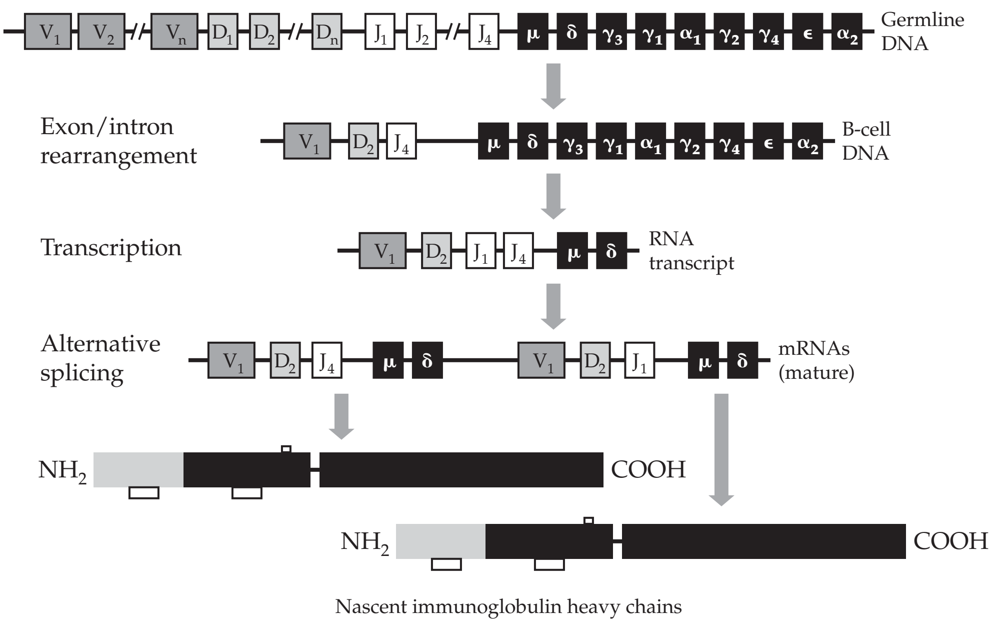

V1 V2 Vn D1 D2 Dn J1 J2 J4 μ δ γ3 γ1 α1 γ2 γ4 ϵ α2
Germline DNA
Exon/intron rearrangement
B-cell DNA
Transcription
RNA transcript
Alternative splicing
mRNAs (mature)
NH2
COOH
Nascent immunoglobulin heavy chains

Figure 31-18 Human immunoglobulin heavy-chain gene structure and gene processing. Exons of the heavy-chain genes that encode the variable regions of the immunoglobulin molecule are labeled V_1, V_2, ... V_n. Selection from these V exons during embryonic development produces the unique sequences of each B-cell clone. The germline genes for the immunoglobulin heavy chains also contain diversity exons, labeled D_1, D_2, ... D_n. Recombination between the V and D regions occurs more frequently than that between the V and J exons in the light-chain exons. Introns between the V and D and between the D and J exons contain signal sequences that regulate synthesis of the Rag1/Rag2 recombinase. This enzyme is responsible for the efficient recombination that gives rise to the epitope-specific B-cell clones with their individual Ig genes. The heavy-chain genes contain exons that encode all of the isotype heavy chains. Class switching, i.e., the change in chain expression that occurs during antibody synthesis after B-cell activation, results from alternative splicing between the J exons and the exons for the various heavy-chain isotypes. Redrawn from Bhagavan.[403] Courtesy of N. V. Bhagavan.

The heptamers have twofold rotational symmetry. The recognition sequences adjacent to one or both sides of the coding segments are separated by 12- or 23-bp spacers, and an empirical rule states that joining can occur only when one pair of recognition sequences are separated by 12 bp and the other by 23. Recombination involves cleavage and rejoining of the DNA, a process also used for reorganization of the T-cell receptor genes. The cleavage is catalyzed by a complex of two proteins RAG1 and RAG2, which are encoded by the adjacent **recombination activating genes** *RAG1* and *RAG2*.[402–406] Cleavage is assisted by HMG chromatin proteins (Chapter 27, Section A,4) and by the level of histone acetylation in an associated enhancer.[407] In the presence of a divalent metal such as Mg^{2+} or Mn^{2+} a single-stranded nick is made between the 3'-end of the DNA coding region and the heptamer sequence. The released 3'–OH group then attacks the phosphodiester bond on the opposite strand, cleaving it in a transesterification reaction. This leaves the coding sequences capped by a hairpin end and a blunt cut end on the RSS.[408] The cut ends are apparently held until synapsis with a nonamer RSS further in the 3' direction and rejoining can occur.[409] Rejoining requires DNA ligase IV.[410] The process resembles that of transposition (Chapter 27, Section D,4).[411,412] The details are still uncertain.[408,413,413a]

The recombinational events are not entirely regular.[401,404,414,415] There are excisions of pieces of DNA, and additional nucleotides may be inserted randomly through the action of **terminal deoxynucleotidyl transferase**.[416,417] All of these mechanisms lead to additional diversity. After the rearrangements are completed, the fused VDJ and VJ gene segments are close enough to suitable constant-region (C) sequences that when the genes are transcribed the intervening sequences in the RNA are spliced out to yield the mature mRNAs for light κ and λ chains and H chains.

2. Somatic Hypermutation and Affinity Maturation

Yet another factor that introduces immunoglobulin diversity is the occurrence of somatic mutation at unusually high rates (**hypermutation**) in the hypervariable regions. Hypermutation consists largely of point mutations in the V-region gene segment and occurs at a rate 10^5- to 10^6-fold higher than for the rest of the genome.[51,418–423] It occurs after B cells have been presented with antigen by dendritic cells in the germinal centers in a process called affinity maturation. The result is generation of an enlarged repetoire of B cells, some of which synthesize antibodies with an increased affinity for the antigen. These will be selected for clonal expansion.[51] Further genetic alteration occurs by **gene conversion** (see p. 1566), which involves copying from a homologous segment of DNA, perhaps from a nearby pseudogene.[423a,b] Some experiments suggested that RNA editing (p. 1642) may also contribute to antibody diversity.[29,424,424a] However, it appears that the observed deamination of cytosine to uracil rings occurs at the DNA level and is initiated by an **activation-induced cytosine deaminase**.[424b] The generation of uracil, which may be removed from the DNA by uracil-DNA glycosylase (p. 1579), apparently triggers both somatic hypermutation and gene conversion. Both processes also depend upon DNA repair via homologous recombination or nonhomologous end joining (p. 1581). The error-prone DNA polymerase τ is also needed.[424c] Class-switching recombination, discussed in the next section, is apparently also initiated by the activation-induced deaminase.[424d] These are affected by **RNA editing** (Chapter 28).[29,424] **Receptor editing**, gene rearrangements that occur in the peripheral immune system, also contributes to affinity maturation.[424,425]

It is important that the genes as finally assembled maintain correct reading frames so that a potentially useful antibody can be made.[426] Although lymphocytes are diploid and therefore contain two sets of immunoglobulin genes, a single cell produces only one kind of immunoglobin (**allelic exclusion**). When the genetic rearrangements produce a light chain able to combine with a heavy chain to form a functional immunoglobulin, a signal may be sent that stops the rearrangements in the other chromosome.[427] However, recent findings suggest that one allele may be marked for inactivation early in development, just as one X chromosome becomes inactive in females (Chapter 32).[428]

3. Immunoglobulin Class (Isotype) Switching

A newly matured B cell produces initially IgM bound to its surfaces. The difference between the C_H domains of bound and secreted antibodies of a given type seems to lie in alternative splicing of the mRNA. More mysterious are the consecutive switches from IgM to other types in the following sequence:

$$\text{IgM} \rightarrow \text{IgG} \rightarrow \text{IgE} \rightarrow \text{IgA}$$

This is the same sequence in which the C_H genes lie (Fig. 31-18). However, the C_δ gene, which codes for IgD, is not utilized in this sequence. A newly matured B cell transcribes the C_μ gene to give IgM. Later, the class switch occurs, apparently by a looping-out recombinational mechanism, allowing a C_γ gene (for IgG) to be expressed.[429–433] The switches are mediated by tandemly repeated DNA sequences in 1- to 10-kbp switch regions and are controlled in part by cytokines. Other complexities are involved in synthesis of the J chain of IgM[434] and in synthesis of IgD. Although the latter is a major surface immunoglobulin on B lymphocytes, its exact functions have been hard to understand. The δ exon of its heavy chain gene is joined to a J exon by alternative splicing of the mRNA (Fig. 31-18).[112] In a similar manner, the difference in the surface-bound and soluble forms of IgM arises by alternative mRNA splicing.[403]

The T-cell receptor gene families are also indicated in Fig. 31-17. Their development is remarkably similar to that of the immunoglobulin genes and involves most of the same mechanism of diversification[365,397,435] with the exception of somatic hypermutation. The same recombinase may cut the DNA to initiate rearrangements of all of these gene families.[436] A single T-cell precursor may give rise to 1000 or more clones with unique β-chain sequences.[437] With a total of ~10^6 different β chains there are potentially ~10^{15} unique T-cell receptor structures that could arise from the 42V and 61J segments of the α-chain gene and the 47V, 2D, and 13J segments of the human β-chain gene.[438] Allelic exclusion is observed, as with the immunoglobulins.[439]

F. Disorders of the Immune System

Many things can go wrong with a system as complex as the human immune system. In immunodeficiency disease some component is missing or has been inactivated. In autoimmune diseases the immune system attacks some component of the body. Of the known problems none is more common than **allergy**,[440–442] which may be described as the inappropriate activation of the immune system by environmental antigens (**allergens**).

1. Allergy

One in 10 persons, ~22 million people, in the United States have allergies. Ten million of these suffer from the nasal discomfort of “hay fever” and six million from the more serious **asthma**. Substantial numbers of people in the United States die of allergic reactions to insect stings (more than 30 per year) or to injections of penicillin (300 per year in 1970). Foods, drugs, pollens, mold spores, mites in house dust, and even heat or cold can evoke serious allergic reactions. Among these **eczema** (atopic dermatitis) is very common. A major cause of allergic reactions has been traced to molecules of immunoglobulins IgE, which bind to the **basophils** in the blood and to the related **mast cells** of tissues. Binding of an antigen to these IgE molecules activates them. These activated antibodies bind (as in Fig. 31-6) to the α subunits of the $F_{c\epsilon}RI$, a transmembrane receptor on basophil or mast cell surfaces.[108,109,443] If two or more IgE molecules bind to a mast cell, they may aggregate and activate the mast cell to release its histamine-containing granules.[444,445] The granules also release cytokines and arachidonate, which is converted primarily into prostaglandin D_2 (Fig. 21-7) and into products of the 5-lipoxygenase pathway (Fig. 21-8). The products include the chemotactic leukotriene B_4 and leukotrienes C_4 and D_4. The latter two constitute the slow-reacting substance of anaphylaxis (Fig. 21-8). The result is a rapid inflammatory response with dilation of blood vessels, increased vascular permeability, infiltration of leukocytes, and destruction of tissues.

What is the normal function of IgE and the mast cells? These cells are located at places where parasites might enter tissues. IgE is involved in killing of schistosomes, and elevated IgE levels are seen in patients infected with various parasites. The killing of schistosomes seems to be mediated by blood platelets as well as by neutrophils and eosinophils with the help of mast cells.[446,447] Allergic persons often have an IgE level over ten times normal. This high level makes the individual especially sensitive to IgE-mediated reactions, a condition called **atopy** (meaning "strange disease").[448,448a] Allergies may also be accompanied by increased B cell levels.[442] This can sometimes be responsible for the sudden and sometimes fatal systemic reaction of **anaphylaxis**. T-cell responses may also cause anaphylaxis.

Most allergy-inducing antigens are proteins, but proteins vary widely in their antigenicity. Only a few natural proteins are major allergens. Many of these are relatively small, with molecular masses of 5–50 kDa. Most are soluble, and some are glycoproteins.[449] Mites are the closest animals to human life and carry allergens that are among the most important causes of asthma and allergic dermatitis. The allergens are 125- to 129-kDa proteins crosslinked by three disulfide bridges.[450,451] Cockroaches[449,452] and other insects also form many allergens. Among them are the hemoglobins of small flies of the chironomid family.[453] Studies of the latter suggest that antigenicity may arise both from flexible regions and also from the presence of a preponderance of amino acids with polar side chains.

Proteins of cat saliva dry and flake off as dander, which contains major indoor allergens linked to asthma. Dogs, horses, cattle, and other animals also provide several allergens,[454] some of which are lipocalins (Box 21-A).[455] Other close-to-home allergens are provided by fungi that live on skin or nails.[456]

Plants provide a host of allergens. Major allergens are found in pollen of rye-grass,[457] of many other grasses, of ragweed,[458] and of olive trees.[459] Natural rubber latex would appear to be a harmless high polymer, but it contains antigenic proteins, which have been blamed for 1100 anaphylactic attacks with at least 15 deaths between 1988 and 1992.[460,461]

Food is a major source of allergens, which are often overlooked. Food allergies may be hard to diagnose and symptoms such as headache, diarrhea, itching, and asthma may be attributed to other causes. However, the occasional rapid death from anaphylactic shock, e.g., from exposure to peanuts,[461a] is a reminder that unrecognized food allergies exist. About 100–200 persons die annually of food allergies. About 90% of recognized food allergies involve milk, eggs, fish, crustacea, peanuts, tree nuts, soybeans, and wheat.[462] There are usually only a few allergenic proteins in any one food. Many of these are resistant to digestion in the stomach. Some but not all are compact proteins with multiple disulfide crosslinkages. However, no structural generalization can be applied to all food allergens.

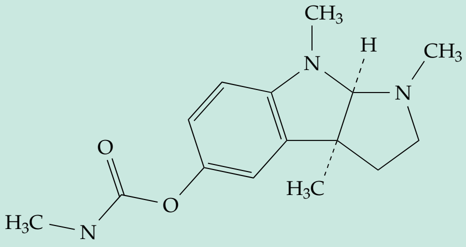
The increasing use of genetic engineering of foods poses both risks and hopes. The accidental incorporation of an allergenic protein into a plant or animal product can make a food previously safe for a person deadly. For this reason, attempts are now made to identify likely allergens and to avoid transferring their genes. However, this process can't be completely reliable, partly because we each have our own personal immune system. This is one reason for requiring accurate labeling of foods. On the positive side recognized major allergens can probably be eliminated by genetic engineering. The muscle protein tropomyosin (Fig. 19-9) is a well-known allergen whose allergenicity varies among different sources. Tropomyosin from beef, pork, and chicken is usually not highly allergenic, but that from shrimp often is.[462] Perhaps safer shrimp can be created.

Allergies are treated in various ways, often with antihistamines or corticosteroid ointments. Injection of adrenaline is an emergency treatment of anaphylaxis or asthma. Also important is **specific immunotherapy**, better known as "allergy shots."[441] Small amounts of allergen in increasing amounts are injected subcutaneously at intervals to desensitize the patient. Use of purer antigens, which may be engineered to decrease their antigenicity, may provide advances in this technique.[463]

Asthma is one of the most common chronic diseases in industrialized countries, affecting 10% or more of young children in some countries. An atopic disease with high IgE levels, asthma is induced by small particles of antigen, which are able to penetrate deep into the lungs.[464] About 80% of asthmatic children are allergic to house mites.[450] Animal danders are another major cause. The incidence of asthma appears to be increasing in many modern societies, but

BOX 31-D MYASTHENIA GRAVIS

One of the best-understood autoimmune diseases is myasthenia gravis, a condition associated with a decrease in the number of functional postsynaptic nicotinic acetylcholine receptors (Fig. 30-23) in neuromuscular junctions.[a–e] The resulting extreme muscular weakness can be fatal. Myasthenia gravis is not rare and affects about one in 10,000 people.[c] An interesting treatment consists of the administration of physostigmine, diisopropylphosphofluoridate (Chapter 12, Section C,1), or other acetylcholinesterase inhibitors (Box 12-E). These very toxic compounds, when administered in controlled amounts, permit accumulation of higher acetylcholine concentration with a resultant activation of muscular contraction. The same compounds are widely used in the treatment of glaucoma.



Alkaloid physostigmine

More than 90% of patients with myasthenia gravis have circulating antibodies directed against a subunit of the acetylcholine receptor.[f] Immunosuppressive drugs and steroids help to cut down on these autoantibodies, and many patients are benefitted by removal of the thymus. Newer approaches involve specific immunotherapy aimed at increasing tolerance to either T cells or to B cells.[c,d] For example, oral ingestion of purified acetylcholine receptors to desensitize the body's response or inhibition of production of Il-2.

A possible cause for the production of the damaging antibodies may be the sharing of common antigenic determinants between the receptor protein and surface proteins of bacteria such as *E. coli*.[f]

---

[a] Fuchs, S. (1980) *Trends Biochem. Sci.* **5**, 259–262

[b] Tzartos, S. J. (1984) *Trends Biochem. Sci.* **9**, 63–67

[c] Drachman, D. B. (1994) *N. Engl. J. Med.* **330**, 1797–1808

[d] Steinman, L., and Mantegazza, R. (1990) *FASEB J.* **4**, 2726–2731

[e] Barnes, D. M. (1986) *Science* **232**, 160–161

[f] Stefansson, K., Dieperink, M. E., Richman, D. P., Gomez, C. M., and Marton, L. S. (1985) *N. Engl. J. Med.* **312**, 221–225

the reasons are unclear,[465,466] and the increase doesn't appear to be linked to air pollution. A chronic atopic condition usually precedes an acute attack of asthma. In addition to IgE and eosinophils there are excessive numbers of neutrophils with high-affinity IgE receptors in the airway tissues.[467] Prostaglandin D_2 released from mast cells may play a role in triggering an attack.[468] Cytokines,[469] nitric oxide,[447] and nerve growth factor[470] may also participate in the response. The presence of a high concentration of glutathione and of glutathione peroxidase, whose concentration increases in asthmatic lungs, may reflect the action of the antioxidant system in combatting inflammation.[471] The surfactant proteins SP-A and SP-D (Box 8-B) are Ca^{2+}-dependent lectins, which serve as regulators of the innate immune response. Their concentrations also increase in asthma.[472,473]

Treatment of asthma has depended upon inhaled glucocorticoids, quick-acting bronchodilators usually β-adrenergic antagonists, and long-acting beta agonists such as theophyllins and leukotriene antagonists.[464]

2. Autoimmune Diseases

There are numerous **autoimmune diseases** in which the body makes antibodies against its own cells (Table 31-3).[363,474–476] In **myasthenia gravis** (Box 31-D) antibodies attack the acetylcholine receptors in postsynaptic membranes.[477] In **Graves disease** aberrant antibodies are directed against receptors for thyrotropin. They have a stimulatory rather than an inhibitory effect and cause hyperthyroidism.[478] **Childhood onset (Type I) diabetes** results from destruction of insulin-secreting cells by an autoimmune reaction triggered by environmental factors in genetically susceptible persons (Box 17-G).[479,480] The principal autoantigen appears to be the 65-kDa form of glutamate decarboxylase (GAD).[481–483] While GAD has an essential function in the formation of γ-aminobutyrate in the brain, its role in the pancreatic islets is not clear. What is established is the presence of specific GAD65-stimulated T cells in diabetic individuals. The stimulating autoantibodies, which may appear in the blood years before diabetes is evident, carry HLA-DR4 type surface MHC antigens.[484,485] After the activated T cells kill enough of the pancreatic β cells, diabetes appears.[486]

Myasthenia gravis, Graves disease, and type I diabetes are organ-specific autoimmune diseases. Another group of autoimmune diseases are systemic, affecting many tissues. For example, in the severe **systemic lupus erythematosus** there are often antibodies against the victim's own DNA.[487–488a] The antibodies may then attack any tissue, e.g., the red blood cells. Antibodies against a variety of other nuclear constitutents such as histones,[207] ribosomal protein L7,[489] ubiquitin,[490] enzymes, cardiolipin,[491] and small nuclear RNAs[492] are also made. The primary defect appears to be an intolerance to chromatin and in particular to nucleosomes.[493] Antibodies to nucleolar and other components of nuclei are also present in progressive systemic sclerosis (**scleroderma**).[494,495] In **rheumatoid arthritis**, a chronic inflammation of joints, the serum and joint fluids contain abnormal complexes, which appear to consist entirely of immunoglobulins. They may be antibody – antiidiotype antibody complexes.[496,496a] Immunization of animals with type II collagen induces a very similar arthritis,[496,497] but this collagen probably doesn't supply the offending human antigen. Susceptibility to rheumatoid arthritis is linked to HLA-DR4 class II MHC genes. Molecules of the class II DR4 subtypes may associate with antigenic peptides of uncertain origin to induce a T-cell response. $CD4^+$ T cells are thought to drive inflammation in the disease.[497] Monocytes are attracted from the blood and become inflammatory macrophages.[498]

Another autoimmune disease in which antibodies attack collagen is **Goodpasture disease**. It is mediated by B cells, which form antibodies directed at the N-terminal domain of the α3 chain of collagen IV (pp. 435–438).[499–501] The antibodies attack the glomerular basement membranes causing a rapidly progressing glomerulonephritis and also lung hemorrhages. Primary **glomerulonephritis**, a major kidney disease, may be caused by a cross-reaction between the membrane of streptococci and the glomerular basement membranes.

TABLE 31-3
Some Autoimmune Diseases

Addison disease	Adrenal glands
Ankylosing spondilitis	
Celiac disease	Upper intestines
Crohn disease	Intestines
Diabetes, type I	Pancreatic islets
Glomerulonephritis	Kidney
Goodpasture disease	Kidney
Graves disease	Thyroid gland
Guillain – Barré syndrome	Gangloisides
Multiple sclerosis (MS)	Peripheral myelin
Myasthenia gravis	
Paroxysmal cold hemoglobinuria	Red blood cells
Primary biliary cirrhosis	
Psoriasis	Skin
Polymyocystitis	
Rasmussen encephalomyelitis	Cerebral cortex
Rheumatoid arthritis	Joints
Scleroderma	Skin
Sjögren syndrome	
Systemic lupus erythematosis	Many tissues
Thyroiditis	Thyroid gland
Ureitis	

In **Sjögren syndrome** autoantibodies are directed against α fodrin (p. 405).[502] In primary **biliary cirrhosis** they are directed at mitochondria and specificially to a pyruvate dehydrogenase subunit (Fig. 15-14).[503] In the inflammatory muscle disease **polymyocystitis** autoantibodies are often directed against cytoplasmic proteins including aminoacyl-tRNA synthetases.[504] In the rare **paroxysmal cold hemoglobinuria** autoantibodies attack red blood cell membranes only when the temperature of an extremity is lowered. **Paroxysmal nocturnal hemoglobinuria**, a serious complement-mediated condition, results from deficiency in the complement decay accelerating factor. This is a result of a defect in the PGI tail on this factor.[505]

Celiac disease (celiac sprue) is an allergic inflammatory condition caused by poorly digested proline-rich sequences of wheat gluten and related proteins (p. 74). The disease is usually not recognized, but it may occur in 3% or more of the United States population. A T-cell response that causes destruction of the smaller intestinal mucosa, celiac disease is characterized by malabsorption and diarrhea.[505a–c] It can cause death by starvation. A primary target of the autoantibodies is a transglutaminase.[505c,d]

Most cells of the immune system are ordinarily kept apart from those of the nervous system by means of the blood–brain barrier. However, allergic encephalomyelitis, in which T cells attack the myelin sheath of brain neurons, can easily be induced in mice.[506] A similar autoimmune process is thought to be involved in human **multiple sclerosis** (see Chapter 30, pp. 1769, 1808, and Fig. 30-9).[507,508] High levels of circulating IgM are found in some demyelinating diseases of peripheral neurons.[508] In **Rasmussen's encephalitis**, which causes brain inflammation and epilepsy, serum antibodies attack a glutamate receptor subunit **GluR3**.[509]

The causes of autoimmune disease doubtless lie largely in the difficulty of developing a repetoire of immunoglobulin-forming B cells and of T-cell receptors that will always reliably distinguish self from a foreign antigen. The problem can lie either with B-cell recognition or with the T-cell receptors. Extensive medical use is made of **immunosuppressants** in treatment of persistent allergic reactions, autoimmune problems, and rejection of transplanted tissues. Among these compounds are the steroidal

BOX 31-E EVADING THE IMMUNE SYSTEM

Parasitic species always have a problem with the antibodies and killer T cells of their hosts, and the chemical makeup of the external coats of parasites tends to reflect this fact.[a] An example is provided by **trypanosomes,** which cause sleeping sickness and which make much of Africa unsuitable for cattle grazing.[b–d] Trypanosomes in the bloodstream evade the immune system by covering the outer surface of their plasma membrane, flagella and all, with a dense 12- to 15-mm thick monolayer of an ~60-kDa **variable surface glycoprotein**.[e,f] The glycoprotein molecules are anchored in the cell membrane by C-terminal glycosylphosphatidylinositol (GPI) anchors (Fig. 8-13).[g] The glycoprotein layer protects the parasite but is soon attacked by the host's immune system. However, the parasite has perhaps 1000 different genes for the variable surface protein, and every ten days or so new clones of trypanosomes appear with new coats that the immune system is not prepared to attack. To accomplish this cells occasionally copy one of the previously unused variable surface glycoprotein genes and place it into a new location in the genome, where it is expressed.[h–j]

Parasitic **nematodes** shed the outer layers of their external cuticle and like trypanosomes reveal a new layer with different antigenic proteins.[k] *Giardia* protects itself in a similar fashion.[l] **Schistosomes**, tiny parasitic flatworms, evade a host's immune system by shedding complex glycoproteins from specialized double outer membranes.[m] Antigenic determinants including MHC antigens characteristic of the mouse have been identified in the membrane of schistosomes from infected mice. Thus, one aspect of the parasite's defense may be to hide behind surface recognition markers stolen from its host.[n,o] Schistosomes also secrete the peptide Thr-Lys-Pro, which inhibits macrophages, as well as a small molecule that inhibits T lymphocytes.[o]

The malaria parasite *Plasmodium* has a complex life cycle with several forms and spends much of its life hiding within red blood cells.[p] It may also suppress the immune system. The unicellular sporozoites, which are injected into the bloodstream by mosquitos, are protected by an external coat protein that is unusual in containing many short repeated sequences. For example, that of *P. falciparum*, which causes the most deadly form of malaria, contains the sequence Asn-Ala-Asn-Pro repeated 37 times.[q] These coat proteins undergo unusually rapid evolution, which makes the preparation of vaccines difficult.[r]

Trypanosomes, schistosomes, and malaria parasites still represent major health problems. Malaria kills two to four million persons a year and endangers almost a third of the world's population. It has been impossible to produce suitable vaccines for any of these parasites. However, the cloning of genes for individual parasite proteins has given hope that effective vaccines can be devised.[e,o–v] One problem is the lack of interest in financing the effort.[w]

Many other protozoan parasites and bacteria invade cells and take up residence in macrophages.[x] These include species of *Salmonella*, *Legionella*, and *Mycobacterium*.[y] Bacteria often employ structural mimicry to gain access to a cell,[z] e.g., by mimicking the type III secretion system (p. 520).[aa] Some bacteria have developed defenses against reactive oxygen species, allowing them to evade the action of phagocytes.[bb] *Borrelia burgdorferi*, the Lyme disease spirochete, synthesizes an unusual single-layer β-sheet outer surface protein,[cc] which becomes coated with complement protein H. This may protect the bacteria and allow them to live for a long time within cells.[dd]

Even the lowly **influenza virus** finds a way around our immunity so that it can strike us repeatedly. As this virus matures, it acquires a lipid membrane by budding from the host cell. Two virally encoded proteins are present in the membrane. One is a trimeric hemagglutinin, which forms small 7.6-nm spikes that protrude from the virus surface.[ee,ff] The hemagglutinin monomer is a 550-residue peptide containing four antigenic regions. The RNA genome of the virus undergoes rapid mutation (Chapter 28, Section E,2). At least one amino acid substitution was found in each antigenic region, when hemagglutinins from influenza viruses causing epidemics in 1972 and 1975 were compared with the strain that caused a worldwide epidemic in 1968. Recently the type A influenza virus that caused the 1918–1919 pandemic, the greatest acute plague of the 20th century, has been "resurrected" and investigated using viral RNA from three victims.[gg–ii] The globular part of the hemagglutinin appears to have come from a pig and the "stalk" from a human lineage. The virus takes advantage of the pool of virus in swine, humans, and birds to vary its structure and create new strains. The reason for the deadly nature of the 1918–1919 strain, which killed 20–40 million people, an unusually large number of whom were young, previously healthy adults, is not clear.

Viruses use a large variety of mechanisms to evade cellular defense mechanisms. Almost every aspect of the innate or adaptive immune systems provides some opportunity for evasion.[jj] The rapid

BOX 31-E (continued)

mutation rate in a population of virus particles contributes greatly to this ability, allowing chronic infections such as those of hepatitis C[kk] or delayed catastrophic infections such as those of HIV.

---

[a] Bloom, B. R. (1979) *Nature (London)* **279**, 21–26

[b] Englund, P. T., Hajduk, S. L., and Marini, J. C. (1982) *Ann. Rev. Biochem.* **51**, 695–726

[c] Rice-Ficht, A. C., Chen, K. K., and Donelson, J. E. (1981) *Nature (London)* **294**, 53–57

[d] Muñoz-Jordán, J. L., Davies, K. P., and Cross, G. A. M. (1996) *Science* **272**, 1795–1797

[e] Metcalf, P., Blum, M., Freymann, D., Turner, M., and Wiley, D. C. (1987) *Nature (London)* **325**, 84–86

[f] Gardiner, P. R., Pearson, T. W., Clarke, M. W., and Mutharia, L. M. (1987) *Science* **235**, 774–777

[g] Ferguson, M. A. J., Haldar, K., and Cross, J. A. M. (1985) *J. Biol. Chem.* **260**, 4963–4968

[h] Raibaud, A., Gaillard, C., Longacre, S., Hibner, V., Buck, G., Bernardi, G., and Eisen, H. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4306–4310

[i] Borst, P., and Greaves, D. R. (1987) *Science* **235**, 658–667

[j] Navarro, M., and Gull, K. (2001) *Nature (London)* **414**, 759–763

[k] Philipp, M., Parkhouse, R. M. E., and Ogilvie, B. M. (1980) *Nature (London)* **287**, 538–540

[l] Papanastasiou, P., Hiltpold, A., Bommeli, C., and Köhler, P. (1996) *Biochemistry* **35**, 10143–10148

[m] Cummings, R. D., and Nyame, A. K. (1996) *FASEB J.* **10**, 838–848

[n] Simpson, A. J. G., and Cioli, D. (1982) *Nature (London)* **296**, 285–287

[o] Kolata, G. (1987) *Science* **227**, 285–287

[p] Kolata, G. (1984) *Science* **226**, 679–682

[q] Young, K. J. F., Hockmeyer, W. T., Gross, M., Ballou, W. R., Wirtz, R. A., Trosper, J. H., Beaudoin, R. L., Hollingdale, M. R., Miller, L. H., Diggs, C. L., and Rosenberg, M. (1985) *Science* **228**, 958–962

[r] de la Cruz, V. F., Lal, A. A., Welsh, J. A., and McCutchan, T. F. (1987) *J. Biol. Chem.* **262**, 6464–6467

[s] Peterson, D. S., Wrightsman, R. A., and Manning, J. E. (1986) *Nature (London)* **322**, 566–568

[t] Balloul, J. M., Sondermeyer, P., Dreyer, D., Capron, M., Grzych, J. M., Pierce, R. J., Cavillo, D., Lecocq, J. P., and Capron, A. (1987) *Nature (London)* **326**, 149–153

[u] Capron, A., DeSaint, J. P., Capron, M., Ouma, J. H., and Butterworth, A. E. (1987) *Science* **238**, 1065–1072

[v] Marshall, E. (1997) *Science* **275**, 299

[w] Enserink, M. (2000) *Science* **287**, 1956–1958

[x] Small, P. L. C., Ramakrishnan, L., and Falkow, S. (1994) *Science* **263**, 637–639

[y] Schorey, J. S., Carroll, M. C., and Brown, E. J. (1997) *Science* **277**, 1091–1093

[z] Vallance, B. A., and Finlay, B. B. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8799–8806

[aa] Stebbins, C. E., and Galán, J. E. (2001) *Nature (London)* **412**, 701–705

[bb] Hassett, D. J., and Cohen, M. S. (1989) *FASEB J.* **3**, 2574–2582

[cc] Huang, X., Nakagawa, T., Tamura, A., Link, K., Koide, A., and Koide, S. (2001) *J. Mol. Biol.* **308**, 367–375

[dd] Hellwage, J., Meri, T., Heikkila, T., Alitalo, A., Panelius, J., Lahdenne, P., Seppälä, I. J. T., and Meri, S. (2001) *J. Biol. Chem.* **276**, 8427–8435

[ee] Wilson, I. A., Skehel, J. J., and Wiley, D. C. (1981) *Nature (London)* **289**, 366–373

[ff] Chen, J., Skehel, J. J., and Wiley, D. C. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8967–8972

[gg] Lederberg, J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2115–2116

[hh] Basler, C. F., Reid, A. H., Dybing, J. K., Janczewski, T. A., Fanning, T. G., Zheng, H., Salvatore, M., Perdue, M. L., Swayne, D. E., Garcîa-Sastre, A., Palese, P., and Taubenberger, J. K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2746–2751

[ii] Gibbs, M. J., Armstrong, J. S., and Gibbs, A. J. (2001) *Science* **293**, 1842–1845

[jj] Ploegh, H. L. (1998) *Science* **280**, 248–253

[kk] Farci, P., Shimoda, A., Coiana, A., Diaz, G., Peddis, G., Melpolder, J. C., Strazzera, A., Chien, D. Y., Munoz, S. J., Balestrieri, A., Purcell, R. H., and Alter, H. J. (2000) *Science* **288**, 339–344

anti-inflammatory agents such as prednisone and the folate antagonist methotrexate. A new vision of the possibilities for immunosuppression came, however, with the discovery of **cyclosporin A** (Box 9-F). This fungal metabolite inhibits lymphokine formation by helper T cells. It alleviates rejection of grafted tissues and prevents graft-versus-host disease. The use of cyclosporin and FK506 (Box 9-F) has permitted organ transplantation to the extent that by 1987 surgeons had transplanted in one year 1200 livers, 1500 hearts, and 9000 kidneys with one-year survival rates of 80% for hearts and over 90% for kidneys.[510] As mentioned in Box 9-F, cyclosporins bind very tightly to **cyclophilins**,[511] which have peptidylprolyl *cis*–*trans* isomerase activity and perhaps other independent functions.[512]

3. Immunodeficiencies

More than 95 different problems of impaired immunity have been identified. They affect about 1 in 10,000 persons born. The defects may involve T cells, B cells, NK cells, or phagocytic cells.[513,514] There may be problems in lymphocyte development.[515] Complement proteins may be lacking,[161,513,516] or their regulation may be faulty.[517] The immune system has specific "blind spots" and fails to recognize all dangerous foreign antigens.[518] In the fatal **X-linked immunoproliferative syndrome**,[519,520] the immune system of susceptible males does not respond to the Epstein–Barr virus-induced mononucleosis by killing the persistently lymphoblastoid cells characteristic of that disease. Some individuals are born with **severe combined immunodeficiency** disease. This condition was made well known to the public by the plight of David, the "bubble boy," who lived 12 years in protective sterile rooms and a plastic bubble-like "space suit."[521] The condition is often caused by a defect in the interleukin-2 receptor, but there are a variety of other causes. About 15% of cases arise from a defect

BOX 31-F AN INSECTICIDAL PROTEIN

During sporulation the bacterium *Bacillus thuringiensis* forms within its own cells large protein crystals, which are highly toxic to some insect larvae. The crystals account for 20–30% of the dry weight of the bacterial spores and contain more than one toxin[a] and, curiously, a 20-kbp piece of DNA.[b] Dusting of garden plants with dried spores from these bacteria has become a popular and effective way of combating cabbage worms and other insects. The toxic protein from one strain of bacteria is encoded by a 4222-bp gene.[c] The corresponding 133-kDa 1176-residue polypeptide protoxin undergoes glycosylation and perhaps other modifications, presumably prior to crystallization. After ingestion by susceptible insect larvae (largely Lepidoptera) the protein is cleaved to form a smaller ~65-kDa protease-resistant core, which is the active toxin. Other strains of bacteria produce toxins specific for Diptera or Coleoptera.[d–f]

X-ray crystallography of the 65-kDa form reveal a three-domain structure. The central domain varies among different strains and is probably involved in recognition and in binding to cell surface receptors.[e–h] The toxin binds to a receptor, apparently an aminopeptidase N,[i] after which the toxin is rapidly inserted into the membrane forming a 1- to 2-nm diameter pore. This leads to cell death.[j]

Because the toxins appear to be harmless to human beings and higher animals the toxin genes have been transferred into various other bacteria, which are symbiotic with plants and into plants themselves. Toxin genes in suitably modified form (Chapter 27) were first transferred into bacteria that live naturally in association with roots of *Zea mays* and into tobacco and tomato plants. The new host organisms expressed the toxin genes and protected the plants from damage by caterpillars.[k,l] Since then the toxin genes have been transferred into many crop plants, which are widely planted.

Two problems must be considered. Insects do develop resistance to the Bt toxin.[m] This problem can be combated by protein engineering[n] and by location of new sources of toxins.[o,p] A second problem deals with the environmental impact.[q] Will Bt toxin kill desirable insects? Will the gene be transferred in nature to other species and into the environment? The latter may seem unlikely, but as the toxins are applied to fight soil organisms such as nematodes, transfer into organisms of the largely unstudied soil ecosphere may pose problems.

---

[a] Donovan, W. P., Dankocsik, C. C., Gilbert, M. P., Gawron-Burke, M. C., Groat, R. G., and Carlton, B. C. (1988) *J. Biol. Chem.* **263**, 561– 567

[b] Clairmont, F. R., Milne, R. E., Pham, V. T., Carrière, M. B., and Kaplan, H. (1998) *J. Biol. Chem.* **273**, 9292 – 9296

[c] Schnepf, H. E., Wong, H. C., and Whiteley, H. R. (1985) *J. Biol. Chem.* **260**, 6264 – 6272

[d] Sekar, V., Thompson, D. V., Maroney, M. J., Bookland, R. G., and Adang, M. J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7036 – 7040

[e] Garfield, J. L., and Stout, C. D. (1988) *J. Biol. Chem.* **263**, 11800 – 11801

[f] Li, J., Koni, P. A., and Ellar, D. J. (1996) *J. Mol. Biol.* **257**, 129 – 152

[g] Li, J., Carroll, J., and Ellar, D. J. (1991) *Nature (London)* **353**, 815 – 821

[h] Grochulski, P., Masson, L., Borisova, S., Pusztai-Carey, M., Schwartz, J.-L., Brousseau, R., and Cygler, M. (1995) *J. Mol. Biol.* **254**, 447– 464

[i] Gill, S. S., Cowles, E. A., and Francis, V. (1995) *J. Biol. Chem.* **270**, 27277– 27282

[j] Burton, S. L., Ellar, D. J., Li, J., and Derbyshire, D. J. (1999) *J. Mol. Biol.* **287**, 1011– 1022

[k] Vaeck, M., Reynaerts, A., Hofte, H., Jansens, S., De Beuckeleer, M., Dean, C., Zabeau, M., Van Montagu, M., and Leemans, J. (1987) *Nature (London)* **328**, 33 – 37

[l] Baum, R. M. (1987) *Chem. Eng. News* **65** **Aug 10**, 9 – 14

[m] Oppert, B., Kramer, K. J., Beeman, R. W., Johnson, D., and McGaughey, W. H. (1997) *J. Biol. Chem.* **272**, 23473 – 23476

[n] Rajamohan, F., Alzate, O., Cotrill, J. A., Curtiss, A., and Dean, D. H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14338 – 14343

[o] Estruch, J. J., Warren, G. W., Mullins, M. A., Nye, G. J., Craig, J. A., and Koziel, M. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5389 – 5394

[p] Guo, L., Fatig, R. O., III, Orr, G. L., Schafer, B. W., Strickland, J. A., Sukhapinda, K., Woodsworth, A. T., and Petell, J. K. (1999) *J. Biol. Chem.* **274**, 9836 – 9842

[q] Pimentel, D. S., and Raven, P. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8198 – 8199

---

in adenosine deaminase, an enzyme of the purine salvage pathway (Fig. 25-17). As mentioned in Chapter 25, genetic therapy for this condition is being used. However, the most reliable treatment for these immunodeficiencies seems to be bone marrow transplantation. By 2000 more than 375 patients worldwide had received this treatment with up to 95% chance of survival.[513] Virus-induced immunodeficiency is the prime characteristic of HIV infection and **AIDS**.[522,523] Both the amounts of autoantibodies and of amyloid deposits increase with age, and immune complex disease is suspected of being a cause of aging.

4. Cancers of the Immune System

A major function of the immune system is thought to be destruction of cancer cells. In this case altered cell surface carbohydrates or proteins elicit an antibody response with destruction of the offending cells.

That this process works imperfectly may explain why the incidence of cancer increases with age and also why the concentration of autoantibodies increases. The immune system is also susceptible to cancers, which include multiple myeloma, leukemias, and lymphomas. Some of these, such as Burkitt's lymphoma, involve rearrangement of chromosome segments that carry immunoglobulin genes.[524,525] These may result from errors in the gene rearrangements involved in the development of lymphocytes.

G. Defense Mechanisms of Plants

Plants make many compounds that repel or poison animals that eat them. Among such compounds are alkaloids, terpenes, calcium oxalate, fluoroacetate, cyanogenic glycosides, and phenolic compounds.[526] The chewing of insects or other wounding of plant tissues releases phenolic glycosides and other reactive compounds from vesicles. Some of these compounds, which are often referred to as **phytoalexins** (see also Box 20-E),[527] are repellent to predators, have antimicrobial activity, and /or participate in chemical cross-linking and strengthening of the plant cell wall.[526,528] Some are protease inhibitors that interfere with a predator's nutrition.[529] Some released compounds attract insects that may assist in defense by feeding on predator eggs or by attracting wasps that deposit eggs in predator larvae.[530] These can all be regarded as part of an innate defense system that in some respects resembles our own innate immune system. For example, plant defensins (Fig. 31-7), most of which are directed against fungi,[531,532] resemble those in our tissues.

A system of **receptor-mediated surveillance**, part of the innate system, triggers both immediate **local responses** and secondary immunity throughout the plant.[530,533,534] Immediate responses include **programmed cell death** (called the **hypersensitive response**[535]), tissue reinforcement, and production of antimicrobial metabolites. Secondary responses, known as **systemic acquired resistance**, develop immunity throughout the plant. The surveillance system of some plants consists of a series of receptors known as **resistance (R) proteins**, which recognize signaling molecules produced by pathogens.[535a,b] The R proteins are thought of as being paired with **avirulence** (**Avr**) **proteins** of the pathogen. If the resistance protein is missing, the plant will be susceptible to attack by the pathogen. The pathogen's Avr protein is thought to be part of the chemical attack on the plant, apparently assisting in the invasion. However, if the *Avr* gene has been lost or is mutated, the R protein won't detect the invasion, and the pathogen may have increased virulence. What are the characteristics of the Avr proteins? They are often small and may be crosslinked by S–S bridges. They may be taken up by plant cells via receptors that resemble the type III translocation system of bacteria (p. 520).[534] Similar small protein **elicitors** are released directly by wounding even in plants that do not have paired *R*–*Avr* genes.[529,536]

The R proteins, which act as receptors for Avr, and other elicitor proteins, are usually leucine-rich-repeat proteins with a characteristic nucleotide binding site attached (NB-LRR proteins).[534,537] Like other cell surface receptors they participate in signaling and utilize both ion channels and Ser/Thr protein kinases.[538] The *Arabidopsis* genome contains ~150 sequences that may represent NB-LRR receptors.[530]

What do these receptors do? Like other cell membrane receptors they may induce both rapid and slower responses. The rapid responses may result from transmembrane flow of ions, just as in neurotransmitter action (Figs. 30-19, 30-20). The first response observed is an oxidative burst, which within minutes generates reduced oxygen intermediates (ROIs; pp. 1072–1074).[533,539–543] These compounds may participate in crosslinking and lignification of cell walls. Together with nitric oxide (NO) and endogenous salicylic acid (Chapter 25, Section B,7),[533,544,545] they promote transcription of defense-related genes and participate in the hypersensitive response. A second pathway, utilized against some pathogens especially those that kill plants to obtain nutrients,[533] involves production of jasmonic acid (Eq. 21-18 and associated text)[545a] and ethylene (Fig. 24-16).[529,546]

Plants also have mechanisms for minimizing the damage from the over 500 known viruses. These don't often kill plants but can cause great damage. There are interferon-like responses[547] and gene-silencing mechanisms.[548] The latter often involve synthesis of dsRNA, cleavage by the enzyme Dicer, and interference with transcription as described on p. 1640.[113a,b,549,550] This defensive reaction can spread between cells and throughout a plant, apparently by transport of RNA through plasmadesmata and the phloem.[550]

References

1. Nossal, G. J. V. (1993) *Sci. Am.* **269** (Sep), 53–62
2. Delves, P. J., and Roitt, I. M. (2000) *N. Engl. J. Med.* **343**, 37–49
3. Bhagovan, N. V. (2001) *Medical Biochemistry*, 4th ed., Harcourt/Academic Press, San Diego, California
4. Cruse, J. M., and Lewis, R. E. (1997) *Atlas of Immunology*, CRC Press, Boca Raton, Florida
5. Van Oss, C. J., and Van Regenmortel, M. H. V. (1994) *Immunochemistry*, Dekker, New York
6. Kuby, J. (1992) *Immunology*, Freeman, New York
7. Male, D., Champion, B., Cooke, A., and Owen, M. (1991) *Advanced Immunology*, 2nd ed., Raven Press, New York
8. Hames, B. D., and Glover, D. M. (1988) *Molecular Immunology*, IRL Press, Oxford
9. Edelson, R. L., and Fink, J. M. (1985) *Sci. Am.* **252**(Jun), 46–53
10. Janeway, C. A., Jr. (1993) *Sci. Am.* **269**(Sep), 73–89
11. Medzhitov, R., and Janeway, C., Jr. (2000) *N. Engl. J. Med.* **343**, 338–344

11a. Medzhitov, R., and Janeway, C. A., Jr. (2002) *Science* **296**, 298–300

11b. Matzinger, P. (2002) *Science* **296**, 301–305

12. Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., Jr., and Ezekowitz, R. A. B. (1999) *Science* **284**, 1313–1318
13. Brown, P. (2001) *Nature (London)* **410**, 1018–1020
14. Janeway, C. A., Jr. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7461–7468
15. Paul, W. E. (1993) *Sci. Am.* **269**(Sep), 91–97
16. Marone, G., Lichtenstein, L. M., and Galli, S. J., eds. (2000) *Mast Cells and Basophils*, Academic Press, San Diego, California
17. Metcalf, D. (1991) *Science* **254**, 529–533
18. Boyington, J. C., Motyka, S. A., Schuck, P., Brooks, A. G., and Sun, P. D. (2000) *Nature (London)* **405**, 537–543
19. Sheriff, S., Chang, C. Y. Y., and Ezekowitz, R. A. B. (1994) *Nature Struct. Biol.* **1**, 789–794
20. Riddihough, G. (1994) *Nature (London)* **372**, 114

20a. Hazen, S. L., and Chisolm, G. M. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12515–12517

21. Hayashi, F., Smith, K. D., Ozinsky, A., Hawn, T. R., Yi, E. C., Goodlett, D. R., Eng, J. K., Akira, S., Underhill, D. M., and Aderem, A. (2001) *Nature (London)* **410**, 1099–1103
22. Banchereau, J., and Steinman, R. M. (1998) *Nature (London)* **392**, 245–252
23. Pulendran, B., Palucka, K., and Banchereau, J. (2001) *Science* **293**, 253–256
24. Thoma-Uszynski, S., Stenger, S., Takeuchi, O., Ochoa, M. T., Engele, M., Sieling, P. A., Barnes, P. F., Röllinghoff, M., Bölcskei, P. L., Wagner, M., Akira, S., Norgard, M. V., Belisle, J. T., Godowski, P. J., Bloom, B. R., and Modlin, R. L. (2001) *Science* **291**, 1544–1547
25. Huang, Q., Liu, D., Majewski, P., Schulte, L. C., Korn, J. M., Young, R. A., Lander, E. S., and Hacohen, N. (2001) *Science* **294**, 870–874
26. Streit, W. J., and Kincaid-Colton, C. A. (1995) *Sci. Am.* **273**(Nov), 54–61

26a. Banchereau, J. (2002) *Sci. Am.* **287** (Nov), 52–59

27. Re, F., and Strominger, J. L. (2001) *J. Biol. Chem.* **276**, 37692–37699

27a. Janssens, S., and Beyaert, R. (2002) *Trends Biochem. Sci.* **27**, 474–482

27b. Mak, T. W., and Yeh, W.-C (2002) *Nature (London)* **418**, 835–836

28. Boissy, G., O'Donohue, M., Gaudemer, O., Perez, V., Pernollet, J.-C., and Brunie, S. (1999) *Protein Sci.* **8**, 1191–1199
29. Fagarasan, S., and Honjo, T. (2000) *Science* **290**, 89–92
30. Macpherson, A. J., Gatto, D., Sainsbury, E., Harriman, G. R., Hengartner, H., and Zinkernagel, R. M. (2000) *Science* **288**, 2222–2226

30a. Khush, R. S., Leulier, F., and Lemaitre, B. (2002) *Science* **296**, 273–275

30b. Choe, K.-M, Werner, T., Stöven, S., Hultmark, D., and Anderson, K. V. (2002) *Science* **296**, 359–362

30c. Christophides, G. K., and 34 other authors. (2002) *Science* **298**, 159–165

31. Hancock, R. E. W., and Scott, M. G. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8856–8861
32. Sawai, M. V., Jia, H. P., Liu, L., Aseyev, V., Wiencek, J. M., McCray, P. B., Jr., Ganz, T., Kearney, W. R., and Tack, B. F. (2001) *Biochemistry* **40**, 3810–3816
33. Ganz, T. (1999) *Science* **286**, 420–421
34. Newman, J. (1995) *Sci. Am.* **273**(Dec), 76–79
35. Zinkernagel, R. M. (2001) *N. Engl. J. Med.* **345**, 1331–1335
36. von Boehmer, H., and Kisielow, P. (1991) *Sci. Am.* **265**(Oct), 74–81
37. Nossal, G. J. V. (2001) *Nature (London)* **412**, 685–686
38. Ahmed, R., and Gray, D. (1996) *Science* **272**, 54–60
39. Fearon, D. T., Manders, P., and Wagner, S. D. (2001) *Science* **293**, 248–250
40. Mackay, C. R., and von Andrian, U. H. (2001) *Science* **291**, 2323–2324
41. O'Garra, A. (2000) *Nature (London)* **404**, 719–720
42. Butcher, E. C., and Picker, L. J. (1996) *Science* **272**, 60–66
43. Service, R. F. (1994) *Science* **265**, 1522–1524

43a. Hayday, A., and Viney, J. L. (2000) *Science* **290**, 97–100

44. Golde, D. W., and Gasson, J. C. (1988) *Sci. Am.* **259**(July), 62–70
45. Socolovsky, M., Lodish, H. F., and Daley, G. Q. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6573–6575
46. Phillips, R. L., Ernst, R. E., Brunk, B., Ivanova, N., Mahan, M. A., Deanehan, J. K., Moore, K. A., Overton, G. C., and Lemischka, I. R. (2000) *Science* **288**, 1635–1640
47. Weissman, I. L., and Cooper, M. D. (1993) *Sci. Am.* **269**(Sep), 65–71
48. Akashi, K., Traver, D., Miyamoto, T., and Weissman, I. L. (2000) *Nature (London)* **404**, 193–197
49. Metcalf, D. (1992) *Trends Biochem. Sci.* **17**, 286–289
50. Delves, P. J., and Roitt, I. M. (2000) *N. Engl. J. Med.* **343**, 108–117
51. Liu, Y.-J. (1997) *Science* **278**, 238–239
52. Burton, D. R. (1990) *Trends Biochem. Sci.* **15**, 64–69
53. Sondermann, P., Kaiser, J., and Jacob, U. (2001) *J. Mol. Biol.* **309**, 737–749
54. Grakoui, A., Bromley, S. K., Sumen, C., Davis, M. M., Shaw, A. S., Allen, P. M., and Dustin, M. L. (1999) *Science* **285**, 221–227
55. Fineschi, B., and Miller, J. (1997) *Trends Biochem. Sci.* **22**, 377–382
56. Williams, N. (1998) *Science* **280**, 198–200

56a. Chen, Y., Chou, K., Fuchs, E., Havran, W. L., and Boismenu, R. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14338–14343

56b. Jameson, J. Ugarte, K., Chen, N., Yachi, P., Fuchs, E., Boismenu, R., and Havran, W. L. (2002) *Science* **296**, 747–749

57. Qi, S. Y., Groves, J. T., and Chakraborty, A. K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6548–6553

57a. Dustin, M. L., and Colman, D. R. (2002) *Science* **298**, 785–789

57b. Lee, K.-H., Holdorf, A. D., Dustin, M. L., Chen, A. C., Allen, P. M., and Shaw, A. S. (2002) *Science* **295**, 1539–1542

58. Trautmann, A., and Vivier, E. (2001) *Science* **292**, 1667–1668

58a. Khan, A. A., Bose, C., Yam, L. S., Soloski, M. J., and Rupp, F. (2001) *Science* **292**, 1681–1686

59. Edelman, G. M. (1973) *Science* **180**, 830–840
60. Porter, R. R. (1973) *Science* **180**, 713–716
61. Edmundson, A. B., Ely, K. R., Abola, E. E., Schiffer, M., and Panagiotopoulos, N. (1975) *Biochemistry* **14**, 3953–3961
62. Putnam, F. W. (1993) *Protein Sci.* **2**, 1536–1542
63. Verfuèse, D. (1972) *Science* **178**, 384–385
64. Putnam, F. W., Florent, G., Paul, C., Shinoda, T., and Shimizu, A. (1973) *Science* **182**, 287–291

64a. Cochran, J. R., Aivazian, D., Cameron, T. O., and Lawrence, L. J. (2001) *Trends Biochem. Sci.* **26**, 304-310

64b. Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A., and Dwek, R. A. (2001) *Science* **292**, 2370–2376

64c. Feinstein, A., and Munn, E. A. (1969) *Nature (London)* **224**, 1307–1309

65. Bork, P., Holm, L., and Sander, C. (1994) *J. Mol. Biol.* **242**, 309–320
66. Schiffer, M., Girling, R. L., Ely, K. R., and Edmundson, A. B. (1973) *Biochemistry* **12**, 4620–4631

66a. Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M., and Poljak, R. J. (1989) *Nature (London)* **342**, 877–883

67. Capra, J. D., and Edmundson, A. B. (1977) *Sci. Am.* **236**(Jan), 50–59
68. Amzel, L. M., and Poljak, R. J. (1979) *Ann. Rev. Biochem.* **48**, 961–997
69. Davies, D. R., Sheriff, S., and Padlan, E. A. (1988) *J. Biol. Chem.* **263**, 10541–10544
70. Lesk, A. M., and Chothia, C. (1988) *Nature (London)* **335**, 188–190
71. Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., and Poljak, R. J. (1986) *Science* **233**, 747–753
72. Lescar, J., Pellegrini, M., Souchon, H., Tello, D., Poljak, R. J., Peterson, N., Greene, M., and Alzari, P. M. (1995) *J. Biol. Chem.* **270**, 18067–18076
73. Sundberg, E. J., Urrutia, M., Braden, B. C., Isern, J., Tsuchiya, D., Fields, B. A., Malchiodi, E. L., Tormo, J., Schwarz, F. P., and Mariuzza, R. A. (2000) *Biochemistry* **39**, 15375–15387
74. Cygler, M., Rose, D. R., and Bundle, D. R. (1991) *Science* **253**, 442–445
75. Novotny, J., and Haber, E. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4592–4596
76. Bruccoleri, R. E., Haber, E., and Novotny, J. (1988) *Nature (London)* **335**, 564–568
77. Boehm, M. K., Woof, J. M., Kerr, M. A., and Perkins, S. J. (1999) *J. Mol. Biol.* **286**, 1421–1447
78. Jung, S., Spinelli, S., Schimmele, B., Honegger, A., Pugliese, L., Cambillau, C., and Plückthun, A. (2001) *J. Mol. Biol.* **309**, 701–716
79. Braden, B. C., and Poljak, R. J. (1995) *FASEB J.* **9**, 9–16
80. Padlan, E. A., Abergel, C., and Tipper, J. P. (1995) *FASEB J.* **9**, 133–139
81. Chatellier, J., Van Regenmortel, M. H. V., Vernet, T., and Altschuh, D. (1996) *J. Mol. Biol.* **264**, 1–6
82. Smith, D. K., and Xue, H. (1997) *J. Mol. Biol.* **274**, 530–545
83. Vargas-Madrazo, E., Lara-Ochoa, F., and Almagro, J. C. (1995) *J. Mol. Biol.* **254**, 497–504

References

84. Strong, R. K., Campbell, R., Rose, D. R., Petsko, G. A., Sharon, J., and Margolies, M. N. (1991) *Biochemistry* **30**, 3739 – 3748

85. Lim, K., Owens, S. M., Arnold, L., Sacchattini, J. C., and Linthicum, D. S. (1998) *J. Biol. Chem.* **273**, 28576 – 28582

86. Guddat, L. W., Shan, L., Broomell, C., Ramsland, P. A., Fan, Z.-c, Anchin, J. M., Linthicum, D. S., and Edmundson, A. B. (2000) *J. Mol. Biol.* **302**, 853 – 872

87. Mol, C. D., Muir, A. K. S., Cygler, M., Lee, J. S., and Anderson, W. F. (1994) *J. Biol. Chem.* **269**, 3615 – 3622

88. Yokoyama, H., Mizutani, R., Satow, Y., Komatsu, Y., Ohtsuka, E., and Nikaido, O. (2000) *J. Mol. Biol.* **299**, 711 – 723

89. Augustine, J. G., de la Calle, A., Knarr, G., Buchner, J., and Frederick, C. A. (2001) *J. Biol. Chem.* **276**, 3287 – 3294

90. Bossart-Whitaker, P., Chang, C. Y., Novotny, J., Benjamin, D. C., and Sheriff, S. (1995) *J. Mol. Biol.* **253**, 559 – 575

91. Berthet-Colominas, C., Monaco, S., Novelli, A., Sibai, G., Mallet, F., and Cusack, S. (1999) *EMBO J.* **18**, 1124 – 1136

92. Landry, R. C., Klimowicz, A. C., Lavictoire, S. J., Borisova, S., Kottachchi, D. T., Lorimer, I. A. J., and Evans, S. V. (2001) *J. Mol. Biol.* **308**, 883 – 893

93. Muyldermans, S., Cambillau, C., and Wyns, L. (2001) *Trends Biochem. Sci.* **26**, 230 – 235

94. Conrath, K. E., Lauwereys, M., Wyns, L., and Muyldermans, S. (2001) *J. Biol. Chem.* **276**, 7346 – 7350

95. Zdanov, A., Li, Y., Bundle, D. R., Deng, S.-J., MacKenzie, C. R., Narang, S. A., Young, N. M., and Cygler, M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6423 – 6427

96. Derrick, J. P., and Wigley, D. B. (1992) *Nature (London)* **359**, 752 – 754

97. Todd, P. E. E., East, I. J., and Leach, S. J. (1982) *Trends Biochem. Sci.* **7**, 212 – 216

98. Westhof, E., Altschuh, D., Moras, D., Bloomer, A. C., Mondragon, A., Klug, A., and Van Regenmortel, M. H. V. (1984) *Nature (London)* **311**, 123 – 126

99. Tainer, J. A., Getzoff, E. D., Alexander, H., Houghten, R. A., Olson, A. J., Lerner, R. A., and Hendrickson, W. A. (1984) *Nature (London)* **312**, 127 – 134

100. Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E., and Lerner, R. A. (1985) *Nature (London)* **318**, 480 – 483

101. Low, T. L. K., Liu, Y.-S. V., and Putnam, F. W. (1976) *Science* **191**, 390 – 392

102. Blank, U., Ra, C., Miller, L., White, K., Metzger, H., and Kinet, J.-P. (1989) *Nature (London)* **337**, 187 – 189

103. Radaev, S., Motyka, S., Fridman, W.-H., Sautes-Fridman, C., and Sun, P. D. (2001) *J. Biol. Chem.* **276**, 16469 – 16477

104. Burmeister, W. P., Huber, A. H., and Bjorkman, P. J. (1994) *Nature (London)* **372**, 379 – 383

105. Weng, Z., Gulukota, K., Vaughn, D. E., Bjorkman, P. J., and DeLisi, C. (1998) *J. Mol. Biol.* **282**, 217 – 225

106. Deisenhofer, J. (1981) *Biochemistry* **20**, 2361 – 2370

107. Sondermann, P., Huber, R., Oosthuizen, V., and Jacob, U. (2000) *Nature (London)* **406**, 267 – 273

108. Garman, S. C., Wurzburg, B. A., Tarchevskaya, S. S., Kinet, J.-P., and Jardetzky, T. S. (2000) *Nature (London)* **406**, 259 – 266

109. Garman, S. C., Sechi, S., Kinet, J.-P., and Jardetzky, T. S. (2001) *J. Mol. Biol.* **311**, 1049 – 1062

110. Teichmann, S. A., and Chothia, C. (2000) *J. Mol. Biol.* **296**, 1367 – 1383

111. Bhagavan, N. V. (2002) *Medical Biochemistry*, 4th ed., Harcourt/Academic Press, San Diego, California (pp. 803 – 837)

112. Blattner, F. R., and Tucker, P. W. (1984) *Nature (London)* **307**, 417 – 422

113. Edelman, G. M. (1976) *Science* **192**, 218 – 226

113a. Klahre, U., Crété, P., Leuenberger, S. A., Iglesias, V. A., and Meins, F., Jr. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11981 – 11986

113b. Zamore, P. D. (2002) *Science* **296**, 1265 – 1269

114. Zhang, L., Benz, R., and Hancock, R. E. W. (1999) *Biochemistry* **38**, 8102 – 8111

115. Wang, Y., Henz, M. E., Gallagher, N. L. F., Chai, S., Gibbs, A. C., Yan, L. Z., Stiles, M. E., Wishart, D. S., and Vederas, J. C. (1999) *Biochemistry* **38**, 15438 – 15447

116. Pütsep, K., Brändén, C.-I., Boman, H. G., and Normark, S. (1999) *Nature (London)* **398**, 671 – 672

117. Park, C. B., Yi, K.-S., Matsuzaki, K., Kim, M. S., and Kim, S. C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8245 – 8250

117a. Kobayashi, S., Hirakura, Y., and Matsuzaki, K. (2001) *Biochemistry* **40**, 14330 – 14335

118. Kragol, G., Lovas, S., Varadi, G., Condie, B. A., Hoffmann, R., and Otvos, L., Jr. (2001) *Biochemistry* **40**, 3016 – 3026

119. Gura, T. (2001) *Science* **291**, 2068 – 2071

120. Ehret-Sabatier, L., Loew, D., Goyffon, M., Fehlbaum, P., Hoffmann, J. A., van Dorsselaer, A., and Bulet, P. (1996) *J. Biol. Chem.* **271**, 29537 – 29544

121. Osaki, T., Omotezako, M., Nagayama, R., Hirata, M., Iwanaga, S., Kasahara, J., Hattori, J., Ito, I., Sugiyama, H., and Kawabata, S.-i. (1999) *J. Biol. Chem.* **274**, 26172 – 26178

122. Silva, P. I., Jr., Daffre, S., and Bulet, P. (2000) *J. Biol. Chem.* **275**, 33464 – 33470

123. Sai, K. P., Jagannadham, M. V., Vairamani, M., Raju, N. P., Devi, A. S., Nagaraj, R., and Sitaram, N. (2001) *J. Biol. Chem.* **276**, 2701 – 2707

124. Lohner, K., Latal, A., Lehrer, R. I., and Ganz, T. (1997) *Biochemistry* **36**, 1525 – 1531

125. Hoover, D. M., Chertov, O., and Lubkowski, J. (2001) *J. Biol. Chem.* **276**, 39021 – 39026

125a. Ganz, T. (2002) *Science* **298**, 977 – 979

126. Jia, H. P., Wowk, S. A., Schutte, B. C., Lee, S. K., Vivado, A., Tack, B. F., Bevins, C. L., and McCray, P. B., Jr. (2000) *J. Biol. Chem.* **275**, 33314 – 33320

127. Park, C. H., Valore, E. V., Waring, A. J., and Ganz, T. (2001) *J. Biol. Chem.* **276**, 7806 – 7810

128. Suemori, S., Lynch-Devaney, K., and Podolsky, D. K. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11017 – 11021

129. Mashimo, H., Wu, D.-C., Podolsky, D. K., and Fishman, M. C. (1996) *Science* **274**, 262 – 265

130. Plaut, A. G. (1997) *N. Engl. J. Med.* **336**, 506 – 507

131. Boix, E., Leonidas, D. D., Nikolovski, Z., Nogués, M. V., Cuchillo, C. M., and Acharya, K. R. (1999) *Biochemistry* **38**, 16794 – 16801

132. Tang, Y.-Q., Yuan, J., Ösapsy, G., Ösapay, K., Tran, D., Miller, C. J., Ouellette, A. J., and Selsted, M. E. (1999) *Science* **286**, 498 – 502

133. Nizet, V., Ohtake, T., Lauth, X., Trowbridge, J., Rudisill, J., Dorschner, R. A., Pestonjamasp, V., Piraino, J., Huttner, K., and Gallo, R. L. (2001) *Nature (London)* **414**, 454 – 457

133a. Nagpal, S., Kaur, K. J., Jain, D., and Salunke, D. M. (2002) *Protein Sci.*, **11**, 2158 – 2167

134. Bauer, F., Schweimer, K., Klüver, E., Conejo-Garcia, J.-R., Forssmann, W.-G., Rösch, P., Adermann, K., and Sticht, H. (2001) *Protein Sci.* **10**, 2470 – 2479

135. Tassin, S., Broekaert, W. F., Marion, D., Acland, D. P., Ptak, M., Vovelle, F., and Sodano, P. (1998) *Biochemistry* **37**, 3623 – 3637

136. Gao, G.-H., Liu, W., Dai, J.-X., Wang, J.-F., Hu, Z., Zhang, Y., and Wang, D.-C. (2001) *Biochemistry* **40**, 10973 – 10978

137. Qi, J., Wu, J., Somkuti, G. A., and Watson, J. T. (2001) *Biochemistry* **40**, 4531 – 4538

138. Tam, J. P., Lu, Y.-A., Yang, J.-L., and Chiu, K.-W. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8913 – 8918

139. Jennings, C., West, J., Waine, C., Craik, D., and Anderson, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10614 – 10619

140. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) *J. Mol. Biol.* **294**, 1327 – 1336

141. Rosengren, K. J., Daly, N. L., Scanlon, M. J., and Craik, D. J. (2001) *Biochemistry* **40**, 4601 – 4609

142. Porter, R. R., and Reid, K. B. M. (1978) *Nature (London)* **275**, 699 – 704

143. Law, S. K. A., and Reid, K. B. M. (1988) *Complement*, IRL Press, Oxford

144. Reid, K. B. M., and Porter, R. R. (1981) *Ann. Rev. Biochem.* **50**, 433 – 464

145. Müller-Eberhard, H. J. (1988) *Ann. Rev. Biochem.* **57**, 321 – 347

146. Winkelstein, J. A., Sullivan, K. E., and Colten, H. R. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3911 – 3941, McGraw-Hill, New York

147. Walport, M. J. (2001) *N. Engl. J. Med.* **344**, 1058 – 1066

148. Walport, M. J. (2001) *N. Engl. J. Med.* **344**, 1140 – 1144

149. Arora, M., Munoz, E., and Tenner, A. J. (2001) *J. Biol. Chem.* **276**, 43087 – 43094

150. Gilbert, T. L., Bennett, T. A., Maestas, D. C., Cimino, D. F., and Prossnitz, E. R. (2001) *Biochemistry* **40**, 3467 – 3475

151. Perkins, S. J. (1985) *Biochem. J.* **228**, 13 – 26

152. Hanson, D. C., Siegel, R. C., and Schumaker, V. N. (1985) *J. Biol. Chem.* **260**, 3576 – 3583

153. Marqués, G., Antón, L. C., Barrio, E., Sánchez, A., Ruiz, S., Gavilanes, F., and Vivanco, F. (1993) *J. Biol. Chem.* **268**, 10393 – 10402

154. Lacroix, M., Rossi, V., Gaboriaud, C., Chevallier, S., Jaquinod, M., Thielens, N. M., Gagnon, J., and Arlaud, G. J. (1997) *Biochemistry* **36**, 6270 – 6282

155. Hoppe, H.-J., and Reid, K. B. M. (1994) *Protein Sci.* **3**, 1143 – 1158

156. Leytus, S. P., Kurachi, K., Sakariassen, K. S., and Davie, E. W. (1986) *Biochemistry* **25**, 4855 – 4863

157. Gaboriaud, C., Rossi, V., Bally, I., Arlaud, G. J., and Fontecilla-Camps, J. C. (2000) *EMBO J.* **19**, 1755 – 1765

158. Luo, C., Thielens, N. M., Gagnon, J., Gal, P., Sarvari, M., Tseng, Y., Tosi, M., Zavodszky, P., Arlaud, G. J., and Schumaker, V. N. (1992) *Biochemistry* **31**, 4254 – 4262

159. Rossi, V., Gaboriaud, C., Lacroix, M., Ulrich, J., Fontecilla-Camps, J. C., Gagnon, J., and Arlaud, G. J. (1995) *Biochemistry* **34**, 7311 – 7321

160. Ruiz, S., Henschen-Edman, A. H., and Tenner, A. J. (1995) *J. Biol. Chem.* **270**, 30627 – 30634

161. Johnson, C. A., Densen, P., Hurford, R. K., Jr., Colten, H. R., and Wetsel, R. A. (1992) *J. Biol. Chem.* **267**, 9347 – 9353

162. Hortin, G. L., Farries, T. C., Graham, J. P., and Atkinson, J. P. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1338 – 1342

163. Sahu, A., and Pangburn, M. K. (1994) *J. Biol. Chem.* **269**, 28997 – 29002

References

164. Plumb, M. E., and Sodetz, J. M. (2000) *Biochemistry* **39**, 13078–13083
164a. Scibek, J. J., Plumb, M. E., and Sodetz, J. M. (2002) *Biochemistry* **41**, 14546 – 14551
165. Hourcade, D. E., Wagner, L. M., and Oglesby, T. J. (1995) *J. Biol. Chem.* **270**, 19716–19722
166. Narayana, S. V. L., Carson, M., El-Kabbani, O., Kilpatrick, J. M., Moore, D., Chen, X., Bugg, C. E., Volanakis, J. E., and DeLucas, L. J. (1994) *J. Mol. Biol.* **235**, 695–708
167. Smith, C. A., Pangburn, M. K., Vogel, C.-W., and Müller-Eberhard, H. J. (1984) *J. Biol. Chem.* **259**, 4582–4588
168. Skriver, K., Radziejewska, E., Silbermann, J. A., Donaldson, V. H., and Bock, S. C. (1989) *J. Biol. Chem.* **264**, 3066–3071
169. Bhakdi, S., and Tranum-Jensen, J. (1983) *Trends Biochem. Sci.* **8**, 134–136
170. Gerard, C., Showell, H. J., Hoeprich, P. D., Jr., Huglis, T. E., and Stimler, N. P. (1985) *J. Biol. Chem.* **260**, 2613–2616
171. Johnson, R. J., and Chenoweth, D. E. (1985) *J. Biol. Chem.* **260**, 10339–10345
172. Humbles, A. A., Lu, B., Nilsson, C. A., Lilly, C., Israel, E., Fujiwara, Y., Gerard, N. P., and Gerard, C. (2000) *Nature (London)* **406**, 998–1001
173. Cain, S. A., Coughlan, T., and Monk, P. N. (2001) *Biochemistry* **40**, 14047–14052
174. Chen, Z., Zhang, X., Gonnella, N. C., Pellas, T. C., Boyar, W. C., and Ni, F. (1998) *J. Biol. Chem.* **273**, 10411–10419
175. Young, J. D.-E., Cohn, Z. A., and Podack, E. R. (1986) *Science* **233**, 184–190
176. Persechini, P. M., Ojcius, D. M., Adeodato, S. C., Notaroberto, P. C., Daniel, C. B., and Young, J. D.-E. (1992) *Biochemistry* **31**, 5017–5021
177. Lichtenheld, M. G., Olsen, K. J., Lu, P., Lowrey, D. M., Hameed, A., Hengartner, H., and Podack, E. R. (1988) *Nature (London)* **335**, 448–451
178. Young, J. D.-E., Cohen, Z. A., and Podack, E. R. (1986) *Science* **233**, 184–190
179. Stenger, S., Hanson, D. A., Teitelbaum, R., Dewan, P., Niazi, K. R., Froelich, C. J., Ganz, T., Thoma-Uszynski, S., Melián, A., Bogdan, C., Porcelli, S. A., Bloom, B. R., Krensky, A. M., and Modlin, R. L. (1998) *Science* **282**, 121–125
180. Young, J. D.-E., and Lowrey, D. M. (1989) *J. Biol. Chem.* **264**, 1077–1083
181. Blom, A. M., Kask, L., and Dahlbäck, B. (2001) *J. Biol. Chem.* **276**, 27136–27144
181a. Aslam, M., and Perkins, S. J. (2001) *J. Mol. Biol.* **309**, 1117 – 1138
182. Chamberlain, D., Ullman, C. G., and Perkins, S. J. (1998) *Biochemistry* **37**, 13918–13929
183. Liszewski, M. K., Tedja, I., and Atkinson, J. P. (1994) *J. Biol. Chem.* **269**, 10776–10779
184. Szakonyi, G., Guthridge, J. M., Li, D., Young, K., Holers, V. M., and Chen, X. S. (2001) *Science* **292**, 1725–1728
185. Oppenheim, J. J., Feldmann, M., Durum, S. K., Hirano, T., Vilcek, J., and Nicola, N. A., eds. (2000) *Cytokine Reference*, Academic Press, San Diego
186. Schindler, C., and Darnell, J. E., Jr. (1995) *Ann. Rev. Biochem.* **64**, 621–651
187. Darnell, J. E., Jr. (1997) *Science* **277**, 1630–1635
188. Beuvink, I., Hess, D., Flotow, H., Hofsteenge, J., Groner, B., and Hynes, N. E. (2000) *J. Biol. Chem.* **275**, 10247–10255
189. Becker, S., Groner, B., and Müller, C. W. (1998) *Nature (London)* **394**, 145–151
190. Williams, J. G. (1999) *Trends Biochem. Sci.* **24**, 333–334
191. Su, L., and David, M. (2000) *J. Biol. Chem.* **275**, 21661–21666
192. Naka, T., Fujimoto, M., and Kishimoto, T. (1999) *Trends Biochem. Sci.* **24**, 394–398
193. Ihle, J. N. (1995) *Nature (London)* **377**, 591–594
194. Ihle, J. N., Witthuhn, B. A., Quelle, F. W., Yamamoto, K., Thierfelder, W. E., Kreider, B., and Silvennoinen, O. (1994) *Trends Biochem. Sci.* **19**, 222–227
195. Chow, D.-c, He, X.-L., Snow, A. L., Rose-John, S., and Garcia, K. C. (2001) *Science* **291**, 2150–2155
196. Revel, M., and Chebath, J. (1986) *Trends Biochem. Sci.* **11**, 166–170
197. Johnson, H. M., Bazer, F. W., Szente, B. E., and Jarpe, M. A. (1994) *Sci. Am.* **270**(May), 68–75
198. Runkel, L., deDios, C., Karpusas, M., Betzenhauser, M., Muldowney, C., Zafari, M., Benjamin, C. D., Miller, S., Hochman, P. S., and Whitty, A. (2000) *Biochemistry* **39**, 2538–2551
199. Li, J., and Roberts, R. M. (1994) *J. Biol. Chem.* **269**, 24826–24833
200. Klaus, W., Gsell, B., Labhardt, A. M., Wipf, B., and Senn, H. (1997) *J. Mol. Biol.* **274**, 661–675
201. Rani, M. R. S., Asthagiri, A. R., Singh, A., Sizemore, N., Sathe, S. S., Li, X., DiDonato, J. D., Stark, G. R., and Ransohoff, R. M. (2001) *J. Biol. Chem.* **276**, 44365–44368
202. Eppstein, D. A., Schryver, B. B., and Marsh, Y. V. (1986) *J. Biol. Chem.* **261**, 5999–6003
203. Sarkar, S. N., Ghosh, A., Wang, H.-W., Sung, S.-S., and Sen, G. C. (1999) *J. Biol. Chem.* **274**, 25535–25542
204. Rebouillat, D., Hovnanian, A., Marié, I., and Hovanessian, A. G. (1999) *J. Biol. Chem.* **274**, 1557–1565
205. Koromilas, A. E., Roy, S., Barber, G. N., Katze, M. G., and Sonenberg, N. (1992) *Science* **257**, 1685–1689
206. Bischoff, J. R., and Samuel, C. E. (1985) *J. Biol. Chem.* **260**, 8237–8239
207. Hardin, J. A., and Thomas, J. O. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7410–7414
208. Sehgal, P. B., May, L. T., Tamm, I., and Vilcek, J. (1987) *Science* **235**, 731–732
209. Gamero, A. M., and Larner, A. C. (2001) *J. Biol. Chem.* **276**, 13547–13553
210. Dinarello, C. A. (1988) *FASEB J.* **2**, 108–115
211. Dinarello, C. A., and Wolff, S. M. (1993) *N. Engl. J. Med.* **328**, 106–112
212. Samad, T. A., Moore, K. A., Sapirstein, A., Billet, S., Allchorne, A., Poole, S., Bonventre, J. V., and Woolf, C. J. (2001) *Nature (London)* **410**, 471–475
213. Yamin, T.-T., Ayala, J. M., and Miller, D. K. (1996) *J. Biol. Chem.* **271**, 13273–13282
214. Vigers, G. P. A., Dripps, D. J., Edwards, C. K., III, and Brandhuber, B. J. (2000) *J. Biol. Chem.* **275**, 36927–36933
215. Simpson, R. J., Hammacher, A., Smith, D. K., Matthews, J. M., and Ward, L. D. (1997) *Protein Sci.* **6**, 929–955
216. Cyster, J. G. (1999) *Science* **286**, 2098–2102
217. Baggiolini, M. (1998) *Nature (London)* **392**, 565–568
218. Mellado, M., Rodríguez-Frade, J. M., Vila-Coro, A. J., Fernández, S., de Ana, A. M., Jones, D. R., Torán, J. L., and Martínez-A, C. (2001) *EMBO J.* **20**, 2497–2507
219. Shao, W., Jerva, L. F., West, J., Lolis, E., and Schweitzer, B. I. (1998) *Biochemistry* **37**, 8303–8313
220. Martinelli, R., Sabroe, I., LaRosa, G., Williams, T. J., and Pease, J. E. (2001) *J. Biol. Chem.* **276**, 42957–42964
221. Ben-Baruch, A., Bengali, K. M., Biragyn, A., Johnston, J. J., Wang, J.-M., Kim, J., Chuntharapai, A., Michiel, D. F., Oppenheim, J. J., and Kelvin, D. J. (1995) *J. Biol. Chem.* **270**, 9121–9128
222. Ben-Baruch, A., Michiel, D. F., and Oppenheim, J. J. (1995) *J. Biol. Chem.* **270**, 11703–11706
222a. Roos, D., and Winterbourn, C. C. (2002) *Science* **296**, 669 – 671
222b. Malkowski, M. G., Wu, J. Y., Lazar, J. B., Johnson, P. H., and Edwards, B. F. P. (1995) *J. Biol. Chem.* **270**, 7077 – 7087
222c. Masuda, M., Suzuki, T., Friesen, M. D., Ravanat, J.-L., Cadet, J., Pignatelli, B., Nishino, H., and Ohshima, H. (2001) *J. Biol. Chem.* **276**, 40486 – 40496
223. Smith, K. A. (1990) *Sci. Am.* **262**(Mar), 50–57
224. Waldmann, T. A. (1991) *J. Biol. Chem.* **266**, 2681–2684
225. Reif, K., Burgering, B. M. T., and Cantrell, D. A. (1997) *J. Biol. Chem.* **272**, 14426–14433
226. Fukushima, K., and Yamashita, K. (2001) *J. Biol. Chem.* **276**, 7351–7356
227. Gadina, M., Sudarashan, C., Visconti, R., Zhou, Y.-J., Gu, H., Neel, B. G., and O'Shea, J. J. (2000) *J. Biol. Chem.* **275**, 26959–26966
227a. Ozaki, K., Spolski, R., Feng, C. G., Qi, C.-F., Cheng, J., Sher, A., Morse, H. C., III, Liu, C., Schwartzberg, P. L., and Leonard, W. J. (2002) *Science* **298**, 1630 – 1634
228. Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., and Clore, G. M. (1992) *Science* **256**, 1673–1677
229. Walter, M. R., Cook, W. J., Zhao, B. G., Cameron, R. P., Jr., Ealick, S. E., Walter, R. L., Jr., Reichert, P., Nagabhushan, T. L., Trotta, P. P., and Bugg, C. E. (1992) *J. Biol. Chem.* **267**, 20371–20376
229a. Taubes, G. (2002) *Science* **296**, 242 – 245
229b. Langen, R. C. J., Schols, A. M. W. J., Kelders, M. C. J. M., Wouters, E. F. M., and Janssen-Heininger, Y. M. W. (2001) *FASEB J.* **15**, 1169 – 1180
229c. Poli, V. (1998) *J. Biol. Chem.* **273**, 29279 – 29282
229d. Dekker, L. V., and Segal, A. W. (2000) *Science* **287**, 982 – 985
229e. Frenette, P. S., and Wagner, D. D. (1996) *N. Engl. J. Med.* **335**, 43 – 45
229f. Albelda, S. M., Smith, C. W., and Ward, P. A. (1994) *FASEB J.* **8**, 504 – 512
229g. Xavier, R. J., and Podolsky, D. K. (2000) *Science* **289**, 1483 – 1484
229h. Ravetch, J. V., and Lanier, L. L. (2000) *Science* **290**, 84 –89
230. Samudzi, C. T., Burton, L. E., and Rubin, J. R. (1991) *J. Biol. Chem.* **266**, 21791–21797
231. Walter, M. R., Windsor, W. T., Nagabhushan, T. L., Lundell, D. J., Lunn, C. A., Zauodny, P. J., and Narula, S. K. (1995) *Nature (London)* **376**, 230–235
231a. Samuel, C. E. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11555 – 11557
231b. Nguyen, K. B., Watford, W. T., Salomon, R., Hofmann, S. R., Pien, G. C., Morinobu, A., Gadina, M., O'Shea, J. J., and Biron, C. A. (2002) *Science* **297**, 2063 –2066
232. Beninga, J., Rock, K. L., and Goldberg, A. L. (1998) *J. Biol. Chem.* **273**, 18734–18742
233. Dhabhar, F. S., Satoskar, A. R., Bluethmann, H., David, J. R., and McEwen, B. S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2846–2851
234. Bazan, J. F. (1995) *Nature (London)* **376**, 217–218
235. Prakash, B., Renault, L., Praefcke, G. J. K., Herrmann, C., and Wittinghofer, A. (2000) *EMBO J.* **19**, 4555–4564
236. Prakash, B., Praefcke, G. J. K., Renault, L., Wittinghofer, A., and Herrmann, C. (2000) *Nature (London)* **403**, 567–571
237. Old, L. J. (1985) *Science* **230**, 630 – 632
238. Argilés, J. M., López-Soriano, J., Busquets, S., and López-Soriano, F. J. (1997) *FASEB J.* **11**, 743–751

References

239. Beutler, B., and Cerami, A. (1988) *Ann. Rev. Biochem.* **57**, 505–518
240. Eck, M. J., Beutler, B., Kuo, G., Merryweather, J. P., and Sprang, S. R. (1988) *J. Biol. Chem.* **263**, 12816–12819
241. Gray, P. W., Aggarwal, B. B., Benton, C. V., Bringman, T. S., Henzel, W. J., Jarrett, J. A., Leung, D. W., Moffat, B., Ng, P., Svedersky, L. P., Palladino, M. A., and Nedwin, G. E. (1984) *Nature (London)* **312**, 721–724
242. Laâbi, Y., and Strasser, A. (2000) *Science* **289**, 883–884
243. Zdanov, A., Schalk-Hihi, C., and Wlodawer, A. (1996) *Protein Sci.* **5**, 1955–1962
244. Scott, P. (1993) *Science* **260**, 496–497
245. Hall, S. S. (1995) *Science* **268**, 1432–1434
246. Sun, H.-W., Bernhagen, J., Bucala, R., and Lolis, E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5191–5196
247. Taylor, A. B., Johnson, W. H., Jr., Czerwinski, R. M., Li, H.-S., Hackert, M. L., and Whitman, C. P. (1999) *Biochemistry* **38**, 7444–7452
248. Nguyen, N. Y., Suzuki, A., Boykins, R. A., and Liu, T.-Y. (1986) *J. Biol. Chem.* **261**, 10456–10465
249. Kirkpatrick, L. L., Matzuk, M. M., Dodds, D. C., and Perin, M. S. (2000) *J. Biol. Chem.* **275**, 17786–17792
250. Coetzee, G. A., Strachan, A. F., van der Westhuyzen, D. R., Hoppe, H. C., Jeenah, M. S., and de Beer, F. C. (1986) *J. Biomol. Struct. Dyn.* **261**, 9644–9651
251. Lowell, C. A., Potter, D. A., Stearman, R. S., and Morrow, J. F. (1986) *J. Biol. Chem.* **261**, 8442–8452
252. Hochepied, T., Van Molle, W., Berger, F. G., Baumann, H., and Libert, C. (2000) *J. Biol. Chem.* **275**, 14903–14909
253. Gehring, M. R., Shiels, B. R., Northemann, W., de Bruijn, M. H. L., Kan, C.-C., Chain, A. C., Noonan, D. J., and Fey, G. H. (1987) *J. Biol. Chem.* **262**, 446–454

253a. Simpson, S. J., and Hines, P. J. (2002) *Science* **296**, 297

254. Levine, M. H., Haberman, A. M., Sant'Angelo, D. B., Hannum, L. G., Cancro, M. P., Janeway, C. A., Jr., and Shlomchik, M. J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2743–2748
255. Otipoby, K. L., Draves, K. E., and Clark, E. A. (2001) *J. Biol. Chem.* **276**, 44315–44322

255a. Malissen, B. (1999) *Science* **285**, 207–208

256. Fischer, M. B., Goerg, S., Shen, L., Prodeus, A. P., Goodnow, C. C., Kelsoe, G., and Carroll, M. C. (1998) *Science* **280**, 582–585
257. Zinkernagel, R. M., and Hengartner, H. (2001) *Science* **293**, 251–253
258. Shaw, S. (1989) *Nature (London)* **338**, 539–540
259. Porcelli, S., Brenner, M. B., Greestein, J. L., Balk, S. P., Terhorst, C., and Bleicher, P. A. (1989) *Nature (London)* **341**, 447–450
260. Feizi, T. (1991) *Trends Biochem. Sci.* **16**, 84–86
261. Barclay, A. N., Birkeland, M. L., Brown, M. H., Beyers, A. D., Davis, S. J., Somoza, C., and Williams, A. F. (1993) *The Leucocyte Antigens Facts Book*, Academic Press, San Diego, California
262. Brady, R. L., Dodson, E. J., Dodson, G. G., Lange, G., Davis, S. J., Williams, A. F., and Barclay, A. N. (1993) *Science* **260**, 979–983
263. Leahy, D. J. (1995) *FASEB J.* **9**, 17–25
264. Veillette, A., Bookman, M. A., Horak, E. M., Samelson, L. E., and Bolen, J. B. (1989) *Nature (London)* **338**, 257–259
265. Grey, H. M., Sette, A., and Buus, S. (1989) *Sci. Am.* **261**(Nov), 56–64
266. von Andrian, U. H., and Mackay, C. R. (2000) *N. Engl. J. Med.* **343**, 1020–1034
267. Cochran, J. R., Aivazian, D., Cameron, T. O., and Lawrence, L. J. (2001) *Trends Biochem. Sci.* **26**, 304–310

267a. Potter, T. A., Grebe, K., Freiberg, B., and Kupfer, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12624–12629

267b. Chen, M., Stafford, W. F., Diedrich, G., Khan, A., and Bouvier, M. (2002) *Biochemistry* **41**, 14539–14545

268. Geleziunas, R., Bour, S., and Wainberg, M. A. (1994) *FASEB J.* **8**, 593–600
269. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) *Nature (London)* **393**, 648–650
270. McMichael, A. J., and Rowland-Jones, S. L. (2001) *Nature (London)* **410**, 980–987

270a. Hill, A. V. S. (1999) *Nature (London)* **398**, 668–669

270b. Welsh, R. M. (2001) *Nature (London)* **411**, 541–542

271. Sieling, P. A., Chatterjee, D., Porcelli, S. A., Prigozy, T. I., Mazzaccaro, R. J., Soriano, T., Bloom, B. R., Brenner, M. B., Kronenberg, M., Brennan, P. J., and Modlin, R. L. (1995) *Science* **269**, 227–230
272. Zeng, Z.-H., Castaño, A. R., Segelke, B. W., Stura, E. A., Peterson, P. A., and Wilson, I. A. (1997) *Science* **277**, 339–345
273. Gumperz, J. E., and Parham, P. (1995) *Nature (London)* **378**, 245–248
274. Maghazachi, A. A., and Al-Aoukaty, A. (1998) *FASEB J.* **12**, 913–924
275. Colonna, M. (1998) *Nature (London)* **391**, 642–643
276. McVicar, D. W., Taylor, L. S., Gosselin, P., Willette-Brown, J., Mikhael, A. I., Geahlen, R. L., Nakamura, M. C., Linnemeyer, P., Seaman, W. E., Anderson, S. K., Ortaldo, J. R., and Mason, L. H. (1998) *J. Biol. Chem.* **273**, 32934–32942
277. Kärre, K., and Schneider, G. (2000) *Nature (London)* **405**, 527–528

277a. MacDonald, H. R. (2002) *Science* **296**, 481–482

277b. Vivier, E., and Biron, C. A. (2002) *Science* **296**, 1248–1249

278. Irwin, M., McClintick, J., Costlow, C., Fortner, M., White, J., and Gillin, J. C. (1996) *FASEB J.* **10**, 643–653
279. Steinmetz, M., and Hood, L. (1983) *Science* **222**, 727–733
280. Lu, C. Y., Khair-El-Din, T. A., Dawidson, I. A., Butler, T. M., Brasky, K. M., Vazquez, M. A., and Sicher, S. C. (1994) *FASEB J.* **8**, 1122–1130
281. Steinmetz, M. (1984) *Trends Biochem. Sci.* **9**, 224–226
282. Flavell, R. A., Allen, H., Burkly, L. C., Sherman, D. H., Waneck, G. L., and Widera, G. (1986) *Science* **233**, 437–443
283. Kappes, D., and Strominger, J. L. (1988) *Ann. Rev. Biochem.* **57**, 991–1028
284. MHC sequencing consortium. (1999) *Nature (London)* **401**, 921–923
285. Klein, J., and Sato, A. (2000) *N. Engl. J. Med.* **343**, 702–709
286. Radley, E., Alderton, R. P., Kelly, A., Trowsdale, J., and Beck, S. (1994) *J. Biol. Chem.* **269**, 18834–18838
287. Nathenson, S. G., Uehara, H., Ewenstein, B. M., Kindt, T. J., and Coligan, J. E. (1981) *Ann. Rev. Biochem.* **50**, 1025–1052
288. Åkerström, B., and Lögdberg, L. (1990) *Trends Biochem. Sci.* **15**, 240–243
289. Glassy, M. C. (1982) *Trends Biochem. Sci.* **7**, 286–288
290. Arguello, R., Avakian, H., Goldman, J. M., and Madrigal, J. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10961–10965
291. Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L., and Wiley, D. C. (1989) *Nature (London)* **342**, 692–696
292. Parham, P. (1989) *Nature (London)* **340**, 426–428
293. Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1994) *Nature (London)* **368**, 215–221
294. Ghosh, P., Amaya, M., Mellins, E., and Wiley, D. C. (1995) *Nature (London)* **378**, 457–462
295. Zinkernagel, R. M. (1979) *Ann. Rev. Microbiol.* **33**, 201–213
296. Morel, P. A., Dorman, J. S., Todd, J. A., McDevitt, H. O., and Trucco, M. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8111–8115
297. Erlich, H. A. (1989) *Nature (London)* **337**, 415
298. Forcione, D. G., Sands, B., Isselbacher, K. J., Rustgi, A., Podolsky, D. K., and Pillai, S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5094–5098
299. Goldberg, A. L., and Rock, K. L. (1992) *Nature (London)* **357**, 375–379
300. Kuttler, C., Nussbaum, A. K., Dick, T. P., Rammensee, H.-G., Schild, H., and Hadeler, K.-P. (2000) *J. Mol. Biol.* **298**, 417–429
301. Shimbara, N., Ogawa, K., Hidaka, Y., Nakajima, H., Yamasaki, N., Niwa, S.-i, Tanahashi, N., and Tanaka, K. (1998) *J. Biol. Chem.* **273**, 23062–23071
302. Heemels, M.-T., and Ploegh, H. (1995) *Ann. Rev. Biochem.* **64**, 463–491
303. Cascio, P., Hilton, C., Kisselev, A. F., Rock, K. L., and Goldberg, A. L. (2001) *EMBO J.* **20**, 2357–2366
304. de la Salle, H., and 16 other authors. (1994) *Science* **265**, 237–241
305. Binder, R. J., Blachere, N. E., and Srivastava, P. K. (2001) *J. Biol. Chem.* **276**, 17163–17171

305a. van Endert, P. M., Saveanu, L., Hewitt, E. W., and Lehner, P. J. (2002) *Trends Biochem. Sci.* **27**, 454–461

306. Vogt, A. B., and Kropshofer, H. (1999) *Trends Biochem. Sci.* **24**, 150–154
307. Fremont, D. H., Hendrickson, W. A., Marrack, P., and Kappler, J. (1996) *Science* **272**, 1001–1004
308. McFarland, B. J., Sant, A. J., Lybrand, T. P., and Beeson, C. (1999) *Biochemistry* **38**, 16663–16670
309. Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 734–738
310. Wang, J.-h, Meijers, R., Xiong, Y., Liu, J.-h, Sakihama, T., Zhang, R., Joachimiak, A., and Reinherz, E. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10799–10804
311. Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1993) *Nature (London)* **364**, 33–39
312. Matsumura, M., Fremont, D. H., Peterson, P. A., and Wilson, I. A. (1992) *Science* **257**, 927–934
313. Fahnestock, M. L., Johnson, J. L., Feldman, R. M. R., Tsomides, T. J., Mayer, J., Narhi, L. O., and Bjorkman, P. J. (1994) *Biochemistry* **33**, 8149–8158
314. Young, A. C. M., Nathenson, S. G., and Sacchettini, J. C. (1995) *FASEB J.* **9**, 26–36
315. Udaka, K. (1996) *Trends Biochem. Sci.* **21**, 7–11
316. Parham, P., and Ohta, T. (1996) *Science* **272**, 67–74
317. Zhang, C., Anderson, A., and DeLisi, C. (1998) *J. Mol. Biol.* **281**, 929–947
318. Parham, P. (1996) *Trends Biochem. Sci.* **21**, 427–433

318a. Rammensee, H.-G. (2002) *Nature (London)* **419**, 443–445

References

319. Wilson, I. A. (1996) *Science* **272**, 973–974
320. Liang, M. N., Lee, C., Xia, Y., and McConnell, H. M. (1996) *Biochemistry* **35**, 14734–14742
321. Bénaroch, P., Yilla, M., Raposo, G., Ito, K., Miwa, K., Geuze, H. J., and Ploegh, H. L. (1995) *EMBO J.* **14**, 37–49
322. Kropshofer, H., Vogt, A. B., Thery, C., Armandola, E. A., Li, B.–C., Moldenhauer, G., Amigorena, S., and Hämmerling, G. J. (1998) *EMBO J.* **17**, 2971–2981
323. Bogyo, M., and Ploegh, H. L. (1998) *Nature (London)* **396**, 625–627
324. Watts, C. (2001) *Science* **294**, 1294–1295
325. Gao, G. F., Tormo, J., Gerth, U. C., Wyer, J. R., McMichael, A. J., Stuart, D. I., Bell, J. I., Jones, E. Y., and Jakobsen, B. K. (1997) *Nature (London)* **387**, 630–634
326. Smith, A. B., III, Benowitz, A. B., Sprengeler, P. A., Barbosa, J., Guzman, M. C., Hirschmann, R., Schweiger, E. J., Bolin, D. R., Nagy, Z., Campbell, R. M., Cox, D. C., and Olson, G. L. (1999) *J. Am. Chem. Soc.* **121**, 9286–9298
327. Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H. N., and Tonegawa, S. (1984) *Nature (London)* **309**, 757–762
328. Arden, B., Klotz, J. L., Siu, G., and Hood, L. E. (1985) *Nature (London)* **316**, 783–787
329. Garcia, K. C., Degano, M., Stanfield, R. L., Brunmark, A., Jackson, M. R., Peterson, P. A., Teyton, L., and Wilson, I. A. (1996) *Science* **274**, 209–219
330. Swan, K. A., Alberola-lla, J., Gross, J. A., Appleby, M. W., Forbush, K. A., Thomas, J. F., and Perlmutter, R. M. (1995) *EMBO J.* **14**, 276–285
330a. Rudolph, M. G., Huang, M., Teyton, L., and Wilson, I. A. (2001) *J. Mol. Biol.* **314**, 1 – 8
331. Abraham, R. T., Karnitz, L. M., Secrist, J. P., and Leibson, P. J. (1992) *Trends Biochem. Sci.* **17**, 434–438
332. Hatada, M. H., Lu, X., Laird, E. R., Green, J., Morgenstern, J. P., Lou, M., Marr, C. S., Phillips, T. B., Ram, M. K., Theriault, K., Zoller, M. J., and Karas, J. L. (1995) *Nature (London)* **377**, 32–38
333. Reinherz, E. L., Tan, K., Tang, L., Kern, P., Liu, J.-h, Xiong, Y., Hussey, R. E., Smolyar, A., Hare, B., Zhang, R., Joachimiak, A., Chang, H.-C., Wagner, G., and Wang, J.-h. (1999) *Science* **286**, 1913–1921
334. Hennecke, J., Carfi, A., and Wiley, D. C. (2000) *EMBO J.* **19**, 5611–5624
335. Germain, R. N. (2001) *J. Biol. Chem.* **276**, 35223–35226
336. Lang, P., Stolpa, J. C., Freiberg, B. A., Crawford, F., Kappler, J., Kupfer, A., and Cambier, J. C. (2001) *Science* **291**, 1537–1540
337. Howe, L. R., and Weiss, A. (1995) *Trends Biochem. Sci.* **20**, 59–64
338. LoGrasso, P. V., Hawkins, J., Frank, L. J., Wisniewski, D., and Marcy, A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12165–12170
338a. Bottomly, K. (1999) *Science* **283**, 1124 – 1125
338b. Edinger, A. L., and Thompson, C. B. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1107 – 1109
338c. Hudrisier, D., and Bongrand, P. (2002) *FASEB J.* **16**, 477 – 486
339. Janis, E. M., Kaufmann, S. H. E., Schwartz, R. H., and Pardoll, D. M. (1989) *Science* **244**, 713–716
340. Steinle, A., Groh, V., and Spies, T. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12510–12515
341. Wingren, C., Crowley, M. P., Degano, M., Chien, Y.-h, and Wilson, I. A. (2000) *Science* **287**, 310–314
342. Allison, T. J., Winter, C. C., Fournié, J.-J., Bonneville, M., and Garboczi, D. N. (2001) *Nature (London)* **411**, 820–823
342a. Belmant, C., Espinosa, E., Poupot, R., Peyrat, M.-A., Guiraud, M., Poquet, Y., Bonneville, M., and Fournié, J.-J. (1999) J. Biol. Chem. **274**, 32079 – 32084
343. Johnson, H. M., Russell, J. K., and Pontzer, C. H. (1992) *Sci. Am.* **266**(April), 92 – 101
344. Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Chi, Y., Stauffacher, C., Strominger, J. L., and Wiley, D. C. (1994) *Nature (London)* **368**, 711–718
345. Abrahmsén, L., Dohlsten, M., Segrén, S., Björk, P., Jonsson, E., and Kalland, T. (1995) *EMBO J.* **14**, 2978–2986
346. Håkansson, M., Petersson, K., Nilsson, H., Forsberg, G., Björk, P., Antonsson, P., and Svensson, L. A. (2000) *J. Mol. Biol.* **302**, 527–537
347. Arcus, V. L., Proft, T., Sigrell, J. A., Baker, H. M., Fraser, J. D., and Baker, E. N. (2000) *J. Mol. Biol.* **299**, 157–168
348. Vath, G. M., Earhart, C. A., Monie, D. D., Iandolo, J. J., Schlievert, P. M., and Ohlendorf, D. H. (1999) *Biochemistry* **38**, 10239–10246
349. Papageorgiou, A. C., Brehm, R. D., Leonidas, D. D., Tranter, H. S., and Acharaya, K. R. (1996) *J. Mol. Biol.* **260**, 553–569
350. Lehnert, N. M., Allen, D. L., Allen, B. L., Catasti, P., Shiflett, P. R., Chen, M., Lehnert, B. E., and Gupta, G. (2001) *Biochemistry* **40**, 4222–4228
350a. Mackay, I. R. (1999) *Science* **284**, 269 – 270
350b. Podulsky, S. H., and Tauber, A. I. (1999) *The Generation of Diversity Clonal Selection*, Harvard Univ. Press, Cambridge, Massachusetts
351. Ada, G. L., and Nossal, G. (1987) *Sci. Am.* **257**(Aug), 62–69
352. Arrhenius, S. (1907) *Immunochemistry*, Macmillan Co., New York
353. Haurowitz, F. (1979) *Trends Biochem. Sci.* **4**, N 268–N 270
353a. Steinman, R. M., and Nussenzweig, M. C. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 351 – 358
354. Howard, J. C. (1985) *Nature (London)* **314**, 494–495
355. Jerne, N. K. (1985) *Science* **229**, 1057– 1059
356. Ban, N., Day, J., Wang, X., Ferrone, S., and McPherson, A. (1996) *J. Mol. Biol.* **255**, 617–627
356a. Greenspan, N. S., and Bona, C. A. (1993) *FASEB J.* **7**, 437 – 444
357. Vaux, D., Tooze, J., and Fuller, S. (1990) *Nature (London)* **345**, 495–502
358. Marshall, E. (1996) *Science* **273**, 174–175
359. Rajewsky, K. (1996) *Nature (London)* **381**, 751–758
360. Forsdyke, D. R. (1995) *FASEB J.* **9**, 164–166
361. Goodnow, C. C., Crosbie, J., Jorgensen, H., Brink, R. A., and Basten, A. (1989) *Nature (London)* **342**, 385–391
362. Boise, L. H., and Thompson, C. B. (1996) *Science* **274**, 67–68
363. Kamradt, T., and Mitchison, N. A. (2001) *N. Engl. J. Med.* **344**, 655–664
364. Van Parijs, L., and Abbas, A. K. (1998) *Science* **280**, 243 – 248
365. Crabtree, G. R. (1989) *Science* **243**, 355–361
366. Mondino, A., Khoruts, A., and Jenkins, M. K. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2245–2252
367. Barton, G. M., and Rudensky, A. Y. (1999) *Science* **283**, 67–70
368. Marrack, P. (1987) *Science* **235**, 1311–1313
369. Janeway, C. A., Jr. (1988) *Nature (London)* **335**, 208–210
370. Ramsdell, F., and Fowlkes, B. J. (1990) *Science* **248**, 1342–1348
371. Schwartz, R. H. (1993) *Sci. Am.* **269**(Aug), 62–71
372. Akdis, C. A., and Blaser, K. (1999) *FASEB J.* **13**, 603–609
373. Weiner, H. L. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10762–10765
373a. Schwartzberg, P. L. (2001) *Science* **293**, 228 – 229
373b. Petty, H. R., and Kindzelskii, A. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3145 – 3149
373c. Germain, R. N. (2001) *Science* **293**, 240 – 245
373d. Lu, Q., and Lemke, G. (2001) *Science* **293**, 306 – 311
374. Rosen, F. S., and Mackay, I. (2001) *N. Engl. J. Med.* **345**, 1343–1344
375. Nabel, G. J. (2001) *Nature (London)* **410**, 1002–1007
376. Ada, G. (2001) *N. Engl. J. Med.* **345**, 1042–1053
377. Clarke, T. (2001) *Nature (London)* **409**, 278–280
377a. Cohen, J. (2001) *Science* **294**, 985
378. Cohen, A. D., Boyer, J. D., and Weiner, D. B. (1998) *FASEB J.* **12**, 1611–1626
379. Shen, X., and Siliciano, R. F. (2000) *Science* **290**, 463–465
379a. Modlin, R. L. (2000) *Nature (London)* **408**, 659 – 660
380. Langridge, W. H. R. (2000) *Sci. Am.* **283**(Sep), 66–71
381. Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S., and Small, P. M. (1999) *Science* **284**, 1520–1523
382. Enserink, M. (2001) *Science* **293**, 234–235
383. Sprent, J., and Tough, D. F. (2001) *Science* **293**, 245–248
384. Barouch, D. H., and 28 other authors. (2000) *Science* **290**, 486–492
385. Riddell, S. R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8933–8935
385a. Rappuoli, R., Miller, H. I., and Falkow, S. (2002) *Science* **297**, 937, 939
386. Simpson, S., and Marshall, E. (2001) *Science* **293**, 233
387. Waldschmidt, T. J., and Noelle, R. J. (2001) *Science* **293**, 2012–2013
388. Martin, S., and Goodnow, C. (2000) *Nature (London)* **407**, 576–577
389. Hayday, A., and Viney, J. L. (2000) *Science* **290**, 97–100
390. Morales, J., Homey, B., Vicari, A. P., Hudak, S., Oldham, E., Hedrick, J., Orozco, R., Copeland, N. G., Jenkins, N. A., McEvoy, L. M., and Zlotnik, A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14470–14475
391. Lanzavecchia, A., and Sallusto, F. (2000) *Science* **290**, 92–97
392. Ku, C. C., Murakami, M., Sakamoto, A., Kappler, J., and Marrack, P. (2000) *Science* **288**, 675–678
393. Tonegawa, S. (1983) *Nature (London)* **302**, 575–581
394. Marx, J. L. (1987) *Science* **238**, 484–485
395. Honjo, T., and Habu, S. (1985) *Ann. Rev. Biochem.* **54**, 803–830
396. Kodaira, M., Kinashi, T., Umemura, I., Matsuda, F., Noma, T., Ono, Y., and Honjo, T. (1986) *J. Mol. Biol.* **190**, 529–541
397. Robertson, M. (1985) *Nature (London)* **317**, 768–771
398. Williams, S. C., Frippiat, J.-P., Tomlinson, I. M., Ignatovich, O., Lefranc, M.-P., and Winter, G. (1996) *J. Mol. Biol.* **264**, 220–232
399. Ignatovich, O., Tomlinson, I. M., Popov, A. V., Brüggemann, M., and Winter, G. (1999) *J. Mol. Biol.* **294**, 457–465
400. Tomlinson, I. M., Cox, J. P. L., Gherardi, E., Lesk, A. M., and Chothia, C. (1995) *EMBO J.* **14**, 4628–4638
401. Halligan, B. D., and Desiderio, S. V. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7019–7023
402. Oettinger, M. A., Schatz, D. G., Gorka, C., and Baltimore, D. (1990) *Science* **248**, 1517–1523

References

402a. Agrawal, A., Eastman, Q. M., and Schatz, D. G. (1998) *Nature (London)* **394**, 744 – 751
402b. Gellert, M. (2002) *Ann. Rev. Biochem.* **71**, 101 – 132
403. Bhagavan, N. V. (2002) *Medical Biochemistry*, 4th ed., Harcourt/Academic Press, San Diego, California
404. Hope, T. J., Aguilera, R. J., Minie, M. E., and Sakano, H. (1986) *Science* **231**, 1141 – 1145
405. Roman, C. A. J., and Baltimore, D. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2333 – 2338
406. Lin, W.-C., and Desiderio, S. (1993) *Science* **260**, 953 – 959
407. McMurry, M. T., and Krangel, M. S. (2000) *Science* **287**, 495 – 498
408. Jones, J. M., and Gellert, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12926 – 12931
409. Tevelev, A., and Schatz, D. G. (2000) *J. Biol. Chem.* **275**, 8341 – 8348
410. Frank, K. M., Sekiguchi, J. M., Seidl, K. J., Swat, W., Rathbun, G. A., Cheng, H.-L., Davidson, L., Kangaloo, L., and Alt, F. W. (1998) *Nature (London)* **396**, 173 – 177
411. Chu, G. (1997) *J. Biol. Chem.* **272**, 24097 – 24100
412. Plasterk, R. (1998) *Nature (London)* **394**, 718 – 719
413. Cherry, S. R., Beard, C., Jaenisch, R., and Baltimore, D. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8467 – 8472
413a. Fugmann, S. D. (2002) *Nature (London)* **416**, 691 – 694
414. Baltimore, D. (1986) *Nature (London)* **319**, 12 – 13
415. Raghavan, S. C., Kirsch, I. R., and Lieber, M. R. (2001) *J. Biol. Chem.* **276**, 29126 – 29133
416. Pandey, V. N., and Modak, M. J. (1989) *J. Biol. Chem.* **264**, 867 – 871
417. Gough, N. (1983) *Trends Biochem. Sci.* **8**, 227 – 228
418. Rada, C., and Milstein, C. (2001) *EMBO J.* **20**, 4570 – 4576
419. Milstein, C., Neuberger, M. S., and Staden, R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8791 – 8794
420. Tomlinson, I. M., Walter, G., Jones, P. T., Dear, P. H., Sonnhammer, E. L. L., and Winter, G. (1996) *J. Mol. Biol.* **256**, 813 – 817
421. Kelsoe, G. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6576 – 6577
422. Papavasiliou, F. N., and Schatz, D. G. (2000) *Nature (London)* **408**, 216 – 221
423. Foster, P. L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7617 – 7618
423a. Arakawa, H. Hauschild, J., and Buerstedde, J.-M. (2002) *Science* **295**, 1301 – 1306
423b. Fugmann, S. D., and Schatz, D. G. (2002) *Science* **295**, 1244 – 1245
424. de Wildt, R. M. T., Hoet, R. M. A., van Venrooij, W. J., Tomlinson, I. M., and Winter, G. (1999) *J. Mol. Biol.* **285**, 895 – 901
424a. Neuberger, M. S., and Scott, J. (2000) *Science* **289**, 1705 – 1706
424b. Di Noia, J., and Neuberger, M. S. (2002) *Nature (London)* **419**, 43 – 48
424c. Faili, A., Aoufouchi, S., Flatter, E., Guéranger, Q., Reynaud, C.-A., and Weill, J.-C. (2002) *Nature (London)* **419**, 944 – 947
424d. Petersen, S., and 15 other authors. (2001) *Nature (London)* **414**, 660 – 665
425. Casellas, R., Shih, T.-A. Y., Kleinewietfeld, M., Rakonjac, J., Nemazee, D., Rajewsky, K., and Nussenzweig, M. C. (2001) *Science* **291**, 1541 – 1544
426. Reth, M., Gehrmann, P., Petrac, E., and Wiese, P. (1986) *Nature (London)* **322**, 840 – 842
427. Ritchie, K. A., Brinster, R. L., and Storb, U. (1984) *Nature (London)* **312**, 517 – 520
428. Mostoslavsky, R., Singh, N., Tenzen, T., Goldmit, M., Gabay, C., Elizur, S., Qi, P., Reubinoff, B. E., Chess, A., Cedar, H., and Bergman, Y. (2001) *Nature (London)* **414**, 221 – 225
429. Noma, Y., Sideras, P., Naito, T., Bergstedt-Lindquist, S., Azuma, C., Severinson, E., Tanabe, T., Kinashi, T., Matsuda, F., Yaoita, Y., and Honjo, T. (1986) *Nature (London)* **319**, 640 – 646
430. Geha, R. S., and Rosen, F. S. (1994) *N. Engl. J. Med.* **330**, 1008 – 1009
431. Fukita, Y., Mizuta, T.-R., Shirozu, M., Ozawa, K., Shimizu, A., and Honjo, T. (1993) *J. Biol. Chem.* **268**, 17463 – 17470
432. Stavnezer, J. (2000) *Science* **288**, 984 – 985
433. Levitzki, A., and Gazit, A. (1995) *Science* **267**, 1782 – 1787
434. Wallin, J. J., Rinkenberger, J. L., Rao, S., Gackstetter, E. R., Koshland, M. E., and Zwollo, P. (1999) *J. Biol. Chem.* **274**, 15959 – 15965
435. Goldwrth, A. W., and Bevan, M. J. (1999) *Nature (London)* **402**, 255 – 262
436. Yancopoulos, G. D., Blackwell, T. K., Suh, H., Hood, L., and Alt, F. W. (1986) *Cell* **44**, 251 – 259
437. Bousso, P., Wahn, V., Douagi, I., Horneff, G., Pannetier, C., Le Deist, F., Zepp, F., Niehues, T., Kourilsky, P., Fischer, A., and de Saint Basile, G. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 274 – 278
438. Arstila, T. P., Casrouge, A., Baron, V., Even, J., Kanellopoulos, J., and Kourilsky, P. (1999) *Science* **286**, 958 – 961
439. Davodeau, F., Difilippantonio, M., Roldan, E., Malissen, M., Casanova, J.-L., Couedel, C., Morcet, J.-F., Merkenschlager, M., Nussenzweig, A., Bonneville, M., and Malissen, B. (2001) *EMBO J.* **20**, 4717 – 4729
440. Buisseret, P. D. (1982) *Sci. Am.* **247**(Aug), 86 – 95
441. Lichtenstein, L. M. (1993) *Sci. Am.* **269**(Sep), 117 – 124
442. Kay, A. B. (2001) *N. Engl. J. Med.* **344**, 30 – 37
443. Cook, J. P. D., Henry, A. J., McDonnell, J. M., Owens, R. J., Sutton, B. J., and Gould, H. J. (1997) *Biochemistry* **36**, 15579 – 15588
444. Corry, D. B., and Kheradmand, F. (1999) *Nature (London)* **402, Supp.**, B18 – B23
445. Turner, H., and Kinet, J.-P. (1999) *Nature (London)* **402, Supp.**, B24 – B30
446. Joseph, M., Auriault, C., Capron, A., Vorng, H., and Viens, P. (1983) *Nature (London)* **303**, 810 – 812
447. Dweik, R. A., Comhair, S. A. A., Gaston, B., Thunnissen, F. B. J. M., Farver, C., Thomassen, M. J., Kavuru, M., Hammel, J., Abu-Soud, H. M., and Erzurum, S. C. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2622 – 2627
448. Cookson, W. (1999) *Nature (London)* **402, Supp.**, B5 – B11
448a. Yazdanbakhsh, M., Kremsner, P. G., and van Ree, R. (2002) *Science* **296**, 490 – 494
449. Pomés, A., Melén, E., Vailes, L. D., Retief, J. D., Arruda, L. K., and Chapman, M. D. (1998) *J. Biol. Chem.* **273**, 30801 – 30807
450. Ichikawa, S., Hatanaka, H., Yuuki, T., Iwamoto, N., Kojima, S., Nishiyama, C., Ogura, K., Okumura, Y., and Inagaki, F. (1998) *J. Biol. Chem.* **273**, 356 – 360
451. Mueller, G. A., Smith, A. M., Chapman, M. D., Rule, G. S., and Benjamin, D. C. (2001) *J. Biol. Chem.* **276**, 9359 – 9365
452. Rosenstreich, D. L., Eggleston, P., Kattan, M., Baker, D., Slavin, R. G., Gergen, P., Mitchell, H., McNiff-Mortimer, K., Lynn, H., Ownby, D., and Malveaux, F. (1997) *N. Engl. J. Med.* **336**, 1356 – 1363
453. Baur, X., Aschauer, H., Mazur, G., Dewair, M., Prelicz, H., and Steigemann, W. (1986) *Science* **233**, 351 – 354
454. Lascombe, M.-B., Grégoire, C., Poncet, P., Tavares, G. A., Rosinski-Chupin, I., Rabillon, J., Goubran-Botros, H., Mazié, J.-C., David, B., and Alzari, P. M. (2000) *J. Biol. Chem.* **275**, 21572 – 21577
455. Rouvinen, J., Rautiainen, J., Virtanen, T., Zeiler, T., Kauppinen, J., Taivainen, A., and Mäntyjärvi, R. (1999) *J. Biol. Chem.* **274**, 2337 – 2343
456. Woodfolk, J. A., Wheatley, L. M., Piyasena, R. V., Benjamin, D. C., and Platts-Mills, T. A. E. (1998) *J. Biol. Chem.* **273**, 29489 – 29496
457. Ansari, A. A., Shenbagamurthi, P., and Marsh, D. G. (1989) *J. Biol. Chem.* **264**, 11181 – 11185
458. Metzler, W. J., Valentine, K., Roebber, M., Marsh, D. G., and Mueller, L. (1992) *Biochemistry* **31**, 8697 – 8705
459. Villalba, M., Batanero, E., Monsalve, R. I., González de la Peña, M. A., and Lahoz, C. (1994) *J. Biol. Chem.* **269**, 15217 – 15222
460. Slater, J. E., Vedvick, T., Arthur-Smith, A., Trybul, D. E., and Kekwick, R. G. O. (1996) *J. Biol. Chem.* **271**, 25394 – 25399
461. Akasawa, A., Hsieh, L.-S., Martin, B. M., Liu, T., and Lin, Y. (1996) *J. Biol. Chem.* **271**, 25389 – 25393
461a. Sampson, H. A. (2002) *N. Engl. J. Med.* **346**, 1294 – 1299
462. Wilkinson, S. L. (1998) *Chem. Eng. News* **Sep 7**, 38 – 40
463. Ferreira, F., Ebner, C., Kramer, B., Casari, G., Briza, P., Kungl, A. J., Grimm, R., Jahn-Schmid, B., Breiteneder, H., Kraft, D., Breitenbach, M., Rheinberger, H.-J., and Scheiner, O. (1998) *FASEB J.* **12**, 231 – 242
464. Busse, W. W., and Lemanske, R. F., Jr. (2001) *N. Engl. J. Med.* **344**, 350 – 361
465. Cookson, W. O. C. M., and Moffatt, M. F. (1997) *Science* **275**, 41 – 42
466. Barnes, P. J. (1991) *Trends Biochem. Sci.* **16**, 365 – 369
467. Gounni, A. S., Lamkhioued, B., Koussih, L., Ra, C., Renzi, P. M., and Hamid, Q. (2001) *FASEB J.* **15**, 940 – 949
468. Matsuoka, T., and 17 other authors. (2000) *Science* **287**, 2013 – 2017
469. Grünig, G., Warnock, M., Wakil, A. E., Venkayya, R., Brombacher, F., Rennick, D. M., Sheppard, D., Mohrs, M., Donaldson, D. D., Locksley, R. M., and Corry, D. B. (1998) *Science* **282**, 2261 – 2263
470. Bonini, S., Lambiase, A., Bonini, S., Angelucci, F., Magrini, L., Manni, L., and Aloe, L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10955 – 10960
471. Comhair, S. A. A., Bhathena, P. R., Farver, C., Thunnissen, F. B. J. M., and Erzurum, S. C. (2001) *FASEB J.* **15**, 70 – 78
472. Mishra, A., Weaver, T. E., Beck, D. C., and Rothenberg, M. E. (2001) *J. Biol. Chem.* **276**, 8453 – 8459
473. Khubchandani, K. R., and Snyder, J. M. (2001) *FASEB J.* **15**, 59 – 69
474. Davidson, A., and Diamond, B. (2001) *N. Engl. J. Med.* **345**, 340 – 350
475. Ridgway, W. M., Fassò, M., and Fathman, C. G. (1999) *Science* **284**, 749 – 751
476. Mitchison, N. A., and Wedderburn, L. R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8750 – 8751
477. Rose, N. R. (1981) *Sci. Am.* **244**, 80 – 98
478. Chen, C.-R., Tanaka, K., Chazenbalk, G. D., McLachlan, S. M., and Rapoport, B. (2001) *J. Biol. Chem.* **276**, 14767 – 14772

References

479. Corper, A. L., Stratmann, T., Apostolopoulos, V., Scott, C. A., Garcia, K. C., Kang, A. S., Wilson, I. A., and Teyton, L. (2000) *Science* **288**, 505 – 511
480. Wilson, S. B., Kent, S. C., Horton, H. F., Hill, A. A., Bollyky, P. L., Hafler, D. A., Strominger, J. L., and Byrne, M. C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7411 – 7416
481. Schwartz, H. L., Chandonia, J.-M., Kash, S. F., Kanaani, J., Tunnell, E., Domingo, A., Cohen, F. E., Banga, J. P., Madec, A.-M., Richter, W., and Baekkeskov, S. (1999) *J. Mol. Biol.* **287**, 983 – 999
482. Yoon, J.-W., Yoon, C.-S., Lim, H.-W., Huang, Q. Q., Kang, Y., Pyun, K. H., Hirasawa, K., Sherwin, R. S., and Jun, H.-S. (1999) *Science* **284**, 1183 – 1187
483. von Boehmer, H., and Sarukhan, A. (1999) *Science* **284**, 1135 – 1137
484. Gianani, R., and Sarvetnick, N. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2257 – 2259
485. Nepom, G. T., Lippolis, J. D., White, F. M., Masewicz, S., Marto, J. A., Herman, A., Luckey, C. J., Falk, B., Shabanowitz, J., Hunt, D. F., Engelhard, V. H., and Nepom, B. S. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1763 – 1768
486. Mathis, D., Vence, L., and Benoist, C. (2001) *Nature (London)* **414**, 792 – 798
487. Kumar, S., Kalsi, J., Latchman, D. S., Pearl, L. H., and Isenberg, D. A. (2001) *J. Mol. Biol.* **308**, 527 – 539
488. Stevens, S. Y., and Glick, G. D. (1999) *Biochemistry* **38**, 560 – 568

488a. Marshall, E. (2002) *Science* **296**, 689 – 691

489. Witte, S., Neumann, F., Krawinkel, U., and Przybylski, M. (1996) *J. Biol. Chem.* **271**, 18171 – 18175
490. Muller, S., Briand, J.-P., and Van Regenmortel, M. H. V. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8176 – 8180
491. Pereira, B., Benedict, C. R., Le, A., Shapiro, S. S., and Thiagarajan, P. (1998) *Biochemistry* **37**, 1430 – 1437
492. Wieben, E. D., Rohleder, A. M., Nenninger, J. M., and Pederson, T. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7914 – 7918
493. Morel, L., Blenman, K. R., Croker, B. P., and Wakeland, E. K. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1787 – 1792
494. Reddy, R., Tan, E. M., Henning, E., Nohga, K., and Busch, H. (1983) *J. Biol. Chem.* **258**, 1383 – 1386
495. Bolívar, J., Guelman, S., Iglesias, C., Ortíz, M., and Valdivia, M. M. (1998) *J. Biol. Chem.* **273**, 17122 – 17127
496. Schultz, G. E., and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York

496a. Vinuesa, C. G., and Goodnow, C. C. (2002) *Nature (London)* **416**, 595 – 598

497. Kotzin, B. L., Falta, M. T., Crawford, F., Rosloniec, E. F., Bill, J., Marrack, P., and Kappler, J. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 291 – 296
498. Nishiura, H., Shibuya, Y., Matsubara, S., Tanase, S., Kambara, T., and Yamamoto, T. (1996) *J. Biol. Chem.* **271**, 878 – 882
499. Hellmark, T., Burkhardt, H., and Wieslander, J. (1999) *J. Biol. Chem.* **274**, 24862 – 25868
500. Netzer.K-O, Leinonen, A., Boutaud, A., Borza, D.-B., Todd, P., Gunwar, S., Langeveld, J. P. M., and Hudson, B. G. (1999) *J. Biol. Chem.* **274**, 11267 – 11274
501. Gunnarsson, A., Hellmark, T., and Wieslander, J. (2000) *J. Biol. Chem.* **275**, 30844 – 30848
502. Haneji, N., Nakamura, T., Takio, K., Yanagi, K., Higashiyama, H., Saito, I., Noji, S., Sugino, H., and Hayashi, Y. (1997) *Science* **276**, 604 – 607
503. Yip, T.-T., Van de Water, J., Gershwin, M. E., Coppel, R. L., and Hutchens, T. W. (1996) *J. Biol. Chem.* **271**, 32825 – 32833
504. Dang, C. V., Tan, E. M., and Traugh, J. A. (1988) *FASEB J.* **2**, 2376 – 2379
505. Miyata, T., Yamada, N., Iida, Y., Nishimura, J., Takeda, J., Kitani, T., and Kinoshita, T. (1994) *N. Engl. J. Med.* **330**, 249 – 255

505a. Farrell, R. J., and Kelly, C. P. (2002) *N. Engl. J. Med.* **346**, 180 – 188

505b. Shan, L., Molberg, O., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., Sollid, L. M., and Khosla, C. (2002) *Science* **297**, 2275 – 2279

505c. Schuppan, D., and Hahn, E. G. (2002) *Science* **297**, 2218 – 2220

505d. Fesus, L., and Piacentini, M. (2002) *Trends Biochem. Sci.* **27**, 534 – 539

506. Shields, D. C., Tyor, W. R., Deibler, G. E., Hogan, E. L., and Banik, N. L. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5768 – 5772
507. Noseworthy, J. H., Lucchinetti, C., Rodriguez, M., and Weinshenker, B. G. (2000) *N. Engl. J. Med.* **343**, 938 – 952
508. Li, Y., Li, H., Martin, R., and Mariuzza, R. A. (2000) *J. Mol. Biol.* **304**, 177 – 188
509. Rogers, S. W., Andrews, P. I., Gahring, L. C., Whisenand, T., Cauley, K., Crain, B., Hughes, T. E., Heinemann, S. F., and McNamara, J. O. (1994) *Science* **265**, 648 – 651
510. Byrne, G. (1988) *Science* **242**, 198
511. Kallen, J., Mikol, V., Taylor, P., and Walkinshaw, M. D. (1998) *J. Mol. Biol.* **283**, 435 – 449
512. Montague, J. W., Hughes, F. M., Jr., and Cidlowski, J. A. (1997) *J. Biol. Chem.* **272**, 6677 – 6684
513. Mackay, I. R., and Rosen, F. S. (2000) *N. Engl. J. Med.* **343**, 1313 – 1324
514. Greenberg, P. D., and Riddell, S. R. (1999) *Science* **285**, 546 – 551
515. Fischer, A., and Malissen, B. (1998) *Science* **280**, 237 – 243
516. Singer, L., Whitehead, W. T., Akama, H., Katz, Y., Fishelson, Z., and Wetsel, R. A. (1994) *J. Biol. Chem.* **269**, 28494 – 28499
517. Xu, C., Mao, D., Holers, V. M., Palanca, B., Cheng, A. M., and Molina, H. (2000) *Science* **287**, 498 – 501
518. Vidovic, D., and Matzinger, P. (1988) *Nature (London)* **336**, 222 – 225
519. Skare, J. C., Milunsky, A., Byron, K. S., and Sullivan, J. L. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2015 – 2018
520. Sayos, J., Wu, C., Morra, M., Wang, N., Zhang, X., Allen, D., van Schaik, S., Notarangelo, L., Geha, R., Roncarolo, M. G., Oettgen, H., De Vries, J. E., Aversa, G., and Terhorst, C. (1998) *Nature (London)* **395**, 462 – 469
521. Rennie, J. (1993) *Sci. Am.* **268**(Jun), 34 – 35
522. Nowak, M. A., and McMichael, A. J. (1995) *Sci. Am.* **273**(Aug), 58 – 65
523. McCune, J. M. (2001) *Nature (London)* **410**, 974 – 979
524. Magrath, I., Erikson, J., Whang-Peng, J., Sieverts, H., Armstrong, G., Benjamin, D., Triche, T., Alabaster, O., and Croce, C. M. (1983) *Science* **222**, 1094 – 1098
525. Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E., and Croce, C. M. (1985) *Science* **229**, 1390 – 1393
526. Dixon, R. A. (2001) *Nature (London)* **411**, 843 – 847
527. Currier, W. W. (1981) *Trends Biochem. Sci.* **6**, 191 – 194
528. Konno, K., Hirayama, C., Yasui, H., and Nakamura, M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9159 – 9164
529. Bergey, D. R., Howe, G. A., and Ryan, C. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12053 – 12058
530. Dangl, J. L., and Jones, J. D. G. (2001) *Nature (London)* **411**, 826 – 833
531. De Samblanx, G. W., Goderis, I. J., Thevissen, K., Raemaekers, R., Fant, F., Borremans, F., Acland, D. P., Osborn, R. W., Patel, S., and Broekaert, W. F. (1997) *J. Biol. Chem.* **272**, 1171 – 1179
532. Fant, F., Vranken, W., Broekaert, W., and Borremans, F. (1998) *J. Mol. Biol.* **279**, 257 – 270
533. McDowell, J. M., and Dangl, J. L. (2000) *Trends Biochem. Sci.* **25**, 79 – 82
534. Parker, J. E., and Coleman, M. J. (1997) *Trends Biochem. Sci.* **22**, 291 – 296
535. Lam, E., Kato, N., and Lawton, M. (2001) *Nature (London)* **411**, 848 – 853

535a. Austin, M. J., Muskett, P., Kahn, K., Feys, B. J., Jones, J. D. G., and Parker, J. E. (2002) *Science* **295**, 2077 – 2080

535b. Nishimura, M., and Somerville, S. (2002) *Science* **295**, 2032 – 2033

536. Fefeu, S., Bouaziz, S., Huet, J.-C., Pernollet, J.-C., and Guittet, E. (1997) *Protein Sci.* **6**, 2279 – 2284
537. Van Der Biezen, E. A., and Jones, J. D. G. (1998) *Trends Biochem. Sci.* **23**, 454 – 456
538. Tang, X., Frederick, R. D., Zhou, J., Halterman, D. A., Jia, Y., and Martin, G. B. (1996) *Science* **274**, 2060 – 2063
539. Chandra, S., and Low, P. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4120 – 4123
540. Legendre, L., Yueh, Y. G., Crain, R., Haddock, N., Heinstein, P. F., and Low, P. S. (1993) *J. Biol. Chem.* **268**, 24559 – 24563
541. Jabs, T., Tschöpe, M., Colling, C., Hahlbrock, K., and Scheel, D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4800 – 4805
542. Chandra, S., and Low, P. S. (1997) *J. Biol. Chem.* **272**, 28274 – 28280
543. Jennings, D. B., Ehrenshaft, M., Pharr, D. M., and Williamson, J. D. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15129 – 15133
544. Wildermuth, M. C., Dewdney, J., Wu, G., and Ausubel, F. M. (2001) *Nature (London)* **414**, 562 – 565
545. Klessig, D. F., Durner, J., Noad, R., Navarre, D. A., Wendehenne, D., Kumar, D., Zhou, J. M., Shah, J., Zhang, S., Kachroo, P., Trifa, Y., Pontier, D., Lam, E., and Silva, H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8849 – 8855

545a. Howe, G. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12317 – 12319

546. Schenk, P. M., Kazan, K., Wilson, I., Anderson, J. P., Richmond, T., Somerville, S. C., and Manners, J. M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11655 – 11660
547. Sela, I. (1981) *Trends Biochem. Sci.* **6**, 31 – 33
548. Waterhouse, P. M., Wang, M.-B., and Lough, T. (2001) *Nature (London)* **411**, 834 – 842
549. Marx, J. (2000) *Science* **288**, 1370 – 1372
550. Jorgensen, R. A. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11561 – 11563

Study Questions

1. Describe major aspects of the vertebrate innate and adaptive immune system. In what ways do they cooperate?

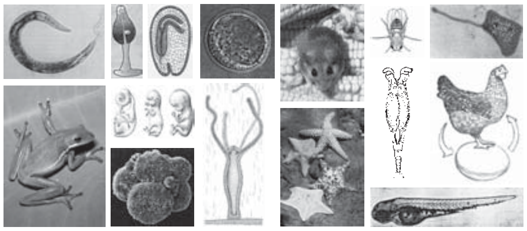
2. Describe briefly the functions of each of the folowing: antibodies, defensins, cytokines, complement, MHC proteins, B cells, T cells, dendritic cells, monocytes, macrophages, and neutrophils.

3. Why do antibodies produced using a native protein tend to bind only weakly to the corresponding denatured protein?

4. Discuss the topic of self-identity.

5. What are autoimmune diseases? How does the body avoid most autoimmune diseases?

6. List some methods by which viruses, bacteria, protozoa, and pathogenic fungi gain access to cells or to tissues.



Single cells develop into an astonishing variety of different species, all of which find their niches in the ecosystem. Whether a rectagular bacterium, a plant, a frog, or a human being, the size, shape, the body construction and metabolic pathways are established by the sequential expression of the organism's genes. Recent investigations have confirmed many similarities among major families of proteins from virtually all species. The same studies also emphasize the profound genetic differences between species. Understanding these differences, as well as the interrelationships among species, provides a continuing challenge to biochemists and biologists. Such understanding may even be essential to the survival of the human species in a changing environment.

Contents

1879 .... **A. Basic Concepts and Molecular Essentials**
1880 .... 1. DNA and Developmental Programs
1880 .... *Programmed alterations and rearrangements of DNA*
1880 .... *Inactivation of genes and imprinting*
1881 .... *Loss of DNA*
1881 .... *Amplification of DNA of chromosomes*
1882 .... 2. Receptors and Signals
1882 .... *Homeostasis*
1882 .... *Transcriptional control in differentiation*
1883 .... 3. Adhesion, Cell–Cell Recognition, and Cell Migration
1884 .... 4. Polarity, Asymmetric Cell Division, and Morphogens
1885 .... 5. Totipotency and Stem Cells
1885 .... *Cloning*
1886 .... *Stem cells*
1888 .... 6. Apoptosis or How the Tadpole Eats Its Tail
1890 .... **B. Differentiation in Prokaryotic Cells and in Simple Eukaryotes**
1890 .... 1. Bacteria
1890 .... *Bacteria with stalks*
1891 .... *Sporulation*
1891 .... *Signaling among bacteria*
1891 .... 2. Yeasts
1892 .... 3. The Cellular Slime Molds
1892 .... 4. The Hydra
1892 .... 5. Cell-Constant Animals
1893 .... **C. Development of Animal Embryos**
1893 .... 1. Germ Cells and Gametes
1893 .... *Sex determination*
1893 .... *The mammalian Y chromosome*
1894 .... *The X chromosomes*
1895 .... 2. Development of the Ovum
1895 .... 3. Fertilization
1897 .... 4. Embryonic Development
1898 .... *The anterior–posterior axis*
1899 .... *The dorsal–ventral (D-V) and right–left axes*
1899 .... *Spemann's organizer*
1899 .... *Patterns, signaling pathways, and homeotic genes*
1900 .... **D. Specialized Tissues and Organs**
1901 .... 1. Blood Cells and the Circulatory System
1901 .... *Globin genes*
1901 .... *Thalassemias*
1902 .... *Blood vessels: vasculogenesis and angiogenesis*
1902 .... 2. Cartilage, Tendons, Bone, Muscle, and Fat
1902 .... 3. Epithelia
1902 .... 4. The Nervous System
1904 .... **E. Development of Green Plants**
1906 .... **F. Aging**
1908 .... **G. Ecological Matters (Author's Personal Postscript)**

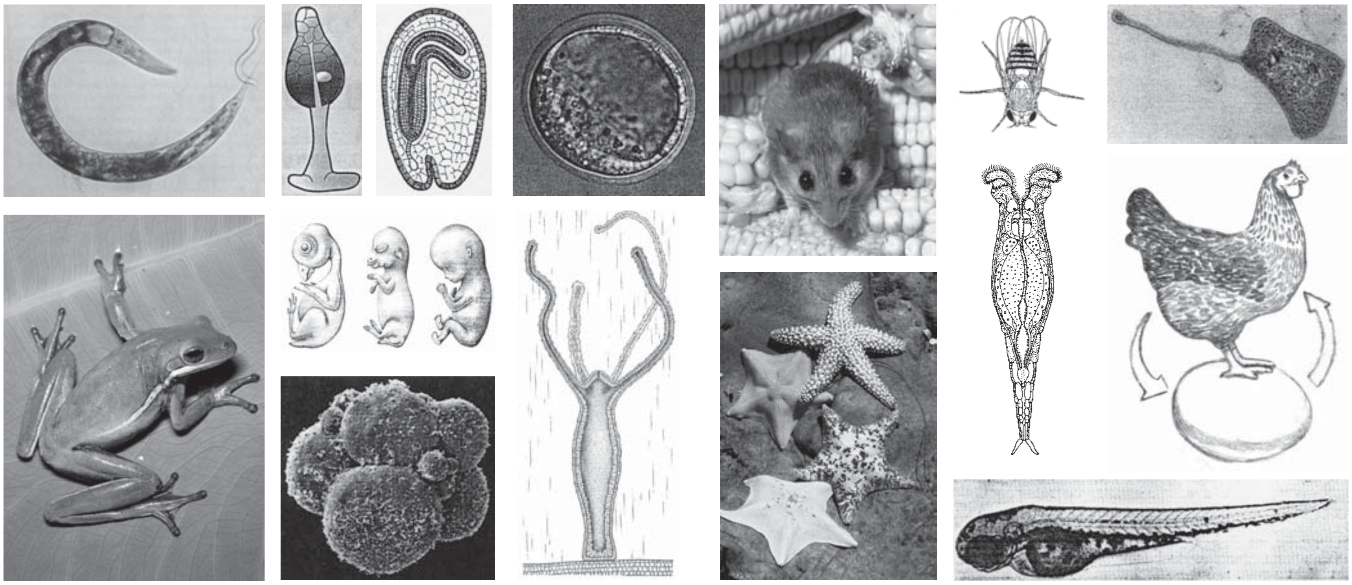
1909 .... **References**
1916 .... **Study Questions**

Tables

1880 .... Table 32-1 Some Essentials for Growth and Differentiation
1888 .... Table 32-2 Some Components of Apoptotic Systems

Growth and Development

32



One of the most fascinating of all biological phenomena is the development of an animal from a fertilized egg. From the early embryonic cells, which appear to be much alike, there arise during the course of a very few cell divisions differentiated organs and tissues such as liver, brain, kidney, muscle, skin, and red blood cells. The biochemical properties of differentiated cells are often highly specialized. Red blood cells make hemoglobin, while muscle cells make large amounts of myosin and actin. The endocrine cells of the pancreas make insulin or glucagon, while the exocrine cells form the digestive enzymes that are secreted into the intestinal tract.

Looking more broadly at the biological world we see many additional specialized features at every level of observation. Developmental patterns differ for every organism. Specialized organs abound. The structures of proteins, lipids, carbohydrates, and nucleic acids all vary, and every species has its own metabolic peculiarities. Even among bacteria we find extreme variation. Furthermore, many unicellular eukaryotic organisms undergo complex development within a single cell. The topics of this chapter are too complex for any detailed discussion. We will examine some aspects of growth and development for a few organisms and will ask whether there are simplifying generalizations. Comprehensive textbooks are available.[1,2]

Enough has been learned about development to make it clear that the DNA contains genetically coded **developmental programs**, which are followed by cells.[3] However, both transcription and translation are controlled by many chemical signals, which influence the execution of the genetic program. Such signals may arise from within a single cell, from the external environment, and from adjacent cells. The tight control is reflected by the fact that in most human tissues at any stage of development no more than about 10% of the total genes are transcribed at any one time. Chemical analysis makes it clear that most specialized cells contain a full complement of DNA but 90% of the genes are turned off.

A. Basic Concepts and Molecular Essentials

Listed in Table 32-1 are some essential aspects of growth and differentiation. Some of them are obvious. **Cohesion** between molecules provides the basis of specificity. **Receptors**, whether they be enzymes, hormone receptors, or receptors for chemotaxis, are essential. They are usually activated by a **conformational change** that accompanies binding. A cell must have receptors that can respond to a variety of **signals**, which may come from within the cell, from the external medium, or from neighboring cells. The receptor–signal pairs are essential to **local control**, which provides the basis for all of development.[4]

To have any kind of spatial differentiation a cell must develop **directionality** (polarity).[5] This permits **asymmetric cell division**[5] and development of poles in a developing cell or embryo. **Adhesion** molecules hold cells together, allowing a cell to have a fixed position relative to other cells. Development of directionality and of developmental **patterns**[6–8] are dependent upon **gradients** of concentration, of foods, heat, light, gravitation, etc., that can be detected by receptors. Gradients of compounds called **morphogens**[9,10] help to provide a **positional identity** to cells.[10–12] **Movement** of molecules of organelles and of intact cells is also essential. In multicellular organisms cells often migrate to new locations by following chemotactic signals.

Growth of individual cells enlarges them and often leads to **cell division**. The **cell cycle** describes this process with emphasis on replication of DNA. **Homeostasis** encompasses adaptation to altered nutrient or other environmental variables and to all processes which influence a cell to change. It provides for defensive reactions to many types of stress.[12a,b]

TABLE 32-1
Some Essentials for Growth and Differentiation

Cohesion of molecules, utilizing specific, hydrogen-bonding, and complementary surface shapes

Recognition, and conformational changes
- **Receptors** and **signals**
- **Local controls**

Polarity (directionality)
- Asymmetric cell division
- Poles

Gradients that can be sensed by receptors
- Food, physical qualities, morphogens

Adhesion between cells is required to hold an organism together and also, together with morphogen gradients, to provide a **positional identity**

Movement of molecules, organelles, and cells

Growth to enlarge cell size and numbers of cells
- **A cell cycle** for **replication of DNA** and **cell division**

Homeostasis to permit adaptation to changes in nutrient concentrations, stress

A developmental program, which is encoded in the genome
- Implementing this program usually requires many preformed compounds
- **Alternative developmental programs** often provide flexibility to an organism

Stem cells of **totipotent**, **pluripotent,** or multipotent nature supply new germ cells and other cells for multicellular organisms when needed

Programmed cell death (**apoptosis**) is part of many developmental programs and provides for removal of unneeded cells without inflammation

1. DNA and Developmental Programs

The genetic developmental program of an organism is encoded in the DNA. The expression of the program is implemented initially at the transcriptional level by a host of transcription factors that act at appropriate times and in appropriate places. Their action usually requires the presence of many preformed compounds. **Alternative developmental programs** may be used by unicellular organisms to adapt to new environmental conditions or by multicellular organisms to differentiate tissues.[6]

Programmed alterations and rearrangements of DNA. Part of the developmental program may involve a temporary or permanent change in the DNA. One of the simplest of these changes makes use of the **transposable recombinational switch** in which a small piece of DNA is present in either of the two possible orientations. (See Chapter 27, Section D,3.) An example, illustrated by Eqs. 27-15 and 27-16, is the variation in "phase" of the flagella produced by *Salmonella*. A somewhat different example is provided by the unicellular yeast *Saccharomyces cerevisiae*, which changes the mating type of its haploid forms in a highly regulated pattern.[13,14] The **a** mating type is expressed constitutively, but this is frequently switched to the **α** type, which produces a different mating pheromone (see Table 30-5) and responds to a pheromone from **a**-type cells. The change occurs through the transposition of different "cassettes" of DNA from "silent sites" into an expression site.[14a,b] The cassettes contain several genes, which are copied into the expression site called MAT. This site always contains genes of either **a** or **α** type. However, both **a** and **α** genes are present in other storage locations. When the mating type is switched, a copy is made of one of the stored cassettes and is placed into the MAT locus replacing the cassette already present. The MAT α genes encode two regulatory proteins, α1 and α2. Protein α1 is a positive regulator of the α-cell-specific genes, while protein α2 is a repressor of the a-cell-specific genes.[15–17] A similar mechanism appears to be employed by trypanosomes in changing their variable cell surface proteins (Box 31-E).[18]

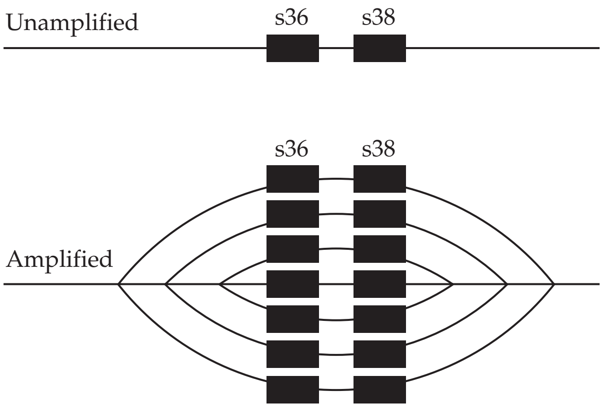
Inactivation of genes and imprinting. Under some circumstances a chromosome or part of a chromosome is permanently inactivated but remains within the cell as compactly folded heterochromatin. Heterochromatin often consists of reiterated sequences of unknown function, but it may also contain groups of inactivated genes. The most impressive case is the total inactivation of one of the two X chromosomes in cells of female mammals.[19–21] The entire chromosome appears as heterochromatin. The inactivation occurs early in embryonic development and is random with respect to the two X chromosomes. In some cells the maternal X chromosome, in others the paternal X chromosome, becomes inactive. However, upon further cell divisions the same chromosome in each clone remains inactive. As a consequence of the random inactivation, the female body is a mosaic with respect to genes in the X chromosomes. Formation of heterochromatin in X chromosomes is described in Section C,1. It depends upon **epigenetic markers**, which consist, at least in part, of methylated cytosines in 5'-CpG-3' sequences. These CG pairs are palindromic and can be methylated on both strands (see Eq. 27-2 and Fig. 32-3).

Selective inactivation of genes by methylation accounts for **genomic imprinting**, which occurs in mammals and which marks a gene of either maternal or paternal origin for silencing.[22–24] About 50 imprinted genes have been identified in mice and in humans.[23,25] Imprinting, known as **epigenetic inheritance**, also appears to arise in part by methylation of cytosines in CpG sequences. Four m^{5}C-methyltransferase genes have been identified in the human genome.[26] One of these presumably provides the initial methylation of the cytosine in one strand of the DNA. A second methyltransferase (apparently encoded by gene *Dnmt1*)[27] methylates the second strand and serves as the **maintenance methylase** that preserves the methylation pattern upon DNA replication (see Eq. 27-2). The faithful maintenance of imprinting has been demonstrated experimentally for genes of mice formed by nuclear transplantation, i.e., mice grown from somatic cell nuclei transplanted into ova.[28] Imprinted genes usually occur in clusters, which also contain imprinted genes for noncoding RNA molecules. The transcription of those genes often is correlated with repression (**silencing**) of protein-coding genes on the same chromosome. This observation is consistent with evidence that genes may be silenced by the binding of small RNA molecules to complementary sequences in the DNA or of mRNAs. The mechanism may be limited to **imprint control regions** or silencer units in the DNA, while the genes to be controlled are set off by **insulator elements** (see also pp. 1882, 1894 and Chapter 28, Section C,1). The significance of genomic imprinting is not clear.[29,30] However, it is very important to nuclear transplantation because the methylation state of imprinted genes is normally reset (**epigenetic reprogramming**) before embryonic development begins.[23,25,27,31–33] See Section 5. Methylation of DNA is not essential to the life of filamentous fungi. Nevertheless, *Neurospora* does have a DNA methyltransferase, which methylates about 1.5% of the cytosines present.

Loss of DNA. While some genes are selectively inactivated, others may be irreversibly lost during development. The extreme case is that of the human red blood cell from which the entire nucleus is expelled. Loss of DNA may result from recombinational events. For example, crossing-over between sister chromatids during mitosis has been demonstrated for some cells and chromosomes. This does not alter the genetics of the progeny cells if equal amounts of genetic material are exchanged. However, if two or more similar base sequences occur in tandem in a DNA molecule, unequal crossing-over (Chapter 27, Section D,2) can occur with the loss of genetic material from a chromosome of one of the progeny cells. This may be a deliberately programmed route of differentiation for some cells.

Alternatively, loss of DNA from a chromosome may occur through a **looping-out excision** mechanism. Like the excision of a prophage from the chromosome of *E. coli* (Chapter 27, Section D,3) this loss of genes occurs at specific sites in the DNA. The best known case is the joining of gene segments during the differentiation of lymphocytes (Fig. 31-18). The extent to which similar changes occur during terminal differentiation of other tissues or in nonmammalian species is uncertain.

Amplification of DNA of chromosomes. During formation of oocytes parts of the DNA are “amplified” by repeated replication. This provides a way for the ovum to accumulate ribosomal RNA and various proteins in large amounts. Similarly, genes for two abundant proteins of the egg shell or **chorion** of insects are amplified. Bidirectional replication initiated at discrete positions yields an “onion skin” structure containing many copies of an ~90-kb sequence containing the two genes. The polyploidy observed in some highly specialized cells such as the Purkinje cells



of the cerebellum and of many cells of Diptera larva (Chapter 26, Section F,3) represent another way of amplifying genes. Polyploid cells of animals generally represent a terminal stage of differentiation and do not divide. They tend to contain their full complement of genes in each copy of their DNA, but most genes are not expressed.

2. Receptors and Signals

Receptors and their ligands are numerous, varied, and essential to all forms of life. Cell-surface receptors on bacteria detect feeding attractants as well as dangerous molecules. From bacteria to humans seven-helix receptors function to detect light, odors, hormones, and other molecules. The numbers of different receptors are impressive. For example, the tiny nematode *C. elegans* has 650 seven-helix transmembrane receptors and 411 protein kinases, many of which may be associated with receptors.[34] Our bodies have thousands.

Every ligand that binds to a receptor is a signal of some kind. How many signaling molecules are there, and what are their structures? The number of proteins and small peptides affecting growth and development of cells may be enormous. Many of these have been considered in Chapter 30. In addition, the catecholamines, serotonin, histamine, and even bicarbonate ions may act as local hormones.[35] For example, HCO_3^- determines whether a thin-walled sporangium or a thick-walled heat-resistant sporangium will appear in the phycomycete water mold *Blastocladiella*. We know that a very large number of different proteins and small peptides are secreted from cells, many in very small amounts.[36] Some protein morphogens act at concentrations so low, e.g., 50 pM, that it has been hard to detect them.[10] T-cell receptors respond to extremely small amounts of short peptides that stream outward from many cells bound to surface MHC molecules. Do some of these same peptides act as signals between other cells during development?

Small RNA molecules are now being found to function in many ways within cells,[37–39] They may also act as messengers within cells and between cells. The small 22 nt RNAs transcribed from genes *lin-4* and *let-7* of the nematode *C. elegans* control key developmental decisions.[38b,d] In green plants small RNAs move between cells and throughout the plants to trigger defensive responses (p. 1869).[38c] Much of the RNA transcribed from genes of any eukaryotic cell lacks any known function. Since evolution tends to act on all of the molecular constituents of a cell, this vast amount of RNA may have acquired vital roles in the control of metabolism and growth.[38e] A poorly understood intracellular structure, the **ribonucleoprotein vault** also remains a biochemical mystery. Differentiated cells may contain 10^4 and embryonic cells over 10^7 of these ~42 x 75 nm hollow objects. Vaults have internal cavities large enough to hold two intact ribosomes. Some vaults appear to be empty, but others contain materials, a fact that suggests some role in transport or storage.[38f]

Homeostasis. A large fraction of the receptors and signaling system of cells is devoted to the maintenance of a constant internal environment. This homeostasis is essential if a cell is to respond to changes in external conditions without disastrous consequences.[12a,b] Some special aspects of these processes are discussed in Chapters 11, 17, and 28. Within mammalian cells the hypothalamus, pituitary, and adrenals have a primary responsibility for maintaining homeostasis.[40]

Transcriptional control in differentiation. Development of an organism depends upon an orderly sequence of transcription of genes. Some genes are transcribed in germ cells, others within cells of an early embryo, and others later. As the embryo develops sequential waves of synthesis of the needed protein are observed.[3,41–43] These are controlled by the actions of numerous transcription factors that act at a variety of **cis-regulatory modules** (CRMs) associated with promoter regions[42,44–46] and also by controls on translation of mRNAs,[47] and by negative feedback loops.[48] The latter may involve hormones, morphogenic proteins, coregulatory proteins of various types, small RNAs, etc.[49] More than 2000 transcription factors are encoded in the human genome. Most are positive-acting, i.e., they *promote* transcription. Several families are specifically involved in regulation of development (see Table 28-2). These include the Zn^{2+}-containing GATA-1 (p. 1634), which regulates globin synthesis, embryonic factors of *Drosophila* (e.g., bicoid), vertebrate homeotic genes (Hox clusters), Pit1, the muscle-specific helix–loop–helix proteins MyoD and Myf5, and several forkhead (winged helix) proteins.[3] A single CRM can bind many different regulatory molecules and single regulators can bind to a variety of CRMs. During development of the simple embryo of the sea urchin *Strongylocentratus purpuratus* a network of 40 highly regulated genes is needed to coordinate growth and development with the production of the proteins needed at each step.[41,43]

Part of the control of differentiation lies in the interaction of proteins that regulate transcription with metabolites and hormones. For example, substrate depletion not only decreases growth rate of bacteria but also alters gene transcription. This occurs in *E. coli* as a result of a rise in internal cAMP concentration. The presence of an alternative energy source such as lactose induces changes in gene transcription (Chapter 28, Section A,1). Such **physiological modulation** of a developmental pattern can also be seen in higher organisms.

More striking is the fact that environmental signals can trigger a cell to switch to an alternative developmental program by which enough new genes are activated to rebuild the cell into a new form. An example is spore formation, a process that occurs with some bacteria when external conditions become unfavorable for vegetative growth (Section B,1). Alternative developmental programs are also evident in eukaryotic organisms that undergo metamorphosis, and they may be important to development. Perhaps persistent states of repression of groups of genes can be passed through several generations of cells until a specific chemical signal triggers the unwrapping of the appropriate nucleosomes and transcription of formerly inactive genes.

3. Adhesion, Cell–Cell Recognition, and Cell Migration

Development of multicellular organisms depends upon both adhesion and on recognition of a correct interaction. Like enzyme–substrate, receptor–hormone, and antibody–antigen binding these interactions of macromolecules on cell surfaces often show a high degree of specificity. They may be accompanied by conformational changes and may trigger signaling cascades. We have already discussed some of these interactions, for example, the binding of a molecule of IgG attached to a surface antigen to protein C1q of the complement system (Fig. 31-8) and the binding of an MHC–antigen complex to a T-cell receptor (Fig. 31-15).

There are many other cell–surface **adhesins**, several of which have been discussed on pp. 402–409. Among them are proteins that contain immunoglobulin-like domains and numerous **glycoproteins**. An example of the latter is the binding of a type of pili found in pathogenic strains of *E. coli* to epithelial cells of the urinary tract. The pilin subunits (Fig. 7-9), like lectins (Box 4-C), bind specifically to the disaccharide group Gal*p*α1→4Gal. A lectinlike protein specific for *N*-acetylglucosamine rings is involved in invasion of erythrocytes by the malaria parasite *Plasmodium*. The unicellular alga *Chlamydomonas* (Fig. 1-11) produces sexual gametes of two mating types. When mixed together, gametes of opposite mating types, prior to fusion, adhere to each other via **agglutinins** present on their flagella. The agglutinins are glycoproteins rich in hydroxyproline, serine, glycine, arabinose, and galactose.[50] As mentioned on p. 29, colored cells of different strains of the marine sponge *Microciona prolifera* find others of the same strain using highly specific proteoglycan-like aggregation factors.[51–53] These compounds are highly polymorphic, and it has been suggested that they are part of a primitive immune system. The aggregation reaction requires calcium ions. In our own bodies Ca^{2+}-dependent lectins, the **selectins** (p. 187, 188), bind leukocytes and help to guide them to their sites of action.

Other adhesins include the **integrins**,[53a] cellular adhesion molecules (**CAMs**), **cadherins**,[53a–c] and **fibronectin** (Fig. 8-19). These are also discussed on pp. 402–409. The CAMs (Fig. 8-18A),[54,55] which are members of the immunoglobulin-like protein family, are glycoproteins bearing large 2,8-linked sialic acid polymers.[56–58] They promote Ca^{2+}-dependent aggregation. However, the effect of NCAM, which is widely distributed in a developmentally regulated fashion, can be antiadhesive if long chains of sialic acid are present. NCAM appears to play a role in remodeling and repair of tissues. Adhesion of molecules within cell membranes and the binding of substances to membrane surfaces provides another driving force in development. Within membranes molecules spontaneously sort themselves into lipid **microdomains**, often called **lipid rafts**.[58a] Related to lipid rafts are caveolae (p. 426). These little craters arise in cholesterol-rich microdomains. They often contain the protein **caveolin** as well as glycosphingolipids and GPI-tailed proteins (Fig. 8-13).[58b] ATP-dependent linking reactions may also occur to provide more permanent bonding. Membrane-associated molecules, in turn, become centers for attachment of cytoskeletal proteins and other protein complexes. As with the cytosol and extracellular fluids homeostatic mechanisms act to provide a relatively constant membrane–lipid environment.[58c]

Several types of cell junctions are associated with adhesion and participate in intercellular communication (Fig. 1-15).[59] The cadherins are transmembrane proteins with large extracellular domains (Fig. 8-18B). They are prominent components of adherens junctions[59–61a] in which they join the exterior surfaces of pairs of cells in a zipper-like manner. Another protein, **β-catenin**, links the short C-terminal tails of cadherin through a-catenin subunits to the actin cytoskeleton.[60] In desmosomal junctions other specialized proteins including desmoglein have functions similar to that of cadherins.[59] Tight junctions, from zebrafish to humans, depend upon a complex of several proteins including those of the claudin family.[62] Significantly, the cohesive powers of some adhesins, e.g., of cadherin, are altered during development. Cadherin E is nonadhesive in a four-cell mouse embryo but becomes adhesive after the eight-cell stage.[59] It is obvious that many other changes in intercellular adhesion must also occur during growth and development.

The integrins (see also p. 405) comprise a large family of adhesive receptors that are found in animals from sponges to humans.[63–65] They have both adhesive and signaling functions. Both subunits of their αβ heterodimeric structures[64] have single transmembrane helices and short C-terminal cytoplasmic tails. The β1 subunit tails interact with cytoplasmic proteins. The distribution of integrins varies among cell types. Human leukocytes contain alpha subunits of types αd, αl, αm, and αx with molecular masses of 150–180 kDa. Two ~95-kDa beta subunits (β1 and β2) are present. However, T lymphocytes express β1, β2, and β7 integrins. Other patterns are observed for other leukocytes,[66] in skin,[67] and in other tissues.[68] Integrin molecules tend to aggregate into clusters, which are found together with other proteins, at the ends of actin stress fibers (p. 370).[63] The largest of these clusters are known as focal adhesions. Signals may be sent through integrins in either of the two directions.[63] The extracellular domains of integrins interact with a variety of proteins of the extracellular matrix. These include fibronectin, fibrinogen, vitronectin, collagen, and entactin.[63,69] Other large cell surface adhesins include laminin and osteopontin (Chapter 8), thrombospondin, von Willebrand factor, and related proteins.[70] These adhesins appear to depend upon the sequence Arg-Gly-Asp (RGD), which binds noncovalently to integrins, which act as cell-surface receptors.[71,71a] See also Chapter 12, Section C,9.

The functioning of the complex network of integrins, adhesins, and other components of the extracellular matrix is not understood in detail. One fundamental question is how the strength of adhesion can vary with time and stage of development. Roseman postulated an association of an oligosaccharide chain of a glycoprotein attached to one cell with a specific glycosyltransferase of another cell.[72] The specific interaction would hold cells together, but addition of another glycosyl unit to the oligosaccharide by the transferase would alter the surface properties of the cell carrying the glycoprotein. This, in turn, could cause disaggregation of the cells. Glycosyltransferases can be found on the outer surfaces of cell membranes, and Roseman's proposal may correctly describe one aspect of cell adhesion.

Other molecules that are abundant on cell surfaces include heparan sulfate proteoglycans. Although they have often been regarded as providing a nonspecific "extracellular fly paper," recent evidence from studies of development in *Drosophila* suggest specific and important functions in signaling and in developmental patterning.[73] Both hyaluronan and chitin also have been proposed to play an important role in vertebrate development.[74,75] Proteoglycans of plant cell surfaces, as well as the hydroxyproline-rich proteins of cell walls, may function in plant development.[76]

Movement of cells from one location to another is essential to embryonic development as well as to wound repair and to the immune response. Many brain structures are composed exclusively of immigrant cells.[77] These **cell migrations** depend upon the cytoskeletal actin filaments, integrins, and focal adhesions.[78,79] Chemotactic signals are also required.[80] A great complexity of underlying chemistry is being elucidated.[79,81] See also Chapter 19, Section C.

4. Polarity, Asymmetric Cell Division, and Morphogens

Cells of *E. coli* usually divide exactly in the center to form two seemingly identical cells.[82–84a] However, under the right conditions some bacteria, e.g., *Caulobacter crescentus* and *Bacillus subtilis*, undergo asymmetric division to form two different types of cells (Fig. 32-1).[83] There is clearly an axial polarity. This polarity is evident even in *E. coli*, which has flagella streaming out at one end and its chemoreceptor-bearing "nose" at the other (Fig. 32-1).[85] Axial polarity is also obvious in other bacteria (Fig. 19-1).

Polarity is evident in eukaryotic cells from protists to higher organisms.[85a,b] Cells of the yeast *S. cerevisiae*, whether haploid or diploid, divide in an asymmetric way by budding.[5,7] Among body cells of higher animals those of the epithelium are among the most polarized (Fig. 32-2; see also Fig. 1-6, Box 8-F). Polarity is always present in ova of eukaryotes, but the ova may initially be radially symmetric.[86–88] The **anterior–posterior** axis, which is formed first, establishes a head-to-tail direction.[89] In bacteria this major axis is determined as perpendicular to the division plane. In the tiny worm *C. elegans* the anterior–posterior axis of an ovum is determined by the position of entrance of the sperm. This marks the posterior end.[90] In higher animals the axis, which is also known as the **animal–vegetal axis**, is established by uneven distribution of materials that include mRNAs and proteins in the unfertilized ovum. During embryonic development of bilaterial species two other axes, the **dorsal–ventral** and **right–left** axes, are also developed and help to establish the body plan. Thoughout development polarized movements allow cells to intercalate between one another to help shape the body.[90a]

Early studies of simple organisms such as Hydra (Fig. 1-13) and planaria (flatworms, Fig. 1-14A) showed that distinct chemical differences can be detected along the anterior–posterior axis. These organisms can be cut into pieces, many of which can regenerate complete individuals.[91,92] Regions near the head regenerate most readily. These and other observations led to the concept of gradients of diffusible **morphogens**, or form-giving molecules.[9,10] In *Drosophila* eggs an mRNA specified by the gene *bicoid* is localized at the anterior pole. The translation product, the bicoid protein, diffuses through the embryo, which in *Drosophila* lacks cell walls at this stage (see also Section C,4).[9,93–95] Bicoid is a transcription factor and also one of a number of established morphogens. Many other morphogens are members of the **TGF-β**

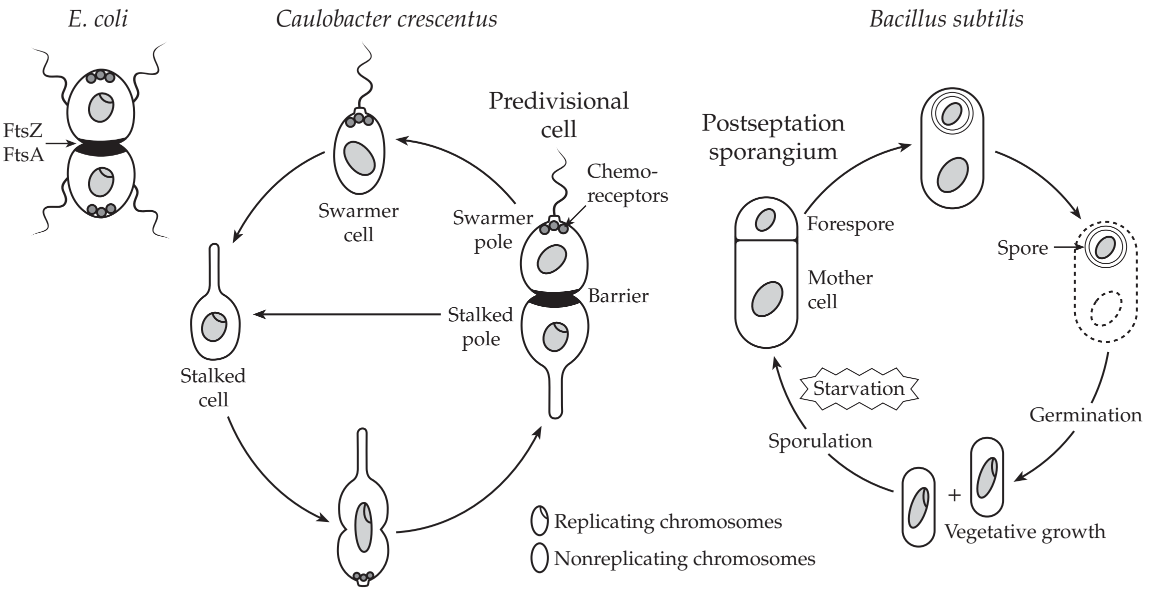

E. coli
Caulobacter crescentus
Bacillus subtilis
FtsZ
FtsA
Predivisional cell
Chemo-receptors
Swarmer cell
Swarmer pole
Barrier
Stalked pole
Stalked cell
Replicating chromosomes
Nonreplicating chromosomes
Postseptation sporangium
Forespore
Mother cell
Spore
Starvation
Germination
Sporulation
+
Vegetative growth

Figure 32-1 Comparison of cell division in three species of bacteria. *Escherchia coli* divides symmetrically after forming a septum in a plane marked by a ring of FtsZ (tubulin-like) and other cell division proteins. *Caulobacter crescentus* divides asymmetrically to give one flagellated swarmer cell and one stalked cell. *Bacillus subtilis*, under starvation conditions, divide to form a mother cell and a forespore. The latter is engulfed by the mother cell, which promotes its conversion to a resistant spore. From Shapiro and Losick.[83] Courtesy of L. Shapiro.

(transforming growth factor beta) family. Among them are proteins that establish the dorsal–ventral axis (Section C,4)[86,96–98] and also bone morphogenic proteins (p. 443). Retinoids also appear to act as morphogens.

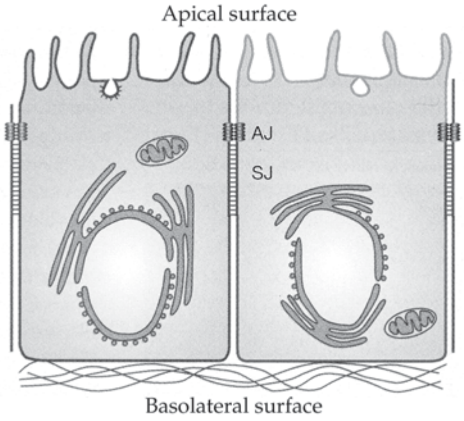

Apical surface
AJ
SJ
Basolateral surface

Figure 32-2 A pair of epithelial cells of *Drosophila*. The apical surface (top), e.g., of epithelial cells of the gut, faces the external surface, while the basolateral surface (bottom) binds to a basal membrane. Adheren junctions (AJ) and septate junctions (SJ) are shown between the cells. From Peifer and Tepass.[84] Drawing by S. Whitfield.

5. Totipotency and Stem Cells

The cambium layer of plant stems (Fig. 1-16) differentiates continuously to form phloem on the outside of the cambium and xylem on the inside. At the same time, cambium cells are retained. Thus, at each cell division one daughter cell becomes a differentiated cell, while another remains the less differentiated cambium. This pattern of continuous differentiation from a line of **stem cells** with constant properties is found in animals as well as in plants. In the differentiation of cambium it appears that chemical signals obtained from the surrounding cells on either the inside or the outside of the cambium layer determine whether the differentiated cell becomes phloem or xylem. Sucrose, auxin, and cytokinins are all involved.

Cloning. Asexual propagation (cloning) of plants ordinarily occurs by virtue of the ability of embryonic meristematic tissue to differentiate into roots and shoots. If isolated phloem cells or other more differentiated cells are cultured, the result is often the formation of a **callus**, a dedifferentiated mass of cells somewhat reminiscent of embryonic cells. Under proper conditions, e.g., in a coconut milk culture and in the presence of the correct auxin-to-cytokinin ratio, some carrot root phloem cells revert to embryonic cells and develop into intact plants.[99] This experiment provided proof that the differentiated carrot phloem cells contained a complete genome for the plant. Its nucleus is **totipotent**, able to generate all cell types. However, the experiment cannot be done easily with most plants, and dedifferentiation is not always automatic. It does occurs often enough to establish the totipotency of the nucleus of many differentiated cells.

Similar considerations apply to animal cells. In the earliest stages differentiation is readily reversible. Later it becomes difficult to convert a differentiated animal cell into one resembling an embryonic cell. However, Gurdon demonstrated that this is sometimes possible. Nuclei from cells of intestinal epithelia and other tissues were substituted, by transplantation, for the nuclei of egg cells. The process is called **nuclear transplantation**[100–101a] or nuclear transfer cloning.[102] The resulting eggs in some cases grew into adult toads. Thus, the full genetic information of the toad was present in the differentiated cells.[103,104] However, it was not possible to accomplish this result with nuclei of neurons, which may have undergone irreversible differentiation. More recently mammalian nuclei have been utilized in the same way to create the famous sheep Dolly as well as mice, calves, pigs, and kittens.[105–108] These animals are commonly said to have been **cloned**, a term that has long been used to denote asexual propagation, e.g., in a colony of dividing cells or in propagation of plants by grafting.

In nuclear transplantation it is the DNA that is hoped to be the same in every individual in a clone. However, the ovum used for the transplantation contains mitochondria. Some mitochondria may also accompany the nucleus during the transfer. If the donors of the ovum and of the nucleus are different individuals the offspring will be mitochondrial hybrids.[105] In addition, there are questions about the methylation state of DNA in the donated nucleus and about the age and health of the donated mitochondria. That these questions are significant is emphasized by a bit of 3000-year-old knowledge from mule breeders: a mare crossed with a donkey yields a mule but a stallion crossed with a donkey yields a hinny, which has shorter ears, a thicker mane and tail, and stronger legs than does a mule.[109] There are worries because Dolly and many other animals produced by nuclear transplantation have not been completely healthy.[107b,110] Is something missing from the transplanted DNA or does it carry something extra, such as methyl groups? Recently it has been recognized that incorrect epigenetic marking of cytosine in CpG pairs that control maternally imprinted genes, especially those on chromosomes 11 or 15, may cause death of embryos or devastating human diseases.[25,111] An important related question for those wishing to clone an animal by nuclear transplantation is "Should the cell that donates the nucleus be in the G_1 state of the cell cycle (Fig. 11-15) or the G_0 or paused state that precedes G_1?"[105] See also Chapter 27, Section B,6.

Stem cells. For many years it has been appreciated that, as shown in Fig. 31-2, both erythrocytes and other blood cells arise throughout life from self-renewing stem cells in the bone marrow.[13,112] Stem cells are also needed for renewal of bones, muscle, skin, neurons, etc. Stem cells appear to be present only in small numbers and in well-protected special **niches** in the body.[53c,113–115] They are able to live throughout an individual's lifetime, dividing quite rarely and always producing one or more highly differentiated cells as well as a new stem cell.[113] A fertilized egg (zygote) is totipotent, able to generate all the cells of an animal including those of the placenta and other tissues that are not part of the embryo. However, the most capable stem cells are **pluripotent**, able to form more than one type of specialized cell.[113,116] Mammalian pluripotent stem cells include tumor cells, **embryonic stem cells**, derived from preimplantation embryos, and **embryonic germ cells**, derived from the primordial germ cells of the postimplantation embryo.[117] These germ cells are not only totipotent but, with good luck, may be immortal.[118]

Recent results indicate that adult-derived somatic cell nuclei may still retain full pluripotency.[107a,117a] Some confusion has arisen because of the discovery that stem cells may sometimes fuse with differentiated cells.[118a,b] It is only recently that it has been possible to locate and to cultivate human stem cells. These cells, which may be recovered from both embryonic and mature tissues, include the blood-cell-forming **hematopoietic** stem cells, fetal **neuronal** stem cells, **melanocyte** stem cells, and **mesenchymal** stem cells (or marrow stromal cells). The last give rise to muscle, bone, cartilage, and tendons.[119–121a] Most stem cells may arise late in development and function principally in tissue renewal.[122] Among the most abundant are those of epthelial tissues, whose cells provide 60% of differentiated tissue types in the mammalian body.[123,124] Epidermal stem cells must provide for regular replacement of the outer skin surface (Box 8-F) but must also provide cells for rapid repair of wounds.[125] The exact locations of epidermal skin cells have been difficult to find. The cells appear to be well-protected in areas deep in the skin. Some are located in hair follicles.[120a,125] Stem cells of plants are present in specialized structures called **meristems**. A seedling typically has two meristems, at the tips of the shoot and root, respectively.[126] See Fig. 32-8B.

Cloning of human stem cells is of great medical interest because of the possibility of replacing defective cells or tissues. Tissue engineering may supply urgently needed differentiated cells for replacement purposes[107,127–128a] and may eventually lead to replacement organs.[129–131] These efforts must be pursued with caution, but most researchers see a bright future for cloning of tissue cells.[128,132–134] At the same time there is nearly universal agreement that nuclear transplantation cloning of human beings should not be attempted.[132,135] One key objection is the near certainty that many seriously defective human embryos would be created.

How can we explain a pattern in which one daughter cell resulting from division of a stem cell undergoes differentiation while the other remains a stem cell? A hypothetical way in which this might happen is illustrated in Fig. 32-3. It depends upon methylation of CpG sequences in DNA (Eq. 27-2). A methyltransferase, E_1 in Fig. 32-3, would modify a site in the DNA that controls differentiation by methylating a base in one of the two strands of a palindromic sequence such as a CpG pair (Eq. 27-2). The maintenance methylase E_2 would further methylate the corresponding cytosine in the palindrome in the second strand. During tissue growth replication would produce a hemimethylated CpG in each daughter cell. These would also be methylated by E_2 (upper part of Fig. 32-3) to complete the replication process. In contrast, a stem cell would need to mark one DNA strand, e.g., by action of a third methyltransferase (E_3) or some other type of DNA-modifying enzyme. As depicted in Fig. 32-3, E_3 would add a methyl group to DNA on some location outside of the palindromic region. Replication in this case would yield one hemimethylated palindome, which would allow one daughter cell to follow the normal growth and replication cycle. However, the other daughter cell would carry the mark designating it as a stem cell. The presence of both E_1 and E_2 in cells would lead to the continuous differentiation of the modified cells from unmodified ones, the situation found in stem line cells. A "maintenance methylase" with the properties of E_2 has been identified (Section A,1).

One question that has been asked is whether the 200 cell types of the vertebrate body all arise as a result of chemical interactions between cells and hormones and other external signals? Alternatively, does a **developmental clock** count the number of cell divisions and at the appropriate time turn off one set of genes and turn on another?[136] Methylases as well as other enzymes might modify DNA at specific times during development. For example, a hydrolase might deaminate the adenine in an AT pair to inosine. Upon replication and cell division one daughter cell would receive an unaltered DNA molecule, but the other would contain in place of the AT base pair an IC pair. Following a second replication, a GC pair would be formed (Eq. 27-20) resulting in AT to GC mutation at a specific site in the DNA of some of the daughter cells. Such a simple change, occurring in response to an enzyme formed at a certain stage of development, could alter the expression of genes in a cell. Schemes involving palindromic sites and modification enzymes that could turn off specific genes after a given number of cell replications have been suggested.[136] In fact, there seems to be little support for such mechanisms.

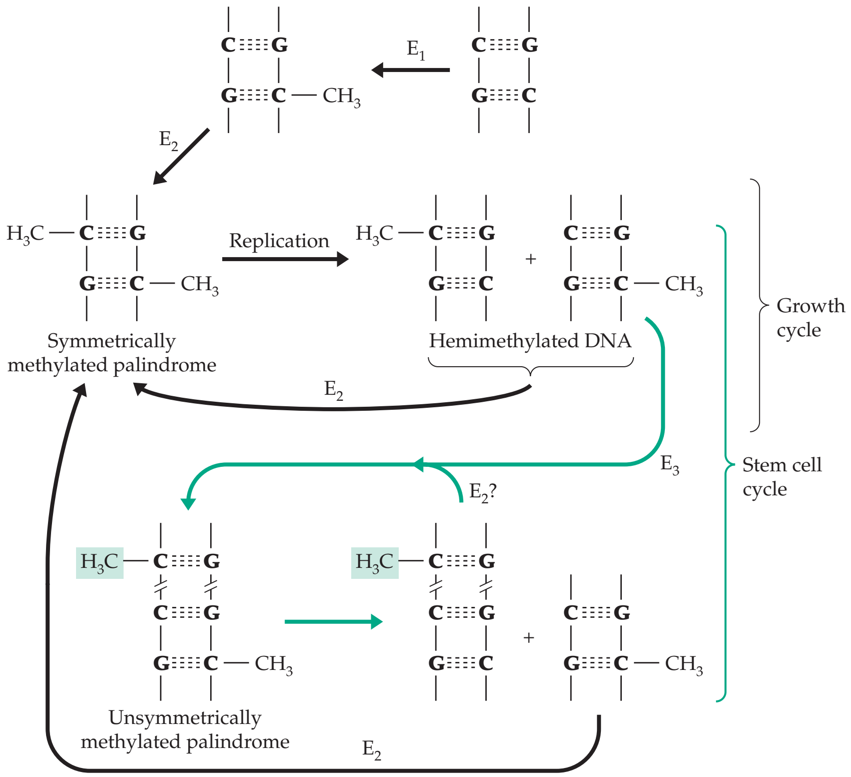


Figure 32-3 Hypothetical way of controlling stem cell replication by methylation or other marking system. Methyltransferases E_1 and E_2 methylate the cytosine in a 5'-CpG-3' or other palindromic sequence. In freely replicating cells these two enzymes keep the CG sequences methylated on both DNA strands. In stem cells another enzyme, perhaps a third methylase (E_3), marks a location outside the palindromic DNA on one strand (🟩). Replication leaves the mark in the duplex, which is retained in a stem cell. The other strand will yield a hemimethylated duplex allowing the cell to follow the normal growth and replication pattern. Based on proposals of Holliday and Pugh.[136]

Transcriptional controls (Section 2) and the sensing of a cell's location may provide adequate control.

If methylation and possibly other covalent modifications of DNA occur, how can one explain the totipotency observed for some nuclei of differentiated cells? During development of the ovum and the sperm there appears to be a "resetting" of the developmental clocks that led to differentiation. At this time all of the methyl groups on the CpG pairs of imprinted genes are removed (Section C,2).[25,137] The mechanism is uncertain. Perhaps marking of newly replicated DNA stops. For example, in the case of the methylated DNA of Fig. 32-3 if E_1 and E_2 were absent in the cytoplasm of the ovum, no further methylation would occur during subsequent cell divisions. However, there may be active enzyme-catalyzed demethylation (see p.1541). Other mechanisms of gene silencing are also known (pp. 1881,1894).[138]

6. Apoptosis or How the Tadpole Eats Its Tail

Observers have long been fascinated by the rapid resorption of a tadpole's tail as it turns into a frog or toad. The process, designated by the Greek word apoptosis (whose "pop" is pronounced),[139] or as **programmed cell death**, plays a major role in many aspects of development in nearly all organisms.[140–142] For example, during human development about one-half of all the neurons generated die.[143,144] Unneeded lymphocytes, some of which produce antibodies or T-cell receptors directed against a person's self, are also killed. Cells may die accidentally from injuries. In many cases the resulting death occurs by **necrosis** rather than apoptosis. Necrotic death is accompanied by swelling and bursting of the cell and a subsequent inflammatory response.[145,146] In contrast, cells dying by apoptosis shrink, break into fragments, and are rapidly eaten by surrounding cells.[146a,b] There is no inflammation. Because of this it has been difficult to determine the extent to which apoptosis contributes to normal development. Apoptosis is also distinguished from **autophagy**, which is intracellular turnover under starvation conditions. Cells may need to scavenge unneeded proteins and organelles, recycling them within the cell.[147]

Cells damaged by disease, e.g., dopaminergic neurons in Parkinson disease, may die by apoptosis.[147a,b] A second form of self-destruction occurs when an axon is cut.[147b] Failure of the elaborate network of mechanism for repair of DNA and maintenance of the genome normally leads to apoptosis. In cancer essential steps in the apoptosis pathway are often inactivated.[147c]

Our view of apoptosis (outlined in Fig. 32-4) changed with the tracing of the origins and fates of all of the ~1000 cells of the nematode *C. elegans*.[13] During development of the adult worm just 131 specific cells die by apoptosis. Studies of mutant worms revealed mutations in several cell-death (*ced*) genes. Three proteins, encoded by genes *ced-3*, *ced-4*, and *egl-1*, are essential for apoptosis.[140,148] Somewhat surprisingly worms with defective Ced proteins are apparently healthy, even though they have 131 extra cells. On the other hand, in the fruit fly *Drosophila* mutations in similar death genes are sometimes fatal.[140]

The nucleotide sequence of the *ced-3* gene revealed that the Ced3 protein is closely related to the **interleukin-converting enzyme ICE**,[149–152] which is discussed on p. 619. ICE is a member of the **caspase family** of thiol proteases (p. 619). At least 14 different caspases are found in the human body. Some of them

TABLE 32-2
Some Components of Apoptotic Systems

Apaf1	Mammalian homolog of Ced4; component of apoptosome
Apoptosome	Cytosolic complex: Apaf1•caspase-9•cytochrome *c*
Bcl2	Mammalian homolog of Ced9 protein of *C. elegans;* inhibitor of apoptosis
Bcl-2 family	Group of regulators of apoptosis, both inhibitory and stimulatory (Bad, Bax, Bik, etc.)
CARD	
CD95 (AP-1, Fas)	One of the most studied death receptors
Ced3	*C. elegans* thiol protease, related to mammalian caspase9
Ced4	Activator of Ced3, related to mammalian Apaf1
Ced9	Inhibitor of Ced3 and related caspases
DD	Death domain of a death receptor
DED	Death effector domain
DISC	Death-inducing signaling complex, formed in plasma membrane
FADD	Fas-associated death domain, an adapter protein
ICE	Interleukin-converting enzyme, structurally related to Ced 3
TNF	Tumor necrosis factor (a family of cytokines secreted by macrophages)
TNFR	Receptors for a TNF family member

function in apoptosis[152a] and others in maturation of pro-inflammatory cytokines.[152–155] Most exist as proenzymes, which must be activated by proteolysis.[156] The mammalian homology of Ced3 is caspase 9.[153,157] The Ced4 protein of *C. elegans* is an activator for Ced3. Its mammalian counterpart is called **Apaf1** (apoptotic protease-activating factor 1).[153,158] Protein **Ced9** is an inhibitor of apoptosis, which probably protects the worm from erroneous deaths.[159] Its mammalian equivalents are proteins of the *Bcl-2* gene family.[142,152a]

It is well established that caspases participate in the final stages of apoptosis (Fig. 32-4), but what initiates the process? There appear to be many ways in which apoptosis can be triggered. If every cell has a proper location in the body, which is determined by signals from adjacent cells, what will happen if the cell becomes detached? There is evidence that such detachment with the loss of survival signals causes apoptosis.[53a,152b] Cell damage is also a major trigger. In other cases the cell is "instructed" to die. An example is the death of unneeded lymphocytes, one of many cellular processes induced by cytokines of the tumor necrosis factor (TNF) family. To allow for this process cells have surface receptors of the TNF receptor (TNFR) superfamily.[159a] Some TNFRs are **death receptors,** which are called by many names.[142,160–163] One of the best known is CD95[164a] (also called Fas[164] or Apo1). CD95 is involved in death of mature T lymphocytes at the end of an immune response and also in the killing of virus-infected cells and cancer cells by cytotoxic T cells or NK cells.

Members of the TNF family that activate CD95 (CD95 ligands or CD95Ls) are trimers. They bind to the cysteine-rich external domains of the transmembrane CD95 molecules inducing them to aggregate (Fig. 32-4). The cytosolic portion of each of these death receptors contain a **death domain** (DD). The bundle of aggregated receptors also bind to an adapter protein such as the Fas-associated death domain (FADD).[156] It is one of many proteins involved in apoptosis whose structures are known.[157,163] The FADD molecule contains a **death effector domain** (DED), which associates with a similar domain in the proenzyme procaspase 8. A rather large membrane-associated molecular complex, the **death-inducing signaling complex** (DISC; Fig. 32-4), is assembled in this way.[161,165] Oligomerization of the procaspase domain causes activation via self-cleavage to give active

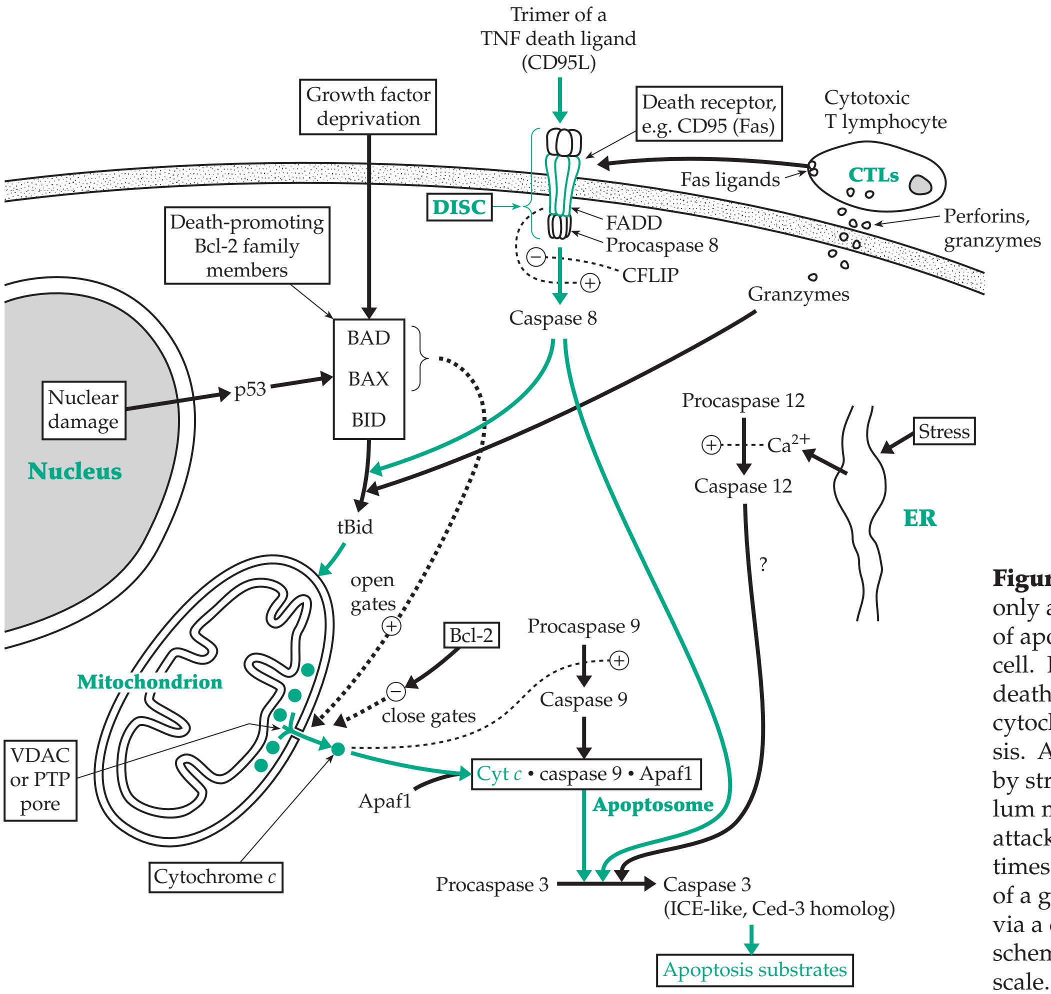

Trimer of a TNF death ligand (CD95L)
Growth factor deprivation
Death receptor, e.g. CD95 (Fas)
Cytotoxic T lymphocyte
CTLs
Fas ligands
DISC
FADD
Procaspase 8
CFLIP
Perforins, granzymes
Granzymes
Death-promoting Bcl-2 family members
Caspase 8
BAD
BAX
BID
Nuclear damage
p53
Nucleus
Procaspase 12
Stress
Ca^{2+}
Caspase 12
ER
?
tBid
open gates
Bcl-2
Procaspase 9
Mitochondrion
close gates
Caspase 9
VDAC or PTP pore
Cyt *c* • caspase 9 • Apaf1
Apaf1
Apoptosome
Cytochrome *c*
Procaspase 3
Caspase 3 (ICE-like, Ced-3 homolog)
Apoptosis substrates

Figure 32-4 Sketch illustrating only a few of the many aspects of apoptosis in a mammalian cell. Emphasis here is on the death receptor pathways and cytochrome *c*-activated apoptosis. A third pathway is initiated by stress in endoplasmic reticulum membranes. In addition, attack by cytolytic T cells sometimes causes apoptosis by action of a granzyme on protein Bid or via a death receptor. Objects in scheme are not drawn to a single scale.

caspase 8, which initiates the apoptotic response. Other "upstream" caspases (caspases 2, 9, and 10) also participate in initiation of apoptosis. In contrast, caspases 3, 6, and 7, the "executioner caspases," are thought to participate directly in demolition of the cell.[154] A caspase-activated DNase also participates by degrading DNA.[166,167]

A second major pathway of activation of apoptosis depends upon mitochondria. Various stresses such as lack of needed growth factors, exposure to ultraviolet light, or other apoptosis-inducing signals apparently open pores or gates in mitochondrial membranes allowing materials that would promote apoptosis to flow out into the cytoplasm and stimulate the effector caspases.[167a] This possibility was supported by the discovery that cytochrome *c* stimulates apoptosis.[168] Cytochrome *c* is a small protein, which is present in inner mitochondrial membranes in a 1:1 ratio with other electron carriers. It is more mobile and less tightly bound than the other components. It carries electrons from complex III to cytochrome *c* oxidase within the intermembrane space (Fig. 18-5). As a result of apoptotic stimuli cytochrome *c* rapidly flows out of the intermembrane space into the cytoplasm both interfering with respiration and triggering other changes in the cell.[169–172] The outflow of cytochrome *c* may occur via the mitochondrial porin VDAC (p. 1047) or under some circumstances via the mitochondrial permeability transition pore (PTP; p. 1049).[173] Within the cytosol the escaped cytochrome *c*, together with caspase 9 and Apaf1, forms a large multimeric complex (cyt*c* • caspase 9 • Apaf1) called an **apoptosome**. The apoptosome catalyzes activation of caspase 3, initiating the caspase cascade.

Control of the gates or pores by which cytochrome *c* escapes from mitochondria is poorly understood.[171,171a,173] Whereas in *C. elegans* a single protein Ced9 has been identified as an inhibitor of apoptosis, vertebrate animals have a large family of proteins that are related to the Ced9 homolog Bcl-2.[157,174] Of these Bcl-2 and Bcl-x$_L$ *inhibit* the flow of cytochrome *c* out of mitochondria, but several other members of the family, e.g., Bad, Bid, Bik, and Bax, *promote* apoptosis.[152a,175,176] Bad carries a signal that indicates a lack of growth factor stimulation. Bid carries a death message from QD95R and other death receptors. Bax carries a signal from p53 (Fig. 11-15) indicating unacceptable DNA damage. However, a truncated form of Bax may prevent apoptosis of neurons.[177] In every one of these pathways there are many complexities. In one of the best known pathways Bid is cleaved by caspase 8 to form a 15-kDa fragment t-Bid that becomes an integral membrane protein in the outer membrane of mitochondria. There it promotes the release of cytochrome *c* (Fig. 32-4).[178]

A quite different source of apoptotic signals are ER membranes, which respond to stress by releasing Ca^{2+} ions that activate caspase 12 (Fig. 32-4).[179] Yet another type of apoptosis is sometimes induced by granzyme B (p. 610), which is released from cytolytic T cells.[180–182]

B. Differentiation in Prokaryotic Cells and in Simple Eukaryotes

Every species undergoes developmental changes. Only a few of these will be considered here briefly.

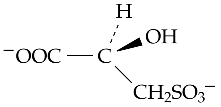
1. Bacteria

Although they are usually regarded as unicellular, some bacteria develop more than one type of cell,[83] and some even form colonies with filamentous growth[183,184] or other distinct morphology.[185] Many bacteria alter their development in response to changes in environment. For example, unfavorable conditions lead bacteria such as *Bacillus subtilis* to form compact endospores inside the vegetative cells.[83,184] Many other bacteria including *E. coli* divide symmetrically. This fact also poses a question. How does a cell locate its center and divide? The answer is only partially understood. In all kinds of bacteria a protein known as **FtsZ** (filamentation temperature-sensitive protein Z), a GTP-binding protein homologous to eukaryotic tubulins (Fig. 7-33), is essential. Prior to division FtsZ accumulates as a **septal ring** at the center of the *E. coli* cell. Contraction of the ring is thought to be an essential step in cell division.[186–188] The FtsZ ring nucleates a growing complex of eight additional proteins known as FtsA, T, K, L, N, Q, W, and ZipA. While ZipA is not highly conserved among bacteria, in *E. coli* it is the first protein to add to the FtsZ ring.[187,189] ZipA is somewhat related to eukaryotic actin. Another group of proteins is also needed for location of the midcell plane. These are known as MinC, MinD, and MinE. A MinC•MinD complex inhibits potential binding sites for FtsZ. MinD is an ATPase, which structurally resembles the Fe protein of nitrogenase (Fig. 24-2) and appears to propel the MinC•MinD complex in an oscillatory fashion from pole-to-pole.[190–193] This behavior is not understood, but in some manner the 10-kDa MinE is able to overcome the inhibition and bind to FtsZ initiating division. Division in *E. coli* follows DNA replication by a constant time period (20 min at 37°C). The timing apparently depends upon diadenosine 5'-tetraphosphate (Ap$_4$A), which acts as a signal to couple division to replication.[194]

Bacteria with stalks. *Caulobacter crescentis* spreads to new areas while retaining a presence at home. As is illustrated in Fig. 32-1, asymmetric division produces two distinctly different cells. One, like the maternal cell, has a stalk, while the other has a rotary flagellum with which it travels. In time the motile "swarmer" cell sheds its flagellum and undergoes metamorphoses into a stalked cell. What controls this process? There is apparently a two-component system similar to the one that controls flagellar movement (Fig. 19-5). Two sensor histidine kinases phosphorylate a central response regulator **CtrA**.[195,195a,b] This represses initiation of replication in the swarmer cells as well as transcription of the cell division gene *ftsZ*. As mentioned on p.1094 changes in DNA methylation may also occur. Flagellar biosynthesis and other steps of differentiation then occur in the swarmer cells but not in the stalked cells in which normal growth and replication take place.[196–201] Action of proteases is also essential.[202]

Sporulation. Bacteria of the genera *Bacillus* and *Clostridium* form metabolically inert spores when deprived of adequate nutrients (Fig. 32-1).[83,203,204] Bacterial spores are remarkably resistant to heat and can survive boiling water for prolonged periods. Their metabolic rate is essentially zero, but they can be revived and can grow even after many years. For example, bacteria have been grown from a 118-year old can of meat. Some data suggest that spores can survive for no more than ~1000 years, but recent reports, not yet fully verified, say that spores have survived when embedded in salt crystals for 250 million years.[203,205] At the onset of sporulation the synthesis of ribosomal RNA is turned off completely, and new classes of mRNA are made. More than 50 genetic loci are affected by mutations that cause spore formation. As was mentioned in Chapter 28, one or more specific forms of the 70-kDa σ subunit of RNA polymerase (Chapter 28, Section A,2) are produced and direct the initiation of the new mRNA molecules encoding new proteins.[204,206–209] Prior to asymmetric cell division the first of the new σ factors, σF, is formed together with two regulatory proteins, Spo0A and Spo0B. Spo0B, a protein kinase, phosphorylates Spo0A, inactivating it.[210] SpoAB also forms an inactive complex with σF. After asymmetric cell division σF remains inactive in the mother cell but is released in the prespore by action of Spo0E. This is a membrane-bound phosphatase, which dephosphorylates Spo0A-*P*, allowing it to form a complex with Spo0B with release of σF. Another protein, SpoIIIE, appears to direct one copy of the replicated DNA into the forespore.[211] The σF factor then directs the transcription of genes in the forespore. In contrast, σE is produced only in the mother cell.

One of the most striking metabolic changes in metabolism during sporulation is the accumulation of large amounts of dipicolinic acid (Fig. 24-14). This requires the appearance of at least one new enzyme. In addition, as the spores develop the bacteria take up large amounts of Ca^{2+} and substantial concentrations of Mn^{2+} and other metal ions. In many bacteria 3-L-sulfolactic acid is also formed.



3-L-Sulfolactate

These components account for the following percentages of the total dry weight of spores of *B. subtilis*: dipicolinic acid, 10%; sulfolactic acid, 3–6%; Ca^{2+}, 3%; and Mn^{2+}, 0.3%. It is often suggested that the dipicolinic acid and other ions protect the proteins from denaturation. However, the heat resistance may arise from the maintenance of the core of the spore in a highly dehydrated state.[212] When conditions become appropriate for growth again, the spore germinates, and the bacterium again follows the cell growth and division program.

More complex alternative developmental programs are followed by colonial forms of bacteria such as the myxobacteria. The life cycle involves aggregation of cells and formation of fruiting bodies as well as sporulation.[185]

Signaling among bacteria. Even bacteria respond to signals from other bacteria. Individuals of a single species often react by secreting pheromones called **autoinducers** using a process called **quorum sensing**. Among the responses are swarming of cells, emission of light by luminous bacteria, synthesis of antibiotics, and formation of biofilms. As mention on p.1758, autoinducers used by gram-negative bacteria are often *N*-acetylhomoserine lactones.[213–215] A furanosyl borate diester (see Box 11-F) may be a more nearly universal autoinducer.[216] Programmed cell death can also be observed among bacterial populations.[217]

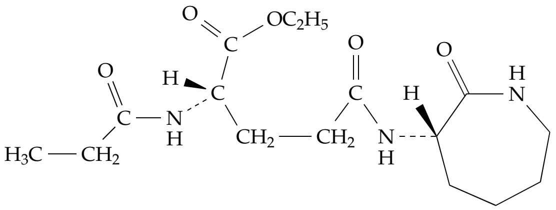
2. Yeasts

The budding yeast *Saccharomyces cerevisiae* and the fission yeast *Schizosaccharomyces pombe* are the best known fungi. Although they usually grow as individual cells, they can grow with a filamentous form under some conditions.[218] Other yeast, notably *Candida albicans*, are important pathogens and can also grow in either yeast or pseudohyphal filamentous forms.[219,220] Like *E. coli*, *S. pombe* undergoes symmetric cell division. However, the strong β(1→3) linked glycans with their β1→6 crosslinkages, mannose polysaccharides, and chitin provide a cell wall very different from those of bacteria.[221] Cells of *S. pombe* grow mainly at their tips and begin early in mitosis to form a ring of actomyosin and other proteins at the center. This corresponds to the mammalian contractile ring. At the end of anaphase the ring contracts, and the septum that separates the two cells develops.[222–223a] In both types of yeast the septum is rich in chitin, which is secreted from the cell membrane or from vesicles known as chitosomes. After the septum is fully formed and thickened, a chitinase partially hydrolyzes the chitin releasing the cells. In *S. cerevisiae* the cell division is asymmetric.[5,221] The position of the bud seems to be directed by the actin cytoskeleton.

All fungi form spores (gametes) during their haploid stage, which follows meiosis (Fig. 1-10). The transcriptional program for *S. cerevisiae* involves at least four sets of genes, which are transcribed consecutively. During spore formation the mRNA levels of more than 1000 of the ~6200 protein-encoding genes are changed. About 50% are elevated and ~50% are depressed.[224] The mating type changes in the haploid state have been mentioned on p. 1574. Similar mating and sporulation pathways are observed for *S. pombe*.[225]

3. The Cellular Slime Molds

The life cycle of *Dictyostelium discoideum* is described briefly in Box 11-C. About 10^5 individual amebas aggregate to form a moving "slug" in response to the chemoattractant cAMP. Some other species of *Dictyostelium* are attracted to a folic acid derivative or to the ethyl ester of *N*-propionyl-γ-L-glutamyl-L-ornithine-δ-lactam.[226]



N-Propionyl-γ-L-glutamyl-L-ornithine-δ-lactam

In all cases the cells also utilize cAMP as an internal second messenger. For *D. discoideum* the components of the chemotactic-aggregation system include a 41-kDa cAMP receptor on the outside, adenylate cyclase, an extracellular diesterase that specifically hydrolyzes the cAMP to AMP, and a diesterase inhibitor protein.[35,227–230] The inhibitor keeps the phosphodiesterase largely inactive initially, but when cAMP concentrations build up synthesis of the inhibitor is repressed and the cAMP is hydrolyzed, a necessary condition for retaining sensitivity of the receptors for the arriving pulses of cAMP.

The slug of aggregated amebas continues to move and to undergo differentiation into two cell types: about 80% of the cells become pre-spores and the remaining 20%, which are at the "head" of the slug, become pre-stalk cells. The front-to-back gradient of cAMP within the moving aggregate seems to be involved in differentiation. However, another "differentiation-inducing factor" as well as NH_3 may be involved in the formation of stalk.[230] As the aggregate forms, the cells become cohesive, an 80-kDa surface glycoprotein being involved.[231] Later other adhesive **discoidins**, 24- to 27-kDa RGD-containing galactose-binding lectins,[232] also participate in holding the colony together. Some cells begin to produce cellulose. Trehalose is also formed and is stored in the spores. New enzymes have to be made to synthesize these materials. An alternative developmental pattern for some strains of *Dictyostelium* is formation of macrocysts between cells of two different mating types. A diffusible 12-kDa inducing factor appears to be released by cells of one strain.[233]

4. The Hydra

A well-fed hydra (Fig. 1-13) appears immortal. Its body cells are sloughed off and replaced at a steady rate so that within a month or so its body has been completely renewed.[35] The hydra contains only ten cell types. These include two kinds of stem cells that give rise to the ectodermal and endodermal cells of the body wall as well as small **interstitial stem cells** (Fig. 1-13) that differentiate nerve cells, germ cells, and the nematocytes or stinging cells. Of the ~10^5 cells in a hydra about 3600 are interstitial stem cells. Each day they generate 400 nerve cells and 1800 nematocyte precursor cells as well as 3500 new interstitial cells. The nematocyte precursors move up the body of the hydra and take up residence in the tentacles. Their movement is thought to be guided by chemotaxis. The head activator peptide (Table 30-5), which was identified following isolation from 3 x 10^6 hydras (3 kg),[234] diffuses from the foot end of the animal forming a gradient. A foot activator may diffuse from the opposite end. The interstitial stem cells of hydra also give rise to clones that develop into the gametes. Female hydra always develop female gametes, but stem cells of male hydra give rise to both male and female gametes.[235] This sex switching is reminiscent of the mating type variation of yeast.

5. Cell-Constant Animals

While the hydra is almost immortal as a result of the continuous differentiation of its stem cell lines, other small invertebrates follow a very different course of development. Both the rotifers and the annelid worms (Fig. 1-14) tend to have a constant number of cells in the adult body. The entire developmental program is specified genetically in strict detail.

The one millimeter long adult nematode *Caenorhabditis elegans* contains only 959 somatic cells. The lineal descent of all of these has been traced.[236–240] The development follows an almost exactly defined pathway with 113 programmed cell deaths during formation of the 558-cell newly hatched larva. In addition, each adult worm contains 302 neurons that make about 8000 synapses. This little nematode also has an alternative developmental pathway. The larvae shed their cuticles in four consecutive molts. If the food supply is inadequate, they enter a persistent nonfeeding state in which they may survive for months and are able to resume development when conditions are appropriate.[241]

C. Development of Animal Embryos

The shapes and body plans of animals vary enormously. Consequently, the study of embryonic development of sea urchins, insects, frogs, chickens, mice, and humans might appear to lead to quite unrelated conclusions. However, there are many similarities as well as variety.

1. Germ Cells and Gametes

Throughout the animal kingdom from protozoa to human beings sexual reproduction predominates. It is true that there are about 1000 species that reproduce asexually.[242,243] Among them are ~350 species of all-female rotifers[242] and even a species of tiny mites, all of which are haploid females.[244] Nevertheless, sex seems to have conferred some advantage on most species. There are two thories that may explain this: (1) Sex brings different combinations of genes together, allowing especially favorable combinations to survive, when changing conditions make life difficult.[242] (2) Sex helps to remove deleterious mutations from a population.[243] A large fraction of human fetuses (at least 10–25%) contain an "incorrect" number of chromosomes and as many as 20% of oocytes are defective. In contrast only 3–4% of sperm are chromosomally abnormal. Female meiosis I appears to be highly error-prone.[243a] Abnormal fertilized eggs or embryos are eliminated later in development.

Sex determination. The sex of an individual is determined by the chromosomes. In humans and other mammals presence or absence of a Y chromosome determines the sex. However, in many organisms including *C. elegans* and *Drosophila* this is not true. Although *Drosophila* males like human males are XY, it is the ratio of the number of X chromosomes to the number of sets of autosome (A) that determines the sex. This is also true for *C. elegans*, which has no Y chromosomes.[244,245] Apparently because of the differing ratios of X:A in the two sexes, organisms utilize a variety of **dosage compensation** methods. In cells of human females only one X chromosome is active. In *Drosophila* the rate of expression of genes from the X chromosome is roughly doubled in males.[245–248] In *C. elegans* the expression from both X chromosomes of the (hermaphroditic) female is roughly halved.[245,246,249,250] The biochemistry underlying these processes is quite complex.

The mammalian Y chromosome. The basic plan of the gonads prior to differentiation is female. However, if a Y chromosome is present (or if genes from a Y chromosome have been translocated to other locations) testes develop and begin to secrete androgen as early as the 60th day of gestation. A male-specific DNA sequence, ***SRY*** (sex determining region Y), constituting the gene for the **testes-determining factor**, is located in the small arm of the Y chromosome (Fig. 32-5).[251–253] A small pseudoautosomal segment at the end of the short arm of the Y chromosome carries other genes and undergoes crossing-over during meiosis.[254,255] The *SRY* gene lies between this and the centromere. The SRY protein is a member of the HMGA subgroup of **HMG** DNA-binding proteins (p. 1535).[256,257] It binds tightly to the sequence AACAA(A / T)(G /C) broadening the minor groove of the B-DNA and bending the DNA by more than 70°.[252,258,259]

Both SRY and the related SOX proteins are critical developmental regulators.[260] In early fetal life the mammalian embryo contains an indifferent gonad, able to differentiate into either a testis or ovary. Adjacent to the gonad are two simple ducts, the Müllerian (female) and Wolffian (male).[261] In the male SRY acts in the developing gonads to induce differentiation into the Sertoli cells of the testis. In the mouse the *Sry* gene is active for only a brief period about ten days after fertilization. During that period cells of the genital ridge start to differentiate. In the absence of protein SRY they develop into the female follicle (granulosa) cells but in the presence of SRY into Sertoli cells.[252,262] This is, in part, a result of production of the **Müllerian inhibitory substance** (MIS), which induces regression of the Müllerian duct, and later production of testosterone. MIS is a glycoprotein of the TGF-β family. Binding of SRY to a site in the Mis gene promoter appears to be involved in activation of the Mis gene.[252] Recent evidence points to a role for both SRY and SOX proteins in pre-mRNA splicing.[260] At least 25 other genes are also involved in spermatogenesis in the mouse.[263] Many of these testis-specific genes have completely unmethylated CpG sequences.[264,265] For example, a cAMP-responsive element present in a promoter sequence for a testis-specific subunit of pyruvate dehydrogenase must be demethylated for

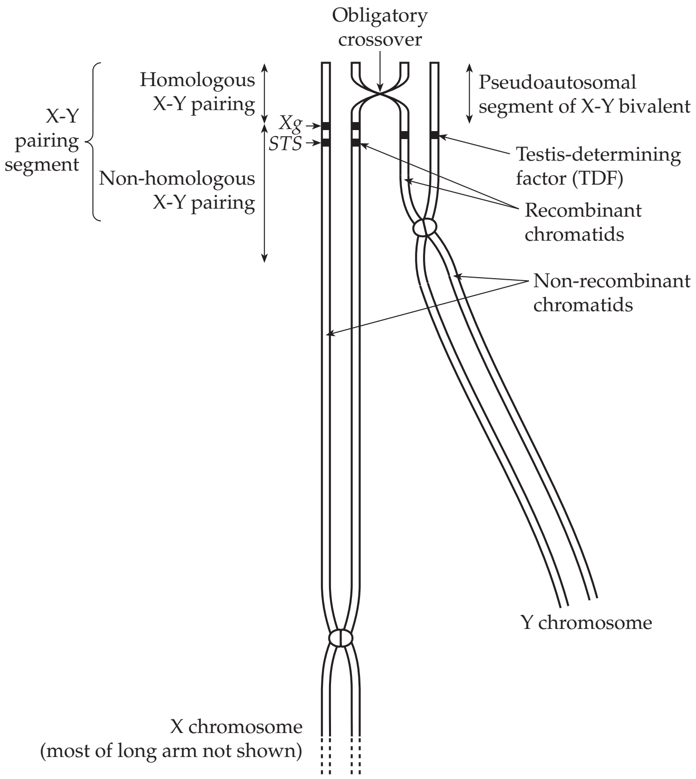

Obligatory crossover
X-Y pairing segment
Homologous X-Y pairing
Non-homologous X-Y pairing
Xg
STS
Pseudoautosomal segment of X-Y bivalent
Testis-determining factor (TDF)
Recombinant chromatids
Non-recombinant chromatids
Y chromosome
X chromosome (most of long arm not shown)

Figure 32-5 Schematic diagram showing crossing-over between the human X and Y chromosome. The pseudoautosomal segment is that part of the X-Y bivalent where there can be X-Y exchange by crossing over. X-Y homology in this segment is maintained by, and may be necessary for, this crossing over. There is always one "obligatory" X-Y crossover, whose position varies. The length of the X-Y pairing segment varies with meiotic stage and can extend well beyond the pseudoautosomal segment into the Y long arm. Much (perhaps all) of the synaptonemal complex formed outside the pseudoautosomal segment represents non-homologous pairing. From Burgoyne.[254]

transcription to occur. The developing germ cells interact with the surrounding Sertoli cells at every stage both through direct cell–cell contacts and via secreted signals.[266] Interstitial cells of the testis differentiate into Leydig cells, which secrete testosterone, promoting development of the Wolffian duct.[261] A small population of germ-line stem cells provide for continuing spermatogenesis. In *Drosophila* their self-renewal depends upon signals from special hub cells.[115,267] Other proteins needed for normal male development include the X-linked androgen receptor, whose absence causes testicular feminization, and dihydrotestosterone reductase (Chapter 22).

The development of spermatozoa is unlike that of somatic cells. Extensive reorganization of chromatin occurs under the direction of cis-regulatory elements that are controlled by cells of the testis.[267a] Among specialized proteins that are synthesized is a testis-specific polyadenylate polymerase.[267b] During the remodeling histones are replaced by arginine- and cysteine-rich protamines.[267a–c] In mammals this occurs in two stages. Small intermediate proteins (TP1 and TP2) replace histones in the first stage and are displaced by protamines in the second.[267a,d] Sulfolipids, which are also present in myelin, are essential to spermatogenesis.[267e] Both the sphingolipid 3-sulfogalactosylceramide and **seminolipid**, a sulfate ester of monogalastosylalkylacylglycerol (structure on p. 387) are present in large amounts. However, their functions are not clear.

Selenium plays a special role in development and protection of spermatozoa (Chapter 15). The selenoprotein **phospholipid hydroperoxide glutathione peroxidase** (PHGPx; Eq. 15-58, Table 15-4) has a high activity in the testis and in spermatids. However, in mature spermatozoa it forms an enzymatically inactive oxidatively crosslinked capsular material around the midpiece of the cell perhaps providing mechanical stability.[268] A similar 34-kDa selenoprotein is present in sperm nuclei and may be essential for condensation of DNA.[269] Sperm tails contain specialized cytoskeletal proteins which form "outer dense fibers."[270] In contrast to mammalian spermatozoa, nematode sperm move by ameboid motility that depends upon a specialized actin-like molecule.[271] Sperm cells are unusually rich in polyamines, most of which are bound to RNA and DNA (Chapter 24).

The X chromosomes. The phenomenon of X chromosome inactivation in mammalian female cells is closely related to imprinting, which has been discussed in Section 1. The inactivation process is quite complex. It involves methylation of 5'-CpG-3' sequences of DNA, as is described in Chapter 27, Section B,6. It also depends upon an **inactivation center**, the *Xist* gene, which is expressed only from the inactivated X_i chromosome, whose *Xist* DNA is unmethylated. On the X_a chromosome this DNA is methylated, and the gene is silent.[272] The *Xist* transcript is a long RNA that may bind to and coat much of the X_i

DNA.[19,21,273–275] The associated chromatin is enriched in a variant of histone H2A and is underacetylated on the tails of histones H2A, H3, and H4 (Fig. 27-4). Also noteworthy is the fact that not all genes on X_i are inactivated. As many as 19% escape this control.[276] Another gene *Tsix*, which is adjacent to *Xist* in the DNA, is also involved. *Tsix* encodes an RNA that is *antisense* to the *Xist* transcript[21,275,277] but is transcribed from the active X chromosome X_a. One hypothesis is that the *Xist* transcript causes X inactivation and that the *Tsix* transcript acts in an opposite way to favor activation of the chromosome. A transcription factor known as **CTCF** has been identified as a possible regulator of the inactivation process.[21,277] This is a trans-acting factor that is encoded on a chromosome other than X or Y. The process also depends upon methylation of histone H3.[278] Methylation of H3 may also be a factor in gene silencing in other organisms.[279] CTCF also regulates a number of other genes, e.g., those of the globin gene cluster (Fig. 27-10). It binds to 60-bp sequences, perhaps in enhancer elements.

2. Development of the Ovum

In the early mammalian female embryo the absence of the Müllerian inhibitory substance MIS permits continuing development of the Müllerian duct, while the absence of testosterone permits the Wolffian duct to degenerate. However, positive developmental signals are also required. Among these is the protein **Wnt-4**, a member of a large family of locally acting signal molecules (Section 4). Wnt-4 may be needed both for oocyte development and for further suppression of male development.[261]

The earliest studies of oocyte development were done with sea urchins (often *S. purpuratus*) and with amphibians (often the South African clawed toad *Xenopus laevis*) whose eggs are as much as 1000 times larger than those of mammals.[280] However, despite the differences in size, modes of fertilization, and ovary development, oocytes of nematodes, sea urchins, frogs, insects, and mammals have much in common. Oocytes of *C. elegans* and of most other animals undergo a temporary arrest in development at the prophase stage of the first meiotic division (Fig. 26-12).[13,281–283a] At this stage oocytes transcribe many genes. In some species chromosomes may develop a “lampbrush” appearance (Fig. 27-6) as a result of the transcriptional activity. Many mRNA molecules are stored in the expanding cytoplasm. Proteins are also synthesized and stored.[284] Among these are specialized proteins of yolk granules and proteins used to construct an outer egg coat. Surrounding **follicle cells** also contribute nutrients to the oocyte.[284a] In insects, whose early embryonic development has some special characteristics, 16 surrounding **nurse cells** are connected to the oocyte by cytoplasmic bridges.[285,286]

Oocytes may remain in arrest at the beginning of meiosis for prolonged periods before continuing through the **maturation** stage to form an ovum (egg). Women and other female mammals are born with thousands of oocytes, but only a few at a time develop into eggs. Maturation is often delayed until sexual maturity, when it is stimulated by hormones.[13] In *X. laevis* progesterone stimulates maturation.[282,287] In *C. elegans* and many other animals a signal from a sperm cell is needed to induce maturation.[281,283] Maturation of the oocyte is often arrested again, this time at metaphase of the second meiotic division. Transcription is halted, and protein synthesis is slowed. Fertilization then induces rapid completion of meiosis. Penetration of the sperm leads to “activation” of the egg and completion of meiosis. In lower organisms activation can often be carried out by chemical or physical treatment in the absence of a sperm cell, with formation of parthenogenetic offspring.

3. Fertilization

Fertilization of the egg is a biochemically complex process.[35,288] It involves recognition of sperm and egg, often in a species-specific manner.[289] The jelly layer around sea urchin eggs contains peptides such as the **sperm-activating peptide** (speract; Table 30-4), which stimulates increased respiration and motility of the sperm cells.[290] Additional chemotactic peptides may also be released from the jelly layer of invertebrate eggs. Chemoattractants for vertebrate eggs are less well known, but a 21-kDa sperm attractant protein from *X. laevis* egg jelly has been characterized.[291] Both in sea urchins and in mammals the jelly layer, which is called the **zona pellucida**, contains sperm cell receptors.[288,292–295] These are glycoproteins that interact with proteins (spermadhesins)[296] of the sperm cell membrane. One of these is the integrin-associated CD9, an integral membrane protein.[297,298] Penetration of the sperm through the zona pellucida often involves a large specialized secretory vesicle, the **acrosome,** as well as the enzyme hyaluronidase.[298a] In some species the acrosome releases a large amount of monomeric G actin, which polymerizes suddenly into a tube of polymeric F actin, which in some way assists the penetration of sperm.[35] In the horseshoe crab *Limulus polyphemus* the acrosome in an unactivated sperm cell contains a twisted bundle of as many as 120 cross-linked actin filaments. When the sperm is activated by contact with the jelly coat of the egg, the acrosome straightens into a 50-µm-long crystalline bundle, which is driven into the egg coat.[299] Of importance to all types of sperm cells are proteases and other materials that are also released from the acrosome and which help to etch a hole that allows the sperm to enter the

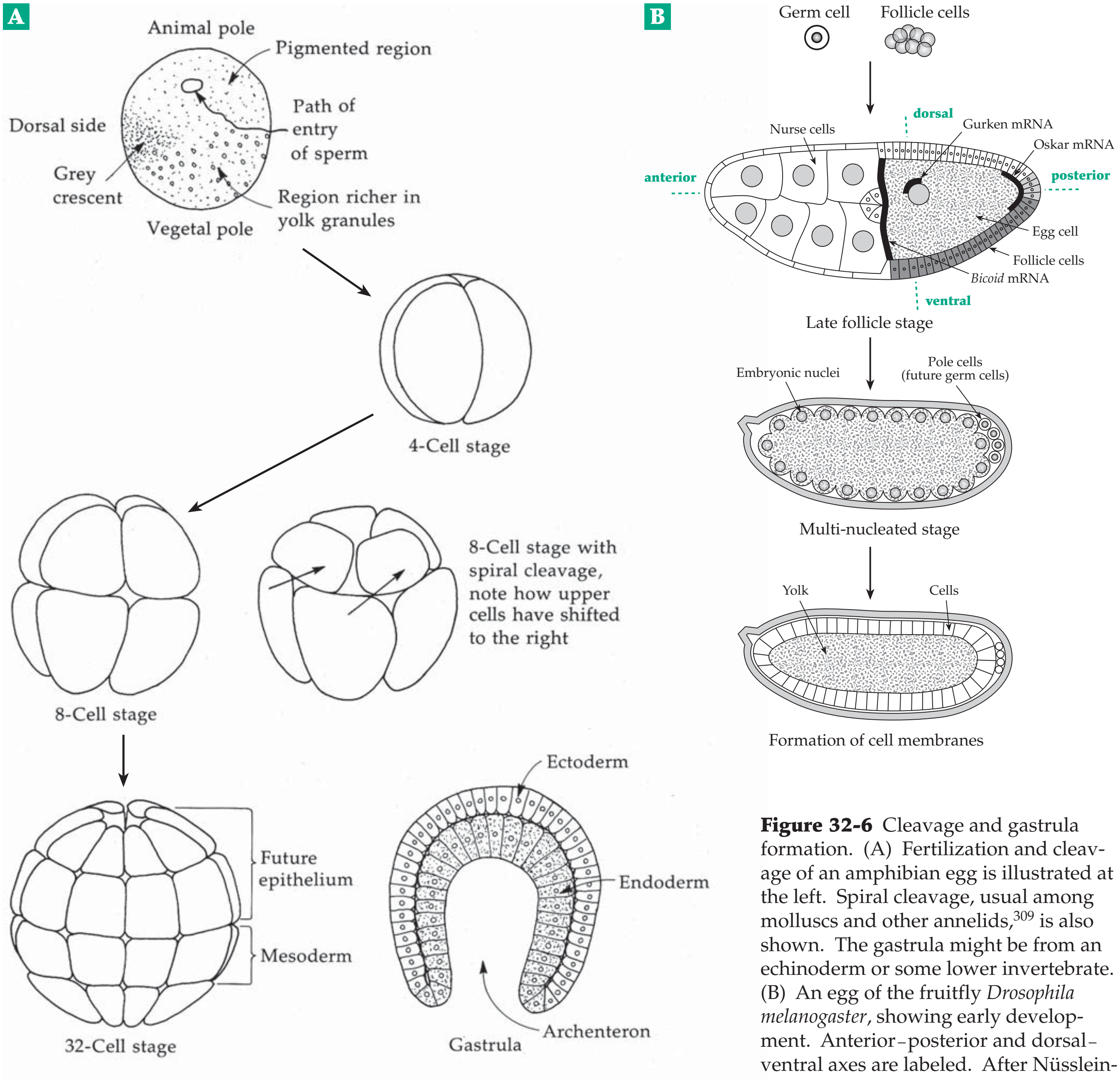

A
Animal pole
Pigmented region
Dorsal side
Path of entry of sperm
Grey crescent
Region richer in yolk granules
Vegetal pole
4-Cell stage
8-Cell stage
8-Cell stage with spiral cleavage, note how upper cells have shifted to the right
32-Cell stage
Future epithelium
Mesoderm
Ectoderm
Endoderm
Gastrula
Archenteron
B
Germ cell
Follicle cells
dorsal
Nurse cells
Gurken mRNA
Oskar mRNA
anterior
posterior
Egg cell
Follicle cells
Bicoid mRNA
ventral
Late follicle stage
Embryonic nuclei
Pole cells (future germ cells)
Multi-nucleated stage
Yolk
Cells
Formation of cell membranes

Figure 32-6 Cleavage and gastrula formation. (A) Fertilization and cleavage of an amphibian egg is illustrated at the left. Spiral cleavage, usual among molluscs and other annelids,[309] is also shown. The gastrula might be from an echinoderm or some lower invertebrate. (B) An egg of the fruitfly *Drosophila melanogaster*, showing early development. Anterior–posterior and dorsal–ventral axes are labeled. After Nüsslein-Volhard (after a drawing by Laurie Grace).[9]

ovum.[300,301] The acrosome reaction also activates the egg, a process that may depend upon NO.[302] In the sea urchin the acrosome contains a large amount of the protein **bindin**, which mediates a species-specific adhesion of the gametes and presumably fusion with the egg membrane.[302a]

Fusion of the sperm membrane with that of the ovum causes a rapid depolarization of the membrane of the ovum and an influx of calcium ions.[303,304] This causes an immediate block to the entrance of any other sperm cells in most species. It also causes the fusion of the membranes of **cortical granules** (several thousand in a mouse ovum) with the cell membrane and release of their contents.[305] The material released to the outside of the ovum includes various proteins and enzymes such as a peroxidase that catalyzes crosslinking of tyrosine side chains (Chapter 25) and hardens the material immediately around the ovum into a tough **fertilization membrane**. Within the ovum a respiratory burst resembling that of activated neutrophils (p. 1074) provides H_2O_2 to the peroxidase.

Fertilization also induces completion of meiosis and formation of a one-cell embryo containing a maternal pronucleus contributed by the egg and a paternal pronucleus derived from the sperm. Each pronucleus undergoes DNA replication and then enters the first mitosis, which yields a two-cell embryo containing one diploid zygotic nucleus per cell.[283,306] Under the influence of the cyclin-dependent kinase cdc2 (see Eq. 26-3) and a hyperphosphorylated form of the protein **nucleoplasmin**, the chromatin of the compact sperm nucleus undergoes decondensation. The sperm basic proteins that coated the DNA are replaced by histones H_2A and H_2B.[307]

Although an amphibian egg is nearly spherical, there is already a strong polarity. The nucleus lies nearer to the **animal pole** than to the **vegetal pole**, which in many eggs is rich in yolk granules. In eggs of amphibians the animal pole is highly pigmented, but the vegetal pole is less so. On one side above the equator, there is a gray crescent. In some animals this marking appears on the opposite side of the egg from the point of sperm cell entry (Fig. 32-6). The gray crescent marks the future back (dorsal) side of the organism and the opposite part of the cell, the future ventral side.[280] The point of sperm entry also marks the ventral side for the mouse, a fact that suggests that the plan of development of mammalian embryos may be basically the same as that of frogs.[308] However, it is the internal components of the cell that actually determine the cell's axes. The cytoplasm of the mature ovum contains an unequal distribution of many materials with a well-developed bilateral symmetry. That this distribution is important is seen from the fact that centrifugation of eggs prior to fertilization often leads to formation of abnormal embryos because of displacement of preformed ribosomes and other materials. It is probably gradients in the concentrations of dormant mRNAs[284] and other metabolites that lead to uneven growth of cells and to the indentation of cells at the vegetal pole, a process that initiates the formation of the endodermal layer of the gastrula (Fig. 32-6). In insects the polarity of the developing ovum (the oocyte) is established by the cytoplasmic bridges from surrounding nurse cells, which are asymmetrically arranged (Fig. 32-6B).[285]

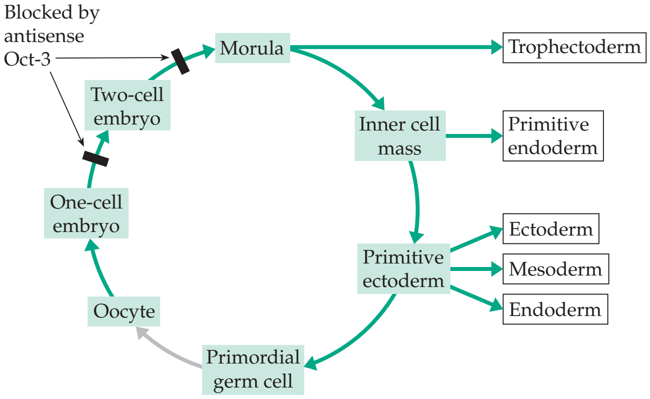


Figure 32-7 Expression pattern of Oct-3 mRNA during mouse development. The green boxes indicate those stages in which Oct-3 is expressed, while the white boxes at the right indicate all types in which little or no Oct-3 mRNA is formed. After Rosner *et al.*[314]

4. Embryonic Development

The fertilized (activated) ovum rapidly undergoes several mitotic divisions, known as **cleavage**, during which no overall growth occurs. The number of cells increases and the DNA replicates at each division, but the overall size of the resulting cell cluster is the same as that of the original ovum (Fig. 32-6). Further development leads quickly to a stage in which a layer of cells (called blastomeres at this stage) surrounds an internal cavity forming a **blastula**. In the sea urchin the blastula, which is released from its protective fertilization membrane ~11h after fertilization,[310a] consists of a single layer of cells In frogs and many other organisms there are two or more layers. In *X. laevis* about 4000 cells are formed in eight hours.[307,310] In mammals a solid cell mass (**morula**) forms first and is later transformed to a **blastocyst**, a hollow ball with an internal cavity.

Early mammalian development has been hard to study because of two facts: the ovum is very small, and a first priority is development of the placenta and of the layers of tissues that surround the embryo.[308,311,312] This occurs in humans within the first week after fertilization. Both the trophectoderm and cells of the inner cell mass (Fig. 32-7) contribute to the extra embryonic tissues.

Development of a mouse beyond the one-cell stage is dependent upon a regulatory gene of the *Oct* family (see p. 1631). The **Oct-3/4** protein, which binds specifically to the DNA motif 5'-ATTTGGAT, consists of two domains, both of which are essential for tight binding to this sequence. One domain is a 75-residue **POU domain** that consists of a helix–turn–helix motif with an amino acid sequence that is highly conserved among mammalian **P**it and **O**ct regulatory proteins as well as some **U**nc proteins of *C. elegans*.[313] The second domain is a 60-residue homeodomain (see p.1900).[313,314] Oct-3/4 appears to be essential not only for cleavage of a one-cell egg but also for progression from a two-cell to a four-cell egg (Fig. 32-7) and also in other embryonic cells.[315] Up to the two-cell stage very little transcription of zygotic genes is observed but further development requires zygotic genes and by the 8-cell stage protein synthesis from maternal mRNAs ends.[283] Oct-3/4 continues to be synthesized in the embryo and is necessary for establishment of pluripotent stem cells in the embryo.[315]

Development in *Drosophila* and other insects follows a somewhat different pathway, as is indicated in Fig. 32-6B. The egg, which is surrounded by follicle cells and 16 nurse cells, does not divide. However, its nucleus divides repeatedly, about once every nine minutes, to form ~6000 nuclei. Only then do separating membranes form to give individual cells.[9,285,316,317] During the first two hours the nuclei form a **syncytium**, in which they are embedded in a common cytoplasm, that allows free diffusion of signaling compounds. At first the nuclei are in the center, but later most migrate to the periphery and form a single layer of cells comparable to the blastoderm of amphibian cells. A few nuclei remain in the central cavity to become yolk cells, and some at the posterior pole become separated into pole cells.

The next stage in embryo formation, which occurs universally, is the invagination of the blastula at the vegetal pole to form a **gastrula**. At this stage the embryo has distinct ectoderm and endoderm cell layers. The cavity, formed in the gastrulation process and connecting to the outside, is referred to as the **archenteron** and is the forerunner of the gastrointestinal tract or enteron. Gastrula formation is more complex in the frog embryo and still more so in the human embryo. In all but the most primitive of animals a third layer of **mesosomal** cells is formed between the endoderm and ectoderm. These three **germ layers** differentiate further as follows. The **ectoderm** yields the skin and nervous sytem; the **mesoderm** the skeleton, muscles, connective tissues, and circulatory system; and the **endoderm** the digestive tract, lungs, and other internal organs and germ cells.

Organ development occurs largely by infoldings of cells from the endoderm and ectoderm. These infoldings appear to be induced by chemical substances secreted by cells of an adjacent germ layer. Thus, ectodermal cells form the **neural plate**, the prospective brain, and spinal cord in response to induction by mesodermal cells lying beneath the neural plate area. The mammary glands also arise from interactions of mesodermal and ectodermal cells, while the formation of the pancreas, liver, and lungs depends upon interactions of groups of cells from endoderm and mesoderm. Because of their transparency zebrafish are especially useful for study of organ formation.[318–321]

During the early stages of embryo growth, development seems to be directed largely by the polarity and gradients of the large amounts of mRNA, yolk and other constituents, which form a prepattern in the ovum.[321a] However, even at very early stages signaling from the nucleus of the ovum to surrounding cells is a necessary part of establishing the cell axes and developmental pattern.

The anterior–posterior axis. Only recently has it become possible to identify some of the specific mRNAs, the signaling molecules, and receptors that are involved in establishing the principal axes of the ovum (Fig. 32-6). Even for this aspect of development the genetic program is very complex, there are many uncertainties, and the details go far beyond the scope of this book. Much of the most important work has been done with *Drosophila* for which numerous mutants have been identified and characterized. Many names of genes and of proteins are derived from a description of the phenotype of a mutant fly. These same names are often used for the corresponding genes or proteins in other organisms. However, a protein may be known by more than one name, depending upon the species.

A *Drosophila* mother deposits mRNA for ~80% of all of her genes in the egg, but not all of the encoded information is used. Because genetic experiments can be done so readily with *Drosophila*, it is possible to ask what **maternal effects** do come from this mRNA and what effects come from the genes of the zygote.[9,88,316] For example, the *bicoid* mRNA that accumulates at the anterior end of the *Drosophila* egg is produced by nurse cells and is transported into the ovum. If the mother has defective *bicoid* genes the eggs die. A normal bicoid gene present in the father does not prevent death. Another maternal effect protein is encoded by *nanos*, whose mRNA accumulates at the posterior pole of the ovum. The maternal *torso* gene, which acts on follicle cells, is also needed for the anterior-posterior axis formation.[316] About 30 of these maternal effect genes are active in organzing the pattern of the embryo.

The *Drosophila bicoid* mRNA forms an anterior-posterior (A-P) concentration gradient, which controls early development along the axis. Another protein called **Staufen**, which forms a complex with the *bicoid* mRNA, assists in moving the RNA along on microtubules to form the gradient.[321b] The microtubules, together with associated transport proteins, are also essential to the prepatterning of the ovum.[9,88,285,322,323] Bicoid protein is a transcription factor that is synthesized at the sites of its mRNA accumulation. It is absent from the eggs when they are laid but soon appears.[316] Bicoid binds to the CRMs of target genes and activates them. Cooperative binding of multiple copies of the transcription factor may be necessary to provide the observed sensitivity to concentration. A transcription factor may bind at many sites on the DNA of the zygote or of nurse cells. For example, a gene called *even-skipped* (*eve*) is expressed in seven stripes in the blastoderm embryo. The enhancer that controls *eve's* second stripe contains at least five binding sites for Bicoid as well as others for the Hunchback, Giant, and Krüppel proteins.[44]

A different gradient along the A-P axis is formed by the **Nanos** protein, whose mRNA localizes in the cytoplasm of the posterior part of the ovum. Like the *bicoid* mRNA, *nanos* mRNA forms a complex with the Staufen protein. Messenger RNA transcribed from a gene called *oskar* is also necessary for development of the posterior region of the ovum. An additional gene, called *gurken*, is also involved in establishing the A-P axis. The nucleus of the ovum secretes *gurken* mRNA (Fig. 32-6), which is translated to the protein **Gurken**, a TGF-α-like protein that carries signals to follicle cells. They, in turn, help to organize the microtubules in the ovum. The nucleus of the ovum moves, sending the Gurken signal first to the posterior pole cells, then to a position that marks the dorsal side of the ovum. Thus, it participates in establishing both the A-P and dorsal–ventral axes.[324]

The animal–vegetal axis of sea urchin eggs is established during oogenesis, but the mechanisms are still unclear. After fertilization a distinct pattern along the axis is established by the 60-cell stage. Signaling from the vegetal pole appears to play an important role.[89] In the presence of lithium chloride, which is known to affect inositol triphosphate (InsP_3) metabolism (Fig. 11-9), embryos develop an excessive fraction of endoderm and mesoderm (tissues of the vegetal half) at the expense of ectoderm (tissues of the animal half).[89,325] Some data suggest that this effect implicates the **Wnt signaling pathway** and the protein **β-catenin**. This is the same protein that links C-terminal tails of cadherin to the actin cytoskeleton.[60] Its cytoskeletal and signaling functions seem to be controlled separately.

In the frog *Xenopus* the ovum accumulates an RNA (Vg1 mRNA) that encodes a growth factor of the TGF-β family (Chapter 30, Section A,6).[280] Initially present throughout the ovum, it moves and forms a gradient of concentration that is highest at the vegetal pole.

The dorsal–ventral (D-V) and right–left axes. Establishment of the D-V axis in *Drosophila* requires the participation of several genes. As mentioned above, *gurken* affects follicle cells. These cells cooperate with the ovum to set up a concentration gradient of the maternal gene **Dorsal**, the protein encoded by the *dorsal* gene.[9,280,326–328] Dorsal is a transcription factor related to NF-κB (Fig. 5-40). In the early embryo of the fly it forms a gradient in which it remains largely in the cytoplasm on the dorsal side but is mostly taken up into the nuclei on the ventral side. There it binds to a series of complex enhancers, each 300–1000 kbp in length, which act along the D-V axis. The enhancers interpret the concentration of Dorsal at five different threshold levels.[328]

The establishment of a second axis in vertebrate embryos is a complex process, which depends upon prior formation of a mesoderm layer. Development of mesodermal cells (not shown in the more primitive gastrula illustrated in Fig. 32-6) is induced by diffusion of a growth factor from the vegetal pole.[329–331] **Activin**, a member of the TGF-β family, has been proposed as the natural inducer of mesoderm. More recently **nodal**, a different TGF-β-like protein, has been suggested.[332] The fact that induction can occur through thin (20 μm) filters without any cell–cell contact indicates that specific chemical agents are responsible. Induction of the mesodermal layer in *X. laevis* appears to be an effect of an epidermal growth-factor-like protein. Additional factors are also needed to establish the D-V axis in vertebrates[326] in which the dorsal side is homologous to the ventral region in *Drosophila*.

Expression of another set of genes establishes the **right–left** axis and characteristic asymmetries of the body.[333–336] In the chick activin 2β, also a member of the TGF-β family, as well as Nodal and **Sonic hedgehog** (Shh) participate in control. The gene *Pitx*2 is a downstream transcription target for this signaling cascade in vertebrates.

Spemann's organizer. In 1924, the German physiologist Hans Spemann with Hilde Mangold transplanted a small piece of tissue from the dorsal lip of a newt blastula to a site on the ventral side of an early blastula of a differently colored species. The embryo developed a small secondary embryo, most of the tissues of which came from the host, not from the transplanted piece. It was concluded from this famous experiment that the transplanted vegetal tissues had supplied a diffusible inducer.[13,14,337–339] This morphogen caused the cells of the ventral surface of the blastula to secrete other morphogens. The signaling center in this part of the blastula surface is known as **Spemann's organizer**, or simply the organizer. It utilizes more than one inducer and a complex set of signaling interactions.[340] The cascade that induces formation of the organizer involves the Wnt–β-catenin and TGF-β pathways and the transcription factor **Smad4**.[41,341,342] The organizer secretes protein such as **noggin**, **folistatin**, and **chordin**, as well as nodal and other members of the TGF-β family.[332,343,344] They establish the D-V axis and also direct the development of the head and the initial patterning of the central nervous system. Noggin seems to be a neural inducer.[338]

Patterns, signaling pathways, and homeotic genes. While gradients in *Drosophila* eggs establish the anterior–posterior axis, products of other genes specify the developmental fates of cells of specific lineages and of cells found in particular spatial domains. Many *Drosophila* genes are needed to establish patterns, e.g., spacings of similar elements such as hairs, components of compound eyes, and whorls of plants.[345] Insects are organized into a pattern of consecutive segments from head to tail along the A-P axis.[13,14] In *Drosophila* there are typically 17 of these segments, some carrying appendages such as antenna, legs, and wings. Each segment develops under a different set of influences from neighboring cells. Development is controlled by ~30 **segmentation genes**, which determine the number of segments and their internal organization. Of these genes one set of at least six **gap genes**, among them *hunchback* (Hb) and *Krüppel,* are expressed first. Both of these genes encode zinc-finger transcription factors (see Fig. 5-38). The Hunchback protein acts mainly on head parts and upper thorax, while Krüppel influences development of the thorax. Also among the segmentation genes are eight "pair-rule" genes and at least 16 segment polarity genes. The interactions of the products of these genes creates a prepattern that provides positional information and guides further development.[346] Somewhat similar to the segmentation of the insect body is the development of skeletal muscle from a series of mammalian embryonic blocks known as somites.[13,347,348] In *Drosophila* signaling pathways involving the secreted proteins EGF, Decapentaplegic (Dpp),[349] Wingless (Wg), Hedgehog (Hh),[349,350] and Notch are used repeatedly to provide positional information. The names of these proteins describe effects of mutations on the limbs of the insect, but the proteins have a much broader significance. They represent an evolutionarily conserved set of intercellular signaling pathways.[346] Wingless is the first member of the previously mentioned Wnt family. Proteins of the Hh family, including the vertebrate Shh, control a large variety of processes that include development both of limbs and of the nervous system.[351–355] The Hh and Shh proteins all carry a molecule of cholesterol covalently bound in ester linkage to the C termini of the biologically active N-terminal domains of these proteins.[350,353] Defects in Shh signaling causes some human birth defects, and it is possible that drugs that inhibit cholesterol synthesis may have adverse effects on embryonic development.[352,355]

The *Drosophila* Notch 1 is a 300-kDa integral membrane protein that contains 36 EGF-like repeats. Its activation by proteolysis produces a 200-kDa N-terminal portion and a 120-kDa C-terminal fragment, which contains a transmembrane domain. The small intracellular domain of this fragment is then released by protease action and travels to the nucleus where it activates several target genes.[356–358] The Notch signaling pathway is conserved in all metazoans and influences many interactions that control cell fate during development.[359] The proteolytic cleavages of the Notch protein parallel those of both the ErbB-4 growth factor receptor (see Table 11-3) and the amyloid precursor protein APP that is pictured in Fig. 30-34. The same type of protease (γ-secretase or presenilin in the human brain) cleaves all three of these proteins.[359–361] Actions of Notch are modulated by posttranslational alteration, e.g., glycosylation by a fucosyltransferase encoded by gene *fringe*.[362,363] A homolog of *fringe*, called *lunatic fringe*, encodes an essential component of somite formation in the mouse.[348,364]

Functioning together with the signaling pathways are **selector genes** that determine which specific pathway is to be influenced. For example, eyes, antennae, legs, or wings of a fly may be selected.[346] *Antennapedia* (*Antp*) is one of these genes. Several of the *Drosophila* selector genes as well as some pair-rule genes[365] are also known as **homeotic genes** (see also Chapter 28, Section C,6). Homeotic genes were first recognized by the fact that mutation causes conversion of one segment of an insect's body into the homologous tissues of another segment.[366,367] For example, a mutation in *antennapedia* (*ante*) changes the antenna into a leg. Similar genes are also active in vertebrates, e.g., in the development of the chick embryo limb bud, the very tip (the last 20 cell diameters of length) contains cells that differentiate into the various elements of the limb in a relatively autonomous manner. If this *progress zone* from one limb bud is grafted onto the end of another limb bud, the bones and cartilaginous elements of the limb are repeated. Both the number and morphology of fingers and toes are determined by homeotic genes[367a] as is the formation of sphincters in the developing gut.[367b] Homeotic genes (also known as *Hox* genes) contain a conserved sequence of 180 bp that specifies a 60-residue protein **homeodomain** (also known as a homeobox). The homeodomain folds into a helix–turn–helix motif characteristic of many transcriptional regulators (Figs. 5-35, 28-3, p. 1631).[365,367–371] *Hox* genes are found among all forms of life. Hundreds have been described.[372] They include genes for the previously mentioned transcription factors of the Oct and Pit families as well as for the yeast mating type proteins MATa1 and MATα2 (p. 1880).[17] The *Drosophila* genome contains eight *Hox* genes, while the human genome has ~40, which are organized into Hox clusters.[373] Hox genes are also abundant in plants.[345,372,374]

Despite intense interest the role of homeodomain proteins in development is not well understood. The highly conserved motif binds to DNA at many places in the genome. Current thinking is that homeodomain proteins interact with other regulatory proteins, and that various combinations of these proteins provide the information needed to direct development.[375]

D. Specialized Tissues and Organs

Here are a few details about development of mammalian tisssues. We'll begin with the blood and connective tissues, which arise from embryonic mesodermal cells.

1. Blood Cells and the Circulatory System

Every second of life a human must produce about 2.5 million red blood cells, about 2 million granulocytes, and many lymphocytes as well as other less numerous leukocytes. All of these arise from **multipotential stem cells** found in the bone marrow.[376–379] Each of these stem cells divides to form one daughter stem cell and one **progenitor cell**.[380] The progenitor cells are also stem cells but have differentiated into **myeloid**,[381] **erythroid**, and **lymphoid**[382,383] cells. These differentiate further as is indicated in Fig. 31-2. Mature blood cells of most types have short lifespans and must be regenerated from stem cells continuously.

At all stages the differentiation process is regulated by the microenvironment which is rich in specific protein growth factors, several of which have been discussed in Chapter 30. Among the 20 interleukins, three stimulate growth of both multipotential progenitor cells and erythroid progenitor cells.[380,384] The acetylated tetrapeptide Ac-Ser-Asp-Lys-Pro inhibits stem cell proliferation. Granulocyte-macrophage colony-stimulating factors stimulate proliferation of both granulocytes and macrophages. The kidney cytokine **erythropoietin**, a 30.4-kDa glycoprotein,[385] is a primary regulator of red blood cell formation. Its action on a differentiated stem cell initiates massive hemoglobin synthesis and terminal differentiation of the erythrocyte. **Thrombopoietin** promotes formation of megakaryocytes and also their maturation and release of platelets to the blood.[386] **Thymopoietin** promotes early T-cell differentiation. Activated macrophages secrete interleukin-1, which stimulates maturation and proliferation of B lymphocytes. Interleukin-2 (T-cell growth factor; Fig. 30-8) is produced by activated T lymphocytes and is needed by T lymphocytes for long-term helper and cytotoxic functions. Differentiation of the stem cells into the erythroid lineage requires transcription factor GATA-1, development into erythrocytes requires GATA-2, while development into T lymphocytes requires GATA-3.[387]

Globin genes. The genes that encode the human globins from which hemoglobin is formed are found in two clusters, the α-like genes on chromosome 16 and the β-like on chromosome 11. They are developmentally regulated, different genes in the clusters being active at different stages of development. Mammalian hemoglobins (Chapter 7) each contain two α chains or two related ζ chains and two other chains, β, γ, δ, or ε. Adult hemoglobin is mainly $\alpha_2\beta_2$ but contains small amounts of $\alpha_2\delta_2$. In early embryos the hemoglobin is $\zeta_2\varepsilon_2$, but during the second to sixth weeks of embryonic life the two fetal hemoglobin chains γ^G and γ^A replace the ε chains. The switch from fetal to adult hemoglobin begins a few weeks before birth and is complete by about ten weeks after birth. The β-like gene cluster contains five genes encoding globins ε, γ^G, γ^A, δ and β (Fig. 27-10). Each gene consists of three exons separated by two introns and has rather similar control signals. These include CACCC at ~–100, CACA at ~– 92, CCAAT at ~– 75, and ATAAAA (TATA sequence) at ~– 30, as well as AATAAA (cleavage and polyadenylation). The CCAAT sequence appears twice in the γ^A promoter. A variety of transcription factors and chromatin modifiers influence the expression of these genes.[388,389]

Only a small amount of fetal hemoglobin ($\alpha_2\gamma_2$) is produced after infancy. There are two genes, γ^G and γ^A, for the β-like chains of fetal hemoglobin. A few adults make large amounts of fetal hemoglobin and this hereditary **persistence of fetal hemoglobin** has survival value for persons carrying thalassemia genes. This condition may result from a single base change in the CCAAT sequence found upstream (~75 bp) of the globin genes. Many other genes are also preceded by the same sequence, which in extended form is often TTGGPyCAAT. In one individual with persistence of fetal hemoglobin the first G in this sequence was replaced by A in one of the two CCAAT sequences present in the γ^A gene.

Thalassemias. In these important hemoglobin diseases the α or β chain either is absent or is present in far less than stoichiometric amounts.[14,390] About 40 point mutations in the β globin gene have been described among patients with β thalassemia, in which β chains are missing (β^0 thalassemia) or are present in reduced amounts (β^+ thalassemia). These mutations sometimes occur in control regions at the 5' end of the gene. For example, a change of the TATA sequence from ATAAAA to ATACAA causes decreased transcriptional efficiency and β^+ thalassemia. Other mutations result in abnormal splicing or in instability of the β globin. Deletions may result in the complete absence of the gene or in a frameshift that results in nonfunctional globin and β^0 thalassemia.

How do embryonic cells choose to transcribe only the embryonic globin genes? The decisions to switch from embryonic to fetal and from fetal to adult at appropriate times appear to be controlled by interactions with appropriately expressed transcription factors. Especially impressive is the total and permanent cessation of transcription of the embryonic globin ε gene at about the seventh week of gestation. Transcription of the β globin gene cluster is controlled by a powerful enhancer, the **locus control region** ~6–22 kbp upstream of the promoter.[391] In addition, ~270 bp upstream of the ε gene promoter is a **silencer**, a DNA sequence to which inhibitory proteins may bind and, in cooperation with the enhancer, may completely silence the ε gene while allowing transcription of other genes in the cluster.[392,393] The ζ gene, in the α globin cluster, is silenced by the binding of an NF-κB transcription factor to a 108-bp segment of DNA located 1.2 kbp to the 3' end of the gene.[394] The globin genes are shielded from action of nearby enhancers by **insulators**, DNA sequences that often contain CpG islands.[395]

Blood vessels: vasculogenesis and angiogenesis. Early in development of an embryo a network of blood vessels is formed from mesenchymal progenitor cells (vasculogenesis).[396–398] Later, in either the embryo or the adult, new capillaries are formed (angiogenesis). As the organism develops these new vessels are "pruned," and the vascular bed is remodeled to a tree-like form with vessels of both large and small diameter.[399,400] A key activator of angiogenesis is the vascular endothelial growth factor **VEGF**.[401,401a] However, a number of other proteins including **endoglin**, a TGF-β binding protein,[402] the clotting factor thrombin,[400,403] and the ribonuclease **angiogenin** (p. 648) exert their influences.[404] There are diseases that result from defective angiogenesis. However, a major interest in this process arises because of the essential role of angiogenesis in the growth of cancer[405] and as a complication of inflammatory processes. Angiogenesis in the retina is a major cause of blindness resulting from diabetes mellitus or from macular degeneration.[405,406] There are natural antagonists of angiogenesis,[407–409] and efforts are being made to utilize them in therapy. A related project is development of completely tissue-engineered blood vessels for surgical use.[410]

2. Cartilage, Tendons, Bone, Muscle, and Fat

Mesenchymal cells differentiate into cartilage, bone, muscle, adipose, and other connective tissues.[411] **Chondrocytes** synthesize the variety of collagens (pp. 431–426) that are needed for synthesis of cartilage and other connective tissues. The 32 or more genes encoding the polypeptides needed for synthesis of the 19 types of vertebrate collagens (Table 8-4) are developmentally regulated in a complex manner.[411] Their promoters contain TATA and CCAAT sequences as well as other presumed regulatory codes. The first intron of several collagen genes has also been identified as a control region containing enhancers.[412] The elastic fibers (p. 436) owe their properties in part to elastin. The control region of elastin genes lacks the TATA sequence but has SP1 binding sites. As with many other mammalian genes, a diversity of protein products, many in small amounts, are made by alternative splicing.

Collagen fibrils provide the scaffolding for formation of bone, whose composition is considered on pp. 440–443. Bone develops under the influence of **bone morphogenic proteins** BMP-2 to BMP-7.[413–415] Most of these are cytokines of the TGF-β family. Noggin (p. 1899) antagonizes the action of the BMPs.[344] A characteristic of bone is rapid remodeling (p. 441) by which ten per cent of skeletal bone is replaced every year. The balance between action of the bone-forming osteoblasts and the bone-resorbing osteoblasts is regulated by surface proteins responding to **c-Fos** and to **interferon-**β.[416]

Muscle, whose structure and function are discussed in Chapter 19, develops in response to four members of the myoD family. These include myoD, **myogenin**, **myf5**, and **MRF4**.[417–419] All are muscle-specific transcription factors of the basic helix–loop–helix class. An unusual aspect of muscle development is formation of multinucleate **myotubes** (muscle fibers; p. 1096).[420] Apoptosis plays an important role in muscle development and can present significant complications in damaged cardiac muscle.[421] Defects in several developmental control genes are responsible for congenital heart diseases.[422]

3. Epithelia

Epithelial tissues, which line both internal and external surfaces, arise from all three cell layers of the blastula. The epidermis (Box 8-F) arises from ectoderm, while the lining of the digestive tract is formed by endodermal cells. Mesoderm provides the linings of blood vessels. About 60% of differentiated tissue types in the mammalian body are epithelia.[123] Stem cells or progenitor cells are present and provide for renewal.[123,124] While epidermal stem cells are located in deep layers of the skin, **keratinocytes** are readily cultured *in vitro* and can give rise to fully differentiated multilayered skin.[423] Development appears to require transcription factors related to **Oct-2**[424] as well as **p63**, a homolog of the tumor suppressor p53.[425,426] Mice deficient in the aspartyl protease **presenilin 1** (which is defective in some forms of Alzheimer disease, Chapter 30) develop characteristic epidermal skin tumors. The β-catenin–Wnt signaling pathway (p. 1899) seems to be involved.[427] The gastrointestinal endoderm develops its highly convoluted surface under some control by the Notch signaling pathway.[428,428a] Endothelial progenitor cells (**angioblasts**) are responsive to many signaling molecules[429] including thrombin.[430] The *Drosophila* eye develops from the epithelium, again through signaling via Notch and other morphogens.[430a] Of outstanding importance to epithelial cells in general is their ability to form complex communicating junctions.[430b]

4. The Nervous System

Development of the vertebrate central nervous systems is initiated during gastrulation through an interaction between the dorsal ectoderm and an infolding of the dorsal mesoderm. Several different diffusible inducers are involved. These include noggin,[344] folistatin, and other members of the TGF-β family as well as thyroid hormones, basic fibroblast growth factor (bFGF), and sonic hedgehog.[329] The nervous system develops over a period of a few days with differentiation of precise numbers of neurons, astrocytes, and oligodendrites in successive waves. The order in which various cell types arise is determined by the order in which transcription factors such as Hunchback, Krüppel, and others are expressed.[431] Multipotential neural stem cells provide the new cells that are required.[432,433] Neural tissue from a region called the **neural plate** develops into a neural fold. The latter is closed to form the **neural tube**,[434] within which the notochord, the precursor to the spine, as well as the neurons, glia, and other cells grow. The **neural crest** forms as an outgrowth from the dorsal surface of the neural tube under the influence of inducers of the Wnt and BMP families.[434a] Cells migrate from this crest to form the peripheral nervous system, melanocytes, and cranial cartilage.[434b] The pituitary, a central component of the neuroendocrine system, develops from tissues from the midline part of the anterior neural ridge.[435]

The **floor plate**, which develops along the midline of the vental surface of the neural tube, is the source of Sonic hedgehog,[436,437] **netrin-1**, and other secreted molecules.[436] Some of these participate in **axon guidance** by which the growing tips (**growth cones**) of axons are able to connect to the correct "targets."[437a] For example, every visual receptor cell in the retina must send its signal to the correct locations in the visual cortex of the brain.[438] How can this be accomplished? Over a century ago Ramon y Cajal proposed that chemoattraction, analogous to chemotaxis of bacteria, might be involved.[439,440] A hundred years later Tessier-Lavigne and coworkers isolated the first of these attractants, **netrin-1** and **netrin-2,** from 25,000 pulverized chick brains.[441,442] Of these two closely related 75- and 78-kDa molecules netrin-1 is produced only in the floor-plate cells. Like the less well understood nerve growth factor (Fig. 30-7), netrins induce outgrowths of neurites and also are chemoattractants for nerve growth cones.[443] The netrin receptor is known as **DCC** (Deleted in Colorectal Cancer).[444] A nematode protein UNC-6 is related to the netrins.[445]

Growth cones are subject to both chemoattractant and chemorepellent effects of guidance molecules and also to attraction or repulsion resulting from cell–cell contacts. To complicate the picture further, the netrins and also the brain-derived neurotrophic factor BDNF (see Fig. 30-6D) may first attract, and then after a period of adaptation or desensitization, repel a growth cone.[443,444,446–448] Consecutive phases of desensitization and resensitization may result in a zig–zag path of growth. The netrins were recognized first by observing growth of **commissural neurons**. These neurons originate within the spinal cord on one side or the other or the midline. They grow down toward the floor plate attracted by the netrin-1 or **BDNF** produced there. The neurons then cross the midline before turning and growing toward the brain. After crossing the midline the growth cones become insensitive to netrin but are repelled by a molecule (first recognized in *Drosophila*) called **slit**. Its receptor is appropriately named roundabout (**Robo**).[449] A similar receptor in zebrafish (called **astray**) is required for retinal axon guidance.[449] An important aspect of neuron guidance is apoptosis induced by misdirected growth.[450] The Notch receptor apparently participates in this decision in the mammalian CNS.[357] Positive signals for axon growth often involve the MAP kinase pathways, while inhibition may involve INS-2*P*–Ca^{2+} signaling.[451]

Chemoattractants that function in development of the cerebral cortex include several **semaphorins**.[445,452,453] A separate family of attractant and repellant compounds, the semaphorins have been identified in insects, chickens, and mammals. They play a role in regulation of communication between neurons. Because of the complexity of the brain, the study of growth of neurons with the brain is difficult. The $>10^{12}$ neurons each contact, on the average, 100 different cells. Some insight comes from mutant mice with names such as *reeler*, *scrambler*, *stargazer*, and *Yotari* (Japanese for tottering).[454] The single defective gene in these mice can be identified and studied. For example, *reeler* mice are defective in **reelin**, a large glycoprotein of the extracellular matrix (ECM).[366,379,455,456] The *reeler* phenotype can also result from mutation of the gene *disabled*-1, which encodes a cytosolic tyrosine kinase. Other mutations in mice implicate **VLDL** and **apoE** receptors (Chapter 21) in these developmental abnormalities.[379,455]

The *stargazer* mutant mouse is ataxic and epileptic. It lacks functional **AMPA receptors** (Fig. 30-1), which apparently are not delivered successfully to the synapses in the cerebellum in which they function.[380,386] Mutation of a transmembrane protein **stargazin**, which may interact with the AMP receptor, causes the symptoms.[457,458] **NMDA receptors** (Fig. 30-20) are involved in synapse formation in the brain. Filopodial extensions on dendrites, triggered by electrical activity, are essential for synapse formation,[459] which occurs rapidly.[459a] Activation of NMDA receptors is apparently also necessary.[379,460] Without this stimulation the excitatory glutamatergic neurons of the developing brain undergo apoptosis.

Why do neurons grow in the embryo but not in most parts of the adult CNS? Two proteins called **Nogo** were isolated from bovine brain. Their sequences were utilized in identifying the *Nogo* gene and three human isoforms of the protein.[389,461] The large 250-kDa Nogo-A is present both in myelin and in the endoplasmic reticulum. Both Nogo-A and the diffusible 35-kDa Nogo-B are inhibitory of neurite outgrowth.[462] This effect, as well as the crowding of regenerating neurons by the chondroitin and other matrix components,[396] may provide obstacles to nerve regeneration.[463]

E. Development of Green Plants

Green plants may have diverged from a common ancestor with animals ~1.6 billion (1.6×10^9) years ago. How do the genomes of present-day plants and animals compare? There are many similarities in basic metabolism. These arise from the intrinsic chemical properties and reactivities of cellular components and from the coevolution of plants and animals. Plants and animals also utilize similar structures and similar control of chromatin. However, in the control of development there are great differences.[464] For example, the *Arabidopsis* genome contains no relative of the *Drosophila* Gurken, no receptor tyrosine kinases, no relatives of transcription factor NF-κB. However, there are similarities in parallel pathways utilized by plants and animals.

The structures and life cycle of angiosperms[465] are described briefly on pp. 29–30. The alternating haploid (n, **gametophytic**) and diploid ($2n$, **sporophytic**) phases of the life cycle[466,467] are diagrammed in Fig. 32-8A. Following flowering a diploid **mother cell** within the **ovule** undergoes meiosis to form four haploid **megaspores**. After mitosis a single egg cell is formed. Within pollen sacs in the anthers of each mother undergoes meiosis to yield four haploid **microspores**. Following mitosis these develop into pollen grains each of which contains two sperm cells as well as a vegetative nucleus. After falling upon the stigma surface and growth of the pollen tube, one of the sperm cells fuses with the egg to give the diploid zygote. The other sperm unites with the specialized diploid **central cell** in the ovule to form a triploid ($3n$) **endosperm nucleus**, which develops into the **endosperm**, the food storage tissue of the seed. Endosperm contains two tissues, a starchy inner layer and a protein- and oil-rich outer layer.[468]

Variations of the life cycle occur. For example, a process called **apomixis** leads to asexual formation of seed.[469] In many plants, including maize, separate flowers form the ovule and the pollen. This is one mechanism for avoiding inbreeding.[470] In many plants systems of **self-incompatibility** have evolved.[471] In some, e.g., *Arabidopsis* and other crucifers, pollen germination is disrupted unless it falls on a stigma possessing a different allele-specific receptor. In other cases development of the pollen tube is disrupted at a later stage. In maize and in more than 150 other species a mitochondrial and therefore maternally inherited trait prevents formation of a functional pollen.[472,473] Male-sterile plants, which carry this trait, are very useful in plant breeding. However, a near disaster occurred in the United States in 1970 when the fungal disease Southern corn leaf blight attacked the male-sterile maize that had been used for production of 85% of commercial hybrid maize. The mitochondrial defect in pollen formation also resulted in an increased sensitivity to the fungal toxin. The target of the toxin is a 115-residue pore-forming polypeptide in the inner mitochondrial membrane.[474] The male sterility can be reversed if the plant carries two nuclear **restorer genes**.[472] One of these encodes an aldehyde dehydrogenase, but its mechanism of action is unclear.

The plant embryo is a juvenile form, the seedling. In *Arabidopsis* the zygote, which is surrounded by maternal diploid tissue, divides asymmetrically. The resulting apical and basal cells (Fig. 32-8B) differ in several ways. The small cytoplasm-rich apical cell is partitioned into eight proembryo cells by two rounds of vertical division and one horizontal division. The larger basal cell contains a vacuole and divides repeatedly horizontally to give 7–9 aligned cells. Only the uppermost of these becomes a part of the embryo. The others form an extra-embryonic suspensor (Fig. 32-8B).[475] The apical part of the embryo develops the shoot meristem and the central part the radial pattern of tissue layers characteristic of plants. The root meristem develops from the basal portion of the embryo. Movements of proteins that provide positional cues are involved in the development of the embryo.[476,477] Early embryonic and endosperm development is largely under maternal control. Most paternal genes may be initially silent.[478]

Many angiosperms develop **fruit** from tissues of the ovary (Fig. 32-8A). The development and ripening of fruit is also complex and highly regulated.[479,480]

Formation of seeds is a slow process. For example, in wheat the mature embryo, which consists of ~10^5 cells, develops over a seven week period. Seeds may live from a few years to 1000 years or more.[481] Subsequent germination of the seed into a seedling requires only two days.[482] The very dry embryo is converted into a highly hydrated plant whose further growth requires uptake of very large amounts of water. Many plants also synthesize large amounts of oxalic acid. This may arise from ascorbate (p. 1135) or via oxidation of glycine (Fig. 24-20). One of the earliest mRNAs to appear during seed germination encodes a 125-kDa glycoprotein called **germin**. This protein, which exists as multiple isoforms, is a copper-dependent oxalate oxidase (Eq. 32-1) which generates hydrogen peroxide. The latter is probably needed to

$$C_2O_4^{2-} + 2\,H^+ + O_2 \rightarrow 2\,CO_2 + H_2O_2 \tag{32-1}$$

crosslink cell-wall polymers. Germin may also be useful to plants in defense against oxalate-forming fungi.[482]

The rapid vegative growth, which includes development of shoots, leaves, and flowers, is controlled by a variety of transcription factors.[483] Among these are homeodomain proteins that control differentiation of meristem cells.[484–486] The induction of flowering is especially complex, involved day length, light quality, and effects of gibberellins.[486] At the ends of their lives plant cells die slowly from **senescence**. In this process many materials are recycled for use by new cells. Other plant cells die via the **hypersensitive response**, a form of programmed cell death.[486a]

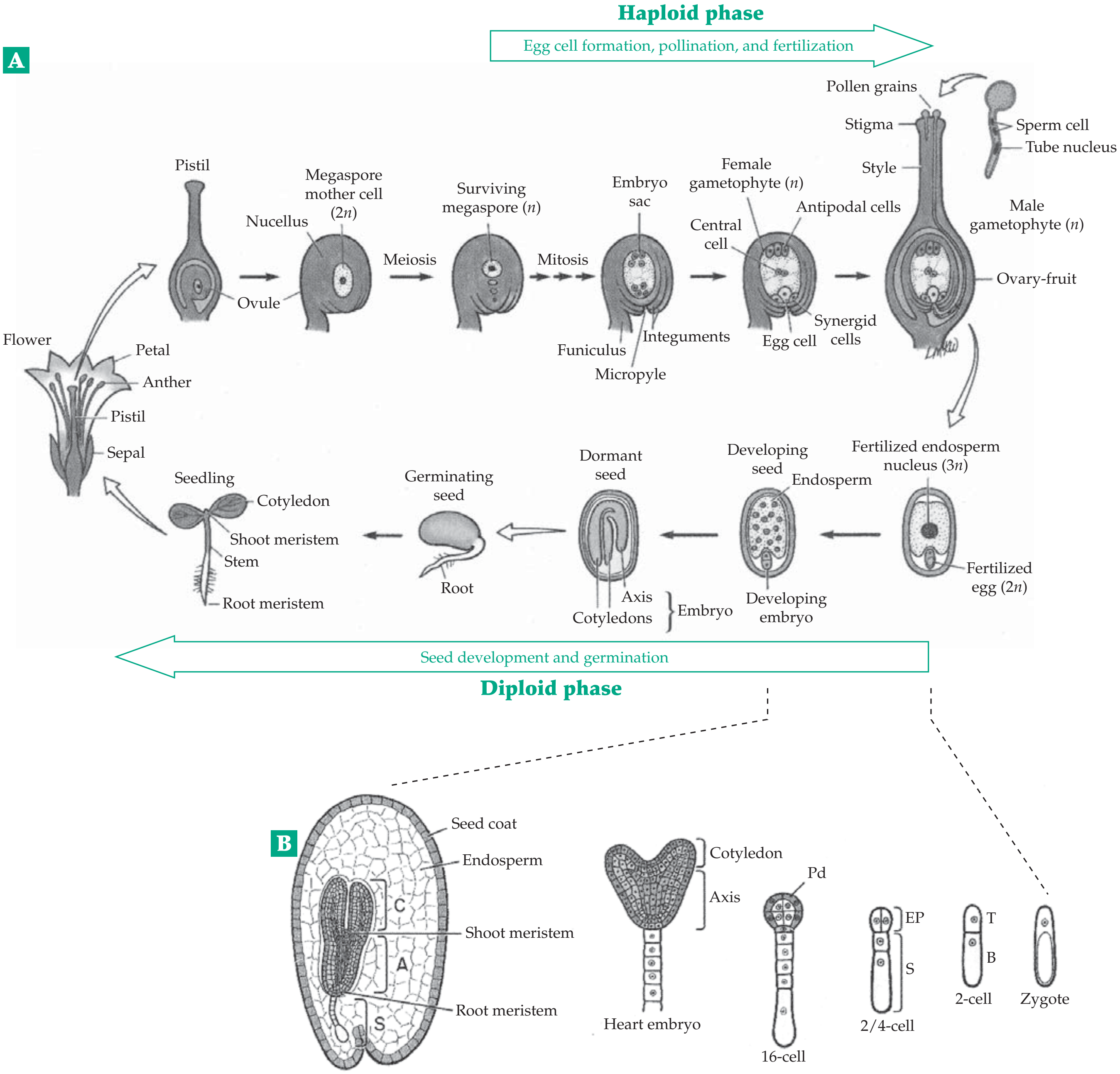

Haploid phase
Egg cell formation, pollination, and fertilization
A
Pollen grains
Stigma
Sperm cell
Tube nucleus
Style
Pistil
Megaspore
mother cell
(2n)
Nucellus
Surviving
megaspore (n)
Embryo
sac
Female
gametophyte (n)
Antipodal cells
Male
gametophyte (n)
Central
cell
Meiosis
Mitosis
Ovary-fruit
Ovule
Synergid
cells
Egg cell
Integuments
Funiculus
Micropyle
Flower
Petal
Anther
Pistil
Sepal
Seedling
Cotyledon
Shoot meristem
Stem
Root meristem
Germinating
seed
Root
Dormant
seed
Axis
Cotyledons
Embryo
Developing
seed
Endosperm
Developing
embryo
Fertilized endosperm
nucleus (3n)
Fertilized
egg (2n)
Seed development and germination
Diploid phase
B
Seed coat
Endosperm
Shoot meristem
Root meristem
C
A
S
Cotyledon
Axis
Heart embryo
Pd
16-cell
EP
S
2/4-cell
T
B
2-cell
Zygote

Figure 32-8 (A) The life cycle of a flowering plant with emphasis on egg-cell formation and seed development. (B) Some further details of embryo development. T, terminal cell; B, basal cell; C, cotyledon; A, axis; SC, seed coat; En, endosperm; EP, embryo proper; S, suspensor; SM, shoot meristem; Pd, protoderm; RM, root meristem. From Goldberg *et al.*[466] with modification.

F. Aging

Why do we age? This question is often asked but the answers are not simple. Do our tissues deteriorate with age as a result of damage to proteins? From an accumulation of mutations in our DNA? From attacks by free radicals? From loss of hormone receptors? From misregulation of mitosis?[487] From loss of telomeres on the ends of replicating DNA (Chapter 27, Section C,8)?[488,489] From an internal genetic program that dictates our life span? All of these possibilities may be partially correct. The simple answer is that "we just wear out." However, different parts wear out at different rates and in different ways.

Perhaps we should be amazed that the human body can live for an average of 75–80 years.[490,491] If we all avoided accidents and could cure all recognized diseases we might live an average of ~90 years.[492] A few very healthy people live for 100 years or more, ~20 years longer than average.[493] The maximum lifespan at present seems to be ~114 years. Long life tends to run in families, indicating a genetic component that can be identified.[493] However, this component is relatively small.[494]

Why do small rodents live only 2–3 years while we often live nearly 100 years?[494,495] Is it because their rate of metabolism is high? But bats have a comparable metabolic rate to mice, yet live ten times longer.[496] Nematodes live only ~20 days and fruitflies ~10 days. At the other extreme fish and some reptiles continue to grow throughout their lifetimes, surviving even longer than mammals.[495] Except for their germ cells nematodes, rotifers, and many insects have no dividing cells in their adult bodies. Their lifespan is presumably determined by the loss of cells through injury or death. In contrast, some simple animals, such as *Hydra*, other coelenterates, and flatworms maintain a pool of pluripotent stem cells that, except for accidental death, seems to make them immortal.[495]

Considerations such as these have helped move contemporary thinking toward an evolutionary view.[488,489,495–499] If the mortality of an animal in the wild (**extrinsic mortality**) is high it will evolve to have rapid development, good reproductive ability, and a short lifetime. If the extrinsic mortality is low the lifetime will be long. Such animals will require development of good protective functions including a highly developed brain.

Many factors must affect aging. It is generally agreed that one of these is the deleterious effects of free radicals derived from oxygen[500–503] (see pp. 1074, 1075). The lowered turnover rate of aging tissues may allow the damage to become lethal. According to this theory we might anticipate that free radical scavengers such as vitamin E could prolong life as might a restriction in food consumption.[504] For example, decreased fat intake might cut down on production of malondialdehyde (p. 1205) and lipid peroxides that may be especially damaging to cell membranes. The life of rodents can be prolonged substantially by a semi-starvation diet. Although there are uncertainties, a convincing case can be made for humans to keep food intake to a minimum and to eat foods rich in antioxidants and other nutrients.[491,504a]

For many years after techniques of cell culture had been developed it was commonly believed that cells in tissue culture were potentially immortal.[505] Challenging this idea, Medvedev[506] and others proposed that cells are internally programmed for a certain lifetime. This might explain why we have short-lived mammals as well as long-lived mammals. Support for the idea was supplied by Hayflick,[507,508] who observed that animal cells in culture have a limited potential for doubling. For example, normal human diploid embryonic fibroblasts grow in culture and double their number aproximately 50 ± 10 times. Regardless of cultural conditions, the cells die after this number of doublings. Cells taken from older humans undergo a smaller number of doublings before dying, as do those taken from shorter lived animals such as the mouse (14–28 doublings).[508] These experiments suggested that there is an internal program by which cells are scheduled to die from **replicative senescence**.[509–511] Malignant transformation overrides this program and transformed cells appear "immortal."[512] However, unlike fibroblasts some glial stem cells have been identified as possibly having unlimited proliferative capacity. These include cultured rat oligodendrocyte[509] and Schwann cells.[513] These results suggest that replicative senescence may not be inevitable.

As mentioned in Chapter 27 (p. 1568), "erosion" of the telomere ends on chromosomes is thought to be a major cause of cell senescence. Old cells have little or no telomerase. However, most human cancer cells, as well as those from immortalized cell cultures, do synthesize telomerase and maintain telomeres of adequate length.[514,515] Inhibition of telomerase activity in immortalized cells causes telomere shortening and cell death.[516] A second pathway for telomere maintenance is based upon homologous recombination.[517] Experimental elongation of telomeres extends the lifespan of cells in culture.[518] Furthermore, apparently healthy calves have been produced by nuclear transfer cloning from senescent fibroblast cells for which four or fewer cell doublings were expected to be possible. The cells of the cloned animals had the capacity for 90 or more cell doublings.[519]

What information about aging can we obtain by study of the "model" organisms *S. cerevisiae*, *C. elegans*, *D. melanogaster*, and the mouse? In every case a variety of mutations may shorten or lengthen the lifespan. In every case dietary energy restriction can lengthen life. Yeast cells grown on 0.5% glucose instead of 2%

glucose may undergo ~25% more cell doublings before a culture becomes senescent. However, mutant strains with a defect in the *SIR*2 gene have a shortened lifespan, which is not increased by caloric restriction.[520–522] *SIR*2 encodes Sir2p, an NAD^+-dependent histone deacetylase (p. 1626).[522,523] It is likely that caloric restriction causes the yeast to switch from anaerobic fermentation to oxidative metabolism. The resulting increase in the $[NAD^+] / [NADH]$ ratio activates Sir2p, thereby altering chromatin and silencing a group of genes. Mechanisms by which this shift in metabolism decreases deleterious mutations, even though respiration is increased, are probably complex.[523]

Nematodes (*C. elegans*) usually live about three weeks but the simultaneous presence of life-extending mutations in two different groups of genes lengthen the lifespan as much as fivefold.[524] One of the genes is the maternal-effect Clock gene *Clk*-1. It has been found to code for a hydroxylase involved in the final step of synthesis of ubiquinone-9 (p. 1429, Fig. 25-4).[525] The lifetime of wild-type nematodes is lengthened by ~60% by elimination of ubiquinone from the diet.[526] The second group of genes that affect longevity regulate an insulin-like signaling system.[527–529] In every case metabolism is slowed, an effect which may reduce the rate of harmful mutations. However, mutant animals may not be healthy. Some of these mutations induce formation of long-lived “dauer” larvae, providing a means for the larvae to survive for up to two months during periods of starvation.[527,530] Others affect sensory cilia.[531] Mutation of a cytosolic catalase gene *reduces* the lifespan.[530] The nematode’s lifetime is also affected negatively by its own germ cells, perhaps via a steroid hormone.[532] The heat shock proteins (p. 1630), by chaperoning newly synthesized proteins and preventing aggregation, also increase life span both in *C. elegans* and *Drosophila*.[532a,b]

Some mutants of *D. melanogaster* with extended lifespans have a defective insulin / IGF signaling pathway.[526,533] The **methuselah** mutant, whose lifespan is 35% greater than average, appears to involve a G-protein-coupled transmembrane receptor.[534] Mutation of an insulin receptor homolog extends lifespan, apparently by causing a juvenile hormone deficiency.[535] *Drosophila* lifespan is also lengthened by mutation of a transmembrane dicarboxylate transporter[536] or by overexpression of a protein repair carboxyl methyltransferase (p. 594).[537]

Some mutant mice have extended lifespans. The Ames dwarf mouse has a mutation in $p66^{shc}$, a cell-surface protein that contains both Src-homology and collagen-homology domains. It lives almost one-third longer than do wild-type mice.[538] Mice deficient in methionine sulfoxide reductase have a reduced lifespan[539] but fruit flies with overexpressed activity of the enzyme are more resistant than wild-type flies to oxidative damage.[540]

In humans 100 or more years in age some mitochondrial mutations are associated with good health and longevity.[541] Dietary factors doubtless play a role. For example, supplementation of rats’ diet with lipoic acid improved mitochondrial function and increased the metabolic rate of old animals.[542]

A number of genetic **progeroid diseases** result in premature aging.[543,544] Several of these arise from deficiencies in repair of DNA (Box 27-A). Among them are some types of cancer, and **Werner syndrome**, which arises from a defect in a 1432-residue protein with a central domain homologous to the RecQ family of DNA helicases (p. 1550).[545] Defects in other RecQ homologs cause **Hutchinson–Gilford progeria** as well as the Bloom syndrome (see Box 27-A).[546] Yet another DNA helicase, a subunit of transcription factor II (TFIIh, p. 1628), is defective in trichothiodystrophy (TTD, see Box 27-A).[547,548] Another gene which helps to prevent aging is *KLOTHO*. First identified in mice, it encodes a transmembrane protein that has sequence similarities to β-glucosidases.[549,550] Some mice with mutations in the tumor suppressor **p53** (Box 11-D) have enhanced resistance to tumors but age rapidly.[551,552]

Aging seems to be inevitably linked to an increase in the incidence of cancer. This uncontrolled growth of cells appears to be allowed by the stepwise accumulation of mutations that affect growth, differentiation, and survival.[553] Several aspects of cancer are discussed in other chapters of this book (see Box 11-D). However, the topic is so complex and research so active that it is hard to give even a thumbnail sketch of more recent discoveries.

Much effort is being dedicated to identifying the many signaling pathways that control growth, the mechanisms that cells employ to recognize problems in the control of growth, and the means by which cells can correct the problems or undergo apoptosis and avoid cancer.[553–558] Some of the complexity arises because of the large number of signaling pathways in which mutations may produce activated proto-onco-genes or faulty tumor suppressors. A large network of these suppressors is present in human cells.[553,559] Among the relevant signaling pathways are the following:

RAS – RAF – ERK (Fig. 11-12)[553–555]
p53[556–558]
the PtdIns 3-kinase – PKB/Akt pathway (Fig 11-9), which is opposed by PTEN[560,561]
EFG receptor (EFGR) signaling
Wnt-Catenin signaling[559,562]
E. Cadherin[559,563]

The importance of oncogenes and tumor suppressors has been demonstrated by conversion of human cells in culture into tumor cells in vitro.[564] Introduction of an activated *ras* gene, an SV40 viral protein that inhibits formation of both p53 and the Rb gene (Fig. 11-15), and an active telomerase gene sufficed. However, there is some doubt about the relevance of this work to human cancer.

Most cancerous cells have extra chromosomes. The karyotype (p. 1472) is rarely normal.[565] This and other evidence suggest that **genomic instability** may be the major cause of cancer.[566,567] In healthy cells stalled RNA polymerase is removed by transcription-coupled repair and lesions in DNA are either repaired (Chapter 27) or the cell undergoes apoptosis. Telomere dysfunction is also a factor.[567] The two breast cancer susceptibility genes *BRCA1* and *BRCA2* are apparently responsible for about half of all hereditary breast and ovarian cancers.[568,569] Protein BRCA1 is an 1863-residue nuclear protein, which is thought to function in transcription. However, recent evidence indicates that BRAC2 is directly involved in repair of double-strand breaks in DNA by homologous recombination.[569,570] Other data implicate the Neu-Ras pathway, proto-oncogenes *c-myc* and *Wnt-1*, and cyclin D1 in breast cancer.[571]

Yet another aspect of cancer is the **aberrant glycosylation** observed for many proteins.[572] The state of glycosylation of cell-surface proteins is one of many factors that affect metastasis, which is critical to growth of tumors.[573–575] The recognition that causes of cancer are numerous has led to a new large-scale project to identify as many cancer-associated mutations as possible within the entire human genome. One early success from this effort is identification of mutations in the gene *BRAF*, one of the three human *RAF* genes. These mutations are present in 15% of human malignant melanomas.[555,576]

G. Ecological Matters (Author's Personal Postscript)

The final section of this chapter deals with interactions among different species. As humans, beset by problems arising from our inability to communicate with other humans, we may feel that ecological relationships are relatively unimportant. However, any careful look at what can be regarded as an extension of metabolic cycles into the biosphere should convince us of the significance of this aspect of biochemistry.

Recall that the original development of eukaryotic creatures may have started with a symbiotic relationship between two prokaryotes and that symbiosis between algae and nonphotosynthetic organisms may have led to development of higher plants. Associations between species are still important today. For example, the bacteria in the protozoa of the digestive tract of ruminant animals are essential to production of meat. Our own bodies play host to bacteria, fungi, and other organisms with whom we have to try to maintain friendly relations. We depend upon antibiotics produced by bacteria or by fungi to fight our bacterial infections. Plants provide both essential nutrients and oxygen. Our environment has been created in large part by other living forms that coexist with us and which are subject to ecological checks and balances. It is therefore important that we learn more about the effects of one group of organisms on another and also about the effects of human activities on plants and animals of all degrees of complexity. This includes the poorly understood world of soil microorganisms. The consequences of environmental pollution, of depletion of atmospheric ozone or other alterations that affect the radiant energy reaching us, and of the availability to humans of excessive amounts of energy must all be considered. Just as a steady state within cells is often essential to the life of organisms, maintenance of a steady state in the chemical cycles of the biosphere may also be a necessity.

Biochemists and molecular biologists are being called upon to play an increasing role in medicine, agriculture, and industry. As such, they must be prepared to help in the making of decisions that may affect the future of life on earth. Biochemical approaches will be required to cope with many important problems. Among these are the long-term effects of the growing number of synthetic compounds in the environment, problems of antibiotic resistance, and effects of bioengineering of plants, fishes, and other organisms in the biosphere. Some of these scientific and ethical questions have been discussed in Chapter 26, and more are considered in the Study Questions that follow in this chapter.

Despite attempts to ignore it, we cannot avoid facing the war problem. The possibility of virtually total destruction of the more complex forms of life by genetic damage from radiation is real. That we have lived with nuclear weapons as long as we have is encouraging but continuing threats to use them as a last resort may bring eventual castastrophe. A race to put weapons into space might result in having computers decide to fight a war in which all people could be destroyed, but one computer might win! Perhaps biochemists, who understand the technical problems of radiation damage and mutation, have a special obligation to point out the hazard to others.

Just as threatening is the possibility of biochemical warfare, e.g., the use of artificial viruses. Biological weapons have been little used because of their lack of discrimination between friend and foe. However, our increasing knowledge of molecular biology makes possible insidious attacks on a population of unvaccinated persons. Since biochemical work does not require elaborate facilities, the development of biological weapons can be carried on by small groups in a clandestine manner. The recent assembly of a viable polio virus from oligonucleotides purchased from a commercial supplier emphasizes the ease with which virus warfare might be launched. Finding a way to protect ourselves may be more difficult.

Should we really worry about such matters? Since biochemistry is unable to ascribe any purpose to life, shouldn't we scientists stick to science? Science is amoral, isn't it? And besides, won't society do just what it wants to regardless of our opinions? Questions like these will always be with us, but most of the best scientists in the world seem to act with a great deal of responsibility. Not only do they want the pleasure and excitement of discovery and recognition for their work, but also they want a world for their children and grandchildren. They tend to feel compassion for other human beings. Many of them will give as a principal motivation for becoming biochemists the desire to contribute to the understanding of living things for the purpose of improving health, medical care, nutrition, etc. Most of them would not like to see the evolution of human beings ended through a disaster with nuclear or biological weapons or by irreversible pollution of land and sea. It will be a strange irony if we use our marvelous inquisitive, ingenous, inventive, and compassionate brains, the pinnacle of biological evolution, to destroy our environment and ourselves.

At a conference in Berkeley in 1971,[577] Joshua Lederberg, discoverer of genetic recombination in bacteria, talked about these matters. Lederberg asked if fairness and objectivity are possible outside the laboratory. He thought so. He pointed out that the nations of the world agreed to stop production of biological weapons and that genuine steps had been taken to decrease some of the hazards facing us. Nevertheless, progress is slow. Some insist on inspection for violation of agreements. But how can one inspect thoroughly enough? Lederberg suggested that the only possible form of control is now evolving. It must come from scientists themselves who must step out of their roles as "pure" scientists and accept the responsibility of preventing foolish uses of new biological discoveries. It may seem impossible that there could be a scientific community which could be counted on always to act in a responsible way, but it may be the only way that the human beings can survive for long on this planet. Lederberg believes it possible (and so do I).

If this book has helped to bring to the reader some awareness of the knowledge and power of molecular biology, I hope that these final words may lead the reader to heed the advice of Professor Lederberg. I sincerely hope that all the young people now studying biochemistry and modern biology will commit themselves to using the fantastic new knowledge available to us for the betterment of mankind and to proceeding with caution and responsibility as they move into positions of influence in the scientific community.

References

1. Gilbert, S. F., and Raunio, A. M., eds. (1997) *Embryology: Constructing the Organism*, Sinauer Assoc., Stamford, Connecticut
2. Wolpert, L., Beddington, R., Jessell, T., Lawrence, P., Meyerowitz, E., and Smith, J. (2002) *Principles of Development*, 2nd ed., Current Biology/Oxford Univ. Press, London
3. Brivanlou, A. H., and Darnell, J. E., Jr. (2002) *Science* **295**, 813–818
4. Britten, R. J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9372–9377
5. Jan, Y. N., and Jan, L. Y. (1998) *Nature (London)* **392**, 775–778
6. Wolpert, L. (1994) *Science* **266**, 571–572
7. Shulman, J. M., and Johnston, D. S. (1999) *Trends Biochem. Sci.* **24**, M60–M64
8. Meyerowitz, E. M. (1999) *Trends Biochem. Sci.* **24**, M65–M68
9. Nüsslein-Volhard, C. (1996) *Sci. Am.* **275**(Aug), 54–61
10. Gurdon, J. B., and Bourillot, P.-Y. (2001) *Nature (London)* **413**, 797–803
11. Tsonis, P. A. (1987) *Trends Biochem. Sci.* **12**, 249
12. Basler, K. (2000) *EMBO J.* **19**, 1169–1175

12a. Hochachka, P. W., and Somero, G. N. (2002) *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*, Oxford Univ. Press, New York

12b. Johnson, G. L., and Lapadat, R. (2002) *Science* **298**, 1911 – 1912

13. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) *Molecular Biology of the Cell*, 3rd ed., Garland, New York
14. Lewin, B. (2000) *Genes VII*, Oxford Univ. Press, New York

14a. Foss, M., McNally, F. J., Laurenson, P., and Rine, J. (1993) *Science* **262**, 1838 – 1844

14b. Arcangioli, B. (2000) *EMBO Reports* **1**, 145 – 150

15. Caldwell, G. A., Wang, S.-H., Xue, C.-B., Jiang, Y., Lu, H.-F., Naider, F., and Becker, J. M. (1994) *J. Biol. Chem.* **269**, 19817–19826
16. Tan, S., and Richmond, T. J. (1998) *Nature (London)* **391**, 660–666
17. Li, T., Stark, M. R., Johnson, A. D., and Wolberger, C. (1995) *Science* **270**, 262–269
18. Acosta-Serrano, A., Vassella, E., Liniger, M., Renggli, C. K., Brun, R., Roditi, I., and Englund, P. T. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1513–1518
19. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., and Brockdorff, N. (1996) *Nature (London)* **379**, 131–137
20. Park, Y., and Kuroda, M. I. (2001) *Science* **293**, 1083–1085
21. Chao, W., Huynh, K. D., Spencer, R. J., Davidow, L. S., and Lee, J. T. (2002) *Science* **295**, 345–347
22. Ferguson-Smith, A. C., and Surani, M. A. (2001) *Science* **293**, 1086–1089
23. Surani, M. A. (2001) *Nature (London)* **414**, 122–128
24. Reik, W., and Murrell, A. (2000) *Nature (London)* **405**, 408–409
25. Surani, M. A. (2002) *Nature (London)* **416**, 491–493
26. Lorincz, M. C., and Groudine, M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10034–10036
27. Reik, W., Dean, W., and Walter, J. (2001) *Science* **293**, 1089–1093
28. Inoue, K., Kohda, T., Lee, J., Ogonuki, N., Mochida, K., Noguchi, Y., Tanemura, K., Kaneko-Ishino, T., Ishino, F., and Ogura, A. (2002) *Science* **295**, 297
29. Burns, J. L., Jackson, D. A., and Hassan, A. B. (2001) *FASEB J.* **15**, 1694–1703
30. Jones, P. A., and Takai, D. (2001) *Science* **293**, 1068–1070
31. Rideout, W. M., III, Eggan, K., and Jaenisch, R. (2001) *Science* **293**, 1093–1098
32. Kang, Y.-K., Koo, D.-B., Park, J. S., Choi, Y.-H., Kim, H.-N., Chang, W.-K., Lee, K.-K., and Han, Y.-M. (2001) *J. Biol. Chem.* **276**, 39980–39984
33. Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E., and Reik, W. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13734–13738

References

34. Plowman, G. D., Sudarsanam, S., Bingham, J., Whyte, D., and Hunter, T. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13603–13610
35. Loomis, W. E. (1986) *Developmental Biology*, Macmillan, New York
36. Schrader, M., and Schulz-Knappe, P. (2001) *Trends in Biotechnology* **19**, S55–S60
37. Hernandez, N. (2001) *J. Biol. Chem.* **276**, 26733–26736
38. Hutvágner, G., McLachlan, J., Pasquinelli, A. E., Bálint, E., Tuschl, T., and Zamore, P. D. (2001) *Science* **293**, 834–838

38a. Storz, G. (2002) *Science* **296**, 1260 – 1262

38b. Lee, R. C., and Ambros, V. (2001) *Science* **294**, 862 – 864

38c. Caplen, N. J., Parrish, S., Imani, F., Fire, A., and Morgan, R. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9742 – 9747

38d. Ambros, V. (2001) *Science* **293**, 811– 813

38e. Dennis, C. (2002) *Nature (London)* **418**, 122 – 124

38f. Suprenant, K. A. (2002) *Biochemistry* **41**, 14447– 14454

39. Lagos–Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001) *Science* **294**, 853–858
40. Hu, R.-M., and 28 other authors. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9543–9548
41. Davidson, E. H., and 24 other authors. (2002) *Science* **295**, 1669–1678
42. Michelson, A. M. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 546–548
43. Yuh, C.-H., Bolouri, H., and Davidson, E. H. (1998) *Science* **279**, 1896–1902
44. Berman, B. P., Nibu, Y., Pfeiffer, B. D., Tomancak, P., Celniker, S. E., Levine, M., Rubin, G. M., and Eisen, M. B. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 757–762
45. Roush, W. (1996) *Science* **272**, 652–653
46. Veenstra, G. J. C., and Wolffe, A. P. (2001) *Trends Biochem. Sci.* **26**, 665–671
47. Richter, J. D., and Theurkauf, W. E. (2001) *Science* **293**, 60–62
48. Freeman, M. (2000) *Nature (London)* **408**, 313–319
49. Mannervik, M., Nibu, Y., Zhang, H., and Levine, M. (1999) *Science* **284**, 606–609
50. Musgrave, A., and van den Ende, H. (1987) *Trends Biochem. Sci.* **12**, 470–473
51. Misevic, G. N., and Burger, M. M. (1990) *J. Biol. Chem.* **265**, 20577–20584
52. Spillmann, D., Hård, K., Thomas-Oates, J., Vliegenthart, J. F. G., Misevic, G., Burger, M. M., and Finne, J. (1993) *J. Biol. Chem.* **268**, 13378–13387
53. Fernàndez-Busquets, X., Gerosa, D., Hess, D., and Burger, M. M. (1998) *J. Biol. Chem.* **273**, 29545–29553

53a. Yamada, K. M., and Clark, K. (2002) *Nature (London)* **419**, 790 – 791

53b. Boggon, T. J., Murray, J., Chappuis-Flament, S., Wong, E., Gumbiner, B. M., and Shapiro, L. (2002) *Science* **296**, 1308 – 1313

53c. Song, X., and Xie, T. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14813 – 14818

54. Edelman, G. M. (1984) *Sci. Am.* **250**(Apr), 118–129
55. Edelman, G. M. (1985) *Ann. Rev. Biochem.* **54**, 135–169
56. Kiss, J. Z., Wang, C., Olive, S., Rougon, G., Lang, J., Baetens, D., Harry, D., and Pralong, W. F. (1994) *EMBO J.* **13**, 5284–5292
57. Nelson, R. W., Bates, P. A., and Rutishauser, U. (1995) *J. Biol. Chem.* **270**, 17171–17179
58. Su, X.-D., Gastinel, L. N., Vaughn, D. E., Faye, I., Poon, P., and Bjorkman, P. J. (1998) *Science* **281**, 991–995

58a. van Meer, G. (2002) *Science* **296**, 855 – 857

58b. Shin, J.-S., and Abraham, S. N. (2001) *Science* **293**, 1447– 1448

58c. Nohturfft, A., and Losick, R. (2002) *Science* **296**, 857– 858

59. Collins, J. E., and Fleming, T. P. (1995) *Trends Biochem. Sci.* **20**, 307–312
60. Daniels, D. L., Spink, K. E., and Weis, W. I. (2001) *Trends Biochem. Sci.* **26**, 672–678
61. Wu, Q., and Maniatis, T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3124–3129

61a. Boggon, T. J., Murray, J., Chappuis-Flament, S., Wong, E., Gumbiner, B. M., and Shapiro, L. (2002) *Science* **296**, 1308–1313

62. Kollmar, R., Nakamura, S. K., Kappler, J. A., and Hudspeth, A. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10196–10201
63. Giancotti, F. G., and Ruoslahti, E. (1999) *Science* **285**, 1028–1032
64. Couzin, J. (2001) *Science* **293**, 1743–1746
65. Calderwood, D. A., Shattil, S. J., and Ginsberg, M. H. (2000) *J. Biol. Chem.* **275**, 22607–22610
66. Harris, E. S., McIntyre, T. M., Prescott, S. M., and Zimmerman, G. A. (2000) *J. Biol. Chem.* **275**, 23409–23412
67. Brakebusch, C., Grose, R., Quondamatteo, F., Ramirez, A., Jorcano, J. L., Pirro, A., Svensson, M., Herken, R., Sasaki, T., Timpl, R., Werner, S., and Fässler, R. (2000) *EMBO J.* **19**, 3990–4003
68. Friedlander, M., Theesfeld, C. L., Sugita, M., Fruttiger, M., Thomas, M. A., Chang, S., and Cheresh, D. A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9764–9769
69. Edwards, S. W. (1995) *Trends Biochem. Sci.* **20**, 362–367
70. Perez-Vilar, J., and Hill, R. L. (1997) *J. Biol. Chem.* **272**, 33410–33415
71. D'Souza, S. E., Ginsberg, M. H., Burke, T. A., Lam, S. C.-T., and Plow, E. F. (1988) *Science* **242**, 91–93

71a. Xiong, J.-P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S. L., and Arnaout, M. A. (2002) *Science* **296**, 151–155

72. Roth, S., McGuire, E. J., and Roseman, S. (1971) *J. Cell Biol.* **51**, 536–547
73. Perrimon, N., and Bernfield, M. (2000) *Nature (London)* **404**, 725–728
74. Zhuo, L., Yoneda, M., Zhao, M., Yingsung, W., Yoshida, N., Kitagawa, Y., Kawamura, K., Suzuki, T., and Kimata, K. (2001) *J. Biol. Chem.* **276**, 7693–7696
75. Varki, A. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4523–4525
76. Knox, J. P. (1995) *FASEB J.* **9**, 1004–1012
77. Rakic, P. (1999) *Nature (London)* **400**, 315–316
78. Horwitz, A. R., and Parsons, J. T. (1999) *Science* **286**, 1102–1103
79. Chicurel, M. (2002) *Science* **295**, 606–609
80. Condliffe, A. M., and Hawkins, P. T. (2000) *Nature (London)* **404**, 135, 137
81. Ng, T., Parsons, M., Hughes, W. E., Monypenny, J., Zicha, D., Gautreau, A., Arpin, M., Gschmeissner, S., Verveer, P. J., Bastiaens, P. I. H., and Parker, P. J. (2001) *EMBO J.* **20**, 2723–2741
82. D'Ari, R., and Bouloc, P. (1990) *Trends Biochem. Sci.* **15**, 191–194
83. Shapiro, L., and Losick, R. (1997) *Science* **276**, 712–718
84. Peifer, M., and Tepass, U. (2000) *Nature (London)* **403**, 611–612

84a. Shapiro, L., McAdams, H. H., and Losick, R. (2002) *Science* **298**, 1942 – 1946

85. Losick, R., and Shapiro, L. (1993) *Science* **262**, 1227–1228

85a. Bourne, H. R., and Weiner, O. (2002) *Nature (London)* **419**, 21

85b. Pellettieri, J., and Seydoux, G. (2002) *Science* **298**, 1946 – 1950

86. Dawid, I. B. (1994) *J. Biol. Chem.* **269**, 6259–6262
87. Rodriguez-Boulan, E., and Nelson, W. J. (1989) *Science* **245**, 718–725
88. St Johnston, D. (2001) *EMBO J.* **20**, 6169–6179
89. Wikramanayake, A. H., Huang, L., and Klein, W. H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9343–9348
90. Wallenfang, M. R., and Seydoux, G. (2000) *Nature (London)* **408**, 89–92

90a. Keller, R. (2002) *Science* **298**, 1950 – 1954

91. Martelly, I., and Franquinet, R. (1984) *Trends Biochem. Sci.* **9**, 468–471
92. Slack, J. M. W. (1987) *Trends Biochem. Sci.* **12**, 200–204
93. Patel, N. H., and Lall, S. (2002) *Nature (London)* **415**, 748–749
94. Houchmandzadeh, B., Wieschaus, E., and Leibler, S. (2002) *Nature (London)* **415**, 798–802
95. Wimmer, E. A., Carleton, A., Harjes, P., Turner, T., and Desplan, C. (2000) *Science* **287**, 2476–2479
96. Vincent, S., and Perrimon, N. (2001) *Nature (London)* **411**, 533–536
97. Sampath, K., Rubinstein, A. L., Cheng, A. M. S., Liang, J. O., Fekany, K., Solnica-Krezel, L., Korzh, V., Halpern, M. E., and Wright, C. V. E. (1998) *Nature (London)* **395**, 185–189
98. Chen, Y., and Schier, A. F. (2001) *Nature (London)* **411**, 607–610
99. Pasternak, C. A. (1970) *Biochemistry of Differentiation*, Wiley (Interscience), New York
100. Vogelstein, B., Alberts, B., and Shine, K. (2002) *Science* **295**, 1237
101. Shin, T., Kraemer, D., Pryor, J., Liu, L., Rugila, J., Howe, L., Buck, S., Murphy, K., Lyons, L., and Westusin, M. (2002) *Nature (London)* **415**, 859

101a. Rossant, J. (2002) *Nature (London)* **415**, 967, 969

102. Lai, L., Kolber-Simonds, D., Park, K.-W., Cheong, H.-T., Greenstein, J. L., Im, G.-S., Samuel, M., Bonk, A., Rieke, A., Day, B. N., Murphy, C. N., Carter, D. B., Hawley, R. J., and Prather, R. S. (2002) *Science* **295**, 1089–1092
103. Gurdon, J. B. (1974) *Nature (London)* **248**, 772–776
104. De Robertis, E. M., and Gurdon, J. B. (1979) *Sci. Am.* **241**(Dec), 74–82
105. Anderson, G. B., and Seidel, G. E. (1998) *Science* **280**, 1400–1401
106. Aldhous, P. (2000) *Nature (London)* **405**, 610–612
107. Gurdon, J. B., and Colman, A. (1999) *Nature (London)* **402**, 743–746

107a. Hochedlinger, K., and Jaenisch, R. (2002) *Nature (London)* **415**, 1035 – 1038

107b. Wilmut, I., Beaujean, N., de Sousa, P. A., Dinnyes, A., King, T. J., Paterson, L. A., Wells, D. N., and Young, L. E. (2002) *Nature (London)* **419**, 583 – 586

108. Kubota, C., Yamakurchi, H., Todoroki, J., Mizoshita, K., Tabara, N., Barber, M., and Yang, X. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 990–995
109. Pennisi, E. (2001) *Science* **293**, 1064–1067
110. De Bie, I., Savaria, D., Roebroek, A. J. M., Day, R., Lazure, C., Van de Ven, W. J. M., and Seidah, N. G. (1995) *J. Biol. Chem.* **270**, 1020–1028
111. Judson, H., Hayward, B. E., Sheridan, E., and Bonthron, D. T. (2002) *Nature (London)* **416**, 539–542
112. Golde, D. W. (1991) *Sci. Am.* **265**(Dec), 86–93
113. Watt, F. M., and Hogan, B. L. M. (2000) *Science* **287**, 1427–1430
114. Spradling, A., Drummond-Barbosa, D., and Kai, T. (2001) *Nature (London)* **414**, 98–104
115. Kiger, A. A., Jones, D. L., Schulz, C., Rogers, M. B., and Fuller, M. T. (2001) *Science* **294**, 2542–2545
116. McKay, R. (2000) *Nature (London)* **406**, 361–364

References

117. Donovan, P. J., and Gearhart, J. (2001) *Nature (London)* **414**, 92–97
117a. Wakayama, T., Tabar, V., Rodriguez, I., Perry, A. C. F., Studer, L., and Mombaerts, P. (2001) *Science* **292**, 740 – 743
118. Ahmed, S., and Hodgkin, J. (2000) *Nature (London)* **403**, 159–164
118a. Wurmser, A. E., and Gage, F. H. (2002) *Nature (London)* **416**, 485 – 487
118b. Blau, H. M. (2002) *Nature (London)* **419**, 437
119. Vogel, G. (1999) *Science* **283**, 1432–1434
119a. Ivanova, N. B., Dimos, J. T., Schaniel, C., Hackney, J. A., Moore, K. A., and Lemischka, I. R. (2002) *Science* **298**, 601– 604
119b. Zhang, Z., Zhang, R., Joachimiak, A., Schlessinger, J., and Kong, X.-P. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7732 – 7737
120. Vogel, G. (2001) *Science* **292**, 1820–1822
120a. Nishimura, E. K., Jordan, S. A., Oshima, H., Yoshida, H., Osawa, M., Moriyama, M. Jackson, I. J., Barrandon, Y., Miyachi, Y., and Nishikawa, S.-i. (2002) *Nature (London)* **416**, 854 – 860
121. Colter, D. C., Sekiya, I., and Prockop, D. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7841–7845
121a. Jiang, Y., Jahagirdar, B. N., Reinhard, R. L., Schwartz, R. E., Keene, C. D., Ortiz-Gonzalez, X. R., Reyes, M., Lenvik, T., Lund, T., Blackstad, M., Du, J., Aldrich, S., Lisberg, A., Low, W. C., Largaespada, D. A., and Verfaillie, C. M. (2002) *Nature (London)* **418**, 41–49
122. van der Kooy, D., and Weiss, S. (2000) *Science* **287**, 1439–1441
123. Slack, J. M. W. (2000) *Science* **287**, 1431–1433
124. Ghazizadeh, S., and Taichman, L. B. (2001) *EMBO J.* **20**, 1215–1222
125. Lavker, R. M., and Sun, T.-T. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13473–13475
126. Weigel, D., and Jürgens, G. (2002) *Nature (London)* **415**, 751–754
127. Langer, R., and Vacanti, J. P. (1993) *Science* **260**, 920–926
128. Solter, D., and Gearhart, J. (1999) *Science* **283**, 1468–1470
128a. Lumelsky, N., Blondel, O., Laeng, P., Velasco, I., Ravin, R., and McKay, R. (2001) *Science* **292**, 1389 – 1394
129. Langer, R., and Vacanti, J. P. (1995) *Sci. Am.* **273**(Sep), 130–133
130. Ferber, D. (1999) *Science* **284**, 422–425
131. Aldhous, P. (2001) *Nature (London)* **410**, 622–625
132. McLaren, A. (2000) *Science* **288**, 1775–1780
133. Lovell-Badge, R. (2001) *Nature (London)* **414**, 88–91
134. Weissman, I. L. (2000) *Science* **287**, 1442–1446
135. Jaenisch, R., and Wilmut, I. (2001) *Science* **291**, 2552
136. Holliday, R., and Pugh, J. E. (1975) *Science* **187**, 226–232
137. Thomassin, H., Flavin, M., Espinás, M.-L., and Grange, T. (2001) *EMBO J.* **20**, 1974–1983
138. Clerc, P., and Avner, P. (2000) *Science* **290**, 1518–1519
139. Georgatsos, J. G. (1995) *Nature (London)* **375**, 100
140. Raff, M. (1998) *Nature (London)* **396**, 119–122
141. Aravind, L., Dixit, V. M., and Koonin, E. V. (2001) *Science* **291**, 1279–1284
142. Ashkenazi, A., and Dixit, V. M. (1998) *Science* **281**, 1305–1308
143. White, K., Grether, M. E., Abrams, J. M., Young, L., Farrell, K., and Steller, H. (1994) *Science* **264**, 677–683
144. Yuan, J., and Yankner, B. A. (2000) *Nature (London)* **407**, 802–809
145. Kim, Y.-M., Talanian, R. V., and Billiar, T. R. (1997) *J. Biol. Chem.* **272**, 31138–31148
146. Bonfoco, E., Krainc, D., Ankarcrona, M., Nicotera, P., and Lipton, S. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7162–7166
146a. Green, D. R., and Beere, H. M. (2002) *Nature (London)* **405**, 28 – 29
146b. Savill, J., and Fadok, V. (2000) *Nature (London)* **407**, 784 – 788
147. Klionsky, D. J., and Emr, S. D. (2000) *Science* **290**, 1717–1721
147a. Xia, X. G., Harding, T., Weller, M., Bieneman, A., Uney, J. B., and Schulz, J. B. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10433 – 10438
147b. Raff, M. C., Whitmore, A. V., and Finn, J. T. (2002) *Science* **296**, 868 – 871
147c. Jones, P. A. (2001) *Nature (London)* **409**, 141–144
148. Meier, P., Finch, A., and Evan, G. (2000) *Nature (London)* **407**, 796–801
149. Härtel, U., Eckel, E., Koch, J., Fuchs, G., Linder, D., and Buckel, W. (1993) *Arch Microbiol* **159**, 174–181
150. Kumar, S. (1995) *Trends Biochem. Sci.* **20**, 198–202
151. Haecker, G., and Vaux, D. L. (1994) *Trends Biochem. Sci.* **19**, 99–100
152. Wolf, B. B., and Green, D. R. (1999) *J. Biol. Chem.* **274**, 20049–20052
152a. Adams, J. M., and Cory, S. (2001) *Trends Biochem. Sci.* **26**, 61– 66
152b. Ruoslahti, E., and Reed, J. (1999) *Nature (London)* **397**, 479 – 480
153. Qin, H., Srinivasula, S. M., Wu, G., Fernandes-Alnemri, T., Alnemri, E. S., and Shi, Y. (1999) *Nature (London)* **399**, 549–557
154. Slee, E. A., Adrain, C., and Martin, S. J. (2001) *J. Biol. Chem.* **276**, 7320–7326
155. Earnshaw, W. C., Martins, L. M., and Kaufmann, S. H. (1999) *Ann. Rev. Biochem.* **68**, 383–424
156. Kumar, S., and Colussi, P. A. (1999) *Trends Biochem. Sci.* **24**, 1–4
157. Liang, H., and Fesik, S. W. (1997) *J. Mol. Biol.* **274**, 291–302
158. Chu, Z.-L., Pio, F., Xie, Z., Welsh, K., Krajewska, M., Krajewski, S., Godzik, A., and Reed, J. C. (2001) *J. Biol. Chem.* **276**, 9239–9245
159. Hengartner, M. O., Ellis, R. E., and Horvitz, H. R. (1992) *Nature (London)* **356**, 494–499
159a. Chen, G., and Goeddel, D. V. (2002) *Science* **296**, 1634 – 1635
160. Aravind, L., Dixit, V. M., and Koonin, E. V. (1999) *Trends Biochem. Sci.* **24**, 47–53
161. Walczak, H., and Sprick, M. R. (2001) *Trends Biochem. Sci.* **26**, 452–453
162. Jeong, E.-J., Bang, S., Lee, T. H., Park, Y. I., Sim, W.-S., and Kim, K.-S. (1999) *J. Biol. Chem.* **274**, 16337–16342
163. Weber, C. H., and Vincenz, C. (2001) *Trends Biochem. Sci.* **26**, 475–481
164. Nagata, S., and Golstein, P. (1995) *Science* **267**, 1449–1456
164a. Krammer, P. H. (2000) *Nature (London)* **407**, 789 – 795
165. Imai, Y., Kimura, T., Murakami, A., Yajima, N., Sakamaki, K., and Yonehara, S. (1999) *Nature (London)* **398**, 777–785
166. Enari, M., Sakahira, H., Yokoyama, H., Okawa, K., Iwamatsu, A., and Nagata, S. (1998) *Nature (London)* **391**, 43–50
167. Hengartner, M. O. (2000) *Nature (London)* **407**, 770–776
167a. Lassus, P., Opitz-Araya, X., and Lazebnik, Y. (2002) *Science* **297**, 1352 – 1354
168. Liu, X., Kim, C. N., Yang, J., Jemmerson, R., and Wang, X. (1996) *Cell* **86**, 147–157
169. Mootha, V. K., Wei, M. C., Buttle, K. F., Scorrano, L., Panoutsakopoulou, V., Mannella, C. A., and Korsmeyer, S. J. (2001) *EMBO J.* **20**, 661–671
170. Bernardi, P., Petronilli, V., Di Lisa, F., and Forte, M. (2001) *Trends Biochem. Sci.* **26**, 112–117
171. Adrain, C., and Martin, S. J. (2001) *Trends Biochem. Sci.* **26**, 390–397
171a. Marsden, V. S., O'Conner, L., O'Reilly, L. A., Silke, J., Metcalf, D., Ekert, P. G., Huang, D. C. S., Cecconi, F., Kuida, K., Tomaselll, K. J., Roy, S., Nicholson, D. W., Vaux, D. L., Bouillet, P., Adams, J. M., and Strasser, A. (2002) *Nature (London)* **419**, 634 – 637
172. Finkel, E. (2001) *Science* **292**, 624–626
173. Qin, Z.-H., Wang, Y., Kikly, K. K., Sapp, E., Kegel, K. B., Aronin, N., and DiFiglia, M. (2001) *J. Biol. Chem.* **276**, 8079–8086
174. Adams, J. M., and Cory, S. (1998) *Science* **281**, 1322–1326
175. Pawlowski, J., and Kraft, A. S. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 529–531
176. Wei, M. C., Zong, W.-X., Cheng, E. H.-Y., Lindsten, T., Panoutsakopoulou, V., Ross, A. J., Roth, K. A., MacGregor, G. R., Thompson, C. B., and Korsmeyer, S. J. (2001) *Science* **292**, 727–730
177. Pozniak, C. D., Radinovic, S., Yang, A., McKeon, F., Kaplan, D. R., and Miller, F. D. (2000) *Science* **289**, 304–306
178. Gross, A., Yin, X.-M., Wang, K., Wei, M. C., Jockel, J., Milliman, C., Erdjument-Bromage, H., Tempst, P., and Korsmeyer, S. J. (1999) *J. Biol. Chem.* **274**, 1156–1163
179. Mehmet, H. (2000) *Nature (London)* **403**, 29–30
180. Podack, E. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8312–8314
181. Alimonti, J. B., Shi, L., Baijal, P. K., and Greenberg, A. H. (2001) *J. Biol. Chem.* **276**, 6974–6982
182. Pinkoski, M. J., Waterhouse, N. J., Heibein, J. A., Wolf, B. B., Kuwana, T., Goldstein, J. C., Newmeyer, D. D., Bleackley, R. C., and Green, D. R. (2001) *J. Biol. Chem.* **276**, 12060–12067
183. Schauer, A., Ranes, M., Santamaria, R., Guijarro, J., Lawlor, E., Mendez, C., Chater, K., and Losick, R. (1988) *Science* **240**, 768–772
184. Youngman, P., Zuber, P., Perkins, J. B., Sandman, K., Igo, M., and Losick, R. (1985) *Science* **229**, 285–291
185. Wireman, J. W., and Dworkin, M. (1975) *Science* **189**, 516–523
186. Sossong, T. M., Jr., Brigham-Burke, M. R., Hensley, P., and Pearce, K. H., Jr. (1999) *Biochemistry* **38**, 14843–14850
187. Mosyak, L., Zhang, Y., Glasfeld, E., Haney, S., Stahl, M., Seehra, J., and Somers, W. S. (2000) *EMBO J.* **19**, 3179–3191
188. Romberg, L., Simon, M., and Erickson, H. P. (2001) *J. Biol. Chem.* **276**, 11743–11753
189. van den Ent, F., and Löwe, J. (2000) *EMBO J.* **19**, 5300–5307
190. Cordell, S. C., Anderson, R. E., and Löwe, J. (2001) *EMBO J.* **20**, 2454–2461
191. Hayashi, I., Oyama, T., and Morikawa, K. (2001) *EMBO J.* **20**, 1819–1828
192. Jacobs, C., and Shapiro, L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5891–5893
193. RayChaudhuri, D., Gordon, G. S., and Wright, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1332–1334
194. Nishimura, A. (1998) *Trends Biochem. Sci.* **23**, 157–159
195. Domian, I. J., Reisenauer, A., and Shapiro, L. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6648–6653
195a. Ouimet, M.-C., and Marczynski, G. T. (2000) *J. Mol. Biol.* **302**, 761–775
195b. Laub, M. T., Chen, S. L., Shapiro, L., and McAdams, H. H. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4632–4637
196. Jenal, U., and Shapiro, L. (1996) *EMBO J.* **15**, 2393–2406

References

197. Brun, Y. V., Marczynski, G., and Shapiro, L. (1994) *Ann. Rev. Biochem.* **63**, 419–450
198. Hecht, G. B., Lane, T., Ohta, N., Sommer, J. M., and Newton, A. (1995) *EMBO J.* **14**, 3915–3924
199. Quon, K. C., Yang, B., Domian, I. J., Shapiro, L., and Marczynski, G. T. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 120–125
200. Wortinger, M., Sackett, M. J., and Brun, Y. V. (2000) *EMBO J.* **19**, 4503–4512
201. Laub, M. T., McAdams, H. H., Feldblyum, T., Fraser, C. M., and Shapiro, L. (2000) *Science* **290**, 2144–2148
202. Jenal, U., and Fuchs, T. (1998) *EMBO J.* **17**, 5658–5669
203. Ponnuraj, K., Rowland, S., Nessi, C., Setlow, P., and Jedrzejas, M. J. (2000) *J. Mol. Biol.* **300**, 1–10
204. Sharp, M. D., and Pogliano, K. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14553–14558
205. Parkes, R. J. (2000) *Nature (London)* **407**, 844–845
206. Arigoni, F., Pogliano, K., Webb, C. D., Stragier, P., and Losick, R. (1995) *Science* **270**, 637–640
207. Lucet, I., Feucht, A., Yudkin, M. D., and Errington, J. (2000) *EMBO J.* **19**, 1467–1475
208. Ducros, V. M.-A., Lewis, R. J., Verma, C. S., Dodson, E. J., Leonard, G., Turkenburg, J. P., Murshudov, G. N., Wilkinson, A. J., and Brannigan, J. A. (2001) *J. Mol. Biol.* **306**, 759–771
209. Shazand, K., Frandsen, N., and Stragier, P. (1995) *EMBO J.* **14**, 1439–1445
210. Lewis, R. J., Brannigan, J. A., Muchová, K., Barák, I., and Wilkinson, A. J. (1999) *J. Mol. Biol.* **294**, 9–15
211. Bath, J., Wu, L. J., Errington, J., and Wang, J. C. (2000) *Science* **290**, 995–997
212. Gould, G. W., and Dring, G. J. (1975) *Nature (London)* **258**, 402–405
213. Dong, Y.-H., Xu, J.-L., Li, X.-Z., and Zhang, L.-H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3526–3531

213a. Zhang, R.-g, Pappas, T., Brace, J. L., Miller, P. C., Oulmassov, T., Molyneaux, J. M., Anderson, J. C., Bashkin, J. K., Winans, S. C., and Joachimiak, A. (2002) *Nature (London)* **417**, 917–974

214. Fuqua, C., and Greenberg, E. P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6571–6572
215. Daniels, R., De Vos, D. E., Desair, J., Raedschelders, G., Luyten, E., Rosemeyer, V., Verreth, C., Schoeters, E., Vanderleyden, J., and Michiels, J. (2002) *J. Biol. Chem.* **277**, 462–468
216. Chen, X., Schauder, S., Potier, N., Van Dorsselaer, A., Pelczer, I., Bassler, B. L., and Hughson, F. M. (2002) *Nature (London)* **415**, 545–549
217. Yarmolinsky, M. B. (1995) *Science* **267**, 836–837
218. Hoffman, M. (1992) *Science* **255**, 1510–1511
219. Magee, P. T. (1997) *Science* **277**, 52–53
220. Herskowitz, I. (1992) *Nature (London)* **357**, 190–191
221. Cabib, E., Roh, D.-H., Schmidt, M., Crotti, L. B., and Varma, A. (2001) *J. Biol. Chem.* **276**, 19679–19682
222. Utzig, S., Fankhauser, C., and Simanis, V. (2000) *J. Mol. Biol.* **302**, 751–759
223. Desautels, M., Den Haese, J. P., Slupsky, C. M., McIntosh, L. P., and Hemmingsen, S. M. (2001) *J. Biol. Chem.* **276**, 5932–5942

223a. Pelham, R. J., Jr., and Chang, F. (2002) *Nature (London)* **419**, 82–86

224. Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O., and Herskowitz, I. (1998) *Science* **282**, 699–705
225. Obara, T., Nakafuku, M., Yamamoto, M., and Kaziro, Y. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5877–5881
226. Shimomura, O., Suthers, H. L. B., and Bonner, J. T. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7376–7379
227. Rutherford, C. L., Taylor, R. D., Merkle, R. K., and Frame, L. T. (1982) *Trends Biochem. Sci.* **7**, 108–111
228. Chiew, Y. Y., Reimers, J. M., and Wright, B. E. (1985) *J. Biol. Chem.* **260**, 15325–15331
229. Rutherford, C. L., and Brown, S. S. (1983) *Biochemistry* **22**, 1251–1258
230. Morrissey, J. H. (1983) *Nature (London)* **303**, 203–204
231. Siu, C.-H., Lam, T. Y., and Choi, A. H. C. (1985) *J. Biol. Chem.* **260**, 16030–16036
232. Berger, E. A., and Armant, D. R. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2162–2166
233. O'Day, D. H., and Lewis, K. E. (1975) *Nature (London)* **254**, 431–432
234. Schaller, H. C., and Bodenmüller, H. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7000–7004
235. Bosch, T. C. G., and David, C. N. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9478–9482
236. Marx, J. L. (1984) *Science* **225**, 40–42
237. Kenyon, C. J. (1983) *Trends Biochem. Sci.* **8**, 349–351
238. Wood, W. B., ed. (1988) *The Nematode Caenorhabditis elegans*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
239. Kenyon, C. (1988) *Science* **240**, 1448–1453
240. Roberts, L. (1990) *Science* **248**, 1310–1313
241. Riddle, D. L., Swanson, M. M., and Albert, P. S. (1981) *Nature (London)* **290**, 668–671
242. Wuethrich, B. (1998) *Science* **281**, 1980–1982
243. Zeyl, C., and Bell, G. (1997) *Nature (London)* **388**, 465–468

243a. Hunt, P. A., and Hassold, T. J. (2002) *Science* **296**, 2181–2183

244. Weeks, A. R., Marec, F., and Breeuwer, J. A. J. (2001) *Science* **292**, 2479–2482
245. Parkhurst, S. M., and Meneely, P. M. (1994) *Science* **264**, 924–932
246. Williams, N. (1995) *Science* **269**, 1826–1827
247. Lee, A. L., Volkman, B. F., Robertson, S. A., Rudner, D. Z., Barbash, D. A., Cline, T. W., Kanaar, R., Rio, D. C., and Wemmer, D. E. (1997) *Biochemistry* **36**, 14306–14317
248. Scott, M. J., Pan, L. L., Cleland, S. B., Knox, A. L., and Heinrich, J. (2000) *EMBO J.* **19**, 144–155
249. Kuroda, M. I., and Kelley, R. L. (1999) *Science* **284**, 1787–1788
250. Carmi, I., Kopczynski, J. B., and Meyer, B. J. (1998) *Nature (London)* **396**, 168–173
251. Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A.-M., Lovell-Badge, R., and Goodfellow, P. N. (1990) *Nature (London)* **346**, 240–244
252. Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1996) *Trends Biochem. Sci.* **21**, 302–308
253. Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) *Cell* **81**, 705–714
254. Burgoyne, P. S. (1986) *Nature (London)* **319**, 258–259
255. Ellis, N. A., Goodfellow, P. J., Pym, B., Smith, M., Palmer, M., Frischauf, A.-M., and Goodfellow, P. N. (1989) *Nature (London)* **337**, 81–84
256. Thomas, J. O., and Travers, A. A. (2001) *Trends Biochem. Sci.* **26**, 167–174
257. Bustin, M. (2001) *Trends Biochem. Sci.* **26**, 431–437
258. Haqq, C. M., King, C.-Y., Ukiyama, E., Falsafi, S., Haqq, T. N., Donahoe, P. K., and Weiss, M. A. (1994) *Science* **266**, 1494–1499
259. Benevides, J. M., Chan, G., Lu, X.-J., Olson, W. K., Weiss, M. A., and Thomas, G. J., Jr. (2000) *Biochemistry* **39**, 537–547
260. Ohe, K., Lalli, E., and Sassone-Corsi, P. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1146–1151
261. Vainio, S., Heikkilä, M., Kispert, A., Chin, N., and McMahon, A. P. (1999) *Nature (London)* **397**, 405–409
262. Swain, A., Narvaez, V., Burgoyne, P., Camerino, G., and Lovell-Badge, R. (1998) *Nature (London)* **391**, 761–767
263. Hurst, L. D. (2001) *Nature (London)* **411**, 149–150
264. Ariel, M., McCarrey, J., and Cedar, H. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2317–2321
265. Iannello, R. C., Gould, J. A., Young, J. C., Giudice, A., Medcalf, R., and Kola, I. (2000) *J. Biol. Chem.* **275**, 19603–19608
266. Akama, T. O., Nakagawa, H., Sugihara, K., Narisawa, S., Ohyama, C., Nishimura, S.-I., O'Brien, D. A., Moremen, K. W., Millán, J. L., and Fukuda, M. N. (2002) *Science* **295**, 124–127
267. Tulina, N., and Matunis, E. (2001) *Science* **294**, 2546–2549

267a. Sassone-Corsi, P. (2002) *Science* **296**, 2176–2178

267b. Kashiwabara, S.-i, Noguchi, J., Zhuang, T., Ohmura, K., Honda, A., Sugiura, S., Miyamoto, K., Takahashi, S., Inoue, K., Ogura, A., and Baba, T. (2002) *Science* **298**, 1999–2002

267c. Càceres, C., Giménez-Bonafé, P., Ribes, E., Wouters-Tyrou, D., Martinage, A., Kouach, M., Sautière, P., Muller, S., Palau, J., Subirana, J. A., Cornudella, L., and Chiva, M. (1999) *J. Biol. Chem.* **274**, 649–656

267d. Kundu, T. K., and Rao, M. R. S. (1996) *Biochemistry* **35**, 15626–15632

267e. Honke, K., Hirahara, Y., Dupree, J., Suzuki, K., Popko, B., Fukushima, K., Fukushima, J., Nagasawa, T., Yoshida, N., Wada, Y., and Taniguchi, N. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4227–4232

268. Ursini, F., Helm, S., Kiess, M., Maiorino, M., Roveri, A., Wissing, J., and Flohé, L. (1999) *Science* **285**, 1393–1396
269. Pfeifer, H., Conrad, M., Roethlein, D., Kyriakopoulos, A., Brielmeier, M., Bornkamm, G. W., and Behne, D. (2001) *FASEB J.* **15**, 1236–1238
270. Shao, X., Tarnasky, H. A., Schalles, U., Oko, R., and van der Hoorn, F. A. (1997) *J. Biol. Chem.* **272**, 6105–6113
271. Haaf, A., Butler, P. J. G., Kent, H. M., Fearnley, I. M., Roberts, T. M., Neuhaus, D., and Stewart, M. (1996) *J. Mol. Biol.* **260**, 251–260
272. Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W., III, Yanagimachi, R., and Jaenisch, R. (2000) *Science* **290**, 1578–1581
273. O'Neill, L. P., Keohane, A. M., Lavender, J. S., McCabe, V., Heard, E., Avner, P., Brockdorff, N., and Turner, B. M. (1999) *EMBO J.* **18**, 2897–2907
274. Gilbert, S. L., and Sharp, P. A. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13825–13830
275. Willard, H. F., and Carrel, L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10025–10027
276. Disteche, C. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14180–14182
277. Percec, I., and Bartolomei, M. S. (2002) *Science* **295**, 287–288
278. Tamaru, H., and Selker, E. U. (2001) *Nature (London)* **414**, 277–283
279. Rice, J. C., and Allis, C. D. (2001) *Nature (London)* **414**, 258–259
280. Melton, D. A. (1991) *Science* **252**, 234–241
281. Miller, M. A., Nguyen, V. Q., Lee, M.-H., Kosinski, M., Schedl, T., Caprioli, R. M., and Greenstein, D. (2001) *Science* **291**, 2144–2147

References

282. Ferrell, J. E., Jr., and Machleder, E. M. (1998) *Science* **280**, 895 – 898
283. Nothias, J.-Y., Majumder, S., Kaneko, K. J., and DePamphilis, M. L. (1995) *J. Biol. Chem.* **270**, 22077 – 22080
283a. Mehlmann, L. M., Jones, T. L. Z., and Jaffe, L. A. (2002) *Science* **297**, 1343 – 1345
284. Kuge, H., and Richter, J. D. (1995) *EMBO J.* **14**, 6301 – 6310
284a. Matzuk, M. M., Burns, K. H., Viveiros, M. M., and Eppig, J. J. (2002) *Science* **296**, 2178 – 2180
285. Cooley, L., and Theurkauf, W. E. (1994) *Science* **266**, 590 – 596
286. Godt, D., and Tepass, U. (1998) *Nature (London)* **395**, 387 – 391
287. Cau, J., Faure, S., Vigneron, S., Labbé, J. C., Delsert, C., and Morin, N. (2000) *J. Biol. Chem.* **275**, 2367 – 2375
288. Wassarman, P. M. (1988) *Sci. Am.* **259**(Dec), 78 – 84
289. Vacquier, V. D. (1998) *Science* **281**, 1995 – 1998
290. Babcock, D. F., Bosma, M. M., Battaglia, D. E., and Darszon, A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6001 – 6005
291. Olson, J. H., Xiang, X., Ziegert, T., Kittelson, A., Rawls, A., Bieber, A. L., and Chandler, D. E. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11205 – 11210
292. Ohlendieck, K., and Lennarz, W. J. (1995) *Trends Biochem. Sci.* **20**, 29 – 32
293. Vilela-Silva, A.-C. E. S., Castro, M. O., Valente, A.-P., Biermann, C. H., and Mourao, P. A. S. (2002) *J. Biol. Chem.* **277**, 379 – 387
294. Litscher, E. S., Juntunen, K., Seppo, A., Penttilä, L., Niemelä, R., Renkonen, O., and Wassarman, P. M. (1995) *Biochemistry* **34**, 4662 – 4669
295. Johnston, D. S., Wright, W. W., Shaper, J. H., Hokke, C. H., Van den Eijnden, D. H., and Joziasse, D. H. (1998) *J. Biol. Chem.* **273**, 1888 – 1895
296. Varela, P. F., Romero, A., Sanz, L., Romao, M. J., Töpfer-Petersen, E., and Calvete, J. J. (1997) *J. Mol. Biol.* **274**, 635 – 649
297. Chen, M. S., Tung, K. S. K., Coonrod, S. A., Takahashi, Y., Bigler, D., Chang, A., Yamashita, Y., Kincade, P. W., Herr, J. C., and White, J. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11830 – 11835
298. Miyado, K., Yamada, G., Yamada, S., Hasuwa, H., Nakamura, Y., Ryu, F., Suzuki, K., Kosai, K., Inoue, K., Ogura, A., Okabe, M., and Mekada, E. (2000) *Science* **287**, 321 – 324
298a. Primakoff, P., and Myles, D. G. (2002) *Science* **296**, 2183 – 2185
299. Sherman, M. B., Jakana, J., Sun, S., Matsudaira, P., Chiu, W., and Schmid, M. F. (1999) *J. Mol. Biol.* **294**, 139 – 149
300. Yamagata, K., Murayama, K., Okabe, M., Toshimori, K., Nakanishi, T., Kashiwabara, S.-i, and Baba, T. (1998) *J. Biol. Chem.* **273**, 10470 – 10474
301. Mengerink, K. J., Moy, G. W., and Vacquier, V. D. (2002) *J. Biol. Chem.* **277**, 943 – 948
302. Kuo, R. C., Baxter, G. T., Thompson, S. H., Stricker, S. A., Patton, C., Bonaventura, J., and Epel, D. (2000) *Nature (London)* **406**, 633 – 636
302a. Glaser, R. W., Grüne, M., Wandelt, C., and Ulrich, A. S. (1999) *Biochemistry* **38**, 2560 – 2569
303. Parrington, J., Swann, K., Shevchenko, V. I., Sesay, A. K., and Lai, F. A. (1996) *Nature (London)* **379**, 364 – 368
304. Fontanilla, R. A., and Nuccitelli, R. (1998) *Biophys. J.* **75**, 2079 – 2087
305. Shapiro, B. M. (1991) *Science* **252**, 533 – 536
306. Lawinger, P., Rastelli, L., Zhao, Z., and Majumder, S. (1999) *J. Biol. Chem.* **274**, 8002 – 8011
307. Leno, G. H., Mills, A. D., Philpott, A., and Laskey, R. A. (1996) *J. Biol. Chem.* **271**, 7253 – 7256
308. Pedersen, R. A. (2001) *Nature (London)* **409**, 473 – 474
309. Shankland, M., and Seaver, E. C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4434 – 4437
310. Blow, J. J. (2001) *EMBO J.* **20**, 3293 – 3297
310a. Nomura, K., Shimizu, T., Kinoh, H., Sendai, Y., Inomata, M., and Suzuki, N. (1997) *Biochemistry* **36**, 7225 – 7238
311. Cross, J. C., Werb, Z., and Fisher, S. J. (1994) *Science* **266**, 1508 – 1517
312. Beddington, R. (1998) *Nature (London)* **395**, 641 – 643
313. He, X., Treacy, M. N., Simmons, D. M., Ingraham, H. A., Swanson, L. W., and Rosenfeld, M. G. (1989) *Nature (London)* **340**, 35 – 42
314. Rosner, M. H., Vigano, M. A., Rigby, P. W. J., Arnheiter, H., and Staudt, L. M. (1991) *Science* **253**, 144 – 145
315. Barnea, E., and Bergman, Y. (2000) *J. Biol. Chem.* **275**, 6608 – 6619
316. Lawrence, P. (1992) *The Making of a Fly: The Genetics of Animal Design*, Blackwell Scientific Publ., Oxford
317. Leptin, M. (1999) *EMBO J.* **18**, 3187 – 3192
318. Roush, W. (1996) *Science* **274**, 1608 – 1609
319. Fishman, M. C. (2001) *Science* **294**, 1290 – 1291
320. Thisse, C., and Zon, L. I. (2002) *Science* **295**, 457 – 462
321. Farber, S. A., Pack, M., Ho, S.-Y., Johnson, I. D., Wagner, D. S., Dosch, R., Mullins, M. C., Hendrickson, H. S., Hendrickson, E. K., and Halpern, M. E. (2001) *Science* **292**, 1385 – 1388
321a. Pearson, H. (2002) *Nature (London)* **418**, 14 – 15
321b. Brendza, R. P., Serbus, L. R., Duffy, J. B., and Saxton, W. M. (2000) *Science* **298**, 2120 – 2122
322. Bullock, S. L., and Ish-Horowicz, D. (2001) *Nature (London)* **414**, 611 – 616
323. Gergely, F., Kidd, D., Jeffers, K., Wakefield, J. G., and Raff, J. W. (2000) *EMBO J.* **19**, 241 – 252
324. Anderson, K. (1995) *Science* **269**, 2189 – 2190
325. McCoon, P. E., Angerer, R. C., and Angerer, L. M. (1996) *J. Biol. Chem.* **271**, 20119 – 20125
326. De Robertis, E. M., and Sasai, Y. (1996) *Nature (London)* **380**, 37 – 40
327. Dissing, M., Giordano, H., and DeLotto, R. (2001) *EMBO J.* **20**, 2387 – 2393
328. Markstein, M., Markstein, P., Markstein, V., and Levine, M. S. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 763 – 768
329. Kessler, D. S., and Melton, D. A. (1994) *Science* **266**, 596 – 604
330. Dawid, I. B., and Sargent, T. D. (1988) *Science* **240**, 1443 – 1448
331. Hemmati-Brivanlou, A., and Melton, D. A. (1992) *Nature (London)* **359**, 609 – 614
332. Schier, A. F., and Shen, M. M. (2000) *Nature (London)* **403**, 385 – 389
333. Ewing, T. (1993) *Science* **260**, 624 – 625
334. Rodriguez Esteban, C., Capdevila, J., Economides, A. N., Pascual, J., Ortiz, A., and Izpisúa, B., JC. (1999) *Nature (London)* **401**, 243 – 251
335. Ryan, A. K., Blumberg, B., Rodriguez-Esteban, C., Yonei-Tamura, S., Tamura, K., Tsukui, T., de la Pena, J., Sabbagh, W., Greenwald, J., Choe, S., Norris, D. P., Robertson, E. J., Evans, R. M., Rosenfeld, M. G., and Belmonte, J. C. I. (1998) *Nature (London)* **394**, 545 – 551
336. Mochizuki, T., Saijoh, Y., Tsuchiya, K., Shirayoshi, Y., Takai, S., Taya, C., Yonekawa, H., Yamada, K., Nihei, H., Nakatsuji, N., Overbeek, P. A., Hamada, H., and Yokoyama, T. (1998) *Nature (London)* **395**, 177 – 181
337. Witkowski, J. (1985) *Trends Biochem. Sci.* **10**, 379 – 381
338. Lamb, T. M., Knecht, A. K., Smith, W. C., Stachel, S. E., Economides, A. N., Stahl, N., Yancopolous, G. D., and Harland, R. M. (1993) *Science* **262**, 713 – 718
339. Bouwmeester, T., Kim, S.-H., Sasai, Y., Lu, B., and De Robertis, E. M. (1996) *Nature (London)* **382**, 595 – 601
340. Ruiz i Altaba, A. (1998) *Nature (London)* **391**, 748 – 749
341. Nishita, M., Hashimoto, M. K., Ogata, S., Laurent, M. N., Ueno, N., Shibuya, H., and Cho, K. W. Y. (2000) *Nature (London)* **403**, 781 – 785
342. Peifer, M., and Polakis, P. (2000) *Science* **287**, 1606 – 1609
343. Feldman, B., Gates, M. A., Egan, E. S., Dougan, S. T., Rennebeck, G., Sirotkin, H. I., Schier, A. F., and Talbot, W. S. (1998) *Nature (London)* **395**, 181 – 185
344. Brunet, L. J., McMahon, J. A., McMahon, A. P., and Harland, R. M. (1998) *Science* **280**, 1455 – 1457
345. Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. M., and Coupland, G. (1997) *Nature (London)* **386**, 44 – 51
346. Affolter, M., and Mann, R. (2001) *Science* **292**, 1080 – 1081
347. Cossu, G., and Borello, U. (1999) *EMBO J.* **18**, 6867 – 6872
348. Zhang, N., and Gridley, T. (1998) *Nature (London)* **394**, 374 – 377
349. Basler, K., and Struhl, G. (1994) *Nature (London)* **368**, 208 – 214
350. Ingham, P. W. (1998) *EMBO J.* **17**, 3505 – 3511
351. Blair, S. S. (1995) *Nature (London)* **373**, 656 – 657
352. Chiang, C., Litingtung, Y., Lee, E., Young, K. E., Corden, J. L., Westphal, H., and Beachy, P. A. (1996) *Nature (London)* **383**, 407 – 413
353. Martin, G. (1996) *Science* **274**, 203 – 204
354. Strauss, E. (1998) *Science* **280**, 1528 – 1529
355. Tozawa, R.-i, Ishibashi, S., Osuga, J.-i, Yagyu, H., Oka, T., Chen, Z., Ohashi, K., Perrey, S., Shionoiri, F., Yahagi, N., Harada, K., Gotoda, T., Yazaki, Y., and Yamada, N. (1999) *J. Biol. Chem.* **274**, 30843 – 30848
356. Kimble, J., Henderson, S., and Crittenden, S. (1998) *Trends Biochem. Sci.* **23**, 353 – 361
357. Chenn, A., and Walsh, C. A. (1999) *Science* **286**, 689 – 690
358. Gupta-Rossi, N., Bail, O. L., Gonen, H., Brou, C., Logeat, F., Six, E., Ciechanover, A., and Israël, A. (2001) *J. Biol. Chem.* **276**, 34371 – 34378
359. Goutte, C., Tsunozaki, M., Hale, V. A., and Priess, J. R. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 775 – 779
360. Heldin, C.-H., and Ericsson, J. (2001) *Science* **294**, 2111 – 2113
361. Gu, Y., Misonou, H., Sato, T., Dohmae, N., Takio, K., and Ihara, Y. (2001) *J. Biol. Chem.* **276**, 35235 – 35238
362. Brückner, K., Perez, L., Clausen, H., and Cohen, S. (2000) *Nature (London)* **406**, 411 – 415
363. Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J., Shao, L., Wilson, R., Wang, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S., and Vogt, T. F. (2000) *Nature (London)* **406**, 369 – 375
364. Evrard, Y. A., Lun, Y., Aulehla, A., Gan, L., and Johnson, R. L. (1998) *Nature (London)* **394**, 377 – 381
365. Hirsch, J. A., and Aggarwal, A. K. (1995) *EMBO J.* **14**, 6280 – 6291
366. De Robertis, E. M., Oliver, G., and Wright, C. V. E. (1990) *Sci. Am.* **263**(Jul), 46 – 52
367. McGinnis, W., and Kuziora, M. (1994) *Sci. Am.* **270**(Feb), 58 – 66

References

367a. Kmita, M., Fraudeau, N., Hérault, Y., and Duboule, D. (2002) *Nature (London)* **420**, 145 – 150

367b. Zákány, J., and Duboule, D. (1999) *Nature (London)* **401**, 761

368. Ades, S. E., and Sauer, R. T. (1995) *Biochemistry* **34**, 14601 – 14608

369. Fraenkel, E., Rould, M. A., Chambers, K. A., and Pabo, C. O. (1998) *J. Mol. Biol.* **284**, 351 – 361

370. Carr, A., and Biggin, M. D. (1999) *EMBO J.* **18**, 1598 – 1608

371. Jabet, C., Gitti, R., Summers, M. F., and Wolberger, C. (1999) *J. Mol. Biol.* **291**, 521 – 530

372. Ippel, H., Larsson, G., Behravan, G., Zdunek, J., Lundqvist, M., Schleucher, J., Lycksell, P.-O., and Wijmenga, S. (1998) *J. Mol. Biol.* **288**, 689 – 703

373. Duboule, D. (2000) *Nature (London)* **403**, 607 – 609

374. Capili, A. D., Schultz, D. C., Rauscher, F. J., III, and Borden, K. L. B. (2001) *EMBO J.* **20**, 165 – 177

375. Wray, G. A. (2001) *Science* **292**, 2256 – 2257

376. Vaziri, H., Dragowska, W., Allsopp, R. C., Thomas, T. E., Harley, C. B., and Lansdorp, P. M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9857 – 9860

377. Orkin, S. H. (1995) *J. Biol. Chem.* **270**, 4955 – 4958

378. Ziegler, B. L., Valtieri, M., Porada, G. A., Maria, R. D., Müller, R., Masella, B., Gabbianelli, M., Casella, I., Pelosi, E., Bock, T., Zanjani, E. D., and Peschle, C. (1999) *Science* **285**, 1553 – 1558

379. Kaufman, D. S., Hanson, E. T., Lewis, R. L., Auerbach, R., and Thomson, J. A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10716 – 10721

380. Nony, P., Hannon, R., Gould, H., and Felsenfeld, G. (1998) *J. Biol. Chem.* **273**, 32910 – 32919

381. Akashi, K., Traver, D., Miyamoto, T., and Weissman, I. L. (2000) *Nature (London)* **404**, 193 – 197

382. Nutt, S. L., Heavey, B., Rolink, A. G., and Busslinger, M. (1999) *Nature (London)* **401**, 556 – 562

383. Natarajan, K., Sawicki, M. W., Margulies, D. H., and Mariuzza, R. A. (2000) *Biochemistry* **39**, 14779 – 14786

384. Metcalf, D. (1992) *Trends Biochem. Sci.* **17**, 286 – 289

385. Wen, D., Boissel, J.-P., Showers, M., Ruch, B. C., and Bunn, H. F. (1994) *J. Biol. Chem.* **269**, 22839 – 22846

386. Metcalf, D. (1994) *Nature (London)* **369**, 519 – 520

387. Lebestky, T., Chang, T., Hartenstein, V., and Banerjee, U. (2000) *Science* **288**, 146 – 149

388. McMorrow, T., van den Wijngaard, A., Wollenschlaeger, A., van de Corput, M., Monkhorst, K., Trimborn, T., Fraser, P., van Lohuizen, M., Jenuwein, T., Djabali, M., Philipsen, S., Grosveld, F., and Milot, E. (2000) *EMBO J.* **19**, 4986 – 4996

389. Razin, S. V., Loudinkova, E. S., and Scherrer, K. (2000) *J. Mol. Biol.* **299**, 845 – 852

390. Weatherall, D. J., Clegg, J. B., Higgs, D. R., and Wood, W. G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3417 – 3484, McGraw-Hill, New York

391. Elnitski, L., Li, J., Noguchi, C. T., Miller, W., and Hardison, R. (2001) *J. Biol. Chem.* **276**, 6289 – 6298

392. Li, Q., Blau, C. A., Clegg, C. H., Rohde, A., and Stamatoyannopoulos, G. (1998) *J. Biol. Chem.* **273**, 17361 – 17367

393. Li, J., Noguchi, C. T., Miller, W., Hardison, R., and Schechter, A. N. (1998) *J. Biol. Chem.* **273**, 10202 – 10209

394. Wang, Z., and Liebhaber, S. A. (1999) *EMBO J.* **18**, 2218 – 2228

395. Chung, J. H., Bell, A. C., and Felsenfeld, G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 575 – 580

396. Hanahan, D. (1997) *Science* **277**, 48 – 50

397. Yancopoulos, G. D., Davis, S., Gale, N. W., Rudge, J. S., Wiegand, S. J., and Holash, J. (2000) *Nature (London)* **407**, 242 – 248

398. Carmeliet, P. (2000) *Nature (London)* **408**, 43,45

399. Risau, W. (1997) *Nature (London)* **386**, 671 – 674

400. Browder, T., Folkman, J., and Pirie-Shepherd, S. (2000) *J. Biol. Chem.* **275**, 1521 – 1524

401. Tao, Q., Backer, M. V., Backer, J. M., and Terman, B. I. (2001) *J. Biol. Chem.* **276**, 21916 – 21923

401a. Funamoto, M., Fujio, Y., Kunisada, K., Negoro, S., Tone, E., Osugi, T., Hirota, H., Izumi, M., Yoshizaki, K., Walsh, K., Kishimoto, T., and Yamauchi-Takihara, K. (2000) *J. Biol. Chem.* **275**, 10561 – 10566

402. Li, D. Y., Sorensen, L. K., Brooke, B. S., Urness, L. D., Davis, E. C., Taylor, D. G., Boak, B. B., and Wendel, D. P. (1999) *Science* **284**, 1534 – 1537

403. Carmeliet, P. (2001) *Science* **293**, 1602 – 1604

404. Leonidas, D. D., Shapiro, R., Subbarao, G. V., Russo, A., and Acharya, K. R. (2002) *Biochemistry* **41**, 2552 – 2562

405. Carmeliet, P., and Jain, R. K. (2000) *Nature (London)* **407**, 249 – 257

406. Otani, A., Slike, B. M., Dorrell, M. I., Hood, J., Kinder, K., Ewalt, K. L., Cheresh, D., Schimmel, P., and Friedlander, M. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 178 – 183

407. Maisonpierre, P. C., Suri, C., Jones, P. F., Bartunkova, S., Wiegand, S. J., Radziejewski, C., Compton, D., McClain, J., Aldrich, T. H., Papadopoulos, N., Daly, T. J., Davis, S., Sato, T. N., and Yancopoulos, G. D. (1997) *Science* **277**, 55 – 60

408. Dawson, D. W., Volpert, O. V., Gillis, P., Crawford, S. E., Xu, H.-J., Benedict, W., and Bouck, N. P. (1999) *Science* **285**, 245 – 248

409. Hohenester, E., Sasaki, T., Olsen, B. R., and Timpl, R. (1998) *EMBO J.* **17**, 1656 – 1664

410. L'Heureux, N., Pâquet, S., Labbé, R., Germain, L., and Auger, F. A. (1998) *FASEB J.* **12**, 47 – 56

411. Alliston, T., Choy, L., Ducy, P., Karsenty, G., and Derynck, R. (2001) *EMBO J.* **20**, 2254 – 2272

412. Liu, Y., Li, H., Tanaka, K., Tsumaki, N., and Yamada, Y. (2000) *J. Biol. Chem.* **275**, 12712 – 12718

413. ten Dijke, P., Yamashita, H., Sampath, T. K., Reddi, A. H., Estevez, M., Riddle, D. L., Ichijo, H., Heldin, C.-H., and Miyazono, K. (1994) *J. Biol. Chem.* **269**, 16985 – 16988

414. Scheufler, C., Sebald, W., and Hülsmeyer, M. (1999) *J. Mol. Biol.* **287**, 103 – 115

415. Harland, R. M. (2001) *Nature (London)* **410**, 423 – 424

416. Alliston, T., and Derynck, R. (2002) *Nature (London)* **416**, 686 – 687

417. Hasty, P., Bradley, A., Morris, J. H., Edmondson, D. G., Venuti, J. M., Olson, E. N., and Klein, W. H. (1993) *Nature (London)* **364**, 501 – 506

418. Zhang, J.-M., Zhao, X., Wei, Q., and Paterson, B. M. (1999) *EMBO J.* **18**, 6983 – 6993

419. Charbonnier, P., Gaspera, B. D., Armand, A.-S., Van der Laarse, W. J., Launay, T., Becker, C., Gallien, C.-L., and Chanoine, C. (2002) *J. Biol. Chem.* **277**, 1139 – 1147

420. Miller, J. B. (1995) *Nature (London)* **377**, 575 – 576

421. Stephanou, A., Brar, B. K., Scarabelli, T. M., Jonassen, A. K., Yellon, D. M., Marber, M. S., Knight, R. A., and Latchman, D. S. (2000) *J. Biol. Chem.* **275**, 10002 – 10008

422. Srivastava, D., and Olson, E. N. (2000) *Nature (London)* **407**, 221 – 226

423. Andreadis, S. T., Hamoen, K. E., Yarmush, M. L., and Morgan, J. R. (2001) *FASEB J.* **15**, 898 – 906

424. Andersen, B., Schonemann, M. D., Flynn, S. E., Pearse, R. V., II, Singh, H., and Rosenfeld, M. G. (1993) *Science* **260**, 78 – 82

425. Mills, A. A., Zheng, B., Wang, X.-J., Vogel, H., Roop, D. R., and Bradley, A. (1999) *Nature (London)* **398**, 708 – 713

426. Yang, A., Schweitzer, R., Sun, D., Kaghad, M., Walker, N., Bronson, R. T., Tabin, C., Sharpe, A., Caput, D., Crum, C., and McKeon, F. (1999) *Nature (London)* **398**, 714 – 718

427. Hartmann, D. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10522 – 10523

428. van den Brink, G. R., de Santa Barbara, P., and Roberts, D. J. (2001) *Science* **294**, 2115 – 2116

428a. Peifer, M. (2002) *Nature (London)* **410**, 274 – 277

429. Risau, W. (1995) *FASEB J.* **9**, 926 – 933

430. Griffin, C. T., Srinivasan, Y., Zheng, Y.-W., Huang, W., and Coughlin, S. R. (2001) *Science* **293**, 1666 – 1670

430a. Cooper, M. T. D., and Bray, S. J. (1999) *Nature (London)* **397**, 526 – 530

430b. Knust, E., and Bossinger, O. (2002) *Science* **298**, 1955–1959

431. Livesey, R., and Cepko, C. (2001) *Nature (London)* **413**, 471,473

432. McKay, R. (1997) *Science* **276**, 66 – 71

433. Ourednik, V., Ourednik, J., Flax, J. D., Zawada, W. M., Hutt, C., Yang, C., Park, K. I., Kim, S. U., Sidman, R. L., Freed, C. R., and Snyder, E. Y. (2001) *Science* **293**, 1820 – 1824

434. Tanabe, Y., and Jessell, T. M. (1996) *Science* **274**, 1115–1123

434a. García-Castro, M. I., Marcelle, C., and Bronner-Fraser, M. (2002) *Science* **297**, 848 – 851

434b. Hatten, M. E. (2002) *Science* **297**, 1660 – 1663

435. Scully, K. M., and Rosenfeld, M. G. (2002) *Science* **295**, 2231 – 2235

436. Dodd, J., Jessell, T. M., and Placzek, M. (1998) *Science* **282**, 1654 – 1657

437. Jeong, J., and McMahon, A. P. (2001) *Nature (London)* **412**, 136 – 137

437a. Dickson, B. J. (2002) *Science* **298**, 1959 – 1964

438. Shatz, C. J. (1992) *Sci. Am.* **267**(Sep), 61 – 67

439. Ramón y Cajal, S. (1892) *La Cellle* **9**, 119

440. Baier, H., and Bonhoeffer, F. (1994) *Science* **265**, 1541 – 1542

441. Serafini, T., Kennedy, T. E., Galko, M. J., Mirzayan, C., Jessell, T. M., and Tessier-Lavigne, M. (1994) *Cell* **78**, 409 – 424

442. Kennedy, T. E., Serafini, T., de la Torre, J. R., and Tessier-Lavigne, M. (1994) *Cell* **78**, 425 – 435

443. Tessier-Lavigne, M., and Goodman, C. S. (1996) *Science* **274**, 1123 – 1133

444. Stein, E., and Tessier-Lavigne, M. (2001) *Science* **291**, 1928 – 1938

445. Marx, J. (1995) *Science* **268**, 971 – 973

446. Ming, G.-L., Wong, S. T., Henley, J., Yuan, X.-b, Song, H.-j, Spitzer, N. C., and Poo, M.-m. (2002) *Nature (London)* **417**, 411 – 418

447. Tear, G. (2001) *Nature (London)* **409**, 472 – 473

448. Dickson, B. J. (2001) *Science* **291**, 1910 – 1911

449. Fricke, C., Lee, J.-S., Geiger-Rudolph, S., Bonhoeffer, F., and Chien, C.-B. (2001) *Science* **292**, 507 – 510

450. Wang, H., and Tessier-Lavigne, M. (1999) *Nature (London)* **401**, 765 – 769

451. Takei, K., Shin, R.-M., Inoue, T., Kato, K., and Mikoshiba, K. (1998) *Science* **282**, 1705 – 1708

452. Polleux, F., Morrow, T., and Ghosh, A. (2000) *Nature (London)* **404**, 567 – 573

References

453. Marín, O., Yaron, A., Bagri, A., Tessier-Lavigne, M., and Rubenstein, J. L. R. (2001) *Science* **293**, 872 – 875
454. Goffinet, A. M. (1997) *Nature (London)* **389**, 668 – 669
455. Bar, I., and Goffinet, A. M. (1999) *Nature (London)* **399**, 645 – 646
456. Yip, J. W., Yip, Y. P. L., Nakajima, K., and Capriotti, C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8612 – 8616
457. Chen, L., Chetkovich, D. M., Petralia, R. S., Sweeney, N. T., Kawasaki, Y., Wenthold, R. J., Bredt, D. S., and Nicoll, R. A. (2000) *Nature (London)* **408**, 936 – 943
458. Nakagawa, T., and Sheng, M. (2000) *Science* **290**, 2270 – 2271
459. Smith, S. J. (1999) *Science* **283**, 1860 – 1861
459a. Cohen-Cory, S. (2002) *Science* **298**, 770–776
460. Ikonomidou, C., Bosch, F., Miksa, M., Bittigau, P., Vöckler, J., Dikranian, K., Tenkova, T. I., Stefovska, V., Turski, L., and Olney, J. W. (1999) *Science* **283**, 70 – 74
461. Goldberg, J. L., and Barres, B. A. (2000) *Nature (London)* **403**, 369 – 370
462. Prinjha, R., Moore, S. E., Vinson, M., Blake, S., Morrow, R., Christie, G., Michalovich, D., Simmons, D. L., and Walsh, F. S. (2000) *Nature (London)* **403**, 33 – 384
463. Olson, L. (2002) *Nature (London)* **416**, 589 – 590
464. Meyerowitz, E. M. (2002) *Science* **295**, 1482 – 1485
465. Buchanan, B. B., Gruissem, W., and Jones, R. L., eds. (2000) *Biochemistry and Molecular Biology of Plants*, American Society of Plant Physiologists, Rockville, Maryland
466. Goldberg, R. B., de Paiva, G., and Yadegari, R. (1994) *Science* **266**, 605 – 614
467. Bewley, J. D., Hempel, F. D., McCormick, S., and Zambryski, P. (2000) in *Biochemistry and Molecular Biology of Plants* (Buchanan, B., Gruissem, W., and Jones, R., eds), American Society of Plant Physiologists, Rockville, Maryland
468. Thompson, R. D. (2000) *Nature (London)* **408**, 39,41
469. Vielle Calzada, J.-P., Crane, C. F., and Stelly, D. M. (1996) *Science* **274**, 1322 – 1323
470. Dellaporta, S. L., and Calderon-Urrea, A. (1994) *Science* **266**, 1501 – 1505
471. Nasrallah, J. B. (2002) *Science* **296**, 305 – 308
472. Cui, X., Wise, R. P., and Schnable, P. S. (1996) *Science* **272**, 1334 – 1336
473. Lichtenstein, C. (1990) *Trends Biochem. Sci.* **15**, 453 – 454
474. Levings, C. S., III. (1996) *Science* **272**, 1279 – 1280
475. Jürgens, G. (2001) *EMBO J.* **20**, 3609 – 3616
476. van den Berg, C., Willemsen, V., Hage, W., Weisbeek, P., and Scheres, B. (1995) *Nature (London)* **378**, 62 – 65
477. Hake, S. (2001) *Nature (London)* **413**, 261 – 264
478. Vielle-Calzada, J.-P., Baskar, R., and Grossniklaus, U. (2000) *Nature (London)* **404**, 91 – 94
479. Ferrándiz, C., Pelaz, S., and Yanofsky, M. F. (1999) *Ann. Rev. Biochem.* **68**, 321 – 354
480. Vrebalov, J., Ruezinsky, D., Padmanabhan, V., White, R., Medrano, D., Drake, R., Schuch, W., and Giovannoni, J. (2002) *Science* **296**, 343 – 346
481. Brown, K. (2001) *Science* **291**, 1884 – 1885
482. Lane, B. G. (1994) *FASEB J.* **8**, 294 – 301
483. Berardini, T. Z., Bollman, K., Sun, H., and Poethig, R. S. (2001) *Science* **291**, 2405 – 2407
484. Jürgens, G. (1997) *Nature (London)* **386**, 17
485. Timmermans, M. C. P., Hudson, A., Becraft, P. W., and Nelson, T. (1999) *Science* **284**, 151 – 153
486. Simpson, G. G., and Dean, C. (2002) *Science* **296**, 285 – 289
486a Dangl, J. L., Dietrich, R. A., and Thomas, H. (2000) in *Biochemistry & Molecular Biology of Plants* (Buchanan, B., Gruissem, W., and Jones, R., eds), pp. 1044 – 1100, American Society of Plant Physiologists, Rockville, Maryland
487. Ly, D. H., Lockhart, D. J., Lerner, R. A., and Schultz, P. G. (2000) *Science* **287**, 2486 – 2492
488. Shay, J. W., and Wright, W. E. (2001) *Science* **291**, 839 – 840
489. Takahashi, Y., Kuro-o, M., and Ishikawa, F. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12407 – 12408
490. Olshansky, S. J., Carnes, B. A., and Désesquelles, A. (2001) *Science* **291**, 1491 – 1492
491. Weindruch, R. (1996) *Sci. Am.* **274**(Jan), 46 – 52
492. Hayflick, L. (2000) *Nature (London)* **408**, 267 – 269
493. Puca, A. A., Daly, M. J., Brewster, S. J., Matise, T. C., Barrett, J., Shea-Drinkwater, M., Kang, S., Joyce, E., Nicoli, J., Benson, E., Kunkel, L. M., and Perls, T. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10505 – 10508
494. Finch, C. E., and Tanzi, R. E. (1997) *Science* **278**, 407 – 411
495. Holliday, R. (2001) *Trends Biochem. Sci.* **26**, 68 – 71
496. Guarente, L., and Kenyon, C. (2000) *Nature (London)* **408**, 255 – 262
497. Kirkwood, T. B. L., and Austad, S. N. (2000) *Nature (London)* **408**, 233 – 238
498. Rusting, R. L. (1992) *Sci. Am.* **267**(Dec), 130 – 141
499. Stearns, S. C., Ackermann, M., Doebeli, M., and Kaiser, M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3309 – 3313
500. Raha, S., and Robinson, B. H. (2000) *Trends Biochem. Sci.* **25**, 502 – 508
501. Finkel, T., and Holbrook, N. J. (2000) *Nature (London)* **408**, 239 – 247
502. Hamilton, M. L., Van Remmen, H., Drake, J. A., Yang, H., Guo, Z. M., Kewitt, K., Walter, C. A., and Richardson, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10469 – 10474
503. Zou, S., Meadows, S., Sharp, L., Jan, L. Y., and Jan, Y. N. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13726 – 13731
504. Murray, C. A., and Lynch, M. A. (1998) *J. Biol. Chem.* **273**, 12161 – 12168
504a. Roth, G. S., Lane, M. A., Ingram, D. K., Mattison, J. A., Elahi, D., Tobin, J. D., Muller, D., and Metter, E. J. (2002) *Science* **297**, 811
505. Witkowski, J. (1985) *Trends Biochem. Sci.* **10**, 258 – 260
506. Medvedev, Z. A. (1972) *Exptl. Gerontol.* **1**, 227 – 238
507. Hayflick, L., and Moorhead, P. S. (1961) *Exp. Cell. Res.* **25**, 585 – 621
508. Hayflick, L. (1980) *Sci. Am.* **242**, 58 – 65
509. Tang, D. G., Tokumoto, Y. M., Apperly, J. A., Lloyd, A. C., and Raff, M. C. (2001) *Science* **291**, 868 – 871
510. Yang, J., Chang, E., Cherry, A. M., Bangs, C. D., Oei, Y., Bodnar, A., Bronstein, A., Chiu, C.-P., and Herron, G. S. (1999) *J. Biol. Chem.* **274**, 26141 – 26148
511. Romanov, S. R., Kozakiewicz, B. K., Hoist, C. R., Stampfer, M. R., Haupt, L. M., and Tisty, T. D. (2001) *Nature (London)* **409**, 633 – 637
512. Freshney, R. I., and Freshney, M. G., eds. (1996) *Culture of Immortalized Cells*, Wiley Liss, New York
513. Mathon, N. F., Malcolm, D. S., Harrisingh, M. C., Cheng, L., and Lloyd, A. C. (2001) *Science* **291**, 872 – 875
514. de Lange, T., and DePinho, R. A. (1999) *Science* **283**, 947 – 949
515. Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S., and Reddel, R. R. (1995) *EMBO J.* **14**, 4240 – 4248
516. Herbert, B. S., Pitts, A. E., Baker, S. I., Hamilton, S. E., Wright, W. E., Shay, J. W., and Corey, D. R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14276 – 14281
517. Kucherlapati, R., and DePinho, R. A. (2001) *Nature (London)* **411**, 647 – 648
518. Wright, W. E., Brasiskyte, D., Piatyszek, M. A., and Shay, J. W. (1996) *EMBO J.* **15**, 1734 – 1741
519. Lanza, R. P., Cibelli, J. B., Blackwell, C., Cristofalo, V. J., Francis, M. K., Baerlocher, G. M., Mak, J., Schertzer, M., Chavez, E. A., Sawyer, N., Lansdorp, P. M., and West, M. D. (2000) *Science* **288**, 665 – 669
520. Lin, S.-J., Defossez, P.-A., and Guarente, L. (2000) *Science* **289**, 2126 – 2128
521. Schweitzer, B. I., Dicker, A. P., and Bertino, J. R. (1990) *FASEB J.* **4**, 2441 – 2452
522. Lin, S.-J., Kaeberlein, M., Andalis, A. A., Sturtz, L. A., Defossez, P.-A., Culotta, V. C., Fink, G. R., and Guarente, L. (2002) *Nature (London)* **418**, 344 – 348
523. Campisi, J. (2000) *Science* **289**, 2062 – 2063
524. Lakowski, B., and Hekimi, S. (1996) *Science* **272**, 1010 – 1013
525. Jonassen, T., Larsen, P. L., and Clarke, C. F. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 421 – 426
526. Larsen, P. L., and Clarke, C. F. (2002) *Science* **295**, 120 – 123
527. Kimura, K. D., Tissenbaum, H. A., Liu, Y., and Ruvkun, G. (1997) *Science* **277**, 942 – 946
528. Lin, K., Dorman, J. B., Rodan, A., and Kenyon, C. (1997) *Science* **278**, 1319 – 1322
529. Strauss, E. (2001) *Science* **292**, 41,43
530. Taub, J., Lau, J. F., Ma, C., Hahn, J. H., Hoque, R., Rothblatt, J., and Chalfie, M. (1999) *Nature (London)* **399**, 162 – 166
531. Apfeld, J., and Kenyon, C. (1999) *Nature (London)* **402**, 804 – 807
532. Arantes-Oliveira, N., Apfeld, J., Dillin, A., and Kenyon, C. (2002) *Science* **295**, 502 – 505
532a. Garigan, D., Hsu, A. L., Fraser, A. G., Kamath, R. S., Ahringer, J., and Kenyon, C. (2002) *Genetics* **161**, 1101– 1112
532b. Tower, J. (2000) *Mech Ageing Dev* **118**, 1– 14
533. Clancy, D. J., Gems, D., Hafen, E., Leevers, S. J., and Partridge, L. (2002) *Science* **296**, 319
534. West, A. P., Jr., Llamas, L. L., Snow, P. M., Benzer, S., and Bjorkman, P. J. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3744 – 3749
535. Tatar, M., Kopelman, A., Epstein, D., Tu, M.-P., Yin, C.-M., and Garofalo, R. S. (2001) *Science* **292**, 107 – 110
536. Rogina, B., Reenan, R. A., Nilsen, S. P., and Helfand, S. L. (2000) *Science* **290**, 2137 – 2140
537. Chavous, D. A., Jackson, F. R., and O'Connor, C. M. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14814 – 14818
538. Guarente, L. (1999) *Nature (London)* **402**, 243,245
539. Moskovitz, J., Bar-Noy, S., Williams, W. M., Requena, J., Berlett, B. S., and Stadtman, E. R. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12920 – 12925
540. Ruan, H., Tang, X. D., Chen, M.-L., Joiner, M. A., Sun, G., Brot, N., Weissbach, H., Heinemann, S. H., Iverson, L., Wu, C.-F., and Hoshi, T. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2748 – 2753
541. De Benedictis, G., Rose, G., Carrieri, G., De Luca, M., Falcone, E., Passarino, G., Bonafé, M., Monti, D., Baggio, G., Bertolini, S., Mari, D., Mattace, R., and Franceschi, C. (1999) *FASEB J.* **13**, 1532 – 1536
542. Hagen, T. M., Ingersoll, R. T., Lykkesfeldt, J., Liu, J., Wehr, C. M., Vinarsky, V., Bartholomew, J. C., and Ames, B. N. (1999) *FASEB J.* **13**, 411 – 418
543. Martin, G. M., and Oshima, J. (2000) *Nature (London)* **408**, 263 – 266

544. Marx, J. (2000) *Science* **287**, 2390
545. Kamath-Loeb, A. S., Johansson, E., Burgers, P. M. J., and Loeb, L. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4603–4608
546. Kusano, K., Johnson-Schlitz, D. M., and Engels, W. R. (2001) *Science* **291**, 2600–2602
547. de Boer, J., Andressoo, J. O., de Wit, J., Huijmans, J., Beems, R. B., van Steeg, H., Weeda, G., van der Horst, G. T. J., van Leeuwen, W., Themmen, A. P. N., Meradji, M., and Hoeijmakers, J. H. J. (2002) *Science* **296**, 1276–1279
548. Hasty, P., and Vijg, J. (2002) *Science* **296**, 1250–1251
549. Kuro-o, M., Matsumura, Y., Aizawa, H., Kawaguchi, H., Suga, T., Utsugi, T., Ohyama, Y., Kurabayashi, M., Kaname, T., Kume, E., Iwasaki, H., Iida, A., Shiraki-Iida, T., Nishikawa, S., Nagai, R., and Nabeshima, Y.-i. (1997) *Nature (London)* **390**, 45–51
550. Arking, D. E., Krebsova, A., Macek, M., Sr., Macek, M., Jr., Arking, A., Mian, I. S., Fried, L., Hamosh, A., Dey, S., McIntosh, I., and Dietz, H. C. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 856–861
551. Tyner, S. D., Venkatachalam, S., Choi, J., Jones, S., Ghebranious, N., Igelmann, H., Lu, X., Soron, G., Cooper, B., Brayton, C., Park, S. H., Thompson, T., Karsenty, G., Bradley, A., and Donehower, L. A. (2002) *Nature (London)* **415**, 45–53
552. Strauss, E. (2002) *Science* **295**, 28–29
553. McCormick, F. (1999) *Trends Biochem. Sci.* **24**, M53–M56
554. Aksan, I., and Stinson, J. A. (2002) *Trends Biochem. Sci.* **27**, 387–389
555. Pollock, P. M., and Meltzer, P. S. (2002) *Nature (London)* **417**, 906–907
556. Ryan, K. M., and Vousden, K. H. (2002) *Nature (London)* **419**, 795, 797
557. Vogelstein, B., Lane, D., and Levine, A. J. (2000) *Nature (London)* **408**, 307–310
558. Evan, G. I., and Vousden, K. H. (2001) *Nature (London)* **411**, 342–348
559. Massagué, J., and Serrano, M. (2000) *EMBO Reports* **1**, 115–119
560. Brazil, D. P., and Hemmings, B. A. (2001) *Trends Biochem. Sci.* **26**, 657–664
561. Mayo, L. D., and Donner, D. B. (2002) *Trends Biochem. Sci.* **27**, 462–467
562. Taipale, J., and Beachy, P. A. (2001) *Nature (London)* **411**, 349–354
563. Huntsman, D. G., Carneiro, F., Lewis, F. R., MacLeod, P. M., Hayashi, A., Monaghan, K. G., Maung, R., Seruca, R., Jackson, C. E., and Caldas, C. (2001) *N. Engl. J. Med.* **344**, 1904–1909
564. Weitzman, J. B., and Yaniv, M. (1999) *Nature (London)* **400**, 401–402
565. Marx, J. (2002) *Science* **297**, 544–546
566. Hoeijmakers, J. H. J. (2001) *Nature (London)* **411**, 366–374
567. Maser, R. S., and DePinho, R. A. (2002) *Science* **297**, 565–569
568. Monteiro, A. N. A. (2000) *Trends Biochem. Sci.* **25**, 469–474
569. Wilson, J. H., and Elledge, S. J. (2002) *Science* **297**, 1822–1823
570. Yang, H., Jeffrey, P. D., Miller, J., Kinnucan, E., Sun, Y., Thomä, N. H., Zheng, N., Chen, P.-L., Lee, W.-H., and Pavletich, N. P. (2002) *Science* **297**, 1837–1848
571. Yu, Q., Geng, Y., and Sicinski, P. (2001) *Nature (London)* **411**, 1017–1021
572. Hakomori, S. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10231–10233
573. Bernards, R., and Weinberg, R. A. (2002) *Nature (London)* **418**, 823
574. Varambally, S., Dhanasekaran, S. M., Zhou, M., Barrette, T. R., Kumar-Sinha, C., Sanda, M. G., Ghosh, D., Pienta, K. J., Sewalt, R. G. A. B., Otte, A. P., Rubin, M. A., and Chinnalyan, A. M. (2002) *Nature (London)* **419**, 624–629
575. Müller, A., Homey, B., Soto, H., Ge, N., Catron, D., Buchanan, M. E., McClanahan, T., Murphy, E., Yuan, W., Wagner, S. N., Barrera, J. L., Mohar, A., Verástegui, E., and Zlotnik, A. (2001) *Nature (London)* **410**, 50–56
576. Davies, H., and 51 other authors. (2002) *Nature (London)* **417**, 949–954
577. Barinaga, M. (2000) *Science* **287**, 1584–1585
578. Raven, P. H. (2002) *Science* **297**, 954–958
579. Myers, N. (2001) *Nature (London)* **410**, 631–632
580. Serageldin, I. (2002) *Science* **296**, 54–58
581. Vitousek, P. M., Mooney, H. A., Lubchenco, J., and Melillo, J. M. (1997) *Science* **277**, 494–499
582. Rojstaczer, S., Sterling, S. M., and Moore, N. J. (2001) *Science* **294**, 2549–2552
583. Rosenzweig, M. L. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5404–5410
584. Krebs, J. R., Wilson, J. D., Bradbury, R. B., and Siriwardena, G. M. (1999) *Nature (London)* **400**, 6111–6112
585. Pimm, S. L., and 32 other authors. (2001) *Science* **293**, 2207–2208
586. Schiermeier, Q. (2002) *Nature (London)* **419**, 662–665
587. Naylor, R. L., Goldburg, R. J., Primavera, J. H., Kautsky, N., Beveridge, M. C. M., Clay, J., Folke, C., Lubchenco, J., Mooney, H., and Troell, M. (2000) *Nature (London)* **405**, 1017–1024
588. Ziman, J. (1996) *Nature (London)* **382**, 751–754

Study Questions

1. Discuss the roles of the following proteins in development: receptors, transcription factors, protein kinases, histones, DNA methylases, adhesion molecules, ubiquitin. How do small RNA molecules participate in development?

2. Are all body cells totipotent?

3. Discuss the roles of apoptosis in various groups of organisms.

4. Compare signaling between bacteria and other unicellular organisms with signaling in higher eukaryotes.

5. Are human beings the most highly developed organisms? If so, in what ways? Has evolution of humans stopped or will it continue? Will it be upward?

6. Is it important for the world to achieve a sustainable state in which the population is constant and the environment stable?[578] How will the world support a projected increase in population from the present 6 billion to 9 billion in 50 years?[579,580]

7. How seriously is the earth's ecosystem dominated by human activity?[581,582] Human activities have greatly reduced the amount of area available to wild species. Will the ensuing extinction of many organisms impoverish future diversity?[583–585] Can the world's fisheries become sustainable?[586,587]

8. Is science losing its objectivity because of an emphasis on monetary gain rather than on meeting social needs?[588]

Complete Index to Volumes 1 and 2

Page numbers set in **boldface** refer to major discussions.
The symbol *s* after a page number refers to a chemical structure.

A

A-band (anisotropic band, muscle) 1096
ABC transporters 417, 419, 839
Abequose (3,6-dideoxy-D-galactose) 180*s*
  from CDP-glucose 1138
Abietic acid 1235s,1236
abl 572, 576
Abortive complexes 466, 475. *See also* Nonproductive complexes
Abrin 1685
Abscisic acid 1235*s*, 1761
Absorption bands
  bathochromic shift 1283
  position in relation to structures 1283
  shapes of 1280
Absorption spectra 122, 123, **1275 – 1286**
  of *N*-acetyltyrosine ethyl ester 123
  of adenosine 205
  of aromatic amino acid derivatives 123
  of aspartate carbamoyltransferase 1285
  bands at 500 nm 750
  of chlorophyll 1304
  of cytidine 205
  of cytochrome *c* 847
  of lycopene 1284
  of phytochrome 1338
  polarized of aspartate aminotransferase 1283
  of proteins 122, 123
  of rhodopsin 1326
  of uridine 205
  of vitamin B_6 coenzymes 749, 750
Abzymes 1842
Acarbose 607, 607*s*
Acatalasemia 852
Acetabularia 21, 22
Acetaldehyde, hydration equilibrium 677
Acetaminophen, toxicity of 1070
Acetan 179
Acetate
  binding constants to metal ions 311
  as carbon source 988
  Gibbs energy of oxidation 297
Acetate kinase 661
Acetic acid
  pK_a value of 99, 293
  utilization by bacteria 1057
Acetic anhydride, acylation by 124
Acetoacetate 946*s*
  concentration in blood 946
  metabolism of 946, 947, 1003
Acetoacetate decarboxylase 705
Acetoacetyl-CoA 700*s*, 945
Acetohydroxy acid isomeroreductase 712
Acetohydroxy acid synthase (acetolactate synthase) 734, 1391
Acetoin 970
Acetolactate 734*s*, 735, 970
Acetone 705*s*
  from fermentation 969
  tautomerism 45
Acetosyringone 1497, 1498*s*
Acetyl adenylate 661, 720
N-Acetylaminopropionaldehyde 1382*s*
Acetylation **720 – 723**
  coenzyme A in 720
  of histones 549
  of sulfanilamide 720
Acetylcholine 720, 1200, 1783*s*
  receptors 422, **1782 – 1786**
Acetylcholinesterase 609, 634, 636
  pseudosubstrate of 636
Acetyl-CoA (acetyl-coenzyme A) 506*s*, 508, **720 – 723**, 881, 940, 952
  as allosteric effector 952
  carboxylation of 724, 978
  in citric acid cycle 950 – 958
  corrinoid-dependent synthesis 876, 877
  formation from fatty acids 511
  reactions of, table 700
  synthesis of 660
  in Wood–Ljungdahl pathway 881, 984
Acetyl-CoA carboxylase **725 – 729**
  biotin subunit 726
Acetyl-CoA: α-glucosaminide *N*-acetyltransferase 1171
Acetyl-CoA-glyoxylate cycle 988
Acetylene hydratase 894
Acetylene reduction test 1360
Acetylenic fatty acids 381
Acetyl-enzyme
  in citrate lyase 704
  from serine proteases 617
N-Acetylgalactosamine 1155
N-Acetylglucosamine (GlcNAc) 164, 599, 1157
N-Acetylglucosamine-1-phosphate 1135
N-Acetylglucosaminidase 1171
N-Acetylglucosaminyl-phosphotransferase 1159
N-Acetylglutamate 1376
N-Acetyllactosamine 183
  biosynthesis 1141
N-Acetylmuramic acid (MurNAc) 165s, 599
N-Acetylneuraminic (sialic) acid (NeuAc) 165*s*, 1136
N-Acetylneuraminic acid 9-phosphate 1137
α-*N*-Acetylneuraminidase deficiency 1171
N-Acetylornithine 742
Acetyl phosphate 799
O-Acetylserine sulfhydrylase 742
N-Acetylspermidine 1382*s*
S-Acetyltransferase 661
N-Acetyltyrosine amide 617*s*
Achondroplasia (human dwarfism) 1753
Acid(s)
  enthalpy change of dissociation, table 293
  Gibbs energy of dissociation, table 293
  pK_a values of, table 293
  strengths of **95, 96**
Acid–base catalysis **469 – 471**, **486 – 491**
  concerted 490
  of mutarotation 487
Acid–base chemistry
  of amino acids 51– 55, 304 – 309
  of enzymatic action 486 – 494, 691– 697
  of nucleic acids 204 – 211
Acid catalysis
  general 487– 491
  specific 489
α1-Acid glycoprotein 1850
Acid lipase 1248
Acid phosphatases 645
Acidity. *See also* Specific compounds and pK_a values
  of benzoic acids 308
  of fatty acids 380
  of groups in enzymes 486
  of ketones 46
  of phenols 308
  of phenylacetic acids 308
Aconitase 676, **686 – 688**, 861, 960
  active site structure 689
  proton retention by 686
cis-Aconitate 516*s*, 686*s*
  isomerization of 697
Acrasins 557
Acridine(s) 1588
Acridine yellow 223*s*
Acrolein 1205*s*, 1588
Acrosome 1895, 1896

  process 369, 370
Acrylochlorin heme 845, 845*s*
Acrylonitrile 115
Acrylyl-CoA 947
ACTH (adrenocorticotropin) 54*s*, 1741, 1744
  adenylate cyclase in action of 556
Actin 406, **1096 – 1099**
  binding to myosin 1106
  filament, structure of 338, 1101
  monomer 1097
  thin filaments 369, 1096, 1097
Actin-based motility **1118 – 1120**
Actin-binding proteins, table 1118
Actinidin 618
α-Actinin 370, 405, 1099
Actinomycetes 6
Actinomycin 1164, 1618, 1619*s*
Actinorhodin 1216
Activation
  of enzyme(s) 471 – 477
  of enzyme synthesis 536
  hormone 1760
  noncompetitive 473 – 475
Activation cascade, pancreatic proteases 609
Active transport 379, 410, **417– 420**
  in mitochondria 1047
  of Na^+ and K^+ 424
Activin 1899
Activity coefficients 288
Activity of a solute 287
Actomyosin **1104** – **1114**
  ATPase cycle 1108, 1109
  chemical mechanisms 1110 – 1114
  Huxley, rowing model 1104
Acute myocardial ischemia 944
Acute phase response 1832, 1850
Acyclovir 1655*s*
Acyl-ACP, dehydrogenation of 766
Acyl adenylate 507
  of tRNA 1672, 1695, 1696
Acyl carrier protein (ACP) 681, 704, 723, 978
Acyl-CoA 526
  derivatives from acetyl-CoA 973
Acyl-CoA dehydrogenase 766, 785, 789, 940
Acyl-dihydrolipoyl intermediate 736
Acyl-enzyme
  in glyceraldehyde phosphate dehydrogenase 775
  intermediates 610 – 613
Acyl groups, table of names 380
Acyl phosphates 507, 509, **660, 661**, 775*s*, 975, 993
Acylation 124
  of proteins 559
  in regulation 543
Acylglycerols 382
N-Acylhomoserine lactone 1758*s*
Acyltransferases 637
Adair equation 328, 354
Adamantyl-1,2 dioxetane phosphate 1342*s*
Ada protein 1583, 1613
Adapter proteins
  Grb2 577
  Shc 577
Addiction 1797
Addison's disease 1254
Addition–elimination sequence 526
Addition reactions 526, 530, **677– 690**
  assisted isomerization 690
  of double bonds, adjacent to carboxylate 683 – 690
  of double bonds, carbon-carbon 681–683, 688
  of double bonds, polarized **677 – 690**
  of enolate to carbon dioxide 705 – 710
  facilitation by carbonyl group 681
  nucleophilic 123
  reversibility of 690
  stereochemistry 680, 681
Adducin 405
Adenine (Ade, A) 199*s*, 203, 1458*s*
  binding constants to metal ions 311
Adenine methylase 1541
Adenine nucleotide carrier
  of mitochondria 1047
Adenosine (Ado) 203, 778*s*, 1752, 1794, 1795
  absorption spectra of 205
  acid–base chemistry 204, 205
  hypermodified 235
  pK_a value of 293
Adenosine 3',5'-bisphosphate 977
Adenosine diphosphate. *See* ADP
Adenosine diphosphate-glucose pyrophosphorylase 1144
Adenosine-diphosphate-ribose. *See* ADP-ribose
Adenosine monophosphate. *See* AMP
Adenosine 5'-phosphate. *See* AMP
Adenosine 5′-phosphosulfate (APS) 1053
  reduction by cytochrome c_3 1056
Adenosine 5′-phosphosulfate reductase 1056
Adenosine triphosphate. *See* ATP
S-Adenosylhomocysteine 592*s*, 778*s*, 1388
S-Adenosylhomocysteine hydrolase 778
S-Adenosylmethionine (AdoMet, SAM) 591, 592*s*, 813, 875, 975*s*, 1388
  decarboxylation 753, 754
  formation of 977
  reaction with catechol
    O-methyltransferase 592
Adenosyltransferase, B_{12s} 870
Adenoviruses 247, 346, 1498, 1559
Adenylate cyclase **556, 557**, 556*s*, 657, 1799
  characterization of 556
  toxic 548
Adenylate kinase 303, 654, 655*s*
  fold 658, 659
  functions of 655
  in intermembrane space 1016
Adenylate system **302 – 304**, 1034
  storage of energy 302
5'-Adenylic acid. *See* AMP
Adenylosuccinate 1454
Adenylosuccinate lyase 685
Adenylylation 545
Adenylyl cyclase. *See* Adenylate cyclase
Adenylyl groups, transfer from ATP 660
Adenylyl sulfate 977
Adenylyltransferase (AT) 1370
Adherens junctions 1883
Adhesion, cell surface 1883, 1884
Adhesion discs (septate desmosomes) 29
Adhesion proteins 187
Adipocyte 379*s*
Adipokinetic hormones 1760
ADP (Adenosine diphosphate) 536, 1752, 1794
  in adenylate system 302 – 304
  complexes with metal ions 296
  dissociation as acid 288
  intracellular concentration 304
  P-31 NMR spectrum 642
  pK_a value of 293
  in regulation 535
ADP-glucose
  formation of 1320
ADP-ribose (ADPR) 315, 778, 780
ADP-ribosylation 545, **778**
ADP-ribosylation factors (ARFs) 559
Adrenal cortical steroid hormones 1252 – 1254, 1747
Adrenal gland, formation of noradrenaline 1064
Adrenaline (epinephrine) 534, 542, 553*s*, 999, 1431*s*, 1433, 1741– 1743, 1789, 1790*s*, 1791
  bound to adrenergic receptor 535, 555
α-Adrenergic receptors 553, 558, 563
β-Adrenergic receptors 553, 554, 1741
  in asthma 553
  in heart failure 553
  receptor kinase 553
  structure (proposed) 534, 555
  topology 555
Adrenergic synapses 1789
  agonists of 1791
  antagonists of 1791
Adrenocorticotropin (ACTH) 54*s*
Adrenodoxin 1065
Adrenodoxin reductase 764, 788
Adrenoleukodystrophy 945
Adsorption chromatography 103
Adsorption isotherm 326
Aequorin 314, 1342
Aerobes, obligate 8
Aerobic respiration, energy source 1013
Affinity chromatography 104 – 106
Affinity labeling 127, 477, 623
Affinity maturation 1862
Aflatoxin 1217, 1588*s*
African sleeping sickness 787
Agammaglobulinemia 572
Agar-agar 170
Agarose 177*s*
  activated 105
  double helices in 177
  gel networks of 175
Agglutination 1839
Agglutinins (lectins) 427, 1883
Aggrecan 181, 1154
Aggregation of cells 187
Aging 1906 – 1908
  cancer in 1907
  of nematodes 1907
  in yeast 1906
Aglycone 167
Agmatine 1440
Agonists 479
Agrin 437, 1154, 1781
Agrobactin 838
Agropine 1498*s*
AIDS 248, 1651, 1654, 1655, **1868**. *See also* HIV
Ajoene 1408
Alamethicin 414, 1774, 1842
Alanine (Ala, A) 52*s*, 1391, 1391– 1397
  fermentation of, scheme 1397
  pK_a values of 293
L-Alanine (Ala, A) 52, 52*s*
  binding constants to metal ions 311
  biosynthesis 517
β-Alanine 1383, 1453
  in coenzyme A 722
  pK_a of 293
D-Alanine 42*s*, 428, 429
Alanine thiosulfonate 1408

L-Alanyl-L-alanine
  pK_a value of 293
D-Alanyl-D-alanyl carboxypeptidase 625
N-β-Alanyldopamine 1434
L-Alanyl-L-valyl-L-methionine 57*s*
Alarmone 1454, 1636
Albinism 887
Alcohol
  addition to carbonyl 677
  dehydrogenation of **765 – 774**
Alcohol dehydrogenase(s) 479, 766 – 774
  isoenzymes 774
  stereospecificity of 479
  structure 773
  zinc in 680, 772 – 774
Alcoholic fermentation 966
Alcoholic liver cirrhosis 1378
Alcoholism 1797
Aldaric acid 164
Aldehyde(s)
  covalent hydrates 775
  dehydrogenation of 766
  oxidation of 775
Aldehyde dehydrogenases 776
Aldehyde oxidases 794, 890
Aldehyde oxido-reductase 891*s*
Aldimines. 434s *See also* Schiff bases
  from pyridoxal phosphate **740*s* – 742*s***
Aldo-keto reductases 774
Aldolase(s) 527, 679, **699 – 703**
  Schiff bases in 732
  type I 699
  type II 699, 700
Aldolase A
  enzyme-substrate complex 699
Aldol cleavage 698, 742
  in epimerization 778
Aldol condensation(s) 527, 699, 700
  drawing 529
  sugar formation mechanism 1136 – 1138
Aldonic acid 164
Aldose-ketose interconversions 692
Aldose reductase 774
Aldosterone 1253, 1256*s*, 1261
Algae 20 – 22
Algin 21
Alginates 170, 178
Alginic acid 1130
Alkali metal ions **206**
Alkaline phosphatases 645, 1491
  active site of 645
  occurrence 645
Alkaloids 1443, 1444, 1447– 1449
  formation, scheme 1444
  from tryptophan 1447– 1449
Alkanes 382
  hydroxylation of 1057
Alkaptonuria 1430
Alkenes, hydration of 683
Alkoxy-phospholipid 1200*s*
Alkylated bases 1583
Alkyl cobalamins 869, 870
Alkyl cobalt 867
Alkyl ethers 382
Alkyl radicals, table of names 380
Alkyl shift, rearrangements with 527
Allantoate 1368, 1368*s*
Allantoic acid 1421, 1458*s*, 1460
Allantoin 1368, 1368*s*, 1421, 1458*s*, 1459
Alleles 184
Allelic exclusion 1858, 1862
Allene oxide 1212
Allene structure 1240
Allergens 1862, 1863
Allergy 1862 – 1864
Allicin 1408
Alliin (garlic) 1408, 1408*s*
Alliinase 742, 1408
Allolactose 168, 1604
Allophanate 1378, 1378*s*
Allophycocyanins 1305
Allopurinol 1074, 1459, 1459*s*
Allose (All) 163*s*
Allosteric activation 475 – 477, 539 – 541
Allosteric constant 476
Allosteric control
  of enzymes 475 – 477, **539 – 541**
  Monod-Wyman-Changeux model 349
  of repressor protein 539
Allosteric effectors 357, 475 – 477, 536
Allosteric inhibition 475 – 477, 539 – 541
Allosteric sites 473
Allostery (allosterism) 475
Alloxan 121
Alloxanthin 1240, 1240*s*
Allylic rearrangement 527, 712
  with condensation 527
Allysine 429
Alpha-carbon plot 64
Alpha effect on reaction rates 590
Alpha oxidation. *See* Oxidation, alpha
Alport disease 438
Alternan 175, 1152
Alternative oxidase, in plants 1023
Alternative splicing of mRNA 538
Altman, Sidney 84
Altrose (Alt) 163*s*
Alu family 1538
Aluminum
  hydroxide complex, soluble 310
  toxicity of **658**
Alzheimer disease 1718, 1812, 1902
Amadori rearrangement 169, 1427, 1460
Amanitins 1618, 1625, 1627
Amantadine 1655*s*
Amber mutations 1606
Ameba 18
Amebic dystentery 187
Amebocytes 23, 25
Ameloblasts 442
Amelogenins 442
Amicyanin 817, 883
Amides
  displacement reactions of 608
  linkages in proteins 55 – 57
  mechanism of formation 661, 662, 1672, **1698 – 1709**
  resonance energy 299
  vibrational frequencies of 1277
  weak basic properties 56
Amidination 124
Amidino transfer 1379
Amidotransferases 1368, 1376
Amiloride 421*s*
Amine(s)
  addition to carbonyl 677
  dehydrogenation of 766
Amine oxidases 766, 782
Amino acid(s) **51– 55**
  abbreviations of, table 52, 53
  alpha (s) 41
  aromatic, light absorption 749
  branched chain 1391– 1397
  catabolism of 1393 – 1397. *See also* Specific amino acids
  configurations **42**
  C-terminal 57
  dehydrogenation of 775
  as dipolar ions 42, 51
  essential 52, 53, 1369
  glucogenic 1397
  ketogenic 1397
  metabolism of, **1359** – **1410**
  N-terminal 57
  pK_a values of 52, 53, 55
  residues (definition) 57
  separation by ion exchange chromatography 104
  sequences of 57
  side chain structures 55
    table 52, 53
  uptake by cells 1369
Amino acid decarboxylase 744
Amino acid dehydrogenase 766
Amino acid oxidases 782
D-Amino acid oxidase 478, 790, 791*s*, 1370
  of kidney 781
Amino acid sequence 57– 59
  of aspartate aminotransferase 57, 202
  of insulin 118
  methods of determination 118, 119
Amino acid side chains
  torsion angles of 61, 62
Aminoacrylate 744*s*
Aminoacyl (A) site in ribosome 1672
Aminoacyladenylate 1672, 1695, 1696
Aminoacyl-tRNA
  codon-specific binding of 1702
  modification reactions of 1697
Aminoacyl-tRNA synthetases 662, 994, **1694 – 1698**, 1696*s*
  action 1696
  functions of 1698
  proofreading 1695
  reaction mechanisms 1695
  recognition of tRNA 1694
  structures 1694
α-Aminoadipate 1386*s*
α-Aminoadipate pathway 1383
α-Aminoadipate semialdehyde 1386*s*
para-Aminobenzoate 473*s*, 1422*s*, 1425 – 1427
para-Aminobenzoate synthase 1425
Aminobutanol 1382*s*
4-Aminobutylcadaverine 1380*s*
γ-Aminobutyrate (GABA) 744, 1785, 1786, 1788
  antagonists of 1788
  concentration in brain 958
  cycle 958, 959
  drugs binding to receptors 1788
  shunt reactions, scheme 959
Aminocyclopropane carboxylate (ACC) 744*s*, 1390
Aminoglycoside 1426*s*
  antibiotics 1139
Amino group reactions 123
α-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid. *See* AMPA
5-Aminoimidazole-4-carboxamide 1449*s*
5-Aminoimidazole ribotide (AIR) 1464*s*
β-Aminoisobutyrate 1453*s*
α-Aminoisobutyric acid 69
δ-Aminolevulinate 1400*s*
5-Aminolevulinate dehydratase 1400*s*
δ-Aminolevulinic acid 742, 745*s*, 1399

  biosynthesis 742, 745
  formation from glutamate-1-semialdehyde 742
Aminomutases (vitamin B_{12}) 872, 874
8-Amino-7-oxo-nonanoate 718, 742
Aminopeptidase(s) 117, 609, 627
β-Aminopropionitrile 438
Aminopterin, inhibition of dehydrofolate reductase 805
Aminotransferases 737, **740 – 743**, 747– 753
  half reactions 737
3-Amino-3,4,6-trideoxyhexose 1138
5-Aminovalerate 1374
Ammonia
  binding constants to metal ions 311
  excretion of 1378
  incorporation into amino acids 1367–1372
  pK_a value of 99, 293
Ammonia oxidation system
  diagram 1052
Ammonium ions 1359, 1366, 1367
Amnesia 1801
Amniocentesis 1172
Amoeba proteus 19
AMP (Adenosine 5'-phosphate) 203, 1455*s*
  in adenylate system 302 – 304
  dissociation as acid 288
  pK_a value of 293
  as protein handle 719
  receptors 1786
AMPA 1786, 1786*s*
AMPA receptors 1786, 1903
Amphetamine (benzedrine) 1791*s*, 1796, 1810
Amphibolic sequences 952
Amphikuemin 1759, 1759*s*
Amphipathic character 380, 401
Amphotericin 1216*s*, 1217
Ampicillin 1164 1164*s*
Amplification
  of DNA 260, 1881
  by polyploidy 1881
  of signals 567
Amygdalin 1442, 1442*s*
α-Amylase **605 – 608**
  action patterns of 606
  calcium and chloride ions in 606
  catalytic groups 603
  pancreatic, 606s
  pH profile of 471
  subsites 606
β-Amylase 603, 608
Amylin 1719
  cause of diabetes 1004
Amylo-1,6-glucosidase/4-α-glucanotransferase 1144
Amyloid diseases 1718 – 1720, 1812 – 1814
Amyloid precursor protein (APP) 1812
Amyloid β-protein (Ab) 1812
Amylopectin 165, 170, 172, 173*s*, 1144*s*
Amyloplast 1144
Amylose 165, 170, **172, 173**, 1144
  double-helical form 173
Amylo-(1,4 → 1,6)transglycosylase 1143
Amyotrophic lateral sclerosis (ALS) 1789, 1812
Amytal 1021*s*
  inhibition of electron transport 1020
Anabolism. *See* Biosynthesis
Anacardic acids 1213, 1216*s*
Anaerobes
  facultative 8
  obligate 8
Anaerobic nucleotide reductases 864
Anaerobic respiration 1013, 1054 – 1057
Analgesic agents 1797
Anandamide 1758, 1794*s*, 1795
Anaphase 1501, 1503
Anaphase-promoting complex (APC) 1503, 1504
Anaphylaxis 385, 1863
Anaplerotic pathways 952
Anatoxin-a 1776, 1783*s*
Anchors, membrane for proteins 402
  diphytanylglyceryl 402
  glycosylphosphatidylinositol (GPI, phosphatidylinositol-glycan) 403*s*, 523, 1722
  myristoyl 402, 522
  palmitoyl 522
  polyprenyl 402
Androgenic (male) sex hormone 1254
Androgens 1254
  anabolic effect 1255
Androstenedione 1254, 1256*s*, 1259, 1260*s*
Anergy 1850, 1859
Anesthetics 1788, 1789
Angel dust (phencylidine) 1798*s*
Angioblasts 1902
Angiogenesis 1902
Angiogenic factors 1755
Angiogenin 648, 1755, 1902
Angiosperms 29
  embryo development, diagram 1905
  life-cycle, diagram 1905
Angiotensin II 563, 1261
Angiotensin-converting enzyme 625, 1261
Anharmonic motion 1276
Anhydrobiotic organisms 168
1,5-Anhydro-D-glucitol 165*s*
Anhydroretinol 1242*s*
1-Anilinonapthalene-8-sulfonate 1294*s*
Animal cells, surroundings of 431– 443
Animal embryo development 1893
Animal pole in egg 1897
Animal tissue, composition of 31
Animal-vegetal axis 1899
Anisotropy (R) 1291
  equation (polarized light) 1291
Ankylosing spondylitis 1854
Ankyrin 405
Annelida (segmented worms) 24
Annexins 316, 422, 571
Anomeric carbon atom 162
Anomeric effect 166
Antagonists 479
Antamanide 414, 1625
  Na^+ complex 415*s*
Antennapedia gene 1900
Anterior-posterior axis 1898
Antheroxanthin 1319
Anthocyanidins 1214, 1215*s*
Anthocyanins 1214
Anthranilate synthase 1425
Anthranilic acid (anthranilate) 747*s*, 1422*s*, **1425 – 1427**
Anthrax 548
Antianxiety drugs 1788, 1789
Antibiotic(s) 414, 479, **1164 – 1166**, 1616, 1618
  aminoglycoside 1690
  anthracycline 223
  inhibition of protein synthesis 1689 – 1692
  macrolide 1690
  steroid 1266
  tetracyclines 1691
  thiopeptide 1691
Antibiotic resistance 1166
Antibodies 1831. *See also* Immunoglobulins
  catalytic 1842
  monoclonal 1841
Anticoagulation systems in blood 634
Anticodon 231, 1672, 1687
  pairing with codon 1693, 1694
Anticodon-dihydrouridine stem loop 1692
Anticodon loop in tRNA 1688
  hypermodified base in 231
Anti conformation of nucleosides 211, 212
Anticooperativity 352
  in binding 329, 497
    to metal ions 307
    of protons 330
Antidotes for organic phosphates 636
Antienzymes 539
Antifreeze proteins 191
Antigen(s) 427, 1130
  presentation **1855, 1856**
    figure 1857
  receptor 1840
O-Antigen(s) 428, 1161
  of *Salmonella* 180
Antigen-presenting cells (ACPs) 1831, 1851, 1857
Antigenic determinants 427
Antihemophilic factor 632
Antihistamines 1863
Antiidiotype set 1858
Antimetabolites 473
Antimycin 1029
Antimycin A 1021*s*
Antioxidants 818 – 820, 822 – 824, 826, 1075
  vitamin E 721, 818, 819*s*, 822, 823
Antioxidant systems, of cells 1074
Antiparallel β structures 62
α_2-Antiplasmin 634
Antiport process 410
Antisense RNA 1614
Antitermination proteins 1616
Antithrombin 634
Antitrypsin 58
Antiviral compounds 1654
Antizyme 1382, 1713
AP-1 (Activator Protein) 576, 577, 1633
Aphidocolin 1561, 1561*s*
Aplasmomycin 1151
Apoferritin 842
Apolipoprotein(s) 1181– 1185
  functions of 1182
  properties of, table 1183
Apolipoprotein A-I 1182
Apolipoprotein B-100 1182
Apolipoprotein E 1182, 1814
Apomixis 1904
Apoptosis 574, 619, 1881, **1888 – 1890**, 1902
  in mammalian cell, figure 1889
Apoptosome 1890
Apoptotic systems
  components of, table 1889
Appican 1813
Apple motif of proteins 367
Apurinic acid 250
Apurinic/apyrimidinic DNA endonucleases 1581
Apyrimidinic acid 255
Aquaporins 411, 412
Aquired systemic resistance 1440

- Ara-A 1655*s*
- *Arabidopsis thaliana*
    - genome 12, 29, **1511, 1512**
- Arabinan 177, 1168
- Arabinogalactan 431, 1150
- D-Arabinose (Ara) 163*s*, 1138
- L-Arabinose (Ara) 163
    - binding protein 418
- L-Arabinose isomerase 693
- Arabinose-phosphate 1136
- Arabinose utilization operon 1613
- araCBAD operon
    - regulatory region, scheme 1614
- Arachidonate 1863
- Arachidonate cascade 1206
- Arachidonic acid 381*s*, 565, 1190*s*, 1206*s*
- Arachidonoylglycerol 1794*s*, 1795
- Arachidonoyl glyceryl ether 1794*s*, 1795
- Arachidonoylphosphatidylethanolamide 1794*s*
- Arachnoid membrane 1765
- Aragonite 440
- ArcA 1614
- Archaea 8
- Archaeobacteria, thermophilic 7, 815
    - CO_2 and sulfur, sources for life 1057
- Archaeosine 1693
- Archenteron 1898
- Area detector for X-rays 134
- Arg-Gly-Asp (RGD) sequence 408
- Arginase 888, 1378
    - manganese in 888
- Arginine (Arg, R) 53*s*, 745, 1680
    - biosynthesis of, scheme 1375
    - catabolism of 1378
    - insecticidal analogs of 1379
    - synthesis of **1374 – 1379**
- Arginine dihydrolase pathway 1378
- Arginine kinase 656
- Arginine phosphate 303
- Arginine vasopressin 1748
- Argininosuccinate 1376, 1377, 1378*s*
- Argininosuccinate lyase 685, 1378
- Argininosuccinic aciduria 1378
- Arginyl-tRNA protein transferase 1728
- Aristolochene 1234, 1235*s*
- Aristolochene synthase 1234
- L-Arogenate (pretyrosine) 1422*s*, 1425, 1425*s*
- Aromatase 1260
- Aromatic clusters in proteins 75
- Aromatic compounds
    - catabolism, bacteria, scheme 1436
    - catabolism, microbial 1434
    - metabolism of, **1421– 1464**
- Aromatic groups, formation by elimination 689
- Arrestin 1331, 1792
- Arrhenius activation energy 483
- Arrhenius equation 483
- Arrhenius, Svante A. 9, 83
- Arrows, drawing 528, 529
- Arsenate
    - phosphorylation, uncoupling of 596
    - similarity to phosphate 596
- Arsenic compounds 387, **596**, **597**
    - toxicity of 596
- Arsenicals in treatment of syphilis, 597
- Arsenite, reaction with thiols 596
- Arsenolysis 596
    - of glucose 1-phosphate 595
- Arsenophospholipid 387*s*
- Arseno-ribofuranoside 596*s*
- Arthritis 1854
- Arthropods 24
- Aryl azides 1297
- Arylsulfatase A, B 1171
- *Ascaris* 25
    - chromosome number of 18
- Ascarylose (3,6-dideoxy-L-mannose) 180*s*
    - from CDP-glucose 1138
- Ascomycetes 20
- Ascorbate antioxidant 1066, 1075
- Ascorbate oxidase 887
- Ascorbate peroxidase 852
- Ascorbic acid 721, 1066s, **1066, 1067**, **1132 – 1135**, 1134s
    - in adrenal gland 1064
    - biosynthesis in plants 1134
    - and common cold 1067
    - dietary requirement 1066
    - free radical form 1066
    - healing of wounds, effect on 1066
    - metabolism of **1132 – 1135**
        - scheme 1133
    - metal complexing properties 1066
    - pK_a 1066
    - nutritional requirement 756
- Ascus 20
- Ash, amount from tissue 31
- Asialoglycoprotein (Gal) receptor 1160
- Asparagine (Asn, N) 53*s*, 1383
- Asparagine synthase 1368
- Aspartame 1800*s*
- Aspartase 526
- Aspartate 737s, 1378, 1793, 1794. *See also* Aspartic acid
    - binding constants to metal ions 311
    - biosynthetic reactions 711, 1383 – 1391
        - scheme 1384
    - effect on biotin requirement 724
    - formation by transamination 952
    - metabolism of 539
    - residues 312
- Aspartate aminotransferase 57*s*, 135*s*, 742 – 744, 747– 753
    - absorption spectra 749
    - active site structure 744
    - atomic structure 750
    - catalytic intermediates, models 752
    - NMR spectra 149
    - polarized spectra 1283
    - quinonoid intermediate 750
    - Ramachandran plot 61
    - sequence 57
    - transamination 742
- Aspartate ammonia-lyase 685
- Aspartate carbamoyltransferase 348*s*
    - active sites 348
    - difference spectrum 1285
    - regulation 540
- Aspartate α-decarboxylase 753, 755
- Aspartate β-decarboxylase 746
- D-Aspartate oxidase 1370
- Aspartate racemase 741
- Aspartate semialdehyde 1384*s*
- Aspartate semialdehyde reductases 1383
- Aspartic acid (Asp, D) 52, 53*s*
    - biosynthesis 517
    - pK_a value of 293, 487
- Aspartic proteases **621– 625**
- Aspartokinases 1383
- β-Aspartyl adenylate 1368*s*
- Aspartyl aminopeptidase 621
- Aspartylglucosaminidase 1170, 1171
- Aspartylglycosaminuria 1171
- β-Aspartyl phosphate 539, 540*s*, 1384*s*
- Aspirin 1211, 1211*s*, 1440
- Assays of enzyme activity 456
- Assembly core of virus shell 365
- Assembly pathway
    - Bacteriophage T4 367
    - for virus ϕX174 366
    - of 30S ribosome 1682
- Assimilatory nitrate reductases 1366
- Association constant. *See* Formation constant
- Astacin 627
- Astaxanthin 1240, 1240*s*
- Asthma 26, 385, 1862, 1863
- Astrocytes 1781
    - micrograph 1778
- Asymmetric unit of crystal 134
- Asymmetry (handedness) of molecules 41– 43
    - in oligomers 344
- Ataxia 1812
- Ataxia telangiectasia (AT) 566, 574, 1585
- Ataxin-2 1702
- Atheromas 1249
- Atherosclerosis 1249, 1250
- Atmungsferment 1022
- Atomic force microscope diagram 131
- Atomic force microscopy 131, 179, 392
- Atomic structures of macromolecules. *See* Three-dimensional structures of
- Atoms, table of sizes **41**
- Atopy 1863
- ATP (Adenosine triphosphate) 536, **719**, **720**, **1043 – 1046**, 1672, 1752, 1776, 1794
    - in acetylation of sulfanilamide 720
    - in adenylate system 302 – 304
    - analogs of 655
    - binding constants to metal ions 311
    - in biosynthesis **973 – 978**
    - in chloroplasts 1318
    - cleavage 977
    - as coenzyme 719
    - complexes with metal ions 296
    - coupling of hydrolysis 977
    - discovery of 1022
    - dissociation as acid 288
    - exchange-inert complex 643
    - formation during fermentation 967
    - formation during oxidation of aldehyde, scheme 775
    - formation by substrate oxidation 962
    - hydrolysis of 507
        - enthalpy of 296
        - Gibbs energy of 293, 297
    - pK_a values of 293
    - proteases, ATP dependent 627, 628
    - proton pumps 1045, 1046
    - in regulation 535
    - screw-sense isomer 643
    - synthesis of 303, **507– 515**
- ATP, biosynthesis of 303, 1014 – 1016, 1032 – 1046
    - in fermentation reactions 661
    - uncoupling 636, 1046, 1047
- ATP-citrate lyase 703
- ATP cleavage in Calvin – Benson Cycle 984
- ATP sulfurylase 1056
- ATP synthase 425, 514, 659, 1014, **1041– 1045**, 1042*s*, 1093
    - Boyer's binding change mechanism 1044, 1045
    - chloroplast 1039

  genes 1041
  mechanism 1043, 1044
  model of 1042
  rotational catalysis 1044
  structure 1041 – 1044
ATPase 638, 1042*s*
  Ca^{2+} 424, 425*s*
  gastric H^+, K^+- 424
  H^+, K^+- 424
  Na^+, K^+- [(Na^+ + K^+)-activated] 423
  P-type 423
  traffic 417
  vacuolar 425, 1046
Atractyloside 1049
Atrazine 1263, 1299*s*
Atrial natriuretic factor (hormone) 561, 1262, 1753
Atropine 1783*s*, 1784
Attenuation (transcription) 1615
  mechanism 1393
  trp operon 1617
Aurocyanin 883
Autoantibodies 1641
Autoantigens, human 1159
Autocatalytic reaction 470
Autocrine effect 385, 574
Autoimmune diseases 1856, **1864 – 1867**
  table 1864
Autoinducers 1891
Autonomic system 1767
Autonomously replicating sequences (ARSs) 1562
Autophagosome 1728
Autophagy 1728, 1888
Autophosphorylation of receptor 562
Autoradiography 1543
Auxilin 427
Auxins 1390, 1446, 1447*s*, 1761
Auxotrophs
  nutritional 1478
Average oxidation state
  of compounds, scheme 983
Avermectin 1217, 1788*s*
Avery 1473
Avidin 257, 728
Avirulence (Avr) proteins 1869
Avogadro's number 283
Axial element 1506
Axis (axes) of symmetry 322, 323, 327– 338
Axon 11, 1763, 1903
Axonal transport 1119, 1776
Axon hillock 1763
L-Azaserine 1371, 1371*s*
Azide, as respiratory inhibitor 1021
Azoferredoxin 1361
Azotoflavin 1362
Azurins 883

B

Bacillus, definition of 6
Bacillus subtilis
  genome 12
Bacillus thuringensis
  as insecticide 1868
Baclofen 1787, 1788*s*
Bacteria 2. *See also* Specific genus and species
  acetic acid 8
  aerobes 10
  anaerobic 8
  autotrophic 8
  binding to cells 186
  branched fatty acids of 381
  chemoheterotrophic 7, 8
  chemolithotrophic 7
  chemotaxis **1093 – 1096**
  classification of **6 – 8**
  coats 431
  composition of 31
  denitrifying 1054, 1359
  electron micrograph of 4
  flagella of 6, 1089 – 1093, 1090*s*
  gram-negative 6, 7, 180
    transporters in 416 – 420
  gram-positive 6, 7, 185, 190, 428, 1161
  green sulfur 1306
  hydrogen-oxidizing 1051
  invasion by 1120
  largest 8
  L forms 6
  methane 1057
  methanogens 7
    coenzymes of **813 – 815**
  methylotrophic 985
  motility **1089 – 1095**
  nitrifying 1051, 1052
  nonsulfur 1306
  nutrition of 8
  photoautotrophic 8
  photosynthetic 7, 9, 982, **1298 – 1320**
  pili 335
  purple photosynthetic 1306
  purple, reaction centers of **1310 – 1315**
  sensing 1798
  signaling among 1891
  with stalks 1890
  sulfate-reducing 1056, 1057
  sulfur-oxidizing 1052 – 1054
  sulfur-reducing 1056, 1057
Bacterial cell walls
  biosynthesis 1160 – 1168
Bacterial dehydratases
  Fe-S clusters in 861
Bacterial diseases 7
Bacterial operon(s)
  control of 1611– 1613
  inducible 1612
  repression 1611, 1612
Bacterial sex factors 1482, 1483
Bacteriochlorophylls 1303*s*, 1304, 1307, 1308, 1310 – 1313, 1311*s*
  biosynthetic pathways 1403
Bacteriocins 1842
Bacteriohopanetetrol 1244, 1245*s*
Bacteriophages
  antibacterial medicine 1167
  chromosome mapping 1477
  filamentous (fd, f1, M13) 129, 334, 335, 1494
  helical 244
  mutant 1477
  recombinant 1478
  temperate 248, 1483 – 1486
Bacteriophage ϕ11 plaques, figure 1477
Bacteriophage ϕX174 244
Bacteriophage MS2 RNA model 1714
Bacteriophage Mu 1576
Bacteriophage P22 tailspike protein 66, 66*s*
Bacteriophage Plkc, transducing 1479
Bacteriophage T4, linkage map 1481
Bacteriophages, T-even 244, 363
Bacteriophage T4(s) 363*s*
  assembly pathway 367
  lysozyme 599
Bacteriorhodopsin 71, 1039, 1325, **1333 – 1337**
  photoreaction cycle of 1335
  structure 1327, 1334
Bacteroids 1360
Bactobilin 1405
Bactoprenol 1152
Baculovirus 247, 1498
Baculovirus vehicles 1498
Baltimore, David 84
Band 3 protein 404
  of erythrocytes 420
Barbital 1788*s*
Barbiturates 1788, 1789, 1796
Barrels
  α,α 71
  α/β 77
  β 65
Basal body of flagella 15, 1089
  electron micrograph 1091
Basal ganglia 1765
Basal metabolic rate 283
  relationship to surface area 283
Basal stem cells 439
Base(s)
  of nucleic acids 199
  strengths of 95, 96
Base catalysis, general 487– 491
Base composition of DNA and RNA 239
Base pairs **207– 211**
  base inclination 217
  of DNA 200
  Hoogsteen 207, 231
  propeller twist 217
  strength of 209
  tautomerism 211
  thermodynamics of 209
  Watson–Crick 207, 231
  wobble 209, 231
Base quartets 207– 211
Base triplets 207– 211
Basement membrane 15, 72, 431, 436
Basic fibroblast growth factor (bFGF) 1903
Basidiomycetes 20
Basophils 26, 188
  in blood 1863
Bathochromic shift 1283
Bathorhodopsin 1329, 1330
Batrachotoxin 1265, 1265*s*, 1775
B cell. *See* Lymphocytes
*bcl*1 (cyclin D1) 572
Bdellovibrio 3
Beer–Lambert law 1275
Beggiotoa
  sulfur precipitates from 1053
Behavior 1805
Bence-Jones proteins 1836, 1837
β Bend **72 – 74**
Benzaldehyde, resonance energy 299
Benz(a)anthracene 1588
Benzene, resonance energy 299
Benzoate 1437, 1437s, 1439s
Benzoic acid 1438
  acidity of 308
  pK_a value of 309
Benzophenones 1297
Benzo(a)pyrene 1588
Benzoquinone 1438
Benzoylformate 1437*s*
Benzoylformate decarboxylase 734

Berenil 225*s*
Bergmann glia 1781
Beriberi 721
Beryllium, toxicity of 299
BES buffer 99
Beta carotene, as antioxidant 1075
Betacelulin 1753
Beta cleavage. *See* Cleavage, beta
Beta condensation. *See* Condensation, beta
Beta helix. *See* β Helix
Betaine 1388, 1389*s*
Beta oxidation 511– 513. *See also* Oxidation, beta
  of fatty acids 511, 512, 939 – 944
  Knoop's theory **513**
  pathway **506**
Beta propeller 67, 560, 764
Beta ribbons 241
Beta sheet 64
Beta spiral 74
Bicarbonate ion
  from carbon dioxide 677
  in carboxylation 710, 711
  incorporation 710, 711
Bicarbonate sensor 557
Bicinchoninic acid 102
BICINE buffer 99
Bicuculline 1787, 1788*s*, 1789
Biglycan 1154
Bile acids
  formation of, scheme 1252
  metabolite of cholesterol 1251
Bile pigments 1404
Bile salts 1251
Biliary cirrhosis 1865
Bilin 1304
Bilin pigments, table 1305
Bilirubin 1404, 1404*s*
meso-Bilirubinogen 1404
Biliverdin 852, 1404, 1404*s*
Binding
  cooperative processes 330 – 332
  equilibria **325 – 332**
  by macromolecules 327– 330
  of metal ions **307**
  multiple sites of 327– 330
  ordered 464
  of oxygen. *See* oxygen (O_2), binding of
  random order 464, 467
Binding constants 325 – 327. *See also* Formation constant
  anticooperativity 329, 330
  intrinsic 328
  for metal complexes, table 311
  microscopic 328, 329
  of protons to dianions, table **328**
Binding curves 326
  sigmoidal 331
Binding energy
  relation to maximum velocity 617
Binding equilibria 325 – 332
  dimerizing protein 350 – 352
Binding proteins 418
  for histidine 419*s*
  for hormones 536
  single-strand 1549
Binding sites, multiple 327
Bioactivation of insecticides 636
Biochemical equilibria **304 – 316**
Biochemical reactions
  table of types **526**, **527**
Biocytin 724, 724*s*
Biodegradative threonine dehydratase 1391
Biological recognition 29
  glycoproteins in 186
Bioluminescence 1340 – 1346
Biomineral precursor 843
Biopolymers, hydrolysis of 292
Biopterin 803, 804*s*, 1460
  as hydroxylation coenzyme 803
Biosensors 122
Biosynthesis 507, **515 – 517**, **972 – 996**. *See also* Specific compounds
  activated groups used in, table 975
  carbonyl groups in 982
  of fatty acids 990
  of hydrocarbon chains 990
  of isoprenoid compounds 689
  parallel with biodegradation 973
  reducing agents for; use of NADPH 978 – 982
  regulation of 996 – 1005
  scheme 506, 509
Biosynthetic pathways
  committed step 535
  feedback inhibition 535
  relationship to catabolism 973
  schemes 506, 509, 974
Biosynthetic threonine dehydratase 1391
Biotin 515, 516, 711, 721, **723 – 730**, 723*s*, 950, 1393
  biosynthesis of 718, 745
  in enzymes, table 724
  mechanism of action 725 – 729
  nutritional requirement 756
Biotin-binding proteins **728**
Biotin carboxylase 724
Biotin carboxyl carrier protein 724
Biotin holoenzyme synthetase 724
Biotinidase 725
Bipolar disorder 1810
2,5-Bisoxo-D-gluconate 1134*s*
Bisoxogulonic acid 1133*s*
  from dehydroascorbic acid 1134
1,7-Bisphosphatase 1320
1,3-Bisphosphoglycerate 596*s*
  in glyceraldehyde-3-phosphate dehydrogenase 775
2,3-Bisphosphoglycerate 357*s*, 361
  effect on hemoglobin 357
Bisphosphoglycerate mutase 654
BIS-TRIS buffer 99
Bisulfite 779
  effect on yeast fermentation 967
  reaction with uridine 251
    with thiamin 731
Biuret 102*s*
Biuret method 102
BK virus 247
Black membrane 393
Blanching hormone 1760
Blastocyst 1897
Blastula 1897
Bleomycin 225*s*, 1589
Bleomycin hydrolase 619
Blood
  carbonic anhydrase 678
  cell(s) 26, 110, 1900, 1901
  complement system 58, 414
  glucose, control of 997– 999
  lipid carriers of 1184
  pressure, regulation of 1261, 1262
  proteins of **58**
Blood-brain barrier 29, 1765
Blood cells
  development of 1901, 1902
  globin genes 1901
  types of 110, 1834, 1835
Blood-cerebrospinal barrier 1765
Blood clotting **631– 634**
  anticoagulation systems 634
  cascade mechanism 632
  cofactors 631
  inhibition of 622
  intrinsic mechanism 633
Blood glucose 997– 999, 1742
Blood group determinants 182, 183, 1155
  ABO system 184, 185
  Cad antigen 185
  Kell antigens 185
  Lewis blood group 184, 185
  MN antigens 185
  P group 185
  Rh antigen 185
Blood plasma
  electrolyte concentrations 206
  proteins **58**
Blood platelets 26, 384
Blood type. *See* Blood group determinants
Bloom syndrome 1550, 1585
Blue denim 1448
Blue-green algae. *See* Cyanobacteria
Blue light responses 1338 – 1340
Blue multicopper oxidase 887
Bohr effect 357
Boltzmann constant 283, 284
Bombardier beetle 1048
  reaction vessel (figure) 1049
Bombesin 1750
Bombyxins 1760
Bond
  angles **39 – 40**
  energies, table 299
  lengths **40**
Bonding
  heterologous 333, 334
  isologous 338
  nonsymmetric 344
Bone 26, 440, 441
  development of 1902
  morphogenic proteins 1902
Bone-derived growth factors 1755
Bone morphogenic factor 443
Bongkrekate 1049
Bordetella 548
Boric acid, 99
Borneol 1233*s*
  biosynthesis 1234
Bornyl diphosphate 1234*s*
Borohydride
  in carbohydrate analysis 190
  reduction by 124, 190
Boron, functions of 1151
Boronic acids 622
Botulinum toxin 1776
Bouton 1777
Bovine spongiform encephalopathy (BSE). *See* Mad cow disease
Boyer, P. 1044
Boyer's binding change mechanism
  ATP synthase 1044, 1045
Bradford method 102
Bradykinin 1262, 1752
Brain 1804
  alpha waves 1802
  conscious thought 1802

  - human (figure) 1742
  - imaging 1801, 1802
  - interconnections of, diagram 1766
  - limbic system 1767
  - lipid mediators 1795
  - mesolimbic dopamine system 1796
  - of molluscs, organization 24
  - nucleus accumbens 1797
  - organization of 1765, 1766
  - peptides 1749, 1750
  - PET image 1806
  - pleasure centers 1767
  - punishing centers 1767
  - rat, diagram 1766
  - reticular system 1767
  - rhythms 1802, 1803
  - stem 1765
  - temporal lobe 1801
  - transporters in 416, 417, 420
  - ventral tegmental area (VTA) 1797
  - visual cortex 1798
- Brain-derived neurotrophic factor (BDNF) 1812
- Branch migration in DNA 229
- Branched carbon chains
  - formation of 992, 993
    - scheme 992
- Branched chain fatty acids 942
- Branching enzyme 994
  - deficiency of 1144
- Brassinolide 1246, 1248*s*
- Brassinosterols 1761
- Breast cancers 572
- Brefeldin A 1159
- Brevetoxin 1770*s*, 1775
- Brevican 1154
- Brine shrimp 168
- Bromelain 618
- Bromoperoxidase(s) 856
- Brönsted equation in basicity 490
- Brönsted plot 309, 490
- Brown, Michael S. 84
- Brown fat 1048
- Brownian movement 461
- Brucellosis 7
- Bryophyta 29
- Bt toxin 1868
- Buchner, Edward 83, 456, 960
- Buffer(s) **97**, **98**
  - redox 980
  - table 99
  - volatile 98
- Buforins 1843
- Bufotenin 1811
- Bufotoxin 1265*s*, 1266
- β Bulge 66, 67*s*
- Bundle sheath 1322
- Bungarotoxin 1776, 1785
- Burkitt's lymphoma 572, 576
- Bursin 1756
- Butadiene 1283
- Butanol-forming fermentations 971
- Butyric acid, fermentation 971, 972
- γ-Butyrobetaine 1063*s*
  - hydroxylation of 1063
- Butyrylcholinesterase 634

C

- C1q 1839
- C_4 compounds, export from mitochondria 981
- C_4 cycle
  - for CO_2 concentration 1322
    - scheme 1323
- Cachetin. *See* Tumor necrosis factor (TNF)
- Cachexia 1756, 1849
- Cacodylic acid 99
- Cadaverine 1380*s*
- Cadherin 187, 407, 1883
  - domains 408*s*
- Cadmium 317
  - toxicity of 680
- cADP ribose (cyclic ADP ribose) 564*s*
- *Caenorhabditis elegans*
  - apoptosis 619
  - cuticle 440
  - development 1893
  - genome 12, 1507
  - number of cells 24
- Caffeine 1458*s*, 1459
- Cage effect 462, 463
- Caged phosphate 137
- Calbindins 313, 1258
- Calcein 1294
- Calciferol (vitamin D) 1257
- Calcineurin 646
- Calciosomes 314, 563
- Calcite 440
- Calcitonin 314, 1258, 1750
- Calcitonin gene-related polypeptide (CGRPP) 1750
- Calcitriol 1257
- Calcium **314**, **315**
  - fluorophores 1294
  - oscillations 315
  - release channel 314
  - in signaling **563 – 566**
- Calcium-binding proteins 312
- Calcium-binding sites 312
- Calcium carbonate 443
- Calcium ion(s) 31, 314, 315
  - activation of metabolic processes 314
  - concentration
    - in tissues 314
    - in serum 314
  - effects on equilibria 297
  - interaction with carboxylate groups 47
  - in muscle 1114
  - rapid uptake mode 1047
- Calcium ion channel(s) 422, 1772
  - dihydropyridine-sensitive 1772
  - release channels 1115
  - voltage sensitive 1114, 1772
- Calcium phosphate granules 1014
- Calcium uniporter 1047
- Caldesmon 1117
- Callose 170, 174, 1148
- Callus 1885
- Calmodulin 313, 315, 317*s*
  - in phosphorylase kinase 544
- Calnexin 188, 520
- Caloric values of food 282, 283
  - table 283
- Calorie, definition of 282
- Calorimeter
  - bomb 282
  - differential scanning 285
- Calorimetry
  - animal 283
  - differential scanning of membranes 394
- Calpactin 316
- Calpains 619
- Calreticulins 313
- Calsequestrin 313, 314, 1114
- Calvin–Benson cycle **985, 986**, 1319 – 1321
- Calvin, Melvin 84
- Cambium layer 29
- Cambrian explosion 9
- cAMP 536, 544, 544*s*, 556, 557
  - second messenger 557, 1743
- cAMP receptor protein (CRP or CAP) 1604, 1612
- Campesterol 1246, 1248*s*
- Camphor 1233*s*
- Camptothecin 1552, 1552*s*
- Canaline 1379, 1379*s*
- Canavanine 1379, 1379*s*
- Canavanine hydrolase 1379
- Cancer **573 – 575**, 1297, 1907, 1908
  - of breast 1263
  - chemotherapy 224, 812
  - p53 as suppressor 581
- Cancer cells
  - communicating junctions in 29
  - lack of contact inhibition 29
  - promoters 573
- Candida albicans 1891
- Cannabidiol 1794s
- Cannabinoid receptors 1795
- Canthaxanthin 1240, 1240*s*
- Capsaicin 1794*s*, 1800
- Capsid 244
- Capsomers 244
- Capsule 6
- Captopril 624, 1261*s*
  - for high blood pressure 1261
- Carbamate group. 80 *See also* Carbamino group
  - in ribulose bisphosphate carboxylase 707– 709
- Carbamate in proteins 80
- Carbamic acid 1376*s*
- Carbamino groups. 678 *See also* Carbamate group
  - in hemoglobin 358
  - in regulation 543
- Carbamoyl aspartate 1451*s*
- Carbamoyl phosphate 975*s*, 1376*s*
- Carbamoyl phosphate synthetases 1376, 1377
- Carbamoylation (carbamylation)
  - of ribulose bisphosphate carboxylase 707
  - by sodium cyanate 124
- Carbamylation. *See* Carbamoylation
- Carbanion mechanism of flavin dehydrogenase 790
- Carbapenems 1164
- Carbinolamine(s) 677
  - configurations 680
  - formation of 526
- 4*a*-Carbinolamine dehydratase 1061
- Carbocation(s)
  - from alkenes 683
  - oxocarbenium ion 598*s*
- Carbocation mechanism of fumarate hydratase 684
- Carbocationic intermediate 595

- Carbodiimide 105*s*
- Carbohydrate(s)
  - analysis 188
  - biosynthesis, 994, 1000, 1001, **1129 – 1172**
  - as biosynthetic precursors 982
  - caloric value 283
  - content, in tissues 31
  - hydrolysis of 189
  - insertion mechanism in growth 1152
  - metabolism, 1002, **1129 – 1172**
  - methylation of 189
  - NMR spectra of 190 – 192
  - periodate oxidation of 190
- Carbohydrate-binding proteins. *See* Lectins
- Carbon-14
  - Calvin–Benson cycle 985
  - properties of 110
- Carbon-carbon bonds
  - addition to 681 – 683
  - cleavage of 530, 982
  - formation 982
  - reduction of 777
- Carbon dioxide 358
  - as biosynthetic precursor 982 – 985
  - carbon units, source of 810
  - hydration of 677, 710
  - photosynthetic reduction of 977
  - reduction of 881
  - reduction, scheme 814
  - resonance energy 299
  - stability of 690
- Carbon monoxide 553, 1756, 1795, 1796
  - guanylate cyclase activation 1757
  - heme iron, binding 1757
  - in hemoglobin 358
  - oxidation of 881, 1051
  - resonance energy 299
  - as respiratory inhibitor 1021
  - vasodilation 1757
- Carbon monoxide dehydrogenase/acetyl-CoA synthase 881, 882
- Carbon monoxide oxidase 893
- Carbon skeletons, assembly of 982
- Carbonic acid 99
- Carbonic anhydrase 443, **676 – 678**, 710
  - active site structure 679
  - mechanism 678
  - turnover number of 458, 678
- Carbonium ion. *See* Carbocation
- 1,1'-Carbonyl-diimidazole 105*s*
- Carbonyl group(s)
  - addition of enolate anion 699 – 704
  - addition to 677
  - in chain formation and cleavage 982
  - in displacement reactions **608 – 637**, 698
  - facilitation of addition or elimination 681
  - in metabolic sequences 681
  - reactivity 608
  - transient formation of 777, 778
- Carbonyl oxide 1060*s*
- *N*-1'-Carboxybiocytin
  - in action of enzyme 725
- *N*-1'-Carboxybiotin 725, 726, 726*s*, 970
  - dimethyl ester of 725
  - intermediate 725
- γ-Carboxyglutamate 821, 821*s*
  - binding of calcium ions 80
  - formation of 820 – 822
- Carboxyl (carboxylate) groups
  - formation **723 – 730**
  - in proteins 312
  - reduction of 776
  - resonance 967
  - resonance energy 299, 967
- Carboxylation 978
  - beta 527, 727– 729
  - in biosynthesis **977– 980**
  - biotin-assisted 724
  - driven by cleavage of ATP 705
  - reductive pentose phosphate pathway 986
- Carboxylesterases 635
- Carboxylic acid
  - activation of 507
  - dehydrogenation of 766
- Carboxylic acid reductase 893
- Carboxyltransferase 724, 971
  - of propionic acid bacteria 725
- Carboxymethylation of proteins 594
- Carboxypeptidase(s) 117, 609, 625
  - zinc ion in 625
- Carboxypeptidase A 64*s*
  - active site 626*s*
  - pK_a values of 626
- Carboxypeptidase Y 610
- Carboxyphosphate 726, 726*s*
- Carcinogenic compounds
  - adducts with nucleic acids 235
- Carcinomas 573
- Cardiolipin 384, 1028, 1198, 1198*s*
- Cardiomyopathy 1113
- Cardiotoxins 1775
- Caricain 618
- Caries 442
- Carnitine 944*s*, 1063*s*, **1383 – 1386**
  - formation 1063
  - mitochondrial permeability 944
- Carotene(s) 1237, 1304
  - functions 1243
  - oxidized 1240
  - photoprotection 1243
  - from phytoene, scheme 1238
  - properties 1243
- β-Carotene 10, 390*s*, 1239*s*
  - antioxidant 1243
- Carotenoid(s) 1227, **1237 – 1243**, 1303, 1304, 1306, 1308, 1309, 1319
  - color 1243
- Carotenoid compounds
  - biosynthetic pathways 1239
  - structures 1239
- Carpal tunnel syndrome 756
- Carrageenan(s) 170, 177, 178*s*
- Carrageenin 22
- Cartilage 26, 436
  - articular 182
  - development of 1902, 1903
  - matrix protein 437
- Cartilage-inducing factor 1755
- Carvone 1233*s*
- Casbene 1235*s*, 1236
- Cascades
  - blood clotting 632
  - dephosphorylation 542
  - MAP kinase 576 – 578
  - pancreatic proteases 609
  - phosphorylation 542
  - regulatory 566, 567
- Casein kinase 544
- Caspase family 619, 1516, 1888, 1889
- Cassava 1442
- Catabolism 536
  - and ATP synthesis **507– 515**
  - pathways of 506, **509**
  - of propionate, scheme 948
  - of propionyl-CoA, scheme 948
  - of sugars 960 – 966
- Catabolite gene activator protein (CAP) 1604, 1613
- Catalase(s) 844, **852 – 855**, 853*s*
  - catalytic cycles, scheme 855
  - compounds I, II 853 – 855
  - function of 852
  - hereditary deficiency 852
  - iron properties 853
  - mechanisms of catalysis 853 – 855
  - molecular activity of 852
  - turnover number of 458
- Catalysis (acid, base) 469, 486
- Catalysis (enzymatic)
  - acid and base 469, **486 – 491**
  - covalent 494
  - mechanisms of **482 – 497**
  - microenvironment 495
  - microscopic reversibility 486
  - multiple attack 606
  - nucleophilic 494
  - proximity factor 495
  - stereoelectronic effect 495
  - strain and distortion in 496, 497
  - tautomeric 490, 491
  - transition state 482 – 486
- Catalyst, polyfunctional 490
- Catalytic antibodies (abzymes) 1842
- Catalytic cycles. *See also* Other specific enzymes and cell cycle
  - for catalase 855
  - for coenzyme M reductase 880
  - common features of 952
  - for copper-containing hydroxylases 1064
  - for cytochrome P450 1068
  - for glycine decarboxylase 809
  - for lysozyme 599
  - for α-oxoacid dehydrogenases 798
  - for peroxidases 855
  - for serine protease 613
  - for vitamin B_{12}-dependent isomerases 872
- Catalytic dyad 620
- Catalytic subunits 348
- Catalytic triad 134, **611– 614**, 619, 635
  - of serine proteases 611– 614
- Cataracts 169, 1131, 1333
- Catchin 1117
- Catechol 838, 1436*s*, 1438
- Catechol 1,2-dioxygenase 1057
- Catechol *O*-methyltransferase (COMT) 591, 592, 1434, 1790
- Catecholamine(s) 553, **1432 – 1434**, 1789, 1792
  - metabolism pathways 1790
- β-Catenin 1883, 1898
- Cathepsin(s) 619, 621
  - G 610
  - K, in osteoclasts 619
- Cations. *See also* Individual ions
  - active transport 422
  - channels 421, 422
- Cauliflower mosaic virus 247
- Cayenne pepper 1439*s*
- CD (Cluster of Differentiation) 187, 407, **1850**
- Cdc20 1504

CDK-activating kinase (CAK) 1502
CDK2 1502
CDP-diacylglycerol 1197, 1198*s*
Cech, Thomas R. 84
Cecropins 1843
CED-3 619
Celiac disease 1854, 1865
Cell(s)
  aggregation of 187
  columnar epithelium 26
  composition of **30 – 33**
  doubling time 470
  fractionation of 98 – 100
  junctions 28
  metal binding sites in 311, 312
  migrations 1884
  sizes of 11
  squamus epithelium 26
Cell adhesion molecules (CAMs) 406
Cell-cell recognition 1883, 1884
Cell constant organisms 24
Cell contacts 28, 29
Cell culture 26
Cell cycle **578 – 581**
  checkpoints 1502
  figure 580
Cell differentiation antigens. *See* CD
Cell division
  asymmetric 1879, 1884, 1885
  in bacteria 1885
  directionality (polarity) 1879
Cell membranes 3
  anions 420, 421
Cell migration 1883, 1884
Cell recognition 29
Cell respiration, inhibition 636
Cell signaling, table of abbreviations 577
Cell structure, compartments 552
Cell types
  animal 25
  plant 29, 30
Cell wall 5, 10, 15, **427– 441**
  of bacteria 427– 431
  of plants 5, 177, 179, 443
    in micrograph 13
Cellobiohydrolases 602
Cellular adhesion molecules (CAMs) 1883
Cellulases **602**, 603
  in bacteria 1148
  genetic engineering of 1148
  structural features 602
Cellulose 170, 171*s*, 172
  biosynthesis of 1130, **1146 – 1150**
  content in spinach 31
  enzymatic degradation 1148
  in plant cell walls 1149
  synthesis in bacteria, regulation 1148
Cellulose I 171*s*, 172
Cellulose II 172
Cellulose-binding domains 602
Cellulose microfibrils
  dark field light micrograph 1146
  formation by bacteria 1146
  in plant cell wall 1146
Cellulose synthases 1146
Cellusome 602
Cementum 442
Centipedes 24
Centrifugation 100
  in CsCl gradients 249
  density gradients 101
  zone 101
Centrins 1503
Centrioles 10, 15, 1502
Centromeres 1502, 1537
  DNA of 228
Centromeric proteins 1538
Centrosome 627, 1119
Cephalopods, mesozoa in kidneys of 23
Cephalosporin 1164, 1713
Cephalosporin C 1164, 1164*s*
Ceramidase 1171
Ceramide(s) 389*s*, 1168, 1202, 1203*s*
Ceramide aminethyl phosphonates 389*s*
Cerebellum 1765
Cerebral cortex 1765
  diagram 1767
Cerebral ganglion in Platyhelminthes 23, 25
Cerebrosides 388, 389*s*, 1168, 1169
Cerebrotendinous xanthomatosis 1251
Cerebrum 1765
Ceruloplasmin 58, 633, 839, 883, 887
CFTR 417
CG doublets 1541
cGMP 561, 1330
Chagas disease 787
Chain elongation process 700
Chain-terminating inhibitors for DNA sequencing 262
Chain termination codons 1475
Chalcones 1214, 1215*s*, 1217
Channeled substrates 552
Chaperone (chaperonin) 339, 1856
  calnexin 188
  DnaK 518
  GroEL 518
  GroES 518
  heat shock proteins 518
Chargaff, Erwin 200, 1474
Charge, effect on metal binding 311
Charge-relay system 612
Charon series 1496
CheA 1094
CheB 1095
Chelate complexes 310*s*
Chelate effect 310
Chelate stability 311
Chemiautotrophic organisms 1057
Chemical force microscopy 131
Chemical shift 138
  hydrogen bonding, effect of 139
  prediction 139
  shielding of nuclei 138
  table of NMR spectra 141
  units of 138
Chemiluminescence 1340
Chemiosmotic theory 1037, 1038
Chemoheterotrophic bacteria 8
Chemokines 1753, 1848
Chemolithotrophic bacteria (chemolithotrophs) 7, 1050
  electron transport in 1050
  slow growth of 1050
Chemoprevention 575
Chemoreceptors 562*s*, 1742
  bacterial 561– 563, **1093 – 1095**
Chemostat 470
Chemotactic factors 1753
  in inflammation 1122
Chemotaxis 1122
  of bacteria 1091, **1093** – **1095**
  carboxylmethylation in 548, 563
  CheA 562, 563
  CheB 563
  CheR 563
  CheW 562, 563
  CheY 562, 563, 1094
  CheZ 563, 1094
  human leukocytes **1122**
Chenodeoxycholic acid 1251
Chiasmata 1504
Chiral acetates 704, 705
  nonenzymatic preparation 704
Chiral methyl group 591
Chirality 42
Chitin 19, 170, 175, 1148, 1884
  biosynthesis of 1130, 1146, 1148, 1149
  in exoskeleton of arthropods 24
  microfibrils from tube-worm 1146
Chitinases 599 – 602, 601, 603, 1149
Chitosan 175, 1148
Chl *a*/*b* binding proteins 1308
Chlamydomonas 20, 22
Chloramines 1073
Chloramphenicol 637*s*, 1444*s*, 1690
Chloramphenicol acetyltransferase (CAT) 637
Chloramphenicol erythromycin 1703
Chlordiazepoxide 1788*s*, 1788, 1789
Chlorella 17, 22
Chloride ion, activation of α-amylases 606
Chloride ion channels 420, 421
  band 3 protein 420
  neurons, ionic environment 1773, 1774
Chlorin(s) 843
  biosynthesis of, scheme 1401
  ring system 1302
β-Chloroalanine 739
Chlorobenzoyl-CoA dehalogenase 682, 683
Chlorocruorins 362
Chlorocruoroheme 845*s*
bis-(2-Chloroethyl)sulfide 1584
Chlorogenic acid 1440, 1440*s*
Chloroperoxidase(s) 855
Chlorophyll(s) 13, 943
  absorption spectrum of 1304
  accessory pigments 1302 – 1305
  biosynthetic pathway 1403
  chelate effect 310
  light-harvesting 1310
  structures of 1303
  synthesis 1402
Chlorophyll a 10, 843*s*
Chlorophyllide(s) 1303*s*
Chlorophyllide *a* 1403*s*
Chlorophyllins 1303*s*
Chlorophyta 20
Chloroplast(s) 10, 14
  ATP formation in 1318
  Calvin cycle in 984
  composition of, table 1302
  electron micrograph of 1302
  ferredoxin 1317
  genomes 1301
  membranes 1301
  metabolism 1320
  in micrograph 13
  oxygen protection 1319
  radiation protection 1319
  ribulose bisphosphate carboxylase in 706
  structure **1301– 1305**
Chlorosomes 1308
Chlorpromazine 1809*s*, 1810
Chlorsulfuron 1392*s*
Cholecalciferol (vitamin D_3) 1257

- Cholera 7
- Cholera toxin **546 – 548**, 546*s*
  - in ADP-ribosylation 546
  - AMF images 131
- Choleragenoid 333*s*
- Cholestanol 1243
- Cholesterol 390*s*, 392, 946, 1243*s*, 1246*s*, 1256*s*
  - abundance of in humans 1247
  - elevated 16
  - esterases 1248
  - formation of 1245
  - gallstones 1251
  - lanosterol converted to 1246
  - metabolism of **1247– 1252**
  - metabolism of, abnormal 1251
  - metabolism to bile acids 1251
  - metabolism to steroid hormones 1252, 1253
    - as precursor of steroid hormones 696
    - from reduction of desmosterol 777
  - serum 1248
  - sterols, precursor to 1244
- Cholesteryl ester storage disease 1251
- Cholesteryl ester transfer protein 1248
- Cholic acid 1251, 1252*s*
- Choline 1199, 1398
  - phosphorylation 995
- Cholinergic synapses 1777, 1785
  - inhibitors of 1783
- Chondroadherin 1154
- Chondrocytes 1902
- Chondroitin 1130
- Chondroitin sulfate(s) 170, 175, 176*s*
- Chordata 25
- Chorion 1881
- Chorionic gonadotropin 1745
- Chorismate 485*s*, 690*s*, 1422*s*, 1424, 1424*s*, 1426*s*
  - biosynthesis of 689, 690
- Chorismate mutase 485, 1424
- Chorismate mutase-prephenate dehydrogenase 1425
- Choroidemia 1332
- Choroid plexus 1766
- Christian, W. 767
- Christmas factor 632
- Chromaffin cells 1767
- Chromatids 1501
  - electron micrograph of 1532
  - figure 1506
- Chromatin 11
  - active 1533
  - fiber models 1533
  - remodeling 1626
- Chromatography **102 – 105**. *See also* specific types
- Chromatophores 1318
- Chromatosome 1531
- Chromium **888**, **889**
  - complex with ATP 643
- Chromophore(s) 1280
  - visual 1326 – 1329
- Chromosome(s) 11, 17, 1534
  - artificial 1562
  - bacterial 3, 1530
  - crossing over, diagram 1894
  - homologous pair 18
  - homopolymeric tails 1491
  - human, numbering, staining patterns 1512
  - human, structure 1543
  - lampbrush 1534
  - mapping of 1477, 1486 – 1490
  - master 1017
  - metaphase 1533
  - micrograph of 371
  - numbers of 18
  - optical mapping 1490
  - physical mapping by electron microscopy 1489, 1490
  - physical maps 1488
  - polytene 1506
  - restriction mapping 1488
  - sizes of 18
  - X-chromosome, defects 1512
- Chromosome map
  - *E. coli* 1486
  - morgan, unit of distance 1504
- Chronic granulomatous disease 1072
- Chronic wasting disease (CWD) 1718
- Chrysanthemic acid 1237*s*
- Chylomicron(s) 1181– 1185
  - composition of, table 1183
- Chylomicronemia 1185
- Chymopapain 618
- Chymosin 621
- Chymotrypsin 66, 609
  - inhibition by TPCK 477
  - specificity 117
  - turnover number of 457
- Chymotrypsinogen 609, 613, 615
  - activation of 609, 615
  - hydropathy index 78
- Cibachron blue in affinity chromatography 106
- Cilia 10, 15. *See also* Flagella
  - cross-sectional views 1121
  - eukaryotic structure 15, **1121**
- Ciliophora 19
- Cineole 1233*s*
- Cinnamate 1429*s*, 1439*s*
- *trans*-Cinnamate 1438
- *trans*-Cinnamic acid 756*s*, 1214
- Cinnamic aldehyde 1439*s*
- Cinnamoyl-CoA thiol esters 682
- Circadian cycle 1339, 1793, **1805 – 1808**
- Circular dichroism (CD) 1286, 1287
  - of bound PLP and PMP 750
  - graph 1287
  - intensity of bands 1287
  - sign of bands 1287
- Circulatory system 1901, 1902
- Circulin 1800
- Cisplatin 218, 224*s*
- Cisternal maturation model 1159
- Cistron 1478
- Citramalate 1372
- Citrate 480*s*, 506*s*, 515, 516*s*, 686*s*
  - binding constants to metal ions 311
  - prochiral centers 480
- Citrate cleaving enzymes **703**, **704**
- Citrate synthase 527, 676, 700 – 703, 702*s*
  - active site 703
  - mechanism 700
- Citric acid 99
  - synthesis of 700
- Citric acid cycle 506, **508**, **515**, **950 – 960**, 981, 1436
  - catabolism of intermediates 957, 958
  - control of 953 – 957
  - discovery of 516
  - in mitochondria 1015
  - priming of 711
  - related oxidative pathways 958 – 960
  - scheme 516, 951
- Citronellal 1233*s*
- Citronellol 1233*s*
- Citrulline 80, 1376
  - biosynthesis of, scheme 1375
- Citrus bioflavonoid 1215
- Citrus fruit, cure for scurvy 1066
- Cityl-CoA 703
- Claisen condensation 699, 946
- Claisen rearrangement 485
- Clathrin 426, 1781
  - light chains 313
- Cleavage
  - aldol 698
  - alpha 730
    - in dicarboxylic acid cycle 960
  - of amino acid side chains 745
  - beta **698 – 711**
    - in dicarboxylic acid cycle 960
  - of ovum 1897
- Clonal selection 1858
- Cloned genes 260 – 264
  - expression of in bacteria 1497
- Clones of cells 26, 1478
- Cloning 1885
  - of DNA 201, 1490 – 1500
    - scheme 1495
  - nuclear transplantation 1886
  - positional 1500
- Cloning plasmids, genetic map of 1493
- Cloning vehicles (vectors) 263, 1490 – 1500
- *Clostridium acetobutylicum* 705
- Clostripain, specificity 117
- Clozapine 1809*s*, 1810
- Clp (caseinolytic protease) 628
- Cluster of differentiation, CD 187, 407, 1850
- Clusterin 1721
- CMP (Cytidine 5'-phosphate) 203, 995, 996
- Cnidaria (Coelenterata) 23
- Coated membrane 426*s*
- Coated pits 426
- Coated vesicles 426
- Coatomer 426
- Cobalamin 867. *See also* Cyanocobalamin, Vitamin B_{12}
  - in ribonucleotide reductase 871
- Cobalamin-binding proteins 58
- Cobalophilin 869
- Cobalt **866 – 877**
  - complex with ATP 643
- Cobaltochelatase 1402
- Cobra venom 1775
- Cocaine 1793, 1793*s*, 1796
- Cocarboxylase 721
- Coccidia 19
- Coccus, definition of 6
- Co-chaperones 1721
- Cockayne syndrome 1585
- Cocoonase 610
- CO dehydrogenase/acetyl-CoA synthase 881
- Coding strand (non-transcribing) of DNA 204, 237
- Codon(s) 4, 236, 1672
  - chain termination 1480
  - initiation 236
  - pairing with anticodon 1693, 1694
  - stop 236
  - tables of 236, 237
- Coelenterazine 1342
- Coenzyme(s) 511, **719 – 757**, **765 – 829**

  ease of dissociation from proteins 719
  oxidative 719
Coenzyme A (CoA) 507, **720 – 723**, 722*s*
  biosynthesis of, scheme 1392
  functions of 722
  transferase 970
Coenzyme A persulfide 790
Coenzyme A transferases **662**
Coenzyme F_{420} 788
Coenzyme M 813*s*, 814, 815, 1410
Coenzyme M reductase
  mechanism of 881
    scheme 880
Coenzyme Q. *See* Ubiquinones
Cofactor F_{430} 879 – 881, 880*s*
Cofilin 1119
Coherent light 1275
Coiled coil
  helical wheel 70
  heptad repeat 70
  superhelix 70
Cointegrate molecule, unknotted
  resolution of, scheme 1577
Colchicine 371, 371*s*, 1443
Colicin(s) 414, **418**, 1482, 1565, 1842
Colicin E1 (ColE1) 248, 1774
  plasmid 418, 1493
  control by antisense RNA 1615
Colicin E3 1686
Colipase 635, 1181
Colitose 180*s*
Collagen 15, **72**, 181, 409, **432 – 441**
  alternative splicing 436
  amino acid composition 72
  biosynthesis of 433
  characteristics, table 434
  crosslinking 433, 434
  gene location, table 434
  genetic defects **438**
  hydrogen bonds in 72
  location, table 434
  sequence 72
  structure 72, 432, 433
  vertebrate types, table 434
Collagen, type IV
  electron micrograph of 437
  structural model of 437
Collagen fibril, electron micrograph 73
Collagenases 627
Collagenlike peptide 73*s*
Collenchyma 30
Colominic acid
  biosynthesis of 1130
Colony-stimulating factors (CSF) 1755
Color reagents 120
Color vision 1328
Combinatorial libraries 128, 129
Combustion, Gibbs energies of 297, 298
Commissural neurons 1903
Communication between cells 26 – 29, 553 – 581, **1740 – 1814**
Competitive inhibitors 471, 472
  slow, tight-binding 472
Complement **1844 – 1846**
  alternative pathway 1844
  chemotactic factors 1844
  classical pathway 1844
  cofactor I 1846
  lectin pathway 1844
  receptor 2 1846
  recognition component C1 1844
Complement system 1832
  activation pathways 1845
  of blood 72, 414
  proteins 1831
  thioester-containing proteins 630
Complementarity
  of amino and carbonyl 679
  of enzyme surfaces 478
  of substrate 478
Complex V. *See* ATP synthase
COMT. *See* Catechol *O*-methyltransferase
Concanavalin A 186, 342
Concatemers 1558, 1559
Concentration gradient(s) 289, 410
Concerted mechanism 490
Condensation
  aldol 699, 700
  alpha 730
  beta 527, **698 – 711**
Cone cells, of eye 1324
Configuration
  of amino acids **41– 43**
  entgegen (*E*) 43
  rectus (*R*) 42
  RS notation **42**
  sinister (*S*) 42
  of sugars 163
  zusammen (*Z*) 43
Confocal micrograph
  of a cerebellum 130
  of metaphase chromosomes 361
Confocal scanning microscopy 129, 1292
Conformations of molecules **43 – 45**
  anti 43
  anti-periplanar 43
  boat (B) 166
  C1 of sugars 166
  chair 166
  eclipsed 44
  envelope 166
  Fisher projection 42
  gauche 44
  of 5- and 6-membered rings 166
  Newman projection 44
  of nucleotides 211
  of polysaccharide chains 170
  quasi-equivalence 348
  random coil 69
  sickle 44
  skew (S) 166
  twist (T) 166
Conformational changes
  cooperative **349 – 353**
  during dehydrogenase action 773
  in enzymes 481
  in proteins 81
  role of buried hydrogen bonds 81
  slow 477
Conformational coupling
  ATP synthase 1044
Conformational maps 60
Conformational substates in proteins 496
Congenital adrenal hyperplasia 1253
Conidia 20
Coniferyl alcohol 1439*s*, 1440
Conjugate bases 486
Connectin. *See* Titin
Connective tissue 72, 431
Connexin images 1782
Connexons 29, 1781, 1782
Conotoxin(s) 1770*s*, 1775, 1783
Consensus sequences in DNA 265
Constitutive enzyme 538
Contact inhibition 573
  of cells 29
Contig 1508
Contraceptive pills 1262
Control elements of metabolic reactions 536
Controllability coefficient 537
CooA 1613
Coomassie brilliant blue 102, 122*s*
Cooperative binding 352
  of protons 331
  of substrate 476
Cooperative changes in conformation **349 – 353**
Cooperative processes 330 – 332, 334
Copper 65, 514, **882 – 887**
  concentration in mitochondria 1019
  in cytochrome *c* oxidase 1028 – 1032
  hydroxylases, containing 1064,1065
  as nutritional essential 882, 883
  uptake regulation 883
Copper amine oxidases **816**, **817**
  mechanisms 817
Copper oxidases 885 – 887
Copper proteins, electron-transferring 883, 884
Copper – sulfide cluster 884*s*
Coproporphyrin III 843, 845*s*
Coproporphyrinogen 1400
Coproporphyrinogen oxidase 1400
β-Coprostanol 1244*s*
Cori cycle 966, 989
Cornified cell envelope 439
Corpus callosum 1765
Corpus luteum 1253
Corrin(s) 1402
  in transmethylation 592
Corrin ring 867, 868
Corrinoid-dependent synthesis
  of acetyl-CoA 876, 877
Cortical granules 1896
Corticosterone 1256*s*
Corticotropin (adrenocorticotropin, ACTH) 1252, 1742, 1743
Corticotropin-releasing hormone 1746
Cortisol 1253, 1256*s*
Cortisone 1254, 1256*s*
Cosmarium 22
Cosmid vectors 1496
COSY-NOESY diagram 143
Cotransport (symport) process 411, 416, 417
Cotton effect 1286
Coulomb 283
  definition of 47
Coulomb's law 47
Coumarate 1429*s*
Coumarin 1439*s*
Coumaroyl (4-hydroxycinnamoyl) group 1336
Coumaroyl-CoA 1440
p-Coumaroylagmatine 1440*s*
Countercurrent distribution 102
Coupled enzyme assay 768
Coupling constants. *See* NMR spectroscopy
Coupling factor(s)
  (F_1) 1014, 1041
  oxidative phosphorylation 1041
Coupling membrane, in mitochondria 1038
Covalent catalysis 494, 591
Covalent modification
  in control of metabolism **541– 545**
  by dephosphorylation 541– 545
  by phosphorylation 541– 545

  reactions, table 543
Covalent radii 40
Covalent structure of proteins 57
Cow, genome 12, 1510
Cozymase 767
Crambin, amide exchange rate 149
Crassulaceae
  acid metabolism, scheme 1324
  metabolism in 1323
C-reactive protein 1850
Creatine 1379
  synthesis 1379
Creatine kinase 656
  transition state structure 640
Creatine phosphate 303, 1117, 1379
Creatinine 1379, 1379*s*
Crepenynic acid 381*s*
Cre recombinase 1501, 1570
Cresol 1436*s*
Cretinism 31, 1432
Creutzfeldt-Jakob disease (CJD) 248, 1718
  sporadic 1718
Crick, Francis H. C. 84, 200, 1474, 1475
C-ring 1093
Cristae 1014
  of mitochondria 14
Crohn disease 1855
Crossbridges, skeletal 1104
Crossing-over 18, 1504, 1505
  unequal 1568
Crosslinking 79
Cross-β structure 1719
Crotonase (enoyl hydratase) 681
Crotonyl-CoA 1386*s*
Crotoxin 1775, 1776
Crowfoot Hodgkin, Dorothy M. 84
Crown galls 1497
Cruciform structure in nucleic acids 229
Crustacea 24
Cruzain 619
Cryo electron microscopy 130, 1669, 1679
Cryoenzymology 469
  elastase 616
Cryoprotectants 191
Cryptochromes 1320
Crystal systems 133
Crystallins 169, 1333
Crystallography 131 – 137
  electron 131
  X-ray 132 – 137
Crystals, liquid 392 – 394
Cubic symmetry
  of oligomers 342
  of pyruvate dehydrogenase 796
Cucurbitacin B 1248*s*
Cupric ion, complexes with ammonia 307
Curdlan 170, 174
Cuticle
  of invertebrates 440
  in micrograph 13
  of plants 15, 30
Cutin 1196
Cutinase 134*s*, 635
Cyanase 717
Cyanate, carbamoylation by 360
Cyanide 1408
  as respiratory inhibitor 1021
Cyanide ion
  adduct with NAD$^+$ 779
  complex with hemoglobin 358
  reactivity of 590
Cyanin 1214
β-Cyanoalanine 438, 1407
Cyanobacteria 7, 10
  genome 12
  photosynthesis 1305, 1313 – 1315
Cyanobacterial heterocysts 1365
Cyanoborohydride 124
Cyanocobalamin 867*s*, 869. *See also* Vitamin B$_{12}$
  reduction 869, 870
Cyanogen bromide 117
  agarose, activation of 105
Cyclic adenosine 3',5'-monophosphate
  as attractant for *Dictyostelium* 20
Cyclic ADP ribose 778
Cyclic AMP. *See* cAMP
Cyclic AMP–protein kinase A cascade 1505
Cyclic diguanylate (c-di-GMP) 1148
Cyclic GMP. *See* cGMP
Cyclic photophosphorylation 1314
Cyclic symmetry 333
Cyclin(s), mitotic 580, 1502, 1503
  cell cycle 524
  D1 572, 1908
  G1 or start 580
Cyclin-dependent protein kinases (CDK's) 1502
Cyclitol 1131
Cyclization
  PLP-dependent enzymes 741
Cycloartenol 1245*s*, 1248*s*
  in sterol biosynthesis 1244
Cyclocreatine 1379
Cyclocreatine phosphate 1380*s*
Cyclodextrins 171
Cyclodextrin glucanotransferase 603, 607
Cyclodienes 1788
Cycloglutamate inhibitors 739*s*
Cycloguanil 805
1,2-Cyclohexanedione, reaction with guanidinium groups, 126
Cyclohexanone oxygenase 1060
Cycloheximide 1216*s*
Cyclooxygenases (COX) 1207, 1208, 1211
Cyclophilin 488, 1867
  human 488*s*
D-Cycloserine 739*s*
Cyclosporin 488, 488*s*, 1713, 1867
β Cylinders 65, 66, 686
Cystathionine, 746*s*, 1383, 1388
  formation 746
Cystathionine β lyase 742, 1388
Cystathionine γ-lyase 1388
Cystathionine β-synthase 744, 1388
Cystathionine γ-synthase 743, 746, 1388
Cystatins 622, 629
Cysteamine disulfide (cystamine) 549
Cysteic acid 1407
Cysteine (Cys, C) 52, 53*s*, 1398
  catabolism 1407
  metabolism, scheme 1406
  microscopic pK_a values 306
  synthesis of 1407
  zinc complex 680
Cysteine desulfurase 1410
Cysteine proteases **618**, **619**
Cysteine residues 53
  crosslinking of 80, 521, 522, 786, 787
Cysteine string proteins 1780
Cysteine sulfinate 1407, 1408
Cysteine synthase 1407
Cystic fibrosis 16, 421, 1513
  alginate-forming bacteria 1153
  transmembrane conductance regulator 421, 1513
Cystine residues, beta elimination 85
Cystinosis 1407
Cystinuria 16, 417, 1407
Cytidine (Cyd) 203
  absorption spectra of 205
  reaction with hydrazine 252
Cytidine diphosphate choline 720, 995
Cytidine diphosphate-glucose 1138
Cytidine 5'-phosphate. *See* CMP
Cytidine 5'-phosphate-*N*-glycolylneuraminic acid 1137
5'-Cytidylic acid. *See* CMP
Cytoadhesins 409
Cytochalasin B 416*s*
Cytochrome(s) **845 – 848**, 1022
  content in mitochondria 1019
Cytochrome *a* 1019, 1027. *See also* Cytochrome *c* oxidase
Cytochrome a_3 1019
Cytochrome aa_3 terminal oxidase 1025
Cytochrome *b* 847, 1019, 1027
Cytochrome b_5, 1016
Cytochrome b_{562} 847
Cytochrome b_6f 1300
Cytochrome bc_1 complex 1027, 1029*s*
Cytochrome *c* 846, 847, 847*s*, 1019
  absorption spectra 846
  structural studies 846
Cytochrome c_1, 1019, 1027
Cytochrome c_3 847*s*
  in sulfate reduction 1056
Cytochrome c_6 1300
Cytochrome *c* oxidase (Complex IV) 513, 848, 887, 1021, 1022, **1028 – 1032**, 1031*s*
  catalytic cycle 1032
  copper centers 1030*s*
Cytochrome *c* peroxidase 849*s*, 852
  active site of 853
Cytochrome cd_1 nitrite reductase 1055
Cytochrome oxidase. *See* Cytochrome *c* oxidase
Cytochrome P450 846, 1056, 1065, 1068, 1253, 1258
  in cholesterol to bile acids 1065
  heme active site 1068
  in lignin synthesis 1065
  in liver 1065
  mechanisms 1068
  stereoscopic view 1069
  in phenylpropanoid pathway 1065
  in steroid metabolism 1065
Cytochrome P450 hydroxylases 1253
Cytochrome P450$_{\text{nor}}$ 1056
Cytokines 188, **1753 – 1756**, 1754*s*, 1832, **1846 – 1850**, 1863
  categories of 1847
  in hemotopoeiesis 1847
  in immune-mediated inflammation 1847
  in lymphocyte regulation 1847
  in natural immunity 1847
Cytokinin(s) 235, 1761
Cytolysins 1846
Cytomegalovirus
  micrograph 246
  protease of 614
Cytophotometry 1474
Cytoplasm, definition of 3
Cytoplasmic inheritance 1507
Cytosine (Cyt, C) 199*s*, 203

  acid–base chemistry 204, 205
  infrared spectrum 1293
Cytosine nucleotides 1451
Cytoskeleton 15, **368 – 372**, 405
  intermediate filaments 369
  microfilaments 369, 370
  microtubules 370
Cytosol, definition of 14
Cytoxic (killer) T cell 1852

D

Daltons 4
Dansyl chloride 118, 118*s*
Dark current 1331
Darwin, Charles 9
Daughter chromosomes 1502
Daunomycin 222, 223*s*, 1589
  intercalation with DNA, figure 223
Davydov splitting 1287
DDT dehydrochlorinase 551
Dead-end complexes 466
DEAE-cellulose 104
DEAE-Sephadex 106
Deamidation
  of asparagine 594
  of asparaginyl residues 84
Deaminase 1135
Death receptors 1889
5-Deazaflavin 1462
  in hydrogenase 879
  nonenzymatic oxidation of NADH 789
Deazariboflavin 5'-phosphate 789
Debranching enzyme 606
  deficiency of 1144
Debye-Hückel equation 288
Decamethonium 1783*s*, 1784
Decapentaplegic (Dpp) 1900
Decaprenaoxanthin 1239*s*
Decarbonylation 1196
Decarboxylases **753 – 755**
  biotin-dependent 729
  dialkylglycine 745
  pyridoxal phosphate-dependent enzymes 744, 745
  pyruvoyl enzymes
  Schiff bases in 732
  thiamin-dependent 730 – 733
Decarboxylation 527, **705 – 710**
  in biosynthesis 977– 980
  catalysis by amines 705
  catalysis by metal ions 705
  of α-hydroxy acids 943
  initiating catabolism 1371
  of β-oxoacids 698, 705, 742
Decarboxylation-transamination reaction 748
Decarboxylative elimination 526
Decoding center
  of ribosomes 1672
Decorin 1154
Defective lambda phage 1483
Defensins 1774, 1832, **1842 – 1844**
Deformylase 1721
Degree of hydration, effect on metal binding 311
Dehydration, of sugars 1137
Dehydroalanine 754 – 756, 757*s*
Dehydroascorbic acid 416, 787, 1064, 1066*s*, 1133*s*, 1134
  transport 416
7-Dehydrocholesterol 1246, 1246*s*
Dehydroepiandrosterone (DHEA) 1255, 1256*s*
Dehydrogenase(s) 514, **765 – 802**, 1025
  conformational changes 771, 772
  3-dimensional structures 768
  mechanisms 768 – 775, **770**, **771**
  nucleotide-dependent 777
  stereospecificity 768 – 775
Dehydrogenase and oxidase complexes 1023, 1026 – 1028
Dehydrogenation 1022
  of alcohols 765 – 767
  of amines 775
  of amino acids 775
  reactions of flavoproteins, table 784
  types of, table 766
Dehydrogenation- decarboxylation system
  of pentose phosphate pathway 963
3-Dehydroquinate 1423*s*, 1424, 1424*s*
3-Dehydroshikimate 1423*s*, 1424*s*, 1438
Deisenhofer, Johan 84
Delayed hypersensitivity 1852
Delphinidin 1214
Dematin 405
Demethylation 1541
Denaturation of proteins 82
  by guanidinium chloride 82
  by sodium dodecyl sulfate 82
  by urea 82
Denaturation loops in DNA 1490
Dendrites 1762, 1802
  diagram 1763
Dendritic cell(s) 1832
  drawing 1835
Denitrification **1054 – 1056**
Denitrifying bacteria 1054, 1359
Dense-core vesicles 1762
Dental plaque 442
Dentin 442
5'-Deoxyadenosyl cobalamin 867, 867*s*, 872
3-Deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) 1422*s*, 1423
Deoxyazidothymidine (AZT) 1655*s*
Deoxyerythronolide B 1217*s*
Deoxygenase, NO-activated 1757
Deoxyhemerythrin, oxygenation of 862
Deoxyhemoglobin 850, 851
Deoxy-D-*manno*-octulosonic acid (KDO) 180*s*, 1136
2-Deoxy-2-methylamino-L-glucose 1139*s*
Deoxyribonuclease **652, 653**
Deoxyribonucleic acid. *See* DNA
Deoxyribonucleotide(s) 995
  in DNA synthesis 1452
2-Deoxyribose 200*s*
Deoxy sugars, synthesis of **1137– 1140**
Deoxythymidine diphosphate 1138
DeoxyUTPase 1581
1-Deoxyxylulose 5-phosphate 735, 736, 1463, 1464*s*
Dephosphorylation, cascade 542
Depression 1808 – 1810
  biogenic amine hypothesis 1808
  monoamine oxidase inhibitors 1809
  MRI images of brains 1810
  serotonin reuptake inhibitor 1810
Depsides 23
Depsipeptide 414
Derepression of enzyme synthesis 536
Derivative spectra 1285
Dermatan, biosynthesis of 1130
Dermatan sulfate 170, 175, 176*s*
Dermatosparaxis 438
Dermis 439
Desaturases 863, 1192
Desaturation 1237
Desmid 22
Desmin 1099
Desmosine 436, 436*s*
Desmosome 10
Desmosterol 777*s*, 1246*s*
  reduction to cholesterol 777
Desulforedoxin 859
Desulfovibrio desulfuricans
  reduction of sulfate 1056
Desulfuration of insectides 636
Dethiobiotin 1393
Development **1879 – 1909**
  basic concepts 1879 – 1883
Developmental clock 1887
Developmental programs, alternative 1880
Dexamethasone 1254, 1254*s*
Dextran(s) 170, 174, 1152
Dextran sucrase 1152
Dextrin(s) 606
DHEA 1255
DHPG 1655*s*
Dhurrum 1442
Diabetes insipidus 1748
Diabetes mellitus 567, **1003 – 1005**
  ketosis in 946
  type I (juvenile) 1864
Diabetic nephropathy 438
Diacylglycerol(s) 563, 564, 1204
  scheme 565
  in signaling 563 – 566, 1743
Diacylglycerol kinases 1204
Diadenosine-5'-tetraphosphate 1636*s*
Dialkylglycine decarboxylase 743, 745
Dialkylnitrosamines 254*s*
Dialysis 100
L-*erythro*-3,5-Diaminohexanoic acid 1388*s*
Diaminopimelate(s) 1385*s*
Diaminopimelate pathway
  lysine synthesis 1383
  scheme 1385
1,4-Diaminopropane 745
Diaminovalerate 1401*s*
Dianionic urate 1460s
Diastereoisomers (diastereomers) 43, 162
Diastereotopic atoms or groups 479
Diatomaceous earth 21
Diatoms 21
  motion of 21
Diazepam 1788*s*, 1788, 1789
6-Diazo-5-oxo-L-norleucine (DON) 1371
Diazotrophs 1359
Dibenamine 1791*s*
Dibenane 1791
DIBOA 1449
1,2-Dibromoethane 1587, 1588
Dicarboxylate transporter 1907
Dicarboxylic acid(s)
  transport in mitochondria 1047
Dicarboxylic acid cycle **958 – 960**
2,6-Dichlorophenol-indophenol (DCIP) 1025*s*
Dichlorophenyldimethylurea 1299, 1299*s*
Dichrograph 1286
Dichromats 1328
Dicoumarol 821, 821*s*
Dictyosome micrograph 13
Dictyostelium discoideum 20, 1892

  cAMP as messenger 557
  genome 12
Dicyanocobalamin 870
Dideoxyadenosine 1655*s*
3,6-Dideoxyhexoses, formation of 747
Dideoxy sugars 1138
Dielectric constant 47, 1275
Diethanolamine 99
Diethyl ether 1789
Diethylmalonic acid 99
Diethylpyrocarbonate reaction with histidine 126
Difference electron density map 136
Difference spectra 1285
Differential scanning calorimetric curves 395
Differentiation
  of cells 1882
  cluster of (CD) 187, 407
  essentials for, table 1881
  among prokaryotes 9
Differentiation antigens 1850
Diffractometer 134
Diffusion 109, 410
  exchange 410
  facilitated 410
  mobile carrier 414
  saturation effect 410
  through membranes 410, 414
Diffusion constant 461
Diffusion-controlled limit on reaction rates 462
Digalactosyl diacylglycerol 1197*s*, 1301
Digestion 505, 507
Digitonin 1021, 1265, 1265*s*
Digoxigenin 257
Dihedral (D_2) symmetry 338
Dihomolinolenic acid 1190
D-*gluco*-Dihydroacarbose 607*s*
Dihydrobiopterin 1061
  quinonoid 1061*s*
Dihydrodipicolinate 1385*s*
Dihydrofolate reductase 804 – 808, 810, 811
  active site structure 807
  inhibitors 805
  mechanism 805 – 808
  structure 805 – 808
Dihydrolipoic acid amide 784*s*
Dihydrolipoyl dehydrogenase 785, 796
Dihydrolipoyl transacetylase 796*s*
Dihydroneopterin 1460
erythro-7,8-Dihydroneopterin triphosphate 1460, 1461*s*
Dihydroorotate 1450, 1451*s*
7, 8-Dihydro-8-oxoguanine 1582, 1582*s*
Dihydropteric acid 1461*s*
Dihydropteridine reductase 1062
Dihydropteroic acid 473
Dihydropyridine receptor 314, 1115
Dihydroriboflavin 794
5α-Dihydrotestosterone 1254, 1255*s*
Dihydro U loop of tRNA 231, 1688
5,6-Dihydrouracil 1453*s*
Dihydrouridine 234
Dihydroxyacetone 164*s*
Dihydroxyacetone phosphate 509, 694*s*, 961, 983, 985
  reduction of 968, 1196
Dihydroxyacid 697
Dihydroxyacid dehydratase 1393
  Fe-S clusters in 861
ortho-Dihydroxybenzene. *See* Catechol
Dihydroxybenzoate 1422, 1425 – 1427, 1426*s*
3,4-Dihydroxycinnamate (caffeate) 1439*s*
Dihydroxyindole 1433*s*, 1434
Dihydroxymandelic acid 1790*s*
Dihydroxyphenylacetaldehyde 1444*s*
3,4-Dihydroxyphenylalanine *See* Dopa
2,3-Dihydroxyproline 80
Diimide 1364
Diiodotyrosine residues 1430
Diiron desaturases 863
Diiron oxygenases 863
Diiron proteins **862 – 864**
Diiron-tyrosinate proteins 862, 863
Diisopropylfluorophosphate (DFP, diisopropylphosphofluoridate) 610*s*
DIMBOA 1448, 1449
Dimethoxyphenylethylamine 1790*s*, 1811
Dimethylallyl diphosphate 712*s*, 1230
Dimethylallyl pyrophosphate, condensation of 527
Dimethylarsine 876
Dimethylbenzimidazole 1460, 1462, 1463*s*
  biosynthesis of 1463
Dimethylglutaric acid 99
Dimethylglycine 1389*s*, 1398, 1399
Dimethylimidazole 868*s*
6,7-Dimethyl-8-ribityllumazine 1462
N-Dimethylserotonin 1811
Dimethylsuberimidate 125*s*
  crosslinking of proteins 125
Dimethylsulfoxide reductase family 836*s*, 890
Dinitrogen, formation by nitrification 1055
Dinitrogenase (nitrogenase) 1360 – 1366
Dinitrogenase reductase 1361
2,4-Dinitrophenol 1033
Dinoflagellate 19
Dioldehydratase 871
  stereochemistry 872, 874
Dioxetane 820
Dioxygen 850*s*
Dioxygenases 1057– 1059, 1437
Dipeptidyl aminopeptidase 117
Dipeptidyl peptidase 619
Diphenylhydantoin 1809*s*
Diphosphatidylglycerol. *See* Cardiolipin
Diphthamide 1450, 1685, 1708, 1708*s*
Diphtheria, toxin 1685
  ADP-ribosylation 548
Diphytanyl group 385
Diphytanyl tetraether lipids 387
Dipicolinic acid 1383 – 1386, 1385*s*
Diploid cell 17
Diploid phase 18
Dipolar ion 41
Disaccharides 167– 169
  hydrolysis of, Gibbs energies 169
Discoidin(s) 186, 1892
Discriminator nucleotide 1694
Diseases *See also* names of specific diseases
  atherosclerosis 1249, 1250
  autoimmune **1864 – 1867**
  bacterial 7
  genetic 216, 945, 1513 – 1517, 1585 – 1587, 1907
  of glycogen metabolism 1145
  hereditary, of muscle 1112, 1113
  lysosomal storage 1170 – 1172
  mitochondrial 1024
  mucopolysaccharidoses 1171
  progeroid 1907
  sphingolipidoses 1171
Displacement reactions, nucleophilic 123, 482, 530, **589 – 670**
  on carbonyl groups 608 – 637, 698
  on 5'-methylene group of ATP 977
  multiple 660
  on phosphorus atom **637– 657** *See also* Kinases, Phosphatases, Transferases
    in-line 638
    permutational rearrangement 638
  proteases 608 – 637
  rates of **589, 590**
  on sulfur atoms 659, 660
Dissimilatory nitrate reductases 1367
Dissimilatory nitrite reductases 1367
Dissipative structures 289
Dissociation constant(s) 304
  apparent 293
Disulfide bridges (crosslinkages) 65, 80
  cleavage of 115, 116
  locating 119
  in proteins 80
  reduction of 115, 785
Diterpenes 1234 – 1236
Dithioerythritol 115
Dithionite ion 779
Dithiothreitol 98, 115, 822
Dityrosine 1434, 1434*s*
  linkages 81, 1074
DMSO reductase 890
DNA 198–270, 1528 – 1589
  Alu sequences, structure 1538
  base composition of 239
  bending of 218
  binding proteins 1551
  cDNA 257, 1499
  chains, elongation of 1556
  chloroplast 1561
  cistron 1478
  cloning of 201, 1490 – 1500. *See also* Cloning
  clustered repeats 1539
  coding capacity 16
  coding strand of 204, 237
  colinearity with protein sequences 1479
  conformational flexibility 216
  consensus sequences 265
  content in tissues 31
  damage to **1578 – 1584**
  denaturation of 255
  denaturation loops 1490
  density gradient centrifugation 101
  discovery 1473 – 1476
  dnaB hexamer, model 1551
  double helical structure 201
  electron micrograph of 213
  electron transfer in 218
  electrophoresis 221, 249
  excision of 1564, 1570 – 1573
  extraction of 249
  fingerprinting 259
  footprinting 266
  germline rearrangements 1859
  glycosylases 1581
  G-segment 1552
  gyrase 220, 1552
  heteroduplex molecule 1489
  Holliday junction **228 – 230, 1565 – 1568**, 1571
  hydration of 216
  hydrolysis of 249, 251
  hypersensitive sites 266

  - interaction with spermidine and spermine 218
  - intercalation **221**, **223**
  - intrachain NOEs 268
  - joining fragments 1490
  - kinetoplast 1561
  - knot, electron micrograph 219
  - ligases 1544, 1549, 1551
  - light absorption 209
  - linking number 220
  - long interspersed repeat sequences (LINES) 1539
  - loss of 1881
  - major groove 201, 213
  - melting curve 256
  - methylation of 1541, 1542
  - minor groove 201, 213
  - minus strand 1482
  - mispaired bases in 218
  - mitochondrial (mtDNA) 14, 1016, 1561. *See also* DNA, mitochondrial
  - nanochemical scale construction 230
  - nearest neighbor analysis 265
  - negative strand 244
  - NMR spectroscopy of **266 – 270**
  - non-coding (transcribing) strand 237
  - open reading frame 237
  - operator 237
  - of organelles 1540
  - organization of 1537– 1542
  - packing of, figure 1530
  - palindromes in 238
  - passenger (in cloning) 1491
  - plus strand 1482
  - polymerization chemistry 1543. *See also* DNA polymerases
  - primer strand 1543
  - probes 1499
  - promoter sequences 1562
  - reassociation curves, figure 256
  - reassociation kinetics 255
  - repetitive 11, **1537– 1539**
  - replication. *See* DNA replication
  - replicative form 244
  - rolling-circle synthesis 1558
  - satellite 1537
  - sequence, decoding 1672
  - sequence determination 260 – 266. *See also* DNA sequencing
  - sequences, understanding 265
  - short interspersed sequences (SINES) 1538
  - SINES, CAT family 1538
  - SINES, homeosequence 1538
  - single-stranded binding (SSB) proteins 239
  - solenoidal 220
  - strand exchange mechanism 1567
  - strand-exchange proteins 1566
  - structures 214, 223, 225, 228, 229, 240 – 243
  - synaptic complex, figure 1572
  - template strand 1543
  - topology of **1529 – 1537**
  - translesion repair 1583
  - transposition, chemistry of 1574
  - T-segment 1552
  - typing 259
  - in viruses 1529
  - Watson–Crick structure 200
- DNA-binding motifs
  - in proteins 241
- DNA-binding proteins 1551, 1632
  - helix-turn-helix motif 239
- DNA, circular 218, 219
  - linking number 220
  - writhe 220
- DNA, forms
  - A form 213, 214*s*
    - within conserved regions 9
  - B form 213, 214*s*
    - hydration of 213
    - right-handed helical structure 213
    - side-by-side structure 213
  - circular, in trypanosomes 219
  - colicin, chi form 1565
  - double helical structure 201
  - H-DNA, proposed structure 226
  - helix parameters 216
    - base inclination 216
    - base roll 216
    - helical twist 216
    - propeller twist 216
  - overstretched 217
  - plectonemically interwound 220
  - polymorphism 259
  - R-DNA 226
  - relaxed 221
  - supercoiled **219 – 222**
  - superhelix density 221
  - tetraplex (quadruplex) structures **227**
  - topological isomers 221
  - topology of **1529 – 1537**
  - toroidal (solenoidal) 220
  - triplex 227
  - underwound 220, 221
  - Z form 214*s*, 215, 239
    - enhancers in 216
- DNA glycosylase/AP lyases 1581
- DNA gyrase 220, 1552
- DNA, mitochondrial 14, 1016, 1561
  - defects of 1024
  - genomic map, figure 1017
  - physical map 262
- DNA photolyase 794, 802, 1339
  - action mechanism, proposed 1341
  - *E. coli* 1340
- DNA polymerase(s) 657, **1544 – 1549**, 1556
  - action on nicked strand, fig. 1549
  - A family 1547
  - B family 1547
  - editing 1544
  - exonuclease activities 1544
  - families three and four 1548
  - Klenow fragment 1547
  - of phage T4 1548
  - polymerase I 1491, 1543
    - ribbon drawing 1546
  - polymerase III 1548, 1549, 1556, 1557
  - polymerase α, δ, ε 1545, 1548, 1581
  - polymerase β 1545, 1548
  - polymerase γ 1545, 1548
  - polymerase ζ, λ, μ, σ, *Umu*, η, ι, κ 1545, 1548
  - processive mechanisms 1547
  - proofreading 1544, 1547
  - in repair 1583, 1584
  - ribbon drawings 1546
  - table 1545
  - *Thermus aquaticus* (*Taq*) 260, 1545
  - zinc in 680
- DNA-protein interactions 266
- DNA recombination
  - cointegrate 1574
  - insertion sequences 1574
  - integration of 1564, **1570 – 1573**
  - invertible sequences 1572
  - RecBCD recombination 1566
  - recombination of 1564
  - recombination mechanism, figure 1565
- DNA repair **1578 – 1584**
  - alternative excision repair 1581
  - base excision repair 1581
  - deficiencies in 1585, 1586
  - excision repair 1581
  - initiation (replicator) regions 1562
  - methyl-directed mismatch repair 1580
  - repair, deficiencies in 1585
    - of double-strand breaks 1583
- DNA replication 1542 – 1552, 1561
  - bacterial **1553 – 1557**
  - bidirectional, figure 1555
  - directions of 1554
  - discontinuous replication 1544
  - eukaryotic **1561– 1564**
  - fork 1543
  - fragments 1544
  - initiation of 1555
    - bidirectional, scheme 1556
  - lagging and leading strands 1544
  - nuclear, replication of 1562
  - Okazaki fragments 1544
  - origin of 1554
  - origin recognition complex (ORC) 1562
  - proteins of, table 1551
  - replication licensing factor (LRF) 1563
  - replication of *Drosophila* genome 1563
  - semiconservative replication 1542
  - synthesis, priming of 1555
  - termination of 1557
  - viral 1557– 1559
- DNA-RNA hybrids 257
- DNA sequencing **260 – 266**
  - chain-terminating inhibitors 262
  - method of Maxam and Gilbert 264
  - Sanger dideoxy method. 262
  - sequencing gel, figure 263
- DNA supercoiling, estimating amounts 222
- DNase. *See* Deoxyribonuclease
- Docking protein for SRP 520, 1722
- Docosahexaenoic acids 1190
- Dolichols 1155, 1231
- Domains in protein structure 74
- Domoate 1786*s*
- Donnan equilibrium 420, 1039
- Dopa 80, 887, 887*s*, **1431– 1434**, 1433*s*, 1790
- Dopachrome 1433, 1433*s*
- Dopachrome tautomerase 1435
- Dopamine 744, 1064*s*, 1431*s*, 1433, 1444*s*, 1782, 1789, 1790*s*
- Dopamine β-hydroxylase 1064
- Dopamine β-monooxygenase 887
- Dopaquinone 1433*s*, 1434
- Dorsal protein 1899
- Dorsal-ventral (D-V) axes 1899
- Double displacement reactions 595
  - ping-pong mechanism 595
  - in ribonuclease 647
- Double helices 201, **213 – 218**
  - conformational flexibility 216
  - hybrid DNA-RNA 230
- Double-layered membranes
  - in *Nitrobacter* 1052
- Double reciprocal plot 460
- Doxorubicin 1217

Drawing arrows 528, 529
Drosophilia melanogaster 24
  bicoid mRNA in 1898
  development of 1898, 1899
  epithelial cells of 1885
  genome 12, 1507, 1511
  maternal effects in 1898
Drosopterins 803, 804*s*, 1460, 1461*s*
Drugs
  addictive 1796 – 1798
  psychotropic 1796 – 1798
  toxic 1796 – 1798
  toxicity of 1070
DSS or Tier's salt 138
Duty ratio, of molecular motors 1108
Dyad axis 134, 213
Dye, reducible 300
Dynamic equilibrium 289
Dynein(s) 370, 1119
  arms, figure 1121
  transport by **1119 – 1121**
β-Dynorphin 1752
Dysentery, *Shigella dysenteriae* 7
Dystroglycan 1112
Dystrophin 405, 1112
Dystrophin-glycoprotein complex
  drawing 1112

E

E (entgegen) configuration 43
E3 binding protein 796
Eadie-Hofstee plot 460
E-cadhedrin 574
Ecdysis-triggering hormone 1760
Ecdysone 1265*s*, 1266, 1760
Echinodermata 25
Eclosian 1760
Ecological matters 1908, 1909
ECoRI restriction endonuclease 1488*s*
Ectoderm 23, 1898
Ectoenzymes 409
Ectopic proteins 573
Eczema 1862
Edelman, Gerald M. 84
Edman degradation 118
EDTA. *See* Ethylenediaminetetraacetic acid
Effector(s) of allosteric enzymes 473 – 475
Efferent fibers 1766
Efflux carriers 1761
EF-hand motif 313, 317
EF-Ts 1702
EGF (epithelial growth factor), definition of 577, 1900
EGF-like domains 367
Egg
  amphibian 1897
  animal pole 1897
  development, figure 1896
  gray crescent 1897
  vegetal pole 1897
Ehlers–Danlos syndrome 438
Eicosanoid 565
Eicosapentaenoic acid 1190
Eigen, Manfred 84
Einstein 1274
Elastase 66, 609, 610, 611*s*
  cryoenzymology 616
  β-cylinder in 78
Elastic fibers 436
Elastin 15, 72, 436
Electrical double layer 400
Electric field jump methods 468
Electrochemical gradient 410
Electrochemical transference 311
Electrode(s)
  calomel 300
  reference 300
Electrode potentials **300 – 302**
  measurement of 302
  of mitochondrial electron carriers 1035
  observed 301
  standard, table 300
Electroencephalograms 1802
Electrogenic exchange 423, 1047
Electromagnetic spectrum, diagram 1274
Electromotive force of electrochemical cell 300
Electron acceptors, artificial, table 1025
Electron carriers 512 – 514, **767– 802**, 843 – 850, 883
  sequence of 1019 – 1026
Electron crystallography 131
Electron flow
  cyclic, in bacteria 1313
  reverse, thermodynamics of 1034 – 1036
Electron micrograph
  of bacteria 4
  of bacterial nucleoid 1531
  of cell junctions 27
  of chloroplast 1302
  of chromatid 1532
  of heteroduplex DNA 1489
  of minichromosome 1532
  of nuclear pores 1536
  of plant cell 13
  of rRNA precursor 1639
  of starch granules 172
  of transcription operon 1611
  of viruses 246
Electron microscope tomography 130
Electron microscopy 129 – 131, 1777
  of bacteriophage 1488, 1489
Electron-nuclear double resonance (ENDOR) 399
Electron paramagnetic resonance (EPR, ESR) spectra **398**, **399**
  first derivative plot 398
  of putidaredoxin 860
  of tetramethylpiperidine-1-oxyl 398
Electron transfer
  biological, mechanism of **848 – 850**
  coupling 849, 850
  gating 849, 850
  ionic equilibria, effects of 850
  pathways 848
Electron transferring flavoprotein (EFT) 514, 794
Electron transport **1019 – 1032**, 1300, 1301
  of bacteria 1025
  in chemolithotrophic organisms 1050
  equilibria 1035
  in fungi 1023
  Gibbs energy changes, table 1035
  in plants 1023
  proton pumps of 1039
  system, composition of 1019
Electron transport chain 512 – 515, 1013, **1019 – 1032**
  components, table 1019
  history of 1022
  of mitochondria 513, 1019
    diagram 1020
  observed potential 1034
Electron-transport proteins 514, 767– 802, 819, 843 – 850
Electroneutral exchange 1047
Electronic absorption band
  Gaussian curves, fitted with 1281
  log normal distribution curve, fitted with 1282
Electronic energy of molecules 1276
Electronic noses 122
Electronic transitions
  classification of 1282
  intensities of 1282
Electropherogram, single erythrocyte 107
Electrophilic center
  definition of 589
  in enzymes 754 – 757
Electrophilic reagents
  reactions of nucleic acids 253 – 255
Electrophoresis 58, **106 – 108**
  agarose gel 106
  capillary 106, 107
  diagonal 120
  pulsed field 249, 250
  SDS-polyacrylamide gel 106, 113
  zone 106, 249
Electrophoretic injection 1777
Electrophoretic molecular sieving 106
Electroporation 1498
Electrostatic effects
  on pK_a values 330
  titration curves of proteins 330
  transmission through aromatic rings 330
Electrostatic theory 311
Element(s)
  essential 33
  in human body 31
  toxic 33
Elemicin (nutmeg) 1439*s*
Elicitin 1187
Elicitors 1149
Elimination reactions 526, 530, **677– 690**
  beta, of cystine residues 85
  conjugative **689**
  decarboxylative **689**
  facilitation by carbonyl group 681
  of γ substituent 746
  of PLP-dependent enzymes 742
  reversibility 690
Ellman's reagent 125, 125*s*
Elongation factor 1673, 1685
Elongation factor EF-Tu 558, 1702
Elongin 1637
Elongin complex 564
E_m, measurement of 1039
  fluorescent probes 1039
  microelectrodes 1039
  permeant ion distribution 1039
Embden, G. 960
Embden–Myerhof–Parnas pathway 508, 510, 969
Embryo(s)
  anterior-posterior axis 1898
  dorsal-ventral (D-V) axes 1899
  green plants 1904
  homeotic genes 1899
  patterns 1899
  plant 30
  right-left axes 1899

  - signaling pathways 1899
  - Spemann's organizer 1899
- Embryonic development **1897 – 1900**
- Embryonic germ cells 1886
- Emission spectrum
  - fluorescence 1288
  - riboflavin tetrabutyrate 1289
- Emphysema 629
  - treatment of 631
- Enamel 442
- Enamelins 442
- Enamelysin 627
- Enamines 526, 528*s*
- Enantiomers 41, 164
- Enantiotopic atoms or groups 479
- Endergonic reactions 286
- Endocrine glands 1743
- Endocytosis 13, 425 – 427
- Endoderm 23, 1898
- Endoglin 1902
- Endoglucanases 602
- Endoglycanases 593, 602
- Endomorphins 1752, 1797
- Endonuclease. *See also* Ribonuclease and Deoxyribonuclease
  - homing 1717
  - restriction 652, 653, 1486 – 1489
- Endonuclease III 1582
- Endophilin I 1781
- Endoplasmic reticulum (ER) 10, 13, 14, 521, 1098
  - cisternae of 14
  - definition of 14
  - in micrograph 13
  - rough 14
  - smooth 14
- Endo ring conformation 212
- Endorphin(s) 1797
- β-Endorphin 1744, 1752
- Endosomes 13, 426
- Endosperm 30
- Endothelial cell-derived relaxing factor. *See* Nitric oxide, NO
- Endothelins 1752, 1753*s*
- Endothelium-derived relaxing factor 1756
- Endotoxins 431
- *cis*-Enediol 693*s*
- Enediol; Enediolate 694*s*, 695*s*
  - intermediate 693
- Energy (*E*) 282
  - from inorganic reactions 1050 – 1057
  - solar 289
  - sources of 507
  - storage and utilization of 302, 303
  - transfer, efficiency of, graph 1292
- Energy-linked processes 1037
  - in mitochondria 1046
- Enhancer(s) 576, 1535, 1628
  - bacterial 1614
  - elements 1631
  - figure 1632
- Enkephalin(s) 54*s*, 1795, 1797
- Enkephalin convertase (carboxypeptidase N) 625
- Enolase 685, 962
  - active site 685*s*
  - metal ion requirement 685
  - relaxation of water protons 685
  - superfamily 686
- Enolate anion 46, 46*s*, 681, 698 – 710
  - formation of 526
  - as nucleophile 527
  - from oxosteroid 696
  - phosphorylation of 706
  - resonance-stabilized 494
- Enolic intermediates 530, **691– 699**
- Enoyl-acyl carrier protein reductase 777
- Enoyl-CoA 940
- Enoyl-CoA hydratase 940
- Enoyl hydratase (crotonase) 681
- Enoyl reductase 1194
- 5-Enoylpyruvylshikimate-3-phosphate (EPSP) 687, 1424*s*, 1425*s*
  - EPSP synthase 1424
- *Entamoeba*, PEP carboxytransphosphorylase of 706
- *Entamoeba histolytica* 19
- Enterobactin (enterochelin) 838*s*, 1425
- Enterokinase (enteropeptidase) 609
- Enteron 24
- Enthalpy change(s) 282, 283
  - of combustion 282, 283
  - estimation of 297
  - of dissociation, table 293
- Entner–Doudoroff pathway 697, 965, 966
- Entropy (*S*) 284, 285
  - of gases 284
  - from heat capacities 285
  - of racemization 284
  - relationship to microscopic disorder 284
  - table 285
  - units, definition of 284
- Environment (thermodynamics), definition of 282
- Enzymatic reactions
  - catalysis 482 – 497
  - classification of 530
  - enolic intermediates **691– 699**
  - first order 457
  - integrated rate equations 457– 460
  - pH, effect on 469 – 471
  - progress curve 456
  - reversible 458, **464 – 468**
  - second order 458
  - stereoelectronic effect 495
  - velocity of 455 – 458
- Enzyme(s) 1, 2, **455 – 503**
  - activation of **471– 478**
  - active site 491– 494
  - biotin-containing 724, 725
    - table 724
  - catalysis by 482 – 497. *See also* Catalysis (enzymatic)
  - classification 497
  - conformational changes in 481
  - constitutive 538
  - dead-end complexes 466
  - debranching 606
  - effectors of 474
  - equilibrium 467
  - feedback inhibition, figure 540
  - history of 456
  - hydrogen bonding in 491– 494
  - induced fit 481
  - inducible 538
  - inhibition of **471– 478**
    - competitive 471, 472
    - irreversible 477, 478
    - noncompetitive 473 – 475
  - inverting 593
  - isologous dimers 349
  - kinetic parameters 459
  - kinetics of action **455 – 471**
  - lysosomal 1159
  - mechanism
    - iso 466
    - ping-pong 465
    - reciprocating (flip-flop) 497
  - modifiers of 474
  - molecular activity of 457
  - in molecular cloning 1491, 1493
  - molybdenum-containing 890 – 893
    - Fe-S clusters in 862
  - multisubstrate, kinetics of 464, 465
  - oligomeric 475, 497
  - pacemaker 535 – 537
  - processive action 602
  - prochiral centers **478 – 481**
  - proofreading 482
  - pyridoxal phosphate, coenzyme table 743
  - rate of encounter with substrate 461–464
  - rate equation for 460
  - reaction rate 455 – 458. *See also* Kinetics
  - regulation. *See* Enzyme regulation
  - retaining 593
  - saturation of 457
  - selenium-containing 822 – 827
  - specific activity of 457
  - specificity of **478 – 482**
  - starch-hydrolyzing 605 – 608
  - stereospecificity 478 – 481
  - substrate binding 463, 464
  - tetrameric 342
  - thiamin-dependent, structures 733
  - thiamin diphosphate-dependent, table 735
- Enzyme-activated inhibitors 682, 864
- Enzyme assay, coupled 768
- Enzyme-inhibitor complex 472
- Enzyme-linked immunoabsorbent assays (ELISA) 1849
- Enzyme regulation 535 – 537, 539 – 553, 571– 581
  - allosteric control 539 – 541
  - by allosteric effectors 475 – 477
  - by covalent modification 541– 545
  - by group transfer modification 545 – 549
  - by hormones 553 – 563
  - by ions 549
  - by thiol-disulfide equilibria 549
- Enzyme-substrate (ES) complexes
  - formation and reaction of 458 – 461
- Enzyme synthesis, genetic control of. *See* Protein synthesis
- Eosinophils 26, 188, 1074, 1843, 1863
- Eotaxin 1754*s*, 1849
- Eperimization
  - of UDP-N-acetylmannosamine (UDP-ManNAc) 1136
- Ephedrine 1791*s*
- Epidermal growth factor (EGF) 544, 571, 1753
- Epidermin 1722
- Epidermis 439
- Epidermolysis bullosa 440
- Epigenetic inheritance 1881
- Epigenetic markers 1881
- Epigenetic reprogramming 1881
- Epilepsy 1811
- Epimer, definition of 163
- Epimerases 692, 778, 1136

Epimerization
  of glycosaminoglycans 177
  of UDP-galactose, mechanisms 778
Epinephrine. *See* Adrenaline
Epiregulin 1753
Episomes. *See* Plasmids
Epithelial cells 29
  development of 1902
Epitheliocytes 25
Epitopes 1850, 1858
Epoxide, alkyation by 254, 1062
Epoxide hydrolases 591
EPR (electron paramagnetic resonance) spectroscopy 398, 399
  of glutamate mutase 873
  in study of phosphotransferases 639
EPSP (enolpyruvoylshikimate-3-phosphate) 687*s*, 1424*s*, 1425*s*
  synthase 687, 1424
Epstein-Barr virus 247
Equilibrium
  in carboxylation reactions 725
  of dehydrogenases in cytoplasm 980
Equilibrium constant(s) 304
  for association **325 – 332**
  for binding of metal ions **307**
  for carboxylation of propionyl CoA 725
  changes with temperature 289
  from concentrations 288
  effect of pH 293 – 297
  for hydrolysis 292
  prediction by Hammett equation 309
  as ratio of rate constants 458
  relationship to Gibbs energy change 287
  stepwise, for complex formation 307
Erabutoxin 1770*s*, 1775
c-erb 571
v-erb 571, 572
Ergocalciferol (vitamin D_2) 1257
Ergosterol 1246, 1248*s*
Erhlich, Paul 473
ERK 577– 579
Erucamide 1758
Erythrocruorins 68, 362
Erythrocytes 26, **404 – 406**
  cytoskeleton 405
  defects 826
  integrin 405, 406
  membrane skeleton 406
  NADP$^+$ from 767
  size of 11
Erythrocyte membranes
  band 3 protein 420
  proteins of 403
Erythroid cells 1901
Erythromycin(s) 1164, 1216, 1217*s*, 1690
Erythropoietin 1755, 1834, 1901
Erythrose 163*s*
Erythrose 4-phosphate 1136, 1422*s*
  in biosynthesis 965
Erythrulose 164*s*
Escherichia coli **3**, 628, 990, 1340, 1673
  chromosome map 1484
  composition of 31
  DNA of 4
  envelope structure 428
  genes of, table 1485
  genetic map of, figure 1484
  genome 12, 1486
  respiration with nitrite 1054
Essential amino acids 1369
Essential trace elements 31, 33, 837, 877– 881, 887– 894
Ester(s)
  biosynthesis of 1185 – 1196
  displacement reactions of 608
  mechanism of formation 661, 662
  resonance energy 299
Esterase(s) 526, **634 – 637**
Estradiol 1256*s*
Estrogen(s) 1260
  stimulates breast cancer 1263
  target tissues for 1260
Estrone 1256*s*
Ethambutol 1194*s*
Ethanol 508, 1796, 1797
  as fermentation product 966
Ethanolamine 99, 995, 1194*s*
Ethanolamine ammonia-lyase 871
Ether lipids 382
Ether phosphatides 383*s*
Ethidium 223*s*
Ethidium bromide 221, 222
Ethionamide 1194*s*
Ethylene 1761
  plant hormone 1389
Ethylenediamine, binding constants to metal ions 311
Ethylenediaminetetraacetic acid (EDTA) 98, 311*s*
Ethylene-responsive elements 1390
N-Ethylmorpholine 99
17-Ethynylestradiol 1262, 1262*s*
ets 576
Eubacterium 3, 6
Eugenol 1439*s*
Euglena 19
Euglenophyta 20
Eukaryote(s)
  differentiation in 1890 – 1893
  unicellular 18 – 22
Eukaryotic cells 11– 15
  definition of 2
  iron uptake 839
  structure 10
Eukaryotic genes
  cloning of 1497– 1499
Eukaryotic organisms, genetics of **1501– 1507**
Eukaryotic plants 1308 – 1310
Eukaryotic polymerase β 1548
Eumelanin 1434
Europeum 1292
Evenimicin 1691
Evolution 9
  of eukaryotes 15 – 18, 1510 – 1513
Exchange, across mitochondrial membranes 1047– 1050
Exchange diffusion 410
Exchange rates, amide protons 148
Exchange reactions of enzymes 595
Excisionase 1572
Excitation spectrum
  riboflavin tetrabutyrate 1289
Exciton 1306
Excitotoxicity 1787
Exergonic reactions 285
Exocytosis 14, 427, 1779
  control of 1780
Exoglycanases 593
Exons 1540, 1628
λ-Exonuclease 1491
Exo ring conformation 212
Expansins 1150
Exponential growth rate constant 470
Expression systems of cloned genes 1497
Exteins 1716
Extensins 1149
  glycosylation of 181
  in plant cell wall 72
Extracellular matrix (ECM) 406, **427– 441**
  reelin 1903
Extradiol aromatic dioxygenases
  intermediates, scheme 1058
Extradiol dioxygenases
  iron ions in 1058
Extrinsic mortality 1906
Eye 1324, 1333
Ezrin 1119

F

Fab fragments. *See* Immunoglobulins , Fab fragments
Fabry disease 1170 – 1172
F-actin 1098
F-actin filament
  with myosin heads 1103
FAD (flavin adenine dinucleotide) 511, 511*s*, 765, 780, 781*s*, 785, 980. *See also* Flavin coenzymes
$FADH_2$ 529, **529*s***
F agent 1482
  integration of, figure 1483
Familial adenomatous polyposis 574
Familial advanced sleep phase syndrome 1808
Familial British dementia 1812
Familial hypercholesterolemia 1249, 1251
Fanconi anemia 1586
Faraday, numerical value of 283
Farber disease 1171
Farnesene 1234, 1234*s*
Farnesyl diphosphate 1236
Farnesyl diphosphate synthase 1230
Farnesyl group 402, 559
Farnesyltransferases
  structure of 1231
Fast axonal transport 1119
Fast-twitch fibers
  of muscle 1096
Fasting state 1002, 1003
Fat(s). *See also* Triacylglycerol (triglyceride)
  composition of 380
  conversion to sucrose in oil seeds 988
  development of 1902
  formation from excess carbohydrate 1003
  hydrolysis of 507
  thermogenic 1048
Fatal familial insomnia (FFI) 1718
Fatty acid(s) **380 – 382**, **1185 – 1196**
  activation of 512
  acyl CoA, derivatives of 507
  anteiso 1189
  biosynthesis of 722, 977– 979, **1185 – 1196**
  biosynthetic reactions, scheme 1188
  branched chain 381
  branches 1189
  control of chain length 1188, 1189
  control mechanism, scheme 1001

  cyclopropane-containing 399, 1193 – 1195
  essential 721, 1190, 1191
  esters 1185 – 1196
  human disorders 944
  iso 1189
  in lipids 380
  names of, table **380**
  oxidation 511, **939 – 947**
    disorders 944, 945
  pK_a values of 380
  polyunsaturated 1190
  stability of 589
  starter pieces for biosynthesis of 1189
  stimulation by carbon dioxide 977
    control mechanism, scheme 1001
  synthesis by oxoacid chain elongation 1189
  synthesis pathways 1192
  unsaturated 1191– 1193
Fatty acid amides 382, 1756, 1792, 1793
Fatty acid amide hydrolase 1758
Fatty acid-binding proteins 1184 – 1186
  figure 1184
Fatty acid esters **1186 – 1196**
Fatty acid synthases 990, 1185 – 1187
Fatty acyl-CoA derivatives 507
Fatty acyl-CoA synthase
  in outer mitochondrial membrane 1016
Fatty alcohols **380 – 382**
Favin 64*s*
Fc fragment 1836
Fecapentaene 1587, 1587*s*
Feedback inhibition 536, 539
Feedback repression 536, 1612
Female sex hormone 1260
FeMo-coenzyme (FeMo-co) 1361
Fermentation(s) 8, 298, 508 – 511, 535, **966 – 969**
  butyric acid 971
  coupling of oxidation and reduction 966
  Embden–Meyerhof pathway **966 – 969**
  energy relationships 966
  Entner–Doudoroff pathway 965, 966
  of glucose, Gibbs energy change 298
  of glutamate 1371
  glyceraldehyde-3-phosphate dehydrogenase in 775, 776
  heterolactic 972
  homolactic 980
  mixed acid 968 – 970
  mixed acid products, table 968
  pentose phosphate pathway 972
  phosphogluconate pathway 972
  phosphorylation 966
  propionic acid 970, 971
Ferns 18
Ferredoxin 793, 857, 858*s*
  from chloroplasts 859
  in methane bacteria 1057
  reduction potentials 857
  in reductive biosynthesis 981
Ferredoxin/thioredoxin system
  in photosynthesis 1320
Ferredoxin:NADP+ oxidoreductase 788, 794, 1317
Ferric-dioxoisobacteriochlorin 1055
Ferrichrome 838*s*, 839
Ferric uptake regulation protein 839
Ferriheme(s) 844
Ferritin 343*s*, 841*s*, 841, 842
Ferrochelatase 843, 1402
Ferroelectric crystals 486
Ferroheme(s) 844
Ferrous ion in aconitase 688
Ferroxidase center 843*s*
  in ferritin 842
Ferrulate 1439*s*
Ferryl iron in peroxidase and catalase 854
Fertilization 1895 – 1897
  meiosis 1895, 1896
  membrane 1896
Fetal hemoglobin 362
α-Fetoprotein 573
Ffh (54 homolog) 520, 1722
FhuA, C, D 839
Fiber cells 30
Fibrillarin 1641
Fibrillin 440
Fibrin 631, 632
  stabilizing factor 632
Fibrinogen 58, 409, 578, 632
Fibroblast(s) 25, 433
  micrograph of 370
Fibroblast growth factors (FGFs) 1755
Fibrocytes 25
Fibromodulin 1154
Fibronectin 407, 409, 1883
  domain structure 409
Fibrous dysplasia 576
Fibulin 437
Ficin 618
Field flow fractionation 103
Fifth disease 244
Filaggrin 370, 439
Filamin 370
Filaria worms 24
Fimbriae 6. *See also* Pili
Finasteride 1255, 1255*s*
Fingerprinting. *See also* Peptide mapping
  of DNA 259
  of proteins 118, 360
First Law of Thermodynamics 282
First order reactions 457
Fischer, Edmond H. 84
Fischer, Emil H. 42, 83
Fischer projection formula 42
  of monosaccharides 163
Fission yeasts 1503
Five-carbon branched units, formation scheme 992
FixL 1614
FK506 488
Flagella
  bacterial 3, 6, 562, **1089 – 1093**
  eukaryotic, structure 15, **1121**
  growth of 1091
  micrograph of 1090, 1091
  rotation of 1091– 1093
Flagellar motor
  bacterial, drawing 1092
  drawing of rotor and stator unit 1093
Flagellin 1089
Flatworms 23
Flavanoid compounds 1365
Flavin(s) 1460
  chromatography of 103
  content in mitochondria 1019
  electrophilic centers 790
  fluorescence 103
  half-reduced 791
  isoalloxazine ring 790
  metal complexes 794
  properties of 793
Flavin adenine dinucleotide. *See* FAD
Flavin adenine diphosphate. *See* FAD
Flavin coenzymes 766, **780 – 795**
  modified 788, 789
  reduced 794
Flavin cofactors 1059
Flavin radicals 792
  color of 794
  formation constant 794
Flavocytochrome b_2 782, 794, 847
Flavocytochrome b_{558} 1072
Flavodoxins 793, 799, 800
Flavones 1214, 1215*s*
Flavonol(s) 1214, 1215*s*
Flavonol pigments 1214
Flavonones 1214, 1215*s*
Flavoprotein(s) 513, 788
  dehydrogenation reactions 782 – 787
    table 784
  reduction potentials of 782
Flavoprotein dehydrogenases 789
  free radicals 793
  mechanisms 789 – 791
Flavoprotein oxygenases
  free radicals 793
FliF protein 1093
Floor plate 1903
Florigen 1761
Flow cytometry 107
Flow of materials within cells 289
Flower pigments 1214, 1215
FLP (Flip) recombinase 1572
Flukes 24
Fluorescamine 120*s*
Fluorescence 1288 – 1294
  action spectrum 1288
  in analysis 120
  analytical applications of 1293
  correlation spectroscopy 1294
  efficiency 1290
  efficiency equation 1291
  emission spectrum 1288
  excitation spectrum 1288
    of LH2 1307
  of flavins 103
  green fluorescent protein 1344
  lifetimes 1290
  probe, polarity-dependent 396
  quantum yield 1290
  quenching, molecular beacons 1291
  recovery after photobleaching 400
  resonance energy transfer (FRET) 1111, 1292, 1741
  time-resolved spectroscopy 1290
  of vitamin B_6 coenzyme 750
Fluorescent
  antibodies to cytochrome *c* 1016
  probes 1294
  tag 1344
Fluoride ion
  inhibition of acid phosphatases 645
Fluoroacetate 957
Fluorographs 103, 252
5-Fluorouracil 812
5-Fluorouridine 574
Fluoxetine 1793*s*
Flurazepam 1788*s*, 1789
Flux
  across membranes 410
  in metabolism (analysis) 536, 537

FMN (Riboflavin 5'-phosphate) 513, 765, 780, 781*s*
fMRI 1802, 1806, 1807
Focal adhesions 405, 406, 1884
Fodrin 405
Folate coenzymes, 803, 804. *See also* Folic acid
Folate receptors 805
3-Fold axis 333
Folding pathway 1727
Folding patterns
  of polypeptides **76**, **77**
    effects of sequence 78
Folic acid (pteroylglutamic acid) 473, 557, 721, **802 – 813**, 1425
  biosynthesis, scheme 1461
  deficiency of 802
  metabolic functions 802
  nutritional requirements 756
  in transmethylation 592
Folistatin 1903
Folitropin (FH) 1745
Follicle cells 1895
Follicle-stimulating hormone (FHS) 1260
Follicular dendritic cells 1834
Folyl polyglutamates 803
Footprinting
  DNA 266
  hydroxyl radical 267
Foraminifera 19
Forces
  acting between molecules **46 – 51**
  electrostatic 47
  van der Waals 46
Formaldehyde 802
  hydration equilibrium 677
  reaction with nucleic acid bases 253
Formate. *See also* Formic acid
  carbon units, source of 810
Formate dehydrogenases 824, 890, 892, 893, 969
  in methane bacteria 1057
  molybdopterin 892
Formation constants 304. *See also* Equilibrium constants
  macroscopic 328
  of metal complexes 312
  for Schiff bases 679
  stepwise 327, 328
Formic acid 99, 293, 802
Formiminoglutamate 810, 1450*s*
Formiminoglycine 810
Formylglycine in sulfatases 659
Formylkynurenine 1059*s*, 1445*s*
Formylmethanofuran 814*s*
N-Formylmethionine 231, 1388, 1698
6-Formylpterin 803
10-Formyltetrahydrofolic acid (10-formyl THF) 810
Formyl-THF synthetase 810
Förster equation 1292
Fos 70, 241, 572, 577
fos 572, 576, 579
c-Fos 1633, 1902
Fossils of bacteria and blue-green algae 17
Fourier transform infrared spectroscopy (FTIR) 1277
  isotope-edited 1278
  of membranes 396
  total reflection 1277
Fourier-transform infrared (FTIR) spectrum
  concanavalin A 1278
  myoglobin 1278
F pili 1482
F plasmids 1559
Fractionation of cells and tissues 98 – 100
Fragile X syndrome 228, 1516
Frame-shifting 538
Frameshift mutations 1580
Franck-Condon principle 848, 1280
Franklin, Rosalind N. 200, 1474
Frataxin 1516
Freckles 1435
Free energy. *See* Gibbs energy
Free radicals. *See also* Superoxide anion radical, nitric oxide (NO)
  in action of peroxidase and catalase 855, 856
  in coenzyme M reductase 880
  flavin 791– 794
  in metalloenzymes 850 – 886
  in pyruvate formate-lyase 800, 801
  in pyruvate oxidase 736, 799
  in reactions of vitamin B_{12} 869 – 876
  in ribonucleotide reductases 864, 865
Freeze-fractured membranes, figure 405
Freeze fracturing 130
Freidreich's ataxia 1516
FreIp metalloreductase 839, 883
FRET 1726
Frictional coefficient 109
  of sphere 109
Fringe 1900
FRN protein 1613
Fructans 1152
Fructofuranose (Fruf) 165*s*
Fructofuranosyl-enzyme 597*s*
Fructokinase 1129
Fructose (Fru) 162*s*, 175, 1131
Fructose-1,6-bisphosphatase 645, 1000, 1144, 1320
Fructose-2,6-bisphosphatase 646, 1320
Fructose-1,6-bisphosphatase-phosphofructokinase cycle 1000
Fructose 2,6-bisphosphate (Fru-2,6-P_2)
  beta form 541*s*
  in regulation of glycogen metabolism 537
Fructose bisphosphate aldolases 699
Fructose 1-phosphate 1129
  cleavage by aldolase 1129
Fructose 6-phosphate 535, 693*s*, 1135*s*
  equilibrium with mannose 6-phosphate 1129
  metabolism of 1135, 1136
Fructose 6-phosphate kinase 656
Fruit fly. *See Drosophilia melanogaster*
FtsZ 1890
Fucose (Fuc) 165*s*, 1155, 1157
L-Fucose (6-deoxy-L-galactose)
  metabolism of 1138
L-Fucose isomerase 693
α-L-Fucosidase 1171
Fucosidosis 1171
Fucosyltransferase 184
Fucoxanthin 22, 1239*s*, 1240, 1308
Fumarase. *See* Fumarate hydratase
Fumarase-aspartase family 685
Fumarate 481*s*, 516*s*, 683*s*
Fumarate hydratase (fumarase) 526, 683, 688
  acid–base catalysis 471
  concerted reaction 685
  fumarase A, B 688
  fumarase C 683
  mechanisms 683 – 685
  pH dependence 684
  rates of substrate exchange 684
  turnover number of 683
Fumarate reductase 785, 1027, 1054
Fumarylacetoacetate 1431*s*
Fumarylacetoacetate hydrolase 1430
Fumarylpyruvate 690*s*, 1436*s*
Functional MRI (fMRI) 1801, 1802, 1806, 1807
Function of state *R* 476
Fungal infections 20
Fungal ribonucleases 1686
Fungi **20**
  cell walls 443
Fur (ferric uptake regulator) 1612
Furanoeudesma-1,3-diene 1235*s*
Furanose rings 161, 175. *See also* Sugars
Furfural 251, 167*s*
β-Furfuryl-β-glucoside 1148*s*
Fusidic acid 1266, 1266*s*, 1703

G

GABA (gamma-aminobutyrate) 1793
GABA receptors 1788
Gabaculline 738*s*
G-actin 1098
GAL4 1725
Galactan 177
Galactanase 602
Galactinol 1132, 1141*s*
  biosynthesis 1141
Galactitol 1130
Galactocerebrosidase 1171
Galactokinase 1129, 1130
Galactolipid 387*s*, 1197
Galactonate 1130
Galactosamine 176
Galactose (Gal) 162, 163*s*, 1129, 1157
  metabolism of 778, 1130, 1131, 1138
Galactose chemoreceptor protein 186
  active site 187*s*
Galactose mutarotase 1131
Galactose (gal) operon 1613
Galactose oxidase 885
  active site 886
  mechanism of 887
  tyrosine-cysteine cofactor 885, 886
Galactose 1-phosphate 1129, 1130
Galactose-1-phosphate uridyltransferase 1130
Galactosemia 16, 169, 1130, 1131
α-Galactosidase 1171
β-Galactosidase 602, 603, 1171, 1475
Galactosyl diacylglycerol 1301
β-Galactosyl hydrolase 1171
Galanin 1750
Galectins 186
Gallate 1438, 1438*s*
Gallic acid 1438
Gallstones, cholesterol 1251
GalNAc 1155
GalNAc kinase 1136
gal operon 1604
Gametes 17, 1504, 1893 – 1895
Gametophyte form of fern 18
Gametophytic cycle of plants 1904
Gangliosides 388, 1168, 1169
  of cholera toxin 546
Gangliosidosis, G_{M1} 1171

Gap genes 1900
Gap junction(s) 29
  image of 1782
Gardner syndrome 1586
Garlic 1408
Garrod, Archibald 16
Gas constant, numerical value of 283
Gastricsin 621
Gastrin 1741, 1749
Gastrin-releasing peptide (GRP) 1750
Gastrointestinal inhibitory peptide (GIP) 1749
Gastrointestinal peptides 1749, 1750
Gastrotropin 1749
Gastrula 1898
  formation, figure 1896
Gated pores 411
Gaucher disease 1170 – 1172, 1518
Gaussian curves
  fitted to electronic absorption band 1281
  parameters 1281
  G-banding, of chromosomes 1512
GDP. *See* Guanosine diphosphate
GDP dissociation inhibitors (GDIs) 559
Gel electrophoresis 112
  relative molecular mass by 112
Gel filtration 100, **112**
  oligosaccharides, separation of 100
  relative molecular mass by 112
Gelonin 1685
Geminal diamine 751, 751*s*
Geminiviruses 244
Gene(s)
  amplification of 538
  antennapedia gene 1900
  blunt ends 1491
  cloning **1490 – 1501**, 1517
  cohesive ends 1490
  gap 1900
  homeotic 1900
  homologous 1510
  human, number of 16
  human variation 1509
  imprinting of 234, 1542
  inactivation of 1880
  linked 1504
  mitochondrial, catalog 1016
  modifying 1500, 1501, 1518, 1519
  nitrogen fixation 1365
  orthologs 1510
  overlapping 1540
  paralogs 1510
  random segregation 1504
  segmentation 1900
  selector 1900
  sequence(s) 1508 – 1513
  sequence for aspartate aminotransferase 202
  structural, definition 1
  suppressor 1480, 1480, 1481
  transferring **1490 – 1501**
Gene A protein 1558
Gene conversion. *See* Recombination: nonreciprocal
Gene therapy 16, 203, 1518
  sickle cell disease 361
Gene transfer
  physical methods 1498
Genetic code 1, **236 – 239**
  deciphering of 1475, 1476
  tables 236, 237
  variations 237
Genetic diseases 2, 16, 1513 – 1517, 1585 – 1587, 1907. *See also* Hereditary diseases
Genetic engineering 1366
Genetic map
  AIDS virus HIV-2 1652
  of cloning plasmids 1493
  of *E. coli*, figure 1484
Genetic methods 1476 – 1481
Genetic recombination 6, 230, 1482, 1577, 1578
  model of 1569
Genome(s)
  bacterial 3 – 5, 201
  changing 17
  definition of 4
  human 201, **1507– 1509**
  mammalian 1017, 1508, 1510
  sequences 12, 1508 – 1513
  table of 12
  of viruses 201, **244 – 248**, 1714, 1715
Genomic imprinting 1881
Genomic libraries 1499
Gentamycins 1139, 1690
Gentisate 1436*s*
Geometrical isomers 43
Geometry of orbitals
  effect on metal binding 311
Geotropic response of plants 1761
Gephyrin 1787
Geranyl diphosphate 1233*s*
Geranylgeranyl diphosphate 1234*s*, 1236
Geranylgeranyl group 402
Geranylgeranyltransferases 1231
Germ cells 1893 – 1895
Germ layers 1898
Germin 1904
Gerstmann–Sträussler–Sheinker syndrome (GSS) 1718
Giardia lamblia 12, 19
Gibberellin(s) 1150, 1234, 1234*s*, 1761
Gibberellin A_1, biosynthetic sequence 1234
Gibbs energy (Free energy) 285, 286
  of activation ($\Delta G^{\ddagger}$) 483
  of association 325 – 332
  of dissociation 293
    of acetic acid 50
  of formation 286, **292 – 298**
    for biochemical substances, table 290
    of RNA helix, table 210
  of hydrolysis 283
    table 294
  in hydrophobic interactions 50
  linear relationships 308
  of oxidation
    for biochemical substances, table 290
    by NAD$^+$, table 290
  partial molar 287
  summing 286
Gibbs energy change(s)
  for dilution 287
  and equilibrium constant 287
  for fermentation 298
  for oxidation of glucose 967
  for oxidation-reduction reactions 300
  for subunit association 333
Gingipains 619
Gla proteins 821
GlcNAc 1155
GlcNAc-6-sulfate sulfatase 1171
Glial cells 26, 1762, 1781
Glial filaments 369
Glicentin 1749
Global regulator 1612
Globin genes 1539, 1540, 1901, 1902
  evolution of 17
  organization, figure 1540
α and β Globins 1539
Globoid cell leukodystrophy 1171
Globular proteins, properties 59
Globulin(s) 101, 58
Glomeruli 1799
Glucagon 542, 1741, 1742, 1743, 1749
  adenylate cyclase in action of 554
  antagonism of insulin 999
  receptor 554
β-Glucanases 602, 603
Glucans **172 – 175**
  β-1,3-linked 174
Glucarate dehydratase 686
D-Glucitol. *See* D-Sorbitol
Glucoamylase 603, 607, 1146
  active site structure 608
  α,α barrel structure, 77*s*
Glucocerebrosidase 603, 1171
Glucocorticoid(s) 999, 1254, 1742
  release during fasting 1002
Glucogenic amino acids 1397
Glucokinase 538, 568, 998
Gluconeogenesis 515, 997– 1000
  adrenaline, effect of 999
  control mechanisms, scheme 1001
  glucagon, effect of 999
  pathway 509
    scheme 989
  regulation of 541
Gluconolactonase 963
5-Gluconolactone 604*s*
Glucopyranose-6-*P* 1132*s*
Glucosamine (GlcN) 164, 176
D-Glucosamine 6-phosphate 1135*s*
  biosynthesis 1135
Glucosamine 6-phosphate synthase 1135
Glucose (Glc) 973
  diastereomers 162
  equilibrium mixture in water 162
  metabolism 997
  ring form 162*s*
Glucose (D-glucose, Glc) 161*s*, 162*s*
  biosynthesis from three-carbon compounds 989, 990
  caloric value 283
  concentration in blood 161
  concentration in blood in diabetes 1004
  conversion to 2-oxo-l-gulonate 1134
  conversion to l-rhamnose 1138
  as energy source 997
  fermentation of 298
L-Glucose (L-Glc), 161*s*
Glucose 1-arsenate as intermediate in arsenolysis 595
Glucose 1,6-bisphosphate 653, 653*s*
Glucose dehydrogenase(s) 814
Glucose oxidase 782
Glucose-6-phosphatase 645
  deficiency of 1144
  glucose release from stores 999
Glucose 1-phosphate 989, 1130
  glycogen, conversion to 989
  stability of ring 992
Glucose 6-phosphate 506*s*, 507, 535, 693, 693*s*, 1131
  Gibbs energy of hydrolysis 293

  - pK_a value of 293
- Glucose-6-phosphate dehydrogenase 770, 776
- Glucose-6-phosphate isomerase
  - efficiency of 692
  - turnover number of 692
- Glucose tolerance factor
  - chromium in 888
- Glucose transporters 415, 416
  - GLUT1, topology diagram 416
  - GLUT2 998
  - GLUT4, response to insulin 416
- Glucosidase
  - alpha 603, 606, 1144, 1171
  - alpha 1,6 606
  - beta 602, 603
- β-Glucoside 1238*s*
- Glucosinolate 1442, 1142*s*
- Glucosinolides 1142
- Glucosylasparaginase (aspartylglucosaminidase) 620
- Glucosyl ceramide 1203*s*
- Glucuronic acid; glucuronate (GlcA) 164*s*, 165*s*, 176
  - epimerization 177
- D-Glucuronic acid 1133*s*
  - catabolism of 1132
  - metabolism of, scheme 1133
    - accumulation 1170
- β-Glucuronidase 1171
- Glucuronides 1141
  - biosynthesis of 1130
- GLUT1 – GLUT7. *See* Glucose transporters
- Glutamate 1782, 1785, 1786*s*, 1793, 1794
  - biosynthesis of 711
    - scheme 1373
  - catabolism of 1371– 1374
  - concentration in brain 958
  - fermentation of 1371
  - formation by transamination 952
  - isomerization of 1372
- L-Glutamate 737*s*
- Glutamate dehydrogenase 770, 775, 1369
  - inhibition by 2-oxoglutarate 780
- Glutamate mutase 871
  - stereochemistry 874
- Glutamate racemase 1370
- Glutamate receptors 1787
  - ionotropic 1786
  - ionotropic, agonists of 1786
  - metabotropic 1786
- Glutamate-1-semialdehyde 742, 1374, 1400*s*
- Glutamate synthase 788, 1369, 1370
- Glutamic acid (Glu, E) 52, 53*s*. *See also* Glutamate
  - pK_a value of 487
- Glutaminase activity 1376
- Glutamine (Gln, Q) 53*s*, 1135
  - biosynthesis of, scheme 1373
  - catabolism of 1371– 1374
  - in glucosamine synthesis 1135
- Glutamine PRPP amidotransferase 620
- Glutamine synthase
  - regulation, scheme 1370
- Glutamine synthetase 342, 662, 1370, 1371
- γ-Glutamylamino acid 1369
- γ-Glutamyl cycle 1369
- Glutaraldehyde 1098
- Glutaredoxin 522, 549, 786, 787, 864
- Glutaryl-CoA 1386, 1386*s*
- Glutathione 521, 549, **550 – 552**, 550*s*, 1588
  - as antioxidants 1075
  - biosynthesis 551, 993
  - protective functions 551, 786, 826, 1075
  - reaction with sulfur 1053
- Glutathione peroxidase 550, 826, 856
  - decomposes superoxide 1074
  - selenium in 823
- Glutathione reductase 785, 785*s*
  - reaction mechanism 791
    - scheme 792
- Glutathione *S*-transferases 550, 682
- Glutathionylspermidine 1382
- Glycans **169 – 180**. *See also* Polysaccharides
  - biosynthesis of **1146 – 1150**
  - β-1,3-linked 1148
  - N-linked 183
- Glycation 169
- Glyceraldehyde 163*s*, 1129
- Glyceraldehyde 3-phosphate 506*s*, 508, 694*s*, 986
  - oxidation of **508 – 510**, 980
- Glyceraldehyde 3-phosphate dehydrogenase 63*s*, 766, 984
  - in fermentation reactions 775, 776
  - formation of NADH-X 780
  - in oxidation of aldehydes 775, 776
  - thiol group of 677
- Glyceraldehyde:pyruvate pathway 1229
  - isoprenoid biosynthesis, scheme 1229
- D-Glycerate dehydrogenase, stereochemistry 768
- Glycerate pathway 960
- L-Glyceric aciduria 1397
- Glycerol 380*s*
  - accumulation in fish and insects 191
  - as fermentation product 967
  - fermentation of, by propionic acid bacteria 954
  - from glucose 968
  - in lipids 380
- Glycerol dehydratase 871
- Glycerolipids, synthesis pathway 1192
- Glycerol kinase 1129
- Glycerol oxidase 886
- Glycerol 1-phosphate 293
- *sn*-Glycerol 3-phosphate 383*s*
- Glycerol phosphate dehydrogenase 1023
- Glycerol-phosphate shuttle 1050
  - diagram 1051
- Glycerolteichoic acids 431*s*
- L-*Glycero*-D-*manno*heptose 180*s*, 1137
- Glycerophospholipids 383
- Glycerylcysteine 428, 428*s*
- Glycine (Gly, G) 52, 52*s*, 745, **1397– 1405**, 1782, 1793
  - antagonists of 1788
  - binding constants to metal ions 311
  - cleavage system 809, 1399
  - decarboxylation of 745
  - drugs binding to receptors 1788
  - metabolism of, scheme 1398
  - as neuroinhibitor 1789
  - pK_a value of 99
  - receptors 1789
  - reductases 753, 755
  - reductase system 1399
- Glycineamide, binding constant to metal ions 311
- Glycine betaine 1142, 1398
- Glycine decarboxylase 743, 745, 809
- Glycine decarboxylase-synthetase system 808
- Glycine *N*-methyltransferase 591, 593*s*
- Glycine reductase 824
- Glycocalyx 6
- Glycocholic acid 1251
- Glycogen 165, 170, 173, 173*s*, 603
  - biosynthesis of 509, 1130, 1143 – 1146
  - breakdown of 537
  - content in rat liver 31
  - conversion to lactate 960
  - debranching 1144
  - digestion of 1146
  - granules 1098
  - metabolism, genetic diseases of 1145
  - methylation analysis 196
  - ^{13}C NMR 174
  - particles in cytoplasm 10, 174
    - STM image of 174
  - storage of 537
  - storage diseases 1145
  - synthesis
  - synthesis in bacteria 1144
  - synthesis in plants 1144
- Glycogen phosphorylase 541– 545, 543*s*, **604**, **605**, 746, 1143
  - catalytic site 605*s*
  - conformational states 604
  - control of **541– 545**
  - deficiency of 1144
  - effect of insulin 571
  - regulation of 539, 604
- Glycogen synthase 1143
  - deficiency of 1144
  - effect of insulin 571
- Glycogenin 174, 1144
- Glycogenolysis 545
- Glycoglycerolipids 387
- Glycolaldehyde 1135*s*
- Glycolate 707
  - production of, scheme 1322
- Glycolate oxidase 782
- Glycolic acid
  - metabolism of 1321, 1322
- Glycolipids **387**, 392
  - biological recognition 161
  - biosynthesis of 1168, 1169, 1196 – 1203
    - scheme 1198
  - breakdown of 1169 – 1172
- Glycolysis 506, **508 – 511**, 699, **960 – 962**
  - control mechanisms, scheme 1001
  - pathway **960 – 962**
    - scheme 961
  - rate of 997
  - regulation of 541, 999
- Glycolytic pathway, scheme 509, 989
- Glycopeptidolipids 1168
- Glycophorin A 404
- Glycophorin C 405
- Glycophosphatidylinositol (GPI) anchors 1168
- Glycoprotein(s) 15, **180 – 188**, 404, 405, **1153 – 1160**, 1883
  - biological recognition 161
  - blood group determinants 184
  - extensins 1150
  - high mannose type 183
  - linkage of saccharides to 181
  - *O*-linked 181, 182
  - oligosaccharides, trimming of 1156, 1157
  - release of oligosaccharides 188
  - synthesis of **1153 – 1160**
  - variable surface 1866
- Glycoprotein hormones 1745

Glycosaminoglycans 181, **1152 – 1155**
    chains in proteins 1154
    modifications of 1153
    repeating disaccharide units of 176
    sulfation of 1153
    synthesis of 1153 – 1155
Glycosidases
    beta 602
    endo 605
Glycosides
    cyanogenic 1442
    displacement reactions of 593
N-Glycosides (glycosylamine) 169
Glycosidic acetal linkages 168
Glycosidic linkages 780
Glycosphingolipids
    biosynthesis of, scheme 1169
    catabolism of, scheme 1169
Glycosylamine 169
Glycosylation definition 169
Glycosyl enzyme intermediates 595
Glycosyl esters 595
Glycosyl exchange reactions 593
Glycosyl groups, active 975
Glycosyl units
    polymerization of, scheme 1130
Glycosylphosphatidylinositol (GPI) anchors 403
Glycosyltransferase(s) 184, 526, 604, 608, 1143, 1150
    catalytic groups, table 603
Glycyl endopeptidase 618
Glycylglycine
    binding constants to metal ions 311
    copper complexes 312*s*
    pK_a value of 99
Glycyl radical 864
Glyoxalase 527, 551, 698
Glyoxylate 705*s*, 1398, 1399, 1458*s*
    formation from isocitrate 988
    metabolism of 959
Glyoxylate carboligase 736, 959
Glyoxylate pathway **987– 989**
Glyoxylic acid
    in polysaccharides 179
Glyoxysomes 14, 988
Glyphosate 687*s*
G_{M3}-*N*-acetylgalactosaminyltransferase 1171
GMP (Guanosine 5'-phosphate) 203
GNRA loops in RNA 1638
Goiter 1432
Gold 318
Goldstein, Joseph L. 84
Golgi, Camillo 83
Golgi apparatus 10, 521, 1157
    definition of 14
    transverse section 1158
Golgi body in micrograph 13
Golgi network, trans 1157
Golgins 1159
Gonadotropin-releasing hormone 1746
    NMR spectrum 145
Gonorrhea 7
Goodpasture disease 438, 1865
Gossypol 1235*s*
Gout 1459
G_0 phase of cell cycle 579
G protein(s) 534, **557– 561**, 1673, 1704, 1741
    bovine brain G_i 560*s*
    coupled receptors 1325
    cycle 1781
    function, scheme 557
    monomeric 558
    Rho family 1148
    three-dimensional structures 559 – 561
G-quartets 208, 227
Gramicidin 414, 1164, 1842
Gramicidin S 54*s*, 1713
    biosynthesis 994
Gramine 1447, 1447*s*
Gram-negative bacteria 6, 7
    outer membrane 428 – 431
Gram-positive bacteria 6, 7, 185, 190, 428
Gram stain 6
Grana 1301
Granulocytes 26
Granzymes 610
Graves disease 1864
Grb2 577, 578, 580*s*
GreA,B 1610
Green fluorescent protein 1111, 1292, 1344
Green plants. *See* Plants (green)
Green sulfur bacteria 1306
Griseofulvin 1216*s*, 1217, 1690
GroEL, GroES 339, 518
Ground substance between cells 26
Group transfer potential(s) 292, 293
    effect of metal ions 296, 297
    of pyrophosphates 636
Group translocation 411
Growth 1879 – 1909
    cones of axons 1903
    control of 578
    essentials for, table 1881
    factors 571– 576, 722, 1751– 1754
    hormone 1742
    rates of cells 470
    regulating hormones 567– 571, 1743, 1751– 1754
Growth-factor receptors 1742
Growth hormone-releasing hormone (somatoliberin) 1746
G-segment 1552
GTP (guanosine triphosphate) 1672
    as coenzyme 719
GTPase-activating center 1687
    antibiotic action site 1691
    of ribosomes 1672
GTPase activating proteins (GAPs) 559
GTP cyclohydrolase 1460
Guam disease 1812
Guanidination 124
Guanidinium chloride 82*s*
Guanidinium groups 53, 55*s*
Guanidinoacetate methyltransferase 591
Guanidinoacetic acid 1379, 1379*s*
Guanine (Gua, G) 199*s*, 203, 1458
    hydration of 215
    quartet 208
Guanine nucleotide binding proteins. *See* G proteins
Guanine nucleotide dissociation stimulators (GDSs) 559
Guanosine (Guo) 203, 234*s*
    absorption spectra of 205
Guanosine diphosphate-L-fucose 1138
Guanosine diphosphate-L-galactose 1134*s*
Guanosine disphosphate-guluronic acid 1130
Guanosine diphosphate-mannose (GDP-Man) 1130, 1134*s*
Guanosine diphosphate-mannuronic acid (GDP-ManUA) 1130
Guanosine monophosphate 1455*s*, 1461*s*
Guanosine 5' phosphate (GMP, Guanosine 5'-guanylic acid) 203
Guanosine polyphosphates (ppGpp, pppGpp) 1715
Guanosine 5'-(β,γ-imido) triphosphate (GppNp) 558*s*
Guanylate cyclase 561, 657
Guanylin 548, 561, 1750
Guanylyl cyclase. *See* Guanylate cyclase
Guard cells 30
Guide RNAs 1017, 1641, 1648
L-Gulonic acid 1132, 1133*s*, 1134
L-Gulonolactone 1133*s*
Gulonolactone oxidase 1134
Gulose (Gul) 163*s*, 1132
Gurken protein 1899
Gustducin 558, 1799
Gutta-percha 389, 1231
Gymnosperms 29

H

Haber process 1359
Haemophilus influenzae
    genome sequence 12
Hageman factor 632
Hairpin ribozyme 649
Haldane, J. B. S. 9
Haldane relationship 464, 465
Half-life 457
Hallucinogens 1798
Haloacid dehalogenase(s) 590
    mechanism of 590
Haloalkane dehalogenase(s) 591
    active site structure 591
Halobacterial transducer I (HtrI) 1335
Halocyanin 883
Haloperidol 1792, 1809*s*
Haloperoxidases 855, 889
Halorhodopsin 1335
    light-dependent cycle 1335
Halothane 1789
Hammerhead ribozyme 649, 651*s*
    mechanism of action 651
Hammett equation 308
    for dehydrogenation reactions 771
    for multisubstituted compounds 309
    table of constants 308
Hammett's constants 1284
Handedness of molecules 41– 43
Haploid
    definition of 5
    phase 18
Haptens 1838
Haptoglobin 58, 1850
Harden, Arthur 83, 960
Harmine 1447
Hawkinsin 1430, 1430*s*
Haworth structural formulas 164
Hay fever 26
Hearing 1801
Heart disease 16, 1113, 1249, 1250
    atherosclerosis 1249, 1250
Heat 282
    of combustion 282
    flow of 284
Heat shock proteins (chaperonins) 518, 1635
Heavy meromysin (HMM) 1101
Hedgehog (Hh) protein 1900
Heisenberg uncertainty principle 138

- HeLa strain of human cancer cells
  - 30 S ribosomal subunit 26
- Helical structure
  - polyethylene 44
  - polyfluoroethylene 44
  - proteins 68 – 71
- Helical wheel 69*s*, 70
- Helicases 1544, 1550, 1551
- Helix (helices). *See also* Superhelix
  - 3_{10} 69, 70
  - amphipathic 79, 401
  - hydrogen bonds in 68
  - one-start 334
  - Pi 69
  - pitch of 68
  - primary 334, 337
  - properties of 69
  - radial projection 334
  - in serum albumin 71
  - stacking of 70
  - structures of 334
  - torsion angles of 68 – 71
- α Helix (helices) 68*s*, 69, 78
  - as helical wheel 69*s*, 70
- β-Helix (helices)
  - parallel 65
- Helix bundles
  - four helix 71, 343, 841, 1752
- Helix-turn-helix motif 76, 1606
- Helper virus 1498
- Hemagglutinin 186, 1650
- Hematoside (G_{M3}) accumulation 1171
- Heme 68, 353, 514, 844, 850, 851, 1404*s*
  - covalent linkage to proteins 719
  - degradation, scheme 1404
  - imidazole group in 851
  - in tryptophan dioxygenase 1059
- Heme *a* 844, 845*s*
- Heme *c* 845, 845*s*
- Heme *d* 845, 845*s*, 1055
- Heme *o* 845*s*
- Heme oxygenases 1404
- Heme-oxygen complexes 851
- Heme proteins **843 – 845**, 1058
  - reactions with hydrogen peroxide **850 – 857**
  - reactions with oxygen **850 – 857**
- Heme-sensitive eIF2 α kinase 1701
- Hemerythrin 362, 852, 862
  - active site, 863*s*
- Hemiacetals 161, 162, 677
  - formation of 526
  - glycosidic linkages 780
- Hemicelluloses 177, 602, 1149
- Hemiketals 677
  - formation of 526
- Hemimercaptals 677
  - formation of 526
- Hemin 844
- Hemochromatosis 840
- Hemochrome(s) 844
- Hemocyanins 885
- Hemoflagellates 19
- Hemoglobin(s) 17, **353 – 359**, 844
  - abnormal 354, 359 – 362
  - 2,3-bisphosphoglycerate, effect 357, 358
  - carbamino groups in 358
  - comparative biochemistry 362
  - competing ligands 358
  - fetal 362, 1901
  - fingerprints of 359
  - folding pattern 354
  - genes (globin) 1539, 1540, 1901, 1902
  - oxygen binding cooperativity 851
  - in oxygen transport 354
  - oxygenation curves 355
  - pK_a values of 357
  - salt bridges 357
  - saturation curve 355
  - structural changes during oxygenation **354 – 357**
  - subunits 353
  - thermodynamic function for oxygen ation, table 355
- Hemoglobin mutants 354, 359 – 362
  - F 362
  - H 354
  - M 359
  - S 359 – 361
- Hemopexin 58, 627, 1850
- Hemophilia A, B 633
- Hemopoietic regulators 1834
- Hemoprotein 1366
- Hemosiderin 842
- Henderson-Hasselbalch equation 97
- Henseleit, K. 1376
- Heparan *N*-sulfatase 1171
- Heparan sulfate 175 – 177, 176*s*
- Heparin 176*s*, 181
  - anticoagulant properties of 177
  - pentasaccharide segment 177*s*
- Heparin cofactor II 177
- Hepatic lipase 1185
- Hepaticae 29
- Hepatitis A virus 247
- Hepatitis B virus 244
- Hepatitis delta virus 247
- HEPES buffer 99
- Heptad in coiled coil 71
- Heptane 1196
- Heptanoyl threonine phosphate (HTP, coenzyme B) 881
- Herbicides 1299
  - glyphosate 687
- Hereditary diseases 2, 16, 1512 – 1517, 1585, 1586, 1907. *See also* Metabolic diseases
  - of blood clotting 633, 634
  - colorectal cancer (HNPCC) 1585
  - of DNA repair 1585, 1586
  - granulomatous disease 1072
  - of hemoglobin 359 – 361
  - of the immune system 1867– 1869
  - muscular dystrophy 1112, 1113
- Hernandulcin 1800, 1800*s*
- Herpesviruses 247
- Hertz 1274
- Hesperidin 1215, 1215*s*
- Heterochromatin 1535, 1625, 1895
- Heterocysts 10, 1366
- Heteroduplex DNA 256, 1489, 1568
  - electron micrograph 1489
- Heterolactic fermentation 972
- Heterologous bonding interaction 333, 334
- Heterologous interactions 342
- Heterologous square 352*s*
- Heteroplasmy 1024
- Heteropolysaccharide(s)
  - of animal body **175 – 177**
  - definition of 169
  - of plants **177– 179**
- Heterotropic interactions 476
- Hexagonal phase
  - of alkanes 394
  - of lipid bilayer 394, 395
- Hexaphosphate (phytic acid) 1132
- Hexokinase 467, 518, 654, 656
  - in regulation 535
- β-Hexosaminidase A and B 1171
- Hexose 162
- Hexose monophosphate shunt. *See* Pentose phosphate pathway
- Hexulose 6-phosphate 986*s*
- High density lipoproteins (HDL) 1181, 1248
  - composition of, table 1183
- High mobility group (HMG) proteins 1535
- High performance liquid chromatography (HPLC) 103
- High potential iron proteins (HIPIP) 857
- Hill, A. V. 960
- Hill plot 331, 332
- Hippocampus 1767, 1799, 1800
- Hippuric acid 513*s*, 1141, 1399, 1399*s*
- Hirudin 634
- Histamine 745, 1450, 1794
- Histidine (His, H) 53*s*
  - binding constants to metal ions 311
  - catabolism of 810, 1449
  - diethylpyrocarbonate, reaction with 127
  - metabolism of **1443 – 1450**
  - pK_a value of 487
  - synthesis regulation 1449
  - titration curve 97
- Histidine ammonia-lyase 755
- Histidine-binding protein 419*s*
- Histidine decarboxylase
  - active site 755
  - mechanism 754
  - proenzyme 754
- Histidine operon 1449
- Histidinohydroxymerodesmosine 434, 434*s*
- Histone(s) 204, **1531– 1535**
  - acetyltransferases (HATs) 1626
  - classes of 1531
  - complex with DNA 221
  - deacetylases 1626
  - H1 1533
  - H2A, ubiquitination of 525
  - linker 1533
  - micromodification 1531
- Histone deacetylase 1542
- Histone fold 1531
- HIV (human immunodeficiency virus) 248, 1651, 1868
  - genetic map 1652
  - replication cycle, figure 1656
  - reverse transcriptase, ribbon drawing 1546
- HIV-1 protease 624*s*
  - catalytic mechanism 625
- HMG proteins 1893
  - as transcription factors 1634, 1635
- Hofmeister series 311
- Holley, Robert W. 84
- Holliday junction 229*s*, 1506, 1572
  - branch migration in 229
  - in DNA **228 – 230**
  - immobile 230
  - processing 1567
- Holliday recombination intermediate 1565
- Holothurin A 1265*s*, 1266
- Homeodomain 1636, 1900
  - proteins 1636
- Homeostasis 1880, 1882
- Homeostatic mechanisms for metal ions 307

Homeotic genes 1636, 1900
Homoaconitase 688
Homocitrate 700*s*, 1361
Homocysteic acid 1793
L-Homocysteine 746, 746*s*, 778*s*, 1798
    metabolism of **1388 – 1391**
Homocysteine lactone 1696*s*
Homogenization 98
Homogentisate 1428, 1429*s*, 1430, 1431*s*
Homolactic fermentation 966, 980
Homology, definition of 59
Homopolysaccharide(s) 175
    definition of 169
L-Homoserine 1384*s*
Homoserine kinase 1383
Homotropic interactions 476
Homovanillate 1790*s*, 1810
Hoogsteen base pair 207
    reversed 208
Hookworms 24
Hopanoid 1244
    lipids 1366
Hopene 1245*s*
Hopene synthase 1244
Hopkins, Frederick G. 83
Hormonal chemoprevention of cancer 1263
Hormone(s). *See also* Specific hormones 1517, 1741– 1762
    biosynthetic pathways 1447
    brain 1749
    of crustaceans 1760
    feedback loops 1742
    gastrointestinal 1749
        table 1751
    hypothalamus **1746 – 1748**
        table 1747
    of insects 1760
    invertebrate, table 1759
    mammalian non-peptide **1756 – 1758**
    mammalian peptide **1750 – 1753**
    microbial, table 1759
    non-peptide of vertebrates, table 1746
    non-vertebrate 1758 – 1760
    pancreatic 1748, 1749
        table 1751
    peptide of vertebrates, table 1744
    of pituitary gland **1743 – 1748**, 1795
    plant 1760 – 1762
    protein of vertebrates, table 1744
    receptor proteins 1741
    receptors for 479, 536, **553 – 563**
    receptors, nuclear 1635
    regulation of enzymes 553 – 563
    release-inhibiting factors 1746
    releasing 1746
    second messengers 1743
    secretion 1762
        constitutive pathway 1762
        regulated pathway 1762
    sequences, table 1751
    signaling cascades 1743
    steroid **1252 – 1265**
    structures 1447
    vertebrate 1743 – 1758
        tables 1744, 1746
Hormone-sensitive lipase 635, 1185
    effect of insulin 571
Horseradish peroxidase 852, 856
Horseshoe crabs 24
Housing silk 67
HPLC. *See* High performance liquid chromatography
HSQC 144
Huber, Robert 84
Human artificial chromosomes 1562
Human body, elements in 31
Human dwarfism 1755
Human genome 12, 201
Human glucose transporters 416
    figure 416
Human immunodeficiency virus. *See* HIV
Humic acid 1443
Hunter syndrome (MPS II) 1169, 1171
Huntingtin 1812
Huntington disease (HD) 1516, 1789, 1812
Hurler syndrome (MPS I) 1169, 1171
Hutchinson–Gilford progeria 1907
Huxley, H. E. 1096
Hyaluronan 170, 175 – 177, 176*s*, 1130, 1884
    hydrogen-bonding 176
    properties of 175
    repeating disaccharide units of 176
Hyaluronan synthases 1152
Hyaluronate lyase 686
Hyaluronic acid 15, 182. *See also* Hyaluronan
Hyaluronidase. *See* Hyaluronate lyase
Hybrid duplexes 256
Hybridomas 1841
Hycanthone 223*s*, 1584
Hydra 23, 24
    development of 1892
    interstitial stem cells 1892
Hydration
    of alkenes, mechanism 683
    effects on biochemical equilibria 50
    of ions 50, 311
    of nucleic acid bases 213
    of polar molecules 50
Hydration layer 51
Hydrazine 738*s*, 1052, 1364
Hydride ion
    shift of 527
    transfer of 765
Hydride-transfer mechanism 789
Hydrocarbon(s) 380 – 382
    chains, synthesis 990
    formation of 1196
    insect 1196
    oxidation of saturated 942
Hydrogen (H_2), *See also* Hydrogenase
    formation by bacteria 300
    photosynthetic formation of 1324
Hydrogen atoms, direct transfer of 768
Hydrogen bond(s) **47– 49**, 69. *See also* Hydrogen bonding
    chains in proteins 56, 64
    in collagen 72
    competition for 82
    in hyaluronan 176
    in ion pairs 75
    lengths of 48, 493, 494
    low-barrier 494
    in polypeptides 56, 61
    in proteins 75
    strengths of 49, 494
    strong, in papain 619
Hydrogen-bonded network
    in carbonic anhydrase 678
    in proteins 75, 1838
Hydrogen bonding **491– 494**. *See also* Hydrogen bond
    chemical shift, effect on 139
Hydrogen dehydrogenase 1051
Hydrogen-oxidizing bacteria 1051
Hydrogen peroxide
    tissue damage 1074
    toxicity of 1072
Hydrogen transfer reactions, table 766
Hydrogenase(s) 861, 878, 879, 879*s*, 969
    iron-only 878
    in methane bacteria 1057
    Ni containing 879
Hydrogenosomes 15, 970
Hydrolases 589
    definition of 498
Hydrolysis 593, 608
    of biopolymers 116 – 118
    of carbohydrates 189, 599 – 609
    of lipids 507, 634 – 637
    Gibbs energies of, table 294
    of nucleic acids 249 – 257, 647– 653
    of proteins 116 – 118, 609 – 631
Hydropathy index 78, 79
Hydroperoxide 1060
Hydroperoxide anion 1059
Hydroperoxy-eicosanoic acid (HPETE)
    biosynthesis of, by 5-lipoxygenase
        pathway 1209
Hydrophobic effect 50
Hydrophobic interaction **50**, **51**
Hydrophobic interaction chromatography 103
Hydrophobicity 51
Hydroquinone(s) 815 – 822, 1438
Hydroxamate 838
Hydroxide ion, binding constants to metal ions 311
Hydroxocobalamin 869
2-Hydroxy acid, formation in plants 942
D-Hydroxy acids, accumulation in green leaves 942
β-Hydroxyacyl-CoA 940
β-Hydroxyacyl-CoA dehydrogenase 940
Hydroxyanthranilate 1445*s*
erythro-β-Hydroxyaspartate
    reaction with aminotransferase 750
para-Hydroxybenzoate 1059, 1422*s*, 1424
4-Hydroxybenzoate hydroxylase 1059
D-3-Hydroxybutyrate 946*s*
γ-Hydroxybutyrate 1798
β-Hydroxybutyrate-acetoacetate couple 981
4-Hydroxycinnamate 1439*s*
β-Hydroxydecanoyl thioester dehydratase 1191
Hydroxydopamine 1790*s*, 1798, 1811
Hydroxy-eicosanoic acids (HETEs)
    biosynthesis of by 5-lipoxygenase
        pathway 1209
Hydroxyguanidine 1379
Hydroxyheme 1404*s*
α-Hydroxyisovalerate 1395*s*
3-Hydroxykynurenine 1333, 1445*s*, 1798
Hydroxylamine 251
    oxidation of 1052
    reactivity of 590
Hydroxylapatite 103, 441*s*
Hydroxylases 1057– 1072. *See also* Monooxygenases
    copper-containing 1064
Hydroxylation 942, 1057– 1072
    with cytochrome P450 1065
Hydroxylation-induced migration 1062
Hydroxyl radicals 856
    reaction 1073
    toxicity of 1072
Hydroxylysine 432*s*

5-Hydroxymethylcytosine 234, 810
Hydroxymethylfurfural 167*s*
3-Hydroxy-3-methylglutaryl-CoA (HMG-CoA) 699, 701*s*, 946, 992
    lyase 699
    reductase 776
    synthesis of 699
5-Hydroxymethyluracil 234
3-D-Hydroxymyristic acid 429
4-Hydroxynonenal 682, 1204, 1204*s*
2-Hydroxy-3-oxoadipate 1399, 1399*s*
2-Hydroxy-6-oxohepta-2,4-dienoate 1437*s*
Hydroxyphenylpyruvate 1422*s*, 1428, 1429*s*, 1431*s*
Hydroxyprogesterone 1256*s*
Hydroxyproline 52*s*, 72
4-Hydroxy-L-proline 432*s*
β-Hydroxypropionate 947
2(α)-Hydroxypyridine 491*s*
    as catalyst 490, 491
    tautomerism 45
Hydroxypyridine crosslink in collagen 434
Hydroxypyrrolidone 1382
Hydroxypyruvate 986, 1397
    reduction to PEP 986
D-10-Hydroxystearate from hydration of oleic acid 688
Hygromycin B 1691
Hyperfine structure of EPR signal 398
Hyperglycinemia 949
Hypericin 1336, 1336*s*
Hypermodified base 1688
Hypermutation, somatic 1861, 1862
Hyperoxaluria, type one 1398
Hyperpolarizability 70
Hyperpolarization 1787, 1788
Hyperprolinemia 1374
Hypersensitive sites 1533
Hyperthyroidism 1432
Hyperuricemia (gout) 1459
Hyphae 20
Hypochromic effect 209, 255, 1285
Hypocretins 1808
Hypoglycemia 949
Hypoglycin A 1394, 1394*s*
Hypophysis 1765
Hypotaurine 1408
Hypothalamus 1742, 1765, 1767
Hypoxanthine (Hyp) 1458*s*
    tautomerization of 203
Hypoxanthine-guanine phosphoribosyltransferase (HGPRT) 1457
Hypusine 1382, 1386, 1386*s* 1702

I

I-bands (muscle) 1096
Ibuprofen 1211*s*
ICAM-1, diagram 408
Ice
    hydrogen bonds in 49
    melting of 284
    reaction rates in 492
    structure of 49
Ice-nucleation proteins 191
I-cell disease (mucolipidosis II) 1170
Ichthyoses 439
Icosadeltahedra 346
Icosahedral viruses 345
Icosahedron 344
Identity-determinant set 1694
Idiopathic pentosuria 1135
Idiotopes 1858
Idiotype set 1858
Idose (Ido) 163*s*
Idoxuridine 1655*s*
Iduronate, iduronic acid 176
    conformation 166
Iduronate 2-sulfate sulfatase 1171
α-L-Iduronidase 1171
IgA molecule. *See* Immunoglobulins, IgG
IgG molecule. *See* Immunoglobulin, IgG
IgM molecule. *See* Immunoglobulin, IgG
Image reconstruction 130
Imidazole
    binding constants to metal ions 311
    pK_a value of 99, 487
Imidazole acetic acid 1450
Imidazole group(s) 46*s*, 53
    hydrogen bonding 48
    numbering of ring atoms 46
    as a polarizable base 590
    in proteins 312
    protonation of 46
    tautomerism of 46
Imidazoleglycerol phosphate 1449*s*
Imidazolinone 1393
    herbicide 1392*s*
Imidazolium cation 46, 46*s*
Imidazolone propionate 1450*s*
Imine(s), 679. *See also* Schiff bases
    in enzymes 699, 740 – 746, 750
Imipenem 1164*s*
Imipramine 1809*s*, 1809
Immune response **1850 – 1859**
    acute phase 1440
    coreceptors 1850
Immune system **1831– 1869**
    adaptive (acquired) 1831, 1832
    cancers of 1868
    cells of, table 1834
    cellular 1831
    diagram 1832
    disorders of **1862 – 1869**
    effectors 1834
    evading 1866, 1867
    humoral 1831
    innate 1831, 1832
    internal mucous membranes 1831
    location of **1831– 1835**
    organization of **1831– 1835**
        table 1833
    proteins of **1840 – 1850**
    self-tolerance 1858, 1859
    in skin 1831
    triggering of 1834, 1835
Immunoassays 1848
Immunodeficiencies 1867, 1868
Immunoelectron microscopy 1680
Immunoglobulin(s) 58, 74, 1834, **1835 – 1840**
    antibody binding responses 1839, 1840
    antigenicity 1839
    complementarity-determining regions (CDRs) 1838
    domain 65, 367
    Fab fragments 1836
    Fc fragments 1836
    Fc receptor 1839
    gene organization 1860
    gene processing 1861
    genes 1539, 1859, 1860
    heavy chain gene structure 1861
    heavy chains 1836
    hypervariable regions 1838
    IgA molecule 1836, 1838
    IgE molecules 1863
    IgG molecule 1836
        ribbon diagrams 1837
    IgG1fragment, structure 1838, 1840
    IgM molecule 1836
        cryo AFM micrograph of 1836
        structure 1836, 1837
    isotype switching 1862
    light chains 1836
    structures **1835 – 1839**
        three-dimensional 1838
    superfamily 187
    synthesis of 1840
    variable regions 1836
Immunoglobulin genes 1859, 1860
    D (diversity) genes 1860
    J (joining) genes 1860
    V genes 1860
Immunologic memory 1833
Immunological networks 1858
    scheme 1858
Immunological synapse 1835
Immunosuppressant 488, 805, 1865
Immunotoxins 1686
Imprinting 1542, 1880
Incoherent light 1275
Indigo 1421
Indigotins 1448
Indolamine 2,3-dioxygenase 1443
Indole 1426*s*, 1445*s*
    resonance energy 299
Indole-3-acetate 734
Indole-3-acetic acid 1446*s*, 1761. *See also* Auxins
Indole-3-acetonitrile 1446*s*
Indole alkaloids
    biosynthetic pathways 1447
    structures 1447
Indole-3-glycerol phosphate 1426*s*, 1446*s*
Indole-5,6-quinone 1433*s*, 1434
Indoleglycerol phosphate 1427
Indoxyl 1444, 1445*s*
    sulfate esters of 1448
Induced fit 481
    model 351
    theory 349
Inducer(s) 539
    of gene expression 539
Inducible enzyme 538
Induction of enzyme synthesis 538, 539
Inflammation 384, 1211
Influenza viruses 1866
    replication of 1650
Infrared linear dichroism 1277
Infrared microspectrophotometry 1298
Infrared (IR) spectra
    of amides 1278
    of cytosine 1293
    of 1-methyluracil 1278
Inhibins 1747
Inhibition
    classical noncompetitive 474
    of enzymes 471– 478
    of enzyme synthesis 536
    noncompetitive 473 – 475
    uncompetitive 475
Inhibition constant 472
Inhibitors
    competitive 471

  enzyme-activated 478, 864
  of respiration, table 1021
  in study of enzyme mechanisms 475
  transition state 484
Initiation codon(s) 236, 1476, 1672, 1698
Initiation factors 1673
Initiation proteins 1551
Initiator proteins 1552, 1553
Inorganic compounds in respiration 1051–1054
Inorganic deposits by cells 15
Inorganic fermentations 1057
Inorganic oxidant
  substitution for oxygen 1051
Inorganic reactions
  as energy source **1050 – 1057**
Inosine (Ino) 203, 1458*s*
Inosine 5'-phosphate (IMP) 1454
  conversion to adenine, scheme 1456
Inosinic acid 203
Inositol 1131, 1132
myo-Inositol 387*s*, 1131, 1132
  conversion to streptidine 1139
myo-Inositol-1-phosphate 1132*s*
Inositol 1,4,5-trisphosphate (InsP$_3$) 542, 563
  receptors 564
Inositol 1,3,4,5-tetrakisphosphate (InsP$_4$) 564
Inositol 1,3,4,5,6-pentakisphosphate (InsP$_5$) 564
Inositol hexakisphosphate (InsP$_6$) 564
Inositol polyphosphates
  scheme 565
  in signaling 563 – 566, 1743
Insecta 24
Insecticides **636**, 1788
Insecticyanin 1186*s*, 1305
Insertion mechanism
  in carbohydrate chain growth 1152
Instantaneous velocity 456
Instincts 1767
Insulator elements in DNA 1881
Insulin 346*s*, 544, **567– 571**, 997– 999, 1742
  effects on acetyl-CoA carboxylase 568
  effects on ATP citrate-lyase 568
  effects on enzymes, table 998
  effects on GLUT4 568
  effects on ribosomal subunit S6 568
  glucose uptake 570
  half-life 571
  hexameric, 347*s*
  human 1004
  metabolic effects of 568
  phosphorylation 571
  response element 578
  second messenger 569 – 571
Insulinlike growth factors (IGF) 1748, 1753
Insulin-like signaling system (IGF) 1907
Insulin receptor 568, 569
  autophosphorylation **569**
  down-regulation 571
  mutant forms 569
  protein kinase of 544
  structure (schematic) 569
  substrates 559
  substrate-1 (IRS-1) 570
  tyrosine kinase domain 569, 570*s*
Integral proteins of membrane 391, 401, 402, 1723, 1724
Integrases 1570, 1651
Integrase (Int) protein 1572
Integrase XerC/XerD
  action of, figure 1571
Integrated rate equations 457– 461
Integrin(s) 187, 405, 406, 1883
Inteins 1716, 1717
Intelligence 1804, 1805
Interaction constant 352
Intercalation in DNA **222**, **223**
Interferon **1847– 1849**
  binding responses, diagram 1847
Interferon-α 1756
Interferon-β 1902
Interferon-γ 1849
Interleukin(s) 1755, 1756, 1847, 1849
Interleukin-1 1847
Interleukin-2 (T-cell growth factor) 1756
Interleukin-4 1849
Interleukin-1β-converting enzyme 619
Intermediate density lipoproteins (IDL) 1181, 1186
  composition of, table 1183
Intermembrane space
  of mitochondria 1014
Internal conversion 1291
Internal energy 282
Internal equilibrium constants 640
Internal image set 1858
Internal ribosome entry sites 1702
Internal signal sequences 1724
International system of units 2
International unit of enzyme activity 457
Interneurons 1766
Interphotoreceptor retinoid-binding protein 1332
Intersystem crossing 1289, 1290
Intervening sequences. *See* Intron
Intosome 1570, 1573, 1574, 1581, 1582, 1583
Intradiol aromatic dioxygenases
  intermediates, scheme 1058
Intragenic complementation 1480
Intrinsic barrier 493, 691
Intrinsic factor in vitamin B$_{12}$ uptake 869
  pernicious anemia 869
Introns 202, 1540
  group I 1643, 1645*s*
Inulin 170, 175
  biosynthesis of 1130
Inversion of configuration
  as criterion of mechanism 590, 591
  by glycosyltransferases 593 – 607
Invertase (sucrase) 1140
Invertebrate vision 1332
Involucrin 439, 1202
Iodacetamide 115
Iodide-starch complex 174*s*
Iodide transporter 1432
Iodine 1430
Iodoacetate 115
Iodothyronine deiodinase 824, 1430
Ion(s)
  hydration of 50
  transport of **420 – 425**
    electrogenic 421
Ion channels 412 – 414, **420 – 425**, **1769 – 1773**
Ion exchange chromatography 103, 104
Ion exchange resins 104
Ion pairs 55
  hydrogen bonded 75
  in proteins 75
Ion pumps 422 – 425
Ionic atmosphere of membranes 400
Ionic radius
  effect on metal binding 311
  for ions, table 310
Ionic strength 288, 297
Ionophores 414
Ipsenol 1232, 1233*s*
Iridial 1233*s*
Iron **837– 866**
  binding of O$_2$ 850
  in hemoglobin 850
  proximal histidine in heme 851
  storage of 841– 843
  uptake by cells 837– 841
    non-transferrin pathway 840, 841
    transferrin pathway 840, 841
Iron protein
  in nitrogenase 1361
Iron regulatory factors 688, 862
Iron-sulfur cluster(s) 514, 857, 857*s*, 1056
  3Fe-4S 859
  4Fe-4S 857
  in flavoproteins 861
  membrane proton pumps 860
  Mossbauer X-ray absorption spectra 860
  oxidation states 857
  properties of 859, 860
  radical generator 862
  reduction of 859
  reduction and oxidation 859 – 862, 1026 – 1030
  single-electron transfer 861
Iron–sulfur proteins 513, 514, 686 – 689, **857– 862**
  adenosine 5'-phosphosulfate reductase 1056
  electron paramagnetic resonance spectrum 860
  functions of 686 – 689, 860 – 862, 1019 – 1023, 1026 – 1030
Iron superoxide dismutase 866*s*
Iron-tyrosinate proteins 1058
Isethionic acid 1408
Isoacceptor tRNA 1694
L-Isoaspartyl/D-aspartyl methyltransferase 594
Isoaspartyl group 84
Isobutyraldehyde, tautomerism 45
Isobutyryl-CoA mutase 872
Isochorismate 1425, 1426*s*, 1429*s*
Isocitrate 516*s*, 686*s*, 704*s*, 952
  cleavage to glyoxylate 988
  oxidation to 2-oxoglutarate 705
Isocitrate dehydrogenase
  phosphorylation of 545
  stereochemistry 768
Isocitrate lyase 704, 988
Isocoumarins 623
Isodesmosine 436, 436*s*
Isoelectric focusing 108, 111
Isoelectric point 101, **106**
  of proteins 83
Isoenzymes (isozymes) 536, 538
Isoeugenol (cloves) 1439*s*
Isoionic point **106**
Isolation of compounds 98 – 108
Isoleucine (Ile, I) 52*s*, 539, 1383, 1391
  biosynthesis 540
    scheme 1392
  branched fatty acids from 381
  catabolism of, scheme 1396
  configuration 43
Isologous interactions 337– 353
  in oligomers 349 – 353
  square 352*s*

Isomaltose 1141
Iso-mechanisms 466
Isomer(s) **41–43**
Isomerases 498, 530, **692–697**
    aldose-ketose interconversions 692
    cis-trans, glutathione-dependent 690
    hydride shift mechanism 695
    3-oxosteroid 696, 697
    peptidyl prolyl *cis-trans* 488, 1721
    ring opening by 693
    triose phosphate 693–697
    xylose 695
Isomerization reactions 526, 527, 530, 690, 712
    of isopentenyl diphosphate 712
    stereochemistry 874
    vitamin B_{12}-dependent 873
        table 871
Isomorphous replacement method 133
Isoniazid 1194*s*
Isonicotinic acyl-NADH 1194*s*
Isonicotinyl hydrazide 738
Isopenicillin N 1064*s*
Isopenicillin N synthase 1064
Isopentenyl diphosphate (prenyl pyrophosphate) 527, 689, 690*s*, 712*s*, 1227, 1229*s*, 1230
Isopentenyladenosine 234*s*
Isopeptidases 525
Isopeptide linkage
    in proteins 80
    to ubiquitin 524
Isopotential groups, in mitochondria 1036
Isoprene formation, isomerization 1230
Isoprenoid alcohol phosphokinase 409
Isoprenoid compounds 389, 390, 1227–1266. *See also* Polyprenyl compounds
    biosynthesis 689, 690, 700, 1227–1240, 1244–1248, 1251–1260
    glyceraldehyde:pyruvate pathway 1229, 1230
Isoprenoid lipids **389**, 390
Isoprenoid quinones **818–823**
    structures of 819
Isoprenylation. *See* Prenylation
Isopropylmalate 700*s*, 1393
Isopropylmalate isomerase 689
Isopropyl-β-D-thiogalactoside (IPTG) 1604
Isoprostanes 1212
Isoproterenol 553, 553*s*
Isoquinoline ring 1443
Isorhodopsin 1330
Isoschizomers 1491
Isospartyl residues in proteins 594
Isotachophoresis 108
Isotope(s)
    properties of 110, 111
    radioactive 110
    stable 110
Isotope-edited FTIR 1278
Isotope effect(s)
    for alcohol dehydrogenases 771
    in displacement reactions
        kinetic 111, 486, 592, 593
Isotope exchange at equilibrium 467, 468
Isotopic labels (tracers) 110, 679
Isotopic tracers
    tricarboxylic acid cycle 954, 956
Isotopomer(s) 111
Isotopomer analysis 111
    of metabolism **954–956**, 1000, 1001
Isovaleric acidemia 1394
Ivermectin 1788*s*, 1789

J

Jacob, Francois 84, 1475, 1603
Jamaican vomiting sickness 1394
Janus kinases (JAKs) 1847
Jasmonate 1149
Jasmonic acid 1212, 1212*s*, 1762
Jaundice 1297, 1404
Joule 282
    definition of 2, 283
Jun, cJun, *jun* 70, 241, 572, 576, 577, 579, 1632, 1633
Juvebione 1234, 1235*s*
Juvenile hormone 1235*s*, 1760
    of insects 1234

K

K absorption edge 1288
Kainate 1786, 1786*s*
Kallikrein 1752
Kanamycins 1139, 1690
Kar3 1107
Karplus equation 139
Karyotype 1472, 1512
Kasugamycins 1691
Katal 457
K^+ channel 413*s*
KDEL receptor 521, 554
KDO (ketodeoxyoctonate) 180*s*, 1137
Kelps 22
Kendrew, John C. 84
Keratan sulfate **175–177**, 176*s*
    disulfide bridges 115
    N-linked 185
    terminal units 181
Keratin 70, 369
Keratinocytes 439, 1902
Keratoglycan 1154
Keshan disease 822
Kethoxal 253
Ketimine 121, 744*s*. *See also* Schiff base
    from pyridoxal phosphate 742
    as electron acceptor 746, 747
α-Ketoacid. *See* α-Oxoacid
Ketoamine 434*s*
Ketodeoxyoctonate. *See* KDO
Ketogenic amino acids 1397
Ketone
    acidity of 46
    bodies **945–947**
Ketosis 946, 949, 1003
Kettin 1099
Khorana, H. Gobind 84
Kidney cells, alkaline phosphatase in 645
Kidney tubules 28
Kilocalorie 282
Kinases 638, **654–657**
    cyclin-dependent 580, 1502
    metal ions in 654
Kinesin(s) 370, **1107–1111**
    ATPase cycles of 1108, 1109
    heads, motion of 1109
    human, ribbon drawing 1107
    transport by **1119**–**1121**
    tubulin complex 1110
Kinetic(s)
    high enzyme concentrations 461
    of enzyme action **455–497**
    of rapid reactions 468, 469
    of reaction centers 1312
Kinetic constants 455
Kinetic energy of molecules 1276
Kinetic equations for complex mechanisms 466, 467
Kinetic isotope effect (KIE) 76, 111, 486, **592**, **593**, 600, 789
Kinetic parameters 459
Kinetic theory of gases 462, 463
Kinetisome 15
Kinetochores 1502, 1503
Kinetoplast 1541
Kininogens 1752
Kinins 1752
Kirromycin 1703
Klenow fragment 1491
    ribbon drawing 1546
Klug, Aaron 84
Knobs-into-holes bonding 71, 334
    in coiled coil 71
Knockout mice 1501, 1805
Knots
    in peptide chains 74
    in DNA 219
Kojibiose 431
Kornberg, Arthur 84
Koshland
    induced fit theory 349
    model 352
Krebs, Hans A. 83, 980, 1376
Krebs Cycle 951. *See also* Citric acid cycle
Kringles 367, 368*s*
K systems 475
Kuru 248, 1718
Kynurenic acid 1444, 1445*s*
Kynurenine 747*s*, 1444, 1445*s*
    conversion to alanine 746

L

Laccase 887
lac control region 1606, 1607
    nucleotide sequence, figure 1605
Lac operator 1606
Lac operon **1603–1607**
    inducer 1604
    schematic representation 1604
Lac repressor 240
    structure 1606
Lachrimatory factor 1408
Lactacystin 620, 620*s*
α-Lactalbumin 1141
β-Lactamases 1165
Lactarazulene 1235*s*
Lactate 506*s*
    binding constants to metal ions 311
    conversion to glycogen 989
    as fermentation product 966
    fermentation by propionic acid bacteria 970
    formation from glucose 966
Lactate dehydrogenase 342, 467, 768, 771
    active site of 769
    inhibition by pyruvate 780
    isoenzymes of 538

Lactate monooxygenase 799, 1059
Lactic acid 508, 510
  pK_a value of 293
Lactic acidemia 1002, 1394
Lactobacillic acid 381*s*
Lactoferrin 840, 840*s*
Lactogen, placental 1745
β-Lactoglobulin 1241
Lactoperoxidase 853, 856
Lactose (Lac) 167
  α-lactose 168*s*
  β-lactose 168
  biosynthesis of 1130, 1141
  sweetness of 168
Lactose (Lac) permease 417, 1475
Lactose synthase 1141
Lactosyl ceramide sulfate 389*s*
Lactosyl ceramidosis 1171
Lactyl-CoA 948
O-Lactyl-GlcNAc. *See* Muramic acid
Lactylthiamin 733*s*
Lag phase of growth curve 470
Lambda repressor 1622, 1623
Lambda transducing phage 1486
Lamellar (smectic) phase 392
Laminin 71, 409, 437, 1112
Lamins 369, 1536
Lampbrush chromosomes 1534
  photomicrograph 1535
Landsteiner, blood groups 184
Langerhans cells 1834
Langmuir – Adam film balance 394
Langmuir – Blodgett layers 392, 394
Langmuir trough 394
Lanosterol 1244, 1245*s*, 1246*s*
  conversion to cholesterol 1245, 1246
  demethylation, scheme 1246
Lanosterol 14α-demethylase 1245
Lantibiotics 1722
Large calorie, numerical value of 283
Lariat structure 1644
Lariet intron 1647
Lateral phase separation 395
Lathyrism 438
α-Latrotoxin 1776
Lavoisier 281
LDL receptors 1185, 1250
Leader peptidase 620, 1722
Leadzyme 649
Learning
  chemistry of 1801– 1814
  molecular coding 1804
Leaving group
  definition of 589
  effect on reaction rates 590
Lebers hereditary optical neuropathy 1024
Lecithin. *See* Phosphatidylcholine
Lecithin:cholesterol acyltransferase 1247
Lecithin:retinol acyltransferase 1332
Lectin(s) 64*s*, 186 – 188, 1154, 1159
  domains in proteins 186
Lederberg, Joshua 83
Leghemoglobin 362, 1360
Legionnaire's disease 7
Legume(s) 1360
  nodules 1365
Leloir, L. F. 1129
Lens, proteins of 1333
Lentivirus 1498, 1655, 1656
Leprosy 7
Leptin 1005, 1750
Lesch – Nyhan syndrome 1457, 1459, 1518
Leucine (Leu, L) 52*s*, 1391
  biosynthesis of 527, 712
    scheme 1392
  branched fatty acids from 381
  catabolism of, scheme 1396
L-Leucine 2,3-aminomutase 875
Leucine aminopeptidase 627
Leucine zipper 69, 71, 241, 576, 1633
Leucodopachrome 1433*s*, 1434
Leucopine 1498
Leucopterin 803, 804*s*
Leucovorin 810
Leu-enkephalin 1752
Leukemias 573
Leukocyte(s) 26, 1833, 1834. *See also* B Cells, T Cells
  adhesion 188
  differentiation antigens 1850
  lymphocytes 1833, 1834
Leukocyte locus A (HLA) 1852
Leukosulfakinin 1749
Leukotriene(s) 1209*s*, 1757, 1863
  biosynthesis of by 5-lipoxygenase pathway 1209
Leukotriene B 1122
Leupeptin 615, 622, 622*s*
Lewis acid, Zn^{2+} 680
Lewis blood group 184, 185
Lewis x antigen 185, 188
Lewy bodies 1812
Lichens 23
Life, origin 9
Ligase 498, 1491
Light **1273 – 1346**
  absorption of 122, 123, 205, 1275 – 1288. *See also* Absorption spectra
  absorption, near infrared 1276
  circularly polarized 1275
  coherent 1275
  electronic spectra **1280 – 1286**
  frequency of 1274
  incoherent 1275
  plane polarized 1275, 1283
  properties of **1273 – 1275**
  quantitative measurement 1275, 1276
  quantum, energy of 1274
Light compensation point 1321
Light energy, exciton transfer 1308
Light-harvesting complexes 1305 – 1310
  chlorosome model 1308
  LH1, LH2 structures 1307
  LHCII structure 1309
Light-induced transformation 1329, 1330
Light meromyosin (LMM) 1101
Lignans 1441
Lignins 1150, 1434, 1440 – 1442, 1441*s*
Lignols 1440
Lily, genome 12
Limonene 1233*s*
Linalyl diphosphate 1234*s*
Lineweaver-Burk plot 460
Link protein (in proteoglycan) 181
Linkage between genes 18, 1477, 1486, 1504
Linkage map
  bacteriophage T4 1481
  E. coli 1486
Linoleic acid 381, 1190*s*, 1193
Linolenic acid 1190*s*, 1193, 1212*s*
Linolenoyl-CoA 1188*s*
Lipase(s) **634 – 637**
  hormone-sensitive 635, 1185
  lipoprotein 635
Lipid(s) **379 – 436**, **1181– 1217**
  bilayer 391*s*, 394
    melting temperature 395
  building blocks of **380 – 389**
  caloric value 283
  content in tissues 31
  ether-linked 1201, 1202
  as hormones 1757
  in membranes 392
    physical properties of 394
  metabolism, **1181– 1217**
  movement between cells 1184 – 1186
  neutral 382
  peroxidation of 1204 – 1206
  in signaling 1202
  of skin 1195, 1196
  surface 1195, 1196
Lipid A 429, 430*s*, 1161, 1198
  synthesis route 1163
Lipid-storage diseases 1251
Lipmann, Fritz A. 83, 720
Lipoamide dehydrogenase
  reaction mechanism, scheme 792
Lipoarabinomannans 431
Lipocalins 1184 – 1186, 1799
Lipo-chitooligosaccharides 1365
Lipocortins 316, 571. *See also* Annexins
Lipogenesis 1003
Lipoic acid **795 – 802**, 1393
  absorption maximum 795
  arsenite, reaction with 596
  chemical reactions 795
  functions of 796 – 798
  proteins, attachment to 795*s*
  reduction of 785
  ring strain in 795
Lipooligosaccharides 1761
  signaling by 1149
Lipophorins 1182, 1186
Lipopolysaccharide 429, 1162*s*
  O-antigen of 180
  structures of 430
  synthesis route 1163
Lipoprotein(s) 58
  bacterial 428
  composition of 1181, 1182
    table 1183
  plasma 1181– 1184
    classes of, table 1183
Lipoprotein(a) (Lpa) 1182
Lipoprotein lipase 1184
Lipopyrophospho-glucose 1148
Liposomes 392 – 394, 1498
  NMR spectra 396
Lipotropin(s) 1743 – 1745
Lipovitellins 1182
Lipoxins 1209*s*, 1210, 1757
  biosynthesis of by 5-lipoxygenase pathway 1209, 1210
Lipoxygenase(s) 1059, 1208 – 1210
  plant 1212
5-Lipoxygenase pathway 1863
Liquid crystals 392 – 394
Liquid crystalline phases 392
Listeria monocytogenes, motility 1120
Lithium salts, in treatment of manic-depressive illness 564
Lithostatine 443
Liver, pentose phosphate pathway in 964
Liverworts 29
Lobry de Bruyn–Alberda van Ekenstein transformation 693

Lock and key theory 478
Log normal distribution curve
    fitted to absorption bands 1282
    parameters of 1282
Log phase of growth 470
Lon protease 628
Loricin 439
Lou Gehrig's disease 884
Lovastatin 1217
Low barrier hydrogen bond 614
Low density lipoprotein (LDL) 1181, 1248
    composition of, table 1183
    receptors for 1186
Lowry method 102
LSD (lysergic acid diethylamide) 1798
Luciferase 795, 1341
Lucifer dyes 1294
Luciferins 1341
    structures of, figure 1343
Luciferyl adenylate 1341, 1343*s*
    anion, 1346*s*
Luciferyl sulfate 1343*s*
Luft disease 1024
Lumazine protein 1345
Lumican 1154
Lumichrome 783, 1762
Luminescence resonance energy transfer (LRET) 1111, 1292
Lumirhodopsin 1329
Lungfish, genome 12
Lupinine 1386
Lupus erythematosus 1641
Luteinizing hormone (LH) 1260
Luteinizing hormone releasing factor (LRF) 54*s*
Lutropin (LH) 1745
    receptor 554
Lyases 530
    definition of 498
Lycopene 1237, 1239*s*
    absorption of light 1283
    absorption spectrum of 1284
Lymph node(s)
    drawing 1832
    germinal centers 1834
Lymphocytes 25, 26, 1831
    B cells 1831, 1833
    cytolytic T cells 1833
    development of **1833 – 1834**
    helper T cells 1833, 1840
    immunoregulatory T cells 1833
    memory cells 1840
    receptors, drawing 1858
    suppressor T cells 1833
    T cells 1831, 1833
    T cells, functions **1851– 1852**
    T cells, micrograph 1851
    T-cells receptors 1835, 1856, 1857, 1860
    T-cells response 1834, 1835
Lymphoid cells 1901
Lymphoid tissues 1833
Lymphokines 1122, 1755, 1847
Lymphomas 573
Lymphotoxin (TNF-β) 1849
Lynen, Feodor 84, 725
Lyotropic series (Hofmeister series) 311
Lysergic acid 1447*s*
Lysine (Lys, K) 700, 745, 1383, 1680
    biosynthesis 53*s* , 540, 1373, 1385
    catabolism of 1386 – 1388
    formation 1383
    pK_a value of 487
D-Lysine 5,6-aminomutase 874
L-Lysine-2,3-aminomutase 1388
Lysine-bradykinin 1752
Lysine monooxygenase 1059
D-α-Lysine mutase 871, 874
L-β-Lysine mutase 871, 874
Lysine tyrosylquinone (LTQ) 817, 817*s*
Lysogenic cycle
    of lambda phage 1496
Lysogenic state 1483
Lysophosphatidic acid (LPA) 1196
Lysophosphatidylcholine 384
Lysosomal deficiency diseases 1145
    enzyme deficiency diseases 1172
Lysosomal enzymes 1159
Lysosomal storage diseases **1169 – 1172**
    table 1171
Lysosome 10, 13, 98
Lysozyme(s) **599 – 602**, 599*s*, 603
    active site structure 600
    catalytic side chain groups 599
    HSQC NMR spectrum 146
    kinetic isotope effect 600, 601
    neighboring group assistance 601, 602
    substrate distortion 601
    transition state 601
    titration curve 98
Lysyl oxidase 434, 886
Lysyl-tRNA synthetase
    active site of 1697
Lytic cycle
    of lambda phage 1496
α-Lytic protease 610
Lyxose (Lyx) 163*s*

M

Macrodipole of helix 70
Macroglobulin(s) (IgM) 58, 629
    α_2 630, 1850
Macrolides 1164
Macromolecular structures
    self-assembly pathways **362 – 368**
Macromolecules. *See also* Polymers
    association of 325
    binding by 327
    packing of **332 – 348**
    synthesis and turnover of **517– 527**
Macronucleus 1578
Macrophage 26
    micrograph 1835
Macrophage migration-inhibition factor (MIF) 1849
Macroscopic dissociation constants 305 – 307
Macular degeneration 1333, 1902
Mad cow disease 1718 – 1720
Magainins 1843
Magnesium ion(s) 31
    in adenylate system 302 – 304
    complex in kinase action 655
    concentrations in tissues **299**
    effect on equilibria 296
    enzymes dependent upon 299
    intracellular concentration 299
Magnesium protoporphyrin chelatase 1402
Magnetic iron oxide 842
Magnetic permeability 1275
Magnetic resonance imaging (MRI) 1806
Magnetic resonance study
    of phosphorus atom displacement 639 – 642
Magnetite in organisms 842
Magnetoencephalography (MEG) 1807
Maintenance methyltransferase 1541
Maize, genome 12, 1512
Major groove of DNA 201, 213
Major histocompatibility complex (MHC) 1833, 1835, 1851, **1852 – 1855**
    class I molecule 1856
    class II molecule 1856
    genes of 1852, 1853
    HMC restriction 1855, 1856
    peptide complex, drawing 1857
    structure 1854, 1855
Malaria 19
    parasite 1866
L-Malate, 516*s*, 683*s*, 705*s*
    cleavage to acetyl-CoA 703
Malate-aspartate shuttle 1050
    diagram 1051
Malate dehydrogenase 497, 768
Malate synthase 988
    in dicarboxylic acid cycle 959
Malathion 636
Male sex hormone 1254
Maleylacetoacetate 1431*s*
Maleylacetoacetate isomerase 1430
Maleylpyruvate 690*s*, 1436*s*
Malic enzyme(s) 705
    in animal mitochondria 958
    regulatory properties 958
Malignant hyperthermia syndrome 1115
Malolactic enzyme 777
Malonate decarboxylase 723
Malondialdehyde 1205*s*, 1588, 1906
Malonic semialdehyde pathways 947
Malonyl-CoA 947
    biosynthesis of fatty acids 950, 978
Maltase 1146
Maltoporin (LamB) 411
Maltose 167*s*, 606
β-Maltose from β-amylase action 608
Malvidin 1214
Mammary gland(s)
    pentose phosphate pathway in 965
S-Mandelate 1437, 1437*s*
Mandelate dehydrogenase 784
Mandelate racemase 691
    active site 690
Mandibular organ-inhibiting hormone 1760
Man-*P*-dolichol 1155*s*
Manganese **887**, **888**
Manganese ion(s)
    effect on equilibria 297
    in enolase 685
Manic-depressive illness 1810
Mannan(s) 170, 175
    biosynthesis of 1130
Mannich reactions 1443
Mannitol 165, 191
Mannoproteins 1157
    of yeast cell walls 185
Mannose (Man) 161*s*, 163*s*
Mannose-6-phosphate 185
    equilibrium with fructose 6-phosphate 1129
    receptors 1159
Mannose-6-phosphate isomerase 693
β-Mannosidase 1171
Mannosidosis 1171, 1172
MAO inhibitor 1792
MAPK 578

  definition of 577
MAPKK, definition of 577
MAP kinase cascade 576 – 578, 1634
  figure 579
  in insulin signaling 570
Maple syrup urine disease 1394
Marcus theory 771, 848
  of proton transfer 493
Marfan’s syndrome 440
Marijuana 1795, 1797
Markovnikov rule 777
Maroteaux-Lamy syndrome (MPS VI) 1171
Mass action ratio 535
Mass spectrometry
  in diagnosis 1517
  electrospray 114
  MALDI 112 – 114
  in proteomics 1728
  in sequencing 119
Mass spectrum
  of Cu/Zn superoxide dismutase 114
  of myoglobin, apo 115
Mast cells 427
  of tissues 1863
Mastigophora 19
Matrix isolation spectroscopy 1293
Matrix metalloproteases 627
Matrix of mitochondria 14, 1014
Max 1633
Maysin 1215, 1215*s*
McArdle’s syndrome 1145
McClintock, Barbara 84
Mechanocytes 25
Mechanoreceptors 1801
Medusa form 23
Megakaryocytes 26
Meiosis 18, 1504, 1895, 1896
  crossing-over 1504
  figure 1505
  pachytene stage 1506
MEK 578, 579
Melanin 439, 887, 1432, 1434, 1435
  synthesis of, scheme 1433
Melanocortins 1748
Melanocyte(s) 439, 1434, 1435
  drawing 439
Melanocyte-stimulating hormone (MSH). *See* Melanotropin
Melanophores 439
Melanosomes 439, 1434
Melanotropin 1435, 1743
Melatonin 1445, 1447*s*, 1792, 1793, 1807
Mellitin 414
Melting curve
  for a diacylglycerol 395
  for DNA 256
Membrane(s) 5, 390 – 410
  active transport 410
  anchors. *See* Anchors, membrane
  artificial 392 – 394
  basement 10
  cation channels 421, **1767– 1773**
  chemical compositions, table 393
  coated 426*s*
  communication 427
  concentration gradient 206
  Davsen–Danielli model 393
  differential scanning calorimetry 394
  double, of gram-negative bacteria 5
  electrical properties 400
  fluid mosaic model 390
  fluidity 397
  freeze fractured 405
  inner side composition 401
  ionic atmosphere 400
  ionic gradients across 400
  lateral diffusion of lipids 399
  lateral phase separation 395
  lateral proton conduction 401
  lipid bilayer structure 390, 393
  lipid composition 392
  lipids, physical properties of 394
  metabolism, enzymes of 409
  NMR spectra 396
  outer side composition 401
  peripheral proteins 407
  permeability to neutral molecules 399
  pretransition temperature 394
  resting potential 400
  structure of **390 – 401**
  transport through **410 – 427**
  transporters 411– 427
Membrane Bohr effect 1040
Membrane cofactor protein (MCP) 1846
Membrane potential
  of mitochondria 1038, 1039
  of nerve axons 1039
Membrane proteins **401– 409**, 1723, 1724
  anchors 402
  glycoproteins 404
  integral 401
  positive-inside rule 401, 1724
Membrane transporters of mitochondria, table 1050
Memory
  chemical changes in synapses 1803
  chemistry of **1801– 1814**
  long-term 1801
  long-term depression (LTD) 1803
  long-term potentiation (LTP) 1803
  short-term 1801
  synapses 29, 1777– 1782. *See also* Synapses
  working 1802
Memory B and T cells 1859
Menadione 818, 1025*s*
Menaquinones (vitamin K_2) 818, 821, 1028, 1056
Mengo virus 247
Menkes’ disease 883
Mental illness 1808 – 1814
Menthol 1233*s*
Meprobamate 1788*s*
Mercaptans 624
Mercaptides 550
Mercaptoethanol 98
β-Mercaptoethylamine in coenzyme A 722*s*
7-Mercaptoheptanoylthreonine phosphate 814*s*
  biosynthesis 1189
3-Mercaptopyruvate 1407, 1408
Mercaptopyruvate sulfurtransferase 1408
Mercapturic acid 550*s*
Mercuric reductase 787
Mercury 317
  methylation of 876
Mercury ions
  binding to proteins 125
  reduction of
Meristems 1886, 1904
Meromycolic acid 1193*s*
Meromyosin 1098, 1102
Merozygote 1482
Merrifield, R. Bruce 84
Merrifield procedure, peptide synthesis 127, 128
MES buffer 99
Mescaline 1791*s*
Mesenchymal cells 1902
Mesoderm 1898
Mesolimbic dopamine system 1797
Mesophyll 1322
Mesosomal cells 1898
Mesosome 3
Mesozoa 23
Metabolic diseases. *See also* names of specific diseases
  of amino acid metabolism 1378, 1394, 1403, 1404, 1408, 1428, 1431, 1432
  anemia 826
  of bone 443, 678
  of carbohydrate metabolism 1002, 1145, 1169, 1170
  diabetes 1003 – 1005, 1748
  of lipid metabolism 943, 945
  lysosomal storage 1170 – 1172
  of mitochondria 1024
  of nucleic acid bases 1457, 1459
Metabolic loops 973
Metabolic pathways **505 – 507**
  scheme 506
Metabolic regulation **535 – 581**
  control elements of 536
  sensitivity coefficient 537
Metabolism. *See also* Specific compounds
  activation 507, 508
  beta oxidation 511, 512
  control by covalent modification reactions
    table 543
  control by substrate cycles 567
  definition of 1, 2
  effects of insulin 568
  fermentation **508 – 511**
  glycolysis 508 – 511
  introduction **505 – 527**
  oxidative phosphorylation 512 – 515
  regulation of. *See* Metabolic regulation
  study with isotopic tracers 110
Metabolite, activation 507, 508
Metachromatic leukodystrophy 1171
Metal-binding proteins, chelate effect 310
Metal-binding sites 312
Metal complexes
  formation constants, table 311
  stability of, transition metal 311
Metal ion(s) 296
  binding of 307
  charge 311
  geometry of orbitals 311
  hydration of 311
  ionic radii 311
    table 310
  toxicity 33
Metals, transition **837– 895**
Metalloenzymes 680
Metalloflavoproteins 794
Metalloproteases **625 – 627**
Metalloproteins 312
Metallothioneins **317**, 317*s*, 680
Metanephrine 1790*s*
Metaphase 1501
Metaphase plate 1502
Metaphosphate ion 639
Metarhodopsin 1329, 1331
Metastasis 573

Metazoa 23 – 29
Met-enkephalin 1752, 1752*s*
Methadone 1796*s*, 1797*s*
Methamphetamine (METH) 1796
Methane bacteria 1057
Methane formation 814
Methanofuran 814*s*
    biosynthesis 1189
Methanogens 7, 813 – 815, 879 – 881
Methanol, utilization by bacteria 1057
Methanol dehydrogenase 816*s*
Methanopterin 803, 804*s*
    biosynthesis 1460
Methemerythrin 862
Methemoglobin 358
Methemoglobin reductase system 359
5,10-Methenyl-THF 986
Methionine (Met, M) 53*s*, 539, 875, 1383
    biosynthesis of 540, 746
    conversion to *S*-adenosyl-methionine 977
    metabolic reactions of 1390
    metabolism of **1388 – 1391**
    in plants 1389
    sulfoxides 85
    toxicity of 1389
Methionine repressor 243
    operator complex 243*s*
Methionine sulfoxide 631*s*
Methionine synthase 813, 874*s*, 875
Method(s) **95 – 150**
    of carbohydrate analysis **188 – 192**
    isolating compounds 98 – 108
    of nucleic acid analysis 249 – 270
    of peptide synthesis 127– 129
Methotrexate 574, 805
8-Methoxypsoralen 1296*s*
5-Methylarginine 881
N-Methyl-D-aspartate (NMDA) 1786, 1786*s*
3-Methylaspartate lyase 685
Methylation
    of arsenic 876
    of carbohydrates 189, 190
    of carboxyl groups 592, 594
    of mercury 876
    of nucleic acids 591, 1541, 1542, 1887, 1888, 1894
    permethylation 189
    of protein side chains 79
    in regulation 543
    of uracil in tRNA 591
Methylcellulose 172
3-Methylcholanthrene
    hydroxylation of 1068
Methylcobalamin 867, 867*s*, 875
Methyl-coenzyme M reductase 879 – 881
Methyl corrinoids 876
β-Methylcrotonyl-CoA, carboxylation of 724
2-Methylcysteine 881
5-Methylcytosine 234, 1541*s*
α-Methyldopa 1791*s*, 1808, 1809
oxo-Methylene analog of transition state 626, 627*s*
β-Methyleneaspartate 1368
Methylene blue 1025*s*
Methylenecycloartenol 1246
Methylenedioxyamphetamine (MDA) 1791*s*
α-Methyleneglutarate mutase 871
Methylene-THF reductase 813
Methylesterase 563
N-Methylglutamate synthase 789
2-Methylglutamine 881
Methylglyoxal 695, 695*s*, 1391, 1588
    metabolism of 697
Methylglyoxal synthase 697
Methyl group(s) 802
    chiral 591
    synthesis of 813
    transfer reactions of 875
O^6-Methylguanine-DNA methyltransferase 1583
Methylhistidines 881, 1450
4-Methylidine-imidazole-5-one (MIO) 757
Methylmalonic aciduria 1394
    vitamin B_{12} level in 949
S-Methylmalonyl-CoA 950, 970
    in branched chain formation 992
    formation from propionyl-CoA 725
        equilibrium constant 725
Methylmalonyl-CoA mutase 871
    mechanism 873
    stereochemistry 874
    three-dimensional structure 874
Methylotrophic bacteria (methylotrophs) 815, 985
Methylphenidate 1809*s*. *See also* Ritalin
Methyl salicylate 1439*s*
6-Methylsalicylic acid 1212*s*, 1213
Methyltetrahydrofolic acid 875
S-Methyl thioacrylate 1409
Methylthioadenosine 1389
5'-Methylthioribose-1-phosphate 1389*s*
Methyltransferases 526, 563, **591– 593**
Mevalonate 993, 1227
    plant carotenoid precursor 1229
    sterol precursor 1229
Mevalonate diphosphate kinase 1227
Mevalonate kinase 1227
Mevalonic acid-5-diphosphate (pyrophosphate) 689, 690*s*
Mevalonic aciduria 1227
Meyerhof, O. 960
Mg Protoporphyrin 1403*s*
MHC. *See* Major histocompatibility complex
MHC restriction 1855, 1856
Micelles 390
Michael addition 681
Michaelis complex 751, 751*s*, 752
Michaelis constant 459, 464
Michaelis–Menten equation 459
Michaelis pH functions **96**, 296
Michel, Hartmut 84
mic RNA 1615
Microbodies. *See* Peroxisomes
Microcin 1722
Micrococcin 1691
Micrococcus denitrificans 1055
Microfilaments 369, 370
    behind desmosomes 29
Microfold (M) cells 1839
Microglia 1781, 1814
Microglobulin 1186
Microiontophoresis 1777
Micronucleus 1578
Microphotolysis 1297, 1298
Micropinocytosis 426
Microsatellites 1537
Microscopic binding constants 328, 329
Microscopic dissociation constants 305
    of pyridoxine 305
Microscopic reversibility, principle of 486
Microscopy
    atomic force 131
    chemical force 131
    confocal scanning optical microscopy 129
    electron 130
    near-field scanning optical 130
    negative contrast 130
    scanning tunneling 131
    STM image, glycogen 174
Microsome, definition of 14
Microtomes 130
Microtubule(s) 15, 369, 370
    associated proteins 372
    growth of 373
    labile 334, 370
    organizing centers 372
    stable 370
Microtubule-associated proteins 1119
Microtubule-organizing center 1119
Microvilli 10, 12
Miescher, Friedrich 199, 1473
Mifepristone (RU486) 1253, 1253*s*
Migraine 1793
Millipedes 24
Mineralocorticoid 1254
Miniature end-plate potentials 1777
Minichromosome(s) 1562
    micrograph of 1532
Minichromosome maintenance (MCM) proteins 1563
Minor groove of DNA 201, 213
Miraculin 1800
Mitchell, Peter 84, 1037, 1038
Mitchell's chemiosmotic theory 1037, 1038
Mitchell's hypothesis 1038
Mites 24
Mitochondria 14, 15, **1013 – 1050**, 1098
    architecture of 1013 – 1019
    chemical reactions in 1015, 1016
    complexes I – IV 1026 – 1032
    composition (table)
    cristae of 1014
    development 1018, 1019
    dimensions 1013
    diseases 1024
    E_m values 1039
    electron transport chain 1019
    energy-linked processes in 1046, 1047
    genes, catalog 1016
    genome 1016 – 1018
    growth of 1018, 1019
    inner membrane 512
    intermembrane space 1014
    matrix of 512, 1014
        definition of 14
    membranes 944, 1014, 1015, 1047–1050
    origin from bacteria 15
    of plant cell, in micrograph 13
    porins 1014
    processing peptidase 1018
    spectrophotometric observation 1033, 1034
    states of 1033
    structure, drawing 1015
    targeting signals 1018
    tightly coupled 1033
    uncouplers in 1046, 1047
Mitochondrial DNA. *See* DNA, mitochondrial
Mitochondrial electron carriers
    electrode potentials of, table 1035
Mitochondrial membrane(s) 1014, 1015
    exchange across 1047– 1050

  permeability 944
  transport across **1047–1050**
  transporters, table 1050
Mitochondrial membrane potential (E_m) 1038
Mitochondrial processing peptidase 625
Mitochondrion 10
Mitogen-activated kinase (MAP kinase) cascade 576–578, 1743
  figure 579
Mitogen-regulated protein (proliferin) 1745
Mitogenic activity 576
Mitomycin 225*s*
Mitosis 18, 371, 1501–1504
  figure 1502
Mitotic spindle
  formation 1503
Mixed acid fermentation products, table 968
M-lines (muscle) 1096, 1098
  proteins of 1099
Mobile defects in proteins 496
Modeccin 1685
Modification of enzyme
  by group transfer 545–549
  by phosphorylation 541–544
Modifiers of enzymes 473
Moesin 1119
Molal solution, definition of 287
Molar ellipticity 1286
Molar extinction coefficient 1275
Molecular activity (of enzyme) 457
Molecular beacons, fluorescence quenching 1291
Molecular biology, definition of 2
Molecular mass
  in daltons 4
  relative 4
Molecular rotation 1286
Molecular size, as basis of separations 100
Molecular weight 4. *See also* Relative molecular mass
Molecules, properties of
  chiral 42
  conformations of **43–45**
  cyclic symmetry 332, 333
  energy levels of 1276
  intersystem crossing 1290
  packing of **41**
  structural principles **39–46**
Molluscs 24
Molting hormone. *See* Ecdysone
Molybdenum **890–893**
  coenzyme forms 890–892
  nutritional need 893
  oxo-transfer reactions 892
Molybdenum-iron cofactor (FeMo-co) 892
Molybdenum-iron protein
  in nitrogenase 1361
  structure 1363
Molybdoenzymes
  active site structures 892
  dimethylsulfoxide reductase family 890
  nitrogenase family 890
  sulfite oxidase family 890
  xanthine oxidase family 890
Molybdoferredoxin 1361
Molybdopterin 803, 804*s*, 825
  biosynthesis 1462
  in formate dehydrogenases 892
Molybdopterin cytosine dinucleotide 891*s*
Monellin 1800
Monensin 414
  K$^+$ complex 415*s*
  synthesis 872
Monoamine oxidase (MAO) 1434, 1790
  in outer mitochondrial membrane 1016
Monoamine oxidase B (MAO-B) 1792
Monochromatic light 1275
Monoclonal antibodies 1841
Monocytes 26, 188
Monod, Jacques 84, 1475, 1603
Monodehydroascorbate reductase 1067
Monod-Wyman-Changeux (MWC) model of allosteric regulation 349, 350, 476
Monoiodotyrosine residues 1430
Monokines 1847
Monomer units, assembly of 982–993
Monooxygenases 1057, **1059–1072**
  cosubstrate of 1057
  external 1059
  flavin-containing 1059
  internal 1059
  pterin-containing 1061
Monosaccharide(s) **161–167**
  characteristic reactions 167
  ^{13}C-NMR spectrum of 162
  interconversion, scheme 1130
  interconversions of **1129–1140**
  reduction of 167
Monoterpenes, biosynthesis of 1232, 1233
Morels 20
Morgan, Thomas H. 83, 1504
Morphine 1443, 1444*s*, 1796*s*, 1796
Morphogen(s) 1880, 1882, 1884, 1885
Morphogenetic factors 1755
Morse curves 1276
Morula 1897
Mössbauer X-ray absorption spectra 1288
  of iron-sulfur clusters 860
Mosses 29
MotA, MotB proteins 1093
Motilin 1750
Motility genes of bacteria 1092
Motion, in nonmuscle cells **1118–1119**
Motion, in proteins 81
Motor cortex 1766
Motor neurons 1766
Mouse, genome 12, 1507, 1510
M-phase of cell cycle 579
MRF4 1902
M-ring 1092
mRNA 5, 230, 257, 536, 1474
  bicoid 1898
  binding proteins 1018
  caps 1642
  polyadenylate tails 1642
  splicing of 11
Mucins 181, 182, 1155
Mucolipidosis 1171
Muconate 1436*s*
cis,cis-Muconate lactonizing enzyme 692
Muconolactone 1436*s*
Mucopolysaccharidoses 659, 1169, 1170
  table 1171
Mucopolysaccharidosis I (MPS I) 1169
Müllerian inhibitory substance 1893
Multicatalytic protease 339, 620. *See also* Proteasome
Multidrug resistance protein 417
Multilamellar vesicles (liposomes) 392
Multiple attack concept 606
Multiple myeloma 1836
Multiple sclerosis 1769, 1865
Multipotential stem cells 1901
Multisubstrate enzymes, kinetics of 464
Munc-13, -18 1780
Muramic acid (Mur) 165*s*
Murein 170, 428, 429*s*. *See also* Peptidoglycan
l(+)-Muscarine 1783*s*
Muscarinic acetylcholine receptors 1785
Musci 29
Muscimol 1788*s*
Muscle(s) **1096–1118**
  calcium ions in 1114
  cardiac 26, 1096
  contraction of 1104–1114
  development of 1902
  electrolyte concentrations 206
  hereditary diseases **1112–1113**
  proteins of **1096–1104**
  rigor 1104
  rowing hypothesis 1104
  skeletal, crossbridges 1104
  smooth 26, 1096
    latch state 1117
  striated (voluntary) 26, 1096, 1767
  thin filaments 336
  T system of membranes 1114
  types of 1096
Muscular disgenesis 1115
Muscular dystrophy 822, 1112, 1113
  nutritional 822
Mutagens 222, 253, 254, 1584
  in environment 1584
Mutan 1152
Mutants
  conditionally lethal 1480
  temperature-sensitive 1480
Mutarotase 486, 526
Mutarotation 486
Mutases 637
  vitamin B$_{12}$ dependent 653, 654, 872
Mutation(s) 2, 1476, 1477
  addition 256, 1578
  autosomal recessive 16
  base substitutions 1578
  causes of 1578, 1579
  conditionally lethal 1480
  conservative 59
  deletions 256, 1476, 1578
  directed 1500, 1501
  frame-shift 1476, 1588
  harmful dominant 16
  heteroduplex formation 256
  insertions 1476
  lethal 16
  locating 1500
  missense 1480
  nonsense 1480
  oligonucleotide-directed 1500
  point 1476
  rate of 16
  transition 1476, 1578
  transversion 1476
Myasthenia gravis 1864
myb 576
myc, Myc 572, 576, 577, 579, 1633
Mycelium 20
Mycobacteria 431, 1166
Mycobacterium tuberculosis 431
Mycocerosic acid 1189, 1194
Mycolic acid(s) 431, 1168, 1193, 1193*s*, 1194
Mycoplasma(s) 2, 3
Mycoplasma genitalium 4

  genome 12
Mycorrhizae 20
Myelin 390, 399
Myelin basic protein 1769
Myelin sheath 1763, 1769
Myeloid cells 1901
Myeloperoxidase 856, 1073
Myf5 1902
Myoclonic epilepsy 1024
MyoD 1902
Myofibrils 1096
Myogenin 1902
Myoglobin 17, 68*s*, 70, **353 – 359**, 354*s*, 844, 1757
  bonding of dioxygen 851
  folding pattern 354
  mass spectra 113, 115
  NMR spectrum 148
  oxygenation curves 355
  packing of side chains 75
Myohemerythrin 71*s*
Myomesin 1099
Myosin 369, 1096, **1099** – **1104**
  coiled coil structure of 71
  diffusion constant of 461
  drawing 1105
  filaments of scallop, figure 1103
  heads, structure of **1105**
  heavy and light chains 1101
  monomer 1097
  photoaffinity label of 1111
  rod protein 1104
  scale drawing 1102
  structure 1106
Myotonia congenita 420
Myotonic dystrophy 1113, 1516
Myotubes 1902
Myrcene 1232, 1233*s*
Myriopoda 24
Myristoyl group 559, 1332
Myrosinase 1142
Myxococcus xanthis, genome 12
Myxomycetes 20

N

NAADP$^+$ 315, 779

Na$^+$, K$^+$-ATPase 423 – 425
  model 424
Nacrein 678
NAD$^+$ (NAD) 63*s*, 507*s*, 535, **765 – 771**, 767*s*, 778
  absorption spectrum of 768
  addition to 779
  containing ^{2}H 768
  determination of 779
  in fermentation 966
  formation from tryptophan 1446
NAD and NADP systems
  regulation of 980, 981
  NAD$^+$/NADH ratio, redox potential 767
NADH (NADH + H$^+$; NADH$_2$) **765 – 771**
  absorption spectrum of 768
  cosubstrate for monooxygenases 1059
  modification in acid 780
NAD^{2}H 771
  oxidation by 298
NADH dehydrogenase 1023
NADH-methemoglobin reductase 826
NADH:ubiquinone oxidoreductase (Complex I) 788, 1021, 1026
  3-dimensional image 1027
  oxidation by ferricyanide 780
NADH-X 780
NADP$^+$ (NADP) 507, **765 – 771**, 767*s*, 779
  in catalase 853
  formation from tryptophan 1446
  isolation of 767
NADPH (NADPH + H$^+$; NADPH$_2$) 507, **765 – 771**
  in photosynthesis 984
NADPH-cytochrome P450 reductase 1065
NAD(P)H dehydrogenase 1023
NADPH/NADP$^+$ ratio 776
NADPH oxidase 1072
Naive T cells 1859
Naloxone 1796*s*, 1797
Nanos protein 1898
Naphthalene, resonance energy 299
Narcolepsy 1808
Narcotics 1796*s*, 1796
Nathans, Daniel 84
Natural killer cells (NK cells) 185, 1073, 1831, 1852
Ncd 1107
  tubulin complex 1110
Near attack conformation 485
Near-field scanning microscope 130
Nebulin **1099**, 1100
Necrosis 1888
Neighboring group assistance in displacement reactions 601, 602
Nekoflavin 783
Nemaline myopathy 1113
Nematocyst 427
Nematodes 24, 25, 1866, 1893
Neocarzinostatin 218, 224*s*, 1589
Neolignans 1441
Neomycin(s) 1139, 1689*s*, 1690
Neopterin 803
Neoxanthin 1240, 1240*s*
Nephrogenic diabetes insipidus 412
Neprilysin 625
Nernst equation 400
Nerve growth factor (NGF) 1755
  effect of 1756
  ribbon drawing of 1754
Nerve impulse 1330 – 1332
  absolute refractory period 1768
  ion conducting channels 1769 – 1775
  in muscle contraction 1114
  propagation of 1767– 1769
Nervous system 1762 – 1813. *See also* Brain
  development of 1902 – 1904
  diseases of 1808 – 1814
Neryl diphosphate 1233*s*
Netrin-1, -2 1903
Netropsin 225*s*
  binding to DNA 225
Neu (protein) 576
neu (oncogene) 572
Neural plate 1898, 1903
Neural tube 1903
Neuraminic acid anion (Neu) 165*s*
Neuraminidase 186, 1650
Neuregulins 1753
Neurexins 1780
Neurocam 1154
Neurochemistry **1762 – 1814**
Neurodegenerative diseases 1811– 1814
Neurofibrillary tangles 1812
Neurofibromatosis 574
Neurofilaments 369
Neuroglia 25
Neurohormones 1742
  of invertebrates 1758 – 1760
  of molluscs 1759
  table 1751
Neuroleukin 1756
Neuromodulators 1782
Neuromuscular junctions 1766, 1781
Neuron(s) 25, 1762. *See also* Axon
  acetylcholine-secreting 1784, 1785
  action potential 1763
  all-or-none response 1763
  anatomy of 1762 – 1765
  axon, action potential 1765
  axon, diagram 1763
  axon, micrograph 1764
  cell body 1762, 1763
  cholinergic receptors 1784
  depolarization 1763
  formation of noradrenaline 1064
  functions of 1762 – 1765
  ionotropic receptors 1774
    signaling pathways 1775
  metabolism 1776
  metabotropic receptors 1774
    signaling pathways 1775
  muscarinic 1785
  neuronal pathways, diagram 1766
  nicotinic 1784, 1785
  NO synthase in 1757
  postsynaptic membrane 1763
  potassium ion channels 1769 – 1773
  presynaptic end 1763
  schematic drawing 1763
  sodium ion channels 1769 – 1772
  of squid 24
Neuropeptide(s) 1795
Neuropeptide Y (NPY) 1750
Neuropeptide YY 1750
Neurophysin II 1748
Neurospora crassa 20
  figure 21
Neurotensin 1750
Neurotoxin(s) 1775, 1776
  produced by body 1798
  Shiga toxin 1685
  structures of 1770
Neurotransmitter(s) 29, 553, 1432, 1762, **1782 – 1796**
  agonists 1784
  amino acids 1785 – 1789
  antagonists 1784
  criteria 1782
  receptors for 479
  release of 1777– 1781
  uptake, G-protein dependent 1779
Neurotrophin(s) 1755
  ribbon drawing of 1754
Neutron diffraction 137
Neutron scattering 1681
Neutrophils 26, 188, 1863
  reactions of reduced oxygen compounds 1073
  respiratory burst of 1072
Nevi (moles) 1435
NF-κB/Rel proteins 1634
Nicastrin 1814
Nickel **877– 882**

  - dietary essential 877
  - in plants 877
- Nickeloplasmin 877
- Nick translation 257
- Nicotinamide 507, 721, 766, 767, 769*s*, 780. *See also* NAD^+, NADH; $NADP^+$, NADPH
  - daily requirements 756, 769
- Nicotinamide adenine dinucleotide. *See* NAD^+, NADH
- Nicotinamide mononuceotide 1549
- Nicotinamide ring 770, 779
- Nicotinate mononucleotide 1446
- Nicotine 769, 1382, 1783*s*, 1796
  - oxidation to nicotinic acid 769
- Nicotinic acetylcholine receptor 1784*s*, 1785
- Nicotinic acid. 769*s*, 1445*s*. *See also* Nicotinamide
- Nicotinic acid adenine dinucleotide (NAAD) 778
- Nicotinic acid adenine dinucleotide phosphate ($NAADP^+$) 315
- Nicotinic acid hydroxylase 825
- Nicotinuric acid 1446
- Nidogen 437
- Niemann-Pick C1 disease 1251
- Niemann-Pick disease 1171
- *nif* genes
  - sequence of 1365
- Night blindness 1241
- Nigrostriatal pathway 1789, 1792
- NIH shift 1062
- Nijmegan breakage syndrome 1586
- Ninhydrin 120, 121*s*
- Nitrate 1366, 1367
  - as electron acceptor 1054
  - as inorganic oxidant 1051
- Nitrate reductase(s) 1054
  - assimilatory 890
  - dissimilatory 890
- Nitrate reduction system scheme 1055
- Nitrate respiration 1367
- Nitric oxide, NO 553, 561, 1069 – 1072, 1753, 1756, 1757, 1795, 1796, 1804
  - binding to hemoglobin 358, 359
  - blood pressure regulation 1756
  - formation of 1055
  - formation by bacteria 1055
  - as hormone 1069
  - in innate immune system 1756
  - metal centers, binds to 1757
  - neurotransmitter functions 1756, 1757
  - reactions 1069
  - transcription regulation 1757
- Nitric oxide reductase 1055
  - tetrahydrobiopterin, requirement for 1071
- Nitric oxide synthase(s) 1071, 1072
  - endothelial 1071
  - inducible 1071
  - mechanism 1071
  - neuronal 1071
  - specialized cytochromes P450 1071
- Nitrification 1359
- Nitrifying bacteria 1051, 1052
- Nitrite 1366, 1367, 1757
  - reduction of **1054 – 1056**
- Nitrite reductase(s) 836*s*, 861, 1055
  - copper in 884, 885
  - cytochrome *cd*$_1$ type 885
- Nitroform 126
- Nitrogen
  - fixation 1359 – 1367
  - metabolism of, **1359** – **1410**
  - reduction of 1359, 1360
- Nitrogen cycle **1359 – 1367**
  - scheme 1360
- Nitrogen fixation genes 1365
- Nitrogen-fixing symbionts 1360
- Nitrogen mustards 574
- Nitrogen pool 1359
- Nitrogen regulatory protein C (NtrC or NrI) 1614
- Nitrogenase(s) 890, **1360 – 1366**
  - components 1361
  - mechanism of 1362
  - structure of 1361, 1363
- Nitromethane, inhibition of ammonia-lyases 756
- *p*-Nitrophenolate anion, as a leaving group 636
- *p*-Nitrophenyl acetate 610, 610*s*
- Nitrosamines 1587
- *Nitrosomonas* 1051
- *S*-Nitrosothiols 1071
- Nitrous acid, reaction with nucleic acids 253
- Nitrous oxide, formation of 1055
- Nitrous oxide reductase(s)
  - copper in 884, 885
- Nitroxyl radical 1060, 1060*s*
- NMDA receptors 1903, 1786, 1787
- NMR spectra
  - of ADP, ^{31}P 642
  - of aspartate aminotransferase 149
  - of carbohydrates 190 – 192
  - of DNA, ^{1}H 269
  - of glycyl-L-histidylglycine 147
  - of gonadatropin-releasing hormone 145
  - of honey, ^{13}C 162
  - of lysozyme 146
  - of muscle, ^{31}P 304
  - of myoglobin 147
  - of phosphatidylcholine, ^{1}H 396
  - of phospholipid phases, ^{31}P 397
  - of pyridoxal 5'-phosphate, ^{1}H 138
  - of pyruvate kinase equilibria 641
  - of runner's arm, ^{31}P 1117
  - of serine proteases, ^{1}H 613
  - of tRNA 268
  - of tRNA, ^{19}F 270
- NMR spectroscopy **137– 149**, 494
  - assignment of resonances 145
  - band widths 138
  - basic principles 137
  - carbon-13 140
  - chemical shift 138
    - table 141
  - correlation time, τ_c 142
  - coupling constants 138, 139, 146
  - deshielded 139
  - deuterium 140
  - distance constraints 145
  - downfield 139
  - DSS or Tier's salt 138
  - exchangable protons 148
  - fluorine-19 140
  - Fourier transform 141
  - free induction decay 141
  - HMQC (15N-multiple quantum coherence) 144
  - HSQC (15N-multiple quantum coherence) 146
  - isotope-edited spectra 144, 145
  - J coupling 139
  - liposomes 396
  - membranes 396
  - mixing time 143
  - multidimensional 143
  - nitrogen-15 140
  - NOE 140
  - NOESY spectra 142, 146
  - nuclear Overhauser effect 140
  - phosphorus-31 140, 295
    - of adenylate system 303
    - of muscle 304
    - of phosphotransferases 640 – 643
  - positional isotope exchange 641
  - protein dynamics 145
  - relaxation times 141
  - residence time, τ_m 142
  - ring current shifts 139
  - saturation-transfer 641
  - scalar coupling 139
  - solid-state 149
  - spin decoupling 139
  - spin-lattice relaxation time, T_1 142
  - spin-spin interaction 139
  - spin-spin relaxation time T_2 142
  - stable isotopes, use of 110, 111
  - study of phosphotransferases 639
  - tetramethylsilane (TMS) 138
  - titrations 147
  - transferred NOE 145
  - two-dimensional 141
    - correlation spectroscopy (COSY) 144
    - TOCSY 144
  - of various nuclei 141
  - of vicinal protons 139
- Nobel prizes, list 83
- Nociceptors 1800
- Nodal protein 1899
- Nodes of Ranvier 29, 1763, 1769
- Nod factors 1148, 1149, 1365
- Noggin protein 1903
- Nogo 1903
- Nojirimycin tetrazole 605, 605*s*
- Noladin ether 1795
- Non-coding (transcribing) strand in DNA 237
- Noncompetitive inhibition 476, 477
- Nonheme iron content in mitochondria 1019
- Nonheme iron proteins. *See* Iron-sulfur and Diiron proteins
- Nonhomologous end-joining DNA repair 1583
- Nonlinear equations 460
- Nonmetallic ions, ionic radii, table 310
- Nonmuscle cells, motion in **1118 – 1122**
- Nonproductive complexes 466, 475
- Nonsulfur bacteria 1306
- Norandrolone phenyl propionate 1255*s*
- Norepinephrine (noradrenaline) 553, 553*s*, 1064*s*, 1431*s*, 1433, 1782, 1789, 1790*s*, 1791
  - in receptor 555*s*
  - synthesis of 1064
- Norethindrone 1253, 1253*s*
- Normetanephrine 1790*s*
- Notch pathway 1814
- Notch protein 1900
- NtrB protein 1614
- Nuclear envelope 11
- Nuclear lamina 1536
  - structure 1535

Nuclear magnetic resonance spectroscopy. *See* NMR spectroscopy
Nuclear matrix 1535
Nuclear membrane 1536
Nuclear Overhauser effect (NOE) 140. *See also* NMR spectroscopy
Nuclear pore 11, 1536
  electron micrograph of 1536
Nuclear proteins
  nonhistone 1535
Nuclear relaxation
  by chromium ions 639
  by manganese ion 639
Nuclear scaffold 1535
Nuclear transporters
  exportins 1536
  importins 1536
Nucleases 250, 1452
  micrococcal, active site 653
Nucleic acid(s) 199 – 213, 218, 1529 – 1542
  abbreviated structures 204
  absorption of light 205
  acid–base chemistry 204, 205
  amount in cells 31
  bases, aromatic character of 204
    stacking of 209
  biosynthesis 995, 1542 – 1563
  catenated 218
  cruciform structure in 229
  dynamics 218
  electrophilic reagents 253 – 255
  homology of 256
  hydrolysis of 249
  interactions with proteins **239 – 244**
  isolation of 249
  knots in 218, 219
  melting curves of 255
  metabolism of bases, **1450** – **1459**
  methods of study **249 – 270**
  pK_a values of 204
  separation of 249
  tautomerism 204, 205
  topology **218 – 230**
  turnover of 527
Nucleic acid bases
  alkylation 253
  aromatic character 207
  chemical reactions of **207**
  epoxides 254
  halogenation 253
  photoreactions of 1296
  reactions 251– 255
    with formaldehyde 253
    with kethoxal 253
    nitrous acid 253
  salvage pathways 1453, 1454
Nucleoid 4
Nucleolin 1537, 1639
Nucleolus 10, 11, 1537
Nucleophilic catalysis 494. *See also* Covalent catalysis
Nucleophilic character and bonding to metal ions 307
Nucleophilic displacement reactions 526, **589 – 662**
  carbocationic intermediates 598
  cellulases 602
  chitinase 599 – 602
  glycogen phosphorylase 604, 605
  β-glycosidases 602
  glycosyltransferases 593 – 598
  inversion 590, 591
  kinetic isotope effects 592, 593
  lysozymes 599 – 602
  starch-hydrolyzing enzymes 605 – 608
  transmethylation 591, 592
Nucleophilic group, definition of 589
Nucleophilic reactivity (nucleophilicity) 589
  alpha effect 590
  effect of polarizability on 590
  of sulfur compounds 590
Nucleoplasmin 1897
Nucleoside(s)
  catabolism 1452, 1453
  conformations of 211, 212
  names and abbreviations, table 203
  torsion angles of 211
Nucleoside diphosphate kinase 655
Nucleoside phosphorylases 1453
Nucleoside Q (Queosine) 234*s*
Nucleoside trialcohols, fluorographs 252
Nucleoside triphosphates 302
  in nucleic acid biosynthesis 995
Nucleoside Y 234*s*
Nucleosome(s) 239, **1531– 1535**
  chains, folding of 1534
  DNA in 221
  model 222
  positioned 1531
  structure 1533
Nucleotide(s) **199 – 213**, 200*s*
  chemistry of **203 – 213**
  conformations of 211– 213
  modified 234
  names and abbreviations, table 203
  pairs 3
  structures of 203 – 213
Nucleotide binding domain 76
  glyceraldehyde phosphate dehydrogenase 63*s*
Nucleotide handles
  of coenzymes 719
  and specificity of metabolism 720
Nucleotide kinases 655
Nucleotide sequence
  of aspartate aminotransferse 202
  of DNA and RNA chains 200
  of genes and genomes 12, 201, 1507– 1516
  of mitochondrial DNA 201
  of viruses 201
Nucleotidylation 545
Nucleotidyltransferases **657**
Nucleotidyltransferase superfamily 1548
Nucleus 10, 11, **1535 – 1537**, 1893 – 1897
  of plant cell, micrograph 13
Nurse cells 1895

O

O-Antigen 180, 428 – 430
Obelia 24
Occlusion zones 29
1-Octadecanol 1195
Octahedral geometry 311
Octopamine 1789, 1790*s*
Octopine 1498*s*
Octopus, intelligence of 24
Oct-3/4 protein 1897
Oculocutaneous albinism 1435
Ocytocin 1748*s*
Odor 1798
Oils, composition of 380
Oil seeds, glyoxylate pathway 988
Okadaic acid 545
Okazaki fragments 1544
Old yellow enzyme 783
Oleamide 1758, 1795, 1808
Oleate hydratase 526, 688
Oleic acid 381
Oleoresin 1236
Olfactory bulb 1799
Olfactory cortex 1799
Olfactory epithelium 1799
Olfactory responses 558
Oligomer(s)
  asymmetry 344 – 347
  cubic symmetry 342 – 344
  with dyad axes 337– 348
  heterologous tetramer 352
  isologous tetramer 352
  quasi-equivalence 344 – 347
Oligomeric enzymes 475
Oligomycin-sensitivity-conferring protein (OSCP) 1041
Oligonucleotides, 200 – 205
  synthesis of 258
  probe 119
Oligosaccharide(s) **167– 170**, 1140 – 1142, 1153 – 1158. *See also* Specific names
  asparagine-linked (*N*-linked) 180 – 185
  biological recognition of 161
  extensions 1157
  glycoprotein, trimming of 1156, 1157
  glycoprotein precursor, biosynthesis 1157
  glycosylation 1155
  metabolism of 1140 – 1142
  N-linked 182 – 185, 192*s*
    assembly of 1155 – 1168
    in eukaryotes 1155
  O-linked 1155
  in plants 1149
  polylactosamine chains 1157
  precursor, trimming of 1158
  release from glycoproteins 188
  separation of 100
  utilization of 1140
Oligosaccharin 1149
Oligosaccharyltransferase 1156, 1722
Omega oxidation 943
Ommochromes 1444
OmpA protein 411, 428
OmpC (osmoporin) 411
OmpF 411, 412*s*
Oncogenes 242, 544, **571– 576**
  in human tumors 571
  table 572
Oncogenic viruses 248
One-start helix 334
Onions 1408
Oocytes
  development 1895
  maturation 1895
Oparin, I. V. 9
Open systems 289
Open tetrapyrroles 1304
Operator, aroH 1612
Operator sequence of DNA
  ^{1}H NMR spectrum 269
Operons 240
  bacterial 1611. *See also* Bacterial operon
  constitutively expressed 1605
  electron micrograph 1610

  - *gal* (galactose) 1613
  - inducible 1604
  - *lac* 1603 – 1607
  - negatively controlled 1604
  - operator (O) 1604
  - promoter (P) 1604
  - regulatory gene 1604
  - repressor (R) 1604
  - *trp* 1616
  - *trp*, attenuation 1617
- Opiates 1797
- Opines 1498
- Opioid receptors 1797
- Opsin(s) 553, 1325
- Opsonization 1839
- Optical coherence tomography 1298
- Optical rotation 42
- Optical rotatory dispersion (ORD) 1286, 1287
- Optical traps **1108**
  - schematic drawing 1110
- Optical tweezers 1298
- Optic nerve 1798
- Optimum rate for enzymes 469
- *d* Orbitals, participation in covalent bond formation 311
- Orcinol 251
- Ordered binding 464
- Ordered sequential mechanism 475
- Orexins 1808
- Organelle(s) 11
  - DNA of 1540
  - separation 101
- Organic mechanisms, use of arrows 528
- Ornithine, synthesis of 1374 – 1379
- Ornithine aminotransferase 1374
- Ornithine decarboxylase 342, 342*s*, 1382
- Ornithine mutases 871, 874
- Ornithine transcarbamoylase 1377
- Orotate 1450
- Orotidine 5'-phosphate 1450, 1451*s*
- Orsellinic acid 1212, 1212*s*
- Orthologs 1510
- *Oscillatoria* 22
- Oscillator strength of electronic transitions 1282
- Osmate ester 393*s*
- Osmium tetroxide 1098
- Osmolarity (OsM) 1142
- Osmoprotectants 1142
- Osmotic adaptation 1142, 1143
- Osmotic pressure 1142
  - adaptation to changes in 1143
- Osmotic shock 417
- Osteoadherin 1154
- Osteoarthritis 438, 1154
- Osteoblasts 26, 441
  - alkaline phosphatase in 645
- Osteocalcin 821, 1259
- Osteoclasts 441
  - acid phosphatase in 645
- Osteocytes 441
  - alkaline phosphatase in 645
- Osteogenesis imperfecta 438
- Osteoglycin 1154
- Osteonectin 313
- Osteopetrosis 572, 678
- Osteopontin 409, 441, 572
- Osteoporosis 443, 619, 1258
- Otoliths 1801
- Ouabain 422, 1265
- Ouavain 1265*s*
- Ovomucoids 183, 629
- Ovoperoxidase 1074
- Ovotransferrin 840
- Ovum
  - cleavage of, figure 1896
  - cortical granules 1896
  - development of 1894, 1895
  - fertilization 1896
  - mitotic divisions 1897
- Oxalate 699*s*
- Oxaloacetate 506*s*, 515, 516*s*, 699*s*, 737*s*, 952
  - acetylhydrolase 698
  - concentration in mitochondria 952
  - decarboxylases 527, 706
  - regeneration of in citric acid cycle 952
- 4-Oxalocrotonate decarboxylase 705
- 4-Oxalocrotonate tautomerase 697
- Oxalosuccinate 516*s*, 705, 952
- Oxaziridine 1060, 1060*s*
- Oxazolidinones 1690
- Oxazoline 601*s*
- Oxidation
  - of biochemical substances, table 290
  - linked 705
  - by NAD$^+$, table 290
- α-Oxidation of fatty acids **942 – 946**
- β-Oxidation **511– 513**, **939 – 944**
  - pathway 506, 511, 981, 1015
  - peroxisomal 941
  - fatty acids 941
- ω-Oxidation 943
- Oxidation-reduction, by dehydration 697
- Oxidation-reduction potentials. *See also* Reduction Potential
  - of pyridine nucleotide system 766
  - table 301
- Oxidation-reduction reactions
  - of coenzymes **765 – 827**
  - internal 982
  - of iron-sulfur clusters 857
- Oxidative damage
  - to tissues 1074
- Oxidative decarboxylation 952
- Oxidative decarboxylation of α-oxoacids 511, 736, **796 – 802**
  - by hydrogen peroxide 799
  - mechanisms 799 – 801
  - with thiamin diphosphate 800
- Oxidative phosphorylation 512, 962, 1024, **1032 – 1047**
  - chemiosmotic theory 1037, 1038
    - scheme 1037
  - coupling factors 1041
  - high energy intermediates 1036, 1037
  - mechanism of 1036 – 1041
  - mitochondrial sites of 1015, 1033, 1034
  - stoichiometry (P/O ratio) 1033, 1034
  - thermodynamics of 1034 – 1036
- Oxidizing power of redox couples 300
- Oxidoreductases 497
- Oxidosqualene cyclase 1244
- Oxo acid(s)
  - chain elongation process **990 – 992**
- α(2)-Oxoacid(s) (α-keto acids). *See also* specific acids
  - cleavage of 802
  - decarboxylation 730
  - oxidative decarboxylation **796 – 802**
- β-Oxoacid(s)
  - decarboxylation 705
- α(2)-Oxoacid dehydrogenases 348
  - sequence of reactions, diagram 798
- β-Oxoacyl-ACP 978*s*
- 3-Oxoacyl-CoA transferase 527
- Oxoadipate 1386*s*, 1436*s*
- α-Oxo alcohol, reaction with thiamin 732
- 5-Oxo-3-aminohexanoate 1388*s*
- 2-Oxobutyrate 1383, 1384*s*
- Oxocarbenium ion 598*s*, 604
- 2-Oxo-3-deoxy-6-phosphogluconate 966*s*
- 4-Oxo-6-deoxy-CDP-glucose 1138
- 2-Oxo-3-deoxy-ᴅ-*glycero*-ᴅ-*galacto*-nononic acid (KDN) 1137
- 3-Oxodipate 1434
- 2-Oxoglutarate (α-ketoglutarate) 516*s*, 737*s*, 1399, 1399*s*
  - as decarboxylating cosubstrate 1063
  - dehydrogenases 796
  - synthase 799
- 2-Oxo-ʟ-gulonate 1134, 1134*s*
- 2-Oxoisovalerate (α-Oxoisovaleric acid) 700*s*
  - as precursor to valine 993
- Oxonium ion. *See* Oxocarbenium ion
- 2-Oxopantoate 810
- Oxophenarsine 597*s*
- 5-Oxoproline, 551*s*, 662*s*
- Oxosteroid isomerase 526, 696
  - molecular activity 696
- Δ^5-3-Oxosteroid isomerase
  - turnover number of 458
- Oxosuberate 1189s
- OXPHOS diseases 1024
- Oxyanion 613
- Oxyanion hole 614, 615*s*
  - formation of in trypsinogen 615
  - of serine proteases 614
- Oxygen (O_2)
  - binding of, to heme proteins 353 – 358, 850 – 852
  - to hemocyanin 885, 886
  - paramagnetism of 850
  - reactivity of 850
- Oxygenases 863, **1057– 1072**
  - diiron 863
  - iron-containing 1068
  - nitric oxide synthases 1069 – 1072
- Oxyhemoglobin 851
- Oxyluciferin 1346*s*
- Oxyperoxidase 854
- OxyR protein 1075
- Oxytocin 54*s*, 1747
- Oysters, glycogen content 31

P

- P1 artificial chromosomes (PACs) 1497
- p53 1568, 1907
- Pacemaker neurons 1802
- Packing
  - of macromolecules **332 – 348**
  - of molecules 41
- Pactamycin 1681, 1691
- Paget's disease 443
- Pain 1800
- Palindromes
  - in DNA 238
  - H- (mirror-repeats) 238
- Pallidin 405
- Palmitic acid 381
- Palmitoleoyl-ACP 1192
- Palmitoyl group **559**
- Palytoxin 1776
- Pancreatic hormones 1748, 1749

Pancreatic lipase 1181
Pancreatic polypeptide 1749, 1749*s*
Pancreatic trypsin inhibitor 629
Pancreozymin-cholecystokinins (CCK) 1749
Pantetheine 722*s*, 1393
Pantetheine 4'-phosphate 722
  in acyl carrier proteins 723
Pantoic acid 722*s*, 723, 1393
  in coenzyme A 722
Pantothenic acid 721, 722, 722*s*
  nutritional requirement of 756
Papain 618
  active site 618
Papaverine 1443, 1444*s*, 1447
Paper chromatography 103
Papovaviruses 244
Parabactin 838, 838*s*
Paracellin 1782
Parallel β structure 62
Paralogs 1510
Paramagnetic ions, effects on NMR signal 640
Paramecium 17, 20
Paramylon 170, 174
Paramyosin 1104
Paraquat 1792, 1811*s*
Parasites 1866
Parasympathetic system 1767
Parathion 636*s*
Parathyroid hormone (parathyrin, PTH) 314, 1258, 1750
Parathyroid hormone-related protein 1750
D-Paratose (3,6-dideoxy-D-glucose) 180*s*
  from CDP-glucose 1138
Pargyline 1792, 1809*s*
Parkinson disease 1789, 1792, 1812
Parnas, J.K. 960
Paromomycin 1690
Paromycin 1689*s*
Paroxysmal cold hemoglobinuria 1865
Partition coefficient 410
Parvalbumin **312 – 316**, 1114, 1788
Parvoviruses 244
Passenger DNA 1491
Pasteur 966
Pathogen-associated molecular patterns (PAMPs) 1831
Pathogenesis-related proteins 1149
Patterson map 135
Pauling, Linus C. 83, 84
Paxillin 406
pBR322 plasmid 1494
  cloning procedure with 1494
p 53 cancer suppressor 574, 581, 1907
P-clusters 1361– 1364
PDZ domain 367
Pectate lyase 65, **686**
Pectate lyase C 686, 686*s*
Pectin(s) 15, 170, 177, 1149, 1150
  biosynthesis 1130
Pellagra 721, 769, 1446
L-Penicillamine 739
Penicillin 20, 1164, 1164*s*
Penicillin acylase 620
Penicillopepsin 621
Pentalenene 1234*s*
Pentalenene synthase 1234
Pentameric acetylcholine-binding protein
  ribbon drawing 1784
Pentose phosphate cycle 1051
Pentose phosphate equilibrium 964
Pentose phosphate pathway 700, **963 – 965**, 972
  carboxylation system of 986
  isomerizing system 963
  nonoxidative 966
    scheme 965
  oxidative 964, 984
    scheme 965
  reductive 517, 984
Pentoses 162
Peonidin 1214
PEP. *See* Phosphoenolpyruvate
PEP carboxykinase. *See* Phosphoenolpyruvate carboxykinase
PEP carboxylase. *See* Phosphoenolpyruvate carboxylase
PEP carboxytransphosphorylase 706
PEP mutase 711
PEP synthase
  mechanism in *E. coli* 990
Pepsin 456, 621
  specificity 117
Pepsinogen 625
Pepstatin 623
Peptidase(s) 526, 609
  leader 620
  procollagen 433
  tripeptidyl 140, 610
  signal 620
Peptide(s) **51– 59**
  as attractants 1753
  β conformation of 60
  endogenous opioid 1752
  knotted chains 74
  torsion angles 60
Peptide deformylase 627
Peptide hormone(s) 1743 – 1756, 1758 – 1760
  amidation of 522
  synthesis of, diagram 1742
Peptide linkage 51, 55*s*
  cis 56*s*
  dimensions 55
  resonance stabilization of 55
  trans 56
Peptide mapping 118, 360
Peptide structure
  torsion angles of, table 61
Peptide synthesis **127– 129**
  nonribosomal 1713
  solid-phase 127, 128
Peptide unit
  hydrogen bonds 56
  partial double bond character 55
  planarity of 55
  properties of **55**, **56**
  structure **55**, **56**
Peptidoglycan(s) 179, 428, 429*s*
  biosynthesis of 1130
    figure 1162
  layer 1160
Peptidyl carrier protein domain 1713
Peptidylglycine α-amidating enzyme 887
  reaction cycle of 1064
Peptidylglycine α-hydroxylating monoxygenase 887
Peptidyl-prolyl-(*cis-trans*) isomerases 488
Peptidyl transferase(s) 234, 637, 994, 1678
  reaction 1672, 1710
Peptidyltransferase center (ribosome) 1672, 1687, 1707
  A loop sequence 1706
  antibiotic action site 1689
  central loop sequence 1706
  intermediate structure 1705
  P loop sequence 1706
  RNA 23S, structure 1706
Peptidyl-tRNA hydrolase 1710
Perforins 1846
Performic acid 115
Perfusion chromatography 100, 101
Peridinin 1308, 1309*s*
Perinuclear space 11
Periodate oxidation
  of carbohydrates 190
Periodental disease 442
Peripherin-RDS 1332
Periplasm 427, 428
Perlecan 1154
Permeability coefficient 410
  for D-glucose 415
Permease 411
  for lactose 417, 1604
  periplasmic 417– 419
Pernicious anemia
  vitamin B_{12} treatment 868, 869
Peroxidase(s) 844, 845, **852 – 855**
  ascorbate 852
  bromo 856
  catalytic cycles, scheme 855
  chloro 855
  compounds I, II 853 – 855
  cytochrome *c* 852
  glutathione 856
  halo 855, 889
  heme linkage 853
  horseradish 852, 856
  iron properties 853
  lacto 853, 856
  mammalian 853
  mechanisms of catalysis **853 – 855**
  myelo 856
  reaction 1208
  structure 853
  thyroid 856
Peroxidation 1204 – 1210
Peroxides 852 – 857, 1072 – 1075
Peroxisomes (microbodies) 10, 14, 941
Peroxynitrite 1069
Persistence of fetal hemoglobin 1901
Perutz, Max F. 84
PEST regions 527
PET 1802, 1806
Petunidin 1214
Peyer's patches 1834
pH **95 – 98**
  chymotrypsin dependence 616
  of cytoplasm 549
  effects on enzymatic action 469
  effects on equilibria **293 – 297**, 305
  effects on metal binding 310
  Michaelis pH functions 296
  observed values of 1039
  within a cell 295
PH (plekstrin homology) domain 102, 367, 566, 571
Phaeomelanins 1434, 1435
  synthesis of, scheme 1433
Phaeophyta 22
Phage display methods 1726
Phage f 29 1559
Phage head 363
Phage lambda 248, **1483 – 1486**
  genome map 1622
  integration of, figure 1570

    lysogenic state 1622, 1623
    repressors 241
Phage tail 363
Phagocytes, killing mechanisms 1074
Phagocytic cells 1831, 1833 – 1835
Phagocytosis 13
Phagosomes 426
Phalloidins 1625
Phantom atoms 42
Phase fluorometers 1290
Phase transition(s) 284
    of lipid bilayer 394
Phenazine methosulfate (PMS) 1025*s*
Phenobarbitol 1809*s*
Phenol(s)
    acidity of 308
    conversion to phenyl glucuronide 1141
    $\mathrm{p}K_a$ value of 293, 309
    resonance energy 299
Phenolsulfonic acid 99
Phenylacetate 1431*s*
Phenylacetic acid 308, 309
Phenylaceturic acid 513*s*
Phenylacetylglutamine 1431*s*
Phenylalanine (Phe, F) 52, 52*s*
    absorption spectrum 1284
    absorption spectrum of derivative 123
    catabolism, microbial 1434
    metabolism of **1428 – 1443**
        scheme 1431
    metabolism of, in plants 1438 – 1443
    synthesis of 1424, 1425
Phenylalanine ammonia-lyase **755**, 1438
Phenylalanine hydroxylase 1061
Phenylephrine 563*s*
Phenyl glucuronide 1141*s*
Phenylisothiocyanate 118
Phenylketonuria (PKU) 16, 1062, 1428
Phenyllactate 1431*s*
Phenylpropanoid pathway 1438
    of plants, scheme 1439
Phenylpyruvate 689, 1422*s*, 1424, 1431*s*
Phenylpyruvate tautomerase 692, 1850
Phenylthiohydantoin 118
Pheophytins 1303
Pheoporphyrin 1302
Pheromones 382, 1196, 1758 – 1760
Phloem 30
Phlorhizin 1214, 1214*s*
Phoborhodopsin 1335
phoR 1614
Phorbol esters 564
Phosducin 558
Phosphagen 303, **1117**, **1118**
Phosphatase(s) 515, 526, 637, **645 – 647**, 984
    for hydrolysis of fructose bisphosphate 989
    specificity of 977
Phosphate anhydride 303
Phosphate, inorganic (P_i)
    in adenylate system 302 – 304
Phosphatides 383, 384, 1397
    structures of 383
Phosphatidic acid 383*s*, 1198*s*
Phosphatidyglycerol 383*s*, 384
Phosphatidyglycerol-3-*P* 1197
Phosphatidylcholine 383, 383*s*, 391*s*, 1398
    in bilayer 391
    packing 391
Phosphatidylethanolamine 383, 383*s*, 1197
    bilayer, AFM image of 392
Phosphatidylinositol 383*s*, 1200
Phosphatidylinositol (PtdIns) 563, 565, 566*s*
    in signalling **563 – 566**
Phosphatidylinositol 3-kinase 565
Phosphatidylinositol 4,5-bisphosphate 563
Phosphatidylserine 383*s*, **564**, 1197, 1198*s*, 1199, 1204, 1398
    decarboxylation of 753
Phosphatidylserine decarboxylase 409, 755
Phosphatidylserine synthase 1720
3'-Phosphoadenosine 5'-phosphosulfate (PAPS) 659, 977*s*
Phosphoadenylation
    of protein side chains 79
Phosphocan 1154
Phospho compounds, active 975
Phospho (phosphoryl) group 292*s*
    in catalysis 962
    chirality of 644
    as a handle 507
    transfer from ATP 660
    transfer of 292
    transfer of, by kinases 654
Phosphocreatine 1379*s*
Phosphodiesterase 556, 638, 1330
    for cGMP 558
Phosphoenolpyruvate (PEP) 510, 510*s*, 685*s*, 952, 973
    key metabolic intermediate 705, 706
    as stabilized enol 697
    sugar chain extension 1136, 1137
Phosphoenolpyruvate carboxykinase (PEPCK) 549, 706, 1144
Phosphoenolpyruvate carboxylase 527, 711, 952
Phosphoenzyme 653, 703
Phosphofructokinase 962
    deficiency of 1144
    glycolysis regulation 999
    in regulation 535
Phosphofructokinase-2 541
Phosphoglucomutase 653
6-Phosphogluconate 964*s*
    in Entner-Doudoroff pathway 966
    ion 164*s*
6-Phosphogluconate dehydratase 697
6-Phosphogluconate dehydrogenase 770, 963
Phosphogluconate pathway 972. *See also* Pentose phosphate pathway
6-Phosphogluconolactone 964*s*
3-Phosphoglyceraldehyde dehydrogenase 770
2-Phosphoglycerate 510*s*, 685*s*
    dehydration 685
    dehydration to PEP 705
3-Phosphoglycerate 506*s*, 707*s*, 986
    in chloroplasts 706
    as key intermediate in biosynthesis 973
Phosphoglycerate kinase 654, 656
Phosphoglycerate mutase 653, 1144
Phosphohexoketolase 972
O-Phosphohomoserine 746
4-Phosphohydroxy-L-threonine 1464*s*
Phosphoinositide(s) 383, 384, 1132, **1199 – 1201**
    cascade 565, 1202
Phosphoketolase 736
Phosphokinases 526
Phospholamban 1116
Phospholipases 635
    A_1 566
    A_2 566, 635
    C **563**, 566, 1202
    D 566, 1204
Phospholipid(s) **379 – 387**, 392, 1197. *See also* Specific substances
    bilayer diagram 391
    biosynthesis 996, 1196 – 1201
        scheme 1198
    head groups, functions of 396
    inverted hexagonal phase 397
    ^{31}P NMR 397
    non-bilayer structures 397
    synthesis regulation 1201
    transfer protein 1248
Phospholipid hydroperoxide glutathione peroxidase 1894
Phosphomannomutase 654
Phosphomutases 526
Phosphonamidate 626*s*
    as metalloprotease inhibitor 626
Phosphonate
    esters 626*s*
    synthesis of 711
Phosphonoethylamine 385*s*
Phosphono group 711
Phosphonolipids 385
N-(Phosphonomethyl)glycine (glyphosate) 1424, 1424*s*
4'-Phosphopantetheine 723, 994
4'-Phosphopantothenylcysteine decarboxylases 753, 755
Phosphophoryns 442
Phosphoporin (PhoE) 411, 412*s*
Phosphorescence 1288 – 1294
Phosphoribosylamine 1454, 1455*s*
Phosphoribosylanthranilate 1426*s*
Phosphoribosyl pyrophosphate (PRPP) 1425, 1425*s*
Phosphoribosyl transferases (ribonucleotide pyrophosphorylases) 1453
Phosphoribulokinase 1320
Phosphoric acid 99
    in lipids 380
Phosphorolysis, 593
Phosphorothioate 1615
    phosphotransferases in study of 642
Phosphorus-32
    properties of 110
Phosphorylase 541– 543, 989, 1143, 1145
Phosphorylase kinase 541, 544, 1144
Phosphorylation 33, 541– 545, 637
    in bacteria 545
    cascade, reactions of 542
    coupling by 977
    oxidative 303, 962 – 1045
    photosynthetic 303, 962. S*ee also* Photosynthetic phosphorylation of protein side chains 79
    substrate level 775, 800, 962, 1036
Phosphorylation-dephosphorylation cycle in biosynthesis 977
    reductive pentose phosphate pathway 986
Phosphorylation, photosynthetic. *See* Photosynthetic phosphorylation
Phosphorylation state ratio (Rp) 980, 1034 – 1036
    adenylate system 302, 303, 1035
O-Phosphoserine 545, 610*s*
Phosphotransferase(s) 637– 659
    bacterial system 419, 420
    in-line displacement 638
    P31 NMR of 642, 643

- pentacovalent intermediate 638
- stereochemistry 642, 643

Phosphotriesterase 646
- mechanism 646

Phosphotyrosine 545
Phosvitin 80
Photoacoustic spectroscopy 1287
Photoactive yellow protein (PYP) (xanthopsin) 1336
- structure 1336

Photoaffinity labels 1297
- of myosin 1111

Photoautotrophic bacteria 8
Photochemistry 1290, **1294 – 1298**
- chemical equilibria in the excited state 1295

Photodynamic therapy 1297
Photofootprinting 266, 1296
Photohydration 1296
Photoinhibition 1319
Photolyases 1296, 1339, 1340
Photons 1273
Photooxidation 1295
Photophosphorylation 1300, 1301
- cyclic 1300

Photoprotein 1342
Photoreactivation 1339
Photoreceptors 1798
- UV-B 1320

Photoreduction 1295
Photoreduction system
- Z scheme, diagram 1299

Photorespiration 707, 1321– 1324
- C_3 plants 1321
- C_4 plants 1321

Photorhodopsin 1329
Photostationary state 1330
Photosynthesis 506, 517, 705, **1298 – 1324**
- bacterial 1301
- control of 1319 – 1321
- control mechanisms, figure 1320
- ferredoxin/thioredoxin system 1320
- hydrogen formation 1324
- light-induced transcription 1320
- overall equation 1298
- reaction centers 1310
- resonant transfer of energy 1292

Photosynthetic bacteria 7, 1303
Photosynthetic phosphorylation 303, 517, 962
Photosynthetic pigments 1302 – 1305
Photosynthetic reaction centers 71
- schematic view 1300

Photosynthetic units 1301
Photosystem I (PSI) 1299, 1309
- of cyanobacteria 1314
- of green plants 1314
- reaction center, models 1315

Photosystem II (PSII) 1299, 1309
- intrinsic proteins of, structures 1316
- manganese cluster of, structure 1317
- oxygen formation 1317

Phototaxis 1273, 1335
Phototherapy 1297
Phototransformation 1337
Phototropism 1273
Phthalate dioxygenase 1059
Phthaldialdehyde 120*s*
Phycobiliproteins 1305, 1306
Phycobiliviolin 1305
Phycobilisome 1305, 1306
Phycocyanins 1304
Phycocyanobilin 1304, 1305, 1305*s*
Phycoerythrins 22, 1304
Phycoerythrobilin 1305, 1305*s*, 1404*s*, 1405
Phycomycetes 20
Phycourobilin 1305
Phylloquinones (vitamin K1) 818, 821
Phylogenetic approach to RNA structure 232, 1676
Phylogenetic tree 7, 8, 1677
Physarum polycephalum 20
Physostigmine 1783*s*
Phytanic acid 943*s*
Phytanyl group 385
Phytate. *See* Inositol hexakisphospate (InsP_6)
Phytin 1132
Phytoalexins 1149, 1232, 1869
Phytochelins 318
Phytochrome 1305, 1320, 1337*s*, 1338
- absorption spectra 1338

Phytocyanin 883
Phytoene 1231, 1236, 1237*s*, 1239*s*
- conversion to carotenes 1237
- synthesis of 1236, 1237

Phytoglycolipids 388, 389*s*
Phytol 943
Phytosphingosine 388, 388*s*, 1203*s*
Picornaviruses 247
Picrotoxin 1235*s*, 1789
Piebaldism 1435
Piericidin A 1021*s*
Pigments(s) 1302 – 1305, 1324 – 1329
- of butterflies 803
- photosynthetic 1302 – 1305
- visual 1324 – 1329

Pigment-dispersing hormone 1760
Pigment epithelium 1332
Pili 3, 6, 335, 336, 364. *See also* F pili
Pimaricin 1217
Pimeloyl-Coenzyme A
- conversion to biotin 719, 723

Pinacol 712*s*
Pinacolone 712*s*
Pinacol-pinacolone rearrangement 712
Pineal body 1793, 1807, 1808
α-Pinene 1232, 1233*s*
Pinene cyclase 1232
Ping-pong mechanism 465, 475
- rate equation for 466
- succinyl-CoA-acetoacetate CoA transferase 662

Pinocytosis 13
Pinoresorcinol 1440
Pinosylvin 1214, 1215*s*
Pipecolate 1386*s*
Piperazine 99
Piperidine 99
Piperine 1439*s*
PIPES 99
Pitch
- of B-DNA 213
- of a helix 68

Pituitary gland 1742
- adenohypophysis 1743
- neurohypophysis 1743
- pars intermedia 1743

Pituitary hormones 54, **1743 – 1748**
- glycoprotein 1745
- growth 1745, 1753

Pit viper venom 1775
pK_a values **95 – 98**. *See also* Specific compounds
- of acids, table 293
- of amino acid side chains 55
- of buffers, table 99
- of buried groups 55
- of dicarboxylic acids, table 328
- effect on ion exchange separation 104
- in hemoglobin 357
- of hydrogen-bonded protons 494
- succinic acid 330
- of terminal – COOH 55
- of terminal – NH3$^+$ 55
- thiamin 330

Plague 7
Planarians 23
Planck's constant 1274
Plane polarized light 1283
Plankton 22
Plant(s) (green)
- C_4 cycle for concentrating CO_2 1322, 1323
- cell, electron micrograph 13
- cell walls of 1149
- composition of 31
- defenses 1149, 1440, 1869
- development of 1904, 1905
- flavonoids, estrogenic activity 1263
- gametophytic phase 1904
- genes, cloning 1497, 1498
- growth regulators 1762
- gums, biosynthesis of 1130
- hypersensitive response 1869
- phenylpropanoid pathway 1439
- photorespiration 1321-1324
- photosynthesis. *See* Photosynthesis
- programmed cell death 1869
- receptor-mediated surveillance 1869
- resistance (R) proteins 1869
- restorer genes 1904
- self-incompatibility 1904
- sporophytic phase 1904
- surface lipids 1196
- systemic acquired resistance 1869, 1440, 1762
- toxins 1685

Plant cells and tissues **29**, **30**
- fiber 11
- thermogenic 1048

Plasma amine oxidase 886
Plasma antithrombin III 177
Plasma cells 1834, 1840
Plasma lipoproteins 1181– 1184
Plasma membrane (plasmalemma) 12, 379
Plasmalogen(s) 383*s*, 384, 1200*s*, 1201
Plasmenylethanolamine 1201
Plasmids 5, 248, 249
- bacterial 1530
- ColE1 248, 1493
- drug resistance 248
- pBR322 1494
- pBR322, cloning procedure 1494

Plasmin 634
Plasminogen 634
Plasmodesmata 10
- in micrograph 13

Plasmodium 19, 20
Plastics
- biodegradable 1200

Plastids 14, 15
Plastochromanols 818
Plastocyanin 65*s*, 883
Plastoquinone(s) 818, 818*s*, 1028, 1427, 1429*s*
- biosynthesis of, scheme 1429

Platelet(s) 26

  cyclic AMP in 557
Platelet activating factor **384**, 384*s*, **385**, 1201, 1757
Platelet-derived growth factor (PDGF) 544, 571, 1755
Platyhelminthes 23
Pleated sheet 62
  structure 62
Plicatic acid 1440, 1441*s*
P loop 425, 648, 658
  consensus 559
  of Ras 559
PLP. *See* Pyridoxal phosphate (PLP)
PLP-dependent reactions. *See* Pyridoxal phosphate-dependent reactions
Pluripotent cells 1833, 1886
PMP. *See* Pyridoxamine phosphate
PNA, 227*s*
Podophyllotoxin 1441, 1441*s*
Poison(s). *See also* Insecticides, Antibiotics, Inhibitors
  cyanide 590
  hydroxylamine 590
  mushrooms 1625
Polar membrane 5
Polar molecules 50
  definition of 48
  hydration of 50
Polarizability 590
Polarization of water 47
Polarization P, equation 1291
Polarized absorption recovery after photobleaching 1044
Polarized light 1275
Poliovirus(es) 247
Poly(A)-binding protein 1702
Polyadenylation of RNA 202
Poly (ADP-ribose) 1584
Poly (ADP-ribose) synthase 545
Polyamines 222, **1380 – 1383**
  biosynthesis, scheme 1381
  catabolism of 1382
  functions 1380
  synthesis, regulation 1381, 1382
Polycarboxylic acid synthases **700 – 703**
  table 700
Polyene antibiotics 1164
Polyenoic acids 1205
Polyethylene, helical structure 45
Polyethyleneglycol (PEG) 102
Polyfluorethylene (Teflon)
  helical structure of 45
Polyglutamates 803
Polyglycine II 72, 73*s*
Polyhistidine “tag” 106
Poly-β-hydroxybutyric acid 3, 1200
Polyisoprenoid compounds 993
Polyketide(s) **1212 – 1217**, 1443
  antibiotics 1216
  origin of, postulated 1212
  plant defensive compounds 1216
β-Polyketone 1212
Polylactosamine chains 1157
Polymer(s). *See also* Proteins, Nucleic acids, Glycogen
  biosynthesis 993 – 996
  catabolism 995
Polymerase(s). *See* DNA, RNA and other specific polymerases
Polymerase chain reaction (PCR) 260, 261
Polymetaphosphate 3
Polymorphonuclear leukocytes 26
Polymyocystitis 1698, 1865
Polynucleotide(s) 199*s*, **226 – 228**. *See also* Nucleic acids, RNA, DNA
  hybridization 255
  hydrazine, reactions with 254
  interactions with ions 218
  light absorption properties of 1284
  melting of 255
  probes 255
  rotational movements of bases, figure 217
  triple-helical **226**
Polynucleotide kinase 258
Polynucleotide polymerases 638
Polyol pathway 1131
Polyomavirus 247
Polypeptide(s) **56 – 59**. *See also* Proteins and Peptides
  antibiotics 66
  chemical synthesis 85
  conformation of **59 – 61**, 78
  definition of 51
  exit tunnel, ribosome 1687
  torsion angles 59 – 62
Polypeptide loop through translocon 1724
Polyphosphates 302, 303, 976
Polyploid 1506
Polyprenyl (isoprenoid) compounds 946, **1227– 1266**, 1230, 1231
  biosynthesis 992
  biosynthetic pathway, scheme 1228
  chain length of 1230
  polymerization of 712
Polyprenyl diphosphate synthases 1230
Polyprenyl synthases 1227
Polyproline 72, 73
Polyproteins 522, 628
Polyribosome 1708
Polysaccharide(s) **169 – 180**, 994, 995
  alternating, list of 170
  of bacterial surfaces 179, 180
  biosynthesis 994
  catabolism of **1143 – 1153**, 1169
  chains, conformation of 170
  lipid-dependent synthesis 1152, 1153
  main linkages, table 170
  structural features of, table 170
  structure patterns 1150 – 1153
  synthesis of **1143 – 1153**
Polysialic acid 187
Polysome. *See* Polyribosome
Polysulfide, linear
  from glutathione 1053
Polytene chromosomes 1507
Polyubiquitin 524
Polyunsaturated fatty acids
  oxidation of 1059
Pompe disease 1145, 1171
P/O ratio 1041
  in oxidative phosphorylation 1033, 1034
  in respiration 1033
Porifera. *See* Sponges
Porins 65, 402, 411, 428
  FepA 839
  FhaA 839
Porphin(s) 843
Porphobilinogen 843*s*, 1399, **1399 – 1405**, 1401*s*
Porphobilinogen deaminase 1400
Porphyrias 1403
Porphyrin(s) 843, **1399 – 1405**
  biosynthesis 742, 1401
  chelate effect 310
  tautomerism of 845
Porphyropsins 1328
Positron emission tomography (PET) 1806
Postganglionic fibers 1767
Postsynaptic density 1781
Postsynaptic membrane 1780, 1781
Postsynaptic potential 1763
Posttranslational modification of proteins 402, 521– 523
Potassium **206**
Potassium ion(s) 31
  activation of enzymes by 206
  binding to tetraplex DNA 206
  concentration in cells 206
  effect on equilibria 297
  replacement by proteins 206
  in tissues 299
Potassium ion channel(s) 1769 – 1774
  inactivation gate 1772
  inward rectifying (Kir) channels 1773, 1774
  model of 1773
  voltage-regulated 1772
Potential energy curves
  graph 1280
Potential energy diagram 1289
Potentiometer 300
Poxvirus 247
ppGpp. *See* Guanosine polyphosphates
PQQ. *See* Pyrroloquinoline quinone
Preangiotensinogen 1261
Precalciferol 1257
Precursor activation 540
Prednisolone 1254, 1254*s*
Pregnenolone 1253, 1256*s*
Prenatal diagnosis 26
Prenyl diphosphate (pyrophosphate) 390*s*
Prenylation 402
  of proteins 559, 1231, 1232
Prenyltransferases 712
Prephenate 485, 690*s*, 1422*s*, 1424
  conjugative elimination 689
Prephenate dehydratase 1424
Prephytoene-diphosphate 1236, 1237*s*
Preproenkephalin A 1752
Preproinsulin 519
Prepro-opiomelanocortin 522, 1743, 1745*s*
Preproparathyrin 1750
Preprorenin 1261
Preproteins **519 – 522**. *See also* Proenzyme(s)
  peroxisomal 521
pre-rRNA precursor 1619
Presenilin-1, -2 1813, 1814, 1902
Presqualene alcohol diphosphate 1236, 1237*s*
Pressure jump method 468
Prestin 1801
Presynaptic membrane proteins 1780
Preterminal protein 1559
Pre-tRNAs 1646
Preuroporphyrinogen (hydroxymethylbilane) 1400, 1401*s*
Primary adaptive immune response 1833
Primary cell wall
  cellulose fibrils in 1149
Primary glomerulonephritis 1865
Primary hydration sphere 311
Primary kinetic isotope effect
  for glucose-6-phosphate isomerase 592, 593, 693
Primary structure

  of nucleic acids 200 – 202
  of protein 56
Primases 1552
Primer for fatty acid synthesis 990
Priming proteins 1551
Priming reaction in metabolism 507
Primosome 1544, 1552, 1558
  assembly sequences 1558
  micrograph 1556
Prion(s) 248, 1718, 1719
  diseases 1718, 1812
  protein (PrP) 1709, 1718
Priority sequence 42
Procapsid 366
Procarcinogen 1588
Processive mechanism
  of α-amylase 606
Processivity factor 1549
Prochiral centers 478, 768
  trigonal 480
Prochlorophytes 10
Procollagen 432, 432*s*, 1063
Procollagen peptidases 433
Proctolin 1760
Product inhibition 464, 475
Proenkephalin B 1752
Proenzyme(s), 519, 609, 1722. *See also* Zymogen
  of histidine decarboxylase 754
Proflavine 223*s*
Progenitor cells 1901
Progesterone 696, 1253, 1256*s*
Progestins 1252, 1253
Proglucagon 1749
Programmed cell death. *See* Apoptosis
Progress curve for enzyme-catalyzed reaction 455
Proinsulin 519
Projection formula
  Fisher 42
  Newman 44
Prokaryotic cells
  differentiation in 1890 – 1893
  definition of 2
Prolactin 1745
Proliferating cell nuclease antigen (PCNA) 1562
Proline (Pro, P) 52*s*
  biosynthesis of, scheme 1373
  catabolism of 1374
  in helices 69
  reductases 753, 755
  synthesis of 1374
Proline oxidase 1374
Proline rings, torsion angles 62
Promethazine 1809*s*
Promoter(s)
  - 35 region 1608
  eukaryotic 1631
  Pribnow sequence 1607
  sequences 1607
  upstream activating region (UAR) 1608
Prontosil 473*s*
Proofreading
  by enzymes 482
1,2-Propanediol 874*s*
β Propeller 67, 560, 764
2-Propene sulfenic acid 1408
S-1-Propenyl-L-cysteine sulfoxide (onions) 1408
Prophage state 1483
Prophase 1501
Propionate
  catabolism 947– 950
  conversion to oxaloacetate 950
Propionic acid, pK_a value of 293
Propionic acid bacteria 725, 954
  fermentation of lactate 970, 971
Propionyl-CoA 940, 950, 970
  in branched chain formation 992
  catabolism of 947– 950
  carboxylation of 724
Propionyl-CoA carboxylase
  deficiency of 949
Propionyl group 992
Propranolol 553*s*, 1791*s*, 1791
Prostacyclin(s) 557, 1206*s*
Prostaglandin(s) 565, 1193, 1757
  effect on blood pressure 1262
  metabolism of 1207, 1208
  oxygenases in formation of 1057
Prostaglandin D2 1795, 1808, 1863
Prostaglandin D synthase 1186
Prostaglandin E_2 1190*s*
Prostaglandin H synthases 1208
Prostaglandin PGE_2 381*s*
Prostanoids 381, 1206
  physiological effects of 1210, 1211
  structures of 1206, 1209
Prostate gland, zinc content 680
Prosthetic groups 312, 719
Protamines 204, 239, 1531– 1535, 1894
Protease (proteinase) 609. *See also* Proteosome
  aspartic 621– 625
  ATP-dependent 627, 628
  in blood clotting 631– 634
  cysteine 618, 619
  in defense systems 628
  E. coli 628
  functions of 628
  metallo- 625 – 627
  multicatalytic 339, 525, 620
  pH dependence of 616
  serine. *See* Serine proteases
  specificities, table 117
  subsites of, nomenclature 617
  substrate specificities 616 – 618
  V-8, specificity 117
  stereoelectronic effect 616
Protease inhibitors **622**, 629 – 631, 1149
  alpha 1 629, 631
Proteasome 523 – 525, 620, 627, 1728
  20S 339, 341*s*
  26S 340
  electron micrograph 340
Protection mapping. *See* Footprinting
Protein(s)
  absorption spectra 1284
  amino acid composition 59, 115 – 123
  anchors for 402
  antioxidant, transcriptional regulation of 1075
  aromatic clusters 75
  caloric value 283
  charged groups in 101
  chemical modification **79 – 81**
  classification of residues 79
  compositions of specialized 59
  constitutive pathway 1724
  content
    from nitrogen analysis 31
    in tissues 31
  cotranslational folding 1721
  cotranslational processing 1721
  covalent structure 57
  crosslinking **79 – 81**
  crossover connections 74*s*
  denaturants of 1021
  disulfide crosslinkages 80
  domains 74
  elongation arrest 1722
  engineering of 85
  families 76
  folding 1727
  functional, formation of 1722
  globular 59
  growth factors 1753 – 1756
  hydration of 50
  hydrogen bonding of side chains 75
  hydrogen bonds in 75
  hydrolysis 116 – 118
  idealized shapes 60
  induction of synthesis 538, 539
  intracellular degradation 523
  isoelectric point 83
  maturation 519
  membrane 1723, 1724
  methylation of 79
  molecular mass 59
  multienzyme complexes 348
  nucleotide-binding domain 76
  oxidation of 85
  oxygen carrying **850 – 853**
  packing density of 75
  pH, effects of 83
  phosphoadenylation of 79
  phosphorylation of 79
  precursor 1722
  prenylation of 1231, 1232
  primary structure 56
  processing of 519, 1720 – 1728
  proteolytic degradation of 1728
  regulated pathways 1727
  repression of genes 538, 539
  ribosomal 1677– 1681
  salting in 101
  secretion of 519, 520, **1720 – 1728**
  separation with ammonium sulfate 101
  selenium-containing, table 825
  side chain groups, reactions of **123 – 127**
  solubility 101
  solvent, effects of 83
  splicing 1716
  sulfate esters of 79
  synthesis rate 1708
  thermostability 85
  translocation through membranes 1722, 1723
  translocation into organelles 1723
  transport of 519 – 521
  turnover 538, 539, 1720 – 1728
  uptake 1019
  water molecules in 76
Protein(s), association
  binding equilibria 350 – 352
  dimerizing 350 – 352
  interfaces 74
  quaternary structure 332
  regulatory subunits 348
  subunits of 74, 332
  tertiary structure 59
  unsymmetrical dimerization 344

Protein(s), conformational states
  compact denatured state 82
  conformational changes in 81
  conformational substates 496
  denaturation of 82
  dynamic properties **81– 85**
  local unfolding 81
  mobile defects in 496
  packing defects 81
Protein(s), folding patterns **76**, **77**, 496, 497
  α/α barrel 77
  α/β barrel 77
  architecture of folded **59 – 76**
  folding intermediates 82
  folding of 81– 83, 518, 519
  four-helix 71, 841
  molten globule 82
  motifs 367
  β sandwich 64
  secondary structure 63
  seven-bladed β propeller 561
  seven-helix 553 – 556
  symmetry in 77
  winged helix 243
Protein A 1839
Protein C 634
Protein coats of viruses 334
Protein concentrations, estimation of 102
Protein data bank 149
Protein disulfide isomerases 83, 659, 787
Protein-DNA interactions 266
Protein G 1839
Protein kinase 541, 656, 659
  cGMP-activated 544
  myosin light chain 1116
Protein kinase C 544, 564, 1204, 1753
Protein kinase, cAMP-dependent active site 657
Protein phosphatases 544, 646
Protein-protein interactions 1725
Protein S 634
Protein sequenators 118
Protein sequences from genes 119
Protein synthesis 3, 538, 539, **1669 – 1729**
  broken transcripts 1710
  codon usage 1710
  elongation 1672
  elongation cycle, figure 1703
  elongation factor EF-G 1705
  errors 1709, 1710
  eukaryotic initiation factors 1700 – 1702
  genetic code expanded 1712
  initiation of 1672, 1698 – 1702
  initiation of, figure 1699
  peptidyltransferase reaction 1704, 1705
  peptidyl-tRNAs, lost 1710
  polypeptide elongation 1702 – 1709
  programmed frameshifts 1712
  prokaryotic initiation factors 1700
  read-through translation 1712 – 1714
  release factors 1709
  ribosomal frame-shifting 1712 – 1714
  ribosome cycle 1698 – 1720
  selenocysteine incorporation 1711
  suppressor genes 1711, 1712
  translation termination 1672, 1709
  translocation 1672
Protein transport machinery, diagram 1018
Protein tyrosine kinases 544
Proteinase. *See* Protease
Proteoglycan(s) 181, 182, **1153 – 1160**, 1884. *See also* Glycosaminoglycans
  aggregates of cartilage 182
  aggregation factor 1155
  binding to proteins 1154
  leucine-rich 1154
  molecular architecture 182
  modular 1154
  terminal unit of polysaccharides 181
Proteolipid protein 1769
  of myelin 401
Proteome 116
Proteomics 1729
Prothoracicotropic hormone 1760
Prothrombin 632, 821
Prothrombinase 631
Proticity 1038
Protista 18 – 22
Protocatechuate 1436*s*, 1438
Protocatechuate dioxygenase(s) 1057
Protofilament 1089
Protoheme(s) 844, 845*s*
Proton(s)
  conduction, lateral 401
  dissociation, diffusion-controlled 493
  intramolecular transfer 693
  inventory technique 493, 618
Proton electrochemical potential 1038
Proton pumps 1037– 1041. *See also* ATPase
  ATP-driven 1045, 1046
  in mitochondria 1038
  stoichiometry of 1040
Proton transfer
  coupled 493
  Marcus theory 493
  rates 492
  ultra-fast 491
Protonmotive force 1038
  in movement of bacterial flagella 1092
Proto-oncogenes **571– 576**, 1907
Protoplast(s) 6
Protoporphyrin IX 1403
Protoporphyrinogen oxidase 1402
Protozoa **18 – 20**
Provirus 248
Proximity factor in catalysis 495
Prozac 1793, 1793*s*, 1810
Pseudoazurin 884
Pseudo first order reactions 457
Pseudogene 1538
Pseudohermaphroditism 1254
Pseudoknot in RNA 231
Pseudomonad(s) 8, 688
Pseudomonic acid 1689*s*, 1691
Pseudorotation 212, 638
Pseudosubstrate of acetylcholinesterase 636
Pseudouridine 234*s*, 1620
  synthases 1620
Pseudo vitamin B_{12} 867
Psicose 164*s*
Psilocybine 1447, 1447*s*
Psoriasin 313
Psoriasis 316, 440, 1297, 1755, 1855
Psychiatric disorders, drugs for 1809
Psychotropic drugs 1798
PTEN 566
Pteridine 803
  in hydroxylases 1061
  ring system 802 – 804
Pteridophyta 29
Pterin(s) 802, 1460
  biosynthesis, scheme 1461
  coenzymes **802 – 813**, 1061, 1062, 1071, 1072
  metabolism 1460
  structures 802 – 804
Puffballs 20
Pufferfish, genome 12, 1507, 1511
Pullulan 170, 175
Pulse fluorometers 1290
Purine(s) 199, 199*s*, 1443
  catabolism of 810
  free bases, catabolism 1458
  nucleosides, catabolism 1458
  nucleotide cycle 1456
  nucleotides
    biosynthesis, scheme 1455
    catabolism 1458
  oxidative metabolism 1459, 1460
  salvage pathway 1456
  synthesis, enzymatic 1454 – 1456
Purine bases
  absorption spectra of 205
  hydrogen-bonding possibilities 207
  names of 203
  pK_a values of 204, 205
Purine hydroxylase 890
Purine nucleoside phosphorylase 1457
Purine salvage pathways, scheme 1457
Purkinje cell(s) 1767
  diagram 1768
Puromycin 1687, 1689, 1689*s*
Purple acid phosphatases 658, 862, 863
Purple Cu$_A$ center 883, 884
Purple membrane 1333
Purple photosynthetic bacteria 1306
  cyclic photophosphorylation 1314
  electron flow in 1313
  reaction centers 1310 – 1314
Putidaredoxin 859, 1068
  EPR spectrum of 860
Putrescine 363, 1380*s*
  biosynthesis 1381
Pycnodysostosis 619
Pyochelin 838*s*
Pyramidal tract 1766
Pyranose rings 161, 162, 175. *See also* Sugars
Pyrazinamide 1194*s*
Pyrenoid 20
Pyrethroid 1788
Pyrethroid insecticides 1775
Pyridine
  binding constants to metal ions 311
  pK_a value of 293
  resonance energy 299
Pyridine nucleotide(s) 765
  chemistry 779, 780
  coenzymes **767– 780**
  content in mitochondria 1019
6-Pyridone(s) 780*s*
Pyridoxal 738
Pyridoxal phosphate (PLP) 121, 679, **737– 753**, 740*s*. *See also* Pyridoxal phosphate-dependent reactions
  absorption spectrum 749
  as active site label 124
  amino acid reactions 740
  in aspartate aminotransferase 57
  as coenzyme, table 743
  dependent enzymes, cyclization 741
  enzymes, atomic structures 543, 605, 750 – 753
  in glycogen phosphorylase 604
  mechanism of action 740, 741, **751– 753**

  - NMR spectrum of 138
  - optical properties 749, 750
  - Schiff bases, reactions of 699, 705, 742
  - transimination 741
- Pyridoxal phosphate-dependent enzymes, list 743
- Pyridoxal phosphate-dependent reactions **741– 747**
  - beta replacement 742 – 744
  - cyclization 741
  - decarboxylation 745
  - elimination 742
  - ketimine intermediate 745, 746
  - nonenzymatic models 737
  - racemization 741
  - side chain cleavage 745
  - stereochemistry of 749
  - transamination 741
- Pyridoxamine 738
  - orange product with ninhydrin 121
- Pyridoxamine phosphate (PMP) 741, 744*s*, 747
  - in biosynthesis of dideoxy sugars 1138
- Pyridoxamine:pyruvate aminotransferase, stereochemistry of 748
- Pyridoxine (pyridoxol, vitamin B_6) 305*s*, 738, 1463
  - microscopic dissociation constants 305
  - pK_a values of 305
  - tautomerism 45
- Pyridoxol phosphate 1464*s*
  - biosynthesis of 1464
- ε-Pyridoxyllysine 750*s*
- Pyrimethamine 805*s*
- Pyrimidine(s) 199*s*, 1383
  - biosynthesis of **1450 – 1454**
  - catabolism of **1450 – 1454**
- Pyrimidine bases
  - absorption spectra of 205
  - hydrogen-bonding possibilities 207
  - names of 203
  - pK_a values of 204, 205
- Pyrimidine nucleotides
  - catabolism 1452, 1453
  - synthesis of 1451, 1452
- Pyrimidine ring
  - assembly of, scheme 1451
- Pyrimidinodiazepine 1461*s*
- Pyrite 1054
- Pyrithiamin 731, 736
- Pyrogen 1211
- Pyroglutamate aminopeptidase 619
- Pyroglutamyl group as N-terminus
  - in proteins 54, 522
- Pyrophosphatases 515, 646, 976
- Pyrophosphate (PP_i) 508, 515
  - in adenylate system 302 – 304
  - hydrolysis of **976, 977**, 994
- Pyrophosphate esters
  - formation of 976
- Pyrophospho group
  - transfer from ATP 660
- Pyrophosphoric acid 99, 293
- Pyrrole, resonance energy 299
- Pyrrolidone 1382*s*
- Pyrroline 1382*s*
- Pyrroline carboxylate 1373*s*, 1374, 1374*s*
- Pyrroloquinoline quinone (PQQ) 764, 765, 815
  - as coenzyme 815, 816
- Pyrrophyta 20
- Pyruvaldehyde 1588. *See also* Methylglyoxal
- Pyruvate 506*s*, 510, 952, 1391
  - anaerobic cleavage 800
  - conversion to acetyl-CoA 962
  - decarboxylation, reversible 799
  - from Entner–Doudoroff pathway 966
  - in fermentation reactions 962
  - formation of 960
  - uridylic acid, conversion to 978
- Pyruvate carboxylase 711, 724, 952
  - control mechanisms 729
  - deficiency of 1144
- Pyruvate decarboxylase 718, 734, 735
- Pyruvate dehydrogenase 511, 735, 796 – 798, 797*s*
  - effect of insulin 571
- Pyruvate:ferredoxin oxidoreductase 799
- Pyruvate:flavodoxin oxidoreductase 1362
- Pyruvate formate-lyase 799 – 801, 875
- Pyruvate kinase 656, 706, 962
  - equilibria by NMR 641
- Pyruvate oxidase 734, 735, 799
  - active site 734
- Pyruvate-phosphate dikinase 990, 1322
  - mechanism 990
  - reaction cycle, scheme 990
- Pyruvic acid, pK_a value of 293
- Pyruvoyl groups in enzymes, **753 – 755**
  - table 753
- 6-Pyruvoyltetrahydropterin 1460
- 6-Pyruvoyltetrahydropterin synthase 1460

Q

- Q cycle 1028, 1040
- Quanine quartet 208*s*
- Quantosomes 1301
- Quantum mechanical tunneling 494, 771, 848
- Quasi-equivalence
  - of flagellin 1090
  - in oligomers 344
  - in virus coats 346
- Quenching, of fluorescence 1291
- Queosine
- Quercitin 1214
- Queuine 1621, 1388
- Queuosine (nucleoside Q) 234, 234*s*, 1621
- Quinaldic acid 1444, 1445*s*
- Quinate 1423*s*
- Quinic acid 1438
- Quinine 1447*s*
- Quinol oxidase 1025
- Quinolinate 1445*s*
  - from aspartate 1446
- Quinolinic acid 1446, 1798
- Quinone(s) **815 – 822**. *See also* Ubiquinone and Plastoquinone
  - as electron carriers 819, 1019 – 1030, 1310 – 1314
- Quinone methide 1433*s*
- Quinone Q_A 1312
- Quinonoid-carbanionic intermediate 741, 744*s*
  - from pyridoxal phosphate 743, 744*s*
- Quinoproteins 815 – 818
  - copper 815
- Quisqualate 1786*s*
  - receptors 1786
- Quorum sensing 1891

R

- Rab proteins 559, 1781
- Rabies 247
- Rabphilin 1781
- Racemase(s) 284
  - amino acid 691, 692, 741– 743
  - aspartate 741
  - mandelate 691, 692
  - PLP-dependent 741– 743
- Racemization
  - of amino acids 284, 741
  - of lens proteins 85
- Rac protein 559
- Radial projection
  - thick filaments 1102
- Radicals, free. *See* Free radicals
- Radioautogram (radioautograph, autoradiograph) 111
- Radioautography 110, 1474
- Radioimmunoassays (RIA) 110, 1849
- Radioisotope 258
- Radiolaria 18
- Radius
  - collision 462
  - covalent **40, 41**
  - van der Waals 40
- Radixin 1119
- Raf 578, 579
- Raffinose 169, 1141
- Raleigh scattering 1278
- Ramachandran plot 60, 61
- Raman spectra **1276 – 1280**
  - of 1-methyluracil 1278
  - of proteins 1279
- Ran 1503
- Rancidity 1204 – 1206
- Random coil conformation 69
- Random order of binding 464
- Rapamycin 488, 1217, 1691
- Rapid equilibrium assumption 467
- Rapid photometric methods 468
- Rapid uptake mode 1047
- Rapsyn 1781
- Ras **577– 579**
  - human c-H ras 560*s*
  - oncogenes (*ras*) 572, 576
  - P loop 559
  - protein superfamily 558
- Rasmussen's encephalitis 1865
- Rates
  - of approach to equilibrium 458
  - of displacement reactions **589**, **590**
  - of enzymatic reactions **455 – 497**. *See also* Kinetics
  - of substrate binding 463
- Rate constants 455
  - apparent first order 458
  - bimolecular 458, 462
  - prediction by Hammett equation 308, 309
- Rate equation for enzymes 455
- Rat liver, composition of 31
- RCC1 (regulator of chromatin condensation) 1503
- Reaction(s)
  - light-induced, ultrafast 1297
- Reaction centers 1299 – 1301, 1310 – 1318
  - kinetics, equation 1312
  - photochemistry 1310
  - properties of, table 1313

  - proteins, structure 1311
- Reaction coordinate 482
- Reaction rates, diffusion controlled limit 462
- Reaction types in metabolism, 530
  - table of **526**, **527**
- Reading frames in nucleic acids **236**, **237**
- Rearrangement reactions 527, 530, 712
- RecA protein 219, 1551, 1566, 1567
  - structures of 1566
- RecBCD pathway
  - DNA recombination 1566
- Receptor(s) 1, 553 – 563. *See also* Specific receptors
  - acetylcholine 422
  - alpha adrenergic. *See* α-Adrenergic receptors
  - autophosphorylation of 562
  - bacterial, for aspartate 561
  - beta adrenergic. *See* β-Adrenergic receptors
  - coagonist 1787
  - on cell surfaces 479
  - down regulation of **571**
  - for hormones **553 – 563**
  - human adrenergic 555
  - LDL 1185
  - for surface proteins 407
  - seven-helix structures 555
  - scavenger 1185
  - VLDL 1185
- Receptor theory 479
- RecG 1568
- Reciprocal plots
  - for kinetics of enzymes 460, 465, 472
- Recognition domains in proteins 367
- Recombinase 1566
- Recombination
  - of DNA 1564
  - mechanisms **1564 – 1568**
  - nonreciprocal 1568
  - site-specific 1564, **1570 – 1573**
  - unequal crossing-over 1568
- Recombination activating genes 1861
- Recoverin 1231, 1331
- Rectus configuration 42
- Red blood cell. *See* Erythrocytes
- Redox buffer 1035
  - system 980
- Redox potential
  - within cells 767
  - of flavoproteins 767
- Reduced flavins, reactions with oxygen 794
- Reduced oxygen compounds, biological effects 1072
- Reducing agents for biosynthesis **978 – 982**
- Reducing equivalents from citric acid cycle 515
- Reducing powers of redox couples 300
- Reduction potential(s)
  - definition of 300
  - for sulfate 1056
  - table, 301
- Reductive carboxylation system 984
- Reductive pentose phosphate pathway 963, 984
- Reductive tricarboxylic acid cycle 982
- Reelin 1903
- Reference electrode 300
- Refinement of X-ray structure 136
- Refolding of proteins 82
- Refractive index 1275
- Refsum disease 943
- Regenerating substrate
  - for citric acid cycle 950 – 952
  - for pentose phosphate cycle 964
- Regulation
  - of aspartate carbamoyltransferase 540
  - control elements in 535 – 537
  - of enzymatic activity **535 – 581**
  - pacemaker enzymes 535 – 537
- Regulators of G-protein signaling (RGS) 559
- Regulatory cascade(s) 566, 567
  - monocyclic (scheme) 567
- Regulatory subunits 348, 540, 541
- Regulon 1612, 1613
- Relative molecular mass (M_r)
  - determination of **108 – 115**
  - estimation by gel filtration 109
  - estimation by PAGE 109
- Relaxation methods 468
- Relaxation time
  - approach to equilibrium 458, 468
  - first order reactions 457
- Relaxin 1748
- Renin 621, 1261
- Renin-angiotensin-aldosterone system
  - blood pressure regulation 1261
- Renin-angiotensin system 1254, 1261
- Reoviruses 248. *See also* front cover
- Repair systems 16
- Replacement vectors 1496
- Replica plating 1478
- Replication
  - of DNA. *See* DNA replication
  - editing 1580
  - fidelity of 1579
  - polymerases role 1579
  - of RNA bacteriophages 1623, 1624
- Replication cycle. *See* Cell cycle
- Replication factor C (RFC) 1562
- Replication protein A (RPA) 1562
- Replicative senescence 1906
- Replicons 1561
- Replisome 1544
  - structure 1556
- Reporter gene products 1497
- Repression of enzyme synthesis 536, 538, 539
- Repressor(s) 76, 239, 1611
  - AraC 1613
  - Arc 1611
  - BirA (biotin) 1612
  - Cro 1611, 1622
  - of gene expression 539
  - ile 1612
  - lac 1611
  - lambda 1611, 1622
  - purine 1612
  - TetR 1612
  - of trehalose biosynthesis 1141
  - trp 1611
  - tyr 1612
- Reserpine 1447*s*, 1448, 1789
- Reservatrol 1439*s*
- Resiniferatoxin 1800
- Resistin 1005
- Resolution (res) sites 1575
- Resolvase 229, 1568
  - Tn3 219
  - gene 1575
- Resolvase/invertase family 1572
- Resonance **45**, **46**
  - energy values, table 299
  - in esters and amides 608
- Resonance Raman spectroscopy 1279
  - of bacteriorhodopsin 1279
- Respiration 300
  - aerobic 1013
  - anaerobic 1013, 1054 – 1057
  - history 1022
  - inorganic compounds as substrates 1051– 1054
- Respiratory burst 1072 – 1074
  - mechanism of 1072
  - of neutrophils (to kill bacteria) 1072
- Respiratory chain 783. *See also* Electron transport chain
- Respiratory control
  - of electron transport 1033
  - of phosphorylation 1033
- Respiratory distress syndrome 386
- Respiratory inhibitors, table 1021
- Response elements 1633, 1634
  - cyclic AMP 1633
  - iron responsive element (IRE) 1635
  - serum 1634
- Response regulator 1094
- Resting potential of cell membrane 400
- Restriction endonuclease(s) 249, 653, 1486
  - table 1487
  - types 1486
- Restriction enzyme mapping 1488
- Restriction fragment length polymorphism (**RFLP**) 1500
- Restriction fragments 1500
- Restriction maps 260
- Restrictocin 1686
- Resurrection plant 168, 1143
- Retention signal 521
- *R*-Reticuline 1443, 1444*s*
- Retina 1332, 1792
  - diseases of 1332, 1333
  - human, pigments 1328
- Retinal 1241, 1243, 1325, 1329*s*
- *cis*-Retinal 1241*s*
- Retinal cycle
  - of mammalian rod cells 1332
- Retinitis pigmentosa 1024, 1332
- Retinoblastoma 1586
- Retinoblastoma protein (Rb) 574, 580, 1907
- Retinochrome 1332
- Retinoic acid 1241, 1243, 1757
- *trans*-Retinoic acid 1241*s*
- Retinoids 1742
- Retinol 1239*s*, 1241, 1242. *See also* Vitamin A
  - reactions of 1332
- Retinol-binding protein 58, 1241
- Retinoyl β-glucuronide 1241, 1241*s*
- Retinyl esters
  - of fatty acids, 1241*s*
- Retrieval signal KDEL 521
- Retrograde axonal transport 1119
- Retroposon 1538
- Retrotransposons 1498, 1577
- Retrovirus 248, 1498, 1651
  - forms of 1653
  - gene expression figure 1653
  - replication of, figure 1652
- Reverse electron flow 1034, 1035
- Reverse electron transport 1052
- Reverse transcriptase 248, 257, 657, 1544, 1548, 1651
- Reversed phase columns 103
- Reversible chemical reactions 284
  - kinetics of 458
- Rhabdoviruses 247
- Rhamnogalacturonan 177, 1150

Rhamnose (Rha) 165*s*, 180, 1138
Rheumatoid arthritis 627, 1211, 1855, 1865
Rhinovirus 247
    icosahedral 344*s*
Rhizobium
    lipopolysaccharides of 1161
Rhizopoda 18
Rhodanese 1409
Rhodophyta 22
Rhodopsin 1324, 1329, 1330
    absorption spectrum 1326
    schematic drawing 1325
    sensory 1335
    topology 1326
Rhodopsin kinase 1331
Rho family of G proteins 1148
Rho protein 558
Ribavirin 1655*s*
Ribbon drawings (example) 64
Ribitol, conformation 44
Ribitolteichoic acids 431*s*
Riboflavin 721, 766, 783*s*, 1463*s*
    absorption spectrum 781
    biosynthesis of 1462, 1463
    fluorescence 783
    light sensitivity 783
    nutritional requirement 756, 783
    properties of 783
Riboflavin 5'-phosphate (flavin mononucleotide, FMN) 513, 765, 780, 781*s*
Ribokinase 1130
Ribonuclease 264, **647**, **648**, 1552
    fungal 648
    pancreatic, diffusion constant of 461
    refolding of 82
Ribonuclease A
    in-line mechanisms 647
    mechanism of catalysis 647
    NMR spectra of 647
    pK_a values of 647
    reaction mechanism 648
    sequence of 647
    X-ray structural studies 647
Ribonuclease H 1548, 1563
Ribonuclease HI 1551
Ribonuclease P 649 – 652
Ribonuclease S 647
Ribonucleic acid. *See* RNA
Ribonucleoprotein domain 244
Ribonucleotide reductase 863 – 865
    active site 865
    cobalamin dependent 871
    enzyme-activated inhibitors 864
Ribose 1, 5-bisphosphate
    in brain 545
Ribose 5-phosphate
    in biosynthesis 965
Ribose 5-phosphate isomerase 693
D-Ribose (Rib) 162, 163*s*, 200*s*
Ribosomal proteins 1677– 1684
    table 1677
Ribosomal RNA 1673 – 1678
    processing of 1619, 1620, 1638 – 1641
Ribosomal RNA genes
    transcription of, micrograph 1639
Ribosomal subunits 1673
    30 S and 50 S 1673
Ribosome(s) 3, 5, 11, 233 – 235, **1669 – 1729.** *See also* RNA
    active sites in 1686, 1687
    assembly map 1682
    assembly of 1681– 1684
    bacterial, structure of 1670
    chemical activity 1677
    composition of, table 1672
    decoding center 1672
    discovery 1474
    eukaryotic, composition 1672
    eukaryotic proteins 1684
    exit (E) site 1673
    formation, figure 1638
    GTPase activating center 1672, 1687
    peptidyltransferase site 1687
    programmed 1700
    recycling factor (RRF) 1709
    regulatory mechanisms 1715 – 1720
    70 S 1673
    structure of 233, 234, 1673 – 1687
    synthesis rate 1708
    three-dimensional structure 1671, 1682
Ribosylthymidine (Thd) 203
Ribosylthymidine 5'-phosphate (Thd-5'P) 203
Ribothymidine 234
    of tRNA 231
Ribothymidylic acid residue 813
Ribozyme(s) 239, **649 – 652**, 1646
    artificial 652
    hairpin 649, 651*s*
    hammerhead 649, 651
    leadzyme, 651*s*
Ribulose 164*s*
Ribulose bisphosphate 707*s*, 984
Ribulose bisphosphate carboxylase (Rubisco) 549, **706 – 710**, 708*s*, 986, 1319
    carbamoylation of 707
    mechanism, scheme 709
    spinach, active site 708
Ribulose monophosphate cycle 985
Ribulose 5-phosphate 964*s*, 984, 986*s*
Rice, genome 12, 1512
    genetic engineering 1240, 1518
Ricin 186, 1679, 1685, 1686
Ricinoleic acid 381, 381*s*
Ricinoleoyl-CoA 1188*s*
Rickets 421, 721, 1257
    X-linked hypophosphatemic 625
Rickettsias 7
Rieske iron-sulfur protein 860, 1027, 1029
Rifampicin 1618
Rifamycin 1217, 1616, 1618
Right-left axes 1899
Ring conformations 332 – 337
    boat 598
    half-chair 598
Ring current 139
Ring formation, terminating chain growth 992
Ritalin 1796, 1809*s*
Riubisco activase 707
RNA 5, **230 – 236**, 1474. *See also* Transcription
    acceptor-TψC-minihelix 1692
    alternative splicing factor 1647
    antisense 237, 1518, 1559, 1614
    aptamers 235
    in bacteria, transcription of 1603
    bacterial, processing of 1619
    base composition of 239
    binding proteins 243
    capping 1624
    conformational properties of 1638
    content in tissues 31
    density gradient centrifugation 101
    double-stranded 1508, 1640, 1847, 1869
    editing of 1624, 1648, 1649
    electrophoresis 249
    extraction of 249
    forms of, table 232
    Gibbs energy of formation 210
    guide (gRNA) 1560, 1615, 1620, 1648
    heterogeneous nuclear (hnRNA) 1624
    hydrolysis of 249, 251
    long terminal repeats 1651
    loops and turns 210*s*, 230
    methylation 1624
    miRNAs 1640
    modification of 1648, 1649
    polyadenylation 1624
    primers 1544
    pseudoknot 230
        figure 231
    rate of synthesis 1607
    ribosomal 1619. *See* rRNA
    self-splicing 649, 650
        figure 1642
    sequence determination 260 – 266
    snoRNAs (small nucleolar) 1640
    snRNAs (small nuclear) 1640
    splicing factors 1647
    splicing of 1642 – 1648
        figure 1647
    stem-loop structures 1638
    synthesis, control of 1608
    three-dimensional structure 1679
    transcription of, **1603 – 1657**
    transfer. *See* tRNA
    viral, in bacteria 1622
    in X-chromosome inactivation 1895
RNA bacteriophages
    replication of 1623
RNA interference 1640, 1869
RNA ligase 1646
RNA polymerase(s) (RNAPs) 657, 1544, 1547
    bacterial 1607
    binding sites 1631
    eukaryotic 1626, 1627*s*
    sigma subunit 1607
    transcription bubble 1609
    zinc in 680
RNA-protein interactions 1681– 1684
RNA viruses 1714, 1715
    replication of 1650 – 1657
RNase H 1651
RNase P 1620
RNase T1 648
RNA world 235
Rocky Mountain spotted fever 7
Rod cells of eye 390, 1324
    diagram of 1325
    retinal cycle of 1332
Rod-cone dystrophy 1332
Root growth hormones 1762
Root hairs, dimensions of 30
Roseoflavin 788, 789*s*
Rossmann fold. *See* Nucleotide-binding domain
Rotamases 488
Rotary diffusion constant 463
Rotation of molecules 462, 463
Rotational barrier 44
Rotational catalysis
    ATP synthase 1044
Rotational energy of molecules 1276
Rotational strength 1286
Rotenone 1021*s*, 1027
    inhibition of electron transport 1020

- Rotifers 24, 25
- Roundworms (Nematoda) 24. *See also* *Caenorhabditis elegans*
- Rous sarcoma virus (RSV) 1651
- Rowing hypothesis, of muscle contraction 1104, 1111
- R*p* (phosphorylation state ratio) 303, 1039
- R plasmids 1559
- rRNA (ribosomal RNA) 5, 230, 233*s*, 1673 – 1677
  - chemical modification 1673
  - comparative sequence analysis 1676
  - crosslinking 1673
  - gene clusters 1539
  - genes for 1539
  - phylogenetic comparisons 1676
  - processing of 1640
  - pseudogenes 1539
  - 5S, processing of 1641, 1642
  - 5S, three-dimensional structure 1680
  - 16S, central domain 1681
  - 16S, 3' domain 1684
  - 16S, 5' domain 1681
  - 16S, secondary structure model 1674
  - 23S 1684
  - structural domains of 5S 1677
  - structural domains of 16S 1676
  - structural domains of 23S 1676
  - structure of 23S – 28S 1678
  - three-dimensional structure 1674
- *RS* notation
  - for configuration **42**
  - for trigonal centers 480
- Rubber 1231
- Rubber elongation factor 1231
- Rubella 247
- Rubidium 31
- Rubisco. *See* Ribulose bisphosphate carboxylase
- Rubredoxin(s) 858*s*, 859, 1068
  - Fe-S cluster 857*s*
- Ruhemann's purple 121*s*
- Rumen
  - propionic acid bacteria in 970
  - fermentations, products of 947
- Ruminant animals, propionate as energy source 947
- Rusticyanin 883
- Rusts 20
- Rutin 1214
- RuvC 1568
- Ryanodine receptors 314, 422, 564, 1115
  - cryo-electron micrograph 1114

S

- Saccharin 1800*s*
- *Saccharomyces cerevisiae*
  - figure 20
  - genome 12, 20
  - mating type 1576, 1880
- Saccharopine 1385, 1386*s*
- SAGA HAT 1626
- Salicin 1440, 1440*s*
- Salicylic acid 1438, 1439*s*, 1762
- *Salmonella* 7, 1880
  - PEP carboxylase in 711
  - serotypes 180
- Salt bridges 47
  - of hemoglobin 357
- Salting in/out of proteins 101
- Salvarsan 597*s*
- SAM. *See* *S*-Adenosylmethionine (Adomet)
- Samuelsson, Bergt I. 84
- Sandhoff disease 1170, 1171
- β Sandwich 64
- Sanfilippo disease 1170, 1171
- Sanger, Frederick 83, 84, 118, 262
- Saponins 1265
- α-Sarcin 1686
- Sarcin/ricin (SR domain)
  - of 23S RNA 1679
- Sarcodina 18
- Sarcoglycans 1112
- Sarcolemma 1096
- Sarcomas 573
- Sarcomere
  - model 1100
  - structure of 1097
- Sarcoplasm 1096
- Sarcoplasmic reticulum 1098, 1114
- Sarcosine 1399
- Sarcosine oxidase 1399
- Sarcosomes 1096
- Satellite tobacco necrosis virus 247
- Saturation in binding 326
- Saturation curve(s) 326 – 331, 458 – 461, 475 – 477
  - hyperbolic 326
  - sigmoidal 326
- Saturation fraction
  - definition of 326
  - for cooperative processes 331
- Savants 1805
- Saxitoxin 1769, 1770*s*, 1775
- Scaffolding proteins of viruses 1559
- Scanning mutagenesis 402
- Scanning tunneling microscope 131
- Scatchard plot 327
- Scavenger receptor 1185, 1251
- *Scenedesmus* 22
- Schiff base(s) 121, 752, 1328, 1329
  - in aminotransferases 750
  - of pyridoxal phosphate 742
  - of pyridoxal phosphate, 740*s*, 742, 748, 750 – 753
  - reduction of
    - by borohydride 124, 679
- Schistosoma 24
- Schistosomes 1866
  - protease inhibitors 622
- Schistosomiasis 24, 222
- Schizophrenia 1810, 1811
- Schizophyllan 170, 174
- *Schizosaccharomyces pombe* 1115, 1891
- Schwann cells 390, 1763
- Sclerenchyma 30
- Scleroderma 1538, 1641, 1865
- Scorpion(s) 24
  - toxin(s) 1770*s*, 1775
- Scrapie 248, 1718
- Screw-sense isomer of ATP complexes 643
- Scurvy 721, 1066
- SDS electrophoresis 111
- SDS-polyacrylamide gel electrophoresis (SDS-PAGE) 106
  - of erythrocyte ghosts, figure 404
- Sea anemone toxins 1775
- Sea cucumbers 25
- Sea snake venom 1775
- Sea squirts. *See* Tunicates
- Sea urchin 25
  - genome 12
- Seasonable affective disorder (SAD) 1810
- Second Law of Thermodynamics 284, 285
- Second messengers 553
  - calcium ions 555
  - cyclic ADP-ribose 555
  - cyclic AMP 555
  - cyclic GMP 555
  - diadenosine 5' tetraphosphate 555
  - guanosine 5' diphosphate, 3-diphosphate (ppGpp) 555
  - inositol-1,4,5 trisphosphate 555
  - for insulin 569 – 571
  - Mn^{2+} 555
  - table of 555
- Second order reactions 458
- Secondary immune response 1833
- Secondary kinetic isotope effect 592, 600
  - on fumarate hydratase 684
- Secondary plots for kinetics of multisubstrate enzymes 465
- Secondary structure 63
  - of a protein 63
- Secretase(s) 1813
- γ-Secretase 1900
- Secretin 1749
- Secretion gene (Se) 185
- Secretion granules 10, 14
- Sedimentation constant 100, 109
  - of nucleic acids 112
  - of proteins 112
- Sedimentation equilibrium 108
- Sedimentation velocity 108
- Sedoheptulose 164*s*, 1320
- Sedoheptulose 7-phosphate 964, 965
- Seeds 30, 1904, 1905
- Segmentation genes 1900
- *S*-Selanylcysteine 893
- Selectins 187, 188, 1883
- Selector genes 1900
- Selenium 1894
  - in enzymes 822 – 827
  - metabolism of 827
  - proteins, incorporation into 827, 1711, 1712
  - in proteins, table 825
- Selenocysteine 237, 823, 824
  - in hydrogenases 878
- Selenocysteine lyase 746
- Selenoprotein P 824
- Selenoprotein W 824
- Self-assembly of macromolecules 332
- Self-splicing RNA 1643
- Seliwanoff's reagent 167
- Semaphorins 1903
- Semiconductor nanocrystals 1294
- Semidehydroascorbate 1064
- Seminalplasmin 1750
- Semiquinone radicals of flavins 792, 794
- Sensitivity coefficient 537
- Sensor kinase/response regulator systems 545
- Sensor-response regulator pairs 1614
- Sensory neuron (afferent fiber) 1766
- Sensory rhodopsins (phoborhodopsin) 1335
- Separations of compounds **98 – 108**
  - by chromatography 100 – 106
  - by partition 102, 103
- Sephadex, gel networks of 175
- Sepiapterin 803, 804*s*, 1460
- Septal ring 1890
- Septate desmosomes 29

Septum 5
Sequence(s)
    of amino acids 57, 118, 119
    of nucleic acids 260 – 265, 1507– 1517
Sequencing primers 260
D-Serine 1787, 1793
L-Serine (Ser, S) 52, 53*s*, 808, **1397– 1405**
    biosynthetic pathways from 1397, 1398
    metabolism of, scheme 1398
    as nucleophile 609, 610
    pK_a value of 487
Serine carboxypeptidases 610
Serine dehydratases 688, 742
Serine hydroxymethyltransferase 745, 1398
    quinonoid intermediate 750
Serine pathway
    one-carbon compound assimilation, scheme 986
Serine proteases **609 – 618**
    acyl-enzyme intermediates 610
    carboxypeptidase 609, 625
    catalytic cycle 613 – 616
    catalytic triad 614
    chymotrypsin 66, 457, 477, 609
    oxyanion hole 614, 615
    papain 618
    pepsin 456, 621
    specificities 117
    thrombin 632
    trypsin 66, 116, 457, 609
Serine *O*-sulfate 739
Serotonin (5-hydroxytryptamine) 744, 1444, 1447*s*, 1782, 1792, 1793, 1793*s*, 1803, 1809
Serotonin *N*-acetyltransferase 1807
Serotonin receptors 1793
Serpins 629
Serum albumin 58
Serum amyloid A 1850
Serum response factor 578
Sesquiterpenes 1234 – 1236
Severe combined immunodeficiency disease 1867
Sex determination 1893
Sex factor 249. *See also* F agent
Sex hormones 1252 – 1260
SH2 and SH3 domains 367, 580
SH2 • phosphopeptide complex 368*s*
Sharp, Phillip A. 84
Shc 579
    as adaptor protein 570, 577
    definition of 577
β Sheet, properties of 64
Shells, mineral 440, 441
Shiga toxin 1685
Shigella 7
Shikimate 1422*s*, 1423*s*, 1424, 1424*s*
Shikimate pathway **1421– 1423**, 1438
    enzymes of 1423, 1424
    scheme 1422
Shikimate 3-phosphate 687*s*, 1424*s*
Shikimic acid 1421*s*
Shine–Dalgarno (SD) sequence 1699, 1700
Shuttle vehicles 1497
Sialic acids 165*s*, 183, 546, 1157. *See also N-*Acetylneuraminic acid
    biosynthesis of 1130
Sialidase 186
Sialidosis 1171
Sialuria 1172
Sickle cell disease 16, 359, 360
Side chains packing 75
Siderophilins 840
Siderophore(s) 838, 838*s*
    chelate effect 310
    receptors 839
Sieve plates 30
Signal peptidase 522, 620
Signal recognition particle (SRP) 520, 1721
Signal sequences 519
Signal transduction (signaling) 1, 427, 553 – 571
Signature sequences 1550
Silaffins 178, 1381
Silencing of genes 1624, 1625, 1741– 1801, 1881
Silica gel, use in separation 103
Silica skeleton
    of radiolarian 178
Silicon
    in biological cross-linking 178
    biological functions 178
Silicon dioxide in “shells” of diatoms 21
Silk 67
    β structure in 63
    fibroin, composition of 59
Simian virus 40 (SV40) 244
Sin3 1626
Sinapate 1439*s*
Single carbon compounds
    in metabolism 808 – 813, 985, 986
    table, according to oxidation state 808
Single molecules, observation of 1108 – 1110
Single-strand binding protein 1549
Singlet oxygen 856, 1074
Singlet excited states 1282
Sinigrin 1142*s*
Sinister configuration 42
Siomycin 1691
Sirenin 1235*s*
Siroheme 794, 845, 845*s*, 861, 1056, 1367
sis, *v-sis* 571, 572 576
Site-directed mutagenesis. *See* Mutation: directed
Sitostanol 1249
Sitosterol 1246, 1248*s*
Sizes of atoms, table **41**
Sjögren syndrome 1865
Skewness 1282
Skin **439**, **440**
    color 1435
Skunk cabbage 1048
S-layers of bacteria 431
Sleep, rapid eye movement (REM) 1808
Sliding clamp 1549
Slime molds 1892. *See also Dictyostelium discoideum*
Slow reacting substance of anaphylaxis (SRS-A) 1210
Slow-twitch muscle fibers 1096
Sly syndrome (MPS VII) 1171
SMADs 1632, 1634, 1899
Small nuclear RNAs (snRNAs) 1640, 1641
Small nucleolar RNAs (snoRNAs) 1640, 1641
Smell 1798, 1799
Smith degradation of carbohydrates 190
Smoluchowski equation 461, 462
Smooth (involuntary) muscle 1098, 1767
    fiber, structure 1097
    latch state 1117
    regulation of 1117
Sm proteins 1640
SNAPS 521
SNAP-25 1779
SNARE proteins 521, 1778, 1779
Snell, Esmond 737
snRNP particles 1640
Sodium arsanilate 597*s*
Sodium borohydride
    action on aldoses and ketoses 167
    radioactive 679
Sodium borotritide 190
Sodium dithionite 767
Sodium dodecyl sulfate 82*s*
Sodium ion 31, **206**
    concentration in cells 206
Sodium ion channel(s) 421, 1769, 1781
    of electric eel 1769
    gating current 1771
    image of 1771
    map of 1771
Sodium-potassium pump 423
Solanesyl diphosphate 1231
Solanidine 1265, 1265*s*
Solubility
    of proteins 101
    separation by 101
Somatic cell fusion 1508
Somatic motor system 1767
Somatomedins. *See* Insulinlike growth factors (IGF)
Somatostatin 1746, 1747*s*, 1795
Somatostatin cell 1046
Somatotropin 1745
Sonication (sonic oscillation) 98
Sonic hedgehog protein 1903
Sorbitol 164, 196, 774, 1131
    pathway 1129
Sorbose 164*s*
Sorghum 1442
Sos 579
    definition of 577
    as exchange protein 578
SOS signal 1583, 1622
Southern blots 260
Southern corn leaf blight 1904
SoxR protein 1075
Soybean lectin 186
Sp1 1628
Space groups 133, 332
Sparsomycin 1689*s*
Specific activity of an enzyme 457
Specific immunotherapy 1863
Specific rotation 1286
Specificity of enzymatic action 478 – 482
Spectinomycin 1691
Spectra
    derivative 1285
    difference 1285
Spectrin 313, 405
Spectrophotometer, dual wavelength 1020
Spectroscopy
    energy-selective 1293
    matrix isolation 1293
Spemann’s organizer 1899
Sperm-activating peptide (speract) 1753, 1760, 1895
Sperm receptors 187, 1895
Spermatozoa
    development 1894
    metabolism of 1131
Spermidine 204, 363, 1380*s*
    interaction with DNA 218
Spermine 1380*s*
    biosynthesis 1381
    interaction with DNA 218
S-phase promoting factor (SPF) 1563

S-phase of cell cycle 579
Spherocytosis 405
Spheroplast 6
Sphinganine (dihydrosphingosine) 388*s*, 1202, 1203*s*
Sphingoglycolipids 388
Sphingolipid(s) **388**, **389**, 1202, 1758
  metabolism, scheme 1203
  structures of 389
  synthesis, scheme 1203
Sphingolipidoses 1170 – 1172
  table 1171
Sphingomyelin(s) 389*s*, 1202, 1203*s*
Sphingomyelinase 1171
Sphingophospholipids 383, 388
Sphingosine (sphingenine) 388*s*, 1203*s*, 1397
  biosynthesis of 745
Sphingosine 1-phosphate 578
Sphingosylphosphocholine 422
Spicules 23
Spiders 24, 67
Spin decoupling. *See* NMR spectroscopy
Spinach (*Spinacia oleracea*)
  composition of 31
  ribulose bisphosphate carboxylase in 706
Spindle 1121, 1502
  assembly checkpoint 580, 1503
Spin labels 398, 399
Spinobulbar muscular atrophy (SBMA) 1516
Spinocerebellar ataxias (SCA) 1516
Spin-orbit coupling 398
Spin-spin interaction. *See* NMR spectroscopy
β Spiral 74
Spiramycin 1690
Spirillum, definition of 6
Spirochetes 7
Spirogyra 21
Spliceosome 1640, 1647
Sponges 23
  aggregation factor 1155
  species-specific reaggregation 187
Spongin 23
Spo0F 1614
Spore(s), bacterial 8
Sporophyte form of fern 18
Sporophytic phase of plants 1904
Sporozoa 19
Sporulation of bacteria 1891
Squalamine 1266, 1266*s*
Squalene 1231, 1236*s*, 1237*s*, 1244
  conversion to cholesterol 1244
  cyclization to cycloartenol 1245
  cyclization to lanosterol 1245
  formation of 1236, 1237
Squalene 2,3-oxide 1244
Squid 24
v-src, *src* 572
Src, definition of 577
SRP receptor 520, 1721
SSB 1550
Stachyose 169, 1141
Standard electrode potential 300
Standard states 286
Starch 172, 173*s*
  biosynthesis of 1130, 1143 – 1146
Starch granules 10
  electron micrographs 172
  layered structure 173
Starch-hydrolyzing enzymes **605 – 608**
Starch-iodine complex 174*s*
  helical structure 173
Starfish 25
Startle disease 1789
STAT (signal transducer-activators of transcription) 1632, 1634, 1847
State, thermodynamic 281
Statherin 442
Statine (Sta) 623*s*
Stationary phase of growth curve 470
Statistical effects
  in binding 328, 329
Staufen protein 1898
Steady state 1, 289
  approximation 459
Stearic acid 380*s*, 381
Stearoyl ACP desaturase 863, 863*s*
Stellacyanin 883
Stem, section of, figure 31
Stem cells 1881, 1885 – 1888
  embryonic 1886
  hematopoietic 1886
  mesenchymal 1886
  neuronal 1886
  pluripotent 1833
  replication control 1887
Stentorin 1336
Stercobilin 1404
Stereochemical numbering system 480
Stereochemical studies
  using chiral acetates 704, 705
Stereochemistry
  of chorismate formation 689
  of enzymatic reactions 111, **478 – 481**
  of phosphotransferases **642**, **643**
  of pyridoxal phosphate-requiring enzymes **747– 749**
Stereocilia 370, 1801
Stereoelectronic control 598
Stereoelectronic effect
  in enzymatic catalysis 495
  in proteases 616
Stereospecificity
  of dehydrogenases 768 – 775
  prochiral centers 478
  Stern-Volmer equation 1291
Steroid(s) 1227
  antibiotic 1266
  compounds 1243 – 1252
  hormones 536, **1252 – 1265**, 1742
    biosynthesis, scheme 1256
  plant, biosynthesis of 1248
  toxic, structures 1265
Steroid receptor family 1263 – 1265
  proteins of, table 1264
Steroidogenic acute regulatory protein 1252
Sterol(s) **389**, 1243
  biosynthesis of **1244 – 1246**
  oxygenases in formation of 1057
Stickland reaction 1374
Stigma 20
Stigmasterol 1246
Stigmatellin 1029
Stilbenecarboxylic acid 1214, 1215*s*
Stilbenes 1217
STM image of glycogen 174
Stoichiometric dissociation constant 305
Stokes-Einstein equation 461
Stokes radius 109
Stomach oxyntic gland, diagram 1046
Stomach ulcer 187
Stomata 30
Stop (termination) codons 236, 237, 1709, 1714, 1715
Stopped flow technique 468
Strand exchange proteins 1566
Stratum corneum 439
Strecker degradation 121
Strength of binding. *See* Formation constants, Dissociation constants, Gibbs energy
Streptavidin 728
Streptidine 1139*s*
Streptococci, group A 1167
Streptogramins 1690
Streptolydigin 1618
Streptomycin 1139*s*, 1164, 1194, 1690, 1703
  biosynthesis of 1139
  inactivation of 1140
Stress-prone pigs 1115
Stress proteins 518, 1635, 1636
Striatum 1765
Stringent response 1608, 1715
Stroma 1301
Stromelysins 627, 1249
Strontium 19
Structural domains, table 367
β Structure 78
  hydrogen bonds in 61
  of proteins **61– 68**
  in small peptides 66
  twist in 63, 64
Strychnine 1788*s*, 1789
Stuart factor 632
Styrene, resonance energy 299
Suberin 1196, 1441
Submitochondrial particles 1020
Substance P 1750, 1795
Substituent constants
  in Hammett equation 308
  table 308
Substitution reactions. *See* Displacement reactions
Substrate
  definition of 332
  primary, for catalytic cycle 950
  regenerating, for catalytic cycle 950
Substrate binding
  rate of 463
  unequal 349, 350
Substrate cycles 567
  in biosynthesis 1000
Substrate level phosphorylation
  in fermentation 966
Substrate saturation
  of an enzyme 457
Substrate specificity
  of chymotrypsin 616
  of serine proteases 616
  of trypsin 617
Substrate surrogate 463
Subtilisin 610
  active site 612
Subunits in protein structures 74
Succinate 478*s*
  binding constants to metal ions 311
Succinate dehydrogenase (succinic dehydrogenase) 478, 766, 785
Succinate-ubiquinone oxidoreductase (Complex II) 1021, 1027
Succinic acid 99, 330. *See also* Succinate
Succinic semialdehyde 958
Succinimide
  racemization 85
Succinimidyl group 84
Succinylacetoacetate 1431*s*
O-Succinylbenzoate 1428, 1429*s*

Succinyl-CoA 970
  formation of GTP 950
Succinyl-CoA– acetoacetate CoA transferase
  ping-pong mechanism 662
Succinyl-CoA synthetase 661
Succinyl-CoA 516*s*
O-Succinylhomoserine 746*s*
  conversion to cystathionine 746
Sucrase (invertase) 1140
Sucrose 167, 168*s*
  biosynthesis 1140
  in dextran synthesis 1152
  as glucosyl donor 1152
  properties of 168, 169, 1140
  as transport sugar 1140
Sucrose-fructose exchange 595
Sucrose phosphorylase 595
Sugar(s) **161– 167**. *See also* Monosaccharides, Oligosaccharides
  catabolism of **960 – 966**
  configurations of 163
  covalent hydrate **162**
  in lipids 380
  metabolism of **1129 – 1140**
Sugar isomerases 526
Sugar nucleotide(s)
  interconversions, scheme 1129, 1130
Sugar phosphates
  interconversions 508
Sugar rearrangement system
  of pentose phosphate pathway 963
Sugar transporters
  fructose 416
  glucose 416
  SGLT1 (Na^+ dependent) 417
Sulfanilamide 473*s*, 720*s*
  acetylation of 720
Sulfatase(s) 526
Sulfate
  as inorganic oxidant 1051
  reduction potential for 1056
Sulfate esters
  in proteins 79
Sulfate-reducing bacteria 1056, 1057
Sulfatides 388
Sulfide
  oxidation by *Thiobacillus* 1052 – 1054
3-Sulfinylpyruvate 1408
Sulfite 1408
  oxidation by *Thiobacillus* 1052 – 1054
Sulfite oxidase 847
  family 890
Sulfite reductase(s) 794, 861, 861*s*, 1056
Sulfo derivative, active 975
Sulfogalactosylglycerolipid 387*s*
Sulfolactaldehyde 1138*s*
Sulfolipid 387, 1406
Sulfometuron methyl 735*s*
Sulfonamides 473, 473*s*
Sulfonolipids 388
Sulfonylureas 421, 1393
  receptor 421
6-Sulfoquinovose 165*s*, 1138*s*, 1406
  biosynthesis 1138
Sulfotransferases 526, 659
Sulfur
  chain molecules of 1053
  compounds of garlic 1408
  compounds of onions 1408
  compounds of skunks 1408
  metabolism 1406
  oxidation by *Thiobacillus* 1052
Sulfur-35, properties of 110
Sulfur displacements **659**
Sulfur-oxidizing bacteria 1052 – 1054
Sulfur-reducing bacteria 1056, 1057
Sulfuraphane 1409, 1409*s*
Sumatriptan 1793, 1793*s*
Summer, James B. 83
SUMO 1583, 1626
Superantigens 1857
Supercoiling, energy of 221
Supercritical fluid chromatography 103
Superhelix 70
Superhelix density in DNA 221
  effect of intercalation 223
Superoxide anion radical 794, 826, 850, 851, 1059, 1060, 1072 – 1074
  reaction with $FADH_2$ 529, 1060
  redox potential 1072
Superoxide dismutase(s) 866, 884, 1059, 1074
  in ALS 1075
  copper, zinc 884
  β-cylinder in 78
  inhibition by 1059
  mass spectrum 113, 114
  superoxide, damage from 1074
  zinc in 680
Suppressor genes 1480, 1481
Suprafacial transfer of protons 696
Surface plasmon resonance 122
Surfactant system of the lungs 386
Surfactin 54*s*
Sutherland, Earl W., Jr. 84
Suzukacillin 414
Svedburg, Theodor 83
Svedberg equation 109
Svedberg units 100
SW1/SNF 1626
Switch helix 1704
Symbiotic relationship 17
Symmetry 133, 134, 338 – 348
  cubic 342 – 344
  dihedral 338
  space groups 133
Sympathetic nervous system 1767
Synapse(s) 29, 1762, **1777– 1782**
  chemical changes in 1803
  cholinergic, inhibitors of 1783
  depolarization of membrane 1777
  depression (habituation) 1803
  diagram 1763
  electrotonic 29, 1765, 1782
  glutamatergic, 1778*s*
  immunological 1781
  potentiation (facilitation) 1803
  proteins of 1778
  retrograde messenger 1804
  schematic drawing 1764
  synaptic plasticity 1803
  vesicle cycle 1777, 1778
  vesicle proteins, table 1780
  vesicles 1779
Synaptic cleft 1763
Synaptic junctions, micrograph 1764
Synaptic knobs (boutons) 1763
Synaptic vesicles 1762
Synaptobrevin (VAMP) 1778, 1779
Synaptonemal complex 1505
Synaptophysin 1781
Synaptotagmin 1778, 1780
Synchrotron radiation 137
  as X-ray source 134
Syn conformation of nucleosides 211, 212
Syncytium 1898
Syndecan 1154
Synemin 370
Synpolydactyly 1516
Syntaxin 1778 – 1780
Synthesis, light directed 1298
Synthetic probe for nucleic acids 257
Syntropism 1423
α-Synuclein 1812
System, thermodynamic, definition of 281
Systemic acquired resistance 1440, 1762, 1869
Systemic amyloidoses 1719
Systemic lupus erythematosus 1864
Systemin(s) 1762

T

T4 DNA polymerase 1491
T4 endonuclease V 1582
T4 polynucleotide kinase 1491
Taft equation 309
  table of constants 308
Tagatose, 164*s*
Tailspike protein
  of bacteriophage 66*s*
Talin 406
Talose (Tal) 163*s*
Tamoxifen 575, 1263, 1263*s*
Tangier disease 1251
Tannins
  condensed 1438
  hydrolyzable 1438
T antigen 1562
TAP (transporter associated with antigen processing) 1723, 1856
Tapetum 680
  zinc in 680
Tapeworms 24
Taq polymerase 260
Tartaric acid 43*s*
  configuration of 43
Tartronic semialdehyde 959*s*, 960
Taste 1799
Tat pathway 1723
TATA-binding protein 1628, 1629
  ribbon drawing 1629
TATA sequence 1628, 1725
Tau protein 372, 1812, 1814
Taurine 1407, 1408, 1793
  chloramine from 1073
Taurocholic acid 1251, 1252*s*
Tautomerases 692
Tautomeric catalysis 490, 491
Tautomeric ratios 45, 305
Tautomerism **45**, **46**
  in amides 45
  in base pairing 211
  of cysteine 306
  equilibrium constant 45
  in imidazole group 46
  in 1-methyluracil 307
  microscopic dissociation constants 305
  of nucleic acid bases 204, 205, 211
  oxo-enol 45
  of papain 306
  of porphyrins 845
  of pyridoxine 305
Taxol 371*s*, 1236
Tay-Sachs disease 1170 – 1172

T cell. *See* Lymphocytes
T-cell receptors 1851
  gene organization 1860
Teeth 440, 441, **442**
  diagram of 442
Teichoic acids
  biosynthesis of 1166
  in cell walls 431
Teichuronic acids
  biosynthesis of 1166
  in cell walls 431
Telomerase 1548, 1560
Telomeres 208, 1538
  DNA in 227
  maintenance 1906
  synthesis, figure 1561
Telophase 1501
Temperature 283, 284
  units of 283
Temperature jump method 468
Tenascin 409, 1154
Tendons, development of 1901, 1902
Tensin 407
Tenuin 407
Terbium 1292
Terminal bars 29
Terminal deoxynucleotidyl transferase 1491, 1861
N-Terminal nucleophile hydrolases 620, 621
  autoactivation of 621
Terminal protein (phage) 1559
Terminase (viral) 1558
Termination factors (ribosomes) 1673
Termites, protozoa in 19
Terpenes 1227, **1232 – 1237**, 1443
  functions 1232
Terpin 1233*s*
α-Terpineol 1232, 1233*s*
Tertiary structure of a protein 59
TES buffer 99
Testes-determining factor 1893
Testican 1154
Testosterone 696, 1254 – 1260, 1256*s*
Tetanus toxin 1776
Tetrachlorodibenzo-*p*-dioxin 1404*s*
1,3,4,6-Tetrachloro-3α,6α-diphenylglycouril 126*s*
Tetracycline(s) 1164, 1213, 1216*s*, 1691
Tetrahedral intermediate
  in displacement reactions 609
  stereochemistry of formation 616
Tetrahedron 342
Tetrahydrobiopterin 1061*s*
Tetrahydrocannabinol 1235*s*, 1794*s*
Tetrahydrofolic acid (THF) 745, **802 – 813**, 803*s*,
  derivatives, scheme 806
Tetrahydroisoquinolines 1811
Tetrahydropicolinate 1385*s*
Tetrahymena 19
Tetrahymena ribozyme 650*s*
Tetrameric enzyme 342
Tetramethylammonium ion 1783*s*
Tetramethyl-*p*-phenylenediamine, oxidized 1025*s*
2,2,6,6-Tetramethylpiperidine-1-oxyl 398*s*
  EPR spectrum 398
Tetramethylsilane (TMS) 138
Tetranitromethane 126
Tetraploid 1506
Tetrapyrrole pigments, open
  formation of, scheme 1404
  structures of 1305
Tetrazolium salts 1025*s*
Tetrodotoxin 736, 1266, 1769, 1770*s*, 1775
TGF-β family 1903
Thalamus 1765
Thalassemias 1500, 1901
Thaumatin 1800, 1800*s*
Theileriosis 544
Thenoyltrifluoroacetone 1021*s*
Thermochemistry 282, 283
  standard states 286 – 289
Thermodynamics **281 – 289**
  base-pairing 209
  biochemical standard, new 297
  of electron transport 1034 – 1036
  first law of **282**, **283**
  irreversible 289
  of life processes 289
  nonequilibrium 289
  second law of **284**, **285**
  system 281
Thermodynamic barrier 492
Thermodynamic temperature scale 284, 285
Thermogenic tissues 1048, 1049
Thermogenin 1048
Thermolysin 625
  specificity 117
Thermostability
  of enzymes 85
Thermus (Taq) polymerase 1547
Thiamin (vitamin B$_1$). *See also* Thiamin diphosphate
  acid–base chemistry (scheme) 331
  adducts, 733*s*
  analogs 733
  biosynthesis of 736, **1460 – 1462**
  chemical properties of 730, 731
  cleavage by bisulfite 731
  cooperative binding of protons 331
  degradation by base 731
  nutritional requirement 756
  pK_a values of 330
  polyphosphates 736
  thiol form 730
  tricyclic form 730, 737
  yellow anion of 331
Thiamin coenzymes 730. *See also* Thiamin diphosphate
  in nerve action 736
  in transport of sodium 736
Thiamin diphosphate (Pyrophosphate) 511, 718, **730 – 737**, 767
  α cleavage 730, 730*s*
  catalytic mechanisms 731– 733
  enzymes, table 735
  enzymatic reactions requiring 734 – 736
  half-reactions, scheme 735
  minor tautomer 733
  pyrimidine, role of 733
  X-ray diffraction 733
Thiamin phosphate 1464*s*
  biosynthesis, scheme 1464
Thiazolidine 1410*s*
Thiazolium dipolar ion (ylid) 732, 733, 733*s*
Thick filaments
  of muscle 1096, 1097, 1103, 1104
  radial projection 1102
Thinking, chemistry of 1801–1814
Thin layer chromatography 103
  of flavins 103
  of maltooligosaccharides 189
  O-methylated glucoses 190
Thin-filaments. *See* Actin
Thiobacillus 1052
Thiobacillus denitrificans 1055
Thiobacillus ferrooxidans
  oxidation of Fe^{2+} 1054
Thiochrome 730, 730*s*
Thiocyanate 1408
Thiocysteine 1410, 1410*s*
Thioesters 559, 775
  displacement reactions of 608
  formation from glyceraldehyde-3-phosphate 677
  function of 700
  mechanism of formation 661, 662
Thioether 559
Thioglutathione 1410
Thioglycine 881
Thiohemiacetal in glyceraldehyde phosphate dehydrogenase 775
Thiol(s) 115
  addition to carbonyl group 677
  adduct, dehydrogenation of 766
  binding of mercury ions 125
Thiol–disulfide equilibria 549
Thiol–disulfide exchange 659
Thiol:disulfide oxidoreductases 786
Thiolases 699, 827, 940
  biosynthetic 699
Thiolate anions 779
Thioltransferases 549, 659, 786
Thioredoxin 522, 549, 659, 786*s*, 787, 864
  disulfide loop 786
  functions of 786
Thioredoxin reductase 785, 786, 787, 824
  mechanism of 791
Thiostrepton 1679, 1687, 1689*s*, 1691
  microccin 1703
Thiosulfate 1408
  oxidation by *Thiobacillus* 1052
Thiosulfate sulfurtransferase (rhodanese) 78, 1408
Thiotaurine 1410
Thiouridine 234
Three-dimensional structures of
  aconitase 689
  actin filament 1101
  adenylate kinase 655
  aldehyde oxido-reductase 891
  D-amino acid oxidase 791
  α-amylase, pancreatic 607
  aspartate aminotransferase 57, 135
    catalytic intermediates 752
  aspartate carbamyltransferase 348
  aspartate chemoreceptor 562
  bacteriophage P22 66
  bacteriorhodopsin 1327
  bc$_1$ complex 1029
  Bchl *b* molecules 1311
  cadherin 408
  calmodulin 317
  carbonic acid anhydrase I 679
  carboxypeptidase A 64
  catalase 853
  cholera toxin 333, 546
  chymotrypsin 611
  citrate synthase 702, 703
  conotoxin 1770, 1783
  cutinase 134
  cyclosporin 488
  cytochrome *c* 847
  cytochrome *c* oxidase 1031

  - cytochrome *c* peroxidase 849
  - dihydrofolate reductase 807
  - DNA 214, 223, 228, 229, 241
  - DNA complex
    - with methionine repressor 243
    - with netropsin 225
    - with Trp repressor protein 240
    - with zinc finger protein 242
  - EcoRI restriction endonuclease 1488
  - elastase 611
  - enolase 686
  - eotaxin-3 1754
  - F-actin 338
  - fatty acid-binding protein 1184
  - favin 64
  - ferredoxin 858
  - ferritin 343, 841
  - flagella 1090
  - glucoamylase 77, 608
  - glutathione reductase 785
  - glyceraldehyde phosphate dehydrogenase 63
  - glycine *N*-methyltransferase 593
  - glycogen phosphorylase 605
  - Grb2 580
  - HIV-1 protease 624
  - hydrogenase 879
  - IgG1 molecule 1840
  - immunoglobulin 1838
  - insecticyanin 1184
  - insulin 347
  - insulin receptor 570
  - iron superoxide dismutase 867
  - K^+ channel 413
  - kringle 2 motif 368
  - lactoferrin 840
  - light-harvesting protein 1309
  - lysozyme 599
  - methanol dehydrogenase 816
  - methionine synthase 874
  - methylmalonyl-CoA mutase 874
  - myohemerythrin 71
  - nitrogenase 1361
  - pancreatic polypeptide 1749
  - P pilus 337
  - pectate lyase C 686
  - phycobiliprotein 1306
  - plastocyanin 65
  - pyridoxal phosphate enzyme 750
  - pyruvate dehydrogenase 797
  - pyruvate oxidase 734
  - reaction center proteins 1311
  - rhinovirus 344
  - ribosome, 30S subunit 1670, 1683
  - ribosome, 50S subunit 1670, 1679
  - ribosome, 70S 1669, 1671, 1682
  - ribulose bisphosphate carboxylase 708
  - rRNA, 5S 1680
  - rRNA, 16S 1675
  - rRNA, 23S, 1678
    - with proteins 1682
  - rRNA loops 1679
  - rubisco 708
  - rubredoxin 858
  - scorpion toxin 1770
  - stearoyl-acyl carrier protein desaturase 863
  - subtilisin 612
  - sulfite reductase 861
  - tissue factor 633
  - trimethylamine dehydrogenase 784
  - triose phosphate isomerase 77, 694
  - tRNA 233
  - trypsin 611
  - trypsinogen 615
  - tubulin 372
  - virus fd protein sheath 335
  - virus ϕX174 365
  - yeast RNA polymerase II 1627
- L-*allo*-Threonine 53*s*
- Threonine (Thr, T) 53*s*, 539, 1383, 1384*s*
  - biosynthesis 540, 1391
  - cleavage to glycine and acetaldehyde 745, 1391
  - configuration 43
- Threonine dehydratase (deaminase) 688, 742
  - biodegradative 1391
  - biosynthetic 1391
- Threonine synthase 746, 1383
- Threose 163*s*
- Thrombin 610, 632
- Thrombocytes 26
- Thrombomodulin 634
- Thrombopoetin 1901
- Thrombospondin 409
- Thromboxanes 1208, 1206*s*
- α-Thujone 1233*s*, 1789
- Thylakoids 1301
- Thylakoid vesicles, pH of 1039
- Thymidine (dThd) 203
- Thymidylate 1452
- Thymidylate synthase 810, 811
  - mechanism of, scheme 811
- Thymine (Thy) 199*s*, 203, 1541*s*
  - formation 810
- Thymine-DNA glycosylase 1583
- Thymineless death 812
- Thymol 1233*s*
- Thymopoietin 1756, 1901
- Thymus 1833
- Thyroglobulin 1430, 1432
- Thyroid hormone(s) **1430 – 1432**, 1903
  - effect on growth 1432
- Thyroid hormone receptors 1432
- Thyroid peroxidase 856
- Thyroid transcription factor 1432
- Thyroperoxidase 1432
- Thyrotropin (TSH) 1430, 1745
- Thyrotropin (thyrotropic hormone)-releasing factor 54*s*, 523
  - receptor 554
- Thyrotropin-releasing hormone 1430, 1746
- Thyroxine 1430, 1432, 1742
- Thyroxine-binding globulin 1430
- Tight junctions 10, 28
- Time-of-entry map 1486
- Ti plasmid 1497
- Tissue(s)
  - adipose 26
  - animal 25
  - collenchyma 29
  - composition of 31
  - connective 26
  - epidermal of plants 30
  - epithelial 26
  - fractionation of 98 – 100
  - meristematic 29
  - muscle 26
  - oxidative damage to 1074
  - parenchyma 29
  - sclerenchyma 29
  - supporting 26
- Tissue factor 631, 633*s*
  - pathway 632
- Tissue plasminogen activator 634
- Titin 74, **1099**
  - location of 1100
- Titration curves **96**, **97**
  - of histidine 97
  - of proteins 330
- tmRNA, structure 1711
- Tobacco mosaic virus 247, 334*s*, 336*s*, 1475
- Tocopherols 392, **815 – 822**, 1427, 1429*s*
  - as antioxidants 822, 1075
  - biosynthesis of, scheme 1429
- α-Tocopherol (vitamin E) 819*s*
- Tocopherolquinone 818, 819, 819*s*
- Tocotrienols 818, 1787
- Togavirus 247
- Toll-like receptors 1832
- Toluene absorption spectrum 1281
- Toluenesulfonyl-amido-2-phenylethyl chloromethyl ketone (TPCK) 623*s*
- Tomaymycin 224*s*
- TonA, B 839
- Tonegawa, Susumu 84
- Tonoplast 11
- Topaquinone (TPQ) 816, 817*s*
- Topogenic signals 1018
- Topoisomerase(s) 219, 575, 638, 657, 659, 1544, 1550, 1551
- Topoisomerase I 1551
  - ribbon drawing 1553
- Topoisomerase II 1552
  - model 1553
- Topological theory of graphs 466
- Topology of proteins 76
  - diagram 76
- Torsion angles 44
  - definition of 59
  - of nucleosides 212
  - of peptides 63, 139
  - of polynucleotides 211
  - of side chain groups 61
  - of sugar rings 212
- Total protonic potential difference 1038
- Total reflection FTIR 1277
- Totipotency **1885 – 1888**
- Toxaphine 1788
- Toxic shock syndrome 1686, 1857
- Toxin-agglutinin fold 1775
- Toxins, channel-forming 414
- Toxopyrimidine 739
- TPQ. *See* Topaquinone
- Tracheids 30
- Tracheophyta 29
- Trachoma 7
- Traffic ATPases 417
- Transacylation 608
- Transaldolase 700, 964, 965
- Transaminases. *See* Aminotransferases
- Transamination 737
  - by PLP-dependent enzymes 741
- Transcobalamin 869
- Transcortin 58, 1253
- Transcription of genetic information 5, **1603 – 1657**
  - activators 1628 – 1632
  - in archaea 1624
  - in bacteria 1603 – 1620
  - of chloroplast genes 1649 – 1657
  - definition of 5
  - elongation in 1610, 1611, 1637, 1638
  - enhancers 1628 – 1632
  - in eukaryotes 1624

  - figure 1630
  - general transcription factors 1629
  - initial transcribing complex 1609, 1610
  - initiation of **1607–1610**
  - light-induced 1320, 1636
  - mediators 1630
  - of mitochondrial genes 1649–1657
  - modification of 1637
  - nutrient control 1635
  - promoters 1628–1632
  - regulation of 536
  - Rho 1616
  - by RNA polymerases I and II 1636, 1637
  - silencer of 1626
  - of SV40 DNA 1628
  - TATA-binding protein 1629
  - TBP-associated factors 1629
  - termination factors 1616
  - termination of 1615, 1616, 1637
  - transcriptosomes 1629
  - of viral genes **1622–1624**, 1628, **1649–1657**
- Transcription bubble, structure of 1609
- Transcription-coupled repair (TCR) 1585
- Transcription factor(s) 76, 539
  - AP-1 576, 1631, 1633
  - binding to enhancer site 576
  - E2F 580, 581, 1633
  - Ets-domain 1634
  - forkhead family 1634
  - GATA family 1634
  - GCN4 70, 241
  - general 1629
  - HMG proteins 1634
  - leucine zipper 1633
  - Max 241, 70*s*
  - NF-κB, 243*s*
  - octamer-binding 1636
  - osteoblast-specific 441
  - Pit-1 1636
  - proto-oncogenes as 576
  - Sp1 202, 1628, 1630
  - table 1632
  - winged helix 1634
  - zinc-containing 1634
- Transcription factor HSF 1636
- Transcriptional responses 1635, 1636
- Transcriptional units 1628
- Transcripts. *See* mRNA
- Transducin 558, 1330
  - cycle, scheme 1331
- Transduction
  - by bacteriophage P1 1486
- Trans fatty acids 1205
- Transfection 1494
- Transferases 498, **589–662**
  - aminotransferases 737, **740–743**, **747–753**
  - definition of 498
- Transferrins 58, 840, 841
- Transfer RNA. *See* tRNA
- Transformation
  - of bacteria 1473
- Transforming growth factor-α (TGF-α) 1753
- Transforming growth factor-β (TGF-β) 1755, 1849
- Transgene 1501
- Transglutaminase 80, 633
- Transhydrogenase reaction 981, 1047
- Transient receptor potential 1800
- Transimination
    - in PLP-dependent enzymes 741, 751
- Transition dipole moment 1283
- Transition metals **837–894**
- Transition state **482–486**
  - diagram 482
  - inhibitors 484
  - theory 483
- Transketolase 733, 736, 964, 986
- Translational repressors 1698
- Translation of genetic information 5, 1673–1728. *See also* Protein synthesis
  - definition of 5
  - nick translation 257
  - regulation of 536, 538, 539, 1698–1710
- Translocase of the mitochondrial inner membrane (Tim) 1018, 1019
- Translocase of the outer mitochondrial membrane (Tom) 1018, 1019
- Translocation
  - in ribosome 1685, 1705, 1708
  - of proteins 519–521, 1720–1723
- Translocon 1721
- Transmembrane proteins 391
- Transmethylation **591**, **592**
  - kinetic isotope effects on 592
- Transmission coefficient κ 483
- Transmission computerized tomography (CT) 1806
- Transpeptidation 1369
- Transplantation immunity 1852
- Transport. *See also* Specific substrates
  - of ions 420–425
  - through membranes 400–415
  - across mitochondrial membranes 1047–1050
  - via pores and channels 401–403
  - retrograde 1159
- Transporter(s)
  - 12-helix 415–417
  - membrane 411–427
  - sugar 415, 416
- Transporter associated with antigen processing (TAP) 1723
- Transposable genetic elements. *See* Transposons
- Transposable recombinational switch 1880
- Transposase 1574
  - gene 1575
- Transposons 249, 1573–1577
  - eukaryotic 1576
  - figure 1574
  - integration of, scheme 1576
  - maize 1577
  - mariner 1577
  - nonreplicative 1575
  - replicative 1575
  - Tn3 219, 1575
- Transsialidase 187
- Transsulfuration sequence 1388
- Transthyretin 1241, 1430, 1719
- Transverse tubules (T-tubules)
  - electron micrograph of 1114
- Treadmilling of microtubules 372
- Trefoil family peptides 1843
- Trehalose 167, 168*s*
  - biosynthesis 1141
  - in fungi 168
  - in insects 168
  - properties 1141
- Tremerogens 1758
- Triacylglycerol(s) 381, 382*s*, 1181–1185, 1192, 1196–1199, 1204–1206
  - biosynthesis of, scheme 1198
    - in human body 1181–1186
  - movement of 1184
  - synthesis of 1181–1185, 1196–1199
- Trialkyl lock 495*s*
- Tricarboxylic acid
  - cycle 950–959
  - transport in mitochondria 1047
- Tricarboxylic transporter
  - of mitochondria 1047
- Trichina 24
- Trichodiene 1234
- Tricholomic acid 739
- *Trichonympha* 19
- Trichothiodystrophy (TTD, brittle hair disease) 1585
- TRICINE buffer 99
- Triclosan 777*s*, 777
- Triethanolamine buffer 99
- Trigger factor 1721
- Triglycosylceramide, accumulation of 1170
- Trigonal bipyramid 638
  - as transition state in pseudorotation 638
- Trigonal carbon atoms 680, 681
- Trigonal prochiral centers 480, 481
- Triiodothyronine 572, 1430, 1430*s*, 1432
- Trimethoprim 805*s*
  - inhibition of dihydrofolate reductase 805
- Trimethylamine dehydrogenase 782, 784*s*
- Trimethylarsonium lactic acid, 387*s*
- Trimethyllysine
  - in ubiquitin 525
- Triose phosphate isomerase 693, 694*s*
  - barrel structure 77*s*
  - high catalytic activity 693
  - target for antitrypanosomal drugs 693
- Tripeptidyl peptidase 140, 610
- Triphosphopyridine nucleotide (TPN$^+$, NADP$^+$) 767. *See also* NADP$^+$, NADPH
- Triple repeat diseases 1113, 1516
  - table 1515
- Triplet state 1290
- Tripolyphosphate, formation of 977
- TRIS buffer 99
- Trisodium phosphonoformate 1655*s*
- Tritium 110
- Triton X 403*s*
- tRNA 230, 231, 231*s*, 233*s*, 1475, 1620, **1687–1698**
  - acceptor specificity 1694–1698
  - aminoacylation of 1694–1698
  - anticodon-dihydrouridine stem loop 1692
  - anticodon loop 1688
  - binding sites in ribosome 1687
  - cloverleaf diagram 1688
  - cognate 1672
  - dihydroU loop 1688
  - ^{19}F NMR spectrum of 270
  - ^{1}H NMR spectrum of 268
  - hypermodified bases 234, 235, 1688
  - initiator 1693
  - L form 1688
  - pre-tRNA 1646
  - presursor, sequence of 1620
  - processing of 1641, 1642
  - structures of 1687
  - TψC loop 1688
  - tunnel 1673
  - variable loop 1688

  - wobble position 231
- tRNA precursors
  - pathways 1645
  - structure 1645
- Trophectoderm 1897
- Tropical macrocytic anemia 802
- Tropoelastin 436
- Tropomodulin 406, 1100
- Tropomyosin 370, 406, 1098, 1099, 1100, 1116
  - coiled coil structure of 71
- Troponin 1098, 1099, 1100, 1116
- Troponin C 313, 314
- Troponin-tropomyosin-actin complex
  - in muscle regulation 1116
- Trp repressor protein 239, 240*s*
- Truffles 20
- Tryosine kinases 657
- Trypanosome 19, 1866
  - DNA circles in 219
  - mitochondria of 14
- Trypanothione 552
- Trypanothione reductase 785
- Tryparsamide 597*s*
- Trypsin 66, 116, 609, 1102. *See also* Chymotrypsin
  - hydrogen-bonding network, structure 612
  - oxyanion hole, structure 615
  - specificity 117
  - turnover number of 457
- Trypsinogen 609, 615*s*
- Tryptases 610
- Tryptathione 1382
- Tryptic peptides 117
- Tryptolines 1811
- Tryptophan (Trp, W) 52*s*, 1059*s*, 1425, 1445*s*, 1809
  - absorption spectrum 123, 1284
  - biosynthesis of, scheme 1426
  - C_β-hydroxylated 853
  - metabolism of **1443 – 1450**
  - nicotinamide activity 769
  - reactions of, scheme 1445
- Tryptophan 2,3-dioxygenase (tryptophan pyrrolase) 1443
- Tryptophan dioxygenase (indoleamine 2,3-dioxygenase)1058
- Tryptophan hydroxylase 1061, 1062
- Tryptophan indole-lyase (tryptophanase) 742
  - quinonoid intermediate 750
- Tryptophan pyrrolase. *See* Tryptophan 2,3-dioxygenase
- Tryptophan synthase 742, 1427*s*
- Tryptophan tryptophanylquinone (TTQ) 817, 817*s*
- T-segment 1552
- T system of membranes, in muscle 1114
- T system of tubules 12
- TTQ. *See* Tryptophan tryptophanylquinone
- T tubules 1098
- Tuberculosis 7, 1194
- D-Tubocurarine 1783*s*, 1785
- Tubular myelin 386
- Tubulin 370, 372*s*, 1110
- Tumor necrosis factor (TNF) 1756, 1849
- Tumor suppressor genes 407, 574, 581, 1502, 1503, 1907
- Tungsten 893
- Tunicamycin 1155, 1155*s*, 1157
- Tunicates
  - vanadium in 25
- Tunichlorin(s) 880*s*, 881
- Tunneling
  - quantum mechanical 494, 771, 848
  - vibration-assisted 494
- Turnover numbers of an enzyme 457
- Turns
  - beta 78
  - in protein structures 72, 74
- Twisted sheets 63
- Twitchin 1117
- Two-carbon units
  - in biosynthesis 990
- Two-fold (dyad) axes 134
  - in oligomers 337, 348
- Two-hybrid system 1725, 1726
- Typhoid fever, *S. typhi* 7
- Tyrocidine 1164, 1842
- Tyrosinase 886, 1065
- Tyrosine (Tyr, Y) 52*s*
  - absorption spectrum 23, 1284
  - catabolism, microbial 1434
  - iodination 126
  - metabolism of **1428 – 1443**
    - in plants 1438 – 1443
    - scheme 1431
  - synthesis of 1424, 1425
- Tyrosine decarboxylase 737
- Tyrosine hydroxylase 1061
- Tyrosine recombinase 1572. *See also* Integrase family
- Tyrosinemia 1430
- Tyrosine phenol-lyase 742
- Tyrosine protein kinases (MEK) 578
- Tyrosine-*O*-sulfate 548
- Tyrosyl radical 864
- D-Tyvelose (3,6-dideoxy-D- mannose) 180*s*
  - from CDP-glucose 1138

U

- Ubichromanol 818, 819*s*
- Ubiquinol-cyt *c* oxidoreductase (Complex II) 1021, 1027, 1028
- Ubiquinone(s) (coenzyme Q) 392, 514, 818, 819*s*, 1021, 1023, 1427, 1907
  - biosynthesis of, scheme 1429
  - content in mitochondria 1019
- Ubiquinone-binding proteins 1023
- Ubiquinonol 1429*s*
- Ubiquitin 81, 524, 524*s*, 525, 1626, 1728
  - activating enzyme **524**, **525**
  - carboxyl-terminal hydrolases 525
  - conjugating enzymes 524
  - genes 525
  - protein ligase 524
- UDP (uridine diphosphate). *See* Uridine 5’ phosphate
- UDP-galactose
  - conversion to UDP glucose 778
- UDP-galactose 4-epimerase 778
- *Ulothrix* 21
- Ultracentrifugation **108**, **109**
- Ultracentrifuge 100
  - analytical 108
  - optical system, figure 109
- Ultrafiltration 100
- Ultrasensitive responses 567
- Ultrasonic sound 468
- Ultraviolet light, damages from 1297
- UMP. *See* Uridine 5'-phosphate
- Uncouplers in mitochondria **1046 – 1051**
- Uncoupling protein 1048
- Undecaprenol 1152
- Unimolecular processes 457
- Uniporters 414
- Units, International System 2
- Unsatisfied ends of hydrogen bonded chains 64
- Uracil (Ura, U) 199*s*, 203
  - tautomerism of 45
- Uracil-DNA glycosylase 1581
- Urate oxidase 886, 1459
- Urea 82*s*, 478*s*, 1368,
  - biosynthesis of, scheme 1375
  - as osmolyte 1143
- Urea carboxylase 730, 1378
- Urea cycle 1374 – 1379
  - defects 1378
  - function, primary 1378
  - with mitochondrial metabolism, scheme 1377
- Urease 478, 877, 878, 1378
  - active site of, 877*s*
  - mechanism of 877
- Ureides 1460
- Ureido anion 726*s*
- Ureidoglycine 1458*s*
- Ureidoglycolate 1458*s*
- β-Ureidopropionate 1453*s*
- Urethan 1588
- Uric acid 203*s*, 1458*s*, 1459
  - formation of 890
- Uridine (Urd) 203, 234*s*
  - absorption spectra of 205
  - tautomer, minor 205*s*
- Uridine 5'- phosphate. *See* UMP
- Uridine diphosphate-*N*-acetylgalactosamine 1136
- Uridine diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) 1135*s*
  - formation 1135
- Uridine diphosphate-*N*-acetyl-mannosamine (UDP-ManNAc)
  - eperimization of 1136
- Uridine diphosphate-*N*-acetylmuramic acid 1135*s*
  - biosynthesis 1135
- Uridine diphosphate galactose (UDPGal) 1130
- Uridine diphosphate galacturonic acid (UDPGal) 1130
- Uridine diphosphate glucose (UDP-Glc) 515*s*, 720, 989, 1129*s*, 1130, 1148
- Uridine diphosphate glucuronate acid 1130, 1133*s*
- Uridine diphosphate L-iduronic acid 1130
- Uridine diphosphate-glucuronosyltransferases 1142
- Uridine diphosphate glucose (UDP-Glc)
  - oxidation of 1130
- Uridine 5'-phosphate (Urd-5'-P or UMP) 200, 200*s*, 203, 1451, 1451*s*
- Uridylate kinase 655
- 5'-Uridylic acid. *See* UMP
- Uridylyltransferase (UT) 1371
- Urobilin 1404, 1404*s*
- Urocanase reaction 778
- Urocanate 755, 756*s*, 1450*s*
- Urogastrone. *See* Epidermal growth factor (EGF)
- Urokinase 634
- Uronic acid 164

- Uroporphyrin(s) 843
- Uroporphyrin I, 845*s*
- Uroporphyrinogen 1401*s*
- Uroporphyrinogen decarboxylase 1400
- Uroporphyrinogen III cosynthase 1400
- Urothione 804*s*, 891
- Urticaria 385
- Usher protein 364
- Usher syndrome 1113, 1332
- Utrophin 1112

V

- Vaccenic acid 1191
- Vaccenoyl-ACP 1188*s*, 1192
- Vaccination 1859
- Vaccines 1517
- Vacuole(s) 10 – 12
  - definition of 3
- Valine (Val, V) 52*s*, 1391
  - biosynthesis of 527, 712
    - scheme 1392
  - branched fatty acids from 381
  - catabolism of, scheme 1396
- Valinomycin 414, 415*s*, 1164
  - effect on potassium transport 414
- Valium 1788*s*
- Vancomycin 1166, 1166*s*
- van der Waals contact surfaces
  - of purines and pyrimidines 207
- van der Waals forces 46
- van der Waals radii **40**, **41**
- Vanadate
  - as inhibitor of ATPases 889
  - insulinlike action 889
- Vanadium 856, **889**
- Vanadocytes 889
- Vanadoproteins 889
- Vanilla 1440
- Vanillate 1790*s*
- Vanillic acid 1431*s*
- Vanillin 1439*s*, 1440
- Vanilloid receptors 1800
- van't Hoff equation 289
- Vapor phase chromatography 103
- Variable loop 1688
- Vascular endothelial growth factor VEGF 1755, 1902
- Vascular plants 29
- Vasculogenesis 1902
- Vasoactive intestinal peptide (VIP) 1749
- Vasopressin 54*s*, 542, 563, 1747, 1748*s*
  - receptor 554
- Vasotocin 1748, 1748*s*
- Vectors (cloning vehicles) 1492 – 1497
  - cosmid 1496
  - lambda 1496
- Veech 980
- Vegetal pole in egg 1897
- Ventricles 1766
- Veratridine 1775
- Verdoheme 1404*s*
- Verotoxin 1685
- Versican 1154
- Very low density lipoprotein (VLDL) 1181
  - composition of, table 1183
  - receptors for 1186
  - remnants 1186
- Vesicles, coated 426
- VHL (van Hippel-Lindau cancer syndrome) 564
- Vibrational circular dichroism 1287
- Vibrational energy of molecules 1276
- Vibrational Raman optical activity 1287
- Vibrations, of molecule 1277
- Vibrio 6, 7
- Vimentin 369, 1099
- Vinblastine 371
- Vincristine 371, 1448
- Vinculin 406
- Vinyl chloride 1588
- Vinylpyridine 116*s*
- Violaxanthin 1240, 1240*s*, 1319
  - abscisic acid, precursor 1243
  - cycle 1319
  - de-epoxidase 1243
- Viomycin 1689*s*
- Viral genomes, packaging of 1559
- Viral oncogenes 573
- Viroid(s) 247, 1646
  - potato spindle tuber structure 1645
- Virion 244
- Virus(es) **244 – 249**, 1866. *See also* Bacteriophage
  - adeno 247
  - baculo 247
  - binding to cells 186
  - BK 247
  - cauliflower mosaic 247
  - characteristics, table 245
  - coat 345
  - DNA in 1529
  - DNA, table 245
  - dsRNA containing 244 – 247
  - electron micrographs of 246
  - Epstein-Barr 247
  - eukaryotic 1562
  - hepatitis A 247
  - hepatitis delta 247
  - herpes 247
  - human immunodeficiency virus (HIV) 248
  - life cycles of 248
  - maedi 1656
  - Mengo 247
  - oncogenic 248
  - papova 244
  - picorna 247
  - plant 346
  - polio 247
  - polyhedral structures 344
  - polyoma 247
  - pox 247
  - protein coat 334
  - rhabdo 247
  - rhino 247
  - RNA containing 247, 248
    - table 245
  - satellite tobacco mosaic 343*s*
  - satellite tobacco necrosis 247, 343*s*, 344
  - simian (SV40) 244
  - ssDNA containing 244
  - tobacco mosaic 247
  - toga 247
  - visna 1656
- Virus fd protein sheath 335*s*
- Virus ϕX174
  - assembly of virion structure 365
- Virusoids 1646
- Viscumin 1685
- Vision 1798
  - invertebrate 1332
- Visual chromophores **1326 – 1329**
- Visual cortex 1766
- Visual pigments 1324 – 1329
  - transmembrane structure 1325
- Vitamin(s) 719. *See also* Individual vitamins
  - deficiency diseases 721
  - discovery of **721**
- Vitamin A 1241, 1242. *See also* Retinol
  - deficiency 1242
  - discovery of 721
  - in reproduction 1241
  - teratogenic 1242
  - in vision 1241
- Vitamin B complex 721
  - nutritional requirements 756
- Vitamin B_1 331. *See also* Thiamin
- Vitamin B_{12} 721, **866 – 877**, 950. *See also* Cobalamin, Cyanocobalamin
  - adenosyltransferase 870
  - blood levels 869
  - cobalt in 866 – 877
  - coenzyme forms 867
  - nutritional requirements 756
- Vitamin B_{12} coenzyme 864. *See also* Cobalamin
  - dependent reactions, table 871
  - enzymatic functions 870 – 877
  - isomerization reactions 872
  - nonenzymatic cleavage 870
  - ribonucleotide reductase 871
- Vitamin B_6 family 721, **738**, 1460. *See also* Pyridoxol, Pyridoxal phosphate
  - functions 737– 753
  - nutritional requirements 756
- Vitamin B_t, 944. *See also* Carnitine
- Vitamin C. *See* Ascorbic acid (ascorbate)
- Vitamin D (calciferol) 721, 1257, 1757
  - deposition of calcium in bones 314
  - function, principal 1257
  - oxygenases in formation of 1057
- Vitamin E 721, 1814, 1906. *See also* Tocopherols
  - nutritional requirement 822
- Vitamin K 721, 818, 819*s*, **820 – 822**, 1425, 1427, 1428, 1429*s*
  - in blood clotting 821
  - biosynthesis of, scheme 1429
  - dihydro 820
  - epoxide 820
  - phylloquinone 821
- Vitamin P. *See* Citrus bioflavonoid
- Vitiligo 1297, 1435
- Vitronectin 409
- VLDL receptor 1185
- Volicitin 1212, 1212*s*
- Volkensin 1685
- Voltage-gated anion-selective channel (VDAC) 1047
- Voltage-gated K^+ channel 412 – 414
- Voltage-sensitive ion channels 412, 1769
- Voltage sensor 1115
- Volvox 21
- Vomeronasal organ (VNO) 1799
- von Ebner's glands 1800
- von Szent-Györgyi, Albert 83
- von Willebrand factor 409, 633
- Voodoo lily 1048
- V system 475

W

Wald, George 84
Warburg, Otto H. 83, 767, 963
Warfarin 821, 822*s*
Water
    addition to carbonyl 677
    clusters of molecules 49
    content in cells 30
    content in tissues 31
    diffusion constant of 461
    hydrogen bonding of 48
    properties of **49 – 51**
    structure of **49**, **50**
Watson–Crick base pair 207, 208, 231
Watson–Crick structure of DNA 200
Watson, James D. 84, 200, 1474
Wave number 1274
Waxes 382, 382*s*
    insect 1196
Werner syndrome 1585, 1907
Wheat, genome 12
Whooping cough 548
Wilkins, Maurice F. 200
Williams, R. J. D. 1040
Wilson's disease 883
Winding number. *See* DNA, circular, linking number
Wingless (Wg) protein 1900
Winogradsky 1051
Wnt signaling pathway 1899
Wnt-4 protein 1895
Wobble hypothesis 1693
Wobble pairing 209
Wood-Ljundahl pathway 881, 984
Work 282
    chemical 282
    to concentrate 1 mol of a substance 283
    electrical 282, 302
    electrochemical 300
    mechanical 282
    to raise 1 kg 1 m 283
    to remove two charges 47
World Wide Web
    protein sequences/structures 148
Wortmannin 566*s*
Wounds, healing of 29

X

Xanthan gum 179, 1153
Xanthine (Xan) 203*s*, 1458*s*, 1459
Xanthine dehydrogenases 794, 825, 890
    molybdenum in 890
Xanthine oxidase 890, 892
Xanthine ribotide 1455
Xanthophyll(s) 1240
Xanthophyll cycle 1238, 1243, 1319
Xanthopsin 1336
Xanthopterin 803, 804*s*
Xanthurenic acid 1444, 1445*s*
X chromosomes 1894
    dosage compensation 1893
    inactivation center 1894
Xenobiotics **550**
Xenoestrogens 1263
XerC, D 1570
Xeroderma pigmentosum 1581, 1585
X-linked ichthyosis 1252
X-linked immunoproliferative syndrome 1867
X-rays, anomolous scattering 135
X-ray absorption fine structure, extended (EXAFS) 1288
X-ray absorption near-edge structure (XANES) 1288
X-ray absorption spectroscopy 1288
X-ray diffraction **132 – 137**, 611, 1669. *See also* Three-dimensional structures
    difference electron density map 136
    electron density map 135
    isomorphous replacement method 133
    MAD phasing 135
    refinement 136
    space groups 133
Xylanases 602, 603
Xylans 165, 170, 175, 1130
Xylem 30
Xylitol 1133*s*, 1134
    catabolism of 1132 – 1135
    sweetness of 1135
Xyloglucan 177, 1149
α-D-Xylopyranose 695*s*
Xylose (Xyl) 163*s*, 165*s*, 175
D-Xylose (Xyl) 162
Xylose isomerase 527, 693, 695
Xylosidases 602
Xylulose 164*s*
D-Xylulose 1133*s*, 1134
L-Xylulose 1133*s*, 1134
Xylulose 5-phosphate 964*s*
D-Xylulose-1-phosphate 1135

Y

Y chromosome 1512, 1893, 1894
    SRY 1893
Yeast(s) 20, 1891, 1892
    genome 1726
    mating types 1725
    two-hybrid system 1725
Yeast artificial chromosomes (YACs) 1497, 1562
Yellow fever 247
Yellow fluorescent protein 1345
Yellow photoactive protein (PYP) 1337
Ylid 733*s*
Young, W. 960

Z

Z (zusammen) configuration 43
Z-discs 1096
    proteins of 1099
    structure 1100
Zeaxanthin 1239*s*, 1243, 1319
Zellweger syndrome 945, 1202
Zinc 317, **680**
    in alcohol dehydrogenases 772 – 775
    content of human 680
Zinc finger domains 1781
Zinc finger protein 241, 242*s*, 243*s*, 680
Zinc ion(s)
    chelation by imidazole groups 625
    in enzymes 773
    replacements 680
Zinc proteases 625 – 627
    FtsH 628
Zinc ribbon motifs 1637
Zingerone (ginger) 1439*s*
Z-lines 1098
Zona pellucida 1895
Zoospore maintenance factor 1758, 1758*s*
Z scheme diagram 1299
Zwitterion 41
Zygote 17
Zymogen 519, 609. *See also* Proenzyme
    granules 609
Zymosterol 1246*s*
Zyxin 406